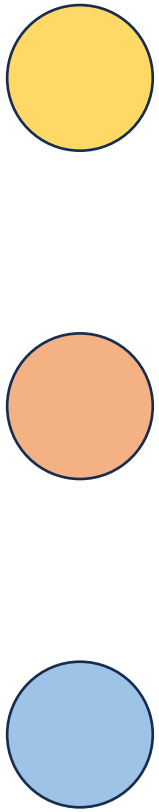
Forget ‘Quality’ oriented Paper Shuffling cupping forms:
Use Sensory Science for Effective Evaluations and Consumer Preferences predictions

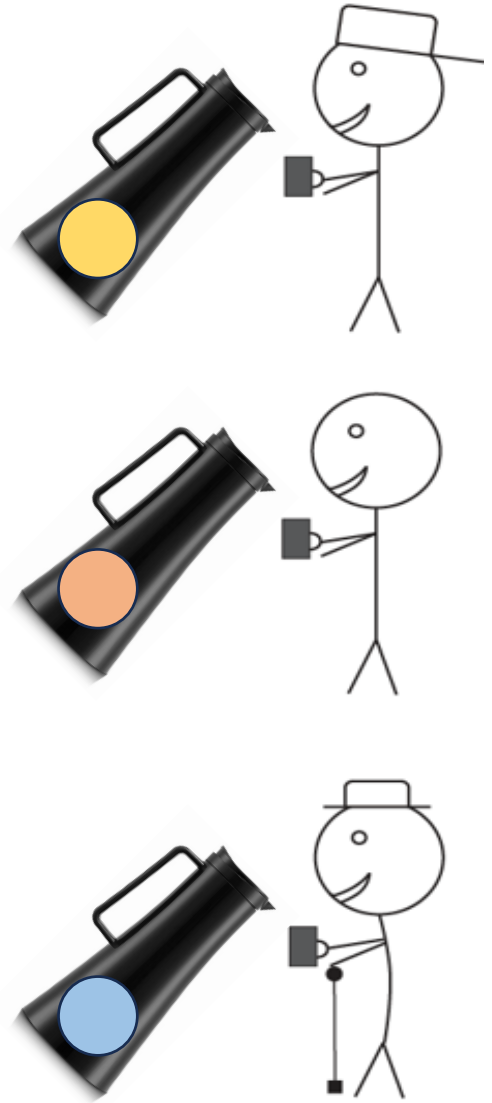
You have a business dream:

Different flavour



Describe flavour

Who likes what?



Principles of Sensory Evaluation of Food by Rose Marie Pangborn, 1965

Different flavour

DISCRIMINATIVE TEST

Is there a difference between products?

Describe flavour

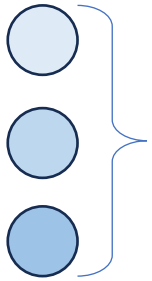
DESCRIPTIVE TEST

What characterises the difference and how strong is it?

Who likes what?

AFFECTIVE TEST

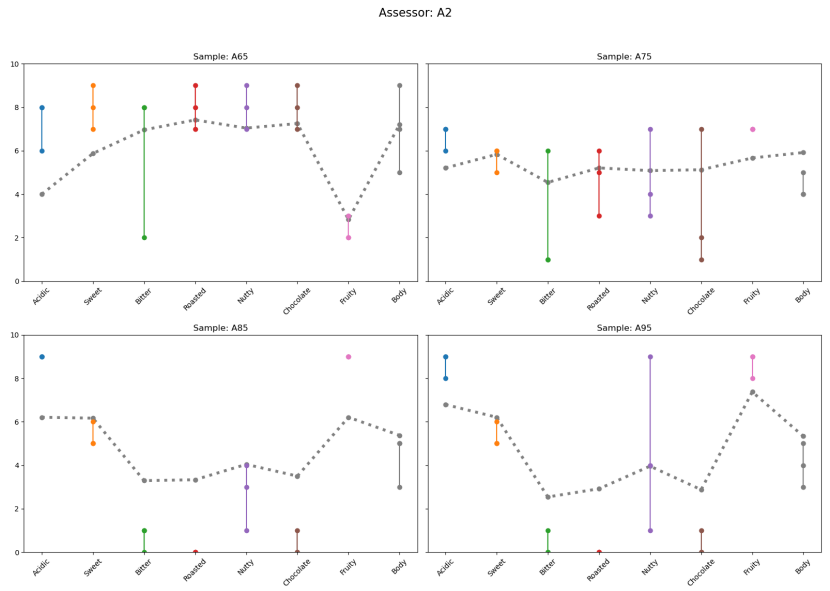
Does the difference have any meaning for the customer?

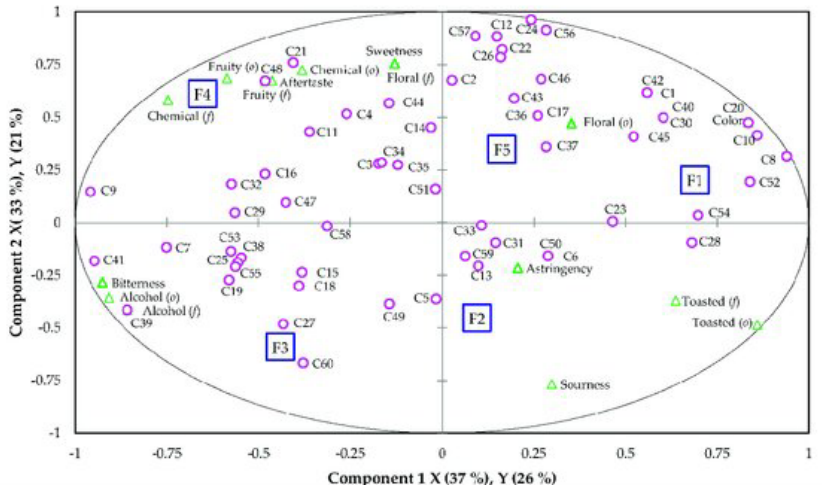


Small differences or precise **scientific investigations**

Assessor	Table 1	Table 2	Table 3
1	3	1	2
10	0	3	2
12	1	1	2
2	1	1	2
3	0	1	2
4	1	1	2
5	0	1	1
6	2	2	1
7	1	3	2
9	2	2	2
Nummer elleve	1	3	2

Not significant (> 5%) Significant (1% - 5%) Highly significant (< 1%)

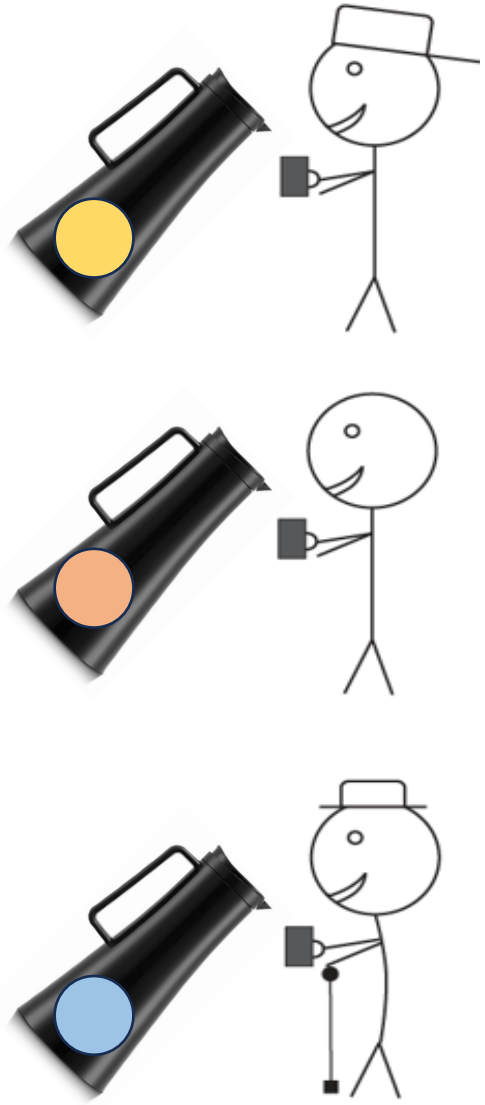




Different flavour

Describe flavour

Who likes what?



Nowhere is a bureaucratic centralised concept of ‘quality’ or even elaborate forms relevant but it seems rather side-tracking

Coffee Roastery startup

Proven sensory Facts

Your Perfect Coffee Roastery

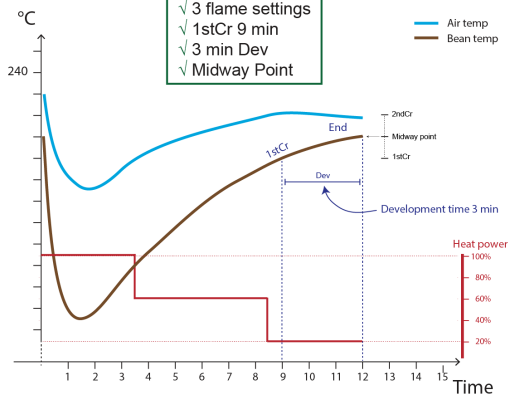
Lean operation: processes optimized and wastes minimized based on customer model

CoffeePreneur
Personal vision
Company vision
Strategy

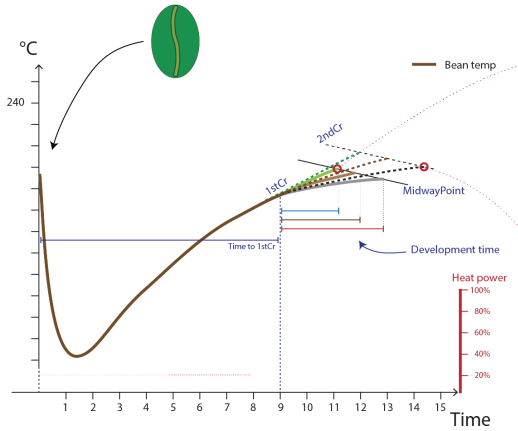
Balanced Influence

Control
Thermodynamic blueprint

Standard Roast
√ 3 flame settings
√ 1stCr 9 min
√ 3 min Dev
√ Midway Point



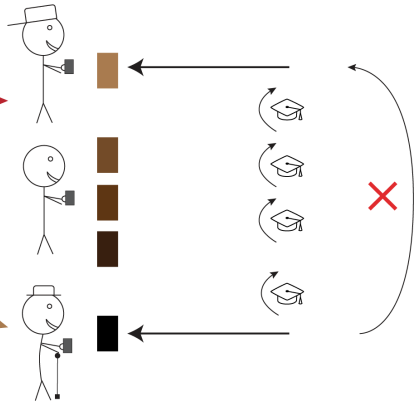
Flavour Experimentation
Bean + Control + Variation



Evaluation
Sensory Intensities

Quality Control

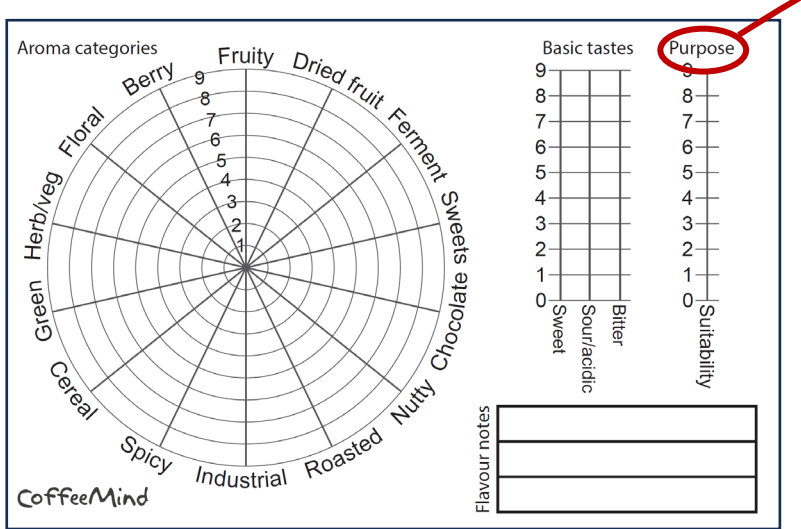
Goal
Consumer preferences

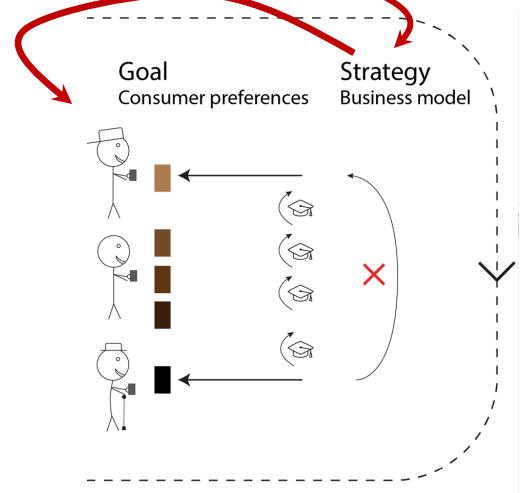


Strategy
Business model

Continous improvement of company processes and relation to CoffeePreneur lifestyle

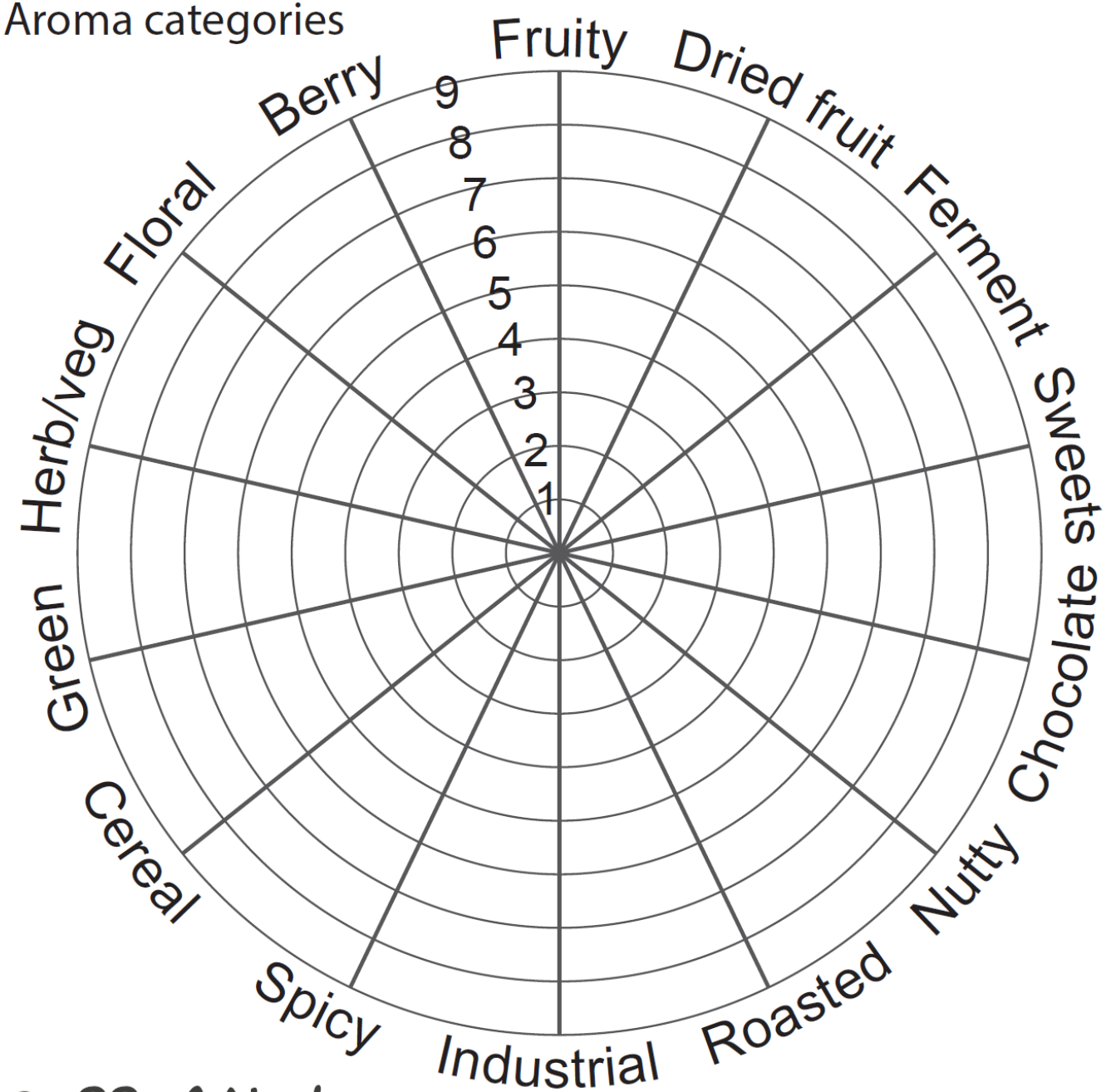
CoffeeMind



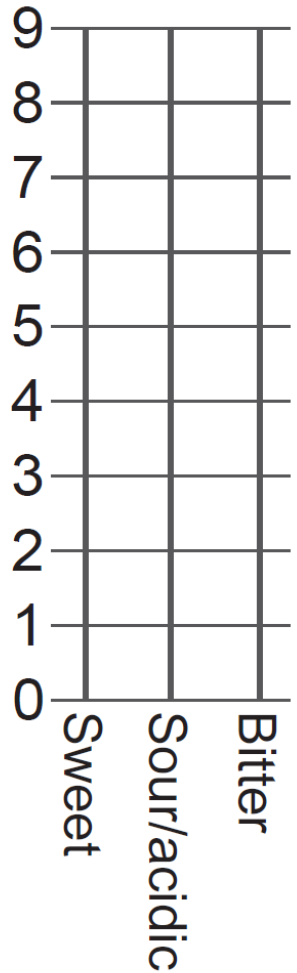


You don't need all the paperwork or the statistical software
Which thermos empties first?

