HCM CITY NATIONAL UNIVERSITY

UNIVERSITY OF LAW AND ECONOMICS

-------- □ □ □ --------

Midterm TEST

PROGRAM PACK IN FINANCE 2

<FACTORS AFFECTING TRADE CREDIT RATES OF COMPANIES LISTED ON THE HNX>

Study program: K20414C_ Fintech

Module: Packages in financial apps

Learner: Nguyen Thanh Phuong Uyen

MSSV: K204141939

Lecturer: Mr. Nguyen Thanh Liem

SUMMARY

Trade credit is a necessary job for economic and legal stakeholders of the business, especially managers. The objective of this study is to determine the factors affecting the commercial credit of enterprises listed on HNX in Vietnam.. Using data from **752** listed companies, the data source is collected at the end of each year of many companies, so the study should randomly take **100** observations to Analysis of factors affecting commercial credit of enterprises. Research results with the model Regressionshows that customer ***receivables*** are affected by the factors: ***revenue, current liabilities, ppe*** của company

DOCUMENTARY REVIEW

1. Theoretical basis and empirical evidence

1.1. Theoretical Foundations of trade credit

Mian and Smith (1992) argue that trade credit (TDTM) is an agreement between a buyer and a seller, in which the seller allows deferred payment for his product instead of paying in cash. According to Hillier et al (2013), from an accounting perspective, commercial credit is a type of credit that enterprises (enterprises) grant to other companies, generating receivables.

It is an important source of finance for all businesses in the world. Rajan and Zingales (1995) reported that in 1991 trade credit accounted for 17.8% of total assets for all US firms, and in European countries such as Germany, France, and Italy, trade credit represented more than a quarter of the company's total assets. Furthermore, trade credit is also important in emerging economies, such as China, where businesses receive limited support from the banking system.

Trade credit is a popular form of credit in the fields of production and business, is an important source of external finance for enterprises, and can replace bank credit in the period of monetary tightening or financial crisis, in developing countries as well as in developed countries (Bougheas et al., 2009). Theories dealing with this issue include the Financing Advantage Theory, the Trade credit as means of price discrimination, and the Transaction cost theory. Transaction Cost Theory).

Financing Advantage Theory was put forward by the American economist Stewart C. Myers in the 1980s. He argued that a business's ability to finance can create a competitive advantage and help it grow. This can be applied in the field of commercial credit to assess the borrowing capacity of businesses and decide to grant credit to those businesses.

Schwart (1974) and Ferris (1981) - the two who successfully formulated the theory of transaction costs, argued that thanks to TM, businesses not only reduce transaction costs because they can better manage inventory, but also reduce the pressure on capital financing needs.

Usually, the seller will have a TDTM policy with different terms and conditions for different customers or "groups" of customers. From there it is possible to supply goods at different prices, or trade credit is a form of price allocation for different “groups” of customers (Brennan et al., 1988, Huyghebaert, 2006). Petersen and Rajan (1997) argued that commercial credit is a "subsidy" for businesses with limited access to credit.

1.2. Relevant empirical evidence

There are a substantial number of studies that have investigated the determinants of trade credit. However, each economy has distinctive factors which affect trade credit. They may be the differences in financial structure or characteristics of the economy, or even the government policy and legislation, etc. This study will present some different points of view on the factors that concern trade credit.

Nadiri (1969) studied the trade credit of manufacturing enterprises in the United States through quarterly data analysis, from the first quarter of 1949 to the fourth quarter of 1964. Regression analysis results of the log-linear model have identified factors that have a statistically significant influence on commercial practice. Revenue is a factor that has a very positive impact on receivables.

Petersen & Rajan (1997) studied the commercial credit of 3,404 small businesses, from the National Survey of Small Business Finances (NSSBF) data in the United States in 1988 - 1989. OLS analysis results. The pool shows that many factors have a significant influence on commercial sports. Factors affecting receivables: ppe, current liabilities, and revenue have a positive effect, while net profit margin (profit after tax/sales) has a negative effect.

Vaidya (2011) analyzed data for the period 1993 - 2006 of 1,522 companies in India to determine the factors affecting smart shopping. Through the Generalized Method of Moments (GMM), the study has demonstrated that many factors have a statistically significant influence. Factors that positively affect receivables: ppe, liquidity (cash and negotiable securities/revenue), and fixed assets. Factors that have a negative impact: Inventory (inventory/sales), profitability (profit before depreciation, and tax/sales).

Giannetti et al. (2011) research on commercial credit of 3,489 small businesses, from the National Survey of Small Business Finances (NSSBF) data in the United States in 1998. The results of the pooled OLS analysis reveal many factors. significant influence on sports practice. Factors affecting receivable: assets and bank loans have a positive effect; fixed assets and distance to the bank have a negative effect.

Akinlo (2012) analyzes panel data of 66 non-financial companies listed on the Nigerian stock market for the period 1999 - 2007, to study the factors affecting commercial credit. The results of the pooled OLS analysis show that many factors have a significant influence on A/R. Which, bank loans (short-term loans/assets) and liquidity (money and negotiable securities/assets) have a positive impact; Return on assets (ROA), size (sales), and inventory (value of inventory/assets) have a negative effect. On the other hand, through the Hausman test (Hausman test), the study shows that the FEM method gives better results than the REM method. The results of the

FEM analysis only confirmed 2 factors affecting receivables: Size has a negative effect and liquidity has a positive effect.

Santos & Silva (2014) analyzed the data of 11,040 industrial enterprises in Portugal for the period 2003 - 2009 using the FEM method to determine the factors affecting the trade credit. Regression analysis also shows the determining factors of receivables. Years of operation, profit margin, equity, and revenue have a positive impact. In addition, accounts payable and official credit shortfall are factors that have a negative impact onreceivable of the companies surveyed.

2. Influential factors & estimation methods

2.1. The factors affecting trade credit

Based on the theoretical basis and empirical evidence just presented above help form a research model of the impact of factors on the commercial credit of enterprises as follows:

<u>Dependent variable</u>: **Receivable (Y)**

Hult et al (2008), Onaolapo and Kajola (2010), Alimadar et al. (2018), Van Cong Nguyen et al (2019)... all said that the credit assessment criteria of enterprises were used. The most commonly used in empirical studies is through financial ratios: Accounts Receivable

The reason for this is that receivables are often one of the most important metrics for assessing the financial health of businesses. Receivables also show the confidence level of customers and the ability of the business to collect money. The use of receivables as the dependent variable allows researchers to focus on the effectiveness of strategies and financial decisions of the enterprise in managing its receivables.

<u>Independent variables:</u>

Based on the business characteristics of enterprises listed on HNX Vietnam as well as inheriting previous studies, the selection of 3 independent variables representing 3 groups of factors: assets and liabilities of enterprises, accounts receivables and payables of the enterprise, and the revenue of the enterprise, specifically: PPE, Current liabilities, Revenue, corresponding to the following hypotheses:

Hypothesis 1:**PPE (POS)** has the same/inverse effect on the receivables of enterprises.

The theory of economic advantage due to this variable can affect receivables through its dependence on the firm's use of fixed assets. If an enterprise invests in large fixed assets, it will limit the capital investment in other current assets thereby reducing the business's trade credit and vice versa if the investment enterprise increases the number of current assets of enterprises, will affect the value of fixed assets, thereby promoting the increase in commercial credit of enterprises (Tran Ai Ket, 2014).

Hypothesis 2:**Currentliabilities (BOR)** have the same/reverse effect on receivables of enterprises.

Short-term debt: If the value of short-term bank loans is low, it means that the enterprise restricts the use of credit sources from commercial banks, the enterprise is using mainly its own capital, this requires the enterprise to must have a fairly strong own capital to meet the capital needs for production and business processes, to maintain capital for the production process, enterprises that have limited activities of granting commercial credit, managing Tightening the customer KPTs, minimizing the situation where customers take over the company's capital to ensure capital whenever necessary for the production process (Giannetti et al., 2011).

According to research papers Research by Yudha Prambudi (2021) and Ahmed Ibrahim Mohammed Al-Matari (2015), Continuous variable Currentliabilities convert to discrete variables assumes that the increase in receivables is negatively correlated with current liabilities. That is when receivables increase, current liabilities will decrease, and vice versa.

Convert the continuous Current liabilities to discrete variables to identify which companies have small, moderate, and large short-term debt. Effective management of current liabilities can increase receivables and increase a company's financial stability. We divide them into clusters as follows: (Unit: Billion VND)

1: $1 \leq x \leq 1000$

2: $1000 > x \leq 5000$

3: $5000 > x \leq 10000$

4: $x > 10000$

Hypothesis 3: **Revenue (GRO)** has the same/inverse effect on the receivables of enterprises.

Revenue growth: when the revenue in the period increases, it means that enterprises do business more efficiently, promoting the development of business and production processes, in this case, businesses tend to expand commercial credit aimed at continued revenue growth and profit maximization (Petersen & Rajan, 1997). On the contrary, when the company's revenue declines, the enterprise will increase the provision of credits to attract customers in order to increase sales volume and increase the value of customer receivables, Which contributes to the increase in revenue achieved

Research results of Tran Ai Ket (2017) show that a company with high revenue growth, the net commercial credit will be large. In contrast, Phan Dinh Nguyen and Truong Thi Hong Nhung (2014) confirm that there is a negative correlation between revenue growth and net trade credit.

