Detailed Project Report

BUDGET SALES ANALYSIS

Archana Bora

1. Problem Statement:

Our "Domain Sale" process is structured to help potential buyers purchase the domain they want immediately without the hassle of contacting the seller directly.

A seller lists a domain for sale at a specific price in our Marketplace. An interested buyer sees this domain for sale and decides to buy it.

2. Objectives:

- The collection includes records for sales orders, customer information, product information, and geographical data.
- In order to deduce important metrics and patterns in the dataset, this project will use the provided data to perform ETL and data analysis.
- Additionally, several visualization's and reports are createdto represent significant linkages.

3. 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.
- Facilitates seamless resource management flow.

4. Data attributes

Customer

CustomerKey	FullName	Birthdate
Maritalstatus	Gender	YearlyIncome
TotalChildren	NumberChildrenAtHome	Education
Occupation	HouseOwnerFlag	NumberCarsOwned
DateFirstPurchase	CommuteDistance	

Product

ProductKey	ProductName	Subcategory
Category	ListPrice	DaysToManufacture
ProductLine	ModelName	ProductDescription
StartDate		

Territory

SalesTerritoryKey	Region	Country
Group		

Sales

ProductKey	OrderDate	ShipDate
CustomerKey	PromotionKey	SalesTerritoryKey
SalesOrderNumber	SalesOrderLineNumber	OrderQuantity
UnitPrice	TotalProductCost	SalesAmount
TaxAmt		

Detailed Project Report

4.1 Dataset information

CustomerKey: Primary key for customer dataset

Birthdate: Birthdate of the customer

MaritalStatus: M- Married / S - Single

Gender: M – Male / F – Female

TotalChildren: Total number of children

NumberChildrenAtHome: Number of children staying along with their parents

Education: Education qualification

Occupation: Present occupation

HouseOwnerFlag: 1– Owns house / 0- Doesn't have a permanent address

NumberCarsOwned: Number of cars owned by the customer

DateFirstPurchase: 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

ProductLine: Product line name

ModelName: Model name of the product

ProductDescription: more details about the product

Detailed Project Report

SalesTerritoryKey: Primary Key of the Territory dataset

Region: Region 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
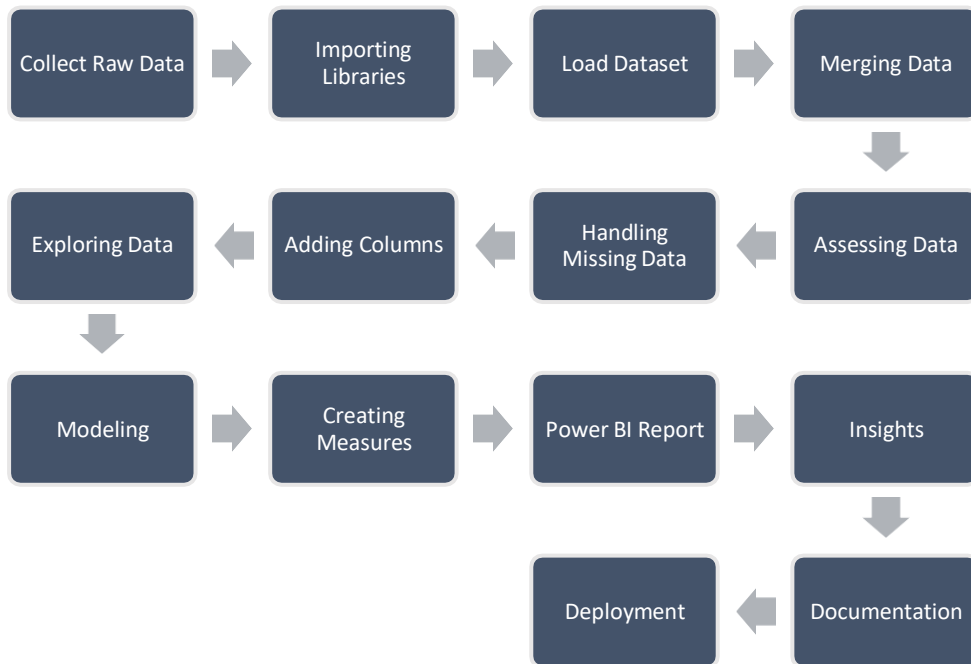
OrderQuantity: Number of quantities ordered for a product

UnitPrice: Per unit sale price of the product

TotalProductCost: Cost of the product

SalesAmount: Total sales price of the product

TaxAmt: Tax collected for the product sold

5. Architecture

```
Collect Raw Data  →  Importing Libraries  →  Load Dataset  →  Merging Data
                                                                  ↓
Exploring Data  ←  Adding Columns  ←  Handling Missing Data  ←  Assessing Data
      ↓
Modeling  →  Creating Measures  →  Power BI Report  →  Insights
                                                          ↓
                          Deployment  ←  Documentation
```

1. **Collect Raw Data -** This step involves extracting the data from different sources relevant to the problem statement or obtaining data from the client

2. **Importing Libraries –** Import analysis related python libraries example – Pandas, Numpy, Plotly, datetime etc

3. **Data Wrangling –** Contains following steps gathering data, assessing data, handling missing data and adding columns

4. **Exploring Data –** Once the data is loaded and pre-processed, we preform data analysis using python libraries and Business Intelligence tools like Power BI

5. **Data Modelling** - Data Modelling is one of the features used to connect multiple data sources in BI tool using a relationship.
A relationship defines how data sources are connected with each other and you can create interesting data visualizations on multiple data sources

