



Budget Sales Analysis

By – Dhanshri Manusmare

PROJECT DETAIL

Project Title	Budget Sales Analysis
Technology	Business Intelligence
Domain	Sales and Retail
Programming Language Used	Python
Tools Used	Jupyter Notebook, MS-Excel, MS-Power BI

Objectives:

This project is to analyse the Customer data, Product data, Sales data and Budget Data of a Retail chain and extract key insights that can be valuable in taking business decisions.

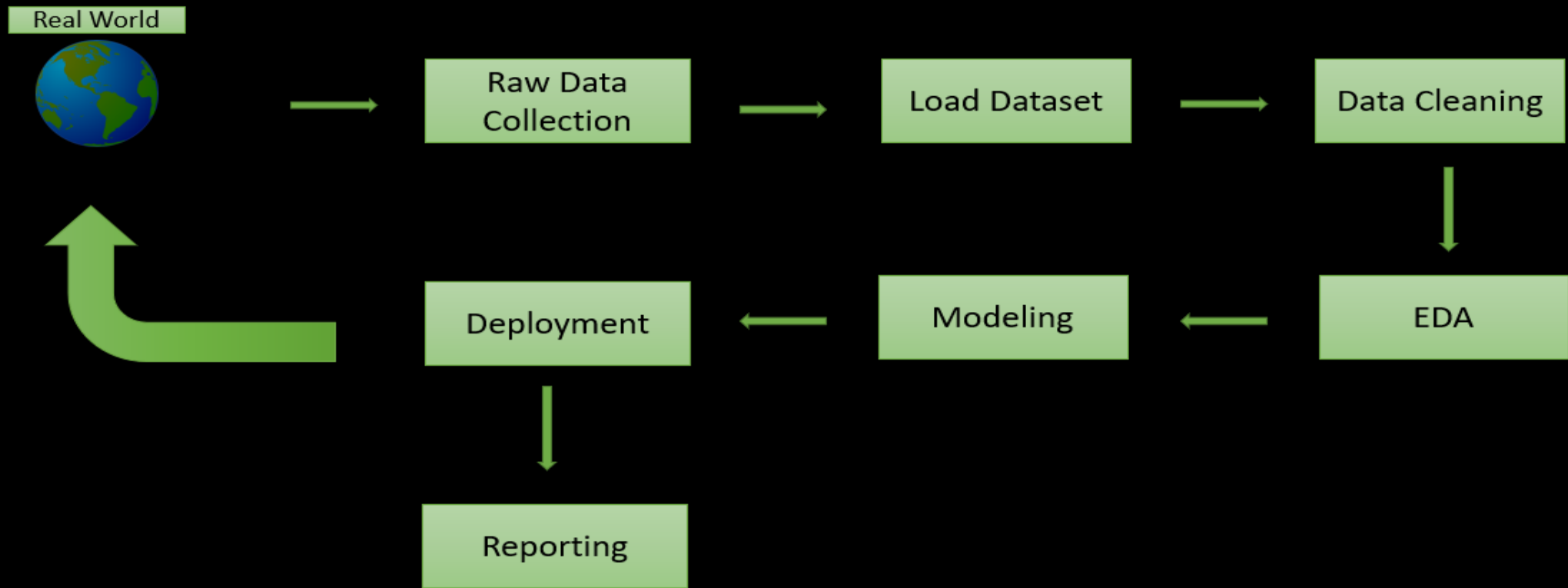
❑ Benefits:

- ▶ Help in making wiser business decisions.
- ▶ Aid in customer satisfaction and trend monitoring, which can serve current consumers and attract new ones.
- ▶ Greater client base understanding is provided.

Problem Statement:

Budget and Sales are by far most important attributes that defines a business's success and failure. Therefore, it is very important to keep a track on various features related to these attributes to keep on increasing the Sales and to allocate the Budget so that it can be utilized wisely and efficiently. So, it is very important for businesses to dig deep into the customer, sales, budget and product data to make better marketing strategy, to know the target customers, to make market friendly product upgrades and to keep a strong track on the budget efficiency. Good data driven systems can help achieve these goals and take the businesses forward towards success.

Architecture:



- Low Level Design Document
- High Level Design Document
- Architecture Document
- Wireframe Document
- Detailed Project Report

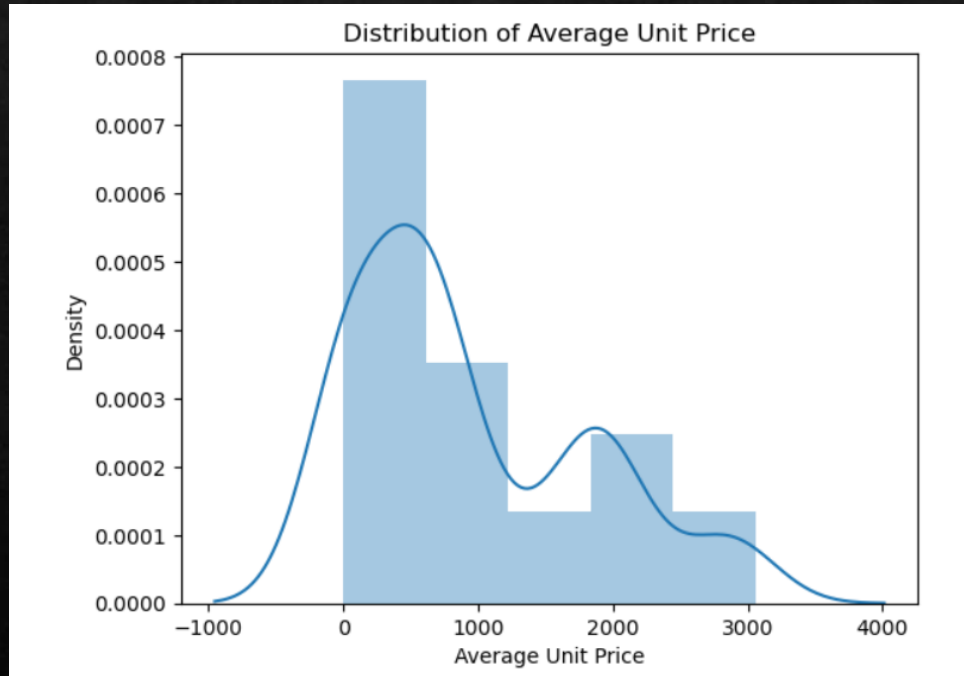
Details of Data:

