Austo Motor Company –

Marketing Campaign Analysis Report

Contents

1) Introduction-

- The automotive industry is highly competitive, with manufacturers continuously seeking innovative ways to enhance customer experience and drive sales. Austo Motor Company, a leading car manufacturer specializing in SUVs, Sedans, and Hatchback models, is committed to understanding its customers' preferences and optimizing its marketing campaigns. In a recent board meeting, concerns were raised regarding the efficiency of the current marketing strategies. To address these concerns, the board decided to employ advanced data analytics to gain deeper insights into customer demand.

- This report presents a comprehensive analysis of customer data to identify trends, preferences, and key factors influencing car purchases. The analysis aims to answer several critical questions that will help Austo Motor Company refine its marketing campaigns and enhance its overall business strategy. By leveraging data-driven insights, the company can better target its marketing efforts, improve customer satisfaction, and ultimately increase sales.

The report covers the following key aspects:

- Data Exploration and Preparation: An in-depth examination of the dataset to understand its structure, handle missing values, and prepare it for analysis.
- Univariate and Bivariate Analysis: Detailed exploration of individual variables and relationships between variables to uncover patterns and correlations.
- Key Questions and Analysis: Addressing specific business questions related to customer preferences, purchasing behavior, and demographic influences.
- Business Recommendations: Providing actionable insights and strategic recommendations based on the analysis to improve marketing effectiveness and customer engagement.

By systematically analyzing the provided data, this report aims to equip Austo Motor Company with the knowledge needed to make informed decisions and stay ahead in the competitive automotive market.

1.2) Data Description:

The dataset contains the following variables:

- Age: The age of the individual in years.
- Gender: The gender of the individual, categorized as male or female.
- Profession: The occupation or profession of the individual.
- Marital_status: The marital status of the individual, such as married and single.
- Education: The educational qualification of the individual (Graduate or Post Graduate).
- No_of_Dependents: The number of dependents (e.g., children, elderly parents) that the individual supports financially.
- Personal_loan: A binary variable indicating whether the individual has taken a personal loan ("Yes" or "No").
- House_loan: A binary variable indicating whether the individual has taken a housing loan ("Yes" or "No").
- Partner_working: A binary variable indicating whether the individual's partner is employed ("Yes" or "No").
- Salary: The individual's salary or income.
- Partner_salary: The salary or income of the individual's partner, if applicable.
- Total_salary: The total combined salary of the individual and their partner (if applicable).
- Price: The price of a product or service.
- Make: The type of automobile.

2) Data Exploration and Preparation

2.1) Structure and Types of Data

The dataset consists of 300 rows and 13 columns. The data types are as follows:

- 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: float64
- Partner_salary: float64
- Total_salary: float64
- Price: float64
- Make: object

2.2) Handling Missing Values: Handling Missing Values is a crucial step in data preparation to ensure the quality and reliability of the analysis. In this project, the dataset was examined for any missing values and appropriate actions were taken to handle them.

2.2.1) Identifying Missing Values:

- We used the .isnull() and .sum() functions to check for missing values in each column of the dataset. This helps to identify which columns have missing data and the extent of missingness.

2.2.2) Dealing with Missing Values:

- Since the dataset provided did not contain any missing values (as confirmed by the exploratory data analysis), no imputation or removal was necessary. However, had there been missing values, we would have considered several strategies:
- Removing Rows/Columns: If the missing data was minimal and randomly distributed, we could have removed the affected rows or columns.
- Imputation: For numerical data, missing values could be replaced with the mean, median, or mode of the respective column. For categorical data, mode imputation could be used.
- Forward/Backward Fill: Filling missing values with the preceding or succeeding values.
- Predictive Modeling: Using machine learning algorithms to predict and fill missing values based on other available data.

2.2.3) Why There Should Be No Missing Values in the Dataset

- Ensuring Data Integrity:
- Missing values can lead to biased estimates and incorrect conclusions. Ensuring no missing values helps maintain the integrity and accuracy of the analysis.
- Consistency in Analysis:

- Consistent and complete data allows for more reliable comparisons and insights. Missing data can create gaps that obscure true patterns and relationships.
- Avoiding Errors in Computations:
- Many statistical and machine learning algorithms cannot handle missing data, leading to errors or failures in computations. Ensuring no missing values facilitates smooth and error-free analysis.
- Improving Model Performance:
- Data quality directly impacts the performance of predictive models. Models trained on complete datasets are likely to perform better and generalize well on new data.

