# **Exploratory Data Analysis Project Documentation**

**TITLE**: Coffee Quality Data Analysis

Name: Adhwaitha A

DA/DS: Data Analytics (DA)

Batch Number: B4 (June - Online)(M) - DA & DS

Online/Offline: Online

# **Table of Contents**

SI/no	TITLE
1	INTRODUCTION
2	Aim
3	Business Problem / Problem Statement
4	Project Workflow
5	Data Understanding
6	Data Cleaning - Missing Values Imputation, Outliers, Handling Inconsistent Values
7	Obtaining Derived Metrics
8	Filtering Data for Analysis
9	EDA – Univariate Analysis
10	Segmented Univariate Analysis
11	Bivariate Analysis
12	Multivariate Analysis
13	Overall Insights Obtained from Analysis
14	Conclusion

#### 1.Introduction:

The objective of this project is to perform the exploratory data analysis in **Coffee Quality** for the underlying patterns and insights lying therein. It will contain cleaning of data, deriving metrics, and other analyses for the purpose of deriving meaningful conclusions. This dataset contains the concept that encompasses a variety of factors influencing the taste, aroma, and overall enjoyment of the beverage. Understanding what contributes to high-quality coffee involves considering the entire journey from the coffee plant to the cup.

#### 2.Aim:

The main goal in making top-notch coffee is to create a drink that gives people an amazing taste experience. This means the coffee should have unique and appealing flavors, smells, and feels in the mouth. To do this, you need to look at the whole process, from growing the beans to brewing the final cup...

#### 3. Problem Statement:

The coffee industry struggles to deliver top-notch coffee . It needs to satisfy different consumer tastes while maintaining ethical and sustainable practices across its supply chain. This issue has many sides, including farming methods, processing techniques roasting accuracy, and brewing skills. It also involves environmental and social duties. Tackling these hurdles is key to staying competitive keeping customers happy, and ensuring the industry lasts long-term.

# 4. Project Workflow

Data Collection --> Data Understanding --> Data Cleaning --> Data Transformation --> Exploratory Data Analysis --> Conclusions

# 5. Data Understanding

Understanding coffee quality data involves collecting and analyzing information from farming, processing, roasting, and brewing stages. This data is cleaned and integrated to ensure consistency, then analyzed to uncover patterns and significant factors affecting quality.

Predictive models are used to forecast outcomes, and continuous monitoring systems are set up

for real-time data analysis. This approach helps businesses make informed decisions, improve quality control, promote sustainability, and enhance the overall coffee experience for consumers.

# 6. Data Cleaning - Missing Values Imputation, Outliers, Handling Inconsistent Values

- 1. Missing Values Imputation: Method used, Mean imputation for numerical variables, mode imputation for categorical variables.
- 2. Outliers: Method used, Z-score method to identify and remove outliers.
- 3. Inconsistent Values : Method used, Correcting inconsistent entries in categorical variables

#### 7. Derived Metrics

Acquiring derived metrics is a method of figuring out new results from the given data that helps in getting a better understanding. In coffee, this may be the metrics such as cupping scores, defect rates per batch, or the correlation in taste quality between the coffee grown at different altitudes. Through the help of these new metrics, many more trends and patterns can easily be recognized and thus this brings about the ability to make better decisions through enhanced factors. For instance, the quality of green coffee beans is a direct result of the altitude at which it is grown.

# 8. EDA – Univariate Analysis

Univariate analysis for coffee quality involves examining each variable individually to summarize and identify patterns. This includes analyzing cupping scores altitude processing

methods and roast levels. This approach helps understand central tendencies, variability, and distribution shapes, providing a foundational insight into the data.

# 9. Segmented univariate analysis

Segmented univariate analysis examines individual variables within specific data segments to uncover patterns. This could involve comparing cupping scores across regions, altitudes, or processing methods. Analyzing average cupping scores by region can reveal distinct flavor profiles, while examining altitudes within each processing method highlights how altitude influences processing. This approach provides deeper insights into the factors affecting coffee quality.

### 10. Filtering data for analysis

Filtering data for analysis involves selecting specific subsets of data to focus on relevant information and remove noise. For coffee quality, this might mean filtering by harvest year, region, processing method, or roast level to analyze trends and patterns within these specific groups. This process ensures that the analysis is more precise and targeted, leading to more accurate insights and conclusions.

# 11. Bivariate analysis

Bivariate analysis examines the relationship between two variables to identify patterns and correlations. In the context of coffee quality, this could involve analyzing the relationship between altitude and cupping scores or between processing methods and flavor profiles. This analysis helps uncover how different factors interact and influence each other, providing a deeper understanding of what drives coffee quality.

# 12. Multivariate analysis

Multivariate analysis explores the relationships among three or more variables simultaneously to understand complex interactions. In coffee quality, this might involve examining how altitude, processing method, and roast level together influence cupping scores. This comprehensive approach helps identify key factors and their combined effects on coffee quality, offering deeper insights for optimizing production processes.

# #Coffee Quality Database

```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
```

df=pd.read\_csv("C:/Users/user/Downloads/coffeeQuality.csv")
df

	Unnamed: 0	Species	Owner	Country.of.Origin
\				
0	0	Arabica	metad plc	Ethiopia
1	1	Arabica	metad plc	Ethiopia
2	2	Arabica	grounds for health admin	Guatemala
3	3	Arabica	yidnekachew dabessa	Ethiopia
4	4	Arabica	metad plc	Ethiopia
1334	1334	Robusta	luis robles	Ecuador
1335	1335	Robusta	luis robles	Ecuador
1336	1336	Robusta	james moore	United States
1337	1337	Robusta	cafe politico	India
1338	1338	Robusta	cafe politico	Vietnam

1 2 2 5			*** O	h., a+ a a a	T arrada 2	0	
1335 laborat	corv		ĽO.	bustasa	Lavado 3	own	
1336			fazenda	cazengo	NaN	ca	fe
cazengo	)						
1337 NaN				NaN	NaN		
1338				NaN	NaN		
NaN				110.11	21021		
Altitud	40	ICO.Number			Comp	pany	
0	ie	2014/2015	metad agri	cultural	developmet	plc	
1950-22	200					I. – .	
1		2014/2015	metad agri	cultural	developmet	plc	
1950-22	200	NT - NT				NT - NT	1.000
2 1800 m		NaN				NaN	1600 -
3		NaN	yidnekachew del	bessa co:	ffee plantat	cion	
1800-22	200						
4	200	2014/2015	metad agri	cultural	developmet	plc	
1950-22	200						
		• • •				• • •	
1334		NaN			robust	tasa	
NaN		27. 27			1		
1335 40		NaN			robust	tasa	
1336		NaN		global o	pportunity 1	fund	795
meters					-		
	L4-1	118-2014-0087			cafe polit	cico	
NaN 1338		NaN			cafe polit	-ico	
NaN		IVAIV			care poir	0100	
		_					,
0 .	• • •		tegory.Two.Defe	cts 0	Expirate April 3rd,		
1 .		Green Green		1	April 3rd,		
2 .		NaN		0	May 31st,		
3.		Green			March 25th,		
4 .		Green		2	April 3rd,	2016	
1001		Blue-Green		 1 Ja	nuary 18th,	2017	
1005	 	Blue-Green			nuary 18th,		
		NaN			ember 23rd,		
		Green			ugust 25th,		
1338 .	• • •	NaN		9 A	ugust 25th,	2015	
		Ce	rtification.Body	у \			
		D Agricultural	Development plo	C			
1 M	AT 31	D Agricultural	Development plo	C			

```
2
            Specialty Coffee Association
3
      METAD Agricultural Development plc
4
      METAD Agricultural Development plc
            Specialty Coffee Association
1334
1335
            Specialty Coffee Association
            Specialty Coffee Association
1336
            Specialty Coffee Association
1337
            Specialty Coffee Association
1338
                          Certification.A dress \
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
      309fcf77415a3661ae83e027f7e5f05dad786e44
1334
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
     ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335
1336
     ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                          Certification.Contact unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                    m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                    m
1335
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                    m
1336
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                    m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1337
                                                                    m
     352d0cf7f3e9be14dad7df644ad65efc27605ae2
1338
     altitude low meters altitude high meters altitude mean meters
0
                   1950.0
                                         2200.0
                                                               2075.0
1
                  1950.0
                                         2200.0
                                                               2075.0
2
                   1600.0
                                        1800.0
                                                               1700.0
3
                   1800.0
                                         2200.0
                                                               2000.0
4
                   1950.0
                                         2200.0
                                                               2075.0
                                            . . .
. . .
                      . . .
                                                                  . . .
1334
                     NaN
                                           NaN
                                                                  NaN
                                           40.0
1335
                     40.0
                                                                 40.0
1336
                    795.0
                                          795.0
                                                                795.0
1337
                      NaN
                                            NaN
                                                                  NaN
1338
                      NaN
                                            NaN
                                                                  NaN
[1339 rows x 44 columns]
```

```
df = df.iloc[:, 1:]
df
      Species
                                  Owner Country.of.Origin \
0
      Arabica
                              metad plc
                                                Ethiopia
1
      Arabica
                              metad plc
                                                Ethiopia
      Arabica grounds for health admin
2
                                                Guatemala
3
      Arabica
                    yidnekachew dabessa
                                                Ethiopia
4
     Arabica
                                                 Ethiopia
                              metad plc
1334 Robusta
                            luis robles
                                                  Ecuador
1335 Robusta
                            luis robles
                                                  Ecuador
1336 Robusta
                            james moore
                                            United States
1337 Robusta
                                                    India
                          cafe politico
1338 Robusta
                          cafe politico
                                                 Vietnam
                                     Farm.Name Lot.Number
Mill \
0
                                     metad plc
                                                                metad
                                                      NaN
plc
                                     metad plc
                                                      NaN
                                                                metad
1
plc
2 san marcos barrancas "san cristobal cuch
                                                      NaN
       yidnekachew dabessa coffee plantation
                                                      NaN
wolensu
4
                                     metad plc
                                                      NaN
                                                              metad
plc
. . .
                                                      . . .
. . .
1334
                                     robustasa Lavado 1 our own
lab
1335
                                     robustasa Lavado 3 own
laboratory
1336
                               fazenda cazengo
                                                      NaN
                                                          cafe
cazengo
1337
                                           NaN
                                                      NaN
NaN
1338
                                           NaN
                                                      NaN
NaN
             ICO.Number
                                                       Company
Altitude \
                             metad agricultural developmet plc
0
              2014/2015
1950-2200
                             metad agricultural developmet plc
              2014/2015
1950-2200
                                                           NaN 1600 -
                    NaN
1800 m
```

3	NaN yidnekachew debessa coffee plantation	
1800-		
4	2014/2015 metad agricultural developmet plc	
1950-	2200	
	•••	
1334	NaN robustasa	
NaN	_	
1335	NaN robustasa	
40		705
1336	NaN global opportunity fund	795
meter		
1337	14-1118-2014-0087 cafe politico	
NaN	NoN 2250 molitica	
1338	NaN cafe politico	
NaN		
	Region Color	
Cater	ory.Two.Defects \	
0	guji-hambela Green	
0	guji-nambeia Green	
1	guji-hambela Green	
1	guji nambela Oleen	
2	NaN NaN	
0	ivalv	
3	oromia Green	
2	oromia orom	
4	guji-hambela Green	
2	5.1.7	
1334	san juan, playas Blue-Green	
1		
1335	san juan, playas Blue-Green	
0		
1336	kwanza norte province, angola NaN	
6		
1337	NaN Green	
1		
1338	NaN NaN	
9		
	Expiration Certification. Body	Λ
0	April 3rd, 2016 METAD Agricultural Development plc	
1	April 3rd, 2016 METAD Agricultural Development plc	
2	May 31st, 2011 Specialty Coffee Association	
3	March 25th, 2016 METAD Agricultural Development plc	
4	April 3rd, 2016 METAD Agricultural Development plc	
	• • •	

```
January 18th, 2017
1334
                                  Specialty Coffee Association
1335 January 18th, 2017
                                  Specialty Coffee Association
1336 December 23rd, 2015
                                  Specialty Coffee Association
        August 25th, 2015
                                  Specialty Coffee Association
1337
      August 25th, 2015
1338
                                  Specialty Coffee Association
                          Certification.Address \
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
      309fcf77415a3661ae83e027f7e5f05dad786e44
     ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1334
1335 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                          Certification.Contact unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
. . .
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1335
                                                                   m
1336
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1337 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
     altitude low meters altitude high meters altitude mean meters
0
                  1950.0
                                        2200.0
                                                               2075.0
1
                  1950.0
                                        2200.0
                                                               2075.0
2
                  1600.0
                                        1800.0
                                                               1700.0
3
                  1800.0
                                        2200.0
                                                               2000.0
                                                               2075.0
4
                  1950.0
                                        2200.0
                                           . . .
                                                                  . . .
1334
                     NaN
                                           NaN
                                                                  NaN
1335
                    40.0
                                          40.0
                                                                 40.0
                   795.0
                                         795.0
                                                                795.0
1336
1337
                     NaN
                                           NaN
                                                                  NaN
1338
                     NaN
                                           NaN
                                                                  NaN
[1339 rows x 43 columns]
df.info()
```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1339 entries, 0 to 1338
Data columns (total 43 columns):

	COLUMNS (LOCAL 43 COLUM		
#	Column	Non-Null Count	Dtype
0	Species	1339 non-null	object
1	Owner	1332 non-null	object
2	Country.of.Origin	1338 non-null	object
3	Farm.Name	980 non-null	object
4	Lot.Number	276 non-null	object
5	Mill	1021 non-null	object
6	ICO.Number	1180 non-null	object
7	Company	1130 non-null	object
8	Altitude	1113 non-null	object
9	Region	1280 non-null	object
10	Producer	1107 non-null	object
11	Number.of.Bags	1338 non-null	float64
12	Bag.Weight	1339 non-null	object
13	In.Country.Partner	1339 non-null	object
14	Harvest.Year	1292 non-null	object
15	Grading.Date	1339 non-null	object
16	Owner.1	1332 non-null	object
17	Variety	1113 non-null	object
18	Processing.Method	1169 non-null	object
19	Aroma	1339 non-null	float64
20	Flavor	1339 non-null	float64
21	Aftertaste	1339 non-null	float64
22	Acidity	1339 non-null	float64
23	Body	1339 non-null	float64
24	Balance	1339 non-null	float64
25	Uniformity	1339 non-null	float64
26	Clean.Cup	1339 non-null	float64
27	Sweetness	1339 non-null	float64
28	Cupper.Points	1339 non-null	
29	Total.Cup.Points	1339 non-null	float64
30	Moisture	1339 non-null	
31	Category.One.Defects	1339 non-null	int64
32	Quakers	1338 non-null	float64
33	Color	1069 non-null	object
34	Category.Two.Defects	1339 non-null	int64
35	Expiration	1339 non-null	object
36	Certification.Body	1339 non-null	object
37	Certification.Address	1339 non-null	object
38	Certification.Contact	1339 non-null	object
39	unit_of_measurement	1339 non-null	object
40	altitude_low_meters	1109 non-null	float64
41	altitude_high_meters	1109 non-null	float64
42	altitude_mean_meters	1109 non-null	float64

dtypes: float64(17), int64(2), object(24)
memory usage: 449.9+ KB

df.des	cri	be ()						
	Nu	mber.of.Bags	Aro	ma	Flav	or	Aftertas	te
Acidit count		1338.000000	1339.0000	00	1339.0000	00	1339.0000	00
1339.0 mean		159.085202	7.7701	87	7.5204	26	7.4010	83
7.5357 std		173.698167	5.5344	40	0.3984	42	0.4044	63
0.3798 min		0.000000	0.0000	00	0.0000	00	0.0000	00
0.0000 25%		14.000000	7.4200	00	7.3300	00	7.2500	00
7.3300 50%		175.000000	7.5800	00	7.5800	00	7.4200	00
7.5800 75%		275.000000	7.7500	00	7.7500	00	7.5800	00
7.7500 max 8.7500		3200.000000	200.0000	00	8.8300	00	8.6700	00
,		Body	Balance	U	niformity		Clean.Cup	Sweetness
count	13	39.000000 1	339.000000	13	39.000000	133	39.000000	1339.000000
mean		7.517498	7.518013		9.834877		9.835108	9.856692
std		0.370064	0.408943		0.554591		0.763946	0.616102
min		0.000000	0.000000		0.000000		0.000000	0.000000
25%		7.330000	7.330000		10.000000		10.000000	10.000000
50%		7.500000	7.500000		10.000000		10.000000	10.000000
75%		7.670000	7.750000		10.000000		10.000000	10.000000
max		8.580000	8.750000		10.000000		10.000000	10.000000
	Cu	pper.Points	Total.Cup.	Poi	nts Mo	ist	ure	

Cuppe	er.Points	Total.Cup.Points	Moisture
Category.One	e.Defects	\	
count 133	39.000000	1339.000000	1339.000000
1339.000000			
mean	7.503376	82.089851	0.088379
0.479462			
std	0.473464	3.500575	0.048287
2.549683			

```
0.00000
                                0.000000
                                              0.00000
min
0.000000
25%
            7.250000
                               81.080000
                                              0.090000
0.00000
50%
            7.500000
                               82.500000
                                              0.110000
0.00000
                               83.670000
                                              0.120000
75%
            7.750000
0.00000
           10.000000
                               90.580000
                                              0.280000
max
63.000000
                     Category. Two. Defects
                                            altitude low meters
           Quakers
       1338.000000
count
                               1339.000000
                                                     1109.000000
          0.173393
                                  3.556385
                                                     1750.713315
mean
                                                     8669.440545
std
          0.832121
                                  5.312541
          0.00000
                                  0.000000
                                                        1.000000
min
25%
          0.00000
                                  0.00000
                                                     1100.000000
50%
          0.00000
                                  2.000000
                                                     1310.640000
75%
          0.00000
                                  4.000000
                                                     1600.000000
         11.000000
                                 55.000000
                                                   190164.000000
max
       altitude high meters
                               altitude mean meters
                 1109.000000
                                        1109.000000
count
                 1799.347775
                                        1775.030545
mean
std
                 8668.805771
                                        8668.626080
min
                    1.000000
                                           1.000000
                 1100.000000
25%
                                        1100.000000
50%
                 1350.000000
                                        1310.640000
75%
                 1650.000000
                                        1600.000000
              190164.000000
                                      190164.000000
max
un = df["Lot.Number"].unique()
a = len(un)
print(a)
228
df.drop(columns=["Lot.Number"],inplace=True)
      Species
                                    Owner Country.of.Origin \
0
      Arabica
                                metad plc
                                                    Ethiopia
1
      Arabica
                                metad plc
                                                    Ethiopia
2
      Arabica
                grounds for health admin
                                                   Guatemala
3
      Arabica
                                                    Ethiopia
                     yidnekachew dabessa
4
      Arabica
                                                    Ethiopia
                                metad plc
                                                         . . .
1334
      Robusta
                             luis robles
                                                     Ecuador
                              luis robles
1335
      Robusta
                                                     Ecuador
1336
      Robusta
                              james moore
                                               United States
```