6. **Deployment -** The prepared visualizations are deployed on the jovian.com site. Where they will be available publicly

7. Insights

1. Distribution of Avg Unitprice of product
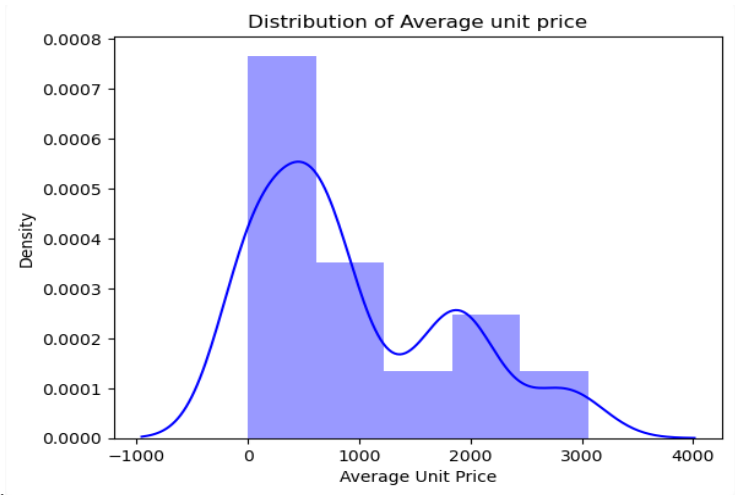
According to the above distribution plot we can conclude that maximum of the product unit price is below $1000

2. Distribution of Order Quantity
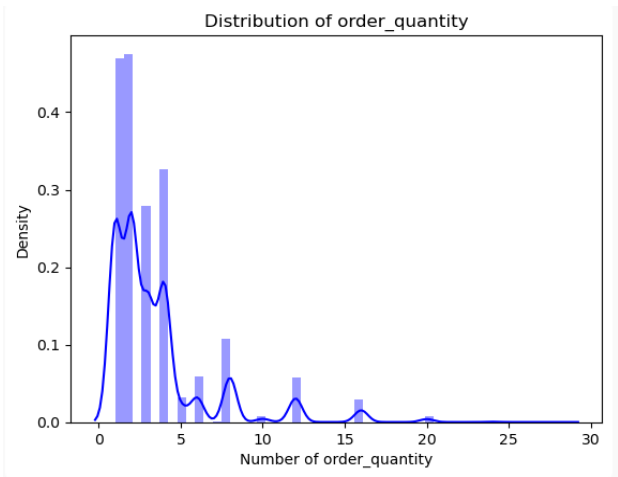
Maximum quantity ordered for a product is below 5

Detailed Project Report

3. Distribution of Sales order line number
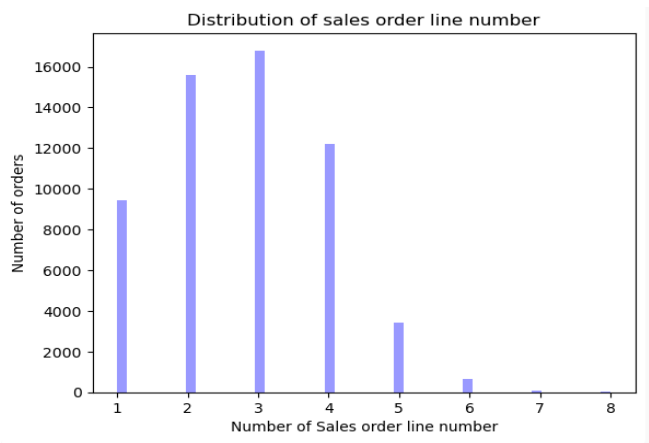
In single order maximum 2 to 3 products are ordered

4. Analysing Sales by Region
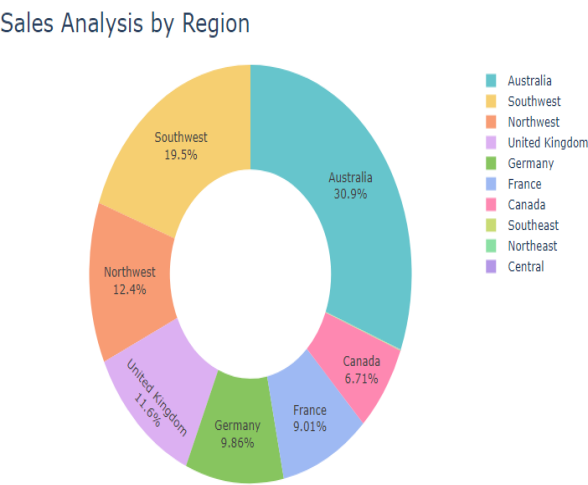
According to the pie chart in terms of Region sales is highest in Australia and and second highest in Southwest region
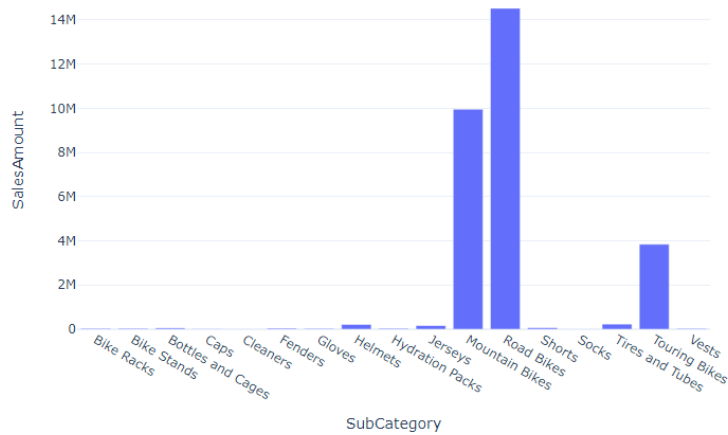
Detailed Project Report

