**ACROPOLIS INSTITUTE OF TECHNOLOGY & RESEARCH, INDORE**

**DEPARTMENT OF COMPUTER SCIENCE**



# CS-605 Data Analytics Lab 3rd Year 6th Semester 2023- 2024

###### SUBMITTED BY - SUBMITTED TO -

**Deepali Arse Prof. Anurag Punde**

###### (0827CS211065)

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Experiment** | **Remarks** |
| 1. | Data Analysis Questions:   1. Data Analysis Principles 2. Statistical Analytics 3. Hypothesis Testing 4. Regression 5. Correlation 6. ANOVA |  |
| 2. | Dashboards:   1. Exploring Car Dataset 2. Cookie Data: Trends and Analysis Report 3. Exploring Loan Dataset 4. Exploring sales on different states of US 5. Store Data Analysis 6. Shop Sale Data Report 7. Sale Samples: A Detailed Report |  |
| 3. | Reports:   1. Exploring Car Dataset 2. Cookie Data: Trends and Analysis Report 3. Exploring Loan Dataset 4. Exploring sales on different states of US 5. Store Data Analysis 6. Shop Sale Data Report 7. Sale Samples: A Detailed Report |  |
| 4. | RAIV Historical Forecast Analysis (2020-2026) |  |

**Data Analysis Principles**

**Assignment-1**

Data Analysis Principles involve systematically applying statistical and logical techniques to describe, condense, and evaluate data. Key principles include understanding the data's source, context, and quality, cleaning the data to remove errors, exploring the data using descriptive statistics and visualization techniques, modeling the data with statistical models for predictions or inferences, and interpreting results to draw meaningful conclusions and make informed decisions.

###### Statistical Analysis

Statistical Analytics uses statistical methods to collect, review, analyze, and draw conclusions from data. This includes descriptive statistics (mean, median, mode, range, variance, standard deviation) to summarize data features, inferential statistics (hypothesis testing, confidence intervals, regression analysis) to extend conclusions beyond immediate data, predictive analytics to forecast future outcomes, and prescriptive analytics to recommend actions based on data analysis.

###### Hypothesis Testing

Hypothesis Testing is a method for making decisions using data from experiments or studies. It involves a null hypothesis (H0) of no effect or difference and an alternative hypothesis (H1) of an effect or difference. The p-value indicates the probability of observing the data if H0 is true, with small p-values suggesting strong evidence against H0. Type I errors (false positives) occur when H0 is wrongly rejected, while Type II errors (false negatives) occur when H0 is wrongly not rejected. The significance level (α), commonly set at 0.05, is the threshold for rejecting H0.

###### Regression

Regression analysis helps understand relationships between dependent and independent variables. Linear regression fits a linear equation to data, multiple regression uses multiple independent variables, logistic regression predicts probabilities for categorical outcomes, and polynomial regression models relationships as nth degree polynomials.

###### Correlation

Correlation measures the strength and direction of relationships between two variables using the correlation coefficient (r), ranging from -1 to 1. A positive correlation means both variables move in the same direction, while a negative correlation means one increases as the other decreases. No correlation indicates no relationship. Importantly, correlation does not imply causation; it simply shows a relationship between variables.

###### ANOVA

ANOVA (Analysis of Variance) is a method for comparing means across multiple groups to determine if at least one group mean differs significantly. One-way ANOVA compares means across one factor with multiple levels, while two-way ANOVA examines the influence of two categorical variables. ANOVA relies on assumptions of normality, homogeneity of variances, and independence of observations. The F-statistic, the ratio of variance between group means to variance within groups, determines the p-value for the test.

SUPER MARKET SALES DASHBOARD



5200

5100

5000

4900

4800

4700

4600

4500

4400

4300

Total Amount of Sales made by each city

5124.651

4998.042

4631.8125

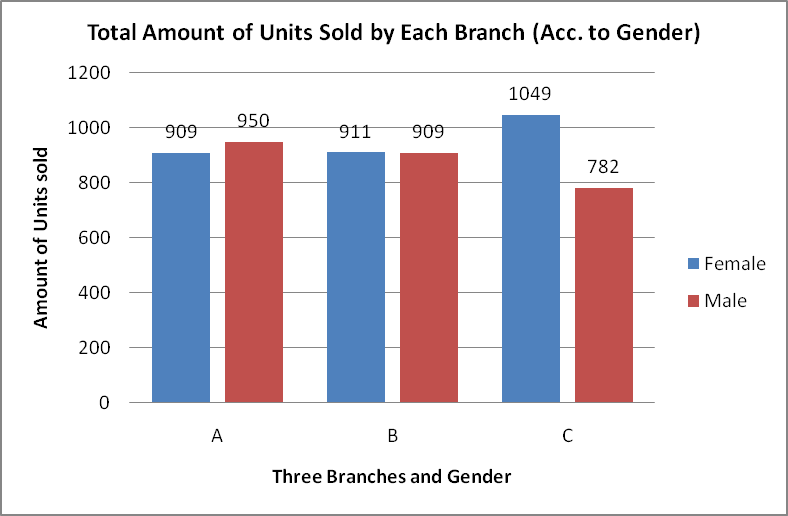
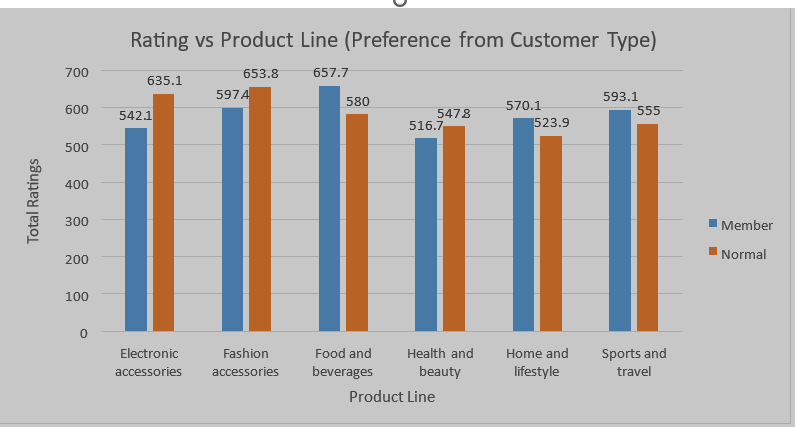
Total

Mandalay

Naypyitaw

City Name

Yangon



Comparison of Lowest and Highest Rating Product on basis of Units Sold

30

26

25

20

18

16

15

10

10

10

10

Electronic accessories Fashion accessories Food and beverages Health and beauty

Sports and travel

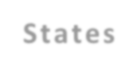
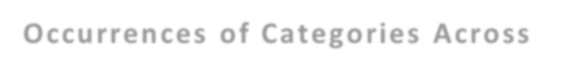
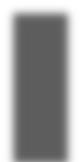
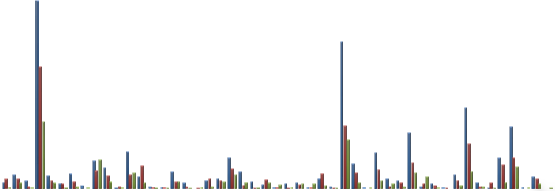
Lowest and Highest Rating Product

**ORDER ANALYSIS DASHBOARD**



SE G ME NT- WI SE SAL E S DI STRI BUTI ON I N SE L E CTE D STATE S

Total



Total Sales by Segment Across U. S.

States

**Occurrences of Categories Across**

**States**

250000

200000

150000

100000

50000

Consumer

Corporate

Home Office

35000000

30000000

25000000

20000000

15000000

10000000

5000000

28852662

10267988

8904250

Total

**AXIS TITLE**



**AVER AGE SALES COM PAR ISON OF CATEGOR IES AN D SUB - CATEGOR IES ACR OSS ALL STATES**

Total

B O O K C A S E S

503.5982243

222419.05

Alabama

California Delaware Georgia Indiana Kentucky Maryland Minnesota Montana

New Hampshire

New York

Ohio Pennsylvania South Dakota

Utah Washington Wyoming

C H A I R S

531.8331647

F U R N I S H I N G S

95.82386466

95360.73

T A B L E S

645.8937197

A P P L I A N C E S

227.9268039

73866.52

A R T

34.01963057

144731.4105

B I N D E R S

134.0675503

E N V E L O P E S

65.03244355

53357.3038

F A S T E N E R S

14.02785047

L A B E L S

34.58746779

**AXIS TITLE**

36333.43

P A P E R

57.4202571

S T O R A G E

263.6338846

79156.003

S U P P L I E S

252.2842826

A C C E S S O R I E S

217.1781746

19854.4984

C O P I E R S

2215.880212

25006.9

M A C H I N E S

1645.553313

P H O N E S

374.1808767

Comparison of Total and Average Sales Across Segments

1500000

1148060.531

1000000

688494.0748

500000

424982.1769

Sum of Sales

Average of Sales2

225.0657775

233.1507195

243.4033086

Consumer

Corporate

Home Office

**ORDER ANALYSIS DASHBOARD**

250000

200000

Total Sales by Segment Across U. S. States

**Occurrences of Categories Across States**

35000000

150000

100000

30000000

25000000

28852662

10267988 8904250 Total

|  |  |  |
| --- | --- | --- |
| 50000 | Consumer | 20000000 |
| 0 | Corporate | 15000000 |
|  | Home Office | 10000000 |
|  |  | 5000000 |
|  |  | 0 |

Furniture Office Supplies

Technology

SE G ME NT- WII ON I N SE L E CTE D STATE S

Total

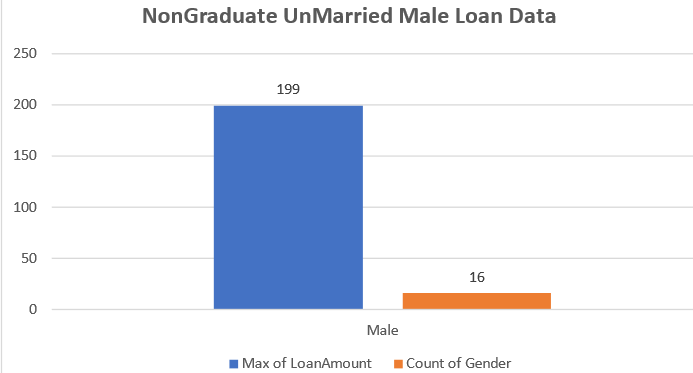
C O N S U M E R C O R P O R A T E H O M E O F F I C E

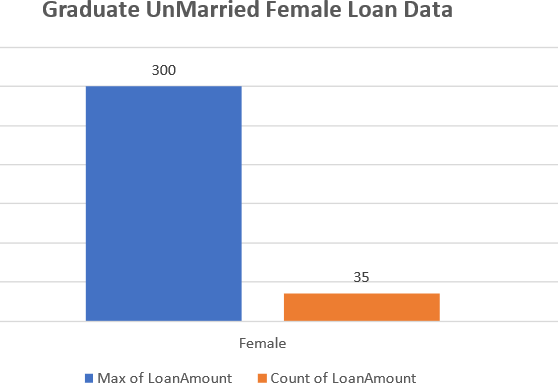
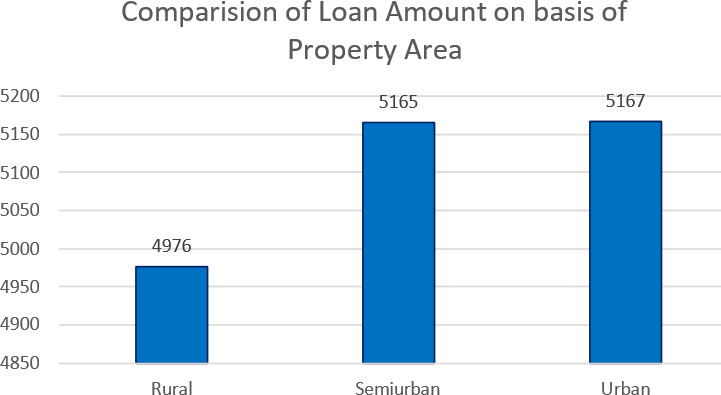
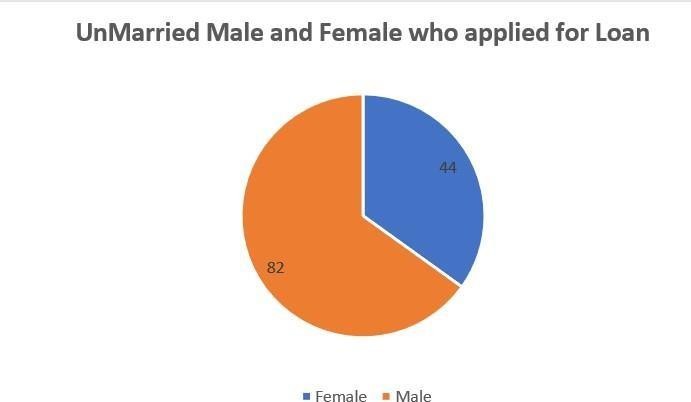
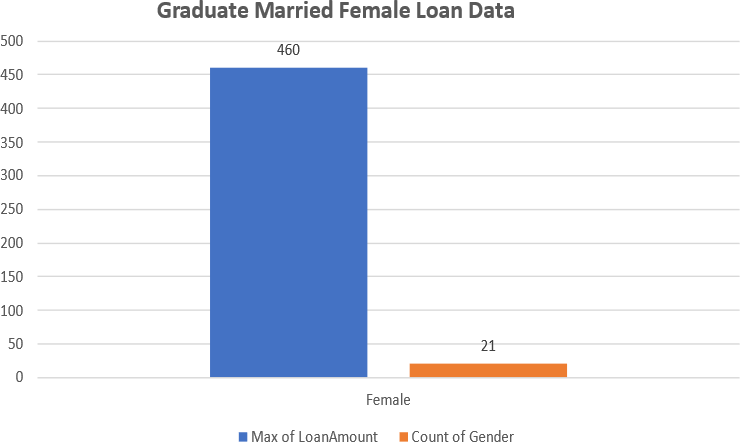
**AVER AGE SALES COM PAR ISON OF CATEGOR IES AN D SUB - CATEGOR IES ACR OSS ALL STATES**

Total

F U R N I T U R E O F F I C E S U P P L I E S T E C H N O L O G Y

**LOAN ANALYSIS DASHBOARD**





Car Model Dashboard

5000

4500

4000

3500

3000

2500

2000

1500

1000

500

Total Cost of Cars Exceeding $2000

4500

4000

3900

4100

3500

3000

2500

2200

Total

Chevrolet

Dodge

Ford

Honda

Nissan

Toyota

2100

3000

3100

3000

3000

Impala

Malibu

Silverado

Charger

Escape

F-150

Fusion

Mustang

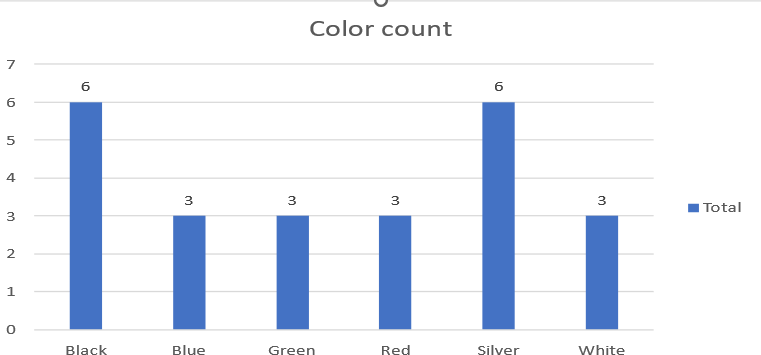
Accord

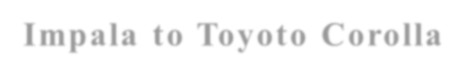
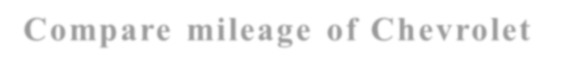
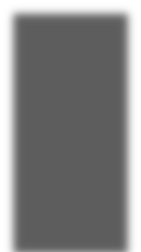
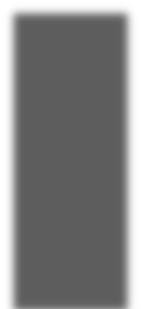
CRV

Altima

Maxima

Corolla





**Compare mileage of Chevrol et Impala to Toyoto Corolla**

120000

114243

100000 92377

80000

60000

40000

Total

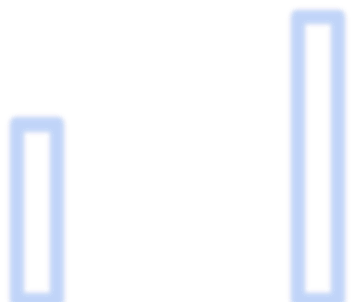
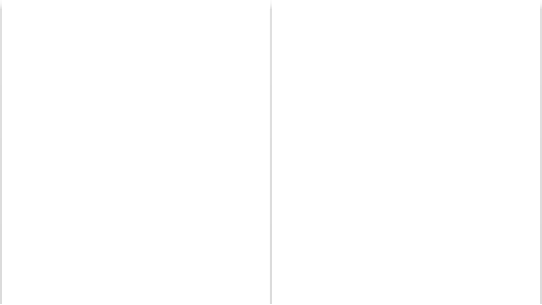
20000

Impala

Corolla

Chevrolet

Toyota



**Buying of any ford car vs honda**

100000

90000

80000

70000

60000

50000

40000

30000

20000

10000

89623.2

55452

Average of Mileage

Average of Price

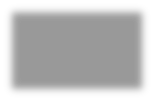
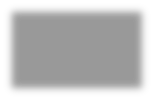
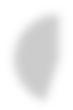
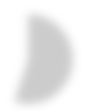
Average of Cost

