Business Report

Objective:

We will put forth the insights that can help the marketing team at Austin Motors make their marketing campaigns more targeted and effective.

Data Information

Let's have a look at features of the customers in the dataset who already have Hatchback, Sedan or SUV.

Variables	Description
Age	Age of the customer
Gender	Gender of the customer (Male or Female)
Profession	Working Profession (Business or salaried)
Marital_status	Customer is married or not
Education	Customer's educational qualification (Graduate or Post-graduate)
No_of_Dependents	Number of family members he/she is taking care of
Personal_loan	Does he has a personal loan (Yes/No)
House_loan	Does he has a House loan (Yes/No)
Partner_working	Is his/her partner working professionally
Salary	Customer's salary either from a job or a business
Partner_salary	Partner's salary
Total_salary	Total Salary (Customers salary + his/her partner's salary)
Price	Price of the vehicle customer owns.
Make	Type of vehicle they own (Hatchback, Sedan or SUV)

We have 14 variables or features that are associated with the customer, who has already bought Hatchback, Sedan or SUV.

We have a record of 1581 such customers available with us in this Dataset.

Let's take a look at what type of features/variable are these -

Categorical Variables (Binary)

Gender

Profession

Marital_status

Education

Personal_loan

House_loan

Partner_working

We have **7 features** where we have binary inputs in the features-
That is - in **Personal_loan** we have - **YES/NO** and in **Gender** we have – **MALE/ FEMALE** etc.

Categorical Variables (Multi-level)

Make

No_of_Dependents

We have **2 features** where we have multiple unique inputs in the features-
That is - in **Make** we have – **HATCHBACK / SEDAN / SUV** etc.

Continuous Variables

Salary

Partner_salary

Total_salary

Price

Age

We have **5 continuous variables/features**

Insights from preliminary analysis of the variables

In initial analysis we found that there are values that are blank in some variables.

These variables are Partner_salary and Gender.

We found that **Partner_salary** has **106 blanks** and **Gender** variable has **53 blank** spaces.

There are 1581 records for each variable apart from **Gender** and **Partner_salary**

Age	1581
Gender	1528
Profession	1581
Marital_status	1581
Education	1581
No_of_Dependents	1581
Personal_loan	1581
House_loan	1581
Partner_working	1581
Salary	1581
Partner_salary	1475
Total_salary	1581
Price	1581
Make	1581

Age	0
Gender	53
Profession	0
Marital_status	0
Education	0
No_of_Dependents	0
Personal_loan	0
House_loan	0
Partner_working	0
Salary	0
Partner_salary	106
Total_salary	0
Price	0
Make	0

We will fill out these blanks values in data with a proper logic-

According to our observation, logic that we have utilized here is that the Partner_salary is the difference between Total_salary and Salary variables in given data.

So we have filled these 106 blanks in Partner_salary with this calculated difference.

We found 2 data points miss-spelled in Gender variable. They are corrected as Female.

Male	1199
Female	327
Femal	1
Femle	1

The Gender column consists of 78.46% Males and 21.53% Females.

Male	0.784686
Female	0.215314

As per the percentage there are more chances that most of the blank values in the Gender variable would be Male.

By this logic we have replaced all the blank spaces in Gender variable as Male.

With this replacement we now have only 20.80% Female customers in data, rest of the customers are Male.

```
Male      0.791904
Female    0.208096
```

Further Insights-Checking out each variable separately-

	Age	Salary	Partner_salary	Total_salary	Price
count	1581.000000	1581.000000	1581.000000	1581.000000	1581.000000
mean	31.922201	60392.220114	19233.776091	79625.996205	35597.722960
std	8.425978	14674.825044	19670.391171	25545.857768	13633.636545
min	22.000000	30000.000000	0.000000	30000.000000	18000.000000
25%	25.000000	51900.000000	0.000000	60500.000000	25000.000000
50%	29.000000	59500.000000	25100.000000	78000.000000	31000.000000
75%	38.000000	71800.000000	38100.000000	95900.000000	47000.000000
max	54.000000	99300.000000	80500.000000	171000.000000	70000.000000

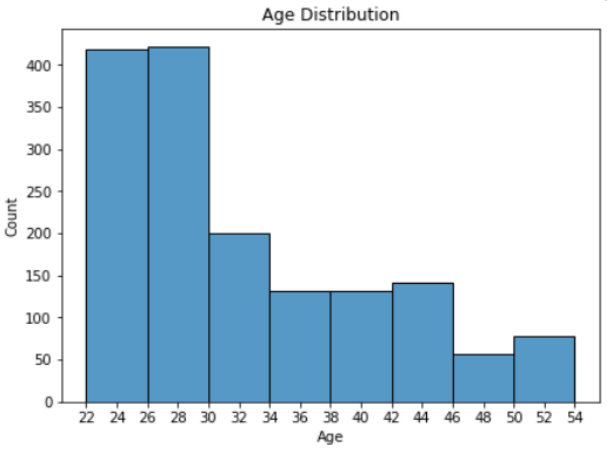
50% of our customers are below 29 years of Age.

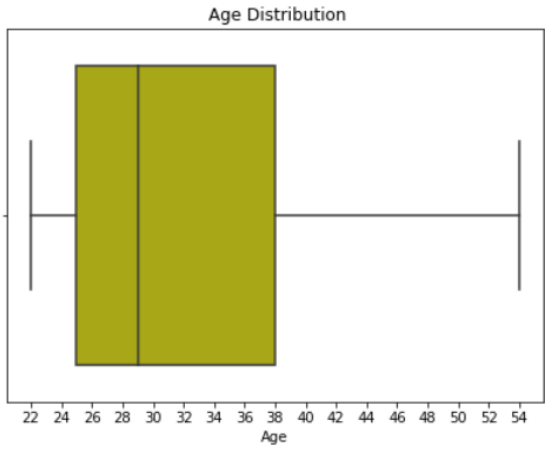
75% of the customers who are our potential buyers do have a working partner.

