

Types of Displays

Bar Graph, Pie Graph, Dotplot, Stemplot, Histogram, Boxplot

1. How would you describe the overall shape of this distribution?

Skewed left

2. Where will the mean fall in respect to the median? (don't try to calculate it)

mean < median

(is less than)

3. What numerical measures would best describe it?
(mean/standard deviation OR median/IQR)

skewed → five # summary → median & IQR

4. Which display best represents categorical data?

pie chart or bar chart

5. What's the main difference between the data bar graphs and pie chart can display?

pie charts cannot display data over 100% but bar graphs can

6. Create a boxplot for the following data. Describe the shape. Calculate and determine if there are any outliers – be sure to show your work.

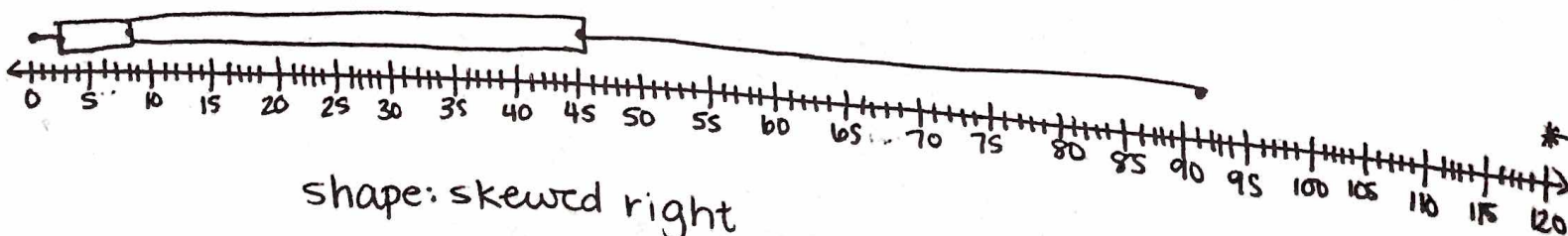
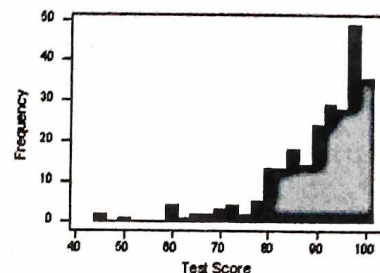
Number of text messages sent by teenagers in the last 24 hours:

0	8	2	28	21	7	4	1	23	88	7	1	22
8	120	82	0	91	42	17	2	3	48	7	52	

min = 0 Q1 = 2.5 med = 8 Q3 = 45 max = 120

LB = -61.25

UB = 108.75 → 120 is an outlier



Vocabulary**Median, IQR, 5 numbers summary, Mean, Standard deviation, Outliers, Variability**

7. A set of data has the following 5-number summary: Min = 15 $Q_1 = 78$ Median = 85 $Q_3 = 90$ High = 95

a. What's the interquartile range?

12

b. Is there an outlier? How do you know?

Since $15 < 60$, there is at least 1 low outlier

Since $95 < 108$, we know there are no high outliers

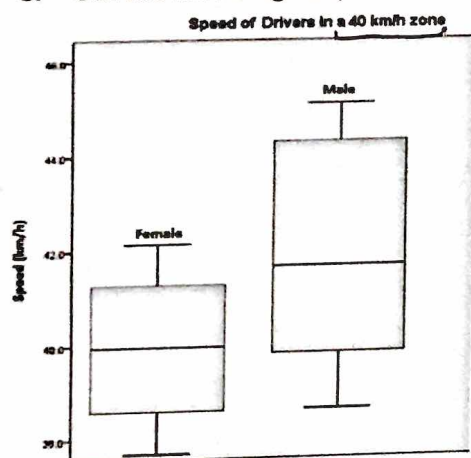
c. What percentage of the data lies above 90?

25%.

d. What percentage of the data lies below 85?

50%.

8. Use the following boxplot



a. Which one shows more variability?

male

b. Roughly what percentage of males are going exactly the speed limit or speeding?
40 km/hr

approximately 75%.

9. At a local university, the mean of the total number of classes taken by students earning their Bachelor's Degree is 35. What does this mean in this situation? the average # of classes is 35

a. If all students graduating with a Bachelor's degree took an extra 3 classes, what would happen to the mean? What would happen to the standard deviation?

mean: increase by 3

S.D.: no change

b. Would anything happen to the median? What about the IQR?

