Analysis of Sensory Data: Making the Case for Digging Deeper

C.C. Gilbert and M. Chambault

Introduction

- More and more, the consumer and sensory community needs to deal with tight deadlines and faster turnaround times.
- Automated analyses can save a lot of time, but nothing can beat the inspiration that builds when ‘playing’ with the data.
- The courage to go beyond the use of ‘routine’ methods can lead to some of the best insights about the data.
- Using real-life examples, we will demonstrate various techniques that can be used to dig deeper into the data in order to make results more meaningful and actionable.

Hedonic Data – Liking, JAR and Preference (1)

- Importance of plotting the raw data to look at the scores distribution (uni- vs. bi-modal?)
- Dataset: CLT, 2 snack products, 117 consumers
- Analyses: summary statistics, 2-way ANOVA without interactions, Newman-Keuls, 1-proportion test (all at 5% significance)
- Below: mean liking scores and preference – with all consumers

Liking Attribute	Sample	N	Mean	Median	Stand. Dev.	ANOVA P-value	Significant at 5%?
Overall	A	117	7.3	8	1.20	0.206	NO
	B	117	7.1	7	1.28		
Appearance	A	117	6.7	7	1.55	0.681	NO
	B	117	6.6	7	1.62		
Flavour	A	117	7.1	7	1.35	0.848	NO
	B	117	7.1	7	1.37		
Texture	A	117	7.0	7	1.34	0.091	NO
	B	117	6.7	7	1.51		

	Preferred		Preferred without no preference	
	Count	%	Count	%
Sample A	57	49	57	53
Sample B	51	44	51	47
No preference	9	78		
Total	117	100	108	100
Proportion test		**P-value**	**0.631**	**NS**

Both samples received fairly high liking scores with neither sample being liked significantly more (ANOVA) or preferred (proportion test)

Hedonic Data – Liking, JAR and Preference (2)

Now significant differences are seen on all modalities:

- Mean liking scores – with respondents who preferred Sample A

Liking Attribute	Sample	N	Mean	Median	Stand. Dev.	ANOVA P-value	Significant at 5%?	NK 5%
Overall	A	57	7.8	8	0.78	<0.001	YES	A
	B	57	6.4	7	1.28			B
Appearance	A	57	6.9	7	1.42	0.001	YES	A
	B	57	6.0	6	1.54			B
Flavour	A	57	7.6	8	1.01	<0.001	YES	A
	B	57	6.4	7	1.48			B
Texture	A	57	7.2	8	1.34	0.012	YES	A
	B	57	6.5	7	1.38			B

- Mean liking scores – with respondents who preferred Sample B

Liking Attribute	Sample	N	Mean	Median	Stand. Dev.	ANOVA P-value	Significant at 5%?	NK 5%
Overall	B	51	7.7	8	0.88	<0.001	0.1%	A
	A	51	6.7	7	1.32			B
Appearance	B	51	7.4	7	1.20	<0.001	0.1%	A
	A	51	6.2	7	1.71			B
Flavour	B	51	7.8	8	0.89	<0.001	0.1%	A
	A	51	6.5	7	1.45			B
Texture	B	51	7.5	8	1.12	<0.001	0.1%	A
	A	51	6.1	7	1.51			B

We would wrongly conclude on no difference if not splitting the data according to respondent preferences!

Hedonic Data – Liking, JAR and Preference (3)

JAR results – looking at Sweetness (same for some other attributes):

- With all respondents

Sweetness	Sample A		Sample B	
	Count	%	Count	%
Much too sweet	1	1	1	1
A little too sweet	19	16	12	10
Just about right	63	**54**	70	**60**
Not quite sweet enough	32	28	28	24
Not at all sweet enough	1	1	6	5
Total	116	100	117	100

We would wrongly conclude on no difference!

- With respondents who preferred Sample A

Sweetness	Sample A		Sample B	
	Count	%	Count	%
Much too sweet	0	0	0	0
A little too sweet	6	10	10	17
Just about right	40	**70**	18	32
Not quite sweet enough	10	18	23	**40**
Not at all sweet enough	1	2	6	**11**
Total	57	100	57	100

For 50% of them, B is clearly not sweet enough

- With respondents who preferred Sample B

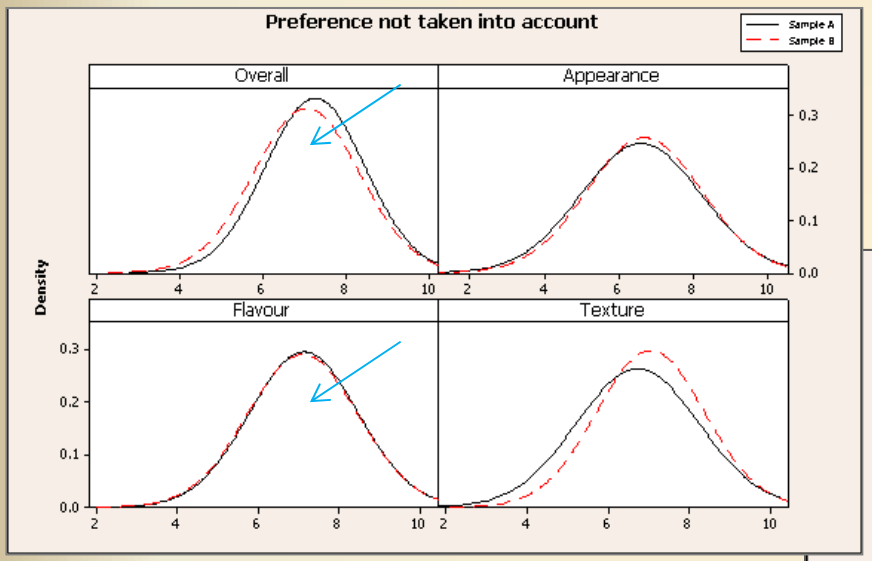
Sweetness	Sample A		Sample B	
	Count	%	Count	%
Much too sweet	1	2	0	0
A little too sweet	13	**26**	2	4
Just about right	20	40	45	**88**
Not quite sweet enough	16	**32**	4	8
Not at all sweet enough	0	0	0	0
Total	50	100	51	100

Less homogeneous group
A slightly too sweet for 30% of them, but also not sweet enough for another 30%!

Hedonic Data – Liking, JAR and Preference (4)

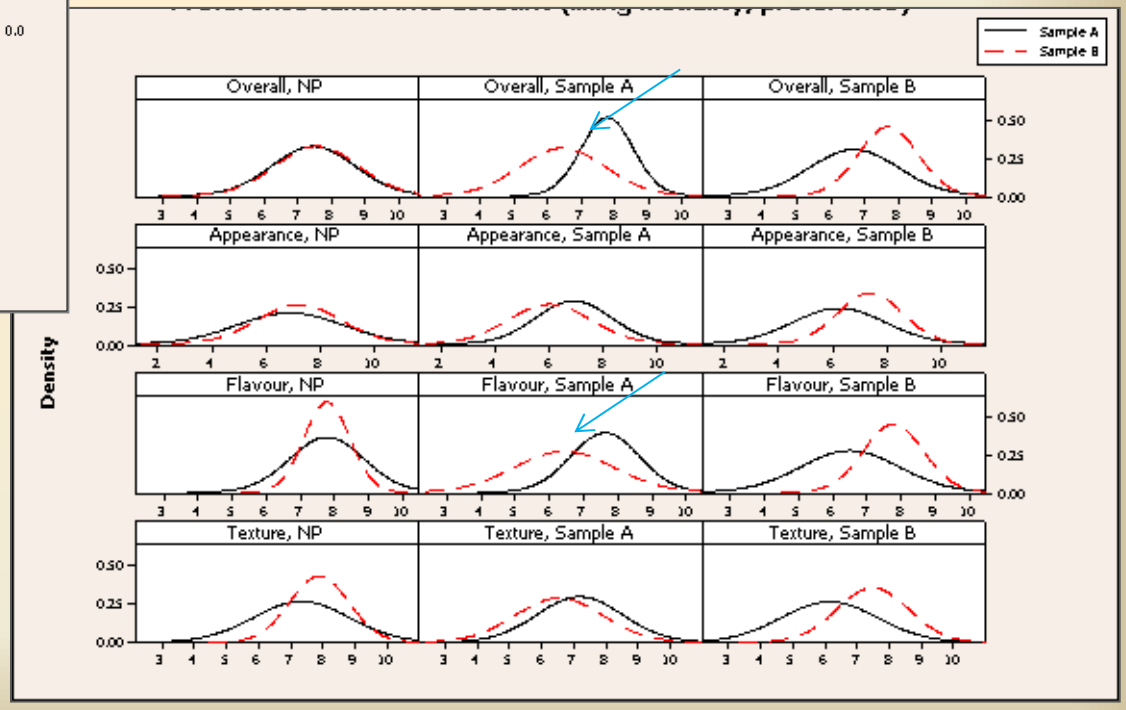
- Data distributions of the liking scores:

Showing the scores of each liking modality for both ‘A likers’ & ‘B likers’ combined



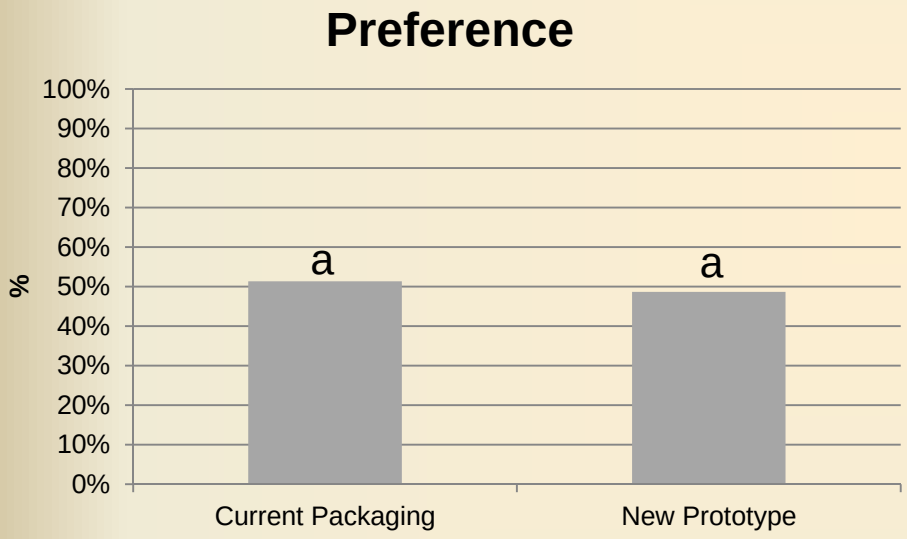
Becomes all clearer when plotting the scores for each liking modality .....

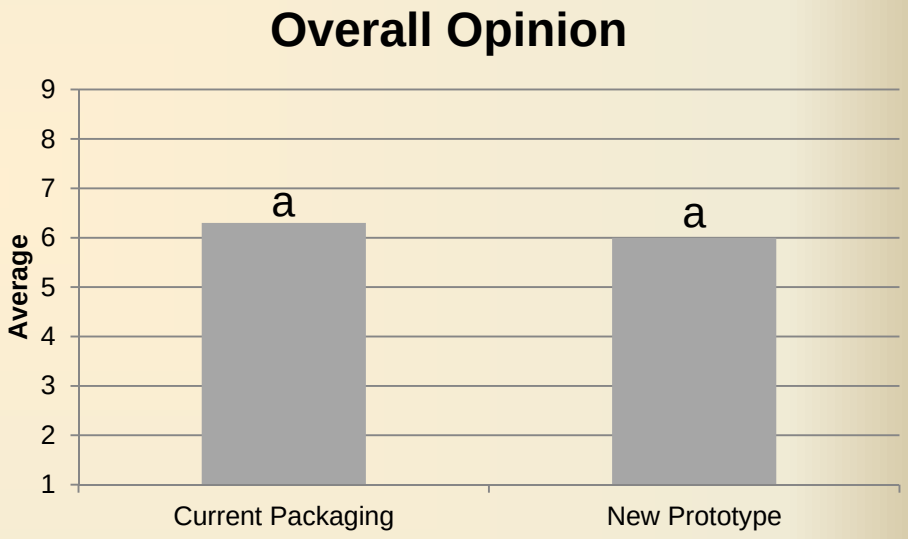
Now distinguishing (column-wise) between ‘A likers’, ‘B likers’ and ‘undecided’ (row-wise)



Alienation

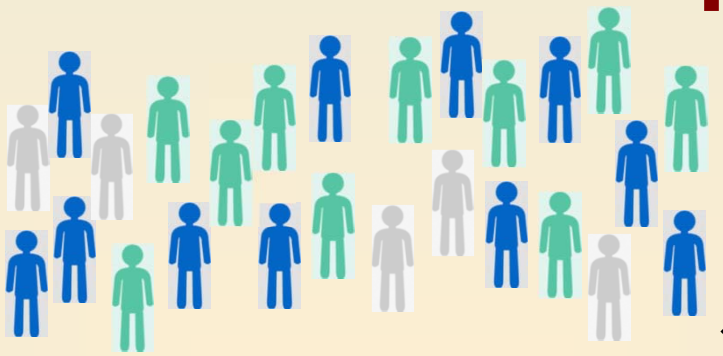
Example:





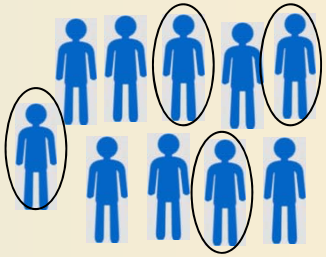
Can we replace the current packaging with the new?

Alienation Principle



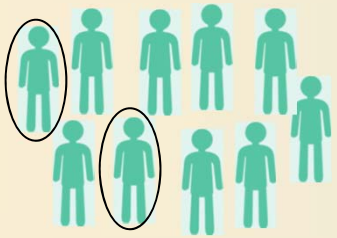
Heavy users of the product

Prefer '**Current**' product



% who rated the '**New**' product negatively (e.g. bottom 2 box PI)

Prefer '**New**' product



% who rated the '**Current**' product negatively (e.g. bottom 2 box PI)