Total salary of 50% of our customers is below 78000/month

Price range of 50% of the vehicles that gets sold is below $31000

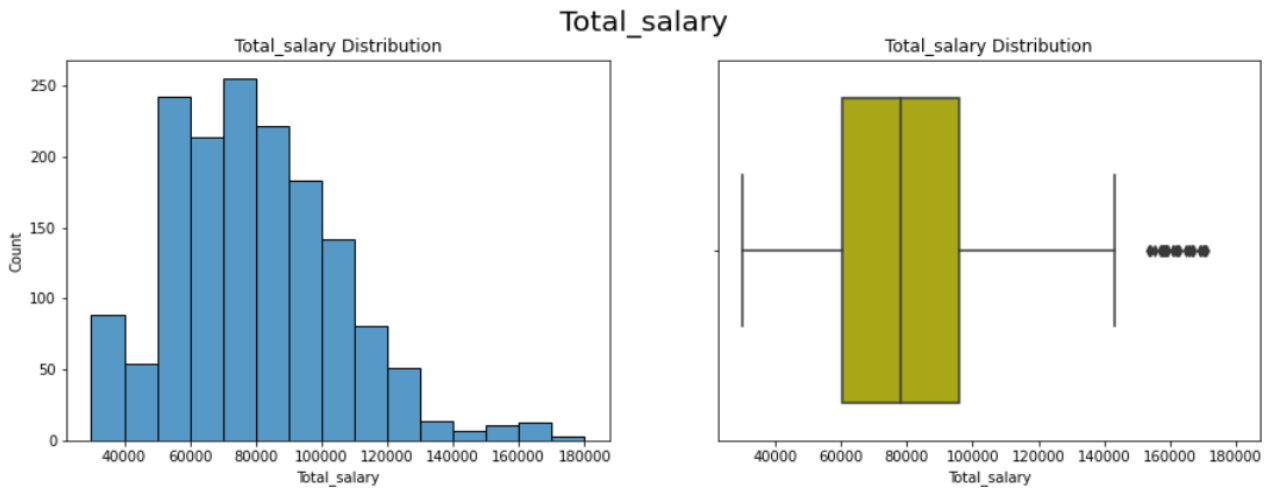
AGE





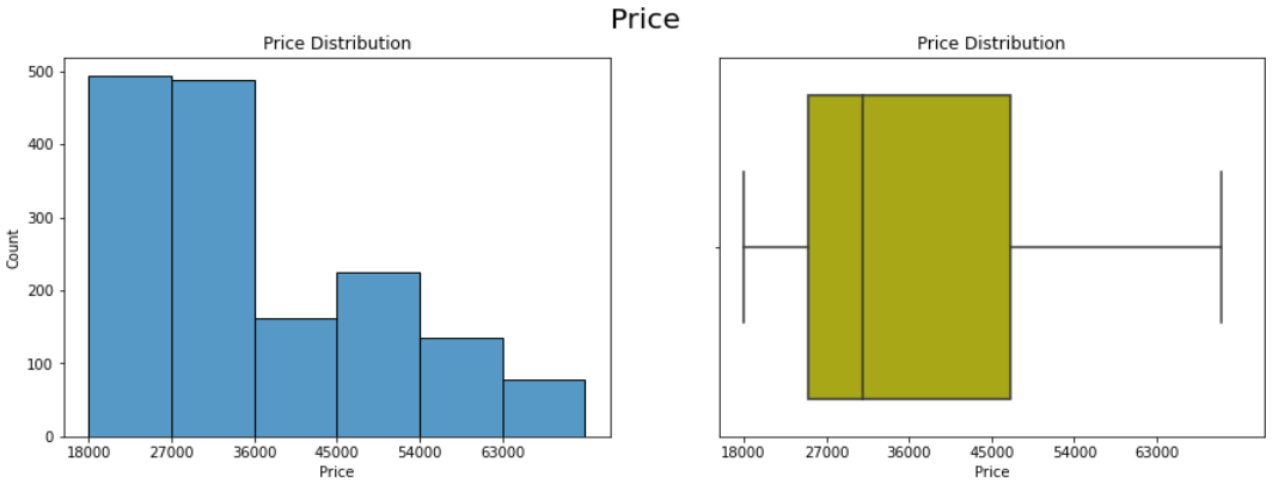
Most of the customers lie in age range 22 to 30.

Almost 50% of customers lie in range 22 to 29.



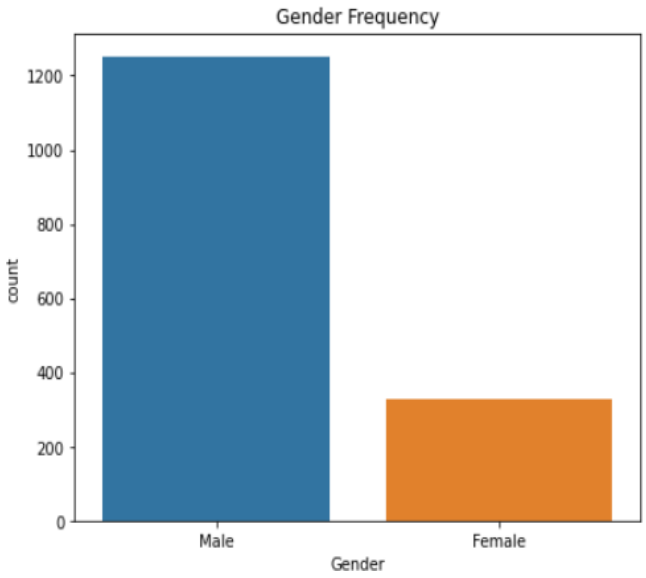
Total salary of almost 50% of our customers lies between 60K to 100K

