Statistical Methods for Business Decision Making Project Report

Problem 1

Context

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 analytics professional to improve the existing campaign.

Objective

They want to analyse the data to get a fair idea about the demand of customers which will help them in enhancing their customer experience, they have some questions to be answered.

Data Description

We have been provided with about 15000 observations or records of Customer Purchase data, some of the attribute's a in the data are from the Austo Motor company's sales records, and other attributes are to do with the demographic characteristics of the said customers.

Key Questions

Do men tend to prefer SUVs more compared to women?

No: Only 10% of Males purchased SUVs and 53% of Female purchased SUVs, The Data Suggests Females prefer SUVs compared to Men. (Refer to Figure 1)

Figure 1: Car Make Purchased by Gender
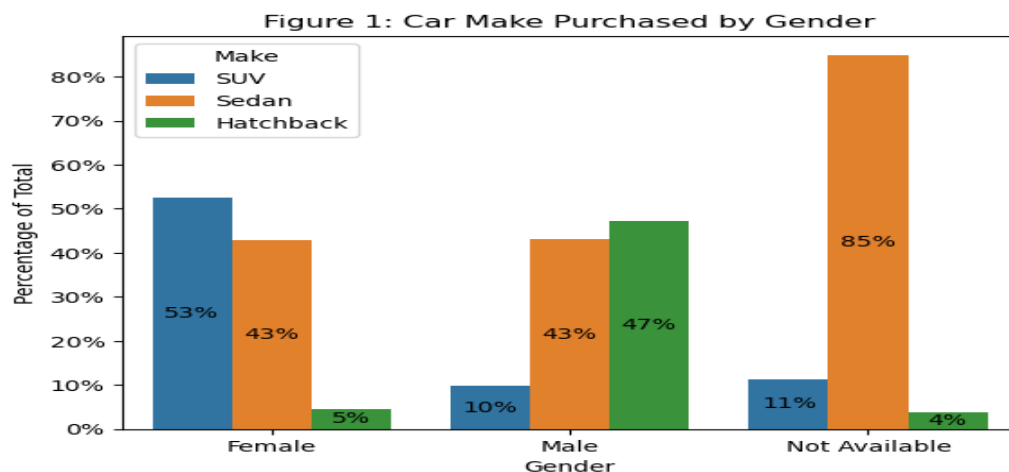
### What is the likelihood of a salaried person buying a Sedan?

44%: Out of 100 salaried individuals 44 are likely to purchase a Sedan, 33 are likely purchase a Hatchback and the rest 22 a SUV. (Refer to Figure 2)

Figure 2: Car Make Purchased by Salaried Professional
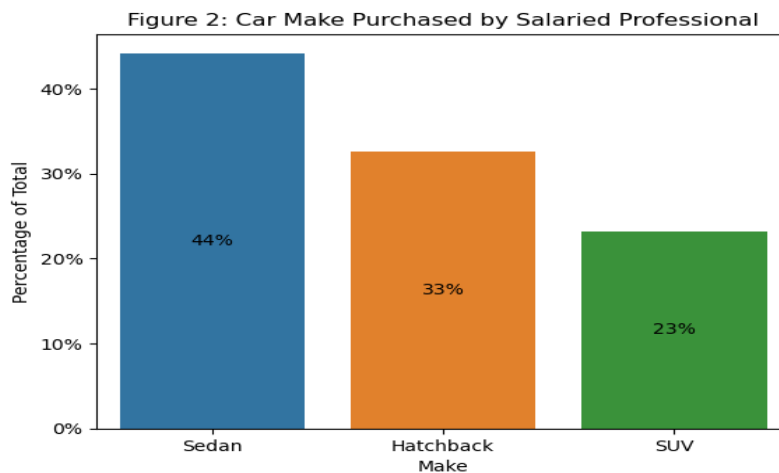
What evidence or data supports Sheldon Cooper's claim that a salaried male is an easier target for a SUV sale over a Sedan sale?

There is no evidence to support the claim***, we do have evidence to support that salaried male is an easier target for a Sedan sale over SUV sale.

1. Males are more likely to purchase Sendan's than SUV's. (Refer Figure 1, above)
2. Salaried Individual are more likely to purchase Sendan's than SUV's. (Refer Figure 2, above)
3. Salaried Male Individuals are more likely to purchase Sendan's than SUV's (Refer Figure 3, below)

