

|  |  |  |  |
| --- | --- | --- | --- |
| Student Number: | Yuvraj Chauhan | | |
| Programme:(e.g. Business Management) | MBA Business Analytics | | |
| Module Title: (e.g. Studying for Business) | Strategic Business Analytics |  |  |
| Module Code: | SG7012 | Word Count | 4232 |

**In submitting this assessment, I confirm that no part of this assignment. except where clearly quoted and referenced. has been copied from material belonging to any other person e.g. from a book. handout, another student. I am aware that it is a breach of UEL regulations to copy the work of another without clear acknowledgement and that attempting to do so renders me liable to disciplinary proceedings.**

**Executive Summary**

Task A of this report involves an analysis of the customers' shopping experience and their preferences and Task B involves an analysis of the factors that influence vehicle ownership in Asian countries. In Task A, the following statistical methods were used for the analysis of customers; descriptive analysis, frequency distribution, and correlation analysis. Using analysis of variables like age, purchase history, and review ratings of the products people buy to market the products people prefer, we find that the results indicated a strong relationship between product features and consumer choices.

In Task B, regression analysis is conducted to find out the vehicle ownership behavior. Two regression models were analyzed defining one rate as vehicle density per thousand per capita and another one as an absolute number of vehicles. The total vehicle ownership model with a relatively higher value of 0.975 was more credible and useful for AutoMobile Inc. because it assisted in understanding the market trends for the company including production strategy.

**Table of Contents**

[1. Introduction 5](#_Toc175772279)

[1.1 Purpose of the Report 5](#_Toc175772280)

[1.2 Audience and Stakeholders 5](#_Toc175772281)

[1.3 Hypothesis or Research Question 5](#_Toc175772282)

[1.4 Definition of Key Terms 6](#_Toc175772283)

[2. Method 6](#_Toc175772284)

[2.1 Description of Data Presentation Methods 6](#_Toc175772285)

[Task A: Customer Shopping Trends 6](#_Toc175772286)

[Task B: AutoMobile Inc. 8](#_Toc175772287)

[2.2 Explanation of Tools and Techniques Used 11](#_Toc175772288)

[Task A: Customer Shopping Trends 11](#_Toc175772289)

[Task B: AutoMobile Inc. 12](#_Toc175772290)

[3. Findings and Results 13](#_Toc175772291)

[3.1 Presentation of Data Analysis Results 13](#_Toc175772292)

[Task A: Customer Shopping Trends 13](#_Toc175772293)

[Task B: AutoMobile Inc. 15](#_Toc175772294)

[3.2 Interpretation of Data 17](#_Toc175772295)

[4. Discussion 18](#_Toc175772296)

[4.1 Interpretation of Findings 18](#_Toc175772297)

[4.2 Exploration of Correlations or Causations 19](#_Toc175772298)

[4.3 Comparison with Expectations or Hypotheses 20](#_Toc175772299)

[5. Conclusion and Recommendations 20](#_Toc175772300)

[5.1 Summary of Key Findings 20](#_Toc175772301)

[5.2 Assessment of Hypothesis 21](#_Toc175772302)

[5.3 Recommendations 21](#_Toc175772303)

[References 22](#_Toc175772304)

[Appendix 25](#_Toc175772305)

**List of Figures**

[Figure 1: Descriptive Statistics 7](#_Toc175772306)

[Figure 2: Frequency Distribution with Histogram 7](#_Toc175772307)

[Figure 3: Correlation Analysis with Bar chart 8](#_Toc175772308)

[Figure 4: Pivot Table of Subscription Status with Donut Chart 8](#_Toc175772309)

[Figure 5: Two different Scatter Plots show the Vehicles for 1000 population with per capita income and population density 9](#_Toc175772310)

[Figure 6: Correlation Analysis with Vehicles per 1000 population 9](#_Toc175772311)

[Figure 7: Regression Analysis 10](#_Toc175772312)

[Figure 8: Cyprus Analysis Comparison 10](#_Toc175772313)

[Figure 9: Pivot table of Payment Method with Bar chart 11](#_Toc175772314)

[Figure 10: Pivot Table of Frequency of purchases with a bar chart 12](#_Toc175772315)

[Figure 11: Correlation with Total vehicles 12](#_Toc175772316)

[Figure 12: Regression with Total Vehicles 13](#_Toc175772317)

[Figure 13: Cross-tabulations Pivot table 14](#_Toc175772318)

[Figure 14: Cross-tabulation Pie Chart 15](#_Toc175772319)

[Figure 15: Scatter plot for total vehicles with population and population in Urban Area 16](#_Toc175772320)

[Figure 16: Scatter plot for total vehicles with Per capita Income and population density 16](#_Toc175772321)

[Figure 17: Scatter plot for Vehicles per 1000 population with population and percentage of population in Urban Areas 17](#_Toc175772322)

# 1. Introduction

## 1.1 Purpose of the Report

The objective of this task is to understand customer shopping patterns for a UK eCommerce market for AutoMobile Inc. and that of evaluating factors affecting vehicle ownership across different Asian countries For Task A, the report seeks to present insights on the clients made through the analysis of the observable attributes such as age, sex, buying price, and the overall rating. Such synthesis of literature will reveal findings on shopping behavior and provide prescriptions for improving consumer experience and business strategies (Eger *et al.* 2021, p.102542). For Task B, the report aims to examine and analyze the factors most correlated with vehicle ownership in the Asian markets including Cyprus.

## 1.2 Audience and Stakeholders

The primary targeted readership of this report is the strategic decision-makers at AutoMobile Inc. namely the senior management and the strategic planners who need to understand factors influencing car purchase in Asian markets to plan strategic expansion. This group will find analysis of market trends and trends forecasting very helpful in their decision-making process (Grashuis *et al.,* 2020, p.5369). Furthermore, various stakeholders involved in the UK eCommerce business such as marketing executives, product managers, and customer experience specialists can use the analysis of the customer shopping trends to improve efforts and strategies towards customer relations.

## 1.3 Hypothesis or Research Question

***Null Hypothesis (H0):***

The current research also finds no moderating variable between customers' shopping patterns and the evaluated attributes (age, gender, purchase volume, and review ratings) in the UK eCommerce industry. Likewise, there is no systematic relation between the ownership of vehicles and the factors (income, population, density, and the agreed urban percent) in Asian countries.

***Alternative Hypothesis (H1):***

A strong correlation between customers' shopping behavior and the analyzed attributes in the analyzed field of the UK eCommerce market is noticeable. For example, there is a high positive relation between factors of car usage and income, population, density, and percentage of urban areas in Asian nations.

***Research Questions:***

***Q1:*** What are the key factors influencing customer shopping trends in the UK eCommerce market, and how can businesses leverage these insights to enhance their strategies?

***Q2:*** How do customer demographics, purchase behavior, and feedback ratings correlate with purchase amounts and shopping frequency?

***Q3:*** What variables most strongly correlate with vehicle ownership in Asian countries, and how do these correlations influence expansion strategies for AutoMobile Inc.?

***Q4:*** How accurate are the regression predictions for vehicle ownership in Cyprus compared to actual values, and what factors contribute to any discrepancies observed?

## 1.4 Definition of Key Terms

In this report, there are a few terms that have been defined to avoid confusion in understanding the analysis made in this report.

