YMCA
(Youth Metrics Cluster Analysis)

Introduction

Marketing Consultancy Firm in Slovakia

Client Industry - Entertainment, Product development, Advertising, Financial, etc.

Background

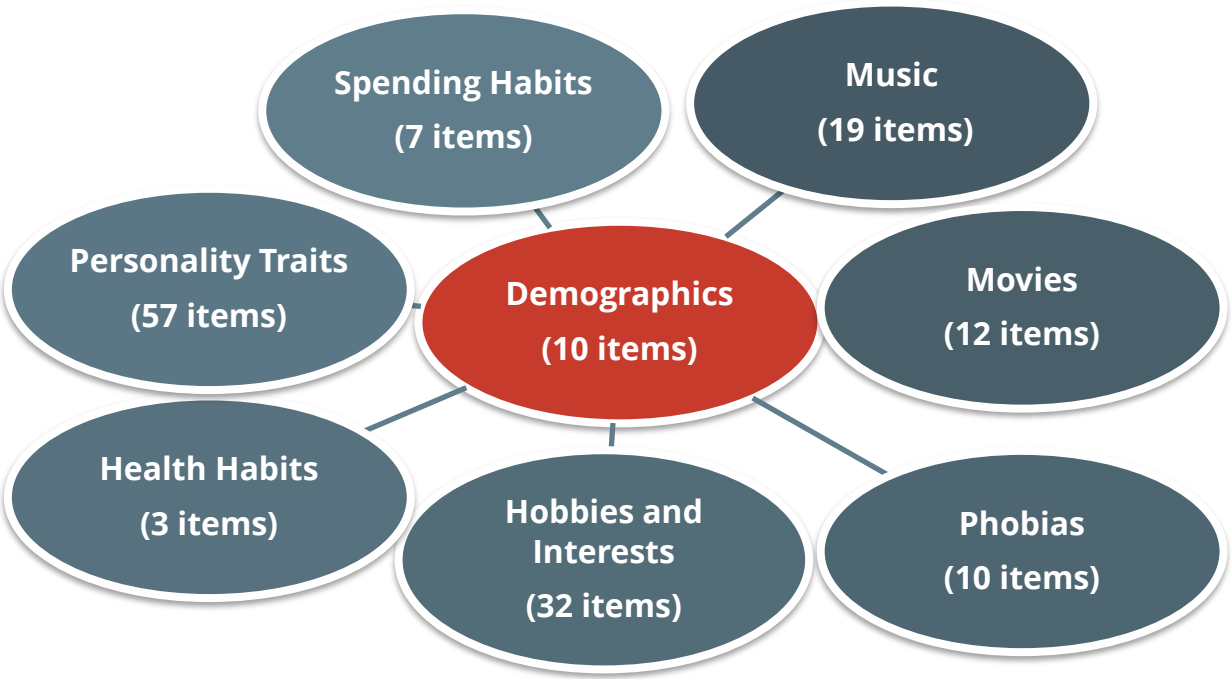
- We do – Segmentation for Targeting and Product Placement
- We use – Clustering, Factor Analysis and Principal Component Analysis
- From – Young People Survey Data, 2013

Problem Statement

"What is the best methodology to segment the customer base to create profiles for targeting and positioning for advertisements and marketing campaigns?"

Data : 'Young People Survey'

- 2013, Friends of Faculty of Social and Economic Sciences Department of the Comenius University in Bratislava, Slovakia
- Participants - Slovakian nationality, aged between 15 - 30
- 1010 rows and 150 columns (139 integer and 11 categorical)



