REPORT: SUPERSTORE SALES ANALYSIS USING PYTHON & POWER BI

1. Objective

The objective of this task was to clean and prepare the Superstore dataset using Python, and then build an interactive Power BI dashboard to analyze sales performance across time, region, and product categories.

2. Data Cleaning (Python – Pandas)

The raw dataset (Sample - Superstore.csv) contained 9,994 rows and 21 columns. Data cleaning steps included:

1. **Handling Missing Values**
    - Replaced missing postal codes with 0.
    - Dropped rows with missing critical IDs (e.g., Order ID, Customer ID).
2. **Removing Duplicates**
    - Used drop_duplicates() to ensure each transaction is unique.
3. **Data Type Conversions**
    - Converted Order Date and Ship Date to datetime format.
4. **Standardizing Text Fields**
    - Removed leading/trailing spaces from categorical columns.
5. **Dropping Irrelevant Columns**
    - Dropped Row ID, which acted as just an index.
6. **Exporting Clean Dataset**
    - Saved the cleaned dataset as Superstore_Cleaned.csv for Power BI import.

3. Data Import (Power BI)

- Imported Superstore_Cleaned.csv into Power BI Desktop.
- Created a new calculated column for **Month-Year**:
- MonthYear = FORMAT([Order Date], "MMM-YYYY")

4. Dashboard Development

Three visuals and one slicer were created:

1. **Line Chart – Sales over Time**
    - Axis: Month-Year
    - Values: Sales
    - Shows sales trend over months with seasonal peaks.
2. **Bar Chart – Sales by Region**
    - Axis: Region
    - Values: Sales
    - Conditional formatting applied:
        - Green = Highest region
        - Orange = Medium
        - Red = Lowest region
3. **Donut Chart – Sales by Category**
    - Legend: Category
    - Values: Sales
    - Shows contribution of Technology, Furniture, and Office Supplies.
4. **Slicer**
    - Filter by Region (or Category) for interactive analysis.

5. Conditional Formatting

- In the **Bar Chart (Sales by Region)**, applied a color scale:
    - Top-performing region highlighted in **green**.
    - Lowest performing region highlighted in **red**.
- This makes it easier to visually identify strong and weak areas.

6. Key Insights

From the analysis, the following insights were identified:

1. **Regional Performance**
    - The West region consistently outperformed other regions throughout the year, especially in Q3.
    - The East region also showed strong performance in Q4, achieving sales of approximately $0.30M.
2. **Category Contribution**
    - The Technology category contributed the highest share of total sales (36.4%), making it the key revenue driver.
    - The Furniture category was most popular in the West region, suggesting regional product preferences.
3. **Monthly Trends**
    - March (Q1), May (Q2), September (Q3), and November (Q4) emerged as the top-performing months in their respective quarters.
    - September recorded the highest peak sales of about $0.31M, making it the strongest month across the year.
    - November–December showed a seasonal sales spike, indicating strong holiday demand.

