

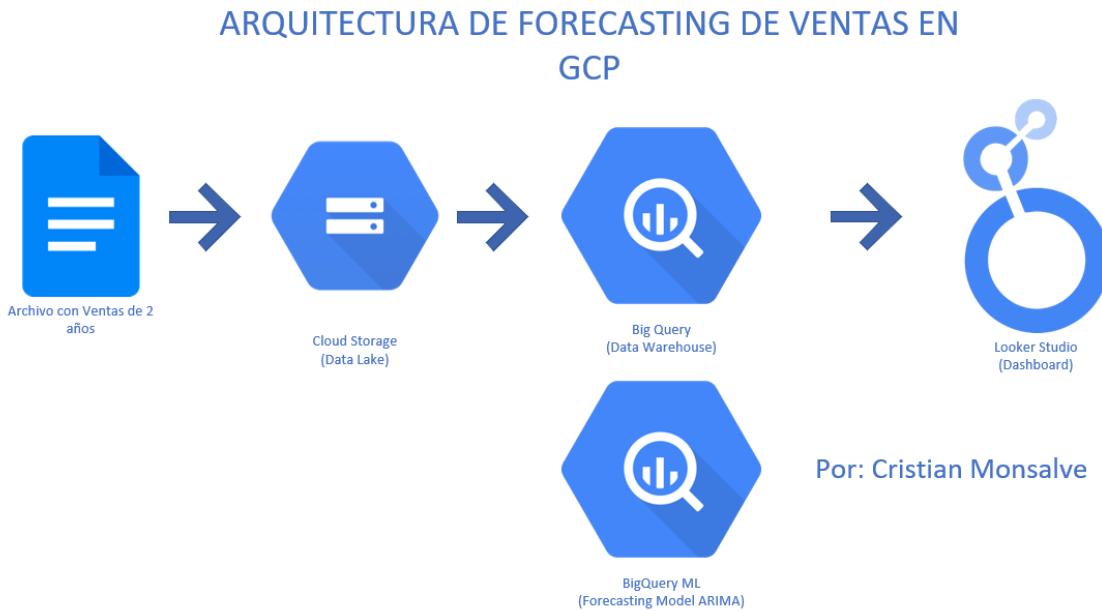
Sistema de Forecasting de Ventas Serverless con BigQuery ML

1. Definición del Problema

Se crea un caso ficticio de una empresa que necesita optimizar su inventario para el año 2026. Basarse en promedios simples provoca quiebres de stock o sobrecompra.

El objetivo es automatizar la predicción de demanda diaria utilizando datos históricos de 2 años para predecir el comportamiento de ventas futuro con un intervalo de confianza del 90%.

2. Arquitectura de la Solución



Servicios Utilizados:

- **Google Cloud Storage:** Data Lake para la ingestión de archivos CSV crudos.
- **BigQuery:** Data Warehouse para el almacenamiento y limpieza de datos (ETL).
- **BigQuery ML:** Motor de Inteligencia Artificial para crear el modelo ARIMA_PLUS (Series de Tiempo) mediante SQL.
- **Looker Studio:** Herramienta de Business Intelligence para la visualización de la proyección a 365 días.

3. Desarrollo Técnico

Paso 1:

Ingesta y ETL: Se cargaron los datos crudos y se creó un proceso de limpieza en SQL para estandarizar fechas y formatos. El archivo CSV por alguna razón solo mostraba una columna, por lo que se separaron las columnas para dejar el formato tabla.

```
CREATE OR REPLACE TABLE `ventas_retail.historial_ventas` AS  
  
SELECT  
  
    PARSE_DATE('%d-%m-%Y', SPLIT(string_field_0, ';')[SAFE_OFFSET(0)]) AS fecha,  
    SPLIT(string_field_0, ';')[SAFE_OFFSET(1)] AS producto,  
    CAST(SPLIT(string_field_0, ';')[SAFE_OFFSET(2)] AS INT64) AS cantidad  
  
FROM  
  
`ventas_retail.historial_ventas`  
  
WHERE  
  
-- Filtramos los encabezados  
string_field_0 NOT LIKE '%Fecha%'  
AND string_field_0 IS NOT NULL;
```

Paso 2:

Entrenamiento del Modelo (Machine Learning): Se entrenó un modelo ARIMA (AutoRegressive Integrated Moving Average) capaz de detectar estacionalidad y tendencias automáticamente.

