

## POWER BI CAPSTONE PROJECT

### Project Objective:

The objective of this project is to analyze the dataset provided by the Coffee Quality Institute (CQI) to gain insights into factors affecting coffee quality. Specifically, the project aims to address the following research questions:

- **Determinants of Coffee Quality:** Identify the key factors influencing coffee quality, focusing on sensory attributes such as aroma, flavor, acidity, etc. This involves understanding which sensory attributes have the greatest impact on overall coffee quality.
- **Correlation Analysis:** Investigate correlations between processing methods, origin regions, and coffee quality scores. Determine if certain processing methods or origin regions are associated with higher coffee quality scores.
- **Defect Occurrences and Impact:** Analyze trends and patterns in defect occurrences (Category One and Category Two defects) and their impact on overall coffee quality. Understand how the presence of defects affects the sensory attributes and Total Cup Points.
- **Interaction Effects:** Explore how different variables interact to influence Total Cup Points, which serve as an overall measure of coffee quality. This involves understanding the combined effect of sensory attributes, processing methods, origin regions, and defect occurrences on Total Cup Points.

By addressing these research questions, the project aims to provide valuable insights to the Coffee Quality Institute and other stakeholders in the coffee industry. These insights can inform decision-making processes related to coffee production, processing, and quality control, ultimately contributing to the promotion of coffee quality standards and the advancement of the specialty coffee industry.

### Data Preparation:

#### **Data Collection:**

- The dataset is provided by the OdinSchool for the Power Bi Capstone Project. The dataset is about the coffee quality and how coffee quality is important in understanding its sales.
- The dataset is of one type of coffee which is Arabica and the file is originally a CVS file and consists of 31 columns with 207 rows.
- The dataset was loaded in the Power Bi to get the insight but before that I have done some cleaning and transformation to get precise information from the data.

**Data Cleaning:** Though the dataset provided is already cleaned but there were some null values in some columns which I removed, which is insignificant to make any effect if removed. Ensured the consistency of data formats across the columns like date, decimals and other units of measures.

**Data Transformation:**

- There are some columns which are transformed like splitting has been done for getting insights like splitting the Altitude column having range using delimiter (like 1700-1900) and taking the average of lower of that range.
- The data type of altitude column is changed from Text to Whole number to get the average of the altitude for creating a new table containing Distinct Countries with other attributes.
- New table created named Countries which consist of Distinct countries with respective average of Altitude and average of Total Cup Points. Which is used to get insight how altitude affects the coffee quality.

