Analysis of factors that affect the Account Receivable of New York Stock Market

Student Name:
Yuxiang Chen
Yuxin Liu
Wenshu Liao
Jinhui Li

Student ID:
301368904
301385527
301326083
301307665

Instructor Name: Derek Bingham

Course: Stat 350 (2020 Fall)

Catalogue

- **Abstract** **3**
- **Introduction** **3**
  - **Dataset** **4**
  - **Response variables and Explanatory variables** **4**
  - **Analysis aims to** **5**
- **Method and Models** **6**
- **Results** **11**
    - **Scatterplot Analysis** **11**
    - **Analysis on full model** **11**
    - **Model Selection** **12**
    - **Analysis on final model** **16**
- **Conclusion and Decisions** **18**
- **Reference** **23**
- **Appendix** **23**

Abstract

In this paper, we will be investigating the evolution of the New York stock market over the years. For the dataset in the study, 1206 observational data were collected with 79 variables, which were the response variables Account Receivable. Among them, the 20 explanatory variables that we selected from all the other variables and wanted to study. By studying the pairs plot and the full linear regression model analysis, we used the AIC model selection steps, decision trees, random forest and stepwise regression to improve the fitted model. Then comparing those new models with ANOVA tests. The final model was determined and showed that the Account Receivable had a relationship with by fourteen explanatory variables, the net income, the adjusted net income, net receivable, the cost of revenue, account payables, operating income, add income and relative net cash flow variables.

The progress on technology has increased our study about the stock market industry. Various statistical models and algorithms have been applied to study the relationship between account receivable and predictor variable. We also focus on independent variables to predict the trend of future market industry. In this paper, we will be determining specific factors which have most impacts on the New York stock market.

Key Words

Accounts payable, accounts receivable, cost of revenue, net income, net cash flow, operating income, add'l income expense items, profit margin.

Introduction

Account receivable is an important element to measure if a company is run well. Since there are many factors which are related to account receivable. Consequently, we decided to use the linear regression model to identify what factors had significant effects on the account receivable. The reason why only chose the fundamental.csv is that the security dataset does not seem to have a strong relationship with the response variable. These categorical variables seem not to have significant effects on the response variables. And the data in two price datasets are too large and always constructs a bad model.

Dataset

The data set, fundamentals.csv, was metrics extracted from annual SEC 10K fillings, (2012-2016) which was uploaded by Dominik Gawlik. (Gawlik, 2017) In the description of the dataset, nine of the 79 variables were considered significantly important and should be contained in the regression models.

Response variables

1.Account.receivable: Account receivable of a company (integer)

Explanatory variables (The first five variables were considered as important explanatory variables and always remain in the model)

1. Net.Receivable: Net receivable of a company (integer)
2. Net.Income.Adjustments: Adjusted net income of a company (integer)
3. Net.Cash.Flow.Operating: Net cash flow of operating of a company (integer)
4. Operating.Income: Operating income of a company (integer)
5. Accounts.Payable: Account payable of a company (integer)
6. Cost.of.Revenue: Cost of revenue of a company (integer)
7. Net.Cash.Flows.Financing: Net cash flows of financing of a company (integer)

8. Net.Income.Cont..Operations: Net income continuous of operation of a company (integer)

9. Gross.Margin: Gross margin of a company (integer)

10. Net.Borrowings: Net borrowings of a company (integer)

11. Net.Income: Net income of a company (integer)

12. Add.l.income.expense.items: Add’l income divides expense items of a company (integer)

13. Cash.Ratio: Cash ratio of a company (integer)

14. Current.Ratio: Current ratio of a company (integer)

15. Operating.Margin: Operating margin of a company (integer)

16. Profit.Margin: Profit margin of a company (integer)

17. Current.Margin: Current margin of a company (integer)

18. Ticker.Symbol: Ticker symbol of a company (integer: 1 = AAL, 2 = AAP, 3 = AAPL, etc.)

19. Quick.Ratio: Quick ratio of a company (integer)

20. Effect.of.Exchange.Rate: Effect of exchange rate of a company (integer)

Analysis aims to

1. To construct an appropriate linear regression model and determine a final model.
2. To check the effects of the final model.
3. To interpret the effects of the explanatory variables (predictors) in the final regression model and test the significance of those explanatory variables in the final model.

Methods and Models

In this paper, the multiple linear regression will be applied to analyze the stock data. In general, we consider the response variable has a multiple linear relationship between those explanatory variables and interactions. They basically follow that:

$Y = X\beta + \varepsilon$.

First of all, we would set a full linear regression model which contains all 20 explanatory models which are most important to the response variable we select from the datafile:

model1.linear: Accounts.Receivable = $\beta_0 + \beta_1$ * Accounts.Payable + β_2 * Cash.Ratio + β_3 * Current.Ratio + β_4 * Net.Borrowings + β_5 * Add.l.income.expense.items + β_6 * Net.Cash.Flow.Operating + β_7 * Net.Cash.Flows.Financing + β_8 * Cost.of.Revenue + β_9 * Net.Income + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Income + β_{12} * Net.Income.Cont..Operations + β_{13} * Net.Receivable + β_{14} * Profit.Margin + β_{15} * Quick.Ratio + β_{16} * Gross.Margin + β_{17} * Effect.of.Exchange.Rate + β_{19} * Ticker.Symbol + β_{20} *Operating.Margin

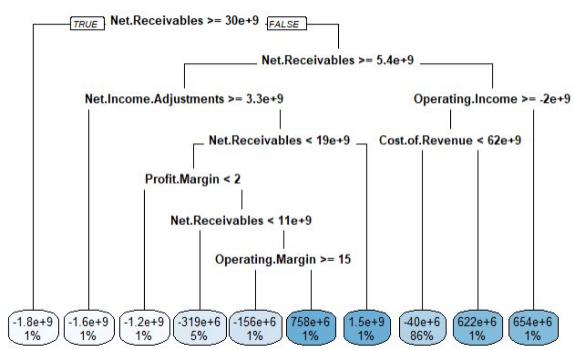
Where the Ticker.Symbol is a categorical variable

The reason why we chose those 20 explanatory variables is that after reading all variables in fundamental dataset, those variables are computing results of other variables in that dataset, choosing them can avoid some collinearity. Cash Flow, income and so on are most relevant variables to the account receivable. And other explanatory variables like ratio and margin, they could stand for the status of those information and data best.

