Budget Sales Analysis

Detailed Project Report

Priyanka Tanpure

Content

Objective

- To analysis the Budget Sales data, extract necessary information about Customers and Products based on combination of features to help potential buyers purchase the domain they want immediately without the hassle of contacting the seller directly.

Project Details

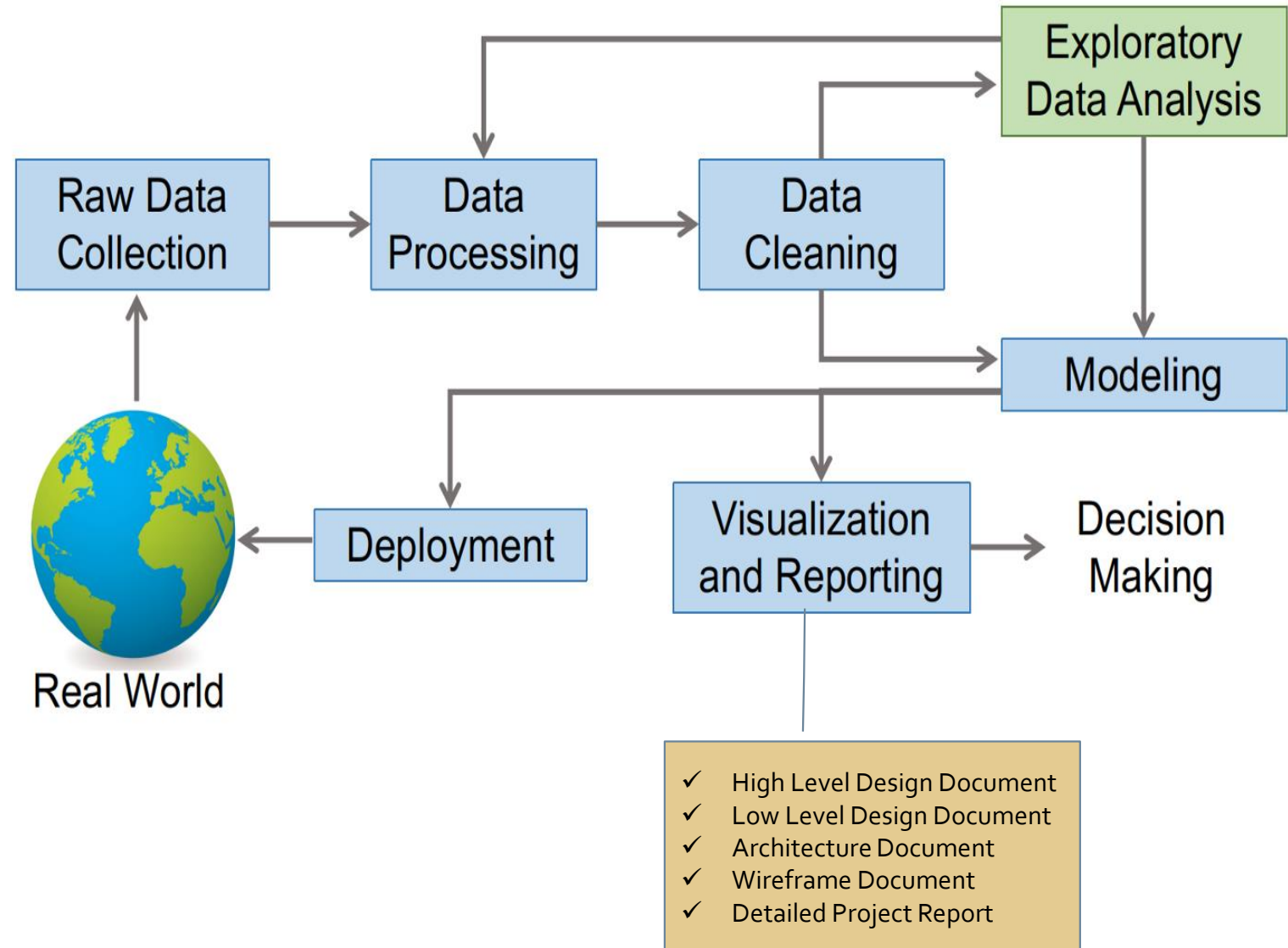|	
Project Title	Budget Sales Analysis
Technology	Business Intelligence
Domain	Retail and Sales
Project Difficulty Level	Advanced
Programming Language Used	Python
Tools Used	Jupyter Notebook, MS-Excel, Power BI

Problem Statement

- Budget and Sales estimates a company's total revenue in a specific period, so they are most important attributes that defines a business's success and failure by focusing on the number of products sold and the price at which they are sold to predict how the company will perform.