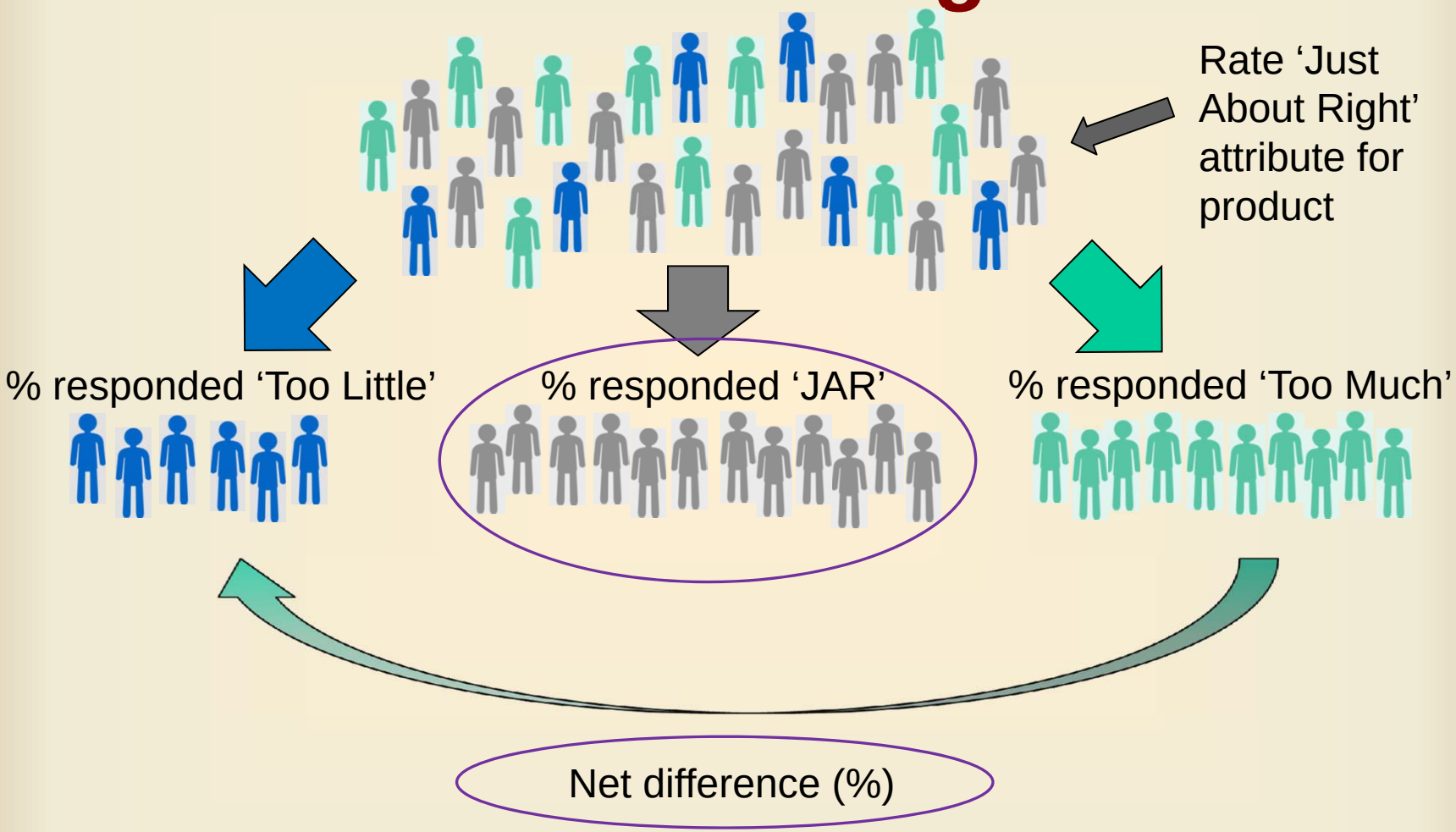
'Alienation' ⟷ 'Noise'

Example – benefits of alienation

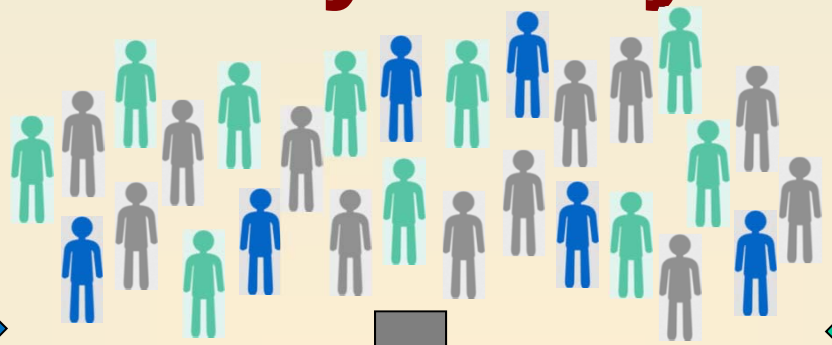
	Packaging study		
	Current Packaging vs New Prototype		
	ALIENATION % of respondents who preferred Current Packaging, who rated New Prototype negatively	**NOISE** % of respondents who preferred New Prototype, who rated Current Packaging negatively	**Significance**
Purchase Intent - Bottom 2 boxes	61.0%	13.7%	Significant alienation ($p<0.05$)
(definitely would not purchase, probably would not purchase)	(47 out of 77)	(10 out of 73)	

- Sometimes liking and preference do not tell the whole story…
- Find a method that best fits your objective

Just About Right

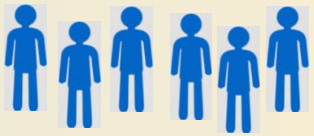


Penalty Analysis



Rate 'Just About Right' attribute for product

% responded 'Too Little'



% responded 'JAR'



% responded 'Too Much'



Mean 'Overall Liking' for 'Too Little' group

Mean 'Overall Liking' for 'JAR' group

Mean 'Overall Liking' for 'Too Much' group

Mean Drop JAR - 'Too Little'

Mean Drop JAR - 'Too Much'

JAR Summary – Ranch Prototype Salad Dressing

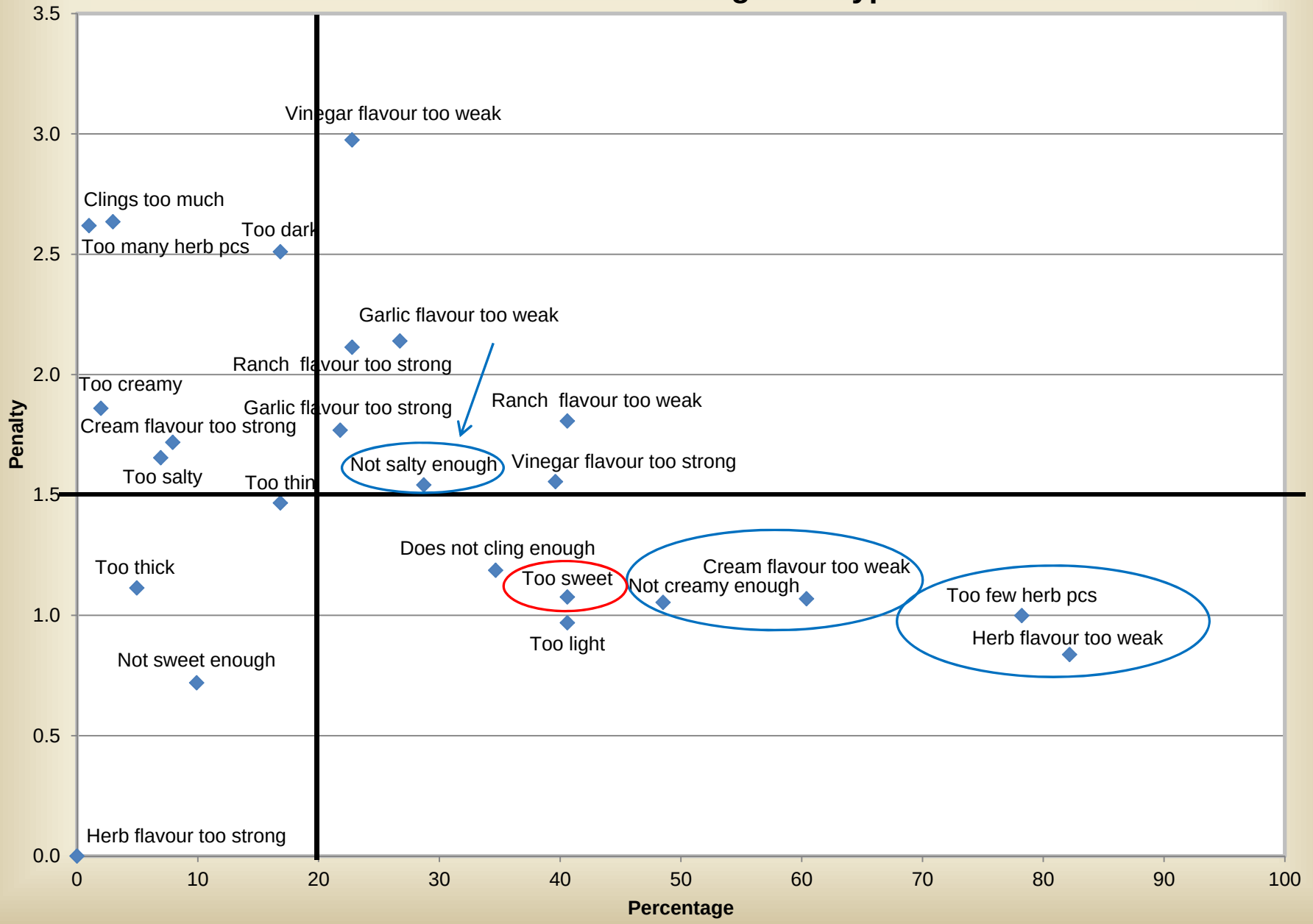
Attribute	JAR % / Net Difference %
Colour	43/-24
Amount of herbs pcs	21/-77
Herb flavour	18/-82
Sweetness	50/+31
Saltiness	50/-47
Vinegar flavour	38/+17
Garlic flavour	51/-5
Ranch flavour	37/-18
Cream flavour	32/-52
Creaminess	64/-22
Consistency/thickness	78/-12
Clings to lettuce	62/-32

