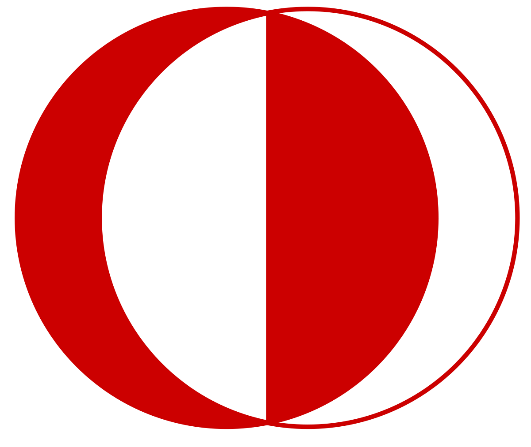
****

**Middle East Technical University**

**Department of Statistics**

STAT 364 – Linear Models II

FINAL REPORT

 “**Credit Card Usage”**

 BY

Alper Tunahan ÖZTÜRK 2290856

Dayanch AKMYRADOV 2347292

Hamza Anıl YILDIRIM 2218386

Kubilay TAŞYÜREK 2218329

June 2021

Table of Contents

[**Introduction** 2](#_Toc74905619)

[**Aim of Research** 3](#_Toc74905620)

[**Data description** 3](#_Toc74905621)

[**Analysis and Results** 4](#_Toc74905622)

[**Research Question 1:** Which variables are likely to best predict average monthly spend? 4](#_Toc74905623)

[**Research Question 2:** Does average monthly travel spend differ among age groups? 8](#_Toc74905624)

[**Research Question 3:** Does the average credit card balance decrease when the annual percentage rate offered is higher? 11](#_Toc74905625)

[**Research Question 4:** Does average monthly balance significantly differ for student loan indicators or income groups? 14](#_Toc74905626)

[**Conclusion** 16](#_Toc74905627)

[**References** 17](#_Toc74905628)

# **Introduction**

Credit cards are an essential and irreplaceable part of our lives because of their advantages and fastness among methods of payment. “Credit cards allow people to take money from the bank with the credit limit without any costs, and this includes all the time and effort involved with obtaining a loan from a financial association. This simplicity allows many consumers to pay more interest rates on exceptional credit card balances, instead of spending time applying for a loan with a lower interest rate. To conclude, credit cards account for a substantial and growing share of people's debt” (Canner and Luckett 1992).

If someone is aged over 18, he/she can easily apply for a credit card and get it. He/she can arrange the credit card limit according to income. Credit card owners can spend their money, however, whenever they want as much as their credit card limit, even without cash, and this is the most important point of the credit card. Then, at the end of the month, people must pay the money they spent whether in one payment or pay by installment with an extra interest rate. Also, if people do not pay their credit card debts for a long time, banks add an additional interest rate to credit cards.

There are many banks with various offers for credit cards such as special offers, method of payments, interest rates. “We need to make a point here that we are not talking against the constant use of some credit cards of banks, on the contrary, we say that these uses of cards may lead to benefits for payment purposes. Hence, we know that banks can offer which is user’s benefits such as points that can be applied to travel, vacation, discount, or other rewards”(Sarofim, S., Chatterjee, P., & Rose, R. (2020)).

     To sum up, credit cards are so important in today’s world and credit cards affect people’s spending behaviors in different areas such as grocery, dining, entertainment, travel, apparel. This project will give enlightenment about credit card spending among different areas, and other variables related to credit card spending.

# **Aim of Research**

Credit cards have become an indispensable part of our lives as people do not like to carry a lot of cash. Even people who do not like credit cards have to use one at some point in their lives. Because the era is evolving towards technology and technology brings innovations. This study analyses credit card data to have insight into certain aspects of people’s behavior. Finding out if there is a relationship between credit card usages and the factors that affect this usage. Also, the spending made on vacation among different age groups to understand the spending difference will be analyzed. There are times that people need extra money, albeit unintentionally. The money loaned from the bank in the short term is called balance. It is aimed to find; would people be willing to borrow money from the bank if the annual percentage were high. Banks apply different interest rates according to different professions and different age groups. Further analysis will be made in this research. According to results, banks and human resources may be able to shape their strategy to reach more customers, their intended use of credit cards. Improvements can be focused on the direction of greatest interest.

# **Data description**

Data used in this study is retrieved from the “Analyttica TreasureHunt Leaps” website[1]. The dataset consists of 1000 observations. Data includes 17 variables, 13 of which this study is interested in. These variables are as follows:

Continuous variables: average monthly spend, average monthly balance, average monthly pay rate, grocery spend, travel spend, apparels spend, dining, and entertainment spend.

Discrete variables: times mailed, number of tradelines.

Categorical variables: age groups, income groups, offered annual percentage rate, student loan indicator.

