Coffee Quality Analysis

Joseph Rios

Vignesh Sivanandha Rao

CSC735: Data Analytics

Missouri State University

Abstract

The idea of this project is to analyze a dataset which has a list of scores given to a wide variety of coffee beans classified based on various categories. This includes a sum of 10 different subjective scoring fields and few other taste testing fields. The dataset has a list of attributes including Species, Country of Origin, Altitude, Region, Harvest Year, Processing Method etc. In this report, we will discuss on what basis the quality of coffee is judged and the scores are given and how external factor analysis, variety analysis, processing method analysis, country of origin analysis, producer analysis can affect the scores of coffee and the overall best and worst factors.

Introduction

In this analysis, we will be exploring the broad question: “What makes a good cup of coffee?” More specifically, what factors can one use to meaningfully predict the quality of a cup of coffee without being an expert coffee taste testers?

Our dataset comes from the Coffee Quality Institute’s review pages. Coffee producers submit their coffee to this organization to be quality tested. If their coffee gets a score of 80 or more, then it is given a special certificate saying it met their standards. The point here is that only producers with already decent coffee will attempt to get the certificate, because there is a cost associated with undergoing each evaluation. For this reason, as will be seen, all the scores are above 50, even though there surely exists coffee with quality lower.

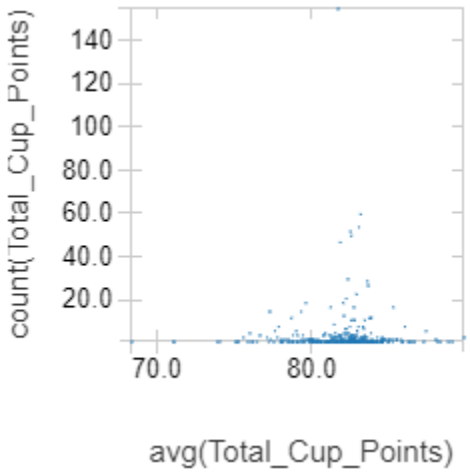
The scoring is done as a sum of 10 different fields: Aroma, Flavor, Aftertaste, Acidity, Body, Balance, Uniformity, Clean Cup, Sweetness, and “Cupper” Points. Most of these are somewhat subjective scorings. The graders follow rigid guidelines when assigning scores, but the final number is usually not discrete. These categories are: Aroma, Flavor, Aftertaste, Acidity,

Body, and Balance. The other group of categories follows a procedure of scoring. In the process of taste testing, 5 cups are tested, and 2 points are awarded if the criteria is met. In this category, there will be a lot 10.0 or 9.0 scores, instead of something like 8.47. These scores are: Uniformity, Clean Cup, and Sweetness. The final score "Cupper" Points is individual score given by the tester. If the tester felt that the individual scores did not adequately reflect the overall quality of the coffee, they may award a higher score and vice versa.

The dataset also includes values about where the coffee came from, how it was processed, and variety. Later in the analysis we will explore, how these external factors affect the scores of the coffee. The two types of coffees, the Coffee Quality Institute grades are Arabica and Robusta. Subspecies of these make up the varieties. Examples include Caturra, Bourbon, and Typica. The processing method deals with how the coffee beans were handled before milling. Different processing methods are known to give vastly different flavor profiles. For example, dry processed coffees tend to give a more full body and wet processed coffees tend to give a lighter, more balanced flavor. The dataset also includes many pieces of information on where the coffee came from. This includes the altitude of the farm, the country, and the year it was harvested.

