CAPSTONE PROJECT POWER BI

Name :- Saiyyam Sushil Jain

ID :- S10159

Introducing The Art of Brewing

List of Contents

Project Overview

This project leverages the Coffee Quality Institute (CQI) dataset to analyze factors that contribute to coffee quality, focusing on sensory evaluation, defects, and processing methods.

The goal is to understand the key determinants of coffee quality and their relationships with processing methods, origin regions, and defects.

The analysis aims to improve decision-making in coffee production and processing to enhance overall coffee quality.

Business Problem

Inconsistent coffee quality due to varying sensory attributes, processing methods, and defects challenges the coffee industry. Understanding these factors is crucial for improving quality control and meeting consumer expectations.

Objective

Sensory attributes Analysis
Defects Analysis
Analysis on Processing Methods and Origin of Regions

Sensory Attributes
Aroma
Flavor
Country of Origin
Aftertaste
Lot Number
Acidity
Altitude
Body
Region
Balance
Number of Bags
Uniformity
Bag Weight
Clean Cup
In-Country Partner
Sweetness
Harvest Year
Grading Date
Variety
Status
Processing Method
Overall
Defects
Total Cup Points
Moisture Percentage
Category One Defects
Quakers
Color
Category Two Defects
Expiration

Dataset Overview

The dataset consists of Coffee Quality Institute (CQI) data consisting of 31 rows and 208 columns.

Approach Used

- Data Cleaning and Preprocessing
- Exploratory Data Analysis
- (EDA)Visualization

Tools & Methods

- Tools: Power BI Desktop, Excel (for preparation).