MUSIC PREFERENCES

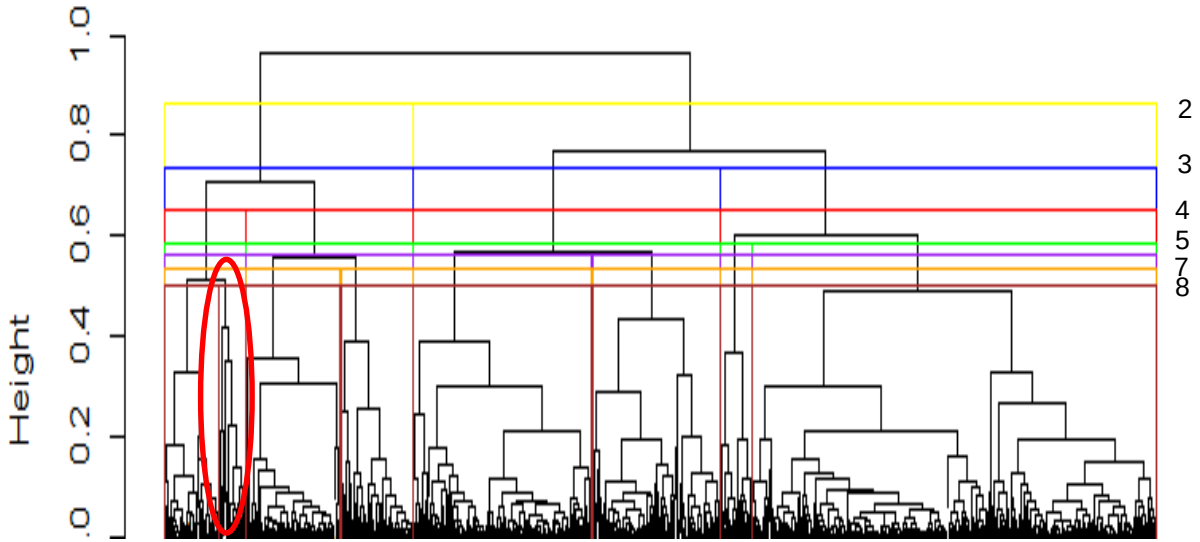
1. **I enjoy listening to music.**: Strongly disagree 1-2-3-4-5 Strongly agree (integer)
2. **I prefer.**: Slow paced music 1-2-3-4-5 Fast paced music (integer)

Demographic Clustering

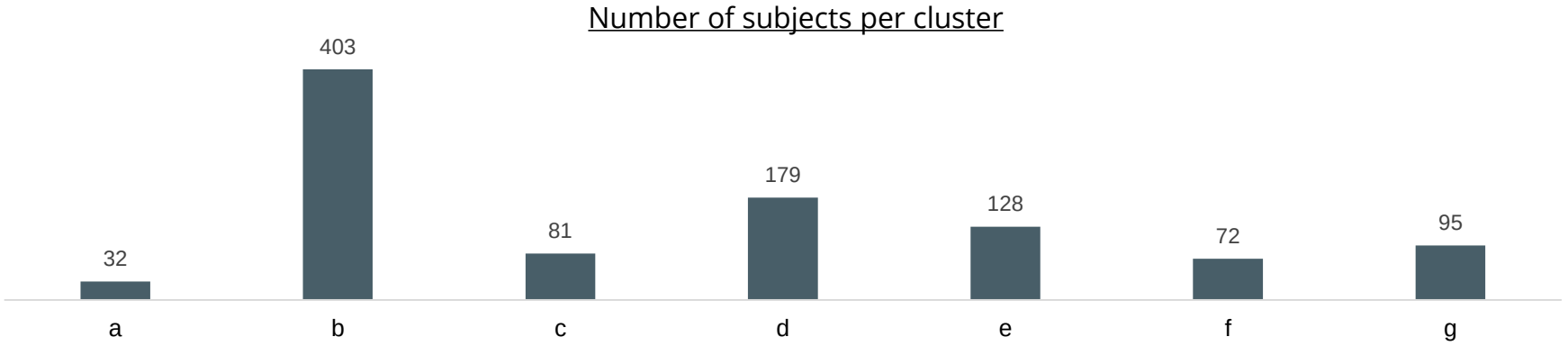
- ❑ Music Preferences
- ❑ Movie Preferences
- ❑ Hobbies & Interests
- ❑ Phobias
- ❑ Health Habits
- ❑ Personality Traits
- ❑ Spending habits
- ❑ Demographics

- Consumer needs, wants, and usage rates vary closely with Demographics
- Easier to measure than other variables.