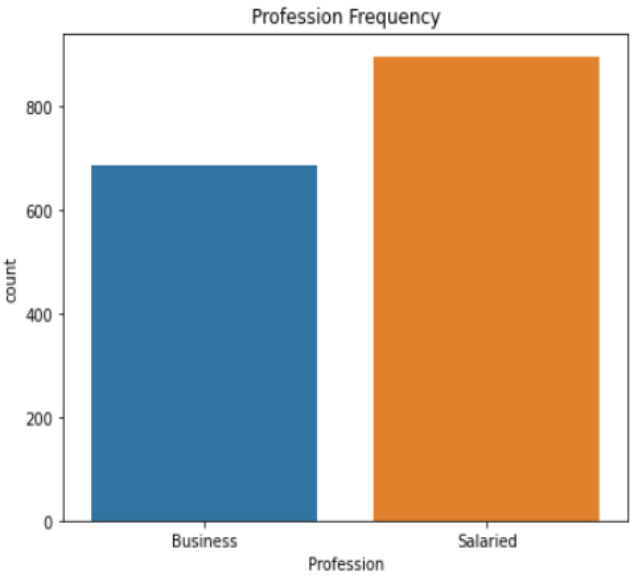
There are some extreme values in Total salary, but they are just due to high partner salaries.



Most of our customers buy cars in range of 18K to 36k.

Almost 50% of our customers buy cars in range of 18k to 31k



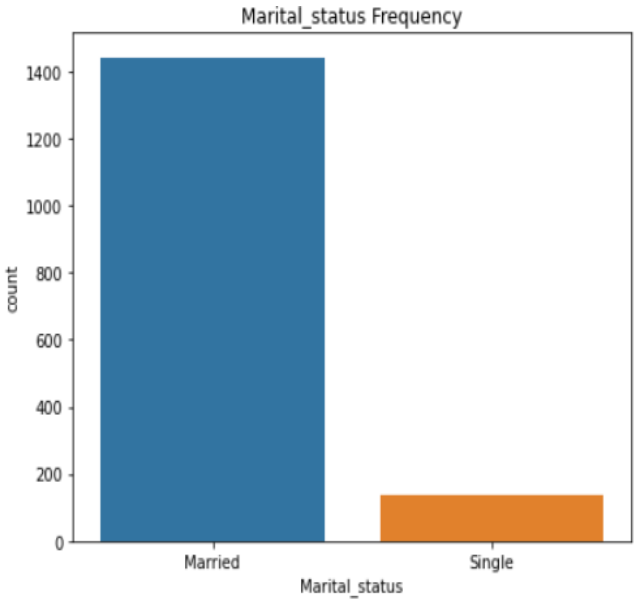


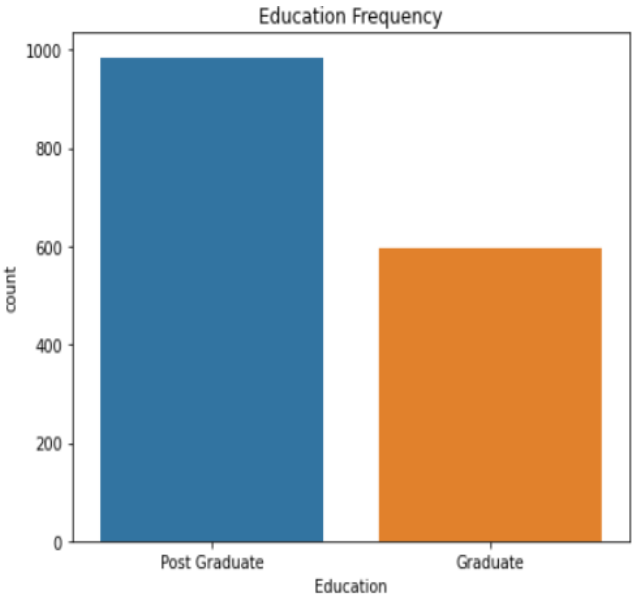

```
Male      0.784686
Female    0.215314
```

```
Salaried    0.56673
Business    0.43327
```

Percentage of male customers is 78.46% where-as for females it's on a lower side 21.52%.

Salaried people buy cars more often than people who have their own business.



