

Chocolate BAR RATINGS

Project 1 - Group 2

Emily Wimmer
Cecilia Rocha
Arya Maredia
Monica Mitry

Objective

The dataset is an evaluation of over 1,700 chocolate bars showcasing their ratings from the years 2006 to 2017.

As chocolate lovers, this dataset helps to determine where the best chocolate bars come from.



Data Engineering

Created new CSV File with clean reference dataframe

- Dropped Bean Type Column
- Removed 1 row of null data
- Cleaned unknown countries for Broad Bean Origin
- Renamed column headers
- Clean up typos for country names
- Created Blended category for country & Cocoa Bean Origin
- Willy Wonka



high level QUESTIONS

Question 1

What countries do the best cocoa beans for chocolate production originate from?

Question 2

What country uses the highest cocoa percentage?

Question 3

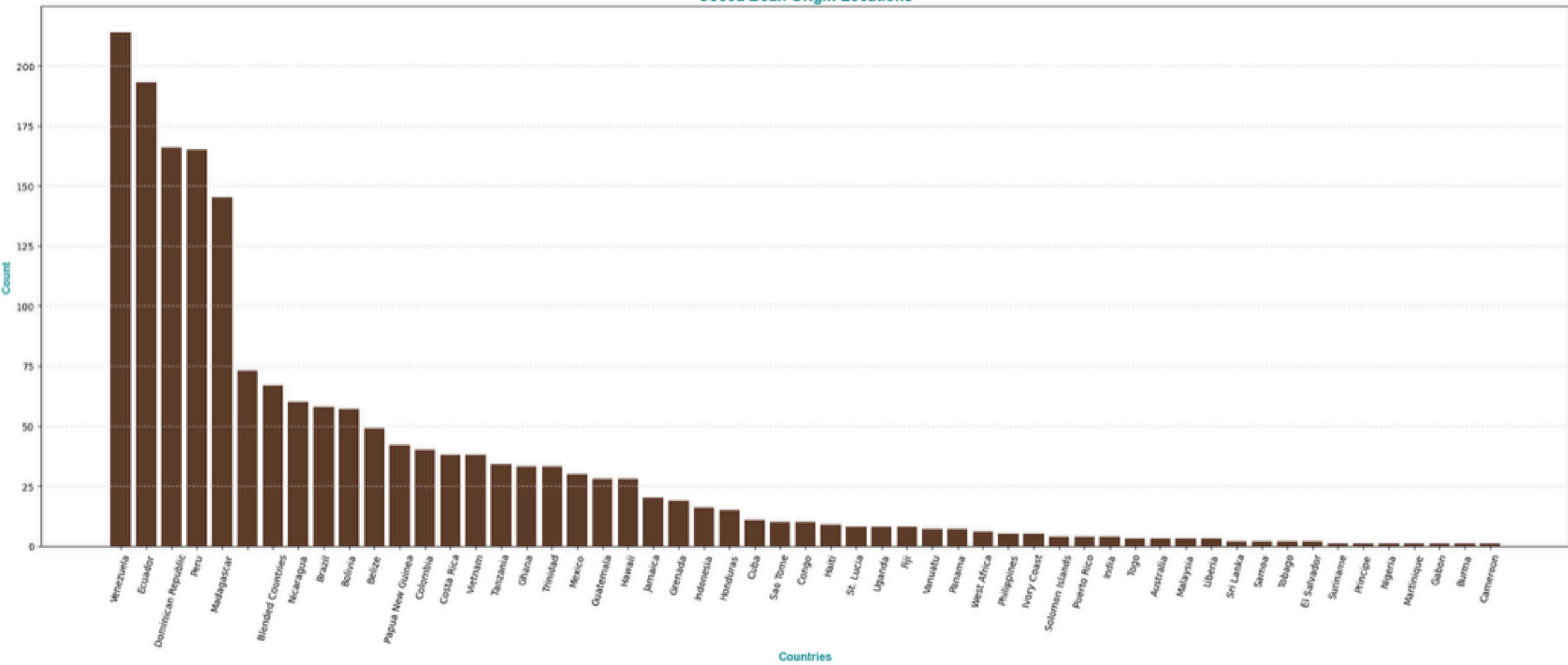
Is there a relationship between the rating of the cocoa and the cocoa solids percentage?

Question #1

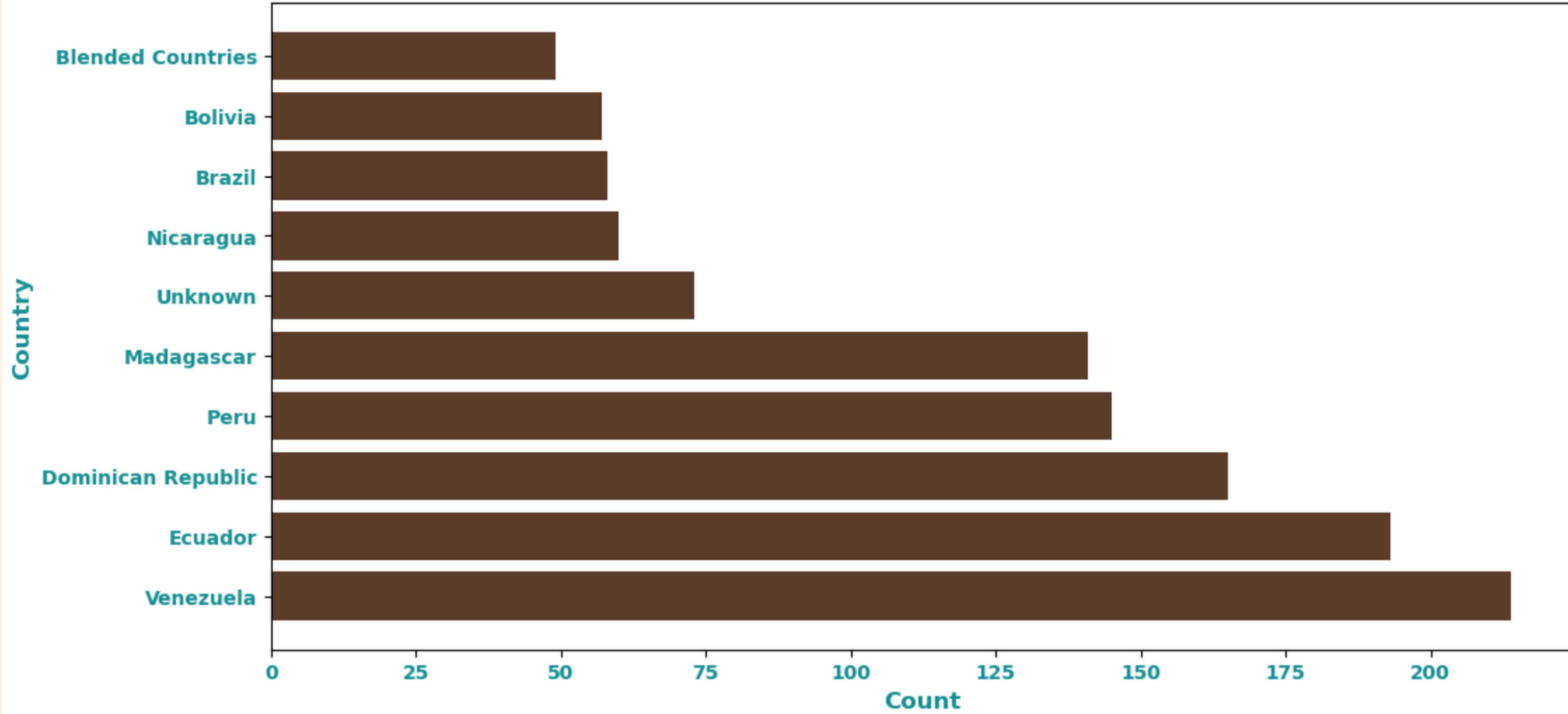
What countries do the best cocoa beans for chocolate production originate from?

