VIETNAM NATIONAL UNIVERSITY HCMC

UNIVERSITY OF ECONOMICS AND LAW

MIDTERM PROJECT

PACKAGE FOR FINANCIAL APPLICATION 2

Course: Package for financial application 2

Lecturer: PhD. Nguyễn Thanh Liêm

Student ID: K214140961

Name: Phạm Duy Tùng

TP. Hồ Chí Minh, ngày 12 tháng 04 năm 2023

CONTENT TABLE

Determine assigned topic:

If your student ID ends with an even number => Trade credit (Payable) + Only firms listed on HOSE

If your student ID ends with an odd number => Trade credit (Receivable) + Only firms listed on HNX.

The assigned student ID number ends with 961, which is part of K214140961. Therefore, the chosen topic for the project is: Trade credit (Receivable) + Only firms listed on HNX.

CHAPTER 1: PERFORM LITERATURE REVIEW

After conducting a thorough review of multiple research articles, I have determined to select two variables that have an impact on trade credit (receivables). The continuous variable chosen is ROA, which stands for Return on Asset, and the discrete variable is firm size.

- ROA, or return on assets, is a financial ratio that measures a company's profitability by dividing its net income by its total assets. It indicates how efficiently a company is using its assets to generate profit. A higher ROA generally indicates better financial performance, while a lower ROA may suggest that the company is not using its assets efficiently. ROA can be used to compare a company's performance against its competitors or to track its own performance over time.

- Firm size refers to the measure of a company's magnitude or scale of operations, usually determined by factors such as the number of employees, revenue, assets, or market capitalization. A larger firm size typically suggests a company with greater resources, bargaining power, and access to diverse financing options. Conversely, a smaller firm size may indicate limited resources, less bargaining power, and reliance on trade credit as a means of financing.

1.1. Literature review

In the paper "Trade credit and SME profitability" by Martínez-Sola, García-Teruel, and Martínez-Solano, the authors investigate the impact of firm size and return on assets (ROA) on trade credit granted to small and medium-sized enterprises (SMEs). The authors find that larger SMEs are more likely to receive higher levels of trade credit, as they have greater bargaining power and more resources to manage their cash flows. In contrast, smaller SMEs rely more heavily on trade credit to finance their operations. Additionally, the authors find that higher levels of ROA are associated with lower levels of trade credit, suggesting that SMEs with better financial performance have less need for external financing.

Martínez-Sola, Cristina, Pedro J. García-Teruel, and Pedro Martínez-Solano. *"Trade credit and SME profitability."* Small Business Economics 42 (2014): 561-577. https://doi.org/10.1007/s11187-013-9491-y

Lawrenz and Oberndorfer explore how firm size affects the supply and demand of trade credit in their paper. They discover that bigger companies are more likely to request trade credit due to their more advanced position and availability of other financial options. Conversely, smaller companies are more likely to offer trade credit as a way of establishing customer relationships and competing with larger enterprises.

Lawrenz, Jochen, and Julia Oberndorfer. *"Firm size effects in trade credit supply and demand."* Journal of Banking & Finance 93 (2018): 1-20. https://doi.org/10.1016/j.jbankfin.2018.05.014

According to García-Teruel and Martínez-Solano's "Determinants of trade credit: A comparative study of European SMEs," SMEs in Europe with higher levels of return on assets (ROA) are less likely to rely on trade credit as a financing option. This is because firms with greater ROA have more internal resources to finance their operations and are therefore less dependent on external sources of financing.

García-Teruel, Pedro Juan, and Pedro Martinez-Solano. *"Determinants of trade credit: A comparative study of European SMEs."* International Small Business Journal 28.3 (2010): 215-233. https://doi.org/10.1177/0266242609360603

Phạm Xuân Quỳnh and Trần Đức Tuấn mentioned in their paper as the size of a company has an inverse impact on trade credit. When the scale of a business increases, there is a tendency for the proportion of outstanding payments to decrease. This result is similar to the research of Ahmed & ctg (2014), Tran Ai Ket (2014), but contrary to many previous studies such as Nadiri (1969), Petersen & Rajan (1997), Giannetti &ctg (2011). Large companies generally have better development, stronger financial resources, and a better reputation on the market, making it easier to access external capital. They also tend to manage their short-term credit more tightly, minimizing risk. This means that they are less likely to offer trade credit, as seen with companies like Vietnam Dairy Products JSC and Masan Group. As more large companies enter the market, the scope of trade credit within the Food industry will shrink.

Phạm Xuân Quỳnh, Trần Đức Tuấn, *"Các nhân tố ảnh hưởng đến tín dụng thương mại của các công ty ngành thực phẩm niêm yết trên HOSE"* Tạp chí Khoa học Kinh tế No.8 (2020) https://vjol.info.vn/index.php/due/article/view/48527

Ahmed, Jaleel, Hui Xiaofeng, and Jaweria Khalid in their study has found that firms with greater size enhance their sales without getting involved in credit transactions. It is found that the size of the non-financial firms is inversely related to trade credit demand. They argue that buyers may divert bank loans instead of purchasing inputs for their projects. Buyers may take inputs on credit and may complete their projects rather than having bank funding. Moreover, it is useful for the sellers to extend trade credit as they may use the accounts receivable as collateral for bank financing.

Ahmed, Jaleel, Hui Xiaofeng, and Jaweria Khalid. *"Determinants of trade credit: The case of a developing economy."* European Researcher 9-2 (2014): 1694-1706. https://doi.org/10.13187/er.2014.83.1694

According to Hill, M., Hill, K., Preve, L., & Sarria-Allende, there is an expectation for the influence of firm size on trade credit is the possibility that size may account for market power along with capital market access. Since greater negotiating ability typically accompanies higher market share, suppliers with higher market power may offer shorter trade credit periods or may not even extend trade credit to customers. Reduced trade credit for high market share suppliers may occur for at least two reasons. First, the level of competition from rivals is lower for suppliers with a larger market share, decreasing the likelihood of losing customers over a reduction in credit terms. Second, powerful suppliers are likely to have longer relationships with customers, implying higher switching costs for those customers. Switching costs include learning and transactions costs. But after evaluated through model, it appeared that the inverse relation between trade credit and firm size is inconsistent with the financial constraint view predicting that larger firms with capital market access provide more credit to customers. However, it was noted earlier that the scale of the firm may also capture suppliers' negotiating ability. These findings indicate substantial variation in the influence of firm size on the provision of trade credit.

M Hill, K Hill, L Preve, V Sarria-Allende *"International evidence on the determinants of trade credit provision."* Managerial Finance 45.4 (2019): 484-498. https://doi:10.1108/mf-07-2018-0295

Oh, and Hwang (2019) conducted a study on Korean manufacturing firms to investigate the association between trade credit and market power, while examining the moderating influence of financial constraint. Their analysis shows that return on assets (ROA) has a significant negative impact on the utilization of trade credit. According to the authors, companies with higher ROA have greater internal resources to fund their operations, hence their demand for trade credit is lower. In contrast, businesses with lower ROA may rely on trade credit more heavily due to limited internal resources and financing difficulties in obtaining external funding. These findings suggest that the link between ROA and trade credit usage is intricate and influenced by various firm-specific factors, such as market power and financial constraints.