- ❑ Hierarchical clustering: Possible to view partitions at different levels of granularities (i.e. can refine or coarsen clusters) using different k
- ❑ K- means can't handle Categorical data



The 7 clusters



Male / Female

City / Village

Only Child / Not

Low / High Income

Method : Principal Component Analysis (PCA)

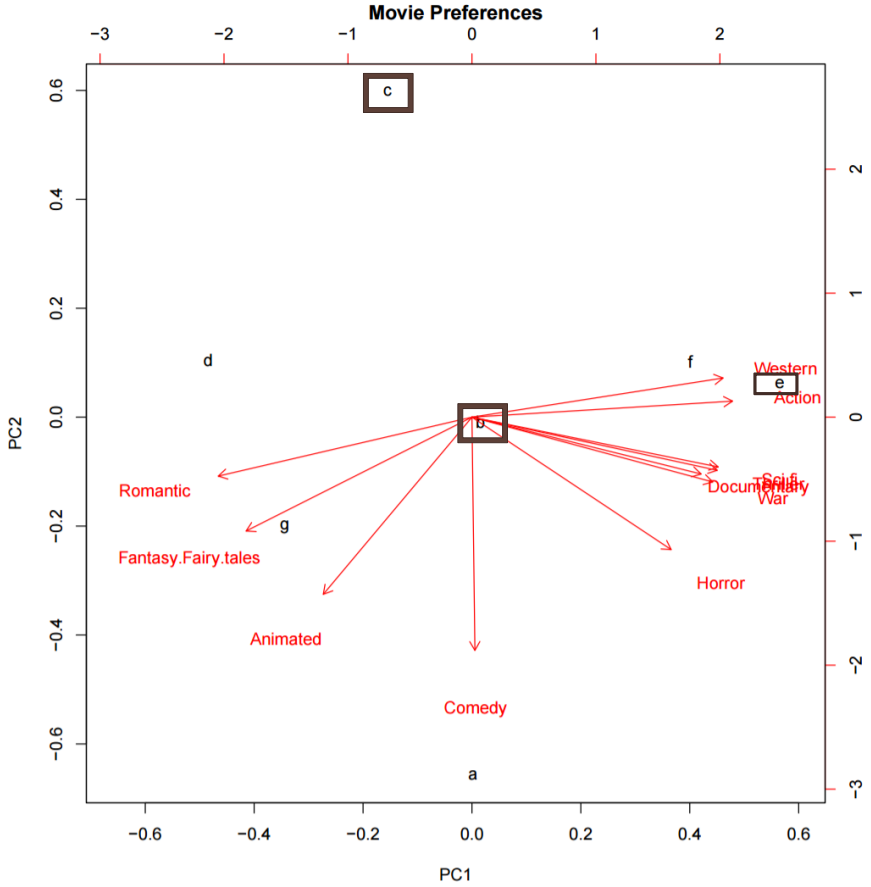
Combine the "preference" variables with the Cluster segmentation for each observation

```
seg.ratings = seg.raw
seg.ratings = seg.ratings[seg.ratings$Gender!="",]
seg.ratings = seg.ratings[seg.ratings$Left...right.handed!="",]
seg.ratings = seg.ratings[seg.ratings$Education!="",]
seg.ratings = seg.ratings[seg.ratings$Only.child!="",]
seg.ratings = seg.ratings[seg.ratings$Village...town!="",]
seg.ratings = seg.ratings[seg.ratings$House...block.of.flats!="",]
seg.ratings = seg.ratings[,1:(length(seg.raw)-9)]
seg.ratings = cbind(seg.ratings, seg.hc.segment)

seg.sc = seg.ratings
seg.sc[,21:31] = scale(seg.sc[,21:31])
seg.sc.music = seg.sc[,c(21:31,142)]
```

Perform PCA analysis and draw the Perceptual Map

```
seg.pc = prcomp(na.omit(seg.sc.music[,1:11]))

seg.pc$center

(seg.mean = aggregate(.~seg.hc.segment, data = seg.sc.music, mean))

seg.mean = seg.mean[,-1]
rownames(seg.mean) = paste("", letters, sep="")[1:7]
seg.mu.pc = prcomp(seg.mean, scale. = TRUE)

biplot(seg.mu.pc, main="Movie Preferences", cex=c(1,1))
```

Method : Principal Component Analysis (PCA)



