Business Report

by

Aaryani Kadiyala

For

SMDM Project

Problem 1

Analysts are required to explore data and reflect on the insights. Clear writing skill is an integral part of a good report. Note that the explanations must be such that readers with minimum knowledge of analytics is able to grasp the insight.

Austo Motor Company is a leading car manufacturer specializing in SUV, Sedan, and Hatchback models. In its recent board meeting, concerns were raised by the members on the efficiency of the marketing campaign currently being used. The board decides to rope in an analytics professional to improve the existing campaign.

1. You as an analyst have been tasked with performing a thorough analysis of the data and coming up with insights to improve the marketing campaign.

The instructions below are given to help you complete the project –

A. Data Overview

Below are the dataset and their data types which are important for a data base administrator
Number of Rows: 1581 entries, 0 to 1580
Data columns (total 14 columns): Number of Variables

#	Column	Non-Null Count	Dtype
0	Age	1581 non-null	int64
1	Gender	1528 non-null	object
2	Profession	1581 non-null	object
3	Marital_status	1581 non-null	object
4	Education	1581 non-null	object
5	No_of_Dependents	1581 non-null	int64
6	Personal_loan	1581 non-null	object
7	House_loan	1581 non-null	object
8	Partner_working	1581 non-null	object
9	Salary	1581 non-null	int64
10	Partner_salary	1475 non-null	float64
11	Total_salary	1581 non-null	int64
12	Price	1581 non-null	int64
13	Make	1581 non-null	object

dtypes: float64(1), int64(5), object(8)

Data types of above variables
Categorical Variables:

Binary:	Multilevel:	Continuous Variable:	Discrete Variable:
Gender	Make	Price	Age
Marital_status	Education	Salary	No_of_Dependents
Personal_loan	Profession	Partner_salary	
House_loan		Total_salary	
Partner_working			

- **Yes, there are discrepancies in the Partner_salary and Gender data, there are null data available in both of these attributes and values need to imputed for these null values.**
- **Female is spelled incorrectly in the gender column for 2 records and are corrected.**
- **Null values in Gender column is replace by Male as maximum count of records is having Gender as Male**

In the given dataset

```
Male      1199
Female     327
Femal        1
Femle        1
Name: Gender, dtype: int64
```

After Imputing the Values

```
Male      1252
Female     329
Name: Gender, dtype: int64
```

- **Partner_salary values to be imputed by 'Total_salary-Salary'.**

```
Data columns (total 14 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Age               1581 non-null   int64
 1   Gender            1528 non-null   object
 2   Profession        1581 non-null   object
 3   Marital_status    1581 non-null   object
 4   Education         1581 non-null   object
 5   No_of_Dependents  1581 non-null   int64
 6   Personal_loan     1581 non-null   object
 7   House_loan        1581 non-null   object
 8   Partner_working   1581 non-null   object
 9   Salary            1581 non-null   int64
 10  Partner_salary    1475 non-null   float64
 11  Total_salary      1581 non-null   int64
 12  Price             1581 non-null   int64
 13  Make              1581 non-null   object
dtypes: float64(1), int64(5), object(8)
```

In the given data

```
Data columns (total 14 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Age               1581 non-null   int64
 1   Gender            1581 non-null   object
 2   Profession        1581 non-null   object
 3   Marital_status    1581 non-null   object
 4   Education         1581 non-null   object
 5   No_of_Dependents  1581 non-null   int64
 6   Personal_loan     1581 non-null   object
 7   House_loan        1581 non-null   object
 8   Partner_working   1581 non-null   object
 9   Salary            1581 non-null   int64
 10  Partner_salary    1581 non-null   float64
 11  Total_salary      1581 non-null   int64
 12  Price             1581 non-null   int64
 13  Make              1581 non-null   object
dtypes: float64(1), int64(5), object(8)
```

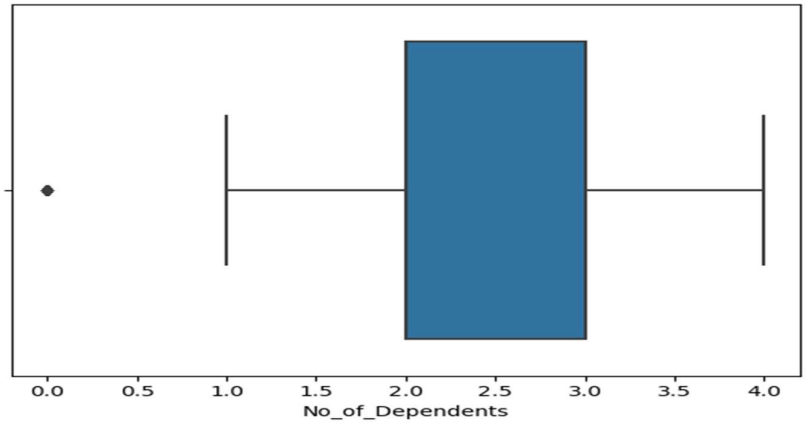
After Imputing the Value

B. Univariant Analysis

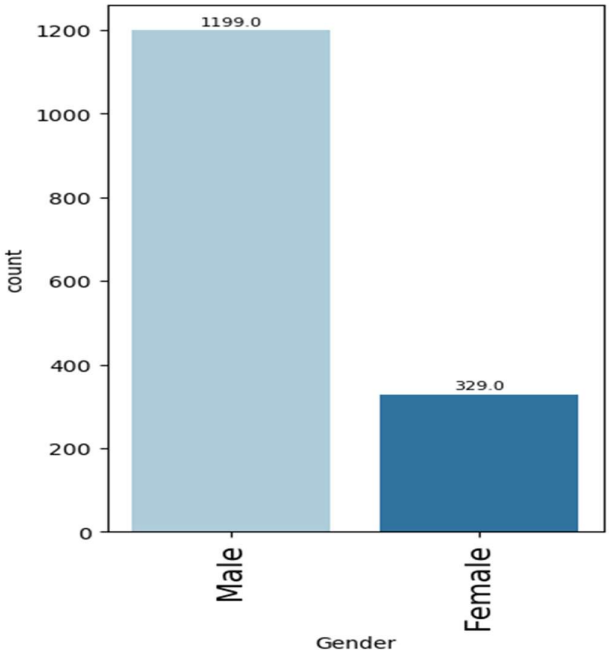
- **70% of the cars are brought by the people below 40 Years**

	Age	No_of_Dependents	Salary	Partner_salary	Total_salary	Price
count	1581.000000	1581.000000	1581.000000	1581.000000	1581.000000	1581.000000
mean	31.922201	2.457938	60392.220114	20585.895003	79625.996205	35597.722960
std	8.425978	0.943483	14674.825044	18952.938643	25545.857768	13633.636545
min	22.000000	0.000000	30000.000000	0.000000	30000.000000	18000.000000
25%	25.000000	2.000000	51900.000000	0.000000	60500.000000	25000.000000
50%	29.000000	2.000000	59500.000000	25600.000000	78000.000000	31000.000000
75%	38.000000	3.000000	71800.000000	38000.000000	95900.000000	47000.000000
max	54.000000	4.000000	99300.000000	80500.000000	171000.000000	70000.000000

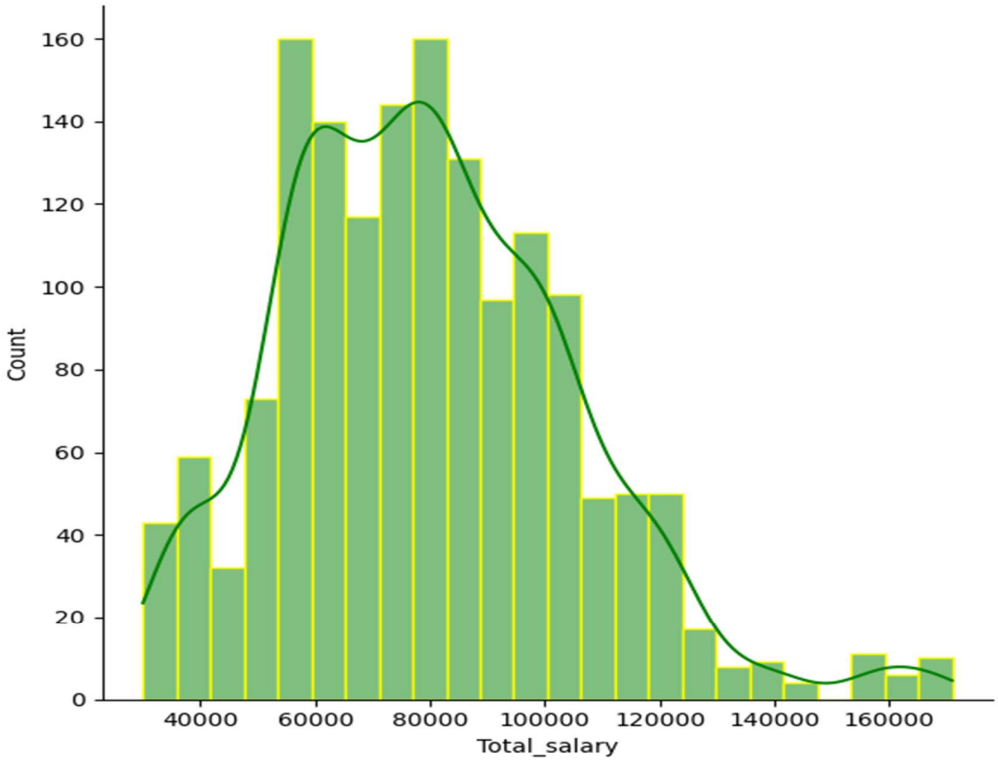
- **Maximum cars are brought by the people where No_of_Dependents are 2 & 3.**

