**Walmart Sales Analysis Report**

**API Selection and Data Fetching Steps**

The dataset used for this analysis is fetched from a GitHub repository. The following code snippet demonstrates how to fetch the data using pandas:

python

Copy code

import pandas as pd

# Original URL

url = "https://github.com/LokeshKumarChauhan/DE\_with\_powerBI/blob/main/Walmart.csv"

# Modify the URL to get the raw file

raw\_url = url.replace("github.com", "raw.githubusercontent.com").replace("/blob/", "/")

# Read the CSV file from the modified URL

df = pd.read\_csv(raw\_url)

**Data Processing and Cleaning Steps**

The dataset contains columns such as 'Order Date' and 'Ship Date' with dates in different formats. The following steps were taken to clean and process the data:

python

Copy code

df['Order Date'] = df['Order Date'].str.replace('-', '/')

df['Ship Date'] = df['Ship Date'].str.replace('-', '/')

# Convert the columns to datetime

df['Order Date'] = pd.to\_datetime(df['Order Date'], format='%d/%m/%Y', errors='coerce')

df['Ship Date'] = pd.to\_datetime(df['Ship Date'], format='%d/%m/%Y', errors='coerce')

print(df)

In the above code, the 'Order Date' and 'Ship Date' columns are initially in the format 'dd-mm-yyyy'. They are first modified to 'dd/mm/yyyy' format and then converted to datetime objects. This ensures that the dates are in a consistent and usable format for analysis.

**Explanation of the Visualizations Created and Insights Derived**

The dashboard contains several visualizations, each providing different insights into the sales data. Below is an explanation of each visualization:

1. **Total Customers**: This card shows the total number of customers. It is calculated as the count of unique Row IDs.
2. **Segment-wise Customer Count**: These cards show the count of customers in each segment: Corporate, Consumer, and Home Office. It provides a quick view of customer distribution across different segments.
3. **Quantity by Category and Ship Mode**: This pie chart shows the quantity of products ordered, broken down by product category and shipping mode. This helps in understanding which categories are more popular and which shipping modes are preferred.
4. **Count by Product Name**: This gauge chart shows the count of orders for different product names. It helps in identifying the most and least popular products.
5. **Sum of Quantity by Category and Sub-Category**: This bar chart shows the sum of quantities ordered for different sub-categories within each category. It helps in identifying which sub-categories are contributing most to the sales.
6. **Sum of Sales and Profit by Order Date**: This line chart shows the trend of sales and profit over time. It helps in understanding how sales and profit have evolved over the years.
7. **State-wise Profit**: This map visualizes the profit across different states. It helps in identifying the geographical regions with the highest and lowest profits.
8. **Order Details**: This table contains detailed information on each order, including OrderID, Customer Name, and Product Name. It provides a granular view of the data, allowing for specific insights into individual orders and customer preferences. This detailed table can be used to:

* Track individual order performance.
* Identify frequent customers and their preferred products.
* Analyze product popularity at a granular level.
* Investigate specific orders for further insights or troubleshooting.