# **Analysis and Results**

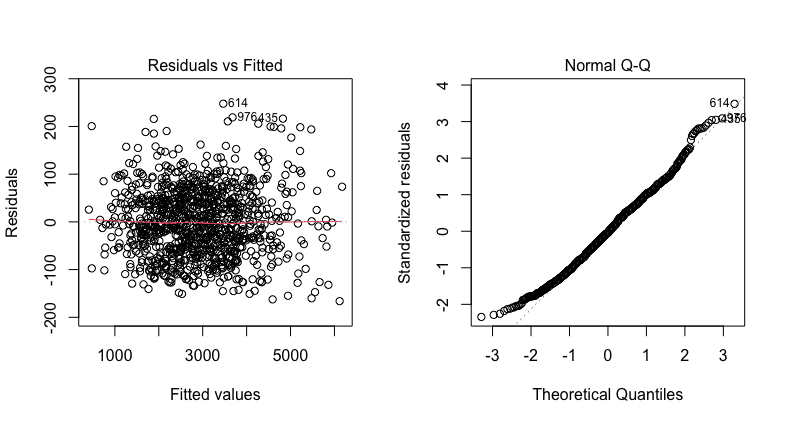
## **Research Question 1:** Which variables are likely to best predict average monthly spend?

In this research question, multiple linear regression is used. In multiple linear regression average monthly spend taken as response and all the other variables except “**Prospect ID**”, “**Channel**”, “**response \_ind**”, and “**Autoloan\_id**”  are taken as covariates.

One of the assumptions of Multiple Linear Regression is multicollinearity, to check the multicollinearity VIF method which is in the below used.

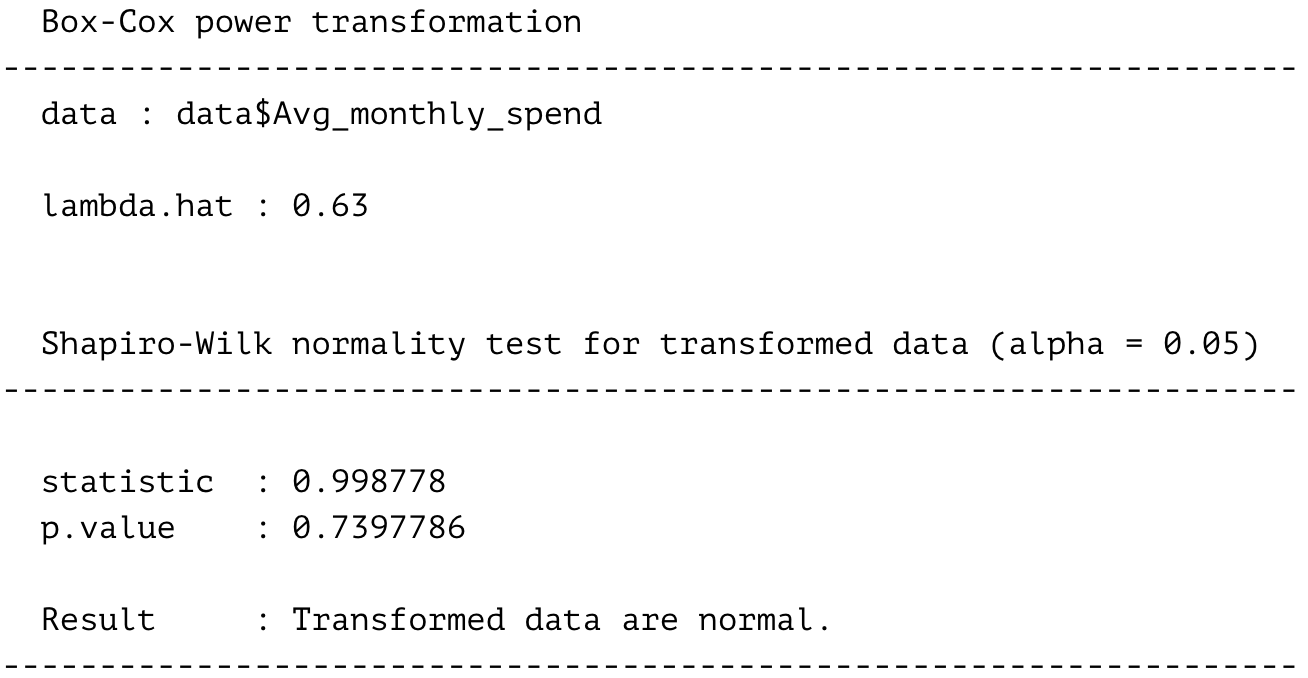
|  |  |  |  |
| --- | --- | --- | --- |
|  | GVIF | df | GVIF^(1/(2\*Df)) |
| Age\_group | 1.043766 | 3 | 1.007165 |
| APR\_offered | 1.069013 | 3 | 1.011185 |
| Grocery\_spend | 1.445107 | 1 | 1.202126 |
| Travel\_spend | 2.001091 | 1 | 1.414599 |
| Apparels\_spend | 1.312185 | 1 | 1.145506 |
| DE\_spend | 1.025421 | 1 | 1.012631 |
| Income\_Group | 1.070876 | 4 | 1.008596 |
| Autoloan\_ind | 1.076440 | 1 | 1.037516 |
| Avg\_Monthly\_balance | 2.527336 | 1 | 1.589760 |
| Avg\_monthyl\_payrate | 1.325159 | 1 | 1.151156 |
| student\_loan\_ind | 1.015379 | 1 | 1.007660 |
| Times\_Mailed | 3.866791 | 1 | 1.966416 |
| No\_of\_tradelines | 3.470460 | 1 | 1.862917 |