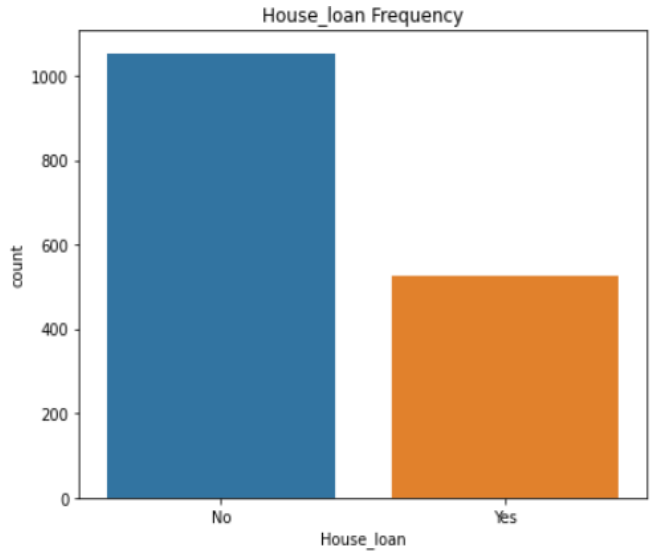


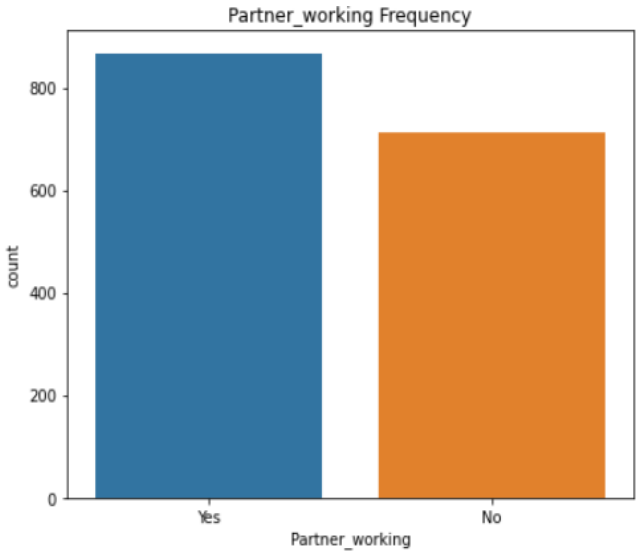

```
Married    0.912713
Single     0.087287
```

```
Post Graduate    0.623023
Graduate         0.376977
```

91.27% of the times, the customer that buys our vehicle is married.

Almost 62.30% people who buy our cars are Post-graduate.



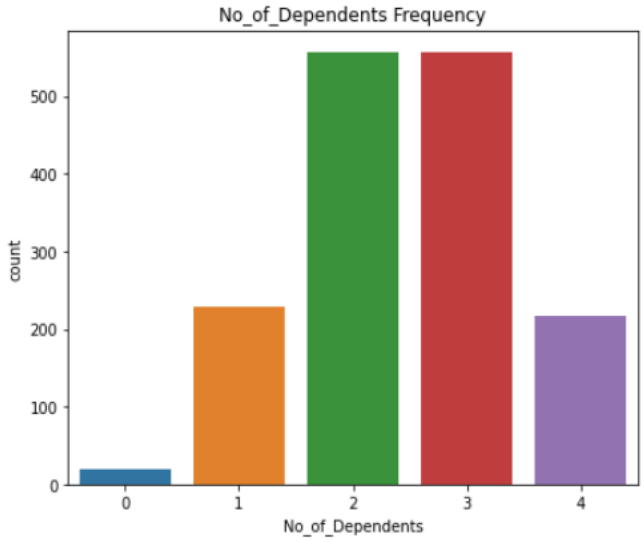


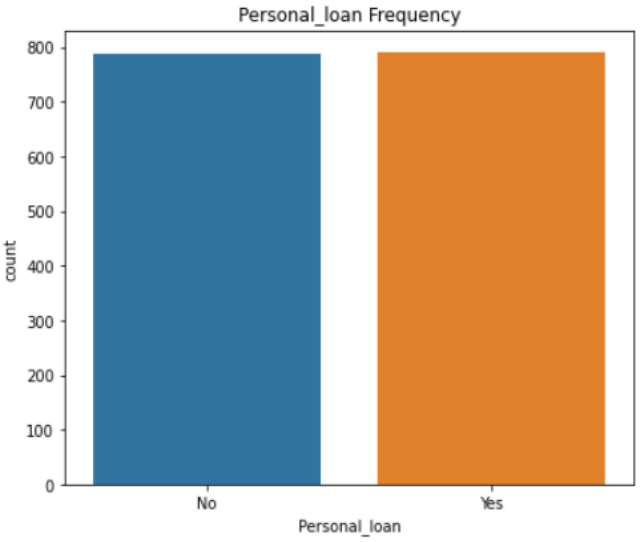

```
No      0.666667
Yes     0.333333
```

```
Yes     0.54902
No      0.45098
```

2/3rd of the customers do not have any kind of house loan, when they buy our vehicle.

Partners of 54.9% of the customers, are employed and working.



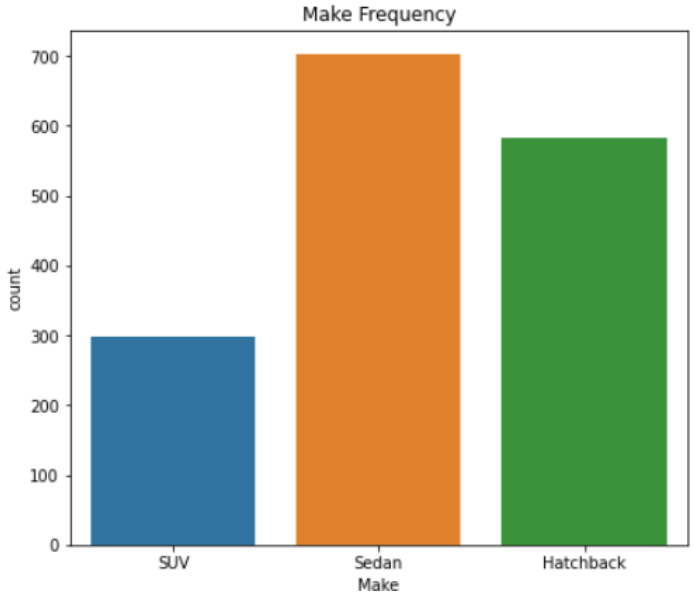



```
3     0.352309
2     0.352309
1     0.144845
4     0.137887
0     0.012650
```

```
Yes     0.500949
No      0.499051
```

Almost 70% of our customers have either 2 or 3 dependents.