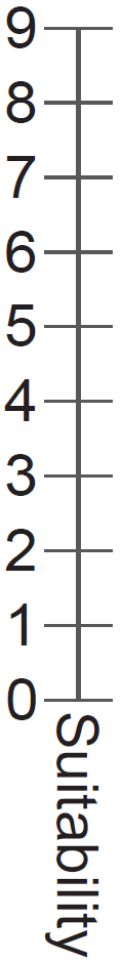
Aroma categories



Basic tastes

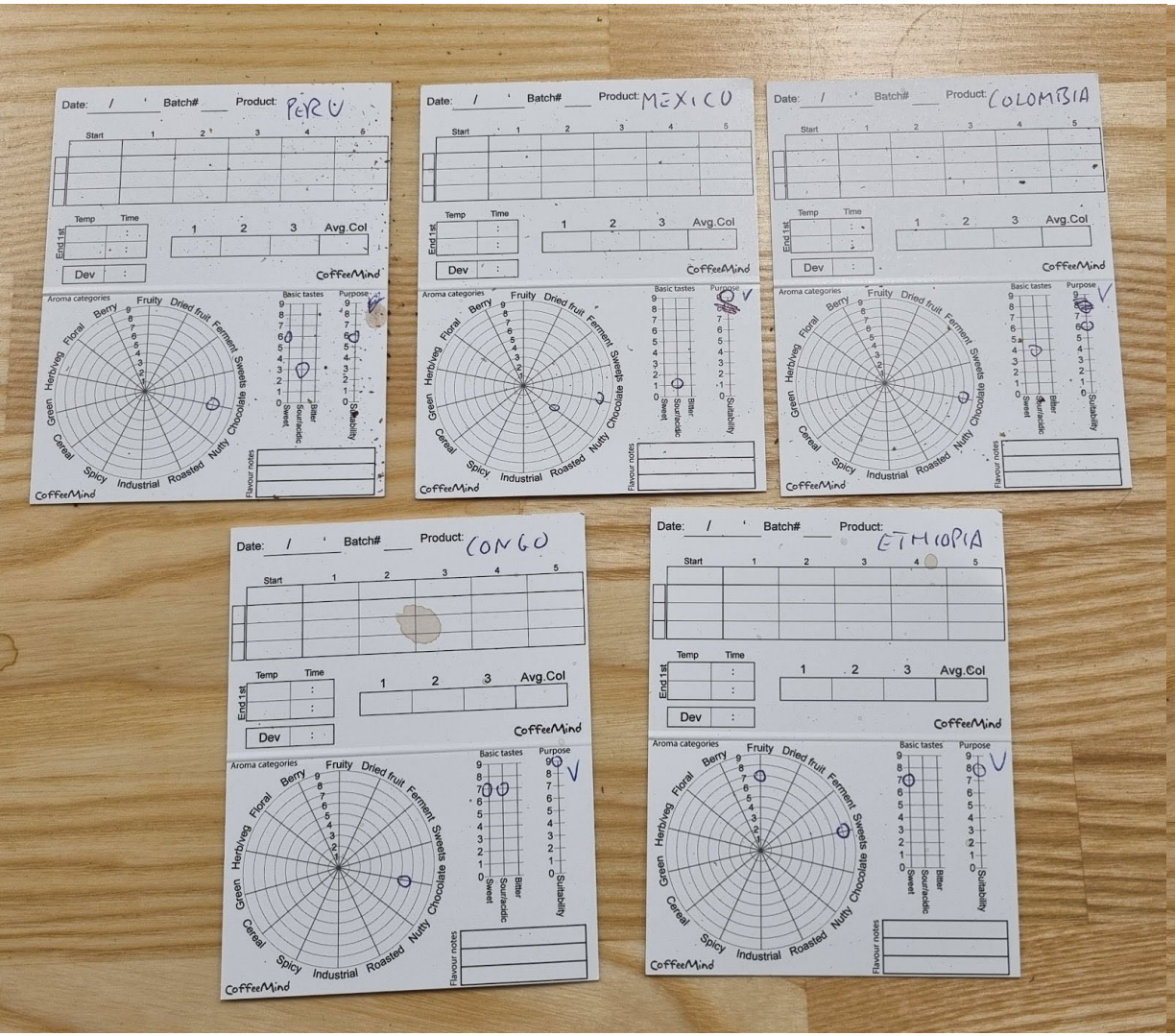


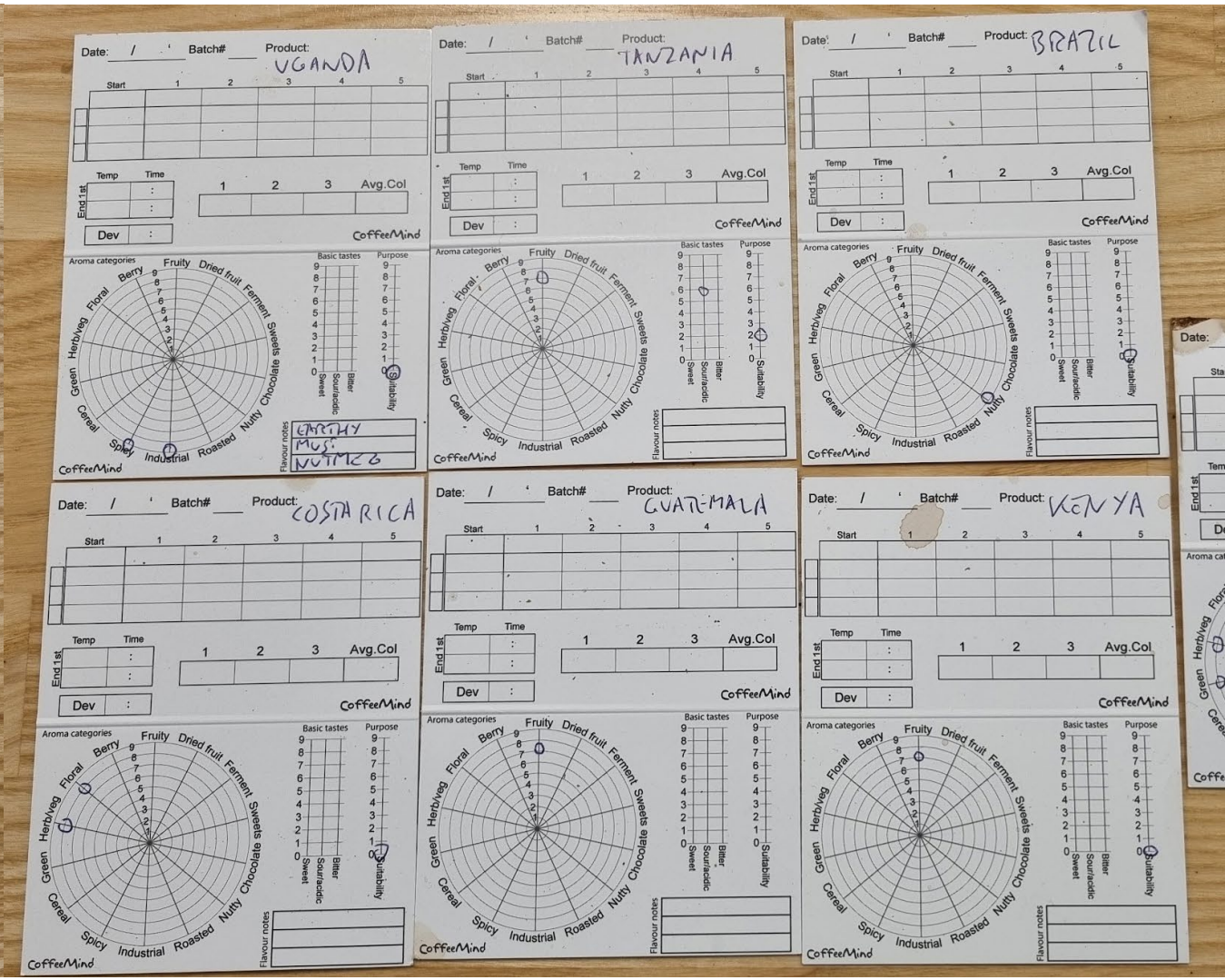
Purpose

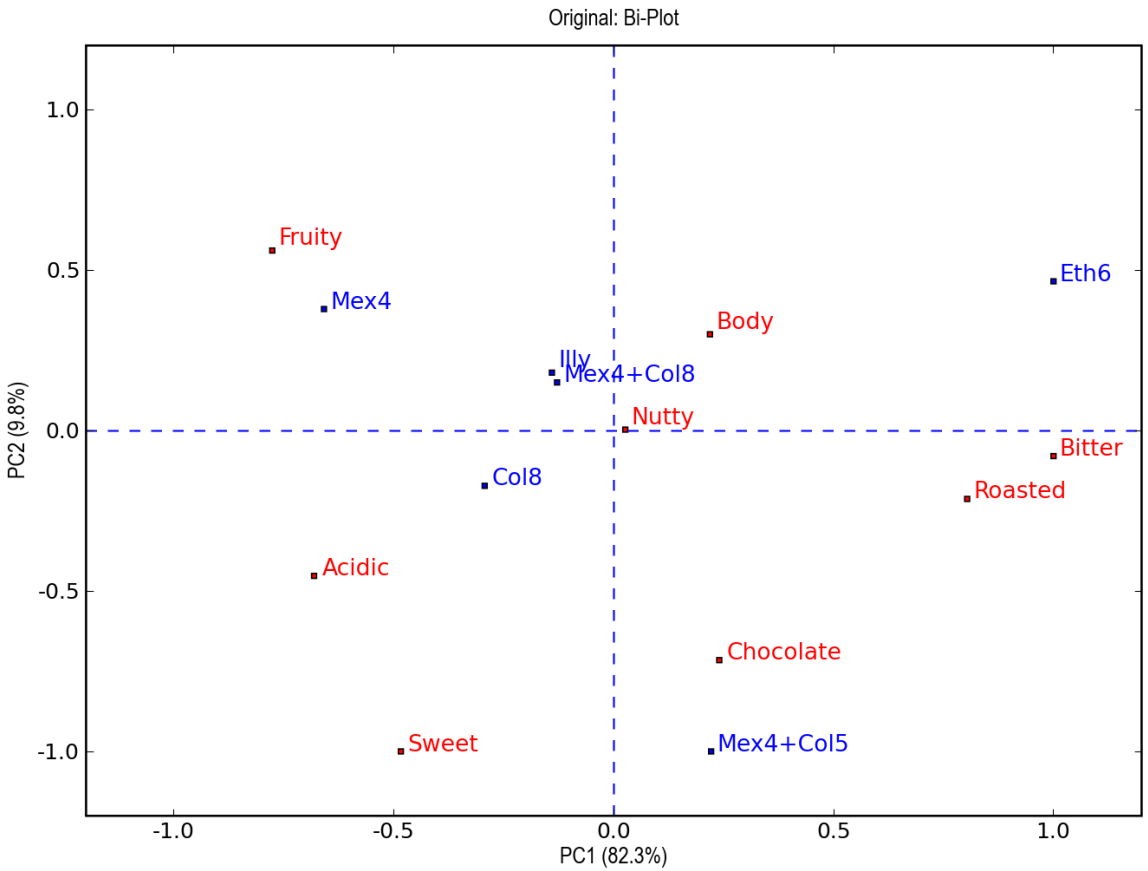


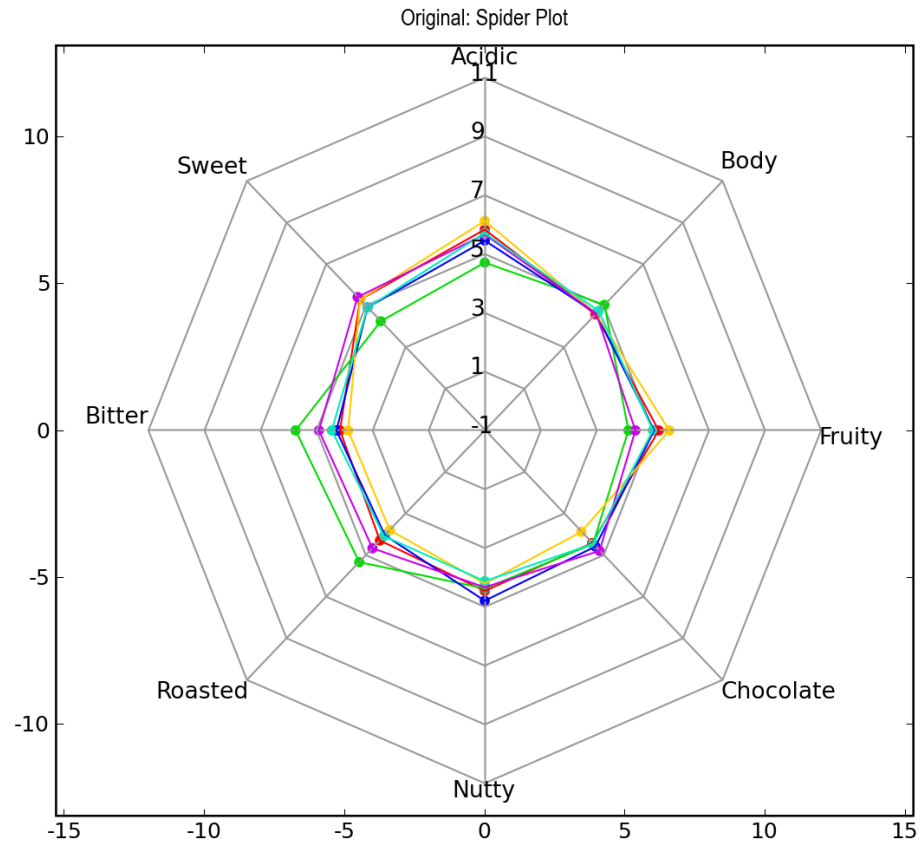
Flavour notes

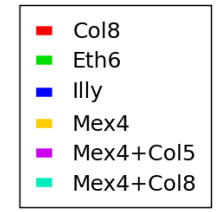
CoffeeMind


Date: / ' Batch# Product: PERU
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: MEXICO
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: COLOMBIA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: CONGO
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: ETHIOPIA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes


Date: / ' Batch# Product: UGANDA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
EARTHY
MUSTY
NUTMEG
Date: / ' Batch# Product: TANZANIA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: BRAZIL
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: COSTA RICA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: GUATEMALA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes
Date: / ' Batch# Product: KENYA
Start 1 2 3 4 5
Temp Time
End 1st : :
1 2 3 Avg.Col
Dev :
CoffeeMind
Aroma categories
Fruity Dried fruit Ferment Sweets Chocolate Nutty Roasted Industrial Spicy Cereal Green Herb/veg Floral Berry
Basic tastes
Sweet Sour/acidic Bitter
Purpose
Suitability
Flavour notes


Original: Bi-Plot
Fruity
Mex4
Eth6
Body
Illy
Mex4+Col8
Nutty
Bitter
Col8
Roasted
Acidic
Chocolate
Sweet
Mex4+Col5
PC1 (82.3%)
PC2 (9.8%)
1.0
0.5
0.0
-0.5
-1.0


Original: Spider Plot
Acidic
Body
Fruity
Chocolate
Nutty
Roasted
Bitter
Sweet
11
9
7
5
3
1
-1
15
10
5
0
-5
-10
-15


Col8
Eth6
Illy
Mex4
Mex4+Col5
Mex4+Col8

Notice ‘Quality’ is a Type 1 error? Why would I do that?

- Foundation of Sensory science (Rose Marie Pangborn’s 3 stage model)
- The data I have been able to find
- The published research about SCA 2004 cupping form

Light

Med

Dark

DISCRIMINATIVE TEST

Is there a difference between products?

DESCRIPTIVE TEST

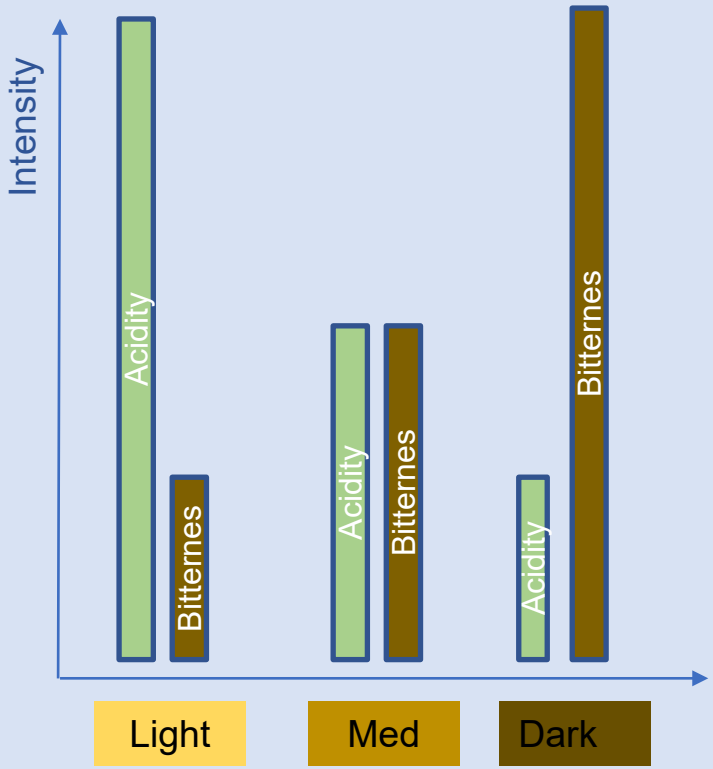
What characterises the difference and how strong is it?

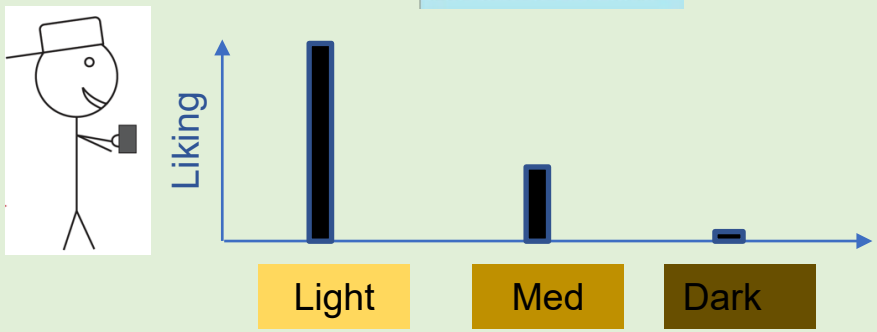
AFFECTIVE TEST

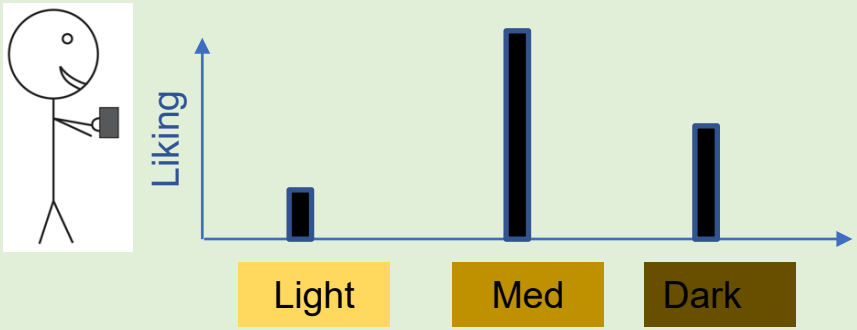
Does the difference have any meaning for the customer?

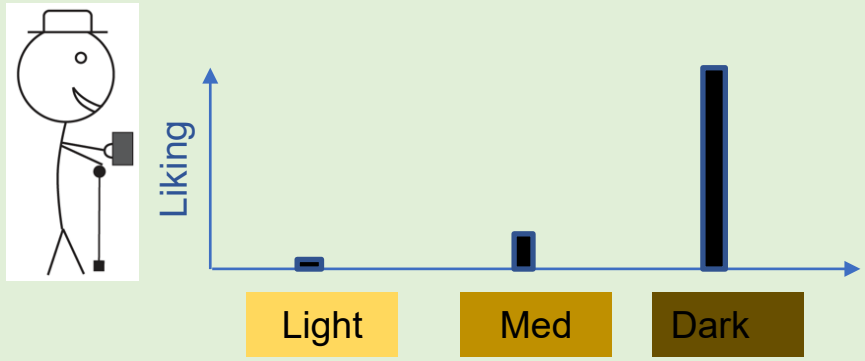
A B C

P < 0,05









‘Quality’ is a Type 1 error if you think it exists as a universal scale indendent of the observer because it is a feature of the observer but interpreted as a feature of the coffee.
There is a fundamental problem and confusion regarding the ‘object of investigation!!

Object of investigation:

Food substance (**Coffee**)	**Persons** preferences

Light

Med

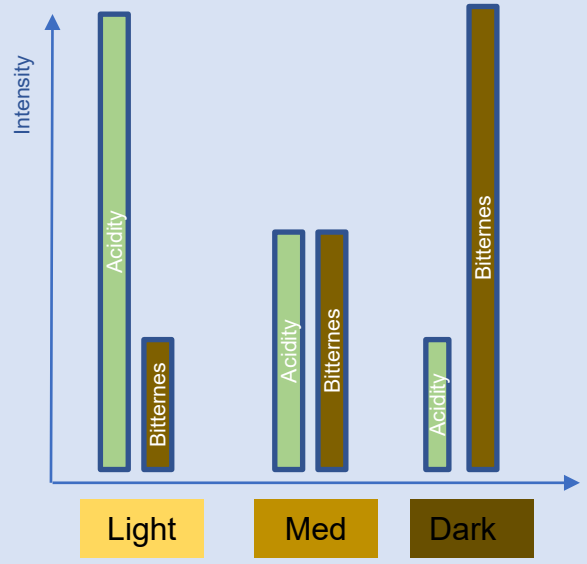
Dark

DISCRIMINATIVE TEST
Is there a difference between products?

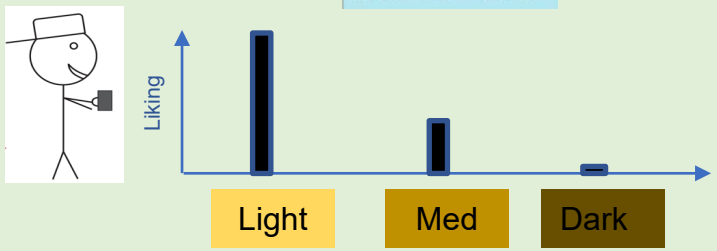
A B C

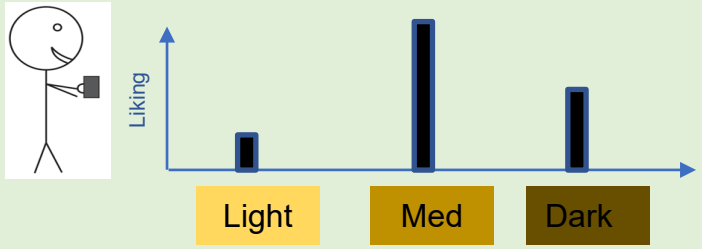
P < 0,05

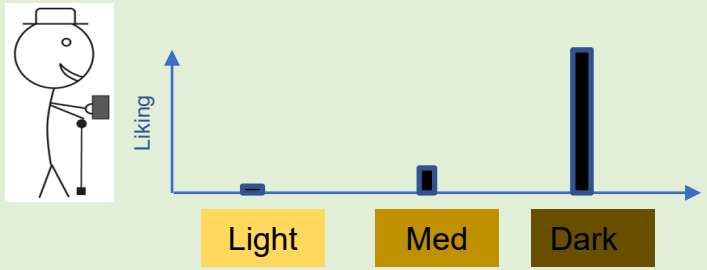
DESCRIPTIVE TEST
What characterises the difference and how strong is it?



AFFECTIVE TEST
Does the difference have any meaning for the customer?







But what if quality assessors are well calibrated?

- If we accept that quality is not a feature of the product but the opinioin by a well calibrated group of people whose opinoin generates the data?
- I have not seen any practical nor scientific documentation of low enough variation between quality scores to justify this

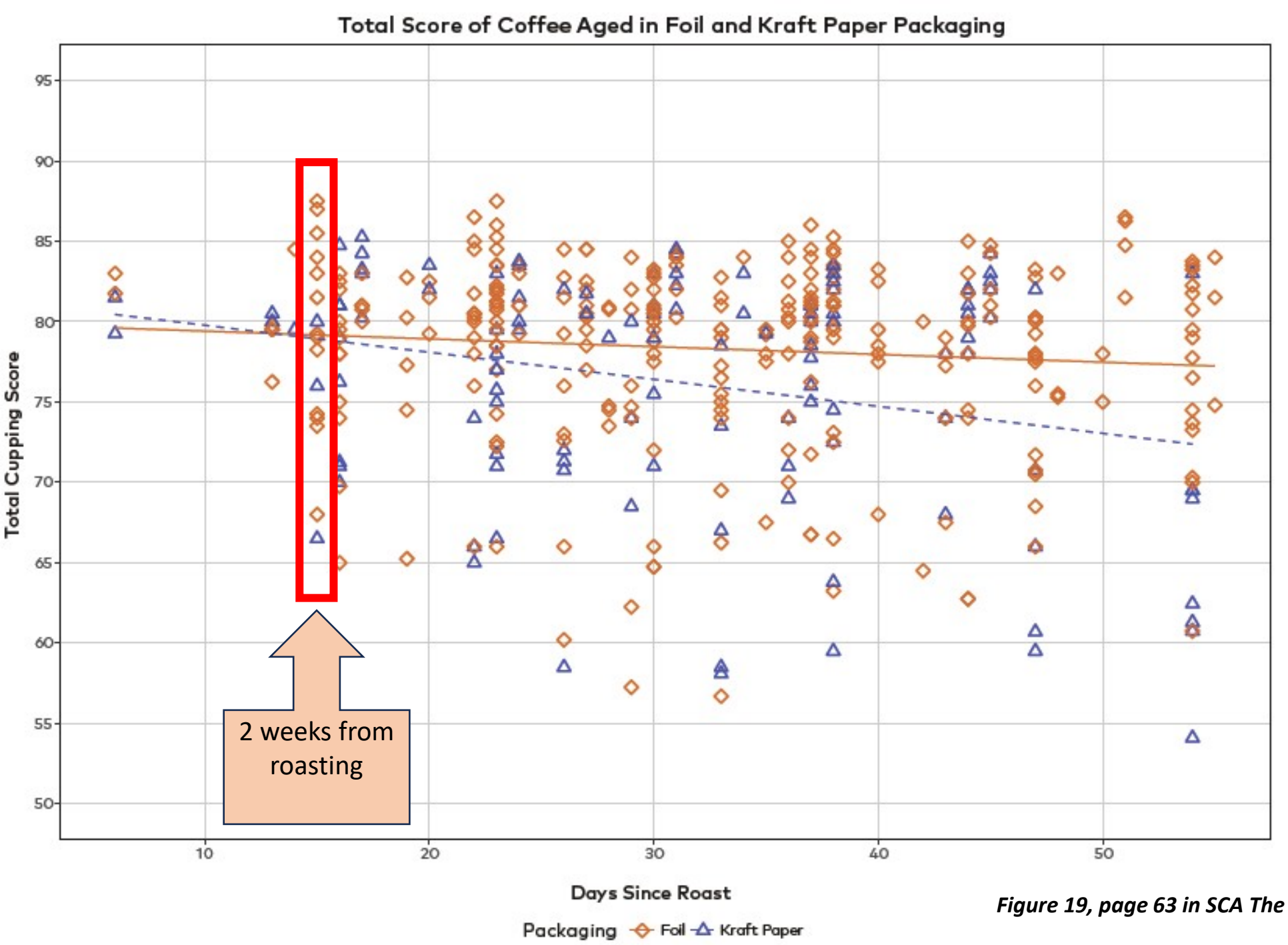
Email from my friend Gilbert in Lebanon. 12th of June 2024

“From your experience, are the below numbers viable? I mean when I buy Green coffee I look at the scoring score from the supplier. Yet after roasting we started to score them in-house.

The variation difference a BIG looking at our scores vs. supplier scores, please advise on the reasons behind the BIG difference”

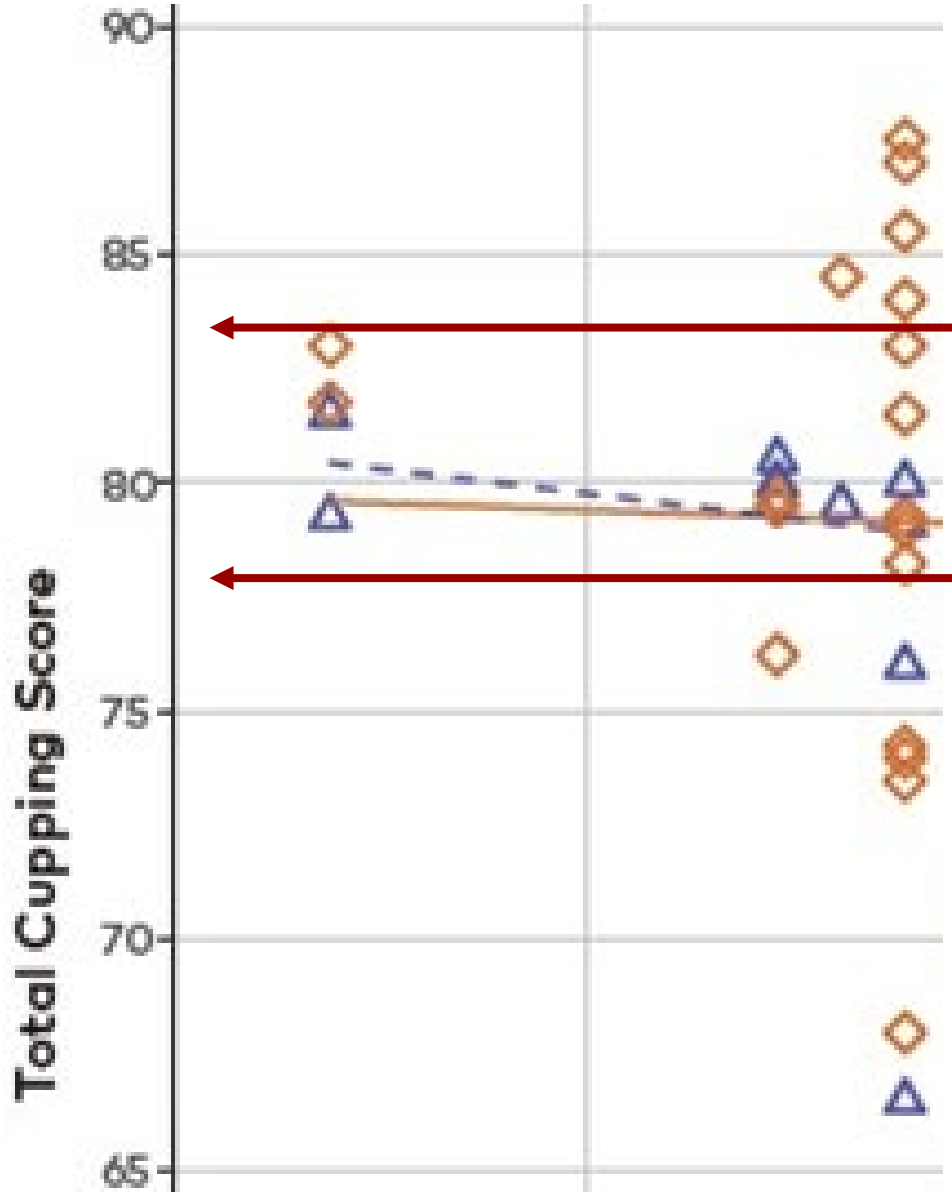
Origin	Supplier ratings	Roaster ratings
KENYA	86	73
GUATEMALA	84.5	70
INDONESIA	83	60.75
BRAZIL	83.5	60.75
COSTA RICA	83.5	65.75
COLOMBIAN	81.75	63.25

But are they calibrated?



