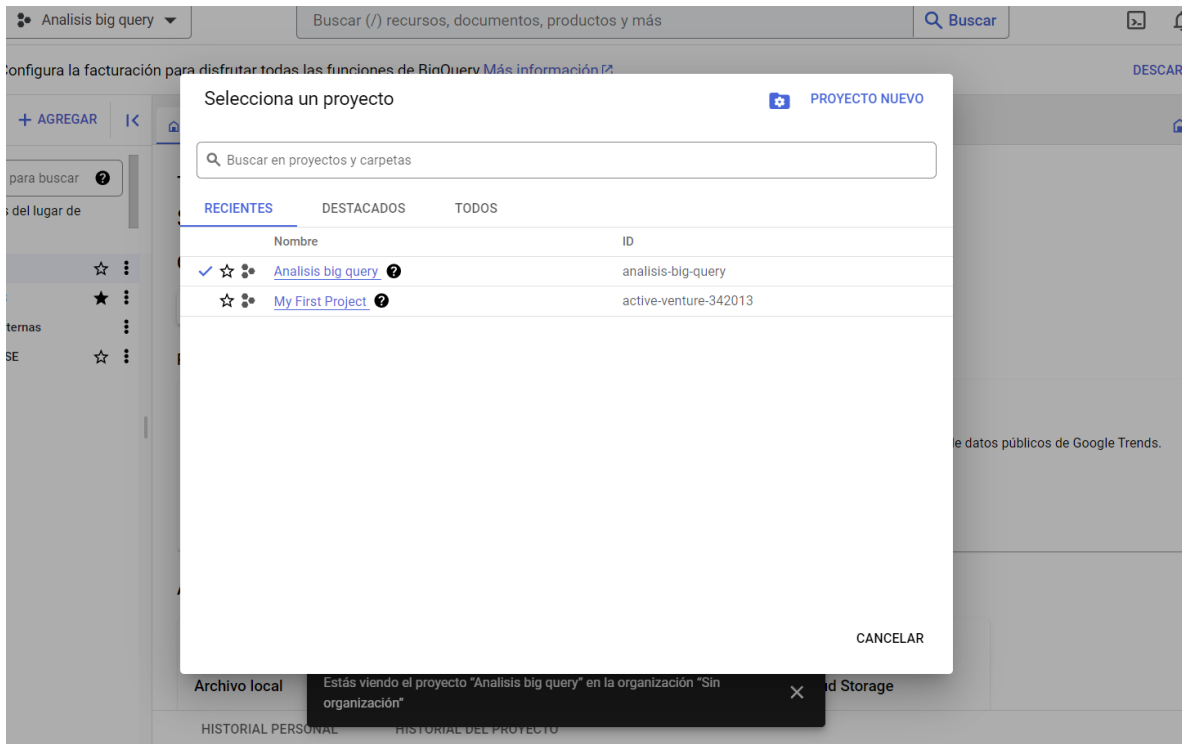


<https://www.kaggle.com/datasets/harlfoxem/housesalesprediction/discussion/23194>

Big query es un almacén de datos en Google, totalmente administrado, que permite extraer analíticas de petabytes de datos, no es necesario gestionar la infraestructura ni contar con un administrador de bases de datos, de esta manera se puede centrar en obtener información valiosa del conocido lenguaje SQL y aprendizaje automático

Se tiene que crear o utilizar un proyecto existente



Después seleccionar crear un conjunto de datos

The screenshot shows the Google Cloud BigQuery Explorer interface. At the top, there's a 'ZONA DE PRUEBAS' (Test Zone) label and buttons for 'DESCARTAR' (Dismiss) and 'ACTUALIZAR' (Update). The main interface is divided into two panels. The left panel, titled 'Explorador' (Explorer), contains a search bar with the placeholder 'Comienza a escribir para buscar' (Start typing to search) and a list of resources. The resources listed are 'analisis-big-query', 'active-venture-342013', 'Conexiones externas' (External connections), and 'NEW\_DATABASE'. The right panel displays a welcome message 'Te damos la bienvenida al lugar de trabajo de SQL' (We welcome you to the SQL workspace) and a 'Comenzar' (Get started) section. This section includes a 'Prueba la consulta de demostración de Google Trends' (Try the Google Trends demo query) button. A context menu is open over the 'active-venture-342013' resource, showing options to 'Crear un conjunto de datos' (Create a dataset) and 'Actualizar contenido' (Update content).

Al momento de crear un conjunto de datos se le tiene que dar un nombre, un lugar donde estarán tus datos y ajustar algunas otras opciones.

**Crea un conjunto de datos**

ID del proyecto  
analisis-big-query CAMBIO

ID del conjunto de datos \*  
Datos\_prueba  
Puede incluir letras, números y guiones bajos

Tipo de ubicación ?  
☐ Región  
Especifica una región para colocar tus conjuntos de datos con otros servicios de Google Cloud.  
☒ Multirregional  
Permite que BigQuery seleccione una región de un grupo para alcanzar límites de cuota más altos.