- To keep a company's profit high it is important to keep track of budget and sales. Wrong budget allocation and ineffective marketing strategies can cause failure of company's business. Therefore, to keep track of budget and sales it is very important to track various features which contribute to increasing sales and to help allocate budget wisely and effectively.

- Thus, a company needs to study these features like customer behaviour, sales pattern, budget track, market friendly products, demands, demography etc to make effective data driven system to achieve business goals.

Business Objectives:

❑ Performing EDA through Python and find insights.

❑ Extract various information such as Budget, Sales, Variance.

❑ Compare attributes of the dataset to extract necessary information about Products and Customers

❑ Make Dashboard with extracted information from data.

❑ Find key metrics and factors and show relationship between attributes.

Architecture

Raw Data Collection → **Data Processing** → **Data Cleaning** → **Exploratory Data Analysis**

Data Cleaning → **Modeling**

Exploratory Data Analysis → **Modeling**

Deployment → **Real World** → **Raw Data Collection**

Visualization and Reporting → **Decision Making**

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

Dataset Information

Various Excel files available in the dataset: **Customer Data, Product Data, Sales Data, Territory Data, Budget Data**
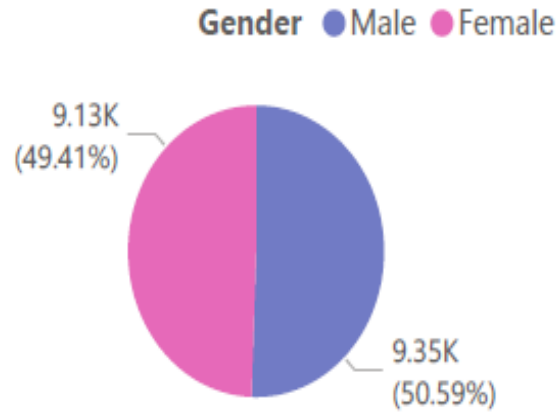
Dataset Information:

- **CustomerKey**: primary key for customer dataser

- **Birthdate**: Birthdate of customer

- **MaritalStatus**: M-Married/S-Single

- **Gender**: M=Male/F-Female

- **TotalChildren**: Total number of children

- **NumberChildrenAtHome**: Number of children staying at home along with their parents

- **Education**: Education qualification

- **Occupation**: Present Occupation

- **HouseOwnerFlag**: 1-Owns House/0-Doesn't own house

- **DateFirstPurchased**: First date of order by the customer

- **ProductKey**: Primary Key for the product dataset

- **ProductName**: Product name with colour of the product

- **SubCategory**: Sub-Category name of the product

- **Category**: Category name of the product

- **ListPrice**: Sale price of the product

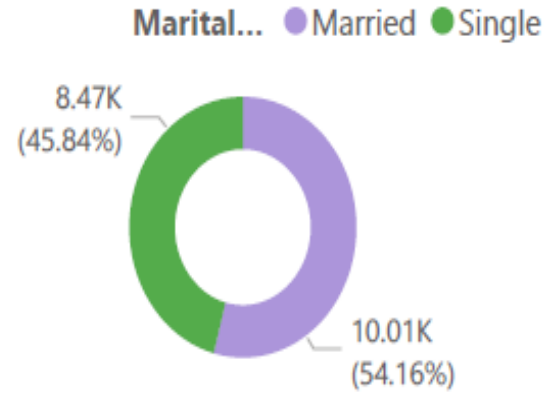- **DaysToManufacture**: Days to manufacture the product after receiving the order

