

# **Sistema de apoyo para la toma de decisiones en Unicauca Virtual utilizando Bodegas de datos y OLAP**



**Trabajo de Grado**

**Lisandro Acosta Mejía  
Jaime Alberto Muñoz Córdoba**

**Directora: MSc. Martha Eliana Mendoza Becerra**

*Uníversidad del Cauca*

**Facultad de Ingeniería Electrónica y Telecomunicaciones**

**Departamento de Sistemas**

**Grupo de I+D en tecnologías de la información**

**Popayán, Junio de 2006**



## TABLA DE CONTENIDO

<b>INTRODUCCION</b> .....	<b>9</b>
<b>PARTE I - MARCO CONCEPTUAL</b> .....	<b>11</b>
<b>CAPITULO 1 - BODEGAS DE DATOS Y OLAP EN UNICAUCA VIRTUAL</b> .....	<b>12</b>
<b>1 DESCRIPCIÓN DEL PROYECTO</b> .....	<b>12</b>
<b>1.1 JUSTIFICACIÓN</b> .....	<b>13</b>
1.1.1 Justificación tecnológica.....	13
1.1.2 Justificación académica .....	13
1.1.3 Justificación social.....	14
1.1.4 Viabilidad del proyecto .....	14
<b>1.2 OBJETIVOS DEL PROYECTO</b> .....	<b>14</b>
<b>CAPITULO 2 - MARCO TEORICO</b> .....	<b>16</b>
<b>2 EDUCACION EN LINEA</b> .....	<b>16</b>
<b>2.1 LA EDUCACIÓN</b> .....	<b>16</b>
<b>2.2 LA EDUCACIÓN EN LÍNEA</b> .....	<b>16</b>
2.2.1 Ventajas de la educación en línea .....	16
2.2.2 Desventajas de la educación en línea.....	17
<b>2.3 SISTEMA ADMINISTRADOR DE APRENDIZAJE (LMS)</b> .....	<b>18</b>
<b>3 SISTEMAS DE INFORMACION</b> .....	<b>18</b>
<b>3.1 DEFINICIÓN</b> .....	<b>18</b>
<b>3.2 Clasificación de los sistemas de información</b> .....	<b>18</b>
3.2.1 Sistemas de procesamiento de transacciones [3] .....	18
3.2.2 Sistemas de información gerencial (MIS) [5].....	18
3.2.3 Sistemas de apoyo a decisiones (DSS) [4].....	19
3.2.4 Sistemas de información para oficinas [3]:.....	20
<b>4 SISTEMAS DE SOPORTE PARA LA TOMA DE DECISIONES (DSS)</b> .....	<b>20</b>
<b>4.1 TOMA DE DECISIONES [6]</b> .....	<b>20</b>
<b>4.2 DEFINICIÓN DSS</b> .....	<b>21</b>
<b>4.3 CARACTERÍSTICAS DSS</b> .....	<b>21</b>
<b>5 BODEGA DE DATOS</b> .....	<b>22</b>
<b>5.1 DEFINICIÓN BODEGA DE DATOS</b> .....	<b>22</b>
<b>5.2 CARACTERÍSTICAS DE LAS BODEGAS DE DATOS</b> .....	<b>23</b>
5.2.1 Orientadas a temas .....	23
5.2.2 Integradas .....	23
5.2.3 No volátiles.....	23
5.2.4 De tiempo variante.....	24
<b>5.3 DIFERENCIAS ENTRE UNA BODEGA DE DATOS Y UN SISTEMAS DE PROCESAMIENTO DE TRANSACCIONES EN LÍNEA (OLTP)</b> .....	<b>24</b>



<b>5.4</b>	<b>ARQUITECTURA GENERAL</b> .....	<b>24</b>
5.4.1	Captura de datos .....	25
5.4.2	Bodega de datos.....	25
5.4.3	Data Mart.....	26
5.4.4	Componentes de acceso.....	26
<b>5.5</b>	<b>BODEGAS DE DATOS EN ORACLE</b> .....	<b>26</b>
5.5.1	Definición.....	26
5.5.2	Diseño lógico y físico en el ambiente de bodega de datos.....	26
5.5.3	Esquemas de la bodega de datos [11].....	28
5.5.4	Objetos de la bodega de datos [11].....	29
<b>6</b>	<b>PROCESAMIENTO ANALITICO EN LINEA (OLAP)</b> .....	<b>31</b>
<b>6.1</b>	<b>DEFINICIÓN</b> .....	<b>31</b>
<b>6.2</b>	<b>CLASIFICACIÓN</b> .....	<b>32</b>
<b>6.3</b>	<b>OLAP EN ORACLE</b> .....	<b>33</b>
6.3.1	Decidiendo cuando usar ROLAP o MOLAP en Oracle [16].....	34
<b>7</b>	<b>METODOLOGIA DE DESARROLLO</b> .....	<b>35</b>
<b>7.1</b>	<b>METODOLOGÍA DE DESARROLLO DE LA BODEGA DE DATOS</b> .....	<b>35</b>
7.1.1	La fase de planeación del proyecto.....	36
7.1.2	Definición de requerimientos del negocio.....	36
7.1.3	Despliegue.....	38
7.1.4	Mantenimiento y crecimiento.....	39
7.1.5	Administración del proyecto.....	39
<b>7.2</b>	<b>METODOLOGIA DE DESARROLLO DEL PROTOTIPO DE HERRAMIENTA OLAP</b> <b>39</b>	
<b>PARTE II – BODEGA DE DATOS</b> .....		<b>41</b>
<b>CAPITULO 3 – DEFINICION DEL PROYECTO Y RECOLECCION DE REQUERIMIENTOS</b> .....		<b>42</b>
<b>8</b>	<b>PLANEACION DEL PROYECTO</b> .....	<b>42</b>
<b>8.1</b>	<b>DEFINICION DEL PROYECTO</b> .....	<b>42</b>
8.1.1	Escenario de la organización.....	42
8.1.2	Definición de requerimientos.....	43
8.1.3	Priorización de los temas de interés para la bodega de datos.....	45
8.1.4	Alcance de la bodega de datos.....	46
8.1.5	Justificación de la bodega de datos.....	47
<b>8.2</b>	<b>PLANEACION DEL PROYECTO</b> .....	<b>47</b>
8.2.1	Identidad del proyecto.....	47
8.2.2	Personal en el proyecto.....	47
8.2.3	Plan del proyecto.....	48
<b>8.3</b>	<b>CONSIDERACIONES ESPECIALES SOBRE LA DEFINICION Y PLANEACION DEL PROYECTO DE BODEGA DE DATOS</b> .....	<b>51</b>
<b>CAPITULO 4 – MODELADO DIMENSIONAL</b> .....		<b>52</b>
<b>9</b>	<b>MATRIZ BUS</b> .....	<b>52</b>



<b>10</b>	<b>DISEÑO DE LAS TABLAS DE HECHOS.....</b>	<b>54</b>
<b>10.1</b>	<b>DATA MART – INTERACCION DEL ESTUDIANTE .....</b>	<b>55</b>
10.1.1	Detalles de la tabla de hechos “Visualización de recurso” .....	56
10.1.2	Detalles de las tablas de dimensión.....	57
<b>10.2</b>	<b>DATA MART – ESTILOS DE APRENDIZAJE .....</b>	<b>59</b>
10.2.1	Detalles de la tabla de hechos “Cambio en el estilo de aprendizaje del estudiante” .....	61
10.2.2	Detalles de la tabla de hechos “Recurso en el estilo de aprendizaje” .....	63
10.2.3	Detalles de las tablas dimensión .....	63
<b>10.3</b>	<b>DATA MART – EVALUACIONES.....</b>	<b>64</b>
10.3.1	Detalles de la tabla de hechos “Evaluación de actividad” .....	65
10.3.2	Detalles de las tablas dimensión .....	66
<b>10.4</b>	<b>CONSIDERACIONES ESPECIALES SOBRE EL MODELADO DIMENSIONAL DE LA BODEGA DE DATOS .....</b>	<b>66</b>
<b>CAPITULO 5 – IMPLEMENTACIÓN DE LA BODEGA .....</b>		<b>69</b>
<b>11</b>	<b>DISEÑO FISICO DE LOS MODELOS DIMENSIONALES.....</b>	<b>69</b>
<b>12</b>	<b>DEFINICION DE LAS TABLAS FISICAS Y OBJETOS DIMENSIONALES .....</b>	<b>71</b>
<b>12.1</b>	<b>TABLAS RELACIONALES .....</b>	<b>71</b>
<b>12.2</b>	<b>OBJETOS DIMENSIONALES .....</b>	<b>73</b>
<b>12.3</b>	<b>PROCESO DE CARGA DE DATOS (ETL) .....</b>	<b>75</b>
<b>12.4</b>	<b>CONSIDERACIONES ESPECIALES EN LA IMPLEMENTACIÓN DE LA BODEGA DE DATOS .....</b>	<b>75</b>
12.3.1	ROLAP .....	75
12.3.2	MOLAP .....	76
12.3.3	¿ROLAP O MOLAP?.....	77
<b>PARTE III – PROTOTIPO HERRAMIENTA OLAP .....</b>		<b>79</b>
<b>CAPITULO 6 – ANALISIS.....</b>		<b>80</b>
<b>13</b>	<b>DEFINICION DE REQUERIMIENTOS .....</b>	<b>80</b>
<b>13.1</b>	<b>USUARIOS.....</b>	<b>80</b>
<b>13.2</b>	<b>METAS .....</b>	<b>80</b>
<b>13.3</b>	<b>FUNCIONES DEL SISTEMA.....</b>	<b>80</b>
<b>13.4</b>	<b>ATRIBUTOS DEL SISTEMA.....</b>	<b>83</b>
<b>14</b>	<b>CASOS DE USO.....</b>	<b>84</b>
<b>14.1</b>	<b>CASO DE USO CARGAR CATALOGO.....</b>	<b>85</b>
<b>15</b>	<b>MODELO CONCEPTUAL PRELIMINAR.....</b>	<b>86</b>
<b>16</b>	<b>DIAGRAMAS DE SECUENCIA DEL SISTEMA.....</b>	<b>88</b>
<b>16.1</b>	<b>DIAGRAMA DE SECUENCIA CARGAR CATALOGO.....</b>	<b>88</b>
<b>CAPITULO 7 – DISEÑO .....</b>		<b>89</b>



<b>17</b>	<b>ARQUITECTURA DE LA APLICACIÓN.....</b>	<b>89</b>
17.1	CLIENTES .....	90
17.2	CAPA DE LÓGICA DE PRESENTACIÓN .....	90
17.3	FACHADA .....	91
17.4	CAPA LÓGICA DEL NEGOCIO.....	91
17.5	CAPA LÓGICA DE SERVICIOS.....	91
<b>18</b>	<b>CASOS DE USO REALES.....</b>	<b>92</b>
18.1	CASO DE USO REAL CARGAR CATALOGO.....	92
<b>CAPITULO 8 – CONSTRUCCIÓN DEL PROTOTIPO DE HERRAMIENTA OLAP.....</b>		<b>95</b>
<b>19</b>	<b>VISTA DE IMPLEMENTACIÓN PROTOTIPO OLAP .....</b>	<b>95</b>
19.1	CLIENTE .....	96
19.2	SERVIDOR DE APLICACIÓN .....	96
19.3	SERVIDOR DE BASE DE DATOS.....	96
<b>20</b>	<b>CONSIDERACIONES ESPECIALES DE LA IMPLEMENTACIÓN DEL PROTOTIPO. ....</b>	<b>97</b>
20.1	VISTAS DEL SISTEMAS PARA EL MANEJO DEL CATALOGO OLAP EN ORACLE 97	
20.1.1	ALL\$OLAP2UCATALOGS .....	97
20.1.2	ALL\$OLAP2UCUBE_DIM_USES .....	98
20.2	EXTENSIONES SQL PARA BODEGAS DE DATOS RELACIONALES EN ORACLE	99
20.2.1	Extensión CUBE para la cláusula GROUP BY .....	99
20.2.2	Funciones GROUPING .....	99
20.3	¿POR QUE UN PROTOTIPO DE HERRAMIENTA OLAP? .....	100
<b>CAPITULO 9 – DESCRIPCION DEL PROTOTIPO .....</b>		<b>101</b>
<b>21</b>	<b>FUNCIONALIDADES DEL PROTOTIPO .....</b>	<b>101</b>
21.1	VISUALIZACIÓN DE COMPONENTES DE LA BODEGA DE DATOS .....	102
21.2	CREACIÓN DE LA CONSULTA ANALÍTICA.....	105
21.3	VISUALIZACIÓN DE DATOS .....	107
21.4	GESTIÓN DE LAS CONSULTAS ANALÍTICAS .....	108
21.5	FILTRADO DE COMPONENTES Y VALORES .....	109
<b>CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO.....</b>		<b>110</b>
<b>REFERENCIAS BIBLIOGRAFICAS .....</b>		<b>113</b>
<b>ANEXOS .....</b>		<b>115</b>



## LISTA DE FIGURAS

Figura 1. Integración de datos desde los sistemas operacionales hacia la bodega de datos. ....	23
Figura 2. Arquitectura general de una Bodega de Datos. ....	25
Figura 3. Esquema en estrella.....	28
Figura 4. Esquema copo de nieve. ....	29
Figura 5. Esquema tercera forma normal. ....	29
Figura 6. Objetos típicos para una bodega de datos. ....	31
Figura 7. Ciclo de vida para la construcción de una bodega de datos según R. Kimball .....	36
Figura 8. Priorización de requerimientos del negocio.....	46
Figura 9. Diagrama de la tabla de hechos “Visualización de Recurso”.....	57
Figura 10. Descripción de la tabla de hechos “Visualización de recurso”.....	57
Figura 11. Diagrama de detalle para la dimensión “Estudiante”.....	58
Figura 12. Diagrama de la tabla de hechos “Cambio en el estilo de aprendizaje del estudiante”.....	62
Figura 13. Descripción de la tabla de hechos “Cambio del estilo de aprendizaje del estudiante”.....	62
Figura 14. Diagrama de la tabla de hechos “Recurso en el estilo de aprendizaje”.....	63
Figura 15. Descripción de la tabla de hechos “Recurso en el estilo de aprendizaje”.....	63
Figura 16. Diagrama de la tabla de hechos “Evaluación de actividad”.....	65
Figura 17. Descripción de la tabla de hechos “Evaluación de actividad”.....	66
Figura 18. Modelo físico de la tabla de hechos “Visualización de recurso”.....	70
Figura 19. Script de creación de la tabla relacional base para la dimensión “Actividad”.....	72
Figura 20. Datos de prueba almacenados en la tabla “Actividad”.....	72
Figura 21. Script de creación de la tabla relacional base para la definición del cubo “Visualización de recurso” ..	73
Figura 22. Sentencias SQL que permiten definir la dimensión “Actividad”.....	74
Figura 23. Sentencias que permiten definir los metadatos para la dimensión “Actividad”.....	74
Figura 24. Diagrama general de casos de uso para el prototipo OLAP.....	84
Figura 25. Arquitectura general del sistema.....	89
Figura 26. Distribución de componentes implementados. ....	95
Figura 27. Pantalla principal del prototipo de herramienta OLAP.....	102
Figura 28. Área modelos dimensionales. ....	103
Figura 29. Área modelos dimensionales – Data Marts, Cubos, Dimensiones, Jerarquías, Datos de nivel. ....	104
Figura 30. Cambiar Jerarquía de una dimensión. Cambiar Jerarquía 1.a. Jerarquía por defecto (género) de la dimensión Estudiante Cambiar Jerarquía 1.b. Pantalla que presenta las dimensiones definidas para la dimensión Estudiante. Cambiar Jerarquía 1.c. Jerarquía pises de residencia para la dimensión estudiante. ....	105
Figura 31. Creación de una consulta analítica.....	106
Figura 32. Abrir una nueva ventana de análisis.....	107
Figura 33. Visualización de datos en forma de matriz de datos. ....	107
Figura 34. Gestión de archivos de análisis.....	108
Figura 35. Filtrado de datos de nivel. Filtrar 1.a. Presenta los datos de nivel sin filtrar. Filtrar 1.b. Muestra la pantalla que presenta al usuario los datos a filtrar. Filtrar 1.c. Presenta en nivel con los datos filtrados.....	109
Figura 36. Modelo físico para la tabla de hechos “Cambio en el estilo de aprendizaje del estudiante”.....	121
Figura 37. Modelo físico para la tabla de hechos “Recurso en estilo de aprendizaje”.....	122
Figura 38. Modelo físico para la tabla de hechos “Evaluación de actividad”.....	123
Figura 39. Diagrama de detalle para la dimensión “Profesor”.....	124
Figura 40. Diagrama de detalle para la dimensión “Recurso”.....	125
Figura 41. Diagrama de detalle para la dimensión “Actividad”.....	126
Figura 42. Diagrama de detalle para la dimensión “Curso”.....	127
Figura 43. Diagrama de detalle para la dimensión “Fecha”.....	128
Figura 44. Diagrama de detalle para la dimensión “Periodo”.....	129
Figura 45. Diagrama de detalle para la dimensión “Localidad”.....	130
Figura 46. Extensión del caso de uso Gestionar consulta analítica.....	143
Figura 47. Extensión para el caso de uso Gestionar área de trabajo.....	147



## LISTA DE TABLAS

Tabla 1 Diferencias entre un DSS y un MIS.....	20
Tabla 2. Diferencias entre un Sistema Operacional y una Bodega de Datos.....	24
Tabla 3. Plan del proyecto. ....	50
Tabla 4. Matriz bus. ....	54
Tabla 5. Descripción detallada de la dimensión “Estudiante”.....	59
Tabla 6. Convenciones. ....	59
Tabla 7. Funciones básicas del prototipo de herramienta OLAP.....	83
Tabla 8. Atributos del prototipo de herramienta OLAP.....	83
Tabla 9. Análisis del caso de uso cargar catalogo.....	86
Tabla 10. Modelo Conceptual Preliminar.....	88
Tabla 11. Diagrama de secuencia caso de uso Cargar catalogo.....	88
Tabla 12. Caso de uso real Cargar Catalogo. ....	94
Tabla 13. Descripción columnas vista ALL\$OLAP2UCATALOGS.....	98
Tabla 14. Descripción columnas vista ALL\$OLAP2UCUBE_DIM_USES.....	99
Tabla 15. Descripción detallada de la dimensión “Profesor”.....	125
Tabla 16. Descripción detallada de la dimensión “Recurso”.....	126
Tabla 17. Descripción detallada de la dimensión “Actividad”. ....	127
Tabla 18. Descripción detallada de la dimensión “Curso”.....	127
Tabla 19. Descripción detallada de la dimensión “Fecha”.....	129
Tabla 20. Descripción detallada de la dimensión “Periodo”.....	130
Tabla 21. Descripción detallada de la subdimensión “Localidad”.....	131
Tabla 22. Análisis del caso de uso diseñar consulta analítica.....	141
Tabla 23. Análisis del caso de uso generar matriz de datos. ....	142
Tabla 24. Análisis del caso de uso entrar al sistema.....	143
Tabla 25. Análisis del caso de uso guardar diseño consulta analítica.....	144
Tabla 26. Análisis del caso de uso abrir diseño consulta analítica. ....	146
Tabla 27. Análisis del caso de uso generar grafico de datos. ....	147
Tabla 28. Análisis del caso de uso guardar área de trabajo.....	148
Tabla 29. Análisis del caso de uso abrir área de trabajo.....	150
Tabla 30. Análisis del caso de uso filtrar niveles.....	151
Tabla 31. Análisis del caso de uso restablecer datos de nivel.....	152
Tabla 32. Análisis del caso de uso cambiar jerarquía. ....	153
Tabla 33. Diagrama de secuencia caso de uso Diseñar Consulta Analítica.....	154
Tabla 34. Diagrama de secuencia caso de uso Generar matriz de datos. ....	155
Tabla 35. Diagrama de secuencia caso de uso Entrar al sistema.....	156
Tabla 36. Diagrama de secuencia caso de uso Guardar diseño consulta analítica.....	157
Tabla 37. Diagrama de secuencia caso de uso Abrir diseño consulta analítica.....	158
Tabla 38. Diagrama de secuencia caso de uso Generar grafico de datos. ....	159
Tabla 39. Diagrama de secuencia caso de uso Guardar área de trabajo.....	160
Tabla 40. Diagrama de secuencia caso de uso Abrir área de trabajo.....	161
Tabla 41. Diagrama de secuencia caso de uso Filtrar niveles.....	162
Tabla 42. Diagrama de secuencia caso de uso Restablecer datos de nivel.....	163
Tabla 43. Diagrama de secuencia caso de uso Cambiar jerarquía. ....	164
Tabla 44. Caso de uso real Diseñar consulta analítica. ....	167
Tabla 45. Caso de uso real Generar matriz de datos. ....	169
Tabla 46. Caso de uso real Entrar al sistema. ....	170
Tabla 47. Caso de uso real Guardar diseño consulta analítica.....	171
Tabla 48. Caso de uso real Abrir diseño consulta analítica. ....	173
Tabla 49. Caso de uso real Generar grafico de datos.....	175
Tabla 50. Caso de uso real Guardar área de trabajo.....	176
Tabla 51. Caso de uso real Abrir área de trabajo.....	178



---

Tabla 52. Caso de uso real Filtrar niveles.....	179
Tabla 53. Caso de uso real Restablecer datos de nivel.....	180
Tabla 54. Caso de uso real Cambiar jerarquía.....	182
Tabla 55. Descripción columnas vista ALL\$OLAP2UCATALOG_ENTITY_USES.....	183
Tabla 56. Descripción columnas vista ALL\$OLAP2UCATALOGS.....	184
Tabla 57. Descripción columnas vista ALL\$OLAP2UCUBE_DIM_USES.....	185
Tabla 58. Descripción columnas vista ALL\$OLAP2UCUBE_MEAS_DIM_USES.....	185
Tabla 59. Descripción columnas vista ALL\$OLAP2UCUBE_MEASURE_MAPS.....	186
Tabla 60. Descripción columnas vista ALL\$OLAP2UCUBE_MEASURES.....	187
Tabla 61. Descripción columnas vista ALL\$OLAP2UCUBES.....	188
Tabla 62. Descripción columnas vista ALL\$OLAP2UDIM_ATTR_USES.....	188
Tabla 63. Descripción columnas vista ALL\$OLAP2UDIM_ATTRIBUTES.....	189
Tabla 64. Descripción columnas vista ALL\$OLAP2UDIM_HIER_LEVEL_USES.....	189
Tabla 65. Descripción columnas vista ALL\$OLAP2UDIM_HIERARCHIES.....	190
Tabla 66. Descripción columnas vista ALL\$OLAP2UDIM_LEVEL_ATTR_MAPS.....	191
Tabla 67. Descripción columnas vista ALL\$OLAP2UDIM_LEVEL_ATTRIBUTES.....	192
Tabla 68. Descripción columnas vista ALL\$OLAP2UDIM_LEVELS.....	193
Tabla 69. Descripción columnas vista ALL\$OLAP2UDIMENSIONS.....	193
Tabla 70. Descripción columnas vista ALL\$OLAP2UENTITY_DESC_USES.....	195
Tabla 71. Descripción columnas vista ALL\$OLAP2UFACT_LEVEL_USES.....	196
Tabla 72. Descripción columnas vista ALL\$OLAP2UFACT_TABLE_GID.....	197
Tabla 73. Descripción columnas vista ALL\$OLAP2UHIER_CUSTOM_SORT.....	198
Tabla 74. Descripción columnas vista ALL\$OLAP2UJOIN_KEY_COLUMN_USES.....	199
Tabla 75. Descripción columnas vista ALL\$OLAP2ULEVEL_KEY_COLUMN_USES.....	200
Tabla 76. Ejemplo reporte cross- tabular.....	202



---

# INTRODUCCION

---

A lo largo de este documento, se desarrollan temáticas y conceptos teóricos importantes para el desarrollo del proyecto. El documento se encuentra organizado de la siguiente manera:

## PARTE 1 – MARCO CONCEPTUAL

*Capítulo 1 – Bodegas de datos y OLAP en Unicauca Virtual.* Ofrece una descripción general del proyecto y presenta la justificación y los objetivos específicos del mismo.

*Capítulo 2 – Marco teórico.* Este capítulo presenta los fundamentos teóricos requeridos para el desarrollo del proyecto. Dentro de estos se encuentran conceptos como la educación en línea, los sistemas de información, los sistemas de soporte a la toma de decisiones, las bodegas de datos, el procesamiento analítico en línea (OLAP) y la metodología de desarrollo del proyecto.

## PARTE 2 – BODEGA DE DATOS

*Capítulo 3 – Definición del proyecto y recolección de requerimientos:* Este capítulo enmarca dos etapas del ciclo de vida para la construcción de bodegas de datos, la definición y la recolección de requerimientos.

*Capítulo 4 – Modelado dimensional:* En este capítulo se muestran los modelos dimensionales creados para el proyecto de bodegas de datos, describiendo detalladamente cada uno de sus componentes.

*Capítulo 5 – Implementación de la bodega:* En este capítulo se especifican los modelos físicos correspondientes a los modelos dimensionales diseñados y se muestran detalles de implementación sobre el motor de base de datos Oracle 10g.

## PARTE 3 – PROTOTIPO DE HERRAMIENTA OLAP.

*Capítulo 6 – Análisis:* Presenta los artefactos obtenidos al realizar el análisis para la construcción del prototipo de herramienta OLAP. Incluye los casos de uso y los diagramas de secuencia.

*Capítulo 7 – Diseño:* Presenta los artefactos obtenidos al realizar el diseño del prototipo de herramienta OLAP. Incluye los casos de uso real y la arquitectura del prototipo.



*Capítulo 8 – Construcción del prototipo de herramienta OLAP:* Presenta una vista de implementación del prototipo y señala algunos aspectos especiales tenidos en cuenta para la construcción.

*Capítulo 9 – Descripción del prototipo:* Ofrece una descripción detallada del prototipo de herramienta OLAP señalando las funcionalidades implementadas y su forma de uso.

*Conclusiones, recomendaciones y trabajo futuro:* Aquí se describen las conclusiones a las cuales se llegó una vez terminado el desarrollo del proyecto. Además se incluyen algunas recomendaciones y aspectos relacionados con el trabajo futuro del proyecto.

*Referencias bibliográficas.* Se presentan todas las referencias bibliográficas tenidas en cuenta para el desarrollo del proyecto.

*Anexos:* Información complementaria de lo tratado a lo largo del documento.



---

# PARTE I - MARCO CONCEPTUAL

---



---

# CAPITULO 1 - BODEGAS DE DATOS Y OLAP EN UNICAUCA VIRTUAL

---

## 1 DESCRIPCIÓN DEL PROYECTO

Sin intención de reemplazar la educación convencional, la educación virtual aparece en el entorno académico como una alternativa para mejorar procesos presenciales de educación y para soportar nuevos paradigmas de enseñanza. Las ventajas que presenta la educación virtual, como el hecho de disminuir las limitaciones geográficas de los estudiantes que pueden tener problemas para desplazarse, abre nuevas puertas para todos aquellos que deseen aprender.

Es por esto que desde diversas áreas de la tecnología se apoyan iniciativas de educación virtual, intentando así mejorar la experiencia académica y los mismos productos o servicios que soportan esta forma de aprendizaje. Algunas de estas áreas tecnológicas se relacionan con los sistemas adaptativos, dentro de estos, los sistemas tutores inteligentes (STI), con los que se busca que el proceso de enseñanza – aprendizaje del estudiante mejore al personalizar sus procesos pedagógicos dentro el ambiente virtual. Otra área que aporta a la educación virtual es la del aprendizaje colaborativo, que pretende mejorar este proceso mediante el trabajo grupal y colaborativo; y en el caso de este proyecto, una tecnología (bodegas de datos y OLAP) muy utilizada en el ámbito empresarial, con la cual se busca permitir a los tutores y directivos contar con una herramienta que les ayude a conocer las actividades que realiza el estudiante y su comportamiento en general, facilitando la identificación de posibles problemas que se puedan generar y aprovechando aspectos positivos de la experiencia del estudiante.

En el ambiente empresarial se utilizan sistemas de información que soportan las operaciones diarias de las empresas, es decir, las transacciones que se pueden generar en una jornada normal de trabajo. Estos sistemas permiten gestionar datos, hacer los procesos del negocio más fácil y se caracterizan por recolectar gran cantidad de información.

La búsqueda de formas para aprovechar la información almacenada en las bases de datos transaccionales, ha llevado a diferentes expertos a plantear nuevas metodologías y estándares centrados en el análisis de información, conocidos en el ambiente tecnológico como Bodegas de Datos (Data Warehouse), Procesamiento Analítico en Línea (OLAP), Minería de Datos (Data Mining), etc. Todos estos terminan convirtiéndose en sistemas informáticos muy robustos denominados Sistemas de



Soporte para la toma de Decisiones (DSS), los cuales son muy utilizados por las empresas en distintas áreas como: ventas, transporte, seguridad, sector financiero, medicina, educación, entre otras. Unicauca Virtual<sup>1</sup> siendo un ambiente de aprendizaje virtual, consideró necesario incluir en su segunda fase un servicio en el cual se utilicen estas tecnologías como soporte al análisis de las diferentes actividades que realizan los estudiantes dentro del ambiente y que ayuden a tomar decisiones importantes que podrían afectar el desempeño del estudiante dentro del sistema.

## 1.1 JUSTIFICACIÓN

### 1.1.1 Justificación tecnológica

Este proyecto, se basa en tecnologías que a nivel mundial están siendo muy utilizadas en el área de la gestión de negocios en general. Estas tecnologías conocidas como bodegas de datos y Procesamiento Analítico en Línea (OLAP), han dado excelentes resultados en grandes cadenas comerciales que requieren de análisis de información que ayude a encontrar nuevas oportunidades de negocio y que apoyen su crecimiento.

Este proyecto buscó aplicar estas mismas tecnologías para el análisis de información generada por un ambiente educativo virtual en línea, facilitando la labor de gestión y buscando obtener el mismo éxito obtenido en la industria. Lo anterior puede representar para la Universidad del Cauca un gran adelanto tecnológico y de innovación al intentar aplicar dentro del ambiente académico técnicas que dentro de las universidades, tanto de la región como del país, no han sido ampliamente utilizadas.

### 1.1.2 Justificación académica

Este proyecto trata un tema novedoso dentro del ambiente académico de la Universidad del Cauca, abre nuevas puertas y posibilidades de investigación en el área, posibilidades que pueden ser utilizadas por estudiantes y docentes que quieran ampliar su bagaje intelectual o crear nuevas áreas de interés dentro de la Facultad de Ingeniería Electrónica y de Telecomunicaciones y en particular en el Departamento de Sistemas.

Así mismo, el desarrollo del proyecto representa para sus autores la posibilidad de aplicar los conocimientos adquiridos a lo largo de su formación como Ingenieros de Sistemas en áreas como: ingeniería del software, bases de datos, interfaz humano-computador entre otras.

---

<sup>1</sup> Proyecto piloto de apropiación tecnológica en el marco de la educación en línea, basado en un sistema de aprendizaje que opera a través de las más avanzadas tecnologías de información, las telecomunicaciones y la calidad en educación, cimentando así la creación de la Universidad Virtual del Cauca como alternativa para aumentar la cobertura de los programas que ofrece el Alma Mater de una forma eficaz y eficiente.



### **1.1.3 Justificación social**

El resultado esperado al finalizar el macroproyecto Unicauca Virtual, es un sistema educativo para ofrecer programas académicos en línea con altos niveles de calidad, que permitan aumentar la cobertura en regiones apartadas del departamento del Cauca.

Este proyecto ayuda en la consecución de los beneficios sociales de Unicauca Virtual al suministrar un mecanismo que apoya la gestión realizada por los tutores y directivos, ofreciendo posibilidades de obtener mayor información sobre el comportamiento de los estudiantes en el Sistema de Gestión de Aprendizaje (Learning Management System, LMS) y de este modo contribuir en el mejoramiento de su proceso de aprendizaje y en su formación integral como futuro profesional.

### **1.1.4 Viabilidad del proyecto**

Este proyecto se enmarcó en el macroproyecto “Unicauca Virtual”, por lo tanto contó con los recursos económicos y de infraestructura tecnológica necesaria para su completo desarrollo, debido al patrocinio que Colciencias y la Universidad del Cauca han dado al proyecto. Por otro lado, se tienen las bases metodológicas y tecnológicas desarrolladas en la primera fase del macroproyecto, dando vía libre a la continuación de la segunda fase y los proyectos contenidos en ellas.

## **1.2 OBJETIVOS DEL PROYECTO**

Con el desarrollo de este proyecto se pretendía en términos específicos, en primera instancia diseñar y construir una bodega de datos que permita almacenar información relacionada con las actividades académicas que desarrolla el estudiante en Unicauca Virtual y que será utilizada por el sistema de soporte a decisiones.

En segunda instancia construir un prototipo de una herramienta OLAP que permita obtener y procesar información contenida en la bodega de datos, para apoyar la toma de decisiones estratégicas de los docentes y directivos de Unicauca Virtual. Este prototipo tendrá las siguientes funcionalidades:

- Gestionar un conjunto de consultas analíticas<sup>2</sup> predefinidas que cumpla con las necesidades de información requerida para apoyar la toma de decisiones estratégica de los usuarios.
- Gestionar nuevas consultas analíticas definidas por los usuarios.

---

<sup>2</sup> Se refiere a un conjunto de informes o reportes que presentan información en un formato determinado (graficas, estadísticas, tablas, etcétera).



Y en tercera instancia Integrar el Sistema de Soporte a la Toma de Decisiones al macroproyecto Unicauca Virtual.

Todo lo anterior dentro del macroproyecto Unicauca Virtual planteado por el Grupo de I+D en Tecnologías de la información (GTI) de la Universidad del Cauca y apoyado por Colciencias.



---

# CAPITULO 2 - MARCO TEORICO

---

## 2 EDUCACION EN LINEA

### 2.1 LA EDUCACIÓN

Para hablar sobre educación en línea, se debe en primera instancia hablar del tema de la educación convencional, la cual se puede entender como un medio privilegiado por el cual la sociedad contribuye al desarrollo intelectual, físico, ético, afectivo, moral y estético de los individuos; prepara para el trabajo productivo y para el disfrute de los bienes materiales, espirituales y culturales; sirve para mejorar la calidad de vida; y forma para la defensa de los recursos naturales y la preservación del ambiente [1].

Algunos de los objetivos de la educación son [1]:

- La educación debe contribuir en forma eficaz y sistemática a la profundización de la democracia, la participación ciudadana, la construcción de una cultura de convivencia y respeto de los derechos humanos y a la conquista de la paz.
- El proceso educativo, inspirado en la vida misma, será integral y estará centrado en el desarrollo de las potencialidades y los talentos de la persona. Cultivará la capacidad de aprender a aprender, la creatividad, la autonomía, el espíritu crítico y reflexivo y el trabajo en equipo. Fomentará un pensamiento más diferenciador que generalizador, más indagante que concluyente, más proactivo que reactivo [1].

### 2.2 LA EDUCACIÓN EN LÍNEA

Se entiende por Educación en Línea el desarrollo e implementación del proceso de educación a distancia (formal o no formal) basado en el uso de las Tecnologías de la Información y las Comunicaciones (TIC), que posibilitan la realización de un aprendizaje interactivo, flexible y accesible a cualquier receptor potencial. De esta manera, permite desarrollar cursos y otras actividades educativas sin que todos los participantes tengan que estar simultáneamente en el mismo lugar. Además la educación en línea se ha utilizado como apoyo a procesos de aprendizaje en esquemas presenciales [2].

#### 2.2.1 Ventajas de la educación en línea



Algunas de las ventajas más importantes de la educación en línea, son:

- **Mayor productividad:** el estudiante utiliza y distribuye su tiempo, aprende a su propio ritmo y enfoca su aprendizaje en temas de su interés.
- **Entrega oportuna:** la enseñanza por medio de la educación en línea esta disponible para un gran número de personas interesadas en aprender.
- **Capacitación flexible:** El estudiante puede elegir el camino a seguir durante su aprendizaje.
- **Acceso y Flexibilidad:** se vencen barreras como las distancias, consistentes en el desplazamiento en ocasiones a lugares lejanos e incluso al exterior del país de origen, se vencen limitaciones de tiempo, ya que en la actualidad no se cuenta con el suficiente para desarrollar una vida con responsabilidades laborales y a la vez seguir estudiando, tiene además materiales de primera y maestros de gran calidad y diversidad de pensamientos [1].
- **Costos:** En la actualidad el costo de la educación en línea disminuye y gana terreno al estudio tradicional, se necesita inicialmente una inversión pero con el tiempo sin lugar a dudas y conforme llegue a más personas será más barata. [1]
- **Autoestudio:** Una cualidad de la enseñanza virtual es estimular el Autoestudio. En la educación presencial el alumnado está acostumbrado a que se le diga lo que viene en el examen y en la educación virtual se convierte en el “centro del proceso de enseñanza” es responsable de su propio desarrollo y el maestro es un facilitador [1].

### 2.2.2 Desventajas de la educación en línea

Algunas de las desventajas que presenta la educación en línea son [1]:

- **Economía:** Aunque es considerada una ventaja, la accesibilidad (en especial de Internet) representa también una enorme desventaja, sobre todo en países donde la mayoría de la población vive en la pobreza.
- **Cultura:** En muchas partes del mundo, todavía se tiene la mentalidad de que para aprender hay que ir a la escuela, sentarse y escuchar a los profesores.
- **Empleo:** Uno de los grandes cuestionamientos de la educación en línea es si los encargados de contratar a los futuros profesionales valorarán de igual manera a los egresados de las aulas de una institución, que a los graduados de una universidad virtual.



- **Motivación:** El hecho de que el estudiante asuma la responsabilidad de su propio aprendizaje implica que, si no encuentra una motivación para seguir o el curso no está bien diseñado, opte por abandonar su proceso educativo.

## 2.3 SISTEMA ADMINISTRADOR DE APRENDIZAJE (LMS)

LMS es la sigla de Learning Management System, la plataforma de formación o plataforma de educación en línea. Estos sistemas integran los cursos y objetos de conocimientos, registran, evalúan y organizan a los usuarios y conectan actividades asincrónicas y sincrónicas. Un LMS es diseñado generalmente para ser utilizado por diferentes editores y proveedores. Generalmente no incluye posibilidades de autoría (creación de cursos propios), en su lugar, se centra en gestionar cursos creados por una gran variedad de fuentes diferentes.

# 3 SISTEMAS DE INFORMACION

## 3.1 DEFINICIÓN

Los sistemas de información son un conjunto formal de procesos que, operando sobre una colección de datos estructurada según las necesidades de la empresa, recopilan, elaboran y distribuyen la información (o parte de ella) necesarias para las operaciones de dicha empresa y para las actividades de dirección y control correspondientes para desempeñar su actividad de acuerdo a su estrategia de negocio.

## 3.2 Clasificación de los sistemas de información

Por lo general los sistemas de información tienen como objetivo procesar datos que pueden ser utilizados para: capturar detalles de las transacciones, permitir que se tomen decisiones o comunicar la información entre personas y localidades. Los sistemas se pueden clasificar de la siguiente manera:

### 3.2.1 Sistemas de procesamiento de transacciones [3]

Estos sistemas procesan datos referentes a las actividades (transacciones) de la empresa, por ejemplo, las ventas, los pedidos e inventario. La función en sí, que se llama “procesamiento de transacciones” se funda en la operación metódica de cualquier empresa u organización. Las cinco razones para el procesado de las transacciones son: la clasificación, el cálculo, la distribución, el resumen y el almacenamiento de los datos.

### 3.2.2 Sistemas de información gerencial (MIS) [5]



Son una clase de sistemas de decisiones administrativas que proporcionan información en forma periódica para ayudar a los gerentes con las decisiones que surjan y que puedan anticiparse; por lo tanto son altamente estructurados y los informes están en un formato predeterminado hecho por el desarrollador del sistema.

Una definición mas formal considera al MIS como un sistema computacional que proporciona información a usuarios que tienen necesidades similares. La información describe a la compañía o a uno de sus principales sistemas en términos de lo que ha sucedido en el pasado, lo que esta sucediendo en el presente y lo que es probable que suceda en el futuro.

### 3.2.3 Sistemas de apoyo a decisiones (DSS) [4]

El DSS (Decision Support Systems) proporciona información para resolver problemas estructurados, y a la vez provee mecanismos de comunicación que permiten resolver problemas que carecen de estructura y no son rutinarios. La información se produce como informes periódicos y especiales definidos por el mismo gerente, el cual trabaja de forma interactiva con el sistema.

Dentro de los objetivos que debe alcanzar un DSS se encuentran:

- Ayudar a los gerentes a tomar decisiones para resolver problemas semiestructurados, es decir problemas de los cuales se tiene poca comprensión.
- Apoyar el juicio del gerente en lugar de tratar de reemplazarlo.
- Mejorar la eficacia del gerente en la toma de decisiones, mas que su eficiencia.

El enfoque principal del concepto de DSS, es usar un computador para apoyar a un *gerente específico* que toma una *decisión específica* para resolver un *problema específico*.

Aunque existen similitudes entre los sistemas de información gerencial (MIS) y los sistemas de soporte a decisiones, también existen diferencias. De forma tradicional los diseñadores de MIS eran técnicos expertos, mientras que los gerentes (quienes debían tomar las decisiones) solo hacían pequeños aportes. En cambio, el DSS se centra en el proceso de tomar decisiones y en los gerentes quienes con la cooperación de los profesionales técnicos diseñan el sistema apropiado para un puesto en particular. Al tener acceso a bases de datos en los DSS, los administradores pueden aprovechar la información e investigar la eficacia de cursos alternativos de acción. La Tabla 1, especifica algunas diferencias entre los DSS y los MIS [6].

Sistemas de información gerencial (MIS)	Sistemas de soporte a decisiones (DSS)
-----------------------------------------	----------------------------------------



Se centra en decisiones estructuradas y decisiones rutinarias.	Se centra en tareas semiestructuradas que requieren de criterio administrativo.
Se hace hincapié en el almacenamiento de información	Se hace hincapié en la manipulación de los datos.
Con frecuencia los administradores tienen acceso a los datos.	Los administradores tienen acceso directo a los datos.
Se apoya en expertos en computación.	Se apoya en el juicio temporal del administrador.
El gerente de MIS no comprende por completo la naturaleza de la decisión.	El gerente conoce el ambiente de la decisión.
Se centra en la eficiencia.	Se centra en la eficacia.

Tabla 1. Diferencias entre un DSS y un MIS.

### 3.2.4 Sistemas de información para oficinas [3]:

Son una clase especial de un sistema de procesamiento de información que puede usarse en el medio de las oficinas. Estos sistemas evolucionaron a partir del procesamiento de palabras, que son sistemas para aumentar la aptitud tanto de gerentes como del personal de operación para elaborar correspondencia, reportes, relaciones y documentos especiales. Aquí se encuentran, los Sistemas Ofimáticos (OA), Sistemas de Transmisión de Mensajería (E-mail y Fax Server) y Coordinación y Control de tareas (Work Flow), entre otros

## 4 SISTEMAS DE SOPORTE PARA LA TOMA DE DECISIONES (DSS)

### 4.1 TOMA DE DECISIONES [6]

La toma de decisiones se define como la selección de un curso de acción entre alternativas y se encuentra enmarcado en el núcleo de la planeación dentro de una organización. Se pueden diferenciar tres tipos de decisiones: programadas, no programadas e intermedias.

Una decisión *programada* (operacional) se aplica a problemas estructurados o rutinarios. Por ejemplo, los operadores de tornos tienen especificaciones y reglas que les señalan si las piezas que han hecho son aceptables, si se tiene que desechar o si se debe reprocesar. Otro ejemplo de una decisión programada es colocar nuevos pedidos de artículos de existencia estándar. Esa clase de decisiones, que se usan para trabajos rutinarios y repetitivos, se basa principalmente en criterios previamente establecidos. Es una toma de decisiones por precedentes.



Las decisiones *no programadas*, se aplican en situaciones no estructuradas, nuevas y mal definidas, de una naturaleza no repetitiva. Las decisiones estratégicas son decisiones no programadas puesto que requieren de juicios subjetivos.

La mayor parte de las decisiones no son completamente programadas ni completamente no programadas, sino una combinación de ambas. La mayor parte de las decisiones no programadas son tomadas por los administradores de niveles más altos, debido a que son ellos quienes tienen que hacer frente a problemas no estructurados. Con frecuencia, los problemas de los niveles inferiores de la organización son rutinarios y bien estructurados y requieren de menos libertad en la toma de decisiones por parte de gerentes y no gerentes.

## 4.2 DEFINICIÓN DSS

Los sistemas de soporte para la toma de decisiones (DSS), han sido estudiados por muchos investigadores y especialistas del área de la informática, por tal motivo se encuentran muchas definiciones de estos sistemas. A continuación se presentan dos de ellas dadas por C. J. Date y W. H. Inmon.

Según C. J. Date [7], los sistemas de apoyo para la toma de decisiones son sistemas que ayudan en el análisis de información de negocios. Su propósito es ayudar a la administración para que “Marque tendencias, señale problemas y tome.... decisiones inteligentes”.

Por otro lado, W. H. Inmon [8] afirma que un sistema de apoyo para la toma de decisiones es un sistema para soportar decisiones de gerencia. Usualmente estos sistemas involucran el análisis de muchas unidades de datos en una forma heurística. Como una regla los procesos DSS no involucran la modificación de datos. Para Inmon, las bodegas de datos son la base del procesamiento DSS dado que presentan una fuerte integración y accesibilidad de datos.

## 4.3 CARACTERÍSTICAS DSS

Existen varias características que deben estar presentes en un sistema para poder considerarlo un sistema de soporte a la toma de decisiones. Algunas de estas características según Daniel Cohen [9], son:

- **Interactividad:** Sistema computacional con la posibilidad de interactuar en forma amigable y con respuestas en tiempo real con el encargado de tomar decisiones.
- **Tipo de decisiones:** Apoya el proceso de toma de decisiones estructuradas y no estructuradas.



- **Frecuencia de uso:** Tiene una utilización frecuente por parte de la administración media y alta para el desempeño de su función.
- **Variedad de usuarios:** Puede emplearse por usuarios de diferentes áreas funcionales como ventas, producción, administración, finanzas y recursos humanos.
- **Flexibilidad:** Permite acoplarse a una variedad determinada de estilos administrativos: autoritativos, participativos, etc.
- **Desarrollo:** Permite que el usuario desarrolle de manera directa modelos de decisión sin la participación operativa de profesionales en informática.
- **Interacción ambiental:** Permite la interacción con información externa como parte de los modelos de decisión.
- **Comunicación ínter organizacional:** Facilita la comunicación de información relevante de los niveles altos, a los niveles operativos y viceversa a través de graficas.
- **Acceso a base de datos:** Tiene capacidad de acceder a información de las bases de datos corporativas.
- **Simplicidad:** Simple y fácil de aprender y utilizar por el usuario final.

## 5 BODEGA DE DATOS

Entre otras Tecnologías de Información existentes en el mercado, se encuentran las Bodegas de Datos, que junto a las “Herramientas de Procesamiento de Transacciones en Línea (OLAP)”, permiten a los analistas, gerentes y ejecutivos sintetizar información sobre la empresa a través de comparaciones, visiones personalizadas, análisis histórico y proyecciones de datos en varios escenarios.

La tendencia en la construcción de sistemas es separar el procesamiento de datos operacional del procesamiento de datos analítico o informativo. Las bodegas de datos son el corazón de este último tipo de procesamiento también conocido como procesamiento DSS o procesamiento de soporte para la toma de decisiones.

### 5.1 DEFINICIÓN BODEGA DE DATOS

W. H. Inmon [8] las define como una colección de datos integrados, orientados a temas, que dan soporte a las funcionalidades del DSS, donde cada unidad de dato es

relevante en algún momento en el tiempo. El diseño y construcción de una Bodega de Datos o Data Warehouse, es de gran importancia en el análisis de datos manejados en un ambiente determinado.

Según R. Kimball [10] una bodega de datos es una copia de datos transaccionales específicamente estructurados para consultas y análisis.

## 5.2 CARACTERÍSTICAS DE LAS BODEGAS DE DATOS

Según la definición de W. H. Inmon [8], cuatro características son importantes para definir una bodega de datos, ellas son:

### 5.2.1 Orientadas a temas

Significa que para la construcción de una bodega de datos se seleccionan áreas importantes del negocio, como por ejemplo el área de “Finanzas” en una organización.

### 5.2.2 Integradas

Porque los datos que contiene, a pesar de que vienen de diferentes fuentes de datos operacionales, están en un formato consistente. La Figura 1 muestra las diferentes formas de representar los datos en los sistemas operacionales. Es el caso de la representación del “genero de las personas”, cuya notación en un sistema puede ser “M y F” y en otro “H y M”, ambos representado el sexo Masculino y Femenino. Este problema de escritura debe ser filtrado a un formato único o estándar al pasar al Data Warehouse.

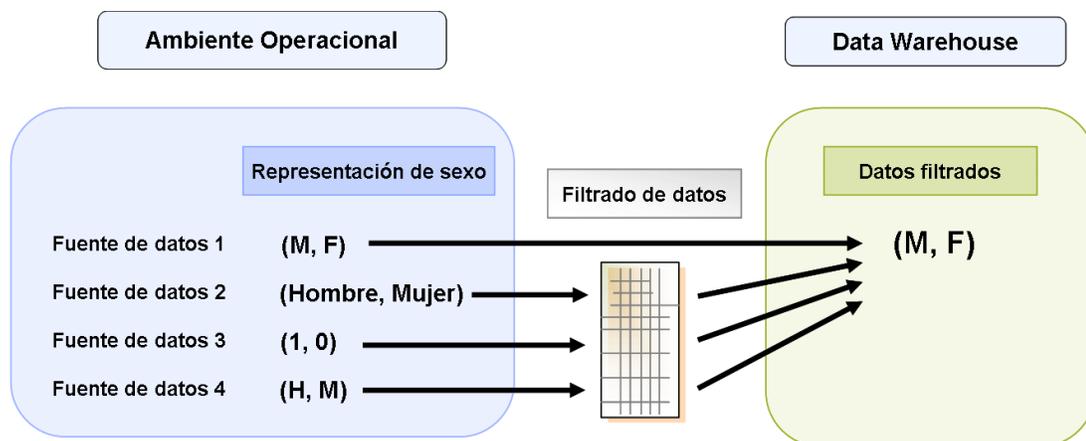


Figura 1. Integración de datos desde los sistemas operacionales hacia la bodega de datos.

### 5.2.3 No volátiles



Significa que los datos que contiene una Bodega de datos no son regularmente accedidos o modificados. Esto es lógico porque el propósito de la bodega es permitir que se analice lo que ha ocurrido.

#### 5.2.4 De tiempo variante

A diferencia de los sistemas de Procesamiento de Transacciones en Línea (OLTP) donde su funcionalidad no depende del cambio sobre los datos en el tiempo (datos históricos), una bodega de datos basa su funcionalidad en la recolección de datos en horizontes de tiempo muy grandes para realizar análisis o pronósticos a través de acontecimientos pasados en la organización.

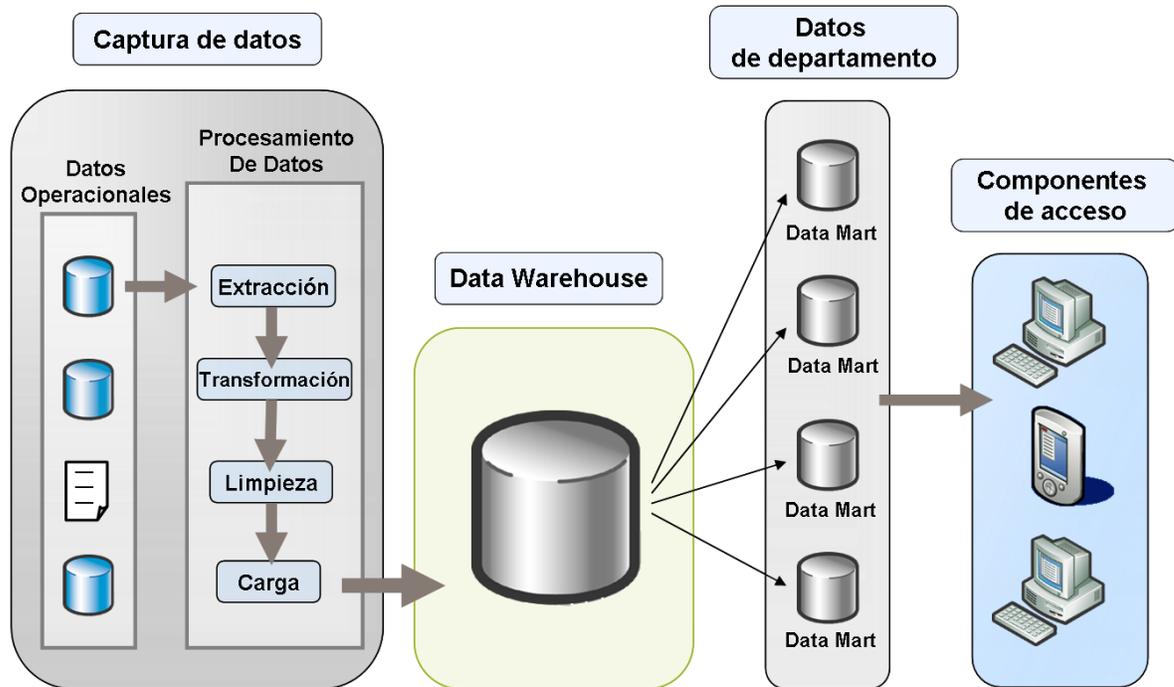
### 5.3 DIFERENCIAS ENTRE UNA BODEGA DE DATOS Y UN SISTEMAS DE PROCESAMIENTO DE TRANSACCIONES EN LÍNEA (OLTP)

Sistema Operacional (OLTP)	Bodega de Datos
Almacena datos actuales	Almacena datos históricos
Almacena datos de detalle	Almacena datos de detalle y datos agregados a distintos niveles
Los datos son dinámicos (actualizables)	Los datos son estáticos
Los procesos (transacciones) son repetitivos	Los procesos no son previsibles
El número de transacciones es elevado	El número de transacciones es menor o medio
Dedicado al procesamiento de transacciones	Dedicado al análisis de datos
Orientado a los procesos de la organización	Orientado a la información relevante
Soporta decisiones diarias (operacionales)	Soporta decisiones estratégicas
Sirve a muchos usuarios (administrativos)	Sirve a técnicos de dirección

**Tabla 2.** Diferencias entre un Sistema Operacional y una Bodega de Datos.

#### 5.4 ARQUITECTURA GENERAL

La arquitectura de la bodega es la organización de cada uno de los componentes que hacen parte del ambiente y se pueden representar de manera general, como se muestra en la Figura 2. Aquí, la arquitectura presenta cuatro capas o divisiones por las que pasan los datos cuando se quiere acceder a información de la bodega de datos.