- **Customer Data:** This file consists of the features related to the data about the customers i.e. 'CustomerKey', 'FirstName', 'LastName', 'FullName', 'BirthDate', 'MaritalStatus', 'Gender', 'YearlyIncome', 'TotalChildren', 'NumberChildrenAtHome', 'Education', 'Occupation', 'HouseOwnerFlag', 'NumberCarsOwned', 'AddressLine1', 'DateFirstPurchase', 'CommuteDistance'.
- **Product Data:** This file consists of the features related to the data about the product i.e. 'ProductKey', 'ProductName', 'Subcategory', 'Category', 'StandardCost', 'Color', 'List Price', 'DaysToManufacture', 'ProductLine', 'ModelName', 'Photo', 'ProductDescription', 'StartDate'.

Details of Data:

- Sales Data: This file consists of the features related to the data about the Sales i.e. 'ProductKey', 'OrderDate', 'ShipDate', 'CustomerKey', 'PromotionKey', 'SalesTerritoryKey', 'SalesOrderNumber', 'SalesOrderLineNumber', 'OrderQuantity', 'UnitPrice', 'TotalProductCost', 'SalesAmount', 'TaxAmt'.
- Territory Data: This file consists of the features related to the data about the Territory i.e. 'SalesTerritoryKey', 'Region', 'Country', 'Group', 'RegionImage'.
- Budget Data: This file consists of the features related to the data about the Budget 2016 i.e., 'Category', 'Subcategory', 'ProductName', 'ProductKey', 'Jan, 2016', 'Feb, 2016', 'Mar, 2016', 'Apr, 2016', 'May, 2016', 'Jun, 2016', 'Jul, 2016', 'Aug, 2016', 'Sep, 2016', 'Oct, 2016', 'Nov, 2016', 'Dec, 2016', 'Grand Total'.

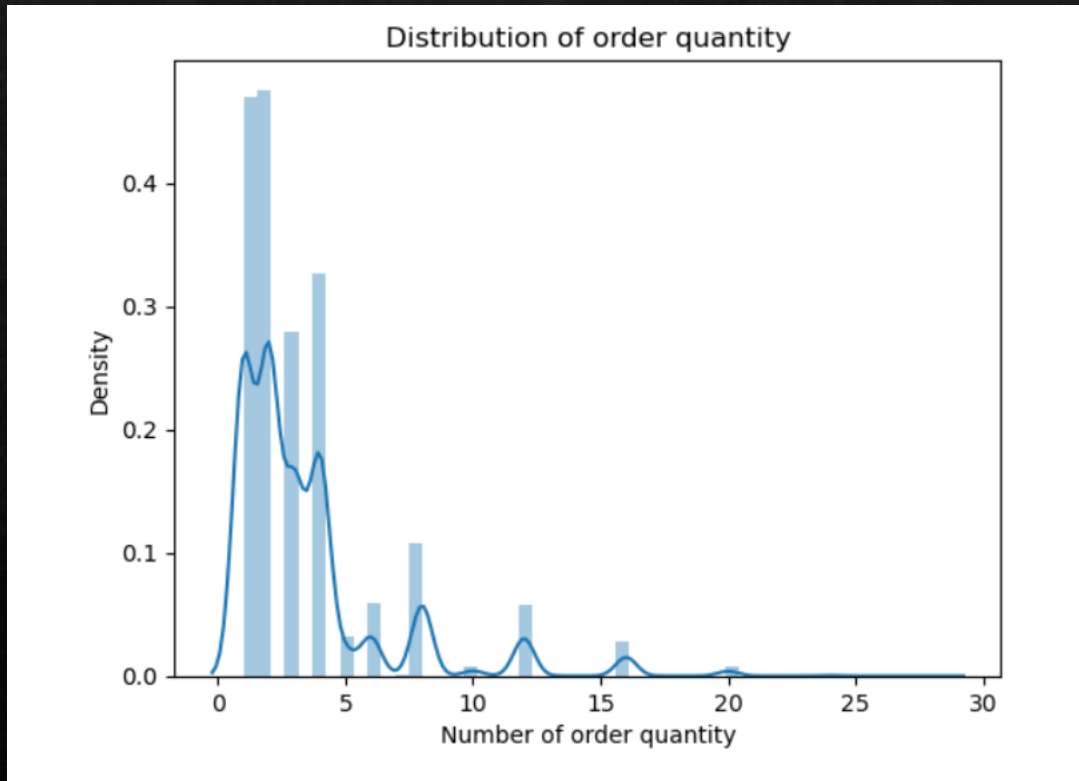
Some Important Insights:



1. Product Price per unit Distribution

According to the above distribution plot, we can conclude that maximum of the product unit price is below \$1000

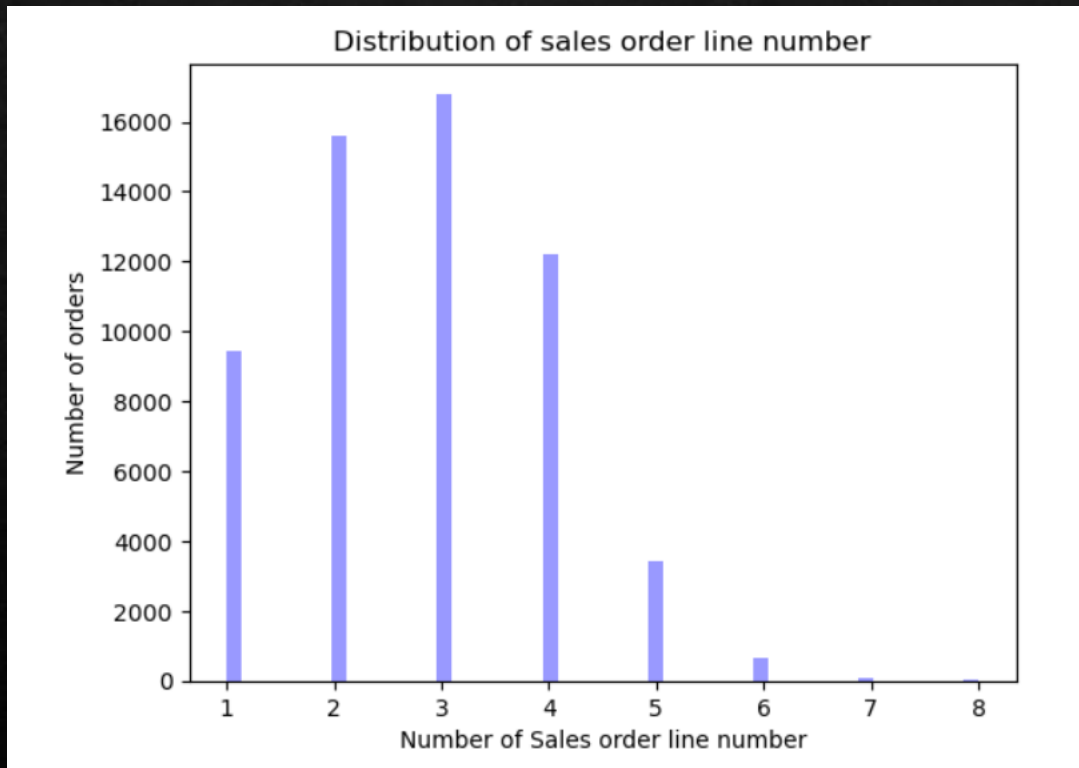
Some Important Insights:



2. Sales order quantity distribution

Maximum quantity ordered for a product is below 5.

Some Important Insights:

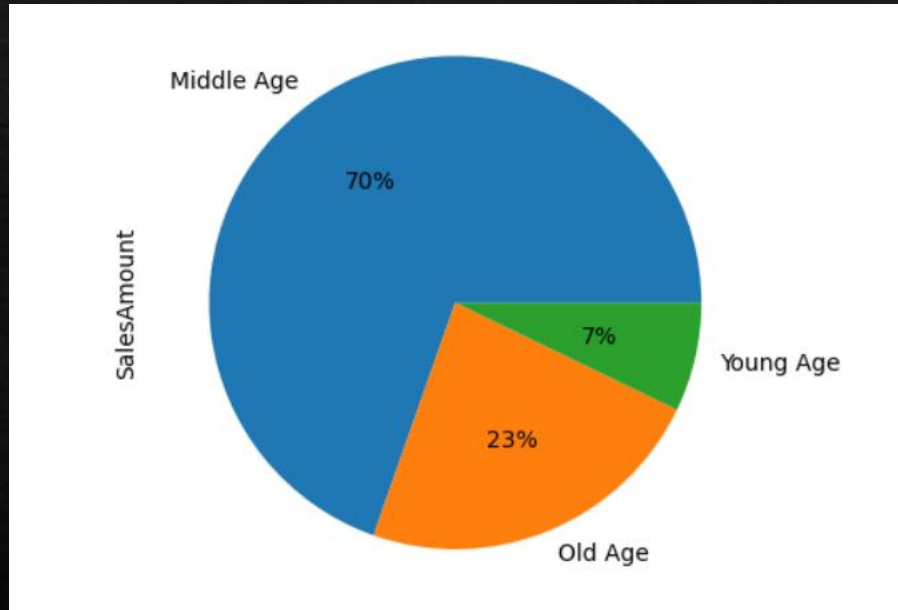


3.Sales order line number distribution

Most of the time three to two products are ordered in a single order

Some Important Insights:

5.Age Distribution



- Middle Age that is 40 to 59 group create more sales than other

Some Important Insights:

sale_year	Category	SubCategory	OrderQuantity
2014	Bikes	Mountain Bikes	616
		Road Bikes	2876
2015	Bikes	Mountain Bikes	1661
		Road Bikes	3284
		Bike Racks	493
		Bike Stands	394
		Bottles and Cages	12055
2016	Accessories	Cleaners	1381
		Fenders	3239
		Helmets	9685
		Hydration Packs	1124
		Tires and Tubes	25518
	Bikes	Mountain Bikes	5490
		Road Bikes	6535
		Touring Bikes	3410
	Clothing	Caps	3178
		Gloves	2143
		Jerseys	5068
		Shorts	1491
		Socks	856
		Vests	824

6. Quantity ordered based on category and subcategory from 2014 to 2016

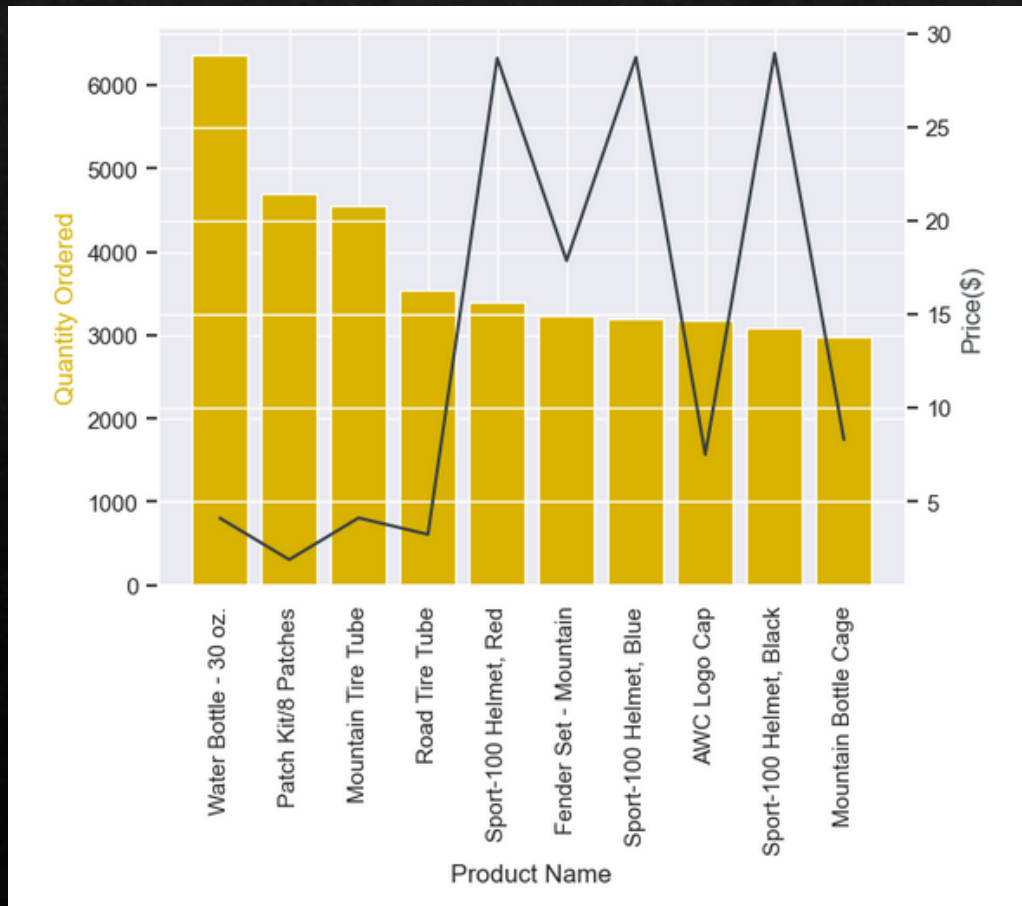
Some Important Insights:

sale_year	Category	SubCategory	profit
2014	Bikes	Mountain Bikes	586874.557600
		Road Bikes	2256280.998300
2015	Bikes	Mountain Bikes	1019388.334900
		Road Bikes	1375064.915000
	Accessories	Bike Racks	23136.960000
		Bike Stands	23689.092000
		Bottles and Cages	34448.978300
		Cleaners	4299.868800
		Fenders	27711.633000
		Helmets	135167.732700
		Hydration Packs	24303.132200
		Tires and Tubes	144793.083200
2016	Bikes	Mountain Bikes	2907361.198000
		Road Bikes	1905953.736400
		Touring Bikes	1454872.695900
	Clothing	Caps	4331.831500
		Gloves	20895.744100
		Jerseys	37965.228300
		Shorts	41973.524600
		Socks	3055.841100
		Vests	20948.777000

7. Overall profit based on order year, category and subcategory

Major Profit is contributed by the Bike Category.