***Customer Shopping Trends:*** Elements noticed in consumer buying, including age, gender, quantities bought, and the review ratings given on the product.

***Descriptive Statistics:*** Techniques employed on characterization, as well as quantification of the basic characteristics of a set of numerical data including Mean, Median, Mode, Standard Deviation, and Variance (Witek and Kuźniar, 2020, p.209).

***Correlation:*** A coefficient that quantifies the degree of association between two or more variables. This report evaluates the extent of the association between car ownership and customers' socioeconomic status and purchase patterns.

***Regression Analysis:*** An inferential procedure employed with the view of estimating the impact of one or several independent variables on a single dependent variable. That is, it makes it possible to prognosticate such results as car ownership or the amount of purchases depending on other effects.

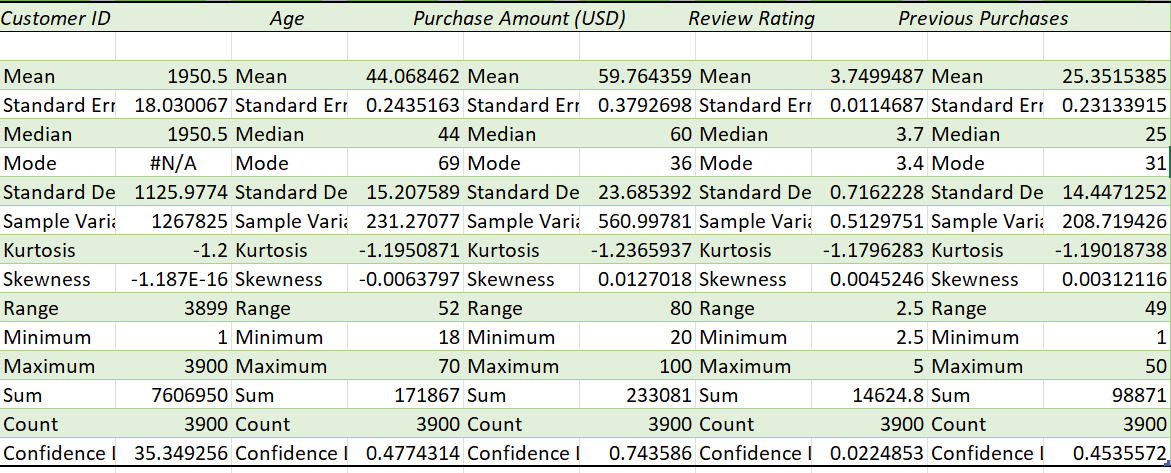
***Per Capita Income:*** This is the total amount of income that is usually generated per head of the population and has been used in this study to make inferences on vehicle ownership.

# 2. Method

## 2.1 Description of Data Presentation Methods

### Task A: Customer Shopping Trends

***Descriptive Statistics***

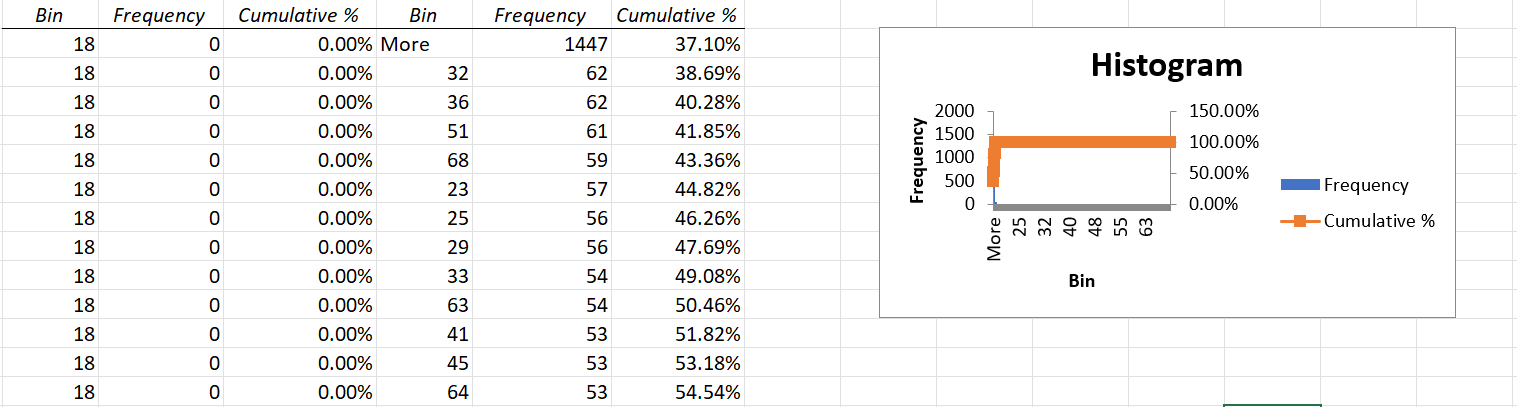


##### Figure 1: Descriptive Statistics

(Source: Extracted from Excel)

Measurement and description were used in tabular form to describe the relevant numerical data and their variability in the Customer Shopping Preferences Dataset. Attributes such as age, purchase amount, review rating, and previous purchases were analyzed, and mean, median, mode, standard deviation, and variance were determined. These statistics helped to get acquainted with the distribution and dispersion of the data to use it for identifying general tendencies of shopping behaviors (Anitha and Patil, 2022, pp.1785-1792).

***Frequency Distributions and Histograms***

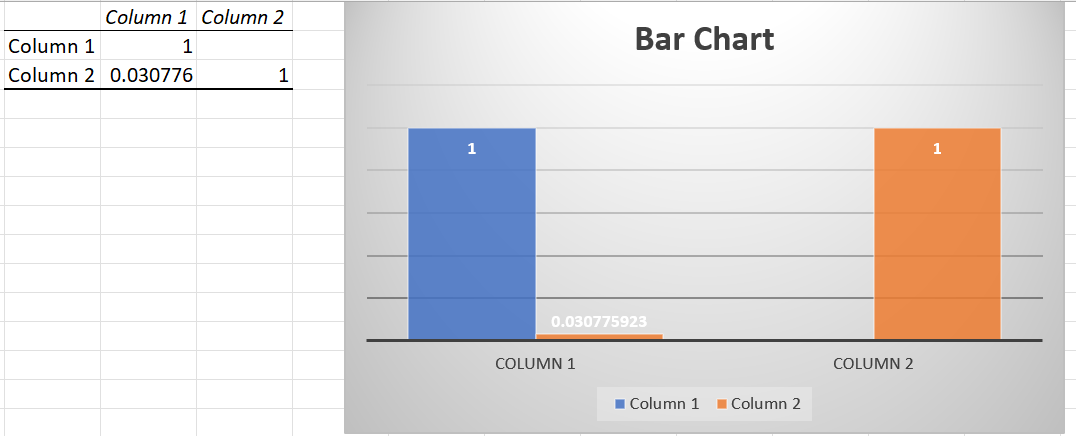


##### Figure 2: Frequency Distribution with Histogram

(Source: Extracted from Excel)

Simple tabular methods known as frequency distributions were used to tabulate and count how often numerical values occur within certain ranges such as age or amount of purchases. These distributions are depicted in histograms that show frequency and accumulative percentages of values. Through this method, patterns and flukes on customers were recognized in terms of their demographics and their expenditure (Srivastava *et al.,* 2024, pp.7-8).