Figure 3: Car Make Purchased by Salaried Professionals by Gender
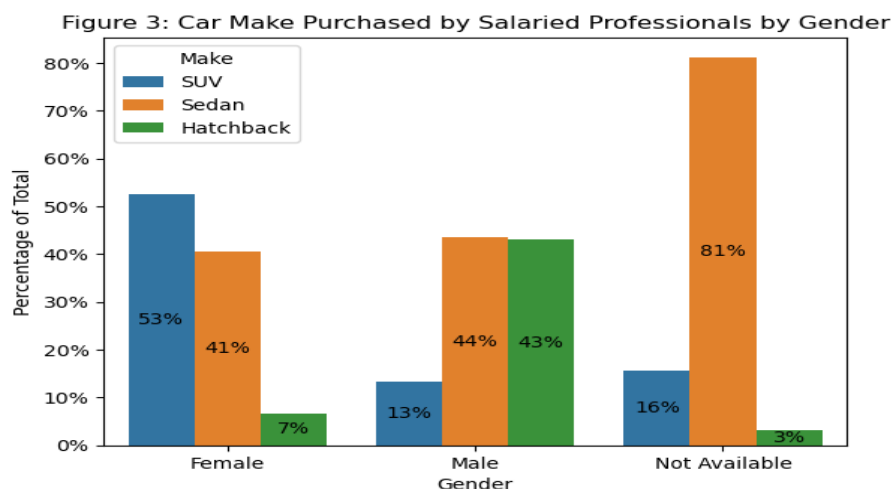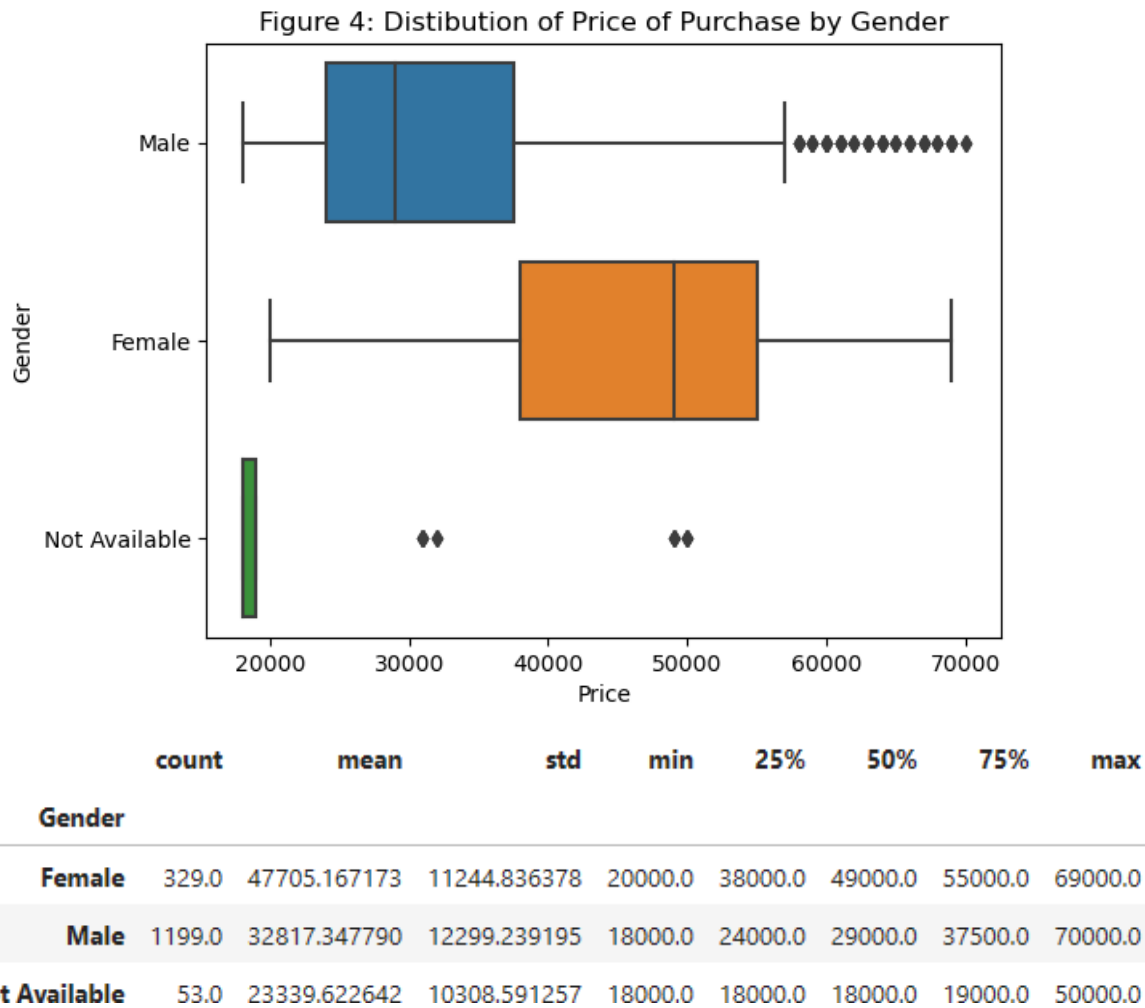
How does the amount spent on purchasing automobiles vary by gender?

