



PLATAFORMA TECNOLOGICA

Google Cloud para almacenar y gestionar nuestros datos

SERVICIOS AI - CRECIMIENTO - CONFIANZA

Google BigQuery como nuestro sistema de almacenamiento en la nube

ESCALABILIDAD - RENDIMIENTO - FACILIDAD USO

PROCESO DE AUTOMATIZACION

Google DataFlow para automatizar ETL

CALIDAD DE DATOS - EFICIENCIA - MANTENIMIENTO

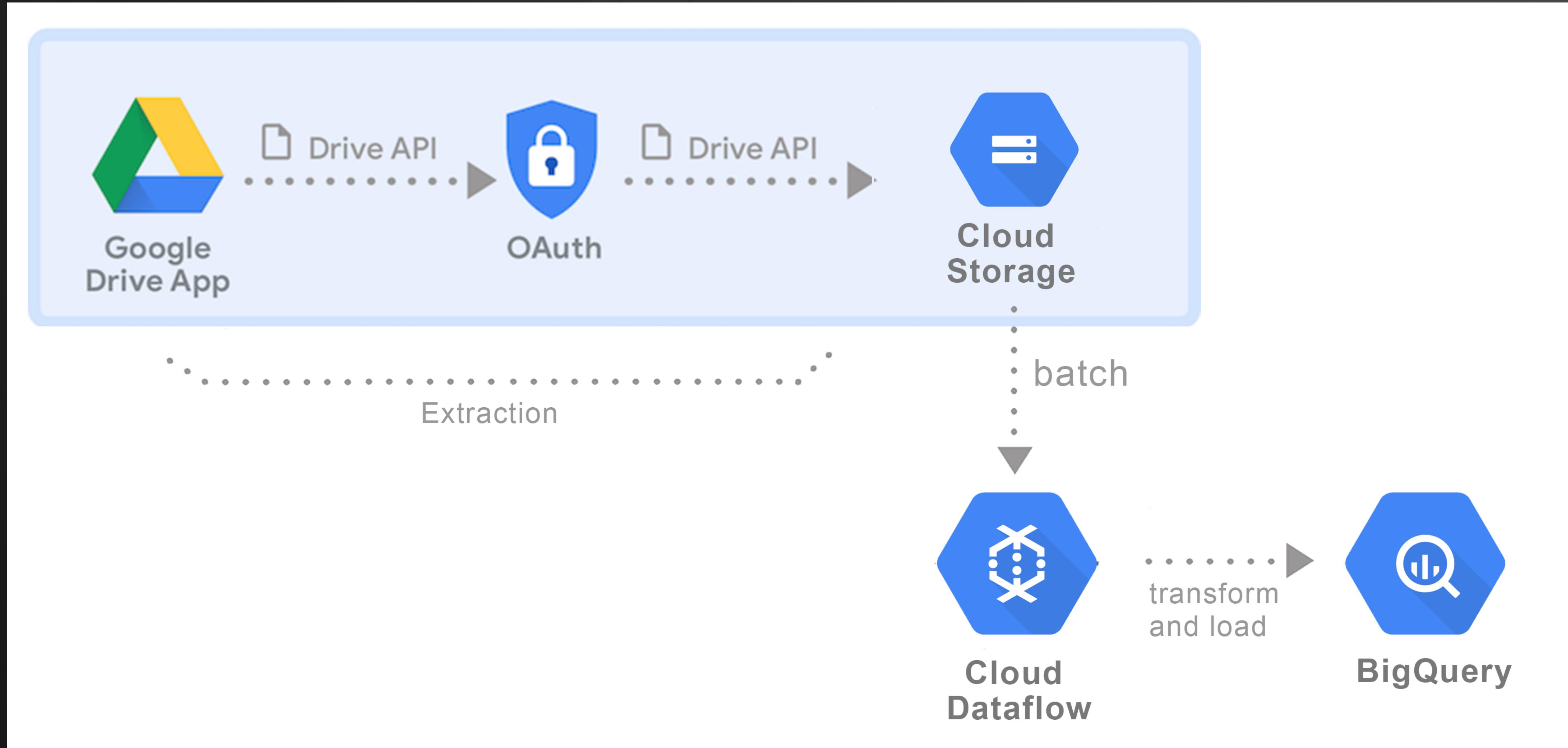
ESTRUCTURA DE DATOS Y RELACION

documentación y comprensión de nuestros datos y las relaciones existentes entre ellos