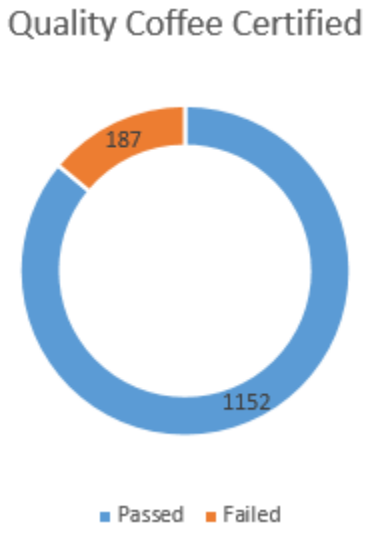
All the data explained above will be explored. There is a lot of other extra information regarding the certification process, but it is not relevant to this analysis. We are forced to assume that the graders doing the testing are reliable. There is a lengthy and expensive process to become a grader, so it can be assumed that there methods are indeed rigorous and objective. There may be some graders who more accurately judge certain coffee, but there is not a good way to check that, since only 1 grader judges the coffee. Maybe some graders give higher scores on average, but without having some baseline to compare it to, one cannot judge its accuracy.

Scores Analysis



	Mean	Median	Std. Dev.
Aroma	7.57	7.58	0.377
Flavor	7.52	7.58	0.398
Aftertaste	7.4	7.42	0.404
Acidity	7.54	7.58	0.38
Body	7.52	7.5	0.359
Balance	7.52	7.5	0.415
Uniformity	9.83	10	0.554
Clean Cup	9.84	10	0.764
Sweetness	9.85	10	0.629
Cupper Points	7.5	7.5	0.473
Total Cup Points	82.09	82.5	3.499

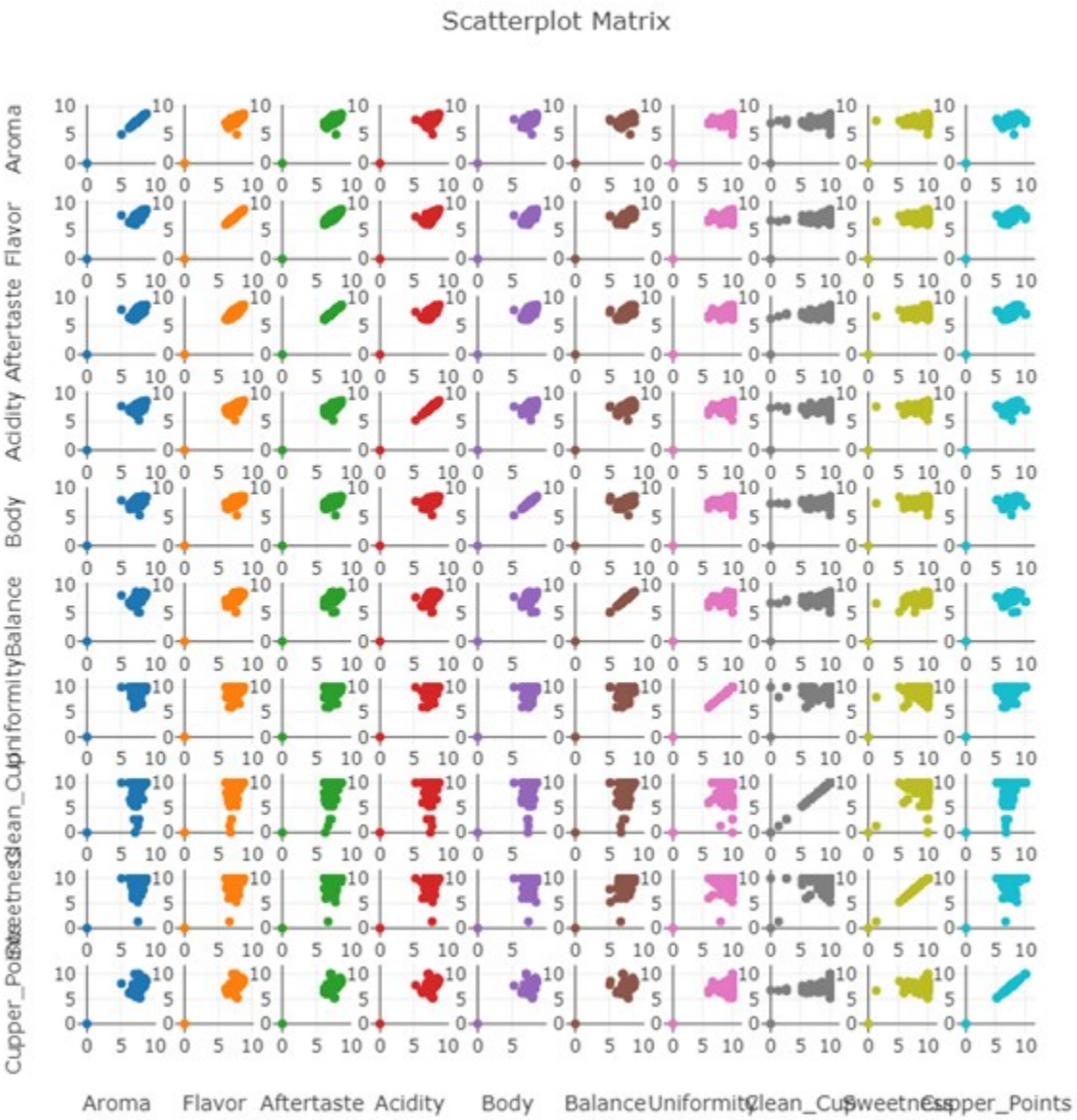
To start, there are the baseline statistics for the each of the scores, 7.5 seems to be around the trend for Aroma, Flavor, Aftertaste, Acidity, Body, Balance, and Cupper Points, whereas 10.0 is the trend Uniformity, Clean Cup, and Sweetness. The main cause of this discrepancy is the way the grading process works for those (See Introduction). There are also probably some comparatively lower scores in the dataset, which are outliers causing the difference between the medians and the means. As can be seen in the above scatter plot, there is one datapoint at 0. The above scatter plot is for owners. As can be seen, there a relationship similar to a bell-curve for the number of gradings has vs. their average score. Owners with few submissions make up the best and worst of submitted gradings.