Oh, S. Y., & Hwang, S. W. (2019). *The effect of market power on trade credit in Korean manufacturing firms*. Asia-Pacific Journal of Accounting & Economics, 26(3-4), 294-314.

In study titled "The Information Content of Trade Credit," by Aktas et al. (2012) explore how the level of profitability of a company, as measured by its return on assets (ROA), affects its use of trade credit. They find that companies with higher ROA are less likely to rely on trade credit as they have access to other sources of financing. On the other hand, companies with lower ROA are more likely to use trade credit to finance their operations and manage their cash flows. This suggests that trade credit plays an important role in helping less profitable firms to maintain liquidity and sustain their business activities.

Aktas, Nihat, et al. "The information content of trade credit." Journal of Banking & Finance 36.5 (2012): 1402-1413. https://doi.org/10.1016/j.jbankfin.2011.12.001

1.2. Discrete variable:

Based on the decree 80/2021/ND-CP issued by the Prime Minister on detailed "Regulations and guidelines for the implementation of some provisions of the Law on Support for Small and Medium-sized Enterprises" on August 26, 2021, the criteria for determining small and medium-sized enterprises are defined as follows:

- Micro-enterprises have an annual total revenue of no more than 10 billion VND or an annual total capital of no more than 3 billion VND.
- Small enterprises have an annual total revenue of no more than 100 billion VND or an annual total capital of no more than 50 billion VND, but not a micro-enterprise as defined in Clause 1 of this Article.
- Medium-sized enterprises have an annual total revenue of no more than 300 billion VND or an annual total capital of no more than 100 billion VND, but not a micro-enterprise or small enterprise as defined in Clause 1 and Clause 2 of this Article.

Any cases not specified in the regulation will be labeled as "Large enterprises" to complete the dataset for analysis and research purposes.

With the available data set, the variables of market capital and total revenue fulfill the conditions for creating the company size variable, therefore the company size variable was created.

1.3. Continuous variable:

According to the above research papers, there is a mixed relationship between ROA and trade credit. Some studies show that there is a positive relationship between ROA and trade credit. Higher ROA indicates better financial performance and profitability, which in turn provides greater access to external financing and bargaining power with suppliers, allowing firms to demand more trade credit. Moreover, firms with higher ROA are perceived as less risky borrowers by suppliers, making them more likely to receive trade credit. On the other hand, lower ROA firms are less likely to demand trade credit as they are perceived as riskier borrowers and have limited access to other sources of financing.

Some studies have found a negative relationship, suggesting that firms with higher profitability may have lower demand for trade credit due to their strict financial policy. As on the other hand, when the return on assets decreases, the company tends to rely more on trade credit as a source of financing for its operations. In other words, companies with lower profitability (measured by ROA) are more likely to use trade credit to manage their cash flow and fund their working capital needs.

Additionally, some studies have found no significant relationship between the two variables. To sum up, the relationship between ROA and trade credit is complex and depends on various factors.

CHAPTER 2: PERFORM CODING TASKS

2. Create dataset

Import library:

```r
library(dplyr) #for manipulate data
library(readxl) #library use to read excel file
library(ggplot2) #library for plotting and data visualization
```

Import file data:

```r
#Task 2: Create Dataset
#import file data
df = read_excel("C:/Users/Admin/Desktop/courses/Gói phần mềm ứng dụng trong tài chính 2/040522 Data Mid-term test Final.xlsx")
View(df)
```

Dataframe have the last row is not an observation, so we drop it out:

```r
#dataframe have the last row is not an observations, so we drop it out
df = df[-(nrow(df)), ]
```

Filter data for just HANOI STOCK EXCHANGE companies to a new dataset called hnx.df:

```r
#filter data for just HANOI STOCK EXCHANGE companies to a new dataset called hnx.df
hnx.df = filter(df, exchangename == "HANOI STOCK EXCHANGE")
```

	no	firmcode	firmname	exchangename	totalasset	roa	totalcurrentasset	totaldebt	totalequity	ppe
1	3	V21.HN	Vinaconex 21 JSC	HANOI STOCK EXCHANGE	3.355799e+11	9.606379e-03	2.428419e+11	4.043566e+10	1.209177e+11	6.7404
2	4	LIG.HN	Licogi 13 JSC	HANOI STOCK EXCHANGE	5.192755e+12	8.785151e-03	3.480836e+12	1.469638e+12	7.095460e+11	1.0637
3	5	MCC.HN	High Grade Brick Tile Corp	HANOI STOCK EXCHANGE	9.295714e+10	2.392523e-02	2.620235e+10	4.000000e+09	7.103377e+10	2.7307
4	6	TET.HN	Northern Textiles And Garments JSC	HANOI STOCK EXCHANGE	1.108181e+11	1.043459e-01	7.309329e+10	0.000000e+00	1.055193e+11	3.7724
5	7	KSD.HN	DNA Investment Joint Stock Corp	HANOI STOCK EXCHANGE	1.322973e+11	1.623126e-01	5.367983e+10	0.000000e+00	1.206204e+11	5.6556
6	8	NBW.HN	Nha Be Water Supply JSC	HANOI STOCK EXCHANGE	2.586037e+11	6.666854e-02	1.051216e+11	NA	1.683347e+11	1.3753
7	11	HTC.HN	HocMon Trade JSC	HANOI STOCK EXCHANGE	5.935884e+11	6.232567e-02	1.864740e+11	2.920000e+10	2.863487e+11	2.8139
8	13	VCM.HN	Vinaconex Trading And Manpower JSC	HANOI STOCK EXCHANGE	9.160112e+10	1.004567e-02	5.293963e+10	0.000000e+00	7.106739e+10	1.4147
9	15	CX8.HN	Constrexim No 8 Investment and Construction JSC	HANOI STOCK EXCHANGE	7.994757e+10	5.472883e-04	7.308603e+10	1.088069e+10	2.758408e+10	6.6588