Not enough of particular attribute (Net Effect < - 20%)

Just Right (JAR > 70%)

Too much of particular attribute (Net Effect > 20%)

Ranch Salad Dressing Prototype



JAR Summary – Ranch Prototype Salad Dressing

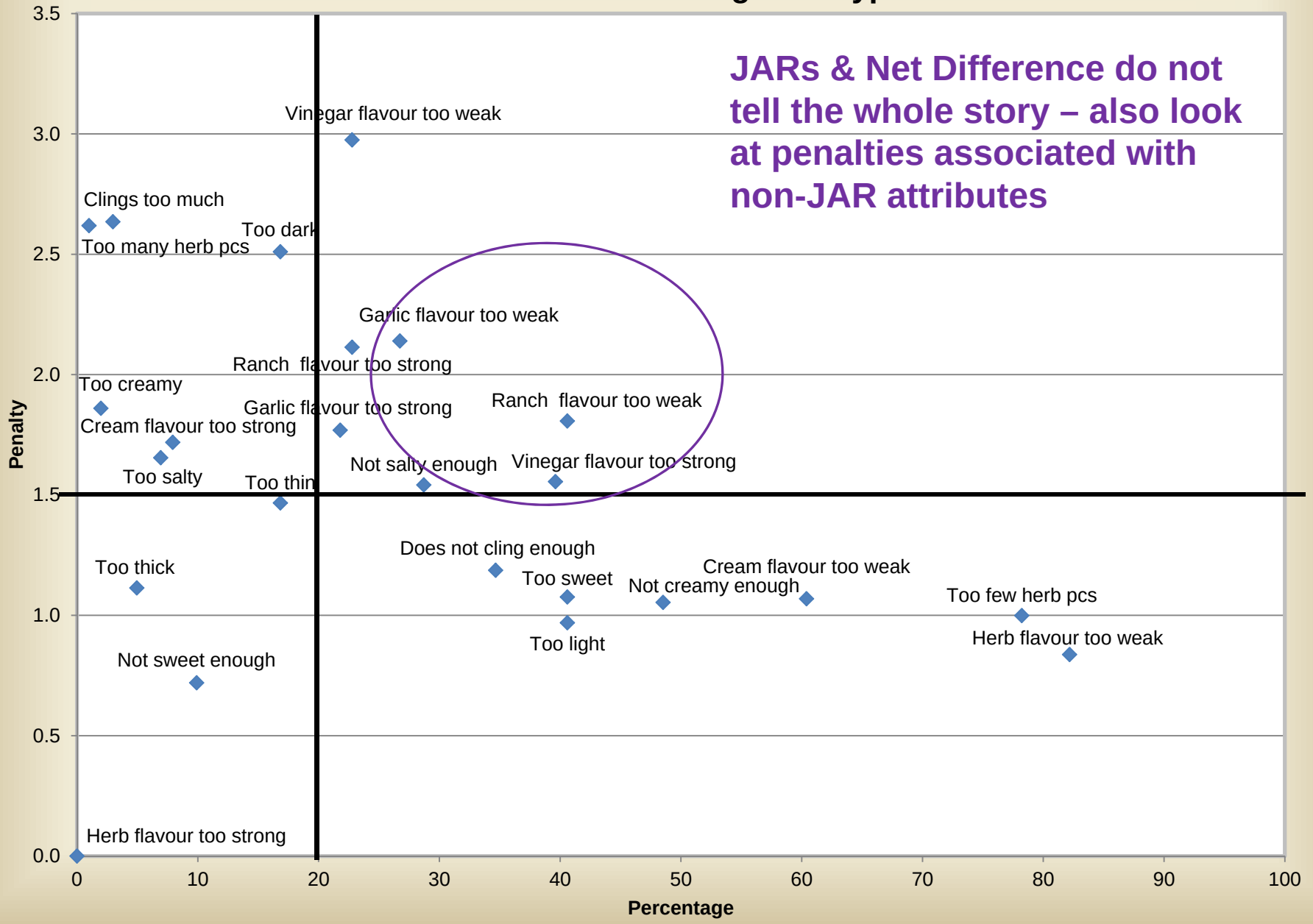
Attribute	JAR % / Net Difference %
Colour	43/-24
Amount of herbs pcs	21/-77
Herb flavour	18/-82
Sweetness	50/+31
Saltiness	50/-47
Vinegar flavour	38/+17
Garlic flavour	51/-5
Ranch flavour	37/-18
Cream flavour	32/-52
Creaminess	64/-22
Consistency/thickness	78/-12
Clings to lettuce	62/-32

Not enough of particular attribute (Net Effect < - 20%)

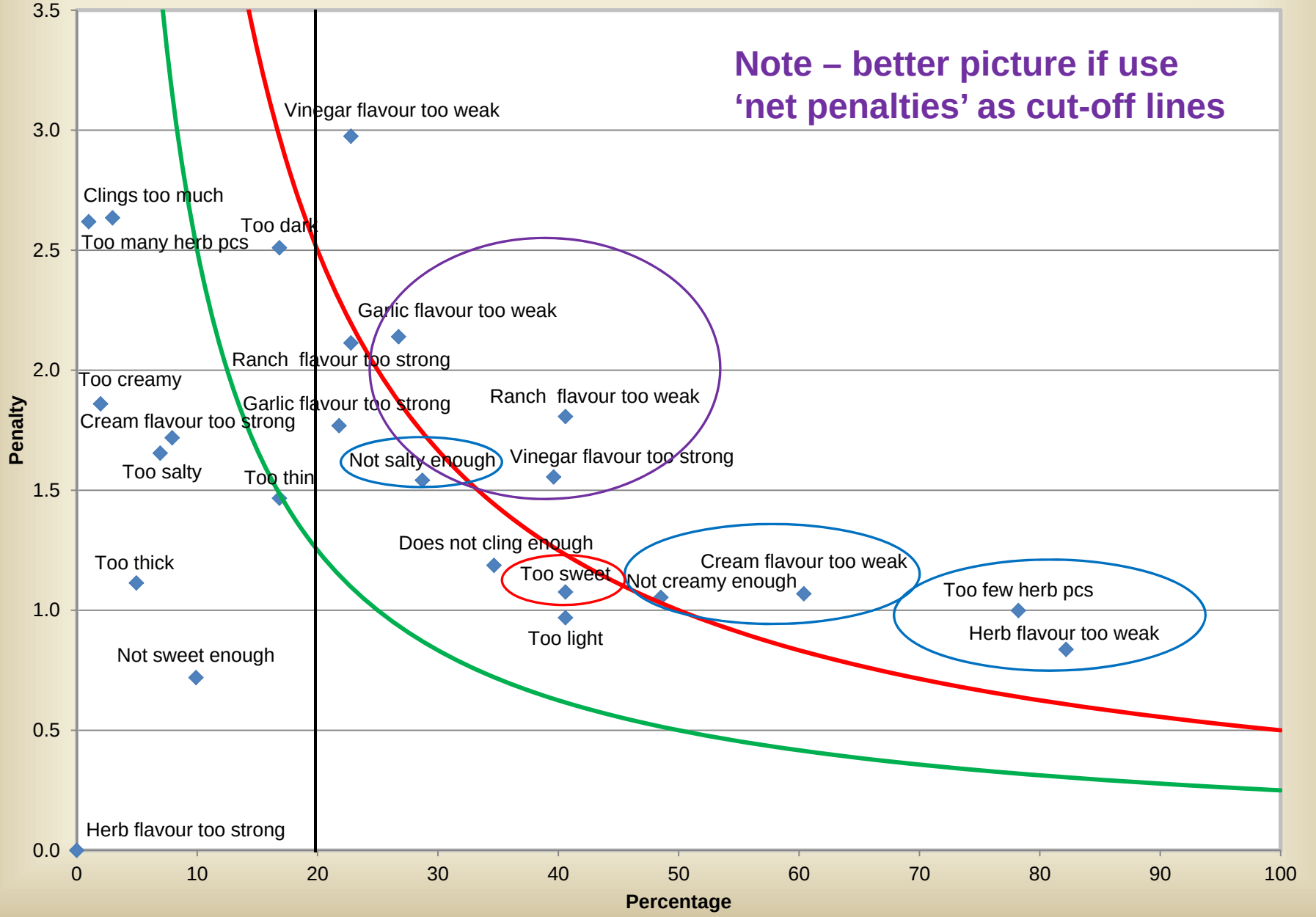
Just Right (JAR > 70%)