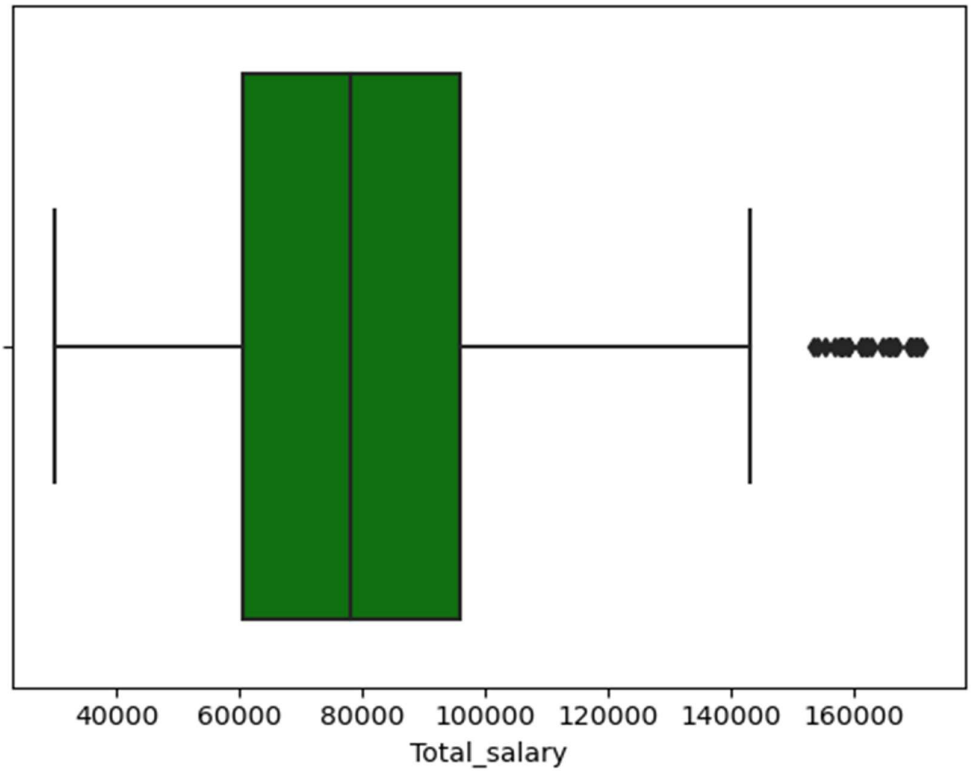


- **Maximum cars are brought by the people with Male Gender**



- **Total _salary is double peaked and lightly rightly skewed and have outliers**

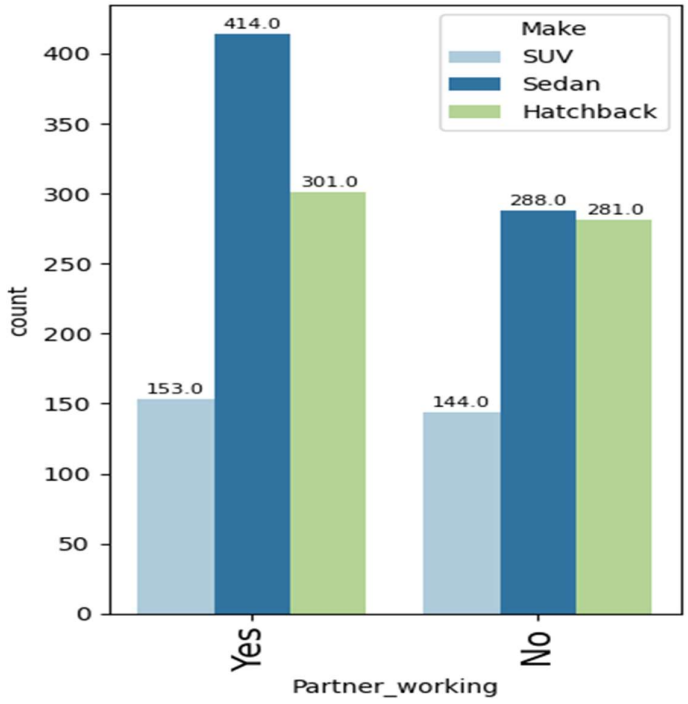




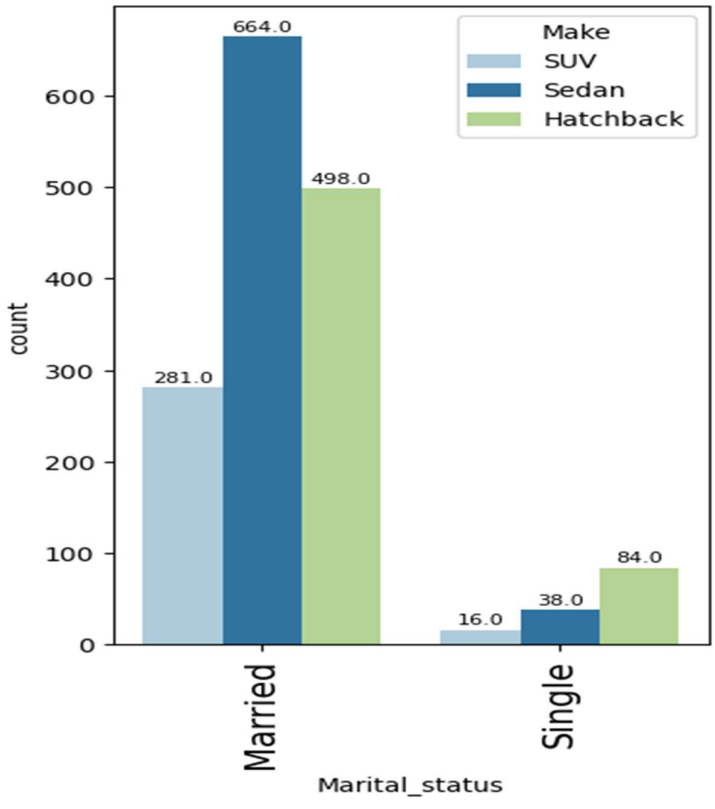
There are 27 records which are greater than the maximum Price from IQR method, these records can be analyzed separately.

C. Bivariant Analysis

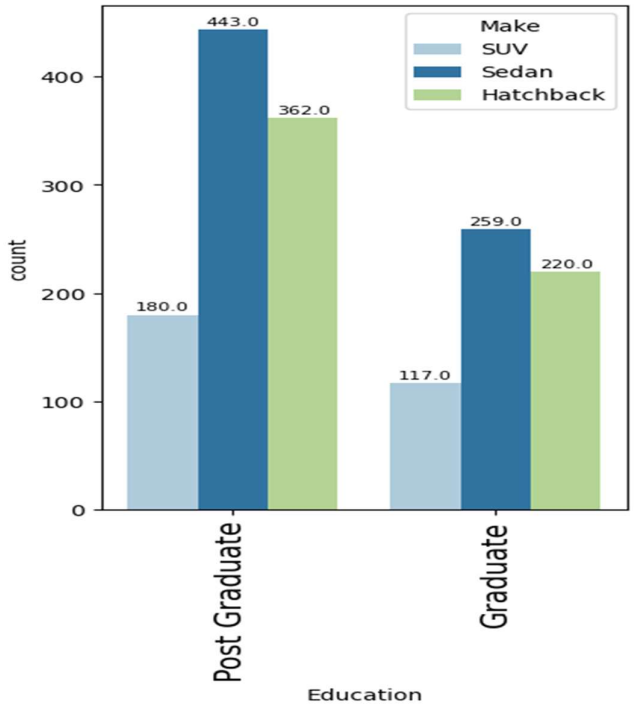
- **Sedan cars are purchased more especially when there are working partners.**