The amount spent on purchasing automobiles varies Greatly between the genders. For the most part Females spend higher amounts on purchasing automobiles than Males. (Refer to Figure 4 and supporting table.)

1. The median value spent by Females is very much higher than Males: 49K compared to 29K respectively.
2. 50% Females purchase automobiles in the price range of 38K to 55K, In comparison 50% Males purchase automobiles in the price range of 24K to 37.5K.
3. 75% of the Males purchase automobiles at a lower price than what 75% of the Females spend on purchasing automobiles.

4. On an average Females spend around 15K more on purchasing automobiles than Males

Figure 4: Distibution of Price of Purchase by Gender

	count	mean	std	min	25%	50%	75%	max
Gender								
Female	329.0	47705.167173	11244.836378	20000.0	38000.0	49000.0	55000.0	69000.0
Male	1199.0	32817.347790	12299.239195	18000.0	24000.0	29000.0	37500.0	70000.0
Not Available	53.0	23339.622642	10308.591257	18000.0	18000.0	18000.0	19000.0	50000.0

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

1. There is not much difference in the spend for purchasing automobiles by individuals who took a personal loan vs who did not.
1. Individuals who took a personal loan spent 28.99 mill vs who did not take a personal loan spent 27.29 mill, 52% vs 48% of the total spend respectively. (Refer to Figure 5)
2. The distribution of price/spend between these two groups are very close and similar. (Refer to Figure 6 and supporting table)
3. 50% Individuals who took a personal loan spent around 42% to 72% of the salary on purchasing automobiles, 50% individuals who did not take a personal loan spent around 44% to 74% of the salary on purchasing automobiles, this tells us that immaterial of whether an individual takes a personal loan or not they spend close to same % of the salary to purchase automobiles. (Refer to

Figure 7 and supporting table)

Figure 5: Total amount in million spent on purchasing automobiles by individuals who have personal loan vs who don't.
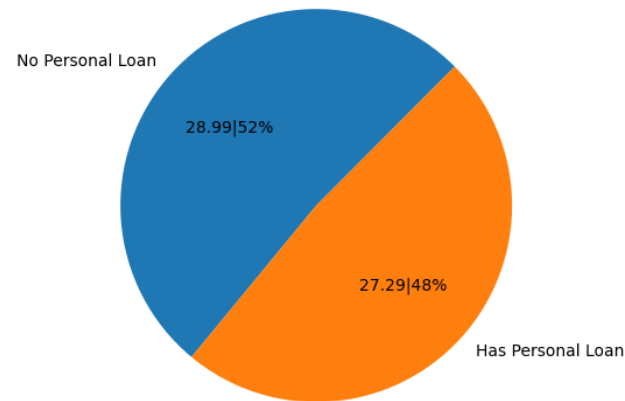

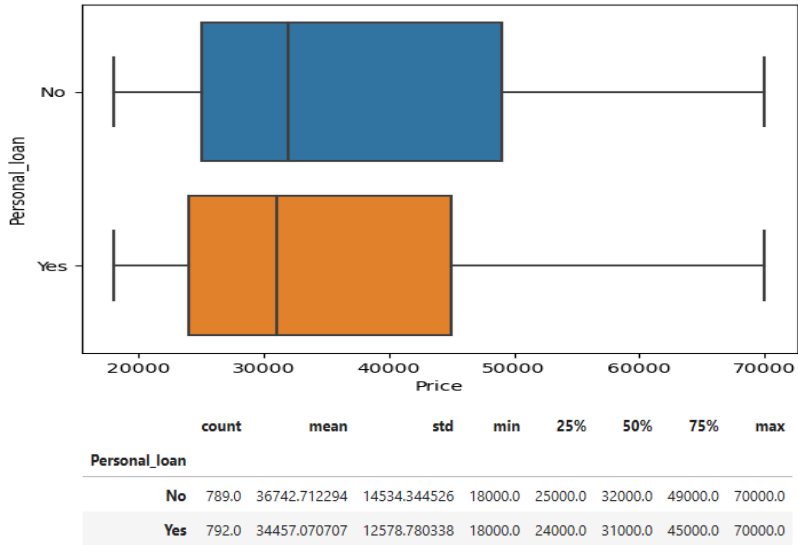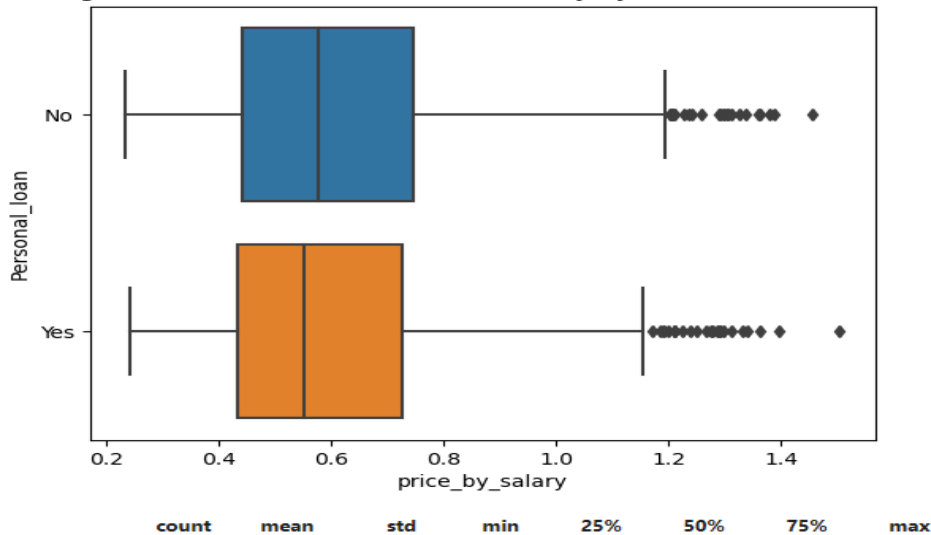Figure 6: Distibution of Price of Purchase by Personal loan status