Dataset Information

- **ProductLine**: Product line name

- **ModelName**: Model name of the product

- **ProductDescription**: more details about the product

- **SalesTerritoryKey**: Primary Key of the territory dataset

- **Region**: Reegion name of the order

- **Country**: Country name of the order

- **OrderDate**: Date of the order received

- **ShipDate**: Date when the order left the factory for export

- **SalesOrderNumber**: Invoice number of the order

- **OrderQuantity**: Number of quantities ordered for a product

- **UnitPrice**: Per Unit sale price of the product

- **TotalProductPrice**: Cost of the product

- **SalesAmount**: Total sales price of the product
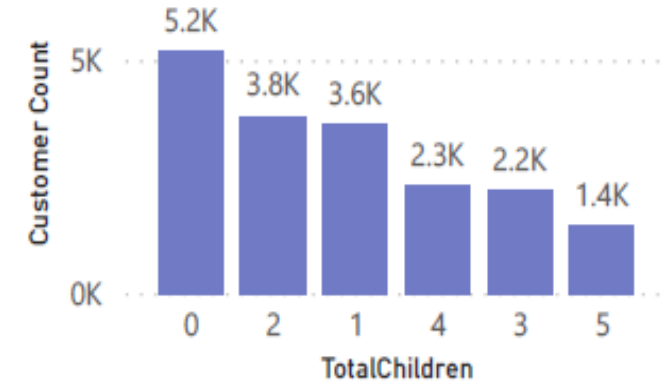
- **TaxAmt**: Tax collected for the product sold

Customer Distribution by Gender

Gender ● Male ● Female

9.13K (49.41%)

9.35K (50.59%)

Customer Distribution by MaritalStatus

Marital... ● Married ● Single

8.47K (45.84%)

10.01K (54.16%)

Customer Distribution by TotalChildren

Customer Count

5.2K, 3.8K, 3.6K, 2.3K, 2.2K, 1.4K

TotalChildren: 0, 2, 1, 4, 3, 5

Insights

Customer Data

Almost 49.41% of customers are Female and 50.59% customers are Male

Almost 45.84% of customers are Single and 54.16% customers are Married

Customers with Total Children between 0-2 are more
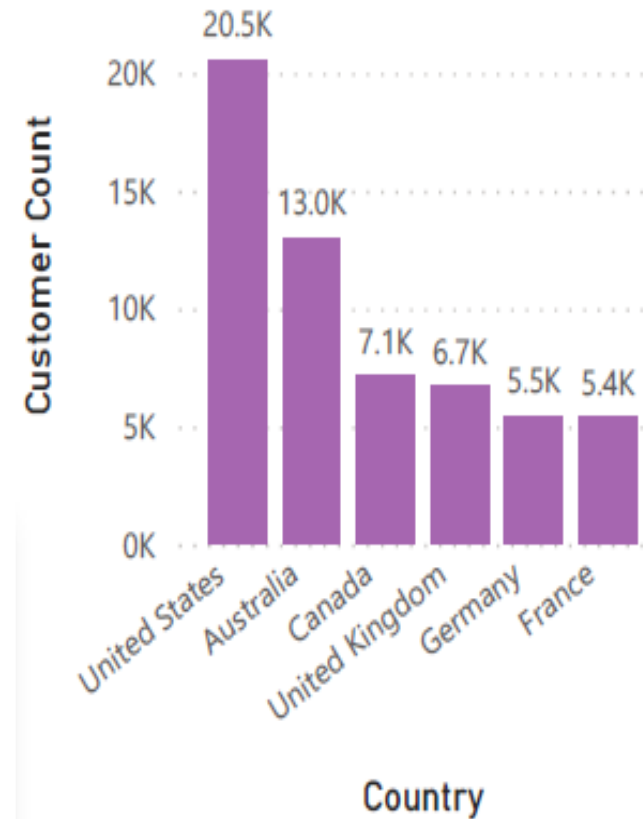
Customer Data