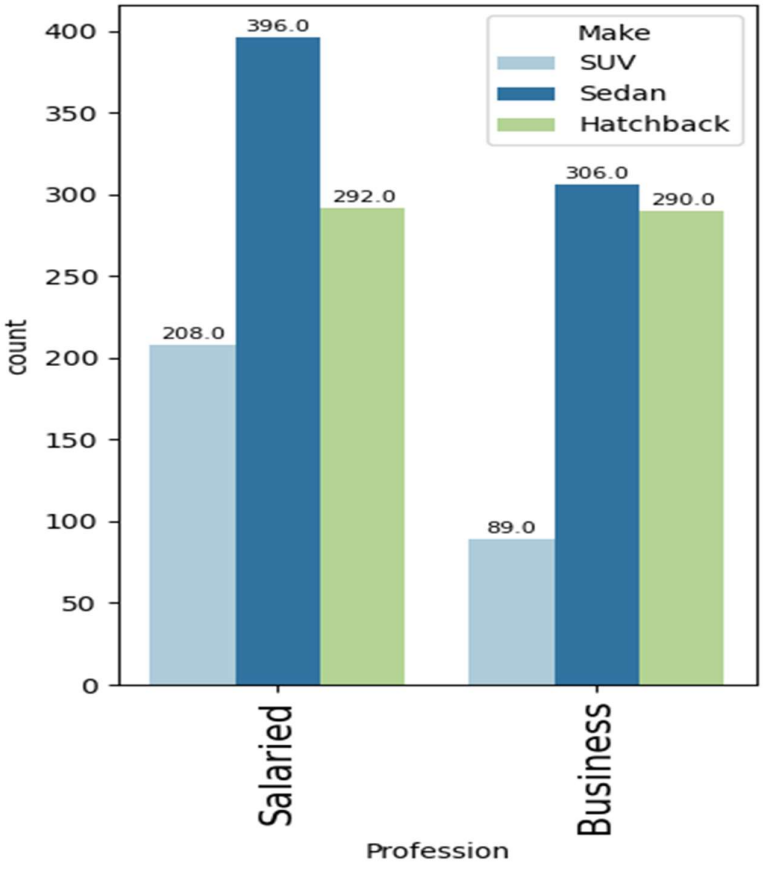
- **Married people purchased more cars especially Sedan Make is the highest among them purchased cars.**



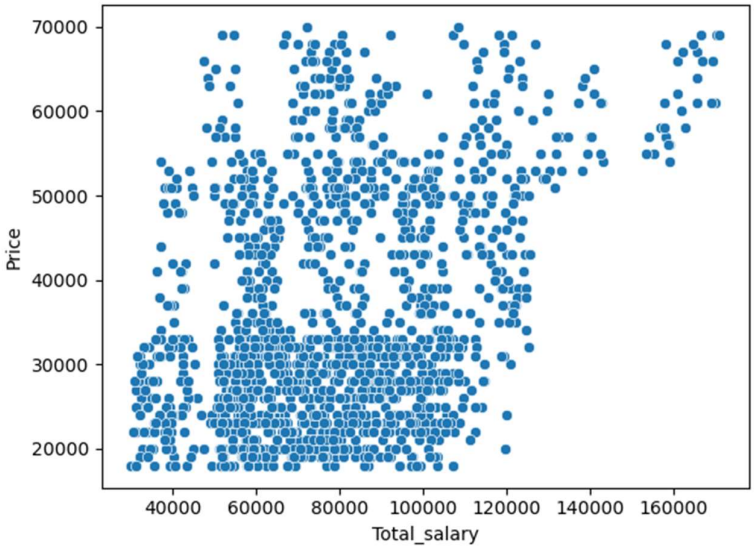
- **Post Graduated persons purchased more cars especially Sedan Make is the highest among them purchased cars.**



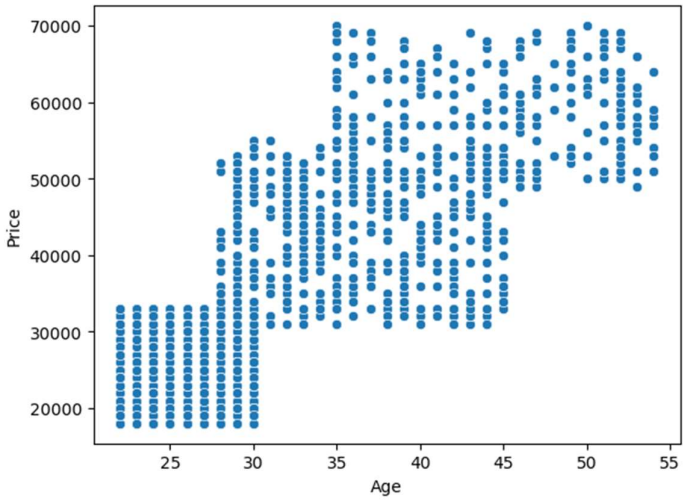
- **Salaried persons purchased more cars especially Sedan Make is the highest among them purchased cars.**
- **SUVs are purchased by salaried than the people in business profession.**



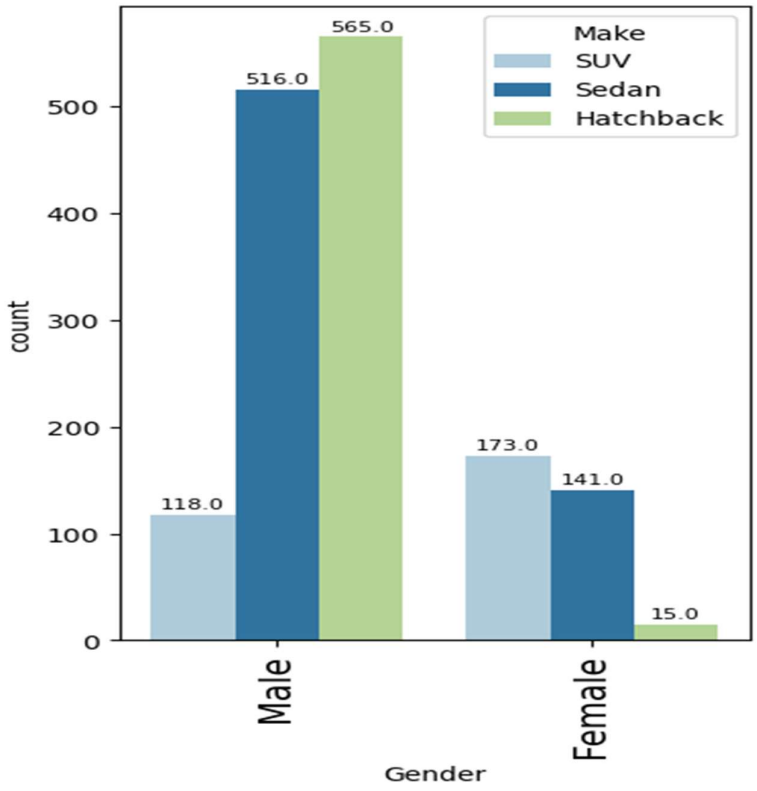
- **No much co-relation between the Total_salary and the price of the car that they purchased.**



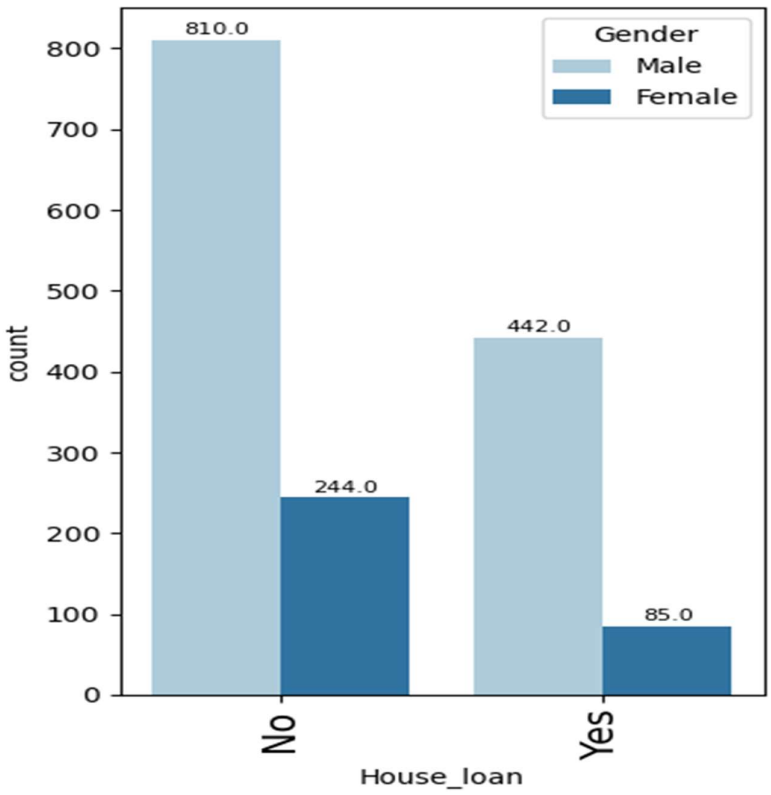
- **Based on the age group the price range of the car purchased in similar.**



- **SUV make are the preferred cars by Females, followed by sedan and then hatchback.**



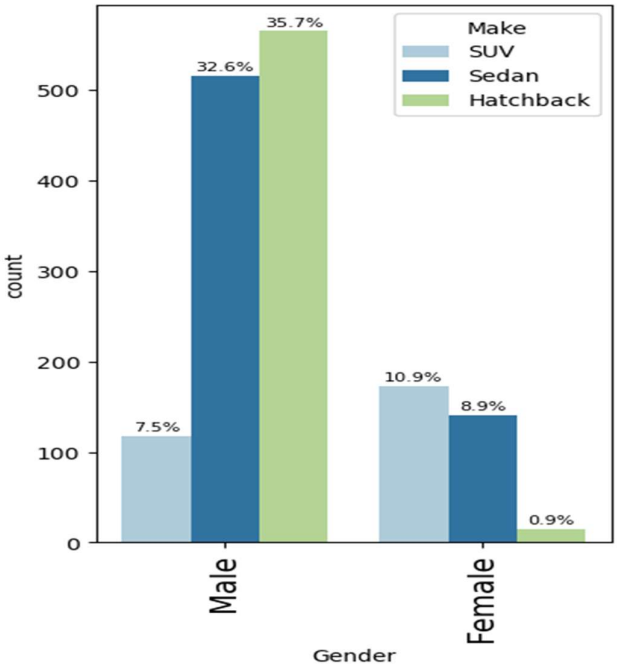
- **Persons having 'No' House_loan have purchased more cars than the people having House_loan.**

