Bases de datos y conexiones

Asociación Mexicana de Inversionistas

¿Qué son las bases de datos?

Una base de datos es una colección de datos ordenada lógicamente y pueden alojarse de forma física o digital.

La mayoría de las bases de datos utilizan un lenguaje de consulta estructurada (SQL) para escribir y consultar datos. Así como una estructura en función de tablas, filas y columnas.

¿Cómo se diferencia de un Excel?

Las hojas de cálculo en Excel se pensaron en el usuario que las utiliza, por lo tanto el manejo de datos y la gestión no está diseñada para extraer y procesar muchos datos por diferentes usuarios al mismo tiempo.

Diferencias:

1. Capacidad de almacenamiento de datos
2. Usuarios que acceden a la información
3. Seguridad de la información.

Datos estructurados y no estructurados

Datos estructurados	Datos semiestructurados	Datos no estructurados
Fichas de clientes Fecha de nacimiento Nombre Dirección Transacciones en un mes Puntos de compra	**Correos electrónicos** Parte estructurada: destinatario, receptores, tema Parte no estructurada: cuerpo del mensaje	**Persona a persona** Comunicaciones en las redes sociales **Persona a máquina** Dispositivos médicos Comercio electrónico Ordenadores, móviles **Máquina a máquina** Sensores, dispositivos GPS Cámaras de seguridad

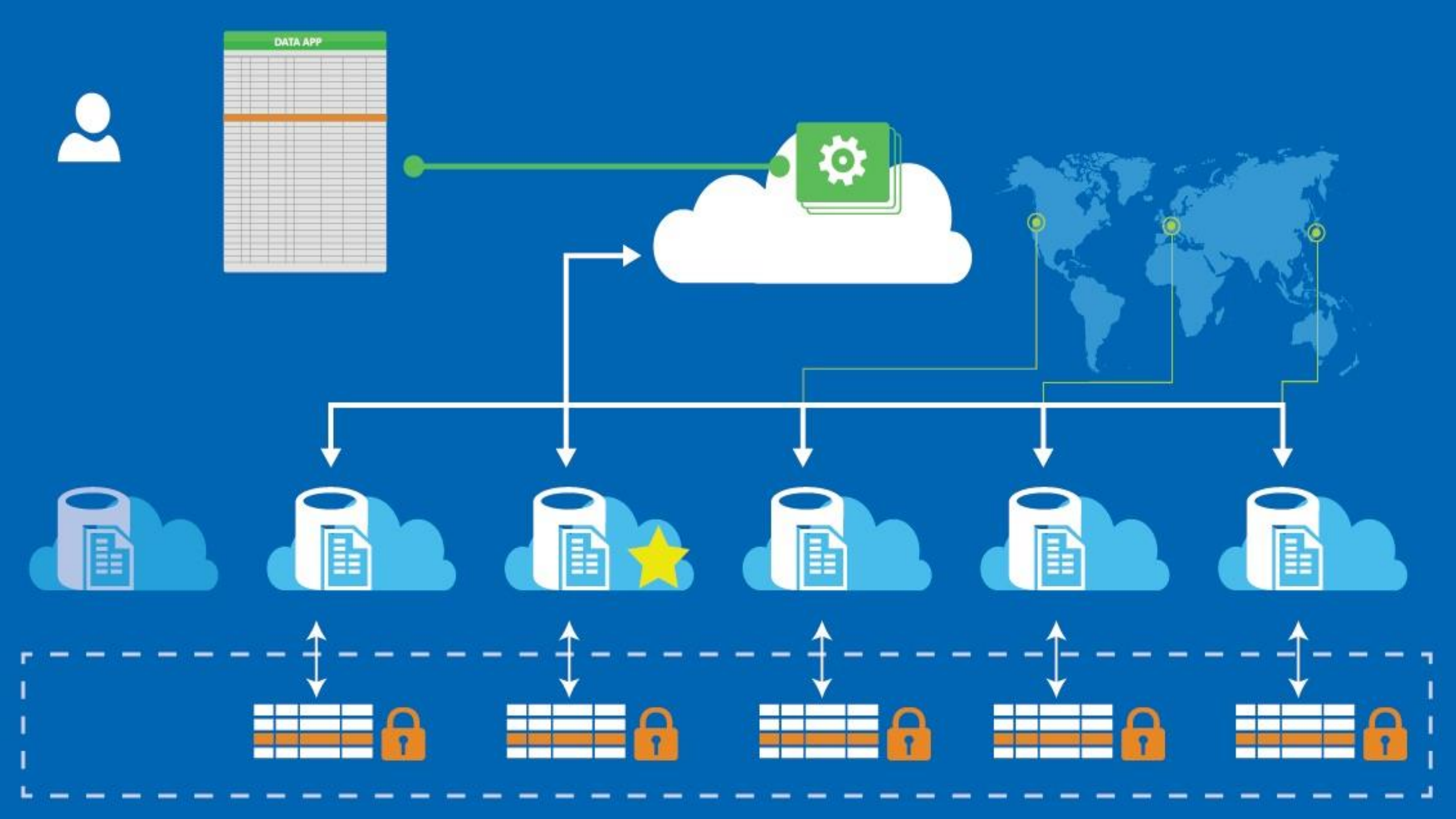
Los **datos estructurados** están altamente organizados y formateados de tal manera **que** se pueden buscar fácilmente en bases de **datos** relacionales. Los **datos** no **estructurados** no tienen un formato u organización predefinidos, lo **que** hace **que** sea mucho más difícil de recopilar, procesar y analizar.