Overview of Dataset

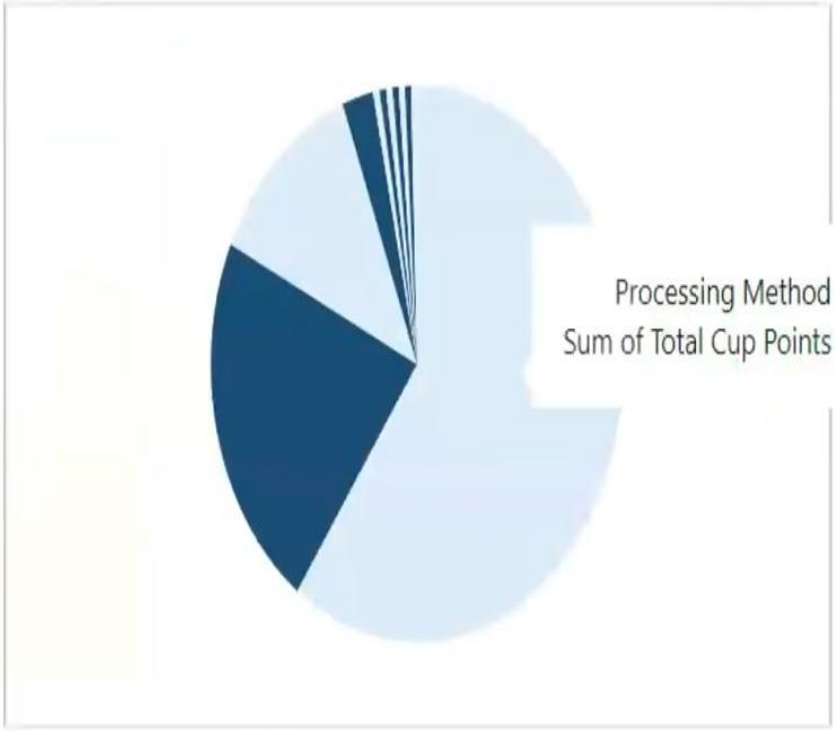
207
Count of ID

22
Count of Country of Origin

115
Count of Region

49
Count of Variety

11
Count of Processing Met...

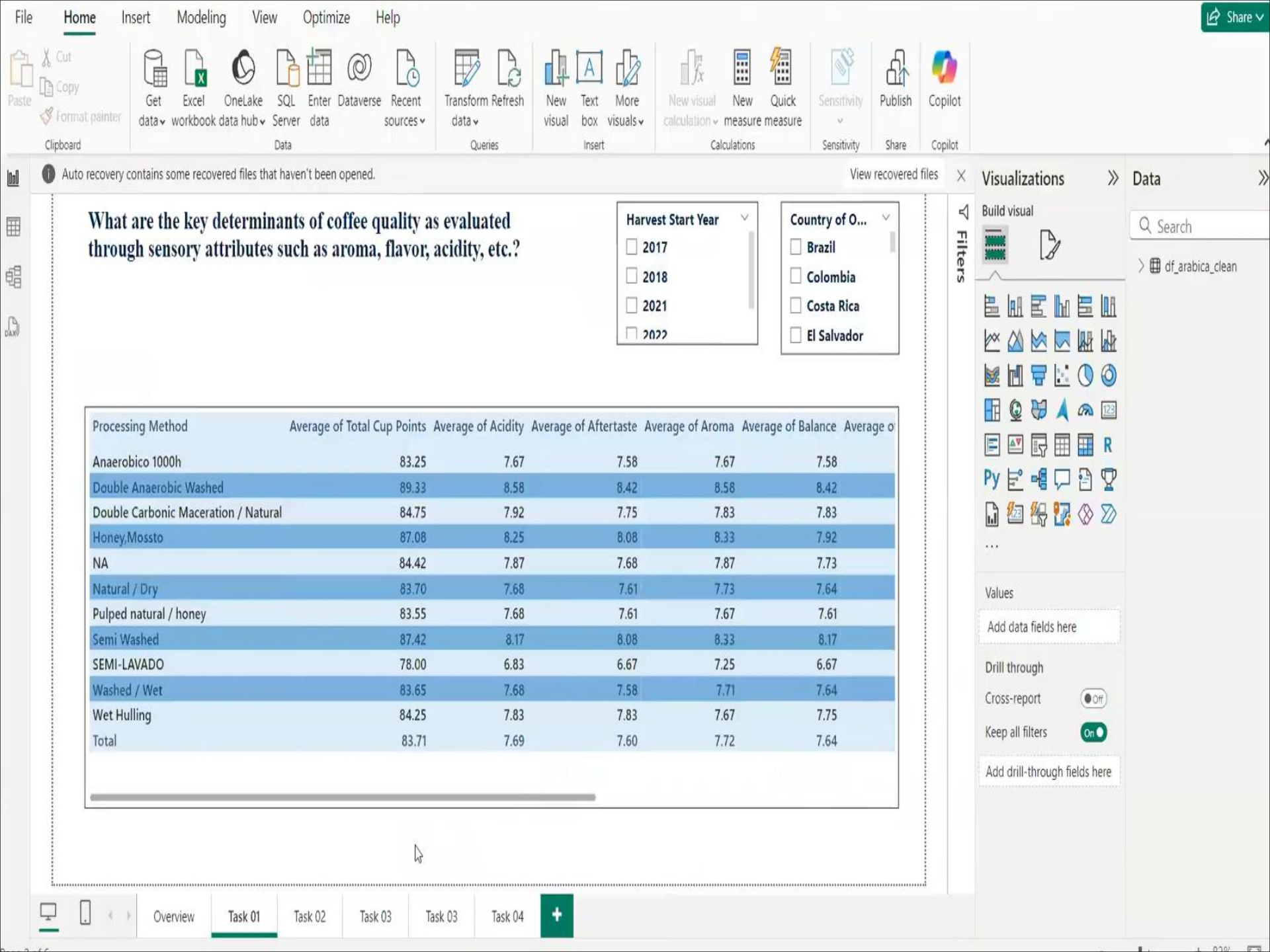


=> Types of Defects

Category 1 Defect
Category 2 Defect

=> Sensory Attributes

Acidity
ftertaste
roma
alance
Body
Clean Cup
Flavor
Sweetness
Uniformity



Processing Method	Average of Total Cup Points	Average of Acidity	Average of Aftertaste	Average of Aroma	Average of Balance	Average o
Anaerobico 1000h	83.25	7.67	7.58	7.67	7.58	
Double Anaerobic Washed	89.33	8.58	8.42	8.58	8.42	
Double Carbonic Maceration / Natural	84.75	7.92	7.75	7.83	7.83	
Honey,Mossto	87.08	8.25	8.08	8.33	7.92	
NA	84.42	7.87	7.68	7.87	7.73	
Natural / Dry	83.70	7.68	7.61	7.73	7.64	
Pulped natural / honey	83.55	7.68	7.61	7.67	7.61	
Semi Washed	87.42	8.17	8.08	8.33	8.17	
SEMI-LAVADO	78.00	6.83	6.67	7.25	6.67	
Washed / Wet	83.65	7.68	7.58	7.71	7.64	
Wet Hulling	84.25	7.83	7.83	7.67	7.75	
Total	83.71	7.69	7.60	7.72	7.64	

File Home Insert Modeling View Optimize Help Format Data / Drill

Share

Paste Cut Copy Format painter — Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources — Data

Transform data Refresh — Queries

New visual Text box More visuals — Insert

New visual calculation New measure Quick measure — Calculations

Sensitivity — Sensitivity

Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

What are the key determinants of coffee quality as evaluated through sensory attributes such as aroma, flavor, acidity, etc.?

