

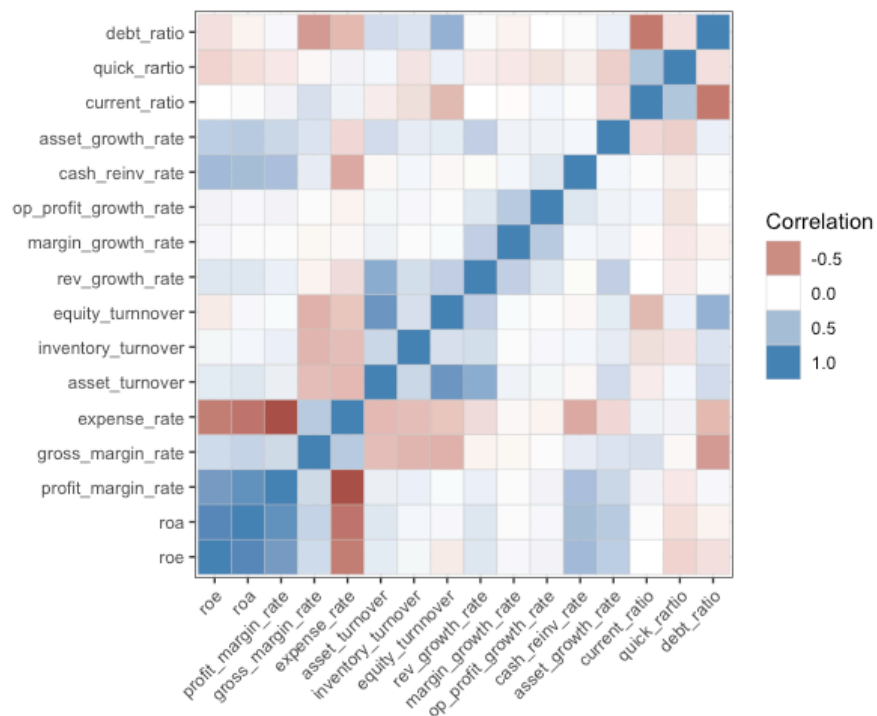
商業分析 HW4
105305072 企管四 許惠甄

1. 資料探索

a. 變數間的 Correlation

	Var1	Var2	value
1	roe	roe	1
2	roa	roe	0.927613722533678
3	profit_margin_rate	roe	0.769836624279952
4	gross_margin_rate	roe	0.263933319779695
5	expense_rate	roe	-0.563964269573648
6	asset_turnover	roe	0.147948736855924
7	inventory_turnover	roe	0.0609948258984017
8	equity_turnnover	roe	-0.0846514271433868
9	rev_growth_rate	roe	0.178890323350809
10	margin_growth_rate	roe	0.0423719956373313
11	op_profit_growth_rate	roe	0.070655719737964
12	cash_reinv_rate	roe	0.518422290791974
13	asset_growth_rate	roe	0.366727395586673
14	current_ratio	roe	0.00161885252532689
15	quick_rartio	roe	-0.178525536129484
16	debt_ratio	roe	-0.125853365081838
17	roe	roa	0.927613722533678
18	roa	roa	1
19	profit_margin_rate	roa	0.849678033161943
20	gross_margin_rate	roa	0.316508191821517
21	expense_rate	roa	-0.602167492346144
22	asset_turnover	roa	0.166463783487875
23	inventory_turnover	roa	0.0498719959855883
24	equity_turnnover	roa	0.0386910904570714
25	rev_growth_rate	roa	0.182344012908911
26	margin_growth_rate	roa	0.0252565080268641
27	op_profit_growth_rate	roa	0.0481880240850674
28	cash_reinv_rate	roa	0.504324078456767
29	asset_growth_rate	roa	0.388442882343791
30	current_ratio	roa	0.010373006903173
31	quick_rartio	roa	-0.133926962915792
32	debt_ratio	roa	-0.0535872922037833

b. Heatmap :



➤ 結論：

➤ 高度正相關變數：

- ✓ roe&roa
- ✓ profit_margin_rate&roa
- ✓ profit_margin_rate&roe
- ✓ asset_turnover&equity_turnover
- ✓ debt_turnover&equity_turnover

➤ 高度負相關變數：

- ✓ expense_rate&profit_margin_rate
- ✓ roa&expense_rate
- ✓ current_ratio&debt_ratio
- ✓ roe&expense_rate
- ✓ debt_ratio&gross_margin_rate