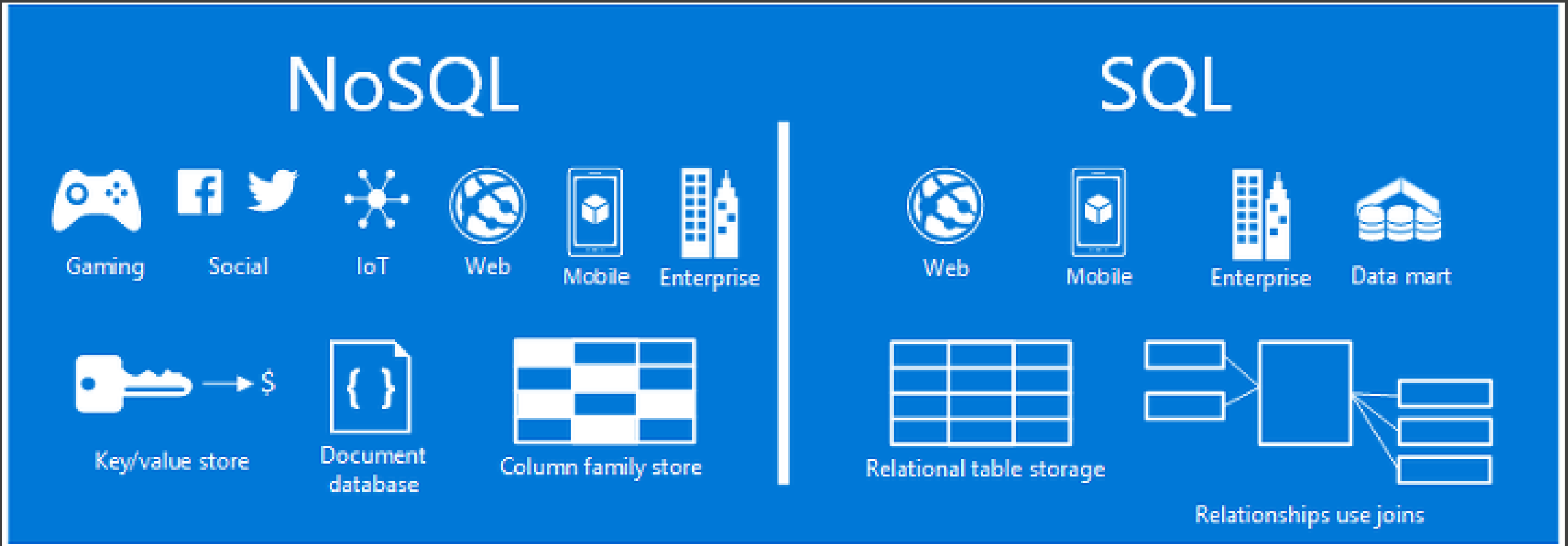
Tipos de bases de datos

- **Base de datos jerárquica:** Aquellas que permitían ligar un elemento estructurado a varios de un subnivel.
- **Bases de datos en red:** Son las bases de datos que permitían conectar un mismo elemento estructurado a otros en el mismo nivel.
- **Bases de datos relacionales:** Los elementos de una base de datos relacional se organizan como un conjunto de tablas con columnas y filas. Es la más usada y eficiente en cuanto a datos estructurados.

Tipos de bases de datos

- Bases de datos distribuidas. Una base de datos distribuida consta de dos o más archivos que se encuentran en sitios diferentes. La base de datos puede almacenarse en varios ordenadores o repartirse en diferentes redes.
- Almacenes de datos. Un data warehouse es un tipo de base de datos diseñado específicamente para almacenar datos, consultas y análisis rápidos.


