Austo Motor Company

Business report

Summary

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. The business report analyzes customer data from Austo Motor Company with the aim of enhancing their marketing campaign. The analysis of the automobile dataset reveals that the majority of customers are male, hold postgraduate degrees, are employed in salaried positions, are married, and show a preference for sedan cars. Evaluations of employees' assumptions about customer preferences indicate that two assumptions are accurate while one is incorrect. Insights into customer spending patterns based on gender, income, and other factors highlight that male customers tend to favour lower-priced cars. Based on these findings, the report recommends focusing the marketing efforts on single males as a strategic approach to boost sales and increase profitability.

Content

1. The important Technical information and attributes of the dataset which are observed, like what is the size of the data, nature of the variables, missing values, duplicate value, outliers and other irregularities which might be present in the data.

2. Treating the Missing values and Irregularities by having a closer look into the Data with the statistical methods.
    a. Missing Values.
    b. Irregularities.

3. Exploring the data with appropriate visuals and visualising the data with Statistical plots.
    a. Plotting with Categorical Values.
    b. Plotting with Numerical Values.

4. Exploring the relationship between the data with appropriate visuals and visualising the data with Statistical plots.

5. Here are some Key Business Question to which we try find the solution by using visualisation and EDA methods.
    a. Do men tend to prefer SUVs more compared to women?
    b. What is the likelihood of a salaried person buying a Sedan?
    c. 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?
    d. How does the amount spent on purchasing automobiles vary by gender?
    e. How much money was spent on purchasing automobiles by individuals who took a personal loan?
    f. How does having a working partner influence the purchase of higher-priced cars?

6. Actionable Insights & Recommendations

Problem:

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.

1. The important Technical information and attributes of the dataset which are observed, like what is the size of the data, nature of the variables, missing values, duplicate value, outliers and other irregularities which might be present in the data.

- The Name of the dataset is “austo_automobile.csv”.
- The Dataset contain a total of 1581 rows and 14 columns.
- There are no Duplicates present in the Data.
- There are some missing values present in “Gender” (53 missing values) and “Partner_salary” (106 missing values).
- There some irregularities found in the “Gender” column with values having spelling mistake as Femal and Femle.
- There are 9 Categorical Columns: Gender, Profession, Marital_status, Education, No_of_Dependents, Personal_loan, House_loan, Partner_working, Make.
- And there are 5 Numerical column: Age, Salary, Partner_salary, Total_salary, Price.
- The Datatypes present are :
    - Int64.
    - Object.
    - Float64.

2. Treating the Missing values and Irregularities by having a closer look into the Data with the statistical methods.

Missing values:

- From the above observation, there are some missing values present in the “Gender” and “Partner_salary” columns.
- Let’s treat the “Gender” column first, the “Gender” column has 53 Missing values which is only 3 percent of the total 1581 rows so we are treating the missing values with the Mode value of the column. The Mode value of “Gender” column is Male so we are filling the missing values with Male.
- For the Partner_salary column, it has 106 missing rows. The partner salary can be found by subtracting the Salary column from Total_salary.

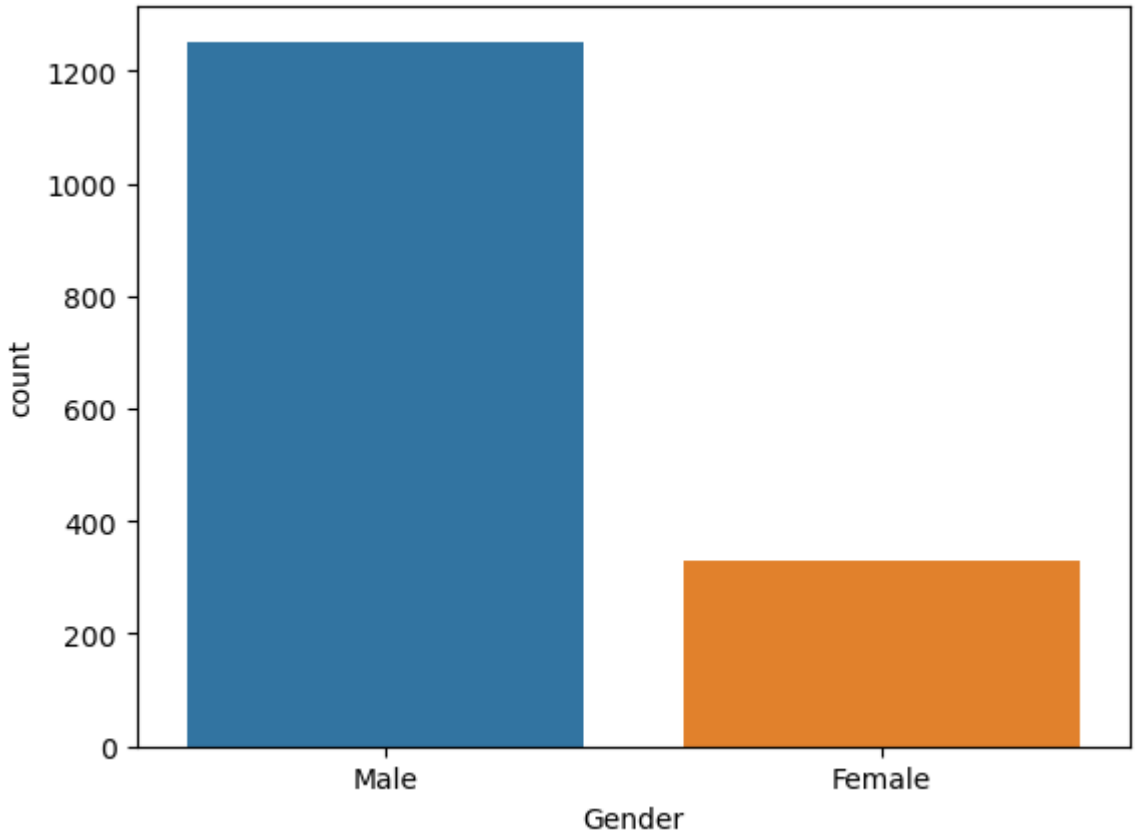
Total_salary = Salary + Partner_salary

Irregularities:

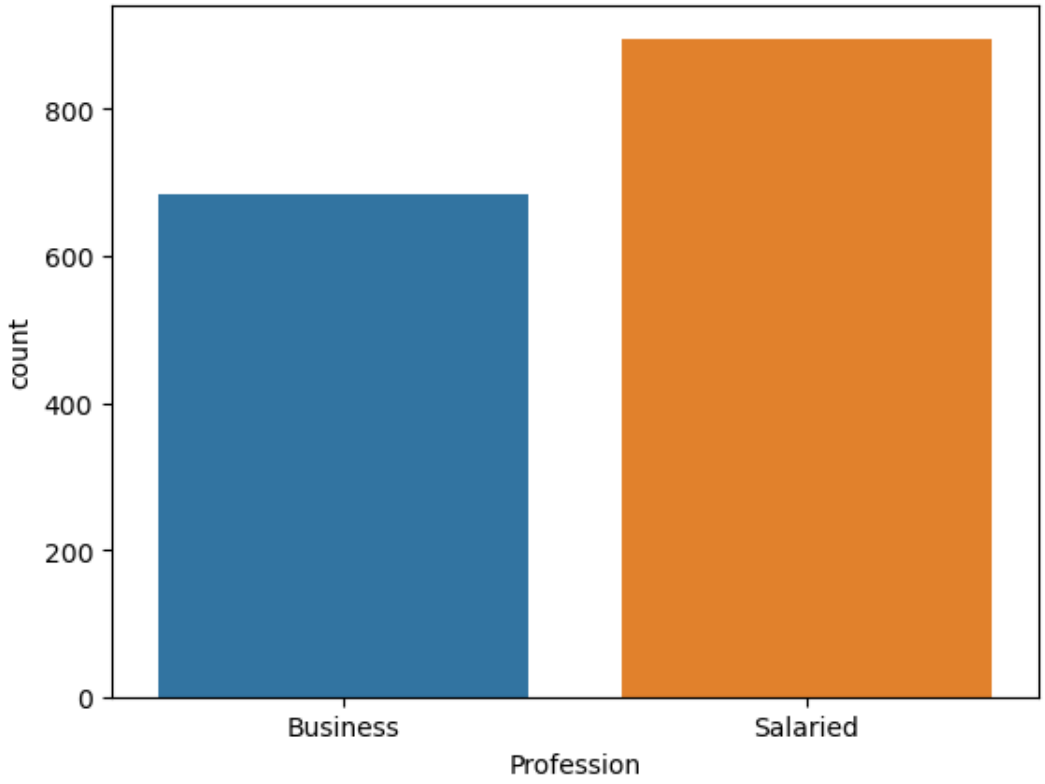
- Also, we found out that there some mistakes in “Gender” column by “Femal” and “Femle” values.
- These values are replaced with “Female” with the help of replace function.