**Figura 2.** Arquitectura general de una Bodega de Datos.

### 5.4.1 Captura de datos

Se identifican en esta capa las fuentes de datos operacionales (OLTP), las cuales se pueden encontrar en cualquier formato (base de datos, archivo XML, archivos planos, etc.). Son estas fuentes las que soportan las operaciones diarias que se realizan en la organización e incluyen software, componentes de seguridad, redes y bases de datos.

El procesamiento de datos también es conocido como proceso “ETL” (Extract/Transform/Load), y tiene como función encontrar en la(s) fuente(s) de dato(s), los datos que se necesitan en la bodega, manipularlos y prepararlos para su posterior uso.

Durante la extracción, los datos deseados son identificados y extraídos de diferentes fuentes, incluyendo sistemas de bases de datos y aplicaciones. En el proceso de transformación se preparan los datos teniendo en cuenta la forma como se van a almacenar en la bodega de datos. Después de la extracción y la transformación los datos son llevados al sistema destino, es decir, se realiza la carga de datos.

### 5.4.2 Bodega de datos



Esta capa contiene los datos integrados, detallados o ligeramente resumidos necesarios en todo el ambiente de la bodega. Estos datos pueden ser enriquecidos por terceras fuentes de información. De esta capa, el usuario final tomara los datos que quiere cargar en los Data Marts para despliegue y consultas.

### **5.4.3 Data Mart**

Lo Data Marts son conjuntos de información de la bodega de datos para un departamento o un área específica (mercadeo, personal, ingeniería, contabilidad, etc.). Algunas personas usan el termino data mart para referirse a pequeñas bodegas de datos. Ellas, crean muchos data marts para conocer las necesidades de muchos departamentos [12] permitiéndose de esta forma conocer a profundidad las necesidades de cada uno de ellos. Para R. Kimball [13] una bodega de datos no es nada mas que la unión de todos los data marts que los constituyen.

### **5.4.4 Componentes de acceso**

La capa de componentes de acceso esta conformada por los clientes o usuarios finales, los cuales cuentan con un sistema de información analítico para acceder a la información contenida en la bodega de datos. Estos componentes no actualizan datos, solo acceden para “leer” información.

## **5.5 BODEGAS DE DATOS EN ORACLE**

### **5.5.1 Definición**

Oracle, define la bodega de datos como “Una base de datos relacional que esta diseñada para consultas y análisis en lugar de procesamiento de transacciones. Usualmente contiene datos históricos derivados de la transacción de datos o de otras fuentes[11]“. Adicionalmente a las bases de datos relacionales, un ambiente de bodegas de datos incluye una aplicación de extracción, transporte, transformación y carga (ETL), un motor analítico OLAP, herramientas de análisis cliente, entre otras aplicaciones.

### **5.5.2 Diseño lógico y físico en el ambiente de bodega de datos**

En comparación con el diseño físico, el diseño lógico es más conceptual y abstracto. En el diseño lógico se observan relaciones lógicas entre objetos y en el físico, se observa el camino más efectivo para almacenar y recuperar los objetos.

Una de las técnicas que se utiliza para modelar los requerimientos de información lógicos en la organización es el modelado Entidad-Relación. El modelado Entidad-Relación involucra la identificación de cosas importantes (entidades), las propiedades de esas cosas (atributos), y como ellas están relacionadas con otras (relaciones). Una



entidad representa una parte de la información y en las bases de datos relacionales a menudo se mapean directamente a una tabla. Un atributo es un componente de una entidad que ayuda a definir de manera única a la entidad. Este atributo se mapea a una columna de la tabla.

Para estar seguro que los datos sean consistentes, se necesitan utilizar identificadores únicos. Estos identificadores se adicionan a las tablas para poder distinguir un ítem cuando aparezcan en diferentes lugares.

Aunque la técnica ha sido utilizada para el modelado de esquemas relacionales, esta técnica es útil también para el diseño de bodegas de datos en forma de modelado dimensional.

Se puede crear un diseño lógico usando un lápiz y papel, o se puede utilizar una herramienta como Oracle Warehouse Builder (especialmente diseñado para soportar el modelado dimensional y los procesos ETL) u Oracle Designer (herramienta de modelado de propósito general).

El resultado del diseño lógico es: un conjunto de entidades y atributos correspondientes a las tablas de hechos y a las tablas de dimensiones y un modelo de datos dimensional en el esquema de bodegas de datos seleccionado (ej. esquema estrella).

Por otro lado, el diseño físico puede verse de manera general como la creación de la bodega de datos utilizando sentencias SQL. Durante este proceso se traslada el esquema obtenido en el diseño lógico dentro a la base de datos, así:

- Entidades a dimensiones o tablas de hechos.
- Relaciones a restricciones de llave foránea.
- Atributos a atributos de dimensión o a niveles.
- Identificadores de llaves foráneas a restricciones de llaves primarias.
- Identificadores únicos a restricciones de llaves únicas.

Una vez se haya mapeado la estructura del diseño lógico a una estructura física teniendo en cuenta las características de la base de datos, se necesita crear alguna de las siguientes estructuras dentro del motor de base de datos:

- Espacios de tablas.
- Tablas particionadas: Típicamente en una bodega de datos se particiona por fecha de transacciones. Por ejemplo, cada mes.
- Compresión de segmento de datos: Para reducir el espacio en disco y memoria, se puede almacenar tablas o tablas particionadas en un formato comprimido dentro de la base de datos. La compresión de segmentos de datos puede aumentar la velocidad en la ejecución de una consulta de datos.
- Vistas.

- Restricciones de integridad: Las restricciones de integridad en un ambiente de bodegas de datos difiere de las restricciones en un ambiente OLTP. En un ambiente OLTP, sobre todo se previene la inserción de datos inválidos dentro de un registro, lo cual no es un gran problema en un ambiente de bodegas de datos porque la exactitud esta garantizada ya. Es este ambiente, las modificaciones solo son usadas para consultas de reescritura. Usualmente se presentan restricciones NOT NULL.

Para mejorar el desempeño se pueden crear las siguientes estructuras:

- Índices e índices particionados: Los índices de mapas de bits (Bitmaps) son muy comunes en ambientes de bodegas de datos.
- Vistas materializadas: Son los resultados de las consultas que se han almacenado por adelantado, sin requerir de cálculos cuando se ejecuta una sentencia SQL. Desde un punto de vista del diseño físico, una vista materializada se asemeja a una tabla o una tabla particionada y se comporta como un índice.

### 5.5.3 Esquemas de la bodega de datos [11].

Un esquema es una colección de objetos de base de datos, incluyendo tablas, vistas, índices y sinónimos. El modelo de su fuente de datos y los requerimientos de los usuarios ayudaran a diseñar el esquema de la bodega de datos.

- Esquema Estrella

El esquema estrella es el esquema mas simple. Se le denomina “estrella”, porque su diagrama es similar a una estrella, con puntos a cierto radio del centro. El centro de un esquema estrella consiste de una o mas tablas de hechos y los puntos de la estrella son tablas de dimensiones (ver Figura 3).



**Figura 3.** Esquema en estrella.

El camino mas natural para modelar una bodega de datos es el esquema estrella, solo un “join” establece la relación entre la tabla de hechos y una de las tablas de dimensiones. Este tipo de esquemas, optimizan el desempeño manteniendo consultas simples y suministrando tiempos rápidos de respuesta. Toda la información sobre cada nivel es almacenado en una fila.

- Esquema copo de nieve (Snowflake Schema)

Es un esquema más complejo de la bodega de datos pero sigue siendo un esquema estrella. El esquema “copo de nieve”, normaliza las dimensiones para eliminar redundancia. Esto significa que los datos dimensionales han sido agrupados dentro de múltiples tablas en lugar de una gran tabla. Mientras se ahorra espacio, se incrementa el número de tablas de dimensiones y se requieren más “joins” de llaves foráneas. El resultado es más complejo y reduce el desempeño de la consulta (ver Figura 4).

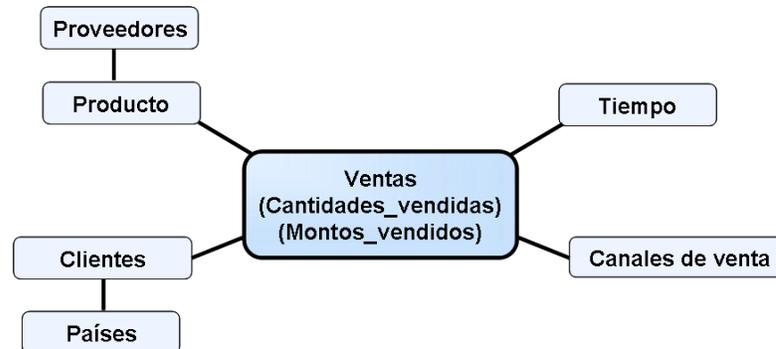


Figura 4. Esquema copo de nieve.

- Tercera forma normal (3FN)

La tercera forma normal es una clásica técnica de modelamiento de bases de datos relacionales que minimiza la redundancia a través de la normalización. Cuando se compara con el esquema estrella, un esquema 3FN tiene un mayor número de tablas debido al proceso de normalización (ver Figura 5).

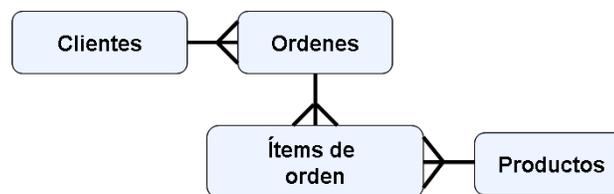


Figura 5. Esquema tercera forma normal.

### 5.5.4 Objetos de la bodega de datos [11]

Los objetos generales de una bodega de datos en Oracle son (ver Figura 6):

- Tabla de hechos: Es una gran tabla dentro de la bodega de datos que almacena medidas de negocio. La tabla de hechos usualmente contiene hechos y llaves foráneas de las tablas de dimensiones; representa datos usualmente numéricos y aditivos que pueden ser analizados y examinados.



Una tabla de hechos usualmente posee dos tipos de columnas: Aquellas que contienen hechos numéricos (también llamados **medidas**) y otra que tiene las llaves foráneas hacia otras tablas.

- Dimensiones: Una dimensión es una estructura, a menudo compuesta de una o más jerarquías. Sus atributos, ayudan a describir el valor de la dimensión y por lo general son de tipo texto. Las dimensiones junto con tablas de hechos, dan la capacidad de responder a preguntas del negocio.
- Jerarquías: Son estructuras lógicas que utilizan niveles como medios para organizar datos. Algunas características de las jerarquías son:
  - Una jerarquía puede ser usada para definir agregación de datos. Por ejemplo, en la dimensión tiempo podría agregar datos desde el nivel MES al nivel TRIMESTRE o al nivel AÑO. Una jerarquía puede ser usada para definir una ruta de navegabilidad y establecer una estructura familiar.
  - Dentro de una jerarquía cada nivel esta conectado lógicamente con el nivel superior e inferior.
  - Una dimensión puede estar compuesta por más de una jerarquía. Ejemplo, la dimensión Productos puede tener una jerarquía por categoría de productos y otra por proveedores.
  - Las herramientas usan jerarquías para permitir ir a través de los diferentes niveles de granularidad.
- Niveles: Un nivel representa una posición en la jerarquía. Los niveles en una dimensión están organizados dentro de uno o más jerarquías.
- Identificadores únicos: Son especificados para un registro distinto en una tabla dimensión.
- Relaciones: La relaciones garantizan la integridad del negocio. Por ejemplo si un negocio vende algo hay obviamente un cliente y un productor. Diseñar una relación entre la información de ventas en la tabla de hechos con la dimensión de productos y clientes refuerza las reblas de negocio en la base de datos.

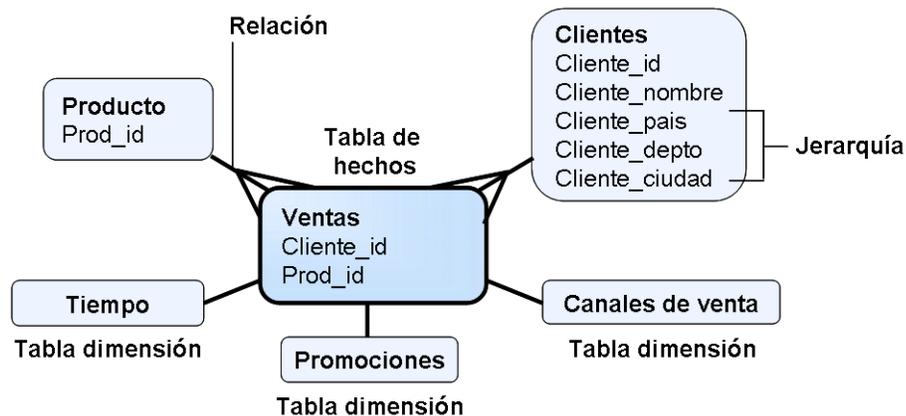


Figura 6 . Objetos típicos para una bodega de datos.

## 6 PROCESAMIENTO ANALITICO EN LINEA (OLAP<sup>3</sup>)

### 6.1 DEFINICIÓN

El termino OLAP fue acuñado por el Dr. E. F. Codd para describir una tecnología que formara un puente entre la computación personal y el manejo empresarial de datos. Los modelos relacionales fallaban al suministrar capacidades analíticas para gerentes, analistas y ejecutivos. Codd reconoció la necesidad de un modelo multidimensional adicional que permitiera análisis más rápidos y de mayor capacidad de las crecientes bases de datos relacionales [12].

Formalmente se puede pensar en OLAP como la actividad general de consultar y presentar datos numéricos y textuales de una bodega de datos [13], o como el proceso interactivo de crear, mantener, analizar y realizar informes sobre datos, añadiendo que los datos en cuestión son percibidos y manejados como si estuvieran almacenados en un “arreglo multidimensional [7].

Algunas características del procesamiento analítico en línea son:

- Cálculos y modelado a través de dimensiones, jerarquías y/o miembros.
- Análisis de tendencias sobre periodos de tiempo.
- Dividir (slicing) subconjuntos de datos para visualizarlos sobre la pantalla.
- Ahondar (drill-down) entre niveles de la jerarquía para obtener mayor detalle de los datos.

<sup>3</sup> Siglas del término en ingles On-Line Analytic Processing.



- Ascender (Drill-up) entre los niveles de la jerarquía para obtener datos resumidos.

## 6.2 CLASIFICACIÓN

Las variaciones de OLAP mas conocidas son ROLAP y MOLAP que se explican a continuación.

*ROLAP (Relational OLAP)*, que puede ser definida como un conjunto de aplicaciones e interfases que le dan a las bases de datos relacionales una tratamiento multidimensional [13], es decir OLAP sobre una base de datos relacional.

Algunas características de ROLAP son [13] [14]:

- Puede manejar gran cantidad de datos: La limitación del tamaño de datos en la tecnología ROLAP depende de los límites de la cantidad de datos que la base de datos pueda manejar.
- Puede apoyar funcionalidades inherentes en la base de datos relacional: A menudo, las bases de datos relacionales ya vienen con un contenedor de funcionalidades. Las tecnologías ROLAP, desde que se ubicaron sobre las bases de datos relacionales, pueden apoyarse en estas funciones.
- El desempeño puede ser bajo: Porque cada reporte ROLAP es esencialmente una consulta SQL (o múltiples consultas SQL), la duración de la consulta puede ser alta si el tamaño de la base de datos es grande.
- Limitado para funcionalidades SQL: La tecnología ROLAP trabaja principalmente con la generación de sentencias SQL para consultar la base de datos relacional. Por lo general las sentencias SQL no cubren con todas las necesidades de análisis, por consiguiente es necesario que la base de datos cuente con extensiones especiales del lenguaje SQL para el manejo de OLAP relacional.

*MOLAP (Multidimensional OLAP)*, suministra capacidad de análisis de datos almacenados con un sistema de bases de datos multidimensional (MDBS) [12]. Algunas características son [13] [14]:

- Mayor desempeño: Los cubos MOLAP son construidos para realizar una rápida recuperación de datos.
- Limitación en la cantidad de datos que puede manejar: Porque todos los cálculos son ejecutados cuando se construye el cubo, no es posible incluir una gran cantidad de datos dentro del mismo.
- Requieren inversión adicional: La tecnología es usualmente propietaria y aun no es muy común en las organizaciones. Por consiguiente, adoptar la tecnología MOLAP, involucra necesidades adicionales de inversión de recursos humanos y económicos.



## 6.3 OLAP EN ORACLE

Oracle permite dentro de su mismo motor de base de datos el manejo de *ROLAP* y *MOLAP*. Para el caso de *ROLAP*, Oracle extiende el lenguaje de consulta SQL para ejecutar funciones de agregación y análisis más robustas. Algunas funciones analíticas adicionadas como RANK, LAG/LEAD, FORST/LAST y operadores de agregación como CUBE y ROLLUP permiten a los usuarios desarrollar cálculos complejos usando consultas SQL estándar. Para mejorar el tiempo de respuesta, se usan tecnologías como Query Rewrite y vistas materializadas para construir resúmenes de datos e incrementar la velocidad de las consultas.

Para *MOLAP* Oracle incluye [15]:

- *Motor de cálculo OLAP*. Soporta selección y cálculos rápidos de datos multidimensionales. El estado de una sesión individual en el motor es persistente para soportar series de consultas típicas de aplicaciones analíticas. La salida o resultado de una o mas consultas, se pueden utilizar como entrada de otra consulta. El motor OLAP corre dentro del kernel de Oracle.
- *Espacios de trabajo analíticos (Analytic Workspace)*. Almacenan objetos de datos multidimensionales y procedimientos escritos en OLAP DML. Dentro de una base de datos sencilla se pueden crear y compartir muchos espacios de trabajo analíticos. Similar a las tablas relacionales, un espacio de trabajo analítico es propio de un usuario en particular y se puede conceder acceso a otros usuarios. Un ambiente de trabajo puede ser temporal (esto es, durante la sesión) o puede ser permanente (guardado) para una próxima sesión. Cuando un espacio de trabajo analítico es persistente, los datos son almacenados como registros LOB's en las tablas de la base de datos.
- *Lenguaje de manipulación de datos multidimensionales OLAP DML*. Es un lenguaje de manipulación de datos con el que trabaja el motor de cálculo OLAP de Oracle. Extiende las capacidades analíticas de los lenguajes de consulta como SQL y del OLAP API incluyendo pronósticos, modelado y escenarios "que pasa si...". Los desarrolladores pueden crear procedimientos almacenados que usan condiciones lógicas y librerías extendidas de funciones y comandos DML para ejecutar análisis de datos complejos.

Los comandos y funciones OLAP DML incluyen las siguientes categorías.

- Agregación.
- Asignación.
- Selección de datos.
- Operaciones con datos y fechas.



- Escritura y lectura de archivos.
- Operaciones financieras.
- Pronósticos y regresiones.
- Manipulación de números.
- Modelos.
- Operaciones estadísticas
- Manipulación de textos.
- Manipulación de series de tiempo.

Al usar OLAP DML, los administradores de bases de datos y desarrolladores de aplicaciones pueden crear objetos de datos multidimensionales que son almacenados dentro de un espacio de trabajo analítico. EL OLAP DML opera sobre datos que están almacenados (temporal o permanentemente) en estos objetos multidimensionales

- *Funciones de tabla SQL.* permiten tomar un conjunto de filas como entrada y producir un conjunto de filas como salida que pueden ser consultadas como a una tabla física convencional de la base de datos. De esta manera, el motor de cálculo y otras fuentes de datos multidimensionales son accesibles para el SQL convencional.
- *OLAP API.* Es una interfase de programación de aplicaciones para OLAP Oracle. Es decir, un lenguaje de consulta que selecciona y manipula datos para desplegarlos en un cliente java. Orientada a objetos y construida completamente en java.
- *Catalogo de metadatos OLAP.* Los metadatos se definen típicamente como "datos sobre datos". Los metadatos del catalogo OLAP son creados y almacenados en tablas relacionales dentro de la base de datos. Las aplicaciones OLAP pueden consultar el repositorio de estos metadatos para encontrar que datos están disponibles para análisis y despliegue. Los metadatos contienen información acerca de la localización física de los datos, es decir, si los datos son almacenados en tablas relacionales o en espacios de trabajo analítico.

### 6.3.1 Decidiendo cuando usar ROLAP o MOLAP en Oracle [16]

Para implementar una solución OLAP en Oracle, se debe crear un almacén de datos usando uno de los siguientes diseños:

- *Esquema Multidimensional:* Un espacio de trabajo analítico.
- *Esquema Relacional:* Un esquema estrella definido en el catalogo OLAP.

Para muchos administradores de bases de datos, las tablas relacionales y el SQL representan un ambiente de desarrollo familiar. Por otro lado, los espacios de trabajo



analíticos requieren transformación de datos y aprender nuevos conceptos. ¿Entonces porque utilizar espacios de trabajo analítico? La respuesta es simple: las grandes capacidades de análisis y el desempeño en tiempo de ejecución representan en muchos casos el mejor escenario para el soporte a decisiones. Sin embargo, los esquemas relacionales son la mejor decisión en algunas situaciones, por ejemplo, si el almacén de datos soporta la creación de reportes rutinarios con carácter exploratorio, y ese uso aparenta ser estable.

Los espacios de trabajo analíticos son una buena opción para los siguientes requerimientos:

- Consultas ad-hoc o no genéricos de todas las áreas de datos
- Cálculos avanzados como modelos, pronósticos, radios de crecimiento y tendencias
- Escenarios “que pasa si”
- Cálculos en series de tiempo como retardos, elementos principales, promedios de movimiento.
- Cálculos complejos en tiempo de ejecución
- Alto desempeño para calcular resúmenes de datos.

Los esquemas relacionales se pueden preferir si se tienen los siguientes requerimientos:

- Patrones de consulta predecibles y reportes preparados.
- Cálculos no avanzados
- Cálculos complejos en tiempo de ejecución infrecuentes
- Medidas con un gran numero de dimensiones
- Dimensiones con pocos niveles de agregación.

## **7 METODOLOGIA DE DESARROLLO.**

### **7.1 METODOLOGÍA DE DESARROLLO DE LA BODEGA DE DATOS**

Se encontraron varios autores quienes proponen metodologías para el desarrollo de una bodega de datos; dentro de ellos R. Kimball [13], brinda conceptos necesarios para llevar a cabo el análisis y diseño de la bodega de datos planteada para Unicauca Virtual. El método presentado por Kimball además de estar extensamente expuesto en su bibliografía, se apoya en gran variedad de ejemplos que facilitan la comprensión de los conceptos relacionados con diseño de bodegas de datos.

Existen otras metodologías que aunque no son tan conocidas podrían convertirse en buenas alternativas a la metodología propuesta por R. Kimball, como por ejemplo la

propuesta por Sergio Lujan Mora y Juan Trujillo denominada “Un método global basado en UML para el diseño de Almacenes de Datos” [16] pero que aun carecen de reconocimiento practico o se encuentran en fase de desarrollo.

La metodología de Kimball [13], presenta un marco de trabajo ilustrado en la Figura 7 en la cual se muestran las diferentes etapas durante todo el proceso de creación de la bodega.

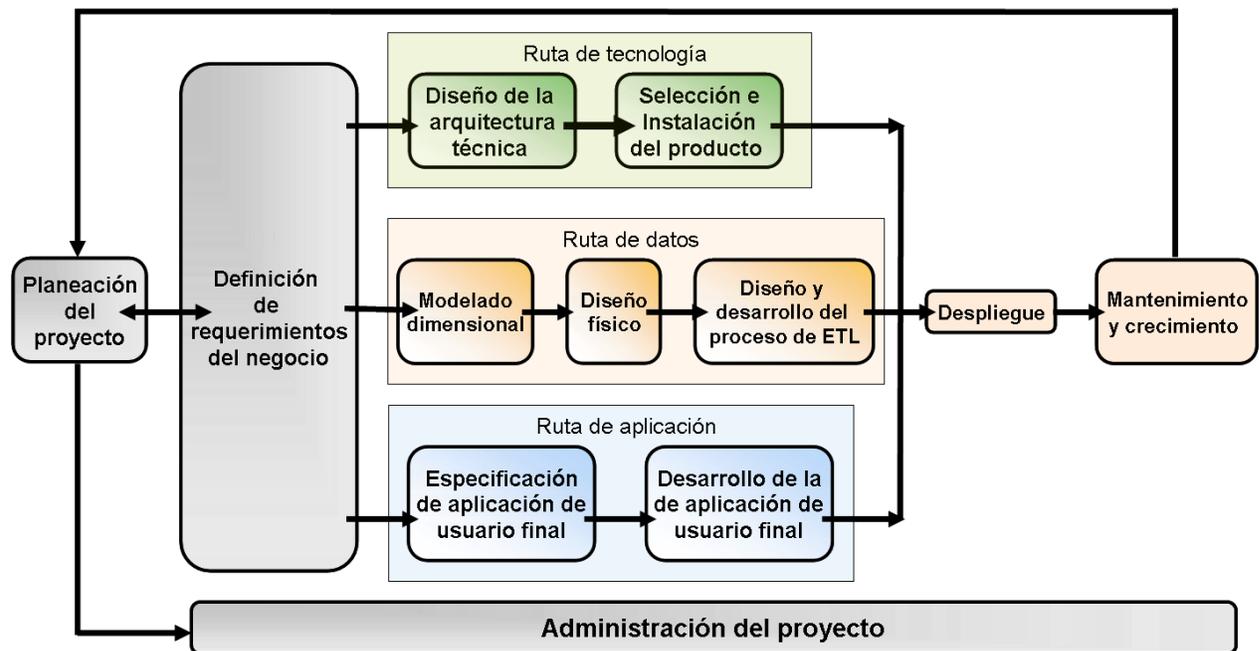


Figura 7. Ciclo de vida para la construcción de una bodega de datos según R. Kimball

### 7.1.1 La fase de planeación del proyecto

La planeación pretende establecer la definición y el alcance del proyecto de la bodega de datos, incluyendo la valoración y justificación del negocio. Estas tareas son importantes ya que permiten obtener una visión general inicial del proyecto. Desde aquí, la planeación del proyecto se enfoca en los recursos y los requerimientos de personal asociado a una tarea con su duración y secuencia. El plan del proyecto, identifica todas las tareas asociadas con el *ciclo de vida dimensional* y señala todas las personas involucradas. La planeación del proyecto depende de los *requerimientos del negocio* como se muestra en la Figura 7 denotado por una flecha entre estas dos actividades.

### 7.1.2 Definición de requerimientos del negocio

La probabilidad de éxito de una bodega de datos se incrementa a medida que se incrementa el entendimiento de los usuarios finales y sus requerimientos. Sin este



entendimiento la bodega de datos podría convertirse en un ejercicio sin resultados para el equipo del proyecto. La forma usada para obtener el conocimiento de las personas que hacen un trabajo analítico difiere de la forma como se hace comúnmente la obtención de requerimientos. Los diseñadores de la bodega de datos deben entender los factores claves que direccionan las actividades del negocio para determinar efectivamente los requerimientos y traducirlos en consideraciones de diseño. Los *requerimientos del negocio* establecen la base para tres rutas paralelas enfocadas en la tecnología, datos y aplicaciones de usuario final.

- **Ruta de datos**

- *Modelado dimensional*: la definición de los requerimientos del negocio determinan los datos que se necesitan para direccionar los requerimientos de los usuarios analíticos. El diseño de modelos para soportar estos tipos de requerimientos, requiere de una aproximación diferente a la utilizada en el modelamiento de sistemas transaccionales convencionales. Se inicia con una construcción de una *matriz* que representa los procesos claves de los procesos del negocio y su dimensionalidad.

Desde allí, se conduce un análisis de datos mas detallado de las fuentes operacionales que sirven como fuente del sistema. Juntando esta etapa de análisis con el entendimiento previo de los requerimientos del negocio, se genera el modelado dimensional. Este modelo identifica la granularidad de las tablas de hechos, las dimensiones asociadas, los atributos, las jerarquías y los hechos. Este conjunto de actividades termina con el desarrollo de un plan inicial de agregación.

- *Diseño físico*: El diseño físico de la bodega de datos se enfoca en definir las estructuras físicas necesarias para soportar el diseño lógico. Dentro de las primeras actividades que se deben realizar se encuentra definir un estándar de nombrado y configurar todo el ambiente de la bodega de datos. Se deben determinar estrategias primarias de indexación y particionamiento.
- *Diseño y desarrollo del proceso ETL*: esta actividad es una de las más subestimadas. Cuenta con tres pasos importantes: la extracción, la transformación y la carga de datos. El proceso de extracción siempre señala aspectos relacionados con la calidad de los datos que se encuentran en las fuentes de datos operacionales. Debido al gran impacto que tiene la calidad de los datos en la credibilidad de la bodega de datos, se deben manejar todos los aspectos relacionados con la calidad durante la etapa ETL.

- **Ruta de tecnología**

- *Diseño de la arquitectura técnica*: en el ambiente de las bodegas de datos requieren la integración de numerosas tecnologías. El diseño de la arquitectura



técnica establece una visión general de la arquitectura y requiere considerar tres factores de manera simultánea: los requerimientos del negocio, el ambiente técnico actual y las directrices técnicas del plan estratégico. El diseño de la arquitectura es importante para mejorar aspectos como:

*La comunicación.* La arquitectura es una excelente herramienta de comunicación entre muchos niveles de la organización. Suministra un mecanismo para ayudar a la gerencia en el entendimiento de la magnitud y la complejidad del proyecto. También ayuda a los desarrolladores a ubicarse dentro del proyecto y a reconocer cuales son sus responsabilidades.

*Flexibilidad y mantenimiento:* la creación de la arquitectura ayuda a anticipar aspectos relacionados con la construcción del sistema que permitan construirlo sin mayores inconvenientes.

*Aprendizaje:* la arquitectura hace parte de la documentación del sistema y puede ayudar a que nuevos miembros del equipo conozcan de una forma más rápida los componentes, los contenidos y las conexiones.

- *Selección del producto e instalación:* usando el diseño de la arquitectura como marco de trabajo se deben evaluar y seleccionar los componentes arquitectónicos como la plataforma hardware, el sistema gestor de base de datos, la herramienta ETL y las herramientas de acceso.

- **Ruta de aplicación.**

- *Especificación de la aplicación de usuario final:* es necesario definir un grupo de usuarios finales de la aplicación más o menos estándar ya que no todos los usuarios del negocio necesitan el mismo tipo de acceso a la bodega de datos. Las especificaciones de la aplicación describen las plantillas de los reportes y los requerimientos de cálculos. Estas especificaciones aseguran que el equipo de desarrollo y los usuarios tengan un entendimiento común de la herramienta que se va a emplear.
- *Desarrollo de la aplicación de usuario final:* después de la especificación de la aplicación, se debe configurar los metadatos de la herramienta y construir los reportes especificados. Estas aplicaciones, por lo general, ofrecen mecanismos para que los usuarios finales modifiquen las plantillas de reportes existentes.

### 7.1.3 Despliegue

El despliegue involucra poner la tecnología, los datos y las aplicaciones de usuario final accesibles desde el escritorio de los usuarios. Para esto es necesaria una planeación extra para garantizar que todas las piezas encajen y funcionen correctamente. Además,



se deben establecer estrategias de comunicación y realimentación por parte de los usuarios finales antes de que cualquiera tenga acceso a la bodega de datos.

#### **7.1.4 Mantenimiento y crecimiento**

Gran parte del trabajo surge después del despliegue inicial de la bodega de datos. Es necesario continuar enfocándose en los usuarios del negocio para poder suministrarles soporte y educación. También es necesario poner atención en las operaciones base de la bodega para asegurarse que los procesos y procedimientos se ejecutan correctamente.

Si se ha seguido todo el ciclo de desarrollo de la bodega de datos hasta este punto, la bodega de datos esta obligada a crecer. Al contrario de los desarrollos de sistemas tradicionales, un cambio se puede ver como un a señal de éxito, no de fallo. Para el correcto manejo del cambio y el crecimiento se deben establecer procesos de priorización para tratar con las nuevas demandas de los usuarios. Después de definir las prioridades, se regresa al inicio del ciclo de desarrollo y se construye sobre lo que ya esta establecido en el ambiente de la bodega enfocándose en los nuevos requerimientos.

#### **7.1.5 Administración del proyecto**

La administración del proyecto garantiza que las actividades de ciclo de vida dimensional permanezcan sincronizadas y encaminadas. Como se muestra en la Figura 7, las actividades de la administración del proyecto se realizan durante todo el ciclo de vida. Estas actividades se enfocan en el monitoreo del estado del proyecto y los controles sobre cambios. Para finalizar, la administración incluye la elaboración de un plan de comunicación del proyecto comprensivo que direcciona al equipo de sistemas de información y a las personas del negocio.

## **7.2 METODOLOGIA DE DESARROLLO DEL PROTOTIPO DE HERRAMIENTA OLAP**

Para el desarrollo del prototipo OLAP se utilizo el Proceso de Desarrollo Unificado (RUP), definido como proceso de desarrollo de software que ayuda a asignar tareas y responsabilidades al equipo de desarrollo, y además sirve como guía del uso del lenguaje de modelado unificado UML [17]. Fue propuesto por Ivar Jacobson, Grady Booch y James Rumbaugh y presenta las siguientes características:

- Guiado o manejado por casos de uso.
- Centrado en la arquitectura.
- Iterativo e incremental.
- Desarrollo basado en componentes.



- Utilización de un único lenguaje de modelamiento (UML).
- Proceso integrado.

El proceso puede ser descrito en dos dimensiones o dos ejes. La primera dimensión representa el tiempo y muestra el aspecto dinámico del proceso. Se expresa en términos de ciclos, fases e iteraciones. La segunda dimensión o eje, representa el aspecto estático del proceso, es decir las actividades y los artefactos.



---

# PARTE II – BODEGA DE DATOS

---



---

# CAPITULO 3 – DEFINICION DEL PROYECTO Y RECOLECCION DE REQUERIMIENTOS

---

El ciclo de vida dimensional del negocio [13], describe la secuencia de tareas de alto nivel requeridas para un diseño, desarrollo y despliegue efectivo de la bodega de datos, como se describe en el capítulo anterior (Ver Figura 7).

Para la construcción de la bodega de datos en Unicauca Virtual, se siguieron algunas de las etapas que se incluyen dentro del ciclo de vida, logrando los siguientes resultados:

## 8 PLANEACION DEL PROYECTO

El objetivo de esta etapa es cumplir con dos aspectos importantes:

- El primero de ellos denominado “la definición del proyecto”, con el cual se pretende definir el escenario en el que se encuentra la organización, el alcance y la justificación del proyecto.
- El segundo aspecto, “la planeación del proyecto”, en el cual se especifica cada una de la actividades que se deben cumplir durante el proceso de construcción de la bodega.

A continuación, se presentan los resultados obtenidos en esta etapa.

### 8.1 DEFINICION DEL PROYECTO

Un punto importante a tener en cuenta, antes de proceder con el desarrollo completo de la bodega de datos, fue identificar el escenario en cual se encontraba la organización (Universidad del Cauca) y evaluar la disposición que tenía la misma para apoyar el proyecto. Además, se realizó un análisis de requerimientos del negocio de alto nivel, con el objetivo de dar claridad al proyecto y servir como base en la definición del alcance y la justificación de la bodega de datos.

#### 8.1.1 Escenario de la organización



Para poder identificar en cual escenario se encontraba la Universidad del Cauca, que es la organización encargada del desarrollo del macroproyecto Unicauca Virtual (incluyendo el proyecto de bodega de datos), debemos explicar alguno de los escenarios más comunes tratados por los expertos [13] en el desarrollo de este tipo de proyectos. Los tres escenarios más comunes en los cuales se puede ubicar a una organización con iniciativas de construir una bodega de datos son:

- *Demanda desde un fanático del negocio*, en el cual un ejecutivo del negocio, con posibilidades de patrocinar el proyecto, tiene una visión de obtener mejor acceso a una mejor información para tomar mejores decisiones.
- El segundo escenario llamado *mucha demanda*, ligeramente mas complejo que el anterior, es donde múltiples ejecutivos del negocio expresan la necesidad de obtener mejor información de los sistemas operacionales.
- *En busca de demanda*, como ultimo escenario, el cual se puede ver como la demanda que se genera desde uno de los Jefe de los Sistemas de Información (CIO por sus siglas en ingles), siendo uno de los escenario mas difícil de tratar.

Basados en lo anterior la Universidad del Cauca se ubicó dentro de un escenario donde existía “*poca demanda de obtener mejor información*”, que pudiera servir como base para un buen análisis de los datos contenidos en los sistemas operacionales sobre los cuales se sustenta. Es decir, que los tutores y directivos del ambiente de aprendizaje en línea, que son algunos de los actores de la plataforma, no tenían una visión acerca de la información que pudiera ser útil para analizar y tomar mejores decisiones relacionadas con el aprendizaje del estudiante.

Identificar el escenario en el cual se encuentra la organización, permite a los desarrolladores de la bodega de datos evaluar cuan difícil es tratar con dicho escenario durante las etapas de construcción.

### **8.1.2 Definición de requerimientos**

El siguiente paso, después de evaluar el escenario, fue identificar los requerimientos del negocio, donde se hizo necesario aplicar las dos técnicas propuestas por R. Kimball [13] en dicha tarea.

La primera técnica utilizada en la tarea de recolección de requerimientos fue la “*entrevista*”, en la cual se formularon una serie de preguntas con el fin de obtener un entendimiento general del negocio. Esta técnica se desarrollo a través de los siguientes pasos:

El primer paso, antes de aplicar la entrevista, fue identificar a las personas quienes serian entrevistadas, es decir, a ejecutivos y personas del sistema de información que



podrían brindar la información necesaria para la construcción de la bodega de datos. Debido a que el proyecto tendría que ir a la par con el desarrollo de los demás proyectos de la fase II de Unicauca Virtual, se dio una condición de construcción especial, en donde se tuvo que elegir como el ejecutivo del negocio al director del macroproyecto y a los desarrolladores de los otros proyectos como las personas encargadas del sistema de información.

El segundo paso fue redactar cada una de los cuestionarios que serían utilizados en cada entrevista. Estas preguntas fueron tomadas y adaptadas de las guías propuestas por R. Kimball [13] en su libro. Los cuestionarios se especifican en el ANEXO 1.

Como último paso, se empezó a realizar cada una de las entrevistas con el personal seleccionado. La primera entrevista se llevó a cabo con el ejecutivo del negocio, quien por su parte trató información relacionada con el desarrollo de los cinco (5) proyectos de la segunda fase de Unicauca Virtual y no acerca del ambiente de aprendizaje en línea como negocio, que era el objetivo de la entrevista. El resultado de esta entrevista se encuentra en el ANEXO 2.

La mayoría de las respuestas dadas en esta entrevista no se ajustaron a los objetivos que cada una de las preguntas tenía en pro de la definición de requerimientos para el desarrollo de la bodega de datos. Por esta razón, el grupo de desarrollo del proyecto decidió no realizar las demás entrevistas creyendo obtener resultados similares.

Como una solución al problema de recolección de requerimientos presentado, se propuso realizar “*sesiones de recolección requerimientos*”<sup>4</sup> [13] (segunda técnica para la recolección de requerimientos). Con cada una de estas sesiones realizadas con todo el grupo a cargo de Unicauca Virtual (directivos y grupos de trabajo de los demás proyectos), se obtuvo información importante del negocio y de esta manera se pudo obtener los requerimientos y definir un alcance y una justificación para el proyecto.

Aunque con la recolección de requerimientos se definieron los “*temas de interés*” [8] sobre los cuales estaría orientada la bodega de datos, fue necesario priorizarlos para elegir los más viables y que potencialmente impactarán más sobre el negocio.

Los temas de interés que se quería pretender atender con la iniciativa de la bodega de datos fueron:

- **Estilos de aprendizaje:** este tema fue seleccionado debido a que los estilos de aprendizaje son el fundamento sobre el cual se realiza la personalización del contenido presentado al estudiante. De esta manera se considero importante

---

<sup>4</sup> Sesiones programadas que permiten reunir a un gran número de personas relacionadas con el proyecto, alentando la lluvia de ideas creativas con el objetivo de definir los temas del negocio a tratar en la bodega de datos.



incluir este tema, para poder analizar la información obtenida de la relación que hay entre ellos y el estudiante.

- **Evaluación de actividades:** la evaluación fue un área que se consideró importante para ser tratada, debido a que es una métrica para medir el rendimiento de los estudiantes dentro del ambiente de aprendizaje virtual y sería útil analizar, de acuerdo a su rendimiento académico, el impacto que tiene aplicar determinados estilos de aprendizaje a un estudiante.
- **Interacción del estudiante dentro de Unicauca Virtual:** el estudiante es el principal actor dentro del sistema, por esto se debía tener en cuenta las actividades que llevaría a cabo el estudiante durante su proceso de aprendizaje. Estas actividades se monitorean con una serie de indicadores con el objetivo de personalizar el contenido académico del estudiante.
- **Éxitos y deserciones de los estudiantes durante su proceso académico:** muchos de los encargados de los ambientes de aprendizaje virtual querían realizar análisis sobre los eventos que causan deserciones dentro del ambiente de aprendizaje en línea.
- **Comportamiento financiero del ambiente virtual:** otro aspecto importante a tener en cuenta en cualquier institución es su comportamiento financiero. En algunas circunstancias este puede impactar negativamente en el crecimiento de la organización y se creyó que sería útil poder analizar su conducta con el objetivo de evitar problemas que se pudieran presentar.

### 8.1.3 Priorización de los temas de interés para la bodega de datos

Después de haber realizado un análisis detallado del impacto y la factibilidad que cada uno de los temas puede tener sobre el negocio, se realiza la priorización de requerimientos [13] como se muestra en la grafica, ver Figura 8.

Los temas que se ubican en el cuadrante superior derecho de la figura, *estilos de aprendizaje, la evaluación de actividades y la interacción del estudiante en Unicauca Virtual*, representaron para el proyecto el objetivo principal de análisis de información. La razón por la cual se eligieron estos temas, como base para el desarrollo de la bodega de datos, fue porque se podía obtener la información necesaria del sistema operacional (OLTP) resultado de la fase II de Unicauca Virtual como soporte para definir los diferentes esquemas multidimensionales.

Las áreas denominados *comportamiento financiero del ambiente virtual y éxitos y deserciones de los estudiantes durante su proceso educativo* no se tuvieron en cuenta en el desarrollo de la bodega de datos.

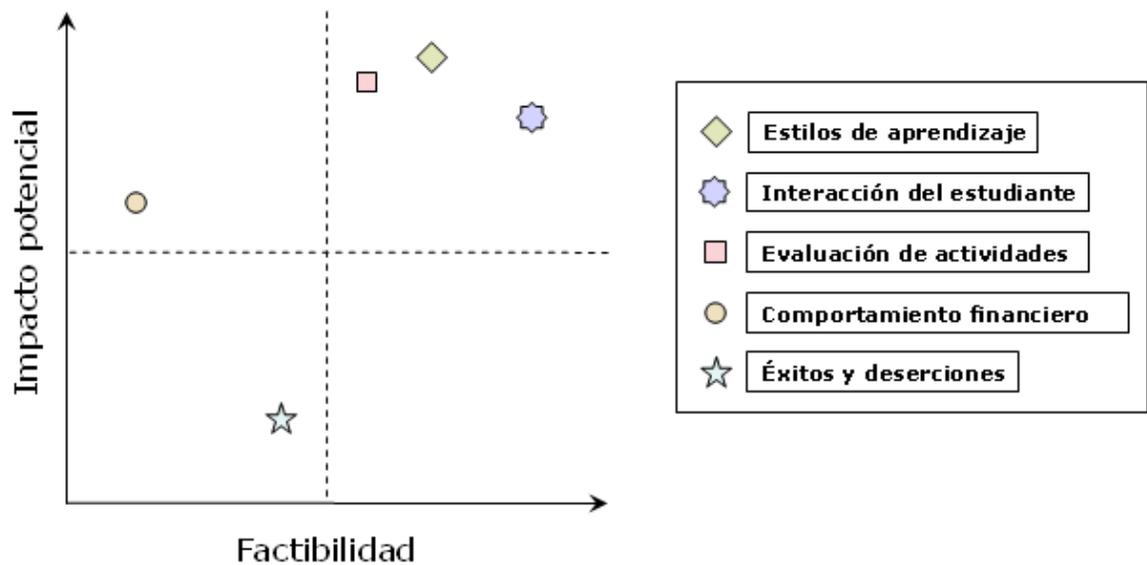


Figura 8. Priorización de requerimientos del negocio.

#### 8.1.4 Alcance de la bodega de datos

Con la definición de los temas a tratar en el proyecto se pudo especificar un alcance para cada uno de ellos, de la siguiente manera:

- **Estilos de aprendizaje:** definir los modelos dimensionales necesarios que permitan recuperar información, relacionada con el cambio en los diferentes estilos de aprendizaje aplicados al estudiante en Unicauca Virtual a través del tiempo. Además, se quiere dar a conocer el impacto que tiene un determinado recurso (SCO ó ASSET)[2] a un estilo de aprendizaje.
- **Interacción del estudiante en Unicauca Virtual:** definir los modelos dimensionales necesarios que permitan recuperar información relacionada con el comportamiento del estudiante en base a una serie de indicadores (tiempo de visualización del recurso, tiempo de descanso en el recurso y número de veces de utilización del recurso), es decir, de esta manera seguir las acciones de los estudiante sobre un recurso determinado a través del tiempo.
- **Evaluación de actividades:** definir los modelos dimensionales necesarios que permitan recuperar información relacionada con los resultados de las evaluaciones presentadas por los estudiantes de las actividades cursadas, de esta manera, poder contar con información relacionada con el grado de comprensión de ciertas actividades.



### 8.1.5 Justificación de la bodega de datos

El proyecto de bodegas de datos esta orientado al desarrollo de modelos dimensionales que permitan obtener información relevante sobre el proceso de aprendizaje del estudiante en Unicauca Virtual, es decir, ayuda en la consecución de los beneficios sociales de Unicauca Virtual al suministrar un mecanismo que apoya la gestión realizada por los tutores y directivos, ofreciendo posibilidades de obtener mayor información sobre el comportamiento de los estudiantes en el LMS y de este modo contribuir en el mejoramiento de su proceso de aprendizaje y en su formación integral como futuro profesional.

Este proyecto por estar enmarcado en un macroproyecto, cuenta con los recursos económicos y de infraestructura tecnológica necesaria para su completo desarrollo, debido al patrocinio que Colciencias y la Universidad del Cauca han dado al proyecto.

## 8.2 PLANEACION DEL PROYECTO

Antes de comenzar con la planeación del proyecto de bodegas de datos es importante definir dos aspectos relevantes: el primero es especificar un nombre para el proyecto que lo identifique a través del desarrollo y el segundo, seleccionar el personal que estará a cargo del proyecto identificando sus respectivos roles.

### 8.2.1 Identidad del proyecto

En todo proyecto es necesario e importante definir un nombre como identidad del mismo. El nombre dado para el proyecto fue *“Bodega de datos y OLAP en Unicauca Virtual”*, para el cual se asignó el acrónimo *“BOUV”*.

### 8.2.2 Personal en el proyecto

Un proyecto de bodegas de datos requiere de personas con diferentes roles y habilidades durante todo su ciclo de vida. Estos roles pueden ser representados por personas del negocio o por personas que pertenecen a la comunidad de sistemas de información de la organización.

Los roles que se identificaron en el proyecto para soportar su desarrollo fueron:

- **Administrador del proyecto:** este rol fue desempeñado por el magíster Miguel Ángel Niño, director de la segunda fase de Unicauca Virtual, quien estaba a cargo de la coordinación de recursos, el seguimiento y la comunicación de progresos y resultados de los demás subproyectos. Esta persona estaba en constante comunicación con el líder o director del proyecto, con el objetivo de informar sobre las decisiones tomadas por los directivos de Unicauca Virtual y que de alguna forma afectaban el desarrollo de la bodega de datos.



- **Líder o director del proyecto:** este papel fue desempeñado por la magíster Martha Eliana Mendoza, quien tenía a su cargo realizar la planeación y seguimiento del proyecto, revisión y validación de los productos y la comunicación entre los desarrolladores y los directivos de Unicauca Virtual. Además, apoyar cada uno de los pasos de construcción del proyecto, desde el entendimiento de los requerimientos hasta la puesta en marcha.
- **Analista del negocio:** este rol fue ejecutado por el líder y autores del proyecto (Jaime Alberto Muñoz C y Lisandro Acosta Mejía). Estas personas fueron los responsables de definir y representar los requerimientos del negocio en modelos dimensionales.
- **Modelador de datos:** esta tarea estuvo a cargo de los autores del proyecto (Jaime Alberto Muñoz C y Lisandro Acosta Mejía), quienes realizaron todos los diseños de los modelos dimensionales para los temas del negocio elegidos.
- **Administrador de la bodega de datos:** este papel fue desempeñado por los dos autores del proyecto (Jaime Alberto Muñoz C y Lisandro Acosta Mejía), quienes tuvieron que trasladar los modelos dimensionales diseñados dentro de las estructuras físicas de la base de datos Oracle 10g (implementación de la bodega de datos).
- **Diseñador y desarrollador de la capa ETL:** estos dos roles fueron ejecutados por los dos autores del proyecto, quienes debían definir y construir los programas encargados de realizar la carga de datos del sistema operacional (OLTP) a la bodega de datos.
- **Desarrolladores de aplicaciones de usuario final:** el desarrollo del prototipo de herramienta OLAP, fue efectuado por los dos encargados del proyecto. El compromiso de estas personas fue construir una aplicación que permitiera visualizar la información contenida en la bodega de datos.
- **Educador de la bodega de datos:** los usuarios del negocio finales deben ser educados sobre el contenido de la bodega de datos y sobre el uso del prototipo de herramienta OLAP, por esta razón, los autores del proyecto se encargaron de realizar una breve capacitación de cada uno de los productos obtenidos.

### 8.2.3 Plan del proyecto

La Tabla 3, muestra la lista de actividades planeadas para la construcción de la bodega de datos.



LEYENDA:	Administrador del proyecto	Director del proyecto	Analista del negocio	Modelador de datos	Administrador de la bodega de datos	Diseñador y desarrollador de la capa ETL	Desarrollador de la aplicación de usuario	Educador de la bodega de datos
<b>ADMINISTRACION DEL PROYECTO Y REQUERIMIENTOS</b>								
<b>DEFINICION DEL PROYECTO</b>								
1. Definición del escenario para el desarrollo del proyecto.			●					
2. Desarrollo del alcance del proyecto.		○	●					
3. Construcción de la justificación del proyecto.		○	●					
<b>PLANEACION DEL PROYECTO Y ADMINISTRACIÓN</b>								
1. Establecer la identidad del proyecto.			●					
2. Identificar roles para el proyecto.		○	●					
3. Definir el plan preeliminar para el desarrollo de la bodega de datos y la construcción del prototipo de herramienta OLAP.		○	●					
<b>DEFINICION DE REQUERIMIENTOS DEL USUARIO</b>								
1. Identificar y preparar al equipo de la entrevista.		○	●					
2. Seleccionar a los entrevistados.		○	●					
3. Realizar las entrevistas programadas.	▶	○	●					
4. Analizar los resultados obtenidos en las entrevistas.	□	●	●					
5. Definir los temas de interés como requerimientos para la bodega de datos.	○	●	●					
6. Priorizar los requerimientos del negocio.	□	●	●					
<b>DISEÑO DE DATOS</b>								
<b>MODELADO DIMENSIONAL</b>								
1. Construcción de la matriz bus.		□	○	●				
2. Escoger los data marts.		□	○	●				
3. Declarar el grano de las tablas de hecho.		□	○	●				
4. Escoger las dimensiones conformadas.	▲	□	○	●				



5. Desarrollar los diagramas dimensionales.		○	○	●				
6. Documentar detalladamente las tablas de hechos.		□	○	●				
7. Documentar detalladamente las dimensiones.	▲	○	○	●				
<b>IMPLEMENTACION</b>								
<b>DISEÑO DE LA BASE DE DATOS FISICA</b>								
1. Definir estándares.	▲	○	◐	◐	●			
2. Diseñar las columnas y tablas físicas.		□	◐	◐	●			
<b>IMPLEMENTACION BASE DE DATOS FISICA</b>								
1. Determinar los parámetros de instalación de la base de datos.		○			●			
2. Instalar la base de datos.	□	□			●			
3. Crear tablas físicas y definir los objetos dimensionales.		□			●			
<b>DISEÑO Y CONSTRUCCION DE LA CAPA DE CARGA DE DATOS</b>								
1. Diseño de los procesos de carga de datos.	▲	○	○	○	◐	●		
2. Diseñar plan de carga para cada tabla.	▲	○	○	○	◐	●		
3. Crear los programas de carga para cada tabla dimensión.		□			○	●		
4. Crear los programas de carga para las tablas de hecho.		□			○	●		
<b>POBLAR LA BASE DE DATOS</b>								
1. Realizar la carga inicial de datos.		□			○	●		
<b>ESPECIFICACION DE LA APLICACIÓN DE USUARIO FINAL (Prototipo de herramienta OLAP)</b>								
1. Identificar y priorizar reportes candidatos.		○	●	●			●	
2. Desarrollar estándares para la aplicación de usuarios final.	▲	□	○	○			●	
<b>DESARROLLO DE LA APLICACIÓN DE USUARIO FINAL</b>								
1. Seleccionar aproximación de implementación.		○	○	○			●	
2. Desarrollar prototipo de herramienta OLAP.		□					●	
3. Documentar la aplicación prototipo.		□					●	
<b>DESPLIEGUE Y CRECIMIENTO</b>								
<b>PLAN DE DESPLIEGUE</b>								
1. Desarrollar la estrategia de educación inicial para el usuario.		○						●
2. Definir el plan de entrega de los productos.	□	●	●	●	●	●	●	●

**Tabla 3.** Plan del proyecto.



### **8.3 CONSIDERACIONES ESPECIALES SOBRE LA DEFINICION Y PLANEACION DEL PROYECTO DE BODEGA DE DATOS**

Durante el desarrollo de estas dos etapas se presentaron varias condiciones especiales, que afectaron el curso normal de las actividades que se encuentran involucradas en la definición y la planeación del proyecto de bodegas de datos.