Create a new dataset (sample.df) with 100 autoselected samples:

```r
#create a new dataset (sample.df) with 100 autoselected samples
set.seed(961)
sample.df = hnx.df[sample(1:nrow(hnx.df), 100, replace = F),]
View(sample.df) #note: 1row no 517, 100row 655
```

	no	firmcode	firmname	exchangename	totalasset	roa	totalcurrentasset	totaldebt	totalequity	ppe
1	517	TST.HN	Telecommunication Technical Service JSC	HANOI STOCK EXCHANGE	2.861844e+11	-8.648038e-04	2.377324e+11	4.294490e+10	8.788588e+10	9.9995
2	527	VCS.HN	Vicostone JSC	HANOI STOCK EXCHANGE	6.892914e+12	2.737142e-01	5.829100e+12	1.625608e+12	4.874197e+12	1.0409
3	164	KSQ.HN	CNC Capital Viet Nam JSC	HANOI STOCK EXCHANGE	3.633148e+11	1.101198e-02	3.633148e+11	3.850000e+10	3.222469e+11	4.1726
4	94	KTS.HN	Kon Tum Sugar JSC	HANOI STOCK EXCHANGE	4.395785e+11	1.352814e-02	2.152207e+11	2.399188e+11	1.631573e+11	2.2051
5	197	FID.HN	Vietnam Enterprise Investment and Development JSC	HANOI STOCK EXCHANGE	2.529269e+11	3.935481e-03	2.465661e+10	5.738207e+10	2.483208e+11	3.9630
6	549	CAP.HN	Yenbai Joint Stocks Forest Agricultural Products and Foodst...	HANOI STOCK EXCHANGE	1.526859e+11	4.257653e-01	1.006671e+11	0.000000e+00	1.194766e+11	4.9197
7	518	VC6.HN	Visicons Construction and Investment JSC	HANOI STOCK EXCHANGE	6.486081e+11	1.314793e-02	5.933375e+11	2.056562e+11	1.145932e+11	4.3717
8	120	INC.HN	IDICO Investment Consultancy JSC	HANOI STOCK EXCHANGE	3.616901e+10	6.977301e-02	2.878113e+10	0.000000e+00	2.593866e+10	4.8799
9	49	VLA.HN	Van Lang Technology Development and Investment JSC	HANOI STOCK EXCHANGE	2.526022e+10	2.757714e-01	1.669993e+10	1.000000e+09	2.073102e+10	2.7410
10	250	CTP.HN	Minh Khang Capital Trading Public JSC	HANOI STOCK EXCHANGE	1.670508e+11	5.346442e-03	1.641271e+11	0.000000e+00	1.502057e+11	0.0000

Scan the dataset to see whether we have NA values:

```r
#scan the dataset to see whether we have NA values
colSums(is.na(sample.df))
sum(is.na(sample.df))
```

```
> #scan the dataset to see whether we have NA values
> colSums(is.na(sample.df))
                     no               firmcode               firmname
                      0                      0                      0
           exchangename             totalasset                    roa
                      0                      1                      3
      totalcurrentasset              totaldebt            totalequity
                      3                      5                      1
                    ppe                cashopr                cashfin
                      2                      1                      5
             cashinvest                    ptb           depreciation
                      1                      1                     26
     currentliabilities                   cash          cash_dividend
                      3                      3                     33
              marketcap   share_held_strategic           freecashflow
                      0                      1                      4
                revenue             receivable                payable
                      3                      3                      2
commonshares_outstanding              industry
                      3                      0
> sum(is.na(sample.df))
[1] 104
```

Impute the missing values with median of the corresponding variable:

```
#impute the missing values with median of the corresponding variable
sample.df %>%
  mutate_if(is.numeric, function(x) ifelse(is.na(x), median(x, na.rm = T), x)) -> sample.df
```

Recheck for the NA values:

```
#recheck for the NA values
colSums(is.na(sample.df))
sum(is.na(sample.df))
```

```
>
> #recheck for the NA values
> colSums(is.na(sample.df))
                     no               firmcode               firmname
                      0                      0                      0
           exchangename             totalasset                    roa
                      0                      0                      0
      totalcurrentasset              totaldebt            totalequity
                      0                      0                      0
                    ppe                cashopr                cashfin
                      0                      0                      0
             cashinvest                    ptb           depreciation
                      0                      0                      0
     currentliabilities                   cash          cash_dividend
                      0                      0                      0
              marketcap   share_held_strategic           freecashflow
                      0                      0                      0
                revenue             receivable                payable
                      0                      0                      0
commonshares_outstanding              industry
                      0                      0
> sum(is.na(sample.df))
[1] 0
>
```

3. Report

3.1. 5 firms with highest trade credit (receivable):

```
#Task 3: Report
#5 firms with highest trade credit (receivable)
max(sample.df$receivable)
highest_receivable = sample.df[order(sample.df$receivable,decreasing=T)[1:5],]
View(highest_receivable)
cat(paste('Top 5 companies with the highest trade credit (receivable): \n', list(highest_receivable$firmname)))
```

	no	firmcode	firmname	exchangename	totalasset	roa	totalcurrentasset	totaldebt	totalequity	ppe
1	527	VCS.HN	Vicostone JSC	HANOI STOCK EXCHANGE	6.892914e+12	0.273714232	5.829100e+12	1.625608e+12	4.874197e+12	1.040927e+12
2	341	THD.HN	Thaiholdings JSC	HANOI STOCK EXCHANGE	1.057201e+13	0.110029267	2.458023e+12	2.296195e+12	5.333790e+12	1.260205e+12
3	394	SDT.HN	Song Da No 10 JSC	HANOI STOCK EXCHANGE	2.755471e+12	-0.004984592	1.984323e+12	1.055374e+12	8.481480e+11	6.305302e+11
4	474	DNP.HN	Dongnai Plastic JSC	HANOI STOCK EXCHANGE	1.404001e+13	0.001913715	4.760223e+12	7.276634e+12	1.534780e+12	6.032829e+12
5	4	LIG.HN	Licogi 13 JSC	HANOI STOCK EXCHANGE	5.192755e+12	0.008785151	3.480836e+12	1.469638e+12	7.095460e+11	1.063767e+12

```
Top 5 companies with the highest trade credit (receivable):
 c("Vicostone JSC", "Thaiholdings JSC", "Song Da No 10 JSC", "Dongnai Plastic JSC", "Licogi 13 JSC")>
```

3.2. 5 firms with lowest trade credit (receivable):

```
#5 firms with lowest trade credit (receivable)
min(sample.df$receivable)
lowest_receivable = sample.df[order(sample.df$receivable,decreasing=F)[1:5],]
View(lowest_receivable)
cat(paste('Top 5 companies with the lowest trade credit (receivable): \n', list(lowest_receivable$firmname)))
```

	no	firmcode	firmname	exchangename	totalasset	roa	totalcurrentasset	totaldebt	totalequity	ppe
1	641	SHS.HN	Saigon Hanoi Securities JSC	HANOI STOCK EXCHANGE	1.091088e+13	0.15652875	1.083904e+13	3.84800e+12	6.055580e+12	12379402980
2	749	PVL.HN	Viet Property Investment JSC	HANOI STOCK EXCHANGE	4.094021e+11	0.04237970	1.598911e+11	0.00000e+00	2.524838e+11	45893599380
3	87	PV2.HN	PV2 Investment JSC	HANOI STOCK EXCHANGE	3.811657e+11	0.01210891	1.748474e+11	0.00000e+00	2.407352e+11	184417187640
4	295	ART.HN	BOS Securities JSC	HANOI STOCK EXCHANGE	1.171495e+12	0.02937352	1.134129e+12	0.00000e+00	1.157513e+12	2733218320
5	492	BVS.HN	Bao Viet Securities JSC	HANOI STOCK EXCHANGE	5.719297e+12	0.06196569	5.502631e+12	2.60589e+12	2.144452e+12	11490519740

```
Top 5 companies with the lowest trade credit (receivable):
 c("Saigon Hanoi Securities JSC", "Viet Property Investment JSC", "PV2 Investment JSC", "BOS Securities JS
C", "Bao Viet Securities JSC")>
```

