

Вправи 1:

- Знайдіть кількість вимірів датафрейму comps.

```
> dim(comps)
[1] 6259 7
```

- Відобразіть перші шість рядків, перші п'ятнадцять рядків, останні шість рядків.

```
> head(comps)
  price  hd ram screen cd multi trend
1  1499   80  4    14 no    no     1
2  1795   85  2    14 no    no     1
3  1595  170  4    15 no    no     1
4  1849  170  8    14 no    no     1
5  3295  340 16    14 no    no     1
6  3695  340 16    14 no    no     1
```

```
> dim(comps)
[1] 6259 7
```

```
> head (comps, 6)
  price  hd ram screen cd multi trend
1  1499   80  4    14 no    no     1
2  1795   85  2    14 no    no     1
3  1595  170  4    15 no    no     1
4  1849  170  8    14 no    no     1
5  3295  340 16    14 no    no     1
6  3695  340 16    14 no    no     1
```

```
> head (comps, 15)
  price  hd ram screen cd multi trend
1  1499   80  4    14 no    no     1
2  1795   85  2    14 no    no     1
3  1595  170  4    15 no    no     1
4  1849  170  8    14 no    no     1
5  3295  340 16    14 no    no     1
6  3695  340 16    14 no    no     1
7  1720  170  4    14 yes   no     1
8  1995   85  2    14 no    no     1
9  2225  210  8    14 no    no     1
10 2575  210  4    15 no    no     1
11 2195  170  8    15 no    no     1
12 2605  210  8    14 no    no     1
13 2045  130  4    14 no    no     1
14 2295  245  8    14 no    no     1
15 2699  212  8    14 no    no     1
```

```
> tail (comps, 6)
  price  hd ram screen cd multi trend
6254 2154  850 16    15 yes   no    35
6255 1690  528  8    15 no    no    35
6256 2223  850 16    15 yes   yes   35
6257 2654 1200 24    15 yes   no    35
6258 2195  850 16    15 yes   no    35
6259 2490  850 16    17 yes   no    35
```

- Відобразіть імена датафрейму.

```
> names(comps)
[1] "price" "hd"    "ram"    "screen" "cd"     "multi" "trend"
```

Вправи 2:

- Скільки змінних у наборі даних flats?

7

```
> str(comps)
'data.frame': 6259 obs. of 7 variables:
 $ price : int  1499 1795 1595 1849 3295 3695 1720 1995 2225 2575 ...
 $ hd     : int   80  85 170 170 340 340 170  85 210 210 ...
 $ ram    : int   4  2  4  8 16 16  4  2  8  4 ...
 $ screen: int  14 14 15 14 14 14 14 14 14 15 ...
 $ cd     : chr  "no" "no" "no" "no" ...
 $ multi  : chr  "no" "no" "no" "no" ...
 $ trend  : int   1  1  1  1  1  1  1  1  1  1 ...
```

- Яка кількість комп'ютерів у наборі даних compscount?

6259

```
> count(comps)
# A tibble: 1 x 1
      n
  <int>
1  6259
```

- Яка кількість комп'ютерів на 16 гб. RAM з діагоналлю екрану 14 дюймів?

393

```
> comps %>%
+   filter(ram == 16) %>%
+   filter(screen == 14) %>%
+   count()
# A tibble: 1 x 1
      n
  <int>
1   393
```

- Яка медіана розміру hdd комп'ютера з дисководом?