3581.63030

3193.62500

Ford

Honda



**Comparision of all the cars of silver color to green color in terms of mileage**

**63797.33,**

**45%**

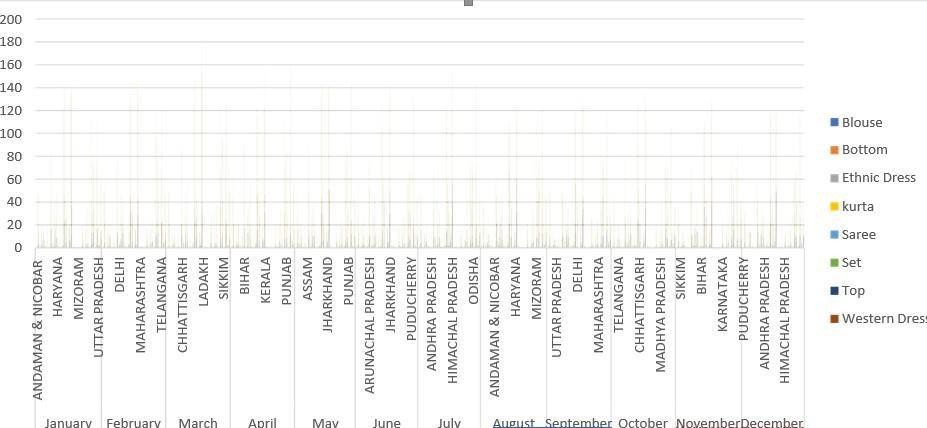
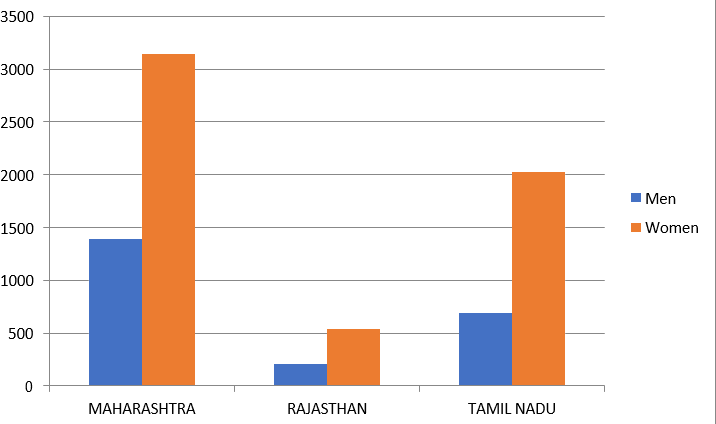
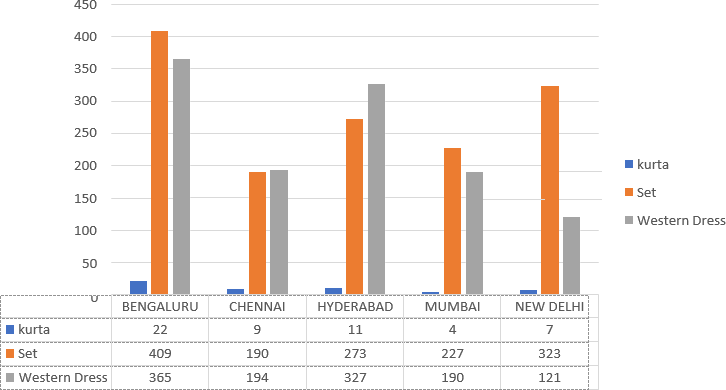
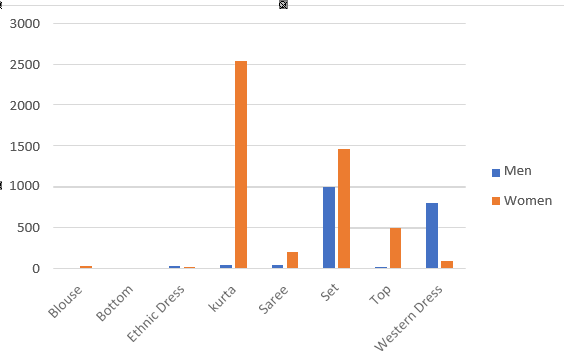
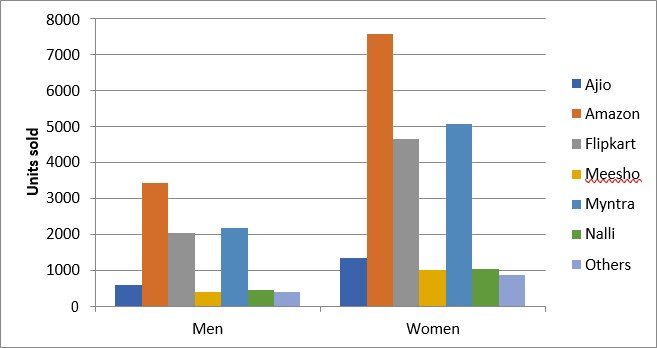
**78103.67,**

**55%**

Green

Silver

# Store Analysis Dashboard



**Shop sales dashboard:-**

**MOST AVERAGE PROFIT**

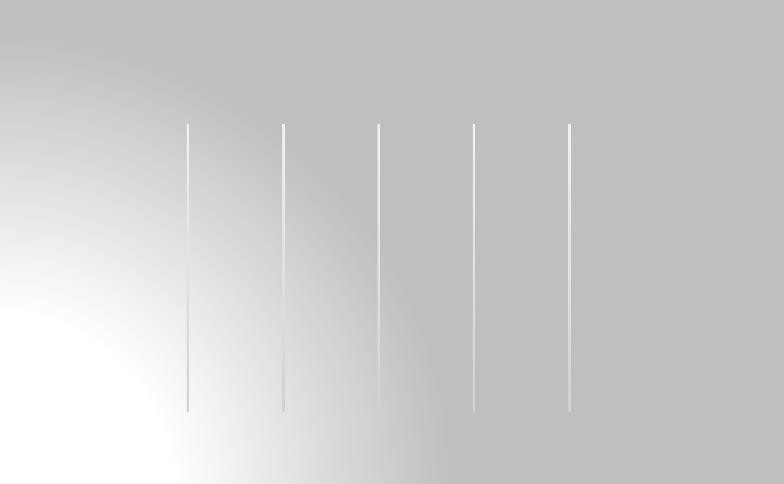
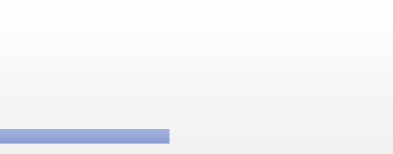
Total



6770.231898

6772.950369

7057.58477



Comparing salesmen on the

basis of Profit Earned

**Most Sold Product**

Vinod

Rohit Ram Rahul

Aman

SEP

AUG JUL

Total

JUN

350000

400000 450000 500000

MAY

Aman Rahul Ram Rohit Vinod

Total 414776.4493541.3476120.4485039.1478167.1

200

400

600

800 1000

**481.0191684**

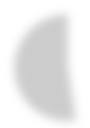
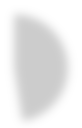
**539.8378925**

Total

**634.6458094**

**712.7183647**

**764.0629212**



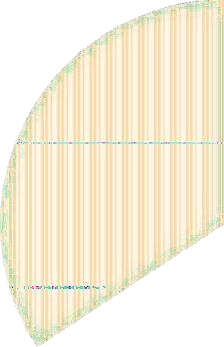
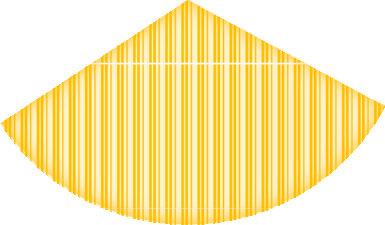
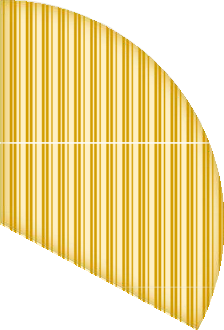
**Most sold**

**52%**

**48**

Computer

Laptop



**AVERAGE SALES**

7057.58477

6770.231898

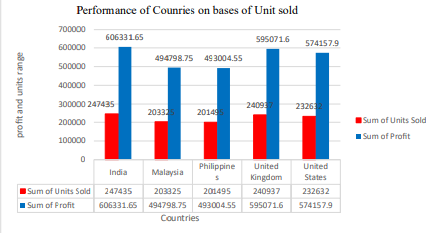
Computer

Laptop

Mobile

6772.950369

### COOKIE DASHBOARD



**PROFIT EARN BY ALL COOKIES IN COUNTRIES**

Total

INDIA

234576

M A L A Y S I A

U N I T E D S T A T E S

189846

208404

INDIA

25192

M A L A Y S I A

25403.2

U N I T E D S T A T E S

25290.4

INDIA

M A L A Y S I A

81852.4

79508.8

U N I T E D S T A T E S

99316

INDIA

M A L A Y S I A

87012.5

68060

U N I T E D S T A T E S

64547.5

INDIA

M A L A Y S I A U N I T E D S T A T E S

72187.5

45725.75

60376.75

INDIA

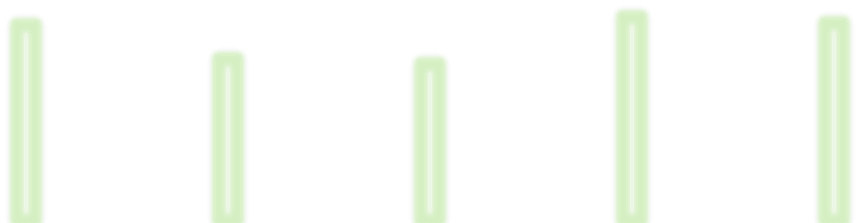
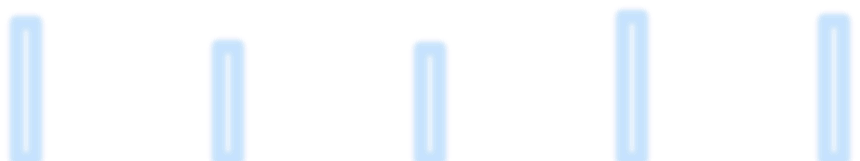
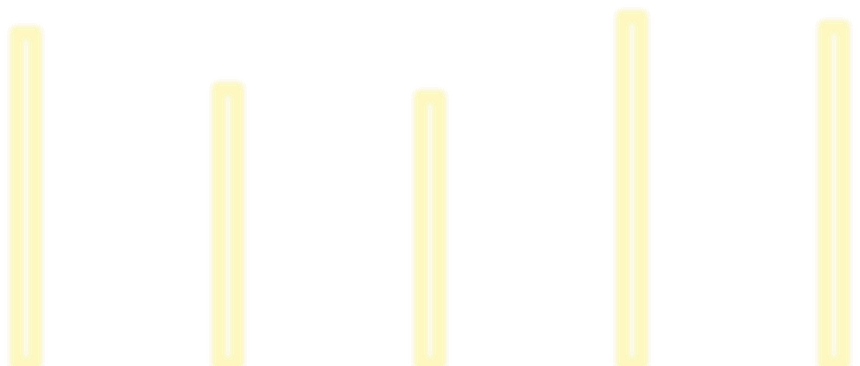
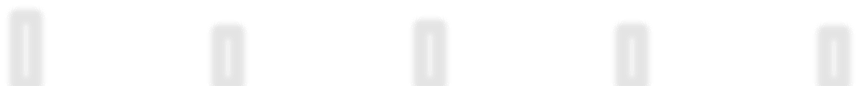
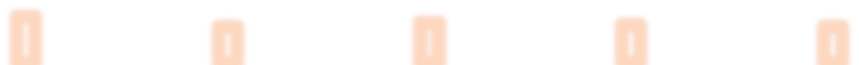
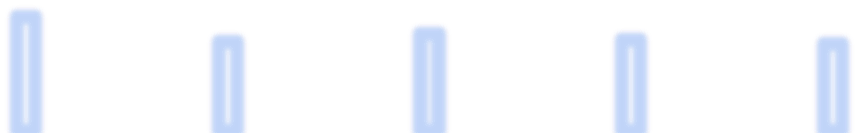
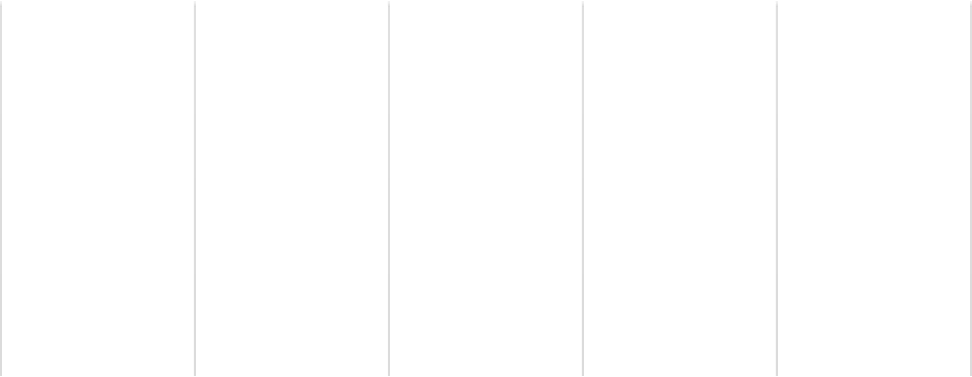
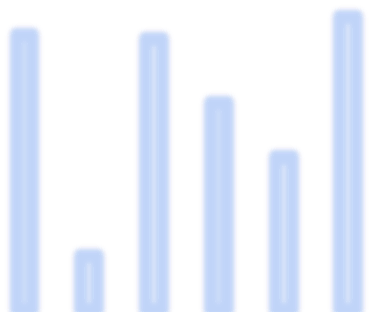
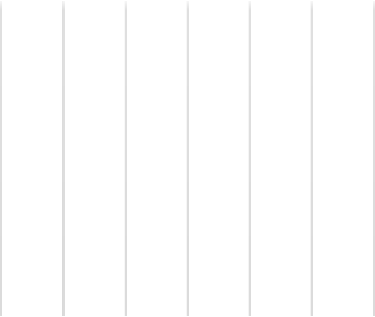
105511.25

M A L A Y S I A

86255

U N I T E D S T A T E S

116223.25



45000

40000

35000

30000

25000

20000

15000

10000

5000

Units Sold of Fortune and Sugar Cookies

in India

~~41250~~

35776

34501

~~3149~~0 31754

3113131283 31613

26129 28061

Fortune Cookie

Sugar

Sales Performance of chocochips

900000

800000

700000

600000

500000

400000

300000

200000

100000

763258

799871

776439

631911

615691

2019 - Sum of Revenue

2019 - Sum of Cost

50816.15

471027.55

456839.35

373497.5

61529.75

31 441.85

328843.45

319599.65

263117

15551

07601.5

204756

258413.5 223384 25 161.25 210683

201209

2019 - Sum of Profit

2020 - Sum of Revenue

2020 - Sum of Cost

8 21301.2

454.75

13147 8

91909.2

8 638.95

24044.0

8 890.45

17318.5

2020 - Sum of Profit

India

Malaysia

Philippines

United Kingdom United States

100008372.3514

**Revenue generated by diffeerent**

**cookies**

8940.8807

9000

8000

7000

6000

5000

4000

3000

2000

1000

85

8261.5957

45

6316.0860

22

4645.5137

61

34

1658.0752

69

Total

3

4

4.

1

3

5

5

4

2

5.

