

Quicksight

Introducción a AWS QuickSight

- Conceptos básicos y características principales
- Integración con el ecosistema de AWS
- Beneficios y casos de uso

Preparación y Transformación de Datos

- Conexión con fuentes de datos (S3, RDS, Redshift, Athena, entre otras.)
- Limpieza y transformación de datos
- Creación de datasets

Visualizaciones Básicas

- Uso de gráficos y tablas en QuickSight
- Tipos de visualizaciones disponibles
- Personalización de gráficos y elementos visuales

Análisis Avanzado y Predictivo

- Uso de "ML Insights" para predicciones automáticas
- Creación de métricas y KPI personalizados
- Incorporación de análisis predictivo en los informes

Creación de Dashboards y Reportes Interactivos

- Diseño de dashboards personalizados
- Uso de filtros y controles interactivos
- Publicación y compartición de dashboards

Colaboración y Automatización

- Compartir informes con equipos y usuarios finales
- Configuración de alertas y reportes automáticos
- Integración con otras herramientas de BI y plataformas externas.

Seguridad del dato en QuickSight Monitorización

Quicksight

Herramienta de Business Intelligence de AWS.

Business Intelligence

Palabrotas en el mundo del dato:

- BBDD: Programas y los datos que gestionan esos programas en los entornos de PRODUCCIÓN de una empresa v ETL: Proceso de extraer datos de una fuente, transformarlos y cargarlos en otra fuente. Procesos online Procesos batch Hay muchas variantes de proceso ETL: ETL: Extraer, Transformar, Cargar ELT: Extraer, Cargar, Transformar ETLT: Extraer, Transformar, Cargar, Transformar TELT: Transformar, Extraer, Cargar, Transformar v Hay poca T... en las ETLs.. guardamos los datos bastante en bruto. v
- Datalake: Es un almacén de datos (muchas veces lo implementamos con un programa de BBDD) históricos y no pensado para la actualización de los datos. Guarda datos en bruto. v ETL: La T es muy fuerte v
- Datawarehouse: Es un almacén de datos históricos, no preparado para su actualización. Pero con una diferencia con respecto al datalake... LOS GUARDO YA CON UN PROPOSITO muy concreto. Muchas veces también usamos programas de gestión de BBDD para almacenar los datos, pero los configuramos de otra forma... y los modelos de datos que montamos son muy diferentes a los de una BBDD de producción (muy orientada a la actualización de los datos)

Los datos cumplen una función principal.. inherente al dato. En ocasiones, podemos usar esos datos para otro tipo de funciones.

| - Business Intelligence Toma de decisiones informadas (basadas en datos) dentro de la empresa | Es la forma más básica de análisis del dato. | Aplicamos técnicas de estadística descriptiva (nivel instituto) | - Se analiza variable a variable y **correlaciones entre variables** | El problema del estudio de relaciones entre datos que hacemos en BI es que | lo hacemos los humanos... y nuestro cerebro está muy limitado | (nos da para pensar en 3/4 variables máximo) | Lo que vamos a querer es usar datos pasados para tomar decisiones | (a día de hoy o de cara al futuro) | - Eso es hacer una predicción: De entrada supongo que el mundo seguirá igual! | Lo que me ofrecen los datos es la capacidad de TASAR la probabilidad de | cometer un error... (para limitarla) | (ojo... siempre y cuando, el mundo siga igual.. que ya es mucho suponer) | Aquí entra el mundo de la estadística INFERENCIAL -----.-----.-----.-----.-----.----- ^^^ como mucho 3/4 variables a la vez en un análisis -----.-----.-----.-----.-----.----- vvv a partir de 4/5 variables... hasta decenas o incluso cientos/miles | - Data mining Es cuando queremos buscar relaciones complejas NO EVIDENTES | entre las variables. NO SE A PRIORI NI LO QUE BUSCO NI LO QUE VOY A ENCONTRAR | - Machine learning Trata de aprender de los datos y hacer predicciones a futuro | - Predecir un dato | - Clasificar un sujeto | - ... v Por orden de complejidad en el análisis de los datos.

- Bigdata
 - Un volumen "grande" de datos.
 - 1 millón de datos es grande?
 - 10 millones de datos es grande?
 - 100 millón de datos es grande?
 - 1000 millón de datos es grande?
 - 100000 millón de datos es grande?

Dónde pongo el corte???

El concepto de bigdata es: Cuando tengo que hacer un trabajo (el que sea) con datos y las técnicas (herramientas de software, computadoras) que hay a mi disposición y que hemos usado tradicionalmente dejan de servirme... tengo un problema. Y entonces busco soluciones alternativas. Los primeros que hicieron esto fue la gente de Google.

Se basa en utilizar granjas de computadoras pequeñas (commodity hardware) y hacer que trabajen como si fueran una. Para qué?

Trabajo el que sea:

- Almacenar un dato
- Analizar un dato
- Transmitir un dato
- ...

4 jugadores: 1 jugador en una partida puede hacer fácilmente 2 movimientos por segundo. De su teléfono deben salir mensajes que al final lleguen a los teléfonos de los otros participantes. 2 mensajes por segundo... por persona... como somos 4 jugando: 24 mensajes por segundo. Mis 2 mensajes emitidos, se convierten en 6 mensajes recibidos. Si tenemos en cuenta que durante muchos años fue el juego más jugado de Apple Store... y de Google play... y que un momento dado podía haber más de 50k partidas en paralelo a nivel mundial: $50k * 24 = 1.2$ millones de mensajes por segundo. ¿Qué computadora es capaz de procesar tal cantidad de mensajes por segundo? NO EXISTE

Si tengo un pincho USB de 16 Gbs... vacío... recién sacado de la caja.. y quiero guardar en él un fichero de 5Gbs, puedo? Depende del formato. FAT16/FAT32... no entra, aunque haya espacio disponible TOTAL NTFS, EXT4.... ahí si... pero incluso esos formatos tienen un límite. Si quiero guardar un fichero de 2Pbs... no puedo... no hay formato tradicional que lo soporte.