Desde el momento de determinar el escenario en la organización, se pudo observar que el proyecto estaba enmarcado en un ambiente de difícil trato. Algunas de las razones por las cuales se llego a esta conclusión fueron: el sistema operacional (OLTP), soporte de la bodega de datos, no se había definido, por ende no existían datos o información que pudiera ser útil al momento de determinar los temas de interés, sobre los cuales diseñar los modelos dimensionales de la bodega; además, como el sistema de aprendizaje en línea (Unicauca Virtual) no se encontraba en producción, no existían usuarios quienes pudieran aportar ideas de lo que se quería plasmar en la bodega de datos en la etapa de definición de requerimientos.

Como solución al inconveniente presentado, se opto por utilizar “*sesiones de recolección requerimientos*” [13] en las etapas tempranas de desarrollo de la segunda fase de Unicauca Virtual. El objetivo de estas sesiones fue apropiar todos los aspectos importantes relacionados con el desarrollo de la segunda fase de Unicauca Virtual y de esta manera poder definir los temas del negocio en la iniciativa de construcción de la bodega de datos.



---

# CAPITULO 4 – MODELADO DIMENSIONAL

---

El modelado dimensional es la etapa donde se definen los modelos dimensionales necesarios a tratar para cada uno de los temas de interés del negocio. Estos modelos deben representar la información necesaria, que permita a los usuarios efectuar análisis relacionados con el tema y así, cumplir con el objetivo por el cual fue creado.

Los resultados obtenidos en esta fase fueron los siguientes:

## 9 MATRIZ BUS

Para construir la matriz bus<sup>5</sup> [13], se hizo necesario definir una lista de data marts o áreas del negocio y una lista de dimensiones conformadas<sup>6</sup> sobre las cuales se iba a trabajar.

Los data marts definidos para el proyecto fueron:

- **Estilos de aprendizaje:** con esta área de negocio se pretendía que los usuarios de la bodega de datos, pudieran obtener información importante sobre los estilos de aprendizaje. Alguna información como el cambio en los estilos de aprendizaje de los estudiantes a través del tiempo o el porcentaje de apoyo de un recurso sobre un estilo determinado, ayudan a seguir el proceso de aprendizaje de los estudiantes en el ambiente de educación virtual.
- **Interacción del estudiante:** el objetivo de este data mart era servir como base para obtener información relacionada con el comportamiento de los estudiantes dentro de ambiente virtual, las actividades que el realiza y la interacción que tiene con recursos (SCO ó ASSETS).
- **Evaluaciones:** este data mart permitiría que los usuarios de la bodega de datos, obtuvieran información relevante como resultado de las evaluaciones que realiza un

---

<sup>5</sup> La matriz bus desarrollada por el equipo de diseño es un diagrama donde se plasma la relación existente entre los data marts seleccionados y las dimensiones que se han definido como conformadas.

<sup>6</sup> Dimensiones conformadas son las dimensiones que significan lo mismo con cada posible tabla de hechos a la cual se encuentra unida.



estudiante, de los temas o actividades en el contenido programado para un curso virtual.

Las dimensiones conformadas fueron:

- **Estudiante:** esta dimensión permite mantener información descriptiva del estudiante que forma parte de Unicauca Virtual.
- **Profesor:** esta dimensión permite mantener información descriptiva del profesor dentro de Unicauca Virtual.
- **Recurso:** en esta dimensión se mantienen datos descriptivos de los recursos (SCO ó ASSET). Los recursos son las unidades de información dentro de Unicauca Virtual con los cuales el estudiante interactúa directamente para llevar a cabo su proceso de aprendizaje.
- **Actividad:** esta dimensión mantiene información que describe ampliamente cada una de las actividades o temas que se encuentran en las estructuras de contenidos<sup>7</sup> y que el estudiante tiene la posibilidad de revisar.
- **Estilo de aprendizaje:** es importante almacenar información sobre los estilos de aprendizaje que poseen los estudiantes y que se gestionan a través del ambiente de aprendizaje virtual.
- **Curso:** conserva información descriptiva de las asignaturas y grupos que se forman para dictar las clases dentro de Unicauca Virtual.
- **Fecha:** guarda datos descriptivos de la fecha del calendario en la cual ocurrieron los hechos dentro del ambiente. Esta dimensión permite obtener información de aspectos importantes ocurridos a través del tiempo.
- **Periodo:** esta dimensión permite mantener datos descriptivos de los diferentes periodos académicos que se presentan dentro del ambiente de aprendizaje virtual.
- **Localidad:** esta subdimensión almacena datos descriptivos sobre las distintas localidades que se manejan para los usuarios del ambiente virtual.

Después de listar los data marts a tratar y las dimensiones conformadas, se realizó la intersección entre ellos (Matriz Bus). Ver Tabla 4.

---

<sup>7</sup> Conjunto de agregaciones y actividades representando los temas que se organizan para el desarrollo de un curso virtual.



	Estudiante	Profesor	Recurso	Actividad	Estilo de aprendizaje	Curso	Fecha	Periodo	Localidad
Estilos de aprendizaje	✓		✓	✓	✓	✓	✓	✓	✓
Interacción del estudiante	✓	✓	✓	✓		✓	✓	✓	✓
Evaluaciones	✓		✓	✓		✓	✓	✓	✓

Tabla 4. Matriz bus.

## 10 DISEÑO DE LAS TABLAS DE HECHOS

El diseño lógico detallado de cada tabla de hechos es conducido por cuatro pasos [13]:

- 1. Seleccionar el data mart:** se escoge uno de los data marts listados en la matriz bus.
- 2. Declarar la granularidad de la tabla de hechos:** declarar la granularidad es equivalente a decir el significado de un registro individual de la tabla de hechos. Los tres estilos mas comunes que se identifican son:
  - Transacciones individuales*, el registro de la tabla de hechos representa una transacción que ocurre en el sistema operacional, por ejemplo, una transacción bancaria.
  - Snapshots (imágenes periódicas)*, en donde se espera hasta el final del día o del mes para registrar un consolidado de las actividades que ocurren durante ese periodo.
  - Línea de ítems*, representa documentos de control como facturas, ordenes de envío, etc. Hay exactamente un registro en cada una de estas tablas de hechos por cada ítem de línea sobre el documento de control.
- 3. Seleccionar las dimensiones:** una buena declaración de la granularidad a menudo hace fácil escoger las dimensiones que se unen a una tabla de hecho



particular. El objetivo de este paso es seleccionar las dimensiones que conformen cada modelo dimensional.

- 4. Escoger los hechos:** el último paso es adicionar cuantos hechos o medidas sea posible dentro del contexto de la granularidad declarada en la tabla de hechos. Una medida o hecho es un indicador de la organización o del negocio que se quiere medir, muchos hechos en el mundo de negocios son numéricos, aunque unos pocos pueden ser valores textuales.

A continuación se especifica el diseño de las tabla de hechos de cada uno de los data marts identificados para la construcción de la bodega de datos en el proyecto.

## 10.1 DATA MART – INTERACCION DEL ESTUDIANTE

### Paso 1.

Como primera instancia se eligió el data mart o área de negocio denominada “*Interacción del estudiante*”. Esta selección se realizó porque esta área era sobre la cual se tenía mayor conocimiento y se contaba con mejores bases para su construcción.

### Paso 2.

Para esta área de negocio se definió una tabla de hechos con el nombre “*Visualización de recurso*”. Con esta tabla se pretendía medir los tiempos que utilizaba un estudiante del ambiente virtual en la visualización de un recurso en un momento determinado (tiempo de visualización y tiempo de descanso en la visualización).

Para la tabla de hechos identificada se definió la granularidad *transaccional (transacciones individuales)*, quiere decir, que cada registro en esta tabla representa los tiempos que utiliza un estudiante cuando visualiza un recurso (SCO ó ASSET) de una actividad programada en un contenido, para un curso de una asignatura en el cual esta matriculado el estudiante, en un periodo académico determinado.

### Paso 3.

La declaración de granularidad en el paso anterior permite la selección fácil de las dimensiones que conforman el esquema estrella para la tabla de hechos. Las dimensiones elegidas para la tabla de hechos “*Visualización de recurso*” fueron:

- Estudiante – Localidad (subdimensión).
- Recurso.
- Curso.
- Actividad.

- Periodo.
- Profesor – Localidad (subdimensión).
- Fecha.

#### Paso 4.

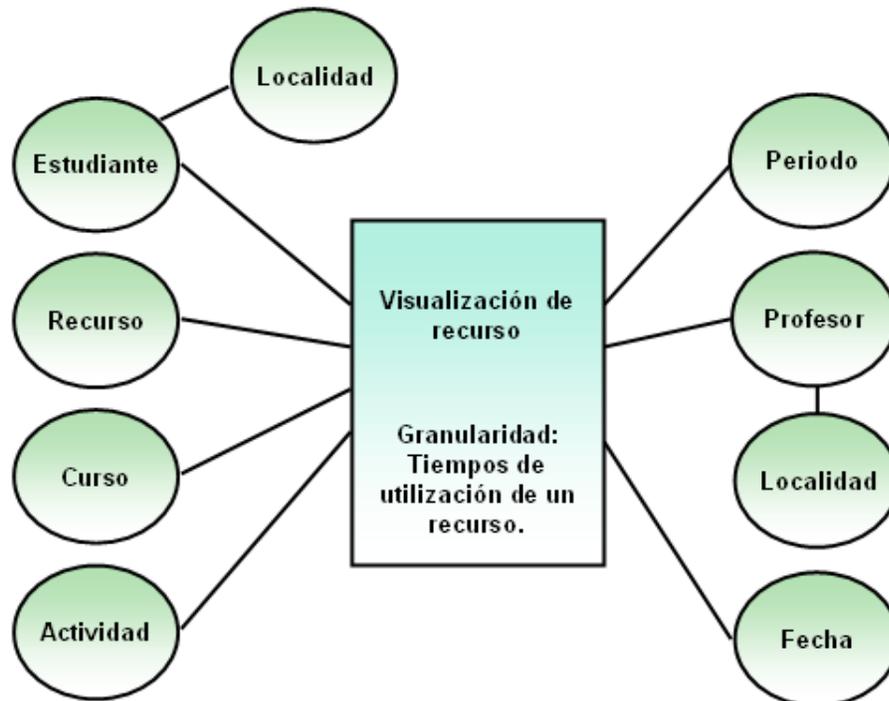
En el contexto de granularidad definido, los hechos (medidas) adicionados para la tabla de hechos fueron:

- *Tiempo de utilización*: este hecho permite medir el tiempo utilizado por un estudiante en visualizar un recurso.
- *Tiempo de descanso*: permite medir el tiempo que se toma un estudiante como descanso cuando se encuentra visualizando un recurso (SCO ó ASSET).
- *Tiempo total*: es el tiempo total medido en la interacción del estudiante con el recurso (tiempo de utilización más tiempo de descanso).
- *Indicador de visualización*: indica que el estudiante visualizo un recurso en un momento dado. Esta medida facilita realizar un conteo de las veces que un estudiante ha revisado un recurso determinado.

#### 10.1.1 Detalles de la tabla de hechos “Visualización de recurso”

- **Diagrama de la tabla de hechos**

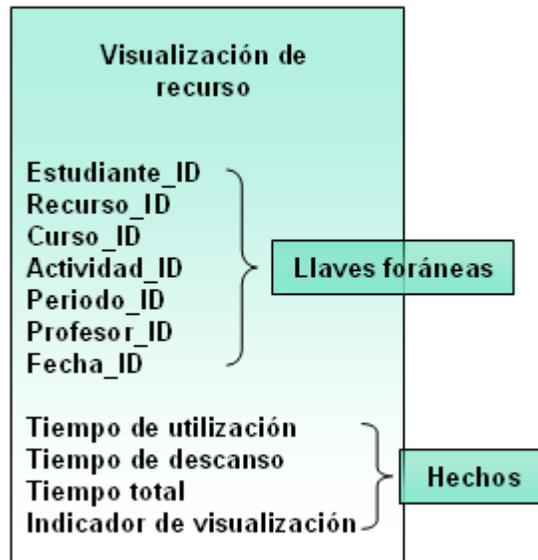
El diagrama estrella de la tabla de hechos “Visualización de recurso” se muestra en Figura 9.



**Figura 9.** Diagrama de la tabla de hechos “*Visualización de Recurso*”.

- **Detalles de la tabla de hechos**

Los detalles de la tabla de hechos “*Visualización de recursos*” se muestran en la Figura 10.



**Figura 10.** Descripción de la tabla de hechos “*Visualización de recurso*”.

### 10.1.2 Detalles de las tablas de dimensión

Para detallar cada una de las dimensiones que conforman la tabla de hechos se hizo un diagrama que indica los atributos, jerarquías y niveles que la conforman; además, se describen detalladamente cada uno de los atributos contenidos en la tabla dimensión.

#### **Dimensión “*Estudiante*”**

- **Diagrama de la dimensión**

La Figura 11, muestra los detalles de la dimensión estudiante.

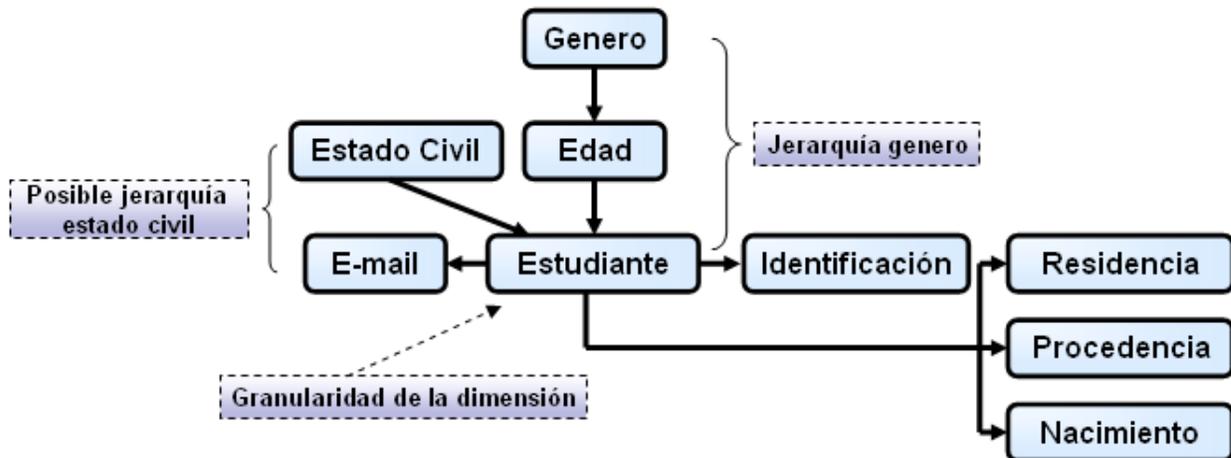


Figura 11. Diagrama de detalle para la dimensión “Estudiante”.

- **Descripción de la dimensión**

La Tabla 5 describe detalladamente el diagrama.

Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...
Código Universidad	Representa el número de identificación del estudiante dentro de Unicauca Virtual.	*...*	No se modifica <sup>8</sup> (Tipo 0)	46001031, 46001022.
Identificación	Representa el número que identifica a un estudiante como ciudadano.	*...*	Se sobrescribe (Tipo 1)	10.293.166, 10.293.710.
Nombres	Representa los nombres del estudiante.	*...*	No se modifica (Tipo 0)	Jaime Alberto, Lisandro.
Apellidos	Representa los apellidos del estudiante.	*...*	No se modifica (Tipo 0)	Muñoz Córdoba, Acosta Mejía.
Género	Género del estudiante.	2	No se modifica (Tipo 0)	Masculino, Femenino.
Estado civil	Representa el estado civil del estudiante.	4	Se sobrescribe (Tipo 1)	Soltero(a), Casado(a), Viudo(a), Divorciado(a).
Edad	Edad del estudiante.	35	Se sobrescribe (Tipo 1)	23 años, 24 años, 27 años.

<sup>8</sup> Las políticas para manejar el cambio en un atributo de la dimensión son: *tipo 0*, donde el valor nunca se modifica, *tipo 1 (sobrescribir)*, se sobrescribe el atributo que se va a modificar, *tipo 2 (nuevo registro creado)*, se crea un nuevo registro en la tabla cuando se detecta un cambio y *tipo 3 (nueva versión del valor a modificar)*, se mantiene solamente una nueva versión del valor a modificar en una columna de la tabla.



Correo electrónico institucional	Representa el e-mail institucional del estudiante.	* ... *	Se sobrescribe (Tipo 1)	jaimem@unicauca.edu.co, lisandroam@unicauca.edu.co.
Correo electrónico personal	Representa el e-mail personal del estudiante.	* ... *	Se sobrescribe (Tipo 1)	jamc63@hotmail.com, lisandro777@hotmail.com.
Sitio Web personal	Representa el sitio Web personal de un estudiante.	* ... *	Se sobrescribe (Tipo 1)	www.unicauca.edu.co/~jaimem, www.unicauca.edu.co/~lisandroam.
Fecha de nacimiento	Representa la fecha de nacimiento del estudiante.	* ... *	No se modifica (Tipo 0)	31/12/1981, 06/07/1981.
ID Residencia	Representa la clave para identificar la subdimensión <sup>9</sup> [13] de localidad. Rol: Residencia.	* ... *	Se sobrescribe (Tipo 1)	1, 2, 3, 4...
ID Procedencia	Representa la clave para identificar la subdimensión de localidad. Rol: Procedencia.	* ... *	Se sobrescribe (Tipo 1)	1, 2, 3, 4...
ID Nacimiento	Representa la clave para identificar la subdimensión de localidad. Rol: Nacimiento.	* ... *	Se sobrescribe (Tipo 1)	1, 2, 3, 4...

**Tabla 5.** Descripción detallada de la dimensión “Estudiante”.

CONVENCIONES	
Convención	Descripción
* ... *	Esta convención indica que puede haber un gran número de valores para este atributo.

**Tabla 6.** Convenciones.

La descripción detallada de las demás dimensiones que conforman la tabla de hechos “Visualización de recurso” se encuentra en el ANEXO 4.

## 10.2 DATA MART – ESTILOS DE APRENDIZAJE

### Paso 1.

En segunda instancia el data mart seleccionado para su desarrollo fue “Estilos de aprendizaje”. Este tema del negocio permite manejar la información relacionada con los

<sup>9</sup> Una subdimensión es un conjunto de atributos de baja cardinalidad que separan de una gran dimensión (esquema permisible de copo de nieve).



estilos de aprendizaje que un estudiante tiene asociados; además, tratar la relación que existe entre los recursos (SCO ó ASSETS) y los estilos de aprendizaje.

## Paso 2.

Para este tema del negocio se definieron dos tablas de hechos:

- La primera denominada “*Cambio en el estilo de aprendizaje del estudiante*”. Esta tabla de hechos permite medir el cambio en el porcentaje de los estilos de aprendizaje asociados al estudiante a través del tiempo. Por ejemplo, un estudiante puede ser catalogado en un modelo de aprendizaje para una asignatura con sus diferentes porcentajes asociados a cada estilo, a través del tiempo y de acuerdo al seguimiento a través de algunos indicadores estos porcentajes pueden cambiar y afectar su proceso su aprendizaje.

La granularidad definida para esta tabla de hechos fue *transaccional (transacciones individuales)*, quiere decir, que cada registro en esta tabla representa el nuevo porcentaje de un estilo de aprendizaje asociado a un estudiante, en un curso asignado de un periodo académico.

- La segunda tabla de hechos “*Recurso en el estilo de aprendizaje*”, en donde se puede visualizar el porcentaje que apoya un determinado recurso a un estilo en particular. Esta asociación se realiza al momento de crear cada recurso y cambia de acuerdo a las recomendaciones de profesores y directivos.

La granularidad identificada para esta tabla fue *transaccional (transacciones individuales)*, es decir, que cada registro de la tabla de hechos representa el porcentaje que apoya un determinado recurso (SCO ó ASSET) a un estilo de aprendizaje.

## Paso 3.

Las dimensiones seleccionadas para cada una de las tablas de hechos fueron:

- Dimensiones en la tabla de hechos “*Cambio en el estilo de aprendizaje del estudiante*”:
  - Estudiante – Localidad (subdimensión).
  - Estilo de aprendizaje.
  - Curso.
  - Periodo.
  - Fecha.
- Dimensiones en la tabla de hechos “*Recurso en el estilo de aprendizaje*”:
  - Estilo de aprendizaje.
  - Recurso.



- Periodo.
- Fecha.

#### **Paso 4.**

En el contexto de granularidad definido para cada tabla de hechos se definieron los hechos ó medidas de la siguiente forma:

- Para la tabla de hechos “*Cambio en el estilo de aprendizaje del estudiante*” se definió el siguiente hecho:
  - *Porcentaje promedio del estilo asociado*: este hecho permite obtener información relacionada con el porcentaje asociado de un estilo de aprendizaje a un estudiante.
- Para la tabla de hechos “*Recurso en el estilo de aprendizaje*” se definió el siguiente hecho:
  - *Porcentaje promedio del recurso al estilo*: este hecho permite medir el porcentaje promedio que apoya un determinado recurso a un estilo de aprendizaje.

#### **10.2.1 Detalles de la tabla de hechos “*Cambio en el estilo de aprendizaje del estudiante*”**

- **Diagrama de la tabla de hechos**

El diagrama estrella de la tabla de hechos “*Cambio en el estilo de aprendizaje del estudiante*” se muestra en Figura 12.

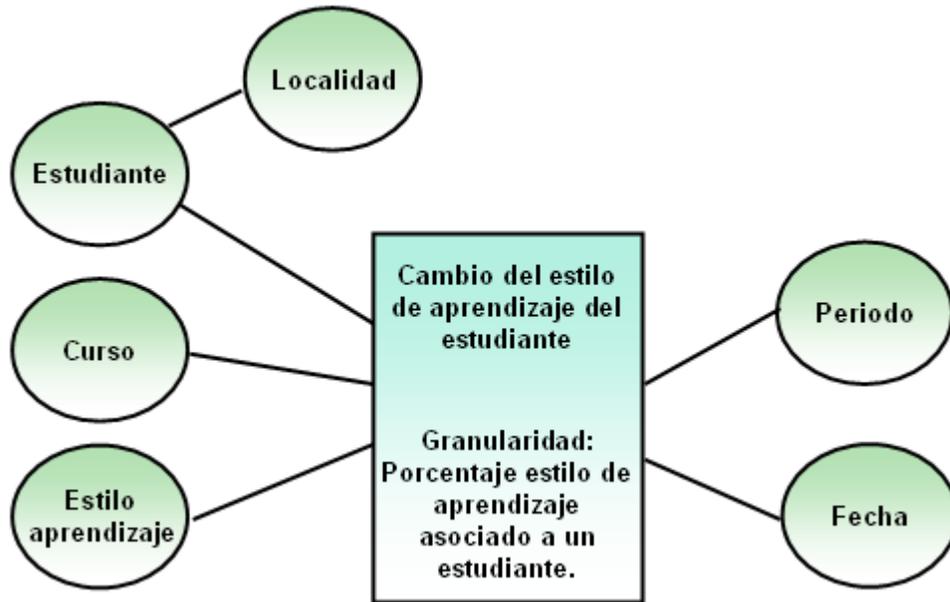


Figura 12. Diagrama de la tabla de hechos "Cambio en el estilo de aprendizaje del estudiante".

- **Detalles de la tabla de hechos**

Los detalles de la tabla de hechos "Cambio en el estilo de aprendizaje del estudiante" se muestran en la Figura 13.

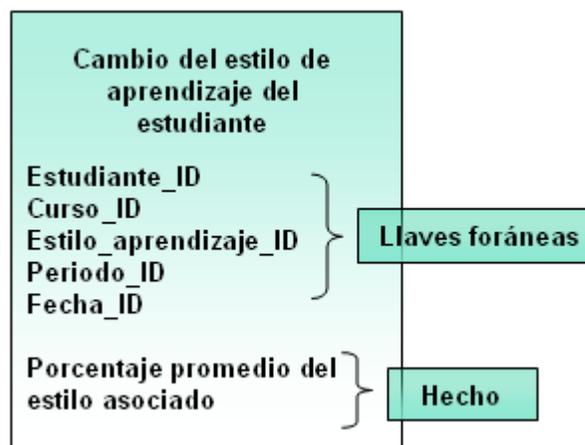


Figura 13. Descripción de la tabla de hechos "Cambio del estilo de aprendizaje del estudiante".

### 10.2.2 Detalles de la tabla de hechos “Recurso en el estilo de aprendizaje”

- Diagrama de la tabla de hechos

El diagrama estrella de la tabla de hechos “Recurso en el estilo de aprendizaje” se muestra en Figura 14.

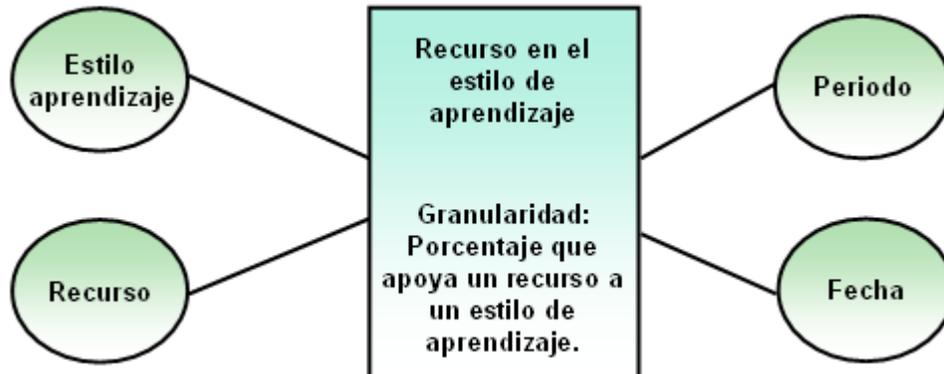


Figura 14. Diagrama de la tabla de hechos “Recurso en el estilo de aprendizaje”.

- Detalles de la tabla de hechos

Los detalles de la tabla de hechos “Recurso en el estilo de aprendizaje” se muestran en la Figura 15.

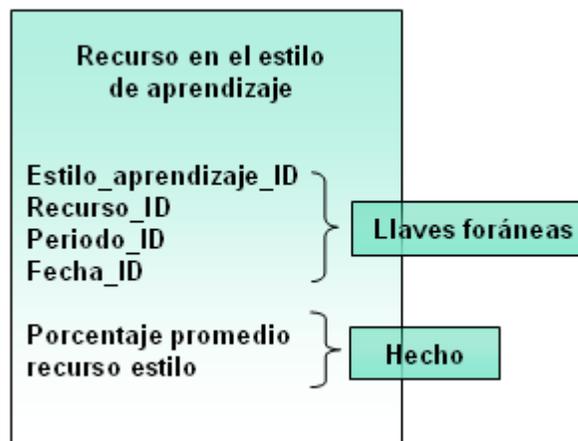


Figura 15. Descripción de la tabla de hechos “Recurso en el estilo de aprendizaje”.

### 10.2.3 Detalles de las tablas dimensión

Para detallar cada una de las dimensiones que conforman la tabla de hechos se hizo un diagrama que indica los atributos, jerarquías y niveles que la conforman; además, se



describen detalladamente cada uno de los atributos contenidos en la tabla dimensión. El detalle de cada una de las dimensiones que conforman la tabla de hechos se muestra en el ANEXO 4.

## 10.3 DATA MART – EVALUACIONES

### Paso 1.

En última instancia se tomó el data mart de “*Evaluaciones*”. Este tema del negocio permite manejar la información relacionada con los resultados obtenidos de las evaluaciones que realiza el estudiante de las distintas actividades (temas) en su proceso de aprendizaje.

### Paso 2.

Para este tema del negocio se definió una tabla de hechos llamada “*Evaluación de actividad*”. Esta tabla permite mantener información de las calificaciones obtenidas al momento de evaluar una cierta actividad, incluida en una estructura de contenido<sup>7</sup> que se presenta al estudiante para que lleve a cabo su proceso de aprendizaje.

La granularidad definida para esta tabla de hechos es *transaccional (transacciones individuales)*, donde cada registro de la tabla representa la nota obtenida de evaluar una actividad contenida en la estructura, definida para un curso virtual en un periodo académico. Se registra cada una de las evaluaciones realizadas por el estudiante de la actividad.

### Paso 3.

Las dimensiones seleccionadas para esta tabla de hechos fueron:

- Estudiante – Localidad (subdimensión).
- Actividad.
- Curso.
- Periodo.
- Profesor – Localidad (subdimensión).
- Fecha.

### Paso 4.

En el contexto de granularidad definido para la tabla de hechos se definieron los siguientes hechos:

- *Nota promedio evaluación actividad*: este hecho permite mantener la nota obtenida de la evaluación realizada para una actividad en un momento específico.
- *Indicador de evaluación*: indica que el estudiante realizó una evaluación para la actividad en un momento dado. Esta medida permite realizar un conteo del número de veces que un estudiante presenta una evaluación.

### 10.3.1 Detalles de la tabla de hechos “Evaluación de actividad”

- **Diagrama de la tabla de hechos**

El diagrama estrella de la tabla de hechos “Evaluación de actividad” se muestra en Figura 16.

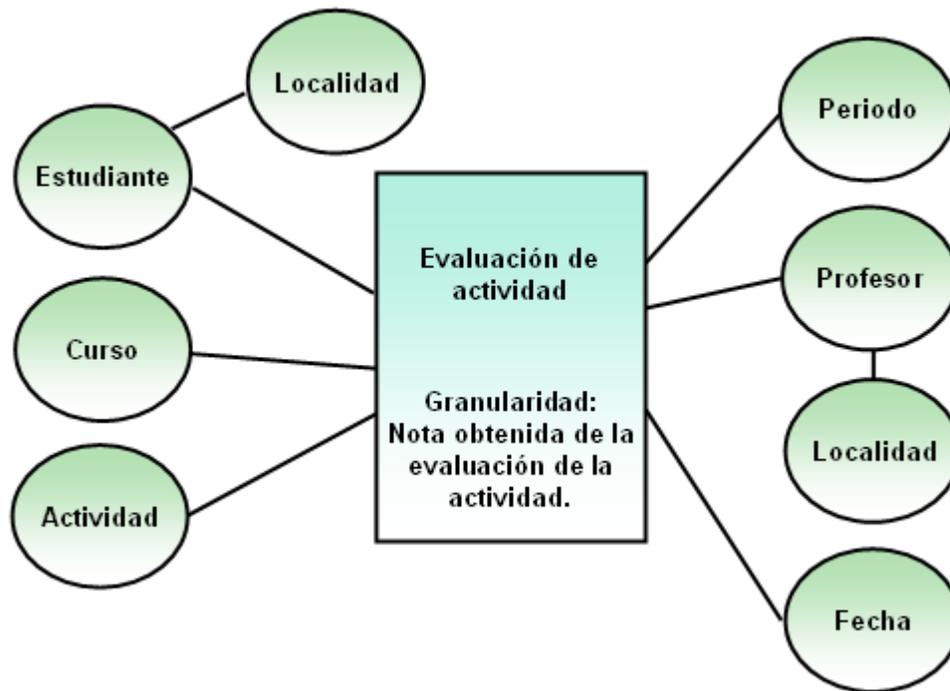


Figura 16. Diagrama de la tabla de hechos “Evaluación de actividad”.

- **Detalles de la tabla de hechos**

Los detalles de la tabla de hechos “Evaluación de actividad” se muestran en la Figura 17.

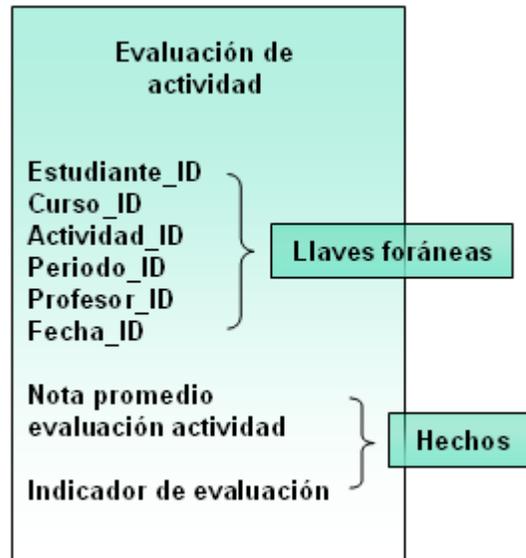


Figura 17. Descripción de la tabla de hechos “Evaluación de actividad”.

### 10.3.2 Detalles de las tablas dimensión

Para detallar cada una de las dimensiones que conforman la tabla de hechos se hizo un diagrama que indica los atributos, jerarquías y niveles que la conforman; además, se describen detalladamente cada uno de los atributos contenidos en la tabla dimensión. El detalle de cada una de las dimensiones que conforman la tabla de hechos se muestra en el ANEXO 4.

## 10.4 CONSIDERACIONES ESPECIALES SOBRE EL MODELADO DIMENSIONAL DE LA BODEGA DE DATOS

El modelado dimensional según expertos en el tema (Kimball, 1998, p.137), conlleva al análisis de casos de modelado especial que se presentan según lo que se quiera representar en un esquema de bodegas de datos. Casos como la representación de una relación “*muchos a muchos*” entre una dimensión y una tabla de hechos o el manejo de “*subdimensiones (outriggers)*” para atributos de dimensión con baja cardinalidad, hacen que los desarrolladores de la bodega piensen en la posible representación física de estos casos en el motor de base de datos.

Debido a que el proyecto de bodega de datos se encontraba enmarcado en el macroproyecto Unicauca Virtual, se eligió construir físicamente la bodega sobre el motor de base de datos Oracle 10g, en el cual se encuentra el soporte de la base operacional del sistema de aprendizaje virtual (OLTP).



Después de analizar una serie de casos presentados en la fase de modelado, los desarrolladores del proyecto pudieron notar que el motor de base de datos elegido no soportaba la representación lógica de algunas de ellos. Por ejemplo, en el momento de intentar especificar “*tablas de hecho y dimensiones personalizadas*” o “*tablas puente*” para relaciones de muchos a muchos, las vistas disponibles de metadatos OLAP<sup>10</sup> en Oracle no permitían representar lógicamente estos casos.

Para dar solución a los problemas presentados, el equipo de desarrollo planteó las siguientes alternativas:

- Para poder representar “*subdimensiones*”, caso que surgió en el momento de especificar información relacionada con la localidad de los estudiantes y profesores, se optó por utilizar la opción que provee Oracle para especificar la unión de dos niveles en jerarquías que se encuentran en diferentes tablas dimensión. La variación que se tuvo en cuenta fue que para cada rol que representaba la dimensión lógica se crearon vistas de datos materializadas.
- En los casos presentados de “*tablas puente*”, como por ejemplo, cuando se trató de representar la relación muchos a muchos que existía entre los recursos y las actividades<sup>11</sup>, se analizó las siguientes opciones como solución:
  - Primero, se intentó representar la misma información con la unión de varios esquemas dimensionales, es decir, se tenía que rediseñar los modelos para poder representar la misma información en el negocio.
  - La segunda opción fue, descartar la dimensión que generaba el problema. Sin embargo, se debía observar cuan importante era el significado del nuevo modelo para poder incluirlo dentro de los temas de interés del negocio.
  - Como última opción, en caso de no obtener buenos resultados con las soluciones planteadas anteriormente, fue eliminar el modelo y plantear otros que pudieran dar significado a otros temas de relevancia para el negocio.
- Para el caso de las “*tablas de hecho y dimensiones personalizadas*”, que se presentó en el momento de tratar de representar el cambio de los diferentes estilos de aprendizaje de un estudiante, se diseñó el modelo de tal forma que se indicara por cada registro el nuevo valor para el estilo de aprendizaje. Aunque el diseño no impactó mucho sobre la idea previamente establecida, era importante tener en cuenta que cuando un estilo de aprendizaje cambiaba su valor, como consecuencia

---

<sup>10</sup> Los metadatos Oracle OLAP son la representación lógica de los esquemas dimensionales implementados en el motor de base de datos. Estos representan dimensiones, jerarquías, niveles, atributos, cubos, medidas, carpetas de medidas (data marts), etc.

<sup>11</sup> Un recurso (SCO ó ASSET) puede apoyar a diferentes actividades en una estructura de contenidos



se tenía un cambio en los demás estilos, lo cual era importante y necesario reflejar los nuevos datos sobre la tabla de hechos.

Por ultimo, el hecho de no tener usuarios que validarán el significado de los modelos dimensionales no favoreció la elaboración de modelos más robusto que respondieran a necesidades reales de análisis. Los modelos propuestos, aunque fueron elaborados teniendo en cuenta que existían necesidades de información, no fueron guiados por requerimientos de análisis de usuarios finales.



---

# CAPITULO 5 – IMPLEMENTACIÓN DE LA BODEGA

---

El objetivo de la etapa de implementación es lograr que los modelos dimensionales propuestos sean representados físicamente en el motor de base de datos elegido. La implementación de cada uno de estos esquemas se realizó en Oracle 10g, como requisito inicial de cada proyecto enmarcado en Unicauca Virtual.

## 11 DISEÑO FÍSICO DE LOS MODELOS DIMENSIONALES

Con el diseño obtenido de cada uno de los esquemas dimensionales en la etapa anterior, se creó el modelo físico que muestra todos los detalles de las tablas de hechos con sus correspondientes dimensiones.

La Figura 18, muestra el diseño físico de la tabla de hechos “*Visualización de recurso*”. En el ANEXO 3, se muestran los modelos físicos para las tablas de hechos restantes.

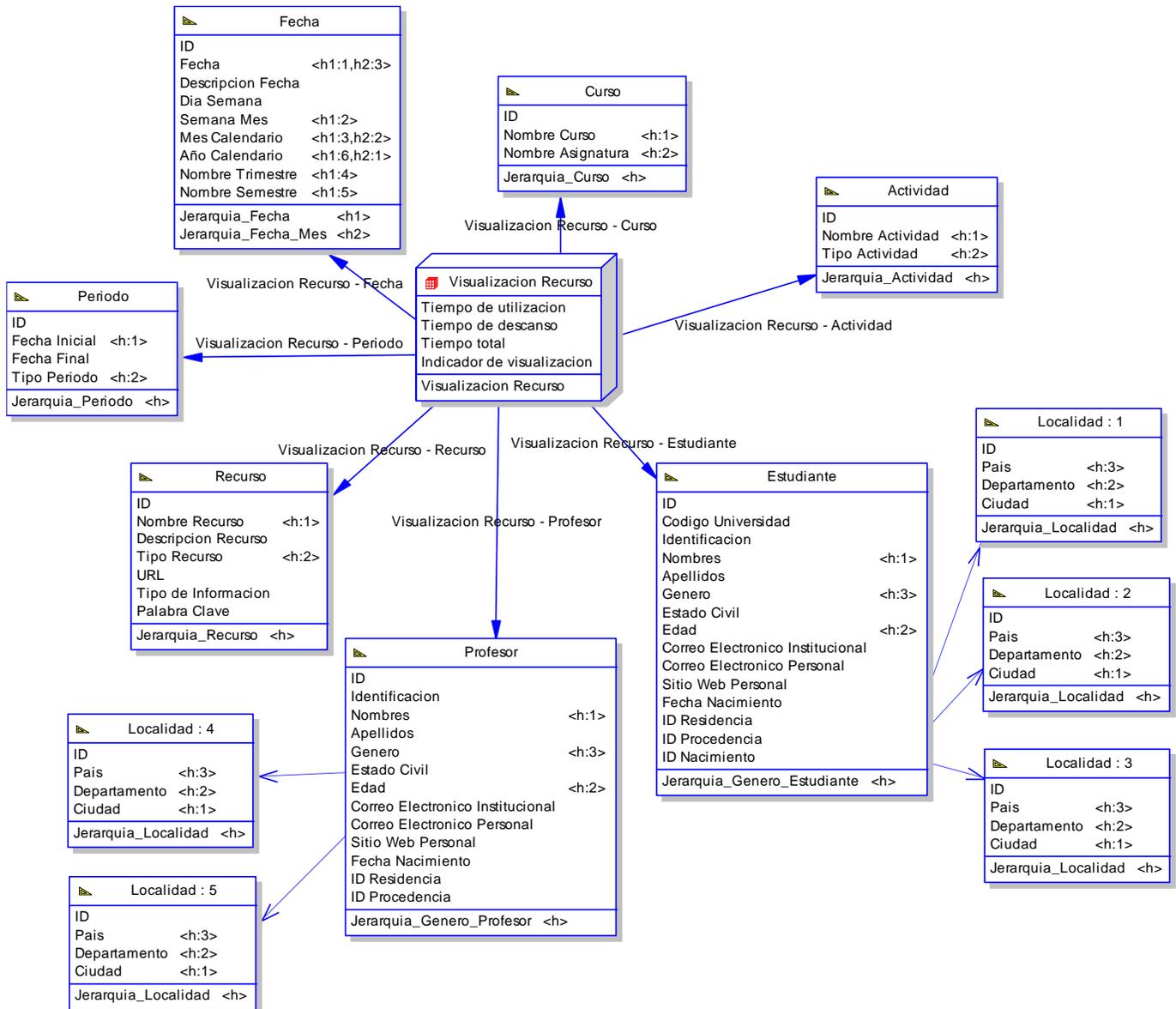


Figura 18. Modelo físico de la tabla de hechos “Visualización de recurso”.

Este modelo muestra detalladamente el esquema dimensional de forma física, cada una de las dimensiones contiene todos los atributos y jerarquías correspondientes y se incluyen las subdimensiones que juegan varios roles dentro del modelo (Localidad).



## 12 DEFINICION DE LAS TABLAS FISICAS Y OBJETOS DIMENSIONALES

### 12.1 TABLAS RELACIONALES

En la exploración tecnológica, con el desarrollo de varias pruebas de implementación, se pudo identificar que la construcción de la bodega de datos requería de una base relacional, la cual sirve como soporte para la creación de objetos multidimensionales, es decir, un esquema relacional distinto al OLTP donde se encuentren los datos “limpios” y estructurados obtenidos después de realizar el proceso ETL.

Para el soporte relacional de la bodega de datos, se definieron dos tipos de tablas relacionales:

- Una tabla relacional con una estructura definida que permita crear las dimensiones del esquema dimensional: cada una de las tablas físicas debe contener las columnas de datos necesarias para poder representar los niveles que forman cada una de las jerarquías de la dimensión. Además, esta estructura puede contener columnas de datos adicionales que permitan almacenar diferentes atributos descriptivos para un nivel.

La estructura definida para soportar la definición de dimensiones, tiene las siguientes especificaciones:

- *Columna que identifica el nivel:* esta columna garantiza la unicidad de los niveles contenidos en la dimensión.
- *Columna que describe el nivel:* esta columna define una descripción corta para el nivel en la dimensión.
- *Columnas adicionales que contienen atributos descriptivos para el nivel (Opcional):* cada una de las columnas permite adicionar información descriptiva para el nivel.

Es importante notar que en la creación de cada tabla relacional, se debe incluir las columnas que representen el nivel superior para la jerarquía. Por ejemplo, en la creación de la tabla física para la dimensión actividad se define una columna que identifica al nivel superior de la actividad (ACT\_TIPOS\_ACTIVIDADES\_ID) y una columna que describe dicho nivel (ACT\_TIPOS\_ACTIVIDAD\_DESC).

La Figura 19, muestra un conjunto de sentencias SQL que permiten crear la tabla relacional física soporte para la dimensión “Actividad”.



```
CREATE TABLE "USER1"."ACTIVIDAD"
(
  "ACT_TIPOS_ACTIVIDADES_ID" NUMBER(1) NOT NULL ,
  "ACT_TIPOS_ACTIVIDADES_DESC" VARCHAR2(30) NOT NULL ,

  "ACT_TIPO_ACTIVIDAD_ID" NUMBER(10) NOT NULL ,
  "ACT_TIPO_ACTIVIDAD_DESC" VARCHAR2(200) NOT NULL ,

  "ACT_ACTIVIDAD_ID" NUMBER(15),
  "ACT_ACTIVIDAD_DESC" VARCHAR2(1024) NOT NULL ,

  CONSTRAINT "PK_ACTIVIDAD" PRIMARY KEY ("ACT_ACTIVIDAD_ID") VALIDATE
)
TABLESPACE "USERS"
```

**Figura 19.** Script de creación de la tabla relacional base para la dimensión “Actividad”.

La Figura 20, muestra algunos datos de prueba almacenados en la tabla “Actividad”.

ACT_TIPOS_ACTIVIDADES_ID	ACT_TIPOS_ACTIVIDADES_DESC	ACT_TIPO_ACTIVIDAD_ID	ACT_TIPO_ACTIVIDAD_DESC	ACT_ACTIVIDAD_ID	ACT_ACTIVIDAD_DESC
1	ACTIVIDADES	2	Lectura	1	500 A.C. HASTA 1949 D.C
1	ACTIVIDADES	2	Lectura	2	1950 D.C. HASTA HOY
1	ACTIVIDADES	5	Juegos	3	SOBRE LA HISTORIA
1	ACTIVIDADES	3	Explicación	4	LOGICA DE PROPOSICIONES
1	ACTIVIDADES	3	Explicación	5	LEYES DEL ALGEBRA DE BOOLE

**Figura 20.** Datos de prueba almacenados en la tabla “Actividad”.

- Una tabla relacional con una estructura definida que permita definir el cubo del esquema dimensional: una tabla física con las columnas necesarias que permitan almacenar las claves foráneas de las dimensiones unidas al cubo. Además, debe contener columnas que permitan almacenar cada uno de los valores de las medidas definidas en el modelado de las tablas de hechos.

La estructura definida para soportar la creación del cubo, tiene las siguientes especificaciones:

- *Columnas que identifican cada una de las claves foráneas del cubo:* cada una de las columnas permite almacenar la clave foránea que une la dimensión con el cubo.
- *Columnas que almacenan valores de medida:* columnas que permiten almacenar los valores de medida definidos para la tabla de hechos.

La Figura 21, muestra el conjunto de sentencias SQL que permiten crear la tabla relacional física soporte para la definición del cubo “Visualización de recurso”.



```
CREATE TABLE "USER1"."VISUALIZACION_RECURSO"
(
  "FEC_FECHA_ID" NUMBER(15),
  "EST_ESTUDIANTE_ID" NUMBER(15),
  "PRO_PROFESOR_ID" NUMBER(15),
  "CUR_CURSO_ID" NUMBER(15),
  "ACT_ACTIVIDAD_ID" NUMBER(15),
  "REC_RECURSO_ID" NUMBER(15),
  "PER_PERIODO_ID" NUMBER(15),

  "T_VISUALIZACION_RECURSO" NUMBER(10, 2),
  "T_DESCANSO_RECURSO" NUMBER(10, 2),
  "INDICADOR_VISUALIZACION" NUMBER(1),
  "T_TOTAL_VISUALIZACION" NUMBER(10, 2),

  CONSTRAINT "FK_ACTIVIDAD_ID" FOREIGN KEY ("ACT_ACTIVIDAD_ID") REFERENCES "USER1"."ACTIVIDAD" ("ACT_ACTIVIDAD_ID")
  VALIDATE ,
  CONSTRAINT "FK_CURSO_ID" FOREIGN KEY ("CUR_CURSO_ID") REFERENCES "USER1"."CURSO" ("CUR_CURSO_ID")
  VALIDATE ,
  CONSTRAINT "FK_ESTUDIANTE_ID" FOREIGN KEY ("EST_ESTUDIANTE_ID") REFERENCES "USER1"."ESTUDIANTE" ("EST_ESTUDIANTE_ID")
  VALIDATE ,
  CONSTRAINT "FK_FECHA_ID" FOREIGN KEY ("FEC_FECHA_ID") REFERENCES "USER1"."FECHA" ("FEC_FECHA_ID")
  VALIDATE ,
  CONSTRAINT "FK_PERIODO_ID" FOREIGN KEY ("PER_PERIODO_ID") REFERENCES "USER1"."PERIODO" ("PER_FECHA_INICIAL_ID")
  VALIDATE ,
  CONSTRAINT "FK_PROFESOR_ID" FOREIGN KEY ("PRO_PROFESOR_ID") REFERENCES "USER1"."PROFESOR" ("PRO_PROFESOR_ID")
  VALIDATE ,
  CONSTRAINT "FK_RECURSO_ID" FOREIGN KEY ("REC_RECURSO_ID") REFERENCES "USER1"."RECURSO" ("REC_RECURSO_ID")
  VALIDATE,
  CONSTRAINT "UN_VISUALIZACION_CLAVES" UNIQUE ("FEC_FECHA_ID", "EST_ESTUDIANTE_ID", "PRO_PROFESOR_ID", "CUR_CURSO_ID",
  "ACT_ACTIVIDAD_ID", "REC_RECURSO_ID", "PER_PERIODO_ID")
)
TABLESPACE "USERS"
```

**Figura 21.** Script de creación de la tabla relacional base para la definición del cubo “*Visualización de recurso*”

Los scripts de creación para las demás tablas que conforman los esquemas dimensionales se encuentran en el ANEXO 5.

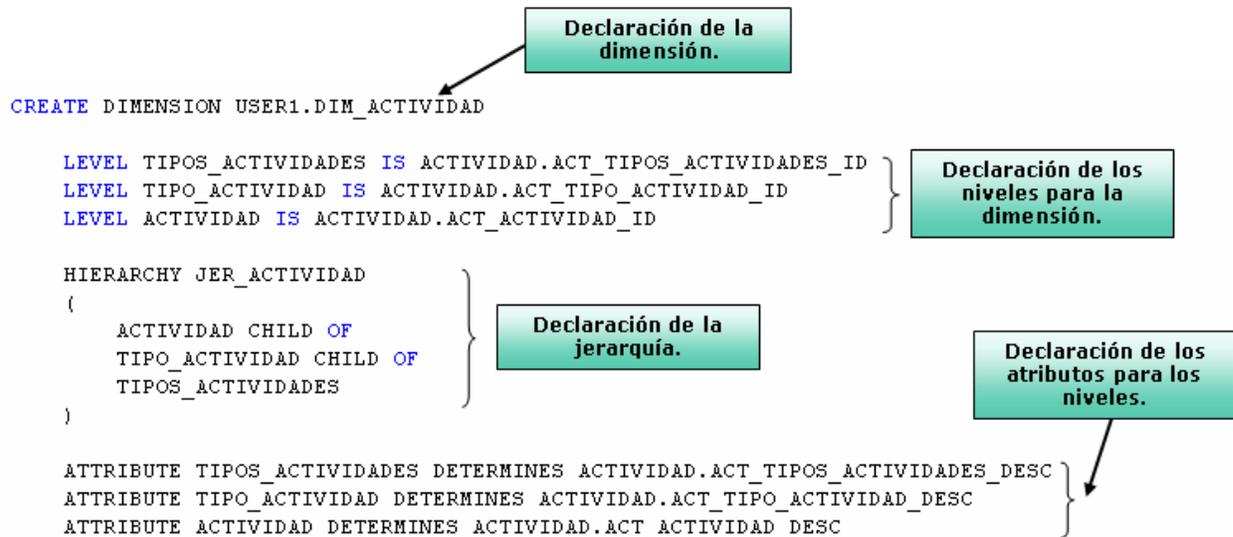
## 12.2 OBJETOS DIMENSIONALES

La definición de los objetos dimensionales involucra la creación de las dimensiones y la especificación de los cubos. Cada una de estas tareas se realizó a través de la herramienta que provee Oracle 10g, llamada “*Oracle Enterprise Manager (OEM)*” en su versión Web. Esta herramienta permitió declarar las relaciones existentes entre: la parte física declarada (tablas relacionales) y los objetos dimensionales (niveles, jerarquías, etc.); y entre objetos dimensionales (nivel - nivel, dimensiones - cubo).

La herramienta, además de crear las relaciones existentes entre objetos, muestra las sentencias SQL con las cuales se definen los objetos. Estas sentencias están conformadas por aquellas que definen los objetos (CREATE DIMENSION) y por aquellas que definen los metadatos CWM1<sup>12</sup> [18] (cwm\_olap\_dimension) para esos objetos.

<sup>12</sup> CWM (Common Warehouse Metamodel) es una especificación de metadatos que permite la comunicación entre bodegas de datos.

El conjunto de sentencias que definen los objetos dimensionales son declaraciones SQL que se ejecutan para crear los diferentes elementos dimensionales de un objeto. En la Figura 22, se muestra la definición de la dimensión “*Actividad*” con sus respectivos elementos (niveles, jerarquías y atributos). El ANEXO 6 muestra detalladamente el proceso y la sintaxis de creación de una dimensión en Oracle 10g[11].



**Figura 22.** Sentencias SQL que permiten definir la dimensión “*Actividad*”.

Por otro lado, a través del OEM también se tiene acceso a las sentencias que permiten declarar los metadatos para los objetos creados. Estos metadatos permiten que otras aplicaciones que manejen estas especificaciones tengan acceso a los objetos multidimensionales. Los metadatos CWM1 fueron el punto de partida para el desarrollo del prototipo OLAP propuesto en el proyecto.

La Figura 23, muestra las sentencias que permiten definir los metadatos para la dimensión “*Actividad*”.

```
begin
cwm_olap_dimension.set_plural_name('USER1', 'DIM_ACTIVIDAD', 'DIM_ACTIVIDAD');
cwm_olap_dimension.set_default_display_hierarchy('USER1', 'DIM_ACTIVIDAD', 'JER_ACTIVIDAD');
cwm_olap_dimension.set_display_name('USER1', 'DIM_ACTIVIDAD', 'ACTIVIDAD');
...
...|
commit;
end;
```

**Figura 23.** Sentencias que permiten definir los metadatos para la dimensión “*Actividad*”.

Igualmente para poder definir el cubo como objeto dimensional dentro del motor de base de datos, se utiliza la sintaxis de metadatos CWM1.



Todos los scripts que se obtuvieron de la definición de los objetos dimensionales para Unicauca Virtual se encuentran en el ANEXO 7.

## 12.3 PROCESO DE CARGA DE DATOS (ETL)

En este proceso se identificaron las fuentes de datos operacionales para extraer los datos y poblar la bodega. Esta fuente operacional es el resultado de la segunda fase del macroproyecto Unicauca Virtual, obtenida de los subproyectos “Agentes Inteligentes para adaptar la presentación de contenidos para cursos de educación en línea” e “Inteligencia Artificial en la generación de estrategias de aprendizaje de un sistema tutor inteligente”.

Para este proceso se realizaron las siguientes tareas:

- Se identificaron los diferentes procesos de carga para las dimensiones conformadas del modelado dimensional.
- Se identificaron los diferentes procesos de carga para las tablas de hechos del modelado dimensional.
- Se utilizó la herramienta “Oracle Warehouse Builder [19]” para realizar los diferentes procesos de carga definidos. Los scripts de carga de datos obtenidos a través de esta herramienta se encuentran en el ANEXO 8.

## 12.4 CONSIDERACIONES ESPECIALES EN LA IMPLEMENTACIÓN DE LA BODEGA DE DATOS

Como ya se ha mencionado, la bodega de datos del proyecto se implementó en el motor de base de datos relacional Oracle; en el cual se pueden crear bodegas de datos relacionales y multidimensionales a partir de la versión 9i.