"Experienced cuppers, who self-identified as "one who cups three or more times per week and scores coffee based on the SCAA quality assessment form"

Figure 19, page 63 in SCA The Coffee Freshness Handbook



Import quality score of 83,25

Upper 95% conf.int = 81,7

Average quality score was 78,6

Lower 95% conf.int = 75,5

With 16 professional cuppers this is an unacceptable broad 95% conf.int. that spans too many critical thresholds

A general symptom for ‘sensory science’

- SCAE’s Freshness handbook contains **PERFECT chemical** data analysis and just **crappy sensory data** and data analysis.
- This is a tendency not only in the coffee business but also in scientific articles! More about that later.
- Seems like BOTH the specialty coffee business AND scientific community in general **do not know and apply sensory science** despite it’s 50 years of existence.
- Ironic as the sensory experience is the whole point of good coffee!!
- What does it mean for your executive **skills**, executive **possibilities** and **planing** if it is not clear for you where you are going and why?

We are lost - but at a GREAT speed!! (and at a HUGE cost)

What does science say about SCA 2004 quality score system?

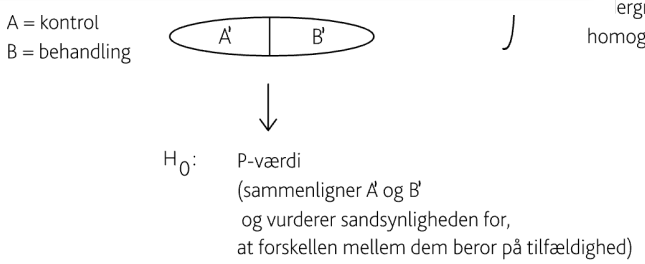
What is statistics in the first place?

A slide I made in 2008 when teaching medical students about research desing

Term	Population Notation	Sample Notation
Mean	μ	$\bar{x}$
Variance	σ^2	s^2
Standard Deviation	σ	s
Number of Observations	N	n

stik er beregninger af relationer
em Population og Stikprøve

ɔgsdesign er strategier for at
sig, at stikprøven er:
evant
ræsentativ mht. population
ergrupper er indbyrdes
homogene (bias)

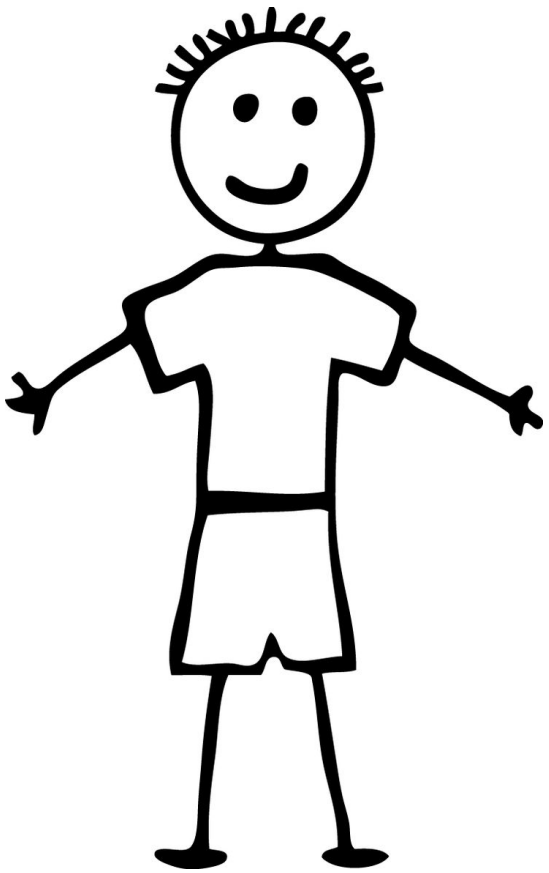


Quote from wikipedia:

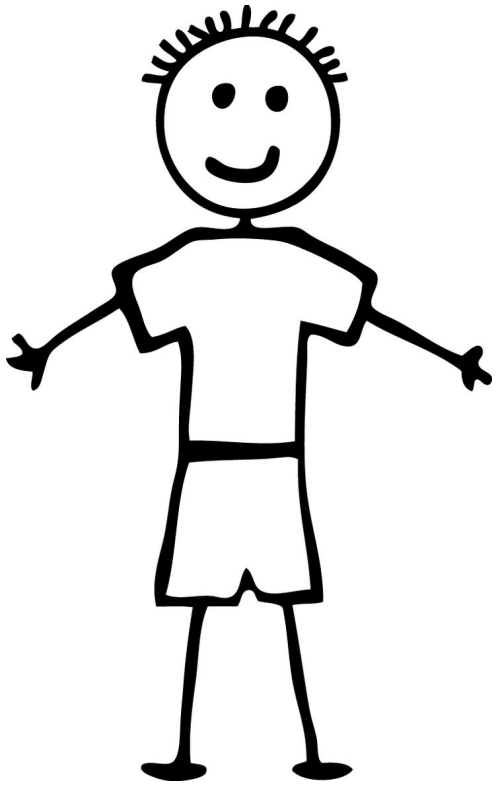
"Representative sampling assures that
inferences and conclusions can reasonably extend from the sample to the population as a whole"

I want to establish mean height and the variation in heights among boys in the 6th grade in Denmark

155 cm



150 cm



Mean ($\bar{x}$)

The mean height is calculated by taking the sum of the heights and dividing by the number of observations.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Where:

- x_i are the individual heights.
- n is the number of observations.

For our two boys:

$$\bar{x} = \frac{155 + 150}{2} = \frac{305}{2} = 152.5 \text{ cm}$$

Variance (s^2)

Variance measures the dispersion of the data points around the mean. It is calculated as:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

For our two boys:

$$s^2 = \frac{1}{2-1}\left[(155 - 152.5)^2 + (150 - 152.5)^2\right]$$

$$s^2 = \left[(2.5)^2 + (-2.5)^2\right] = 6.25 + 6.25 = 12.5$$

Standard Deviation (s)

Standard deviation is the square root of the variance:

$$s = \sqrt{s^2} = \sqrt{12.5} \approx 3.54 \text{ cm}$$

Conclusion: The Average Height of 6th grade boys in Denmark is 152,5 cm and the Standard Deviation is 3,54 cm

Scientific survey

Rest assured, providing your email here **will NOT subscribe you to any mailing lists**. It's used solely for scientific data verification purposes.

This survey is serious business! Although it involves tooth fairies, it is designed to uncover significant insights about the state of coffee education. So please **ONLY GIVE THRUTHFUL ESTIMATES**

The findings will be revealed at the World of Coffee event in Copenhagen this June. See you there or online as all our presentations will be available online after the event.

You will receive the result of the research as well as a link to a presentation of the research on the email below but nothing more.

Mailadresse *

Gyldig mailadresse

Denne formular indhenter mailadresser. Rediger indstillinger

In which geographical region did your family experience the tooth fairy tradition of receiving money *

- Africa
- Antarctica
- Asia
- Australia
- Europe
- Middle East
- North America
- South America

How many children are there in the household? *

- 1
- 2
- 3
- 4
- 5
- More than 5

Where is the household situated?

Flere svarmuligheder

- Countryside
- Village
- City
- Capital

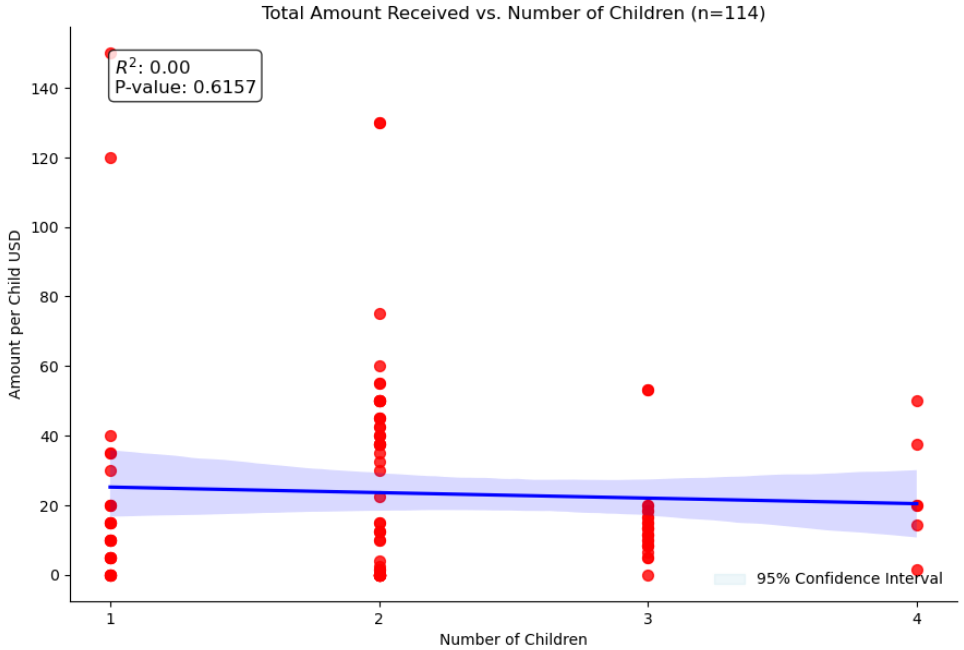
Tilføj mulighed eller tilføj "Andet"

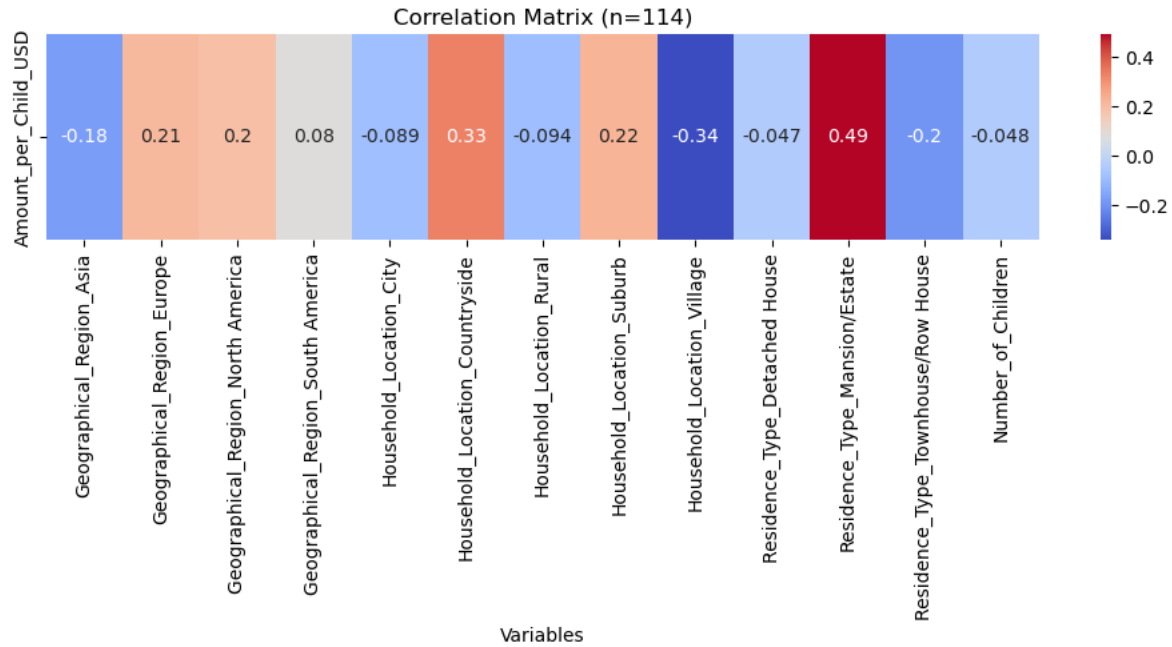
Påkrævet

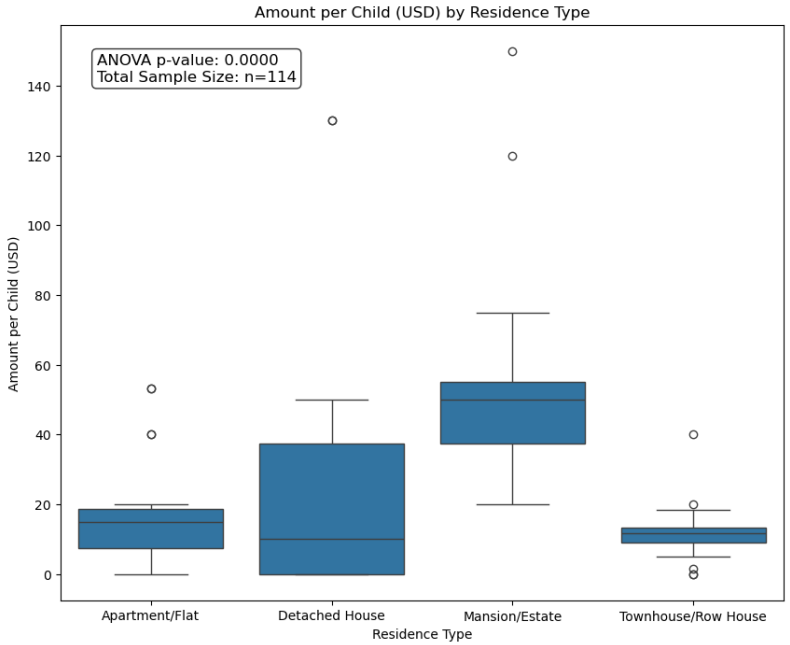
Total Amount Recei	In which geographic	How many children	Does the tooth fairy seem to give more to the firstborn a	Where is the housel	What is the type of r
10	Asia	1	Does not apply (in case of only 1 child in household)	City	Detached House
0	Asia	1	Does not apply (in case of only 1 child in household)	City	Detached House
25	Europe	2	Yes	City	Townhouse/Row House
20	Europe	2	No	City	Townhouse/Row House
30	Europe	2	No	City	Detached House
3	Europe	2	No	Village	Detached House
5.8	Europe	4	No	Countryside	Townhouse/Row House
260	Europe	2	Yes	City	Detached House
150	Europe	1	Does not apply (in case of only 1 child in household)	Countryside	Mansion/Estate
60	Europe	3	No	City	Apartment/Flat
100	Europe	2	No	City	Detached House
80	South America	2	No	Capital	Apartment/Flat
0	Europe	1	Does not apply (in case of only 1 child in household)	City	Townhouse/Row House
160	Europe	3	No	Capital	Apartment/Flat
50	Europe	3	No	Capital	Apartment/Flat
75	Europe	2	No	City	Mansion/Estate
10	Asia	1	Does not apply (in case of only 1 child in household)	City	Detached House
0	Asia	1	Does not apply (in case of only 1 child in household)	City	Detached House
25	Europe	2	Yes	City	Townhouse/Row House
20	Europe	2	No	City	Townhouse/Row House
30	Europe	2	No	City	Detached House
3	Europe	2	No	Village	Detached House
5.8	Europe	4	No	Countryside	Townhouse/Row House
260	Europe	2	Yes	City	Detached House
120	Europe	1	Does not apply (in case of only 1 child in household)	Countryside	Mansion/Estate
60	Europe	3	No	City	Apartment/Flat
100	Europe	2	No	City	Detached House
80	South America	2	No	Capital	Apartment/Flat
0	Europe	1	Does not apply (in case of only 1 child in household)	City	Townhouse/Row House
160	Europe	3	No	Capital	Apartment/Flat
50	Europe	3	No	Capital	Apartment/Flat
110	Europe	2	No	City	Mansion/Estate
0	Asia	2	No	City	Apartment/Flat
20	Europe	1	Does not apply (in case of only 1 child in household)	City	Townhouse/Row House
15	Asia	3	No	Rural	Detached House
5	North America	1	Does not apply (in case of only 1 child in household)	City	Apartment/Flat
150	North America	2	Yes	Suburb	Mansion/Estate
35	North America	3	No	City	Townhouse/Row House
0	Africa	2	No	Village	Detached House
200	Europe	4	Yes	Suburb	Mansion/Estate
5	Asia	1	Does not apply (in case of only 1 child in household)	City	Apartment/Flat
85	Europe	2	No	City	Detached House
0	Africa	1	Does not apply (in case of only 1 child in household)	Rural	Detached House
25	North America	2	No	City	Apartment/Flat
40	Europe	3	No	Suburb	Townhouse/Row House
120	Europe	2	Yes	City	Mansion/Estate
75	Europe	2	No	City	Detached House
15	North America	3	No	Suburb	Townhouse/Row House
30	Europe	1	Does not apply (in case of only 1 child in household)	City	Detached House
100	Asia	2	Yes	Suburb	Mansion/Estate
5	North America	2	No	City	Apartment/Flat
150	North America	4	Yes	Suburb	Mansion/Estate
0	Africa	1	Does not apply (in case of only 1 child in household)	Village	Detached House
8	Asia	2	No	City	Apartment/Flat
25	Europe	3	No	Suburb	Detached House
90	Europe	2	Yes	City	Mansion/Estate
40	Asia	1	Does not apply (in case of only 1 child in household)	City	Townhouse/Row House
0	Africa	3	No	Rural	Detached House
45	North America	2	No	Suburb	Detached House
80	Europe	4	Yes	City	Mansion/Estate
15	Asia	1	Does not apply (in case of only 1 child in household)	City	Apartment/Flat

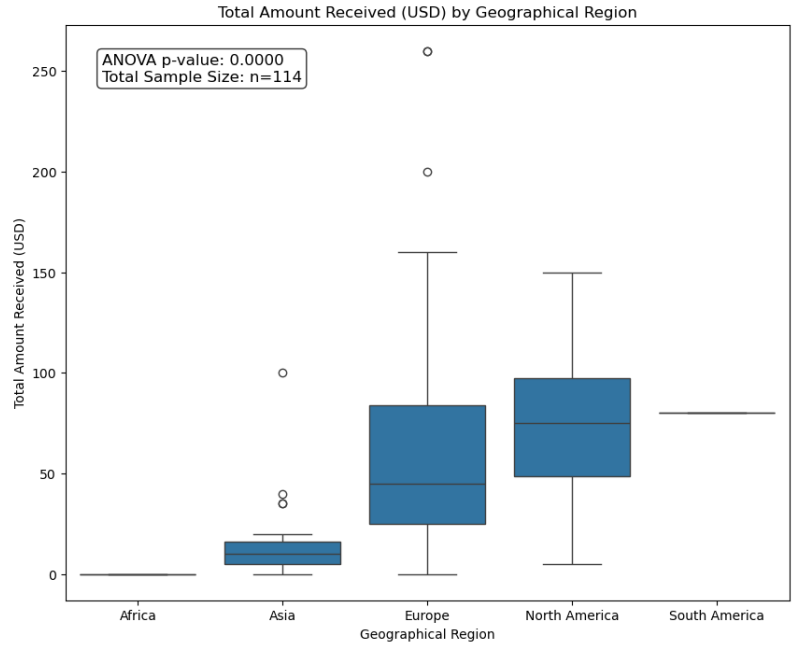
243 lines of Python code

```python
# Residence Type Analysis
# Check if each group has enough data
groups_residence = [survey_data[survey_data['Residence_Type'] == res_type]['Total_Amount_Received_USD'].dropna() for res_type in survey_data['Residence_Type'].unique()]
groups_residence = [group for group in groups_residence if len(group) > 0]  # Ensure no empty groups

if len(groups_residence) > 1:  # Ensure there is more than one group to compare
    f_val_residence, p_val_anova_residence = stats.f_oneway(*groups_residence)
    print('ANOVA result for Residence_Type: F-statistic = {}, p-value = {}'.format( *args: f_val_residence, p_val_anova_residence))
else:
    p_val_anova_residence = float('nan')
    print('Not enough groups for ANOVA test on Residence_Type')

# Calculate the total sample size
total_sample_size = len(survey_data)

# Plotting 'Amount_per_Child_USD' by 'Residence_Type'
plt.figure(figsize=(10, 8))
sns.boxplot(x='Residence_Type', y='Amount_per_Child_USD', data=survey_data)
plt.title('Amount per Child (USD) by Residence Type')
plt.xlabel('Residence Type')
plt.ylabel('Amount per Child (USD)')

# Adding the ANOVA p-value and total sample size to the plot
plt.text( x: 0.05,  y: 0.95,  s: f'ANOVA p-value: {p_val_anova_residence:.4f}\nTotal Sample Size: n={total_sample_size}',
         transform=plt.gca().transAxes, fontsize=12,
         verticalalignment='top', bbox=dict(boxstyle='round', facecolor='white', alpha=0.8))
```


Total Amount Received vs. Number of Children (n=114)
R^2: 0.00
P-value: 0.6157
Amount per Child USD
Number of Children
95% Confidence Interval


Correlation Matrix (n=114)
Amount_per_Child_USD
-0.18
0.21
0.2
0.08
-0.089
0.33
-0.094
0.22
-0.34
-0.047
0.49
-0.2
-0.048
Geographical_Region_Asia
Geographical_Region_Europe
Geographical_Region_North America
Geographical_Region_South America
Household_Location_City
Household_Location_Countryside
Household_Location_Rural
Household_Location_Suburb
Household_Location_Village
Residence_Type_Detached House
Residence_Type_Mansion/Estate
Residence_Type_Townhouse/Row House
Number_of_Children
Variables


Amount per Child (USD) by Residence Type
ANOVA p-value: 0.0000
Total Sample Size: n=114
Amount per Child (USD)
Apartment/Flat
Detached House
Mansion/Estate
Townhouse/Row House
Residence Type


Total Amount Received (USD) by Geographical Region
ANOVA p-value: 0.0000
Total Sample Size: n=114
Total Amount Received (USD)
Africa
Asia
Europe
North America
South America
Geographical Region

Eur Food Res Technol (2017) 243:1545–1554
DOI 10.1007/s00217-017-2863-9

ORIGINAL PAPER

The consistency in the sensory analysis of coffees using Q-graders

Lucas Louzada Pereira[1] · Wilton Soares Cardoso[2] · Rogério Carvalho Guarçoni[3] · Aymbiré Francisco Almeida da Fonseca[4] · Taís Rizzo Moreira[5] · Carla Schwengber ten Caten[6]

Purpose: This study aims to examine whether Q-graders provide consistent coffee scores under different conditions, specifically analyzing the effect of shift times (morning vs. afternoon) and dialogue/noise during tastings.

