

Documentacion

1. Claridad y Justificación del Proyecto

- **Título del Proyecto:** Integración de datos de campañas de marketing en Google BigQuery
- **Descripción:**

Este proyecto tiene como objetivo extraer datos de una API de marketing, procesarlos y cargarlos en Google BigQuery para su análisis. Las empresas de marketing manejan grandes volúmenes de datos provenientes de diferentes plataformas (Google Ads, Facebook Ads, etc.), y requieren una solución eficiente para centralizar y analizar estos datos. Este proyecto facilita la integración y análisis eficiente, automatizando la carga de datos y mejorando la toma de decisiones basada en análisis.
- **Relevancia del problema:**

La necesidad de tomar decisiones basadas en datos es crítica para optimizar campañas publicitarias. Sin una solución que permita integrar datos de múltiples fuentes, los equipos de marketing perderían tiempo en procesos manuales y análisis fragmentados, afectando la efectividad de sus decisiones. Este proyecto resuelve la necesidad de una plataforma única de análisis que centralice los datos para una visión completa.

2. Diseño de la Solución

- **Descripción técnica:**

Se desarrolló una solución en Python para automatizar la extracción de datos desde la API de Mockaroo (simulando diferentes plataformas de anuncios), que se guarda en archivos CSV y posteriormente se sube a Google BigQuery. Utilizamos `pandas_gbq` para interactuar con BigQuery y `pandas` para procesar los datos localmente antes de cargarlos.
- **Simplicidad y funcionalidad:**

La solución fue diseñada para ser lo más eficiente posible, manteniendo el proceso simple pero eficaz. Se usaron componentes existentes de Python y Google Cloud para evitar la complejidad innecesaria. La selección de herramientas como

`pandas_gbq` y Google BigQuery se basa en su eficiencia y facilidad para manejar grandes volúmenes de datos.

- **Justificación del enfoque:**

El uso de una API simulada y el posterior procesamiento en BigQuery permite la escalabilidad del proyecto. Este enfoque asegura que el proceso sea fácilmente replicable en un entorno real de marketing con datos masivos y múltiples fuentes.

3. Desarrollo Técnico

- **Tecnologías utilizadas:**

- **Python:** Para la extracción de datos de la API y su procesamiento.
- **pandas:** Para manipulación de datos.
- **pandas_gbq:** Para cargar datos a Google BigQuery.
- **Google BigQuery:** Para almacenamiento y análisis de datos.
- **Mockaroo API:** Para simular los datos de campañas de marketing.

- **Explicación técnica detallada:**

1. **Extracción de datos:** Se utiliza la librería `requests` para obtener los datos desde la API.
 2. **Transformación de datos:** Los datos se guardan en CSV y se formatean con `pandas` para asegurar que las fechas y las campañas estén correctas.
 3. **Carga en BigQuery:** Finalmente, se cargan los datos en una tabla predefinida en Google BigQuery utilizando `pandas_gbq`.
-