EXPLORATORY DATA ANALYSIS

- For a SUPERSTORE

Assigned as TASK 5 of the Internship with SPARK Foundation

Introduction to Dataset

- The dataset represents the Sales of its THREE CATEGORY of Goods across UNITED States – of a SUPERSTORE.

- The goods are primarily sold to THREE Segments of Customers .

- The Profitability is accounted against each sales made.

We will be attempting to understand the distribution of Profitability across Category of goods as well as the Sales distribution across the cities in USA.

DATA Report -

The total Transactions in the Superstore can be classified, Category wise as –

- Furniture ➔ 2121 Transactions
- Office Supplies ➔ 6026 Transactions
- Technology Goods ➔ 1847 Transactions.

SUB-Category wise –

- Furniture has 04 Sub-categories of goods – Bookcases; Chairs; Furnishings & Tables.
- OFFICE SUPPLIES has 09 Sub-Categories of goods – Appliances; Art; Binders; Envelops; Fasteners; Labels; Paper; Storage & Supplies.
- TECHNOLOGY has 04 Sub-Category of goods – Accessories; Copiers; Machines & Phones.

SALES –

- Sales has taken place across 531 Cities spread out in 49 States of the UNITED States.
- Sales measures indicate MAXIMUM Sales from the states of Washington; New York; California & Pennsylvania.
- REGION Wise sales – All the FOUR Regions have been covered proportionately – SOUTH Region having the least transactions & WEST region having the Maximum transactions.

Univariate Analysis -

- REGION Wise Sales & Profit Analysis-

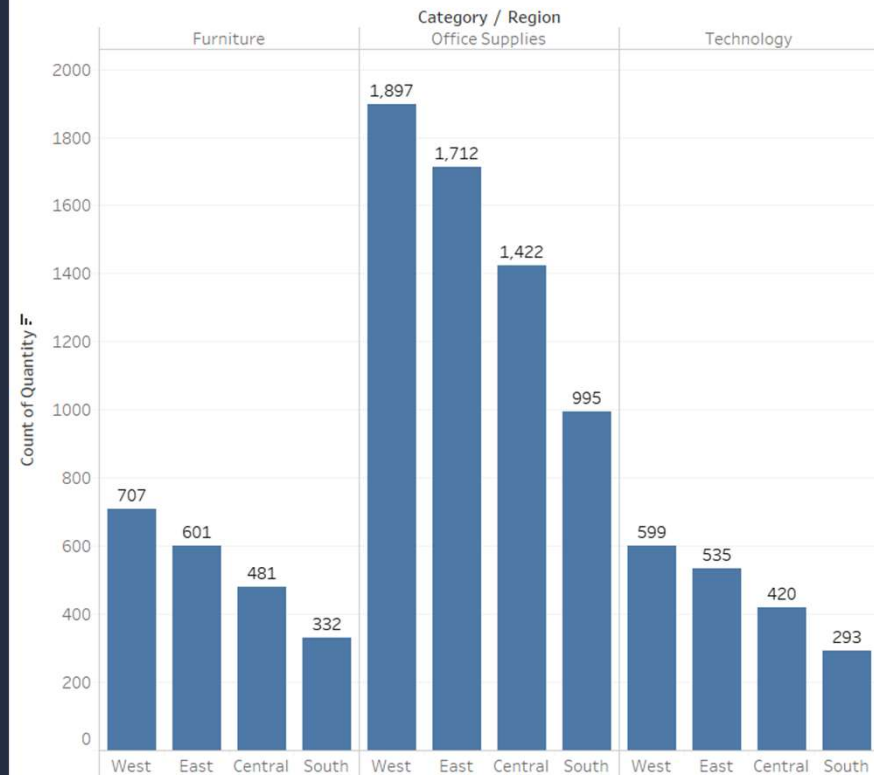| Category of Goods	MAXIMUM SALES	MINIMUM SALES	MAXIMUM PROFIT	MINIMUM PROFIT	LOSS MAKING
FURNITURE	WEST Region	SOUTH Region	WEST Region	EAST Region	CENTRAL Region
OFFICE Supplies	WEST Region	SOUTH Region	WEST Region	CENTRAL Region	NIL
TECHNOLOGY	EAST Region	SOUTH Region	EAST Region	SOUTH Region	NIL
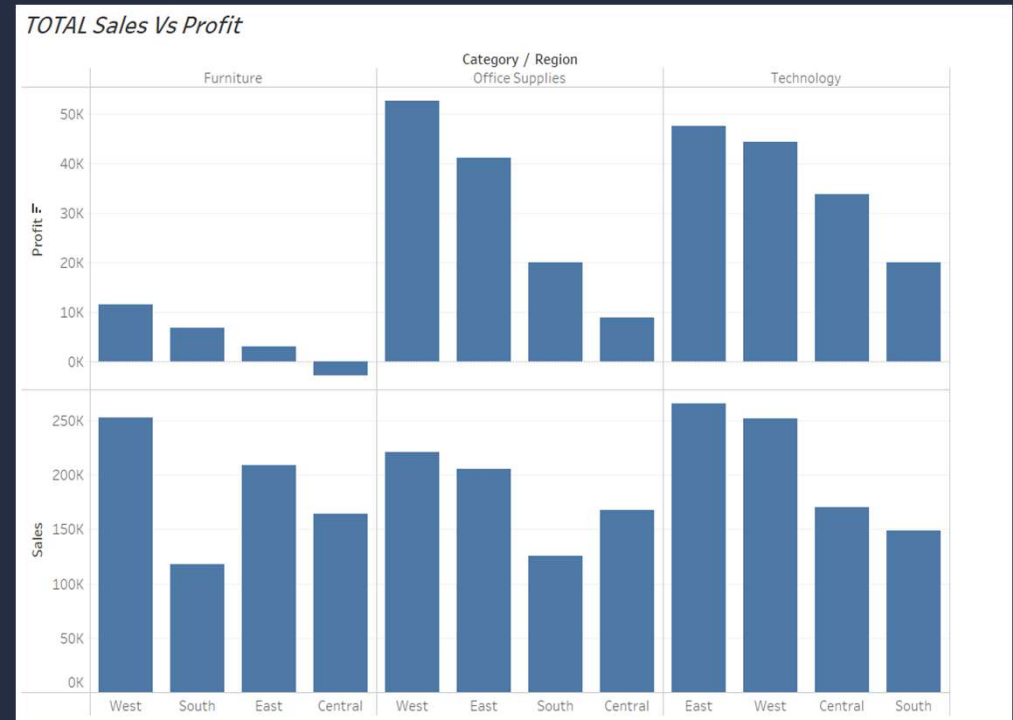
Click for Live Visualisation