- To answer this question we must look at two factors, where most cocoa beans originate from and the locations of the manufacturing companies
- We will look at the top ten countries where beans originate from and the top ten manufacturing companies location

Cocoa Bean Origin Locations

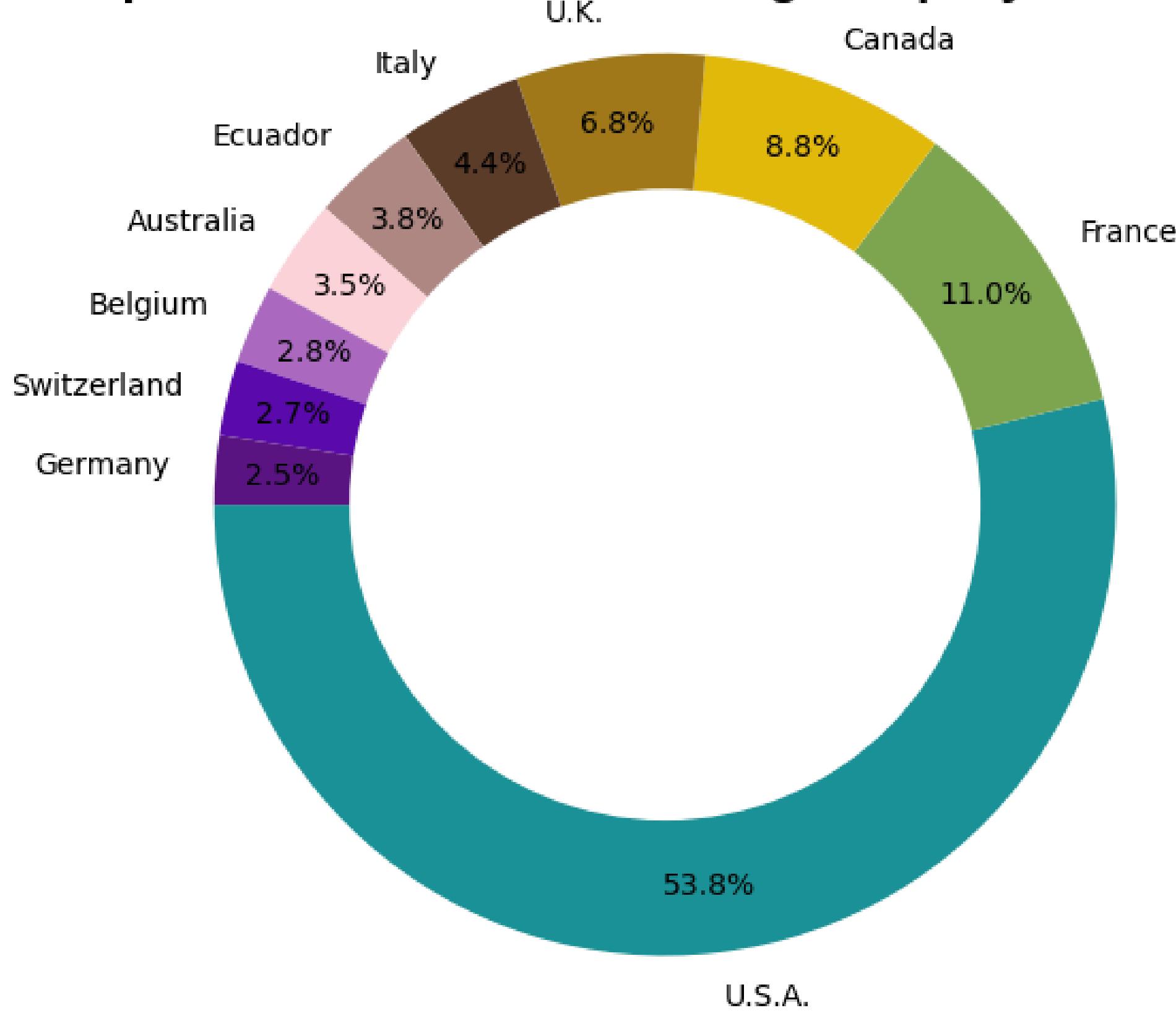


Top 10 Cocoa Bean Origin Locations



Venezuela is the leading source of cocoa bean origins used by companies in chocolate production, with 214 companies sourcing their cocoa beans from Venezuela, as indicated by the data.

Top 10 Chocolate Manufacturing Company Locations



The United States leads in the number of chocolate manufacturing companies, accounting for 58.3% of the total, with 763 companies in operation.

