

# Sprint #2

## Data Engineering



- ETL
- PIPELINE
- VISUALIZACION
- MODELO ML



# ETL

## Estático, Google

**Área** de inversión "hoteles": carga completa, filtrado y normalizado en sus datos.

**Ubicación:** eliminamos columnas no necesarias y normalización.

**Tiempo:** normalización del tipo de dato.

**Review:** optimización de los review, análisis de sentimiento



## ESTRUCTURA DE DATOS IMPLEMENTADA

### **-Escalabilidad y Administración Simplificada:**

se encarga de las tareas operativas, como aprovisionamiento de hardware, escalado automático, copias de seguridad y actualizaciones de software.

Rápido aprovisionamiento de la base de datos.

**-Rendimiento y Fiabilidad:** Alto rendimiento y confiabilidad con infraestructura redundante

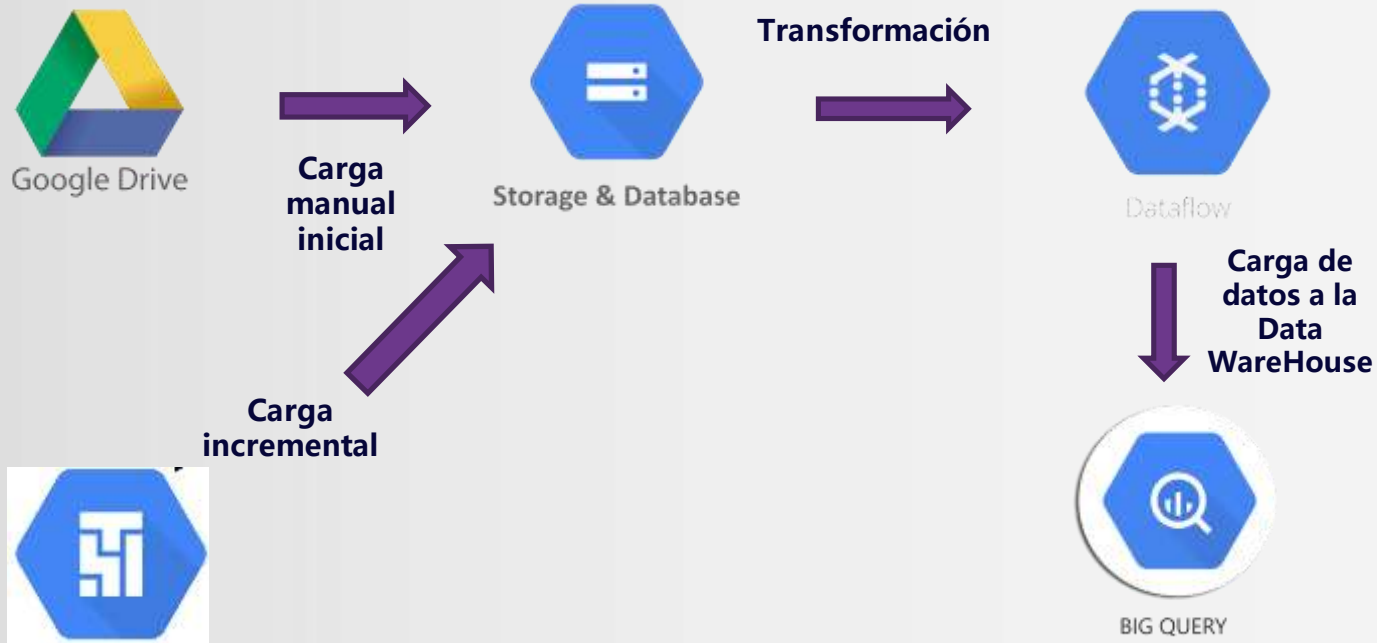
**-Durabilidad y disponibilidad:** Durabilidad y alta disponibilidad de los datos