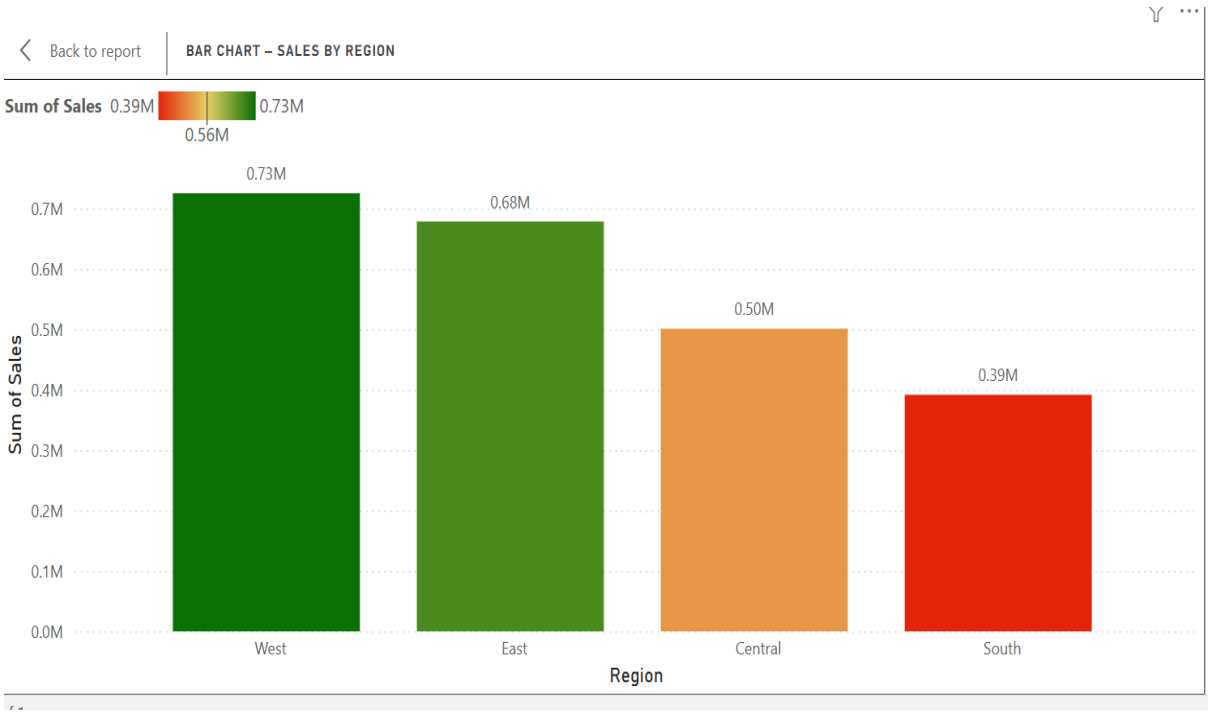
7. Conclusion

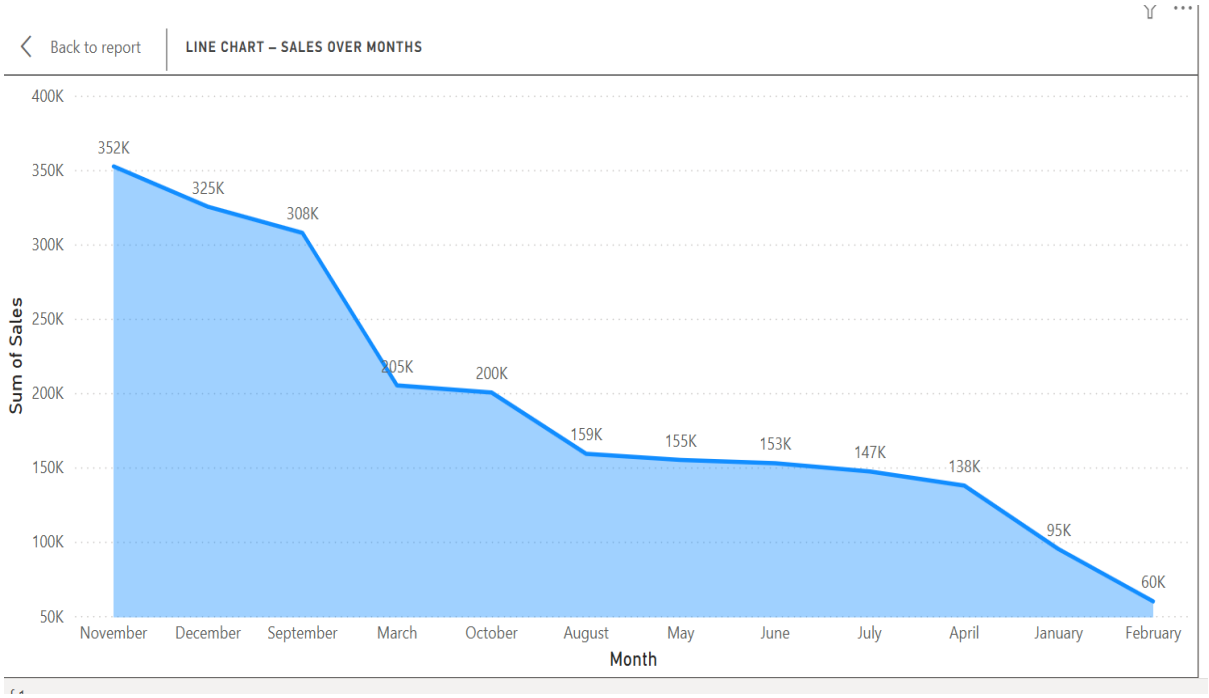
By combining Python and Power BI:

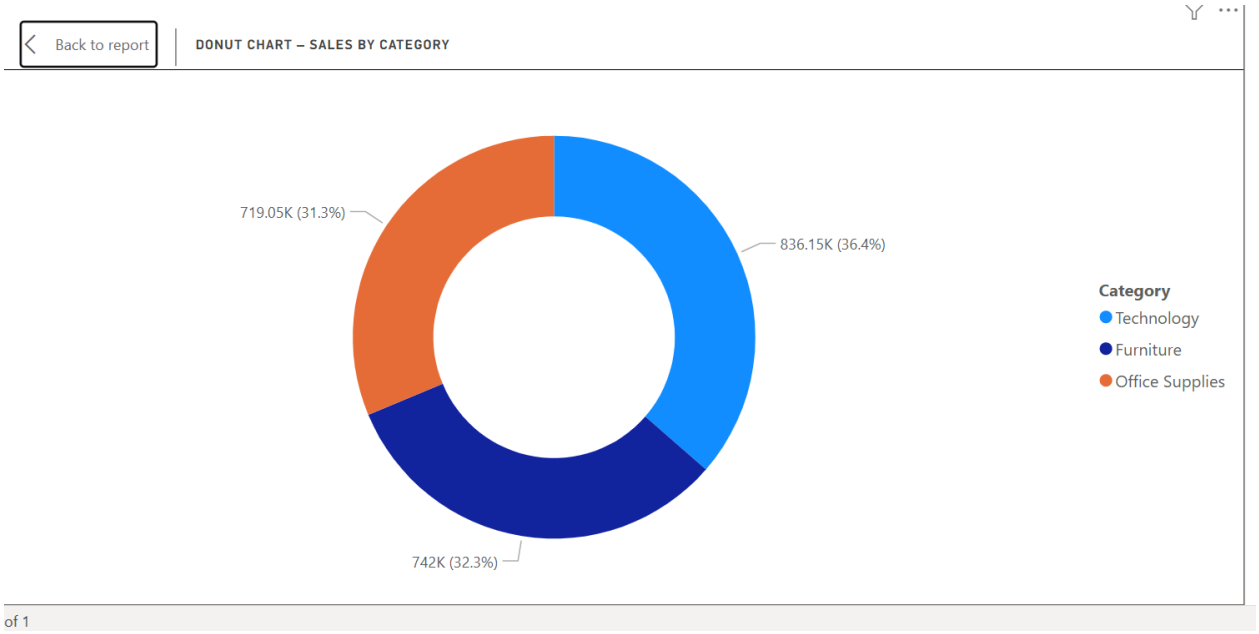
- **Python** handled the heavy lifting of data cleaning and preparation.
- **Power BI** provided interactive dashboards, filters, and conditional formatting for business insights.

This workflow ensures **accurate, clean data** and **interactive visualization**, enabling better decision-making for sales strategy.

