

## 1. Predictor Table

Variable	Include for TransCount	Effect on Trans Count	Include for Trans Amount	Effect on Trans Amount	Rationale
Month	No	()	No	()	
ClientNum	No	()	NO	()	It acts as a primary key to identify unique customers and won't be helpful in my analysis
WealthTag	No	()	No	()	I have created another numeric factor grouping for ease of computation and would be using that
CardType	No	()	No	()	I have created another numeric factor grouping for ease of computation and would be using that
RevolvingIndicator	No	()	No	()	I have created another numeric factor grouping for ease of computation and would be using that
SpendCategory	No	()	No	()	I have created another numeric factor grouping for ease of computation and would be using that
TransAmount	Yes	+/-	Yes	()	Y variable in case for computing transamount, in computing the transaction count, it is
TransCount	Yes	()	Yes	+/-	Y variable in case for computing transCount,
SpendGroup , RevolvingGroup, WealthGroup, CardGroup	Yes	+/-	Yes	+/-	These variables are included as they represent different categorical factors that may influence transaction counts and amounts. Their inclusion allows for the analysis of how different customer segments, spending behaviors, card types, and revolving indicators affect transaction counts and amounts

## 2. Models

### A. Transaction Count

**Model1:**  $\text{lm}(\text{TransCount} \sim \text{TransAmount} + \text{SpendGroup} + \text{CardGroup} + \text{WealthGroup} + \text{RevolvingGroup}, \text{data} = \text{data})$

**Model2:**  $\text{lm}(\text{TransCount} \sim \log(\text{TransAmount}) + \text{SpendGroup} + \text{CardGroup} + \text{WealthGroup} + \text{RevolvingGroup}, \text{data} = \text{data})$

**Model 2 is a better fit and I will be using this model.**

### B. Transaction Amount

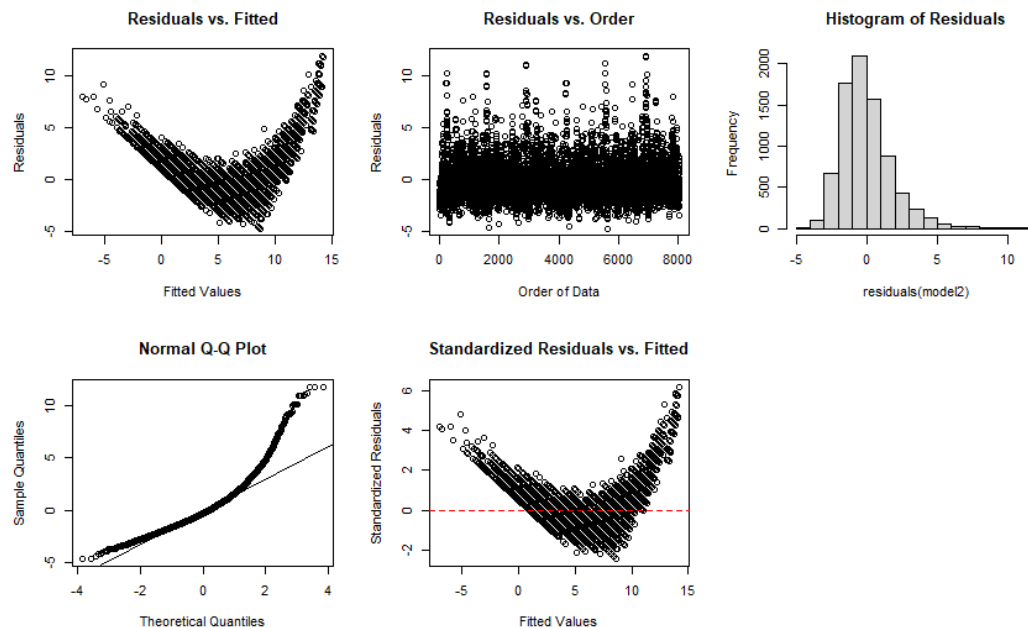
**Model3:**  $\text{lm}(\log(\text{TransAmount}) \sim \text{TransCount} + \text{SpendGroup} + \text{CardGroup} + \text{WealthGroup} + \text{RevolvingGroup}, \text{data} = \text{data})$

**Model4:**  $\text{lm}(\text{Log}(\text{TransAmount}) \sim \text{TransCount} + \text{SpendGroup} + \text{CardGroup} + \text{WealthGroup} + \text{RevolvingGroup}, \text{data} = \text{data})$