Problems:

- **Small Sample Size:** The study uses only 4 Q-graders, which is insufficient to generalize findings to the entire population of Q-graders.
- **Statistical Weakness:** With only two tasters per group, the dataset is too small for meaningful statistical analysis, rendering any conclusions about variability unreliable.
- **Misleading Title:** The title suggests a broader analysis of Q-grader consistency, but the study only examines intra-grader consistency under specific conditions, failing to address inter-grader variation.
- **Assumptions about Normality:** The use of Pearson's correlation implies assumptions of normality and linearity, which may not be appropriate for sensory data.

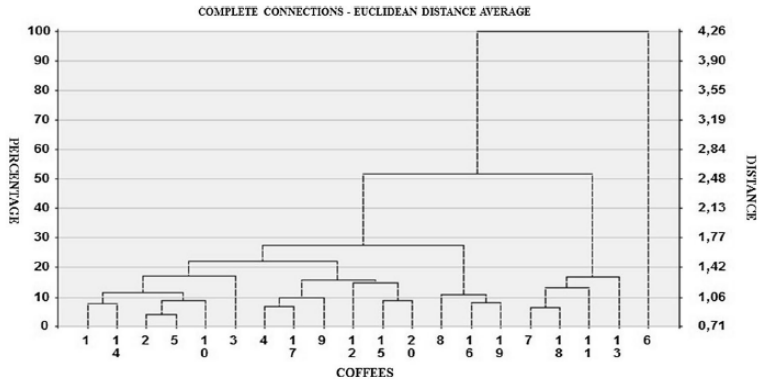


Fig. 1 Dendrogram among coffees obtained based on the sensory analyses carried out in the morning

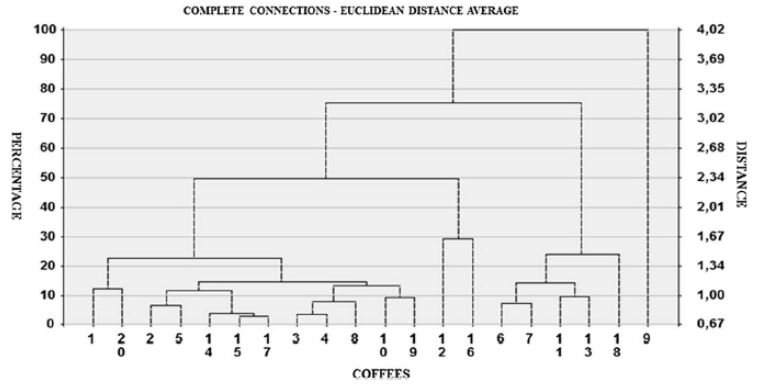


Fig. 2 Dendrogram among coffees obtained based on the sensory analyses carried out in the afternoon

Table 1 Averages and standard deviations for the characteristics: fragrance/aroma, uniformity, absence of defects, sweetness, flavor, acidity, body, aftertaste, balance, overall and total evaluation, before (morning) and after (afternoon)

Characteristics	Before (morning)		After (afternoon)	
	Average[a]	Standard deviation[b]	Average[a]	Standard deviation[b]
Fragrance/Aroma	7.5125 a	0.6952	7.8375 a	0.7663
Uniformity	9.7500 a	0.9104	9.7000 a	0.9787
Absence of defects	9.8000 a	0.8944	9.7000 a	0.9787
Sweetness	9.8000 a	0.8944	9.7000 a	0.9787
Flavor	7.7562 a	0.7018	7.4375 a	0.8316
Acidity	7.6500 a	0.7193	7.3875 a	0.7640
Body	7.5687 a	0.6804	7.5375 a	0.7653
Aftertaste	7.5062 a	0.6852	7.5938 a	0.6999
Balance	7.3312 a	0.5866	7.6750 a	0.7145
Overall	7.5125 a	0.7081	7.6250 a	0.7333
Total	82.1875 a	6.3541	82.1937 a	7.2955

[a]The averages of the attributes measured before and after followed by a letter in the line do not differ by the t test at 5% probability

[b]The standard deviation before and after

Received: 9 May 2022 | Revised: 10 November 2022 | Accepted: 16 November 2022
DOI: 10.1111/joss.12805

REVIEW

Journal of Sensory Studies WILEY

Adherence and concordance among Q-Graders in the sensory analysis of coffees

Deusélio Bassini Fioresi[1] | Alessandro Coutinho Ramos[2] | Amanda Azevedo Bertolazi[2] | Lucas Louzada Pereira[1]

Purpose: This study examines the correlation and concordance among Q-graders in sensory evaluation, aiming to understand the degree of agreement in their scoring and how it relates to the Q-grader calibration system.

Problems:

- **Sample Size Limitation:** With only 6 Q-graders, the study cannot provide reliable conclusions about the broader Q-grader population.
- **Flawed Statistical Assumptions:** The study's reliance on correlation and concordance measures without a larger, representative sample leads to weak conclusions. Assumptions about Q-graders being a homogeneous group are not supported by data.
- **Overestimation of Calibration:** The assumption that Q-graders are inherently calibrated and that this small sample can represent the entire system is flawed. There is a need to first measure broader inter-Q-grader variation.
- **Methodological Oversights:** The study does not sufficiently address the diversity and potential variability among Q-graders, which could significantly impact the findings and their generalizability.

2.3 | Shapiro–Wilk test

The Shapiro–Wilk (SW) test (Shapiro & Wilk, 1965) is a test of data adherence to the normal probability distribution and is indicated in cases of small samples ($n < 30$) (Fávero, Belfiore, Silva, & Chan, 2009). This test was used in this study due to the comparison among the tasters, who performed 25 cup tastings each.

The test statistic is denoted by "W" and obtained by:

$$W = \frac{b^2}{\sum_{i=1}^{n} (x_i - \overline{X})^2}$$

for $i = 1, 2, 3, \ldots, n$, and

$$b = \sum_{i=1}^{m} a_{n-i+1} \times (x_{n-i+1} - x_i)$$

where $m = \frac{n}{2}$ if n is even, or $m = \frac{n+1}{2}$ if n is odd and the terms a_{n-i+1} are constants generated by means, variances and covariances of the order statistics for samples of the same size "n" from a normal distribution.

TABLE 3 Pearson's coefficient (ρ)

		ρ: Arabica					
		nTP1	nTP2	nTP3	nTP4	nTP5	nTP6
Conilon: ρ	nTP1		0.16	0.47[b]	0.53[a]	0.38	0.34
	nTP2	0.38		0.04	0.53[a]	0.64[a]	0.50[b]
	nTP3	0.56[a]	0.49[b]		0.30	0.10	0.36
	nTP4	0.42[b]	0.50[b]	0.63[a]		0.54[a]	0.56[a]
	nTP5	0.31	0.44[b]	0.51[a]	0.55[a]		0.42[b]
	nTP6	0.49[b]	0.29	0.31	0.26	0.26	

Note: nTPj: Total score of the taster j, for $j = 1, 2, \ldots, 6$.
[a]Significant at 1%.
[b]significant a 5%.

TABLE 4 Lin's coefficient (ρ_c)

		ρ_c: Arabica					
		nTP1	nTP2	nTP3	nTP4	nTP5	nTP6
Conilon: ρ_c	nTP1		0.14	0.43[b]	0.43[a]	0.27	0.33
	nTP2	0.32		0.03	0.51[a]	0.56[a]	0.48[b]
	nTP3	0.55[a]	0.45[b]		0.29	0.09	0.35
	nTP4	0.42[b]	0.41[b]	0.61[a]		0.53[a]	0.52[a]
	nTP5	0.30	0.31[b]	0.46[a]	0.54[b]		0.36[b]
	nTP6	0.46[b]	0.15	0.22	0.22	0.23	

Note: nTPj: Total score of the taster j, for $j = 1, 2, \ldots, 6$.
[a]Significant at 1%.
[b]Significant a 5%.

Hindawi
Journal of Food Quality
Volume 2018, Article ID 3285452, 7 pages
https://doi.org/10.1155/2018/3285452

WILEY | Hindawi

Research Article

Propositions on the Optimal Number of Q-Graders and R-Graders

Lucas Louzada Pereira,[1] **Rogério Carvalho Guarçoni**,[2] **Gustavo Soares de Souza**,[2] **Dério Brioschi Junior**,[3] **Taís Rizzo Moreira**,[4] **and Carla Schwengber ten Caten**[1]

Purpose: This article seeks to determine the optimal number of Q-graders necessary for reliable sensory evaluation, using bootstrap simulations to estimate the required number for accurate sensory tests.

Problems:

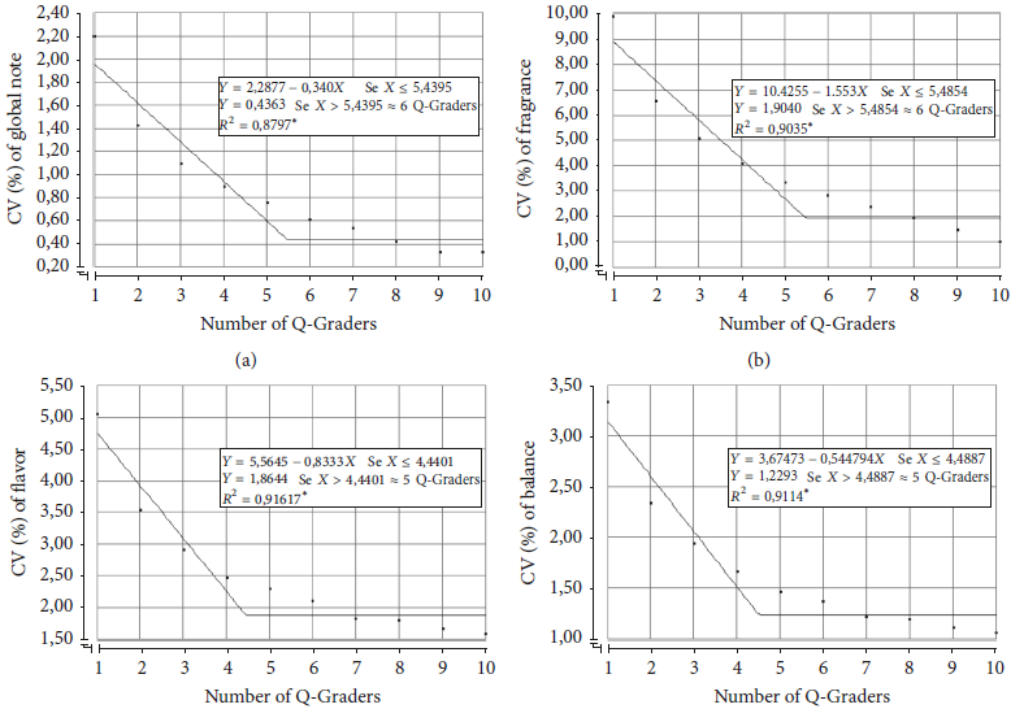
- **Bootstrap Simulation Assumptions:** While bootstrap methods are used to estimate confidence intervals and variability, the initial sample of 6 Q-graders may not be representative, leading to biased results.
- **Inadequate Description of Participants:** The lack of detailed profiles for the Q-graders undermines the reliability of the conclusions about the optimal number of graders needed.
- **Overgeneralization:** The study assumes the variation measured in a small group can be extrapolated to the larger population, which is a flawed approach without a large, representative sample.
- **Precision Concerns:** The significant reduction in the coefficient of variation with increasing graders is not adequately explored, and the practical implications of these findings are not clearly discussed.

2.4. Ideal Number of Tasters. For the grouping of the pair numbers of tasters and their respective coefficients of variation $[X, CV(X)]$, we used the bootstrap method, where 1000 sample simulations were performed with 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 tasters [13].

In order to determine the optimal size of the tasting room, the linear regression method of plateau response was used [14]. The optimal number of tasters was proposed when the linear model becomes a plateau:

$$Y_i = \begin{cases} \beta_0 + \beta_1 X_i + \varepsilon_i & \text{se } X_i \leq X_0 \\ P + \varepsilon_i & \text{se } X_i > X_0. \end{cases} \quad (1)$$

Here Y_i is the response variable, β_0 is the linear coefficient of the linear model of the segment before the plateau, β_1 is the angular coefficient of this same segment, ε_i is the error associated with the ith observation, P is the plateau, and X_0 is the point of attachment of the two segments. P and X_0 should be estimated.



Received: 19 June 2018 | Revised: 11 January 2019 | Accepted: 18 January 2019
DOI: 10.1111/jtxs.12390

RESEARCH ARTICLE

Journal of Texture Studies WILEY

Very beyond subjectivity: The limit of accuracy of Q-Graders

Lucas Louzada Pereira[1] | Rogério Carvalho Guarçoni[2] | Taís Rizzo Moreira[3] | Luiz Henrique Bozzi Pimenta de Sousa[4] | Wilton Soares Cardoso[4] | Aldemar Polonini Moreli[4] | Samuel Ferreira da Silva[3] | Carla Schwengber ten Caten[5]

Purpose: The goal is to evaluate the accuracy and efficiency of Q-graders in distinguishing specialty coffees from non-specialty ones, assessing their precision in scoring different quality levels.
Problems:
•**Limited Sample Size:** Only 7 Q-graders are involved, which is too few to reliably infer about the wider population of Q-graders.
•**Unexpected Findings:** The claim that Q-graders show high precision for top-quality coffees but make errors on lower-quality ones is counterintuitive and likely influenced by the small sample size.
•**Subjectivity and Representation:** The study does not address the potential for significant subjectivity and variation among Q-graders, nor does it ensure a representative sample.
•**Methodological Issues:** There is no robust justification for the inclusion of specific Q-graders or their sensory skills, weakening the study's validity.

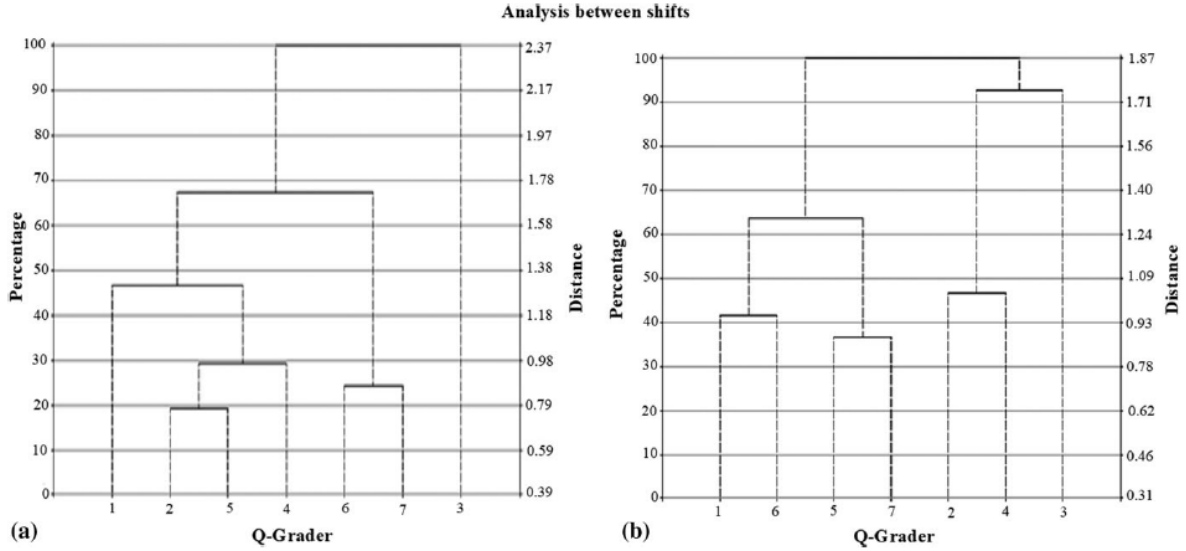


FIGURE 5 Dendrograms obtained between the Q-Graders from the sensory analysis performed between all the replicates shifts based on the paired *t*-test—morning (4A) and afternoon (4B) relative to Experiment 2 on the second day

TABLE 9 Averages of the sensory variables evaluated in the Experiment 3 in the samples of the morning and afternoon shifts

Sensory characteristics and standard deviation			Sensory characteristics and standard deviation		
Attributes	Average ± *SD* (*n* = 7 Q-Graders) [a]		Attributes	Average ± *SD* (*n* = 7 Q-Graders)[a]	
	Sample 1 (morning)	Sample 2 (morning)		Sample 1 (afternoon)	Sample 2 (afternoon)
Fragrance	7.46 A/0.5089	7.18 A/0.3740	Fragrance	7.07 A/0.5147	6.96 A/0.4661
Flavor	7.39 A/0.3493	7.36 A/0.2835	Flavor	7.14 A/0.3780	7.07 A/0.1220
Aftertaster	7.14 A/0.2440	7.07 A/0.3134	Aftertaster	6.79 A/0.4880	6.71 A/0.466
Acidity	7.18 A/0.5147	7.14 A/0.3780	Acidity	6.86 A/0.3780	6.82 A/0.5147
Body	7.07 A/0.5345	7.39 A/0.5175	Body	6.82 A/0.4261	6.79 A/0.5669
Uniformity	10	10	Uniformity	10	10
Balance	7.43 A/0.3450	7.50 A/0.4330	Balance	7.11 A/0.4532	6.89 A/0.4046
Clean cup	10	10	Clean cup	10	10
Sweetness	10	10	Sweetness	10	10
Overall	7.14 A/0.4046	7.39 A/0.2835	Overall	7.04 A/0.5089	6.96 A/0.3660
Global score	80.82 A/1.5660	81.04 A/1.6673	Global score	78.82 A/1.7000	78.21 A/2.0383

[a] The averages of the characters measured before and after, followed by the same letter in the line do not differ according to the *t*-test, at 5% probability.

Descriptive Cupping – a Rapid Coffee Flavour Profiling Method Using the Specialty Coffee Association of America (SCAA) Cupping Protocol

M.R. FERNANDEZ-ALDUENDA, K. LUSK, P. SILCOCK, J. BIRCH

University of Otago, Department of Food Science,
PO Box 56, Dunedin 9054, New Zealand.

Purpose: The study aims to investigate whether the SCAA Cupping Protocol, traditionally used for quality assessment, can be adapted for descriptive flavor profiling. It evaluates natural coffee samples to see if descriptive sensory data can be obtained rapidly using this protocol, proposing it as a cost-effective method for regions without advanced sensory facilities.

Problems:

- **Small Number of Assessors:** The study involves only 7 cuppers, an insufficient sample size for generalizing findings to the broader population of coffee cuppers, limiting the robustness of the conclusions.
- **Cost Misrepresentation:** While the article implies the SCAA cupping form is cost-efficient, the education required for Q-graders is expensive. Moreover, the actual descriptive analysis conducted in the study involves rigorous training and advanced statistical techniques, making it as costly and complex as any other scientific sensory methodology, thereby negating any proposed cost savings.
- **Discrepancy with Title:** The study title suggests using the SCAA cupping protocol for descriptive profiling. However, the actual research primarily uses a separate descriptive analysis with additional training in flavor descriptors, failing to demonstrate that the SCAA cupping form alone can be effectively used for descriptive flavor profiling.
- **Insufficient Validation:** The study does not robustly validate the extracted descriptive data against traditional descriptive sensory analysis methods, raising concerns about the reliability and accuracy of the flavor profiles generated.
- **Incomplete Integration:** There is a lack of strong correlation or integration between the descriptive data and SCAA quality scores, undermining the claim that the SCAA protocol can serve as a standalone method for descriptive analysis.

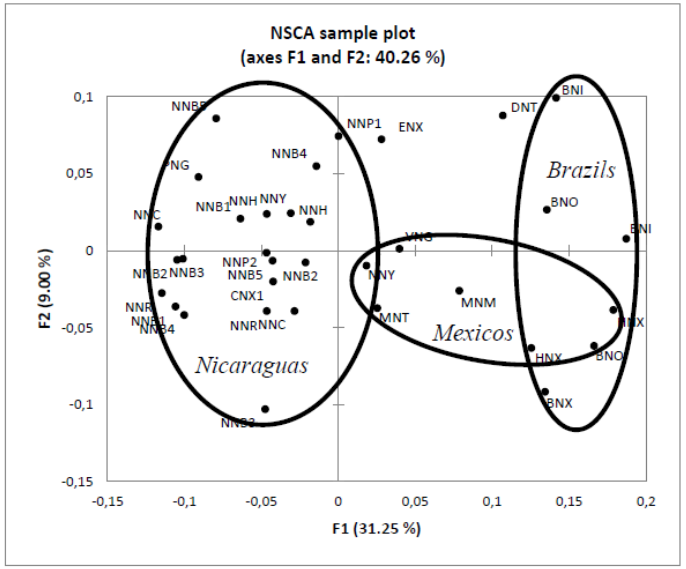


Figure 1. Non-symmetrical correspondence analysis (NSCA) map representing the projection on F1 and F2 of natural coffees for 22 samples, evaluated using the descriptive cupping method. Samples with the same name are duplicates.

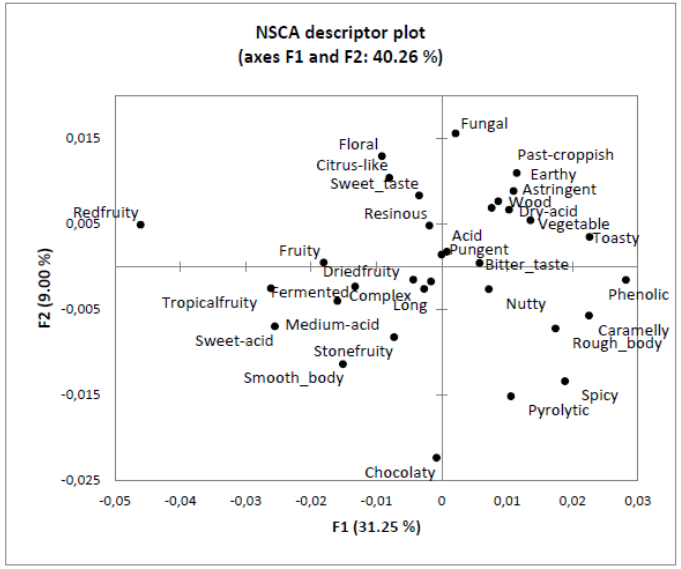


Figure 2. Non-symmetric correspondence analysis (NSCA) map representing the projection on F1 and F2 of descriptor categories for 22 coffee samples, evaluated using the descriptive cupping method.

How Do Cuppers Cup?