A continuación se presenta una descripción de las herramientas que ofrece Oracle, para la creación de bodegas de datos relacionales (ROLAP) y multidimensionales (MOLAP).

### 12.3.1 ROLAP

Si se desea implementar una bodega de datos relacional en Oracle, se cuenta con las siguientes opciones:

- Oracle Warehouse Builder (OWB) [19], suministra un completo entorno de trabajo para realizar un proyecto desde la definición misma de las tablas



relacionales hasta la creación de los objetos lógicos, como por ejemplo dimensiones, jerarquías, niveles, cubos, medidas, etc. Además, presenta un entorno gráfico que permite realizar el proceso de extracción, transformación y carga de datos desde el sistema transaccional al esquema de la bodega de datos. Una desventaja del OWB es que carece de la creación directa de los metadatos CWM para los objetos.

- Oracle Enterprise Manager (OEM) en las versiones de escritorio y Web. Permite crear los objetos lógicos con sus correspondientes metadatos CWM (Common Warehouse Metamodel) [20], es decir, quedan visibles para que aplicaciones de terceros que soporten el estándar puedan trabajar con los objetos definidos. La consola de escritorio necesita un componente adicional (parche) para que queden completamente habilitadas las opciones de bodegas de datos. Estos componentes no son necesarios si se usa el OEM versión Web donde no se presentan restricciones de ningún tipo y permite crear todos los objetos lógicos necesarios.

### 12.3.2 MOLAP

Si se desea implementar una bodega de datos multidimensional en Oracle, se tienen las siguientes opciones:

- Analytic Workspace Manager (AWM) [21]. Permite crear la definición de los objetos multidimensionales y además realizar la carga de datos. Los objetos creados desde el AWM quedan con sus correspondientes metadatos CWM. Usando el AWM, la creación de los objetos dimensionales (dimensiones, jerarquías, niveles, cubos, medidas, etc.) se realiza de una manera muy cómoda ayudado por los asistentes. Sin embargo, si se encuentran problemas en el momento de hacer el mapeo de los datos, se presentan dificultades para corregirlos debido a que no es fácil identificar el motivo de los errores. Los objetos que se crean en la “vista de diseño” <sup>13</sup> [21] del AWM cumplen con la forma estándar para *espacios de trabajo analíticos*, los cuales, tienen una gran cantidad de extensiones que no están documentadas para el uso público, dificultando así la comprensión de la definición de los objetos
- OLAP DML [22]. Lenguaje de manipulación de objetos multidimensionales de Oracle que extiende las capacidades de análisis de lenguajes de consulta como el SQL. La implementación directa en lenguaje OLAP DML no es sencilla, teniendo en cuenta que es un lenguaje relativamente nuevo, extenso y la bibliografía que se encuentra al respecto es bastante genérica. Lo que se recomienda según expertos [23] en el manejo de bodegas de datos en Oracle, es usar AWM para la definición de los objetos más genéricos como dimensiones,

<sup>13</sup> El AWM presenta dos opciones de implementación una basada en la vista de diseño y la otra basada en la vista de objeto.



jerarquías, etc. y el lenguaje OLAP DML para definir objetos más complejos y específicos como cálculos o algún tipo de operación de pronósticos.

### 12.3.3 ¿ROLAP O MOLAP?

Para muchos administradores de bases de datos, las tablas relacionales y el SQL representan un ambiente de desarrollo familiar. Por otro lado, los espacios de trabajo analíticos requieren transformación de datos y aprender nuevos conceptos. ¿Entonces porque utilizar espacios de trabajo analítico? La respuesta es: las grandes capacidades de análisis y el desempeño en tiempo de ejecución representan en muchos casos el mejor soporte para la toma a decisiones. Sin embargo, los esquemas relacionales son la mejor decisión en algunas situaciones, por ejemplo, si el almacén de datos soporta la creación de reportes rutinarios con carácter exploratorio, y ese uso aparenta ser estable [21].

Los espacios de trabajo analíticos son una buena opción para los siguientes requerimientos:

- Consultas ad-hoc o “no genéricas” de todas las áreas de datos.
- Cálculos avanzados como modelos, pronósticos, radios de crecimiento y tendencias.
- Escenarios “que pase si”.
- Cálculos en series de tiempo como retardos, elementos principales, promedios de movimiento.
- Cálculos complejos en tiempo de ejecución.
- Alto desempeño para calcular resúmenes de datos.

Los esquemas relacionales se pueden preferir si se tienen los siguientes requerimientos:

- Patrones de consulta predecibles y reportes preparados.
- Cálculos no avanzados.
- Cálculos complejos en tiempo de ejecución infrecuentes.
- Medidas con un gran número de dimensiones.
- Dimensiones con pocos niveles de agregación.

Después de revisadas las consideraciones anteriores y como resultado de la exploración tecnología llevada a cabo en el proyecto, se opto por crear la bodega de datos enteramente relacional. Los puntos a favor de esta decisión son los siguientes:

- El proyecto de bodega de datos para Unicauca Virtual, representa una etapa inicial de lo que puede ser en un futuro un sistema de soporte para la toma de decisiones completo, que incluya todas las necesidades de análisis que se puedan presentar una vez el ambiente virtual se encuentre en fase producción. Por lo tanto, es



recomendable realizar una implementación relacional que soporte las necesidades de análisis que hasta el momento se presentan, las cuales, no se encuentran definidas completamente y carecen de complejidad.

- Por el estado actual del macroproyecto Unicauca Virtual, los modelos dimensionales obtenidos hasta el momento y los que se obtendrán a corto plazo, están sujetos a modificaciones, producto de las nuevas necesidades de análisis de los usuarios finales. Estas modificaciones permitirán extender el significado de los modelos y la cantidad de información disponible para análisis.



---

# PARTE III – PROTOTIPO HERRAMIENTA OLAP

---



---

# CAPITULO 6 – ANALISIS

---

## 13 DEFINICION DE REQUERIMIENTOS

En términos generales, dentro del proyecto se creó un prototipo de herramienta de Procesamiento Analítico en Línea (OLAP), que le permite a los tutores y directivos del ambiente de aprendizaje virtual “Unicauca Virtual” realizar consultas analíticas sobre los datos relacionados con el estudiante que se encuentran almacenados en la Bodega de Datos construida en Oracle 10g Release2.

### 13.1 USUARIOS

Los usuarios podrán interactuar directamente con el prototipo sin ninguna restricción. Solamente se presenta una restricción a nivel de datos que depende de los privilegios en la base de datos que tienen los usuarios para acceder a los objetos ahí definidos.

- *Tutores*: acceden sin restricción al prototipo y pueden utilizar todas las funcionalidades definidas en él. Los datos con los que puede interactuar este usuario están vinculados directamente a las actividades que realizan los estudiantes dentro de Unicauca Virtual.
- *Directivos*: acceden sin restricción al prototipo y pueden utilizar todas las funcionalidades definidas en él. Los datos con los que interactúa el directivo se vinculan con las actividades de los estudiantes, los tutores y ambiente virtual.

### 13.2 METAS

El desarrollo del prototipo de herramienta OLAP tiene como meta fundamental, facilitar el acceso por parte de tutores y directivos de Unicauca Virtual a la información contenida en la Bodega de Datos para realizar análisis multidimensionales.

### 13.3 FUNCIONES DEL SISTEMA

Funciones del sistema son las acciones que el sistema debe hacer y pueden priorizarse de acuerdo a su grado de importancia o visibilidad en el sistema, es decir, se puede establecer cuales de ellas son esenciales para el funcionamiento adecuado del sistema



y cuales por el contrario pasarían inadvertidas pero que consumen tiempo y recurso [24].

Las funciones se pueden categorizar en:

- *Evidente*: la función debe realizarse y el usuario debería saber que se ha realizado.
- *Ocultas*: la función debe realizarse, aunque no es visible para los usuarios.
- *Superflua*: son funciones opcionales; su inclusión no repercute significativamente en el costo del proyecto ni en otras funciones.

De acuerdo a esta categorización, las funciones del prototipo se muestran a continuación en la Tabla 7.

Referencia	Función	Categoría
R.1	Solicitar y procesar la información del usuario para entrar al sistema.	Evidente
R.2	Conectarse al sistema gestor de base de datos.	Ocultas
R.3	Registrar información de salida del usuario.	Ocultas
R.4	Consultar la información de los objetos que forman el catalogo OLAP dentro del repositorio (dimensiones, medidas, jerarquías. etc.).	Ocultas
R.5	Mostrar la estructura que forma cada uno de los esquemas dimensionales contenidos en la bodega de datos.	Evidente
R.6	Consultar los diferentes valores que contiene cada uno de los niveles de la jerarquía en la dimensión.	Ocultas
R.7	Mostrar los diferentes valores que contiene cada uno de los niveles de la jerarquía en la dimensión.	Evidente
R.8	Relacionar los diferentes elementos de un esquema para formar el diseño de la consulta analítica.	Ocultas
R.9	Seleccionar y arrastrar un objeto de un cubo del catalogo al área de diseño.	Evidente
R.9.1	Colocar el objeto seleccionado en el área de diseño.	Evidente
R.9.2	Mostrar el diseño de la consulta analítica.	Evidente
R.10	Construir la sentencia para la consulta analítica.	Ocultas
R.11	Ejecutar la consulta analítica.	Ocultas



Referencia	Función	Categoría
R.12	Transformar los datos obtenidos (formato plano) de la consulta a un formato matricial.	Oculto
R.13	Presentar el análisis dimensional de forma matricial.	Evidente
R.14	Establecer relación entre los datos matriciales y la representación gráfica	Oculto
R.15	Presentar el análisis dimensional de forma gráfica	Evidente
R.16	Guardar un diseño de consulta analítica realizada por el usuario.	Evidente
R.16.1	Seleccionar ubicación donde se guardara del archivo de diseño.	Evidente
R.16.2	Crear archivo de diseño de consulta analítica.	Oculto
R.16.3	Crear estructura del archivo de diseño de consulta.	Oculto
R.17	Abrir un diseño de consulta analítica realizada por el usuario.	Evidente
R.17.1	Seleccionar ubicación donde se encuentra el archivo de diseño.	Evidente
R.17.2	Abrir archivo de diseño de consulta analítica.	Oculto
R.17.3	Leer estructura del archivo de diseño de consulta.	Oculto
R.18	Guardar espacios de trabajo que incluyen varias consultas analíticas realizadas por los usuarios.	Evidente
R.18.1	Seleccionar la ubicación donde almacenar el área de trabajo	Evidente
R.18.2	Crear el archivo del área de trabajo	Oculto
R.18.3	Crear la estructura del archivo del área de trabajo	Oculto
R.19	Abrir espacios de trabajo que incluyen varias consultas analíticas realizadas por los usuarios.	Evidente
R.19.1	Seleccionar la ubicación desde donde abrir el área de trabajo	Evidente
R.19.2	Abrir el archivo del área de trabajo	Oculto
R.19.3	Leer la estructura del archivo del área de trabajo	Oculto
R.20	Filtrar los datos de los niveles de una	Evidente



Referencia	Función	Categoría
	jerarquía en una dimensión	
R.20.1	Presentar una lista de los datos de nivel	Evidente
R.20.2	Seleccionar datos de nivel	Evidente
R.21	Restablecer el estado de los niveles de la jerarquía de una dimensión	Evidente
R.22	Cambiar la jerarquía de una dimensión	Evidente
R.22.1	Presentar las jerarquías de la dimensión	Evidente
R.22.2	Establecer la jerarquía	Evidente

Tabla 7. Funciones básicas del prototipo de herramienta OLAP

### 13.4 ATRIBUTOS DEL SISTEMA

Los atributos del sistema<sup>14</sup> se presentan en la Tabla 8 a continuación:

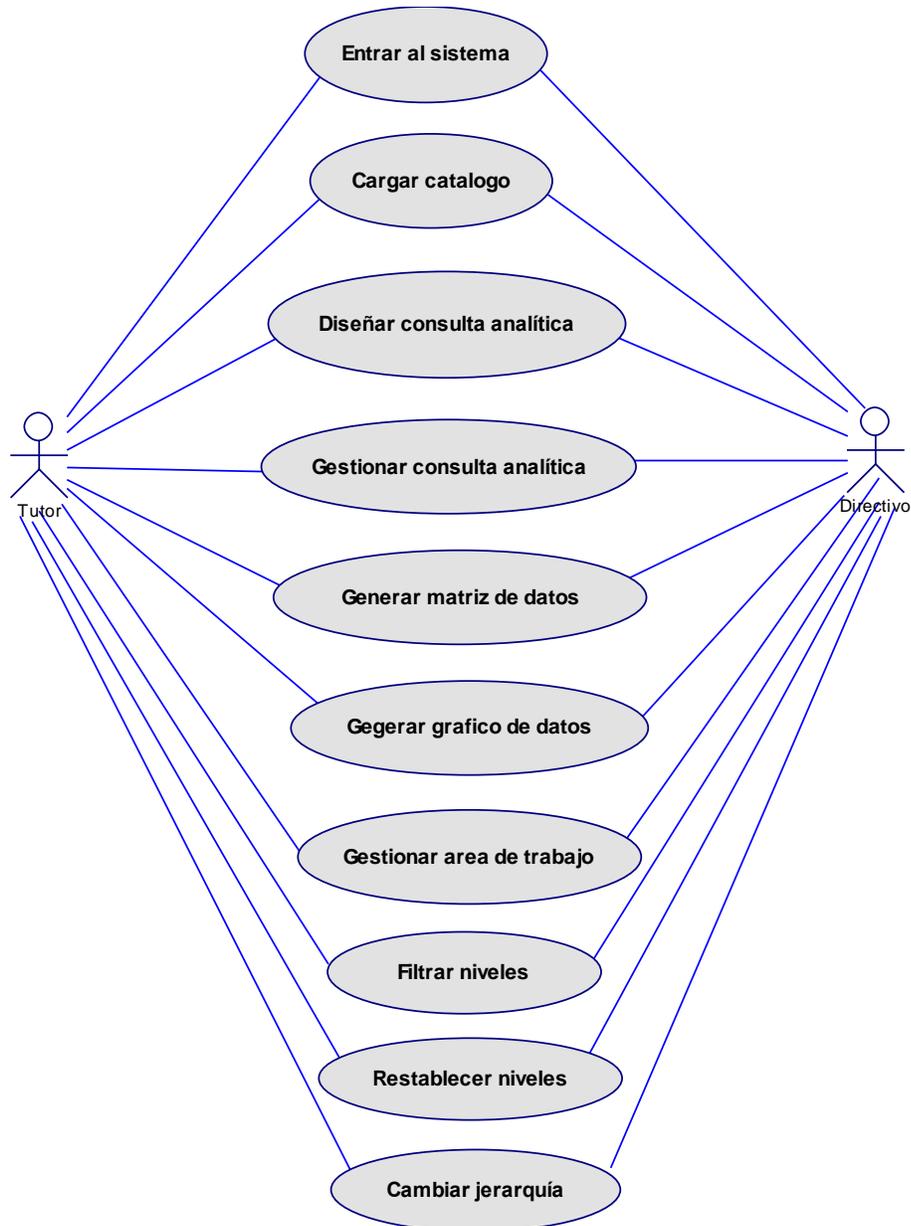
Atributo	Detalles y restricciones de frontera
Facilidad de uso	El usuario de la herramienta OLAP podrá realizar los análisis dimensionales sin requerir de un amplio conocimiento técnico de la herramienta y de la estructura de los datos contenidos dentro de la bodega de datos.
Metáfora de interfaz	Suministrar una fácil interacción del usuario y la herramienta a través de los periféricos del computador monitor, mouse y teclado. Además, presentar los diferentes análisis dimensionales requeridos por el usuario a través de grillas de datos y gráficos para un mejor entendimiento; proporcionando retroalimentación de las diferentes actividades que realiza dentro de la herramienta.
Tolerancia a fallos	Responder a través de mensajes explicativos a cada una de las acciones erradas por el usuario dentro de la herramienta.
Tiempo de respuesta	Algunas de las actividades que se llevan a cabo a través de la herramienta requieren de intervalos de tiempo un poco extensos, debido a la gran cantidad de información que se maneja y de las diferentes operaciones que se deben realizar.
Plataformas	Microsoft Windows XP professional, .Net Framework 1.1, Oracle 10g Release 2.

Tabla 8. Atributos del prototipo de herramienta OLAP.

<sup>14</sup> Los atributos del sistema están sujetos al ambiente del prototipo OLAP y no al ambiente del sistema global.

## 14 CASOS DE USO

La Figura 24 presenta un diagrama de los casos de uso generales realizados para cada uno de los usuarios del prototipo.



**Figura 24.** Diagrama general de casos de uso para el prototipo OLAP.

A continuación se describe uno de los casos de uso más importantes del prototipo, los demás se podrán encontrar en el ANEXO 9 de este documento.



## 14.1 CASO DE USO CARGAR CATALOGO

<b>Nombre del Caso de Uso:</b>		Cargar Catalogo
<b>Actores</b>		
<ul style="list-style-type: none"> <li>Tutor o directivo.</li> </ul>		
<b>Propósito:</b>	Cargar el catalogo OLAP desde la bodega de datos, es decir la estructura lógica que incluye los data marts, cubos, medidas, dimensiones, jerarquías por defecto y niveles.	
<b>Resumen:</b>	Una vez el usuario ingresa al sistema y decide utilizar la herramienta OLAP, el sistema presenta los nombre de los data marts que se encuentran en el catalogo OLAP. El usuario solicita al sistema cargar los cubos de cada uno de los data marts y luego la estructura de cada uno de los cubos, para ser desplegada y manipulada. El caso de uso finaliza, cuando se presenta completamente el catalogo OLAP al usuario.	
<b>Tipo:</b>	Primario.	
<b>Referencias Cruzadas:</b>	R.4, R.5, R.6, R.7.	
<b>Curso normal de los eventos</b>		
<b>Acción del Actor</b>		<b>Respuesta del Sistema</b>
1. El usuario ingresa al prototipo de herramienta OLAP		2. Cargar data marts contenidos en el catalogo OLAP con sus respectivos cubos.
3. Solicita visualizar los cubos que se encuentran definidos en un data mart particular.		4. Presenta los nombres de los cubos definidos dentro del catalogo OLAP para el data mart seleccionado.
5. Solicita visualizar las dimensiones y medidas que contiene un cubo seleccionado.		6. Carga y presenta las dimensiones y medidas del cubo seleccionado.
7. Solicita visualizar la estructura una dimensión seleccionada, es decir sus jerarquías por defecto con sus correspondientes valores de nivel.		8. Carga y presenta la estructura de la dimensión seleccionada.
<b>Curso alterno 1:</b>		
<b>Acción del Actor</b>		<b>Respuesta del Sistema</b>
		2, 4, 6, 8. El catalogo OLAP dentro de la base de datos, no se encuentra o esta mal definido. Se presenta al usuario un mensaje de error.



Trazabilidad	
Artefactos Anteriores	Artefactos Posteriores
<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>

Tabla 9. Análisis del caso de uso cargar catalogo.

## 15 MODELO CONCEPTUAL PRELIMINAR

El modelo conceptual preliminar obtenido para el desarrollo del prototipo de herramienta OLAP (ver Tabla 10) se compone de los siguientes conceptos:

- *Catalogo OLAP:* permite acceder a todos los objetos que conforman el catalogo OLAP definido en la base de datos.
- *Catalogo:* representa cada uno de los data marts con su estructura (cubos, medidas, dimensiones, jerarquías y niveles) definidos en la bodega de datos.
- *Cubo:* representa un cubo dentro de un data mart definido en la bodega de datos.
- *Medida:* representa una medida definida para un cubo en la bodega de datos.
- *Dimensión:* representa una dimensión relacionada a un cubo en la bodega de datos.
- *Jerarquía:* representa una jerarquía definida en la bodega de datos para una dimensión.
- *Nivel:* representa un nivel dentro de una jerarquía definida para una dimensión en la bodega de datos.
- *Atributo:* representa un valor descriptivo asociado a un nivel.
- *Reporte OLAP:* permite interactuar con los objetos lógicos obtenidos de la bodega para crear un reporte.
- *Sentencia Consulta:* construye una consulta SQL que será ejecutada en la bodega de datos al realizar un reporte.
- *Matriz de datos:* representación matricial de los datos obtenidos como resultado de ejecutar una consulta en la bodega de datos.
- *Dimensión Eje:* representa la estructura de la dimensión que se visualiza en un eje (horizontal o vertical) de la matriz al momento de diseñar la consulta analítica.
- *Relación Cubo-Tabla Dimensional:* mantiene información relacionada con las tablas dimensionales a las cuales esta vinculado un cubo.
- *Relación Dimensión-Tabla Dimensional:* mantiene información relacionada con la tabla dimensional a la cual esta vinculada la dimensión.



- *Relación Dimensión-Subdimension (Vista Materializada)*: mantiene información de la subdimension a la cual se hace referencia desde una dimensión
- *Relación Subdimension (Vista Materializada)-Dimensión*: mantiene información de la dimensión a la cual se referencia desde la subdimension.

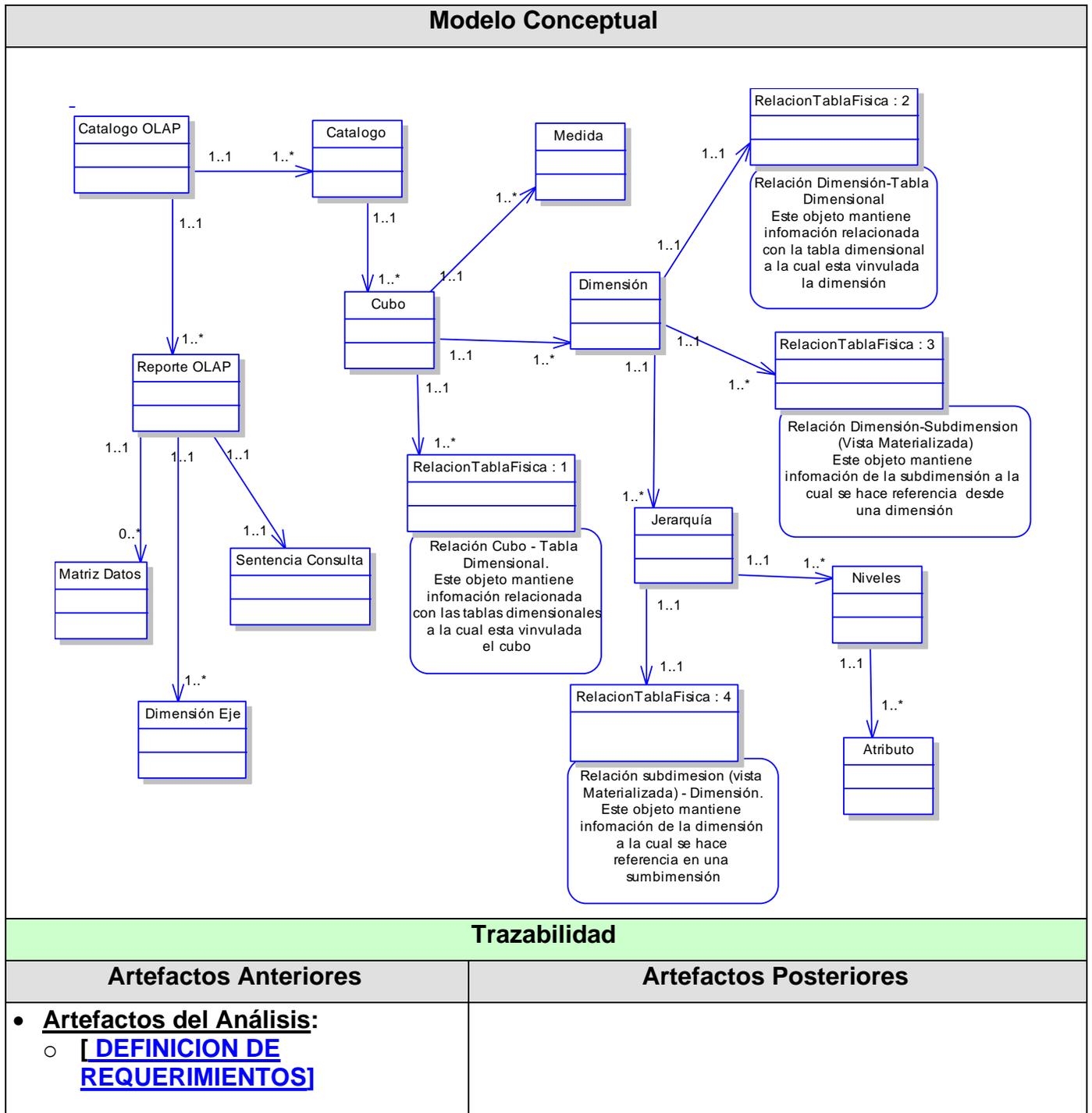


Tabla 10. Modelo Conceptual Preliminar.

## 16 DIAGRAMAS DE SECUENCIA DEL SISTEMA

A continuación se presenta el diagrama de secuencia correspondiente al caso de uso expuesto en el apartado *casos de uso* de este documento. Los demás diagramas se pueden encontrar en el ANEXO 10.

### 16.1 DIAGRAMA DE SECUENCIA CARGAR CATALOGO

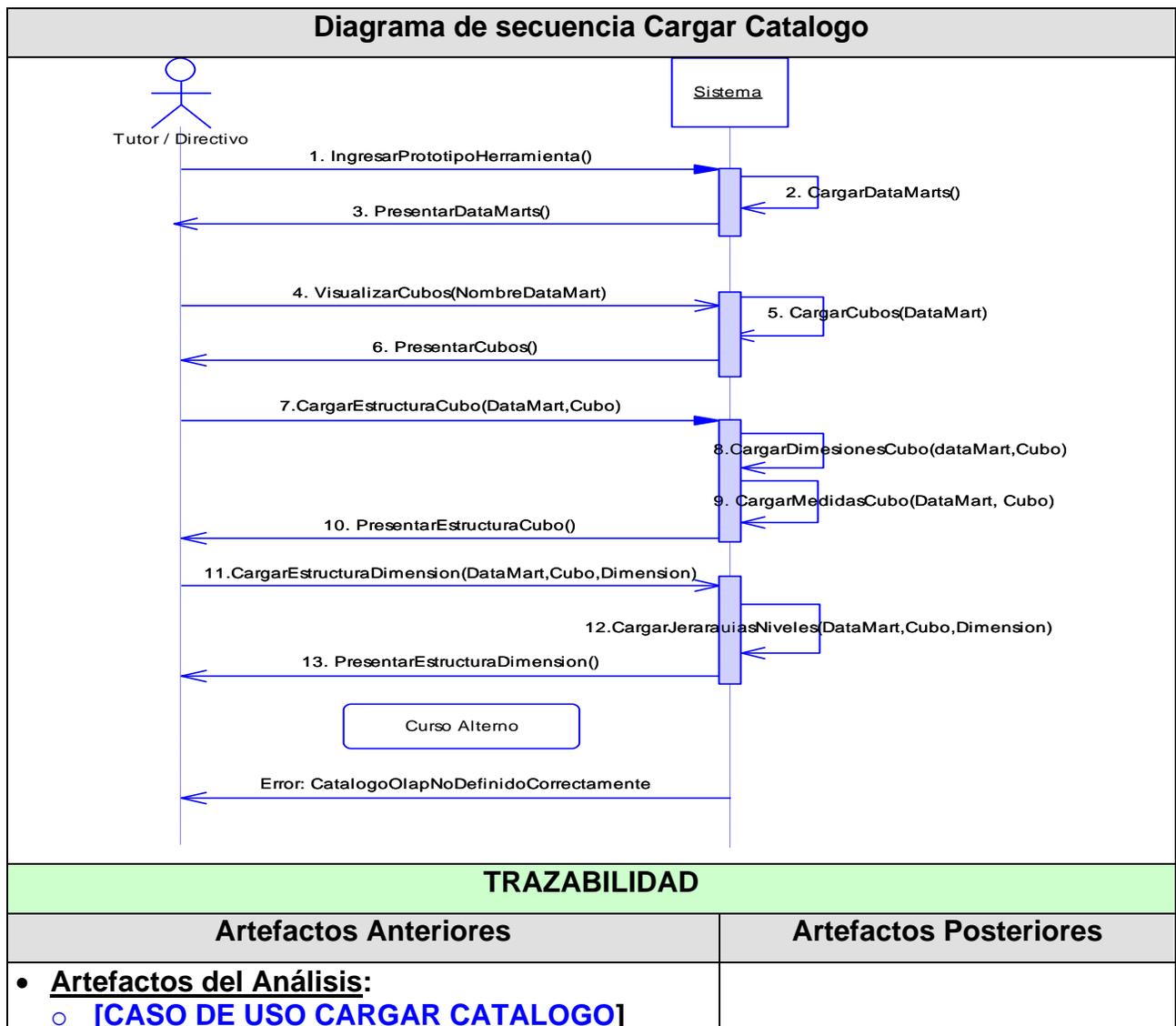


Tabla 11. Diagrama de secuencia caso de uso Cargar catalogo.

# CAPITULO 7 – DISEÑO

## 17 ARQUITECTURA DE LA APLICACIÓN

El prototipo de herramienta OLAP se integra a la aplicación Sistema Tutor Inteligente del tutor y del directivo desarrollados en otros subproyectos del macroproyecto Unicauca Virtual. Dado este hecho, se adoptó la misma arquitectura utilizada en estas aplicaciones buscando desde el diseño, la integración global del sistema.

La arquitectura cuenta con tres capas (ver Figura 25), lo que representa un factor de mucha utilidad en la construcción de una aplicación, dado que permite división de responsabilidades, flexibilidad, escalabilidad y descomposición de las capas en otras menos densas. Las tres capas de la arquitectura son:

- *Presentación*: componentes de interfaz de usuario.
- *Lógica del Negocio*: objetos que representan conceptos del dominio del problema, que cumplen con los requisitos de la aplicación.
- *Lógica de Servicios*: objetos que prestan servicios de soporte, como el acceso a datos.

En cada capa se definen componentes necesarios para un óptimo funcionamiento.

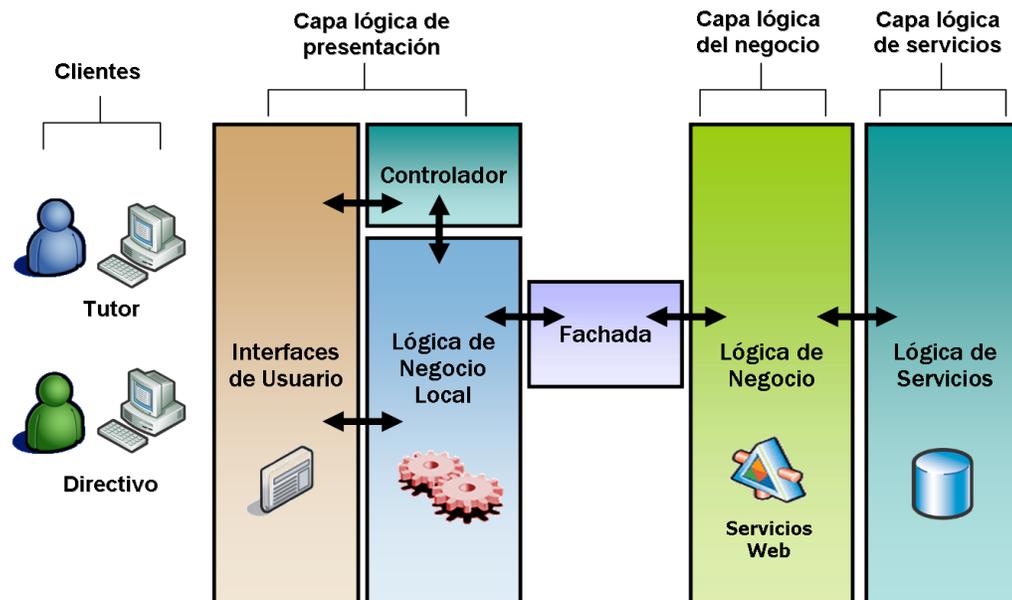


Figura 25. Arquitectura general del sistema.



## 17.1 CLIENTES

- *El Directivo:* su funcionalidad es ofrecer al tutor las herramientas adecuadas para el establecimiento de los parámetros que este último desea establecer en un curso determinado, además juega como el vocero de todos los tutores, es el encargado de establecer las reglas del juego generales y de dar a los tutores los puntos exactos donde estos puedan entrar a trabajar. El directivo se comporta como un administrador pero no de la herramienta sino del proceso pedagógico llevado a cabo por el estudiante y el tutor.
- *El Tutor:* personaje activo en el proceso de aprendizaje del estudiante, pues es quien establece las condiciones iniciales que según su experiencia sean las adecuadas para dar inicio a dicho proceso. Es el tutor quien tendrá la potestad de establecer la estructura que todo estudiante deberá seguir para lograr los objetivos del curso, pues aunque el aprendizaje en línea dependa casi exclusivamente de las condiciones bajo las cuales quiera trabajar el estudiante, no significa que éste no deba cumplir ciertos parámetros que como se dijo anteriormente son establecidos por el profesor.

## 17.2 CAPA DE LÓGICA DE PRESENTACIÓN

La capa de lógica de presentación expone los servicios de la capa de lógica de negocios a los usuarios. Sabe cómo procesar una petición de cliente, cómo interactuar con la capa de lógica de negocios, y cómo seleccionar la siguiente vista a mostrar.

Esta capa se estructura siguiendo el patrón Modelo-Vista-Controlador, que permite separar el modelo que se está manejando, de la forma como se muestra (la vista). La vista (un control en un formulario, un gráfico), puede que esté interesada en conocer cuándo cambia el modelo subyacente.

En la aplicación, este patrón se representa así:

- *Modelo:* representado por la lógica de negocios local, que contiene la lógica de los datos que permite interactuar con la capa de lógica de Negocio a través de la interfaz común Fachada.
- *Vista:* contiene los componentes de interfaz, como: formularios Windows, controles de usuario, archivos xml, entre otros.
- *Controlador:* el código subyacente de los formularios Windows y de las clases que se encarga de hacer los llamados respectivos a los métodos de la lógica de negocio local, dependiendo de las acciones de los usuarios.

Esta capa se compone de las siguientes partes:



- *Interfases de usuario:* aquí se distinguen claramente tantas aplicaciones como usuarios finales existentes, es decir, cada cliente posee una aplicación independiente que le permite realizar las labores exclusivas de su rol y a la vez las labores que guardan cierta relación con las demás aplicaciones (labores definidas en conjunto). Esta capa es con la que el usuario interactúa directamente y a la que hace las peticiones de funcionalidad.
- *Controlador:* Actúa como un controlador de la lógica de interfaz
- *Lógica de negocio local:* La lógica de negocios local cuenta con un conjunto de clases (una por cada elemento del dominio) que representa el modelo de negocio de la aplicación, implementando métodos para acceder a los elementos del dominio, cada una de estas clases ha sido creada con el objetivo de dar organización a todas y cada una de las funcionalidades que provee el sistema, haciendo una clasificación de las mismas de acuerdo al usuario que las implementa o al elemento del dominio que dicha funcionalidad afecta.

### 17.3 FACHADA

Sirve como intermediaria entre la instancia de la lógica local de negocio y la lógica de negocio. La clase fachada instancia cada uno de los elementos del dominio que se encuentran en la lógica de negocio.

### 17.4 CAPA LÓGICA DEL NEGOCIO

Contiene los objetos y servicios de negocio de la aplicación. Recibe peticiones de la capa de presentación, procesa la lógica de negocio basada en las peticiones, y media en los accesos a los recursos de la capa de lógica de servicios. Esta capa está basada en servicios Web XML y contiene un conjunto de objetos que permiten encapsular la lógica del dominio de la bodega de datos.

### 17.5 CAPA LÓGICA DE SERVICIOS

La capa de lógica de servicios representa el mecanismo por el cual se manipula y persiste la información, es decir, esta capa se encarga de recibir solicitudes de almacenamiento o recuperación de información desde la capa lógica del negocio a través de diferentes métodos que permiten interactuar con el servidor de la base de datos Oracle 10g.

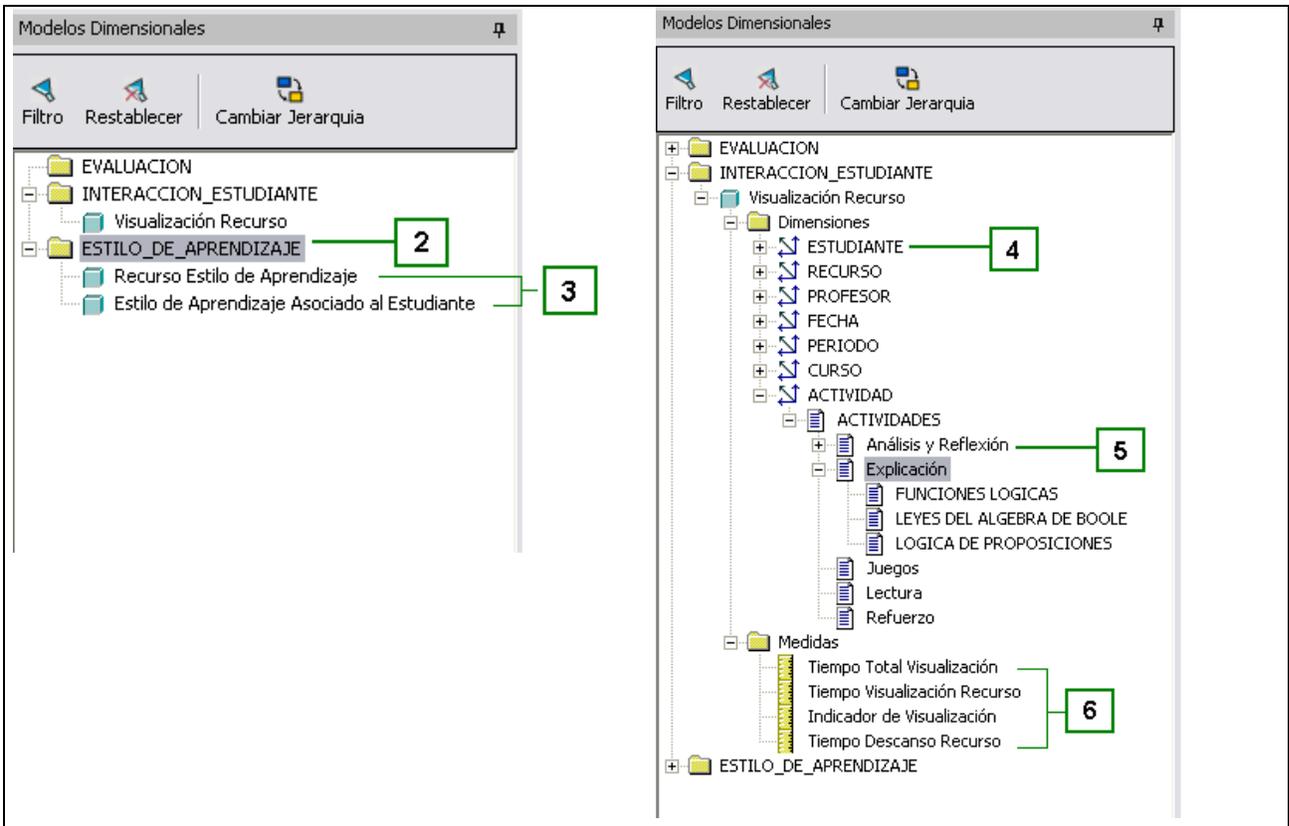


## 18 CASOS DE USO REALES

A continuación se presentan uno de los casos de uso reales más importantes correspondiente a los casos de uso expuestos en la etapa de análisis; los demás casos de uso reales se pueden encontrar en el ANEXO 11.

### 18.1 CASO DE USO REAL CARGAR CATALOGO

The screenshot displays the OLAP tool interface for the 'Caso de uso real Cargar catalogo' scenario. The title bar reads 'HERRAMIENTA OLAP - [ANÁLISIS DE DATOS 1]'. The menu bar includes 'Análisis' and 'Ventana'. The toolbar contains icons for 'Generar matriz de datos', 'Generar gráfico de datos', 'Abrir análisis', and 'Guardar análisis'. The main workspace is divided into several sections: a left sidebar with 'Modelos Dimensionales' and a tree view showing 'EVALUACION', 'INTERACCION\_ESTUDIANTE', and 'ESTILO\_DE\_APRENDIZAJE'; a top navigation bar with 'Análisis Matricial' and 'Análisis Gráfico'; and a main area with a 'Resumen' section containing fields for 'Nombre del Data Mart', 'Nombre del Cubo', and 'Descripción del Cubo'. A green box labeled '1' highlights the 'INTERACCION\_ESTUDIANTE' folder in the tree view. A large green box labeled 'Area Catalogo OLAP' is positioned in the lower-left area of the main workspace.



**Curso normal de los eventos**

Acción del Actor	Respuesta del Sistema
1. El usuario ingresa al prototipo de herramienta OLAP.	2. Carga los data marts [1] contenidos en el catalogo OLAP.
3. Dando doble clic sobre un data mart [2], solicita visualizar los cubos que se encuentran definidos en él.	4. Presenta en el área del catalogo OLAP los nombres de los cubos [3] definidos en la bodega de datos para el data mart seleccionado.
5. Dando doble clic sobre un cubo, visualiza las dimensiones [4] y medidas [6] que contiene.	6. Carga y presenta las dimensiones y medidas del cubo seleccionado.
7. Al dar clic sobre cada una de las dimensiones [4], se visualiza su estructura, es decir sus jerarquías por defecto con sus correspondientes valores de nivel [5].	8. Carga y presenta la estructura de la dimensión seleccionada.

**TRAZABILIDAD**

Artefactos Anteriores	Artefactos Posteriores
-----------------------	------------------------



---

<ul style="list-style-type: none"><li>▪ <b>Artefactos del Análisis:</b><ul style="list-style-type: none"><li>○ <b><u>[CASO DE USO CARGAR CATALOGO]</u></b></li></ul></li></ul>	
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

**Tabla 12.** Caso de uso real Cargar Catalogo.

# CAPITULO 8 – CONSTRUCCIÓN DEL PROTOTIPO DE HERRAMIENTA OLAP

## 19 VISTA DE IMPLEMENTACIÓN PROTOTIPO OLAP

Esta vista describe cómo se implementaron los componentes agrupándolos en subsistemas organizados en capas y jerarquías, ilustra, además las dependencias entre éstos.

En la Figura 26 se muestra cómo el patrón de diseño en capas planeado en la arquitectura del sistema puede llevarse a una implementación desarrollando componentes físicos:

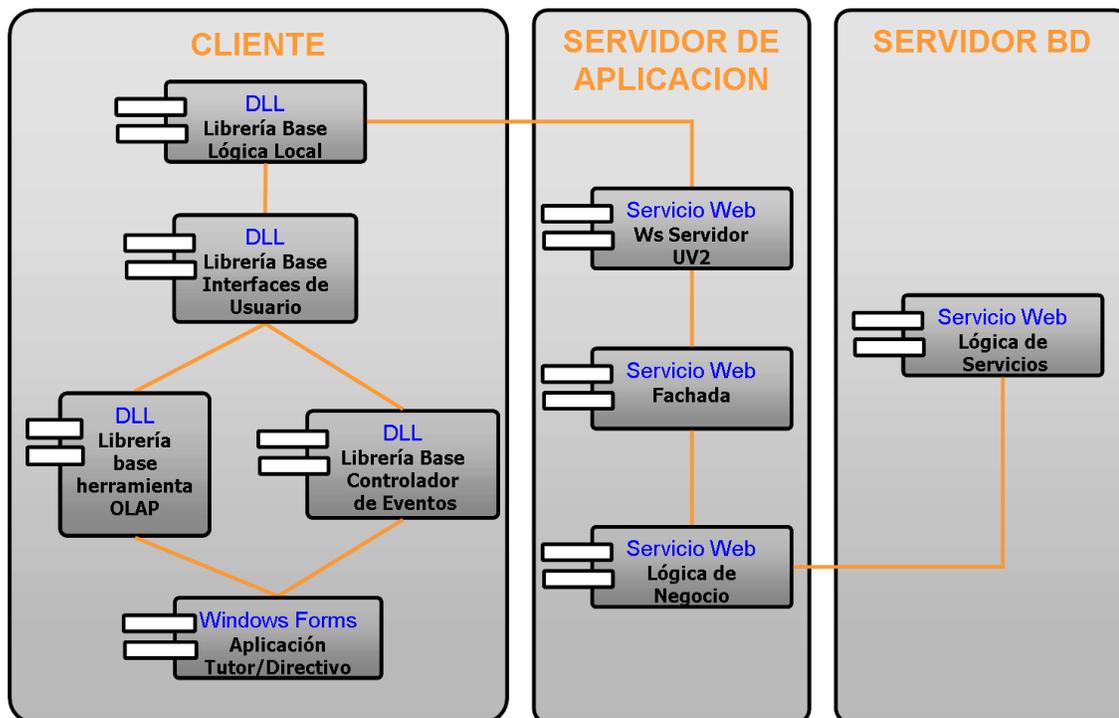


Figura 26. Distribución de componentes implementados.



## 19.1 CLIENTE

- *Aplicaciones Windows*: Aplicación del Tutor o Aplicación del Directivo. Cada una esta implementada a través de un Proyecto *Windows Forms* compuesto por una interfaz de múltiples documentos (MDI).
- *Librería Base controlador de Eventos*: se implemento como una DLL que contiene clases que consiguen desacoplar las interfaces de usuario y la lógica de negocio logrando la separación vista controlador en el sistema.
- *Librería Base Interfaces de Usuario*: Es una DLL que contiene controles de usuario y Formularios necesario para la recolección y visualización de datos del sistema.
- *Librería base herramienta OLAP*: Es un componentes que contiene una forma la cual controla las demás formas del prototipo OLAP.

## 19.2 SERVIDOR DE APLICACIÓN

- *Servidor UV2*: Contiene los *Servicios Web* correspondientes a la fachada del sistema y la lógica de negocio.
- *Fachada*: Simplifica el acoplamiento entre el *Cliente* y la *Lógica de Negocio* ocultando el sistema complejo detrás de este servicio que hace las veces de pantalla, consigue aislar los posibles cambios que se puedan producir en alguna de las partes.
- *Lógica de Negocio*: Esta parte de la arquitectura se desarrollo utilizando servicios Web que permiten crear y garantizar las reglas de los clientes, y del sistema en general.

## 19.3 SERVIDOR DE BASE DE DATOS

- *Lógica de Servicios*: Se implementa a través de servicios Web y permite una separación del motor de base de datos sin impactar la estabilidad del sistema.



## 20 CONSIDERACIONES ESPECIALES DE LA IMPLEMENTACIÓN DEL PROTOTIPO.

Para desarrollar el prototipo de herramienta OLAP, se hizo necesario tener en cuenta aspectos de la bodega de datos y del motor de base de datos en el cual se construyo.

La tecnología Oracle para la creación de bodegas de datos permite manejar Metadatos CWM para que aplicaciones de terceros, como el prototipo de herramienta OLAP construido, pueda acceder e interactuar con los objetos dimensionales que componen la bodega. Esto se logra a través de un paquete de vistas del sistema [15], las cuales relacionan los objetos físicos de la bodega de datos (tablas dimensionales, tablas de hechos) con los objetos lógicos (dimensiones, cubos, jerarquías, niveles, medidas).

Por otro lado, a nivel de la base de datos, Oracle extiende el lenguaje SQL para permitir el manejo de datos dimensionales contenidos dentro de una bodega de datos relacional. Estas extensiones se relacionan con el manejo de grupos de datos y la ejecución de cálculo en estos grupos.

A continuación se presenta una descripción de algunas vistas y extensiones SQL, las demás se pueden encontrar en el ANEXO 12 y en el ANEXO 13 respectivamente.

### 20.1 VISTAS DEL SISTEMAS PARA EL MANEJO DEL CATALOGO OLAP EN ORACLE

#### 20.1.1 ALL\$OLAP2UCATALOGS

Sinónimo: ALL\_OLAP2\_CATALOGS

Cada fila representa una carpeta de medidas OLAP (catalogo o data Mart). Las carpetas de medidas son un medio de agrupación de medidas relacionadas para un tema de negocio o data mart. Por ejemplo, todas las medidas que almacenan información acerca de una línea de producto dado podrían ser recopiladas en una carpeta de medidas “línea de productos”.

Las carpetas de medidas son esquemas independientes. Todos los usuarios pueden ver todas las carpetas de medidas definidas en la bodega de datos, aun si ellos no tienen privilegios de acceso a las medidas.

Las carpetas de medidas pueden ser anidadas dentro de otras carpetas de medidas.

Columna	Tipo de dato	NULL	Descripción
CATALOG_ID	NUMBER	NOT	Identificador de la carpeta de



Columna	Tipo de dato	NULL	Descripción
		NULL	medidas.
CATALOG_NAME	VARCHAR2(30)	NOT NULL	Nombre de la carpeta de medidas.
PARENT_CATALOG_ID	NUMBER		Identificador de la carpeta de medidas padre. Esta columna es NULL para carpetas que no tengan raíz como la carpeta padre.
DESCRIPTION	VARCHAR2(2000)		Descripción de la carpeta de medidas.

Tabla 13. Descripción columnas vista ALL\$OLAP2UCATALOGS.

### 20.1.2 ALL\$OLAP2UCUBE\_DIM\_USES

Sinónimo: ALL\_OLAP2\_CUBE\_DIM\_USES

Cada fila representa una asociación entre un cubo y una dimensión. Una dimensión puede ser asociada más de una vez en el mismo cubo, pero cada asociación es especificada en una fila separada, bajo su propio alias de dimensión único.

Columna	Tipo de dato	NULL	Descripción
CUBE_DIMENSION_USE_ID	NUMBER	NOT NULL	Identificador de la asociación entre un cubo y una dimensión.
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo.
DIMENSION_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
DIMENSION_ALIAS	VARCHAR2(30)		Alias de la dimensión, provee identificación única de la dimensión usada dentro del cubo.
DEFAULT_CALC_HIERARCHY_NAME	VARCHAR2(30)		La jerarquía por defecto usada para hacer "drilling"



Columna	Tipo de dato	NULL	Descripción
DEPENDENT_ON_DIM_USE_ID	NUMBER		down” o bajar dentro de la dimensión. ID de la asociación cubo/dimensión de la cual depende la asociación cubo/dimensión. (solamente para OLAP 1 Catálogo (CWM1))

Tabla 14. Descripción columnas vista ALL\$OLAP2UCUBE\_DIM\_USES.

## 20.2 EXTENSIONES SQL PARA BODEGAS DE DATOS RELACIONALES EN ORACLE

La agregación es una parte fundamental de la bodega de datos. Oracle ofrece las siguientes extensiones para el manejo de agregaciones [11]:

- Extensiones CUBE y ROLLUP para la cláusula GROUP BY.
- Tres funciones GROUPING.
- La expresión GROUPING SETS.

### 20.2.1 Extensión CUBE para la cláusula GROUP BY

CUBE toma un conjunto específico de columnas agrupadas y crea subtotales para todas las posibles combinaciones. En términos de análisis multidimensional, CUBE genera todos los subtotales que podrían ser calculados para un cubo de datos con las dimensiones específicas. Si se especifican “n” columnas en la cláusula CUBE, habrán  $2^n$  combinaciones de subtotales retornados.

La extensión CUBE puede ser útil en cualquier situación que requiera reportes “cross-tabular”, generación de tablas resumidas y consultas que usen columnas de múltiples dimensiones.

### 20.2.2 Funciones GROUPING

Se presentan dos escenarios especiales con el uso de las extensiones CUBE o ROLLUP. El primero, determinar programáticamente, cual conjunto de resultados obtenidos es un subtotal y como relacionar exactamente un nivel de agregación a un



subtotal dato; y el segundo, como determinar que los valores NULOS obtenidos son NULOS almacenados o creados por las cláusulas ROLLUP o CUBE.

Para manejar este tipo de escenarios se utiliza la función *GROUPING*, donde usando una columna como argumento, devuelve “1” cuando encuentra un valor NULO creado por una operación ROLLUP o CUBE. Es decir, si el nulo indica que la fila es un subtotal, *GROUPING* retorna un 1, de cualquier otra forma, retorna 0.

### 20.3 ¿POR QUE UN PROTOTIPO DE HERRAMIENTA OLAP?

En el mercado existen diferentes herramientas OLAP, que permiten consultar la información de la Bodega de Datos. Oracle, por ejemplo, tiene una aplicación llamada “*Oracle Business Intelligence Spreadsheet Add-In*”, la cual se adiciona a Microsoft Excel y permite a los usuarios finales desplegar y navegar a través de los datos Oracle OLAP utilizando una hoja de cálculo de Excel.

Otras casas de software como “CNS International” o “Cognos”, líderes en inteligencia de negocios, ofrecen soluciones OLAP muy completas, “*Data Warehouse Explorer*” por ejemplo, para acceder a datos basados en sistemas Microsoft, o “*Power Play*” para otras bases de datos como Oracle, Microsoft, SAP BW e IBM. Aunque existen muchas soluciones, la mayoría de ellas involucran elevados costos adicionales, no estimados para la elaboración de este proyecto. El mercado también ofrece otras alternativas enmarcadas dentro del software libre como PALO [25], Mondrian [26] y JPALO [27] que a pesar de no representar incrementos en los costos del proyecto no permiten conectividad con el motor de base de datos ORACLE. Dado lo anterior el grupo del proyecto decidió iniciar el desarrollo de una aplicación prototipo de herramienta OLAP, que se encuentre integrada a los demás módulos de la Fase 2 de Unicauca Virtual y que permite consultar los datos que se encuentran dentro de los esquemas dimensionales implementados en Oracle 10g Release 2.



---

# CAPITULO 9 – DESCRIPCION DEL PROTOTIPO

---

## 21 FUNCIONALIDADES DEL PROTOTIPO

Para el desarrollo del prototipo de herramienta OLAP se tuvieron en cuenta las funciones que además de permitir cumplir con los objetivos del proyecto facilitara a los usuarios finales un mediano nivel de interactividad.