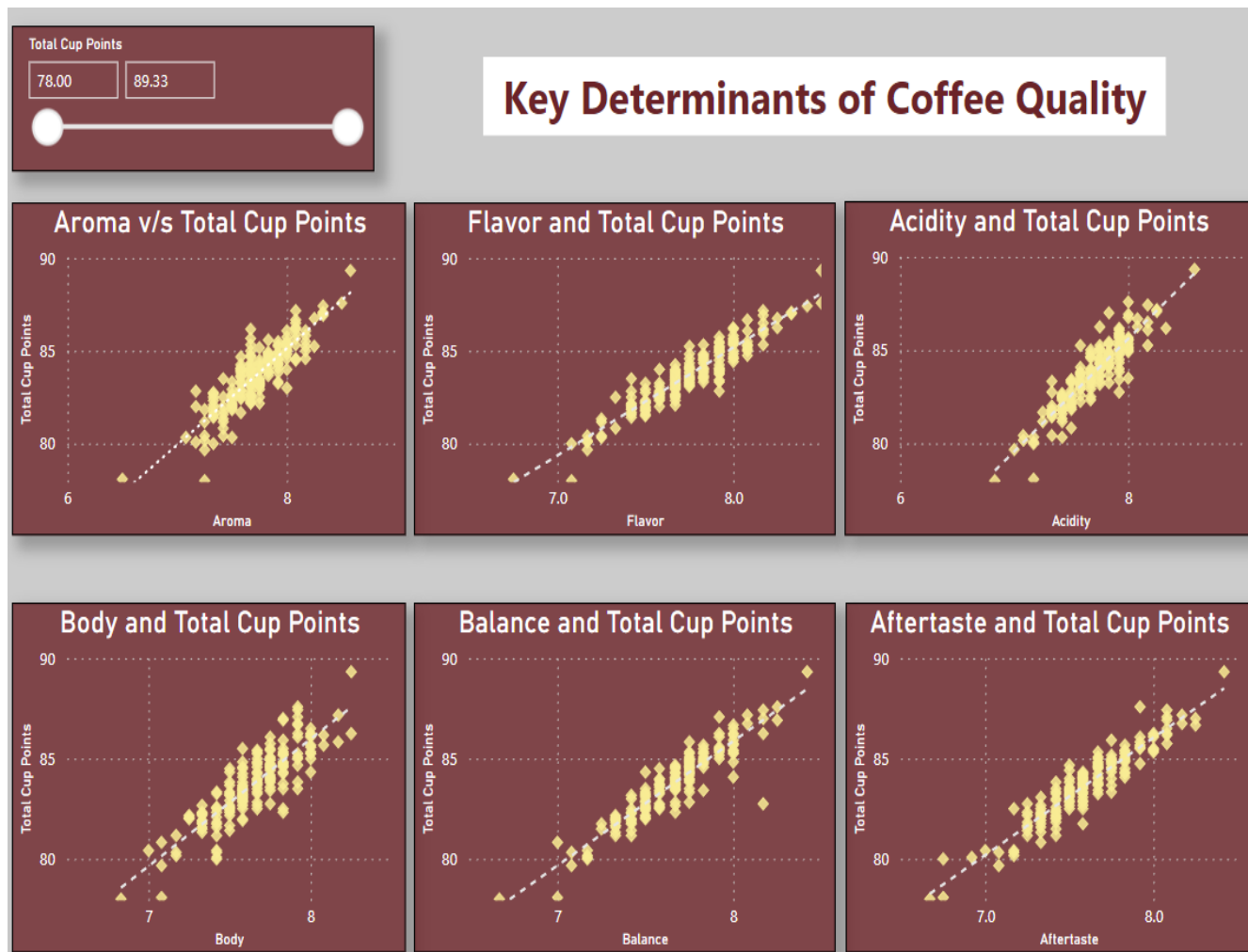
**Visualization and Insights:**

- The primary goal is to gain insights from the data, revealing how coffee quality scores are affected and identifying the key elements that need to be understood to provide high-quality coffee.

**The most important factor to consider is Total Cup Points:**

1. Cupping is a standardized tasting method used to assess the quality and characteristics of coffee beans. It involves brewing coffee samples and evaluating them in a controlled environment. This process allows cuppers (professional coffee tasters) to objectively analyze and compare various coffees based on specific criteria.
2. Typically, a cupping score ranges from 0 to 100, with 100 being the highest possible score. However, it's important to note that a perfect score is quite rare and often elusive. In practice, the majority of high-quality specialty coffees usually fall within the range of 80 to 90 points.
3. During a cupping session, several attributes are considered when assigning a score to a coffee. These attributes include fragrance and aroma, acidity, body, flavor, aftertaste, sweetness, uniformity, and overall balance. Each attribute is evaluated based on its intensity, clarity, and overall sensory experience.

