



European Open
Business School

TÍTULO EXPERTO

Parte 1

Business
intelligence y
fundamentos
del Big Data



ÍNDICE

1. FUNDAMENTOS DEL BUSINESS INTELLIGENCE	4
1.1 Datos vs Información.....	4
1.2 Historia del análisis de datos.....	4
1.3 Niveles de aplicación de la analítica de datos.....	6
1.4 Tipos de análisis de datos.....	7
1.5 ¿Qué es el Business Intelligence?.....	9
1.6 ¿Qué es el Data Analytics?.....	9
1.7 ¿Cuál es la diferencia entre Data Analytics y Business Intelligence?.....	10
1.8 La analítica de datos en los negocios	10
2. BASES DE DATOS. INTRODUCCIÓN.....	13
2.1 Definición de bases de datos.....	13
2.2 Componentes clave de una base de datos	13
2.3 Tipos de bases de datos.....	13
2.4 Operaciones básicas en una base de datos	14
2.5 Beneficios de las bases de datos.....	14
3. MODELOS RELACIONALES DE DATOS.....	15
3.1 ¿Qué son las bases de datos relacionales?	15
4. CONOCIENDO SQL Y SUS SUBLENGUAJES.....	17
4.1 ¿Por qué SQL es importante?	17
4.2 Sublenguajes en SQL.....	17
5. CONSULTAS Y QUERIES	20
5.1 Consulta.....	20
5.2 Query	20

1.1 DATOS VS INFORMACIÓN

Datos

Característica o atributo sin procesamiento, el cual no informa nada por sí solo.

Los datos son esos atributos que le dan identidad a un objeto o ser viviente. Ejemplo: nombre, apellido, DNI, marca, modelo.

Información

Unión de datos procesados que se complementan para informar un hecho.

La información es la unión de varios datos; permite complementar datos que no tienen sentido por sí solos para brindar contexto completo a una situación. Ejemplo: Layla Scheli, Argentino de 31 años.

1.2 HISTORIA DEL ANÁLISIS DE DATOS

Analytics 1.0

Se destaca por su año de inicio en 1990. Esta fue la era del dataware-house empresarial, utilizado para almacenar información, y del software de business intelligence, utilizado para consultar y reportar.

Las nuevas tecnologías informáticas fueron clave. Los sistemas de información fueron en un principio hechos a medida por empresas cuya gran escala justificaba la inversión; luego, fueron comercializados por proveedores externos en formas más genéricas. Ésta fue la era del almacén de datos empresarial, que se utiliza para capturar información, y del software de inteligencia empresarial, que se utiliza para consultar y reportar.

También se requerían nuevas competencias, comenzando con la capacidad de administrar datos. Los conjuntos de datos eran lo suficientemente pequeños en volumen y lo suficientemente estáticos en velocidad como para ser separados en almacenes para su análisis. Sin embargo, fue difícil preparar un conjunto de datos para su inclusión en un almacén. Los analistas dedicaron gran parte de su tiempo a preparar datos para el análisis y relativamente poco tiempo al análisis en sí.

Analytics 2.0

Aparece por primera vez el término Big Data, cuando empresas de redes sociales y basadas en Internet, principalmente en Silicon Valley (Google, eBay, etc.), comenzaron a acumular y analizar nuevos tipos de información. Aunque el término “big data” no se acuñó de



inmediato, la nueva realidad que significó, cambió muy rápidamente el papel de los datos y el análisis en esas empresas.

Los macrodatos también llegaron a distinguirse de los pequeños porque no fueron generados exclusivamente por los sistemas de transacciones internos de una empresa. También se obtuvo de fuentes externas, provenientes de Internet, sensores de varios tipos, iniciativas de datos públicos como el proyecto del genoma humano y capturas de grabaciones de audio y video.

Cuando la analítica entró en la fase 2.0, rápidamente se hizo evidente la necesidad de nuevas herramientas poderosas y la oportunidad de obtener ganancias proporcionándolas. Las empresas se apresuraron a desarrollar nuevas capacidades y adquirir clientes. El amplio reconocimiento de la ventaja que podría obtener un pionero llevó a un nivel impresionante de publicidad, pero también provocó una aceleración sin precedentes de nuevas ofertas.

Analytics 3.0

Las organizaciones integran grandes y pequeños volúmenes de datos de fuentes internas y externas y en formatos estructurados y no estructurados para producir nuevos conocimientos en modelos predictivos y prescriptivos.