3.3. The name of industries which the firms belong to:

```
#the name of industries which the firms belong to
select(lowest_receivable, firmname, industry)
select(highest_receivable, firmname, industry)
a=highest_receivable$industry
b=lowest_receivable$industry
c=c(a,b)
cat(paste('Name of industries which the firms belong to: \n', list(unique(c))))
```

```
Name of industries which the firms belong to:
 c("Consumer Cyclicals", "Basic Materials", "Industrials", "Financials", "Real Estate")
```

3.4. Provide descriptive statistics with median, mean, max, min, standard deviation of trade credit of:

Add company size into dataset:

```
#add company size in to dataset
company.size=c()
for (i in rownames(sample.df)) {
  if (sample.df[i, 'totalequity'] <= 3000000000 | sample.df[i, 'revenue'] <= 10000000000) {
    company.size <- append(company.size, 'Micro')
  } else if ((10000000000 < sample.df[i, 'revenue'] & sample.df[i, 'revenue'] < 100000000000) | sample.df[i, 'totalequity'] < 50000000000) {
    company.size <- append(company.size, 'Small')
  } else if ((100000000000 > sample.df[i, 'totalequity']) & (300000000000 > sample.df[i, 'revenue'])) {
    company.size <- append(company.size, 'Medium')
  } else {
    company.size <- append(company.size, 'Large')
  }
}
company.size
#count number of company types
table(company.size)

#append into sample.df as new column named companysize
sample.df$companysize = company.size
```

```
> #count number of company types
> table(company.size)
company.size
 Large Medium  Micro  Small
    65      7      1     27
```

	cash_dividend	marketcap	share_held_strategic	freecashflow	revenue	receivable	payable	commonshares_outstanding	industry	companysize
09	373300	6.000000e+10	2812000	1333424600	9.509088e+10	89141995220	8.862217e+10	4800000	Technology	Small
12	642014742700	1.872000e+13	141219228	1216993587730	7.070129e+12	2401752857940	2.523126e+11	160000000	Consumer Cyclicals	Large
08	7147580000	2.760000e+11	14807800	10645686850	3.479133e+10	270460800	3.278830e+10	30000000	Industrials	Small
09	7147580000	1.135680e+11	1362690	17278934070	2.482185e+11	160526353380	1.890367e+10	5070000	Consumer Non-Cyclicals	Large
09	7147580000	2.273868e+11	3755532	4027104070	1.403618e+11	5933876000	4.464355e+09	23539960	Industrials	Large
10	20944092000	4.555340e+11	2674198	24061565170	5.179688e+11	5310273690	4.027380e+09	5236020	Basic Materials	Large
10	4800000000	1.152000e+11	5816580	6950196920	9.386389e+11	386125884790	2.567763e+11	8000000	Industrials	Large
09	2735141640	2.400000e+10	1571100	383506690	2.421950e+10	6801942850	2.057718e+09	2000000	Industrials	Small
10	540000000	3.078000e+10	704800	2135641910	1.445424e+10	817814600	4.158212e+08	1080000	Technology	Small

Group data by the discrete variable and calculate median, mean, max, min, standard deviation of trade credit (receivable), different categories of the discrete variable:

```
#group data by the discrete variable
sample.df %>%
  mutate(companysize = factor(companysize, levels = c("Large", "Medium", "Small", "Micro"))) %>%
  group_by(companysize) %>%
  summarise(
    median_tc = median(receivable),
    mean_tc = mean(receivable),
    max_tc = max(receivable),
    min_tc = min(receivable),
    sd_tc = sd(receivable)
  )
```

```
# A tibble: 4 × 6
  companysize    median_tc        mean_tc          max_tc          min_tc          sd_tc
  <fct>              <dbl>          <dbl>           <dbl>           <dbl>          <dbl>
1 Large        89031671830  260897295967.   2401752857940   -425539669650  445628414937.
2 Medium       39502025370   39230734649.     75378556130      6053073710   26869913508.
3 Small         9544929990   16205282025.     94364399580    -48606751860   28930940166.
4 Micro       -40672951490  -40672951490    -40672951490    -40672951490             NA
>
```

Comment: Based on the output table, it seems that there is a positive relationship between company size and trade credit. The median and mean trade credit are highest for large companies, followed by medium and small companies. Micro companies have a negative value for all trade credit variables, it repeats itself though all columns, so it can be viewed as an outlier classified as a micro enterprise by our code, therefore should not be included in our analysis. The standard deviation of trade credit also appears to be higher for larger companies, suggesting that there is more variation in the amount of trade credit extended by larger companies.

The initial observations and analysis from the dataset indicate that they are consistent with our literature review.

Group data by the continuous variable and calculate median, mean, max, min, standard deviation of trade credit (receivable), groups of above/below median of the continuous variable:

```
> median(sample.df$roa)
[1] 0.02937352
```

```r
#group data by the continuous variable and calculate median, mean, max, min, standard deviation of trade credit
sample.df %>%
  mutate(x=cut(sample.df$roa,breaks=c(min(sample.df$roa),
                                       median(sample.df$roa),
                                       max(sample.df$roa)),
               labels=c("Below Median","Above Median"))) %>%
  group_by(x) %>%
  filter(!is.na(x)) %>%  # filter out any rows where x is NA
  summarise(mean_tc=mean(receivable),
            median_tc=median(receivable),
            max_tc=max(receivable),
            min_tc=min(receivable),
            std_tc=sd(receivable)
  )
```

```
# A tibble: 2 × 6
  x                  mean_tc   median_tc        max_tc        min_tc        std_tc
  <fct>                <dbl>       <dbl>         <dbl>         <dbl>         <dbl>
1 Below Median 222295378546. 85043666420 1350271581700  -40672951490 305715012069.
2 Above Median 130475082869. 21989987435 2401752857940 -425539669650 441663105875.
>
```