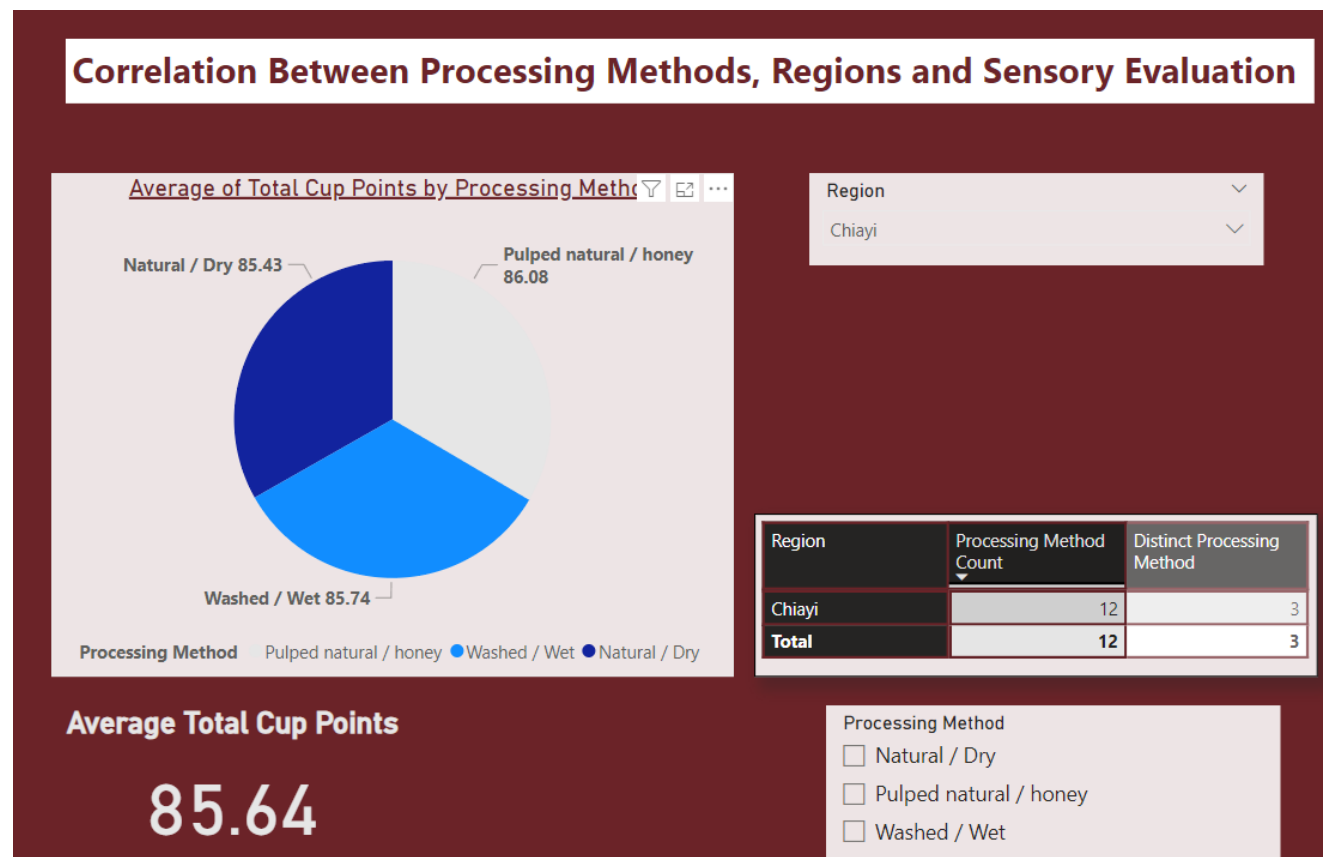
## 1. Key Determinants of Coffee Quality:



The major factors affecting the coffee quality is Aroma, Flavor, Acidity, Body, Balance and Aftertaste. But after analyzing the correlation in between these factors and the Total Cup Points we got the insight which says the Flavor is most important factor followed by Aftertaste and Balance.

We can see in the visuals that Total Cup Points are increasing linearly with the Flavor. But we cannot neglect other factors also.