Analytics 3.0 marca el punto en el que otras grandes organizaciones comenzaron a hacer lo mismo. Hoy en día, no solo las empresas de información y las empresas en línea pueden crear productos y servicios a partir del análisis de datos, son todas las empresas de todas las industrias.

Si una empresa fabrica cosas, mueve cosas, consume cosas o trabaja con clientes, tiene cantidades cada vez mayores de datos sobre esas actividades. Cada dispositivo, envío y consumidor deja un rastro. Tiene la capacidad de analizar esos conjuntos de datos en beneficio de los clientes y los mercados.

También tiene la capacidad de integrar el análisis y la optimización en cada decisión comercial que se tome en la primera línea de sus operaciones.

1.3 NIVELES DE APLICACIÓN DE LA ANALÍTICA DE DATOS

Los niveles operativo, táctico y estratégico son términos que se utilizan para describir diferentes niveles de toma de decisiones dentro de una organización. Cada nivel implica un enfoque y un alcance diferentes en el proceso de gestión. Empecemos:

Nivel Operativo:

- **Enfoque:** El nivel operativo se centra en las actividades y operaciones diarias de una organización. Implica la ejecución de tareas y procesos rutinarios para mantener las operaciones en funcionamiento.
- **Decisiones:** En este nivel, las decisiones son específicas, detalladas y están orientadas a resolver problemas prácticos inmediatos. Se enfocan en la asignación de recursos, la supervisión del personal y la gestión de actividades diarias.
- **Responsables:** Los responsables de las decisiones a nivel operativo suelen ser los empleados de primera línea y los supervisores que están directamente involucrados en las operaciones diarias.

Nivel Táctico:

- **Enfoque:** El nivel táctico se enfoca en la implementación de estrategias y planes a mediano plazo. Implica la coordinación de recursos y actividades para alcanzar los objetivos establecidos por la alta dirección.
- **Decisiones:** Las decisiones a nivel táctico son más amplias que las decisiones operativas y se centran en la asignación de recursos a corto y mediano plazo, la planificación de proyectos y la supervisión de departamentos o unidades específicas.
- **Responsables:** Los gerentes de nivel medio suelen ser los responsables de tomar decisiones a nivel táctico. Estos gerentes supervisan equipos o departamentos y se aseguran de que las estrategias se implementen eficazmente.

Nivel Estratégico:

- **Enfoque:** El nivel estratégico se enfoca en la formulación de objetivos a largo plazo y en la planificación global de la organización. Implica la toma de decisiones fundamentales que afectan la dirección y el futuro de la empresa.
- **Decisiones:** Las decisiones estratégicas son de gran alcance y tienen un impacto a largo plazo en la organización. Involucran la definición de la misión y visión de la empresa, la identificación de nuevas oportunidades de mercado y la asignación de recursos a largo plazo.
- **Responsables:** Los altos directivos, como CEO (Director Ejecutivo) y otros miembros del consejo de administración, son responsables de tomar decisiones estratégicas. Estos líde-

res definen la dirección estratégica de la organización y establecen los objetivos generales a largo plazo.

En resumen, el nivel operativo se centra en las operaciones diarias, el nivel táctico se enfoca en la implementación de estrategias a mediano plazo y el nivel estratégico se concentra en la formulación de objetivos a largo plazo y la dirección global de la organización. Estos niveles están interrelacionados y trabajan juntos para lograr los objetivos organizacionales.

1.4 TIPOS DE ANÁLISIS DE DATOS

Análisis descriptivo

El análisis descriptivo, como su nombre lo indica, consiste en describir las tendencias claves en los datos existentes y observar las situaciones que conduzcan a nuevos hechos. Respondiendo a la pregunta: ¿Qué?

Este método se basa en una o varias preguntas de investigación y no tiene una hipótesis. Además, incluye la recopilación de datos relacionados; posteriormente, los organiza, tabula y describe el resultado.

Un análisis básico descriptivo implica el calcular las medidas simples de composición y distribución de variables. Dependiendo del tipo de datos, pueden ser proporciones, tasas, razones o promedios. Además, cuando sea necesario, como en el caso de las encuestas por muestreo, pueden utilizarse medidas de asociación entre variables para decidir si las diferencias observadas entre mujeres y hombres son estadísticamente significativas o no.

El análisis descriptivo proporciona una base de conocimiento que puede ser una base para realizar posteriormente un análisis cuantitativo. Si se interpreta de forma correcta, los datos pueden ofrecer perspectivas útiles que lleven a la creación de una hipótesis.