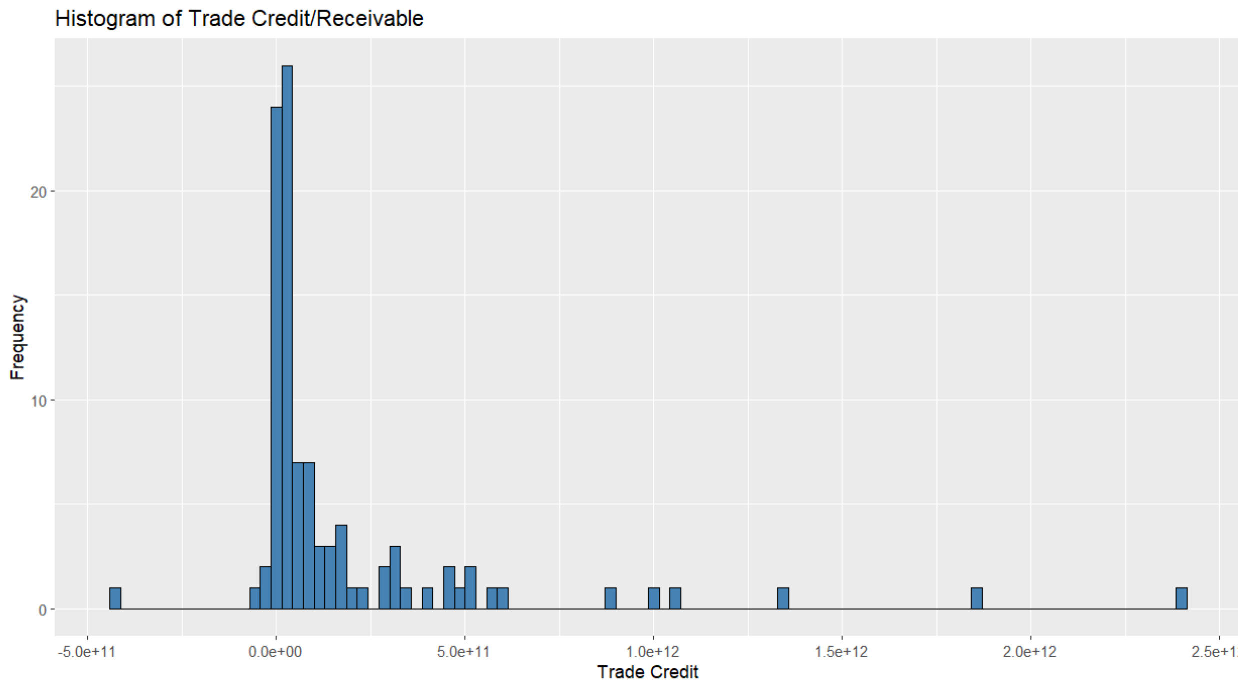
Comment: Based on the output table, we can see that there is a significant difference in the mean and median trade credit for companies below and above the median ROA. Companies with ROA below the median have a higher mean and median trade credit compared to those with ROA above the median. This suggests that there may be a possible link between lower profitability (as indicated by ROA) and higher trade credit.

The initial observations and analysis indicate that it is possible that our dataset aligns with the hypothesis of an inverse relationship between ROA and trade credit in empirical studies. This suggests that companies with low profitability tend to increase their trade credit to manage their cash flow and fund their working capital needs.

4. Data visualization

4.1. Provide histogram of trade credit:

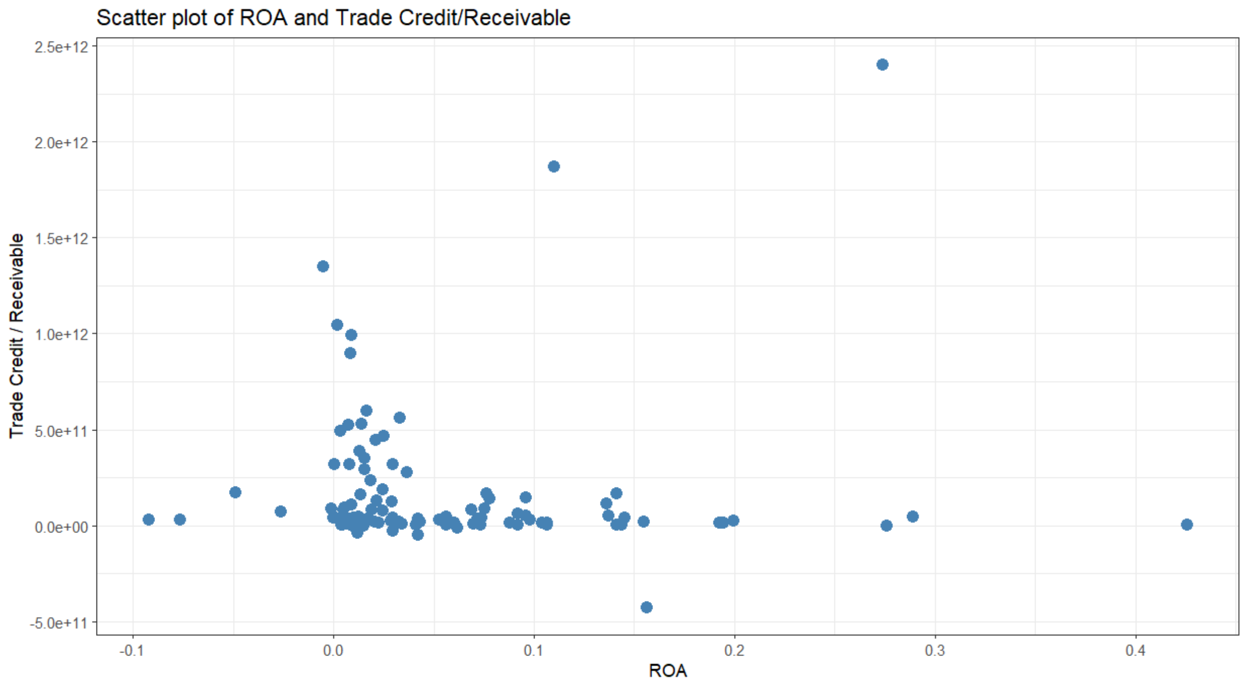
```r
#Task 4: Data visualization
#histogram plot for trade credit receivable
ggplot(sample.df, aes(x = receivable)) +
  geom_histogram(fill = "steelblue", color = "black", bins = 100) +
  labs(title = "Histogram of Trade Credit/Receivable", x = "Trade Credit", y = "Frequency")
```



Comment: The histogram chart is divided into 100 equally sized intervals (bins=100 in the code). The trade credit receivables of the companies are mainly concentrated in the range from 0 to 200 billion, then scattered and gradually decrease towards 500 billion. There are scattered outliers ranging from 500 billion to the highest at 2500 billion. The mode of the histogram is located at the first interval on the right of interval 0 which has 26 companies.