	count	mean	std	min	25%	50%	75%	max
Personal_loan								
No	789.0	36742.712294	14534.344526	18000.0	25000.0	32000.0	49000.0	70000.0
Yes	792.0	34457.070707	12578.780338	18000.0	24000.0	31000.0	45000.0	70000.0

Figure 7: Distibution of Price % of Salary by Personal loan status

	count	mean	std	min	25%	50%	75%	max

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

1. 14% of the individuals who have a working partner buy higher-priced cars vs 16% of individuals who do not have a working partner. (Refer to Figure 8)
2. This would seem to suggest that having a working partner may lead to a drop in the preference of higher-priced cars slightly.

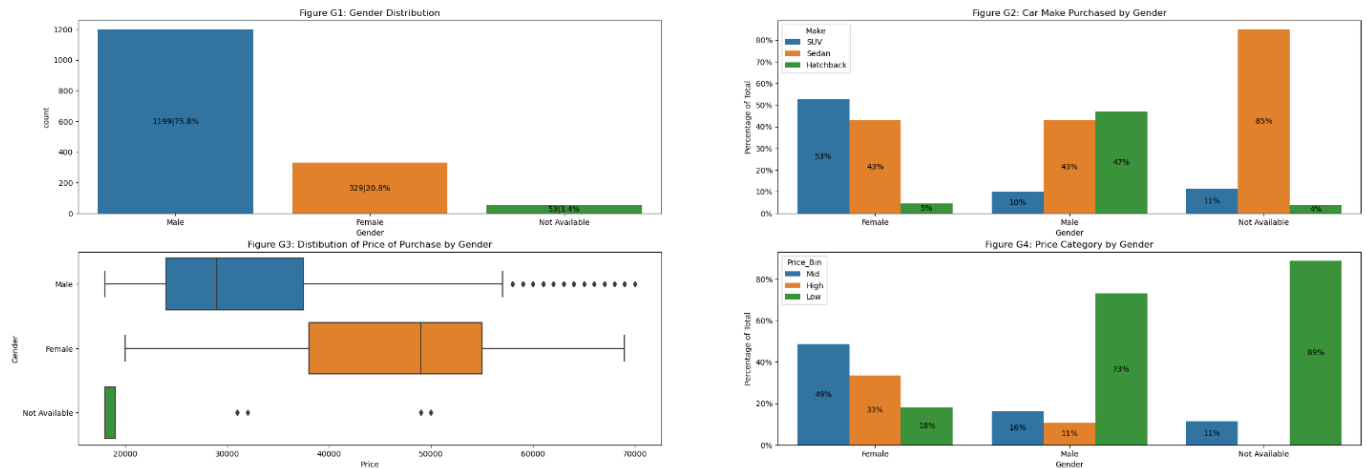Figure 8: Price Category by individuals with or with out partners
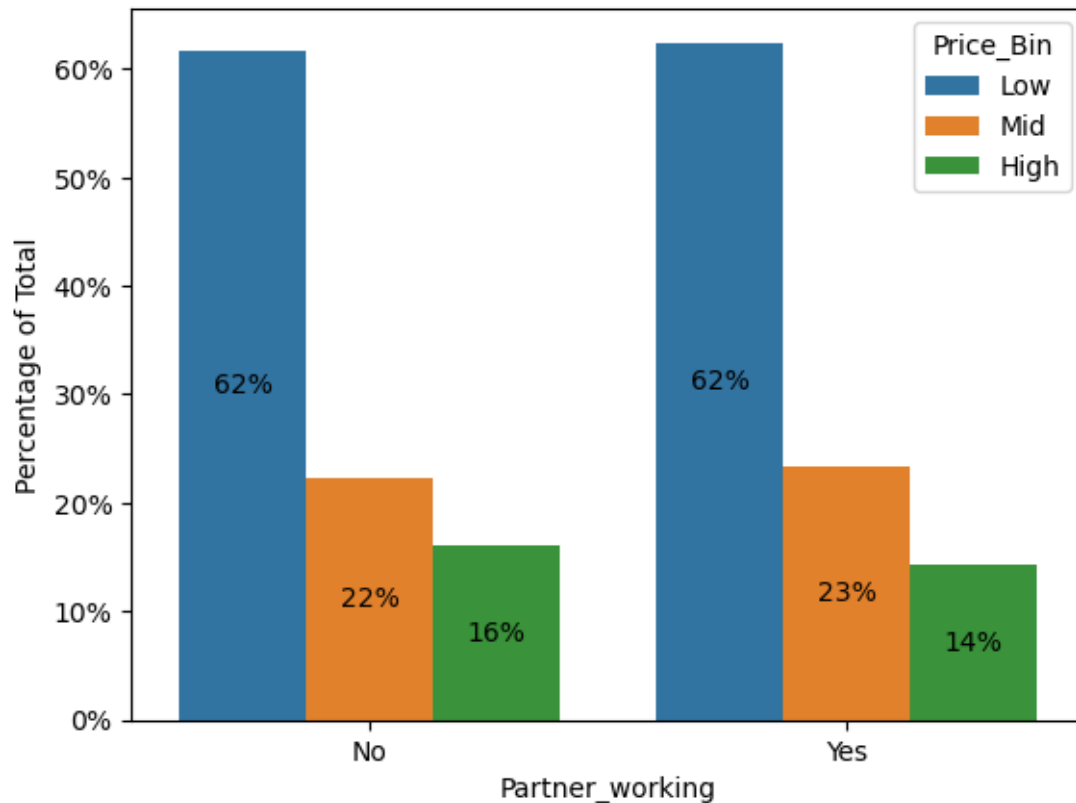
Actionable Insights & Recommendations