1337 Ro	busta	cafe politico	India	
	busta	cafe politico	Vietnam	
0 1 2 sa 3 4  1334 1335 1336 1337 1338		Farm.l metad metad s "san cristobal o ssa coffee plantat metad robus robus fazenda caze	plc metad plc plc metad plc cuch NaN tion wolensu plc metad plc tasa our own lab tasa own laboratory	
	TOO March		0	
Altitude	ICO.Number		Company	
0 1950-220	2014/2015	metad agricul	tural developmet plc	
1	2014/2015	metad agricul	tural developmet plc	
1950-220 2			NaN	1600 -
1800 m	NaN		NaN	1000 -
3 1800-220		yidnekachew debes	sa coffee plantation	
4	2014/2015	metad agricul	tural developmet plc	
1950-220	0			
• • •	• • •		• • •	
1334	NaN		robustasa	
NaN 1335	NaN		robustasa	
40 1336	NaN	αlo	bal opportunity fund	795
meters		,	11 1	
1337 14 NaN	-1118-2014-0087		cafe politico	
1338 NaN	NaN		cafe politico	
		Region		
Producer				
0	gu	ji-hambela		METAD
PLC 1	an a	ji-hambela		METAD
PLC	gu	JI Hambera		HILLIAD
2		NaN		
NaN			lea ele con P. I	
3		oromia Yidne	kachew Dabessa Coffee	9

```
Plantation
4
                        guji-hambela
                                                                     METAD
PLC
. . .
1334
                    san juan, playas
                                               Café Robusta del Ecuador
S.A.
1335
                    san juan, playas
                                               Café Robusta del Ecuador
S.A.
1336
     kwanza norte province, angola
                                                                 Cafe
Cazengo
1337
                                  NaN
NaN
1338
                                  NaN
NaN
                                                        Expiration \
      . . .
                Color Category. Two. Defects
0
                                           0
                                                   April 3rd, 2016
                Green
1
                                           1
                Green
                                                   April 3rd, 2016
2
                 NaN
                                           0
                                                    May 31st, 2011
3
                                           2
                                                  March 25th, 2016
                Green
      . . .
4
                                           2
                                                   April 3rd, 2016
      . . .
                Green
      . . .
. . .
                                         . . .
                                                January 18th, 2017
1334
      ... Blue-Green
                                           1
1335
      ... Blue-Green
                                           0
                                                January 18th, 2017
                                             December 23rd, 2015
1336
                  NaN
                                           6
      . . .
1337
                Green
                                           1
                                                 August 25th, 2015
      . . .
1338
                   NaN
                                           9
                                                 August 25th, 2015
      . . .
                       Certification.Body \
0
      METAD Agricultural Development plc
1
      METAD Agricultural Development plc
2
             Specialty Coffee Association
3
      METAD Agricultural Development plc
4
      METAD Agricultural Development plc
1334
             Specialty Coffee Association
1335
             Specialty Coffee Association
1336
             Specialty Coffee Association
1337
             Specialty Coffee Association
1338
             Specialty Coffee Association
                          Certification.Address \
      309fcf77415a3661ae83e027f7e5f05dad786e44
0
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
      309fcf77415a3661ae83e027f7e5f05dad786e44
1334 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
```

```
1335 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                          Certification.Contact unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
1334 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1335
                                                                   m
1336 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1337
                                                                   m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
     altitude low meters altitude high meters
                                                 altitude mean meters
0
                  1950.0
                                         2200.0
                                                                2075.0
1
                  1950.0
                                         2200.0
                                                                2075.0
2
                  1600.0
                                         1800.0
                                                                1700.0
3
                  1800.0
                                         2200.0
                                                                2000.0
4
                  1950.0
                                         2200.0
                                                                2075.0
                     . . .
                                            . . .
                                                                   . . .
1334
                     NaN
                                            NaN
                                                                   NaN
1335
                    40.0
                                           40.0
                                                                  40.0
                   795.0
                                          795.0
                                                                 795.0
1336
1337
                     NaN
                                            NaN
                                                                   NaN
1338
                     NaN
                                            NaN
                                                                   NaN
[1339 rows x 42 columns]
df.isnull().sum()
                            0
Species
                            7
Owner
                            1
Country.of.Origin
                          359
Farm.Name
Mill
                          318
ICO.Number
                          159
                          209
Company
                          226
Altitude
                          59
Region
Producer
                          232
Number.of.Bags
                            1
Bag.Weight
                            0
                            0
In.Country.Partner
                           47
Harvest.Year
Grading. Date
                            0
```

```
Owner.1
                          7
Variety
                        226
Processing.Method
                        170
                          0
Aroma
Flavor
                          0
Aftertaste
                          0
                          0
Acidity
Body
                          0
                          0
Balance
Uniformity
                          0
                          0
Clean.Cup
Sweetness
                          0
                          0
Cupper.Points
Total.Cup.Points
Moisture
                          0
Category.One.Defects
                          0
Ouakers
                          1
Color
                        270
                        0
Category. Two. Defects
Expiration
                          0
Certification. Body
Certification.Address
Certification.Contact
unit of measurement
                         0
altitude low meters
                        230
altitude high meters
                        230
altitude mean meters
                        230
dtype: int64
c = ['Farm.Name', 'Mill', 'ICO.Number', 'Company', 'Altitude',
'Variety',
                  'Processing.Method', 'Color',
'altitude low meters', 'altitude high meters', 'altitude mean meters']
df.drop(columns=c, inplace=True)
df
      Species
                                Owner Country.of.Origin \
0
     Arabica
                             metad plc Ethiopia
1
    Arabica
                             metad plc
                                              Ethiopia
2
     Arabica grounds for health admin
                                             Guatemala
3
    Arabica
               yidnekachew dabessa
                                              Ethiopia
    Arabica
                            metad plc
                                              Ethiopia
                                  . . .
                                                    . . .
1334 Robusta
                          luis robles
                                               Ecuador
1335 Robusta
                           luis robles
                                               Ecuador
1336 Robusta
                          james moore
                                          United States
1337 Robusta
                        cafe politico
                                                  India
1338 Robusta
                         cafe politico
                                               Vietnam
```

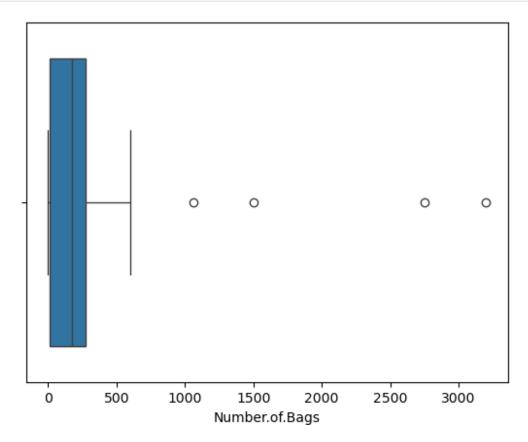
D			Regi	Lon			
Producer \ 0 PLC		guji-	hambe	ela			METAD
1		guji-	hambe	ela			METAD
PLC 2			7	I o NI			
NaN			Γ	laN			
3			oron	nia Yio	dnekachew Dak	pessa Coffee	е
Plantation 4		guji-	·hambe	ela			METAD
PLC		5 5					
1334	sa	n juan,	play	as .	Café I	Robusta del	Ecuador
S.A. 1335	9.3	n duan	nlar	7.0.C	Cofó I	Robusta del	Equador
S.A.	Sa	n juan,	ртау	as	Cale	Robusta dei	ECUACOL
	norte pr	ovince,	ango	la			Cafe
Cazengo 1337			N	NaN			
NaN			•	· carv			
1338 NaN			1	laN			
Number 0	.of.Bags 300.0	_	_	METAD	In Agricultural	n.Country.P	
1	300.0		) kg		Agricultura		
2	5.0		1	MEET	Specialty Co		
3 4	320.0 300.0		) kg ) kg		Agricultural Agricultural	-	-
• • •	• • •				_	_	• • •
1334 1335	1.0		kg kg		Specialty Co		
1336	1.0		. kg		Specialty Co		
1337	1.0		lbs		Specialty Co		
1338	1.0	5	lbs		Specialty Co	oliee Associ	lation
Harvest	.Year	Gr	ading	g.Date	Total.	Cup.Points	Moisture
0	2014	April	4th,	2015		90.58	0.12
		-					
	~ ~ 4 .						0 10
1	2014	April	4th,	2015		89.92	0.12
2	2014 NaN	_		2015		89.92 89.75	0.12
		_	31st,	2010			

				• • •	• • •		• • •
1334	2016	January	19th,	2016		78.75	0.00
1335	2016	January	19th,	2016		78.08	0.00
1336	2014	December	23rd,	2014		77.17	0.00
1337	2013	August	25th,	2014		75.08	0.10
1338	2013	August	25th,	2014		73.75	0.12
0 1 2 3 4  1334 1335 1336 1337 1338	Category.One	.Defects 0 0 0 0 0 0 0 20 63	0 0 0 0 0 0 0 0 0	rs Cat .0 .0 .0 .0 .0 .0 .0	egory.Two	.Defects \     0	
0 1 2 3 4  1334 1335 1336	April 3rd April 3rd April 3rd May 31sd March 25td April 3rd January 18td January 18td December 23rd	d, 2016 Pt, 2011 h, 2016 Pt d, 2016 Pt h, 2017 h, 2017	METAD 2 METAD 2 METAD 2	Agricul Special Agricul Agricul Special Special	tural Deve tural Deve ty Coffee tural Deve tural Deve ty Coffee ty Coffee	ication.Body elopment plc elopment plc Association elopment plc elopment plc Association Association Association	
1337 1338	August 25ti	h, 2015 h, 2015	C h	Special Special	ty Coffee ty Coffee	Association Association	
0 1 2 3 4  1334 1335 1336	309fcf77415a3 309fcf77415a3 36d0d00a37243 309fcf77415a3 309fcf77415a3 ff7c18ad303d4 ff7c18ad303d4 ff7c18ad303d4	3661ae83e0 3661ae83e0 338ba79370 3661ae83e0 3661ae83e0 4b603ac3f8	027f7e5 027f7e5 027f7e5 027f7e5 027f7e5 8cff7e6	5f05dad 3d085f2 5f05dad 5f05dad 511ffc73	786e44 786e44 172daa 786e44 786e44  35e720		

```
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                         Certification.Contact
                                                unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                  m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                  m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                  m
      19fef5a731de2db57d16da10287413f5f99bc2dd
4
                                                                  m
1334
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                  m
1335 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                  m
1336 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                  m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                  m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
[1339 rows x 31 columns]
df.isnull().sum()
```

Species	0
Owner	7
Country.of.Origin	1
Region	59
Producer	232
Number.of.Bags	1
Bag.Weight	0
In.Country.Partner	0
Harvest.Year	47
Grading.Date	0
Owner.1	7
Aroma	0
Flavor	0
Aftertaste	0
Acidity	0
Body	0
Balance	0
Uniformity	0
Clean.Cup	0
Sweetness	0
Cupper.Points	0
Total.Cup.Points	0
Moisture	0
Category.One.Defects	0
Quakers	1
Category.Two.Defects	0
Expiration	0
Certification.Body	0
Certification.Address	0
Certification.Contact	0

```
unit_of_measurement 0
dtype: int64
sns.boxplot(x=df['Number.of.Bags'])
<Axes: xlabel='Number.of.Bags'>
```



```
df['Number.of.Bags'].sort_values(ascending=True).head()
704     0.0
1206     1.0
379     1.0
1188     1.0
444     1.0
Name: Number.of.Bags, dtype: float64

median=df['Number.of.Bags'].median()
median

175.0

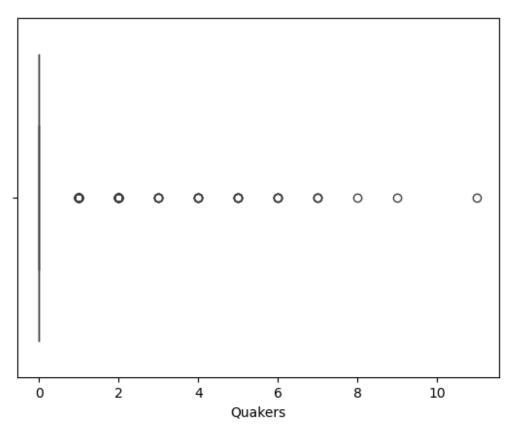
df["Number.of.Bags"].fillna(median, inplace=True)
df
```

0 1 2 3 4  1334 1335 1336 1337 1338	Species Arabica Arabica Arabica Arabica Arabica Consta Robusta Robusta Robusta Robusta Robusta	grounds f yidn	me for healt ekachew me luis luis jame cafe p	tad plc tad plc h admin	Ethiopia Guatemala Ethiopia Ethiopia Ecuador Ecuador United States	
			Reg	ion		
Produ	cer \					
0		g	uji-hamb	ela		METAD
PLC				-		
1		g	uji-hamb	ela		METAD
PLC 2				NaN		
NaN				INAIN		
3			oro	mia Yid	nekachew Dabessa C	offee
Plant	ation					
4		g	uji-hamb	ela		METAD
PLC						
1334		san i	uan, pla	172 C	Café Robusta	del Fauador
S.A.		San	aan, pra	yas	care nobasea	der hedddor
1335		san j	uan, pla	yas	Café Robusta	del Ecuador
S.A.						
1336		orte provi	nce, ang	ola		Cafe
Cazen	go					
1337				NaN		
NaN 1338				NaN		
NaN				IVAIV		
	Number.o	f.Bags Bag	_			ry.Partner \
0		300.0	60 kg		Agricultural Devel	
1 2		300.0 5.0	60 kg 1	ME'I'AD	Agricultural Devel Specialty Coffee A	
3		320.0	60 kg	METAD	Agricultural Devel	
4		300.0	60 kg		Agricultural Devel	
			• • •		,	• • •
1334		1.0	2 kg		Specialty Coffee A	
1335		1.0	2 kg		Specialty Coffee A	
1336		1.0	1 kg		Specialty Coffee A	
1337		1.0	5 lbs		Specialty Coffee A	
1338		1.0	5 lbs		Specialty Coffee A	SSUCTACTOIL

## Harvest.Year   Grading.Date Total.Cup.Points   Moisture								
0 2014 April 4th, 2015 90.58 0.12  1 2014 April 4th, 2015 89.92 0.12  2 NAN May 31st, 2010 89.75 0.00  3 2014 March 26th, 2015 89.00 0.11  4 2014 April 4th, 2015 88.83 0.12	\	Harvest.Year	Gi	rading	.Date	To	tal.Cup.Points	Moisture
NaN   May 31st, 2010     89.75   0.00	0	2014	April	l 4th,	2015		90.58	0.12
3 2014 March 26th, 2015 89.00 0.11 4 2014 April 4th, 2015 88.83 0.12	1	2014	April	l 4th,	2015		89.92	0.12
4 2014 April 4th, 2015 88.83 0.12	2	NaN	May	31st,	2010		89.75	0.00
	3	2014	March	26th,	2015		89.00	0.11
1334 2016 January 19th, 2016 78.75 0.00  1335 2016 January 19th, 2016 78.08 0.00  1336 2014 December 23rd, 2014 77.17 0.00  1337 2013 August 25th, 2014 75.08 0.10  1338 2013 August 25th, 2014 73.75 0.12  Category.One.Defects Quakers Category.Two.Defects \ 0 0 0.0 0 1 2 0 0.0 0 1 0 0.0 1 2 0 0.0 0 3 0 0.0 2 4 0 0.0 2 4 0 0.0 2 4 0 0.0 2 4 0 0.0 2 5 0 0.0 1 6 1337 0 0.0 1 1338 0 0 0.0 1 1338 0 0 0.0 0 1 1339 0 0 0.0 0 1 1339 0 0 0.0 0 1 1339 0 0 0.0 0 0 1331 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1333 0 0 0.0 0 0 1334 0 0 0.0 0 0 1335 0 0 0.0 0 0 14 0 0 0.0 0 0 15 0 0 0 0 16 0 0 0 0 0 17 0 0 0 0 18 0 0 0 0 0 0 19 0 0 0 0 0 10 0 0 0 0 0 10 0 0 0 0 0 11 0 0 0 0	4	2014	April	l 4th,	2015		88.83	0.12
1335   2016   January 19th, 2016     78.08   0.00     1336   2014   December 23rd, 2014     77.17   0.00     1337   2013   August 25th, 2014     75.08   0.10     1338   2013   August 25th, 2014     73.75   0.12     Category.One.Defects   Quakers   Category.Two.Defects   0   0   0   0     1   0   0.0   0   1     2   0   0.0   0   0     3   0   0.0   0   2     4   0   0.0   2     5   0   0.0   0   1     1334   0   0.0   0   1     1335   0   0.0   0   6     1336   0   0.0   0   6     1337   20   0.0   0   1     1338   63   0.0   0   0								
1336   2014   December 23rd, 2014     77.17   0.00     1337   2013   August 25th, 2014     75.08   0.10     1338   2013   August 25th, 2014     73.75   0.12     Category.One.Defects   Quakers   Category.Two.Defects	1334	2016	January	19th,	2016		78.75	0.00
1337 2013 August 25th, 2014 75.08 0.10  1338 2013 August 25th, 2014 73.75 0.12  Category.One.Defects Quakers Category.Two.Defects \ 0	1335	2016	January	19th,	2016		78.08	0.00
Category.One.Defects Quakers Category.Two.Defects \ 0	1336	2014	December	23rd,	2014		77.17	0.00
Category.One.Defects Quakers Category.Two.Defects \ 0	1337	2013	August	25th,	2014		75.08	0.10
Category.One.Defects Quakers Category.Two.Defects \ 0	1338	2013	-				73.75	
0 0.0 0.0 1 1 0 0.0 0.0 1 2 0 0.0 0.0 2 4 0 0.0 2 4 0 0.0 2 5 0 0.0 2 6 0 0.0 2 7 0 0.0 1 8 0 0.0 1 8 0 0.0 1 8 1 0 0.0 1 8 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			. 5	,				
1 0 0.0 1 2 0 0.0 0 3 0 0.0 2 4 0 0.0 2 4 0 0.0 2 1334 0 0.0 0.0 1 1335 0 0.0 0.0 1 1336 0 0.0 0 0 1337 20 0.0 6 1337 20 0.0 1 1338 63 0.0 9  Expiration Certification.Body \ April 3rd, 2016 METAD Agricultural Development plc		Category.One	.Defects			tegory.		
2 0 0.0 0.0 2 4 0 0.00 2								
3 0 0.0 2 4 0 0.0 2 1334 0 0.0 0.0 1 1335 0 0.0 0.0 0 1336 0 0.0 6 1337 20 0.0 6 1338 63 0.0 9  Expiration Certification.Body \ April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc Specialty Coffee Association 1334 January 18th, 2017 Specialty Coffee Association 1335 January 18th, 2017 Specialty Coffee Association 1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association								
4 0 0.0 2	3							
1334 0 0.0 0.0 1 1335 0 0.0 0.0 6 1336 0 0.0 6 1337 20 0.0 1 1338 63 0.0 9  Expiration Certification.Body \ April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc Specialty Coffee Association 1334 January 18th, 2017 Specialty Coffee Association 1335 January 18th, 2017 Specialty Coffee Association 1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association								
1335 0 0.0 0.0 6 1336 0 0.0 6 1337 20 0.0 1 1338 63 0.0 9  Expiration Certification.Body \ 0 April 3rd, 2016 METAD Agricultural Development plc 1 April 3rd, 2016 METAD Agricultural Development plc 2 May 31st, 2011 Specialty Coffee Association 3 March 25th, 2016 METAD Agricultural Development plc 4 April 3rd, 2016 METAD Agricultural Development plc 5 April 3rd, 2016 METAD Agricultural Development plc 6 April 3rd, 2016 METAD Agricultural Development plc 7 April 3rd, 2016 METAD Agricultural Development plc 8 April 3rd, 2016 METAD Agricultural Development plc 9 April 3rd, 2016 METAD Agricultural Development plc								
1336 0 0.0 6 1337 20 0.0 1 1338 63 0.0 9  Expiration Certification.Body \ April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc Specialty Coffee Association 1334 January 18th, 2017 Specialty Coffee Association 1335 January 18th, 2017 Specialty Coffee Association 1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association								
20 0.0 1 1338 63 0.0 9  Expiration Certification.Body \ April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc Specialty Coffee Association 1334 January 18th, 2017 Specialty Coffee Association 1335 January 18th, 2017 Specialty Coffee Association 1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association								
Expiration Certification.Body \  April 3rd, 2016 METAD Agricultural Development plc  April 3rd, 2016 METAD Agricultural Development plc  May 31st, 2011 Specialty Coffee Association  March 25th, 2016 METAD Agricultural Development plc  April 3rd, 2016 METAD Agricultural Development plc  April 3rd, 2016 METAD Agricultural Development plc   1334 January 18th, 2017 Specialty Coffee Association  1335 January 18th, 2017 Specialty Coffee Association  1336 December 23rd, 2015 Specialty Coffee Association  1337 August 25th, 2015 Specialty Coffee Association			-					
Expiration Certification.Body \ April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc  Specialty Coffee Association								
April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc  Specialty Coffee Association	2000				• •			
April 3rd, 2016 METAD Agricultural Development plc May 31st, 2011 Specialty Coffee Association March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc  Specialty Coffee Association January 18th, 2017 Specialty Coffee Association Specialty Coffee Association August 25th, 2015 Specialty Coffee Association Specialty Coffee Association Specialty Coffee Association Specialty Coffee Association		Exp					-	\
2 May 31st, 2011 Specialty Coffee Association 3 March 25th, 2016 METAD Agricultural Development plc 4 April 3rd, 2016 METAD Agricultural Development plc 5 Specialty Coffee Association 6 Specialty Coffee Association 7 Specialty Coffee Association		_			_			
March 25th, 2016 METAD Agricultural Development plc April 3rd, 2016 METAD Agricultural Development plc  Specialty Coffee Association January 18th, 2017 Specialty Coffee Association		_			_			
April 3rd, 2016 METAD Agricultural Development plc  Specialty Coffee Association January 18th, 2017 Specialty Coffee Association		<del>-</del>			-	_		
Specialty Coffee Association 1335 January 18th, 2017 Specialty Coffee Association 1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association			· ·		_			
January 18th, 2017 Specialty Coffee Association January 18th, 2017 Specialty Coffee Association December 23rd, 2015 Specialty Coffee Association August 25th, 2015 Specialty Coffee Association		Whiti 21	u, 2010 F	THU A	agricu.	rtural	peverobment bic	
1335 January 18th, 2017 Specialty Coffee Association 1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association		Januarv 18	th, 2017	Ç	Special	tv Cof	fee Association	
1336 December 23rd, 2015 Specialty Coffee Association 1337 August 25th, 2015 Specialty Coffee Association		_						
		_			_	_		
1338 August 25th, 2015 Specialty Coffee Association					_	_		
	1338	August 25t	h, 2015	C	Specia	lty Cof	fee Association	

```
Certification.Address \
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
      309fcf77415a3661ae83e027f7e5f05dad786e44
1334
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                         Certification.Contact
                                                 unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                    m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
. . .
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                    m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1335
                                                                    m
1336
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                    m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1337
                                                                    m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                    m
[1339 rows x 31 columns]
df.isnull().sum()
Species
                            0
Owner
                            7
                           1
Country.of.Origin
Region
                           59
                          232
Producer
                           0
Number.of.Bags
                           0
Bag.Weight
                           0
In.Country.Partner
                           47
Harvest.Year
Grading.Date
                            0
Owner.1
                            7
Aroma
                            0
Flavor
                            0
Aftertaste
                            0
                            0
Acidity
                            0
Body
                            0
Balance
                            0
Uniformity
Clean.Cup
                            0
```

```
Sweetness
Cupper.Points
                           0
Total.Cup.Points
Moisture
Category.One.Defects
                           0
                           1
Quakers
                           0
Category. Two. Defects
Expiration
Certification.Body
Certification.Address
                           0
Certification.Contact
unit of measurement
                           0
dtype: int64
sns.boxplot(x=df['Quakers'])
<Axes: xlabel='Quakers'>
```



```
df['Quakers'].sort_values(ascending=True).head()

0      0.0
892      0.0
891      0.0
890      0.0
```

```
889
       0.0
Name: Quakers, dtype: float64
mean=df['Quakers'].mean()
mean
0.17339312406576982
df["Quakers"].fillna(mean, inplace=True)
df
      Species
                                   Owner Country.of.Origin \
0
      Arabica
                                                  Ethiopia
                              metad plc
1
      Arabica
                              metad plc
                                                 Ethiopia
2
     Arabica grounds for health admin
                                                Guatemala
3
      Arabica
                    yidnekachew dabessa
                                                 Ethiopia
4
     Arabica
                              metad plc
                                                 Ethiopia
. . .
                                     . . .
1334 Robusta
                            luis robles
                                                   Ecuador
1335 Robusta
                                                  Ecuador
                            luis robles
1336 Robusta
                                             United States
                            james moore
1337 Robusta
                          cafe politico
                                                     India
1338 Robusta
                                                   Vietnam
                          cafe politico
                             Region
Producer
                       guji-hambela
0
                                                                  METAD
PLC
1
                       guji-hambela
                                                                  METAD
PLC
2
                                NaN
NaN
                             oromia Yidnekachew Dabessa Coffee
Plantation
                       guji-hambela
                                                                  METAD
PLC
. . .
1334
                   san juan, playas
                                              Café Robusta del Ecuador
S.A.
1335
                   san juan, playas
                                              Café Robusta del Ecuador
S.A.
1336 kwanza norte province, angola
                                                               Cafe
Cazengo
1337
                                NaN
NaN
1338
                                NaN
NaN
      Number.of.Bags Bag.Weight
                                                  In.Country.Partner \
```