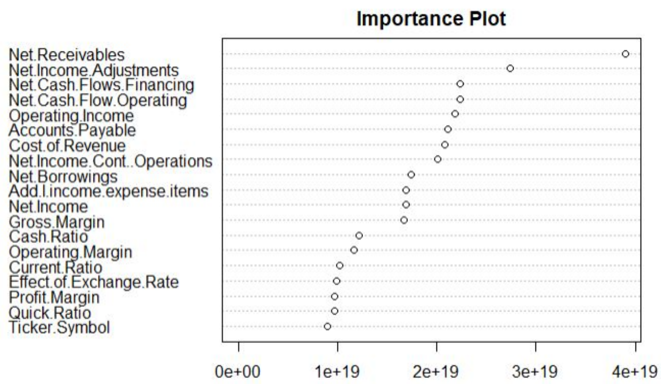
To test the effectiveness of the full model, the residual plots, normal Q-Q plots and the Residual vs Leverage plot will be applied to see whether the hypothesis of the linear regression model will be rejected, and to see if any variables need further transformation. According to the summary of the full model, all 20 explanatory variables are involved in the model, but some of them do not fit well into the model and have no significant relationship between the Account Receivable.

Therefore, we need to reduce some explanatory variables and add some new interactions so that we can construct a better linear model. AIC select method, decision tree and random forest are applied. And the first five explanatories should always be kept in all models we use.

We used the decision tree and the importance plot of random forest to decide and tell which explanatory variables have the strongest relationship with Account receivable. So according to these two figures, five of the most important variables among explanatory variables were found.



Plot of decision tree of the full model



Importance Plot of random forest for all explanatory variables of the full model

So according to the importance plot, the categorical variable, Ticker.Symbol is with the least importance, and after applying this explanatory variable into the model, it will influence the model and the plot of decision tree greatly, both model summary and decision tree will be hard to read and with no value to do further research.

The next step is using stepwise regression in both directions to reduce some explanatory variables to improve the model, which use Akaike's Information Criterion (AIC) to tell which variables should remain in the model. Multiple R-squared was applied to check if the removal of some variables makes the model better. Then some most irrelevant variables were deleted from the original model.

Here is one of the stepwise models which we created:

model2.step: Accounts.Receivable = $\beta_0 + \beta_1$ * Accounts.Payable + β_2 * Net.Borrowings + β_3 * Add.l.income.expense.items + β_4 * Net.Cash.Flow.Operating + β_5 * Net.Cash.Flows.Financing + β_6 * Cost.of.Revenue + β_7 * Net.Income.Adjustments + β_8 * Operating.Income + β_9 * Net.Income.Cont..Operations + β_{10} * Net.Receivables β_{11} * Gross.Margin + β_{12} * Effect.of.Exchange.Rate

At the same time, the T-test was applied to make a further decision. We used P-value to select reject variables which were not related to the response variable. The new model was be created:

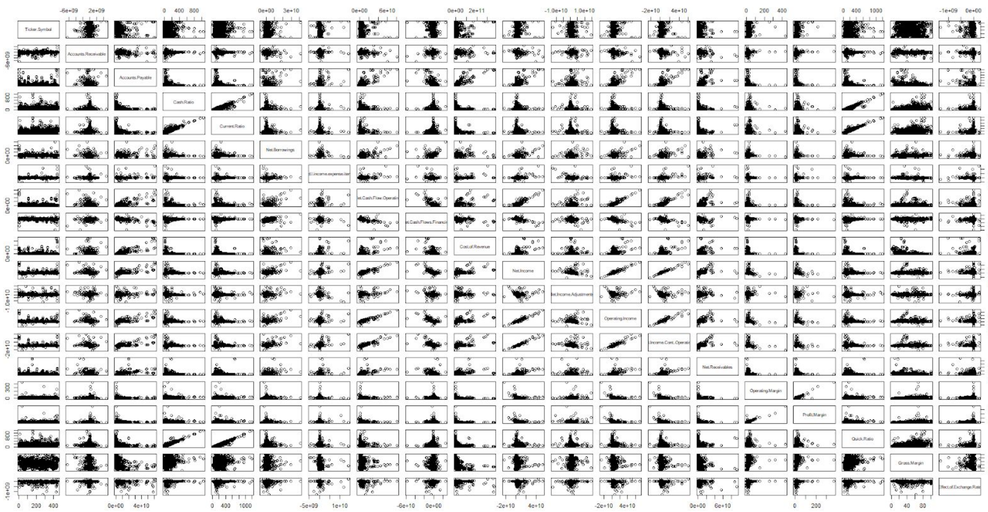
model1.linear3: Accounts.Receivable = $\beta_0 + \beta_1$ * Accounts.Payable + β_2 * Net.Cash.Flow.Operating:Net.Cash.Flows.Financing + β_3 * Net.Income : Operating.Income + β_4 * Operation.Margin * Profit.Margin+ β_5 * Add.l.income.expense.items + β_6 * Net.Cash.Flow.Operating + β_7 * Net.Cash.Flows.Financing + β_8 * Cost.of.Revenue + β_9 * Net.Income + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Income + β_{12} * Net.Income.Adjustments:Net.Income.Cont..Operations + β_{13} * Net.Receivables + β_{14} * Net.Income:Net.Income.Adjustments + β_{15} * Net.Income:Net.Income.Cont..Operations + β_{16} * Gross.Margin + β_{17} * Effect.of.Exchange.Rate

We repeated these steps many times to find better models. Also, ANOVA test was used to check if the newest model is the better one. We used ANOVA test to compare model1.linear3 and model.final. As a result, model.final was chosen as the best model. This is the final model we get:

model.final: Accounts.Receivable = β_1 * Accounts.Payable + β_2 * Add.l.income.expense.items + β_3 * Net.Cash.Flow.Operating + β_4 * Net.Cash.Flow.Operating:Net.Cash.Flows.Financing + β_5 * Net.Cash.Flows.Financing + β_6 * Cost.of.Revenue + β_7 * Net.Income + β_8 * Net.Income:Net.Income.Cont..Operations + β_9 * Net.Income.Adjustments:Net.Income.Cont..Operations + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Margin:Profit.Margin + β_{12} * Operating.Income + β_{13} * Net.Income : Operating.Income + β_{14} * Net.Receivables