Too much of particular attribute (Net Effect > 20%)

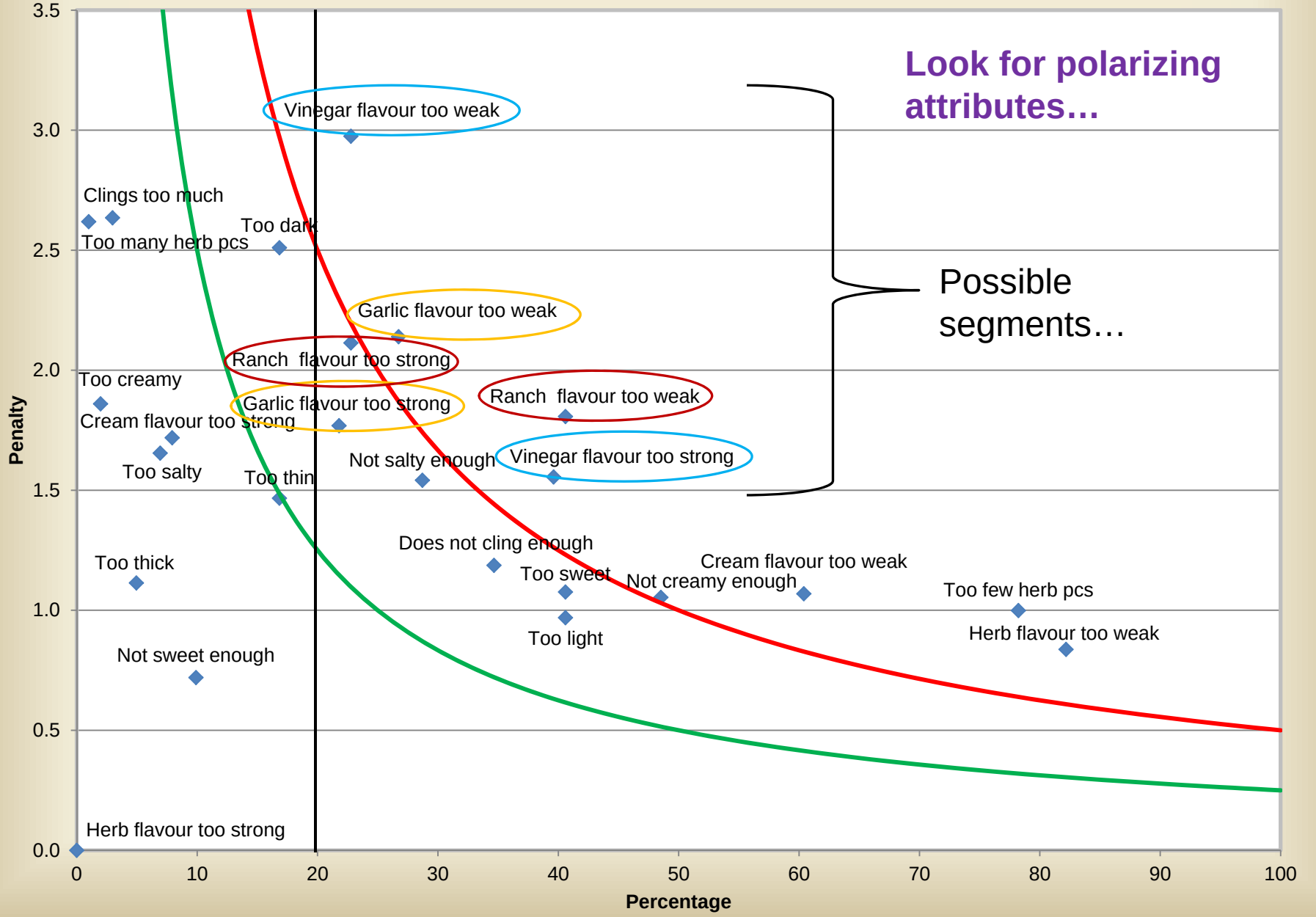
Ranch Salad Dressing Prototype



Ranch Salad Dressing Prototype



Ranch Salad Dressing Prototype



Segmentation

- Clustering is a great ‘digging deeper’ method…
- We’ve seen many examples over the years showing its benefit: results with little to no discrimination between samples when examined overall, but with clear segments of consumers with different liking patterns when results are clustered.
- But sometimes you need to dig a little deeper even within the method itself!

Clustering Example # 1

- Sometimes, clustering consumers using the raw data works great!

Raw Data
(cons x sample scores)

↓

Cluster consumers
(AHC or k-means)

↓

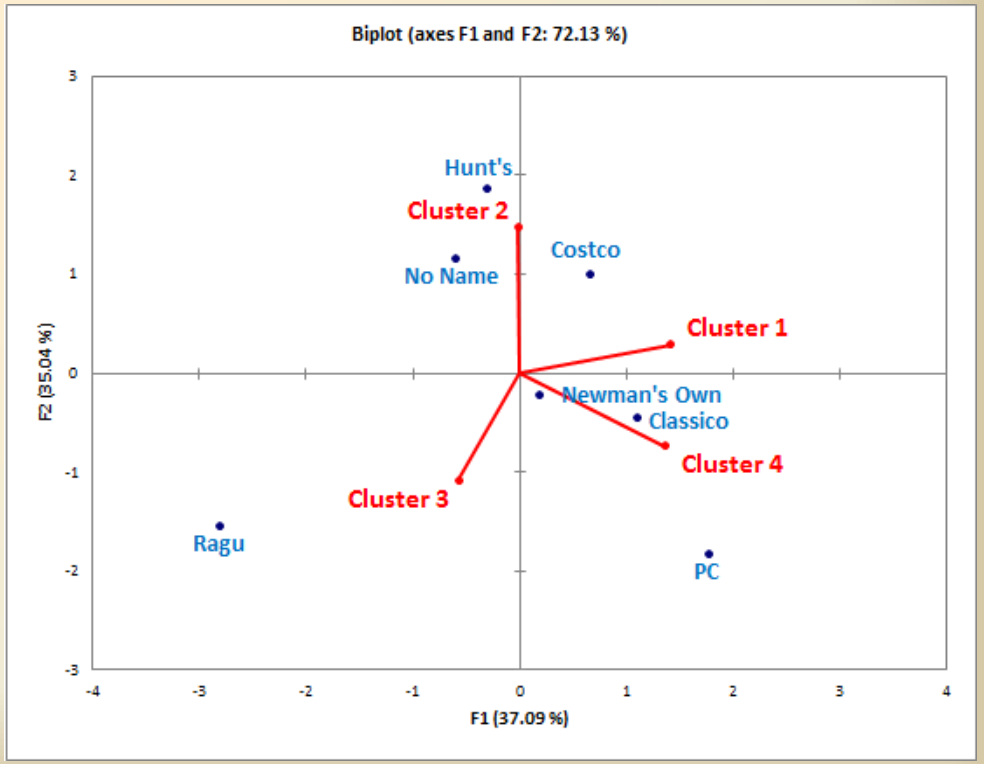
Compute means
(samp x cluster means)

