

# Business Report

Austo Automobile



- Ruchi Rikta

# Contents

1. About.....	1
2. Data Structure.....	2 - 6
3. Univariate Analysis.....	7 - 18
4. Bivariate Analysis.....	19 - 21
5. Key Questions.....	22 - 27
6. Conclusions.....	28
8. Recommandations.....	29

# About

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

## Objective

They want to analyze the data to get a fair idea about the demand of customers, which will help them in enhancing their customer experience. Suppose you are a Data Scientist at the company, and the Data Science team has shared some of the key questions that need to be answered. Perform the data analysis to find answers to these questions that will help the company improve the business.

# 1. Data Information

There are 1581 Rows and 14 Columns.

Column	Data Type
Age	int64
Gender	object
Profession	object
Marital_status	object
Education	object
No_of_Dependents	int64
Personal_loan	object
House_loan	object
Partner_working	object
Salary	int64
Partner_salary	float64
Total_salary	int64
Price	int64
Make	object

- The Partner\_salary column is being read as a 'float64' data type, but it should be "int64" data type.

- Converted 'Partner\_Salary' column from 'float' to 'int' by filling the null values with 0.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1581 entries, 0 to 1580
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        1581 non-null   int64
11  Total_salary          1581 non-null   int64
12  Price                 1581 non-null   int64
13  Make                  1581 non-null   object
dtypes: int64(6), object(8)
memory usage: 173.1+ KB
```

## Checking for missing values :

	0
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	0
Total_salary	0
Price	0
Make	0
dtype: int64	

- The "Gender" column contains 53 missing values, which constitute approximately 3% of the data. To prevent any potential errors in the future and ensure consistency, we will drop the rows with missing values in the "Gender" column, ensuring that no new invalid gender values are introduced.

## Checking for Data Irregularities :

- While checking unique values present in the data set, it was noticed that the 'Gender' column has data irregularities. It contains records that mentions - 'Male' 'Femal' 'Female' nan 'Femle'.

