A Brief Overview of Perceptual Mapping with Applications

Prepared for

NWGIC

by

Tom Carr

March 11, 2010

Perceptual Mapping

- A statistical technique for summarizing sensory descriptive evaluations of products.
  - Groups large numbers of attributes onto a small number of key sensory dimensions based on how the attributes correlate with each other.
  - Mapping explains maximum amount of sensory variability with fewest number of dimensions.
  - Locates the products on the key dimensions to illustrate similarities and differences in their sensory properties.
- Perceptual mapping uses Factor Analysis (FA) to analyze the sensory descriptive data.

What Does the Product Landscape Look Like?

- How are products in the category similar and different from each other?
- How uniform is my product?
  - Plant to Plant?
  - Within-Plant?
- How does my product age?
  - Do different storage conditions follow the same path over time?
- How does distribution affect my product?
  - Is my product the same regardless of where it is purchased?

How are Products Distributed in the Category Space?

- Are my products so similar that they are competing with each other?
- Are there “holes” in the product space that offer niche opportunities?
- Does a competitor “own” a part of the product space where I could compete?

How Do Wines Differ?

- What are the differences within and between the varieties?
- What are the difference among growing regions?
- What are the differences among vintages?

Creating the Perceptual Map

1. Perform factor analysis to ID key sensory dimensions.
2. Obtain coordinates of samples for mapping.
3. Plot samples on perceptual maps.

Identify Key Sensory Dimensions

- Steps in Factor Analysis.
  - Select variables (i.e., sensory attributes) to analyze.
  - Control output with six “options”.
    - » Generate ***Scree Plot*** to help select number of dimensions
    - » ***Rotate*** solution to make it easier to interpret
    - » ***Reorder*** factor loadings to make it easier to interpret solution
    - » Display only meaningfully large factor loadings using ***Fuzz***
    - » Select ***Number*** of dimensions
    - » ***Output*** factor scores of samples

What Attributes Should Be Included?

- No universally accepted criteria.
- Options:
  - Include them all.
  - Include only attributes with significant sample differences.
  - Include only attributes with meaningful ranges and some intensities above threshold.
- Try different approaches and pick the one that makes the most sense.

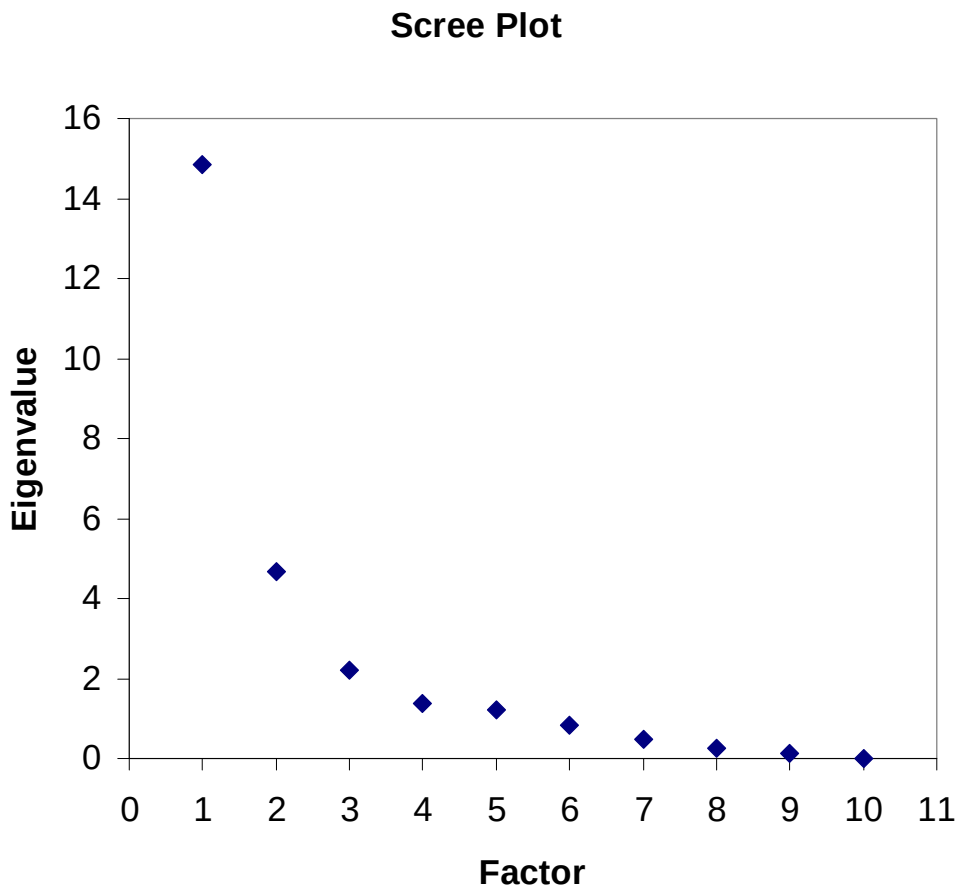
Selecting Number of Factors

- Total Variance Explained
  - Initial Eigenvalues
    - » Look for big drops.
    - » Cumulative > 75%.
    - » Individual eigenvalues > 1.0.
- Results suggest 2 to 5 factors should be retained.

