Homework for Chapter 3: Describing Variables

1. What is a va­riable?
2. For each of the following variables, what types of variables are they (continuous, count, ordinal, categorical, qualitative)?
   1. Age
   2. Gender
   3. The number of times that the President has tweeted in the past day
   4. Income
   5. Number of Instagram posts about statistics in the past month
   6. The number of unemployment claims filed in US last week
   7. The university or college that a student attends
   8. A therapist’s written assessment of a patient’s symptoms of depression
   9. Whether a soccer team is in its league’s A division (highest), B division (next highest), or C division (lowest)
3. Which of the following provides a description of the probability that each possible value of a variable will occur?
   1. Variation
   2. Distribution
   3. Range
   4. Mean
4. Below is a frequency table depicting the salaries of Economics professors employed at a university. The column named Salary contains the salary, and the column named Frequency contains the number of professors who earn the stated salary.

|  |  |
| --- | --- |
| Salary | Frequency |
| $85,000 | 5 |
| $90,000 | 4 |
| $100,000 | 1 |
| $120,000 | 2 |
| $125,000 | 3 |
| $130,000 | 2 |

* 1. Calculate the average salary earned by professors in the Economics department.
  2. Calculate the median.
  3. Calculate the minimum and maximum.
  4. Calculate the interquartile range.

1. Which of the following commonly represents the *truth* we are trying to estimate in statistics?
   1. English/Latin letters like and
   2. Modifications of English letters like
   3. Greek letters like and
   4. Modifications of Greek letters like
2. Which of the following reduces impact of extreme values if the distribution of a variable is skewed?
   1. Log transforming the variable
   2. Mean centering the variable
   3. Standardizing the variable
   4. Ignoring the variable
3. Order the following distributions from the one with lowest to the one with highest variability:  
   Shape, histogram

   Description automatically generated
4. The following graph represents the final exam scores for 1000 students who took an Introduction to Statistics course at a university.  
   Chart, line chart

   Description automatically generated
   1. Describe the distribution.
   2. Is there skewness to the data?
   3. Would the mean or the median be a better measure to describe the center of the distribution?
   4. What measure would you use to describe the variability in the distribution?
5. Which of the following statement is correct regarding a *left-skewed distribution*?
   1. The mean is greater than the median
   2. The mean is less than the median
   3. The mean and median are equal
6. The table below shows data on how many students in a university are Freshman, Sophomore, Junior and Senior.

|  |  |
| --- | --- |
| Student Standing | Frequency |
| Freshman | 1000 |
| Sophomore | 1200 |
| Junior | 900 |
| Senior | 1500 |

* 1. For an individual student, what kind of variable is their student standing (continuous, count, ordinal, categorical, qualitative)?
  2. For the university itself, what kind of variable is the number of students with a given student standing (continuous, count, ordinal, categorical, qualitative)? (Yes, the answer *is* different from part a!)
  3. What kind of graph would best visualize this distribution for the university? Create the graph either by hand or with software.

1. Collect any continuous data from your classmates, family members, and/or friends. For example, you can collect data about their age or height.
   1. What is the distribution of the data? Describe the shape of the distribution. Summarize the data using measures like mean, median, range, standard deviation.
   2. What theoretical distribution may have generated the data that you sampled?
   3. Does your data give you a good idea about what the underlying theoretical distribution might be? Explain your reasoning.