```
array(['Male', 'Femal', 'Female', nan, 'Femle'], dtype=object)
```

- Fixed it by replacing it with 'Female'.

## Checking Statistical Data :

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Age	1581.0	NaN	NaN	NaN	31.922201	8.425978	22.0	25.0	29.0	38.0	54.0
Gender	1528	4	Male	1199	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Profession	1581	2	Salaried	896	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Marital_status	1581	2	Married	1443	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Education	1581	2	Post Graduate	985	NaN	NaN	NaN	NaN	NaN	NaN	NaN
No_of_Dependents	1581.0	NaN	NaN	NaN	2.457938	0.943483	0.0	2.0	2.0	3.0	4.0
Personal_loan	1581	2	Yes	792	NaN	NaN	NaN	NaN	NaN	NaN	NaN
House_loan	1581	2	No	1054	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Partner_working	1581	2	Yes	868	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Salary	1581.0	NaN	NaN	NaN	60392.220114	14674.825044	30000.0	51900.0	59500.0	71800.0	99300.0
Partner_salary	1581.0	NaN	NaN	NaN	18869.512966	19570.644035	0.0	0.0	24900.0	38000.0	80500.0
Total_salary	1581.0	NaN	NaN	NaN	79625.996205	25545.857768	30000.0	60500.0	78000.0	95900.0	171000.0
Price	1581.0	NaN	NaN	NaN	35597.72296	13633.636545	18000.0	25000.0	31000.0	47000.0	70000.0
