MAS291 - HOMEWORK CHAP 6

Nguyen Dang Loc - SE160199

6-8

In Applied Life Data Analysis (Wiley, 1982), Wayne Nelson presents the breakdown time of an insulating fluid between electrodes at 34 kV. The times, in minutes, are as follows: 0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, and 72.89. Calculate the sample mean and sample standard deviation

Solution:

Sample mean:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{\sum_{i=1}^{19} x_i}{19} = \frac{272.82}{19} \approx \boldsymbol{14.359\ min}$$

Sample variance:

$\sum_{i=1}^{19} x_i = 272.82$ $\qquad$ $\sum_{i=1}^{19} x_i^2 = 10333.9$

$$s^2 = \frac{\sum_{i=1}^{19} x_i^2 - \frac{\left(\sum_{i=1}^{19} x_i\right)^2}{19}}{19-1} = \frac{10333.9 - \frac{272.82^2}{19}}{18} \approx 356.4718$$

Sample standard deviation:

$$s = \sqrt{356.471} \approx \boldsymbol{18.88\ min}$$

Use R:

```
x <- c(0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50,
       7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71, 72.89)
mean(x) // 14.35895
sd(x) // 18.88
```

6-50

Construct frequency distributions and histograms with 8 bins and 16 bins for the motor fuel octane data in Exercise 6-30. Compare the histograms. Do both histograms display similar information?

(6-30) An article in Technometrics (1977, Vol. 19, p. 425) presented the following data on the motor fuel octane ratings of several blends of gasoline:

Solution:

8 bins:

Construct frequency distributions: create 8 intervals of equal width

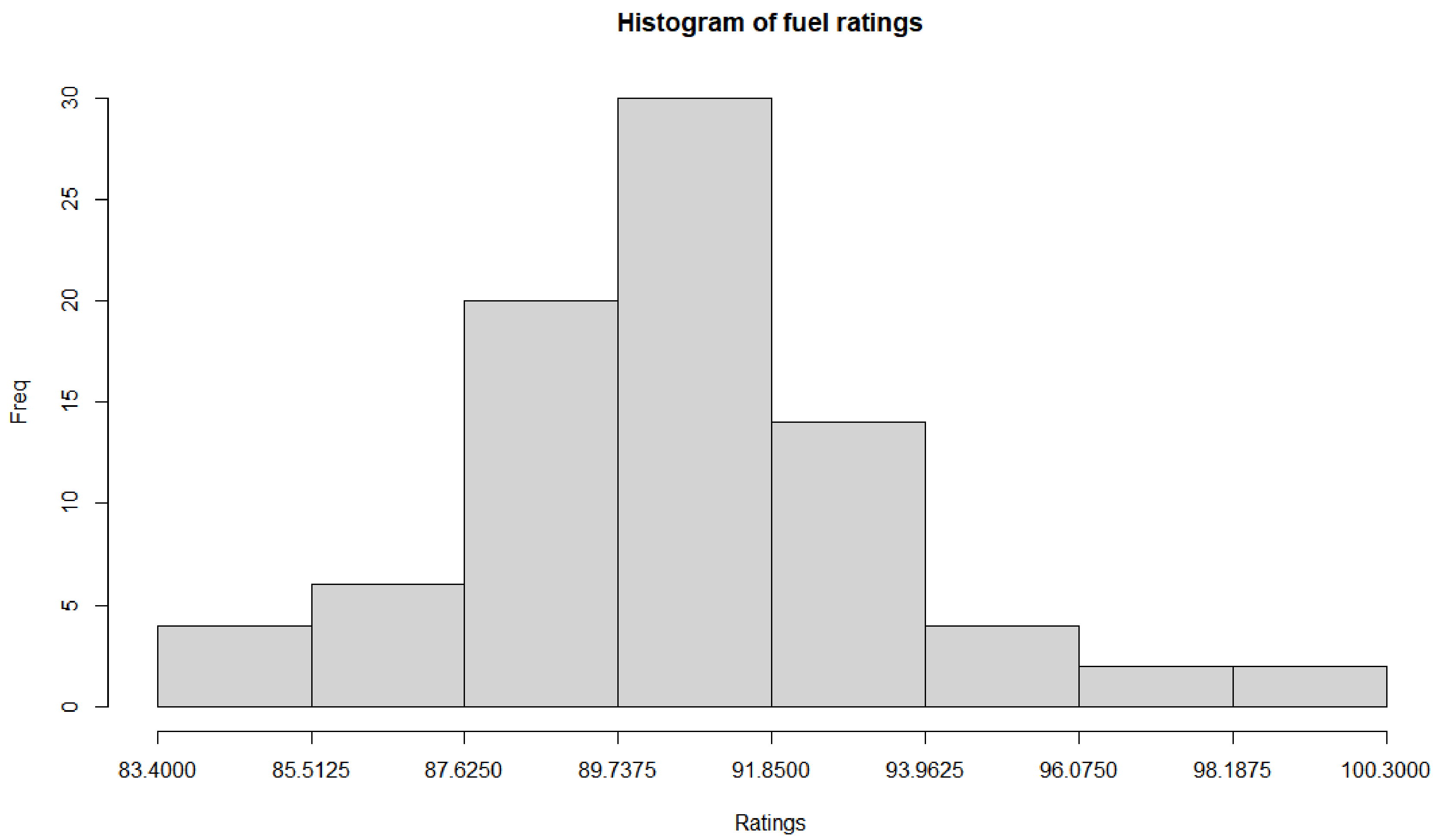
```
data <- read.table("D://MAS//650.txt", sep="", header=F,
                   na.strings="", stringsAsFactors=F) # read data
dt <- as.numeric(data) # convert to list of numeric variable
d <- (max(dt) - min(dt)) / 8
bins <- seq(min(dt), max(dt), by=d) # list of boundaries
fd <- cut(dt, bins) # group data into bins
transform(table(fd))

# Result
           fd Freq
1 (83.4,85.5]    3
2 (85.5,87.6]    6
3 (87.6,89.7]   20
4 (89.7,91.8]   30
```

```
5   (91.8,94]   14
6   (94,96.1]    4
7 (96.1,98.2]    2
8  (98.2,100]    2
```