## 2. Correlation Between Processing Methods, Origin Regions and Coffee Quality Scores:

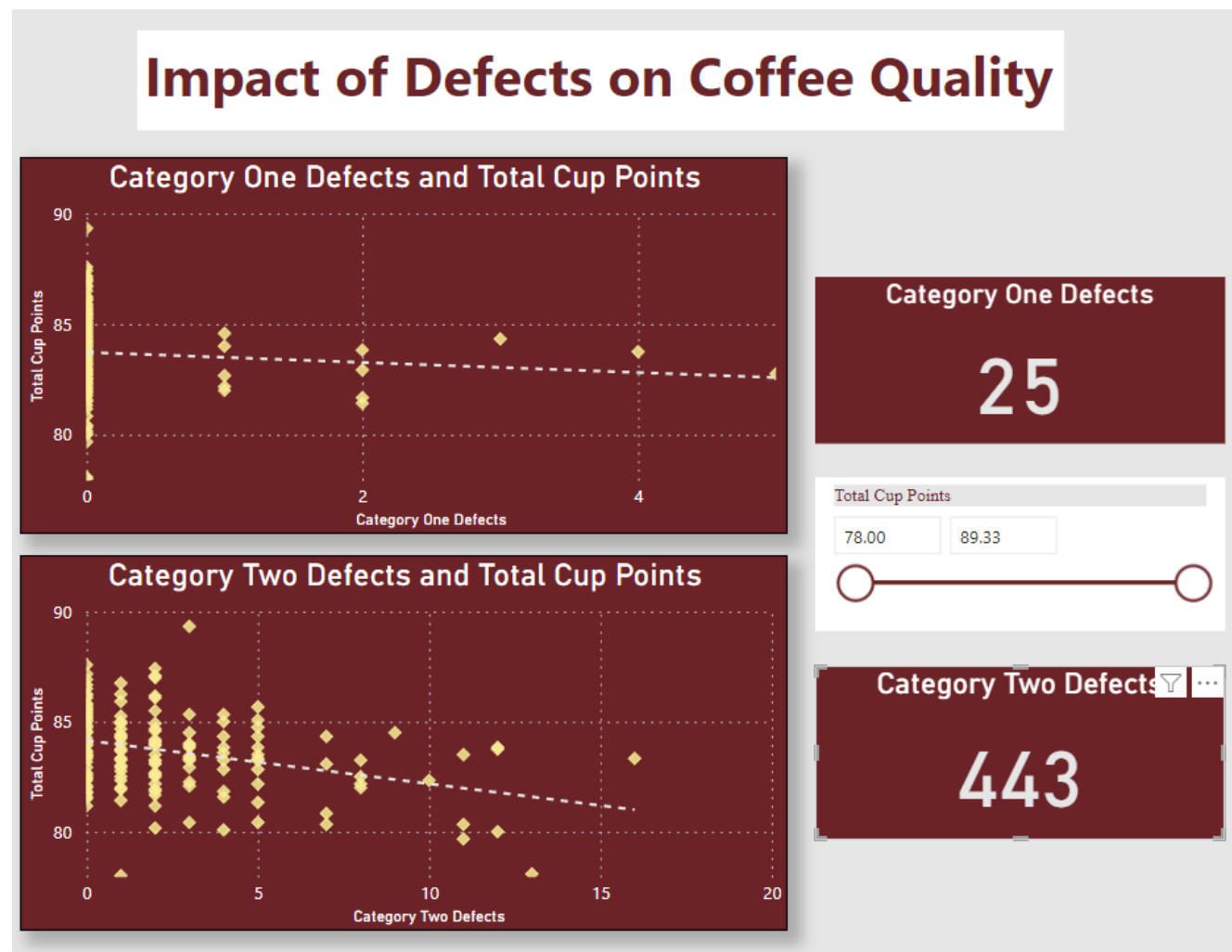


There are regions like Piendamo, Cauca (89.33), Laos Borofen Plateau, Popayan, Chimaltenango (87) having high Total Cup Points means these are the region which gives high quality coffee. Likewise if we talk about the the low quality coffee than its Minas Gerais, La Paz, El Salvador (78).

If we talk about the Processing Methods, so method like Double Anaerobic Washed or Semi Washed process the coffee beans finely to give high quality beans but process like Semi Lavado not able to match the former process quality.

The relation says that the Region like Piendama, Cauca are using the Processing Methods like Double Anaerobic Washed to produce high quality coffee beans.