Univariate Analysis –

- Category Wise Transaction Volume analysis –

| Category of Goods	NUMBER OF TRANSACTION			
	WEST Region	EAST Region	SOUTH Region	CENTRAL Region
FURNITURE	707	601	332	481
OFFICE Supplies	1897	1712	995	1422
TECHNOLOGY	599	535	293	420

Number of Transaction - Category & Region wise

Click for Live Visualisation

Bivariate Analysis -

Customer Segment wise Sales ; Profit & Discount analysis–

- In the **FURNITURE** category of goods , while maximum SALES is generated by the Consumer segment of customers, this segment gives the Least profitability. The maximum discount is also given to the Consumer segment.

- The 'Home Office' segment of customers gives sizable profitability as compared to the 'Corporate' segment , even after having lesser discounts availed.
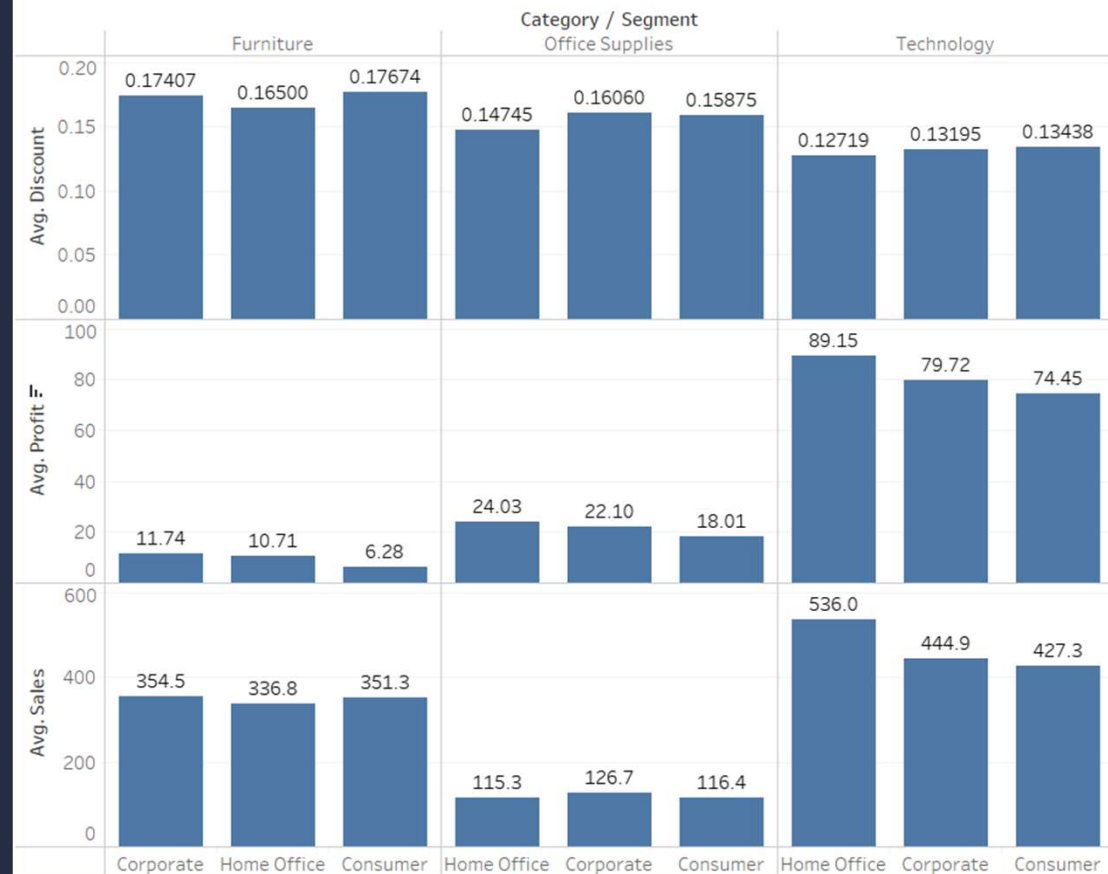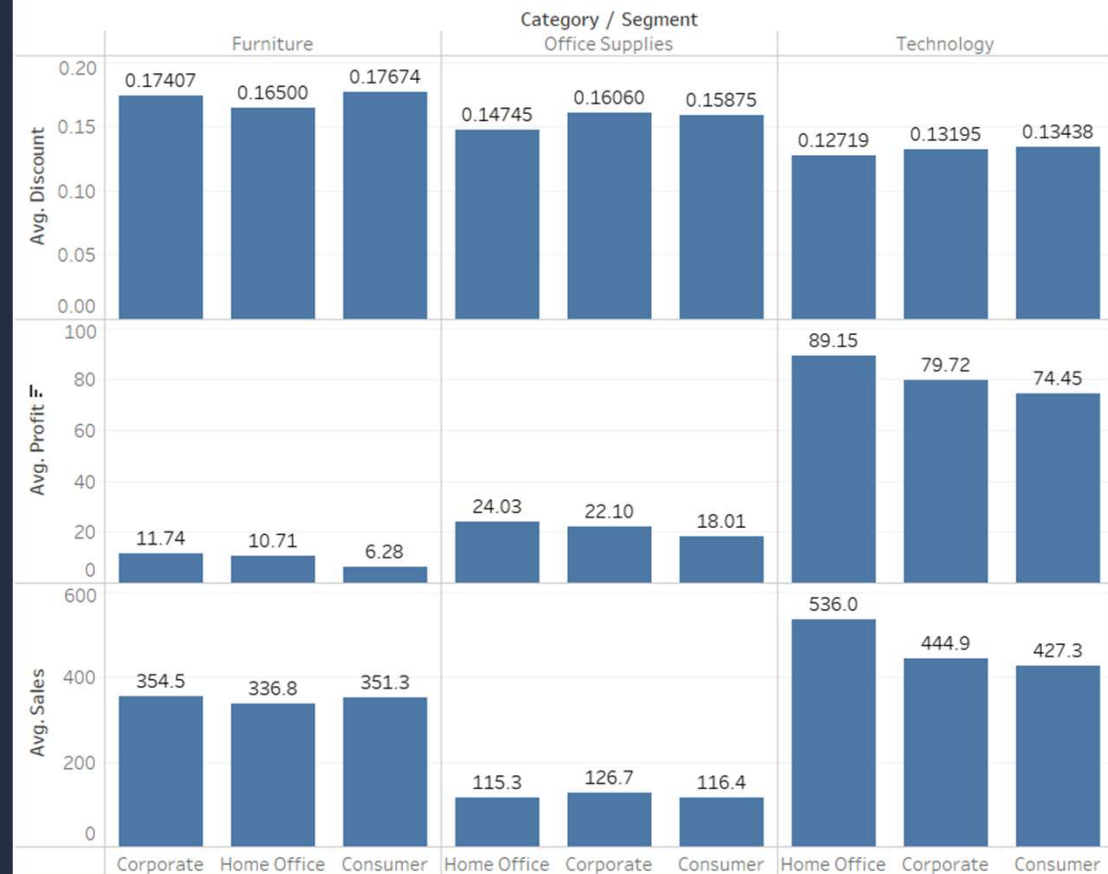

