Table of Contents

Problem Statement 2

Introduction 2

Research Objectives 4

Analysis Results & Interpretations 5

Questions 2: Factor Analysis (Principle Component Analysis) 12

Conclusion & Recommendation 19

References 20

Figure 1: SCAA Coffee Flavour Wheel 3
Figure 2: Generated Spreadsheet 5
Figure 3: Scree Plot 15

Table 1: Summary of Variables 4
Table 2: Hypothesis of Research Objectives 4
Table 3: Descriptive Statistics 5
Table 4: Correlation Table 6
Table 5: Variable Entered 7
Table 6: Model Summary 8
Table 7: ANOVA Analysis 9
Table 8: Coefficients 10
Table 9: Excluded Variables 11
Table 10: Correlations Table 12
Table 11: PCA Descriptive Statistics 13
Table 12: KMO & Bartlett's Test 13
Table 13: Communalities 14
Table 14: Total Variance 14
Table 15: Component Matrix 16
Table 16: Rotated Component Matrix 17
Table 17: Component Transformation Matrix 17

Problem Statement

What is the effect and cause of independent variables toward dependent variable in the dataset provided by Coffee Quality Institute (CQI). How aroma, flavour, aftertaste, acidity, body, balance, uniformity, clean-cup, sweetness and cupper points affect the total cupper point which indicates the quality of a coffee. Furthermore, how multivariate methods for data analysis such as multiple linear regression and factor analysis can help researcher to identify the best coffee quality with the help of IBM SPSS statistical software.

Introduction

According to (Yerkes, 2014) CQI is an international non-profit organisation who has a mission to enhance quality of coffee and the living quality of the coffee farmers. Most of the farmers do not have access to the right tools and support to fathom the quality of their harvested coffee nor they have access to the market that rewards them based on the quality of the produced coffee. Also, CQI focuses on farmers' self-sufficiency with the help of coffee experts and their decades of experience. Plus, CQI wants to advance market connections that provide healthier supply chains and exuberant coffee communities.

(Luca, 2020) states there are hundred species of coffee bean, the most common types are Arabica and Robusta, this research focuses on Arabica coffee due to its high quality is suitable to make specialty coffee. In order to assess the quality of a coffee, coffee cupping is introduced. The Specialty Coffee Association of America (SCAA) established a systematic solution to assess and record the results of coffee cupping with a standard cupping form. The SCAA's cupping protocols to assess the quality of coffee are aroma, flavour, aftertaste, acidity, body, balance, uniformity, clean cup, sweetness and cupper points. Also, they set a standard operating procedure for coffee cupping brewing method, grind size and roasting method.

Aroma can be classified to dry fragrance and wet fragrance. Dry fragrance is the aroma of the dry coffee beans and ground coffee, the aroma can be nutty, flowery, herbal or fruity. Wet fragrance is the vapor releases from the coffee after boiling water is poured into ground coffee and stirred. Next, the flavour of coffee involves both senses of smell and taste, SCAA introduces a Coffee Flavour Wheel Print as shown in Figure 1, the nine basic flavours are roasted, spices, nutty, sweet, floral, fruity, sour, green and other.

Figure 1: SCAA Coffee Flavour Wheel

Aftertaste provides coffee drinker a lingering coffee taste after swallowing, the longer it lasts the higher the score. A right balance of acidity provides taster liveliness, brightness, sweetness and sparkling sensation while imbalance acidity can cause sour, sharp or bitter taste which leads to lower cupper points. The body of coffee is the concentration that causes taster's mouth feel grainy, oily or watery, in detail it is the density, consistency, weight and texture of the coffee. Balance means the flavour, aftertaste, body and acidity of the coffee is not overpower one another. Sweetness is the amino acids, sugars and alcohols from the coffee, lighter roast is sweeter than darker roast in general. Furthermore, during coffee cupping there are 5 samples for each coffee, uniformity is how consistent the coffee taste throughout these 5 coffees. Clean cup means there is no unpleasant taste such as burnt flavour.

Table 1: Summary of Variables

Variables	Description
Aroma	Dry fragrance and wet fragrance of the coffee
Flavour	Roasted, spices, nutty/cocoa, sweet, floral, fruity, sour/fermented, green/vegetative, chemical/papery
Aftertaste	Lasting taste at the back of the palate
Acidity	The pleasantness of the acid tastes
Body	Density, consistency, weight, texture
Balance	The equilibrium of flavour, aftertaste, body, acidity
Uniformity	Consistency of 5 samples
Clean cup	Absence of unpleasant taste
Sweetness	Intensity of the coffee sugars
Cupper points	Average points
Total cupper points	Overall quality of coffee

Table 1 concludes the variables and its description, all metric independent variables have impact on metric dependent variable: Total cupper points.

