

109-2 電腦概論與程式設計

作業 (9)

學號: 410978040 姓名: 黃冠翔

17 六月 2021

Contents

注意事項	1
寫作要點	1
上傳答題檔案	1
其它事項	2
臺北市所得收入者每人所得，本業薪資（依行業別）	2
臺北市所得收入者每人所得，本業薪資（依性別）	4

注意事項

寫作要點

- 繳交期限: 2021/06/21(二), 24:00 前上傳完畢。
- 於課程網站 (<http://www.hmwu.idv.tw>) 下載題目卷。
- 可參考課本、上課講義（包含電子檔）及其它資料。
- 自己親手做，儘量不要與別人（或同學）討論，不可參考同學的答案，不可抄襲。
- 程式設計題，若程式碼直接複製（或照抄）講義上的以不給分為原則。
- 有問題者，請 FB 群組或私訊問助教或老師。
- 程式直接寫在本 Rmd 檔。經 knitr 編譯，產生.html、.pdf、.doc 檔，需印出 R 程式碼及執行結果。
- 不按照規定作答者，酌量扣分。

上傳答題檔案

- 於教師網站首頁登入 [作業考試上傳區]，帳號: r1092。密碼: xxx。
- 選取「正確的」資料夾上傳，若傳錯，請最終要上傳一份正確的答題檔案。
- 請上傳「學號-姓名-R-HW9.Rmd」、「學號-姓名-R-HW9.html」、「學號-姓名-R-HW9.pdf」及「學號-姓名-R-HW9.doc」。(學號及姓名，改成自己)
- 若上傳檔案格式錯誤，內容亂碼，空檔等等問題。請自行負責。
- 若要重覆上傳（第 2 次以上），請在檔名最後加「-2」、「-3」，例如: 「學號-姓名-R-HW7-2.Rmd」等等。
- 上傳兩次（含）以上、格式不合等等酌量扣分。

- 如果上傳網站出現「You can modify the html file, but please keep the link www.wftpservers.com at least.」，請將滑鼠移至「網址列」後，按「Enter」即可。若再不行，請換其它瀏覽器（IE/Edge/Firefox/Chrome）。
- 有問題者，請 FB 私訊老師。

其它事項

- 若有題目不會寫、或只會寫一半、或結果是有錯的，導致 **knit** 無法編譯產生文件，則可以「不執行有錯的程式碼」，但必需列印此段程式碼。助教會依照狀況部份給分。
- 此份作業，助教以 **pdf** 檔批改為準。**Rmd** 則是做為比對之用（比對同學們的 **Rmd** 是否相似）。
- 總分 100 分，由助教決定每題配分。全部改完會上傳，答案卷同學們可自由下載。

臺北市所得收入者每人所得，本業薪資（依行業別）

- 資料敘述：臺北市所得收入者每人所得 - 行業別-年依項目，年別與行業
- 資料來源：台北市家庭收支資料庫查詢系統
- 資料網址：https://statdb.dbas.gov.taipei/pxweb2007-tp/dialog/statfile9_Fl.asp
- 資料檔：fi00123y2a2021523444471.xlsx

讀取資料，並以 **ggplot2** 畫出下圖。提示：

- **read_excel {readxl}** 會使用到的參數（引數）為 **range** 和 **na**。
- 檢查資料之結構，及每個欄位之資料類別是否正確。
- 讀入之資料，第一欄為年份（命名 **year**），並修正為有順序之因子。
- 選取繪圖所需之欄位資料，造出一個為 **tidy** 型式的資料框（使用 **stack {utils}** 或 **melt {reshape2}**），並重新對各欄位命名。
- 圖形中，有主標題，縱軸 **tick** 之標記有 **\$** 之符號。

```
# your source code here
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.0.4
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
library(scales)
```

```
## Warning: package 'scales' was built under R version 4.0.5
```

```
library(showtext)
```

```
## Warning: package 'showtext' was built under R version 4.0.5
```

```
## Loading required package: sysfonts
```

```
## Warning: package 'sysfonts' was built under R version 4.0.5
```

```
## Loading required package: showtextdb
```

```
## Warning: package 'showtextdb' was built under R version 4.0.5
```

```
showtext.auto()
```

```
## 'showtext.auto()' is now renamed to 'showtext_auto()'
## The old version still works, but consider using the new function in future code
```

```
job <- read_excel("fi00123y2a2021523444471.xlsx", range = "B4:S15", na="-")
```

```
## New names:
## * `` -> ...1
```