4.2. Provide scatter plot of trade credit with the continuous variable:

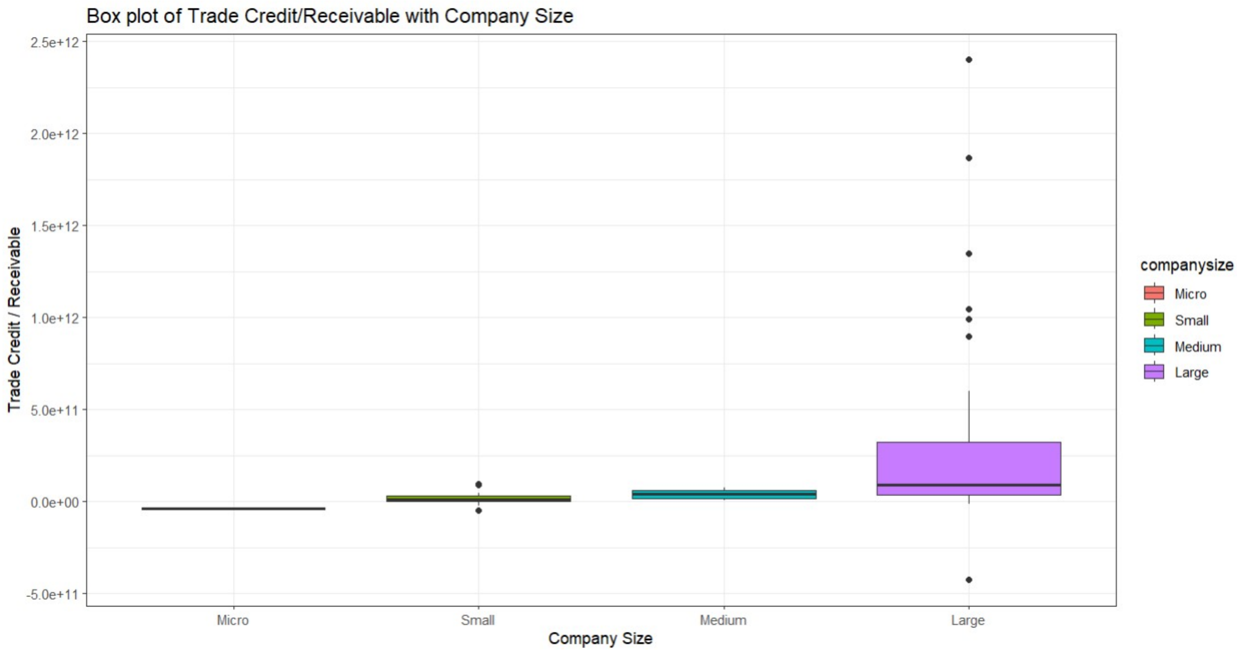
```
#scatter plot of ROA and receivable
ggplot(sample.df, aes(x=roa, y=receivable)) +
  geom_point(size=3, color="steelblue") +
  labs(title = "Scatter plot of ROA and Trade Credit/Receivable", x="ROA", y="Trade Credit / Receivable") +
  theme_bw()
```



Comment: The scatter plot illustrates the relationship between ROA and trade credit, and for the selected dataset, there appears to be no clear linear relationship between ROA and trade credit. Companies with high profitability still maintain their receivable levels within the range of 0 to 200 billion, while some low-profit companies still have very high receivables, ranging from 0 to 1500 billion. Companies are mainly concentrated in the region with ROA (0,0.05) and Receivable (0,250 billion). They are scattered without following any specific order.

4.3. Provide boxplot of trade credit with the discrete variable (different colour for different categories of discrete variable):

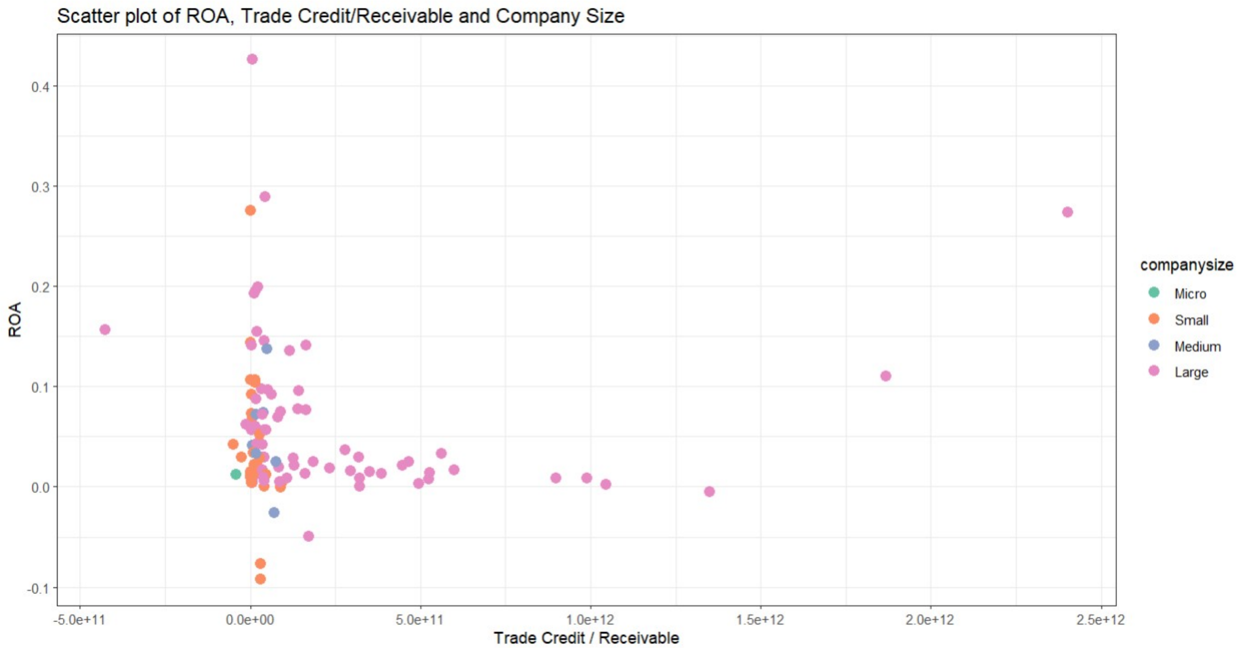
```
#reorder company by company size
sample.df$companysize <- factor(sample.df$companysize, levels = c("Micro", "Small", "Medium", "Large"))

#box plot of Trade credit with company size
ggplot(sample.df, aes(x=companysize, y=receivable, fill=companysize)) +
  geom_boxplot() +
  labs(title = "Box plot of Trade Credit/Receivable with Company Size", x="Company Size", y="Trade Credit / Receivabl
  theme_bw()
```



Comment: The boxplot illustrates the trade credit of four types of companies: Micro, Small, Medium, and Large, which is a discrete variable. The receivable of large companies, although at the same level as the other three types, has much greater variation. Moreover, there are more outliers for large companies than for the other three types, where no outliers are present. The other three types have relatively similar and lower levels of receivable, with no significant difference among them. The ranges of trade credit receivable increase gradually with the company's scale, and the width is also widened, indicating that larger companies are more likely to negotiate for better and more flexible trade credit and manage it more flexibly.

4.4. Provide a plot that allow the combination of continuous, discrete variables and trade credit:

```
#Scatter plot of ROA, Trade Credit/Receivable and Company Size
ggplot(sample.df, aes(x = receivable, y = roa , color = companysize)) +
  geom_point(size=3) +
  scale_color_manual(values = c("#66c2a5", "#fc8d62", "#8da0cb", "#e78ac3")) +
  labs(title = "Scatter plot of ROA, Trade Credit/Receivable and Company Size", x = "Trade Credit / Receivable", y =
  theme_bw()
```



Comment: The scatter plot presented in this figure depicts the relationship between ROA and trade credit for four different types of companies: Micro, Small, Medium, and Large. Each type of company is represented by a different color on the plot. Upon analysis, it appears that there is no clear linear relationship between ROA and trade credit in the selected dataset. Companies with high profitability still maintain their receivable levels within the range of 0 to 200 billion, while some low-profit companies still have very high receivables, ranging from 0 to 1500 billion. The companies are mainly concentrated in the region with ROA (0,0.05) and Receivable (0,250 billion), and they are scattered without following any specific order.