DATA APP



Tipos de bases de datos

Bases de datos NoSQL. Una base de datos no relacional permite almacenar y manipular datos no estructurados y semiestructurados. Hoy son la base de la tecnología Blockchain dato su capacidad de ampliación de capacidad y replicaciones de seguridad.

Sistema de Gestión de Base de Batos

Una base de datos requiere un software de gestión de bases de datos que sirve como interfaz entre la base de datos y sus programas o usuarios finales, lo que permite a los usuarios recuperar, actualizar y gestionar cómo se organiza y se optimiza la información.

Un SGBD también facilita la supervisión y el control de las bases de datos, lo que permite una variedad de operaciones administrativas como la supervisión del rendimiento, el ajuste, la copia de seguridad y la recuperación.

Retos para la gestión de bases de datos

1. Conectividad y aumento en la tasa de escritura de datos y su volumen derivados de la 4ta revolución industrial y el Internet de las Cosas IOT
2. Seguridad de los datos y accesibilidad balanceada.
3. Mantenimiento de la base de datos y la infraestructura
4. Eliminación de los límites de la escalabilidad.

El futuro de las bases de datos

Bases de datos de autogestión. El tipo de base de datos más nuevo e innovador, las bases de datos autónomas están basadas en la nube y utilizan el Machine Learning para automatizar el ajuste de la base de datos, la seguridad, las copias de seguridad, las actualizaciones y otras tareas de gestión rutinarias que tradicionalmente realizan los administradores de bases de datos.

Estructuración de datos

Asociación Mexicana de Inversionistas

Tipos de variable

Nombre	Tipo	Ocupa	Rango aprox.
byte	Entero	1 byte	-128 a 127
short	Entero	2 bytes	-32768 a 32767
int	Entero	4 bytes	2,000,000,000
long	Entero	8 bytes	Muy grande
float	Decimal simple	4 bytes	Muy grande
double	Decimal doble	8 bytes	Muy grande
char	Caracter simple	2 bytes	----
String	Cadena de texto	----	----
boolean	Valor true o false	1 byte	----

Tablas

Para saber identificar el tipo de tabla, es necesario validar la naturaleza de los elementos que la conforman y su grado de actualización. Existen dos segmentos comunes que son las tablas fijas y las dinámicas.

Tablas

Para las bases de datos relacionales se pueden observar tablas que van relacionadas entre sí pero pueden adoptar una de las siguientes funciones:

1. Tablas de definiciones o catálogos
2. Tablas de segmentación de datos
3. Tablas de registros y subregistros
4. Tablas de vinculación/puente

Asociación Mexicana de Inversionistas

Relaciones

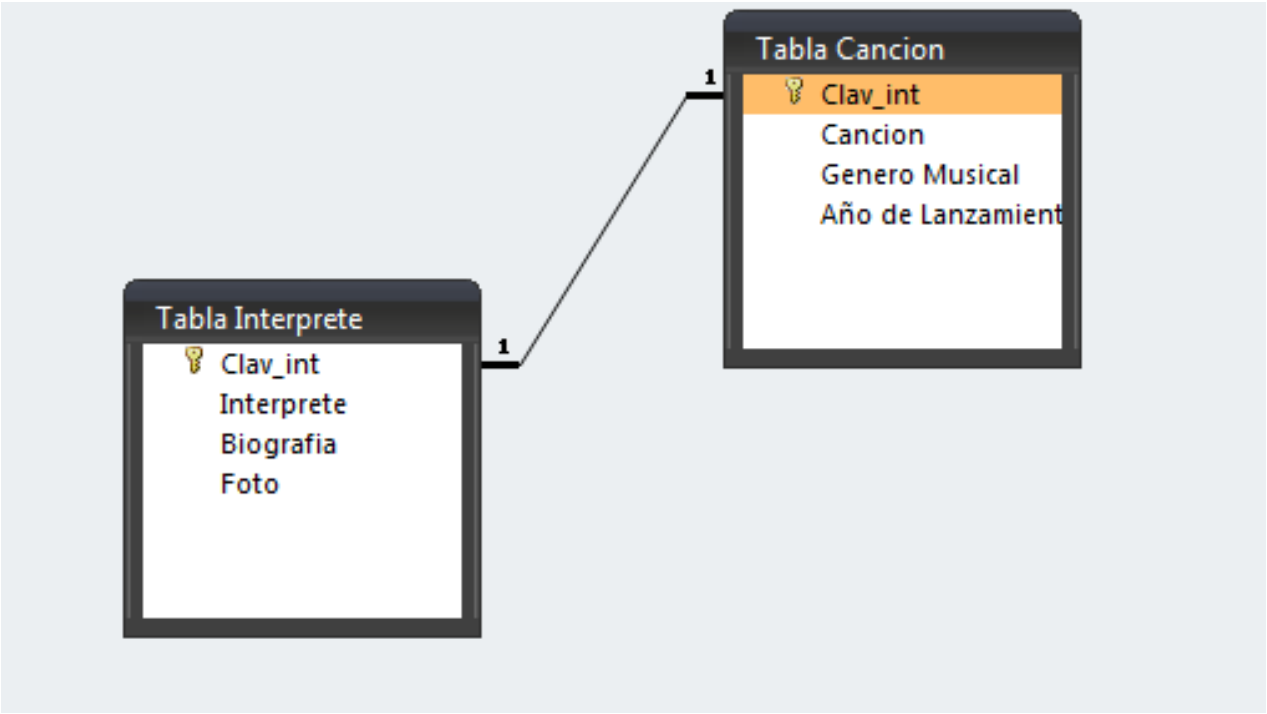
Las relaciones en las bases de datos nos permiten gestionar la información con facilidad. Permiten vincular dos o más tablas por medio de llave primaria o dato clave que generan una integridad y concordancia lógica entre estos. Siempre debe considerarse el mismo tipo de variable para estas relaciones entre tablas.