3. Exploring the data with appropriate visuals and visualising the data with Statistical plots.

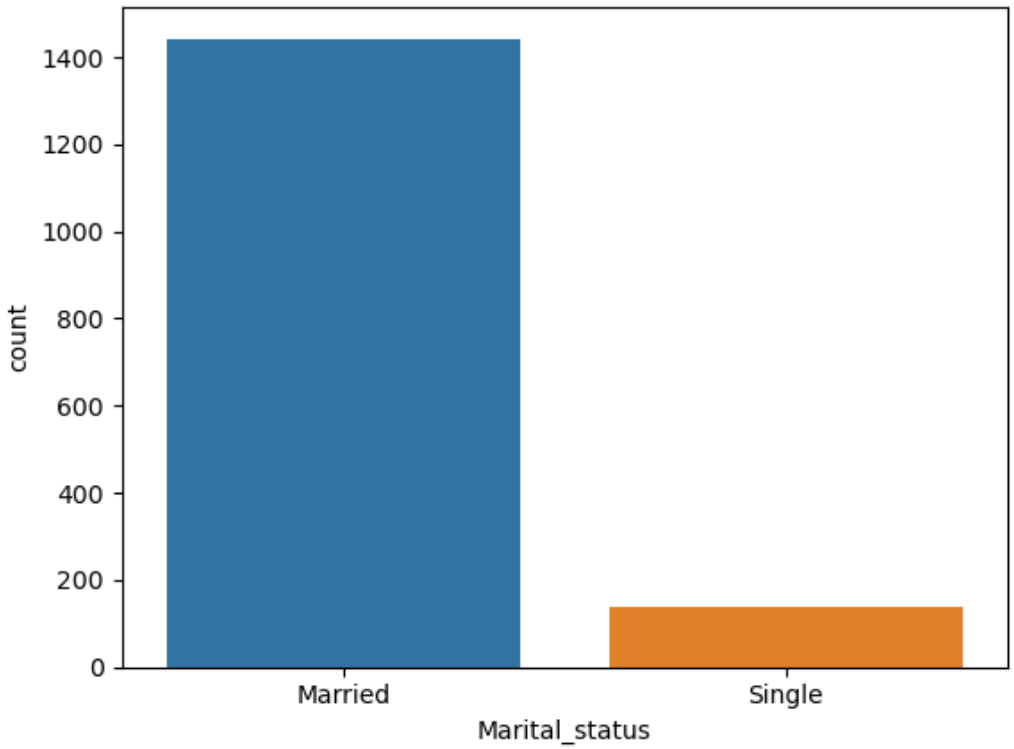
A. **Plotting with categorical values:**



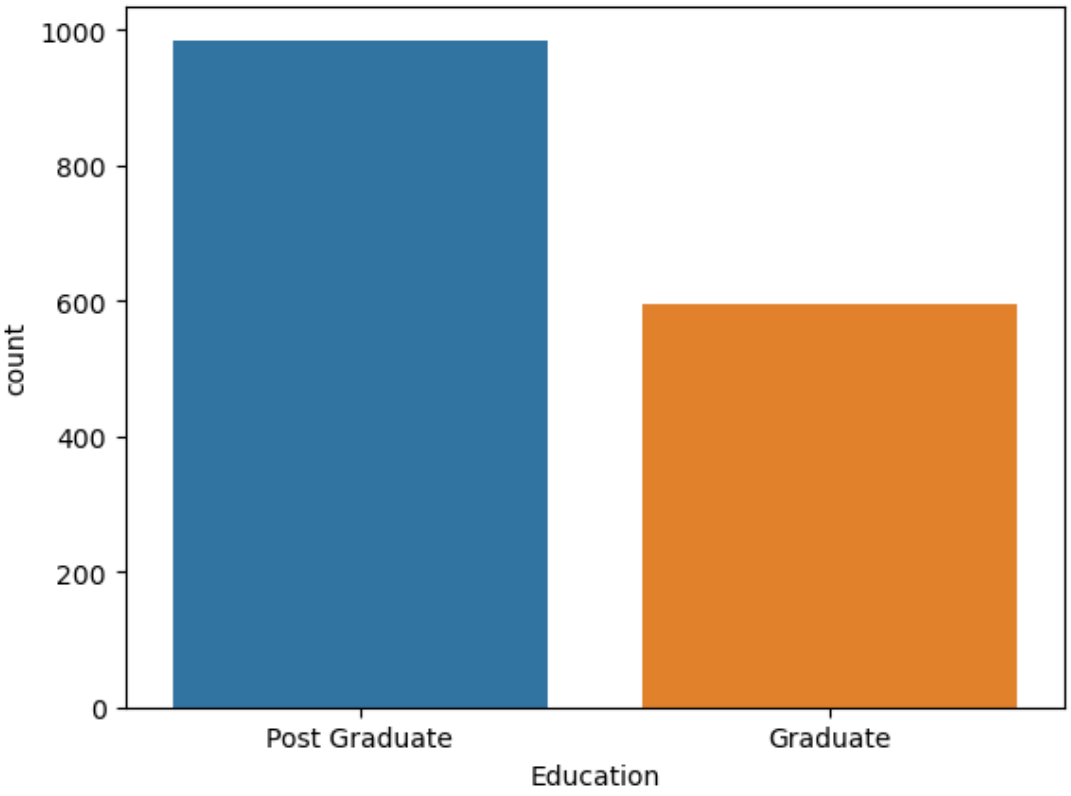
- *From the Above plots we can clearly see that the Male are more in count than Female. By which we can say that there are more number of Male customers than Female customers.*



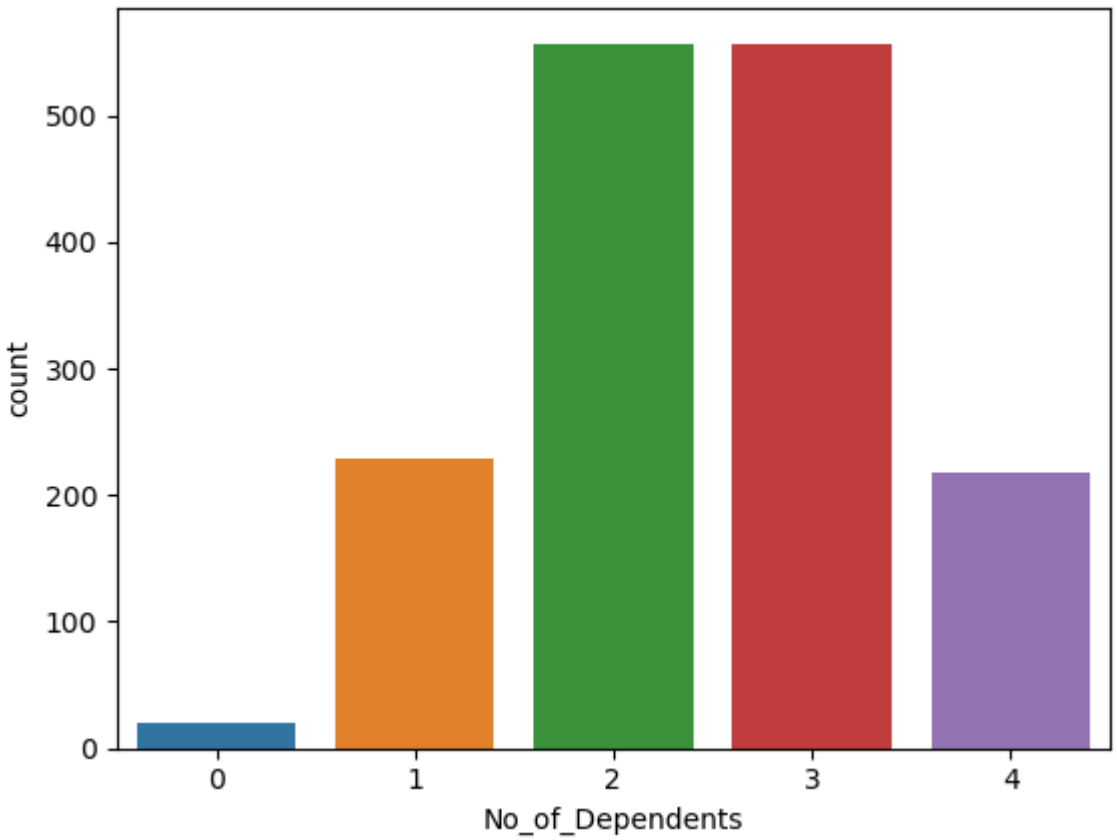
- *The people who bought a car are most likely doing a Salaried job and there are less people doing Business has bought a car.*



