

College Score Card Data

18) The College Scorecard (<https://collegescorecard.ed.gov/>), a Department of Education website, gives data on various variables regarding a school's performance. One may use this site to search for information by programs/degrees, location, size, religious affiliations, etc. The College Score Card dataset selected represents a subset of data from the College Score Card website. The dataset represents the colleges/universities in the Southern Association of Colleges and Schools Commission on Colleges (SACS-COC) region. SACS-COC is the regional body for the accreditation of degree-granting higher education institutions in the Southern states. It serves as the common denominator of shared values and practices among the diverse institutions in Alabama, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia that award associate, baccalaureate, master's, or doctoral degrees.

A link to the data set is provided in the Checkpoint Assignment #1 posting along with a description of each of the variables contained in the data set.

Based on this description of the data and research setup, answer the following questions.

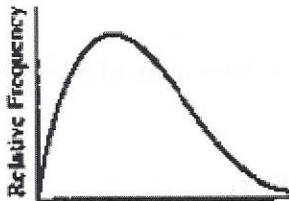
- Would this be considered an observational or experimental research study? Explain your choice. (2 points)
- Could you expect the conclusions you reach in analyzing this data to apply to all colleges/universities in the United States? Explain why or why not. (2 points)
- Determine the level of measurement for each of the 20 variables listed below. (Ordinal, Nominal, Interval, or Ratio) (Copy the table onto your paper) (20 points)

Variable	Level of Measurement	Explanation
INSTNM		
CITY		
STABBR		
CONTROL		
CCBASIC		
ADM_RATE		
SAT_AVG		
UGDS		
UGDS_WHITE		
UGDS_BLACK		
UGDS_HISP		
UGDS_ASIAN		
UGDS_AIAN		
UGDS_NHPI		
PPTUG_EF		
COSTT4_A		
INEXPFTE		
PFTFAC		
C150_4		
RET_FT4		

14) What is the difference between a bar chart and a histogram?

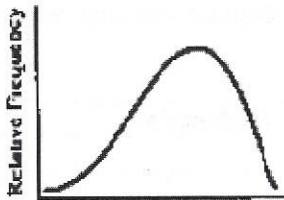
- A) There is no difference between these two graphical displays.
- B) The bars in a bar chart are all the same width while the bars of a histogram may be of various widths.
- C) The bars in a bar chart may be of various widths while the bars of a histogram are all the same width.
- ☒ D) The bars on a bar chart do not touch while the bars of a histogram do touch.

15) Describe the shape of the distribution.



- ☒ A) skewed to the right
- B) symmetric
- C) uniform
- D) skewed to the left

16) Describe the shape of the distribution.



- ☒ A) skewed to the right
- B) uniform
- C) skewed to the left
- D) symmetric

In exercises 17 - 19 you must show all of the required work to receive full credit. Follow the directions indicated on each problem.

17) The preschool children at Elmwood Elementary School were asked to name their favorite color. The results are listed below.

orange	yellow	purple	purple	red
red	red	pink	red	blue
pink	blue	purple	purple	purple
blue	yellow	purple	purple	yellow
orange	pink	red	red	blue
yellow	red	red	red	pink

- a) **By hand**, construct one frequency distribution table containing both the frequency distribution and the relative frequency distribution for each color (Put the colors in alphabetical order). Make sure that you give the table an appropriate title. Round relative frequency three decimal places as needed. (8 points)
- b) **Using a statistical program** (Excel, StatCrunch, etc.) construct a pie chart for the data. Make sure to add an appropriate title to the pie chart. Include a printed copy of the pie chart with your assignment. (7 points)

- 19) A random sample of 50 college students was selected. Their weights (in pounds) were obtained from their medical records. The resulting data are listed below.

Weights (in lbs) College Students

98	150	108	158	162	112	118	167	170	120
177	186	191	128	135	195	137	205	190	120
188	176	118	168	115	115	162	157	154	148
101	143	145	108	155	110	154	116	161	165
142	184	120	170	195	132	129	215	176	183

Find the following values for the data set

- Maximum data value (1 point)
- Minimum data value (1 point)
- Range of the set of data (1 point)
- Class width for the set of data used to construct a frequency distribution using 10 classes. (2 points)
- By hand**, construct a frequency distribution table for the data above, using 10 classes. The table should contain the following columns. (Copy the table on your paper—Make sure you give the table a title) (42 points)

Class	Frequency	Relative Frequency	Class Midpoints

You will need to enter the data set for this problem into a statistical program (Excel, StatCrunch, etc.).

- Using technology**, construct a frequency histogram for the data. Be sure to use a starting point equal to the minimum data in the set and use the same class width as you used in the frequency table. Be sure to print out the histogram and correct the labels and title as needed to properly reflect the data. (15 points)
- By hand**, organize the data above using a stem-and-leaf plot. Make sure you order the leaves on each of the stem-and-leaf plots from smallest to largest, include a key and a title for the stem-and-leaf plot. (15 points)
- Describe the distribution of the data. (2 points)

- d) Complete the frequency table below for Control, round relative frequency three decimal places. (Copy the table to your paper) (6 points)

Control	Frequency	Relative Frequency
Public		
Private Non-Profit		
Private For-Profit		

Prepare a frequency distribution for the enrollment of undergraduate certificate/degree-seeking students (UGDS column) data for the 2006 colleges/universities reporting this information (this will exclude any NULL values where this information was not submitted). Use a starting point of 1 and determine what class width will give you 10 classes.

- e) Calculate the Class Width (show calculations) (2 points)

This class width is _____.

- f) Now fill out the frequency table below, round relative frequency three decimal places: (Copy the table on your paper—Make sure you give the table a title) (32 points)

Class	Frequency	Relative Frequency
1 -		

- g) Use a statistical program (Excel, StatCrunch, etc.) to construct a histogram of the UGDS data. Be sure to use a starting point of 1 and the same class width as you used in the frequency table. Be sure to print out the histogram and correct the labels as needed to properly reflect the data. (15 points)
- h) Describe the shape of this distribution using proper statistical terminology. (2 points)
- i) Which classe(s) are the largest? Estimate the average enrollment of undergraduate certificate/degree-seeking students based to the histogram. (3 points)