Research Objectives

The objectives of this research are:

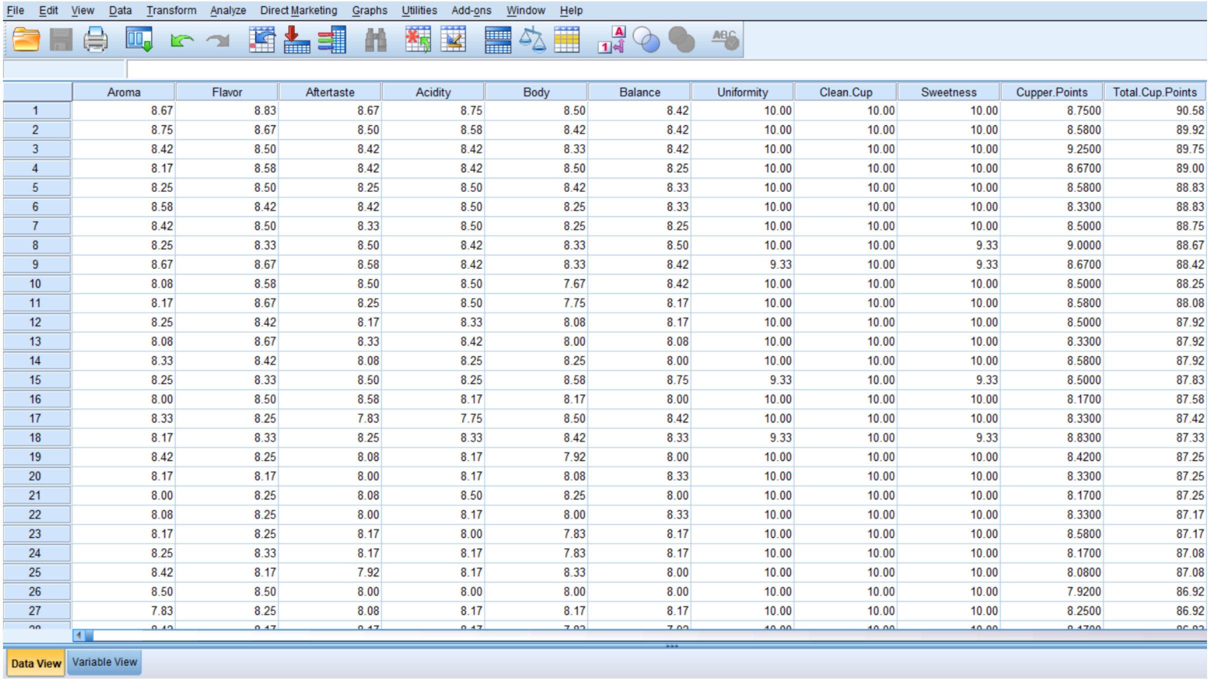
1. To observe the relationship between aroma and total cupper points.
2. To investigate the relationship between balance and total cupper points.
3. To identify the relationship between uniformity and total cupper points.

Table 2: Hypothesis of Research Objectives

H_0	H_1
There is no relationship between aroma and total cupper points	There is relationship between aroma and total cupper points
There is no relationship between balance and total cupper points	There is relationship between balance and total cupper points
There is no relationship between uniformity and total cupper points	There is relationship between uniformity and total cupper points

Table 2 indicates the hypothesis of the research objectives, lets observe how aroma, balance and uniformity or other independent variables will affect total cupper points.

Analysis Results & Interpretations



File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	Aroma	Flavor	Aftertaste	Acidity	Body	Balance	Uniformity	Clean.Cup	Sweetness	Cupper.Points	Total.Cup.Points
1	8.67	8.83	8.67	8.75	8.50	8.42	10.00	10.00	10.00	8.7500	90.58
2	8.75	8.67	8.50	8.58	8.42	8.42	10.00	10.00	10.00	8.5800	89.92
3	8.42	8.50	8.42	8.42	8.33	8.42	10.00	10.00	10.00	9.2500	89.75
4	8.17	8.58	8.42	8.42	8.50	8.25	10.00	10.00	10.00	8.6700	89.00
5	8.25	8.50	8.25	8.50	8.42	8.33	10.00	10.00	10.00	8.5800	88.83
6	8.58	8.42	8.42	8.50	8.25	8.33	10.00	10.00	10.00	8.3300	88.83
7	8.42	8.50	8.33	8.50	8.25	8.25	10.00	10.00	10.00	8.5000	88.75
8	8.25	8.33	8.50	8.42	8.33	8.50	10.00	10.00	9.33	9.0000	88.67
9	8.67	8.67	8.58	8.42	8.33	8.42	9.33	10.00	9.33	8.6700	88.42
10	8.08	8.58	8.50	8.50	7.67	8.42	10.00	10.00	10.00	8.5000	88.25
11	8.17	8.67	8.25	8.50	7.75	8.17	10.00	10.00	10.00	8.5800	88.08
12	8.25	8.42	8.17	8.33	8.08	8.17	10.00	10.00	10.00	8.5000	87.92
13	8.08	8.67	8.33	8.42	8.00	8.08	10.00	10.00	10.00	8.3300	87.92
14	8.33	8.42	8.08	8.25	8.25	8.00	10.00	10.00	10.00	8.5800	87.92
15	8.25	8.33	8.50	8.25	8.58	8.75	9.33	10.00	9.33	8.5000	87.83
16	8.00	8.50	8.58	8.17	8.17	8.00	10.00	10.00	10.00	8.1700	87.58
17	8.33	8.25	7.83	7.75	8.50	8.42	10.00	10.00	10.00	8.3300	87.42
18	8.17	8.33	8.25	8.33	8.42	8.33	9.33	10.00	9.33	8.8300	87.33
19	8.42	8.25	8.08	8.17	7.92	8.00	10.00	10.00	10.00	8.4200	87.25
20	8.17	8.17	8.00	8.17	8.08	8.33	10.00	10.00	10.00	8.3300	87.25
21	8.00	8.25	8.08	8.50	8.25	8.00	10.00	10.00	10.00	8.1700	87.25
22	8.08	8.25	8.00	8.17	8.00	8.33	10.00	10.00	10.00	8.3300	87.17
23	8.17	8.25	8.17	8.00	7.83	8.17	10.00	10.00	10.00	8.5800	87.17
24	8.25	8.33	8.17	8.17	7.83	8.17	10.00	10.00	10.00	8.1700	87.08
25	8.42	8.17	7.92	8.17	8.33	8.00	10.00	10.00	10.00	8.0800	87.08
26	8.50	8.50	8.00	8.00	8.00	8.00	10.00	10.00	10.00	7.9200	86.92
27	7.83	8.25	8.08	8.17	8.17	8.17	10.00	10.00	10.00	8.2500	86.92