5. Analysing Sales By Category

Sales Analysis by Category

Bikes
96.6%

Accessories
2.28%

Legend:
- Bikes
- Accessories
- Clothing
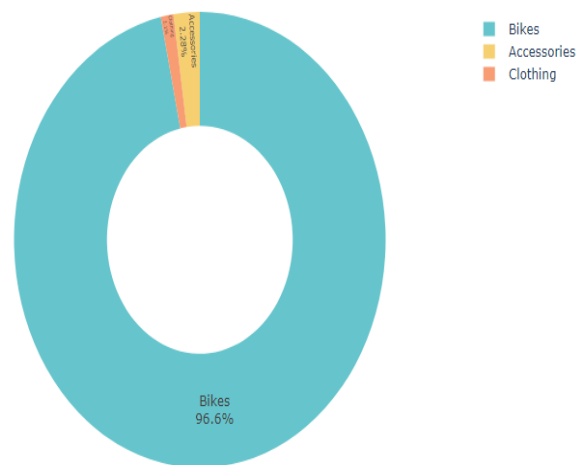
According to the pie chart Bikes
contributes 96.6% in sales

## 6. Analysing Sales By sub category

Sales Analysis by SubCategory

According to the above bar chart the aggregate sales for the
sub category Tires and Tubes has the highest sales
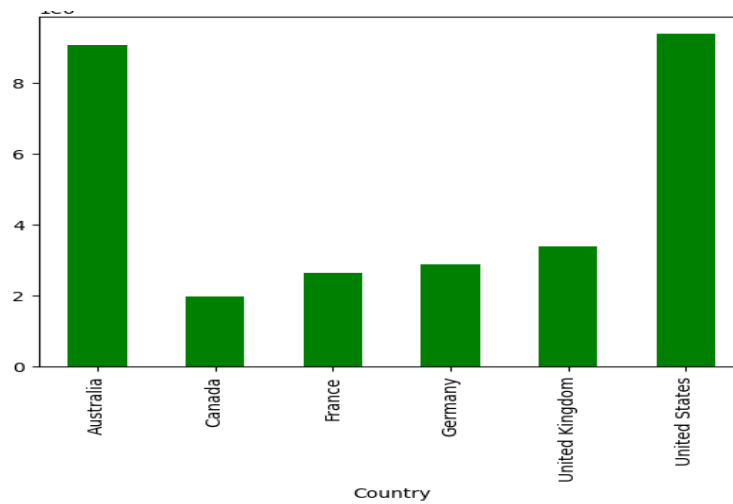
7. Which day of the week has maximum sales?

Daily Sales by Day of the Week
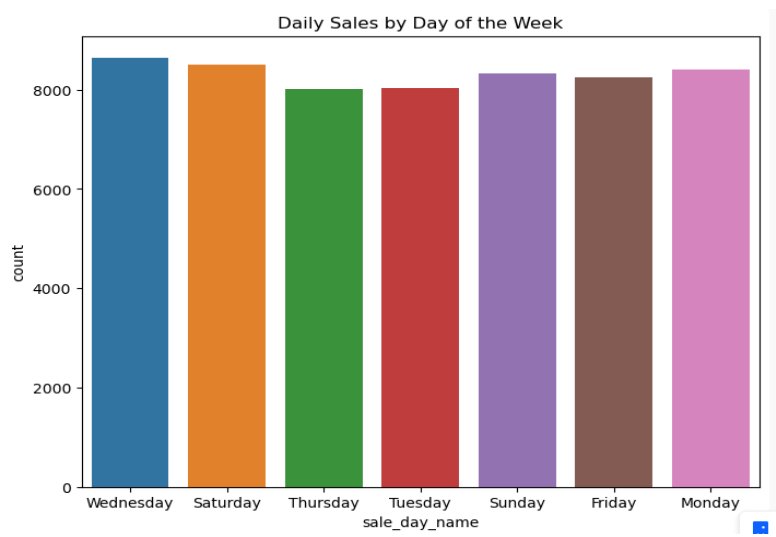
Sales order is highest on Wednesday and Saturday. Tuesday and Thursday has the lowest sales order.
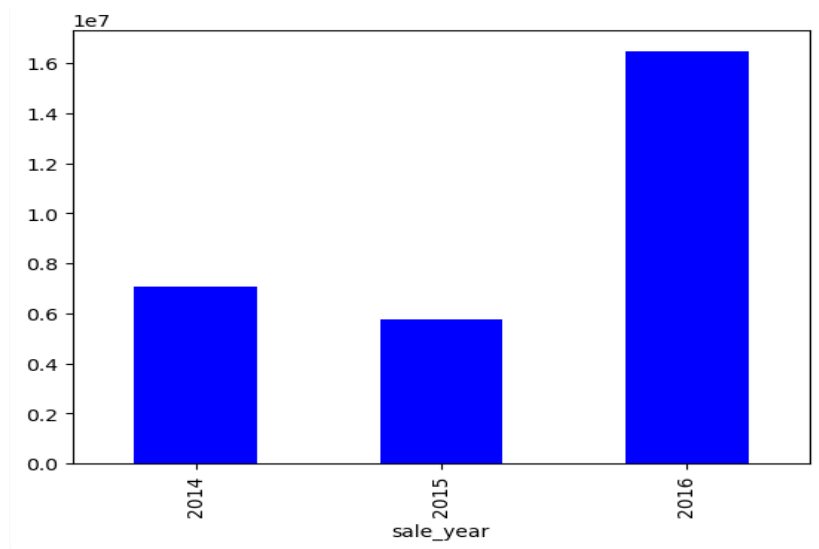
8.Which country ordered the most?