0 1 2 3 4  1334 1335 1336 1337 1338	300.0 300.0 5.0 320.0 300.0  1.0 1.0	60 60 60 2 2 2 1 5 1	kg 1 kg kg kg kg kg kg kg kg	METAD METAD	Agricultural I Agricultural I Specialty Coff Special Spec	Developme fee Assoc Developme Developme fee Assoc fee Assoc fee Assoc fee Assoc	nt plc iation nt plc nt plc iation iation iation iation
	Harvest.Year	Gra	ding	.Date	Total.Cup	.Points	Moisture
0	2014	April	4th,	2015		90.58	0.12
1	2014	April	4th,	2015		89.92	0.12
2	NaN	May 3				89.75	0.00
3	2014	March 2			• • • • • • • • • • • • • • • • • • • •	89.00	0.11
4	2014	April	4th,	2015	• • •	88.83	0.12
	• • •			• • •		• • •	
1334	2016	January 1	9th,	2016		78.75	0.00
1335	2016	January 1	9th,	2016		78.08	0.00
1336	2014 г	ecember 2	3rd,	2014		77.17	0.00
1337	2013	August 2	5th,	2014		75.08	0.10
1338	2013	August 2	5th,	2014		73.75	0.12
0 1 2 3 4  1334 1335 1336 1337 1338	Category.One.I	Defects Q 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 0 0 0	rs Ca .0 .0 .0 .0 .0 .0 .0	tegory.Two.Def	fects \     0     1     0     2     2      1     0     6     1     9	
	Expir	ation			Certificat	cion.Body	\
	1					7	

```
0
          April 3rd, 2016 METAD Agricultural Development plc
1
          April 3rd, 2016 METAD Agricultural Development plc
2
           May 31st, 2011
                                  Specialty Coffee Association
3
         March 25th, 2016 METAD Agricultural Development plc
4
          April 3rd, 2016 METAD Agricultural Development plc
. . .
       January 18th, 2017
                                  Specialty Coffee Association
1334
       January 18th, 2017
                                  Specialty Coffee Association
1335
1336
      December 23rd, 2015
                                  Specialty Coffee Association
1337
        August 25th, 2015
                                  Specialty Coffee Association
        August 25th, 2015
1338
                                  Specialty Coffee Association
                         Certification.Address \
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
1334
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335
1336 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                         Certification.Contact
                                                 unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
      19fef5a731de2db57d16da10287413f5f99bc2dd
4
                                                                   m
1334
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1335
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1336
                                                                   m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1338
[1339 rows x 31 columns]
df.isnull().sum()
                            0
Species
                            7
Owner
Country.of.Origin
                           1
                           59
Region
                          232
Producer
Number.of.Bags
                            0
                            0
Bag.Weight
In.Country.Partner
                            0
```

```
Harvest.Year
                          47
Grading.Date
                           0
Owner.1
                           7
Aroma
                           0
                           0
Flavor
                           0
Aftertaste
                           0
Acidity
Body
                           0
                           0
Balance
Uniformity
                           0
                           0
Clean.Cup
                           0
Sweetness
Cupper.Points
                           0
Total.Cup.Points
                           0
Moisture
                           0
Category.One.Defects
                           0
Quakers
Category. Two. Defects
                           0
Expiration
Certification.Body
                           0
Certification.Address
                           0
Certification.Contact
unit of measurement
                           0
dtype: int64
df['Owner'].sort values(ascending=True).head()
899 acacia hills ltd
         adam ciruli ye
1212
1082
          adam ciruli ye
623
             adam kline
              adam kline
673
Name: Owner, dtype: object
mode=df['Owner'].mode()
mode
0 juan luis alvarado romero
Name: Owner, dtype: object
df["Owner"].fillna(mode[0], inplace=True)
df
      Species
                                 Owner Country.of.Origin \
0
     Arabica
                             metad plc
                                            Ethiopia
1
     Arabica
                             metad plc
                                               Ethiopia
2
     Arabica grounds for health admin
                                               Guatemala
3
    Arabica
                  yidnekachew dabessa
                                               Ethiopia
4
     Arabica
                             metad plc
                                               Ethiopia
                                    . . .
1334 Robusta
                            luis robles
                                                Ecuador
```

1336 1337	Robusta Robusta Robusta Robusta			Ecuador United States India Vietnam	
Produc	er \	Regi	on		
0	,	guji-hambe	ela		METAD
PLC 1		guji-hambe	ela		METAD
PLC					
2 NaN		N	IaN		
3		orom	nia Yidneka	chew Dabessa Co	offee
Planta 4	tion	guji-hambe	ıl a		METAD
PLC		guji-nambe	:Ia		METAD
1334		san juan, play	as	Café Robusta	del Ecuador
S.A. 1335		aan duan nlar	7.2.G	Café Robusta	dol Equador
S.A.		san juan, play	as	cale Robusta	der Ecuador
		province, ango	la		Cafe
Cazeng	0	<b>N</b>	laN		
NaN		1	ian		
1338		Ν	laN		
NaN					
0 1 2 3 4	Number.of.Bag 300. 300. 5. 320. 300.	0 60 kg 0 60 kg 0 1 0 60 kg	METAD Agri Spec METAD Agri	In.Counts Cultural Develo Cultural Develo Cialty Coffee As Cultural Develo Cultural Develo	opment plc ssociation opment plc
1334 1335 1336 1337 1338	1. 1. 1. 1.	0 2 kg 0 2 kg 0 1 kg 0 5 lbs	Spec Spec Spec	cialty Coffee Ascialty Coffee	ssociation ssociation ssociation
	Harvest.Year	Grading	g.Date	Total.Cup.Poir	nts Moisture
0	2014	April 4th,	2015	90	.58 0.12
J		_			
1	2014	April 4th,	2015	89	.92 0.12
2	NaN	May 31st,	2010	89	.75 0.00

3	2014	March 26	5th, 2015		89.00	0.11
4	2014	April 4	1th, 2015		88.83	0.12
1334	2016	January 19	9th, 2016	• • •	78.75	0.00
1335	2016	January 19	9th, 2016		78.08	0.00
1336	2014	December 23	Brd, 2014		77.17	0.00
1337	2013	August 25	5th, 2014		75.08	0.10
1338	2013	August 25	5th, 2014		73.75	0.12
0 1 2	Category.One.	Defects Qu 0 0 0	0.0 0.0 0.0 0.0	tegory.Tv	wo.Defects \ 0 1 0	
3 4		0 0	0.0		2 2 •••	
1334 1335 1336 1337 1338		0 0 0 20 63	0.0 0.0 0.0 0.0		1 0 6 1 9	
0 1 2 3 4	April 3rd April 3rd May 31st March 25th	, 2016 MET , 2011 , 2016 MET	TAD Agricu Specia TAD Agricu	ltural De ltural De lty Coffe ltural De	ification.Body evelopment plc evelopment plc ee Association evelopment plc evelopment plc	
1334 1335 1336 1337 1338	January 18th January 18th December 23rd August 25th August 25th	, 2017 , 2015 , 2015	Specia Specia Specia	lty Coffe lty Coffe lty Coffe	ee Association ee Association ee Association ee Association ee Association ee Association	
0 1 2 3 4	309fcf77415a3 309fcf77415a3 36d0d00a37243 309fcf77415a3 309fcf77415a3	661ae83e027 661ae83e027 38ba7937c52 661ae83e027	f7e5f05da 2a378d085f 7f7e5f05da	d786e44 d786e44 2172daa d786e44		

```
ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1334
1335 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                         Certification.Contact
                                                 unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
      19fef5a731de2db57d16da10287413f5f99bc2dd
4
                                                                   m
. . .
                                                                  . . .
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                   m
1335
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1336
                                                                   m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1338
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
[1339 rows x 31 columns]
df.isnull().sum()
Species
                            0
                            0
Owner
Country.of.Origin
                           1
                           59
Region
Producer
                         232
Number.of.Bags
                           0
                           0
Bag.Weight
In.Country.Partner
                           0
                          47
Harvest.Year
Grading.Date
                           0
Owner.1
                           7
                           0
Aroma
Flavor
                           0
Aftertaste
                           0
                           0
Acidity
                           0
Body
Balance
                           0
Uniformity
                           0
```

0

0

0

0

0

0

Clean.Cup

Sweetness

Moisture

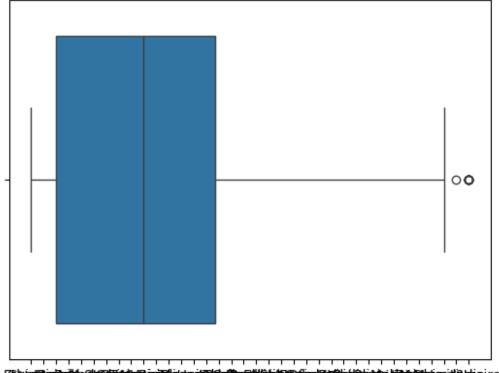
Quakers

Cupper.Points

Total.Cup.Points

Category.One.Defects

Category. Two. Defects



**Ethings in the Act of Control of the Act of Control o** 

```
df['Country.of.Origin'].sort_values(ascending=True).head()

1075     Brazil
730     Brazil
483     Brazil
987     Brazil
Name: Country.of.Origin, dtype: object

mode=df['Country.of.Origin'].mode()
mode

0     Mexico
Name: Country.of.Origin, dtype: object
```

```
df["Country.of.Origin"].fillna(mode[0], inplace=True)
df
      Species
                                   Owner Country.of.Origin \
0
      Arabica
                               metad plc
                                                   Ethiopia
1
      Arabica
                               metad plc
                                                  Ethiopia
2
      Arabica grounds for health admin
                                                 Guatemala
3
      Arabica
                     yidnekachew dabessa
                                                   Ethiopia
4
      Arabica
                               metad plc
                                                   Ethiopia
                                                        . . .
. . .
                                      . . .
1334 Robusta
                             luis robles
                                                    Ecuador
1335 Robusta
                             luis robles
                                                    Ecuador
1336 Robusta
                                              United States
                             james moore
1337 Robusta
                           cafe politico
                                                      India
                                                    Vietnam
1338 Robusta
                           cafe politico
                              Region
Producer \
0
                        guji-hambela
                                                                    METAD
PLC
1
                        guji-hambela
                                                                    METAD
PLC
2
                                 NaN
NaN
                              oromia Yidnekachew Dabessa Coffee
3
Plantation
                        guji-hambela
                                                                    METAD
PLC
. . .
. . .
                                               Café Robusta del Ecuador
1334
                    san juan, playas
S.A.
1335
                    san juan, playas
                                               Café Robusta del Ecuador
S.A.
     kwanza norte province, angola
                                                                 Cafe
1336
Cazengo
1337
                                 NaN
NaN
1338
                                 NaN
NaN
      Number.of.Bags Bag.Weight
                                                  In.Country.Partner \
0
               300.0
                           60 kg
                                  METAD Agricultural Development plc
1
                           60 kg METAD Agricultural Development plc
               300.0
2
                                         Specialty Coffee Association
                  5.0
                              1
3
                           60 kg METAD Agricultural Development plc
               320.0
4
               300.0
                           60 kg
                                  METAD Agricultural Development plc
                  . . .
                             . . .
. . .
                                         Specialty Coffee Association
1334
                 1.0
                            2 kg
1335
                 1.0
                                         Specialty Coffee Association
                            2 kg
```

1336 1337 1338	1.0 1.0 1.0	1 kg 5 lbs 5 lbs		Specialty (	Coffee Associa Coffee Associa Coffee Associa	ation
	Harvest.Year	Grading	.Date	Total	.Cup.Points N	Moisture
0	2014	April 4th,	2015		90.58	0.12
1	2014	April 4th,	2015		89.92	0.12
2	NaN	May 31st,	2010		89.75	0.00
3	2014 M	March 26th,	2015		89.00	0.11
4	2014	April 4th,	2015		88.83	0.12
					• • •	
1334	2016 Jar	uary 19th,	2016		78.75	0.00
1335	2016 Jan	uary 19th,	2016		78.08	0.00
1336	2014 Dece	ember 23rd,	2014		77.17	0.00
1337	2013 Au	igust 25th,	2014		75.08	0.10
1338	2013 Au	igust 25th,	2014		73.75	0.12
0 1 2 3 4  1334 1335 1336 1337 1338	Category.One.Defe	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	rs Ca .0 .0 .0 .0 .0 .0 .0	ategory.Two	.Defects \     0	
0 1 2 3 4  1334 1335	Expirati April 3rd, 20 April 3rd, 20 May 31st, 20 March 25th, 20 April 3rd, 20 January 18th, 20 January 18th, 20	016 METAD 2 016 METAD 2 011 S 016 METAD 2 016 METAD 2	Agricu Specia Agricu Agricu Specia	altural Deve alty Coffee altural Deve altural Deve alty Coffee	elopment plc Association elopment plc elopment plc elopment plc elopment plc elopment plc Association Association	

```
1336 December 23rd, 2015
                                Specialty Coffee Association
        August 25th, 2015
1337
                                 Specialty Coffee Association
1338 August 25th, 2015 Specialty Coffee Association
                         Certification.Address \
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
. . .
1334 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                                                unit of measurement
                         Certification.Contact
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                   m
1335
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1336
                                                                   m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
[1339 rows x 31 columns]
df.isnull().sum()
Species
                           0
                           0
Owner
                           0
Country.of.Origin
                          59
Region
Producer
                         232
Number.of.Bags
                           0
                           0
Bag.Weight
                           0
In.Country.Partner
Harvest.Year
                          47
Grading.Date
                           0
Owner.1
                           7
                           0
Aroma
                           0
Flavor
                           0
Aftertaste
                           0
Acidity
Body
                           0
```

```
0
Balance
Uniformity
                           0
                           0
Clean.Cup
Sweetness
                           0
Cupper.Points
                           0
Total.Cup.Points
                           0
Moisture
Category.One.Defects
                           0
Quakers
                           0
Category. Two. Defects
Expiration
                           0
Certification.Body
Certification.Address
Certification.Contact
                           0
unit of measurement
                           0
dtype: int64
df['Region'].sort values(ascending=True).head()
628
       52 narino (exact location: mattituy; municipal...
57
                                              acatenango
129
                                              acatenango
139
                                              acatenango
638
                                                    aceh
Name: Region, dtype: object
mode=df['Region'].mode()
mode
0 huila
Name: Region, dtype: object
df["Region"].fillna(mode[0], inplace=True)
df
                                  Owner Country.of.Origin \
      Species
0
      Arabica
                                             Ethiopia
                              metad plc
1
     Arabica
                                                Ethiopia
                              metad plc
2
     Arabica grounds for health admin
                                                Guatemala
3
     Arabica
                    yidnekachew dabessa
                                                Ethiopia
    Arabica
                              metad plc
                                                Ethiopia
          . . .
. . .
                                                      . . .
1334 Robusta
                            luis robles
                                                 Ecuador
1335 Robusta
                            luis robles
                                                  Ecuador
1336 Robusta
                            james moore
                                            United States
1337 Robusta
                          cafe politico
                                                    India
1338 Robusta
                          cafe politico
                                                 Vietnam
                             Region
Producer \
                       guji-hambela
                                                                 METAD
```