Data View Variable View

Figure 2: Generated Spreadsheet

Figure 2 show the spreadsheet of the data, it contains 11 columns and 1310 rows, all data is metric. Multiple stepwise regression is performed between the independent variables and dependent variable, the outputs are as shown in Table 3, Table 4, Table 5, Table 6, Table 7, Table 8 and Table 9.

Table 3: Descriptive Statistics

Descriptive Statistics

	Mean	Std. Deviation	N
Total.Cup.Points	82.1159	3.51576	1311
Aroma	7.5638	.37867	1311
Flavor	7.5181	.39998	1311
Aftertaste	7.3977	.40512	1311
Acidity	7.5331	.38160	1311
Body	7.5177	.35921	1311
Balance	7.5175	.40632	1311
Uniformity	9.8334	.55934	1311
Clean.Cup	9.8331	.77135	1311
Sweetness	9.9033	.53083	1311
Cupper.Points	7.497864	.4746100	1311

Table 3 indicates the mean and standard deviation, which are not so important for stepwise multiple linear regression analysis.

Table 4: Correlation Table

Correlations

		Total.Cup. Points	Aroma	Flavor	Aftertaste	Acidity	Body	Balance	Uniformity	Clean.Cup	Sweetness	Cupper. Points
Pearson Correlation	Total.Cup.Points	1.000	.797	.878	.866	.801	.777	.836	.658	.661	.617	.802
	Aroma	.797	1.000	.814	.777	.725	.696	.717	.366	.335	.328	.693
	Flavor	.878	.814	1.000	.895	.817	.762	.803	.409	.396	.362	.797
	Aftertaste	.866	.777	.895	1.000	.792	.761	.823	.399	.386	.343	.788
	Acidity	.801	.725	.817	.792	1.000	.733	.742	.373	.299	.331	.700
	Body	.777	.696	.762	.761	.733	1.000	.768	.339	.287	.327	.671
	Balance	.836	.717	.803	.823	.742	.768	1.000	.405	.374	.343	.741
	Uniformity	.658	.366	.409	.399	.373	.339	.405	1.000	.525	.538	.359
	Clean.Cup	.661	.335	.396	.386	.299	.287	.374	.525	1.000	.526	.357
	Sweetness	.617	.328	.362	.343	.331	.327	.343	.538	.526	1.000	.300
	Cupper.Points	.802	.693	.797	.788	.700	.671	.741	.359	.357	.300	1.000
Sig. (1-tailed)	Total.Cup.Points	.	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Aroma	.000	.	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Flavor	.000	.000	.	.000	.000	.000	.000	.000	.000	.000	.000
	Aftertaste	.000	.000	.000	.	.000	.000	.000	.000	.000	.000	.000
	Acidity	.000	.000	.000	.000	.	.000	.000	.000	.000	.000	.000
	Body	.000	.000	.000	.000	.000	.	.000	.000	.000	.000	.000
	Balance	.000	.000	.000	.000	.000	.000	.	.000	.000	.000	.000
	Uniformity	.000	.000	.000	.000	.000	.000	.000	.	.000	.000	.000
	Clean.Cup	.000	.000	.000	.000	.000	.000	.000	.000	.	.000	.000
	Sweetness	.000	.000	.000	.000	.000	.000	.000	.000	.000	.	.000
	Cupper.Points	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.
N	Total.Cup.Points	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Aroma	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Flavor	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Aftertaste	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Acidity	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Body	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Balance	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Uniformity	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Clean.Cup	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Sweetness	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
	Cupper.Points	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311

As shown in Table 4, flavour has the largest magnitude or correlation with total cupper points, hence this variable will be first included in the multiple stepwise linear regression and then search for other variables to be included after controlling the effects of flavour.

Table 5: Variable Entered

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	Flavor	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	Clean.Cup	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	Uniformity	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
4	Balance	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
5	Sweetness	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
6	Cupper. Points	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
7	Acidity	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
8	Aroma	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
9	Body	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
10	Aftertaste	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Total.Cup.Points

Table 5 indicates the independent variables that are selected into the stepwise multiple linear regression model, all independent variables are selected as all of them are significant to dependent variable.