Almost more than half i.e 33.5K customers are from United States and Australia

Nearly 57.5% of customers belong to United States and Australia
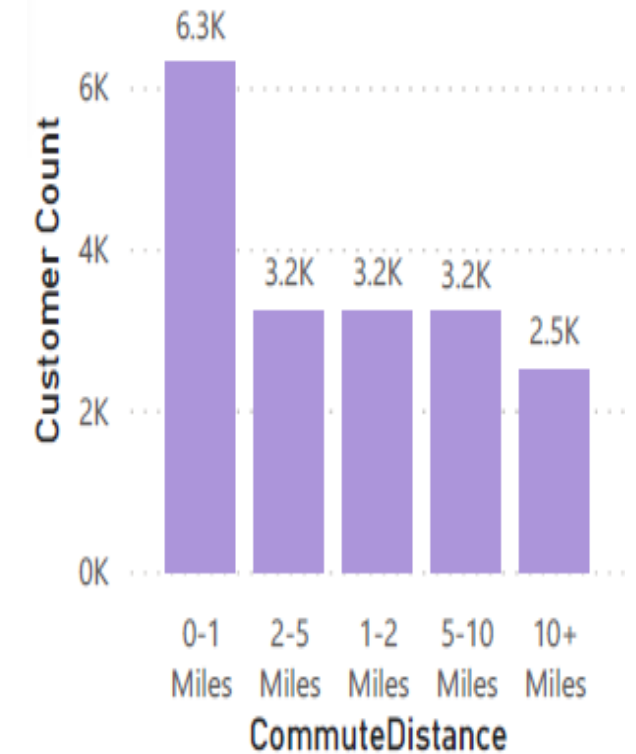
More than 35% of customer alone belong to United States

Customers with commute distance 0-1 Miles are more in number
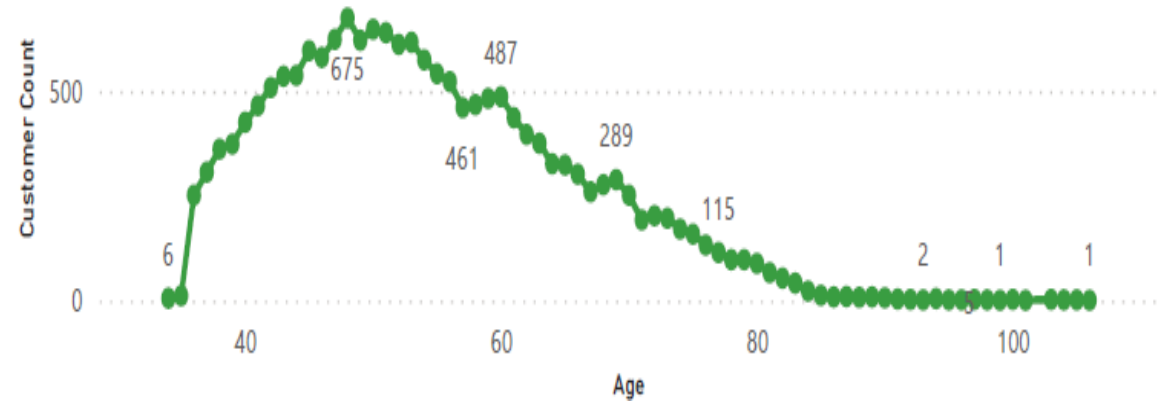

