



UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE CUENCA
CARRERA DE INGENIERÍA DE SISTEMAS

**ESTUDIO COMPARATIVO DE HERRAMIENTAS DE INTELIGENCIA DE
NEGOCIOS Y SUS PRESTACIONES**

Trabajo de titulación previo a la obtención del
título de Ingeniera de Sistemas

AUTORA: LILIANA MAGALY CAJAS CHUQUI

TUTOR: ING. GERMÁN ERNESTO PARRA GONZÁLEZ

Cuenca - Ecuador

2022

**CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE
TITULACIÓN**

Yo, Liliana Magaly Cajas Chuqui con documento de identificación N° 0302647615, manifiesto que:

Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Cuenca, 31 de agosto del 2022

Atentamente,



Liliana Magaly Cajas Chuqui

0302647615

**CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE
TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA**

Yo, Liliana Magaly Cajas Chuqui con documento de identificación N° 0302647615, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud que soy autora del Proyecto técnico: “Estudio comparativo de herramientas de inteligencia de negocios y sus prestaciones”, el cual ha sido desarrollado para optar por el título de: Ingeniera de Sistemas, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Cuenca, 31 de agosto del 2022

Atentamente,



Liliana Magaly Cajas Chuqui

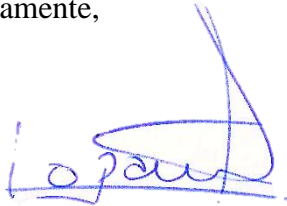
0302647615

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Germán Ernesto Parra González con documento de identificación N° 0102287190, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: ESTUDIO COMPARATIVO DE HERRAMIENTAS DE INTELIGENCIA DE NEGOCIOS Y SUS PRESTACIONES, realizado por Liliana Magaly Cajas Chuqui con documento de identificación N° 0302647615, obteniendo como resultado final el trabajo de titulación bajo la opción Proyecto técnico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Cuenca, 31 de agosto del 2022

Atentamente,



Ing. Germán Ernesto Parra González

0102287190

DEDICATORIA

Dedico este proyecto de titulación a:

Principalmente a mi familia, mis padres quienes se preocuparon de mi bienestar y me brindaron su apoyo moral, a mis hermanos mayores que cada uno aportó de diferentes maneras en mi educación e hicieron de mí una persona de bien, a mis hermanos menores que son mi motivación día a día para mejorar y que sin duda alguna, cuento con ellos hoy y hasta que Dios lo permita.

También a mi tutor de tesis, Ing. German Parra quien estuvo pendiente y fue una persona paciente y ayudó a que este proyecto se realice y culmine satisfactoriamente.

Así también dedico este trabajo a la única persona que fue un gran amigo, gracias por tu paciencia y tu apoyo ya que con tu ayuda se hizo posible uno de los anhelos más deseados en mi vida.

Liliana Magaly Cajas Chuqui

AGRADECIMIENTO

Quiero agradecer a Dios por darme a personas maravillosas que estuvieron a mi lado apoyándome y por la fortaleza para enfrentarme a los retos que se presentaron en todo el transcurso y finalización de mis estudios, así también permitirme obtener no solo conocimientos sino también una experiencia satisfactoria en la Universidad Politécnica Salesiana.

Necesitaría más de un párrafo para agradecer a mi familia, a mi gran amigo y a mi tutor, sin embargo, quiero expresar que agradezco a mis padres a quienes amo que estuvieron conmigo durante esos momentos difíciles, me han inculcado valores en beneficio a mi formación tanto personal como profesional.

De la misma manera agradezco a la Universidad Politécnica Salesiana, a los docentes de la carrera de Ingeniería de Sistemas, en especial a mi Tutor Ing. German Parra, por la ayuda brindada para la realización de este proyecto siendo un guía y fuente de conocimientos, que permitió el desarrollo y finalización de este proyecto.

Por último, quiero dar gracias a mis hermanos, por el apoyo incondicional que me han brindado, en los momentos más difíciles que se me han presentado, a mis familiares, amigo y demás personas que fueron parte de este proceso.

Liliana Magaly Cajas Chuqui

RESUMEN

La investigación está enfocada para aquellas empresas que adopten una solución eficaz al momento de tomar decisiones acertadas, para ello hay que tomar en cuenta que una empresa exitosa debe llevar la información adecuada en tiempos acorde a la gestión de las operaciones de la empresa, para ello es necesario determinar qué es la inteligencia de negocios y cómo aplicarlo para hacer que un negocio sea inteligente, ya que al implementarla es más visible darse cuenta de que la información es concisa y oportuna en base a los datos analizados.

El presente proyecto tiene como objetivo realizar un estudio comparativo, de las herramientas de la inteligencia de negocios que ofrecen diferentes soluciones, por lo que no solo se enfocará a la tecnología para el análisis de datos, sino también a cómo se presenta la información procesada mediante el uso y opciones que nos ofrece cada de una las herramientas para determinar de manera eficiente cómo aporta cada herramienta a una empresa.

Para el desarrollo del proyecto se optó por la metodología M3S de cual se determinó los criterios de evaluación para la elección de las herramientas de BI según un análisis comparativo en base a la versión gratuita, versión de prueba, costo de adquisición, análisis, aceptación de archivos, tamaño de archivo, tipo de fuentes de datos, facilidad de uso y compatibilidad con sistema operativo Windows.

En base a los criterios anteriormente definidos se logró determinar que tres herramientas como Tableau Desktop, Microsoft Power BI Desktop y IBM Cognos como resultado de sus comparativas demostraron cumplir con las expectativas en cuanto a un análisis de datos a profundidad y sus interfaces intuitivas, sin embargo, entre la comparativa la herramienta con mayor relevancia fue Microsoft Power BI por su manera de almacenamiento de procesos.

Palabras clave: Inteligencia de negocios, análisis de datos, herramientas, implementación, comparativa, criterio de evaluación.

ABSTRACT

The research is focused for those companies that adopt an effective solution at the time of making the right decisions, for this it is necessary to take into account that a successful company must carry the right information in times according to the management of the company's operations, for this it is necessary to determine what business intelligence is and how to apply it to make a business intelligent, since by implementing it is more visible to realize that the information is concise and timely based on the analyzed data.

The objective of this project is to carry out a comparative study of the business intelligence tools offered by different solutions, so it will not only focus on the technology for data analysis, but also on how the processed information is presented through the use and options offered by each of the tools to efficiently determine how each tool contributes to a company.

For the development of the project, the M3S methodology was used to determine the evaluation criteria for the choice of BI tools according to a comparative analysis based on free version, trial version, acquisition cost, analysis, file acceptance, file size, type of data sources, ease of use and compatibility with Windows operating system.

Based on the criteria previously defined, it was determined that three tools such as Tableau Desktop, Microsoft Power BI Desktop and IBM Cognos, as a result of their comparative tests, proved to meet the expectations in terms of in-depth data analysis and intuitive interfaces; however, among the comparative tests, the tool with the highest relevance was Microsoft Power BI due to its way of storing processes.

Keywords: business intelligence, data analysis, tools, implementation, benchmarking, evaluation criteria.

ÍNDICE

1. INTRODUCCIÓN	14
2. PROBLEMÁTICA DE ESTUDIO	15
3. OBJETIVOS DE LA INVESTIGACIÓN	16
3.1 Objetivo General	16
3.2 Objetivo Específico	16
4. REVISIÓN DE LA LITERATURA O FUNDAMENTOS TEÓRICOS	17
4.1 INTELIGENCIA DE NEGOCIOS	17
4.1.1 ESTRUCTURA DE INTELIGENCIA DE NEGOCIOS	21
4.1.1.1 Fuentes de información	21
4.1.1.2 Proceso ETL	22
4.1.1.3 Almacenamiento y agregación	24
4.1.1.3.1 Data Mart	24
4.1.1.3.2 Bodegas de datos (DataWarehouse)	25
4.1.1.3.3 Lago de datos	27
4.1.1.4 Agregación	27
4.1.1.5 Visualización	28
4.1.2 IMPORTANCIA DE LA INTELIGENCIA DE NEGOCIOS	29
4.1.3 POR QUÉ IMPLEMENTAR	31
4.1.4 INDICADORES QUE CONSIDERAR PARA CONOCER CUANDO SE DEBE IMPLEMENTAR	32
4.1.5 ANÁLISIS DE DATOS	33
4.2 METODOLOGIA M3S	35
4.2.1 CONCEPTO M3S	35
4.2.2 BÚSQUEDA Y MINERÍA DE DATOS	35
4.2.3 ELECCIÓN DE LA MEJOR HERRAMIENTA DE BI	36
4.2.4 ESTABLECER LOS REQUERIMIENTOS DE HARDWARE	38

4.2.5 PLANTEAMIENTO DE LAS PRUEBAS	39
5. MARCO METODOLOGICO	40
5.1 BÚSQUEDA Y MINERÍA DE DATOS	41
4.2.6 Búsqueda de las mejores herramientas ofrecidas para BI	43
5.2 ELECCIÓN DE LAS HERRAMIENTAS DE BI	45
5.3 REQUERIMIENTO DE HARDWARE	48
5.3.1 Pruebas a las herramientas	50
5.4 PLANTEAMIENTO DE PRUEBAS	50
5.5 DISEÑO DE DASHBOARD	52
6 RESULTADOS	55
6.1 EVALUACIÓN DE DATOS	55
6.2 ELECCIÓN DE LA MEJOR HERRAMIENTA DE BI	57
6.3 ESTABLECER LA INFRAESTRUCTURA	57
6.4 ANÁLISIS COMPARATIVO	58
6.4.1 Resultados de facilidad de uso de la herramienta	58
6.4.2 Tiempos de carga	62
6.5 PRESENTACIÓN DE DASHBOARD	65
6.5.1 Dashboard y reporte de Power BI	65
6.5.2 Dashboard y reporte de Tableau	70
6.5.3 Dashboard y reporte de IBM Cognos	72
7. CRONOGRAMA	73
8. PRESUPUESTO	75
9. CONCLUSIONES	76
10. RECOMENDACIONES	77
REFERENCIAS	78
ANEXOS	80

ÍNDICE DE FIGURAS

Figura 1 La importancia de la inteligencia de negocios	18
Figura 2 Diagrama de conocimiento y relación al entorno	20
Figura 3 Estructura de inteligencia de negocios y relación al entorno	21
Figura 4 Diagrama de almacén de datos comparado con data mart	25
Figura 5 Diagrama de almacén de datos comparado con un data warehouse	25
Figura 6 Esquema del lago de datos	27
Figura 7 Procesos para el análisis de datos	33
Figura 8 Cuadro mágico de GARTNER	43
Figura 9 Filtración de carreras del periodo 2020 al 2021 con mayor acogida	52
Figura 10 Filtración de calificaciones del periodo 2019 al 2021 de las tres sedes de la Universidad Politécnica Salesiana	53
Figura 11 Filtración de carreras del periodo 2019 al 2021 con menor acogida	53
Figura 12 Resultados de la limpieza automática de Tableau	55
Figura 13 Portada de los datos de fuentes externas no estructuradas	56
Figura 14 menú de los datos de fuentes externas no estructuradas	56
Figura 15 Figura Bigotes de power BI	58
Figura 16 Figura de Bigotes IBM Cognos	59
Figura 17 Gráfico de Bigotes Tableau	59
Figura 18 Paneles de graficas Power BI/Tableau/ IBM Cognos	59
Figura 19 Relación de tablas	60
Figura 20 Cruce de tablas	60
Figura 21 Administrar relaciones	60
Figura 22 Histograma de Resultados comparativos	64
Figura 23 Consulta del promedio total por sede y periodo del 2019 al 2022	66
Figura 24 Consulta del promedio total por sede y sexo de los estudiantes en el periodo 2019 al 2021	66
Figura 25 Consulta del número de cedula del promedio total por sede y periodo académico	66
Figura 26 Consulta del promedio total por carrera y sede	67
Figura 27 Consulta del promedio total por periodo y sede	67
Figura 28 Consulta del promedio total por carrera y sede	68
Figura 29 Dashboard, general de las consultas	68
Figura 30 Reporte General de consultas Power BI	69
Figura 31 Reporte General de consultas Tableau	70
Figura 32 Dashboard de consultas de Tableau	71
Figura 33 Menú de dashboard	71
Figura 34 Reporte General de consultas	72

ÍNDICE DE TABLAS

Tabla 1 Aspectos que considerar al adquirir una herramienta	17
Tabla 2 Ventajas de usar almacenamiento data Marts	24
Tabla 3 Ventajas de usar data warehouse	26
Tabla 4 Tipos de OLAP según la implementación del data warehouse	28
Tabla 5 Requerimiento de Análisis de información	45
Tabla 6 Resumen comparativo de elección de herramientas	47
Tabla 7 Requerimientos técnicos de hardware para Tableau Desktop	48
Tabla 8 Requerimientos técnicos de hardware para Power BI Desktop	49
Tabla 9 Requerimientos técnicos de hardware para IBM Cognos Desktop	49
Tabla 10 Pruebas de interfaz de usuario	51
Tabla 11 Pruebas de rendimiento	51
Tabla 12 Cuadro resumen de facilidad de uso	61
Tabla 13 Especificación de los tiempos de carga de datos	62
Tabla 14 Cuadro de resumen de resultados de las herramientas	65
Tabla 15 Cronograma de actividades	74
Tabla 16 Presupuesto	75

ANEXOS

Anexo 1. Manual de instalación de herramientas

1. INTRODUCCIÓN

Debemos considerar que nos encontramos en una era donde la tecnología avanza rápidamente, por lo que el mundo empresarial no se queda fuera, siendo importante incorporar en la empresa técnicas y herramientas necesarias, como es el caso de la inteligencia de negocios (en sus siglas BI) como apoyo en la toma de decisiones estratégicas.

La inteligencia de negocios hoy en día es indispensable, si no se cuenta con las herramientas y soluciones necesarias, las compañías no se encontrarán al nivel de competitividad de la Industria, considerando que BI simplifica la mecánica de trabajo para así facilitar la investigación, combinación y consulta de datos confiables. Por lo que una empresa exitosa tiene como prioridad la de ayudar a ejecutivos, gerentes y otro tipo de usuario final a la toma de decisiones en beneficio y crecimiento de su negocio, por lo que BI ofrece una variedad de herramientas y aplicaciones tanto para extraer datos internos como de fuentes externas.

Es importante determinar las técnicas y herramientas que permitan analizar una gran cantidad de datos que manejan constantemente las compañías, muy independientemente de su tamaño en la industria; se tiene que identificar las áreas que requieren de una mejora, partiendo de la determinación de estrategias funcionales que permitan analizar de manera concisa los datos, visualizando así los reportes generados con dichas herramientas.

La inteligencia de negocios no es solo una herramienta sino un conjunto de estrategias, análisis del negocio, visualización de datos y mecanismos para manejar grandes volúmenes de datos. Con la finalidad de ayudar a las empresas con la obtención, administración y generación de conocimientos en este proyecto, se ha definido la metodología "M3S" la cual utiliza modelos matemáticos para la obtención de información a partir de varias fuentes de datos. Las empresas deben buscar la mejor herramienta que les permita crear conocimiento a partir de sus datos, por lo que este trabajo se enfocara en una comparativa entre diferentes herramientas.

