



CAMO

3-Way PLS Regression in Consumer Preference Analysis

Valérie Lengard

Sensometrics, Davis 2004

vl@camo.no



Outline

- Data description
- Analysis
 1. Segment selection
 2. Extraction of positive and negative drivers of liking
 3. Key demographics, attitudinal and usage characteristics
- Conclusions



Data description

- 17 tomato varieties
- 11 sensory attributes, 14 sensory panelists
- 379 consumers from 3 cities. Overall consumer liking on a scale 0-10 + attitudinal questionnaire
- Data from 2001 was used



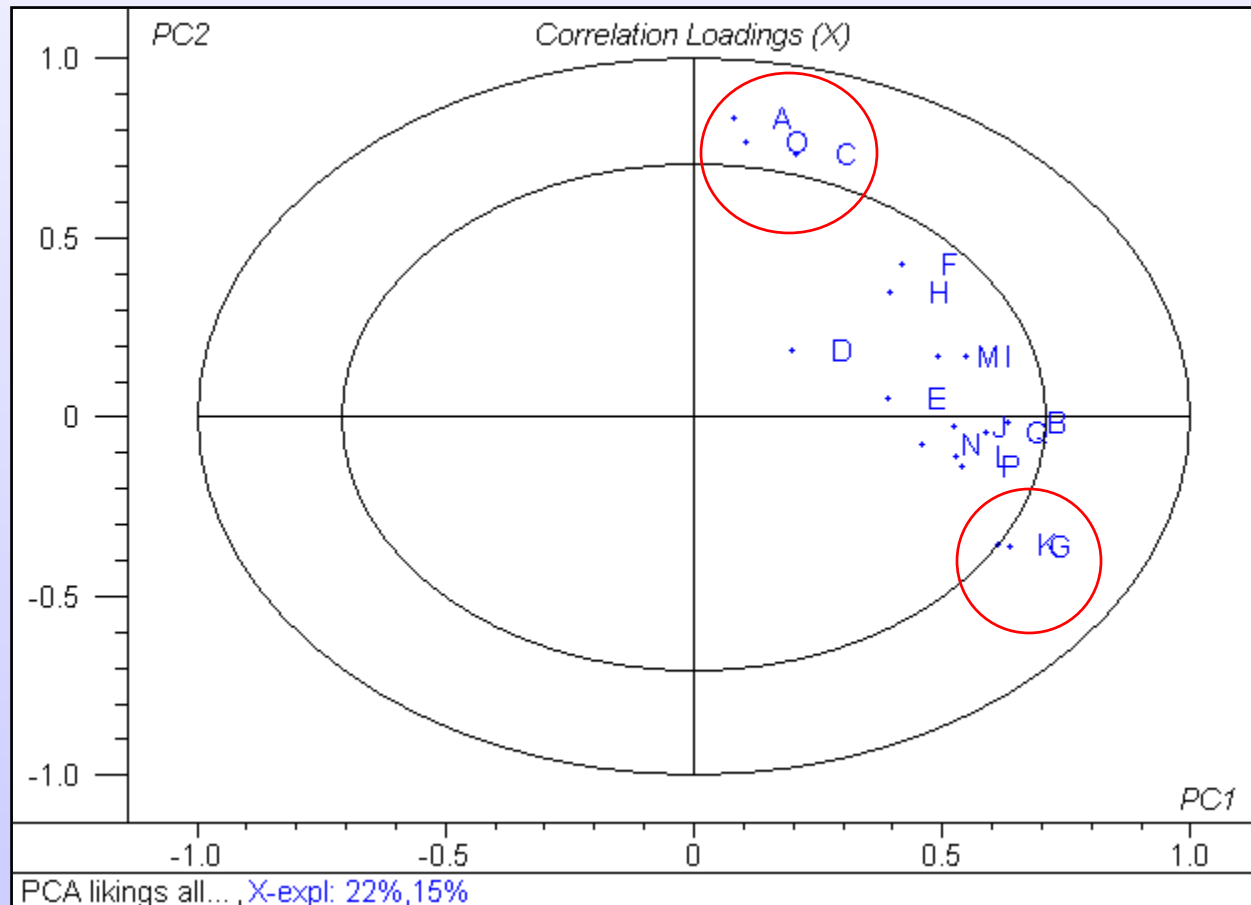
CAMO

1. Segment Selection



Segment selection (PCA)