***Correlation Analysis***

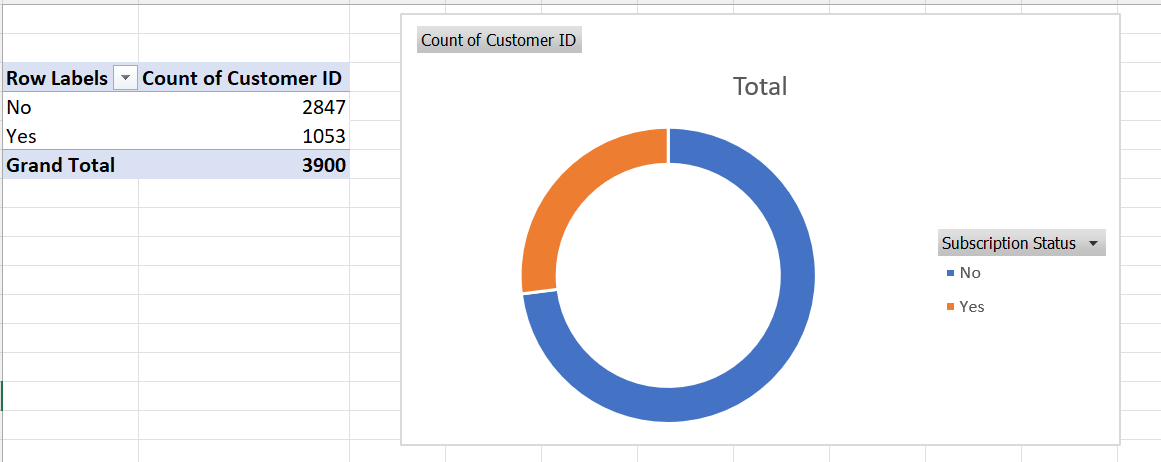


##### Figure 3: Correlation Analysis with Bar chart

(Source: Extracted from Excel)

The results of a Pearson's correlation matrix were used to determine the nature and degree of association of items such as review rating and the amount spent by the consumer. Comparing levels of correlation in this analysis offered an understanding of the relevance of the attributes that affect customers' satisfaction and their expenditure.

***Pivot Tables and Charts***



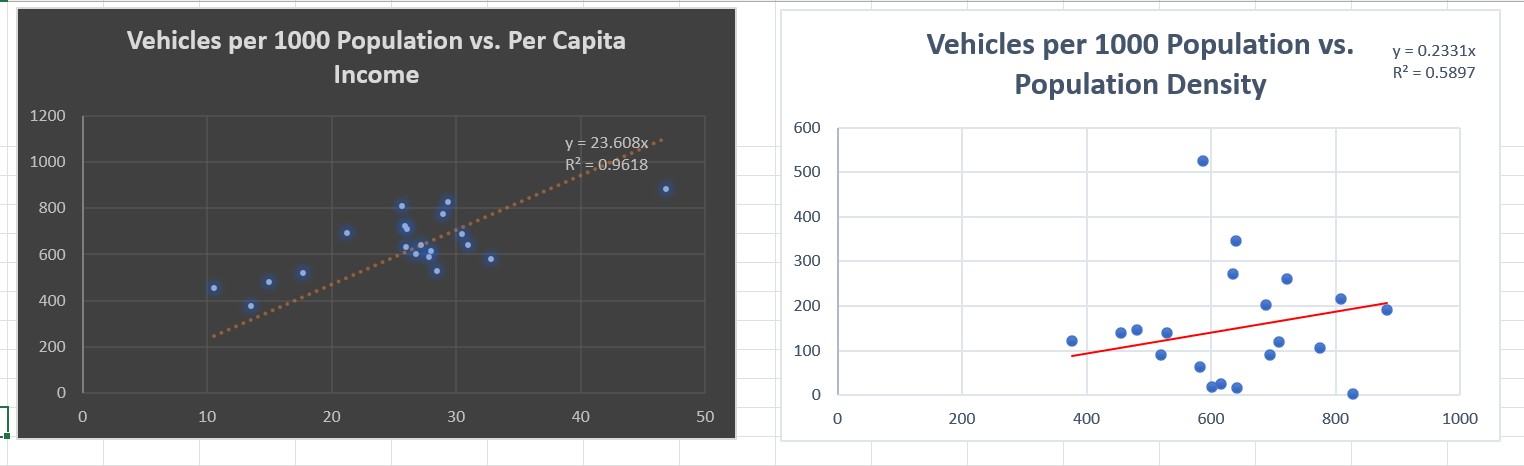
##### Figure 4: Pivot Table of Subscription Status with Donut Chart

(Source: Extracted from Excel)

Using pivot tables, purchase frequencies, amounts, and subscription status were summarized to present the data in a more aggregate format. For this aggregated information, different types of charts such as bar charts, donut charts, and pie charts were developed to enable the analysis of the trends and distributions within attributes.

### Task B: AutoMobile Inc.

***Scatter Plots***

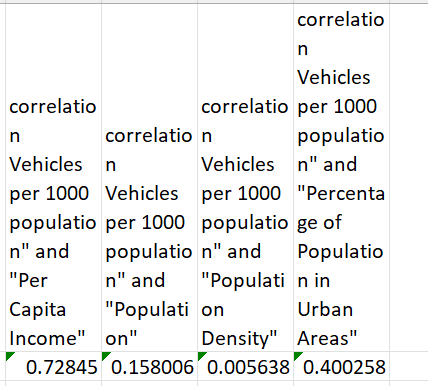


##### Figure 5: Two different Scatter Plots show the Vehicles for 1000 population with per capita income and population density

(Source: Extracted from Excel)

In this case, Bunisha employed the scatter plots to compare the vehicles with the various socio-economic indicators including income, number of population, population density, and the percentage of people in urban areas. These plots showed the extent and the nature of the relationship between variables; such plots assisted in probing the important causative determinants of car ownership.

***Correlation Analysis***

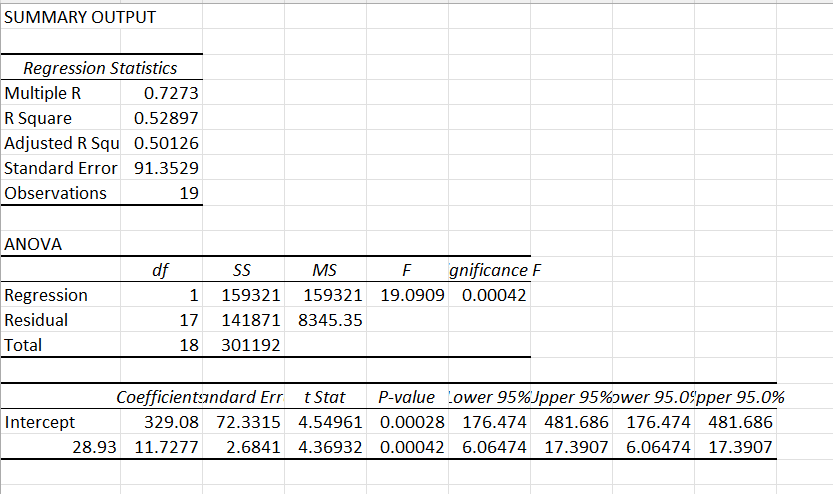


##### Figure 6: Correlation Analysis with Vehicles per 1000 population

(Source: Extracted from Excel)

Pearson product-moment coefficients of correlation were used in a computation of the strength and direction between vehicle ownership and other predictor variables. Such identification informed the following regression analysis which sought to determine the variables most closely related to vehicle ownership (Fernandes *et al.,* 2022, p.103066).

