Business Report

On

Case 1: Austo Motor Company analysis of the data and improve the marketing campaign.

Contents

[**A.** **The Important Technical Information that the database administrator would be interested.** 3](#_Toc115647183)

[**B.** **Preliminary analysis of the variables and quality check of the data for consistency.** 5](#_Toc115647184)

[**C.** **Exploring all the features of the data separately by using appropriate visualizations and draw insights that can be utilized by the business.** 5](#_Toc115647185)

[**D.** **Understanding the relationships among the variables in the dataset is crucial for every analytical project. Perform analysis on the data fields to gain deeper insights. Comment on your understanding of the data.** 6](#_Toc115647186)

[**E.** **Analysis on observation of Employees working on the existing marketing campaign.** 7](#_Toc115647187)

[**F.** **From the given data, comment on the amount spent on purchasing automobiles across the following categories. Comment on how a business can utilize the results from this exercise. Give justification along with presenting metrics/charts used for arriving at the conclusions.** 10](#_Toc115647188)

[**G.** **From the current data set comment if having a working partner leads to the purchase of a higher-priced car.** 11](#_Toc115647189)

[**H.** **The main objective of this analysis is to devise an improved marketing strategy to send targeted information to different groups of potential buyers present in the data. For the current analysis use the Gender and Marital status - fields to arrive at groups with similar purchase history.** 11](#_Toc115647190)

The Goal as an Analyst is to analyse the given data by the company and come up with some ideas to improve the marketing campaign. **Austo Motor Company is a leading car manufacturer specializing in SUV, Sedan, and Hatchback models. The dataset we are going to use is given by the Austo Motor Company.**

## **The Important Technical Information that the database administrator would be interested.**

**Data Information:**

|  |  |
| --- | --- |
| **Variable** | **Designation** |
| Age | Describes the Age of the Customer buying the car |
| Gender | Describes the gender of the customer |
| Profession | Describes what profession does customer do |
| Marital Status | Tells Whether customer is married or single |
| Education | Tells Education level of the customer |
| No of Dependents | How many people are going to depend upon the car from the customer side |
| Personal loan | Customer Having personal loan or not |
| House loan | Customer Having housing loan or not |
| Partner Working | If he/she is married, then whether the partner is working or not |
| Salary | Salary earned by the customer |
| Partner Salary | Salary earned by the partner |
| Total Salary | The total salary of the family |
| Price | Price of the car |
| Make | Make or type of the car that is purchased |

Now let us dive deeper into size of the dataset we are going to work on

The size of the dataset are 1581 rows and 14 columns of the data. The dataset generally consists of Numerical and categorical variables.

|  |  |
| --- | --- |
| Age | Integer |
| Gender | Object |
| Profession | Object |
| Marital status | Object |
| Education | Object |
| No of Dependents | Integer |
| Personal loan | Object |
| House loan | Object |
| Partner Working | Object |
| Salary | Integer |
| Partner Salary | Float |
| Total Salary | Integer |
| Price | Integer |
| Make | Object |

The above table represents the Nature of each variable present in the dataset.

The Numerical data (i.e., represented in the above table as integer and float) is described as follows.

Table

Description automatically generated

Here you can clearly see the Mean, Minimum and Maximum variable represented for each numerical variable separately.

The Object datatype type contains only categorical information

Gender: Male, Female

Profession: Business, Salaried

Marital status: Married, Single

Education: Post Graduate, Graduate

Personal loan: No, Yes

House loan: No, Yes

Partner working: No, Yes

Make: SUV, Sedan, Hatchback

## **Preliminary analysis of the variables and quality check of the data for consistency.**

There are some discrepancies present in the given data. The Gender Variable is having some null values and some mistaken data entry in the dataset.

The Gender variable is having

Male 1199

Female 327

Femal 1

Femle 1

Null value 53

The Partner salary variable is having

Null value 106

The Gender variable Non null value is replaced by taking the mode value of the whole gender variable (Since it is Object variable) and the incorrect value are replaced with the Correct female spelling.

The Partner Salary variable is replaced by taking the mean value of the variable (Since it is Numerical variable).