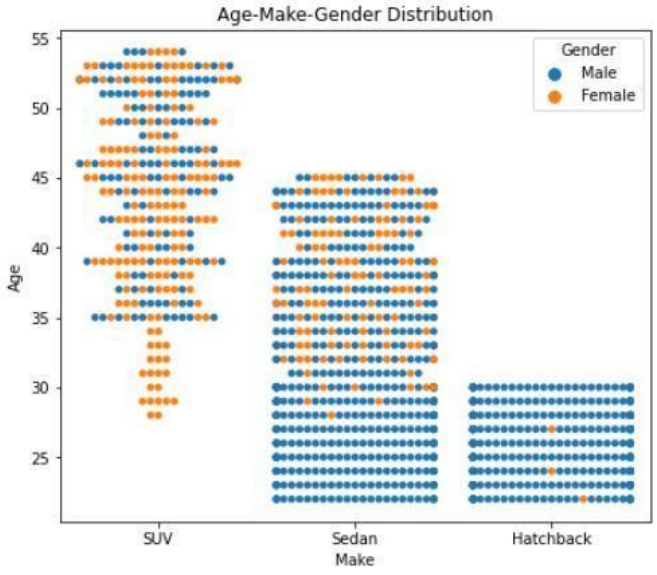
50% of our customers have personal loan and 50% don't have it.




```
Sedan        0.444023
Hatchback    0.368121
SUV          0.187856
```

Almost 44% of all the vehicles that we sell are Sedan's, 36% are Hatchback's and rest 18% are SUV's.

Let's figure out some interesting trends for our business-



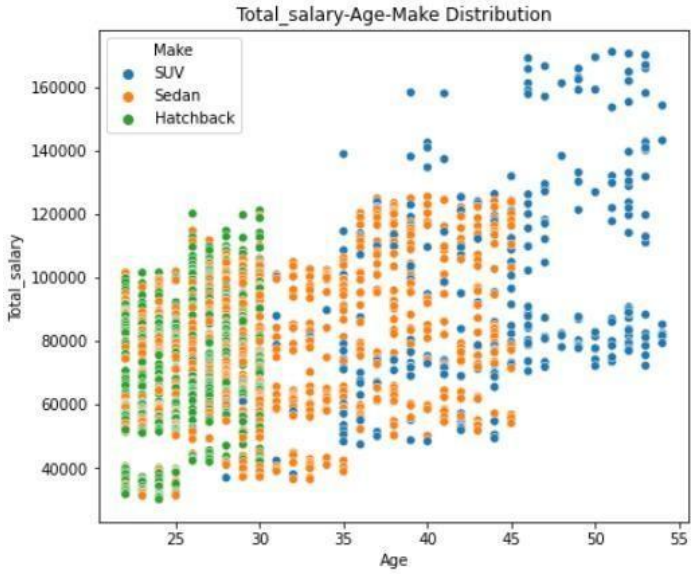
We can clearly see that Male customers below age 30, would only prefer either Sedan or Hatchback.

Males or Females with age above 45 would only prefer to have SUV.

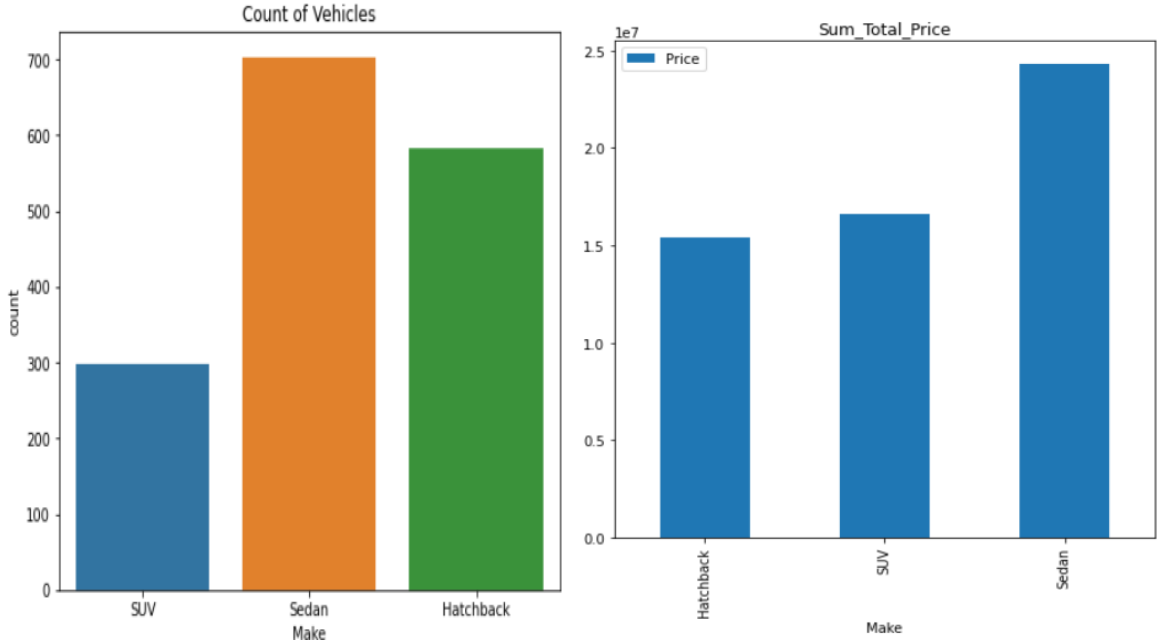
Females with age above 27 would prefer to have either SUV or sedan.

Females don't prefer to go with Hatchback.

Category	Make
M, Age<30	Sedan or Hatchback
M, Age>45	SUV
M, 45>Age>30	SUV, Sedan
F, 45>Age>27	SUV, Sedan
F, Age>45	SUV



Customer is mostly only interested in SUV when his Total Salary exceeds 120000.

