

Coffee Quality Institute:

The Coffee Quality Institute (CQI) is a non-profit organization that works to improve the quality and value of coffee worldwide. It was founded in 1996 and has its headquarters in California, USA. CQI's mission is to promote coffee quality through a range of activities that include research, training, and certification programs. The organization works with coffee growers, processors, roasters, and other stakeholders to improve coffee quality standards, promote sustainability, and support the development of the specialty coffee industry.

Data:

The data includes a range of information on coffee production, processing, and sensory evaluation. It also contains data on coffee genetics, soil types, and other factors that can affect coffee quality.

Sensory evaluations (coffee quality scores)

- **Aroma:** Refers to the scent or fragrance of the coffee.
- **Flavor:** The flavor of coffee is evaluated based on the taste, including any sweetness, bitterness, acidity, and other flavor notes.
- **Aftertaste:** Refers to the lingering taste that remains in the mouth after swallowing the coffee.
- **Acidity:** Acidity in coffee refers to the brightness or liveliness of the taste.
- **Body:** The body of coffee refers to the thickness or viscosity of the coffee in the mouth.
- **Balance:** Balance refers to how well the different flavor components of the coffee work together.
- **Uniformity:** Uniformity refers to the consistency of the coffee from cup to cup.
- **Clean Cup:** A clean cup refers to a coffee that is free of any off-flavors or defects, such as sourness, mustiness, or staleness.
- **Sweetness:** It can be described as caramel-like, fruity, or floral, and is a desirable quality in coffee.

PLEASE NOTE: 'Total Cup Points' is literally the total of 10 features given above. There were some notebooks trying to predict the total cup points given these features. We know the exact function underlying the total cup points.

Defects:

Defects are undesirable qualities that can occur in coffee beans during processing or storage.

Defects can be categorized into two categories: Category One and Category Two defects.

Category One defects are primary defects that can be perceived through visual inspection of the coffee beans. These defects include Black beans, sour beans, insect-damaged beans, fungus-damaged beans, etc.

Category Two defects are secondary defects that are more subtle and can only be detected through tasting. These defects include Over-fermentation, staleness, rancidness, chemical taste, etc.

Objective:

The primary goal of this project is to leverage the rich dataset provided by CQI to understand the factors that contribute to coffee quality. Specifically, we aim to explore the following research questions:

1. What are the key determinants of coffee quality as evaluated through sensory attributes such as aroma, flavor, acidity, etc.?
2. Is there a correlation between processing methods, origin regions, and coffee quality scores?
3. Can we identify any trends or patterns in defect occurrences and their impact on overall coffee quality?

4. How do different variables interact to influence the Total Cup Points, which represent an overall measure of coffee quality?