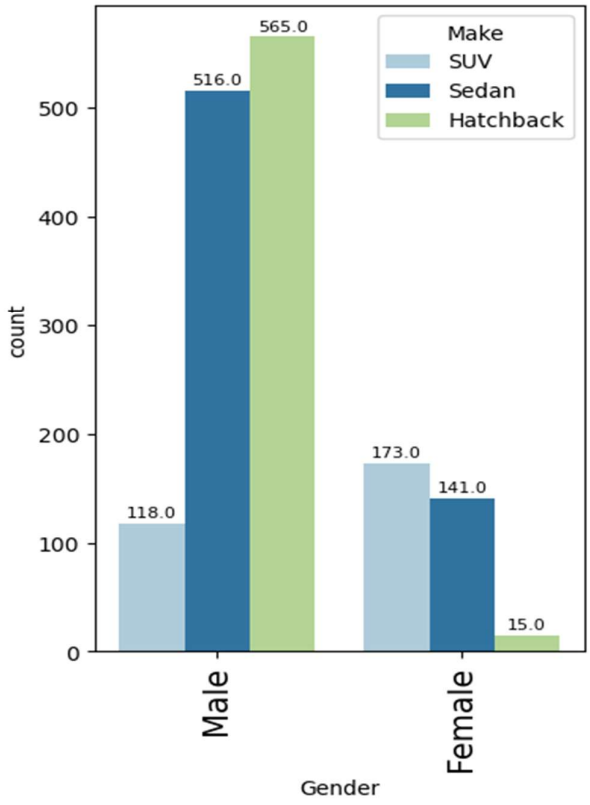


D. Key Questions

Q1) Do men tend to prefer SUVs more compared to women?

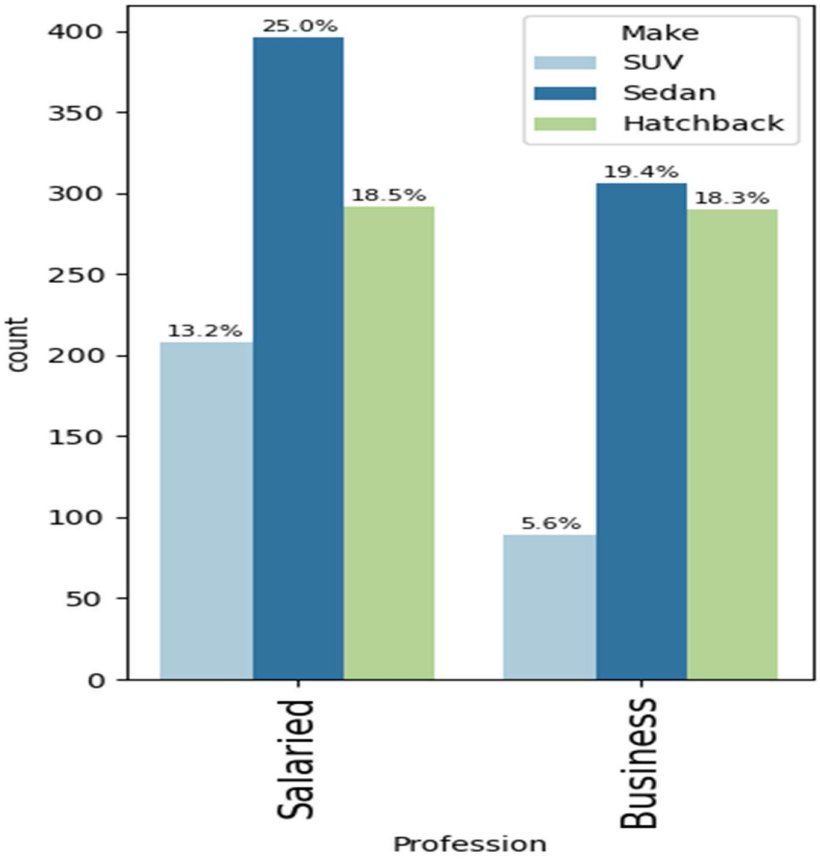
Yes, Men tend to prefer SUVs more compared to women.

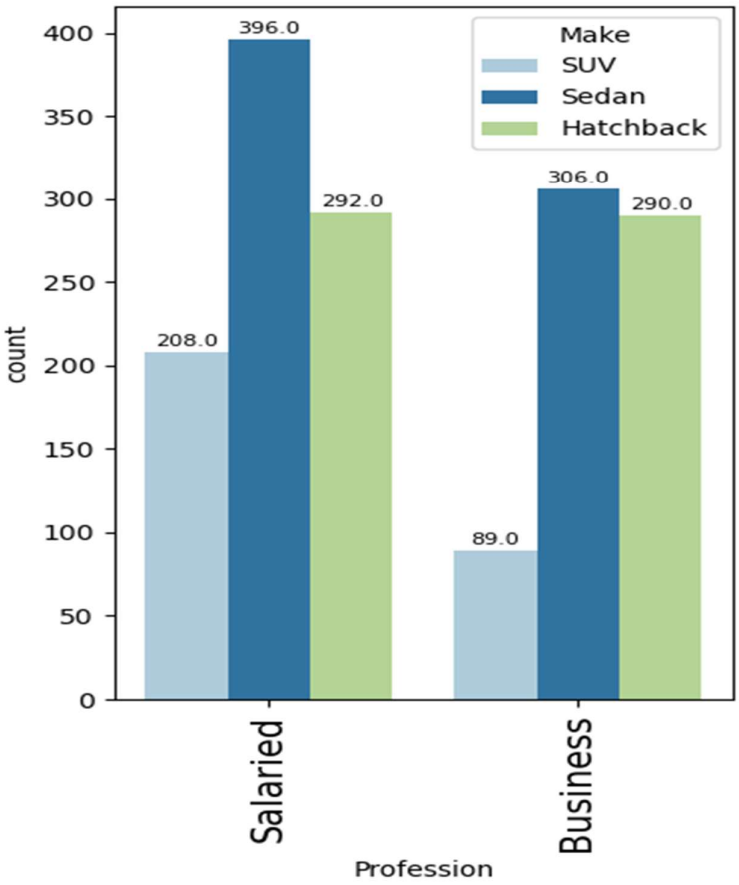




Q2) What is the likelihood of a salaried person buying a Sedan?

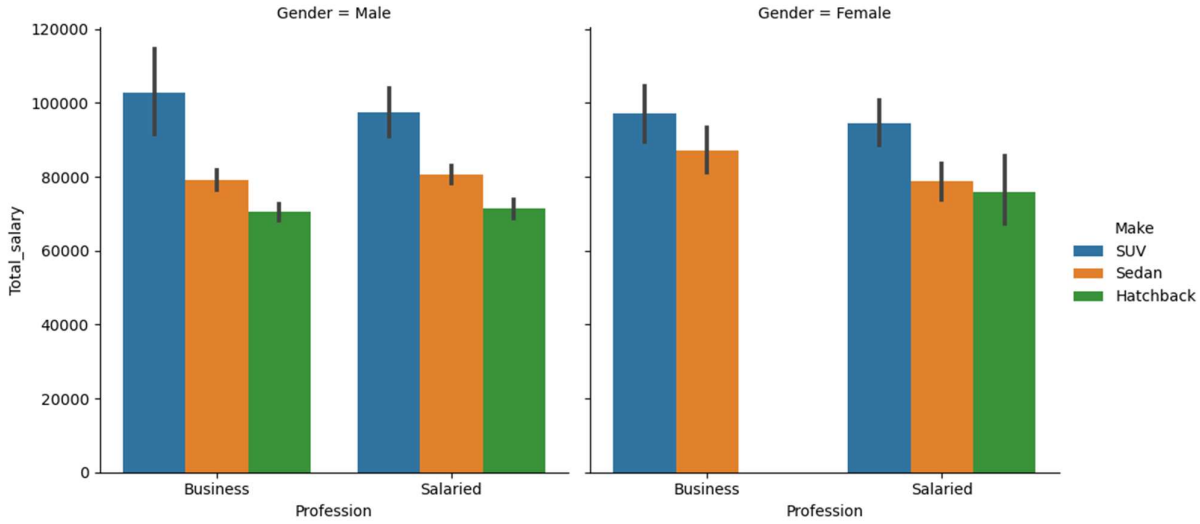
Yes, salaried person is more likely to buy a Sedan.





Q3) What evidence or data supports Sheldon Cooper's claim that a salaried male is an easier target for a SUV sale over a Sedan sale?