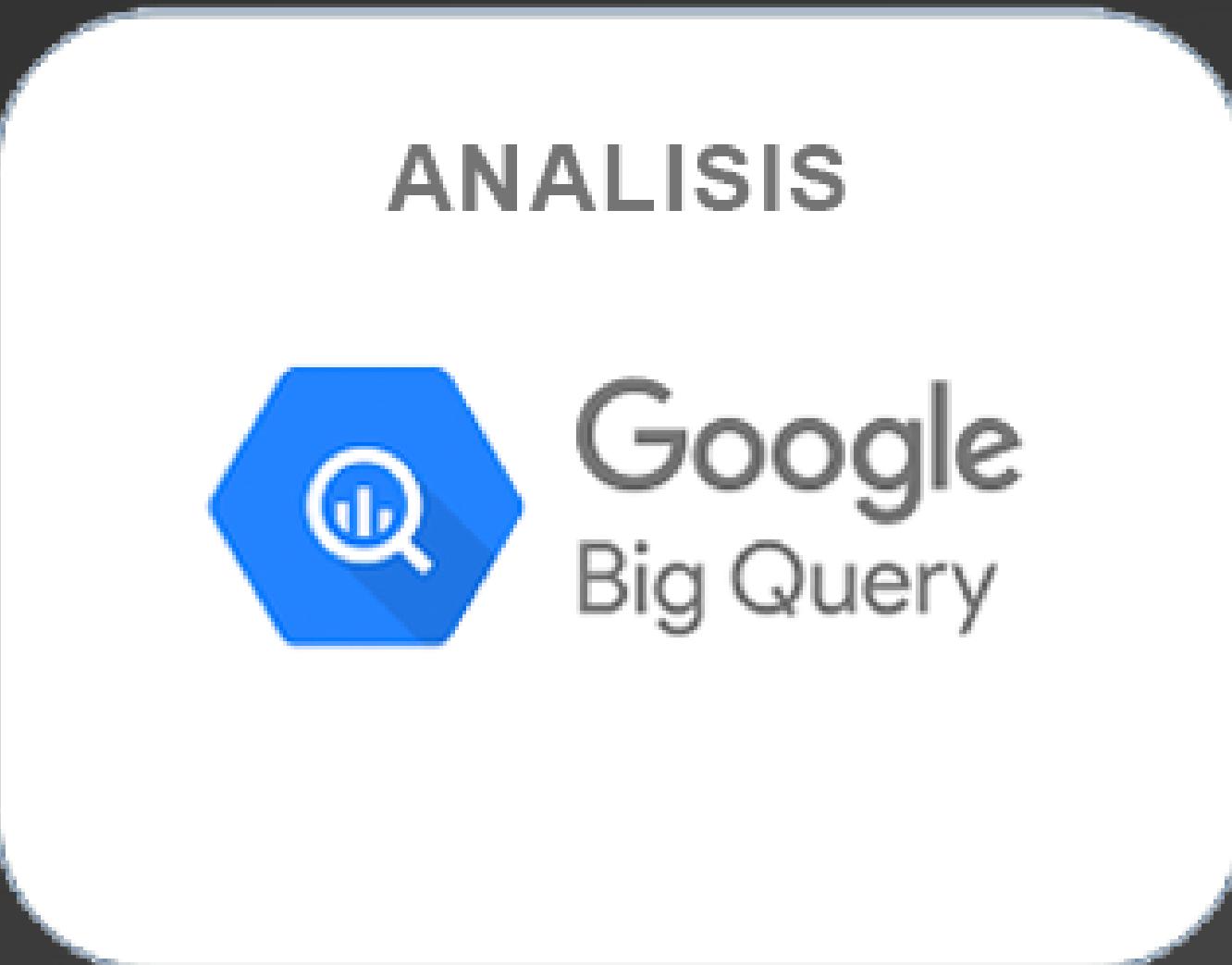
FLUJO DE TRABAJO

Google Cloud Platform



AUTOMATIZACIÓN

ETL





DATAWAREHOUSE



Google Cloud My First Project ▾ Buscar (/) recursos, documentos, productos y más

Explorador +AGREGAR ▾ gmap_metadata ▾ +

Comienza a escribir para buscar ?

Se muestran los recursos del lugar de trabajo.

cedar-pottery-388916

- Conexiones externas
- conjunto_2
 - gmap_metadata
 - gmap_reviews
 - yelp_business
 - yelp_reviews
 - yelp_user

gmap_metadata

ESQUEMA DETALLES VISTA PREVIA LINAJE

Filtro Ingresar el nombre o el valor de la propiedad

Nombre del campo	Tipo	Modo	Clave	Intercalación	Valor predeterminado	Etiquetas de proyecto
name	STRING	NULLABLE				
address	STRING	NULLABLE				
gmap_id	STRING	NULLABLE				
latitude	FLOAT	NULLABLE				
longitude	FLOAT	NULLABLE				
category	STRING	NULLABLE				
avg_rating	FLOAT	NULLABLE				
num_of_reviews	INTEGER	NULLABLE				
hours	STRING	NULLABLE				
MISC	STRING	NULLABLE				
state	STRING	NULLABLE				
relative_results	STRING	NULLABLE				
url	STRING	NULLABLE				

