

Ch1 Homework 01

2021103523 黄毓乾

Ex.11

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

stem: tens
leaf: ones

This display shows that there is no score on the "7H", and score have a larger proportion in "8L", "7L", "6H" and "8H".

Ex.14 a)

Stem	Leaf
2	2 3
3	2 3 4 4 5 6 7 7 8 9
4	0 1 3 5 6 8 8 9
5	0 0 0 0 1 1 1 4 4 5 5 6 6 6 7 8 9
6	0 0 0 0 1 2 2 2 2 3 3 4 4 4 5 6 6 6 7 7 8 9 9 9 9
7	0 0 0 1 2 2 2 3 4 5 5 5 5 5 6 6 8
8	0 2 2 3 3 4 4 8
9	0 1 2 2 3 3 3 3 5 6 6 6 7 8 8
10	2 3 4 4 4 5 5 6 8 8
11	2 3 3 5 9 9 9
12	3 7
13	8
14	3 6
15	0 0 3 5
16	
17	
18	9

Stem: 1.0
Leaf: 0.1

- b) The representative flow rate is about 6.5
c) The data at "5-7" is highly concentrated, in other value, it spreads out.

d) No, it's not reasonably symmetric. This data is positively skewed.

e) The value "18.9" is an outlier, because it far away from the previous value more than two stem.

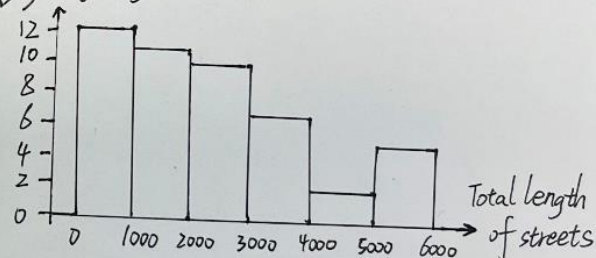
Ex.20

stem	leaf
0	1 2 3 3 3 4 5 5 5 5 9 9
1	0 0 1 2 2 2 3 4 6 8 8
2	1 1 1 2 3 4 4 4 7 7
3	0 1 1 3 3 3 8
4	3 7
5	2 3 7 7 8

stem: thousands
leaf: hundreds

The representative data value is somewhere is the "2000" (lower). And the display is bimodal, which is "0" and "5". The display has a positive skew.

b) Frequency



- ① The proportion of subdivisions with total length less than 2000: $\frac{12+11}{47} \times 100\% = 48.9\%$
 ② The proportion between 2000 and 4000: $\frac{10+7}{47} \times 100\% = 36.2\%$
 ③ The general shape is bimodal, which is "0-1000" and "5000-6000", and it has a positive skew.

Ex. 34

a) ① urban homes, $\bar{x}_u = 21.55$ EU/mg

② farm homes, $\bar{x}_f = 8.56$ EU/mg

\bar{x}_u is more than double \bar{x}_f .

$$\bar{x}_u = 2.518 \bar{x}_f$$

b) ① Urban homes, $\tilde{x}_u = 17.00$ EU/mg

② Farm homes, $\tilde{x}_f = \frac{8.97}{8.90}$ EU/mg

\tilde{x}_u is more than double \tilde{x}_f .

The mean and median endotoxin concentration for urban homes are so different, because there are some very large values, such as the extreme value of 80.0, which raise the mean but not the median.

c)

① For urban homes:

we deleting the smallest ($x=4.0$) and the largest ($x=80.0$) one, and get the trimmed mean of $\bar{x}_{tr} = \frac{153}{9} = 17$ EU/mg.

The corresponding trimming percentage: $\frac{1}{11} \times 100 \times 100\%$

$$= \frac{1}{11} \times 100 = 9.0909\%$$

Because the sample has a positive skew, the trimmed mean is less than the entire one.

The median (total) is equal the trimmed mean.

② For farm homes:

we delete the smallest and largest one, and get the trimmed mean of $\bar{x}_{tr} = \frac{107.1}{13} = 8.24$ EU/mg.

The corresponding trimming percentage:

$$100 \times \frac{1}{15} \times 100\% = 6.66\%$$

The trimmed mean is less than the mean and median of the entire sample.

Ex. 40

① because the data is ~~star~~ sorted from smallest to largest, so the sample median is 92. ($\frac{91+93}{2} = 92$).

② find 25% trimmed mean

we delete 11, 14, 20, 23, 31, 36, 39 and 44, 47, 50, 59, 61 and 148, 158, 161, 168, 184, 206, 248, 263, 289, 322, 388, 513.

$$25\% \text{ trimmed mean} = \frac{65+67+68+\dots+123+136+139+141}{50-(12 \times 2)} = 95.38$$

③ find 10% trimmed mean:

we delete 11, 14, 20, 23, 31 and 263, 289, 322, 388, 513.

$$10\% \text{ trimmed mean} = \frac{36+39+\dots+206+248}{50-(5 \times 2)}$$

$$= 102.225 \approx 102.2$$

$$\text{④ Sample mean} = \frac{11+14+20+\dots+388+513}{50}$$

$$= 119.26$$

Ex. 44

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

b)

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	x_i^2	
29.5	-1.53	2.3409	870.25	
49.3	18.27	333.7929	2430.49	
30.6	-0.43	0.1849	936.36	
28.2	-2.83	8.0089	795.24	
28.0	-3.03	9.1809	784.00	
26.3	-4.73	22.3729	691.69	
33.9	2.87	8.2369	1149.21	
29.4	-1.63	2.6569	864.36	
23.5	-7.53	56.7009	552.25	
31.6	0.57	0.3249	998.56	
310.3	0	443.801	10072.41	Total

$$\bar{x} = 31.03$$

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = \frac{443.801}{9} = 49.3112$$

$$c) S = \sqrt{s^2} = \sqrt{49.3112} = 7.0222$$

$$d) s^2 = \frac{\sum_{i=1}^n x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1} = \frac{10072.41 - \frac{(310.3)^2}{10}}{9} = 49.3112$$

Ex. 56

median = 19.2%

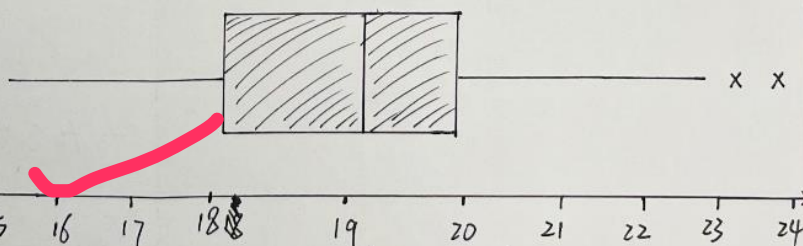
Smallest $x_i = 15.30$

Largest $x_i = 23.78$

Lower fourth = $\frac{18.00 + 18.68}{2} = 18.34$

Upper fourth = $\frac{19.90 + 19.62}{2} = 19.76$

There are two outliers 23.25% and 23.78%, which two port wines is very abnormally high alcohol content.



A+