Harvest Start Year
- 2017
- 2018
- 2021
- 2022

Country of O...
- Taiwan

Filters

dity	Average of Aftertaste	Average of Aroma	Average of Balance	Average of Body	Average of Clean Cup	Average of Flavor	Average of Sweetness	Average of
3.00	7.75	7.75	7.92	7.83	10.00	7.92	10.00	
3.00	7.75	7.75	7.92	7.83	10.00	7.92	10.00	

Visualizations

Build visual

Field

Harvest Start Year

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview Task 01 Task 02 Task 03 Task 03 Task 04

File Home Insert Modeling View Optimize Help

Share

Cut Copy Format painter — Clipboard

Get data | Excel workbook | OneLake data hub | SQL Server | Enter data | Dataverse | Recent sources — Data

Transform data | Refresh — Queries

New visual | Text box | More visuals — Insert

New visual calculation | New measure | Quick measure — Calculations

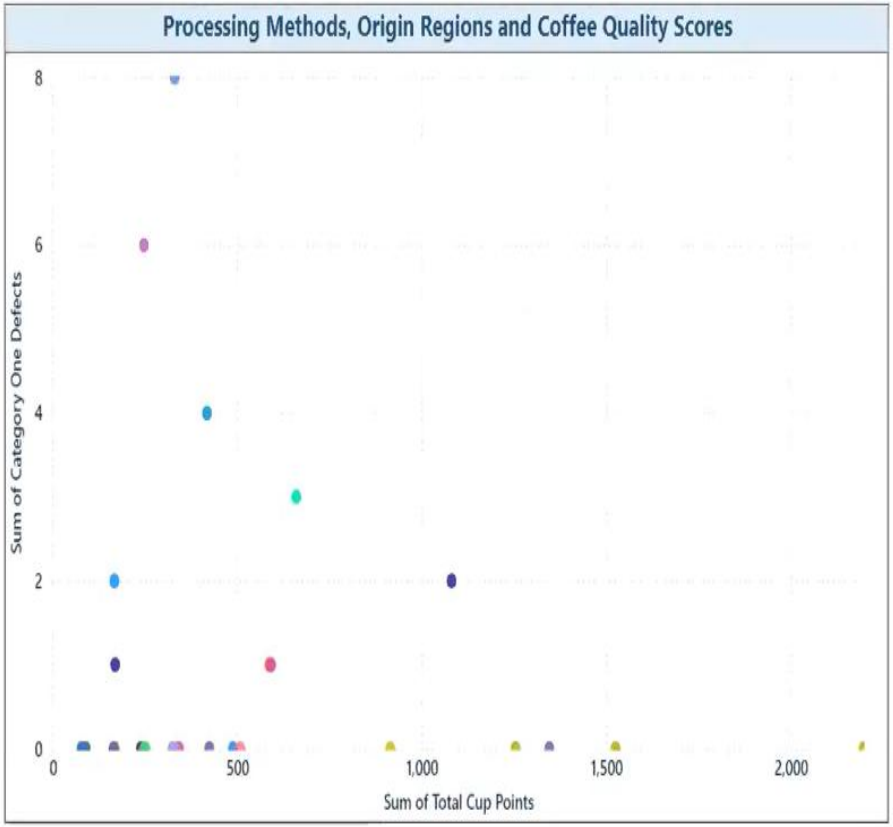
Sensitivity — Sensitivity

Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Is there a correlation between processing methods, origin regions, and coffee quality scores?