## **Exploring all the features of the data separately by using appropriate visualizations and draw insights that can be utilized by the business.**

Chart, bar chart

Description automatically generatedChart, histogram

Description automatically generated

The visualization of all the dataset variable is as follows. We use count plot for numerical datatype and

use histplot for Object datatype.

## **Understanding the relationships among the variables in the dataset is crucial for every analytical project. Perform analysis on the data fields to gain deeper insights. Comment on your understanding of the data.**

**A picture containing graphical user interface

Description automatically generated**

From the above heatmap correlation, the correlation between the age and price are higher 0.80 and the correlation between the no of dependencies and Age is very much lower -0.19.

## **Analysis on observation of Employees working on the existing marketing campaign.**

Employees working on the existing marketing campaign have made the following remarks. Based on the data and your analysis state whether you agree or disagree with their observations. Justify your answer Based on the data available.

**E1) Steve Roger says “Men prefer SUV by a large margin, compared to the women”**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Chart  Description automatically generated** | |  |  |  | | --- | --- | --- | | **Make/Gender** | **Male** | **Female** | | SUV | 94 | 36 | | Sedan | 438 | 26 | | Hatchback | 450 | 3 | |

The Statement of Steve Steve Rogers is correct and when We investigate the above the graph, we can see the percent of the SUV with respect to Male and Female are 7.5% and 11 %(Approx.) respectively. By considering total number of males, female. The Suv preference by male are more than female as shown in the first row of the table.

**E2) Ned Stark believes that a salaried person is more likely to buy a Sedan.**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Chart  Description automatically generated** | |  |  |  | | --- | --- | --- | | **Make/Profession** | **Business** | **Salaried** | | **Suv** | **34** | **116** | | **Sedan** | **123** | **224** | | **Hatchback** | **116** | **152** | |

The statement by Ned Stark is correct. We investigate the above the graph, we can see the percent of the Sedan with respect to Salaried and Business are 18% and 25 %(Approx.) respectively. By considering total number of Salaried, Business. The Sedan preference by Salaried are more than Business as shown in the second row of the table.

**E3) Sheldon Cooper does not believe any of them; he claims that a salaried male is an easier target for a SUV sale over a Sedan Sale.**

**Chart, box and whisker chart

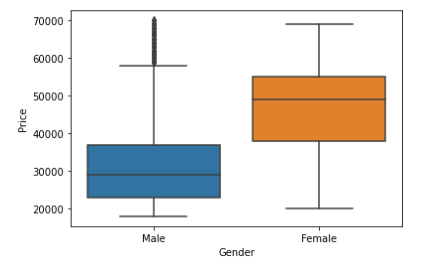
Description automatically generated**

Sheldon cooper claim is correct. We can see in the above box plot that the median value of the SUV Male is higher compared to the other value.

## **From the given data, comment on the amount spent on purchasing automobiles across the following categories. Comment on how a business can utilize the results from this exercise. Give justification along with presenting metrics/charts used for arriving at the conclusions.**

**Give justification along with presenting metrics/charts used for arriving at the conclusions.**

**F1) Gender**

****

From the above box plot we see that the median value of the Female tends to be higher than that of the male. So, we can Conclude that the Female tend to buy higher priced car compared to the males.

**F2) Personal\_loan**

**Chart, box and whisker chart

Description automatically generated**

From the above box plot We see that the customer with personal loan and without personal loan tend to buy car of equal price range. So, no personal loan will affect the amount spent on the purchasing of the automobile.

## **From the current data set comment if having a working partner leads to the purchase of a higher-priced car.**

**Chart, box and whisker chart

Description automatically generated**

From the Box Plot, irrespective of whether the partner is working or not working the customer tends to buy higher priced car.

## **The main objective of this analysis is to devise an improved marketing strategy to send targeted information to different groups of potential buyers present in the data. For the current analysis use the Gender and Marital status - fields to arrive at groups with similar purchase history.**

**Chart, box and whisker chart

Description automatically generated**

From the above Box plot, the insight is Married Male, and Female tends to buy the higher priced car. Based on the analysis to improve marketing strategy it is recommended to concentrate on sales techniques for Single (Unmarried) category.