Result

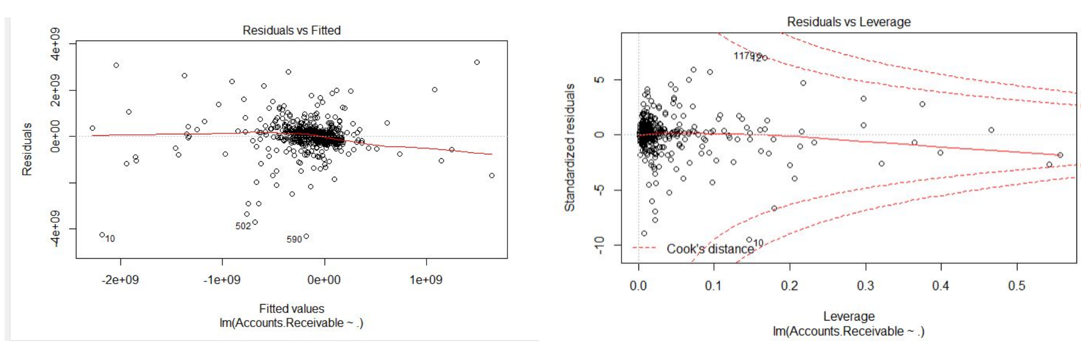
Scatter plot matrix



Paris plot of all explanatory variables of the full model

The first plot is the pairs plot of the full model. By the plot, there are still some explanatory variables that have collinearity with each other. Also, the plot matrix shows that Net.Receivable, Operating.Income and Net.Income.Adjustments might have linear relationship with the response variable. And no influential outliers are observed, so the data does not need further transformation.

Analysis on the full model



Residual vs Fitted Plot and Residual vs Leverage Plot for the full model

Estimated σ	507612249		Significant Predictors	Coefficient	P-value
R-square	0.1606		Net.Borrowing	-0.003461	0.0001
F-statistics	13.37		Net.Receivable	0.009167	1.03e-7
F-stat P-value	<2.2e-16		Net.Income.Cont..Operations	-0.002576	1.12e-12
			Gross.Margin	-2.656e6	0.000405

Table for Summary of Significant Predictors

According to the Multiple R-square, the full model could only explain 16% of the variability of the data. But among all those explanatory variables, four of them are significant with p-value<5%. They are Net.Borrowing, Net.Receivable, Net.Income.Cont..Operationd and Gross.Margin. Therefore, the model did follow the first assumption of the linear regression model.

Model Selection

It is necessary to reduce some explanatory variables and add some new interactions, so that a better linear model can be constructed. The AIC select method, decision tree and random forest were applied. And the first five explanatories should always be kept in all models we use.

Although the pairs plot of the dataset and Normal Q-Q of the full model shows that the distribution of both explanatory variables and response variables are not very good for linear regression, but log transformation is not applied because the many variables in the model contains zero and negative value, which makes it impossible to do the log transformation. Even though an inverse transformation can be applied, all of the test result shows that it is not even as good as the original linear regression. As a result, no transformation could be applied in such a situation.

We used the decision tree and the importance plot of random forest to decide and tell which explanatory variables have the strongest relationship with Account receivable. So according to these two figures, Cost of Revenues, Operating income, Net Receivables, Operating income and account payable are five of the most important variables among explanatory variables.

At first, the variable Ticker.Symbol was considered as a categorical variable that might have a strong relation between Account Receivable. However, it is of the least important and it also disturbs the result of the decision tree greatly. Thus, Ticker.Symbol was decided to be removed.

Therefore, the next step is using stepwise regression to reduce some explanatory variables to improve the model, which use Akaike's Information Criterion (AIC) in both directions to tell which variables should remain in the model.

As a result, this is the new model we get:

model2.step: Accounts.Receivable = β_0 + β_1 * Accounts.Payable + β_2 * Net.Borrowings + β_3 * Add.l.income.expense.items + β_4 * Net.Cash.Flow.Operating + β_5 * Net.Cash.Flows.Financing + β_6 * Cost.of.Revenue + β_7 * Net.Income.Adjustments + β_8 * Operating.Income + β_9 * Net.Income.Cont..Operations + β_{10} * Net.Receivables + β_{11} * Gross.Margin + β_{12} * Effect.of.Exchange.Rate

But even after we used stepwise regression and reduced some variables, the Multiple R-squared is still very low, which is about 0.2215. There is a reason to doubt that collinearity has a strong impact on the model. On the other hand, it is hard to predict real-life factors, so predicting these models will tend to have R-squared values less than the R-squared values of models in theoretical study. For these reasons, we added more interactions into the model, removed some most irrelevant variables and came up with a new linear model.

model1.linear2: Accounts.Receivable = β_0 + β_1 * Accounts.Payable + β_2 * Cash.Ratio:Current.Ration + β_3 * Current.Ratio:Quick.Ratio + β_4 * Net.Borrowings + β_5 * Add.l.income.expense.items + β_6 * Net.Cash.Flow.Operating + β_7 * Net.Cash.Flows.Financing + β_8 * Cost.of.Revenue + β_9 * Net.Income + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Income + β_{12} * Net.Income.Cont..Operations + β_{13} * Net.Receivables + β_{14} * Profit.Margin + β_{15} * Quick.Ratio:Cash.Ratio + β_{16} * Gross.Margin + β_{17} * Effect.of.Exchange.Rate + β_{18} * Net.Cash.Flow.Operating:Net.Cash.Flows.Financing + β_{19} * Net.Income:Net.Income.Cont..Operations + β_{20} * Netncome:Net.Income.Adjustments + β_{21} * Net.Income.Adjustments:Net.Income.Cont..Operations + β_{22} * Net.Income : Operating.Income + β_{23} * Operating.Margin:Profit.Margin

And again, using stepwise regression to reduce some explanatory variables and interactions. Then we got a new model.

model1.step2: Accounts.Receivable = β_0 + β_1 * Accounts.Payable + β_2 * Cash.Ratio:Current.Ration + β_3 * Current.Ratio:Quick.Ratio + β_4 * Net.Borrowings + β_5 * Add.l.income.expense.items + β_6 * Net.Cash.Flow.Operating + β_7 * Net.Cash.Flows.Financing + β_8 * Cost.of.Revenue + β_9 * Net.Income + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Income + β_{12} * Net.Income.Cont..Operations + β_{13} * Net.Receivables + β_{14} * Profit.Margin + β_{15} * Quick.Ratio:Cash.Ratio + β_{16} * Gross.Margin + β_{17} * Effect.of.Exchange.Rate + β_{18} * Net.Cash.Flow.Operating:Net.Cash.Flows.Financing + β_{19} * Net.Income:Net.Income.Cont..Operations + β_{20} * Netncome:Net.Income.Adjustments + β_{21} * Net.Income.Adjustments:Net.Income.Cont..Operations + β_{22} * Net.Income : Operating.Income