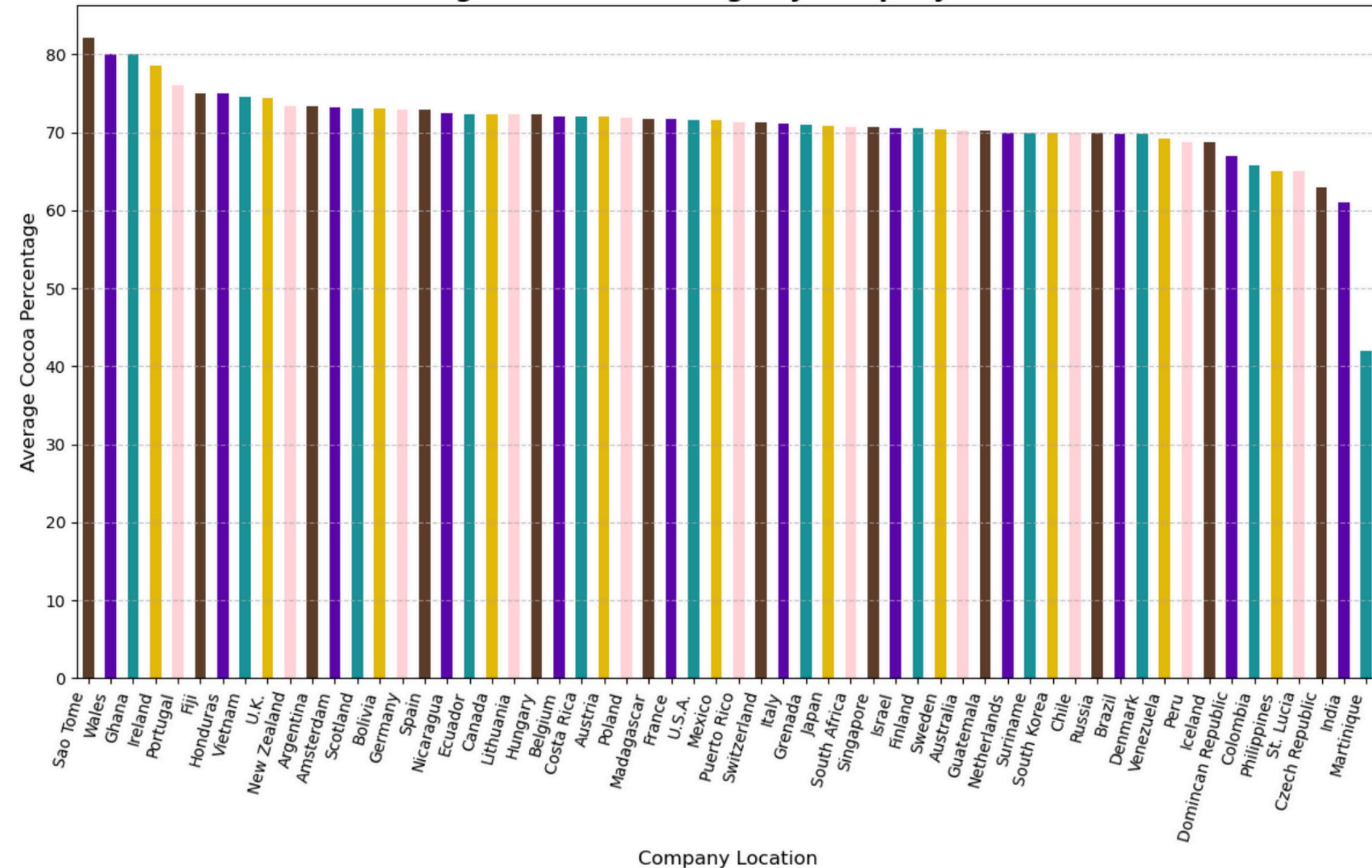
company_location	count
U.S.A.	763
France	156
Canada	125
U.K.	96
Italy	63
Ecuador	54
Australia	49
Belgium	40
Switzerland	38
Germany	35

Question #2

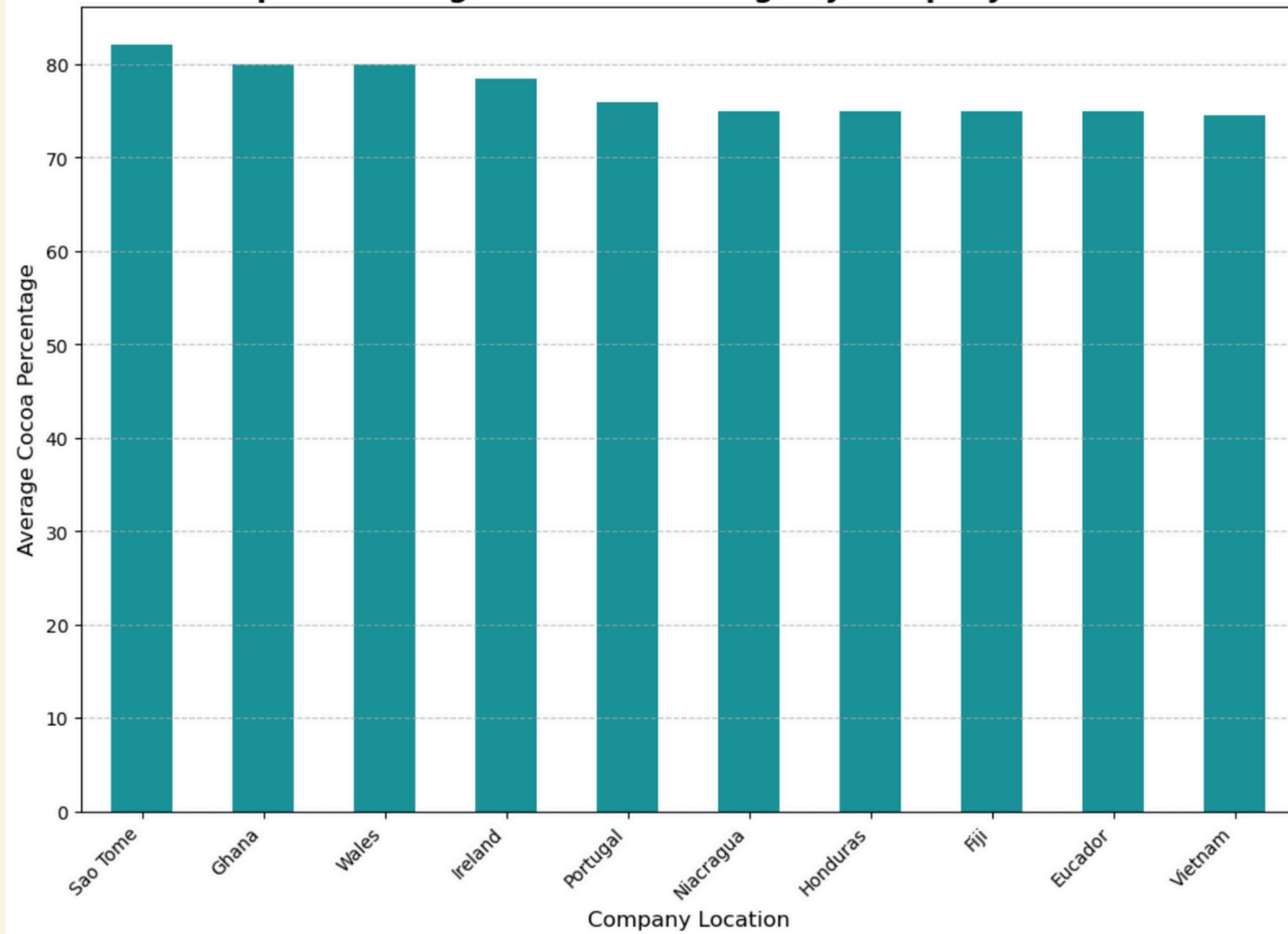
What country uses the highest cocoa percentage?