### 3. Defects and Their Impacts of Coffee Quality:

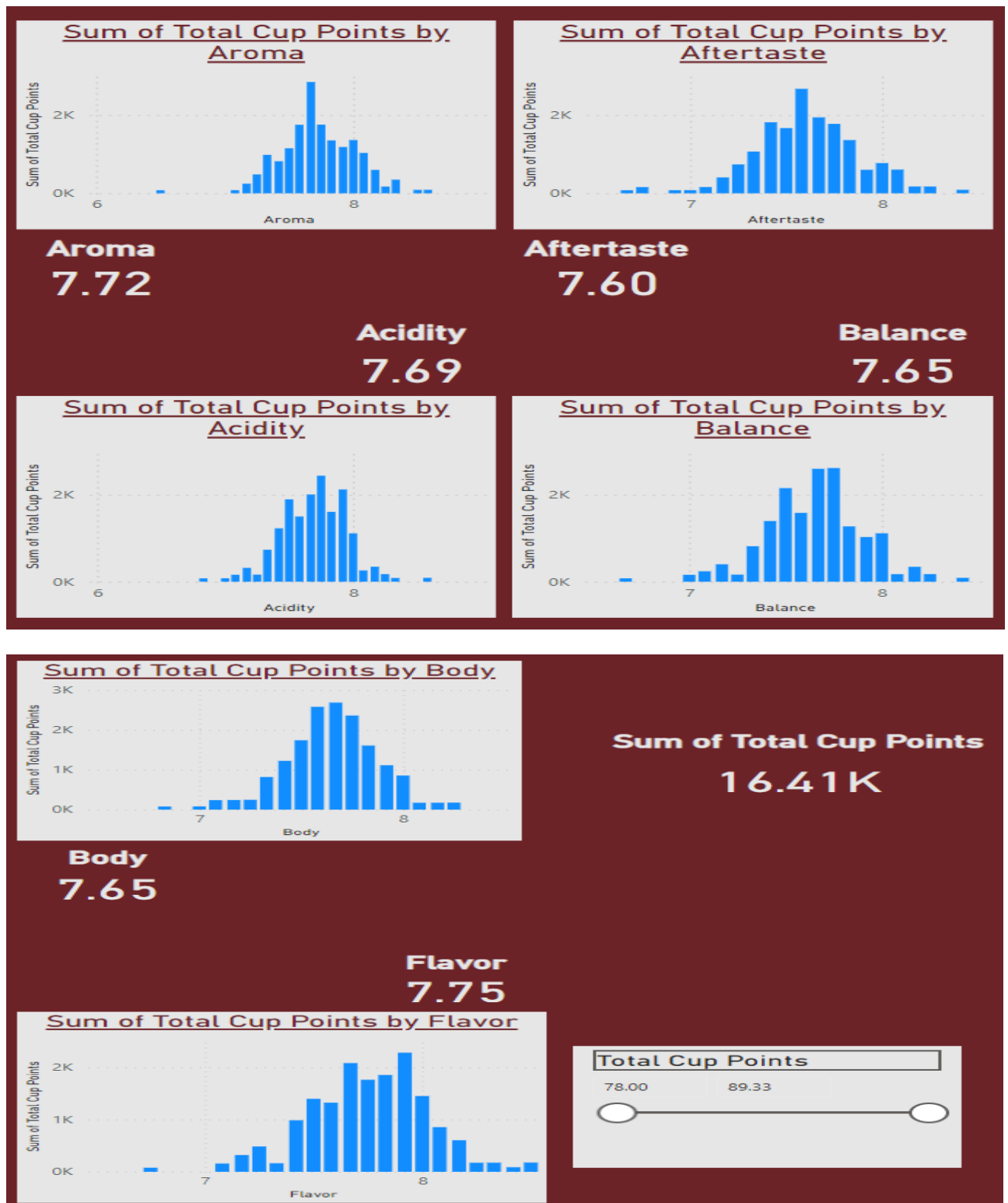


The impact of defects are obviously not positive on the coffee quality, but if we talk about how category 1 and 2 defects affects the quality, then we can say that the increase in Category One defects decreases the coffee quality significantly than Category Two defects.

Although the insights are saying the increase in Category Two defects decreases the Total Cup Points but we also need to see scale of increase in both the defects and how they impact.