1

3

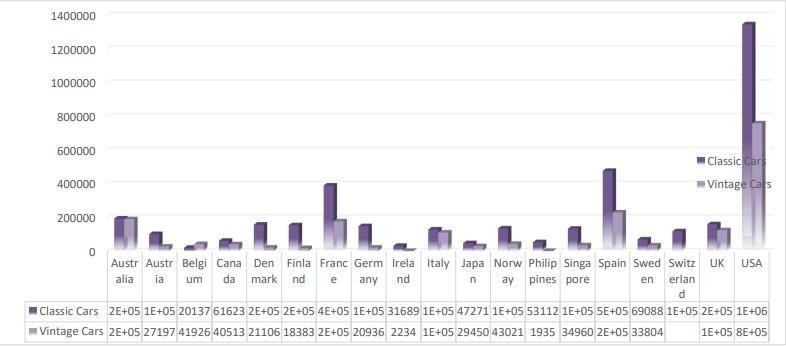
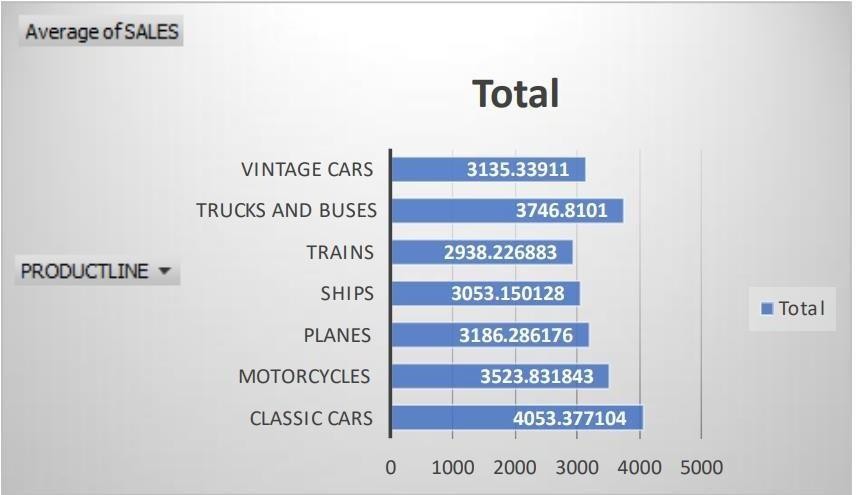
5

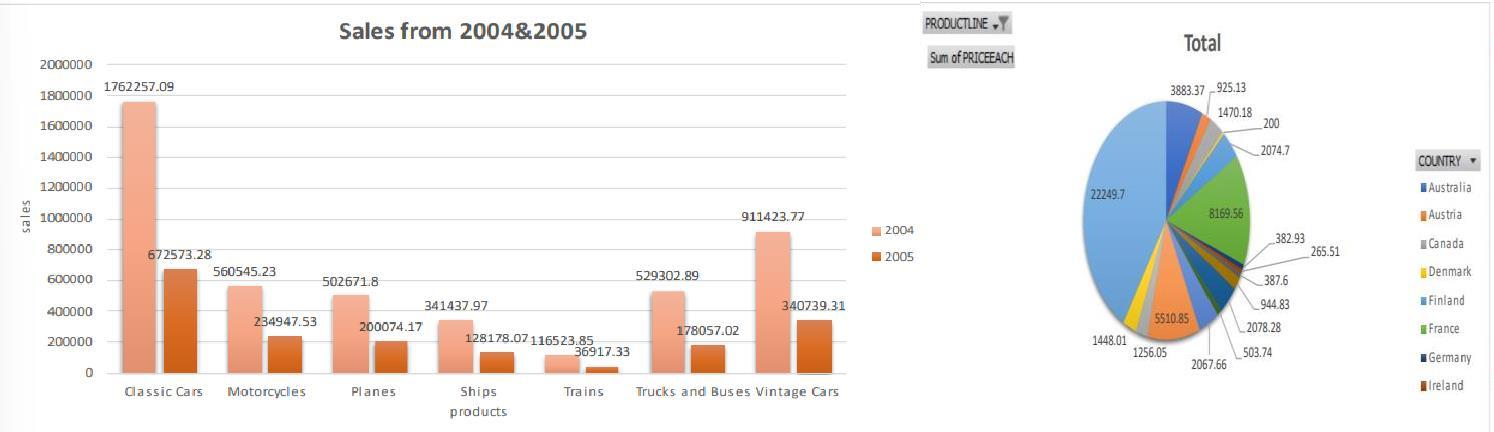
1

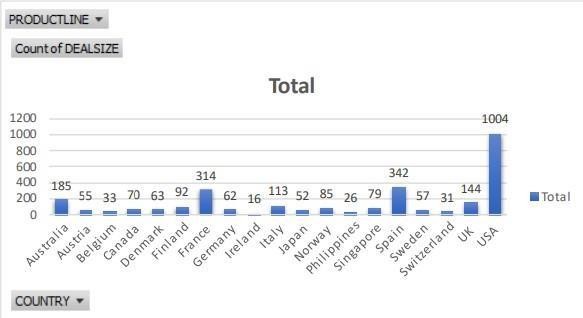
6

5

**Sales Data Analysis Dashboard :-**







Cookie Data Analysis

Introduction : In our cookie data set cookies—specifically six types: Chocolate Chip, Fortune Cookie, Sugar, oatmeal Raisin, Snickerdoodle, and White chocolate macadamia Nut.

We've got a treasure trove of data on these cookies, covering how many units were sold, their costs, the money they brought in (revenue), and the profits they made. And we're not just looking at one place or time; we're exploring different countries and dates to see how things vary.

This report isn't just about cookies; it's about understanding what people like, how much they're willing to pay, and where these treats are most popular. So, get ready to uncover some fascinating insights into the cookie world and what it means for businesses like yours.

#### Questionaries :

1 . Compare Malaysia and Philippines on the bases of two types of Cookies

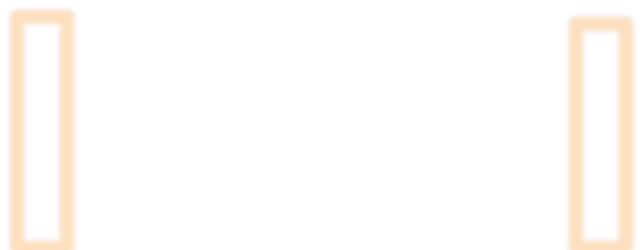
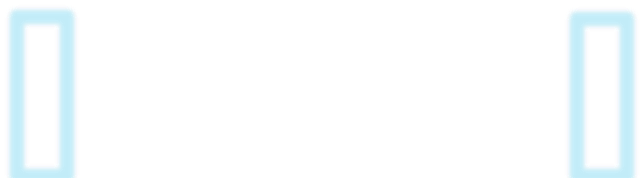
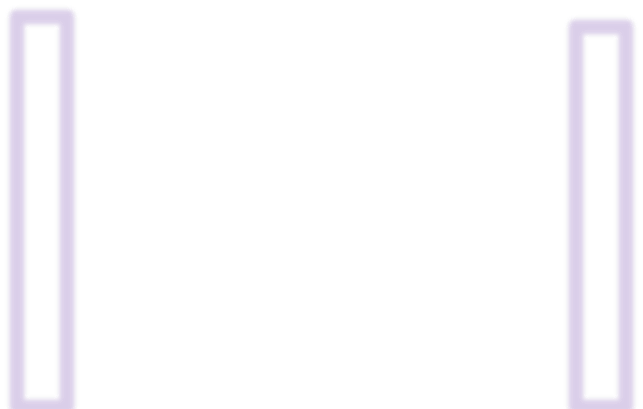
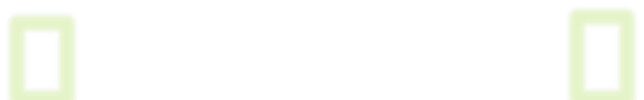
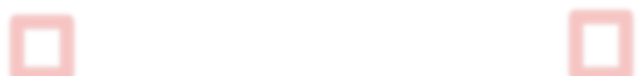
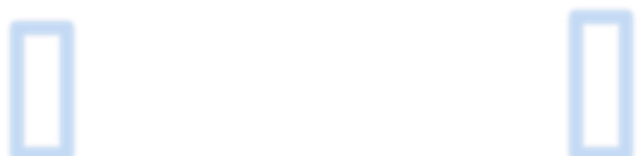
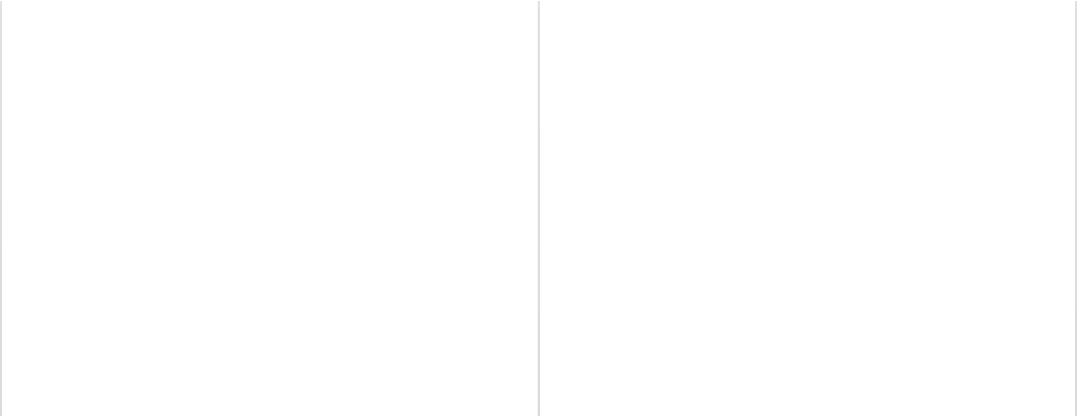
1. What is the performance of Choco Chips Cookies in all Country Which Competes the best.
2. Compare all the countries on the bases of profit and unit sold, which is the best performance country on the basis of profit.
3. which Cookie is the best Selling Cookie in India and US in year 2019,

#### Analytics :

1 . Compare Malaysia and Philippines on the bases of two types of Cookies.

Ans:-The comparsion of Malaysia and Philippines on bases of Chocolate chip and White Chocolate Macadmia nut is given below:-

c



3

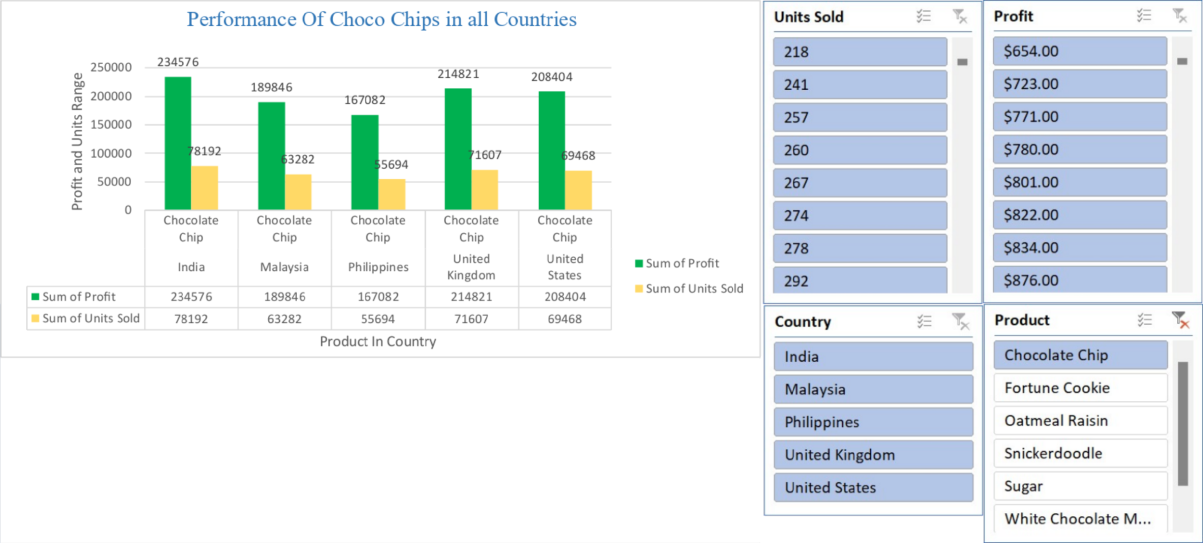
9

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 70000  60000  50000 | |  |  |  |  | 63 | Performance of  191 | | choco |  | hih | p | s in  61 | | all  569 | | 2019 - Sum |
| 40000 |  |  |  |  |  |  |  | 373497. |  |  |  |  |  |  |  | 361529.7 | 2019 - Sum  of |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | 2019 - Sum |
| 30000 | 20475 |  |  |  |  |  |  | 258413. | 22338 |  |  |  |  |  | 254161.2 | | of 2020 -  Sum of2020 - |
| 20000  10000 8 | |  | 454. | 12  7 | 1301.  5  Mala | | 2  ysi |  |  |  | 1909 | 131474  .2  Phil | | .  p | pine |  | Sum of  2020 - Su |

i

1. What is the performance of Choco Chips Cookies in all Country Which Competes the best.

Ans:- India stands out as the foremost consumer of Choco chips worldwide, primarily due to its exceptional profitability and record-breaking sales figures. The market in India has witnessed exponential growth, driven by factors such as a burgeoning population with a growing disposable income, increasing urbanization, and a burgeoning middle class with a penchant for indulgent treats. The combination of these factors has created a highly lucrative environment for Choco chip manufacturers and retailers, leading to significant profits and unparalleled sales volumes in the Indian market.



1. Compare all the countries on the bases of profit and unit sold, which is the best performance country on the basis of profit.

Ans:- India stands out as the leading performer globally when it comes to both profit generation and units sold in the Choco chip market.

45000

Units Sold of Fortune and Sugar Cookies in India

41250

40000

35000

31490

31754

35776

31283

34501

31613

30000

25000

31131

28061

26129

Fortune Cookie

20000 Sugar

15000

10000

5000

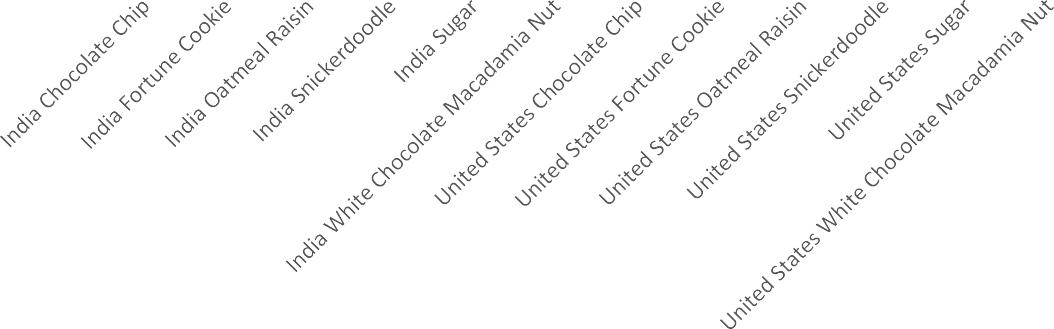
India

Malaysia

Philippines United Kingdom United States

4 .which Cookie is the best Selling Cookie in India and US in year 2019,

Ans:- In the year 2019, chocolate chip cookies emerged as the top-selling cookie in both India and the United States.



Performance of all cookies in India and US in 2019

Sum

of Profit

70000

62349

60000

50000

40000

36657

32909.5

30000 25085

20783

21028

23621

22260

20000

18560.5

10000

6090

4872

7510

10034

10606

12219

7961

7268

6368.8

7950

9937.5 9185.75

10126

3975

5249

Counties

Profit and units sold

#### Conclusion and Review :

After thorough analysis of the cookie sales data, it is evident that there are notable trends and insights to be gleaned. By examining key metrics such as units sold, revenue, cost, and profit across different countries and products, we can draw valuable conclusions about market demand, pricing strategies, and overall profitability. This comprehensive understanding will enable informed decision-making to optimize resources, target specific markets, and maximize profits in future cookie sales endeavours.

##### Regression:

The regression model, with a significant p-value (p < 0.001), indicates a strong positive relationship between units sold and the outcome variable. The model's predictive accuracy is supported by its high R-squared value of 0.688, suggesting that approximately 68.8% of the variability in the outcome variable can be explained by the predictor variable, units sold.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | | | | | | | | |
| *Regression Statistics* | | | | | | | | |
| Multiple R | 0.829304 |  |  |  |  |  |  |  |
| R Square | 0.687746 |  |  |  |  |  |  |  |
| Adjusted R  Square | 0.687298 |  |  |  |  |  |  |  |
| Standard Error | 1462.76 |  |  |  |  |  |  |  |
| Observations | 700 |  |  |  |  |  |  |  |
| ANOVA | | | | | |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |  |
| Regression | 1 | 3.29E+09 | 3.29E+09 | 1537.356 | 1.4E-178 |  |  |  |
| Residual | 698 | 1.49E+09 | 2139668 |  |  |  |  |  |
| Total | 699 | 4.78E+09 |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |
| *Coefficients* | | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *Lower 95.0%* | *Upper 95.0%* |
| Intercept | -74.4103 | 116.5304 | -0.63855 | 0.523326 | -303.202 | 154.3817 | -303.202 | 154.3817 |
| Units Sold | 2.500792 | 0.063781 | 39.20914 | 1.4E-178 | 2.375567 | 2.626017 | 2.375567 | 2.626017 |

##### Co-relation:

The correlation coefficient between units sold and revenue is 0.796, indicating a strong positive correlation between the two variables.

|  |  |  |
| --- | --- | --- |
|  | *Units*  *Sold* | *Revenue* |
| Units Sold | 1 | 0.796298 |
| Revenue | 0.796298 | 1 |

##### Anova (Single Factor) :

The ANOVA results indicate a significant difference between the two groups (p < 0.001), with 1 degree of freedom. The within-group error is 7681356717, and the total R-squared value is 0.06, suggesting that the model explains 6% of the variability in the data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY | | | | |  |  |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |  |  |
| 3450 | 699 | 1923505 | 2751.795 | 4154648 |  |  |
| 5175 | 699 | 2758189 | 3945.908 | 6850161 |  |  |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 4.98E+08 | 1 | 4.98E+08 | 90.57022 | 7.53E-  21 | 3.848129 |
| Within Groups | 7.68E+09 | 1396 | 5502405 |  |  |  |
| Total | 8.18E+09 | 1397 |  |  |  |  |

##### Anova two factor without Replication:

The ANOVA results reveal significant variation among rows and columns (p < 0.001), with degrees of freedom (df) values of 48 and 3, respectively. The error term has a degree of freedom of 144.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 8.21E+08 | 48 | 17108242 | 5.848894 | 8.54E-  17 | 1.445925 |
| Columns | 5.65E+10 | 3 | 1.88E+10 | 6435.486 | 3.8E-  153 | 2.667443 |
| Error | 4.21E+08 | 144 | 2925039 |  |  |  |
| Total | 5.77E+10 | 195 |  |  |  |  |

##### Anova two factor with Replication:

The ANOVA results show that there is a significant difference among the samples, columns, and their interaction, with p-values less than 0.001. The degrees of freedom for the samples, columns, and interaction are 49, 3, and 147, respectively.