2. PROBLEMÁTICA DEL ESTUDIO

La gran problemática que se presenta en las empresas es no poder identificar qué tipo de problemas o debilidades se trata de solucionar al implementar inteligencia de negocios. Para *Scott Schlesinger (2018)* "las organizaciones no deben apurarse ni precipitarse en el aprovechamiento de las herramientas de BI, al menos que tengan un caso de negocio distinto".

Para tener más claro lo antes mencionado; la implementación de herramientas de inteligencia de negocios permite garantizar que la obtención de datos se convierta en información útil para la toma de decisiones, sin embargo, el aplicarlo de una forma apresurada o incorrecta, provocara problemas en la empresa como pérdida de información o informes que no se correlaciona con los datos analizados.

Lo más recomendable a primera instancia es proyectar el caso del negocio y determinar el tipo de fuentes de información, para el análisis de datos se requiere de herramientas complementarias de las cuales existe una amplia variedad de herramientas en el mercado.

Al elegir una herramienta, está debe cumplir con las expectativas de mejorar la productividad y facilidad de uso, por lo que el objetivo al implementar este tipo de herramientas siempre debe ser anticipar el comportamiento del consumidor y el mercado. Mediante las herramientas de BI podemos determinar datos importantes, anticipar comportamientos, optimizar el trabajo y detectar posibles errores presentes en una empresa, hoy en día, se maneja una gran cantidad de datos que requieren ser almacenadas y analizados al momento de tomar una decisión sobre el rumbo que la empresa debe seguir. Antes de implementar una herramienta es necesario saber si realmente se ajusta a las necesidades de una empresa y cómo será la aceptación del usuario al implementarla, para ello es necesario un ambiente de pruebas destinado a detectar y corregir posibles errores y capacitar a los usuarios.

Por lo que este proyecto tiene como objetivo realizar un estudio comparativo de las herramientas de BI en funcionamiento y los requisitos necesarios para implementar Inteligencia de negocios.

3. OBJETIVOS DE LA INVESTIGACION

3.1 Objetivo General

Realizar un estudio comparativo de herramientas de inteligencia de negocios y sus prestaciones.

3.2 Objetivos Específicos

1. Conocer acerca de la inteligencia de negocios y la analítica de los datos.
2. Realizar un estudio comparativo y descriptivo de las ventajas y desventajas de las herramientas de inteligencia de negocio disponibles en el mercado.
3. Realizar una prueba práctica mediante la selección de una muestra de las herramientas analizadas.
4. Determinar la facilidad de uso, así como el tiempo de implementación de las diferentes herramientas.

4. REVISIÓN DE LA LITERATURA O FUNDAMENTOS TEORICOS

4.1 INTELIGENCIA DE NEGOCIOS

La inteligencia de negocios (por sus siglas en ingles BI: Business Intelligence) es una combinación de procesos que incluye análisis, minería y visualización de datos que permite a las empresas facilitar la toma de decisiones, sin embargo, para impulsar de manera oportuna la aplicabilidad de inteligencia de negocios es necesario tener una visión clara de los datos de la organización, que puede ser la clave entre el éxito y el fracaso de la organización.

Para Aguilar, L. J. (2019) *“Anteriormente los sistemas de apoyo a la decisión (o por sus siglas DSS) eran el termino dominante en la gestión empresarial y comenzaba a utilizarse una nueva disciplina conocida como inteligencia de negocios, la cual ha ido ganando fuerza y en la que se han integrado los DSS, dado que la idea central gira en torno a los datos de las empresas”.*

En la actualidad las empresas manejan un gran volumen de datos diariamente, por lo que constantemente están buscando un software que permita la administración de estos datos y los conviertan en información útil

El aprovechar eficazmente una herramienta de BI depende de tener un buen sistema dentro de la empresa; el segundo paso es buscar la herramienta correcta que se ajuste a los requerimientos de la empresa, y finalmente permitir a la administración la toma de decisiones en base a la información presentada por la herramienta.

Las herramientas de BI comprenden una gran variedad en el mercado sin embargo se debe considerar aspectos como:

Gestión de datos	Obtener información, procesarla y finalmente exportarla a otro sistema.
Descubrimiento	Analizar la información mediante aplicaciones para realizar predicciones o proyecciones.
Reporte	Representar de forma visual la información obtenida y procesada.

Tabla 1: Aspectos que considerar para adquirir una herramienta

BI permite mediante análisis y minería de datos transformar los datos en información relevante, base para apoyar la toma de decisiones por parte de la administración.

BI se podría resumir en 5 pilares o beneficios para su implementación dentro de una empresa, los cuales son:



Figura 1: La importancia de la inteligencia de negocios (*Fuentes: propias*)

a. Visibilidad mejorada:

La vinculación de sus BI por parte de las organizaciones refleja un mayor control de sus procesos y procedimientos. BI no solamente visualiza todas las áreas de la organización sino también identifica fácilmente las áreas de mejora.

b. Información procesable:

Una herramienta de BI eficaz se usa como medio para identificar patrones y tendencias, es decir que este le permite entender la implicación de varios procesos y cambios organizativos.

c. Mejoras en la eficiencia:

Las herramientas de BI permiten compartir información entre departamentos con la finalidad de ahorrar tiempo, este tipo de intercambio permite enfocarse más en el trabajo que en el procesamiento de los datos.

d. Datos en tiempo real:

Hace referencia al tiempo de espera de los informes compilados por las diferentes áreas de la organización que debe esperar el equipo directivo quienes son los responsables de la toma de decisiones.

e. Conocer al cliente:

Mediante el análisis BI de tendencias y comportamiento del cliente se garantiza conocer la conducta del comprador y esto permite orientar las acciones de la empresa en función de las predicciones identificadas de los clientes.

Para Benítez Y. (2019, p.4), la Inteligencia de Negocios “hace referencia al manejo optimizado de los datos que almacena, recopila y analiza una organización, permitiendo transformarlos en decisiones estratégicas que permiten alcanzar el éxito profesional”.

La inteligencia de negocios abarca procesos para mejores prácticas que ayudan a las empresas a tomar decisiones basadas en la recopilación de datos tanto internas como externas, las cuales son analizadas mediante herramientas de software que responde a las necesidades de quien lo adquiera, por lo que se trata de una evolución de los sistemas de soporte a las decisiones (o DSS).

Para ello es necesario tomar decisiones acertadas en base a la propia experiencia y conocimiento adquirido mediante BI, sin embargo, es necesario que se interprete la información de manera adecuada, siendo necesario considerar que los datos se transforman en información, y esta información en conocimiento, el cual permite conocer cómo está la organización en relación con el entorno en el que se encuentran.

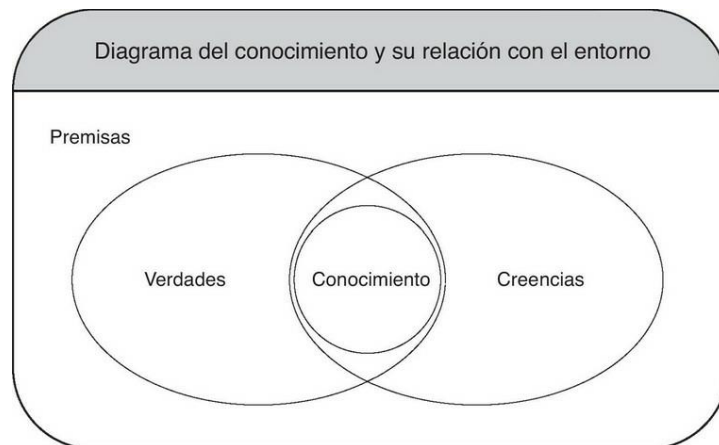


Figura 2: Diagrama de conocimiento y relación al entorno.

Las empresas generan nuevos datos que continuamente que deben ser transformados en información útil. Por lo que métodos tradicionales ya no son eficientes y la implementación de BI permite automatizar el proceso.

4.1.1 ESTRUCTURA DE INTELIGENCIA DE NEGOCIOS

La inteligencia de negocios es una solución de Tecnología de la Información (o TI) que se compone de un conjunto de estrategias, procesos y herramientas. Para tener más claro como es el funcionamiento e implementación de BI en cuanto a su aplicabilidad para la optimización de procesos analíticos, consideramos la siguiente estructura:

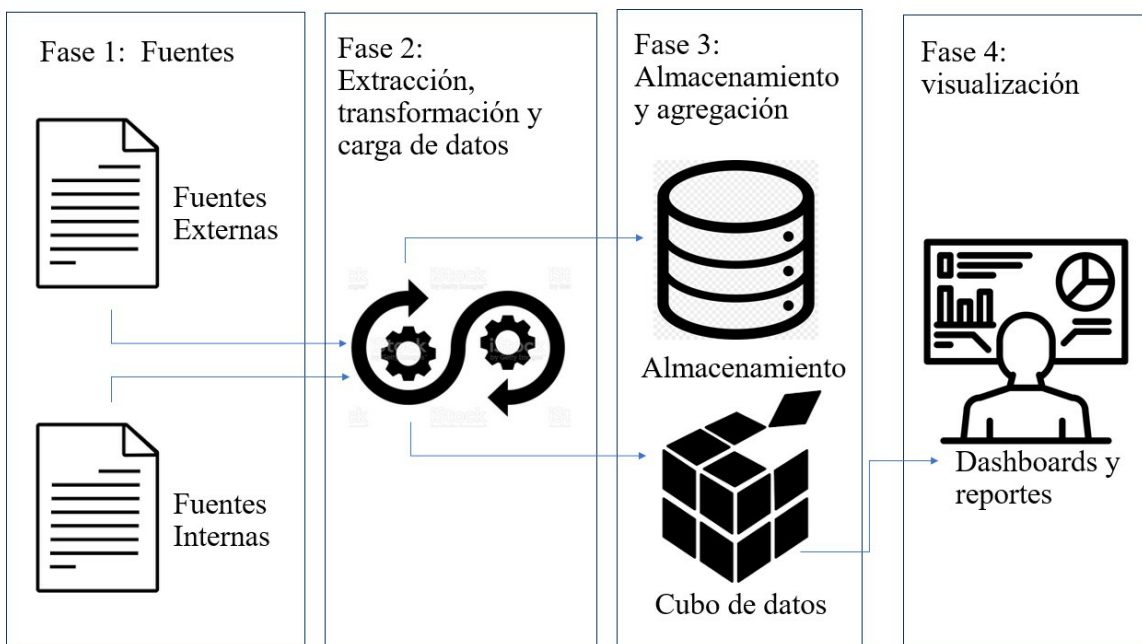


Figura 3: Estructura de Inteligencia de negocios y relación al entorno (Fuente: Alberto Rodríguez Elizabeth Bernal Gamboa)

La siguiente estructura nos permite conocer mediante fases como se aplica BI en la transformación estructural de los procesos analíticos por lo que es necesario tomar en cuenta la figura 3.

4.1.1.1 Fuentes de información

Generalmente el primer paso es identificar las fuentes de información que son recursos que contienen gran variedad de datos de distintos tipos de áreas de una empresa, por ello la primera fase consta de la identificación de las fuentes de datos primarias cuyos datos originales se encuentran en estado bruto, o fuentes secundarias que comprenden a la información reestructurada partiendo de las fuentes primarias para determinar quién va a

utilizar la información, podrían ser los supervisores de ventas, el gerente general, o el CEO de la empresa, etc. Un ejemplo a este paso puede ser una pregunta por parte del gerente de cuál es el producto más vendido o el resumen de ventas del trimestre anterior.

Una vez definida la clasificación de la fuente que se analizará se requiere identificar el tipo de fuente de datos mismo que puede ser interna o externa:

- **Fuentes de datos interna:** son todos aquellos datos generados dentro de la misma empresa, procede de la recopilación de datos como resúmenes de ventas, inventarios, etc., cabe destacar la importancia del CRM en cuanto a la gestión de la información, considerando que este tipo de fuentes procede de la investigación propia de la misma empresa que puede intercambiarse tanto a clientes como proveedores.
- **Fuentes de datos externa:** son aquellos datos compuestos por sitios ajenos a la empresa que provienen de agencias o empresas de investigación, por tal razón son cada vez más relevantes, puesto que complementan un mejor análisis juntamente con los datos internos.

Hay que resaltar y considerar que cada fuente de datos se clasifica en fuentes de datos primarias como secundarias así que es necesario identificar la fuente a analizar.

Según la madurez del negocio tendremos inteligencia de negocios o análisis de datos avanzados. Dentro de la inteligencia de negocios podremos hacer un análisis descriptivo (que pasó) y un análisis diagnóstico (por qué pasó). Sin embargo, si una empresa ya utilizado inteligencia de negocios por varios años podrá hacer un análisis predictivo (que va a pasar) y un análisis prescriptivo (que quiero q pase).

4.1.1.2 Proceso ETL

La siguiente fase hace énfasis al proceso de extracción, transformación y carga de datos o en sus siglas ETL, es decir es el proceso mediante el cual se obtiene información de una o varias fuentes, transforma datos brutos para cargarlos en un sistema de destino como un Data Warehouse.

ETL es parte esencial dentro del ecosistema de datos de cualquier empresa, así también este proceso se usa en el procedimiento OLAP (almacén de datos).

Debido a que los datos provienen de diferentes tipos de fuentes, éstos pueden tener dos

tipos; los convencionales donde la información podrá estar almacenada en archivos de Excel, archivos de texto o base de datos y los no convencionales donde podrán estar las redes sociales, grabaciones o conversaciones con nuestros clientes. En este punto se implementará el proceso visto anteriormente ETL por lo que debe transformarse de manera diferente antes de aplicar para BI o ser analizadas.

Estos procesos no son herramientas o técnicas específicas, sino una colección de lo que se debe realizar para obtener información y conocimiento de los datos. ETL está conformado por 4 procedimientos los cuales son:

- Limpieza de datos: Se cargarán datos solo relevantes para el negocio y se evitara datos contradictorios.
- Integración de datos: Permite unificar las diferentes fuentes de datos.
- Carga de datos: Se copia o transfiere datos al Data Warehouse.
- Actualización de datos: Los datos se actualizarán según se reciba nueva información.

4.1.1.3 Almacenamiento y Agregación

Es considerado el paso más importante durante la transformación de la información es donde se utiliza aproximadamente el 70% del tiempo aquí se depura los datos descartando datos innecesarios y toman en cuenta clasificándoles en dos tipos cuantitativos y cualitativos.

La tercera fase comprende al almacenar de datos de tres diferentes tipos como el Data Mart, Bodega de datos (Data Warehouse DW) y el lago de datos.

4.1.1.3.1 Data Mart

El data mart es una manera simple de almacenamiento de datos del cual solo se enfoca en un solo tema ya sea del área de ventas, finanzas o marketing la razón es porque el data mart puede extraer datos de menor cantidad de fuentes que pueden incluir sistemas operativos internos como externos.