2.2. Estimation method:

To test and evaluate the impact of factors on receivables of 100 companies listed on HNX in Vietnam, through estimation of regression model on panel data with the support of Rstudio software. At the same time, tests related to the reliability of the regression model with panel data are also performed appropriately such as the multicollinearity test through the variance exaggeration factor (VIF). and the variance of the error changes

PERFORMANCE RESEARCH

1. Selection of independent variables suitable for model research

```
# Create discrete variable currentliabilities
> currentliabilities_categories <- cut(df$currentliabilities,
+                                     breaks = c(1000000000, 1000000000000, 5000000000000, 10000000000000, Inf),
+ labels = c("Here 1 ty den 1000 ty", "Here 1000 ty day 5000 ty", "Here 5000 ty den 10000 ty", "Here 10000 ty tro len"))
# Create new vector
> new_col <- ifelse(df$currentliabilities >= 1000000000 & df$currentliabilities <= 1000000000000, 1,
+                   ifelse(df$currentliabilities > 1000000000000 & df$currentliabilities <= 5000000000000, 2,
+                                  ifelse(df$currentliabilities > 5000000000000 & df$currentliabilities <= 10000000000000, 3,
+                                  ifelse(df$currentliabilities > 10000000000000, 4, df$currentliabilities))))
# Assign the new vector to the corresponding column in df
> df$currentliabilities <- new_col
> View(df)
```

	firmcode	exchangename	ppe	currentliabilities	revenue	receivable	industry
1	PDB.HN	HANOI STOCK EXCHANGE	6.589725e+10	1	2.630341e+11	109030815090	Basic Materials
2	VC9.HN	HANOI STOCK EXCHANGE	2.770724e+10	2	2.098815e+11	602384133870	Industrials
3	SHN.HN	HANOI STOCK EXCHANGE	4.423497e+09	2	3.745725e+12	165298564230	Industrials
4	DTC.HN	HANOI STOCK EXCHANGE	1.686762e+11	1	2.741834e+11	-3742134340	Basic Materials
5	BAB.HN	HANOI STOCK EXCHANGE	1.744740e+11	1	3.642946e+11	45309102560	Financials
6	LUT.HN	HANOI STOCK EXCHANGE	4.328143e+10	1	7.916624e+10	90134379780	Industrials

Comment:

The results show that the continuous variable currentliabilities has turned into a discrete variable, this action will help the regression modelbecome more flexible and accurate in predicting dependent values on the basis of independent variables

2. Create Dataset

```
install.packages("readxl")

library(readxl)

install.packages("dplyr")

library(dplyr)
```

Comment:

Install the R software package called "readxl" to read and process Excel files and activate the installed "readxl" software package for use in the current session. Similar to "dplyr"

```
> View(X040522_Data_Mid_term_test_Final)
```

firmname	exchangename	totalasset	roa
1369 Construction JSC	HANOI STOCK EXCHANGE	8.987719e+11	0.0244983471
40 Investment and Construction JSC	HANOI STOCK EXCHANGE	1.934882e+11	0.0016107849
577 Investment Corp	HOCHIMINH STOCK EXCHANGE	4.490729e+12	0.0787585888
A Cuong Mineral Group JSC	HANOI STOCK EXCHANGE	5.335485e+11	-0.0246308975
ACC Binh Duong Investment and Construction JSC	HOCHIMINH STOCK EXCHANGE	1.175563e+12	0.0368648905
Additives and Petroleum Products JSC	HANOI STOCK EXCHANGE	1.003384e+11	0.0327380912
Agribank Securities Joint-Stock Corp	HOCHIMINH STOCK EXCHANGE	2.739269e+12	0.1552543916

Display the object's contents in a data table in the RStudio interface.

```
# Select the columns to get
> dataset <- select(X040522_Data_Mid_term_test_Final, firmcode,
exchangename, ppe, currentliabilities, revenue, receivable, industry)
# Filter values with exchangename by "HANOI STOCK EXCHANGE"
> data <- filter(dataset, exchangename == "HANOI STOCK EXCHANGE")
> View(data)
```

	firmcode	exchangename	ppe	currentliabilities	revenue	receivable	industry
1	V21.HN	HANOI STOCK EXCHANGE	6.740464e+10	1.888458e+11	1.255000e+11	64096912220	Industrials
2	LIG.HN	HANOI STOCK EXCHANGE	1.063767e+12	3.276834e+12	2.439778e+12	989997938420	Industrials
3	MCC.HN	HANOI STOCK EXCHANGE	2.730796e+09	2.119091e+10	4.095361e+10	6293668100	Basic Materials
4	TET.HN	HANOI STOCK EXCHANGE	3.772485e+10	5.298855e+09	2.259313e+10	1320000	Consumer Cyclicals
5	KSD.HN	HANOI STOCK EXCHANGE	5.655679e+10	1.167689e+10	6.882658e+10	41068616300	Consumer Non-Cyclicals

Comment:

Create a new data table name data by filtering out the rows in a dataset where the value of the column exchange name with "HANOI STOCK EXCHANGE".

```
# Set seed value to generate reproducible random results

> set.seed(939)

# Extract random sample with 100 observations

> df <- data[sample(nrow(data), 100, replace = FALSE), ]
```

Comment:

According to the reporting requirements, based on the last 3 digits of the Student Number (K204141**939**), generate a replicable random sample, then create a new data table name df by extracting a random sample of 100 rows from the data table data. replace = FALSE which means each row is extracted only once.

```
# Replace missing values with the corresponding variable's median

> df$ppe[is.na(df$ppe)] <- median(df$ppe, na.rm = TRUE)

> df$currentliabilities[is.na(df$currentliabilities)] <-
median(df$currentliabilities, na.rm = TRUE)

> df$receivable[is.na(df$receivable)] <- median(df$receivable, na.rm
= TRUE)

> df$revenue[is.na(df$ revenue)] <- median(df$revenue, na.rm = TRUE)

> View(df)
```

firmcode	exchangename	ppe	currentliabilities	revenue	receivable	industry
PDB.HN	HANOI STOCK EXCHANGE	6.589725e+10	8.080499e+10	2.630341e+11	109030815090	Basic Materials
VC9.HN	HANOI STOCK EXCHANGE	2.770724e+10	1.085833e+12	2.098815e+11	602384133870	Industrials
SHN.HN	HANOI STOCK EXCHANGE	4.423497e+09	1.900895e+12	3.745725e+12	165298564230	Industrials
DTC.HN	HANOI STOCK EXCHANGE	1.686762e+11	1.186064e+11	2.741834e+11	-3742134340	Basic Materials
BAB.HN	HANOI STOCK EXCHANGE	1.744740e+11	1.263576e+11	3.642946e+11	45309102560	Financials
LUT.HN	HANOI STOCK EXCHANGE	4.328143e+10	3.059100e+11	7.916624e+10	90134379780	Industrials

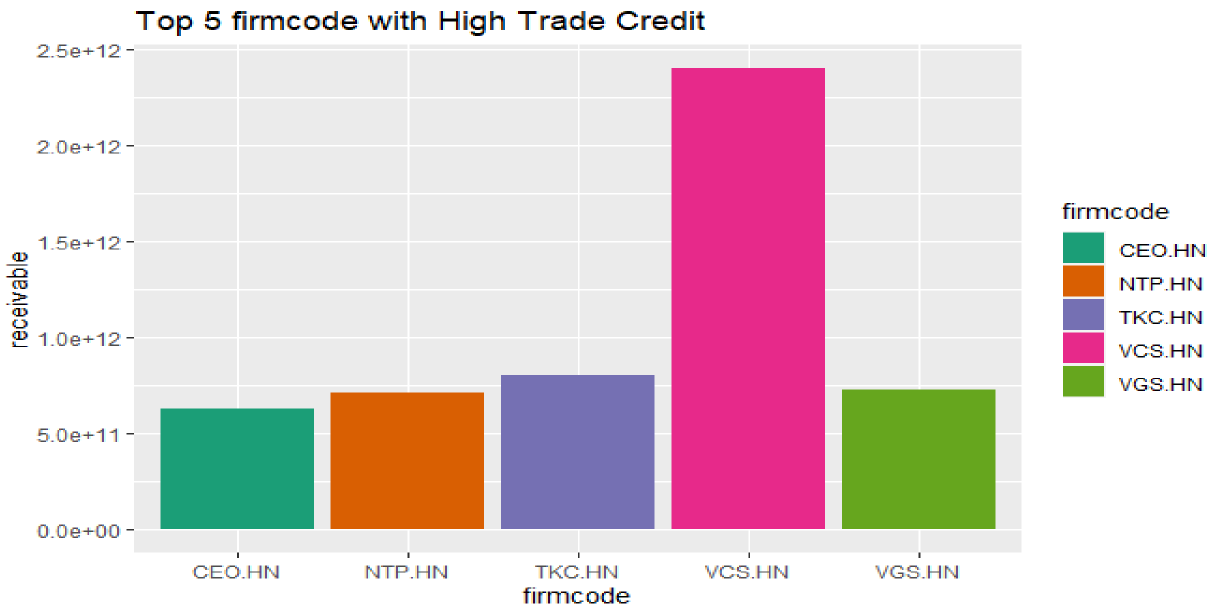
Comment:

Replace the missing values (NA) with the median of the corresponding variable help improve data quality and reliability.