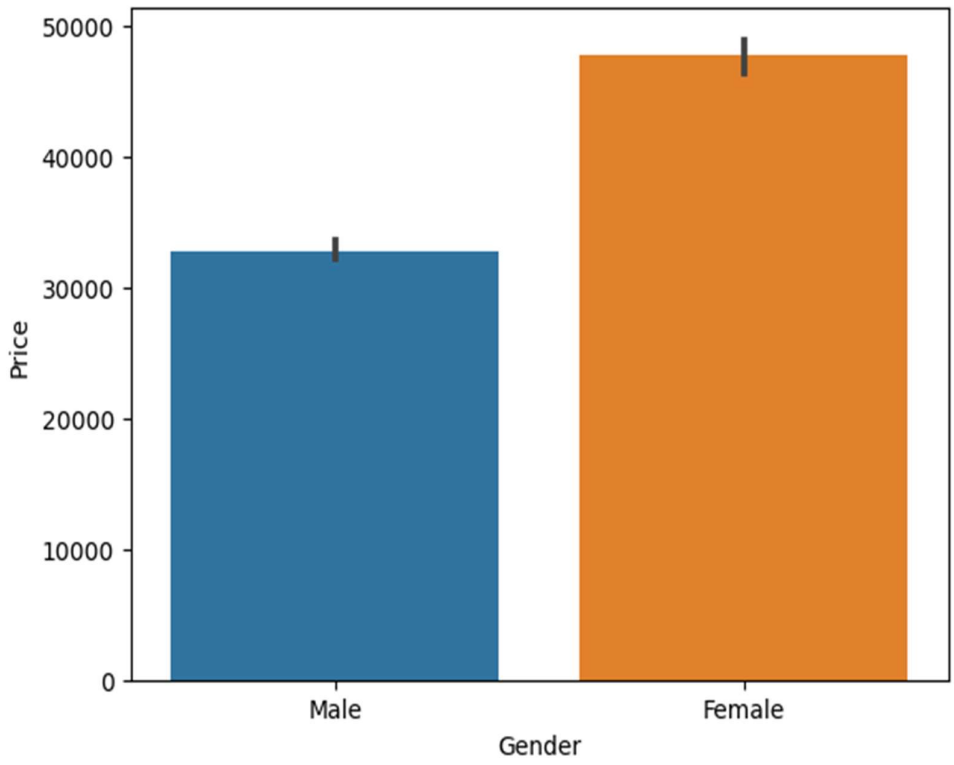
Yes, salaried men are easier target for suv sale over sedan sale.



Q4) How does the the amount spent on purchasing automobiles vary by gender?

Male spent more amount in purchasing cars than women.

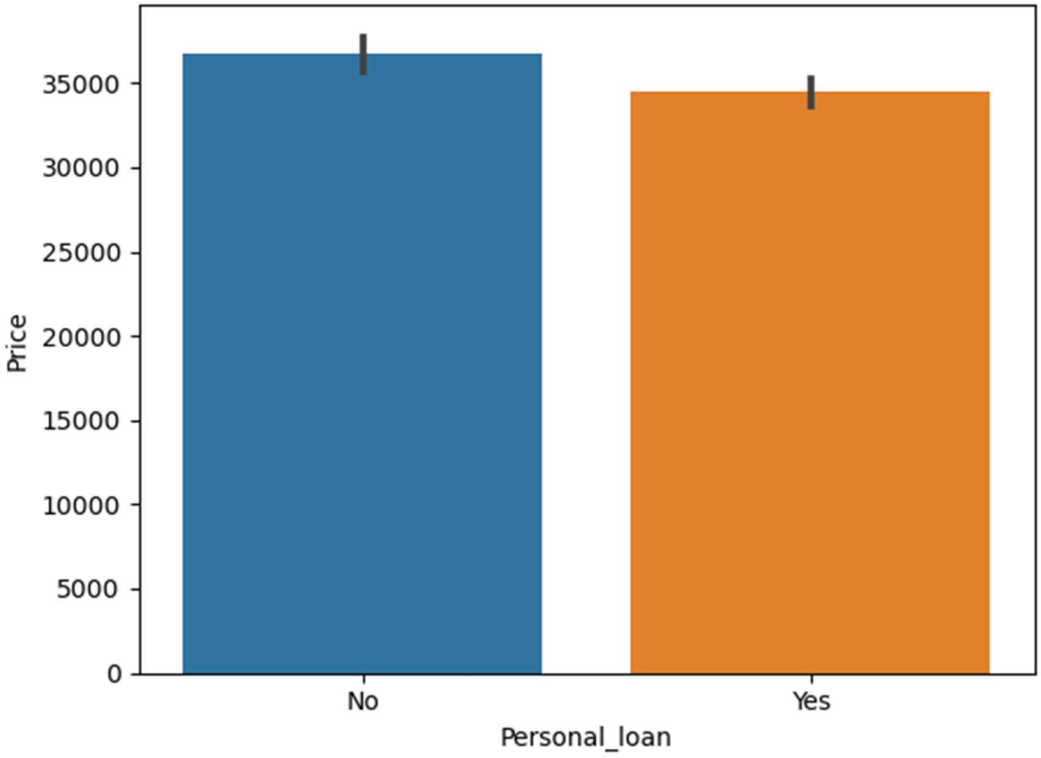
```
Gender
Female    15695000
Male      40585000
Name: Price, dtype: int64
```



Q5) How much money was spent on purchasing automobiles by individuals who took a personal loan?

Similar range of amount is spent on purchasing automobiles by both the people having and not having the personal loan

```
Personal_loan
No     28990000
Yes    27290000
Name: Price, dtype: int64
```

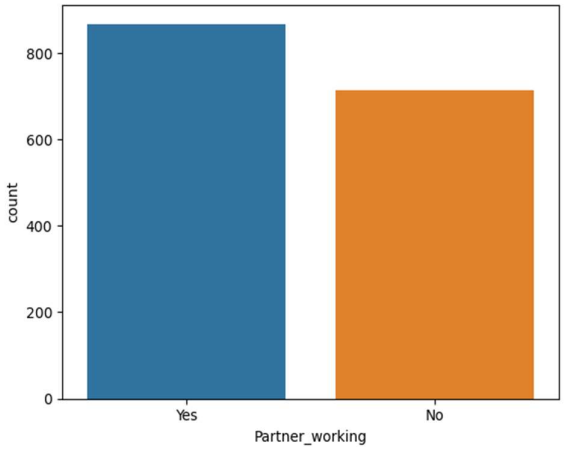


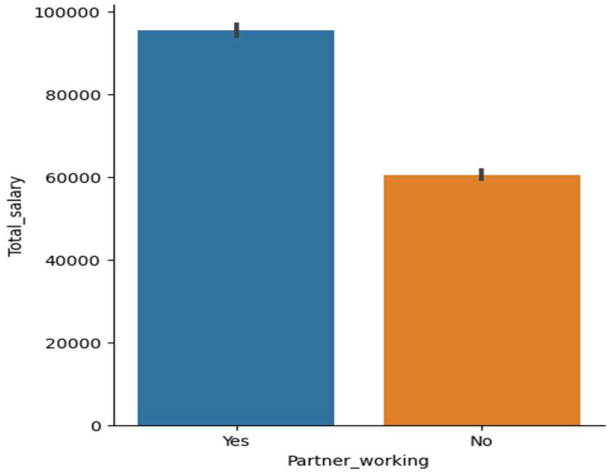
Q6) How does having a working partner influence the purchase of higher-priced cars?

Maximum Price, Median of price for partner working and not working is same value, of price for partner working and not working is not having much difference.

```
Partner_working
No     31000.0
Yes    31000.0
Name: Price, dtype: float64
```

Median Partner_working & Price


800
600
400
200
0
count
Yes
No
Partner_working

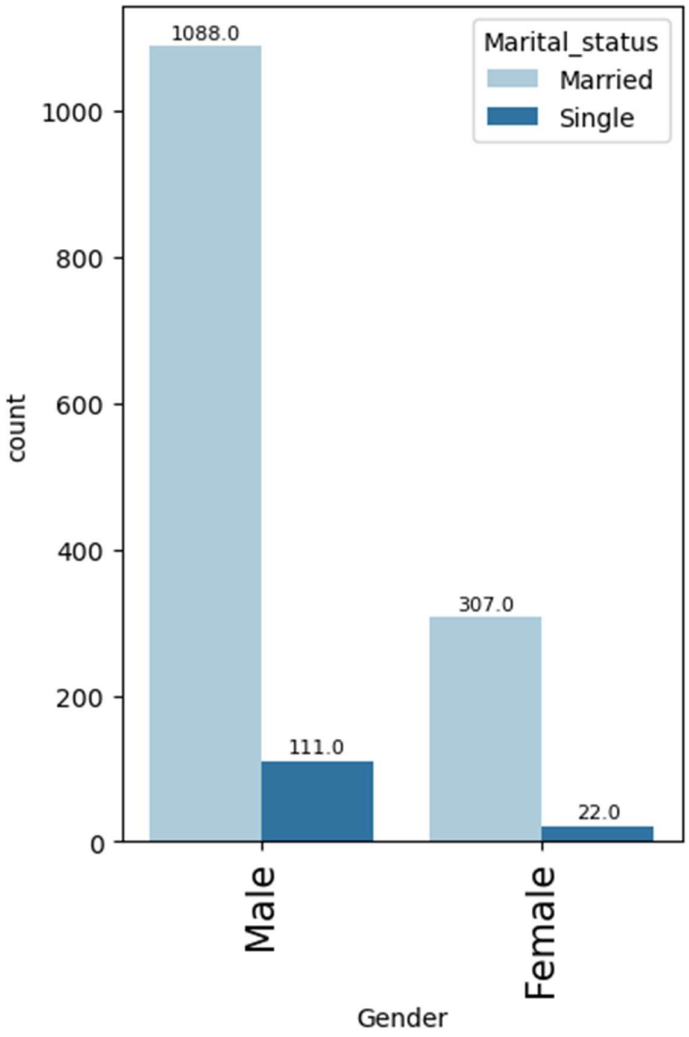

100000
80000
60000
40000
20000
0
Total_salary
Yes
No
Partner_working

E. Actionable Insights & Recommendations

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.