Factor	Eigenvalue	Difference	Proportion	Cumulative
1	14.86	10.19	0.57	0.57
2	4.67	2.45	0.18	0.75
3	2.22	0.83	0.09	0.84
4	1.38	0.18	0.05	0.89
5	1.20	0.39	0.05	0.94
6	0.82	0.35	0.03	0.97
7	0.47	0.20	0.02	0.99
8	0.27	0.15	0.01	1.00
9	0.12	0.12	0.00	1.00
10	0.00	0.00	0.00	1.00

Selecting Number of Factors

- Scree Plot
  - Look for bend (or elbow) at bottom of drop.
- Results suggest 2 or 3 factors should be retained.
- Next Steps
- Rerun analyses with 2 & 3 factors.
  - Pick the solution that makes the most sense based on both sensory and statistical criteria.



Factor Loadings – Three Factor Solution (84%)

- Three Key Dimensions
  - *"Artificial to Natural"*
  - *"Clarity to Viscosity"*
  - *"Oily"*

ATTR#	Attribute	Artificial to Natural	Clarity to Viscosity	Oily
X8	FL Meat Identity	0.89	.	.
X9	FL Brothy	0.88	.	.
X28	TXT Meat Firm	0.84	.	.
X6	APP Potato Size	0.82	.	.
X7	APP Green Herb AMT	0.80	.	.
X4	APP Solid Size (no Potato)	0.79	.	.
X30	TXT Potato Firm	0.79	.	.
X15	FL Carrot	0.79	.	.
X16	FL Celery	0.78	0.62	.
X21	FL Wheat	0.74	.	.
X29	TXT Vegetable Firm	0.73	.	.
X12	FL Green Herbs	0.72	0.64	.
X10	FL Meat Non-Natural	-0.85	.	.
X26	TXT Viscosity (Sauce)	.	0.89	.
X14	FL Black Pepper	.	0.88	.
X13	FL Spice Blend	.	0.87	.
X23	FL Bitter	.	0.86	.
X25	FL Sour	.	0.86	.
X11	FL Filler	.	0.86	.
X3	APP Solid AMT (no Potato)	.	0.76	.
X17	FL Onion	0.66	0.69	.
X2	APP Sauce Clarity	.	-0.92	.
X1	APP Surface Oil	.	.	0.92
X27	TXT Oily MF	.	.	0.87
X22	FL Salty	.	.	.
X5	APP Potato AMT	.	.	.

Three Dimensions Define the Sensory Space

- Three dimensions explain 84% of the sensory variability among the ten samples.
  - *"Artificial to Natural"* 38%
  - *"Clarity to Viscosity"* 35%
  - *"Oily"* 11%

Dimension 1: *"Artificial to Natural"*

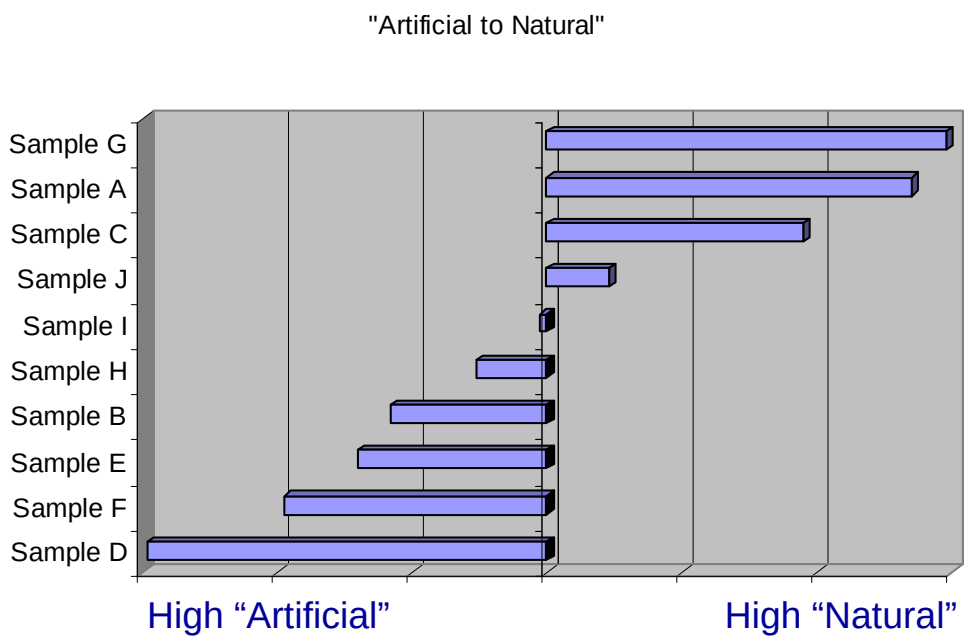
- The first sensory dimension is bipolar. Samples range from high "*Artificial*" to high "*Natural*" character. Samples high in "*Artificial*" character are low in "*Natural*" character and *vice versa*.