3. Report

3.1. Five firms with highest trade credit

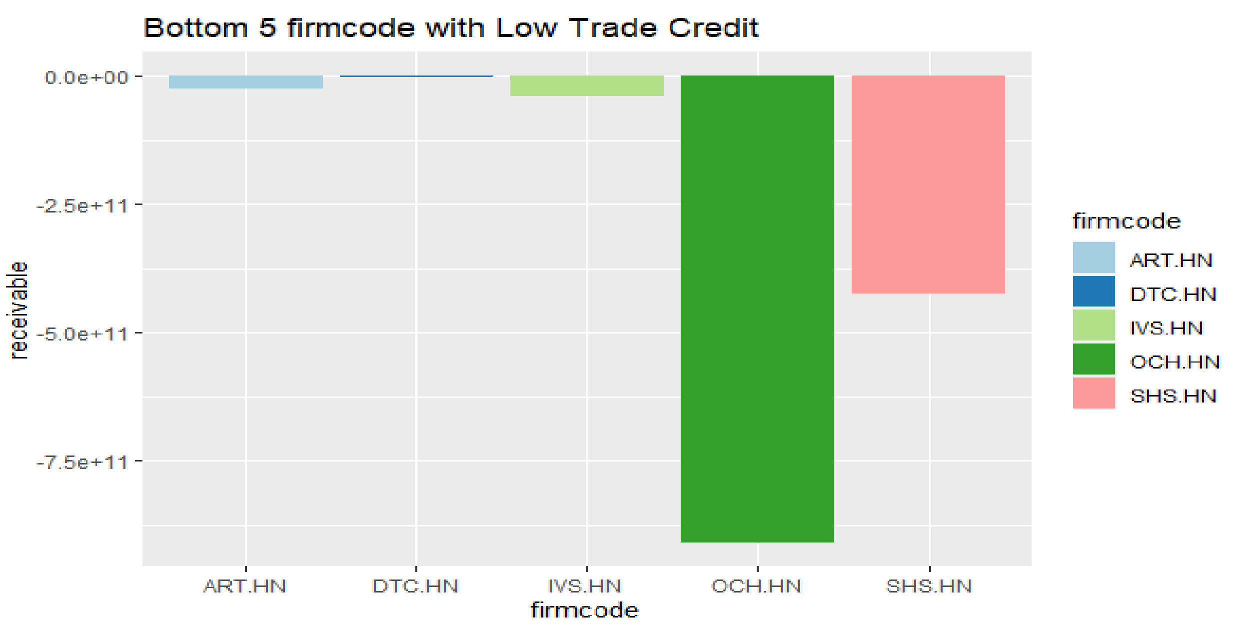
```
# Filter out the 5 companies with the highest commercial credit by the "receivable" objective variable

> top_5_firmcode <- head(df[order(-df$receivable), ], 5)

> view(top_5_firmcode)

# Plot a bar chart showing the value of the variable "receivable" for these companies

> library(ggplot2)

> library(RColorBrewer)

> ggplot(top_5_firmcode, aes(x = firmcode, y = receivable, fill = firmcode)) +

+  geom_bar(stat = "identity") +

+  scale_fill_brewer(palette = "Dark2") +

+  ggtitle("Top 5 firmcode with High Trade Credit") +

+  xlab("firmcode") +

+ ylab("receivable")
```



The results show that the company VCS (Pink) has the highest trade credit, while the CEO has the lowest

3.2. Five firms with lowest trade credit

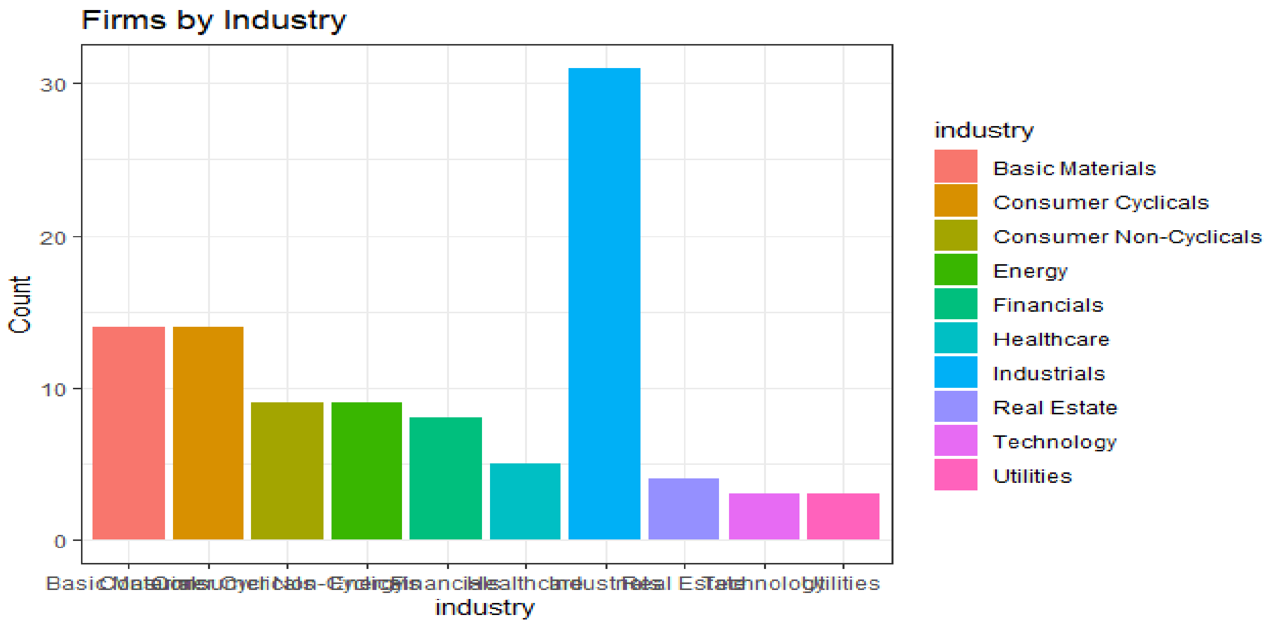
```
# Filter out the 5 companies with the lowest commercial credit by
item variable title "receivable"

> bottom_5_firmcode <- head(df[order(df$receivable), ], 5)

# Plot a bar chart showing the value of the variable "receivable" for
these companies

> ggplot(bottom_5_firmcode, aes(x = firmcode, y = receivable, fill =
firmcode)) +

+  geom_bar(stat = "identity") +

+  scale_fill_brewer(palette = "Paired") +

+  ggtitle("Bottom 5 firmcode with Low Trade Credit") +

+  xlab("firmcode") +

+ ylab("receivable")
```



The results show that OCH (Green) company has the lowest trade credit

3.3. Provide the name of industries which the firms belong to

```
# The name of industries which the firms belong to

> ggplot(df, aes(x = reorder(industry, -as.numeric(industry)), fill =
industry)) +

+ geom_bar() +

+  theme_bw() +

+  xlab("industry") +

+ ylab("Count") +

+  ggtitle("Firms by Industry")
```



Comment:

The results show that there are a total of 10 industries mentioned in the data, in which industrials is the field with the most companies operating, while businesses participating in the Utilities sector are the least.

3.4. Descriptive statistics

3.4.1. Different categories of the discrete variable

```
df %>%
+  group_by(currentliabilities) %>%
+  summarise(mean_lev=mean(receivable),
+            median_lev=median(receivable),
+            max_lev=max(receivable),
+            min_lev=min(receivable),
+            std_lev=ifelse(is.na(sd(receivable)), runif(1, 0, 1),
sd(receivable)))
# A tibble: 2 x 6
currentliabilities  mean_lev  median_lev  max_lev  min_lev   std_lev
 <dbl>                <dbl>      <dbl>     <dbl>     <dbl>     <dbl>
1            1 64275987410.  35570236875   5.04It is11  -9.10It is11
1.47It is11
2            2 567549974856. 533839822955  2.40It is12  -4.26It is11
6.18It is11
```

Comment:

The results show that the average of "receivable" for current liabilities of 1 is 64275987410 and for current liabilities of 2 is 567549974856. The median value of "receivable" for current liabilities of 1 is 35570236875 and for current liabilities of 2 is 533839822955. The maximum value of "receivable" for current liabilities of 1 is 5.04e11 coins and for current liabilities of 2 is 2.40e12 coins. The minimum value of "receivable" for current liabilities of 1 is -9.10e11 coins and for current liabilities of 2 is -4.26e11 coins. The standard deviation of "receivable" for current liabilities of 1 is 1.47e11 coins and for current liabilities of 2 is 6.18e11 coins.

However, the standard deviation of "receivable" is equal to NA (the standard deviation cannot be determined), the code above uses the run if (1, 0, 1) function to generate a random standard deviation value between 0 to 1. Therefore, this standard deviation value is not exact and may not accurately reflect the correctness of the data.