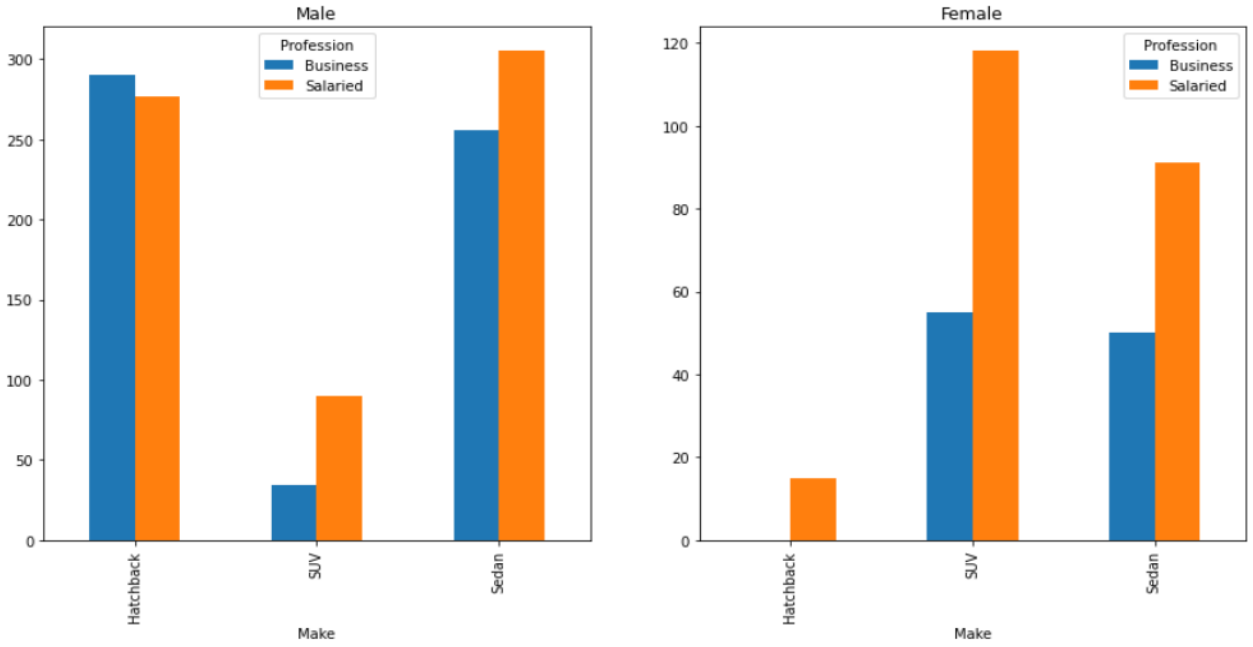


Sedan is the most sold vehicle type and it generates maximum money inflow.

After Sedan, SUV is the segment that generates greater money inflow than Hatchback.

Although the number of SUV sold are less in number than Hatchback, still it generates greater money inflow.

Let's check some of Assumptions if they are True or False-

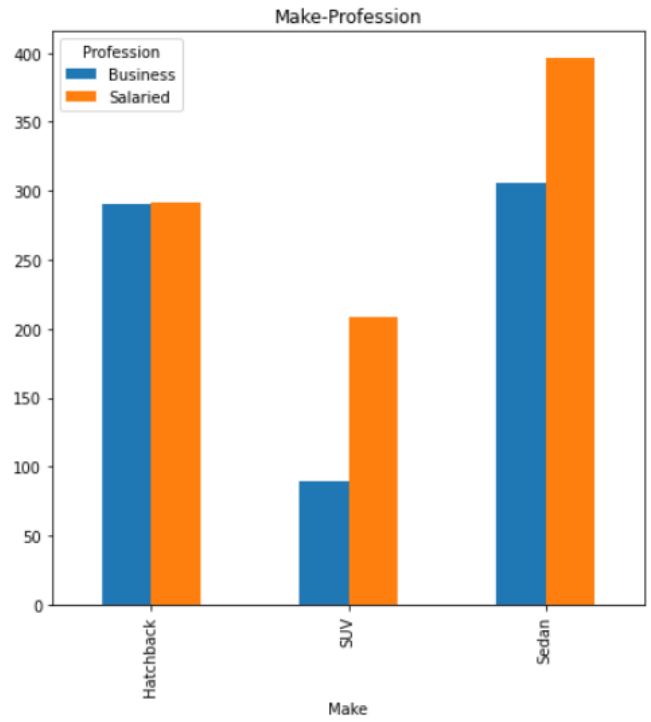


It could clearly be seen that Women prefer SUV by a large margin, compared to the Men

Assumption that Men prefer SUV by a large margin, compared to the women is False

It could clearly be seen that Salaried male prefer Sedan by a large margin, compared to the SUV

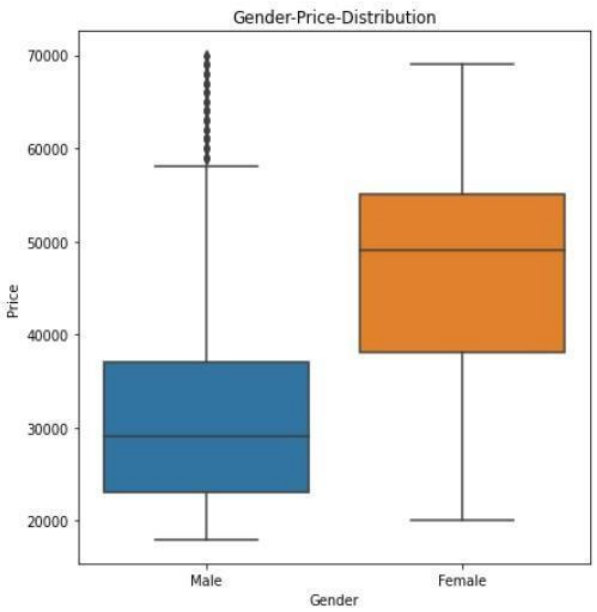
Assumption that salaried male is an easier target for a SUV sale over a Sedan Sale is False.