The pie chart to the right shows how many gradings have passed and failed. The vast majority of submitted coffee get a total score greater than 80 and are given their coffee quality certification. The success rate is a staggering 86%.

Below is a scatter plot matrix of each of the individual scores for all of the graded coffee, generated using Python and plotly. This was generated to see if there were any identifiable trends among individual scores that are visible just on inspection. It is already pretty obvious that Uniformity, Clean Cup, and Sweetness exhibit the least linear relationship among other individual scores. To measure how well the different individual scores correlate, one can apply a linear regression to each pair of categories. For each pair, the root mean squared error (RMSE)

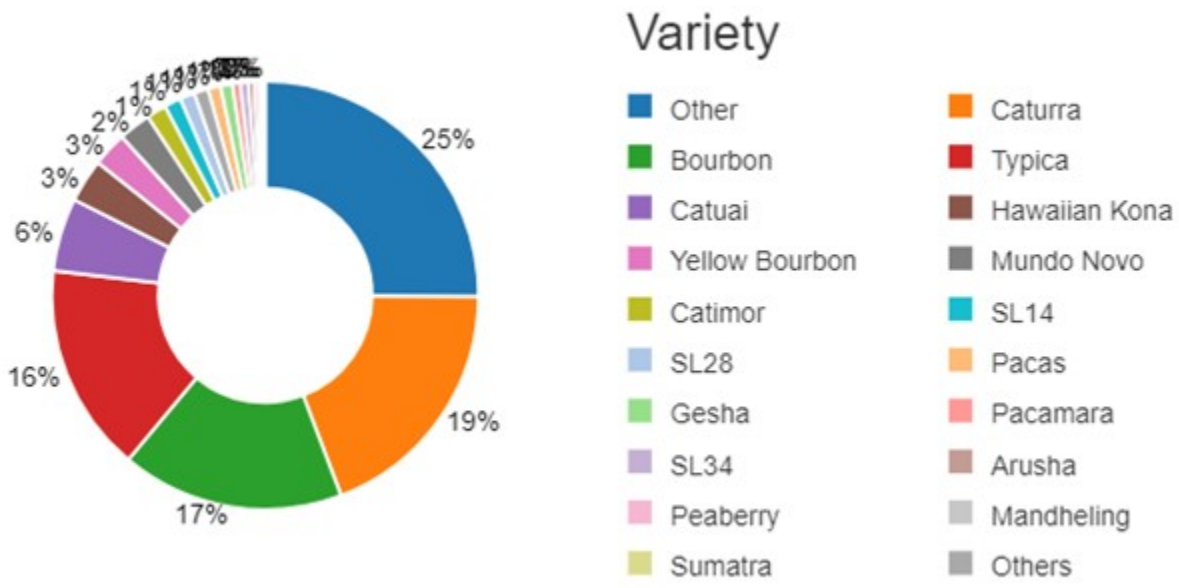
was recorded and collated into the table below the scatter plot. The yellow squares indicate the top 25% most correlative categories and red the bottom 25%. The lower RMSE indicates that the line of best fit generated by the linear regression fits with a greater accuracy. As expected, Uniformity, Clean Cup, and Sweetness exhibited the least linear relationships. Flavor and Aftertaste were the overall most related categories, but the most interesting part are the Yellow squares in the “Cupper Points” row. It is important, because this shows that Flavor and Aftertaste most directly affect the overall score given by the grader. These categories can be used to predict the overall score, without having/knowing every individual score.