3.4.2. Groups of above/below the median of the continuous variable

```
# For continuous variables

# Revenue

> median(df$revenue)

[1] 364294626490

> summary_revenue <- df %>%

+  mutate(x= ifelse(df$revenue >= median(df$revenue), "above",
"below"))  %>%

+  group_by(x) %>%

+  summarise(mean_tradecredit=mean(receivable),

+             median_tradecredit=median(receivable),

+             max_tradecredit=max(receivable),

+             min_tradecredit=min(receivable),

+             std_tradecredit=sd(receivable))

> summary_revenue

> View(summary_revenue)
```

	x	mean_tradecredit	median_tradecredit	max_tradecredit	min_tradecredit	std_tradecredit
1	above	220149122615	139633586990	2.401753e+12	-910332864040	416395125587
2	below	45833251262	17944311100	6.023841e+11	-39635105470	92854651150

```
# ppe

> median(df$ppe)

[1] 61024466555

> summary_ppe <- df %>%

+  mutate(x= ifelse(df$ppe >= median(df$ppe), "above", "below"))  %>%

+  group_by(x) %>%

+  summarise(mean_tradecredit=mean(receivable),

+             median_tradecredit=median(receivable),
```

```
+              max_tradecredit=max(receivable),
+              min_tradecredit=min(receivable),
+              std_tradecredit=sd(receivable))
> summary_ppe
> View(summary_ppe)
```

	x	mean_tradecredit	median_tradecredit	max_tradecredit	min_tradecredit	std_tradecredit
1	above	198224843406	111421385760	2.401753e+12	-910332864040	403170242581
2	below	71243847899	24205719545	8.004590e+11	-425539669650	173474504034

```
# currentliabilities
> median(df$currentliabilities)
[1] 1
> summary_currentliabilities <- df %>%
+  mutate(x= ifelse(df$currentliabilities >=
median(df$currentliabilities), "above", "below"))  %>%
+  group_by(x) %>%
+  summarise(mean_tradecredit=mean(receivable),
+              median_tradecredit=median(receivable),
+              max_tradecredit=max(receivable),
+              min_tradecredit=min(receivable),
+              std_tradecredit=sd(receivable))
> summary_currentliabilities
> View(summary_currentliabilities)
```

	x	mean_tradecredit	median_tradecredit	max_tradecredit	min_tradecredit	std_tradecredit
1	above	134734345652	45309102560	2.401753e+12	-910332864040	315307242899

Comment:

The results show that descriptive statistics with mean, maximum, minimum, trade credit standard deviation of continuous variables

3.5. Do you have any comments on the possible link between the discrete and continuous variables with trade credit?

Based on the numbers from the above descriptive statistics, we can see the relationship between current liabilities, revenue, assets including plant and equipment (PPE), and trade credit receivables.

CTrade credit receivables can be affected by sales and PPE. As a business generates more revenue, it can extend more credit to its customers, which can lead to higher trade credit receivables. Similarly, if a business invests in more PPE, it can increase production capacity and generate more revenue, which can also lead to higher trade credit receivables.

Also, if a business has high levels of short-term debt, it may indicate that the business is relying heavily on trade credit to finance its operations. This can become a problem if businesses are unable to collect trade credit receivables in a timely manner or if their suppliers decide to tighten their credit policies.

3.6. Is it in line with your literature review?

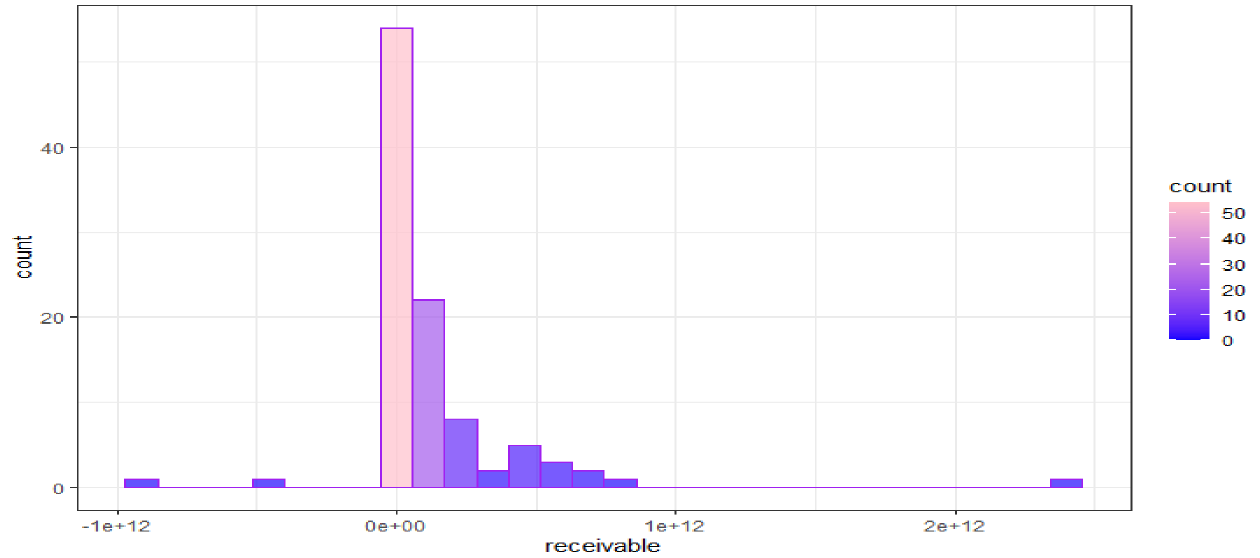
Based on the relevant evidence mentioned above, it is completely consistent with the theoretical basis laid out based on a lot of articles and research papers, especially typical clear views such as Santos & Silva (2014), Ahmed et al. (2014), Nadiri (1969), or Petersen & Rajan (1997)

4. Data visualization

```
#Install and download the necessary software packages for data analysis and visualization in CHEAP

install.packages("ggplot2")

install.packages("tidyverse")

install.packages("forcats")

install.packages("scales")

library(ggplot2) #for plotting

library(tidyverse) #for dataframe manipulation

library(forcats) #for handling factors

library(scales) #for axis scale formatting
```

4.1. Provide a histogram of trade credit

```
> ggplot(df, aes(x = receivable, fill = ..count..)) +
+  geom_histogram(color = "purple", alpha = 0.7) +
+  scale_fill_gradient(low = "blue", high = "pink") +
+  theme_bw()
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Comment:

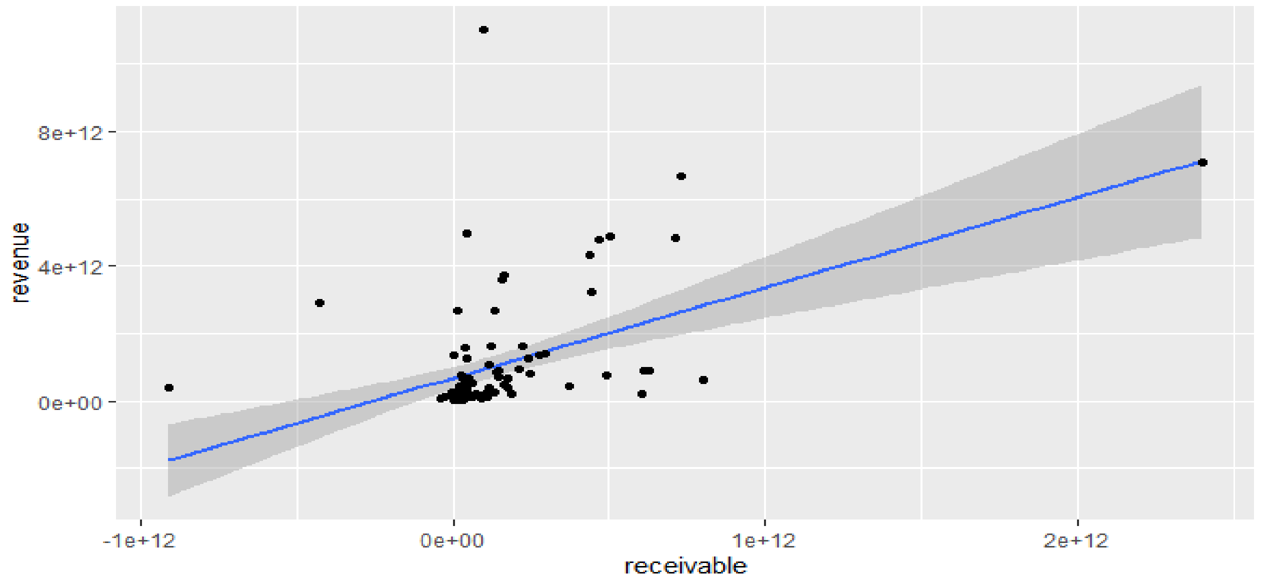
Between the colors in the column, there are 50 accounts receivable at level 0e+00, which means that the business has no significant debt from customers. This can have many reasons, such as a newly established business, not many customers, or a business that has fully settled debts in that accounting period.

However, in this case, the business needs to ensure that there is no loss in its business because receivables are an important part of the business's financial assets and income. This may mean that the business needs to find a way to recover money from existing customers or find new customers to keep its business afloat.

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

```
> ggplot(df, aes(x = receivable, y = revenue)) +
+  geom_smooth(method = "lm")+
```

```
+   geom_point()