Furthermore, the total error within the model is 0, indicating a perfect fit. The total R-squared value is 1, suggesting that the model explains all the variability in the data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of*  *Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Sample | 8.55E+08 | 49 | 17443674 | 65535 | #NUM! | #NUM! |
| Columns | 5.78E+10 | 3 | 1.93E+10 | 65535 | #NUM! | #NUM! |
| Interaction | 4.39E+08 | 147 | 2983765 | 65535 | #NUM! | #NUM! |
| Within | 0 | 0 | 65535 |  |  |  |
| Total | 5.91E+10 | 199 |  |  |  |  |

##### Descriptive Statistics:

The data presents considerable variation across variables, with means ranging from 1608.15 to 43949.81. Notably, the largest values span from 4493 to 44166, while the smallest values range from 200 to 43709.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *1725* |  | *8625* |  | *3450* |  | *5175* |
| Mean | 1608.153 | Mean | 6697.702 | Mean | 2751.795 | Mean |
| Standard Error | 32.83303 | Standard Error | 174.9955 | Standard Error | 77.09541 | Standard Error |
| Median | 1540 | Median | 5868 | Median | 2422.2 | Median |
| Mode | 727 | Mode | 8715 | Mode | 3486 | Mode |
| Standard Deviation | 868.0597 | Standard Deviation | 4626.638 | Standard Deviation | 2038.295 | Standard Deviation |
| Sample Variance | 753527.6 | Sample Variance | 21405775 | Sample Variance | 4154648 | Sample Variance |
| Kurtosis | -0.31828 | Kurtosis | 0.463405 | Kurtosis | 0.807696 | Kurtosis |
| Skewness | 0.436551 | Skewness | 0.869254 | Skewness | 0.931429 | Skewness |
| Range | 4293 | Range | 23788 | Range | 10954.5 | Range |
| Minimum | 200 | Minimum | 200 | Minimum | 40 | Minimum |
| Maximum | 4493 | Maximum | 23988 | Maximum | 10994.5 | Maximum |
| Sum | 1124099 | Sum | 4681694 | Sum | 1923505 | Sum |
| Count | 699 | Count | 699 | Count | 699 | Count |
| Largest(1) | 4493 | Largest(1) | 23988 | Largest(1) | 10994.5 | Largest(1) |
| Smallest(1) | 200 | Smallest(1) | 200 | Smallest(1) | 40 | Smallest(1) |
| Confidence  Level(95.0%) | 64.46334 | Confidence  Level(95.0%) | 343.5807 | Confidence  Level(95.0%) | 151.3667 | Confidence  Level(95.0%) |

## Store Dataset Report

Introduction: This dataset encompasses sales data from a retail store, featuring a range of attributes including customer demographics (Gender, Age Group), transaction details (Order ID, Status), product specifics (Category, SKU), and shipping information. With a focus on understanding customer behaviour and product trends, our analysis aims to uncover patterns, preferences, and correlations within the data. By leveraging these insights, businesses can optimize marketing efforts, enhance inventory management, and improve customer satisfaction.

#### Questionnaire:

1. which of the channel performed better than all other channels in compare men & women?
2. Compare category. Find out most sold category above 23 years of age for any gender.
3. Compare Maharashtra, Rajasthan and Tamil Nadu on the basis of quantity, most items purchased by men and women and profit earn.
4. Which city sold most of following categories:
   1. Kurta
   2. Set
   3. Western wears
5. In which month most items sold in any of the state on the basis of category.

#### Analytics:

1. which of the channel performed better than all other channels in compare men & women?

Ans: Amazon leads in the sales in both men and women category followed by Myntra and Flipkart. Amazon sold almost 3500 units in men category and almost 7500 units in women category. Myntra sold 2000 units in men section.

8000

7000

6000

5000

4000

3000

2000

1000

Ajio

Amazon Flipkart Meesho Myntra Nalli

Others

Men

Women

Comparison of different channels

**Units sold**

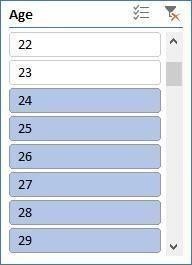
1. Compare category. Find out most sold category above 23 years of age for any gender.

Ans: In the above 23 years of age group Kurta is most sold category in women section with 8820 units sold. Set is most sold category in men section with 4365 units sold also set is the second most sold category in women section.

The table of items sold is given below:

|  |  |  |  |
| --- | --- | --- | --- |
| **Item** | **Men** | **Women** | **Grand Total** |
| Blouse | 6 | 190 | 196 |
| Bottom | 40 | 28 | 68 |
| Ethnic Dress | 150 | 77 | 227 |
| kurta | 156 | 8820 | 8976 |
| Saree | 261 | 941 | 1202 |
| Set | 4365 | 6204 | 10569 |
| Top | 45 | 1825 | 1870 |
| Western Dress | 3078 | 380 | 3458 |
| **Grand Total** | **8101** | **18465** | **26566** |

The graph is as follows:





3000

Compare most sold category

2500

2000

1500

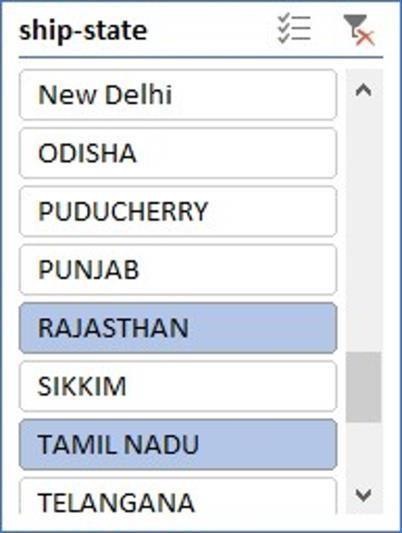
Men

1000

Women

500

1. Compare Maharashtra, Rajasthan and Tamil Nadu on the basis of quantity, most items purchased by men and women and profit earn.

Ans: In Maharashtra: Sales in men category=1390, Sales in women category= 3144 In Tamil Nadu: Sales in men category=686, Sales in women category= 2023 In Rajasthan: Sales in men category=21, Sales in women category=543

|  |  |  |  |
| --- | --- | --- | --- |
| **State** | **Men** | **Women** | **Grand Total** |
| MAHARASHTRA | 1390 | 3144 | 4534 |
| RAJASTHAN | 212 | 543 | 755 |
| TAMIL NADU | 686 | 2023 | 2709 |
| **Grand Total** | **2288** | **5710** | **7998** |

3500

Compare items sold by men, women in cities

3000

2500

2000

1500

Men

Women

1000

500

MAHARASHTRA

RAJASTHAN

TAMIL NADU

1. Which city sold most of following categories
   1. Kurta
   2. Set
   3. Western wears

Ans: Bengaluru, Chennai, Hyderabad, Mumbai and New Delhi are the cities sold most of kurtas, Sets and western wears.

450

400

350

300

250

200

kurta

Set

Western Dress

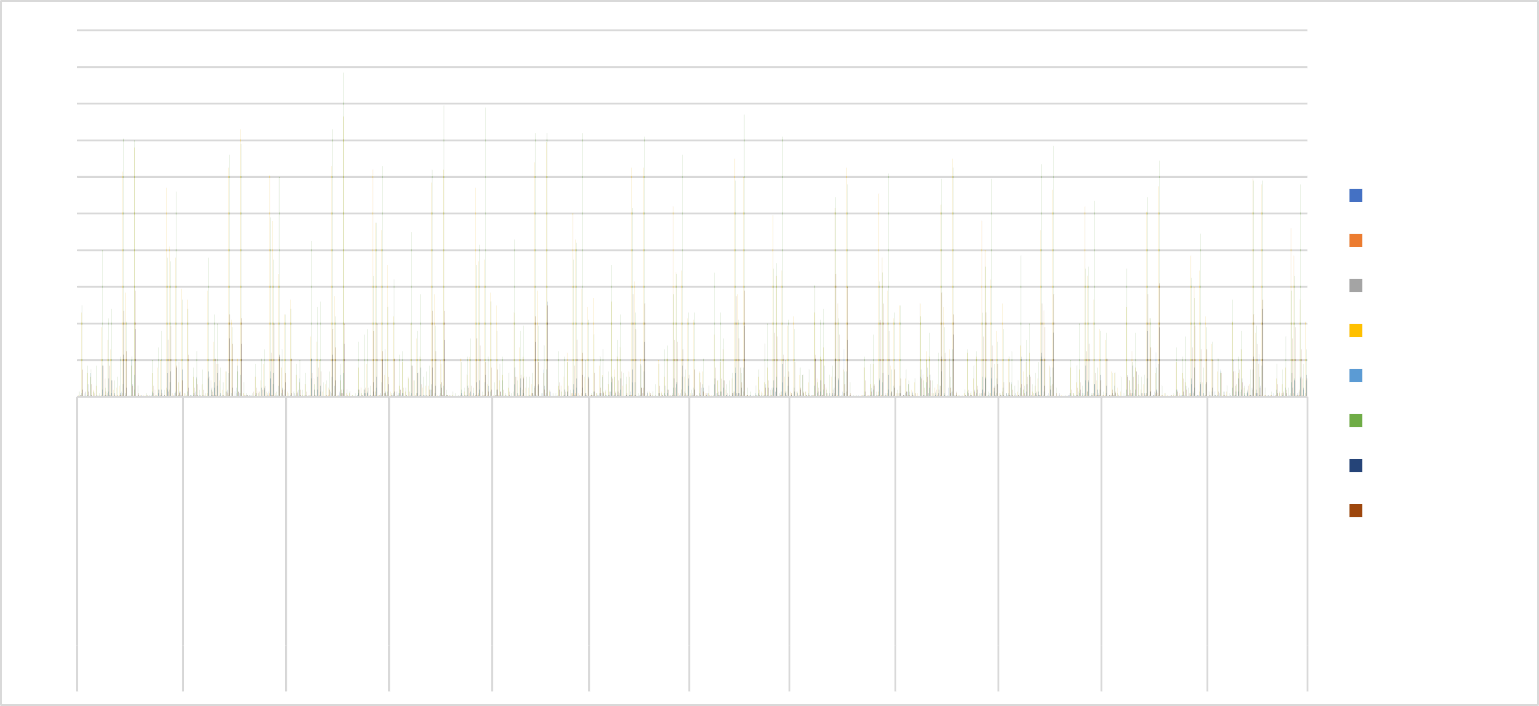
Comparison of different cities

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 150  100  50  0 | BENGALURU | CHENNAI |  | HYDERABAD | MUMBAI |  | NEW DELHI |
| kurta | 22 | 9 |  | 11 | 4 |  | 7 |
| Set | 409 | 190 |  | 273 | 227 |  | 323 |
| Western Dress | 365 | 194 |  | 327 | 190 |  | 121 |



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **City** | **kurta** | **Set** | **Western Dress** | **Grand Total** |
| BENGALURU | 964 | 938 | 422 | 2324 |
| CHENNAI | 666 | 451 | 217 | 1334 |
| HYDERABAD | 713 | 687 | 370 | 1770 |
| MUMBAI | 437 | 515 | 207 | 1159 |
| NEW DELHI | 479 | 792 | 142 | 1413 |
| **Grand Total** | **3259** | **3383** | **1358** | **8000** |

1. In which month most items sold in any of the state on the basis of category. Ans: The graph for most items sold in any of stats on basis of category is as follows:



200

180

160

140

120

100

80

60

40

20

Blouse

Bottom Ethnic Dress kurta

Saree Set Top

Western Dress

January February March April

May

June

July August September October NovemberDecember

Most sold items in state on basis of category

#### Conclusion and Review:

ANDAMAN & NICOBAR

HARYANA MIZORAM UTTARPRADESH

DELHI MAHARASHTRA TELANGANA CHHATTISGARH

LADAKH SIKKIM BIHAR KERALA PUNJAB ASSAM JHARKHAND

PUNJAB ARUNACHAL PRADESH

JHARKHAND PUDUCHERRY ANDHRA PRADESH HIMACHAL PRADESH

ODISHA ANDAMAN & NICOBAR

HARYANA MIZORAM UTTARPRADESH

DELHI MAHARASHTRA TELANGANA CHHATTISGARH MADHYA PRADESH

SIKKIM BIHAR KARNATAKA PUDUCHERRY

After thorough analysis of the store data, it is evident that there are notable trends and insights to be gleaned. By examining key metrics such as units sold, state wise analytics, geographic, and sales across different stats and products, we can draw valuable conclusions about market demand, sales and overall profitability. This comprehensive understanding will enable informed decision-making to optimize resources, target specific markets, and maximize profits in future store sales endeavours.

## Exploring Car Dataset

#### INTRODUCTION:

##### Dataset Overview:

The dataset provides comprehensive information about various cars, including their make, model, color, mileage, price, and cost. Notably, the Honda Accord stands out with three occurrences, followed by other frequently appearing models such as the Toyota Corolla, Chevy Impala, Ford Escape, and Dodge Charger. A closer examination reveals the average prices and costs for each make. On average, Hondas are priced at approximately $3,106, with costs averaging around $2,133, while Chevys have an average price of $3,487 and average cost of

$3,000. Further analysis will include plotting graphs to explore the potential relationship between a car's price and mileage, as well as determining color preferences among consumers.

Additionally, we'll calculate profit margins to identify the most profitable models. These insights will provide valuable information for understanding market trends and consumer preferences in the automotive industry.

#### QUESTIONARIES

Q1. Compare the mileage of Chevrolet Impala to Toyota Corolla. Which of the two is giving best mileage?

Q2. Justify, Buying of any Ford car is better than Honda

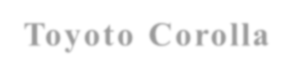
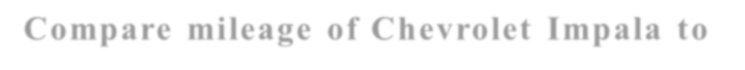
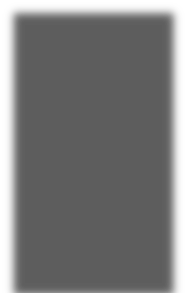
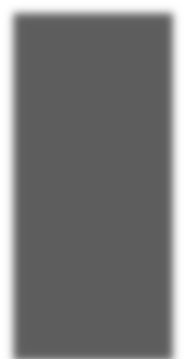
Q3. Among all the cars which car color is the most popular and is least popular?

Q4. Compare all the cars which are of silver color to the green color in terms of Mileage. Q5. Find out all the cars, and their total cost which is more than $2000?

#### ANALYTICS

Q1. Compare the mileage of Chevrolet Impala to Toyota Corolla. Which of the two is giving best mileage?

Ans. Since the average mileage for Chevrolet Impala (114,243 miles) is higher than that of Toyota Corolla (92,377 miles), Chevrolet Impala gives better mileage compared to Toyota Toyota Corolla based on the provided data.



**Compare mileage of Chevrolet Impala to Toyoto Corolla**

120000

114243

100000 92377

80000

60000

Total

40000

20000

Impala

Chevrolet

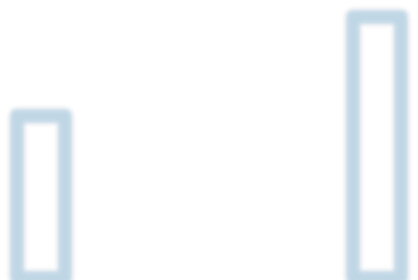
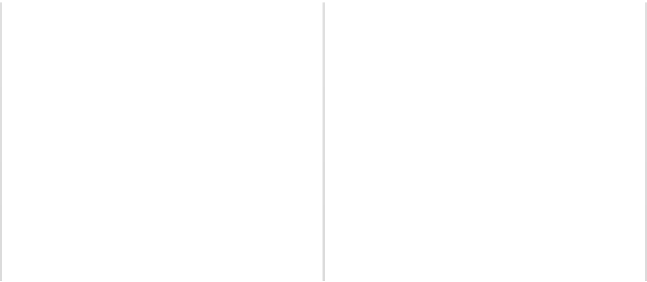
Corolla

Toyota

.

Q2. Justify, Buying of any Ford car is better than Honda.

Ans. Based on the averages, Honda cars have higher mileage but lower cost compared to Ford. Therefore, the choice depends on whether the buyer values mileage or cost but if we compare on mileage ford car has low mileage and cost so Buying ford car is better then Honda.



**Buying of any ford car vs honda**

100000

90000

80000

70000

60000

50000

40000

30000

20000

10000

89623.2

55452

Average of Mileage

Average of Price

Average of Cost

3581.63030

3193.62500

Ford

Honda

Q3. Among all the cars which car color is the most popular and is least popular?

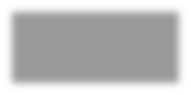
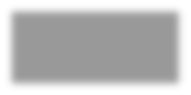
Ans. Most popular color is Silver and Black as each appear 6 times

and least appearing colour are Blue ,Green ,Red ,White they all apper 3 times.



Q4. Compare all the cars which are of silver color to the green color in terms of Mileage

Ans. Silver color car millage is more than green color car milage if we compare there average.