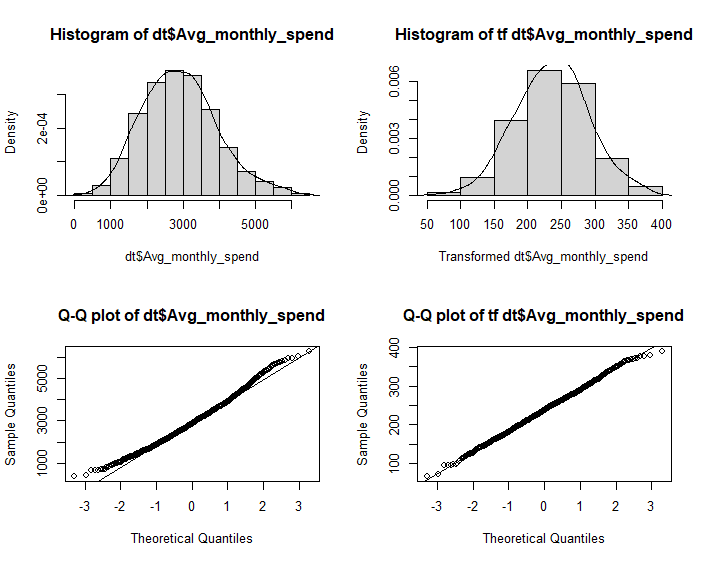
Since the VIF values are less than 10 (or 5), there is no multicollinearity problem.



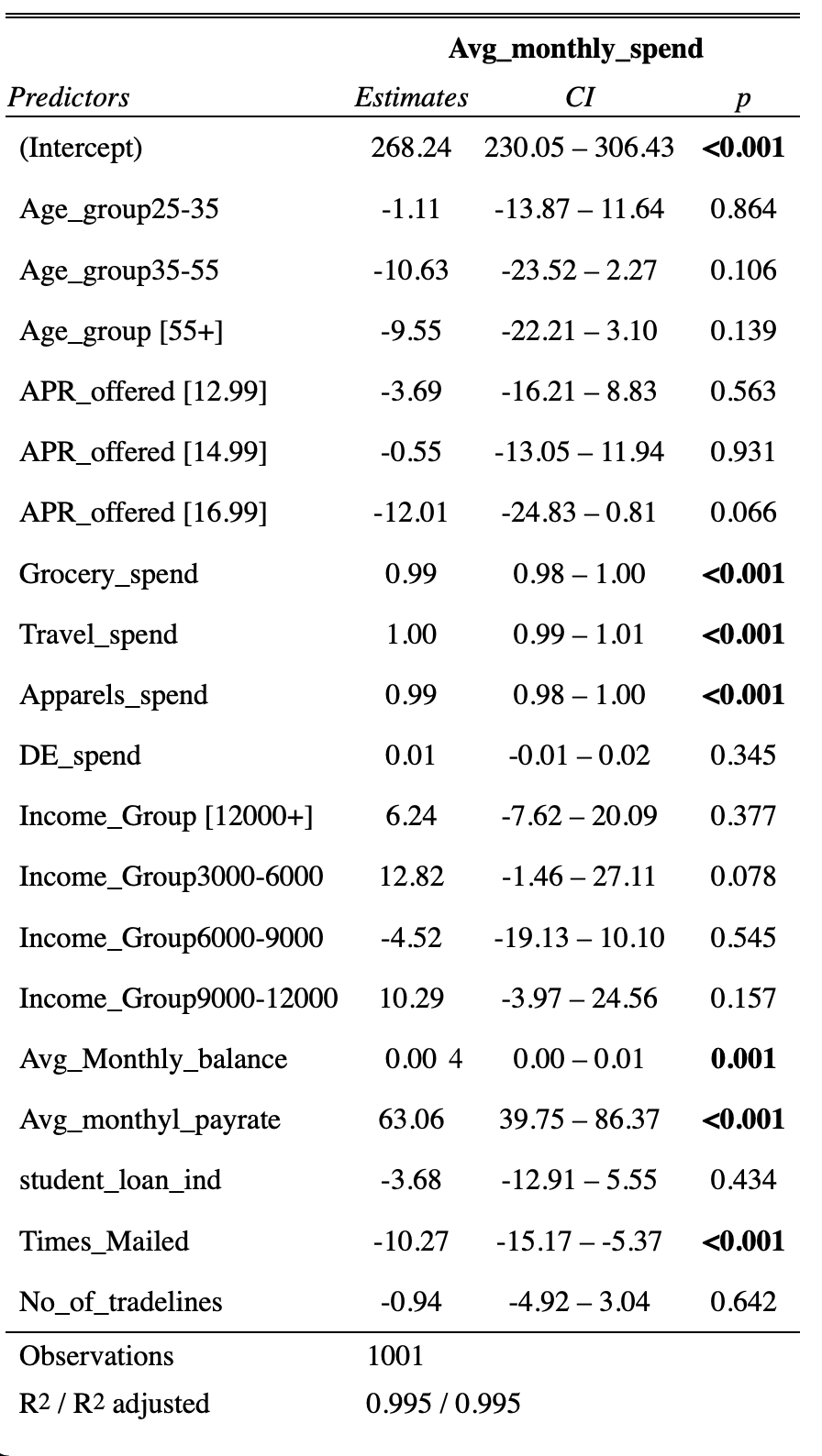
There is no pattern in the residuals versus fitted values plot, so residuals seem to have constant variance.