***Regression Analysis***

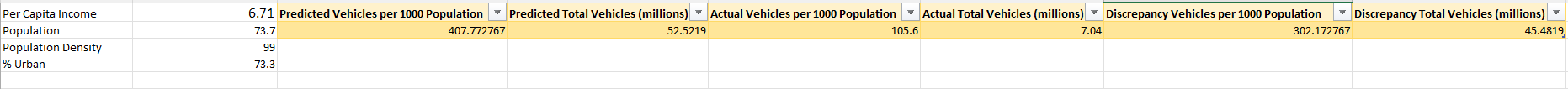


##### Figure 7: Regression Analysis

(Source: Extracted from Excel)

Regression analysis entailed finding the degradation line equation for variables that had the most relationship with vehicle ownership. This technique sought to establish the correlation between dependent and independent variables to get a predictive model on ownership of vehicles based on socio-economic characteristics.

***Predictive Analysis for Cyprus***



##### Figure 8: Cyprus Analysis Comparison

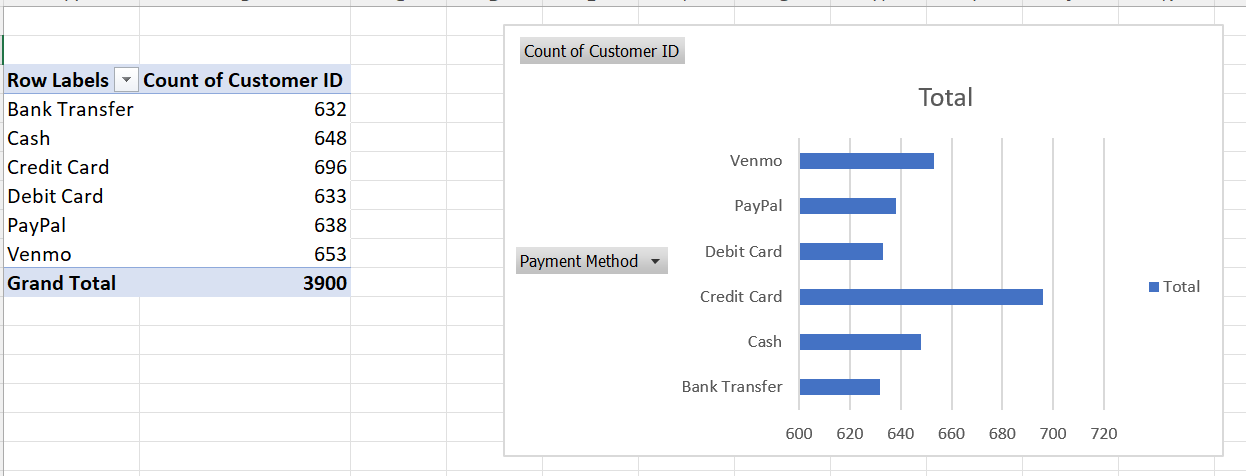
(Source: Extracted from Excel)

Applying the regression equations, American vehicle ownership and vehicles per 1000 population were also estimated. To assess the accuracy of the regression models and identify such discrepancies, the information discussed above was compared to the actual data obtained. This process also gave a comprehension of the market capacity and market verification of the forecast models.

## 2.2 Explanation of Tools and Techniques Used

### Task A: Customer Shopping Trends

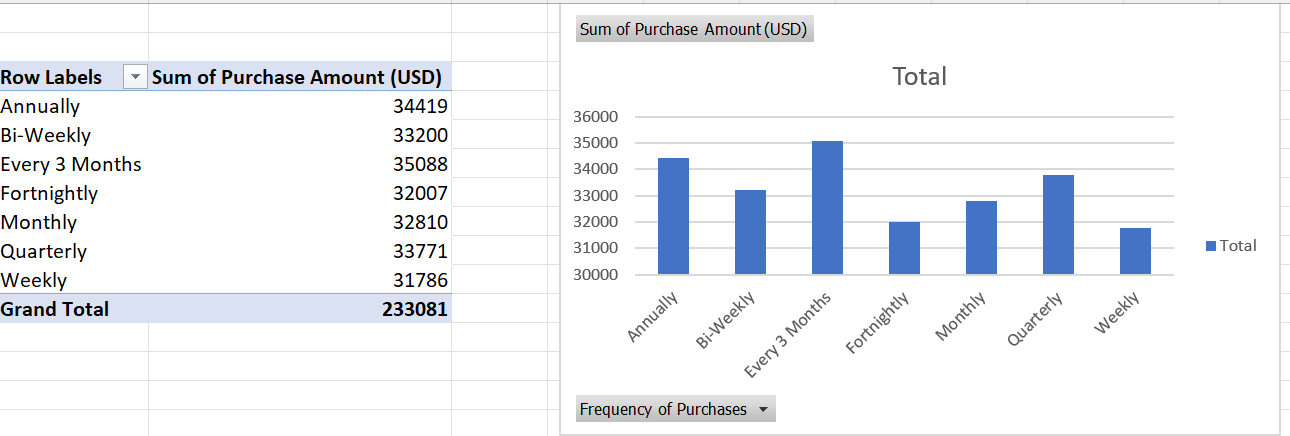
Several approaches and methods were employed in the course of Task A to examine the customer shopping patterns. Exploratory analysis of the data was conducted by calculating the descriptive statistics on the given data set by using Excel functions in order to get the basic characteristics of the data. It used basic statistics like mean, median, mode, standard deviation, and variance to analyze the overall trends of the customer's variables including age, amount of purchase, and the ratings given by the customers (Ling *et al.,* 2021, p.11719).



##### Figure 9: Pivot table of Payment Method with Bar chart

(Source: Extracted from Excel)

Frequency distributions and histograms were used to sort and illustrate raw data mostly for quantitative variables such as age and amount of the purchase. The histogram feature in Excel helps in showing the frequency of a particular value in relation to others; hence, various patterns in customer spending and the number of customers within a given age bracket were observed (Kalthaus and Sun, 2021, pp.473-510).



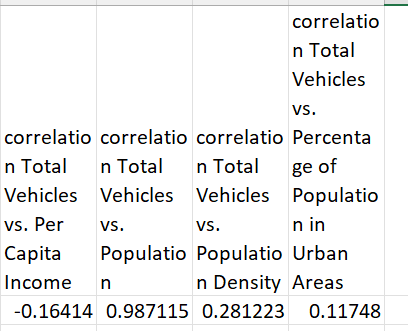
##### Figure 10: Pivot Table of Frequency of purchases with a bar chart

(Source: Extracted from Excel)

Purchase frequency and amount and the subscription status were analyzed using a pivot table in order to summarize a large data set. Using bar, donuts, and pie charts on the side, were employed to create these summaries with the intention of making the analysis and interpretation of the data easier. These charts made it easier to present trends, compare categories, and present the results in a more easily digestible fashion.