Customer Distribution by Country

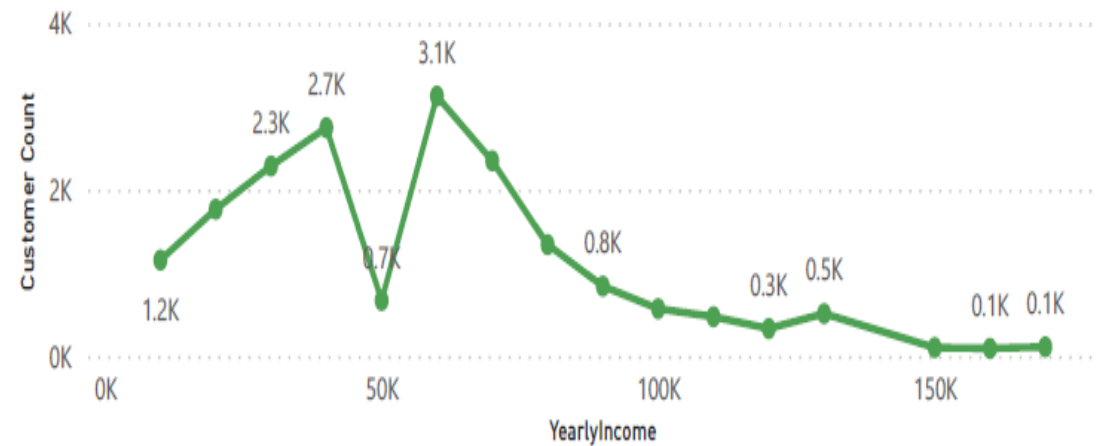Customer Distribution by CommuteDistance

Customer Data

Customers Vs Age

Customers in the Middle Age group 40-60 years are more

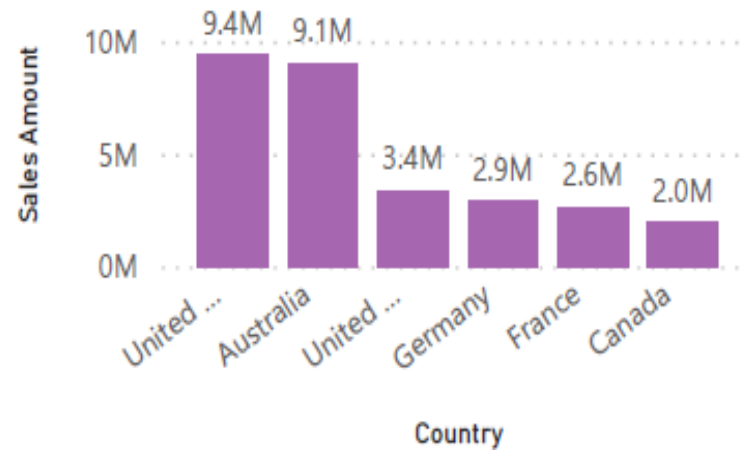**Customer Distribution by YearlyIncome**

Customers with Yearly Income between 10K-40K and 60K-70K are more

Yearly Income of most of the customers is below 100K

Sales Data

Sales Distribution by Country

Sales Distribution by CommuteDistance
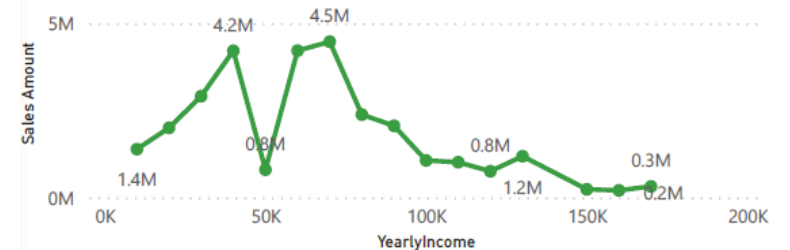

Sales Distribution by YearlyIncome

- Most of the Sales are from United States and Australia i.e. almost 18.5 million of sales

- Nearly 10.9 million of sales are from UK, Germany, France, Canada

- Sales are high for commute distance between 0-1 Miles

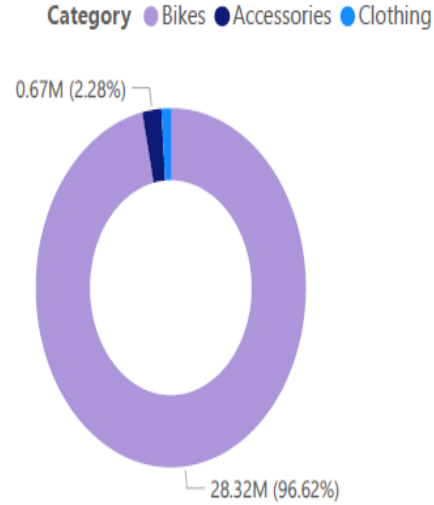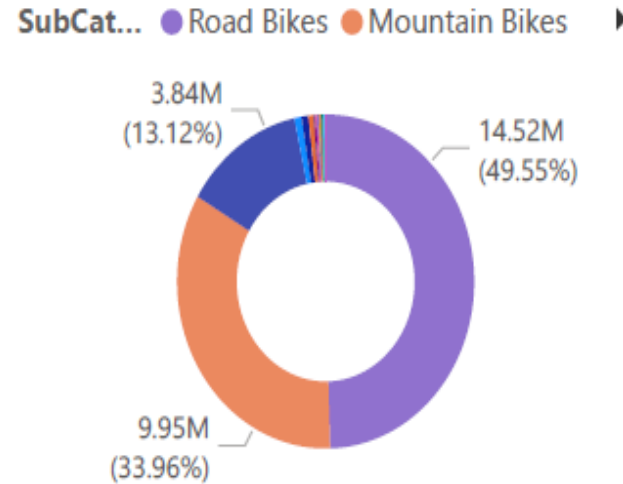- Sales are high for Yearly Income between 10K-40K and 60K-70K

Sales Based on Category

Category ● Bikes ● Accessories ● Clothing

0.67M (2.28%)

28.32M (96.62%)

Sales Based on SubCategory

SubCat... ● Road Bikes ● Mountain Bikes ▶

3.84M (13.12%)

