**SDM A3 Credit Card Transactions**

1. **PREDICTOR TABLE: Y= Transaction amount**

|  |  |  |
| --- | --- | --- |
| **Predictor** | **Effect** | **Rationale** |
| WealthTag | + | People with higher incomes or greater wealth may be able to afford more expensive goods and services, which could lead to larger transaction amounts |
| CardType | + | Cards with higher-tier designations such as gold or platinum typically come with higher credit limits which enable users to make larger purchases or complete transactions with a higher value |
| RevolvingIndicator | +/- | They may not affect directly but being a transactor, the credit limit may increase and possible to do more transactions and a delinquent may not have more much limit based on their past credit history |
| SpendCategory | +/- | Higher spends need higher transactions and lower need lower transactions |
| *Excluded Variables* | | |
| Month | 0 | This attribute does not fit in this data |
| ClientNumber | 0 | This attribute does not affect transaction amount or transaction count |

**PREDICTOR TABLE: Y= Transaction count**

|  |  |  |
| --- | --- | --- |
| **Predictor** | **Effect** | **Rationale** |
| WealthTag | + | People with higher incomes or greater wealth may be able to do more purchases |
| CardType | + | Different cards are used for more number of transactions like people with gold and platinum are tend to do more number of transactions |
| RevolvingIndicator | +/- | They may not affect directly but being a transactor, the credit limit may increase and possible to do more number of transactions and a delinquent may not have more much limit based on their past credit history |
| SpendCategory | +/- | Some spending category needs high number of transactions like reoccurring transactions and some need single transactions |
| *Excluded Variables* | | |
| Month | 0 | This attribute does not fit in this data |
| ClientNumber | 0 | This attribute does not affect transaction amount or transaction count |

* Here our dependent variables are “transaction amount and transaction count” and independent variables are “Wealthtag”, “cardtype”, “Revolvingindicator”, “spendcategory”.
* Wealthtag shows the income level and the wealth level of an individual and having higher wealth leads to having gold or platinum type of card.
* This may show that they may be correlated with each other. Let us consider one of the both.
* The data have 4 null values in the “transactionamount” column and needed preprocessing.

The histogram plot of the dependent variables is as follows.

Chart, histogram

Description automatically generated Chart, histogram

Description automatically generated

Both of them are not normally disributed and hence need to be transformed. Let us use log transformation.

Chart, histogram

Description automatically generated Chart, histogram

Description automatically generated

These curves looks better than the last curves with a bell shape distribution.

Regression Models.

sc\_model = lm(log(transamount)~wealthtag+cardtype+revolvingindicator+spendcategory, data = df\_clean)

sc\_model\_2 = lm(log(transamount)~spendcategory, data = df\_clean)

sc\_model\_3 = lm(log(transamount)~wealthtag+spendcategory+revolvingindicator, data = df\_clean)

Text

Description automatically generated

The model looks stable as the coefficient values does not change significantly with change in variables.

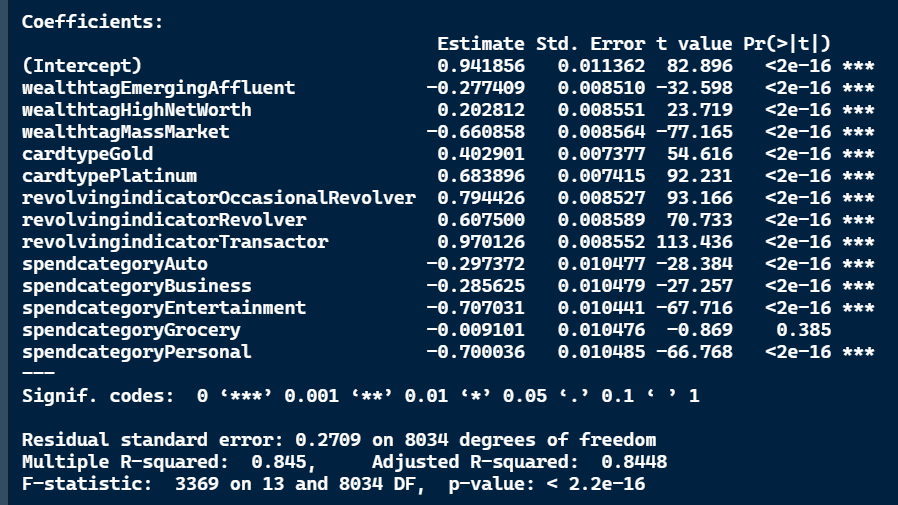
Interpretation:

1. Based on the coefficiant values, Airlines have higher affect on the transaction amount and next is grocery transactions. Airlines spend change the transaction amount by (e^ 3.9). similarly for grocery items with (e^3.9-0.03) from the exclusive spendcategory model
2. In model 1, it is shown that the base category of cardtype is “Blue” and the **platinum** card has more change with value (e^0.800) and the “Blue” has the minimum affect on the transactionamount.
3. People with “highnetworth” have more transaction amount when compared to the “Affluent” with positive affect. Wealth tag with “massmarket” has the lowest value compared to the base and this shows that “massmarket” category people are accountable to low “transactionamount” with value (e^0.780) compared to the affluent.
4. The transactor is accountable for higher transaction amount and the delliquent for the lowest transactionamount. If the deliquent is accountable for “e^x” change in transaction amount, then transactor is accountable for “e^(x+1.140)” change which is the highest in this category

Regression models for transaction count:

With inclusion of all the attributes, the model is :

sc\_model\_count = lm(log(transcount)~wealthtag+cardtype+revolvingindicator+spendcategory, data = df\_clean)



The intrepretation similar to the transaction amount.

1. Mass market have made low number of transactions and highnetworth with high transactions
2. Platinum card holders made more number of transactions and blue the lowest
3. Transactor have made more number of transactions and deliquent being the lowest, occasional revolver accounts for second highest number of transactions
4. Airlines transactions are made in higher count and entertainment and personal being the lowest. Groceries account for transaction count almost equal to airlines transactions

Let us see how the model fits.

1. Transaction amount:

A picture containing timeline

Description automatically generated Chart

Description automatically generated with low confidence

1. Transaction count

Chart, scatter chart

Description automatically generated Chart, line chart

Description automatically generated