AUSTO AUTOMOBILE BUSINESS REPORT

PGP DSA 18 MAY 2025

ANJU SAINI

CONTENT

1) Introduction:-

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.

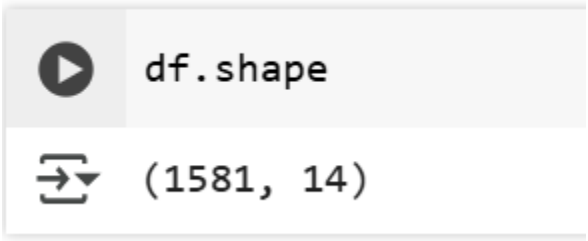
2) Objective:-

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 to improve the business.

3) Data Overview:-

Understanding the data and Structure

- Data Contains 1581 Rows and 14 Columns



- Data Type of Different columns

```
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)
```

4) Statistical Summary

```
df.describe().T
```

	count	mean	std	min	25%	50%	75%	max
Age	1581.0	31.922201	8.425978	22.0	25.0	29.0	38.0	54.0
No_of_Dependents	1581.0	2.457938	0.943483	0.0	2.0	2.0	3.0	4.0
Salary	1581.0	60392.220114	14674.825044	30000.0	51900.0	59500.0	71800.0	99300.0
Partner_salary	1475.0	20225.559322	19573.149277	0.0	0.0	25600.0	38300.0	80500.0
Total_salary	1581.0	79625.996205	25545.857768	30000.0	60500.0	78000.0	95900.0	171000.0
Price	1581.0	35597.722960	13633.636545	18000.0	25000.0	31000.0	47000.0	70000.0

Observation

- 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
- Average price at which the vehicle is bought is ~36000 and highest price is 70000

5) Checking Duplicate Data

```
[12] df.duplicated().sum()
```

```
np.int64(0)
```

There are no duplicate rows in the Data

6) Checking Missing Data

```
df.isnull().sum()
```

	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	106
Total_salary	0
Price	0
Make	0

Based on the given response, data for two columns, Gender and Partner_salary, appears to be missing

6.1) Analysis Partner_salary missing values

Partner_salary	
0.0	603
NaN	103
40500.0	28
40200.0	27
40000.0	25
...	...
22300.0	1
100.0	1

From the give response its show that there are103 fields that is nan

Total_Salary=salary+Partner_Salary so lets add 1 more column that get the correct value of 103 missing partner_salary

7) Data Correction

In the Gender column, there is some corrupted data. After analyzing the responses, it appears that "female" is misspelled in some rows (e.g., as "Femal" or "Femle")

Gender	
Male	1199
Female	327
Femal	1
Femle	1

Gender	
Male	1199
Female	329

7.1) Checking null value count for Gender

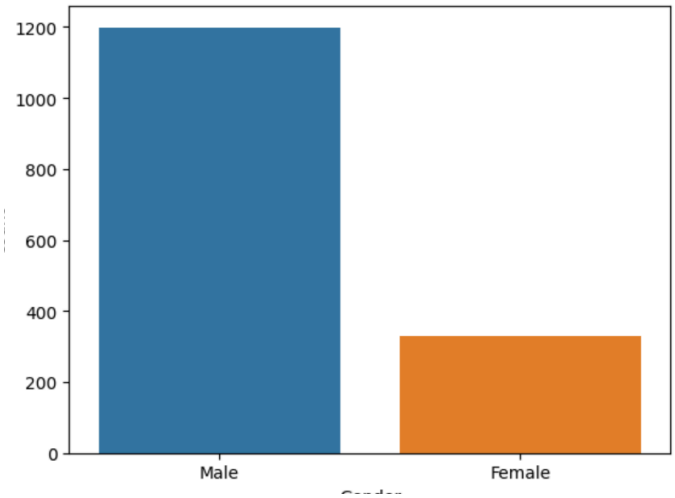
Gender	
Male	1199
Female	329
NaN	53

Total 53 Values is missing in gender

Drop data from a file where gender is nan . because in subject it's given there is only 2 type of gender Male and Female