High quantity of products is ordered from Australia and United States
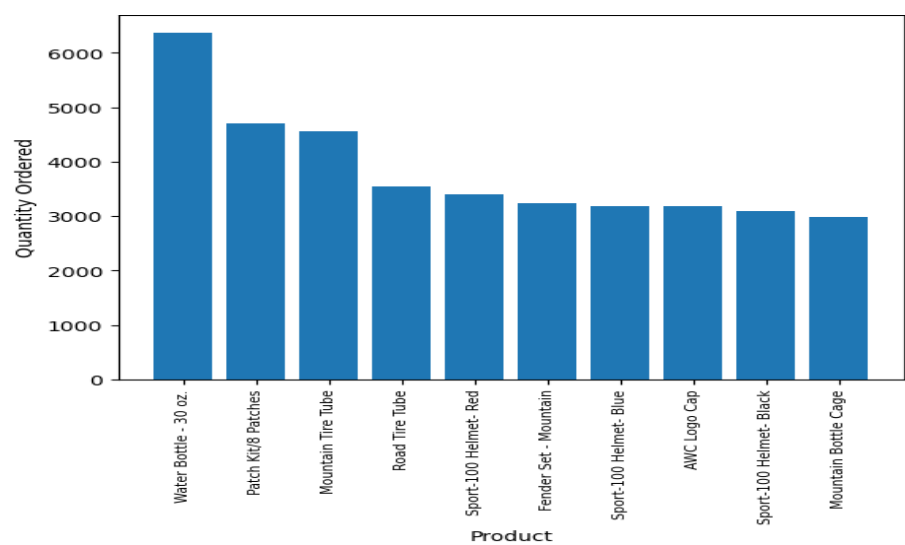
12
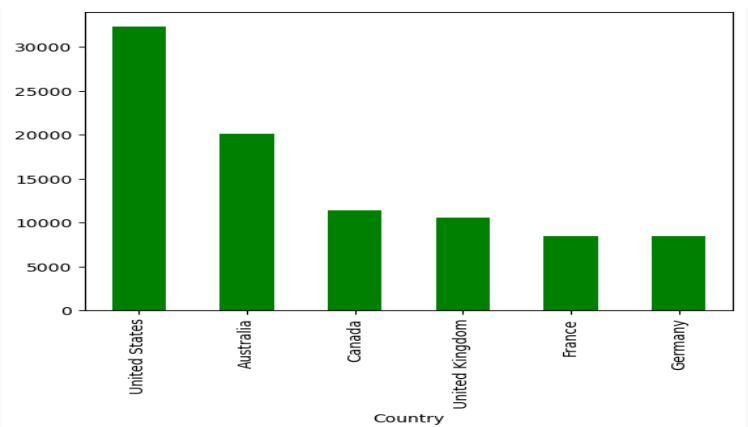
9. Which country has maximum sales?

Maximum sales are from Australia and United States

10. Which year has the maximum sales?

The year 2016 saw an exponential surge in sales

11. What Product sold the most and why?

We can see that what product sold the most, its Water bottle-30 oz, We can also see that patch kit/8 patches, Mountain Tire tube are sold more than other products.

Why are they sold the most. We can assume that they are cheaper than other products so lets prove this hypothesis.
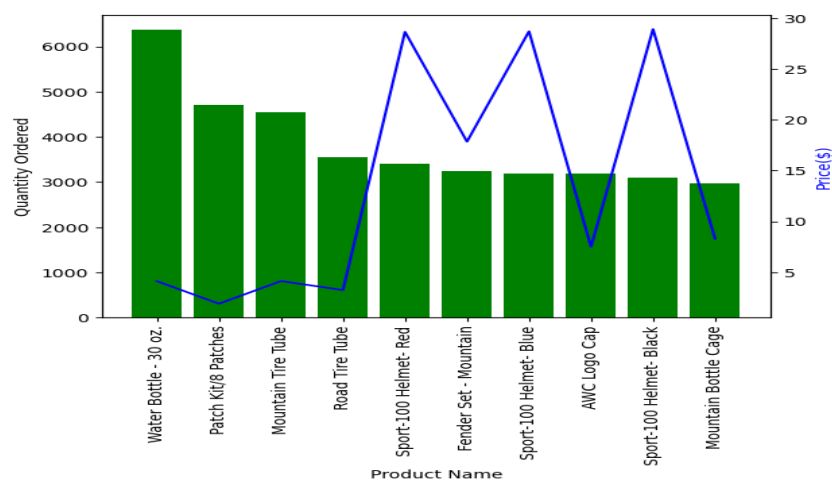
Now we will interpret our results. Our hypothesis is true if the high sold products have low price. From the graph we can see it is the case for Water Bottle-30 oz and all products except the Sports and Helmet. They have decent orders even though they are expensive. We can say that there are many people in the world need sports and Helmet product. So the sports & Helmet products are the exception because they have high demand.

