

108208004 經濟三 白植允

Homework 1

#1

```
1 client <- read.csv("client_list.csv", sep = ",")
2 product <- read.csv("product_list.csv", sep = ",")
3 sales <- read.csv("salesdata.csv", sep = ",")
4 library(tidyverse)
5
6 #1
7 product %>%
8   separate(Item, c("Product", "Item"), sep = "_") -> product
```

	Product	Item
1	101	iPhone
2	102	iPad
3	103	MacBook
4	104	iMac
5	105	AirPods
6	106	AppleWatch

#2

```
10 #2
11 product$Product <- as.numeric(product$Product)
12 full.table <- client %>%
13   inner_join(sales) %>%
14   inner_join(product)
```

	Client	Age	Membership	Gender	salesID	Store	Product	UnitPrice	Quantity	Region	Item
1	1	22	basic	male	1	B	105	4	13	Taiwan	AirPods
2	1	22	basic	male	2	B	106	5	10	Taiwan	AppleWatch
3	1	22	basic	male	3	A	103	4	11	Taiwan	MacBook
4	1	22	basic	male	4	B	102	6	2	Taiwan	iPad
5	2	35	silver	female	5	A	101	4	44	USA	iPhone
6	2	35	silver	female	6	A	102	5	3	USA	iPad
7	2	35	silver	female	7	A	103	6	8	USA	MacBook
8	2	35	silver	female	8	B	104	7	4	USA	iMac
9	2	35	silver	female	9	B	105	8	6	USA	AirPods
10	2	35	silver	female	10	B	106	9	10	USA	AppleWatch
11	3	33	gold	male	11	A	101	4	22	Brazil	iPhone
12	3	33	gold	male	12	A	102	5	35	Brazil	iPad
13	3	33	gold	male	13	B	103	6	67	Brazil	MacBook
14	4	33	diamond	female	14	B	101	4	3	Korea	iPhone
15	4	33	diamond	female	15	B	102	5	5	Korea	iPad
16	4	33	diamond	female	16	B	103	6	7	Korea	MacBook
17	4	33	diamond	female	17	B	104	7	34	Korea	iMac
18	4	33	diamond	female	18	A	105	8	64	Korea	AirPods

#3

```
16 #3
17 full.table %>%
18   mutate(spend = UnitPrice * Quantity) -> full.table
19
```

	Client	Age	Membership	Gender	salesID	Store	Product	UnitPrice	Quantity	Region	Item	spend
1	1	22	basic	male	1	B	105	4	13	Taiwan	AirPods	52
2	1	22	basic	male	2	B	106	5	10	Taiwan	AppleWatch	50
3	1	22	basic	male	3	A	103	4	11	Taiwan	MacBook	44
4	1	22	basic	male	4	B	102	6	2	Taiwan	iPad	12
5	2	35	silver	female	5	A	101	4	44	USA	iPhone	176
6	2	35	silver	female	6	A	102	5	3	USA	iPad	15
7	2	35	silver	female	7	A	103	6	8	USA	MacBook	48
8	2	35	silver	female	8	B	104	7	4	USA	iMac	28
9	2	35	silver	female	9	B	105	8	6	USA	AirPods	48
10	2	35	silver	female	10	B	106	9	10	USA	AppleWatch	90
11	3	33	gold	male	11	A	101	4	22	Brazil	iPhone	88
12	3	33	gold	male	12	A	102	5	35	Brazil	iPad	175
13	3	33	gold	male	13	B	103	6	67	Brazil	MacBook	402
14	4	33	diamond	female	14	B	101	4	3	Korea	iPhone	12
15	4	33	diamond	female	15	B	102	5	5	Korea	iPad	25
16	4	33	diamond	female	16	B	103	6	7	Korea	MacBook	42
17	4	33	diamond	female	17	B	104	7	34	Korea	iMac	238
18	4	33	diamond	female	18	A	105	8	64	Korea	AirPods	512
19	5	52	basic	male	19	A	101	4	23	China	iPhone	92
20	5	52	basic	male	20	A	102	5	45	China	iPad	225

#4

```
20 #4
21 full.table %>%
22   mutate(group = ifelse(Membership == "gold" | Membership == "diamond", "gold & diamond", "others")) %>%
23   group_by(group) %>%
24   summarise(mean(Age), male = sum(Gender == "male"), mean(spend))
25
```

```
group          `mean(Age)`    male  `mean(spend)`
<chr>          <dbl>    <int>         <dbl>
1 gold & diamond    27.3      6         242.
2 others           32.3     10         165.
> |
```