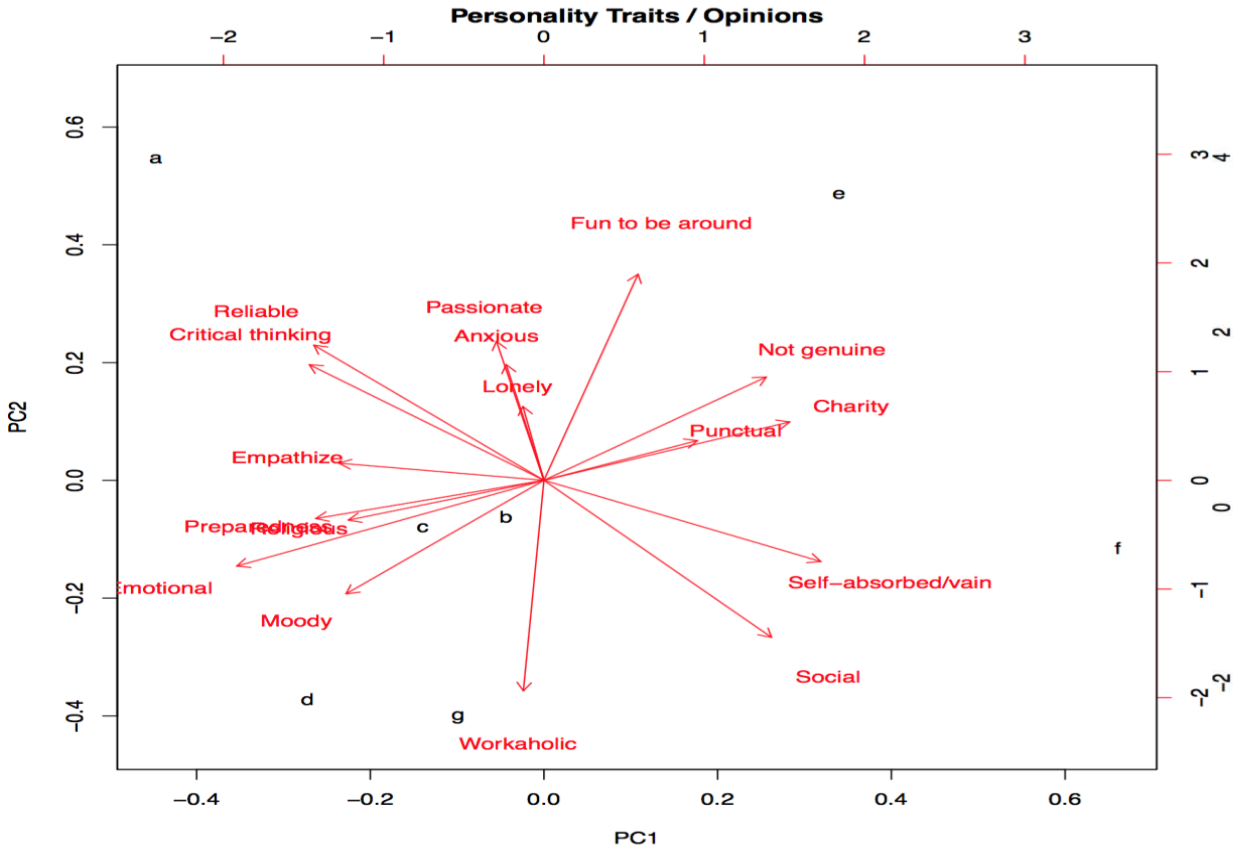
Mean per Cluster for Movie Preferences			
Movie	**B**	**C**	**E**
Action	-0.0080	-0.0378	0.5827
Animated	-0.0337	-0.0941	-0.0775
Comedy	0.1827	-0.1135	-0.0209
Documentary	-0.0303	-0.1674	0.3617
Fantasy/FairyTale	-0.0451	-0.0181	-0.2698
Horror	0.1096	-0.1414	0.1279
Romantic	-0.0545	0.0037	-0.5192
SciFi	-0.0323	-0.1128	0.4284
Thriller	0.0842	-0.2252	0.3397
War	-0.0605	-0.2413	0.5683
Western	-0.0225	0.0717	0.5983

- Placement is the aggregated scores per movie preference across PC1/PC2
- Cluster B average highest for Comedy, Horror, Thriller
- Cluster C doesn't really like anything
- Cluster E likes Action, Western, and War; dislikes Romantic and Fantasy/FairyTale

