**Power BI Project Documentation: Coffee Quality Analysis (CQI Dataset)**

**Project Title:**

Coffee Quality Analysis using Power BI

**Organization:**

Coffee Quality Institute (CQI)

**1. Project Overview:**

The Coffee Quality Institute (CQI), a non-profit based in California, USA, aims to improve the quality and value of coffee worldwide. This project leverages data from CQI to analyze various factors that influence coffee quality using Power BI.

**2. Objectives:**

* Identify key determinants of coffee quality through sensory attributes.
* Examine correlation between processing methods, origin regions, and coffee quality scores.
* Identify patterns in defect occurrences and their impact on overall quality.
* Understand how different variables interact to influence Total Cup Points.

**3. Dataset Description:**

The dataset contains the following components:

**Sensory Evaluation (Score-based):**

* **Aroma**: Scent/fragrance of the coffee.
* **Flavor**: Taste profile (sweetness, bitterness, acidity, etc.).
* **Aftertaste**: Lingering taste.
* **Acidity**: Brightness or liveliness.
* **Body**: Thickness or mouthfeel.
* **Balance**: Harmony of all flavors.
* **Uniformity**: Consistency across cups.
* **Clean Cup**: Absence of defects.
* **Sweetness**: Pleasant sugary or fruity taste.
* **Overall**: General impression.
* **Total Cup Points**: Sum of the above 10 attributes (known formula).

**Other Attributes:**

* **Processing Method**: Washed, Natural, Honey, etc.
* **Origin Region/Country**: Source of coffee.
* **Harvest Year**: Year the coffee was harvested (cleaned to first year in cases like 2021/2022).
* **Defects**:
  + **Category One Defects**: Black beans, sour beans, etc.
  + **Category Two Defects**: Over-fermentation, staleness, etc.

**4. Tools Used:**

* **Power BI Desktop**: For data modeling, analysis, and visualization.
* **Power Query**: For data cleaning and transformation.
* **Bookmarks & Selections**: For interactive navigation and enhanced user experience.

**5. Methodology:**

* **Data Cleaning**: Removed nulls, standardized columns (e.g., Harvest Year).
* **Data Modeling**: Created relationships between tables where applicable.
* **Visualization**:
  + Bar Charts, Radar Charts: For comparing sensory scores.
  + Scatter Plots with Trend Lines: For correlation analysis.
  + Heatmaps: To analyze patterns and correlations.
  + Stacked/Clustered Columns: To analyze defects.
  + **Bookmarks & Selection Pane**: Used to switch between different views dynamically.
  + Slicers & Filters: For interactivity.

**6. Key Insights & Analysis:**

**Q1: What are the key determinants of coffee quality?**

* Used bar charts and scatter plots to compare each sensory attribute against Total Cup Points.
* Found strong positive correlation with attributes like **Flavor**, **Aroma**, and **Acidity**.

**Q2: Correlation between processing methods/origin & quality?**

* Used bar charts, stacked columns, and scatter plots.
* Observed that certain processing methods like **Washed** had higher average scores.
* Some regions consistently performed better (e.g., Ethiopia, Colombia).

**Q3: Patterns in defects and impact on quality?**

* Plotted count of Category 1 and 2 defects.
* High defects correlated with lower scores in **Clean Cup**, **Uniformity**, and **Total Points**.

**Q4: Interaction of variables influencing Total Cup Points?**

* Created scatter plots and heatmaps to visualize interactions.
* Highlighted the additive impact of all sensory scores.
* Suggested a possible predictive model using selected features.

**7. Challenges:**

* Some records had Chinese characters which required proper font and encoding support (handled successfully).
* **Complex Navigation Requirement**: Used Power BI **Bookmarks** & **Selection Pane** to improve user interactivity and make dashboard switching smooth.

**8. Conclusion:**

This Power BI project successfully explored various determinants of coffee quality using rich CQI data. It revealed key sensory attributes that influence total scores, exposed how defects and processing affect quality, and helped build a foundational understanding for further research or dashboard-based decision making.

**9. Next Steps (Future Scope):**

* Integrate machine learning models to predict Total Cup Points.
* Include real-time data from growers or quality labs.
* Create an interactive dashboard for farmers or exporters.

**10. Screenshots of Key Visuals:**

**Screenshot 1:**

**A screenshot of a computer

AI-generated content may be incorrect.**

**A screenshot of a computer

AI-generated content may be incorrect.**

**Screenshot 2:** Used bookmarks to switch between different views dynamically and

Slicers & Filters for better interactivity.

**Prepared by:** Angeline Joyce J

**Date:** 5/4/2025

**Tool Used:** Microsoft Power BI Desktop