PLC 1 PLC	guji-hambela		METAD
2 NaN	huila		
3	oromia Y	Yidnekachew Dabessa Coff	ee
Plantation 4 PLC	guji-hambela		METAD
•••			
1334 S.A.	san juan, playas	Café Robusta del	Ecuador
1335	san juan, playas	Café Robusta del	Ecuador
S.A. 1336 kwanza norte Cazengo	province, angola		Cafe
1337	huila		
NaN 1338	huila		
NaN			
0 300 1 300 2 5 3 320 4 300 	.0 60 kg METAI .0 1 .0 60 kg METAI	In.Country.P D Agricultural Developme D Agricultural Developme Specialty Coffee Assoc D Agricultural Developme D Agricultural Developme Specialty Coffee Assoc	nt plc nt plc iation nt plc nt plc iation iation iation iation
Harvest.Year	Grading.Date	e Total.Cup.Points	Moisture
0 2014	April 4th, 2015	5 90.58	0.12
1 2014	April 4th, 2015	5 89.92	0.12
2 NaN	May 31st, 2010	99.75	0.00
3 2014	March 26th, 2015	5 89.00	0.11
4 2014	April 4th, 2015	5 88.83	0.12
1334 2016	January 19th, 2016	6 78.75	0.00

```
1335
             2016
                     January 19th, 2016
                                                          78.08
                                                                      0.00
1336
             2014
                    December 23rd, 2014
                                                          77.17
                                                                      0.00
1337
                                                          75.08
                                                                      0.10
             2013
                      August 25th, 2014
1338
             2013
                      August 25th, 2014
                                                          73.75
                                                                      0.12
      Category.One.Defects
                             Quakers
                                      Category. Two. Defects
0
                          0
                                  0.0
                                                           \Omega
1
                          0
                                  0.0
                                                           1
2
                          0
                                  0.0
                                                           0
3
                                  0.0
                                                           2
                          0
4
                                  0.0
                                                           2
                          0
                                  . . .
                        . . .
                                                          . . .
1334
                          0
                                  0.0
                                                           1
1335
                          0
                                  0.0
                                                           0
1336
                          0
                                  0.0
                                                           6
                                                           1
1337
                                  0.0
                         20
1338
                         63
                                  0.0
                                                           9
                Expiration
                                             Certification.Body \
0
          April 3rd, 2016 METAD Agricultural Development plc
1
          April 3rd, 2016 METAD Agricultural Development plc
2
           May 31st, 2011
                                   Specialty Coffee Association
3
         March 25th, 2016 METAD Agricultural Development plc
          April 3rd, 2016 METAD Agricultural Development plc
4
1334
       January 18th, 2017
                                   Specialty Coffee Association
       January 18th, 2017
                                   Specialty Coffee Association
1335
1336
      December 23rd, 2015
                                   Specialty Coffee Association
1337
        August 25th, 2015
                                   Specialty Coffee Association
1338
       August 25th, 2015
                                  Specialty Coffee Association
                          Certification.Address \
      309fcf77415a3661ae83e027f7e5f05dad786e44
0
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
      309fcf77415a3661ae83e027f7e5f05dad786e44
. . .
1334 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                          Certification.Contact unit of measurement
```

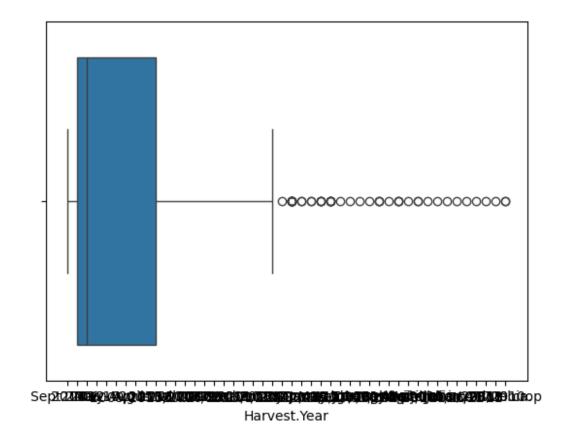
```
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
. . .
                                                                  . . .
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                   m
1335
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1336
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
[1339 rows x 31 columns]
df.isnull().sum()
                           0
Species
                           0
Owner
Country.of.Origin
                           0
Region
                           0
Producer
                         232
Number.of.Bags
                           0
                           0
Bag.Weight
                           0
In.Country.Partner
Harvest.Year
                          47
Grading.Date
                           0
                           7
Owner.1
Aroma
                           0
                           0
Flavor
                           0
Aftertaste
Acidity
                           0
                           0
Body
Balance
                           0
Uniformity
                           0
Clean.Cup
                           0
                           0
Sweetness
                           0
Cupper.Points
Total.Cup.Points
                           0
                           0
Moisture
Category.One.Defects
                           0
                           0
Quakers
Category. Two. Defects
                           0
Expiration
Certification.Body
                           0
                           0
Certification.Address
Certification.Contact
                           0
unit of measurement
dtype: int64
df['Producer'].sort values(ascending=True).head()
```

```
624
537
255
1269
1061
       ADONIS JOSE ORTEZ BELTRAN
Name: Producer, dtype: object
mode=df['Producer'].mode()
mode
0 La Plata
Name: Producer, dtype: object
df["Producer"].fillna(mode[0], inplace=True)
df
                                  Owner Country.of.Origin \
      Species
0
      Arabica
                              metad plc
                                                 Ethiopia
1
     Arabica
                              metad plc
                                                 Ethiopia
2
     Arabica grounds for health admin
                                                Guatemala
3
     Arabica
                    yidnekachew dabessa
                                                 Ethiopia
                              metad plc
4
     Arabica
                                                 Ethiopia
. . .
                                     . . .
1334 Robusta
                            luis robles
                                                  Ecuador
1335 Robusta
                            luis robles
                                                  Ecuador
1336 Robusta
                                            United States
                            james moore
1337 Robusta
                          cafe politico
                                                     India
1338 Robusta
                          cafe politico
                                                  Vietnam
                             Region
Producer \
0
                       quji-hambela
                                                                  METAD
PLC
                       guji-hambela
                                                                  METAD
PLC
2
                              huila
                                                                   La
Plata
                             oromia Yidnekachew Dabessa Coffee
Plantation
                       guji-hambela
                                                                  METAD
PLC
. . .
. . .
                   san juan, playas
1334
                                             Café Robusta del Ecuador
S.A.
1335
                                             Café Robusta del Ecuador
                   san juan, playas
S.A.
1336 kwanza norte province, angola
                                                               Cafe
Cazengo
1337
                              huila
                                                                   La
```

Plata			hui	.la			La
Plata							
0 1 2 3 4	320 300	.0 60 .0 60 .0 60	kg kg l kg kg	METAD METAD	In.C Agricultural De Agricultural De Specialty Coffe Agricultural De Agricultural De	evelopment ee Associa evelopment	plc plc tion plc
1334 1335 1336 1337 1338	1 1 1 1 1	.0 2 .0 1	kg kg kg kg lbs		Specialty Coffe Specialty Coffe Specialty Coffe Specialty Coffe Specialty Coffe	ee Associa ee Associa ee Associa	ition ition ition
	Harvest.Year	Gr	ading	.Date	Total.Cup	.Points M	loisture
0	2014	April	4th,	2015		90.58	0.12
1	2014	April	4th,	2015		89.92	0.12
2	NaN	May	31st,	2010		89.75	0.00
3	2014	March	26th,	2015	• • •	89.00	0.11
4	2014	April	4th,	2015		88.83	0.12
1334	2016	January	19th,	2016		78.75	0.00
1335	2016	January	19th,	2016	• • •	78.08	0.00
1336	2014	December	23rd,	2014		77.17	0.00
1337	2013	August	25th,	2014		75.08	0.10
1338	2013	August	25th,	2014		73.75	0.12
0 1 2 3 4  1334 1335	Category.One	.Defects 0 0 0 0 0 0 0	000000000000000000000000000000000000000	ers Co	ategory.Two.Def	ects \     0     1     0     2     2      1     0	

```
1336
                          0
                                 0.0
                                                          6
1337
                                                          1
                         20
                                 0.0
1338
                         63
                                 0.0
                                                          9
               Expiration
                                             Certification.Body \
0
          April 3rd, 2016
                            METAD Agricultural Development plc
1
          April 3rd, 2016
                            METAD Agricultural Development plc
2
           May 31st, 2011
                                  Specialty Coffee Association
3
         March 25th, 2016
                            METAD Agricultural Development plc
4
          April 3rd, 2016
                            METAD Agricultural Development plc
. . .
1334
                                  Specialty Coffee Association
       January 18th, 2017
1335
       January 18th, 2017
                                  Specialty Coffee Association
      December 23rd, 2015
1336
                                  Specialty Coffee Association
1337
        August 25th, 2015
                                  Specialty Coffee Association
1338
        August 25th, 2015
                                  Specialty Coffee Association
                          Certification.Address \
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
4
      309fcf77415a3661ae83e027f7e5f05dad786e44
. . .
1334
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                          Certification.Contact
                                                  unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                    m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                    m
      19fef5a731de2db57d16da10287413f5f99bc2dd
4
                                                                    m
1334
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                     m
1335
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                     m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1336
                                                                     m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                    m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
[1339 rows x 31 columns]
df.isnull().sum()
Species
                           0
                           0
Owner
                           0
Country.of.Origin
```

```
Region
                          0
Producer
                          0
Number.of.Bags
                          0
Bag.Weight
                          0
                         0
In.Country.Partner
                         47
Harvest.Year
Grading.Date
                         0
Owner.1
                         7
Aroma
                          0
Flavor
                          0
                          0
Aftertaste
Acidity
                          0
Body
                          0
Balance
                          0
Uniformity
                          0
Clean.Cup
                          0
Sweetness
                          0
                          0
Cupper.Points
                          0
Total.Cup.Points
Moisture
                          0
Category.One.Defects
                          0
Quakers
                          0
Category. Two. Defects
                          0
                          0
Expiration
Certification.Body
                         0
Certification.Address
                         0
Certification.Contact
                         0
unit of measurement 0
dtype: int64
sns.boxplot(x=df['Harvest.Year'])
<Axes: xlabel='Harvest.Year'>
```



df['Harvest.Year'].sort values(ascending=True).head() 08/09 crop 1261 1191 08/09 crop 1T/2011 437 857 1t/2011 2009 - 2010 Name: Harvest. Year, dtype: object mode=df['Harvest.Year'].mode() mode 2012 Name: Harvest.Year, dtype: object df["Harvest.Year"].fillna(mode[0], inplace=True) df Species Owner Country.of.Origin \ Arabica 0 metad plc Ethiopia Ethiopia 1 Arabica metad plc 2 Arabica grounds for health admin Guatemala 3 Arabica yidnekachew dabessa Ethiopia 4 Arabica metad plc Ethiopia . . . . . .

1337	Robusta Robusta Robusta Robusta		lu: jar cafe	is robled is robled mes moore politice politice	s e United o	Ecuador Ecuador d States India Vietnam	
Produ	cer \		Re	egion			
0 PLC	CCI	g.	uji-han	nbela			METAD
1 PLC		g.	uji-han	nbela			METAD
2			ľ	nuila			La
Plata 3			01	romia Yid	dnekachew Da	abessa Coffee	Э
Planta	ation	g.	uji-ham	mbela			METAD
PLC							
1334		san j	uan, pl	ayas	Café	Robusta del	Ecuador
S.A. 1335		san j	uan, pl	ayas	Café	Robusta del	Ecuador
S.A. 1336	kwanza noi	rte provi	nce, an	igola			Cafe
Cazeno 1337	do .		ŀ	nuila			La
Plata			1	Idild			Ц
1338			ŀ	nuila			La
Plata							
	Number.of	Rage Rag	Weight	_	-	In.Country.Pa	artner \
0		.вауз вау 300.0	60 kg			al Developmen	
1		300.0	60 kg		_	al Developmen	_
2	`	5.0	1		_	Coffee Assoc	_
3	:	320.0	60 kg			al Developmen	
4		300.0	60 kg	f	_	al Developmen	_
			• •		3	•	• • •
1334		1.0	2 kg		specialty (	Coffee Assoc	iation
1335		1.0	2 kg			Coffee Assoc	
1336		1.0	1 kc	f		Coffee Assoc	
1337		1.0	5 lbs			Coffee Assoc	
1338		1.0	5 lbs			Coffee Assoc	
	Harvest.Yea	ar	Gradi	ing.Date	Total.	Cup.Points	Moisture
\	0.0	1 /		-b 0015		00 50	0 10
0	201	14 A	pril 4t	th, 2015		90.58	0.12
1	201	14 A	pril 4t	ch, 2015		89.92	0.12

2	2012	Мау	31st,	2010		89.	75	0.00
3	2014	March	26th,	2015		89.	00	0.11
4	2014	Apri	l 4th,	2015		88.	83	0.12
				• • •			• •	
1334	2016	January	19th,	2016		78.	75	0.00
1335	2016	January	19th,	2016		78.	08	0.00
1336	2014	December	23rd,	2014		77.	17	0.00
1337	2013	August	25th,	2014		75.	08	0.10
1338	2013	August	25th,	2014		73.	75	0.12
0 1 2 3 4  1334 1335 1336 1337 1338	Category.One	.Defects 0 0 0 0 0 0 0 20 63	0 0 0 0 0 0 0 0 0 0 0 0	rs Ca .0 .0 .0 .0 .0 .0 .0 .0 .0 .0 .0 .0 .0	tegory	.Two.Defects 0 1 0 2 2 1 0 6 1 9		
0 1 2 3 4  1334 1335 1336 1337 1338	April 3rd April 3rd April 3rd May 31st March 25th April 3rd January 18th January 18th January 25th August 25th August 25th 309fcf77415at 36d0d00a3724t 309fcf77415at	d, 2016 M t, 2011 n, 2016 M d, 2016 M n, 2017 n, 2017 d, 2015 n, 2015 n, 2015 a, 2015 a, 3661ae83e(338ba79376)	METAD METAD METAD rtific 027f7e 027f7e c52a37	Agricu Specia Agricu Agricu Specia Specia Specia Specia Specia ation. 5f05da 8d085f	ltural lty Co: ltural lty Co: lty Co: lty Co: lty Co: lty Co: d786e4 2172da	4 4 a	plc plc ion plc plc ion ion ion	

```
4
      309fcf77415a3661ae83e027f7e5f05dad786e44
. . .
1334 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1335
1336 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1338 ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                         Certification.Contact
                                                 unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                   m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
4
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                   m
1334
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1335
                                                                   m
1336
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1337
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
1338 352d0cf7f3e9be14dad7df644ad65efc27605ae2
                                                                   m
[1339 rows x 31 columns]
df.isnull().sum()
                         0
Species
                         0
Owner
Country.of.Origin
                         0
Region
                         0
Producer
                         0
Number.of.Bags
Bag.Weight
                         0
                         0
In.Country.Partner
Harvest.Year
                         0
Grading.Date
                         0
                         7
Owner.1
Aroma
                         0
Flavor
                         0
                         0
Aftertaste
                         0
Acidity
Body
                         0
Balance
                         0
Uniformity
                         0
```

0

0

0

0

Clean.Cup

Sweetness

Moisture

Ouakers

Cupper.Points
Total.Cup.Points

Category.One.Defects

```
Category. Two. Defects
                        0
Expiration
                        0
Certification.Body
                        0
Certification.Address
                        0
Certification.Contact
                        0
unit of measurement
dtype: int64
df['Owner.1'].sort values(ascending=True).head()
       ADRIANA TORRES RICO QUEVEDO
1176
402
49
                              AFCA
593
                              AFCA
939
                              AFCA
Name: Owner.1, dtype: object
mode=df['Owner.1'].mode()
mode
0
    Juan Luis Alvarado Romero
Name: Owner.1, dtype: object
df["Owner.1"].fillna(mode[0], inplace=True)
      Species
                                 Owner Country.of.Origin \
0
     Arabica
                             metad plc
                                           Ethiopia
                                               Ethiopia
1
     Arabica
                             metad plc
2
     Arabica grounds for health admin
                                               Guatemala
3
    Arabica
                   yidnekachew dabessa
                                               Ethiopia
4
    Arabica
                             metad plc
                                                Ethiopia
                                    . . .
                                                     . . .
1334 Robusta
                           luis robles
                                                 Ecuador
1335 Robusta
                           luis robles
                                                 Ecuador
1336 Robusta
                           james moore
                                           United States
1337 Robusta
                          cafe politico
                                                   India
1338 Robusta
                          cafe politico
                                                 Vietnam
```

1334		san juan, pla	ıvas	Café R	obusta del	Ecuador
S.A.						
1335 S.A.		san juan, pla	ıyas	Café R	obusta del	Ecuador
1336		province, ang	ola			Cafe
Cazen 1337	go	hu	ila			La
Plata 1338		le e e				T o
Plata		nu	iila			La
	Manula a a f. Da	na Dan Maink		T	Q	
0 1 2 3 4  1334 1335 1336 1337 1338	300 300 5 320 300 • 1 1	.0 60 kg .0 1 .0 60 kg	IETAD IETAD	Agricultural Agricultural Specialty Co Agricultural Agricultural Specialty Co Specialty Co Specialty Co Specialty Co Specialty Co Specialty Co	Development Development Development Fiee Associaties A	nt plc iation nt plc nt plc nt plc iation iation iation iation
	larvest.Year	Gradin	ıg.Date	Total.C	up.Points	Moisture
0	2014	April 4th	2015		90.58	0.12
		-		•••		
1	2014	April 4th	, 2015	• • •	89.92	0.12
2	2012	May 31st	2010		89.75	0.00
3	2014	March 26th	, 2015		89.00	0.11
4	2014	April 4th	2015		88.83	0.12
1334	2016	January 19th	, 2016		78.75	0.00
1335	2016	January 19th	, 2016		78.08	0.00
1336	2014	December 23rd	2014		77.17	0.00
1337	2013	August 25th	2014		75.08	0.10
1338	2013	August 25th	2014		73.75	0.12

```
Category.One.Defects
                             Quakers
                                     Category. Two. Defects
0
                                 0.0
                                                           0
                          0
1
                          0
                                 0.0
                                                           1
2
                          0
                                 0.0
                                                           0
3
                                                           2
                          0
                                 0.0
4
                          0
                                 0.0
                                                           2
                                  . . .
                                                         . . .
1334
                          0
                                 0.0
                                                           1
1335
                          0
                                 0.0
                                                           0
1336
                          0
                                 0.0
                                                           6
                                                           1
1337
                         20
                                 0.0
                                                           9
1338
                         63
                                 0.0
                                             Certification. Body
               Expiration
0
          April 3rd, 2016 METAD Agricultural Development plc
1
          April 3rd, 2016 METAD Agricultural Development plc
2
           May 31st, 2011
                                   Specialty Coffee Association
3
         March 25th, 2016 METAD Agricultural Development plc
4
          April 3rd, 2016 METAD Agricultural Development plc
1334
       January 18th, 2017
                                   Specialty Coffee Association
1335
     January 18th, 2017
                                   Specialty Coffee Association
      December 23rd, 2015
1336
                                  Specialty Coffee Association
       August 25th, 2015
                                   Specialty Coffee Association
1337
1338
       August 25th, 2015
                                  Specialty Coffee Association
                          Certification.Address
0
      309fcf77415a3661ae83e027f7e5f05dad786e44
1
      309fcf77415a3661ae83e027f7e5f05dad786e44
2
      36d0d00a3724338ba7937c52a378d085f2172daa
3
      309fcf77415a3661ae83e027f7e5f05dad786e44
      309fcf77415a3661ae83e027f7e5f05dad786e44
. . .
     ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1334
1335
     ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1336
     ff7c18ad303d4b603ac3f8cff7e611ffc735e720
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
1337
1338
      ff7c18ad303d4b603ac3f8cff7e611ffc735e720
                          Certification.Contact
                                                  unit of measurement
0
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                     m
1
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                     m
2
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
                                                                     m
3
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                     m
      19fef5a731de2db57d16da10287413f5f99bc2dd
                                                                     m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1334
                                                                     m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1335
                                                                     m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1336
                                                                     m
      352d0cf7f3e9be14dad7df644ad65efc27605ae2
1337
```

[1339 rows x 31 columns]

df.isnull().sum()

Species	0
Owner	0
Country.of.Origin	0
Region	0
Producer	0
Number.of.Bags	0
Bag.Weight	0
In.Country.Partner	0
Harvest.Year	0
Grading.Date	0
Owner.1	0
Aroma	0
Flavor	0
Aftertaste	0
Acidity	0
Body	0
Balance	0
Uniformity	0
Clean.Cup	0
Sweetness	0
Cupper.Points	0
Total.Cup.Points	0
Moisture	0
Category.One.Defects	0
Quakers	0
Category.Two.Defects	0
Expiration	0
Certification.Body	0
Certification.Address	0
Certification.Contact	0
unit_of_measurement	0
dtype: int64	

df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1339 entries, 0 to 1338
Data columns (total 31 columns):

#	Column	Non-Null Count	Dtype
0	Species Owner	1339 non-null 1339 non-null	object object
2	Country.of.Origin Region	1339 non-null	object object

```
4
         Producer
                                                      1339 non-null
                                                                                     object
  5 Number.of.Bags
                                                    1339 non-null float64
 6 Bag.Weight
      Bag.Weight 1339 non-null object In.Country.Partner 1339 non-null object
 7
  8 Harvest.Year
                                                   1339 non-null object
  9 Grading.Date
                                                   1339 non-null object
                                                   1339 non-null object
 10 Owner.1
 11 Aroma
                                                   1339 non-null float64

      11 Aroma
      1339 non-null
      float64

      12 Flavor
      1339 non-null
      float64

      13 Aftertaste
      1339 non-null
      float64

      14 Acidity
      1339 non-null
      float64

      15 Body
      1339 non-null
      float64

      16 Balance
      1339 non-null
      float64

      17 Uniformity
      1339 non-null
      float64

      18 Clean.Cup
      1339 non-null
      float64

      19 Sweetness
      1339 non-null
      float64

      20 Cupper.Points
      1339 non-null
      float64

      21 Total.Cup.Points
      1339 non-null
      float64

      22 Moisture
      1339 non-null
      float64

 22 Moisture
                                                    1339 non-null float64
 23 Category.One.Defects 1339 non-null int64
24 Quakers 1339 non-null float64
 25 Category. Two. Defects 1339 non-null int64
 26 Expiration 1339 non-null object 27 Certification.Body 1339 non-null object
 28 Certification.Address 1339 non-null object
 29 Certification.Contact 1339 non-null object
 30 unit of measurement 1339 non-null object
dtypes: float64(14), int64(2), object(15)
memory usage: 324.4+ KB
```

df.shape

(1339, 31)