Dimension 1: (37%)
"Artificial to Natural"

- The attributes that define the dimension and their factor loadings are:
  - FL Meat Non-Natural -0.85
  - FL Green Herbs 0.72
  - TXT Vegetable Firmness 0.73
  - FL Wheat 0.74
  - FL Celery 0.78
  - FL Carrot 0.79
  - TXT Potato Firmness 0.79
  - APP Solid Size 0.79
  - APP Green Herb Amount 0.81
  - APP Potato Size 0.82
  - TXT Meat Firmness 0.84
  - FL Brothy 0.88
  - FL Meat Identity 0.89

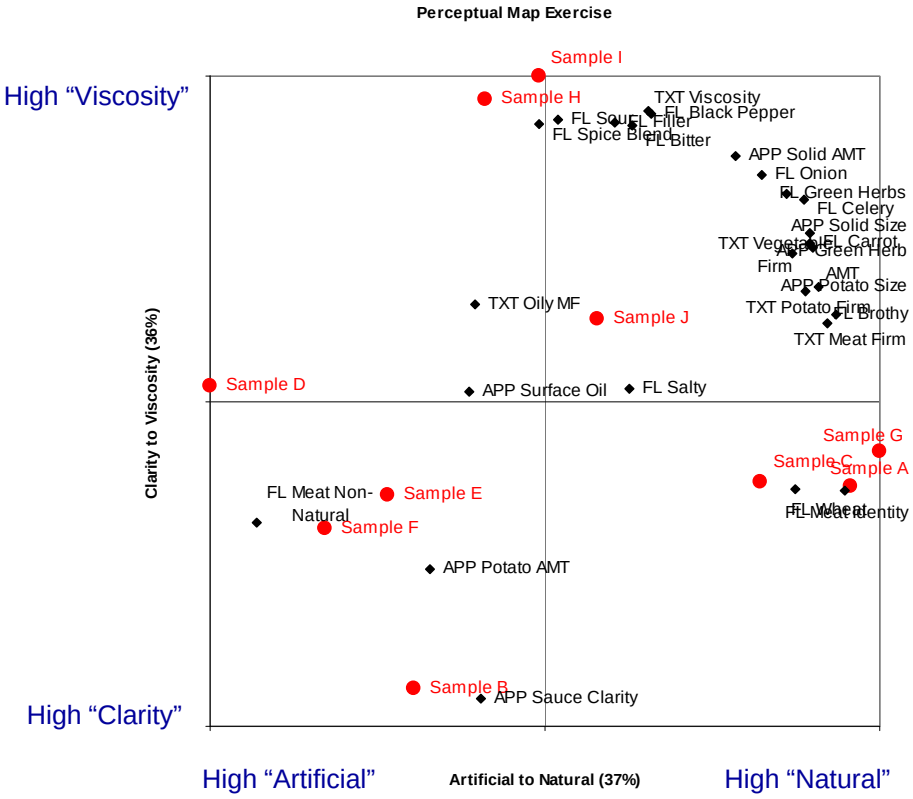
"Artificial to Natural" Dimension

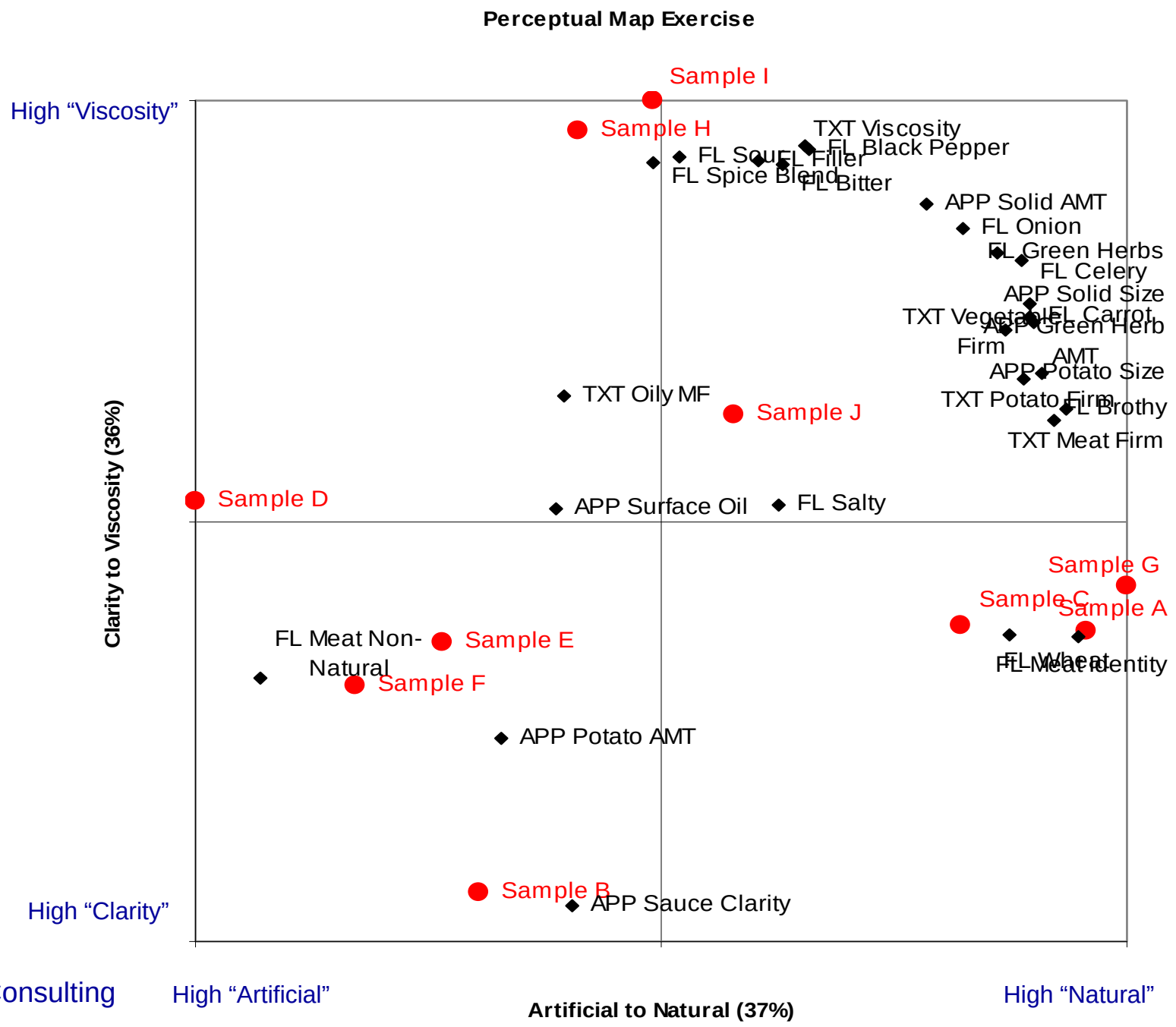
- Sample D is lowest in this dimension.
  - i.e. High in "*Artificial*"/Low in "*Natural*"
- Samples A, C and G are high in this dimension.
  - i.e. Low in "*Artificial*"/High in "*Natural*"