`geom_smooth()` using formula = 'y ~ x'
```

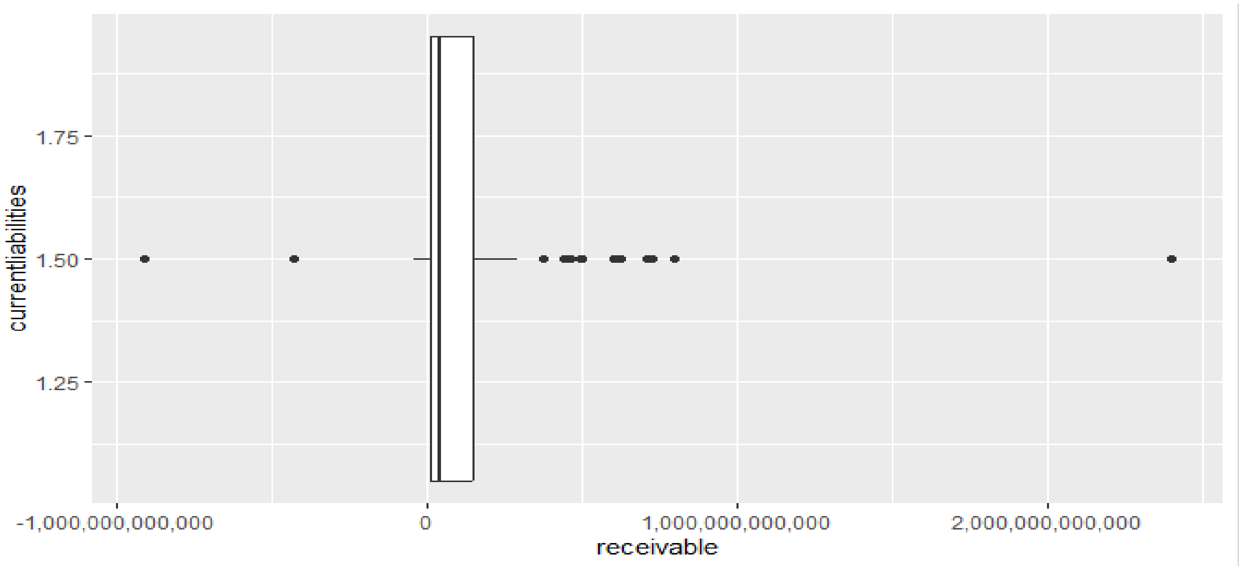


Comment:

The results show that two independent variables and dependent variables are highly correlated if revenue increases, receivables will also increase, according to the original theoretical basis.

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

```
> df %>%
+  filter(!is.na(currentliabilities), !is.na(receivable)) %>%
+  ggplot(aes(x = currentliabilities, y = receivable, fill =
currentliabilities)) +
+  geom_boxplot() +
+  coord_flip() +
+  scale_y_continuous(labels = scales::comma)
```



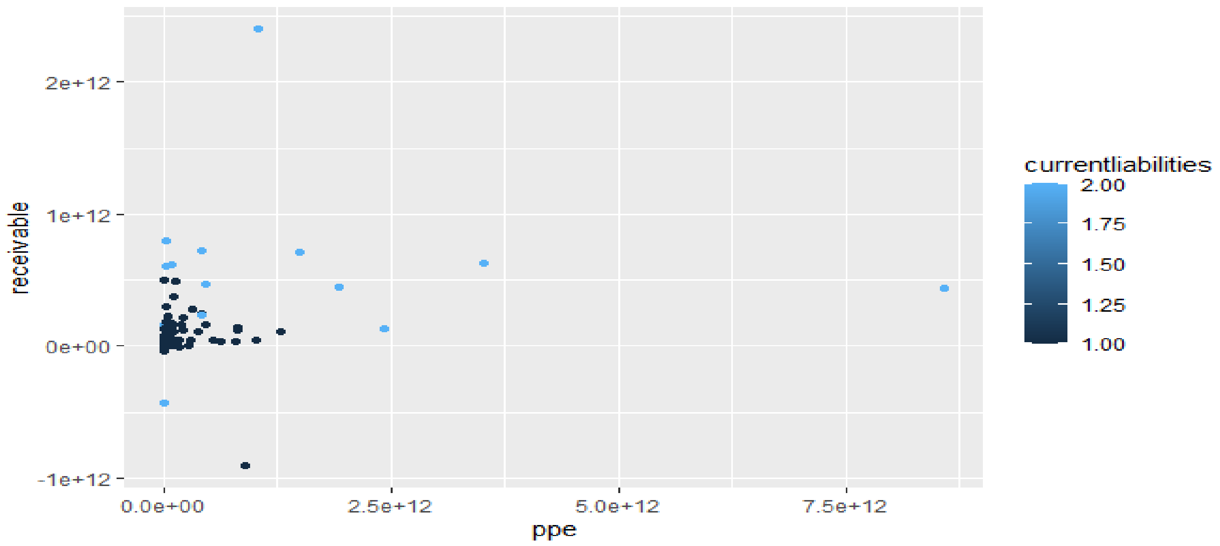
Comment:

In general, the results show that most of the receivables that are not high are concentrated in class 1 debts, the process does not fluctuate much. Where the right skewed graph shows Mode <Median<Mean, i.e. the data is asymmetric, the average short-term debt ratio is at 1.5

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

4.4.1. With Aesthetics

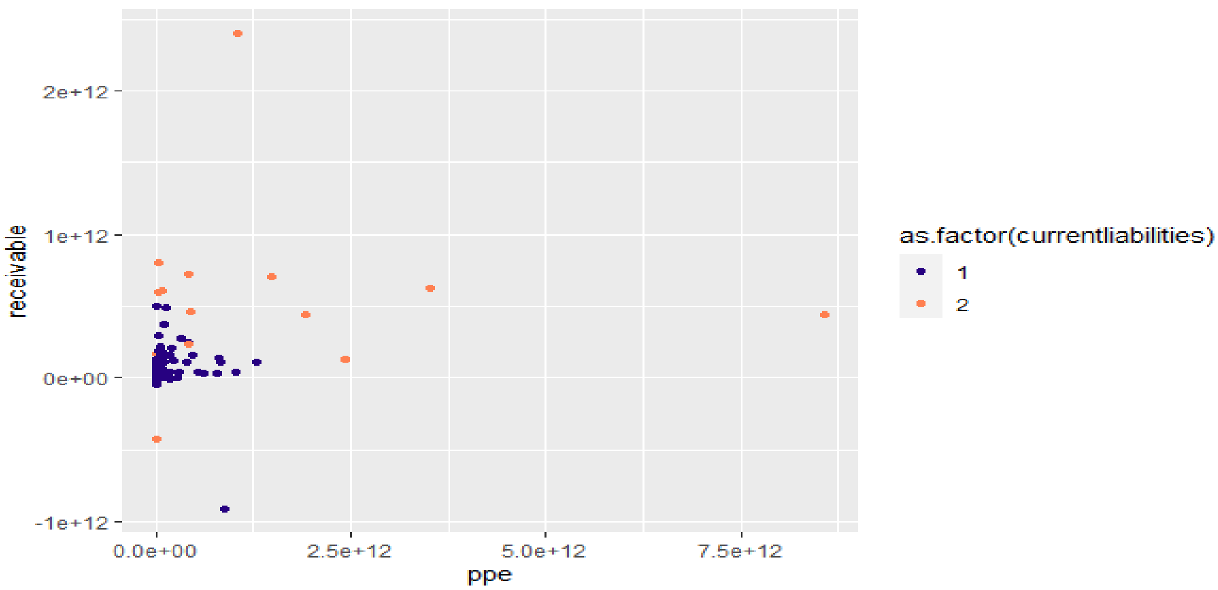
```
#With Aesthetics
#3 variables
> ggplot(df, aes(x = ppe, y = receivable, color = currentliabilities)) +
+    geom_jitter(width = .2)
```



Comment:

Since current liabilities are numeric, ggplot creates a legend with a continuous color scale. To change this, either 1) make female a character or factor vector, or 2) temporarily specify it as such when plotting.