Estos son los tipos de relaciones más comunes:

1. Uno a Uno
2. Uno a Varios
3. Varios a Varios

Relaciones uno a uno

Este tipo de relaciones nos permite amarrar y conciliar datos únicos. Nos permite gestionarlos de manera ordenada. Este tipo de relación conecta tablas con un solo elemento clave y sin duplicidad.



Asociación Mexicana de Inversionistas

Relaciones uno a varios

Son el tipo de relaciones más utilizadas ya que conecta una tabla con una llave primaria con otra tabla que tiene el mismo valor repetido (llave foránea) en cada registro.

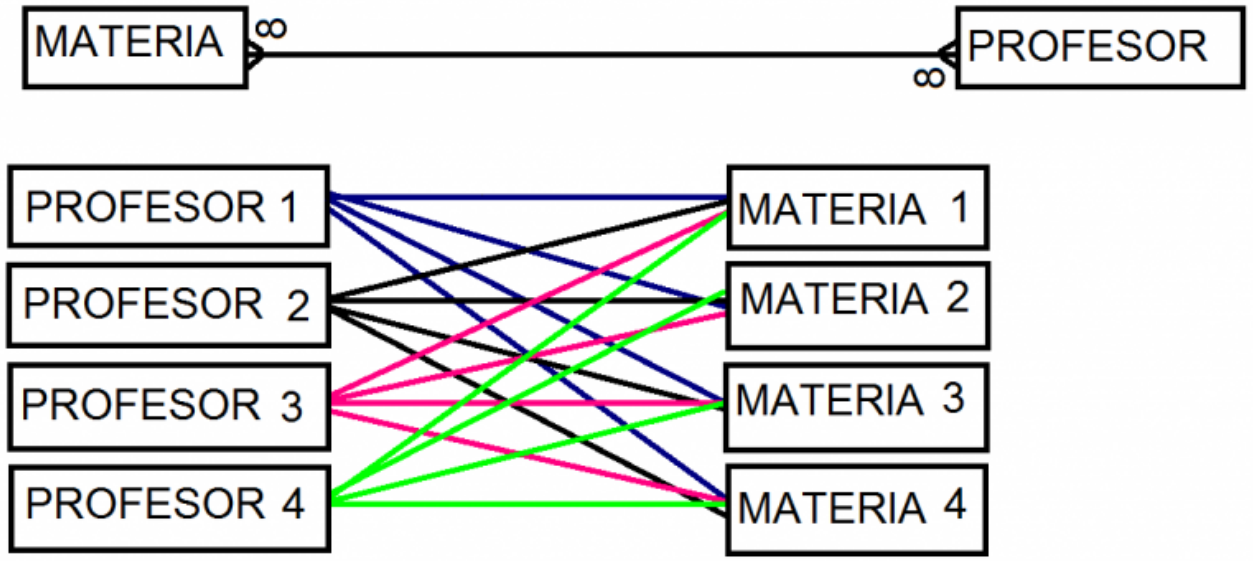
Ejemplo:

Catalogo de productos → Registro de ventas

Asociación Mexicana de Inversionistas

Relaciones varios a varios

Es muy raro e ineficiente utilizar este tipo de relaciones en bases de datos relacionales ya que permite unir con ciertas limitantes tablas que contienen los mismos datos principales en dos o más registros.