5. Regression

5.1. Conduct regression analysis of the determinants of trade credit:

```
#Task 5: Regression
#regression
y = sample.df$receivable
x1 = sample.df$roa
x2 = sample.df$companysize
model = lm(y ~ x1 + x2, data = sample.df)

summary(model)
```

```
> summary(model)

Call:
lm(formula = y ~ x1 + x2, data = sample.df)

Residuals:
       Min         1Q     Median         3Q        Max 
-6.786e+11 -2.126e+11 -1.938e+10  1.785e+10  2.159e+12 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -3.963e+10  3.661e+11  -0.108    0.914
x1          -8.576e+10  4.860e+11  -0.176    0.860
x2Small      5.939e+10  3.731e+11   0.159    0.874
x2Medium     8.323e+10  3.918e+11   0.212    0.832
x2Large      3.061e+11  3.698e+11   0.828    0.410

Residual standard error: 3.661e+11 on 95 degrees of freedom
Multiple R-squared:  0.09523,	Adjusted R-squared:  0.05713 
F-statistic:   2.5 on 4 and 95 DF,  p-value: 0.04762
```

Comment: Based on the model summarize, we can draw the following conclusions regarding the relationship between ROA (x1) and Receivables (y):

- The coefficient estimate for x1 (ROA) is negative (-8.576e+10), indicating that there is a negative relationship between ROA and Receivables.

- The coefficient estimates for x2Small, x2Medium, and x2Large are all positive, indicating that there is a positive relationship between these variables and receivable (y).

- The p-value for x1 is 0.860, which is greater than the typical significance level of 0.05. This suggests that the coefficient for x1 may not be statistically significant, and we cannot conclude with confidence that there is a significant relationship between ROA and Receivables.

- The same senerio happen with the p-values of x2Small, x2Medium, and x2Large coefficients, these are all greater than 0.05, which suggests that these coefficients are not statistically significant. In other words, we cannot conclude with confidence that there is a significant relationship between x2Small, x2Medium, x2Large and receivable (y).

- The adjusted R-squared value is 0.05713, which indicates that only 5.7% of the variation in Receivables can be explained by the variation in ROA and the other independent variables (x2Small, x2Medium, and x2Large) included in the model.

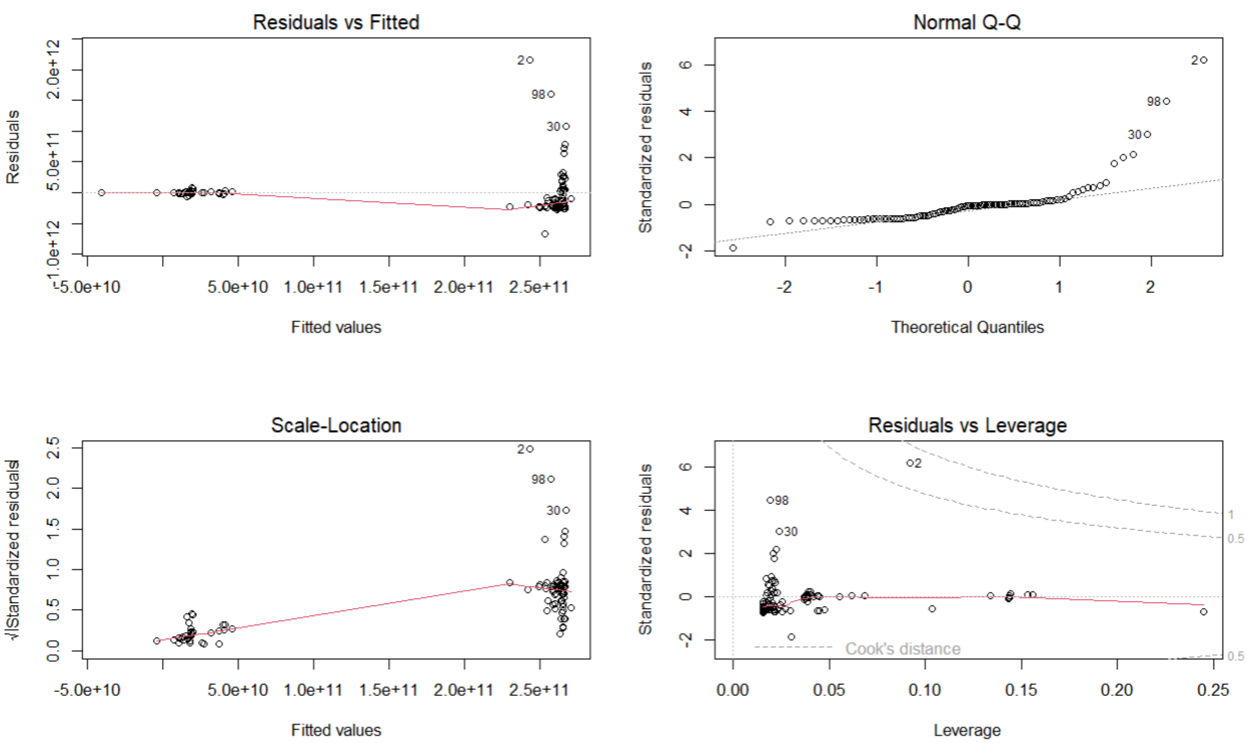
To sum up, the analysis suggests that there may be a negative relationship between ROA and Receivables, but we cannot conclude with certainty whether this relationship is statistically significant, as well as there appears to be a positive relationship between x2Small, x2Medium, x2Large and receivable (y), but this relationship is not strong statistically. Therefore, the low adjusted R-squared value indicates that there may be other factors influencing Receivables beyond the ones included in the model.

5.2. Perform test of multicollinearity, heteroskedasticity, then determine if any correction measure is required to obtain reliable estimates.

Because there is no statistically significant relationship between the variables and the receivable, then there should not be any issue of multicollinearity as multicollinearity occurs when there is a strong linear relationship between two or more predictor variables. However, it is still possible to have heteroskedasticity in the data even if there is no relationship between the predictor variables, as heteroskedasticity refers to the unequal variance of the residuals across different values of the predictor variables. But to ensure a comprehension of the paper, we will proceed to analyze the presence of multicollinearity and heteroskedasticity of the model, assuming it exists.

Heteroscedasticity

```
#plot the graph of the model
par(mfrow = c(2, 2))
plot(model)
```



Comment: The graphs above show the "Residual vs. Fitted"-plot and the "Scale-Location"-plot for a regression model with heteroscedastic residuals. The "Residuals vs. Fitted" plot shows an increase in variance across the fitted values. For the left cluster, the residuals are not much different. Whereas, for the right cluster, the difference between the residuals (i.e., variability) becomes bigger. This is a clear indicator of heteroscedasticity.