Table 6: Model Summary

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.878[a]	.771	.771	1.68357
2	.942[b]	.888	.887	1.17933
3	.963[c]	.927	.927	.94843
4	.977[d]	.955	.955	.74639
5	.986[e]	.971	.971	.59741
6	.991[f]	.981	.981	.48215
7	.994[g]	.989	.989	.37262
8	.997[h]	.994	.994	.27385
9	.999[i]	.998	.998	.16139
10	1.000[j]	1.000	1.000	.01621

a. Predictors: (Constant), Flavor

b. Predictors: (Constant), Flavor, Clean.Cup

c. Predictors: (Constant), Flavor, Clean.Cup, Uniformity

d. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance

e. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness

f. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points

g. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity

h. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma

i. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma, Body

j. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma, Body, Aftertaste

Table 6 is the model summary, all independent variables are selected as candidates for regression equation, so total 10 predictors of the analysis. R square increases gradually as the inclusion of variables, it increases slowly starting from model 6 until model 10 which eventually reaches r square of 1. After inclusion of flavour, clean cup is selected for predictors although it is not the second largest correlation magnitude, instead the stepwise regression model is looking for biggest partial correlation. On the other hand, the standard error of the estimate decreases gradually as the percentage of variance for independent variable increases.

Table 7: ANOVA Analysis

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	12482.097	1	12482.097	4403.759	.000[b]
	Residual	3710.254	1309	2.834		
	Total	16192.351	1310			
2	Regression	14373.155	2	7186.577	5167.142	.000[c]
	Residual	1819.196	1308	1.391		
	Total	16192.351	1310			
3	Regression	15016.681	3	5005.560	5564.715	.000[d]
	Residual	1175.670	1307	.900		
	Total	16192.351	1310			
4	Regression	15464.779	4	3866.195	6939.869	.000[e]
	Residual	727.571	1306	.557		
	Total	16192.351	1310			
5	Regression	15726.605	5	3145.321	8813.049	.000[f]
	Residual	465.746	1305	.357		
	Total	16192.351	1310			
6	Regression	15889.212	6	2648.202	11391.655	.000[g]
	Residual	303.139	1304	.232		
	Total	16192.351	1310			
7	Regression	16011.436	7	2287.348	16474.105	.000[h]
	Residual	180.915	1303	.139		
	Total	16192.351	1310			
8	Regression	16094.710	8	2011.839	26827.138	.000[i]
	Residual	97.640	1302	.075		
	Total	16192.351	1310			
9	Regression	16158.464	9	1795.385	68928.846	.000[j]
	Residual	33.887	1301	.026		
	Total	16192.351	1310			
10	Regression	16192.009	10	1619.201	6160255.314	.000[k]
	Residual	.342	1300	.000		
	Total	16192.351	1310			

a. Dependent Variable: Total.Cup.Points

b. Predictors: (Constant), Flavor

c. Predictors: (Constant), Flavor, Clean.Cup

d. Predictors: (Constant), Flavor, Clean.Cup, Uniformity

e. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance

f. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness

g. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points

h. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity

i. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma

j. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma, Body

k. Predictors: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma, Body, Aftertaste

As shown in Table 7 there are 10 models, the F-values are testing for statistical significance for the model R and all variables are statistically significant.