	Aroma	Flavor	Aftertaste	Acidity	Body	Balance	Uniformity	Clean Cup	Sweetness	Cupper Points
Aroma		0.23177	0.25449	0.26207	0.25852	0.29723	0.51579	0.71935	0.60907	0.34379
Flavor	0.21962		0.1798	0.21911	0.23252	0.25947	0.50541	0.70092	0.60305	0.28822
Aftertaste	0.23756	0.17713		0.23168	0.23278	0.24917	0.50775	0.70426	0.60765	0.29313
Acidity	0.26051	0.22985	0.24671		0.24387	0.28746	0.51384	0.72826	0.60818	0.33885
Body	0.2719	0.25808	0.26227	0.25803		0.28186	0.52121	0.73127	0.61098	0.35113
Balance	0.27039	0.2491	0.24283	0.26308	0.2438		0.50976	0.71143	0.6002	0.33472
Uniformity	0.35115	0.36311	0.3703	0.35192	0.33738	0.38149		0.64938	0.56441	0.44162
Clean Cup	0.35552	0.36557	0.37287	0.36209	0.34363	0.3865	0.47142		0.5675	0.4421
Sweetness	0.36555	0.38196	0.39069	0.36721	0.34867	0.39598	0.49758	0.68917		0.46333
Cupper Points	0.27415	0.24255	0.25041	0.27184	0.26624	0.29341	0.51729	0.71334	0.61561	