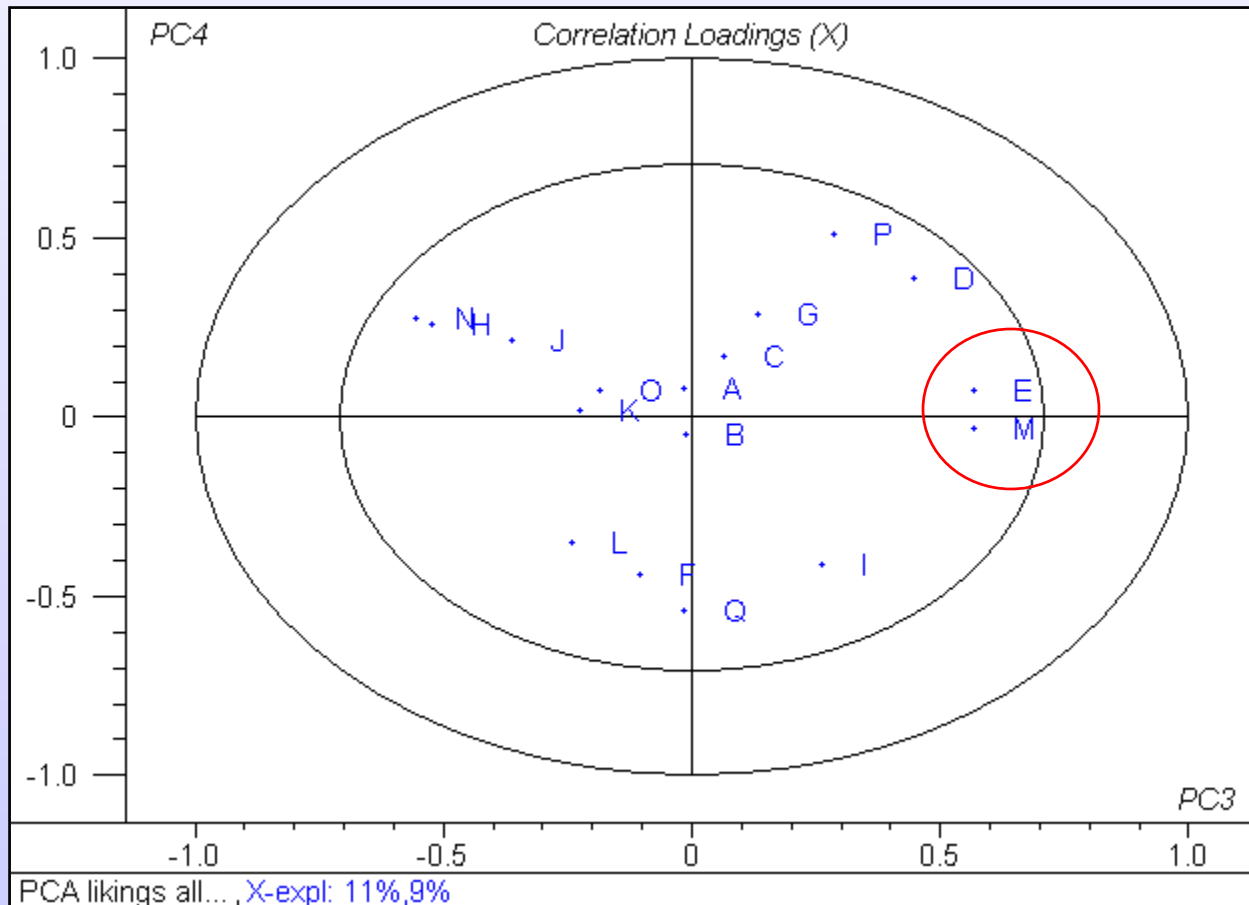
- ACO and KG groups on PC1/PC2





Segment selection (PCA)

- EM group on PC3





Segment definition

- Overall consumer liking results were used to place consumers into each segment.
- For the ACO-group, a segment was formed from consumers with a liking above 7 for at least one of the tomato varieties A, C or O
- The two other groups, KG and EM were created in a similar fashion only with the threshold lowered to 5
- Some consumers were placed in two segments
- Some consumers were not included in any of the three segments



Most liked and most disliked tomatoes


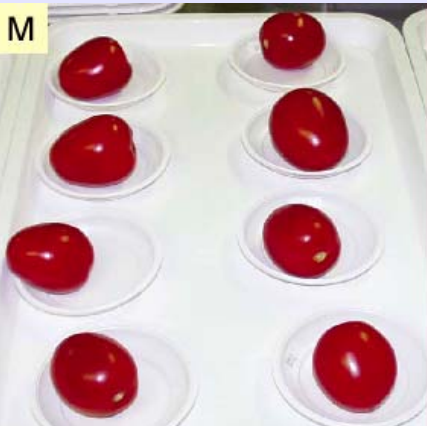

- For each of the segments, the average liking for the 17 tomato varieties was computed.
- Each segment was renamed according to its first three favorite tomatoes
- The three least liked tomatoes of each segment were identified

	ACO	EMA	GAK
Size	44%	29%	27%
Disliked	EPG	PDK	EDP



CAMO

Tomato shapes

<u>ACO</u>	<u>EMA</u>	<u>GAK</u>
<p data-bbox="295 753 330 796">A</p>  <p data-bbox="349 1182 614 1243">Cocktail</p>	<p data-bbox="765 753 799 796">M</p>  <p data-bbox="865 1182 1058 1243">Ovoid</p>	<p data-bbox="1300 753 1335 796">G</p>  <p data-bbox="1315 1182 1580 1243">Regular</p>

