Project Proposal: OLYMPICS

I.  Introduction
    To look at the medal counts per country and what factors will influence the outcome.

II. Data Source
    ❖ Primary Data Sources
    https://www.kaggle.com/samruddhim/olympics-althlete-events-analysis
    athlete_events.csv
    https://www.kaggle.com/piterfm/olympic-games-medals-19862018
    olympic_medals.csv
    ❖ Supporting Resources
    https://www.kaggle.com/themlphdstudent/country-wise-gdp-from-1994-to-2017
    Country wise GDP from 1994 to 2017.csv

III. Research Questions
    a. Is the relationship between participating Country success and major historical event /World War II?
    b. Is Country GDP & Population size contributing factors for Olympic success?
    c. Is there a high success rate for specialty games in some countries?
    d. Team vs individual sports, which one has highest success rate per country?
    e. Which game is more popular in which continent?

IV. Brainstorm/Hypothesis
    Statistical Test
            1. ANOVA Test or Chi Test
            - Gold, Silver, Bronze categories
            - Determine which one would be more appropriate/accurate.
            2. T-Test: Country vs Country
            3. Regression
            - Predict the total medal count for each country
            - New Olympic records could be set
    Hypothesis
    ❖ We predict the GDP has a major influence on the number of medals the USA achieved.
    ❖ We predict population has a major influence on the number of medals both China/Russia achieved.

V.   Inspirations
- ❖ The "Year" of the Olympics
- ❖ Due to the recent experiences we all had with social distancing and being away from family, friends, and entertainment our group wanted to celebrate the event that brings everyone together and back to some form of normalcy.

VI.   Role Breakdown
- ❖ Cleaning Data: Kidist/Alisha
- ❖ Data Visualizations/Regression/Live Demo: Alfred/Group
- ❖ Slide Deck/Analytical Write Up: Lucas