#Outlier Dectation

new df=df.select dtypes(exclude=['object'])

new df

	Number.of.Bags	Aroma	Flavor	Aftertaste	Acidity	Body
Balanc	:e \					
0	300.0	8.67	8.83	8.67	8.75	8.50
8.42						
1	300.0	8.75	8.67	8.50	8.58	8.42
8.42						
2	5.0	8.42	8.50	8.42	8.42	8.33
8.42						
3	320.0	8.17	8.58	8.42	8.42	8.50
8.25						
4	300.0	8.25	8.50	8.25	8.50	8.42

8.33								
•		1 0	7 75	7. 50	7 22	7. 50	···	• •
1334 7.83		1.0	7.75	7.58	7.33		5.08	
1335 5.25		1.0	7.50	7.67	7.75	7.75	5.17	
1336 7.17		1.0	7.33	7.33	7.17	7.42	7.50	
1337		1.0	7.42	6.83	6.75	7.17	7.25	
7.00 1338		1.0	6.75	6.67	6.50	6.83	6.92	
6.83								
Total.	Uniformity Cup. Points		an.Cup	Sweetness	Cupper.	Points		
0 90.58	10.00		10.00	10.00		8.75		
1	10.00		10.00	10.00		8.58		
89.92	10.00		10.00	10.00		9.25		
89.75	10.00		10.00	10.00		8.67		
89.00	10.00		10.00	10.00		8.58		
88.83								
	10.00		10.00			7.00		• •
1334 78.75	10.00		10.00	7.75		7.83		
1335 78.08	10.00		10.00	8.42		8.58		
1336 77.17	9.33		9.33	7.42		7.17		
1337 75.08	9.33		9.33	7.08		6.92		
1338	9.33		9.33	6.67		7.92		
73.75								
0 1 2 3 4	Moisture 0.12 0.12 0.00 0.11 0.12	ateg	ory.One	.Defects 0 0 0 0	0.0 0.0 0.0 0.0 0.0	ategory.Tv	vo.Defec	ts 0 1 0 2 2
1334 1335 1336	0.00			0 0 0	0.0			1 0 6

```
1337
          0.10
                                   20
                                           0.0
1338
          0.12
                                   63
                                           0.0
                                                                    9
[1339 rows x 16 columns]
Q1=new df.quantile(0.25)
Q1
                         14.00
Number.of.Bags
                          7.42
Aroma
                          7.33
Flavor
Aftertaste
                          7.25
Acidity
                          7.33
                          7.33
Body
                         7.33
Balance
Uniformity
                         10.00
Clean.Cup
                         10.00
                         10.00
Sweetness
Cupper.Points
                         7.25
Total.Cup.Points
                         81.08
Moisture
                          0.09
Category.One.Defects
                          0.00
Quakers
                          0.00
Category. Two. Defects
                          0.00
Name: 0.25, dtype: float64
Q3=new df.quantile(0.75)
Q3
Number.of.Bags
                         275.00
                           7.75
Aroma
Flavor
                           7.75
                           7.58
Aftertaste
Acidity
                           7.75
Body
                           7.67
                          7.75
Balance
Uniformity
                          10.00
Clean.Cup
                          10.00
Sweetness
                         10.00
                          7.75
Cupper.Points
Total.Cup.Points
                          83.67
Moisture
                           0.12
Category.One.Defects
                           0.00
Quakers
                           0.00
Category. Two. Defects
                           4.00
Name: 0.75, dtype: float64
IQR=Q3-Q1
```

```
IQR
                       261.00
Number.of.Bags
Aroma
                         0.33
                         0.42
Flavor
Aftertaste
                         0.33
                         0.42
Acidity
                         0.34
Body
                         0.42
Balance
Uniformity
                         0.00
                         0.00
Clean.Cup
Sweetness
                         0.00
                         0.50
Cupper.Points
                         2.59
Total.Cup.Points
Moisture
                         0.03
Category.One.Defects
                         0.00
Quakers
                         0.00
Category. Two. Defects
                         4.00
dtype: float64
a = (new df < (Q1-1.5*IQR)) | (new df > (Q3+1.5*IQR))
     Number.of.Bags Aroma Flavor Aftertaste Acidity Body
Balance \
0
              False True
                             True
                                         True
                                                  True
                                                        True
True
              False True
                             True
                                                 True True
                                         True
True
2
              False True
                             True
                                         True
                                                  True
                                                        True
True
3
              False False
                             True
                                         True
                                                  True True
False
              False True
                             True
                                         True
                                                 True
                                                        True
False
. . .
1334
              False False False
                                        False False True
False
1335
              False False
                            False
                                        False False
                                                        True
True
1336
              False False False
                                        False False
False
1337
              False False False
                                         True False False
False
1338
              False True True
                                         True False False
False
     Uniformity Clean.Cup Sweetness Cupper.Points
Total.Cup.Points \
```

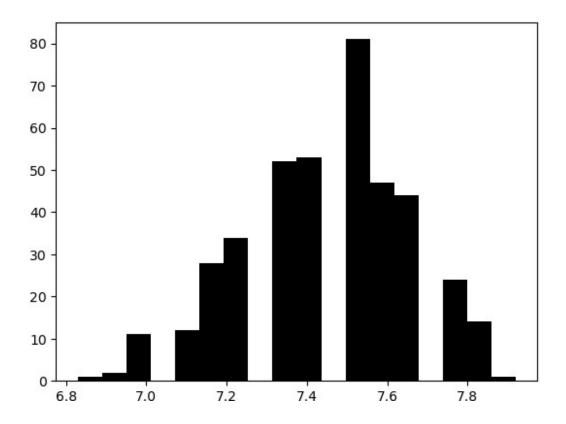
0	False	False	False	9	True	
True 1	False	False	False	2	True	
True						
2	False	False	False	9	True	
True 3	False	False	False	٥	True	
True	raise	raise	raiso	-	iiuc	
4	False	False	False	€	True	
True						
• • •	• • •	• • •	• •	•	• • •	• •
1334	False	False	True	3	False	
False						
1335	False	False	True	9	True	
False	True	True	True	٥	False	
True	TTUC	TTUC	IIu	_	14150	
1337	True	True	True	9	False	
True			m		n-1	
1338 True	True	True	True	2	False	
1140						
	Moisture	:ategory.One		)uakers	Category.T	wo.Defects
0	False False		False False	False False		False False
2	True		False	False		False
3	False		False	False		False
4	False		False	False		False
1 2 2 4						 Ealas
1334 1335	True True		False False	False False		False False
1336	True		False	False		False
1337	False		True	False		False
1338	False		True	False		False
[1339	rows x 16	columnsl				
		COLUMNIS				
a.sum	()					
Number	r.of.Bags	4				
Aroma	_	75				
Flavor		4 4				
Aftert Acidit		87 25				
Body	~ <u>Y</u>	34				
Baland	ce	40				
Unifor	_	187				
Clean	_	120				
Sweetr	iess	121				

```
Cupper.Points
                        33
Total.Cup.Points
                        72
                       305
Moisture
Category.One.Defects
                       202
Quakers
                        95
Category. Two. Defects
                        94
dtype: int64
filter=df[\sim((new df<Q1-1.5*IQR)) | (new df>Q1+1.5*IQR)).any(axis=1)]
filter
     Species
                                       Owner Country.of.Origin \
148
     Arabica
                   juan luis alvarado romero
                                               Guatemala
160 Arabica
                      yunnan coffee exchange
                                                         China
196 Arabica
                                                         China
                      yunnan coffee exchange
203 Arabica
                   juan luis alvarado romero
                                                    Guatemala
206 Arabica
                       jesus salazar velasco
                                                       Mexico
. . .
                                                           . . .
1152 Arabica
                     calixto guillen vazquez
                                                        Mexico
                   juan luis alvarado romero
1160 Arabica
                                                     Guatemala
1166 Arabica armando luis pohlenz martinez
                                                        Mexico
1167 Arabica exportadora de cafe condor s.a
                                                     Colombia
                      lin, che-hao krude 林哲豪 Taiwan
1182 Arabica
                  Region
Producer \
148
                                                        ROBERTO
                  jalapa
MONTERROSO
160
                  yunnan
                                 Menglian Ban'an Coffee Processing
Factory
                  yunnan Menglian County Nayun Town Mangzhang Coffee
196
Farm
203
                 oriente
                                                            AMILCAR
LAPOLA
206 san pedro cotzilnam
                                                     JESUS SALAZAR
VELASCO
. . .
. . .
1152
                 chiapas
                                                        VARIOS
PRODUCTORES
1160
                 oriente
                                                 AGRIPEC DE LA VEGA Y
CIA.
1166
      la concordia
                                             ARMANDO LUIS POHLENZ
MARTINEZ
1167
                   huila
                                                                  La
Plata
1182
                                                             WU SHU
           natou county
YI 巫叔憶
```

	Number.of.Bags	Bag.Weigl	nt	In.Coun	try.Partr	ner
	st.Year \					
148	250.0	69	kg	Asociacion Nacion	al Del Ca	ıfé
2014						
160	3.0	60	kg	Yunnan Coff	ee Exchar	nge
2015						
196	3.0	60	kg	Yunnan Coff	ee Exchar	nge
2015			_			_
203	250.0	69	kg	Asociacion Nacion	al Del Ca	afé
2014						
206	10.0	1	kg		AMECA	AFE
2012			_			
1152	380.0	1	kq		AMECA	AFE
2013						
1160	250.0	69	kq	Asociacion Nacion	al Del Ca	ıfé
2013			,			
1166	250.0	1	kα		AMEC <i>A</i>	AFF.
2012	200.0	_	. 5		01	
1167	250.0	70	kg		Almaca	afé
4T/10			_			
182	50.0	20	kg	Specialty Coffee	Associati	on
2014				1 1		
	Grading	g.Date .		Total.Cup.Points	Moisture	\
148	May 29th,	2014 .		84.58	0.11	
60	April 6th,	2016 .		84.50	0.10	
96	April 6th,	2016 .		84.25	0.10	
03	August 21st,	2014 .		84.25	0.10	
206	September 11th,	2012 .		84.25	0.12	
L152	March 29th,	2013 .		79.75	0.12	
1160	June 27th,			79.67	0.11	
1166	August 30th,			79.58	0.13	
L167	February 9th,			79.58	0.10	
182	November 7th,			79.25	0.11	
	Category.One.De	fects Qu	ıak	ers Category.Two.	Defects	\
L48		0		0.0	1	
.60		0		0.0	1	
L96		0		0.0	1	
203		0		0.0	1	
206		0		0.0	1	
				• • •		
1152		0		0.0	5	
1160		0		0.0	6	
1166		0		0.0	5	
1107		0		0.0	4	
167						

```
0.0
1182
                                      Certification.Body \
                Expiration
148
            May 29th, 2015 Asociacion Nacional Del Café
           April 6th, 2017
160
                                 Yunnan Coffee Exchange
196
           April 6th, 2017
                                 Yunnan Coffee Exchange
203
         August 21st, 2015
                           Asociacion Nacional Del Café
      September 11th, 2013
206
                                                 AMECAFE
. . .
1152
          March 29th, 2014
                                                 AMECAFE
1160
           June 27th, 2014
                           Asociacion Nacional Del Café
         August 30th, 2013
1166
                                                 AMECAFE
1167
        February 9th, 2012
                                                Almacafé
        November 7th, 2015
1182
                            Specialty Coffee Association
                         Certification.Address \
148
     b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
160
      683fa6014608fc10ff681b0435b0b2dbe6df988f
196
      683fa6014608fc10ff681b0435b0b2dbe6df988f
203
      b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
      59e396ad6e22a1c22b248f958e1da2bd8af85272
206
1152
      59e396ad6e22a1c22b248f958e1da2bd8af85272
1160 b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
      59e396ad6e22a1c22b248f958e1da2bd8af85272
1166
      e493c36c2d076bf273064f7ac23ad562af257a25
1167
1182 36d0d00a3724338ba7937c52a378d085f2172daa
                         Certification.Contact
                                                unit of measurement
148
      724f04ad10ed31dbb9d260f0dfd221ba48be8a95
                                                                  ft
160
      f6d87a6c04653c569d4911a66f89d5e30ce83b93
                                                                   m
196
      f6d87a6c04653c569d4911a66f89d5e30ce83b93
                                                                   m
203
      724f04ad10ed31dbb9d260f0dfd221ba48be8a95
                                                                  ft
      0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
206
                                                                   m
                                                                 . . .
1152
      0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
                                                                   m
      724f04ad10ed31dbb9d260f0dfd221ba48be8a95
1160
                                                                  ft
1166
      0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
                                                                  m
      70d3c0c26f89e00fdae6fb39ff54f0d2eb1c38ab
1167
                                                                   m
1182
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
[404 rows x 31 columns]
filter.info()
<class 'pandas.core.frame.DataFrame'>
Index: 404 entries, 148 to 1182
Data columns (total 31 columns):
                            Non-Null Count Dtype
    Column
```

```
0
    Species
                          404 non-null
                                         object
1
    Owner
                          404 non-null
                                         object
 2
    Country.of.Origin
                          404 non-null
                                         object
 3
    Region
                          404 non-null
                                         object
 4
   Producer
                          404 non-null
                                         object
 5
    Number.of.Bags
                          404 non-null
                                        float64
 6
   Bag.Weight
                          404 non-null
                                        object
 7
   In.Country.Partner
                          404 non-null
                                        object
   Harvest.Year
                          404 non-null
                                         object
 8
 9
   Grading.Date
                          404 non-null
                                         object
10 Owner.1
                          404 non-null
                                         object
                                         float64
11 Aroma
                          404 non-null
12 Flavor
                          404 non-null
                                        float64
13 Aftertaste
                          404 non-null
                                        float64
14 Acidity
                         404 non-null
                                        float64
                          404 non-null
                                        float64
15 Body
16 Balance
                         404 non-null
                                        float64
17 Uniformity
                         404 non-null
                                       float64
18 Clean.Cup
                         404 non-null
                                        float64
19 Sweetness
                         404 non-null
                                        float64
20 Cupper.Points
                          404 non-null float64
                        404 non-null float64
404 non-null float64
21 Total.Cup.Points
22 Moisture
 23 Category.One.Defects 404 non-null
                                        int64
                          404 non-null
24 Quakers
                                        float64
25 Category. Two. Defects 404 non-null
                                        int64
                                       object
26 Expiration
                          404 non-null
27 Certification.Body
                          404 non-null object
28 Certification.Address 404 non-null
                                        object
29 Certification.Contact 404 non-null
                                         object
30 unit of measurement 404 non-null
                                         object
dtypes: float64(14), int64(2), object(15)
memory usage: 101.0+ KB
filter.shape
(404, 31)
                                                 #Univariate
plt.hist(filter['Balance'],bins=18,color='Black')
(array([ 1., 2., 11., 0., 12., 28., 34., 0., 52., 53., 0., 81.,
47.,
       44., 0., 24., 14., 1.]),
                , 6.89055556, 6.95111111, 7.01166667, 7.07222222,
 array([6.83
       7.13277778, 7.19333333, 7.25388889, 7.31444444, 7.375
       7.43555556, 7.49611111, 7.55666667, 7.61722222, 7.67777778,
       7.73833333, 7.79888889, 7.85944444, 7.92 ]),
<BarContainer object of 18 artists>)
```



<pre>df.groupby(['Country.of.Origin']).count()</pre>				
	Species	Owner Red	gion Produ	ucer \
Country.of.Origin	-			
Brazil	132	132	132	132
Burundi	2	2	2	2
China	16	16	16	16
Colombia	183	183	183	183
Costa Rica	51	51	51	51
Cote d?Ivoire	1	1	1	1
Ecuador	3	3	3	3
El Salvador	21	21	21	21
Ethiopia	44	44	44	44
Guatemala	181	181	181	181
Haiti	6	6	6	6
Honduras	53	53	53	53
India	14	14	14	14
Indonesia	20	20	20	20
Japan	1	1	1	1
Kenya	25	25	25	25
Laos	3	3	3	3
Malawi	11	11	11	11
Mauritius	1	1	1	1
Mexico	237	237	237	237
Myanmar	8	8	8	8

Nicaragua Panama Papua New Guinea Peru Philippines Rwanda Taiwan Tanzania, United Republic Of Thailand Uganda United States United States (Hawaii) United States (Puerto Rico) Vietnam Zambia	26 26 4 4 1 1 10 10 5 5 1 1 75 75 40 40 32 32 36 36 10 10 73 73 4 4 8 8 1 1		26 4 1 10 5 1 75 40 32 36 10 73 4 8 1	
<pre>In.Country.Partner \ Country.of.Origin</pre>	Number.of.Bags	Bag.Weight		
Brazil	132	132		
132	-			
Burundi	2	2		
2				
China	16	16		
16 Colombia	183	183		
183	183	183		
Costa Rica	51	51		
51				
Cote d?Ivoire	1	1		
1				
Ecuador	3	3		
3 El Salvador	21	21		
21	21	21		
Ethiopia	44	44		
44				
Guatemala	181	181		
181				
Haiti 6	6	6		
6 Honduras	53	53		
53	55	55		
India	14	14		
14				
Indonesia	20	20		
20	1	7		
Japan	1	1		

1			
1	٥٢	٥٦	
Kenya	25	25	
25			
Laos	3	3	
3			
Malawi	11	11	
11			
Mauritius	1	1	
1			
Mexico	237	237	
237			
Myanmar	8	8	
8			
Nicaragua	26	26	
26			
Panama	4	4	
4			
Papua New Guinea	1	1	
1			
Peru	10	10	
10	10	10	
Philippines	5	5	
5	9	9	
Rwanda	1	1	
1	1	1	
Taiwan	75	75	
75	7.5	7.5	
Tanzania, United Republic Of	40	40	
40	40	40	
	2.0	2.0	
Thailand 32	32	32	
	2.0	2.0	
Uganda	36	36	
36	1.0	1.0	
United States	10	10	
10	5.0	<b>5</b> 0	
United States (Hawaii)	73	73	
73			
United States (Puerto Rico)	4	4	
4			
Vietnam	8	8	
8			
Zambia	1	1	
1			
	Harvest.Year	Grading.Date	Owner.1
\			
Country.of.Origin			
	1 2 0	122	132
Brazil	132	132	102

Burundi	2	2	2	
China	16	16	16	
Colombia	183	183	183	
Costa Rica	51	51	51	
Cote d?Ivoire	1	1	1	
Ecuador	3	3	3	
El Salvador	21	21	21	
Ethiopia	44	44	44	
Guatemala	181	181	181	
Haiti	6	6	6	
Honduras	53	53	53	
India	14	14	14	
Indonesia	20	20	20	
Japan	1	1	1	
Kenya	25	25	25	
Laos	3	3	3	
Malawi	11	11	11	
Mauritius	1	1	1	
Mexico	237	237	237	
Myanmar	8	8	8	
Nicaragua	26	26	26	
Panama	4	4	4	
Papua New Guinea	1	1	1	
Peru	10	10	10	
Philippines	5	5	5	
Rwanda	1	1	1	

Taiwan	75	75	75	
Tanzania, United Republic Of	40	40	40	
Thailand	32	32	32	
Uganda	36	36	36	
United States	10	10	10	
United States (Hawaii)	73	73	73	
United States (Puerto Rico)	4	4	4	
Vietnam	8	8	8	
Zambia	1	1	1	

	Total.Cup.Points	Moisture \
Country.of.Origin		
Brazil	132	132
Burundi	2	2
China	16	16
Colombia	183	183
Costa Rica	51	51
Cote d?Ivoire	1	1
Ecuador	3	3
El Salvador	21	21
Ethiopia	44	4 4
Guatemala	181	181
Haiti	6	6
Honduras	53	53
India	14	14
Indonesia	20	20
Japan	1	1
Kenya	25	25
Laos	3	3
Malawi	11	11
Mauritius	1	1
Mexico	237	237
Myanmar	8	8
Nicaragua	26	26
Panama	4	4
Papua New Guinea	1	1
Peru	10	10
Philippines	5	5
Rwanda	1	1
Taiwan	75	75
Tanzania, United Republic Of	40	40