Shapiro-Wilk Normality test for response variable (Average monthly spend) gives W = 0.98993 with p-value = 2.322e-06, which indicates that the variable is not normal. It might be due to the test being too sensitive for a large sample size since Normal Q-Q suggests it is acceptable, nonetheless, transformation is applied to see if it can be normalized. Using Box-Cox transformation:



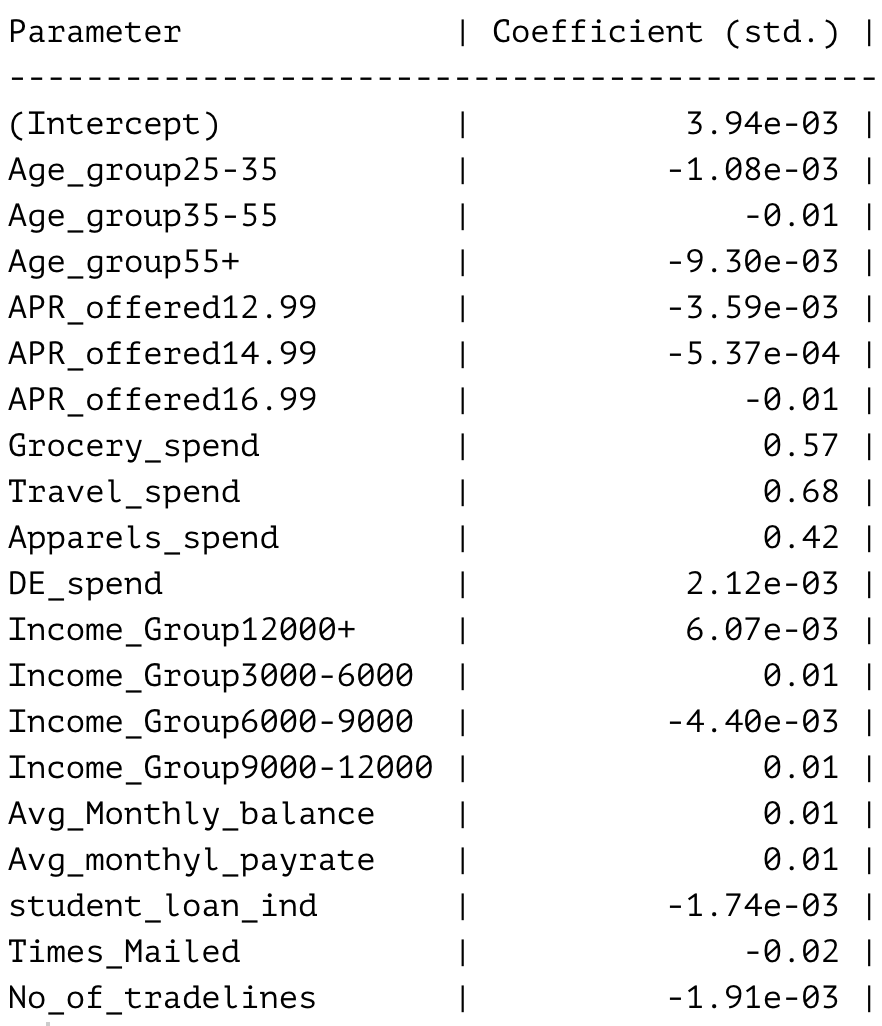


Normal Q-Q plot and normality test results indicate that transformed data is normalized. But there is not much difference between actual data and transformed data and constant variance of residuals are violated in the transformed data, so a model with original data will be carried out.



There are some variables significantly contributing to the model. Insignificant variables will not be considered. Significant variables are grocery spend, travel spend, apparels spend, average monthly balance, average monthly pay rate, and times mailed.

To compare which variables, contribute the most, the model is standardized. Results are as follows:



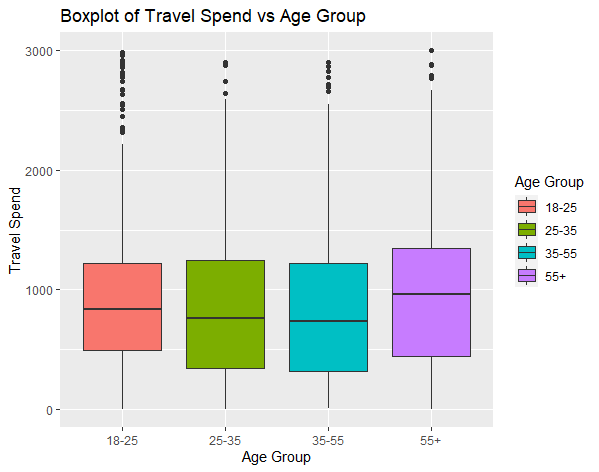
As seen from the above table, in descending order the most contributions come from travel spend, grocery spend, and apparels spend.

