

**MÉTRICA DE SIMILITUD MIXTA Y  
PONDERADA PARA  
COMPARAR CARACTERÍSTICAS DE  
HERRAMIENTAS DE UN  
PROYECTO DE INTELIGENCIA DE NEGOCIOS**

**ANGELICA SOFIA PINTO REBOLLEDO  
IDUAR ALBEIRO SILVA MUÑOZ**

**TRABAJO DE GRADO PARA OPTAR AL TÍTULO DE INGENIERO  
DE SISTEMAS**

**DIRECTOR: PhD. MARTHA ELIANA MENDOZA BECERRA**

**CO-DIRECTOR: PHD. CARLOS ALBERTO COBOS LOZADA**

**Universidad del Cauca**  
**Facultad de Ingeniería Electrónica y Telecomunicaciones**  
Departamento de Sistemas  
**Línea de Investigación: Gestión de la Información,**  
**Bodega de Datos**  
Popayán, julio de 2023

# **MÉTRICA DE SIMILITUD MIXTA Y PONDERADA PARA COMPARAR CARACTERÍSTICAS DE HERRAMIENTAS DE UN PROYECTO DE INTELIGENCIA DE NEGOCIOS**



**ANGELICA SOFIA PINTO REBOLLEDO  
IDUAR ALBEIRO SILVA MUÑOZ**

Director: PhD. Martha Eliana Mendoza Becerra

Co-Director: PhD. Carlos Alberto Cobos Lozada

**Universidad del Cauca**  
**Facultad de Ingeniería Electrónica y Telecomunicaciones**  
**Departamento de Sistemas**  
**Línea de Investigación: Gestión de la Información,**  
**Bodega de Datos**  
Popayán, julio de 2023

## **AGRADECIMIENTOS**

### **Angélica Pinto**

Quiero expresar mi profundo agradecimiento al concluir esta etapa, y en especial a mis padres, cuyo esfuerzo y apoyo durante estos años me han brindado la oportunidad de iniciar y culminar mis estudios. Su amor incondicional y dedicación han sido fundamentales para sostenerme a lo largo de este camino. Mi familia y mi hermana Elizabeth han sido un pilar constante de amor y respaldo, y siempre he sabido que puedo contar con ellos en cualquier circunstancia.

A mi pareja, Gabriel, quiero expresar mi gratitud por su constante atención a los detalles más pequeños, especialmente durante esas largas noches dedicadas al trabajo universitario.

También quiero agradecer a mis compañeros, Leyder Inagan y Albeiro Silva, por su amistad y el tiempo que generosamente me han brindado en momentos en que necesitaba apoyo.

Agradezco a la ingeniera Martha por haber estado a nuestro lado a lo largo de estos años, guiándonos y realizando un excelente trabajo de dirección en nuestro proyecto de grado. También quiero extender mi agradecimiento a todos los docentes que han dejado una huella en mi proceso de aprendizaje.

Gracias a cada una de estas personas, he logrado llegar a este punto con éxito y estoy ansiosa por lo que el futuro me depara.

### **Albeiro Silva**

A mis padres y hermanos por el esfuerzo y confianza durante todo el proceso.

A mi pareja, Diana, por el cariño y apoyo constante.

A la profe Martha por el apoyo y experiencia brindada durante el desarrollo del proyecto.

A mi compañera de tesis, Angélica, por la paciencia y compañerismo durante la carrera.

Al Albeiro de 17 años que soñaba con ser Ingeniero.

## TABLA DE CONTENIDO

1	INTRODUCCIÓN .....	8
1.1	PLANTEAMIENTO DEL PROBLEMA .....	8
1.2	OBJETIVOS .....	9
1.2.1	Objetivo General.....	9
1.3	APORTE.....	10
2	MARCO TEÓRICO Y ESTADO DEL ARTE .....	11
2.1	MARCO TEÓRICO .....	11
2.1.1	Inteligencia de negocios .....	11
2.1.2	Métrica.....	12
2.2	ESTADO DEL ARTE .....	13
3	MÉTRICA DE SIMILITUD MIXTA Y PONDERADA PROPUESTA .....	16
3.1	Organización de las características de las herramientas .....	16
3.2	Preprocesamiento valores de categorías y subcategorías .....	21
3.2.1	Tipo de dato texto.....	21
3.2.2	Tipo de dato numérico .....	23
3.2.3	Tipo de dato binario .....	23
3.3	Método de puntuación para categorías y subcategorías .....	23
3.3.1	Cálculo del puntaje por tipo de dato.....	23
3.3.2	Cálculo de puntaje por categoría .....	24
3.4	Métrica mixta y ponderada.....	26
4	APLICACIÓN WEB .....	27
4.1	ARQUITECTURA .....	27
4.2	BACKEND .....	28
4.2.1	Microservicio para calcular la métrica .....	28
4.2.2	Microservicio para almacenar los resultados .....	29
4.2.3	Base de datos relacional .....	29
4.2.4	Base de datos noSQL.....	32
4.3	FRONTEND.....	32
4.3.1	Formulario para conocer la identidad de la empresa .....	32
4.3.2	Interfaz para seleccionar los servicios .....	32
4.3.3	Formularios para obtener los requerimientos de la organización .....	33
4.3.4	Interfaz de resultados .....	40
5	EVALUACIÓN.....	43

5.1	Utilidad de la aplicación web.....	43
5.1.1	Objetivo de la evaluación.....	43
5.1.2	Normatividad .....	43
5.1.3	Implementación del test de utilidad.....	43
5.1.4	Aplicación del test de utilidad.....	44
5.2	Adecuación herramientas sugeridas por la aplicación web .....	45
5.2.1	Objetivo de la evaluación.....	45
5.2.2	Normativa .....	45
5.2.3	Implementación del test de adecuación de herramientas .....	45
5.2.4	Aplicación del test de adecuación de herramientas .....	45
5.3	PROCESO DE EVALUACIÓN .....	46
5.4	RESULTADOS OBTENIDOS DEL TEST DE UTILIDAD DE LA APLICACIÓN WEB 46	
5.5	RESULTADO OBTENIDO DEL TEST DE ADECUACIÓN DE HERRAMIENTAS DE BI SUGERIDAS A UNA EMPRESA .....	48
6	CONCLUSIONES Y TRABAJO FUTURO .....	50
6.1	CONCLUSIONES .....	50
6.2	TRABAJO FUTURO .....	51
7	BIBLIOGRAFÍA .....	53

## TABLAS

Tabla 1. Pregunta de investigación .....	14
Tabla 2. Cadena de búsqueda.....	14
Tabla 3. Herramientas seleccionadas .....	16
Tabla 4. Categorías y subcategorías por servicio .....	19
Tabla 5. Valores categorías y subcategorías, Servicio SGBD, MySQL community .....	20
Tabla 6. Valores categorías y subcategorías, Servicio ETL, Tableau Prep .....	20
Tabla 7. Categorías y subcategoría con su tipo de dato .....	22
Tabla 8. Ejemplo de preprocesamiento.....	23
Tabla 9. Número de subcategorías de tipo texto y binario por cada servicio .....	26
Tabla 10. Característica y subcaracterísticas de usabilidad: Estética de la interfaz de usuario.....	43
Tabla 11. Calificación de los cuestionarios según la escala de Likert .....	44
Tabla 12. Característica y subcaracterísticas de usabilidad: Reconocibilidad de la adecuación .....	45
Tabla 13. Grado de satisfacción según niveles de puntuación.....	46
Tabla 14. Cantidad de preguntas evaluadas en los test.....	47
Tabla 15. Matriz de calidad .....	47
Tabla 16. Cantidad de preguntas evaluadas en el test para la empresa Ferragro.....	48
Tabla 17. Matriz de calidad – Ferragro .....	48
Tabla 18. Cantidad de preguntas evaluadas en el test para la empresa Club Independiente Argelia.....	48
Tabla 19. Matriz de calidad – Club Independiente Argelia. ....	49

## FIGURAS

Figura 1. Arquitectura de la aplicación web.....	28
Figura 2. Diagrama de servicios .....	29
Figura 3. Modelo relacional.....	31
Figura 4. Interfaz para identificar una empresa.....	32
Figura 5. Interfaz para seleccionar los servicios de BI. ....	33
Figura 6. Formulario SGBD – Sección de Porcentaje por categoría.....	34
Figura 7. Formulario SGBD – Sección de Categoría Costo.....	35
Figura 8. Formulario SGBD – Sección de Categoría Documentación .....	35
Figura 9. Formulario SGBD – Sección de Categoría Características técnicas .....	36
Figura 10. Formulario SGBD – Sección de Categoría Funcionalidades .....	36
Figura 11. Formulario ETL – Sección de Categoría Funcionalidades.....	37
Figura 12. Formulario OLAP – Sección de Categoría Funcionalidades.....	38
Figura 13. Formulario Reportes – Sección de Categoría Funcionalidades.....	39
Figura 14. Resultados - Información de ayuda.....	40
Figura 15. Resultados - Puntaje de las herramientas.....	41
Figura 16. Resultados - Puntaje por categoría de cada herramienta.....	42

# 1 INTRODUCCIÓN

---

La inteligencia de negocios (BI) es una disciplina que ha cobrado gran relevancia en el mundo empresarial, gracias a su capacidad para transformar los datos en información valiosa para la toma de decisiones. En este contexto, existe una amplia variedad de herramientas disponibles para los servicios de gestión de base de datos relacional, gestión del cubo, proceso ETL y reportes; que forman parte de un proyecto de BI. Sin embargo, la selección de la herramienta adecuada puede resultar un desafío, ya que cada una tiene diferentes versiones con sus propias características y beneficios.

En este trabajo de grado, se propone una solución innovadora para sugerir herramientas para proyectos de BI que se adecuen a las necesidades o requerimientos del usuario, usando una métrica de similitud mixta y ponderada que compare estas necesidades con las herramientas on-premise del mercado más representativas. Además, se propone una aplicación web que permite a los usuarios responder un cuestionario y recibir una recomendación personalizada.

Este trabajo presenta una importante contribución al campo de la BI, al proponer una alternativa sencilla para la selección de herramientas y ofrecer una aplicación web que puede ser utilizada en diferentes contextos empresariales, como apoyo al desarrollo de proyectos de BI.

Finalmente, la organización de este documento cuenta con cinco capítulos, en el primero se presenta la introducción, el planteamiento del problema y los objetivos; el segundo capítulo expone el marco teórico que describe en detalle los conceptos relacionados con el desarrollo del presente trabajo de grado, además de la revisión del estado del arte con respecto a las tareas que utilizan medidas de similitud textual; en el tercer capítulo se presenta la métrica de similitud mixta y ponderada propuesta; en el cuarto se presenta la arquitectura y el diseño de la aplicación web; y en el quinto capítulo se evalúa el grado en que se adecuan las herramientas sugeridas a las necesidades de dos empresas.

## 1.1 PLANTEAMIENTO DEL PROBLEMA

El aumento de la competencia en el mercado y la enorme cantidad de datos recopilados por las actividades comerciales aumentan la importancia de utilizar un sistema de inteligencia empresarial (BI) [1], el cual, es un conjunto de herramientas y procesos que apoyan la toma de decisiones de las empresas basado en datos precisos, ahorrando tiempo y esfuerzo. La idea principal de un proyecto de BI es la posibilidad de analizar fácilmente datos basados en conceptos comerciales sin tener conocimientos técnicos sobre herramientas de bases de datos u otras fuentes que contienen los datos. Además, proporcionan un análisis de datos para respaldar productos, servicios y procesos relacionados en su desarrollo, mejora o redefinición. Esto permite que cualquier organización o empresa se beneficie de la aplicación de BI independientemente de su tamaño, negocio o ubicación en los mercados nacionales e internacionales [2].

Por otro lado, las empresas Micro, Pequeñas y Medianas (MiPyMe's) son la columna vertebral de la economía mundial actual, compuestas por un 95% de estas [1]. En Colombia, las MiPyMe's se han convertido a lo largo de los años como parte fundamental del tejido económico del país, esto debido a su gran capacidad de generación de empleo; representando entre el 96% y 99% de la estructura empresarial [3]. Según [4] para tomar



decisiones a tiempo los directivos de estas empresas utilizan principalmente su experiencia debido a la falta de recursos humanos, lo que implica un alto riesgo de fracaso. Para implementar un proyecto de BI se necesita integrar herramientas para: (1) la gestión de la base de datos relacional; (2) el motor para el procesamiento analítico en línea (OLAP), para realizar consultas analíticas complejas; (3) el proceso de extracción, transformación y carga de datos (ETL), que permite la organización de datos; y (4) los reportes, para presentar la información almacenada de manera que pueda ser entendida y analizada para la posterior toma de decisiones.

Los principales desafíos para implementar un proyecto de BI según [4] son la falta de tiempo para administrar el sistema, la falta de recursos humanos, factores económicos y la variedad de herramientas que no coinciden con las necesidades comerciales. El creciente uso de BI en el mercado ha generado esta gran oferta de herramientas, lo cual, dificulta seleccionarlas de acuerdo con las necesidades técnicas de las empresas. Una forma automática de abordar este último problema es haciendo una comparación textual de las características de cada una de las herramientas con las necesidades técnicas de una empresa. En la literatura se encontraron trabajos que presentan que las medidas de similitud textual más eficientes y precisas en comparación con otras para realizar tareas de comparación, clasificación y agrupación de texto son: Coseno y Jaccard [5]. Estos estudios también proporcionan evidencia sobre que la similitud de coseno mejora su desempeño en temáticas con un alto vocabulario técnico. Sin embargo, no hay estudios que usen estas medidas de similitud en el proceso de comparación de necesidades técnicas específicas de una empresa con las características de un conjunto de herramientas on-premise disponibles en el mercado.

Teniendo en cuenta lo anterior, surge la siguiente pregunta de investigación que se abordará en este proyecto: ¿Es útil usar una métrica de similitud mixta y ponderada, para ordenar las herramientas de gestión de bases de datos relacional, gestión del cubo, proceso ETL y reportes de un proyecto de inteligencia de negocios de acuerdo con las necesidades técnicas de la empresa?

## **1.2 OBJETIVOS**

### **1.2.1 Objetivo General**

Proponer una métrica de similitud mixta y ponderada, que permita comparar características de un conjunto de herramientas On-premise para un proyecto de inteligencia de negocios, y ordenarlas de acuerdo con las necesidades técnicas y económicas de una empresa.

### **1.2.2 Objetivos Específicos**

- Definir una métrica de similitud mixta y ponderada, basada en el tipo y estructura de las características de un conjunto herramientas de gestión de bases de datos relacional, gestión del cubo, proceso ETL y reportes de un proyecto de inteligencia de negocios, que permitan hacer la comparación con las necesidades técnicas y económicas de una empresa.
- Construir un prototipo de una aplicación web con la métrica mixta y ponderada definida para un conjunto de herramientas de inteligencia de negocios, que permita ordenar las herramientas a partir de las necesidades técnicas y económicas de una empresa.

- Evaluar el grado en que se adecuan las herramientas sugeridas a las necesidades de dos empresas, por medio de un test<sup>1</sup> aplicado a dos profesionales con conocimiento en proyectos de inteligencia de negocios.

### **1.3 APORTE**

Este trabajo de investigación contribuye con la definición de una métrica de similitud mixta y ponderada que tiene en cuenta características cuantitativas y cualitativas de un conjunto de herramientas de un proyecto de inteligencia de negocios (servicios de gestión de bases de datos relacional, gestión del cubo, proceso ETL y reportes); que permita la comparación con las necesidades técnicas y económicas de una empresa. Teniendo en cuenta, que no se encontró en la literatura un reporte con una métrica de similitud mixta y ponderada para abordar este problema.

Además, se presenta un prototipo de una aplicación web que permite recolectar las solicitudes específicas de una empresa y ordenar un conjunto de herramientas de BI disponibles en el mercado de acuerdo con las necesidades específicas de la empresa.

---

<sup>1</sup> Test de adecuación de herramientas de BI sugeridas a una empresa

## 2 MARCO TEÓRICO Y ESTADO DEL ARTE

---

En este capítulo se presentan los conceptos necesarios para el desarrollo de este trabajo de grado y los artículos relevantes encontrados relacionados con las tareas que utilizan medidas de similitud textual.

### 2.1 MARCO TEÓRICO

En esta sección se describen los conceptos teóricos relacionados con los servicios y herramientas de un proyecto de inteligencia de negocios. Además, el concepto de métrica mixta, ponderada y de similitud de texto, importantes para el desarrollo de este proyecto.

#### 2.1.1 Inteligencia de negocios

El término Inteligencia de negocios (BI) fue acuñado originalmente por consultores de negocios de Gartner Group desde 1996. BI es un término general que incluye arquitecturas, herramientas, bases de datos, aplicaciones y metodologías con el objetivo de analizar datos para apoyar la toma de decisiones de los gerentes de negocios, proporcionando beneficios a cualquier organización independientemente de sus factores económicos [6].

Un sistema de BI es un conjunto de herramientas, tecnologías y productos programados que se utilizan para recopilar, integrar, agregar y hacer que los datos estén disponibles. BI comprende una variedad de software analítico que proporciona la información que necesitan las empresas y su énfasis está en la información en tiempo real que respalda la presentación de informes en todos los niveles de la organización, principalmente, conformado por herramientas ETL, técnicas de almacenamiento de datos y cubos dimensionales que trabajan de forma interoperable ya que la conectividad entre estas ha sido garantizada [7]:

- **Bases de datos relacionales.** Una base de datos relacional es un tipo de base de datos que almacena y proporciona acceso a puntos de datos relacionados entre sí. Las bases de datos relacionales se basan en el modelo relacional, una forma intuitiva y directa de representar datos en tablas. En una base de datos relacional, cada fila en una tabla es un registro con una ID única, llamada *clave*. Las columnas de la tabla contienen los atributos de los datos y cada registro suele tener un valor para cada atributo, lo que simplifica la creación de relaciones entre los puntos de datos [8]. Entre las más conocidas están: SQL Server, Oracle, PostgreSQL, MySQL y Firebird.
- **Cubo dimensional.** Un cubo OLAP es una estructura de datos que supera las limitaciones de las bases de datos relacionales y proporciona un análisis rápido de datos. Los cubos pueden mostrar y sumar grandes cantidades de datos, a la vez que proporcionan a los usuarios acceso mediante búsqueda a los puntos de datos. De este modo, los datos se pueden agregar y segmentar, según sea necesario para controlar la variedad de preguntas que son relevantes para el área de interés de un usuario [9]. Las herramientas más conocidas son: SQL Server Analysis Services, Pentaho Schema Workbench Community, Iccube y SQL Power Architect.
- **Proceso ETL.** El proceso de extracción, transformación y carga (ETL) permite la integración y el análisis de los datos almacenados en diferentes bases de datos hacia la bodega de datos relacional. Las tres etapas que conforman este proceso son: (1) Extracción, consiste en acceder a los datos de origen y recopilar los valores necesarios

para la bodega de datos, convirtiéndolos a un formato adecuado; (2) Transformación, aplica reglas o funciones para validar los registros y garantizar la calidad de los datos, los procesos más comunes para la transformación son: conversión, borrar duplicados, estandarizar, filtrar, ordenar, traducir y buscar o verificar si las fuentes de datos son inconsistentes; (3) Carga, consiste en cargar los datos extraídos y transformados en un repositorio de destino como tablas de dimensiones y tablas de hechos [10]. Algunas herramientas más usadas son: SQL Server Integration Services, Pentaho Data Integration, Cloverdx Designer y Talend For Data Integration.