Asociación Mexicana de Inversionistas

Fuentes de datos

En la actualidad todo genera datos, la información relevante nace cuando somos capaces de extraer datos y relacionarlos para obtener informes que aporten mayor valor.

¿Cómo iniciar?

Para desarrollar modelos de datos es importante entender la lógica de estos y su finalidad.

1. Generar las definiciones de la estructura de la base de datos.
2. Identificar las variables y sus usos
3. Desarrollar un plan de expansión de datos (uso y almacenamiento)
4. Limpieza y depuración de datos.

¿Cuánto cuesta?

Depende de la dimensión del proyecto y la tecnología asociada, así como los costos operativos, licenciamiento, almacenamiento y de mantenimiento.

Hay opciones que nos permiten generar bases de datos y modelo de datos de forma económica y con gran alcance

Alternativas simples

Hoy en día existen software de código abierto, licenciamiento lite o SAS flexible y económico que nos permiten hacer gran cantidad de actividades ligadas a datos y a su explotación para la toma de decisiones.

Alternativas complejas

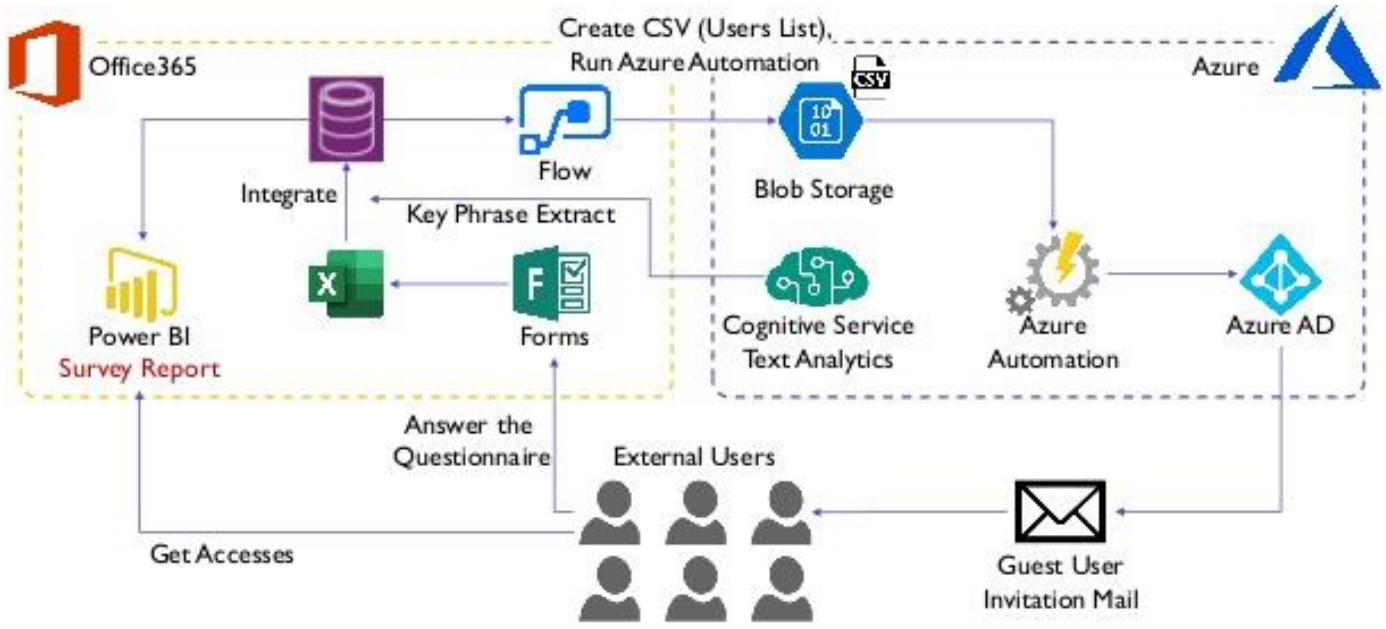
Establecer modelos de datos y gestión de bases de datos más estructuradas y robustas con elementos de seguridad y vínculos con sistemas satelitales que permiten explorar nuevas capas de información.

Ejemplo: Un ERP conectado con aplicaciones de gestión más específicas:

1. CRM
2. WMS
3. S&OP

REST API

Al final de trata de ser creativos



Google Workspace

Si se tienen identificados los alcances del desarrollo y requerimientos clave, se cuenta con datos clasificados y estructurados. Es posible generar alternativas que no representen una alta inversión pero si agregan valor.