Both in Male and Female Gender the maximum number of cars are purchased by the Married people.

```
Gender  Marital_status
Female  Married            307
        Single              22
Male    Married            1088
        Single              111
Name: count, dtype: int64
```



Problem 2

A bank can generate revenue in a variety of ways, such as charging interest, transaction fees and financial advice. Interest charged on the capital that the bank lends out to customers has historically been the most significant method of revenue generation. The bank earns profits from the difference between the interest rates it pays on deposits and other sources of funds, and the interest rates it charges on the loans it gives out.
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 also cross-sells asset products to its existing customers through personal loans, auto loans, business loans, etc., and to do so they use various communication methods including cold calling, e-mails, recommendations on the net banking, mobile banking, etc.
GODIGT Bank also has a set of customers who were given credit cards based on risk policy and customer category class but due to huge competition in the credit card market, the bank is observing high attrition in credit card spending. 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. (Higher intent means consumers would want to use the card and hence not be attrite.)

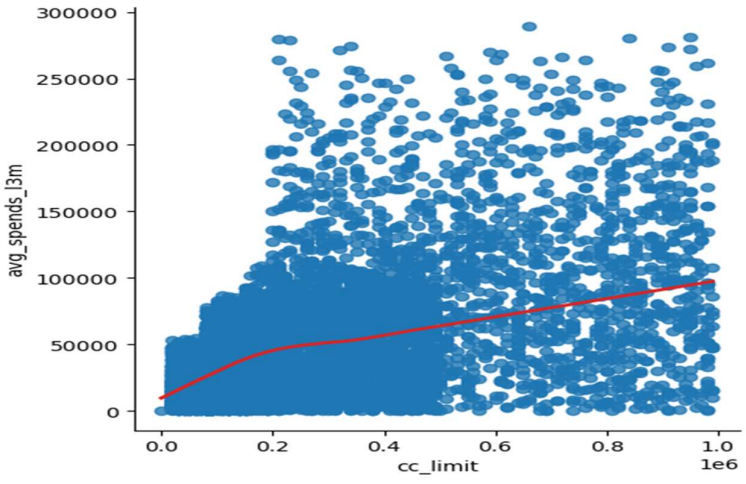
Problem 2 Question: (Analyze the dataset and list down the top 5 important variables, along with the business justifications. (10 Points) Data Dictionary - Link)

#ANS
Top 5 important variables are and justifications are as given below.

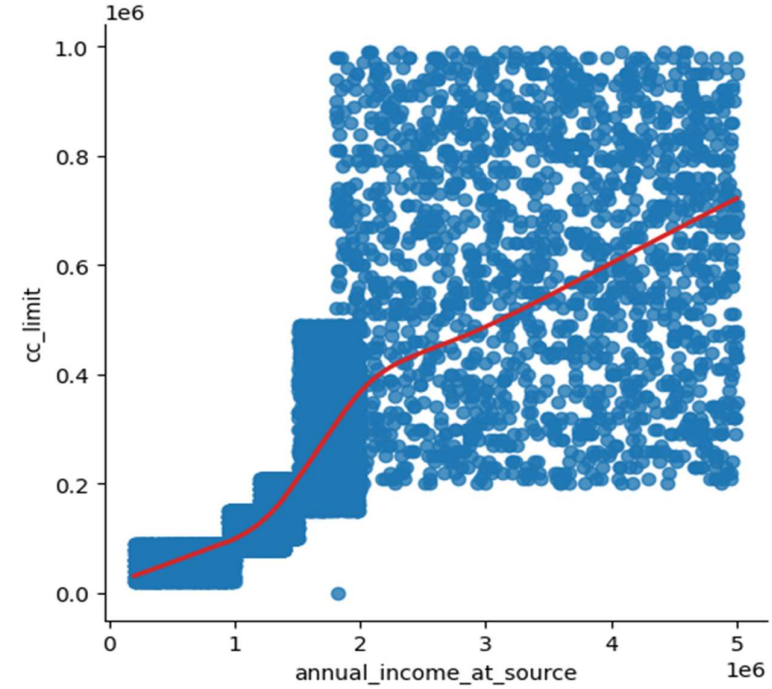
- **annual_income_at_source**
- **avg_spends_l3m**
- **cc_limit**
- **Occupation_at_source**
- **card_type**

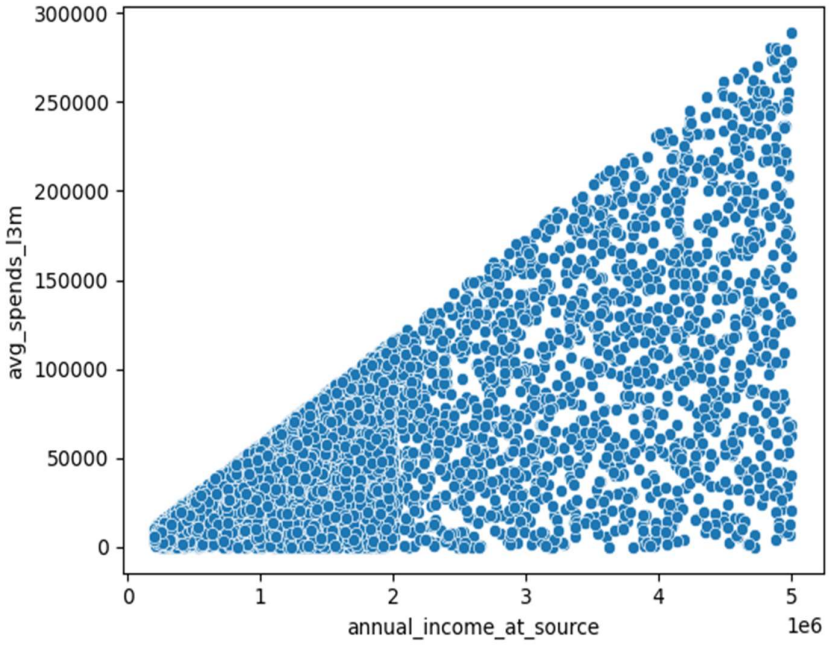
The bank will make a profit only through the customers that show higher intent towards a recommended credit card (Higher intent means consumers would want to use the card and hence not be attrite)
Higher intent will have more average spends in last 3 months (avg_spends_l3m)

To have more average appends in last 3 months user should have more 'cc_limit'



There is positive co-relation between 'annual_income_at_source' and 'cc_limit' therefore the user having more 'annual_income_at_source' will spend more and hence not be attrite