External Factors Analysis

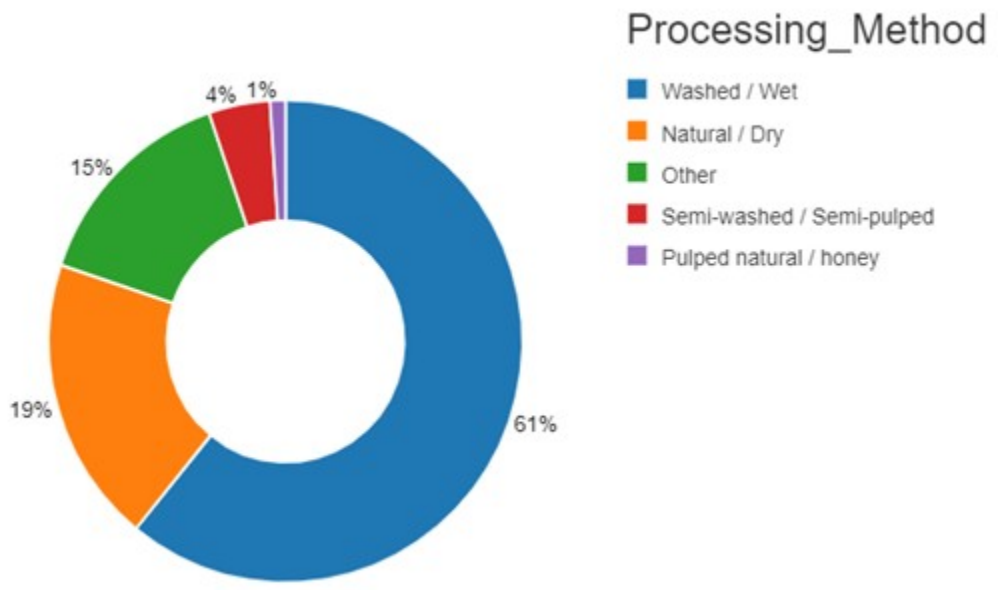
The first external factor that will be analyzed is variety. The variety of the coffee is subspecies of bean. This section is used to see which varieties score better. Firstly, there is the popularity of each variety. As can be seen, “Other” is the most popular. This is not a variety, but rather the absence of a variety on the grading. This could mean that producers are secretive about the specific varieties they use or that the breeding done by producers has made a variety that has no official name. It is worth noting that the coffee that is submitted by producers is usually high quality from the beginning. Therefore, it may be enough to say that Caturra, Bourbon, and Typica yield high quality coffee, because these varieties are most popular among producers who submit coffee to the Coffee Quality Institute.



Next, we have a breakdown of highest scoring varieties on average. Ethiopian Yirgacheffe seems to be the big winner. However, the detail not show, is that that variety only has 2 submitted gradings. It does not seem fair to draw conclusions about the variety as a whole with so few data points. This is what the second table takes into account. It only shows varieties who already have more than 10 submitted gradings. It is also only showing the individual categories that matter, since the first table showed that many varieties get an average score of 10.0 for Uniformity, Clean Cup, and Sweetness. Thus, it looks like SL28 and SL14 are the winners.

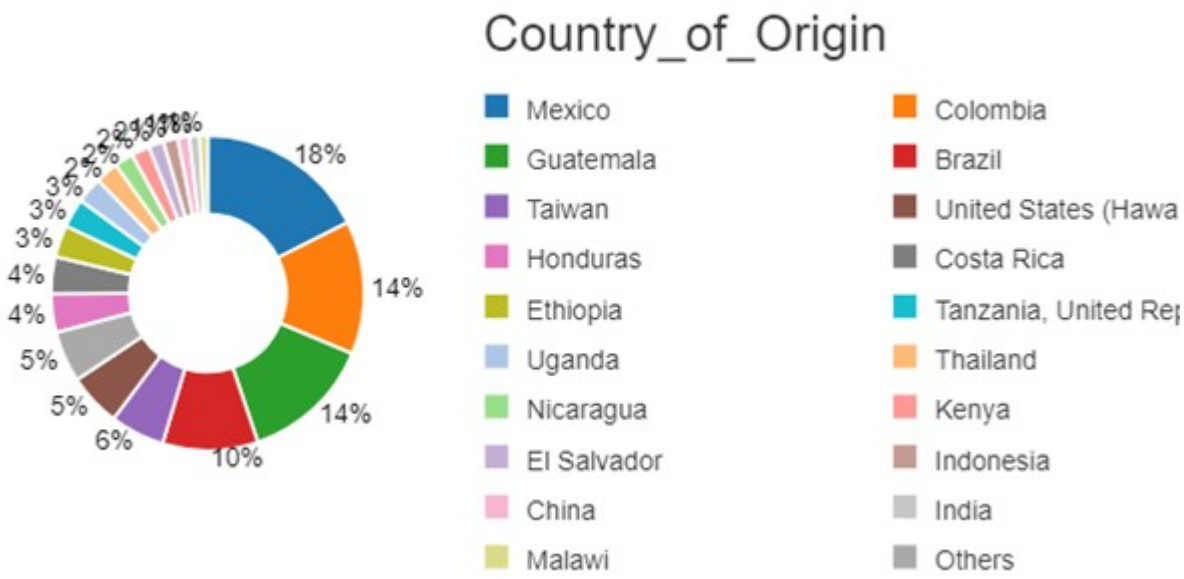
Variety Analysis				
Individual Score	Variety with Highest Average Score	Highest Average Score	Variety with Highest Max Score	Maximum Score
Aroma	Ethiopian Yirgacheffe	8.125	Other	8.75
Flavor	Ethiopian Yirgacheffe	8	Other	8.83
Aftertaste	Ethiopian Yirgacheffe	7.875	Other	8.67
Acidity	Sumatra Lintong	8	Other	8.75
Body	Sumatra Lintong	8.08	Other	8.58
Balance	Ethiopian Yirgacheffe	8	Other	8.75
Uniformity	(Many tied)	10	(Many tied)	10
Clean Cup	(Many tied)	10	(Many tied)	10
Sweetness	(Many tied)	10	(Many tied)	10
Cupper Points	Ethiopian Yirgacheffe	8.125	Other	10