Análisis de diagnóstico

Se definen conclusiones basadas en los datos históricos identificados en el análisis descriptivo. Respondiendo a la pregunta: ¿Por qué?

El análisis descriptivo, es el paso inicial en la toma de decisiones de la mayoría de las empresas, es un proceso más simple que narra los hechos de lo que ya sucedió, por lo que se complementa con el análisis descriptivo. El análisis de diagnóstico va un paso más allá para descubrir el razonamiento detrás de ciertos resultados.

Este tipo de análisis generalmente se realiza mediante técnicas como el descubrimiento de datos, el desglose, la extracción de datos y las correlaciones. En el proceso de descubrimiento, los analistas identifican las fuentes de datos que les ayudarán a interpretar los resultados. Profundizar implica centrarse en una determinada faceta de los datos o un widget en particular. Este desglose se realiza fácilmente utilizando una plataforma de BI.

Análisis predictivo

El análisis predictivo es el dominio de los científicos de datos, a quienes se les asignan pasos en el flujo de trabajo analítico representados por las siguientes cinco categorías:

- **Identificar los resultados comerciales:** determina qué preguntas deben responderse mediante el análisis predictivo. Si no se identifican los resultados correctos, ejecutar análisis predictivos es como lanzar dardos en la oscuridad. También identifique los impulsores (variables independientes) que probablemente afectarán el resultado (variable dependiente).
- **Determinar los datos necesarios para capacitar:** el análisis predictivo requiere datos de múltiples fuentes, por lo que los analistas deben identificar las fuentes de datos actuales. Si las fuentes existentes son insuficientes, deben adquirir datos de otras fuentes para garantizar que los modelos puedan entrenarse con precisión.
- **Determinar métodos de análisis:** diferentes técnicas son adecuadas para responder diferentes preguntas dependiendo de la cantidad y tipo de datos disponibles. El análisis estadístico, las redes neuronales, el aprendizaje automático y la minería de datos son ejemplos de técnicas sofisticadas que pueden predecir resultados.
- **Validar resultados:** la analítica avanzada no se puede utilizar como una caja negra. Los datos de entrenamiento incorrectos, los algoritmos incorrectos y las suposiciones deficientes son solo algunos de los escollos que pueden resultar en predicciones falsas. Los científicos de datos deben trabajar en estrecha colaboración con analistas y líderes de líneas de negocio para garantizar que los modelos predictivos tengan sentido comercial.
- **Probar las predicciones sobre el rendimiento:** los modelos predictivos deben ajustarse continuamente para mejorar la precisión. Si un modelo falla, los analistas deben identificar la causa raíz y volver a capacitarse y probar para mejorar los modelos.

Las soluciones modernas de análisis predictivo deben proporcionar a los científicos de datos un banco de trabajo de productividad que admita todas estas funciones. Pero el análisis predictivo no se trata solo de la productividad de los científicos de datos. El valor de la analítica predictiva proviene de su implementación dentro de otras aplicaciones en el flujo de negocios. Eso significa que una solución de análisis predictivo debe admitir la ejecución de puntuaciones predictivas a escala.

Análisis prescriptivo

El análisis prescriptivo se complementa con el análisis predictivo ya que toma como fuente de información, la salida de la predicción, combinado con reglas y optimización basada en restricciones, permite tomar mejores decisiones sobre qué hacer. La decisión podría ser enviar una tarea automatizada a un tomador de decisiones humano junto con un conjunto de recomendaciones de acción siguiente, o enviar un comando de acción siguiente preciso a otro sistema.

Por lo tanto, el análisis prescriptivo es más adecuado para situaciones en las que las restricciones son precisas. Esto suele suceder con las elecciones tácticas, en las que se deben tomar muchas decisiones dentro de un período determinado. Algunos ejemplos de esto incluyen compras de publicidad programática, negociación de acciones y detección de fraudes. Sin embargo, el universo de situaciones en las que se están aplicando análisis prescriptivo continúa expandiéndose y eventualmente estará en muchos tipos de procesos de toma de decisiones.

La adopción más amplia de análisis prescriptivo a menudo se ve obstaculizada no por la funcionalidad de las soluciones de análisis prescriptivo, sino más bien por factores externos como la regulación gubernamental, el riesgo de mercado o el comportamiento organizacional. Esta es la

situación en la atención médica, por ejemplo, donde existen algunos éxitos tempranos, pero la adopción generalizada de análisis prescriptivos aún llevará años. Independientemente del plazo para la adopción generalizada de análisis prescriptivos, todas las empresas deben comenzar a evaluar la aplicabilidad de este tipo de análisis para sus propias operaciones.