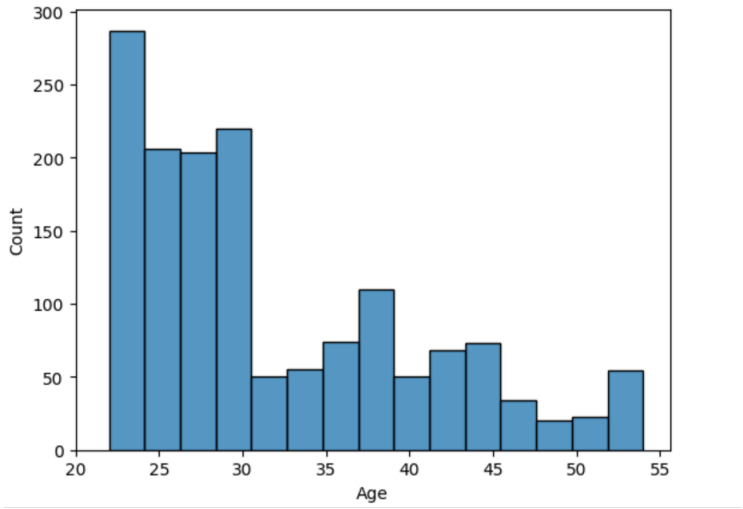
8) Exploratory Data Analysis

8.1) Gender



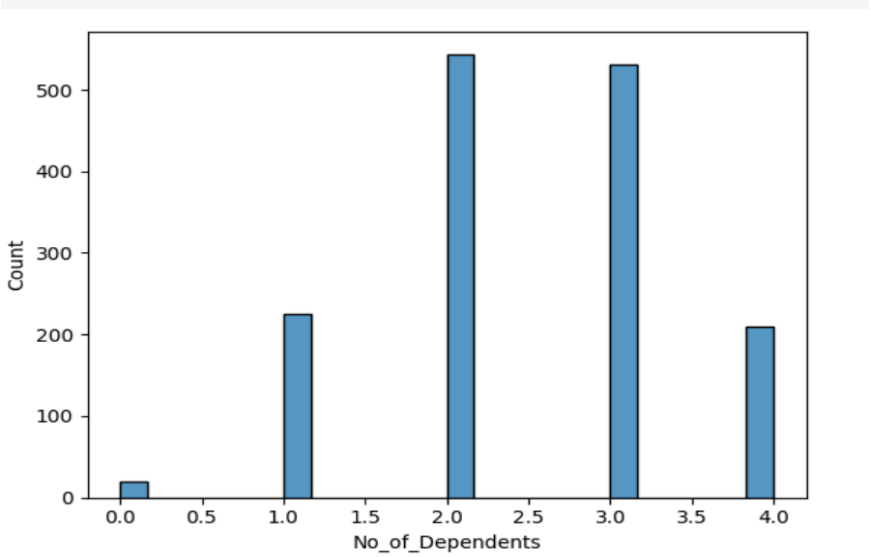
The number of Male is more then Number of Female.

8.2) Age



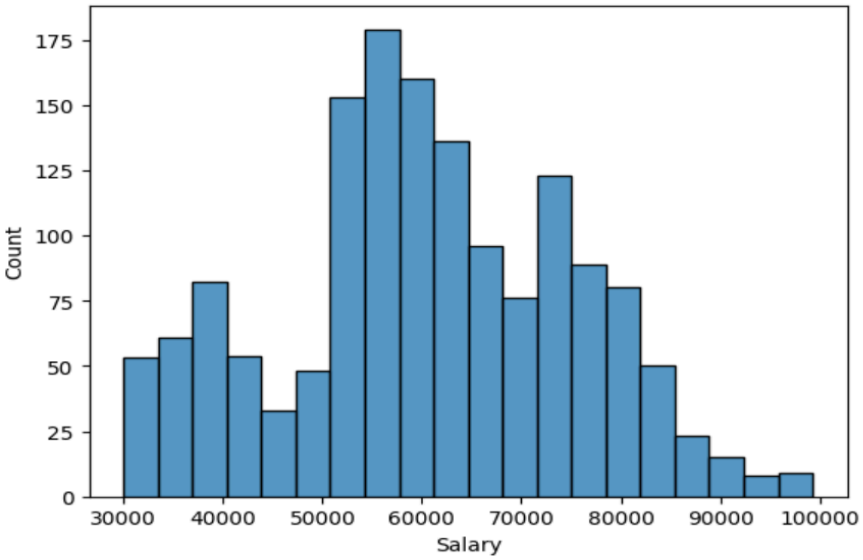
Based on the given histogram, the age range is between 23 and 54. Most of the target audience or buyers are between 23 and 30 years old, indicating that this age group shows the highest interest in purchasing the product or service