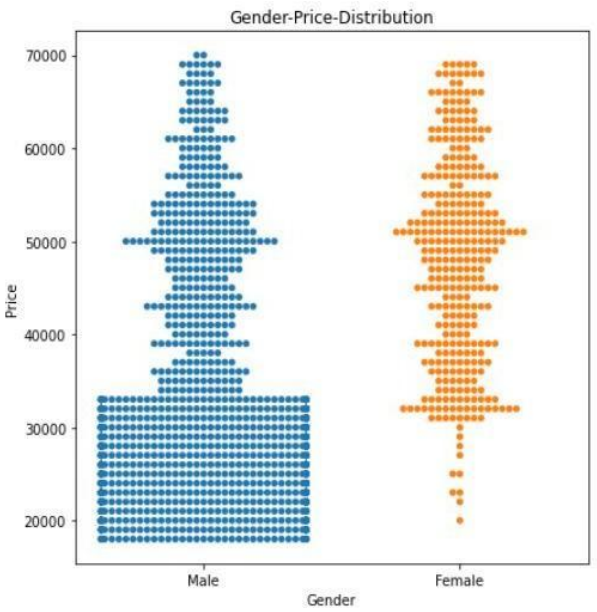


It could clearly be seen that salaried person is more likely to buy a Sedan over SUV and Hatchback.

Assumption that salaried person is more likely to buy a Sedan is True

Checking if Gender has any impact on price distribution





Gender	Female	Male
count	329.000000	1252.000000
mean	47705.167173	32416.134185
std	11244.836378	12366.253107
min	20000.000000	18000.000000
25%	38000.000000	23000.000000
50%	49000.000000	29000.000000
75%	55000.000000	37000.000000
max	69000.000000	70000.000000

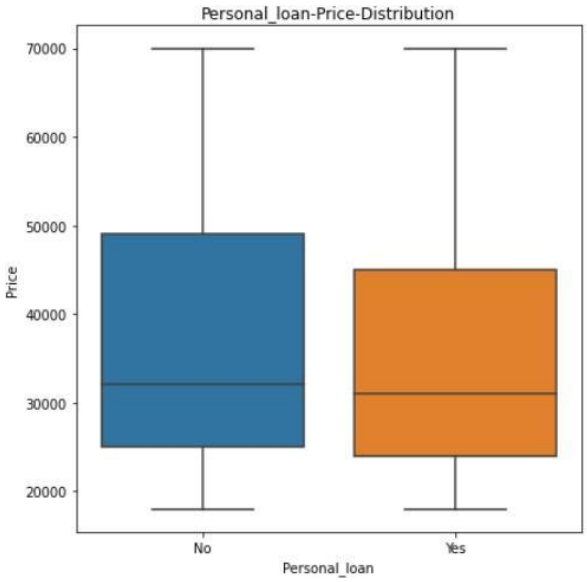
It could easily be inferred that 75% of the times when female customers spends money it's above 38K.

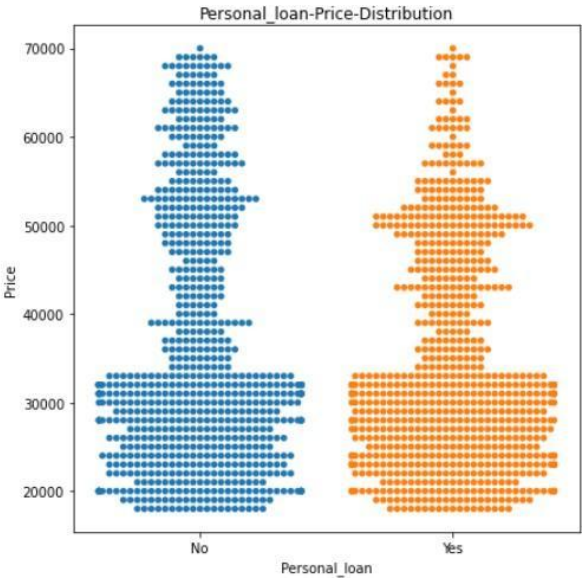
It could also be inferred that 75% of the times when male customers spends money it's below 38K.

From business point of view, female customer is more likely to buy a high priced vehicle then male customer.

Female customer is likely to spend more than male customer.