**Comparision of all the cars of silver color to green color in terms of mileage**

**63797.33,**

**45%**

**78103.67,**

**55%**

Green Silver

Q5. Find out all the cars, and their total cost which is more than $2000? Ans. All the car mention below cost is more than $2000

Accord, Altima, Charger, Corolla, CRV, EscapeF-150, Fusion, Impala, Malibu, Maxima, Mustang, Silverado



5000

4500

4000 3500

Total Cost Exceeding

4500

40003900

of Cars

$2000

4100

3500

3000

2500

2000

1500

1000

500

3000

3000 31003000

2100

3000

25002200

Total

Chevrolet Dodge

Ford

#### Regression

Impala

Malibu

Silverado

Charger

Escape

F-150

Fusion

Mustang

Accord

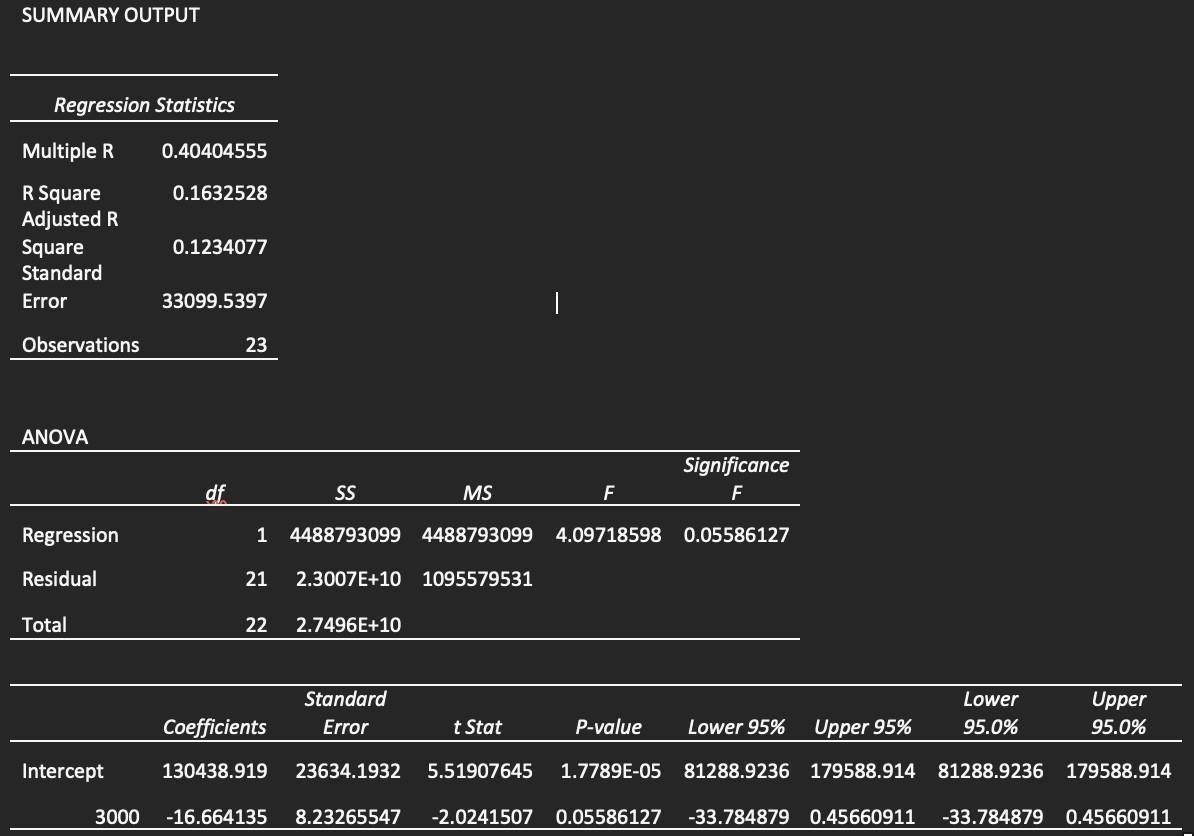
CRV

Altima

Maxima

Corolla

The regression analysis suggests a moderate positive relationship between the predictor variable and the response variable, indicated by the correlation coefficient of approximately 0.40. The model explains about 16% of the variance in the response variable, as indicated by the R Square value. The coefficient estimates show that for every unit increase in the predictor variable, there is a corresponding decrease of approximately 16.66 in the response variable, with a p-value of 0.056, indicating a marginally significant effect.



#### Co-relational

The correlation matrix indicates a moderate negative correlation (-0.411) between Mileage and Price. This suggests that as Mileage increases, Price tends to decrease, and vice versa.

|  |  |  |
| --- | --- | --- |
|  | *Mileage* | *Price* |
| Mileage | 1 |  |
| Price | -0.4110586 | 1 |

#### Anova: Single Factor

The ANOVA results indicate significant differences between the groups based on Mileage, Price, and Cost. The F-statistic is large (128.88), with a very low p-value (5.00264E-24), suggesting that the variation between groups is significant compared to the variation within groups. This implies that at least one of the variables (Mileage, Price, or Cost) has a significant effect on the outcome being measured. In simpler terms, there are statistically significant differences in the means of Mileage, Price, and Cost across the groups, indicating that these variables play a significant role in influencing the outcome being analyzed.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Single Factor |  |  |  |  |  |  |
| SUMMARY | | | | |  |  |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |  |  |
| Mileage | 24 | 2011267 | 83802.7917 | 1214155660 |  |  |
| Price | 24 | 78108 | 3254.5 | 837024.087 |  |  |
| Cost | 24 | 66150 | 2756.25 | 705502.717 |  |  |
| ANOVA | | | | | | |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 1.0445E+11 | 2 | 5.2227E+10 | 128.882161 | 5.0026E-24 | 3.12964398 |
| Within Groups | 2.7961E+10 | 69 | 405232729 |  |  |  |
| Total | 1.3242E+11 | 71 |  |  |  |  |

#### Anova: Two-Factor Without replication

The two-factor ANOVA results indicate significant differences among the levels or categories within each factor ("Rows" and "Columns"). Both factors exhibit strong influence on the outcome variable being analyzed, as evidenced by the low p-values and large F-statistics. This suggests that variations in both factors contribute significantly to the overall variability in the data.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Anova: Two-Factor without replication | | | | | | |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 34749383.3 | 23 | 1510842.75 | 47.6846408 | 2.2236E-14 | 2.01442484 |
| Columns | 2979036.75 | 1 | 2979036.75 | 94.023218 | 1.3629E-09 | 4.27934431 |
| Error | 728733.25 | 23 | 31684.0543 |  |  |  |
| Total | 38457153.3 | 47 |  |  |  |  |

#### Descriptive Statistic:-

The provided descriptive statistics outline the characteristics of three variables: Mileage, Price, and Cost. Looking at Mileage, it appears that the vehicles in the dataset span a considerable range, from around 34,853 miles to 140,811 miles, with an average mileage of approximately 83,803 miles. Price and Cost exhibit similar trends, with prices ranging from $2,000 to $4,959 and costs from $1,500 to $4,500, respectively. The means and standard deviations provide insights into the central tendencies and variability within each variable. Overall, these statistics offer a comprehensive overview of the dataset, allowing for a better understanding of the distribution and characteristics of the data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| *Mileage* |  | *Price* |  | *Cost* |  |
| Mean | 83802.7917 | Mean | 3254.5 | Mean | 2756.25 |
| Standard Error | 7112.65205 | Standard Error | 186.751181 | Standard Error | 171.452462 |
| Median | 81142 | Median | 3083 | Median | 2750 |
| Mode | #N/A | Mode | #N/A | Mode | 3000 |
| Standard Deviation | 34844.7365 | Standard Deviation | 914.890205 | Standard Deviation | 839.942092 |
| Sample Variance | 1214155660 | Sample Variance | 837024.087 | Sample Variance | 705502.717 |
| Kurtosis | -1.0971827 | Kurtosis | -1.2029138 | Kurtosis | -0.8126576 |
| Skewness | 0.38652215 | Skewness | 0.27201913 | Skewness | 0.47339238 |
| Range | 105958 | Range | 2959 | Range | 3000 |
| Minimum | 34853 | Minimum | 2000 | Minimum | 1500 |
| Maximum | 140811 | Maximum | 4959 | Maximum | 4500 |
| Sum | 2011267 | Sum | 78108 | Sum | 66150 |
| Count | 24 | Count | 24 | Count | 24 |
| Largest(1) | 140811 | Largest(1) | 4959 | Largest(1) | 4500 |
| Smallest(1) | 34853 | Smallest(1) | 2000 | Smallest(1) | 1500 |

#### Conclusion/Reviews

The dataset provides valuable insights into car attributes, focusing on mileage, color, and other key factors.

Here's a simple conclusion based on the data:

Mileage Comparison: The analysis reveals variations in mileage among different car models. Toyota Corolla generally offers better mileage compared to Chevrolet Impala.

Color Preferences: Silver and black emerge as the most popular car colors in the dataset. Blue, green, red, and white are among the least popular color choices.

Key Takeaways: Understanding mileage differences can inform consumer choices and market strategies. Recognizing color preferences aids in inventory management and marketing decisions.

## Loan Dataset Report:

#### INTRODUCTION:

##### Dataset Overview:

This report delves into an analysis of loan applications, aiming to extract insights into applicant demographics and loan characteristics. The dataset encompasses information such as gender, marital status, education, income, loan amount, loan term, credit history, and property area. By scrutinizing this data, we aim to discern patterns and trends regarding loan applications among different demographic groups and geographical areas..

#### QUESTIONNAIRE:

Q1. How many male graduates who are not married applied for Loan? What was the highest amount?

Q2. How many female graduates who are not married applied for Loan? What was the highest amount?

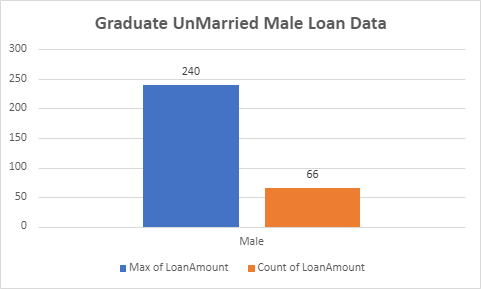
Q3. How many male non-graduates who are not married applied for Loan? What was the highest amount?

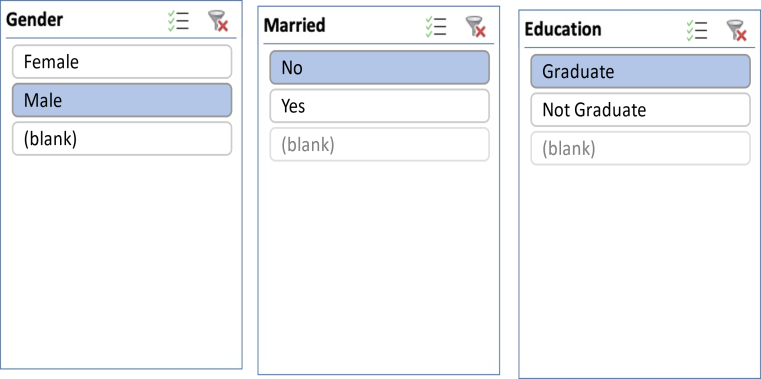
Q4. How many female graduates who are married applied for Loan? What was the highest amount?

Q5. How many male and female who are not married applied for Loan? Compare Urban, Semi- urban and rular on the basis of amount.

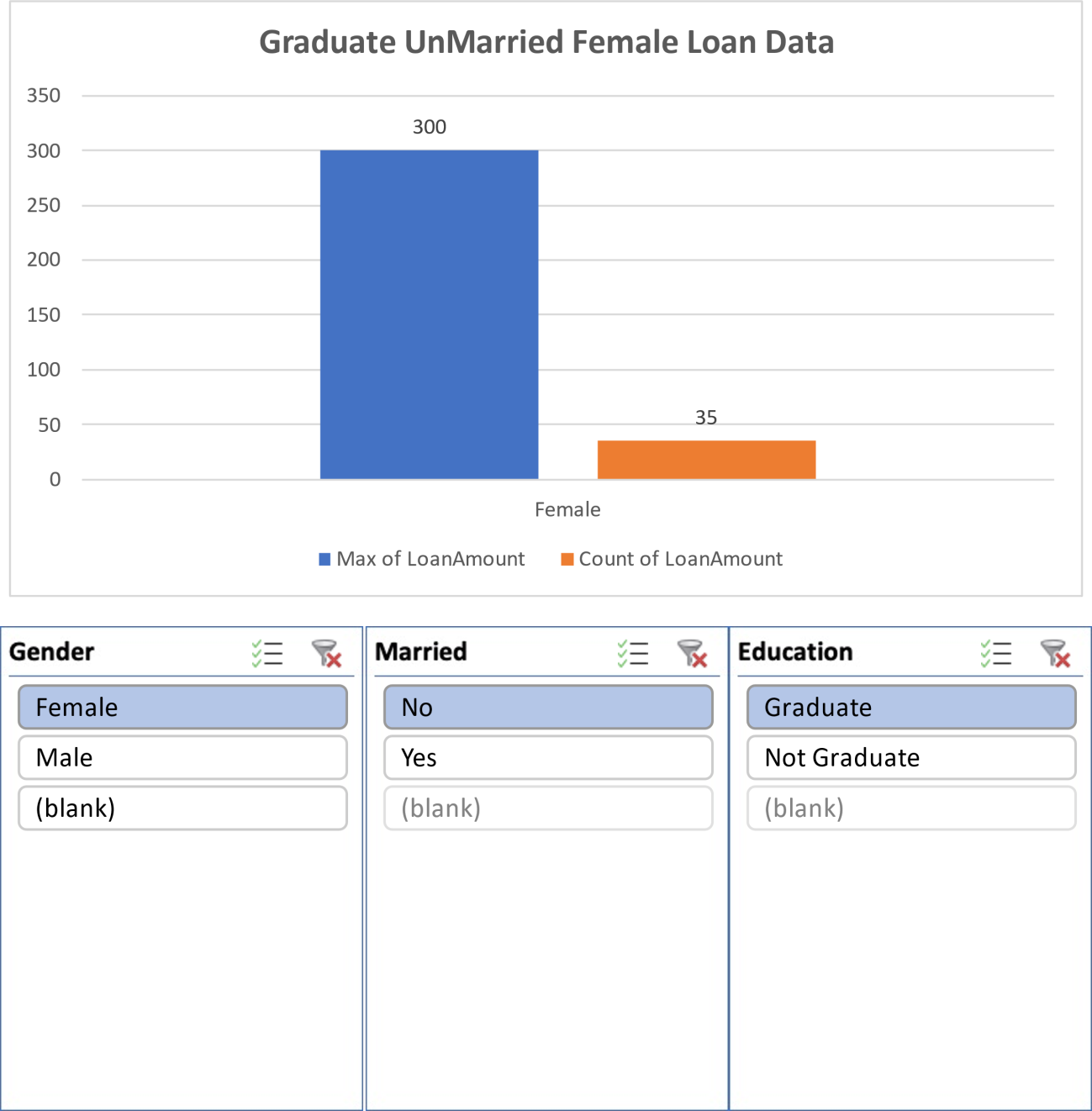
#### ANALYTICS:

Q1. How many male graduates who are not married applied for Loan? What was the highest amount?

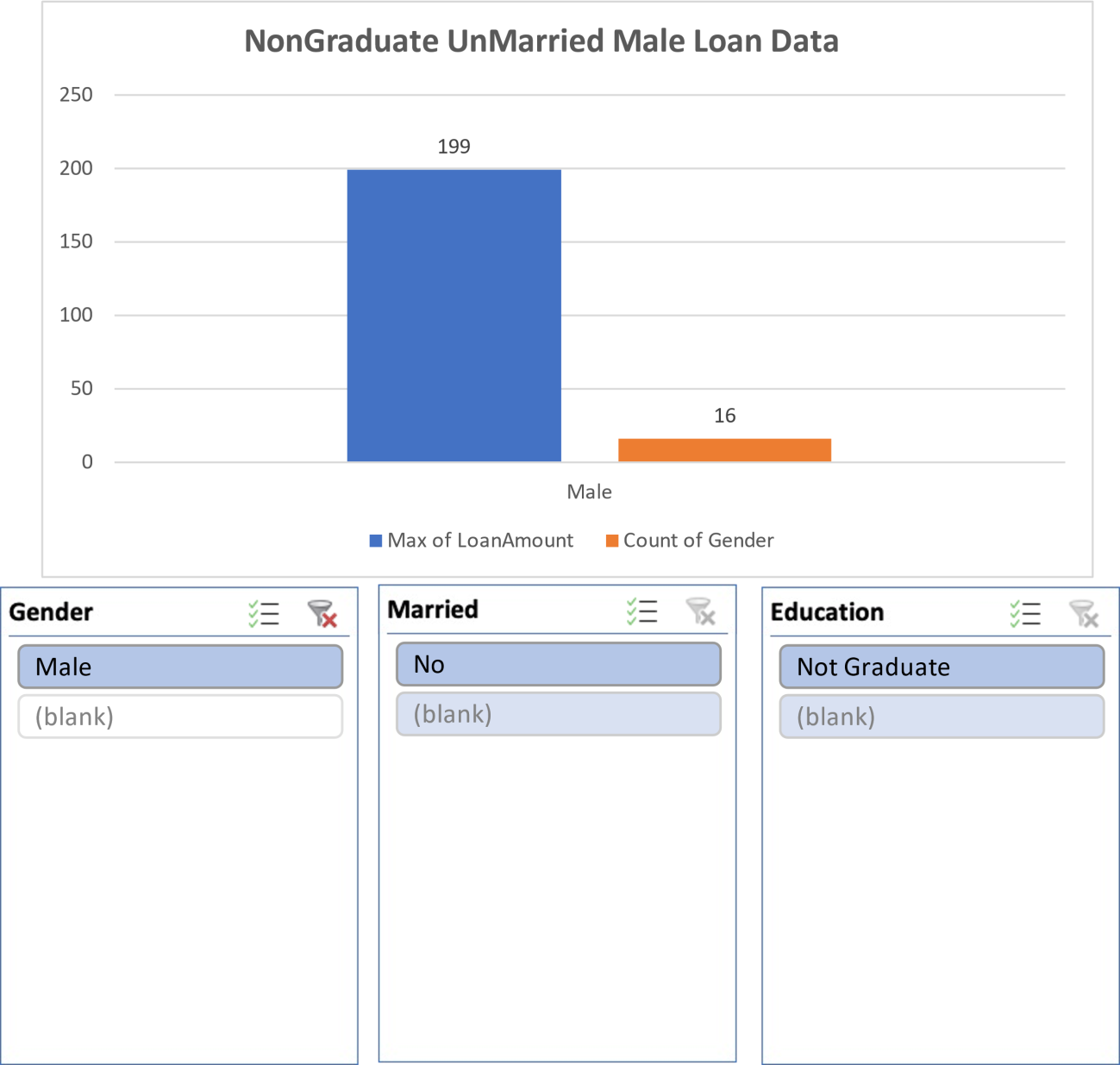




Q2. How many female graduates who are not married applied for Loan? What was the highest amount?