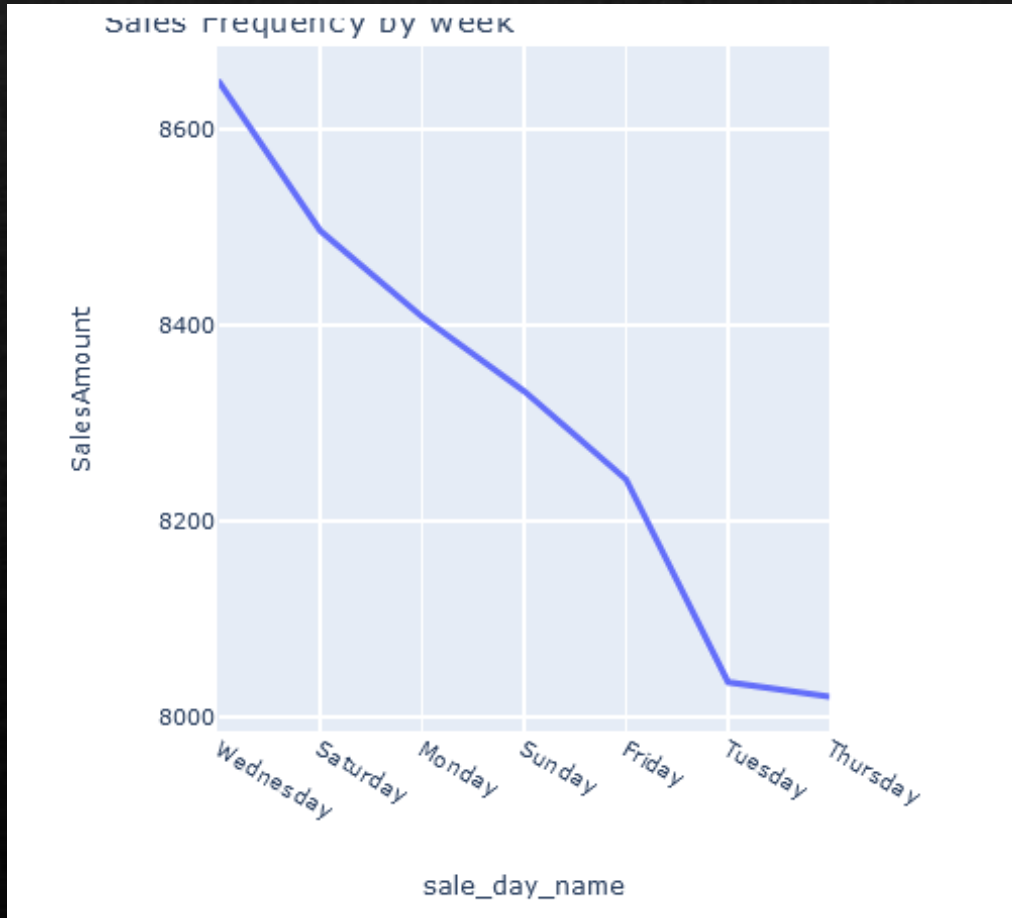
Some Important Insights:



8. Which product sold the most? why do you think it sold the most?

- There is a high negative correlation between Price and number of Quantity ordered
- we can conclude that low price product has high demand

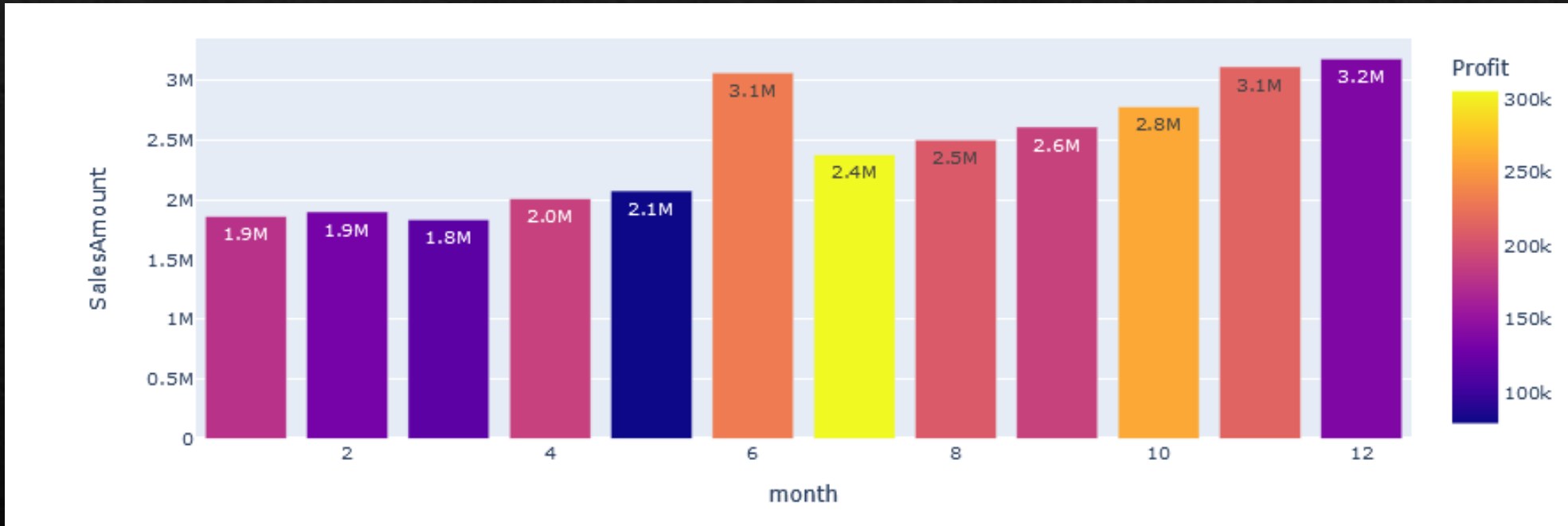
Some Important Insights:



9. What time should we display advertisement to maximize likelihood of customer is buying product?

> High sales orders are seen on Wednesday and Saturday; therefore, we can promote our product during these workweek

Some Important Insights:



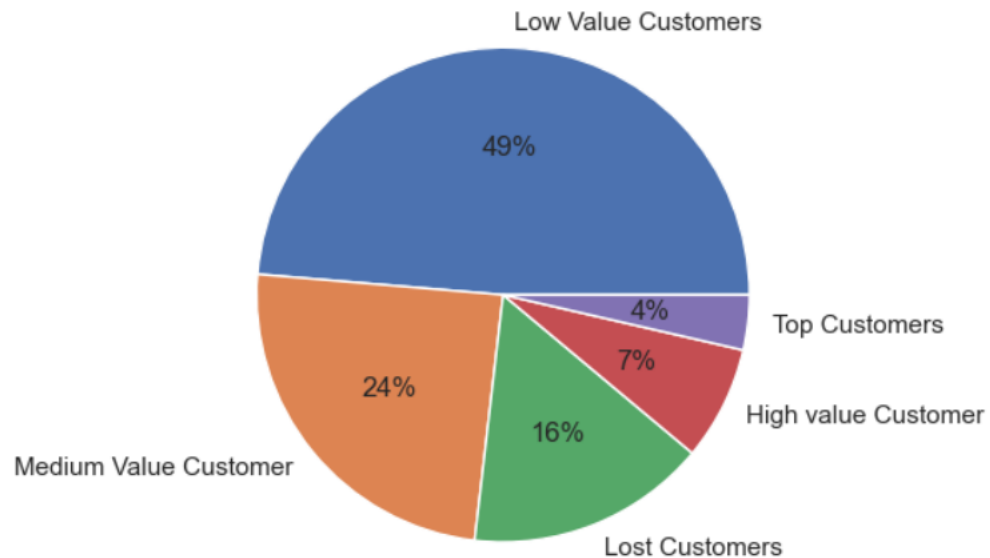
10. What was the best month for sales? How much was earned that month?

➤ Maximum profit earned in the months of June, November, and December

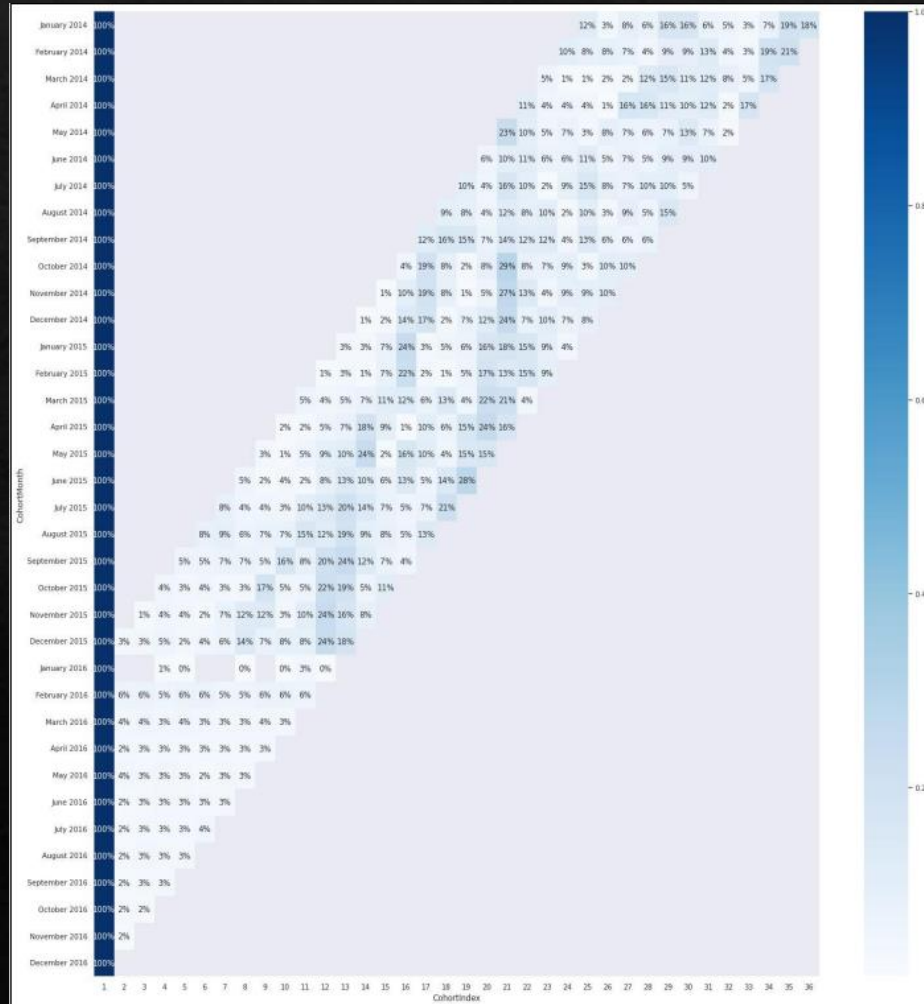
Some Important Insights:

11. Customer segmentation

According to the customer segmentation described above, approximately 15% of our clients are high-value clients, whereas the majority of our clientele are lowvalue and lost clients.