Filters

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview | Task 01 | Task 02 | Task 03 | Task 03 | Task 04 | +

File Home Insert Modeling View Optimize Help Format Data / Drill

Share

Cut Copy Format painter — Clipboard

Get data | Excel workbook | OneLake data hub | SQL Server | Enter data | Dataverse | Recent sources — Data

Transform data | Refresh — Queries

New visual | Text box | More visuals — Insert

New visual calculation | New measure | Quick measure — Calculations

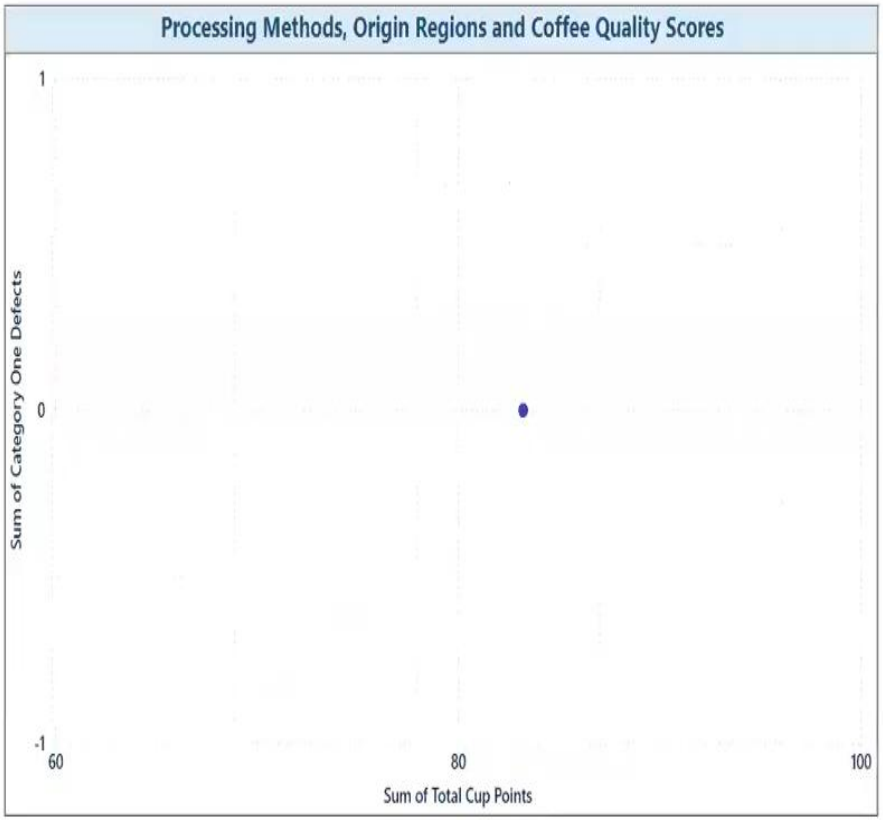
Sensitivity — Sensitivity

Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Is there a correlation between processing methods, origin regions, and coffee quality scores?



Processing Method
- Anaerobico 1000h
- Double Anaerobic Washed
- Double Carbonic Maceration / N...
- Honey,Mossto
- NA
- Natural / Dry
- Pulped natural / honey
- Semi Washed
- SEMI-LAVADO
- Washed / Wet
- Wet Hulling

Filters

Visualizations

Build visual

Field: Processing Method

Drill through

Cross-report: Off

Keep all filters: On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview | Task 01 | Task 02 | Task 03 | Task 03 | Task 04 | +

File Home Insert Modeling View Optimize Help

Share

Paste Cut Copy Format painter — Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources — Data

Transform data Refresh — Queries

New visual Text box More visuals — Insert

New visual calculation New measure Quick measure — Calculations

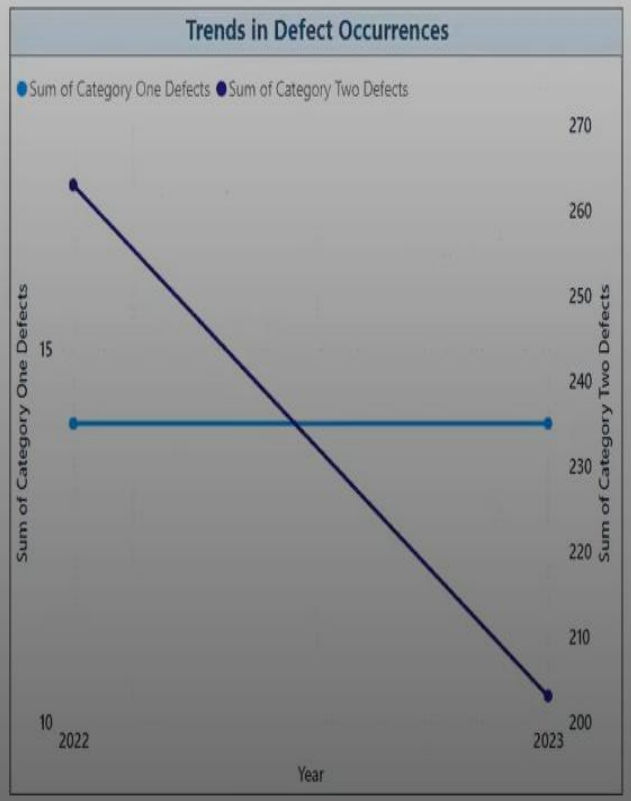
Sensitivity — Sensitivity

Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Can we identify any trends or patterns in defect occurrences?