The result is not very good, because only one explanatory variable is removed. But by the summary and AIC of this new model, some variables could be deleted which have large p-value and make the AIC large. Multiple R-squared also be used to tell if the removal of some variables makes the model better. And the following result is this model.

model1.linear3: Accounts.Receivable = β_0 + β_1 * Accounts.Payable + β_2 * Net.Cash.Flow.Operating:Net.Cash.Flows.Financing + β_3 * Net.Income : Operating.Income + β_4 * Operation.Margin * Profit.Margin+ β_5 * Add.l.income.expense.items + β_6 * Net.Cash.Flow.Operating + β_7 * Net.Cash.Flows.Financing + β_8 * Cost.of.Revenue + β_9 * Net.Income + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Income + β_{12} *

Net.Income.Adjustments:Net.Income.Cont..Operations + β_{13} * Net.Receivables + β_{14} * Net.Income:Net.Income.Adjustments + β_{15} * Net.Income:Net.Income.Cont..Operations + β_{16} * Gross.Margin + β_{17} * Effect.of.Exchange.Rate

After repeating stepwise regression and comparing AIC and multiple R-squared of the model, the final model is set and determined.

model.final: Accounts.Receivable = β_1 * Accounts.Payable + β_2 * Add.l.income.expense.items + β_3 * Net.Cash.Flow.Operating + β_4 * Net.Cash.Flow.Operating:Net.Cash.Flows.Financing + β_5 * Net.Cash.Flows.Financing + β_6 * Cost.of.Revenue + β_7 * Net.Income + β_8 * Net.Income:Net.Income.Cont..Operations + β_9 * Net.Income.Adjustments:Net.Income.Cont..Operations + β_{10} * Net.Income.Adjustments + β_{11} * Operating.Margin:Profit.Margin + β_{12} * Operating.Income + β_{13} * Net.Income : Operating.Income + β_{14} * Net.Receivables

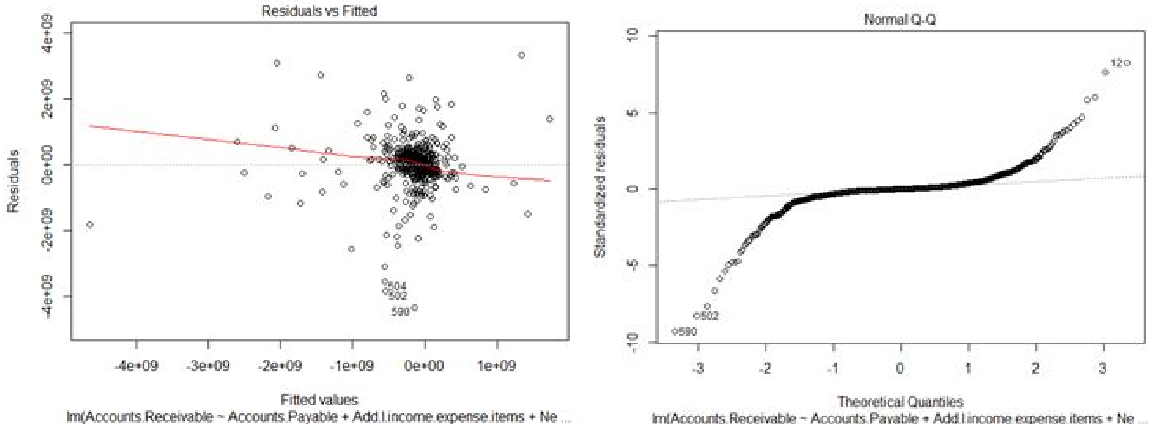
In model.final, Operating.Margin:Profit.Margin and Net.Income : Operating were added. Net.Income : Operating.Income, Operation.Margin * Profit.Margin , Net.income:Net.Income.Adjustments, Gross.Margin and Effect.of.Exchange.Rate were rejected from model1.linear3. ANOVA test was used to compared model1.linear3 and model.final. The following table is ANOVA test table:

Model	Res.DF	RSS	DF	Sum of Sq	F	Pr (> F)
model.final	1193	3.1396e+20				
model1.linear3	1184	3.0869e+20	7	5.2711e+18	2.8882	0.005381

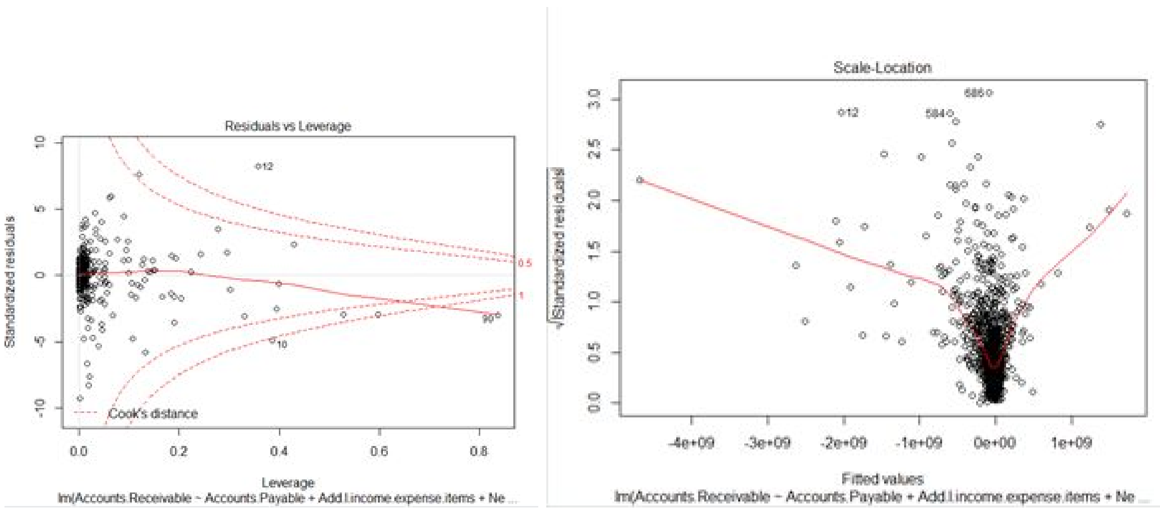
Degrees of Freedom refers to the maximum number of logically independent values, which are values that have the freedom to vary, in the data sample. Through the null hypothesis of the F test in the ANOVA table, the model model.final has nine more interaction terms significant better than the model model1.linear3. Meanwhile, the p-value was less than 1%, which means we had significant evidence to reject the null hypothesis. Therefore, the model model.final was significantly better than the model model1.linear3.