Algoritmo: ARIMA_PLUS

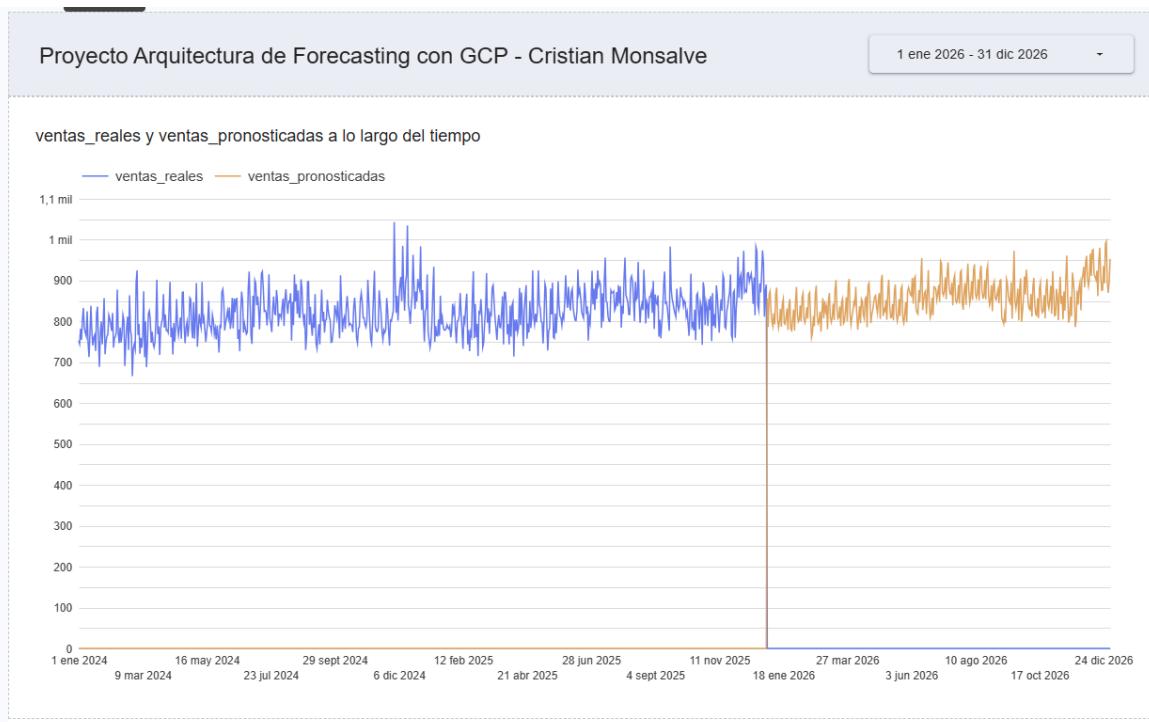
Métrica objetivo: total_ventas

```
CREATE OR REPLACE MODEL `ventas_retail.modelo_forecast_ventas`  
OPTIONS(  
    model_type='ARIMA_PLUS',  
    time_series_timestamp_col='fecha',  
    time_series_data_col='total_ventas'  
) AS  
SELECT  
    fecha,  
    SUM(cantidad) as total_ventas  
FROM  
    `ventas_retail.historial_ventas`  
GROUP BY  
    fecha;
```

Paso 3:

Proyección (Forecasting): Se generó una predicción para el horizonte de 365 días (Año 2026), unificando los datos históricos y futuros en una tabla maestra para el reporte.

4. Resultados



El gráfico muestra la tendencia histórica y cómo el modelo proyecta el comportamiento para el próximo año, permitiendo tomar decisiones anticipadas de compra de stock.

5. Conclusión

La solución implementada es 100% Serverless, lo que reduce los costos operativos y de mantenimiento. El uso de BigQuery ML permitió democratizar el acceso a la Inteligencia Artificial usando SQL estándar, acelerando el tiempo de desarrollo de semanas a horas.