### Task B: AutoMobile Inc.

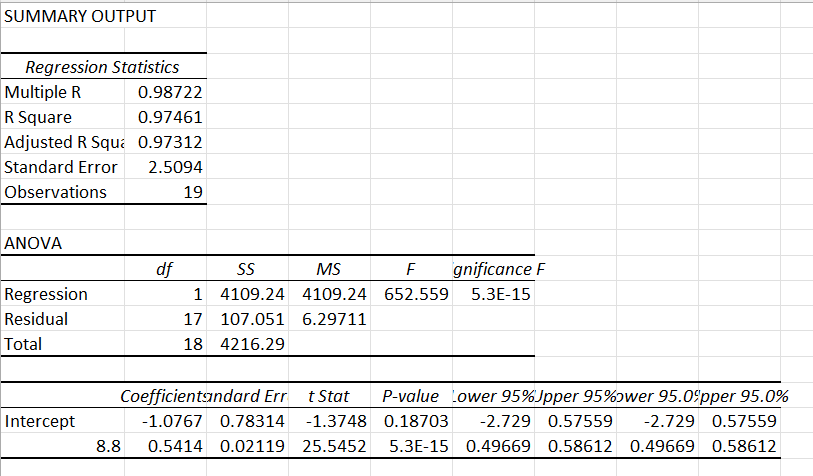
In the process of data analysis for Task B, several complex statistical methods and methods were used in relation to vehicle ownership data. This helped in drawing some hypotheses about the results on the basis of patterns that might exist in the data (Ouyang *et al.* 2021, p.112023).



##### Figure 11: Correlation with Total vehicles

(Source: Extracted from Excel)

Correlation analysis was conducted with a view of ascertaining the mutual strength and direction of the linear relationships that exist between vehicle ownership and predictor variables.



##### Figure 12: Regression with Total Vehicles

(Source: Extracted from Excel)

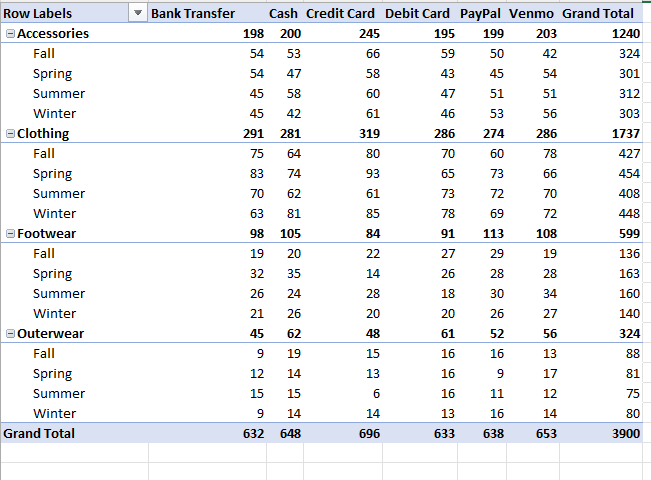
Multiple regression analysis was also performed in order to determine the regression equations for the variables that will have a high correlation with the ownership of vehicles. This included applying the method of regression analysis to estimate the equation and make predictions on owning a vehicle given appropriate socio-economic characteristics (Ling *et al.,* 2021, p.11719). For Cyprus, the predictive analysis was done by employing the regression equations earned in the previous steps. This approach also enabled checking how well the models forecasted the vehicle ownership and offered some clues into possible market prospects.

# 3. Findings and Results

## 3.1 Presentation of Data Analysis Results

### Task A: Customer Shopping Trends

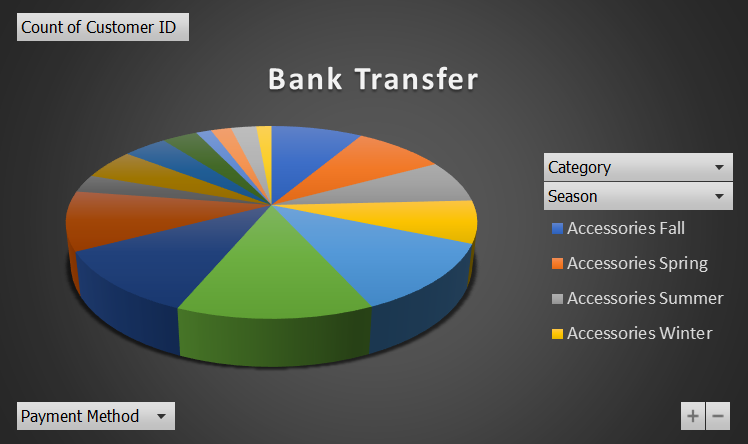
Descriptive and analytical tools were applied in the analysis of the data for Task A with the aim of evaluating customer shopping patterns. A descriptive analysis of basic customer characteristics was carried out to give the focused numerical summaries of the variables such as mean, median, and mode of customer's age, the amount of purchase, the rating of the product, and the previous purchase records. It also reveals that the average age of the customers is 44 and that the dispersion of the age is quite large considering a standard deviation of about 15. For the purchase amount it was found that it has an average of $60 and, a standard deviation of $24 depicting the variation in the amount (Acheampong and Siiba, 2020, pp.2557-2580).



##### Figure 13: Cross-tabulations Pivot table

(Source: Extracted from Excel)

Histograms shown below demonstrate where most of the purchase falls within a certain range indicating the purchasing trends or habits as well as age levels of the customers. For instance, prominent frequencies were obtained in the purchase amount of $ 36 to $ 51 per order, the cumulative percentages of which suggest that most customers tend to fall within this category (Demircan Çakar *et al.,* 2021, pp.45693-45713).



##### Figure 14: Cross-tabulation Pie Chart

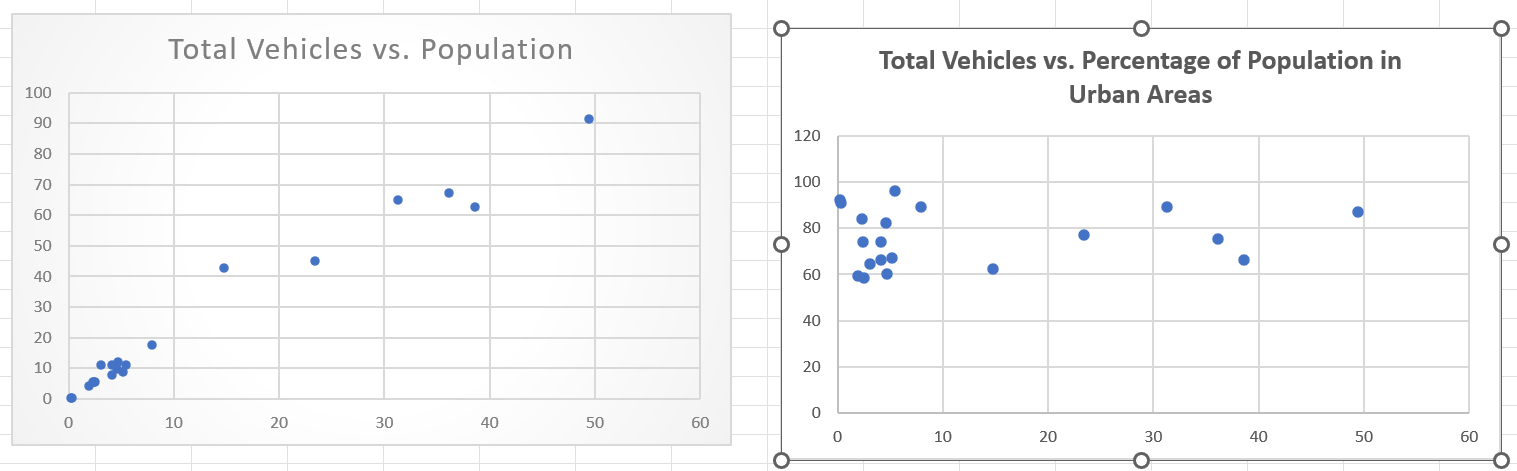
(Source: Extracted from Excel)