1. The Purchase behaviour between **Females** and **Males** certainly stands out. (Refer to Figure G1, G2, G3 and G4)
   a. **Females** seem to prefer SUV followed by Sedan and Hatchback, **Males** prefer Hatchback followed by Sedan and SUV.
   b. 50% **Females** purchase automobiles in the price range of 38K to 55K, In comparison 50% **Males** purchase automobiles in the price range of 24K to 37.5K.
   c. 75% of the *Males** purchase automobiles at a lower price than what 75% of the *Females** spend on purchasing automobiles.
   d. On an average **Females** spend around 15K more on on purchasing automobiles than **Males**
      i. 82% **Females** preferred price points in the Mid-low range, where as 73% **Males** preferred price points in the low range.
   e. **Recommendations**:
      i. Marketing campaign targeted to *Gender**, should be looked into.
      ii. 76% of the customer base are **Men**,**73%** preferred low range price points, **43% of Men** prefer Sedans, marketing campaigns for **Men targeting low-range price point Sedans** should be looked into.

iii. 82% **Females** preferred price points in the Mid-High range, **96% of Women** prefer SUV's and Sedans with over 50% preferring SUV's, marketing campaigns for **Women targeting Mid-High range price point SUVs and Sedans** should be looked into.

iv. Look into features of cars by **Make** and **Model**, purchased by different genders to check if some features are preferred by one over the other.

Analysis by Gender.

Figure G1: Gender Distribution

Figure G2: Car Make Purchased by Gender

Figure G3: Distibution of Price of Purchase by Gender

Figure G4: Price Category by Gender

2. The Purchase behaviour between **Different Age groups** also stands out.
   a. Since we have established that **Gender** is an important factor, we look at Age groups behaviour which are gender specific.
      i. The Analysis of **Age Groups** for **Women** give us the following insights. (Refer to Figure W1, W2, W3 and W4)
         1. Nearly 70% of the Female Population is in the Age group of **30-45 years**, Majority of these prefer Sedans with the **42-45 years** group preferring SUVs.
         2. The age groups above **45 years**, all prefer SUV's. The youngest age group of **22-25 years** all prefer Hatchbacks. **26-37 years** lean heavily towards Sedans and **26-37 years** show an almost equal preference to Sedans and SUVs with no preference for Hatchbacks.
         3. From a price point stand point, The youngest age group of **22-25 years** all prefer price points on the lower range, age group of **30-45 years** prefer Mid-High range prices 70-80% of the time, age group of **46 and above** prefer High range prices about most of the time 77% or greater to be specific, and in the case of age groups above **54 years** High range prices are preferred 100% of the time.
      ii. **Recommendations**:
         1. Marketing campaign targeted to **Women of different age groups**, should be looked into.
         2. Campaigns for **Hatchbacks with Lower Price range** should be targeted towards **Women between the age groups of 22-25 years**.