```
> ggplot(df, aes(x = ppe, y = receivable, color =
as.factor(currentliabilities))) +
+  geom_point() +
+  scale_color_manual(values = c("#270181","coral"))
```

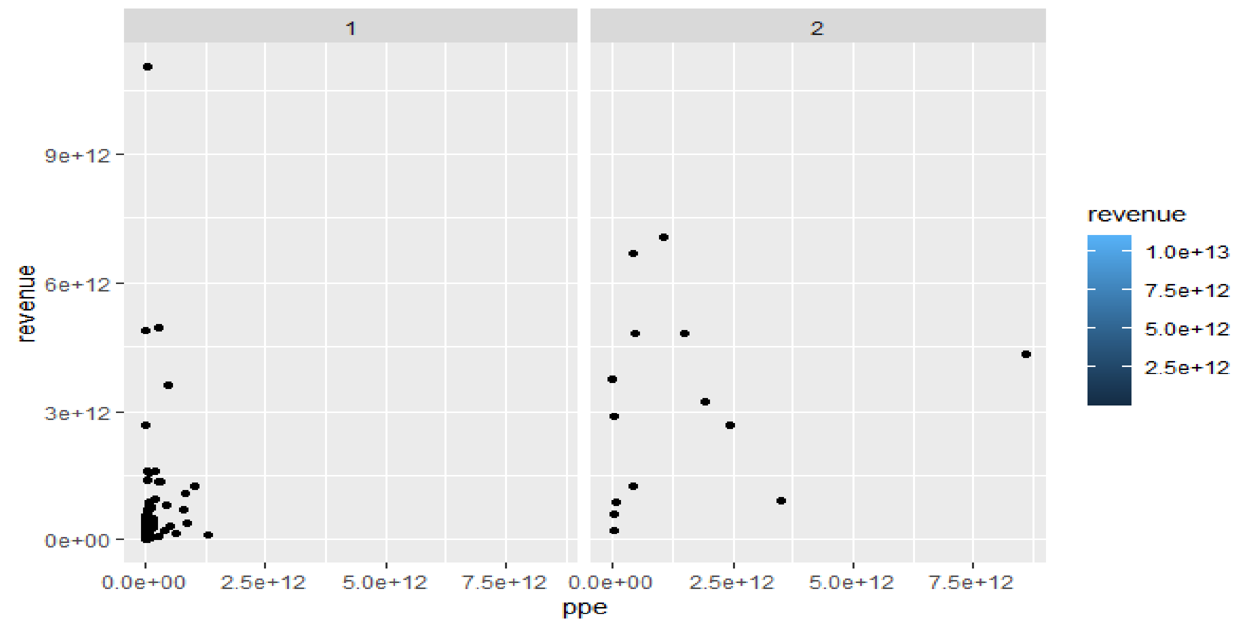


Comment:

The results show a short-term debt of type 1 (from 1 billion to 1000 billion) - purple, highly concentrated at the level of receivables is not large and at the same time, ppe is only 0. This shows all 3 variables are positively correlated with each other. Which, types of short-term debt of type 2 (from 1000 to 5000 billion) are very few and scattered, moreover all are in the position that receivable & ppe are higher than that of type 1. This shows that for short-term loans The higher the term, the higher they are in terms of receivables and fixed assets.

4.4.2. With Facets

```
# With Facets
> ggplot(df, aes(x = ppe, y = revenue, fill = revenue)) +
+   geom_point() +
+   facet_wrap(~ currentliabilities) #try ncol=, nrow=
```



Comment:

Facets help to visualize categorical variables with many categories. Facets split our plot into several smaller plots along a categorical variable.

5. Regression

Load the necessary R packages

```
library(tidyverse)

library(ggplot2)
```

5.1. Multiple regression

a. Use cor() to test the relationship between independent variables

```
> cor(df$revenue, df$ppe)

[1] 0.2853752

> cor(df$revenue, df$currentliabilities)

[1] 0.4827275

> cor(df$ppe, df$currentliabilities)

[1] 0.463865
```

Comment:

Use cor() to test the relationship between independent variables shows that they aren't highly correlated, so we can include both variables in the same model.

The correlation between sales and fixed assets is 0.285, indicating that a change in the value of fixed assets is not synonymous with a similar change in a company's revenue. (Assuming a value between -0.3 and 0.3 indicates a weak or no correlation).

The correlation between sales and short-term payables is 0.483, indicating a stronger relationship between sales and short-term liabilities than fixed assets. However, the correlation between these two variables is not high, just average. (Assuming values between 0.3 and 0.7 show a moderate correlation)

The correlation between fixed assets and short-term payables is 0.464, indicating a significant correlation between these two variables, but this level is still lower than the relationship between sales and payables. Short-term.

b. Use the variance Inflation Factor test for more formal testing of multicollinearity (from "car" package).

```
install.packages("stargazer")

library(stargazer)

install.packages("car")

library(car)
```

```
> summary(tradecredit.lm<-lm(receivable ~ ppe +
revenue+currentliabilities, data = df))


Call:

lm(formula = receivable ~ ppe + revenue + currentliabilities, data =
df)

Residuals:

       Min           1Q      Median           3Q          Max

-1.003e+12 -3.395e+10 -1.715e+10 6.439e+10 1.637e+12

Coefficients:

                      Estimate  Std. Error  t value Pr(>|t|)

(Intercept)         -3.706e+11  9.745e+10  -3.803 0.000251 ***

ppe                 -1.553e-02  2.947e-02  -0.527 0.599531

revenue              4.877e-02  1.649e-02   2.958 0.003898 **

currentliabilities  4.036e+11  9.091e+10   4.439 2.41e-05 ***

---

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 2.545e+11 on 96 degrees of freedom

Multiple R-squared:  0.368,Adjusted R-squared:  0.3483

F-statistic: 18.63 on 3 and 96 DF,  p-value: 1.321e-09
```

Comment:

The estimated coefficients (Estimate) show that for the variable ppe, there is no significant relationship with the value of receivables since the p-value of ppe is 0.599531 > 0.05. Meanwhile, the coefficients for the revenue and current liabilities variables both have p-value < 0.05, showing that they have a significant relationship with the value of receivables.

R-squared is used to evaluate the explanatory strength of the model, and the R-squared value of this model is 0.368, that is, the model explains 36.8% of the variation in the value of the receivables. The F-statistic is 18.63, with a very low p-value of 1.321e-09, indicating that this linear regression model has high accuracy.

```
> stargazer(tradecredit.lm,type ="text")

                          Dependent variable:

                      -----------------------------

                              receivable

-------------------------------------------------

ppe                             -0.016

                                (0.029)

revenue                        0.049***

                                (0.016)

currentliabilities      403,583,791,560.000***

                          (90,907,827,498.000)

Constant                -370,624,748,274.000***

                          (97,454,400,587.000)

-------------------------------------------------

Observations                     100

R2 0.368

Adjusted R2                     0.348

Residual Std. Error   254,549,219,383.000 (df = 96)
```

```
F Statistic              18.634*** (df = 3; 96)
=================================================
Note:                  *p<0.1; **p<0.05; ***p<0.01
```

Comment:

In this case, R^2 = 0.368, that is, 36.8% of the variation of the dependent variable can be explained by the independent variables in the model. However, because the number of independent variables in the model is not much, we should pay more attention to the Adjusted R^2 index. Adjusted R^2 adds an adjustment to the number of independent variables in the model to avoid overfitting. In this case, Adjusted R^2 = 0.348, lower than R^2, but still acceptable.

The results show that revenue is significantly positively correlated with receivables at a 1% level (p < 0.01). This means that as sales increase, so do receivables.

Other variables, including ppe and current liabilities, were not significantly correlated with receivables. The model is valid with significant p-values and the F-statistic shows that the model is built better than a model with only one constant.

```
> car::vif(tradecredit.lm)
             ppe           revenue      currentliabilities
        1.282207          1.312057            1.535626
```

Comment:

The results show that the input variables in your model (tradecredit.lm) do not have a serious multicollinearity problem, because the VIF values of all variables are within a safe range from 1 to 2.

5.2. Estimate model and interpret results with ppe, revenue và current Liabilities

```
> tradecredit.lm<-lm(receivable ~ ppe + revenue + currentliabilities,
data = df)
> summary(tradecredit.lm)

Call:
lm(formula = receivable ~ ppe + revenue + currentliabilities,
    data = df)
```

```
Residuals:
       Min         1Q     Median         3Q        Max
-1.003e+12 -3.395e+10 -1.715e+10 6.439e+10 1.637e+12

Coefficients:
                    Estimate  Std. Error t value Pr(>|t|)
(Intercept)       -3.706e+11  9.745e+10  -3.803 0.000251 ***
ppe               -1.553e-02  2.947e-02  -0.527 0.599531
revenue            4.877e-02  1.649e-02   2.958 0.003898 **
currentliabilities  4.036e+11  9.091e+10   4.439 2.41e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.545e+11 on 96 degrees of freedom
Multiple R-squared:  0.368,Adjusted R-squared:  0.3483
F-statistic: 18.63 on 3 and 96 DF,  p-value: 1.321e-09
```

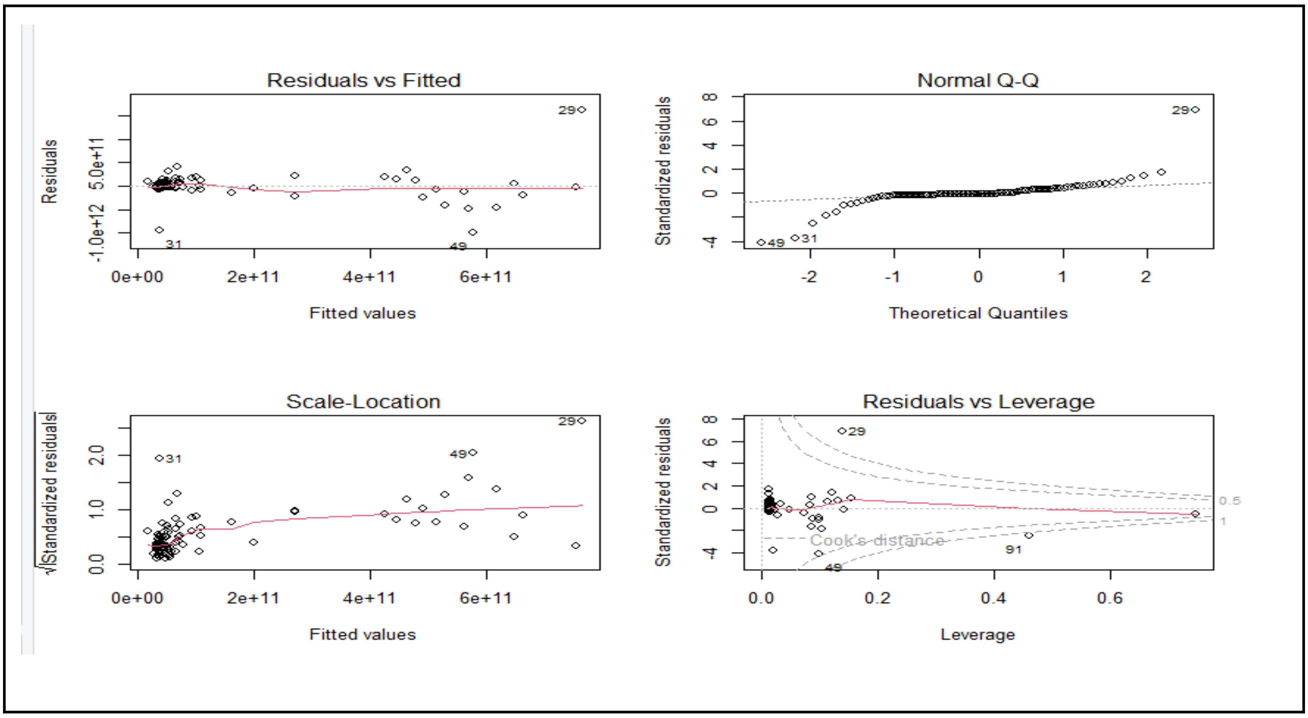
Comment:

The linear regression model shows that the variable ppe is not statistically significant in this model (p-value is greater than 0.05), while the revenue and current liabilities variables are statistically significant because of the p-value. Their values are all less than statistical significance at the 0.05 level.

The model's F-statistic is 18.63, with a p-value of 1.321e-09, indicating that the overall model is statistically significant.

5.3. Check important assumptions for linear regression