Detailed Project Report
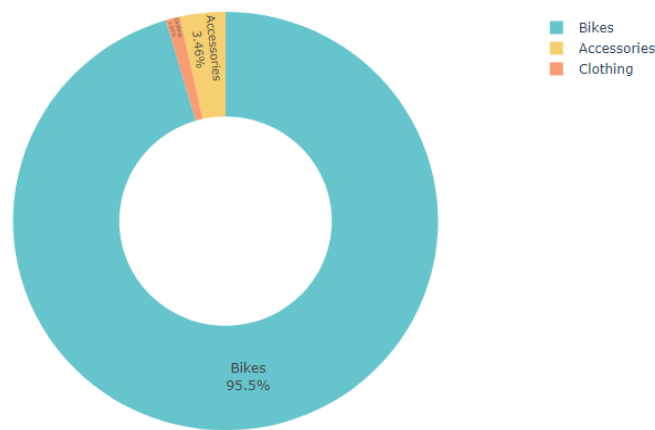
12. Analysing Profit by Region

Profit Analysis by Region
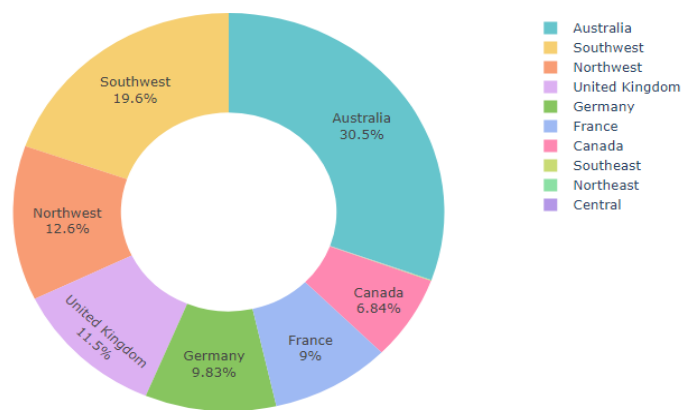
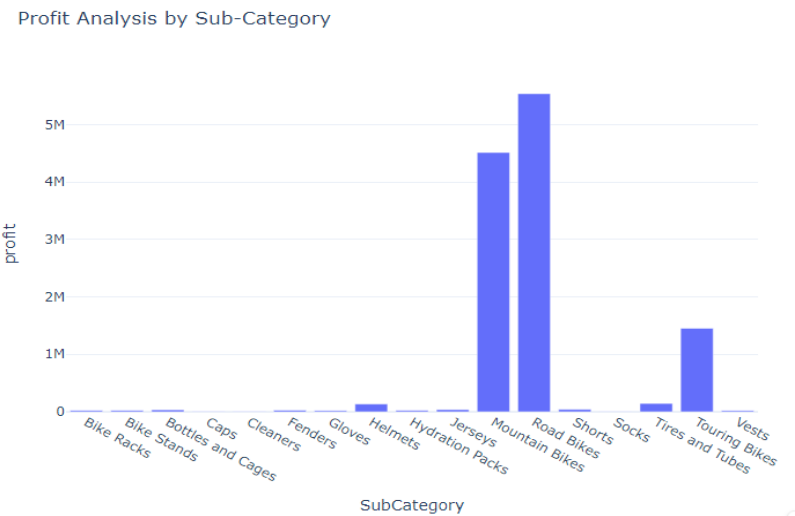In terms of region profit Australia contributes with 30.5% and Southwest 19.6%

13. Analysing Profit by category

Profit Analysis by Category

In terms of category highest profit earned by Bikes category which is 95.5%

Detailed Project Report

14. Analysing Profit by sub category

Profit Analysis by Sub-Category

In terms of subcategory highest profit earned by road bikes and second highest by mountain bikes

Detailed Project Report

15. Quantity with highest Order based on category, subcategory over the years 2014, 2015 and 2016

sale_year	Category	SubCategory	OrderQuantity
2014	Bikes	Mountain Bikes	616
		Road Bikes	2876
2015	Bikes	Mountain Bikes	1661
		Road Bikes	3284
2016	Accessories	Bike Racks	493
		Bike Stands	394
		Bottles and Cages	12055
		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

16. Product name with highest Order Quantity based on category, subcategory over the years 2014, 2015 and 2016

Category	SubCategory	ProductName	OrderQuantity
	Bottles and Cages	Water Bottle - 30 oz.	6370
		Patch Kit/8 Patches	4705
	Tires and Tubes	Mountain Tire Tube	4551
Accessories		Road Tire Tube	3544
	Helmets	Sport-100 Helmet- Red	3398
	Fenders	Fender Set - Mountain	3239
	Helmets	Sport-100 Helmet- Blue	3193
Clothing	Caps	AWC Logo Cap	3178
Accessories	Helmets	Sport-100 Helmet- Black	3094
	Bottles and Cages	Mountain Bottle Cage	2977