Construct histogram

```
hist(dt, xlab = "Ratings", ylab = "Freq", breaks = seq(min(dt), max(dt),
     length.out = 9), xaxt='n', main="Histogram of fuel ratings")
axis(1, at = seq(min(dt), max(dt))
```



16 bins:

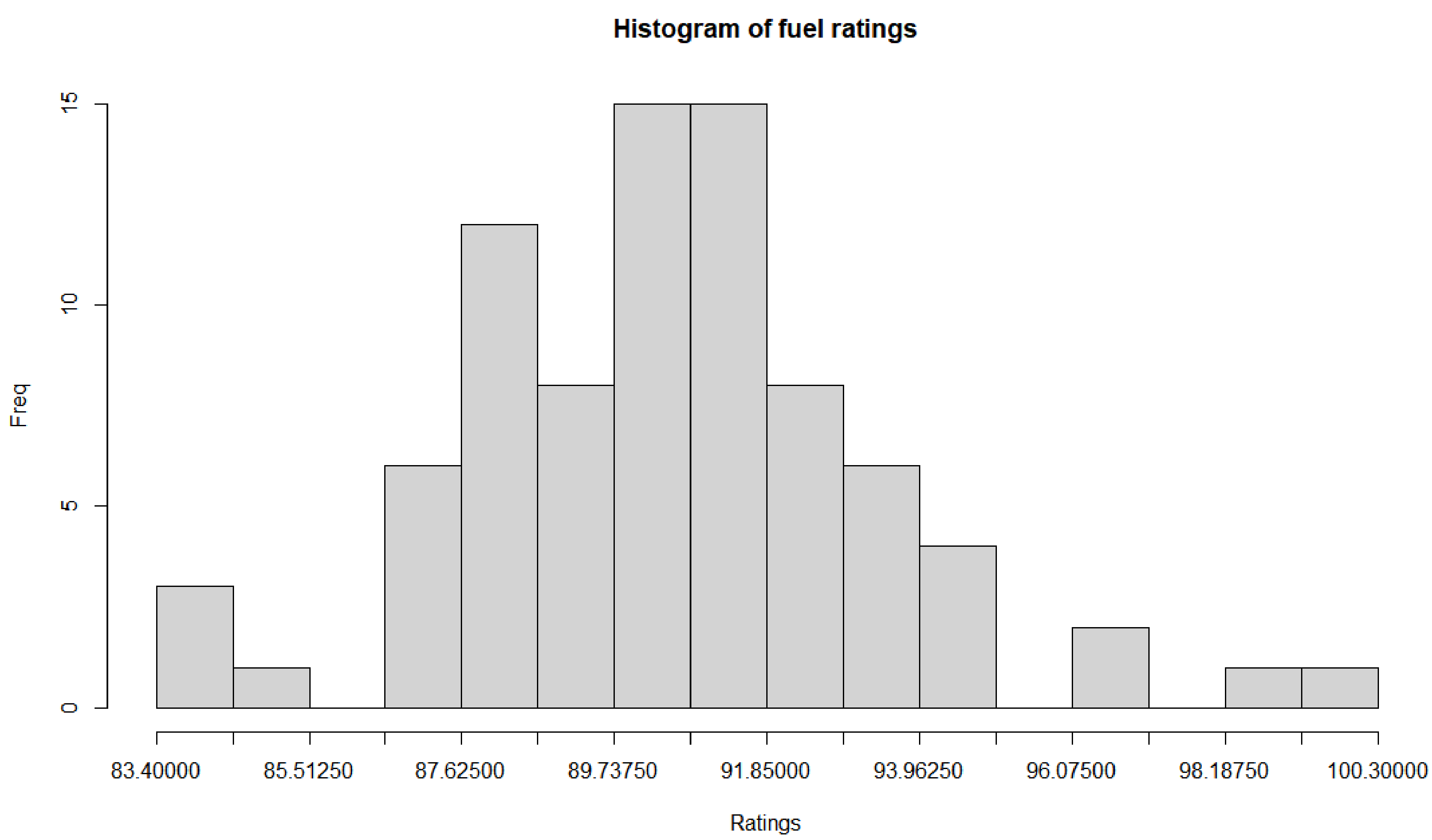
Construct frequency distributions: create 16 intervals of equal width

```
data <- read.table("D://MAS//650.txt", sep="", header=F,
                   na.strings="", stringsAsFactors=F) # read data
dt <- as.numeric(data) # convert to list of numeric variable
d <- (max(dt) - min(dt)) / 16
bins <- seq(min(dt), max(dt), by=d)
fd <- cut(dt, bins)
transform(table(fd))

# Result
            fd Freq
1  (83.4,84.5]    2
2  (84.5,85.5]    1
3  (85.5,86.6]    0
4  (86.6,87.6]    6
5  (87.6,88.7]   12
6  (88.7,89.7]    8
7  (89.7,90.8]   15
8  (90.8,91.8]   15
9  (91.8,92.9]    8
10   (92.9,94]    6
11     (94,95]    4
12   (95,96.1]    0
13 (96.1,97.1]    2
14 (97.1,98.2]    0
15 (98.2,99.2]    1
16  (99.2,100]    1
```

Construct histogram

```
hist(dt, xlab = "Ratings", ylab = "Freq", breaks = seq(min(dt), max(dt),
     length.out = 17), xaxt='n', main="Histogram of fuel ratings")
axis(1, at = seq(min(dt), max(dt), by=d))
```



The shapes of the 2 histograms are quite the same. The two histograms are likely to display similar information.

6-83

The pull-off force for a connector is measured in a laboratory test. Data for 40 test specimens follow (read down, then left to right). Construct and interpret either a digidot plot or a separate stem-and-leaf and time series plot of the data.