8.3) Number of Dependents



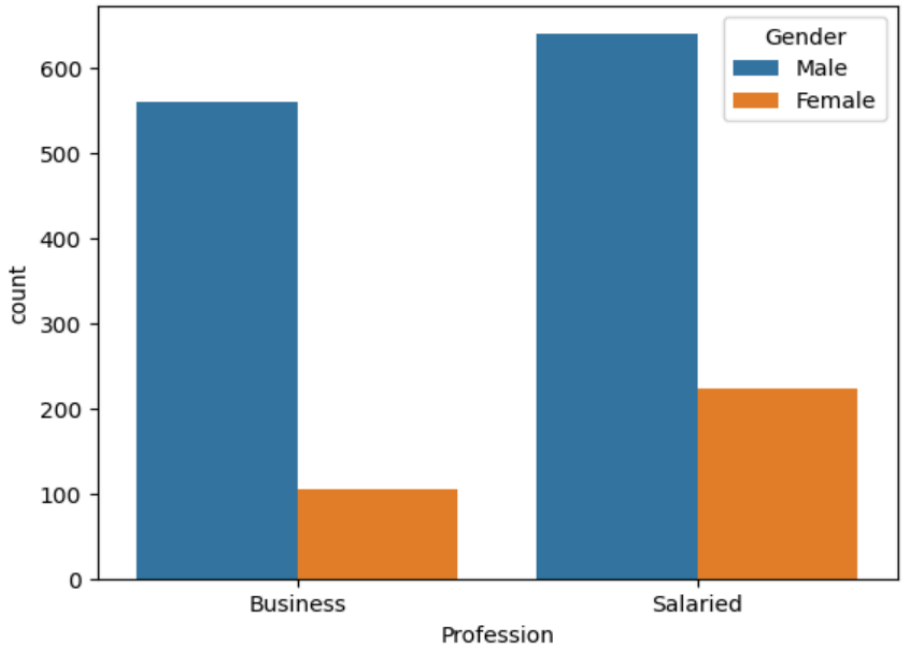
From The given response we can say that Average no of dependents are 2 -3

8.4) Salary



From the Above response we can get that majority of the Employee Salary is in between to 50000 to 65000

8.5) Gender Profession



from the given response we can get that most of the people are salaried employee for both the category in business man count is more than women

8.6) Loan vs Price

House Loan vs Car Price 0 = No House Loan, 1 = Has House Loan

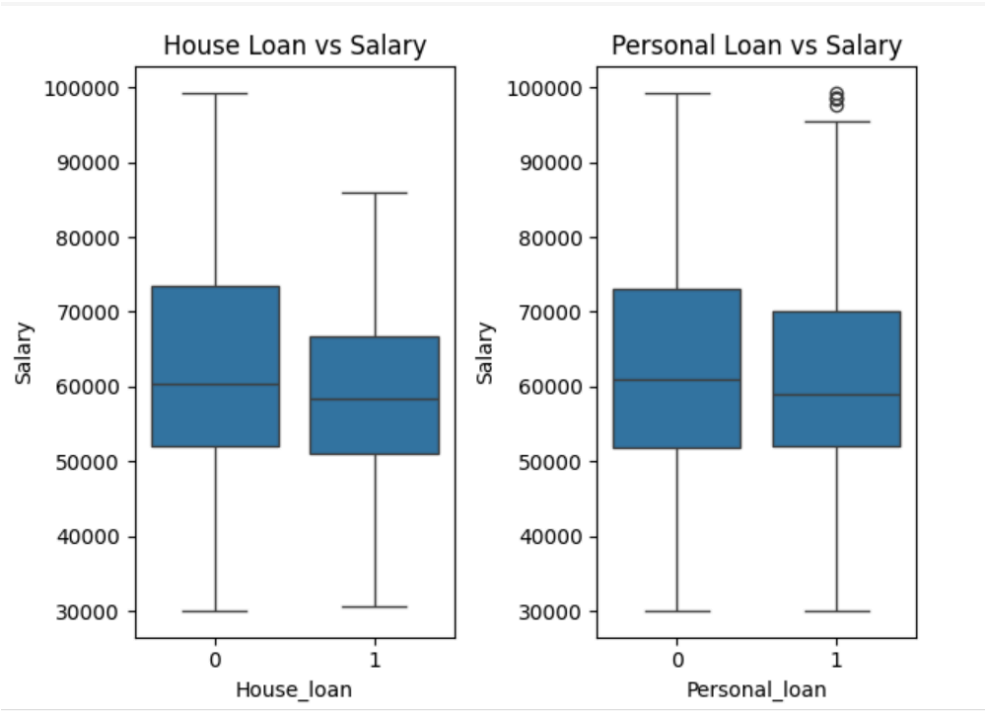
Median price for people without a house loan (0) is higher than those with a house loan (1).