median: increase by 3

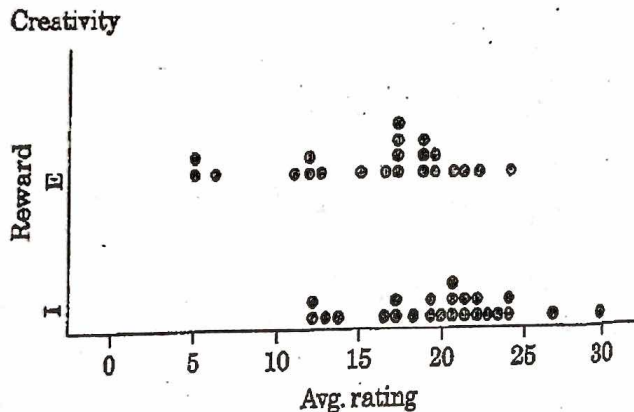
IQR: no change

Measures of center and spread

Median/IQR, Mean/Standard Deviation

10. Psychologists designed a study involving 47 experienced creative writers who were college students. Students were divided into two groups using a chance process (like drawings name from a hat). One group were given external reasons for writing (like public recognition/making money/pleasing their parents). The other group were given a list of statements about internal reasons (expressing yourself/enjoying working with words). Both groups were then instructed to write a poem about laughter. Each student's poem was rated separately by 12 different poets using a creativity scale.

Here's the data collected during this study:



- a) What's the median of the students who were given external reasons?

18

- b) What's the median of the students who were given internal reasons?

21

- d) Which one has the most variability? How do you know?

internal reasons → more spread out

- e) Did either one have one or more outliers? How do you know?

external reasons is more likely to have outlier(s) ↓

large gap between min & other cluster of scores

Statistical Reasoning
Unit 2 – Quick Review

Name:

Below is data from a sample of 28 people about what age they were when they got their first job. Use this data to answer all the questions below.

12	17	15	16	21	20	16
18	18	19	23	25	24	22
18	19	20	21	16	16	17
30	23	22	20	18	18	17

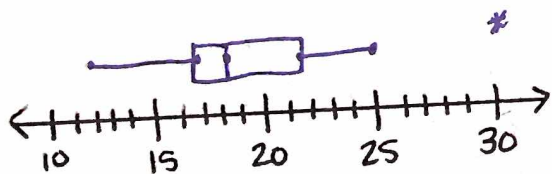
1) Find the five number summary.

$$\begin{aligned} \min &= 12 & Q3 &= 21.5 \\ Q1 &= 17 & \max &= 30 \\ \text{med} &= 18.5 \end{aligned}$$

2) Determine whether there are any outliers.

$$\begin{aligned} LB &= 10.25 \\ UB &= 28.25 \end{aligned} \quad \left. \vphantom{\begin{aligned} LB &= 10.25 \\ UB &= 28.25 \end{aligned}} \right\} 30 \text{ is an outlier}$$

3) Create a box plot.



4) Find the mean, standard deviation, range, and IQR.

$$\begin{aligned} \text{mean} &= 19.32 \text{ or } 19.3 & \text{range} &= 18 \\ \text{s.d.} &= 3.642 & \text{IQR} &= 4.5 \end{aligned}$$

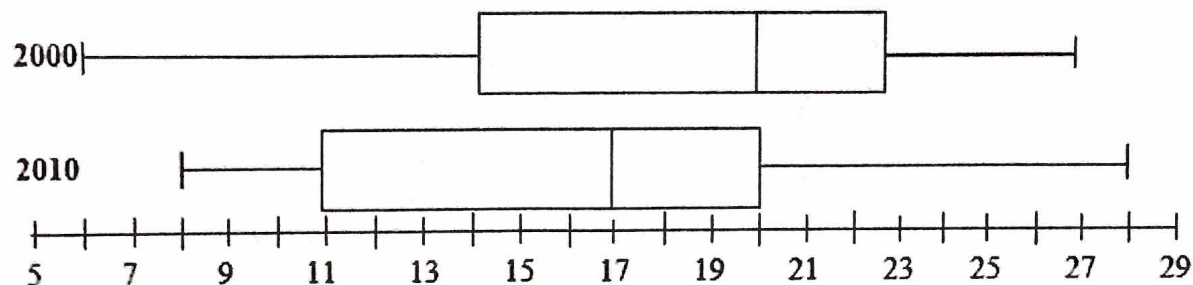
5) If we were to add 5 years to each age, how would the mean, median, standard deviation and IQR change?

$$\begin{aligned} \text{mean} &: \uparrow \text{ by } 5 & \text{s.d.} &: \text{no change} \\ \text{median} &: \uparrow \text{ by } 5 & \text{IQR} &: \text{no change} \end{aligned}$$

6) If we were to double each age, how would the mean, median, standard deviation and IQR change?

$$\begin{aligned} \text{mean} &: \text{double} & \text{s.d.} &: \text{double} \\ \text{median} &: \text{double} & \text{IQR} &: \text{double} \end{aligned}$$

The results from a pre-test for students for the year 2000 and the year 2010 are illustrated in the box plots below.



Identify each statement below as true or false.

- The median score for 2000 was higher than the median score for 2010.

True

- The scores from 2000 have less variety than the scores from 2010.

False

- 2000 had the lowest score.

True

- The scores from 2010 have a larger IQR.

True

Use the box plot from 2010 to answer the questions below.

What percent of students scored above 17?

50%

What percent of students scored between 11 and 20?

50%

What percent of students scored less than 11?