- To find the country with the highest cocoa percentage we must find the average of cocoa percentage used by companies in the country.
 - After grouping the data to find the mean it is clear that Sao Tome had the highest (aka Dark Chocolate) with 82%.
 - Martinique was the lowest with 42%.
-

Average Cocoa Percentage by Company Location



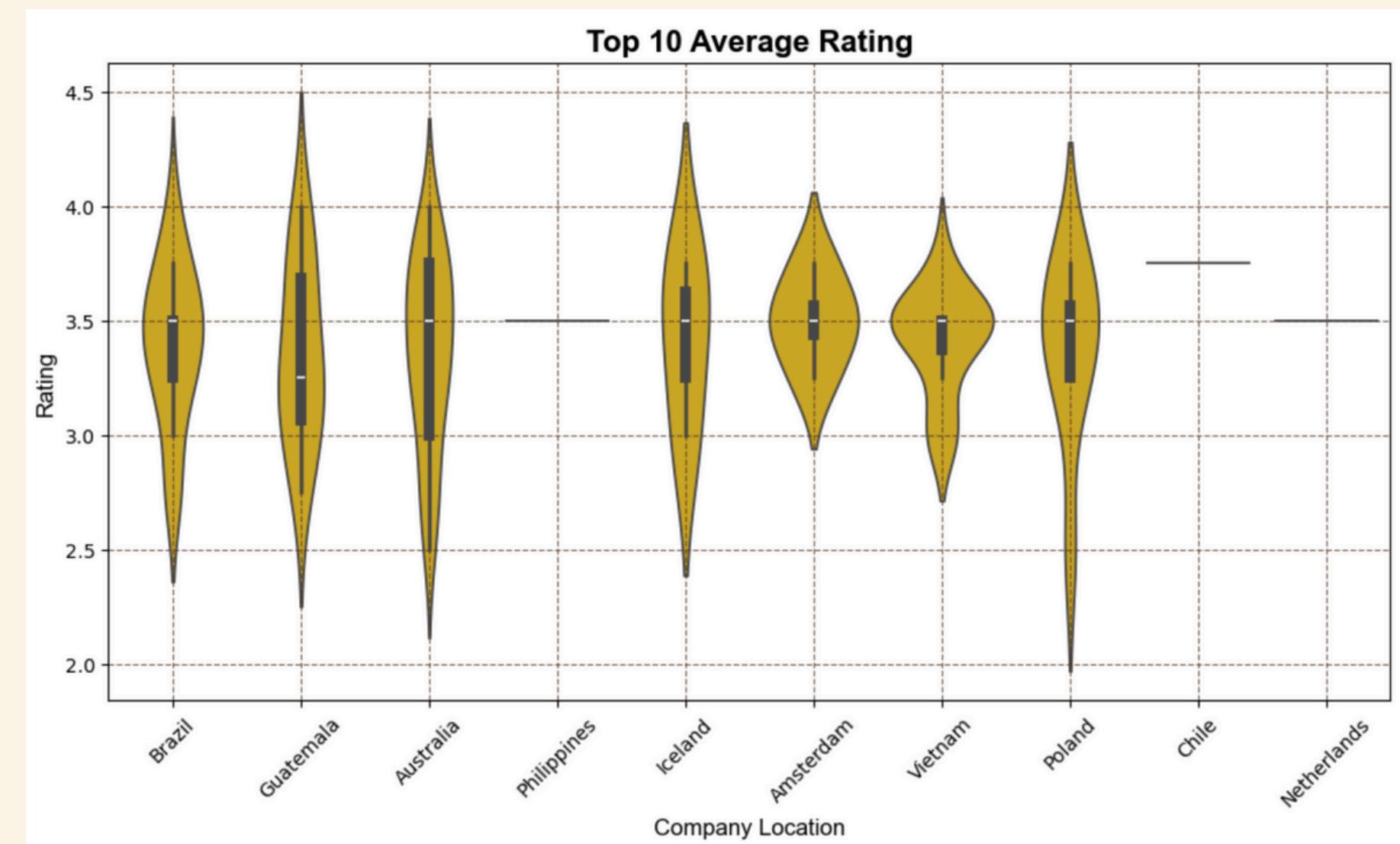
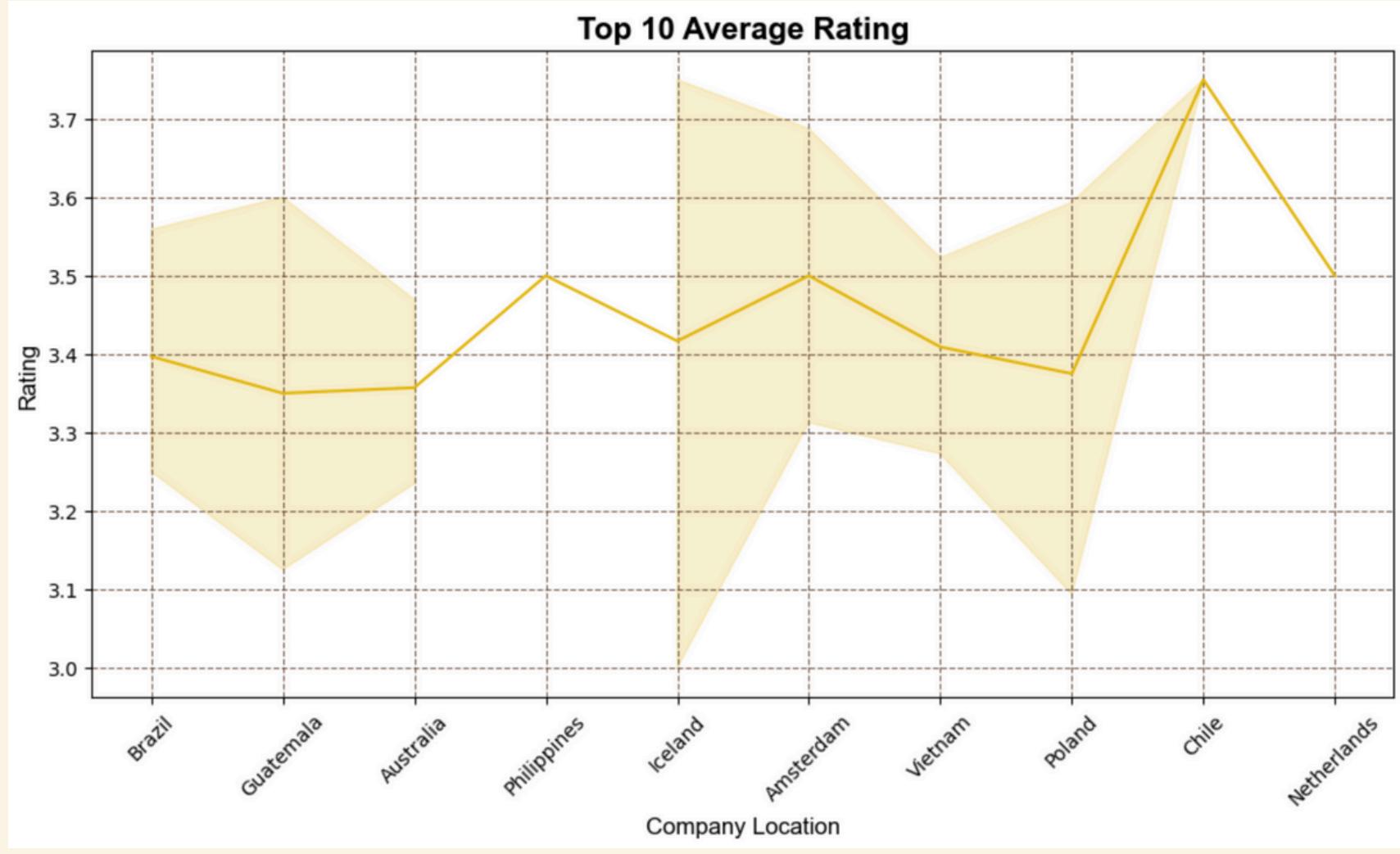
Top Ten Average Cocoa Percentage by Company Location