Thailand Uganda United States United States (Hawaii) United States (Puerto Rico) Vietnam Zambia	32 36 10 73 4 8 1	32 36 10 73 4 8
24	_	_
	Category.One.Defects	Ouakore \
Country.of.Origin	category.one.berects	Quakers (
	1 2 2	1 2 2
Brazil	132	
Burundi	2	
China	16	
Colombia	183	
Costa Rica	51	
Cote d?Ivoire	1	
Ecuador	3	
El Salvador	21	
Ethiopia	44	4 4
Guatemala	181	
Haiti	6	
Honduras	53	
India	14	
Indonesia	20	20
Japan	1	1
Kenya	25	25
Laos	3	3
Malawi	11	11
Mauritius	1	1
Mexico	237	237
Myanmar	8	8
Nicaragua	26	26
Panama	4	4
Papua New Guinea	1	1
Peru	10	10
Philippines	5	5
Rwanda	1	1
Taiwan	75	75
Tanzania, United Republic Of	40	
Thailand	32	32
Uganda	36	36
United States	10	10
United States (Hawaii)	73	73
United States (Puerto Rico)	4	4
Vietnam	8	8
Zambia	1	1
	Category. Two. Defects	Expiration \
Country.of.Origin		

Brazil	132	132
Burundi	2	2
China	16	16
Colombia	183	183
Costa Rica	51	51
Cote d?Ivoire	1	1
Ecuador	3	3
El Salvador	21	21
Ethiopia	44	44
Guatemala	181	181
Haiti	6	6
Honduras	53	53
India	14	14
Indonesia	20	20
Japan	1	1
Kenya	25	25
Laos	3	3
Malawi	11	11
Mauritius	1	1
Mexico	237	237
Myanmar	8	8
Nicaragua	26	26
Panama	4	4
Papua New Guinea	1	1
Peru	10	10
Philippines	5	5
Rwanda	1	1
Taiwan	75	75
Tanzania, United Republic Of	40	40
Thailand	32	32
Uganda	36	36
United States	10	10
United States (Hawaii)	73	73
United States (Puerto Rico)	4	4
Vietnam	8	8
Zambia	1	1
Contification Address \	Certification.Body	
(Cortification Addross )		

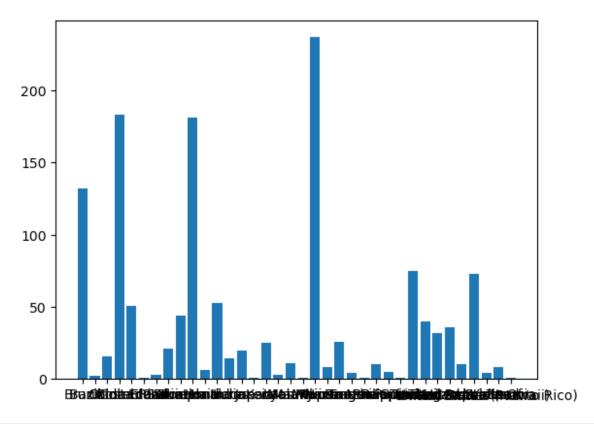
	001011000101110001
Certification.Address \	_
Country.of.Origin	
Brazil	132
132	
Burundi	2
2	
China	16
16	
Colombia	183
183	

51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 55 Laos 3 3 Malawi 11 11 Mauritius 1 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1		
Cote d?Ivoire       1         1       1         Ecuador       3         3       2         El Salvador       21         21       44         44       44         44       44         44       4         40       6         6       6         Honduras       53         53       53         India       14         14       1         Indonesia       20         20       20         Japan       1         Kenya       25         Laos       3         3       3         Malawi       11         11       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         Philippines       5         5 <t< td=""><td>Costa Rica</td><td>51</td></t<>	Costa Rica	51
Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 55 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 Philippines 5 5 Rwanda 1 1 Taiwan 75 75	51	
Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 55 Laos 3 3 Malawi 11 11 Mexico 237 Myanmar 8 Nicaragua 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 Rwanda 1 1 Taiwan 75 75		1
3       21       21       21       21       21       21       21       21       21       21       21       21       21       21       21       21       24       44       44       44       44       42       42       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44       44		_
El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 66 Honduras 53 53 India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 55 Rwanda 1 1 Taiwan 75		3
21         Ethiopia       44         44         44         44         44         44         181         Haiti       6         6         Honduras       53         53         India       14         14       14         India       14         14       14         India       14         14       14         India       14         Kenya       25         25       25         Laos       3         3       3         Malawi       11         11       11         Mexico       237         237       237         Myanmar       8         8       Nicaragua       26         26       2         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         Philippines       5         5       5         Rwanda       1 <t< td=""><td></td><td>0.1</td></t<>		0.1
Ethiopia 44 44 Guatemala 181 181 Haiti 66 Honduras 53 53 India 14 Ifficial 20 Indonesia 20 Indonesia 20 India 15 India 16 India 17 Indonesia 20 India 17 India 18 India 19 India 19 India 10 India 10 India 10 India 10 India 11 Ind		21
44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		4.4
Guatemala       181         181       181         Haiti       6         Honduras       53         53       53         India       14         14       14         Indonesia       20         20       20         Japan       1         Kenya       25         25       25         Laos       3         Malawi       11         11       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       10         0       Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75		44
181         Haiti       6         6       6         Honduras       53         53       1         India       14         14       14         14       14         Indonesia       20         20       20         Japan       1         Kenya       25         25       1         Laos       3         3       3         Malawi       11         11       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75		1 8 1
Haiti       6         Honduras       53         53       53         India       14         14       14         14       14         Indonesia       20         20       20         Japan       1         Kenya       25         25       1         Laos       3         3       3         Malawi       11         11       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75		101
6 Honduras 53 53 India 14 14 Indonesia 20 20 Japan 1 I Kenya 25 25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		6
## Honduras		Ç
53 India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		53
India 14 14 Indonesia 20 20 Japan 1 1 Kenya 25 25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75	53	
14         Indonesia       20         20       20         Japan       1         1       1         Kenya       25         25       3         Laos       3         3       3         Malawi       11         11       1         Mauritius       1         1       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75	India	14
20 Japan 1 1 Kenya 25 25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 Nicaragua 26 Panama 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75	14	
Japan       1         Kenya       25         25       5         Laos       3         3       3         Malawi       11         11       1         Mauritius       1         1       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75	Indonesia	20
1       Kenya       25         25       25         Laos       3         3       3         Malawi       11         11       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       2         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75	20	
Kenya       25         25       25         Laos       3         3       3         Malawi       11         11       1         Mauritius       1         1       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		1
25 Laos 3 3 Malawi 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75	1	
Laos 3 3 Malawi 11 11 11 Mauritius 1 1 Mexico 237 237 Myanmar 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		25
3         Malawi       11         11         Mauritius       1         1       1         Mexico       237         237       237         Myanmar       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       1         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		
Malawi       11         11       11         Mauritius       1         1       1         Mexico       237         237       237         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		3
11 Mauritius 1 1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		7.7
Mauritius 1  Mexico 237 237  Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		11
1 Mexico 237 237 Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		1
Mexico       237         237       37         Myanmar       8         8       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		1
237  Myanmar 8 8 Nicaragua 26 26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		227
Myanmar       8         Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       10         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		231
8         Nicaragua       26         26         Panama       4         4       4         Papua New Guinea       1         1       10         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		8
Nicaragua       26         26       26         Panama       4         4       4         Papua New Guinea       1         1       10         Peru       10         10       10         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		9
26 Panama 4 4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		26
Panama       4         4       4         Papua New Guinea       1         1       1         Peru       10         10       1         Philippines       5         5       5         Rwanda       1         1       1         Taiwan       75         75       75		
4 Papua New Guinea 1 1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75	Panama	4
1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75		
1 Peru 10 10 Philippines 5 5 Rwanda 1 1 Taiwan 75	Papua New Guinea	1
10 Philippines 5 5 Rwanda 1 1 Taiwan 75	1	
Philippines 5 5 Rwanda 1 1 Taiwan 75	Peru	10
5 Rwanda 1 1 Taiwan 75 75		
Rwanda 1 1 Taiwan 75 75		5
1 Taiwan 75 75		
Taiwan 75		1
75		
		75
Tanzania, United Republic Of 40		4.0
	ranzania, United Republic Ui	40

36				
32 Uganda 36 36 United States 10 10 10 United States (Hawaii) 73 73 United States (Puerto Rico) 4 4 Vietnam 8 8 Zambia 1 1  Certification.Contact Unit_of_measurement Country.of.Origin  Brazil 132 132 Buzundi 2 China 16 16 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ecuador 3 3 3 3 5 El Salvador 21 Elthiopia 44 44 44 44 44 44 44 44 44 44 44 44 44	40			
Uganda 36 36 36 10 10 10 10 10 10 10 10 173 73 United States (Hawaii) 73 73 United States (Puerto Rico) 4 4 Vietnam 8 8 8 10 11 11 11 11 11 12 132 132 132 132 132 1	Thailand		32	
36 United States 10 10 United States (Hawaii) 73 73 United States (Puerto Rico) 4 4 Vetnam 8 8 Zambia 1 1  Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 132 Burundi 2 2 China 16 16 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ceuador 3 3 3 El Salvador 21 Ethiopia 44 44 44 44 44 44 44 44 44 44 44 44 44	32			
United States 10 10 10 United States (Hawaii) 73 73 United States (Puerto Rico) 4 4 Vietnam 8 8 2ambia 1 1 Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 132 Burundi 2 2 China 16 6 Colombia 183 183 183 183 183 183 183 183 183 183	Uganda		36	
10 United States (Hawaii) 73 73 73 United States (Puerto Rico) 4 4 Vietnam 8 8 Zambia 1 1  Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 132 Burundi 2 2 China 16 Colombia 16 6 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 Ecuador 3 3 3 El Salvador 21 El Salvador 21 El Salvador 21 El Salvador 21 El Salvador 31 El Salvado	36			
United States (Hawaii) 73 73 United States (Puerto Rico) 4 4 4 Vietnam 8 8 Zambia 1 1	United States		10	
73 United States (Puerto Rico) 4 4 Vietnam 8 8 Zambia 1	10			
United States (Puerto Rico) 4 4 Vietnam 8 8 8 Zambia 1 1  Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 132 Burundi 2 2 China 16 6 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 21 Ethiopia 44 44 Guatemala 181 181 181 181 184 184 44 Guatemala 181 181 184 184 186 Haiti 6 6 Honduras 53 53 India 14 14 14 17 Indonesia 20 20		(Hawaii)	73	
Vietnam       8         8       8         Zambia       1         1       Certification.Contact         unit_of_measurement       Country.of.Origin         Brazil       132         132       Burundi       2         China       16         16       Colombia       183         183       Costa Rica       51         51       51	United States	(Puerto Rico)	4	
8 Zambia 1 1  Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 32 Burundi 2 2 China 16 6 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Cote d?Ivoire 1 1 Etuador 3 3 3 E1 Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 14 Indonesia 20 20			0	
Zambia 1  Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132  Burundi 2  China 16  Colombia 183  183  Costa Rica 51  51  Cote d?Ivoire 1  1  Ecuador 3  3  Sell Salvador 21  Ethiopia 44  44  44  Guatemala 181  181  Haiti 6  Guatemala 181  181  Haiti 6  Honduras 53  53  India 14  14  17  Indonesia 20  20			8	
Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 Burundi 2 2 China 16 Colombia 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Godon Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 14 Indonesia 20			1	
Certification.Contact unit_of_measurement Country.of.Origin  Brazil 132 Burundi 2 2 China 16 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 China 16 Colombia 183 Costa Rica 51 Cote d?Ivoire 1 1 Cote d?Ivoir			I	
unit_of_measurement         Country.of.Origin         Brazil       132         132         Burundi       2         2       2         China       16         16       6         Colombia       183         183       5         Costa Rica       51         51       51         51       51         51       1         10       2         20       2         Ecuador       3         3       3         El Salvador       21         21       21         21       21         21       21         21       3         8       44         44       44         Guatemala       181         181       6         6       6         Honduras       53         53       3         India       14         14       14         14       14         14       14         14       14         15       16         16	1			
unit_of_measurement         Country.of.Origin         Brazil       132         132         Burundi       2         2       2         China       16         16       6         Colombia       183         183       5         Costa Rica       51         51       51         51       51         51       1         10       2         20       2         Ecuador       3         3       3         El Salvador       21         21       21         21       21         21       21         21       3         8       44         44       44         Guatemala       181         181       6         6       6         Honduras       53         53       3         India       14         14       14         14       14         14       14         14       14         15       16         16			Contification Contest	
Country.of.Origin  Brazil 132  Burundi 2  China 16  Colombia 183  183  Costa Rica 51  51  Cote d?Ivoire 1  1	unit of monguin	oom on t	Certification.Contact	
Brazil 132 Burundi 2 2 China 16 Colombia 183 R3 Costa Rica 51 S1 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 Ethiopia 44 44 Guatemala 181 Haiti 6 G Honduras 53 S1 India 14 Indonesia 20				
132         Burundi       2         2         China       16         16       18         Colombia       183         183       51         State       51         51       1         Cote d?Ivoire       1         1       1         Ecuador       3         3       21         21       21         21       21         21       44         44       44         Guatemala       181         181       181         Haiti       6         6       6         Honduras       53         53       53         India       14         14       14         14       14         15       14         16       14         17       14         18       14         19       14         14       14         15       15         16       16         17       17         18       18         19       18 <td>Country.of.Ori</td> <td>ıgın</td> <td></td> <td></td>	Country.of.Ori	ıgın		
132         Burundi       2         2         China       16         16       18         Colombia       183         183       51         State       51         51       1         Cote d?Ivoire       1         1       1         Ecuador       3         3       21         21       21         21       21         21       44         44       44         Guatemala       181         181       181         Haiti       6         6       6         Honduras       53         53       53         India       14         14       14         14       14         15       14         16       14         17       14         18       14         19       14         14       14         15       15         16       16         17       17         18       18         19       18 <td>Brazil</td> <td></td> <td>132</td> <td></td>	Brazil		132	
Burundi 2 2 China 16 16 Colombia 183 183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 Indonesia 20 20			132	
2         China       16         16       183         Colombia       183         183       51         Costa Rica       51         51       51         Cote d?Ivoire       1         1       1         Ecuador       3         3       3         El Salvador       21         21       21         Ethiopia       44         44       44         Guatemala       181         181       8         Haiti       6         6       6         Honduras       53         53       53         India       14         14       14         Indonesia       20			2	
China       16         16       183         Colombia       183         183       183         Costa Rica       51         51       1         Cote d?Ivoire       1         1       1         Ecuador       3         3       21         21       21         21       21         Ethiopia       44         44       44         Guatemala       181         181       181         Haiti       6         6       6         Honduras       53         53       53         India       14         14       14         Indonesia       20			2	
16         Colombia       183         183       51         Costa Rica       51         51			16	
Colombia       183         183       51         Sota Rica       51         51       51         Cote d?Ivoire       1         1       1         Ecuador       3         3       3         El Salvador       21         21       21         Ethiopia       44         44       44         Guatemala       181         181       6         Haiti       6         6       6         Honduras       53         53       53         India       14         14       14         Indonesia       20         20       20			10	
183 Costa Rica 51 51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 S3 India 14 14 Indonesia 20			183	
Costa Rica       51         51       1         Cote d?Ivoire       1         1       1         Ecuador       3         3       3         El Salvador       21         21       21         Ethiopia       44         44       44         Guatemala       181         181       6         6       6         Honduras       53         53       53         India       14         14       14         Indonesia       20         20       20			103	
51 Cote d?Ivoire 1 1 Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20			5.1	
Cote d?Ivoire       1         1       1         Ecuador       3         3       3         El Salvador       21         21       4         Ethiopia       44         44       44         Guatemala       181         181       6         6       6         Honduras       53         53       53         India       14         14       14         Indonesia       20         20       20			31	
1 Ecuador 3 3 3 3 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5			1	
Ecuador 3 3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20			Í	
3 El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20			3	
El Salvador 21 21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20			3	
21 Ethiopia 44 44 Guatemala 181 181 Haiti 6 Honduras 53 53 India 14 14 Indonesia 20			21	
Ethiopia 44 44 Guatemala 181 181 Haiti 6 Honduras 53 India 14 14 Indonesia 20			2.1	
44 Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20			4.4	
Guatemala 181 181 Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20 20			11	
181 Haiti 6 6 Honduras 53 India 14 14 Indonesia 20			1 8 1	
Haiti 6 6 Honduras 53 53 India 14 14 Indonesia 20			101	
6 Honduras 53 53 India 14 14 Indonesia 20			6	
Honduras 53 53 India 14 14 Indonesia 20 20			0	
53 India 14 14 Indonesia 20 20			53	
India 14 14 Indonesia 20 20			33	
14 Indonesia 20 20			1.4	
Indonesia 20 20			11	
20			20	
			20	
			1	
	1.		-	

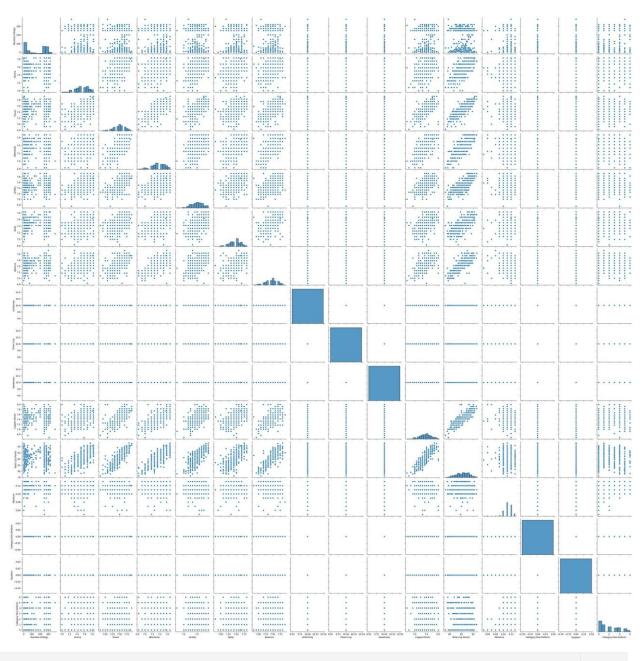
1	
Kenya	25
25	
Laos	3
3	
Malawi	11
11	
Mauritius	1
1	
Mexico	237
237	
Myanmar	8
8	
Nicaragua	26
26	
Panama	4
4	1
Papua New Guinea 1	1
Peru	10
10	10
Philippines	5
5	3
Rwanda	1
1	_
Taiwan	75
75	
Tanzania, United Republic Of	40
40	
Thailand	32
32	
Uganda	36
36	
United States	10
10	
United States (Hawaii)	73
73	
United States (Puerto Rico)	4
4	0
Vietnam 8	8
Zambia	1
1	1
1	
[36 rows x 30 columns]	
<pre>b=df.groupby(['Country.of.Origin']).size().rese ename(columns={'Country.of.Origin':'CO'})</pre>	et_index(name="count").r
b	

```
CO count
0
                           Brazil
                                    132
1
                                     2
                          Burundi
2
                                     16
                            China
3
                         Colombia
                                    183
4
                                     51
                       Costa Rica
5
                                     1
                   Cote d?Ivoire
6
                          Ecuador
7
                     El Salvador
                                     21
8
                                     44
                         Ethiopia
9
                       Guatemala
                                    181
10
                            Haiti
                                      6
11
                        Honduras
                                      53
12
                           India
                                     14
13
                       Indonesia
                                     20
14
                            Japan
                                      1
15
                            Kenya
                                      25
                                      3
16
                            Laos
17
                          Malawi
                                     11
18
                       Mauritius
                                      1
19
                          Mexico
                                     237
20
                         Myanmar
                                      8
21
                       Nicaragua
                                     26
22
                           Panama
                                      4
23
                                      1
                Papua New Guinea
24
                                     10
                             Peru
25
                      Philippines
                                      5
26
                                      1
                           Rwanda
27
                           Taiwan
                                      75
28
    Tanzania, United Republic Of
                                      40
29
                         Thailand
                                     32
30
                           Uganda
                                     36
31
                   United States
                                     10
32
          United States (Hawaii)
                                      73
33
     United States (Puerto Rico)
34
                          Vietnam
                                       8
35
                           Zambia
plt.bar(b['CO'],b['count'])
<BarContainer object of 36 artists>
```



b['count%']=b[ b	'count']/sum(b['cou	nt'])*1	00
	CO	count	count%
0	Brazil	132	9.858103
1	Burundi	2	0.149365
2	China	16	1.194922
3	Colombia	183	13.666916
4	Costa Rica	51	3.808813
5	Cote d?Ivoire	1	0.074683
6	Ecuador	3	0.224048
7	El Salvador	21	1.568335
8	Ethiopia	44	3.286034
9	Guatemala	181	13.517550
10	Haiti	6	0.448096
11	Honduras	53	3.958178
12	India	14	1.045556
13	Indonesia	20	1.493652
14	Japan	1	0.074683
15	Kenya	25	1.867065
16	Laos	3	0.224048
17	Malawi	11	0.821509
18	Mauritius	1	0.074683
19	Mexico	237	17.699776
20	Myanmar	8	0.597461

```
21
                    Nicaragua
                                26 1.941748
22
                      Panama
                                4 0.298730
23
              Papua New Guinea
                                1 0.074683
                                10 0.746826
24
                        Peru
25
                                5 0.373413
                  Philippines
26
                                1 0.074683
                       Rwanda
27
                       Taiwan
                                75 5.601195
28
   Tanzania, United Republic Of
                               40 2.987304
                               32 2.389843
29
                     Thailand
30
                       Uganda
                                36 2.688574
31
                United States
                               10 0.746826
                               73 5.451830
32
        United States (Hawaii)
33
   United States (Puerto Rico)
                                4 0.298730
34
                                8 0.597461
                      Vietnam
35
                      Zambia 1 0.074683
c=sns.pairplot(filter)
С
<seaborn.axisgrid.PairGrid at 0x2306b7a0d50>
```



filt	er				
	Species		Owner	Country.of.Origin	\
148	Arabica	juan luis	alvarado romero	Guatemala	
160	Arabica	yunnan	coffee exchange	China	
196	Arabica	yunnan	coffee exchange	China	
203	Arabica	juan luis	alvarado romero	Guatemala	
206	Arabica	jesus	salazar velasco	Mexico	
1152	Arabica	calixto	guillen vazquez	Mexico	
1160	Arabica	juan luis	alvarado romero	Guatemala	
1166	Arabica	armando luis p	pohlenz martinez	Mexico	