- **Reportes.** Los informes de BI permiten preparar y analizar los datos para encontrar y compartir conocimientos prácticos, de esta forma, estos informes ayudan a los usuarios a mejorar las decisiones y el rendimiento empresarial. Esta es la fase final de todo proyecto de BI [10]. Existen diferentes tipos de informes, desde estáticos hasta paneles interactivos y análisis integrados. Algunas de las herramientas más conocidas para generación de reportes son: SQL Server Reporting Services, Pentaho Reports Designer, SQL Power Wabbit Community Edition y Jaspersoft [6].

### 2.1.2 Métrica

Es una función matemática que define una distancia entre cada par de elementos de un conjunto. Una métrica en un conjunto  $X$  es una función, llamada función de distancia o simplemente distancia [11]. Ver Ecuación (1).

$$d: X \times X \rightarrow [0, \infty) \quad (1)$$

- **Métrica mixta**

Una métrica mixta está compuesta por variables que hacen referencia a variables cuantitativas y variables cualitativas; independientemente de que se encuentren tipificadas en formato binario, ordinal, o nominal o se encuentren estas en formato alfanumérico [11].

- **Métrica ponderada**

Es un tipo de métrica que otorga diferentes pesos a los distintos valores sobre los que se calcula. El peso es el valor que se utiliza para otorgar mayor o menor importancia, cuanto mayor sea este valor mayor importancia tendrá el elemento [12].

- **Medidas de similitud**

Al trabajar con sistemas de Procesamiento de Lenguaje Natural (NLP) es frecuente encontrar la necesidad de comparar diferentes palabras o frases entre sí. En muchos casos es de interés no solamente las coincidencias exactas entre dos cadenas de texto, sino también tener una medida de aproximación o similitud entre estas cuando la coincidencia no es perfecta [5].

La medición de la similitud de texto es la base de las tareas de NLP, que desempeñan un papel importante en la comparación, clasificación y agrupación de textos para la recuperación de información, detección de temas, la traducción automática, respuesta a preguntas, calificación de ensayos, calificación de respuestas cortas, resumen de texto y otros [13]. Para dichos casos resulta de utilidad apoyarse en las medidas de similitud o distancia en cadenas de caracteres.

Cuando los elementos comparados son cadenas de texto, la métrica es una métrica de cadenas. A continuación, se describen las medidas de Coseno y Jaccard, por ser de interés para el desarrollo de este proyecto:

- **Medida de similitud de coseno.** Es una medida de similitud entre dos vectores de un espacio de producto interno que mide el coseno del ángulo entre ellos [13]. Esta medida fue originalmente usada en aplicaciones de recuperación de información para encontrar la semejanza entre una consulta y un documento. Los documentos y la consulta son representados por vectores y la métrica calcula su similitud. La forma más simple de los vectores es tomar cada elemento  $d_i=1$  si la palabra  $i$ -ésima aparece en el documento o tomar  $d_i=0$  en caso contrario [13]. Esta es una métrica que utiliza el modelo de representación *tf-idf* para la representación de los datos. La forma general de esta métrica es como se muestra en la Ecuación (2):

$$\text{Cos}(q, d) = \frac{\sum_{i=1}^n d_i q_i}{\sqrt{(\sum_{i=1}^n (d_i)^2 * \sum_{i=1}^n (q_i)^2)}} \quad (2)$$

donde  $q$  es el vector consulta y  $d$  es el vector documento.

- **Medida Índice Jaccard.** El índice Jaccard opera sobre conjuntos, por lo cual, es común que se utilice para comparar sentencias o párrafos completos como un conjunto de palabras. Sin embargo, también puede ser utilizado para comparar palabras considerándolas como conjuntos de letras o caracteres. Cualquiera que sea el nivel de tokenización sobre el que se utilice, es interesante notar que la posición que ocupa el elemento no tiene relevancia y que los elementos repetidos son considerados como uno solo dentro del conjunto [14].

Esta medida está definida como la intersección sobre la unión de los conjuntos y es una medida de la similitud entre ambos. Es decir, es la división entre el número de elementos en común que tienen los dos conjuntos sobre el número de elementos únicos que tiene la unión de ambos conjuntos [12]. La fórmula del índice Jaccard es como indica la Ecuación (3).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

Por su parte, la distancia de Jaccard es el resultado de restarle a uno el valor de la similitud, ver Ecuación (4)

$$d_j(A, B) = 1 - J(A, B) \quad (4)$$

## 2.2 ESTADO DEL ARTE

Se realizó una revisión sistemática siguiendo algunos lineamientos planteados por [15], en diferentes fuentes de información para la investigación científica (ACM Digital Library, IEEE, Science Direct, Scopus SpringerLink). Para esto, se planteó una pregunta (ver **Tabla 1**) con el objetivo de identificar las principales medidas de similitud utilizadas en el tema de investigación.

**Tabla 1. Pregunta de investigación**

---

**Pregunta**

---

¿Cuáles son las medidas de similitud de texto que actualmente reportan los mejores resultados en tareas de comparación, clasificación y agrupación?

---

A partir de la pregunta planteada, se definió una cadena de búsqueda con las palabras claves que se presentan en la **Tabla 2**.

**Tabla 2. Cadena de búsqueda**

---

similarity AND text AND measure

---

Para obtener los artículos de investigación se definieron los siguientes criterios de inclusión: artículos publicados desde el año 2016 hasta la fecha, título y resumen escritos en inglés. Y como criterios de exclusión: artículos que no usan medidas para la similitud de texto y no presentan resultados sobre el rendimiento de estas.

Luego, se aplicó la cadena de búsqueda y en total se encontraron 22 artículos relacionados con el tema de investigación, a los cuales se les aplicaron los criterios de inclusión y exclusión, quedando 7 trabajos relevantes para esta investigación; los cuales se agruparon en las tareas de recuperación de la información de comparación, clasificación y agrupación.

- **Comparación**

En el año 2017, se realizó una investigación que se centra en el efecto del uso de técnicas superficiales de procesamiento del lenguaje natural junto con una combinación de medidas de similitud para la detección de plagio en documentos de texto en Malayalam [16]. Se usan tres medidas de similitud textual: Jaccard, coeficiente de Dice y Coseno; y se realiza la comparación de los corpus usando PLN y también sin usar PLN en combinación con estas medidas. En ambos casos la medida de *similitud de coseno* muestra una mejor precisión.

Más adelante el año 2019, se propone un sistema de recomendación automática para ayudar a los autores a elegir un medio de publicación más apropiado para enviar un manuscrito [17]. El estudio examina la idoneidad de tres medidas de similitud de texto (coseno, unigrama y BM25) para la comparación del contenido del manuscrito y los artículos de revistas existentes en dos corpus temáticos sobre las ciencias sociales y la medicina. Los resultados proporcionan evidencia para argumentar que la *similitud de coseno* mejora su desempeño en temáticas y dominios de materias con un vocabulario técnico más alto.

En el mismo año, en España [18], implementaron nuevos algoritmos inteligentes capaces de determinar la similitud de documentos de texto en un entorno empresarial, comparando la similitud textual del documento de las solicitudes de los clientes de propiedades inmobiliarias con las características de las propiedades disponibles en el servidor web de la empresa, para hacer coincidir la mejor opción para el cliente.

Recientemente, en el año 2020, se realizó una comparación de tres métodos para estimar la similitud semántica entre dos artículos de noticias sobre el mismo tema o evento en dos idiomas diferentes (hindi e inglés) [5]. En el experimento se utilizaron las medidas de

similitud de Coseno con tf-idf, Jaccard con tf-idf y distancia euclidiana con Bag of words, sobre los conjuntos de datos de GoogleNews. La *similitud de coseno* utilizando tf-idf mostró una mayor precisión, recuerdo y medida F; que las otras dos medidas.

- **Clasificación**

En el año 2018 se realiza un trabajo [19] donde clasifican las opiniones de reseñas textuales de películas en función de si son positivas o negativas, los autores deciden usar la similitud de coseno en combinación con algoritmos de clasificación y demuestran empíricamente que la precisión mejora en el conjunto de datos cuando se usa *similitud de coseno* en comparación con el uso del producto escalar.

Un estudio en el mismo año para optimizar la búsqueda y el proceso de recomendación de artículos de investigación [20], utiliza algoritmos de clasificación no lineal (rpart, random forest y aprendizaje automático) y combinaciones de medidas de texto: coseno, distancia euclidiana, Jaccard y coeficiente de Pearson; para comparar las características importantes de un artículo con las de otros documentos. La *medida de coseno* se comportó mejor en comparación con las otras.

- **Agrupación**

En el año 2016, un trabajo de investigación crea varios modelos para agrupar documentos de proyectos finales utilizando tres técnicas de agrupación [21]: k medias, k-medias rápidas y k-medoids utilizando un conjunto de datos de la biblioteca de la Facultad de Ciencias de la Información y la Computación de la Universidad King Saud, Riyadh. Se probaron tres medidas de similitud: coseno, Jaccard y coeficiente de correlación. Los resultados indicaron que el mejor rendimiento de agrupación se logra utilizando k-medias y k-medoids combinados con *similitud de coseno*.

Los artículos mencionados en el estado del arte permiten apreciar que la medida de similitud Coseno y Jaccard obtuvieron un mejor rendimiento en tareas de comparación, clasificación y agrupación.

## 3 MÉTRICA DE SIMILITUD MIXTA Y PONDERADA PROPUESTA

### 3.1 ORGANIZACIÓN DE LAS CARACTERÍSTICAS DE LAS HERRAMIENTAS

La implementación de un proyecto de Inteligencia de negocios (BI) requiere la integración de diversas herramientas para gestionar la base de datos relacional, realizar el procesamiento analítico en línea (OLAP), llevar a cabo el proceso de extracción, transformación y carga de datos (ETL) y generar reportes que permitan el análisis y toma de decisiones.

Sin embargo, seleccionar las herramientas adecuadas puede ser un desafío debido a la falta de tiempo, recursos humanos, consideraciones económicas y la amplia variedad de opciones disponibles en el mercado. Para abordar este problema, se utiliza un enfoque automatizado que compara las características de cada herramienta con las necesidades técnicas de una empresa.

Los criterios utilizados para evaluar y seleccionar las herramientas incluyen su funcionalidad, costo, documentación y características técnicas relevantes. Para evaluar las herramientas de OLAP, ETL y reportes, se utilizó el Cuadrante Mágico de Analítica e Inteligencia de Negocios Plataformas de Gartner [22], una herramienta que evalúa y clasifica a los proveedores en función de su capacidad de ejecución y visión completa. Además, se consideraron herramientas de software libre debido a sus beneficios como el ahorro de costos, flexibilidad, participación de una comunidad activa, interoperabilidad e innovación continua [23]. Para seleccionar los motores relacionales, se consideró que estos fueran relevantes en el mercado [referencia], se buscaron características de seguridad y control de acceso robustas para garantizar su idoneidad en entornos empresariales y se valoró la escalabilidad y la integridad de los datos [24]. En la **Tabla 3** se presentan las herramientas seleccionadas por cada servicio.

**Tabla 3. Herramientas seleccionadas**

Fuente usada	SGBD	ETL	OLAP	REPORTES
<b>Cuadrante de Gartner</b>	—	Power Query Tableau Prep	Power BI Tableau	Power BI Tableau
<b>Software libre</b>	Postgresql Mysql	Pentaho Jaspersoft	Pentaho Jaspersoft	Pentaho Jaspersoft
<b>Bases de datos</b>	Oracle Sql server	—	—	—

Para las herramientas seleccionadas se evaluaron sus características y capacidades, teniendo en cuenta tanto las necesidades técnicas como las consideraciones económicas; buscando proporcionar a las empresas las herramientas más adecuadas para lograr sus objetivos en el ámbito de la inteligencia de negocios.



Este proyecto tomó como base las cuatro categorías establecidas para cada servicio (SGBD, ETL, OLAP y reportes) en el estudio [25], y se agregan nuevas subcategorías que se consideraron relevantes. Estas categorías son:

- **Costo:** Se refiere al precio de la herramienta, incluyendo licencias o suscripciones necesarias para acceder a ella. El costo es un factor importante al elegir una herramienta, ya que varía significativamente entre las opciones disponibles en el mercado.
- **Documentación:** Tiene en cuenta la calidad y cantidad de recursos disponibles para ayudar a los usuarios a comprender y utilizar la herramienta. Esto puede incluir manuales de usuario, tutoriales, videos explicativos, foros de discusión y otros recursos educativos.
- **Características técnicas:** Se refiere a las especificaciones técnicas necesarias para garantizar las capacidades necesarias para la instalación y el correcto funcionamiento de la herramienta.
- **Funcionalidades específicas:** Comprende las características y funcionalidades específicas para cada servicio que ofrece la herramienta.

Durante el proceso de investigación se exploró una variedad de foros y documentación oficial de las herramientas seleccionadas que permitieron identificar otras subcategorías con el objetivo de agregar las funcionalidades más relevantes relacionadas con los servicios de SGBD, ETL, OLAP y reportes.

Dentro del servicio SGBD se incorporó la siguiente subcategoría:

- **Funciones de cifrado:** Las funciones de cifrado en una base de datos es importante para salvaguardar la confidencialidad, la integridad y la disponibilidad de los datos sensibles, protegiendo así la información y cumpliendo con las regulaciones de protección de datos.

Dentro del servicio ETL se incorporaron las siguientes subcategorías:

- **Depurador con puntos de interrupción:** Esta característica facilita la identificación y solución de errores, permite pruebas y validaciones más efectivas y ayuda en la optimización del rendimiento del flujo de ETL.
- **Recorrer hojas Excel:** Esto permite la integración de datos provenientes de archivos de Excel, automatiza tareas, mejora la eficiencia y brinda flexibilidad para adaptarse a diferentes formatos de datos.
- **Actualización programada de datos:** Permite mantener los datos actualizados, asegurar la consistencia de la información, ahorrar tiempo y recursos, cumplir con plazos y requerimientos, y mejorar la toma de decisiones basada en datos precisos y actualizados.

Dentro del servicio reportes se incorporaron las siguientes subcategorías:

- Plantillas predeterminadas para la realización de informes: Estas plantillas agilizan el proceso de creación, aseguran la coherencia visual, facilitan el uso y personalización, y mejoran la comunicación de la información.
- Gráficos 3D: Permite mejorar la visualización de datos, un análisis más profundo, brinda una perspectiva espacial y crea presentaciones impactantes. Todo esto contribuye a una mejor comprensión de la información y facilita la toma de decisiones fundamentadas.
- Minería de datos: Ayuda a descubrir patrones, predecir eventos futuros, personalizar estrategias, detectar fraudes/anomalías y optimizar procesos, lo que contribuye a la toma de decisiones informadas y al éxito empresarial.
- Notificaciones: Proporciona actualizaciones en tiempo real, mejora la eficiencia, ofrece relevancia personalizada y permite la detección temprana de eventos críticos, lo que contribuye a una toma de decisiones más informada y ágil.
- Mobile: Permite una mayor accesibilidad, agilidad en la toma de decisiones, colaboración efectiva y análisis en tiempo real, lo que se traduce en una mejor utilización de los datos y una ventaja competitiva para las organizaciones.
- Visualizaciones interactivas: Mejoran la comprensión de los datos, facilitan la toma de decisiones informada, ayudan a descubrir información oculta y promueven una comunicación efectiva.

También se eliminaron algunas subcategorías de cada servicio como se presenta a continuación.

Dentro del servicio SGBD se eliminaron las siguientes subcategorías porque para algunos motores no se encontró información relacionada:

- Monitorear las acciones del usuario por cada esquema.
- Particiones y compresión de datos.

Dentro del servicio ETL se eliminaron las siguientes subcategorías, porque en algunas herramientas no se pudo comprobar si tenía o no la funcionalidad:

- Número máximo de componentes por transformación.
- Máxima cantidad de datos que se pueden migrar.
- Carga incremental.

Dentro del servicio OLAP se eliminaron las siguientes subcategorías, porque en algunas herramientas no se pudo comprobar si tenía o no la funcionalidad:

- Relación muchos a muchos entre dimensión y tabla de hechos.
- Juego de roles.
- Representación de esquema en archivo XML.

Dentro del servicio reportes se eliminaron las siguientes subcategorías, porque todas las herramientas presentan estas características, por lo tanto, al momento de realizar el cálculo no se presentaría una diferencia entre las herramientas:

- Tabla de contenido.
- Parámetros estáticos.
- Botones de radio.

Cómo resultado se obtuvo la **Tabla 4** donde se obtiene las categorías y subcategorías para cada servicio. Como se observar las categorías de Costo, Documentación y Características técnicas son iguales para todos los servicios, pero la de Funcionalidades si tiene en cuenta las particularidades que ofrece la herramienta para cada servicio.

**Tabla 4. Categorías y subcategorías por servicio**

SGBD	ETL	OLAP	Reportes
<b>Costo</b>			
<b>Documentación</b>			
<b>Características técnicas</b>			
<ul style="list-style-type: none"> <li>• Espacio en disco.</li> <li>• Uso de memoria.</li> <li>• Multiplataforma.</li> </ul>			
<b>Funcionalidades</b> <ul style="list-style-type: none"> <li>• APIs y otros métodos de acceso.</li> <li>• Métodos de replicación.</li> <li>• Funciones nativas de cifrado.</li> </ul>	<b>Funcionalidades</b> <ul style="list-style-type: none"> <li>• Extracción.</li> <li>• Carga.</li> <li>• Conexión a SGBD.</li> <li>• Búsqueda.</li> <li>• Concatenar valores.</li> <li>• Dividir campos.</li> <li>• Filtro.</li> <li>• Agregación.</li> <li>• Merge.</li> <li>• Normalización.</li> <li>• Ordenamiento.</li> <li>• Remoción duplicados.</li> <li>• Desnormalización.</li> <li>• Depurador.</li> </ul>	<b>Funcionalidades</b> <ul style="list-style-type: none"> <li>• Tipo de dimensiones.</li> <li>• Medidas calculadas.</li> <li>• Jerarquía.</li> <li>• Traducciones.</li> <li>• Creación de varios cubos en un mismo esquema.</li> <li>• Múltiples orígenes de datos en un mismo esquema.</li> <li>• Permisos por medio de roles.</li> </ul>	<b>Funcionalidades</b> <ul style="list-style-type: none"> <li>• Conexión a motor OLAP.</li> <li>• Conexión a SGBD.</li> <li>• Varios orígenes de datos en un mismo reporte.</li> <li>• Plantillas predeterminadas.</li> <li>• Gráficos 3D.</li> <li>• Minería de datos.</li> <li>• Notificaciones.</li> <li>• Soporte dedicado.</li> <li>• Mobile.</li> <li>• Visualizaciones interactivas.</li> </ul>

Para revisar los posibles valores de cada categoría y subcategoría, se consultó los sitios web oficiales de cada herramienta. Para las funcionalidades específicas, también se revisaron foros y sitios relevantes en el sector. En algunos casos, fue necesario descargar y probar la herramienta para confirmar si contaba con la funcionalidad en cuestión. En la **Tabla 5** y en la **Tabla 6** se presenta a modo de ejemplo los valores para las categorías y subcategorías en el servicio de SGBD para una herramienta community y ETL para una propietaria.