Make	1581	3	Sedan	702	NaN	NaN	NaN	NaN	NaN	NaN	NaN

## Observations :

- Mean age is 31.9 years and max is 54 years
- Average number of dependents 2.45 ~ 2-3 dependents
- Average salary is ~60,000
- Average partner salary is ~ 19000
- Average total salary is ~80000 with maximum as 171000
- The average price at which the vehicle is bought is ~36000, and the highest price is 70000

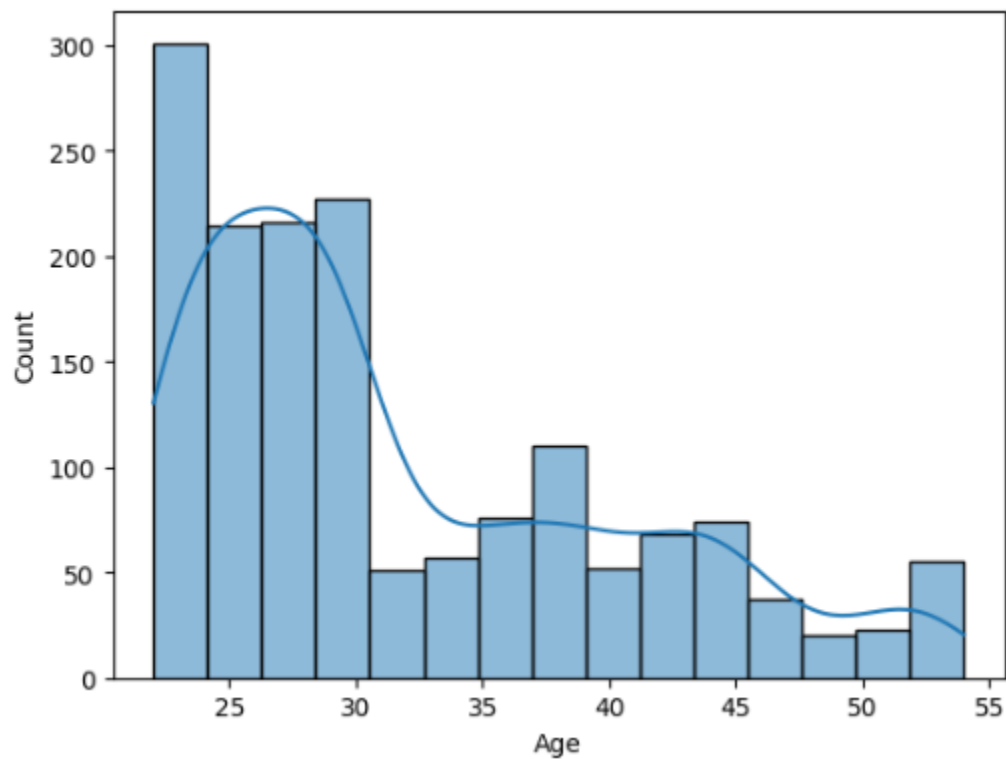


## 2. Univariate Analysis

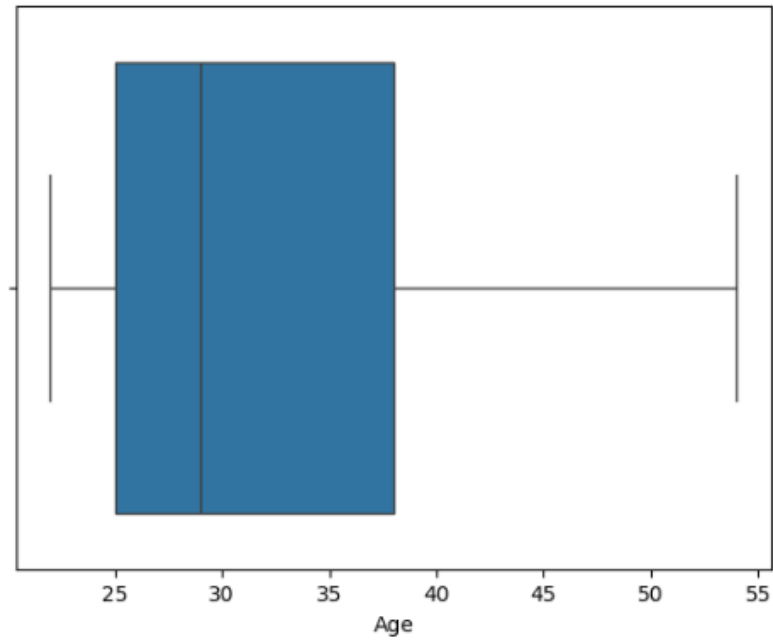
### 2.1. Numeric Variables

We can check the distribution of observations by plotting Histograms and Boxplots.

#### 2.1.1. Observations on Age

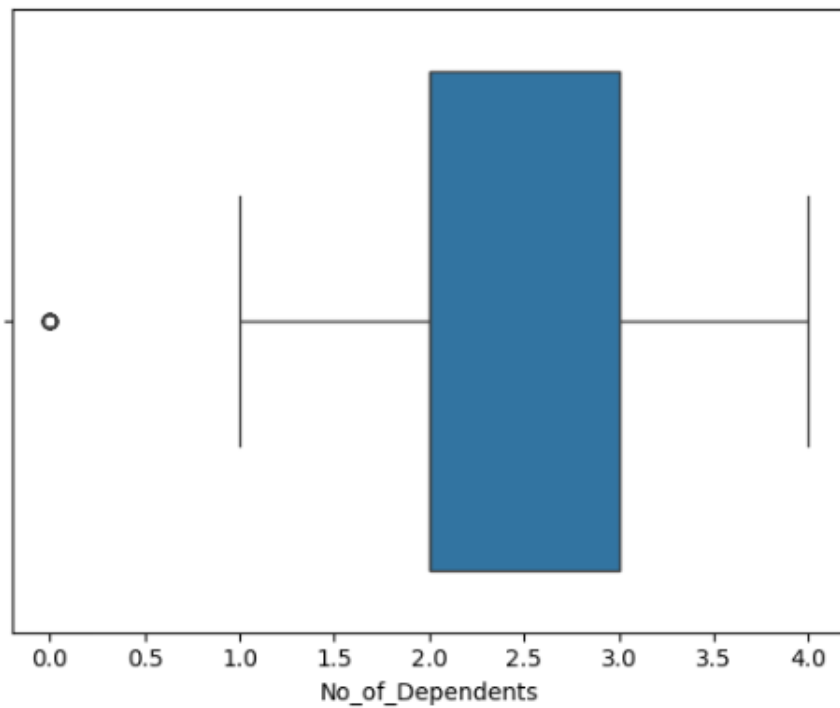
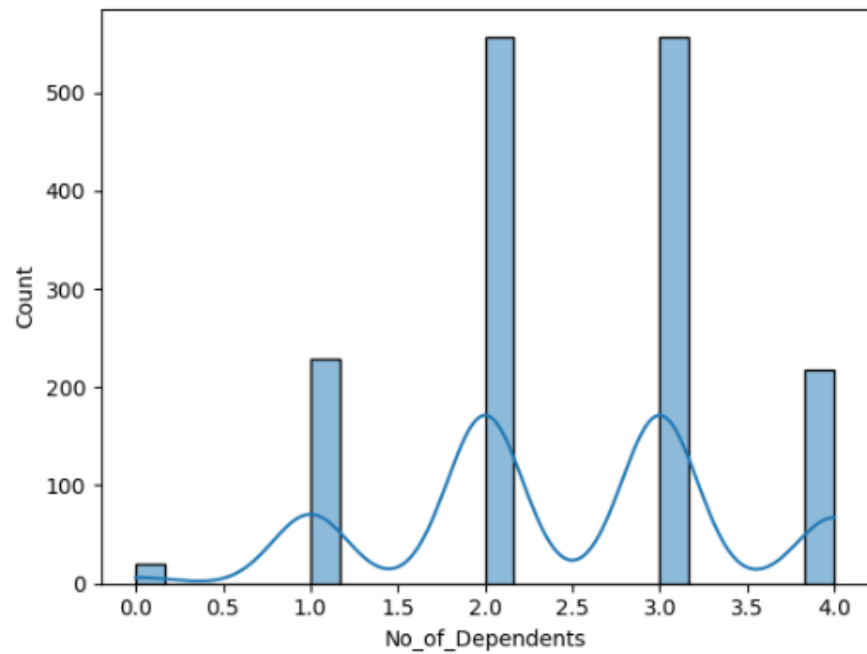


- The curve is more skewed towards the right.



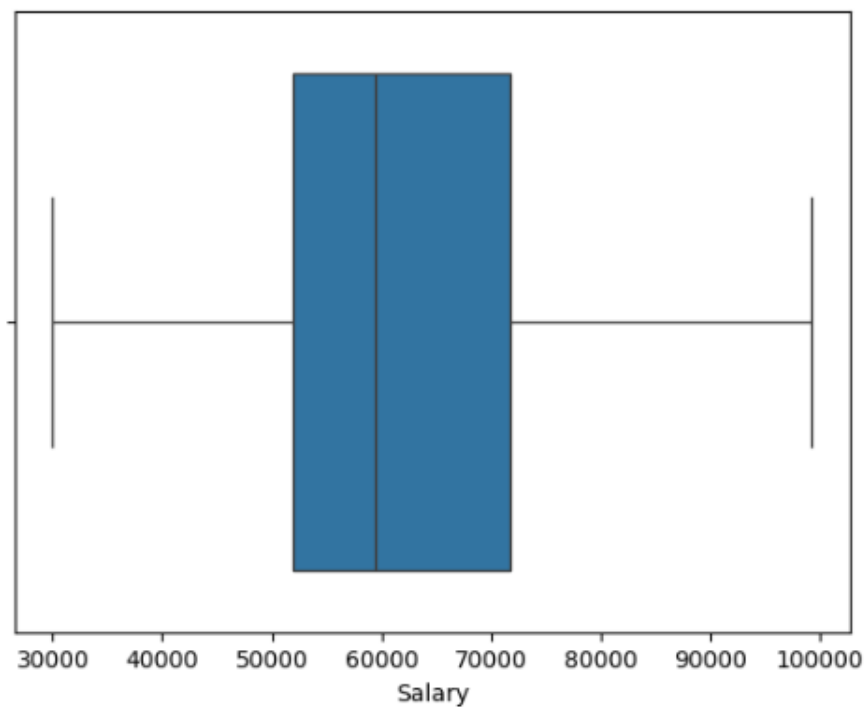
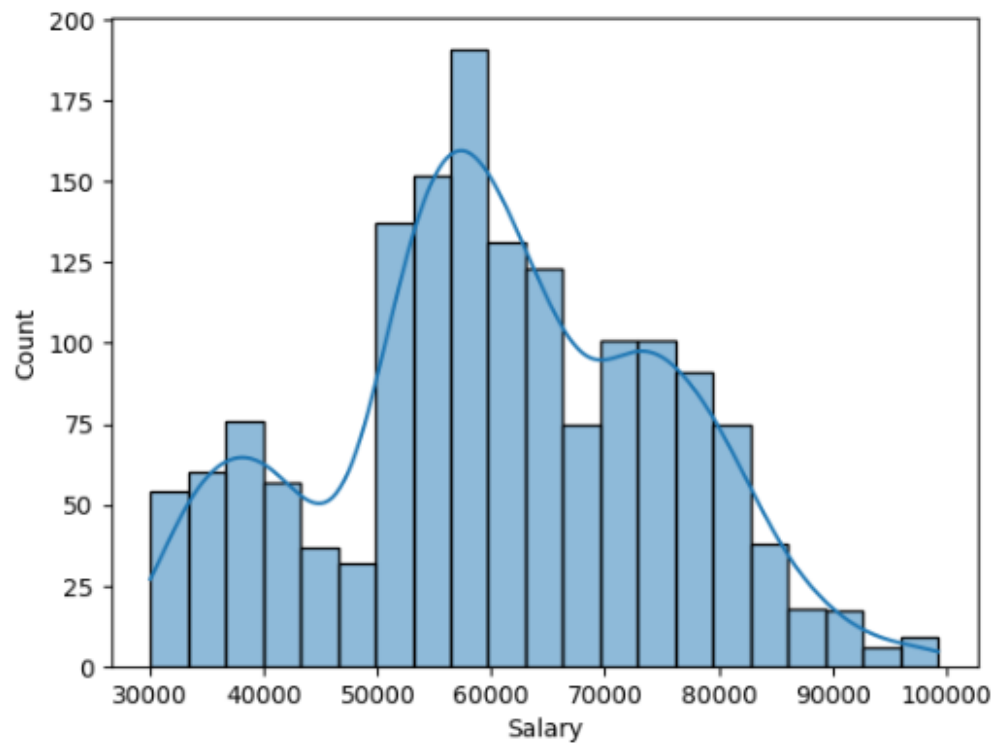
- About 60% of the people in the dataset are less than 30 years old.

### 2.1.2 Observations on No.of Dependents



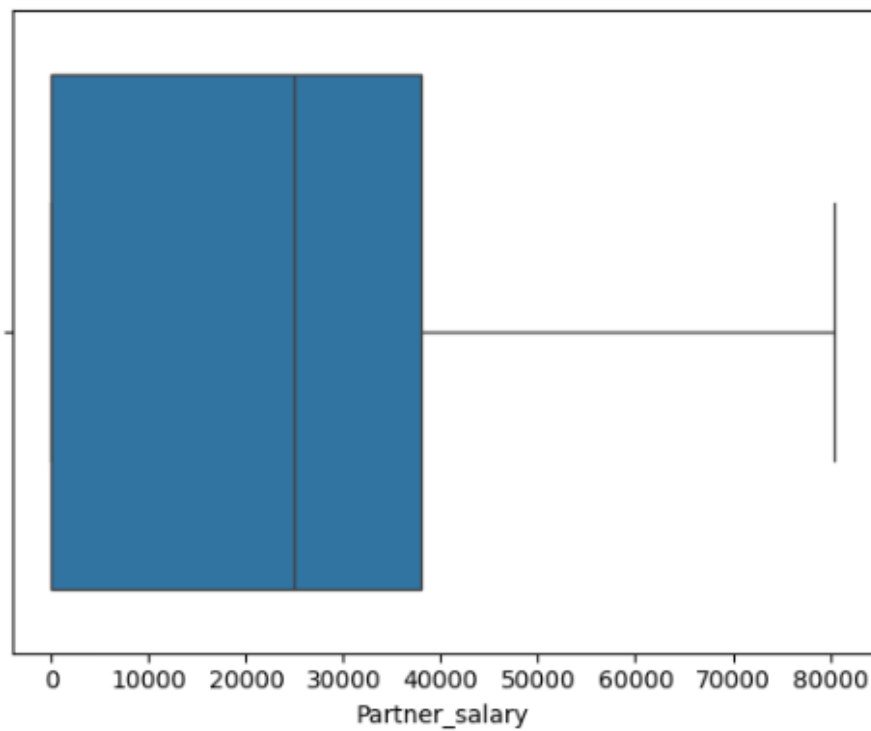
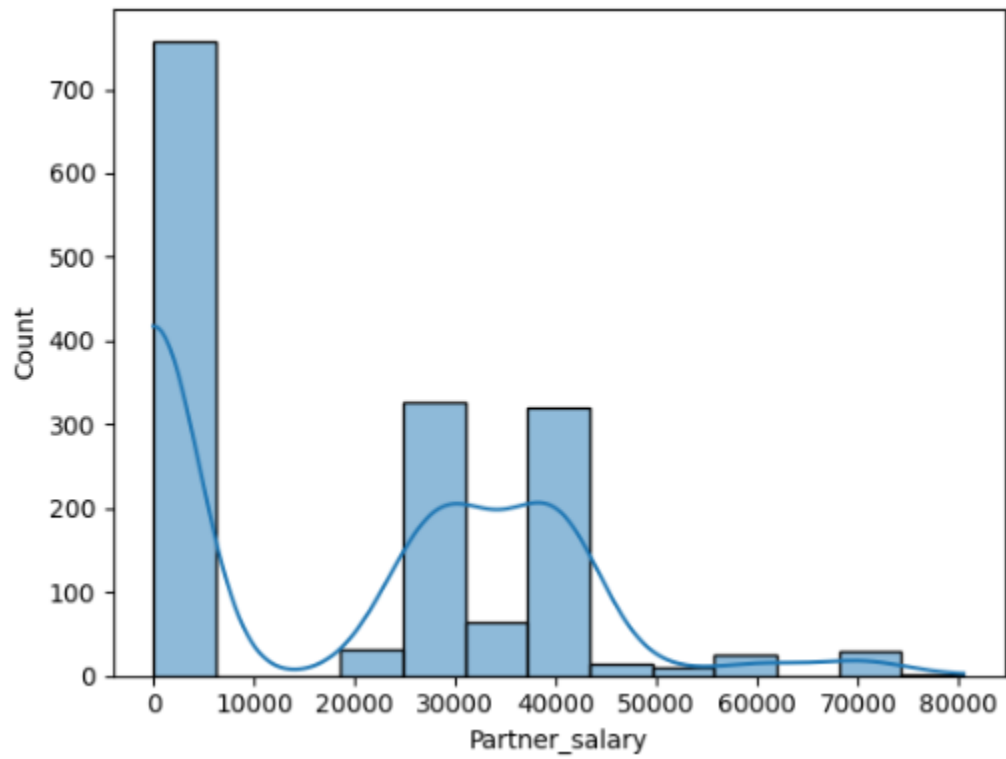
- There is a big difference between min and max. There are outliers present here.

### 2.1.3. Observations on Salary



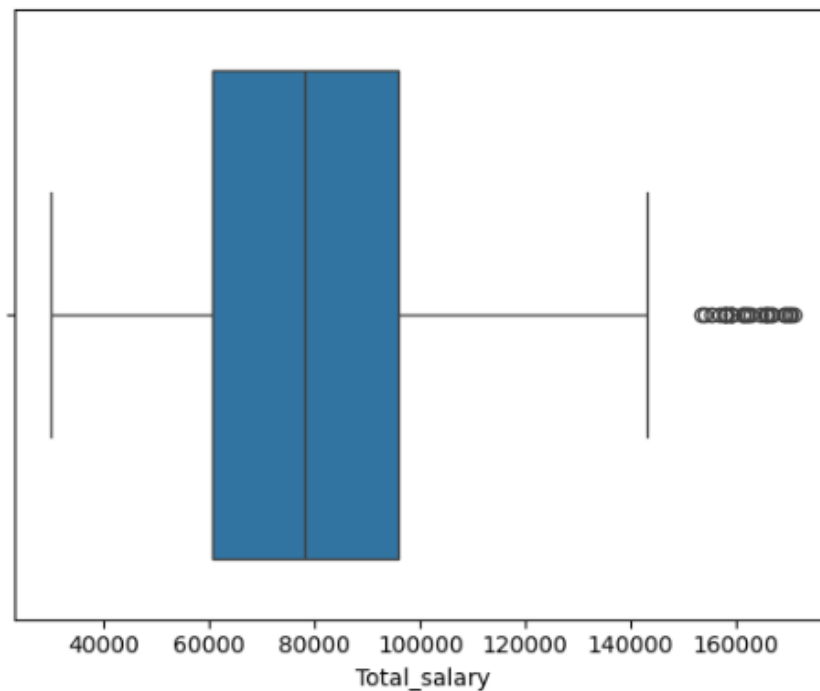
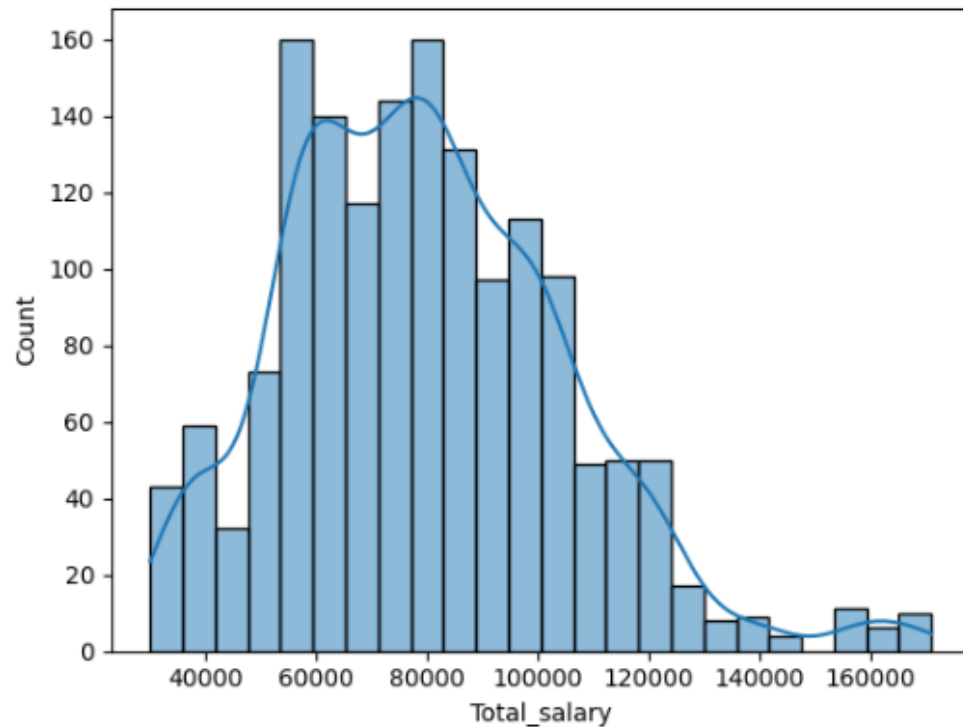