Table 8: Coefficients

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	24.096	.876		27.521	.000			
	Flavor	7.717	.116	.878	66.361	.000	.878	.878	.878
2	(Constant)	17.152	.642		26.735	.000			
	Flavor	6.423	.089	.731	72.399	.000	.878	.895	.671
	Clean.Cup	1.696	.046	.372	36.874	.000	.661	.714	.342
3	(Constant)	10.667	.570		18.711	.000			
	Flavor	5.915	.074	.673	80.128	.000	.878	.912	.597
	Clean.Cup	1.220	.041	.268	29.707	.000	.661	.635	.221
	Uniformity	1.524	.057	.242	26.747	.000	.658	.595	.199
4	(Constant)	8.214	.457		17.977	.000			
	Flavor	4.037	.088	.459	45.836	.000	.878	.785	.269
	Clean.Cup	1.176	.032	.258	36.362	.000	.661	.709	.213
	Uniformity	1.387	.045	.221	30.765	.000	.658	.648	.180
	Balance	2.440	.086	.282	28.361	.000	.836	.617	.166
5	(Constant)	4.326	.393		11.011	.000			
	Flavor	3.922	.071	.446	55.526	.000	.878	.838	.261
	Clean.Cup	.949	.027	.208	34.849	.000	.661	.694	.164
	Uniformity	1.052	.038	.167	27.556	.000	.658	.606	.129
	Balance	2.401	.069	.277	34.861	.000	.836	.694	.164
	Sweetness	1.069	.039	.161	27.085	.000	.617	.600	.127
6	(Constant)	4.415	.317		13.924	.000			
	Flavor	3.067	.066	.349	46.805	.000	.878	.792	.177
	Clean.Cup	.919	.022	.202	41.775	.000	.661	.757	.158
	Uniformity	1.043	.031	.166	33.857	.000	.658	.684	.128
	Balance	1.978	.058	.229	34.205	.000	.836	.688	.130
	Sweetness	1.096	.032	.166	34.389	.000	.617	.690	.130
	Cupper.Points	1.284	.049	.173	26.448	.000	.802	.591	.100
7	(Constant)	2.708	.252		10.756	.000			
	Flavor	2.267	.057	.258	39.511	.000	.878	.738	.116
	Clean.Cup	.980	.017	.215	57.220	.000	.661	.846	.168
	Uniformity	1.004	.024	.160	42.108	.000	.658	.759	.123
	Balance	1.679	.046	.194	36.645	.000	.836	.712	.107
	Sweetness	1.048	.025	.158	42.447	.000	.617	.762	.124
	Cupper.Points	1.192	.038	.161	31.669	.000	.802	.660	.093
	Acidity	1.450	.049	.157	29.670	.000	.801	.635	.087
8	(Constant)	1.244	.190		6.545	.000			
	Flavor	1.659	.046	.189	36.097	.000	.878	.707	.078
	Clean.Cup	.985	.013	.216	78.251	.000	.661	.908	.168
	Uniformity	.994	.018	.158	56.708	.000	.658	.844	.122
	Balance	1.545	.034	.179	45.565	.000	.836	.784	.098
	Sweetness	1.029	.018	.155	56.688	.000	.617	.844	.122
	Cupper.Points	1.127	.028	.152	40.649	.000	.802	.748	.087
	Acidity	1.294	.036	.140	35.717	.000	.801	.703	.077
	Aroma	1.182	.035	.127	33.323	.000	.797	.678	.072
9	(Constant)	-.192	.116		-1.658	.097			
	Flavor	1.475	.027	.168	53.955	.000	.878	.831	.068
	Clean.Cup	1.014	.007	.222	136.275	.000	.661	.967	.173
	Uniformity	1.006	.010	.160	97.371	.000	.658	.938	.123
	Balance	1.201	.021	.139	56.729	.000	.836	.844	.072
	Sweetness	.987	.011	.149	91.995	.000	.617	.931	.117
	Cupper.Points	1.095	.016	.148	66.967	.000	.802	.880	.085
	Acidity	1.088	.022	.118	50.055	.000	.801	.811	.063
	Aroma	1.066	.021	.115	50.655	.000	.797	.815	.064
	Body	1.075	.022	.110	49.474	.000	.777	.808	.063
10	(Constant)	-.014	.012		-1.216	.224			
	Flavor	.999	.003	.114	327.353	.000	.878	.994	.042
	Clean.Cup	1.001	.001	.220	1336.884	.000	.661	1.000	.170
	Uniformity	1.003	.001	.160	966.120	.000	.658	.999	.123
	Balance	1.002	.002	.116	456.038	.000	.836	.997	.058
	Sweetness	.998	.001	.151	925.705	.000	.617	.999	.118
	Cupper.Points	.997	.002	.135	598.411	.000	.802	.998	.076
	Acidity	.998	.002	.108	453.668	.000	.801	.997	.058
	Aroma	1.000	.002	.108	471.365	.000	.797	.997	.060
	Body	1.001	.002	.102	456.717	.000	.777	.997	.058
	Aftertaste	1.003	.003	.116	357.244	.000	.866	.995	.046

a. Dependent Variable: Total.Cup.Points

As shown in Table 8 all independent variables are significant as p is lower than 0.05. Standardised beta weight is decreasing in magnitude as more predictors are included in the model causes less variance to predict, SPSS is algorithmically trying to find the subsequence best predictor therefore magnitude is decreasing across the models.

Table 9: Excluded Variables

Excluded Variables[a]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	Aroma	.245[b]	11.233	.000	.297	.337
	Aftertaste	.403[b]	14.653	.000	.376	.199
	Acidity	.252[b]	11.546	.000	.304	.333
	Body	.257[b]	13.403	.000	.348	.420
	Balance	.370[b]	18.819	.000	.462	.356
	Uniformity	.359[b]	34.001	.000	.685	.833
	Clean.Cup	.372[b]	36.874	.000	.714	.843
	Sweetness	.345[b]	32.816	.000	.672	.869
	Cupper.Points	.280[b]	13.676	.000	.354	.365
2	Aroma	.231[c]	15.772	.000	.400	.337
	Aftertaste	.347[c]	18.700	.000	.459	.197
	Acidity	.279[c]	19.817	.000	.481	.332
	Body	.270[c]	22.111	.000	.522	.420
	Balance	.315[c]	24.244	.000	.557	.352
	Uniformity	.242[c]	26.747	.000	.595	.676
	Sweetness	.226[c]	24.648	.000	.563	.696
	Cupper.Points	.239[c]	17.218	.000	.430	.363
3	Aroma	.212[d]	18.487	.000	.455	.336
	Aftertaste	.323[d]	22.701	.000	.532	.197
	Acidity	.246[d]	22.192	.000	.523	.329
	Body	.252[d]	27.344	.000	.603	.418
	Balance	.282[d]	28.361	.000	.617	.348
	Sweetness	.166[d]	20.026	.000	.485	.621
	Cupper.Points	.230[d]	21.595	.000	.513	.362
4	Aroma	.166[e]	18.057	.000	.447	.325
	Aftertaste	.209[e]	15.959	.000	.404	.167
	Acidity	.185[e]	19.994	.000	.484	.308
	Body	.176[e]	20.359	.000	.491	.349
	Sweetness	.161[e]	27.085	.000	.600	.620
	Cupper.Points	.166[e]	18.372	.000	.453	.335
5	Aroma	.158[f]	22.659	.000	.532	.325
	Aftertaste	.214[f]	21.736	.000	.516	.167
	Acidity	.171[f]	24.350	.000	.559	.307
	Body	.158[f]	23.621	.000	.547	.346
	Cupper.Points	.173[f]	26.448	.000	.591	.334
6	Aroma	.145[g]	27.208	.000	.602	.323
	Aftertaste	.175[g]	21.559	.000	.513	.162
	Acidity	.157[g]	29.670	.000	.635	.305
	Body	.148[g]	29.558	.000	.634	.345
7	Aroma	.127[h]	33.323	.000	.678	.317
	Aftertaste	.146[h]	23.655	.000	.548	.158
	Body	.122[h]	32.197	.000	.666	.330
8	Aftertaste	.129[i]	31.648	.000	.660	.157
	Body	.110[i]	49.474	.000	.808	.326
9	Aftertaste	.116[j]	357.244	.000	.995	.155