528

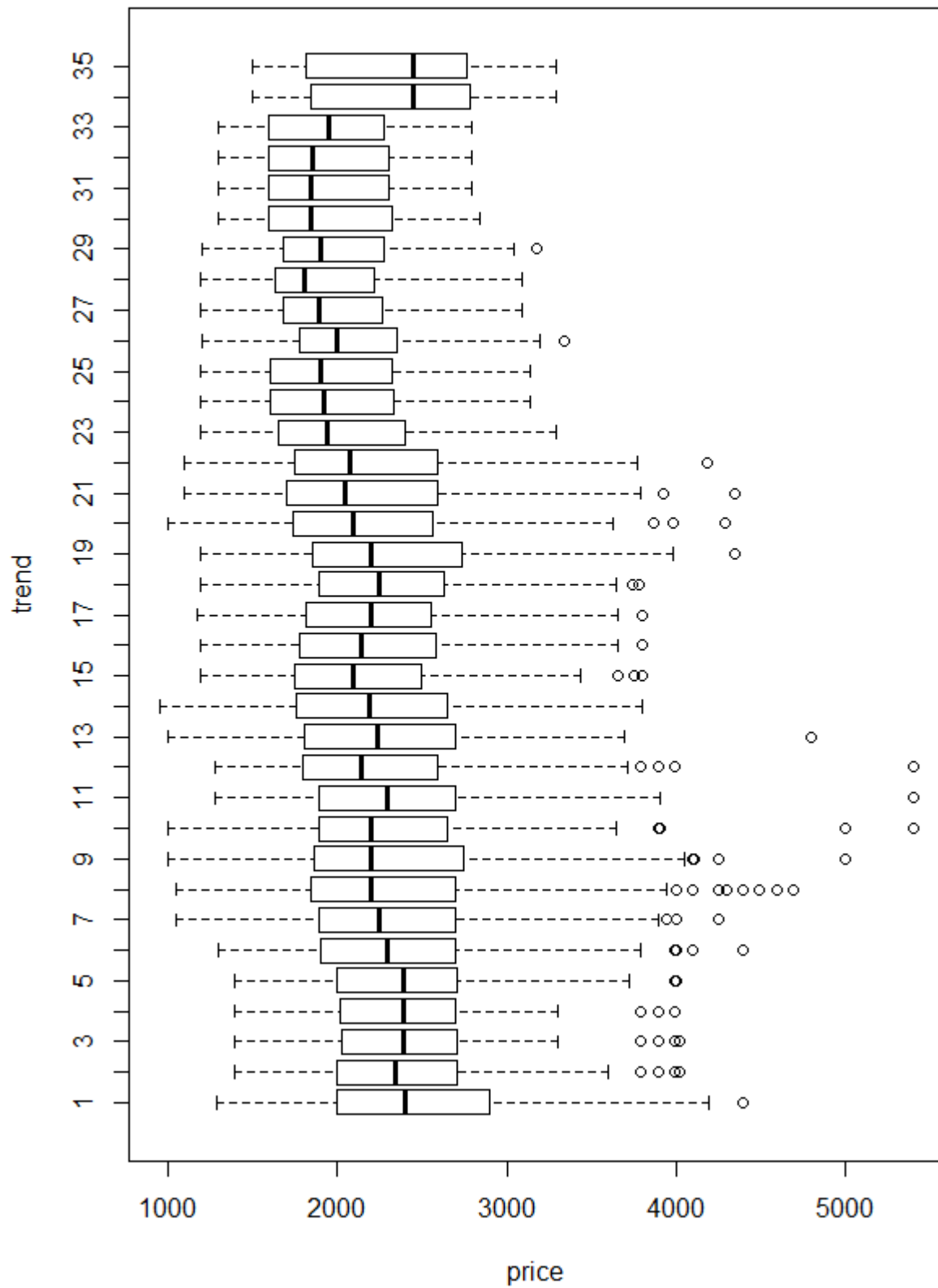
```
> comps %>%
+   filter(cd == "yes") %>%
+   summarise(median(hd))
  median(hd)
1         528
```

Вправи 3:

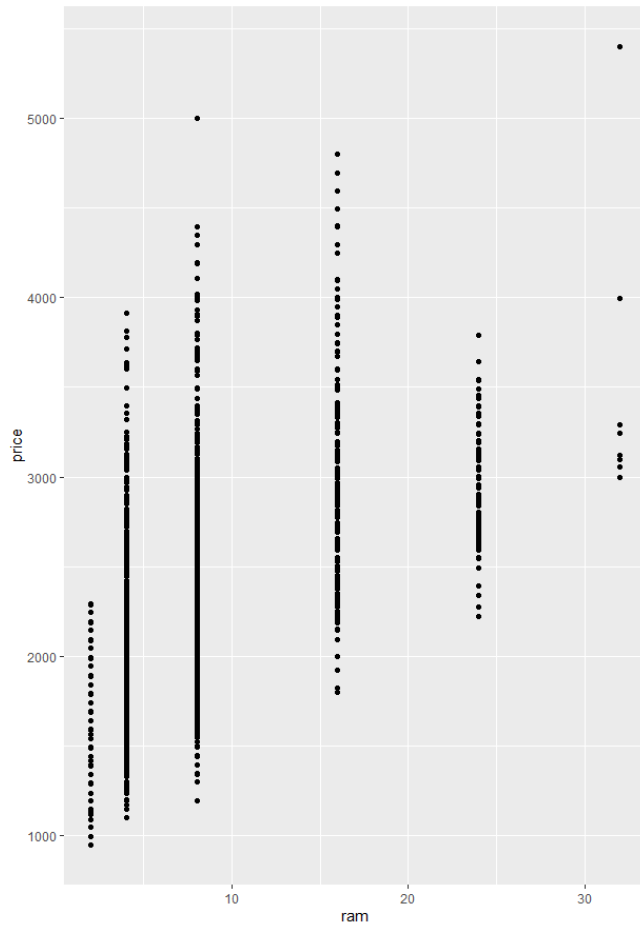
- Побудуємо коробчатую діаграму для візуалізації популярності в залежності від ціни

```
boxplot(price ~ trend,
        ylab = "trend",
        xlab = "price",
        main = "Залежність популярності від ціни",
        horizontal = TRUE,
        data = comps)
```

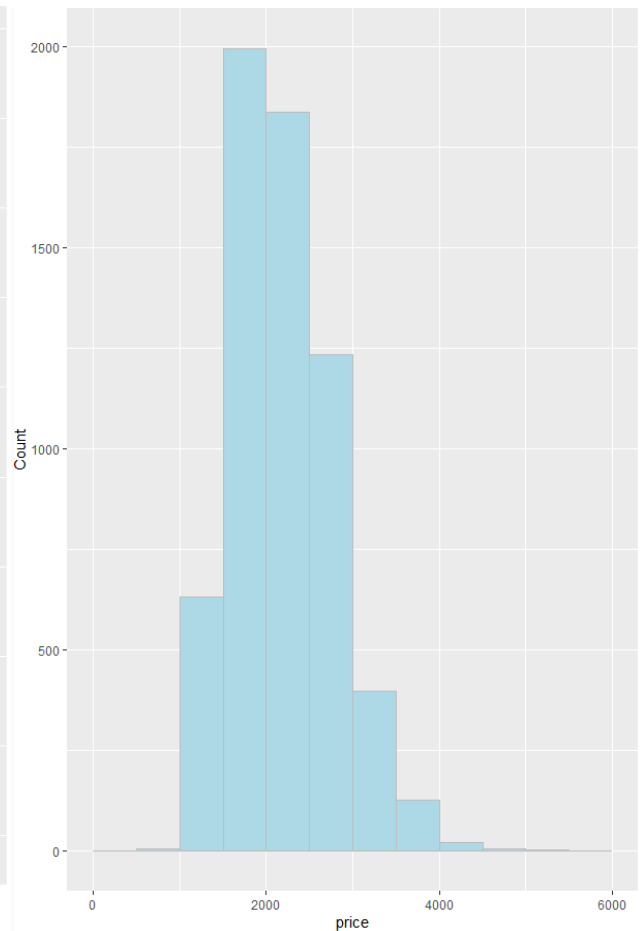
Залежність популярності від ціни



- Побудуємо графік розсіювання, який відображатиме залежність ціни від об'єму RAM (1)
- Побудуйте гістограму для оцінки розподілу ціни комп'ютерів (2)



1



2

Висновки

Я навчився працювати з таблицями даних за допомогою мови R, трансформувати та візуалізувати дані, зокрема знайшов медіану, фільтрував інформацію для отримання потрібних даних, будував діаграми для порівняння різних характеристик таблиці, завдяки чому навчився оцінювати дані за коробчатою, точковою діаграмою та гістограмою,