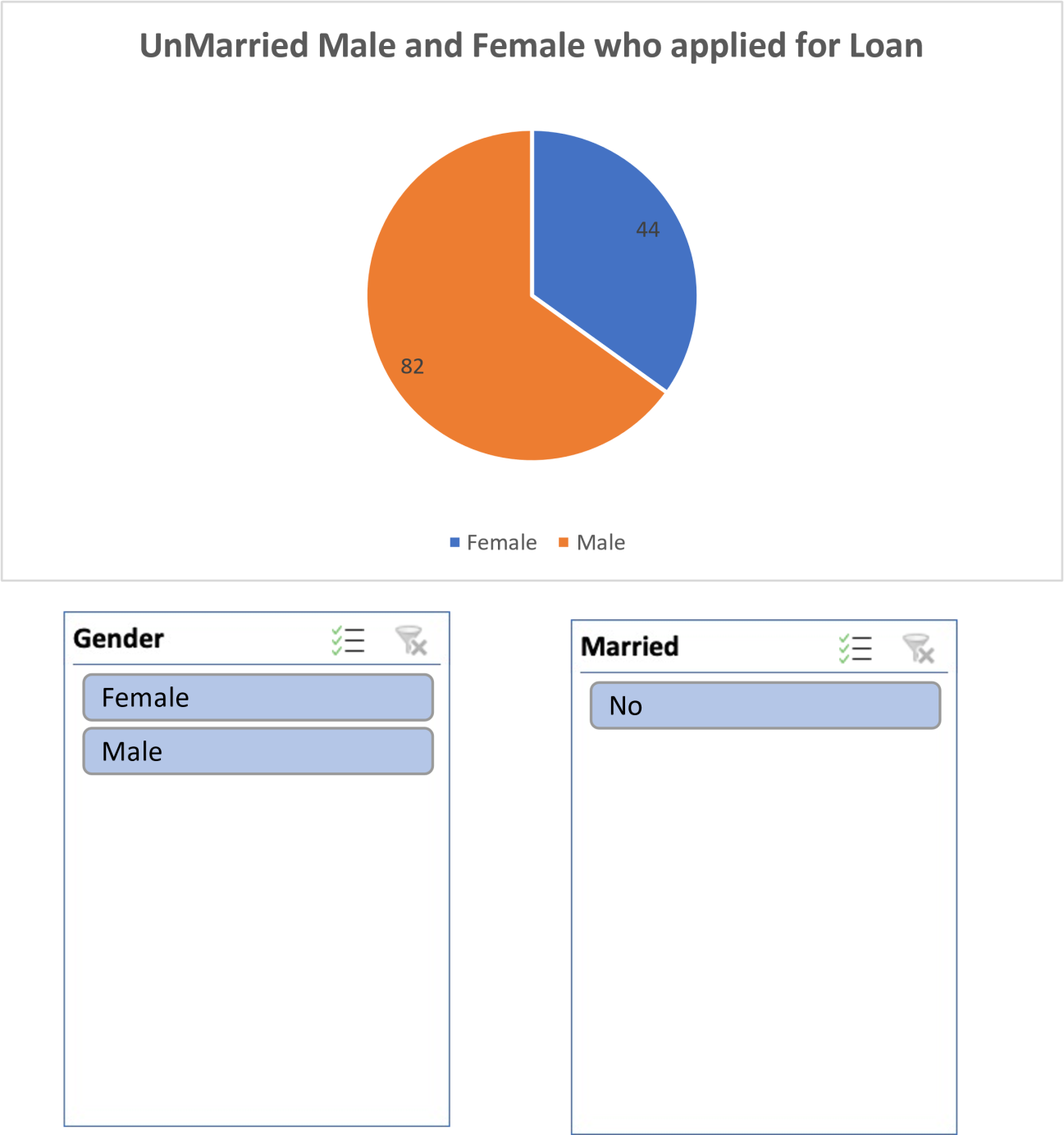
Q3. How many male non-graduates who are not married applied for Loan? What was the highest amount?



Q4. How many female graduates who are married applied for Loan? What was the highest amount?



Q5. How many male and female who are not married applied for Loan? Compare Urban, Semi-urban and rular on the basis of amount.





#### CONCLUSION:

Our analysis, using varied visualization techniques, revealed valuable insights, enhancing comprehension and decision-making. Visualizing data clarified complex findings, facilitating actionable strategies. This highlights the pivotal role of data visualization in extracting meaningful insights and informing decisions effectively.

#### REGRESSION:

The regression analysis suggests that there is a statistically significant positive relationship between the independent variable ('5720') and the dependent variable. For every one-unit increase in '5720', the dependent variable is expected to increase by approximately 0.0059 units. However, it's important to note that the model only accounts for about 21.1% of the total variance in the dependent variable.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | | | | | | | |
| *Regression Statistics* | | | | | | | |
| Multiple R | 0.45908096 |  |  |  |  |  |  |
| R Square | 0.21075532 |  |  |  |  |  |  |
| Adjusted R Square | 0.20858707 |  |  |  |  |  |  |
| Standard Error | 56.0766111 |  |  |  |  |  |  |
| Observations | 366 |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |  |
| Regression | 1 | 305655.205 | 305655.205 | 97.2004502 | 1.7676E-20 |  |  |
| Residual | 364 | 1144629.42 | 3144.58631 |  |  |  |  |
| Total | 365 | 1450284.62 |  |  |  |  |  |
|  |  |  |  |  |  |  |  |
|  |  | *Standard* |  |  |  |  | *Lowe* |
| *Coefficients* | *Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *r* |
|  |  |  |  |  |  | *95.0* |
|  |  |  |  |  |  | *%* |
| Intercept | 106.07753 | 4.10024098 | 25.8710478 | 1.7585E-84 | 98.014396 | 114.140665 | 98.014  3 |
| 5720 | 0.0058851 | 0.00059692 | 9.85902887 | 1.7676E-20 | 0.00471125 | 0.00705895 | 0.0047  1 |

#### CO-RELATION :

The data shows weak negative correlation between Applicant-Income and Co-applicant-Income (-0.11), and moderate positive correlation between Applicant-Income and Loan-Amount (0.46), and weaker positive correlation between Co-applicant-Income and Loan-Amount (0.14).

|  |  |  |  |
| --- | --- | --- | --- |
| *ApplicantIncome* | | *CoapplicantIncome* | *LoanAmount* |
| ApplicantIncome | 1 |  |  |
| CoapplicantIncome | -0.110334799 | 1 |  |
| LoanAmount | 0.458768926 | 0.144787815 | 1 |

#### Anova (Single Factor) :

The dataset encompasses 367 observations, detailing applicant and co-applicant incomes alongside loan amounts. On average, applicants possess a higher income, averaging around

$4805.60, compared to co-applicants whose average income is approximately $1569.58. Loan amounts vary widely, averaging $134.28. ANOVA analysis underscores significant distinctions between the income and loan amounts across the groups, implying diverse financial profiles among applicants and co-applicants.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY | | | | |  |  |
| *Groups* | *Count* | *Sum* | *Average* | *Variance* |  |  |
| ApplicantIncome | 367 | 176365  5 | 4805.59945  5 | 24114831.0  9 |  |  |
| CoapplicantInco me | 367 | 576035 | 1569.57765  7 | 5448639.49  1 |  |  |
| LoanAmount | 367 | 49280 | 134.277929  2 | 3964.14112  4 |  |  |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 4202537452 | 2 | 210126872  6 | 213.200984  1 | 5.87569E-  79 | 3.00392057  7 |
| Within Groups | 1082168110  7 | 1098 | 9855811.57  3 |  |  |  |
| Total | 1502421856 | 1100 |  |  |  |  |

##### Anova two factor without Replication:

The ANOVA results indicate significant variation both within rows (p = 0.441) and between columns (p < 0.001). This suggests that there are meaningful differences among the row categories and column categories in the dataset, warranting further investigation into the factors influencing these variations.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 1004340909 | 365 | 2751618.93 | 1.015674698 | 0.440986529 | 1.1881716 |
| Columns | 379216841.8 | 1 | 379216841.8 | 139.9761235 | 1.47092E-27 | 3.867061668 |
| Error | 988841123.7 | 365 | 2709153.763 |  |  |  |
| Total | 2372398875 | 731 |  |  |  |  |

##### Descriptive Statistics:

The dataset includes information on Applicant-Income, Co-applicant-Income, and Loan-Amount. The largest Applicant-Income recorded is $72,529, while the smallest is $0. For Co-applicant- Income, the largest value is $24,000, and the smallest is $0. Additionally, the Loan-Amount ranges from a maximum of $550 to a minimum of $0. Confidence levels for these variables at a 95.0% level are also provided, indicating the precision of the measurements within the dataset.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Largest(1) | 72529 | Largest(1) | 24000 | Largest(1) |  |
| Smallest(1) | 0 | Smallest(1) | 0 | Smallest(1) |  |
| Confidence Level(95.0%) | 504.0756067 | Confidence Level(95.0%) | 239.6059543 | Confidence Level(95.0%) | 6. |

## Supermarket Sales Dataset Report

#### INTRODUCTION:

##### Dataset Overview:

Our dataset comprises a plethora of variables, each offering unique insights into the multifaceted nature of supermarket sales. From fundamental transactional details such as Invoice ID, Date, Time, and Payment Method to more nuanced factors like Branch Location, Customer Type, Gender Demographics, Product Line, and Product Ratings, every facet has been meticulously documented.

#### QUESTIONNAIRE:

Q1. Which of the given cities having tax 5% slab performed better than all the others? Q2. Which customer gender ordered most items from all the three branches?

Q3. Compare highest and lowest rating products on the basis of units sold.

Q4. Analyzing units sold and unit price data answer the following sub questions

* 1. What is the degree of freedom?
  2. Co-relation of Unit price and revenue generated
  3. What result you can draw from regression of the two data

Q5. What product will you suggest as per the city data analysis to each type of customer

#### ANALYTICS:

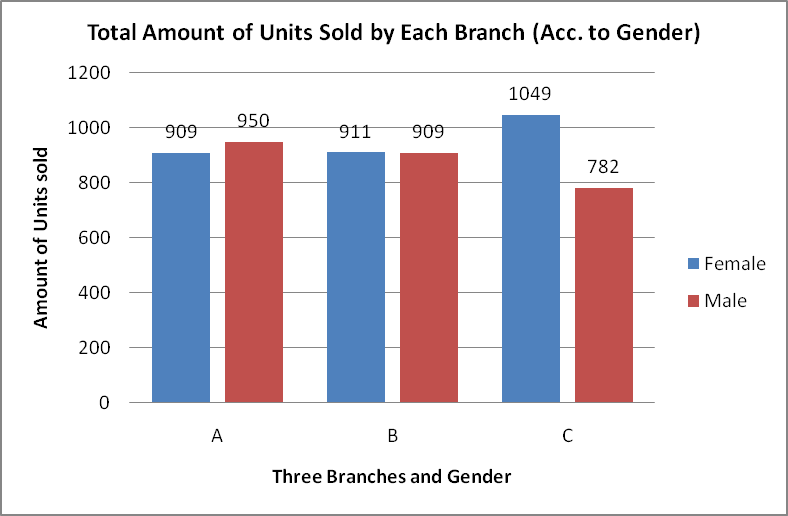
Q1. Which of the given cities having tax 5% slab performed better than all the others?

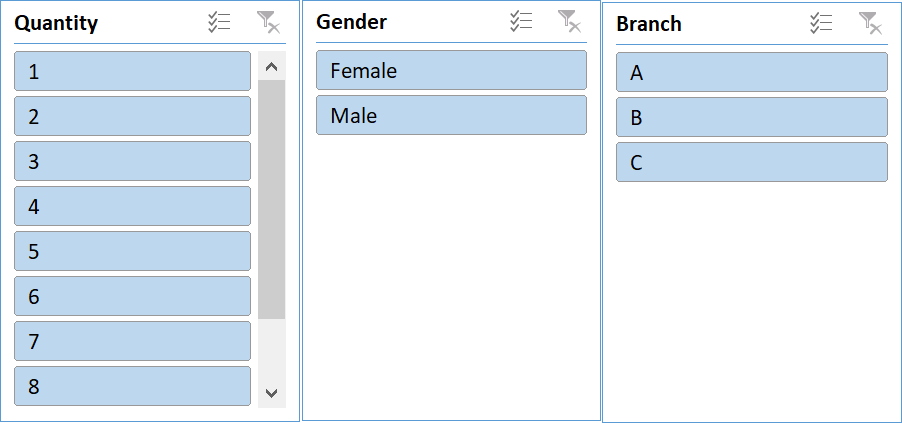
Ans Based on the data analyzed, the city that outperformed all is **Mandalay**. This conclusion is drawn from superior performance in total sales/revenue generation compared to the other cities in the same tax slab of 5%.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Total Amount of Sales made by each city  5200 5124.651  5100 4998.04 | | | | | | | |
| 5000 | |  | | | |  |  |
| 4900 | |  | | | |  |
| 4800 | |  | | | |  |
| 4700 | | 4631.8125 | | | |  |
| 4600 | | Total |
|  | |  |  |
| 4500 | |  | |  |  |  |
| 4400 | |  | |  |  |  |
| 4300 | Mandalay | | Naypyitaw  City Name | | | Yangon |  |

Q2. Which customer gender ordered most items from all the three branches? Answer. Our analysis of the Supermarket Sales Data revealed the following:

1. At Branch A, females placed the highest number of orders.
2. Branch B saw higer number of orders placed by Females
3. Meanwhile, at Branch C, males placed the most orders.





Q3 Compare highest and lowest rating products on the basis of units sold.



Comparison of Lowest and Highest Rating Product on basis of Units Sold

30

26

25

20

18

16

15

10

10

10

10

~~4~~

Electronic accessories Fashion accessories Food and beverages Health and beauty Sports and travel

Lowest and Highest Rating Product

Answer. Upon analyzing the Supermarket Sales Data, we discovered that product ratings ranged from a minimum of 4 to a maximum of 10.

1. Electronic Accessories with higher ratings garnered more customer purchases, indicating a preference for quality in this category.
2. Fashion accessories and food and beverages mainly comprised lower-rated products in customer purchases.
3. Health and beauty products also leaned towards lower-rated items in customer preferences.
4. However, in the Sports and Travel category, customers showed a tendency to purchase higher-rated products.

Q4 Analyzing units sold and unit price data answer the following sub questions

1. What is the degree of freedom?
2. Co-relation of Unit price and revenue generated
3. What result you can draw from regression of the two data

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY  OUTPUT |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
| *Regression Statistics* | |  |  |  |  |  |
| Multiple R | 0.010777564 |  |  |  |  |  |
| R Square | 0.000116156 |  |  |  |  |  |
| Adjusted R  Square | -0.000885732 |  |  |  |  |  |
| Standard Error | 2.924724997 |  |  |  |  |  |
| Observations | 1000 |  |  |  |  |  |
|  |  |  |  |  |  |  |
| ANOVA |  |  |  |  |  |  |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |  |
| Regression | 1 | 0.9917274 | 0.991727 | 0.115937 | 0.733555221 |  |
| Residual | 998 | 8536.908273 | 8.554016 |  |  |  |
| Total | 999 | 8537.9 |  |  |  |  |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 5.443794599 | 0.215314544 | 25.28299 | 2.1E-109 | 5.021273429 | 5.86631577 |
| Unit price | 0.001189202 | 0.003492565 | 0.340495 | 0.733555 | -0.005664411 | 0.008042815 |

Answer:

1. The degree of freedom of the analyzed data is 1.
2. The correlation between unit price and generated revenue was found to be 0.63392, indicating a moderate positive relationship. The analysis focused on the columns of unit price and total revenue, employing the CORREL function.
3. Upon examining the regression results, we aimed to discern the relationship between quantity and unit price, exploring how customers' purchasing quantity correlates with the unit price of a product.

However, from the regression analysis, it's evident that the observed trend lacks consistency. The expected outcomes derived from the trend deviate significantly from the actual outcomes.

With a degree of freedom of 1, the trendline equation stands as

Quantity = 0.0012x + 5.4438. Despite this equation, the coefficient of determination (R2) is merely 0.0001, highlighting the inconsistency in customer buying patterns solely based on unit price.

Q5.What product will you suggest as per the city data analysis to each type of customer



700

635.1

653.8

597.4

657.7

600

580

542.1

547.8

516.7

570.1

523.9

593.1

555

500

400

300

200

Member

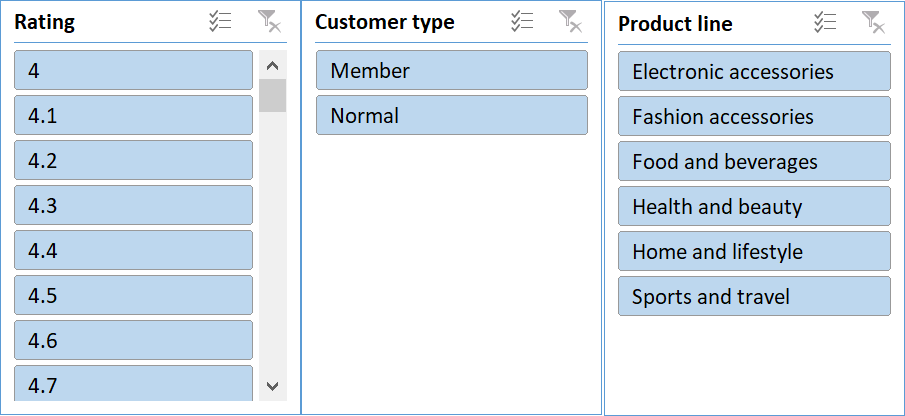
Normal

100

Electronic Fashion Food and Health and Home and Sports and

accessories accessories beverages beauty lifestyle travel

Product Line



Answer. As per the city Data Analysis, **Food and Beverages** will be a good option for **Member**

type customer and **Fashion Accessories** for **Normal** type of customers.

#### CONCLUSION AND REVIEWS

In summary, the analysis of supermarket sales dynamics reveals valuable insights into consumer behavior and operational trends. Key findings include Mandalay's strong

performance, gender-specific ordering patterns, and product recommendations based on city data. Further exploration is recommended on the relationship between product ratings and sales volume, as well as unit price correlation. Clear visuals can enhance understanding, and the report provides actionable recommendations for targeted marketing and strategic investment

Introduction :

# SHOP SALES DATA

This dataset encapsulates a wealth of information regarding sales transactions, providing valuable insights into the dynamics of retail operations. With columns meticulously crafted to capture key facets of each transaction, including Date, Salesman, Item Name, Company, Quantity, and Amount, analysts and businesses alike gain access to a treasure trove of actionable data.