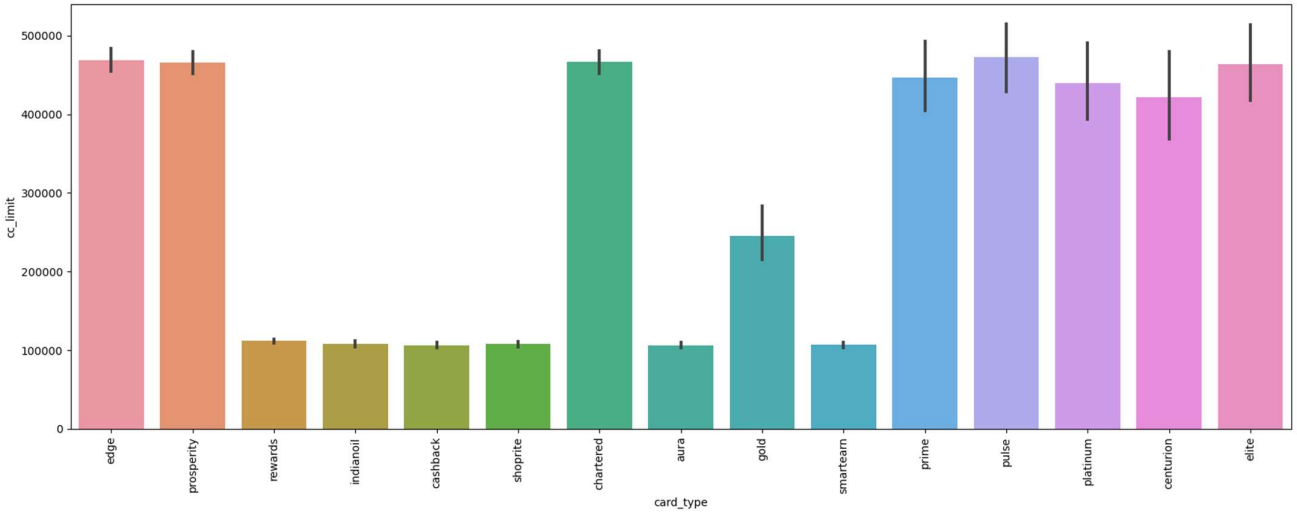




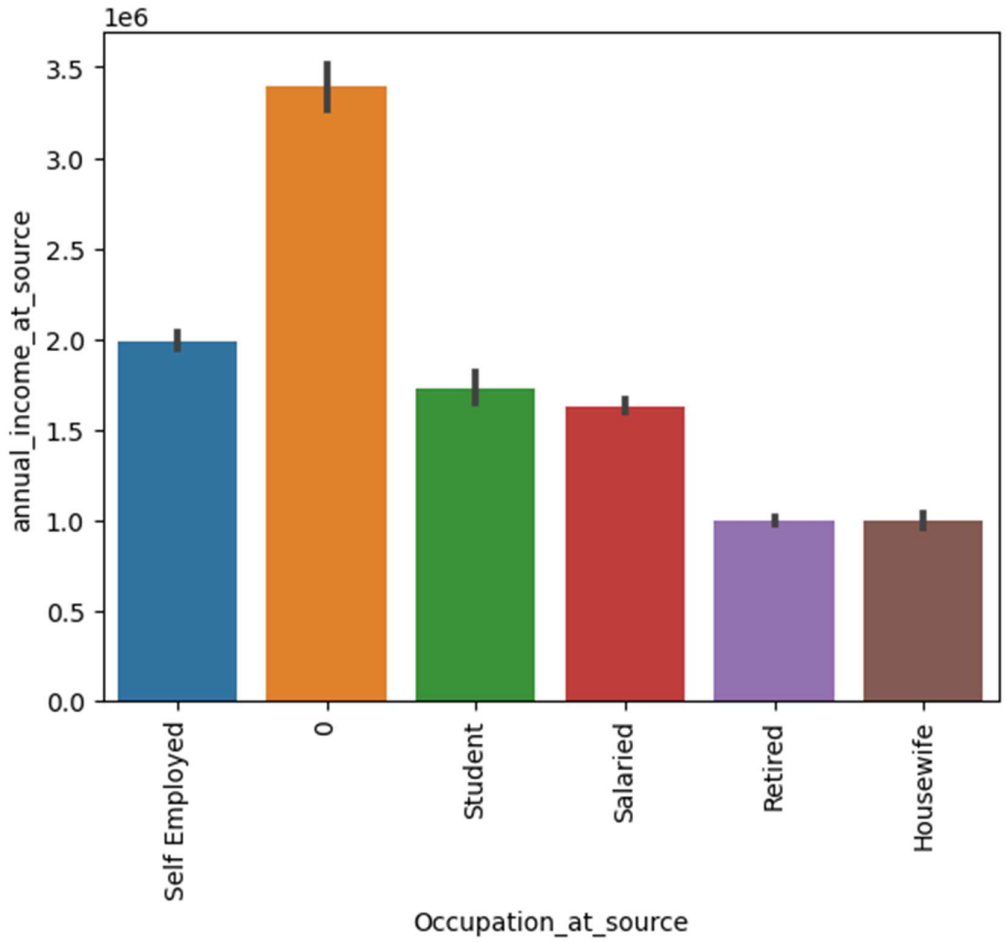
Top 4 card types having more 'cc_limit' are

- **Chartered**
- **prosperity**
- **edge**
- **rewards**

More 'annual_income_at_source' is there for the 'Salaried' and 'Self Employed' occupation

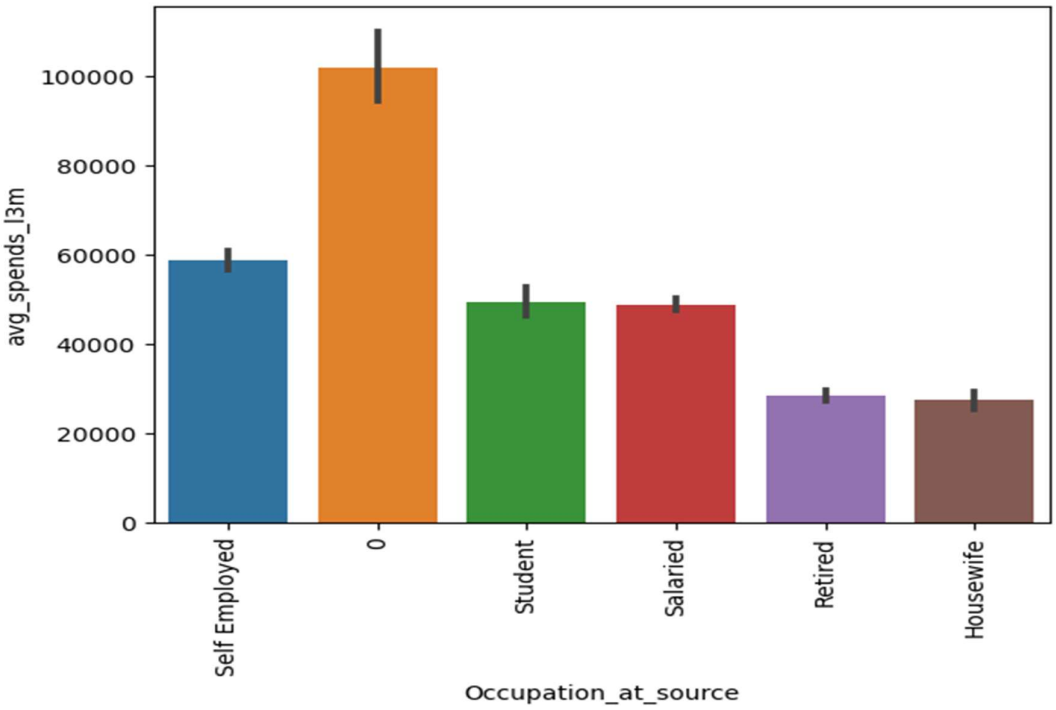


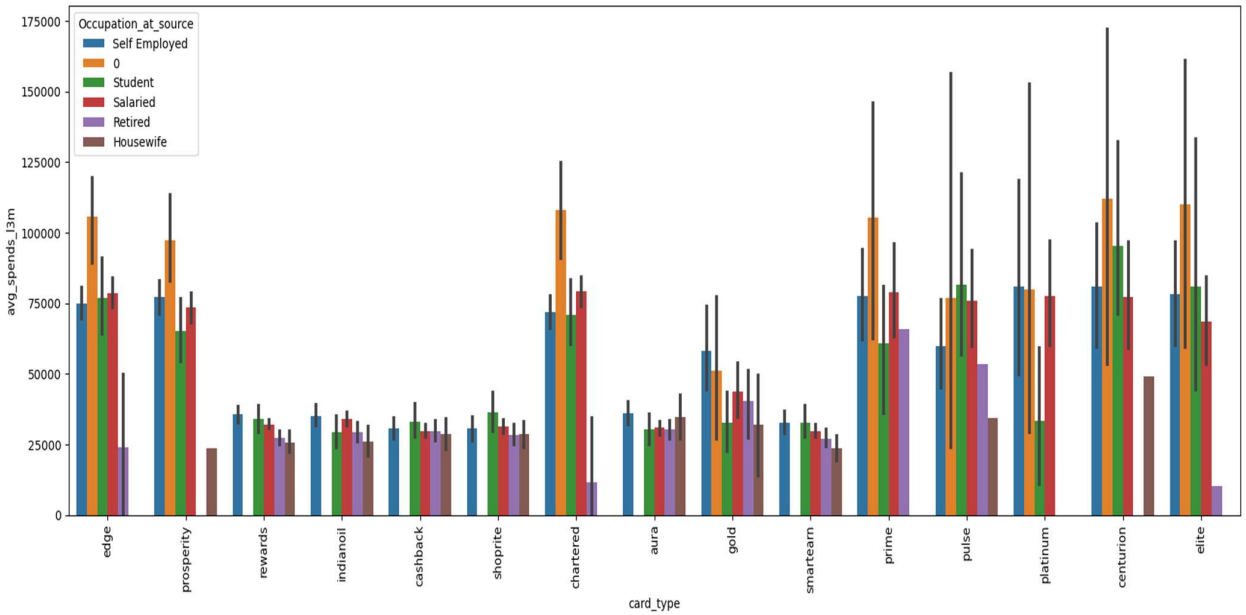

```
Occupation_at_source
0                885682388
Housewife        381504991
Retired         1085732289
Salaried        6394329966
Self Employed   4328466549
Student         1071260172
Name: annual_income_at_source, dtype: int64
```