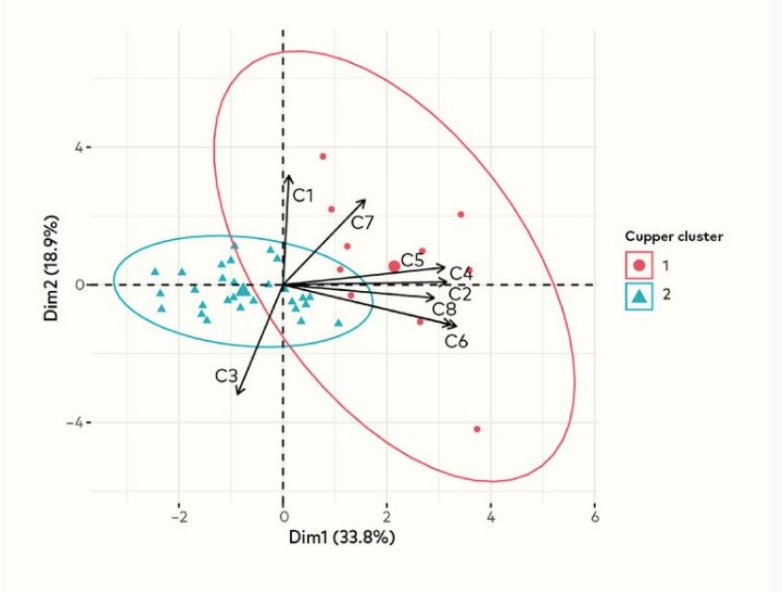
Evaluating and Evolving Elements of the SCA Cupping Protocol

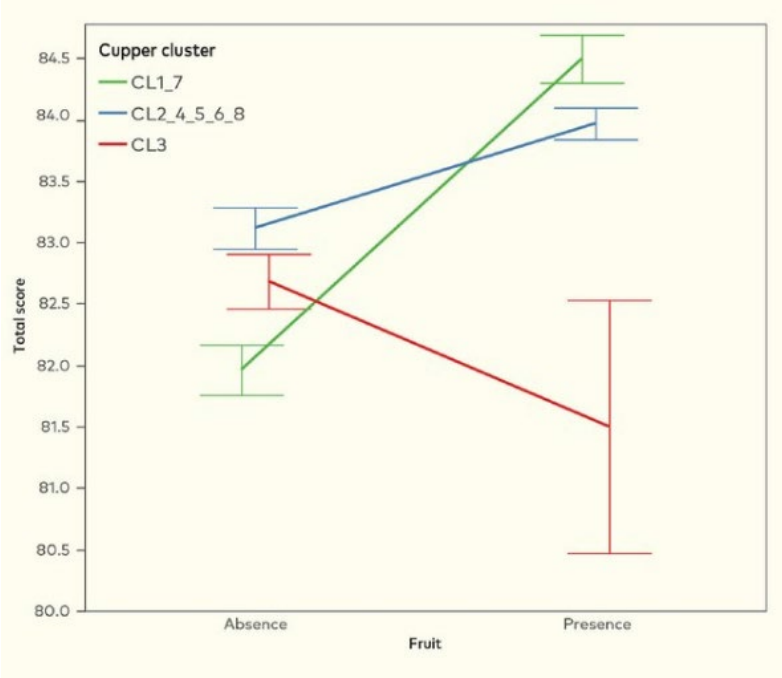
Dr. JORGE BERNY and Dr. MARIO FERNÁNDEZ-ALDUENDA share initial results of a collaborative study examining how cuppers cup and exploring the potential impacts of a proposed component of the reengineered cupping protocol.

Purpose: The study examines how effectively cuppers use the SCA cupping protocol and explores improvements by testing a descriptive “check-all-that-apply” (CATA) component for flavor attributes. It aims to assess the alignment of cuppers' evaluations and refine the cupping protocol to better capture coffee quality.

Problems:

- **Small Sample Size:** Only 8 cuppers were used, limiting the ability to generalize findings. Disagreement among such a small group suggests even greater variability and lack of consensus in the broader Q-grader population.
- **Internal Inconsistency:** Significant variability in scores among the 8 cuppers indicates a lack of agreement, questioning the reliability of the protocol across a larger population.
- **Outdated Method Presentation:** The article claims the use of "recent advances" in sensory science, yet it applies methods established in the 1970s by Rose Marie Pangborn, potentially misleading readers about the novelty of the approach.
- **Limited Descriptive Analysis:** The integration of descriptive data without employing a traditional descriptive sensory analysis limits the depth of the sensory insights, reducing the effectiveness of the proposed improvements to the SCA protocol.
- **Weak statistical power:** Not impressive Dim1 and Dim2 % of explained variance





Rose Marie Pangborn, credited as being the leading pioneer within sensory science started her career at UC Davis in 1955 (Meiselman et al., 2022), wrote the fundamental learning book: *Principles of Sensory Evaluation of Food* by Rose Marie Pangborn, 1965. […] Two other fundamental books to mention are: *Sensory Evaluation Techniques* by Meilgaard, Carr, and Civille, 1987; principles *of Sensory Evaluation Handbook* by Heymann, 1998.

All three books have served as fundamental resources for sensory professionals, providing practical methodologies and valuable insights into sensory analysis. Many different sensory methods have been developed hereof the Time Intensity method, 1954, Qualitative Descriptive Analysis (QDA), 1974, Free-Choice profiling, 1984, Napping 2012, CATA, 2007 and Rate all that apply (RATA) in 2014. These milestones show sensory science's ongoing development, with improvements in methodology, education, and professional networking contributing to the growth of sensory science in various industries.
[…]
Despite these efforts, there are still ambiguities as the hedonic and descriptive scoresheets can be combined; when experts score the coffees, and not consumers, the approach is irrelevant for preference predictions. Instead of pointing back to consumer preferences that drive purchasing behavior and ultimately determine the coffee's value, the focus remains on the coffee professionals' personal opinion.
This thesis aims to add to the conversation and encourage the use of more scientific approaches in the coffee industry.

- No universal quality yard stick
- Small differences are not relevant

Laboratory analysis: DESCRIPTIVE

Consumer testing: PREFERENCES

Descriptive analysis

Overall impression

Intensity

A little — A lot

Basic taste

Acidic

A little — A lot

Sweet

Nothing — A lot

Bitter

A little — A lot

Flavour

Burnt

A little — A lot

Tobacco

Nothing — A lot

Licorice

Nothing — A lot

Chocolate

Nothing — A lot

Dark berries

Nothing — A lot

Consumer survey

How much do you like this coffee?

Dislike strongly — Neither like nor dislike — Like Strongly

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

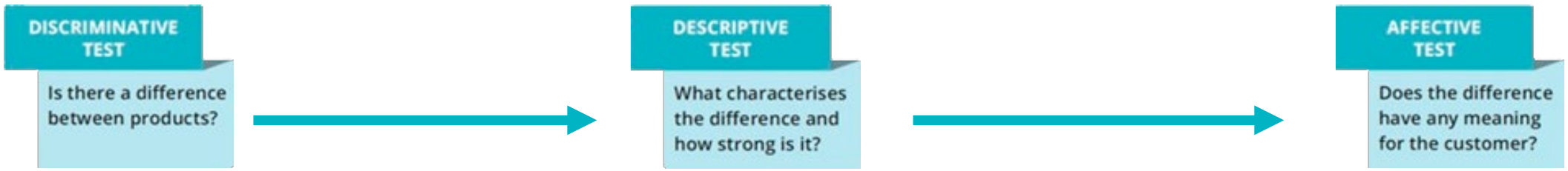
Place a mark by <u>all</u> the words that you think describe this coffee.

- ☐ Sharp
- ☐ Harmonic
- ☐ Intense
- ☐ Bland
- ☐ Thin
- ☐ Roasted ryebread
- ☐ Astringent
- ☐ Tobacco
- ☐ Caramel
- ☐ Long aftertaste
- ☐ Hey/straw
- ☐ Bitter
- ☐ Grass
- ☐ Sweet
- ☐ Earthy
- ☐ Citrus
- ☐ Mild
- ☐ Licorice
- ☐ Complex
- ☐ Dark berries (e.g. black currant or blue berries)
- ☐ Delicate
- ☐ Acidic
- ☐ Rich
- ☐ Chocolate
- ☐ Strong/powerfull
- ☐ Pleasent
- ☐ Unpleasent
- ☐ Nutty
- ☐ Balanced
- ☐ Burnt

It's scientific, specific, sustainable, and specialty. And it's yours to use.

Be a part of the redefinition of specialty coffee.

The first step then – according to Pangborn would be to focus on reminding the general community with NOT wasting time scoring differences that they can't pick up in a discriminative test, no?



SCA Coffee Value Assessment

Descriptive Form

Name .....

Date .....

Purpose .....

SAMPLE NO. _____ **ROAST LEVEL** _____

Fragrance Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Aroma Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Notes

☐ Floral
☐ Fruity ☐ Berry ☐ Dried Fruit ☐ Citrus Fruit
☐ Sour/Fermented ☐ Sour ☐ Fermented
☐ Green/Vegetative
☐ Other ☐ Chemical ☐ Musty/Earthy ☐ Woody
☐ Roasted ☐ Cereal ☐ Burnt ☐ Tobacco
☐ Nutty/Cocoa ☐ Nutty ☐ Cocoa
☐ Spice
☐ Sweet ☐ Vanilla/Vanillin ☐ Brown Sugar

Flavor Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Aftertaste Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Notes

☐ Floral
☐ Fruity ☐ Berry ☐ Dried Fruit ☐ Citrus Fruit
☐ Sour/Fermented ☐ Sour ☐ Fermented
☐ Green/Vegetative
☐ Other ☐ Chemical ☐ Musty/Earthy ☐ Woody
☐ Roasted ☐ Cereal ☐ Burnt ☐ Tobacco
☐ Nutty/Cocoa ☐ Nutty ☐ Cocoa
☐ Spice
☐ Sweet ☐ Vanilla/Vanillin ☐ Brown Sugar

Main Tastes (2)
☐ Salty ☐ Bitter
☐ Sour ☐ Umami
☐ Sweet

Acidity Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Notes

Sweetness Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Notes

Mouthfeel Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

Notes

☐ Rough (Gritty, Chalky, Sandy) ☐ Smooth (Velvety, Silky, Syrupy) ☐ Metallic
☐ Oily ☐ Mouth-Drying

Affective Form

Date .....

Purpose .....

IMPRESSION OF QUALITY
① EXTREMELY LOW ② VERY LOW ③ MODERATELY LOW ④ SLIGHTLY LOW ⑤ NEITHER HIGH NOR LOW ⑥ SLIGHTLY HIGH ⑦ MODERATELY HIGH ⑧ VERY HIGH ⑨ EXTREMELY HIGH

SAMPLE NO. _____

Fragrance ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Aroma ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Flavor ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Aftertaste ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Acidity ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Sweetness ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

SAMPLE NO. _____

Fragrance ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Aroma ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Flavor ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Aftertaste ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Acidity ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Sweetness ① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Concerns Regarding the New SCA Coffee Value Assessment Protocol

The distinction between generic (homogeneous) and differentiated products was most clearly introduced in "The Theory of Monopolistic Competition" (1933) by Edward Chamberlin and "Economics of Imperfect Competition" (1933) by Joan Robinson. Both works, published independently in the same year, laid the groundwork for understanding product differentiation in economic theory.

Hedonic/affective values are consumer PREFERENCES. Not expert opinion but data that predicts how easy it is for you to sell your coffee in your local market. So having “IMPRESSION OF QUALITY ” as the parameter on the Affective form is in direct and extreme violation of the whole point behind Rose Marie Pangborn’s division of Discriminative (expert), Descriptive (expert) and Hedonic (consumer).

Quality of ‘acidity’, ‘sweetness’. Consumer are naïve (not trained and should not score attributes other than ‘liking’). CATA is often used in consumer studies – not descriptive analysis!

Not having 'Bitterness’ as a sensory attribute is a sign that it still hangs in the old 2004 form universe and does not map one of the most important attribute variations and consumer preference differentiation parameters.

Bias is avoided by “To avoid introducing bias, no assessors undertaking the physical, sensory descriptive, or affective assessments should compile this information until after those assessments have been completed.“ But it is still the same assessors undertaking it just separated in time. Not the real end users/buyers?

All data are Objective! If they are repeatable by anybody using the same method on the same objects reporting 1st person data (intensities or preferences)

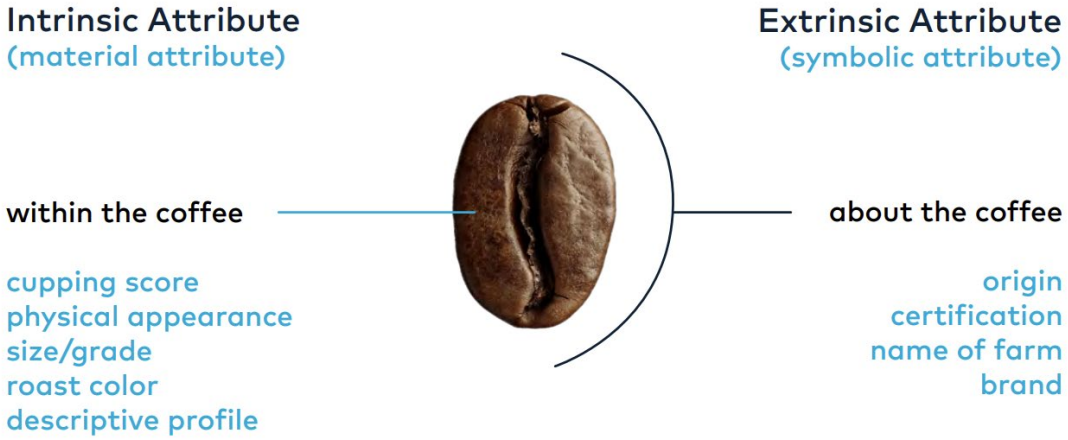
1970'ies

Advances in sensory science frame the "objectivity" or "subjectivity" question faced by cuppers with the 2004 SCA Cupping System as two separate and distinct sensory tests. Analytical quality measurements, like taste intensity or body level, are understood as objective (i.e., trainable with sensory references such as the World Coffee Research Sensory Lexicon). Value judgement (like grade, preference, liking, or acceptability) is subjective. In the 2004 SCA Cupping Form, these kinds of tests are mixed alongside several "discriminatory" (yes/no) sensory tests. This explains why the form's users were so split as to whether or not the form was subjective or objective.

1930’ies – not 2015 as cited in A System to Assess Coffee Value

The dominant theory of value in modern economics is termed the “subjective theory of value.”42 In simple terms, the subjective theory states that the value of a good or service is derived from the subjective needs and preferences of the individual who is judging it. This theory is a key concept in modern economic thinking and explains why markets behave the way they do—for example, why a particular product might fetch different prices in different situations.

A System to Assess Coffee Value, June 2024



Towards a Deinition of Specialty Coffee, An SCA White Paper 2021

	INNER	OUTER
INDIVIDUAL	**PSYCHOLOGY (UL)** Thoughts, feelings	**THINGS (UR)** Coffee tree, bean, roast profile, brew, cup, science facts
COLLECTIVE	**CULTURE (LL)** Group identity, tribe membership, ethics	**SYSTEMS (LR)** Protocols, workflow, recipes, QC

Fig. 20 Ken Wilber’s Integral Approach.
Psychology (UL) = Upper Left
Culture (LL) = Lower Left
Things (UR) = Upper Right
Systems (LR) = Lower Right

Table 4 Target readings of cupping level roast for different roast meters

Instrument or Scale	Target reading at cupping roast level
Agtron "Gourmet"*	63.0
Agtron "Commercial"	48.0
Colorette 3b by Probat	96.0
Colortrack:	62.0

Hence, the cupping score is now the result of the following equation, rounded to the nearest 0.25 points:

$$S = 0.65625 \sum_{i=1}^{i=8} h_i + 52.75 - 2u - 4d$$

Where:

S is the cupping score prior to rounding

h_i is the 9-point score of each affective section, from $i = 1$ (fragrance) to $i = 8$ (overall)

u is the number of non-uniform cups

d is the number of defective cups

SCA Coffee Value Assessment

Combined Form

Name ..... Date .....

Purpose ..... Sample No. .....

IMPRESSION OF QUALITY

① EXTREMELY LOW ④ SLIGHTLY LOW ⑦ MODERATELY HIGH
② VERY LOW ⑤ NEITHER HIGH NOR LOW ⑧ VERY HIGH
③ MODERATELY LOW ⑥ SLIGHTLY HIGH ⑨ EXTREMELY HIGH

PART 1: DESCRIPTIVE ASSESSMENT

ROAST LEVEL

PART 2: AFFECTIVE ASSESSMENT

Fragrance Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Aroma Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

☐ Floral
☐ Fruity ☐ Berry ☐ Dried Fruit ☐ Citrus Fruit
☐ Sour/Fermented ☐ Sour ☐ Fermented
☐ Green/Vegetative
☐ Other ☐ Chemical ☐ Musty/Earthy ☐ Woody
☐ Roasted ☐ Cereal ☐ Burnt ☐ Tobacco
☐ Nutty/Cocoa ☐ Nutty ☐ Cocoa
☐ Spice
☐ Sweet ☐ Vanilla/Vanillin ☐ Brown Sugar

Notes

Notes

Flavor Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Aftertaste Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

☐ Floral
☐ Fruity ☐ Berry ☐ Dried Fruit ☐ Citrus Fruit
☐ Sour/Fermented ☐ Sour ☐ Fermented
☐ Green/Vegetative
☐ Other ☐ Chemical ☐ Musty/Earthy ☐ Woody
☐ Roasted ☐ Cereal ☐ Burnt ☐ Tobacco
☐ Nutty/Cocoa ☐ Nutty ☐ Cocoa
☐ Spice
☐ Sweet ☐ Vanilla/Vanillin ☐ Brown Sugar

Main Tastes (2)
☐ Salty
☐ Sour
☐ Sweet
☐ Bitter
☐ Umami

Notes

Notes

Acidity Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

Notes

Notes

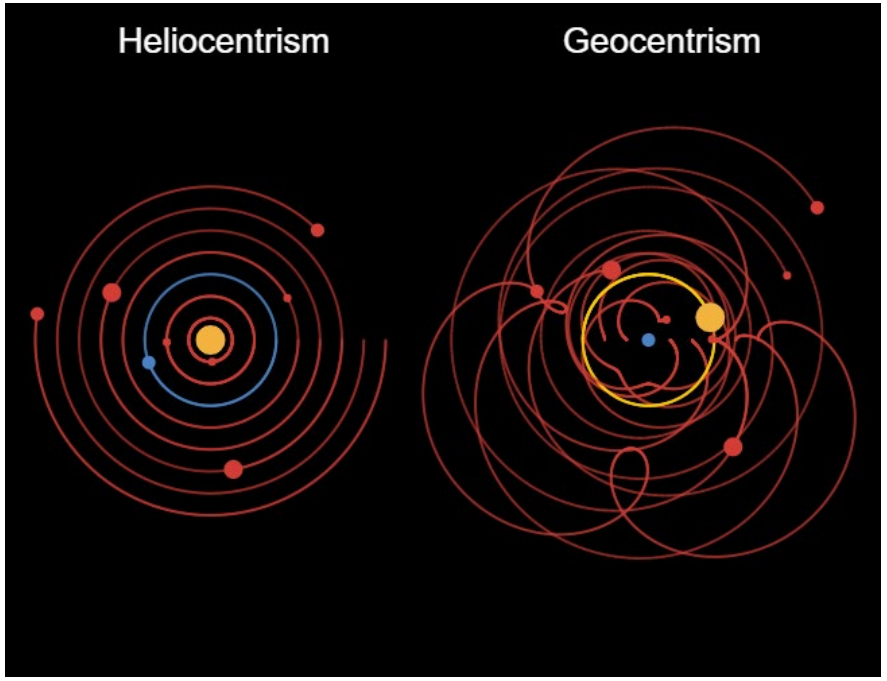
Sweetness Intensity — LOW / MEDIUM / HIGH — 0 5 10 15

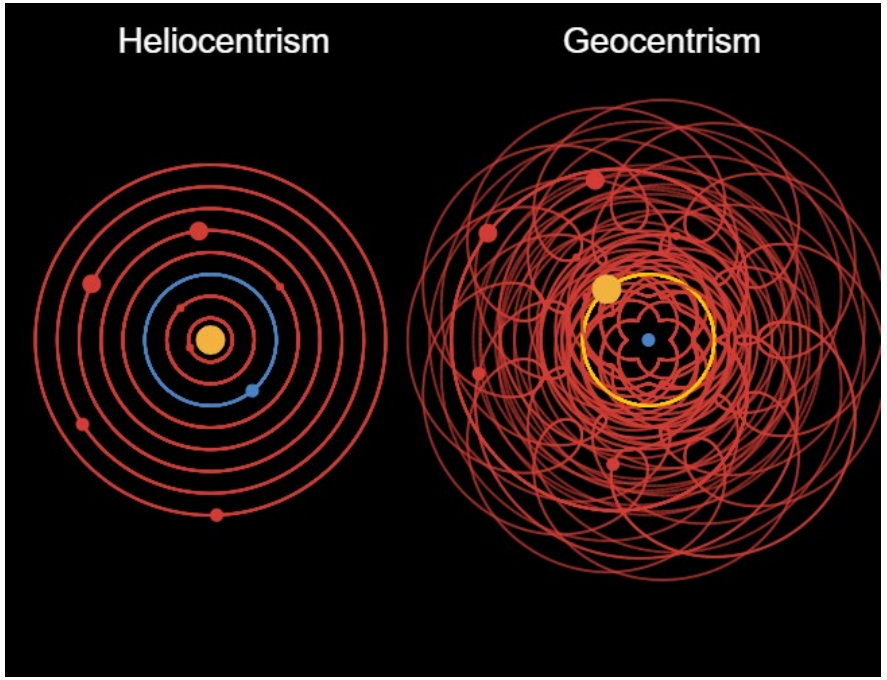
① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ FINAL

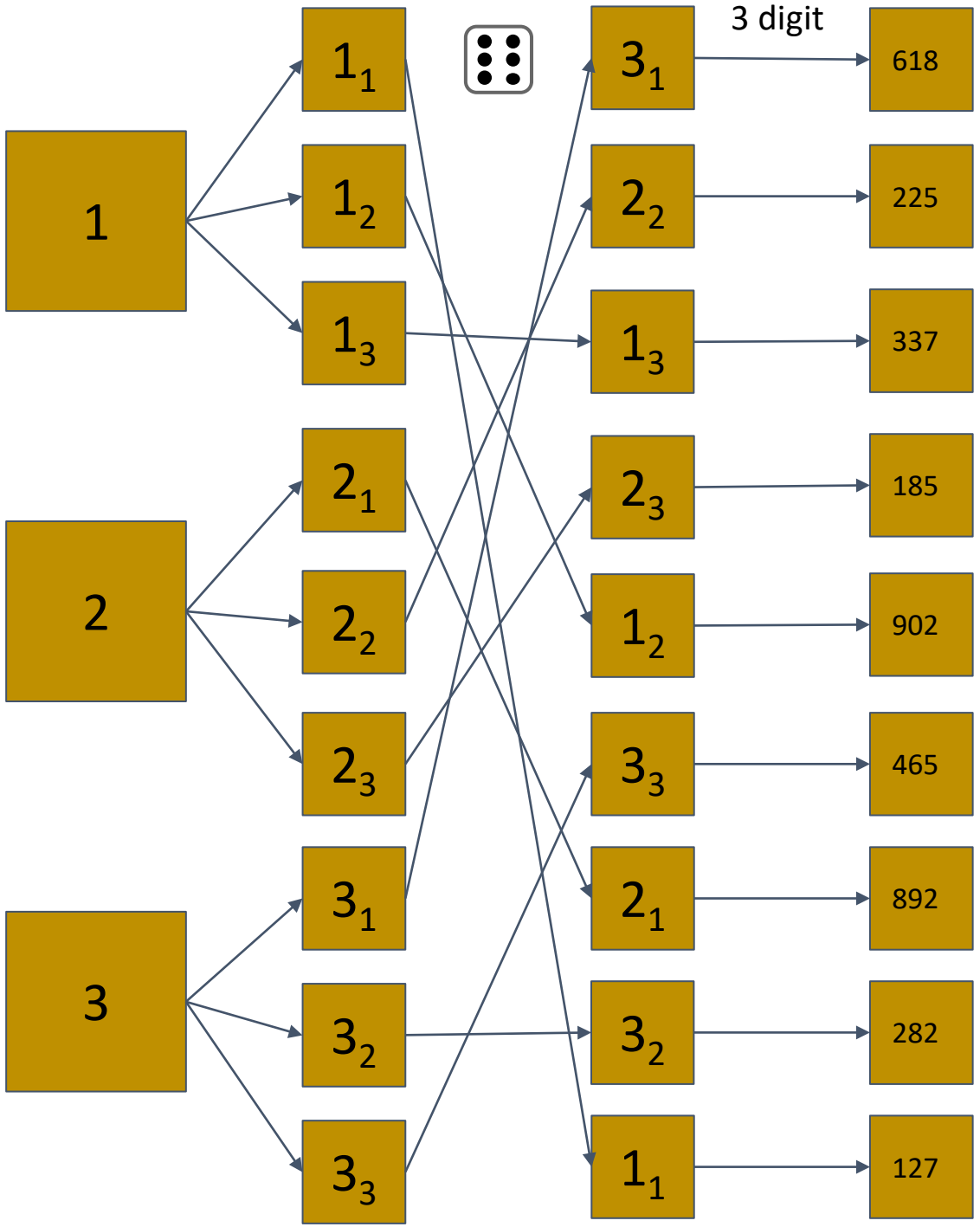
Notes

Notes

How does a good theory look like?