Moreover, in the "Scale-Location"-plot, the increase in the square root of the standardized residuals across the fitted values is evident. Therefore, combining the "Residuals vs. Fitted"-plot and the "Scale-Location"-plot, we can conclude that this model violates the homoscedasticity assumption.

Multicollinearity

Note: In order to comply with the requirements of the task, using external packages for data analysis is not permitted. During the research process, I was able to evaluate heteroskedasticity by reading the graph, without using external packages. However, when it comes to assessing the multicollinearity of the model, I couldn't find another method that doesn't involve using external packages. In order to complete the report thoroughly, I will run the package and evaluate the results displayed in this report. And in the code file, I will turn the code running package into a comment to ensure that it does not affect the grading process of the lecture.

```
install.packages("olsrr")
library("olsrr")
ols_vif_tol(model)
```

```
> ols_vif_tol(model)
  Variables  Tolerance       VIF
1        x1 0.97756051  1.022955
2   x2Small 0.04885265 20.469720
3  x2Medium 0.13409821  7.457221
4   x2Large 0.04308101 23.212083
>
```

The outcome table indicates the multicollinearity diagnostics for four variables: x1 (as for ROA variable), x2Small, x2Medium, and x2Large.

- Tolerance: It measures the proportion of variance in one variable that is not accounted for by the others. A tolerance value closer to 1 indicates less correlation with other variables.

- VIF: It stands for variance inflation factor, and it measures the degree of correlation between each independent variable and other independent variables in the model. A VIF value greater than 1 indicates the presence of correlation, and a value greater than 5 or 10 suggests high multicollinearity.

Based on the table, x1 (ROA) and x2Medium have tolerance values of 0.977 and 0.13409821, and VIF of 1.02 and 7.457221, which suggests low correlation with the other variables (the tolerance value closer to 1 and VIF value lower that 5, indicate low correlation). On the other hand, x2Small and x2Large have low tolerance values (0.049 and 0.043, respectively) and high VIF values (20.47 and 23.21, respectively), (the compare value for tolerance is below 0.1 and VIF value is above 10) which indicates a high degree of correlation with other variables, indicating the presence of multicollinearity.

Handle method of multicollinearity and heteroscedasticity

In general, there are 3 options to handle multicollinearity:

- Do nothing. As multicollinearity only affects the regression coefficients and not the actual predictions, it is not always necessary to do something. If you are only interested in the prediction of your model and not in the significance of the coefficients, you must do nothing.

- Remove one or more variables. However, if you are interested in the regression coefficients, multicollinearity is a problem. Therefore, the easiest way to deal with it is by removing the variable(s) that cause this issue.

- Combing variables. Instead of removing variables, you could also combine the variables that cause multicollinearity into a new variable.

And in case of heteroscedasticity, there are several ways to address it in a regression analysis:

- Transform the data: Transforming the data by taking the logarithm or square root of the dependent variable or predictor variable can help to reduce the effect of heteroscedasticity.

- Use weighted least squares: Weighted least squares regression gives more weight to observations with smaller variances and less weight to those with larger variances, allowing for a better fit of the model to the data.

- Use robust standard errors: Robust standard errors take into account the heteroscedasticity of the residuals and provide more accurate estimates of the standard errors of the regression coefficients.

- Use a different regression method: Some regression methods, such as quantile regression, are less sensitive to heteroscedasticity than ordinary least squares regression.

- Include additional variables in the model: Adding additional predictor variables that explain the variance in the dependent variable can help to reduce heteroscedasticity.

6. Using LOOP

Count the number of firms in an industry (if you are given an industry name, you can count the number of firms in that industry)

```
#Task 6: Using LOOP
#the LOOP 1
# choose the industry you want to count how many firm, example here is "Industrials"
chosen.industry01 = "Industrials"
firm.count01 = 0
for (i in 1:nrow(df)) {
  if (df[i, "industry"] == chosen.industry01) {
    firm.count01 = firm.count01 + 1
  }
}
#print the firm count
cat(paste("The number of companies in", chosen.industry01, "is", firm.count01))
```

```
> cat(paste("The number of companies in", chosen.industry01, "is", firm.count01))
The number of companies in Industrials is 200
```

Count the number of firms in an industry and with trade credit above a certain value (if you are given an industry name and a specific value of trade credit, you can count the number of firms in that industry and above that certain value)

```
#the LOOP 2
# choose the industry you want to count how many firm, example here is "Technology"
# and choose the trade credit limit for counting number of company above of, example here is "10ty"
chosen.industry02 = "Technology"
tradecredit.above.limit = 10000000000
firm.count02 = 0
for (i in 1:nrow(df)) {
  if (df[i, "industry"] == chosen.industry02 & df[i, "receivable"]> tradecredit.above.limit) {
    firm.count02 = firm.count02 + 1
  }
}
#print the firm count
cat(paste("The number of companies in", chosen.industry02, "and have trade credit/receivable above", tradecredit.above.limit,

#end-------
```

```
> cat(paste("The number of companies in", chosen.industry02, "and have trade credit/receivable above", tradecredit.
above.limit, "is", firm.count02))
The number of companies in Technology and have trade credit/receivable above 1e+10 is 16
```

REFERENCES

1. Martínez-Sola, Cristina, Pedro J. García-Teruel, and Pedro Martínez-Solano. *"Trade credit and SME profitability."* Small Business Economics 42 (2014): 561-577. https://doi.org/10.1007/s11187-013-9491-y
2. Lawrenz, Jochen, and Julia Oberndorfer. *"Firm size effects in trade credit supply and demand."* Journal of Banking & Finance 93 (2018): 1-20. https://doi.org/10.1016/j.jbankfin.2018.05.014
3. García-Teruel, Pedro Juan, and Pedro Martinez-Solano. *"Determinants of trade credit: A comparative study of European SMEs."* International Small Business Journal 28.3 (2010): 215-233. https://doi.org/10.1177/0266242609360603
4. Phạm Xuân Quỳnh, Trần Đức Tuấn, *"Các nhân tố ảnh hưởng đến tín dụng thương mại của các công ty ngành thực phẩm niêm yết trên HOSE"* Tạp chí Khoa học Kinh tế No.8 (2020) https://vjol.info.vn/index.php/due/article/view/48527
5. Ahmed, Jaleel, Hui Xiaofeng, and Jaweria Khalid. *"Determinants of trade credit: The case of a developing economy."* European Researcher 9-2 (2014): 1694-1706. https://doi.org/10.13187/er.2014.83.1694
6. M Hill, K Hill, L Preve, V Sarria-Allende *"International evidence on the determinants of trade credit provision."* Managerial Finance 45.4 (2019): 484-498. https://doi:10.1108/mf-07-2018-0295
7. Oh, S. Y., & Hwang, S. W. (2019). *The effect of market power on trade credit in Korean manufacturing firms*. Asia-Pacific Journal of Accounting & Economics, 26(3-4), 294-314.
8. Aktas, Nihat, et al. "The information content of trade credit." Journal of Banking & Finance 36.5 (2012): 1402-1413. https://doi.org/10.1016/j.jbankfin.2011.12.001