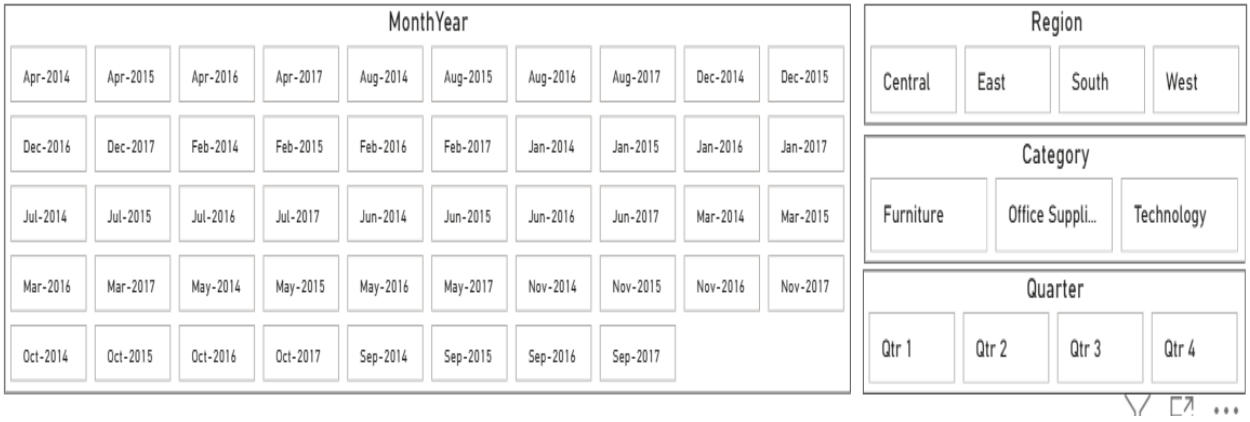
Charts:


Back to report
BAR CHART – SALES BY REGION
Sum of Sales 0.39M 0.56M 0.73M
0.73M
0.68M
0.50M
0.39M
0.7M
0.6M
0.5M
0.4M
0.3M
0.2M
0.1M
0.0M
Sum of Sales
West
East
Central
South
Region