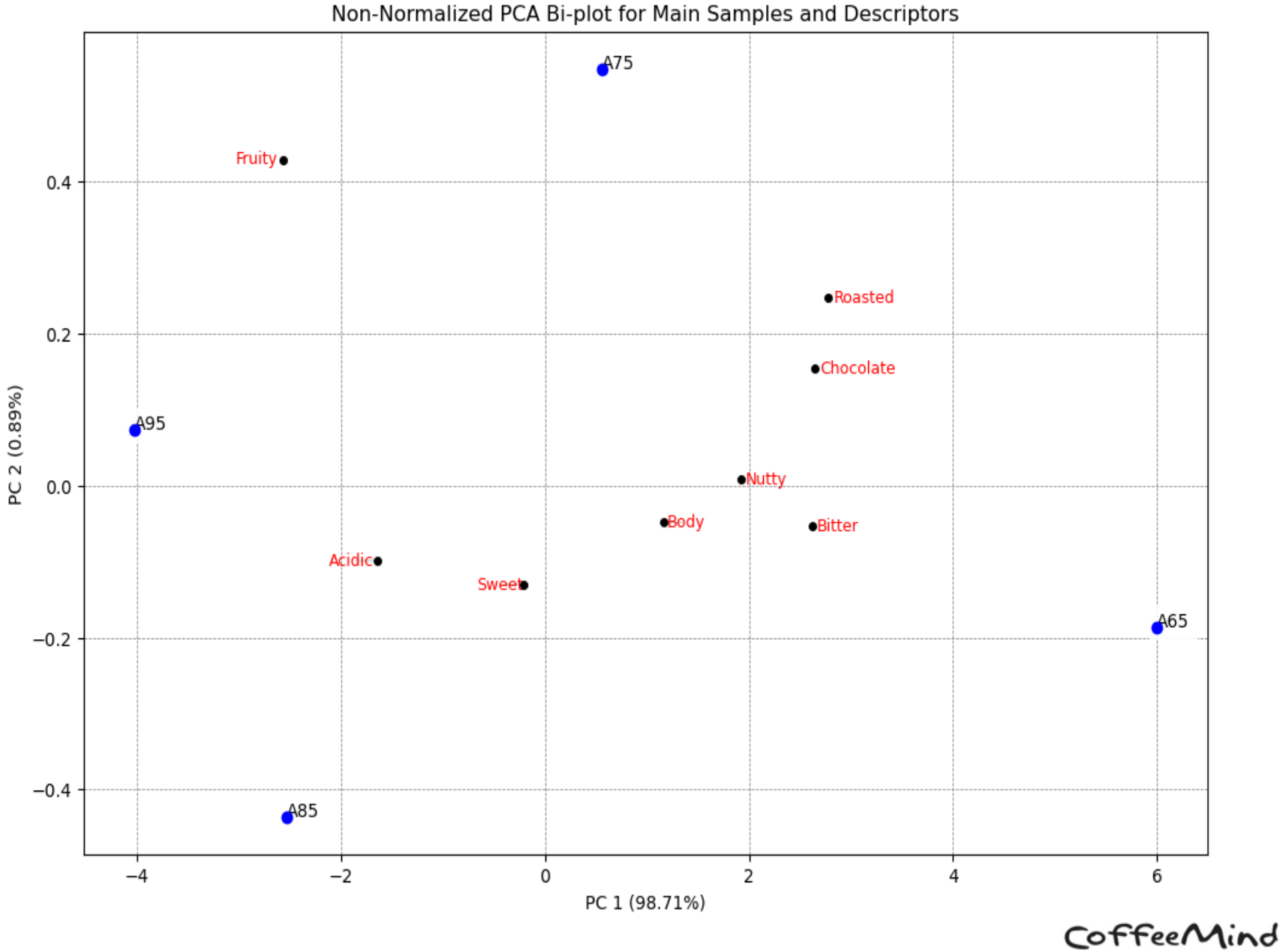




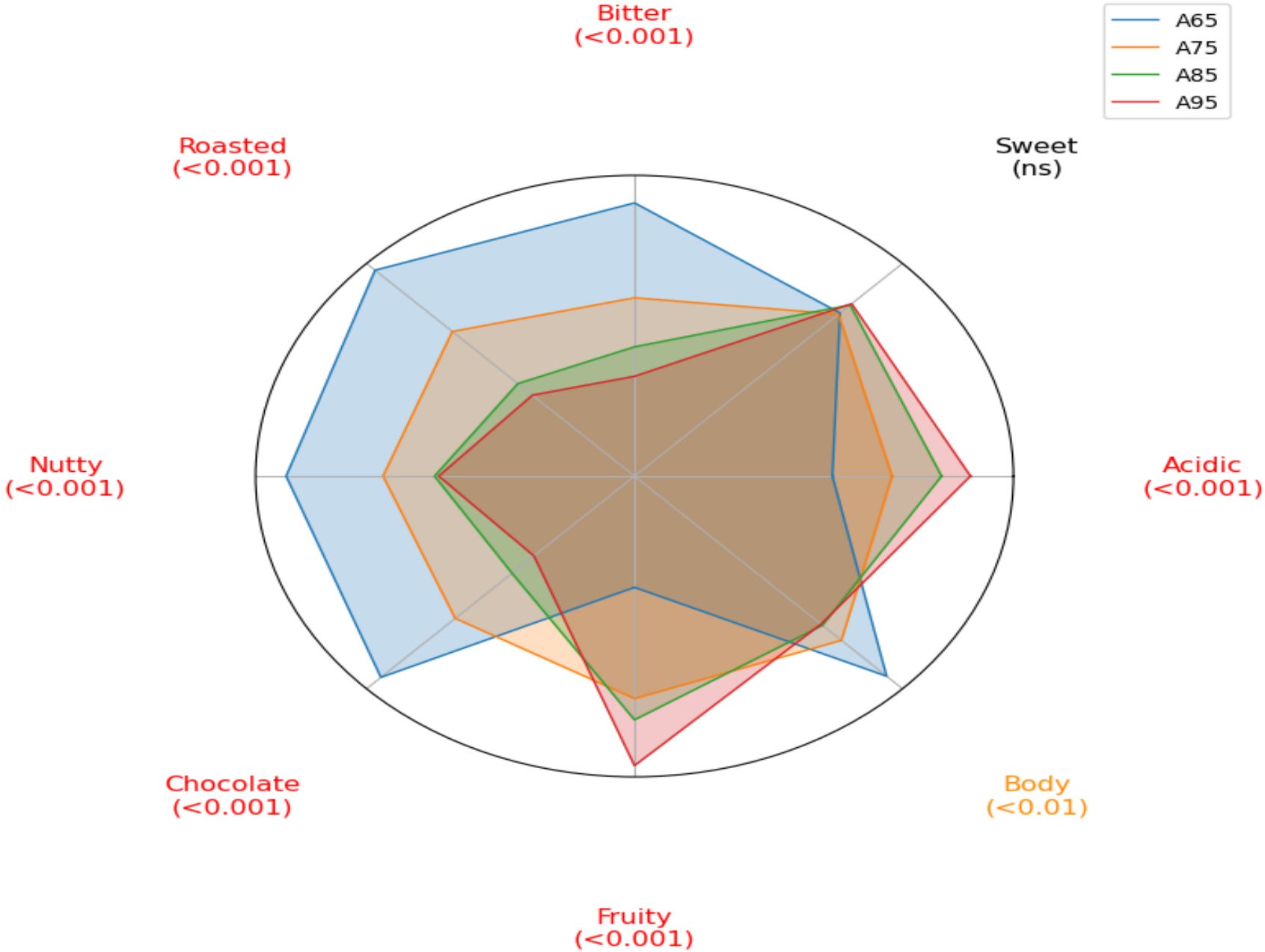


10-20 descriptors pr sample scored from 0-15

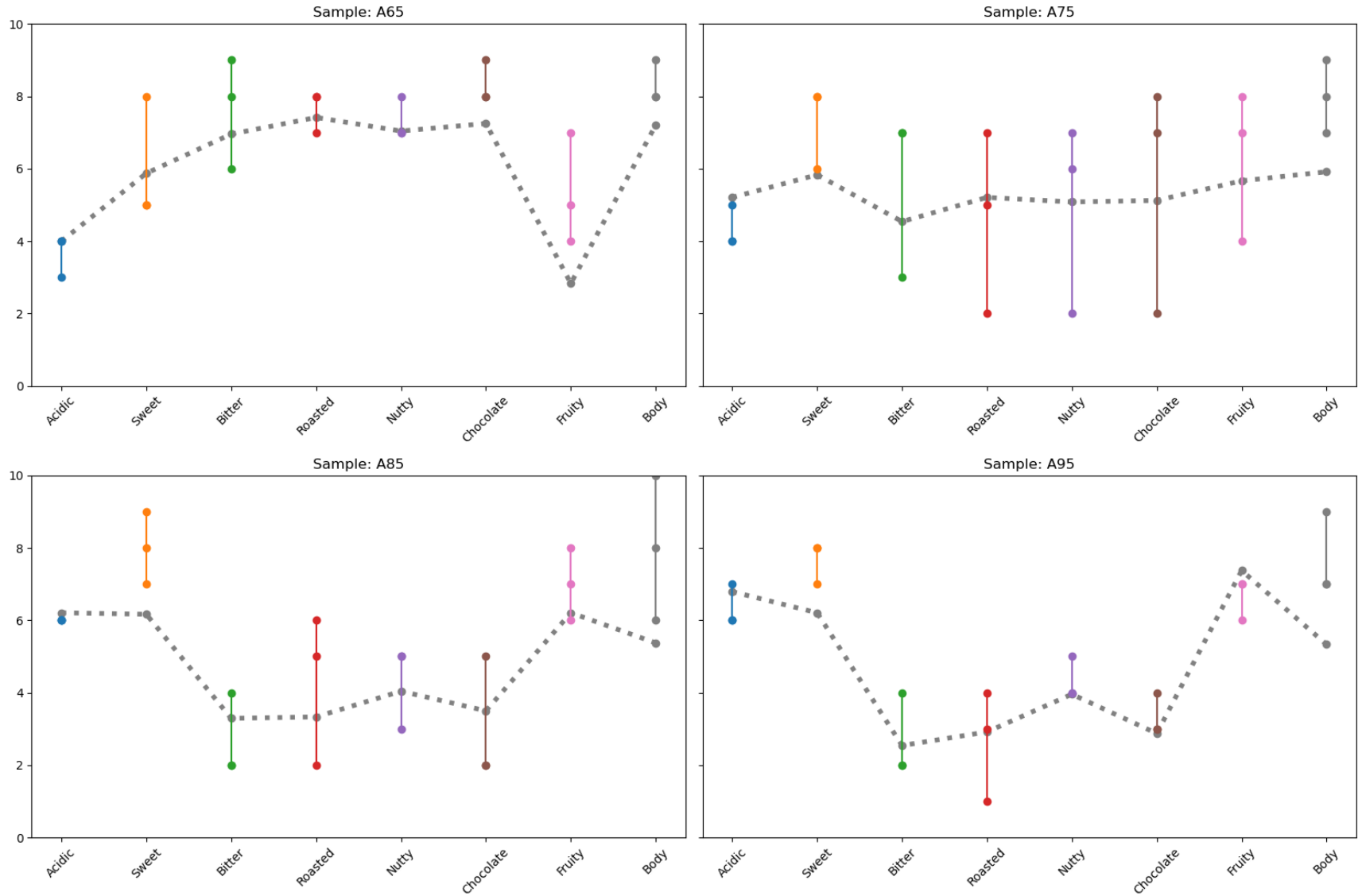
	1	2	3	4	5	6	7	8	9
1	185	618	892	225	337	902	465	127	282
2	465	902	892	185	618	337	127	225	282
3	127	892	282	618	185	902	225	337	465
4	185	465	892	337	618	225	902	282	127
5	225	465	127	902	337	618	282	892	185
6	225	465	127	902	337	618	282	892	185
7	282	337	127	465	892	225	185	618	902
8	892	618	337	185	127	225	282	465	902
9	282	185	892	225	127	618	465	337	902
10	185	127	618	902	465	337	892	225	282


Non-Normalized PCA Bi-plot for Main Samples and Descriptors
A75
Fruity
Roasted
Chocolate
A95
Nutty
Body
Bitter
Acidic
Sweet
A65
A85
PC 2 (0.89%)
PC 1 (98.71%)
0.4
0.2
0.0
−0.2
−0.4
−4
−2
0
2
4
6

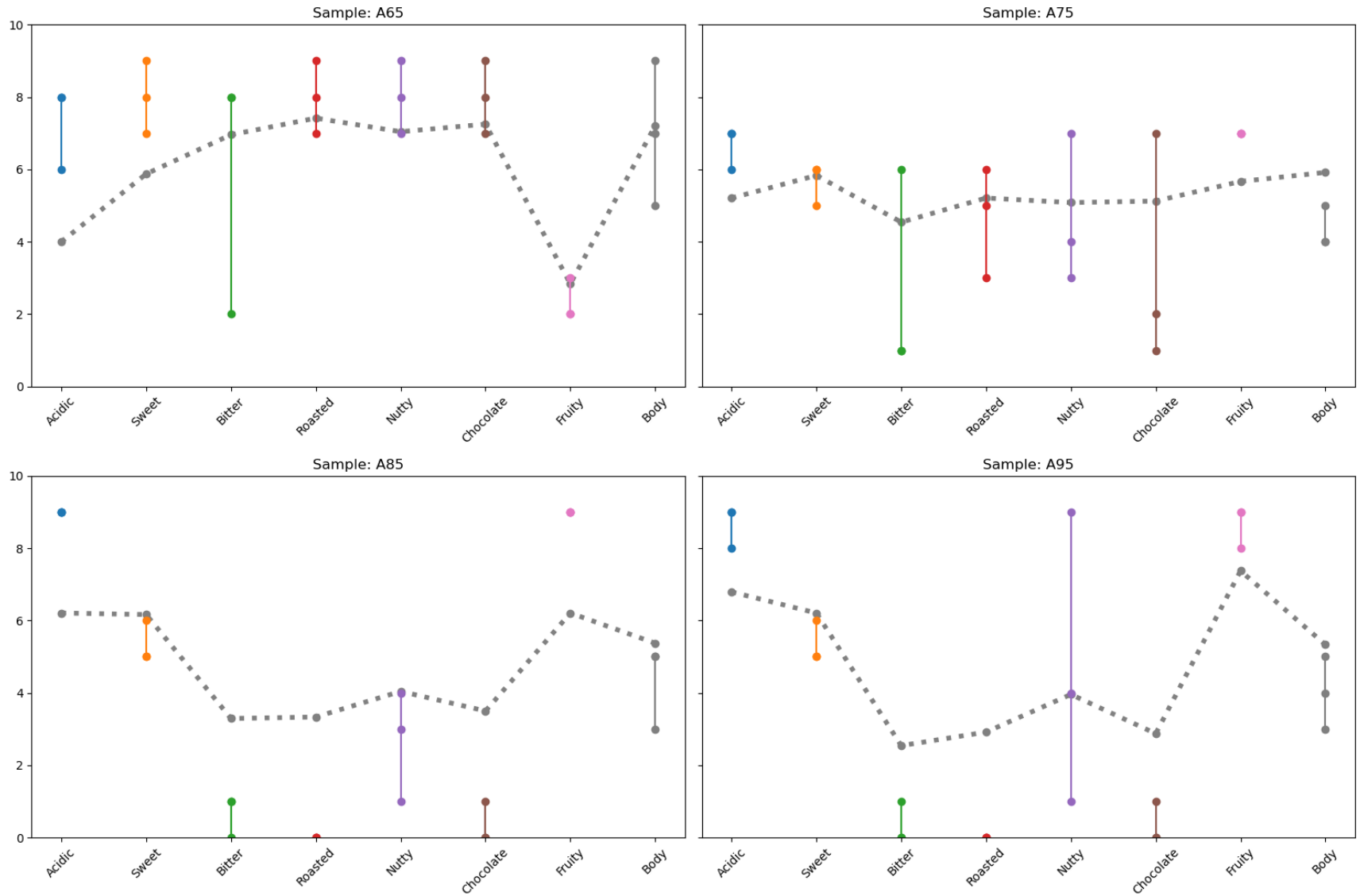
CoffeeMind


Bitter
(<0.001)
Sweet
(ns)
Acidic
(<0.001)
Body
(<0.01)
Fruity
(<0.001)
Chocolate
(<0.001)
Nutty
(<0.001)
Roasted
(<0.001)
A65
A75
A85
A95

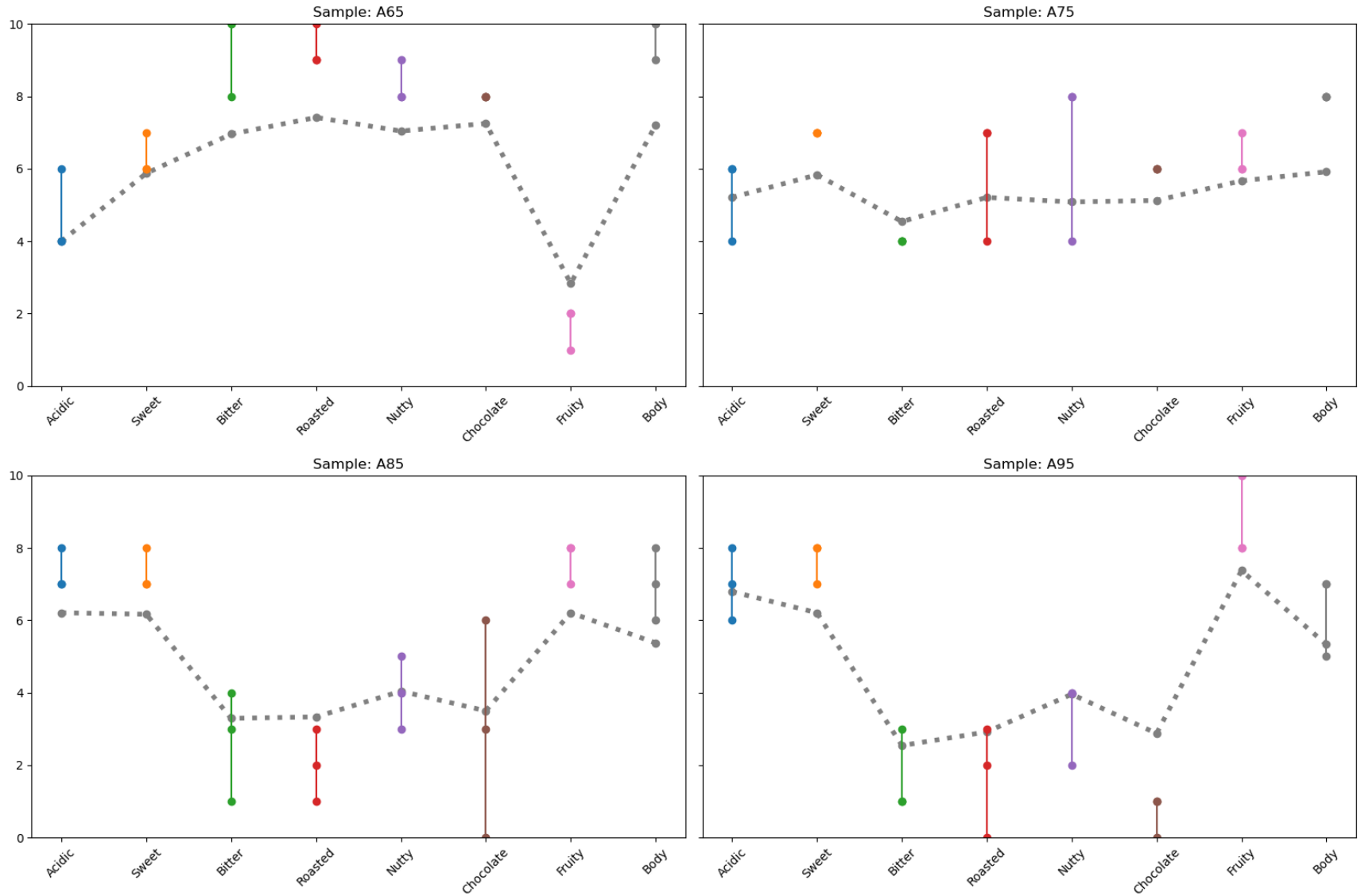
Assessor: A4


Sample: A65
Sample: A75
Sample: A85
Sample: A95
0
2
4
6
8
10
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body

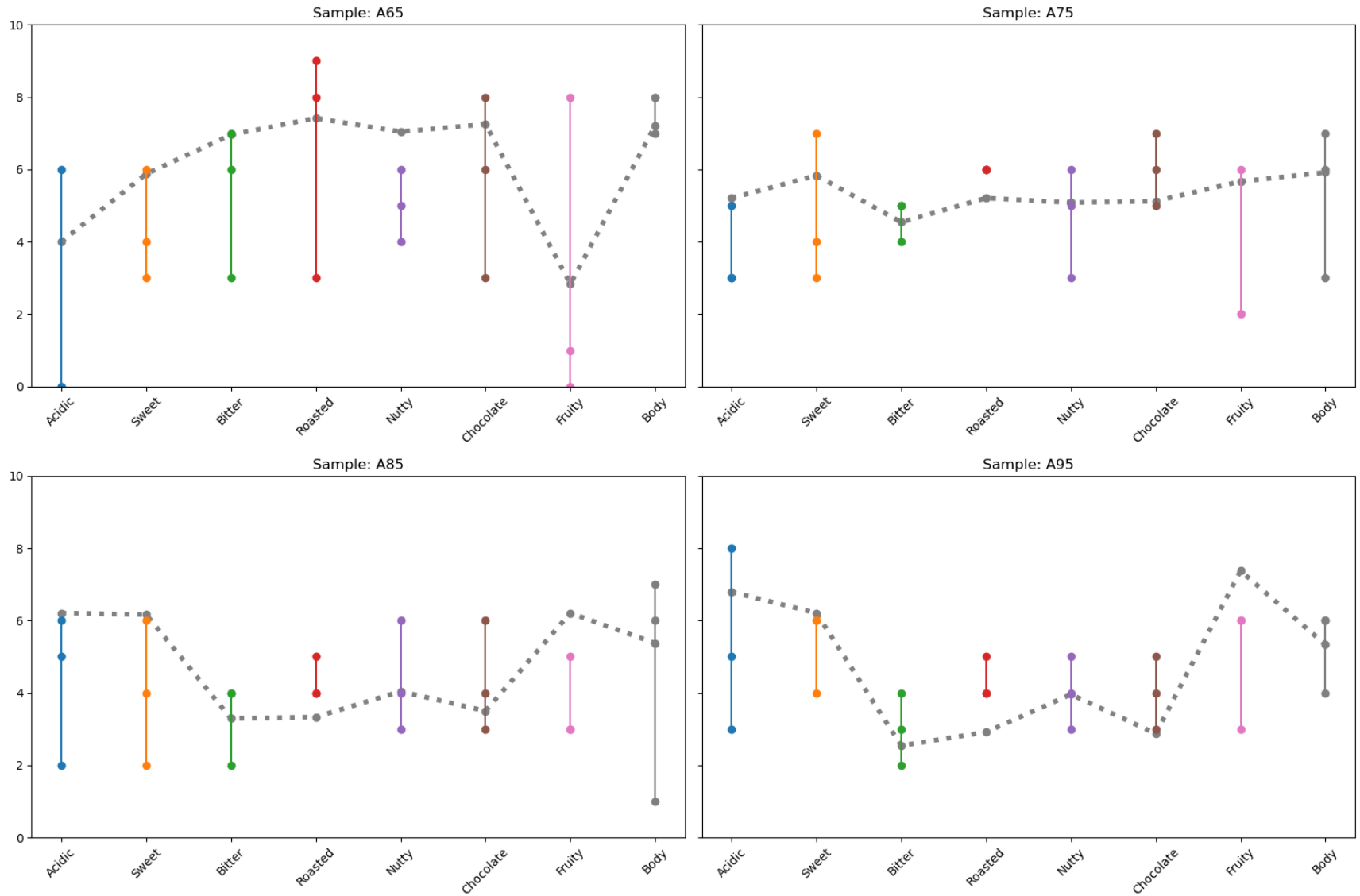
Assessor: A2


Sample: A65
Sample: A75
Sample: A85
Sample: A95
0
2
4
6
8
10
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body