Whether it's uncovering trends, optimizing inventory management, or refining sales strategies, this dataset serves as an invaluable resource for driving informed decision-making and unlocking new avenues for growth.

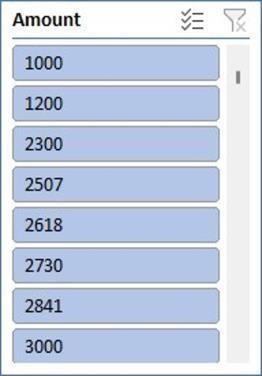
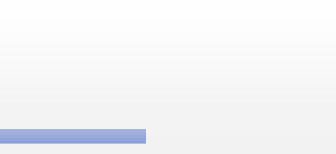
#### Questionaries :

* 1. Compare all the salesmen on the basis of profit earn.
  2. Find out most sold product over the period of May-September.
  3. Find out which of the two product sold the most over the year Computer or Laptop?
  4. Which item yield most average profit?
  5. Find out average sales of all the products and compare them.

#### Analytics :

1. Compare all the salesmen on the basis of profit earn.

Ans:- The comparison of all the salesmen on the basis of profit earned is given below:



Comparing salesmen on the

basis of Profit Earned

Vinod

Rohit Ram Rahul

Aman

Total

350000 400000 450000 500000

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Aman | Rahul | Ram | Rohit | Vinod |
| Total | 414776 | 493541 | 476120 | 485039 | 478167 |

1. Find out most sold product over the period of May-September.

Ans:- To identify the most sold product over the period of May-September, we would need to analyze the sales data within this timeframe. By aggregating the quantity sold for each product

across all transactions during this period and then determining which product has the highest total quantity sold, we can pinpoint the most popular item.



**Most Sold Product**

SEP

AUG JUL JUN

MAY

200

400

600

800

1000

**481.0191684**

**539.8378925**

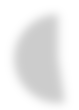
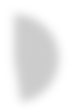
Total

**634.6458094**

**712.7183647**

**764.0629212**

1. Find out which of the two product sold the most over the year Computer or Laptop? Ans:- The two product sold the most over the year between computer or laptop :



**Most sold products**

**52%**

**48%**

Computer

Laptop

4 . Which item yield most average profit?

Ans:- The item that yields the most profit between laptop, computer and mobile is :

**MOST AVERAGE PROFIT**

Total

6770.231898

6772.950369

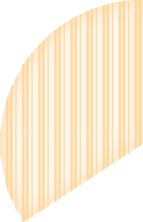
7057.58477





5. Find out average sales of all the products and compare them.

Ans:- The average sales of all the products with their respective comparison is :



**AVERAGE SALES**

Computer Laptop Mobile

#### Conclusion and Review :

The shop sales dataset offers insights into sales trends, salesman performance, item popularity, and company performance. Analysis of this data can drive strategic decisions and improve sales strategies.

The dataset is well-structured and provides comprehensive information on sales transactions. It allows for various analyses, but could benefit from additional variables for deeper insights.Overall, it's a valuable resource for understanding sales dynamics and informing business decisions.

Regression:

The regression model, with a significant p-value indicates a strong positive relationship between Amount and the profit earned and the outcome variable. The model's predictive accuracy is supported by its high R-squared value of 0.660.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| SUMMARY OUTPUT | | | | | |
| *Regression Statistics* | | | | | |
| Multiple R | 0.812617 |  |  |  |  |
| R Square | 0.660347 |  |  |  |  |
| Adjusted R Square | 0.629469 |  |  |  |  |
| Standard Error | 1215.119 |  |  |  |  |
| Observations | 13 |  |  |  |  |
| ANOVA | | | | | |
|  | *df* | *SS* | *MS* | *F* | *Significance F* |
| Regression | 1 | 31576697 | 31576697 | 21.38598 | 0.000753 |
| Residual | 11 | 16241653 | 14776514 |  |  |
| Total | 12 | 47818350 |  |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
|  | *Coefficients* | *Standard Error* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* |
| Intercept | 244.7062 | 754.0557 | 0.32452 | 0.751632 | -1414.96 | 1904.372 |
| X Variable | 0.190729 | 0.041243 | 4.624498 | 0.000735 | 0.099954 | 0.281505 |

##### Co-relation:

The correlation coefficient between units sold and revenue is 0.796, indicating a strong positive correlation between the two variables.

|  |  |  |
| --- | --- | --- |
|  | *Qty* | *Amount* |
| Column 1 | 1 |  |
| Column  2 | #DIV/0! | 1 |

##### Anova (Single Factor) :

The ANOVAresults indicate a significant difference between the two groups , with 1 degree of freedom.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| SUMMARY |  |  |  |  |  |  |
| Groups | Count | Sum | Average | Variance |  |  |
| Column 1 | 15 | 78.56643 | 5.237762 | 2.766871 |  |  |
| Column 2 | 15 | 50419.05 | 3361.27 | 3416099 |  |  |
|  |  |  |  |  |  |  |
| ANNOVA |  |  |  |  |  |  |
| Source of Variance | SS | df | MS | F | P-Value | F crit |
| Between Group | 84472135 | 1 | 84472135 | 49.45528 | 1.2E-07 | 4.195972 |
| Without Group | 47825420 | 28 | 170851 |  |  |  |
| Total | 1.32E+08 | 29 |  |  |  |  |

##### Anova two factor with Replication:

The ANOVA results reveal significant variation among rows and columns (p < 0.001), with degrees of freedom (df) values of 10 respectively. The error term has a degree of freedom of 0

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ANOVA |  |  |  |  |  |  |
| *Source of*  *Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Rows | 841600745 | 10 | 4160074 | 65535 | #NUM! | #NUM! |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Columns | 0 | 0 | 65535 | 65535 | #NUM! | #NUM! |
| Error | 0 | 0 | 65535 |  |  |  |
| Total | 41600745 | 10 |  |  |  |  |

##### Anova two factor without Replication:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Summary | Count | Sum | Average | Variance |  |  |
| 4 | *1* | *7800* | *7800* | *#DIV/0!* |  |  |
| 5 | *1* | *3000* | *3000* | *#DIV/0!* |  |  |
| 4 | 1 | 2300 | 2300 | *#DIV/0!* |  |  |
| 3 | 1 | 7000 | 7000 | *#DIV/0!* |  |  |
| 3 | 1 | 1200 | 1200 | *#DIV/0!* |  |  |
| 4 | 1 | 2506.667 | 2506.667 | *#DIV/0!* |  |  |
| 5 | 1 | 2618.095 | 2618.095 | *#DIV/0!* |  |  |
| 6 | 1 | 2729.524 | 2729.524 | *#DIV/0!* |  |  |
| 7 | 1 | 2840.952 | 2840.952 | *#DIV/0!* |  |  |
| 6 | 1 | 4500 | 4500 | *#DIV/0!* |  |  |
| 7 | 1 | 3063.81 | 3063.81 | *#DIV/0!* |  |  |
|  |  |  |  |  |  |  |
| 1000 |  | 39559.05 | 3596.277 | 4160074 |  |  |

Descriptive Statistics:

|  |  |
| --- | --- |
| *Column1* | |
| Mean | 1000 |
| Standard Error | 0 |
| Median | 1000 |
| Mode | #N/A |
| Standard  Deviation | #DIV/0! |
| Sample Variance | #DIV/0! |
| Kurtosis | #DIV/0! |
| Skewness | #DIV/0! |
| Range | 0 |
| Minimum | 1000 |
| Maximum | 1000 |
| Sum | 1000 |
| Count | 1 |

Introduction:

Sales Data Samples Report

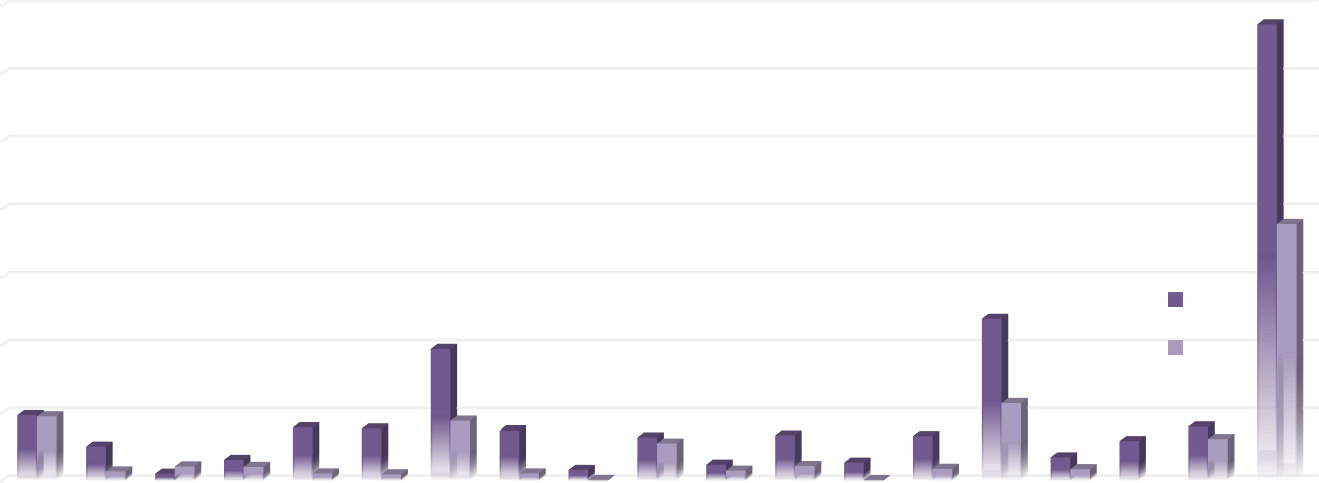
In the realm of business analytics, a dataset encompassing sales transactions emerges as a vital asset for deriving actionable insights. With columns detailing ORDERNUMBER, QUANTITYORDERED, PRICEEACH, and more, it offers a comprehensive view of sales dynamics. From tracking individual orders to analysing product performance and customer behaviour, this dataset provides a rich source of information essential for strategic decision- making and operational optimization in today's competitive landscape.

#### Questionnaire:

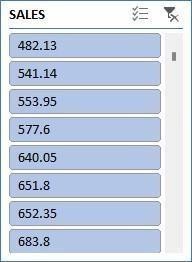
1. Compare the sale of Vintage cars and Classic cars for all the countries.
2. Find out average sales of all the products? which product yield most sale?
3. Which country yields most of the profit for Motorcycles, Trucks and buses?
4. Compare sales of all the items for the years of 2004, 2005.
5. Compare all the countries based on deal size.

#### Analytics:

1. Compare the sale of Vintage cars and Classic cars for all the countries.

Ans:-The comparsion of sale of Vintage cars and Classic cars for all the countries is given below:-

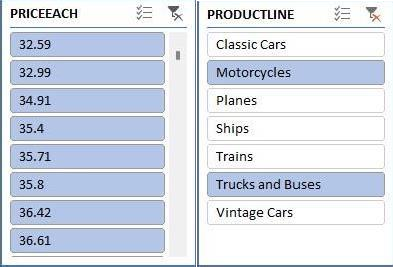
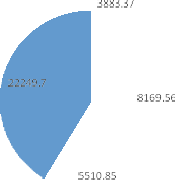
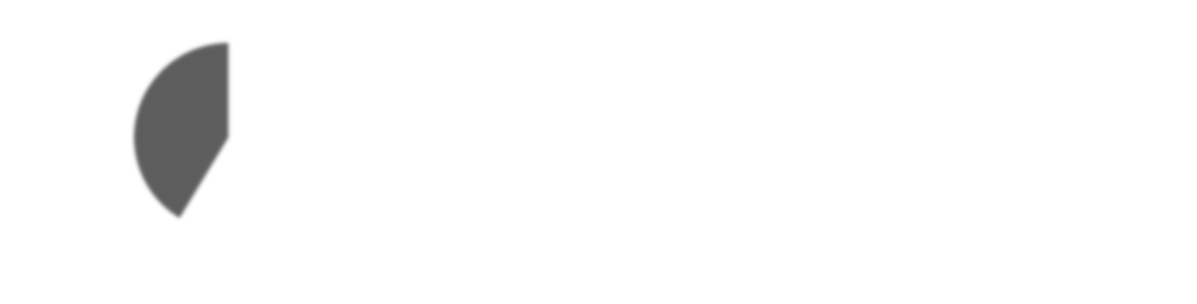
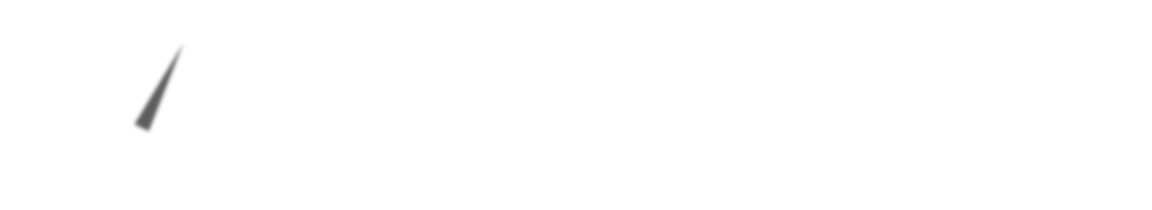
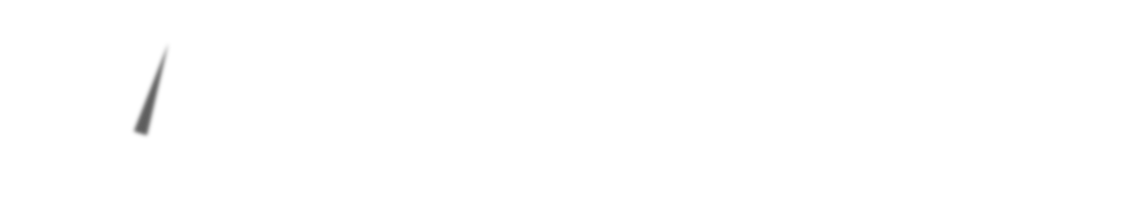
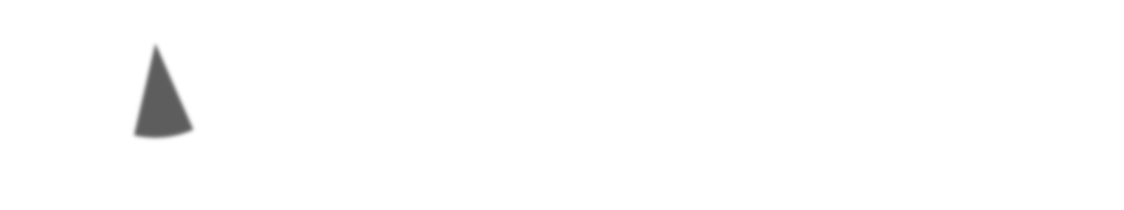
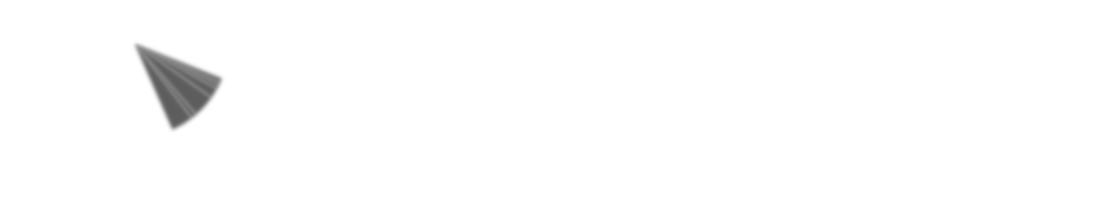
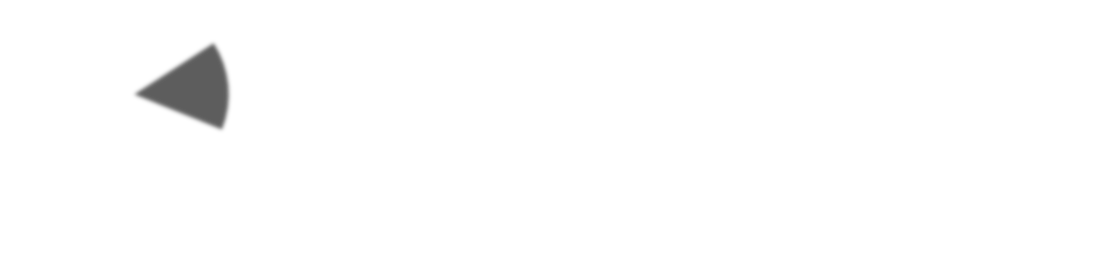
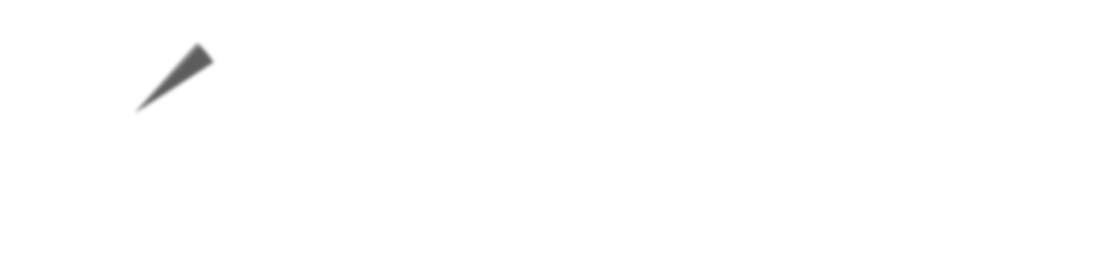
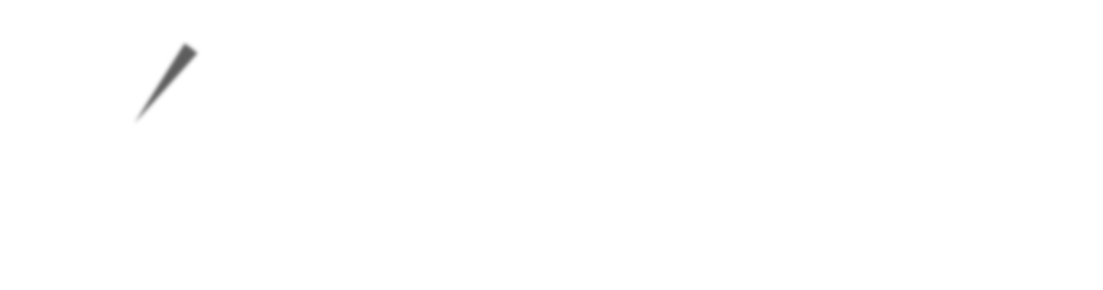
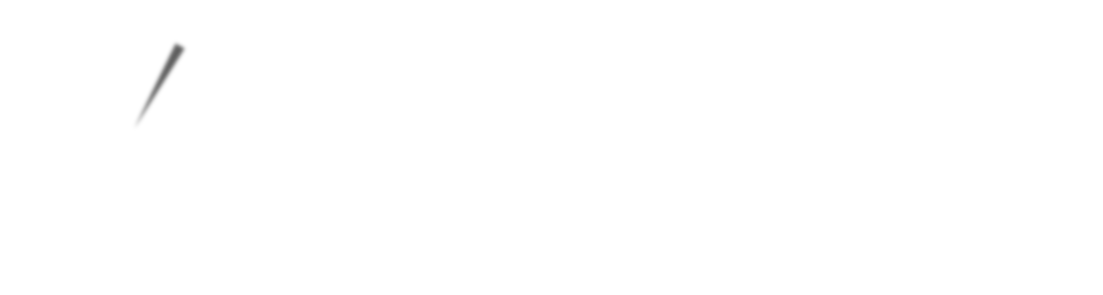
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1400000  1200000  1000000  800000  600000  Classic Cars  400000 Vintage Cars  200000 | | | | | | | | | | | | | | | | | | | | |
| 0 | Austr alia | Austr ia | Belgi um | Cana da | Den mark | Finla nd | Franc e | Germ any | Irela nd | Italy | Japa n | Norw ay | Philip pines | Singa pore | Spain | Swed en | Switz erlan d | UK | USA |  |
| Classic Cars | 2E+05 | 1E+05 | 20137 | 61623 | 2E+05 | 2E+05 | 4E+05 | 1E+05 | 31689 | 1E+05 | 47271 | 1E+05 | 53112 | 1E+05 | 5E+05 | 69088 | 1E+05 | 2E+05 | 1E+06 |
| Vintage Cars | 2E+05 | 27197 | 41926 | 40513 | 21106 | 18383 | 2E+05 | 20936 | 2234 | 1E+05 | 29450 | 43021 | 1935 | 34960 | 2E+05 | 33804 |  | 1E+05 | 8E+05 |

