Objectives: This activity highlights the *statistical thinking* mindset, where we observe several ways to quantitatively summarize data and then draw justified, qualitative conclusions about a population. I **don't** expect you to have already mastered the skills and techniques used in this activity—we will continue to practice and develop them throughout the semester.

Assignment You will be randomly assorted into groups of 3 or 4. Once you are in your breakout room, introduce yourselves and **work together** to answer the following questions. Write your answers on your group's shared google doc (link will be posted in your Slack classroom channel).

Problem 1

In 2012, Trayvon Martin was shot to death by George Zimmerman in Florida. Zimmerman was later acquitted based on Florida's 'Stand Your Ground' law. This prompted the Tampa Bay Times to investigate whether the conviction rate of those who use the 'Stand Your Ground' defense varies by the defendant's race. Below are the 220 cases where the defense was used, along with some information about the defendant's race and the trial outcome.

Table 1

	Minority	White
Aquitted	60	86
Convicted	29	45
Total	89	131

(a) Explain why it would be misleading to simply compare the conviction counts of 29 and 45 in order to conclude that defendant race affects conviction rate.

(b) Calculate the proportion of white and minority defendants who were convicted and compare the proportions. Which group is convicted at a higher rate?

(c) What other factors may we want to consider before concluding that whites are convicted at a higher rate than minorities?

(d) One possible factor we have access to is the race of the victim. Below you will find a table that includes this information as well.

Table 2

	Minority Defendant Convicted	Minority Defendant Acquitted	White Defendant Convicted	White Defendant Acquitted
Minority Victim	19	45	5	19
White Victim	10	15	40	67

Explain how to obtain the data in **Table 1** using the data in **Table 2**.

(e) Explore the data in **Table 2** by computing several more proportions. In particular compute the conviction rate for minority defendants with minority victims and for minority defendants with white victims. Compute similar conviction rates for white defendants. Then compute the conviction rate among all defendants when the victim is a minority, and when the victim is white.

(f) Based on the proportions computed in the previous problem, why might the direction of the trend reversed itself when including the victim's race? (from a statistical perspective, not necessarily a sociological perspective)

Problem 2

Statistics professor Tom Short investigated whether cancer pamphlets were written at the appropriate level to be understood by cancer patients. The data collection included two components: a sample of 63 cancer patients were given a reading test, and a sample of 30 cancer pamphlets were analyzed for their readability. Both data sources were scored on the same scale, where an 8 corresponds to an 8th grade reading level. Below is a graphical summary of the data:

Graph 1
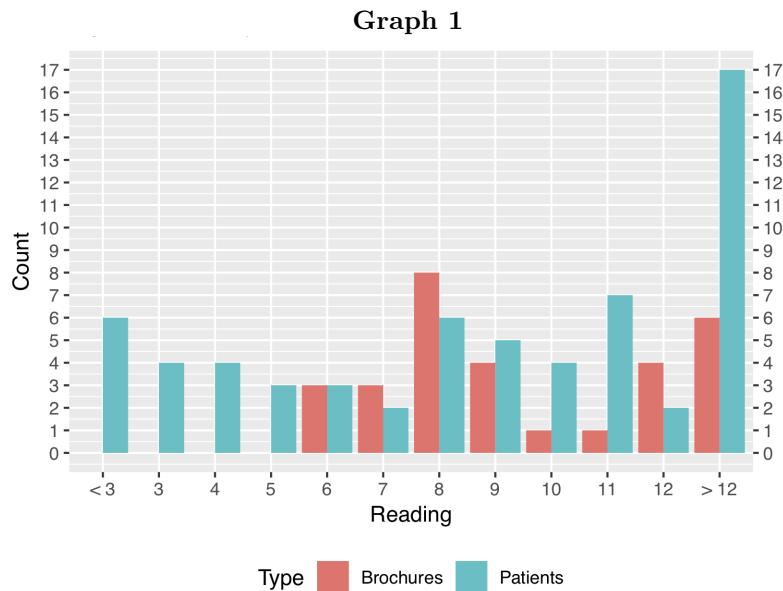
(a) Explain why we cannot compute the mean (average) reading level of a patient based on how the data is presented here.

(b) The **median** is another measurement of 'typical' value for a data set, and is a number greater than or equal to half of all values in the data set and less than or equal to half of all values in the data set. Calculate the median reading level for both patients and pamphlets.

(c) Compare the two medians. Does this suggest that patient and pamphlet reading levels are well-matched with each other?

(d) What proportion of patients have reading level below the reading level of the simplest pamphlet?

(e) The **distribution** of a data set describes the distinct values in the data set and how frequently they occur. **Graph 1** gives a visual depiction of the distributions of patient and brochure reading levels. Compare the shapes of the distributions for patient and brochure reading levels.

(f) Based on your observations, do you think cancer pamphlets are written at an appropriate level for patients? Why or why not?