```
> par(mfrow=c(2,2))
> plot(tradecredit.lm)
```



Analyze residuals to test hypothetical analyzes in linear regression

Comment:

After estimating, The first left plot (a) shows that the residuals are centered around the y=0 line, or e1 has a mean of 0 so this assumption is acceptable. The first right graph (b) plots the residuals and the expected values, we see that the residuals are concentrated very close to the values on the standard curve, so we assume (b), i.e. e1 is distributed according to the law. normally distributed, can also satisfy. The last left plot (c) plots the root of the standard remainder and the value of y, showing that there is no difference between the standard residuals for values of y1, so assume (d) the graph the last right, i.e. e1 with fixed variance for all x1, can also satisfy.

6. Using LOOP

6.1 Count the number of firms in an industry

```
# Check important assumptions for linear regression
> par(mfrow=c(2,2))
> plot(tradecredit.lm)

# Specify the industry name to count
> industry_name <- "Consumer Cyclicals"

# Initialize a variable to count the number of firms in the industry
> count <- 0

# LOOP through each element of the industry variable
> for (i in df$industry) {

# If the element is the same as the specified industry name, add 1 to
the count variable
+  if (i == industry_name) {
+    count <- count + 1}}

# Print the count of firms in the specified industry
> cat("Number of firms in", industry_name, ":", count)

Number of firms in Consumer Cyclicals : 14
```

The results show that the number of companies operating in the Consumer Cyclicals field is 14

6.2 Count the number of firms in an industry and with trade credit above a certain value

```
> unique(df$industry)
[1] "Basic Materials"          "Industrials"    "Financials"
[4] "Technology"               "Utilities"      "Consumer Cyclicals"
[7] "Consumer Non-Cyclicals"   "Real Estate"    "Energy"
[10]"Healthcare"
> unique(df$receivable)
 [1]  109030815090  602384133870  165298564230   -3742134340   45309102560   90134379780
 [7]   30119317200   88702365450     -50000000  109621304680  132799080170   99963168510
[13]   44299845420   28444434450    1959295210    1467059900  162119316880   22012503530
[19]   30539831300  116462437320    2474387590    3405457290   25894257760   37353364860
[25]  612656597950   46318359700  465295512040   17944311100 2401752857940  727950604420
[31] -910332864040   38211616560  212549205150    3173747640   31154608890   25125281450
[37]  -26051486480    3977340690   46703631920   14123501750   16715358930   29462474397
[43]   64096912220   23279790480  -39635105470    5334684140  494298728270   34871203120
[49] -425539669650  109245482860  174369974870   24997531820  503555626500   36269270630
[55]   14730575450  159694936140  131457374170   20856622670  800459015300  439209127150
[61]   11815692920  139633586990  144755406540    7113253460    -120272490    6801942850
[67]  174985973010    5828423960   18715013300    5960358220  143767657820   30004732970
[73]  710167343110   42174810110  113221466840   21510120590  116461432270  446741702360
[79]    1906595750   22806294530   77094945240   51709072530  247188785770   12441632260
[85]   47695094490   38456155980  241359330290  186011398750   60742838520   98787646690
[91]  278879311600   11882663130   47515617950  626507154800   14905648990  121112444270
[97]       1320000   23286157640    2669109680
```

```
> industry <- c("Basic Materials", "Utilities",
"Technology","Industrials", "Financials", "Consumer Cyclicals",
"Consumer Non-Cyclicals", "Real Estate", "Energy", "Healthcare")

> receivable <- df$receivable
```

Aggregate results of industries and receivables values included in the data

```
# Set industry name and receivable threshold
> industry_name <- "Consumer Cyclicals"
> receivable_value <- 77094945240
# Initialize count variable
> count <- 0
# Loop through each element of the industry vector
> for (i in 1:length(industry)) {
+
# Check if industry and receivable value meet conditions
+      if    (industry[i]    ==    industry_name    &&    receivable[i]    >
receivable_value) {
+
# If conditions are met, increment count
+     count <- count + 1
+   }
+ }
# Print result
>  cat(paste("Number  of  firms  in",  industry_name,  "with  receivable
above", receivable_value, "is", count))

Number   of   firms   in   Consumer   Cyclicals   with   receivable   above
77094945240 is 1
```

Number of Consumer Cyclicals companies with higher accounts receivable77094945240đồng is 1

REFERENCES

https://thitruongtaichinhtiente.vn/nhan-to-tac-dong-toi-chinh-sach-tin-dung-thuong-mai-cua-doanh-nghiep-22901.html

https://tapchitaichinh.vn/nang-cao-hieu-qua-kinh-doanh-cua-doanh-nghiep-vien-thong-niem-yet-tren-thi-truong-chung-khoan.html

https://m.hvtc.edu.vn/Portals/0/CFO/1.TTTV_TAN.pdf

http://213.226.126.9/fc/2017/fc21/fc2117-1258.pdf

https://www.researchpublish.com/upload/book/FACTORS%20AFFECTING%20ACCOUNTS-6702.pdf

https://tapchitaichinh.vn/tin-dung-thuong-mai-va-hieu-qua-hoat-dong-cua-cac-cong-ty-niem-yet-o-viet-nam.html

https://tapchicongthuong.vn/bai-viet/cac-yeu-to-anh-huong-toi-tin-dung-thuong-mai-cua-doanh-nghiep-nganh-giao-thong-van-tai-niem-yet-o-hose-va-hnx-51103.htm4

https://gs.ctu.edu.vn/96LD4AMFnN8YuCGGi3zo3lsy6ClShi9EjxGD9mUfsYGUb0rT2DKE7cCedZehzbBmxPzdkurNqCoeU9Y2hgg4-AVCA6tAr-qy

ine:///C:/Nguy%E1%BB%85n%20Thanh%20Ph%C6%B0%C6%A1ng%20Uy%C3%AAn_K204141939/R%20CODE/The-Determinants-of-Trade-Credit_Vietnam-Experience.pdf

file:///C:/Users/Admin/Documents/48527-Article%20Text-152294-1-10-20200617.pdf

APPENDIX

```
install.packages("readxl")

library(readxl)

install.packages("dplyr")

library(dplyr)

View(X040522_Data_Mid_term_test_Final)

# Select the columns to get

dataset <- select(X040522_Data_Mid_term_test_Final, firmcode, exchangename, ppe, currentliabilities, revenue, receivable, industry)

# Filter values with exchangename by "HANOI STOCK EXCHANGE"

data <- filter(dataset, exchangename == "HANOI STOCK EXCHANGE")

View(data)

# Set seed value to generate reproducible random results

set.seed(939)

# Extract random sample with 100 observations

df <- data[sample(nrow(data), 100, replace = FALSE), ]

# Replace missing values with the corresponding variable's median

df$ppe[is.na(df$ppe)] <- median(df$ppe, na.rm = TRUE)

df$currentliabilities[is.na(df$currentliabilities)] <- median(df$currentliabilities, na.rm = TRUE)

df$receivable[is.na(df$receivable)] <- median(df$receivable, na.rm = TRUE)

df$revenue[is.na(df$ revenue)] <- median(df$revenue, na.rm = TRUE)

View(df)

# Create discrete variable currentliabilities

currentliabilities_categories <- cut(df$currentliabilities,

                        breaks = c(1000000000, 1000000000000, 5000000000000, 10000000000000, Inf),

                        labels = c("Here 1 ty den 1000 ty", "Here 1000 ty day 5000 ty", "Here 5000 ty den 10000 ty", "Here 10000 ty tro len"))