Some Important Insights:



12.Cohort Analysis

- From the heatmap client retention in 2014 was subpar
- Since August of 2015, we have noticed some customers returning, though not in large numbers
- 2016 brought about a slight improvement in retention

Dashboard:

Budget Sales Analysis



CUSTOMER ANALYSIS



SALES ANALYSIS



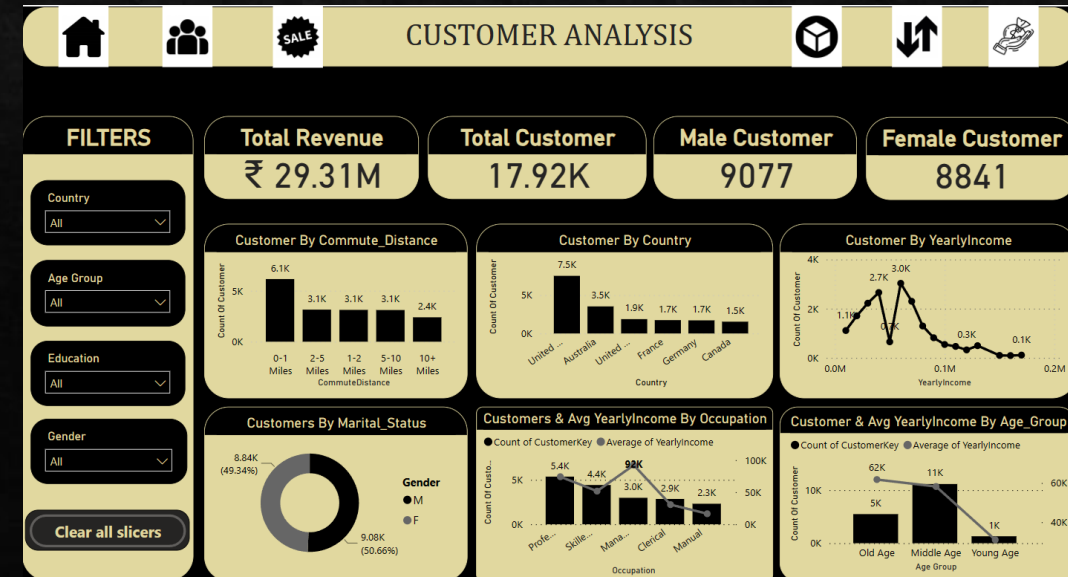
PRODUCT ANALYSIS



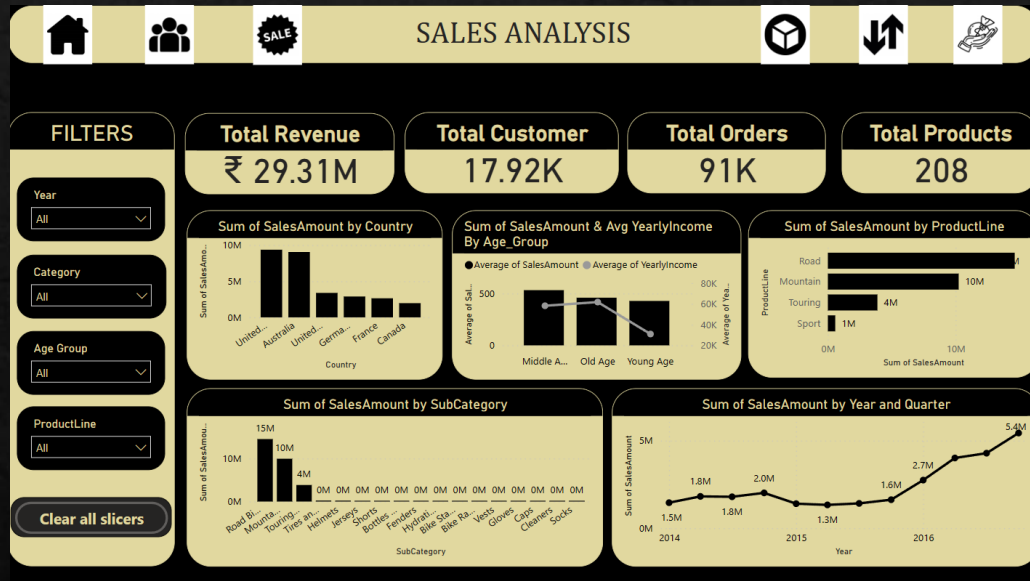
VARIANCE ANALYSIS



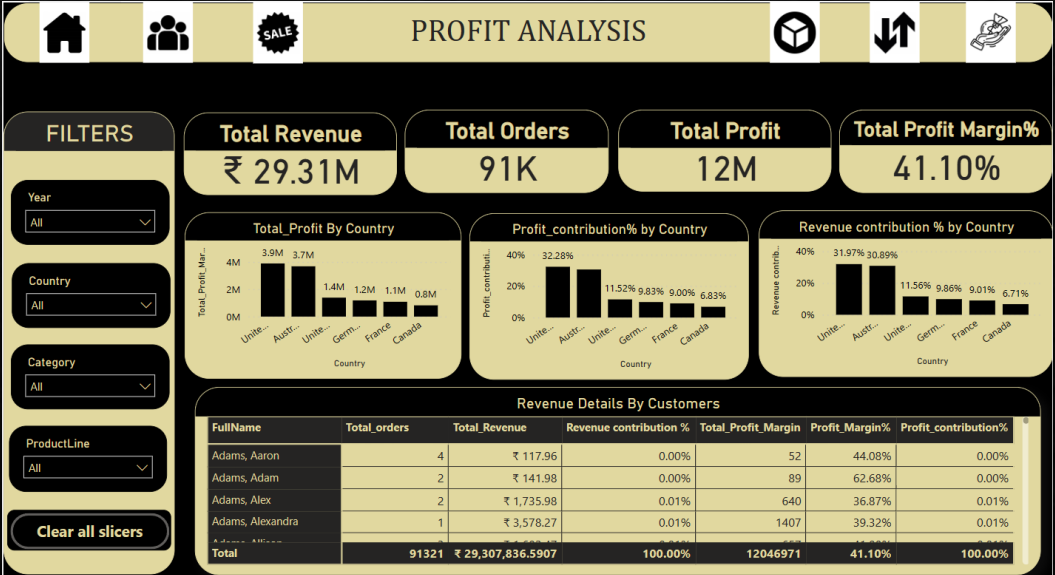
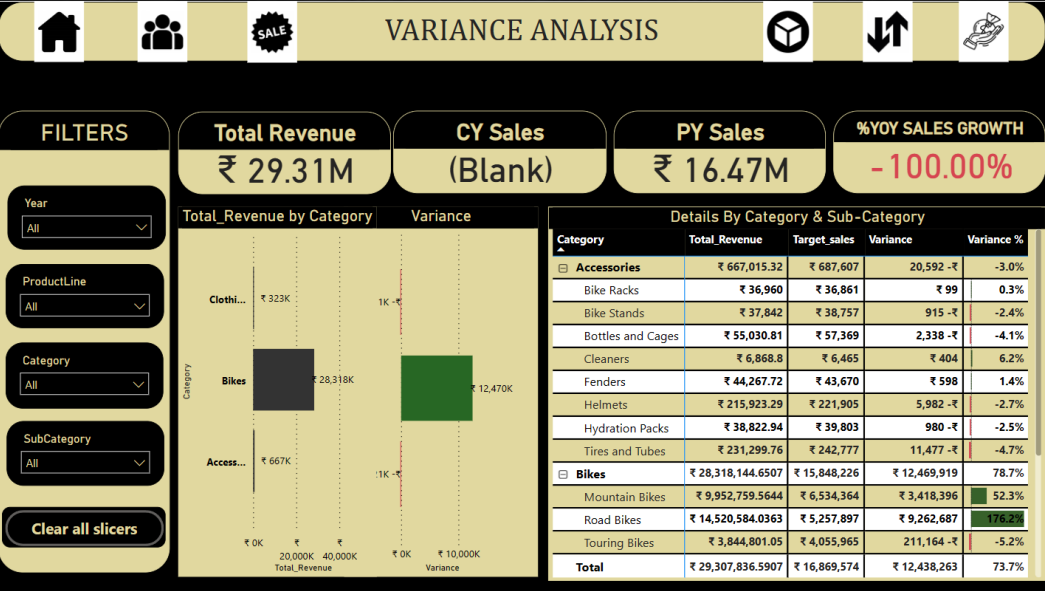
PROFIT ANALYSIS



Dashboard:



Dashboard:



Key Performance Indicator:

- Customer By Commute_Distance
- Customer By Country
- Customer By YearlyIncome
- Customer By Marital_Status
- Customer & Avg YearlyIncome By Occupation
- Customer & Avg YearlyIncome By Age_Group
- Sales By Country
- Sales & Avg YearlyIncome By Age Group
- Sales By ProductLine
- Sales By Sub Category
- Sales By Year,Quarter,Month
- Top 5 Products By Order Quantity
- Top 5 Products By Order Profit
- Bottom 5 Products By Order Profit
- Top 5 Products By Order Sales Amount
- Monthly Details By Category & Sub Category
- Variance to target comparison by category
- Actual sales and target sales matrix
- Revenue contribution by country
- Profit contribution by country
- Profit % by country
- Revenue Details By Customer

Conclusion:

- A sizable portion of the clientele is made up of people between the ages of 40 and 59
- The year 2016 saw an exponential surge in sales
- High quantity of products is ordered from Australia and United States
- Major Profit is contributed by the Bike Category
- The average order has a gap of 7 days between the day the order is ready for export from the factory and the date it was shipped