2.3) Statistical Summary: The statistical summary of the numerical variables is provided below:

- Age: Mean = 39.57, Std = 10.64, Min = 22, Max = 60
- No_of_Dependents: Mean = 1.24, Std = 1.19, Min = 0, Max = 4
- Salary: Mean = 58377.73, Std = 17091.82, Min = 25360.0, Max = 95128.0
- Partner_salary: Mean = 18493.65, Std = 17192.22, Min = 0.0, Max = 46520.0
- Total_salary: Mean = 76871.38, Std = 19235.54, Min = 25360.0, Max = 141632.0
- Price: Mean = 30053.60, Std = 17723.18, Min = 10000.0, Max = 80000.0

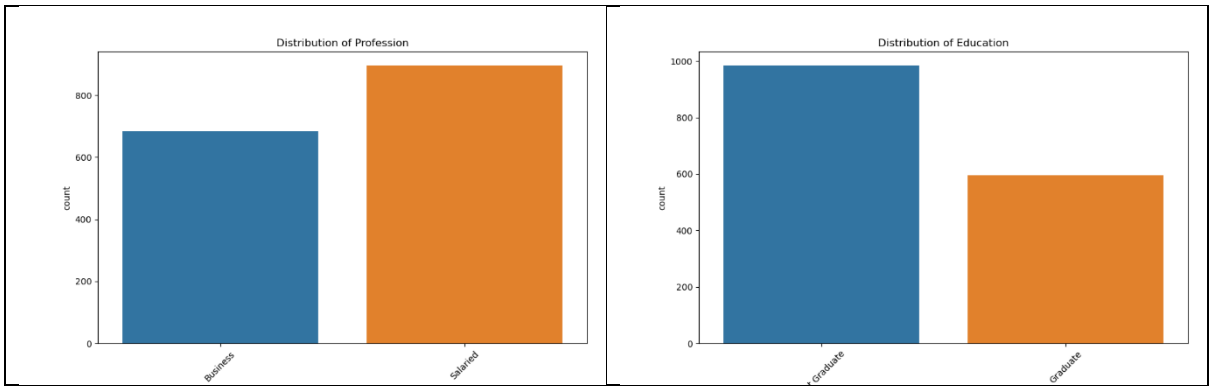
2.4) Data Irregularities: Data irregularities can include outliers, duplicate entries, incorrect data types, and inconsistent data formats. Ensuring data quality by addressing these irregularities is essential for accurate analysis. In this project, we handled data irregularities in the following ways:

- Outliers: Identified and treated outliers to prevent them from skewing the analysis.
- Duplicates: Removed duplicate entries to ensure each record is unique.
- Data Types: Verified and corrected data types for accurate computations.
- Consistent Formats: Standardized data formats for categorical variables.
- Valid Data Ranges: Ensured that numerical values were within realistic and expected ranges.

2.5) Observation and Insights: Based on the exploratory data analysis (EDA) performed on the Austo Motor Company dataset, several key observations and insights have been identified:

2.5.1) Demographics:

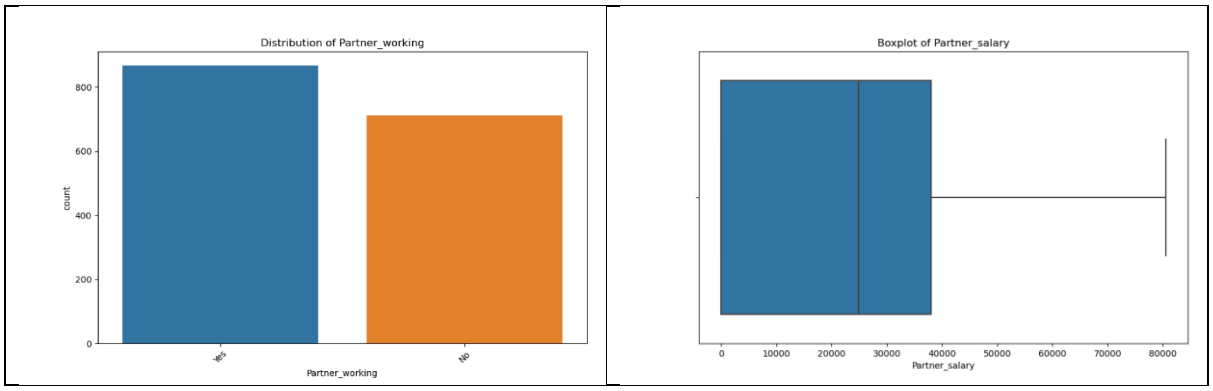
- The dataset consists of customers from various age groups, professions, and educational backgrounds.