Distribution is wider when no house loan — suggests more expensive cars purchased.

Personal Loan vs Car Price Similar pattern: people without personal loans tend to buy higher-priced cars.

Median price for group 0 (no loan) is slightly higher

8.7) Loan vs Salary



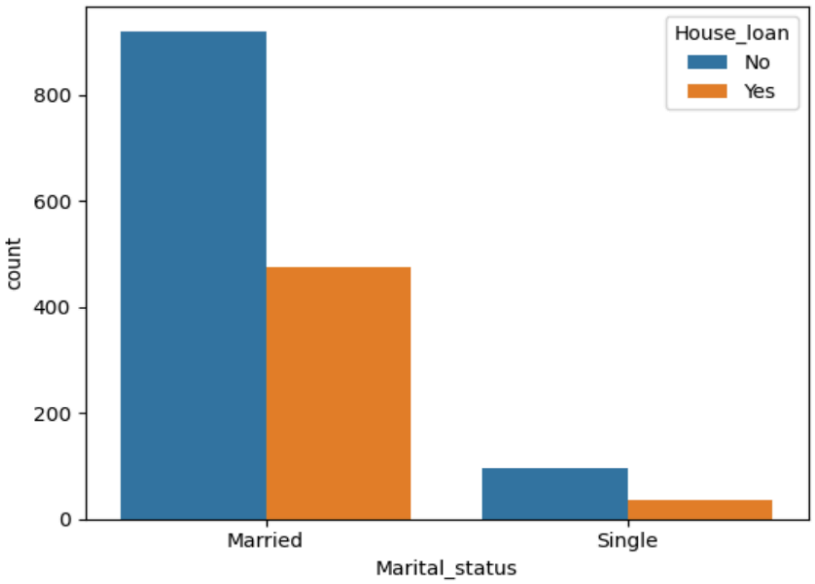
House Loan vs Salary Individuals without a house loan (0) tend to have a higher median salary and wider salary range.

Those with a house loan (1) have lower median salaries, indicating that individuals with lower incomes may still opt for housing loans, possibly due to long-term investment goals or housing needs.

Personal Loan vs Salary Similarly, individuals without a personal loan (0) tend to have slightly higher salaries than those with one.

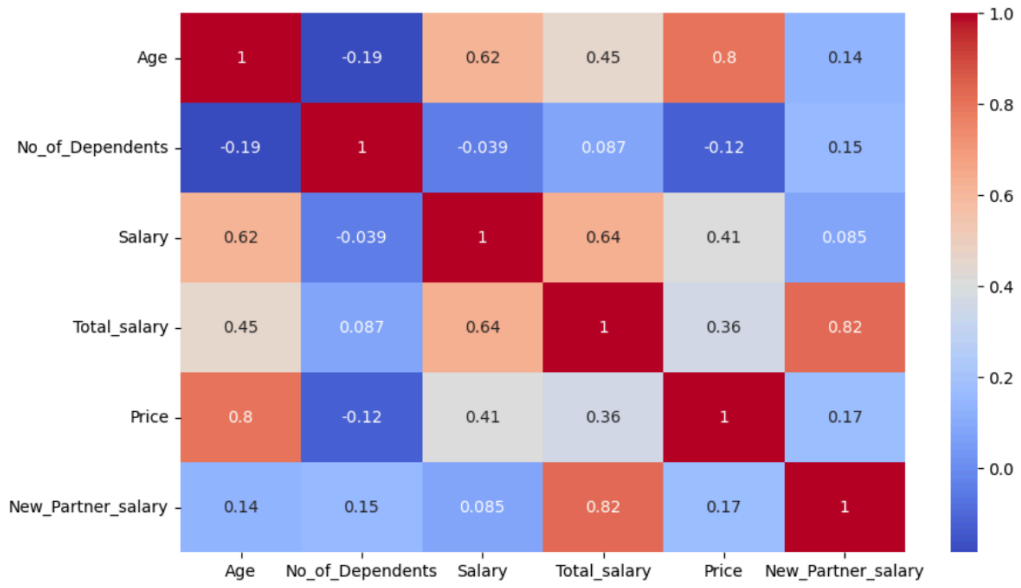
The difference is not as prominent as in the case of house loans, but it still shows that personal loans are more common among slightly lower-salaried individuals, perhaps for immediate financial needs.