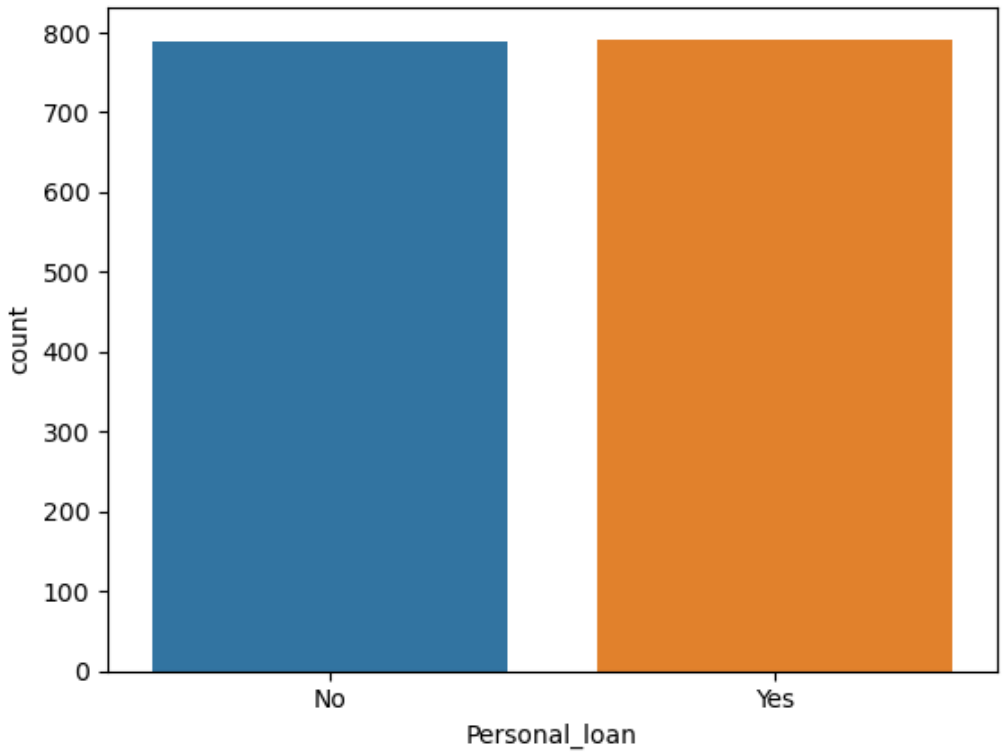
- *From the above plot we can see there are more number of Married people who bought a car than Single.*



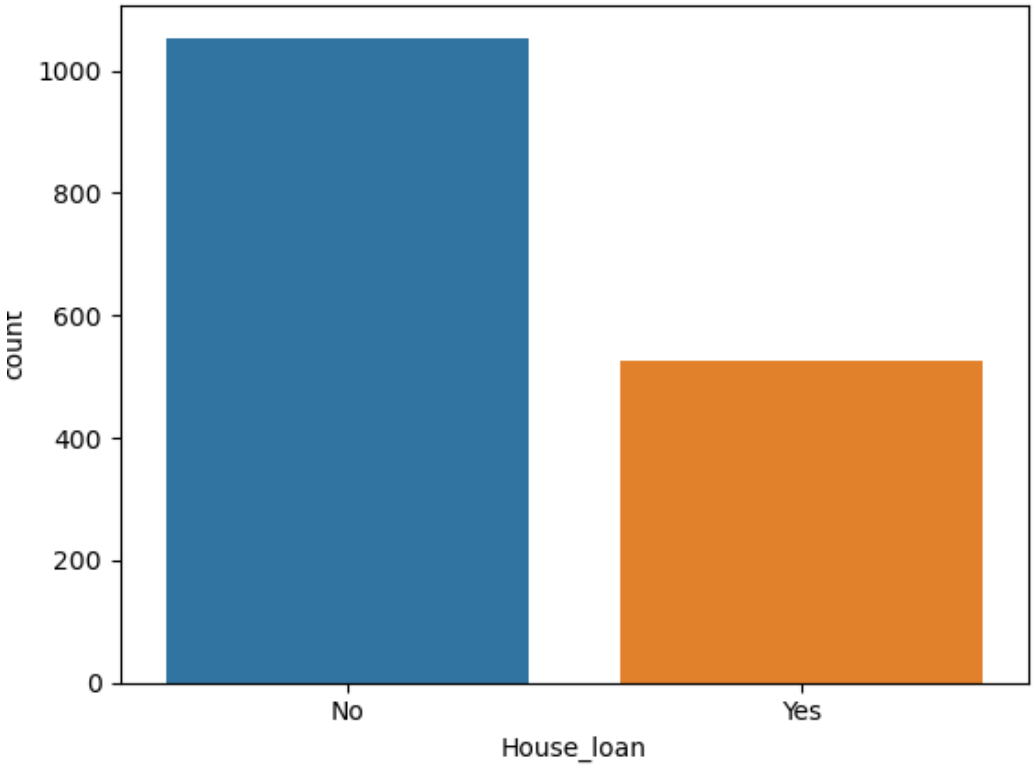
- *From the above plot the people with Post Graduation more likely bought a car than Graduates*.



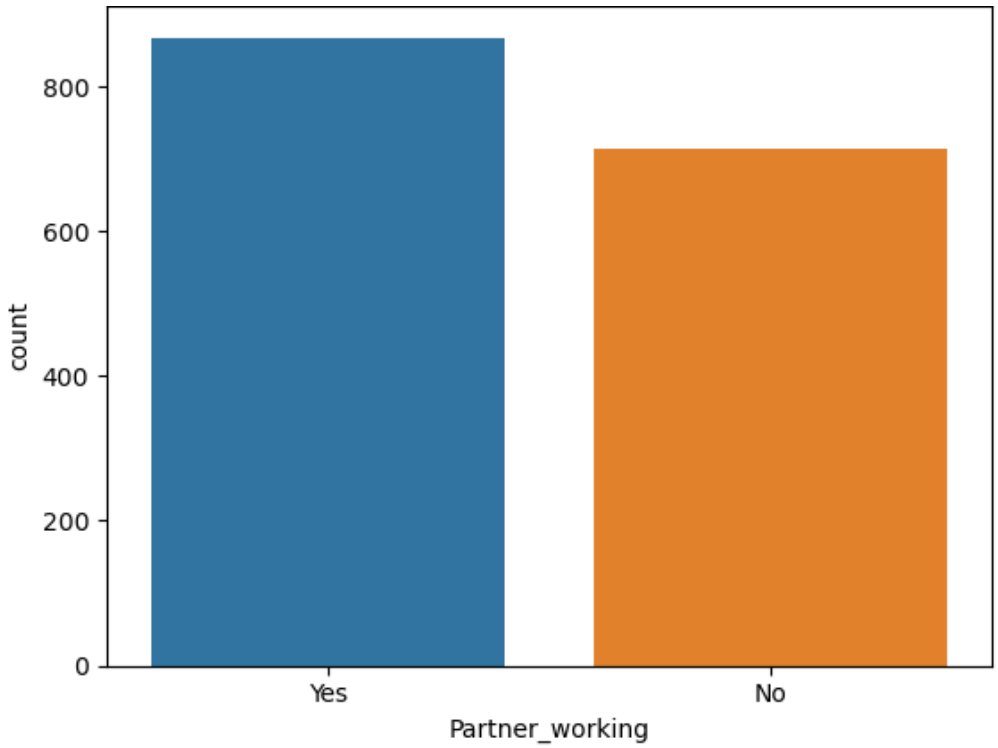
- *People who has No children are comparatively very less with people having children.*
- *People having 1 children and People having 4 children both are almost similar but there is a very small difference.*
- *People having 2 children and people having 3 children are equal in count.*
- *People with children most likely bought a car.*



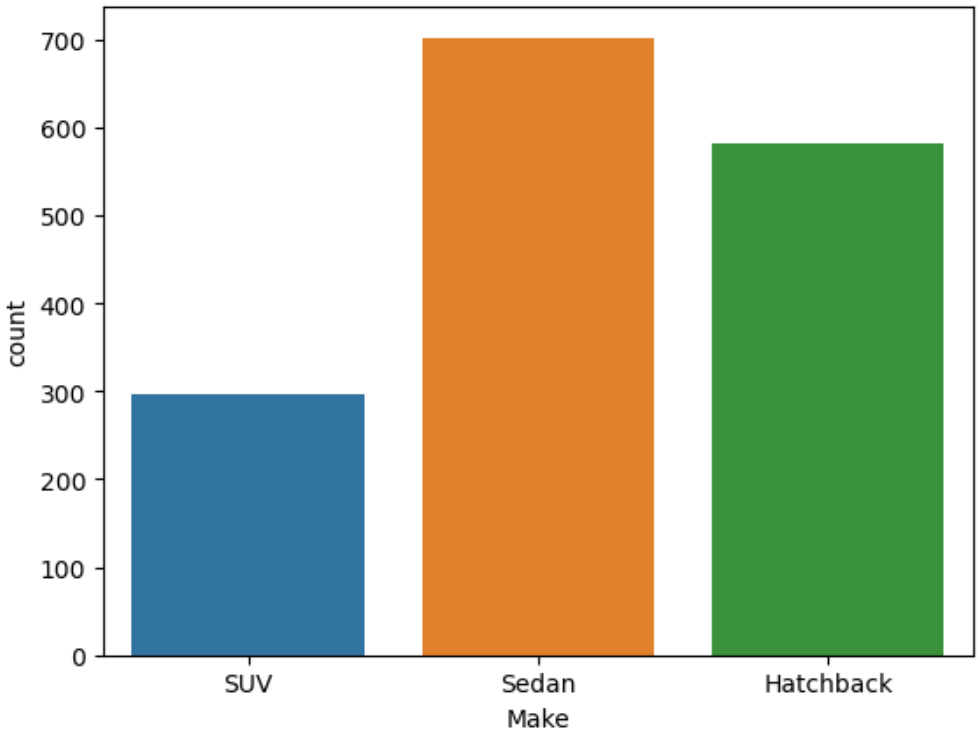
- *There is no difference in people having personal loan and people who are not having any personal loan both are equal in number.*