1.5 ¿QUÉ ES EL BUSINESS INTELLIGENCE?

El Business Intelligence (BI), o inteligencia empresarial en español, se refiere al conjunto de tecnologías, procesos y herramientas que ayudan a las empresas a recopilar, analizar y presentar información empresarial para facilitar la toma de decisiones estratégicas. El objetivo principal del BI es convertir los datos brutos en información significativa y útil para apoyar la toma de decisiones informadas en una organización.

Algunos aspectos clave del Business Intelligence incluyen:

- **Recopilación de datos:** Se recopilan datos de diversas fuentes, tanto internas como externas a la empresa. Estos datos pueden incluir información sobre clientes, operaciones, finanzas, recursos humanos, entre otros.
- **Procesamiento y transformación:** Los datos recopilados se procesan y transforman para convertirlos en información significativa. Esto puede implicar la limpieza de datos, la consolidación de información y la creación de modelos de datos coherentes.
- **Almacenamiento de datos:** Los datos procesados se almacenan en almacenes de datos o bases de datos diseñadas para facilitar la consulta y el análisis.
- **Ánalisis y consulta:** Se utilizan herramientas de BI para analizar y consultar los datos almacenados. Estas herramientas permiten a los usuarios explorar tendencias, patrones y relaciones en los datos.
- **Generación de informes y dashboards:** El BI facilita la creación de informes y dashboards interactivos que proporcionan visualizaciones gráficas de la información. Estos informes ayudan a los usuarios a entender fácilmente los resultados del análisis.
- **Toma de decisiones:** Basándose en la información proporcionada por el BI, los líderes y tomadores de decisiones pueden tomar decisiones más informadas y estratégicas para mejorar el rendimiento empresarial.

1.6 ¿QUÉ ES EL DATA ANALYTICS?

Conjunto de métodos y técnicas de medición, que permiten gestionar la información en tres grandes etapas, recolección, transformación y visualización.

El Data Analytics no es un software, sino que se trata de un grupo de tareas para identificar la necesidad del negocio a fin de trabajar en su solución. Para eso, las personas que tienen habilidades para el análisis de datos utilizan herramientas y tienen criterio para determinar las métricas que deben ser tenidas en cuenta ya que no todas son válidas.

La tendencia a trabajar en base al análisis de datos está en aumento ya que las organizaciones que trabajan de esta manera pueden tomar mejores decisiones operativas y gestionar los riesgos.

1.7 ¿CUÁL ES LA DIFERENCIA ENTRE DATA ANALYTICS Y BUSINESS INTELLIGENCE?

Data Analytics y Business Intelligence son dos campos relacionados pero distintos dentro del ámbito de la gestión y análisis de datos. Aunque a menudo se utilizan de manera intercambiable, tienen enfoques y objetivos ligeramente diferentes. Aquí te explico las diferencias clave entre ellos:

Data Analytics (Analítica de Datos):

- **Enfoque:** Se centra en el análisis estadístico y matemático de los conjuntos de datos para descubrir patrones, tendencias, correlaciones y obtener información valiosa.
- **Objetivo:** El objetivo principal de la analítica de datos es descubrir significados ocultos en los datos para tomar decisiones informadas, identificar oportunidades y optimizar procesos.
- **Herramientas:** Utiliza técnicas avanzadas de estadísticas, matemáticas y programación para analizar datos. También implica el uso de algoritmos de machine learning y modelado predictivo para hacer pronósticos y generar insights.

Business Intelligence (Inteligencia de Negocios):

- **Enfoque:** Se centra en la presentación y visualización de datos para proporcionar información detallada y orientada a los negocios.
- **Objetivo:** El objetivo principal de la inteligencia de negocios es ayudar a las organizaciones a tomar decisiones basadas en datos mediante la creación de informes, dashboards interactivos y herramientas de análisis que permitan a los usuarios finales comprender fácilmente los datos.
- **Herramientas:** Utiliza herramientas específicas de Business Intelligence para crear informes, dashboards, cubos OLAP (Procesamiento Analítico en Línea) y herramientas de consulta para explorar datos de manera intuitiva.