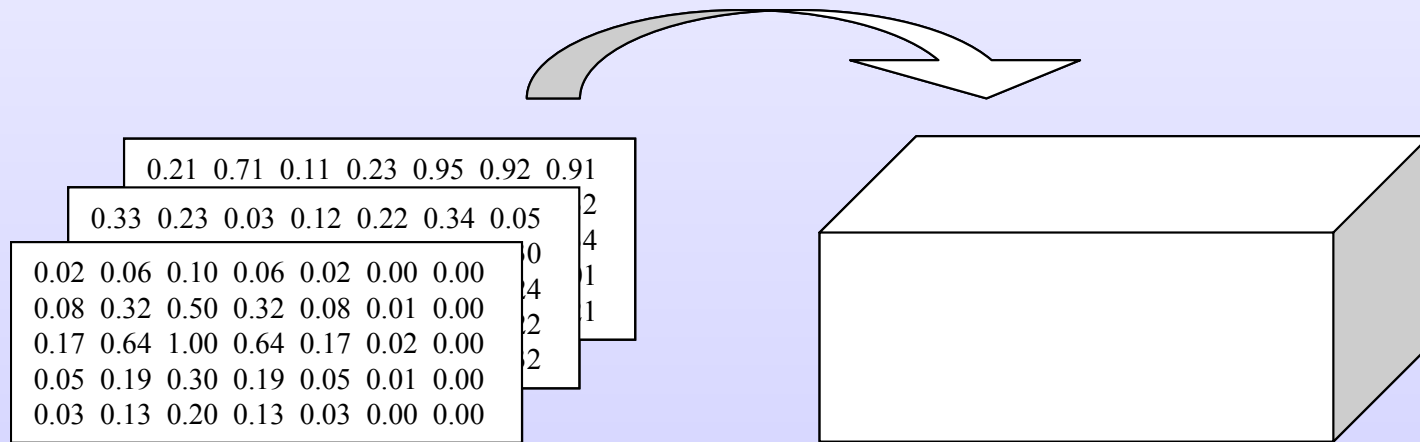


2. Extraction of positive and negative drivers of liking



3-way Sensory Data

- The sensory data was used in its 3-way structure:
17 tomatoes x 11 attributes x 14 panelists



- The model will allow to extract the drivers of liking + to check panel agreement on the evaluations



3-way PLS regression

- Samples: 17 tomatoes
- Predictors (X): Sensory attributes x Panelists
- Y: Average tomato likings per segment

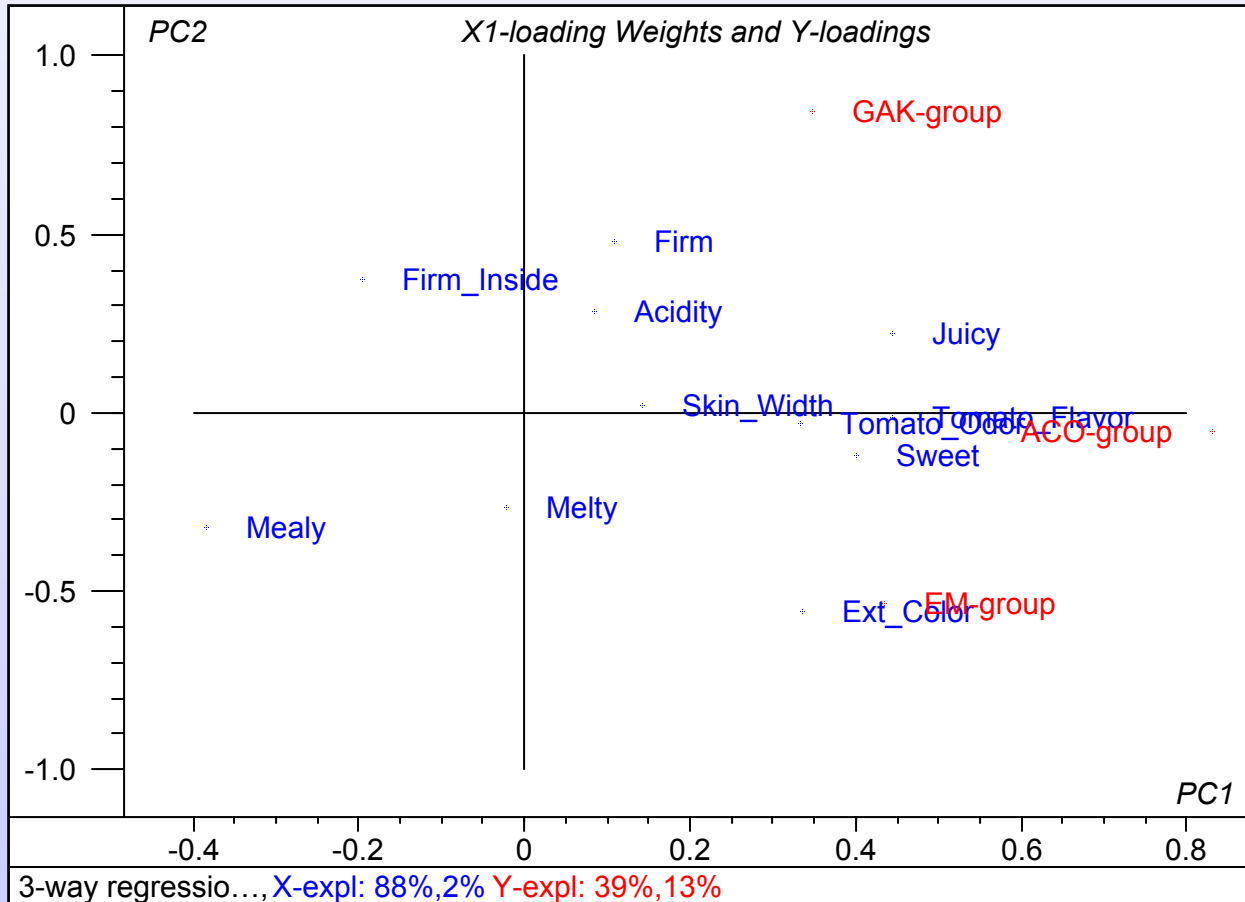
The screenshot shows the 'Regression (Three-Way PLS)' dialog box with the following settings:

- Y-Variable File:** AverageSegmentsLiking (with a 'Browse...' button)
- Variable Set:** All Variables [3] (with a 'Define...' button)
- Keep Out of Calculation:** (empty field with a 'Select...' button)
- Weights:** All 1.0/SDev (with a checked 'Include Weights' checkbox)
- Validation Method:** Cross Validation (with 'Setup...' buttons for 'Leverage Correction', 'Uncertainty test: Opt PCs', and 'Test Set')
- Model Size:** Full (dropdown menu)
- Num PCs:** 4 (spin box)
- Center Data
- Add Start Noise
- Issue Warnings (with a 'Warning Limits...' button)

Buttons on the right side: OK, Cancel, Help.

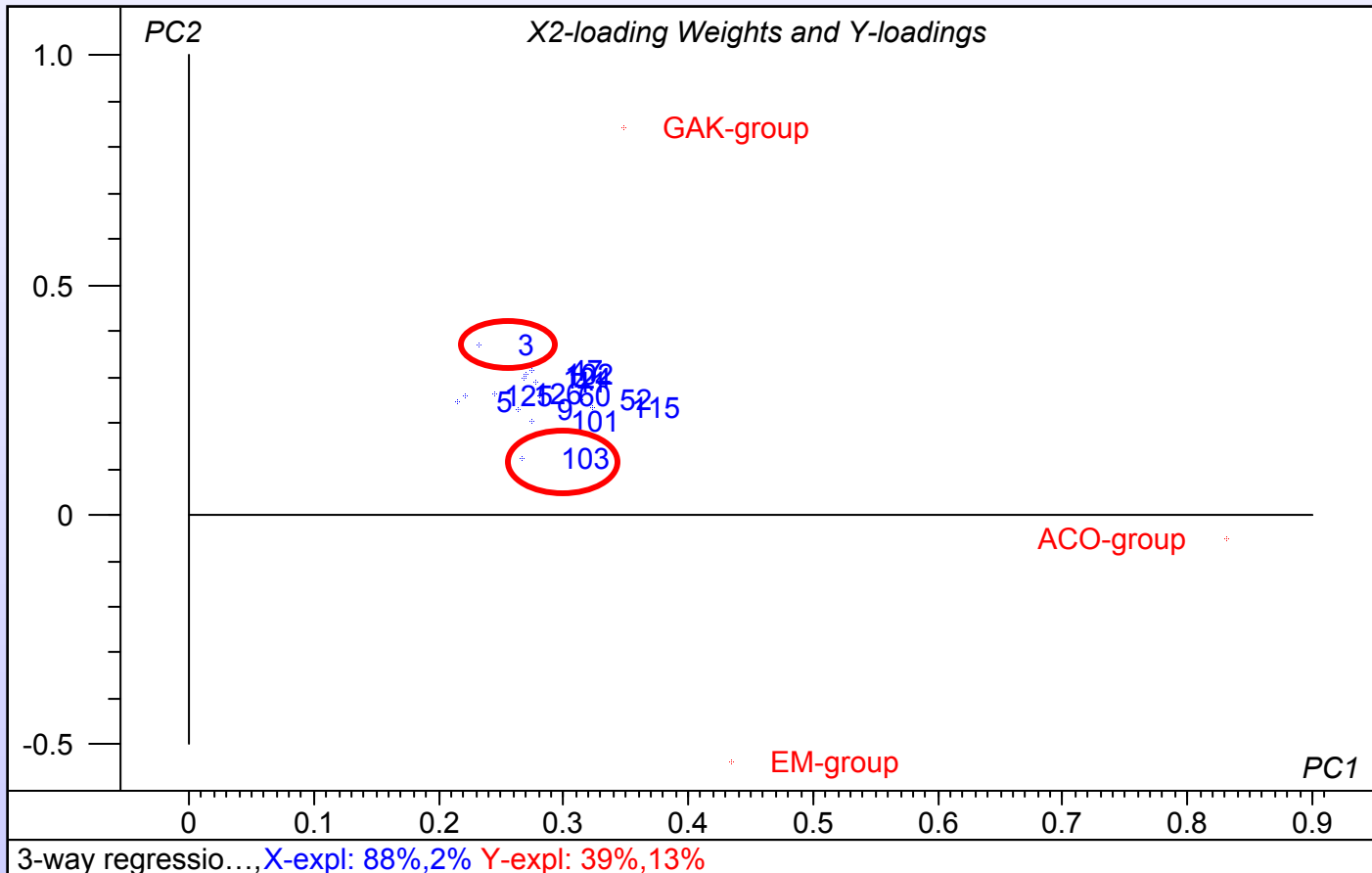


Positive and negative drivers of liking



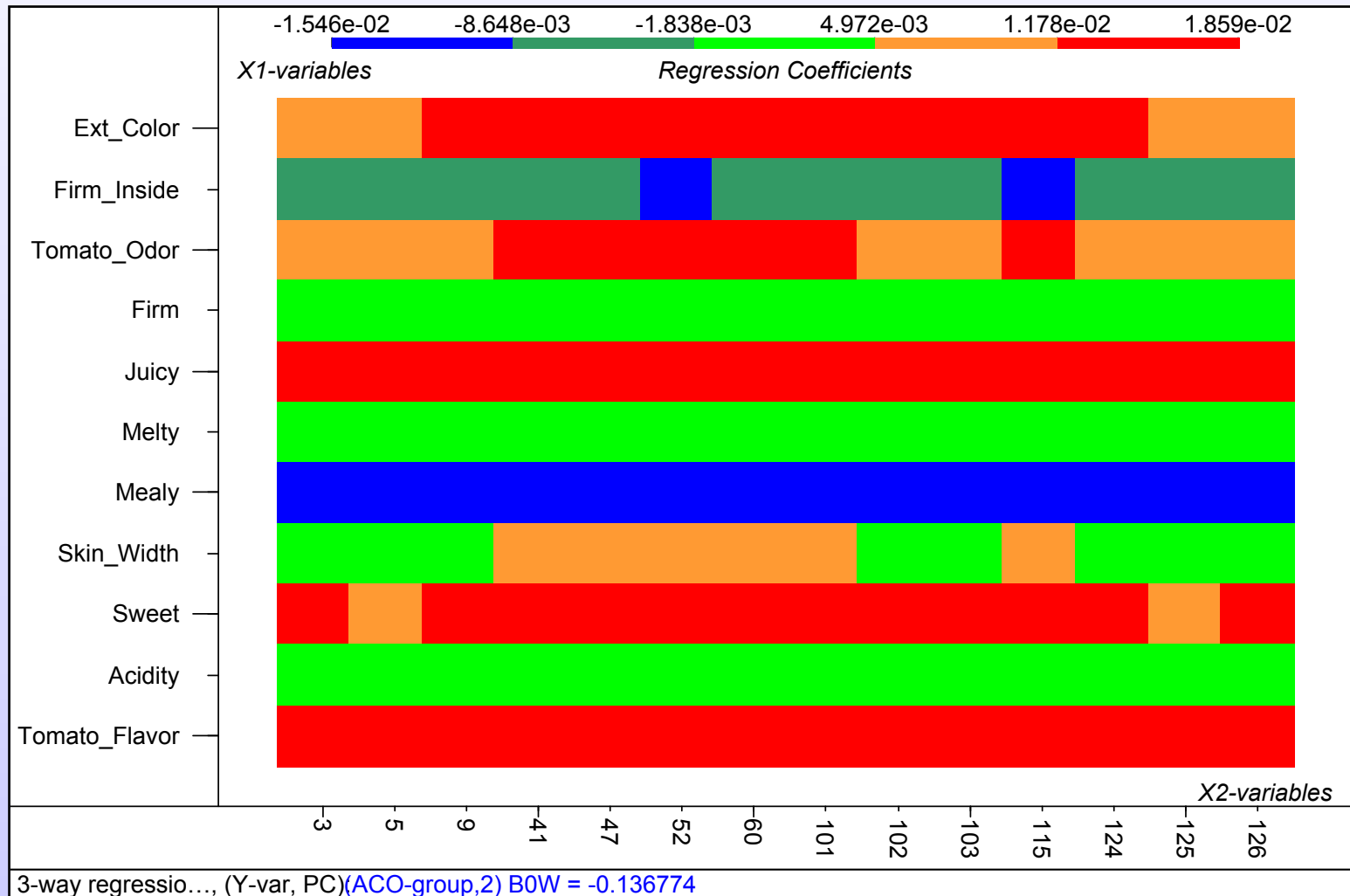


Panel agreement



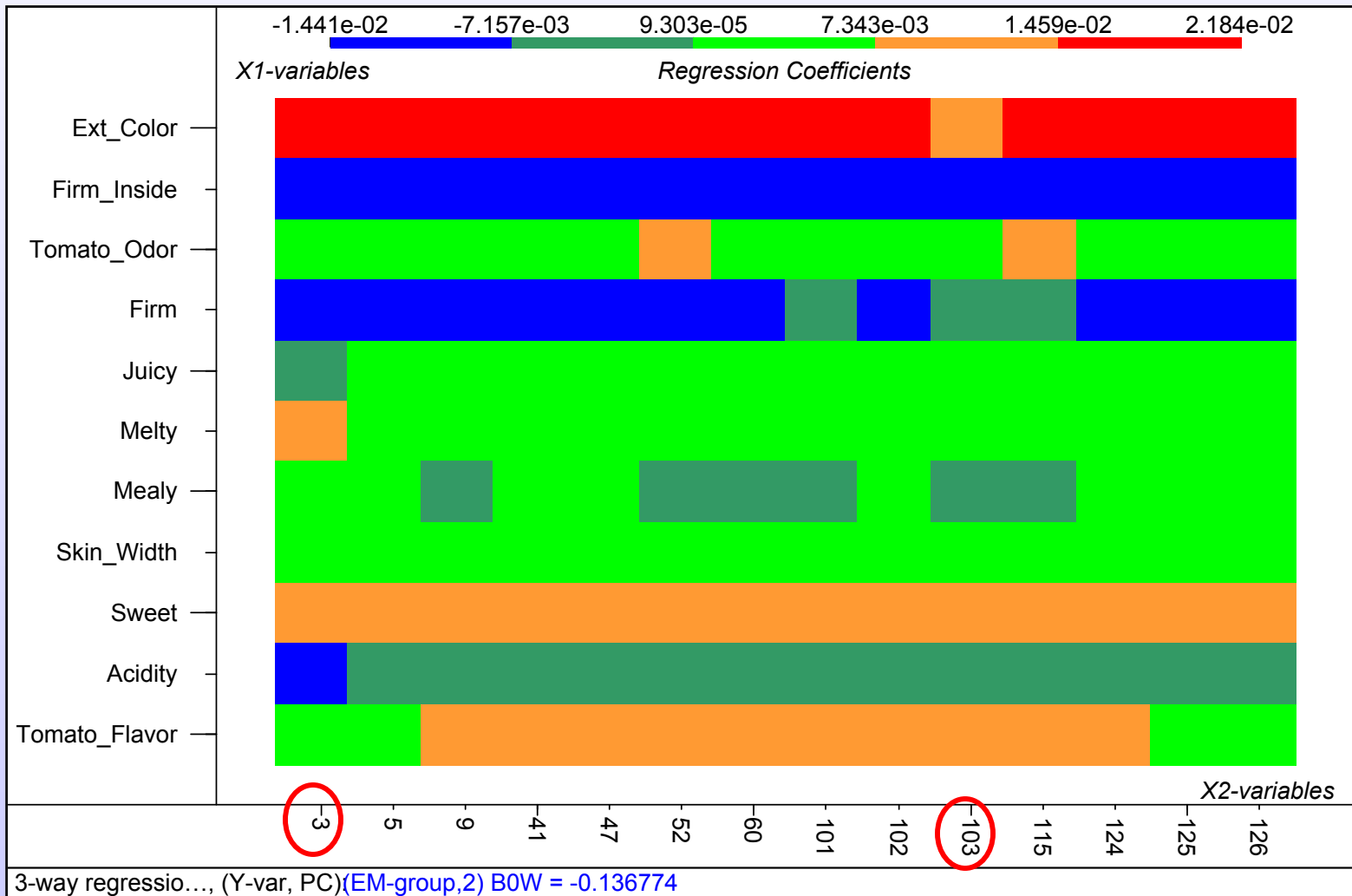


Drivers for ACO group



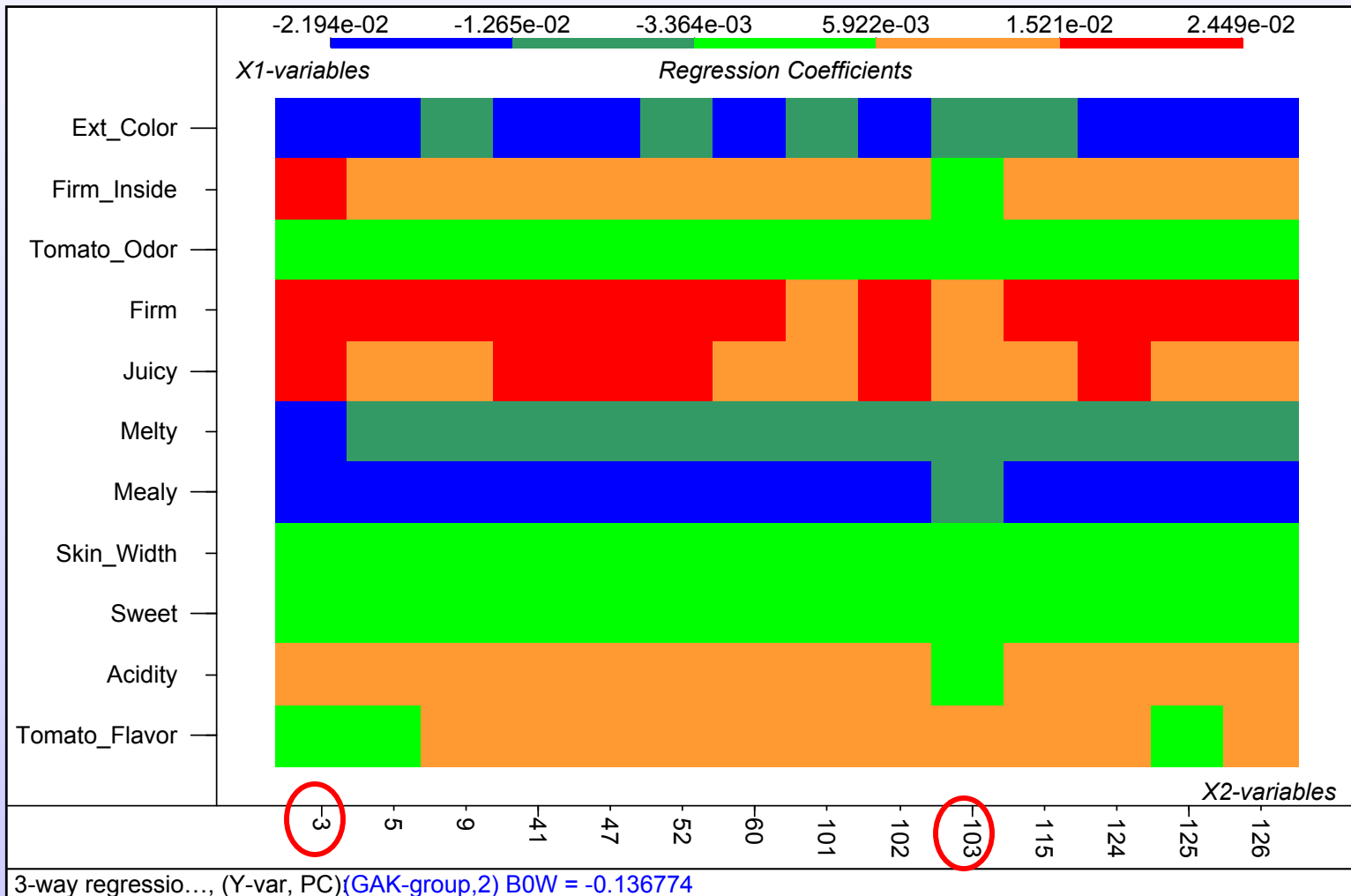


Drivers for EMA group





Drivers for GAK group