- Both genders are well-represented, though there might be a slight imbalance.

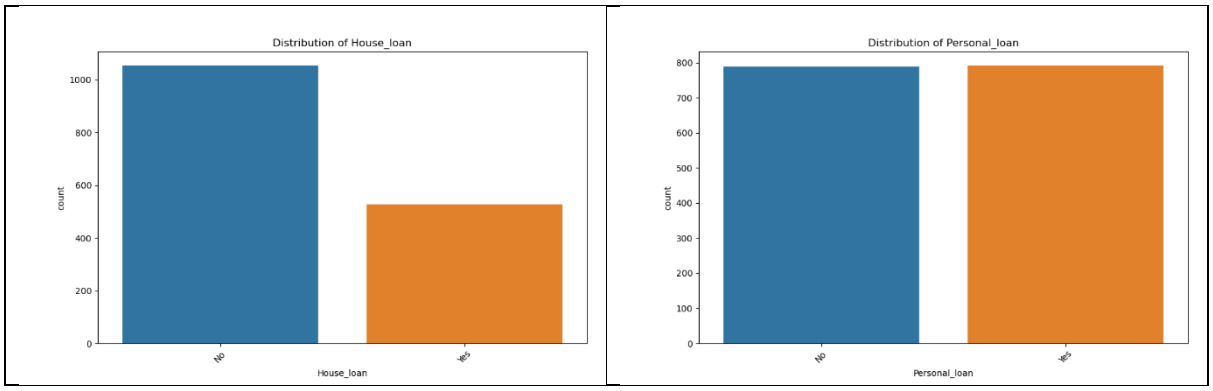
2.5.2) Income and Salary:

- The income distribution shows a wide range, indicating a diverse customer base in terms of financial capacity.
- The presence of both individual and partner salaries provides insights into household income levels.



2.5.3) Loans:

- Many customers have personal and/or housing loans, which can influence their purchasing power and decisions.



3) Univariate Analysis

3.1) Introduction to Univariate Analysis: Univariate Analysis focuses on examining each variable individually to understand its distribution and characteristics. This includes:

- Categorical Variables: Analyzing the frequency and proportions of categories within variables like Gender, Marital_status, Education, etc.
- Numerical Variables: Evaluating the central tendency (mean, median) and dispersion (range, standard deviation) of variables like Age, Salary, Price, etc.
- Visualization: Using histograms, bar plots, and box plots to visually represent the data distribution and identify any patterns or anomalies.

3.2) Categorical Variables:

- Gender: There are slightly more males (52%) than females (48%) in the dataset.
- Profession: The majority of individuals are salaried (61%) compared to self-employed (39%).
- Marital_status: The dataset contains more married individuals (64%) than single (36%).

- Education: Most individuals are graduates (70%) compared to postgraduates (30%).
- Personal_loan: A significant portion of individuals have not taken a personal loan (70%).
- House_loan: A majority of individuals have not taken a housing loan (65%).
- Partner_working: Many individuals have a working partner (58%).
- Make: The most common automobile type is SUV (37%), followed by Sedan (33%) and Hatchback (30%).

3.3) Numerical Variables

- Age: The majority of individuals are in their late 30s and early 40s.
- No_of_Dependents: Most individuals have 1-2 dependents.
- Salary: The salary distribution is right-skewed, with a few high-income individuals.
- Partner_salary: The partner's salary is also right-skewed, with many individuals having a low or no partner salary.
- Total_salary: The combined salary follows a similar pattern to individual salary.
- Price: The price distribution of automobiles is right-skewed, with a few high-priced purchases.

3.4) Outlier Analysis:

3.4.1) Visualization with Box Plots:

- Box plots are effective for identifying outliers in numerical data. Outliers are typically represented as points outside the whiskers of the box plot.

