Week 02 Assignment

Problem 1.

A list of 5 pulse rates is 70, 64, 80, 74, 92. What is the median for this list?

a. 74

b. 76

c. 77

d. 80

Answer A.

The middle value of the sorted data: 64, 70, 74, 80, 92

Problem 2.

If the variance of a data set is correctly computed with the formula using n - 1 in the denominator, which of the following is true?

a. the data set is a sample
b. the data set is a population
c. the data set could be either a sample or a population
d. the data set is from a census
e. None of the above answers is correct.

Answer: A

Use the definition of sample variance.

Problem 3

Based on the following sample of ages (in months) of 18 children at a daycare:

18 19 22 22 24 24 25 26 28 29 29 30 31 32 35 36 36 42

The 25th percentile, 75th percentile and the interquartile range for this data set are:

a) 23.5 , 34, 11.5

b) 24.5, 33.5 , 9

c) 24, 32, 8

d) 24, 32.5 8.5

e) 23, 32, 9,

Answer: C

Calculate the two quartiles. The range = maximum - minimum

Problem 4.

A national random sample of 20 ACT scores from 2010 is listed below. Calculate the sample mean and standard deviation.

12 12 13 14 17 17 17 18 18 19 22 23 23 25 26 26 26 29 30 30

a. 20.50, 5.79
b. 20.50, 5.94
c. 20.85, 5.79
d. **20.85, 5.94**

Answer D. You can use IntroStatsApps (descriptive statistics) to get his answer.

Problem 5.

Calculate the mean number of children per family for the sample from the following table.

Number of children	Number of families
0	8
1	16
2	22
3	14
4	6
5	4
6	2

a. 1.91

b. 2.47

c. 3.14

d. 2.19

Answer D.

You need to understand the frequency table to find the mean. For example, 8 families have no child, 16 families have one child, and 22 families have 2 children, following this pattern, you can find the total number of children and the total number of families.

(8x0 + 16x1 + 22x2 + 14x3 + 6x4 + 4x5 + 2x6)/ (8+16+22+14+6+4+2) = 2.19.

Problem 6.

A list of 5 pulse rates is 70, 64, 80, 74, and 92. What is the median for this list?

a. 74

b. 76

c. 77

d. 80

Answer A. Using the definition of the median.

Problem 7

What measure of the center was most affected by the outlier?

A. mean
B. median
C. mode
D. IQR

Answer: A. Understand the definition of mean and other measures

Problem 8.

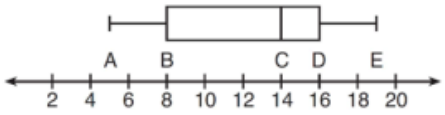
Which of the following lists all parts of the five-number summary?

A. Mean, Median, Mode, Range, and Total
B. Minimum, Quartile 1, Median, Quartile 3, and Maximum
C. Smallest, Q1, Q2, Q3, and Q4
D. Minimum, Maximum, Range, Mean, and Median

Answer B. Based on the definition of 5-number-summary.

Problem 9

The box-and-whisker plot shown below represents the number of magazine subscriptions sold by members of a club.



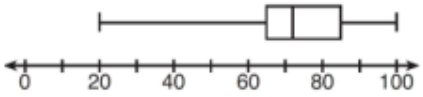
Which statistical measures do points B, D, and E represent, respectively?

A. minimum, median, maximum
B. first quartile, median, third quartile
C. first quartile, third quartile, maximum
D. median, third quartile, maximum

Answer: C. Understand the structure of the box-plot.

Problem 10

The box-and-whisker plot below represents the results of test scores in a math class.



What do the scores 65, 85, and 100 represent?

A. Q1 , median, Q3
B. Q1 , Q3 , maximum
C. median, Q1 , maximum
D. minimum, median, maximum

Answer: B. Understand the structure of the box-plot.

Problem 11

On an intelligence test with a mean of 100 and a standard deviation of 15, Jamie scored 85. What is Jamie's z-score?

a. -2
b. -1
c. 1
d. 2

Answer: B. Using the definition of z-score

Problem 12

Which of the following is not a measure of variability in a data set?

A. IQR (Inter Quartile range)

B. Variance

C. Upper quartile

D. Standard deviation

Answer: C. Understand the definition of variability measures.

Problem 13

If the mean is larger than the median when

A. The distribution is symmetric (bell-shaped)

B. The distribution is skewed to the right (positively skewed)

C. The distribution is skewed to the left (negatively skewed)

D. The distribution is approximately normal.

Answer: B. Understanding the definition of the three central tendency measures.