1.8 LA ANALÍTICA DE DATOS EN LOS NEGOCIOS

La analítica de datos ha emergido como una herramienta invaluable en la industria moderna, brindando una variedad de beneficios y aplicaciones que transforman la forma en que las empresas operan y toman decisiones. Estos son algunos de los principales beneficios y aplicaciones de la analítica de datos en la industria:

1. **Mejora de la Toma de Decisiones:** La analítica de datos permite a las empresas tomar decisiones informadas basadas en evidencia y análisis profundo de datos. Esto conduce a decisiones más precisas y estratégicas que impulsan el crecimiento y la eficiencia operativa.



2. Optimización de Operaciones:

Al analizar grandes volúmenes de datos, las organizaciones pueden identificar áreas de ineficiencia en sus procesos y operaciones. Al optimizar estas áreas, las empresas pueden reducir costos, mejorar la calidad y aumentar la productividad.

3. Predicción y Pronóstico:

La analítica de datos utiliza técnicas de modelado predictivo para prever tendencias futuras y comportamientos del mercado. Esto permite a las empresas anticipar la demanda del cliente, identificar oportunidades y mitigar riesgos potenciales.

4. Personalización y Experiencia del Cliente:

Analizar los datos del cliente ayuda a las empresas a comprender mejor las preferencias y necesidades individuales. Esto conduce a la personalización de productos y servicios, lo que mejora la satisfacción del cliente y fortalece la lealtad a la marca.

5. Gestión de Inventarios:

La analítica de datos optimiza la gestión de inventarios al prever la demanda y evitar excesos o escaseces. Esto garantiza que las empresas tengan los productos adecuados en el lugar correcto y en el momento adecuado, reduciendo los costos de almacenamiento y pérdida de ventas.

6. Mantenimiento Predictivo:

En la industria manufacturera, la analítica de datos se utiliza para predecir el momento óptimo para el mantenimiento de maquinaria y equipos. Esto reduce el tiempo de inactividad no planificado y prolonga la vida útil de los activos.

7. Seguridad y Detección de Fraude:

Las empresas utilizan la analítica de datos para identificar pa-

trones sospechosos y anomalías que podrían indicar actividades fraudulentas. Esto es especialmente relevante en sectores como servicios financieros y comercio electrónico.

8. Innovación y Desarrollo de Productos: Al analizar datos de mercado y retroalimentación del cliente, las empresas pueden identificar oportunidades de innovación y desarrollar nuevos productos y servicios que satisfagan las necesidades cambiantes de los consumidores.

Para concluir, la analítica de datos en la industria no solo proporciona información valiosa, sino que también impulsa la innovación, mejora la eficiencia operativa y fortalece la ventaja competitiva. Al aprovechar los poderes analíticos, las empresas pueden adaptarse rápidamente a un entorno empresarial en constante cambio y tomar decisiones estratégicas fundamentadas para el éxito a largo plazo.

2.

BASES DE DATOS. INTRODUCCIÓN

Las bases de datos son un componente fundamental en el ámbito de la informática y la gestión de la información. Proporcionan un medio organizado y eficiente para almacenar, gestionar y recuperar datos de manera estructurada. Aquí hay una introducción básica a las bases de datos:

2.1 DEFINICIÓN DE BASES DE DATOS

Una base de datos es un conjunto organizado de datos que se almacena de manera estructurada para que pueda ser fácilmente gestionado, accedido y actualizado. Está diseñada para almacenar información de manera coherente y eficiente.

2.2 COMPONENTES CLAVE DE UNA BASE DE DATOS

- **Tablas:** Las tablas son la principal estructura de almacenamiento en una base de datos. Representan entidades específicas, como clientes, productos o empleados, y están organizadas en filas y columnas.
- **Registros:** Cada fila en una tabla se conoce como un registro. Representa una instancia única de la entidad que la tabla describe.
- **Campos:** Cada columna en una tabla se llama campo. Define el tipo de datos que se puede almacenar en esa columna, como texto, números o fechas.
- **Claves primarias:** Una clave primaria es un campo o conjunto de campos que identifica de manera única cada registro en una tabla. Garantiza la integridad de los datos y facilita las relaciones entre tablas.
- **Relaciones:** Las relaciones establecen conexiones entre tablas basadas en claves primarias y externas. Permiten la vinculación de datos entre entidades relacionadas.

