

Entregable Intermedio — TS2

Tema: Analítica de Datos y Business Intelligence aplicado a datos farmacéuticos

Periodo: 1 sep – 28 nov 2025 — Entregable intermedio (10/oct/2025)

Autor: Luis Mario Ayala Castellanos

1. Resumen / Abstract

Este proyecto aplica técnicas de Analítica de Datos y Business Intelligence (BI) a un conjunto de ventas farmacéuticas con el objetivo de generar KPIs accionables y un pipeline reproducible para la toma de decisiones en inventarios y marketing. En este entregable se presenta el marco teórico, la definición del caso, objetivos SMART, dataset y diccionario de datos, propuesta de KPIs, y un avance de ETL con visualizaciones preliminares basadas en una muestra (50 registros) como prueba de concepto.

2. Marco teórico y estado del arte

Analítica descriptiva, diagnóstica y predictiva: La analítica descriptiva se enfoca en resumir qué ocurrió en el pasado, la diagnóstica busca explicar causas, y la predictiva intenta anticipar comportamientos futuros. La arquitectura BI integra datos mediante procesos ETL en un data warehouse, seguido por análisis y dashboards. Se incluyen KPIs como ventas netas, ticket promedio, rotación de inventario y ventas por categoría.

3. Definición del caso y objetivos (SMART)

Caso: Optimizar la gestión comercial de una farmacia mediante analítica de ventas históricas.

Problema: Falta de métricas estandarizadas que permitan priorizar surtido/inventario y detectar productos con baja rotación.

Objetivo general: Para el 28 de noviembre de 2025, desarrollar un paquete reproducible de BI que permita reducir desabasto y optimizar inventario mediante KPIs y dashboards en Python.

Objetivos específicos:

1. Definir y calcular KPIs operativos antes del 10/oct.
2. Implementar un pipeline ETL reproducible.
3. Construir 2–4 vistas en dashboards con narrativa ejecutiva.

4. KPIs propuestos

1. Ventas Totales = $\Sigma (\text{Cantidad} \times \text{PrecioUnitario})$.
2. Número de Transacciones = conteo de filas.
3. Ticket Promedio = Ventas Totales / Número de Transacciones.
4. Clientes Únicos = `COUNT_DISTINCT(ClienteID)`.
5. Ventas por Categoría = $\Sigma \text{TotalVenta por CategoríaProducto}$.
6. Top N Productos por ventas = orden descendente por TotalVenta.
7. Frecuencia de Compra promedio = intervalo promedio entre compras por cliente.
8. % Ventas por Ciudad = $(\text{Ventas Ciudad} / \text{Ventas Totales}) \times 100$.

5. Dataset y data dictionary

Fuente: Dataset propio de la farmacia (original 30,000 registros). Para este entregable se utilizó una muestra de 50 registros.

Columnas incluidas: ClienteID, NombreCliente, Edad, Ciudad, ProductoID, NombreProducto, CategoríaProducto, FechaCompra, Cantidad, PrecioUnitario, TotalVenta.

Ejemplo de data dictionary:

ClienteID (string) — ID cliente. Ej: C001.

NombreCliente (string) — Nombre completo (PII, se recomienda anonimizar).

Edad (int) — Edad en años.

Ciudad (string) — Ciudad de residencia.

ProductoID (string) — ID producto.

NombreProducto (string) — Descripción producto.

CategoríaProducto (string) — Categoría.

FechaCompra (datetime) — Fecha de compra.

Cantidad (int) — Unidades vendidas.

PrecioUnitario (float) — Precio por unidad.

TotalVenta (float) — $\text{Cantidad} \times \text{PrecioUnitario}$.

6. Avance ETL / limpieza

Operaciones realizadas:

- Carga del archivo Excel con pandas.
- Limpieza de nombres de columnas.
- Conversión de fechas a datetime.
- Conversión de variables numéricas (Cantidad, PrecioUnitario).
- Creación del campo TotalVenta.

- Chequeo de nulos y exportación de la muestra procesada.

7. Visualizaciones preliminares

Se generaron gráficos de ventas por ciudad, ventas por categoría y ventas por fecha. En la muestra se observó concentración de ventas en algunas ciudades y categorías.