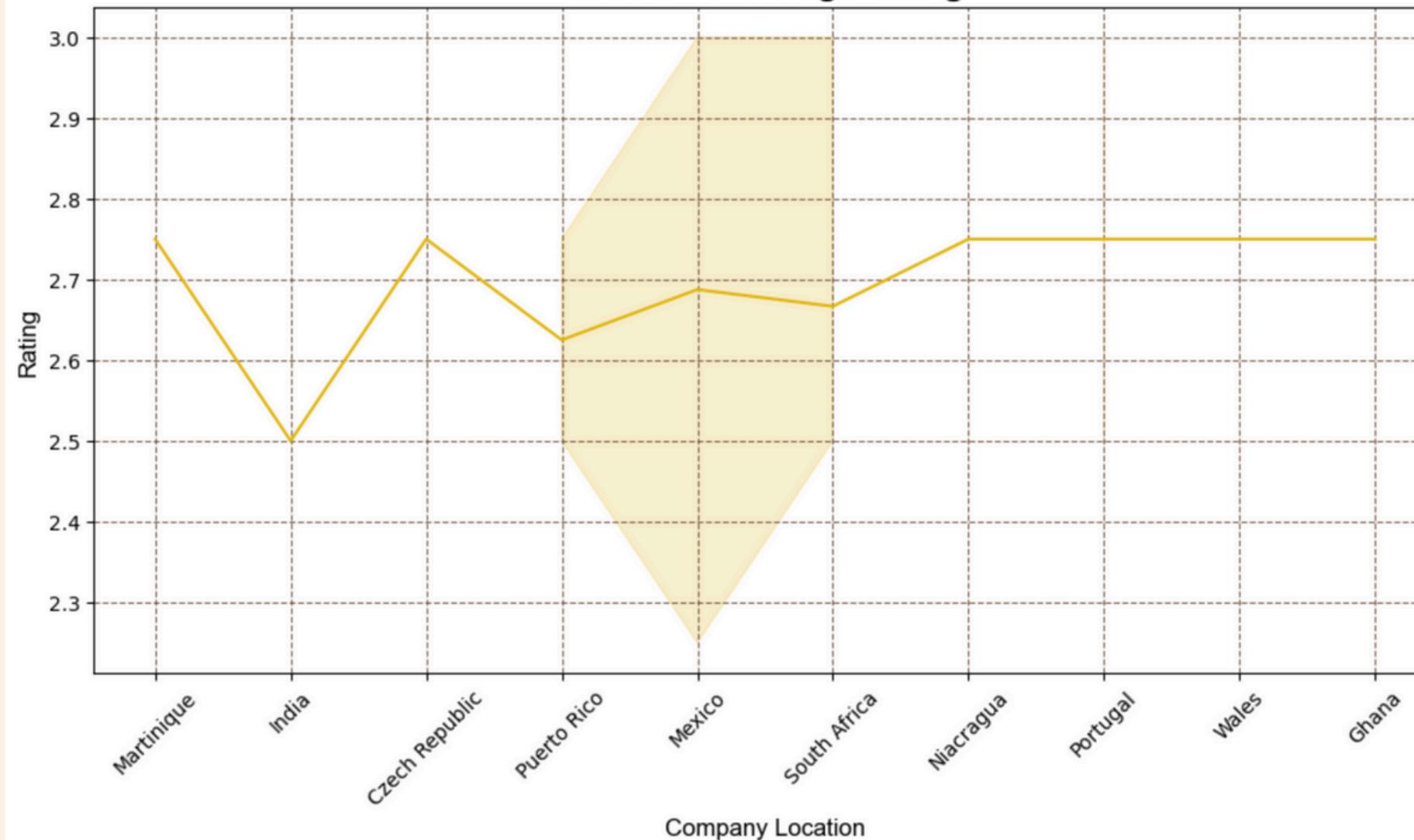
Question #3

Is there a relationship between the rating of the cocoa and the company location?

