

Cloud para Data Science

Herramientas, características y precios



15 de agosto de 2024

Alvaro jesus castro pizaña

Centro universitario UAEM Zumpango

Diagrama

Descripción generada automáticamenteUn **data warehouse** es un repositorio centralizado que agrupa datos de varias fuentes, como sistemas transaccionales, bases de datos relacionales y fuentes externas. Su principal propósito es soportar actividades de inteligencia empresarial (BI), permitiendo análisis profundos, reportes y toma de decisiones.

Algunas de las características que se debe contemplar a la hora de contratar un buen servicio de **data warehouse** podrían ser:

**Procesamiento en tiempo real**: Esto permite a que los data warehouse la integración y análisis de datos en tiempo real, permitiendo obtener conocimientos inmediatos​.

**Integrado**: permite combinar datos de diversas fuentes, asegurando uniformidad en convenciones de nombres, formatos y codificación.

**Variable en el tiempo**: esto es de gran utilidad ya que mantienen datos históricos, permitiendo el análisis a lo largo del tiempo como tendencias y comparaciones.

Algunos de los servicios mas populares son los siguientes:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Servicio** | **Funciones Principales** | **Precios Estimados** | **Ventajas** | **Logo** | **Página Oficial** |
| Amazon Redshift | Almacenamiento y análisis de datos a escala de petabytes. Integración con AWS. Consultas SQL y análisis en tiempo real. | Almacenamiento: $0.024 por GB/mes. Consultas: $0.25/hora por nodo. | Escalable, altamente integrado con AWS, costo flexible. | Logotipo  Descripción generada automáticamente | <https://aws.amazon.com/redshift/> |
| Google BigQuery | Servicio sin servidor completamente gestionado. Separación de almacenamiento y cómputo. Soporte para SQL y ML integrado. | Almacenamiento: $0.02 por GB/mes. Consultas: $5 por TB procesado. | Modelo sin servidor, ideal para análisis en tiempo real, alto rendimiento. | Imagen que contiene Aplicación  Descripción generada automáticamente | <https://cloud.google.com/bigquery> |
| Snowflake | Arquitectura de procesamiento en paralelo. Multi-cloud. Integración de datos estructurados y semiestructurados. | Almacenamiento: $23 por TB/mes. Cómputo: Pago por segundo de uso. | Multi-cloud, escalabilidad flexible, soporte para datos complejos. | Logotipo, nombre de la empresa  Descripción generada automáticamente | [https://www.snowflake.com](https://www.snowflake.com/) |
| Azure Synapse Analytics | Integración con Azure. Mezcla datos de Data Lakes y Data Warehouses. Consultas SQL y no SQL. | Almacenamiento: $23 por TB/mes. Cómputo: $1.20 por hora por DWU. | Altamente integrado con Azure, ideal para grandes volúmenes de datos. | Imagen que contiene Icono  Descripción generada automáticamente | <https://azure.microsoft.com/en-us/services/synapse-analytics/> |
| IBM Db2 Warehouse | Almacenamiento de datos en la nube y análisis avanzado. Soporte para SQL y ML. Optimizado para alto rendimiento. | Almacenamiento: $0.65 por GB/mes. Cómputo: Pago por instancias usadas. | Alto rendimiento, seguridad avanzada, optimizado para análisis. | Un conjunto de letras blancas en un fondo blanco  Descripción generada automáticamente con confianza media | <https://www.ibm.com/cloud/db2-warehouse> |
| Oracle Autonomous Data Warehouse | Servicio autónomo en la nube con IA. Escalabilidad automática. Seguridad avanzada y recuperación automática. | Almacenamiento: $0.016 por GB/mes. Cómputo: $0.1125 por OCPU/hora. | Escalabilidad y seguridad automáticas, ideal para cargas de trabajo variables. | Imagen que contiene nombre de la empresa  Descripción generada automáticamente | <https://www.oracle.com/autonomous-database/> |

"What is a Data Warehouse? - Data Warehouse Explained," AWS Documentation, [En línea]. Available: https://aws.amazon.com/data-warehouse/. [Consultado el 14 de Ago. 2024].

S. Dutta, "Characteristics of Data Warehouses: Key Features for Business Intelligence," Sprinkle Data, Jul. 24, 2023. [En línea]. Available: https://www.sprinkledata.com/blogs/characteristics-of-data-warehouse-understanding-the-core-features. [Consultado el 14 de Ago. 2024].

J. Richman, "Modern Data Warehouses: Functions, Architecture, & Examples," Estuary, [En línea]. Available: https://estuary.dev/modern-data-warehouse. [Consultado el 14 de Ago. 2024].