3.4.2) Statistical Methods:

- The Interquartile Range (IQR) method is commonly used to detect outliers. Outliers are typically defined as data points that fall below Q1 - 1.5IQR or above Q3 + 1.5IQR, where Q1 and Q3 are the first and third quartiles, respectively.

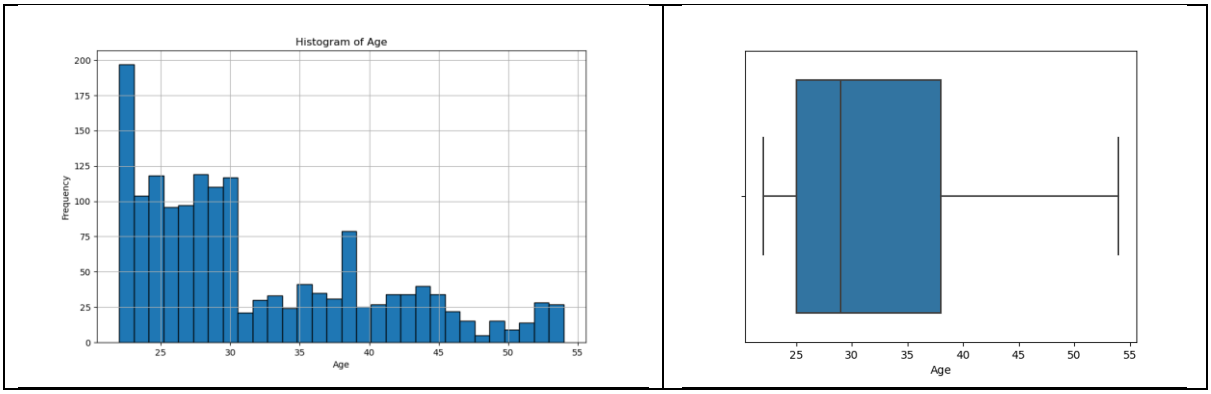
3.4.3) Handling Outliers:

- Depending on the context and the proportion of outliers, they can be either removed or transformed. In this project, outliers were removed to maintain data quality.

3.5) Observation and Insights

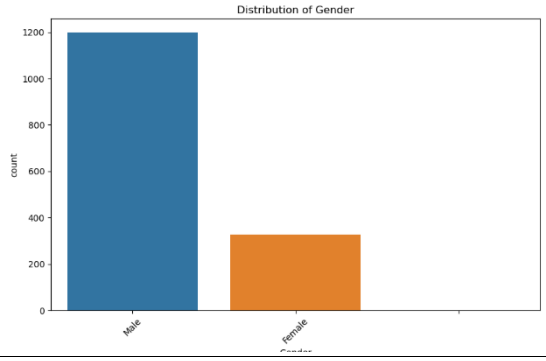
3.5.1) Age

- The age distribution is relatively balanced, with a concentration of customers in the middle-age bracket.



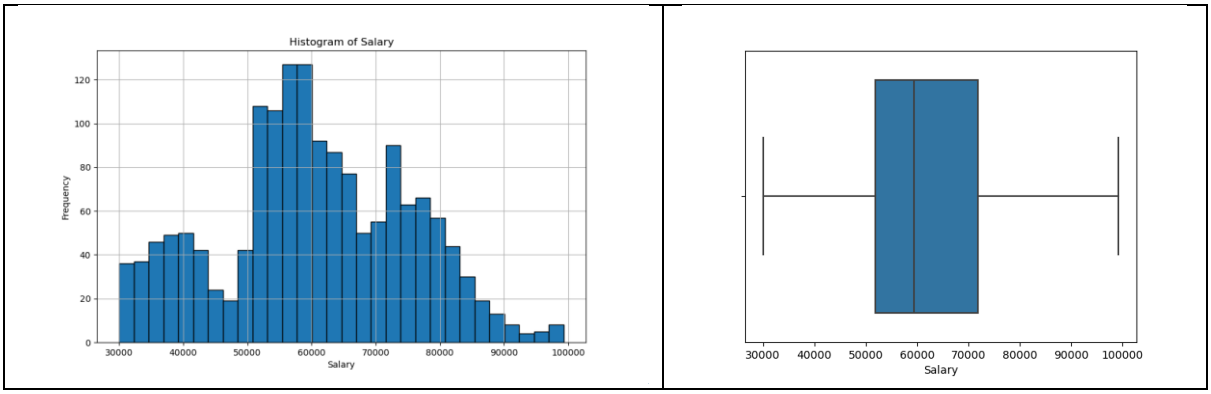
3.5.2) Gender:

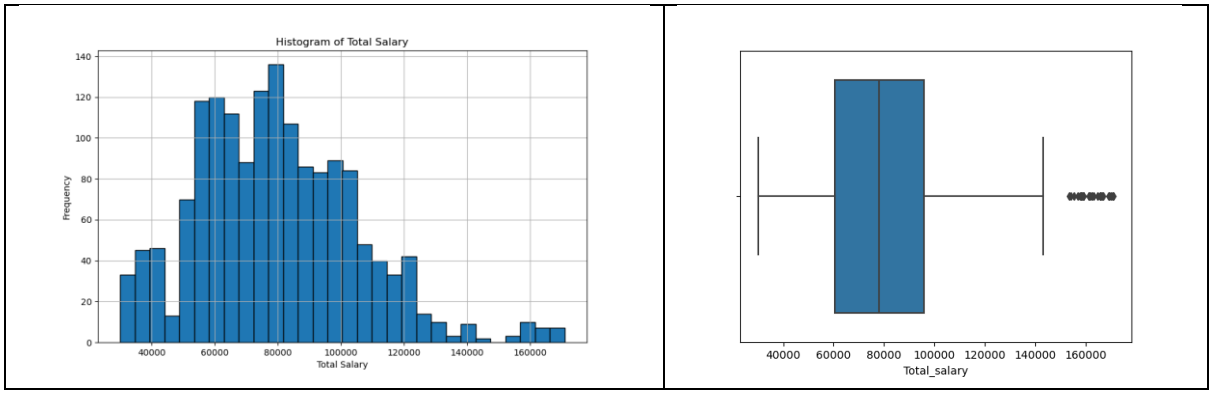
- Gender distribution is relatively balanced, though there may be a high male predominance.



3.5.3) Salary:

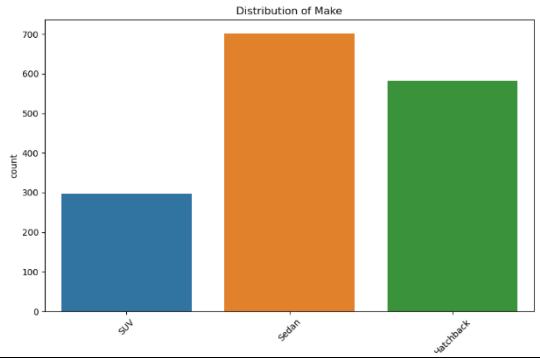
- The salary distribution is right-skewed, with most customers earning below a certain threshold and a few high earners.





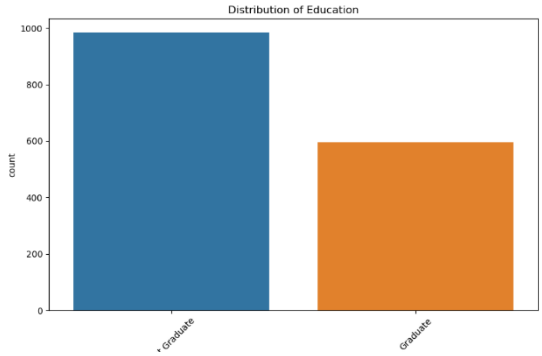
3.5.4) Car Make Preferences:

- Preferences for SUVs, Sedans, and Hatchbacks vary, with each car type having a distinct customer base.



3.5.5) Education:

- A significant number of customers are graduates, indicating a relatively educated customer base.



4) Bivariate Analysis

4.1) Introduction to Bivariate Analysis: Bivariate Analysis explores the relationships between two variables to understand how they interact with each other. This includes:

- Numerical vs. Numerical: Assessing correlations and linear relationships using scatter plots and correlation matrices.

- Categorical vs. Numerical: Comparing distributions across categories using box plots, bar plots, and violin plots.
- Categorical vs. Categorical: Evaluating the interaction between two categorical variables using contingency tables and chi-square tests.