8.8) Hoser loan vs Marital Status



Analysis of the dataset reveals a significant pattern—**married individuals are more likely to take house loans compared to single individuals**. This trend is consistent across income brackets and appears to be influenced by both financial and lifestyle factors.

9)Relationship Between Numerical Column



Age strongly influences car price: There's a strong positive relationship (0.8) between age and the price of a car, suggesting that older individuals tend to purchase more expensive vehicles.

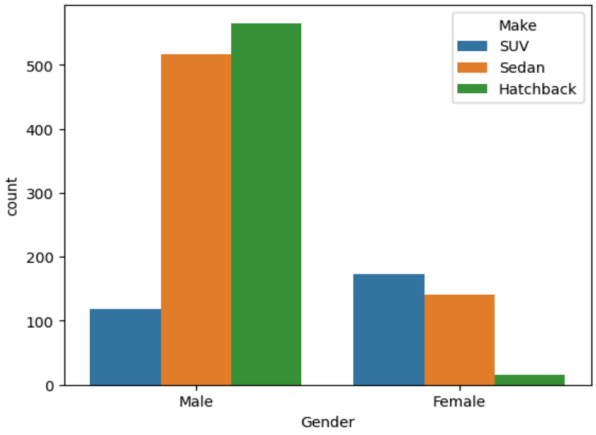
Total salary and new partner salary are closely linked: As expected, total salary and the new partner's salary show a high positive correlation (0.82), likely due to how total salary is calculated.

Individual salary contributes to total earnings: There's a positive relationship between an individual's salary and their total salary.

Price is driven by age and income: The price of a car has a notable positive correlation with age (0.8), total salary (0.36), and individual salary (0.41). Key

10) Key Insights

1: Do men tend to prefer SUVs more compared to women?

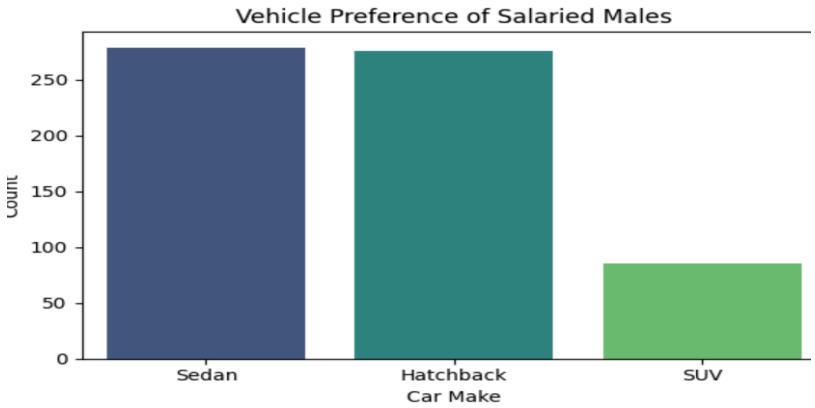


Most of the male like Hatchback and female like Suv

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

```
Make
Sedan        279
Hatchback    276
SUV           85
Name: count, dtype: int64
```

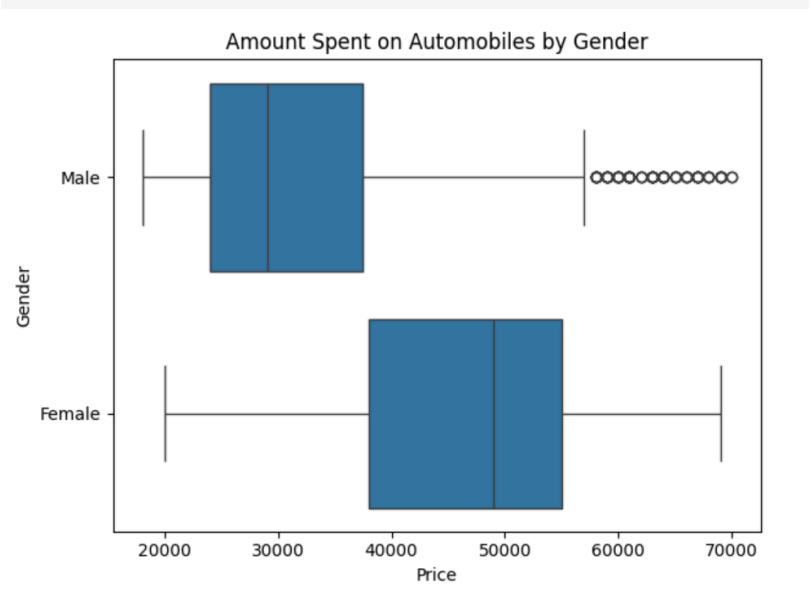
MOST OF THE Mans that are salaried employee is like to purchase Sedan as compared to SUV



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?