↓

PCA

PCA on matrix of sample by cluster means

↓

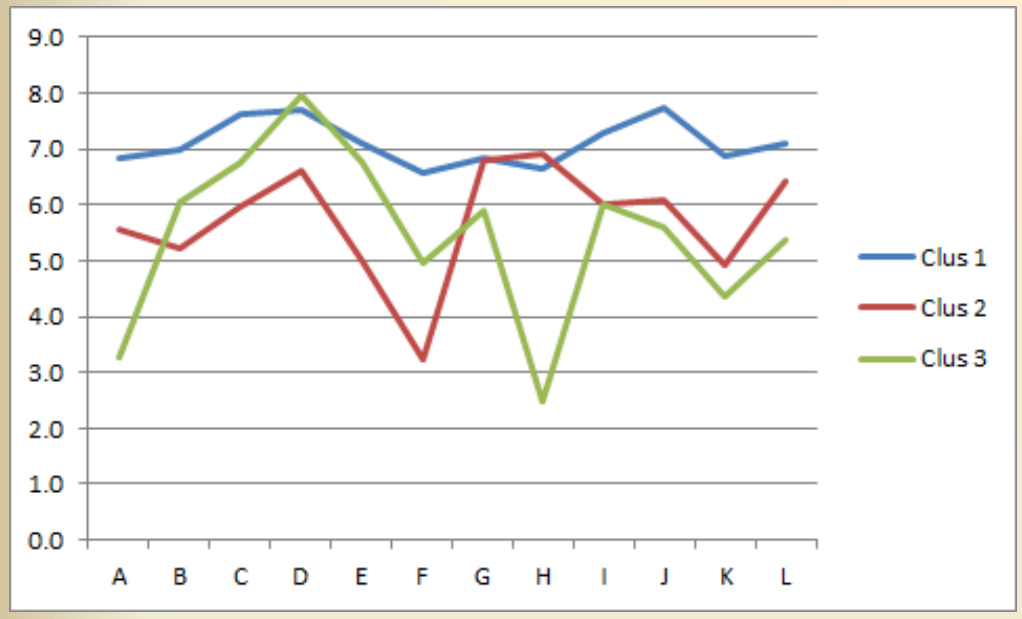


Clustering Example # 2

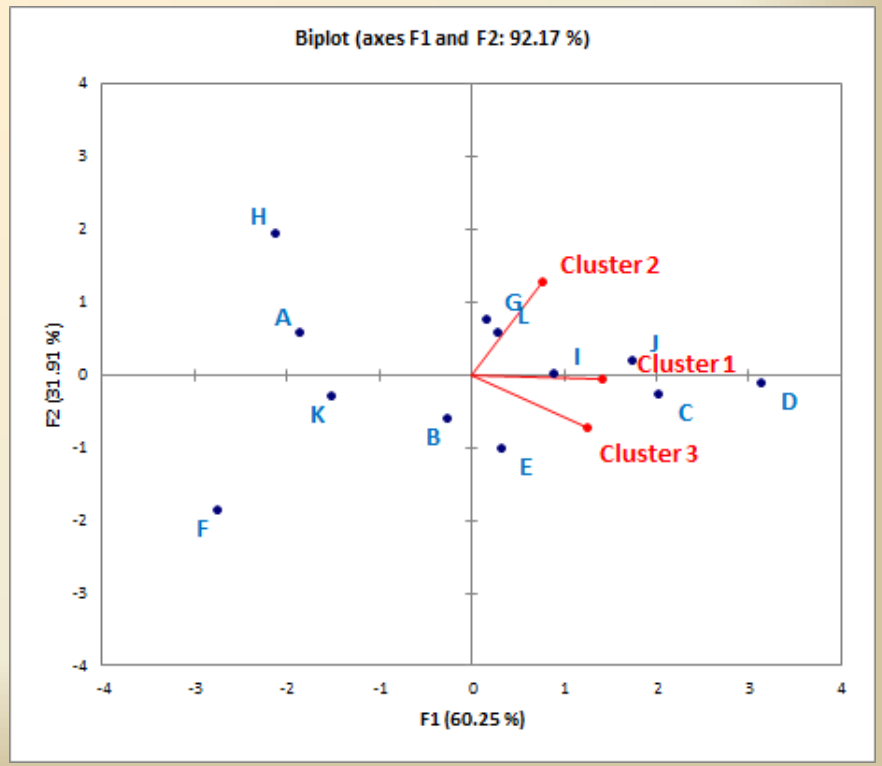
- …other times, it only captures similar clusters with scaling differences…

Example:

AHC clusters (12 samples)



PCA on matrix of sample by cluster means



Clustering Example # 2 con't

- …in this case, it may be better to use a different approach!

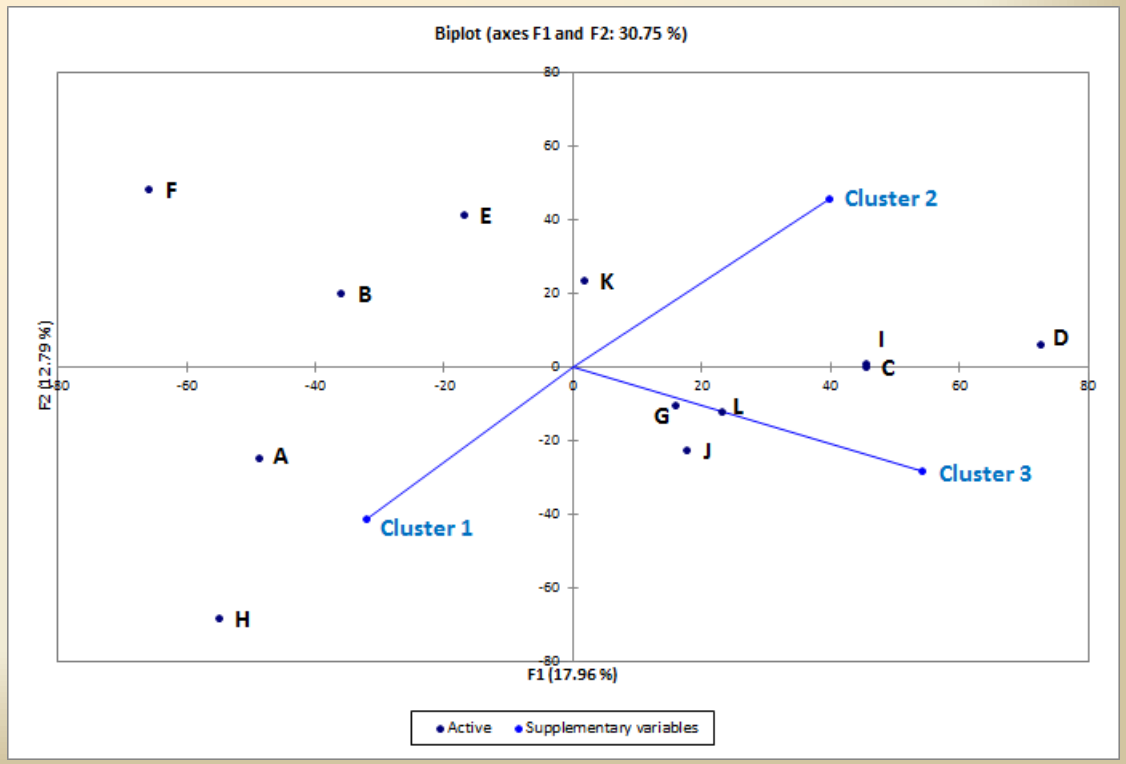
Raw Data (T)
(samp x cons scores)

PCA

Use PCA loadings to cluster consumers (AHC or k-means)