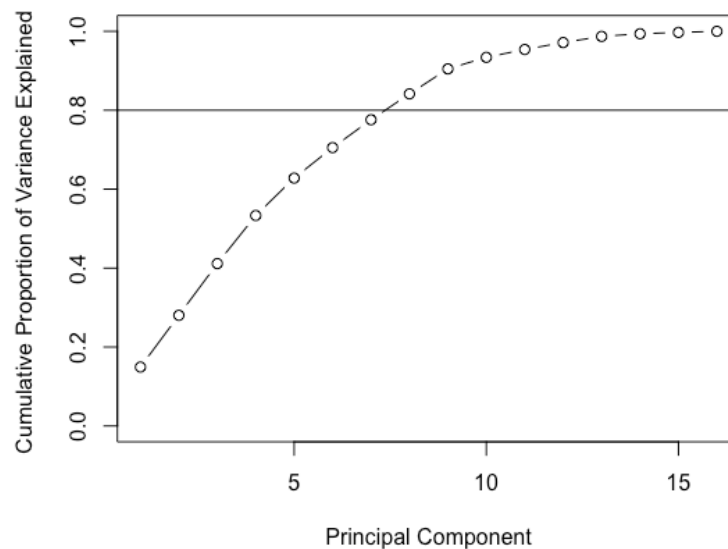
2. 以 SPCA 分析：

a. 每個主成份解釋多少變異：

Importance of components:											
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Standard deviation	1.0621	0.9953	0.9949	0.9595	0.84638	0.76393	0.7304	0.70373	0.69153	0.47041	0.38745
Proportion of Variance	0.1493	0.1311	0.1310	0.1218	0.09481	0.07724	0.0706	0.06555	0.06329	0.02929	0.01987
Cumulative Proportion	0.1493	0.2804	0.4114	0.5333	0.62808	0.70532	0.7759	0.84147	0.90476	0.93405	0.95392
	PC12	PC13	PC14	PC15	PC16						
Standard deviation	0.36419	0.34324	0.22416	0.16236	0.14516						
Proportion of Variance	0.01756	0.01559	0.00665	0.00349	0.00279						
Cumulative Proportion	0.97148	0.98707	0.99372	0.99721	1.00000						

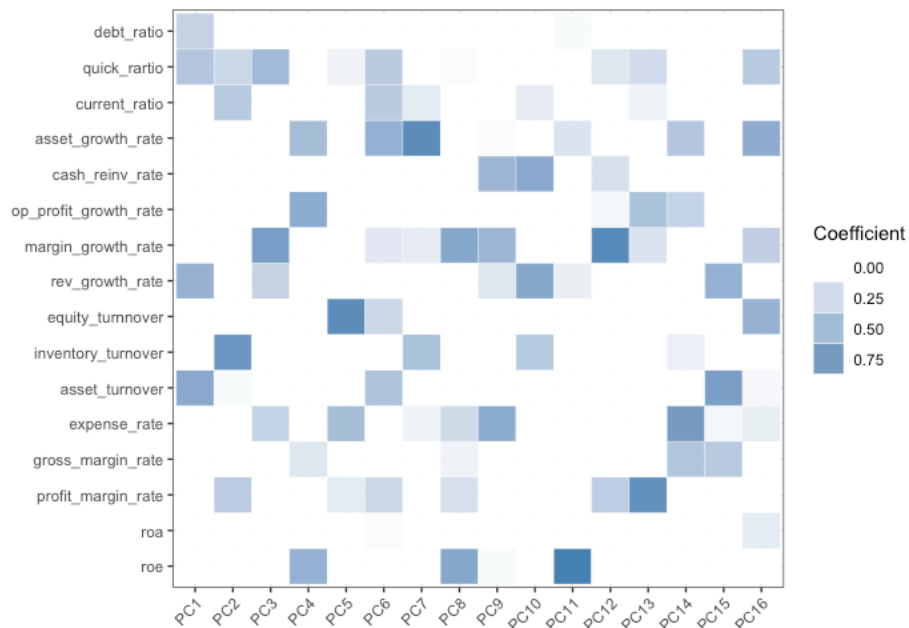
PC	解釋多少變異	PC	解釋多少變異
PC1	14.93%	PC9	6.329%
PC2	13.11%	PC10	2.929%
PC3	23.10%	PC11	1.987%
PC4	12.18%	PC12	1.756%
PC5	9.481%	PC13	1.559%
PC6	7.724%	PC14	0.665%
PC7	7.060%	PC15	0.349%
PC8	6.555%	PC16	0.279%

b. 大概需要多少個 PC 來解釋這筆資料：



➤ 結論：至少需要 8 個 PC 才能解釋 80%以上的變異。

3. 找出前三個主成份分別重要變數為何並解釋



a. PC1 :

