1. Project Overview

- **Title:** *Warehouse and Retail Sales Analytics*
- **Description:** Hands-on analytics project using Python to compare retail vs. warehouse sales, analyze item types, supplier efficiency, and monthly trends. Designed to demonstrate applied data storytelling and business impact.

2. Data Preparation

- **Dataset Size:** 307,645 rows
- **Cleaning:**
  - Missing suppliers (<1%) → filled with "Unknown"
  - Retail sales missing (<5 rows) → filled with 0
  - Item type missing (1 row) → filled with "Unknown"
- **Impact:** Minimal missing data, no effect on analysis validity.

3. Key Metrics Created

- **Net_Sales_Quantity** = Retail Sales + Warehouse Sales – Retail Transfers
- **Sales_Per_Supplier** = Retail Sales ÷ Number of Suppliers
- **Quarter Column** = Q1–Q3 (based on month)

4. Analysis & Insights

- **Item Type with Highest Avg Net Sales:** Beer dominates overall sales.
- **Supplier Efficiency:** ARCHER ROOSE LLC shows highest consistency.
- **Quarterly Retail Sales:** Q2 shows stronger averages compared to Q1/Q3.
- **Fastest Growing Month:** July shows peak retail sales growth.
- **Statistical Tests:**
  - *t-test:* Retail vs. Warehouse Sales differ significantly.
  - *ANOVA:* Item Type drives differences in sales.
  - *Chi-Square:* Supplier allocation is not independent of item type.

5. Visualizations

- 📈 **Monthly Retail Sales Trend** → Line chart (peak in July)
- 🏷️ **Retail Quantities by Quarter** → Bar chart
- 📊 **Retail, Warehouse, Transfers by Item Type** → Stacked bar chart
- 🎨 **Heatmap of Item Type vs. Month** → Seasonal sales patterns
- 🗂️ **Overall Sales Representation** → Combined chart for storytelling