**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 |
| * In Math, conclusions are generally certain, whereas in Statistics conclusions are determined within levels of confidence * Statistics uses inductive reasoning * Math is more problem solving, while Statistics is more contextual and data based * Math can be used as a tool in solving Statistical problems, but not the other way around * Statistics looks more at complex relationships between variables, while Mathematics focuses on the values of such variables * Statistics involve both experiments and observational studies, whereas Math tends not to focus on the procurement of the data at hand and requires much less analysis and interpretation * Statistics revolves around the interpretation of collected data, while Math is not reliant on data collection | * Both involve numbers * Mathematics and Statistics can be used simultaneously to solve a problem, as they both serve as problem-solving mechanisms * Both utilize formulas * Both are often formatted in word problems or expressions with variables * Both use consistent numbers (like pi and alpha) |

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

The word “Statistics” is a mathematical practice that revolves around collection, organization, analysis, and interpretation of data. Statistics involves experimental methods, like randomization and replication. Through the use of “Statistics” we can arrive at a “statistic”. The word “statistics” can be numerical facts, such as the number of COVID cases or percentage of deaths, that are used to support conclusions.

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?**

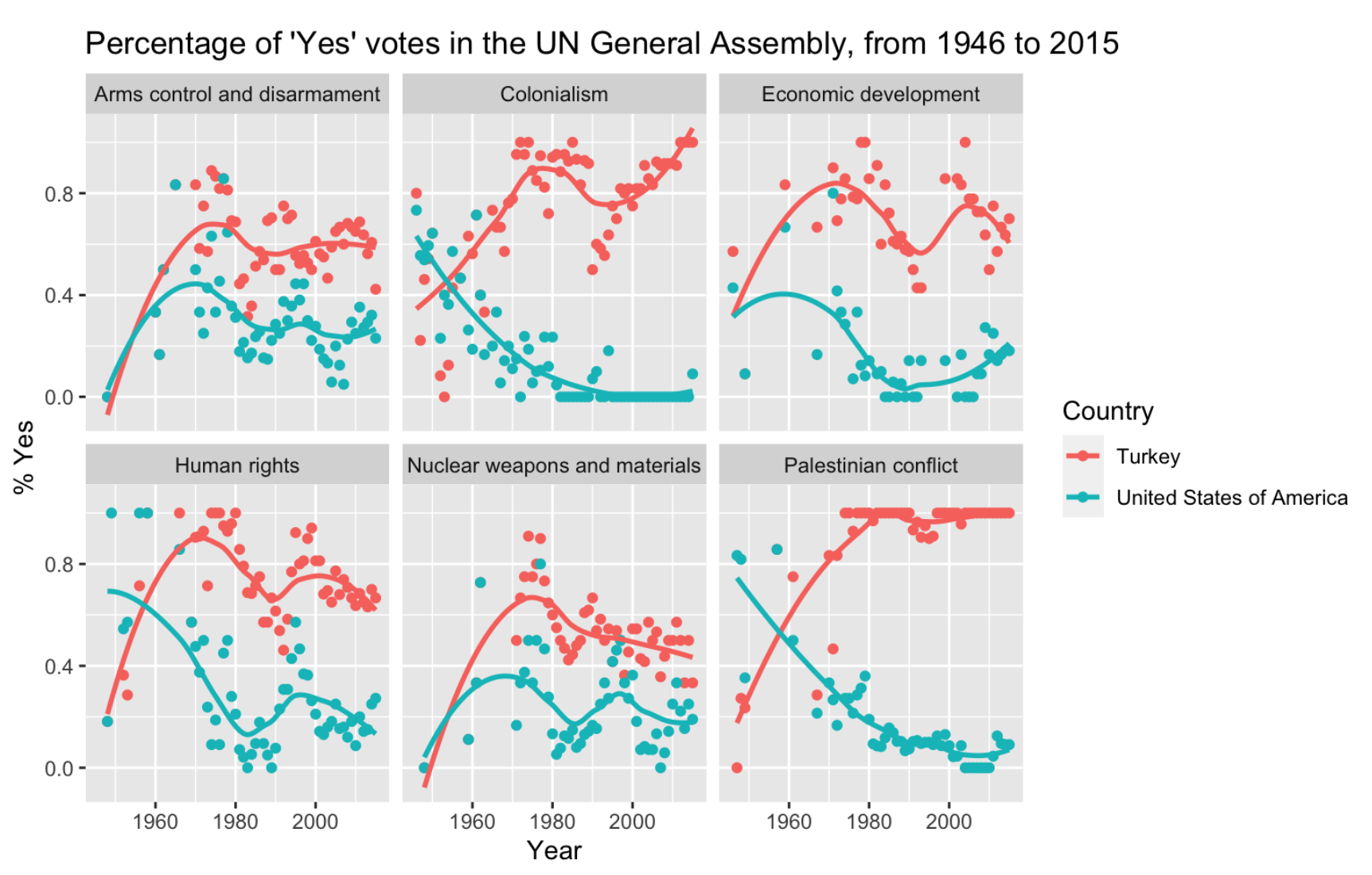
All variables are either numerical or categorical. Numerical values are measurable and have exact values. Examples of numerical values are height, weight, age. You can take averages of all these things and conduct mathematical equations with them. Numerical variables can be either continuous or discrete. Discrete variables involve countable values at positive whole number increments, while continuous variables involve measurable values to any level of accuracy. Categorical variables are variables that fit into a category. Examples include gender, race, and hair color. These things cannot be added or subtracted from each other, unlike a numerical variable. Categorical variables can be either ordinal or nominal. Ordinal variables are categorical with an order to them, like education level (some high school, high school, some college, bachelor's degree). A nominal variable is a categorical variable without an order to them, like house color. You can determine whether a variable is numerical or categorical by checking if the variable is a measure of something, in the form of a number, or if it is a variable in a category.