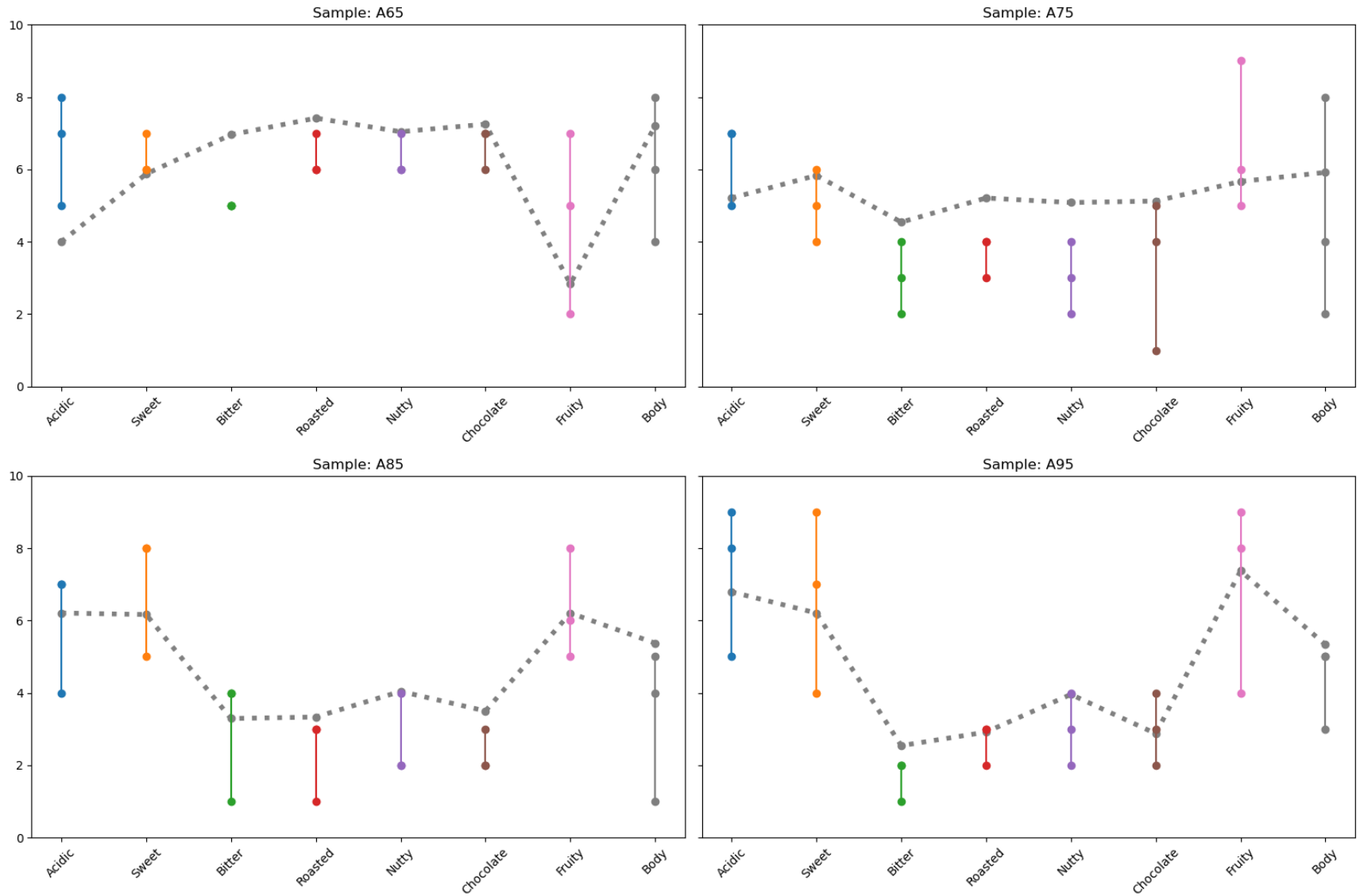
Assessor: A7


Sample: A65
Sample: A75
Sample: A85
Sample: A95
0
2
4
6
8
10
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body

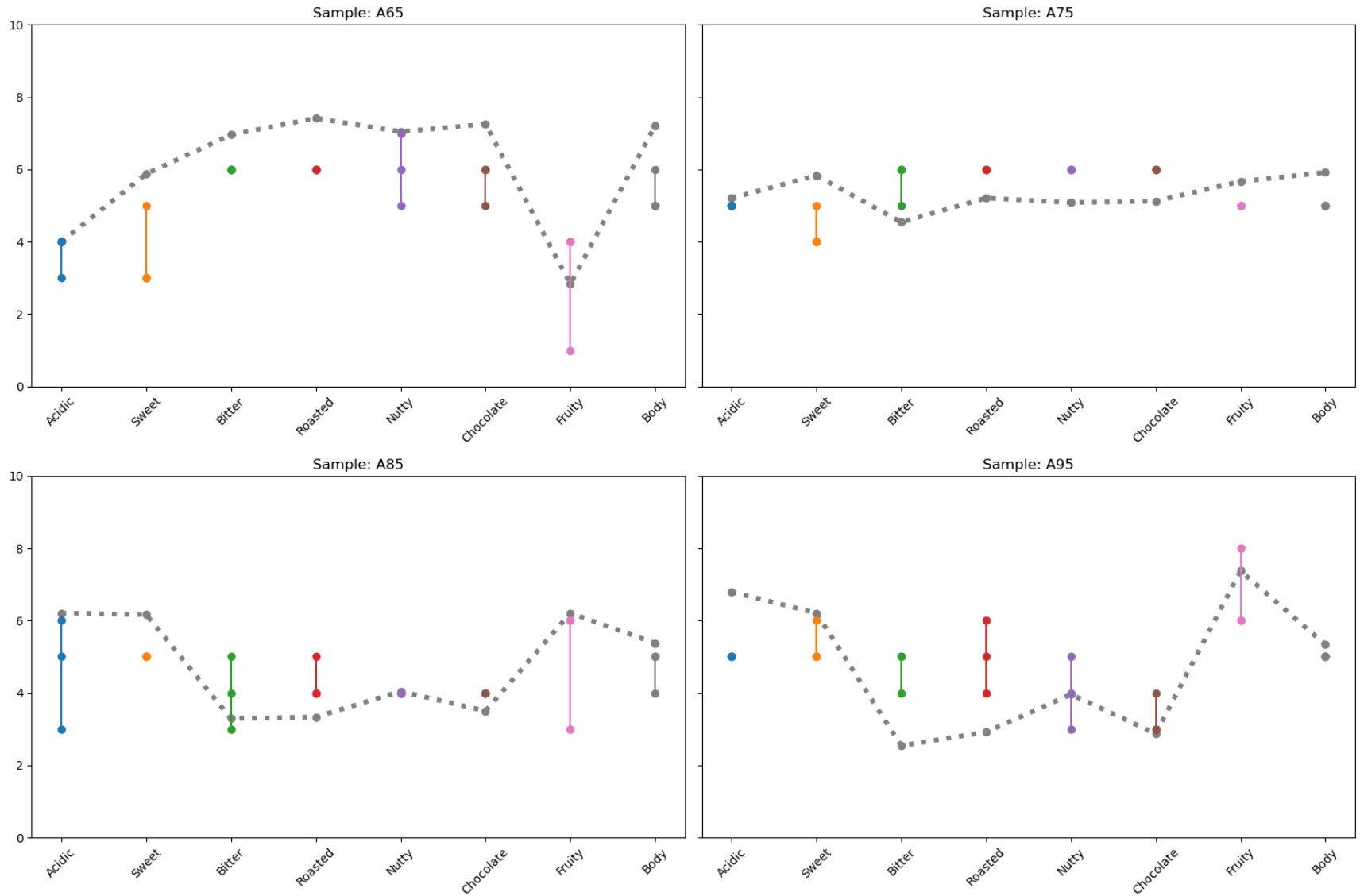
Assessor: A1


Sample: A65
Sample: A75
Sample: A85
Sample: A95
0
2
4
6
8
10
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body

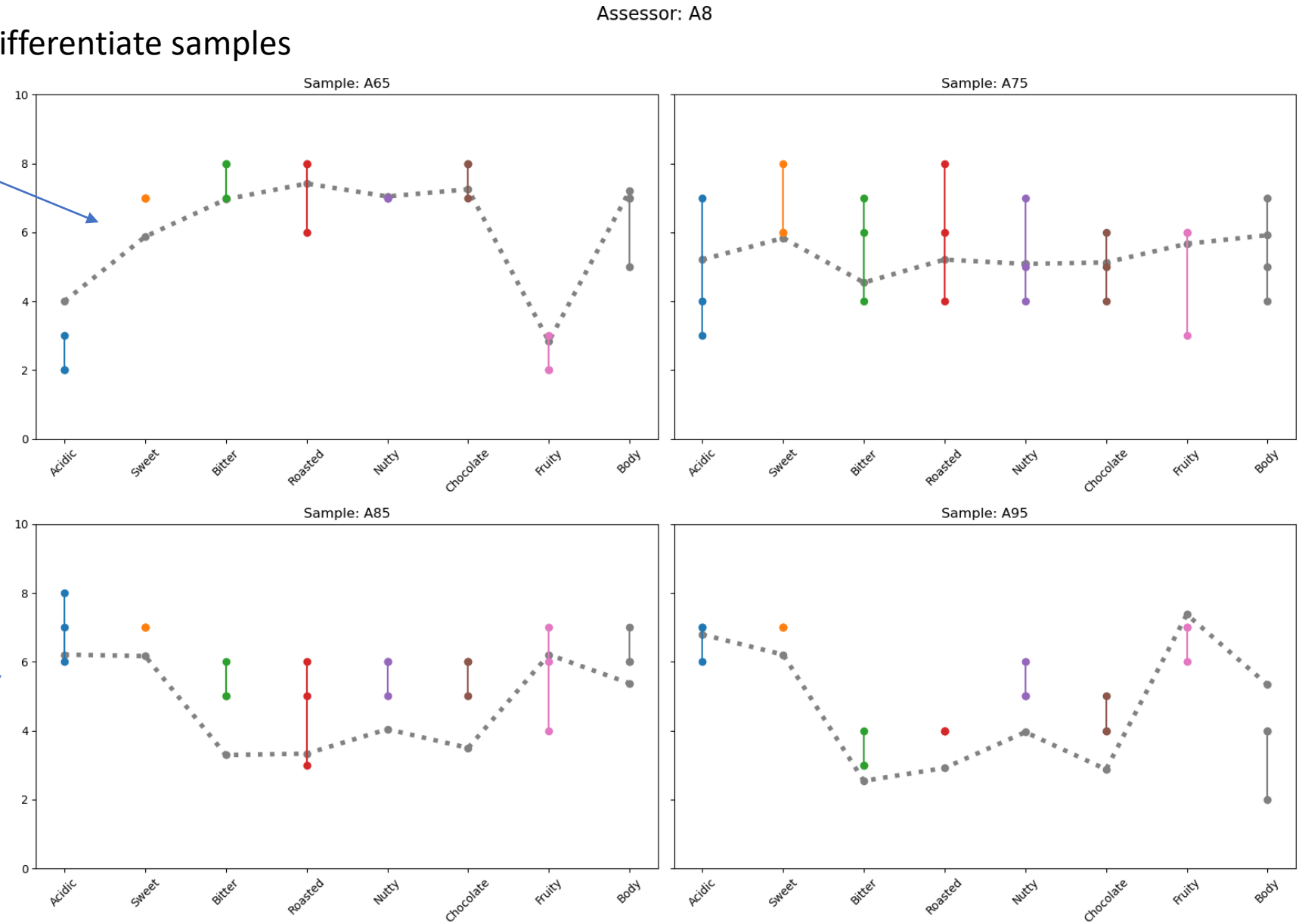
Assessor: A6


Sample: A65
Sample: A75
Sample: A85
Sample: A95
0
2
4
6
8
10
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body

Assessor: A3


Sample: A65
Sample: A75
Sample: A85
Sample: A95
0
2
4
6
8
10
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body

Low p-value:
Use descripters to differentiate samples



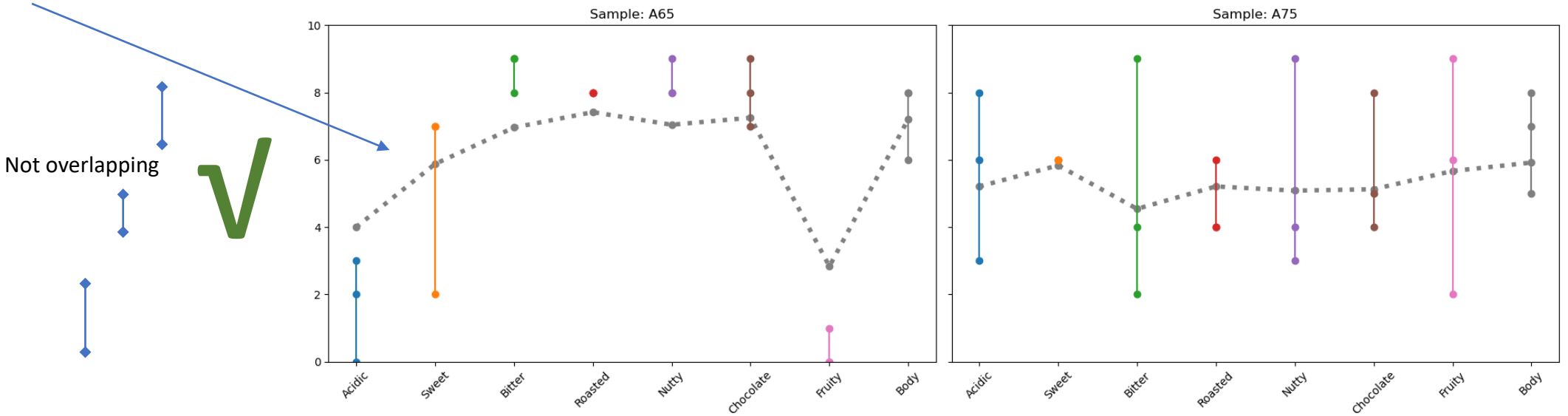
High variation
(High MSE)

Low variation
(Low MSE)

High p-value:
Don’t use descripters well to differentiate samples

Low p-value:
Use descripters to differentiate samples

Assessor: A5

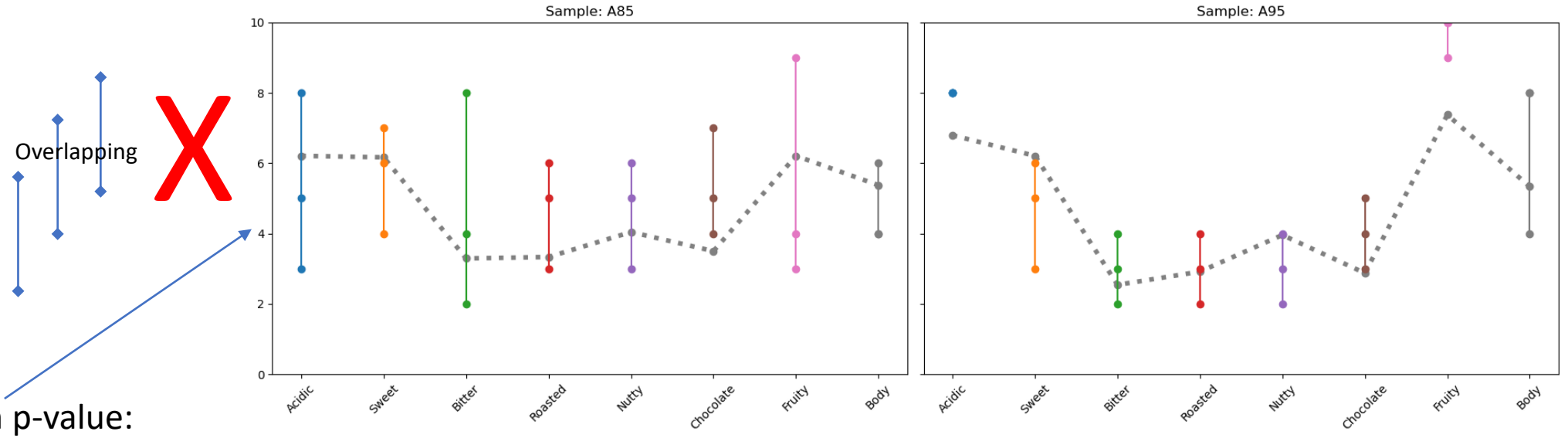


Not overlapping

√

X

High variation
(High MSE)