Multirregional \*  
US (varias regiones en Estados Unidos) ▼

Vencimiento predeterminado de la tabla  
☐ Habilitar el vencimiento de la tabla ?  
Máxima antigüedad predeterminada de la tabla Days

Opciones avanzadas ^

Encriptación ?  
☒ Clave de encriptación administrada por Google  
No se requiere configuración  
☐ Clave de encriptación administrada por el cliente (CMEK)  
Se administra a través de [Google Cloud Key Management Service](#)

**CREAR CONJUNTO DE DATOS** CANCELAR

Después tenemos que crear una tabla en el conjunto de datos que acabamos de crear, en la tabla podremos subir nuestros datos, elegir en donde los vamos a almacenar, si deseamos que se seleccione el esquema de manera automática, y de que manera se realizaran las particiones y clústeres (para mejorar la velocidad de las consultas).

Big query vs Mysql

Big query in minute

<https://www.youtube.com/watch?v=CFw4peH2UwU>

+Big query es la empresa de nube en Google data warehouse, ellos toman cuidado de la infraestructura

+Diseñada para ingestar, almacenar, analizar y visualizar data hasta los petadatos.

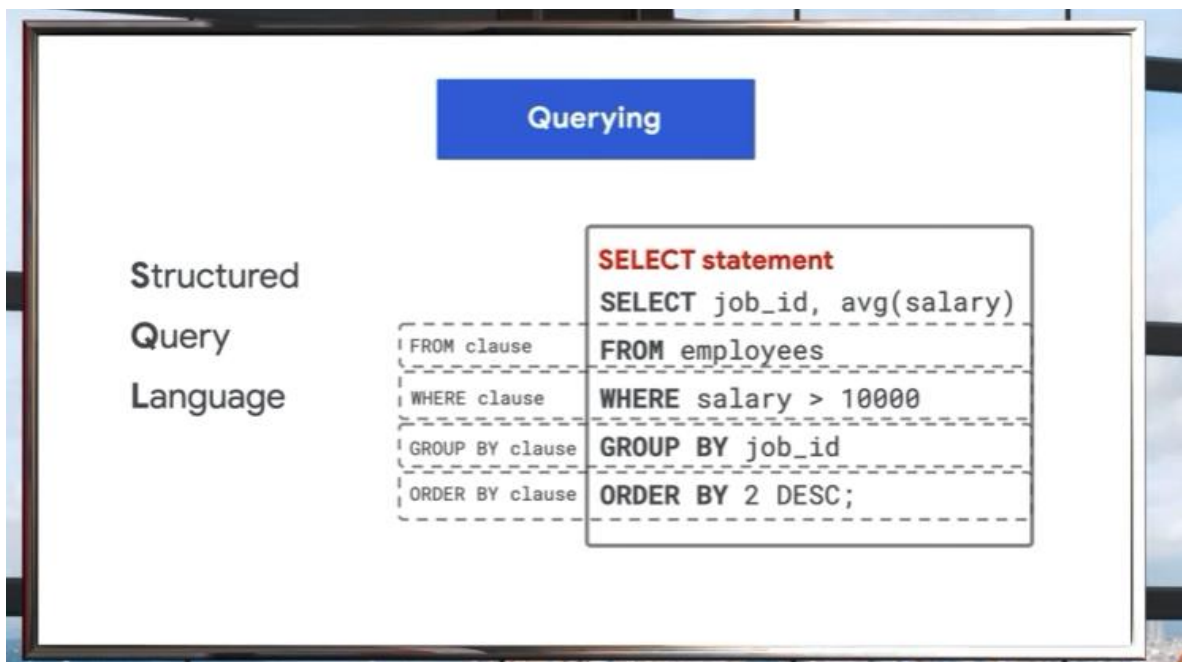
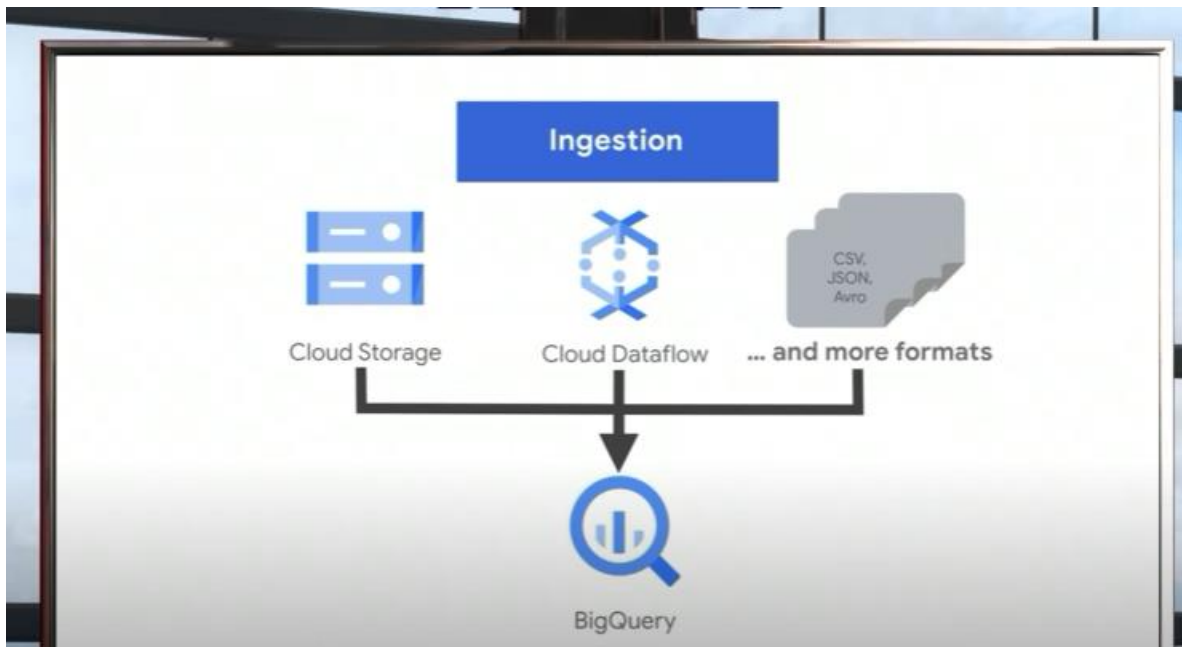
+Soporta SQL