## **Research Question 2:** Does average monthly travel spend differ among age groups?

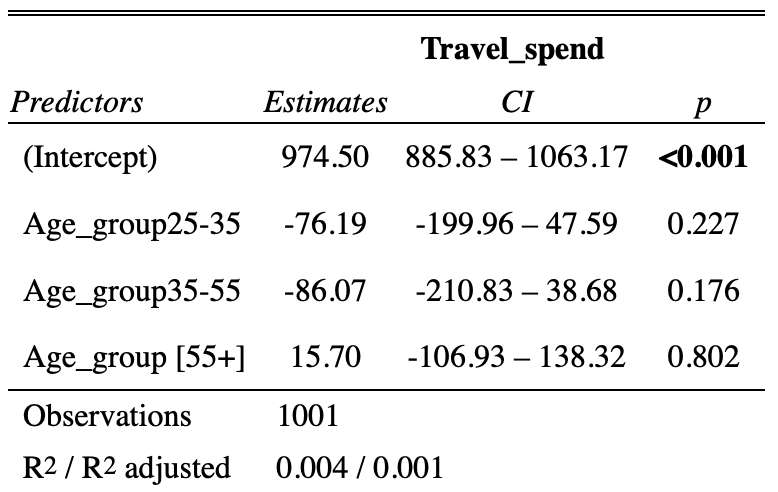
In this research question, simple linear regression is used. In regression average monthly travel spend is taken as predictor and age groups taken as the dependent variable.

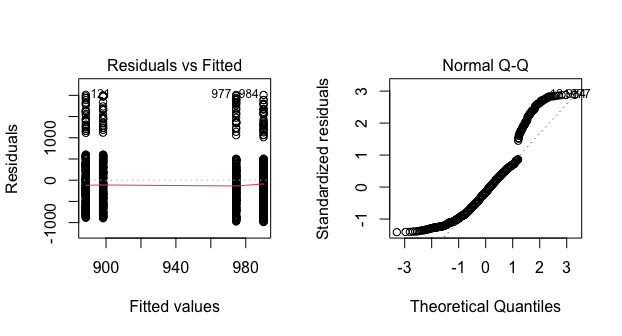
H0: Average monthly travel spend does not differ among age groups.

H1:  Average monthly travel spend differs for at least one age group.