1. Microsoft 365 – Dataverse & PowerPlatform
2. G Suite & Workspace
3. Software de código abierto
4. Excel y RPAs

Asociación Mexicana de Inversionistas

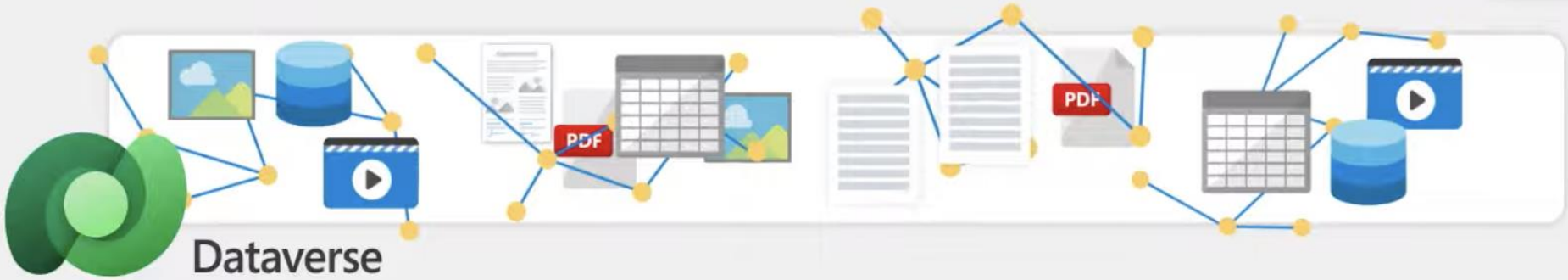
Power BI

Power Apps

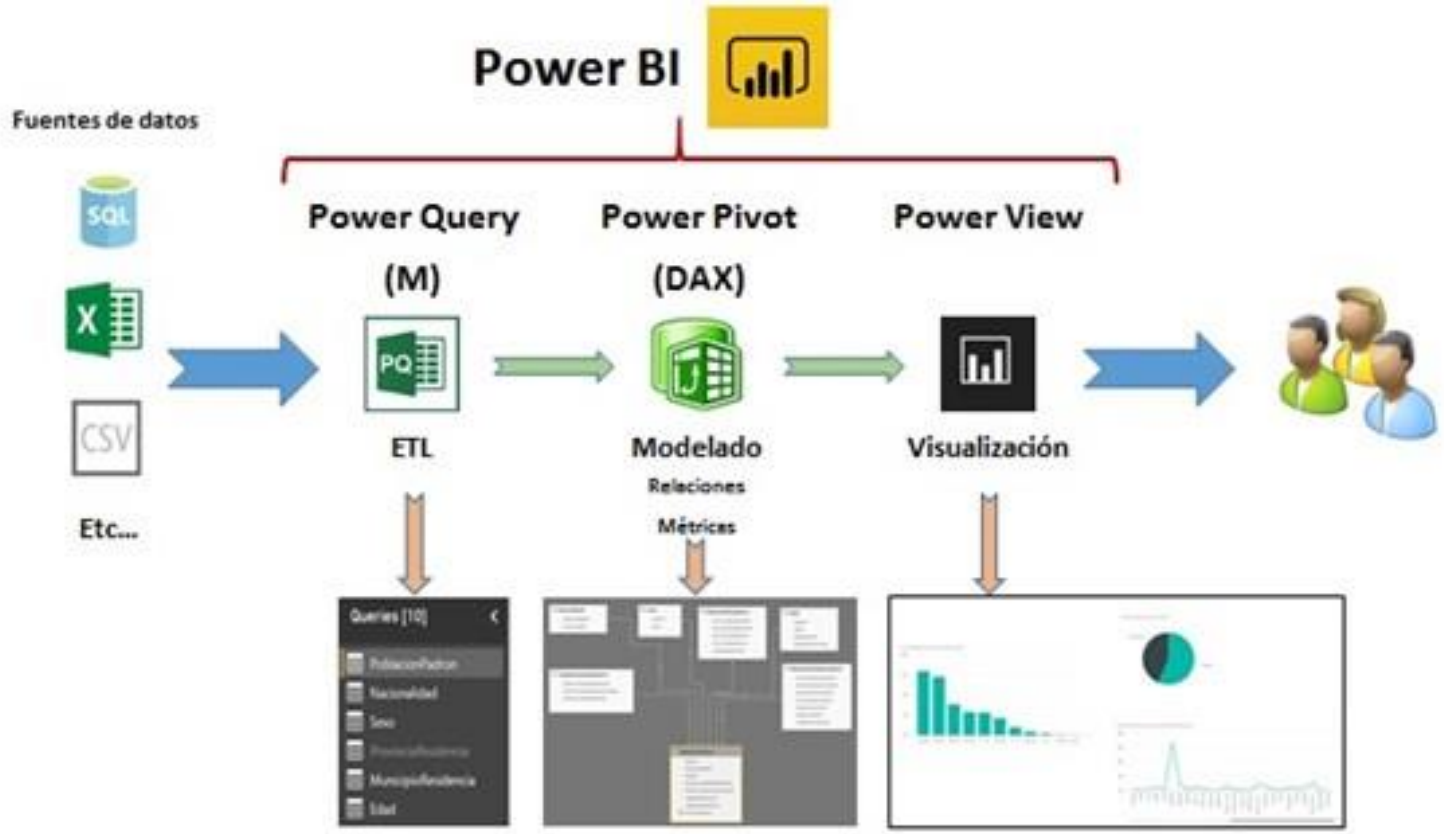
Power Automate

Power Virtual Agents

Power Platform



Modelado de datos básico



El modelar datos es llevarlos de una fuente posiblemente no estructurada a un esquema de base de datos relacional con sus correctas definiciones de variables y conexiones. El objetivo es obtener información relevante.

1. Fuente u origen
2. Limpieza de datos
3. Modelado de datos
4. Conexiones
5. Informes

Matrices dinámicas

Un cambio en el Motor de cálculos internos de Excel permite trabajar y obtener datos provenientes de funciones que hagan relación a un rango o matriz.