**Tabla 5. Valores categorías y subcategorías, Servicio SGBD, MySQL community**

<b>Categorías/Subcategorías</b>	<b>MySQL Community</b>
<b>Costo</b>	0
<b>Documentación</b>	Oficial, foros
<b>Características técnicas</b>	
Espacio en disco	500 MB
Uso de memoria	512 MB
<b>Funcionalidades</b>	
APIs métodos de acceso	<ul style="list-style-type: none"> <li>• ADO.NET Driver</li> <li>• ODBC Driver</li> <li>• JDBC Driver</li> <li>• Node.js Driver</li> <li>• Python Driver</li> <li>• C++ Driver</li> <li>• C Driver</li> <li>• C API</li> </ul>
Funciones nativas de cifrado	Algoritmo AES

**Tabla 6. Valores categorías y subcategorías, Servicio ETL, Tableau Prep**

<b>Categorías/Subcategorías</b>	<b>Tableau Prep</b>
<b>Costo</b>	70 USD
<b>Documentación</b>	Oficial, foros
<b>Características técnicas</b>	
Espacio en disco	4000 MB
Uso de memoria	2000 MB
<b>Funcionalidades</b>	
Extracción (lectura de archivos)	<ul style="list-style-type: none"> <li>• Microsoft excel</li> <li>• Archivo de texto</li> <li>• Microsoft access</li> <li>• Archivo json</li> <li>• Archivo xml</li> </ul>
Carga de archivos	<ul style="list-style-type: none"> <li>• CSV</li> </ul>
Conexión a bases de datos representativas	<ul style="list-style-type: none"> <li>• MySQL</li> <li>• SAP HANA</li> <li>• Oracle</li> <li>• SQLServer</li> </ul>
Búsqueda	Si
Concatenar valores	Si
Dividir campos	Si
Filtro	No
Agregación	Si
Merge	Si
Normalización	Si
Ordenamiento	Si
Remoción duplicados	Si
Desnormalización	Si

Depurador	Si
-----------	----

Debido a que las herramientas cuentan con diferentes versiones, la información se presenta en un archivo Excel para mejor visualización, ver anexo *Categorías y subcategorías con sus valores*.

### 3.2 PREPROCESAMIENTO VALORES DE CATEGORÍAS Y SUBCATEGORÍAS

Después de identificar los valores para cada categoría y subcategoría, se estableció un tipo de dato para cada una. En este caso, se han clasificado en tres tipos de datos que abarcan todas las posibilidades, que son:

- **Numérico:** Cuando el valor es un número.
- **Texto:** Para valores textuales.
- **Binario:** Indica si la herramienta cuenta con una funcionalidad determinada o no.

En la **Tabla 7** se presenta las categorías y subcategorías con su tipo de dato para cada servicio.

#### 3.2.1 Tipo de dato texto

Las categorías y subcategorías que contienen valores de tipo texto requirieron preprocesamiento ya que contenían múltiples opciones que debían ser almacenadas en un solo campo textual. Por ejemplo: la subcategoría *Conexión a SGBD* de la categoría *funcionalidades* del servicio OLAP, puede contener los siguientes valores:

- Oracle.
- Mysql.
- Postgres.
- María db.

Los cuales se convierten en un texto con su respectivo delimitador de la siguiente manera:

- "Oracle, Mysql, Postgres, Maria db".

Adicional a esto, se realizó un segundo preprocesamiento, que consistió en filtrar de los valores de las subcategorías almacenadas, con sólo los valores que seleccionó el usuario, esta parte es importante para que el cálculo que se explicará en la sección 3.3.1, no se realice sobre la totalidad de los valores. Por ejemplo: Si el usuario seleccionó para la subcategoría *Conexión a SGBD* los valores MySQL y María db, pero en la herramienta se tienen almacenados Oracle, MySQL, SQLServer y Postgresql. El cálculo se hará con el valor de la subcategoría almacenada MySQL, como se muestra en la **Tabla 8**.

**Tabla 7. Categorías y subcategoría con su tipo de dato**

SGBD	ETL	OLAP	Reportes	Tipo de dato
<b>Costo</b>				<b>Numérico</b>
<b>Documentación</b>				<b>Texto</b>
<b>Características técnicas</b>				
<ul style="list-style-type: none"> <li>• Espacio en disco.</li> <li>• Uso de memoria.</li> </ul>				<b>Numérico</b>
<ul style="list-style-type: none"> <li>• Multiplataforma.</li> </ul>				<b>Texto</b>
<b>Funcionalidades</b>				
<ul style="list-style-type: none"> <li>• APIs y otros métodos de acceso.</li> <li>• Métodos de replicación.</li> <li>• Funciones nativas de cifrado.</li> </ul>	<ul style="list-style-type: none"> <li>• Extracción.</li> <li>• Carga.</li> <li>• Conexión a SGBD.</li> </ul>	<ul style="list-style-type: none"> <li>• Tipo de dimensiones.</li> <li>• Conexión a SGBD.</li> <li>• Conexión archivos</li> </ul>	<ul style="list-style-type: none"> <li>• Conexión a motor OLAP.</li> <li>• Conexión a SGBD.</li> <li>• Varios orígenes de datos en un mismo reporte.</li> </ul>	<b>Texto</b>
	<ul style="list-style-type: none"> <li>• Búsqueda.</li> <li>• Concatenar valores.</li> <li>• Dividir campos.</li> <li>• Filtro.</li> <li>• Agregación.</li> <li>• Merge.</li> <li>• Normalización.</li> <li>• Ordenamiento.</li> <li>• Remoción duplicados.</li> <li>• Desnormalización.</li> <li>• Depurador.</li> </ul>	<ul style="list-style-type: none"> <li>• Medidas calculadas.</li> <li>• Jerarquía.</li> <li>• Traducciones.</li> <li>• Creación de varios cubos en un mismo esquema.</li> <li>• Múltiples orígenes de datos en un mismo esquema.</li> <li>• Permiso por medio de roles</li> </ul>	<ul style="list-style-type: none"> <li>• Plantillas predeterminadas.</li> <li>• Gráficos 3D.</li> <li>• Minería de datos.</li> <li>• Notificaciones.</li> <li>• Soporte dedicado.</li> <li>• Mobile.</li> <li>• Visualizaciones interactivas.</li> </ul>	<b>Binario</b>

**Tabla 8. Ejemplo de preprocesamiento**

Ingresado por el usuario	Almacenado para la herramienta	Resultado por el preprocesamiento
MySQL, MariaDB	Oracle, MySQL, SQLServer, PostgreSQL	MySQL

### 3.2.2 Tipo de dato numérico

Para las categorías y subcategorías que contienen valores numéricos, tales como espacio en disco y uso de memoria, se estableció utilizar como unidad de medida megabyte (MB) con el fin de simplificar las operaciones de cálculo en la aplicación.

### 3.2.3 Tipo de dato binario

Para las categorías y subcategorías que contienen valores binarios, se realizó un preprocesamiento parecido al realizado a los tipos de dato texto, donde se filtran los valores almacenados teniendo en cuenta los valores ingresados por el usuario.

## 3.3 MÉTODO DE PUNTUACIÓN PARA CATEGORÍAS Y SUBCATEGORÍAS

Para realizar el cálculo, primero se estableció la escala a utilizar, en este caso, una escala de 0 a 1, donde 0 representa la ausencia completa de la característica o habilidad, mientras que 1 representa su presencia completa. Los puntajes intermedios se ubicaron en una escala que va desde la ausencia completa hasta la presencia completa de la característica o habilidad evaluada.

### 3.3.1 Cálculo del puntaje por tipo de dato

Para los diferentes tipos de datos identificados, se utilizaron diferentes escalas de medición. En el caso de los datos numéricos, se utilizó una escala binaria de 0 o 1 para indicar su presencia o ausencia. Para los datos de texto y binarios, se utilizó una escala de medición que va del 0 al 1, con el fin de reflejar la presencia y relevancia de dichos datos.

#### 1. Tipo de dato numérico

Si el tipo de dato numérico es de la categoría *Costo* se usó el siguiente condicional:

- Si el valor que tiene la herramienta es cero, entonces tendrá un puntaje de 1, adicionalmente tendrá una bonificación de 1 porque se ahorraría la totalidad de su dinero.
- Si el valor ingresado por el usuario es mayor al valor que tiene la herramienta, el puntaje será de 1, adicional a esto se aplica la Ecuación (5) para dar una bonificación, debido al ahorro que se tendría con esta herramienta.

$$\left(\frac{(\text{valorAlmacenado} - \text{valorEntrada})}{\text{valorAlmacenado}}\right)^* - 1 \quad (5)$$

Si, por el contrario, el valor ingresado es menor que el valor que tiene la herramienta, el puntaje será de 0, debido a que no se ajusta a los requerimientos de costo del usuario.

Si el tipo de dato numérico es de las subcategorías *Espacio en disco* y *Uso de memoria* se usa el siguiente condicional:

- Si el valor ingresado por el usuario es mayor al valor que tiene la herramienta, el puntaje será de 1, adicional a esto se aplica también la Ecuación (5) para dar una bonificación, debido al ahorro que tendría en espacio en disco o en memoria con esta herramienta.

Si, por el contrario, el valor ingresado es menor que el valor que tiene la herramienta, el puntaje será de 0, debido a que no se ajusta a los requerimientos de espacio en disco y memoria del usuario.

## 2. Tipo de dato texto

Para comparar textos, se utilizó la medida de similitud de Coseno. Antes de aplicar estas medidas, se realizó un preprocesamiento que se explicó en la sección 3.2. La medida de similitud de coseno está dada por la Ecuación (2) descrita en el capítulo 2.

## 3. Tipo de dato binario

Para el cálculo de subcategorías de tipo binario, se usó la siguiente operación:

1. Primero se realizó el preprocesamiento descrito en la sección 3.2.4.
2. Se obtiene el número de funcionalidades que el usuario ingresó y que además posee la herramienta.
3. El resultado del paso 2 se divide entre la cantidad de funciones ingresadas por el usuario. Como se observa en el Ecuación (6)

$$\text{Resultado} = \frac{\text{numeroFuncionalidadesAlmacenadas}}{\text{cantidadFuncionalidadesIngresadas}} \quad (6)$$

### 3.3.2 Cálculo de puntaje por categoría

El cálculo del puntaje por categoría se determina considerando el tipo de dato y el número de subcategorías asociadas a cada una. En el caso de las categorías de costo y documentación, las cuales no cuentan con subcategorías, se realiza un único cálculo para cada una. Por otro lado, en las categorías de características técnicas y funcionalidades, se efectúa el cálculo por cada subcategoría, para luego obtener el puntaje total. A continuación, se muestra en detalle el cálculo para cada categoría.

1. **Costo.** Como esta categoría es de tipo numérico, se usa el proceso descrito en la sección 3.2.2.



2. **Documentación.** Esta categoría es de tipo texto, por esta razón se usa el proceso descrito en la sección 3.2.1.
3. **Características técnicas.** Esta categoría tiene 3 subcategorías: espacio en disco, uso de memoria y multiplataforma.

Para las subcategorías *Espacio en disco* y *Uso de memoria* se usa el proceso descrito en la sección 3.2.2, que corresponde a tipo de datos numéricos. Para la subcategoría *Multiplataforma* se usa el proceso descrito en la sección 3.2.1, que corresponde a tipo de datos texto.

Para obtener el puntaje final, se calcula un promedio, sumando estos 3 puntajes y dividiendo entre 3, como se observa en la Ecuación (7)

$$\left( \frac{(\text{puntajeEspacioDisco} + \text{puntajeUsoMemoria} + \text{puntajeMultiplataforma})}{3} \right) \quad (7)$$

#### 4. Funcionalidades

Esta categoría es diferente para algunos servicios, por lo tanto, se explicará cómo se realizó el cálculo agrupando los servicios que se califican de la misma forma.

- **Cálculo para el servicio de base de datos.** En este servicio todas las subcategorías son de tipo texto, entonces a cada una se le aplica el proceso descrito en la sección 3.2.1. Los resultados se suman y se dividen entre el número de funcionalidades, como se observa en la Ecuación (8).

$$\left( \frac{(\sum \text{puntajesSubcategorias})}{\text{numero de subcategorias}} \right) \quad (8)$$

- **Cálculo para el servicio de ETL, OLAP y reportes.** En estos servicios se tienen subcategorías con tipo de dato texto y binario, por lo tanto, el cálculo se realiza de la siguiente forma:

1. Para el cálculo del puntaje de las subcategorías de tipo de dato texto se usó el proceso descrito en la sección 3.2.1. Los puntajes obtenidos por cada subcategoría se suman.
2. Para el cálculo del puntaje de las subcategorías de tipo de dato binario se usó el proceso descrito en la sección 3.2.3, obteniendo un único valor.
3. Luego se realiza la suma de los puntajes obtenido por las subcategorías de tipo texto más el puntaje de las subcategorías de tipo binario.
4. Por último, esta suma se divide entre el número de subcategorías de tipo texto más 1 (el valor obtenido para las subcategorías de tipo binario) como se muestra en la Ecuación (9).

$$\frac{\sum \text{puntajesSubcategoriasTexto} + \text{puntajeSubcategoriasBinario}}{\# \text{subcategoriasTexto} + 1} \quad (9)$$

En la **Tabla 9** se muestra la información del número de subcategorías con valores de tipo texto y binario por cada servicio.

**Tabla 9. Número de subcategorías de tipo texto y binario por cada servicio**

Servicio	ETL	OLAP	Reportes
Número de subcategorías de tipo texto	3	3	2
Número de subcategorías de tipo binario	12	7	8

### 3.4 MÉTRICA MIXTA Y PONDERADA

La métrica mixta y ponderada propuesta en este trabajo de grado, se obtiene al multiplicar los puntajes de cada categoría, que fueron calculados en la sección 3.3 multiplicado por el porcentaje asignado por el usuario. Este porcentaje permite que el usuario pueda ponderar dando un valor a la categoría que sea más importante para él, para tomar una decisión sobre la herramienta que más se adecua a sus necesidades.

Para entender la métrica propuesta, que se aplica a cada herramienta, se definieron las siguientes siglas:

- %CU = Porcentaje asignado por el usuario para la categoría Costo.
- %DU = Porcentaje asignado por el usuario para la categoría Documentación.
- %CTU = Porcentaje asignado por el usuario para la categoría Características técnicas.
- %FU = Porcentaje asignado por el usuario para la categoría Funcionalidades.
- PCC = Puntaje de la categoría Costo.
- PCD = Puntaje de la categoría Documentación.
- PCCT = Puntaje de la categoría Características técnicas.
- PCF = Puntaje de la categoría Funcionalidades.
- MMP = Métrica mixta y ponderada por herramienta.

En la Ecuación (10) se muestra la métrica mixta y ponderada:

$$MMP = PCC * \%CU + PCD * \%DU + PCCT * \%CTU + PCF * \%FU \quad (10)$$

## 4 APLICACIÓN WEB

---

En este capítulo se realiza una descripción de la aplicación web que se construyó para permitir comparar las necesidades técnicas y económicas de una empresa con las características que ofrecen diferentes herramientas del mercado para un proyecto de BI.

### 4.1 ARQUITECTURA

La arquitectura de la aplicación web está basada en microservicios, éste es un enfoque de diseño de software que consiste en dividir una aplicación en componentes más pequeños y autónomos, llamados microservicios, que se comunican entre sí a través de una interfaz bien definida y generalmente a través de API.

El flujo de la arquitectura de la aplicación web se muestra en la Figura 1, está dada por los siguientes pasos:

- **Cliente web:** Representa la empresa que completa el formulario.
- **Microservicio que realiza el cálculo:** Este microservicio hace uso de la base de datos relacional (donde se almacenan los datos de las herramientas) y de la API TRM (Tasa Representativa del Mercado) para realizar el cálculo. Además, llama al microservicio encargado de persistir los resultados para almacenarlos en el base de datos No SQL.
- **Microservicio para persistir los resultados:** Este microservicio se encarga de almacenar los resultados junto con los datos ingresados por el cliente web.
- **Base de datos relacional:** Es la encargada de almacenar los datos de las herramientas.
- **Base de datos NoSQL:** Es la encargada de almacenar los resultados junto con los datos ingresados por el cliente web.
- **API TRM:** Es la encargada de proveer la información de tasa de cambio.