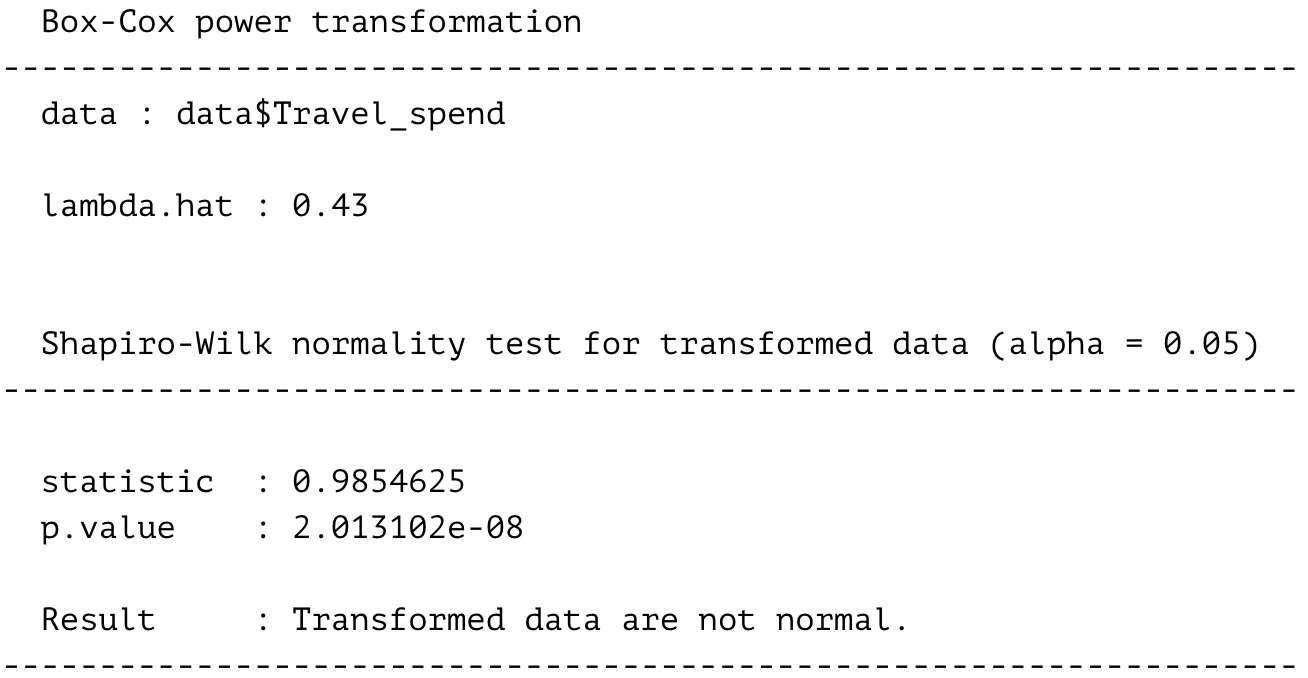


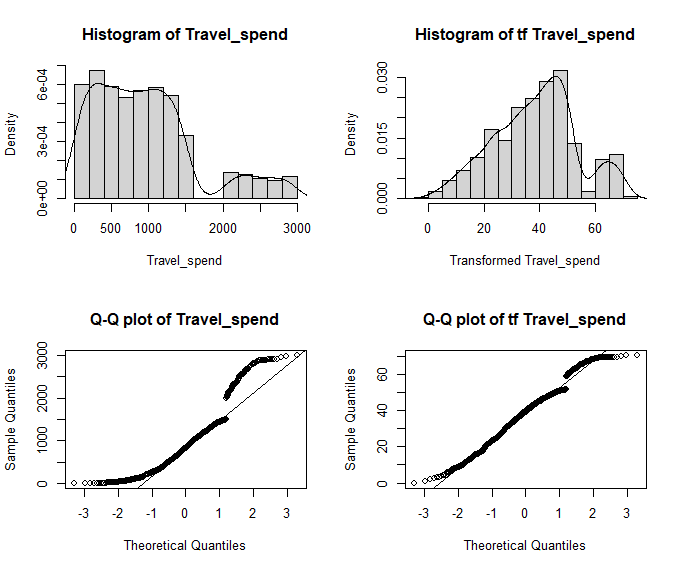
Considering boxplot there does not seem to be significant differences among age groups for travel spend.





Residuals do not seem to follow the normal distribution. There is no problem with constant variance. For normalization, a transformation of data is needed. Applying Box-Cox:



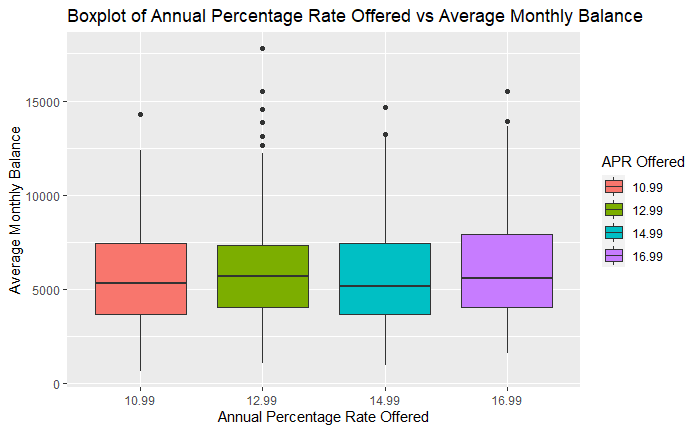


Transformed data did not fix the normality problem. So, robust regression was considered to see if there is a significant change in the models. But robust regression did not differ in terms of the significance of estimates. Hence, since the model indicates that estimates are insignificant, the null hypothesis is not rejected. There is no evidence to support that age groups make any difference regarding average monthly travel spend.

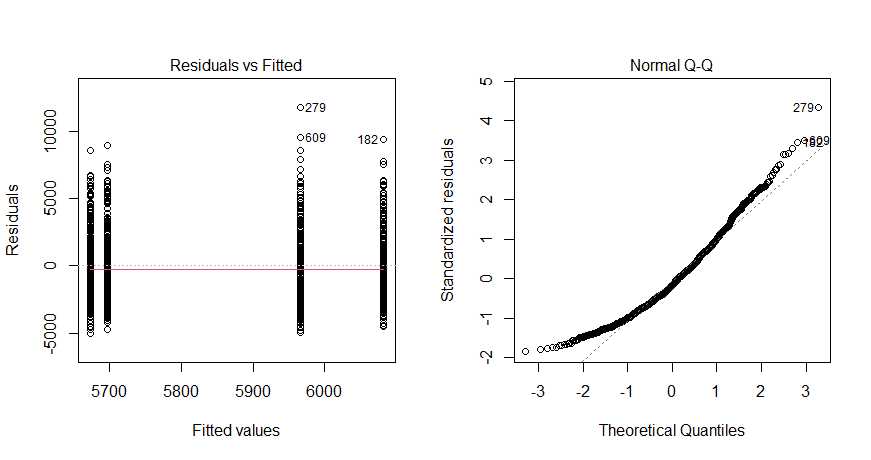
## **Research Question 3:** Does the average credit card balance decrease when the annual percentage rate offered is higher?

H0: Average credit card balance does not decrease when APR offered increases.