Method : Factor Analysis

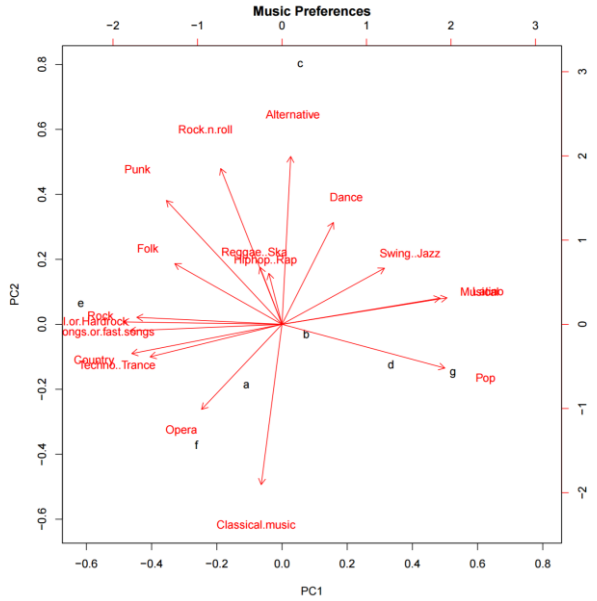
Personality Traits/Opinions:

- 57 features – too many
- Eigenvalues: 17 above 1.00
- Naming convention: values >0.25 included to describe the factor
- Incorporated these factors as the final variables in the analysis



Analysis : Segmentation

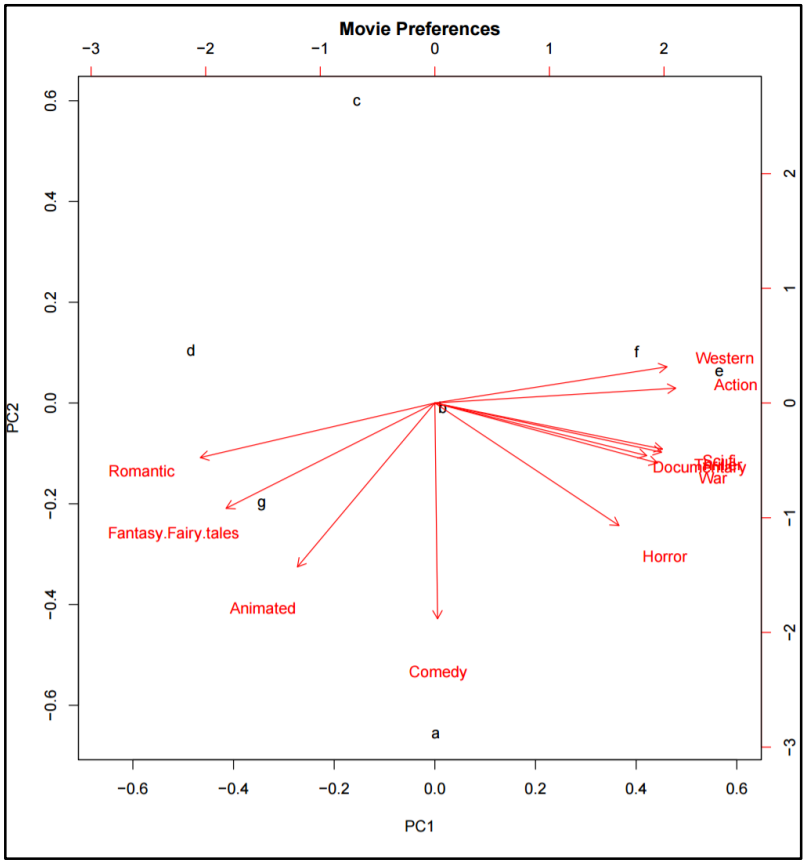
Results: 7 Perceptual Maps, 7 Customer Groups, Multiple Positioning Strategies

