

Name: 黃子瑋

Student ID: 410475001

---

Part I: Complete the following table using `str_count`

Word	Frequency
黃捷	26
韓國瑜	23
發大財	4
發財	4

Source codes for Part I:

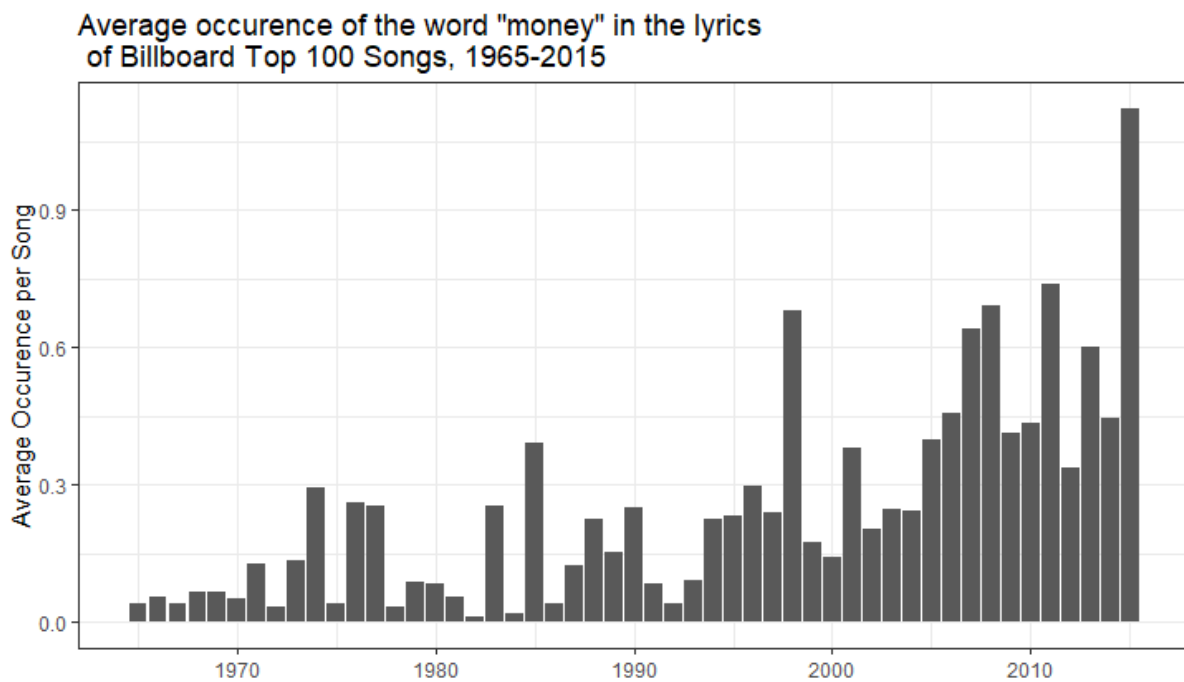
```
library(tidyverse)

#讀取檔案資料
setwd(getwd())
transcript<-read_file('transcript.txt')

#計算字符出現頻率
str_count(transcript,'黃捷')
str_count(transcript,'韓國瑜')
str_count(transcript,'發大財')
str_count(transcript,'發財')
```

---

Part II: Bar plot: Word Frequency of Billboard Top Songs over Time



Question: What do you see in your bar plot? Briefly comment your plot (<50 words):

1990 年代前，只有零星幾年 money 出現的頻率較高，從 2000 年起，money 出現的頻率有逐漸上升的趨勢，在 2015 年達到最高峰

Source Code for Part II

```
library(tidyverse)

#讀取檔案資料
setwd(getwd())
songs<-read_csv('billboard_lyrics.csv')

songs%>%
  mutate(word_count=str_count(Lyrics,'money'))%>%
  group_by(Year)%>%
  summarise(mean=mean(word_count,na.rm=TRUE))%>%
  ggplot(aes(x=Year,y=mean))+
  geom_col()+
  labs(x="y='Average Occurence per Song',
       title='Average occurence of the word \"money\" in the lyrics
of Billboard Top 100 Songs, 1965-2015')+
  theme_bw()
```