En la tabla 3 se presenta las ventajas del uso de data Mart.

Acceso rápido	Una vez establecida la conexión con el origen de la fuente los usuarios y equipos pueden acceder rápidamente a los datos deseados.
Estadísticas más rápidas	El data mart permite un análisis de datos a nivel de departamento, esto permite centrarse a retos y oportunidades de áreas específicas como finanzas, recursos humanos, etc.
Fácil implementación	Data mart se centra en las necesidades de determinados equipos por lo q se necesita menos datos y no requiere de mucho tiempo y esfuerzo.
Gestión de datos ágiles	Utiliza información recopilada de proyectos anteriores q ayudan a las tareas actuales es decir que el data mart varía según el análisis de un proyecto nuevo o en evolución.
Análisis de datos	El análisis de datos es de corta duración ya que se centra en un tipo específico de datos de un área.

Tabla 2: ventajas de usar almacenamiento data Marts

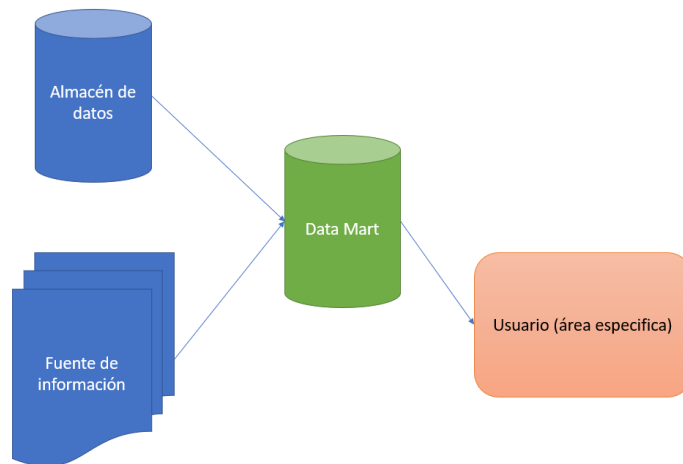


Figura 4: Diagrama de almacén de datos comparado con data mart. (Fuente: Sap) [/https://www.sap.com/latinamerica/insights/what-is-a-data-warehouse.html](https://www.sap.com/latinamerica/insights/what-is-a-data-warehouse.html)

4.1.1.3.2 Bodegas de datos (Data Warehouse)

Según Benítez Y. L “el Data Warehouse DW es un contenedor que almacena datos procedentes de fuentes con distintas funcionalidades, integrándolas desde las actividades operacionales y rutinarias de la organización”.

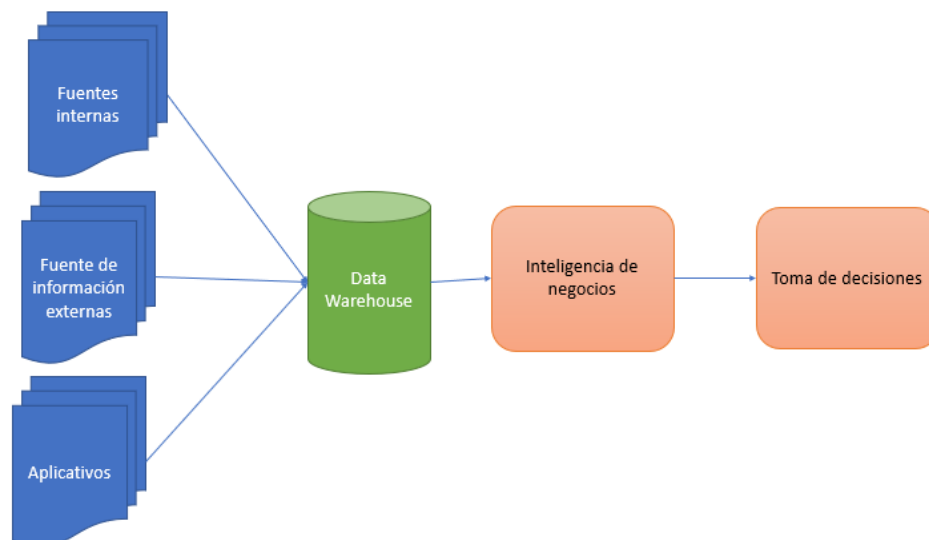


Figura 5: Diagrama de almacén de datos comparado con un data warehouse. (Fuente: Sap) [/https://www.sap.com/latinamerica/insights/what-is-a-data-warehouse.html](https://www.sap.com/latinamerica/insights/what-is-a-data-warehouse.html)

Data Warehouse en esencia es una arquitectura que soporta un gran volumen de datos actuales e históricos asumiendo el papel de integrador de todas las diferentes fuentes de datos, los usuarios pueden acceder a ella mediante una herramienta, por lo que su objetivo es recopilar y respaldar el proceso de toma de decisiones.

Los almacenes de datos se encuentran diseñados para manejar datos estructurados y no estructurados que son aprovechados para acceder a los datos en tiempo real impulsando confianza al momento de tomar decisiones, en la siguiente tabla se determina las ventajas del almacenamiento Data Warehouse.

Analítica	La responsabilidad de la toma de decisiones debe basarse en datos de fuentes con información completa.
Consultas rápidas	Con DW se puede realizar consultas a gran cantidad de datos al menor tiempo posible.
Calidad de datos	Los casos de limpieza de datos se crean y se ingresan en listas de trabajo para su posterior procesamiento antes de cargarlos en DW, lo que garantiza que los datos se conviertan en un formato coherente para respaldar el análisis.
Información histórica	Al almacenar datos históricos esto permite a los responsables de la toma de decisiones aprender de las tendencias y los desafíos del pasado, permitiendo hacer predicciones e impulsar un mejoramiento.

Tabla 3: ventajas de usar Data Warehouse

4.1.1.3.3 Lago de datos

Un lago de datos hace referencia a los repositorios centralizados que permiten almacenar datos estructurados y no estructurados, a su vez se pueden ejecutar diferentes tipos de análisis desde paneles y visualizaciones hasta el procesamiento de big data con el fin de realizar un aprendizaje automático para la toma de decisiones.

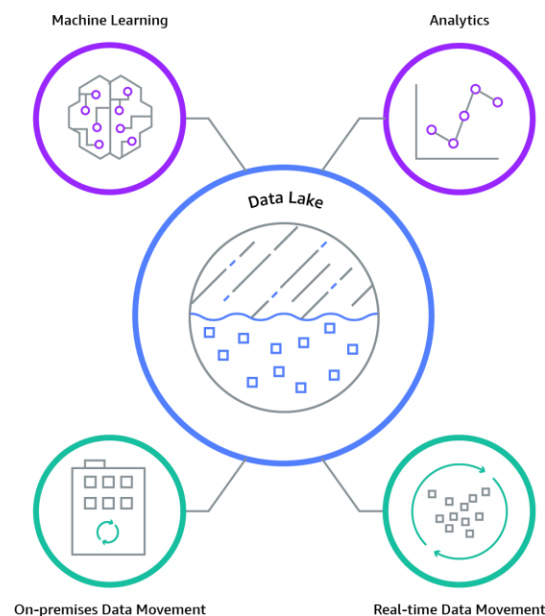


Figura 6: Esquema del lago de datos (Fuente: Aws) /<https://aws.amazon.com/es/big-data/datalakes-and-analytics/what-is-a-data-lake/>

Al almacenar una gran cantidad de datos estos pueden encontrarse sin procesar o no han sido analizados para un propósito específico, por lo que los datos no se colocan en el lago de datos hasta que se consultan.

4.1.1.4 Agregación:

La fase de agregación hace énfasis a los cubos de datos OLAP (procesamiento Analítico en Línea) son una arquitectura de datos que supera las limitaciones de las bases de datos, los cubos pueden visualizar y agregar gran cantidad de datos, esto permite a los usuarios realizar búsquedas apropiadas.

Los Cubos OLAP son una característica propia del administrador de servicios que utiliza la infraestructura de almacenamiento de datos existentes que proporciona funcionalidades de inteligencia empresarial a los usuarios finales.

La presentación de datos analizados con OLAP se denomina visualización y utiliza herramientas y formas de representación gráfica que permite efectuar análisis de tendencias, puntos de equilibrio y sensibilidad.

La técnica OLAP es conocida como Análisis Multidimensional y contrapuesta a la de Procesamiento Transaccional en Línea (OLTP). Por lo que existe tres tipos de OLAP, cuya variación radica en la implementación del Data Warehouse, en la siguiente tabla se da a conocer los tipos de OLAP según el data Warehouse implementado:

Tipos de OLAP	Implementación de data Warehouse
ROLAP	Este tipo de Olap se efectúa sobre DW implementados como tablas relacionales (usando modelos star schema y snowflake).
MOLAP	Olap sobre DWs implementados sobre bases de datos multidimensionales.
HOLAP	Este tipo de Olap es un enfoque híbrido de ROLAP y MOLAP.

Tabla 4: Tipos de OLAP según la implementación data Warehouse.

4.1.1.5 Visualización

Finalmente, la fase cuatro comprende a la visualización de la información se la realiza mediante reportes o dashboards, los cuales deben ser fáciles de entender para los directivos, estos resultados son obtenidos por parte de una herramienta, por tal razón se debe considerar que el diseño y construcción del panel donde se refleja la información debe comprender alguno de los dos tipos como:

- **Reportes:** Es un documento impreso o digital acompañado de graficas que permite resumir la información.
- **Dashboard:** La construcción del Dashboard debe ser un resumen de todo el análisis que hemos realizado, sin embargo, no debe ser complejo ni estar sobrecargado de información, debe estar presentado de una forma sencilla de gráficos para que la alta gerencia pueda entenderla con rapidez y aprovechar la información.

4.1.2 IMPORTANCIA DE INTELIGENCIA DE NEGOCIOS

La aplicación de la técnica de Inteligencia de Negocios permite manejar, procesar y actualizar una gran cantidad de datos y relacionarlos entre sí, con la finalidad de extraer información útil con gran valor empresarial.

En el marco empresarial, donde los negocios compiten en un entorno económico complejo, el valor de una empresa está determinada por la información que esta posee sobre sí misma y el mercado.

La inteligencia de negocios aplicada a cualquier tipo de organización genera nuevas oportunidades de negocio, esto influye directamente en las decisiones que se toman a nivel superior o estratégico y determinan el rumbo de la empresa.

Para entender mejor como influye la inteligencia de negocios dentro de la organización hay que considerar los niveles de gestión empresarial que permiten planificar los movimientos y como se mantendrá el rumbo empresarial:

- **Nivel Estratégico:** este nivel trata de la visión que mueve las acciones de la empresa, es parte del rol principal a la hora de definir estrategias para la organización, por ello deben contar con información, analizarla, monitorear las tendencias y evoluciones, valorando los objetivos estratégicos de la organización para la toma de decisiones rápidas.
- **Nivel Táctico:** los departamentos son los encargados de crear la planificación específica que haga realidad la estrategia para llegar a soluciones oportunas.
- **Nivel Operativo:** este último nivel proporciona a los empleados información directa y clara sobre los procedimientos a seguir mediante instrucciones precisas: La información se presenta mediante reportes con datos actualizados.

BI busca un negocio inteligente, en el que las tomas de decisiones sean en base a la información obtenida tanto por datos internos como externos.

Por lo que al terminar la implementación de BI se busca que la empresa obtenga los siguientes resultados:

- El reconocimiento de la experiencia; mientras más datos se obtengan más crecerá la empresa.
- El análisis de datos contextualizados; los datos se conviertan en información.
- La capacidad de extraer e integrar datos de múltiples fuentes; ser capaces de minar datos.
- El procesamiento de los registros obtenidos en información útil para conocimiento del negocio.
- Búsqueda de causas y efectos, trabajando con hipótesis y desarrollando estrategias y acciones competitivas.

En la práctica, la implementación de la inteligencia de negocios significa tener una comprensión integral de todos los datos de la organización. Además, implica el uso de estos datos para impulsar el cambio, eliminar ineficiencias y adaptarse rápidamente a los cambios en el mercado o la demanda.

4.1.3 PORQUÉ IMPLEMENTAR

Las aplicaciones, herramientas y prácticas de BI mejoran el acceso y el análisis de la información

La implementación de BI proporciona diversos beneficios entre los que podemos destacar:

- Permitir una visión única, conformada, histórica, persistente y de calidad de la información.
- Crear, manejar y mantener métricas, indicadores y claves de rendimiento.

Los departamentos y áreas de una organización más beneficiadas son:

- **Finanzas y contabilidad.** Las soluciones de BI no solo ayudan a trazar cifras de presupuesto e inversión, también brindan un dashboard sobre problemas financieros, salud general, etc. según lo que deseamos visualizar.
- **Recursos humanos.** Permite rastrear la productividad de los empleados, permitirá trazar objetivos y recompensas, análisis de satisfacción de colaboradores y más.

“Con la información ingresada a la base de datos que realiza cada departamento, los administradores pueden determinar cuáles son las debilidades o puntos críticos que generan mayores gastos con el propósito de aplicar estrategias adecuadas para reducir dichos gastos”. Cordero-Naspud, E. I. (10 de Julio-Diciembre de 2020).

4.1.4 INDICADORES QUE CONSIDERAR PARA CONOCER CUANDO SE DEBE IMPLEMENTAR

Lo óptimo o recomendable es implementar BI al iniciar operaciones, para sembrar el hábito de utilizar todos los datos obtenidos. Sin embargo, muchas empresas no lo realizan dicha implementación por presupuesto o desconocimiento,

Por lo que se han planteado varios indicadores para saber cuándo una empresa necesita implementar BI:

- Cuando no se conoce del estado de salud financiera y de los resultados del negocio.
- Cuando no hay certeza del efectivo que se dispone para gastos diarios.
- Cuando no se sabe el estado de venta de los productos.
- Cuando no se conoce a los clientes.
- Cuando se pierde la cuenta de cuánto nos deben y desde cuándo.
- Cuando no se tienen claros los gastos que realiza la empresa.
- Cuando no se conocen los ingresos y márgenes de utilidad.

En resumen, se debe implementar BI cuando no conocemos las actividades que realiza nuestra empresa y queremos aprovechar los datos obtenidos para hacer crecer el negocio.

4.1.5 ANÁLISIS DE DATOS

El análisis de datos es una parte fundamental en la implementación de inteligencia de negocios, permite transformar los datos obtenidos de fuentes heterogéneas en información para lograr un correcto análisis, para ello se necesita comprender cuál es el objetivo o la pregunta que se requiere responder; Como por ejemplo ¿cuál es el producto más vendido?, o ¿cuál es la satisfacción del cliente con nuestro producto?