"Artificial to Natural" by *"Clarity to Viscosity"*

- Sample B is high in *"Clarity"* and mid-low in *"Artificial to Natural".*
- Sample D is high in "*Artificial*" and mid-range in *"Clarity to Viscosity".*
- Samples H and I are high in "Viscosity" is mid-range in "*Artificial to Natural*".
- Samples A, C and G are high in *"Natural"* and mid-range in *"Clarity to Viscosity".*




Perceptual Map Exercise
Sample I
High "Viscosity"
Sample H
TXT Viscosity
FL Black Pepper
FL Soup
FL Filler
FL Spice Blend
FL Bitter
APP Solid AMT
FL Onion
FL Green Herbs
FL Celery
APP Solid Size
TXT Vegetable
FL Carrot
APP Green Herb
Firm
AMT
APP Potato Size
TXT Oily MF
TXT Potato Firm
Sample J
FL Brothy
TXT Meat Firm
Sample D
APP Surface Oil
FL Salty
Clarity to Viscosity (36%)
Sample G
Sample C
Sample A
FL Meat Non-
Natural
Sample E
FL Wheat
FL Meat Identity
Sample F
APP Potato AMT
Sample B
High "Clarity"
APP Sauce Clarity
High "Artificial"
Artificial to Natural (37%)
High "Natural"

Cluster Analysis to Group Products

Cluster Analysis Overview

- Factor analysis groups attributes.
- Cluster analysis groups samples or respondents.

Hierarchical Clustering

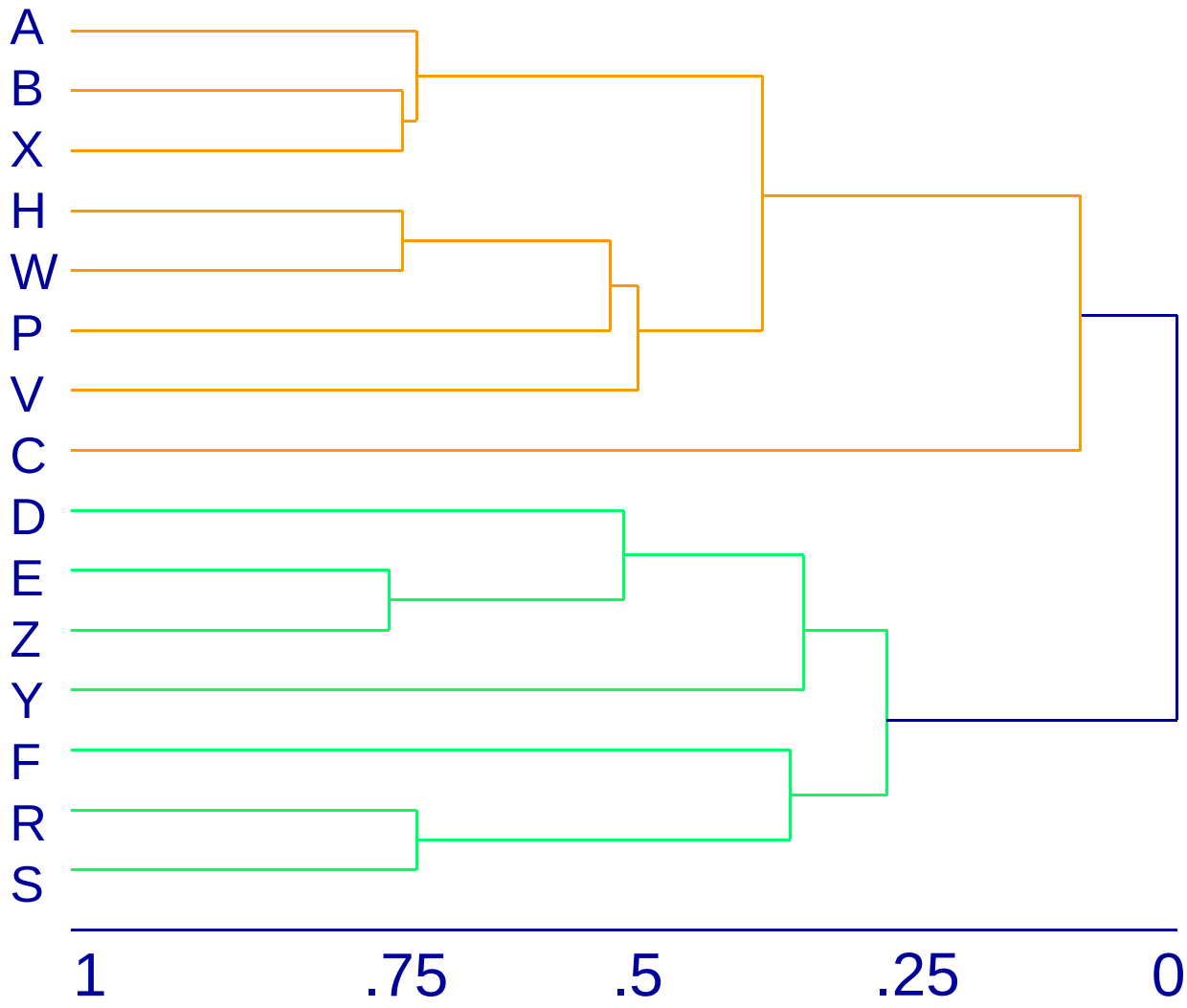
- Distance Measures
  - Euclidean (normal or squared)
  - Pearson (normal or squared)
  - Manhattan (city block approach)
- Linkage Techniques
  (i.e., How clusters are joined)
  - Single Linkage (Nearest Neighbor)
  - Complete Linkage (Farthest Neighbor)
  - Average Linkage
  - Centroid Linkage
  - Median Linkage
  - McQuitty's Linkage
  - Ward's Linkage ← RECOMMENDED