25%

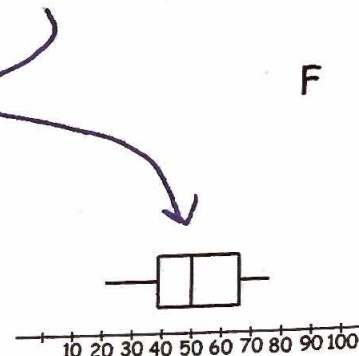
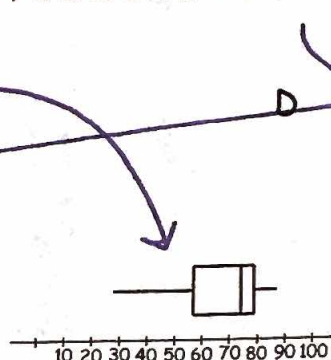
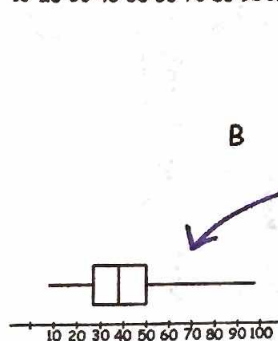
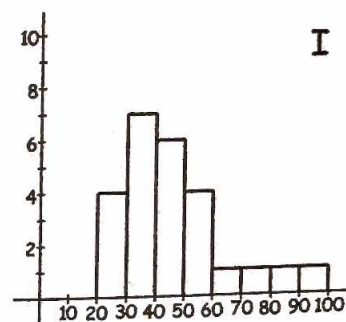
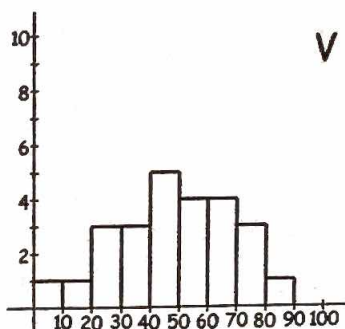
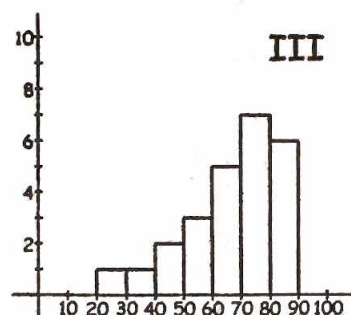
If there were 300 students enrolled in the class in 2010, how many scored above 20?

$$(300)(.25) = 75 \text{ students}$$

Statistical Reasoning
Chapter 2 Review – Part 2

Name: _____

Match the following histograms to the corresponding box plot based on shape.



Describe the relationship between the mean and the median for each of the following types of distributions.

For a symmetric distribution, the mean is approximately equal to the median.

For a skewed right distribution, the mean is greater than the median.

For a skewed left distribution, the mean is less than the median.

What are the two main measures of center that we use? When do we use each?

mean: when data is symmetric/no outliers

median: when data is skewed/has outliers

What are the two main measures of spread that we use? When do we use each?

standard deviation: when data is symmetric/no outliers

IQR (interquartile range): when data is skewed/has outliers

Given the following information, determine whether there are any outliers.

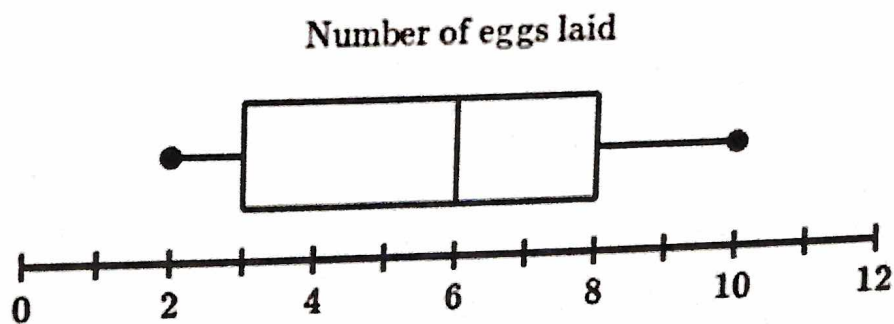
Min = 2 Q1 = 10 Med = 14 Q3 = 20 Max = 31

LB = -5 → nothing less than -5

UB = 35 → nothing over 35

No Outliers

Use the box plot below to answer the following questions.



What percent of chickens lay more than 2 eggs?

100%

What percent of chickens lay between 2 and 8 eggs?

75%

What percent of chickens lay at most 3 eggs?

25%

What percent of chickens lay at least 8 eggs?

25%

What percent of chickens lay between 3 and 8 eggs?

50%

If your data doubles, how does the mean change? How does the standard deviation change? How does the median change? How does the IQR change?

They all double

If your data all increases by 7, how does the mean change? How does the standard deviation change? How does the median change? How does the IQR change?

mean & median increases by 7

Standard dev. & IQR increases by 7

Identify the types of graphs below.

