



**INSTITUTO SUPERIOR
TECNOLÓGICO QUITO**

Formamos tu PROPÓSITO DE VIDA



GUÍA GENERAL DE INTELIGENCIA DE NEGOCIOS

AUTOR: MÓNICA RAMÍREZ

ISBN: 978-9942-562-17-3



9 789942 562173

ITQ
WWW.ITQ.EDU.EC
INVESTIGACIÓN





GUÍA GENERAL DE INTELIGENCIA DE NEGOCIOS

AUTOR: MÓNICA ALEXANDRA RAMÍREZ VELASTEGUI

EDICIÓN: PRIMERA EDICIÓN

AÑO: 2025

TRABAJO EN EDICIÓN:



EQUIPO EDITORIAL:

Camila Díaz Díaz

Keyerman Toapanta Cisneros

Este material está protegido por derechos de autor. Queda estrictamente prohibida la reproducción total o parcial de esta obra en cualquier medio sin la autorización escrita de los autores y el equipo editorial. El incumplimiento de esta prohibición puede conllevar sanciones establecidas en las leyes de Ecuador.

Todos los derechos están reservados.

ISBN: 978-9942-562-17-3

ISBN: 978-9942-562-17-3



SOBRE EL AUTOR



Mónica Ramírez es una profesional con más de 10 años de experiencia en el apasionante mundo de las Tecnologías de la Información (TI). A lo largo de su carrera, se ha destacado por liderar proyectos innovadores y fomentar una comunicación efectiva en equipos multidisciplinarios. Su amor por la educación la llevó a obtener una Maestría en Gerencia y Liderazgo Educativo, y actualmente se encuentra cursando una Maestría en Inteligencia Artificial Aplicada, lo que refleja su compromiso con el aprendizaje continuo.

Con más de 8 años de experiencia como docente universitaria, Mónica ha dejado una huella significativa en el desarrollo académico y personal de sus estudiantes. Su enfoque pedagógico busca formar profesionales íntegros, combinando conocimientos técnicos con habilidades interpersonales esenciales para el ámbito laboral.

La sólida trayectoria de Mónica en TI, junto con su interés en áreas emergentes como la Inteligencia Artificial, le permite ofrecer una educación actualizada y relevante. Ella cree firmemente que la educación tiene el poder de transformar vidas, preparándolos para enfrentar los desafíos del futuro con confianza, conocimiento y creatividad.

Su carrera comenzó como Ingeniera de Software, donde adquirió habilidades técnicas esenciales. Luego, se desempeñó como Analista de Software y Administradora de Base de Datos, así como Analista de Ciencia de Datos, aplicando técnicas de análisis para extraer información valiosa que respalda la toma de decisiones estratégicas. Su capacidad de liderazgo la llevó a convertirse en Gerente de Desarrollo de Software, donde impulsó proyectos innovadores y mejoró la eficiencia de los equipos. Además, ha realizado importantes contribuciones en el ámbito educativo, trabajando como Analista Técnica en evaluación educativa a nivel nacional y como Docente Investigador.

Gracias a su experiencia en el sector público y privado, así como a sus valiosas contribuciones a la educación y su liderazgo académico, Mónica se ha consolidado como una referente en el ámbito de las Tecnologías de la Información y la Educación. Su compromiso con la excelencia y su anhelo de influir positivamente en las futuras generaciones de profesionales en tecnología son palpables en su trayectoria y en los logros que ha conseguido a lo largo de su carrera



TABLA DE CONTENIDO

GUÍA GENERAL DE INTELIGENCIA DE NEGOCIOS.....	7
1. DESCRIPCIÓN DE LA ASIGNATURA	7
2. BIBLIOGRAFÍA.....	7
2.1. Básica.....	7
2.2. Complementaria.....	8
3. COMPETENCIAS GENÉRICAS Y ESPECÍFICAS.....	9
4. OBJETIVO GENERAL.....	10
5. FORMACIÓN CIUDADANA, VALORES Y HABILIDADES BLANDAS.....	11
6. NORMAS DE CLASE.....	11
7. SISTEMA DE EVALUACIÓN.....	12
8. UNIDADES.....	12
UNIDAD 1 (INTRODUCCIÓN A LA INTELIGENCIA DE NEGOCIOS).....	13
Temas y Subtemas.....	13
Conceptos básicos de BI.....	13
Definición de BI.....	14
Importancia de BI.....	15
Aplicaciones de BI.....	15
Arquitectura o componentes de BI.....	16
Data Warehousing.....	18
Modelos Multidimensionales.....	19
Modelo Estrella.....	20
Modelo Copo de Nieve.....	21
Bases de datos y Data Warehouse.....	22
Diferencias Clave entre Bases de Datos y Data Warehouses.....	23
Transformación de una Base de Datos Transaccional a un Data Warehouse.....	23
Autoevaluación 1.....	25
Resumen de la Unidad 1.....	28
UNIDAD 2 (HERRAMIENTAS Y TÉCNICAS DE TRANSFORMACIÓN).....	29
Temas y Subtemas.....	29
ETLs.....	29
Componentes del ETL:.....	30
Herramientas utilizadas en los procesos ETL.....	30
Instalación de Pentaho Data Integration.....	30



Requisitos Previos	31
Pasos de Instalación	32
Configuración de PDI.....	32
Familiarizándose con Pentaho Data Integration.....	32
Creación de transformaciones.....	33
El Componente Input en Pentaho Data Integration	35
Tipos de input en Pentaho:	36
El Componente Transform en Pentaho Data Integration	37
Tipos de transform en Pentaho:.....	38
El Componente Utility en Pentaho Data Integration	39
Tipos de Utility en Pentaho:.....	40
El Componente Joins en Pentaho Data Integration	41
El Componente Output en Pentaho Data Integration	42
Tipos de output en Pentaho:.....	43
Otros Componentes en Pentaho Data Integration	44
Diseño de Transformación.....	46
Ejecución de la transformación	48
Autoevaluación 2.....	50
Resumen de la Unidad 2.....	53
UNIDAD 3. (DASHBOARD PARA LA TOMA DE DECISIONES).....	54
Temas y Subtemas.....	54
Reportes y Visualización de datos	54
Importancia de reportes y Visualización de datos	55
Herramientas y Técnicas de Visualización.....	55
Power BI.....	56
Principios Esenciales de Power BI	56
Característica Principales de Power BI.....	56
Bloques de Power BI	57
Diseño y Desarrollo de Dashboard	64
Uso de DAX para Cálculos y Medidas	69
Introducción a DAX.....	70
Columnas Calculadas.....	70
Medidas.....	70
Funciones DAX.....	70
Publicación y Compartición de Dashboards	72
Exportar Dashboards.....	72
Publicación de Dashboards en Power BI Service	73
Configuración de Permisos y Seguridad.....	74

Compartición de Dashboards con Usuarios Internos.....	74
Integración de Dashboards en Aplicaciones y Sitios Web	74
Uso de Aplicaciones Móviles de Power BI.....	75
Autoevaluación 3.....	76
Resumen de la Unidad 3.....	79
9. REFERENCIAS.....	80
10. SOLUCIONARIO	81

ÍNDICE DE FIGURAS

Figura 1 Componentes de BI.....	17
Figura 2 Modelo Estrella.....	21
Figura 3 Modelo Copo de nieve.....	22
Figura 4 Pantalla principal de Pentaho-Spoon.....	33
Figura 5 Menú de File de Pentaho.....	34
Figura 6 Panel Design.....	34
Figura 7 Ventana de Pentaho del componente Input	35
Figura 8 Ventana de Pentaho del componente Transform	38
Figura 9 Ventana de Pentaho del componente Utility	40
Figura 10 Ventana de Pentaho del componente Joins	41
Figura 11 Ventana de Pentaho del componente Output	43
Figura 12 Ventana de Pentaho del componente varios	45
Figura 13 Componentes de una transformación de datos	47
Figura 14 Hoja de trabajo de la transformación	48
Figura 15 Ventana de ejecución del Job	49
Figura 16 Bloques de Power BI	57
Figura 17 Gráfico de barras.....	60
Figura 18 Gráfico de líneas	60
Figura 19 Gráfico de dispersión	61
Figura 20 Gráfico de pastel y anillo	61
Figura 21 Mapa de calor o matriz de correlación	62
Figura 22 Gráfico tipo mapa	62
Figura 23 Tarjetas en dashboards en Power BI	63
Figura 24 Pantalla inicial de conectar a fuente de datos.....	65
Figura 25 Pantalla de Power Query	65



Figura 26 Ventana de visualizaciones	66
Figura 27 Ventana de Datos.....	67
Figura 28 Dashboard personalizado	68
Figura 29 Ventana de Panel de formato.....	68
Figura 30 Ventana de Etiquetas de datos.....	69
Figura 31 Pantalla de modelado de datos	71
Figura 32 Pantalla para exportar Dashboards	73

ÍNDICE DE TABLAS

Tabla 1: Requerimientos de software para instalar Pentaho	31
Tabla 2: Visualización tipo Tabla Pivot.....	63



GUÍA GENERAL DE INTELIGENCIA DE NEGOCIOS

1. DESCRIPCIÓN DE LA ASIGNATURA

La asignatura de Inteligencia de Negocios es esencial para formar a los tecnólogos en Desarrollo de software, dotándolos de habilidades para convertir datos en insights estratégicos que facilitan la toma de decisiones empresariales. A través del estudio de técnicas y herramientas avanzadas, los estudiantes aprenderán a recopilar, analizar y presentar datos, utilizando sistemas de reportes, análisis en línea, inteligencia artificial, modelado de negocio y visualización de datos. Este conocimiento es crucial en un contexto marcado por el big data, el aprendizaje automático y la analítica avanzada, donde las tensiones sobre privacidad y seguridad de datos presentan desafíos éticos y técnicos significativos.

La inteligencia de negocios, al proporcionar soporte para decisiones estratégicas mediante el análisis de datos, es fundamental para mejorar la eficiencia y competitividad organizacional. La asignatura prepara a los estudiantes para diseñar e implementar sistemas de inteligencia de negocios, haciéndolos capaces de enfrentar problemas complejos y ofrecer soluciones innovadoras.

Integrando la inteligencia de negocios en el perfil de egreso, esta asignatura enriquece el desarrollo profesional de los tecnólogos en desarrollo de software, ampliando sus oportunidades laborales en diversos sectores y proporcionando una base sólida para especializaciones futuras en TI, como análisis y ciencia de datos. Así, se convierte en un componente clave para la formación de futuros líderes tecnológicos, equipados con un conjunto de habilidades altamente demandadas en el mercado laboral.

2. BIBLIOGRAFÍA

2.1. Básica

- Sánchez, R. (2009). Business intelligence (BI) - To BI or not to BI.: Editorial El Cid Editor | apuntes.

Este texto explora la importancia y el impacto de la inteligencia de negocios en las organizaciones modernas. Se describe cómo BI puede transformar datos en información valiosa, mejorando la toma de decisiones y la competitividad empresarial. El libro abarca



desde los conceptos básicos hasta las aplicaciones avanzadas de BI, incluyendo técnicas de análisis de datos, herramientas de software y casos prácticos. Además, se discuten los desafíos éticos y técnicos relacionados con la privacidad y la seguridad de los datos, proporcionando una visión integral y práctica del uso de BI en el entorno empresarial actual.

- Curto, J. (2016). *Introducción al business intelligence*. Editorial UOC

Este texto ofrece una visión detallada sobre los fundamentos y aplicaciones de la inteligencia de negocios, se aborda cómo BI permite a las organizaciones transformar datos brutos en información útil para la toma de decisiones estratégicas. A través de ejemplos prácticos y casos de estudio, Curto explica las principales herramientas y técnicas de análisis de datos, la importancia de la visualización de la información y cómo implementar proyectos de BI de manera efectiva. Además, se destacan los beneficios de BI en mejorar la eficiencia operativa y competitividad en el entorno empresarial contemporáneo.

2.2. Complementaria

- Fuentes, D. & Carrión, G. (Dir.). (2021). *Modelo integrado de inteligencia de negocios como soporte a la toma de decisiones en la gestión comercial de las mypes: (1 ed.)*. D - Universidad Señor de Sipán. <https://elibro.net/es/lc/itq/titulos/228906>.

El texto presenta un enfoque para mejorar la toma de decisiones comerciales en MYPES mediante BI, destacando su implementación y beneficios en la eficiencia operativa y competitividad.

- Rodríguez, J. R. González Farran, X. & Guitart, I. (2016). *¿Cómo planificar un proyecto de inteligencia de negocio?: (ed.)*. Editorial UOC. <https://elibro.net/es/lc/itq/titulos/58548>

Es una guía práctica para la planificación de proyectos de BI. El texto aborda desde la identificación de necesidades hasta la implementación, proporcionando herramientas y metodologías esenciales para asegurar el éxito de proyectos de inteligencia de negocios.

- Perez, M. (2024). *Business Intelligence. (1 ed.)*. Editorial ICB. <https://elibro.net/es/lc/itq/titulos/283858>



Este libro explora desde los conceptos básicos hasta los aspectos más avanzados, ofreciendo a los lectores una comprensión profunda de cómo transformar datos en información valiosa para la toma de decisiones empresariales. A lo largo de sus capítulos, se abordan los pilares de la inteligencia de negocio, incluyendo la transformación y análisis de datos, el uso de sistemas como OLTP y OLAP, y la implementación de datawarehouses. Además, el libro ofrece una guía detallada sobre el proceso de minería de datos, desde la preparación y modelado hasta su despliegue. El texto también presenta el proceso de ETL (Extracción, Transformación y Carga), así como las herramientas y técnicas de visualización de datos, con un enfoque especial en plataformas como Power BI.

- Dávila, L. (2007). *Hacia la inteligencia del negocio con Excel 2003*: (ed.). Editorial Politécnico Grancolombiano. <https://elibro.net/es/lc/itq/titulos/70765>

Este libro trata los principios fundamentales de la Inteligencia de Negocios (Business Intelligence), enfocándose en la administración, gestión y explotación de la información dentro de las organizaciones. Explica conceptos clave como las fuentes de datos, los sistemas transaccionales, las bases de datos, los procesos de extracción, consolidación, análisis y visualización de la información, así como el uso de herramientas prácticas para el análisis de datos. Profundiza en el manejo de tablas y gráficos dinámicos, la consolidación de datos, la creación de consultas, el uso de formularios y macros, y la construcción de cubos OLAP usando Excel, permitiendo transformar datos en conocimiento útil para la toma de decisiones.

- Carrasco Gómez, F. J. (2024). *Power BI: curso práctico*: (1 ed.). RA-MA Editorial. <https://elibro.net/es/lc/itq/titulos/279004>

Este libro es una guía práctica y aplicada sobre Power BI, orientada al análisis y visualización de datos para la Inteligencia de Negocios. Trata temas desde la preparación, limpieza y modelado de datos usando Power Query, hasta el diseño de informes interactivos, uso de visualizaciones avanzadas, creación de medidas y cálculos con DAX, y la publicación y distribución de reportes.

3. COMPETENCIAS GENÉRICAS Y ESPECÍFICAS

Las competencias genéricas y específicas de la materia de Business Intelligence (BI) se resumen de la siguiente manera:

- Aplicar herramientas de administración de empresas digitales que permitan visualizar



datos significativos para respaldar la toma de decisiones ejecutivas, operativas y gerenciales.

- Utilizar herramientas de visualización de datos que ofrecen la capacidad de importar datos de diversas fuentes y crear informes interactivos y personalizables.
- Apoyar en la toma de decisiones ejecutivas, gerenciales y operativas mediante el correcto uso de estas herramientas en empresas digitales permite mejorar la eficiencia operativa, identificar tendencias clave y adaptarse rápidamente a un entorno empresarial en constante cambio.
- Crear y desarrollar nuevos sistemas a través de la investigación de necesidades de los usuarios.
- Evaluar y seleccionar sistemas de hardware y software con base a la información tecnológica existente y su correlación con los requerimientos de las empresas instituciones.

Estas competencias permiten a los profesionales en desarrollo de software utilizar herramientas de business intelligence para analizar y visualizar datos relevantes, lo que a su vez respalda la toma de decisiones estratégicas y operativas en la organización.

4. OBJETIVO GENERAL

Proporcionar a los estudiantes el conocimiento y las habilidades necesarias para recopilar, procesar y analizar datos empresariales, utilizando herramientas y técnicas de última generación en BI. Esto se logra mediante la enseñanza de conceptos teóricos fundamentales junto con la aplicación práctica de modelado de negocios, análisis de datos y visualización.

El propósito es capacitar a los estudiantes para que puedan interpretar grandes volúmenes de datos y convertirlos en información estratégica que apoye la toma de decisiones y la formulación de estrategias en diversas organizaciones. De esta manera, la asignatura busca dotar a los futuros tecnólogos de Desarrollo de Software no solo de competencias técnicas sino también de una visión estratégica, preparándolos para enfrentar los retos del mercado laboral actual y contribuir efectivamente al crecimiento y éxito de las empresas en la era digital.



5. FORMACIÓN CIUDADANA, VALORES Y HABILIDADES BLANDAS

La asignatura desempeña un papel crucial en el desarrollo integral de los alumnos al proporcionarles herramientas esenciales para su crecimiento personal y social, enfocándose en:

- Educación ambiental: biodiversidad
- Valores & habilidades blandas: amor: comunicación asertiva y escucha activa
- Valores & habilidades blandas: compromiso social: adaptabilidad
- Valores & habilidades blandas: cultura: creatividad
- Valores & habilidades blandas: lealtad: liderazgo
- Valores & habilidades blandas: optimismo: planificación y gestión del tiempo
- Valores & habilidades blandas: orgullo nacional: pensamiento crítico
- Valores & habilidades blandas: respeto: inteligencia emocional
- Valores & habilidades blandas: solidaridad: trabajo en equipo
- Valores & habilidades blandas: tolerancia: flexibilidad
- Valores & habilidades blandas: verdad y honradez: proactividad

6. NORMAS DE CLASE

Con relación a las normas de clase, es importante destacar que la evaluación de los componentes de gestión académica se compone de tres notas sumativas, cada una con una puntuación máxima de 6.60/6.60, así como un proyecto práctico, como evaluación formativa que se valora con 3.40/3.40, lo que da un total de 10/10 para la calificación del módulo.

Los parciales se califican en una escala de hasta 6.60 puntos, representando cada uno el 2.22 de la calificación total de 6.6 puntos.

Para presentarse al proyecto final, el estudiante debe haber obtenido al menos 4.50 puntos sumando las tres primeras notas. En caso de no alcanzar este mínimo en el proyecto, se otorga una oportunidad de recuperación dentro de las 48 horas laborables siguientes, según el calendario académico oficial.





La nota mínima acumulada requerida para aprobar la asignatura es 7/10, y es esencial mantener al menos un 70% de asistencia a las clases. Los docentes deben informar a los estudiantes sobre sus notas individuales antes de registrarlas en el sistema, y se espera que los alumnos confirmen su aceptación y conformidad con estas calificaciones. Además, los docentes deben entregar un reporte de notas y asistencia a través del SGA y notificado a la coordinación de carrera y registrar las calificaciones en el sistema en un plazo máximo de 5 días posteriores a la recepción del proyecto final.

7. SISTEMA DE EVALUACIÓN

- PARCIAL I - 22%
- PARCIAL II - 22%
- PARCIAL III - 22%
- PROYECTO FINAL - 34%

8. UNIDADES

1. Introducción la Inteligencia de Negocios
2. Herramientas y técnicas de transformación.
3. Dashboards para la toma de decisiones.