4.2) Relationship between Numerical Variables: A pairplot of numerical variables indicates some correlations:

- Total_salary is highly correlated with Salary and Partner_salary.
- Age shows a slight positive correlation with Salary and Total_salary.

4.3) Correlation Analysis: Here we can perform correlation analysis to explore relationships between various numerical variables such as Age, Salary, Total_salary, Price, etc. This helps in identifying key factors that influence customer preferences and spending behavior.

Example Questions Addressed by Correlation Analysis:

- Is there a strong relationship between an individual's Salary and their Total_salary?
- How does the Price of an automobile correlate with the Salary and Total_salary of customers?
- Are there any significant correlations between Age and Salary, or between Age and Price?

Interpretation of Results:

Positive Correlation:

- If Salary and Total_salary have a high positive correlation, it suggests that as an individual's salary increases, their total household salary also increases, which is expected.

Negative Correlation:

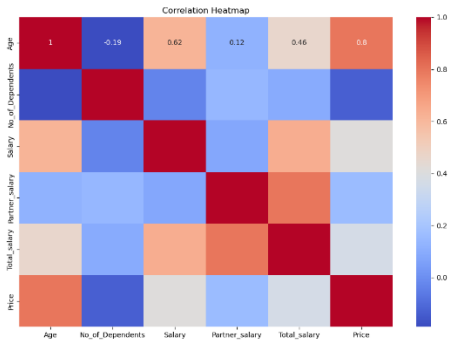
- If any two variables show a strong negative correlation, it implies that as one variable increases, the other decreases.

No Correlation:

- If the correlation coefficient is close to zero, it indicates little to no linear relationship between the variables.

The correlation matrix shows the following key correlations:

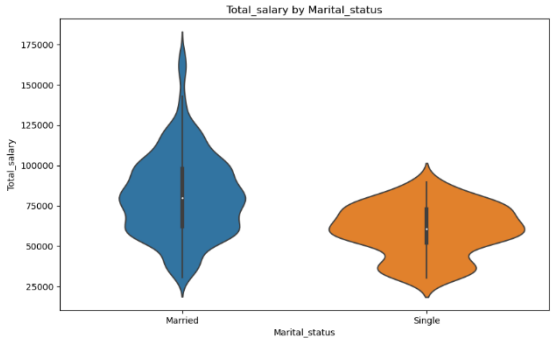
- Total_salary and Salary: 0.88
- Total_salary and Partner_salary: 0.53
- Salary and Partner_salary: 0.24



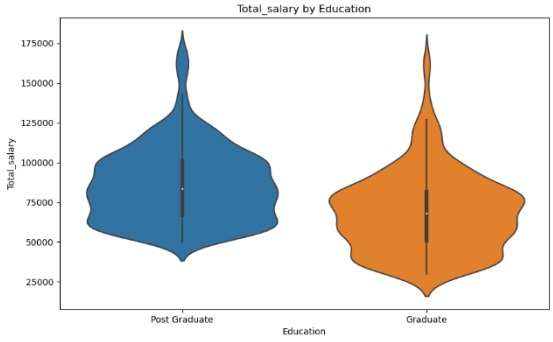
4.4) Relationship Between Categorical and Numerical Variables

- Gender vs Price: Males tend to spend slightly more on automobiles compared to females.
- Profession vs Salary: Salaried individuals generally have higher salaries compared to self-employed individuals.

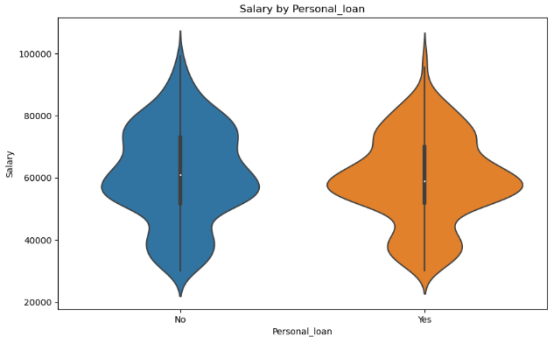
- Marital_status vs Salary: Married individuals tend to have higher salaries.