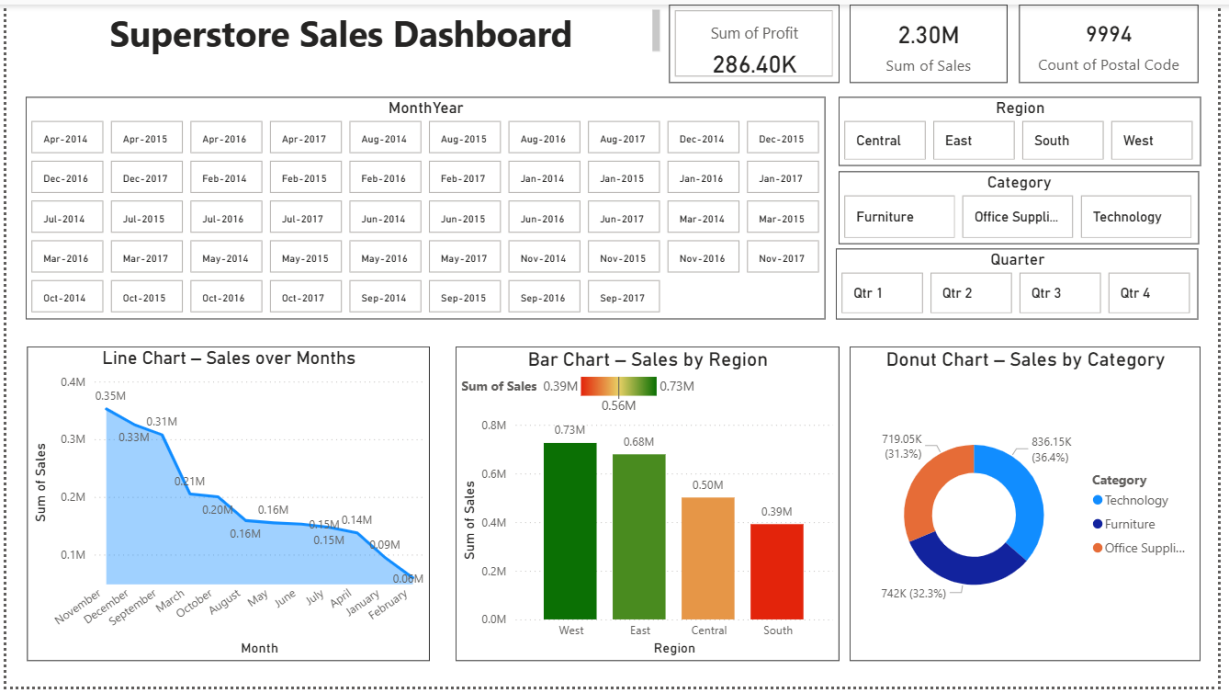

Back to report
LINE CHART – SALES OVER MONTHS
400K
350K
300K
250K
200K
150K
100K
50K
Sum of Sales
352K
325K
308K
205K
200K
159K
155K
153K
147K
138K
95K
60K
November
December
September
March
October
August
May
June
July
April
January
February
Month


Back to report
DONUT CHART – SALES BY CATEGORY
719.05K (31.3%)
836.15K (36.4%)
742K (32.3%)
Category
Technology
Furniture
Office Supplies
of 1

Slicer:


MonthYear
Apr-2014 Apr-2015 Apr-2016 Apr-2017 Aug-2014 Aug-2015 Aug-2016 Aug-2017 Dec-2014 Dec-2015
Dec-2016 Dec-2017 Feb-2014 Feb-2015 Feb-2016 Feb-2017 Jan-2014 Jan-2015 Jan-2016 Jan-2017
Jul-2014 Jul-2015 Jul-2016 Jul-2017 Jun-2014 Jun-2015 Jun-2016 Jun-2017 Mar-2014 Mar-2015
Mar-2016 Mar-2017 May-2014 May-2015 May-2016 May-2017 Nov-2014 Nov-2015 Nov-2016 Nov-2017
Oct-2014 Oct-2015 Oct-2016 Oct-2017 Sep-2014 Sep-2015 Sep-2016 Sep-2017
Region
Central East South West
Category
Furniture Office Suppli... Technology
Quarter
Qtr 1 Qtr 2 Qtr 3 Qtr 4

Dashboard


Superstore Sales Dashboard
Sum of Profit
286.40K
2.30M
Sum of Sales
9994
Count of Postal Code
MonthYear
Apr-2014 Apr-2015 Apr-2016 Apr-2017 Aug-2014 Aug-2015 Aug-2016 Aug-2017 Dec-2014 Dec-2015
Dec-2016 Dec-2017 Feb-2014 Feb-2015 Feb-2016 Feb-2017 Jan-2014 Jan-2015 Jan-2016 Jan-2017
Jul-2014 Jul-2015 Jul-2016 Jul-2017 Jun-2014 Jun-2015 Jun-2016 Jun-2017 Mar-2014 Mar-2015
Mar-2016 Mar-2017 May-2014 May-2015 May-2016 May-2017 Nov-2014 Nov-2015 Nov-2016 Nov-2017
Oct-2014 Oct-2015 Oct-2016 Oct-2017 Sep-2014 Sep-2015 Sep-2016 Sep-2017
Region
Central East South West
Category
Furniture Office Suppli... Technology
Quarter
Qtr 1 Qtr 2 Qtr 3 Qtr 4
Line Chart – Sales over Months
Sum of Sales
0.4M 0.3M 0.2M 0.1M
0.35M 0.33M 0.31M 0.21M 0.20M 0.16M 0.16M 0.15M 0.15M 0.14M 0.09M 0.05M
November December September March October August May June July April January February
Month
Bar Chart – Sales by Region
Sum of Sales 0.39M 0.56M 0.73M
0.8M 0.6M 0.4M 0.2M 0.0M
0.73M 0.68M 0.50M 0.39M
West East Central South
Region
Donut Chart – Sales by Category
719.05K (31.3%)
836.15K (36.4%)
742K (32.3%)
Category
Technology
Furniture
Office Suppli...