- It has a normal distribution.

### 2.1.4. Observations on Partner Salary



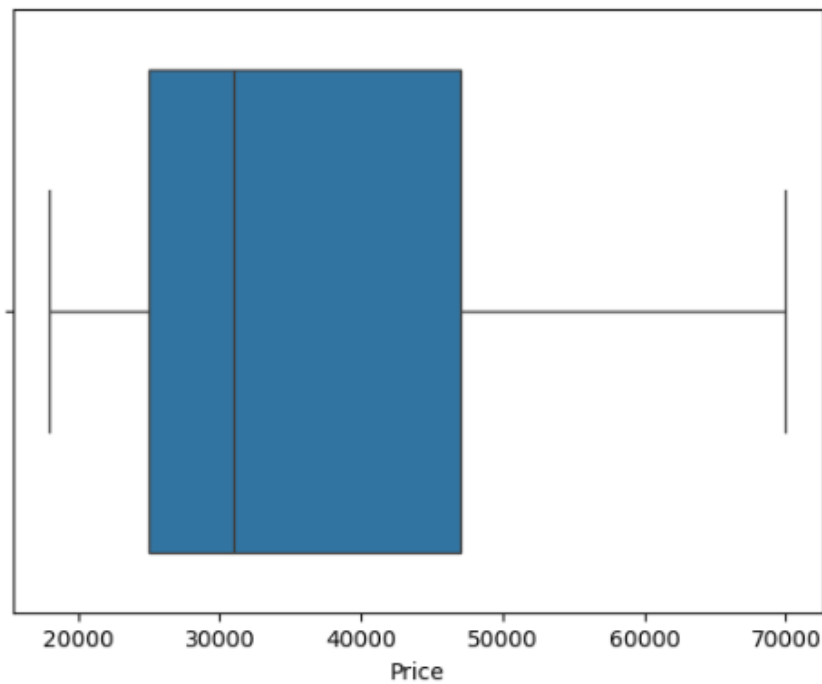
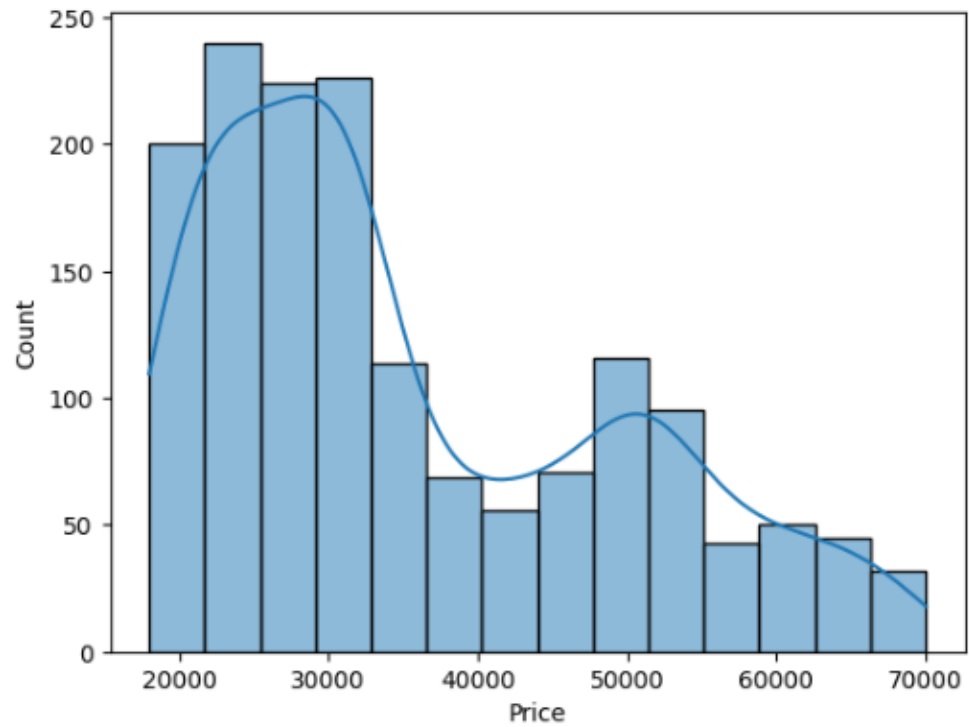
- It skewed towards the right side.

### 2.1.5. Observations on Total Salary



- The difference between min and max is huge, that's why we have a lot of outliers here.

### 2.1.6. Observations on Price

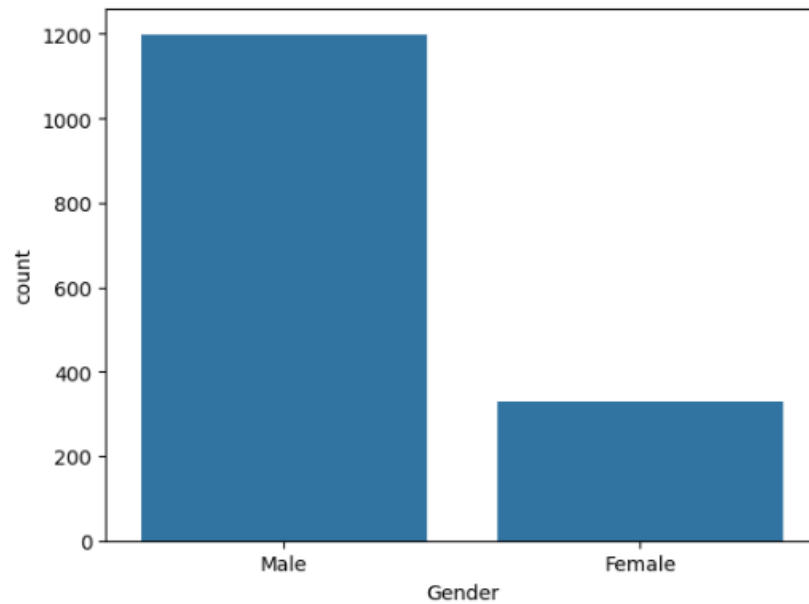


- It is right-skewed.

## 2.2. Categorical Variables

Bar Charts can be used to explore the distribution of Categorical Variables.

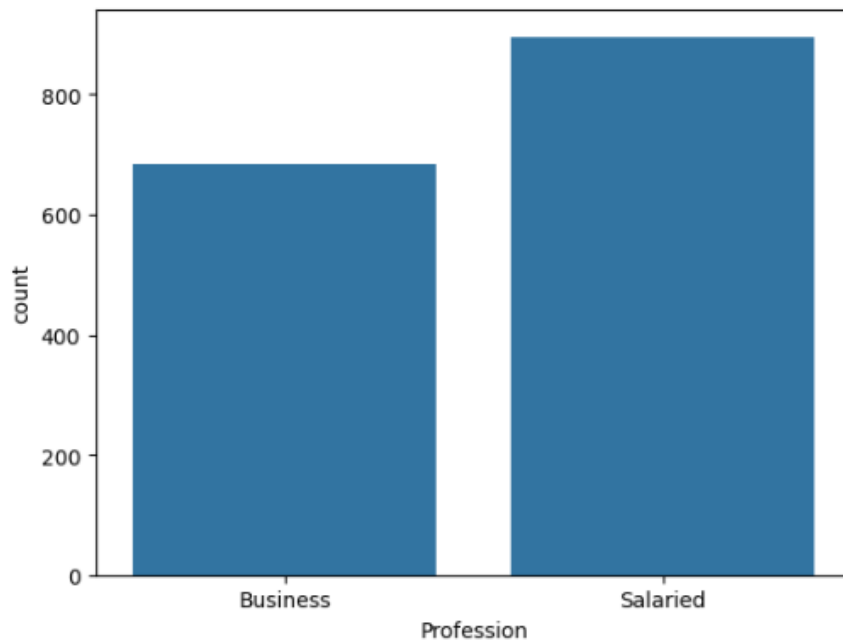
### 2.2.1 Observations on Gender



- There are more males than females in the dataset.

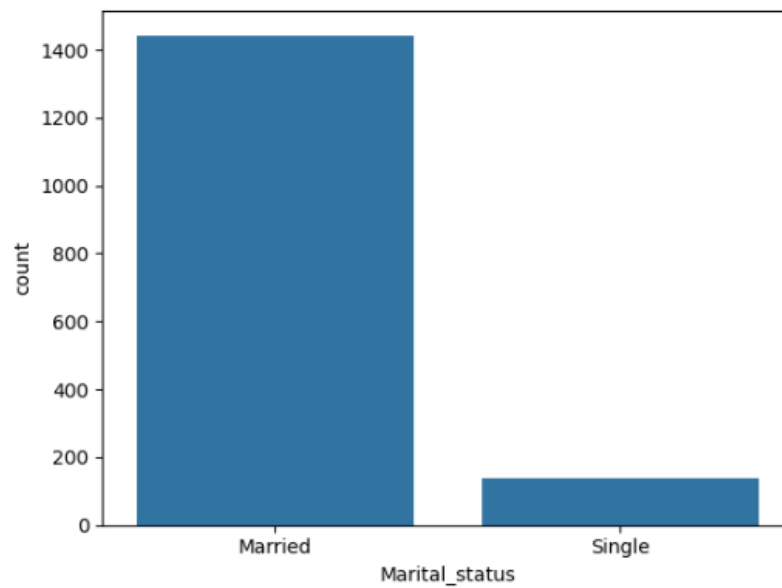


### 2.2.2. Observations on Profession



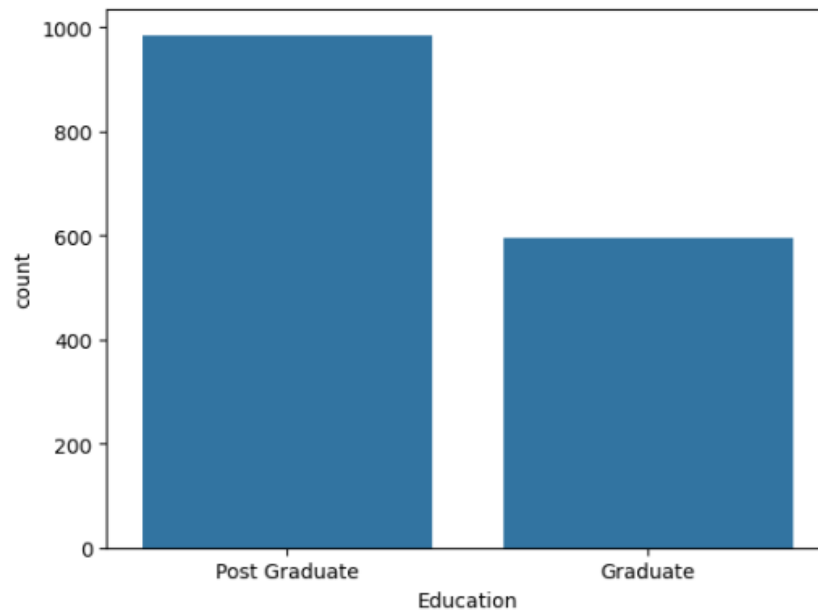
- Most people are salaried in the dataset.

### 2.2.3. Observations on Marital status



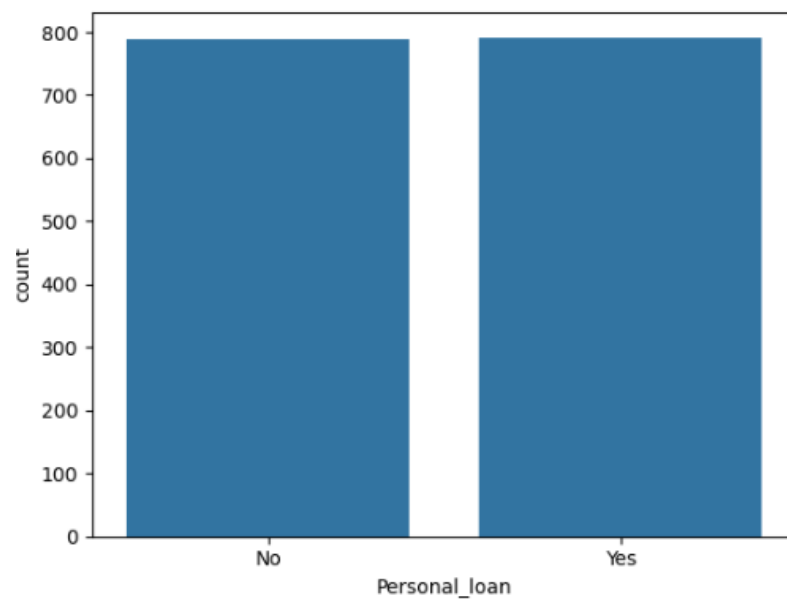
- Most of the people are married.

#### 2.2.4. Observations on Education



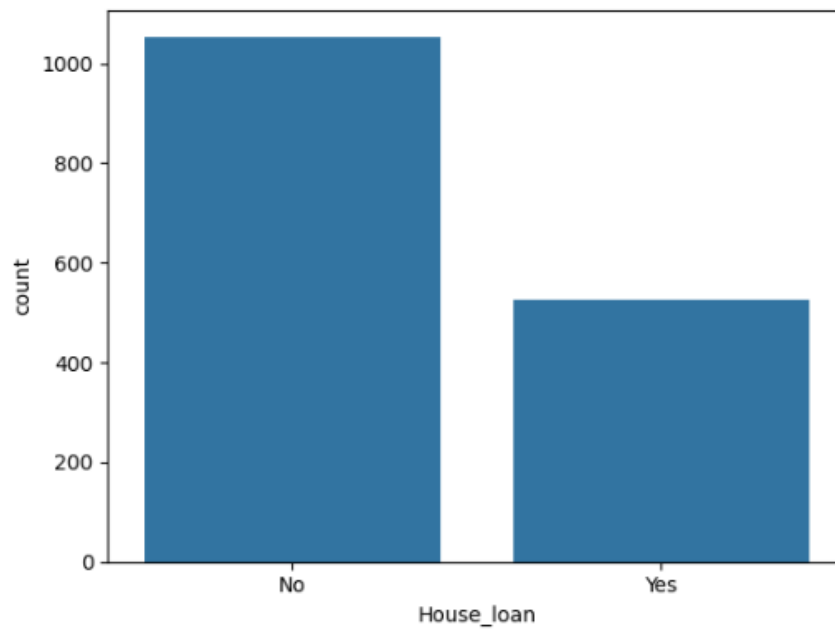
- More people hold post-graduate degrees in the dataset.

#### 2.2.5. Observations on Personal Loans



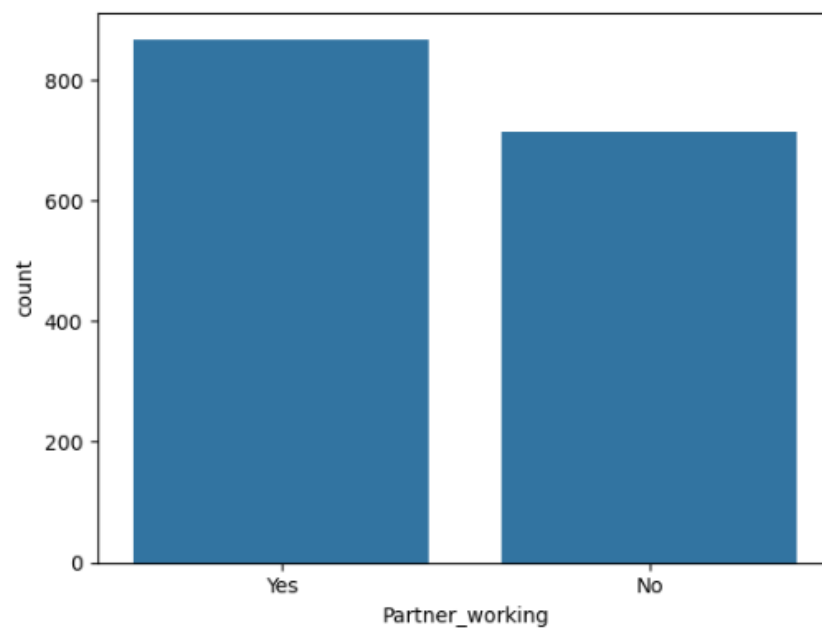
- 50% of the people have taken a Personal Loan.