Final model analysis

The final model determined model.final. Although the multiple R-square that we retrieved from the model can only explain 29.57% variability, that was the largest multiple R-square that we could retrieve from so many combinations of variables. So it is the best option among all of the "bad" models. What's more, most terms in this model has a p-value that were much smaller than 0.05, we still concluded that the model.final is quite an appropriate model. In addition, the following three plots generated for model.final showed that the model satisfied the conditions of constructing the linear regression well. In conclusion, model.final was a good selected final model.



Risidual vs Fitted plot and Normal Q-Q plot for final model



Residual vs Leverage plot and Scale-Location plot for final model

Conclusion and discussion

According to the final model, we can conclude that the Account Receivable is linear related to following explanatory variables:

1. Accounts.Payable
2. Add.l.income.expense.items
3. Net.Cash.Flow.Operating
4. Net.Cash.Flows.Financing
5. Cost.of.Revenue
6. Net.Income
7. Net.Income.Adjustments
8. Operating.Income
9. Net.Receivables
10. The interaction of Net Cash Flow Operating and Net Cash Flow Financing (Net.Cash.Flow.Operating:Net.Cash.Flows.Financing)
11. The interaction of Net Income and Net Income Cont Operations (Net.Income:Net.Income.Cont..Operations)
12. The interaction of Net Income Adjustments and Net Income Cont Operations (Net.Income.Adjustments:Net.Income.Cont..Operations)
13. The interaction of Net.Income and Operating.Income (Net.Income:Operating.Income)
14. The interaction of Operating Margin and Profit Margin (Operating.Margin:Profit.Margin)

The t-tests in the summary table shows that ten of the above fourteen explanatory variables are significant, which are Add.l.income.expense.items, Net.Cash.Flow.Operating, Net.Cash.Flows.Financing, Cost.of.Revenue, Net.Income.

And the interaction term Operating.Income, Net.Receivables,

Net.Cash.Flow.Operating:Net.Cash.Flows.Financing,Net.Income:Net.Income.Cont..Operations,

Net.Income:Operating.Income, Operating.Margin:Profit.Margin

For the three significant terms, their estimated coefficients are

```
Coefficients:
                                                Estimate Std. Error t value Pr(>|t|)
Accounts.Payable                              -1.543e-02  4.904e-03  -3.147 0.001690 **
Add.l.income.expense.items                    -1.013e-01  2.463e-02  -4.114 4.16e-05 ***
Net.Cash.Flow.Operating                        6.110e-02  8.574e-03   7.126 1.78e-12 ***
Net.Cash.Flows.Financing                      -5.630e-02  8.287e-03  -6.794 1.72e-11 ***
Cost.of.Revenue                                3.732e-03  8.676e-04   4.302 1.83e-05 ***
Net.Income                                     6.220e-02  1.854e-02   3.355 0.000819 ***
Net.Income.Adjustments                        -3.611e-02  1.573e-02  -2.296 0.021844 *
Operating.Income                              -1.235e-01  1.357e-02  -9.101  < 2e-16 ***
Net.Receivables                               -2.833e-02  3.137e-03  -9.030  < 2e-16 ***
Net.Cash.Flow.Operating:Net.Cash.Flows.Financing  3.077e-12  2.973e-13  10.351  < 2e-16 ***
Net.Income:Net.Income.Cont..Operations         3.888e-12  4.998e-13   7.780 1.57e-14 ***
Net.Income.Adjustments:Net.Income.Cont..Operations  2.990e-12  9.600e-13   3.115 0.001886 **
Net.Income:Operating.Income                   -2.240e-12  4.210e-13  -5.320 1.24e-07 ***
Operating.Margin:Profit.Margin                -1.206e+04  3.393e+03  -3.554 0.000394 ***
```

Based on the result above, we can get following conclusions:

1. The net receive has a negative relationship with the account receivable, so the growth of net receive will cause the decrease of the account receive amount.

2. To keep the trend of the account receivable increasing, it had better to keep the cost of revenue under 2 * 10^11, the account receive will basically increase by below 2 * 10^11. So it is very important to control the cost on revenues to keep the account receivable from decreasing.

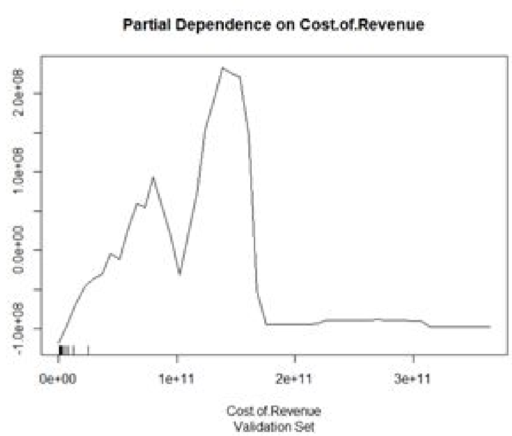
3. Net Income is one of the important explanatory variables that account receivable increases with it, because the net income contributes to part of the account receivable.

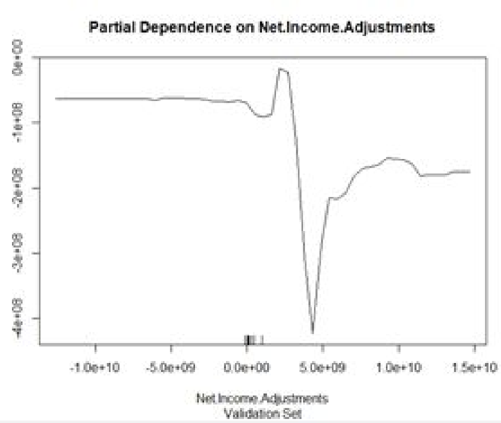
4. Account receivable is also affected by net cash flow, in general, net cash flow is in positive correlation with account receivable, both the interaction and difference of the coefficient of net cash flow for operating and net cash flow for financing is positive, so

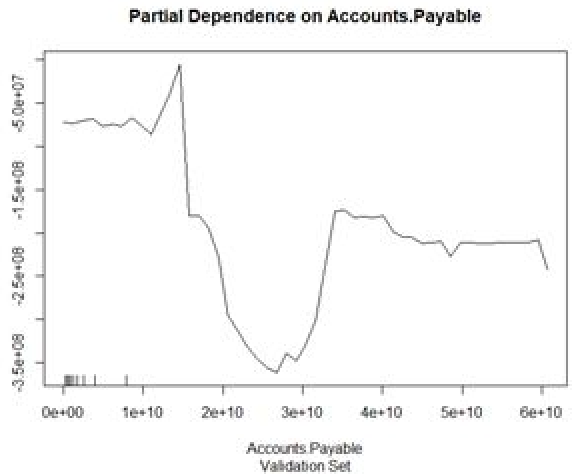
strengthen the investment on net cash flow for operating and for financing would help the growth the receive of the account.