UNIDAD 1 (INTRODUCCIÓN A LA INTELIGENCIA DE NEGOCIOS)

*“La vida es una elección constante,
solo el tiempo dirá si fue la correcta o no”*

D.O

Temas y Subtemas

Introducción a BI

Importancia de BI y aplicaciones

Arquitectura de BI

Data Warehouse

Base de datos transaccional - Data Warehouse

La inteligencia de negocios (BI, por sus siglas en inglés Business Intelligence) es un proceso tecnológico que permite analizar datos y presentar información a la que se puede acceder para ayuda a los ejecutivos, gerentes y usuarios finales de la empresa a tomar decisiones informadas. Esta unidad se centra en los conceptos básicos de la inteligencia de negocios, su importancia y sus aplicaciones en el mundo empresarial actual.

Conceptos básicos de BI

La inteligencia de negocios abarca el conjunto de tecnologías, de aplicaciones y todas las mejores prácticas para la recopilación, integración, análisis y presentación de información con carácter empresarial. El objetivo de BI es apoyar la toma de decisiones en las organizaciones.



Definición de BI

BI es el conjunto de metodologías, de procesos, de arquitecturas y de tecnologías que transforman a los datos en información significativa y útil para un fin. Esto permite a las organizaciones mejorar la toma de decisiones estratégicas, tácticas y operativas (Rud, 2009).

En este contexto es importante mencionar que un *insight* se define como una comprensión significativa de los datos, que son el resultado del análisis y que proporcionan a los mandos gerenciales datos procesados que apoyen a la toma de decisiones, esta información es de suma importancia para la estrategia empresarial, los insights permiten a las organizaciones identificar patrones, tendencias y/o relaciones que no son evidentes a simple vista sino que necesitan o requieren de un análisis y pre-pocesamiento de datos para facilitar la identificación de oportunidades y poder mitigar riesgos dentro de una organización.

En el contexto de la inteligencia de negocios (BI), los insights son cruciales porque transforman grandes volúmenes de datos en información valiosa que puede ser utilizada para mejorar la eficiencia operativa, optimizar procesos y fortalecer la competitividad en el mercado. Mediante el uso de técnicas de análisis de datos, la minería de datos y el procesamiento analítico de información; las organizaciones pueden descubrir patrones y relaciones que pueden ser aprovechados para tomar decisiones estratégicas (Rud, 2009) se puede decir que los insights se refieren a comprensiones profundas y accionables que se derivan del análisis de datos.

Entre las características que presenta un *insight* se encuentran:

- **Accionable:** El insight debe ser práctico y debe permitir la toma de decisiones basada en la información obtenida. (Chaudhuri, 2011)
- **Relevante:** Debe ser pertinente al contexto específico del negocio o de la empresa, alineados a los objetivos, a la visión y misión de la empresa. (Negash, 2004)
- **Oportuno:** Debe obtenerse en el momento adecuado para que las decisiones se tomen tomarse con la información actualizada. (Turban, 2011).
- **Comprensible:** Debe ser claro y conciso para que pueda ser fácilmente entendido e interpretable por los encargados en tomar decisiones empresariales (Watson, 2009).

Importancia de BI

La inteligencia de negocios es de suma importancia y crucial en las organizaciones de vanguardia por varias razones:

- Mejora la Toma de Decisiones ya que BI proporciona a los mandos directivos datos relevantes y precisos, lo que les permite tomar decisiones basadas en hechos y caso reales en lugar de suposiciones. (Chaudhuri, 2011).
- Aumenta la Eficiencia Operativa, mediante la automatización de procesos de análisis de datos y reduce el tiempo y el esfuerzo necesarios para extraer información útil, lo que mejora la eficiencia operativa (Turban, 2011).
- Ventaja Competitiva para las organizaciones ya se puede identificar tendencias y patrones del mercado antes que sus competidores, lo que les proporciona una ventaja competitiva (Ranjan, 2009).
- Mejoramiento del Servicio al Cliente ya que permite a las empresas comprender mejor las necesidades y preferencias de sus clientes identificando tendencias y patrones, lo que facilita la personalización de productos y de los servicios. (Negash, 2004).

Aplicaciones de BI

La inteligencia de negocios tiene una gama muy diversa de aplicaciones en varios sectores:

Sector de Administración de empresas: En este ámbito, BI se aplica para mejorar la toma de decisiones estratégicas y operativas de la empresa. Las herramientas de BI permiten a los administradores analizar datos de desempeño, identificar áreas de mejora y monitorear indicadores clave de rendimiento (KPI). Además, BI facilita la planificación y previsión financiera, optimización de recursos y mejora de la eficiencia organizacional (Rud, 2009), por ejemplo, en:

- Análisis de Desempeño al ayudar a monitorear y evaluar el desempeño de diferentes departamentos y unidades de la empresa.
- Planificación Estratégica utilizando datos históricos y proyecciones que apoyan en la formulación de estrategias a largo plazo.



- Optimización de Recursos ya que permite la identificación de áreas donde se pueden optimizar los recursos y por ende reducir costos y mejorar la eficiencia.
- Gestión de la Relación con el Cliente (CRM) ya que integra los datos de diversas fuentes, mejora la comprensión del comportamiento y las preferencias del cliente, lo que facilita la personalización de servicios y productos.

Sector Financiero: Las instituciones financieras utilizan BI para detectar fraudes, gestionar riesgos y cumplir con las regulaciones nacionales e internacionales (Chen, 2012).

Sector de Retail: Se utiliza para gestionar inventarios, analizar patrones de compra y optimizar precios. (Inmon, 2008).

Sector de Salud: Ayuda en la gestión de pacientes, análisis de costos y mejora de la atención al paciente. (Wang, 2018).

Sector de Manufactura: Las empresas manufactureras utilizan BI para optimizar sus procesos de producción y mejorar la calidad del producto. (Loshin, 2012).

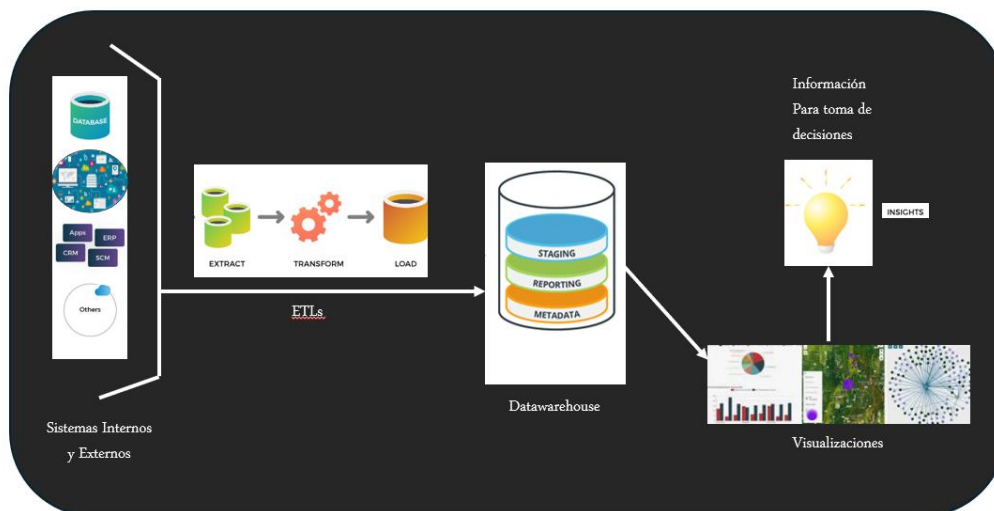
Sector de Marketing: Permite a los equipos de marketing analizar campañas, segmentar mercados y evaluar el retorno de la inversión (ROI) de sus estrategias.

Arquitectura o componentes de BI

La arquitectura de BI se compone de varios componentes que trabajan entre si en conjunto para proporcionar como objetivo una visión integral de los datos de la empresa. Estos componentes incluyen desde los sistemas externos, los procesos ETLs, el almacenamiento de datos, las herramientas de análisis y de visualización, y las plataformas de datos en la nube (Loshin, 2012)



Figura 1
Componentes de BI



Nota. En la ilustración se puede observar los componentes de Inteligencia de negocios en la gestión de datos en una organización. *Fuente:* Elaboración propia.

Es importante destacar que las fuentes de datos pueden ser de diversos orígenes que pueden provenir de bases de datos relacionales o no relacionales como PostgreSQL, MySQL, MongoDB, o cualquier gestor de base de datos, archivos de Excel, archivos CSV, de texto, entre otros. También pueden estos datos integrarse desde múltiples sistemas externos, como ERP o CRM, que proporcionan información crítica sobre las operaciones y el rendimiento de una empresa (Chaudhuri, Dayal, & Narasayya, 2011).

Los procesos ETLs cuya siglas vienen de (Extract, Transform, Load) son indispensables para la integración de los datos, estos procesos implican la extracción de datos de diversas fuentes, luego del análisis exploratorio de datos (EDA) que permite conocer los datos a través de la exploración, investigación y aprendizaje se puede realizar su transformación o limpieza, esto incluye adaptar los datos a un formato adecuado limpiar el conjunto de datos para poder realizar la carga en un almacén de datos llamado Data Warehouse, existen herramientas útiles para realizar este tipo de procesos, entre los que podemos citar son Pentaho o Apache Nifi que son comunes para llevar a cabo estas tareas (Kimball & Caserta, 2011). Se puede decir que los procesos ETLs extrae datos de las fuentes, los transforma según las necesidades del negocio y los carga en un Data Warehouse.

Un Data Warehouse en resumen son un repositorio centralizado que permite almacena datos integrados desde diversas fuentes. Este almacén de datos está diseñado mediante un

modelo de datos que agiliza las consultas y mejora los análisis, es importante mencionar que el data warehouse es el núcleo de cualquier arquitectura de BI y que las bases de datos como PostgreSQL, MySQL, SQL Server, entre otras son típicamente utilizadas para este propósito (Inmon, 2005).

Las herramientas de análisis de datos permiten a las organizaciones realizar análisis detallados y descubrir patrones ocultos en sus datos, mediante la minería de datos que es el proceso analítico que explora grandes cantidades de datos para encontrar estos patrones significativos.

Y finalmente las aplicaciones de reportes y visualización que son cruciales para presentar los datos de manera comprensible y útil, estas herramientas pueden ser Power BI, Tableau y QlikView, entre otras que se encuentran en el mercado y que son ampliamente utilizadas para crear dashboards interactivos y visualizaciones que permiten a los usuarios explorar y entender los datos fácilmente (Few, 2012).

Data Warehousing

Un Data Warehouse es un sistema utilizado para el reporte y análisis de datos, actuando como un repositorio centralizado de información que integra estos datos desde múltiples fuentes. Kimball y Ross (2013) lo definen como “un conjunto de datos orientado a temas, integrado, no volátil y variante en el tiempo, que apoya la toma de decisiones”. A diferencia de las bases de datos transaccionales, los DW están diseñados para responder consultas complejas de análisis histórico y tendencias, proporcionando una visión consolidada y coherente de los datos de la empresa.

Entre las características clave de un DW se puede destacar:

- **Orientación a temas:** Los datos se organizan por áreas de interés, como ventas, finanzas o recursos humanos.
- **Integración:** La información proviene de diversas fuentes y se consolida en un formato uniforme dentro de un modelo de datos estructural.
- **No volatilidad:** Una vez cargados, los datos no se alteran, garantizando estabilidad para el análisis.
- **Variante en el tiempo:** Los datos almacenados permiten observar tendencias y realizar

comparaciones temporales (Inmon, 2005).

El principal valor de los Data Warehouses radica en su capacidad para proporcionar información precisa y oportuna, mejorando la toma de decisiones en distintos niveles de la organización. Algunos beneficios destacados incluyen:

- **Consolidación de datos dispersos:** Los DW integran información de múltiples sistemas, eliminando redundancias y garantizando la coherencia de los datos.
- **Análisis histórico:** La capacidad de almacenar datos históricos permite identificar patrones, tendencias y relaciones que de otra manera podrían pasar desapercibidos.
- **Mejoras en la eficiencia operativa:** Los reportes y análisis se generan de manera rápida y precisa, reduciendo el tiempo invertido en tareas manuales de consolidación de datos.
- **Apoyo a la planificación estratégica:** Con información confiable, las empresas pueden anticipar escenarios futuros y tomar decisiones fundamentadas (Golfarelli, 2009).

En resumen, los Data Warehouses son una herramienta necesaria para las empresas actuales que buscan mejorar la toma de decisiones mediante el análisis de datos. Si bien la implementación puede ser desafiante, los beneficios en términos de eficiencia, análisis histórico y apoyo estratégico son invaluable. A medida que la tecnología avanza, el DW seguirá evolucionando, integrándose con nuevas soluciones y desempeñando un papel crucial en la gestión de datos empresariales.

La arquitectura de un Data Warehouse (DW) es esencial para garantizar su eficacia en la organización y análisis de datos. Dentro de este campo, los modelos de diseño de bases de datos multidimensionales, como el Modelo Estrella y el Modelo Copo de Nieve, juegan un papel crucial en la estructuración y optimización de los datos para su consulta y análisis. Este ensayo explora las características, ventajas y desventajas de estos modelos, así como su aplicación en distintos escenarios empresariales.

Modelos Multidimensionales

Al ser un Data Warehouse un sistema de almacenamiento de datos diseñado para facilitar el análisis y la toma de decisiones. Los modelos multidimensionales, como el Modelo Estrella y el Modelo Copo de Nieve, son estructuras que permiten organizar los datos de manera eficiente para consultas analíticas, centrándose en las dimensiones y las medidas clave.





Modelo Estrella

El Modelo Estrella es una estructura en la que los datos se organizan alrededor de una tabla central, conocida como tabla de hechos, que contiene las medidas cuantitativas del negocio. Las tablas dimensión, que describen los contextos de las medidas, están directamente conectadas a la tabla de hechos. Esta configuración se asemeja a una estrella, de ahí su nombre (Kimball & Ross, 2013). Entre las ventajas del modelo estrella se puede mencionar:

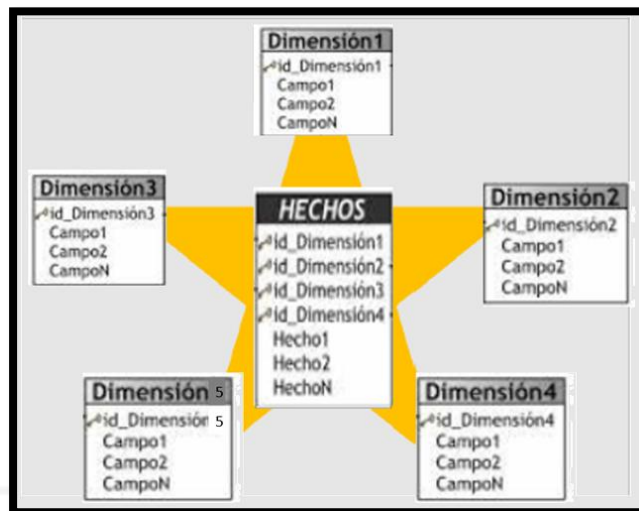
- **Simplicidad:** Su diseño es fácil de entender y manejar para los usuarios.
- **Eficiencia en consultas:** Las uniones entre tablas son directas, lo que acelera el tiempo de respuesta en consultas.
- **Compatibilidad:** Funciona bien con herramientas de visualización y generación de reportes.

Entre las desventajas se puede mencionar:

- **Redundancia de datos:** Las tablas dimensión pueden contener información repetida, lo que incrementa el tamaño del almacenamiento.
- **Escalabilidad limitada:** No es la mejor opción para sistemas con datos extremadamente complejos o grandes.



Figura 2
Modelo Estrella



Nota. En la ilustración se puede observar el modelo estrella que trabaja con la tabla hecho y las tablas dimensiones. *Fuente:* Elaboración propia.

Modelo Copo de Nieve

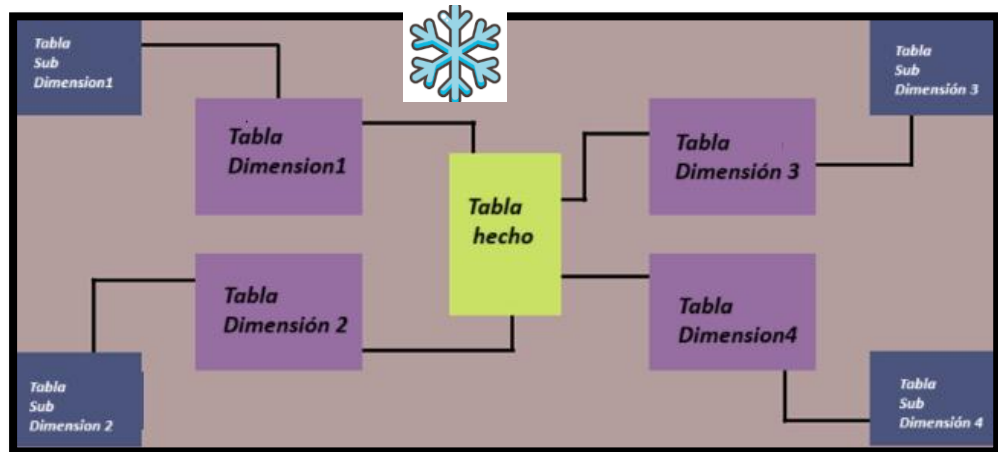
El Modelo Copo de Nieve es una extensión del Modelo Estrella, donde las tablas dimensión están normalizadas. Esto significa que los datos redundantes se dividen en múltiples tablas relacionadas, formando una estructura que se asemeja a un copo de nieve (Golfarelli & Rizzi, 2009). Entre las ventajas del modelo estrella se puede mencionar:

- **Menor redundancia:** La normalización reduce el almacenamiento duplicado, optimizando el espacio.
- **Consistencia:** Facilita la integridad y coherencia de los datos.
- **Escalabilidad:** Mejor adaptado para sistemas complejos y grandes volúmenes de datos.

Entre las desventajas se puede mencionar:

- **Complejidad en las consultas:** Las uniones entre múltiples tablas pueden ralentizar el rendimiento.
- **Mayor dificultad de uso:** Es menos intuitivo para usuarios finales y desarrolladores.

Figura 3
Modelo Copo de nieve



Nota. En la ilustración se puede observar el modelo copo de nieve que trabaja con la tabla hecho y las tablas dimensiones que a su vez se relacionan con tablas subdimensiones. *Fuente:* Elaboración propia.

Tanto el Modelo Estrella como el Modelo Copo de Nieve son fundamentales para el diseño de Data Warehouses eficaces. Cada modelo ofrece ventajas y desventajas que deben evaluarse según las necesidades específicas de la organización. La combinación de ambos enfoques puede ofrecer soluciones óptimas, permitiendo a las empresas maximizar el valor de sus datos y mejorar la toma de decisiones.

Bases de datos y Data Warehouse

El manejo de datos en las organizaciones modernas ha evolucionado significativamente, llevando a la implementación de sistemas que permiten el almacenamiento y análisis eficiente de información. En este contexto, las bases de datos y los Data Warehouses (DW) desempeñan roles complementarios pero distintos. Este ensayo analiza los conceptos fundamentales de ambos sistemas y detalla el proceso de transformación de una base de datos transaccional a un Data Warehouse.

Una base de datos es un sistema que permite el almacenamiento y la gestión de datos estructurados. Está diseñada principalmente para operaciones transaccionales, como la inserción, actualización y eliminación de registros. Según Connolly y Begg (2015), las bases de datos transaccionales son esenciales para las operaciones diarias de una organización, proporcionando una gestión eficiente de los datos actuales.

Por otro lado, un Data Warehouse es una colección de datos orientada a temas, integrada, no volátil y variante en el tiempo, diseñada para el soporte de decisiones (Inmon, 2005). Su función principal es consolidar datos de múltiples fuentes y ofrecer una visión histórica coherente que facilite el análisis y la generación de reportes.