VENTAJAS DE BIGQUERY

ESCALABILIDAD

RENDIMIENTO

FACILIDAD DE USO

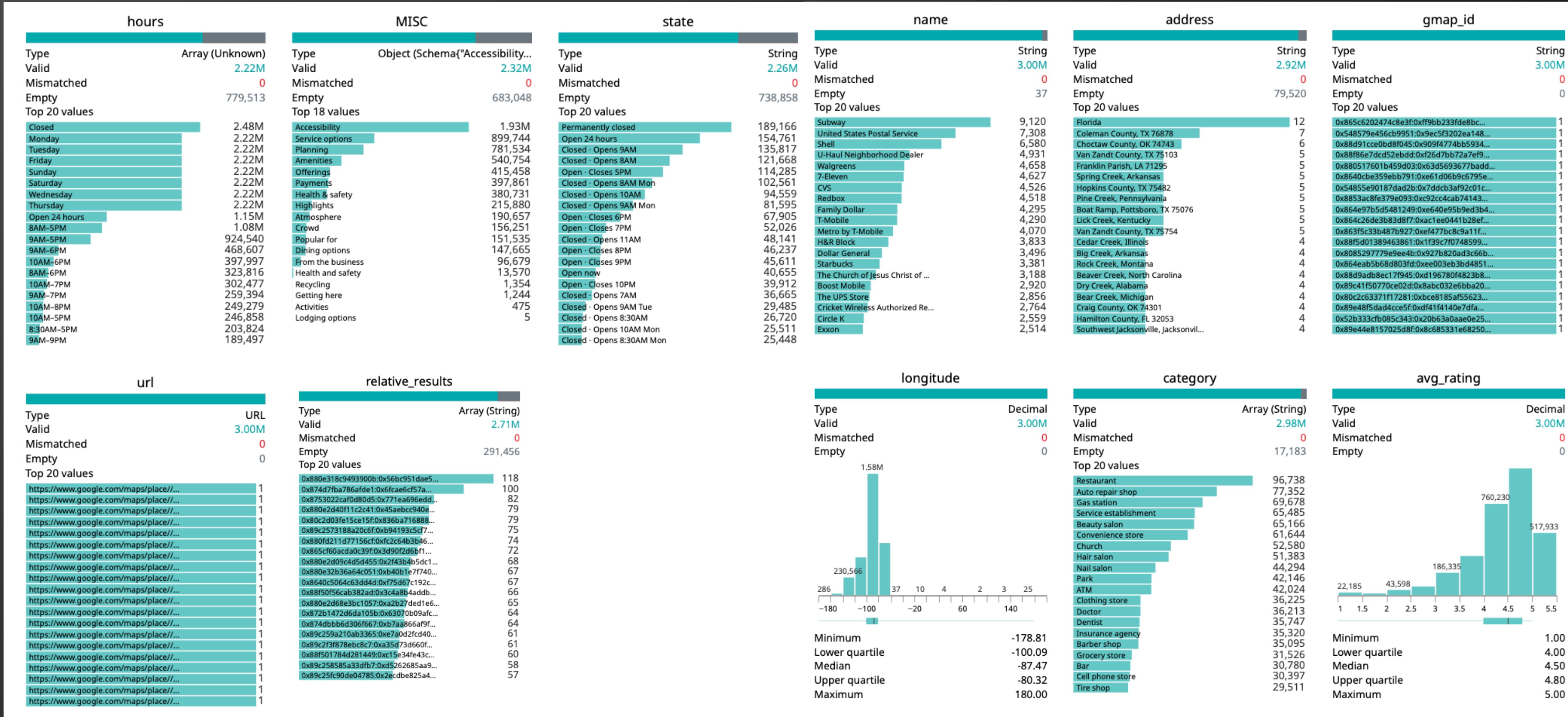
INTEGRACIÓN CON EL ECOSISTEMA
DE GOOGLE CLOUD

PRECIO BASADO EN EL CONSUMO

BigQuery se integra con herramientas populares de análisis y visualización de datos, como Google Data Studio, lo que facilita el análisis y la generación de informes a partir de los datos almacenados.