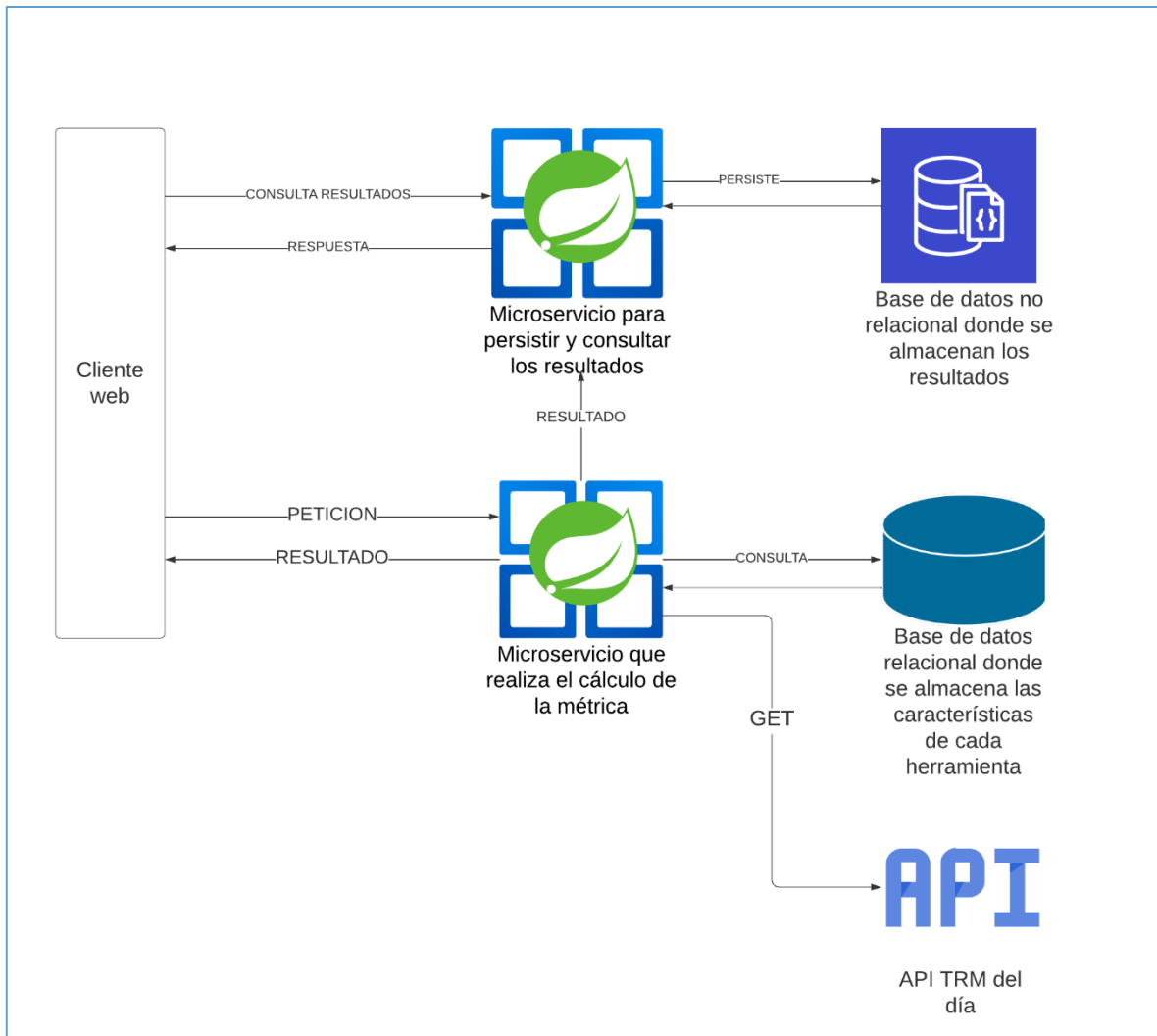


Figura 1. Arquitectura de la aplicación web

## 4.2 BACKEND

Los microservicios están escritos en lenguaje Java 11 usando el framework Spring Boot, con una arquitectura en capas.

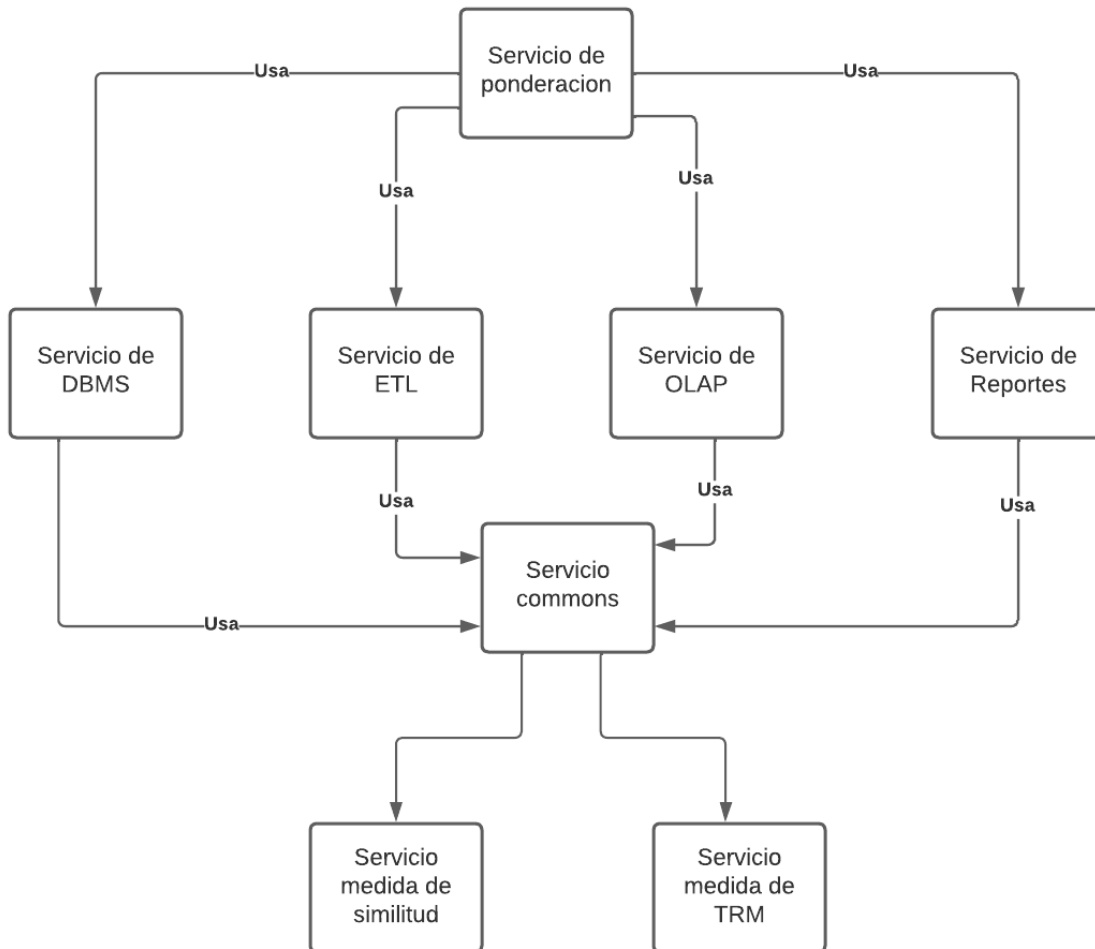
### 4.2.1 Microservicio para calcular la métrica

Este microservicio está compuesto por ocho (8) servicios internos, independientes los cuales realizan una función muy puntual. El microservicio tiene cuatro puertos de conexión, uno para recibir la petición desde la interfaz web, otro para la comunicación con la base de datos relacional, otro para la API TRM y el último para llamar al microservicio que persiste los resultados.

El *Servicio de ponderación* es el componente más externo del microservicio, encargado de recopilar los cálculos de los cuatro (4) servicios principales: *DBMS*, *ETL*, *OLAP* y *Reportes*. Cada uno de estos servicios utiliza el *Servicio commons*, el cual contiene las operaciones comunes necesarias para realizar los cálculos. A su vez, el servicio commons se basa en dos (2) servicios adicionales: el servicio de *medida de similitud textual*, donde se encuentra

la fórmula de la medida y otros métodos relacionados; y el servicio de *TRM*, utilizado para consultar la tasa de cambio de moneda del día, ya que el costo de las herramientas se encuentra en dólares.

En la **Figura 2** se observa el diagrama de servicios internos para el microservicio y la manera en que se usan.



**Figura 2. Diagrama de servicios**

#### 4.2.2 Microservicio para almacenar los resultados

Este microservicio está compuesto por un servicio interno y tiene tres puertos de conexión, uno para recibir la petición desde la interfaz web, otro para la comunicación con la base de datos noSQL y el último para recibir la petición del microservicio que calcula la métrica.

#### 4.2.3 Base de datos relacional

Se diseñó una base de datos para almacenar los valores de cada categoría y subcategoría de las herramientas seleccionadas. Para la creación de la base de datos se usó el motor relacional PostgreSQL, que presenta características de escalabilidad, flexibilidad y seguridad. En la Figura 3 se puede observar el modelo entidad relación.

La base de datos está compuesta por siete (7) tablas descritas a continuación:

- Características generales: Almacena información común entre todas las herramientas, como son: espacio en memoria, costo, documentación, espacio en disco y multiplataforma.
- Características bd: Contiene información propia de las características del servicio de base de datos.
- Características etl: Almacena información específica de las características del servicio ETL.
- Características olap: Contiene información propia de las características del servicio OLAP.
- Características reportes: Almacena información específica de las características del servicio de reportes.
- Herramientas: Contiene información de las herramientas, nombre y edición.
- Servicios: Almacena información de cada servicio.

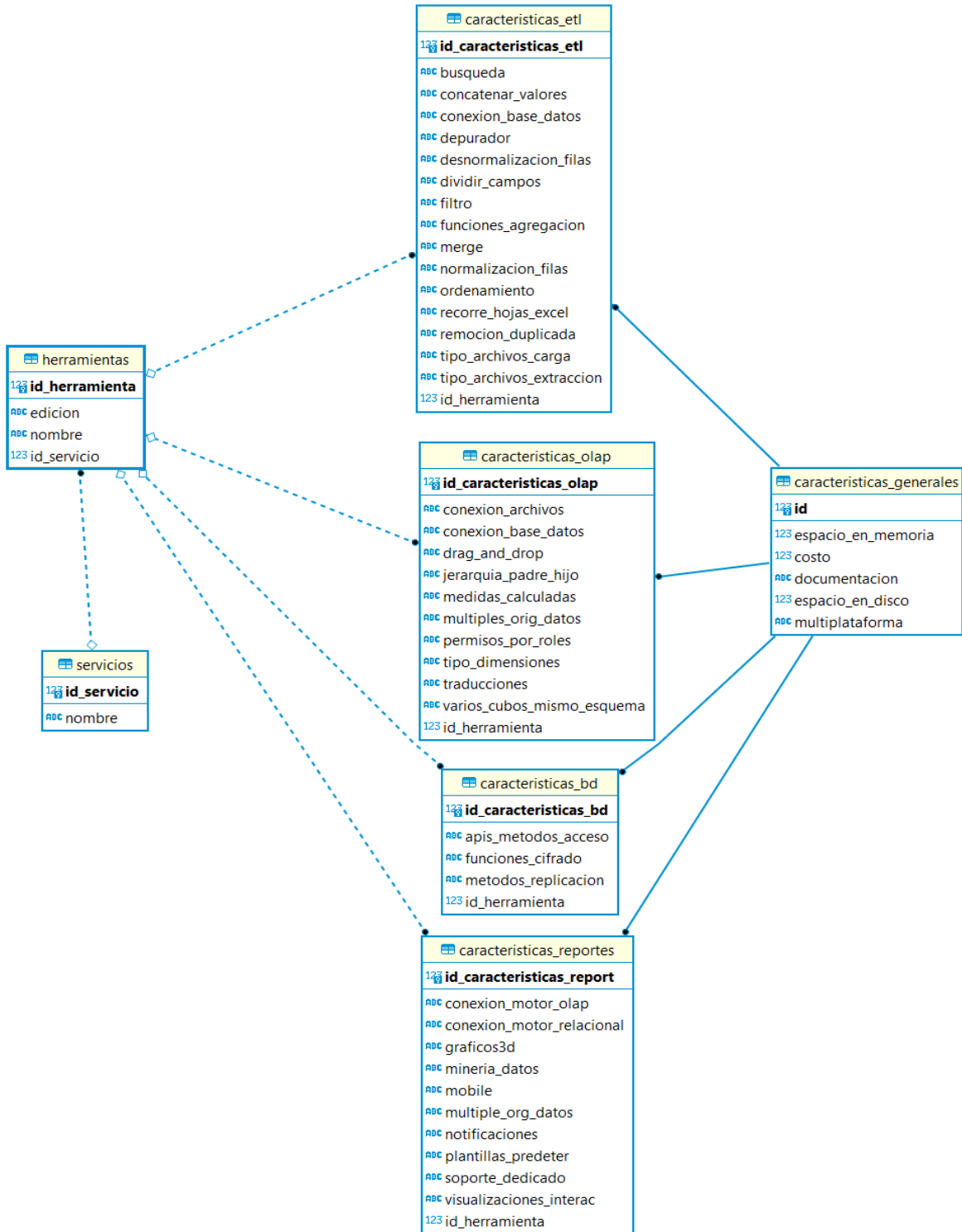


Figura 3. Modelo relacional

#### 4.2.4 Base de datos noSQL

Las bases de datos NoSQL están diseñadas específicamente para modelos de datos específicos y tienen esquemas flexibles para crear aplicaciones modernas. Las bases de datos NoSQL son ampliamente reconocidas porque son fáciles de desarrollar, por su funcionalidad y el rendimiento a escala.

En la aplicación se usó la base de datos DynamoDB de aws debido a la facilidad para conectar y administrar. Se optó por este tipo de base de datos debido a que en la aplicación fue necesario almacenar un modelo de datos complicado para modelar en una base de datos sql, mientras que esta base de datos ofrece flexibilidad en el almacenamiento de datos lo que significa que no hay una dependencia de ningún modelo específico. Esto proporciona una mayor versatilidad al sistema y facilita la integración de futuras actualizaciones y mejoras.

#### 4.3 FRONTEND

La aplicación web está compuesta por las interfaces que se presentan a continuación.

- Interfaz de resultados para mostrar los puntajes obtenidos en cada herramienta para cada servicio.

##### 4.3.1 Formulario para conocer la identidad de la empresa

En la **Figura 4** se observa el formulario que solicita a la empresa el nombre, tipo y número de documento con el fin de almacenar los datos ingresados y asociarlos a los resultados obtenidos.

The screenshot shows a web interface for identifying a company. The title is "Herramientas de inteligencia de negocios (BI)". On the left, there is explanatory text about BI tools. On the right, there is a form with the following fields and values:

- Button: Ver resultados
- Greeting: ¡Hola!
- Instruction: Por favor, ingrese su nombre y número de documento de identidad
- Field: Nombre \* (Value: Empresa - Droguería La Gloria)
- Field: Tipo de documento \* (Value: NIT)
- Field: Número de documento \* (Value: 9876543210)
- Button: Continuar

**Figura 4. Interfaz para identificar una empresa**

##### 4.3.2 Interfaz para seleccionar los servicios

En la **Figura 5** se solicita a la empresa seleccionar los servicios que le interesan para hacer la comparación de herramientas.





The image shows a user interface for selecting BI services. At the top, there is a blue header with the text "Herramientas de inteligencia de negocios (BI)". Below this, a white rounded rectangle contains the following content:

- Escoger servicios**
- Seleccione los servicios que desee para obtener herramientas sugeridas por cada uno
- SELECCIONAR TODOS
- Sistema Gestor de Base de Datos (SGBD)
- Extraer, Transformar y Cargar (ETL) Procesamiento
- Analítico en Línea (OLAP)
- Reportes
- 

**Figura 5. Interfaz para seleccionar los servicios de BI.**

### **4.3.3 Formularios para obtener los requerimientos de la organización**

En cada formulario se puede observar que se solicita ingresar los requerimientos de la empresa para los servicios de BI (SGBD, ETL, OLAP y Reportes) en cinco (5) secciones: Porcentaje por categoría, Categoría Costo, Documentación, Características técnicas y Funcionalidades.

## 1. Formulario para el servicio SGBD.

El servicio SGBD tiene las preguntas presentadas en las Figuras 6, 7, 8, 9 y 10, las cuales se agruparon en las siguientes secciones: Porcentaje por categoría, Categoría costo, Categoría documentación, Categoría características técnicas y Categoría funcionalidades. Las preguntas son similares en cada servicio para las tres primeras categorías, mientras que en la categoría funcionalidades presentan diferencias específicas.

1 SGBD      2 ETL      3 OLAP      4 REPORTES

# SGBD

Sistema Gestor de Base de Datos

Responda las siguientes preguntas con respecto a las herramientas de inteligencia de negocios, cada una está agrupada en las categorías principales (Costo, Documentación, Características técnicas y Funcionalidades).

### PORCENTAJE POR CATEGORÍA

En las siguientes cuatro (4) opciones ingrese un valor para indicar el nivel de importancia de la categoría que influirá al momento de realizar la evaluación de la herramienta.

Costo *	<input type="text" value="25"/>	%	!
Documentación *	<input type="text" value="25"/>	%	!
Características Técnicas *	<input type="text" value="25"/>	%	!
Funcionalidades *	<input type="text" value="25"/>	%	!

Figura 6. Formulario SGBD – Sección de Porcentaje por categoría

**CATEGORÍA COSTO**

---

Responda las preguntas relacionadas al costo de la herramienta.

Ingrese la cantidad aproximada de usuarios que se conectarán al motor de base de datos: \*  !

Digite la cantidad aproximada de procesadores donde será instalada la base de datos: \*  !

¿Cuánto presupuesto aproximado mensual tiene para el sistema gestor de bases de datos? \*  !

**Figura 7. Formulario SGBD – Sección de Categoría Costo**

**CATEGORÍA DOCUMENTACIÓN**

---

¿Qué tipo de documentación prefiere leer sobre las herramientas de bases de datos? (Seleccione una o múltiples opciones) ?

- Seleccionar todo
- Documentación oficial en línea
- Foros, tutoriales

**Figura 8. Formulario SGBD – Sección de Categoría Documentación**

**CATEGORÍA CARACTERÍSTICAS TÉCNICAS**

Indique el espacio aproximado en disco disponible para la instalación: **8 GB** <sup>!</sup>

GB

Indique el espacio en memoria RAM disponible para la ejecución de la base de datos: **8 GB** <sup>!</sup>

GB

Seleccione los sistemas operativos en los cuales le interesa instalar la base de datos (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- Linux
- Mac
- Windows

**Figura 9. Formulario SGBD – Sección de Categoría Características técnicas**

**CATEGORÍA FUNCIONALIDADES**

¿Qué métodos de acceso le interesa elegir para conectarse a su base de datos? (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- ADO.NET (Active Data Objects.NET)
- C API
- C Driver
- C++ Driver
- JDBC (Java Database Connectivity)
- Node.js Driver
- ODBC (Open Database Connectivity)
- pgJDBC driver
- Python Driver
- SQLserver Driver

¿Qué funciones de cifrado necesita para la base de datos? (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- Cifrado de datos transparente (TDE)
- Cifrado simétrico: Estándar de cifrado avanzado (AES)
- Funciones hash: MD5 (Message Digest Algorithm 5)
- Funciones hash: SHA-256 (Secure Hash Algorithm 256-bit)

**Figura 10. Formulario SGBD – Sección de Categoría Funcionalidades**

## 2. Formulario para el servicio ETL.

Para el servicio ETL se tienen las preguntas que se observan en la **Figura 11** para la categoría funcionalidades.

### CATEGORÍA FUNCIONALIDADES

Seleccione los tipos de archivos de donde va a extraer la información (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- CSV
- Excel
- JSON
- Microsoft Access
- XML

Debe seleccionar al menos una opción

Seleccione los tipos de archivos en los cuales necesita cargar la información (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- Excel
- CSV
- JSON

Debe seleccionar al menos una opción

¿Con cuáles bases de datos requiere conexión? (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- Amazon Redshift
- MySQL
- Oracle
- PostgreSQL
- SAP HANA
- SQL Server

Debe seleccionar al menos una opción

Seleccione las opciones que desea encontrar en el software (Seleccione una o múltiples opciones)

- Seleccionar todo
- Actualización programada de datos
- Búsqueda
- Concatenar valores
- Depurador con puntos de interrupción
- Desnormalización de filas
- Dividir campos
- Filtro de filas
- Funciones de agregación
- Merge / Fusión de datos
- Normalización de filas
- Ordenamiento
- Recorrer hojas excel
- Remover duplicados

Debe seleccionar al menos una opción

Listo

**Figura 11. Formulario ETL – Sección de Categoría Funcionalidades**

### 3. Formulario para el servicio OLAP.

Para el servicio OLAP se tienen las preguntas que se observan en la **Figura 12** para la categoría funcionalidades.