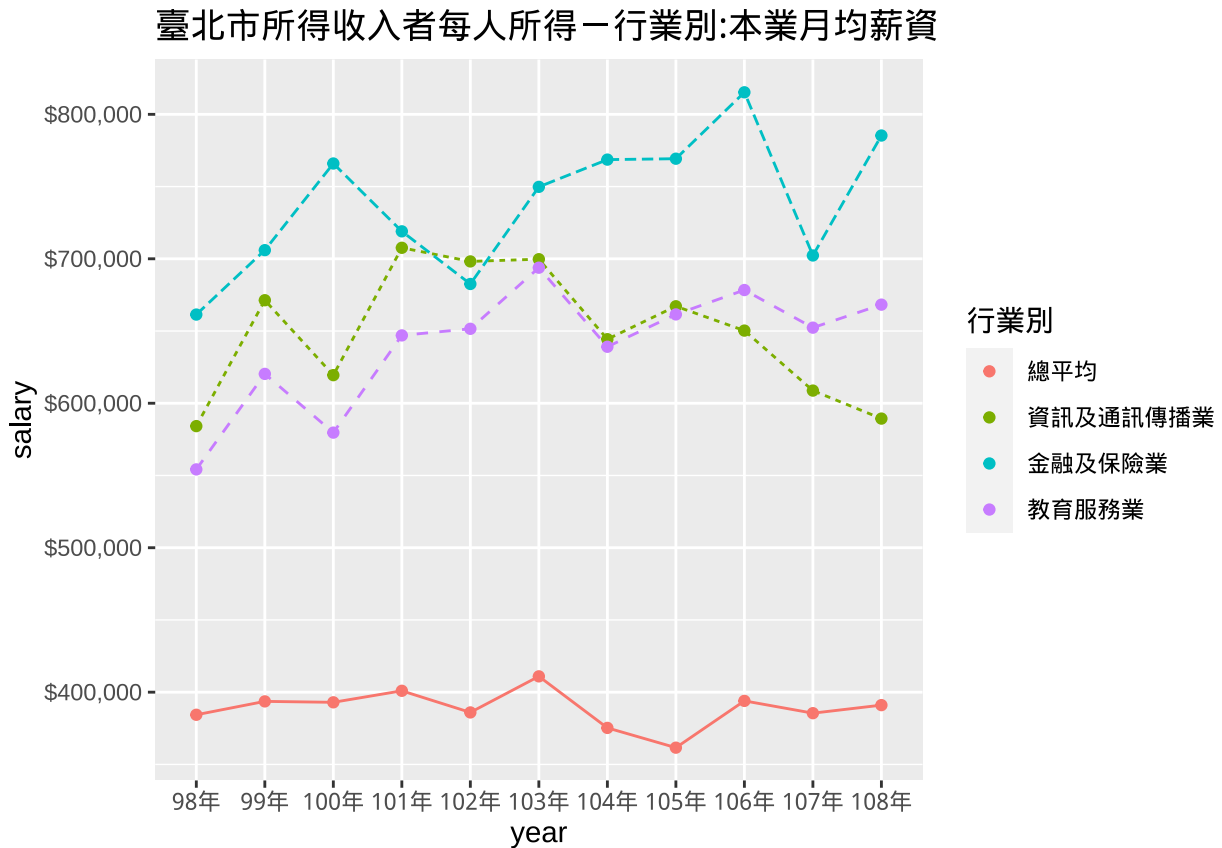
```
str(job)
```

```
## tibble [11 x 18] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:11] "98年" "99年" "100年" "101年" ...
## $ 總平均 : num [1:11] 384362 393629 392977 400942 385983 ...
## $ 農林漁牧業 : num [1:11] 181119 NA NA 18471 180762 ...
## $ 工業 : num [1:11] 534587 548791 531035 545619 607659 ...
## $ 批發及零售業 : num [1:11] 387149 362585 357534 422184 398104 ...
## $ 運輸及倉儲業 : num [1:11] 458256 439096 435056 426898 472565 ...
## $ 住宿及餐飲業 : num [1:11] 256377 261758 271990 303493 287234 ...
## $ 資訊及通訊傳播業 : num [1:11] 584172 671331 619455 707620 698195 ...
## $ 金融及保險業 : num [1:11] 661405 705985 765977 719002 682576 ...
## $ 不動產業 : num [1:11] 578425 648006 483807 506087 643146 ...
## $ 專業、科學及技術服務業 : num [1:11] 598956 630438 557103 519730 524620 ...
## $ 支援服務業 : num [1:11] 371212 349055 362878 406428 401331 ...
## $ 公共行政及國防 : num [1:11] 622175 607745 655204 632209 639553 ...
## $ 教育服務業 : num [1:11] 554212 620317 579683 646987 651541 ...
## $ 醫療保健及社會工作服務業 : num [1:11] 597467 677586 675178 740450 543259 ...
## $ 藝術、娛樂及休閒服務業 : num [1:11] 398387 391337 521664 383413 401794 ...
## $ 其他服務業 : num [1:11] 229712 246252 257831 241582 270133 ...
## $ 無業者 : num [1:11] 1184 NA 308 127 NA ...
```

```
year.f <- as.factor(job$...1)
year <- factor(year.f, levels = c("98 年", "99 年", "100 年", "101 年", "102 年", "103 年", "104 年", "105 年"))
p <- stack(job, select = c(" 總平均", " 資訊及通訊傳播業", " 金融及保險業", " 教育服務業"))
a <- cbind(year, p)
pic <- ggplot(p, aes(x=a$year, y=a$values, group=a$ind, color=a$ind)) + geom_line(aes(linetype=a$ind),
  labs(title = " 臺北市所得收入者每人所得 - 行業別: 本業月均薪資", x="year", y="salary", col=" 行業別") +
  scale_y_continuous(labels = dollar)
```

```
## Warning: `show_guide` has been deprecated. Please use `show.legend` instead.
```