EDA

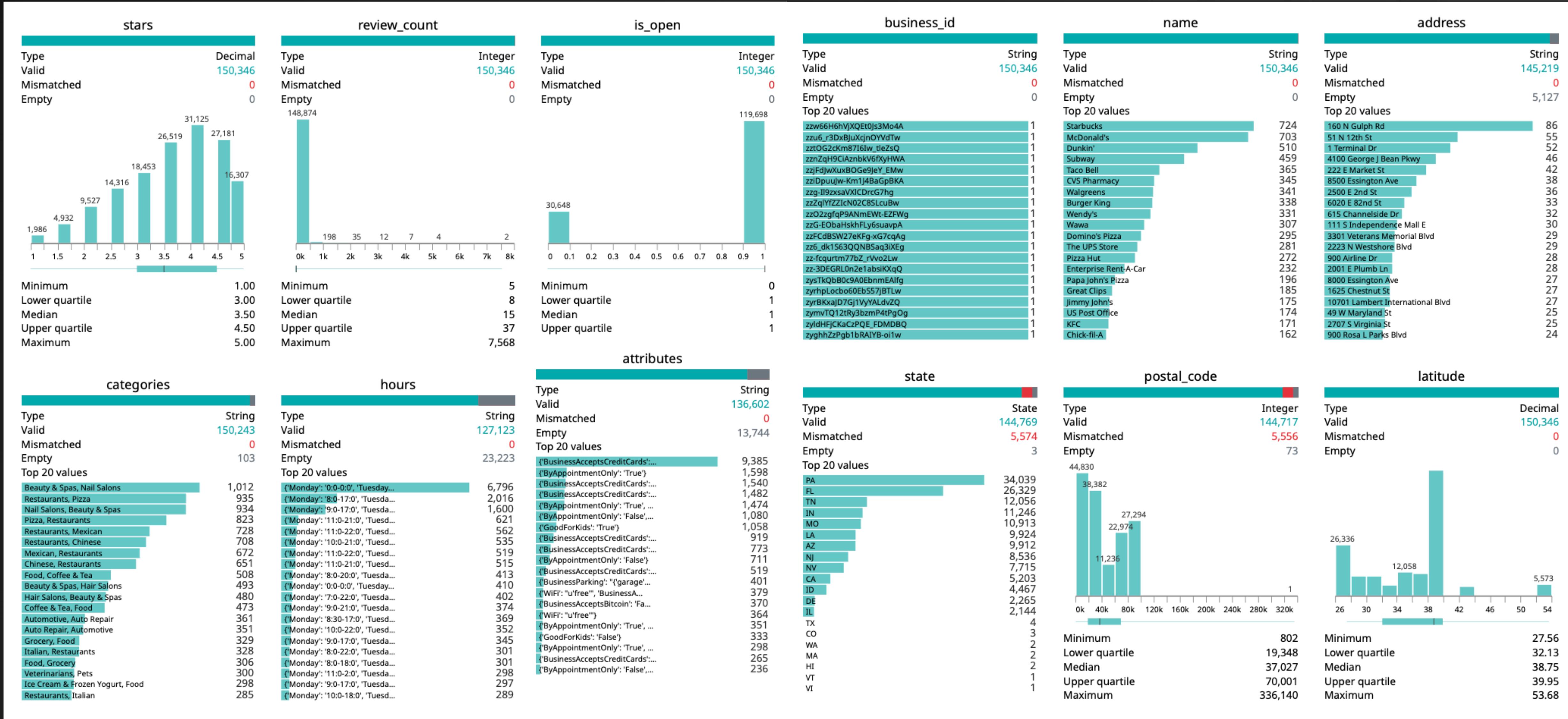
Google Maps dataset metadata-sitios





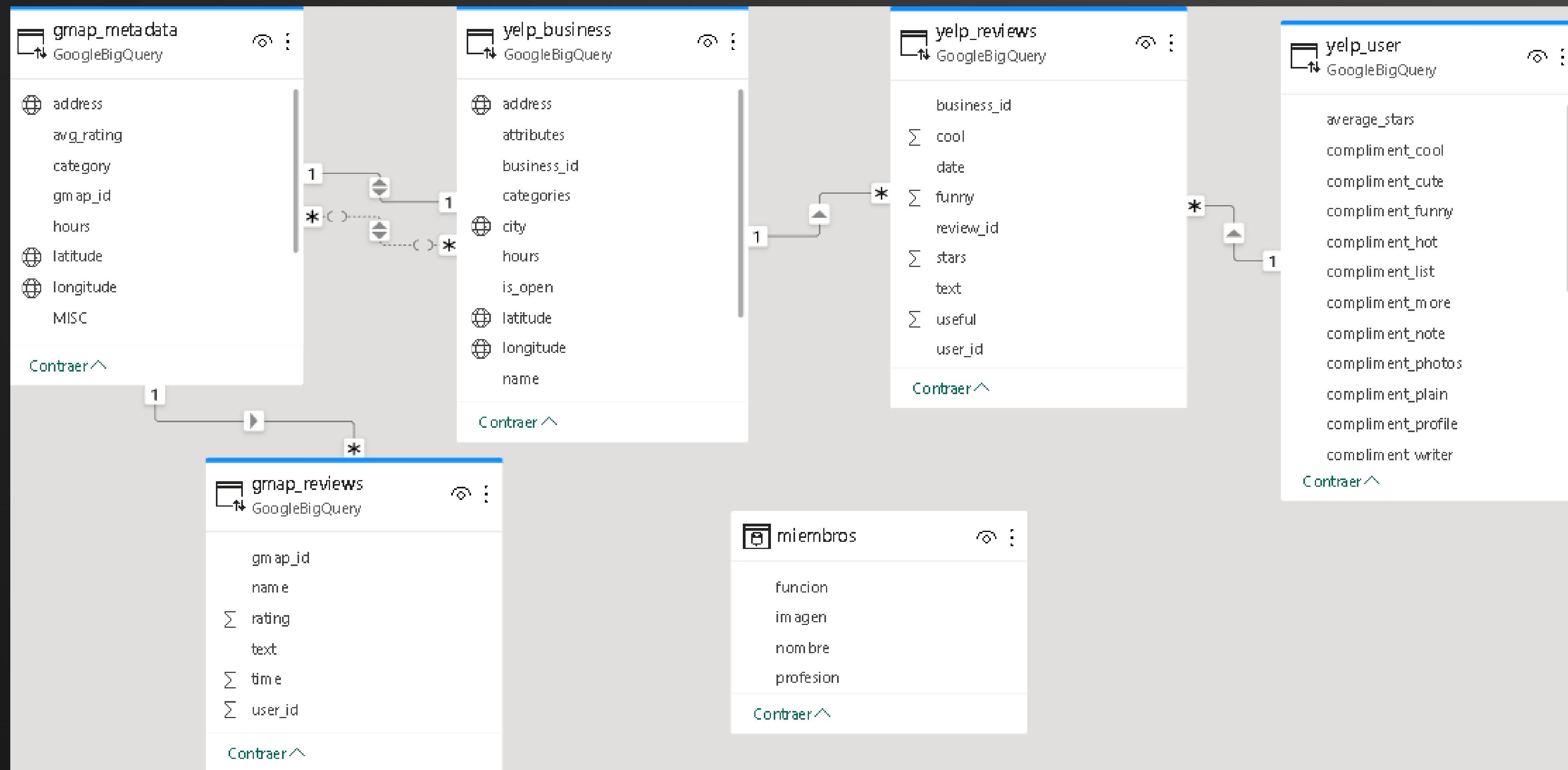
EDA

Yelp dataset business





DISEÑO MODELO ENTIDAD - RELACIÓN



✓ DICCCIONARIO DE DATOS

✓ Google Maps metadata-sitios

'name': String, Nombre del local.

'address': String, Nombre del establecimiento, Número y nombre de la calle,Ciudad.

'gmap_id': String, ID único de la ubicación en Google Maps.

'description': Breve descripción del comercio.

'latitude': Float, Latitud

'longitude': Float, Longitud

'category': String, Categoría que clasifica el tipo de comercio.

'avg_rating': Float, Promedio del puntaje de las reseñas.

'num_of_reviews': Entero, Cantidad de reseñas.

'price': Entero, Precios.

'hours': String, dia/hora de atención

'MISC': Detalles adicionales:

 '**'Service options'**: String, Opciones de servicio.

 '**'Health & safety'**: String, Protocolos sanitarios.

 '**'Accessibility'**:String, Accesibilidad a nivel físico.

 '**'Planning'**: String, Tipo de plan semejante.

 '**'Payments'**: String, Tipo de pago disponible.

'state': String, Estado de funcionamiento.

'relative_results': String, códigos de ubicación geográfica.

'url': String, Link de la búsqueda en Google Maps.

✓ Google Maps metadata-reviews

'user_id': String, ID único de usuario.

'name': String, Nombre y Apellido.

'time': Date, fecha de la reseña.

'rating': Entero, Número de calificación entre 1 y 5.

'text': String, Comentario de la reseña.

'pics': String, Lista de imágenes asociadas a la reseña vinculada cada una a una URL.

'resp': String, Respuesta proporcionada por el propietario del negocio o entidad relacionada con la reseña

'gmap_id': String, ID único de la ubicación en Google Maps.