#### CATEGORÍA FUNCIONALIDADES

¿Con cuáles bases de datos requiere conexión? (Seleccione una o múltiples opciones) <sup>1</sup>

- Seleccionar todo
- Apache Derby
- Aplicaciones empresariales como Google Analytics, Google Docs y Java Message Services
- Firebird
- Hsqldb
- IBM DB2
- JTDS
- MariaDB
- Microsoft Access
- MySQL
- Pervasive se conecta a BD no SQL
- PostgreSQL
- SAP HANA
- SQL Server Analysis Services
- SQL Server
- Sybase

Debe seleccionar al menos una opción

¿Desde cuáles tipos de archivos requiere conexión? (Seleccione una o múltiples opciones) <sup>1</sup>

- Seleccionar todo
- CVS
- Excel
- JSON
- PDF
- XML

Debe seleccionar al menos una opción

¿Cuáles tipos de dimensiones principales requiere? (Seleccione una o múltiples opciones) <sup>1</sup>

- Seleccionar todo
- Canal
- Clientes
- Cuantitativo
- Cuenta
- Escenario
- Geografía
- Lista de materiales
- Moneda
- Normal
- Organización
- Productos
- Promoción
- Tarifas
- Tiempo
- Utilidad

Debe seleccionar al menos una opción

Seleccione las opciones que desea encontrar en el software (Seleccione una o múltiples opciones)

- Seleccionar todo
- Configurar permisos de roles
- Creación de varios cubos en un mismo esquema
- Drag and drop
- Jerarquía padre/hijo
- Medidas calculadas
- Múltiples orígenes de datos en un mismo esquema
- Traducciones

Debe seleccionar al menos una opción

Listo

Figura 12. Formulario OLAP – Sección de Categoría Funcionalidades

#### 4. Formulario para el servicio Reportes.

Para el servicio Reportes se tienen las preguntas que se observan en la **Figura 13** para la categoría funcionalidades.

### CATEGORÍA FUNCIONALIDADES

¿Con cuáles bases de datos requiere conexión? (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- MySQL
- Oracle
- PostgreSQL
- Redshift
- SAP ERP
- SAP HANA
- SparkSQL
- SQLite
- SQL Server

Debe seleccionar al menos una opción

¿Con cuáles motores OLAP requiere conexión? (Seleccione una o múltiples opciones) <sup>!</sup>

- Seleccionar todo
- Jaspersoft OLAP
- Pentaho Analysis
- Sql server Analysis Services Database
- Teradata OLAP Data

Debe seleccionar al menos una opción

Seleccione las opciones que desea encontrar en el software (Seleccione una o múltiples opciones)

- Seleccionar todo
- Conexión a varios orígenes de datos en un mismo reporte
- Gráficos 3D
- Minería de datos
- Notificaciones
- Plantillas predeterminadas para la realización de informes
- Soporte dedicado
- Visualización en mobile
- Visualizaciones interactivas

Debe seleccionar al menos una opción

Listo

**Figura 13. Formulario Reportes – Sección de Categoría Funcionalidades**

#### 4.3.4 Interfaz de resultados

Las siguientes figuras presentan un ejemplo de los resultados para una empresa de la puntuación que se obtuvo en cada herramienta para cada servicio SGBD, ETL, OLAP y/o Reportes.

En la **Figura 14** se muestra información que detalla cómo interpretar los resultados obtenidos para cada una de las empresas.

Herramientas de inteligencia de negocios (BI)

RESULTADOS DE  
**FERRAGRO**

Seleccione una opción\*  
FERRAGRO

Leer información

Los resultados a continuación contienen los datos que has ingresado previamente para comparar con los datos de las herramientas. Además, observarás un listado de herramientas que corresponden a cada servicio seleccionado, y sobre cada una se encuentra la siguiente información:

- **Nombre de la herramienta.**
- **Puntaje:** Un valor de cien (100) indica que la herramienta cumple con los requerimientos ingresados por el usuario, al igual que en la categoría cuando el puntaje es el mismo que el porcentaje ingresado por el usuario.
- **Bonificación:** Es un valor que se adiciona al puntaje de la herramienta, cuando se obtiene un ahorro en la categoría Costo o Características técnicas debido al espacio en disco o en memoria. Por ejemplo, se obtiene una bonificación cuando se posee más espacio en disco del requerido por la herramienta para su instalación.
- **Categorías:** Se refiere a las siguientes cuatro (4) categorías principales: Costo, Documentación, Características técnicas y Funcionalidades, las cuales agrupan subcategorías.
- **Subcategorías:** Es cada subcategoría agrupada en las categorías principales. Ejemplo: espacio en disco y espacio en memoria son subcategorías de Características técnicas.
- **Valores ingresados por el usuario:** Se refiere a los datos ingresados en los formularios iniciales.
- **Datos de la herramienta:** Es información de la herramienta sobre cada categoría y subcategoría, sirven para comparar con los datos ingresados por el usuario.
- **Porcentaje:** Es el porcentaje ingresado por el usuario que representa el nivel de importancia que tiene cada categoría.
- **Resultado:** Es el valor obtenido por cada subcategoría. Un resultado de uno (1) en la subcategoría representa que cumple con el requerimiento solicitado, por el contrario si es un valor de cero (0) entonces indica que no cumple, además, si el valor está entre cero (0) y uno (1) significa que el requerimiento se aproxima en cualquiera de esos valores.

**Figura 14. Resultados - Información de ayuda**

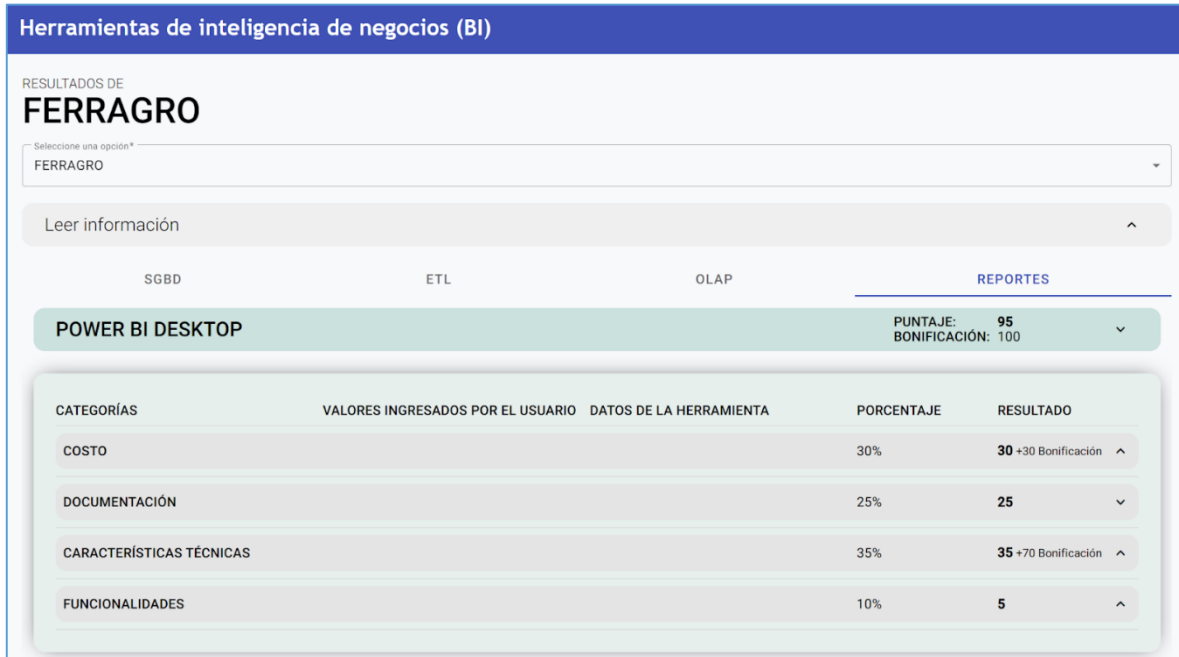


La **Figura 15** muestra los puntajes totales y las bonificaciones obtenidas por cada herramienta. El cálculo de la bonificación se describe en detalle en el Capítulo 3, en la sección 3.3.1.

Herramientas de inteligencia de negocios (BI)			
RESULTADOS DE FERRAGRO			
Seleccione una opción* FERRAGRO			
Leer información			
SGBD	ETL	OLAP	REPORTES
POWER BI DESKTOP			PUNTAJE: 95 BONIFICACIÓN: 100
POWER BI PRO			PUNTAJE: 90 BONIFICACIÓN: 126
POWER BI PREMIUM			PUNTAJE: 90 BONIFICACIÓN: 83
TABLEAU REPORT CREATOR COMERCIAL			PUNTAJE: 90 BONIFICACIÓN: 34
PENTAHO COMMUNITY			PUNTAJE: 95 BONIFICACIÓN: 854
PENTAHO ENTERPRISE			PUNTAJE: 91 BONIFICACIÓN: 854
JASPERSOFT REPORT COMMUNITY			PUNTAJE: 71 BONIFICACIÓN: 30

**Figura 15. Resultados - Puntaje de las herramientas**

La **Figura 16** muestra la puntuación de las categorías correspondiente a cada herramienta. Si la puntuación coincide con el porcentaje ingresado por el usuario, significa que esa categoría cumple con los requisitos establecidos. Para las subcategorías, se utiliza una escala del 0 al 1, donde 1 indica un cumplimiento completo del requisito, 0 indica que no se cumple y los valores entre 0 y 1 reflejan un grado de adecuación dentro de ese rango.



**Figura 16. Resultados - Puntaje por categoría de cada herramienta.**

## 5 EVALUACIÓN

En este capítulo se describe el proceso de evaluación para medir el grado en que se adecuan las herramientas sugeridas con respecto a la puntuación proporcionada por la aplicación web sobre un conjunto de herramientas on-premise a las necesidades de dos empresas.

Primero, se realiza una descripción acerca del objetivo de la evaluación y la normativa ISO/IEC 25010, luego se detallan los test y su aplicación, posteriormente una explicación del proceso de evaluación para finalizar con los resultados obtenidos.

### 5.1 UTILIDAD DE LA APLICACIÓN WEB

#### 5.1.1 Objetivo de la evaluación

Medir la utilidad de la aplicación web con respecto a las herramientas sugeridas de inteligencia de negocios en los servicios de Sistema Gestor de Base de Datos (SGBD), Extracción - Transformación - Carga (ETL), Procesamiento Analítico en Línea (OLAP) y Reportes.

#### 5.1.2 Normatividad

La norma ISO/IEC 25010 establece un conjunto de requisitos y recomendaciones para la evaluación de la calidad del software y sistemas. En particular, se enfoca en la calidad del producto de software, proporcionando un modelo para evaluar la calidad del software desde la perspectiva del usuario final.

Dentro de la norma, la característica de usabilidad se define como la capacidad de un software para ser comprendido, aprendido, usado y atractivo para el usuario, en un contexto específico de uso. Esta característica se divide en subcaracterísticas, entre las cuales se elige para esta evaluación la *Estética de la interfaz de usuario* permitiendo evaluar si la interfaz es agradable y genera satisfacción en la interacción con el usuario (ver **Tabla 10**).

**Tabla 10. Característica y subcaracterísticas de usabilidad: Estética de la interfaz de usuario**

USABILIDAD	
SUBCARACTERÍSTICAS	DESCRIPCIÓN
Estética de la interfaz de usuario	Capacidad de la interfaz de usuario de agrandar y satisfacer la interacción con el usuario.

#### 5.1.3 Implementación del test de utilidad

El test de utilidad de la aplicación web está compuesto por las siguientes cinco (5) secciones:

1. **Primera:** se indica cual es el objetivo de la evaluación, además de solicitar algunos datos de la empresa como: Nombre de la empresa, Nombre del usuario, y la Ocupación del usuario.

2. **Segunda:** se realizan dos preguntas generales acerca de la facilidad de uso de aplicación para obtener los resultados y el entendimiento de la puntuación total de las herramientas en cada servicio.
3. **Tercera:** preguntas enfocadas al proceso para ingresar datos en los formularios de cada servicio.
4. **Cuarta:** preguntas relacionadas a la visualización de los resultados de cada herramienta de BI sugerida a la empresa.
5. **Quinta:** preguntas abiertas por si el usuario tiene alguna sugerencia adicional sobre el proceso del ingreso de datos y los resultados obtenidos.

Para cada una de las preguntas se usa la escala de calificación de Likert (Ver **Tabla 11**, ya que permite establecer el nivel de conformidad de quien es encuestado y es la escala más empleada en cuestionarios y encuestas para la investigación.

**Tabla 11. Calificación de los cuestionarios según la escala de Likert**

Valor numérico	Conformidad
5	Totalmente de acuerdo
4	Parcialmente de acuerdo
3	Indiferente
2	Algo en desacuerdo
1	Totalmente en desacuerdo

#### 5.1.4 Aplicación del test de utilidad

El Test de utilidad de la aplicación web es aplicado a los siguientes representantes de cada empresa:

- Johan Daniel Pinto Rebolledo, representante legal del Club Independiente Argelia.
- Yesid Muñoz Ome, representante de empresa Ferragro.

Debido a la poca disponibilidad de los representantes, se realiza una sesión de socialización con cada uno sobre el funcionamiento de la aplicación web y el significado de los puntajes en cada herramienta. Posteriormente, se realiza el proceso de evaluación por parte de cada representante sobre la usabilidad de aplicación.

Los puntajes que se evalúan de las herramientas sobre cada servicio corresponden a las siguientes categorías:

- Costo.
- Documentación
- Características técnicas
- Funcionalidades

El puntaje de cada categoría contribuye a la suma del puntaje total de la herramienta.

## 5.2 ADECUACIÓN HERRAMIENTAS SUGERIDAS POR LA APLICACIÓN WEB

### 5.2.1 Objetivo de la evaluación

El objetivo de esta evaluación es evaluar el grado en que se adecuan las herramientas sugeridas por una aplicación web desarrollada (que usa la métrica propuesta), a las necesidades técnicas y económicas de dos empresas; por medio de dos test aplicados a dos profesionales con conocimiento en proyectos de inteligencia de negocios, y a dos empresas.

### 5.2.2 Normativa

Teniendo en cuenta nuevamente la norma ISO/IEC 25010 dentro de la característica de usabilidad, se toma para esta evaluación la subcaracterísticas de *Reconocibilidad de la adecuación* permitiendo evaluar si los resultados obtenidos por la aplicación se adecuan a las necesidades de la empresa.

**Tabla 12. Característica y subcaracterísticas de usabilidad: Reconocibilidad de la adecuación**

USABILIDAD	
SUBCARACTERÍSTICAS	DESCRIPCIÓN
Reconocibilidad de la adecuación	Capacidad del producto que permite al usuario entender si el software es adecuado para sus necesidades.

### 5.2.3 Implementación del test de adecuación de herramientas

El test para medir el grado de adecuación de herramientas de inteligencia de negocios (BI) sugeridas por una aplicación web, a las necesidades técnicas y económicas de una empresa; está compuesto en seis (6) secciones:

- **Primera:** indica cual es el objetivo de la evaluación.
- **Segunda:** solicita algunos datos de los evaluadores como: Nombre completo, profesión, Nombre de la empresa en la que trabaja y Cargo desempeñado en la empresa.
- **Tercera:** solicita el nombre de la empresa que va a evaluar.
- **Cuarta:** preguntas sobre el puntaje de cada categoría con respecto a cada servicio.
- **Quinta:** preguntas sobre el puntaje total de las herramientas en cada servicio.
- **Sexta:** preguntas abiertas con el objetivo de recopilar sugerencias adicionales sobre los puntajes obtenidos de las herramientas en cada una de sus categorías y subcategorías.

### 5.2.4 Aplicación del test de adecuación de herramientas

El Test de adecuación de herramientas de BI sugeridas a una empresa es aplicado a dos profesionales con conocimiento en proyectos de inteligencia de negocios:

- David Antonio Fuentes Vargas, ingeniero de datos.
- Daisy Yisel Meneses Lopez, ingeniera de datos.

Debido a la poca disponibilidad de los ingenieros, se envía un material previo que explicaba el proceso para puntuar las herramientas de BI según los requerimientos de la empresa, luego se realiza una sesión de socialización sobre los resultados obtenidos para cada

empresa. Posteriormente, se realiza el proceso de evaluación por parte de cada profesional sobre el grado de adecuación de herramientas de BI.

Los puntajes que se evalúan de las herramientas sobre cada servicio corresponden a las siguientes categorías:

- Costo
- Documentación
- Características técnicas
- Funcionalidades

El puntaje de cada categoría contribuye a la suma del puntaje total de la herramienta.

### 5.3 PROCESO DE EVALUACIÓN

Para responder a la pregunta “¿Qué tan satisfecho está el usuario?”, se hace uso de la métrica de calidad de *Nivel de satisfacción*, definida en [26]. Para medir esta métrica se utiliza la siguiente fórmula (11):

Donde:

$$x = \frac{A}{B} \tag{11}$$

A: Representa el número de preguntas con respuestas satisfactorias.

B: Representa la cantidad de preguntas realizadas en el cuestionario aplicado,  $B > 0$ .

El valor deseado de X está en el rango:  $0 \leq X \leq 1$ , donde 1 es el mayor valor de satisfacción.

Para tener una interpretación más clara del resultado de X se utiliza una escala de medición de 0 a 10, donde el valor esperado para los resultados de la evaluación es 10. En la **Tabla 13** se muestra la clasificación de escalas, con su nivel de puntuación y su respectivo grado de satisfacción.

**Tabla 13. Grado de satisfacción según niveles de puntuación.**

Escala de medición	Niveles de puntuación	Grado de satisfacción
8,75 - 10	Cumple con los requisitos	Muy satisfactorio
5 - 8,74	Aceptable	Satisfactorio
2,75 - 4,9	Mínimamente aceptable	Insatisfactorio
0 - 2,74	Inaceptable	

### 5.4 RESULTADOS OBTENIDOS DEL TEST DE UTILIDAD DE LA APLICACIÓN WEB

El resultado de la evaluación de satisfacción se basa en el número de respuestas de la empresa en cada valor numérico de la escala de Likert. En la **Tabla 14** se presenta por cada empresa el total de preguntas evaluadas y la cantidad de preguntas por cada calificación dada.

**Tabla 14. Cantidad de preguntas evaluadas en los test**

Empresa	Cantidad preguntas	Calificación				
		5	4	3	2	1
Ferragro	37	32	5	0	0	0
Club Deportivo Independiente Argelia	37	0	33	4	0	0
<b>Total</b>	<b>74</b>	<b>32</b>	<b>38</b>	<b>4</b>	<b>0</b>	<b>0</b>

De acuerdo con [26] para realizar el análisis de calidad del producto software se utiliza una matriz de calidad, la cual permite evaluar el producto de manera completa y concisa. Se adaptó la matriz al propósito de evaluación con la característica elegida denominada *Usabilidad* y la subcaracterística *Estética de la interfaz de usuario*.