17. which is the most profitable country?
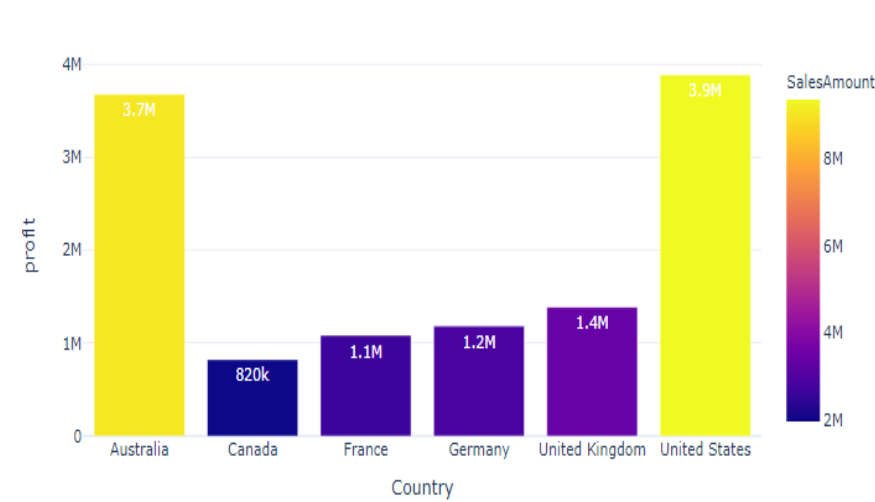
Australia contributes 3.7 Million and US 3.9 Million sales amount

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

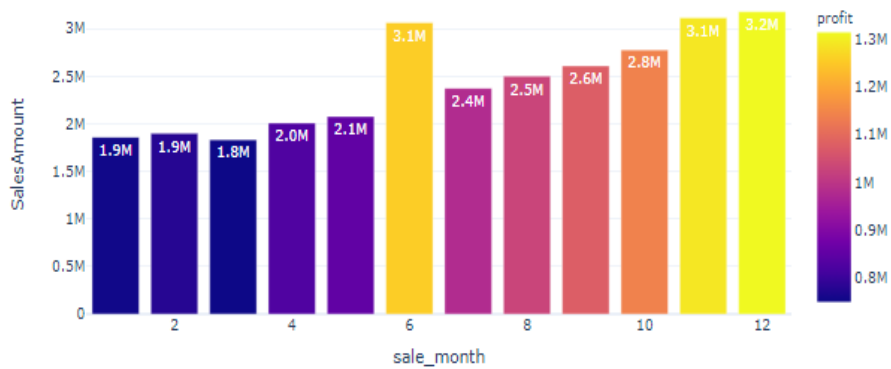| sale_year	Category	SubCategory	profit
2014	Bikes	Mountain Bikes	586874.557600
		Road Bikes	2256280.998300
2015	Bikes	Mountain Bikes	1019388.334900
		Road Bikes	1375064.915000
2016	Accessories	Bike Racks	23136.960000
		Bike Stands	23689.092000
		Bottles and Cages	34448.978300
		Cleaners	4299.868800
		Fenders	27711.633000
		Helmets	135167.732700
		Hydration Packs	24303.132200
		Tires and Tubes	144793.083200
	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

Major Profit is contributed by the Bike Category

19. Which product has the lowest profit?

| Category	SubCategory	ProductName	profit
Clothing	Socks	Racing Socks- L	1474.4574
		Racing Socks- M	1581.3837
Accessories	Cleaners	Bike Wash - Dissolver	4299.8688
	Tires and Tubes	Patch Kit/8 Patches	4314.8350
Clothing	Caps	AWC Logo Cap	4331.8315
Accessories	Tires and Tubes	Touring Tire Tube	4363.8089
Clothing	Jerseys	Long-Sleeve Logo Jersey- XL	4495.6007
		Short-Sleeve Classic Jersey- L	4544.8782
		Long-Sleeve Logo Jersey- S	4610.5777
		Short-Sleeve Classic Jersey- M	4793.2322

Lowest profit contributing products are 'clothing' category

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

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

Detailed Project Report

**21. what is the time difference between order date and shipping date?**
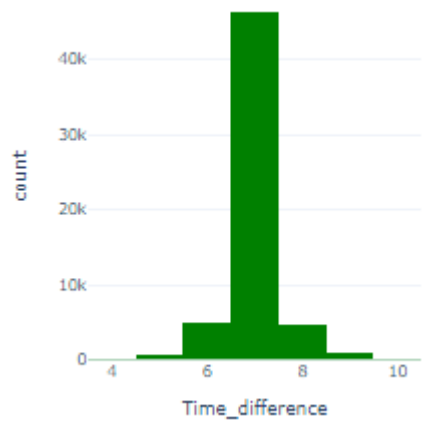
The difference between order date and ship date is maximum 7 days which needs to decrease

**22. which group of Gender contributed more on product?**
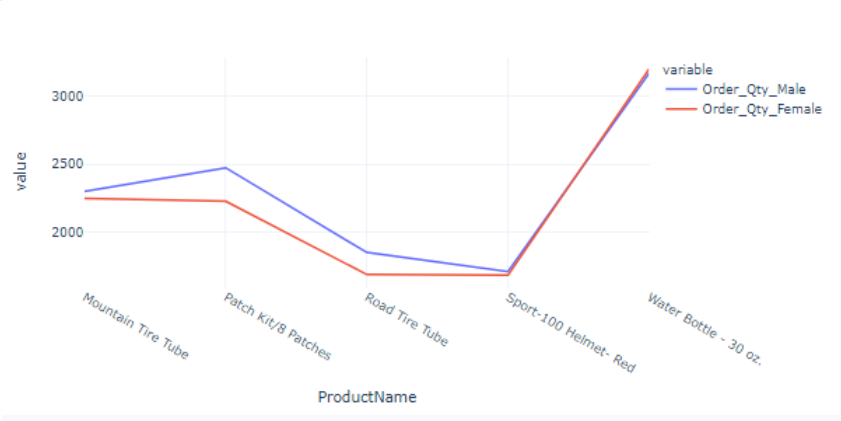
Both male and female contributed more on Water Bottle, male also seen slight high in contributing patch kit/8 patches

