

Homework Assignment Expectations

Bio 273: Biostatistics

- For each of the following variables, indicate what kind/level of variable it is. Also indicate if it's categorical or numerical. If it's numerical, indicate if it's discrete or continuous.
 - Class year (Sophomore, Junior, Senior) of the members of this class
 - Diagnosis of patients admitted to the Emergency Department at MGH
 - Weights (masses) of babies born at Winchester Hospital (MA) in 2007
 - Core temperature of green sea urchins collected in Cobscook Bay, Maine
 - Range of motion of the right elbow joint for members of the Suffolk Women's Hockey team
 - The number of steps it takes for a member of this class to get from the Sawyer Library to the Red Hat.
- A survey sponsored by the American Laser Centers included responses from 575 adults, and 24% of the respondents said that their face was their favorite body part (based on data from *USA Today*). What's wrong with this sample?
- A nutritionist wanted to understand how much soda Americans drink. They sampled 1473 Instagram users with the question "How much soda do you drink per day?" What's wrong with that sample?
- Open the `lynx` dataset in R using the function `data(lynx)`. What kind of variable is this? What is the most appropriate graph to represent those data? Also make that graph. Use the script demonstrated in class or the data visualization video to help to display the number of lynxes caught per year.
- Create a variable in R with these 10 measurements of lung capacity (in mL) from women who are 168 cm (5'6") in height. What kind of variable is this? What is the best graph to represent this sample? Make that graph. Use the combine function `c()` to create the object in R.

3233, 4650, 3466, 2766, 4400, 4000, 3300, 2800, 2900, 3800, 3100, 3533, 3267, 3600, 3600

6. What is this?

Type of cancer	Frequency
Skin	31
Breast	22
Brain	16
Cervical	15
Lung	8
Bone	4
Esophageal	1

Choices: Raw data A frequency table Two samples Seven samples histogram

7. A ____ is a frequency table used to describe outcomes of two or more categorical variables.

8. If you wanted to describe the relationship between two numerical variables like height and lung capacity, what would be the correct plot to make?

Choices:

Scatter plot

Box-and-whisker plot

Strip chart/dot plot

Histogram

Line chart

9. What's wrong with this graph? Setting aside your politics, how is it deceptive?



10. This plot is ostensibly about how often people communicate online. Assume that the data are correct and describe the flaws of this monstrosity.