Las herramientas OLAP que se encuentran en el mercado, ofrecen una gran variedad de funcionalidades que usualmente se complementan con otros tipos de sistemas de información como CRM (Customer Relationship Management) o ERP (Enterprise Resource Planning) para ofrecer una solución bastante robusta a las personas que toman decisiones en las empresas, incrementando el numero de funcionalidades involucradas. Sin embargo, se reconocen algunas tareas generales como:

- *Visualización de componentes de la bodega de datos:* como cubos, medidas, dimensiones, jerarquías, niveles.
- *Creación de la consulta analítica:* lo cual involucra la construcción previa de un diseño, es decir que objetos dimensionales se tendrán en cuenta para la consulta.
- *Visualización de datos:* por lo general el resultado de la ejecución de una consulta analítica se visualiza en una grilla de datos o en un grafico.
- *Gestión de las consultas analíticas:* facilitar almacenamiento permanente de los reportes creados en la herramienta.
- *Filtrado de componentes y valores:* filtrado de componentes de la bodega de datos y de los resultados presentados en la grilla.
- *Cálculos sobre los resultados:* permitir aplicar funciones (estadísticas, probabilísticas, etc.) sobre los resultados presentados.
- *Funciones especiales sobre los resultados:* aplicar sobre los resultados obtenidos funcionalidades relacionadas con la forma como se visualizan los datos, por ejemplo, ordenar de mayor a menor, o supresión de valores nulos o valores ceros, comparación de dos o mas columnas, etc.
- *Interoperabilidad con otras aplicaciones:* los resultados se pueden exportar a formato .pdf, hojas de cálculo de Excel, .doc (Microsoft Word), .html, etc. y pueden servir datos de entrada para otros sistemas de información dentro de la organización.

El prototipo de herramienta OLAP desarrollado, implemento algunas de las características generales descritas con algunas restricciones. Las funcionalidades se pueden visualizar en la pantalla principal de la herramienta, mostrada en la Figura 27.

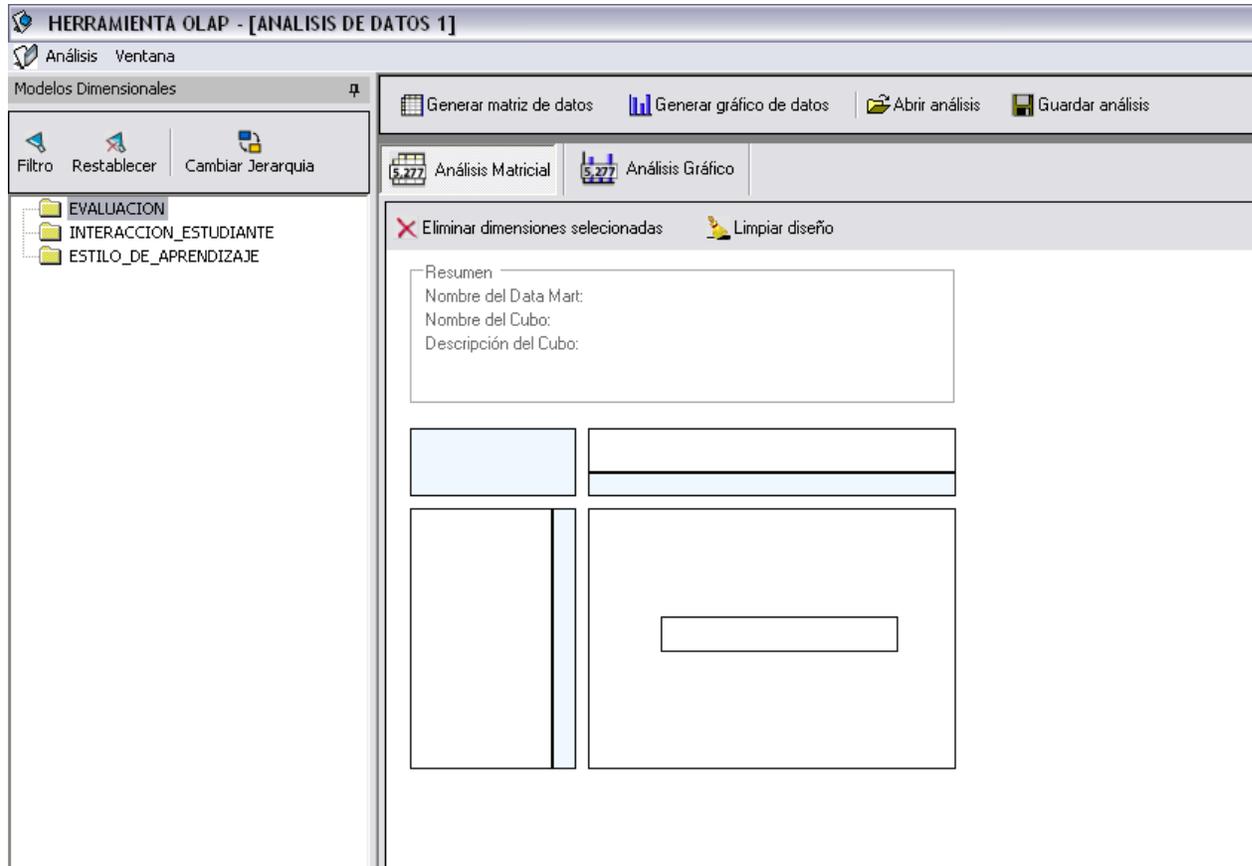
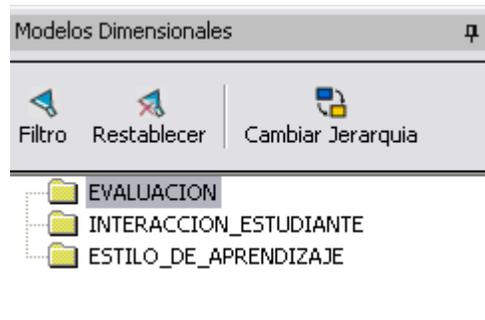


Figura 27. Pantalla principal del prototipo de herramienta OLAP.

## 21.1 VISUALIZACIÓN DE COMPONENTES DE LA BODEGA DE DATOS

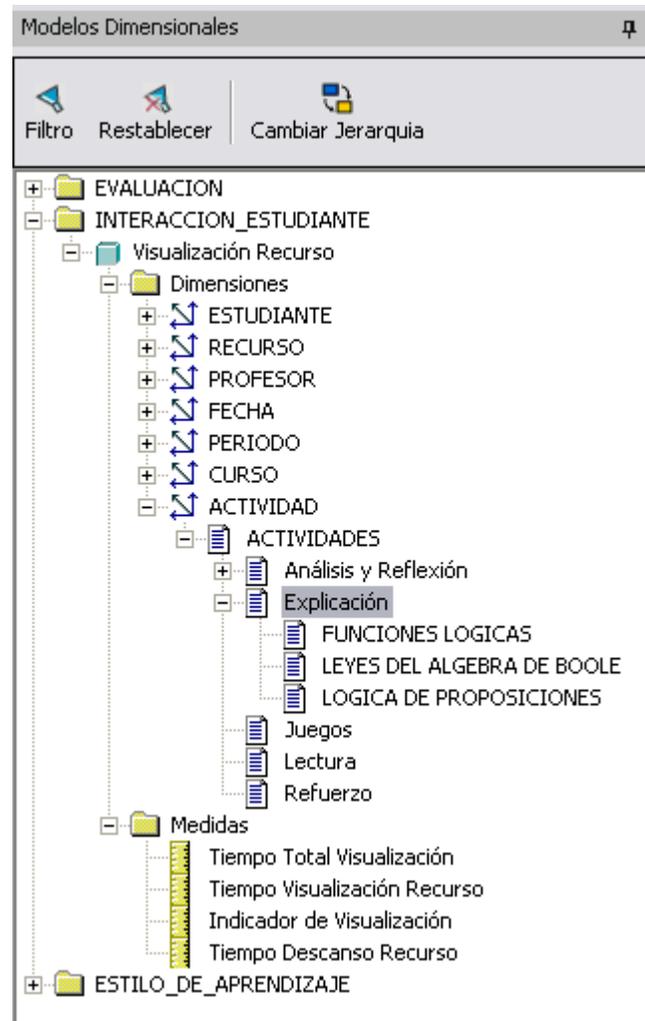
En el prototipo desarrollado, los componentes definidos en la bodega de datos se pueden visualizar en el área denominada “Modelos Dimensionales” (ver Figura 28).



**Figura 28.** Área modelos dimensionales.

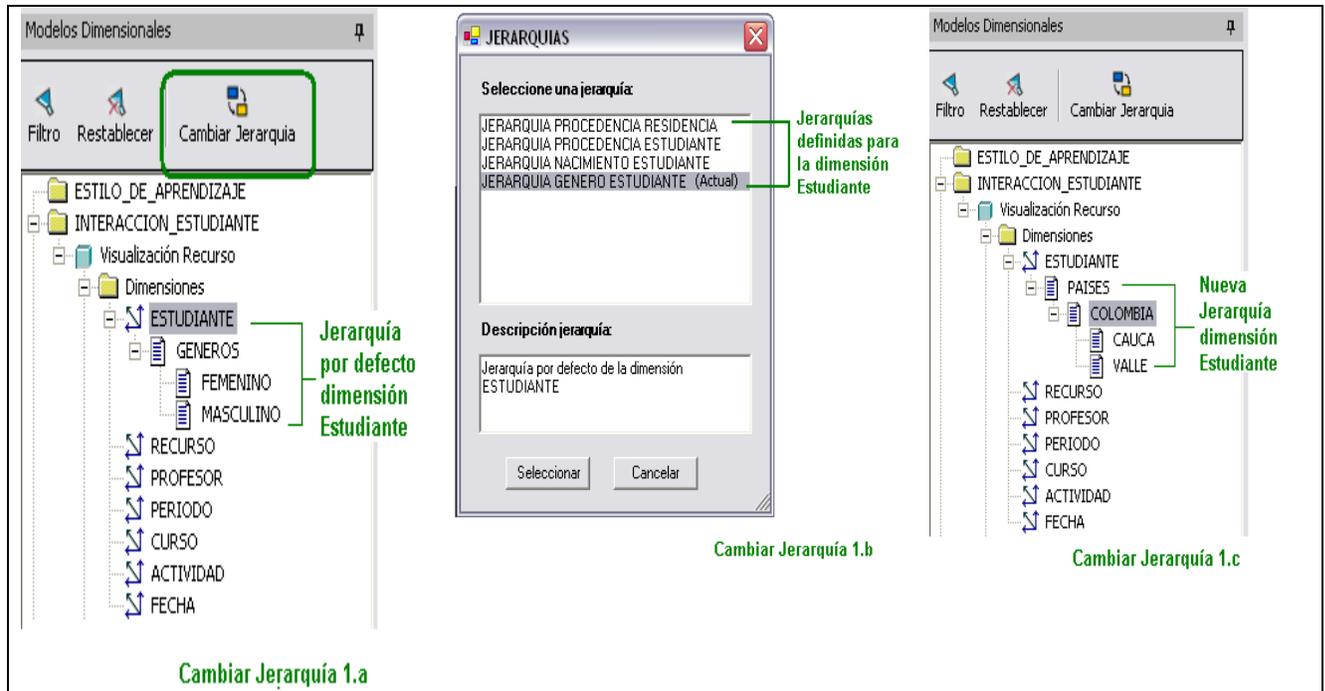
Esta área presenta en primera instancia los data marts definidos en la bodega de datos, posteriormente con la interacción del usuario, se pueden visualizar los cubos contenidos dentro de un data mart y de igual forma las dimensiones y la estructura de cada dimensión.

La Figura 29 presenta los data marts construidos para Unicauca virtual (Ver capítulo 4 de este documento) y los objetos dimensionales definidos en el data mart *Interacción Estudiante*. Los objetos que aparecen con el símbolo cubo (■) representan cubos dentro del data mart, los objetos con el símbolo dimensión (↕) representan las dimensiones dentro del cubo y los que tienen el símbolo dato de nivel (☰) representan los datos de nivel de una jerarquía en la dimensión. Las medidas definidas dentro del cubo, se representan con el símbolo medida (■).



**Figura 29.** Área modelos dimensionales – Data Marts, Cubos, Dimensiones, Jerarquías, Datos de nivel.

Las dimensiones en el prototipo, presentan las jerarquías definidas por defecto al momento de la creación de la dimensión. Si la dimensión tiene especificadas mas jerarquías, el usuario puede cambiar entre ellas usando la opción cambiar jerarquía (ver Figura 30)



**Figura 30.** Cambiar Jerarquía de una dimensión. **Cambiar Jerarquía 1.a.** Jerarquía por defecto (género) de la dimensión Estudiante **Cambiar Jerarquía 1.b.** Pantalla que presenta las dimensiones definidas para la dimensión Estudiante. **Cambiar Jerarquía 1.c.** Jerarquía *paises de residencia* para la dimensión estudiante.

## 21.2 CREACIÓN DE LA CONSULTA ANALÍTICA

La herramienta OLAP permite crear consultas analíticas a través del diseño de las mismas. Antes de que el usuario pueda visualizar los datos, debe crear un diseño que represente, qué objetos dimensionales harán parte de la consulta analítica. Esto lo puede hacer utilizando el área *Modelos Dimensionales* y el área de *Diseño de Consulta* que se presenta en cada ventana de análisis dimensional (ver Figura 31)

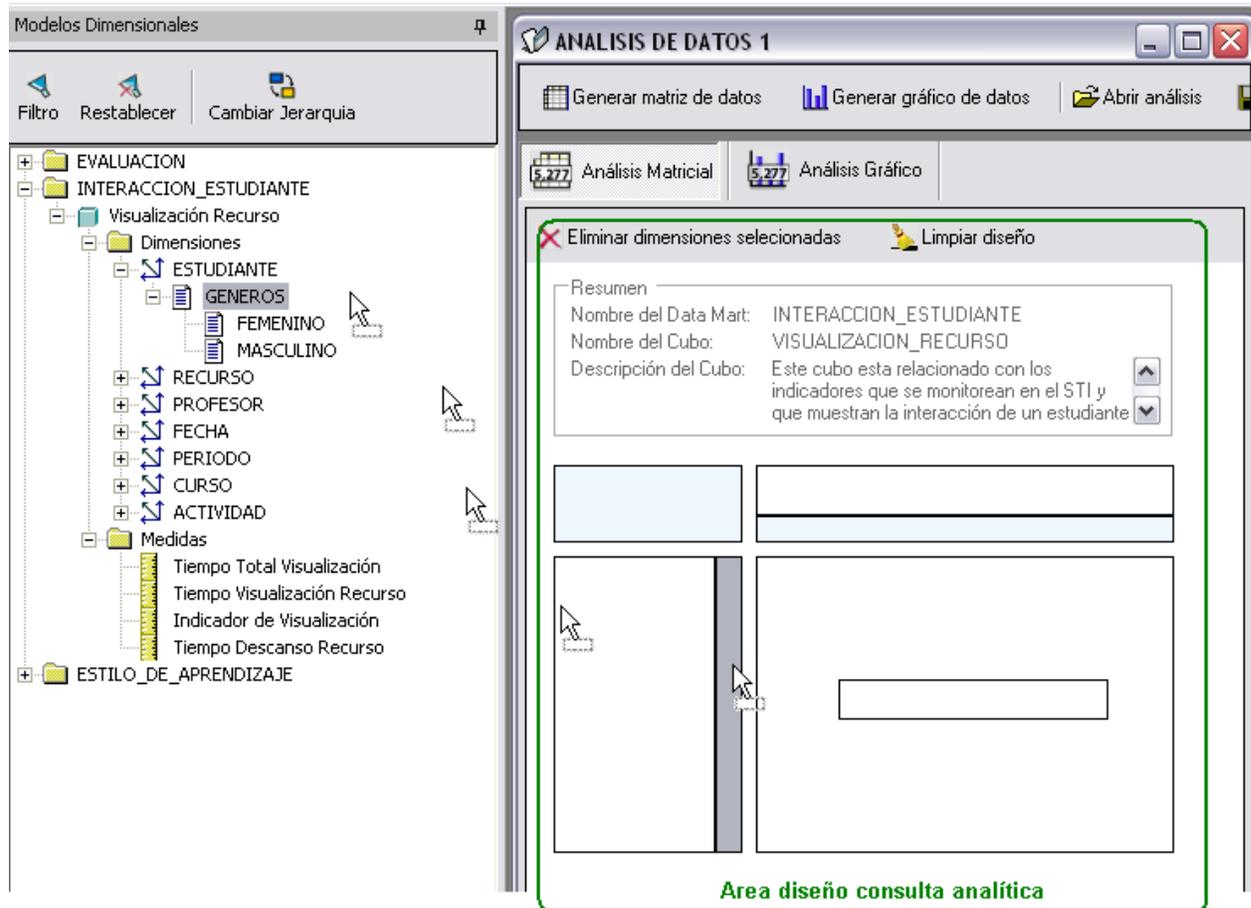


Figura 31. Creación de una consulta analítica.

El usuario debe utilizar la opción arrastrar y colocar del prototipo para ubicar los objetos dimensionales en el área de diseño. Estos objetos son dimensiones u objetos de la dimensión (como datos de nivel en la jerarquía) o medidas. Cada objeto tiene una sección determinada en el diseño, es decir existe una sección para ubicar las dimensiones y otra para las medidas. Las dimensiones deben ubicar en alguno de los dos ejes (vertical u horizontal), no se pueden poner más de tres dimensiones en cada eje y solamente se puede manejar una medida por diseño.

El área de diseño de consulta analítica, presenta una sección denominada *Resumen* donde se presenta una descripción general del cubo en el cual se esta trabajando en el momento de diseñar y dos opciones para el manejo del diseño: *Eliminar dimensiones seleccionadas* y *limpiar diseño*. Estas opciones permiten que el usuario, en tiempo de diseño, cambie los objetos involucrados en el mismo o inicie un diseño desde cero.

Por otro lado, se debe tener en cuenta que en un diseño solo se pueden vincular objetos del mismo data mart. Si se quiere trabajar con otro data mart, es necesario abrir una nueva ventana de análisis utilizando el menú Análisis del prototipo.

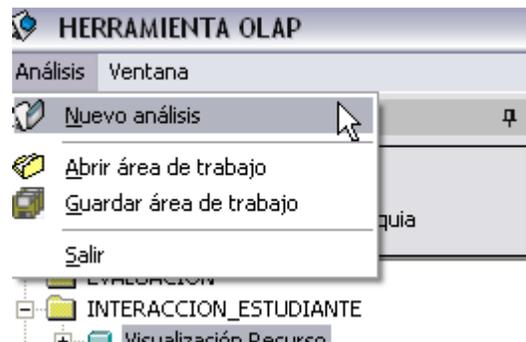


Figura 32. Abrir una nueva ventana de análisis.

## 21.3 VISUALIZACIÓN DE DATOS

Una vez el diseño este terminado, es decir cuente con mínimo una dimensión en cada eje y una medida, se pueden visualizar los datos obtenidos, basado en el diseño realizado. Para esto el usuario debe oprimir el botón *Generar matriz de datos* presente en cada ventana de análisis o sobre el botón *generar gráfico de datos* (ver Figura 33). De esta manera se permite al usuario contar con dos formas distintas de visualizar la misma información

Timeo Visualización Recurso	FEMENINO	MASCULINO	TOTAL GENEROS	ESTUDIANTE
BASES DE DATOS	NA	NA	NA	
INTRODUCCION A LA INFORMATICA	31,96	156,17	188,13	
PROGRAMACION ORIENTADA A OBJETOS	60,23	112,77	173	
TOTAL TODAS LAS ASIGNATURAS	92,19	268,94	361,13	
CURSO				

Figura 33. Visualización de datos en forma de matriz de datos.

La Figura 33 presenta los datos obtenidos como resultado de generar la matriz de datos basado en el diseño realizado. La matriz de datos suministra información sobre los objetos involucrados en el diseño y los valores que toma la medida para cada combinación de estos objetos incluidos los valores totales.

## 21.4 GESTIÓN DE LAS CONSULTAS ANALÍTICAS

El prototipo de herramienta OLAP permite que los usuarios *guarden* y *abran* diseños de análisis de forma individual o en forma conjunta (esto es: varios diseños individuales que pueden ser de distintos data marts) en archivos denominados archivos de análisis (.anlx) o en áreas de trabajo (.wsx) respectivamente. Estos archivos son útiles ya que les permiten a los usuarios ahorrar tiempo al momento de hacer análisis de información al no tener que recordar como construyeron un diseño que suministre información significativa en un momento dado.

La Figura 34 muestra las opciones que se deben usar si se quieren guardar archivos de diseños de análisis individuales o en conjunto (áreas de trabajo).

The screenshot shows the OLAP tool interface. On the left, there is a dimension hierarchy tree with categories like ESTUDIANTE, GENEROS, FEMENINO, MASCULINO, RECURSO, PROFESOR, FECHA, PERIODO, CURSO, and ACTIVIDAD. The central area displays a data table with columns for FEMENINO, MASCULINO, TOTAL GENEROS, and ESTUDIANTE. The right sidebar shows a menu with options like 'Nuevo análisis', 'Abrir área de trabajo', 'Guardar área de trabajo', and 'Salir'. The 'Abrir análisis' and 'Guardar análisis' buttons are highlighted in green, and the 'Abrir área de trabajo' and 'Guardar área de trabajo' buttons are highlighted in orange.

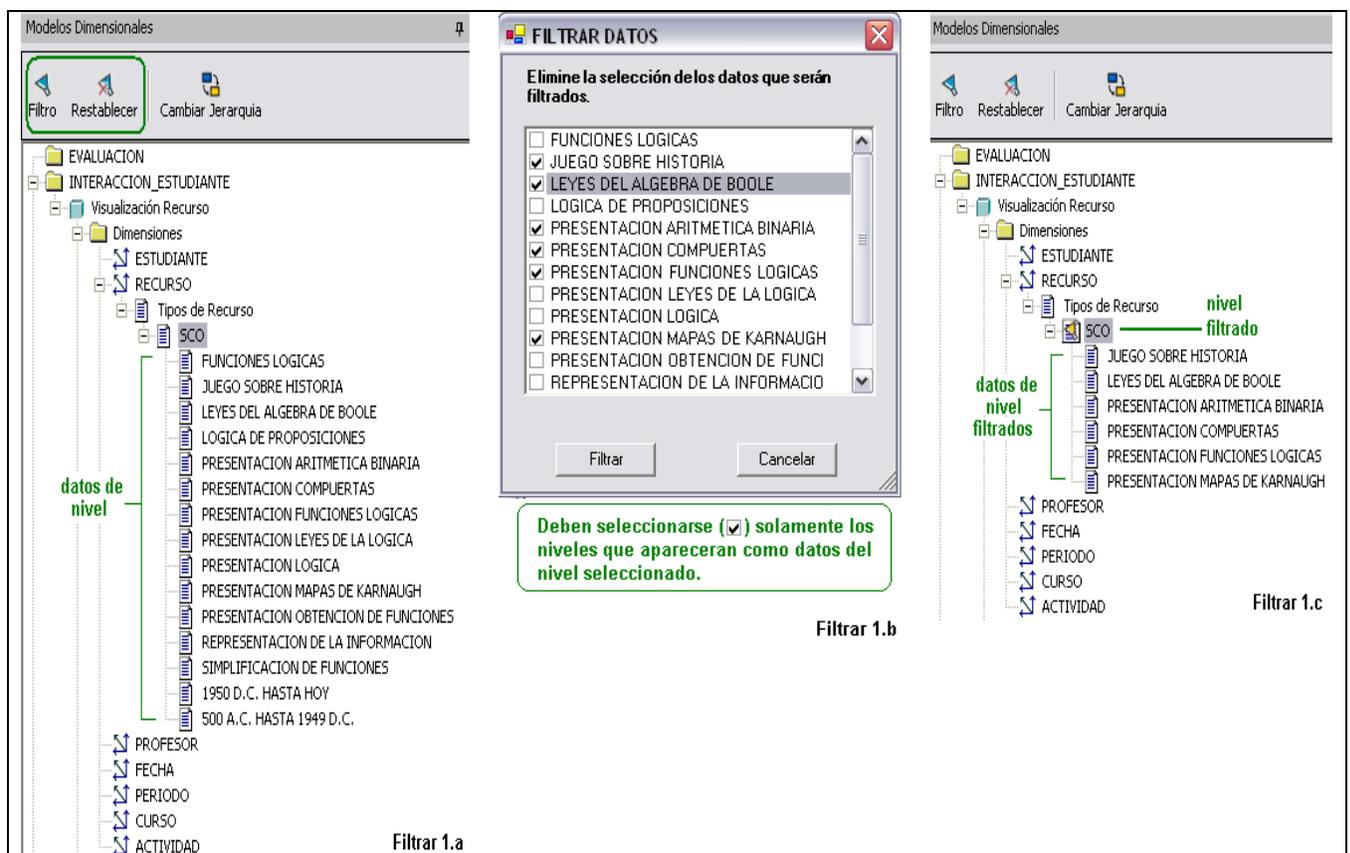
Medidas	FEMENINO	MASCULINO	TOTAL GENEROS	ESTUDIANTE
ANÁLISIS Y REFLEXIÓN	33,13	107,83	140,96	
Explicación	20,01	50,56	70,57	
Juegos	22,22	26,73	48,95	
Lectura	3,48	24,21	27,69	
Refuerzo	13,35	59,61	72,96	
TOTAL ACTIVIDADES	92,19	268,94	361,13	

Figura 34. Gestión de archivos de análisis.

## 21.5 FILTRADO DE COMPONENTES Y VALORES

La herramienta solamente permite realizar filtrado a componentes dimensionales, es decir, permite filtrar los datos de nivel de una jerarquía en una dimensión. Esta característica es útil si se presentan en un nivel de la jerarquía muchos datos de nivel y no se quieren incluir todos en un diseño de un análisis.

La Figura 35 muestra el proceso que se debe seguir para filtrar un nivel. Para este hecho, el usuario debe seleccionar un nivel específico y utilizar el botón *filtro* para filtrar los datos; en caso de que se requieran los datos de nivel en su estado original, se debe seleccionar un nivel filtrado (que se distingue por el símbolo nivel filtrado (🗑️)) y dar clic en el botón *Restablecer*.



**Figura 35.** Filtrado de datos de nivel. **Filtrar 1.a.** Presenta los datos de nivel sin filtrar. **Filtrar 1.b.** Muestra la pantalla que presenta al usuario los datos a filtrar. **Filtrar 1.c.** Presenta en nivel con los datos filtrados.



---

# CONCLUSIONES, RECOMENDACIONES Y TRABAJO FUTURO

---

- Como resultado del proyecto se logró construir una bodega de datos con los modelos dimensionales necesarios para soportar tres áreas de interés para el negocio, Estilos de Aprendizaje, Interacción del Estudiante en el ambiente virtual y Evaluación.
- Como resultado del proyecto se implementó un prototipo de herramienta OLAP que permite visualizar la información contenida dentro de la bodega de datos y permite gestionar los diseños de las consultas analíticas predefinidas y creadas por los usuarios.
- Como resultado del proyecto se desarrolló un modulo de carga de datos para la bodega en la herramienta Oracle Warehouse Builder (OWB).
- Se integro los productos obtenidos (Bodega de Datos y Prototipo de herramienta OLAP) a la solución global del macroproyecto.
- Se logró publicar dos artículos en revistas nacionales:
  - MUÑOZ, Jaime, ACOSTA, Lisandro, MENDOZA y Martha. Bodega de datos para Unicauca Virtual: una primera visión. Enlace Informático. Revista de Ciencia y Tecnología Departamento de Sistemas, Facultad de Ingeniería Electrónica y Telecomunicaciones, Universidad del Cauca, Cuarta Edición, Junio 2006. ISSN: 1692-374X. <http://enlaceinformatico.unicauca.edu.co/>.
  - MENDOZA, Martha, COBOS, Carlos, MUÑOZ, Jaime, ACOSTA, Lisandro y GOMEZ, Luis. Bodega de datos y OLAP en Unicauca Virtual. Revista Ingenierías. Facultad de Ingenierías Físico Mecánicas. Universidad Industrial de Santander, Volumen 5, No.1. En turno de publicación Julio 2006. ISSN: 1657-4583.
- La exploración tecnológica es una excelente alternativa cuando se inicia un proyecto relacionado con bodegas de datos y OLAP, y de gran utilidad cuando no se cuenta con el escenario ideal para su desarrollo. La exploración tecnológica no solo sirve para conocer a fondo la tecnología en la que se va a implementar la bodega de datos y se desarrollará las consultas OLAP, sino que también facilita el entendimiento de los conceptos teóricos generales relacionados con la tecnología.



- Guiar el modelado dimensional por medio de una metodología planteada para este tipo de modelado es muy importante, sobretodo para el diseño dimensional, sin embargo, se debe tener en cuenta que la tecnología de bodegas de datos que se va a utilizar para construirla, permita implementar los casos que se modelan en el diseño, o de otra forma redefinir el modelado de acuerdo a las limitaciones de la tecnología usada.
- La experiencia obtenida al desarrollar el proyecto, mostró que cuando se esta iniciando el desarrollo de este tipo de proyectos es importante, empezar a modelar los temas del negocio de los que se tiene mayor información. Además, una vez seleccionado los hechos o medidas, es aconsejable iniciar con un modelado básico, sin pretender de una vez vislumbrar e incluir todos los casos posibles que se obtienen de un modelado dimensional (hechos conformados, relaciones muchos a muchos, jerarquías, subdimensiones, minidimensiones, roles, etc.). Estos se pueden incluir en una próxima versión que rápidamente puede convertirse en un modelo dimensional completo.
- Las herramientas de Oracle disponibles para implementar una bodega de datos, no soportan la especificación de algunos casos especiales que surgen en el modelado dimensional. En el proyecto, el hecho de no poder representar casos como relaciones de muchos a muchos entre dimensiones y tablas de hechos y personalización de hechos y dimensiones, obligó a la creación de nuevos modelos dimensionales reduciendo la posibilidad de obtener información más significativa para los usuarios de la bodega.
- Las tecnologías de bodegas de datos y OLAP no son de uso exclusivo de sistemas relacionados con ambientes empresariales, estas tecnologías son útiles en cualquier área donde se requiera el análisis de información para mejorar la toma de decisiones. Algunos ambientes “no tan comunes” donde se pueden utilizar estas tecnologías son:
  - Ambiental: Manejo de información relacionada con procesos de crecimiento de plantas dentro de un vivero inteligente, control de especies y recursos naturales, etc.
  - Medicina: Manejo de información relacionada con el control de enfermedades en diferentes zonas del país.
  - Comportamiento Social: Manejo de información relacionada con producción y consumo de alimentos, destinos y preferencias turísticas en el país, etc.
- Un aspecto importante a tener en cuenta al momento de construir una Bodega de Datos, es contar con la base de datos que soporta el proceso operacional del negocio. En el caso de Unicauca Virtual, solamente una parte del sistema OLTP se



ha definido hasta el momento, lo cual limita la definición de nuevos temas y modelos dimensionales útiles para los usuarios.

- Oracle 10g Release 2, permite la implementación de las Bodegas de Datos de dos formas: Relacional y Multidimensional. Aunque se ha hecho un gran esfuerzo por brindar una solución fuerte que soporte la inteligencia de negocios, la etapa de transición desde la versión 9i a 10g produjo una serie de problemas:
  - ODP NET<sup>15</sup> para Oracle 10g R1 no funcionaba correctamente. Después de instalado no se podía conectar a la base de datos desde Visual Studio .NET utilizando este componente. Este error solo se corrigió hasta la realización por parte de Oracle de la versión R2 de su motor de base de datos con su correspondiente ODP NET.
  - El contenido de alguna documentación que se encuentra en el sitio oficial de Oracle sobre el tema, difiere entre versiones de la base de datos, haciendo que algunos de los ejemplos que se encuentran no puedan ser implementados a través de las herramientas. Además, esta documentación presenta el contenido de forma muy general limitando el entendimiento de algunos de los casos especiales que se presentan durante el desarrollo de una Bodega de Datos como por ejemplo el manejo de subdimensiones, relaciones de muchos a muchos entre la tabla de hechos y las dimensiones, dimensiones y tablas de hechos personalizados y tablas puente para el manejo de jerarquías.
- Como trabajo futuro, se puede decir que es necesario identificar e implementar otros modelos dimensionales que permitan conocer más sobre el comportamiento del estudiante y su experiencia dentro del ambiente virtual. Además, el prototipo de herramienta OLAP podría complementar sus funcionalidades mejorando la interacción con el usuario y la presentación de los datos. Algunas funciones que se que pueden adicionar son:
  - Presentar los atributos descriptivos adicionales de los niveles definidos para las jerarquías en la bodega de datos.
  - Supresión de los valores “NA” (valores que no aplican para análisis) obtenidos en la generación de la matriz y el grafico de datos.
  - Ordenar los valores de las columnas y realizar cálculos sobre ellos.
  - Eliminación de columnas en la grilla de datos.
  - Cambio del componente utilizado para generar gráficos de datos por otro que permita mejor definición, visualización e interacción.

---

<sup>15</sup> ODP NET es el proveedor de datos de Oracle para Visual Studio .NET.



# REFERENCIAS BIBLIOGRAFICAS

- [1]. Colombia. Ministerio de Educación. (1996 - 2005). *Plan Decenal de Educación*.
- [2]. Universidad del Cauca. (2003). Información relacionada con Unicauca Virtual y demás proyectos que lo integran. <http://uv.unicauca.edu.co/>.
- [3]. Senn, J.A. (1995). *Sistemas de Información para la Administración*. México: Grupo Editorial Iberoamérica.
- [4]. McLeod, R. (2000). *Sistemas de Información Gerencial* (7ma ed). México: Prentice Hall Hispanoamérica.
- [5]. Senn, J.A. (1992). *Análisis y Diseño de Sistemas de Información*. México: McGraw-Hill.
- [6]. Koontz, H., Weihrich, H. (1994). *Administración, una perspectiva global* (10ma ed). Toma de decisiones (pp. 206). México: McGrawHill.
- [7]. Date, C.J. (2001). *Introducción a los Sistemas de Bases de Datos* (7ma ed). Apoyo para la toma de decisiones (pp. 694-724). Prentice Hill.
- [8]. Inmon, W.H. (1996). *Building the Data Warehouse* (2da ed). Wiley Computer Publishing.
- [9]. Cohen, D. (1994). *Sistemas de Información para la Toma de Decisiones* (2da ed). México: McGrawHill.
- [10]. Kimball, R. (1998). *The Data Warehouse Toolkit: Practical techniques for building dimensional Data Warehouse*. Wiley Computer Publishing.
- [11]. Oracle Corporation. (2002). *Oracle9i Data Warehousing Guide* (Release 2 9.2). Oracle Corporation: Autor.
- [12]. Dodge, G., Gorman, T. (2000). *Essential Oracle8i Data Warehousing*. Wiley Computer Publishing.
- [13]. Kimball, R. (1998). *The Data Warehouse Lifecycle Toolkit*. Wiley Computer Publishing.
- [14]. 1keydata. (n.d.). Información relacionada con los conceptos de OLAP, ROLAP y MOLAP. <http://www.1keydata.com/datawarehousing/datawarehouse.html>.
- [15]. Oracle Corporation. (2002). *Oracle9i OLAP User's Guide* (Release 2 9.2.0.2). Oracle Corporation: Autor.
- [16]. Luján Mora, S. (2005). *Data Warehouse Design with UML*, PhD Thesis. *Department of Software and Computing Systems*. University of Alicante.
- [17]. Gornik, D. (2001). *IBM Rational Unified Process: Best Practices for Software Development Teams*.
- [18]. Oracle Corporation. (2005). *Oracle OLAP Application Developer's Guide 10g* (Release 2 10.1.0.4). Oracle Corporation: Autor.
- [19]. Oracle Corporation. (2003). *Oracle Warehouse Builder User's Guide 10g* (Release 1 10.1) (pp. 2 – 28). Oracle Corporation: Autor.



- 
- [20]. Chang, D. T. (n.d.). *Common Warehouse Metamodel specification*. <http://www.omg.org/cwm/>.
- [21]. Oracle Corporation. (2005). *Oracle OLAP Application Developer Guide 10g* (Release 1 10.1.0.4). (pp. 1 – 22). Oracle Corporation: Autor.
- [22]. Oracle Corporation. (2005). *Oracle OLAP DML Reference 10g* (Release 2 10.2). (pp. 5 – 21). Oracle Corporation: Autor.
- [23]. Eriksen, H. H. (n.d.). *Oracle OLAP Forum*. <http://forums.oracle.com/forums/thread.jspa?messageID=1085783&#1085783>
- [24]. Craig, L. (1999). *UML y Patrones Introducción al Análisis y Diseño Orientado a Objetos* (1ra Edición). México, Naucalpan de Juárez: Prentice Hall.
- [25]. Jedox. (n.d.), *PALO – Open Source MOLAP*. <http://www.jedox.com/show.php?index=content/Open%20Source%20-%20Palo.07>.
- [26]. Pentaho. (n.d), *Mondrian OLAP*. <http://mondrian.sourceforge.net>.
- [27]. Tensegrity Software. (n.d.), *PALO Eclipse Client*. <http://www.tensegrity-software.com/jpalo.html>



## ANEXOS

### ANEXO 1. CUESTIONARIOS PARA RECOLECCIÓN DE REQUERIMIENTOS EN LA CONSTRUCCIÓN DE LA BODEGA DE DATOS PARA UNICAUCA VIRTUAL

Los siguientes cuestionarios han sido adaptados de las preguntas guía que ofrece R. Kimball en su libro *“The Data Warehouse Toolkit: Practical techniques for building dimensional Data Warehouse”* [13].

#### CUESTIONARIOS PARA LAS ENTREVISTAS - UNICAUCA VIRTUAL

##### I. Cuestionario para los “Ejecutivos del Negocio”

Objetivos:

- Proveer una visión de los procesos claves de Unicauca Virtual.
- Entender las métricas de éxito de Unicauca Virtual.
- Identificar los problemas que se puedan presentar dentro de Unicauca Virtual.

1. ¿Por qué un ambiente virtual para la enseñanza?
2. ¿Cuáles son los objetivos de Unicauca Virtual?
3. ¿Cómo se pretende medir el éxito de Unicauca Virtual?
4. ¿Se ha pensado como medir el cumplimiento de los objetivos en Unicauca Virtual durante su producción?
5. ¿Cuáles son las áreas estratégicas de mayor relevancia dentro de Unicauca Virtual?  
¿Por que?
6. ¿Qué problemas podrían evitar que se cumplan los objetivos de Unicauca Virtual?
7. ¿Cuál es el impacto de estos problemas dentro de Unicauca Virtual?
8. ¿Han definido dentro de Unicauca Virtual algún conjunto de estrategias pedagógicas que se vayan a aplicar durante la producción?
9. ¿Cómo se pretende medir el éxito de estas estrategias pedagógicas aplicadas?
10. ¿Qué podría evitar que se aplique de la mejor forma una estrategia pedagógica?



11. ¿Cómo se identifica que una determinada estrategia pedagógica tiene o podría tener problemas?
12. ¿Considera que si se tiene mayor información respecto al comportamiento de los estudiantes dentro de Unicauca Virtual, se impactaría la definición de las estrategias pedagógicas aplicadas?
13. ¿Cómo considera usted que se podría mejorar el manejo de la información dentro de Unicauca Virtual?
14. ¿Qué tipos de usuarios de Unicauca Virtual interactuarán directamente con la información?

## II. Cuestionario para los “Administradores o Analistas del Negocio”

Objetivos:

- Proveer una visión de los aspectos claves de Unicauca Virtual por parte de los tutores.
- Entender las métricas de éxito de los tutores en su labor como docentes de Unicauca Virtual.
- Identificar problemas que se puedan presentar dentro de Unicauca Virtual.

1. ¿Cuáles son sus objetivos como docente dentro de Unicauca Virtual?
2. ¿Cuáles de estos objetivos son los más importantes en su labor?
3. ¿Qué hace usted para cumplir sus objetivos?
4. ¿Cuáles son las métricas que utiliza para identificar el éxito al desempeñar su labor?
5. ¿Cómo sabe usted que desempeña bien su labor?
6. ¿Con qué frecuencia usted evalúa su trabajo?
7. ¿Actualmente encuentra dificultades en su labor como docente de Unicauca Virtual?
8. ¿Qué factor(es) no le permite(n) alcanzar un mejor desempeño en su labor?
9. Realice una descripción del tipo de estudiantes a los cuales usted orienta. ¿Cómo distingue usted entre estos estudiantes? ¿Con qué frecuencia cambia esta categorización?



10. ¿Suponiendo que usted cuenta con toda la información de todos los estudiantes, cómo hace usted para encontrar información de un estudiante con características específicas?
11. ¿Qué tipo de análisis utiliza actualmente sobre las estrategias pedagógicas? ¿Qué datos utiliza para el análisis? ¿Cómo obtiene esos datos?
12. ¿Qué hace con la información del análisis una vez la obtiene?
13. ¿Cómo usted puede mejorar el análisis que hace actualmente?
14. ¿Qué tipo de análisis le gustaría realizar sobre las estrategias pedagógicas?
15. ¿Qué tipo de análisis realiza usted basado en su experiencia? ¿Qué personas solicitan este tipo de análisis? ¿Qué hacen estas personas con este tipo de análisis?
16. ¿Se presentan “cuellos de botella” en el momento de obtener información que usted necesita?
17. ¿Qué tanta información histórica requiere para hacer sus análisis?
18. ¿Considera que si se tiene mayor información respecto al comportamiento de los estudiantes dentro de Unicauca Virtual, se impactaría en la definición de las estrategias pedagógicas aplicadas?
19. ¿La información con la que cuenta usted actualmente la obtiene en algún formato o reporte específico? ¿Qué datos son los más importantes en estos reportes? ¿Cómo usa la información contenida en estos reportes?
20. ¿Qué diferencia encontraría, si los reportes se generan de una manera dinámica?

### **III. Cuestionario para las “Personas del SI (Sistema de Información)”**

#### **Objetivos:**

- Verificar si los datos existen en el SI para soportar los “temas” que se van a trabajar en el proyecto.
  - Obtener una visión global de los sistemas operacionales.
  - Entender mejor las definiciones de los campos, la granularidad, volumen de los datos, datos demográficos etc.
1. Por favor describa de manera general la información almacenada en las fuentes de datos operacionales de Unicauca Virtual.



2. ¿Cuál es el proceso actual usado para entregar la información a los usuarios de los datos dentro de Unicauca Virtual?
3. ¿Qué herramientas se usan para acceder y analizar la información?
4. ¿Realizan algún tipo de análisis sobre la información de manera rutinaria?
5. ¿Qué tipos de informes son requeridos por los usuarios del sistema?
6. ¿Cuántos y cuáles sistemas de información soportan el manejo de los datos en Unicauca Virtual?
7. ¿Existen problemas en la calidad de los datos?
8. ¿Cuál es el mayor “cuello de botella” o problema en el proceso actual de acceso a datos?
9. ¿Qué piensa usted que limitaría el éxito de Unicauca Virtual una vez este en producción?
10. ¿Qué quiere que el DW logre dentro de Unicauca Virtual?

#### **IV. Cuestionario general enfocado hacia “Tutores de un Ambiente de Aprendizaje Virtual”**

Objetivos:

- Proveer una visión de los aspectos claves de un ambiente de aprendizaje virtual por parte de los tutores.
  - Entender las métricas de éxito de los tutores en su labor como docentes del ambiente de aprendizaje virtual.
  - Identificar problemas que se puedan presentar dentro del ambiente de aprendizaje virtual.
1. ¿Cuáles son sus objetivos como docente dentro de un ambiente de aprendizaje virtual?
  2. ¿Cuáles de estos objetivos son los más importantes en su labor?
  3. ¿Qué hace usted para cumplir sus objetivos?
  4. ¿Cuáles son las métricas que utiliza para identificar el éxito al desempeñar su labor?
  5. ¿Cómo sabe usted que desempeña bien su labor?



6. ¿Con qué frecuencia usted evalúa su trabajo?
7. ¿Actualmente encuentra dificultades en su labor como docente de un ambiente de aprendizaje virtual?
8. ¿Qué factor(es) no le permite(n) alcanzar un mejor desempeño en su labor?
9. Realice una descripción del tipo de estudiantes a los cuales usted orienta. ¿Cómo distingue usted entre estos estudiantes? ¿Con que frecuencia cambia esta categorización?
10. ¿Suponiendo que usted cuenta con toda la información de todos los estudiantes, cómo hace usted para encontrar información de un estudiante con características específicas?
11. ¿Qué tipo de análisis realiza en su labor como docente de un ambiente de aprendizaje virtual? ¿Qué datos utiliza para el análisis? ¿Cómo obtiene esos datos?
12. ¿Qué hace con la información del análisis una vez la obtiene?
13. ¿Cómo usted puede mejorar el análisis que hace actualmente?
14. ¿Qué tipo de análisis le gustaría realizar relacionado con la labor que desempeña?
15. ¿Qué tipo de análisis realiza usted basado en su experiencia? ¿Qué personas solicitan este tipo de análisis? ¿Qué hacen estas personas con este tipo de análisis?
16. ¿Se presentan “cuellos de botella” en el momento de obtener información que usted necesita?
17. ¿Qué tanta información histórica requiere para hacer sus análisis?
18. ¿La información con la que cuenta usted actualmente la obtiene en algún formato o reporte específico? ¿Qué datos son los más importantes en estos reportes? ¿Cómo usa la información contenida en estos reportes?
19. ¿Qué diferencia encontraría, si los reportes se generan de una manera dinámica?



## **ANEXO 2. RESULTADO DE LA ENCUESTA APLICADA AL EJECUTIVO DEL NEGOCIO**

El archivo de audio que contiene la entrevista realizada al ejecutivo del negocio se encuentra en el **CD-ROM** que se anexa a este documento.

## ANEXO 3. MODELADO FÍSICO DE LOS MODELOS DIMENSIONALES

- Tabla de hechos “*Cambio en el estilo de aprendizaje del estudiante*”

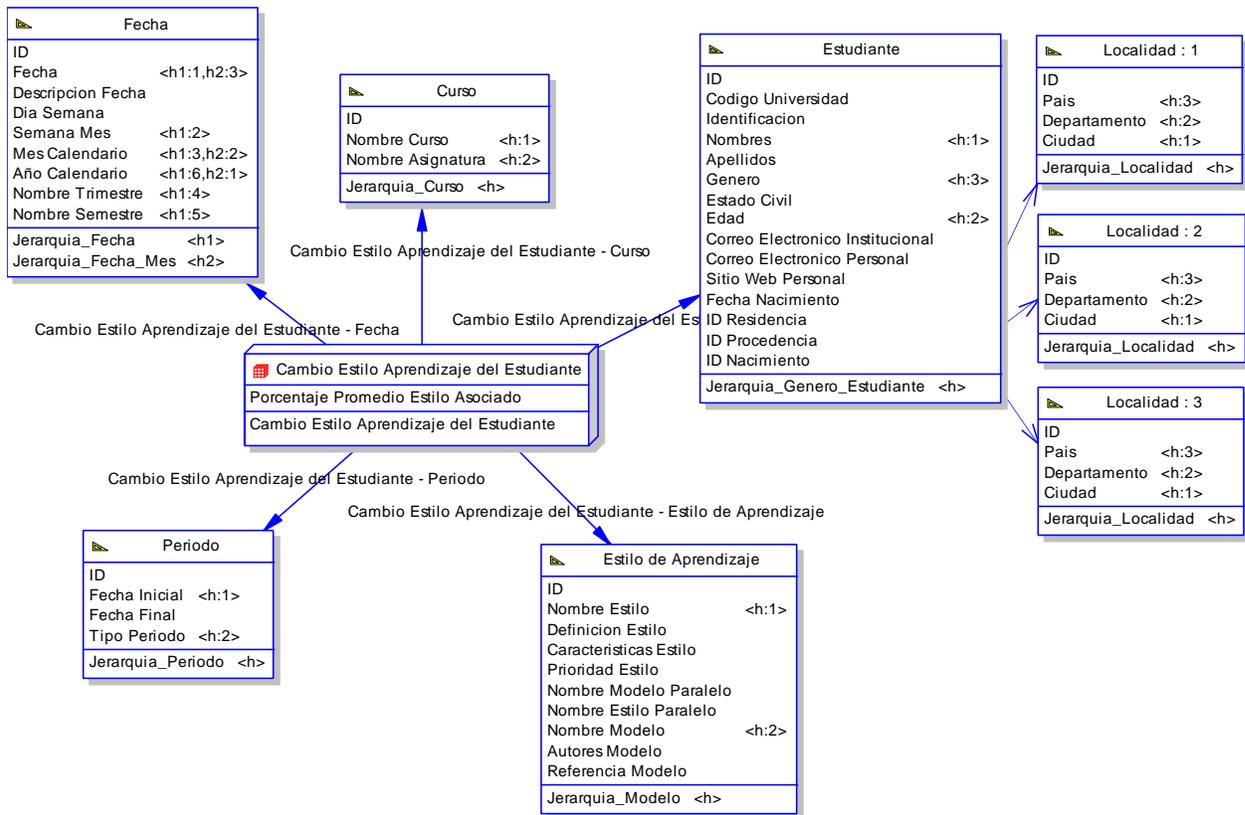
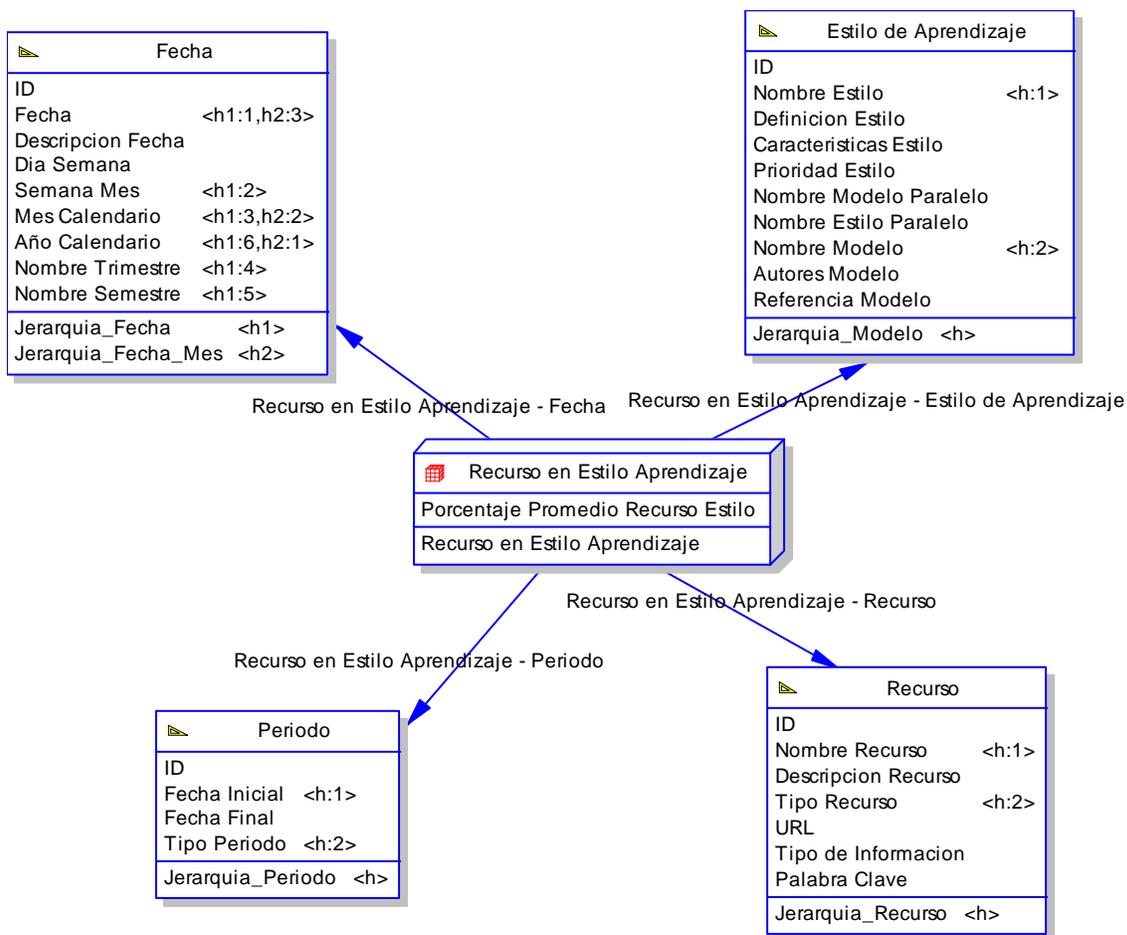


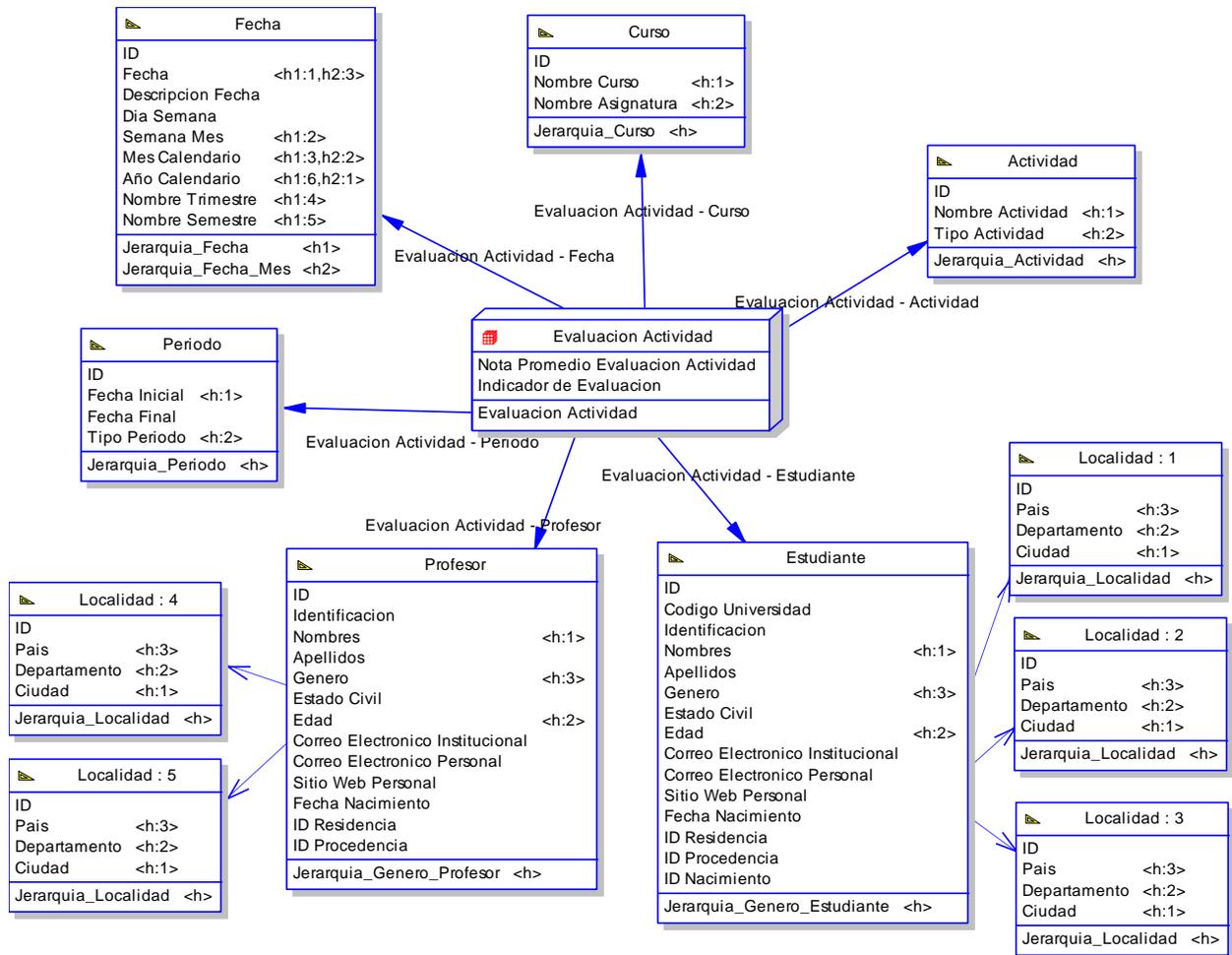
Figura 36. Modelo físico para la tabla de hechos “*Cambio en el estilo de aprendizaje del estudiante*”.