En la **Tabla 15** se presenta una adaptación de la matriz de calidad donde se muestran las fórmulas empleadas en la evaluación de la métrica de Nivel de Satisfacción. Debido a que las respuestas de los representantes son *Totalmente de acuerdo*, con calificación 5, *Parcialmente de acuerdo*, con calificación 4 e *Indiferente*, con calificación 3, se definen las fórmulas para medir la proporción de estos valores en la escala de Likert. El valor  $X_1$ , es la proporción de la cantidad de preguntas que son calificadas con 5 ( $A_1$ ) sobre el número total de preguntas aplicadas en los formularios ( $B$ ); el valor  $X_2$ , es la proporción de las preguntas que son calificadas con 4 ( $A_2$ ) sobre el número total de preguntas aplicadas en los formularios ( $B$ ); el valor  $X_3$ , es la proporción de las preguntas que son calificadas con 3 ( $A_3$ ) sobre el número total de preguntas aplicadas en los formularios ( $B$ ). Los valores obtenidos en las fórmulas se presentan en la columna Valor obtenido por variable y el Valor total obtenido en la métrica  $X$  correspondería a la sumatoria de valores  $X_1$ ,  $X_2$  y  $X_3$ . El Valor deseado para la medida es 1, el cual se traduce en un porcentaje que representa el nivel de satisfacción.

**Tabla 15. Matriz de calidad**

Métrica	Fórmula	Valor obtenido por variable	Valor total obtenido	Valor deseado
Nivel de satisfacción	$X_1 = A_1/B$	$A_1 = 32$ $B = 74$ $X_1 = 0.43$	$X = X_1 + X_2 + X_3$ $X = 0.99$	1
	$X_2 = A_2/B$	$A_1 = 38$ $B = 74$ $X_2 = 0.51$		
	$X_3 = A_3/B$	$A_3 = 4$ $B = 74$ $X_3 = 0.05$		

En la **Tabla 15** se observa que en el 43% ( $X_1$ ) de las preguntas de las empresas, manifiestan estar *Totalmente de acuerdo*, en el 51% ( $X_2$ ) estar *Parcialmente de acuerdo* debido a que los representantes de las empresas manifestaron que algunas preguntas fueron demasiado técnicas y requerían apoyo de expertos en BI y en el 5% ( $X_3$ ) es *Indiferente*, obteniendo un valor total de 99% ( $X$ ). Por lo tanto, con base en la **Tabla 13**,

para el *valor total obtenido* el nivel de puntuación corresponde a *Cumple con los requisitos* y el grado de satisfacción es *Muy satisfactorio*.

### 5.5 RESULTADO OBTENIDO DEL TEST DE ADECUACIÓN DE HERRAMIENTAS DE BI SUGERIDAS A UNA EMPRESA

El resultado de la evaluación del grado de adecuación se basa en el número de respuestas de la empresa en cada valor numérico de la escala de Likert. En la **Tabla 16** se presenta por cada empresa el total de preguntas evaluadas y la cantidad de preguntas por cada calificación dada.

**Tabla 16. Cantidad de preguntas evaluadas en el test para la empresa Ferragro**

Profesionales	Cantidad preguntas	Calificación				
		5	4	3	2	1
David Antonio Fuentes Vargas – Ferragro.	20	20	0	0	0	0
Daisy Yisel Meneses Lopez - Ferragro.	20	20	0	0	0	0
<b>Total</b>	<b>40</b>	<b>40</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>

**Tabla 17. Matriz de calidad – Ferragro**

Métrica	Fórmula	Valor obtenido por variable	Valor total obtenido	Valor deseado
Grado de adecuación	$X1 = A1/B$	A1 = 40 B = 40 X1 = 1	X = X1 X = 1	1

En la **Tabla 17** se observa que en el 100% (X1) de las preguntas a los profesionales, manifiestan estar *Totalmente de acuerdo*, en el 0% (X2) estar *Algo de acuerdo* y en el 0% (X3) es *Indiferente*, obteniendo un valor total de 100% (X). Por lo tanto, con base en la **Tabla 13**, para el *valor total obtenido* el nivel de puntuación corresponde a *Cumple con los requisitos* y el grado de adecuación es *Muy satisfactorio*.

**Tabla 18. Cantidad de preguntas evaluadas en el test para la empresa Club Independiente Argelia.**

Profesionales	Cantidad preguntas	Calificación				
		5	4	3	2	1
David Antonio Fuentes Vargas - Club Independiente Argelia	20	20	0	0	0	0
Daisy Yisel Meneses López - Club Independiente Argelia	20	17	3	0	0	0
<b>Total</b>	<b>40</b>	<b>37</b>	<b>3</b>	<b>0</b>	<b>0</b>	<b>0</b>



**Tabla 19. Matriz de calidad – Club Independiente Argelia.**

Métrica	Fórmula	Valor obtenido por variable	Valor total obtenido	Valor deseado
Grado de adecuación	$X1 = A1/B$	A1 = 37 B = 40 X1 = 0.925	X = X1 + X2 X = 0.925 + 0.075	1
	$X2 = A2/B$	A2 = 3 B = 40 X2 = 0.075		

En la **Tabla 19** se observa que en el 92.5% (X1) de las preguntas a los profesionales, manifiestan estar *Totalmente de acuerdo*, en el 7.5% (X2) estar *Algo de acuerdo* y en el 0% (X3) es *Indiferente*, obteniendo un valor total de 100% (X). Por lo tanto, con base en la **Tabla 13**, para el *valor total obtenido* el nivel de puntuación corresponde a *Cumple con los requisitos* y el grado de adecuación es *Muy satisfactorio*.

Siguiendo las sugerencias de los evaluadores, es importante considerar los siguientes aspectos para obtener resultados óptimos:

- Popularidad de la herramienta de BI: Es importante tener en cuenta la popularidad y el nivel de adopción de la herramienta en el mercado nacional e internacional. Esto brinda una idea de su uso generalizado y confianza en la industria.
- Utilizar colores, como el rojo, para indicar herramientas con menor puntaje: Para facilitar la visualización y comprensión de los resultados, se recomienda emplear colores distintivos. Asignar el color rojo a las herramientas con menor puntaje resalta claramente aquellas opciones que pueden ser menos favorables en términos de rendimiento.
- Adaptar la aplicación a nuevas herramientas de BI: Con la evolución constante de las herramientas de BI, es fundamental mantenerse actualizado y adaptar la aplicación para incorporar nuevas opciones conforme surjan en el mercado. Esto garantiza que la evaluación se base en las últimas herramientas disponibles y se mantenga relevante en un entorno en constante cambio.
- Organizar las herramientas mediante la agrupación por proveedor, teniendo en cuenta los puntajes asignados a cada herramienta de BI según su servicio específico (SGBD, ETL, OLAP, REPORTES). De esta manera, se obtendrá una visión clara y ordenada de las opciones disponibles facilitando la toma de decisiones y la comparación entre las soluciones ofrecidas por cada proveedor.

## 6 CONCLUSIONES Y TRABAJO FUTURO

---

### 6.1 CONCLUSIONES

- Este trabajo de grado se centra en la selección y evaluación de herramientas de inteligencia de negocios, así como en el desarrollo de una aplicación web para proporcionar recomendaciones personalizadas a las empresas con base en sus requerimientos técnicos y económicos. Esta investigación identifica las categorías y subcategorías más relevantes de un conjunto de herramientas para los servicios de BI (SGBD, ETL, OLAP, Reportes) y seleccionar las que mejor se ajustaban a los objetivos del proyecto.
- La métrica de similitud definida es de naturaleza mixta, dado que considera para su cálculo tipos de datos numéricos, de texto y binarios. Además, es ponderada porque contempla la asignación por parte del usuario de un porcentaje a cada categoría según su nivel de importancia. Para el cálculo de los puntajes de cada subcategoría se utiliza una escala de 0 a 1, y a partir de estos, se calcula el puntaje para cada categoría, obteniendo finalmente el puntaje total de la herramienta evaluada, el cual se presenta en una escala de 0 a 100 para una mejor interpretación del resultado por parte del usuario.
- Para asignar el puntaje a las categorías y subcategorías con valores de texto, se utiliza la medida de similitud coseno, debido a que en la revisión del estado del arte se encontró que esta presenta un mejor rendimiento en tareas de comparación, clasificación y agrupación de textos. Al aplicar esta medida de similitud en la métrica propuesta, se encontró que la puntuación que se asignaba a la comparación entre los valores ingresados por el usuario y los almacenados en la aplicación fue la adecuada.
- En la puntuación de las subcategorías con valores numéricos como costo, espacio en disco y espacio en memoria, se asigna un puntaje de uno si se cumple con el requerimiento del usuario, de lo contrario un cero. Sin embargo, cuando el requerimiento del usuario es mayor al que ofrece la herramienta, se asigna una bonificación de acuerdo a la diferencia que haya entre los dos valores, dado que esto representa un ahorro para el usuario.
- En la puntuación de las subcategorías con valores de tipo binario, se divide el número de funcionalidades requeridas por el usuario y que tiene la herramienta, entre el número de funcionalidades requeridas por el usuario, obteniendo un único puntaje entre cero y uno, donde cero indica que la herramienta no cumple con las funcionalidades requeridas, uno que cumple con todas, y un valor entre cero y uno que cumple con algunas de estas.
- La aplicación web desarrollada se basa en una arquitectura de microservicios, lo que facilita la adición de nuevas funcionalidades, brindando flexibilidad y escalabilidad al sistema. Además, se optó por utilizar una base de datos NoSQL, debido a que modelar en una base de datos relacional los resultados de cada herramienta junto a los datos ingresados por el usuario era bastante complejo, lo cual se simplifica usando una base de datos NoSQL.

- Después de evaluar los resultados y recibir la retroalimentación de los profesionales con conocimiento en proyectos de inteligencia de negocios, se concluye que existe un alto grado de adecuación de las herramientas sugeridas por la aplicación web a las necesidades técnicas y económicas de las empresas que hicieron uso de esta. En el caso de la empresa Ferragro, se alcanzó un porcentaje del 100% para *Totalmente de Acuerdo*, y de la empresa Club Independiente Argelia se obtuvo un 92.5% para *Totalmente de Acuerdo*.
- En la evaluación de la utilidad de la aplicación web, se observó que los representantes de las empresas mostraron un 43% estar *Totalmente de acuerdo* y un 51% estar *Parcialmente de acuerdo*. Estos resultados reflejan que para responder de forma más acertada las preguntas planteadas en la aplicación web, se hace necesario el apoyo de personal con conocimiento técnico de la empresa y de un experto en inteligencia de negocios.
- En proyectos de esta naturaleza, que involucran la evaluación de numerosas herramientas, se recomienda enfocarse en soluciones para entornos homogéneos, es decir, basadas en nube o en el sitio (on-premise). De esta forma se busca simplificar la tarea de comparación de costos, considerando la diversidad de modalidades de pago existentes, como la variación en función de la cantidad de usuarios, los modelos de pago por uso y la especificación del número de procesadores.
- La confirmación de la presencia de ciertas funcionalidades en las herramientas evaluadas fue una tarea difícil, debido a que, en muchos casos, la información no estaba disponible en la documentación o en foros, teniendo que considerar la opción de instalar la herramienta para su revisión. No obstante, esto implicaba la adquisición de la licencia, ya que algunas funcionalidades solo estaban disponibles en versiones de pago. Esta limitación no permitió validar la existencia de ciertas funcionalidades, lo que condujo a la exclusión de varias herramientas en este proyecto.
- Para el despliegue de la aplicación web se utilizó tecnología basada en la nube con una suscripción gratuita con créditos limitados para el uso de los servicios de bases de datos. Por esta limitante, fue necesario realizar varios despliegues (uno de ellos para la etapa de evaluación por parte de los profesionales con conocimiento en inteligencia de negocios), usando diferentes cuentas con este tipo de suscripción, lo cual requirió una mayor cantidad de esfuerzo para esta actividad. Por lo tanto, se recomienda en este tipo de soluciones en la nube disponer de un presupuesto para adquirir una suscripción con una mayor cantidad de créditos que permita un mayor uso de los servicios de base de datos en un tiempo más amplio.

## 6.2 TRABAJO FUTURO

- Incluir un módulo de administración en la aplicación web que permita la fácil incorporación de nuevas herramientas de BI, para evitar tener que agregarlas directamente en la base de datos.
- Incluir una nueva funcionalidad en la aplicación web para que varios especialistas en inteligencia de negocios puedan asignar diferentes ponderaciones a las categorías de cada servicio de BI (SGBD, ETL, OLAP, Reportes), e ingresar sus propios requerimientos. Esta funcionalidad permitirá a cada especialista seleccionar los requerimientos de la empresa en cada servicio según su experticia.

- Incluir en la métrica mixta y ponderada propuesta nuevas categorías, como la popularidad de las herramientas de BI, la cual fue sugerida por uno de los profesionales con conocimiento en inteligencia de negocios. Esto permitiría evaluar el nivel de adopción de las herramientas en el mercado nacional e internacional, brindando una perspectiva más completa en la toma de decisiones.
- Proponer una métrica mixta y ponderada para herramientas de inteligencia de negocios basada en la nube, con sus correspondientes categorías y subcategorías, incluyendo aspectos como la escalabilidad, la disponibilidad y las diferentes alternativas de suscripción disponibles.

## 7 BIBLIOGRAFÍA

---

- [1] E. Ayoubi and S. Aljawarneh, "Challenges and opportunities of adopting business intelligence in SMEs: Collaborative model," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Oct. 2018. doi: 10.1145/3279996.3280038.
- [2] K. Sharma, A. Shetty, A. Jain, and R. K. Dhanare, "A Comparative Analysis on Various Business Intelligence (BI), Data Science and Data Analytics Tools," in *2021 International Conference on Computer Communication and Informatics, ICCCI 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021. doi: 10.1109/ICCCI50826.2021.9402226.
- [3] N. Liliana, C. Cristancho, and M. A. Contreras Buitrago, "UN SALVAVIDAS A LAS MIPYMES DE COLOMBIA EN TIEMPOS DEL COVID-19," 2020.
- [4] E. Papachristodoulou, M. Koutsaki, and E. Kirkos, "Business intelligence and SMEs: Bridging the gap," 2017. [Online]. Available: <https://ojs.hh.se/>
- [5] R. Singh and S. Singh, "Text Similarity Measures in News Articles by Vector Space Model Using NLP," *Journal of The Institution of Engineers (India): Series B*, vol. 102, no. 2, pp. 329–338, Apr. 2021, doi: 10.1007/s40031-020-00501-5.
- [6] K. Kasemsap, "The Fundamentals of Business Intelligence," *International Journal of Organizational and Collective Intelligence*, vol. 6, no. 2, pp. 12–25, Apr. 2016, doi: 10.4018/ijoci.2016040102.
- [7] M. Murugesan, K. Karthikeyan, and K. Karthikeyan Assistant Professor, "Business Intelligence Market Trends and Growth in Enterprise Business," 2006, [Online]. Available: <http://www.ijritcc.org>
- [8] "Qué es una base de datos relacional | Oracle Colombia," 2021. <https://www.oracle.com/co/database/what-is-a-relational-database/> (accessed Oct. 25, 2021).
- [9] "Uso de cubos OLAP para análisis avanzado | Microsoft Docs," 2021. <https://docs.microsoft.com/es-es/system-center/scsm/olap-cubes-overview?view=sc-sm-2019> (accessed Oct. 25, 2021).
- [10] W. Yeoh and A. Popovič, "Extending the understanding of critical success factors for implementing business intelligence systems," *J Assoc Inf Sci Technol*, vol. 67, no. 1, pp. 134–147, Jan. 2016, doi: 10.1002/asi.23366.
- [11] Carlos Santos Mangudo, "'Two-step Cluster' en SPSS y técnicas relacionadas," 2015.
- [12] "▷ Media ponderada ¿Qué es?," 2022. <https://www.sdelsol.com/glosario/media-ponderada/> (accessed Feb. 19, 2022).

- [13] J. Wang and Y. Dong, "Measurement of text similarity: A survey," *Information (Switzerland)*, vol. 11, no. 9. MDPI AG, pp. 1–17, Sep. 01, 2020. doi: 10.3390/info11090421.
- [14] "Métricas de similitud para cadenas de texto. Parte III: Métricas de conjuntos y emparejamiento de caracteres. | by Diego Campos Sobrino | Medium." <https://medium.com/@diego.campos.sobrino/m%C3%A9tricas-de-similitud-para-cadenas-de-texto-83bb09cd8591> (accessed Nov. 28, 2021).
- [15] "Guidelines for performing Systematic Literature Reviews in Software Engineering," 2007.
- [16] L. Sindhu and S. M. Idicula, "Plagiarism detection in malayalam language text using a composition of similarity measures," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Feb. 2017, pp. 456–460. doi: 10.1145/3055635.3056655.
- [17] M. Wijewickrema, V. Petras, and N. Dias, "Selecting a text similarity measure for a content-based recommender system: A comparison in two corpora," *Electronic Library*, vol. 37, no. 3, pp. 506–527, Aug. 2019, doi: 10.1108/EL-08-2018-0165.
- [18] R. Rodriguez-Cardos, F. P. R. Chicharro, J. A. O. Varela, and J. Serrano-Guerrero, "A method for the semantic analysis of documents in a business context in Spanish," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 803–810. doi: 10.1016/j.procs.2019.12.053.
- [19] T. Thongtan and T. Phienthrakul, "Similarity measurement for sentiment classification on textual reviews," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Mar. 2018, pp. 24–28. doi: 10.1145/3206185.3206204.
- [20] Institute of Electrical and Electronics Engineers, *A Comparative Analysis of Text Similarity Measures and Algorithms in Research Paper Recommender Systems*. 2018.
- [21] S. Al-Anazi, H. Almahmoud, and I. Al-Turaiki, "Finding Similar Documents Using Different Clustering Techniques," in *Procedia Computer Science*, Elsevier B.V., 2016, pp. 28–34. doi: 10.1016/j.procs.2016.04.005.
- [22] J. Richardson, K. Schlegel, R. Sallam, A. Kronz, and J. Sun, "Magic Quadrant for Analytics and Business Intelligence Platforms," 2021.
- [23] N. Leite, I. Pedrosa, and J. Bernardino, "Open source business intelligence platforms' assessment using osspal methodology," in *ICETE 2018 - Proceedings of the 15th International Joint Conference on e-Business and Telecommunications*, SciTePress, 2018, pp. 190–196. doi: 10.5220/0006910101900196.
- [24] K. Kolonko, "Master of Science in Software Engineering Performance comparison of the most popular relational and non-relational database management systems," 2018. [Online]. Available: [www.bth.se](http://www.bth.se)

- [25] I. Alejandra Tapia Ortega Lucero Cruz López Director and M. Eliana Mendoza Becerra, “Modelo Dimensional para la División de Gestión de Salud Integral y Desarrollo Humano de la Universidad del Cauca.”
- [26] J. R. A. T. (1),Adriana C. R. (2) Jaime García Sánchez (1), “Guía técnica para la construcción de escalas de actitud,” 2011.