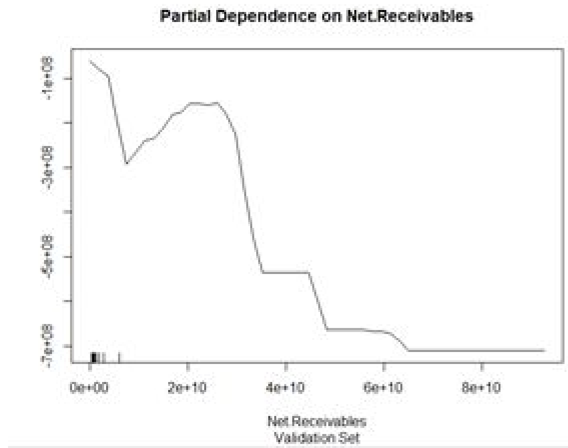
5. For some reason, when the net income adjustment reaches 5 * 10^9, the relationship with account receivable gets weak, that is an important phenomenon that needed further investigation.

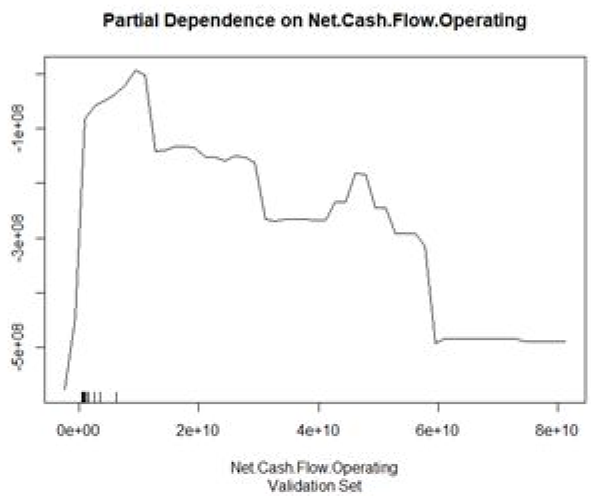
6. The interaction of operating margin and profit margin is the predictor with the largest absolute value of regression coefficient, while the coefficient is negative, so it is vital to keep this interaction of these two factors low so that the receivable of accounts would not be seriously influenced by it.











Partial Dependence Plot for top five Important Explanatory Variables in Final Model

Therefore, in order to increase the Account Receivables, we suggest that:

1.Limit the cost of revenue, under the line of 2 * 10 ^11 limit, the relationship keeps growing stronger, but falls weaker beyond the limit 2* 10^11, so to control and limit the cost of revenue will be very necessary for the increment of the Account Receivable.

2.Invest more on Net Cash Flow Operating and Net Cash Flow Financing, this will return more account receivable, so it will be a wise decision to increase the investment on Cash Flow Financing and Operating.

3.Try to increase the net income will be the most direct way to increase the account receivable.

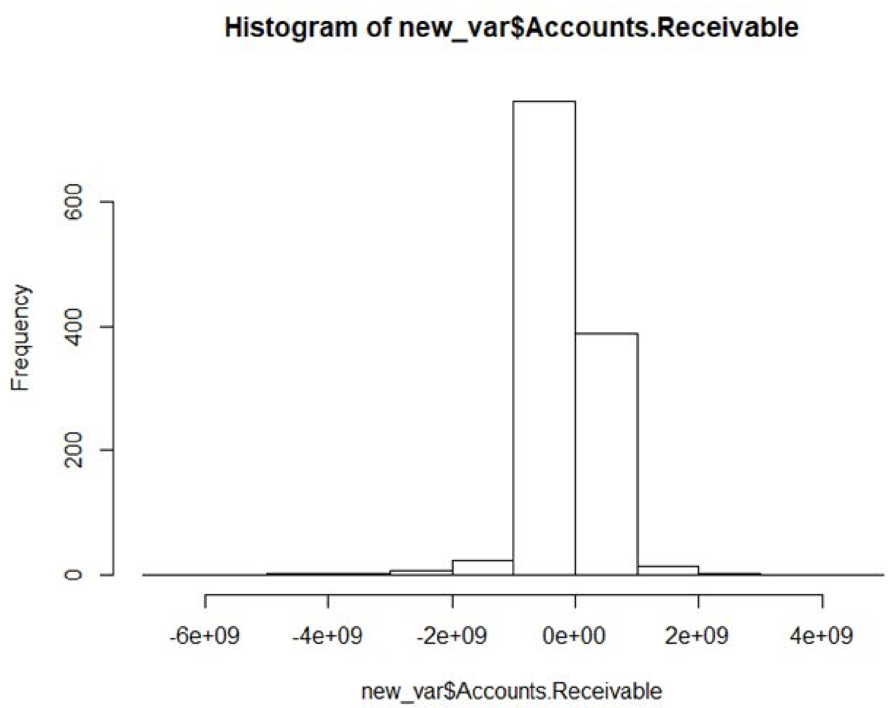
For the further investigation, we suggest to collect more data and add three more variable to study, the baby' s gender, and whether drink alcohol during the pregnancy. In addition, the study from the Boston Children's Hospital also observed that drinking alcohol also significant increases the risk of low birth weight babie s. Consequently, these two variables are highly recommended to be investigated in the future

Reference

1. Gawlik, D. (2017, February 22). New York Stock Exchange. Retrieved December 08, 2020, from https://www.kaggle.com/dgawlik/nyse

2. Montgomery, D. C., & Ryan, A. G. (2013). *Introduction to linear regression analysis, fifth edition*. Hoboken, NJ: Wiley.