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

The difference between observational studies and experiments is how they are conducted as well as what types of conclusions can be drawn from them. In experiments, researchers are looking to determine the relationship between an explanatory and response variable. Researchers manipulate variables, usually with a treatment and a control group, while attempting to control the differences between groups. In experiments, cases can be randomized, controlled, replicated, or blocked in order to try and eliminate the external effects of unaccounted variables. Observational studies are simply observational and do not influence the way the data actualize. Due to this, observational studies cannot determine causal relationships, whereas experiments can conclude cause. Observational studies can also be used where experiments would be unethical. It would be unethical to assign people to treatment groups like brain damage, so observational studies offer an ethical way of gathering data.

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?**

Stratified random, simple random, cluster, and multistage are the different methods for sampling from a population. Simple random sampling would involve taking a random sample from the entire population with all cases having an equal chance of being selected. You could use a number generator or something similar to ensure randomness. Stratified would be taking your sample population, splitting them into strata, and then doing a simple random sample from each stratum. For example, if your population was college students, you could stratify the cases by grade and then conduct a simple random sample from each grade. A stratified random sample would better represent the entire population because it would guarantee that each sub-group is accurately represented. Cluster sampling involves breaking the population into a fixed number and then sampling a fixed number of the clusters. All cases within the clusters are included in this cluster sampling. Multistage sampling is similar, but goes one step further. The population is broken up into a fixed number of clusters, the clusters are chosen at random, and cases are chosen within each cluster through simple random sampling.

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

Turkey has a higher percentage of “Yes” votes in the UN General Assembly across all topics in more recent years than the United States. However, America generally had a higher percentage of “Yes” votes in earlier years (pre-1960). From 1960 until present, Turkey and the U.S. followed a similar trend for arms control and disarmament, human rights, economic development, and nuclear weapons and materials. In contrast, Turkey and the U.S. seem to be increasingly voting in opposite directions for both the Palestinian conflict and Colonialism. There is a positive correlation between the percent “Yes” votes and time for Turkey when it comes to colonialism and the Palestinian conflict. In both of these topics, there is a negative correlation between time and percent “Yes” votes for the United States. In addition, the percentage of “Yes” votes for Turkey regarding arms control and disarmament, colonialism, economic development, human rights, and nuclear weapons and materials all peak at around 1970-1975. The plots also show how much Turkey and the US differed on certain issues. The gap between percentage of “Yes” votes on issues regarding colonialism, economic development, human rights and Palestinian conflict are large, with at least a 45% difference between the two countries.