Diferencias Clave entre Bases de Datos y Data Warehouses

Existen diferencias muy marcadas entre una DB y un DW según tres conceptos su propósito, su estructura y su temporalidad.

- Según su propósito, las bases de datos se enfocan en operaciones transaccionales diarias, mientras que los DW están diseñados para el análisis y la toma de decisiones.
- Según su estructura, las bases de datos suelen estar normalizadas para reducir redundancias, mientras que los DW utilizan modelos multidimensionales para optimizar el análisis.
- Según la temporalidad, las bases de datos contienen datos actuales, mientras que los DW almacenan información histórica.

Transformación de una Base de Datos Transaccional a un Data Warehouse

El proceso de transformación de una base de datos transaccional a un Data Warehouse implica varias etapas. Estas etapas garantizan la integridad, consistencia y utilidad de los datos para el análisis. A continuación, se describen los pasos principales del proceso:

Extracción de Datos: En esta etapa, los datos se extraen de las bases de datos transaccionales y otras fuentes de datos relevantes. Esto incluye sistemas ERP, CRM y archivos planos. La extracción debe considerar la calidad y relevancia de los datos para el propósito del DW (Kimball & Ross, 2013).

Transformación de Datos: La transformación es una etapa crítica en la que los datos extraídos se limpian, integran y convierten al formato requerido por el DW. Esto incluye:

Limpieza: Eliminación de datos duplicados o inconsistentes.

Integración: Combinación de datos provenientes de diferentes fuentes.

Conversión: Formateo de datos para cumplir con los estándares del DW.



Carga de Datos: En esta etapa, los datos transformados se cargan en el Data Warehouse. La carga puede realizarse de manera incremental o completa, dependiendo de los requisitos del sistema y la periodicidad del análisis.

Modelado Multidimensional: Una vez que los datos están en el DW, se organizan utilizando modelos como el Modelo Estrella o el Modelo Copo de Nieve. Estos modelos permiten estructurar los datos en tablas de hechos y tablas de dimensiones, facilitando el análisis rápido y eficiente (Golfarelli & Rizzi, 2009).

Optimización y Generación de Reportes: Finalmente, se implementan herramientas de Business Intelligence (BI) para permitir la consulta y generación de reportes. Estas herramientas aprovechan la estructura del DW para ofrecer análisis visuales y detallados.

Se concluye que tanto las bases de datos como los Data Warehouses desempeñan un papel crucial en la gestión de datos empresariales. Mientras que las bases de datos son esenciales para las operaciones o transacciones diarias, los DW son indispensables para el análisis y la toma de decisiones a largo plazo. La transformación de bases de datos transaccionales a un DW es un proceso complejo, pero altamente beneficioso, que permite a las organizaciones maximizar el valor de sus datos.



Autoevaluación 1

Lea detenidamente la pregunta y marque la respuesta correcta entre las opciones dadas.

1. ¿Qué es BI?

- a) Un sistema que permite el almacenamiento de datos empresariales.
- b) Un proceso para transformar datos en información significativa y útil.
- c) Una herramienta para la comunicación entre departamentos empresariales.
- d) Una tecnología que permite realizar la gestión de administración de

empresas.

2. ¿Qué es un Insight?

- a) Un conjunto de datos crudos sin procesar.
- b) Un informe detallado de las ventas mensuales.
- c) Una herramienta para la gestión de proyectos empresariales.
- d) Una comprensión profunda y significativa derivada del análisis de datos.

3. ¿Cuáles son las características de un Insight?

- a) Debe ser confuso y difícil de entender.
- b) No necesita ser relevante para el negocio.
- c) Debe ser accionable, oportuno y comprensible.
- d) Solo debe ser comprensible no requiere ser relevante.

4. Identifique la importancia de Inteligencia de negocios en una empresa.

- a) Mejora la toma de decisiones al basarse en hechos y datos precisos.
- b) Reduce la eficiencia operativa.
- c) No tiene impacto en la ventaja competitiva.





d) No afecta el servicio al cliente.

5. Son componentes claves de BI, excepto:

- a) El almacenamiento de datos.
- b) La gestión de recursos humanos.
- c) La presentación y visualización de datos.
- d) El proceso de transformación de datos.

6. ¿Qué significan las siglas BI?

- a) Business Information.
- b) Business Integration.
- c) Business Innovation.
- d) Business Intelligence.

7. ¿Qué modelo de diseño de Data Warehouse organiza los datos en tablas de hechos y dimensiones conectadas directamente??

- a) Modelo Copo de Nieve.
- b) Modelo Dimensional Relacional.
- c) Modelo Estrella.
- d) Modelo de red.

8. ¿Qué componente de la arquitectura de BI se encarga de transformar datos en un formato adecuado para su análisis?

- a) Almacenamiento de datos.
- b) Visualización de datos.
- c) Procesos ETLs.
- d) Modelos multidimensionales.





9. ¿Qué significa ETL?

- a) Extract, Transform, Load.
- b) Extract, Translate, Load.
- c) Extract, Transport, Load.
- d) Extract, Transform, Link.

10. ¿Qué función tiene un Data Warehouse en la arquitectura de BI?

- a) Guardar datos en tiempo real para procesos transaccionales.
- b) Almacenar datos integrados de múltiples fuentes para análisis rápidos.
- c) Transformar datos para adaptarlos a las necesidades del negocio.
- d) Ejecutar aplicaciones y procesos de mensajería.

Si ha concluido la autoevaluación:

Verifique sus respuestas en el solucionario que se encuentra al final de la presente Guía.



Si alcanzó un porcentaje alto de respuestas correctas en las preguntas de evaluación, es hora de continuar con la Unidad 2, caso contrario, vuelva a revisar los temas en los que ha tenido dificultad.



Resumen de la Unidad 1

La Unidad 1 trata sobre los fundamentos teóricos de la inteligencia de negocios (BI) explora los conceptos básicos, la importancia y las aplicaciones de BI en el contexto empresarial moderno. BI, o Business Intelligence es un proceso tecnológico que permite a las organizaciones transformar datos en información útil para tomar decisiones estratégicas, tácticas y operativas. Esta disciplina combina herramientas, aplicaciones y metodologías para recopilar, integrar, analizar y presentar datos empresariales. Los insights derivados de BI son esenciales para identificar patrones, tendencias y oportunidades, facilitando la personalización de servicios y la mejora de la eficiencia organizacional. Adicional el Data Warehouse (DW) es un componente central de BI. Funciona como un repositorio centralizado que almacena datos integrados de diversas fuentes, permitiendo el análisis histórico y la generación de reportes. Las características distintivas del DW incluyen su orientación temática, integración, estabilidad y capacidad para almacenar datos históricos. Estas propiedades son fundamentales para mejorar la toma de decisiones, optimizar procesos operativos y proporcionar una ventaja competitiva. El proceso de creación de un DW desde bases de datos transaccionales implica varias etapas clave: extracción, transformación y carga (ETL) de datos. Estas etapas aseguran que los datos sean consistentes, relevantes y utilizables para análisis avanzados. Además, el modelado multidimensional, a través de modelos como el Estrella y el Copo de Nieve, organiza los datos de manera eficiente, optimizando las consultas y el análisis. BI encuentra aplicaciones en múltiples sectores, incluyendo administración, finanzas, retail, salud, manufactura, marketing entre otros. En cada ámbito, facilita la toma de decisiones basada en hechos, mejora el servicio al cliente y fortalece la competitividad. Además, las herramientas de BI, como Power BI, proporcionan visualizaciones interactivas que transforman datos complejos en información accesible. En conclusión, BI y los Data Warehouses son esenciales para las organizaciones que buscan maximizar el valor de sus datos. Aunque su implementación puede ser desafiante, los beneficios en términos de eficiencia, competitividad y toma de decisiones estratégica justifican ampliamente el esfuerzo. A medida que las tecnologías avanzan, BI y DW continuarán evolucionando, consolidándose como pilares de la gestión empresarial moderna.



¡Siga adelante! ¡Muchos éxitos!





UNIDAD 2 (HERRAMIENTAS Y TÉCNICAS DE TRANSFORMACIÓN)

“El éxito llega para aquellos que están dispuestos a trabajar un poco más que el resto”

Og Mandino.

Temas y Subtemas

ETLs

Componentes del ETL

Herramientas utilizadas en los procesos ETL

Pentaho Data Integration

Creación de transformaciones

Diseño de transformaciones

Ejecucion de transformaciones

La importancia de saber gestionar los datos de una organización impulsa el desarrollo de tecnologías especializadas en la transformación de datos en información útil para la toma de decisiones.

ETLs

Los procesos ETL (Extract, Transform, Load) son un conjunto de procedimientos que permiten extraer datos de diversas fuentes, para luego transformarlos según las necesidades del negocio y finalmente cargarlos en un almacén de datos llamado Data Warehouse (DW).





Componentes del ETL:

1. **Extracción:** Obtención de datos desde diferentes fuentes como bases de datos, archivos planos, sistemas ERP o CRM.
2. **Transformación:** Limpieza, integración y adaptación de los datos para cumplir con los estándares del DW.
3. **Carga:** Transferencia de los datos transformados al DW para su almacenamiento y análisis.

Herramientas utilizadas en los procesos ETL

Existen diversas herramientas para implementar procesos ETL, tanto de código abierto como comerciales. Algunas de las más populares incluyen:

- **Pentaho Data Integration (PDI):** Herramienta de código abierto que ofrece funcionalidades robustas para ETL.
- **Apache Nifi:** Diseñado para la automatización del flujo de datos entre sistemas.
- **Talend:** Plataforma de integración de datos con soporte para grandes volúmenes.
- **Informatica PowerCenter:** Solución comercial ampliamente utilizada en empresas grandes.
- **Microsoft SQL Server Integration Services (SSIS):** Ideal para entornos basados en Microsoft.

Instalación de Pentaho Data Integration

Pentaho Data Integration, también conocido como Kettle, es una herramienta de código abierto que se utiliza para generación de procesos ETL. A continuación, se describe el proceso de instalación:

- **Descarga:** Obtener el archivo desde el sitio oficial de Pentaho o desde repositorios confiables.
- **Requisitos previos:** Asegurarse de contar con Java Development Kit (JDK) instalado, ya que





PDI depende de Java.

- **Extracción del paquete:** Descomprimir el archivo descargado en una ubicación adecuada.
- **Ejecución:** Abrir el archivo Spoon.bat (Windows) o Spoon.sh (Linux/MacOS) para iniciar la interfaz gráfica de PDI.

A continuación, se describe el proceso de instalación de Pentaho en los diferentes Sistemas Operativos:

Tabla 1:
Requerimientos de software para instalar Pentaho

Sistemas Operativos	Windows	Linux / MAC
Requisitos Previos	<p>Java Development Kit (JDK): Pentaho requiere JDK 8 o superior. Se lo puede instalar siguiendo el siguiente link:</p> <p>https://www.java.com/es/download/manual.jsp</p>	<p>Java Development Kit (JDK): Pentaho requiere JDK 8 o superior. Se lo puede instalar siguiendo el siguiente link:</p> <p>Para Linux:</p> <p>https://www.java.com/es/download/linux_manual.jsp</p> <p>para MAC:</p> <p>https://docs.oracle.com/en/java/javase/21/install/installation-jdk-macos.html</p>
	<p>Descargar Pentaho Data Integration desde la página oficial de Pentaho</p> <p>https://privatefilesbucket-community-edition.s3.us-west-2.amazonaws.com/9.4.0.0-343/ce/client-tools/pdi-ce-9.4.0.0-343.zip</p>	





Pasos de Instalación	Descomprime el archivo descargado en una ubicación de su preferencia.	
Configuración de PDI	Navegar en la carpeta descomprimida de Pentaho Data Integration.	
	Ejecutar el archivo Spoon.bat para iniciar la herramienta Spoon, que es la interfaz gráfica de PDI.	Ejecutar el archivo spoon.sh (Linux/Mac) para iniciar la herramienta Spoon, que es la interfaz gráfica de PDI.

Nota. En la tabla se visualiza los requisitos y los links de instalación dependiendo del Sistema Operativo.

Fuente: Elaboración propia.

Familiarizándose con Pentaho Data Integration

Los componentes de PDI incluyen:

- **Spoon:** Es la interfaz gráfica para diseñar y ejecutar trabajos (jobs) y transformaciones (transformations).
- **Transformations:** Son los procesos que definen cómo extraer, transformar y cargar datos.
- **Jobs:** Son los procesos que coordinan la ejecución de múltiples transformaciones y otras tareas.

Además, Pentaho ofrece una interfaz gráfica intuitiva para diseñar y gestionar procesos ETL mediante la herramienta Spoon. Algunos conceptos importantes para familiarizarse con PDI incluyen:

- **Transformaciones:** Definen las operaciones realizadas en los datos, como limpieza o conversión.
- **Jobs:** Secuencias de transformaciones y tareas automatizadas.

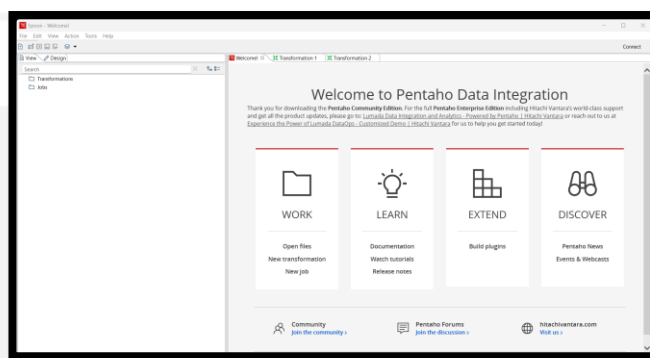


- **Conexiones:** Configuración de acceso a las fuentes y destinos de datos.

Spoon al ser la interfaz gráfica de usuario (GUI) de Pentaho Data Integration (PDI), se utiliza para diseñar, ejecutar y monitorear transformaciones (ETL) y trabajos que se realizan sobre los datos. Spoon permite a los usuarios conectar, transformar y visualizar datos de diversas fuentes de manera intuitiva y eficiente.

La pantalla principal de Pentaho ofrece varias opciones, como abrir un archivo existente, crear uno nuevo, acceder a tutoriales, y más tal como se muestra en la siguiente ilustración:

Figura 4
Pantalla principal de Pentaho-Spoon



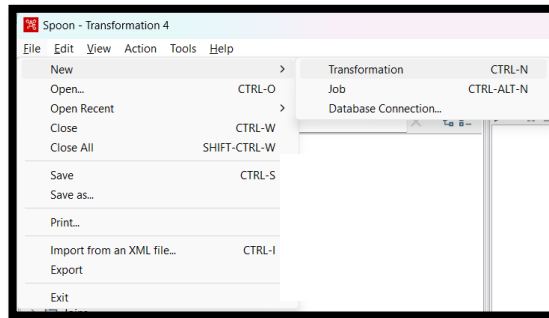
Nota. En la ilustración se puede observar la interfaz de usuario del software Pentaho Data Integration, también conocido como Spoon. *Fuente:* Elaboración propia.

Creación de transformaciones

La creación de transformaciones en Pentaho es un proceso intuitivo gracias a su interfaz gráfica Spoon, que permite a los usuarios diseñar flujos ETL arrastrando y soltando componentes en un lienzo de trabajo. Para comenzar, es necesario establecer el objetivo de la transformación, como la integración de datos, limpieza, o cálculo de métricas. Posteriormente, se agregan y configuran los pasos, que incluyen componentes de entrada, transformación y salida. Cada paso puede ser vinculado a otro, definiendo un flujo lógico que representa cómo los datos serán procesados. Durante la creación, Spoon permite probar y depurar los flujos en tiempo real, ayudando a identificar errores o cuellos de botella en el procesamiento. Este enfoque garantiza que las transformaciones creadas no solo cumplan con los requisitos técnicos, sino que también sean eficientes y escalables.

Se puede realizar la creación de una Nueva Transformación seleccionando *New* en el menú principal de *File*, tal como se visualiza en la siguiente ilustración.

Figura 5
Menú de File de Pentaho



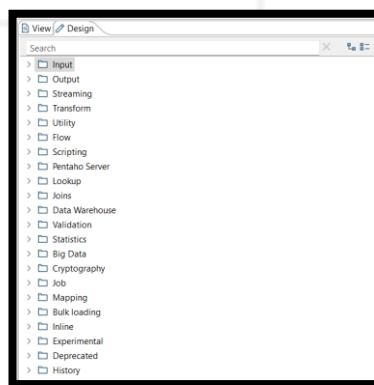
Nota. En la ilustración se puede observar el menú desplegable *File* de la aplicación Pentaho Data Integration.

Fuente: Elaboración propia.

En esta interfaz se puede realizar el diseño de la transformación, donde se pueden agregar los pasos que son los componentes básicos de una transformación. Para agregar un paso, se arrastra y suelta el paso deseado desde el panel izquierdo *Design* al área de trabajo.

Existen varios pasos como se muestran en la ilustración siguiente:

Figura 6
Panel Design



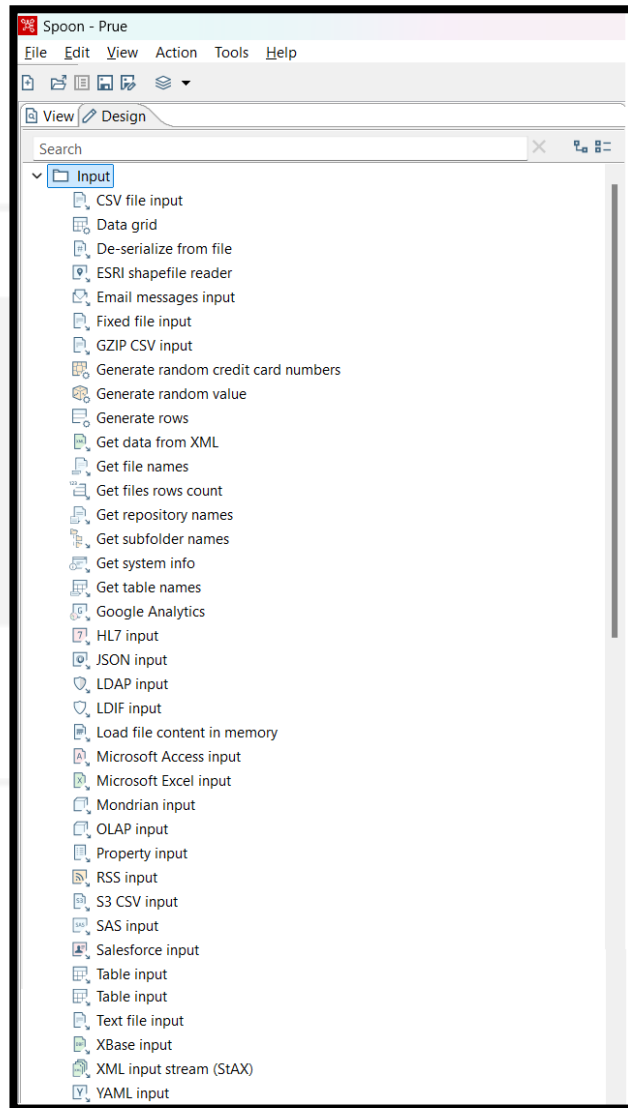
Nota. En la ilustración se puede observar una vista de la interfaz del software Pentaho Data Integration, donde se presenta una lista de categorías y componentes disponibles en el panel Design. *Fuente:* Elaboración Propia.

El Componente Input en Pentaho Data Integration

Dentro de Pentaho Data Integration, el componente de entrada o input es fundamental para iniciar los procesos ETL. Este paso define cómo se obtienen los datos desde sus fuentes originales para ser procesados.