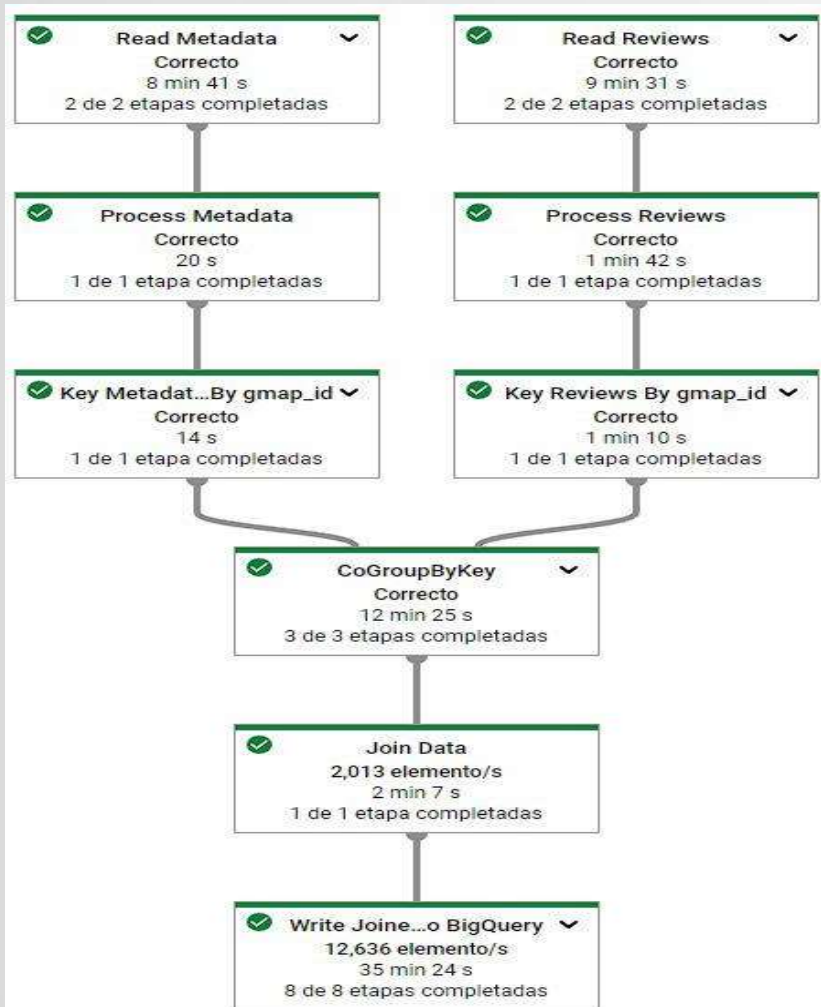
**-Seguridad:** Encriptación de datos en reposo y en tránsito, gestión de accesos y redes privadas virtuales (VPC).