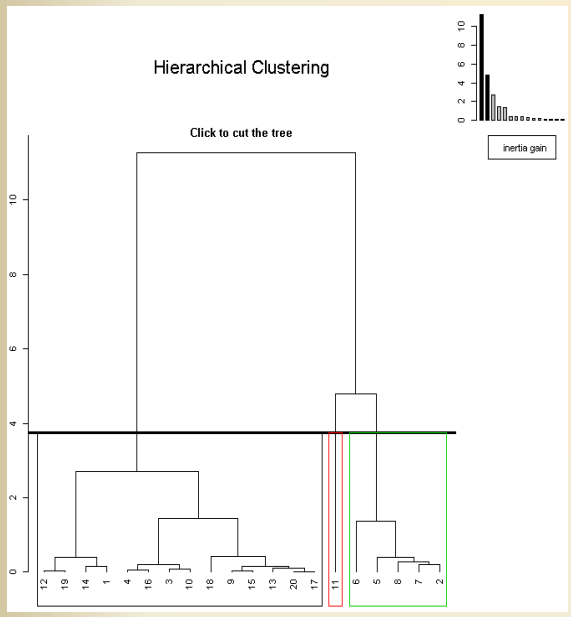
Correlate clusters onto PCA

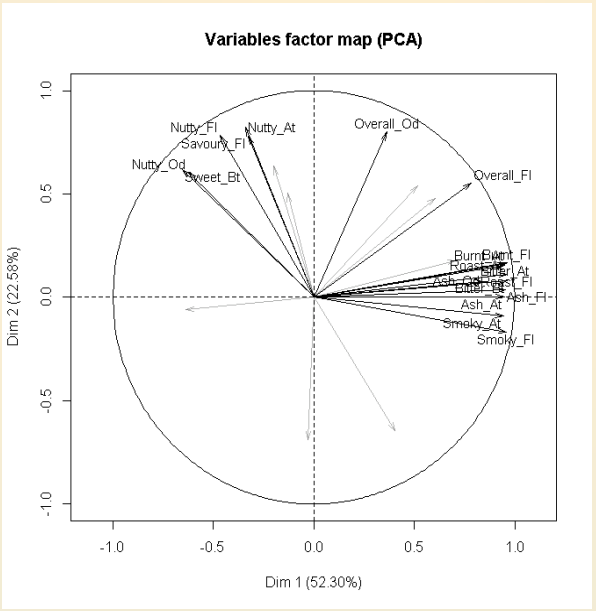
PCA on raw data with clusters as suppl. variables

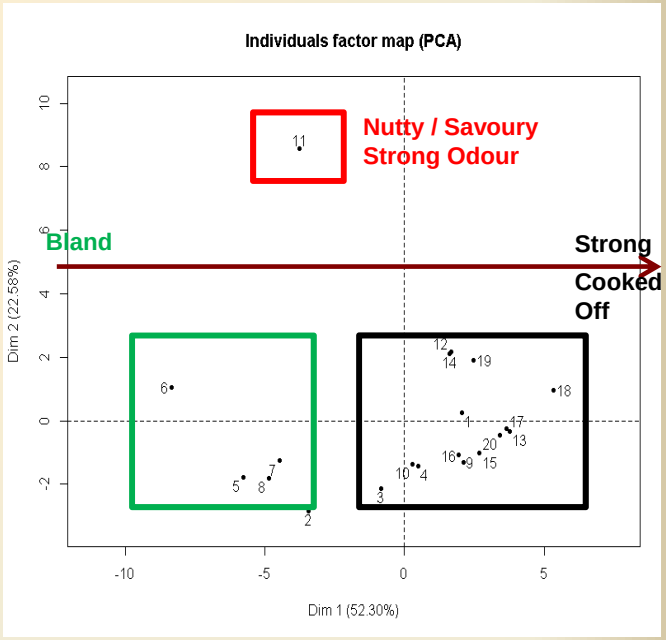


Sensory Data – Dealing with ‘extreme’ samples (1)

- Sample differences can be masked/misinterpreted when the sample set is not defined carefully
- Dataset: sensory assessment, 20 samples, 10 judges, 35 attributes
- Analyses: summary statistics, 2-way ANOVA (5%) with interactions, PCA on transposed data, PCA , HCPC, MFA (scaled)
- Below: all samples and modalities – only significant attributes represented (x 26)

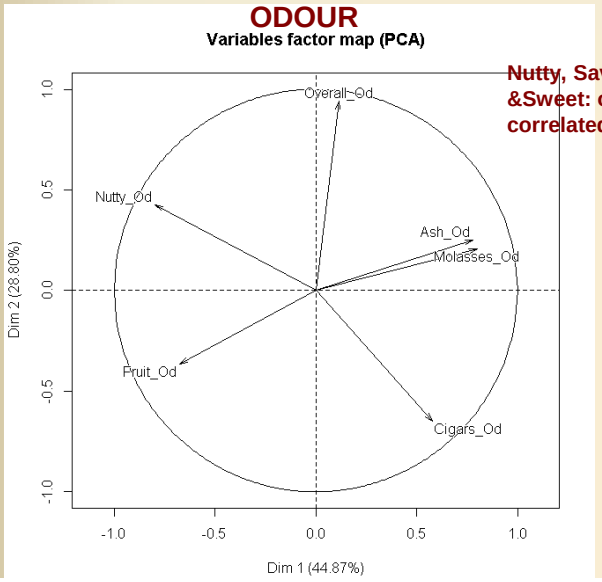


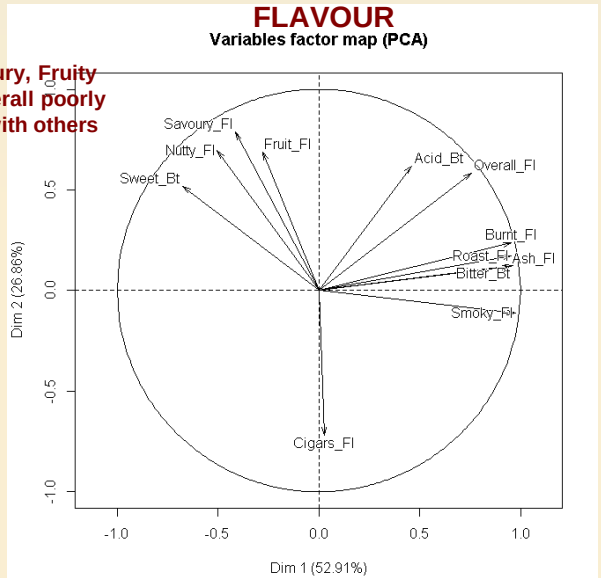


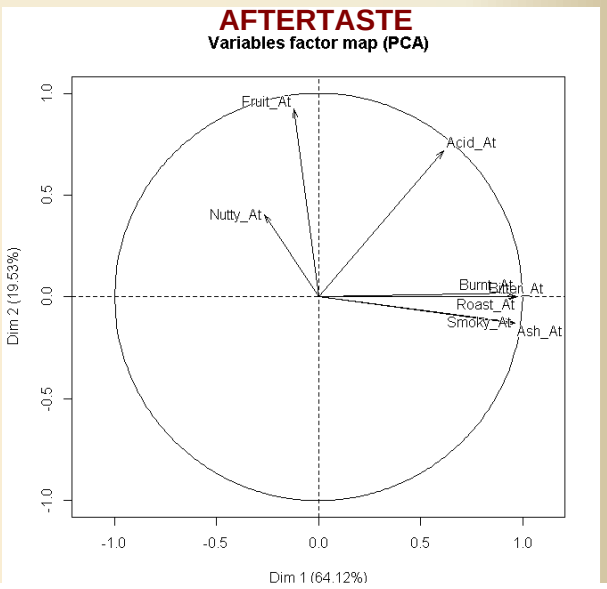


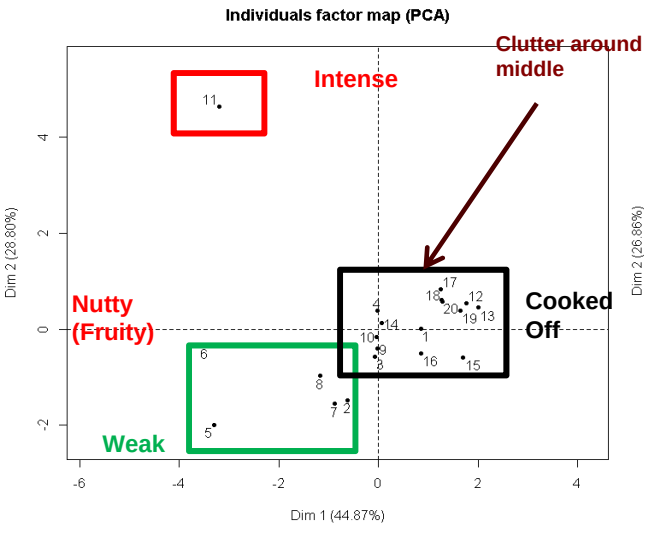
Sensory Data – Dealing with ‘extreme’ samples (2)