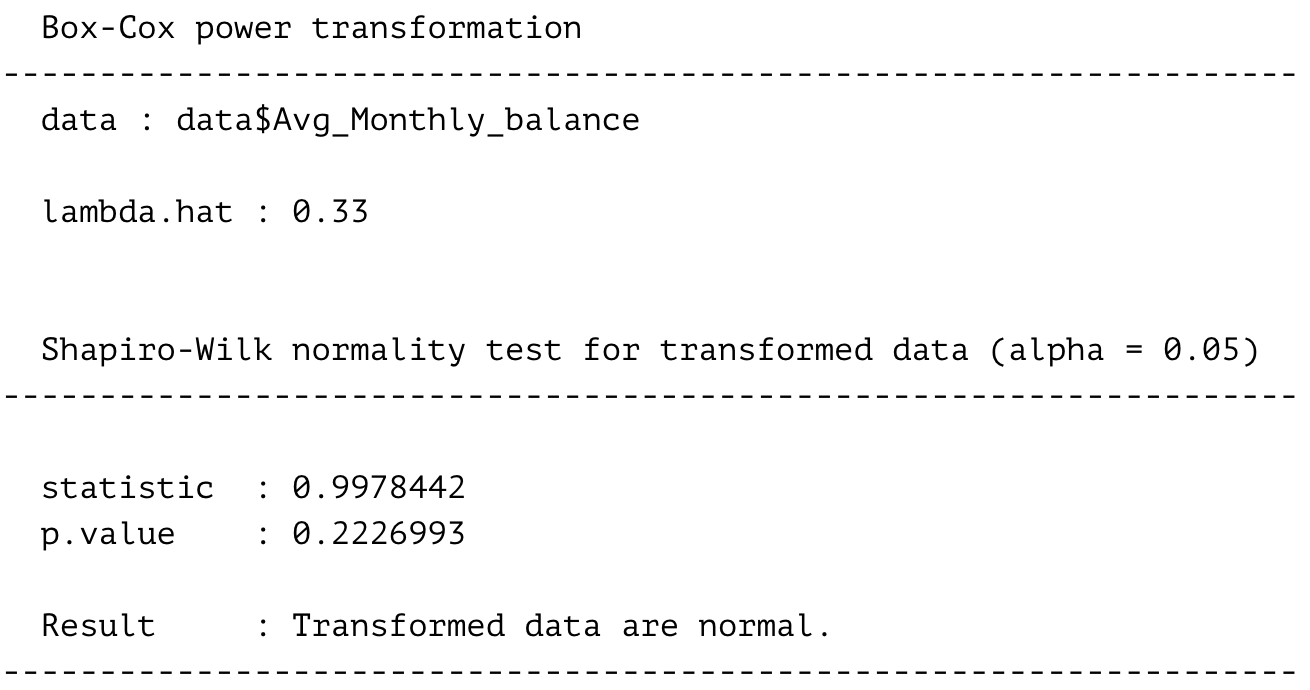
H1:  Average credit card balance decreases when APR offered increases.

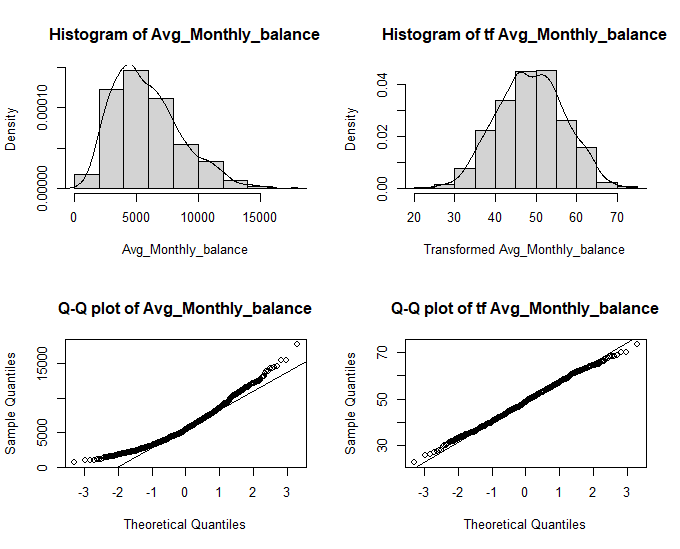


Boxplot indicates there is no significant difference in average monthly balance among APR offers.

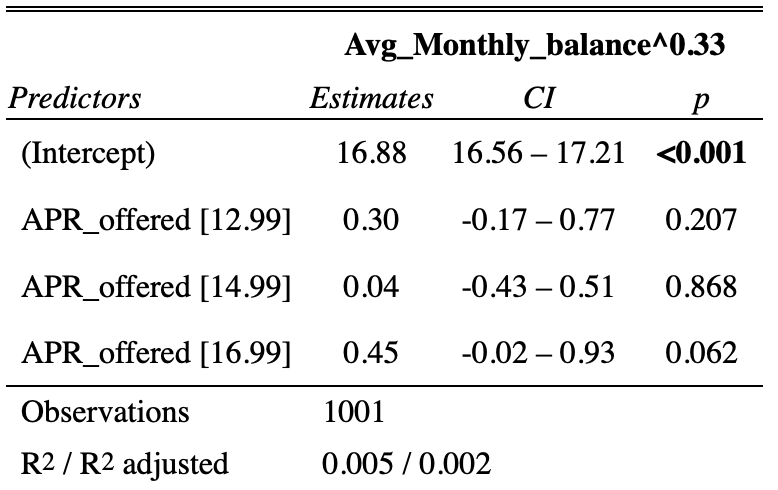


Residual versus fitted plot seems to indicate constant variance. The normal Q-Q plot seems to indicate there may be a need for transformation. Shapiro-Wilk normality test gives a significant p-value, indicating data do not follow the normal distribution. Applying Box-Cox:





Transformation on data fixed the normality problem and constant variance was not violated. The model is fitted with the transformed response variable.