Dendrogram

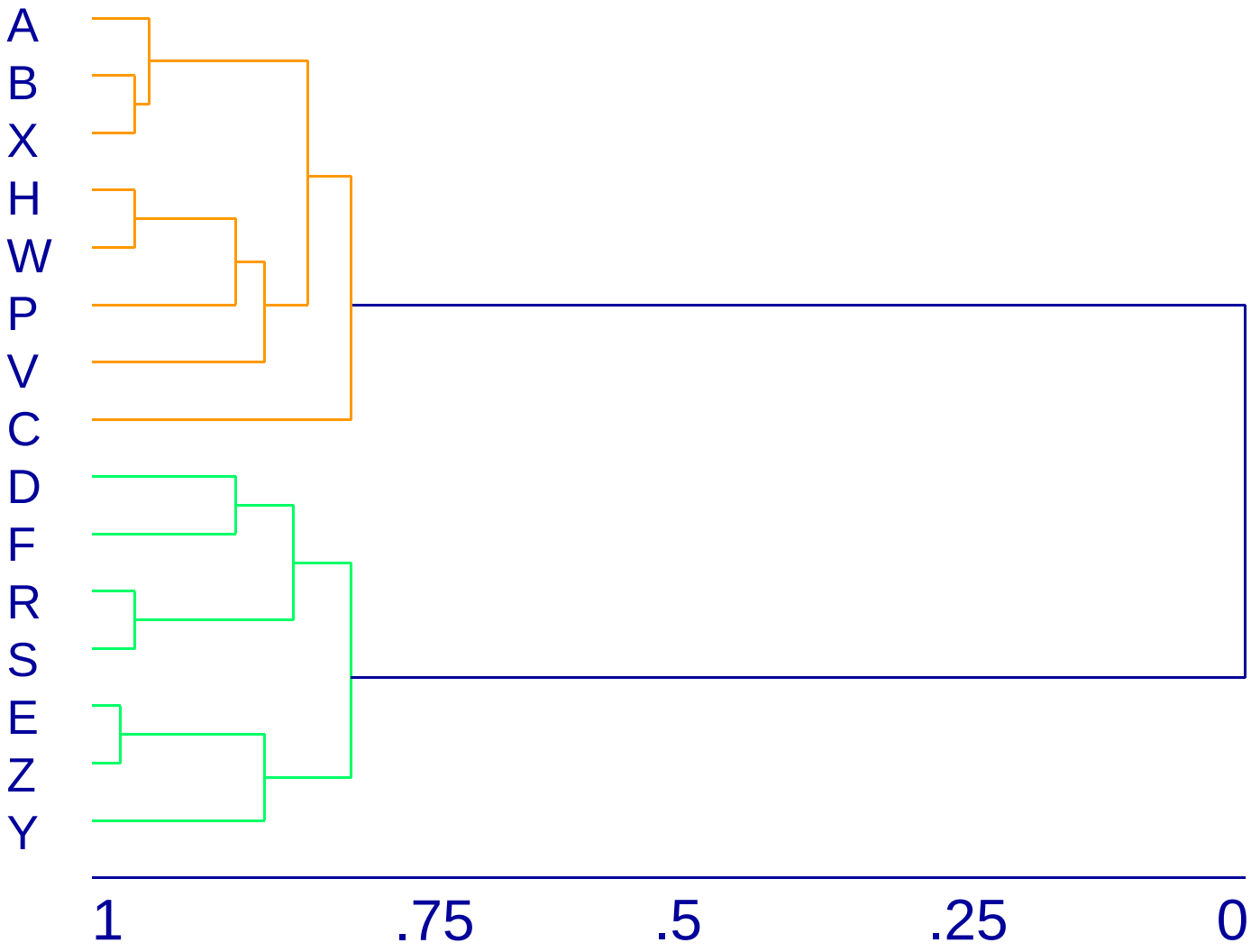
- Graphical depiction of linkages.
- Displays when samples or respondents are merged.
- Used to decide how many clusters.

