Project 1, Group 2
Monica Mitry, Emily Wimmer, Cecilia Rocha, Arya Maredia

1. **Introduction**
    a. Our dataset is a summary of chocolate bar ratings from the years 2006 to 2017
    b. We chose this dataset because we love chocolate and we would love to know where the best chocolate comes from
    c. https://www.kaggle.com/datasets/rtatman/chocolate-bar-ratings?select=flavors_of_cacao.csv
2. **High Level Questions**
    a. What countries do the best cocoa beans for chocolate production originate from?
    b. What country uses the highest cocoa percentage?
    c. Is there a relationship between the rating of the cocoa and the cocoa solids percentage?
3. **Inspiration** - We are just here for the sweets. We want to know the bean origin and company that makes the best chocolate.
    a. Code:
       https://www.kaggle.com/code/tibhar940/chocolate-bar-ratings-python-eda-dataviz
4. **Visuals**
    a. Map geographical data
        i. Geographical distribution of bean origin
        ii. Geographical distribution of manufacturing companies
    b. Bar chart/Donut Chart
        i. Used to view which countries produce more cocoa beans
        ii. View the number of manufacturing companies per countries
    c. Line chart
        i. Number of reviews per year
    d. Violin Plot
        i. Comparison of company location and the cocoa solids percentage.
    e. Color Scheme
        i. Willy Wonka Branding
        ii. https://coolors.co/199296-e4b90c-5e3d2b-fed2d9-5c08ac-5c265f
5. **Linear Regression** Is there a regression in the rating of the cocoa raing and the percentage
    a. Is there a correlation between the cocoa percentage and the cocoa rating?
    b. Is there a negative relationship between the cocoa percentage and cocoa rating for manufacturing companies in the U.S.A.?
6. **Roles and Responsibilities**
    a. Question 1 - Cecilia
    b. Question 2 - Emily
    c. Question 3 - Monica
    d. Regression - Arya
    e. Slides, Write Up and Data Cleaning - Collaborative

f. Slide Presentation
   i. Intro - Cecilia
   ii. Objective - Emily
   iii. Our Dataset - Arya
   iv. Question 1 - Cecilia
   v. Question 2 - Emily
   vi. Question 3 - Monica
   vii. Regression - Arya
   viii. Limitations & Conclusion - Arya