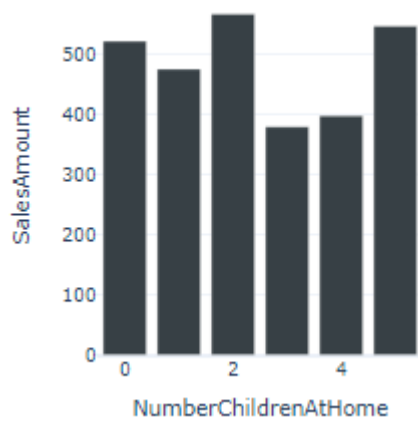**23. Relationship between Education, Occupation and Purchase**
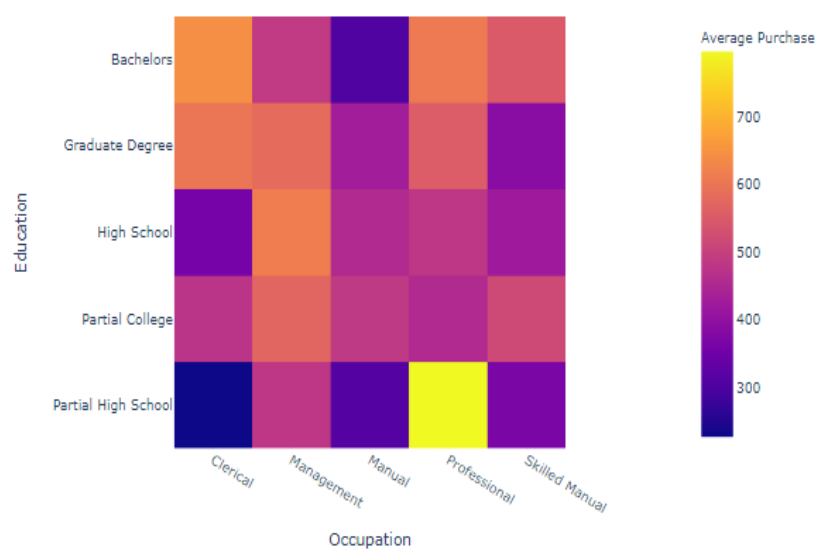
There is a positive correlation with professional and average purchase. Also positive correlation can be seen from Bachelors and average purchase.

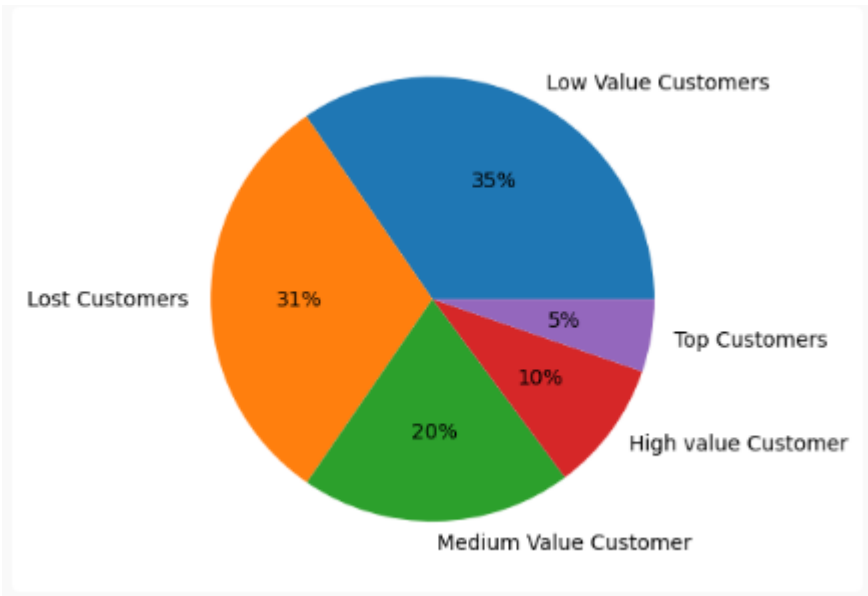**24. Correlation of Income and Purchase amount**
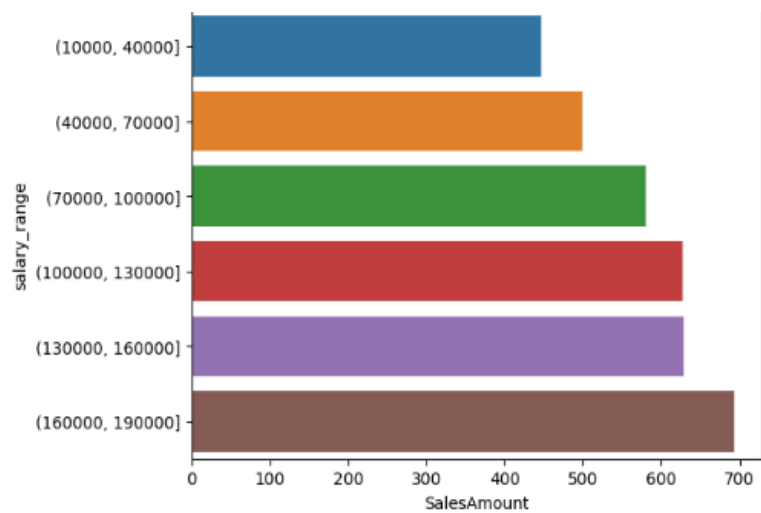
Person with high salary have higher purchase

25. Relationship between Number of Children and Purchase


There is a positive correlation between sales and having 2 numbers of children at home

26. Customer segmentation in percentage using Pie chart

From this pie chart we have seen that 5% are Top Customers and 10% are High Value Customers, whereas the majority of our clients are low value and lost clients.