Appendix



Histogram of Response Variable (Account.Receivable)

Summary of Full Model

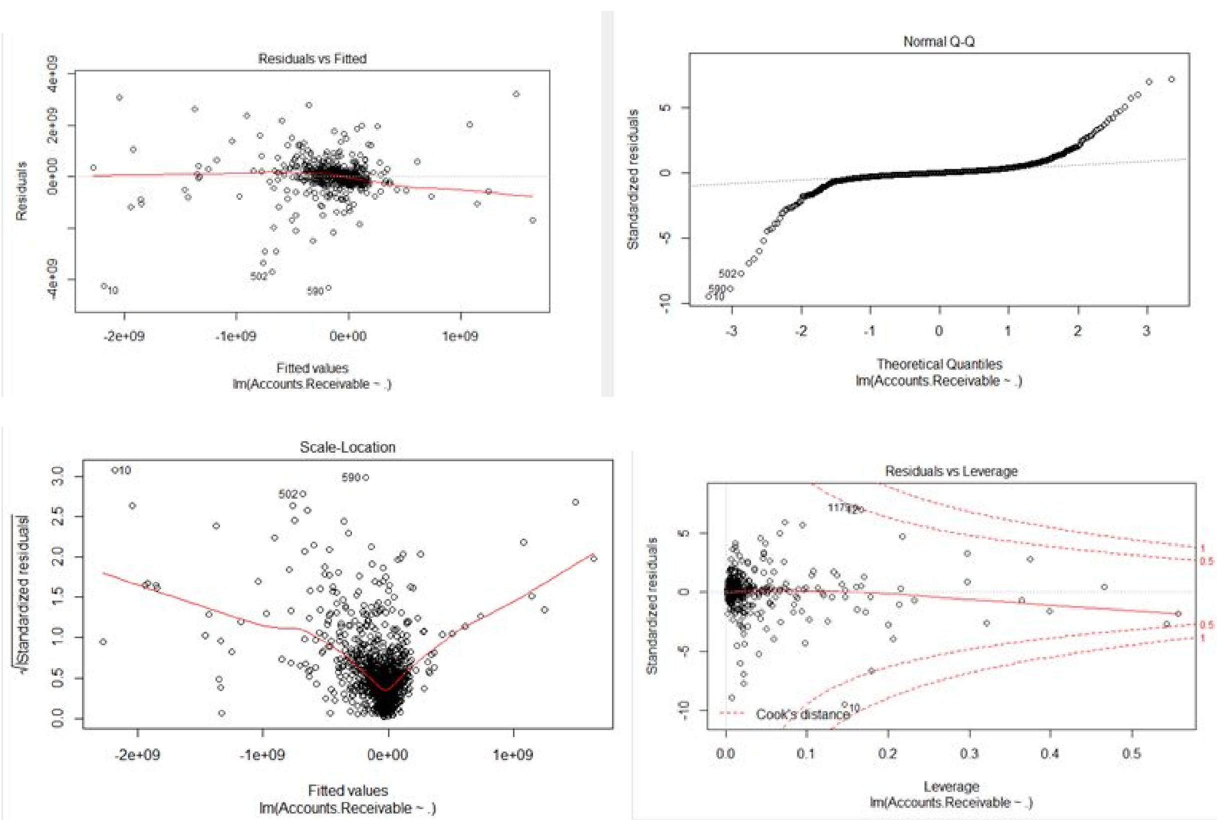
```
> summary(model1.linear)

Call:
lm(formula = Accounts.Receivable ~ Accounts.Payable + Cash.Ratio:Current.Ratio +
    Net.Borrowings + Net.Cash.Flow + Net.Cash.Flow.Operating +
    Net.Cash.Flows.Financing + Net.Cash.Flows.Investing + Net.Income +
    Net.Income.Adjustments + Net.Income.Applicable.to.Common.Shareholders +
    Net.Income.Cont..Operations + Net.Receivables + Profit.Margin +
    Quick.Ratio + Gross.Margin + Effect.of.Exchange.Rate + Operating.Margin +
    Profit.Margin, data = new_data)

Residuals:
       Min         1Q     Median         3Q        Max
-4.333e+09 -8.687e+07 -4.382e+06  7.308e+07  4.614e+09

Coefficients:
                                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                                  1.158e+08  4.620e+07   2.508 0.012291 *
Accounts.Payable                            -6.798e-03  3.806e-03  -1.786 0.074344 .
Net.Borrowings                              -3.461e-02  8.868e-03  -3.903 0.000100 ***
Net.Cash.Flow                                4.612e-02  5.617e-02   0.821 0.411722
Net.Cash.Flow.Operating                     -3.615e-02  5.579e-02  -0.648 0.517130
Net.Cash.Flows.Financing                    -2.602e-02  5.603e-02  -0.464 0.642472
Net.Cash.Flows.Investing                    -5.858e-02  5.620e-02  -1.042 0.297478
Net.Income                                   1.915e-01  3.456e-01   0.554 0.579572
Net.Income.Adjustments                      -2.225e-02  1.532e-02  -1.452 0.146751
Net.Income.Applicable.to.Common.Shareholders -2.961e-01  3.455e-01  -0.857 0.391730
Net.Income.Cont..Operations                  9.167e-02  1.712e-02   5.354 1.03e-07 ***
Net.Receivables                             -2.576e-02  3.581e-03  -7.193 1.12e-12 ***
Profit.Margin                                5.270e+06  2.154e+06   2.447 0.014568 *
Quick.Ratio                                  9.747e+04  2.622e+05   0.372 0.710141
Gross.Margin                                -2.656e+06  7.486e+05  -3.547 0.000405 ***
Effect.of.Exchange.Rate                     -3.976e-01  1.667e-01  -2.385 0.017226 *
Operating.Margin                            -2.597e+06  1.790e+06  -1.451 0.147015
Cash.Ratio:Current.Ratio                    -3.769e+02  3.871e+02  -0.974 0.330445
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 507600000 on 1188 degrees of freedom
Multiple R-squared:  0.1606,	Adjusted R-squared:  0.1486
F-statistic: 13.37 on 17 and 1188 DF,  p-value: < 2.2e-16
```



Summary Information and Plot of the Full Model

Model Selection Step

Information Criteria	Selection Direction	Selected Model	AIC / R-square For Selected Model
AIC	**Both**	**model1.linear**	**51791.44 0.1606**
AIC	**Forward**	**model2.step**	**51784.77 0.1582**
AIC	**Both**	**model1.linear2**	**51601.89 0.2862**
AIC	**Backward**	**model2.step2**	**51596.6 0.2847**
AIC and R-square	**Both**	**model1.linear3**	**51596.55 0.2916**
AIC and R-square		**model.final**	**51596.52 0.2957**