- *There is an interesting insight that people who bought a car and are having Housing loan are less in count.*
- *People without House loan and bought a car are more in number.*

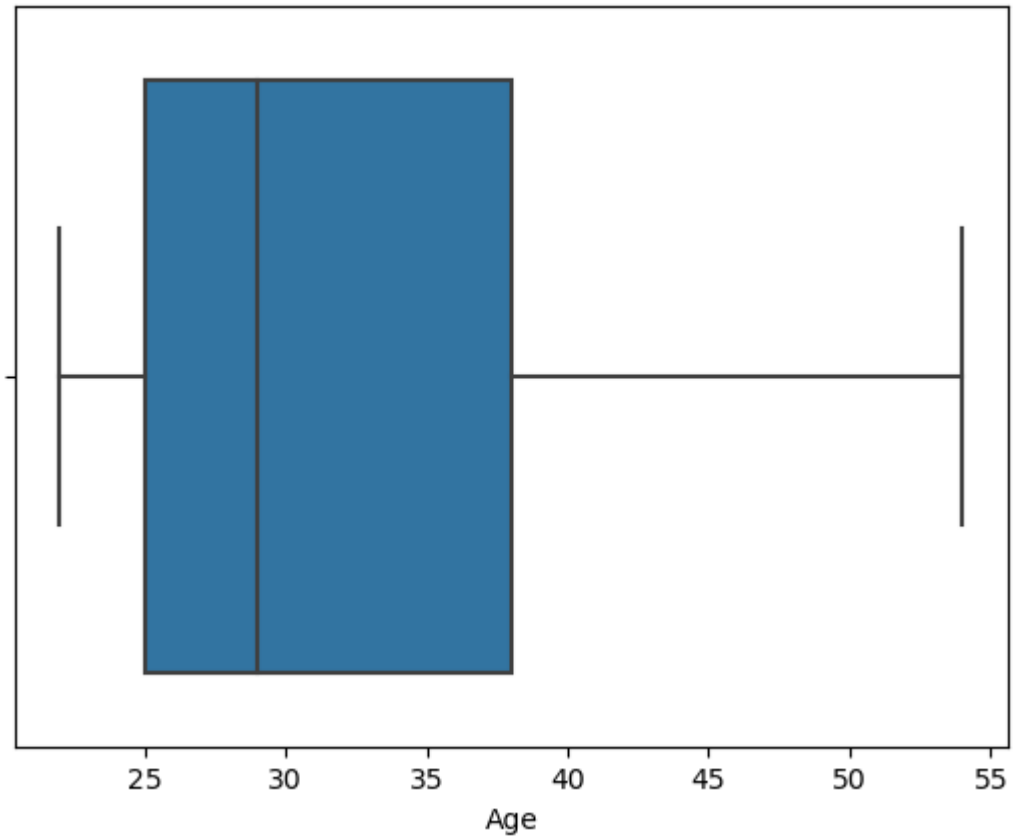


- *The count of people with their partner working and bought a car are more.*
- *It is interesting that the Partner working has less impact in people buying the car, the difference is less compared to Partner not working.*

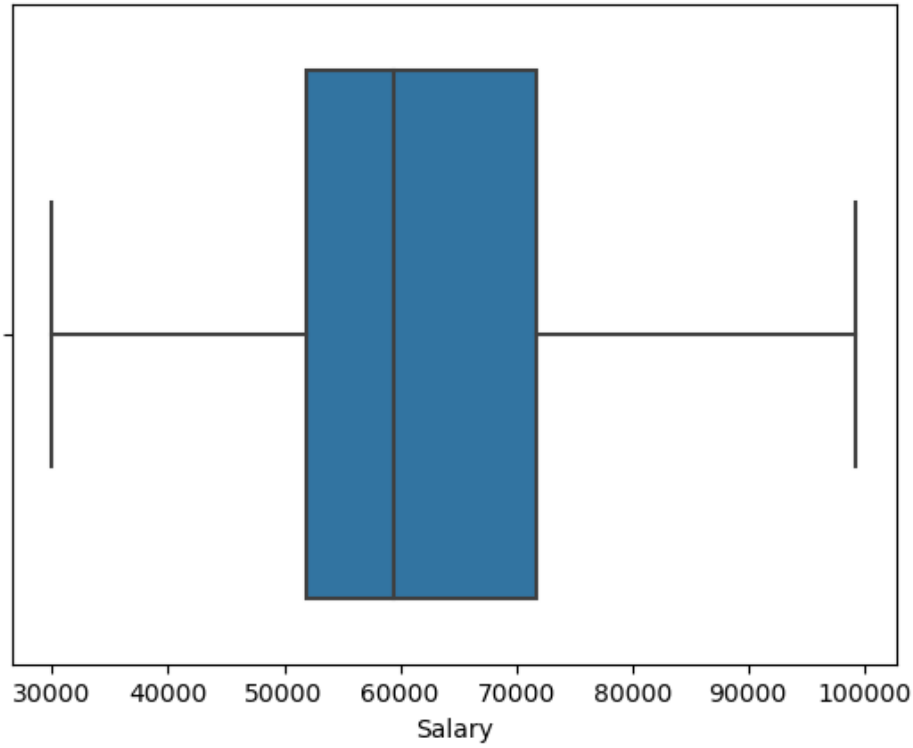


- *Sedan is the most bought car compared to other cars, then Hatchback is most sold car and SUV is the least sold car compared to other two cars.*

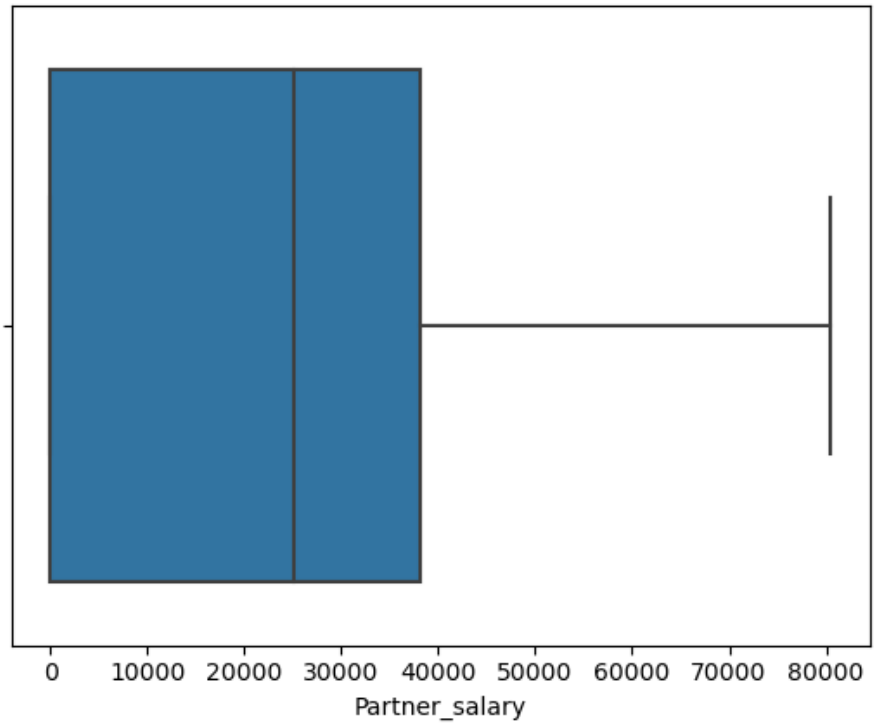
B. **Plotting with Numerical values:**



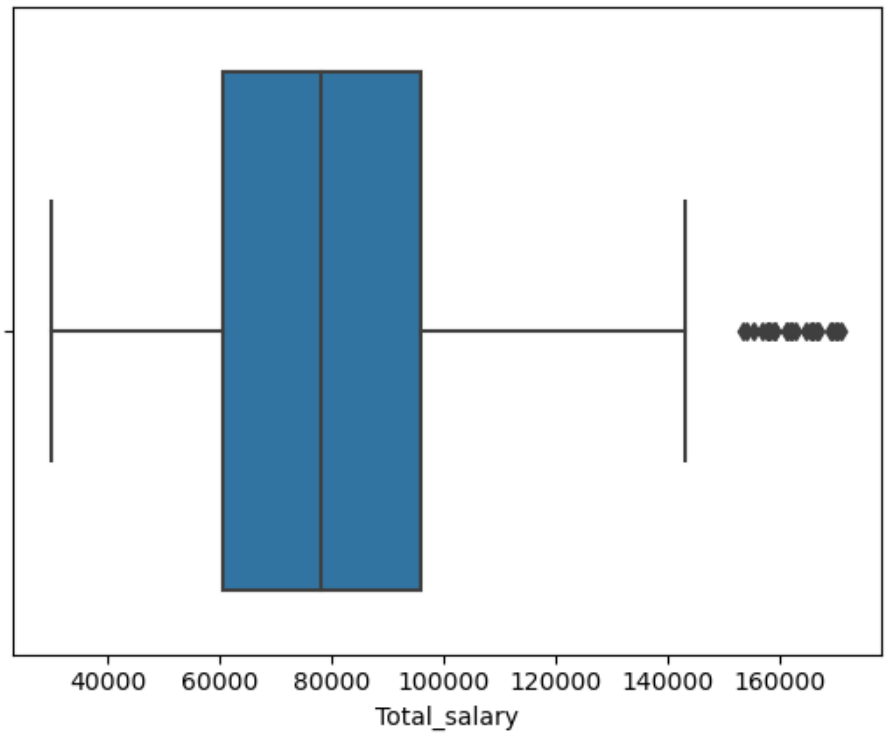
- *The Maximum number of customers are from Age category of 25 to somewhere around 37.*
- *The plot is skewed towards the right (Positively skewed).*
- *Median value of Age (is 29 from the .describe table) lies somewhere near 30.*