Detailed Project Report

27. Comparing sales and profit by customer segment
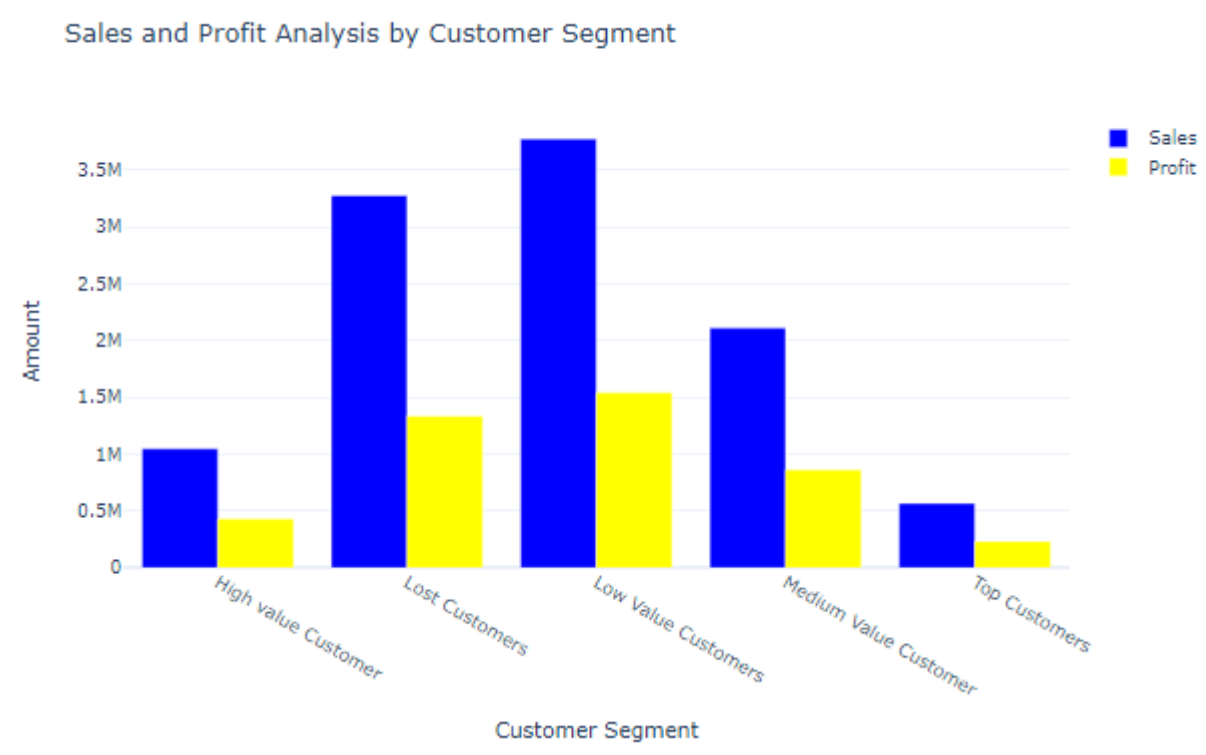
Sales and Profit Analysis by Customer Segment

Profits are higher from the low value customers, lost customers and medium value customers, but the profit from Top customers and High value customer's sales is better in the sales-to-profit ratio. To validate the below table is shown:

	Customer_segment	Sales_to_Profit_Ratio
0	High value Customer	2.435406
1	Lost Customers	2.453406
2	Low Value Customers	2.447959
3	Medium Value Customer	2.440391
4	Top Customers	2.440457

7. Key Performance Indicator

- Sales area chart
- Cost area chart
- Sales by country & Fullname
- Revenue generated by Subcategory
- Sales by Product Line
- Revenue contribution by region
- Profit contribution by region
- Profit % by region
- Sales by Country using map
- Total orders
- Total revenue
- Total Cost
- Total Sales
- Variance to target sales & subcategory
- Actual sales and target sales matrix
- Sales by Year & Month
- Cost by Year & Month
- Sales by Product Name
- Sales by Quarter & year
- Sales by Region using Donut chart
- **Sales and profit by customer segment**

8. Conclusion

- 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 difference between order date and ship date is maximum 7 days which needs to decrease

- 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 positive correlation with professional and average purchase. Also positive correlation can be seen from Bachelors and average purchase.

Detailed Project Report

- High salary range leads to increase in revenue

- There is a positive correlation between sales and having 2 numbers of children at home

- From this pie chart we have seen that 5% are Top Customers and 10% are High Value Customers, whereas the majority of our clients are low value and lost clients.

- Profits are higher from the low value customers, lost customers and medium value customers, but the profit from Top customers and High value customer's sales is better in the sales-to-profit ratio.

9. Q & A

Q1) What's the source of data?

➢ The Dataset was taken from AdventureWorks

Q2) What was the type of data?

➢ The data was the combination of numerical and Categorical values

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

➢ Refer page 4 for better Understandings

Q4) What techniques were you using for data?

➢ Removing unwanted attributes
➢ Visualizing relation of independent variables with eachother
➢ Cleaning data by removing column with missing values
➢ Converting Numerical data into Categorical value

Q 6) What were the libraries that you used in Python?

➢ I used Pandas, NumPy, Matplotlib, Seaborn and Plotly libraries