Summary of Week #2 Quiz

Types of Descriptive Statistics

Numerical Measures

comma separated numerical data

10,8.5,8.5,10,10,5.5,10,10,10,9,9.5,10,9,7

Measure Types

5-number summary and boxplot

Report bugs to C. Peng

Five Number Summary and Boxplot

The data values are:
10, 8.5, 8.5, 10, 10, 5.5, 10, 10, 10, 9, 9.5, 10, 9, 7.5, 10, 7.5, 8.5, 9, 10, 8.5, 10, 7, 10, 8, 10, 10, 9.5, 7.5, 9, 9.5, 10, 6.5, 8, 8, 9, 8, 8.5, 10, 10, 10, 8, 7.5, 7.5, 10, 8.5, 8.5, 6, 9.5, 10, 7.5, 8, 9, 10, 9, 10, 6.5, 9, 8.5, 9, 8, 9, 8

1. Five Number Summary :

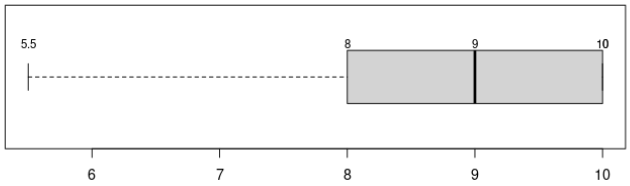
The five-number summary is use used to describe the shape of the distribution of a given numerical data. It consists of five numbers: minimum data value, first quartile, median, the third quartile, and the maximum dat value.

The five-number summary of this given data set is:

stats	value
Min.	5.50
1st Qu.	8.00
Median	9.00
3rd Qu.	10.00
Max.	10.00

2. Boxplot :

The boxplot is a geometric representation of the five-number summary. The boxplot of the given data set is given below.



Types of Descriptive Statistics

Table and Chart: Numerical Data

comma separated numeric raw data

10,8.5,8.5,10,10,5.5,10,10,10,9,9.5,10,9,7.

Summary Types

histogram

Boundary [must be equally spaced]

5.01, 6.01, 7.01, 8.01, 9.01, 10.01

Report bugs to C. Peng

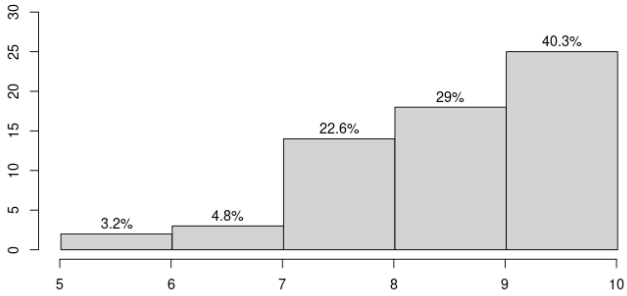
The input data values:

10, 8.5, 8.5, 10, 10, 5.5, 10, 10, 10, 9, 9.5, 10, 9, 7.5, 10, 7.5, 8.5, 9, 10, 8.5, 10, 7, 10, 8, 10, 10, 9.5, 7.5, 9, 9.5, 10, 6.5, 8, 8, 9, 8, 8.5, 10, 10, 10, 8, 7.5, 7.5, 10, 8.5, 8.5, 6, 9.5, 10, 7.5, 8, 9, 10, 9, 10, 6.5, 9, 8.5, 9, 8, 9, 8

The sorted input data values:

5.5, 6, 6.5, 6.5, 7, 7.5, 7.5, 7.5, 7.5, 7.5, 7.5, 8, 8, 8, 8, 8, 8, 8, 8, 8.5, 8.5, 8.5, 8.5, 8.5, 8.5, 8.5, 8.5, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9.5, 9.5, 9.5, 9.5, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10

Probability Distribution Histogram



Types of Descriptive Statistics

Table and Chart: Numerical Data

comma separated numeric raw data

10,8.5,8.5,10,10,5.5,10,10,10,9,9.5,1

Summary Types

Frequency Tables

Boundary [must be equally spaced!]

5,6,7,8,9,10

Report bugs to C. Peng

The input data values:

10, 8.5, 8.5, 10, 10, 5.5, 10, 10, 10, 9, 9.5, 10, 9, 7.5, 10, 7.5, 8.5, 9, 10, 8.5, 10, 7, 10, 8, 10, 10, 9.5, 7.5, 9, 9.5, 10, 6.5, 8, 8, 9, 8, 8.5, 10, 10, 10, 8, 7.5, 7.5, 10, 8.5, 8.5, 6, 9.5, 10, 7.5, 8, 9, 10, 9, 10, 6.5, 9, 8.5, 9, 8, 9, 8

The sorted input data values:

5.5, 6, 6.5, 6.5, 7, 7.5, 7.5, 7.5, 7.5, 7.5, 7.5, 8, 8, 8, 8, 8, 8, 8, 8, 8.5, 8.5, 8.5, 8.5, 8.5, 8.5, 8.5, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9, 9.5, 9.5, 9.5, 9.5, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10, 10

The class boundary is: 5,6,7,8,9,10

cut.data	Freq	midpts	rel.freq	cum.freq	rel.cum.freq
[5,6]	2	5.50	0.03	2	0.03
(6,7]	3	6.50	0.05	5	0.08
(7,8]	14	7.50	0.23	19	0.31
(8,9]	18	8.50	0.29	37	0.60
(9,10]	25	9.50	0.40	62	1.00