Customer Segment wise Sales ; Profit & Discount

Bivariate Analysis -

Customer Segment wise Sales ; Profit & Discount analysis–

- In the **TECHNOLOGY** category of goods , the 'Home Office' segment has shown the maximum sales with maximum profitability while the Discount released were the least.

- The 'Consumer' segment has the minimum Sales, and the Least profitability.

Customer Segment wise Sales ; Profit & Discount

Click for Live Visualisation

Bivariate Analysis -

Customer Segment wise Sales ; Profit & Discount analysis–

- In the **Office Supplies** category of goods , the 'Consumer' segment has shown the maximum sales but with Least profitability while the Discount released were comparatively high.

- The 'Home Office' segment has Less Sales but given maximum profitability with Least discounts shelled out

Customer Segment wise Sales ; Profit & Discount
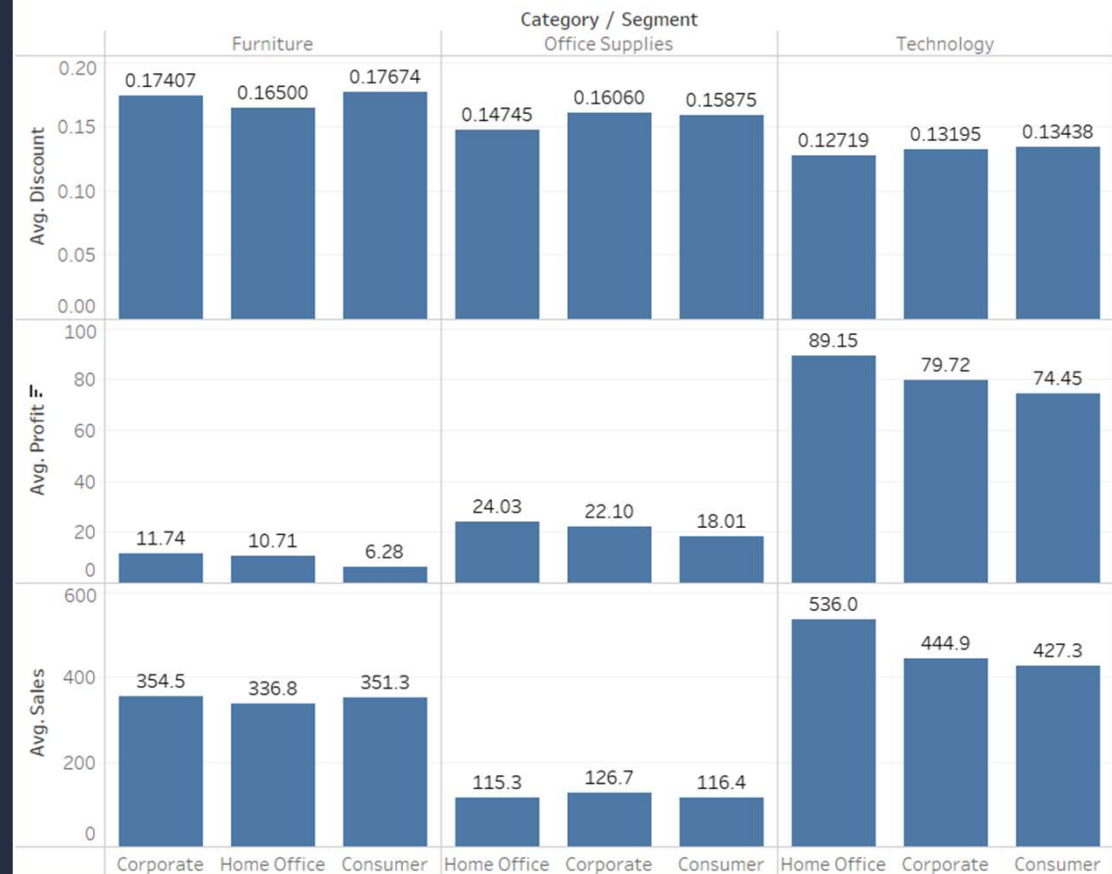
Insights from EDA-