- *The Average salary of people lies between 50000 and 72000.*
- *There are no outliers in this data.*

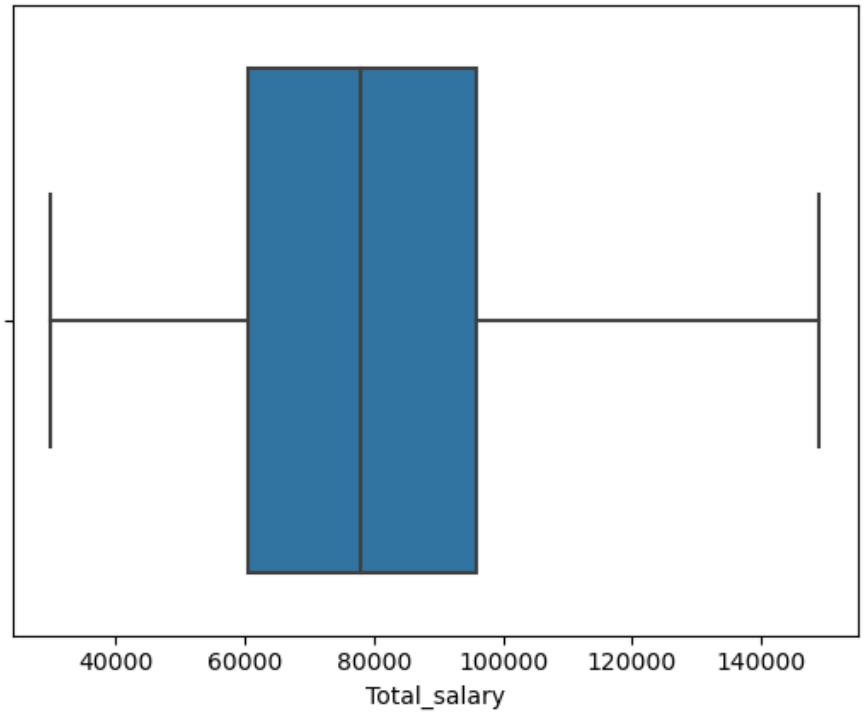


- *Since there are more people who's partner are not working the majority of data lies between 0 to 40,000.*
- *It will be interesting to see the relationship between `Partner_salary` and `Price` of the car in the Bivariate analysis.*



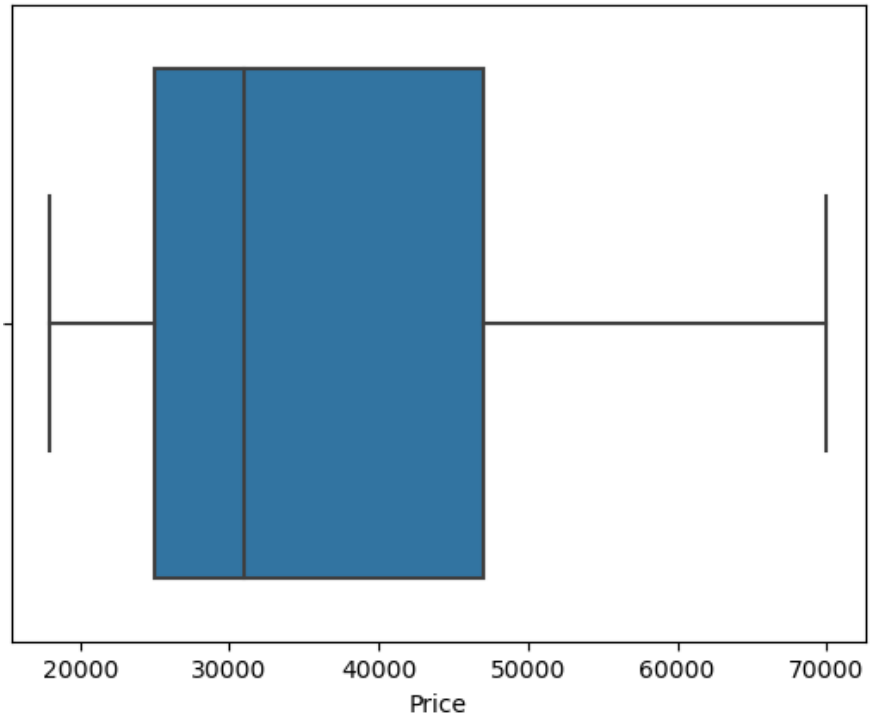
- *There are outliers in `Total_salary` that needs to be treated.*
- *Most of the Total salary lies between 60,000 to 95,000.*

After Treating the outlier:



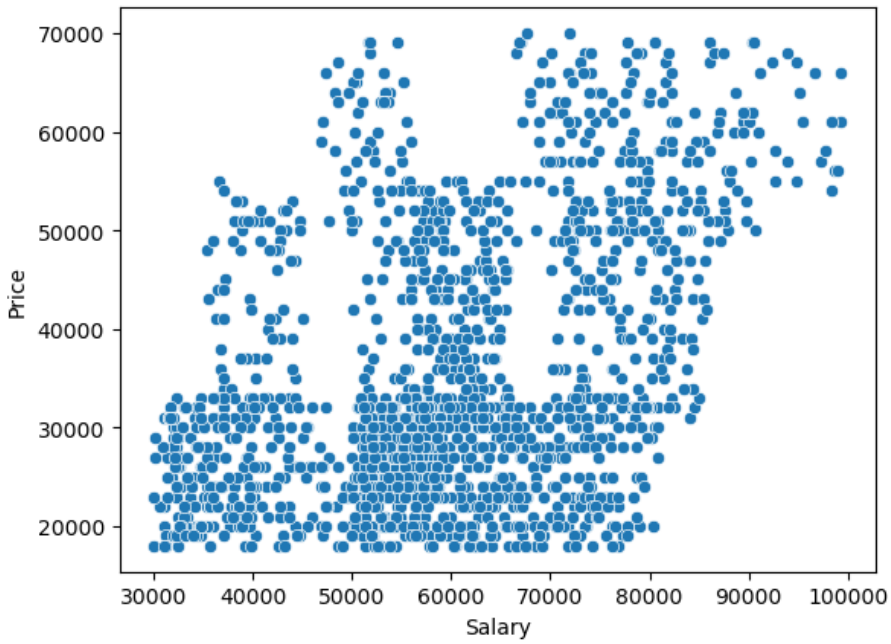
- *The outliers are treated and the values are capped. The values are capped by:*

  *Q3 + 1.5 * IQR*

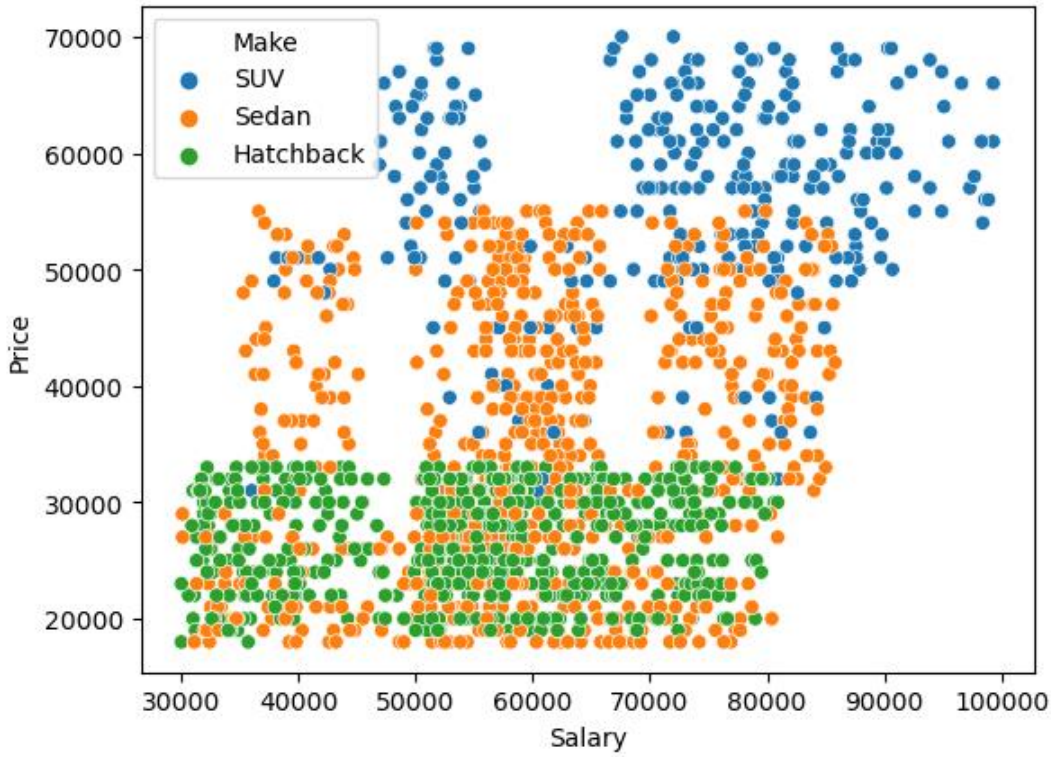


- *The Median price of the car between 25,000 to 47,000.*
- *The plot is Right skewed which means most of the car price lies on left.*

