Chocolate Bar Ratings Analysis

Data

Source of Data:

The data is available on one of open source platform – kaggle.com and downloaded from following link https://www.kaggle.com/datasets/rtatman/chocolate-bar-ratings

Context
Chocolate is one of the most popular candies in the world. Each year, residents of the United States collectively eat more than 2.8 billions pounds. However, not all chocolate bars are created equal! This dataset 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.

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)

Data Collection:

- The dataset used here have been acquired from Rachael Tatman's Chocolate Bar Ratings dataset on Kaggle.
- The original ratings were compiled by Brady Brelinski, Founding Member of the Manhattan Chocolate Society. For up-to-date information, as well as additional content (including interviews with craft chocolate makers), please see his website: Flavors of Cacao

.

Data Content:

This dataset 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.

Columns and its descriptions

company_maker_if_known : Name of the company manufacturing the bar.

specific_bean_origin_or_bar_name : The specific geo-region of origin for the bar.

ref : A value linked to when the review was entered in the database. Higher = more recent.

review_date : Date of publication of the review.

cocoa_percent : Cocoa percentage (darkness) of the chocolate bar being reviewed.

company_location : Manufacturer base country.

rating : Expert rating for the bar.

bean_type : The variety (breed) of bean used, if provided.

broad_bean_origin : The broad geo-region of origin for the bean

Data understanding

Columns	Qualitative/Quantitative	Discrete / Continuous	Nominal/Ordinal/Binary
Company	Qualitative		Nominal
Bar Name	Qualitative		Nominal
REF	Quantitative	Discrete	
Review Date	Quantitative	continuous	
Cocoa Percent	Quantitative	continuous	
Company Location	Qualitative		Nominal
Rating	Quantitative	continuous	
Bean Type	Qualitative		Nominal
Broad Bean Origin	Qualitative		Nominal

Limitation and ethics

A rating here only represents an experience with one bar from one batch which could led to biasness

The database is narrowly focused on plain dark chocolate with an aim of appreciating the flavors of the cacao when made into chocolate.

Ethics: I don't see any ethical issues

Questions to explore:

- Where are the best cocoa beans grown?
- Which countries produce the highest-rated bars?
- What's the relationship between cocoa solids percentage and rating?

⍰