**Directions**

In your group you are expected to collaborate on addressing each question. If you are the first person to post your ideas, do not fear! This is intended to provide you with a safe space to put your ideas on paper.

If you come to the document and there are ideas that you disagree with or would like to revise / clarify, feel free to use comments to discuss with your group members.

My hope is that each group member contributes equally, but that may look different for each person. I expect for each of you to edit and add to posts from the rest of your group. This **is not** an individual assignment, so please don’t write individual responses and fail to engage with your group.

Your final responses are due on Friday, January 8 by 11:59pm. You will be required to download this Google Doc as a Word file to submit in Canvas. Only one submission per group is required.

**Question(s)**

1. What do you believe are similarities and differences between Statistics and Mathematics?

|  |  |
| --- | --- |
| Differences | Similarities |
| * Math has more concrete, definite answers while statistics is more based on reasoning and answers are more abstract * Statistical examples can be cherry-picked or strategically unmentioned to promote a particular argument or worldview, while math is unassailable | * Use numbers * Both provide detailed observations about the world * Both lead to inferences using logistics * Can be applied in real life scenarios |

1. From your perspective, what is Statistics? Is there a difference between “statistics” and “Statistics”? If so, what?

Statistics is a field that collects data with the intended use to provide conclusions through analysis and interpretation of this said data. Generally, a statistical model or population will be studied to start the study on the data. There is a slight formality difference between statistics and Statistics. When statistics is thought of, data collection on sports teams is a familiar thought as it seems to be informal but may be for pure joy and entertainment for an individual to record statistics when at a game. This word could be shortened to stats. Contrary, Statistics may be the formal use of these ‘stats’ by an organization. A team’s management can assess how the team is doing in many different metrics such as crowd attendance, batting averages, and throwing velocity. A more formal approach may be useful.

1. Just because a variable has numeric values, does not mean it is a numeric variable. What are the different types of variables that can appear in a dataset? How can you determine if a variable is numerical versus categorical?

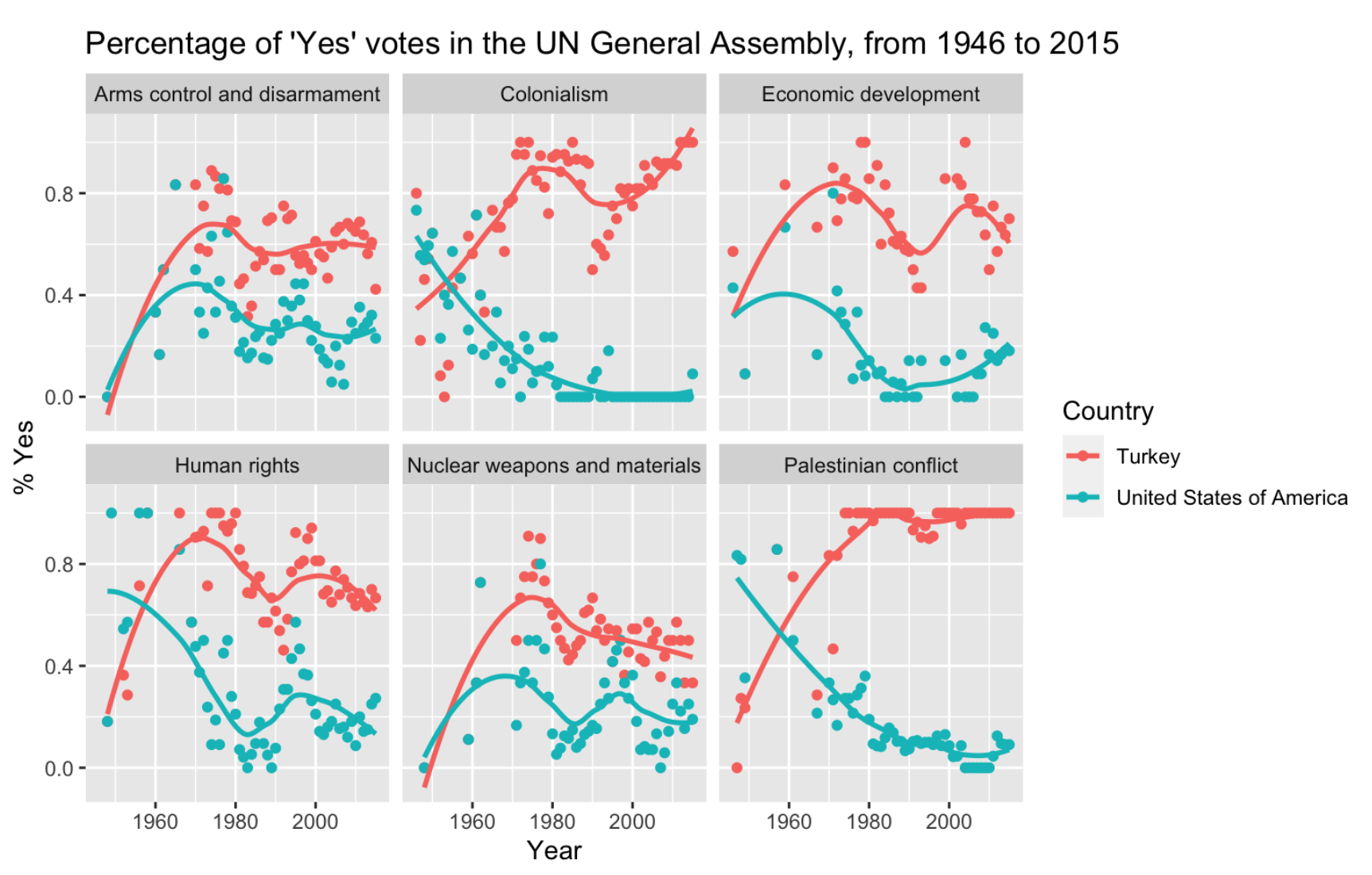
*The two main types of variables are numerical and categorical. A variable is numerical if it can have a wide range of values, and it makes sense to add, subtract, or take the average of those values. Categorical variables are represented by fixed categories or classifications and don’t necessarily have an order to them. Numerical variables can be continuous or discrete, and categorical variables can be ordinal or nominal.*

What are the differences between observational studies and experiments?

*In observational studies, the investigator does not have control over the exposure factor. The investigator also cannot randomly assign subjects to the conditions of an observational study.. These studies can’t show a causal connection. Experiments have explanatory and response variables, and they can show a causal connection. Experimental designs enable the investigator to control who is exposed to a factor of interest and can randomly assign subjects into the groups in the study.*

1. What are different methods for sampling from a population? Why would a researcher choose a stratified random sample instead of a simple random sample?

There are four methods for sampling from a population. They are simple random sampling, stratified sampling, cluster sampling and multistage sampling. In stratified random samples the population is divided into groups based on similarities, called strata. This ensures that each case has been included in the sample. It also protects against having a case that is not represented in the sample and an equal representation of the population.

1. Describe what you see in the plot below. 

The graphs represent 6 strata for policy priorities that represent ‘yes’ votes between Turkey and the United States of America in the UN General Assembly from 1946 to 2015. The policy priorities shown are Arms control and disarmament, colonialism, economic development, human rights, nuclear weapons and materials, and Palestinian conflict. All 6 strata are set upon the same large XY plane, with time ranging from 1960-2000 on the x-axis and percentage of “yes” votes on the y-axis. There is also a legend showing that Turkey’s response data correlates with the color red and the USA’s response data correlates with the color turquoise. The data consists of individual points on the graph along with a trendline for each nation showing the aggregate direction of their responses.