**MÉTRICA DE SIMILITUD MIXTA Y  
PONDERADA PARA  
COMPARAR CARACTERÍSTICAS DE  
HERRAMIENTAS DE UN  
PROYECTO DE INTELIGENCIA DE NEGOCIOS**

**ANGELICA SOFIA PINTO REBOLLEDO  
IDUAR ALBEIRO SILVA MUÑOZ**

**TRABAJO DE GRADO PARA OPTAR AL TÍTULO DE INGENIERO  
DE SISTEMAS**

**DIRECTOR: PhD. MARTHA ELIANA MENDOZA BECERRA**

**CO-DIRECTOR: PHD. CARLOS ALBERTO COBOS LOZADA**

**Universidad del Cauca**  
**Facultad de Ingeniería Electrónica y Telecomunicaciones**  
Departamento de Sistemas  
**Línea de Investigación: Gestión de la Información,**  
**Bodega de Datos**  
Popayán, julio de 2023



# **MÉTRICA DE SIMILITUD MIXTA Y PONDERADA PARA COMPARAR CARACTERÍSTICAS DE HERRAMIENTAS DE UN PROYECTO DE INTELIGENCIA DE NEGOCIOS**



**ANGELICA SOFIA PINTO REBOLLEDO  
IDUAR ALBEIRO SILVA MUÑOZ**

Director: PhD. Martha Eliana Mendoza Becerra

Co-Director: PhD. Carlos Alberto Cobos Lozada

**Universidad del Cauca**  
**Facultad de Ingeniería Electrónica y Telecomunicaciones**  
**Departamento de Sistemas**  
**Línea de Investigación: Gestión de la Información,**  
**Bodega de Datos**  
Popayán, julio de 2023

## TABLA DE CONTENIDO

---

2	CATEGORÍAS Y SUBCATEGORÍAS CON SUS VALORES .....	7
3	FORMULARIOS CON PREGUNTAS Y RESPUESTAS .....	19
3.1	FORMULARIO PARA EL TEST DE UTILIDAD – EMPRESA CLUB INDEPENDIENTE ARGELIA.....	19
3.2	FORMULARIO PARA EL TEST DE UTILIDAD – FERRAGRO .....	25
3.3	FORMULARIO DEL Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa – DAVID FUENTES – FERRAGRO .....	31
3.4	FORMULARIO DEL Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa – DAVID FUENTES – CLUB INDEPENDIENTE ARGELIA	36
3.5	FORMULARIO DEL Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa – DAISY MENESES – FERRAGRO.....	41
3.6	FORMULARIO DEL Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa – DAISY MENESES – CLUB INDEPENDIENTE ARGELIA	46

## **TABLAS**

Tabla 1. Categorías y subcategorías con sus valores para el servicio SGBD .....	8
Tabla 2. Categorías y subcategorías con sus valores para el servicio SGBD .....	9
Tabla 3. Categorías y subcategorías con sus valores para el servicio ETL.....	11
Tabla 4. Categorías y subcategorías con sus valores para el servicio OLAP .....	13
Tabla 5. Categorías y subcategorías con sus valores para el servicio de Reportes .....	16

## FIGURAS

Figura 1. Respuestas de Club Independiente Argelia – Preguntas generales.....	19
Figura 2. Respuestas de Club Independiente Argelia – Preguntas sobre el ingreso de datos.....	20
Figura 3. Respuestas de Club Independiente Argelia – Preguntas sobre la visualización de resultados – Parte 1.....	21
Figura 4. Respuestas de Club Independiente Argelia – Preguntas sobre la visualización de resultados – Parte 2.....	22
Figura 5. Respuestas de Club Independiente Argelia – Preguntas sobre la visualización de resultados – Parte 3.....	23
Figura 6. Respuestas de Club Independiente Argelia – Preguntas abiertas.....	24
Figura 7. Respuestas de Ferragro – Preguntas generales.....	25
Figura 8. Respuestas Ferragro – Preguntas sobre ingreso de datos .....	26
Figura 9. Respuestas Ferragro - Preguntas sobre visualización de resultados – Parte 1.	27
Figura 10. Respuestas Ferragro - Preguntas sobre visualización de resultados – Parte 2	28
Figura 11. Respuestas Ferragro - Preguntas sobre visualización de resultados – Parte 3	29
Figura 12. Respuestas Ferragro - Preguntas abiertas.....	30
Figura 13. Respuestas de David Fuentes – Datos Generales.....	31
Figura 14. Respuestas de David Fuentes – Puntaje de cada categoría – Parte 1.....	32
Figura 15. Respuestas de David Fuentes – Puntaje de cada categoría – Parte 2.....	33
Figura 16. Respuestas de David Fuentes – Puntaje total.....	34
Figura 17. Respuestas de David Fuentes – Preguntas abiertas.....	35
Figura 18. Respuestas de David Fuentes – Datos Generales.....	36
Figura 19. Respuestas de David Fuentes – Puntaje por categoría – Parte 1 .....	37
Figura 20. Respuestas de David Fuentes – Puntaje por categoría – Parte 2 .....	38
Figura 21. Respuestas de David Fuentes – Puntaje total.....	39
Figura 22. Respuestas de David Fuentes – Preguntas abiertas.....	40
Figura 23. Respuestas de Daisy Meneses – Datos Generales .....	41
Figura 24. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 1 .....	42
Figura 25. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 2 .....	43
Figura 26. Respuestas de Daisy Meneses – Puntaje total .....	44
Figura 27. Respuestas de Daisy Meneses – Preguntas abiertas .....	45
Figura 28. Respuestas de Daisy Meneses – Datos Generales .....	46
Figura 29. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 1 .....	47

Figura 30. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 2 ..... 48  
Figura 31. Respuestas de Daisy Meneses – Puntaje total ..... 49  
Figura 32. Respuestas de Daisy Meneses – Preguntas abiertas ..... 50

# **1 CATEGORÍAS Y SUBCATEGORÍAS CON SUS VALORES**

---

**Tabla 1. Categorías y subcategorías con sus valores para el servicio SGBD**

Características	Oracle Express Edition - XE	Oracle Standard Edition 2 - SE2	Oracle Enterprise Edition - EE	MySQL Community Edition	MySQL Standard Edition	MySQL Enterprise Edition
<b>Costo</b>	0	Named User Plus 350 Processor License 17500	Named User Plus 950 Processor License 47500	0	1-4 socket server - suscripción mensual - 70 5+ socket server - suscripción anual - 4280	1-4 socket server - suscripción mensual - 150 5+ socket server - suscripción anual - 10700
<b>Documentación</b>	Oficial	Foros	Foros	Oficial	Oficial	Oficial
<b>Características técnicas</b>						
Espacio en disco - mínimo - instalación	11GB	7,2GB	7,2GB	200MB	500MB	500MB
Uso de memoria - mínimo - instalación	1GB	2GB	2GB	512MB	512 MB	2GB
Multiplataforma	Windows, Linux	Windows, Linux	Windows, Linux	Linux, Mac and Windows.	Linux, Mac and Windows.	Linux, Mac and Windows.
<b>Funcionalidad</b>						
APIs y otros métodos de acceso	JDBC	JDBC	JDBC	ADO.NET Driver ODBC Driver JDBC Driver Node.js Driver Python Driver C++ Driver	ADO.NET Driver ODBC Driver JDBC Driver Node.js Driver Python Driver C++ Driver C Driver C API	ADO.NET Driver ODBC Driver JDBC Driver Node.js Driver Python Driver C++ Driver C Driver C API

				C Driver C API		
Funciones nativas de cifrado de datos	Cifrado de datos transparente (TDE)	Cifrado de datos transparente (TDE)	Cifrado de datos transparente (TDE)	Algoritmo de AES	Algoritmo de AES	Algoritmo de AES

**Tabla 2. Categorías y subcategorías con sus valores para el servicio SGBD**

Características	SQL Server Express	SQL Server Estándar	SQL Server Enterprise	PostgreSQL Community
<b>Costo</b>	0	118 USD mes	452 USD mes	0
<b>Documentación</b>	Oficial	Foros	Foros	Oficial
<b>Características técnicas</b>				
Espacio en disco - mínimo - instalación	6 GB	6 GB	6 GB	1 GB
Uso de memoria - mínimo - instalación	512 MB	1 GB	1 GB	512 MB
Multiplataforma	Windows, Linux	Windows, Linux	Windows, Linux	Windows, Linux y Mac
<b>Funcionalidad</b>				
APIs y otros métodos de acceso	ADO.NET Driver ODBC Driver JDBC Driver	ADO.NET Driver ODBC Driver JDBC Driver	ADO.NET Driver ODBC Driver JDBC Driver	pgJDBC



Funciones nativas de cifrado de datos	Cifrado de datos transparente (TDE)	Cifrado de datos transparente (TDE)	Cifrado de datos transparente (TDE)	Funciones hash: MD5 (Message Digest Algorithm 5), SHA-256 (Secure Hash Algorithm 256-bit). Cifrado simétrico: AES (Advanced Encryption Standard) y DES (Data Encryption Standard)
---------------------------------------	-------------------------------------	-------------------------------------	-------------------------------------	---

**Tabla 3. Categorías y subcategorías con sus valores para el servicio ETL**

<b>Características</b>	<b>Microsoft (Power BI) - Power query</b>	<b>Salesforce - Tableau Prep</b>	<b>Pentaho (Pentaho data integration) Community</b>	<b>Pentaho (Pentaho data integration) Enterprise</b>	<b>Jaspersoft (Jaspersoft ETL) community</b>	<b>Jaspersoft (Jaspersoft ETL)</b>
<b>Costo</b>	0	Tableau Creator at \$70.00 per user/month.	0	160 USD usuario / MES	0	Empieza en 150 USD por licencia/mes
<b>Documentación</b>	Oficial, Foros	Oficial, Foros	Oficial, Foros	Oficial	Oficial, Foros	Oficial
<b>Características técnicas</b>						
Uso de memoria - minimo - instalación	4 GB	4 GB	128Mb o superior	128Mb o superior	4GB	4GB
Espacio en disco - minimo - instalación	2 GB	2 GB	200Mb	200Mb	10 GB	10 GB
Multiplataforma	Windows, Mac	Windows, linux, mac	Windows, Linux, Mac corre sobre la plataforma java	Windows, Linux, Mac corre sobre la plataforma java	mac, windows, linux	mac, windows, linux
<b>Funcionalidades</b>						
Extracción (lectura de archivos)	Excel JSON PDF CSV XML	Microsoft Excel Archivo de texto Microsoft Access Archivo JSON Archivo PDF	CSV EXCEL JSON	CSV EXCEL JSON	CSV EXCEL JSON	NO
Carga (carga en archivos)	Excel	.csv	CSV EXCEL JSON	CSV EXCEL JSON	CSV EXCEL JSON	NO

Conexión a SGBD representativas	Oracle SQL Server MySQL PostgreSQL	SAP HANA MYSQL AMAZON REDSHIFT	ORACLE POSTGRESQL MYSQL SAP HANA AMAZON REDSHIFT	ORACLE POSTGRESQL MYSQL SAP HANA AMAZON REDSHIFT	SQL SERVER ORACLE MYSQL POSTGRESQL	SQL SERVER ORACLE MYSQL POSTGRESQL
Búsqueda	SI	SI	SI	SI	NO	NO
Concatenar valores	SI	SI	SI	SI	SI	SI
Dividir campos	SI	SI	SI	SI	NO	SI
Filtro de filas	SI	NO	SI	SI	SI	SI
Funciones de agregación	SI	SI	SI	SI	NO	SI
Merge (fusión de datos)	SI	SI	SI	SI	NO	NO
Normalización de filas	SI	SI	SI	SI	NO	NO
Ordenamiento	SI	si	SI	SI	NO	NO
Remoción duplicada	SI	SI	NO	NO	NO	NO
Des-normalización de filas	SI	SI	SI	SI	NO	NO
Depurador con puntos de interrupción.	SI	no	NO	NO	NO	NO
Recorre hojas excel	SI	NO	SI	SI	NO	NO
Actualización programada de datos	SI	SI	SI	SI	NO	NO

**Tabla 4. Categorías y subcategorías con sus valores para el servicio OLAP**

Características	POWER BI - Pro	POWER BI - Premium	Pentaho Analysis Services Mondrian COMMUNITY	Pentaho Analysis Services Mondrian ENTERPRISE	Jaspersoft OLAP COMMUNITY	Jaspersoft OLAP ENTERPRISE
<b>Costo</b>	8,40 Euros por usuario/mes	16,90 E por usuario/mes 4212,30 E por capacidad/mes	0	160 USD usuario / MES	0	Empieza en 150 USD por licencia
<b>Documentación</b>	Oficial, Foros	Oficial	Oficial, Foros	Oficial	Oficial, Foros	Oficial
<b>Características técnicas</b>						
Multiplataforma	Mac, Windows	Mac, Windows	Windows, Linux, Mac corre sobre la plataforma java	Windows, Linux, Mac corre sobre la plataforma java	Mac, Windows, Linux	Mac, Windows, Linux
Espacio en disco - mínimo - instalación	Power BI Desktop - 1GB	1GB	100 GB	100 GB	10GB	40GB
Uso de memoria - mínimo - instalación	Mínimo: 1 GB Recomendado: 4 GB como mínimo	Mínimo: 1 GB Recomendado: 4 GB como mínimo	4 GB RAM	4 GB RAM	4 GB	8 GB
Medidas calculadas	SI	SI	SI	SI	NO	SI
<b>Funcionalidades</b>						
Jerarquía Padre/Hijo	SI	SI	NO	SI	NO	NO
Examinar funcionalidad de los esquemas (drag and drop)	NO	NO	NO	SI	SI	SI

Traducciones	SI	SI	NO	NO	SI	SI
Creación de varios cubos en un mismo esquema	NO	SI	SI	SI	NO	SI
Múltiples orígenes de datos en un mismo esquema	SI	SI	NO	SI	NO	SI
Permisos por medio de roles	NO	SI	NO	SI	SI	SI
Conexión a motores de bases de datos	SQL Server SQL Server Analysis Services Oracle MySQL PostgreSQL SAP HANA MariaDB	SQL Server SQL Server Analysis Services Oracle MySQL PostgreSQL SAP HANA MariaDB"	Apache Derby Firebird hsqldb IBM DB2 Microsoft Access Microsoft SQL Server MySQL Oracle PostgreSQL Sybase	Apache Derby Firebird hsqldb IBM DB2 Microsoft Access Microsoft SQL Server MySQL Oracle PostgreSQL Sybase	Mysql, sql server, JTDS, oracle, postgresSQL, pervasive Se conecta a BD no SQL	Mysql, sql server, JTDS, oracle, postgresSQL, pervasive Se conecta a BD no SQL
Conexión a archivos	Excel CSV XML JSON PDF	Excel CSV XML JSON PDF	CSV, JSON	CSV, JSON	json, xml, csv	json, xml, csv, .xls, texto

Dimensión	Normal Time Organización Geography Lista de materiales Cuentas Clientes Productos Escenario Cuantitativo Utilidad Moneda Tarifas Canal Promoción	Normal Time Organización Geography Lista de materiales Cuentas Clientes Productos Escenario Cuantitativo Utilidad Moneda Tarifas Canal Promoción	Dimensión tiempo Normal Time Organización Geography Lista de materiales Cuentas Clientes Productos	Dimensión tiempo Normal Time Organización Geography Lista de materiales Cuentas Clientes Productos	Normal Time Organización Geography Lista de materiales Cuentas Clientes Productos Dimensión tiempo	Normal Time Organización Geography Lista de materiales Cuentas Clientes Productos Dimensión tiempo
-----------	---	---	---	---	--	---

**Tabla 5. Categorías y subcategorías con sus valores para el servicio de Reportes**

<b>Características</b>	<b>POWER BI Desktop</b>	<b>POWER BI Pro</b>	<b>POWER BI Premium</b>	<b>Salesforce (Tableau Report Creator)</b>	<b>Pentaho COMMUNITY</b>	<b>Pentaho ENTERPRISE</b>	<b>Jaspersoft (Jaspersoft report) COMMUNITY</b>	<b>Jaspersoft (Jaspersoft report) ENTERPRISE</b>
<b>Costo</b>	Libre	8,40 Euros por usuario/mes	16,90 E por usuario/mes 4212,30 E por capacidad/mes	120 USD por mes	0	160 USD usuario / MES	0	Empieza en 150 USD por licencia
<b>Documentación</b>	Oficial, Foros	Oficial	Oficial	Oficial	Oficial, Foros	Oficial	Oficial,Foros	Oficial
<b>Características técnicas</b>								
Multiplataforma	Windows, linux, mac	Windows, linux, mac	Windows, linux, mac	Windows, linux, mac	Windows, linux, mac	Windows, linux, mac	Windows, linux, mac	Windows, linux, mac
Espacio en disco - mínimo - instalación	1GB	1GB	1GB	4 GB	128Mb o superior	128Mb o superior	10GB	40GB
Uso de memoria - mínimo - instalación	2GB	2GB	2GB	2GB	200Mb	200Mb	4 GB	8 GB
<b>Funcionalidades</b>								
Plantillas predeterminadas	SI	SI	SI	SI	NO	SI	NO	SI
Permite gráficos 3D	SI	SI	SI	SI	NO	SI	NO	SI

Permite minería de datos?	NO	SI	SI	SI	NO	SI	NO	SI
notificaciones	NO	SI	SI	SI	NO	SI	NO	SI
Soporte dedicado	NO	SI	SI	SI	NO	SI	NO	SI
Mobile	NO	SI	SI	SI	NO	SI	NO	SI
Visualizaciones interactivas	NO	SI	SI	SI	NO	SI	NO	SI
Conexión a motor OLAP	SQL SERVER ANALYSIS SERVICES DATABASE	SQL SERVER ANALYSIS SERVICES DATABASE	SQL SERVER ANALYSIS SERVICES DATABASE	Teradata OLAP data	Pentaho Analysis (Mondrian)	Pentaho Analysis (Mondrian)	Jaspersoft OLAP	Jaspersoft OLAP



Conexión motores de bases de datos relacionales.	SQL Server Oracle database MySQL Database SAP HANA POstgreSQL database	SQL Server Oracle database MySQL Database SAP HANA POstgreSQL database	SQL Server Oracle database MySQL Database SAP HANA POstgreSQL database	Oracle database MySQL Database SAP HANA POstgreSQL database	Oracle PostgreSQL SQLite Redshift SparkSQL SAP HANA SAP ERP MySQL SQL Server Aplicaciones empresariales como Google Analytics, Google Docs y Java Message Services	Oracle PostgreSQL SQLite Redshift SparkSQL SAP HANA SAP ERP MySQL SQL Server Aplicaciones empresariales como Google Analytics, Google Docs y Java Message Services	DB2, MySQL, Oracle, PostgreSQL	DB2, MySQL, Oracle, PostgreSQL
Conexión a varios orígenes de datos en un mismo reporte	SI	SI	SI	SI	NO	SI	NO	SI

## 2 FORMULARIOS CON PREGUNTAS Y RESPUESTAS

### 2.1 FORMULARIO PARA EL TEST DE UTILIDAD – EMPRESA CLUB INDEPENDIENTE ARGELIA

No se pueden editar las respuestas

### Test de utilidad de la aplicación web

El objetivo del presente test es medir la utilidad de la aplicación web con respecto a las herramientas sugeridas de inteligencia de negocios en los servicios de Sistema Gestor de Base de Datos (SGBD), Extracción - Transformación - Carga (ETL), Procesamiento Analítico en Línea (OLAP) y Reportes.