- +Puede ingresarse por batch o ingresar datos directamente
- +Se puede interactuar con la consola UI , con la consola de comandos, o utilizar la API
- +Integra además grandes herramientas para visualizar datos de manera sencilla
- +Se paga por los datos guardados, por la información que se sube mediante streaming insert y por la información que se consulta, la información que se sube es gratis

## Big query

- +large scale data analyze
- + Cuando tus datos crecen del tamaño grande como gigabytes, terabytes, etc necesitas un dataware house, información que es inútil si no se puede analizar
- +Ejemplos del tipo de datos que se pueden analizar son por ejemplos logs de los sistemas, o datos de los sensores de los vehículos a lo largo del mundo.
- +Es totalmente manejada y serverless dataware house
- +Evita el problema de los data silo que son creados por grupos individuales que tienen sus data marts, también evitando los conflictos entre el versionado de los distintos grupos analíticos
- +Se pueden asignar accesos a usuarios, grupos o proyectos manteniendo segura la información
- +fully management services implica que no se tiene que instalar nada o tener un administrador de base de datos
- + los datos están guardados en en estructuras de tablas, replicadas y distribuidas





Introduction to BigQuery → What is BigQuery?

## Benefits of using BigQuery

Query results SAVE RESULTS EXPLORE DATA

JOB INFORMATION RESULTS JSON EXECUTION DETAILS

For help debugging or optimizing your query, check our documentation. [Learn more.](#)


Elapsed time	Slot time consumed	Bytes shuffled	Bytes spilled to disk
1 min 48 sec	29 min 24 sec	33.54 GB	0 B

SELECT \* FROM Table <10GB Storage 19,176,586 Rows x 60 Columns == 1,140,000,000 Cells Free

Slot time consumed es el tiempo usado en una cpu virtual, trabajando en paralelo , se tardaría 29 min trabajando en un solo core, es escalable

Introduction to BigQuery → What is BigQuery?

## Benefits of using BigQuery

Query results SAVE RESULTS 

JOB INFORMATION RESULTS JSON EXECUTION DETAILS

For help debugging or optimizing your query, check our documentation. [Learn more.](#)

Elapsed time	Slot time consumed	Bytes shuffled	Bytes spilled to disk
4 hr 26 min	1,156 days 11 hr	85.3 TB	36.57 TB

100 dolares

<https://www.youtube.com/watch?v=NqTdOfXsROk>



What is Google BigQuery - Explained under 3 minutes [2021]

**IF THIS IS THE CASE, GOOGLE BIGQUERY IS A GREAT TOOL WHICH SOLVES THE PURPOSE OF STORING AND ANALYZING THE MASSIVE AND RAPIDLY GROWING DATA.**





**"GOOGLE BIGQUERY IS ONE  
OF THE BEST TOOL FOR  
BIG DATA ANALYSIS"**

[Subscribe](#)

Google  
BigQuery



**ACCESS  
MECHANISM**



**"IDENTITY AND ACCESS MANAGEMENT"**



read&write

[Subscribe](#)



## 3 PRIMARY PARTS OF GOOGLE BIGQUERY