2.3 TIPOS DE BASES DE DATOS

- **Relacionales:** Utilizan tablas para almacenar datos y establecen relaciones entre ellas. Ejemplos incluyen MySQL, PostgreSQL y Microsoft SQL Server.
- **NoSQL:** Diseñadas para manejar grandes volúmenes de datos no estructurados o semi-estructurados. Incluyen bases de datos de documentos, de grafos, clave-valor, entre otras. Ejemplos son MongoDB, Cassandra y Redis.

- **Objeto-Relacionales:** Combina características de las bases de datos relacionales y de las bases de datos orientadas a objetos.

2.4 OPERACIONES BÁSICAS EN UNA BASE DE DATOS

- **CRUD (Create, Read, Update, Delete):**

- Create: Inserción de nuevos datos.
- Read: Consulta y recuperación de datos.
- Update: Modificación de datos existentes.
- Delete: Eliminación de datos.

- **Consultas SQL:**

- Uso del lenguaje SQL (Structured Query Language) para realizar operaciones en bases de datos relacionales.

2.5 BENEFICIOS DE LAS BASES DE DATOS

- **Organización Eficiente:** Facilitan la organización lógica y eficiente de grandes cantidades de datos.
- **Integridad de Datos:** Mantienen la integridad y consistencia de los datos mediante el uso de claves primarias y restricciones.
- **Recuperación Rápida:** Permiten recuperar información de manera rápida y eficiente mediante consultas optimizadas.
- **Compartir Datos:** Facilitan el acceso y compartición de datos entre usuarios y aplicaciones autorizadas.
- **Seguridad:** Ofrecen medidas de seguridad para proteger los datos sensibles.

INFORMACIÓN DE INTERÉS:

<https://www.w3schools.com/sql/default.asp>

En resumen, las bases de datos son esenciales en el mundo de la informática para gestionar datos de manera organizada, permitiendo a las organizaciones almacenar, acceder y gestionar información de manera eficiente.

Importante: En las clases en vivo profundizaremos en el detalle y explicación asociada a esta temática.

3

MODELOS RELACIONALES DE DATOS

3.1 ¿QUÉ SON LAS BASES DE DATOS RELACIONALES?

Las bases de datos relacionales son un tipo de sistema de gestión de bases de datos (SGBD) que organiza y estructura la información en forma de tablas relacionadas entre sí. Este modelo de base de datos se basa en los principios de la teoría de conjuntos y relaciones, y utiliza claves para establecer conexiones entre los diferentes conjuntos de datos.

Aquí hay algunos conceptos clave asociados con las bases de datos relacionales:

1. Tablas:

- Las tablas son la estructura fundamental en las bases de datos relacionales.
- Cada tabla representa una entidad específica, como clientes, productos o empleados.
- Las tablas están organizadas en filas y columnas.

2. Registros:

Cada fila en una tabla se denomina “registro”.

Un registro representa una instancia única de la entidad que la tabla describe.

3. Campos:

- Cada columna en una tabla se llama “campo”.
- Cada campo tiene un tipo de datos específico, como texto, número o fecha.

4. Claves Primarias:

- Una clave primaria es un campo o conjunto de campos que identifican de manera única cada registro en una tabla.
- Garantiza que no haya duplicados y facilita las relaciones entre tablas.

5. Relaciones:

- Las relaciones establecen vínculos entre tablas utilizando claves primarias y claves externas.
- Permiten la consulta de datos relacionados en múltiples tablas.

6. SQL (Structured Query Language):

- Se utiliza el lenguaje SQL para realizar operaciones en bases de datos relacionales.
- Incluye comandos como SELECT, INSERT, UPDATE y DELETE para realizar operaciones CRUD.

7. Normalización:

- La normalización es un proceso que se utiliza para organizar datos de manera eficiente, reducir redundancias y mejorar la integridad de los datos.

EJEMPLO PRÁCTICO:

Supongamos que tienes dos tablas: una para clientes y otra para pedidos. La tabla de pedidos podría tener una clave externa que hace referencia a la clave primaria de la tabla de clientes, estableciendo así una relación entre ellas.

The screenshot shows a database interface with two tables:

Tabla de Clientes:

ID_Cliente	Nombre	Edad	Ciudad
1	Juan	30	Buenos Aires
2	Maria	25	Madrid

Tabla de Pedidos:

ID_Pedido	Fecha	Producto	ID_Cliente
101	2023-01-15	Producto A	1
102	2023-02-03	Producto B	2

En este ejemplo, la relación se establece a través de la columna ID_Cliente, que es una clave externa en la tabla de pedidos que referencia la clave primaria en la tabla de clientes.