Variety Analysis with Adjusted Averages (Having more than 10 gradings)		
Individual Relevant Scorings	Variety with Highest Adjusted Average Score	Highest Adjusted Average Score
Aroma	SL14	7.88
Flavor	SL28	7.73
Aftertaste	SL28	7.69
Acidity	SL28	7.82
Body	SL28	7.7
Balance	SL28	7.81
Cupper Points	SL14	7.74



The second external factor to analyze is the processing method. This is deals with specifically how the fruit are handled after being harvested. For example, in “Natural/Dry” the whole fruit is left to dry in the sun, before being cleaned and sorted. Essentially, what is being asked is: what processing methods are good for high quality coffee? One can first look at the popularity of each method. Washed/Wet is the most popular among coffee that was submitted, so just like before, it may be enough to say that coffee processed this way is high quality on average. Then, there is also the individual category breakdown. It seems the lesser used methods yield higher scores on average, but the most popular method yields the highest maximum scores for most relevant categories.

Processing Method Analysis				
Individual Score	**Processing Method with Highest Average Score**	**Highest Average Score**	**Processing Method with Highest Max Score**	**Maximum Score**
Aroma	Semi-Washed / Semi-Pulped	7.6	Washed / Wet	8.75
Flavor	Semi-Washed / Semi-Pulped	7.59	Washed / Wet	8.83
Aftertaste	Pulped Natural / Honey	7.52	Washed / Wet	8.67
Acidity	Other	7.56	Washed / Wet	8.75
Body	Pulped Natural / Honey	7.63	Other	8.58
Balance	Natural / Dry	7.59	Other	8.75
Uniformity	Pulped Natural / Honey	10	(Tied)	10
Clean Cup	Pulped Natural / Honey	10	(Tied)	10
Sweetness	Pulped Natural / Honey	10	(Tied)	10
Cupper Points	Other	7.63	Other	10



The third external factor to be analyzed is country of origin. The most popular country for gradings to come from seems to be Mexico. It is hard to draw generalities about the country of origin, but one can say that the Coffee Quality Institute is most popular in Mexico. The most interesting part here is in the individual score breakdown. In the same way as with variety, there are some countries which have only 1 or 2 gradings, so it is hard to draw conclusions about the whole country based on that. However, the best country based on a simple average

is Papua New Guinea, even though they only have a single grading. If one filters by countries with more than 10 gradings, Ethiopia becomes the decisive winner. Not only does it have the maximum score for nearly every category, but it also is the best on average.

