No.

Date

概率统计 Probability Statistics

Homework 01

22cst 薛子翔

Section 1.2

Ex.11

Stem	Leaf
9H	8 5
9L	3 0
8H	9 9 8 7 5 5 5
8L	4 4 3 2 2 1 1 1 1 0 0
7H	
7L	4 4 2 2 2 1 0 0
6H	9 9 8 7 6 6
6L	4 3 0

Stem: "H" represents high tens digits.
"L" represents low tens digits.
leaf: ones digits.

Feature: ① A typical value is "81".
② There is a gap in "7H".
③ The value is not symmetric.
④ There is only one peak "8L".
⑤ It seems no outlying value exis

Ex.14

a.) Stem	Leaf
18	9
15	5 3 0 0
14	6 3
13	8
12	7 3
11	9 9 9 5 3 3 2
10	8 8 6 5 5 4 4 4 3 2
9	8 8 7 6 6 6 5 3 3 3 3 2 2 1 0
8	8 4 4 3 3 2 2 0
7	8 6 6 5 5 5 5 5 4 3 3 2 2 1 0 0 0
6	9 9 9 9 8 7 7 6 6 6 5 4 4 4 3 3 2 2 2 2 1 0 0 0 0
5	9 8 7 6 6 6 5 5 4 4 1 1 1 0 0 0 0
4	9 8 8 6 5 3 1 0
3	9 8 7 7 6 5 4 4 3 2
2	3 2

stem: tens and ones digits
leaf: decimals digits

No.

Date . .

b.) The typical flow rate is "7.5".

c.) The display appears to be spread out.

d.) It does not appear to be symmetric actually. It seems like skewed.

e.) "18.9" seems like an outlying data.

Ex.20

a.)

Stem	leaf
5	850 770 700 320 220
4	770 390
3	870 380 350 330 150 150 060
2	730 700 460 400 400 320 250 120 109 100
1	890 850 670 419 320 280 250 240 120 050 000
0	960 960 540 530 510 500 450 396 360 340 240 100

Stem: thousands digit

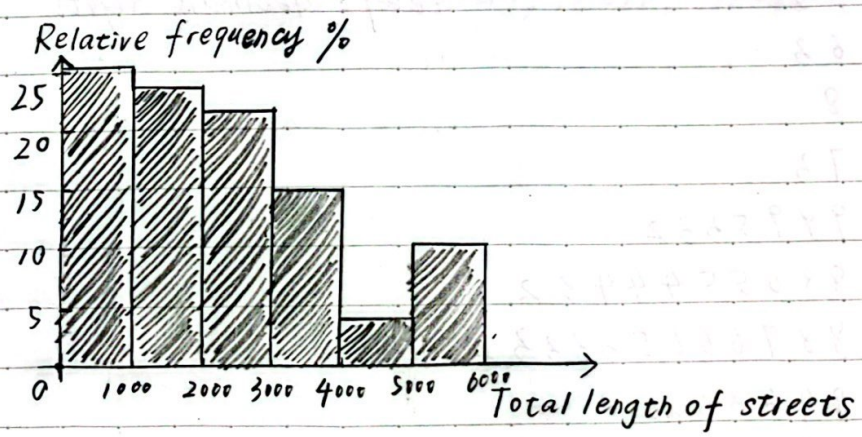
leaf: hundreds digit

Features: ① It seems not very symmetric but kind of skewed.

② It has only one peak at "0"

③ Seems no outlying value exists.

b.) histogram:



Proportion: less than 2000: (11+12)/47 × 100% = 48.9%.

The shape of the histogram is "bimodal" and kind of positively skewed.

No.

Date . .

Section 1.3

Ex.34

a.) $\bar{U} = (6.0+5.0+11.0+33.0+4.0+5.0+80.0+18.0+35.0+17.0+23.0)/11 \approx 21.55$

$\bar{F} = (4.0+14.0+11.0+9.0+9.0+8.0+4.0+20.0+5.0+8.9+21.0+9.2+3.0+2.0+0.3)$

$/15 = 8.56$

Comparison: $\bar{u}$ is larger than $\bar{F}$, which means the environment in farm home is better than urban home's.

b.) $\tilde{U} = 17$, $\tilde{F} = 8.9$, & $\tilde{U}$ is still larger than $\tilde{F}$.

Because there exists a extremely large value "80" among them

b. c.) $\bar{U}' = (6.0+5.0+11.0+33.0+5.0+18.0+35.0+17.0+23.0)/9 = 17$

$\bar{F}' = (4.0+14.0+11.0+9.0+9.0+8.0+4.0+20.0+5.0+8.9+9.2+3.0+2.0)/13 = 8.2$

trimming percentage: As for urban: $(21.55-17)/21.55 \times 100\% = 21.1\%$

As for farm: $(8.56-8.24)/8.56 \times 100\% = 3.73\%$

The urban's trimmed mean is close to its median value, which becomes more accurate. As for farm's, it does change a lot.

Ex.4a

Sample median: 92

25% trimmed mean: 94.75

10% trimmed mean: 102.23

mean = 119.26

mean > 10% trimmed mean > 25% trimmed mean > median

No.

Date

Section 1.4

Ex.44

a) The sample range: $49.3 - 23.5 = 25.8$

b) $\bar{x} = (49.3 + 33.9 + 31.6 + 30.6 + 29.5 + 29.4 + 28.2 + 28 + 26.3 + 23.5)/10 = 31.03$

$S_{xx} = \sum_{i=1}^{10} (x_i - \bar{x})^2 \approx 443.80$

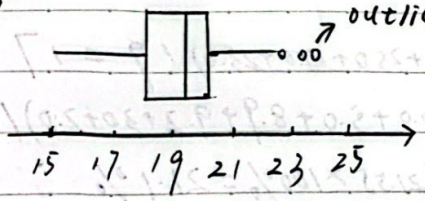
$s^2 = \frac{S_{xx}}{10-1} = 49.31$

c) The sample standard deviation: $s = \sqrt{s^2} = 7.02$

d) $S_{xx} = \sum_{i=1}^{10} (x_i - \bar{x})^2 = \sum_{i=1}^{10} x_i^2 - \frac{(\sum_{i=1}^{10} x_i)^2}{10} \approx 443.80$

$s^2 = \frac{S_{xx}}{10-1} = 49.31$

Ex.56



smallest: 15.3

lower fourth: 18

median: 19.2

upper fourth: 19.9

largest: 23.78

$f_s = 19.9 - 18 = 1.9$

$1.5 f_s = 2.85$

① spread is 1.9

② The data seems to be symmetric. The data's extent is [15.3, 23.78]