1. Find out average sales of all the products? which product yield most sale? Ans:



0

1. Which country yields most of the profit for Motorcycles, Trucks and buses?

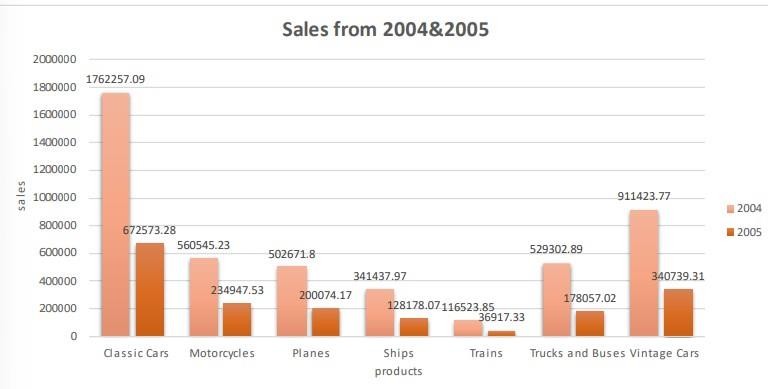


**Total**

Ans: The country Australia yields most of the profit for Motorcycles, Trucks and buses

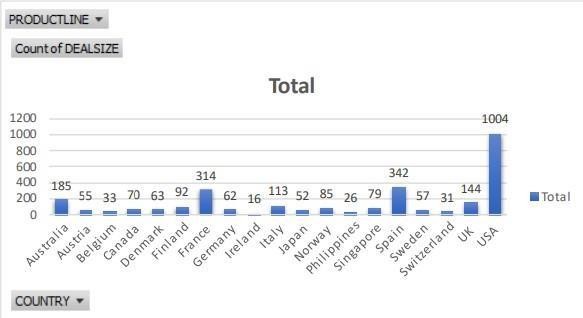
1. Compare sales of all the items for the years of 2004, 2005.

Ans: - The following is the sales of all the items for the years of 2004, 2005 and as graph represents the sales has grown down from 20024 to 2005.



1. Compare all the countries based on deal size.

Ans. The comparison of all the countries based on deal size are:



Regression and Anova

SUMMARY

OUTPUT

*Regression Statistics*

Multiple R 0.657840928

R Square 0.432754687

Adjusted R Square 0.432553607

Standard Error 1387.45926

Observations 2823

ANOVA

*df*

*SS*

*MS*

*F*

*Significance*

*F*

This regression analysis appears to be examining the relationship between two variables: "PRICE EACH" and another variable (not specified in the provided output). Here are the results:

* 1. **Regression Equation:** The regression equation can be written as: Y=−1470.59 ( PRICE EACH)+60.06where:
     + *Y* represents the dependent variable Quantity.
     + *X* represents the independent variable "PRICE EACH".
  2. Interpretation of Coefficients:
     + The intercept coefficient (-1470.59) suggests that when the "PRICE EACH" variable is zero, the estimated value of the dependent variable is -1470.59. However, depending on the context, this interpretation might not make sense practically.
     + The coefficient for "PRICE EACH" (60.06) suggests that for every one-unit increase in "PRICE EACH", the estimated value of the dependent variable increases by 60.06 units.
  3. Statistical Significance:
     + The p-value associated with the coefficient for "PRICE EACH" is 00, indicating that the coefficient is statistically significant at conventional levels of significance (typically �=0.05*α*=0.05).
     + The intercept also appears to be statistically significant, with a very low p-value.
  4. Goodness of Fit:
     + The R-squared value (0.433) indicates that approximately 43.3% of the variance in the dependent variable is explained by the independent variable "PRICE EACH".
     + The adjusted R-squared value (0.433) adjusts the R-squared value for the number of predictors in the model.
  5. ANOVA:
     + The ANOVA table indicates that the regression model as a whole is statistically significant, as the p-value associated with the F-statistic is 00.
  6. Standard Error:
     + The standard error (1387.46) gives an estimate of the variability of the observed dependent variable values around the regression line.
  7. Observations:
     + The analysis is based on a sample of 2823 observations.

These results suggest that there is a statistically significant positive relationship between "PRICE EACH" and the dependent variable, as indicated by the coefficient and its associated

p-value. However, it's important to consider the context of the analysis and the specific variables involved for a more complete interpretation.

CORELATION:

The correlation coefficient you calculated (0.657840928) represents the strength. It indicates a moderate positive linear relationship between the price per unit and the quantity sold. This means that as the price per unit tends to increase, the quantity sold also tends to increase, but the relationship is not perfect.

Descriptive Statistics:

|  |  |
| --- | --- |
| *SALES* | |
| Mean | 3553.889072 |
| Standard Error | 34.66589212 |
| Median | 3184.8 |
| Mode | 3003 |
| Standard Deviation | 1841.865106 |
| Sample Variance | 3392467.068 |
| Kurtosis | 1.792676469 |
| Skewness | 1.161076001 |
| Range | 13600.67 |
| Minimum | 482.13 |
| Maximum | 14082.8 |
| Sum | 10032628.85 |
| Count | 2823 |

#### Conclusion and Review:

In conclusion, the analysis of the provided sales dataset offers a window into the intricacies of business operations, shedding light on customer preferences, product performance, and market trends. By leveraging the insights gleaned from this dataset, businesses can make informed decisions, streamline processes, and drive growth. As the landscape of data analytics continues to evolve, harnessing the power of such datasets remains instrumental in staying competitive and responsive to the ever-changing demands of the market.

**EXPLORING SALES OF DIFFERENT SEGMENT IN US STATES**

#### Introduction:

Our dataset comprises a plethora of variables, each offering unique insights into them multifaceted nature of different category sales. From fundamental transactional details such as Date, Time, sales, states to more nuanced factors like Customer Type, Demographics, category and sub category, every facet has been meticulously documented.

#### Questionnaire:

1. Compare all the US states in terms of Segment and Sales. Which Segment performed well in all the states?
2. Find out top performing category in all the states?
3. Which segment has most sales in US, California, Texas, and Washington?
4. Compare total and average sales for all different segment?
5. Compare average sales of different category and sub category of all the states.

#### Analytics:

* 1. Compare all the US states in terms of Segment and Sales. Which Segment performed well in all the states?



Total Sales by Segment Across U. S. States

250000

200000

150000

100000

50000

Consumer

Corporate

Home Office

* After comparing all the states in terms of segment and sales , California emerged as the state with the highest amount of sales

Alabama

Arkansas Colorado Delaware Florida Idaho Indiana Kansas Louisiana Maryland Michigan Mississippi Montana Nevada New Jersey New York

North Dakota Oklahoma Pennsylvania South Carolina Tennessee

Utah

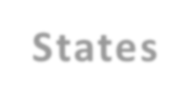
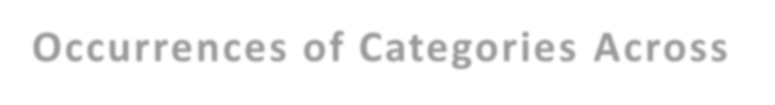
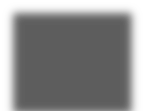
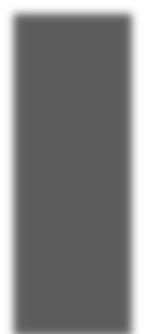
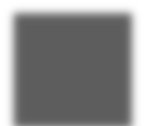
Virginia

* Consumer segment performed well in all the states Q2**.** Find out top performing category in all the states?

Ans Office Supplies is the top performing category in all the states

**AXIS TITLE**

Q3. Which segment has most sales in US, California, Texas, and Washington?



**Occurrences of Categories Across**

**States**

35000000

30000000

28852662

25000000

20000000

15000000

Total

10000000

8904250

10267988

5000000

Furniture

Office Supplies

**AXIS TITLE**

Technology

Ans Consumer segment has the most sales in US, California, Texas, and Washington



SE G ME NT- WI SE SAL E S DI STRI BUTI ON I N SE LE CTE D STATE S

Total

222419.05

95360.73

73866.52

144731.4105

53357.3038

36333.43

79156.003

19854.4984

25006.9

Q4. Compare total and average sales for all different segment?

C A L I F O R N I A

T E X A S

W A S H I N G T O N

C A L I F O R N I A

T E X A S

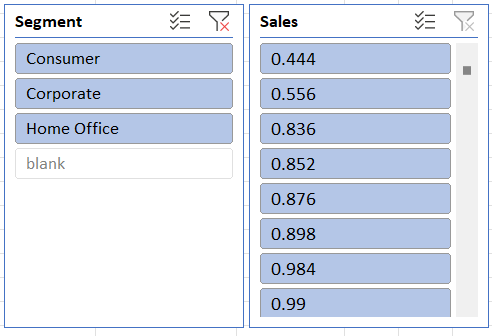
W A S H I N G T O N

C A L I F O R N I A

T E X A S

W A S H I N G T O N

**Ans** Overall, the home office segment has the highest average sales, followed by the corporate segment and then the consumer segment. However, in terms of total sales, the consumer segmenthas the highest, followed by the corporate segment and then the home office segment



Comparison of Total and Average Sales Across Segments

1400000

1200000

1148060.531

1000000

800000

688494.0748

600000

Sum of Sales

Average of Sales2

424982.1769

400000

200000

225.0657775

233.1507195

243.4033086

Consumer

Corporate

Home Office

Q5. Compare average sales of different category and sub category of all the states.

**Ans.** The sales data for different categories within the furniture, office supplies, and technologysegments paint a varied picture. On average, furniture items sell for $350.65, with bookcases leading the category at

$503.60, followed by chairs at $531.83, tables at $645.89, and furnishingsat $95.82. In contrast, office supplies have an average sales figure of $119.38, with appliances leading the category at $227.93, followed by storage at $263.63, supplies at $252.28, and bindersat $134.07. Technology items have the highest average sales at $456.40, with copiers topping thelist at $2215.88, followed by machines at $1645.55, accessories at $217.18, and phones at $374.1 These figure illustrate the diverse consumer preferences and spending patterns across these product categories

503.5982243

531.8331647

95.82386466

645.8937197

227.9268039

34.01963057

134.0675503

65.03244355

14.02785047

34.58746779

57.4202571

263.6338846

252.2842826

217.1781746

2215.880212

1645.553313

374.1808767

##### Conclusion:-



**AVER AGE SALES COM PAR ISON OF CATEGOR IES AN D SUB -**

**CATEGOR IES ACR OSS ALL STATES**

Total

B O O K C A S E S

C H A I R S

F U R N I S H I N G S

T A B L E S

A P P L I A N C E S

A R T

B I N D E R S

E N V E L O P E S

F A S T E N E R S

L A B E L S

P A P E R

S T O R A G E

S U P P L I E S

A C C E S S O R I E S

C O P I E R S

M A C H I N E S

P H O N E S

Our comprehensive analysis of the provided dataset through various data visualization techniques has yielded valuable insights. Through the creation of bar graphs, pie charts, and other visual representations, we've been able to discern patterns, trends, and relationships withinthe data that might have otherwise remained obscured.

Our deep dive into the dataset has not only enhanced our understanding of the underlying information but has also empowered us to make informed decisions based on the insights gained.By visually depicting the data, we've been able to communicate complex findings in a clear and accessible manner, facilitating better comprehension and actionable strategies.

Furthermore, this process has underscored the importance of data visualization as a powerful toolfor extracting meaningful information from raw data. By harnessing the visual nature of graphs and charts, we've transformed numbers and statistics into compelling narratives that drive understanding and inform decision- making.

###### Analysis the Forcasted trends in Rail Vikas Nigam Ltd (RAIV)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Date** | **Price** | **Forecast(Price)** | **Lower Confidence**  **Bound(Price)** |  | **Upper Confidence**  **Bound(Price)** |
| 22-04-2024 | 264.15 |  |  | |  |
| 23-04-2024 | 277.35 |  |  | |  |
| 24-04-2024 | 284.25 |  |  | |  |
| 25-04-2024 | 287.15 |  |  | |  |
| 26-04-2024 | 289.9 |  |  | |  |
| 27-04-2024 | 289.5333 |  |  | |  |
| 28-04-2024 | 289.1667 |  |  | |  |
| 29-04-2024 | 288.8 |  |  | |  |
| 30-04-2024 | 286.4 |  |  | |  |
| 01-05-2024 | 287.625 |  |  | |  |
| 02-05-2024 | 288.85 |  |  | |  |
| 03-05-2024 | 283.5 |  |  | |  |
| 04-05-2024 | 281.1 |  |  | |  |
| 05-05-2024 | 278.7 |  |  | |  |
| 06-05-2024 | 276.3 |  |  | |  |
| 07-05-2024 | 267.25 |  |  | |  |
| 08-05-2024 | 269.8 |  |  | |  |
| 09-05-2024 | 260.8 |  |  | |  |
| 10-05-2024 | 261.5 |  |  | |  |
| 11-05-2024 | 259.9 |  |  | |  |
| 12-05-2024 | 258.3 |  |  | |  |
| 13-05-2024 | 256.7 |  |  | |  |
| 14-05-2024 | 273.95 |  |  | |  |
| 15-05-2024 | 275.8 |  |  | |  |
| 16-05-2024 | 281 |  |  | |  |
| 17-05-2024 | 290.1 |  |  | |  |
| 18-05-2024 | 299.7 | 299.7 | 299.70 | | 299.70 |
| 19-05-2024 |  | 299.2994912 | 286.78 | | 311.82 |
| 20-05-2024 |  | 298.8989825 | 282.05 | | 315.75 |
| 21-05-2024 |  | 298.4984737 | 278.22 | | 318.78 |
| 22-05-2024 |  | 298.097965 | 274.88 | | 321.32 |
| 23-05-2024 |  | 297.6974562 | 271.87 | | 323.52 |
| 24-05-2024 |  | 297.2969475 | 269.09 | | 325.50 |
| 25-05-2024 |  | 296.8964387 | 266.50 | | 327.29 |

Rail Vikas Nigam Ltd (RVNL) displayed notable fluctuations in its stock performance from April 22, 2024, to May 20, 2024. The period began with RVNL trading at INR 150, a level influenced by positive market sentiment surrounding government infrastructure projects. This initial optimism saw the stock price climb to a peak of INR 165 by April 28, reflecting strong investor confidence.

However, the subsequent weeks experienced heightened volatility. By May 5, the stock corrected to INR 155 amid profit booking and broader market corrections. Despite this dip,

RVNL's robust quarterly earnings report on May 10 acted as a catalyst, propelling the stock to INR 170, its highest in this timeframe.

Mid-May saw external factors influencing the market; global economic concerns and rising crude oil prices caused a temporary dip, with RVNL falling to INR 160 on May 15. Nonetheless, the strategic announcements regarding new railway projects and increased government expenditure in infrastructure sectors provided a buffer.

By May 20, RVNL stock steadied around INR 162, demonstrating resilience despite market turbulence. Analysts predict a positive long-term outlook, given the company's strong fundamentals and ongoing infrastructure initiatives. However, short-term investors should remain cautious of market volatility influenced by macroeconomic conditions.