Las bases de datos relacionales son ampliamente utilizadas en aplicaciones empresariales y sistemas de información debido a su capacidad para gestionar datos de manera estructurada, eficiente y con integridad. Ejemplos de sistemas de gestión de bases de datos relacionales (RDBMS) incluyen MySQL, PostgreSQL, Microsoft SQL Server y Oracle Database.

4

CONOCIENDO SQL Y SUS SUBLENGUAJES

SQL, o **Structured Query Language** (Lenguaje de Consulta Estructurado), es un lenguaje de programación diseñado para gestionar y manipular datos almacenados en sistemas de gestión de bases de datos relacionales (RDBMS). Fue creado en la década de 1970 por IBM, y desde entonces ha sido adoptado como un estándar de facto en la interacción con bases de datos.

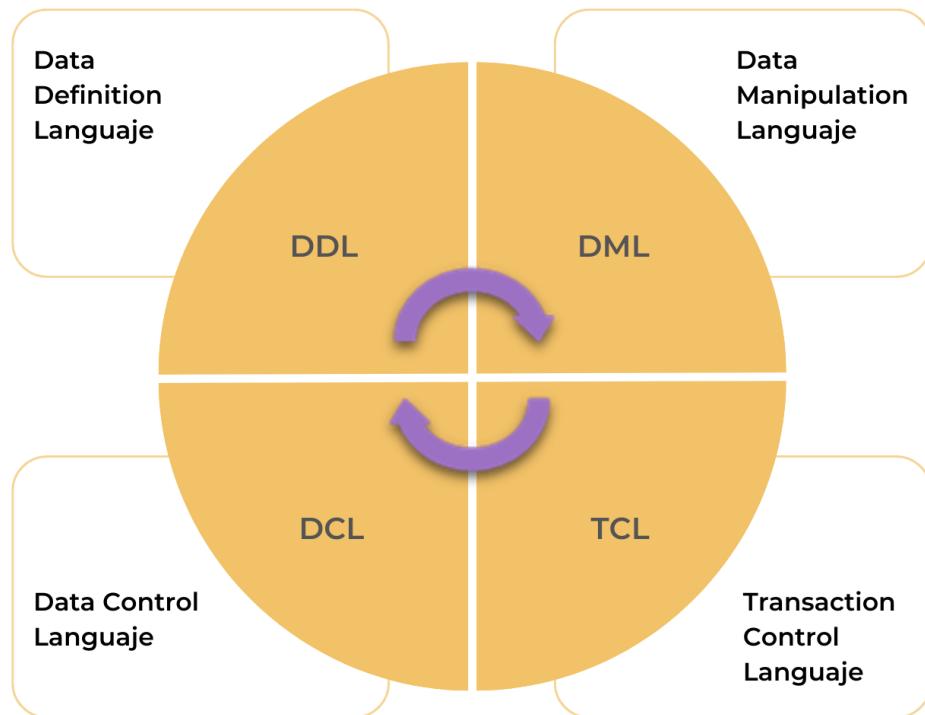
4.1 ¿POR QUÉ SQL ES IMPORTANTE?

Realmente existen un sinfín de motivos y razones por lo que SQL es una herramienta clave e indispensable que todo analista de datos debería de conocer. Comentemos juntos, algunos puntos esenciales:

- **Interacción con bases de datos relacionales:** SQL se utiliza para realizar operaciones en bases de datos relacionales. Le permite crear, modificar y eliminar datos en estas bases de datos.
- **Estándar de la industria:** SQL es el estándar de facto de la industria de bases de datos. La mayoría de los sistemas de gestión de bases de datos relacionales, como MySQL, PostgreSQL, SQL Server y Oracle, utilizan SQL como lenguaje principal.
- **Facilidad de uso:** SQL es fácil de aprender y usar. Su sintaxis es clara y comprensible, lo que permite a los desarrolladores escribir consultas complejas con relativa facilidad.
- **Seguridad:** SQL proporciona mecanismos para controlar el acceso a los datos.
- **Escalabilidad:** Las bases de datos SQL son altamente escalables. Pueden manejar grandes cantidades de datos y una gran cantidad de usuarios simultáneos, lo que los hace adecuados para aplicaciones comerciales y sitios web de alto tráfico.
- **Ánalisis de datos:** SQL se utiliza no solo para gestionar datos sino también para realizar operaciones analíticas complejas. Las consultas SQL se pueden combinar y procesar para extraer información significativa de grandes conjuntos de datos.
- **Aplicaciones comerciales:** Muchas aplicaciones comerciales, como sistemas de gestión de inventario, sistemas CRM y sistemas de recursos humanos, utilizan bases de datos SQL para almacenar y administrar datos críticos para el negocio.