Checking if Personal loan has any impact on price distribution



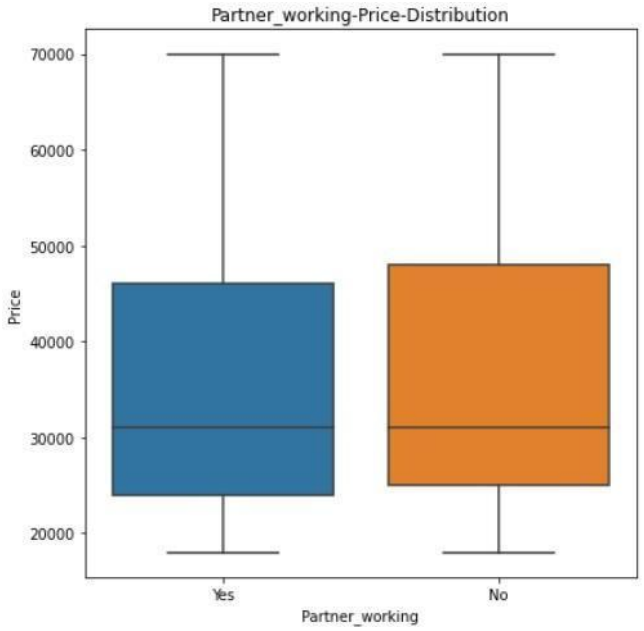


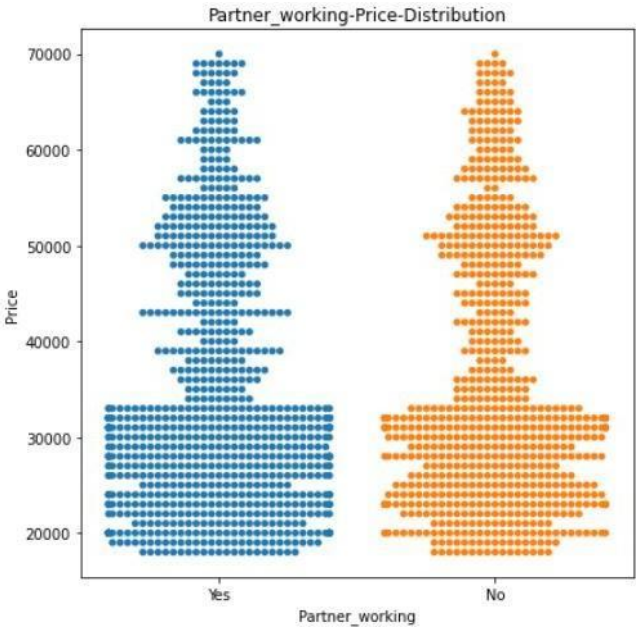
Personal_loan	No	Yes
count	789.000000	792.000000
mean	36742.712294	34457.070707
std	14534.344526	12578.780338
min	18000.000000	18000.000000
25%	25000.000000	24000.000000
50%	32000.000000	31000.000000
75%	49000.000000	45000.000000
max	70000.000000	70000.000000

Customers with no personal loan spend a bit more than person with personal loan.

There is not a huge difference in their spending power.

If having a working partner leads to the purchase of a higher-priced car?

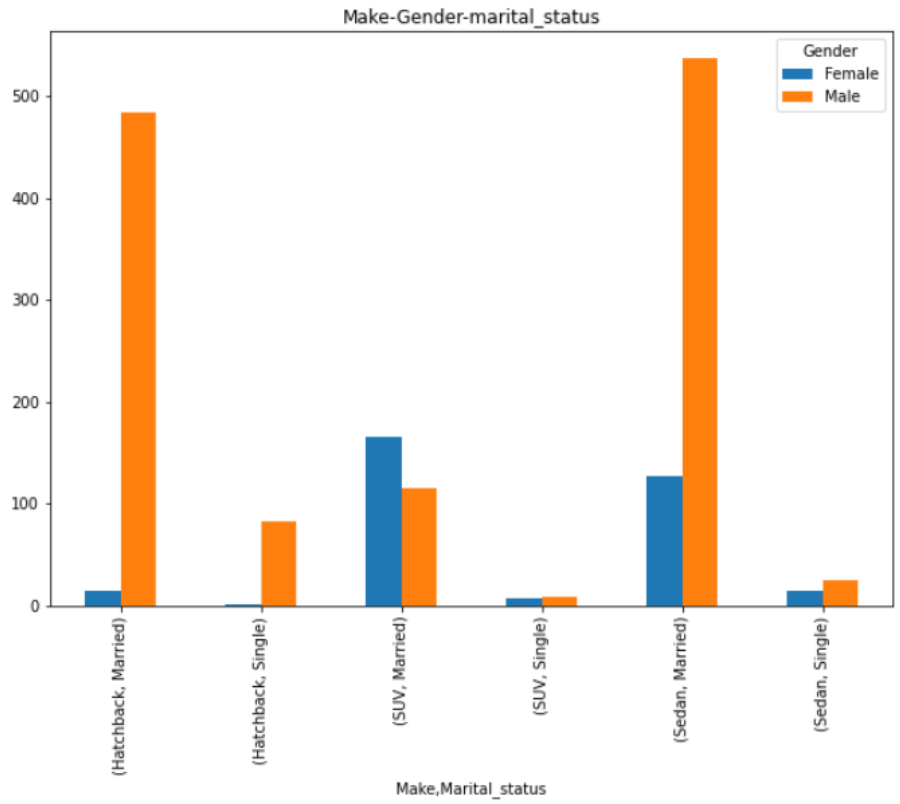




Partner_working	No	Yes
count	713.000000	868.000000
mean	36000.000000	35267.281106
std	13817.734086	13479.532555
min	18000.000000	18000.000000
25%	25000.000000	24000.000000
50%	31000.000000	31000.000000
75%	48000.000000	46000.000000
max	70000.000000	70000.000000

Partner working or not has no significant impact on how much a customer spends.

Checking if marital status and Gender can give us some common purchase pattern



Make	Marital_status	Female	Male
Hatchback	Married	14	484
	Single	1	83
SUV	Married	166	115
	Single	7	9
Sedan	Married	127	537
	Single	14	24

Single male would either buy a Hatchback or Sedan

Married man would prefer buying Sedan and Hatchback

Single female would prefer buying Sedan or SUV

Married female would buy SUV or sedan

Category	Most Preferred Vehicle
Single Male	Hatchback , Sedan
Married Male	Sedan, Hatchback

Single Female	Sedan, SUV
Married Female	SUV, Sedan