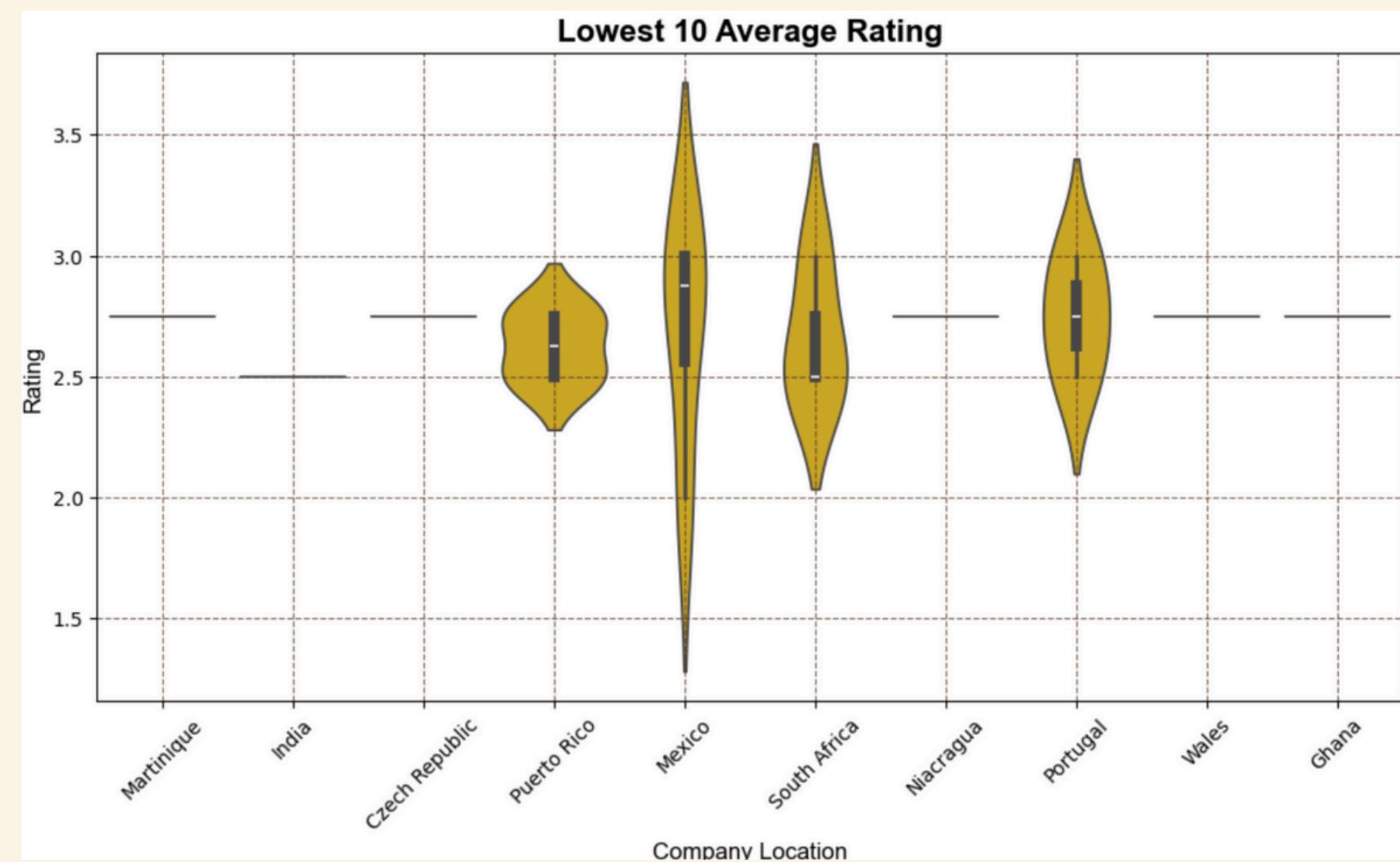
- While narrowing it down to the top 10 average chocolate bars Chile, Netherlands, Philippines have a higher distribution but some of the countries with their one bar/company have a higher average rating.
- On the other side of the spectrum is the lower average bars where some countries like Martinique have narrow distributions and lower rated bars.