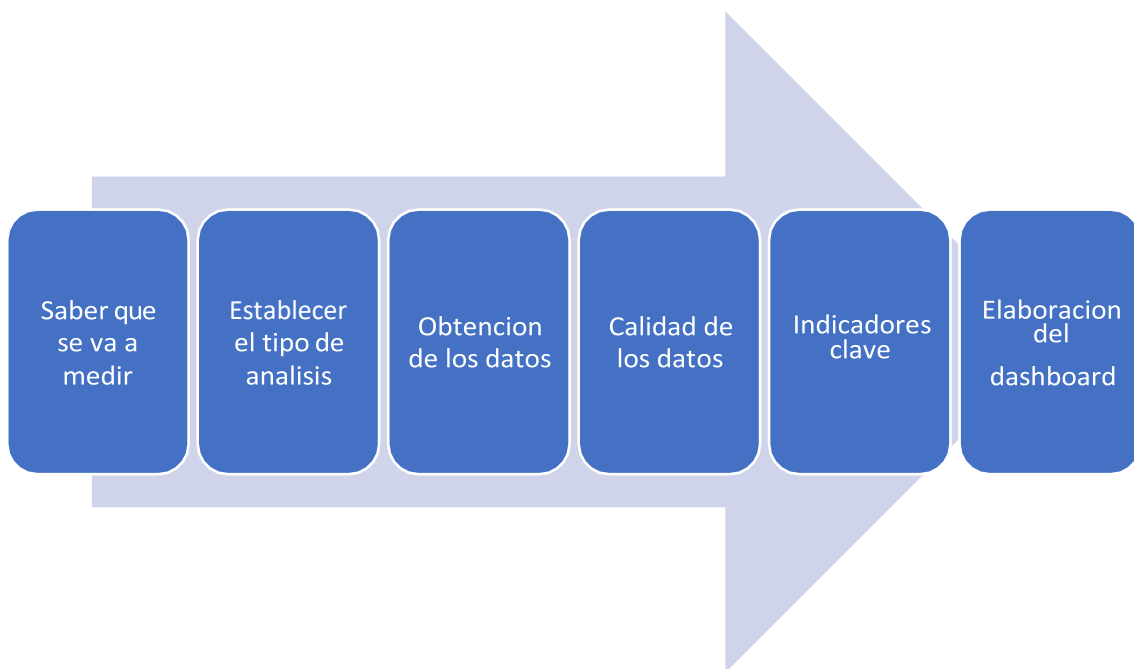


Figura 7: Procesos para el análisis de datos (*Fuente: propia*)

El análisis de datos permite obtener información, para lo cual las empresas recolectan datos constantemente. Por ejemplo, en una empresa de venta de neumáticos se almacenará datos de los clientes a los cuales vendemos y el producto vendido; sin embargo, si los datos recabados solamente son almacenados y no analizados lo que generaremos es un registro histórico. Al no aprovechar los datos recolectados estamos perdiendo oportunidades de crecimiento.

El negocio que simplemente almacena sus datos para un resumen anual se verá perjudicado. La toma de decisiones se da constantemente por parte de la gerencia por lo que dicha área debe estar actualizada con el comportamiento de la empresa y el mercado para tomar las decisiones más acertadas.

La gerencia y directiva de la empresa deben estar al tanto de cómo se mueve el mercado y la aceptación de los productos o servicios por parte del cliente. Por lo que tener la información de forma oportuna ayuda a la toma de decisiones, y por lo tanto al crecimiento del negocio. Para ayudar con esta tarea se deben implementar controles y objetivos a cumplir del crecimiento esperado de la empresa, en caso de no cumplirlos se deberá revisar el proceso y mejorarlo.

Las herramientas de BI nos ayudan con este proceso, ya que permiten administrar de una mejor manera la transformación de los datos, y nos permiten interactuar entre las diferentes áreas de la empresa permitiendo un análisis colaborativo de la información y obtener resúmenes de la empresa de una forma clara y rápida.

El análisis de datos no es un proceso nuevo, sin embargo, si aprovecha nuevas técnicas como es el caso de la inteligencia artificial, aprovechando el rendimiento de los ordenadores modernos para simular el razonamiento humano al momento de leer un dato y predecir cuál será el comportamiento del usuario.

Por ejemplo, la red social Twitter en sus inicios fue pionera en implementar la inteligencia artificial para entender e interpretar lo que quieren sus usuarios. Para ello decidieron desarrollar una IA capaz de leer el tuit y entender la emoción que intentaban transmitir.

.

4.2 METODOLOGIA M3S

4.2.1 CONCEPTO M3S

Se eligió la metodología M3S para el proyecto de estudio comparativo de herramientas de inteligencia de negocios.

M3S es una metodología usada en la implementación de proyectos de inteligencia de negocios (I.N.), la cual permite obtener los mejores resultados de extracción de datos, determinado por un proceso adecuado, permitiendo tener mayor conocimiento y valor en la información.

Esta metodología se sustenta en el algoritmo de ID3 el cual comprende a los árboles de decisiones. La metodología M3S propuesta para la comparación de herramientas de BI que existen en el mercado, busca resaltar los beneficios que presenta cada una de ellas, por lo que se utilizara una fuente de datos internas y externas.

Para el desarrollo de este proyecto, la metodología M3S es la más adecuada para determinar resultados de casos de estudio específicos en base a herramientas de extracción de información, en base a la metodología propuesta el proceso de evaluación para las herramientas de BI se basará en los siguientes requisitos:

4.2.2 BÚSQUEDA Y MINERÍA DE DATOS

Hace referencia al conjunto de técnicas y tecnologías para la exploración de grandes bases de datos, tal como tal como se muestra en la figura 8 mediante la búsqueda y minería de datos podemos obtener datos, por lo que este proceso debe encontrar anomalías, patrones de grandes conjuntos de datos que permitan predecir resultados, todo con la finalidad de aportar información valiosa y así tomar decisiones futuras.

La minería de datos contiene ciertas ventajas como, por ejemplo:

- Mediante combinaciones de los algoritmos se puede obtener información inesperada.

- Capaces de analizar grandes cantidades de datos
- Fácil interpretación de resultados por lo que no es necesario tener conocimientos previos en informática.

4.2.3 ELECCIÓN DE LA MEJOR HERRAMIENTA DE BI

El proceso de evaluación y selección de las diferentes herramientas de BI comienza con el cumplimiento de ciertos requisitos o factores que determinan los aspectos prioritarios en tomar en cuenta en la elección.

Una vez definida la lista de requisitos, hay que evaluar adecuadamente a los proveedores y descartar aquellos que no ofrezcan alguno de los requisitos principales.

Recordemos que no existen requisitos ya definidos, es decir que para la elaboración de una lista de requisitos se puede desarrollar desde áreas con requisitos comunes o mediante al cuadro mágico de Gartner, partiendo del alcance, la funcionalidad, experiencia de uso, costo etc. Ninguna herramienta es buena o mala, puesto que se pueden ajustar a las necesidades organizacionales, por lo que deben ser evaluadas con objetividad durante el proceso de comparación.

La elección de las herramientas de BI se realiza considerando los siguientes requerimientos considerados como aspectos necesarios de cumplir:

- Herramientas que disponen de versiones gratuitas
- Herramientas que disponen de 30 días de pruebas
- Herramientas basadas en análisis estadísticos
- Tipo de archivos con extensión .xlsx y .csv que acepta la herramienta
- Tamaño máximo de datos de carga
- Tipos de carga fuentes de datos que admite la herramienta (Hojas de cálculo, Bases de datos)
- Facilidad de uso
- Compatibilidad con sistema operativo Windows

La lista de requerimientos comprende aspectos comunes que presenta cada herramienta actualmente, es decir la mayoría de las aplicaciones para pc presentan versiones gratuitas

que contienen limitaciones a las versiones premium, así también las pruebas por 30 días permiten conocer si es recomendable adquirir o no el aplicativo, sin embargo, en caso de las herramientas basadas en análisis estadísticos no aplica para todas las herramientas de BI, así también estas herramientas admiten fuentes de datos ya sea de archivos con extensiones específicas como de bases de datos, esto resulta más comprensible al momento de determinar la facilidad de uso y la compatibilidad con el sistema operativo.

La selección de herramientas basadas en los criterios antes mencionados y también basados en el cuadro mágico de Gartner cuyos datos reflejan la posición de las herramientas con mayor aceptación

4.2.4 ESTABLECER LOS REQUERIMIENTOS DE HARDWARE

Los requisitos de hardware son aquellas características que debe tener un ordenador para ejecutar una aplicación, así también es necesario que los requisitos de software instalado sean tomados en cuenta ya que las características determinaran si existe compatibilidad con la aplicación, para ello hay que determinar características físicas y lógicas al momento de realizar pruebas comparativas para el análisis de datos.

Los requerimientos funcionales de hardware describen cualquier actividad que este realice, este comportamiento o función en particular de una aplicación determina las condiciones que cumple.

Los criterios para cumplir en cuanto a los requerimientos físicos y lógicos del ordenador serán:

- Laptop o PC de escritorio
- Procesador Intel i7
- Memoria RAM de máximo 16GB
- Almacenamiento en disco de mínimo 500GB
- Sistema operativo Windows 7 o más actualizado

4.2.5 PLANTEAMIENTO DE LAS PRUEBAS

El planteamiento de pruebas es parte de la selección que determinará que herramientas cumplen con los requisitos o criterios planteados, por lo que, en esta sección, el primer paso es determinar cuáles herramientas disponen de una versión o prueba gratuita en la que se puede instalar el aplicativo y descartar las herramientas que no cumplen el requisito ya sea por incompatibilidad o dificultad de instalación.

Una vez descartadas las herramientas que no cumplen el requisito anterior, se procede a identificar las fuentes de datos que serán analizadas mediante las herramientas seleccionadas, esto quiere decir que las herramientas deberán generar informes, cuadros de mando y consultas, con la finalidad de comprobar la facilidad de uso. Así también los conjuntos de datos serán exclusivamente de archivos de extensión (.csv / .xlsx) para un mejor muestreo, por lo que se determinará la capacidad, tiempo de carga y tamaño de archivo que admitirá cada herramienta.

Para ello se determinan los mismos escenarios para todas las herramientas a evaluar, con la finalidad de observar el comportamiento, tiempos de carga, resultado final de cada herramienta al analizar los mismos datos, la misma cantidad de datos tanto con las mismas condiciones físicas como lógicas.

5. MARCO METODOLÓGICO

Para este estudio comparativo de herramientas de BI, se plantea como una necesidad estratégica el análisis de la falta de carreras ofertadas en las universidades del Ecuador, para ello se pretende elaborar dashboards que reflejen las carreras con menor y mayor acogida por los estudiantes mediante sus calificaciones.

En base a la necesidad de la falta de carreras ofertadas en las universidades del Ecuador, se establece realizar el análisis de los siguientes datos:

- **estructuraDatos_v3** cuyo archivo contiene un registro de tablas de las calificaciones de los estudiantes de las tres sedes de la Universidad Politécnica Salesiana, sin embargo, este documento no contiene datos concretos de los estudiantes y docentes por el motivo de mantener confidencialidad y anonimato,
- **Oferta Vigente** cuyo archivo no contiene un registro estructurado de tablas de las carreras ofertadas cuya fuente fue obtenida de la Senecyt, mediante un proceso más o menos complejo de compilado, se elaborará un dashboard con el objetivo de determinar en qué herramienta se acoplan mejor.

Una vez definida la necesidad estratégica y los datos a analizar, se aplicará los pasos de la metodología definida, para realizar una comparativa de herramientas, en base a las especificaciones y características, y como procesan los datos en cada una de ellas.

La metodología M3S permite transformar los datos en información, siguiendo su proceso. Primero se analizan los datos y se normalizan para su correcta implementación. Segundo se estableció los requerimientos mínimos de las aplicaciones y se empleó un equipo que no genere cuello de botella para las pruebas. M3S establece en su cuarto punto que debe existir un diseño de las pruebas, por lo que se diseñó un ejercicio que permita implementar el dashboard.

5.1 BÚSQUEDA Y MINERÍA DE DATOS

El análisis por realizar es con almacenamiento data mart ya que su enfoque de datos es del tipo académico, por lo que, en base a la necesidad de información de la falta de carreras ofertadas, se obtuvo un conjunto de datos de hojas electrónicas de calificaciones y ofertas académicas tanto de fuentes internas como fuentes externas.

En este escenario se revisaron y analizaron un total de 13 hojas de cálculo con más de 500 mil registros, sin embargo, solamente dos archivos se tomaron en consideración para el análisis estos fueron **estructuraDatos_v3** el archivo detalla los siguientes campos en de acuerdo con las siguientes hojas las cuales son:

- **Estudiantes:** Cédula (código secuencial), Estudiante (código generado secuencialmente), Sexo, Fecha de nacimiento, Colegio de origen, Ciudad del colegio, fecha de graduación de bachiller, Quintil socioeconómico.
- **Docentes:** Cedula (código generado secuencialmente), Sexo, Título tercer nivel, Título Cuarto nivel, Dedicación y calificación.
- **Sede:** Matriz Cuenca, Quito Guayaquil.
- **Carreras:** Sede, Carrera, tipo de proyecto, Modalidad.
- **Tipo de proyecto:** Unificado, Nueva o Rediseño, ajuste curricular/ proyecto de carrera nuevos.
- **Materias:** Código, Carrera, Materia.
- **Periodo académico:** Código, periodo académico.
- **Calificaciones:** Periodo académico, sede, carrera, materia, nivel, grupo, tipo de inscripción, Nota primer y segundo Inter ciclo, total y estado.

A si también el archivo de **Oferta_Vigente** detalla los siguientes campos de acuerdo con las siguientes hojas:

- **Becas por área:** Área de estudio, periodo académico.
- **Oferta:** Nombre de la institución, Tipo de institución, tipo de financiamiento, Carrera, Área de estudio Nivel de formación, tipo de formación, provincia, total Oferta.

Estos archivos pertenecen a los datos internos de la Universidad Politécnica Salesiana y datos de las ofertas vigentes a la Secretaria Nacional de Educación Superior, Ciencia y Tecnología Senescyt.

Tal como se planteó en el proceso de análisis de datos (Figura 8) dentro de la búsqueda y minería de datos se determina los registros de los estudiantes de la Universidad Politécnica Salesiana y los registros de la secretaria de educación el cual responde al primer paso del proceso de análisis de datos que se va a medir y el tipo de análisis establecido.

Fuentes de datos interna: La Universidad Politécnica Salesiana ha proporcionado datos para su análisis. Dicho archivo pesa 46,4MB.

1. Primer archivo: Contiene ocho tablas con el registro de datos de estudiantes y docentes en el periodo 2019 – 2021 de las tres sedes.
2. Segundo archivo: Script con el cual se extrajo la información
3. Tercer archivo: Datos de calificaciones, periodo académico, profesores, estudiantes, materias, carreras de las 3 sedes de la universidad.

Fuentes de datos externa: Se investigó en diferentes sitios web para la extracción de información relacionada con el tema.

1. Universidades del Ecuador: De la página de la Senescyt se extrajo información de títulos registrados, oferta académica, docentes, matriculas y becas. Peso 12.9MB
2. Universidades Internacionales: Registro de egresados en universidades de España. 284KB

Se usarán dos archivos con extensión .xlsx para las pruebas de rendimiento de las herramientas de BI, ya que el mismo cuenta con un peso de 1.20GB.