4.2 SUBLENGUAJES EN SQL

SQL se compone de 4 sublenguajes, los cuales se detallan y se explican a continuación, a partir de la siguiente imagen de referencia:



Fuente: Elaboración propia

- **DDL:** El Lenguaje de Definición de Datos, se ocupa de modificar la estructura de objetos de una bb.dd. Lo conforman diferentes sentencias que nos permiten crear, modificar, borrar o definir la estructura de las tablas que almacenan datos.

Las sentencias disponibles a través de DDL, son:

CREATE	ALTER	DROP
Creación de bases de datos por ejemplo, tablas, vistas.	Modificación de bases de datos por ejemplo, tablas, vistas.	Eliminación de bases de datos por ejemplo, tablas, vistas.

- **DML:** Este sublenguaje de SQL, permite a los usuarios manipular datos en una base de datos. Esta manipulación implica por ejemplo insertar datos en nuestras tablas, recuperar datos existentes, eliminar registros e incluso modificar los datos existentes.

Las sentencias más relevantes por destacar dentro del mundo de DML son:

INSERT	UPDATE	DELETE	SELECT
Insertión de registros en tablas.	Actualización y/o modificación de registros en tablas.	Eliminación de registros en tablas.	Consulta de registros en tablas.

- **DCL:** Es un lenguaje proporcionado por el sistema de gestión de base de datos que incluye una serie de diferentes comandos en SQL que nos permiten controlar el acceso a los datos contenidos en nuestras bases de datos, tiene un enfoque de seguridad.

Las sentencias más importantes son:

GRANT	REVOKE
Otorgar accesos a los usuarios.	Revoca los accesos concedidos previamente.

- **TCL:** Este sublenguaje en SQL, nos permite administrar diferentes transacciones que ocurren dentro de una base de datos. Al igual que DCL, tiene un enfoque orientado a la seguridad de datos.

Las instrucciones claves son:

COMMIT	ROLLBACK
Empleado para guardar el trabajo hecho.	Utilizado para deshacer la modificación que hice desde el último COMMIT.

Importante: En las clases en vivo profundizaremos en el detalle y explicación asociada a esta temática.

5

CONSULTAS Y QUERIES

En el contexto de bases de datos y SQL, el término “consulta” y “query” son utilizados de manera intercambiable para referirse a instrucciones o solicitudes que se envían a una base de datos con el objetivo de recuperar, manipular o gestionar datos almacenados. Aquí hay una breve descripción de ambos términos:

5.1 CONSULTA

- Una consulta es una solicitud realizada a una base de datos para obtener información específica.
- Puede referirse a una variedad de operaciones, como la recuperación de datos, la actualización de registros o la eliminación de información.
- En términos generales, cualquier interacción con la base de datos que implique recuperar o modificar datos se puede llamar “consulta”.
- Puede incluir consultas simples como la recuperación de todos los registros de una tabla o consultas más complejas que involucran múltiples tablas y criterios de filtrado.

EJEMPLO DE CONSULTA:

```
SELECT Nombre, Salario FROM Empleados WHERE Salario > 50000;
```

5.2 QUERY

- El término “query” es sinónimo de “consulta” en el contexto de bases de datos y SQL.
- En algunas regiones o contextos específicos, es posible que se utilice “query” con más frecuencia que “consulta”.
- Puede referirse tanto a consultas de lectura (SELECT) como a consultas de escritura (INSERT, UPDATE, DELETE).

EJEMPLO DE QUERY:

```
UPDATE Empleados SET Salario = 55000 WHERE Nombre = 'Juan';
```

En resumen, tanto “consulta” como “query” se refieren a solicitudes realizadas a una base de datos para interactuar con la información almacenada. Pueden implicar operaciones de lectura o escritura, y ambas expresiones son comúnmente utilizadas en el ámbito de la gestión de bases de datos y el lenguaje SQL.

Mas información:

<https://thedataschools.com/sql/select/>

Importante: En las clases en vivo profundizaremos en el detalle y explicación asociada a esta temática.



European Open
Business School

Teléfono : +34 919 058 955
Whatsapp : +34 627 861 440
info@europeanopen.es
europeanopen.es