Drivers of liking

	ACO	EMA	GAK
Positive	Tomato flavor and odor, sweet, juicy, external color	Ext. color, tomato flavor and sweet	Firm, firm inside, juicy, acidity, tomato flavor
Negative	Mealy and firm inside	Mealy, acidity, firm and firm inside	Ext. color, melty and mealy

Skin width is not a driver of liking for these segments



3. Key demographics, attitudinal and usage characteristics



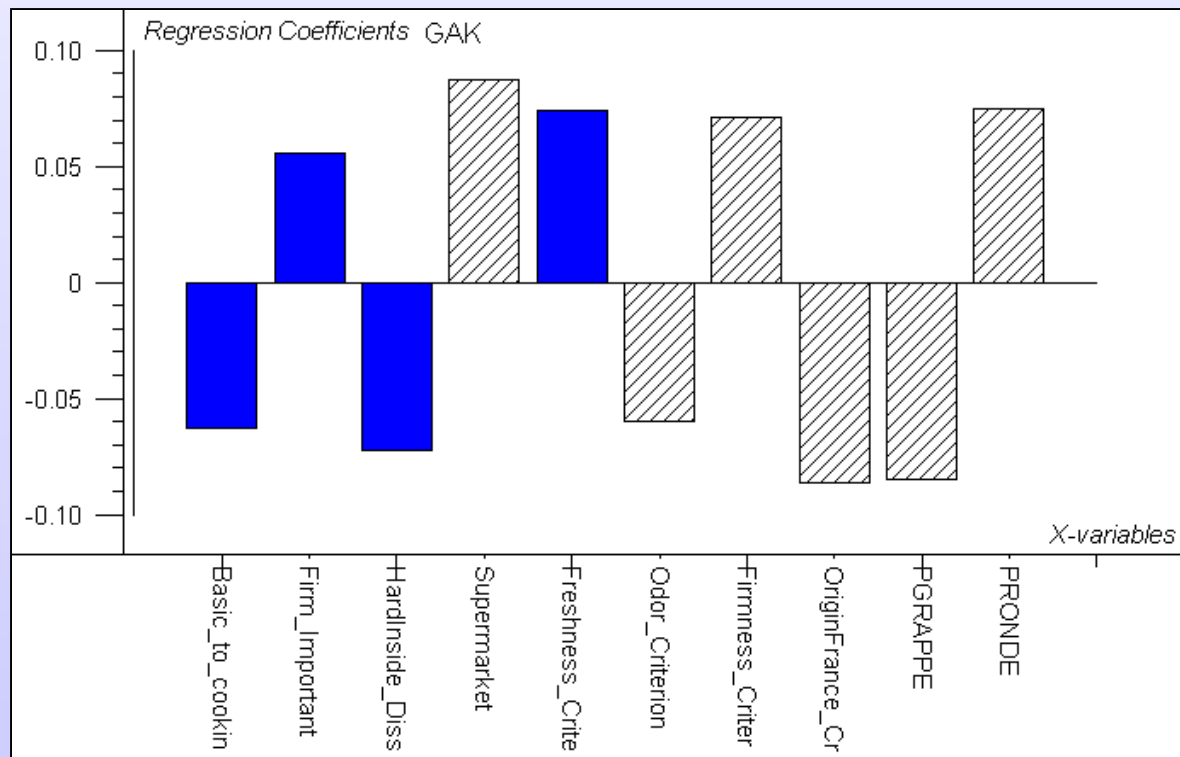
Key consumer characteristics per Segment

- The demographics & attitudinal questionnaire was recoded into interpretable numbers
- For each segment, a PLS1 model was built with the answers from the consumer questionnaire as X and the group belonging information as Y in a binary (1/0) variable
- A PLS1 approach was preferred to a PLS2 approach so as to focus the analysis on characterizing each segment, rather than discriminating each segment from the others.