### 2.2.6. Observations on House Loans



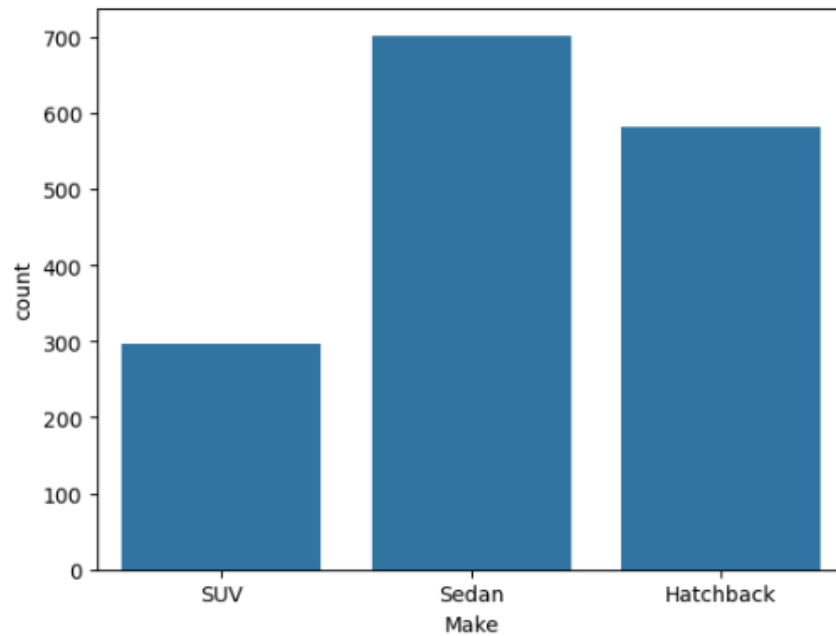
- 33% of the people have taken a House Loan.

### 2.2.7. Observations on Partner working



- More than 800 people have a working partner.

### 2.2.8. Observations on Make



- According to the dataset, people prefer sedans, followed by hatchbacks, more than SUVs.

## 3. Bivariate Analysis

### 3.1. Relationship between all Numeric Variables

We can analyse this by using a Scatterplot, Pairplot, and Correlation.

- There is a linear relationship between all the numerical columns (Age, No. of Dependents, Salary, Partner salary, Total salary, and Price) of the dataset.

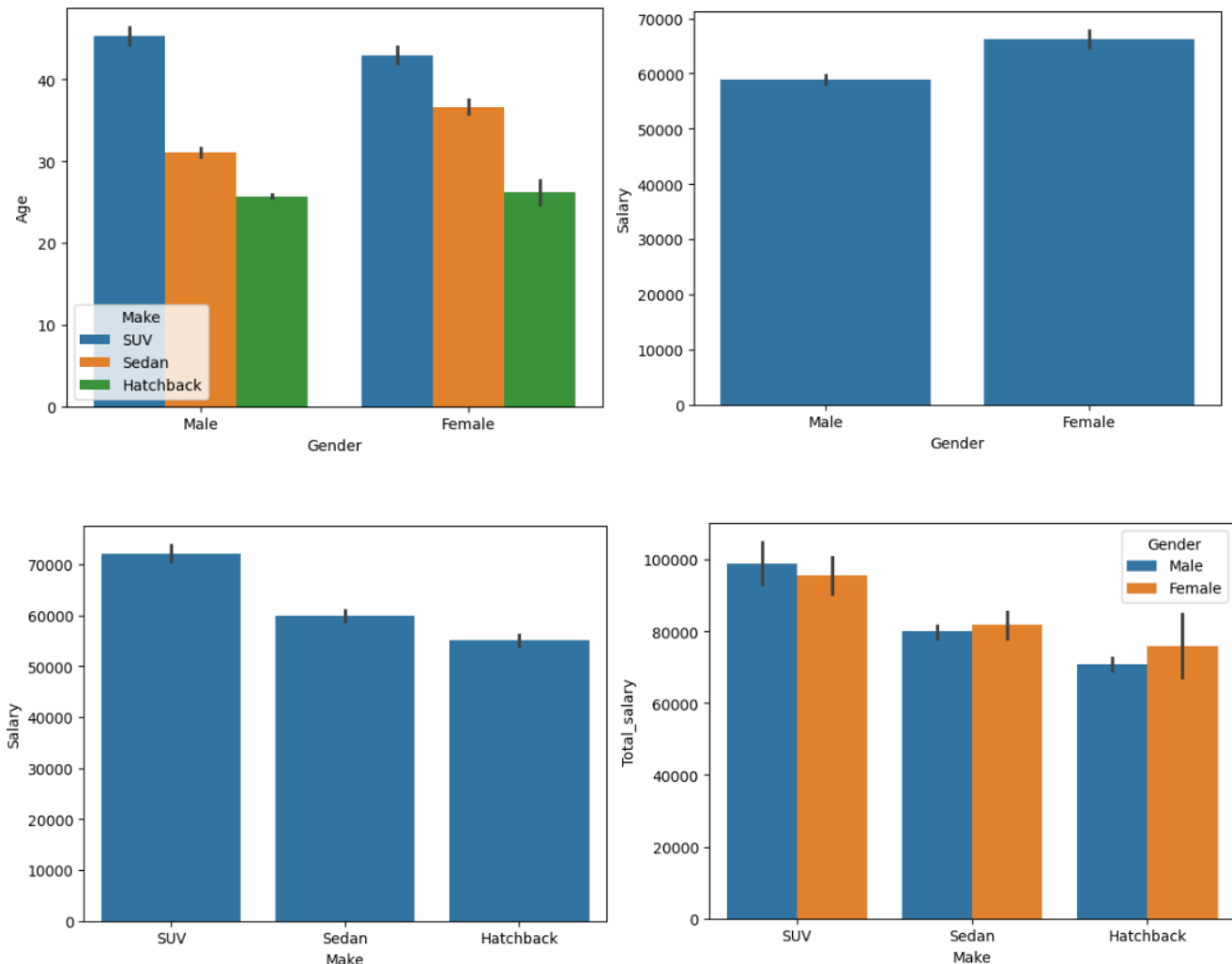
### 3.2 Correlation between all Numeric Variables

We can analyse correlation by using Heatmap.



- There are strong linear correlations between 'Salary' vs 'Total salary' and 'partner salary' vs 'Total salary'.
- 'No.o Dependents' seem to have discrete values instead of a continuous range.

### 3.3 Relationship between Categorical vs Numeric variables

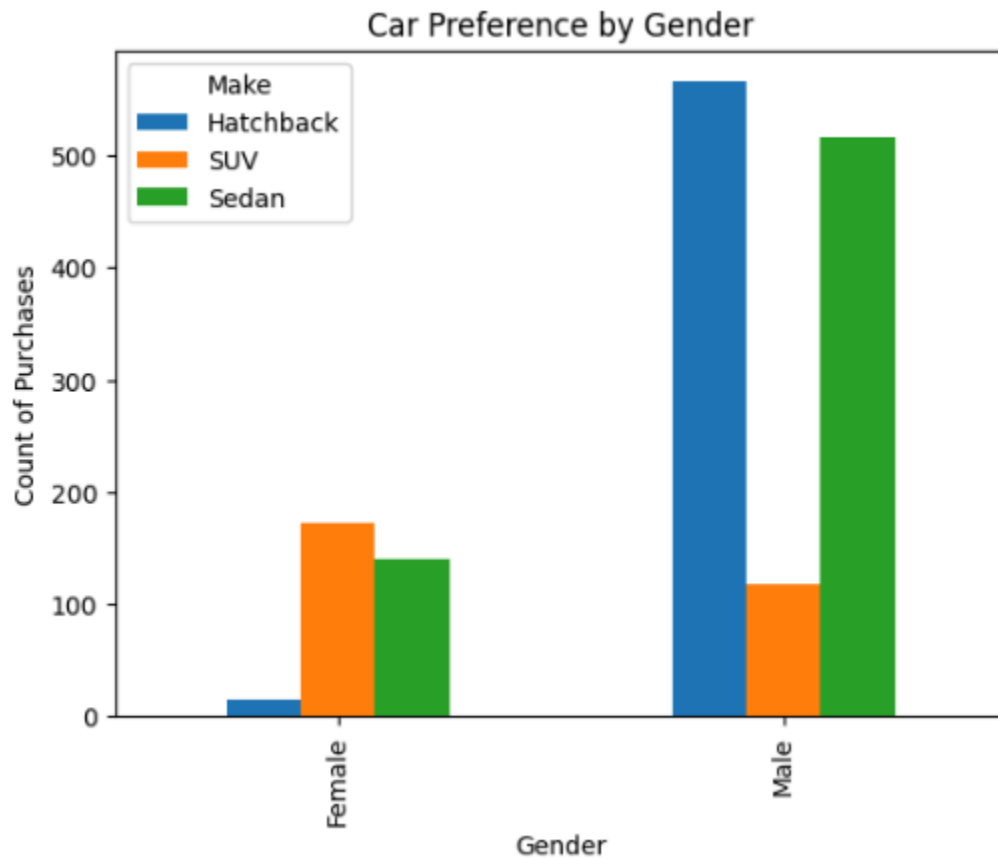


#### Observations :

- Older individuals prefer SUVs, middle-aged individuals prefer Sedans, and younger individuals prefer Hatchbacks.
- Salary distribution is fairly equal between males and females.