1167 Arabic	-		fe condor s.a Colombia
1182 Arabic	a li	n, che-	hao krude 林哲豪 Taiwan
	Region		
Producer \			
148	jalapa		ROBERTO
MONTERROSO			
160	yunnan		Menglian Ban'an Coffee Processing
Factory			
196	yunnan	Mengl	ian County Nayun Town Mangzhang Coffee
Farm			
203	oriente		AMILCAR
LAPOLA			
206 san pe	dro cotzilnam		JESUS SALAZAR
VELASCO			
	• • •		
1152	chiapas		VARIOS
PRODUCTORES	Сптарав		V111(1 0 0
1160	oriente		AGRIPEC DE LA VEGA Y
CIA.	Offelice		AGRIFEC DE LA VEGA I
1166	la concordia		ARMANDO LUIS POHLENZ
MARTINEZ	ia concolula		AINMANDO DOIS TOTTENZ
	1 ' 7		-
1167	huila		La
Plata			
1182	natou county		WU SHU
YI 巫叔憶			
Numbor	of.Bags Bag.	Woight	In.Country.Partner
Harvest.Year		Neight	In.Country.Farther
148	250.0	60 100	Aggaingian Magianal Dal Cafá
2014	230.0	69 Kg	Asociacion Nacional Del Café
160			
	2 0	60 100	Vunnan Coffee Euchange
	3.0	60 kg	Yunnan Coffee Exchange
2015		-	
196	3.0	60 kg	Yunnan Coffee Exchange Yunnan Coffee Exchange
196 2015	3.0	60 kg	Yunnan Coffee Exchange
196 2015 203		-	Yunnan Coffee Exchange
196 2015 203 2014	3.0	60 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café
196 2015 203 2014 206	3.0	60 kg	Yunnan Coffee Exchange
196 2015 203 2014	3.0	60 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café
196 2015 203 2014 206	3.0	60 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café
196 2015 203 2014 206 2012	3.0 250.0 10.0	60 kg 69 kg 1 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café  AMECAFE
196 2015 203 2014 206 2012 	3.0	60 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café
196 2015 203 2014 206 2012  1152 2013	3.0 250.0 10.0  380.0	60 kg 69 kg 1 kg 1 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café  AMECAFE   AMECAFE
196 2015 203 2014 206 2012  1152 2013 1160	3.0 250.0 10.0	60 kg 69 kg 1 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café  AMECAFE
196 2015 203 2014 206 2012 1152 2013 1160 2013	3.0 250.0 10.0  380.0 250.0	60 kg 69 kg 1 kg 1 kg 69 kg	Yunnan Coffee Exchange  Asociacion Nacional Del Café  AMECAFE   AMECAFE  Asociacion Nacional Del Café
196 2015 203 2014 206 2012 1152 2013 1160 2013 1166	3.0 250.0 10.0  380.0	60 kg 69 kg 1 kg 1 kg	Yunnan Coffee Exchange Asociacion Nacional Del Café  AMECAFE   AMECAFE
196 2015 203 2014 206 2012 1152 2013 1160 2013 1166 2012	3.0 250.0 10.0  380.0 250.0	60 kg 69 kg 1 kg 1 kg 69 kg 1 kg	Yunnan Coffee Exchange  Asociacion Nacional Del Café  AMECAFE   AMECAFE  Asociacion Nacional Del Café  AMECAFE
196 2015 203 2014 206 2012 1152 2013 1160 2013 1166	3.0 250.0 10.0  380.0 250.0	60 kg 69 kg 1 kg 1 kg 69 kg	Yunnan Coffee Exchange  Asociacion Nacional Del Café  AMECAFE   AMECAFE  Asociacion Nacional Del Café

```
4T/10
1182
                 50.0 20 kg Specialty Coffee Association
2014
              Grading.Date ... Total.Cup.Points Moisture \
            May 29th, 2014 ...
                                             84.58
                                                         0.11
148
           April 6th, 2016
                                             84.50
                                                         0.10
160
                            . . .
           April 6th, 2016 ...
196
                                             84.25
                                                         0.10
         August 21st, 2014
203
                                             84.25
                                                         0.10
                             . . .
206
      September 11th, 2012
                                             84.25
                                                         0.12
                             . . .
                                              . . .
. . .
                        . . .
                              . . .
                                                          . . .
1152
         March 29th, 2013
                                             79.75
                                                         0.12
                            . . .
1160
           June 27th, 2013 ...
                                             79.67
                                                        0.11
        August 30th, 2012
                                             79.58
1166
                                                         0.13
                             . . .
1167
       February 9th, 2011
                             . . .
                                             79.58
                                                         0.10
     November 7th, 2014
                                            79.25
1182
                                                    0.11
      Category.One.Defects
                             Quakers Category. Two. Defects \
148
                          0
                                  0.0
                                                           1
                          0
                                 0.0
                                                           1
160
                          0
                                 0.0
                                                           1
196
203
                          0
                                 0.0
                                                           1
                          0
206
                                 0.0
                                                           1
. . .
                                  . . .
                                                         . . .
1152
                          0
                                 0.0
                                                           5
1160
                          0
                                 0.0
                                                           6
                                                           5
1166
                          0
                                 0.0
1167
                          0
                                 0.0
                                                           4
1182
                                  0.0
                Expiration
                                        Certification.Body \
148
            May 29th, 2015 Asociacion Nacional Del Café
           April 6th, 2017
April 6th, 2017
160
                                  Yunnan Coffee Exchange
           April 6th, 2017
196
                                   Yunnan Coffee Exchange
203
         August 21st, 2015 Asociacion Nacional Del Café
206
      September 11th, 2013
                                                    AMECAFE
1152
          March 29th, 2014
                                                    AMECAFE
1160
           June 27th, 2014 Asociacion Nacional Del Café
1166
         August 30th, 2013
                                                   AMECAFE
        February 9th, 2012
1167
                                                   Almacafé
        November 7th, 2015 Specialty Coffee Association
1182
                          Certification.Address \
      b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
148
      683fa6014608fc10ff681b0435b0b2dbe6df988f
160
196
      683fa6014608fc10ff681b0435b0b2dbe6df988f
203
      b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
206
     59e396ad6e22a1c22b248f958e1da2bd8af85272
. . .
```

```
1152 59e396ad6e22a1c22b248f958e1da2bd8af85272
1160 b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
1166 59e396ad6e22a1c22b248f958e1da2bd8af85272
1167 e493c36c2d076bf273064f7ac23ad562af257a25
1182 36d0d00a3724338ba7937c52a378d085f2172daa
                                            Certification.Contact unit of measurement
148
          724f04ad10ed31dbb9d260f0dfd221ba48be8a95
160
          f6d87a6c04653c569d4911a66f89d5e30ce83b93
                                                                                                                     m
          f6d87a6c04653c569d4911a66f89d5e30ce83b93
196
                                                                                                                    m
203
          724f04ad10ed31dbb9d260f0dfd221ba48be8a95
                                                                                                                   ft
          0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
206
                                                                                                                    m
                                                                                                                  . . .
          0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
1152
                                                                                                                    m
1160 724f04ad10ed31dbb9d260f0dfd221ba48be8a95
                                                                                                                   ft
          0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
1166
                                                                                                                     m
1167 70d3c0c26f89e00fdae6fb39ff54f0d2eb1c38ab
1182 0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
[404 rows x 31 columns]
filter.info()
<class 'pandas.core.frame.DataFrame'>
Index: 404 entries, 148 to 1182
Data columns (total 31 columns):
 # Column Non-Null Count Dtype
O Species 404 non-null object
Country.of.Origin 404 non-null object
Region 404 non-null object
Producer 404 non-null object
Number.of.Bags 404 non-null float64
Bag.Weight 404 non-null object
In.Country.Partner 404 non-null object
Harvest.Year 404 non-null object
Grading.Date 404 non-null object
Owner.1 404 non-null object
Aroma 404 non-null float64
Flavor 404 non-null float64
 12 Flavor
                                               404 non-null float64

      12
      Flavor
      404 non-null
      float64

      13
      Aftertaste
      404 non-null
      float64

      14
      Acidity
      404 non-null
      float64

      15
      Body
      404 non-null
      float64

      16
      Balance
      404 non-null
      float64

      17
      Uniformity
      404 non-null
      float64

      18
      Clean.Cup
      404 non-null
      float64

      19
      Sweetness
      404 non-null
      float64

      20
      Cupper.Points
      404 non-null
      float64

      21
      Total.Cup.Points
      404 non-null
      float64
```

```
22 Moisture
                             404 non-null
                                             float64
 23 Category.One.Defects
                             404 non-null
                                             int64
 24 Quakers
                             404 non-null
                                             float64
 25 Category. Two. Defects
                             404 non-null
                                             int64
 26 Expiration
                             404 non-null
                                             object
    Certification.Body
                             404 non-null
                                             object
 27
 28 Certification.Address 404 non-null
                                             object
29 Certification.Contact
                            404 non-null
                                             object
30 unit of measurement
                             404 non-null
                                             object
dtypes: float64(14), int64(2), object(15)
memory usage: 101.0+ KB
s=filter.select dtypes(exclude=['object'])
S
      Number.of.Bags Aroma
                             Flavor Aftertaste Acidity Body
Balance \
148
                       7.75
                                7.83
                                            7.67
               250.0
                                                     7.92 7.75
7.83
160
                                7.92
                 3.0
                       7.75
                                            7.67
                                                      7.83 7.75
7.67
196
                 3.0
                       7.83
                                7.83
                                            7.58
                                                      7.75 7.75
7.75
203
               250.0
                       7.83
                                7.83
                                            7.50
                                                      7.92 7.58
7.75
206
                10.0
                       7.83
                                7.92
                                            7.67
                                                      7.83
                                                            7.50
7.75
. . .
                  . . .
                         . . .
                                 . . .
                                              . . .
1152
               380.0
                       7.25
                                7.08
                                            7.00
                                                      7.17 7.17
7.08
1160
               250.0
                       7.50
                                7.00
                                            6.83
                                                      7.08 7.08
7.17
1166
               250.0
                       7.25
                                6.92
                                            6.92
                                                      7.33 7.25
7.00
1167
               250.0
                       7.25
                                7.17
                                            7.00
                                                      6.75 7.17
7.33
1182
                50.0
                       7.08
                                6.83
                                            6.83
                                                      7.25 7.42
7.08
      Uniformity Clean.Cup
                             Sweetness Cupper.Points
Total.Cup.Points
                       10.0
                                   10.0
148
            10.0
                                                  7.83
84.58
160
            10.0
                       10.0
                                   10.0
                                                   7.92
84.50
196
            10.0
                       10.0
                                   10.0
                                                   7.75
84.25
203
            10.0
                       10.0
                                   10.0
                                                   7.83
```

84.25							
206	10.0	10.0	10.	0	7.75		
84.25							
	• • •	• • •	• •	•	• • •		• •
· 1152	10.0	10.0	10.	0	7.00		
79.75	2000	2000			, • • •		
1160	10.0	10.0	10.	0	7.00		
79.67							
1166	10.0	10.0	10.	0	6.92		
79.58	10.0	10.0	10.	Ω	6.92		
79.58	10.0	10.0	10.	O	0.32		
1182	10.0	10.0	10.	0	6.75		
79.25							
	Moisture	'ategory.One	Dofoata	)uakers	Category.Tw	o Dofosts	
148	0.11	acegory.one	0	0.0	category.iw	1	
160	0.10		0	0.0		1	
196	0.10		0	0.0		1	
203	0.10		0	0.0		1	
206	0.12		0	0.0		1	
1152	0.12		0	0.0		5	
1160	0.11		0	0.0		6	
1166	0.13		0	0.0		5	
1167	0.10		0	0.0		4	
1182	0.11		0	0.0		0	

## [404 rows x 16 columns]

new=s.drop(columns=['Uniformity','Clean.Cup','Sweetness','Category.One
.Defects','Quakers'])

## new

Baland	Number.of.Bags	Aroma	Flavor	Aftertaste	Acidity	Body	
148	250.0	7.75	7.83	7.67	7.92	7.75	
7.83 160	3.0	7.75	7.92	7.67	7.83	7.75	
7.67 196	3.0	7.83	7.83	7.58	7.75	7.75	
7.75 203	250.0	7.83	7.83	7.50	7.92	7.58	
7.75 206	10.0	7.83	7.92	7.67	7.83	7.50	
7.75	2000	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	7,52	. • • •	, • 00	,,,,,	
	•••	• • •	• • •	•••	•••	• • •	••

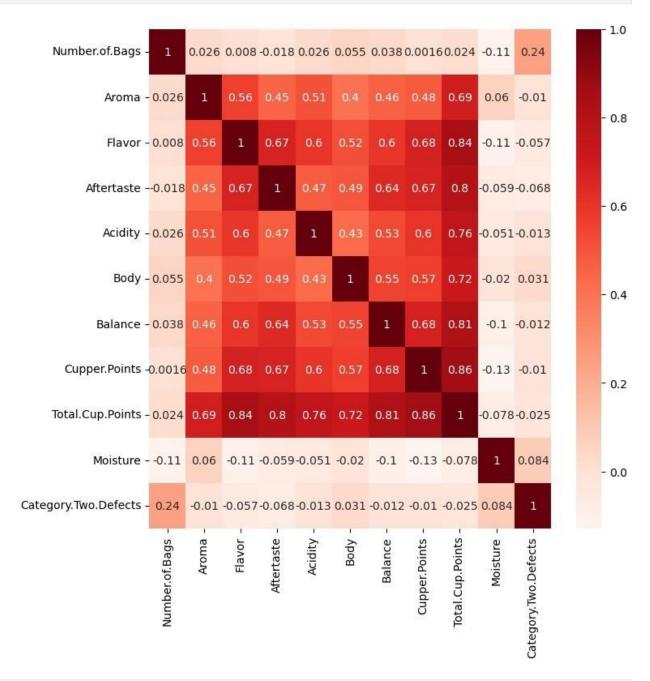
1152 7.08	;	380.0	7.25	7.08	7.00	7.17	7.17	
1160	:	250.0	7.50	7.00	6.83	7.08	7.08	
7.17 1166	;	250.0	7.25	6.92	6.92	7.33	7.25	
7.00 1167	:	250.0	7.25	7.17	7.00	6.75	7.17	
7.33 1182		50.0	7.08	6.83	6.83	3 7.25	7.42	
7.08		30.0	7.00	0.03	0.00	7.23	7 • 12	
	Cupper.Po	ints I	otal.Cup	.Points	Moisture	e Categor	y.Two.Defect	ts
148		7.83		84.58	0.11	L		1
160		7.92		84.50	0.10	)		1
196		7.75		84.25	0.10	)		1
203		7.83		84.25	0.10	)		1
206		7.75		84.25	0.12	2		1
1152		7.00		79.75	0.12	2		5
1160		7.00		79.67	0.11			6
1166		6.92		79.58	0.13			5
1167		6.92		79.58	0.10			4
1182		6.75		79.25	0.11	L		0
[404	rows x 11 d	columns]						
g=new	.corr()							
g								
		1	Number.ot	f.Bags	Aroma	Flavor		
After	taste \		1 (	000000	0.025602	0.007954	-0.018019	9

	Number.of.Bags	Aroma	Flavor	
Aftertaste \				
Number.of.Bags	1.000000	0.025602	0.007954	-0.018019
Aroma	0.025602	1.000000	0.559678	0.449410
Flavor	0.007954	0.559678	1.000000	0.670064
Aftertaste	-0.018019	0.449410	0.670064	1.000000

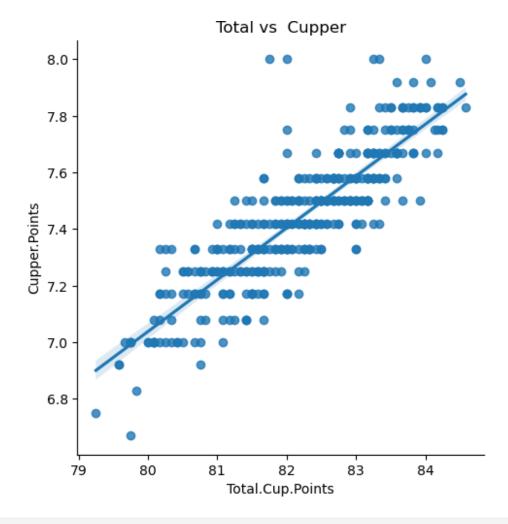
Acidity	0.026252 0.508095 0.597453 0.473935
Body	0.054575 0.404418 0.518042 0.494965
Balance	0.037832 0.457744 0.597673 0.643990
Cupper.Points	0.001627 0.483042 0.678251 0.672963
Total.Cup.Points	0.023737 0.694559 0.841711 0.800253
Moisture	-0.114628 0.060143 -0.110836 -0.058805
Category.Two.Defects	0.241994 -0.010320 -0.057406 -0.067737
Number.of.Bags Aroma Flavor Aftertaste Acidity Body Balance Cupper.Points Total.Cup.Points Moisture Category.Two.Defects	Acidity Body Balance Cupper.Points \ 0.026252 0.054575 0.037832 0.001627  0.508095 0.404418 0.457744 0.483042  0.597453 0.518042 0.597673 0.678251  0.473935 0.494965 0.643990 0.672963  1.000000 0.431081 0.531185 0.604851  0.431081 1.000000 0.550884 0.565472  0.531185 0.550884 1.000000 0.677059  0.604851 0.565472 0.677059 1.000000  0.756864 0.717805 0.814135 0.863134  -0.050692 -0.019634 -0.099618 -0.128718  -0.013021 0.030588 -0.012185 -0.010187
	Total.Cup.Points Moisture Category.Two.Defect
Number.of.Bags	0.023737 -0.114628 0.24199
Aroma	0.694559 0.060143 -0.01032
Flavor	0.841711 -0.110836 -0.05740
Aftertaste	0.800253 -0.058805 -0.06773
Acidity	0.756864 -0.050692 -0.01302
Body	0.717805 -0.019634 0.03058
Balance	0.814135 -0.099618 -0.01218
Cupper.Points	0.863134 -0.128718 -0.01018
Total.Cup.Points	1.000000 -0.078231 -0.02522
Moisture	-0.078231 1.000000 0.08435

```
Category.Two.Defects -0.025220 0.084359 1.000000

plt.figure(figsize=(8,8))
sns.heatmap(g,annot=True,cmap='Reds')
plt.show()
```