a. Dependent Variable: Total.Cup.Points

b. Predictors in the Model: (Constant), Flavor

c. Predictors in the Model: (Constant), Flavor, Clean.Cup

d. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity

e. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity, Balance

f. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness

g. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points

h. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity

i. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma

j. Predictors in the Model: (Constant), Flavor, Clean.Cup, Uniformity, Balance, Sweetness, Cupper.Points, Acidity, Aroma, Body

As shown in Table 9 as the model number increases the number of variables excluded from the model decreases. Due to the fact that all independent variables are statistically significant, all Ps are less than 0.05, the multiple regression creates another model until there is no more independent variable.

Questions 2: Factor Analysis (Principle Component Analysis)

Table 10: Correlations Table

Correlations

		Aroma	Flavor	Aftertaste	Acidity	Body	Balance	Uniformity	Clean.Cup	Sweetness	Cupper. Points
Aroma	Pearson Correlation	1	.814**	.777**	.725**	.696**	.717**	.366**	.335**	.328**	.693**
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Flavor	Pearson Correlation	.814**	1	.895**	.817**	.762**	.803**	.409**	.396**	.362**	.797**
	Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Aftertaste	Pearson Correlation	.777**	.895**	1	.792**	.761**	.823**	.399**	.386**	.343**	.788**
	Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Acidity	Pearson Correlation	.725**	.817**	.792**	1	.733**	.742**	.373**	.299**	.331**	.700**
	Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Body	Pearson Correlation	.696**	.762**	.761**	.733**	1	.768**	.339**	.287**	.327**	.671**
	Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Balance	Pearson Correlation	.717**	.803**	.823**	.742**	.768**	1	.405**	.374**	.343**	.741**
	Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Uniformity	Pearson Correlation	.366**	.409**	.399**	.373**	.339**	.405**	1	.525**	.538**	.359**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Clean.Cup	Pearson Correlation	.335**	.396**	.386**	.299**	.287**	.374**	.525**	1	.526**	.357**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000		.000	.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Sweetness	Pearson Correlation	.328**	.362**	.343**	.331**	.327**	.343**	.538**	.526**	1	.300**
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000		.000
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311
Cupper.Points	Pearson Correlation	.693**	.797**	.788**	.700**	.671**	.741**	.359**	.357**	.300**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	N	1311	1311	1311	1311	1311	1311	1311	1311	1311	1311

**. Correlation is significant at the 0.01 level (2-tailed).

As stated by (Darlington, 1973) and (Bruin, 2006) the purpose of performing factor analysis is for data reduction or data simplification technique by reducing number of independent variables. Principal component analysis is the most commonly used factor analysis method and it is also by default factor analysis in SPSS. The main idea of factor analysis is to include number of variables and identify the correlation between the variables through a smaller number of factors and components. As a result, a more parsimonious and succinct solution that explains the relationship between these variables. Table 10 shows the correlations of all independent variables, all of them are significance with p-value less than 0.05.

Table 11: PCA Descriptive Statistics

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Aroma	7.5638	.37867	1311
Flavor	7.5181	.39998	1311
Aftertaste	7.3977	.40512	1311
Acidity	7.5331	.38160	1311
Body	7.5177	.35921	1311
Balance	7.5175	.40632	1311
Uniformity	9.8334	.55934	1311
Clean.Cup	9.8331	.77135	1311
Sweetness	9.9033	.53083	1311
Cupper.Points	7.497864	.4746100	1311

Table 12: KMO & Bartlett's Test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.938
Bartlett's Test of Sphericity	Approx. Chi-Square	11111.887
	df	45
	Sig.	.000

As shown in Table 12, the Bartlett's Test of Sphericity is significant as it is less than 0.05 and the chi-square is distributed. This test is to observe whether the variables in correlation matrix as shown in Table 10 are correlated significantly different than zero, df = 45 is the number of variables in the triangle under the Pearson Correlation = 1 diagonal. The test assesses these 45 correlations are significantly differ from zero or not, more precisely it is testing whether the correlation matrix significantly different than an identity matrix, as a results it proves all the variables are significantly correlated.

Table 13: Communalities

Communalities

	Initial	Extraction
Aroma	1.000	.752
Flavor	1.000	.891
Aftertaste	1.000	.875
Acidity	1.000	.779
Body	1.000	.745
Balance	1.000	.802
Uniformity	1.000	.681
Clean.Cup	1.000	.684
Sweetness	1.000	.697
Cupper.Points	1.000	.743

Extraction Method: Principal Component Analysis.