- Higher-income individuals prefer SUVs, middle-income individuals prefer Sedans, and lower-income people prefer Hatchbacks because they are budget-friendly.
- Gender based salary gap is minimal across vehicle categories.

### 3 . Key Questions

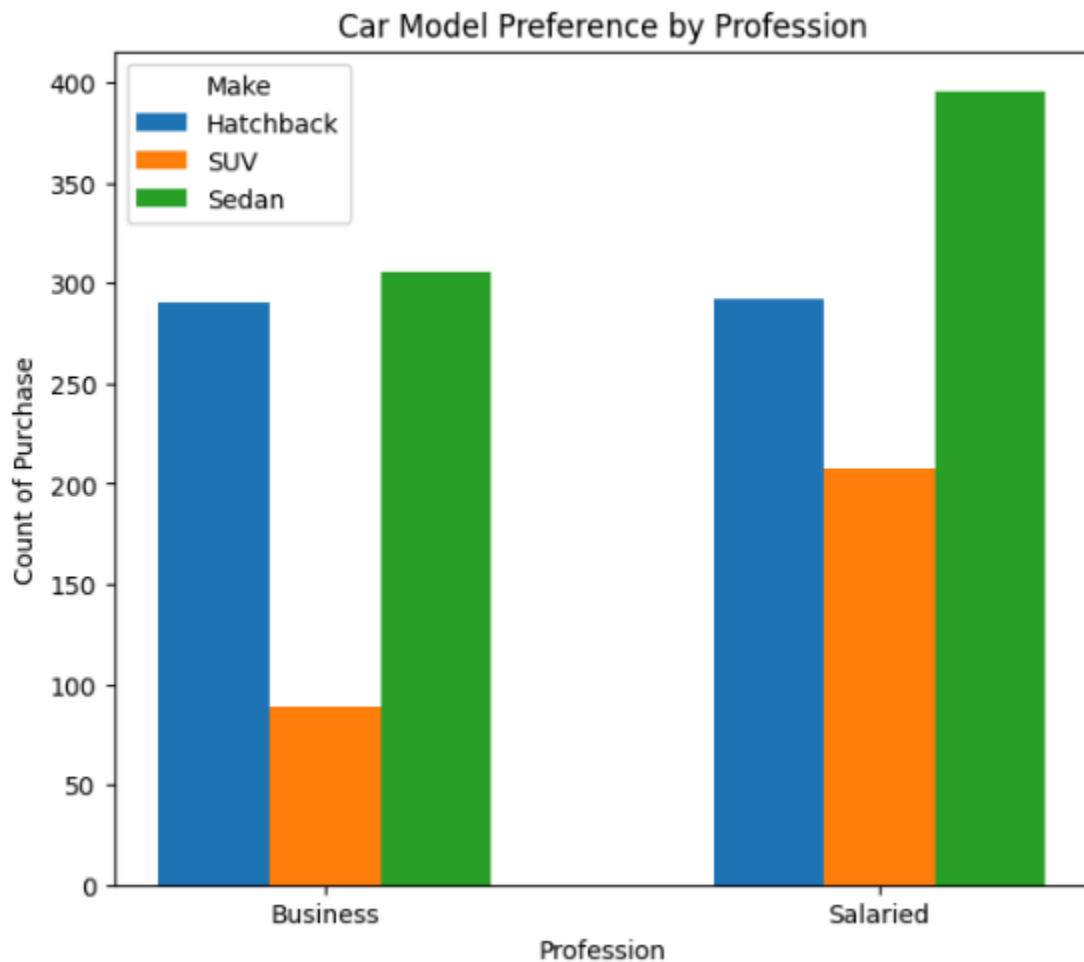
#### 1. Do men tend to prefer SUVs more than women?



- Females prefer SUVs more than Males.
- Males prefer Hatchbacks more than SUVs and Sedans.
- Females prefer Sedans after SUVs.
- Males buy more cars overall.

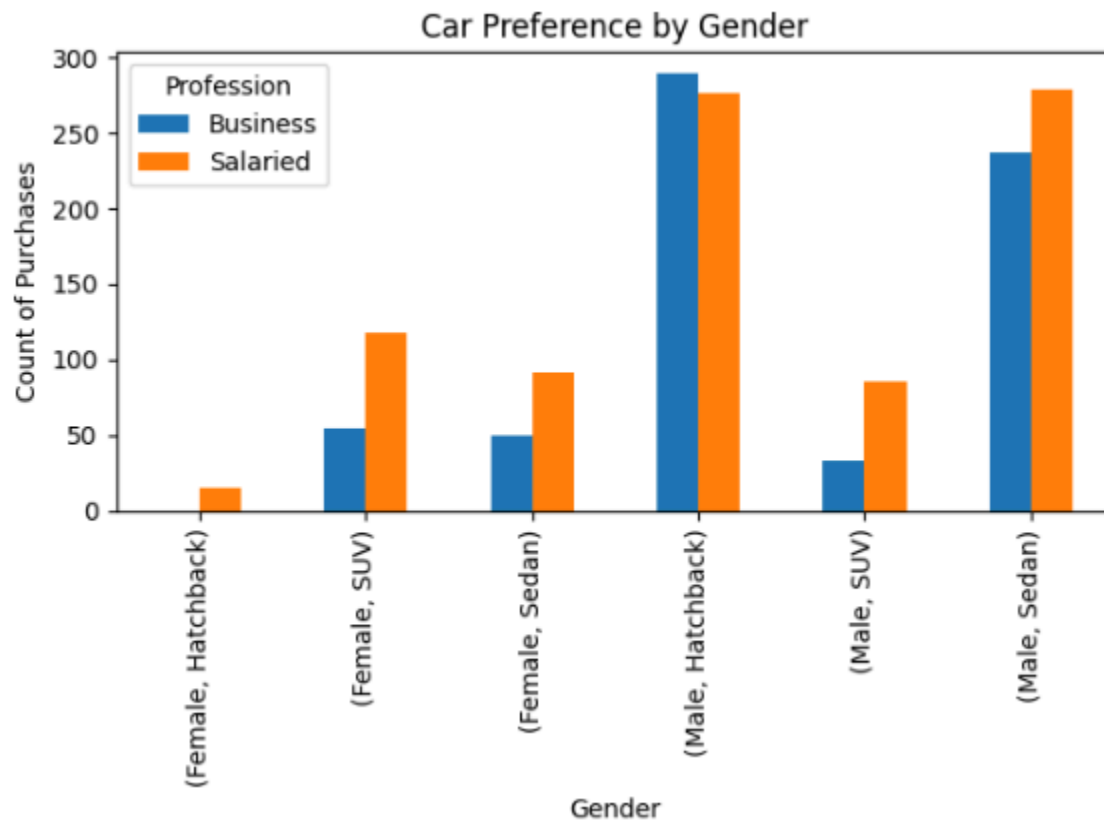


## 2. What is the likelihood of a salaried person buying a Sedan?



- Likelihood of a salaried person buying a Sedan: 0.44
- Salaried people prefer Sedans more than SUVs and Hatchbacks.
- Hatchback is popular in both professions.
- SUVs are least preferred in both professions.
- Business professionals purchase almost equal no. of Hatchbacks.

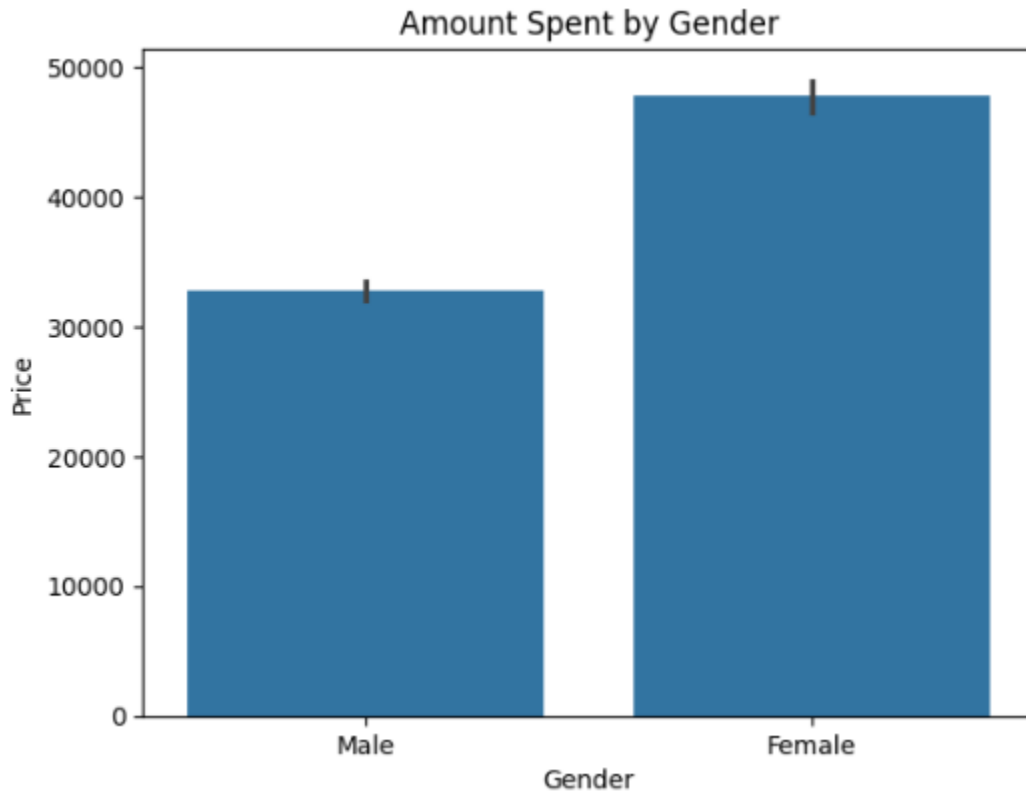
**3. What evidence or data supports Sheldon Cooper's claim that a salaried male is an easier target for a SUV sale over a Sedan sale?**



**Sheldon Cooper's Statement is False as**

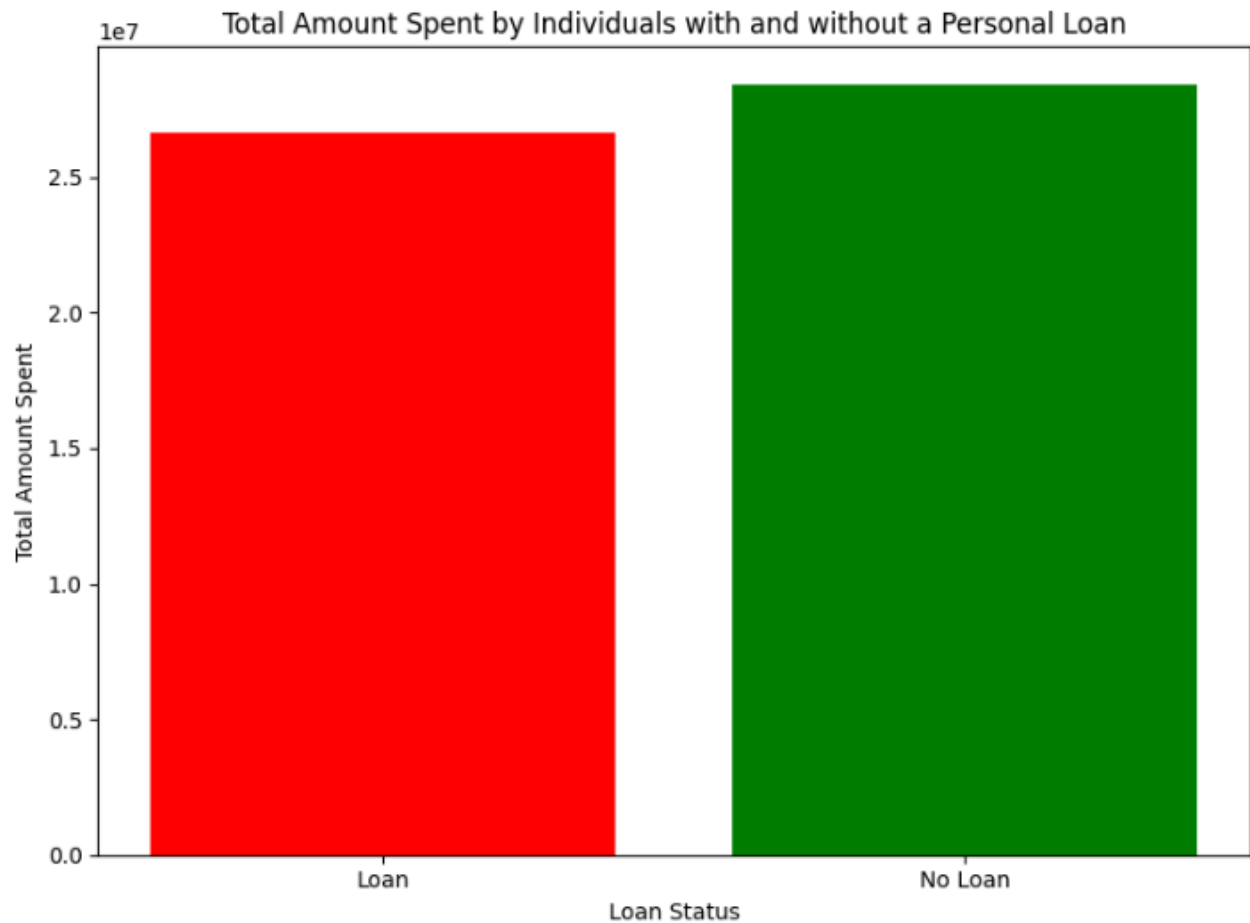
- Salaried Males prefer Sedans over SUVs.