Figura 7

Ventana de Pentaho del componente Input



Nota. En la ilustración se puede observar en la interfaz de usuario del software Pentaho Data Integration la ventana de Design el componente Input. *Fuente:* Elaboración propia

Tipos de input en Pentaho:

En Pentaho, los tipos de inputs son los componentes que permiten la extracción de datos desde diversas fuentes para iniciar el proceso ETL. Cada tipo está adaptado para necesidades específicas, asegurando flexibilidad y eficiencia en la integración de datos. Entre los tipos destacados se puede mencionar:

- **Table Input:** Extrae datos desde bases de datos relacionales utilizando consultas SQL. Este paso permite conectarse a bases de datos como PostgreSQL, MariaDB, SQLite, Sybase, MS SQL Server, Azure SQL, Oracle, entre otras, y ejecutar consultas que seleccionen los datos necesarios para la transformación.
- **Text File Input:** Carga datos desde archivos planos como TXT, o delimitados por tabulaciones. Es ideal para manejar grandes volúmenes de datos exportados desde otros sistemas o generados por procesos automatizados.
- **Excel Input:** Permite la extracción de datos desde hojas de cálculo en formatos XLS o XLSX. Es útil para empresas que manejan reportes y datos de clientes en formatos tabulares de fácil manejo.
- **JSON Input:** Carga datos desde archivos o servicios web en formato JSON. Este paso es crucial en la integración de datos provenientes de APIs modernas o sistemas de intercambio de datos.
- **XML Input:** Permite extraer datos de archivos o flujos en formato XML, comúnmente utilizado en integraciones con sistemas de legado o herramientas que exportan datos en este formato estructurado.
- **CSV File Input:** Permite la lectura y extracción de datos desde archivos delimitados, típicamente en formato CSV (valores separados por comas). Este tipo de archivo es ampliamente utilizado debido a su simplicidad y compatibilidad con múltiples sistemas.

Existen otros tipos de input que amplían las capacidades de entrada de datos, adaptándose a diversas necesidades. Ejemplos de estos incluyen el Data Grid, que permite generar datos manualmente dentro de la herramienta; Email Message Input, para extraer datos directamente desde correos electrónicos; Fixed File Input, diseñado para leer archivos de ancho fijo; GZIP CSV Input, que procesa archivos CSV comprimidos en formato GZIP.



Adicional también existen Generate Random Credit Card Number, que genera números de tarjetas de crédito para pruebas, Generate Row, que crea filas de datos predefinidas. Es importante mencionar que Pentaho soporta formatos especializados como Get Data from XML, para extraer información de archivos XML; LDAP Input, para integraciones con directorios LDAP; OLAP Input para datos multidimensionales, SAS Input para datos provenientes de sistemas SAS, XBase Input para archivos dBASE, y YAML Input para datos estructurados en YAML, entre otros. Todos estos tipos de inputs hacen de Pentaho una solución versátil capaz de manejar múltiples formatos y fuentes de datos en cualquier entorno empresarial

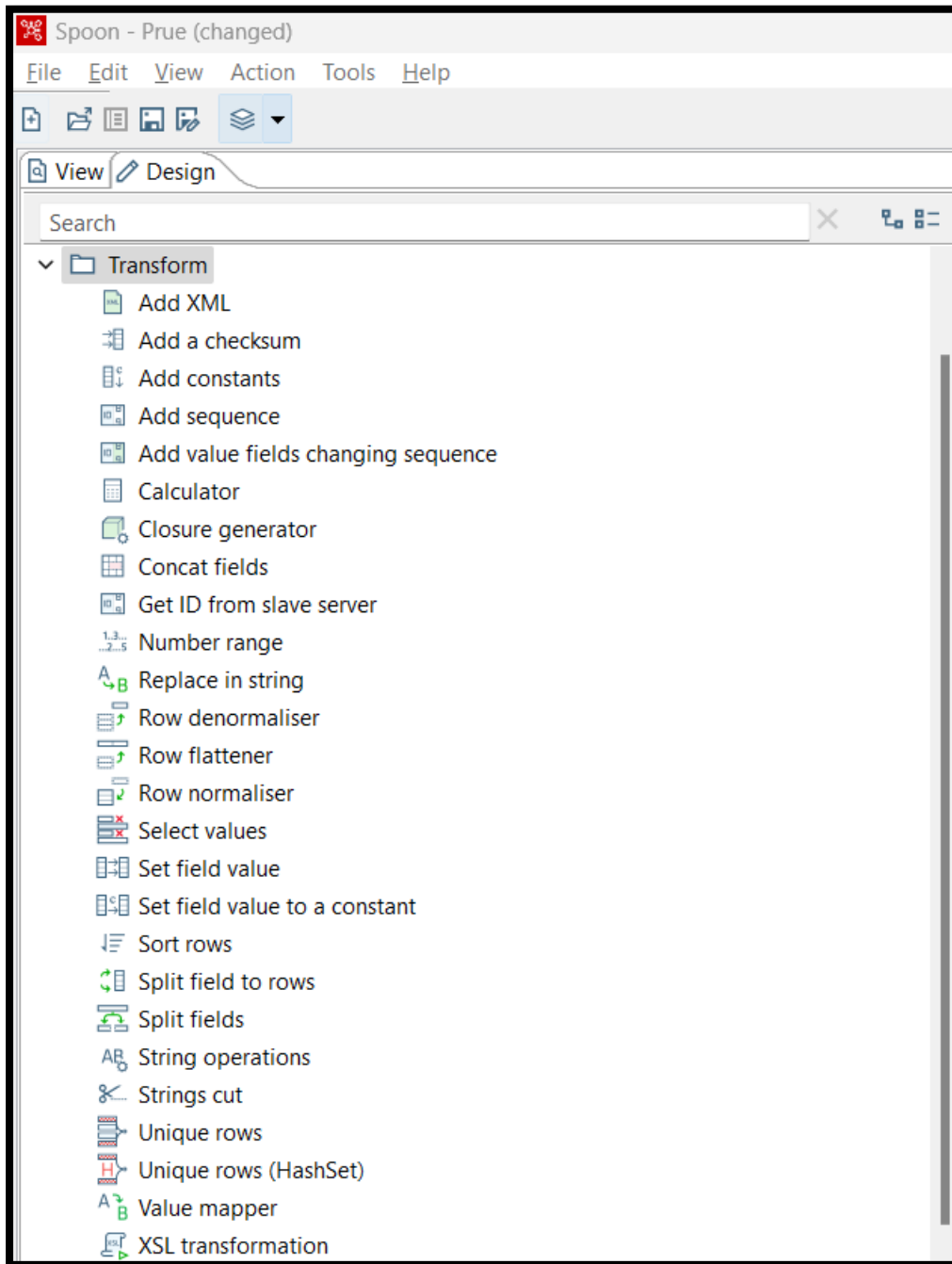
El paso de entrada asegura que los datos se extraigan correctamente y de manera eficiente desde diversas fuentes, sentando las bases para la transformación y carga posteriores. Cada tipo de input está diseñado para adaptarse a diferentes necesidades y formatos, permitiendo a las organizaciones manejar datos desde múltiples orígenes sin comprometer la calidad ni la consistencia.

El Componente Transform en Pentaho Data Integration

El componente de transformación o transform en Pentaho se utiliza para realizar operaciones sobre los datos extraídos, asegurando que estén en el formato adecuado para su análisis o almacenamiento. Este paso es esencial para limpiar, combinar, calcular y estructurar los datos.



Figura 8
Ventana de Pentaho del componente Transform



Nota. En la ilustración se puede observar en la interfaz de usuario del software Pentaho Data Integration la ventana de Design el componente Transform. *Fuente:* Elaboración propia.

Tipos de transform en Pentaho:

- **Sort Rows:** Ordena los datos según una o más columnas en orden ascendente o descendente.

- **Add Constants:** Agrega columnas con valores constantes a los datos.
- **Calculator:** Realiza cálculos matemáticos o lógicos sobre los datos, como sumas, restas, fechas o expresiones complejas.
- **Replace in String:** Reemplaza textos específicos dentro de los datos, útil para limpieza y estandarización.
- **Row Normalizer:** Convierte datos de formato ancho a largo, redistribuyendo valores en columnas separadas.

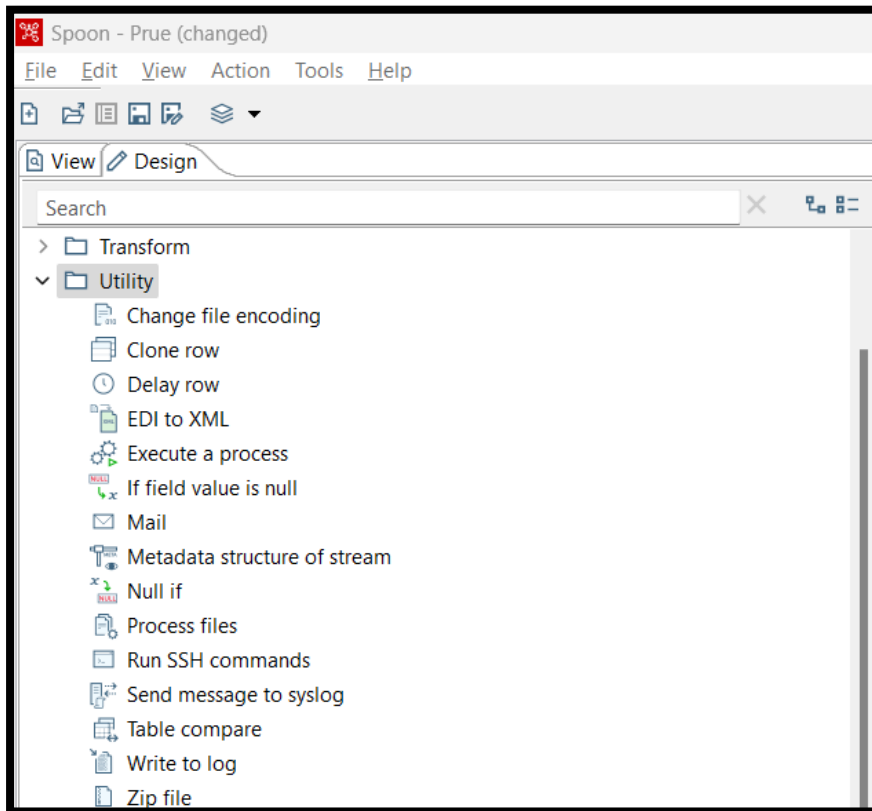
En las transformaciones comunes, se pueden emplear otras opciones avanzadas como XML Transformation, que permite trabajar con datos en formato XML, estructurándolos o transformándolos según sea necesario; String Cut, para dividir cadenas de texto en fragmentos más pequeños; Unique Row, que filtra y elimina filas duplicadas en un flujo de datos; Number Range, que clasifica valores numéricos en rangos definidos; Add Constants, que añade columnas con valores constantes a los datos; Add XML, que genera estructuras XML dentro del flujo ETL; Add a Checksum, que crea sumas de verificación para validar la integridad de los datos, entre otros. Estas opciones amplían la capacidad de Pentaho para manejar transformaciones complejas y específicas en diferentes escenarios empresariales.

El Componente Utility en Pentaho Data Integration

El componente de utilidades o utility en Pentaho proporciona herramientas auxiliares que facilitan la gestión y control de los procesos ETL. Aunque no están directamente relacionadas con la extracción, transformación o carga de datos, estas utilidades son clave para automatizar, monitorear y optimizar flujos ETL.



Figura 9
Ventana de Pentaho del componente Utility



Nota. En la ilustración se puede observar en la interfaz de usuario del software Pentaho Data Integration la ventana de Design el componente Utility. *Fuente:* Elaboración propia.

Tipos de Utility en Pentaho:

- **Delay row:** Introduce pausas temporales en la ejecución del flujo, útil para sincronizar procesos.
- **Mail:** Envía correos electrónicos, útil para notificaciones automáticas al completar procesos ETL.
- **Write to Log:** Registra información personalizada en los logs de ejecución, útil para auditorías y debugging.

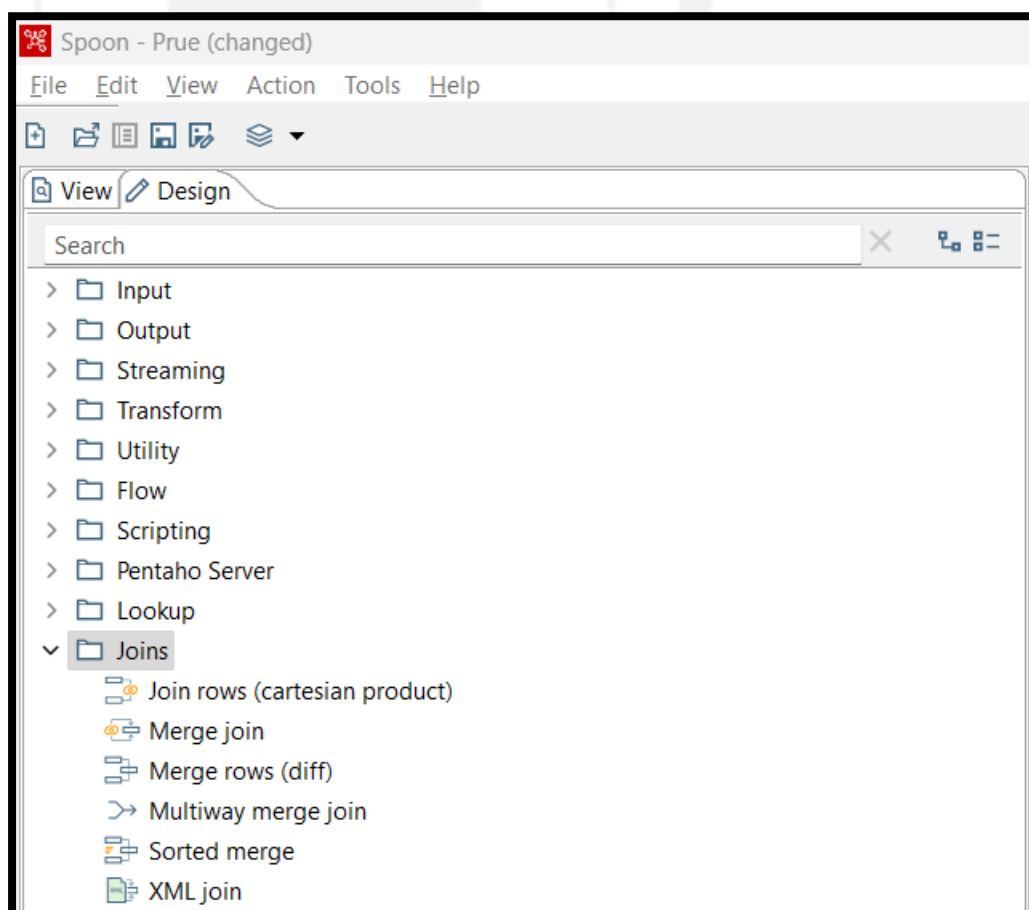
Existen componentes avanzados en la categoría de utilidades que enriquecen los procesos ETL, como Clone Row, que genera duplicados de filas para pruebas o análisis paralelos; Execute a Process, que permite ejecutar procesos externos dentro del flujo ETL; If Field Value is Null, para manejar valores nulos con lógicas condicionales; Null If, que reemplaza valores

específicos por nulos; Process File, que permite manipular archivos directamente desde el flujo; Table Compare, para comparar datos entre tablas y detectar cambios; y Zip File, que facilita la compresión de archivos generados o utilizados. Estos pasos proporcionan herramientas clave para personalizar, automatizar y optimizar flujos complejos de datos.

El Componente Joins en Pentaho Data Integration

El componente de uniones o **joins** en Pentaho permite combinar datos provenientes de diferentes flujos o fuentes en un solo conjunto, basándose en claves comunes o condiciones específicas. Este paso es crucial cuando los datos relevantes están dispersos en múltiples tablas, archivos o sistemas.

Figura 10
Ventana de Pentaho del componente Joins



Nota. En la ilustración se puede observar en la interfaz de usuario del software Pentaho Data Integration la ventana de Design el componente Joins. *Fuente:* Elaboración propia.



Tipos de joins en Pentaho:

- **Join Rows (Cartesian Product):** Realiza un producto cartesiano, combinando cada fila de un flujo de datos con todas las filas de otro. Este paso se utiliza principalmente en situaciones donde no se necesitan claves comunes y se requiere explorar todas las combinaciones posibles.
- **Merge Rows (Diff):** Compara dos flujos de datos y detecta diferencias entre ellos, como registros añadidos, modificados o eliminados. Es ideal para tareas de sincronización de datos o auditoría.
- **Multiway Merge Join:** Combina múltiples flujos de datos ordenados en uno solo, utilizando una clave común para realizar las uniones. Este paso es eficiente en flujos que requieren combinar más de dos conjuntos de datos.
- **Sorted Merge:** Fusiona dos flujos de datos previamente ordenados según una clave común, asegurando que el resultado mantenga el orden. Es útil en escenarios donde el orden de los datos es crítico.
- **XML Join:** Combina datos provenientes de flujos estándar con estructuras en formato XML. Este paso permite integrar datos estructurados en XML dentro de flujos ETL, facilitando la interoperabilidad con sistemas que utilizan este formato.

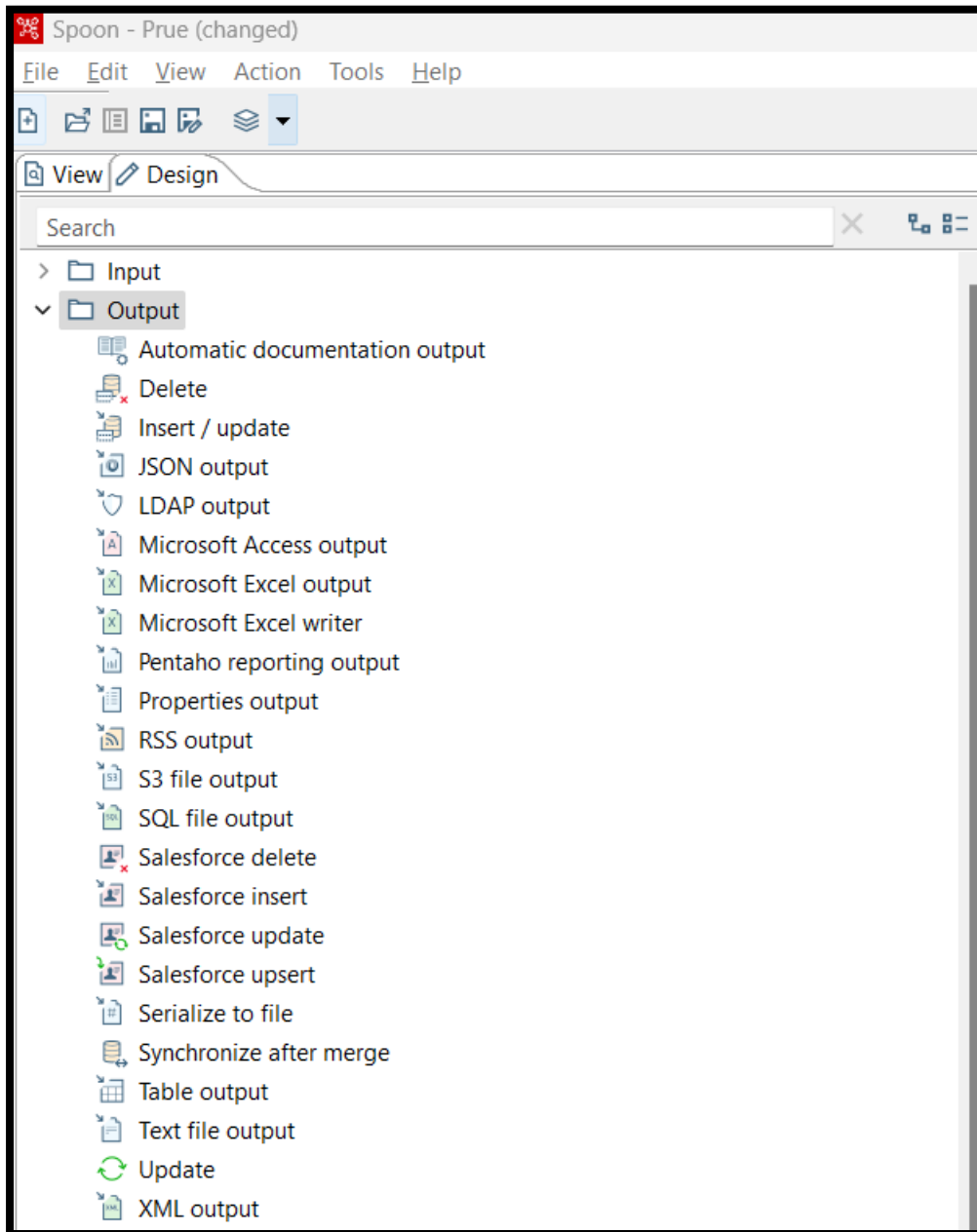
El componente joins es fundamental para consolidar datos dispersos en un conjunto coherente, facilitando análisis más completos y precisos. Su flexibilidad permite adaptarse a diferentes escenarios, desde la integración de datos históricos con actuales, hasta la validación de cambios entre versiones de datos. La variedad de opciones disponibles en Pentaho garantiza que los usuarios puedan manejar tanto escenarios simples como complejos de combinación de datos con eficiencia.