• **Tabla de hechos “Recurso en estilo de aprendizaje”**



**Figura 37.** Modelo físico para la tabla de hechos “Recurso en estilo de aprendizaje”.

• **Tabla de hechos “Evaluación de actividad”**



**Figura 38.** Modelo físico para la tabla de hechos “Evaluación de actividad”.

## ANEXO 4. DESCRIPCIÓN DETALLADA DE LAS DIMENSIONES CONFORMADAS PARA LAS TABLAS DE HECHOS

### DIMENSIÓN “PROFESOR”

- Diagrama de la dimensión

La Figura 39, muestra los detalles de la dimensión profesor.

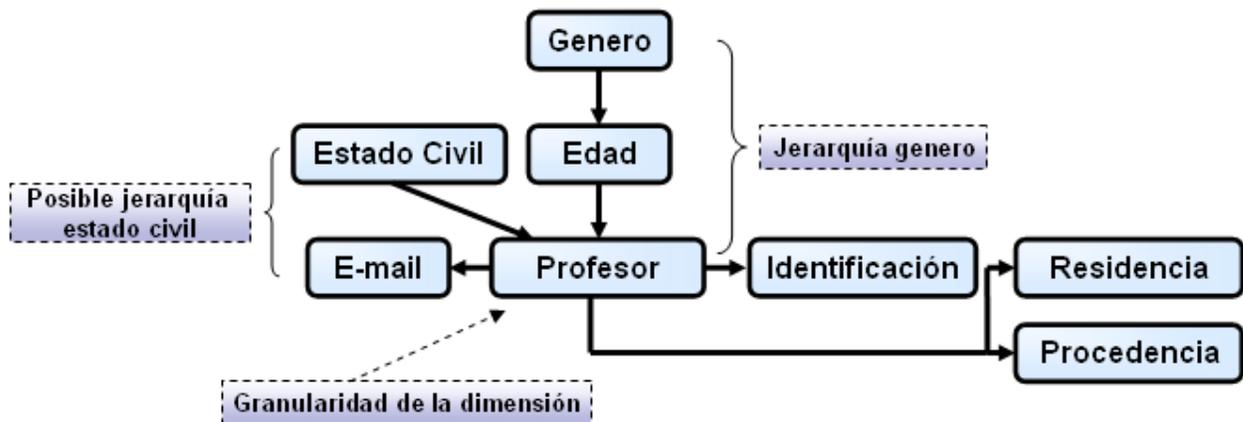


Figura 39. Diagrama de detalle para la dimensión “Profesor”.

- Descripción de la dimensión

La Tabla 15 describe detalladamente el diagrama.

Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...
Identificación	Representa el número que identifica a un profesor como ciudadano.	* ... *	Se sobrescribe (Tipo 1)	34.496.111, 22.236.412.
Nombres	Representa los nombres del profesor.	* ... *	No se modifica (Tipo 0)	Martha Eliana, Carlos Alberto.
Apellidos	Representa los apellidos del profesor.	* ... *	No se modifica (Tipo 0)	Mendoza Becerra, Cobos Lozada.
Genero	Genero del profesor.	2	No se modifica (Tipo 0)	Masculino, Femenino.
Estado civil	Representa el estado civil del profesor.	4	Se sobrescribe (Tipo 1)	Soltero(a), Casado(a), Viudo(a), Divorciado(a).



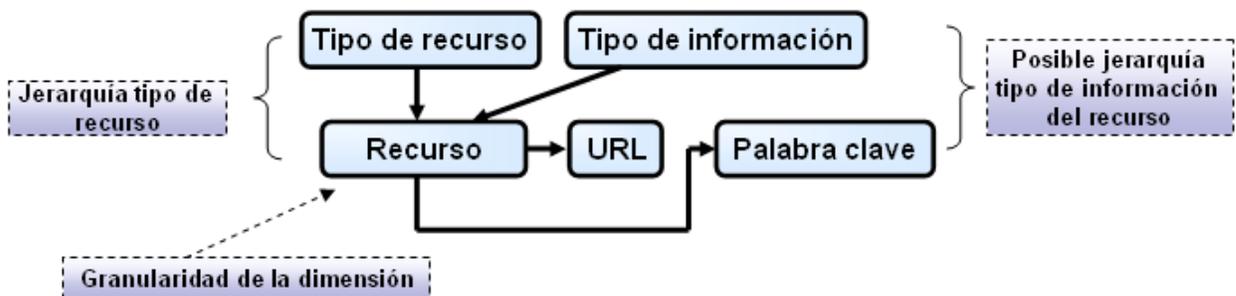
Edad	Edad del profesor.	35	Se sobrescribe (Tipo 1)	26 años, 28 años, 31 años.
Correo electrónico institucional	Representa el e-mail institucional del profesor.	* ... *	Se sobrescribe (Tipo 1)	mmendoza@unicauca.edu.co, ccobos@unicauca.edu.co.
Correo electrónico personal	Representa el e-mail personal del profesor.	* ... *	Se sobrescribe (Tipo 1)	mendoza_martha_eliana@hotmail.com, ccobos@hotmail.com.
Sitio Web personal	Representa el sitio Web personal de un profesor.	* ... *	Se sobrescribe (Tipo 1)	www.unicauca.edu.co/~mmendoza, www.unicauca.edu.co/~ccobos.
Fecha de nacimiento	Representa la fecha de nacimiento del profesor.	* ... *	No se modifica (Tipo 0)	10/09/1978, 22/10/1976.
ID Residencia	Representa la clave para identificar la subdimensión de localidad. Rol: Residencia.	* ... *	Se sobrescribe (Tipo 1)	1, 2, 3, 4...
ID Procedencia	Representa la clave para identificar la subdimensión de localidad. Rol: Procedencia.	* ... *	Se sobrescribe (Tipo 1)	1, 2, 3, 4...

**Tabla 15.** Descripción detallada de la dimensión “*Profesor*”.

## DIMENSIÓN “*RECURSO*”

- **Diagrama de la dimensión**

La Figura 40, muestra los detalles de la dimensión recurso.



**Figura 40.** Diagrama de detalle para la dimensión “*Recurso*”.

- **Descripción de la dimensión**

La Tabla 16 describe detalladamente el diagrama.

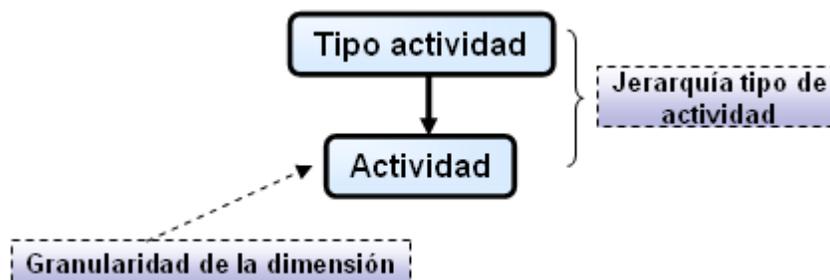
Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...
Nombre recurso	Nombre dado al recurso.	* ... *	No se modifica (Tipo 0)	Leyes del álgebra de BOOLE, funciones lógicas.
Descripción recurso	Describe detalladamente el recurso creado.	* ... *	No se modifica (Tipo 0)	Este recurso SCO permite comprender las LEYES DE BOOLE, además...
Tipo recurso	Clasificación del recurso SCO ó ASSET.	2	No se modifica (Tipo 0)	SCO, ASSET.
URL	Ubicación del recurso.			http://uv.unicauca.edu.co/Recursos/Fase1/.../leyes_de_l_algebra_de_boole1.htm.
Tipo de información	Indica el tipo de información que presenta el recurso.	10	No se modifica (Tipo 0)	Texto, video, diagrama.
Palabra clave	Palabra clave que representa el recurso.	* ... *	Se sobrescribe (Tipo 1)	BOOLE, Funciones Lógicas.

**Tabla 16.** Descripción detallada de la dimensión “Recurso”.

## DIMENSIÓN “ACTIVIDAD”

- Diagrama de la dimensión

La Figura 41, muestra los detalles de la dimensión actividad.



**Figura 41.** Diagrama de detalle para la dimensión “Actividad”.

- **Descripción de la dimensión**

La Tabla 17 describe detalladamente el diagrama.

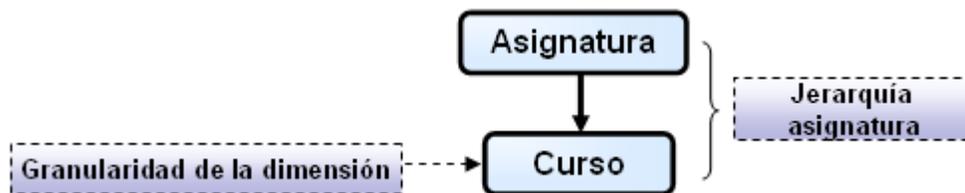
Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...
Nombre actividad	Nombre dado a la actividad creada.	* ... *	No se modifica (Tipo 0)	500 A.C. hasta 1949 D.C, Funciones lógicas.
Tipo actividad	Indica el tipo de actividad que representa.	20	No se modifica (Tipo 0)	Lectura, juegos, explicación.

**Tabla 17.** Descripción detallada de la dimensión “Actividad”.

### DIMENSIÓN “CURSO”

- **Diagrama de la dimensión**

La Figura 42, muestra los detalles de la dimensión curso.



**Figura 42.** Diagrama de detalle para la dimensión “Curso”.

- **Descripción de la dimensión**

La Tabla 18 describe detalladamente el diagrama.

Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...
Nombre curso	Nombre dado para el curso de la asignatura.	5	No se modifica (Tipo 0)	Curso A, Curso B, Curso C, Curso D.
Nombre asignatura	Nombre dado para la asignatura.	* ... *	No se modifica (Tipo 0)	Informática, Programación Orientada a Objetos.

**Tabla 18.** Descripción detallada de la dimensión “Curso”.

## DIMENSIÓN “FECHA”

- Diagrama de la dimensión

La Figura 43, muestra los detalles de la dimensión fecha.

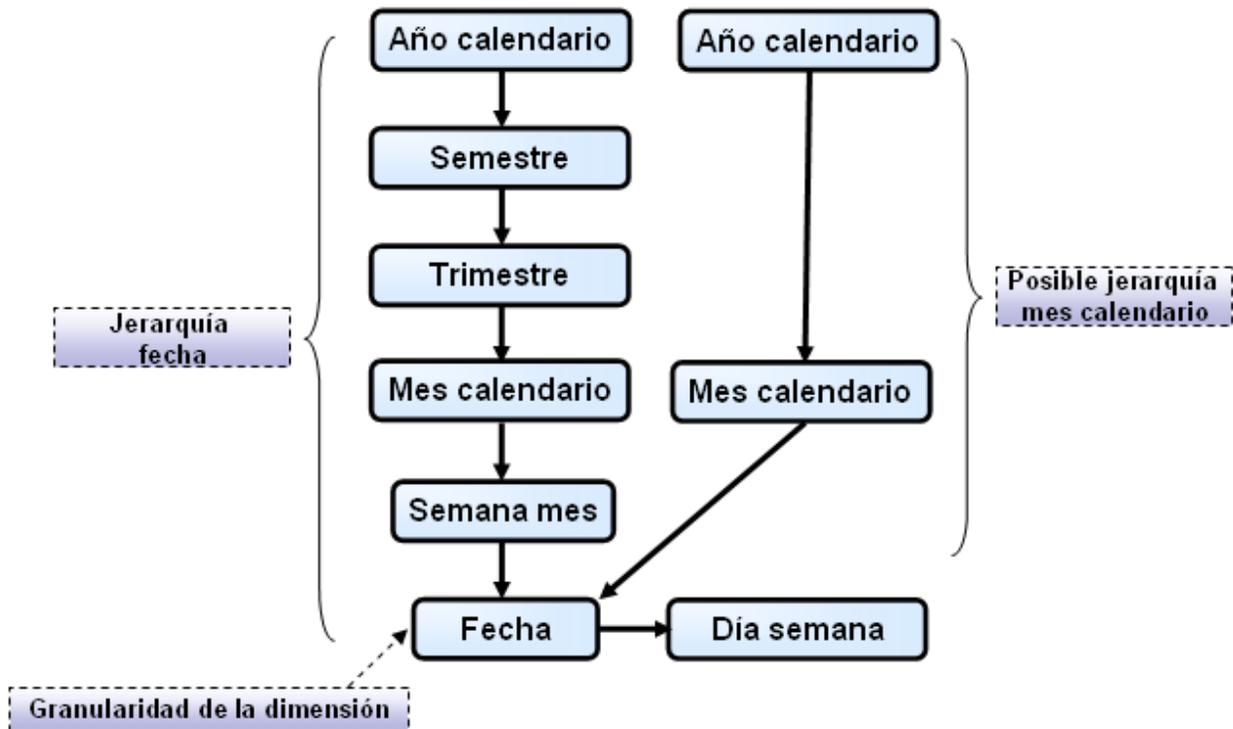


Figura 43. Diagrama de detalle para la dimensión “Fecha”.

- Descripción de la dimensión

La Tabla 19 describe detalladamente el diagrama.

Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...
Fecha	Representa una fecha específica en la que ocurre un hecho.	366	No se modifica (Tipo 0)	07/05/2005, 10/01/2006.
Descripción fecha	Descripción completa de la fecha.	366	No se modifica (Tipo 0)	07 Mayo de 2005, 10 Enero de 2006.



Día semana	Nombre del día de la semana.	7	No se modifica (Tipo 0)	Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo.
Semana mes	Representa la semana en el mes.	4	No se modifica (Tipo 0)	1ra Semana, 2da Semana, 3ra Semana, 4ta Semana.
Mes calendario	Representa el mes del calendario.	12	No se modifica (Tipo 0)	Enero, Febrero, Marzo.
Nombre trimestre	Representa el trimestre de un año específico.	4	No se modifica (Tipo 0)	1er Trimestre, 2do Trimestre, 3er Trimestre, 4to Trimestre.
Nombre semestre	Representa el semestre del calendario para un año específico.	2	No se modifica (Tipo 0)	1er Semestre, 2do Semestre.
Año calendario	Representa el año del calendario.	1	No se modifica (Tipo 0)	2005, 2006.

Tabla 19. Descripción detallada de la dimensión “Fecha”.

## DIMENSIÓN “PERIODO”

- Diagrama de la dimensión

La Figura 44, muestra los detalles de la dimensión periodo.

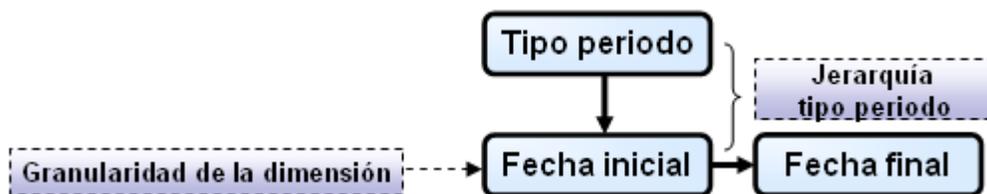


Figura 44. Diagrama de detalle para la dimensión “Periodo”.

- Descripción de la dimensión

La Tabla 20 describe detalladamente el diagrama.

Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...

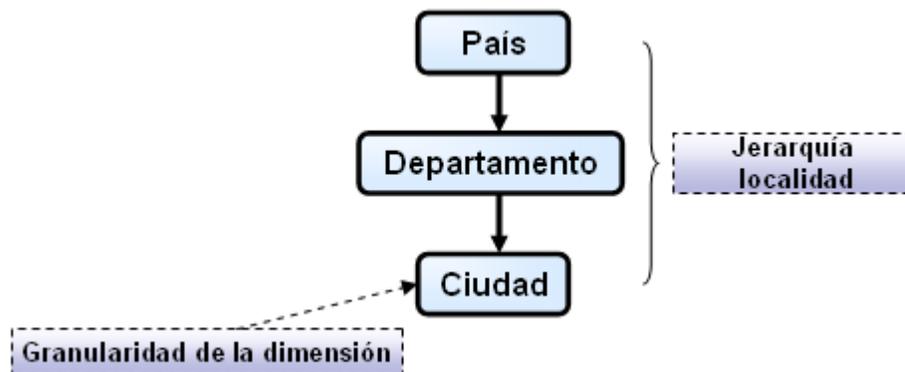
Fecha inicial	Representa la fecha inicial del periodo académico.	4	No se modifica (Tipo 0)	01/03/2005, 01/07/2005, 01/05/2005.
Fecha final	Representa la fecha final del periodo académico.	4	No se modifica (Tipo 0)	28/06/2005, 28/11/2005, 28/06/2005.
Tipo periodo	Representa el tipo del periodo académico.	2	No se modifica (Tipo 0)	Semestral, Trimestral.

**Tabla 20.** Descripción detallada de la dimensión “Periodo”.

## SUBDIMENSIÓN “LOCALIDAD”

- Diagrama de la subdimensión

Esta subdimensión juega varios roles para el estudiante y profesor. Estos roles son: *localidad de residencia*, *localidad de nacimiento* y *localidad de procedencia*. Ver la Figura 45, que indica los detalles de la dimensión.



**Figura 45.** Diagrama de detalle para la dimensión “Localidad”.

- Descripción de la subdimensión

La describe detalladamente el diagrama.

Nombre atributo	Descripción atributo	Cardinalidad	Política de cambio lento	Valores de ejemplo
ID	Representa la clave sustituta para la dimensión.			1, 2, 3, 4,...



Ciudad	Representa la ciudad en la dimensión de localidad. De acuerdo al rol que juegue esta subdimensión este atributo puede ser ciudad de residencia, procedencia o nacimiento.	* ... *	No se modifica (Tipo 0)	Popayán, Cali, Bolívar, Santander, Pasto.
Departamento	Representa el departamento en la dimensión localidad. De acuerdo al rol que juegue esta subdimensión este atributo puede ser departamento de residencia, procedencia o nacimiento.	32	No se modifica (Tipo 0)	Cauca, Valle del Cauca, Nariño.
País	Indica el país en la dimensión localidad. De acuerdo al rol que juegue esta subdimensión este atributo puede ser el país de residencia, procedencia o nacimiento.	1	No se modifica (Tipo 0)	Colombia.

**Tabla 21.** Descripción detallada de la subdimensión “Localidad”.



## **ANEXO 5. SCRIPTS DE CREACIÓN DE LAS TABLAS RELACIONALES FÍSICAS COMO SOPORTE DE LOS ESQUEMAS DIMENSIONALES**

Los scripts de creación de las tablas relacionales se encuentran anexos en el **CD-ROM**.



## ANEXO 6. DIMENSIONES EN ORACLE 10G

### ● DIMENSIONES

Una dimensión es una estructura que permite organizar o clasificar datos por categorías, además permite a usuarios responder preguntas del negocio.

Las dimensiones más comunes son: clientes, productos tiempo entre otras.

En Oracle 9i la información dimensional se almacena en la tabla dimensión y el objeto de la base de datos “dimensión” ayuda a organizar y agrupar información dentro de jerarquías. Estas jerarquías representan relaciones naturales de **1:n** entre columnas y grupos de columnas (los niveles de una jerarquía) que no pueden ser representados con condiciones de restricciones. Ir hacia arriba sobre los niveles de datos se llama “rolling up” e ir hacia abajo a través de los niveles se llama “drilling down”.

### Ejemplo:

- En la dimensión tiempo, desde los meses se sube a los cuartos, de los cuartos se sube a los años y de los años a todos los años.
- Dentro de la dimensión productos, desde los productos se sube a subcategoría, desde subcategoría a categoría y desde categoría a todos los productos.
- Dentro de la dimensión cliente, desde cliente se sube a ciudad, desde ciudad se sube a estado, de estado a país y de país a subregiones o regiones.

El análisis de datos empieza en el nivel superior de la jerarquía y gradualmente desciende si la situación necesita tal análisis.

### ● CREANDO DIMENSIONES

Antes de crear un objeto dimensión, la tabla física dimensión debe existir en la base de datos, la cual debe contener datos de la dimensión. Por ejemplo, si se quiere crear la dimensión “Cliente”, una o mas tablas deben existir en la base de datos que contengan información de la ciudad, estado y el país.

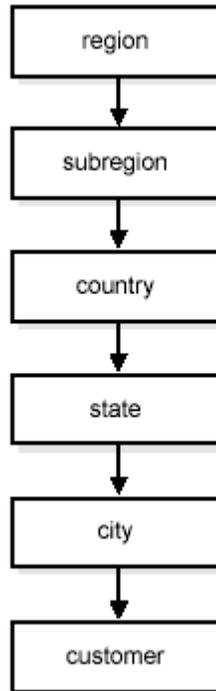
Una típica implementación relacional para una bodega de datos es un esquema estrella. La información de hechos es almacenada en una tabla llamada tabla de hechos, mientras la información dimensional es almacenada las tablas llamadas dimensión.

Use la sentencia CREATE DIMENSION para crear la dimensión y utilice la sentencia LEVEL para identificar los nombres de los niveles de la dimensión.



En las jerarquías se puede notar que para cualquier HIJO hay uno y solo un PADRE.

### Ejemplo:



### Ejemplo de creación de una dimensión:

Se declara la dimensión “products\_dim” que contiene los niveles producto, subcategoría y categoría.

```
CREATE DIMENSION products_dim
  LEVEL product          IS (products.prod_id)
  LEVEL subcategory      IS (products.prod_subcategory)
  LEVEL category         IS (products.prod_category) ...
```

Cada nivel en la dimensión DEBE corresponder a uno o más columnas en la tabla que esta en la base de datos. *En el ejemplo la tabla esta desnormalizada y existen todas las columnas en la misma tabla. Sin embargo este no es un prerrequisito para crear una dimensión (se puede crear usando tablas normalizadas).*

El próximo paso es declarar la relación de los niveles con la jerarquía. La relación jerárquica es una dependencia funcional de un nivel de la jerarquía al próximo nivel. Usando los nombres de los niveles definidos anteriormente la relación CHILD OF denota que cada valor del nivel de hijo esta asociado con uno y solo un nivel de padre.

### Ejemplo:



```
HIERARCHY prod_rollup
(product          CHILD OF
subcategory      CHILD OF
category)
```

En adición a las relaciones jerárquicas **1:n**, las dimensiones también incluyen relaciones de atributos **1:1** entre los niveles de jerarquía y sus dependencias.

### Ejemplo:

```
LEVEL fis_month  IS TIMES.FISCAL_MONTH_DESC
...
ATTRIBUTE fis_month DETERMINES
(fiscal_month_name, days_in_fiscal_month)
```

La cláusula anterior relaciona “fis\_month” a fiscal\_month\_name y days\_in\_fiscal\_month.

Un ejemplo de dimensión es el siguiente:

```
CREATE DIMENSION products_dim
LEVEL product          IS (products.prod_id)
LEVEL subcategory      IS (products.prod_subcategory)
LEVEL category         IS (products.prod_category)
HIERARCHY prod_rollup (
    product            CHILD OF
    subcategory        CHILD OF
    category)
ATTRIBUTE product DETERMINES
(products.prod_name, products.prod_desc,
prod_weight_class, prod_unit_of_measure,
prod_pack_size, prod_status, prod_list_price, prod_min_price)
ATTRIBUTE subcategory DETERMINES
(prod_subcategory, prod_subcat_desc)
ATTRIBUTE category DETERMINES
(prod_category, prod_cat_desc);
```

Una vez creada la dimensión se debe chequear que cumpla lo siguiente:

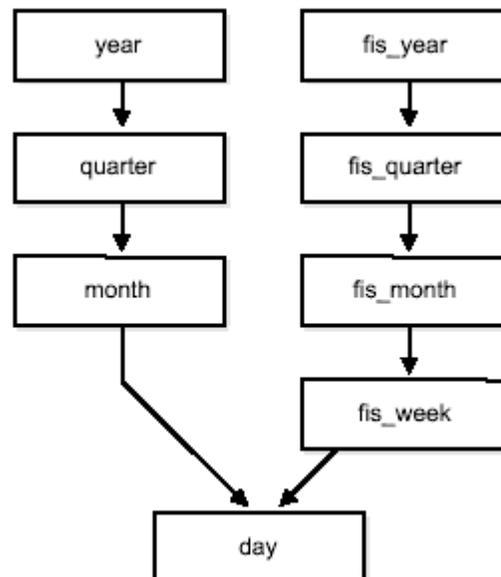
- Debe haber una relación **1:n** entre un padre y un hijo. Un padre puede tener uno o más hijos, pero un hijo debe tener un solo padre.
- Debe haber una relación entre atributos **1:1** entre los niveles de la jerarquía y sus atributos de dimensión que dependen. Por ejemplo, si hay una columna “fiscal\_month\_desc”, entonces una relación posible de atributos podría ser “fiscal\_month\_desc” o “fiscal\_month\_name”.
- Si las columnas de un nivel padre y un nivel hijo están en diferentes relaciones, entonces la conexión entre ellos requiere una relación **1:n**. Cada fila de la tabla hijo debe unirse con una y solo una fila de la tabla padre.



- Hay que asegurarse (por medio de restricciones si es necesario) que las columnas de cada nivel de jerarquía sean no nulas y que la integridad jerárquica se mantenga.
- Las jerarquías de una dimensión pueden traslaparse o desconectarse de cualquier otra. Sin embargo, las columnas de un nivel de jerarquía no pueden asociarse con más de una dimensión.
- Relaciones unidas que forman ciclos en los gráficos de dimensión no son soportados. Por ejemplo, un nivel de jerarquía no puede ser unida a si misma directamente o indirectamente.

Se puede evaluar cualquier dimensión con el procedimiento llamado "DBMS\_MVIEW.VALIDATE\_DIMENSION".

### • MÚLTIPLES JERARQUÍAS



```
CREATE DIMENSION times_dim
  LEVEL day IS TIMES.TIME_ID
  LEVEL month IS TIMES.CALENDAR_MONTH_DESC
  LEVEL quarter IS TIMES.CALENDAR_QUARTER_DESC
  LEVEL year IS TIMES.CALENDAR_YEAR
  LEVEL fis_week IS TIMES.WEEK_ENDING_DAY
  LEVEL fis_month IS TIMES.FISCAL_MONTH_DESC
  LEVEL fis_quarter IS TIMES.FISCAL_QUARTER_DESC
  LEVEL fis_year IS TIMES.FISCAL_YEAR
  HIERARCHY cal_rollup (
    day CHILD OF
    month CHILD OF
    quarter CHILD OF
    year
  )
)
```



```
HIERARCHY fis_rollup (
    day          CHILD OF
    fis_week     CHILD OF
    fis_month    CHILD OF
    fis_quarter  CHILD OF
    fis_year
) <attribute determination clauses>...
```



## **ANEXO 7. SCRIPTS DE CREACIÓN DE LOS OBJETOS DIMENSIONALES**

Los scripts de creación de los objetos dimensionales se encuentran anexos en el **CD-ROM**.



## **ANEXO 8. SCRIPTS PARA LA CARGA DE DATOS EN LA BODEGA DE UNICAUCA VIRTUAL**

Los scripts para la carga de datos de la bodega se encuentran anexos en el **CD-ROM**.



## ANEXO 9. CASOS DE USO PROTOTIPO HERRAMIENTA OLAP

### CASO DE USO DISEÑAR CONSULTA ANALÍTICA

<b>Nombre del Caso de Uso:</b>	Diseñar consulta analítica
<b>Actores</b>	
• Tutor o directivo.	
<b>Propósito:</b>	Permitir al usuario diseñar la consulta analítica al interactuar con los objetos registrados en el catalogo OLAP.
<b>Resumen:</b>	El usuario desea obtener información de la bodega de datos y para esto debe hacer una consulta sobre ella, que le presente la información requerida. El usuario selecciona un objeto de un cubo en el catalogo (dimensión, nivel, medida) y la arrastra hacia el área de diseño. El sistema despliega la información correspondiente al objeto arrastrado.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.8, R.9, R.9.1, R.9.2.
<b>Curso normal de los eventos</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
1. El usuario selecciona un objeto de un cubo del catalogo.	
2. Ubica el objeto seleccionado en el área de diseño de la consulta analítica	3. Presenta el diseño de la consulta analítica que se lleva hasta ese momento.
<b>Curso alterno 1:</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
2. El área de diseño de la consulta analítica esta dividida en dos partes. Una parte para las medidas y otras para las dimensiones. El usuario coloca una medida en la parte destinada para las dimensiones o viceversa	3. Se presenta al usuario un mensaje de error indicando el hecho.
<b>Trazabilidad</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>



<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>
---------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 22. Análisis del caso de uso diseñar consulta analítica.

## CASO DE USO GENERAR MATRIZ DE DATOS

<b>Nombre del Caso de Uso:</b>		Generar matriz de datos
<b>Actores</b>		
<ul style="list-style-type: none"> <li>• Tutor o directivo.</li> </ul>		
<b>Propósito:</b>	Permitir al usuario visualizar los datos de forma matricial, basado en el diseño de una consulta analítica.	
<b>Resumen:</b>	El usuario después de realizar el diseño de una consulta analítica, decide ver de forma matricial los datos que se obtienen como resultado de realizar la consulta basado en el diseño. El sistema ejecuta la consulta, obtiene los datos de la bodega de datos y los presenta al usuario de forma matricial.	
<b>Tipo:</b>	Primario.	
<b>Referencias Cruzadas:</b>	R.10, R.11, R.12, R.13.	
<b>Curso normal de los eventos</b>		
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>	
1. El usuario decide ver de forma matricial, los datos que se pueden obtener basado en un diseño de una consulta analítica.	2. Construye la consulta basado en el diseño.	
	3. Ejecuta la consulta en la bodega de datos.	
	4. Convierte los datos obtenidos en una representación matricial.	
.	5. Presenta los datos en forma matricial basado en diseño de la consulta analítica creado por el usuario.	
<b>Curso alterno 1:</b>		



Acción del Actor		Respuesta del Sistema
		2. si no se ha definido ningún diseño de consulta analítica no e puede construir la consulta ni continuar con las demás operaciones. Se indica al usuario al respecto con un mensaje de información.
Trazabilidad		
Artefactos Anteriores	Artefactos Posteriores	
<b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>[Diagrama de Casos de Uso]</li> </ul>	<ul style="list-style-type: none"> <li><b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>[Diagrama de Secuencia del Sistema]</li> <li>[Modelo Conceptual]</li> </ul> </li> <li><b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>[Caso de Uso Real]</li> </ul> </li> </ul>	

Tabla 23. Análisis del caso de uso generar matriz de datos.

## CASO DE USO ENTRAR AL SISTEMA

<b>Nombre del Caso de Uso:</b>	Entrar al sistema
Actores	
<ul style="list-style-type: none"> <li>Tutor o directivo.</li> </ul>	
<b>Propósito:</b>	Permitir al usuario ingresar al sistema.
<b>Resumen:</b>	Este caso de uso comienza cuando el sistema verifica que el usuario esta registrado en el sistema. Para tal fin el usuario se registra con un nombre de usuario y una clave o contraseña y el sistema verifica que el usuario se haya registrado previamente.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.1, R.2.
Curso normal de los eventos	
Acción del Actor	Respuesta del Sistema
1. El usuario ingresa su nombre de usuario y clave.	2. El sistema recibe los datos y comprueba la validez de los mismos.

		3. Si el usuario esta registrado, el sistema crea una sesión y visualiza la interfaz principal de la aplicación de acuerdo a su rol de usuario.
<b>Curso alternativo 1:</b>		
<b>Acción del Actor</b>		<b>Respuesta del Sistema</b>
		3. Si el usuario no se encuentra registrado, no se le permite el ingreso al sistema y se muestra un mensaje de error al usuario.
<b>Trazabilidad</b>		
<b>Artefactos Anteriores</b>		<b>Artefactos Posteriores</b>
<u>Artefactos del Análisis:</u> <ul style="list-style-type: none"> <li>[Diagrama de Casos de Uso]</li> </ul>		<ul style="list-style-type: none"> <li><u>Artefactos del Análisis:</u> <ul style="list-style-type: none"> <li>[Diagrama de Secuencia del Sistema]</li> <li>[Modelo Conceptual]</li> </ul> </li> <li><u>Artefactos del Diseño:</u> <ul style="list-style-type: none"> <li>[Caso de Uso Real]</li> </ul> </li> </ul>

Tabla 24. Análisis del caso de uso entrar al sistema.

## CASO DE USO GESTIONAR CONSULTA ANALÍTICA

Este Caso de uso permite al usuario abrir y guardar diseños de consultas analíticas realizadas durante la sesión. Para facilitar su descripción se divide en dos casos de uso más pequeños *Guardar diseño consulta analítica* y *Abrir diseño consulta analítica*.

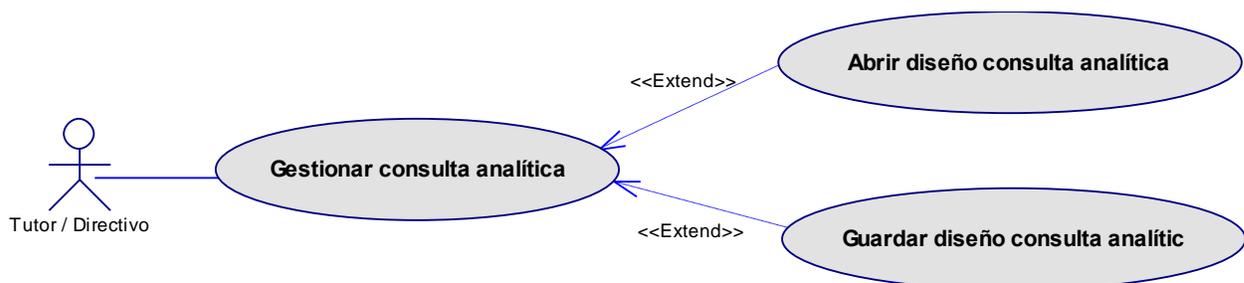


Figura 46. Extensión del caso de uso Gestionar consulta analítica.

- Caso de uso Guardar diseño consulta analítica**

<b>Nombre del Caso de Uso:</b>	Guardar diseño consulta analítica
--------------------------------	-----------------------------------



<b>Actores</b>	
<ul style="list-style-type: none"> <li>Tutor o directivo.</li> </ul>	
<b>Propósito:</b>	Permitir al usuario guardar un archivo que representa el diseño de una consulta analítica.
<b>Resumen:</b>	El usuario desea guardar un diseño de una consulta analítica, selecciona el lugar donde desea guardarla y luego el sistema crea un archivo que representa el diseño guardado por el usuario.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.16, R.16.1, R.16.2, R.16.3
<b>Curso normal de los eventos</b>	
Acción del Actor	Respuesta del Sistema
1. El usuario decide guardar un diseño de consulta analítica.	2. presenta una interfaz que le permite seleccionar la ubicación donde guardar diseño.
3. Selecciona la ubicación donde guardar el archivo de diseño	4. Verifica que el archivo de diseño no exista en esa ubicación
	5. Crea el archivo de diseño con la estructura que representa el diseño de la consulta analítica en la ubicación seleccionada.
<b>Curso alternativo 1:</b>	
Acción del Actor	Respuesta del Sistema
	5. Si el usuario no tiene permisos de escritura en la ubicación seleccionada, el archivo no se puede crear. Se le informa al usuario con un mensaje de error.
<b>Trazabilidad</b>	
Artefactos Anteriores	Artefactos Posteriores
<b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li><a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li><b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li><a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li><a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li><b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li><a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>

**Tabla 25.** Análisis del caso de uso guardar diseño consulta analítica.



- **Caso de uso Abrir diseño consulta analítica**

<b>Nombre del Caso de Uso:</b>	Abrir diseño consulta analítica
<b>Actores</b>	
• Tutor o directivo.	
<b>Propósito:</b>	Permitir al usuario abrir un archivo que representa el diseño de una consulta analítica desde una ubicación determinada.
<b>Resumen:</b>	El usuario desea abrir un archivo de diseño de una consulta analítica y selecciona el lugar donde se encuentra el archivo. El sistema abre el archivo que el usuario desea, lee su estructura y presenta al usuario el diseño guardado.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.17, R.17.1, R.17.2, R.17.3, R.4, R.5, R.6, R.7, R.9.2.
<b>Curso normal de los eventos</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
1. El usuario decide abrir un diseño de una consulta analítica desde una ubicación determinada.	2. presenta una interfaz que le permite seleccionar la ubicación desde donde desea abrir el archivo.
3. Selecciona la ubicación donde se encuentra el archivo que desea abrir.	4. Abre el archivo desde la ubicación seleccionada.
	5. Lee la estructura del archivo que representa el diseño de la consulta analítica.
	6. Presenta el diseño de la consulta analítica guardada en el archivo.
<b>Curso alternativo 1:</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
<b>Trazabilidad</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>



<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>
---------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 26. Análisis del caso de uso abrir diseño consulta analítica.

## CASO DE USO GENERAR GRAFICO DE DATOS

<b>Nombre del Caso de Uso:</b>	Generar grafico de datos
<b>Actores</b>	
• Tutor o directivo.	
<b>Propósito:</b>	Permitir al usuario visualizar los datos de forma grafica, basado en el diseño de una consulta analítica.
<b>Resumen:</b>	El usuario después de realizar el diseño de una consulta analítica, decide ver de forma grafica los datos que se obtienen como resultado de realizar la consulta basado en el diseño. El sistema ejecuta la consulta, obtiene los datos de la bodega de datos y los presenta al usuario de forma grafica.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.10, R.11, R.12, R.14, R.15.
<b>Curso normal de los eventos</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
1. El usuario decide ver en forma de grafico, los datos que se pueden obtener basado en un diseño de una consulta analítica.	2. Construye la consulta basado en el diseño.
	3. Ejecuta la consulta en la bodega de datos.
	4. Convierte los datos obtenidos en una representación matricial.
	5. Establece una relación entre los datos matriciales y su forma grafica.
.	6. Presenta los datos en forma grafica basado en diseño de la consulta analítica creado por el usuario.

Curso alternativo 1:	
Acción del Actor	Respuesta del Sistema
	2. si no se ha definido ningún diseño de consulta analítica no e puede construir la consulta ni continuar con las demás operaciones. Se indica al usuario al respecto con un mensaje de información.
Trazabilidad	
Artefactos Anteriores	Artefactos Posteriores
<b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>[Diagrama de Casos de Uso]</li> </ul>	<ul style="list-style-type: none"> <li><b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>[Diagrama de Secuencia del Sistema]</li> <li>[Modelo Conceptual]</li> </ul> </li> <li><b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>[Caso de Uso Real]</li> </ul> </li> </ul>

Tabla 27. Análisis del caso de uso generar grafico de datos.

## CASO DE USO GESTIONAR ÁREA DE TRABAJO

En este caso de uso el usuario puede guardar varios diseños de consultas analíticas simultáneamente en un archivo denominado área de trabajo. El área de trabajo puede contener diseños de consultas analíticas de varios data marts y de diferentes cubos. Este caso de uso se divide en: guardar área de trabajo y abrir área de trabajo.

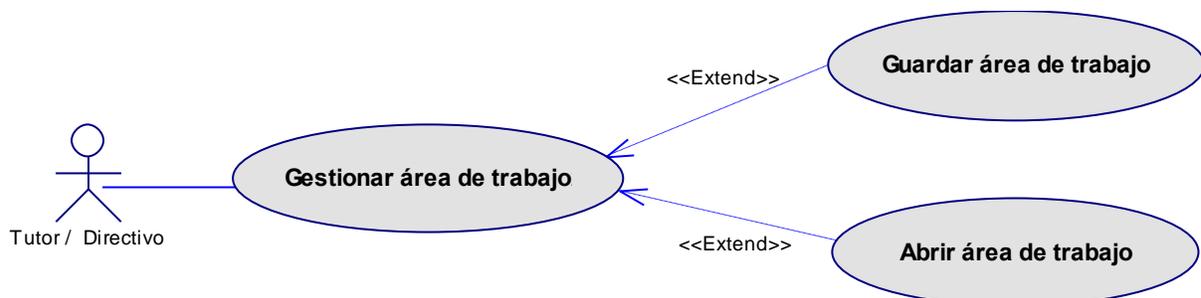


Figura 47. Extensión para el caso de uso Gestionar área de trabajo.

- **Caso de uso Guardar área de trabajo**

<b>Nombre del Caso de Uso:</b>	Guardar área de trabajo
Actores	



• Tutor o directivo.	
<b>Propósito:</b>	Permitir al usuario guardar un área de trabajo (varios diseños de consultas analíticas)
<b>Resumen:</b>	El usuario después de crear varios diseños de consultas analíticas, decide guardar su área de trabajo. Selecciona la ubicación donde desea almacenar el área de trabajo para que el sistema cree el archivo que la va a contener su representación.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.18, R.18.1, R.18.2, R.18.3.
<b>Curso normal de los eventos</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
1. El usuario decide guardar el área de trabajo.	2. Presenta una interfaz que facilita ubicar el lugar donde se quiere guardar el área de trabajo.
3. Selecciona la ubicación donde guardar el área de trabajo	4. Verifica que el área de trabajo no exista en esa ubicación
	5. Crea un archivo que representa el área de trabajo que se presenta en el prototipo.
<b>Curso alterno 1:</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
	5. Si el usuario no tiene permisos de escritura en la ubicación seleccionada, el archivo no se puede crear. Se le informa al usuario con un mensaje de error.
<b>Trazabilidad</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>
<b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>

**Tabla 28.** Análisis del caso de uso guardar área de trabajo.



- **Caso de uso Abrir área de trabajo**

<b>Nombre del Caso de Uso:</b>		Abrir área de trabajo
<b>Actores</b>		
• Tutor o directivo.		
<b>Propósito:</b>	Permitir al usuario abrir un área de trabajo (varios diseños de consultas analíticas) de una ubicación específica.	
<b>Resumen:</b>	El usuario decide abrir un área de trabajo seleccionándola de una ubicación específica. El sistema abre el archivo y presenta al usuario el área de trabajo especificada.	
<b>Tipo:</b>	Primario.	
<b>Referencias Cruzadas:</b>	R.19, R.19.1, R.19.2, R.19.3, R.4, R.5, R.6, R.7, R.9.2	
<b>Curso normal de los eventos</b>		
<b>Acción del Actor</b>		<b>Respuesta del Sistema</b>
1. El usuario decide abrir un área de trabajo desde una ubicación determinada.		2. Presenta una interfaz que le permite seleccionar la ubicación donde se encuentra el área de trabajo.
3. Selecciona la ubicación donde se encuentra el área de trabajo.		4. Abre el área de trabajo desde la ubicación seleccionada.
		5. Lee la estructura del archivo que representa el área de trabajo.
		6. Presenta el área de trabajo guardada.
<b>Curso alterno 1:</b>		
<b>Acción del Actor</b>		<b>Respuesta del Sistema</b>
<b>Trazabilidad</b>		
<b>Artefactos Anteriores</b>		<b>Artefactos Posteriores</b>



<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>
---------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 29. Análisis del caso de uso abrir área de trabajo.

## CASO DE USO FILTRAR NIVELES

<b>Nombre del Caso de Uso:</b>		Filtrar niveles
<b>Actores</b>		
<ul style="list-style-type: none"> <li>• Tutor o directivo.</li> </ul>		
<b>Propósito:</b>	Permitir al usuario filtrar el conjunto de niveles de la jerarquía de una dimensión con los que quiere interactuar	
<b>Resumen:</b>	El usuario selecciona el nivel de la jerarquía en cual quiere realizar el filtrado. El sistema presenta los datos correspondientes a ese nivel que el usuario puede filtrar. El usuario selecciona los datos de nivel con los que quiere interactuar y los filtra.	
<b>Tipo:</b>	Primario.	
<b>Referencias Cruzadas:</b>	R.20, R.20.1, R.20.2.	
<b>Curso normal de los eventos</b>		
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>	
1. El usuario decide filtrar los datos de nivel de un nivel de una jerarquía, para esto selecciona un nivel determinado.	2. Carga y presenta un listado con los datos de nivel correspondientes al nivel seleccionado.	
3. Selecciona los datos de nivel con los cuales quiere interactuar.	4. Presenta la jerarquía con los datos de nivel filtrados para el nivel seleccionado	
<b>Curso alterno 1:</b>		
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>	
<b>Trazabilidad</b>		
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>	



<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>
---------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 30. Análisis del caso de uso filtrar niveles.

### CASO DE USO RESTABLECER DATOS DE NIVEL

<b>Nombre del Caso de Uso:</b>		Restablecer datos de nivel
<b>Actores</b>		
• Tutor o directivo.		
<b>Propósito:</b>	Permitir al usuario restablecer el estado original de los datos de un nivel de una jerarquía en una dimensión.	
<b>Resumen:</b>	El usuario selecciona el nivel que ha sido filtrado y quiere restablecer. Solicita al sistema restablecer el nivel al estado original y éste lo presenta.	
<b>Tipo:</b>	Primario.	
<b>Referencias Cruzadas:</b>	R.21	
<b>Curso normal de los eventos</b>		
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>	
1. El usuario decide restablecer los datos de nivel de un nivel que ha sido filtrado, para esto selecciona un nivel filtrado determinado.	2. Presenta el nivel con sus datos correspondientes en su forma original (como se encontraban antes de ser filtrados).	
<b>Curso alterno 1:</b>		
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>	
<b>Trazabilidad</b>		
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>	



<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"> <li>• <a href="#">[Diagrama de Casos de Uso]</a></li> </ul>	<ul style="list-style-type: none"> <li>• <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Diagrama de Secuencia del Sistema]</a></li> <li>○ <a href="#">[Modelo Conceptual]</a></li> </ul> </li> <li>• <b>Artefactos del Diseño:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de Uso Real]</a></li> </ul> </li> </ul>
---------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 31. Análisis del caso de uso restablecer datos de nivel.

### CASO DE USO CAMBIAR JERARQUÍA

<b>Nombre del Caso de Uso:</b>	Cambiar jerarquía
<b>Actores</b>	
• Tutor o directivo.	
<b>Propósito:</b>	Permitir al usuario cambiar la jerarquía de una dimensión
<b>Resumen:</b>	El usuario selecciona la dimensión en la cual quiere cambiar la jerarquía. El sistema presenta las jerarquías definidas para esa dimensión y el usuario selecciona la que quiere utilizar para sus análisis.
<b>Tipo:</b>	Primario.
<b>Referencias Cruzadas:</b>	R.22, R.22.1, R.22.2.
<b>Curso normal de los eventos</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
1. El usuario decide cambiar la jerarquía de una dimensión seleccionada.	2. Carga y presenta un listado con las jerarquías definidas para la dimensión seleccionada.
3. Selecciona la jerarquía con la que quiere realizar análisis para la dimensión seleccionada.	4. Presenta la dimensión con la jerarquía seleccionada
<b>Curso alterno 1:</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
<b>Trazabilidad</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>



<p><b>Artefactos del Análisis:</b></p> <ul style="list-style-type: none"><li>• <b>[Diagrama de Casos de Uso]</b></li></ul>	<ul style="list-style-type: none"><li>• <b>Artefactos del Análisis:</b><ul style="list-style-type: none"><li>○ <b>[Diagrama de Secuencia del Sistema]</b></li><li>○ <b>[Modelo Conceptual]</b></li></ul></li><li>• <b>Artefactos del Diseño:</b><ul style="list-style-type: none"><li>○ <b>[Caso de Uso Real]</b></li></ul></li></ul>
----------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Tabla 32. Análisis del caso de uso cambiar jerarquía.

## ANEXO 10. DIAGRAMAS DE SECUENCIA DEL SISTEMA

### DIAGRAMA DE SECUENCIA DISEÑAR CONSULTA ANALÍTICA

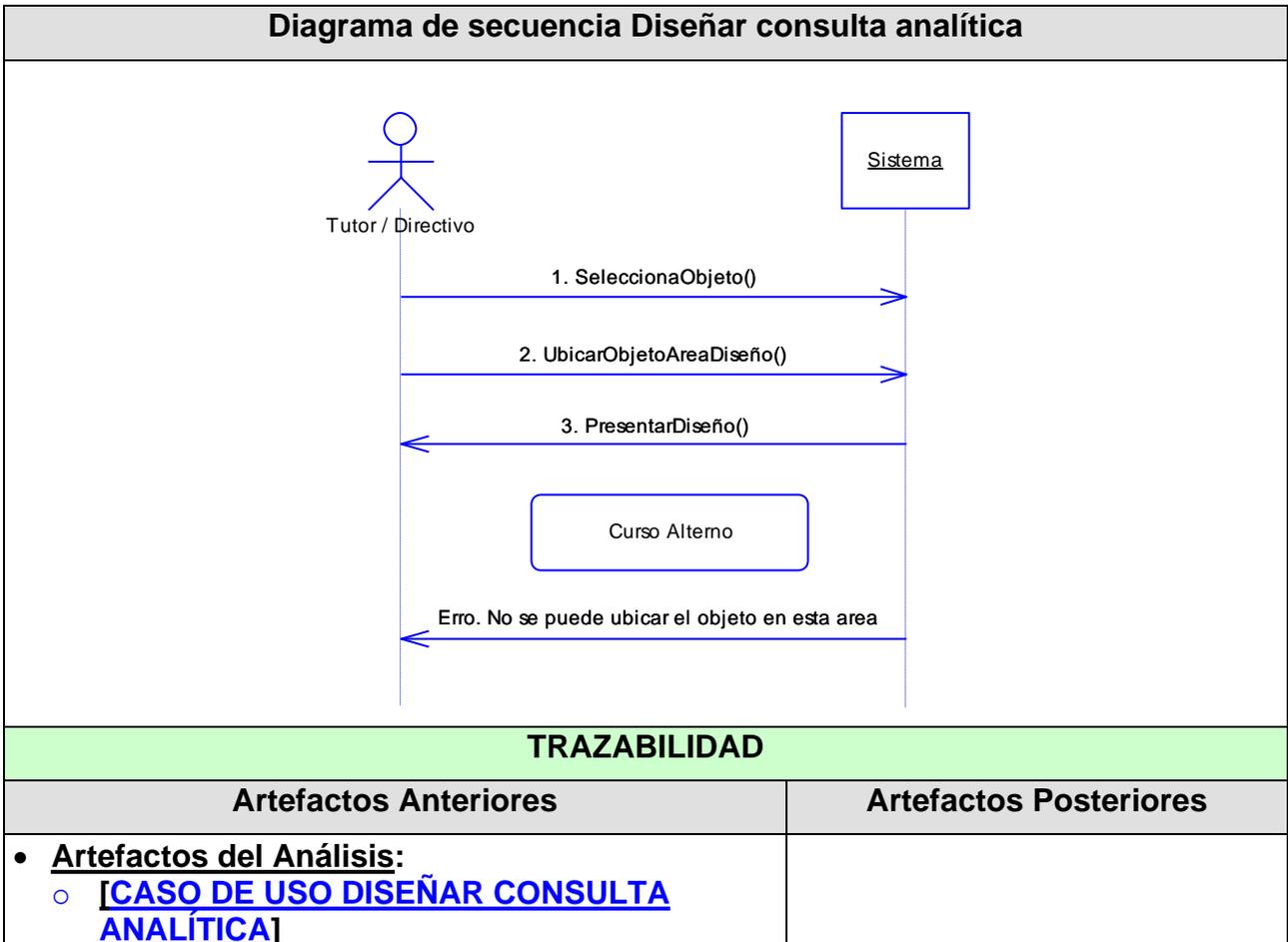


Tabla 33. Diagrama de secuencia caso de uso Diseñar Consulta Analítica.



## DIAGRAMA DE SECUENCIA GENERAR MATRIZ DE DATOS

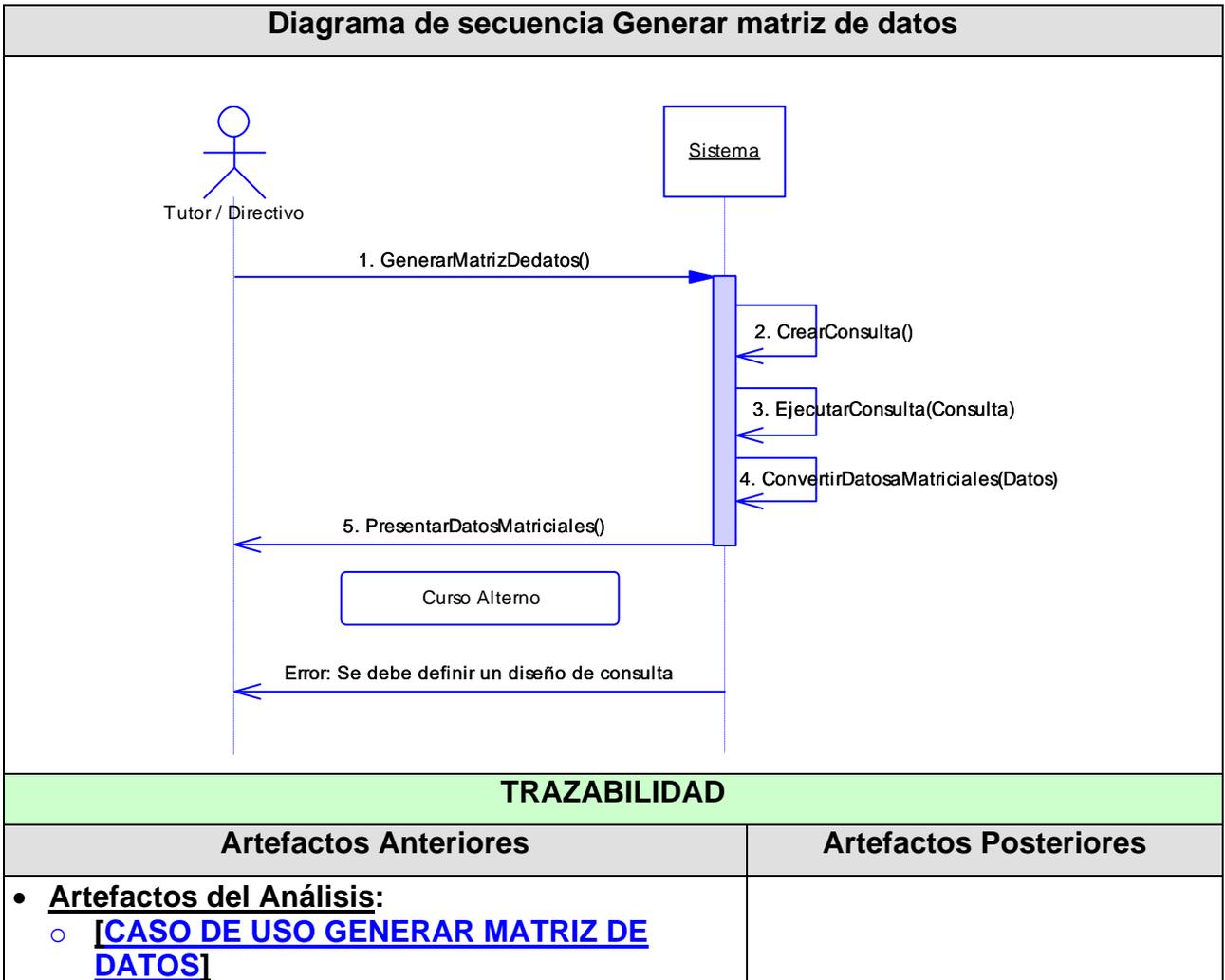


Tabla 34. Diagrama de secuencia caso de uso Generar matriz de datos.