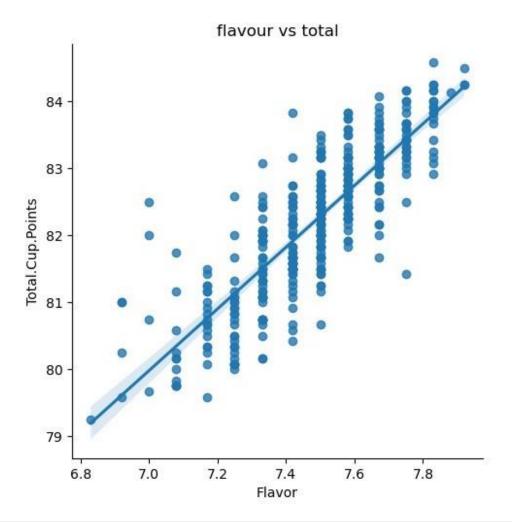


sns.lmplot(x='Total.Cup.Points', y='Cupper.Points', data=new)
plt.title("Total vs Cupper")



sns.lmplot(x='Flavor', y='Total.Cup.Points', data=new)
plt.title("flavour vs total")

Text(0.5, 1.0, 'flavour vs total')



```
X = filter[['Flavor', 'Acidity', 'Aftertaste', 'Cupper.Points',
'Aroma', 'Body']]
Y = filter['Total.Cup.Points']
Χ
               Acidity Aftertaste Cupper.Points
      Flavor
                                                       Aroma
                                                              Body
148
         7.83
                  7.92
                                7.67
                                                7.83
                                                        7.75
                                                              7.75
                                                        7.75
160
         7.92
                  7.83
                                7.67
                                                7.92
                                                              7.75
         7.83
                  7.75
                                7.58
                                                7.75
                                                        7.83
196
                                                              7.75
         7.83
                  7.92
                                                7.83
                                                        7.83
203
                                7.50
                                                              7.58
206
         7.92
                  7.83
                                7.67
                                                7.75
                                                        7.83
                                                              7.50
                                . . .
. . .
         . . .
                                                 . . .
                                                        7.25
1152
         7.08
                  7.17
                                7.00
                                                7.00
                                                              7.17
1160
        7.00
                  7.08
                                6.83
                                                7.00
                                                        7.50
                                                              7.08
                                                6.92
                                                        7.25
1166
         6.92
                  7.33
                                6.92
                                                              7.25
                                7.00
                                                6.92
                                                        7.25
1167
         7.17
                   6.75
                                                              7.17
1182
         6.83
                  7.25
                                6.83
                                                6.75
                                                        7.08
                                                              7.42
[404 rows x 6 columns]
```

```
Y
148
       84.58
160
        84.50
196
        84.25
203
        84.25
206
        84.25
        . . .
1152
        79.75
1160 79.67
1166
       79.58
1167
       79.58
1182
       79.25
Name: Total.Cup.Points, Length: 404, dtype: float64
from sklearn.feature selection import f classif
u=f classif(X,Y)
(array([17.96140216, 9.76119633, 12.8141253, 21.34390575,
7.47525214,
         8.51363175]),
array([1.01246262e-74, 1.89407214e-45, 1.15813405e-57, 3.64833219e-
84,
        6.23268531e-35, 7.13981932e-40]))
filter.groupby(['Flavor','Total.Cup.Points']).count()
                         Species Owner Country.of.Origin Region
Producer \
Flavor Total.Cup.Points
6.83 79.25
1
6.92 79.58
       80.25
                                                                 1
1
       81.00
7.00 79.67
                                                                 1
1
. . .
7.83
       84.25
                                      2
                                                                 2
       84.58
7.88 84.13
1
```

7.92	84.25	2	2		2	2
2	84.50	1	1		1	1
1	01.00	_	_		_	_
		Number of Da	aa Baa	Wojah+		
In.Cou	ntry.Partner \	Number.of.Ba	ys bay	·weight		
	Total.Cup.Points					
6.83	79.25		1	1		
1			_			
6.92 1	79.58		1	1		
	80.25		1	1		
1	81.00		2	2		
2	01.00		2	۷		
7.00	79.67		1	1		
1						
	0.4.05		•	0		
7.83 2	84.25		2	2		
	84.58		1	1		
1 7.88	84.13		1	1		
1						
7.92 2	84.25		2	2		
۷	84.50		1	1		
1						
		Harvest.Year	Gradin	ng.Date .		
	.Points \					
Flavor	Total.Cup.Points				• • •	
6.83	79.25	1		1		
1 6.92	79.58	1		1		
1						
1	80.25	1		1	• • •	
	81.00	2		2		
2 7.00	79.67	1		1		
1	19.01	1		Τ	•••	
7.83	84.25	2		2		
2						

1	84.58		1	1		
1 7.88	84.13		1	1		
1						
7.92 2	84.25		2	2		
۷	84.50		1	1		
1	04.50		1	Τ.	• • •	
_						
		Moisture	Category.On	e.Defects	s Quakers \	
Flavor	Total.Cup.Points		J _			
6.83	79.25	1			1	1
6.92	79.58	1			1	1
	80.25	1			1	1
	81.00	2			2	2
7.00	79.67	1			1	1
				•		
7.83	84.25	2			2 2	2
	84.58	1				1
	84.13	1				1
7.92		2				2
	84.50	1			1	1
		~ .				
~	' · · · · · · · · · · · · · · · · · · ·	Category	.Two.Defects	Expirati	ion	
	ication.Body \					
Flavor	Total.Cup.Points					
6.83	79.25		1		1	
1						
6.92	79.58		1		1	
1						
	80.25		1		1	
1						
	81.00		2		2	
2						
	79.67		1		1	
1						
• • •						
	0.4.05					
	84.25		2		2	
2	0.4 5.0		1		1	
1	84.58		1		1	
1 7.88	0112		1		1	
1	84.13		1		1	
7.92	84.25		2		2	
2	U1.4J		۷.		۷	
۷	84.50		1		1	
1	01.50		1		_	
_						

<b>\</b>		Certification.Addres	SS	Certifi	cation.Con	tact
\ Flavor	Total.Cup.Points					
6.83	79.25		1			1
6.92	79.58		1			1
	80.25		1			1
	81.00		2			2
7.00	79.67		1			1
7.83	84.25		2			2
	84.58		1			1
7.88	84.13		1			1
7.92	84.25		2			2
	84.50		1			1
6.83 6.92 7.00  7.83 7.88 7.92		unit_of_measurement  1 1 1 2 1 2 1 1 1 2 1 ,'Acidity']).count()		Danian	Dung dung an A	
Flavor	Species Acidity	Owner Country.of.Oric	gin	Region	Producer \	
6.83 6.92	7.25 1 7.25 1 7.33 3 7.08 1	1 1 3 1		1 1 3 1	1 1 3 1	1 1 3 1
7.00	7.00	1		_	1	_

7.83 7.88 7.92	7.25 7.83 7.92 7.75 7.75	1 6 3 1	1  6 3 1		1  6 3 1	1 6 3 1	1 6 3 1 1
	7.83 t.Year \ Acidity	2 Number.of.Bac	2 gs Bag.We	eight In.Co	2 ountry.Pa	2 artner	2
6.83 1 6.92	7.25 7.25		1	1		1	
1	7.33		3	3		3	
3 7.00 1	7.08		1	1		1	
1	7.25		1	1		1	
7.83	7.83		6	6		6	
6	7.92		3	3		3	
3 7.88 1	7.75		1	1		1	
7.92 1	7.75 7.83		2	1		1	
2	7.03		2	۷		Ζ	
		Grading.Date	Tota	l.Cup.Poir	nts Moist	ture \	
	Acidity						
6.83	7.25	1	• • •		1	1	
6.92	7.25	1	• • •		1	1	
	7.33	3	• • •		3	3	
7.00	7.08 7.25	1	• • •		1	1	
7.83	7.83 7.92	6			6 3	6 3	
7.88 7.92	7.75 7.75 7.83	1 1 2			1 1 2	1 1 2	
		Category.One	.Defects	Quakers	Category	.Two.Defect	s \

6.83 6.92 7.00  7.83	Acidity 7.25 7.25 7.33 7.08 7.25 7.83 7.92 7.75 7.75 7.83		1 1 3 1 1  6 3 1 1 2	1 1 3 1 1  6 3 1 1 2	1 1 3 1 1  6 3 1 1 2
\	2 1 21	Expiration	Certification.B	Body Certifi	cation.Address
	Acidity	1		1	1
6.83	7.25	1		1	1
6.92	7.25	1		1	1
	7.33	3		3	3
7.00	7.08	1		1	1
	7.25	1		1	1
7.83	7.83	6		6	6
	7.92	3		3	3
7.88	7.75	1		1	1
7.92	7.75	1		1	1
	7.83	2		2	2
Flavor	Acidity	Certification	on.Contact unit_	of_measureme	nt
6.83 6.92 7.00	7.25 7.25 7.33 7.08 7.25		1 1 3 1		1 1 3 1
7.83 7.88	7.83 7.92 7.75		6 3 1		6 3 1

7.92	7.75 7.83			1 2		1 2	
[89 ro	ws x 29 colu	mns]					
filter	<pre>filter.groupby(['Flavor','Aftertaste']).count()</pre>						
\		Species	Owner	Country.of.	Origin	Region	Producer
Flavor	Aftertaste						
6.83	6.83	1	1		1	1	1
6.92	6.92	1	1		1	1	1
	7.00	1	1		1	1	1
	7.17	1	1		1	1	1
	7.33	1	1		1	1	1
			• • •		• • •		• • •
7.83	7.50	4	4		4	4	4
	7.58	6	6		6	6	6
	7.67	3	3		3	3	3
7.88	7.56	1	1		1	1	1
7.92	7.67	3	3		3	3	3
		Number	f Daga	Bag.Weight I	n Countr	Danta	o. 70
Flavor	Aftertaste		_			y.Partin	
6.83 6.92	6.83 6.92		1 1	1 1			1 1
0.52	7.00		1	1			1
	7.17 7.33		1 1	1 1			1 1
7.83	7.50 7.58		4 6	4			4 6
	7.67		3	3			3
7.88 7.92	7.56 7.67		1 3	1 3			1 3
Harvest.Year Grading.Date							
Total.Cup.Points \							
Flavor	Aftertaste				• • •		

6.83	6.83		1	1			1
6.92	6.92		1	1			1
	7.00		1	1			1
	7.17		1	1			1
	7.33		1	1			1
7.83	7.50		4	4			4
	7.58		6	6			6
	7.67		3	3			3
7.88	7.56		1	1			1
7.92	7.67		3	3			3
Flavor 6.83 6.92  7.83	Aftertaste 6.83 6.92 7.00 7.17 7.33 7.50 7.58 7.67 7.56 7.67	Moisture Ca 1 1 1 1 1  4 6 3 1 3			1 1 1 1 1  4 6 3 1 3	1 1 1 1 1 1 1 4 6 3 1 3	
~		Category.T	wo.Defect	s Expirat	cion		
	ication.Body Aftertaste	\					
6.83 1	6.83			1	1		
6.92	6.92			1	1		
1	7.00			1	1		
1	7.17			1	1		
1	7.33			1	1		

1				
		• • •	• • •	
7.83 4	7.50	4	4	
6	7.58	6	6	
3	7.67	3	3	
7.88 1	7.56	1	1	
7.92	7.67	3	3	
773		ldress Certif	Fication.Contact \	
6.83	Aftertaste 6.83 6.92 7.00 7.17 7.33	1 1 1 1	1 1 1 1	
	7.50 7.58 7.67 7.56 7.67	4 6 3 1 3	4 6 3 1 3	
	unit_of_measurer	ment		
6.83	Aftertaste 6.83 6.92 7.00 7.17 7.33	1 1 1 1		
7.83	7.50 7.58 7.67	4 6 3		
7.88 7.92	7.56 7.67	1 3		
[77 row	ws x 29 columns]			
			#Anova	
filter	['Flavor'].unique()			

```
array([12, 14, 11, 13, 10, 7, 9, 8, 6, 5, 2, 3, 4, 1, 0],
dtype=int64)
from sklearn.preprocessing import LabelEncoder
label=LabelEncoder()
filter['Flavor']=label.fit transform(filter['Flavor'])
C:\Users\user\AppData\Local\Temp\ipykernel 5156\2628298509.py:3:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row indexer,col indexer] = value instead
See the caveats in the documentation:
https://pandas.pydata.org/pandas-docs/stable/user guide/indexing.html#
returning-a-view-versus-a-copy
 filter['Flavor']=label.fit transform(filter['Flavor'])
     Species
                                      Owner Country.of.Origin \
148
     Arabica
                   juan luis alvarado romero
                                                  Guatemala
160 Arabica
                      yunnan coffee exchange
                                                       China
196 Arabica
                                                        China
                      yunnan coffee exchange
203 Arabica
                   juan luis alvarado romero
                                                  Guatemala
206 Arabica
                       jesus salazar velasco
                                                      Mexico
1152 Arabica
                     calixto quillen vazquez
                                                      Mexico
1160 Arabica
                   juan luis alvarado romero
                                                  Guatemala
1166 Arabica armando luis pohlenz martinez
                                                     Mexico
1167 Arabica exportadora de cafe condor s.a
                                                     Colombia
1182 Arabica lin, che-hao krude 林哲豪 Taiwan
```

MARTIN				
1167	hı	ıila		La
Plata				
1182	natou cou	unty		WU SHU
YI 巫赤	双憶			
	Number.of.Bags H	Bag.Weight	In.Country.Partner	
Harves	st.Year \	- 9 - 9 -		
148	250.0	69 kg	Asociacion Nacional Del Café	
2014	200.0	03 119	noonanna ber dare	
160	3.0	60 kg	Yunnan Coffee Exchange	
2015	J. U	oo ng	ramian correct Exemange	
196	3.0	60 kg	Yunnan Coffoo Eyghango	
2015	3.0	ou kg	Yunnan Coffee Exchange	
	0.5.00	CO 1	7 ' ' 7 ' 1 5 1 6 6 6 6	
203	250.0	69 kg	Asociacion Nacional Del Café	
2014	10.0	4 1		
206	10.0	1 kg	AMECAFE	
2012				
	• • •		• • •	
1152	380.0	1 kg	AMECAFE	
2013				
1160	250.0	69 kg	Asociacion Nacional Del Café	
2013				
1166	250.0	1 kg	AMECAFE	
2012		2		
1167	250.0	70 kg	Almacafé	
4T/10				
1182	50.0	20 kg	Specialty Coffee Association	
2014		. ,		
	Grading	.Date	Total.Cup.Points Moisture \	
148	May 29th,		84.58 0.11	
160	April 6th,		84.50 0.10	
196	April 6th,		84.25 0.10	
203	August 21st,		84.25 0.10	
206	September 11th,		84.25 0.12	
	september IItii,			
1152	March 29th,	2012	79.75 0.12	
1160	June 27th,		79.67 0.11	
	·			
1166	August 30th,		79.58 0.13	
1167	February 9th,		79.58 0.10	
1182	November 7th,	2014	79.25 0.11	
		c		
	Category.One.De:		kers Category.Two.Defects \	
148		0	0.0	
160		0	0.0	
196		0	0.0	
203		0	0.0	
206		0	0.0	

```
. . .
1152
                         0
                                0.0
                                                         5
1160
                         0
                                0.0
                                                         6
1166
                         0
                                0.0
                                                         5
                                                         4
1167
                         0
                                0.0
1182
                                0.0
                Expiration
                                      Certification.Body \
            May 29th, 2015 Asociacion Nacional Del Café
148
160
           April 6th, 2017 Yunnan Coffee Exchange
196
           April 6th, 2017
                                 Yunnan Coffee Exchange
203
         August 21st, 2015 Asociacion Nacional Del Café
206
      September 11th, 2013
                                                 AMECAFE
. . .
         March 29th, 2014
1152
                                                 AMECAFE
1160
           June 27th, 2014
                           Asociacion Nacional Del Café
1166
         August 30th, 2013
                                                 AMECAFE
1167
        February 9th, 2012
                                                Almacafé
1182
        November 7th, 2015 Specialty Coffee Association
                         Certification.Address
    b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
148
      683fa6014608fc10ff681b0435b0b2dbe6df988f
160
196
      683fa6014608fc10ff681b0435b0b2dbe6df988f
      b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
203
206
      59e396ad6e22a1c22b248f958e1da2bd8af85272
. . .
1152
      59e396ad6e22a1c22b248f958e1da2bd8af85272
1160 b1f20fe3a819fd6b2ee0eb8fdc3da256604f1e53
1166
      59e396ad6e22a1c22b248f958e1da2bd8af85272
      e493c36c2d076bf273064f7ac23ad562af257a25
1167
1182 36d0d00a3724338ba7937c52a378d085f2172daa
                         Certification.Contact
                                                unit of measurement
      724f04ad10ed31dbb9d260f0dfd221ba48be8a95
148
                                                                  ft
160
      f6d87a6c04653c569d4911a66f89d5e30ce83b93
                                                                   m
196
      f6d87a6c04653c569d4911a66f89d5e30ce83b93
                                                                   m
203
      724f04ad10ed31dbb9d260f0dfd221ba48be8a95
                                                                  ft
206
      0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
                                                                   m
1152
      0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
                                                                   m
      724f04ad10ed31dbb9d260f0dfd221ba48be8a95
1160
                                                                  ft
      0eb4ee5b3f47b20b049548a2fd1e7d4a2b70d0a7
1166
                                                                   m
      70d3c0c26f89e00fdae6fb39ff54f0d2eb1c38ab
1167
                                                                   m
      0878a7d4b9d35ddbf0fe2ce69a2062cceb45a660
1182
[404 rows x 31 columns]
df1=filter[['Number.of.Bags','Aroma','Aftertaste','Acidity','Flavor']]
df1
```

```
Number.of.Bags Aroma Aftertaste Acidity Flavor
                               7.67
148
             250.0 7.75
                                       7.92
                                                 12
160
               3.0 7.75
                               7.67
                                       7.83
                                                14
                   7.83
                                       7.75
                                                 12
196
               3.0
                               7.58
             250.0 7.83
203
                               7.50
                                       7.92
                                                12
             10.0 7.83
206
                               7.67
                                       7.83
                                                14
              . . .
                    . . .
                               . . .
                                       . . .
                                                . . .
             380.0 7.25
1152
                               7.00
                                       7.17
                                                 3
             250.0 7.50
                                       7.08
1160
                               6.83
                                                 2
1166
             250.0 7.25
                               6.92
                                       7.33
                                                 1
                   7.25
                                                 4
1167
             250.0
                               7.00
                                       6.75
1182
                                                0
             50.0 7.08
                               6.83
                                       7.25
```

## [404 rows x 5 columns]

```
u=df1[['Number.of.Bags','Aroma','Aftertaste','Acidity']]
v=df1['Flavor']
```

u

	Number.of.Bags	Aroma	Aftertaste	Acidity
148	250.0	7.75	7.67	7.92
160	3.0	7.75	7.67	7.83
196	3.0	7.83	7.58	7.75
203	250.0	7.83	7.50	7.92
206	10.0	7.83	7.67	7.83
1152	380.0	7.25	7.00	7.17
1160	250.0	7.50	6.83	7.08
1166	250.0	7.25	6.92	7.33
1167	250.0	7.25	7.00	6.75
1182	50.0	7.08	6.83	7.25

## [404 rows x 4 columns]

V

Name: Flavor, Length: 404, dtype: int64

```
from sklearn.feature selection import f classif, SelectKBest
p values=f classif(u,v)
p values
(array([ 0.85242197, 14.52699572, 26.60334715, 18.27662554]),
array([6.11543311e-01, 4.34008194e-28, 2.18511309e-48, 6.86968975e-
35]))
selector=SelectKBest(score func=f classif, k=4)
new=selector.fit transform(u,v)
array([[250. , 7.75, 7.67, 7.92],
      [ 3. , 7.75, 7.67, 7.83],
      [ 3. , 7.83, 7.58, 7.75],
      . . . ,
    [250. , 7.25, 6.92, 7.33], [250. , 7.25, 7. , 6.75],
   [ 50. , 7.08, 6.83, 7.25]])
p value=pd.Series(p values[1])
p_value.index=u.columns
p_value
Number.of.Bags 6.115433e-01
Aroma
                 4.340082e-28
Aftertaste 2.185113e-48
Acidity 6.869690e-35
dtype: float64
#after taste only coffee has more taste
```