El Componente Output en Pentaho Data Integration

El componente de salida o output en Pentaho es crucial para almacenar y distribuir los datos procesados al destino final. Este paso completa el flujo ETL al garantizar que los datos transformados se escriban correctamente en el formato o sistema requerido.



Figura 11
Ventana de Pentaho del componente Output



Nota. En la ilustración se puede observar en la interfaz de usuario del software Pentaho Data Integration la ventana de Design el componente Output. *Fuente:* Elaboración propia.

Tipos de output en Pentaho:

- **Table Output:** Escribe datos procesados en una tabla de base de datos relacional como MySQL, PostgreSQL o Oracle.
- **Text File Output:** Guarda datos en archivos planos como CSV, TXT, o delimitados por



tabulaciones, ideal para compartir datos procesados entre sistemas.

- **Microsoft Excel Output:** Exporta los datos transformados a hojas de cálculo Excel en formato XLS o XLSX, útil para reportes y análisis manual.
- **JSON Output:** Convierte los datos en formato JSON para integrarlos con aplicaciones modernas o servicios web.
- **XML Output:** Escribe los datos en archivos XML estructurados, comúnmente utilizados en integraciones con sistemas de legado.
- **Insert/Update:** Realiza operaciones mixtas de inserción y actualización en bases de datos, dependiendo de si los registros ya existen.

Existen otros tipos de salida como LDAP Output, que permite escribir datos directamente en directorios LDAP, facilitando la integración con sistemas de autenticación y directorios empresariales; SQL File Output, que genera scripts SQL para insertar o actualizar datos en bases de datos, útil para entornos donde no se permite conexión directa; y Serialize to File, que guarda objetos Java serializados en archivos, ideal para escenarios donde se necesita persistencia o transferencia de estructuras complejas.

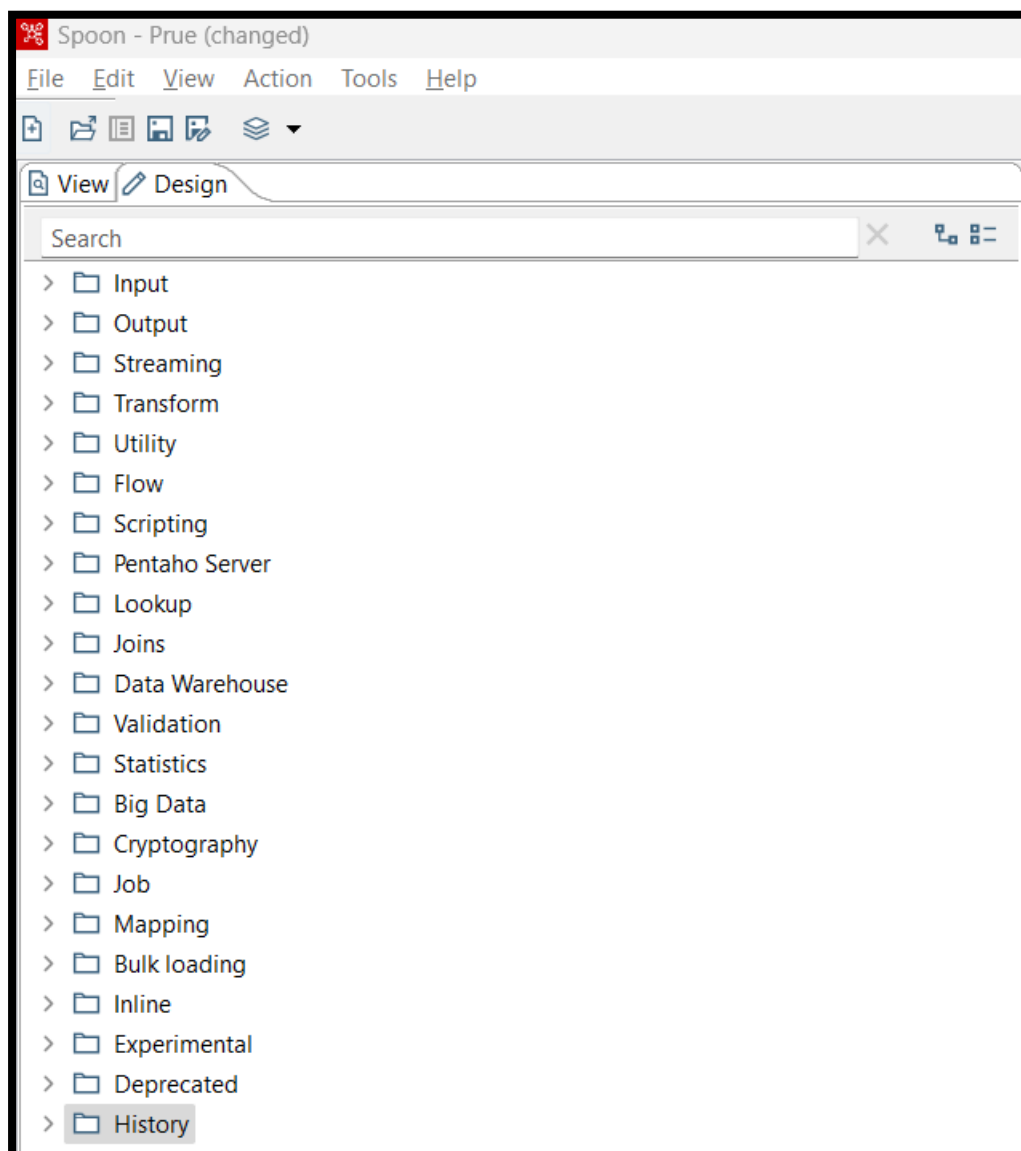
Es importante mencionar que el componente de salida asegura que los datos procesados estén disponibles en el formato o sistema adecuado para su análisis, almacenamiento o distribución. Su flexibilidad permite adaptarse a diferentes escenarios, desde la generación de reportes hasta la integración con aplicaciones empresariales o plataformas en la nube.

Otros Componentes en Pentaho Data Integration

Además de los componentes principales como Input, Transform, Joins y Output, Pentaho ofrece una variedad de herramientas adicionales que amplían las capacidades de los procesos ETL.



Figura 12
Ventana de Pentaho del componente varios



Nota. En la ilustración se puede observar en la interfaz de usuario del software Pentaho Data Integration la ventana de Design de varios componentes. *Fuente:* Elaboración propia.

Entre los componentes que se puede mencionar constan:

- **Lookup:** Permite buscar y recuperar valores específicos desde un flujo de datos secundario o una base de datos.
- **Streaming:** Facilita el manejo de flujos de datos en tiempo real, integrando y procesando datos dinámicamente.

- **Flow:** Ayuda a controlar el flujo lógico de las transformaciones mediante bifurcaciones y condiciones.
- **Scripting:** Permite incluir scripts personalizados para operaciones avanzadas utilizando lenguajes como JavaScript.
- **Data Warehouse:** Contiene pasos específicos para gestionar estructuras de DW, como esquemas en estrella o copo de nieve.
- **Validation:** Proporciona herramientas para validar datos, asegurando que cumplan con reglas y restricciones definidas.
- **Statics:** Calcula métricas estadísticas como promedios, medianas o distribuciones a partir de los datos.
- **Job:** Facilita la automatización de tareas complejas al coordinar múltiples transformaciones y procesos.
- **Mapping:** Permite reutilizar transformaciones comunes mediante subtransformaciones encapsuladas.
- **Inline:** Proporciona herramientas para crear o modificar datos directamente dentro de los flujos.
- **History:** Permite gestionar datos históricos para auditorías o análisis de cambios en el tiempo.
- **Experimental:** Incluye componentes en fase de prueba para funciones innovadoras o específicas.

Estos componentes adicionales fortalecen las capacidades de Pentaho, haciendo que los flujos ETL sean más flexibles, robustos y adaptables a diversas necesidades empresariales.

Diseño de Transformación

El diseño de transformaciones en Pentaho Data Integration es una etapa clave en la creación de procesos ETL eficientes y efectivos. Una transformación en Pentaho consiste en una serie de pasos conectados que procesan y manipulan datos desde su extracción hasta su preparación para la carga mediante los componentes analizados anteriormente. Durante el

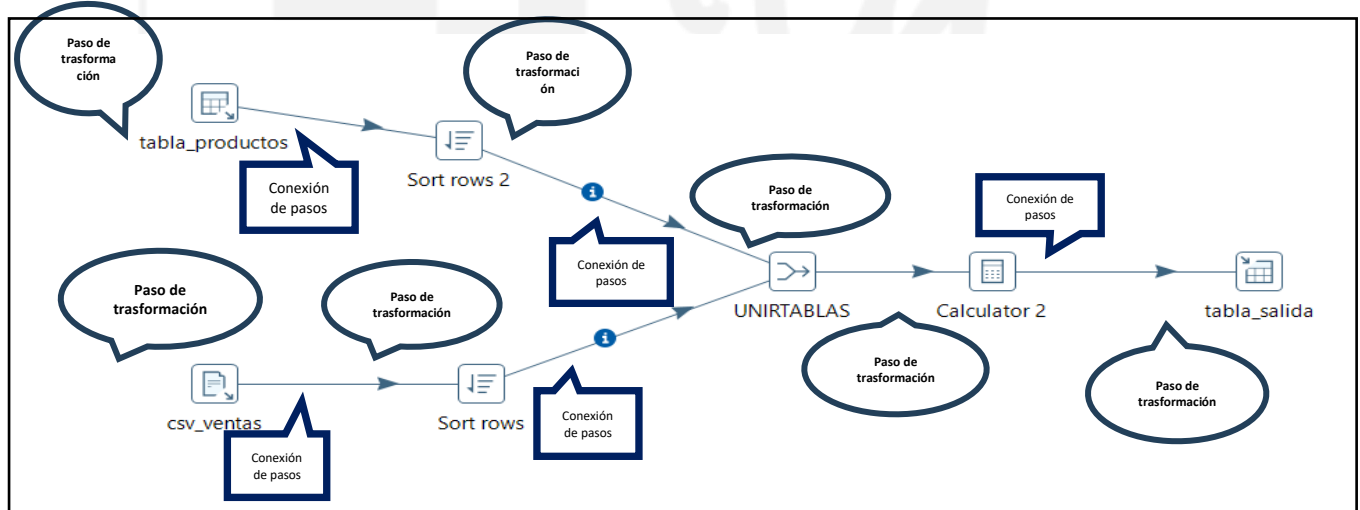
diseño, el usuario puede definir pasos específicos como filtros, cálculos, uniones, o conversiones, y configurarlos para que interactúen de manera lógica dentro del flujo.

La interfaz gráfica de Pentaho, conocida como Spoon, facilita este proceso mediante herramientas de arrastrar y soltar, permitiendo a los desarrolladores visualizar y estructurar cada paso. Adicionalmente, el diseño de transformaciones incluye pruebas y validaciones en tiempo real, lo que asegura que los datos procesados cumplan con los requisitos del negocio antes de ser implementados en un entorno de producción.

En el diseño de la transformación luego de agregar los pasos se debe configurar cada uno de estos, para esto se debe realizar doble clic en cada paso. Y finalmente luego de agregar todos los pasos se deben conectar arrastrando la flecha desde el conector de salida de un paso hasta el conector de entrada de siguiente paso, y de esta manera se define el flujo de datos entre cada paso de la transformación.

En la siguiente ilustración se podrá visualizar los componentes de una transformación de datos formada por varios pasos.

Figura 13
Componentes de una transformación de datos



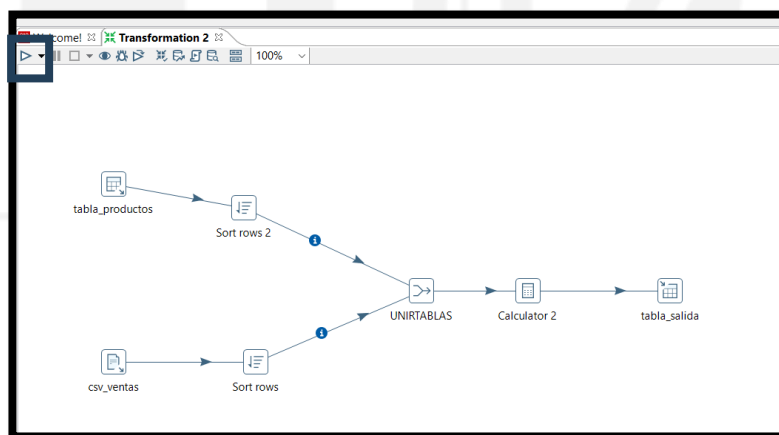
Nota. En la ilustración se puede observar el espacio de trabajo de Pentaho Data Integration donde se puede colocar la serie de pasos para la transformación de datos, iniciando desde la fuente de origen, los pasos de transformación y la data de salida. *Fuente:* Elaboración propia.

Ejecución de la transformación

La ejecución de una transformación en Pentaho Data Integration es el proceso mediante el cual se activan los pasos diseñados en un flujo ETL para procesar datos según las configuraciones establecidas. Una vez creada la transformación en Spoon, el usuario puede ejecutarla directamente desde la interfaz gráfica, lo que permite visualizar en tiempo real el flujo de datos y detectar posibles errores. Durante la ejecución, cada paso procesa los datos entrantes y genera salidas según las instrucciones configuradas, mientras que los indicadores en Spoon proporcionan estadísticas sobre filas procesadas y tiempos de ejecución. Este enfoque dinámico no solo facilita el monitoreo y la depuración, sino que también garantiza la eficiencia y precisión en el procesamiento de datos.

Es decir, una vez que se realiza todo el proceso de transformación se procede con la ejecución de la misma para validar la transformación, para este fin se debe hacer clic en el botón de ejecución del ícono de la flecha hacia la derecha que se encuentra en la barra de herramientas.

Figura 14
Hoja de trabajo de la transformación



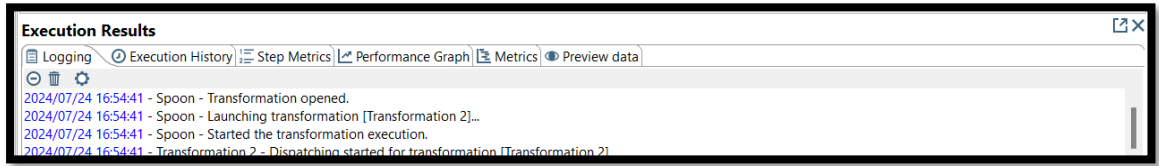
Nota. En la ilustración se puede observar el espacio de trabajo de la transformación donde se colocan la serie de pasos para la transformación de datos y se resalta el icono de ejecución. *Fuente:* Elaboración propia.

Al ejecutar la transformación aparece una ventana de ejecución donde se muestra el progreso y cualquier notificación sea de éxito o error que ocurra durante la ejecución en la parte inferior de donde se diseña la transformación tal como se muestra en la siguiente ilustración:



Figura 15

Ventana de ejecución del Job



Nota. En la ilustración se puede observar la ventana donde se registrar los eventos y procesos ejecutados en la transformación de datos. *Fuente:* Elaboración propia.





Autoevaluación 2

Lea detenidamente la pregunta y marque la respuesta correcta entre las opciones dadas.

1. ¿Qué significa ETL?

- a) Extract, Transform, Load.
- b) Extract, Translate, Load.
- c) Extract, Transport, Load.
- d) Extract, Transform, Link.

2. ¿Qué función tiene un Data Warehouse en la arquitectura de BI?

- a) Guarda datos en tiempo real para procesos transaccionales.
- b) Almacenar datos integrados de múltiples fuentes para análisis rápidos.
- c) Transformar datos para adaptarlos a las necesidades del negocio.
- d) Ejecutar aplicaciones y procesos de mensajería.

3. ¿Qué herramientas se mencionan para la realización de procesos ETL?

- a) Microsoft Excel y Google Sheets.
- b) PostgreSQL y SQL Server.
- c) SAP HANA y MongoDB.
- d) Pentaho y Apache Nifi.

4. ¿Cuál es el propósito de las herramientas de análisis de datos en la arquitectura de BI?

- a) Realizar transacciones financieras y administrativas.
- b) Realizar estudios profundos y encontrar patrones ocultos en los datos.
- c) Almacenar datos en tiempo real.
- d) Ejecutar aplicaciones y procesos de administración empresarial.





5. ¿Qué herramientas se mencionan para la creación de dashboards interactivos?

- a) Power BI, Tableau, y QlikView.
- b) Microsoft Word, Google Docs, y LibreOffice.
- c) Microsoft Excel, Google Docs
- d) Salesforce, HubSpot, y Zoho CRM.

6. ¿Cuál es el propósito de los procesos ETL en la integración de datos?

- a) Ejecutar aplicaciones y procesos de administración empresarial.
- b) Almacenar datos en tiempo real de transacciones organizacionales.
- c) Extraer, transformar y cargar datos en un Data Warehouse
- d) Realizar transacciones financieras y administrativas.

7. ¿Qué es Pentaho Data Integration (PDI)?

- a) Un sistema de gestión de bases de datos.
- b) Una herramienta de ETL para la integración de datos.
- c) Una aplicación para la creación de documentos de BI.
- d) Un software de diseño gráfico estadístico para datos.

8. ¿Qué permite Power BI en términos de visualización de datos?

- a) Ejecutar transacciones financieras y administrativas.
- b) Almacenar datos en tiempo real y realizar cambios de Insight.
- c) Conectar datos en tiempo real y crear visualizaciones interactivas.
- d) Realizar procesos de administración de recursos humanos.





9. ¿Cuál es el primer paso para la creación de un reporte en Power BI?

- a) Conectar la herramienta a las fuentes de datos relevantes.
- b) Crear gráficos y visualizaciones.
- c) Transformar y limpiar los datos.
- d) Publicar el reporte en Power BI Service.

10. ¿Qué son los Jobs en Pentaho Data Integration?

- a) Almacenes de datos centralizados de las transformaciones
- b) Interfaces gráficas para diseñar informes e integraciones
- c) Herramientas para la visualización de datos integrales
- d) Procesos que coordinan la ejecución de transformaciones

Si ha concluido la autoevaluación:

Verifique sus respuestas en el solucionario que se encuentra al final de la presente Guía.