## DIAGRAMA DE SECUENCIA ENTRAR AL SISTEMA

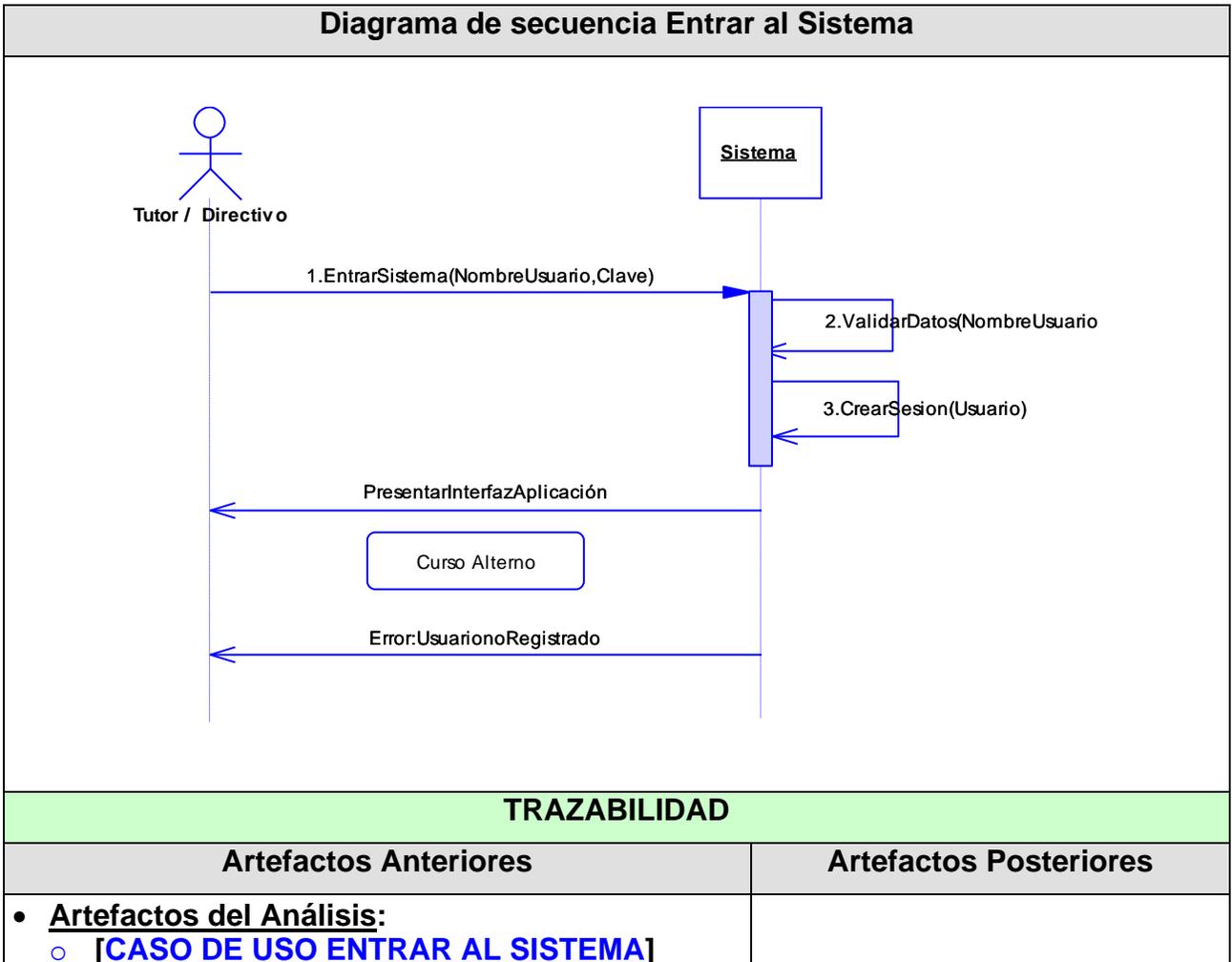


Tabla 35. Diagrama de secuencia caso de uso Entrar al sistema.

## DIAGRAMAS DE SECUENCIA GESTIONAR CONSULTA ANALÍTICA

- Diagrama de secuencia Guardar diseño consulta analítica

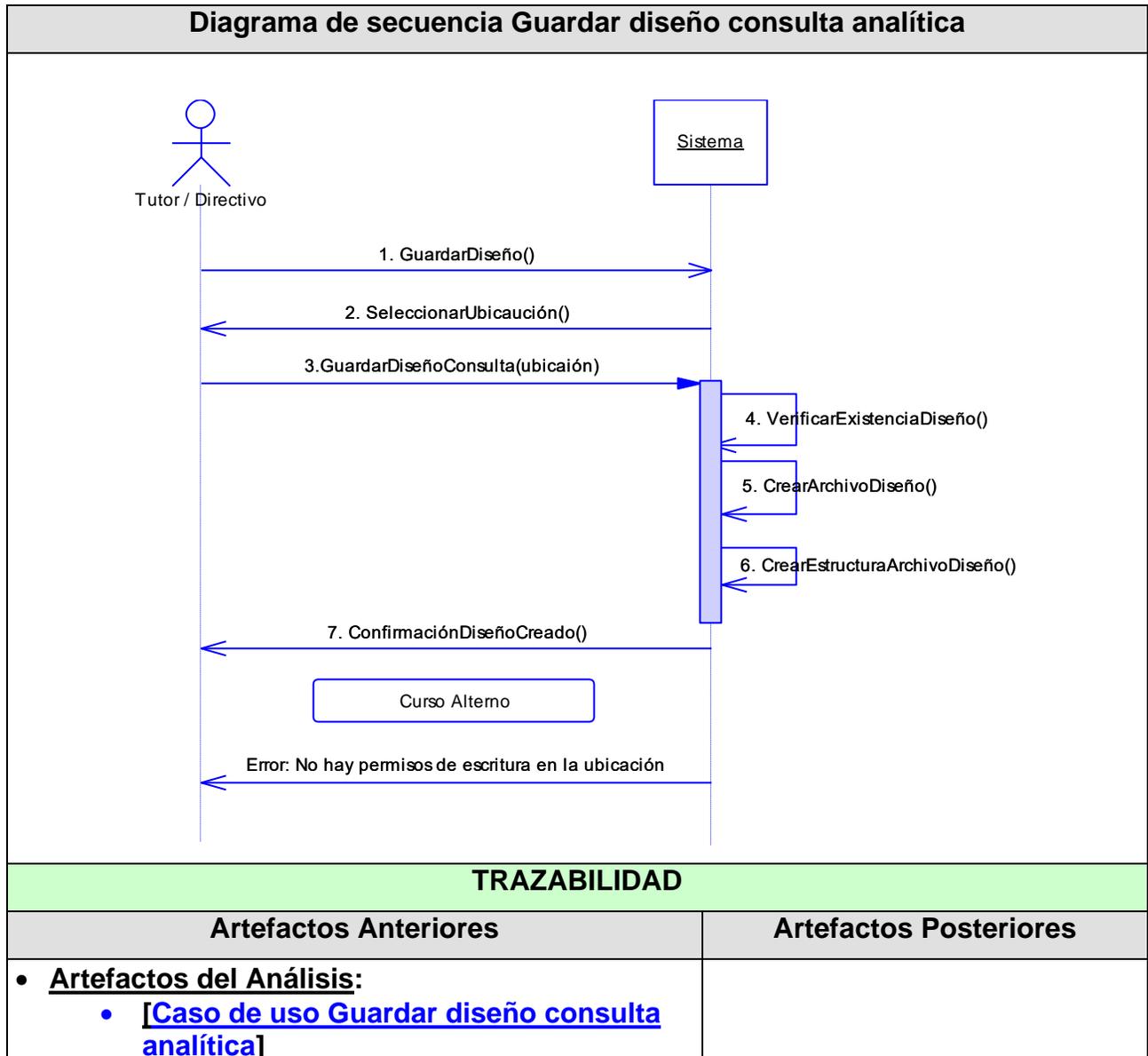


Tabla 36. Diagrama de secuencia caso de uso Guardar diseño consulta analítica.



- Diagrama de secuencia Abrir diseño consulta analítica

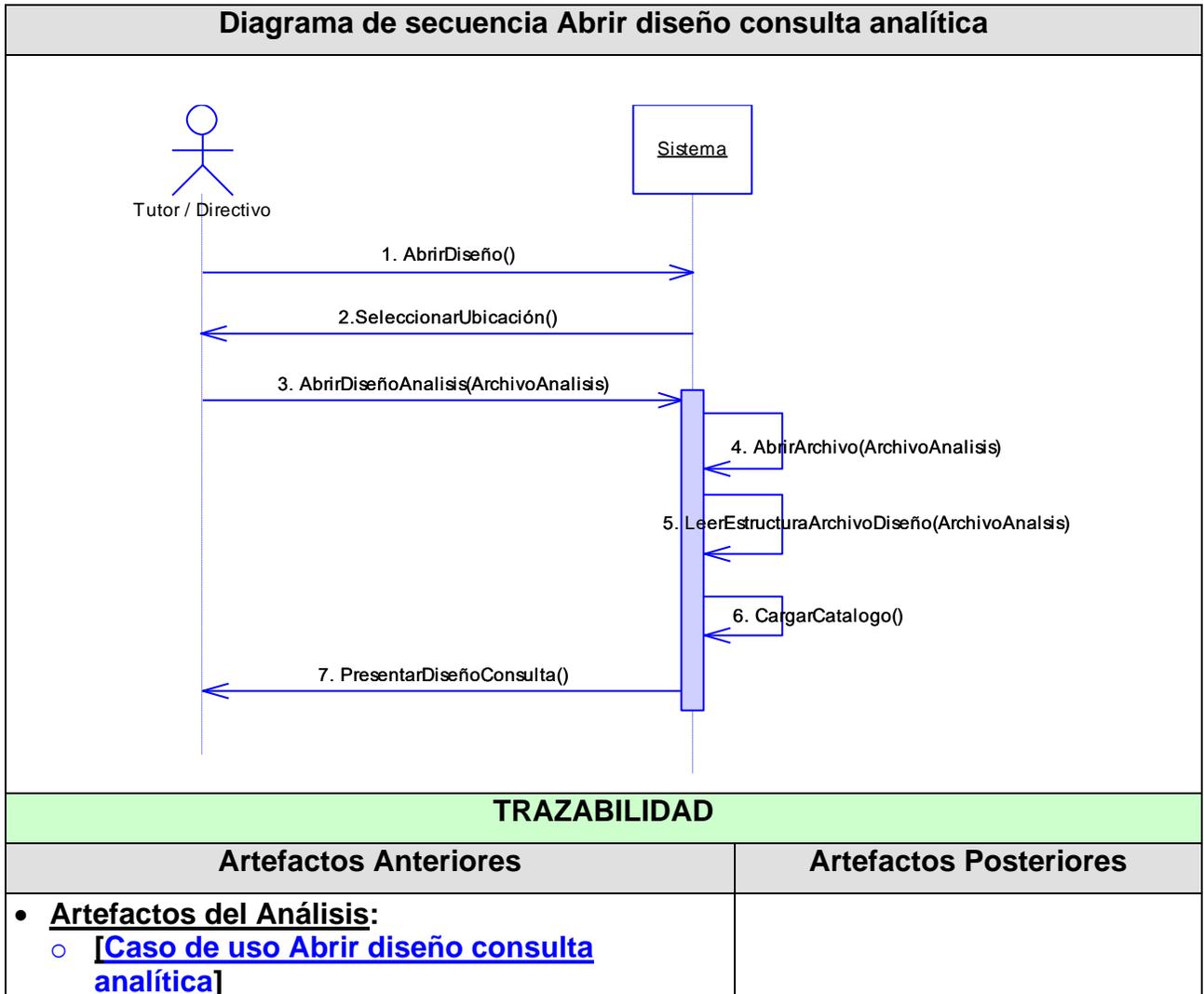


Tabla 37. Diagrama de secuencia caso de uso Abrir diseño consulta analítica.



## DIAGRAMA DE SECUENCIA GENERAR GRAFICO DE DATOS

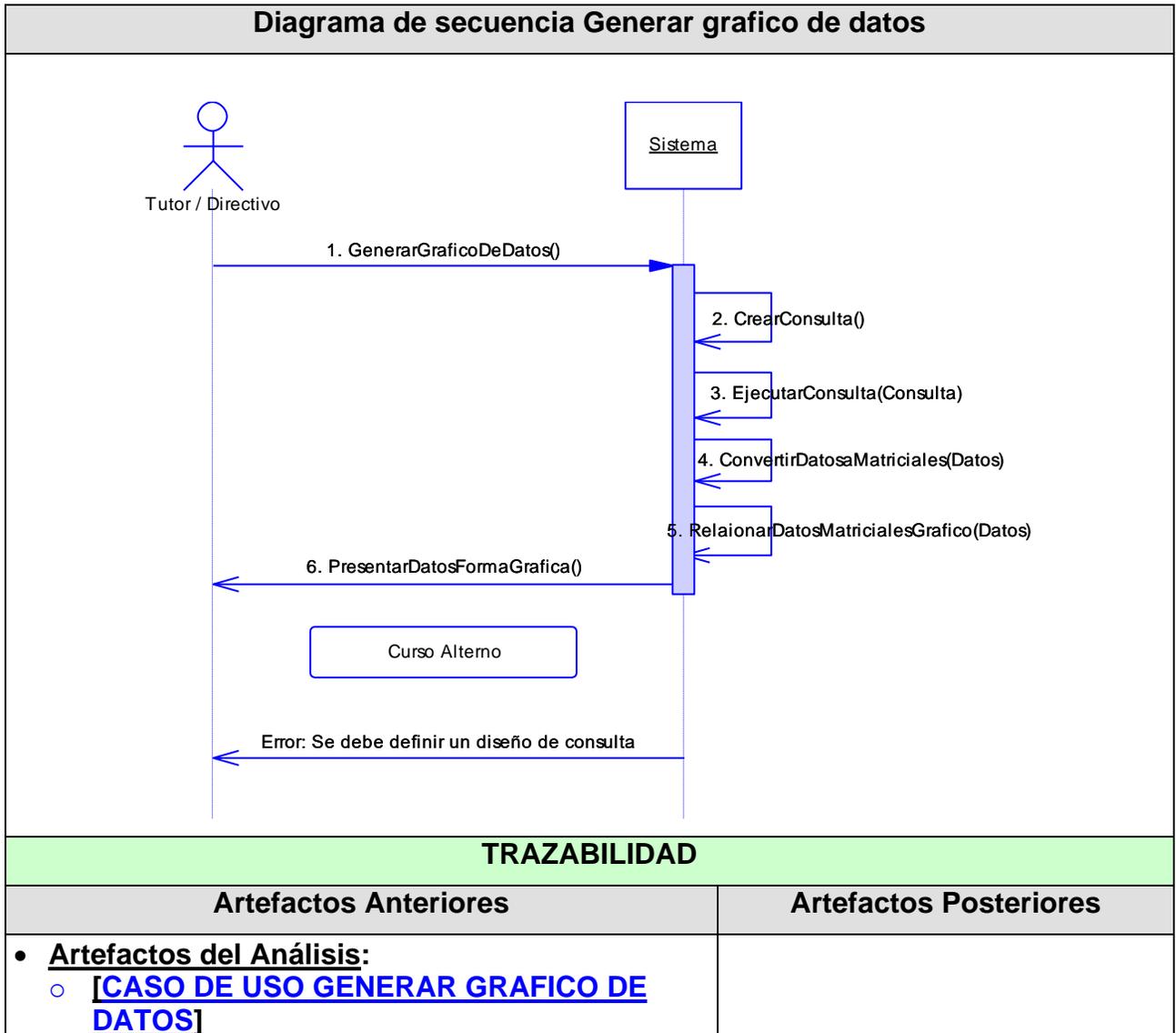


Tabla 38. Diagrama de secuencia caso de uso Generar grafico de datos.

## DIAGRAMAS DE SECUENCIA GESTIONAR ÁREA DE TRABAJO

- Diagrama de secuencia Guardar área de trabajo

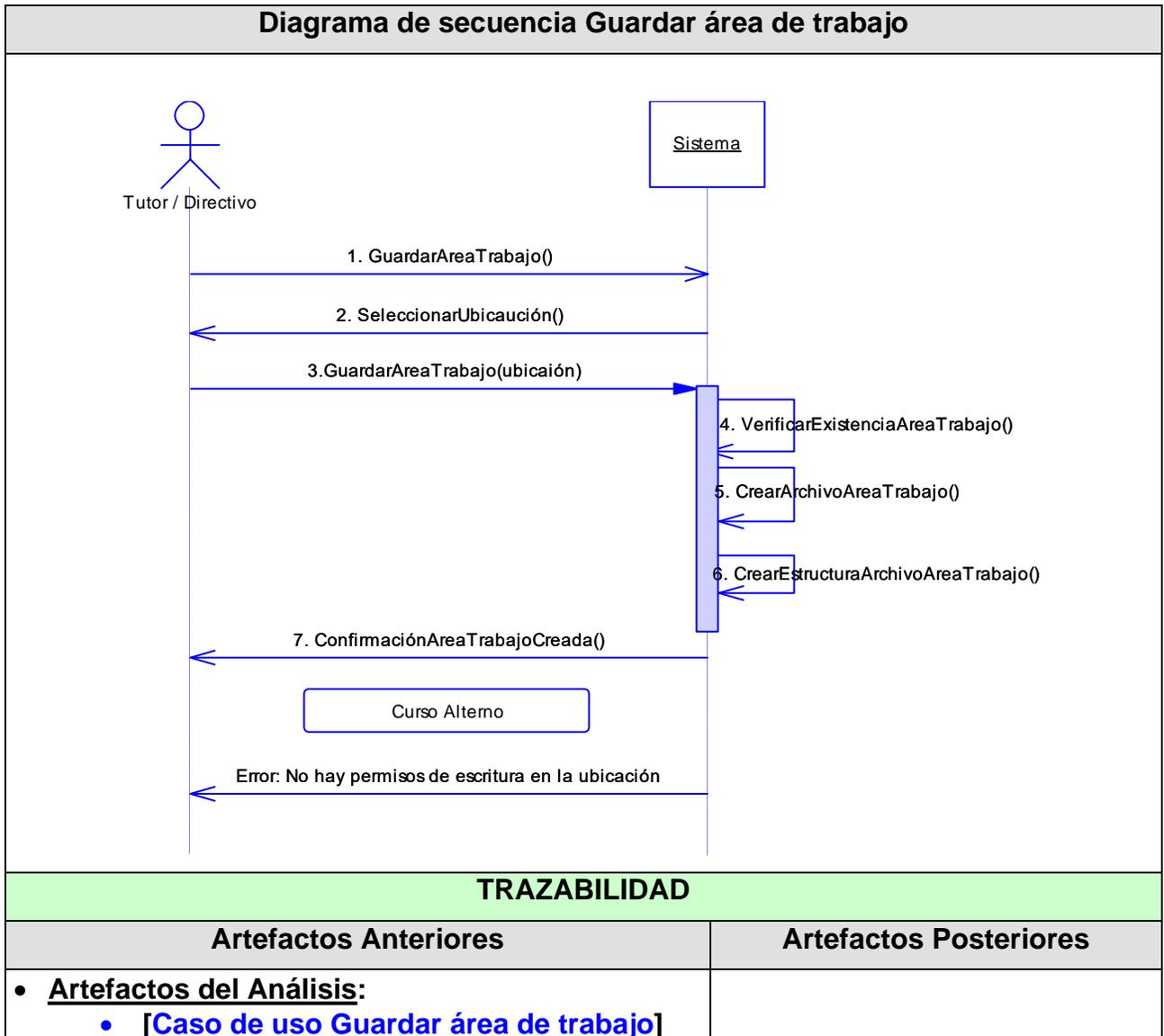


Tabla 39. Diagrama de secuencia caso de uso Guardar área de trabajo.

- **Diagrama de secuencia Abrir área de trabajo**

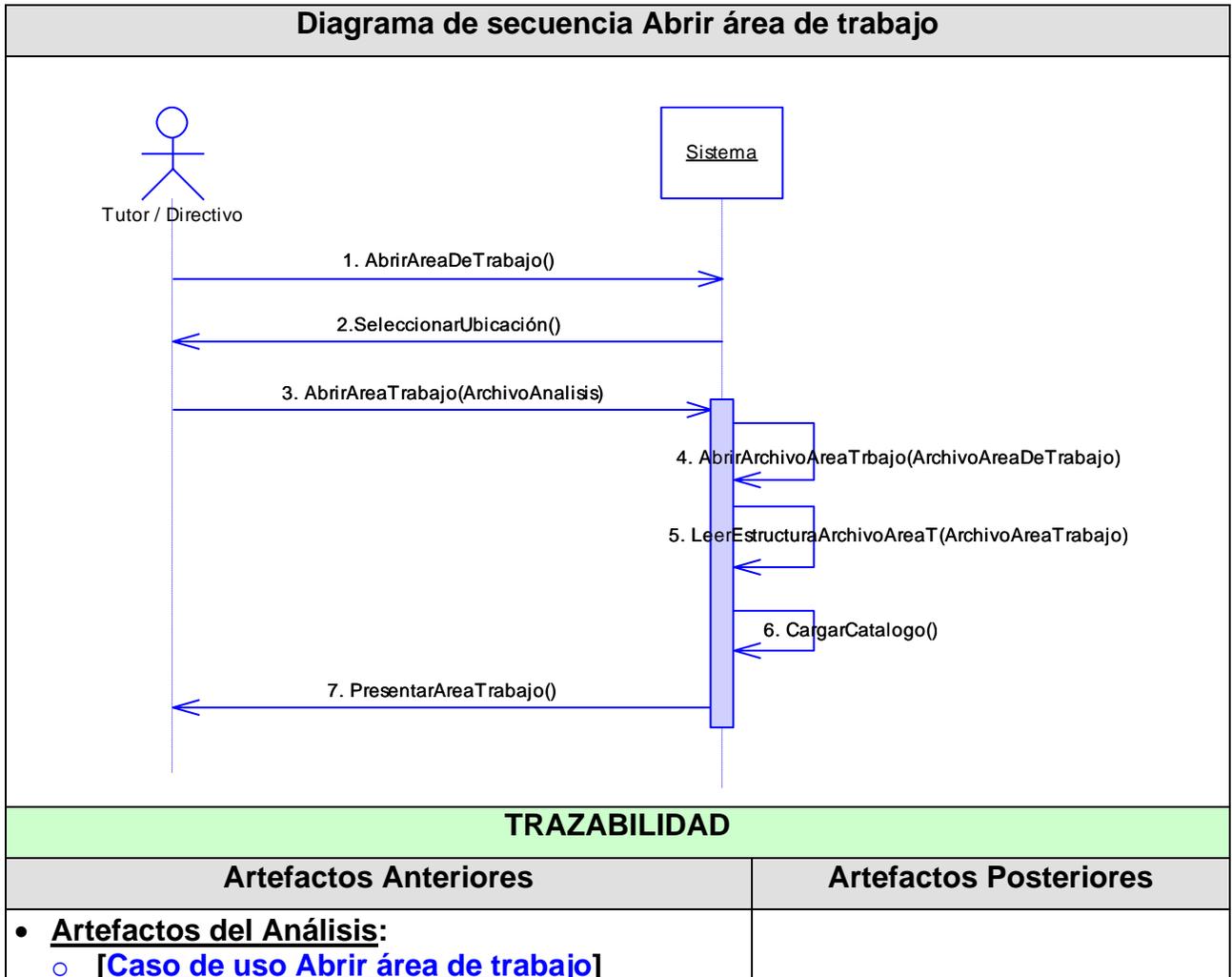
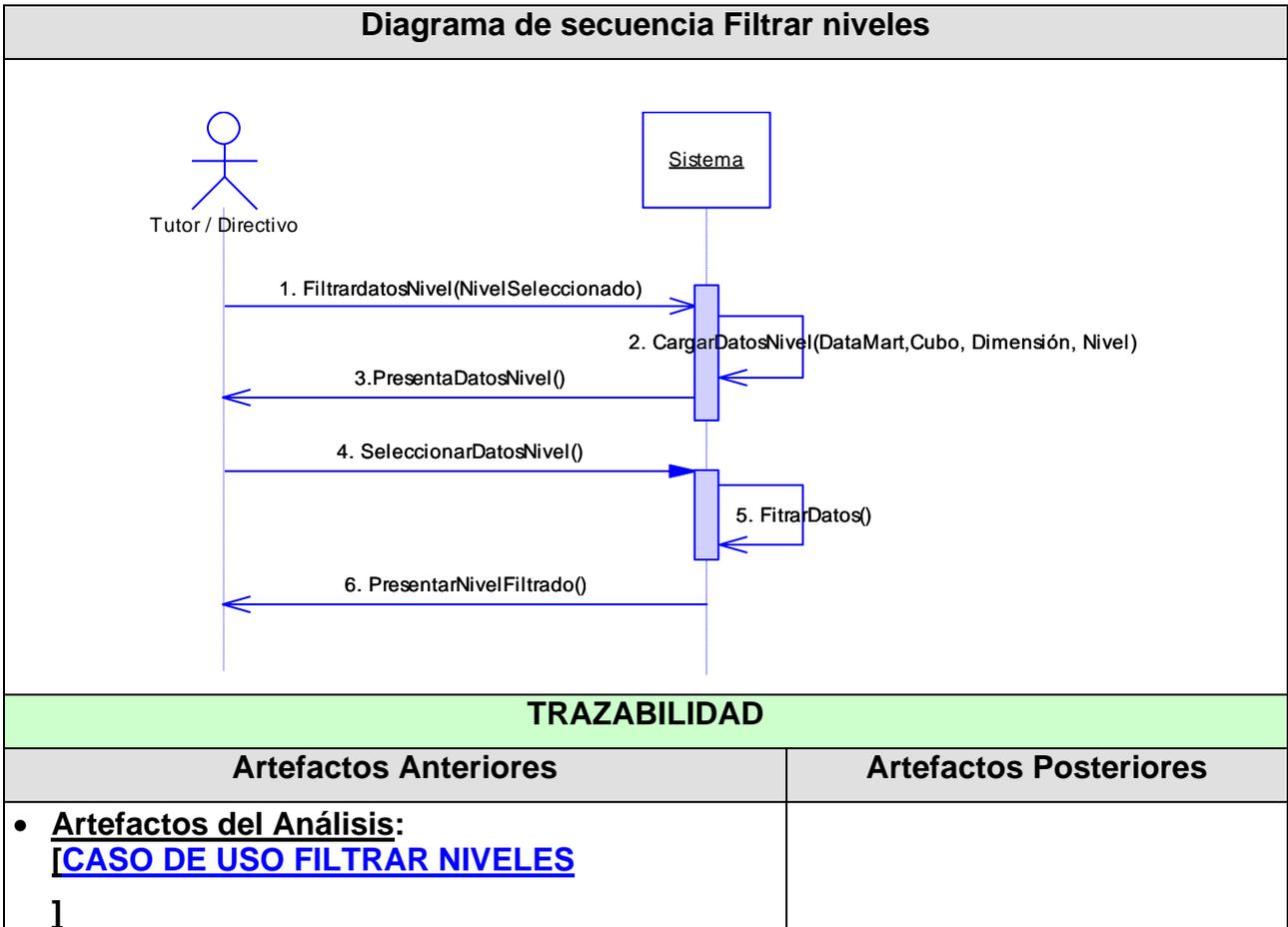


Tabla 40. Diagrama de secuencia caso de uso Abrir área de trabajo.



## DIAGRAMA DE SECUENCIA FILTRAR NIVELES



**Tabla 41.** Diagrama de secuencia caso de uso Filtrar niveles.



## DIAGRAMA DE SECUENCIA RESTABLECER DATOS DE NIVEL

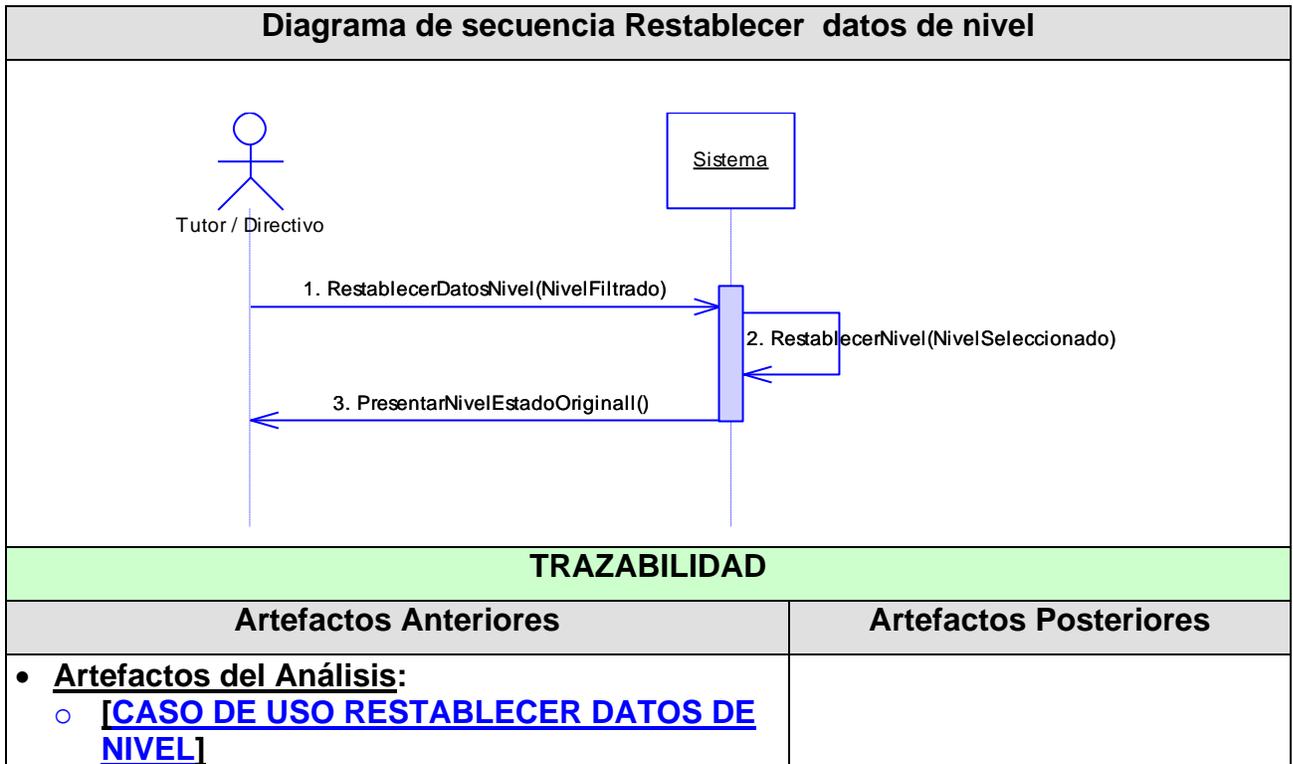


Tabla 42. Diagrama de secuencia caso de uso Restablecer datos de nivel.

## DIAGRAMA DE SECUENCIA CAMBIAR JERARQUÍA

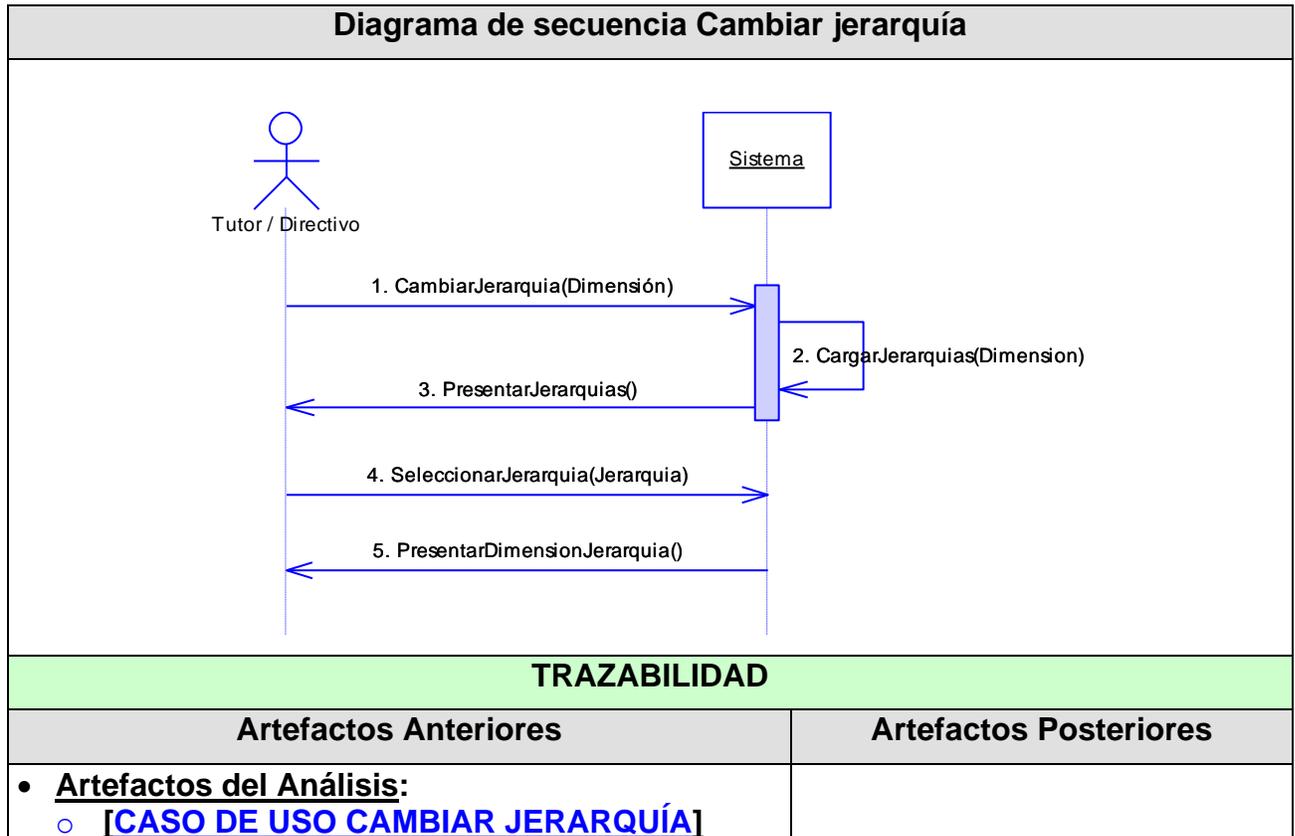


Tabla 43. Diagrama de secuencia caso de uso Cambiar jerarquía.



## ANEXO 11. CASOS DE USO REALES

### CASO DE USO REAL DISEÑAR CONSULTA ANALÍTICA

**Caso de uso real Diseñar consulta analítica**

The screenshot displays a Business Intelligence tool interface. On the left, a tree view titled 'Modelos Dimensionales' shows a hierarchy: 'EVALUACION' > 'INTERACCION\_ESTUDIANTE' > 'Visualización Recurso' > 'Dimensiones' > 'ESTUDIANTE' > 'GENEROS'. The 'GENEROS' node is highlighted with a green box labeled '1'. Below it, 'FEMENINO' and 'MASCULINO' are listed. Other dimensions include 'RECURSO', 'PROFESOR', 'FECHA', 'PERIODO', 'CURSO', and 'ACTIVIDAD'. Under 'Medidas', several metrics are listed, with 'Tiempo Total Visualización' highlighted by a green box labeled '1.c'. A mouse cursor is shown hovering over '1.a' (near 'FEMENINO') and '1.b' (near 'PERIODO').

On the right, a window titled 'ANÁLISIS DE DATOS 1' is open. It contains buttons for 'Generar matriz de datos', 'Generar gráfico de datos', and 'Abrir análisis'. Below these are tabs for 'Análisis Matricial' and 'Análisis Gráfico'. A section titled 'Eliminar dimensiones seleccionadas' and 'Limpiar diseño' is visible. A 'Resumen' box shows: 'Nombre del Data Mart: INTERACCION\_ESTUDIANTE', 'Nombre del Cubo: VISUALIZACION\_RECURSO', and 'Descripción del Cubo: Este cubo esta relacionado con los indicadores que se monitorean en el STI y que muestran la interacción de un estudiante'. Below the summary are two panels: 'Area medidas' and 'Area Dimensiones'. A mouse cursor is shown hovering over a box in the 'Area Dimensiones' panel, labeled '1.d'. Another mouse cursor is shown hovering over a box in the 'Area Dimensiones' panel, labeled '1.e'.



The screenshot shows the OLAP tool interface. On the left, the 'Modelos Dimensionales' pane displays a hierarchy: EVALUACION > INTERACCION\_ESTUDIANTE > Visualización Recurso > Dimensiones > ESTUDIANTE > GENEROS > FEMENINO and MASCULINO. Other dimensions include RECURSO, PROFESOR, FECHA, PERIODO, CURSO, ACTIVIDAD, and Medidas (Tiempo Total Visualización, etc.). On the right, the 'ANÁLISIS DE DATOS 1' window shows the 'Análisis Matricial' view. The 'Resumen' section indicates the Data Mart is 'INTERACCION\_ESTUDIANTE' and the Cube is 'VISUALIZACION\_RECURSO'. A table below shows the selected dimensions: FEMENINO, MASCULINO, TOTAL GENEROS, and ESTUDIANTE. A green box with the number '2' highlights the 'TOTAL GENEROS' row.

This screenshot is similar to the one above but includes annotations. A green box with the number '3' highlights the 'Medidas' folder in the dimension hierarchy. A green box with '3.a' points to the 'Tiempo Visualización Recurso' measure. A green box with '3.b' points to the 'TOTAL GENEROS' row in the matrix table. A green box with '3.c' points to a blue button labeled '3.c' in the interface.

The screenshot shows the 'ANÁLISIS DE DATOS 1' interface. On the left, a tree view under 'Modelos Dimensionales' shows the hierarchy: EVALUACION > INTERACCION\_ESTUDIANTE > Visualización Recurso > Dimensiones > ESTUDIANTE > GENEROS (FEMENINO, MASCULINO). Other dimensions include RECURSO, PROFESOR, FECHA, PERIODO, CURSO, and ACTIVIDAD. Measures include 'Tiempo Total Visualización', 'Tiempo Visualización Recurso', 'Indicador de Visualización', and 'Tiempo Descanso Recurso'. The right pane shows the 'ANÁLISIS DE DATOS 1' window with buttons for 'Generar matriz de datos', 'Generar gráfico de datos', and 'Abrir análisis'. Below these are 'Análisis Matricial' and 'Análisis Gráfico'. A 'Resumen' section displays: 'Nombre del Data Mart: INTERACCION\_ESTUDIANTE', 'Nombre del Cubo: VISUALIZACION\_RECURSO', and a description. The visualization design area shows a table with columns for 'Tiempo Visualización Recurso' and 'ESTUDIANTE', with a green box highlighting the number '4' in the 'ESTUDIANTE' column.

**Curso normal de los eventos**

Acción del Actor	Respuesta del Sistema
1. El usuario selecciona un objeto dimensional [1] o medida [3] de un cubo del catalogo.	
2. Utilizando la opción arrastrar y colocar ([1.a,...,1.e], [3.a,...,3.c]) ubica el objeto seleccionado en el área correspondiente.	3. Presenta el diseño de la consulta analítica que se lleva hasta ese momento [2] [4].

**TRAZABILIDAD**

Artefactos Anteriores	Artefactos Posteriores
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <b>[CASO DE USO DISEÑAR CONSULTA ANALÍTICA]</b></li> </ul> </li> </ul>	

Tabla 44. Caso de uso real Diseñar consulta analítica.



## CASO DE USO REAL GENERAR MATRIZ DE DATOS

### Caso de uso real Generar matriz de datos

**Modelos Dimensionales**

Generar matriz de datos | Generar gráfico de datos | Abrir análisis | Guar

Filtro | Restablecer | Cambiar Jerarquía

**Dimensiones:** ESTUDIANTE (GENEROS: FEMENINO, MASCULINO), RECURSO (PROFESOR, FECHA, PERIODO, CURSO), ACTIVIDAD (Tiempo Total Visualización, Tiempo Visualización Recurso, Indicador de Visualización, Tiempo Descanso Recurso)

**Resumen:** Nombre del Data Mart: INTERACCION\_ESTUDIANTE, Nombre del Cubo: VISUALIZACION\_RECURSO, Descripción del Cubo: Este cubo esta relacionado con los indicadores que se monitorean en el STI y que muestran la interacción de un estudiante

**Tabla de Resumen:**

Tiempo Visualización Recurso	FEMENINO	MASCULINO	TOTAL GENEROS	ESTUDIANTE
INTRODUCCION A LA INFORMATICA	31,96	156,17	188,13	
PROGRAMACION ORIENTADA A OBJETOS	60,23	112,77	173	
TOTAL TODAS LAS ASIGNATURAS	92,19	268,94	361,13	

**Curso normal de los eventos**



Acción del Actor	Respuesta del Sistema
1. Una vez terminado el diseño [1] El usuario decide ver los datos en forma de matriz de datos [2] que se pueden obtener basado en el diseño realizado.	2. Construye la consulta basado en el diseño.
	3. Ejecuta la consulta en la bodega de datos.
	4. Convierte los datos obtenidos en una representación matricial.
.	5. Presenta los datos en forma matricial basado en diseño de la consulta analítica creado por el usuario [3].
TRAZABILIDAD	
Artefactos Anteriores	Artefactos Posteriores
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[CASO DE USO GENERAR MATRIZ DE DATOS]</a></li> </ul> </li> </ul>	

Tabla 45. Caso de uso real Generar matriz de datos.

## CASO DE USO REAL ENTRAR AL SISTEMA

### Caso de uso real Entrar al sistema

**Curso normal de los eventos**



Acción del Actor	Respuesta del Sistema
1. El usuario ingresa su nombre de usuario [1] y clave [2] y presiona el botón “conectar” [3].	2. El sistema recibe los datos y comprueba la validez de los mismos.
	3. Si el usuario esta registrado, el sistema crea una sesión y visualiza la interfaz principal de la aplicación [4].
TRAZABILIDAD	
Artefactos Anteriores	Artefactos Posteriores
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[CASO DE USO ENTRAR AL SISTEMA]</a></li> </ul> </li> </ul>	

Tabla 46. Caso de uso real Entrar al sistema.

## CASO DE USO REAL GUARDAR DISEÑO CONSULTA ANALÍTICA

**Caso de uso real Guardar diseño consulta analítica**

Dimensionales

Restablecer | Cambiar Jerarquía

CCION  
CCION\_ESTUDIANTE  
alización Recurso  
Dimensiones

- ESTUDIANTE
  - GENEROS
    - FEMENINO
    - MASCULINO
- RECURSO
- PROFESOR
- FECHA
- PERIODO
- CURSO
- ACTIVIDAD
  - ACTIVIDADES
    - Análisis y Reflexión
    - Explicación
    - Juegos
    - Lectura
    - Refuerzo

Medidas

- Tiempo Total Visualización
- Tiempo Visualización Recurso

Generar matriz de datos | 
 Generar gráfico de datos | 
 Abrir análisis | 
 Guardar análisis

5,277 Análisis Matricial | 
 5,277 Análisis Gráfico

✖ Eliminar dimensiones seleccionadas | 
 🧹 Limpiar diseño

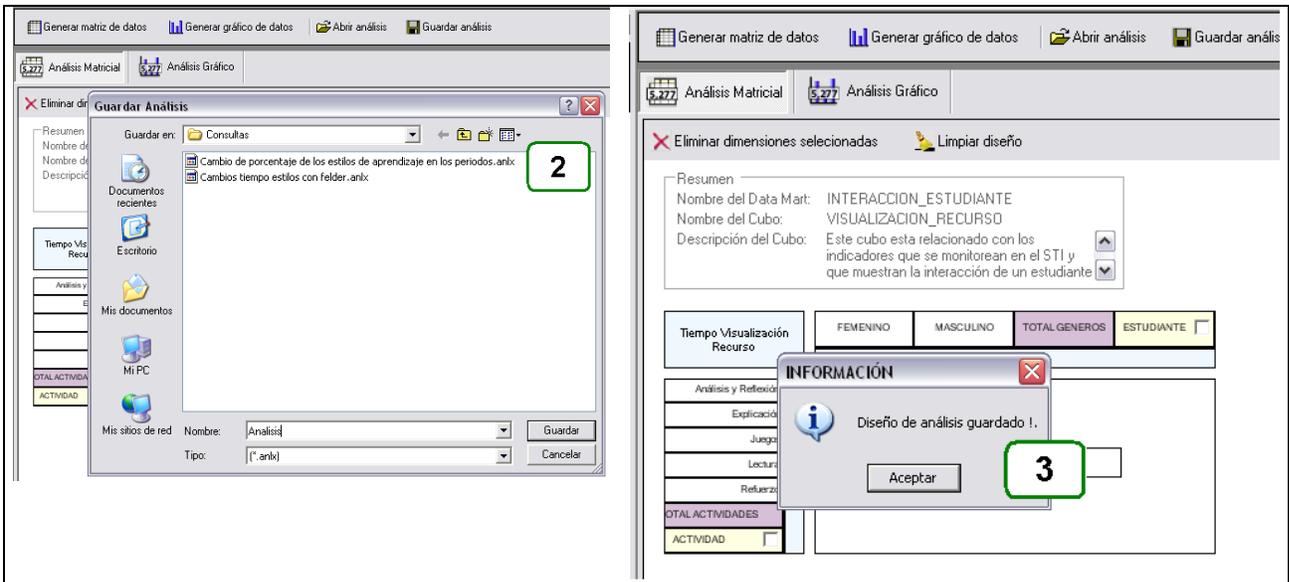
Resumen

Nombre del Data Mart: INTERACCION\_ESTUDIANTE

Nombre del Cubo: VISUALIZACION\_RECURSO

Descripción del Cubo: Este cubo esta relacionado con los indicadores que se monitorean en el STI y que muestran la interacción de un estudiante

Tiempo Visualización Recurso	FEMENINO	MASCULINO	TOTAL GENEROS	ESTUDIANTE <input type="checkbox"/>
Análisis y Reflexión	33,13	107,83	140,96	
Explicación	20,01	50,56	70,57	
Juegos	22,22	26,73	48,95	
Lectura	3,48	24,21	27,69	
Refuerzo	13,35	59,61	72,96	
<b>TOTAL ACTIVIDADES</b>	<b>92,19</b>	<b>268,94</b>	<b>361,13</b>	



### Curso normal de los eventos

Acción del Actor	Respuesta del Sistema
1. El usuario decide guardar un diseño de consulta analítica y da clic sobre el botón [1].	2. presenta una pantalla de dialogo que le permite seleccionar la ubicación donde guardar diseño [2].
3. Selecciona la ubicación donde guardar el archivo de diseño en la pantalla de dialogo [2]	4. Verifica que el archivo de diseño no exista en esa ubicación
	5. Crea el archivo de diseño con la estructura que representa el diseño de la consulta analítica en la ubicación seleccionada y confirma la creación del mismo [3].

### TRAZABILIDAD

Artefactos Anteriores	Artefactos Posteriores
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de uso Guardar diseño consulta analítica]</a></li> </ul> </li> </ul>	

Tabla 47. Caso de uso real Guardar diseño consulta analítica.

## CASO DE USO REAL ABRIR DISEÑO CONSULTA ANALÍTICA

### Caso de uso real Abrir diseño consulta analítica

Curso normal de los eventos

Acción del Actor	Respuesta del Sistema
<p><b>1.</b> El usuario decide abrir un diseño de una consulta analítica desde una ubicación determinada y da clic sobre el botón <b>[1]</b>.</p>	<p><b>2.</b> presenta una pantalla de dialogo <b>[2]</b> que le permite seleccionar la ubicación desde donde desea abrir el archivo.</p>

Sistema de apoyo para la toma de decisiones en Unicauca Virtual utilizando Bodegas de Datos (Data Warehouse) y OLAP  
Unicauca Virtual Fase II

- 172 -



3. Selecciona la ubicación donde se encuentra el archivo que desea abrir y presiona el botón de abrir [3].	4. Abre el archivo desde la ubicación seleccionada.
	5. Lee la estructura del archivo que representa el diseño de la consulta analítica.
	6. Presenta el diseño de la consulta analítica guardada en el archivo [4].
<b>TRAZABILIDAD</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de uso Abrir diseño consulta analítica]</a></li> </ul> </li> </ul>	

Tabla 48. Caso de uso real Abrir diseño consulta analítica.

### CASO DE USO REAL GENERAR GRAFICO DE DATOS

#### Caso de uso real Generar grafico de datos

Modelos Dimensionales

Filtro Restablecer Cambiar Jerarquia

- [-] EVALUACION
  - [-] INTERACCION\_ESTUDIANTE
    - [-] Visualización Recurso
      - [-] Dimensiones
        - [-] ESTUDIANTE
          - [-] GENEROS
            - [-] FEMENINO
            - [-] MASCULINO
          - [-] RECURSO
          - [-] PROFESOR
          - [-] FECHA
          - [-] PERIODO
          - [-] CURSO
          - [-] ACTIVIDAD
            - [-] ACTIVIDADES
              - [-] Análisis y Reflexión
              - [-] Explicación
              - [-] Juegos
              - [-] Lectura
              - [-] Refuerzo
- [-] Medidas
  - [-] Tiempo Total Visualización
  - [-] Tiempo Visualización Recurso

Generar matriz de datos Generar gráfico de datos Abrir análisis Guardar

Análisis Matricial Análisis Gráfico

Eliminar dimensiones seleccionadas Limpiar diseño

Resumen

Nombre del Data Mart: INTERACCION\_ESTUDIANTE

Nombre del Cubo: VISUALIZACION\_RECURSO

Descripción del Cubo: Este cubo esta relacionado con los indicadores que se monitorean en el STI y que muestran la interacción de un estudiante

Tiempo Visualización Recurso	FEMENINO	MASCULINO	TOTAL GENEROS	ESTUDIANTE
Análisis y Reflexión	33,13	107,83	140,96	
Explicación	20,01	50,56	70,57	
Juegos	22,22	26,73	48,95	
Lectura	3,48	24,21	27,69	
Refuerzo	13,35	59,61	72,96	
TOTAL ACTIVIDADES	92,19	268,94	361,13	



**Curso normal de los eventos**

Acción del Actor	Respuesta del Sistema
1. El usuario decide ver en forma de grafico, los datos que se pueden obtener basado en un diseño de una consulta analítica y para esto da clic en el botón [1].	2. Construye la consulta basado en el diseño.
	3. Ejecuta la consulta en la bodega de datos.
	4. Convierte los datos obtenidos en una representación matricial.
	5. Establece una relación entre los datos matriciales y su forma grafica y presenta la grafica [2].

**TRAZABILIDAD**

Artefactos Anteriores	Artefactos Posteriores
-----------------------	------------------------



- **Artefactos del Análisis:**
  - **[CASO DE USO GENERAR GRAFICO DE DATOS]**

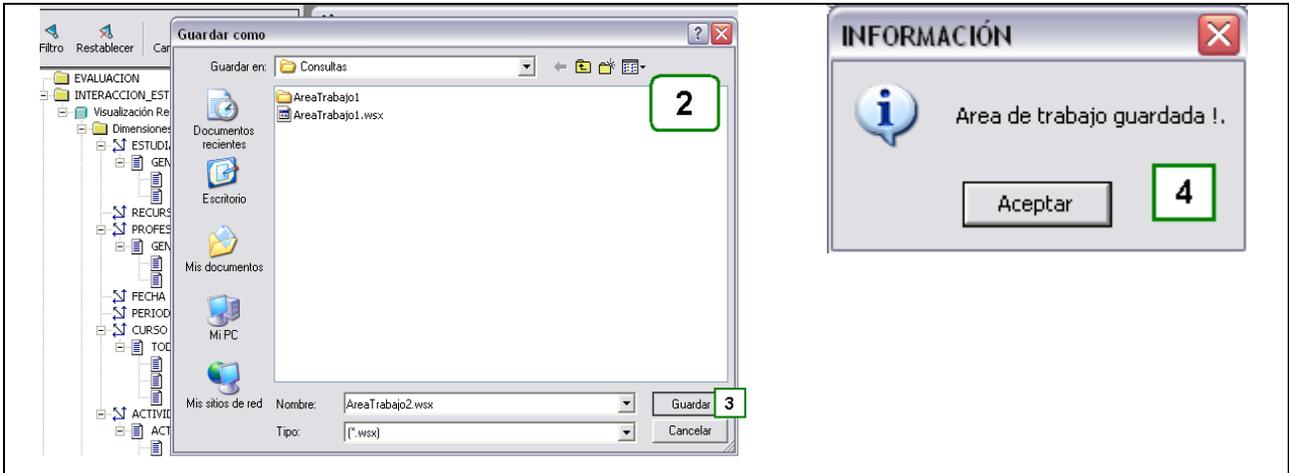
Tabla 49. Caso de uso real Generar grafico de datos.