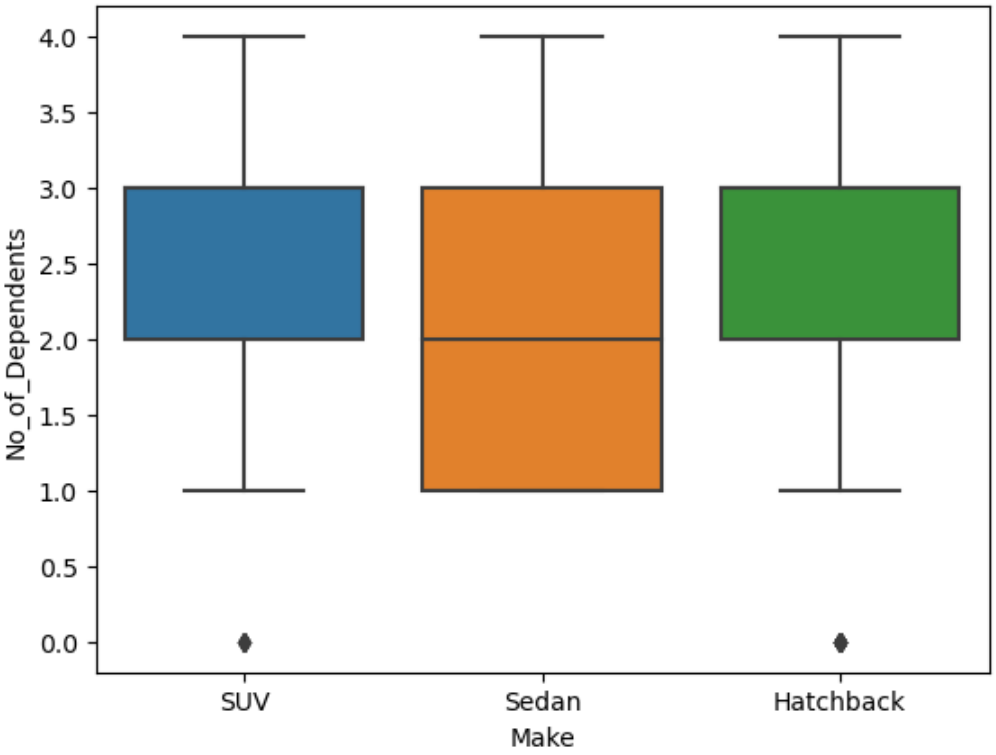
4. Exploring the relationship between the data with appropriate visuals and visualising the data with Statistical plots.



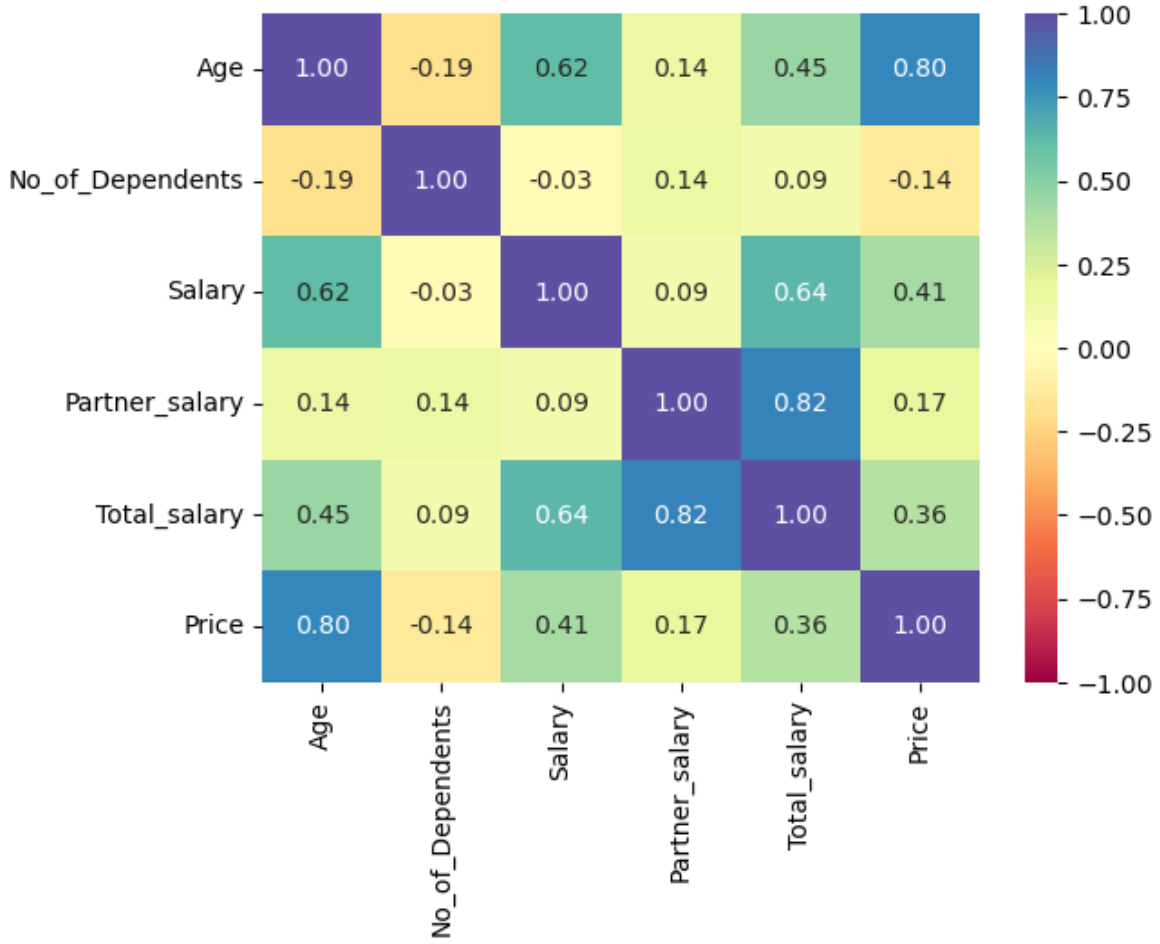
- *The above plot shows that the people with higher salary can buy higher priced car.*
- *Hence it is Positively correlated.*



- *People with higher salary buy SUV and then comes Sedan and Hatchback.*



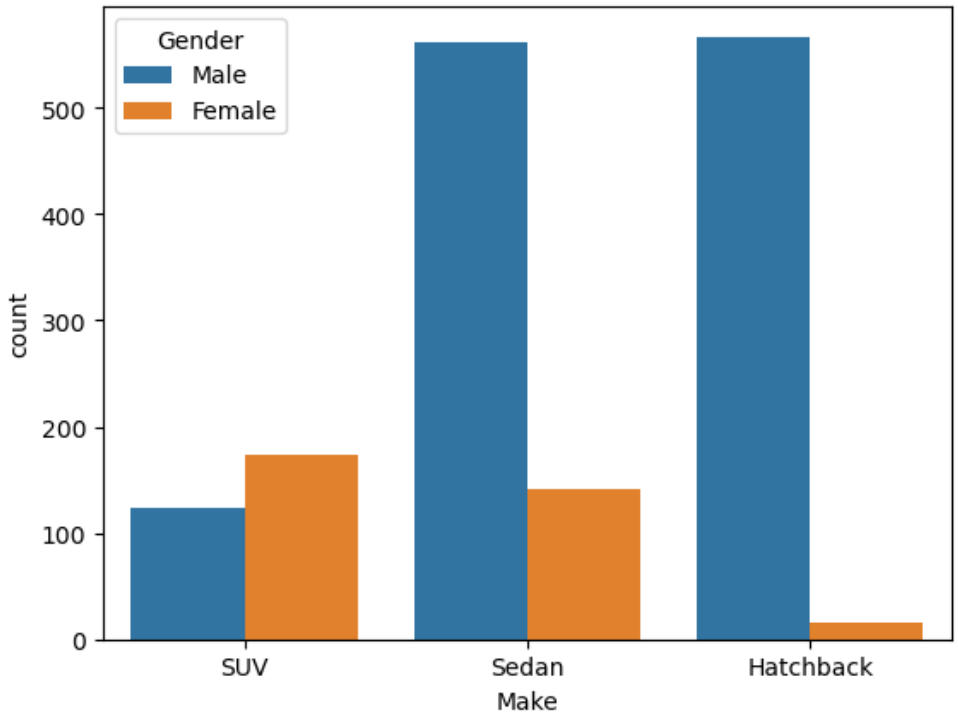
- *There are some outliers present in the plot but the No_of_Dependents value can be zero since there might not be any dependents for a particular individual.*