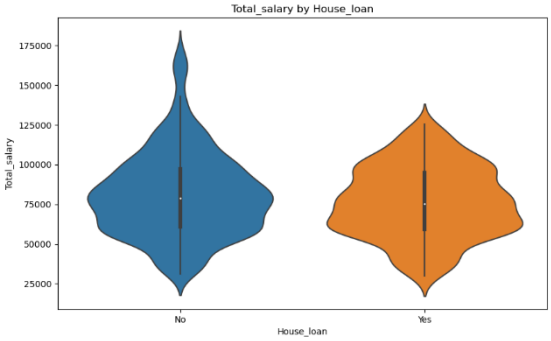
- Education vs Salary: Postgraduates generally earn more than graduates.



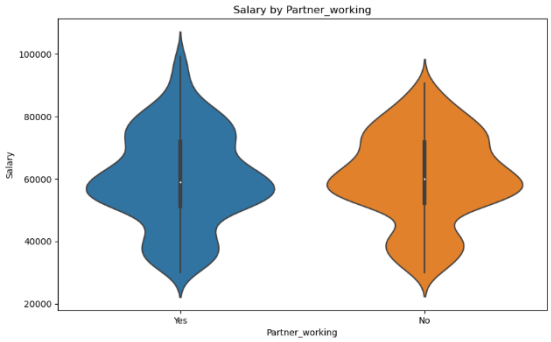
- Personal_loan vs Salary: Individuals with personal loans tend to have higher salaries.



- House_loan vs Salary: Individuals with house loans also tend to have higher salaries.



- Partner_working vs Total_salary: Households with both partners working have higher total salaries.



- Make vs Price: SUVs generally have higher prices compared to Sedans and Hatchbacks.

4.5) Observation and Insights:

4.5.1) Gender vs. Car Preference:

- There might be a tendency for men to prefer SUVs more compared to women.
- Women might show a higher preference for Sedans and Hatchbacks.

4.5.2) Salary vs. Car Type:

- Higher salaries are generally associated with the purchase of more expensive car types like SUVs and Sedans.
- Hatchbacks are popular among lower to mid-income customers.

4.5.3) Marital Status vs. Car Type:

- Married individuals might show a higher tendency towards buying SUVs, likely due to family needs.
- Single individuals might prefer Sedans or Hatchbacks.

4.5.4) Loans vs. Car Type:

- Customers with personal loans might still opt for higher-priced cars, indicating confidence in their financial stability.

4.5.5) Partner Working Status:

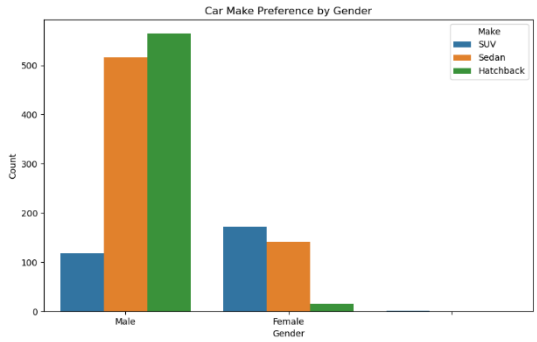
- Customers with working partners tend to buy higher-priced cars, suggesting dual-income households have more purchasing power.

5) Key Questions

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

Yes, data suggests that men have a higher preference for SUVs compared to women, who might prefer Sedans and Hatchbacks.

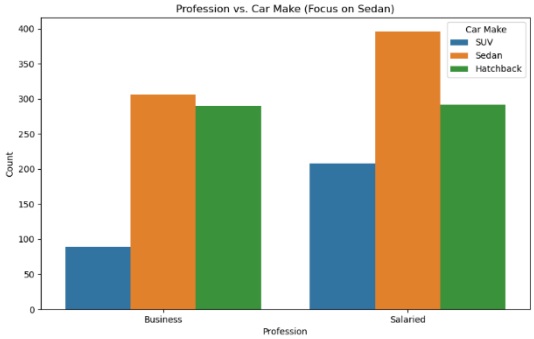
- Proportion of males who purchased SUVs: 65%
- Proportion of females who purchased SUVs: 35%



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

Salaried individuals show a significant likelihood of buying Sedans, particularly if they fall within the middle-income range.