El correo electrónico del destinatario (jdpinto109@misena.edu.co) se registró al enviar el formulario.

\* Indica que la pregunta es obligatoria

Nombre de la empresa \*  
Club Independiente Argelia

Nombre usuario \*  
Daniel Pinto

Ocupación / Rol \*  
Presidente

#### Preguntas generales

¿Cree que la aplicación es fácil de utilizar para obtener los resultados de las herramientas sugeridas? \*

Totalmente de acuerdo  
 Parcialmente de acuerdo  
 Indiferente  
 Algo en desacuerdo  
 Totalmente en desacuerdo

¿Es fácil entender la puntuación total de las herramientas en cada servicio? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 1. Respuestas de Club Independiente Argelia – Preguntas generales

Preguntas sobre los formularios de ingreso de datos

¿Las preguntas en los formularios de cada servicio son entendibles? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Las ayudas presentes en los formularios de cada servicio son útiles en el momento de ingresar los datos? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 2. Respuestas de Club Independiente Argelia – Preguntas sobre el ingreso de datos**

**Preguntas sobre la visualización de los resultados**

¿Es fácil entender la información que se despliega en cada servicio con respecto a las **categorias** (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Es fácil entender la información que se despliega en cada servicio con respecto a las **sub categorias** de costo, documentación, características técnicas y funcionalidades? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 3. Respuestas de Club Independiente Argelia – Preguntas sobre la visualización de resultados – Parte 1**

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio SGBD** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio ETL** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 4. Respuestas de Club Independiente Argelia – Preguntas sobre la visualización de resultados – Parte 2**

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio OLAP** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio Reportes** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 5. Respuestas de Club Independiente Argelia – Preguntas sobre la visualización de resultados – Parte 3**

**Preguntas abiertas**

¿Considera que faltó algún elemento a tener en cuenta con respecto al costo?, de ser así, ¿cuál? \*

no

¿Considera que faltó algún elemento a tener en cuenta con respecto a la documentación?, de ser así, ¿cuál? \*

no

¿Considera que faltó algún elemento a tener en cuenta con respecto a las características técnicas?, de ser así, ¿cuál? \*

no

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio SGBD?, de ser así, ¿cuál? \*

no

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio ETL?, de ser así, ¿cuál? \*

no

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio OLAP?, de ser así, ¿cuál? \*

no

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio de Reportes?, de ser así, ¿cuál? \*

no

Enviado el 12/6/23, 17:56

**Figura 6. Respuestas de Club Independiente Argelia – Preguntas abiertas**

## 2.2 FORMULARIO PARA EL TEST DE UTILIDAD – FERRAGRO

No se pueden editar las respuestas

### Test de utilidad de la aplicación web

El objetivo del presente test es medir la utilidad de la aplicación web con respecto a las herramientas sugeridas de inteligencia de negocios en los servicios de Sistema Gestor de Base de Datos (SGBD), Extracción - Transformación - Carga (ETL), Procesamiento Analítico en Línea (OLAP) y Reportes.

El correo electrónico del destinatario ([inge.yesidmunozome@gmail.com](mailto:inge.yesidmunozome@gmail.com)) se registró al enviar el formulario.

\* Indica que la pregunta es obligatoria

Nombre de la empresa \*  
FERRAGRO

Nombre usuario \*  
YESID MUÑOZ OME

Ocupación / Rol \*  
Ingeniero Civil

#### Preguntas generales

¿Cree que la aplicación es fácil de utilizar para obtener los resultados de las herramientas sugeridas? \*

Totalmente de acuerdo  
 Parcialmente de acuerdo  
 Indiferente  
 Algo en desacuerdo  
 Totalmente en desacuerdo

¿Es fácil entender la puntuación total de las herramientas en cada servicio? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 7. Respuestas de Ferragro – Preguntas generales



**Preguntas sobre los formularios de ingreso de datos**

¿Las preguntas en los formularios de cada servicio son entendibles? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Las ayudas presentes en los formularios de cada servicio son útiles en el momento de ingresar los datos? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 8. Respuestas Ferragro – Preguntas sobre ingreso de datos**

**Preguntas sobre la visualización de los resultados**

¿Es fácil entender la información que se despliega en cada servicio con respecto a las **categorías** (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Es fácil entender la información que se despliega en cada servicio con respecto a las **sub categorías** de costo, documentación, características técnicas y funcionalidades? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 9. Respuestas Ferragro - Preguntas sobre visualización de resultados – Parte

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio SGBD** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio ETL** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 10. Respuestas Ferragro - Preguntas sobre visualización de resultados – Parte 2**

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio OLAP** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Las herramientas de BI resaltadas en color verde en el listado del **servicio Reportes** son las que se adecuan a los requerimientos que fueron ingresados en las categorías (costo, documentación, características técnicas y funcionalidades)? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 11. Respuestas Ferragro - Preguntas sobre visualización de resultados – Parte 3**

**Preguntas abiertas**

¿Considera que faltó algún elemento a tener en cuenta con respecto al costo?, de ser así, ¿cuál? \*

No \_\_\_\_\_

¿Considera que faltó algún elemento a tener en cuenta con respecto a la documentación?, de ser así, ¿cuál? \*

No \_\_\_\_\_

¿Considera que faltó algún elemento a tener en cuenta con respecto a las características técnicas?, de ser así, ¿cuál? \*

No \_\_\_\_\_

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio SGBD?, de ser así, ¿cuál? \*

No \_\_\_\_\_

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio ETL?, de ser así, ¿cuál? \*

No \_\_\_\_\_

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio OLAP?, de ser así, ¿cuál? \*

No \_\_\_\_\_

¿Considera que faltó una funcionalidad a tener en cuenta en el servicio de Reportes?, de ser así, ¿cuál? \*

No \_\_\_\_\_

**Figura 12. Respuestas Ferragro - Preguntas abiertas**

## 2.3 FORMULARIO DEL TEST DE ADECUACIÓN DE HERRAMIENTAS DE BI SUGERIDAS A LAS NECESIDADES DE UNA EMPRESA – DAVID FUENTES – FERRAGRO

No se pueden editar las respuestas

### Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa

El objetivo de este test es medir el grado de adecuación de herramientas de inteligencia de negocios (BI) sugeridas por una aplicación web, a las necesidades técnicas y económicas de una empresa. Teniendo en cuenta, que estas herramientas corresponden a los servicios de gestión de base de datos relacionales (SGBD), procesamiento analítico en línea (OLAP), proceso de extracción, transformación y carga de datos (ETL) y la generación de reportes; necesarios para el desarrollo de un proyecto de BI.

El correo electrónico del destinatario ([davidafuentesv@gmail.com](mailto:davidafuentesv@gmail.com)) se registró al enviar el formulario.

\* Indica que la pregunta es obligatoria

#### Datos del evaluador

Nombre completo \*

David Antonio Fuentes Vargas

Profesión \*

Ingeniero electrónico

Nombre de la empresa en la que trabaja \*

Fundación Cardiovascular de Colombia

Cargo desempeñado en la empresa \*

Ingeniero de datos

#### Datos de la empresa a evaluar

Nombre de la empresa \*

FERRAGRO

Figura 13. Respuestas de David Fuentes – Datos Generales

**Preguntas relacionadas sobre el puntaje de cada categoría con respecto a cada servicio.**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de SGBD** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de ETL** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 14. Respuestas de David Fuentes – Puntaje de cada categoría – Parte 1**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de OLAP** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de Reportes** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 15. Respuestas de David Fuentes – Puntaje de cada categoría – Parte 2



**Pregunta relacionada sobre el puntaje total de las herramientas en cada servicio.**

¿Considera que la puntuación total de cada herramienta en los servicios se adecuan a los requerimientos de la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 16. Respuestas de David Fuentes – Puntaje total**

**Preguntas abiertas**

¿Considera que se debe agregar alguna categoría para la evaluación de los resultados de cada herramienta? \*

Popularidad en el mercado, una métrica relacionada con el uso de la herramienta en el mundo, algo similar a lo que hace Gartner

¿Considera que se debe agregar alguna subcategoría para la evaluación de los resultados de cada herramienta? \*

popularidad nacional e internacional

¿Tiene alguna sugerencia adicional sobre la visualización del puntaje total de las herramientas en cada servicio? \*

ubicarlas en orden o en su defecto utilizar barra de colores, siendo verde claro el mas sugerido y rojo el menos

¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las categorías de las herramientas? \*

ninguna

¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las subcategorías de las herramientas? \*

manejar colores en las calificaciones de verde a rojo para fijar mejor la atención en lo verde

¿Tiene alguna sugerencia adicional? \*

Ninguna

Enviado el 30/6/23, 12:24

**Figura 17. Respuestas de David Fuentes – Preguntas abiertas**

## 2.4 FORMULARIO DEL TEST DE ADECUACIÓN DE HERRAMIENTAS DE BI SUGERIDAS A LAS NECESIDADES DE UNA EMPRESA – DAVID FUENTES – CLUB INDEPENDIENTE ARGELIA

No se pueden editar las respuestas

### Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa

El objetivo de este test es medir el grado de adecuación de herramientas de inteligencia de negocios (BI) sugeridas por una aplicación web, a las necesidades técnicas y económicas de una empresa. Teniendo en cuenta, que estas herramientas corresponden a los servicios de gestión de base de datos relacionales (SGBD), procesamiento analítico en línea (OLAP), proceso de extracción, transformación y carga de datos (ETL) y la generación de reportes; necesarios para el desarrollo de un proyecto de BI.

El correo electrónico del destinatario ([davidafuentesv@gmail.com](mailto:davidafuentesv@gmail.com)) se registró al enviar el formulario.

\* Indica que la pregunta es obligatoria

#### Datos del evaluador

Nombre completo \*

David Antonio Fuentes Vargas

Profesión \*

Ingeniero electrónico

Nombre de la empresa en la que trabaja \*

Fundación Cardiovascular de Colombia

Cargo desempeñado en la empresa \*

Ingeniero de Datos

#### Datos de la empresa a evaluar

Nombre de la empresa \*

CLUB INDEPENDIENTE ARGELIA

Figura 18. Respuestas de David Fuentes – Datos Generales

**Preguntas relacionadas sobre el puntaje de cada categoría con respecto a cada servicio.**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de SGBD** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de ETL** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 19. Respuestas de David Fuentes – Puntaje por categoría – Parte 1**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de OLAP** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de Reportes** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 20. Respuestas de David Fuentes – Puntaje por categoría – Parte 2

**Pregunta relacionada sobre el puntaje total de las herramientas en cada servicio.**

¿Considera que la puntuación total de cada herramienta en los servicios se adecuan a los requerimientos de la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 21. Respuestas de David Fuentes – Puntaje total**

**Preguntas abiertas**

¿Considera que se debe agregar alguna categoría para la evaluación de los resultados de cada herramienta? \*

Popularidad de la herramienta a nivel internacional, midiendo con alguna escala estándar

¿Considera que se debe agregar alguna subcategoría para la evaluación de los resultados de cada herramienta? \*

Popularidad nacional consultando con empresas del país y popularidad internacional con escalas estándar

¿Tiene alguna sugerencia adicional sobre la visualización del puntaje total de las herramientas en cada servicio? \*

Ninguna

¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las categorías de las herramientas? \*

Ninguna

¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las subcategorías de las herramientas? \*

me gustaria que la bonificación de 0 no apareciera, en lugar de aparecer bonificación +0 puntos

¿Tiene alguna sugerencia adicional? \*

Excelente trabajo, como mencioné en la evaluación de la herramienta anterior me gustaria que aparecieran las herramientas en orden de la mejor a la peor. Especificar la forma de actualización ya que a medida que pasan los años aparecen nuevas herramientas. Respecto a la presentación sugeriría decir porque se elige la formula de documentación, tambien hay una diapositiva donde se hacen afirmaciones del tipo "número porque es un número, texto porque es un texto", siento que suena redundante. Todos son aspectos por mejorar pequeños y sugerencias pero en general es un muy buen trabajo y muy útil, la empresa en la que trabajo seguramente le hubiera interesado usar esto cuando eligió sus herramientas.

Enviado el 30/6/23, 12:23

Figura 22. Respuestas de David Fuentes – Preguntas abiertas

## 2.5 FORMULARIO DEL TEST DE ADECUACIÓN DE HERRAMIENTAS DE BI SUGERIDAS A LAS NECESIDADES DE UNA EMPRESA – DAISY MENESES – FERRAGRO

No se pueden editar las respuestas

### Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa

El objetivo de este test es medir el grado de adecuación de herramientas de inteligencia de negocios (BI) sugeridas por una aplicación web, a las necesidades técnicas y económicas de una empresa. Teniendo en cuenta, que estas herramientas corresponden a los servicios de gestión de base de datos relacionales (SGBD), procesamiento analítico en línea (OLAP), proceso de extracción, transformación y carga de datos (ETL) y la generación de reportes; necesarios para el desarrollo de un proyecto de BI.

El correo electrónico del destinatario ([daisyysel@gmail.com](mailto:daisyysel@gmail.com)) se registró al enviar el formulario.

\* Indica que la pregunta es obligatoria

#### Datos del evaluador

Nombre completo \*

Daisy Yisel Meneses Lopez

Profesión \*

Msc (C) Maestría en computación

Nombre de la empresa en la que trabaja \*

Cooemeva

Cargo desempeñado en la empresa \*

Ingeniera de Datos

#### Datos de la empresa a evaluar

Nombre de la empresa \*

Cooemeva

Figura 23. Respuestas de Daisy Meneses – Datos Generales



**Preguntas relacionadas sobre el puntaje de cada categoría con respecto a cada servicio.**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de SGBD** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de ETL** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 24. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 1**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de OLAP** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de Reportes** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 25. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 2**

**Pregunta relacionada sobre el puntaje total de las herramientas en cada servicio.**

¿Considera que la puntuación total de cada herramienta en los servicios se adecuan a los requerimientos de la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 26. Respuestas de Daisy Meneses – Puntaje total**

Preguntas abiertas	
¿Considera que se debe agregar alguna categoría para la evaluación de los resultados de cada herramienta? *	No
¿Considera que se debe agregar alguna subcategoría para la evaluación de los resultados de cada herramienta? *	si, podría ser la cantidad de usuarios que van a utilizar el servicio (costo beneficio), dado que para algunas aplicaciones ejemplo BI se requiere licencia por cada usuario .
¿Tiene alguna sugerencia adicional sobre la visualización del puntaje total de las herramientas en cada servicio? *	no
¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las categorías de las herramientas? *	poner en otro color el pontaje de bonificación
¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las subcategorías de las herramientas? *	no
¿Tiene alguna sugerencia adicional? *	Organizar los valores ejemplo si son pesos en los valores ingresados por los usuarios asi mismo convertir el valor a peso en el costo de la herramienta, igual para espacio en disco, que los datos queden ya sean en GB o en MB

**Figura 27. Respuestas de Daisy Meneses – Preguntas abiertas**

## 2.6 FORMULARIO DEL TEST DE ADECUACIÓN DE HERRAMIENTAS DE BI SUGERIDAS A LAS NECESIDADES DE UNA EMPRESA – DAISY MENESES – CLUB INDEPENDIENTE ARGELIA

No se pueden editar las respuestas

### Test de adecuación de herramientas de BI sugeridas a las necesidades de una empresa

El objetivo de este test es medir el grado de adecuación de herramientas de inteligencia de negocios (BI) sugeridas por una aplicación web, a las necesidades técnicas y económicas de una empresa. Teniendo en cuenta, que estas herramientas corresponden a los servicios de gestión de base de datos relacionales (SGBD), procesamiento analítico en línea (OLAP), proceso de extracción, transformación y carga de datos (ETL) y la generación de reportes; necesarios para el desarrollo de un proyecto de BI.

El correo electrónico del destinatario (daisyysel@gmail.com) se registró al enviar el formulario.

\* Indica que la pregunta es obligatoria

#### Datos del evaluador

Nombre completo \*

Daisy Yisel Meneses Lopez

Profesión \*

Msc (c) Maestría en computación

Nombre de la empresa en la que trabaja \*

Coomeva

Cargo desempeñado en la empresa \*

Ingeniera de datos

#### Datos de la empresa a evaluar

Nombre de la empresa \*

CLUB INDEPENDIENTE ARGELIA

Figura 28. Respuestas de Daisy Meneses – Datos Generales

**Preguntas relacionadas sobre el puntaje de cada categoría con respecto a cada servicio.**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de SGBD** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de ETL** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 29. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 1**

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de OLAP** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

¿Considera que las herramientas con mayor puntaje con respecto al **servicio de Reportes** se adecuan al costo, documentación, características técnicas y funcionalidades definidas por la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
Costo	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Documentación	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Características técnicas	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Funcionalidades	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figura 30. Respuestas de Daisy Meneses – Puntaje de cada categoría – Parte 2

**Pregunta relacionada sobre el puntaje total de las herramientas en cada servicio.**

¿Considera que la puntuación total de cada herramienta en los servicios se adecuan a los requerimientos de la empresa? \*

	Totalmente de acuerdo	Parcialmente de acuerdo	Indiferente	Algo en desacuerdo	Totalmente en desacuerdo
SGBD	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ETL	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
OLAP	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
REPORTES	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Figura 31. Respuestas de Daisy Meneses – Puntaje total**



Preguntas abiertas	
¿Considera que se debe agregar alguna categoría para la evaluación de los resultados de cada herramienta? *	no
¿Considera que se debe agregar alguna subcategoría para la evaluación de los resultados de cada herramienta? *	cantidad de licencias por usuario
¿Tiene alguna sugerencia adicional sobre la visualización del puntaje total de las herramientas en cada servicio? *	no
¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las categorías de las herramientas? *	no
¿Tiene alguna sugerencia con respecto a la forma en que se visualizan los puntajes de las subcategorías de las herramientas? *	color de las bonificaciones que se diferencie del resultado
¿Tiene alguna sugerencia adicional? *	no

Figura 32. Respuestas de Daisy Meneses – Preguntas abiertas