3. Campaigns for **Sedans and SUVs with Mid-High Price range** should be targeted towards **Women between the age groups of 26-45 years**
4. Campaigns for **SUVs with High Price range** should be targeted towards **Women of age groups of 46 or older years**

Age Group Analysis for Women.
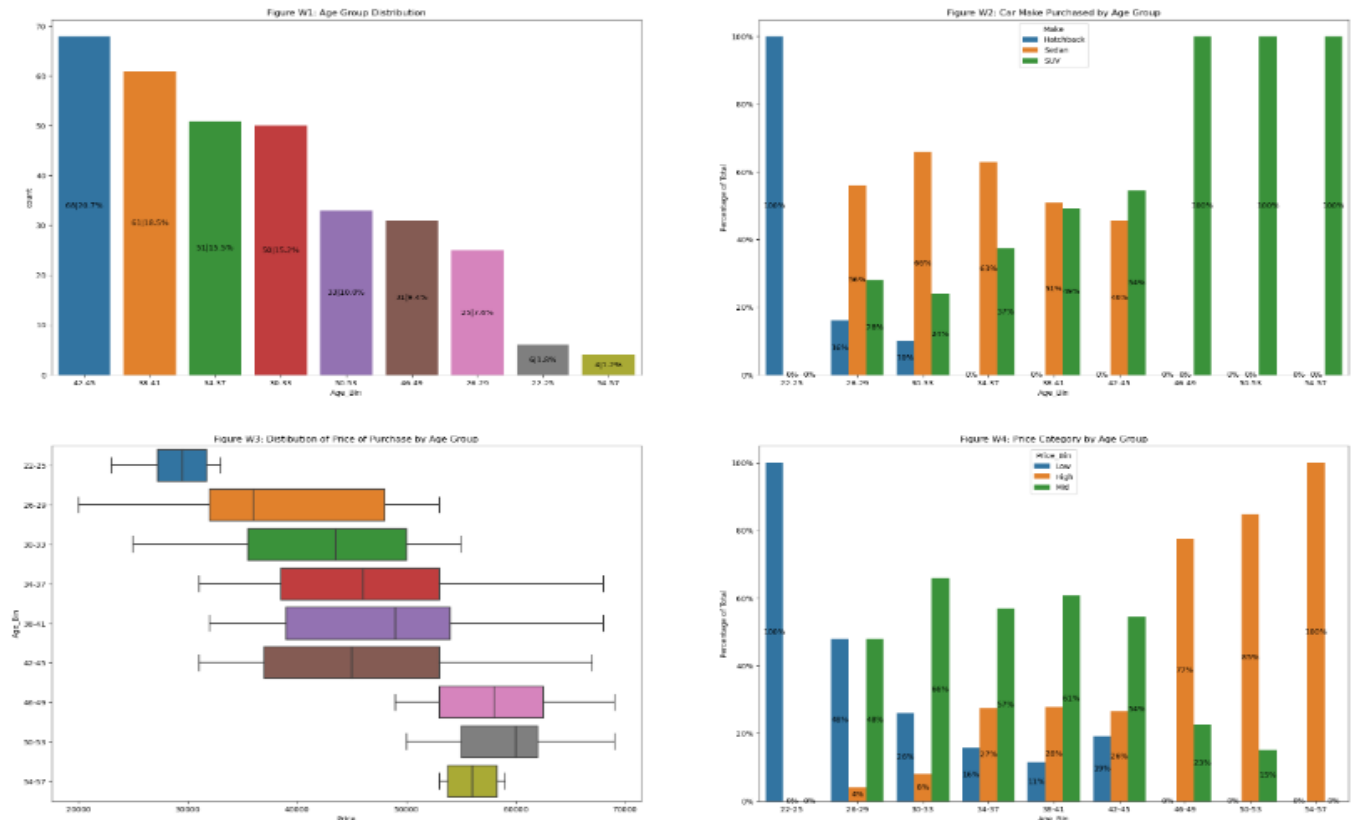
iii. The Analysis of **Age Groups** for **Men** give us the following insights. (Refer to Figure M1, M2, M3 and M4)
   1. Nearly 70% of the Male Population is in the Age group of **30-45 years**, Majority of these prefer Sedans with the **34-45 years** group preferring SUVs about 25% of the time and the younger age group **30-33 years** preferring Hatchbacks about 33% of the time.
   2. The age groups above **45 years**, all prefer SUV's. The younger age group of **22-29 years** prefer Hatchbacks most of the time and Sedans the rest of the time, They do not show interest in SUVs for the most part.