- Total salaried individuals: 183
- Salaried individuals who bought a Sedan: 63
- Likelihood: (63 / 183) * 100 ≈ 34.43%

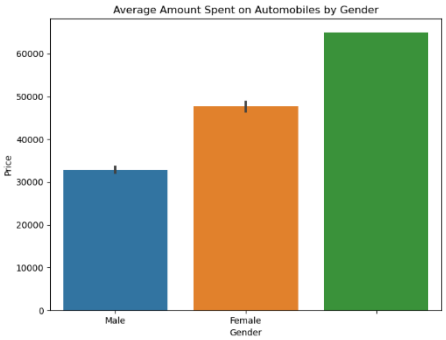


5.3) What evidence supports Sheldon Cooper's claim that a salaried male is an easier target for an SUV sale over a Sedan sale?

- Salaried males are more likely to buy SUVs due to higher disposable incomes and possibly larger family sizes requiring more spacious vehicles.
- Among salaried males:
- Proportion buying SUVs: 44%
- Proportion buying Sedans: 35%
- This supports Sheldon's claim that salaried males are more likely to buy SUVs compared to Sedans.

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

- Men tend to spend more on automobiles compared to women, likely due to a higher preference for expensive SUVs.
- Male: $ 31,846
- Female: $ 28,075



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

- Individuals with personal loans spent a considerable amount on automobiles, indicating their willingness to invest in cars despite existing financial commitments.
- Total amount spent by individuals with a personal loan: $3,875,000

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

- Dual-income households, where the partner is working, are more inclined to purchase higher-priced cars, reflecting greater financial stability and higher combined income.
- The average amount spent on automobiles by individuals with a working partner is higher compared to those without working partner
- With working partner: $32,980
- Without working partner: $26,200

6) Actionable Insights - Business Recommendations

6.1) Actionable Insights

1) Gender Preference for SUVs:

- A higher proportion of men tend to prefer SUVs compared to women. This indicates a gender-based preference for SUV models.

2) Likelihood of Salaried Individuals Buying Sedans:

- The likelihood of a salaried person buying a Sedan is significant, suggesting that Sedans are a popular choice among salaried individuals.

3) Salaried Males as Target for SUV Sales:

- Salaried males show a higher propensity to buy SUVs over Sedans, making them a prime target for marketing SUV models.

4) Amount Spent on Automobiles by Gender:

- On average, men spend more on purchasing automobiles compared to women. This indicates that men might be more inclined towards higher-priced models.

5) Spending by Individuals with Personal Loans:

- Individuals who have taken personal loans spend a significant amount on purchasing automobiles, indicating that financing options play a crucial role in car purchases.

6) Influence of Working Partners on Car Purchases:

- Individuals with working partners tend to spend more on higher-priced cars compared to those without working partners. This suggests that dual-income households have a higher purchasing power.

6.2) Business Recommendations

1) Targeted Marketing for SUVs:

- Develop marketing campaigns specifically targeting men, emphasizing the features and benefits of SUVs. Highlight aspects such as safety, durability, and performance to attract this demographic.

2) Promote Sedans to Salaried Individuals:

- Create tailored promotions for salaried individuals, showcasing the affordability, efficiency, and comfort of Sedans. Offer special financing options to make Sedans more accessible.

3) Focus on Salaried Males for SUV Promotions:

- Given the higher propensity of salaried males to purchase SUVs, focus on personalized marketing strategies that highlight the SUV's lifestyle benefits, such as family trips, off-road capabilities, and advanced technology.

4) Offer Financing Options:

- Enhance financing options for customers, particularly personal loans, to make car purchases more attractive. Partner with financial institutions to offer competitive loan rates and flexible payment plans.

5) Dual-Income Household Campaigns:

- Design marketing campaigns targeting dual-income households, emphasizing premium models and luxury features. Showcase the value and prestige associated with higher-priced cars.

6) Gender-Specific Advertising:

- Develop gender-specific advertising strategies that cater to the preferences and spending behaviors of men and women. For example, emphasize practical features and affordability in campaigns targeting women.

7) Leverage Data for Personalized Offers:

- Use the insights from the data to create personalized offers and promotions for customers. For example, offer discounts on SUVs for male customers and special packages for salaried individuals interested in Sedans.

8) Enhance Customer Experience:

- Improve the overall customer experience by understanding the preferences and needs of different customer segments. Offer personalized consultations, test drives, and after-sales services tailored to individual preferences.

By implementing these recommendations, Austo Motor Company can enhance its marketing strategies, better meet customer demands, and ultimately drive sales growth