由分析可知 gold&diamond 組的平均年紀、平均消費較高，男性會員較少

#5

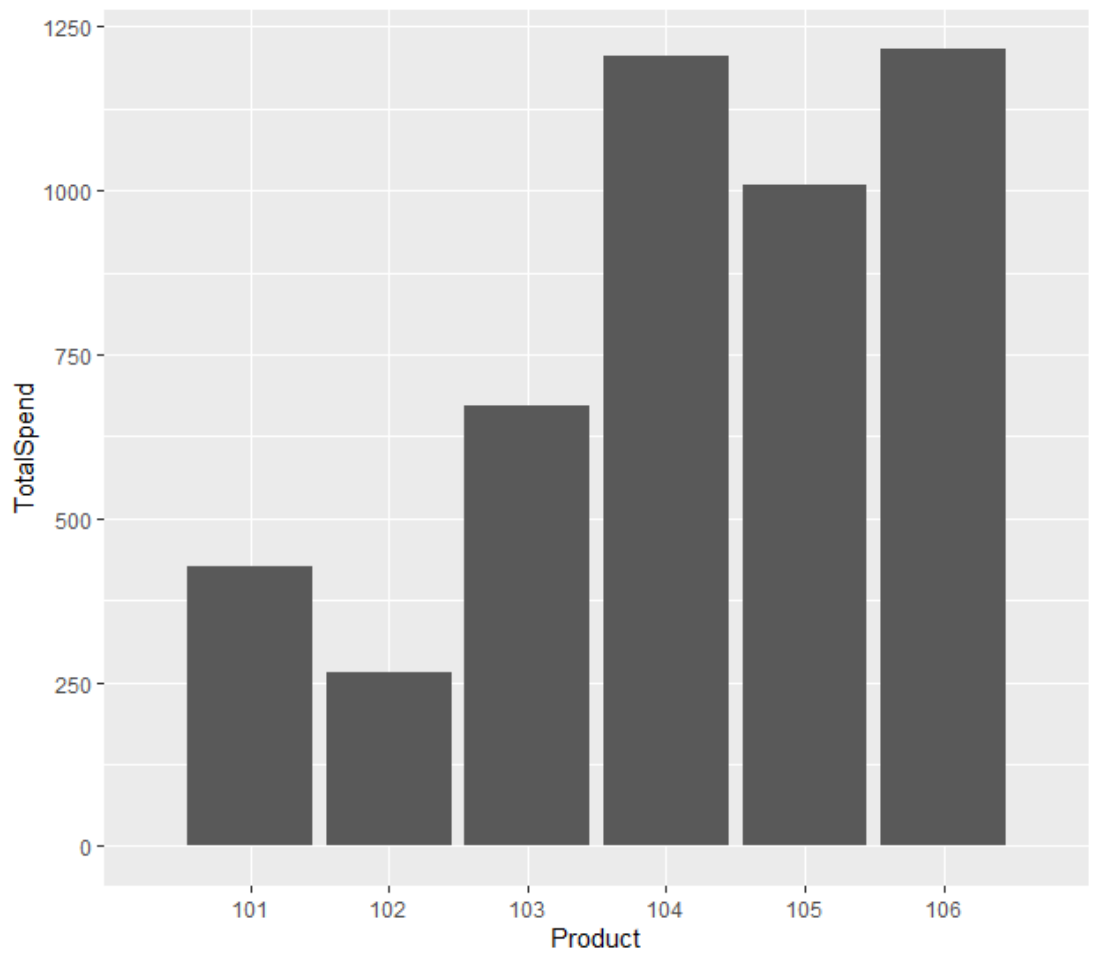
```
26 #5
27 full.table %>%
28   filter(Gender == "female") %>%
29   summarise(mean(Age), mean(spend), median(spend), sum(spend))
30
31 Totals <- full.table %>%
32   group_by(Product) %>% |
33   filter(Gender == "female") %>%
34   summarise(TotalSpend = sum(spend))
35
36 Totals %>%
37   ggplot(aes(x = Product, y = TotalSpend), fill = Product) +
38   geom_bar(stat = 'identity') +
39   scale_x_discrete(limits = Totals$Product)
40
```

```

mean(Age) mean(spend) median(spend) sum(spend)
1 27.78261 208.3478 216 4792
> |

```

可知女性客戶的平均年齡約 27 歲，平均消費約 208 元，總消費額 4792 元



由此可知女性花在產品 106 和 104 的錢最多，102 的最少

- 程式碼

```

client <- read.csv("client_list.csv", sep = ",")
product <- read.csv("product_list.csv", sep = ",")
sales <- read.csv("salesdata.csv", sep = ",")
library(tidyverse)

#1
product %>%
  separate(Item, c("Product", "Item"), sep = "_") -> product

#2
product$Product <- as.numeric(product$Product)

```

```
full.table <- client %>%  
  inner_join(sales) %>%  
  inner_join( product)
```

#3

```
full.table %>%  
  mutate( spend = UnitPrice * Quantity) -> full.table
```

#4

```
full.table %>%  
  mutate( group = ifelse( Membership == "gold" | Membership  
== "diamond", "gold & diamond", "others")) %>%  
  group_by(group) %>%  
  summarise(mean(Age), male = sum(Gender=="male"), mean(spend))
```

#5

```
full.table %>%  
  filter(Gender == "female") %>%  
  summarise(mean(Age), mean(spend), median(spend), sum(spend))
```

```
Totals <- full.table %>%  
  group_by(Product) %>%  
  filter(Gender == "female") %>%  
  summarise (TotalSpend= sum(spend))
```

```
Totals %>%  
  ggplot(aes(x= Product, y= TotalSpend), fill = Product)+  
  geom_bar(stat = 'identity') +  
  scale_x_discrete(limits = Totals$Product)
```