AFM - Unit 2, Statistics

Notes 2-1 Central Tendency and Box Plots

Name ______

Date 2-14-2020 Period 2nd

Measures of Central Tendency

- Measures of central tendency, also referred as measures of center, refer to different types of averages.
- The most common measures of central tendency are mean, median, and mode.

MEAN

- The symbol for the mean is $\overline{X}$, which is read as x bar.
- Another symbol for the mean is μ, which is read as mew.

MEDIAN

- Median refers to the middle value of a set of data once it has been ordered from least to greatest. The median of a set of data with an even number of values is the mean of the two middle values.

1, 4, 11, 15

$\overline{X}$ = 7.5 = median

MODE

- Mode refers to the number that appears most frequently in a set of data. Data sets with two modes are said to be bi-modal. Sets have no mode when each item of the set has equal frequency

Ex. 1: Salary Data

Find the mean, median, and mode of the salaries for the corporate employees listed below. Which measure of central tendency appears to most accurately represent the set of data?

Allen: $40,000
Baker: 42,000
Chase: 59,000
Deitz: 60,000
Eckerd: 62,000
Francis: 65,000

How do extreme values (outliers) affect the measures of central tendency?

- Mean - $54,666.66
  $\overline{X}$
- Median - $59,500
  Med
- Mode - None

the mean is a good representation of the data because there is no outlier

Notes 2-1 STAT → Edit → Data in L1 → STAT → CALC → #1

Ex. 2: Backpack Weights

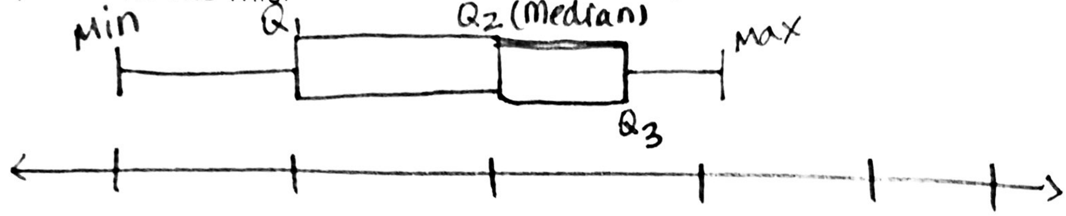
Owen is a member of the student council and wants to present data about backpack safety to the school board. He collects data on the weights of backpacks of 30 randomly chosen students. How much does the typical backpack weigh at Owen's school?

{ 3, 4, 4, 4, 6, 7, 7, 7, 7, 7, 8, 8, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10, 10, 13, 15, 15, 16, 17, 20, 33 }

Mean: 10.2 lbs   median: 9 lbs   mode: 10 lbs

Box and Whisker Plots and the Five Number Summary

Next we will look at Box and Whisker Plots (aka Box Plots). They are used to summarize a data set and to visually illustrate the variability (spread) of the data. A Box and Whisker plot looks like this:



The five parts of a Box and Whisker plot for a particular data set correspond to the Five Number Summary for that data set. The five numbers in the Five Number Summary are the Q1, Q2 (Median), Q3, Minimum, and maximum.

1st: Arrange the data in order and find the median. This separates the data into 2 groups.
2nd: Find the median of the 1st half and 2nd half of the data set.
Now your data set is divided into four groups, and each of these four groups is called a Quartiles. There are 3 points called Quartile points, (Q_1, Q_2, and Q_3) that denote the breaks in the data for each quartile.

- Q_1 is the median of the 1st half of the data.
- Q_2 is the median of the entire set of data
- Q_3 is the median of the 2nd half of the data
- The difference between Q1 and Q3 (i.e., Q3-Q1) is called the Interquartile Range (IQR)
- The difference between the maximum and minimum values is called the range

Box-and-Whiskers plots...

- can be drawn vertically or horizontally
- consists of a rectangular box with the ends, or Medians, located at the first and third quartiles

- the segments extending from the ends of the box are called whiskers
- the whiskers stop at the minimum and maximum values of a data set unless it contains outlier.

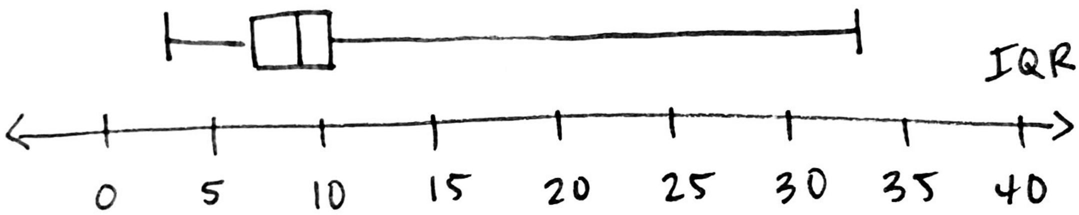
Outliers

- Outliers are outside values
- The technical definition of an outlier is a data point that is more than 1.5 of the interquartile range beyond the upper or lower quartiles. That is, any number less than Q_1 - 1.5(IQR) or greater than Q_3 + 1.5(IQR) is considered an outlier.
- Outliers are values represented by single points on a box plot.
- If outliers exist, each whisker is extended to the last value of the data set that is not an outlier.

Box and whisker with outlier

$Q_1 = 7$ med $= 9$ $Q_3 = 10$

min $= 3$ max $= 33$



IQR: $10 - 7 = 3$

Find values to determine if we have any outliers

$Q_1 - 1.5(IQR)$

$Q_3 + 1.5(IQR)$

$7 - 1.5(3) = 2.5$

$10 + 1.5(3) = 14.5$

Outliers: 15, 15, 16, 17, 20, 33