Country of Origin Analysis				
Individual Score	**Country with Highest Average Score**	**Highest Average Score**	**Country with Highest Max Score**	**Maximum Score**
Aroma	Papua New Guinea	8.33	Ethiopia	8.75
Flavor	Papua New Guinea	8.42	Ethiopia	8.33
Aftertaste	Ethiopia	7.89	Ethiopia	8.67
Acidity	Papua New Guinea	8.33	Ethiopia	8.75
Body	Japan	8.08	Ethiopia	8.58
Balance	Papua New Guinea	8.25	Ethiopia	8.75
Uniformity	(Many Tied)	10	(Many Tied)	10
Clean Cup	(Many Tied)	10	(Many Tied)	10
Sweetness	(Many Tied)	10	(Many Tied)	10
Cupper Points	Japan	8.08	Taiwan	10

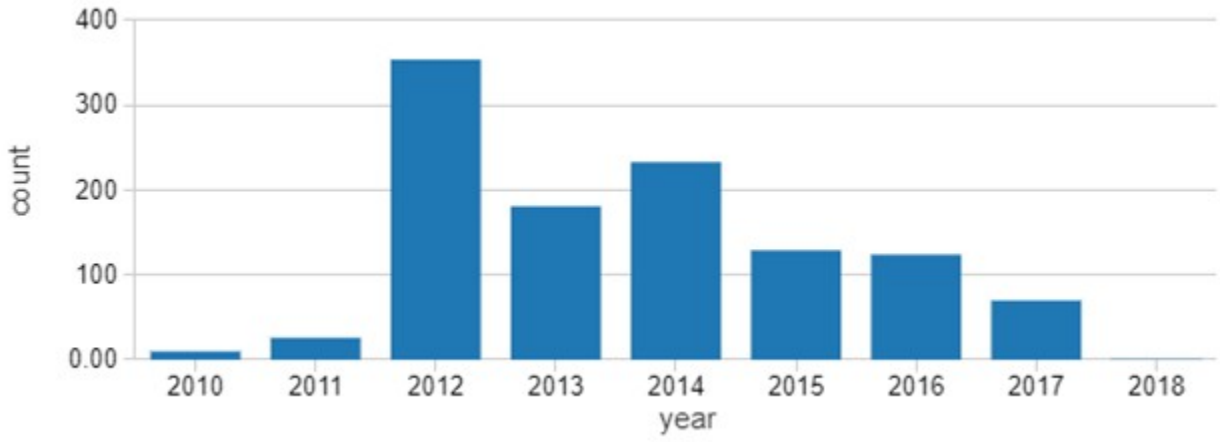
Country of Origin Analysis with		
Individual Relevant Scorings	Country with Highest Adjusted Average Score	Highest Adjusted Average Score
Aroma	Ethiopia	7.9
Flavor	Ethiopia	8.01
Aftertaste	Ethiopia	7.89
Acidity	Ethiopia	8.04
Body	Ethiopia	7.92
Balance	Ethiopia	7.97
Cupper Points	Ethiopia	8.05

Top 5 Most Prolific Producers (Owners)	
Juan Luis Alvarado Romero	155
Racafe & Cia S.C.A	60
Exportadora de Cafe Condor S.A	54
Kona Pacific Farmers Cooperative	52
Ipanema Coffees	50

The fourth external factor to be analyzed is the producer. The most prolific producer is Juan Luis Alvarado Romero, which explains why Mexico was the country with the most submitted coffee. Since, there were so many producers and one graph has already shown about many that is there will be no pie chart. Also, it has already been shown that producers who produce very few make up the best and worst of submitted coffee. Therefore, there will also be no table of simple averages of Producers. The table to the right skips that only shows producers with more than 10 gradings. Once again, Ethiopia shows that its coffee is superior, because the Ethiopia Commodity Exchange makes up the top averages of all the important categories.

The fifth external factor is harvest year. As can be seen from the bar graph, 2012 was the year most submitted coffees were harvested. This would probably be the year that the Coffee Quality Institute started to gain popularity. Over the years, the number of submitted coffees has decreased. This may be because certifications given by the Coffee Quality Institute gives enough clout, despite expiring after some time. So producers who would normally want their coffee graded, do not re-submit since they already got the necessary certification. Now, some harvest years may be better for different geographic regions, so the next tables show

groupings of countries and harvest years to see where and when the best coffee was harvested.