Example Clustering Samples

Dendrogram – Average Linkage



Dendrogram – Ward's Method



Average Attribute Ratings Distinguish the Clusters

- Cluster 1: Products A, B, C, H, P, V, W, X
- Cluster 2: Products D, E, F, R, S, Y, Z

Intensity	Cluster 1	Cluster 2	
Greasy	2.4	1.8	***
Tacky	2.1	1.9	***
Coated	2.5	2.2	***
Smooth	3.7	3.6	
Firm	3.0	3.1	
Silky	2.7	2.7	

*** = Significantly Different

Summary

- Both Methods Yielded the Same Clusters.
  - Average linkage suggests three clusters with C by itself.
  - Ward's method more strongly suggests two clusters.
- Clusters Differed in Greasy/Tacky/ Coated Intensities.
  - Samples in Cluster 1 are creams.
  - Samples in Cluster 2 are lotions.

Example of Perceptual Mapping of Wines

Steps in the Analysis

- Factor Analysis to Create the Perceptual Map.
- Cluster Analysis Group Products on the Perceptual Map.
- ANOVA to Understand the Differences among the products.

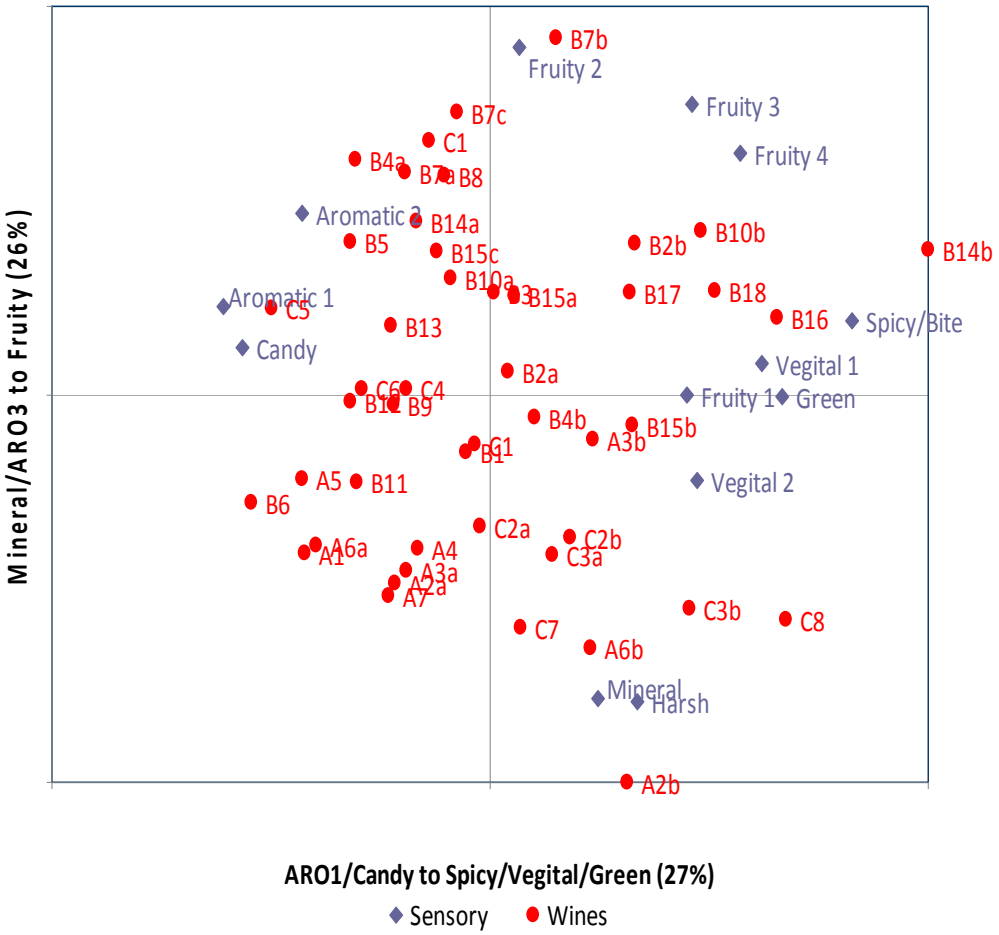
The Analyses

- Performed a FA on the average profiles from the full descriptive panel.
  - Two factors explained 53% of the variability in the sensory profiles.
- Performed Cluster Analysis on the factor scores from the FA.
  - Identified two primary clusters.
    - » One primary cluster had two sub-clusters; one had three sub-clusters.