- Both business and salaried males prefer Hatchbacks the most.
- Salaried females prefer SUVs more than Hatchbacks and Sedans.

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



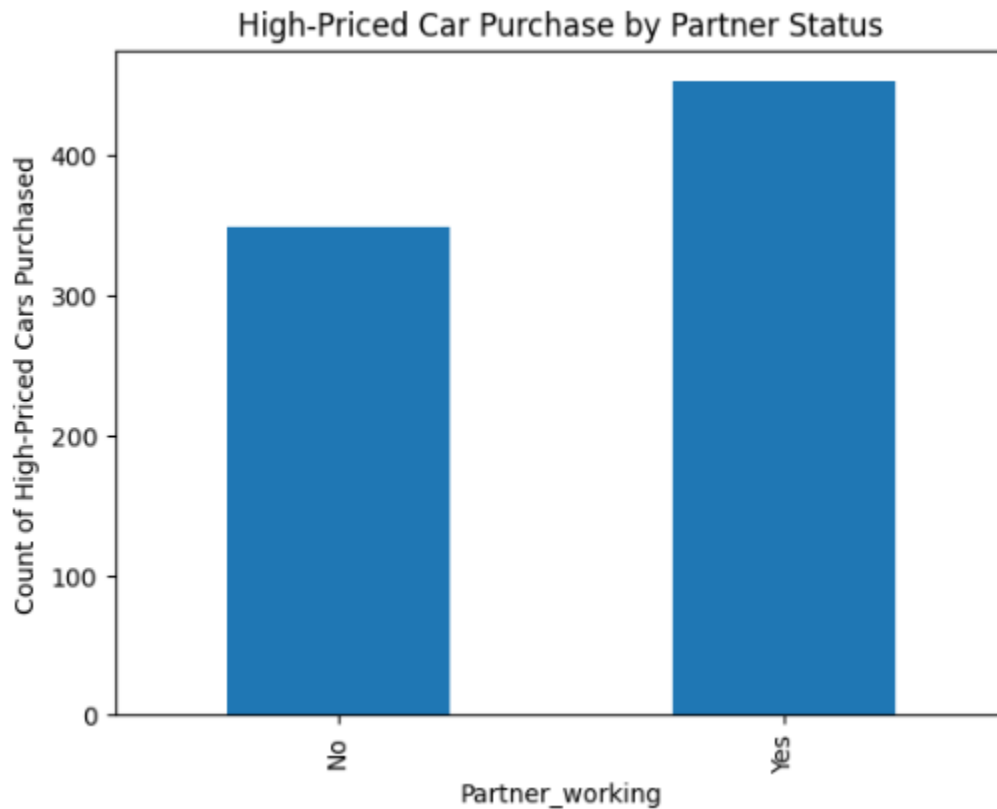
- Female spends approximately = 47705.167173 and Male = 32817.347790
- Females spend more money on cars than Males.

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



- Total amount spent by individuals who took a loan: \$26644000
- Individuals without a personal loan tend to spend slightly more on average than those with a personal loan.

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



- Partner\_working  
No = 349  
Yes = 452
- Individuals with a working partner tend to purchase more high-priced cars compared to those without a working partner.
- Dual-income households are more likely to afford high-priced cars.

## Conclusions

- Car preferences are influenced by various factors, including age, income, gender, profession, and whether the individual has a working partner.
- Females tend to spend more and prefer SUVs, while Males prefer Hatchbacks.
- Salaried individuals generally prefer Sedans, while Business professionals show a stronger preference for Hatchbacks.
- People with dual incomes tend to buy higher-priced cars, and individuals without loans generally spend more on their car purchases than those with loans.
- Females generally spend more money on cars than Males. The average spending for Females is significantly higher, reflecting their preference for higher-priced cars compared to Males, who generally spend less.

## Recommendations

- The company should personalise offers for women on key occasions, highlighting luxury and long-term value for SUVs.
- Highlight the unique features of the cars through distinctive advertising to effectively attract potential buyers.
- Target Older & High-Income Buyers for SUVs by emphasizing luxury and safety.
- Highlight Hatchbacks as affordable, budget-friendly cars for younger buyers and lower-income groups.
- Promote Sedans to salaried and middle-income individuals with a Focus on value and fuel efficiency.
- Provide additional offers on car servicing and offer free or extra discounts on car accessories to attract more buyers.
- Promote Family-Friendly Features in SUVs to attract families looking for space and safety.
- Focus on Dual-Income Households by offering higher-priced vehicles like SUVs with financing options.
- Offer Flexible Financing to those with personal loans, making car purchases easier.