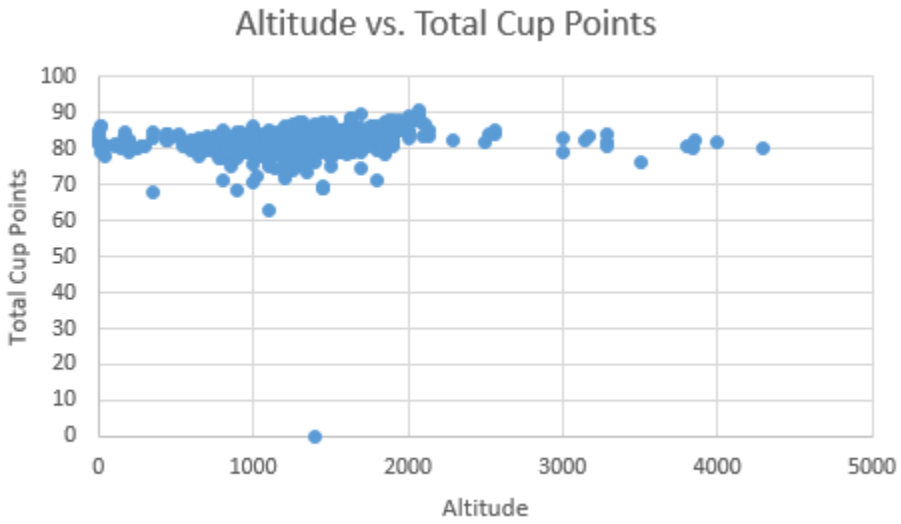
Where & When the Most Beans were harvested for grading (Top 5)		
Country of Origin	**Harvest Year**	**Number of Gradings**
Mexico	2012	182
Guatemala	2012	46
Colombia	2013	40
Guatemala	2014	33
Colombia	2012	31

Top 5 Best Harvest Years for Different Countries (Having more than 10 gradings)		
Country of Origin	**Harvest Year**	**Average Adjusted Score**
Ethiopia	2014	86.59
Brazil	2011	84.52
Colombia	2015	83.79
China	2015	83.77
Colombia	2014	83.54

Once again, Mexico shows that they submit the most coffee, and now it is known that this happened in 2012. Also, Ethiopia has the average score in 2014. The hypothesis about harvest years being better for different regions proves accurate, since the top scores are fairly unique.

The sixth and final external factor is altitude. This refers to the altitude of the farm in which the coffee was harvested. The above graph shows every grading with a valid mean altitude in meters by its total score. There does not appear to be any obvious trends between the 2 factors. However, it does appear that the 2000m mark is something of a sweet spot, before dropping off at values past it. It also appears that altitudes between the values of 1000m and 2000m are the most popular among producers who submit to the Coffee Quality Institute.

The last table is a summary of the best and worst external factors. The table uses simple averages, which may be misleading without recognizing how few gradings some have. See the above tables that take that into account.



Overall Best and Worst Factors		
	Factor	**Total Cup Points**
Best Average Variety	Ethiopian Yirgacheffe	85.96
Worst Average Variety	Pacas	80.91
Variety of Best	Other	90.58
Best Average Processing Method	Pulped Natural / Honey	82.81
Worst Average Processing Method	Washed / Wet	81.97
Processing Method of Best	Washed / Wet	90.58
Best Average Country (For Coffee)	Papua New Guinea	85.75
Worst Average Country (For Coffee)	Haiti	77.18
Country of Best (Coffee)	Ethiopia	90.58
Best Average Owner	Metad plc	89.78
Worst Average Owner	Juan Carlos Garcia Lopez	68.33
Owner of Best	Metad plc	90.58

Summary

Based on the input dataset we have done various analysis on the conditions, external factors, variety, processing method, country of origin that affect the scores of coffee. Additionally we have also analyzed the top prolific producers/owners of the coffee with and

without adjusted averages, time and place of the harvest for top gradings, years for different countries for top gradings. Based on all the individual analysis done, we have also categorized the factors and listed the total cup points received for each category.