- *Most of the values are Positively correlated which means we have the trend going upwards.*
- *The values Partner_salary and Total_salary have 0.82 and have high positive correlation.*
- *The values Age and Price have 0.80 are the two which have high positive correlation.*
- *From the above plot we can clearly see that there are only two negatively correlated plots one is one is Age and No_of_Dependents and the other is Price and No_of_Dependents.*
- *There is a good correlation between Age and Salary.*
- *All the Salary, Partner_salary and Total_salary column are highly correlated with each other.*
- *The Negative correlation for the plots is very minimal with the values -0.14 and -0.19.*
- *It is interesting that Age and No_of_Dependents are negatively correlated (-0.19) but the correlation is close to 0.*

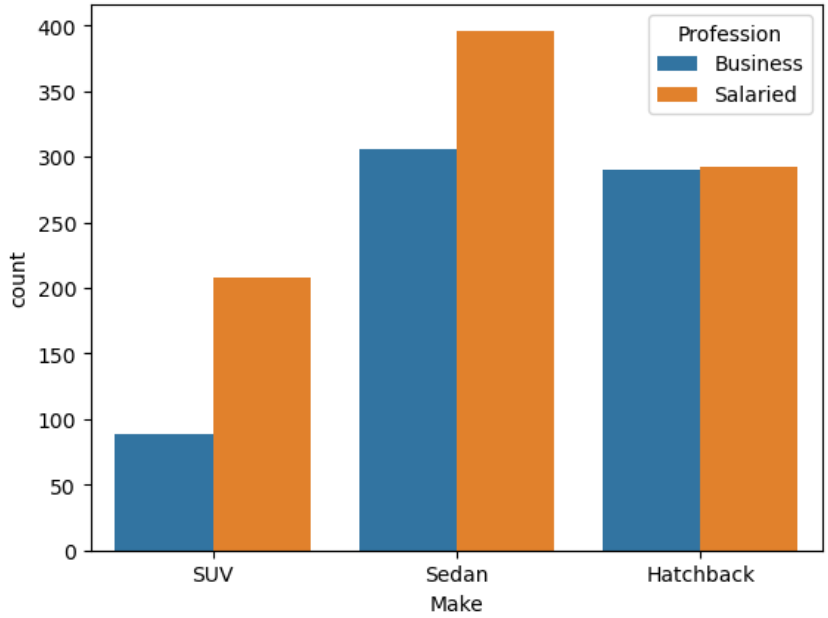
5. **Here are some Key Business Question to which we try find the solution by using visualisation and EDA methods.**

   a. **Do men tend to prefer SUVs more compared to women**



- *No men prefer Sedan and Hatchback, from the above plot we can clearly see that Women tend to prefer more SUV than Men.*

   b. **What is the likelihood of a salaried person buying a Sedan?**



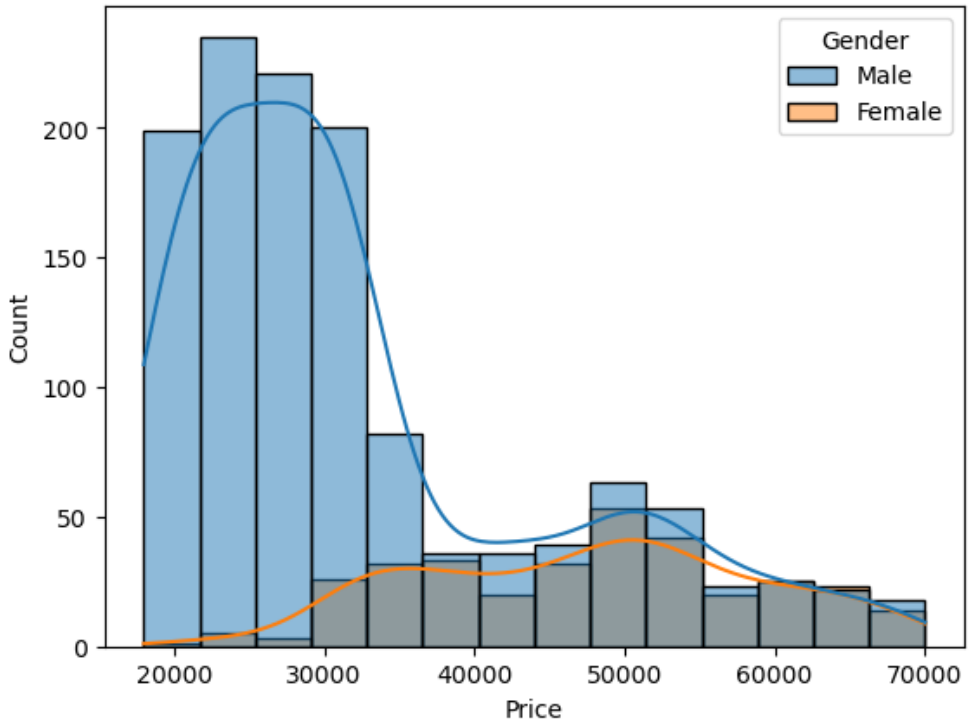
- *As we can see from the above plot Salaried person likes to prefer a Sedan over SUV and Hatchback.*

c. 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?

```
Profession  Gender  Make
Business    Female  SUV          55
                    Sedan        50
            Male    Hatchback   290
                    SUV          34
                    Sedan       256
Salaried    Female  Hatchback    15
                    SUV         118
                    Sedan        91
            Male    Hatchback   277
                    SUV          90
                    Sedan       305
Name: Make, dtype: int64
```

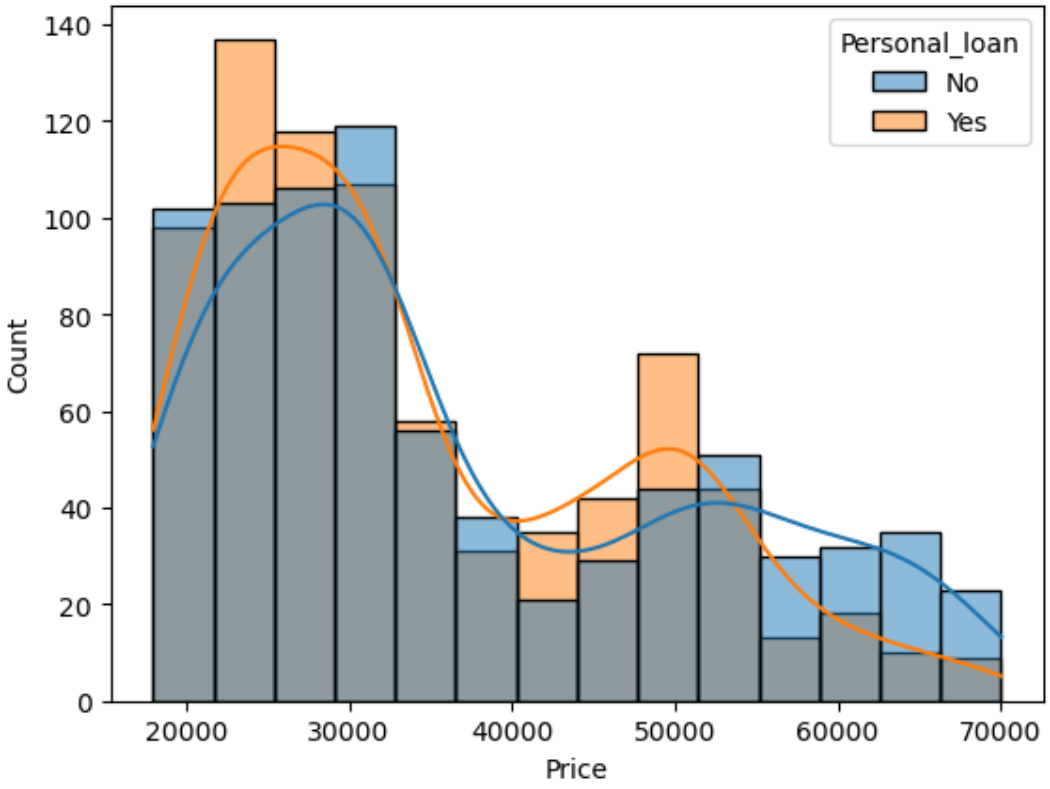
- *As per the above Table Sheldon Cooper's claim seem to be wrong that the Salaried male prefers Sedan over SUV.*
- *The above Observation can be used for the analysis for the question asked as answers the previous business questions too.*

d. How does the amount spent on purchasing automobiles vary by gender?