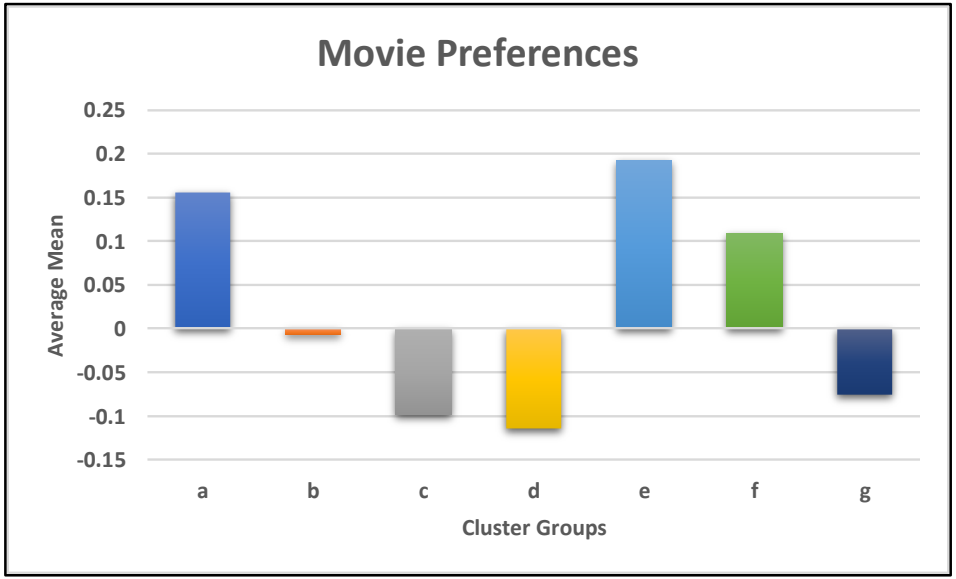


- Demographics: only child with high income
- Music: Alternative
- Movies: Doesn't like movies
- Hobbies: Medicine/Biology
- Phobia: Snakes
- Habits: Smoker
- Personal: Emotional/Moody
- Spending: Enjoys shopping