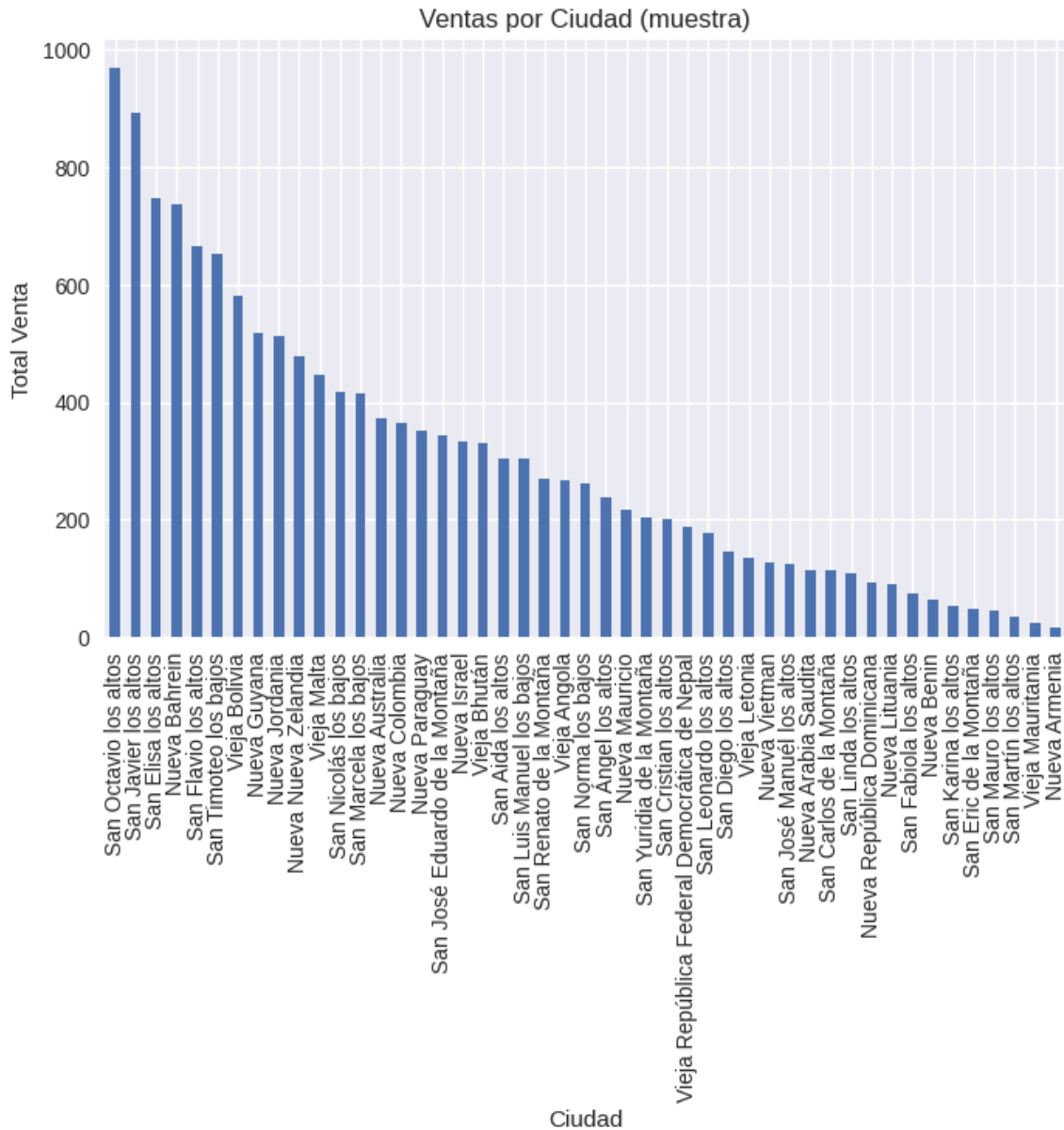


Fig 1. Ventas por Ciudad

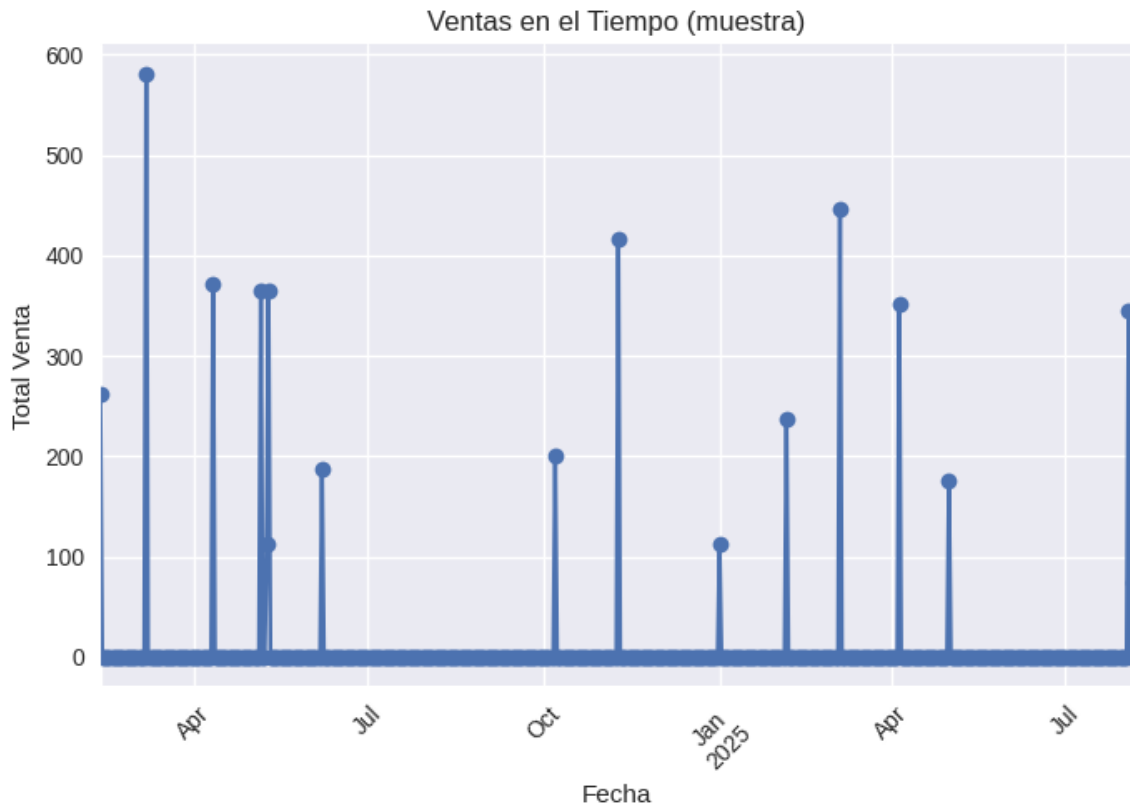


Fig 2. Ventas por fechas

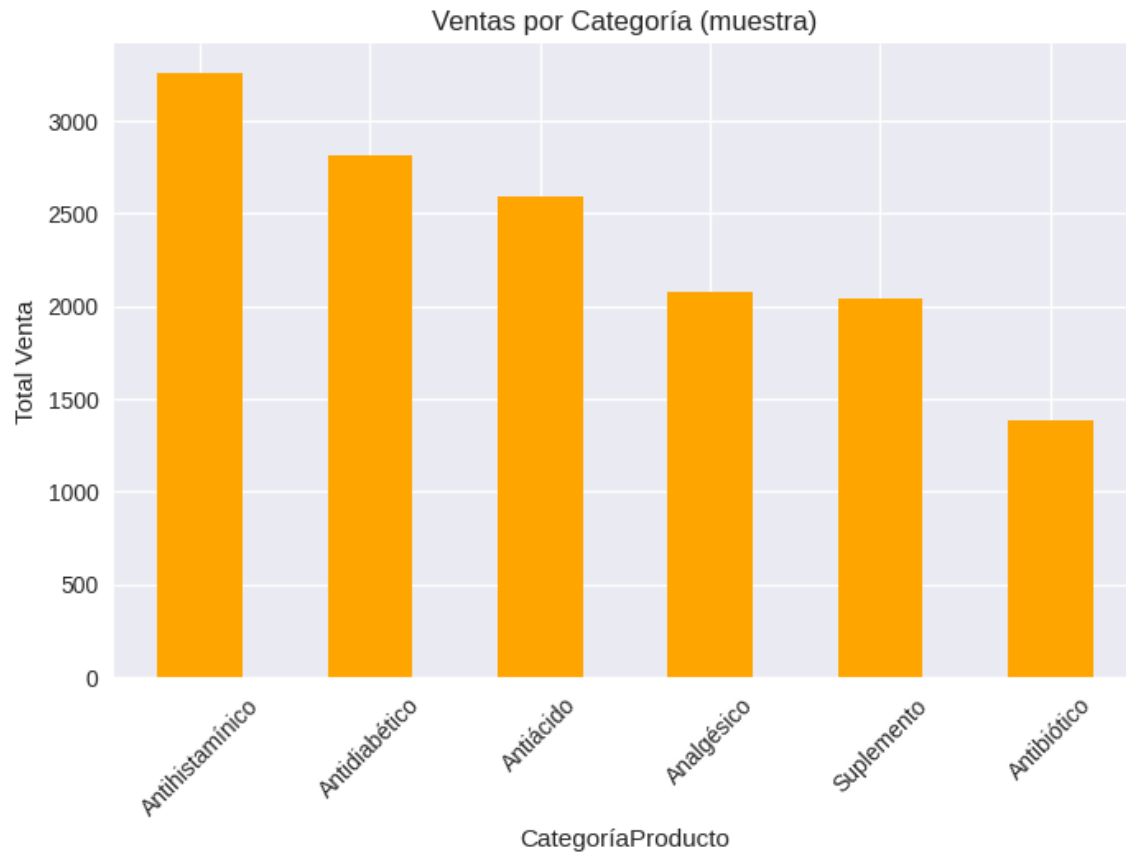


Fig 3. Ventas por Categoría

Referencias

- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice (2nd ed.). OTexts.
- Few, S. (2012). Show Me the Numbers: Designing Tables and Graphs to Enlighten. Analytics Press.
- Provost, F., & Fawcett, T. (2013). Data Science for Business. O'Reilly.