Table 14: Total Variance

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6.216	62.163	62.163	6.216	62.163	62.163	5.368	53.684	53.684
2	1.431	14.313	76.476	1.431	14.313	76.476	2.279	22.792	76.476
3	.501	5.011	81.487						
4	.463	4.631	86.118						
5	.328	3.283	89.401						
6	.311	3.113	92.513						
7	.262	2.620	95.133						
8	.216	2.163	97.296						
9	.173	1.728	99.024						
10	.098	.976	100.000						

Extraction Method: Principal Component Analysis.

Table 14 indicates components for analysis, because of the eigenvalue = 1 is set in the SPSS factor analysis, any component that has eigenvalue less than 1 will not be selected. The table shows only components 1 and 2 are selected, that means 10 components are reduced to 2 components. To assess how well this analysis is performed, percentage of variance is used to decide the performance, the first component achieves 62.163% of variance which is strong while component 2 only achieves 14.313% which is weak. Also, the sum of the eigenvalue is equal to the number of components.

The % of variance is:

$$\text{Percentage of Variance} = (\text{eigenvalue} / \text{no. of component}) * 100\%$$

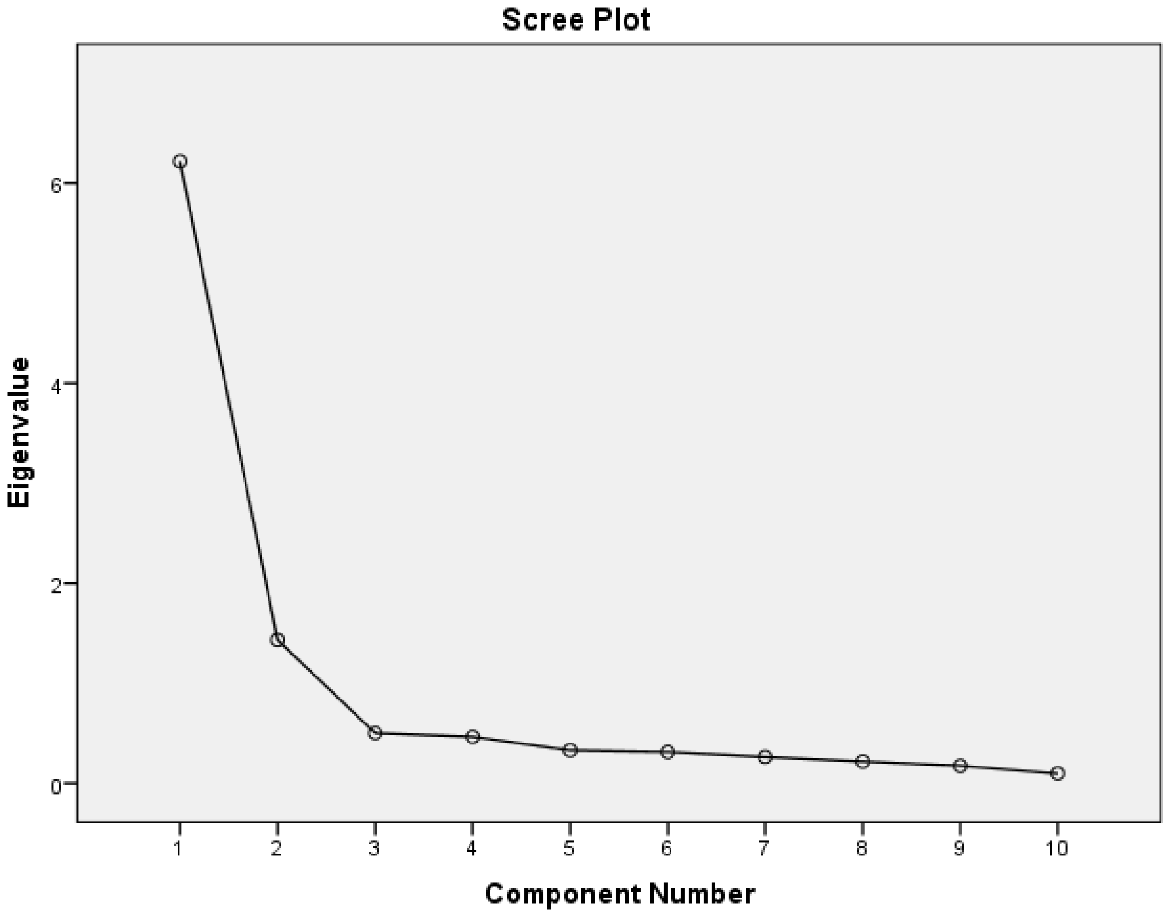


Figure 3: Scree Plot

Figure 3 is a scree plot, component number represents in x axis while eigenvalue represents in y axis. As the number of component increases, the eigenvalue decreases, also the components above the 'scree' are retained.

Table 14 and Figure 3 establish two rules, the first rule is the eigenvalue = 1 rule and the second rule is the components above the 'scree' are retained. As the number of components retained increases, the stated two rules tend to agree less often whereas with fewer components retained, the rules tend to agree more often.

Table 15: Component Matrix

Component Matrix[a]

	Component	
	1	2
Aroma	.852	-.160
Flavor	.932	-.151
Aftertaste	.921	-.162
Acidity	.863	-.186
Body	.840	-.197
Balance	.885	-.134
Uniformity	.550	.615
Clean.Cup	.517	.645
Sweetness	.502	.667
Cupper.Points	.847	-.161

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Communalities

	Initial	Extraction
Aroma	1.000	.752
Flavor	1.000	.891
Aftertaste	1.000	.875
Acidity	1.000	.779
Body	1.000	.745
Balance	1.000	.802
Uniformity	1.000	.681
Clean.Cup	1.000	.684
Sweetness	1.000	.697
Cupper.Points	1.000	.743

Extraction Method: Principal Component Analysis.