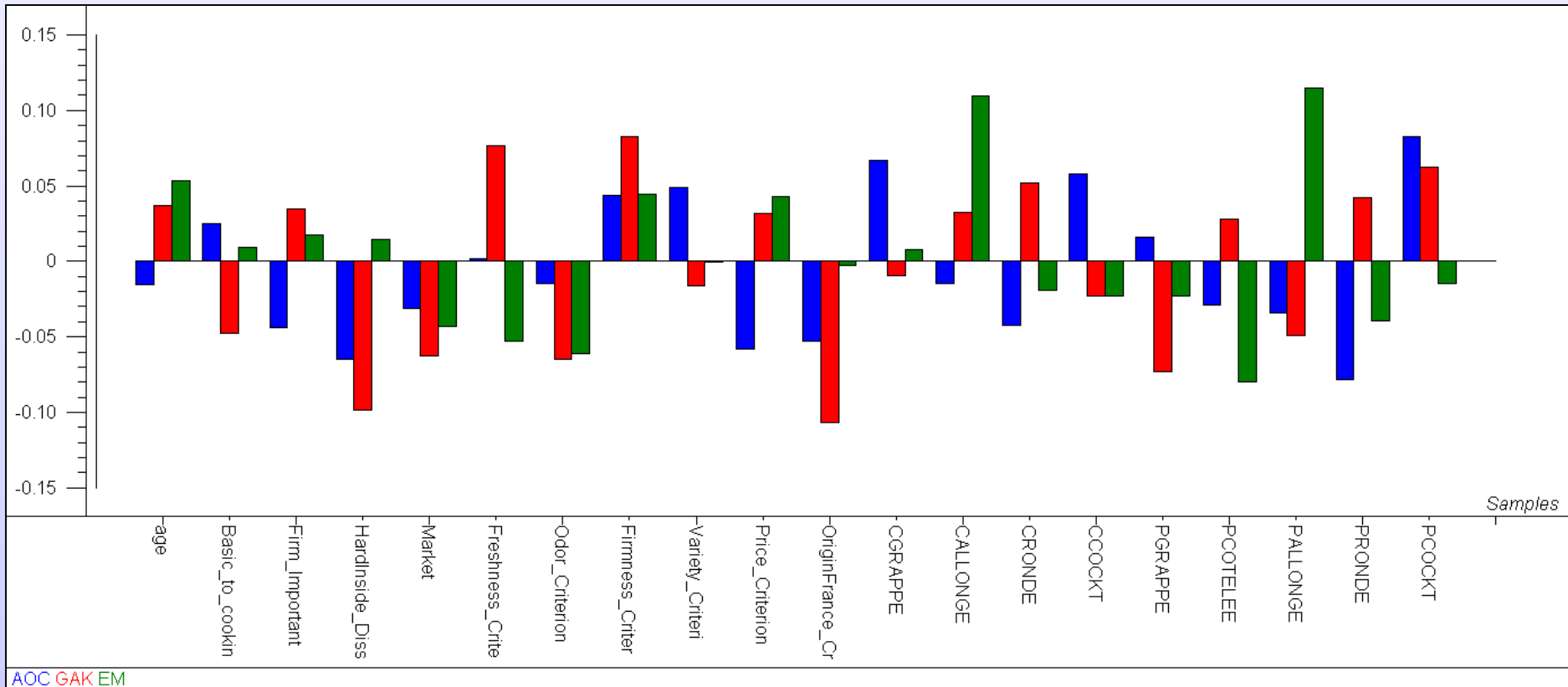
Key consumer characteristics per Segment

- Uncertainty Testing (Jack-knifing) in cross-validation was utilized to identify the significant parameters





Key consumer characteristics per Segment





Key consumer characteristics per Segment

	ACO	EMA	GAK
Key characteristics	<ul style="list-style-type: none">•Tend to be younger•Do not consider the price as a purchase criterion•Consume vine attached and cocktail type tomatoes•Prefer cocktail type tomatoes	<ul style="list-style-type: none">•Do not purchase tomatoes at the market•Do not consider the odor as a purchase criterion•Consume and prefer oval type tomatoes	<ul style="list-style-type: none">•Indicate firmness and freshness as important purchase criterions•Buy tomatoes in supermarkets•Do not require tomatoes to be originated from France



Main findings from the 2001 data set

	Segment 1: ACO 44%	Segment 2: EMA 29%	Segment 3: GAK 27%
Segment size (%)			
Tomato varieties liked most	A, C and O	E, M and A	G, A and K
Tomato varieties liked least	E, P and G	P, D and K	E, D and P
Positive drivers of liking	Tomato flavor and odor, sweet, juicy, external color	Ext. color, tomato flavor and sweet	Firm, firm inside, juicy, acidity, tomato flavor
Negative drivers of liking	Mealy and firm inside	Mealy, acidity, firm and firm inside	Ext. color, melty and mealy
Key demographics, attitudinal and usage characteristics	Do not consider the price as a purchase criterion; consume vine attached and cocktail type tomatoes; prefer cocktail type tomatoes	Do not purchase tomatoes at the market, do not consider the odor as a purchase criterion, consume and prefer oval type tomatoes	Indicate firmness and freshness as important purchase criterions, buy tomatoes in supermarkets, do not require tomatoes to be originated from France



Conclusions

- Exploratory Analysis (PCA) allowed to detect groups of tomato likings without prejudice
- With 3-way PLS regression, we linked the preference data to the sensory profiles and could identify positive and negative drivers of liking for each segment
- PLS-R with Uncertainty Testing pointed out key attitudinal characteristics for each segment
- More groups could be studied



Aknowledgements

- Dr. Martin Kermit for the group selection
- Dr. Dongsheng Bu for building the 3-way data array
- The collection of these data was a part of a collaborative research program between Centre Technique des Fruits et Légumes (CTIFL) and Institut National de la Recherche Agronomique (INRA) in France sponsored by the Ministère de la Recherche and organized by the Association des Centres Technique Agricoles (ACTA).



CAMO

Thank you for your attention

camo@camo.no

www.camo.com