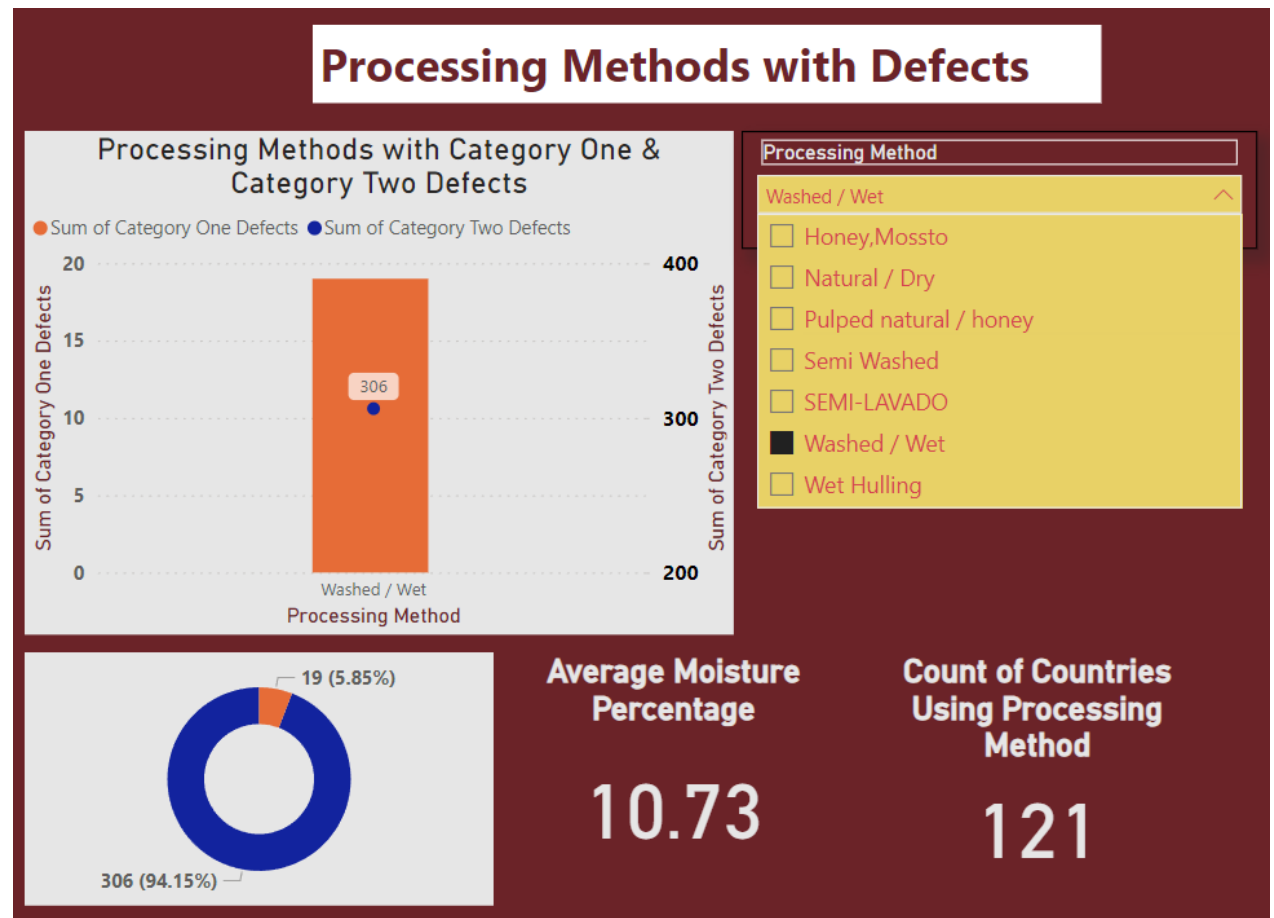
Small increase in Category One defects makes the major impact on the quality as compare to Category Two defects.

#### 4. Different Variables Interact to Influence Total Cup Points:



Out of 9 different variables contributing to the Total Cup Points 3 have been eliminated as they either were constant or insignificantly affecting the quality.