- *From the Above Histogram we can clearly see that Male prefer less priced cars compared to Female*

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

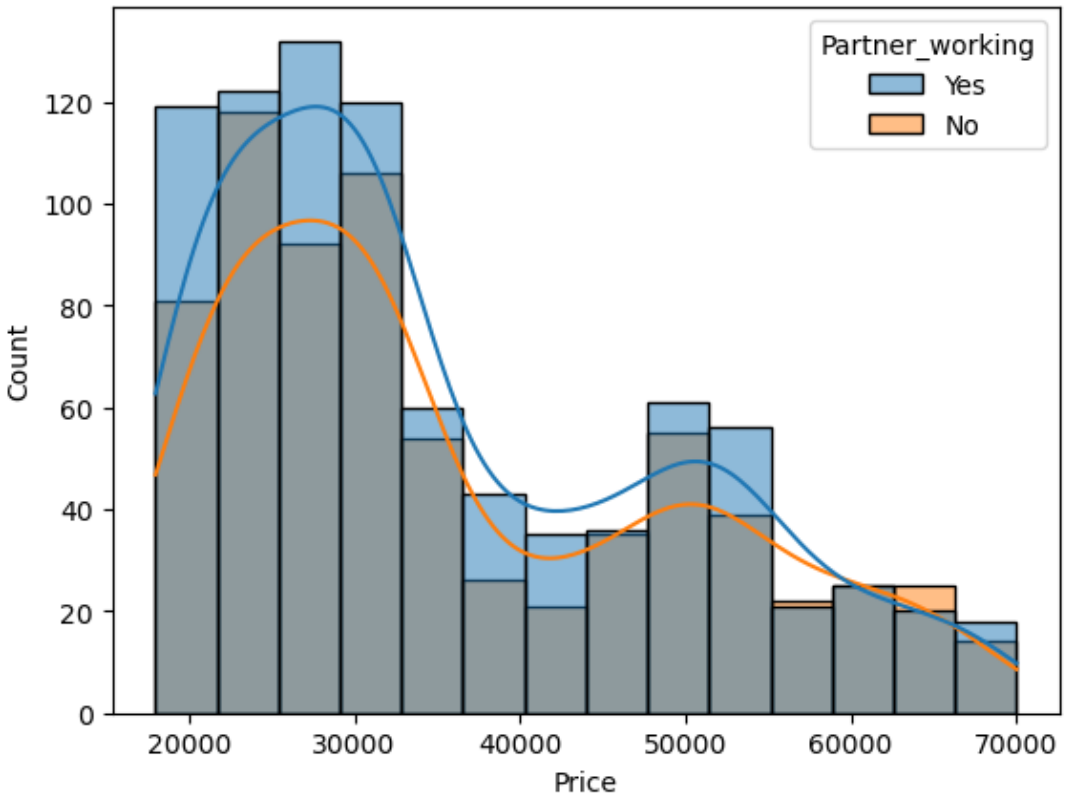


- *From the above histogram and Barplot, we can see that people having Personal loan prefer Low priced cars but the difference is very light.*
- *Higher salaried people haven't taken personal loan to buy a higher priced car.*
- *Lower salaried people might have take Personal Loan to buy a car.*

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

- *The money which is spent on purchasing the Automobiles by individuals who took personal loan is 27290000.*

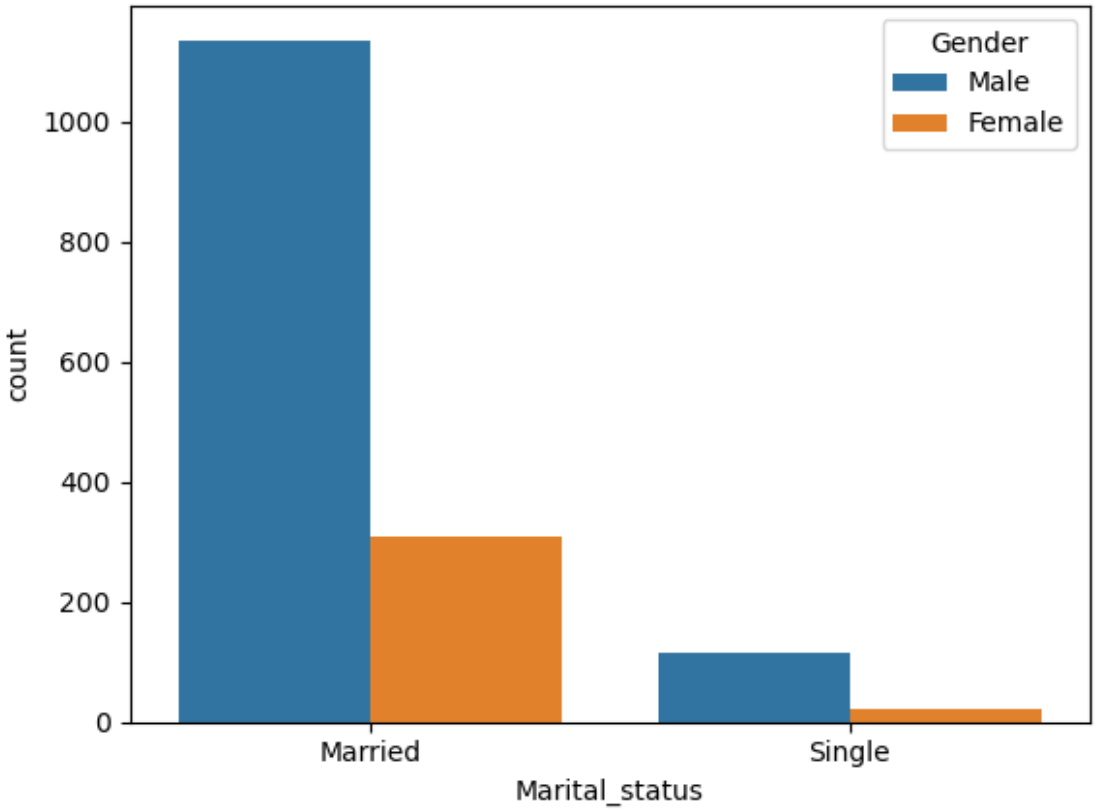
f. How does having a working partner influence the purchase of higher-priced cars?



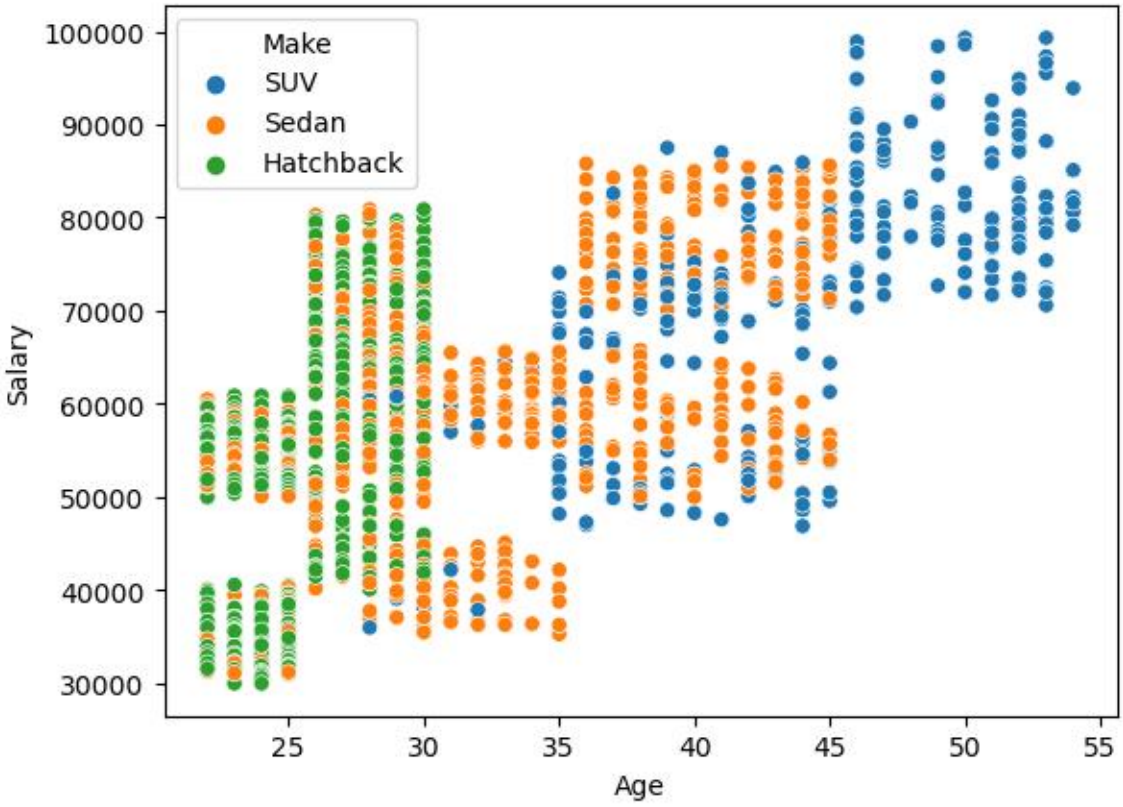
- *From the above histogram we can see that individuals having working partner tends to buy Higher priced cars and individuals with partner who is not working likely to buy lower priced car.*
- *We can come to a conclusion that People with Working partner tends to buy higher priced car.*

6. Actionable Insights & Business recommendations

A. From the above analysis we can clearly see that Male customers are more in number than Female customers, so further analysis is required



- *From the above plot we can see that Married Men tend to buy car than Married Female and both single Male and Female*
- *Since Male customers are more in number the marketing team can focus in Single Male customer to gain more profit*



- *From the above plot we see that people of young age prefer Sedan and Hatchback than SUV.*
- *People of Age category around 31 to 45 mostly prefer Sedan and SUV.*
- *Aged people above 45 prefer SUV.*
- *This may be due to young people have Less salary o don't prefer SUV and Middle aged and Aged people might have more salary and No_of_Dependents so they prefer SUV but it is an interesting observation*
- *The Marketing team can focus on attracting people with different Age category by the above observation.*