Google Cloud

# DIAGRAMA DEL FLUJO DE TRABAJO





## Diagrama de flujo pipeline

# WORKFLOW TECNOLOGIAS



Google Maps



Dataflow



BIG QUERY



Looker

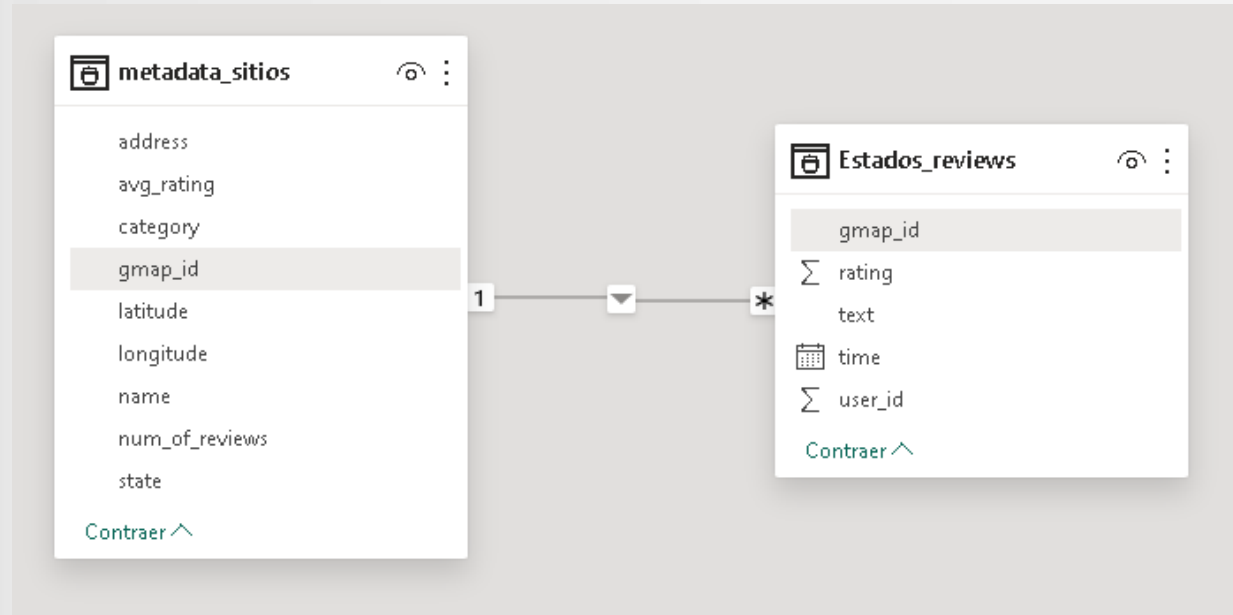


Streamlit

# Diseño del Modelo Entidad - Relación

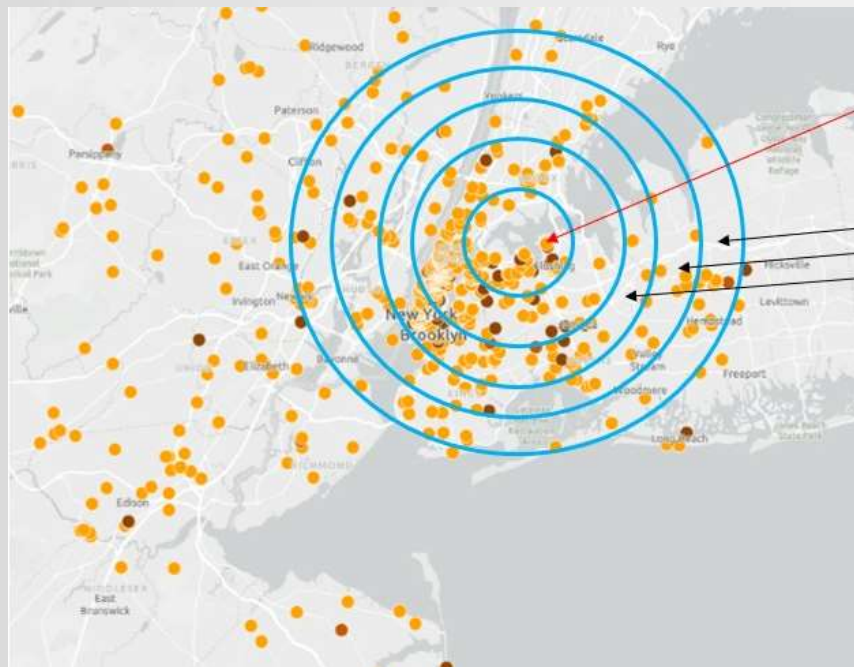
Nuestras entidades se relacionan con **gmap\_id**. Esto significa que cada sitio tiene un identificador único en Google Maps asociado.

Se muestran las diferentes entidades, atributos y relaciones que existen entre los metadatos de los sitios web, los estados de las reseñas y los identificadores de Google Maps.





# Modelo de Machine Learning



Estadio

Zonas de inversión alrededor del estadio

Dispersión por zona: 3,1 millas

Total zonas por estadio : 10

Total zonas para el modelo: 110

## Variables por zona para el modelo ML

1. Densidad de hoteles por milla cuadrada
2. Densidad de negocios por milla cuadrada
3. Calificación de rating en hoteles
4. Sentimiento de hoteles por zona

# Looker Studio

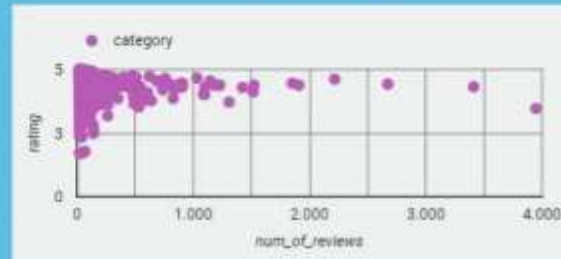
## opportunity hunters



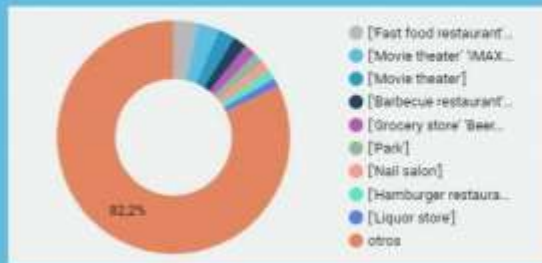
### Categorías



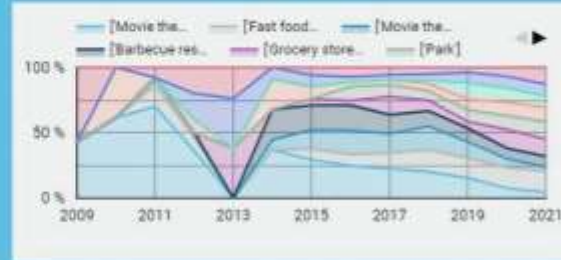
### Reviews



### Participación



### Categorías



# ¡Gracias!



**Más que un servicio, una experiencia**