📄 Executive Summary – TerpBuy Data Analytics (Part II)

🗒 Introduction

This executive summary presents key findings from the analysis performed in Part II of the TerpBuy Data Science project. The objective was to extract actionable insights using transactional and reference data sourced from TerpBuy's order, product, and customer systems.

Due to technical limitations in establishing a live connection to MySQL, the analysis was conducted using local CSV files that replicate the structure and content of the original database. Python was utilized for data processing and visualization using Pandas, Matplotlib, Seaborn, and Plotly.

🛠 Approach

The following CSV datasets were analyzed: `orders`, `order_line`, `product`, `department`, `category`, and `customer`. These files were joined and processed to simulate SQL queries and provide visual business insights. The main focus areas included product movement by department and annual trends in order volume.

📦 Insight 1: Department-wise Quantity Sold

Items sold were grouped by department to identify high-performing areas. This revealed that some departments significantly outperform others in terms of sales volume (e.g., 'Fitness' and 'Technology'). These areas may be strategically important for the business.

Low-performing departments may require attention in terms of inventory optimization, marketing focus, or product strategy revision.

📈 Insight 2: Orders Per Year (Simulated)

Since actual order dates were not available, realistic order dates were simulated from 2018 to 2023. This allowed for the creation of an interactive line graph to observe potential trends in order volume over time.

Though based on simulated data, the chart demonstrated how time-based analysis can be valuable for identifying peaks, trends, and seasonal behaviors. With actual dates, the same methodology can be applied for operational forecasting.

🪃 Strategic Recommendations

Based on the analysis, the following recommendations are proposed:

- • Continue investing in high-performing departments to sustain growth.
- • Use time-based sales data for strategic planning once real order dates are available.
- • Reevaluate underperforming departments to optimize SKU performance and marketing efforts.

☑ Conclusion

This project demonstrates the power of local data analysis using Python tools as an alternative to SQL environments. It delivers quick, replicable, and insightful outputs that are essential for agile business decision-making.

Prepared by: [Sahil Verma
Date: July 21, 2025
Project: TerpBuy – Data Science Toolkit Assignment