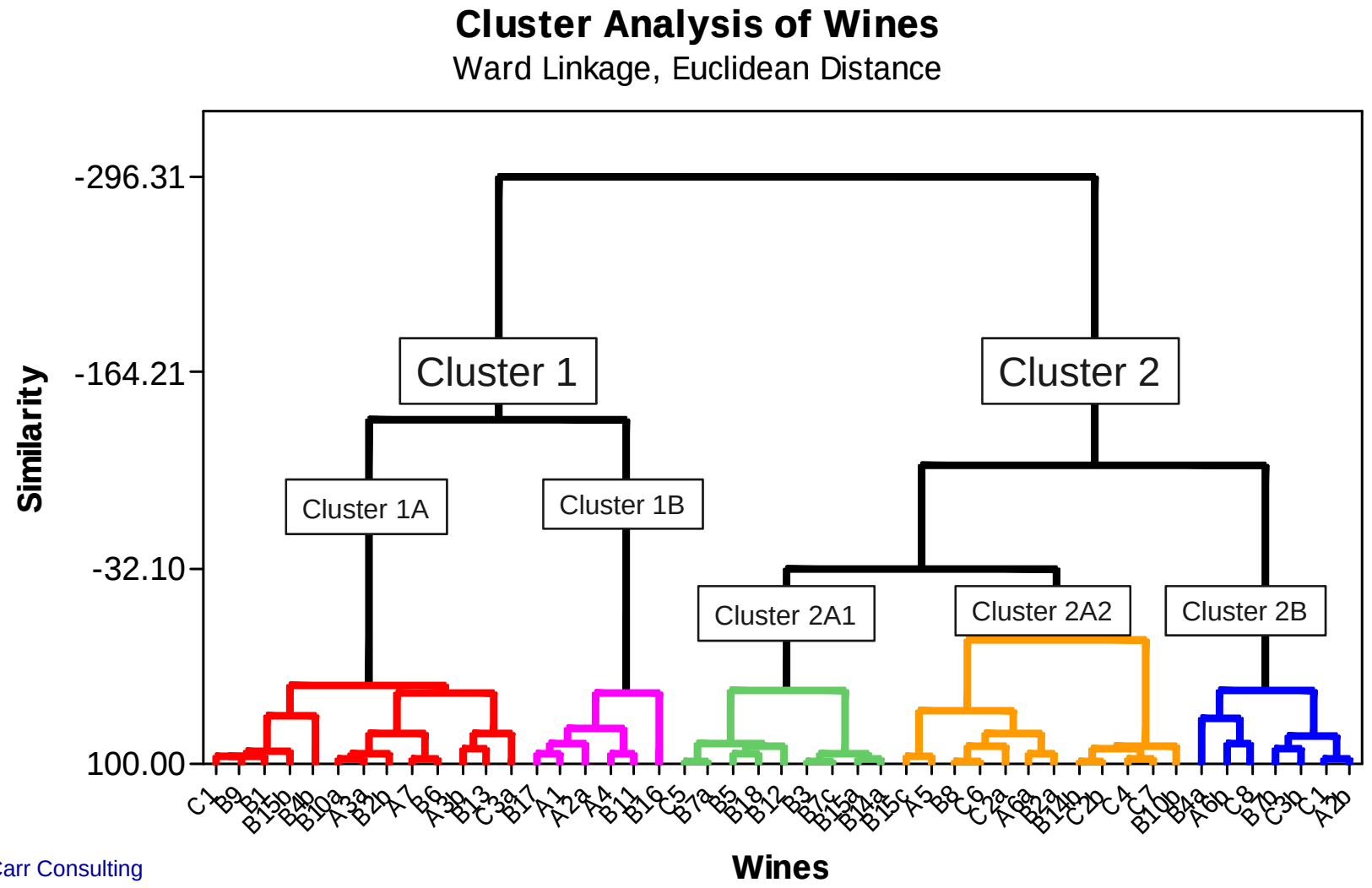
Factor Loadings from the PCA

Attribute	ARO1/Candy to Spicy/Vegital/Green	Mineral/ARO3 to Fruity
Spicy/Bite	**0.83**	0.19
Green	**0.67**	-0.01
Vegital 1	**0.62**	0.08
Vegital 2	**0.47**	-0.22
Fruity 1	**0.45**	0.00
Candy	**-0.57**	0.12
Aromatic 1	**-0.61**	0.22
Fruity 2	0.07	**0.89**
Fruity 3	0.46	**0.75**
Fruity 4	0.57	**0.62**
Aromatic 2	-0.43	**0.46**
Mineral	0.25	**-0.78**
Harsh	0.34	**-0.79**
Variability Explained	27%	26%

Perceptual Map of the Wines



Cluster Analysis Results



Cluster 1 Wines

- Wines in Cluster 1 are predominantly from Region B.
- The sub-clusters distinguish the 2004 and 2005 vintages.