1. Con este nuevo motor de cálculo interno, todas las funciones se tratan de la misma forma. Ya no se requiere el uso de Control + Shift + Enter.
2. Aparece un nuevo concepto Intervalo o Rango de Desbordamiento.
   - Si una fórmula devuelve un solo valor, este se mostrará en la celda donde se resuelve la función.
   - Si la fórmula devuelve más de un resultado respuesta, se usará el Rango de desbordamiento para mostrar todos los resultados, el cual puede ocupar varias filas y/o columnas.

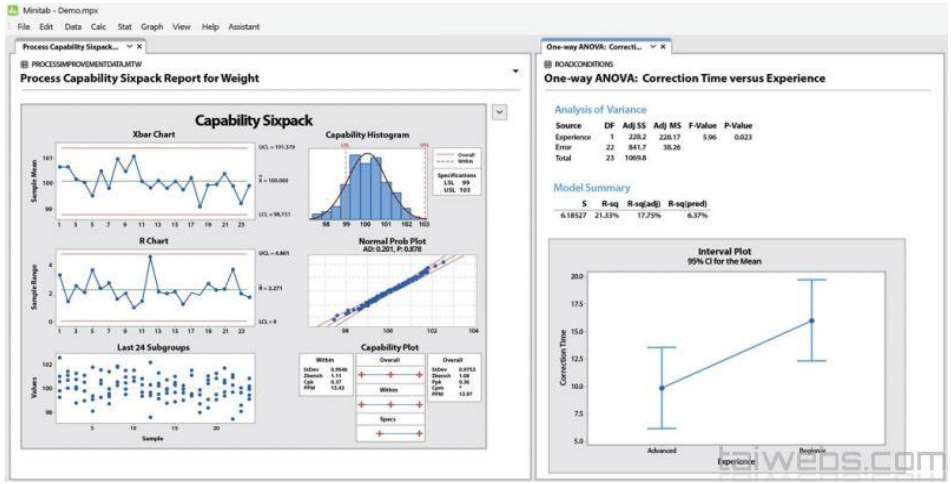
Matrices dinámicas

Funciones y aplicaciones principales

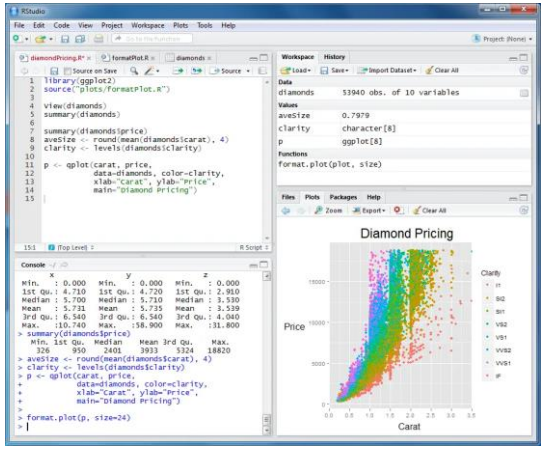
1. Ordenar
2. Ordenar por
3. Únicos
4. Filtrar
5. Secuencias
6. Matrizleat
7. Relleno rápido