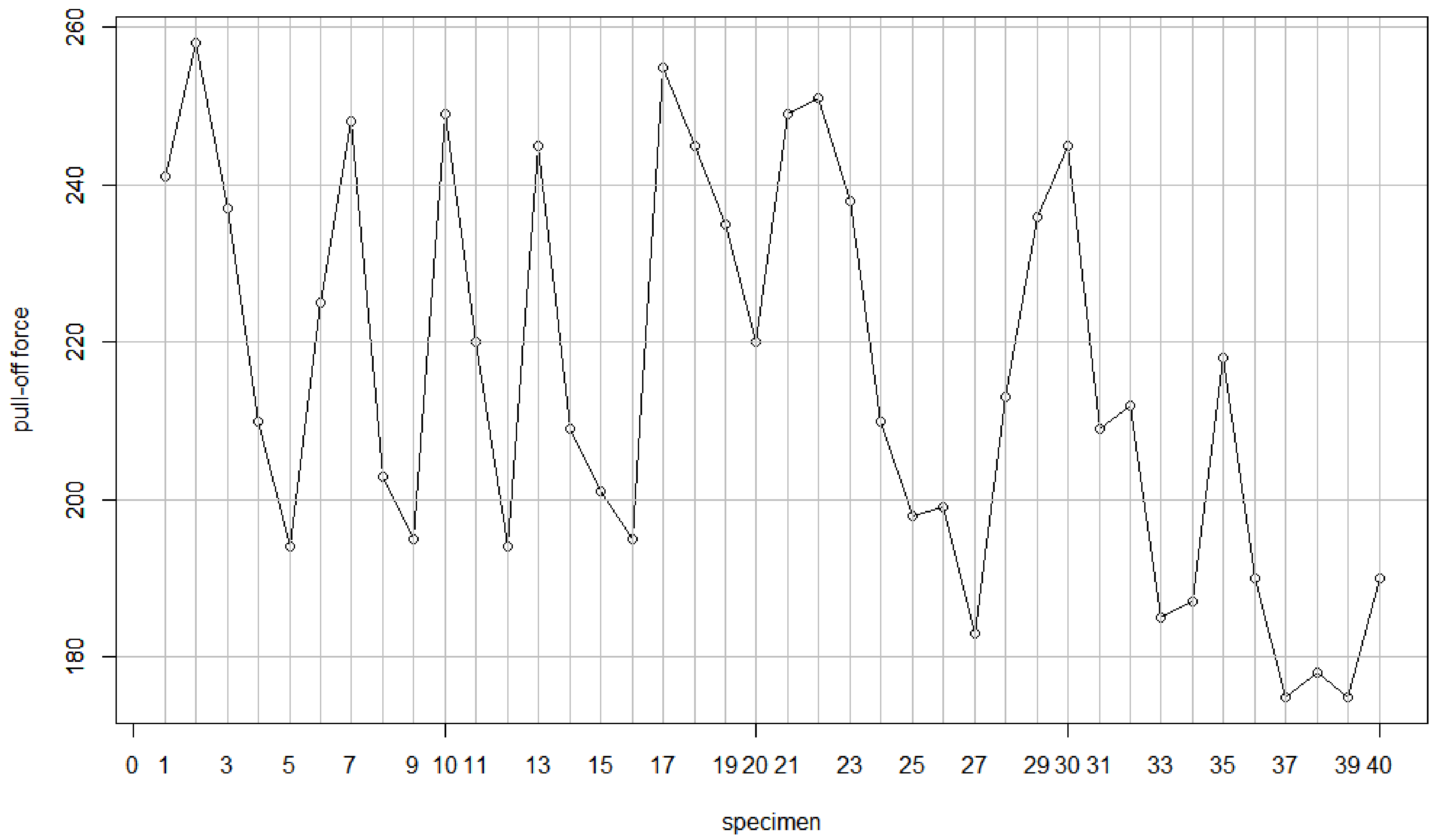
Solution:

Time series plot:

The specimen is on y-axis and the pull-off force is on the x-axis.

```
data <- c(241,258,237,210,194,225,248,203,195,249,220,194,245,
           209,201,195,255,245,235,220,249,251,238,210,198,199,
           183,213,236,245,209,212,185,187,218,190,175,178,175,190)

xValue <- 1:length(data)
yValue <- data
plot(data, type='o',xlab='specimen', ylab='pull-off force')
axis(1, at = seq(1, 40, by = 1), tck = 1, lty = 1, col = "gray")
axis(2, at = seq(min(yValue), max(yValue), by = 20), tck = 1, lty = 1, col = "gray")
```



Stem-and-leaf plot

```
stem(data)
  The decimal point is 1 digit(s) to the right of the |
  17 | 558
  18 | 357
  19 | 00445589
  20 | 1399
  21 | 00238
  22 | 005
  23 | 5678
  24 | 1555899
  25 | 158
```

In the time-series plot, it does not likely to present upwards trend or downwards trend

→ there is no obvious pattern or conclusion drawn from the plot.