5.1.1 Búsqueda de las mejores herramientas ofrecidas para BI

El análisis de datos es una parte fundamental en la implementación de inteligencia de negocios; para facilitar este análisis se han creado diversas herramientas las cuales tienen sus ventajas y desventajas, sin embargo, todas permiten transformar datos en información de una forma sencilla y rápida, para que las empresas puedan aprovecharla de una forma oportuna y tomar las mejores decisiones en base a ellas.

Para facilitar la elección de herramientas, éste estudio se basa en el análisis anual de las herramientas en el mercado que realiza la consultora Gartner Inc. Con el objetivo de saber en qué punto de innovación y nivel de desarrollo están las organizaciones dedicadas a la tecnología, se aplica el denominado cuadro mágico de Gartner, el cual nos permite dar una idea de en donde se encuentra la herramienta en cuanto a calidad y desempeño.

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms



Source: Gartner (March 2022)

Figura 8: cuadro mágico de GARTNER (Fuente: Gartner)/ <https://softwarehardware.com/software/software-empresarial/cuadrante-magico-de-gartner/>

Según Gartner, el cuadro mágico está dividido en 4 zonas a calificar a la empresa y su herramienta (figura 9), las cuales son:

- Los aspirantes (Challengers) tienen una buena ejecución del negocio actualmente y son capaces de dominar un gran segmento de mercado, pero aún no demuestran un real entendimiento de la dirección en que va el mercado.
- Los líderes (Leaders) desarrollan bien su negocio con una adecuada visión actual del mercado y están bien posicionados para el futuro.
- Los jugadores de nichos específicos (Niche Players), se enfocan con éxito en un segmento de mercado específico, pero muchos de ellos no adquieren en su desempeño una visión global y no se caracterizan por hacer grandes innovaciones o por superar a sus competidores.
- Los visionarios (Visionaires), entienden hacia dónde van los mercados e incluso pueden tener una idea para cambiar las reglas y paradigmas, pero aún no son capaces de llevar a cabo estas ideas por completo o con éxito.

De acuerdo con el cuadro mágico de Gartner (Figura 9), se puede destacar que Microsoft y Salesforce ambas poseen herramientas que son líderes en el mercado por lo que Power BI y Tableau son las dos opciones más recomendables, a su vez como se menciona en los puntos anteriores se ha demostrado las razones de la elección de estas herramientas.

Sin embargo, IBM también posee una herramienta la IBM Cognos analítica que se posiciona en entre las mejores herramientas visionarias.

5.2 ELECCIÓN DE LAS HERRAMIENTAS DE BI

Para el proceso de análisis de datos (figura 8), en base a la información recopilada de las herramientas de BI y los datos obtenidos de las universidades, se ha realizado el análisis de las herramientas bajo los siguientes criterios establecidos para la elección de las herramientas según el manejo de datos y el funcionamiento de las aplicaciones en las diferentes herramientas:

	Power BI	Tableau	Domo	Sisense	IBM Cognos	Google analítica	Oracle BI
Versión gratuita		Dispone de versión gratuita	Dispone de versión Demo			Dispone de versión gratuita	
Prueba por 30 días	Dispone de prueba por 30 días			Dispone de prueba por 30 días			Dispone de prueba por 30 días
Costo mensual de adquisición	Costo de \$10	Costo de \$70	Sin costo	Costo de \$10	Costo de \$10	Gratuito	Costo de \$150
Basada según análisis estadístico	Herramientas basadas en análisis estadísticos y matemáticos						
Admite archivos .xlsx o .csv	Admiten ambos tipos de extensión			Admiten ambos tipos de extensión			
Tamaño máximo de carga de datos	1GB	300MB	100MB	100MB	100MB	100MB	1GB
Tipos de fuente (hojas de cálculo, base de datos)	Hojas de cálculo			Hojas de cálculo			
Facilidad de uso	Intuitiva	Intuitiva		Intuitiva	Intuitiva	Intuitiva	
Compatibilidad con Windows	Compatible Windows 10/11		Compatible Windows 10	Compatible Windows 10/11		Compatible Windows 10	

Tabla 5: Requerimiento de Análisis de información

Mediante el proceso de selección y la experiencia adquirida al usar herramientas BI, se descartan las siguientes:

- **Domo** por su interfaz poco intuitiva y más complicada de implementar, sin embargo, se reconoce que es una de las herramientas más utilizadas.
- **Oracle BI** por problemas presentados en la página al momento de descargar su versión de prueba.
- **Sisense** es un caso especial ya que la plataforma es intuitiva, fácil de usar, sin embargo, al momento de salir de la aplicación generaba un conflicto con la contraseña y recuperar la cuenta también se hacía complicado ya que los correos

no siempre llegaban.

- **Google analítica** es intuitiva sin embargo esta herramienta tiene como prioridad la ejecución y planeación de análisis de sitios web el cual no aplica en este proceso.

Dado que las herramientas mencionadas no entran en este proceso de selección, se optó por herramientas que cumplieron las expectativas de manera que su análisis fue más completo que las otras herramientas cuya limitación por ser versiones de prueba no aportaban con solucionar problemas.

Por lo que, como referencia al cuadro mágico de Gartner, y la información obtenida en base al tema de las universidades se pudo comprobar que las siguientes herramientas cumplieron con las expectativas deseadas de acuerdo con los procesos de analítica, carga de datos, y facilidad de uso estas corresponden a:

- **Tableau:** Tableau realiza un análisis de datos completo y su proceso de instalación de la herramienta del Tableau es sencilla. (Anexo Manual de Usuario). Para las pruebas se optó por Tableau Desktop el cual comprende a la última actualización del 2022.2 por lo que se da a conocer la instalación de la herramienta, el tamaño máximo de datos es de 300MB de carga mediante comprobaciones con datos reales.
- **Microsoft Power BI:** También realiza un análisis completo de datos, sin embargo, el límite de tamaño del conjunto de datos es de 1 GB. Para conjuntos de datos de mayor tamaño, puede usar Power BI Premium. Con relación a los valores distintos en una columna: cuando se almacenan datos en caché en un conjunto de datos de Power BI (a veces denominado modo "de importación"), el número de valores distintos que se pueden almacenar en una columna es de 1999 999 997.

- **IBM Cognos:** Su tipo de análisis a diferencia de los anteriores permite acelerar y mejorar la fiabilidad al momento de preparar los datos previos, también consta de pruebas gratuitas antes de la adquisición para manejar la herramienta sin embargo al ser una versión de pruebas puede ser limitada y con tiempos de carga mayores, por lo que como toda herramienta de paga conecta datos, crea visualizaciones y permite reflejar el estado actual de la compañía al tiempo que consta de aplicación móvil, el tamaño máximo de datos cargados que admite es de 100MB.

Como tal mediante el siguiente cuadro de resumen comparativo se puede notar las siguientes diferencias que presentaron mediante los criterios de evaluación:

	Tableau	Power Bi	IBM Cognos
Conversión datos	Análisis completo	Análisis completo	Análisis completo
Tamaño conjunto de datos	Admite un máximo de 300MB	Admite un máximo de 1GB	Admite un máximo de 100MB
Interpretación de datos	Permite limpiar los datos una vez cargados	Presenta alerta de presencia de títulos, o notas	Presenta alerta de presencia de títulos, o notas
Actualizaciones	Tableau debe instalar las actualizaciones en cuanto exista una nueva	Power BI tiene una actualización automática no se requiere de nuevas instalaciones	IBM contiene las versiones más recientes cuando se usa por su portal web.
Tipo de herramienta	Desktop	Desktop	Portal web

Tabla 6: Resumen comparativo de elección de herramientas

5.3 REQUERIMIENTOS DE HARDWARE

En el análisis de la implementación y facilidad de uso se eligieron 3 herramientas para realizar pruebas comparativas las cuales son Microsoft Power BI una de las herramientas más utilizadas y completas que existen en el mercado. La segunda herramienta es Tableau igual una herramienta muy utilizada casi a la par de Microsoft sin embargo por la semejanza que muestran estas dos herramientas se decidió utilizar una tercera herramienta IBM Cognos la cual a diferencia de sus competidoras no se usa una versión de escritorio sino mediante su portal web.

De tal manera que se consideró las siguientes especificaciones técnicas en cuanto a requerimientos físicos y lógicos de los cuales se menciona:

WINDOWS	MacOS
Compatible con Microsoft 7 o versiones recientes de 64bits	Compatible con equipos iMac/MacBook2009 o más recientes
Microsoft Server 2008 R2 o una versión más reciente	OSX 10.11 o una versión más reciente
Procesador Intel Pentium 4 AMD Opteron o una versión más reciente	Espacio libre mínimo en disco de 1,5GB
Memoria de 2GB	Resolución de la pantalla de 1366x768 o superior
Espacio libre mínimo en disco de 1,5GB	
Resolución de la pantalla de 1366x768 o superior	

Tabla 7: Requerimientos técnicos de hardware para Tableau Desktop

WINDOWS
Windows 7 y Windows server 2008 R2 o posterior
Memoria de 1GB
Espacio libre mínimo en disco de 1,5GB
Resolución de pantalla de 1440x900 o 1600x900
Microsoft Excel versión 2010 o superior
No compatible con Mac

Tabla 8: Requerimientos técnicos de hardware para Power BI Desktop

Sistema operativo	Windows, Unix y Linux
RAM	Mínimo de 10GB
Espacio libre en disco	Mínimo de 3,5 GB para instalación y 5GB de espacio libre en la unidad para los componentes de IBM Cognos.
Java Runtime Environment (JRE)	Se proporciona un IBM JRE como parte de la instalación con IBM Cognos Analytics.
Base de datos	Oracle, IBM Db2, Microsoft SQL Server e Informix.

Tabla 9: Requerimientos técnicos de hardware para IBM Cognos Desktop

5.3.1 Pruebas a las herramientas

Luego de la selección de las 3 herramientas que presentan las mejores características en base a la investigación y el cuadro mágico de Gartner, se establecerá una serie de pruebas que corroboren estas afirmaciones, así como obtener las impresiones sobre su uso.

Como parte del proceso de análisis de datos, la calidad de datos recopilados de distintas fuentes y la transformación de estas son punto clave en las pruebas comparativas con las herramientas ya que el objetivo de las pruebas es establecer cuál es la mejor opción al momento de implementar BI en una institución educativa.

5.4 PLANTEAMIENTO DE PRUEBAS

Se plantea como una necesidad estratégica el análisis de la falta de carreras ofertadas en las universidades del Ecuador, para ello se pretende elaborar un dashboard que refleje que carreras son las más acogidas por los estudiantes mediante sus calificaciones. Y un segundo dashboard que permita identificar las carreras con menor acogida.

La primera sección de pruebas corresponde a la realización de 2 dashboard con las siguientes características:

Característica	Fuente de datos	Evaluación
Grafica calificaciones promedio por sexo, carrea, profesor, dependiendo de cada ciclo académico y sede.	Interna	Facilidad de relacionar tablas.
Gráfico de bigotes para analizar la distribución de calificaciones.	Interna	Atractivo visual que presenta la herramienta.
Gráfico lineal comparativa de cada carrera dependiendo de la sede.	Interna	Capacidad de interpretación de datos por parte de la herramienta.
Gráfico comparativo de los estudiantes dependiendo de la calificación obtenida.	Interna	
Gráfico de las ofertas académica de universidades del Ecuador	Externa	Manejo de archivos.
Gráfico de las ofertas académica de universidades de la UPS	Interno	

Gráfico de las becas por área.	Externa	
--------------------------------	---------	--

Tabla 10: Pruebas de interfaz de usuario

La segunda sección de las pruebas corresponde a la parte técnica por lo que se tomara en cuenta los tiempos empleados:

Actividad	Evaluación
Carga de archivos	Tiempo de carga.
Cruce de datos entre 3 tablas.	Tiempo de cruce
Filtro de datos con una gráfica de barras.	Tiempo de filtrado
Tiempo de generación del dashboard con graficas del ejercicio anterior.	Tiempo de generación

Tabla 11: Pruebas de rendimiento

La comparativa técnica utilizara los mismos archivos para una comparativa correcta entre las herramientas seleccionadas. Los datos obtenidos se mostrarán mediante una tabla comparativa.

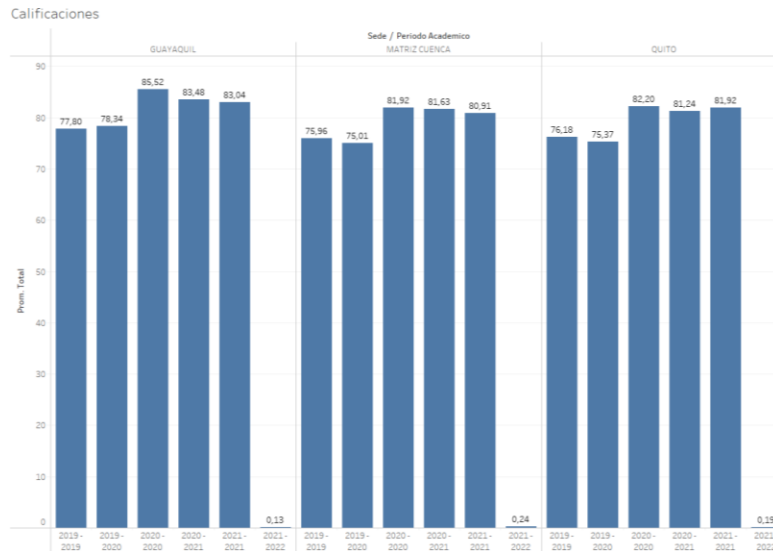


Figura 10: Filtración de calificaciones del periodo 2019 al 2021 de las tres sedes de la Universidad Politécnica Salesiana (Fuente Propia)

2. Se emplearán las consultas de las ofertas académicas para mostrar cual permite una mejor interpretación de los datos, y su manejo de varias fuentes de datos con el objetivo de determinar carreras que presenten menor acogida.

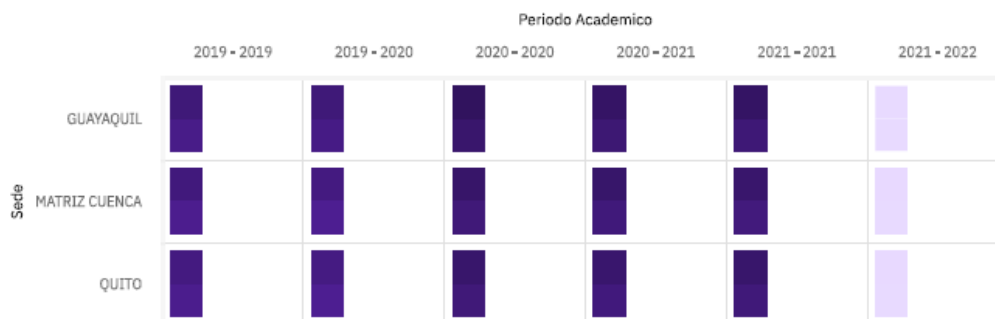


Figura 11: Filtración de carreras del periodo 2019 al 2021 con menor acogida (Fuente Propia)

Para el proceso de análisis de datos y la elaboración de los dashboards como resultado final, es necesario describir el procedimiento y pasos seguir durante el proceso empleado se podrá entender más acerca del enfoque del almacenamiento data mart, los pasos son:

1. Carga de datos del archivo de Excel previamente evaluados: la carga se realizó en las 3 herramientas, Tableau, Power BI, IBM Cognos.
2. Relación entre campos de las tablas.
3. Generación de graficas independientes en función de los requerimientos del dashboard.