Si alcanzó un porcentaje alto de respuestas correctas en las preguntas de evaluación, es hora de continuar con la Unidad 3, caso contrario, vuelva a revisar los temas en los que ha tenido dificultad.



Resumen de la Unidad 2

La Unidad 2 aborda las herramientas y técnicas clave en los procesos de transformación de datos, enfatizando la implementación de ETLs (Extract, Transform, Load) dentro del contexto de la inteligencia de negocios (BI). Los ETLs son fundamentales para extraer datos de diversas fuentes, transformarlos según los requisitos del negocio y cargarlos en un almacén de datos (Data Warehouse). Este proceso asegura que los datos sean consistentes, relevantes y listos para el análisis estratégico.

Se analiza la herramienta Pentaho Data Integration (PDI), conocida como Kettle, destacada en la implementación de ETLs, ofreciendo una interfaz gráfica intuitiva llamada Spoon. A través de Spoon, los usuarios pueden diseñar, ejecutar y monitorear transformaciones (procesos ETL) y trabajos (coordinación de múltiples tareas). El diseño de transformaciones permite configurar pasos de entrada, transformación y salida, asegurando flexibilidad y eficiencia en la integración de datos. En esta unidad, se destacan componentes específicos de PDI: Input, que define la extracción de datos desde múltiples fuentes (bases de datos, archivos planos, APIs, entre otros); Transform, que realiza operaciones como cálculos, limpieza y estructuración de datos; Joins, para combinar datos dispersos; Output, que distribuye los datos procesados a diversos formatos y sistemas; y Utility, que ofrece herramientas auxiliares para monitorear y automatizar procesos ETL.

En esta unidad también se detalla el proceso de instalación y configuración de PDI, adaptándose a diversos sistemas operativos. Además, se abordan herramientas adicionales como Talend, Apache Nifi y SSIS, ampliando el panorama de opciones para los ETLs.

En conclusión, la Unidad 2 proporciona un enfoque integral sobre las herramientas y técnicas de transformación, destacando la versatilidad de Pentaho y su capacidad para manejar datos complejos de manera eficiente, lo que lo convierte en una solución clave para la gestión empresarial moderna.



¡Continúe así! ¡Le deseo mucho éxito!





UNIDAD 3. (DASHBOARD PARA LA TOMA DE DECISIONES)

“Si no persigues lo que quieres, nunca lo tendrás. Si no vas hacia delante, siempre estarás en el mismo lugar.”

Nora Roberts

Temas y Subtemas

Herramientas y Técnicas de Visualización

Familiarizandose con Power BI

Diseño y Desarrollo de Dashboard

Uso de DAX para Cálculos y Medidas

Publicación y Compartición de Dashboards

Reportes y Visualización de datos

La visualización de datos y la generación de reportes son componentes esenciales en la arquitectura de inteligencia de negocios (BI). Estas herramientas permiten transformar datos brutos en información útil y accionable, facilitando la toma de decisiones informadas. En esta sección, se tratará temas de estudio en Power BI, como una plataforma popular para la visualización de datos y generación de reportes. También se mencionará otras herramientas disponibles en el mercado útiles para la visualización de información.

La visualización de datos implica la representación gráfica de información compleja para hacerla más comprensible y accesible.





Importancia de reportes y Visualización de datos

Los reportes y la visualización de datos incluyen:

- Gráficos y tablas que facilitan a los usuarios entender rápidamente grandes volúmenes de datos, mejorando la comprensión de estos.
- Las visualizaciones pueden mostrar fácilmente patrones y tendencias que podrían no ser evidentes en tablas de datos.
- Las visualizaciones son herramientas clave en la administración de empresas, ya que permiten presentar datos de manera clara y concisa. Estas representaciones gráficas ayudan a los tomadores de decisiones en los mandos gerenciales a analizar información de manera más efectiva, facilitando la comprensión de datos complejos, permitiendo una comunicación efectiva, ayudando a identificar oportunidades y desafíos, y facilitando la toma de decisiones basada en datos reales de la empresa.
- Las visualizaciones de datos sin duda son una herramienta poderosa para comunicar o informar hallazgos y resultados a diferentes audiencias o clientes.

Herramientas y Técnicas de Visualización

Las herramientas y técnicas de visualización de datos permiten transformar grandes volúmenes de información en representaciones gráficas interactivas y comprensibles.

Las herramientas de visualización de datos incluyen una amplia variedad de software diseñados para transformar datos en gráficos, tablas, mapas y paneles interactivos. Algunas de las más utilizadas son:

Power BI: Plataforma de Microsoft que permite crear informes dinámicos y paneles interactivos con integración a diversas fuentes de datos.

Tableau: Herramienta avanzada que ofrece potentes capacidades analíticas y visualizaciones interactivas.

Google Data Studio (Looker Studio): Solución gratuita de Google para la creación de informes basados en datos de diversas fuentes.



Excel: Aunque es más tradicional, permite crear gráficos dinámicos, tablas y reportes con análisis de datos.

Qlik Sense: Plataforma de visualización y análisis de datos con capacidades de inteligencia artificial.

D3.js: Biblioteca de JavaScript para la creación de visualizaciones personalizadas y altamente interactivas en la web.

Matplotlib y Seaborn: Librerías de Python utilizadas en entornos científicos y de análisis de datos para generar gráficos detallados.

En esta materia se enfocará en la herramienta de Microsoft Power BI.

Power BI

Power BI útil para el análisis de negocios creada por Microsoft que permite a los usuarios visualizar sus datos y compartir información en toda la organización además de transformar datos de manera general. Con Power BI, los usuarios pueden conectarse a múltiples fuentes de datos, transformarlas, y crear dashboards y reportes interactivos.

Principios Esenciales de Power BI

Esta herramienta es usada en la inteligencia de negocios que permite analizar, visualizar y compartir datos de manera eficiente. Su interfaz intuitiva facilita la creación de paneles interactivos y reportes personalizados sin necesidad de conocimientos avanzados en programación. Para comenzar a usar Power BI, es fundamental entender sus principales componentes: Power BI Desktop, donde se diseñan y modelan los informes; Power BI Service, una plataforma en la nube para compartir y colaborar en proyectos; y Power BI Mobile, que permite acceder a los reportes desde dispositivos móviles.

Características Principales de Power BI

- **Conectividad de Datos:** Power BI puede conectarse a una amplia variedad de fuentes de datos, incluyendo bases de datos, archivos planos, servicios en la nube y aplicaciones SaaS, ya que cuenta con funcionalidades incorporadas que únicamente se debe escoger **que tipo de fuentes se usa.**
- **Transformación y Limpieza de Datos:** La funcionalidad de Power Query en Power BI

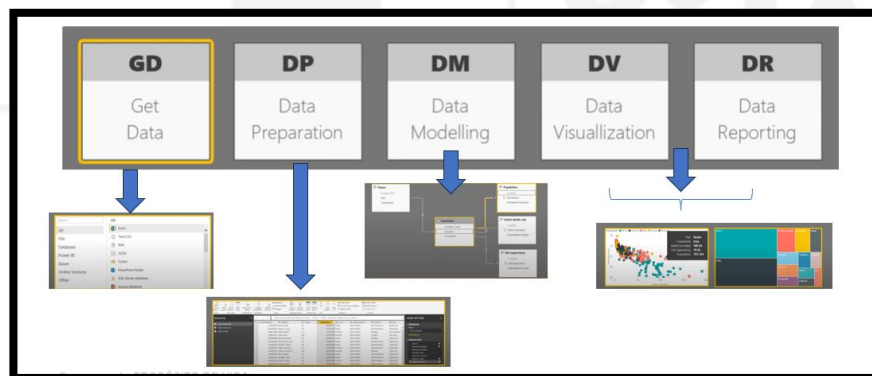


permite a los usuarios transformar y limpiar datos antes de visualizarlos, aunque no es tan potente como Pentaho tiene funcionalidades básicas para realizar transformaciones de datos.

- **Modelado de Datos:** Power BI permite crear modelos de datos que integran múltiples tablas y establecen relaciones entre ellas, lo que facilita el trabajar con tablas de diferentes fuentes.
- **Visualizaciones Interactivas:** Power BI ofrece una amplia gama de visualizaciones interactivas que pueden personalizarse para satisfacer las necesidades del usuario.
- **Dashboards:** Los usuarios pueden crear dashboards personalizados que combinan múltiples visualizaciones y métricas clave en una sola vista.
- **Actualización de Datos en Tiempo Real:** Power BI puede mostrar datos en tiempo real, lo que permite a los usuarios monitorear el rendimiento y tomar decisiones informadas rápidamente.

Bloques de Power BI

Figura 16
Bloques de Power BI



Nota. En la ilustración se puede observar el proceso estructurado de gestión de datos dividido en varias etapas que presenta Power BI. *Fuente:* Elaboración propia.

Power BI es una herramienta integral de análisis y visualización de datos que permite a los usuarios conectar, transformar y visualizar datos provenientes de diversas fuentes. Su arquitectura se puede dividir en varios bloques fundamentales, cada uno desempeñando un

papel crucial en el proceso de convertir datos brutos en información útil y accionable. A continuación, se describen los bloques principales de Power BI:

Obtener Datos (Get Data - GD)

El primer paso en cualquier proyecto de Power BI es la obtención de datos. Power BI ofrece una amplia variedad de conectores que permiten acceder a datos desde múltiples fuentes, como bases de datos, archivos de Excel, servicios en la nube y APIs.

Funcionalidades Clave:

- Conexión a diferentes fuentes de datos.
- Importación de datos desde archivos locales o fuentes en la nube.
- Posibilidad de combinar datos de diferentes orígenes.

Preparación de Datos (Data Preparation - DP)

Una vez que los datos han sido obtenidos, es necesario prepararlos para el análisis. Esto implica limpiar, transformar y estructurar los datos para que sean adecuados para el modelado y la visualización.

Funcionalidades Clave:

- Limpieza de datos, incluyendo la eliminación de duplicados y la gestión de valores nulos.
- Transformación de datos, como la modificación de tipos de datos y la creación de nuevas columnas.
- Integración y combinación de datos de diferentes fuentes.

Modelado de Datos (Data Modelling - DM)

El modelado de datos implica estructurar los datos en un modelo relacional que facilite el análisis. En esta etapa, se definen relaciones entre diferentes tablas y se crean medidas y columnas calculadas utilizando DAX (Data Analysis Expressions).

Funcionalidades Clave:

- Creación de relaciones entre tablas.



- Definición de medidas y columnas calculadas.
- Optimización del modelo de datos para mejorar el rendimiento de las consultas.

Visualización de Datos (Data Visualization - DV)

La visualización de datos es donde los datos se convierten en gráficos interactivos y visualizaciones que permiten a los usuarios explorar y analizar la información. Power BI ofrece una amplia gama de visualizaciones, desde gráficos de barras y líneas hasta mapas y gráficos de dispersión.

Funcionalidades Clave:

- Creación de gráficos y visualizaciones interactivas.
- Personalización de visualizaciones con colores, etiquetas y formatos.
- Uso de filtros y segmentaciones para explorar los datos desde diferentes perspectivas.

Para la creación de reportes en Power BI primeramente se debe tener una conexión a fuentes de datos, este primer paso en la creación de un reporte en Power BI es conectar la herramienta a las fuentes de datos relevantes. Power BI soporta una amplia gama de conectores, lo que facilita la integración de datos de diferentes orígenes sean estos archivos de Excel, csv, texto, o bases de datos.

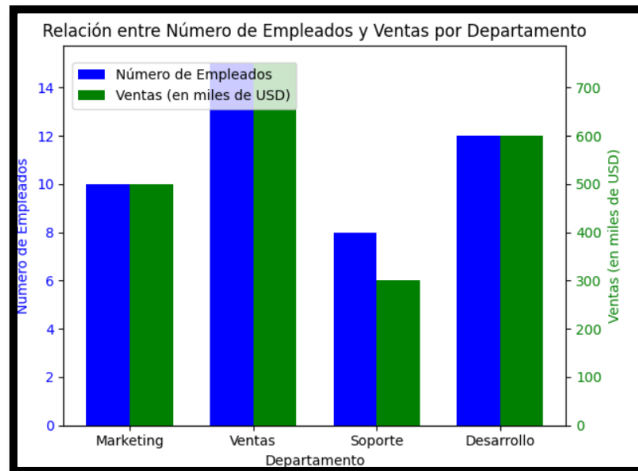
Una vez que los datos están conectados, se procede a su transformación y modelado. Power BI permite limpiar, transformar y organizar los datos utilizando el editor de Power Query. Este paso es crucial para asegurar que los datos sean precisos y estén en un formato adecuado para el análisis.

Y finalmente con los datos transformados y modelados, se pueden crear visualizaciones. Power BI ofrece una variedad de gráficos, tablas y otros elementos visuales que pueden personalizarse. Algunos tipos comunes de visualizaciones incluyen:

- **Gráficos de Barras y columnas:** Este tipo de gráficos son útiles para comparar cantidades entre diferentes categorías.



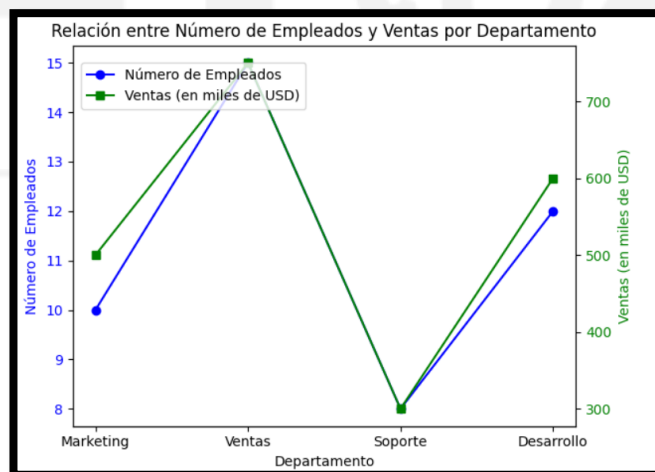
Figura 17
Gráfico de barras



Nota. En la ilustración se puede observar la relación entre el número de empleados y las ventas en miles de USD para los diferentes departamentos de una empresa. *Fuente:* Elaboración propia.

- **Gráficos de Líneas:** Ideales para mostrar tendencias a lo largo del tiempo.

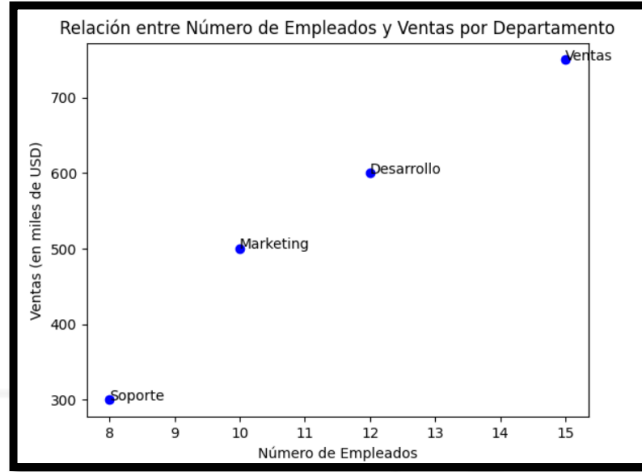
Figura 18
Gráfico de líneas



Nota. En la ilustración se puede observar la relación entre el número de empleados y las ventas en miles de USD por departamento dentro de una empresa. *Fuente:* Elaboración propia.

- **Gráficos de Dispersión:** Ayudan a identificar relaciones entre dos variables.

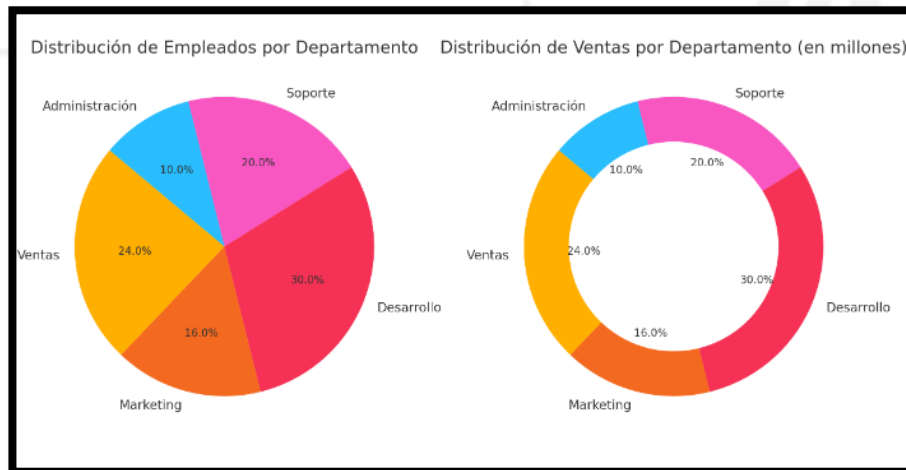
Figura 19
Gráfico de dispersión



Nota. En la ilustración se puede visualizar rápidamente cómo varían las ventas en relación con el tamaño del equipo en cada departamento. *Fuente:* Elaboración propia.

- **Gráficos circulares y de anillos:** Son muy útiles y eficaces para representar proporciones y porcentajes

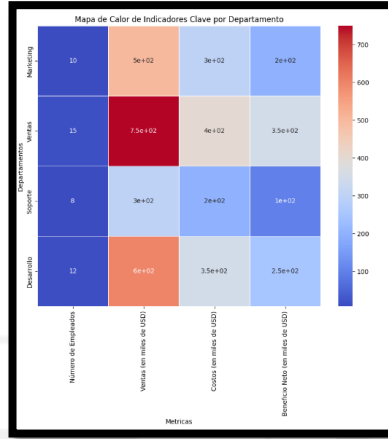
Figura 20
Gráfico de pastel y anillo



Nota. En la ilustración se puede visualizar dos gráficos circulares que comparan la distribución de empleados y de ventas por departamento dentro de una organización. *Fuente:* Elaboración propia.

- **Mapas de Calor:** Representan datos en una matriz y utilizan colores para indicar la intensidad de los valores.

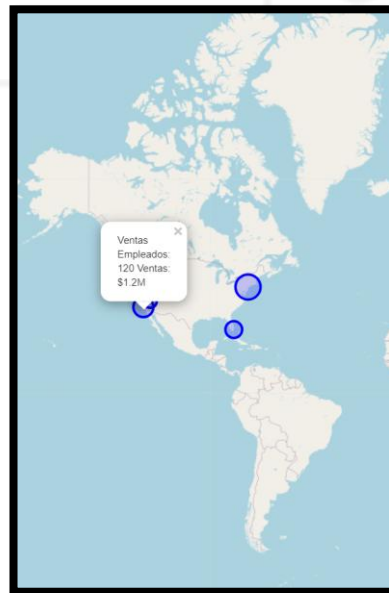
Figura 21
Mapa de calor o matriz de correlación



Nota. En la ilustración se puede visualizar diferentes indicadores clave por departamento. *Fuente:* Elaboración propia.

- **Mapas:** Los mapas permiten la visualización de datos geográficos.

Figura 22
Gráfico tipo mapa



Nota. En la ilustración se puede visualizar la distribución geográfica de ciertas métricas de negocio en América del Norte y Central. *Fuente:* Elaboración propia.

- **Tablas y matrices:** Este tipo de visualizaciones son útiles para mostrar datos detallados en formato tabular.

Tabla 2:

Visualización tipo Tabla Pivot

Departamento	Empleados	Ventas (millones USD)
Administración	50	0.5
Desarrollo	150	1.5
Marketing	80	0.8
Soporte	100	1.0
Ventas	120	1.2

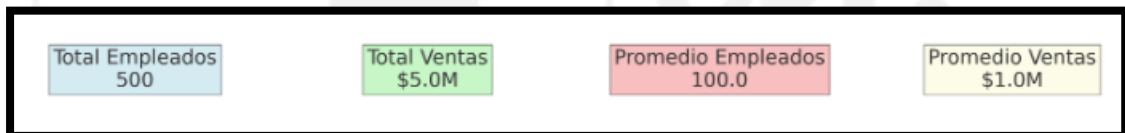
Nota. Tabla pivote que resume datos sobre empleados y ventas por departamento dentro de una organización.

