**11.3 Describing and Analyzing Data – Overview**

**Measures of Center**

**Mean (Average Value)**

* Simple, arithmetic average of the data.
  + Sum all numbers and divide by the sample size (*n*).

**Example 1 – Mean**

Data: 1, 5, 2, 9, 3

* Same calculation for the population and sample mean

(just different notation).

* + Sample mean = (pronounced ”x-bar”)
  + Population mean = (Greek letter mu)
* Mean is NOT a resistant measure.
  + This means it is heavily affected by outliers.

**Example 2 – Median**

Case 1 – Odd *n*

7 Obs: 10, 5, 6, 1, 3, 9, 8

Sorted:

Case 2 – Even *n*

8 Obs: 10, 5, 6, 1, 3, 9, 8, 3

Sorted:

**Median (Middle Value)**

* The middle value in an ordered list.
* Median IS a resistant measure.
  + NOT affected by outliers.

**Mode (Most Common Value)**

* The most frequently occurring value(s).
  + Unimodal data has one mode.
  + Bimodal data has 2 modes.
  + Multimodal data has more than 2 modes.
  + Can be no modes (every value is distinct).
* This is the only measure of center that can be used with categorical data.
  + Ex) Most common favorite color (can’t average this)

**Measures of Spread (Dispersion)**

**Range**

* Range = Max - Min
* Gives idea of the entire ”range” of values, how much distance do they span in total.
* Ex) Case 2 above: Range =

**Standard Deviation**

**Graphical user interface, text, application

Description automatically generated**

* Complex formula that measures the average distance that each data point is from the mean.

Chart, box and whisker chart

Description automatically generated

**Using your Calculator!**

Using TI-83/84 (and TI-30 XS MultiView / XIIS) to calculate mean, median, sample / population st dev.

Steps for the TI-30 XIIS

1. 2nd 🡪 STAT 🡪 1-VAR (Enter)

2. DATA

X1 = # (scroll down)

FRQ = 1 (for ALL Xs, scroll down)

X2 = # (scroll down)

…

3. STATVAR (scroll across)

4. To exit this menu: 2nd 🡪 EXIT STAT 🡪 Y

b) FRQ: L2

c) CALC

Steps for the TI-30XS MultiView

1. Data 🡪 Enter data in L1

2. 2nd 🡪 stat 🡪 1-Var Stats

a) DATA: L1

b) FRQ: ONE

c) CALC

Steps for the TI-83/84

1. Enter data: STAT → Edit → Enter data in L1

*(Demo dataset: 10, 23, 4, 6, 9, 3, 15, 6)*

2. Calculate: STAT → CALC → 1-Var Stats

* 1. List is L1.
  2. Leave FreqList blank.
  3. Calculate!

**A picture containing diagram

Description automatically generated**

Data here:

10, 23, 4, 6, 9, 3, 15, 6, 12, 11, 19, 10, 6, 8, 15

**Example 3**

Find the mean, median, mode and sample standard deviation of the following dataset.

* Data (7 obs): 35, 70, 31, 37, 65, 38, 38

**Other Considerations**

**Outliers**

* Data values that are extreme when compared to the rest of the data.
* Can significantly impact measures of center and spread.

**Types of Distributions**

Chart, histogram

Description automatically generated

Best measure of center:

**Measures of Relative Position**

**Percentiles**

* A **percentile** tells you the percent of observations/individuals you are higher than.
  + Interpreting example: You are told you scored in the 90th percentile on GRE. This means you have a score that is higher than 90% of all others that took the test.
  + Range from 0th to 100th percentile.
  + There is complement aspect to percentiles as well; for example, if you are the 80th percentile, there is \_\_\_\_\_\_ greater than you!
* Best way to remember!

A graph with arrows and a few words

Description automatically generated with low confidence

* Notation: Xth Percentile = Px

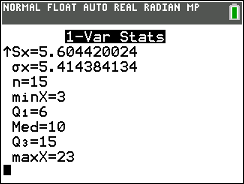
**5-Number Summaries and Boxplots**

* **Quartiles** are specific percentiles.
  + Q1 is the 25th Percentile.
  + Q3 is the 75th Percentile.
  + Q2 is the 50th Percentile = Median.
* **Inner Quartile Range (IQR)**
  + Another measure of variation, less informative than the standard deviation.
  + Uses quartiles to measure how far data is spread out around the median. Specifically, it measures the range of the middle 50% of the data
    - IQR = Q3 – Q1
  + Visualized very well in boxplots! It is the length of the box!
* **5-number summary**
  + Min, Q1, Med, Q3, Max 🡪 Points of a boxplot

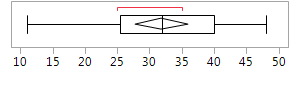


**Example 4**

1. Using this output from a 1-Var Stat, what is the IQR?



1. Find the IQR from this boxplot.



**Example 5**

1. Calculate the 5-number summary of the following dataset (20 numbers):

38, 33, 5, 5, 47, 29, 24, 42, 3, 18, 30, 46, 25, 44, 40, 42, 39, 44, 29, 13

1. Draw a boxplot based on the 5-number summary from (a).