pic



臺北市所得收入者每人所得，本業薪資（依性別）

- 資料敘述: 臺北市所得收入者每人所得 - 性別-年依項目, 年別與性別
- 資料來源: 台北市家庭收支資料庫查詢系統
- 資料網址: https://statdb.dbas.gov.taipei/pxweb2007-tp/dialog/statfile9_FL.asp
- 資料檔: fi00128y2a2021523591831.xlsx

讀取資料，並以 `ggplot2` 畫出下圖。提示: `geom_bar` · `geom_line` · `geom_segment` · `geom_text`。

```
# your source code here
```

```
library(readxl)
library(ggplot2)
library(scales)
library(RColorBrewer)
```

```
## Warning: package 'RColorBrewer' was built under R version 4.0.3
```

```
sex <- read_excel("fi00128y2a2021523591831.xlsx", range = "B4:E15")
```

```
## New names:
## * `` -> ...1
```

```

year.f <- as.factor(sex$...1)
year <- factor(year.f, levels = c("98 年", "99 年", "100 年", "101 年", "102 年", "103 年", "104 年", "105 年", "106 年", "107 年", "108 年"))
p <- stack(sex, select = c(" 男", " 女"))
a <- cbind(year, p)
sex <- as.data.frame(sex)
b <- cbind(sex$總平均, p)
sexp <- ggplot(p, aes(x=a$year, y=a$values, fill=a$ind, group=a$ind)) + geom_bar(stat="identity", position="dodge") +
  geom_line(b, mapping = aes(x=a$year, y=`sex$ 總平均`), color=brewer.pal(11, "PiYG")[11], show_guide = FALSE) +
  geom_point(b, mapping = aes(x=a$year, y=`sex$ 總平均`), color=brewer.pal(11, "PiYG")[11], show_guide = FALSE) +
  geom_segment(aes(x="106 年", y=450000, xend="107 年", yend=450000), color=brewer.pal(11, "PiYG")[11], show_guide = FALSE) +
  geom_text(aes(x="108 年", y=450000, label=" 總平均")) +
  labs(title = " 臺北市所得收入者每人所得 - 性別: 本業月均薪資", x="year", y="salary", fill=" 性別") +
  scale_y_continuous(labels = dollar)

```

Warning: `show_guide` has been deprecated. Please use `show.legend` instead.

Warning: `show_guide` has been deprecated. Please use `show.legend` instead.

sexp