Fuente: Elaboración propia.

- **Tarjetas y KPI:** Estos recursos son utilizadas para mostrar métricas clave de rendimiento.

Figura 23

Tarjetas en dashboards en Power BI



Nota. En la ilustración se puede visualizar tarjetas que permiten que el dashboard sea claro donde se puede mostrar los resúmenes y las KPI de la data. *Fuente:* Elaboración propia.

Informes de Datos (Data Reporting - DR)

El último bloque implica la creación de informes que consolidan las visualizaciones y análisis en un formato coherente y accesible. Estos informes se pueden compartir con otros usuarios a través de Power BI Service, incrustar en aplicaciones o sitios web, o exportar a otros formatos.

Funcionalidades Clave:

- Creación de informes y dashboards interactivos.

- Publicación y compartición de informes en Power BI Service.
- Integración de informes en aplicaciones y sitios web.

Power BI se presenta como una solución poderosa y versátil para el análisis y la visualización de datos. Su capacidad para conectar diversas fuentes de información, junto con sus herramientas de transformación y visualización, permite a las organizaciones tomar decisiones informadas basadas en datos. Al comprender y utilizar estos bloques fundamentales, los usuarios pueden maximizar el potencial de Power BI y transformar datos complejos en insights valiosos.

Diseño y Desarrollo de Dashboard

Power BI se ha convertido en una herramienta fundamental para las organizaciones que desean maximizar el uso de sus datos, gracias a su facilidad de uso y potentes capacidades de análisis y visualización. Esta plataforma permite a analistas de datos y profesionales de negocios crear dashboards interactivos de manera sencilla, promoviendo así un acceso democratizado a la información. Al transformar la forma en que las empresas visualizan y utilizan sus datos, Power BI se posiciona como un recurso estratégico que impulsa la eficiencia, la innovación y el crecimiento.

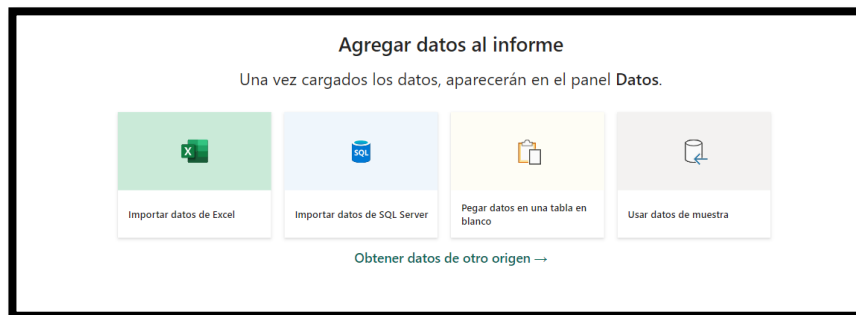
El diseño y desarrollo de dashboards en Power BI es un viaje que une la ciencia de datos con el arte del diseño. Un buen dashboard no solo muestra información de forma atractiva, sino que también se convierte en una herramienta clave para facilitar decisiones informadas. Power BI empodera a los profesionales, permitiéndoles crear visualizaciones impactantes que transforman la forma en que las organizaciones analizan y aprovechan sus datos, convirtiendo la información en un aliado estratégico para el éxito.

El diseño de dashboards en Power BI es un proceso que permite a los usuarios transformar datos en visualizaciones efectivas y comprensibles. A continuación, se presentan los pasos fundamentales para crear un dashboard exitoso:

- **Conectar a la fuente de datos:** El primer paso consiste en seleccionar Obtener datos y elegir la fuente de datos adecuada. Esto establece la base sobre la cual se construirán las visualizaciones.



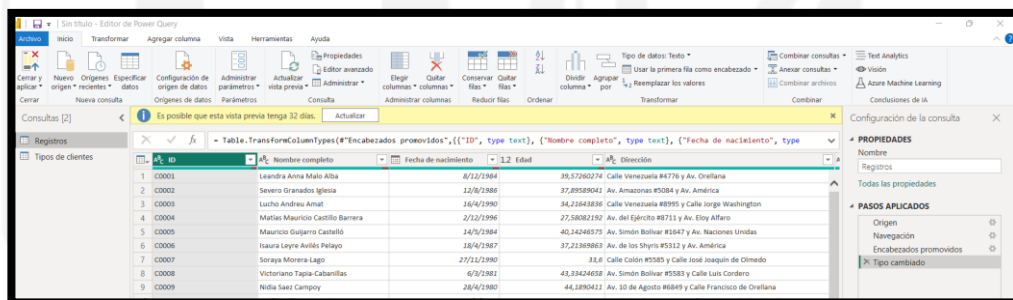
Figura 24
Pantalla inicial de conectar a fuente de datos



Nota. En la ilustración se puede visualizar las diferentes fuentes a las que se puede conectar Power BI. *Fuente:* Elaboración propia.

- **Transformar y modelar los datos:** Para este paso se utiliza el Editor de consultas o Editor de Power Query que sirve para limpiar y preparar los datos. Este proceso es crucial para asegurar que la información sea precisa y relevante para el análisis.

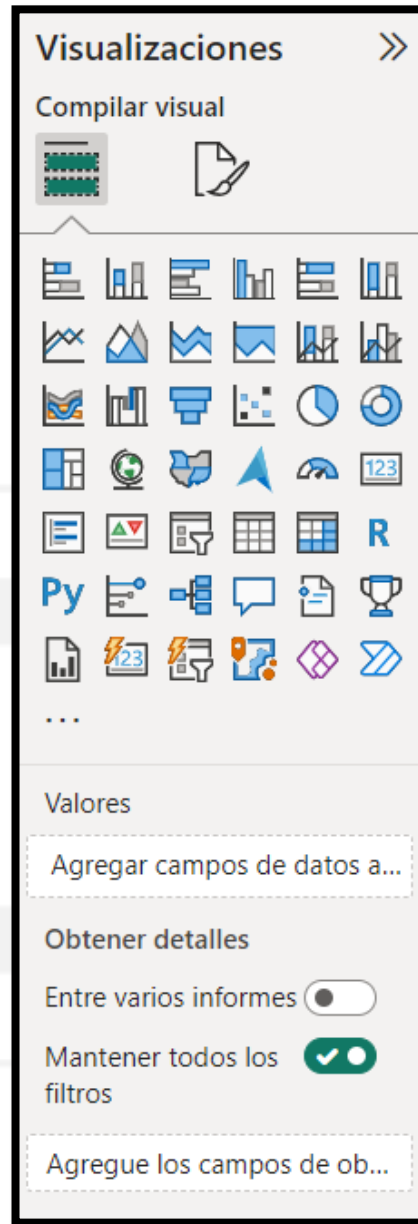
Figura 25
Pantalla de Power Query



Nota. En la ilustración se puede visualizar la pantalla principal del editor de consultas donde se puede realizar las transformaciones a la data. *Fuente:* Elaboración propia.

- **Seleccionar visualización:** En la interfaz de Power BI, elige el tipo de visualización que mejor se adapte a tus necesidades desde el panel de visualizaciones. La selección adecuada de visualizaciones es clave para comunicar la información de manera efectiva.

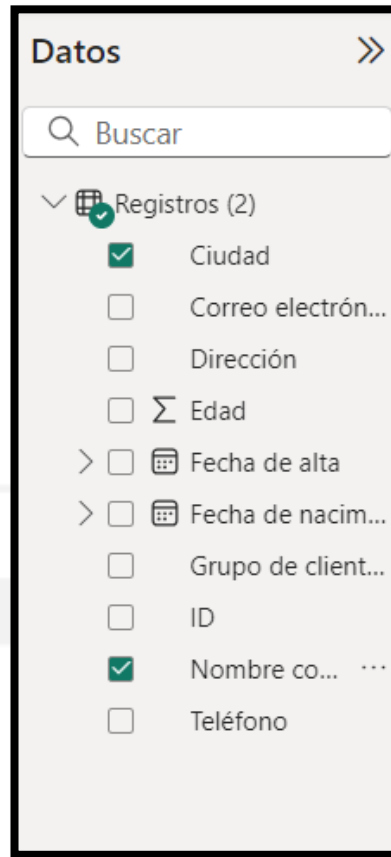
Figura 26
Ventana de visualizaciones



Nota. En la ilustración se puede observar la ventana de visualizaciones donde se puede escoger el tipo de gráfico que se puede agregar al dashboard. *Fuente:* Elaboración propia.

- **Agregar campos:** Arrastra y suelta los campos desde el panel de campos a los diferentes roles de la visualización, como ejes, valores y leyendas. Esto permite que los datos se representen de manera clara y estructurada.

Figura 27
Ventana de Datos

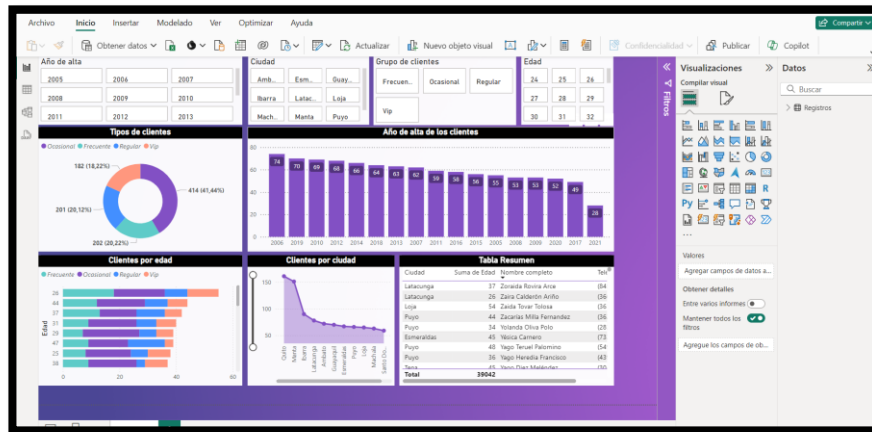


Nota. En la ilustración se observa la ventana de datos, desde la cual es posible seleccionar los campos que alimentan el gráfico del dashboard. *Fuente:* Elaboración propia.

Siguiendo estos pasos, se podrá crear dashboards en Power BI que no solo sean visualmente atractivos, sino que también faciliten la toma de decisiones informadas en la organización.

La personalización de las visualizaciones en Power BI es fundamental para mejorar la interpretación y presentación de los datos. A continuación, se describen los pasos necesarios para personalizar una visualización de manera efectiva:

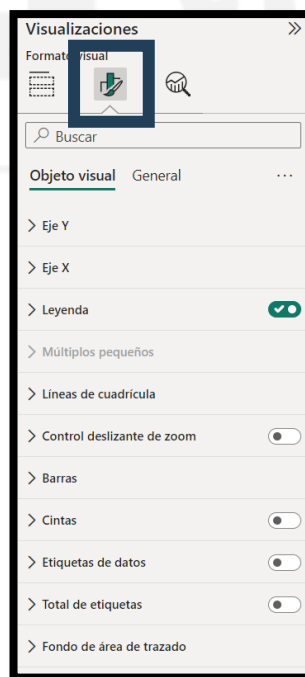
Figura 28
Dashboard personalizado



Nota. En la ilustración se puede visualizar un dashboard personalizado con varios componentes entre ellos tarjetas, tablas y gráficos. *Fuente:* Elaboración propia.

- **Acceda al panel de formato:** Se debe dar clic en el icono de formato representado por un pincel en el panel de visualizaciones. Este panel le permitirá acceder a diversas opciones de personalización tal como se muestra en la siguiente ilustración.

Figura 29
Ventana de Panel de formato



Nota. En la ilustración se puede visualizar la ventana que permite dar formato a cualquier tipo de gráfico en el dashboard. *Fuente:* Elaboración propia.

- **Ajuste los colores:** Permite cambiar los colores de las barras, líneas y otros elementos seleccionando la opción de Colores de datos. La elección de colores adecuados puede hacer que la visualización sea más atractiva y fácil de interpretar, esta opción la encuentra en la personalización de cada grafico tal como muestra la ilustración 20.
- **Modifique las etiquetas:** Se utiliza la sección de Etiquetas de datos para activar, desactivar o cambiar el estilo de las etiquetas. Las etiquetas claras y bien diseñadas son cruciales para la comprensión de los datos presentados.

Figura 30

Ventana de Etiquetas de datos



Nota. En la ilustración se puede visualizar la ventana que permite agregar o quitar las etiquetas del grafico en el dashboard. etiquetas como títulos, valores, porcentajes, entre otros. *Fuente:* Elaboración propia.

- **Configure los títulos y leyendas:** Ajuste los títulos, subtítulos y leyendas desde las opciones correspondientes en el panel de formato. Títulos descriptivos y leyendas claras ayudan a contextualizar la información. También se puede agregar otras opciones de personalización, como sombras, bordes y transparencia, para mejorar el diseño visual. Estos detalles pueden hacer que el dashboard sea más profesional y atractivo, tal como muestra la Ilustración 20.

Uso de DAX para Cálculos y Medidas

DAX (Data Analysis Expressions) es un lenguaje de fórmulas utilizado en Power BI, Power Pivot y Analysis Services para realizar cálculos y crear medidas y columnas calculadas. DAX es una herramienta poderosa que permite a los usuarios realizar análisis de datos avanzados y obtener información valiosa a partir de sus datos. Con DAX, es posible crear cálculos complejos y personalizar los análisis según las necesidades específicas del negocio.

Introducción a DAX

DAX es un lenguaje funcional que se utiliza para definir cálculos personalizados en Power BI. Su sintaxis es similar a la de Excel, pero DAX está diseñado específicamente para trabajar con datos relacionales y modelos de datos. Los elementos básicos de DAX incluyen columnas calculadas, medidas y tablas calculadas, que se utilizan para realizar diversos tipos de cálculos.

Columnas Calculadas

Las columnas calculadas se crean en la tabla de datos y se utilizan para agregar nuevas columnas con valores calculados a partir de otras columnas en la misma tabla. Las fórmulas DAX para columnas calculadas son similares a las fórmulas de Excel, y se recalculan automáticamente cada vez que los datos de la tabla cambian.

Ejemplo de columna calculada:

$$\text{TotalCostos} = [\text{PrecioUnitario}] * [\text{Cantidad}]$$

Medidas

Las medidas son cálculos que se realizan sobre los datos agregados, y se utilizan para obtener resultados dinámicos basados en el contexto de filtrado y agrupación en los informes y visualizaciones. A diferencia de las columnas calculadas, las medidas no se almacenan en la tabla de datos, sino que se calculan en tiempo de consulta.

Ejemplo de medida:

$$\text{TotalVentas} = \text{SUM}([\text{VentasTotales}])$$

Funciones DAX

DAX incluye una amplia variedad de funciones que se pueden utilizar para realizar cálculos y transformaciones en los datos. Estas funciones se dividen en varias categorías, incluyendo funciones de agregación, funciones lógicas, funciones de texto, funciones de fecha y hora, y funciones de inteligencia de tiempo.

Ejemplo de funciones DAX:

CALCULATE, FILTER, SUMX, AVERAGE, RELATED

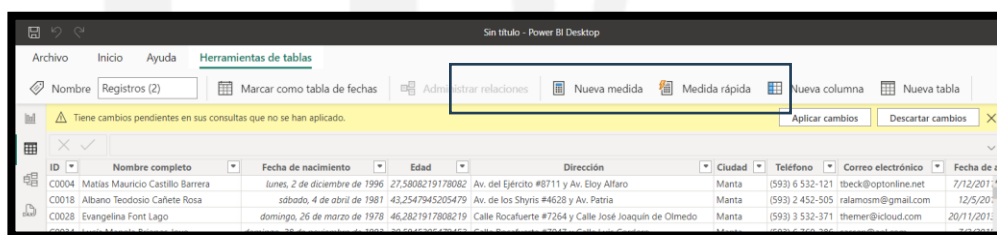


La creación de medidas y columnas calculadas en Power BI es un proceso esencial para el análisis de datos, ya que permite realizar cálculos personalizados que enriquecen la información presentada. A continuación, se describen los pasos para crear medidas y columnas calculadas:

Se debe cargar los datos en el modelo de datos. Este paso es fundamental para comenzar a trabajar con su información.

- **Seleccionar la tabla:** Se elige la tabla en la que se desea crear la columna calculada o medida. Se debe asegurar que la tabla contenga los datos necesarios para sus cálculos.
- **Acceder a la pestaña Modelado:** En la pestaña Modelado, se debe hacer clic en Nueva columna si se desea crear una columna calculada, o en Nueva medida si se desea crear una medida. Esta opción le permitirá iniciar el proceso de creación.

Figura 31
Pantalla de modelado de datos



Nota. En la ilustración se puede visualizar la pantalla del modelado de Power Bi donde se puede agregar columnas o medida a la data existente. *Fuente:* Elaboración propia.

- **Ingrese la fórmula DAX:** Se debe escribir la fórmula DAX correspondiente en la barra de fórmulas y presionar Enter. DAX (Data Analysis Expressions) es el lenguaje utilizado para realizar cálculos en Power BI.
- **Utilice las nuevas columnas y medidas:** Una vez creadas, las nuevas columnas calculadas y medidas, estas estarán disponibles para su uso en visualizaciones y análisis.

Esto permite enriquecer los informes con cálculos específicos que se alineen con las necesidades analíticas.

El uso de DAX en Power BI permite realizar análisis de datos avanzados y obtener información valiosa a partir de los datos. Mediante el uso de columnas calculadas, medidas y funciones DAX, los usuarios pueden crear cálculos personalizados y adaptar los análisis a las

necesidades específicas del negocio. Dominar DAX es esencial para aprovechar al máximo las capacidades de Power BI y crear dashboards interactivos y potentes.

Publicación y Compartición de Dashboards

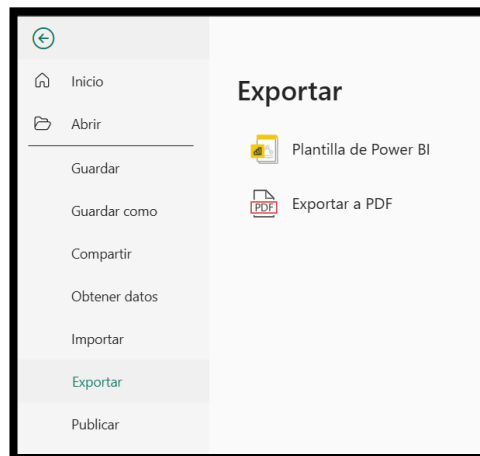
La publicación y el compartir dashboards en Power BI es un paso fundamental para garantizar que las visualizaciones e informes lleguen a las personas correctas en el momento preciso. Power BI ofrece múltiples opciones para compartir dashboards, lo que facilita la colaboración y la toma de decisiones basadas en datos dentro de una organización. Desde la publicación en Power BI Service hasta la integración en aplicaciones empresariales y sitios web, las funcionalidades de compartición de Power BI aseguran que los usuarios tengan acceso a la información que necesitan de manera segura y eficiente. Esto no solo mejora la comunicación entre equipos, sino que también potencia la capacidad de respuesta ante desafíos y oportunidades en el entorno empresarial.

Exportar Dashboards

Se puede exportar los dashboards generados en Power BI a diferentes formatos con el fin de facilitar su reutilización y distribución. Al exportar un dashboard como archivo plantilla (.PBIX), se guarda la estructura del informe, incluyendo las visualizaciones, medidas y relaciones entre datos, pero sin los datos originales, permitiendo que otros usuarios lo utilicen con diferentes fuentes de información.