4. Generación del dashboard mediante la selección de las mejores graficas de cada aplicación.

El dashboard al ser solo una representación gráfica interactiva, se eligió por su fácil interpretación, ya que se puede dar a conocer mejor los resultados de las carreras con mejor valoración y mejor aceptadas, por ello no fue necesario realizar reportes como parte de los resultados de visualización ya que no son fáciles de comprender y el resumen es más extenso resultando no agradable para el usuario final.

6. RESULTADOS

Mediante el análisis realizado se han cumplido los objetivos planteados, gracias a la implementación de M3S se ha desarrollado un dashboard que permitió aprovechar todas las utilidades de la herramienta para obtener una comparativa clara.

A medida que realizábamos los ejercicios nos percatamos que las herramientas tenían limitaciones que impedían un correcto funcionamiento en entornos reales.

6.1 EVALUACIÓN DE LOS DATOS

Los datos obtenidos de fuentes internas y externas se encuentran en archivos .xlsx los cuales deberán ser cargados a la plataforma.

La evaluación de los archivos se realizó de manera manual buscando inconsistencias y campos vacíos y comprobar que no existan archivos corruptos, que impidan la correcta carga de los datos.

Se resalta que la herramienta Tableau cuenta con un limpiado de interprete de datos, esto quiere decir que el archivo puede contener títulos, encabezados, notas o comentarios, razón por la cual no se podrá interpretar los datos correctamente, por ello Tableau permite realizar el limpiado tal como se muestra en las siguientes figuras:

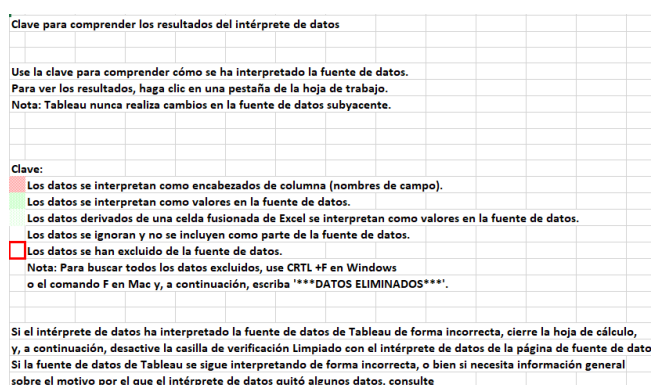


Figura 12: Resultados de la limpieza automática de Tableau (Fuente: Propia)

Los datos proporcionados de forma interna no tuvieron inconvenientes sin embargo los datos de fuentes externas, en este caso de la página de la Senescyt cuentan con menús y descripciones de la información por lo que los datos útiles se tuvieron mover a archivos independientes.

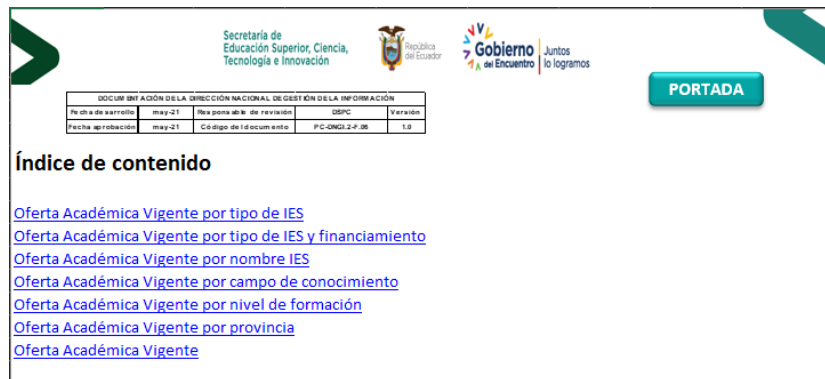


Figura 13: Portada de los datos de fuentes externas no estructuradas (Fuente: Senecyt) <https://educacion.gob.ec/datos-abiertos/>

NOMBRE IES	TIPO INSTITUCIÓN	TIPO FINANCIAMIENTO	NOMBRE CURSO	CAMPO AMPLIO	NIVEL FORMACIÓN	TIPO FORMACIÓN	PROVINCIA	TOTAL OFERTA
CONSERVATORIO SUPERIOR DE MÚSICA JAIMÉ MOLA	INSTITUTO	PARTICULAR AUTOFINANCIADA	EJECUCION DE INSTRUMENTO MUSICAL CON NIVEL EQUIVALENTE A TECNOLOGIA SUPERIOR	ARTES Y HUMANIDADES	TERCER NIVEL TECNOLÓGICO SUPERIOR	TERCER NIVEL TECNOLÓGICO SUPERIOR	PICHINCHA	3

Figura 14: menú de los datos de fuentes externas no estructuradas (Fuente: Senecyt) <https://educacion.gob.ec/datos-abiertos/>

6.2 ELECCIÓN DE LA MEJOR HERRAMIENTA DE BI

Luego de la evaluación de las diferentes opciones que existen en el mercado se seleccionó 3 herramientas:

- Tableau: se encuentran entre las herramientas con una mayor visión de acuerdo con el cuadro mágico de Gartner.
- Power BI: al ser desarrollada por Microsoft es compatible con todo su catálogo de programas de ofimática y se encuentra entre los más recomendados.
- IBM Cognos: cuenta con un entorno en la nube que permite la integración con otros programas que van desde el análisis de datos hasta inteligencia artificial

6.3 ESTABLECER LA INFRAESTRUCTURA

En base a los requerimientos establecidos por las aplicaciones para su funcionamiento y la información obtenida para el análisis, se estableció que la siguiente laptop cumple con los requerimientos para dar los resultados deseados sin ocasionar un cuello de botella en el rendimiento.

Para ello se da a conocer las especificaciones del equipo del cual se utilizó para la realización de las pruebas:

Hardware utilizado durante la prueba

Laptop:	Asus Tuf Dash 2022
Procesador:	Intel i7-12850H
Ram:	16GB
Almacenamiento:	SSD 500GB
Sistema operativo:	Windows 11 21h2

6.4 ANÁLISIS COMPARATIVO

El análisis comparativo se realizó en base al diseño de pruebas establecido en el capítulo anterior, con la finalidad de determinar los tiempos de carga, facilidad de uso de la herramienta y presentación de dashboard.

6.4.1 Resultados de Facilidad de uso de la herramienta.

Las 3 herramientas cuentan con características similares, sin embargo, la interfaz de usuario tiene diferencias clave para su facilidad de uso. Además de contar con licencias de prueba diferentes para cada aplicación.

Tableau cuenta con una interfaz sencilla, la cual se diferencia por tener una página para cada una de las gráficas las cuales se podrán unir en una sección dedicada al dashboard, contando con un total de 15 graficas de cuales elegir.

Power BI cuenta con una interfaz familiar para los usuarios que han utilizado Word o Excel, la cual mantiene una interfaz continuista. A diferencia de la anterior herramienta, Power BI se divide en páginas donde se podrán insertar varias gráficas y no tiene una sección separada para el dashboard. Cuenta con 37 opciones de gráficas, y, además, con la ventaja frente a sus competidoras de poder descargar más graficas de su tienda en caso de requerirlas.

IBM Cognos cuenta con la interfaz más complicada de utilizar entre las 3 herramientas, ya que tiene una interfaz basada en módulos, los cuales se deberán acceder de forma separada, aumentando el tiempo que se necesita para encontrar funciones como el dashboard o el tener que guardar una relación para poder utilizarla.

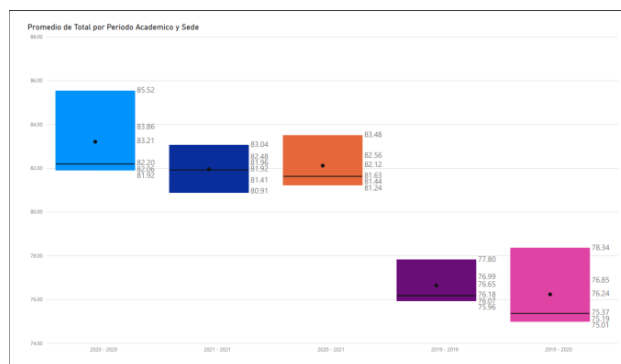


Figura 15: Figura Bigotes de power BI (Fuente: propia)



Figura 16: Figura de Bigotes IBM Cognos (Fuente: propia)

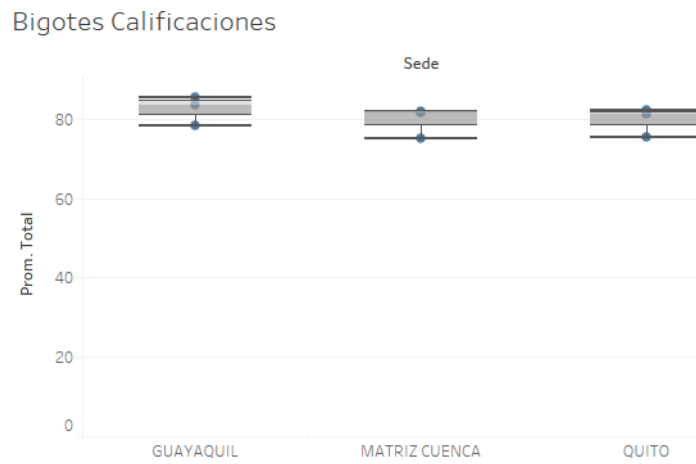


Figura 17: Gráfico de Bigotes Tableau (Fuente: propia)

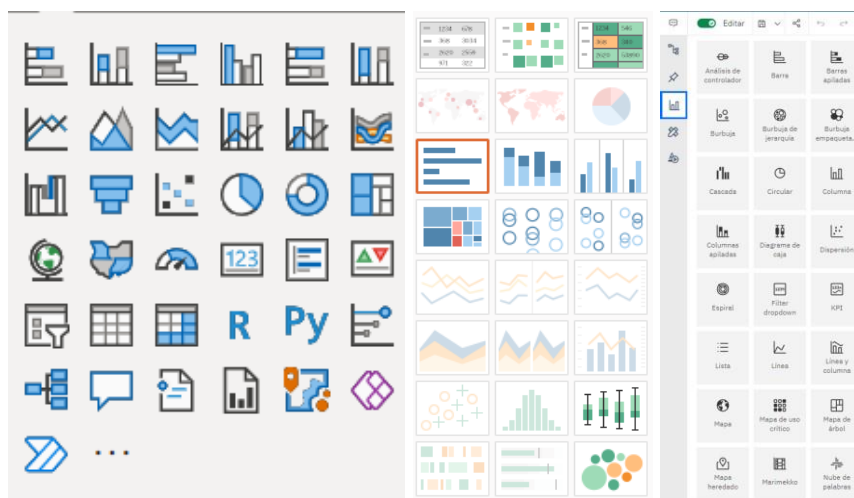


Figura 18: Paneles de graficas Power BI/Tableau/ IBM Cognos (Fuente: propia)



Figura 19: Relación de tablas (*Fuente: propia*)

The 'Editar relación' dialog box is open, showing the relationship between 'Calificaciones' and 'Estudiante'. It displays the primary key 'CedulaEstud_ok' in the 'Estudiante' table and the foreign key 'CedulaEstud_ok' in the 'Calificaciones' table. The cardinality is set to 'Varios a uno (1:1)' and the cross-filtering direction is 'Única'. There are checkboxes for 'Activar esta relación' (checked) and 'Aplicar filtro de seguridad en ambas direcciones'. Buttons for 'Aceptar' and 'Cancelar' are at the bottom right.

Figura 20: Cruce de tablas (*Fuente: propia*)

The 'Administrar relaciones' dialog box is open, showing a table with columns 'Activo', 'Desde: tabla (columna)', and 'A: tabla (columna)'. The first row is checked and shows 'Calificaciones (CedulaEstud_ok)' and 'Estudiante (Cedula_ok)'. At the bottom, there are buttons for 'Nuevo...', 'Detección automática...', 'Editar...', 'Eliminar', and 'Cerrar'.

Activo	Desde: tabla (columna)	A: tabla (columna)
<input checked="" type="checkbox"/>	Calificaciones (CedulaEstud_ok)	Estudiante (Cedula_ok)

Figura 21: Administrar relaciones (*Fuente: propia*)

Mediante el siguiente cuadro de resumen de resultados de facilidad de uso de las herramientas destacamos los siguientes aspectos:

	Tableau	Power BI	IBM Cognos
Interfaz	Intuitiva, su panel contiene una cantidad limitada de graficas que permite generar	Intuitiva su panel contiene una cantidad ilimitada de graficas que permite generar	Para la generación de graficas se requiere buscar el panel y contiene una cantidad ilimitada.
Gráficas	Graficas con tonos más oscuros	Graficas con tonos más coloridos	Graficas con tonos coloridos
Cruce de tablas	Se genera un reporte en relación con los campos que se desea consultar	Se genera un reporte en relación con los campos que se desea consultar	Se genera un reporte en relación con los campos que se desea consultar

Tabla 12: Cuadro resumen de facilidad de uso

6.4.2 Tiempos de carga.

Para la comparativa de las diferentes herramientas se analizó los tiempos para diferentes actividades, tomando como referencia un tiempo promedio. Se realizaron pruebas utilizando las mismas fuentes de datos y generando los mismos gráficos y dashboard de las pruebas anteriores.

La diferencia más significativa se da entre una versión desktop y una versión de nube como lo es IBM. Donde los tiempos de carga se incrementan en función del ancho de banda de internet.