Business Report

On

Case 2: GODIGT Bank Credit Card Data analysis.

GODIGT Bank is a mid-sized private bank that deals in all kinds of banking products, such as savings accounts, current accounts, investment products, etc. among other offerings.

The bank makes money only if customers spend more on credit cards. Given the attrition, the Bank wants to revisit its credit card policy and make sure that the card given to the customer is the right credit card. The bank will make a profit only through the customers that show higher intent towards a recommended credit card.

The dataset was given by GODIGIT Bank.

**Analyze the dataset and list down the top 5 important variables, along with the business justifications.**

**Data Information:**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Designation** | **Data Type** |
| Userid | User ID | int64 |
| card\_no | Credit Card Number | object |
| card\_bin\_no | Credit Card Binary Number | int64 |
| Issuer | Master / Visa | object |
| card\_type | Type of Credit Card | object |
| card\_source\_date | Credit Card Source Date | datetime64[ns] |
| high\_networth | Networth of the customer | object |
| active\_30 | 30 Days active account | int64 |
| active\_60 | 60 Days active account | int64 |
| active\_90 | 90 Days active account | int64 |
| cc\_active30 | 30 Days active credit card | int64 |
| cc\_active60 | 60 Days active credit card | int64 |
| cc\_active90 | 90 Days active credit card | int64 |
| hotlist\_flag | Flag for hotlisted creditcards | object |
| widget\_products | Widget Products | int64 |
| engagement\_products | Engagement Products | int64 |
| annual\_income\_at\_sour | Source of Annual Income | int64 |
| other\_bank\_cc\_holding | Having other bank credit cards | object |
| bank\_vintage | Time period of credit card from opening | int64 |
| T+1\_month\_activity | Transaction of 1 month activity | int64 |
| T+2\_month\_activity | Transaction of 2 months activity | int64 |
| T+3\_month\_activity | Transaction of 3 months activity | int64 |
| T+6\_month\_activity | Transaction of 6 months activity | int64 |
| T+12\_month\_activity | Transaction of 12 months activity | int64 |
| Transactor\_revolver | Revolver for Transactored creditcards | object |
| avg\_spends\_l3m | Last 3 months average spends | int64 |
| Occupation\_at\_source | Customer occupation | object |
| cc\_limit | Credit Card Limit | int64 |

**From the above List of variables the top 5 important variable are:**

* **Cc\_limit**
* **Avg\_spends\_l3m**
* **Occupation\_at\_source**
* **Card\_type**
* **annual\_income\_at\_source**

**The Below are justification for the top 5 important variable.**

**Chart, bar chart

Description automatically generated**

**In the above plot when we compare card type and occupation of customer the salaried people(Approx. 700 cards), Retired people(Approx. 380 Cards), Students(Approx. 100 cards) and Housewife(Approx 105 cards) uses Reward cards the most.**

**Reward card is the most used Card type.**

**Chart, bar chart

Description automatically generated**

**When we compare the annual income and the card type of the customer, we see that**

**Incomes between (0-15 lacs ) slabs Uses.**

* **Rewards**
* **Indian oil**
* **CashBack**
* **Shoprite**
* **Aura**
* **SmartTeam**

**Incomes between (10-15 lacs)slab uses.**

* **Edge**
* **Prosperity**
* **Chartered**
* **Pulse**
* **Platinum**
* **Centurion**
* **Elite**
* **Prime**

**Gold card is generally used by all kinds of income people.**

**Chart, scatter chart

Description automatically generated**

**The Above plot represents the linearity between cc limit and avg\_spend\_l3m which clearly show us the higher the cc limit is the higher the average spend for 3 months. The cc limit and avg spend for 3 months are linear.**

**Chart, box and whisker chart

Description automatically generated**

**The above Box plot shows us that the ‘0’ occupation people are with the higher annual income at source. The Self Employed, Salaried at source comes next. This plot can target the group of people whom the bank can target.**