**Introduction**

Chocolate is a gift from nature and a multisensory product. It appeals to all five senses (sight, touch, hearing, smell and taste) thanks to all its attributes: colour, snap, mouthful and complexity of flavors and aromas. People all around the globe love it. It’s one of the favorites – if not THE favorite – sweet product around. But what makes people rate one chocolate bar above another? Is it the Cacao beam origin? Or is it the percentage of Cocoa in the bar? We will use the insights from this analysis to figure it all out.

**The Data**

The data used for this analysis was gotten from Kaggle under this [license](https://creativecommons.org/publicdomain/zero/1.0/). It contains expert ratings of over 1,700 individual chocolate bars, along with information on their regional origin, percentage of cocoa, the variety of chocolate bean used and where the beans were grown.

### Flavors of Cacao Rating System:

5= Elite (Transcending beyond the ordinary limits)

4= Premium (Superior flavor development, character and style)

3= Satisfactory(3.0) to praiseworthy(3.75) (well made with special qualities)

2= Disappointing (Passable but contains at least one significant flaw)

1= Unpleasant (mostly unpalatable)

**Project Tool Used: Excel**

**Data Preparation**

The business questions to look into in this project are the following:

1. Where are the best cocoa beans grown?
2. Which countries produce the highest-rated bars?
3. What’s the relationship between cocoa solids percentage and rating?

**Availability of Data**

- The provided data would be good enough for the analysis to answer the business questions asked.

**Data Cleaning Process**

I corrected some data with wrong spelling most especially in the Company Location and Broad Bean Origin columns.

The missing values in the Bean Type column were replaced with N/A while all the rows with missing values in the Broad Bean column were deleted due to their small number. When checking for duplicates, only one duplicate was found and removed.

A column was created with the name Rating Category to group the rating values into categories while unnecessary columns like REF, Review Date, Bean Type and Specific Bean Origin were removed due to their irrelevances to the business questions.

**Analysis and Insights**

- A pivot table was created with all the cleaned data and graphs were plotted using the sum of Rating for each Broad Bean Origin and Country.

- The rating category was used to filter the data to only praiseworthy, premium and elites and the best ten (10) regions/countries were selected.

- A scatter chart was plotted to show the relationship between the cacao percentage and rating.

- It was found that cacao beans were best grown in Venezuela, Madagascar, Peru, Ecuador, and Papua New among others. While the country with the highest-rated bars is USA which is far ahead of France and Canada.

- The cocoa solids percentage does not have any correlation with the rating as shown in the visualization.

**Dashboard**