Pivot tables were created in order to identify how many times and for what amount the purchases were made depending on the payment method selected and the subscription status of the customer. From the bar chart depicting the frequencies of purchases made by the different methods, annual purchases stood out at the highest total amount followed by the 'bi-weekly' and the 'monthly' purchase. A donut chart of the subscription status provided a clear envision of the company where the largest portion was formed by customers who do not have subscriptions while a smaller portion comprised of customers who have subscriptions (Gonçalves *et al.,* 2021, p.113452).

Pivot tables were used to show the distribution of purchase methods with various product categories of interest. The pie chart of payment modes provided a clear picture of the mode of payment used indicating that credit cards received the most acceptance. The correlation between the reviews' rating and the amount of purchases made was again computed to be a fairly low value of correlation coefficient which was 0. 031 indi- indicating that there is very little actual relationship between the customer satisfaction ratings and the amount spent on purchasing.

### Task B: AutoMobile Inc.

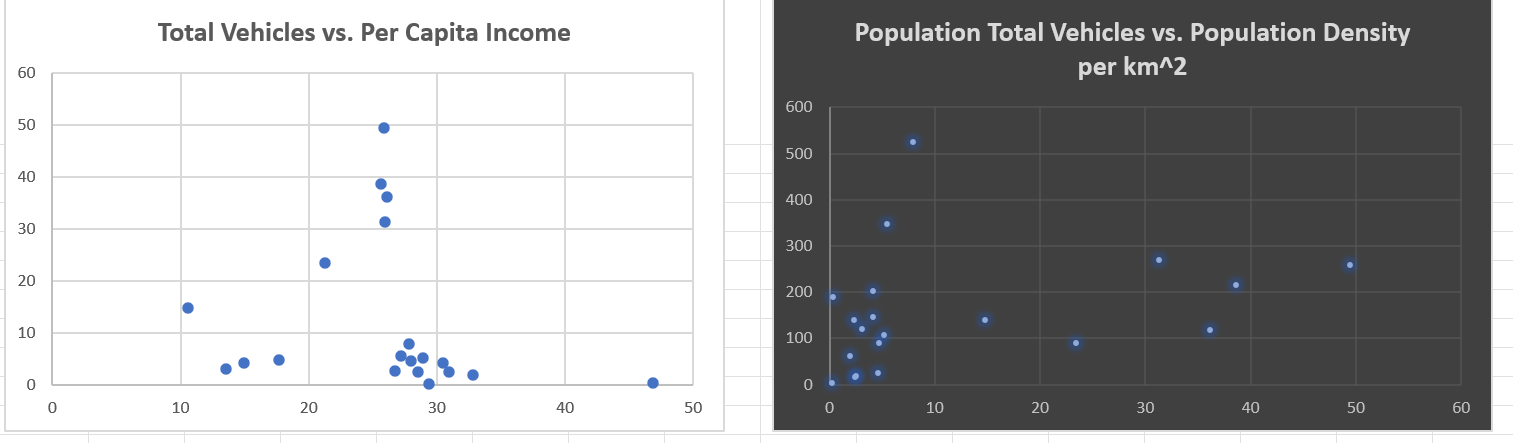
In Task B, the countries of interest were chosen to identify the factors that could determine Vehicle Ownership. Following the initial identification of the research questions, scatter plots were first employed to compare the variables of interest: vehicle ownership per 1000 population to the control variables including per capita income, population, population density, and proportion of the population that resides in cities.



##### Figure 15: Scatter plot for total vehicles with population and population in Urban Area

(Source: Extracted from Excel)

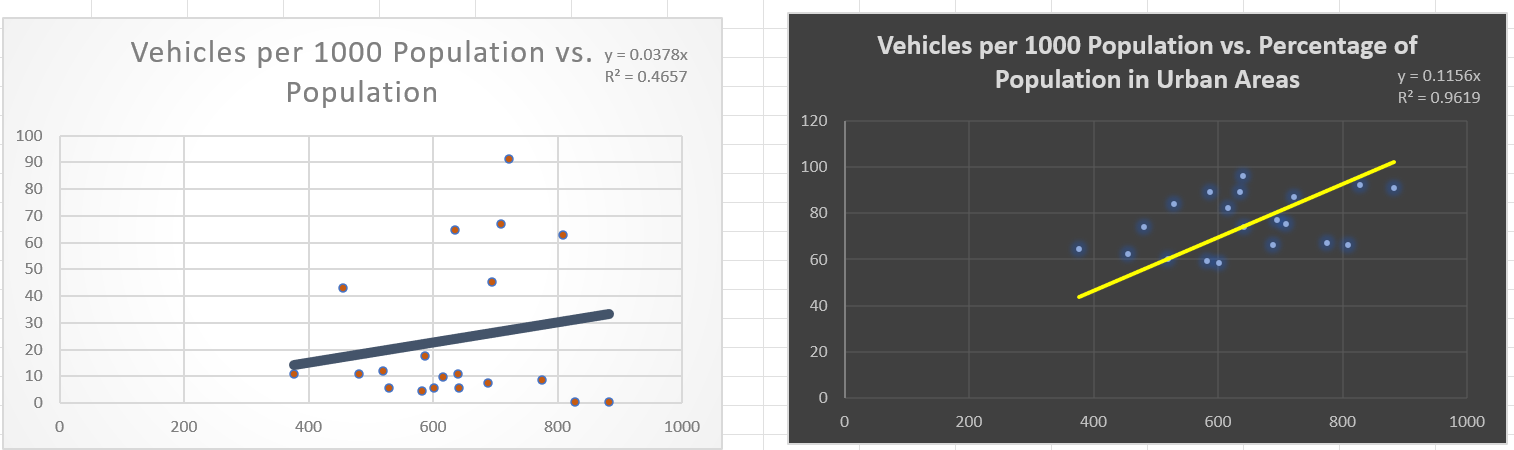
These plots offered information on how well they would fit correlation, thus a per capita income had the highest coefficient value of 0. 728 to the vehicle ownership and thus bore an extremely positive relationship (Di Leo *et al,.* 2020, p.117058).



##### Figure 16: Scatter plot for total vehicles with Per capita Income and population density

(Source: Extracted from Excel)

A regression analysis was then performed to obtain regression equations for estimating vehicle ownership against per capita income. The regression results found a moderate variance in the dependent variable explained by the independent variable which is 0. 529, this implies that the extent of variation in vehicle ownership was 53 % of the total variation and was attributed to per capita income.



