

General Instructions: All answers should be in full sentences. Answers that are not sufficiently justified or explained will not be given full credit (see scoring rubric).

Directions, Exercise 1: Enter and analyze the following data set in Excel. You should complete all of the following steps:

- a. Add a column to the table for relative frequency. Relative frequencies should be expressed as percentages to the nearest hundredth of a percent (e.g., 3.75%)
 - b. Make a relative frequency bar graph and/or histogram, using percents.
 - c. Discuss the shape of the data set and features of the distribution; include at a minimum: symmetry or skewness, direction of any skew, location and number of peaks, and a summary of where most of the values in the distribution are found.
 - d. Predict whether the median is higher or lower than the mean, and tell why.
1. The table below gives the frequencies of birth weights for nearly 4 million infants born in the United States in 1986.

Birth Weight (grams)	Number of Newborns
0 – 499	4,843
500 – 999	17,487
1000 – 1499	23,139
1500 – 1999	49,112
2000 – 2499	160,919
2500 – 2999	597,738
3000 – 3499	1,376,008
3500 – 3999	1,106,634
4000 – 4499	344,390
4500 – 4999	62,769
5000 – 5500	8,236

Directions – Exercise 2: Enter and analyze the following data set in Excel. You should complete all of the following steps.

- a. Enter and sort the data in Excel.
 - b. Create 3 separate histograms of the data: one with bin size 10, one with bin size 5, and one with a different bin size of your choice.
 - c. Discuss the shape of the data set and features of the distribution; include at a minimum: symmetry or skewness, direction of any skew, location and number of peaks, and a summary of where most of the values in the distribution are found.
 - d. Compute the median and the first and third quartiles.
 - e. Compute the IQR and determine if there are any outliers.
If so, which values are the outliers?
 - f. Calculate the mean and standard deviation.
2. For all former US presidents who are deceased, below are the ages at which they died:

67	90	83	85	73	80	78	79	68	71	53	65	74
64	77	56	66	63	70	49	57	71	67	58	60	72
67	57	60	90	63	88	78	46	64	81	93	93	

Directions, Exercises 3-4: Enter and analyze the following data sets in Excel. You should complete all of the following steps:

- a. Make a summary table including a column for relative frequency. Relative frequencies should be expressed as percentages to the nearest hundredth of a percent (e.g., 3.75%)
 - b. Make a relative frequency bar graph and/or histogram, using percents.
 - c. Discuss the shape of the data set and features of the distribution; include at a minimum: symmetry or skewness, direction of any skew, location and number of peaks, and a summary of where most of the values in the distribution are found.
 - d. Find the median and lower and upper quartiles of the data set
 - e. Determine whether the data set has outliers. If so, name them.
 - f. Calculate the mean and standard deviation.
3. Researchers conducted a community household survey that asked, among other things, number of rooms per household. The following table shows the responses.

3	2	4	4	1	6	3	6	6	6	7	6	4	3	6	6	4	3	6	5
5	7	5	2	7	5	4	6	8	4	5	7	5	5	2	5	8	6	6	3
6	1	5	6	5	4	3	4	3	6	5	4	7	7	7	5	5	6	7	6

4. Over 700 bus drivers employed by public corporations participated in a survey to determine the number of traffic accidents each bus driver was involved in during a 4 year period. The results are given below:

Number of Accidents	0	1	2	3	4	5	6	7	8	9	10	11
Bus Drivers	117	157	158	115	78	44	21	7	6	1	3	1

Directions – Exercise 5: Enter into a spreadsheet and analyze the following data set.

Remember to type all your responses as text in a worksheet in your spreadsheet file with the rest of your work. You should complete all of the following steps.

- a. Enter the data into your worksheet.
 - b. Compute a 5 Number Summary.
 - c. Check for outliers using the appropriate Box-Plot; list any outliers that you find.
 - d. List the mean, range, and standard deviation of the data set. How far is the mean from the median, and which is larger?
 - e. Without making a histogram or stem-and-leaf plot, tell whether you think data is skewed left, skewed right or neither and explain your reasoning.
5. The US Department of Agriculture maintains a record of the sugar content in breakfast cereals. For 34 cereals marketed primarily as children's cereals, the table below shows the proportion of sugar in the cereal as a percentage of the cereal ingredients.

56.0	54.6	48.0	46.0	46.0	45.6	45.5	44.0	43.7
43.5	43.3	43.0	42.6	42.5	42.2	41.0	41.0	40.7
40.1	40.1	40.0	39.5	38.0	37.2	37.0	35.9	33.3
32.2	26.0	22.0	21.0	7.8	4.8	3.0		