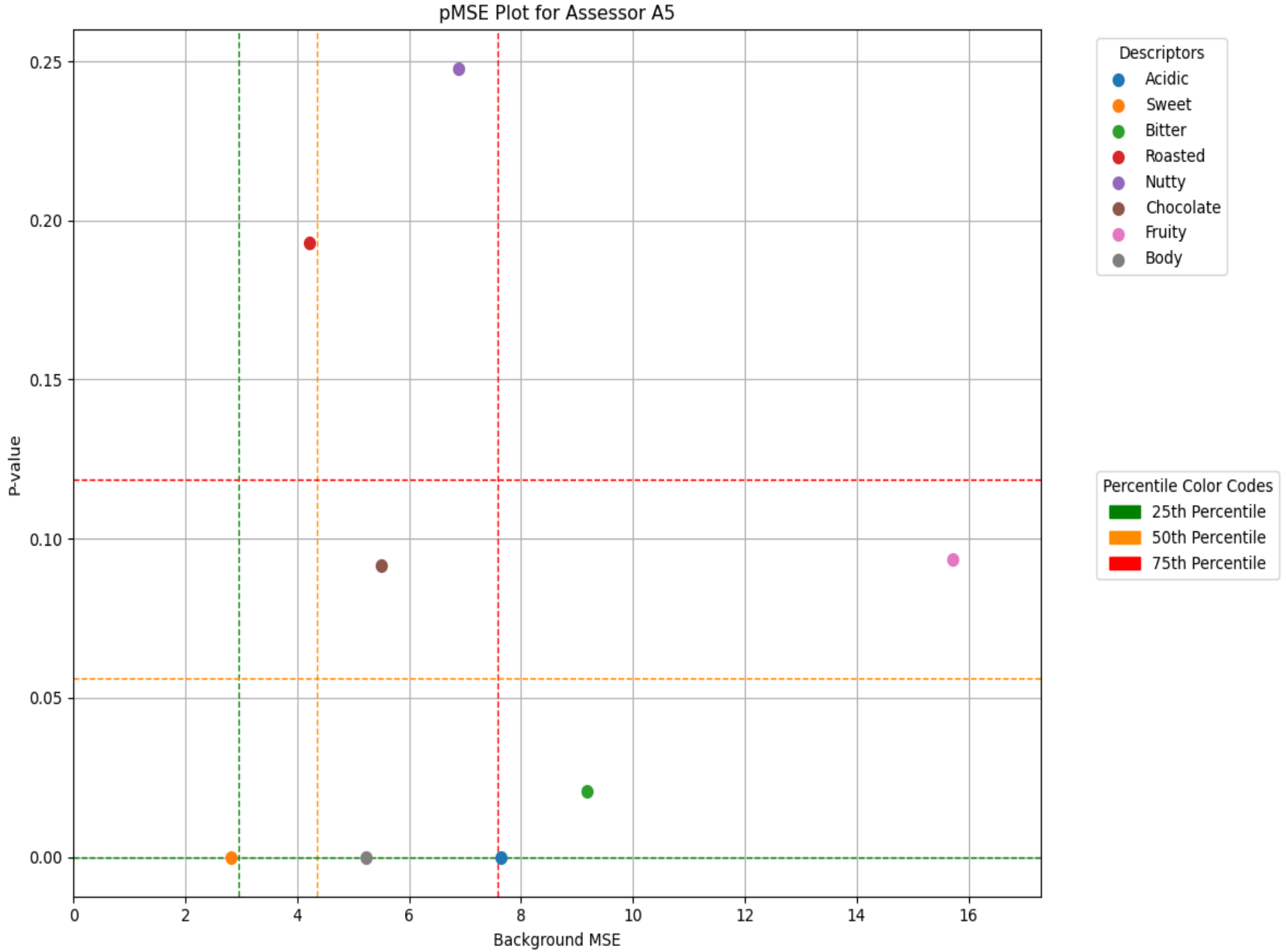
Overlapping

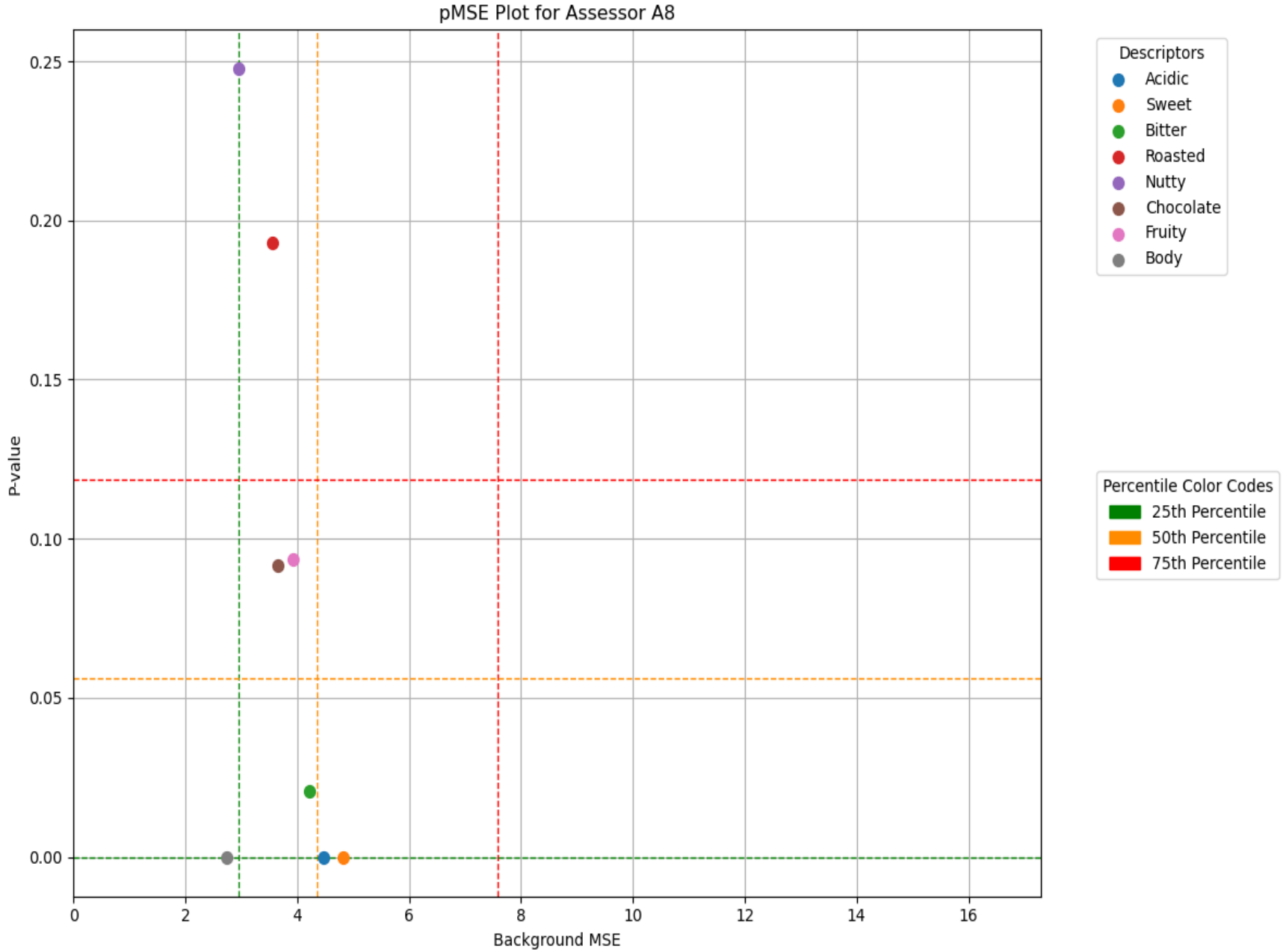
X

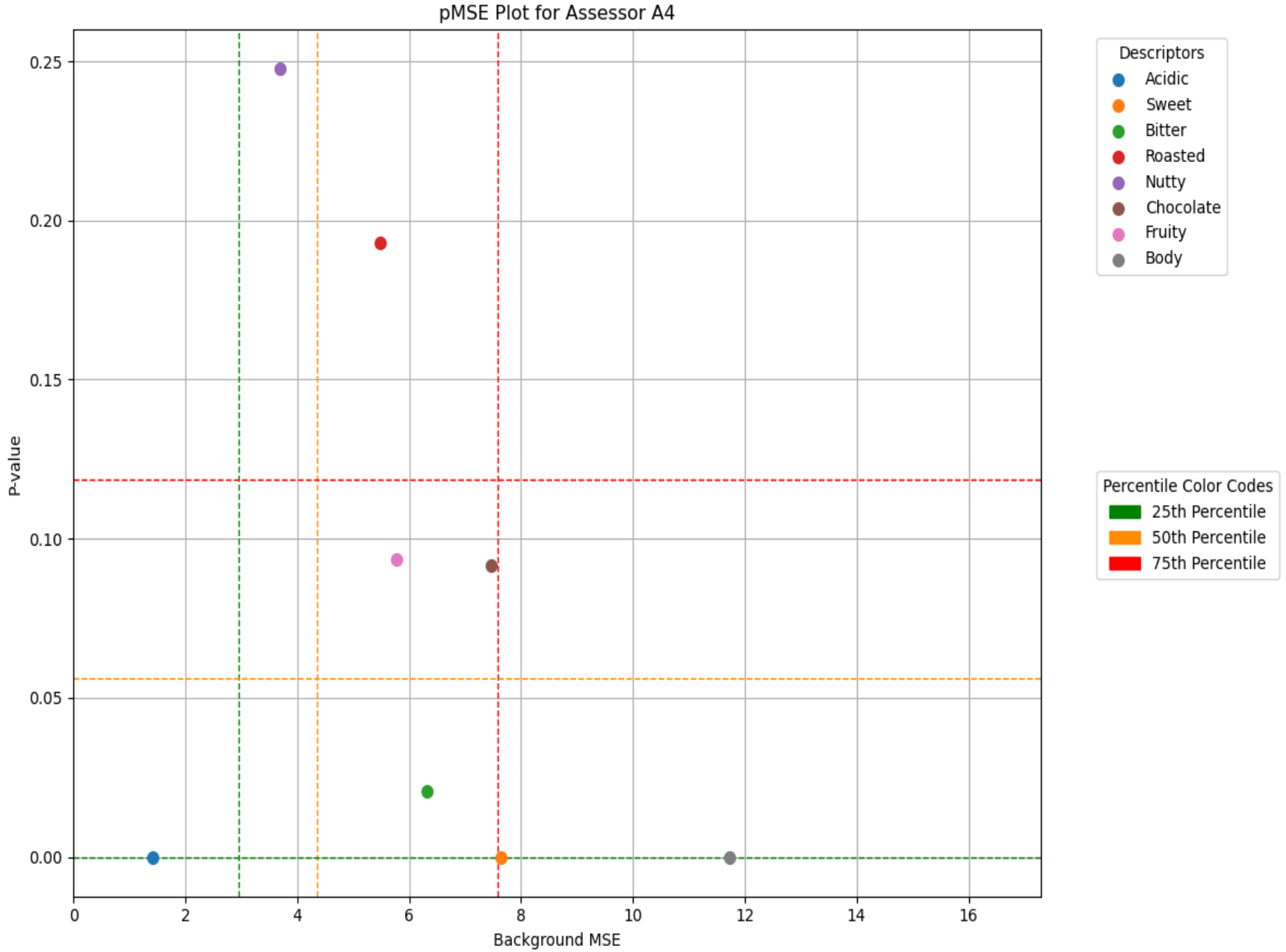
√

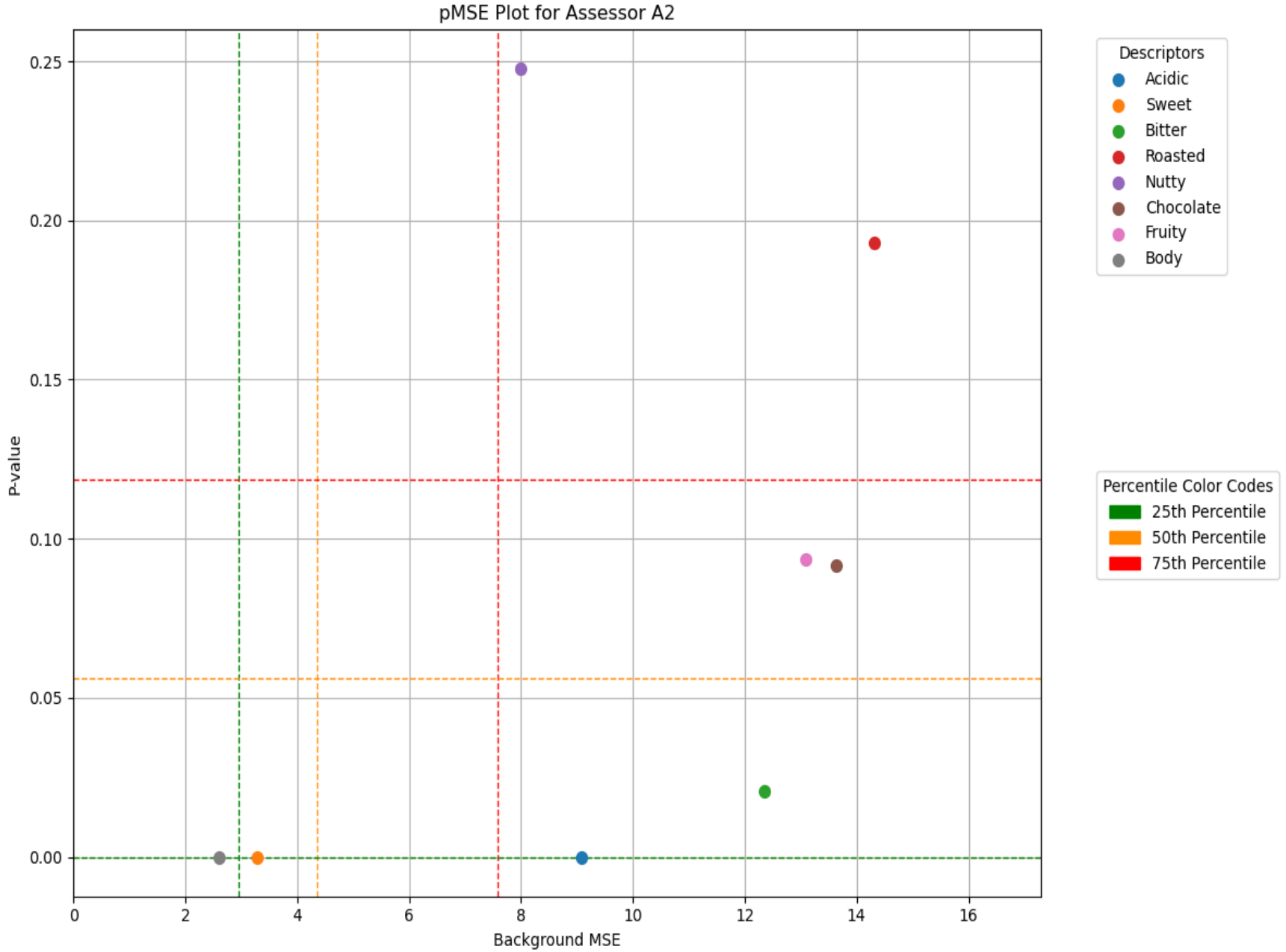
Low variation
(Low MSE)

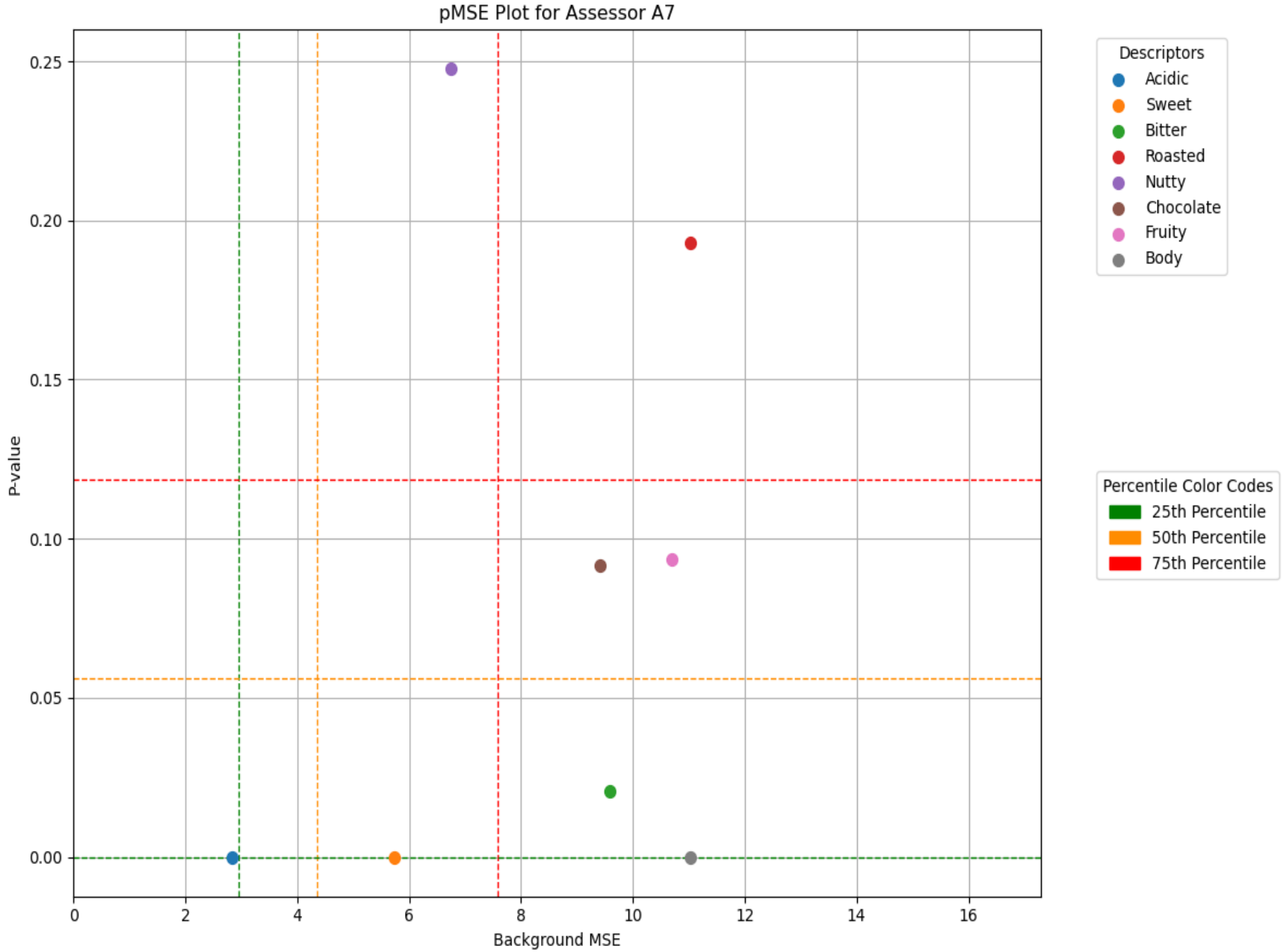
High p-value:
Don’t use descripters well to differentiate samples

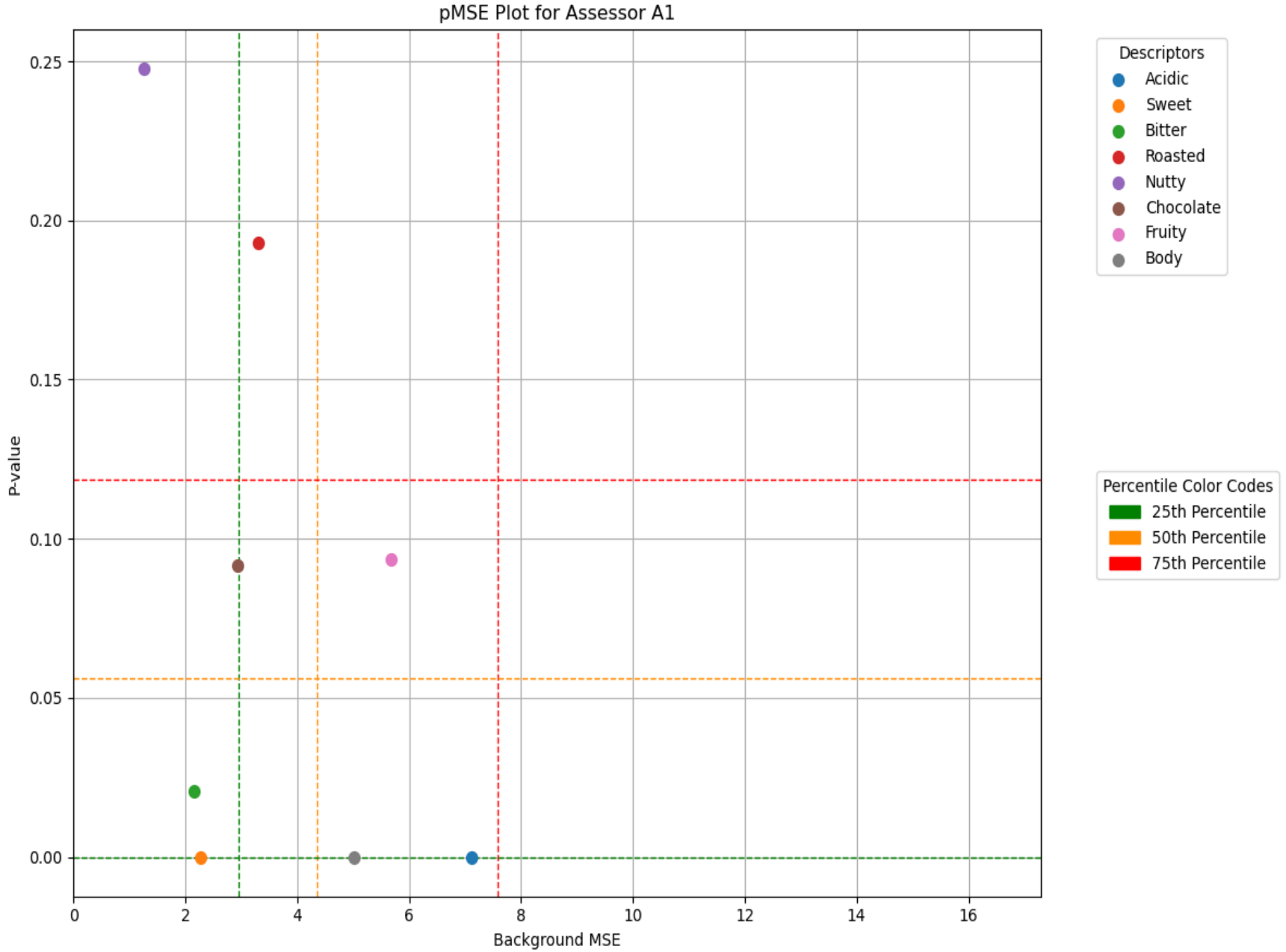

pMSE Plot for Assessor A5
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile

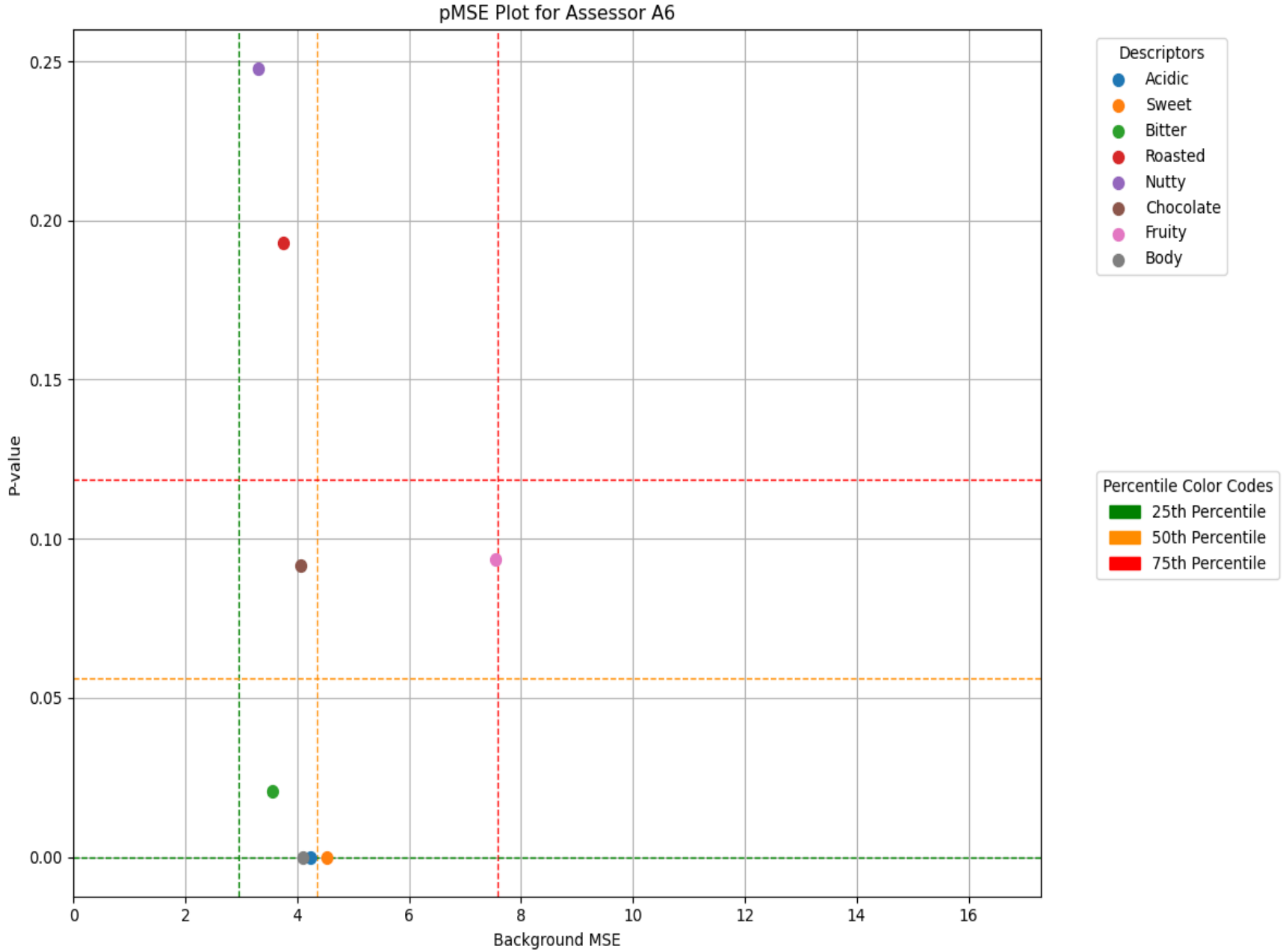

pMSE Plot for Assessor A8
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile

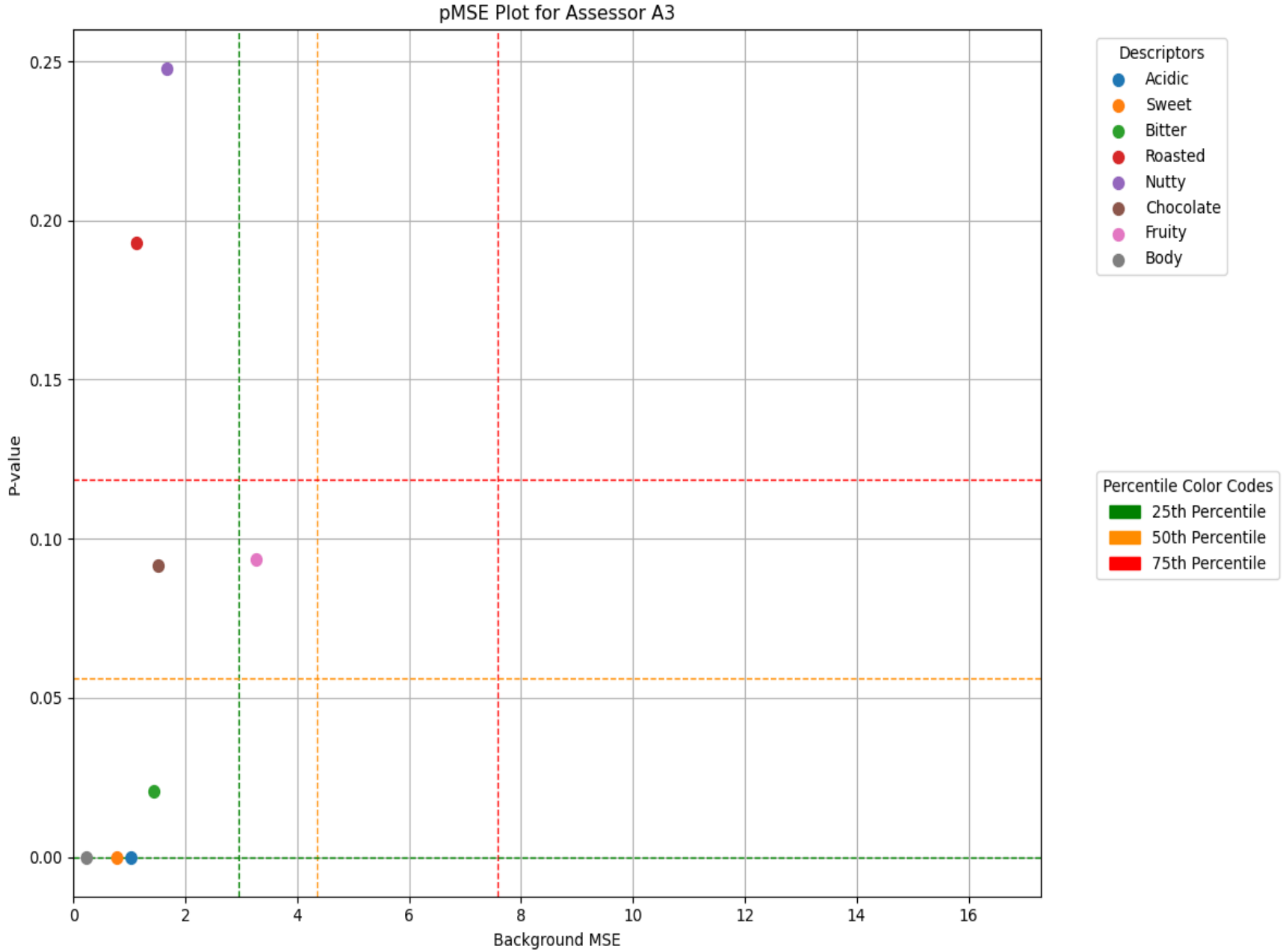

pMSE Plot for Assessor A4
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile


pMSE Plot for Assessor A2
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile


pMSE Plot for Assessor A7
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile


pMSE Plot for Assessor A1
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile


pMSE Plot for Assessor A6
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile


pMSE Plot for Assessor A3
P-value
Background MSE
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
16
Descriptors
Acidic
Sweet
Bitter
Roasted
Nutty
Chocolate
Fruity
Body
Percentile Color Codes
25th Percentile
50th Percentile
75th Percentile