		IBM Cognos	Power BI	Tableau
Tiempo de carga	Calificaciones	Tiempo de carga 45 segundos	Tiempo de carga 35 segundos	Tiempo de carga 35 segundos
	Oferta (Senscyl)	Tiempo de carga 45 segundos	Tiempo de carga 30 segundos	Tiempo de carga 30 segundos
Tiempo de carga	Selección de los campos relacionados	Tiempo estimado para relacionar campos es de 3 minutos	Tiempo estimado para relacionar campos es de 3 minutos	Tiempo estimado para relacionar de 4 minutos
	Consulta de datos	Tiempo respuesta de 35 segundos	Tiempo respuesta de 25 segundos	Tiempo de respuesta de 30 segundos
Tiempo de generación de grafico	Gráficos individuales	No genera gráficos individuales	Tiempo respuesta de 2 segundos	Tiempo de respuesta de 3 segundos
Tiempo de generación de dashboard	Dashboard 1	Tiempo de respuesta de 2 segundos por grafico en la misma hoja	Se recrea arrastrando las gráficas individuales toma alrededor de 1 minuto.	Se genera mediante una hoja nueva y se cruzan los datos nuevamente toma alrededor de 3 minutos.
Tiempo de exportación a PDF	Exportación reportes	Tiempo estimado es de 20 segundos	Tiempo estimado es de 20 segundos	Tiempo estimado es de 20 segundos
	Exportación dashboard	Tiempo estimado es de 1 minuto.	Tiempo estimado es de 45 segundos	Tiempo estimado es 1 minuto

Tabla 13: Especificación de los tiempos de carga de datos

En base a los tiempos de carga analizados, tanto por el tamaño del archivo cargado en las herramientas como los resultados obtenidos en base a las comparativas de rendimiento y funcionamiento de las herramientas, es claramente evidente que existen similitudes, sin embargo, cada herramienta tiene sus pro y contras de las que se destacan:

- **En tiempos de carga** IBM Cognos ocupa un mayor índice de tiempo y el problema radica que al utilizar mediante su plataforma web su tiempo de carga sea mayor a las herramientas desktop.
- Así también en cuanto al **cruce de tablas** IBM Cognos y Tableau presenta un mayor índice de tiempo esto porque ambas duplican su tiempo de generación.
- En cuanto al **tiempo de generación de gráficos**, IBM Cognos nuevamente presenta un índice de tiempo más que las otras herramientas y esto es más por que todo se ejecuta en la plataforma web.
- **Tiempo de generación de dashboard** presenta que la herramienta Tableau duplica el tiempo de generación de esto se debe a que la herramienta cuenta con una sección separada para su generación mientras, Power BI e IBM Cognos utilizan la hoja de trabajo como dashboard.
- Finalmente, el **tiempo de exportación a PDF** define que las tres herramientas tienen el mismo índice de tiempo de exportación del documento.

Un dato para destacar es que se utilizaron archivos con un peso menor a 100Mb debido a que este es el límite de carga de IBM Cognos por lo que utilizar archivos más pesados hubiera consistido en una desventaja para las otras herramientas.

Para que sea más claro el resultado de evaluación se dará a conocer mediante un histograma los resultados obtenidos de la comparativa, mediante un rango del 1 al 5, siendo 1 la valoración más baja y 5 la valoración más alta tal como se muestra en la siguiente imagen.

RESULTADOS COMPARATIVOS DE HERRAMIENTAS DE BI

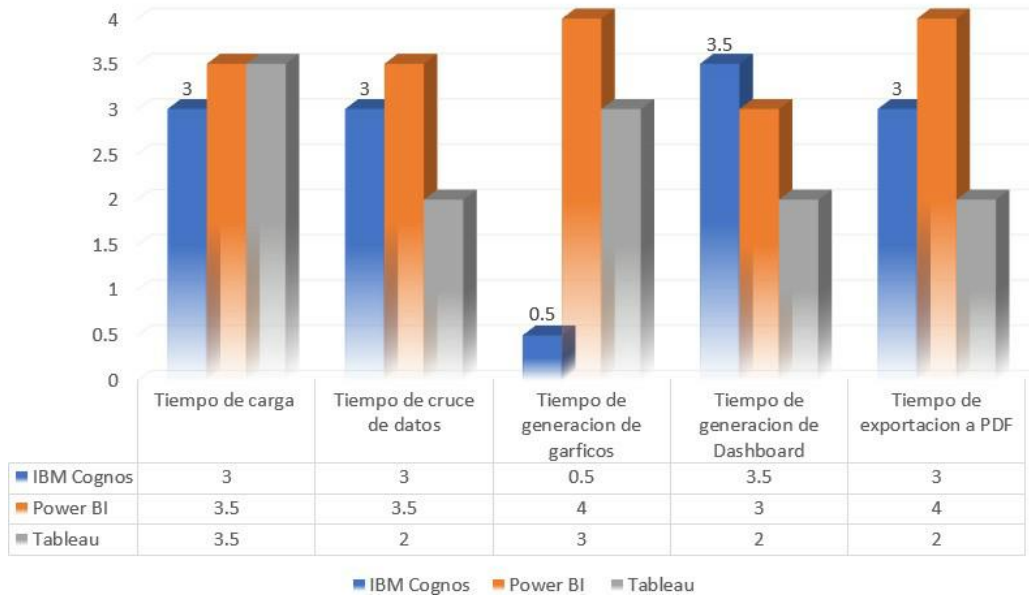


Figura 22: Histograma de Resultados comparativos (*Fuente: propia*)

Como resultado de la comparativa en las tres herramientas se puede decir que según mi experiencia propia, la mejor herramienta es Microsoft Power BI, ya que esta herramienta trabaja con un mejor tiempo de respuesta, y al ser una herramienta escalable, sus procesos se van almacenando según la acción que se realice es decir, en caso de eliminar un cruce de tabla o un dashboard por error no es necesario realizar nuevamente desde el principio, solo se requiere ir al panel de actividades de Power BI y eliminar los procesos siguientes después del que necesitamos.

En la siguiente tabla se presentan las ventajas y desventajas encontradas al momento de utilizar cada una de las herramientas seleccionadas, por ello es necesario recalcar que las herramientas con licenciamiento para esta selección fueron Tableau y Power BI, a diferencia de las dos mencionadas para IBM Cognos se usó una prueba gratis, tal como describe la siguiente tabla de resultados.

Herramientas	Ventajas	Desventajas	Versiones	Pruebas	Experiencia de uso
IBM Cognos	Facilidad de generar dashboard	El límite de carga de datos es de 100MB	no dispone de una versión gratuita	prueba gratis por 30 días.	Fácil e intuitiva al momento de cargar y cruzar datos.
Power BI	Power BI almacena los pasos en su panel de actividades y su límite de carga de datos es de 10GB	No es posible compartir informes con otros usuarios a pesar de que se integra a otras plataformas.	no dispone de una versión gratuita	prueba gratis por 30 días.	Intuitiva, se integra con Office 365 y a su vez tiene mucha similitud en análisis de datos con Tableau.
Tableau	Análisis de datos a profundidad , más allá de lo visual	Una de las opciones más caras ya que su versión gratuita es limitante.	dispone de una versión gratuita	prueba gratis por 30 días.	Intuitiva, fácil de usar, pero con mayor tiempo de respuesta

Tabla 14: Cuadro de resumen de resultados de las herramientas

6.5 PRESENTACIÓN DE DASHBOARD

En base al diseño de pruebas se estableció un dashboard que utiliza los datos de la universidad para la generación de dashboard con el objetivo demostrar la claridad de la presentación y la interpretación de los datos.

6.5.1 Dashboard y reporte de Power BI

Power BI representa los datos mediante un esquema de colores que lo hace muy vistoso como se observa en la figura 17, sin embargo, lo que resalta es su forma de presentar los datos en secciones que se pueden desplazar con la ayuda una barra por cada sección por lo generar se separa en 4 secciones como se observa en la figura 18.

En los siguientes dashboards se da a conocer las siguientes consultas:

Mediante una consulta general se determina el promedio total por cada una de las sedes del periodo académico 2019 – 2022, en el cual determina que en el año 2020 la sede Guayaquil tubo un mayor índice de valoración.

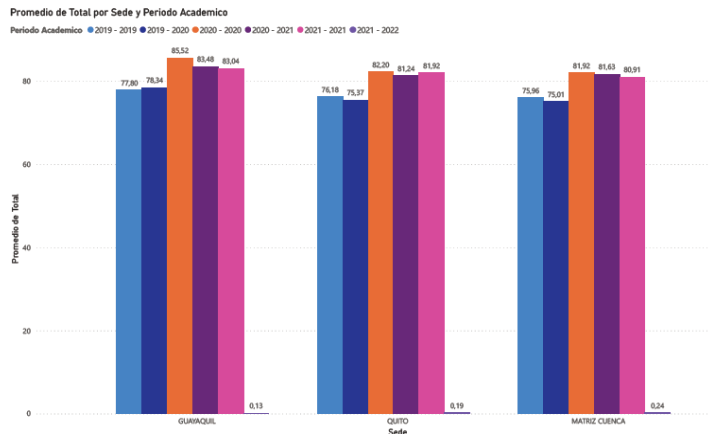


Figura 23: Consulta del promedio total por sede y periodo del 2019 al 2022 (Fuente: propia)

Consulta del promedio total por sexo del estudiante del periodo académico 2019-2021.



Figura 24: Consulta del promedio total por sede y sexo de los estudiantes en el periodo 2019 al 2021 (Fuente: propia)

Consulta del promedio total por número de cedula de estudiantes que cursaron el periodo académico 2019 – 2022, se determina que efectivamente la sede la sede Guayaquil tubo un mayor índice de valoración en el año 2020 como muestra la figura 24 y 25.

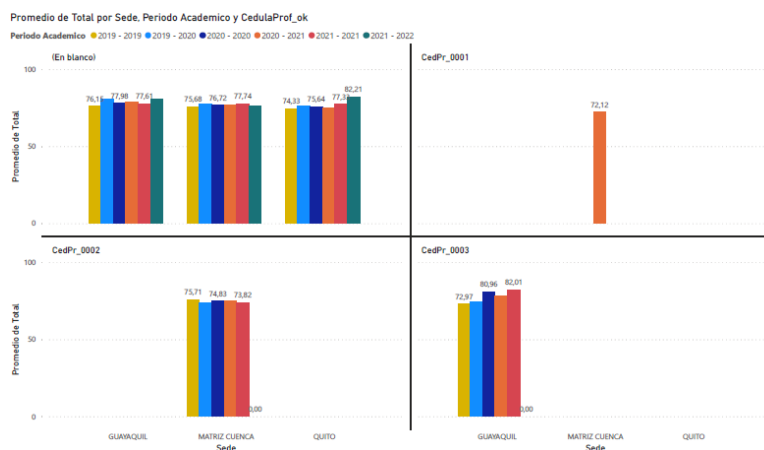


Figura 25: Consulta del número de cedula del promedio total por sede y periodo académico (Fuente: propia)

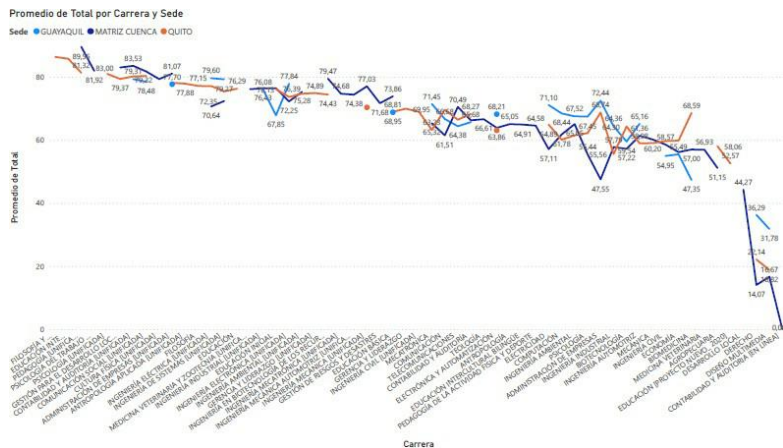


Figura 28: Consulta del promedio total por carrera y sede (Fuente: propia)

Para tener un mayor enfoque se realizó una gráfica que determine el promedio total de valoración por carreras de las tres sedes de la Universidad Politécnica Salesiana.

En base a las consultas anteriores se realizó un dashboard general donde se muestra el promedio total de los estudiantes que cursaron el periodo 2019 – 2022 por carreras, sexo y número de identificación del estudiante de las tres sedes.

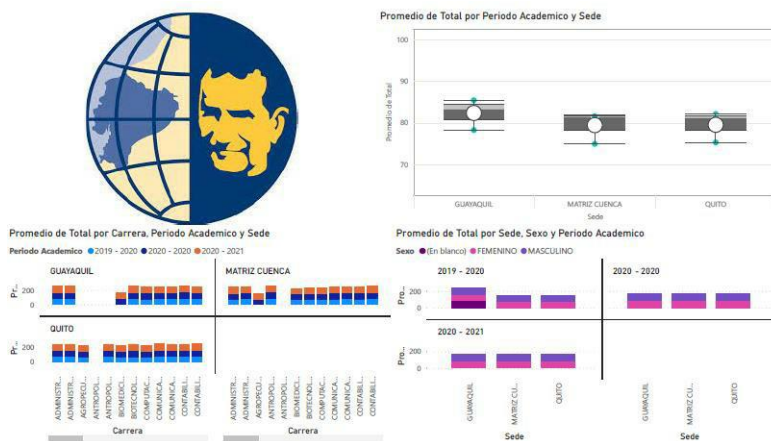


Figura 29: Dashboard, general de las consultas (Fuente: propia)

El siguiente reporte da a conocer que área de estudio tiene mayor aceptación en periodos anteriores 2015 - 2019.

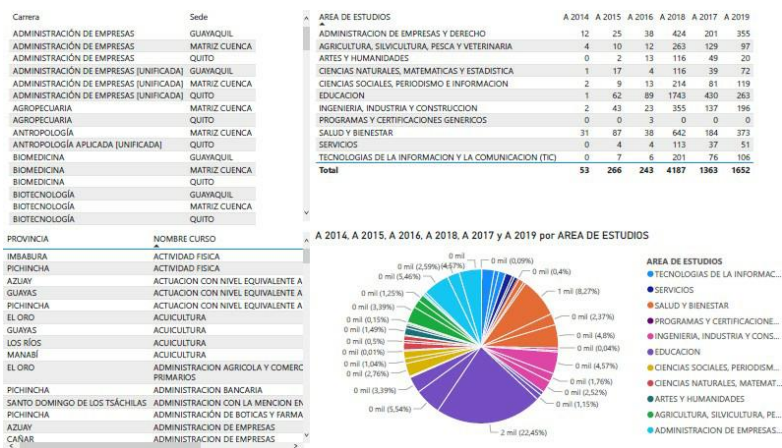


Figura 30: Reporte General de consultas Power BI (Fuente: propia)

6.5.2 Dashboard y reporte de Tableau

Tableau representa los datos de forma predeterminada en tonos azules, sin embargo, se lo puede modificar como se observa en la figura 24 y 25 a diferencia de Power BI, Tableau distingue entre hojas de trabajo y el dashboard. En cada hoja de trabajo solo se podrá ingresar una gráfica. y posteriormente se generará el dashboard, el cual presentará un menú como se observa en la figura 26 permitiendo la selección de las gráficas que deseamos mostrar. El diseño permite mantener un trabajo ordenado.

Para ello se determinó una consulta del promedio total por carreras cursadas en el periodo académico 2019 - 2022 y determinar el área de estudio con mayor valoración y oferta académica.