- 重點變數：
asset_turnover、rev_growth_rate、quick_ratio、debt_ratio
- 解釋：PC1 注重在公司在運用資產使用效率、營收成長幅度，去除掉存貨的償債能力，以及公司的財務槓桿。

b. PC2

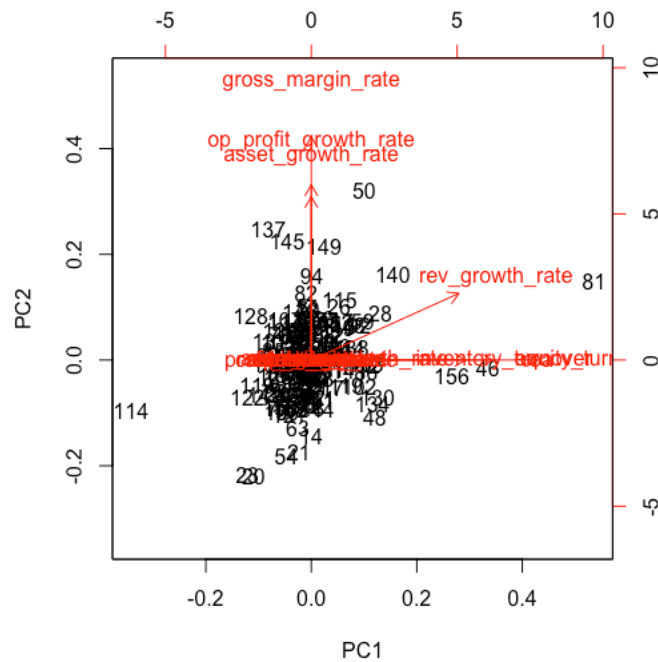
- 重點變數：
inventory_turnover、current_ratio、profit_margin_rate、quick_ratio
- 解釋：PC2 注重在公司使用存貨的效率、整體的償債能力以及企業的經營能力。

c. PC3

- 重點變數：
margin_growth_rate、quick_ratio、rev_growth_rate、expense_rate

- 解釋：PC3 注重在除存貨以外的流動資產使用效率，公司營收與毛利的成長率，以及營業費用率，總體而言詮釋公司費用及資產投入是否可以創造營收且年復一年成長。

4. 找出適合投資的公司



- 結論：因為資產與收益面指標有較佳表現，而選擇投資右上角的公司比較適合。

```

finance <- read.csv("financialdata.csv")
data <- finance[,-1] #去掉第一行公司id

#把Factor型態轉成numeric型態
data$op_profit_growth_rate <- as.numeric(data$op_profit_growth_rate)
data$current_ratio <- as.numeric(data$current_ratio)
data$quick_rartio <- as.numeric(data$quick_rartio)
str(data)

#1.
M <- cor(data) #求出數據及內各個變數的Corr

#用melt函數Reshape資料
library(reshape2)
melted_Corrmat <- melt(M)
head(melted_Corrmat)
write.csv(melted_Corrmat, 'melted_Corrmat.csv')

#畫Heatmap
library(ggplot2)
ggplot(data = melted_Corrmat,
       aes(Var1, Var2)) +
  geom_tile(aes(fill = value), colour = "grey") +
  scale_fill_gradient2(low = "firebrick4", high = "steelblue",
                      mid = "white", midpoint = 0) +
  guides(fill=guide_legend(title="Correlation")) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, vjust = 1),
        axis.title = element_blank())

#2.
library(nsprcomp)
spca <- nscumcomp(data, k=80, nneg=T, scale=T) #每組4個非零係數*16個變數
summary(spca)
screeplot(spca)

pve=(spca$sdev)^2 / (sum(spca$sdev^2))
plot(cumsum(pve), xlab="Principal Component", ylab="Cumulative Proportion
of Variance Explained ", ylim=c(0,1),type='b')
abline(h=0.8)

#3.
ggplot(melt(spca$rotation), aes(Var2, Var1)) +
  geom_tile(aes(fill = value), colour = "white") +
  scale_fill_gradient2(low = "white", high = "steelblue") +
  guides(fill=guide_legend(title="Coefficient")) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1, vjust = 1),
        axis.title = element_blank())

#4.
biplot(spca,scale=T)

```