   1. From a price point stand point, The younger age group of **22-29 years** all prefer price points on the lower range for the most part, age group of **34-45 years** prefer Mid-High range prices 80% of the time, age group of ***46 and above** prefer High range prices.
b. **Recommendations**:

1. Marketing campaign targeted to **Men of different age groups**, should be looked into.
2. Campaigns for **Hatchbacks with Lower Price range** should be targeted towards **Men between the age groups of 22-33 years**.
3. Campaigns for **Sedans and SUVs with Mid-High Price range** should be targeted towards **Men between the age groups of 34-45 years**
4. Campaigns for **SUVs with High Price range** should be targeted towards **Men of age groups of 46 or older years**

Age Group Analysis for Men.
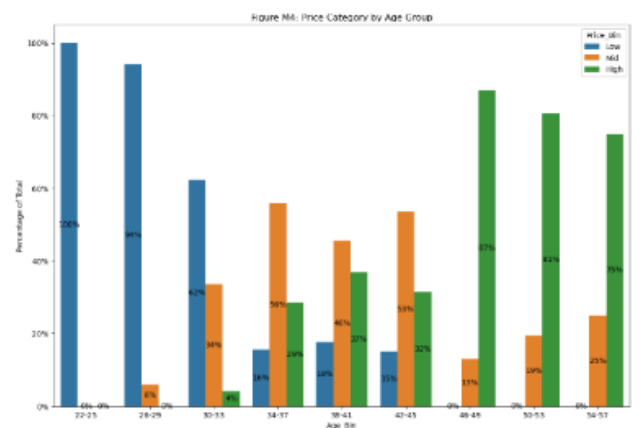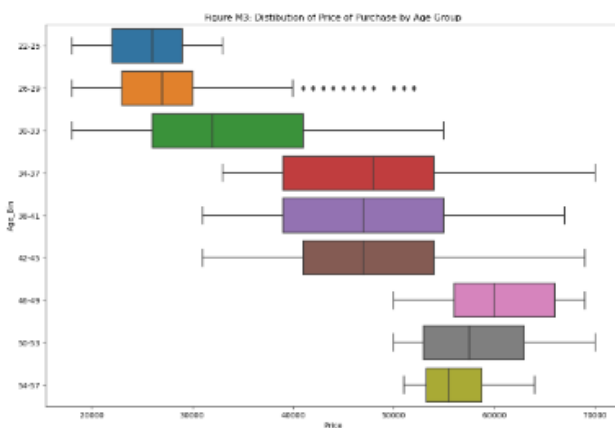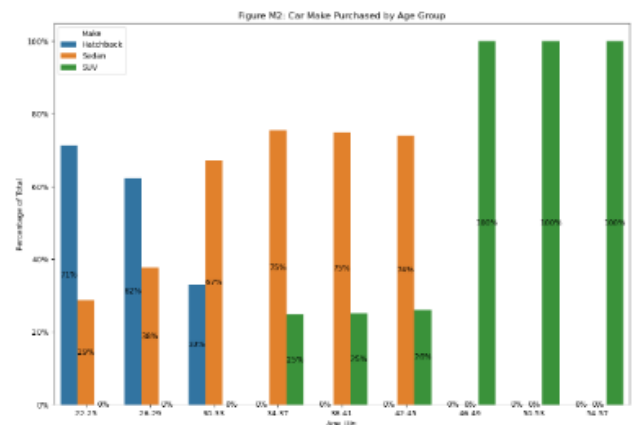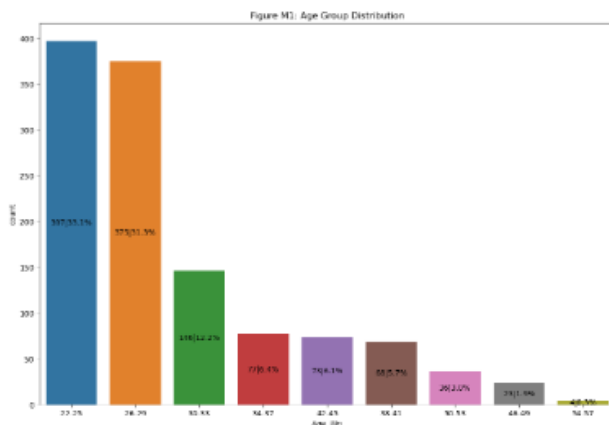
Data Overview

Attribute	Description
Age	The Age of the customer
Gender	The Gender of the customer
Profession	Is the customer being salaried or a business person
Marital Status	Is the customer is married
Education Level	Is the customer a Graduate or Postgraduate
Number of Dependants	How many dependants does the customer have
Has Personal Loan	Does the customer have a personal loan
Has House Loan	Does the customer have a home loan
Has a Working Partner	Is the customers partner working
Salary	What is the salary of the customer

Partner Salary	What is the salary of the customers partner
Total Salary	What is the total salary (Salary + Partner Salary)
Price	What is the price at which the automobile was purchased
Make	What is the make of the automobile

1. The data provided has an overall of **1581 observations**.
2. The data has **no duplicate values**.
3. The values in **Gender** variable had few **data entry errors**, which were corrected.
4. Not all the observations in the **Gender** field had values, these were updated to "**Not Available**".
5. Missing values were seen in the "**Partner Salary**" variable, these were updated to "**0.0**"
6. There was a calculation error seen where "**Total salary**" was greater than the sum of Salary and partner salary, for 16 records.
7. The records in point "6" were treated to populate the **"Partner Salary"** as a difference of **"Total Salary"** and **"Salary".**
8. The variable **"Number of Dependants** "has numerical values in it, but for this data set it makes sense to treat this as a Category.
9. The variable **"Age"** is numerical, but we can also derive an **"Age_Bin"** as a Category column.