Por otro lado, la exportación a PDF genera una versión estática del informe, ideal para presentaciones o compartir análisis con personas que no tienen acceso a Power BI, asegurando que la información visualizada sea exactamente la misma sin riesgo de modificaciones.

Figura 32
Pantalla para exportar Dashboards



Nota. En la ilustración se puede visualizar la opción de Exportar que se encuentra en el menú de Inicio. *Fuente:* Elaboración propia.

Publicación de Dashboards en Power BI Service

Power BI Service es una plataforma en línea que facilita a los usuarios compartir sus dashboards y reportes con otros miembros de la organización, promoviendo la colaboración y el acceso a información crítica. A continuación, se describen los pasos necesarios para publicar un dashboard en Power BI Service:

- **Acceder a Power BI Desktop:** Se debe iniciar Power BI Desktop y asegurarse de que el dashboard esté listo para ser compartido.
- **Publicar el dashboard:** Se debe hacer clic en Publicar que se encuentra en la barra de herramientas. Este botón inicia el proceso de publicación del dashboard en la plataforma en línea.
- **Seleccionar la cuenta de Power BI Service:** En el cuadro de diálogo que aparece se puede seleccionar la cuenta de Power BI Service en la que se desea publicar el dashboard. Se debe elegir la cuenta correcta para asegurarse de que el dashboard se comparta con los usuarios adecuados.
- **Confirmación de publicación:** Una vez que se completa el proceso de publicación, el dashboard estará disponible en Power BI Service. Desde allí, otros usuarios de la organización podrán acceder al dashboard para verlo y analizarlo.



Configuración de Permisos y Seguridad

Una vez publicado el dashboard, es importante configurar los permisos adecuados para asegurarse de que solo los usuarios autorizados tengan acceso. Power BI permite configurar permisos a nivel de informe y dataset, asegurando que la información sensible esté protegida. Puede compartir dashboards con usuarios específicos, grupos de seguridad, o incluso con toda la organización, dependiendo de las necesidades del negocio.

Compartición de Dashboards con Usuarios Internos

Compartir un dashboard con usuarios internos de la organización es un proceso sencillo que permite la colaboración y el acceso a información relevante. A continuación, se detallan los pasos para compartir un dashboard en Power BI Service:

- **Abrir el dashboard:** Se debe iniciar sesión en Power BI Service y acceder al dashboard que se desea compartir.
- **Haga clic en Compartir:** En la parte superior derecha del dashboard, se encuentra el botón Compartir. Se debe dar clic en este botón para iniciar el proceso de compartición.
- **Ingresar las direcciones de correo electrónico:** En el cuadro de diálogo que aparece se debe ingresar las direcciones de correo electrónico de los usuarios con los que se desea compartir el dashboard. Se debe asegurar de incluir a todos los colaboradores relevantes.
- **Configurar los permisos:** Antes de enviar las invitaciones, se debe ajustar los permisos de visualización y edición según sea necesario. Esto permite controlar quién puede ver o editar el dashboard compartido.
- **Haga clic en Compartir:** Finalmente, se debe hacer clic en el botón Compartir para enviar las invitaciones a los usuarios seleccionados. Los usuarios recibirán un correo electrónico con el enlace al dashboard, permitiéndoles acceder a la información compartida.

Integración de Dashboards en Aplicaciones y Sitios Web

Power BI ofrece capacidades de integración que permiten incrustar dashboards en aplicaciones empresariales y sitios web, facilitando así el acceso a la información dentro de los



sistemas que los usuarios ya están utilizando. A continuación, se describen los pasos para integrar un dashboard en una aplicación o sitio web:

- **Abrir el dashboard en Power BI Service:** Se debe iniciar sesión en Power BI Service y acceder al dashboard que se desea integrar en la aplicación o sitio web.
- **Publicar en la web:** Se debe dar clic en el botón Archivo en la parte superior izquierda y se debe seleccionar la opción Publicar en la web. Esta opción permite generar un código de incrustación para el dashboard.
- **Copiar el código de incrustación:** Una vez que se haya seleccionado Publicar en la web, se proporciona un código de incrustación, este código se debe copiar en el sitio web, ya que será necesario para la integración.
- **Pegar el código en el HTML:** Se debe insertar el código de incrustación en el HTML de la aplicación o sitio web, en el lugar donde se desea que aparezca el dashboard. Se debe asegurar de que el entorno de la aplicación o sitio web soporte la incrustación de contenido.

Uso de Aplicaciones Móviles de Power BI

Power BI también proporciona aplicaciones móviles para iOS y Android, lo que permite a los usuarios acceder a los dashboards desde cualquier lugar. Las aplicaciones móviles de Power BI ofrecen una experiencia de usuario optimizada para dispositivos móviles, permitiendo la interacción con los dashboards y la visualización de datos en tiempo real. Para utilizar las aplicaciones móviles, simplemente se la debe descargar desde la tienda de aplicaciones correspondiente e iniciar sesión con la cuenta de Power BI.





Autoevaluación 3

Lea detenidamente la pregunta y marque la respuesta correcta entre las opciones dadas.

1. ¿Cuál es el primer paso en el diseño y desarrollo de dashboards en Power BI?

- a) Agregar campos.
- b) Seleccionar visualización.
- c) Conectar a la fuente de datos.
- d) Configurar títulos y leyendas.

2. ¿Qué permite realizar Power BI mediante el uso de DAX?

- a) Crear presentaciones en PowerPoint
- b) Calcular medidas y columnas calculadas
- c) Enviar correos automáticos
- d) Programar reuniones

3. ¿Cómo se agregan los campos a las visualizaciones en Power BI?

- a) Copiando y pegando
- b) Arrastrando y soltando
- c) Escribiendo código DAX
- d) Usando comandos de voz

4. ¿Qué es Power BI?

- a) Un servicio de email empresarial
- b) Un programa de procesamiento de textos
- c) Un software de control de versiones de datos.
- d) Una herramienta de análisis y visualización de datos.





5. ¿Qué tipo de visualizaciones se pueden crear en Power BI?

- a) Gráficos, mapas y tablas
- b) Documentos de textos
- c) Presentaciones de diapositivas
- d) Formularios tabulados

6. ¿Qué es un dashboard en Power BI?

- a) Una colección de datos sin procesar
- b) Un conjunto de herramientas de programación
- c) Una pantalla que muestra visualizaciones de datos
- d) Un tipo de archivo de texto

7. ¿Qué tipo de software es Power BI?

- a) Herramienta de inteligencia de negocios
- b) Software de desarrollo web
- c) Software antivirus
- d) Plataforma de redes sociales

8. ¿Cuál es un beneficio clave de usar Power BI en una empresa?

- a) Aumenta el consumo de recursos de la empresa
- b) Reduce la capacidad de almacenamiento de datos
- c) Mejora la toma de decisiones mediante el análisis de datos
- d) Mejora la visibilidad de marca

9. ¿Qué es una visualización en Power BI?

- a) Un documento de texto descriptivo





- b) Una representación gráfica de datos
- c) Un video explicativo
- d) Un organigrama

10. ¿Qué función tiene el Editor de consultas (Power Query) en Power BI?

- a) Conectar y transformar datos de diversas fuentes.
- b) Crear presentaciones y diapositivas de BI.
- c) Desarrollar aplicaciones móviles empresariales.
- d) Gestionar las comunicaciones institucionales.

Si ha concluido la autoevaluación:

Verifique sus respuestas en el solucionario que se encuentra al final de la presente Guía.



Si alcanzó un porcentaje alto de respuestas correctas en las preguntas de evaluación, ha finalizado este módulo de Inteligencia de Negocios, caso contrario, vuelva a revisar los temas en los que ha tenido dificultad.



Resumen de la Unidad 3

En la Unidad 3 se aborda la Visualización de Datos a través de Dashboards, centrándose en cómo Power BI se ha vuelto esencial para las organizaciones que desean aprovechar al máximo sus datos. Esta herramienta permite a los profesionales crear dashboards interactivos y visualizaciones efectivas, lo que facilita la toma de decisiones informadas.

El diseño de dashboards en Power BI combina la ciencia de datos con el arte del diseño. Primero, se debe conectar a la fuente de datos seleccionando 'Obtener datos' y eligiendo la fuente adecuada. Luego, es importante transformar y modelar los datos utilizando el 'Editor de consultas' para garantizar que sean precisos y relevantes. Después, se elige el tipo de visualización que mejor se ajuste a las necesidades desde el panel correspondiente y se arrastran y sueltan los campos necesarios en los roles de visualización, como ejes, valores y leyendas.

La personalización de las visualizaciones es clave para mejorar la interpretación de los datos. Esto incluye acceder al panel de formato, ajustar colores de elementos visuales y modificar etiquetas según sea necesario. También es importante configurar títulos y leyendas para que la visualización sea clara y efectiva.

DAX, o Data Analysis Expressions, es un lenguaje de fórmulas que permite realizar cálculos avanzados en Power BI. A través de columnas calculadas y medidas, DAX ayuda a personalizar los análisis según las necesidades del negocio, permitiendo crear cálculos complejos que enriquecen la información presentada en los dashboards.

Finalmente, la publicación y compartición de los dashboards son cruciales para que la información llegue a las personas adecuadas. Power BI Service facilita compartir estos dashboards dentro de la organización, configurando permisos y asegurando la protección de la información sensible. Además, ofrece la posibilidad de integrar dashboards en aplicaciones empresariales y sitios web, así como acceso a través de aplicaciones móviles.



¡Sigue brillando! ¡Le deseo lo mejor en sus futuros proyectos!



9. REFERENCIAS

- Chaudhuri, S. D. (08 de 2011). An overview of business intelligence technology. *CommuniCations of the aCm*. doi:10.1145/1978542.1978562
- Chen, H. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. doi:<https://doi.org/10.2307/41703503>
- Golfarelli, M. (2009). Data warehouse design : modern principles and methodologies. New York : McGraw-Hill. Obtenido de <https://archive.org/details/datawarehousedes0000golf>
- Inmon, W. H. (2008). *DW 2.0: The architecture for the next generation of data warehousing*. Morgan Kaufmann Publishers. Obtenido de <https://pdfs.semanticscholar.org/f711/a63ce446e38e52c66d3a82583d29a4a0db53.pdf>
- Loshin, D. (2012). Business intelligence: The savvy manager's guide. doi:<https://doi.org/10.1016/B978-1-55860-916-7.X5000-3>
- Negash, S. (2004). Business Intelligence. doi:10.17705/1CAIS.01315
- Ranjan, J. (2009). Business intelligence: Concepts, components, techniques and benefits. *Journal of Theoretical and Applied Information Technology*. Obtenido de <https://www.jatit.org/volumes/Vol9No1/9Vol9No1.pdf>
- Rud, O. P. (2009). *Business Intelligence Success Factors: Tools for Aligning Your Business in the Global Economy*. John Wiley & Sons. Obtenido de <https://fliphtml5.com/lnym/wvkj/basic>
- Sauter, V. (2011). *Decision Support Systems for Business Intelligence*. John Wiley & Sons, Inc. . doi:10.1002/9780470634431
- Turban, E. S. (2011). *Decision support and business intelligence systems (9ª ed.)*. Pearson Education International. Obtenido de <https://es.scribd.com/document/355912027/Decision-Support-and-Business-Intelligence-Systems-9th-Ed-Efraim-Turban>
- Wang, Y. K. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. doi:10.1016/j.techfore.2015.12.019
- Watson, H. (2009). Tutorial: Business Intelligence – Past, Present, and Future. Obtenido de <https://aisel.aisnet.org/cgi/viewcontent.cgi?article=3490&context=cais>

10. SOLUCIONARIO

Autoevaluación 1		
Pregunta	Respuesta	Retroalimentación
1	b	BI implica el proceso de analizar datos para obtener información valiosa que pueda guiar la toma de decisiones.
2	d.	Un insight es el resultado de analizar datos para obtener una comprensión significativa.
3	c.	Las características clave de un buen Insight son: relevante, accionable, oportuno, comprensible
4	a.	BI mejora la toma de decisiones al proporcionar datos precisos y análisis.
5	b.	La gestión de recursos humanos no es un componente clave de BI, los componentes claves del BI son el análisis de datos, la minería de datos y el procesamiento analítico, y presentación o visualizaciones de datos significativos.
6	d.	BI tiene las siglas de Business Intelligence o Inteligencia de negocios
7	c.	En el modelo Estrella se organiza los datos alrededor de una tabla de hechos conectada directamente con tablas de dimensiones, facilitando consultas rápidas y eficientes.
8	c.	En la arquitectura de BI, los procesos ETL (Extract, Transform, Load) son fundamentales para preparar datos





		provenientes de diversas fuentes y convertirlos en un formato adecuado para análisis.
9	a.	ETL es un proceso que se utiliza en la integración de datos, esencial para la arquitectura de Inteligencia de Negocios. Significa Extract (Extracción de datos de diversas fuentes), Transform (Transformación de datos en un formato adecuado) y Load (Carga de datos en un almacén de datos o Data Warehouse).
10	b.	Un Data Warehouse actúa como un repositorio centralizado que almacena datos integrados de diversas fuentes. Su función principal es facilitar consultas y análisis rápidos, proporcionando una base sólida para la toma de decisiones informadas en la organización.





Autoevaluación 2		
Pregunta	Respuesta	Retroalimentación
1	a.	ETL es un proceso que se utiliza en la integración de datos, esencial para la arquitectura de Inteligencia de Negocios. Significa Extract (Extracción de datos de diversas fuentes), Transform (Transformación de datos en un formato adecuado) y Load (Carga de datos en un almacén de datos o Data Warehouse).
2	b.	Un Data Warehouse actúa como un repositorio centralizado que almacena datos integrados de diversas fuentes. Su función principal es facilitar consultas y análisis rápidos, proporcionando una base sólida para la toma de decisiones informadas en la organización.
3	d.	Las herramientas Pentaho y Apache Nifi son mencionadas como soluciones comunes para realizar procesos ETL. Estas herramientas permiten extraer datos de diversas fuentes, transformarlos según las necesidades del negocio y cargarlos en un Data Warehouse, asegurando la integridad y calidad de los datos para el análisis.
4	b.	Las herramientas de análisis de datos en la arquitectura de BI permiten realizar análisis detallados y descubrir patrones ocultos en los datos. Estas herramientas facilitan la toma de decisiones informadas al proporcionar insights valiosos y ayudar a identificar tendencias y relaciones en los datos empresariales.
5	a.	Power BI, Tableau y QlikView son herramientas ampliamente utilizadas para la creación de dashboards interactivos. Estas plataformas permiten a los usuarios visualizar datos de manera intuitiva y explorar





		información de forma interactiva, facilitando la toma de decisiones basada en datos.
6	c.	El propósito de los procesos ETL es extraer datos de diversas fuentes, transformarlos en un formato adecuado y cargarlos en un Data Warehouse. Este proceso asegura la integración, calidad y consistencia de los datos, preparándolos para análisis y consultas eficientes
7	b.	Pentaho Data Integration (PDI), también conocido como Kettle, es una herramienta poderosa y flexible de ETL. PDI se utiliza para la extracción, transformación y carga de datos, facilitando la integración de datos de diversas fuentes y asegurando la calidad y consistencia necesaria para el análisis en sistemas de Inteligencia de Negocios.
8	c.	Power BI permite a los usuarios conectar datos en tiempo real y crear visualizaciones interactivas. Esto facilita la exploración dinámica de los datos y la obtención de insights valiosos de manera rápida y eficiente, lo que ayuda en la toma de decisiones informadas basadas en datos actualizados.
9	a.	El primer paso para la creación de un reporte en Power BI es conectar la herramienta a las fuentes de datos relevantes. Esta conexión permite importar los datos necesarios para el análisis y la visualización, y es fundamental para asegurar que los datos sean precisos y completos antes de proceder con la transformación y el modelado.
10	d.	En Pentaho Data Integration (PDI), los "Jobs" son procesos que coordinan la ejecución de múltiples transformaciones y otras tareas. Los Jobs permiten gestionar y orquestar la secuencia en la que se ejecutan





		las transformaciones y pueden incluir pasos como la transferencia de archivos, la ejecución de scripts, y la interacción con sistemas externos, facilitando un flujo de trabajo ETL completo y automatizado.
--	--	--



Autoevaluación 3		
Pregunta	Respuesta	Retroalimentación
1	c.	El primer paso en el diseño y desarrollo de dashboards en Power BI es conectar a la fuente de datos. Este paso es crucial porque establece la base sobre la cual se construirán todas las visualizaciones y análisis posteriores, asegurando que los datos que se utilizan son los adecuados y están correctamente integrados en el dashboard.
2	b.	Power BI permite realizar cálculos dinámicos sobre los datos mediante el uso de DAX (Data Analysis Expressions). DAX es un lenguaje de fórmulas avanzado que ayuda a crear medidas personalizadas, columnas calculadas y tablas calculadas, permitiendo análisis más profundos y específicos para obtener información valiosa y apoyar la toma de decisiones basada en datos.
3	b.	En Power BI, los campos se agregan a las visualizaciones mediante un proceso de arrastrar y soltar. Esta funcionalidad facilita la configuración de visualizaciones personalizadas, permitiendo a los usuarios colocar campos específicos en los ejes, valores y leyendas de sus gráficos y otros elementos visuales de manera rápida y eficiente.
4	d.	Power BI es una herramienta de análisis y visualización de datos desarrollada por Microsoft. Permite a los usuarios conectar a diversas fuentes de datos, transformar esos datos y crear visualizaciones interactivas y reportes para facilitar el análisis y la toma de decisiones en una organización.



5	a.	En Power BI, se pueden crear una variedad de visualizaciones, incluyendo gráficos, mapas y tablas. Esto permite a los usuarios representar datos de maneras que facilitan la comprensión y el análisis, haciendo que sea más sencillo identificar patrones, tendencias y outliers, lo cual es esencial para tomar decisiones informadas basadas en datos.
6	c.	Un dashboard en Power BI es una pantalla que compila varias visualizaciones de datos, como gráficos y tablas, proporcionando una vista integral y actualizada de la información clave. Esto permite a los usuarios monitorear métricas importantes en tiempo real y tomar decisiones estratégicas basadas en datos consolidados y fácilmente accesibles.
7	a.	Power BI es una herramienta de inteligencia de negocios. Permite a los usuarios visualizar datos, realizar análisis y compartir insights. Con Power BI, las empresas pueden transformar datos de diversas fuentes en reportes y dashboards interactivos, facilitando el acceso y la comprensión de la información para tomar decisiones informadas.
8	c.	Un beneficio clave de usar Power BI en una empresa es que mejora la toma de decisiones mediante el análisis de datos. Esta herramienta permite a los usuarios visualizar datos complejos de forma sencilla, identificar tendencias y obtener insights que pueden impulsar estrategias y operaciones más efectivas, ayudando a la empresa a ser más data-driven y competitiva.
9	b.	Una visualización en Power BI es una representación gráfica de datos que permite entender mejor la información mediante gráficos, mapas, tablas y otros elementos visuales. Estas visualizaciones son clave





		para analizar tendencias, identificar patrones y comunicar insights de manera efectiva a los stakeholders, ayudando a tomar decisiones basadas en datos.
10	a.	La función principal del Editor de consultas (Power Query) en Power BI es conectar y transformar datos provenientes de diversas fuentes antes de cargarlos en Power BI. Esto permite limpiar, reorganizar y preprocesar los datos, asegurando que sean adecuados para el análisis y la visualización en tus reportes y dashboards. Esta herramienta es esencial para optimizar la eficiencia de tus proyectos de BI al permitirte manipular los datos según tus necesidades específicas