## CASO DE USO REAL GUARDAR ÁREA DE TRABAJO

### Caso de uso real Guardar área de trabajo

The screenshot shows a software interface for data analysis. On the left is a tree view with categories like 'INTERACCION\_ESTUDIANTE', 'RECURSO', 'PERIODO', 'CURSO', 'ACTIVIDAD', and 'ACTIVIDADES'. The 'Guardar área de trabajo' option in the 'Análisis' menu is highlighted with a green box and the number '1'. On the right, there are three overlapping windows titled 'ANALISIS DE DATOS 3', 'ANALISIS DE DATOS 4', and 'ANALISIS DE DATOS 5'. The 'ANALISIS DE DATOS 4' window is the most prominent and contains a summary and a data table.

Resumen	FEMENINO	MASCULINO	TOTAL GENEROS	ESTUDIANTE
Nombre del Data Mart: INTERACCION_ESTUDIANTE				
Nombre del Cubo: VISUALIZACION_RECURSO				
Descripción del Cubo: Este cubo esta relacionado con los indicadores que se monitorean en el STI y que muestran la interacción de un estudiante				
<b>Tiempo Visualización Recurso</b>				
BASES DE DATOS	NA	NA	NA	
INTRODUCCION A LA INFORMATICA	31,96	156,17	188,13	
PROGRAMACION ORIENTADA A OBJETOS	60,23	112,77	173	



<b>Curso normal de los eventos</b>	
<b>Acción del Actor</b>	<b>Respuesta del Sistema</b>
1. El usuario decide guardar el área de trabajo dado clic sobre la opción guardar área de trabajo [1].	2. Presenta una ventana de dialogo [2] que facilita ubicar el lugar donde se quiere guardar el área de trabajo.
3. Selecciona la ubicación donde guardar el área de trabajo y da clic en el botón guardar de la ventana de dialogo [3].	4. Verifica que el área de trabajo no exista en esa ubicación
	5. Crea un archivo que representa el área de trabajo y presenta una ventana de confirmación [4].
<b>TRAZABILIDAD</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <a href="#">[Caso de uso Guardar área de trabajo]</a></li> </ul> </li> </ul>	

**Tabla 50.** Caso de uso real Guardar área de trabajo.



## CASO DE USO REAL ABRIR ÁREA DE TRABAJO

**Caso de uso real Guardar área de trabajo**

**Curso normal de los eventos**

Acción del Actor	Respuesta del Sistema
------------------	-----------------------



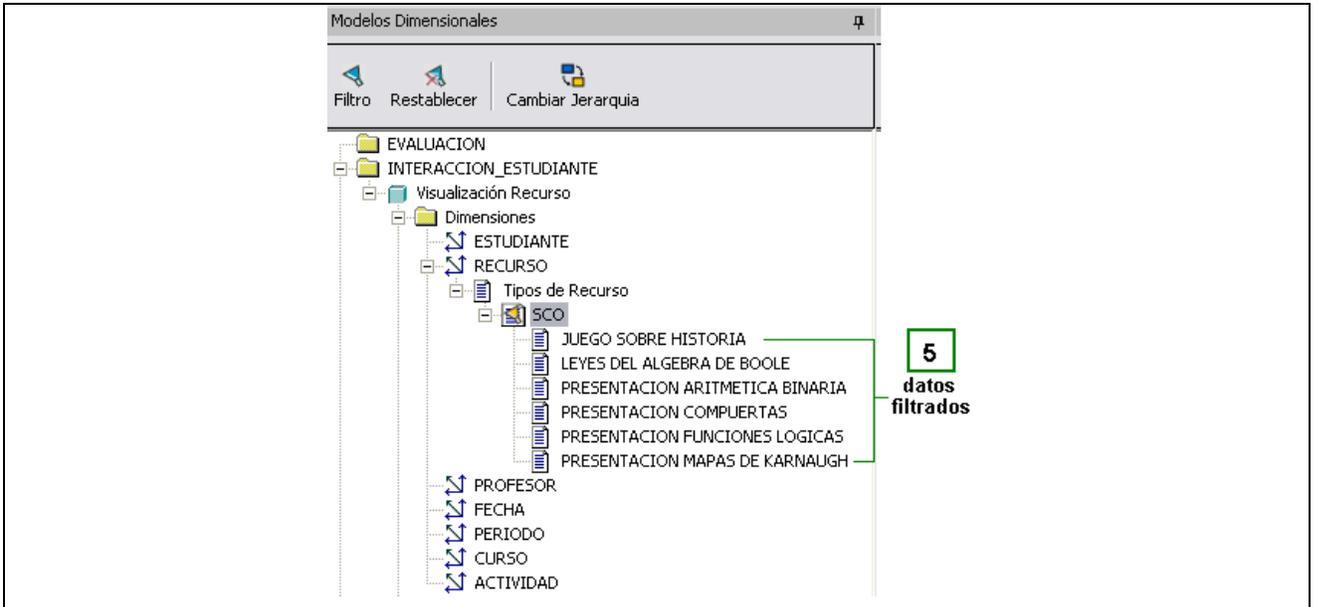
1. El usuario decide abrir un área de trabajo desde una ubicación determinada y para esto da clic en el botón abrir área de trabajo [1].	2. Presenta una ventana de dialogo que permite seleccionar la ubicación donde se encuentra el área de trabajo [2].
3. Selecciona la ubicación donde se encuentra el área de trabajo y da clic en el botón abrir [3].	4. Abre el área de trabajo desde la ubicación seleccionada.
	5. Lee la estructura del archivo que representa el área de trabajo.
	6. Presenta el área de trabajo guardada [4].
<b>TRAZABILIDAD</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ [Caso de uso Abrir área de trabajo]</li> </ul> </li> </ul>	

Tabla 51. Caso de uso real Abrir área de trabajo.

## CASO DE USO REAL FILTRAR NIVELES

### Caso de uso real Filtrar niveles

The screenshot shows the 'Modelos Dimensionales' interface. On the left, a tree view shows the hierarchy: EVALUACION > INTERACCION\_ESTUDIANTE > Visualización Recurso > Dimensiones > ESTUDIANTE > RECURSO > Tipos de Recurso > SCO. The 'SCO' level is highlighted with a green box labeled '1'. Below it, a list of items is shown, including 'FUNCIONES LOGICAS', 'JUEGO SOBRE HISTORIA', 'LEYES DEL ALGEBRA DE BOOLE', etc. A green box labeled 'datos de nivel' points to this list. On the right, the 'FILTRAR DATOS' dialog box is open, showing a list of items with checkboxes. The 'LEYES DEL ALGEBRA DE BOOLE' item is checked and highlighted with a green box labeled '3'. The dialog has a 'Filtrar' button highlighted with a green box labeled '4' and a 'Cancelar' button. A green box labeled '2' highlights the dialog box itself.



Curso normal de los eventos	
Acción del Actor	Respuesta del Sistema
1. El usuario decide filtrar los datos de nivel de un nivel de una jerarquía, para esto selecciona un nivel [1].	2. Carga y presenta un listado con los datos de nivel correspondientes al nivel seleccionado [2].
3. Selecciona los datos de nivel con los cuales quiere interactuar [3] y hace clic sobre el botón filtrar [4].	4. Presenta la jerarquía con los datos de nivel filtrados para el nivel seleccionado [5]
TRAZABILIDAD	
Artefactos Anteriores	Artefactos Posteriores
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b>  <a href="#">[CASO DE USO FILTRAR NIVELES</a>            ]</li> </ul>	

Tabla 52. Caso de uso real Filtrar niveles.



## CASO DE USO REAL RESTABLECER DATOS DE NIVEL

Caso de uso real Restablecer datos de nivel	
Curso normal de los eventos	
Acción del Actor	Respuesta del Sistema
<p>1. El usuario decide restablecer los datos de nivel [2] de un nivel que ha sido filtrado, para esto selecciona un nivel filtrado [1] cualquiera.</p>	<p>2. Presenta el nivel [3] con sus datos correspondientes en su forma original [4] (como se encontraban antes de ser filtrados).</p>
TRAZABILIDAD	
Artefactos Anteriores	Artefactos Posteriores
<ul style="list-style-type: none"> <li>▪ <b>Artefactos del Análisis:</b> <ul style="list-style-type: none"> <li>○ <b>[CASO DE USO RESTABLECER DATOS DE NIVEL]</b></li> </ul> </li> </ul>	

Tabla 53. Caso de uso real Restablecer datos de nivel.

## CASO DE USO REAL CAMBIAR JERARQUÍA

**Caso de uso real Cambiar jerarquía**

The screenshot shows the 'Modelos Dimensionales' application. On the left, a tree view shows the hierarchy: ESTILO\_DE\_APRENDIZAJE > INTERACCION\_ESTUDIANTE > Visualización Recurso > Dimensiones > ESTUDIANTE (highlighted with box 1). The 'ESTUDIANTE' dimension has sub-dimensions: GENEROS (FEMENINO, MASCULINO), RECURSO, PROFESOR, PERIODO, CURSO, ACTIVIDAD, and FECHA. On the right, a second view shows the same hierarchy with 'ESTUDIANTE' expanded to show 'Genero' (Femenino, Masculino) and 'Recurso' (Estudiante, Profesor, Fecha, Periodo, Curso, Actividad). A dialog box 'JERARQUIAS' is open, listing several hierarchy options. The selected option is 'JERARQUÍA GENERO ESTUDIANTE (Actual)' (highlighted with box 3). The 'Seleccionar' button (highlighted with box 4) is at the bottom of the dialog.

The screenshot shows the 'Modelos Dimensionales' application after the hierarchy change. The tree view on the left now shows: ESTILO\_DE\_APRENDIZAJE > INTERACCION\_ESTUDIANTE > Visualización Recurso > Dimensiones > ESTUDIANTE (highlighted with box 5). The 'ESTUDIANTE' dimension has sub-dimensions: PAISES (COLOMBIA, CAUCA, VALLE), RECURSO, PROFESOR, PERIODO, CURSO, ACTIVIDAD, and FECHA. The 'COLOMBIA' dimension is highlighted.

### Curso normal de los eventos

Acción del Actor	Respuesta del Sistema
------------------	-----------------------



1. El usuario decide cambiar la jerarquía de una dimensión seleccionada [1].	2. Carga y presenta un listado con las jerarquías definidas para la dimensión seleccionada [2].
3. Selecciona la jerarquía [3] con la que quiere realizar análisis para la dimensión seleccionada dando clic en el botón seleccionar [4].	4. Presenta la dimensión con la jerarquía seleccionada [5].
<b>TRAZABILIDAD</b>	
<b>Artefactos Anteriores</b>	<b>Artefactos Posteriores</b>
▪ <b><u>Artefactos del Análisis:</u></b> <ul style="list-style-type: none"><li>○ <b><u>[CASO DE USO CAMBIAR JERARQUÍA]</u></b></li></ul>	

Tabla 54. Caso de uso real Cambiar jerarquía.



## ANEXO 12. VISTAS DEL SISTEMAS PARA EL MANEJO DEL CATALOGO OLAP EN ORACLE

- **ALL\$OLAP2UCATALOG\_ENTITY\_USES**

Sinónimo: ALL\_OLAP2\_CATALOG\_ENTITY\_USES

Cada fila representa una entidad dentro de un carpeta de medidas OLAP (catalogo). Las medidas son las únicas entidades de metadatos OLAP que pueden ser recopiladas en carpetas.

El sistema de clasificación, usado para administrar las carpetas de medidas y clasificar varias entidades de metadatos OLAP, es implementado en el catalogo OLAP1 y referenciado desde el catalogo OLAP2. Ambos catálogos son desplegados por esta vista.

El termino **Catalogo** cuando se usa en el contexto del sistema de clasificación de metadatos OLAP, se refiere a carpetas de medidas. Este no debe ser confundido con el término **Catalogo OLAP**, el cual se refiere a la colección de tablas que implementan el modelo de metadatos OLAP.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
CATALOG_ID	NUMBER	NOT NULL	Identificador de la carpeta de medidas.
ENTITY_OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo de medidas.
ENTITY_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo de medidas.
CHILD_ENTITY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la medida dentro de la carpeta.

Tabla 55. Descripción columnas vista ALL\$OLAP2UCATALOG\_ENTITY\_USES.

- **ALL\$OLAP2UCATALOGS**

Sinónimo: ALL\_OLAP2\_CATALOGS

Cada fila representa una carpeta de medidas OLAP (catalogo). Las carpetas de medidas son un medio de agrupación de medidas relacionadas para un área del negocio (data mart). Por ejemplo, todas las medidas que almacenan información acerca de una línea de producto dado podrían ser recopiladas en una carpeta de medidas.



Las carpetas de medidas son **esquemas independientes**. Todos los usuarios pueden ver todas las carpetas de medidas definidas en la base de datos, aun si ellos no tienen privilegios de acceso para las medidas dentro de las carpetas.

Las carpetas de medidas pueden ser anidadas dentro de otras carpetas de medidas.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
CATALOG_ID	NUMBER	NOT NULL	Identificador de la carpeta de medidas.
CATALOG_NAME	VARCHAR2(30)	NOT NULL	Nombre de la carpeta de medidas.
PARENT_CATALOG_ID	NUMBER		Identificador de la carpeta de medidas padre. Esta columna es NULL para carpetas que no tengan raíz como carpeta padre.
DESCRIPTION	VARCHAR2(2000)		Descripción de la carpeta de medidas.

Tabla 56. Descripción columnas vista ALL\$OLAP2UCATALOGS.

- **ALL\$OLAP2UCUBE\_DIM\_USES**

Sinónimo: ALL\_OLAP2\_CUBE\_DIM\_USES

Cada fila representa una asociación entre un cubo y una dimensión. Una dimensión puede ser asociada más de una vez en el mismo cubo, pero cada asociación es especificada en una fila separada, bajo su propio alias de dimensión único.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
CUBE_DIMENSION_USE_ID	NUMBER	NOT NULL	Identificador de la asociación entre un cubo y una dimensión.
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo.
DIMENSION_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
DIMENSION_ALIAS	VARCHAR2(30)		Alias de la dimensión, provee identificación única de la dimensión usada dentro del cubo.
DEFAULT_CALC_HIERARCHY_NAME	VARCHAR2(30)		La jerarquía por defecto



			usada para “drilling down” o bajar dentro de la dimensión.
DEPENDENT_ON_DIM_USE_ID	NUMBER		ID de la asociación cubo/dimensión (solamente para el catalogo OLAP1).

**Tabla 57.** Descripción columnas vista ALL\$OLAP2UCUBE\_DIM\_USES.

- **ALL\$OLAP2UCUBE\_MEAS\_DIM\_USES**

Sinonimo: ALL\_OLAP2\_CUBE\_MEAS\_DIM\_USES

Cada fila representa la asociación de una medida con una de sus dimensiones, y especifica como los datos de medida pueden ser agregados sobre la dimensión. Si ningún método de agregación es especificado, los datos se adicionan (SUM).

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo que define esta medida.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo que define esta medida.
MEASURE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la medida.
DIMENSION_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la asociación dimensión - medida.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
DIMENSION_ALIAS	VARCHAR2(30)		Alias de la dimensión.
DEFAULT_AGGR_FUNCTION_USE_ID	NUMBER		Identificador del método de agregación por defecto usado para agregar el dato de esta medida sobre esta dimensión. Si esta columna es nula, el método de agregación es adición (SUM).

**Tabla 58.** Descripción columnas vista ALL\$OLAP2UCUBE\_MEAS\_DIM\_USES.

- **ALL\$OLAP2UCUBE\_MEASURE\_MAPS**

Sinonimo: ALL\_OLAP2\_CUBE\_MEASURE\_MAPS



Cada fila representa el mapeo de una medida a la columna en una tabla de hechos.

En el catalogo OLAP2, las medidas son mapeadas separadamente por cada combinación de jerarquías de dimensión. Por ejemplo, si una medida tiene tres dimensiones y cada dimensión tiene dos jerarquías, entonces la medida tiene ocho mapeos separados a la tabla de hechos. Estas ocho columnas deben existir dentro de la misma tabla de hechos o en tablas separadas. Cada una de las ocho columnas mapeadas debe aparecer como una fila separada en la vista.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo.
MEASURE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la medida definida para ese cubo.
DIM_HIER_COMBO_ID	NUMBER	NOT NULL	Identificador de la asociación entre la medida y una combinación de sus jerarquías de dimensión. (Solamente en el Release 2 de metadatos)
FACT_TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla de hechos.
FACT_TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla de hechos.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna en la tabla de hechos donde este dato de medida es almacenado. En el Release 2 de metadatos, el dato de medida es para una combinación de sus jerarquías.

**Tabla 59.** Descripción columnas vista ALL\$OLAP2UCUBE\_MEASURE\_MAPS.

- **ALL\$OLAP2UCUBE\_MEASURES**

Sinonimo: ALL\_OLAP2\_CUBE\_MEASURES

Cada fila representa una medida.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo que define la medida.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo que define la medida.



MEASURE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la medida.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue para la medida.
DESCRIPTION	VARCHAR2(2000)		Descripción de la medida.

**Tabla 60.** Descripción columnas vista ALL\$OLAP2UCUBE\_MEASURES.

- **ALL\$OLAP2UCUBES**

Sinonimo: ALL\_OLAP2\_CUBES

Cada fila representa un cubo.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo.
INVALID	VARCHAR2(1)	NOT NULL	Si el cubo esta o no esta en un estado invalido. Un cubo es valido si este tiene al menos una dimensión, todas sus dimensiones son validas, y todos los mapeos de las tablas de hechos son validos.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue para el cubo.
DESCRIPTION	VARCHAR2(2000)		Descripción del cubo. Si este cubo tiene una vista materializada en el Catalogo OLAP2, el código de resumen de la vista materializada especifica si este esta en la forma GS (Conjunto agrupado) o RU(Rolled Up)
MV_SUMMARYCODE	VARCHAR2(2)		La forma RU significa que todas las columnas llaves de la dimensión son pobladas, y los datos pueden solo ser accedidos cuando su linaje completo es especificado.  La forma GS significa que las columnas llaves de la dimensión contienen valores nulos, y los datos pueden ser accedidos simplemente especificando uno o más niveles.



**Tabla 61.** Descripción columnas vista ALL\$OLAP2UCUBES.

- **ALL\$OLAP2UDIM\_ATTR\_USES**

Sinonimo: ALL\_OLAP2\_DIM\_ATTR\_USES

Cada fila representa la asociación de un atributo nivel con un atributo dimensión. Un atributo dimensión esta compuesto de un conjunto de atributos nivel. El mismo atributo nivel puede ser incluido en mas de un atributo dimensión.

Un atributo nivel designa una columna en una tabla dimensión que almacena información descriptiva acerca de un nivel en la dimensión. Por ejemplo, podría haber un atributo color que pertenece a un nivel ID de producto.

Un atributo de dimensión es una colección de atributos nivel. Por ejemplo, el atributo dimensión TIME\_SPAN almacena el número de días asociados con cada periodo de tiempo en una dimensión tiempo. Los periodos de tiempo son definidos como niveles, y cada nivel tiene su propio atributo nivel TIME\_SPAN asociado, pero todos los atributos nivel TIME\_SPAN (por ejemplo, MONTH\_TIME\_SPAN, QUARTER\_TIME\_SPAN y YEAR\_TIME\_SPAN) están definidos como un simple atributo dimensión TIME\_SPAN.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
DIM_ATTRIBUTE_NAME	VARCHAR2(30)	NOT NULL	Nombre del atributo dimension.
LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre de un nivel dentro de la dimensión.
LVL_ATTRIBUTE_NAME	VARCHAR2(30)	NOT NULL	Nombre de una atributo para este nivel. Este atributo nivel esta incluido en el atributo dimensión.

**Tabla 62.** Descripción columnas vista ALL\$OLAP2UDIM\_ATTR\_USES.

- **ALL\$OLAP2UDIM\_ATTRIBUTES**

Sinonimo: ALL\_OLAP2\_DIM\_ATTRIBUTES

Cada fila representa un atributo de dimensión que es un atributo lógico el cual proporciona una agrupación de atributos de nivel dentro de la dimensión. Los atributos nivel dentro de la agrupación de atributos dimensión pueden ser determinados desde la vista ALL\_OLAP2\_DIM\_ATTR\_USES.



COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
ATTRIBUTE_NAME	VARCHAR2(30)	NOT NULL	Nombre del atributo nivelado.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue para el atributo dimensión.
DESCRIPTION	VARCHAR2(2000)		Descripción del atributo dimensión.
DESC_ID	NUMBER		Una clasificación para el atributo dimensión. Si un atributo dimensión es clasificado, este puede ser del tipo LONG_DESCRIPTION, SHORT_DESCRIPTION, END_DATE, TIME_SPAN, PRIOR_PERIOD, or YEAR_AGO_PERIOD.

**Tabla 63.** Descripción columnas vista ALL\$OLAP2UDIM\_ATTRIBUTES.

- **ALL\$OLAP2UDIM\_HIER\_LEVEL\_USES**

Sinonimo: ALL\_OLAP2\_DIM\_HIER\_LEVEL\_USES

Cada fila representa una relación jerárquica entre dos niveles en una jerarquía de dimensión. Dentro de jerarquías separadas, el mismo nivel padre puede ser jerárquicamente relacionado a diferentes niveles hijo.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la jerarquía.
PARENT_LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre del nivel padre.
CHILD_LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre del nivel hijo.
POSITION	NUMBER	NOT NULL	Posición de esta relación Padre-Hijo dentro de la jerarquía, con la posición uno (1) como la mas detallada.

**Tabla 64.** Descripción columnas vista ALL\$OLAP2UDIM\_HIER\_LEVEL\_USES.



- **ALL\$OLAP2UDIM\_HIERARCHIES**

Sinonimo: ALL\_OLAP2\_DIM\_HIERARCHIES

Cada fila representa una jerarquía de dimensión. Las relaciones entre niveles por cada jerarquía son representadas por ALL\_OLAP2\_DIM\_HIER\_LEVEL\_USES.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la jerarquía.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue de la jerarquía.
DESCRIPTION	VARCHAR2(2000)		Descripción de la jerarquía.
SOLVED_CODE	VARCHAR2(2)	NOT NULL	El código podría ser uno de los siguientes:  UNSOLVED LEVEL-BASED, para una jerarquía que no contiene totales no embebidos y son almacenados en tablas dimensión de Nivel-Base. Jerarquías de versión 1 son siempre de este tipo.  SOLVED LEVEL-BASED, para jerarquías que contienen totales embebidos, tienen un ID de agrupación, y son almacenadas en tablas dimensión de Nivel-Base. (Solo jerarquías de versión 2)  SOLVED VALUE-BASED, para una jerarquía que contiene totales embebidos para todas las combinaciones de niveles y son almacenados en una tabla dimensión Padre/Hijo. (Solo jerarquías de versión 2).

Tabla 65. Descripción columnas vista ALL\$OLAP2UDIM\_HIERARCHIES.

- **ALL\$OLAP2UDIM\_LEVEL\_ATTR\_MAPS**

Sinonimo: ALL\_OLAP2\_DIM\_LEVEL\_ATTR\_MAPS

Cada fila representa el mapeo de un atributo nivel a su nivel asociado.



Cada nivel mapea a una o más columnas en una tabla dimensión. Cada atributo nivel mapea a una simple columna en la misma tabla dimensión con su nivel asociado.

El mapeo de atributos nivel a niveles es dependiente de la jerarquía. El mismo nivel podría tener diferentes atributos cuando este es usado en diferentes jerarquías.

Todos los niveles definidos como metadatos OLAP son representados por la vista ALL\_OLAP2\_DIM\_LEVELS.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)		Nombre de la jerarquía que contiene este nivel.
ATTRIBUTE_NAME	VARCHAR2(30)		El nombre de una agrupación de atributos de dimensión que contiene este atributo nivel.
LVL_ATTRIBUTE_NAME	VARCHAR2(30)	NOT NULL	Nombre del atributo nivel, o nombre de la columna si el nombre del nivel no es especificado.
LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre del nivel.
TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla dimensión que contiene el nivel y el atributo nivel.
TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla dimensión que contiene el nivel y el atributo nivel.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna que contiene el atributo nivel.
DTYPE	VARCHAR2(10)	NOT NULL	Tipo de dato de la columna que contiene el atributo nivel.

**Tabla 66.** Descripción columnas vista ALL\$OLAP2UDIM\_LEVEL\_ATTR\_MAPS.

- **ALL\$OLAP2UDIM\_LEVEL\_ATTRIBUTES**

Sinonimo: ALL\_OLAP2\_DIM\_LEVEL\_ATTRIBUTES

Cada fila representa un atributo nivel. Cada atributo nivel es una columna en una tabla dimensión. La columna almacena información descriptiva acerca de un nivel definido dentro de la misma tabla dimensión. Si el nivel atributo no es nombrado, se usa el nombre de la columna.



El mapeo de la columna del atributo nivel asociado a la columna nivel se describe en ALL\_OLAP2\_DIM\_LEVEL\_ATTR\_MAPS.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión que contiene el atributo nivel.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión que contiene el atributo nivel.
ATTRIBUTE_NAME	VARCHAR2(30)		Nombre del atributo nivel. Si el nombre del atributo no es especificado, se usa el nombre de la columna.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue para el atributo nivel.
DESCRIPTION	VARCHAR2(2000)		Descripción del atributo nivel.
DETERMINED_BY_LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre del nivel al cual este atributo nivel es mapeado.

**Tabla 67.** Descripción columnas vista ALL\$OLAP2UDIM\_LEVEL\_ATTRIBUTES.

- **ALL\$OLAP2UDIM\_LEVELS**

Sinonimo: ALL\_OLAP2\_DIM\_LEVELS

Cada fila representa un nivel dentro de una dimensión. Un nivel es mapeado a uno o más columnas dentro de una tabla dimensión. En un esquema estrella, todos los niveles de dimensión son mapeados a columnas dentro de la misma tabla. En un esquema “snowflake”, los niveles dimensión son mapeados a columnas en tablas separadas.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	El dueño de la dimensión que contiene este nivel.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión que contiene este nivel
LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre del nivel.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue del nivel.
DESCRIPTION	VARCHAR2(2000)		Descripción del nivel.
LEVEL_TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla dimensión que contiene las columnas de este nivel.
LEVEL_TABLE_NAME	VARCHAR2(30)	NOT	Nombre de la tabla dimensión que



COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
		NULL	contiene las columnas para este nivel.

Tabla 68. Descripción columnas vista ALL\$OLAP2UDIM\_LEVELS.

- ALL\$OLAP2UDIMENSIONS**

Sinonimo: ALL\_OLAP2\_DIMENSIONS

Cada fila representa una dimensión. En metadatos OLAP1, las dimensiones son basadas sobre objetos dimensión Oracle. En metadatos OLAP2, las dimensiones son completamente independientes de objetos dimensión Oracle.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
PLURAL_NAME	VARCHAR2(30)		El nombre plural para la dimensión. Usado para despliegue.
DISPLAY_NAME	VARCHAR2(30)		Nombre de despliegue para la dimensión.
DESCRIPTION	VARCHAR2(2000)		Descripción de la dimensión.
DEFAULT_DISPLAY_HIERARCHY	VARCHAR2(30)	NOT NULL	Jerarquía de despliegue por defecto para la dimensión.
INVALID	VARCHAR2(1)	NOT NULL	Si la dimensión es o no es valida. Una dimensión es valida cuando todos sus niveles y atributos nivel son mapeados a columnas existentes, y los atributos dimensión son definidos con conjunto de atributos nivel validos, y las jerarquías son definidas con niveles validos.
DIMENSION_TYPE	VARCHAR2(10)		No usado.

Tabla 69. Descripción columnas vista ALL\$OLAP2UDIMENSIONS.

- ALL\$OLAP2UENTITY\_DESC\_USES**



Sinonimo: ALL\_OLAP2\_ENTITY\_DESC\_USES

Cada fila representa una asociación entre una entidad metadata OLAP y su descriptor. Las entidades de metadata OLAP y descriptores son definidas en el sistema de clasificación de Catalogo OLAP1 y referenciadas desde el Catalogo OLAP2.

Las siguientes entidades de metadatos OLAP son representados en esta vista:

- Dimensiones cuyos descriptores son tiempo.
- Atributos dimensión cuyos descriptores son: *Long Description, Short Description, or Description*.
- Atributos dimensión (solo para dimensiones tiempo) cuyos descriptores son: *End Date, Time Span, Prior Period, or Year Ago Period*.
- Atributos nivel (solo para dimensión tiempo) cuyos descriptores son: *Day, Month, Quarter, or Year*.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
DESCRIPTOR_ID	NUMBER	NOT NULL	ID del descriptor, derivado desde el sistema de clasificación OLAP1.
ENTITY_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la entidad. Si la entidad es un atributo dimensión o atributo nivel, el propietario es el propietario de la dimensión.
ENTITY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la entidad. La entidad podría ser una dimensión, un atributo dimensión, o un atributo nivel.
CHILD_ENTITY_NAME	VARCHAR2(30)		Nombre de la entidad hijo (si aplica). Un atributo dimensión es una entidad hijo de una dimensión. Un atributo nivel es una entidad hijo de un atributo dimensión
SECONDARY_CHILD_ENTITY_NAME	VARCHAR2(30)		Nombre de la entidad del hijo secundario (si aplica). Un atributo dimensión es una entidad hijo de una dimensión. Un atributo nivel es una entidad hijo de un atributo dimensión. Un



			atributo nivel podría ser la entidad hijo secundario de una dimensión.
--	--	--	------------------------------------------------------------------------

**Tabla 70.** Descripción columnas vista ALL\$OLAP2UENTITY\_DESC\_USES.

- **ALL\$OLAP2UFACT\_LEVEL\_USES**

Sinonimo: ALL\_OLAP2\_FACT\_LEVEL\_USES

Cada fila representa una relación unida entre una tabla de hechos y una tabla dimensión. La relación unida es derivada desde una simple columna llave en la tabla de hechos.

En los metadatos OLAP2, la tabla de hechos es siempre mapeada en el contexto de una jerarquía dimensión específica.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo.
DIMENSION_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	NUMBER	NOT NULL	Nombre de la dimensión.
DIMENSION_ALIAS	VARCHAR2(30)		Alias de la dimensión (si aplica).
HIERARCHY_NAME		NOT NULL	Nombre de la jerarquía.
DIM_HIER_COMBO_ID	NUMBER	NOT NULL	ID de la combinación jerarquía de dimensión asociada con esta tabla de hechos. (Solo metadatos version2)
LEVEL_NAME	VARCHAR2(30)		Nombre del nivel dentro de la jerarquía donde ocurre el mapeo. Esto representa el nivel mas bajo de agregación definido por la llave foránea/primaria para una dimensión específica del cubo. En la versión 1 de los metadatos, este siempre es el nivel de la hoja (todas las jerarquías de la dimensión comparte el mismo nivel hoja).
FACT_TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla de hechos.



COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
FACT_TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla de hechos.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna de llave foránea en la tabla de hechos.
POSITION	NUMBER		Posición de esta columna dentro de una llave multi-columna.
DIMENSION_KEYMAP_TYPE	VARCHAR2(30)	NOT NULL	<p>Tipo de llave mapeada por la tabla de hechos. Los valores podrían ser:</p> <p>LL (Lowest Level), cuando se guardan sólo miembros de dimensión de bajo-nivel en la columna clave. Para versiones de metadatos 1, el tipo de mapeo es siempre LL.</p> <p>ET (Embedded Totals), cuando los miembros de dimensión para todas las combinaciones nivel son almacenados en la columna clave. Si la a tabla de hechos es resuelta (contiene totales embebidos para todas las combinaciones de niveles). (Solo para versiones 2 de metadatos).</p> <p>RU (Rolled Up), cuando los miembros de una dimensión para cada nivel son almacenados en columnas claves separadas (llave multi-columna). (Solo para versiones 2 de metadatos).</p>
FOREIGN_KEY_NAME	VARCHAR2(30)		Nombre de la restricción de llave foránea (Solo Catalogo OLAP1) aplicado a la columna de llave foránea. Las restricciones no son usadas en el Catalogo OLAP2.

**Tabla 71.** Descripción columnas vista ALL\$OLAP2UFACT\_LEVEL\_USES.



- **ALL\$OLAP2UFACT\_TABLE\_GID**

Sinonimo: ALL\_OLAP2\_FACT\_TABLE\_GID

Cada fila representa información de una columna que agrupa ID (GID) en la tabla de hechos.

En este caso de solución, la tabla de hechos total embebida (donde el total para cada combinación de niveles esta incluida en la tabla), hay una columna GID correspondiente a cada dimensión clave en la tabla de hechos. Por ejemplo, una tabla de hechos total embebida para datos de ventas, la cual es dimensionada por Producto, Geografía y Tiempo, tiene tres columnas GID (una por cada de las dimensiones).

**Nota:** La vista solo pertenece a metadatos OLAP2 (CWM2). Columnas GID no son soportadas por metadatos OLAP1 (CWM).

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario del cubo.
CUBE_NAME	VARCHAR2(30)	NOT NULL	Nombre del cubo.
DIMENSION_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la jerarquía.
DIM_HIER_COMBO_ID	NUMBER	NOT NULL	ID de la asociación Dimensión - Jerarquía.
FACT_TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla de hechos.
FACT_TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla de hechos.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna GDI.

**Tabla 72.** Descripción columnas vista ALL\$OLAP2UFACT\_TABLE\_GID.

- **ALL\$OLAP2UHIER\_CUSTOM\_SORT**

Sinonimo: ALL\_OLAP2\_HIER\_CUSTOM\_SORT

Cada fila proporciona la información sobre la ordenación específica para una jerarquía de una dimensión dada. Esta información de la ordenación específica es opcional.



La información de ordenación indica como ordenar los miembros de una jerarquía basados en columnas de tabla dimensión asociada.

La información de ordenación puede especificar que las columnas pueden ser ordenadas de forma ascendente o descendente, con nulos en primera posición o nulos al final. El ordenamiento puede ser aplicado en múltiples niveles de una dimensión.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la jerarquía.
TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla dimensión que contiene la columna a ser ordenada.
TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla dimensión que contiene la columna a ser ordenada.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna a ser ordenada.
POSITION	NUMBER	NOT NULL	Representa la posición dentro de una multi-columna SORT_POSITION. En muchos casos, una simple columna representa SORT_POSITION, y el valor de POSITION es 1.
SORT_POSITION	NUMBER	NOT NULL	Posición dentro del orden de clasificación del nivel a ser ordenado.
SORT_ORDER	VARCHAR2(4)	NOT NULL	Orden de clasificación. Puede ser ascendente o descendente.
NULL_ORDER	VARCHAR2(5)	NOT NULL	Donde insertar valores nulos en el orden de clasificación. Puede ser nulos primero o nulos al final.

**Tabla 73.** Descripción columnas vista ALL\$OLAP2UHIER\_CUSTOM\_SORT.

- **ALL\$OLAP2UJOIN\_KEY\_COLUMN\_USES**

Sinonimo: ALL\_OLAP2\_JOIN\_KEY\_COLUMN\_USES.

Cada fila representa la información clave que une dos niveles en una jerarquía. Si el nivel es mapeado a más de una columna, cada columna mapeada se representa en una fila separada.



En un esquema “snowflake”, donde los niveles son definidos en tablas dimensiones separadas, los niveles en una jerarquía tienen una relación de llave foránea lógica. En un esquema estrella, donde los niveles definidos dentro de la misma tabla dimensión, la clave del nivel hijo especifica su posición en la jerarquía.

COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)	NOT NULL	Nombre de la jerarquía.
CHILD_LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nivel hijo en la jerarquía.
TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla dimensión.
TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla dimensión.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna nivel hijo en la tabla dimensión. En un esquema estrella, esta es la columna asociada con CHILD_LEVEL_NAME. En un esquema Snowflake, esta es la columna padre de CHILD_LEVEL_NAME en la misma tabla dimensión
POSITION	NUMBER		Posición de la columna dentro de la clave. Aplica a solo llaves multi-columna (donde el nivel es mapeado a más de una columna).
JOIN_KEY_TYPE	VARCHAR2(30)	NOT NULL	La clave es de tipo SNOWFLAKE si la clave unida es una llave foránea lógica. La llave es de tipo STAR si la llave unida se refiere a una columna dentro de la misma tabla.

**Tabla 74.** Descripción columnas vista ALL\$OLAP2UJOIN\_KEY\_COLUMN\_USES.

- **ALL\$OLAP2ULEVEL\_KEY\_COLUMN\_USES**

Sinonimo: ALL\_OLAP2\_LEVEL\_KEY\_COL\_USES

Cada fila representa la llave lógica enlazando un nivel de dimensión a su columna subyacente en una tabla dimensión. Si el nivel es mapeado a más de una columna, cada columna mapeada se representa en una fila separada de la vista.



COLUMNA	TIPO DE DATO	NULL	DESCRIPCIÓN
OWNER	VARCHAR2(30)	NOT NULL	Propietario de la dimensión.
DIMENSION_NAME	VARCHAR2(30)	NOT NULL	Nombre de la dimensión.
HIERARCHY_NAME	VARCHAR2(30)		Nombre de la jerarquía que incluye este nivel.
CHILD_LEVEL_NAME	VARCHAR2(30)	NOT NULL	Nombre del nivel.
TABLE_OWNER	VARCHAR2(30)	NOT NULL	Propietario de la tabla dimensión.
TABLE_NAME	VARCHAR2(30)	NOT NULL	Nombre de la tabla dimensión.
COLUMN_NAME	VARCHAR2(30)	NOT NULL	Nombre de la columna que almacena CHILD_LEVEL_NAME.
POSITION	NUMBER		Posición de la columna dentro de la llave. Aplica solo a llaves multi-columna (donde el nivel es mapeado a más de una columna).

**Tabla 75.** Descripción columnas vista ALL\$OLAP2ULEVEL\_KEY\_COLUMN\_USES.



## **ANEXO 13. EXTENSIONES SQL PARA BODEGAS DE DATOS RELACIONALES EN ORACLE**

La agregación es una parte fundamental de la bodega de datos. Oracle ofrece las siguientes extensiones para el manejo de esta:

- Extensiones CUBE y ROLLUP para la cláusula GROUP BY.
- Tres funciones GROUPING.
- La expresión GROUPING SETS.

ROLLUP calcula agregaciones como SUM, COUNT, MAX, MIN y AVG en niveles crecientes de agregación, desde el más detallado hasta el gran total. CUBE es una extensión similar a ROLLUP, permitiendo una simple declaración para calcular todas las posibles combinaciones de agregaciones. CUBE, ROLLUP y GROUPING producen un conjunto simple de resultado que es equivalente a un UNION ALL de diferentes filas de grupos.

Las tres funciones de GROUPING ayudan a identificar el grupo al que pertenece cada fila y permite ordenar subtotales de filas y filtrar resultados.

### **HACIENDO ANALISIS A TRAVÉS DE MULTIPLES DIMENSIONES**

Uno de los conceptos claves en los sistemas de soporte a la toma de decisiones es el análisis multidimensional, es decir, examinar la organización desde todas las combinaciones dimensionales. Se usa el termino “dimensión” para dar significado a cada categoría usada en preguntas específicas. Entre las dimensiones mas comunes están tiempo, geografía, producto, departamento y canal de distribución, pero las dimensiones potenciales pueden ser una variedad interminable, como las actividades de la organización. Los eventos o entidades asociados con un conjunto particular de valores dimensionales son usualmente llamados “hechos”. Los hechos podrían ser ventas en unidades, ganancia, cuentas de clientes, volumen de producción o cualquier otro valor de rastreo.

Algunos ejemplos de preguntas multidimensionales son:

- Mostrar las ventas totales a través de todos los productos, agregando en los niveles para una dimensión geografía, desde un estado a un país y a una región; para el año 1999 y 2000.
- Crear un análisis “cross-tabular” de nuestras operaciones, mostrando gastos por territorio en Sur América para 1999 y 2000. Incluyendo todos los posibles subtotales.
- Listar las 10 ventas mas representativas en Asia de acuerdo al promedio vendido en el 2000 para productos de automotor.



Todas estas peticiones involucran múltiples dimensiones. Muchas preguntas multidimensionales requieren datos agregados y comparación de conjunto de datos, a menudo a través del tiempo, geografía o presupuestos.

## UN ESCENARIO DE AGREGACION

Para ilustrar el uso de la extensión GROUP BY se presenta un ejemplo donde una compañía hipotética tiene ventas a través del mundo y rastrea las ventas por dólares y cantidad de información. Debido a que hay muchas filas de datos, las consultas mostradas típicamente tienen restricciones fuertes sobre su cláusula WHERE para limitar los resultados a un pequeño número de filas.

La siguiente tabla es un ejemplo de un reporte “cross-tabular” mostrando el total de ventas para “country\_id” y “channel\_desc” para Estados Unidos y El Reino Unido a través de la Internet y Ventas directas en Septiembre de 2000.

Channel	Country		
	UK	US	Total
Direct Sales	1,378,126	2,835,557	4,213,683
Internet	911,739	1,732,240	2,643,979
Total	2,289,865	4,567,797	6,857,662

**Tabla 76.** Ejemplo reporte cross- tabular

Considere que aun un reporte simple como el anterior, con justamente nueve (9) valores en esta grilla, genera cuatro (4) subtotales y un gran total. La mitad de los valores requeridos para este reporte no serán calculados con una sola consulta que solicita SUM (amount\_sold) y ejecuta GROUP BY(channel\_desc, country\_id). Para obtener agregaciones de alto nivel se podrían requerir consultas adicionales. Comandos de la base de datos que ofrecen mejoramiento en cálculos y subtotales trae mejor beneficio a las consultas, reportes y operaciones de análisis.



```
SELECT channel_desc, country_id,  
       TO_CHAR(SUM(amount_sold), '9,999,999,999') SALES$  
FROM sales, customers, times, channels  
WHERE sales.time_id=times.time_id AND  
       sales.cust_id=customers.cust_id AND  
       sales.channel_id= channels.channel_id AND  
       channels.channel_desc IN ('Direct Sales', 'Internet') AND  
       times.calendar_month_desc='2000-09'  
       AND country_id IN ('UK', 'US')  
GROUP BY CUBE(channel_desc, country_id);
```

CHANNEL_DESC	CO	SALES\$
-----	--	-----
Direct Sales	UK	1,378,126
Direct Sales	US	2,835,557
Direct Sales		4,213,683
Internet	UK	911,739
Internet	US	1,732,240
Internet		2,643,979
	UK	2,289,865
	US	4,567,797

## INTERPRETANDO VALORES NULOS EN EL EJEMPLO

Los nulos retornados por las extensiones del GROUP BY no son siempre el significado de los nulos tradicionales con valores desconocidos. En lugar de eso, un nulo podría indicar que esta fila es un subtotal. Para evitar introducir otro valor en el sistema de la base de datos, estos valores de subtotales no tienen una etiqueta especial.

## EXTENSION ROLLUP PARA GROUP BY

ROLLUP permite a una sentencia SELECT calcular múltiples niveles de subtotales a través de grupos específicos de dimensiones y además, calcula un gran total. ROLLUP es una simple extensión a la cláusula GROUP BY, por lo tanto, es considerada de fácil uso.

La acción de ROLLUP es simple: crea subtotales que van desde el nivel mas detallado a un gran total, siguiendo una lista de agrupación especificada en la cláusula ROLLUP. ROLLUP toma como argumentos una lista ordenada de columnas agrupadas. Primero, calcula los valores estándar agregados que se especifican en la cláusula GROUP BY. Luego, crea progresivamente subtotales de alto nivel, moviéndose desde la derecha hacia la izquierda a través de la lista de columnas agrupadas. Finalmente, crea un gran total.



ROLLUP crea subtotales en  $n+1$  niveles, donde  $n$  es el número de columnas agrupadas. Por ejemplo, si una consulta especifica ROLLUP sobre columnas agrupadas de tiempo, región y departamento ( $n=3$ ), el conjunto resultado incluirá filas en cuatro (4) niveles de agregación.

### **Cuando usar ROLLUP**

La extensión ROLLUP es conveniente en tareas que involucran subtotales.

- Es muy útil para subtotales a lo largo de dimensiones jerárquicas tales como tiempo o geografía. Por ejemplo, una consulta podría especificar un ROLLUP (year, month, day) o ROLLUP (country, state, city).
- Para administradores de bodegas de datos que usan tablas resumidas, ROLLUP pueden simplificar y acelerar el mantenimiento de tablas resumidas.

### **Sintaxis ROLLUP**

ROLLUP aparece en la cláusula GROUP BY en una sentencia SELECT de la siguiente forma:

```
SELECT ... GROUP BY ROLLUP(grouping_column_reference_list)
```

### **Ejemplo – ROLLUP.**

Este ejemplo usa los datos almacenados en el esquema “sh” de Oracle.

```
SELECT channel_desc, calendar_month_desc, country_id,  
       TO_CHAR(SUM(amount_sold), '9,999,999,999') SALES$  
FROM sales, customers, times, channels  
WHERE sales.time_id=times.time_id AND  
       sales.cust_id=customers.cust_id AND  
       sales.channel_id= channels.channel_id AND  
       channels.channel_desc IN ('Direct Sales', 'Internet') AND  
       times.calendar_month_desc IN ('2000-09', '2000-10')  
       AND country_id IN ('UK', 'US')  
GROUP BY ROLLUP(channel_desc, calendar_month_desc, country_id);
```



CHANNEL_DESC	CALENDAR	CO	SALES\$
Direct Sales	2000-09	UK	1,378,126
Direct Sales	2000-09	US	2,835,557
Direct Sales	2000-09		4,213,683
Direct Sales	2000-10	UK	1,388,051
Direct Sales	2000-10	US	2,908,706
Direct Sales	2000-10		4,296,757
Direct Sales			8,510,440
Internet	2000-09	UK	911,739
Internet	2000-09	US	1,732,240
Internet	2000-09		2,643,979
Internet	2000-10	UK	876,571
Internet	2000-10	US	1,893,753
Internet	2000-10		2,770,324
Internet			5,414,303
			13,924,743

Esta consulta retorna el siguiente conjunto de filas:

- Filas de agregaciones regulares que podrían ser producidas por GROUP BY sin usar ROLLUP.
- Subtotales de primer nivel de agregación a través de country\_id para cada combinación de channel\_desc y calendar\_month.
- Subtotales de segundo nivel de agregación a través de calendar\_month\_desc y country\_id para cada valor de channel\_desc.
- Una fila de un gran total.

## EXTENSION CUBE DENTRO DE GROUP BY

CUBE toma un conjunto específico de columnas agrupadas y crea subtotales para todas las posibles combinaciones. En términos de análisis multidimensional, CUBE genera todos los subtotales que podrían ser calculados para un cubo de datos con dimensiones específicas. Por ejemplo, si se especifica CUBE (time, region, departament), el conjunto resultado incluirá todos los valores que podrían ser incluidos en una sentencia ROLLUP con algunas combinaciones adicionales.

Si se especifican n columnas en la sentencia CUBE, habrán dos (2) a la n combinaciones de subtotales retornados.

## Cuando usar CUBE



Se considera considerar usar CUBE en cualquier situación que requiera reportes “cross-tabular”. También en consultas que usen columnas de múltiples dimensiones, es decir, columnas que representan diferentes niveles de una simple dimensión.

### Sintaxis de CUBE

CUBE aparece en la cláusula GROUP BY en una sentencia SELECT de la siguiente forma:

```
SELECT ... GROUP BY CUBE (grouping_column_reference_list)
```

### Ejemplo – CUBE

```
SELECT channel_desc, calendar_month_desc, country_id,  
       TO_CHAR(SUM(amount_sold), '9,999,999,999') SALES$  
FROM sales, customers, times, channels  
WHERE sales.time_id=times.time_id AND  
       sales.cust_id=customers.cust_id AND  
       sales.channel_id= channels.channel_id AND  
       channels.channel_desc IN ('Direct Sales', 'Internet') AND  
       times.calendar_month_desc IN ('2000-09', '2000-10')  
       AND country_id IN ('UK', 'US')  
GROUP BY CUBE(channel_desc, calendar_month_desc, country_id);
```



CHANNEL_DESC	CALENDAR	CO	SALES\$
Direct Sales	2000-09	UK	1,378,126
Direct Sales	2000-09	US	2,835,557
Direct Sales	2000-09		4,213,683
Direct Sales	2000-10	UK	1,388,051
Direct Sales	2000-10	US	2,908,706
Direct Sales	2000-10		4,296,757
Direct Sales		UK	2,766,177
Direct Sales		US	5,744,263
Direct Sales			8,510,440
Internet	2000-09	UK	911,739
Internet	2000-09	US	1,732,240
Internet	2000-09		2,643,979
Internet	2000-10	UK	876,571
Internet	2000-10	US	1,893,753
Internet	2000-10		2,770,324
Internet		UK	1,788,310
Internet		US	3,625,993
Internet			5,414,303
	2000-09	UK	2,289,865
	2000-09	US	4,567,797
	2000-09		6,857,662
	2000-10	UK	2,264,622
	2000-10	US	4,802,459
	2000-10		7,067,081
		UK	4,554,487
		US	9,370,256
			13,924,743

Esta consulta ilustra la agregación CUBE a través de tres dimensiones.

## CALCULANDO SUBTOTALES SIN CUBE

En cuanto a ROLLUP, múltiples sentencias SELECT combinadas con sentencias UNION ALL podrían proveer la misma información que se obtiene con la sentencia CUBE. Sin embargo, esto podría requerir muchas sentencias SELECT. Para un cubo de n dimensiones, se necesitarían 2<sup>n</sup> sentencias SELECT. En un ejemplo de tres dimensiones, esto podría significar sentencias SELECT unidas con UNION ALL.

## Funciones Grouping

Dos aspectos importantes se deben tener en cuenta con el uso de ROLLUP y CUBE. Primero, cómo determinar de manera programática cuales conjuntos de resultado son



subtotales, y como encontrar el nivel exacto de agregación para un subtotal dado. Segundo, qué pasa si los resultados de la consulta contiene valores NULOS tanto almacenados como creados por ROLLUP o CUBE. ¿Como se puede diferenciar entre los dos?

## **Función GROUPING**

GROUPING maneja estos problemas. Usando una simple columna como argumento, GROUPING devuelve uno (1) cuando este encuentra un valor NULO creado por una operación ROLLUP o CUBE. Es decir, si el nulo indica que la fila es un subtotal, GROUPING retorna un 1, en cualquier otro caso, retorna 0.

## **Sintaxis GROUPING**

GROUPING aparece en la lista de selección de la sentencia SELECT de la siguiente forma:

```
SELECT ... [GROUPING(dimension_column)...] ...  
GROUP BY ... {CUBE | ROLLUP | GROUPING SETS} (dimension_column)
```

## **Ejemplo – GROUPING**

Este ejemplo usa GROUPING para crear un conjunto de columnas mascara para el conjunto de resultados. Las columnas mascara son fáciles de analizar programáticamente.

```
SELECT channel_desc, calendar_month_desc, country_id,  
       TO_CHAR(SUM(amount_sold), '9,999,999,999') SALES$,  
       GROUPING(channel_desc) as Ch,  
       GROUPING(calendar_month_desc) AS Mo,  
       GROUPING(country_id) AS Co  
FROM sales, customers, times, channels  
WHERE sales.time_id=times.time_id AND  
       sales.cust_id=customers.cust_id AND  
       sales.channel_id= channels.channel_id AND  
       channels.channel_desc IN ('Direct Sales', 'Internet') AND  
       times.calendar_month_desc IN ('2000-09', '2000-10')  
       AND country_id IN ('UK', 'US')  
GROUP BY ROLLUP(channel_desc, calendar_month_desc, country_id);
```



CHANNEL_DESC	CALENDAR	CO	SALES\$	CH	MO	CO
Direct Sales	2000-09	UK	1,378,126	0	0	0
Direct Sales	2000-09	US	2,835,557	0	0	0
Direct Sales	2000-09		4,213,683	0	0	1
Direct Sales	2000-10	UK	1,388,051	0	0	0
Direct Sales	2000-10	US	2,908,706	0	0	0
Direct Sales	2000-10		4,296,757	0	0	1
Direct Sales			8,510,440	0	1	1
Internet	2000-09	UK	911,739	0	0	0
Internet	2000-09	US	1,732,240	0	0	0
Internet	2000-09		2,643,979	0	0	1
Internet	2000-10	UK	876,571	0	0	0
Internet	2000-10	US	1,893,753	0	0	0
Internet	2000-10		2,770,324	0	0	1
Internet			5,414,303	0	1	1
			13,924,743	1	1	1

Un programa puede fácilmente identificar filas por una mascara (“0 0 0”). El primer nivel de subtotales tiene una mascara de “0 0 1”, el segundo nivel de subtotales tiene una mascara de “0 1 1”, y la fila del total general tiene una mascara de “1 1 1”.

Se puede mejorar la legibilidad de un conjunto de resultados usando las funciones GRUPING y DECODE, así:

### Ejemplo – GROUPING para mejorar legibilidad de los datos

```
SELECT DECODE(GROUPING(channel_desc), 1, 'All Channels', channel_desc)
      AS Channel,
      DECODE(GROUPING(country_id), 1, 'All Countries', country_id)
      AS Country, TO_CHAR(SUM(amount_sold), '9,999,999,999') SALES$
FROM sales, customers, times, channels
WHERE sales.time_id=times.time_id AND
      sales.cust_id=customers.cust_id AND
      sales.channel_id= channels.channel_id AND
      channels.channel_desc IN ('Direct Sales', 'Internet') AND
      times.calendar_month_desc= '2000-09'
      AND country_id IN ('UK', 'US')
GROUP BY CUBE(channel_desc, country_id);
```



CHANNEL	COUNTRY	SALES\$
Direct Sales	UK	1,378,126
Direct Sales	US	2,835,557
Direct Sales	All Countries	4,213,683
Internet	UK	911,739
Internet	US	1,732,240
Internet	All Countries	2,643,979
All Channels	UK	2,289,865
All Channels	US	4,567,797
All Channels	All Countries	6,857,662

Para entender la sentencia anterior, considere la primera línea de la sentencia previa:

```
SELECT DECODE(GROUPING(channel_desc), 1, 'All Channels', channel_desc)AS Channel
```

En esta, el valor channel\_desc esta determinando con una función DECODE que contiene una función GROUPING. La función GROUPING retorna un 1 si un valor de fila es una agregación creada por ROLLUP o CUBE, en otro caso este retorna 0. La función DECODE entonces opera sobre los resultados de la función GROUPING. Este retorna “ALL Channels” si este recibe 1 y el valor de channel\_desc desde la base de datos si este recibe un 0. Los valores desde la base de datos serán valores reales como “Internet” o valores NULOS almacenados. La especificación de la segunda columna, despliega country\_id, trabajando de la misma forma.

### **Cuando usar GROUPING**

La función GROUPING se usa no solo es útil para identificar NULOS, esta también permite ordenar filas de subtotales y filtrar resultados.