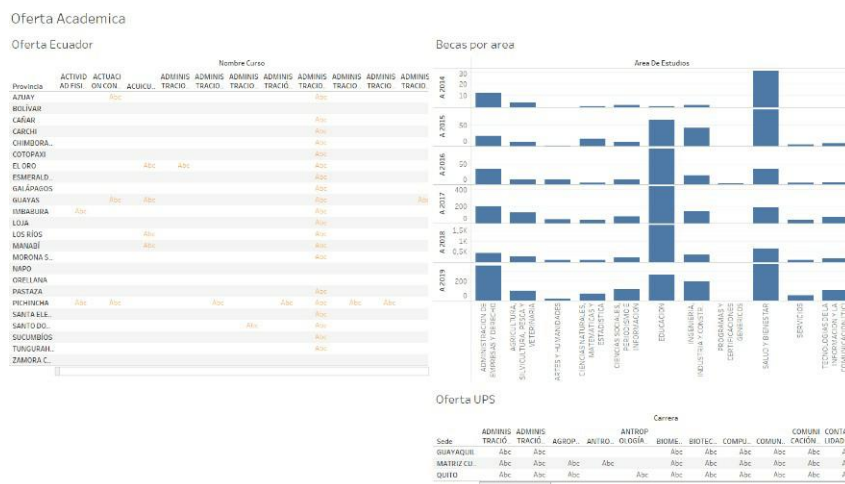


Figura 31: Reporte General de consultas Tableau (Fuente: propia)



Figura 32: Dashboard de consultas de Tableau (Fuente: propia)

La figura 32 da a conocer mediante una consulta realizada al periodo académico 2019 - 2021 acerca del rendimiento de los estudiantes en base a la carrera y sede cursada en el periodo.

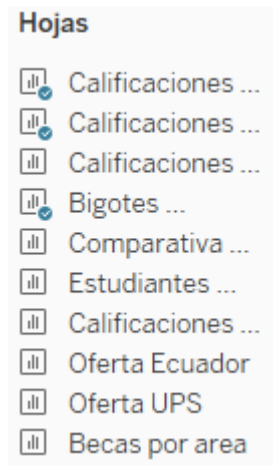


Figura 33: Menú de dashboard (Fuente: propia)

La figura 33 da a conocer aquellas hojas de trabajo donde se destacan las consultas, gráficos, y reporte final.

6.5.3 Dashboard y reporte de IBM Cognos

El dashboard de IBM Cognos es similar al de Power BI ya que se trabaja directamente en el dashboard, sin embargo, su manejo es más tardado por lo que se está aplicando mediante su plataforma web ya que si una relación no fue creada se deberá, salir del dashboard y entrar al módulo de relación, lo que resulta en un trabajo engorroso, con respecto a la presentación de dashboard sus colores se mantienen en derivados del morado. El redimensionar la hoja de trabajo resulta inútil ya que al cerrar y volver a abrir vuelve a sus valores por defecto lo que obstaculiza el tener una buena presentación.

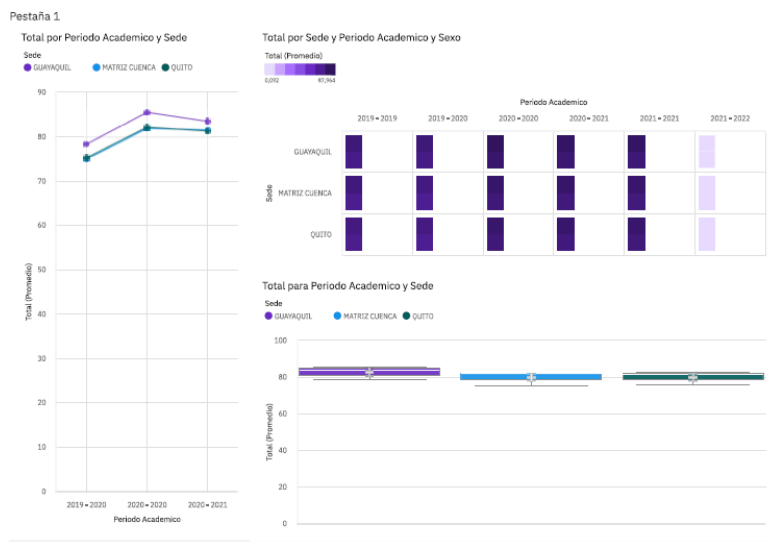


Figura 34: Reporte General de consultas (Fuente: propia)

7. CRONOGRAMA

1 ACTIVIDADES POR DESARROLLAR

1.1 Objetivo 1

- Recopilación de información a evaluar en base a la primera fase de la metodología M3S.
 - Describir e identificar los tipos de herramientas de inteligencia de negocios en el mercado.
 - Revisar artículos y foros relacionados al tema
 - Determinar el tipo de herramienta según un criterio de evaluación
 - Aplicar los principios de la metodología propuesta.

1.2 Objetivo 2

- Determinar los aspectos a evaluar para la elección de herramientas BI.
 - Definir los aspectos necesarios para el criterio de evaluación, en base a la metodología planteada
 - Comparar las herramientas y realizar un descarte de herramientas que no pasen desde su compatibilidad e instalación.
 - Detectar el tipo de herramienta cumple con los requisitos, considerando precio, facilidad de uso y aceptabilidad.
 - Determinar el tipo de datos para las pruebas de elección de las herramientas seleccionadas.

1.3 Objetivo 3

- Análisis comparativo de las herramientas seleccionadas
 - Selección de tres herramientas que cumplan con los requisitos de evaluación
 - Aplicar a cada herramienta el mismo caso de análisis de datos para determinar tiempos de carga y resultados
 - Registrar y capturar los tiempos de respuesta de las herramientas bajo gran cantidad de datos a analizar.

1.4 Objetivo 4

- Obtención de resultados
 - Comparar tiempos de respuesta de las pruebas de las herramientas en base a los datos obtenidos.
 - Aplicar pruebas en base a consultas simples y más complicadas que permitan determinar la respuesta inmediata de las herramientas
 - Determinar que herramienta presenta mayor aceptación en caso de facilidad de uso y presentación de una mejor solución en representar los datos.

2 CRONOGRAMA DE ACTIVIDADES

	Actividad	Duración	Fecha Inicio	Fecha Fin
Objetivo 1	Actividad 1	97 días	19/04/2021	13/11/2021
	Actividad 1.1	30	19/04/2021	21/06/2021
	Actividad 1.2	35	21/06/2021	30/08/2021
	Actividad 1.3	25	30/08/2021	05/10/2021
	Actividad 1.4	7	05/10/2021	13/11/2021
Objetivo 2	Actividad 2	84 días	13/11/2021	07/03/2022
	Actividad 2.1	35	13/11/2021	20/12/2021
	Actividad 2.2	14	20/12/2021	06/01/2022
	Actividad 2.3	22	06/01/2022	22/02/2022
	Actividad 2.4	13	22/02/2022	07/03/2022
Objetivo 3	Actividad 3	57 días	07/03/2022	25/06/2022
	Actividad 3.1	22	07/03/2022	15/04/2022
	Actividad 3.2	17	15/04/2022	21/05/2022
	Actividad 3.3	20	21/05/2022	25/06/2022
Objetivo 4	Actividad 4	55 días	10/06/2022	28/08/2022
	Actividad 4.1	15	10/06/2022	10/07/2022
	Actividad 4.2	14	10/07/2022	30/07/2022
	Actividad 4.3	26	30/07/2022	28/08/2022

Días Laborales	293
Horas Totales	450
Fecha Inicio	19/04/2021
Fecha Fin	28/08/2022

Tabla 15: Cronograma de Actividades

8. PRESUPUESTO

Para establecer el presupuesto se ha considerado el tiempo aplicado en el desarrollo del proyecto técnico, el cual se estima de \$2 por hora de trabajo, y debido a que se trabajó 5 días a la semana en un horario de 2 horas, el tiempo resultante es de 450 horas y un valor en el desarrollo del proyecto de \$310, a continuación, se da a conocer el detalle en la siguiente tabla:

DENOMINACIÓN	CANT.	COSTO UNITARIO	COSTO TOTAL
Ordenador de alta gama con las mejores especificaciones técnicas	1	1400	1400
Tarjeta de video NVidia GTX	1	500	500
Material Bibliográfico y Fotocopias	1	20	20
Varios e Imprevistos	1	30	30
TOTAL	4	1940	1940

Tabla 16: Presupuesto

9. CONCLUSIONES

Mediante las pruebas realizadas de la elección y la comparativa entre herramientas el resultado reveló que las aplicaciones de Tableau y Power BI presentan una ventaja en cuanto a análisis de datos a profundidad y usabilidad, por lo que se puede decir que existe tal coincidencia según la (figura 9) del cuadro mágico de Gartner.

Tableau y Power BI son dos de las herramientas más utilizadas a nivel empresarial, y como se pudo observar en el cuadro mágico de Gartner también presentan una mayor innovación, las herramientas no son muy diferentes, sin embargo, al ser Microsoft la desarrolladora detrás de Power BI la misma cuenta con un catálogo más amplio de objetos visuales.

La comparativa también permitió determinar las ventajas y desventajas de utilizar una herramienta en la nube o en desktop. Siendo la nube una buena opción para las empresas que no cuentan con equipo propio que pueda ejecutar las herramientas, sin embargo, la misma tiene limitantes como el peso de carga de archivos o el rendimiento que se verá limitado por el proveedor.

En conclusión, las herramientas de inteligencia de negocios no tienen una dificultad elevada de implementación, sin embargo, no se podrá aprovechar todo su potencial si no existe un proceso preestablecido para transformar los datos en información. Las herramientas de BI cuentan con pocas diferencias entre sí, pero estas diferencias son claves a la hora de decidir la mejor opción para el negocio, por un lado, se tienen las que presentan una ligera ventaja de rendimiento, mientras Power BI cuenta con un dashboard mucho más vistoso y IBM Cognos presenta la ventaja de la nube.

Los criterios de evaluación y selección aplicadas mediante la metodología M3S permitió como resultado descartar herramientas que no cumplían las expectativas, por ello las herramientas más relevantes permitieron realizar un análisis comparativo en el cual se determinó cual es la mejor herramienta de inteligencia de negocios con mejores prestaciones.

10. RECOMENDACIONES

Una vez concluida la tesis, se recomienda investigar más allá acerca de otros aspectos a considerar acerca de la inteligencia de negocios y sus prestaciones por ello se propone:

- Extender la investigación realizada en este proyecto técnico con la finalidad de determinar que tipo de herramienta es aplicable según el tipo de negocio o empresa.
- Trabajar en mejoramiento de las técnicas de evaluación de herramientas BI para identificar y determinar los tipos de herramientas que se pretende utilizar.
- Realizar con mayor detenimiento pruebas comparativas con casos más realistas con datos más complejos como el sector financiero cuyo comportamiento es más complejo.

REFERENCIAS

- [1] B. Rababah, S. Zhou y M. Bader, «Evaluation the Performance of DMZ, » *Modern Education and Computer Science*, vol. 8, n° 1, p. 13, 2018.
- [2] M. Keri, «DICOM Honeypot, » *Github*, 22 Octubre 2020. [En línea]. Available: <https://github.com/nsmfoo/dicompot>. [Último acceso: 19 agosto 2021].
- [3] Benítez, Y. L. (2019). *Business Intelligence. ADGG102PO*
- [4] Valderrama Triviño, F. A., & Garcés Bohada, A. S. (2018). *Diseño e implementación de un datamart para las no-tas históricas de los estudiantes de la Universidad Distrital Francisco José Caldas. (Proyecto de grado para optar al Título de Ingeniero de Sistemas). Universidad Distrital Francisco José Caldas.*
- [5] Perugachi, M. (2019). *Aplicación de la metodológica M3Sde inteligencia de negocios en un sistema pecuario en el manejo reproductivo y productivo de haciendas lecheras.*
- [6] A. A. Gómez, «Análisis de vulnerabilidades en IoT para el despliegue de Honeypots,» *Lálaga*, 2018
- [7] Santiago Morales, G. T. (2020). *metodología m3s para proyectos de inteligencia de negocios.*
- [8] Aguilar, L. J. (2019). *Inteligencia de negocios y analítica de datos: Una visión global de Business Intelligence & Analytics. Alpha.*
- [9] Aguilar, L. J. (2020). *Inteligencia de negocios y analítica de datos. Marcombo.*
- [10] Jones, H. (2019). *Analítica de datos: La guía definitiva de análisis de Big Data para empresas, técnicas de minería de datos, recopilación de datos y conceptos de inteligencia empresarial. Bravex Publications.*
- [11] Jones, H. (2019). *Ciencia de Datos para Empresas: Modelo Predictivo, Minería de Datos, Análisis de Datos, Análisis de Regresión, Consulta de Bases de Datos y Aprendizaje Automático para Principiantes (Spanish Edition). Bravex Publications.*
- [12] Ramos, S. (2016). *Data Warehouse, Data Marts y Modelos Dimensionales un pilar fundamental en la toma de decisiones. SolidQ Press.*
- [13] Benítez, Y. L. (2019). *Business Intelligence ADGG102PO. IC Editorial.*
- [14] Cordero-Naspud, E. I. (10 de Julio-Diciembre de 2020). *Soluciones corporativas de BI. Soluciones corporativas de BI. Venezuela: ISSN: 2542-3088.*

- [15] Cordero Avila, P. A. (2019). *Propuesta metodologica para procesar y analizar informacion de costos y cantidades de obra de proyectos de construccion (B.S. thesis)*. Universidad del Azuay.
- [16] Cuarto Diaz, J. (2017). *Introduccion al business intelligence. Introduccion al business intelligence*,1–315.
- [17] Supply Chain, R. (2021, Enero). *La tecnologia OLTP: que es y que puede hacer por ti*. <https://retos-operaciones-logistica.eae.es/la-tecnologia-oltp-que-es-y-que-puede-hacer-por-ti/>.
- [18] Jangla, K. (2018). *Accelerating development velocity using docker: Docker across microservices*. Apress.
- [19] Fuerza comercial consultoría (2020, febrero) *Fuentes de información para una buena gestión comercial*<https://www.fuerzacomercial.es/fuentes-info-gestion>
comercial/#:~:text=Como%20fuentes%20internas%20se%20conocen,primarias%20y%20fuentes%20internas%20secundarias.
- [20] Workbench, P. S. (2021, October). *Mondrian Schema Workbench*. https://mondrian.pentaho.com/documentation/schema_workbench.pdf.
- [21] Rojas Munoz, C., Saquicela Galarza, V., y cols. (2017). *Hacia un sistema de ayuda a la decision para universidades: caso de uso de la universidad de cuenca*.
- [22] Gartner. (2020). *Los 9 errores a evitar en la transformaci3n digital empresarial*. <https://www.gartner.es/es/articulos/errores-a-evitar-en-la-transformacion-digital-empresarial>.
- [23] Duque, A. B., Romero, R. D. E., Rivillas, C. A. R., y Orejuela, A. R. (s.f.). *Investigaci3n en administraci3n y su impacto en comunidades academicas internacionales*.
- [24] Workbench, P. S. (2021, October). *Mondrian Schema Workbench*. https://mondrian.pentaho.com/documentation/schema_workbench.pdf.
- [25] Nieves Lahaba, Y., y Leon Santos, M. (2001). *La gestion del conocimiento: una nueva perspectiva en la gerencia de las organizaciones*. *Acimed*, 9(2), 121–126.
- [26] SAP /<https://www.sap.com/latinamerica/insights/what-is-a-data-warehouse.html>

Anexos

Anexo 1 Manual de instalación de herramientas