Table 15 shows the factor cross-loading, it tells researcher how strong relationship between the item and component, in other words it is the Pearson correlation between the item and components/factors. Flavour loads the highest on the component 1, others are quite good as well except uniformity, clean cup and sweetness. All the loads on component 2 are bad except uniformity, clean cup and sweetness.

Table 15 component matrix can relate to Table 13 the communality, the extraction can be calculated using equation below (Bruin, 2006):

$$\text{Extraction} = (\text{Component 1})^2 + (\text{Component 2})^2$$

In simple regression, R-squared is square of the correlation magnitude and the R-squared indicates the amount of variance that is accounted for. The same goes for factor analysis, for example using variable Aroma the sum of component loadings 0.852^2 and -0.16^2 is 0.752 extraction, that indicates the amount of variance that two components accounted for in Aroma. In conclusion, 75.2% of the variance in Aroma is justified by the retained components, 80.2% of the variance in Balance is justified by the retained components and 68.1% of the variance in Uniformity is justified by the retained components.

Table 16: Rotated Component Matrix

Rotated Component Matrix[a]

	Component	
	1	2
Aroma	.840	.214
Flavor	.908	.256
Aftertaste	.904	.241
Acidity	.861	.195
Body	.845	.175
Balance	.860	.251
Uniformity	.240	.789
Clean.Cup	.197	.803
Sweetness	.174	.816
Cupper.Points	.836	.211

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Table 16 is the estimation of correlation between independent variables and the estimated components.

Table 17: Component Transformation Matrix

Component Transformation Matrix

Component	1	2
1	.907	.421
2	-.421	.907

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Table 17 displays the correlations among the components before and after rotation as stated by (Starkweather & Herrington, 2018).

Factorability:

The factorability can be enhanced by examining correlation of the independent variables and applying factor analysis on independent variables with high correlation.

Factor Cross-Loading:

Factor cross-loading is the happening of a variable in more than one factor, the factor in this case is related to more factors.

Reducing Cross-Loading:

The cross-loading can be mitigated using rotation method, it minimises the eigenvalues and maximises for different factors.

Conclusion & Recommendation

The first part of the research is to perform multiple linear regression with 10 independent variables which will affect the total cupper point dependent variable. According to correlation table, flavour has the highest correlation magnitude. All the independent variables are selected into the stepwise multiple linear regression model. R square increases gradually until model 6 then increases slowly until it reaches 1 at model 10. In a nutshell, all independent variables are significant and all hypotheses are accepted.

The second part performs factor analysis or principal component analysis for data simplification. In the Bartlett's Test of Sphericity, the significance is less than 0.05 and chi-square is distributed. Component with eigenvalue less than one is not selected hence only two components are produced, the factor analysis achieves 62.163% for component 1 and 14.313% for component 2. Scree plot shows that as the number of component increases, the eigenvalue decreases. Flavour has the highest factor cross-loading in component 1 whereas sweetness has the highest factor cross-loading in component 2. Aroma achieves 75.2% of the variance, 80.2% in Balance and 68.1% in Uniformity.

More independent variables are recommended for stepwise multiple linear regression and factor analysis. Variables such as country of origin, altitude, processing method and moisture should be included to achieve more accurate analysis result on arabica coffee quality.

References

Yerkes, T. (2014). *About Us - Coffee Quality Institute*. [Online]. 2014. Coffee Quality Institute. Available from: https://www.coffeeinstitute.org/about-us/. [Accessed: 1 March 2020].

Luca, L. (2020). *Coffee Cupping - Must Know You Can't Miss*. [Online]. 2020. SoloEspresso.net. Available from: https://soloespresso.net/coffee-cupping/. [Accessed: 16 March 2020].

Bruin, J. (2006). Principal Components (PCA) and Exploratory Factor Analysis (EFA) with SPSS. [Online]. 2006. Stats.idre.ucla.edu. Available from: https://stats.idre.ucla.edu /spss/seminars/efa-spss/. [Accessed: 1 March 2020].

Starkwheather, J. & Herrington, R. (2018). DSA SPSS Short Course Module 9 Principal ComponentsAnalysis 1. [Online]. 2018. Bayes.acs.unt.edu. Available from: http://bayes.acs.unt.edu:8083/BayesContent/class/Jon/SPSS_SC/Module9/M9_PCA/SPSS_M9_PCA1.htm. [Accessed: 1 March 2020].

Darlington, R. (1973). Factor Analysis. [Online]. 1973. Node101.psych.cornell.edu. Available from: http://node101.psych.cornell.edu/Darlington/factor.htm. [Accessed: 1 March 2020].

Bruin, J. (2006). A Practical Introduction to Factor Analysis: Exploratory Factor Analysis. [Online]. 2006. Stats.idre.ucla.edu. Available from: https:/ /stats.idre.ucla.edu/ spss/ seminars/ introduction-to-factor-analysis /a-practical-introduction-to-factor-analysis /. [Accessed: 1 March 2020].