DICCIONARIO DE DATOS

↗ Yelp! business.pkl

"business_id": string, 22 caracteres id del negocio
"name": string, nombre del negocio
"address": string, dirección completa del negocio
"city": string, ciudad
"state": string, código de 2 letras del Estado donde se ubica el negocio
"postal code": string, el código postal
"latitude": float, latitud
"longitude": float, longitud
"stars": float, rating en estrellas,
"review_count": entero, número de reseñas
"is_open": entero, 0 si está cerrado, 1 si está abierto
"attributes": objeto, atributos del negocio como valores.
Algunos valores de atributos también pueden ser objetos.
"categories": lista de categorías de los negocios
"hours": objeto, dia/hora de atención

↗ Yelp! checkin.json

Registros en el negocio
"business_id": string, 22 caracteres id del negocio
"date": string que es una lista de fechas separados por coma, en formato YYYY-MM-DD
HH:MM:SS

↗ Yelp! review.json

"review_id": string, 22 caracteres id de reseña
"user_id": string, 22 caracteres id único de usuario, refiere al usuario en user.json
"business_id": string, 22 caracteres id del negocio, refiere al negocio en business.json
"stars": entero, puntaje en estrellas de 1 al 5
"date": string, fecha formato YYYY-MM-DD
"text": string, la reseña en inglés
"useful": entero, números de votos como reseña útil
"funny": entero, número de votos como reseña graciosa
"cool": entero, número de votos como reseña cool.

↗ Yelp! tip.json

Tips escritos por el usuario. Los tips son más cortas que las reseñas.
"text": string, texto del tip
"date": string, fecha cuando se escribió el tip
YYYY-MM-DD
"compliment_count": entero, cuantos cumplidos totales tiene
"business_id": string, 22 caracteres, id del negocio que se refiere al negocio en business.json
"user_id": string, 22 caracteres de id de usuario, que se refieren al usuario en user.json

↗ Yelp! user.parquet

Data del usuario incluyendo referencias a otros usuarios amigos y a toda la metadata asociada al usuario.
"user_id": string, 22 caracteres, id de usuario que refiere al usuario en user.json
"name": string, nombre del usuario
"review_count": entero, número de reseñas escritas
"yelping_since": string, fecha de creación del usuario en Yelp en formato YYYY-MM-DD
"friends": lista con los id de usuarios que son amigos de ese usuario
"useful": entero, número de votos marcados como útiles por el usuario
"funny": entero, número de votos marcados como graciosos por el usuario
"cool": entero, número de votos marcados como cool por el usuario
"fans": entero, número de fans que tiene el usuario
"elite": lista de enteros, años en los que el usuario fue miembro elite
"average_stars": float, promedio del valor de las reseñas



✓ NUESTRO DATAWAREHOUSE ESTÁ CONSTRUIDO EN GOOGLE BIGQUERY Y SIGUE UNA ESTRUCTURA ORGANIZADA CON TABLAS RELACIONADAS Y UN ESQUEMA DEFINIDO. UTILIZAMOS TRANSFORMACIONES PREVIAS A LA CARGA Y PROGRAMACIÓN AUTOMATIZADA PARA MANTENER LOS DATOS ACTUALIZADOS Y LIMPIOS. ESTO NOS PERMITE REALIZAR CONSULTAS Y ANÁLISIS EFICIENTES PARA OBTENER INFORMACIÓN VALIOSA Y RESPALDAR LA TOMA DE DECISIONES.



**ESTAMOS EMOCIONADOS DE
COMENZAR LA PRÓXIMA SEMANA,
APLICAR LO QUE HEMOS APRENDIDO
Y LLEVAR NUESTRO NEGOCIO AL SIGUIENTE
NIVEL BASADO EN DATOS SÓLIDOS Y ANÁLISIS
INFORMADOS. ¡EL FUTURO SE VE PROMETEDOR!**



GRACIAS 

