

ACTIVITY SHEET: DESCRIBING DATA NUMERICALLY

This activity sheet includes exercises to assess students' understanding of important concepts presented in the *Describing Data Numerically* lesson.

Describing Data *Numerically*

The data for these exercises are in the Minitab file ***DescribingDataNumerically_Activity.mtw***.

Exercise 1

Below is a random sample of $n = 5$ bowling scores for "Twinkle Toes" Fred Flintstone from his Slate Rock and Quarry Company bowling league games.

212, 232, 250, 204, 261

- (a) By hand, calculate the **sample mean** \bar{x} of these scores. Write down the formula that you use.
- (b) By hand, calculate the **sample median** m of these scores.
- (c) By hand, calculate the **sample range** R of these scores.
- (d) By hand, calculate the **sum of the squared deviations** from the mean of these scores. Write down the formula you are using. Note: If you didn't square the deviations, what would the sum of just the deviations be?
- (e) Calculate the **sample variance** s^2 of these scores by using the sum of the squared deviations from part (d).
- (f) Calculate the **sample standard deviation** s for the scores.
- (g) Verify the statistics you computed in parts (a) – (c), (e) and (f) using Minitab. Put the five bowling scores in a column in Minitab, such as C1, and name it "Bowling Scores."

- 1 Choose **Stat > Basic Statistics > Display Descriptive Statistics**.
- 2 In **Variables**, enter 'Bowling Scores'.
- 3 Click **Statistics** and check **Mean, Standard deviation, Variance, Median, Range, and N total**.
- 4 Click **OK** in each dialog box.

(h) Suppose we randomly select another one of his scores from the past month; it is 101. In Minitab, calculate sample statistics for the mean, median, range, variance, and standard deviation using this additional data value. Provide their values below.

(i) We find out later that his score of 101 was from a day when Fred was sick from eating too many Brontosaurus burgers. In general, do outliers have a larger impact on the **mean** or **median**?

(j) Given these 6 data points, what additional 7th data point could you add that would keep:

- The median at the same value?
- The mean at the same value?

Exercise 2

Go to the <http://www.freeonlinegames.com/game/sheep-reaction> website to collect some reaction time data. You shouldn't need to sign up for an account.

Start with a practice round of 5 sheep to get used to the game. If you shoot at a sheep before it begins running, you will be penalized with a score of 3 seconds. After you tranquilize 5 sheep, the game will show your 5 reaction times and the average for those 5 times.

In a Minitab worksheet, label a column "Reaction Times." Play the game 2 rounds in order to collect 10 reaction times (in seconds). Type your 10 reaction times (and not the averages) into the column "Reaction Times."

(a) By hand or with Minitab, calculate sample statistics for the mean, median, range, variance, and standard deviation using the reaction time data. See **Exercise 1** for instructions on how to compute these values in Minitab.

(b) Look at the mean and the median for your data. What is the relationship between the mean and the median (e.g. is the mean greater than the median)? Provide a brief explanation as to why you think this has happened.

Exercise 3

The following table shows the running times (in minutes) of a sample of videotape versions of $n = 22$ movies directed by Alfred Hitchcock. Movies are listed in alphabetical order. The data are in the Minitab columns "Hitchcock Movies" and "Running Times."

Film	Time	Film	Time
The B irds	119	P sycho	108
D ial M for Murder	105	R ear Window	113
F amily Plot	120	R ebecca	132
F oreign Correspondent	120	R ope	81
F renzy	116	S hadow of a Doubt	108
I Confess	108	S pellbound	111
The M an Who Knew Too Much	120	S trangers on a Train	101
M arnie	130	T o Catch a Thief	103
N orth by Northwest	136	T opaz	126
N otorious	103	U nder Capricorn	117
The P aradise Cane	116	V ertigo	128

(a) By hand or with Minitab, calculate sample statistics for the mean, median, range, variance, and standard deviation using the running time data. See **Exercise 1** for instructions on how to compute these values in Minitab.

(b) The sample range is almost 1 hour. What data point is the main influence behind this large range?

(c) When plotted, the movie running times are fairly symmetric. How can you tell this without graphing the running times?

Exercise 4

The following are the top 20 player salaries (in millions of dollars) for 2014-15 NBA teams, according to ESPN. They are ordered from largest to smallest.

Reference: <http://espn.go.com/nba/salaries>

23.5	23.4	23.2	22.5	21.4	20.6	20.6	20.1	19.8	19.3
19.0	18.9	17.7	16.5	16.0	15.9	15.8	15.7	15.7	15.7

(a) Using only the data above (without Minitab), what is the median salary of the top 20 player salaries?

(b) What is the mode salary of the top 20 player salaries?

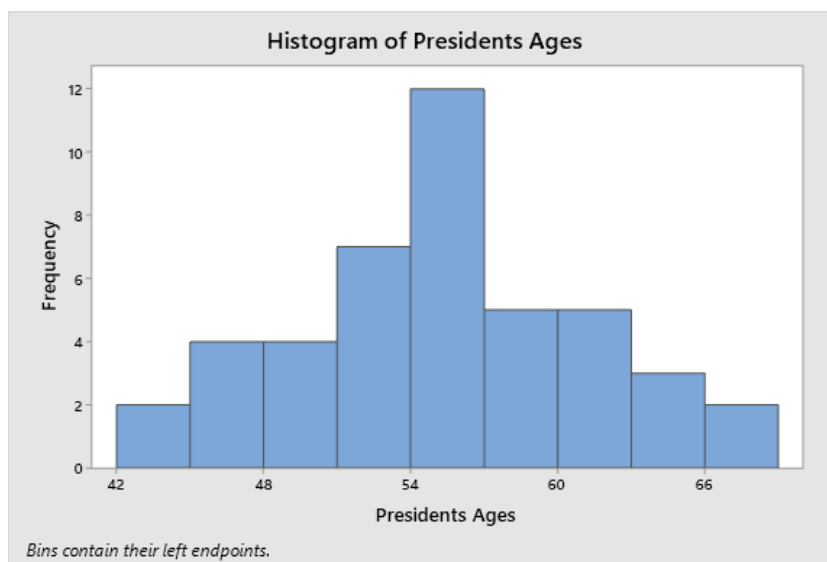
Exercise 5

Below are the ages at which U.S. presidents began their first terms, increasing in order from George Washington to Barack Obama, with Grover Cleveland serving 2 nonconsecutive terms. This is an entire population of data, not a sample. This data is in column "Presidents Ages" in the Minitab worksheet.

57	61	57	57	58	57	61	54	68	51	49	64	50
48	65	52	56	46	54	49	51	47	55	55	54	42
51	56	55	51	54	51	60	62	43	55	56	61	52
69	64	46	54	47								

(a) We determine the mean, median, and mode of a population in the same way that we determine the mean, median, and mode of a sample. By hand or in Minitab, find the mean, median, and mode ages of Presidents when they began their first term in office.

(b) Below is a histogram displaying the ages at which U.S. presidents began their first term. Considering the results that you obtained in part (a), is the shape and center of the histogram surprising to you?



(c) Using only the histogram from part (b), can you determine the age of President John F. Kennedy when he first took office? Why or why not?