Filters

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview Task 01 Task 02 Task 03 Task 03 Task 04

File Home Insert Modeling View Optimize Help

Share

Paste Cut Copy Format painter — Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources — Data

Transform data Refresh — Queries

New visual Text box More visuals — Insert

New visual calculation New measure Quick measure — Calculations

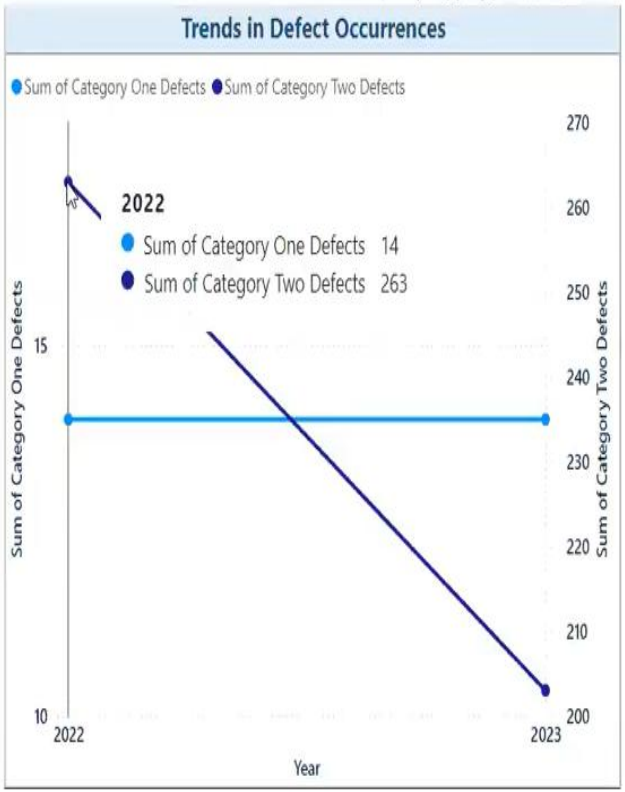
Sensitivity — Sensitivity

Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Can we identify any trends or patterns in defect occurrences?



Filters

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview Task 01 Task 02 Task 03 Task 03 Task 04

File Home Insert Modeling View Optimize Help

Share

Paste Cut Copy Format painter — Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources — Data