##### Figure 17: Scatter plot for Vehicles per 1000 population with population and percentage of population in Urban Areas

(Source: Extracted from Excel)

A more detailed examination included scatter and regression between total vehicle ownership and the relevant factors. This led to the identification of the population as the most closely related variable with a coefficient = 0. 987. For the case of the total vehicle, the regression analysis yielded an R-squared of 0. 975, which means that the model is well-fitted with the data collected. A regression equation was able to predict vehicle ownership with a high degree of accuracy though it indicated minor disparities when it was compared with the actual values of Cyprus (Szász *et al.,* 2021, p.108085).

## 3.2 Interpretation of Data

The identification of customer shopping frequency in Task A and the factors that influence car ownership in Task B helped to understand the future tendencies in customers’ behavior and car ownership in the regions (Li *et al.,* 2021, p.107924).

The results show that the customers were on average 44 years old, and had moderate variability in spending habits; the mean amount spent was 60 $ and the standard deviation was 24. The fact that the degree of relationship between the review ratings and the purchase amounts is a mere 0. 031 points to the idea that customer satisfaction does not directly determine the amount to be spent.

The cross-tabulation for purchase frequency by payment methods and purchase amount by payment methods had the following findings: it was found that an annual payment made the largest contribution to the total buying profile and there is also variation in the use of different payment methods (Simester *et al.,* 2020, pp.2495-2522).

The comparison of the vehicle ownership of different countries in Task B, based on many factors including, per capita income, population, population density as well as the level of urbanization also presented clear patterns. In fact, a close relationship between per capita income and vehicle holdings of 0. 728 implicates the economic consideration in vehicle acquisition pointing to the fact that; higher amounts of income yield corresponding levels of holdings in vehicles. The regression model as stated above has an R-squared of =0. The study made use of 529 which helped in explaining the variance in the ownership of vehicles, though the study also revealed that there were some sort of restrictions and it was also an indication that there might be more variables that were needed for incorporating into the model (Khrais, 2020, p.226).

An examination of each factor revealed that the population was the most proximity related to car ownership with a correlation coefficient of 0. 987. The regression analysis values have been shown in the above table and the R-squared value is quite high at 0. 975, also implying a positive and significant correlation with total vehicle ownership and population size. Thus, disparities between forecasted and square values, especially in Cyprus, show that the incorporation of more variables or improvement of existing variables might improve the results (Seyedan and Mafakheri, 2020, p.53).

# 4. Discussion

## 4.1 Interpretation of Findings

The first one, which was assigned the letter A was to conduct an examination of the customer shopping tendency. Descriptive statistics indicate that the age of the/customer is over 44 years, with an average purchase amount of $ 60. The spent fluctuation ratio is moderate, proved by the standard deviation of $23 as seen from the data. 69. The low positive correlation of 0. 031 between the ratings given in the review and the number of purchases made by the customers indicates that satisfaction has little effect on the purchase volume indicating that there are other factors that compel customers to purchase such as promotions or availability. The use of frequency distributions and histograms on age and the purchase amount helped to identify the specific spending behavior, and thereby marketing segments (Anwar *et al.,* 2021, pp.722-742).

The analysis of Pivot tables and bar charts made deeper elucidations that "Annually" and "Every 3 Months" are the most common purchase frequencies.

"Annually" however indicates the highest total purchase amount of $34,419. As a result of subscription, a possible option to enhance subscriber interest through subscription-based benefits, it is found that there are 2,847 non-subscribers and 1,053 subscribers (Yilmazer and Kocaman, 2020, p.104889).

***Reason for useful of Regression***

Of the two regression equations presented; the regression for total vehicle ownership is more appropriate for AutoMobile Inc The regression model for total vehicle ownership has a high 𝑅2 of 0. 975 which shows that total vehicles have a very strong positive relation with the independent variables. This high 𝑅2 means that the model accounts for about 97% of fluctuation in total vehicle ownership and as such should be appropriate for preliminary purposes. The probability value or p-value of the coefficients is also insignificant thus affirming the statistical reliability of the model prediction.

On the other hand, the regression of vehicles per thousand per capita with driver deaths is comparatively lower having an 𝑅2 value of 0. , which represents 529, this explicating 53% of the variance in the data set. As effective as this is, it does not encompass as much variability of the dependent variable as the total vehicle ownership model above does.

***Findings for Task B***

Under the payment method analysis, the findings showed that credit cards are the most popular form of Payment with a total of 696. Such preference requires the fine-tuning of the payment processing mechanisms and approaches. The analysis of payment types by-products also revealed that consumers have diverse preferences of payment methods when it comes to the various products and credit cards are preferred for accessories and clothing (Onyele and Nwadike, 2020, pp.18-26).

Task B attempted to analyze vehicle possession within Asian nations. The findings of correlation analysis also projected economic variables as influential correlates of vehicles through the direct positive correlation (0. 728) with per capita income and vehicle ownership. In regression analysis, the coefficient of determination is 0. In both studies, 1 wrote that per capita income is a significant factor associated with car ownership/favicon. Additional regression of total vehicle ownership as compared with the population had a very close relationship with a coefficient of determination of 0. 987. 975; The obtained results suggest that population number has a direct positive relationship with the number of cars owned. This high accuracy increases the confidence in the use of population metrics in estimating the trend of vehicle ownership (Tripathi *et al.* 2021, pp. 315-320).

## 4.2 Exploration of Correlations or Causations

The relatively low coefficient of 0. 031 of review ratings and purchase amount implies that customers' satisfaction may not affect spending immensely. This means that factors, for instance, product promotions or even customer incentives, may have a stronger inclination than this concerning the amount of the purchase.

In Task B, we personally establish a high correlation (0. 728) of per capita income with vehicle ownership indicating the presence of strong economic factors leading to the purchase of vehicles. The result of the regression of the composite score is as follows: The model established an R-squared of 0. 529, which also shows that as the income level rises the units of vehicles owned also correspondingly increase. On the other hand, the weak link between car ownership and a population density coefficient of 0. 0056 makes it rather evident that those economic factors are crucial, though population density itself is not very influential (Camilleri, 2020, pp.50-63).

The near unity or perfect positive correlation coefficient of 0. 987 assigns the total vehicle population and the numerical strength of a country's population. This indicates the high reliability of the model with the R-squared value of 0 975 thereby validating the notion that population metrics are accurate indicators of ownership of automobiles.

## 4.3 Comparison with Expectations or Hypotheses

Regarding Task A, the null hypothesis which assumed that customer review ratings would have a significant relationship with purchase amounts did not find support as the correlation was found to be 0. 031. Therefore, it can be concluded that elements other than satisfaction are far more significant in shaping a customer's behavior, these may include marketing techniques or special offers during the Christmas period (Camilleri, 2020, pp.50-63).

Concerning Task B, the hypothesis that postulated that economic variables would have a huge influence on vehicle ownership was supported. The graphs and regression analysis also showed that as income per capita rose, the number of vehicles also rose, and also proved with the correlation coefficient of 0. 728 proved that there was a strong direct relationship between the two.