Region -

- The Central Region of the STORES is doing Bad. In the Furniture category of goods, it is even eating away the profits of the company.
- For the CENTRAL Region – an exclusive Sales Strategy must be developed to INCREASE SALES and thereby increase profitability – for all category of goods.
- The WEST Region sales and profitability performance can be taken as benchmark and strategies developed for the rest of the regions – for Furniture & Office Supplies categories.
- The EAST Region Sales & Profitability for the Technology category can be adapted with regional changes – for the rest of the three regions.

Category of Customers	FURNITURE	OFFICE SUPPLIES	TECHNOLOGY
HOME OFFICE	Least Sales; Less Profitable	Less Sales; Most Profitable	Most Sales; Most Profitable
CORPORATE	Most Sales; Most Profitable	Most Sales; Less Profitable	Less Sales; Less Profitable
CONSUMER	Less Sales; Least Profitable	Less Sales; Least Profitable	Least Sales; Least Profitable

REGION	FURNITURE	OFFICE SUPPLIES	TECHNOLOGY
WEST REGION	Most Sales; Most Profitable	Most Sales; Most Profitable	Less Sales; Less Profitability
SOUTH REGION	Least Sales; Less Profitable	Least Sales; Low Profitability	Least Sales ; Least Profitability
CENTRAL REGION	Low Sales; LOSS MAKING	Low Sales; Least Profitable	Low Sales; Low Profitability
EAST REGION	Less Sales; Least Profitable	Low Sales; Low Profitability	Most Sales; Most Profitable

Insights from EDA-

Category of Customers -
- Even with variable discounts offered across categories of goods, NO CUSTOMER Segment is bringing Loss to the STORES.
- SALES to the Consumer segment is not benefitting the STORES as such. This means that the Consumer segment is not attracted to the STORES for buying ANY-OF-THE there category of goods.
- High Pricing / Negative Sales strategy can be reasons for this LOW Sales.
- An overhauling of SALES strategy with respect to the competitors in the market, needs to be adapted by the STORES – to increase sales and profitability in the CONSUMER Segment.
- Fine-Tuning is the sales strategy is required with the 'HOME OFFICE' segment of customers to optimize sales.
- When it comes to the CORPORATE Segment, a Sales strategy with Association rules can be thought of for increasing the sales of Technology products.

Category of Customers	FURNITURE	OFFICE SUPPLIES	TECHNOLOGY
HOME OFFICE	Least Sales; Less Profitable	Less Sales; Most Profitable	Most Sales; Most Profitable
CORPORATE	Most Sales; Most Profitable	Most Sales; Less Profitable	Less Sales; Less Profitable
CONSUMER	Less Sales; Least Profitable	Less Sales; Least Profitable	Least Sales; Least Profitable

REGION	FURNITURE	OFFICE SUPPLIES	TECHNOLOGY
WEST REGION	Most Sales; Most Profitable	Most Sales; Most Profitable	Less Sales; Less Profitability
SOUTH REGION	Least Sales; Less Profitable	Least Sales; Low Profitability	Least Sales ; Least Profitability
CENTRAL REGION	Low Sales; LOSS MAKING	Low Sales; Least Profitable	Low Sales; Low Profitability
EAST REGION	Less Sales; Least Profitable	Low Sales; Low Profitability	Most Sales; Most Profitable