- Maximum profit earned in the months of June, November, and December
- High sales orders are seen on Wednesday and Saturday, when compared to other weekdays
- There is a high negative correlation between Price and number of Quantity ordered

Conclusion:

- The average amount spent by men without permanent addresses is low, whilst the average amount spent by women without permanent addresses is higher
- Age range of 40-49 and 50-59 is shows high demand compared to other age group
- High salary range leads to increase in revenue
- Customers with a high school diploma and modest annual income buy more products than people with bachelor's degrees
- According to the customer segmentation described above, approximately 15% of our clients are high value clients, whereas the majority of our clientele are low value and lost clients
- Client retention in 2014 was subpar
- 2016 brought about a slight improvement in retention

Q & A

Q1) What's the source of data?

Ans) The Dataset was taken from iNeuron's Provided Project Description Document.

[DhanshriM24/Budget_analysis \(github.com\)](https://github.com/DhanshriM24/Budget_analysis)

Q2) What was the type of data?

Ans) The data was the combination of numerical and Categorical values.

Q 3) What's the complete flow you followed in this Project?

Ans) Refer slide 5th for better Understanding

Q4) What techniques were you using for data?

Ans) -Removing unwanted attributes -Visualizing relation of independent variables with each other and output variables - Removing outliers -Cleaning data and imputing if null values are present. -Converting Numerical data into Categorical values.

Q 5) What were the libraries that you used in Python?

Ans) I used Pandas, NumPy and Matplotlib and Seaborn libraries in Pandas.



Thank You.