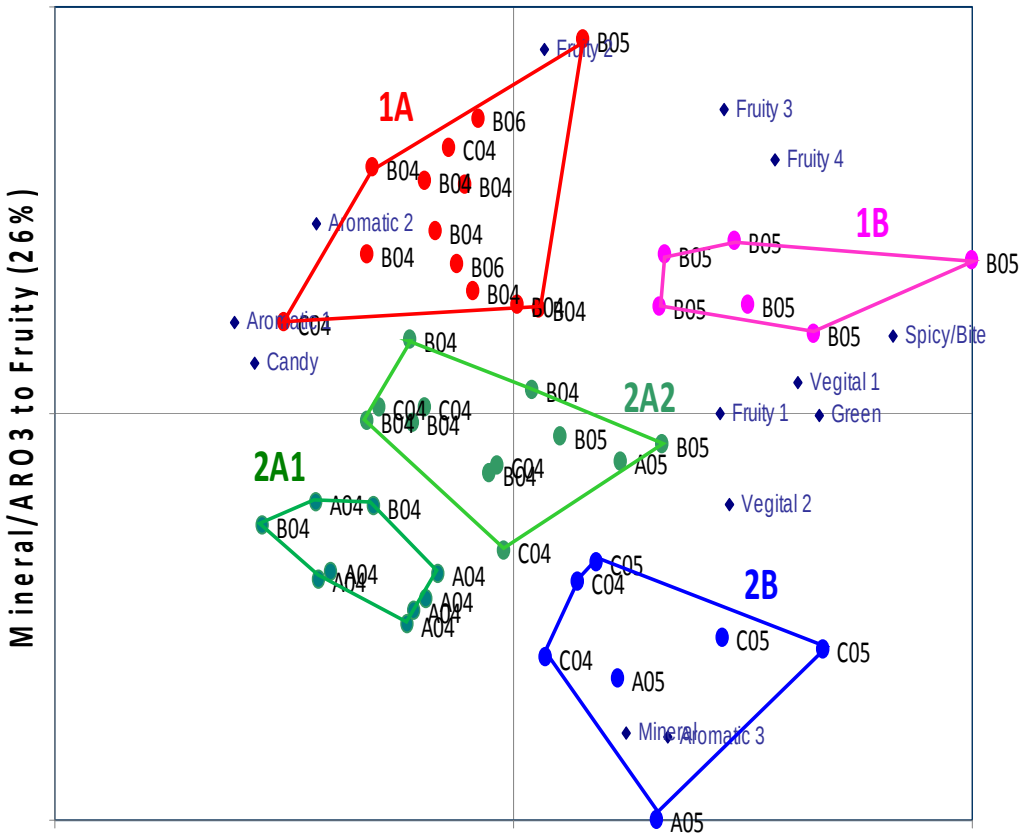
Primary Cluster	Clusters	Wine	Vintage
1	1A	C1	2004
1	1A	C5	2004
1	1A	B3	2004
1	1A	B4a	2004
1	1A	B15a	2004
1	1A	B15c	2006
1	1A	B14a	2004
1	1A	B5	2004
1	1A	B7a	2004
1	1A	B7b	2005
1	1A	B7c	2006
1	1A	B8	2004
1	1A	B10a	2004
1	1B	B17	2005
1	1B	B18	2005
1	1B	B14b	2005
1	1B	B10b	2005
1	1B	B2b	2005
1	1B	B16	2005

Cluster 2 Wines

- Wines in Cluster 2 are mainly from Region A and Region C.
  - Although, sub-cluster 2A2 has a high proportion of wines from Region B.
- Sub-clusters 2A1 and 2A2 are predominantly 2004 vintage.
  - The wines in sub-cluster 2A1 are mainly from Region A.
  - The wines in sub-cluster 2A2 are mainly from Regions B and C.
- Sub-cluster 2B is predominantly 2005 vintage with wines mainly from Region C.

Primary Cluster	Clusters	Wine	Vintage
2	2A1	A7	2004
2	2A1	A5	2004
2	2A1	A3a	2004
2	2A1	A2a	2004
2	2A1	A4	2004
2	2A1	A1	2004
2	2A1	A6a	2004
2	2A1	B11	2004
2	2A1	B6	2004
2	2A2	A3b	2005
2	2A2	B2a	2004
2	2A2	B13	2004
2	2A2	B12	2004
2	2A2	B1	2004
2	2A2	B4b	2005
2	2A2	B15b	2005
2	2A2	B9	2004
2	2A2	C4	2004
2	2A2	C6	2004
2	2A2	C2a	2004
2	2A2	C1	2004
2	2B	A2b	2005
2	2B	A6b	2005
2	2B	C7	2004
2	2B	C3a	2004
2	2B	C2b	2005
2	2B	C3b	2005
2	2B	C8	2005

The Clusters are Well Separated in the Sensory Space



NOTE: Plot symbols are Region Code-Vintage

Comparing the Primary Clusters

- Primary Cluster 1 (essentially Region B wines) is higher in Fruity 2, Fruity 3, Fruity 4, Spicy/Bite and Aromatic 2; while Primary Cluster 2 (essentially Regions A and C) is higher in Mineral and Aromatic 3.
  - The primary clusters differ on the attributes associated with the second sensory dimension.
  - The first sensory dimension is primarily distinguishes the 2004 and 2005 vintages.

	Primary Cluster				
Attribute	Region B	Region A & Region C	P-Value	LSD	
Fruity 3	63 A	52 B	0.000	3	**
Spicy/Bite	32 a	30 b	0.078	2	*
Aromatic 2	43 A	37 B	0.010	5	**
Fruity 4	51 A	44 B	0.000	3	**
Green	26	25	0.663	na	
Mineral	20 B	27 A	0.000	3	**
Harsh	20 B	26 A	0.000	3	**
Candy	23	25	0.320	na	
Fruity 2	37 A	24 B	0.000	3	**
Fruity 1	37	36	0.392	na	
Aromatic 1	17	16	0.757	na	
Vegital 2	10	11	0.436	na	
Vegital 1	20	18	0.215	na	

Green shading = highest intensities
Red shading = lowest intensities

Summary and Conclusions

- A two-dimensional solution provides an adequate summary of the perceptible differences among the wines.
- Cluster analysis of the factor scores from the FA uncovers two primary clusters and five sub-clusters.
  - The primary clusters distinguish wines from Region B from wines from Region A and Region C.
  - One pair of sub-clusters distinguish the 2004 and 2005 vintages of wines from Region B.
  - Two of the remaining three sub-clusters distinguish the 2004 and 2005 vintages of wines from Region C.
  - The last sub-cluster is composed primarily of 2004 wines from Region A.