# Create new vector

new_col <- ifelse(df$currentliabilities >= 1000000000 & df$currentliabilities <= 1000000000000, 1,

          ifelse(df$currentliabilities > 1000000000000 & df$currentliabilities <= 5000000000000, 2,
```

```
        ifelse(df$currentliabilities > 5000000000000 & df$currentliabilities <= 10000000000000, 3,
        ifelse(df$currentliabilities > 10000000000000, 4, df$currentliabilities))))

# Assign the new vector to the corresponding column in df
df$currentliabilities <- new_col
View(df)
# Print the categories
currentliabilities_categories
df %>%
transmute(currentliabilities_categories, currentliabilities)
view(df)
# Filter out the 5 companies with the highest commercial credit by the "receivable" objective variable
top_5_firmcode <- head(df[order(-df$receivable), ], 5)
view(top_5_firmcode)
# Plot a bar chart showing the value of the variable "receivable" for these companies
library(ggplot2)
library(RColorBrewer)
ggplot(top_5_firmcode, aes(x = firmcode, y = receivable, fill = firmcode)) +
      geom_bar(stat = "identity") +
      scale_fill_brewer(palette = "Dark2") +
      ggtitle("Top 5 firmcode with High Trade Credit") +
      xlab("firmcode") +
      ylab("receivable")
# Filter out the 5 companies with the lowest commercial credit by the "receivable" target variable
bottom_5_firmcode <- head(df[order(df$receivable), ], 5)
# Plot a bar chart showing the value of the variable "receivable" for these companies
ggplot(bottom_5_firmcode, aes(x = firmcode, y = receivable, fill = firmcode)) +
      geom_bar(stat = "identity") +
      scale_fill_brewer(palette = "Paired") +
      ggtitle("Bottom 5 firmcode with Low Trade Credit") +
      xlab("firmcode") +
```

```
        ylab("receivable")

# The name of industries which the firms belong to
ggplot(df, aes(x = reorder(industry, -as.numeric(industry)), fill = industry)) +
        geom_bar() +
        theme_bw() +
        xlab("industry") +
        ylab("Count") +
        ggtitle("Firms by Industry")
# Calculate descriptive statistics for the currentliabilities variable by group currentliabilities_categories
currentliabilities_summary <- aggregate(df$currentliabilities, by=list(currentliabilities_categories), FUN=function(x) c(mean=mean(x), median=median(x), max=max(x), min=min(x), sd=sd(x)))
view(currentliabilities_summary)
# Calculate descriptive statistics for the receivable variable
receivable_mean <- mean(df$receivable)
receivable_median <- median(df$receivable)
receivable_max <- max(df$receivable)
receivable_min <- min(df$receivable)
receivable_sd <- sd(df$receivable)
# Calculate the descriptive statistics for the continuous variable ppe
ppe_mean <- mean(df$ppe)
ppe_median <- median(df$ppe)
ppe_max <- max(df$ppe)
ppe_min <- min(df$ppe)
ppe_sd <- sd(df$ppe)
# Calculate descriptive statistics for continuous variable revenue
revenue_mean <- mean(df$revenue)
revenue_median <- median(df$revenue)
revenue_max <- max(df$revenue)
revenue_min <- min(df$revenue)
```

```
revenue_sd <- sd(df$revenue)

# For continuous variables
# Revenue
median(df$revenue)
summary_revenue <- df %>%
mutate(x= ifelse(df$revenue >= median(df$revenue), "above", "below"))  %>%
group_by(x) %>%
        summarise(mean_tradecredit=mean(receivable),
        median_tradecredit=median(receivable),
        max_tradecredit=max(receivable),
        min_tradecredit=min(receivable),
        std_tradecredit=sd(receivable))
summary_revenue
View(summary_revenue)
# Ppe
median(df$ppe)
summary_ppe <- df %>%
mutate(x= ifelse(df$ppe >= median(df$ppe), "above", "below"))  %>%
group_by(x) %>%
        summarise(mean_tradecredit=mean(receivable),
        median_tradecredit=median(receivable),
        max_tradecredit=max(receivable),
        min_tradecredit=min(receivable),
        std_tradecredit=sd(receivable))
summary_ppe
View(summary_ppe)
# Currentliabilities
median(df$currentliabilities)
summary_currentliabilities <- df %>%
```

```
mutate(x= ifelse(df$currentliabilities >= median(df$currentliabilities), "above", "below"))  %>%
group_by(x) %>%
summarise(mean_tradecredit=mean(receivable),
          median_tradecredit=median(receivable),
          max_tradecredit=max(receivable),
          min_tradecredit=min(receivable),
          std_tradecredit=sd(receivable))
summary_currentliabilities
View(summary_currentliabilities)
# Descriptive statistics
# For discrete variable
df %>%
group_by(currentliabilities) %>%
summarise(mean_lev=mean(receivable),
          median_lev=median(receivable),
          min_lev=min(receivable),
          std_lev=ifelse(is.na(sd(receivable)), runif(1, 0, 1), sd(receivable)))
df %>%
group_by(currentliabilities) %>%
summarise(mean_lev=mean(receivable),
          median_lev=median(receivable),
          max_lev=max(receivable),
          min_lev=min(receivable),
# Data visualization
install.packages("ggplot2")
install.packages("tidyverse")
install.packages("forcats")
install.packages("scales")
library(ggplot2) #for plotting
library(tidyverse) #for dataframe manipulation
```

```
library(forcats) #for handling factors

library(scales) #for axis scale formatting

# Provide histogram of trade credit

ggplot(df, aes(x = receivable, fill = ..count..)) +

        geom_histogram(color = "purple", alpha = 0.7) +

        scale_fill_gradient(low = "blue", high = "pink") +

        theme_bw()

# Provide scatter plot of trade credit with the continuous variable

ggplot(df, aes(x = receivable, y = revenue)) +

        geom_smooth(method = "lm")+

        geom_point()

# Provide boxplot of trade credit with the discrete variable (different colour for different categories of discrete variable)

df %>%

filter(!is.na(currentliabilities), !is.na(receivable)) %>%

ggplot(aes(x = currentliabilities, y = receivable, fill = currentliabilities)) +

        geom_boxplot() +

        coord_flip() +

        scale_y_continuous(labels = scales::comma)

        scale_fill_manual(values = c("pink", "blue"))

# Provide a plot that allow the combination of continuous, discrete variables and trade credit

#With Aesthetics

ggplot(df, aes(x = ppe, y = receivable, color = currentliabilities)) + #3 variables

        geom_jitter(width = .2)

ggplot(df, aes(x = ppe, y = receivable, color = as.factor(currentliabilities))) +

        geom_point() +

scale_color_manual(values = c("#270181","coral"))

# With Facets

ggplot(df, aes(x = ppe, y = revenue, fill = revenue)) +

        geom_point() +

        facet_wrap(~ currentliabilities) #try ncol=, nrow=
```

```r
        scale_color_manual(values = c("#270181","coral"))



# LINEAR REGRESSION

# Load packages

library(tidyverse)

library(ggplot2)

# Multiple regression (df)

cor(df$revenue, df$ppe)

cor(df$revenue, df$currentliabilities)

cor(df$ppe, df$currentliabilities)

# Use Variance Inflation Factor test for more formal testing of multicollinearity (from “car” package).

install.packages("stargazer")

library(stargazer)

install.packages("car")

library(car)

summary(tradecredit.lm<-lm(receivable ~ ppe + revenue + currentliabilities, data = df))

stargazer(tradecredit.lm,type ="text")

car::vif(tradecredit.lm)

# Estimate model and interpret results with  PPE, revenue và Current Liabilities

tradecredit.lm<-lm(receivable ~ ppe + revenue + currentliabilities, data = df)

summary(tradecredit.lm)

# Check important assumptions for linear regression

par(mfrow=c(2,2))

plot(tradecredit.lm)

#Loop

#Count the number of firms in an industry

#Count the number of firms in Financials

# Specify the industry name to count

industry_name <- "Consumer Cyclicals"
```

```
# Initialize a variable to count the number of firms in the industry

count <- 0

# LOOP through each element of the industry variable

for (i in df$industry) {

# If the element is the same as the specified industry name, add 1 to the count variable

if (i == industry_name) {

        count <- count + 1}}

# Print the count of firms in the specified industry

cat("Number of firms in", industry_name, ":", count)

# Count the number of firms in an industry and with trade credit above a certain value

unique(df$industry)

unique(df$receivable)

industry <- c("Basic Materials", "Utilities", "Technology", "Industrials", "Financials", "Consumer Cyclicals", "Consumer Non-Cyclicals", "Real Estate", "Energy", "Healthcare")

receivable <- df$receivable

# Set industry name and receivable threshold

industry_name <- "Consumer Cyclicals"

receivable_value <- 77094945240

# Initialize count variable

count <- 0

# Loop through each element of the industry vector

for (i in 1:length(industry)) {

# Check if industry and receivable value meet conditions

if (industry[i] == industry_name && receivable[i] > receivable_value) {

# If conditions are met, increment count

count <- count + 1

 }

# Print result

cat(paste("Number of firms in", industry_name, "with receivable above", receivable_value, "is", count))
```