

Project: Super Store Sales Analysis

Data Overview

Data consists of 9800 rows and 17 columns

Project Plane and Roles

Data Exploration and Cleaning by Python: Mohamed Adel

EDA questions: Khalid Habib

Data Analysis by SQL : Bassant Maged

Data Visualization by Tableau: Mai Mamdooh

Insights and KPI: Asmaa Mahmoud

Risks

By exploring data using python we found that the data almost in a clean form except:

- 1- Postal code column has 11 null values.
- 2- Dates columns (order and ship) were in wrong data format .
- 3- There were only 1 duplicated row.

Solutions

- 1- All postal code null values found to be to the same state which is “Vermont” by a quick google search found it “05401” and we can fill it manually:
- 2- Change the data type to date for Order and Ship columns using python code:
- 3- Drop duplicated row

KPIs

By a quick look at the data we can detect the top insight which can be useful to the business like:

- 1- Total Customers
- 2- Total Orders
- 3- Returns Orders
- 4- Total Sales
- 5- Total Profit
- 6- Discount Value
- 7- Returned Orders Cost
- 8- Average Delivery Duration

We can find more by continue the analysis