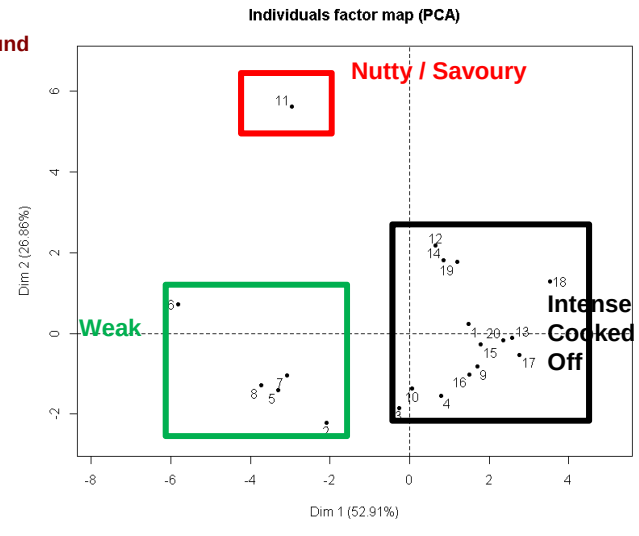
- Per modality – only significant attributes (PCA):

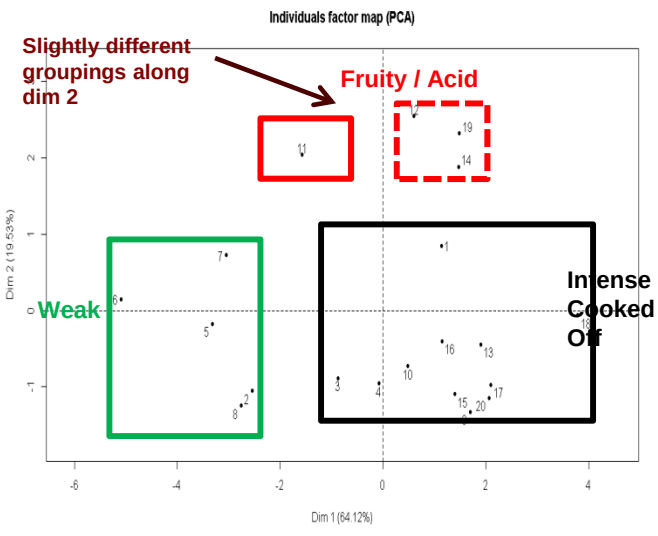






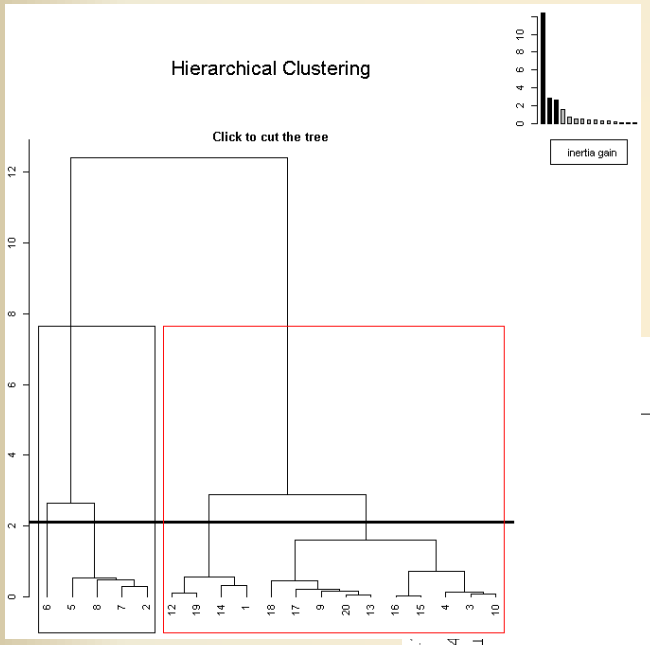


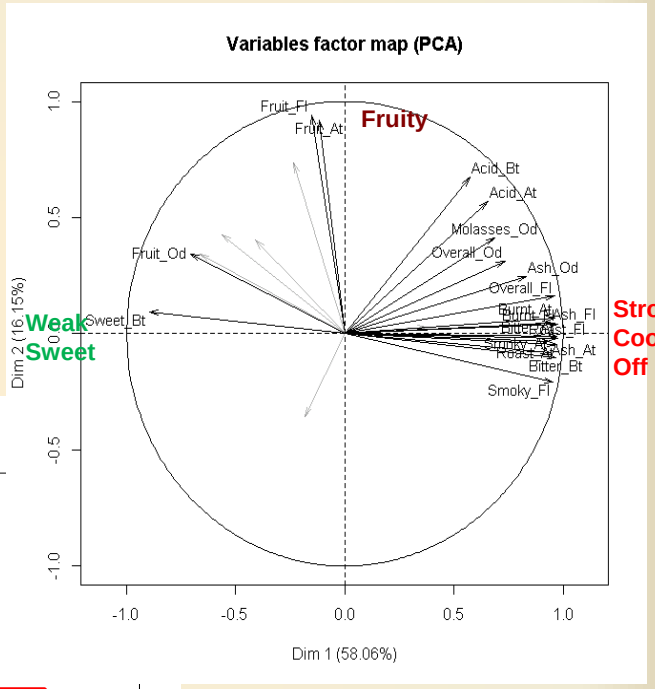


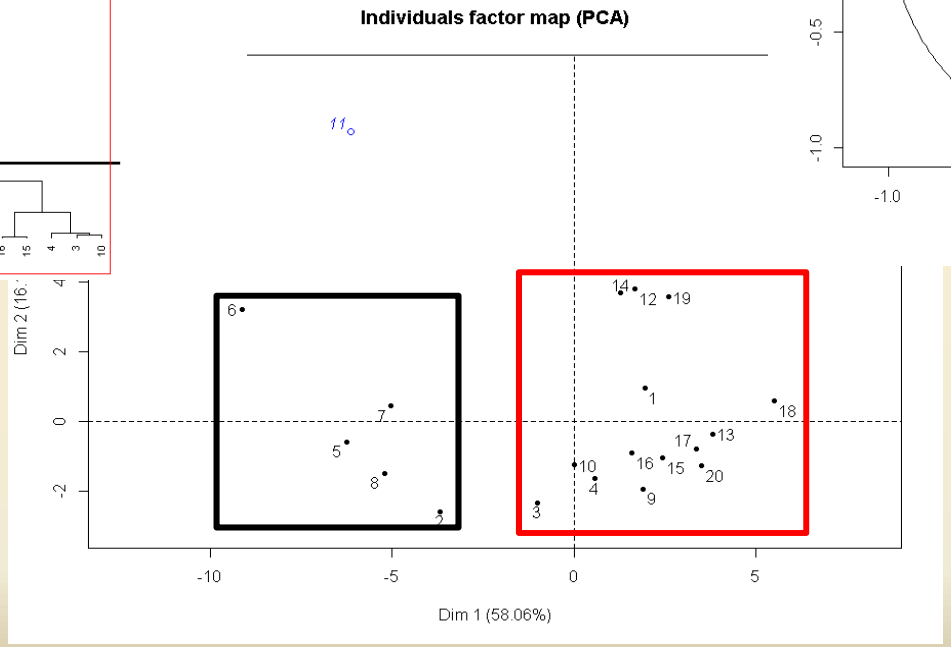


Sensory Data – Dealing with ‘extreme’ samples (3)

- With extreme sample (P11) as supplementary (PCA):







Conclusions

- Study potential differences in liking/disliking between the samples.
- Deal with consumer segments and extreme or polarizing samples.
- Explore different ways of clustering the data.
- Ensure your method of analysis addresses your true objective.
- Take the time to dig a little deeper!

Thank you for your attention!

cgilbert@acceintl.com marleen.chambault@campdenbri.co.uk