**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 |
| Statistics and mathematics are fundamentally different. Mathematics is based on a strict regime of calculations, and tends to be more abstract as it is commonly removed from practical applications.  Statistics uses validated testing methods in order to extract an answer that can be further interpreted in order to explain the probabilities of the outcome happening by random chance.  Mathematical computations come from a foundation of laws and theories, thus forming standardized equations, and statistics revolves around the foundation of uncertainty and chance. | Statistics and mathematics have to work together in order to determine a conclusion.  Both statistics and mathematics work with variables.  Statistics USES math to get to its conclusions. |

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

Statistics (with a capitalized “S”) refers to the course that we, as students, are enrolled in at Cal Poly. This Statistics course will further present the students with statistics from datasets, and ask us to interpret the meaning and present the data illustrating our outcome conclusion. This type of statistics (with a lowercase “s”) may be presented in quantitative values in order to analyze data and predict the probabilities of an outcome. These statistics may be presented in everyday life such as in a news article, a one-sheet for real estate, or presenting demographic information to make a statement on social media. Thus, this grammatical perspective may change the meaning of statistics-- one being raw data and the other being presented in a standard statistical way.

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?**

Numbers do not always hold numeric values associated with measurements such as an individual’s telephone number or bank account number. These are numeric values. Numeric values are used in standard mathematical calculations because the numbers represent an amount of some measurement. Numerical values can be discrete or continuous. Discrete numbers include non-negative, whole numbers. Contrastingly, continuous numbers include non-whole and negative numbers.

There may also be categorical data, which presents levels of which data can fall into. If these categories are usually ordered in a specific way they would be considered ordinal. This could be seen in a dataset ordering the planets according to sequence around the sun.

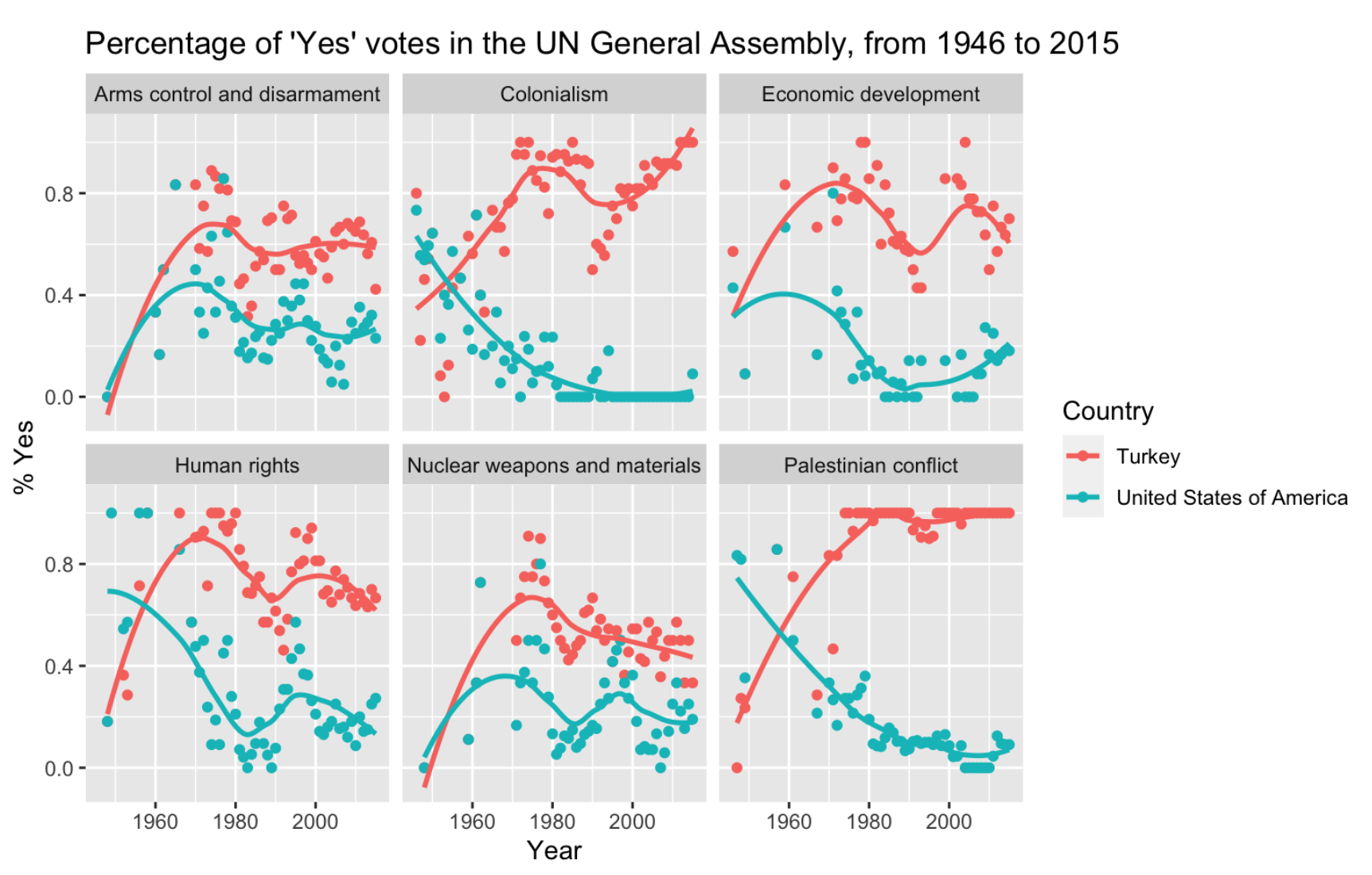
1. **What are the differences between observational studies and experiments?**

An experimental study divides the population into treatment and control groups in order to test the effect of a specific variable. Conversely, data for observational studies are taken from phenomena in pre-existing populations rather than from populations with specifically administered treatments. Observational studies do not divide the population into groups, but rather might measures risk factors by analyzing whether the subject was or wasn’t exposed. Observational studies might not even have distinct groups- they may have one sample and use statistics to find correlations and trends within that group.

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 random sampling methods: simple, stratified, cluster and multistage.

In stratified random sampling, the population gets divided into groups based on what each case has in common. These groups are called strata. Strata are specifically chosen to ensure that all desired groups are represented in the data since subgroups are specifically designated (unlike simple random where all population members are considered equal). Stratified random sampling might be used as opposed to simple random sampling if the cases in each stratum are very similar. It is also advantageous to use stratified random sampling to see how results differ in particular strata, and to clearly see the results within the different strata. This would allow the researchers to obtain more precise estimates within each subpopulation.

1. **Describe what you see in the plot below.** 

In the plot above we see the percentage of ‘yes’ votes given in the United Nations General Assembly by Turkey and the United States, from 1946 to 2015. The trendlines shows how the percentage of the vote of ‘Yes’ has changed from 1940 to 2020 for both the United States and Turkey for the following issues: arms control and disarmament, colonialism, economic development, human rights, nuclear weapons and materials, and palestinian conflict.

The data points show that in a majority of years, The US voted a lower percentage of yes votes than Turkey for all issues. Additionally, the US data trended negatively over time with relatively strong negative trends for Colonialism and Palestinian conflict issues. This suggests that US approval of the issues displayed in the plots lowered over time.