Using methods of descriptive statistics to analyse the data:

Checking each of the Attributes (Univariate Analysis):

Attributes that are categories

Below is listed the proportion of each value in the attributes that make up the **1581** records**.**

1. Gender: Male (76%), Female (21%) and Not Available (3%)
2. Profession: Salaried (57%), Business (43%)
3. Marital Status: Married (91%), Single (9%)
4. Education: Post Grad (62%), Grad (38%)
5. Personal loan: Yes (50.1%), No (49.9%)
6. Home loan: No (67%), Yes (33%)
7. Partner working: Yes (55%), No (45%)
8. Make: Sedan (44%), Hatchback (36%), SUV (18%)
9. Age_Bin: 26-29(27%), 22-25(26%), 30-33(13%), 42-45(9%), 34-37(8%), 38-41(8%), 50-53(4%), 46-49(4%), 54-57(1%)
10. No_of_Dependants: 3(35%), 2(35%), 1(15%), 4(14%), 0(1%)

Attributes that are numeric that can be looked at as a continuous series:
1. Age: Minimum = 22   Maximum = 54   Mean = 31.9   Median = 29   std = 8.4
2. Salary: Minimum = 30000 Maximum = 993000 Mean = 60392.2 Median = 59500 std = 14674.8
3. Partner salary: Minimum = 0.0   Maximum = 80500 Mean = 18869.5 Median = 24900 std = 19570.6
4. Total salary: Minimum = 30000 Maximum = 171000 Mean = 79625.9 Median = 78000 std = 25545.8
5. Price: Minimum = 18000 Maximum = 70000 Mean = 35597.7 Median = 31000 std = 13633.6

Explore the relationship between the categories and the numerical attributes
1. Distribution of Age, Salary and Price by Genders is very different and we may want to Analyse the data by "Gender" separately go forward.

2. Salaried Individuals have a slightly bigger appetite for more expensive vehicles and purchase vehicles 1 to 3 years later as compared to Business Individuals
3. Married Individual may purchase slightly costlier vehicles than Single Individuals.
4. Post Graduate Individuals have a much higher income and spend capability than Graduate Individuals, yet Graduate Individuals have a slightly bigger appetite than Post Graduate Individuals to purchase more expensive vehicles.
5. Individuals who do not have a Personal loan have a slightly higher appetite to purchase costlier vehicles but not by much.
6. income and spend capability of Individuals without a House loan is slightly higher than Individuals with a House loan, Individuals without a House loan higher appetite to purchase costlier vehicles than Individuals with a House loan.
7. Younger individuals prefer Hatchback, middle aged individuals prefer Sedan's and older individuals prefer SUV's, the preference of Make seems to be tied closely to the income and spend capability of an individual.
8. We see Sedan's in the Hatchback's price range and vice versa, Also SUVs in Sedan's price range and vice versa, This would indicate Individuals have a preference of Make, we may need to Analyse the data by "Make" separately go forward.
9. Individuals with bigger families or no family purchase vehicles earlier in life, higher % of Individuals with 1 dependant have a higher income and spend capabilities compared to the others with no or more than 1 dependant, Individuals with 1 or 2 dependants may have an appetite of more expensive vehicles, Than Individuals with 0,3,4 dependants.
10. Income and spend capability increase with age, appetite of more expensive vehicles increases with age.

Problem 2

Context

A bank generates revenue through interest, transaction fees, and financial advice, with interest charged on customer loans being a significant source of profits.

GODIGT Bank, a mid-sized private bank, offers various banking products and cross-sells asset products to existing customers through different communication methods.

However, the bank is facing high credit card attrition, leading them to reevaluate their credit card policy to ensure customers receive the right card for higher spending and intent, resulting in profitable relationships.

Data Description

We have been provided with about 8400 observations or records of Customer Credit Card and Bank Account usage data, all of the attribute's a in the data are from the GODIGT company's banking records.

Actionable Insights & Recommendations

1. Top 5 important variables are high_networth, avg_spends_l3m, cc_limit,card_type and cc_active fields.

2. *Business*
   Justifications high_networth,avg_spends_l3m, cc_limit , card_type and overall_future_month_activity are the top 5 important variables as the distributions of avg_spends_l3m and *cc_limit* vary in most cases for

*high_networth,card_type and overall_future_month_activity.(Refer to Figure P2.1 through to P2.6 below).