Lowest 10 Average Rating



Lowest 10 Average Rating



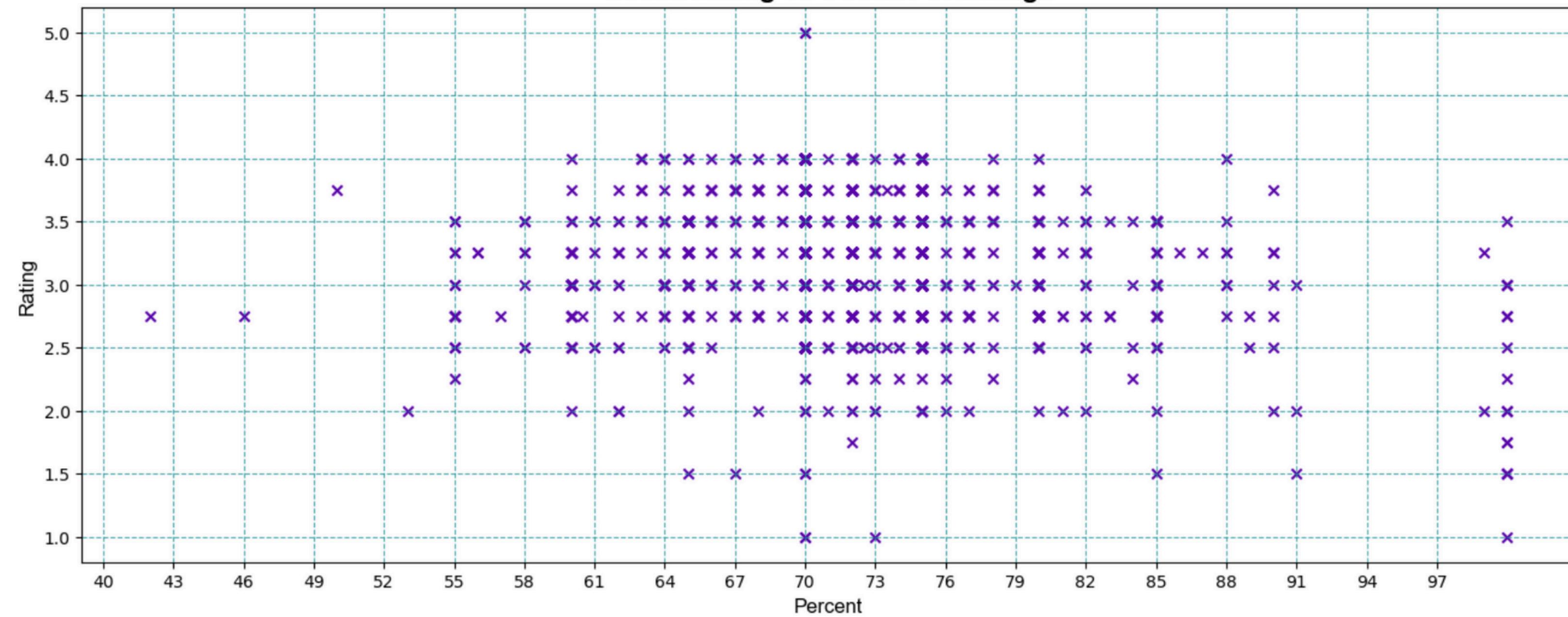
Regression

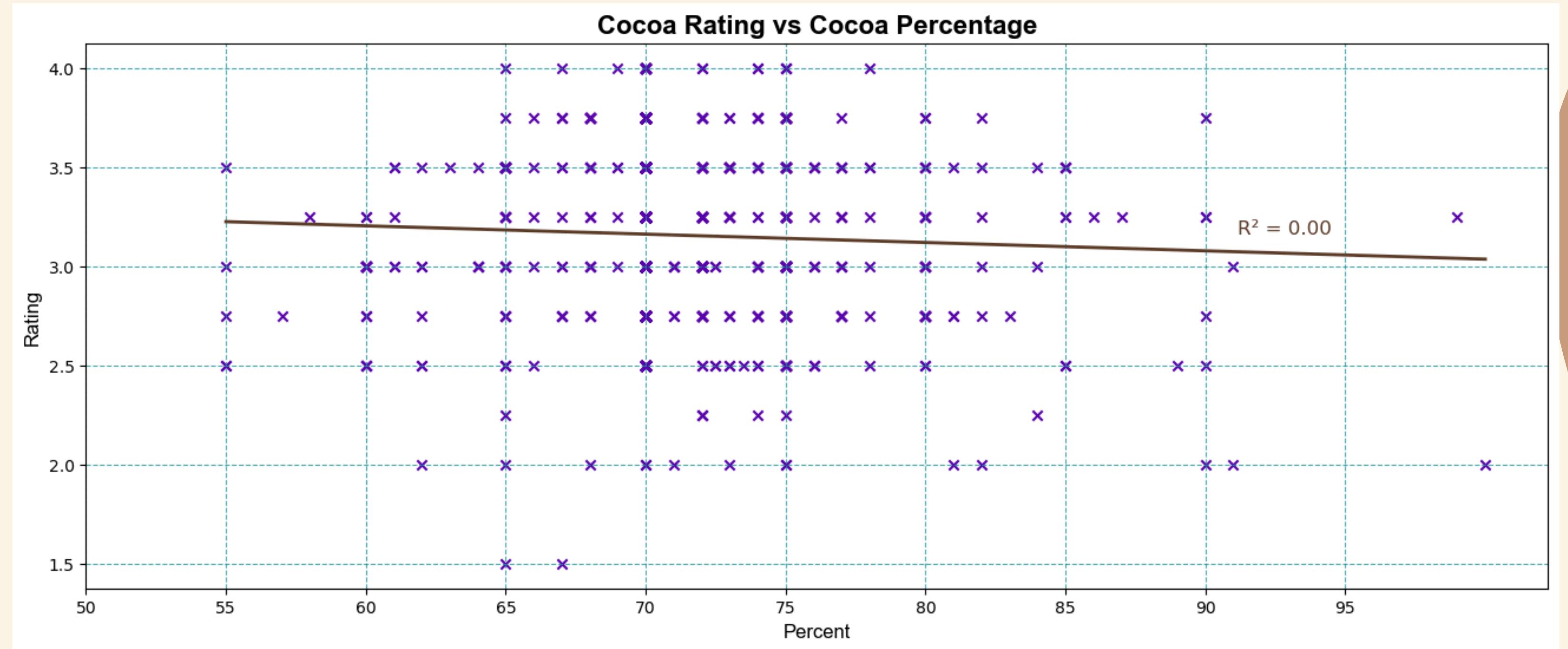
Is there a correlation between cocoa rating and cocoa percentage?

Is there a negative correlation between cocoa rating and cocoa percentage in the US?

- Scatter plot for continuous numerical values.
- Locate chocolate bars only manufactured in the US.