14.52M (49.55%)

9.95M (33.96%)

Sales Based on Color

Color ● Black ● Red ● Silver ● Yellow ● Blue ● NA ● Multi ○ White

8.8M	7.7M	5.1M	4.9M	2.3M	0.4M	0.1M	0.0M
Black	Red	Silver	Yellow	Blue	NA	Multi	White

SalesAmount — Color
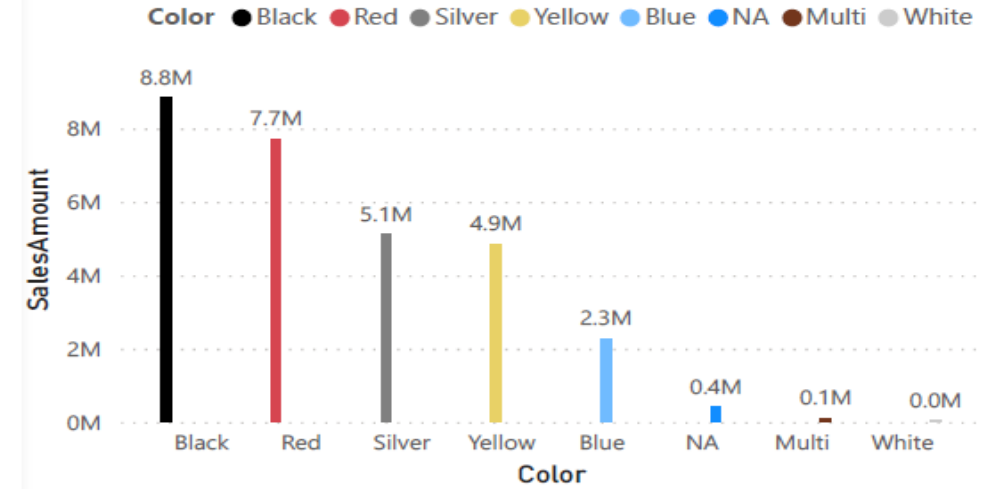
Product Data

Almost 96.62% of sales comes from the product category Bikes

Others are 2.28% of sales from Accessories and Cloth category

Subcategory sale of Road Bike is highest i.e. 49.55%

Black and Red color products are more liked by the customers

Almost more than half of the sales are formed by 16.5M products of color Black and Red

Next sold colors are Silver and Yellow

Key Performance Indicators (KPIs)

Key indicators displaying a summary of the Budget Sales Analysis and its relationship with different metrics:

1. Customer Distribution by Age, Marital Status, Total Children

2. Customer and Sales Relation

3. Sales and Customer Distribution based on Countries and Commute Distance

4. Category and Sub-Category wise Distribution of Sales

5. Customers and Sales Distribution by Education, Occupation

6. Yearly Income of Customers and Sales

7. Profit and Budget Analysis

Conclusion

» Nearly 61% of Sales comes from only two Countries i.e. United States and Australia

» Customers from Middle Age group i.e. 40-60 years contribute more to Sales generation

» Nearly 56% of the Sale happens of Black and Red color Products and 34% of Sale from products of color Silver and Yellow

» Both Gender customers are almost same in number i.e. 50%

» Marital Status segmentation of customers is nearly same i.e. Married 51.72% and Single 48.28%

» Sales is more from customers of Professional and Skilled Manual Occupation

» Customers with education as Bachelors and Partial College contribute more in Sales

» Sales come from category Bike nearly 96.62%

» Profit and Sales is increasing significantly throughout in the year 2016

Q & A

Q.1. What's the source of Data?

➢ The Dataset was taken from iNeuron's provided project description document.

➢ **Database list: heart_disease_dataset.csv**

   Link: https://drive.google.com/drive/folders/165Pjmfb9W9PGyorZjHEA22LWoLt3YQ8?usp=sharing

Q.2. What was the type of Data?

➢ The Data has both Numerical and Categorical values

Q.3. What is the complete flow followed in this project?

➢ Refer slide for better understanding

Q.4. What techniques were used for Data?

➢ Data Handling and Cleaning, Removing Outliers

➢ Converting Numerical data into Categorical data

➢ Visualization and Report making

Q.5. What libraries were used in Python?

➢ Pandas, Numpy, Matplotlib, Seaborn

Thank You!