Transform data Refresh — Queries

New visual Text box More visuals — Insert

New visual calculation New measure Quick measure — Calculations

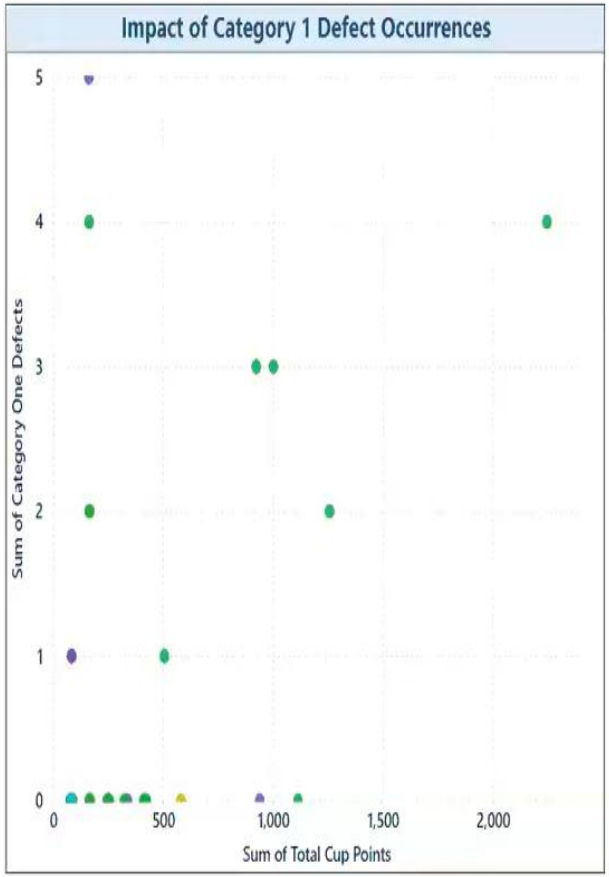
Sensitivity — Sensitivity

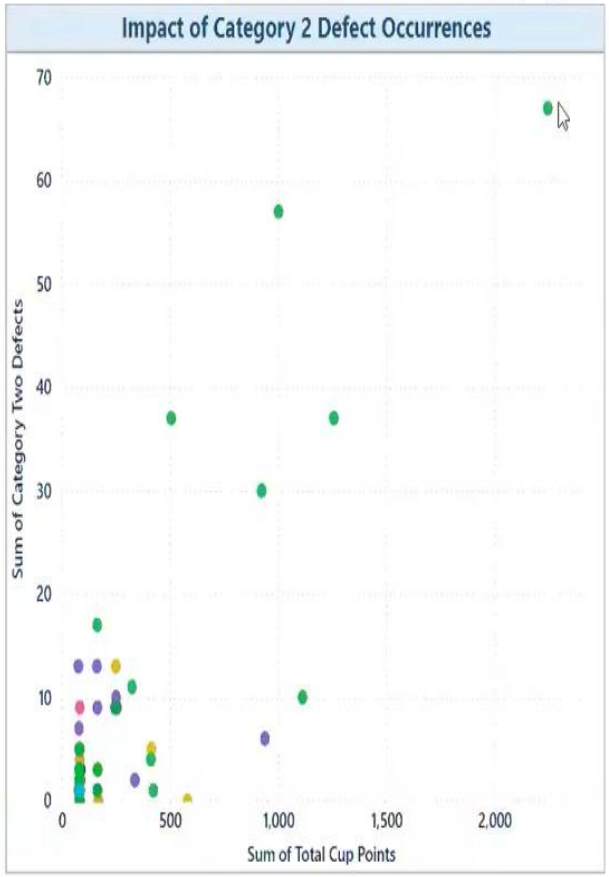
Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Can we identify the impact of defect occurrences on overall coffee quality?





Filters

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview Task 01 Task 02 Task 03 Task 03 Task 04

File Home Insert Modeling View Optimize Help

Share

Paste Cut Copy Format painter — Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources — Data