**Model 2&4 is a better fit and I will be using this model.**

## 2.Assumptions

## A. Transaction Count

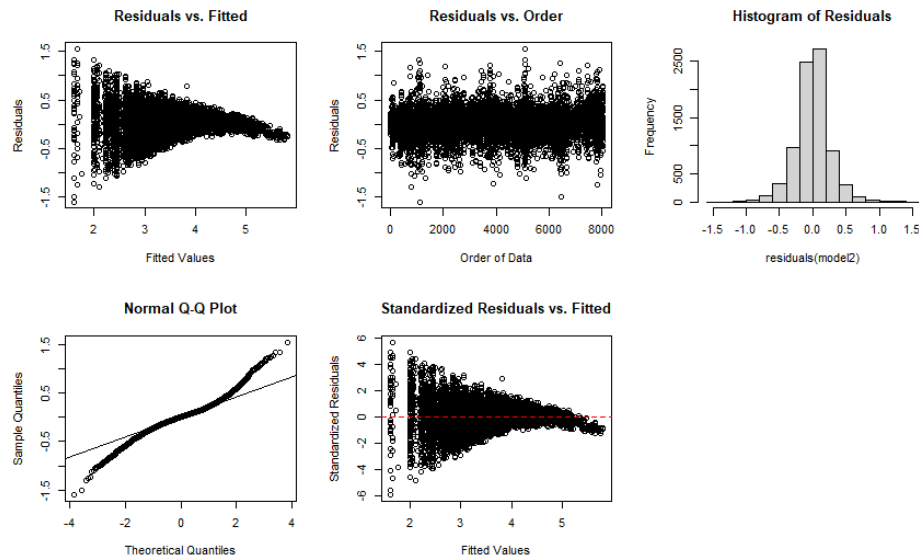


- **Kolmogorov-Smirnov Test for Normality (Model 2):**  $D = 0.16006$ ,  $p\text{-value} < 2.2e-16$   
alternative hypothesis: two-sided, we can **reject the null hypothesis of normality** for the residuals in Model 2. This suggests the residuals are not normally distributed, violating the normality assumption. **Normality is not met**
- **BP for Homogeneity of Variances (for model2):**  $BP = 1105.4$ ,  $df = 14$ ,  $p\text{-value} < 2.2e-16$ , indicating we can **reject the null hypothesis of homoscedasticity** (constant variance) of the residuals. This suggests the variance of the residuals is not constant across the fitted values, **violating the homoscedasticity assumption**.
- **"Variance Inflation Factors (VIF) for Model 2:"** Potential issue with RevolvingGroup.

$GVIF\ Df\ GVIF^{1/(2*Df)}$

<i>log(TransAmount)</i>	7.573891	1	2.752070
<i>SpendGroup</i>	2.331081	5	1.088318
<i>CardGroup</i>	2.293659	2	1.230643
<i>WealthGroup</i>	2.710461	3	1.180793
<i>RevolvingGroup</i>	3.233998	3	1.216065

## B.Transaction Amount



### 1. Normality:

- **The Kolmogorov-Smirnov test p-value ( $< 2.2e-16$ )** is highly significant, indicating we can reject the null hypothesis of normality for the residuals in Model 2. This suggests the residuals are not normally distributed, **violating the normality assumption**.

### 2. Homoscedasticity:

- **The Breusch-Pagan test p-value ( $< 2.2e-16$ )** is still highly significant, indicating we can reject the null hypothesis of homoscedasticity. However, the test statistic (BP) is 2155.6, which is lower than the previously reported values (1105.4 and 1455.1). This suggests a potential improvement compared to the previous outputs, but **the violation might still be present**.

### 3. Multicollinearity:

- **VIF values are all below 3**, which is generally considered a low risk of multicollinearity. This is an improvement compared to the previous outputs where RevolvingGroup had a VIF closer to 5.

### 3. Interpretations

- **How do customers' spending pattern vary by spending category, while controlling for other variables?**  
Customers typically have slightly fewer transactions in the auto and business categories (around 0.16-0.18 fewer transactions on average). Customers have significantly fewer transactions in the entertainment category (around 0.48 fewer transactions on average). There's a minimal decrease in the number of transactions in the grocery category, but this difference is not statistically significant. Customers have significantly fewer transactions in the personal care category (around 0.47 fewer transactions on average).
- **What type of customers have the highest and lowest spending by card type? By what amount?**

**CardGroup1(Platinum) but based on the interpretations of other groups:**

- **CardGroup2(Gold)** has a negative coefficient in both models **2.016 units less** indicating lower spending compared to CardGroup1.
- **CardGroup3** has a **significantly lower coefficient compared to CardGroup1(Blue)**, suggesting the lowest spending among the three card groups. **They spent 4.04 units less than Platinum.**

- **How do their spending patterns vary by wealth tag?**

Emerging Affluent & Affluent Mass Market customers tend to have lower spending compared to the High net income (baseline) as indicated by the negative coefficients. Mass Market customers have the lowest spending among the listed wealth groups.

- **How do their spending patterns vary by revolving indicators?**

Occasional Revolvers: Compared to delinquent customers, they tend to have approximately 0.5873958 more transactions. Revolvers: They also have approximately 0.469635 more transactions than delinquent customers. Transactors: They have the highest number of transactions, with approximately 0.6926591 more transactions than delinquent customers.