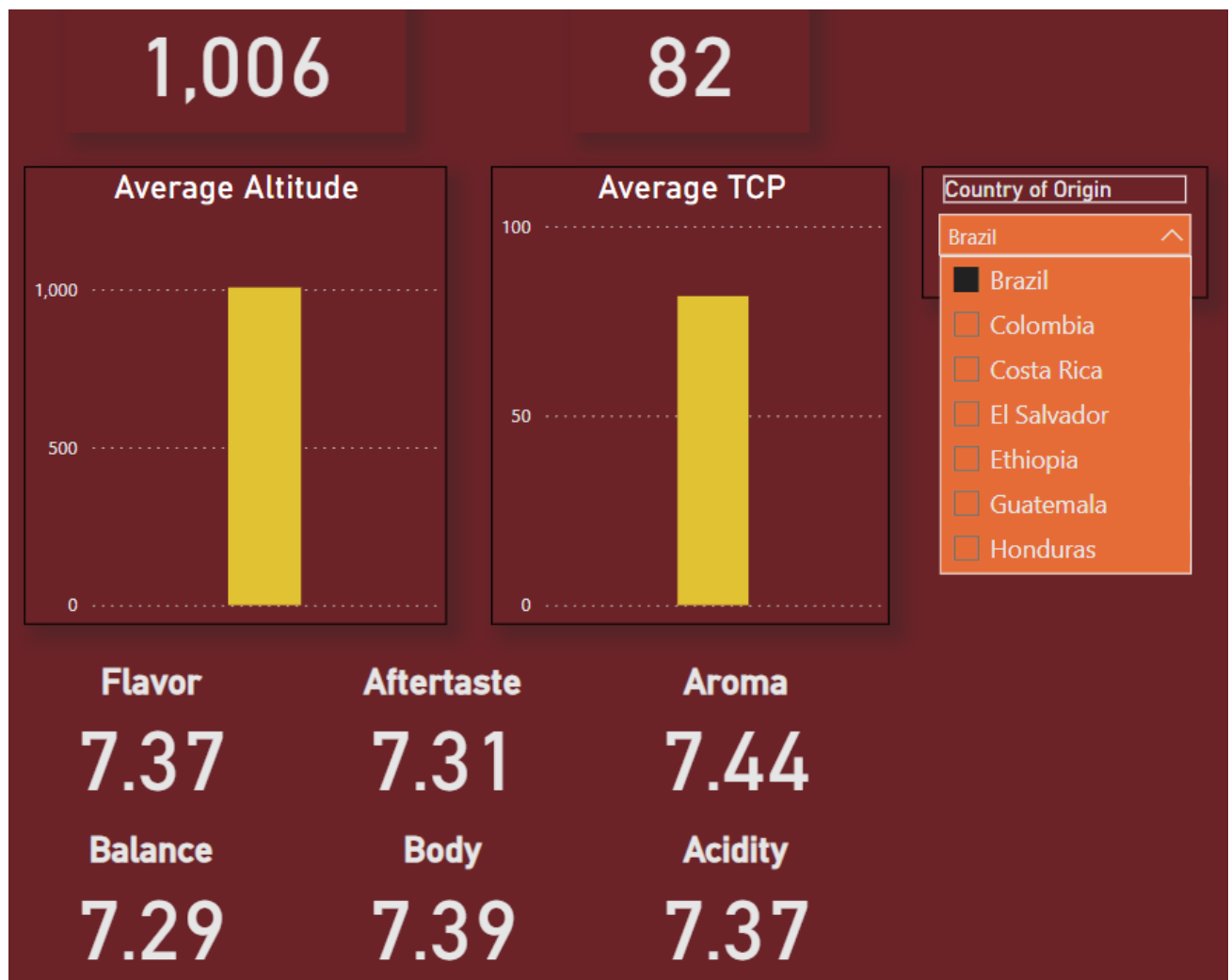
## 5. Relation Between Processing Methods and Category of Defects:



There are methods which are not providing quality beans and increases the defects and moisture percentage in the coffee beans.

Majority of the countries are using such methods like Washed/Wet method which is cost effective but not much effective when talk about detecting defects or removing the moisture.

## 6. Countries with their Altitude Affecting Total Cup Points:



Countries having Tropical and Semi-tropical climate is favorable for coffee production. It requires significant amount of rainfall but water should not be stagnant. Though it can tolerate dry climate but direct sunlight is not good for good coffee quality.

Coffee is generally planted in hilly areas where hot rainy climate is present and fog is good as it protects the beans from direct sun light.

So high altitude with high rainfall and high moisture is required to get the high quality coffee beans.



## **Recommendations:**

Based on the analysis conducted, the following recommendations can be made:

- **Focus on Flavor:** Given that Flavor has the strongest correlation with Total Cup Points, it is essential to prioritize efforts to enhance and maintain the flavor profile of coffee beans.
- **Consider Multiple Factors:** While Flavor is critical, it is important not to overlook other factors such as Aroma, Aftertaste, and Balance, as they also contribute to overall coffee quality.
- **Identify High-Quality Regions:** Regions such as Piendamo, Cauca, and Laos Borofen Plateau consistently produce high-quality coffee beans. Consider sourcing beans from these regions to ensure quality.
- **Optimize Processing Methods:** Methods like Double Anaerobic Washed have shown to produce high-quality beans, while others like Semi Lavado may not meet the desired quality standards. Focus on refining processing methods to improve bean quality.
- **Address Defects:** Reduce the occurrence of Category One defects, as they have a more significant impact on coffee quality compared to Category Two defects. Implement measures to detect and mitigate defects during processing.
- **Evaluate Climate Conditions:** Favorable climate conditions, including high altitude, ample rainfall, and proper moisture levels, are crucial for producing high-quality coffee beans. Ensure that coffee plantations are located in suitable regions with the right environmental conditions.
- **Implement Quality Assurance Measures:** Invest in quality assurance measures throughout the production process to maintain consistency and ensure that only the highest quality beans are selected for further processing.

By implementing these recommendations, coffee producers can enhance the quality of their beans and meet the demands of consumers who seek premium coffee products.