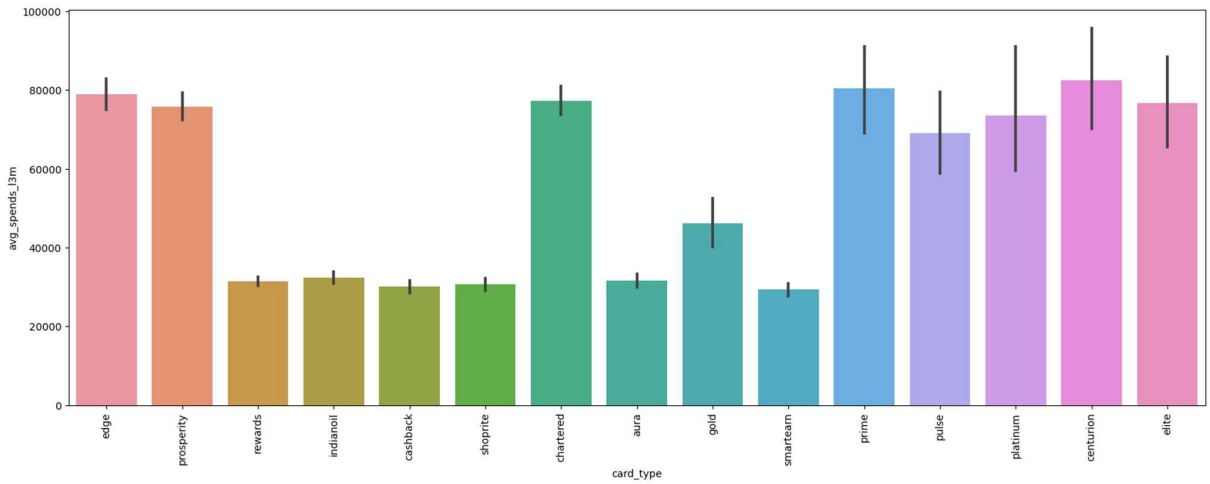


The cards having higher cc_limit should be given to the users having 'Salaried' and 'Self Employed' occupation as more number of users are from these occupation, so that there is increase in avg_spends_l3m and and made sure that the card given to the customer is the right credit card.

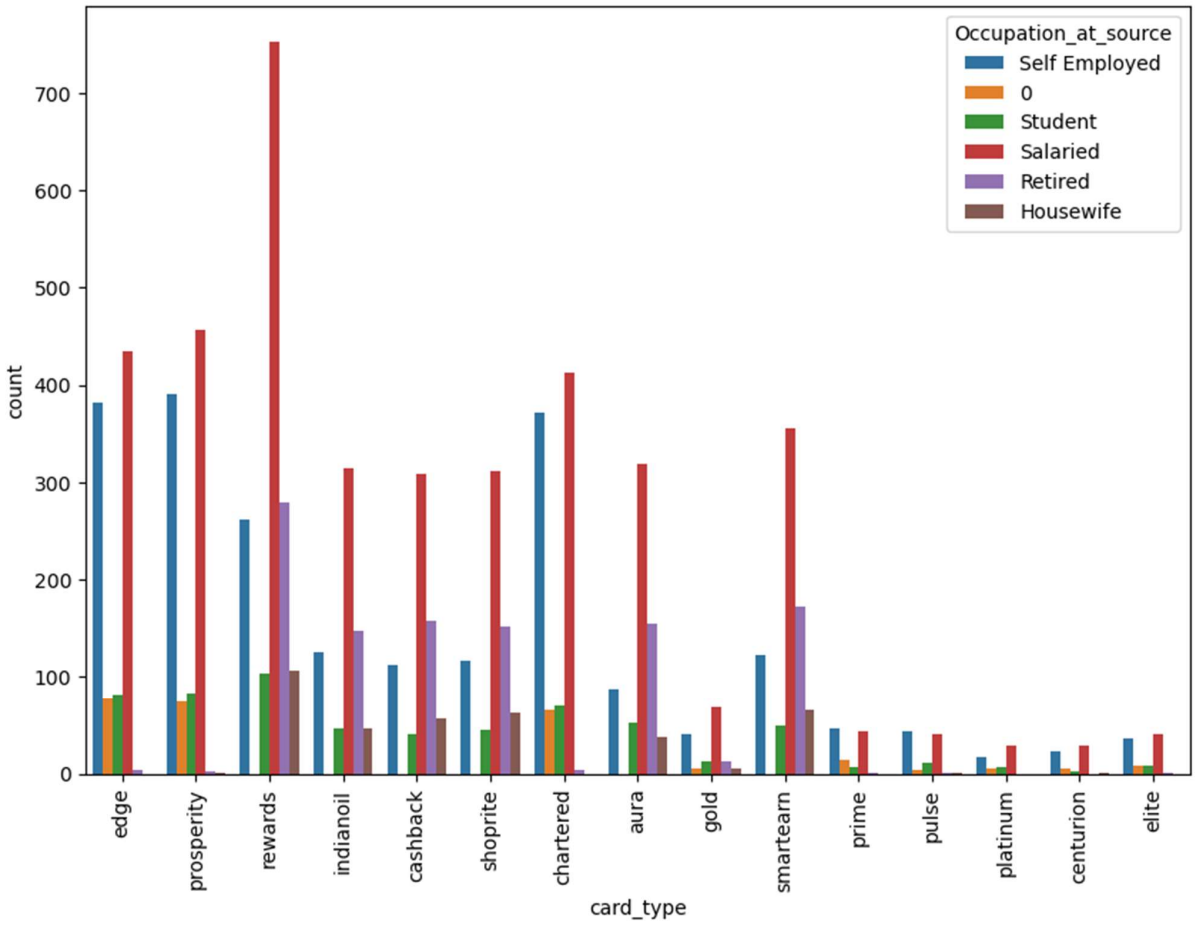
```
Occupation_at_source
0                 26636048
Housewife         10531574
Retired           31118023
Salaried         191554231
Self Employed    127884163
Student           30683145
Name: avg_spends_l3m, dtype: int64
```





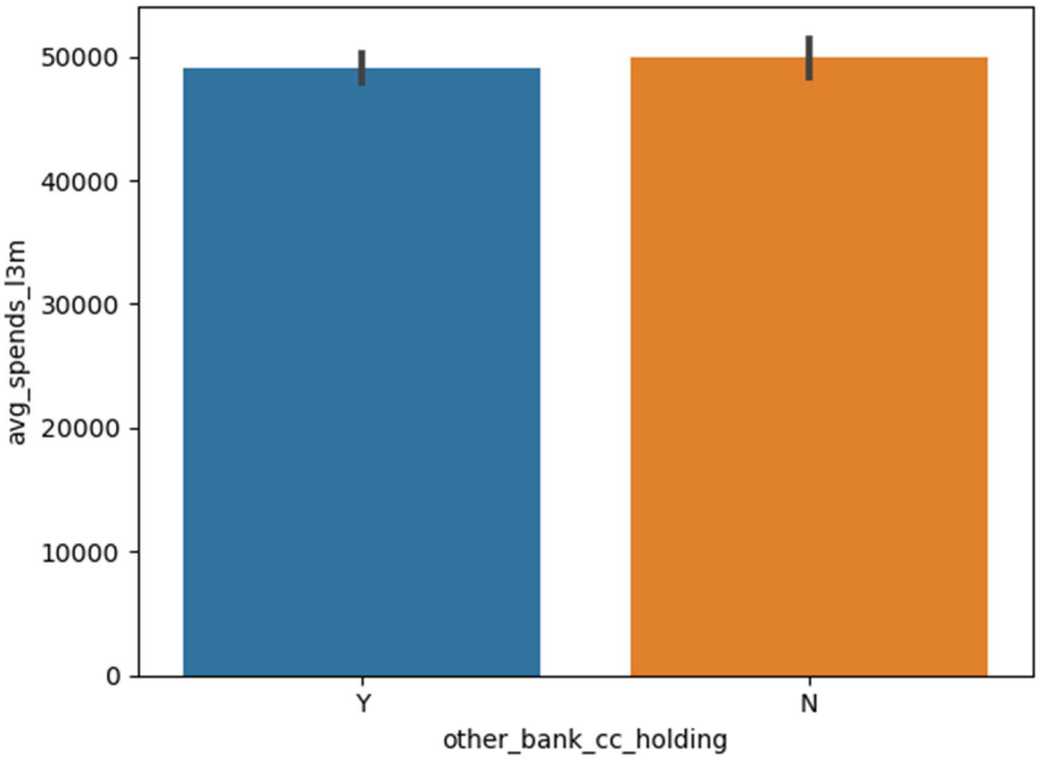


Based on the above analysis Bank must to revisit its credit card policy for the user having occupation other than Salaries and Self Employed. So that the user from other occupations also increase the spends with the revise credit card policies and thus bank will make a profit



Users having credit card from other banks are spending more and there is increase in avg_spends_l3m therefore bank can target the other users having credit card from other banks by providing attractive credit card policies

```
other_bank_cc_holding
N    186135917
Y    232271267
Name: avg_spends_l3m, dtype: int64
```



From the below plot we see that 75% of customers are utilizing the credit upto the specified limit (350000), but few customer are having a credit card limit upto 1000000, So attract new customer by proving the credit limit of 350000 so that avg_spends_l3m will increase and bank will profit