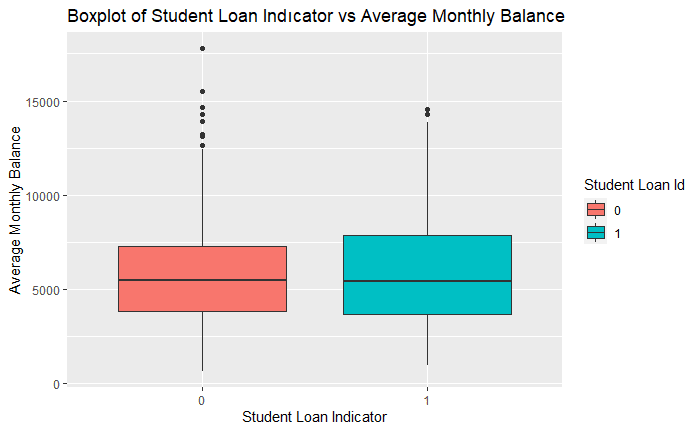


None of the estimates are significant since p-values for all covariates are less than 0.05. The null hypothesis is not rejected. There is no evidence to conclude that the high APR offered decreases the average monthly balance.

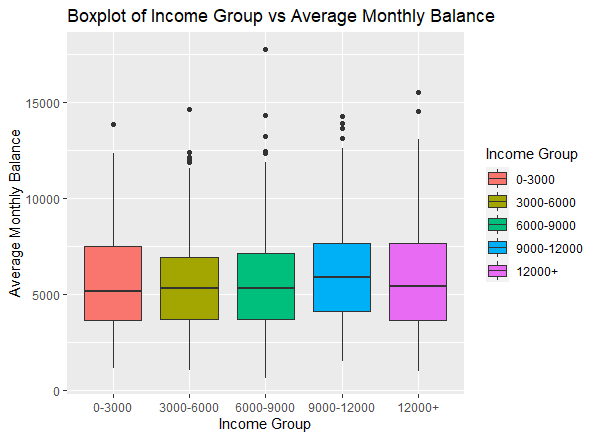
## **Research Question 4:** Does average monthly balance significantly differ for student loan indicators or income groups?

H0: Average monthly balance does not differ for student loan indicators or income groups.

H1:  Average monthly balance differs for student loan indicators or income groups.

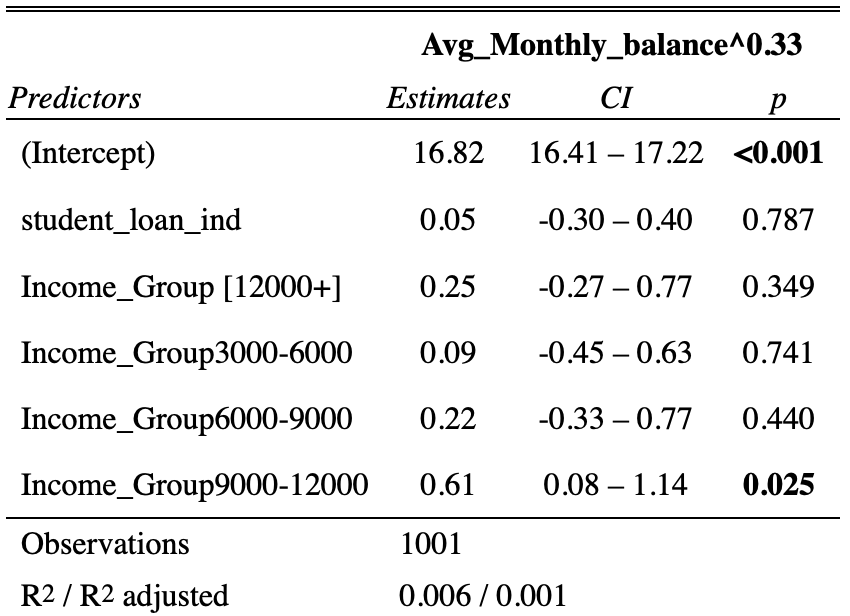


Boxplot indicates student loan does not make a significant difference on average monthly balance.



The plot suggests insignificant differences in the average monthly balance of income groups.

For modeling, the transformed response variable from research question 3 is used.



The model gives insignificant estimates for estimates except for the income group of 9000-1200. So, the income group of 9000-12000 has a higher average monthly balance compared to other groups.

# **Conclusion**

To begin with, out of 12 variables 6 has a significant relationship with average monthly spending. The best-related variable is travel spending, followed by grocery, apparel spending, balance, and pay rate. Using a credit card for travel spending does not seem to differ for various age groups. High annual percentage rates do not seem to lower the customers’ credit balance. Customers’ average monthly balance does not seem to significantly differ for whether they have a student loan or not. And only the 9000 to 12000 income group have a significantly higher balance on average.

# **References**

[1] Data retrieved from <https://leaps.analyttica.com/sample_cases/7>, last accessed 18.06.2021.

Lee, J., & Kwon, K. N. (2002). Consumers’ use of credit cards: store credit card usage as an alternative payment and financing medium. Journal of Consumer Affairs, 36(2), 239-262.

Sarofim, S., Chatterjee, P., & Rose, R. (2020). When store credit cards hurt retailers: The differential effect of paying credit card dues on consumers' purchasing behavior. Journal of Business Research, 107, 290-301.