Cocoa Rating vs Cocoa Percentage



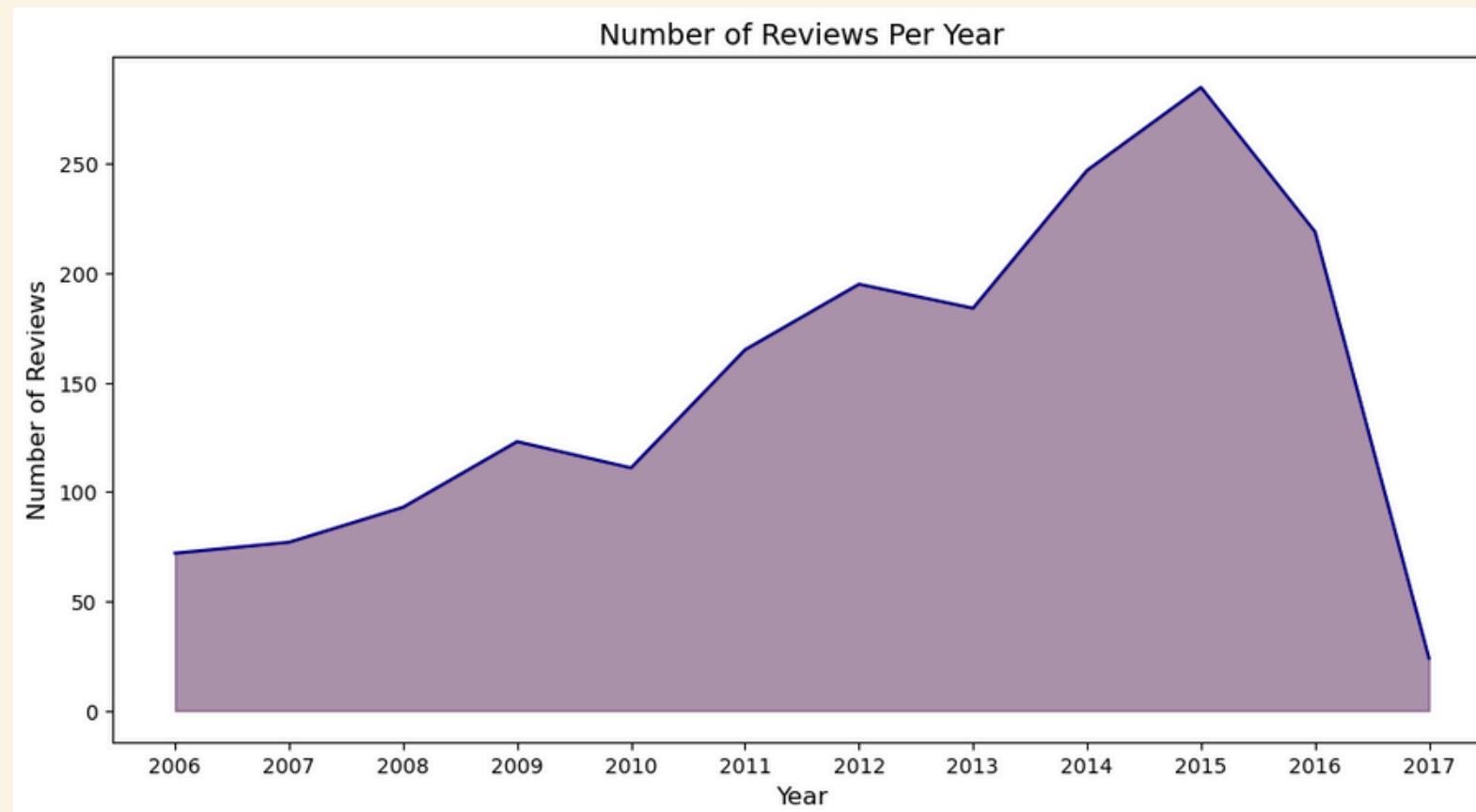


Bias

- 🍫 2017 had significantly less data points than previous years.
- 🍫 Countries that had one company that produced one chocolate bar skewed average ratings and cocoa percentage.

Limitations

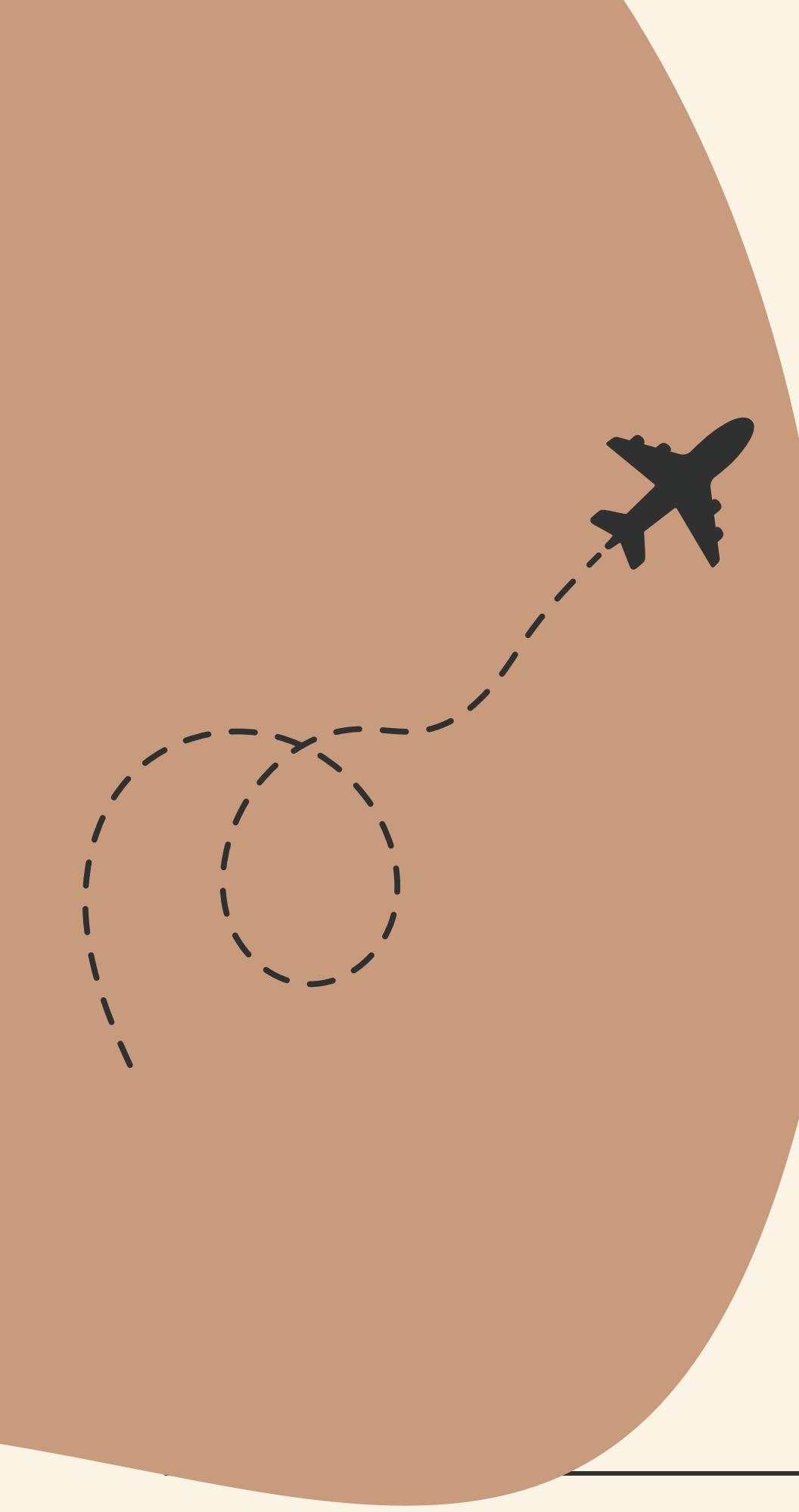
- 🍫 Bean type column had to be dropped due to a high quantity of missing data (~50%).
- 🍫 Bean Type Origin with blended countries and unknown data.



FINAL THOUGHTS

- United States has the most manufacturing companies
- Most cocoa beans originate from the Amazon Rainforest
- Africa uses the highest percent of cocoa
- Dark chocolate bars were rated just as high as milk chocolate bars, even in the US
- Chile with one submission had the highest rated chocolate bar





What's next?

- ─ Go to Chile - try this AMAZING chocolate bar
- ─ Group by continent

WORK CITED

- <https://www.kaggle.com/code/tibhar940/chocolate-bar-ratings-python-eda-dataviz>
 - Chat GPT
-

questions!