Transform data Refresh — Queries

New visual Text box More visuals — Insert

New visual calculation New measure Quick measure — Calculations

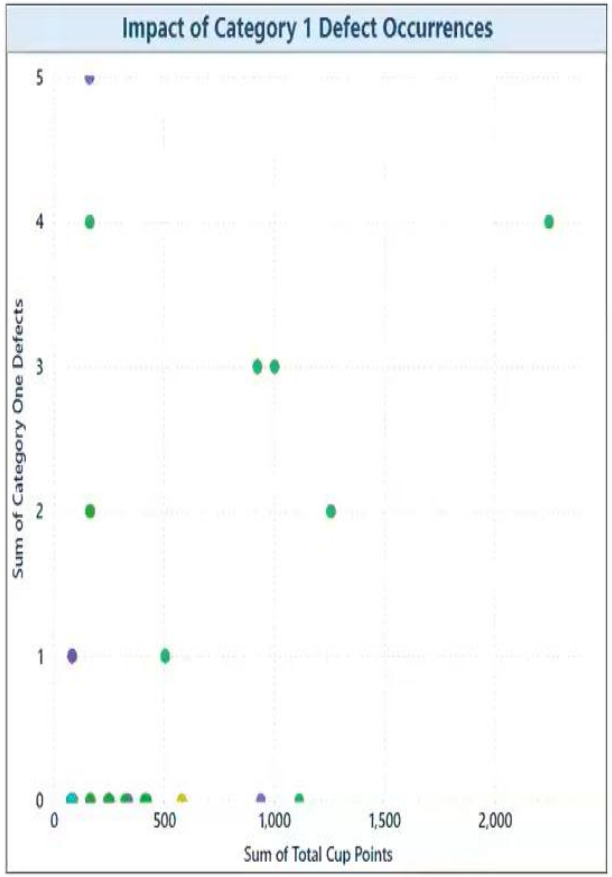
Sensitivity — Sensitivity

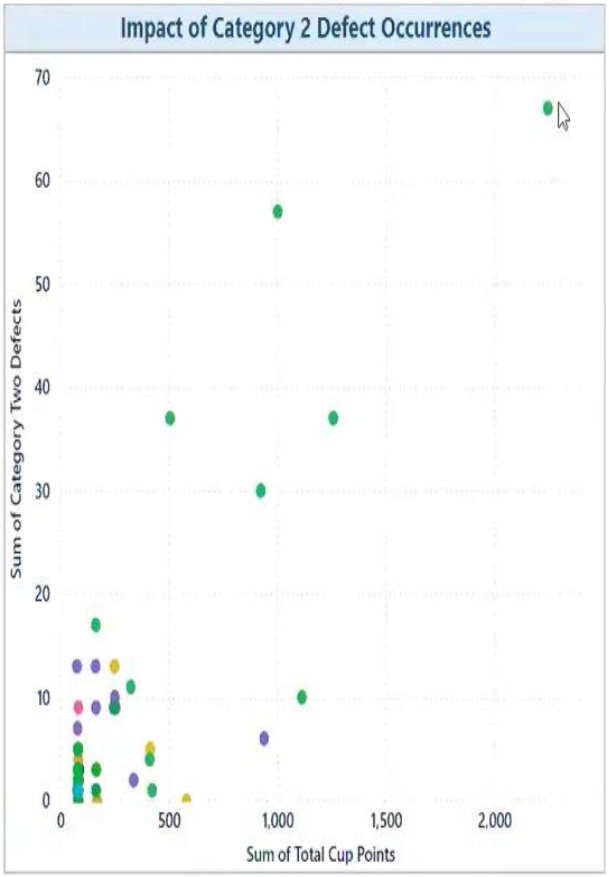
Publish — Share

Copilot — Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Can we identify the impact of defect occurrences on overall coffee quality?





Filters

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview Task 01 Task 02 Task 03 Task 03 Task 04

File Home Insert Modeling View Optimize Help

Share

Cut Copy Format painter

Clipboard

Get data | Excel workbook | OneLake data hub | SQL Server | Enter data | Dataverse | Recent sources

Data

Transform data | Refresh

Queries

New visual | Text box | More visuals

Insert

New visual calculation | New measure | Quick measure

Calculations

Sensitivity

Sensitivity

Publish

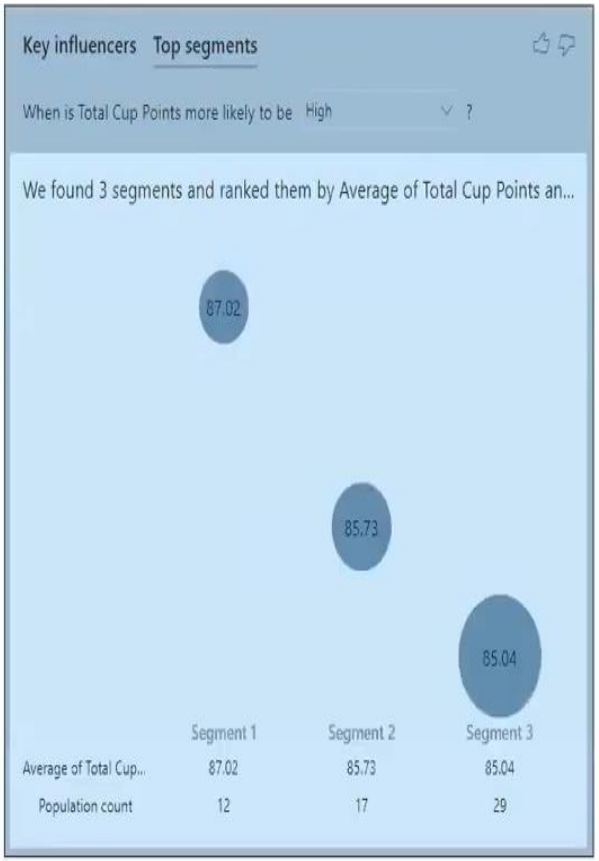
Share

Copilot

Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

How do different variables interact to influence the Total Cup Points, which represent an overall measure of coffee quality?