This data does not support Sheldon Cooper's claim. Sedans and Hatchbacks are far more popular among salaried males than SUVs**.**

4: **How does the amount spent on purchasing automobiles vary by gender?**



Females tend to spend more on automobiles than males.

The median purchase price is higher for females.

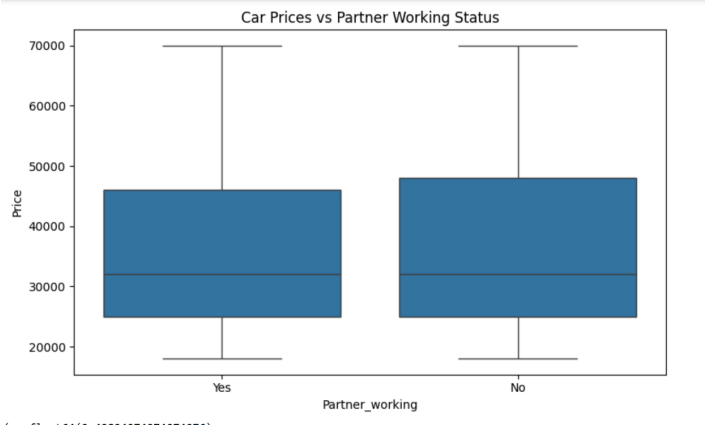
Males have more extreme outliers, but their average or typical purchase is lower.

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

```
[ ]  loan_total_spent = df[df['Personal_loan'] == 'Yes']['Price'].sum()
     print(loan_total_spent)
```

```
26644000
```

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



The median price is the same (₹31,000) in both groups.

People without a working partner spend slightly more on average.

11) Observations:

From the above chart the marketing strategy can be devised to send targeted information to different groups of potential buyers present in the data i.e.

1. As Married Males tend to buy more of 'Sedan' & 'Hatchback' may be due to budget constraints and other liabilities like 'Personal loans' & 'house loans . Multiple schemes can be provided to them for discounts on 'SUVs' . Which can come under their budget and is more spacious as per the family requirements.

2. Single Males tend to buy more 'Hatchback' cars as compared to other two categories. Marketing strategy can be devised in a way to target only single males for increasing sales of 'hatchback' cars.

3. Married females tend to buy more 'SUVs' & 'Sedans' as cars as compared to 'hatchback'. Marketing strategy can be devised in a way to target only married females for increasing sales of 'SUVs' & 'Sedan

12) Business Recommendations

1. Targeted Marketing by Gender & Marital Status

- Married males prefer Sedans & Hatchbacks, likely due to budget and loan obligations.
- Single males lean heavily towards Hatchbacks.
- Married females show interest in SUVs & Sedans.

2. Income-Based Product Positioning

- Individuals with higher total salaries tend to purchase higher-priced cars.
- However, loan obligations affect this pattern.

3. Focus on Family-Oriented Campaigns

- Number of dependents correlates with vehicle type (SUVs preferred for larger families).

4. Segmented Campaigns by Age & Profession

- Buyers aged 23–30 dominate the market.
- Salaried professionals are more likely to purchase Sedans and Hatchbacks.

5. Broaden Outreach Across Gender

- Males dominate the customer base.

6. Leverage Partner Income & Working Status

- Partner's working status impacts vehicle choice and price.

13) Conclusion

Following are the suggestions provided to the Marketing team for improving the campaign.

1.Defining the Target Audience - The first step is to define the Target Audience for the Product. Based On the observation made above, it is clear, the Target audience were skewed below the age of 30.The marketing team should make sure their Target audience is fixed and try to showcase all the type of cars and wider age group

2.The Target audience was gender specific (dominated by Male). The Marketing team should in Future Marketing concentrate on Gender equality and not gender sensitivity.

3.The Sample data inferred that Hatchback and Sedan Cars were sold the maximum. Moreover, the sample data also suggested, based on total salary and Car type, there can be cross selling and upselling done with higher total salary.

4.The future marketing campaign should segment the cars and do different Market campaign based on the age group, income levels and number of dependents

5.The marketing drive should target the entire family to sell the car as collective decision plays a major role, if both the family members are working.

Thank You!

Prepared by:

Anju Saini

PGP-DSA Batch | 18 May 2025