Asociación Mexicana de Inversionistas

Estadística descriptiva, analítica y proyectiva



pandas



TensorFlow

Asociación Mexicana de Inversionistas

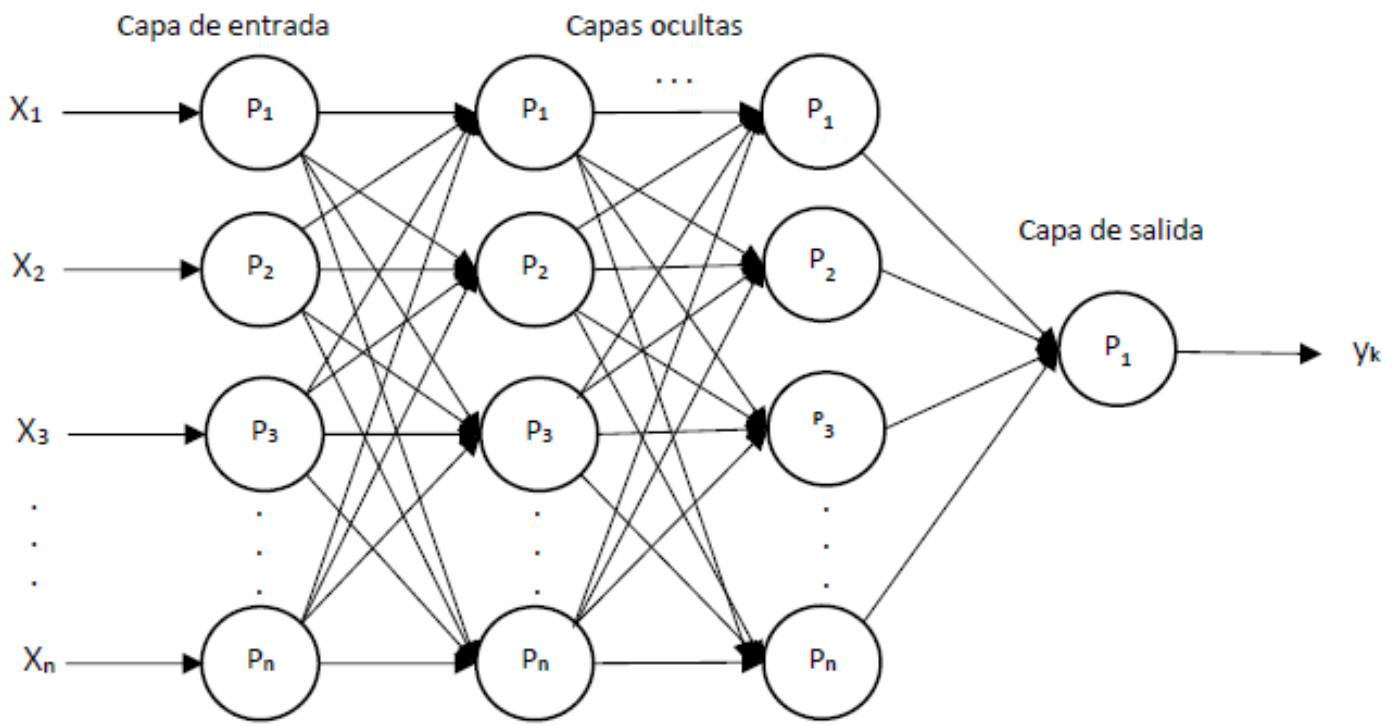
Redes neuronales

Una **red neuronal** es un modelo simplificado que emula el modo en que el cerebro humano procesa la información: Funciona simultaneando un número elevado de unidades de procesamiento interconectadas que parecen versiones abstractas de **neuronas**.

Machine Learning

Se pueden distinguir **tres** tipos de capas:

1. De entrada: reciben datos o señales procedentes del entorno.
2. De salida: proporcionan la respuesta de la **red** a los estímulos de la entrada.
3. Ocultas: no reciben ni suministran información al entorno (procesamiento interno de la **red** y proyección).



BI
KNOWLEDGE MANAGEMENT
BENCHMARKING
DATA MINING
DATA VISUALIZATION
MEASUREMENT ANALYSIS
REPORTING
COLLABORATION PLATFORM

Asociación
Mexicana de
Inversionistas