The hypothesis that there would be a high positive correlation between population size and vehicle ownership was also supported logically and statistically as noted by a near perfect positive correlation coefficient of (0. 987) and a fairly high coefficient of determination of (0. 975) respectively. This result supports the hypothesis that demographic factors are important to understanding vehicle trends, as population size is directly proportional to vehicle ownership.

# 5. Conclusion and Recommendations

## 5.1 Summary of Key Findings

In Task A, the major concern was to identify customers' shopping characteristics. The descriptive statistics summary gave the details on the customer's age, gender, purchase value, rating, and the-frequency of their purchases. These are; the mean purchase amount which stands at a mean of 59 dollars. , 76, and a mean of 3 for the review ratings. 75 representing reasonable satisfaction among Ethereum's clients. From the histograms of age and purchase amounts it is clear that most customers fell within the 32 to 68 age range and there were more amounts that were more than the average purchase amount of $200 in this age range. In Task B, the focus was on aspects of car acquisition in various Asian countries and territories. The correlation analysis also pointed to per capita income as the best predictor of vehicle ownership with a high correlation coefficient of 0. 728. The regression analysis also supported this conclusion indicating that there is a positive relationship between the per capita income and vehicle possession. Furthermore, the close relationship between total vehicles per state and state population size (r=0. 987) thus supports the hypothesis that the more the population, the more the number of vehicles on the road.

## 5.2 Assessment of Hypothesis

The hypothesis that per capita income has a large impact on vehicle ownership was confirmed by the analysis of data. The results that were obtained from the correlation and regression analysis of Task B revealed that per capita income was closely related to the number of vehicles per 1000 people and the value of the coefficient of correlation was 0. ,728, and R-squared of 0. The coefficient of determination, R²=0.529, shows that the income variables explained a significant proportion of total variance in vehicle ownership. Moreover, there was a significant relation between total vehicles and the population size of the state which was evidenced by the high-value Co-efficient of correlation (0. 987) proving that vehicle ownership did change with the population density.

## 5.3 Recommendations

It is suggested that the attention of the policymakers be shifted towards the economic factors that determine ownership of cars. For instance, the effort that aims at enhancing the per-capita income may stand a better chance of enhancing the proportion of vehicle ownership or automobiles per one thousand individuals. Further, investment into infrastructure and the public transport domain should also be looked at as a consideration since population and level of urbanization also affect car usage. Appearance and improvement of public transport can enable the reduction of Private car ownership gaps in densely populated areas. Also, arising from policies that aim at enhancing economic growth and higher per capita income in the developing zones, there will be a corresponding rise in the ownership of vehicles in consonance with the trends shown.

# References

Acheampong, R.A. and Siiba, A., 2020. Modelling the determinants of car-sharing adoption intentions among young adults: The role of attitude, perceived benefits, travel expectations and socio-demographic factors. Transportation, 47(5), pp.2557-2580.

Anitha, P. and Patil, M.M., 2022. RFM model for customer purchase behavior using K-Means algorithm. Journal of King Saud University-Computer and Information Sciences, 34(5), pp.1785-1792.

Anwar, A., Thongpapanl, N. and Ashraf, A.R., 2021. Strategic imperatives of mobile commerce in developing countries: the influence of consumer innovativeness, ubiquity, perceived value, risk, and cost on usage. Journal of Strategic Marketing, 29(8), pp.722-742.

Camilleri, M.A., 2020. The use of data-driven technologies for customer-centric marketing. International Journal of Big Data Management, 1(1), pp.50-63.

Demircan Çakar, N., Gedikli, A., Erdoğan, S. and Yıldırım, D.Ç., 2021. A comparative analysis of the relationship between innovation and transport sector carbon emissions in developed and developing Mediterranean countries. Environmental Science and Pollution Research, 28(33), pp.45693-45713.

Di Leo, S., Caramuta, P., Curci, P. and Cosmi, C., 2020. Regression analysis for energy demand projection: An application to TIMES-Basilicata and TIMES-Italy energy models. Energy, 196, p.117058.

Eger, L., Komárková, L., Egerová, D. and Mičík, M., 2021. The effect of COVID-19 on consumer shopping behaviour: Generational cohort perspective. Journal of Retailing and consumer services, 61, p.102542.

Fernandes, S., Panda, R., Venkatesh, V.G., Swar, B.N. and Shi, Y., 2022. Measuring the impact of online reviews on consumer purchase decisions–A scale development study. Journal of Retailing and Consumer Services, 68, p.103066.

Gonçalves, J.N., Cortez, P., Carvalho, M.S. and Frazao, N.M., 2021. A multivariate approach for multi-step demand forecasting in assembly industries: Empirical evidence from an automotive supply chain. Decision Support Systems, 142, p.113452.

Grashuis, J., Skevas, T. and Segovia, M.S., 2020. Grocery shopping preferences during the COVID-19 pandemic. Sustainability, 12(13), p.5369.

Kalthaus, M. and Sun, J., 2021. Determinants of electric vehicle diffusion in China. Environmental and Resource Economics, 80(3), pp.473-510.

Khrais, L.T., 2020. Role of artificial intelligence in shaping consumer demand in E-commerce. Future Internet, 12(12), p.226.

Li, Y., Chu, X., Tian, D., Feng, J. and Mu, W., 2021. Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm. Applied Soft Computing, 113, p.107924.

Ling, Z., Cherry, C.R. and Wen, Y., 2021. Determining the factors that influence electric vehicle adoption: A stated preference survey study in Beijing, China. Sustainability, 13(21), p.11719.

Onyele, K.O. and Nwadike, E.C., 2020. Dynamics of stock returns amidst COVID-19 lockdown: A descriptive analysis of global stock markets. International Journal of Scientific Research in Multidisciplinary Studies, 6(12), pp.18-26.

Ouyang, D., Zhou, S. and Ou, X., 2021. The total cost of electric vehicle ownership: a consumer-oriented study of China's post-subsidy era. Energy Policy, 149, p.112023.

Seyedan, M. and Mafakheri, F., 2020. Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. Journal of Big Data, 7(1), p.53.

Simester, D., Timoshenko, A. and Zoumpoulis, S.I., 2020. Targeting prospective customers: Robustness of machine-learning methods to typical data challenges. Management Science, 66(6), pp.2495-2522.

Srivastava, P.R., Eachempati, P., Panigrahi, R., Behl, A. and Pereira, V., 2024. RETRACTED ARTICLE: Analyzing online consumer purchase psychology through hybrid machine learning. Annals of Operations Research, 337(Suppl 1), pp.7-8.

Szász, L., Csíki, O. and Rácz, B.G., 2021. Sustainability management in the global automotive industry: A theoretical model and survey study. International Journal of Production Economics, 235, p.108085.

Tripathi, A., Bagga, T., Sharma, S. and Vishnoi, S.K., 2021, January. Big data-driven marketing enabled business performance: A conceptual framework of information, strategy and customer lifetime value. In 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 315-320). IEEE.

Witek, L. and Kuźniar, W., 2020. Green purchase behavior: The effectiveness of sociodemographic variables for explaining green purchases in emerging market. Sustainability, 13(1), p.209.

Yilmazer, S. and Kocaman, S., 2020. A mass appraisal assessment study using machine learning based on multiple regression and random forest. Land use policy, 99, p.104889.

# Appendix