Filters

Visualizations

Build visual

Values

Add data fields here

Drill through

Cross-report Off

Keep all filters On

Add drill-through fields here

Data

Search

df_arabica_clean

Overview | Task 01 | Task 02 | Task 03 | Task 03 | Task 04

File Home Insert Modeling View Optimize Help Format Data / Drill

Share

Paste Cut Copy Format painter — Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources — Data

Transform data Refresh — Queries

New visual Text box More visuals — Insert

New visual calculation New measure Quick measure — Calculations

Sensitivity — Sensitivity

Publish — Share

Copilot — Copilot

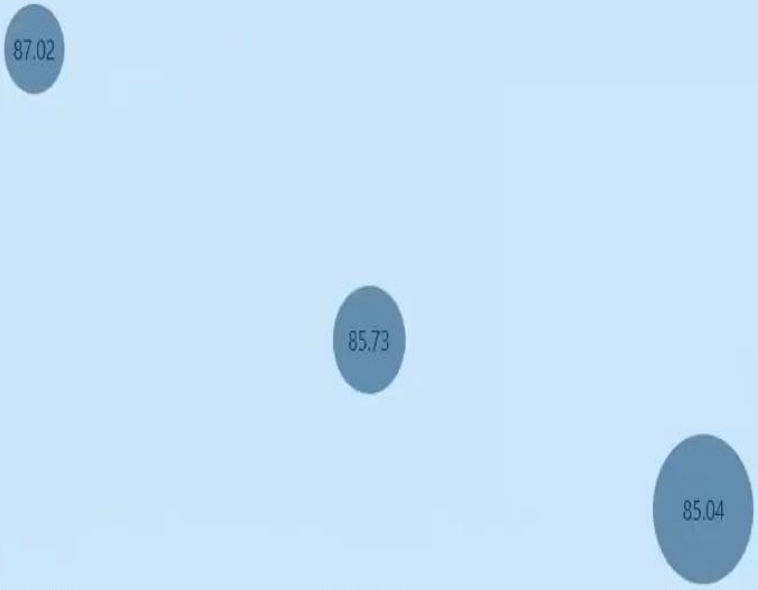
Auto recovery contains some recovered files that haven't been opened. View recovered files

Back to report

Key influencers Top segments

When is Total Cup Points more likely to be High ?

We found 3 segments and ranked them by Average of Total Cup Points and population size. Select a segment to see more details.



	Segment 1	Segment 2	Segment 3
Average of Total Cup Points	87.02	85.73	85.04
Population count	12	17	29

Filters

Visualizations

Build visual

Analyze

Total Cup Points

Explain by

Sum of Acidity
Sum of Aftertaste
Sum of Aroma
Sum of Balance
Sum of Body
Sum of Clean Cup
Sum of Flavor
Sum of Sweetness

Data

Search

df_arabica_clean

Page 6 of 6

100%

File Home Insert Modeling View Optimize Help Format Data / Drill

Share

Cut Copy Format painter Paste

Clipboard

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources

Data

Transform data Refresh

Queries

New visual Text box More visuals

Insert

New visual calculation New measure Quick measure

Calculations

Sensitivity

Sensitivity

Publish

Share

Copilot

Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Back to report

Key influencers Top segments

When is Total Cup Points more likely to be High ?

87.02 85.73 85.04

Segment 1

Sum of Aftertaste is greater than 7.83

Sum of Flavor is greater than 8.08

In segment 1, the average Total Cup Points is 87.02. This is 3.31 units higher than the overall average, 83.71.

Segment 1 contains 12 data points (5.8% of the data).

Filters

Visualizations

Build visual

Analyze

Total Cup Points

Explain by

Sum of Acidity
Sum of Aftertaste
Sum of Aroma
Sum of Balance
Sum of Body
Sum of Clean Cup
Sum of Flavor
Sum of Sweetness

Data

Search

df_arabica_clean

Page 6 of 6

100%

File Home Insert Modeling View Optimize Help Format Data / Drill

Share

Cut Copy Format painter

Get data Excel workbook OneLake data hub SQL Server Enter data Dataverse Recent sources

Transform data Refresh

New visual Text box More visuals

New visual calculation New measure Quick measure

Sensitivity

Publish

Copilot

Clipboard Data Queries Insert Calculations Sensitivity Share Copilot

Auto recovery contains some recovered files that haven't been opened. View recovered files

Back to report

Key influencers Top segments

What influences Total Cup Points to Increase ?

When...	...the average of Total Cup Points increases by
Sum of Flavor goes up 0.28	0.35
Sum of Aftertaste goes up 0.28	0.33
Sum of Balance goes up 0.26	0.32
Sum of Acidity goes up 0.26	0.32
Sum of Aroma goes up 0.29	0.32
Sum of Body goes up 0.23	0.26

Filters

Visualizations

Build visual

Analyze

Total Cup Points

Explain by

Sum of Acidity
Sum of Aftertaste
Sum of Aroma
Sum of Balance
Sum of Body
Sum of Clean Cup
Sum of Flavor
Sum of Sweetness

Data

Search

df_arabica_clean

100%

Thank You