Table of AIC and R-square for models

Summary of Final Model

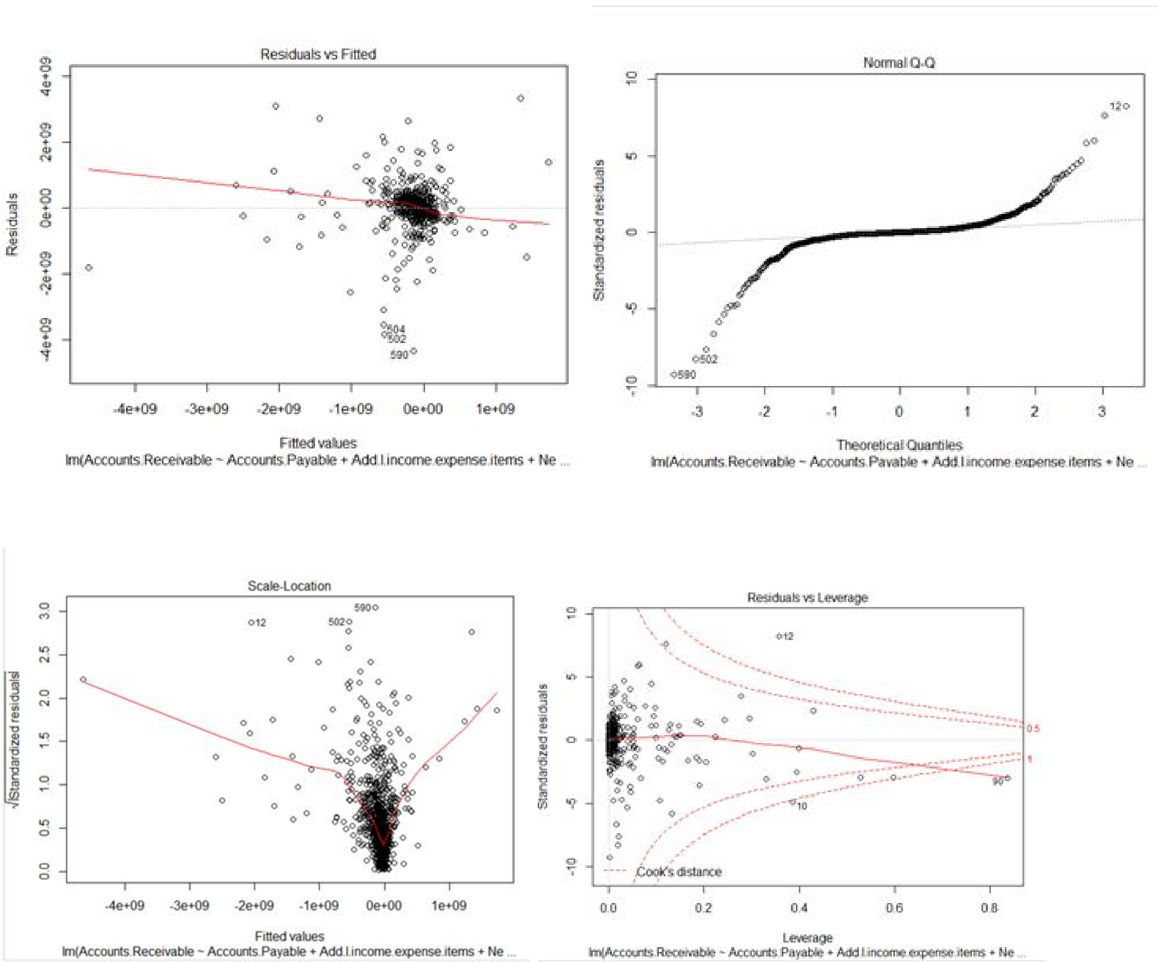
```
> summary(model.final)

Call:
lm(formula = Accounts.Receivable ~ Accounts.Payable + Add.l.income.expense.items +
    Net.Cash.Flow.Operating + Net.Cash.Flows.Financing + Cost.of.Revenue +
    Net.Income + Net.Income:Net.Income.Cont..Operations + Net.Income.Adjustments:Net.Income.Cont..Operations +
    Net.Income.Adjustments + Operating.Income + Net.Cash.Flow.Operating:Net.Cash.Flows.Financing +
    Net.Receivables + Net.Income:Operating.Income + Operating.Margin:Profit.Margin -
    1, data = new_var)

Residuals:
       Min         1Q     Median         3Q        Max
-4.351e+09 -6.156e+07  3.618e+06  8.873e+07  3.348e+09

Coefficients:
                                                   Estimate Std. Error t value Pr(>|t|)
Accounts.Payable                                 -1.543e-02  4.904e-03  -3.147 0.001690 **
Add.l.income.expense.items                       -1.013e-01  2.463e-02  -4.114 4.16e-05 ***
Net.Cash.Flow.Operating                           6.110e-02  8.574e-03   7.126 1.78e-12 ***
Net.Cash.Flows.Financing                         -5.630e-02  8.287e-03  -6.794 1.72e-11 ***
Cost.of.Revenue                                   3.732e-03  8.676e-04   4.302 1.83e-05 ***
Net.Income                                        6.220e-02  1.854e-02   3.355 0.000819 ***
Net.Income.Adjustments                           -3.611e-02  1.573e-02  -2.296 0.021844 *
Operating.Income                                 -1.235e-01  1.357e-02  -9.101  < 2e-16 ***
Net.Receivables                                  -2.833e-02  3.137e-03  -9.030  < 2e-16 ***
Net.Income:Net.Income.Cont..Operations            3.888e-12  4.998e-13   7.780 1.57e-14 ***
Net.Income.Cont..Operations:Net.Income.Adjustments  2.990e-12  9.600e-13   3.115 0.001886 **
Net.Cash.Flow.Operating:Net.Cash.Flows.Financing  3.077e-12  2.973e-13  10.351  < 2e-16 ***
Net.Income:Operating.Income                      -2.240e-12  4.210e-13  -5.320 1.24e-07 ***
Operating.Margin:Profit.Margin                   -1.206e+04  3.393e+03  -3.554 0.000394 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.69e+08 on 1192 degrees of freedom
Multiple R-squared:  0.2957,    Adjusted R-squared:  0.2874
F-statistic: 35.75 on 14 and 1192 DF,  p-value: < 2.2e-16
```



Summary and Plot for Final Model