C

Analysis : Segmentation

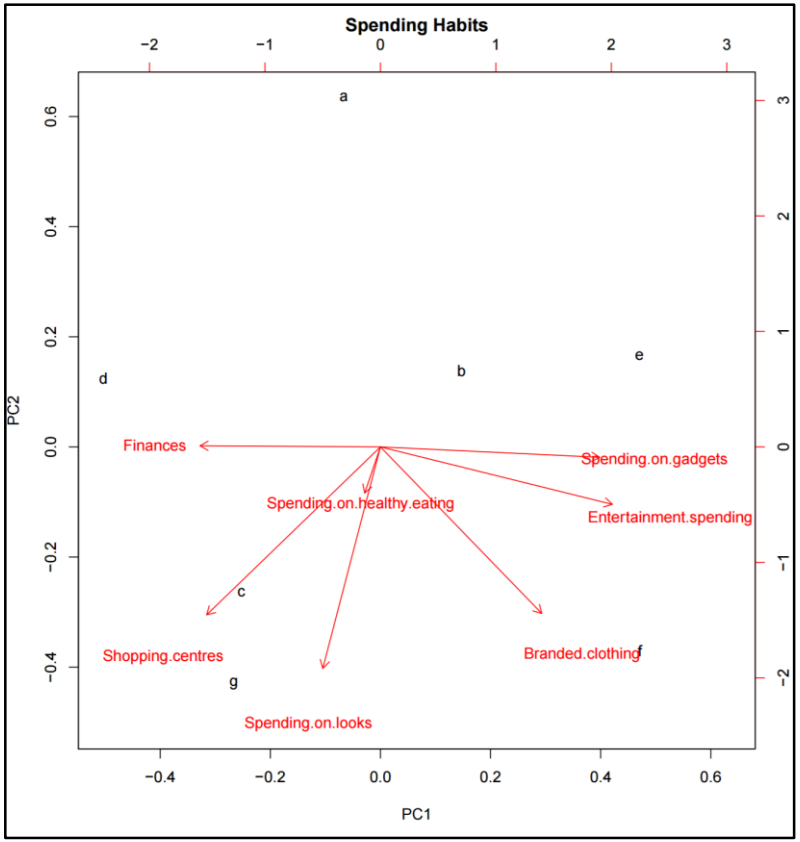


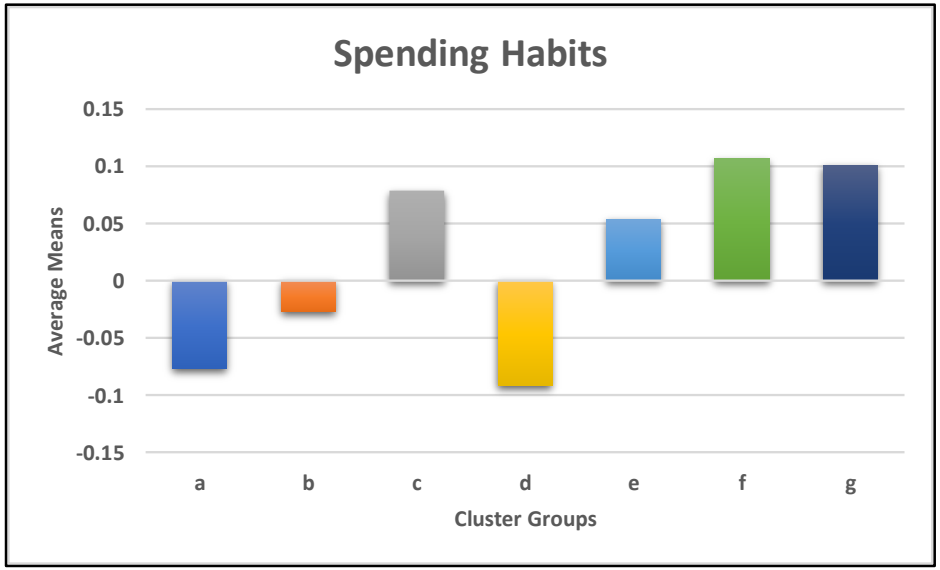


Interpretation:

- Perceptual Map shows cluster positioning
- Chart of the average means shows which Cluster group has a stronger movie preference overall

Analysis : Segmentation





Interpretation:

- Cluster f and g have a stronger affiliation with spending trends
- Cluster d has a stronger desire to save

Analysis : Positioning Strategy

- Demographics: male, with siblings, high income
- Music: Rock
- Movies: Action/Western
- Hobbies: Geography/X-treme Sports
- Phobia: None
- Habits: Alcohol
- Personal: Fun to be around
- Spending: Buys gadgets

e

Healthy grocery stores; latest designer bag trends

Latest Apple watch; Tickets to Moto-Racing

- Demographics: city girl, only child, high income
- Music: Pop
- Movies: FairyTales/Romantic
- Hobbies: Foreign Languages / Fun with Friends
- Phobia: Darkness/Flying
- Habits: Healthy eater
- Personal: Workaholic
- Spending: Likes to shop for looks/healthy food

g

Summary, Learnings and Implications

Summary & Learnings

❑ What is our objective?

✓ What is the best technique to create a profile structure, using statistical approaches, that will enable clients to devise actionable plans

❑ Why are we analyzing it?

✓ Market segmentation allows us to speak to customers in terms that will get their attention and in-turn they will be more likely to do business

❑ How have we done market segmentation?

✓ A combination of Cluster Analysis, Factor Analysis and PCA

❑ What was the most challenging part?

✓ Categorical Missing values (20 rows – Clustering, additional 139 rows – Factor Analysis)

7

Identifiable Segments

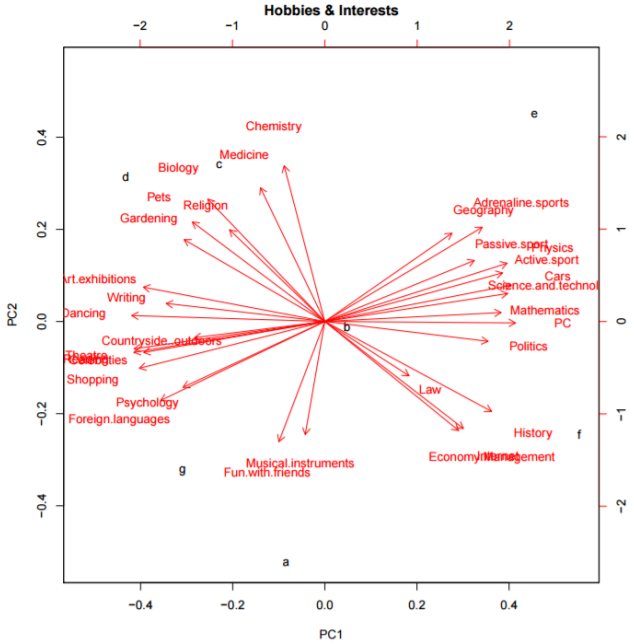
Summary, Learnings and Implications

Implications

Bernie: Marketing Director of XYZ Advertising Company | **Team B-05: Analytics Consultants**

Bernie: *We need a strategy on what to advertise to users on the Pandora Slovakia Rock channel. Can your team help?*

Team B-05: *We have an analysis report ready that will pinpoint what are the favorite movie genres, spending habits, personality traits and phobias of the youth users listening in on the Pandora Slovakia Rock channel based on their demographics.*



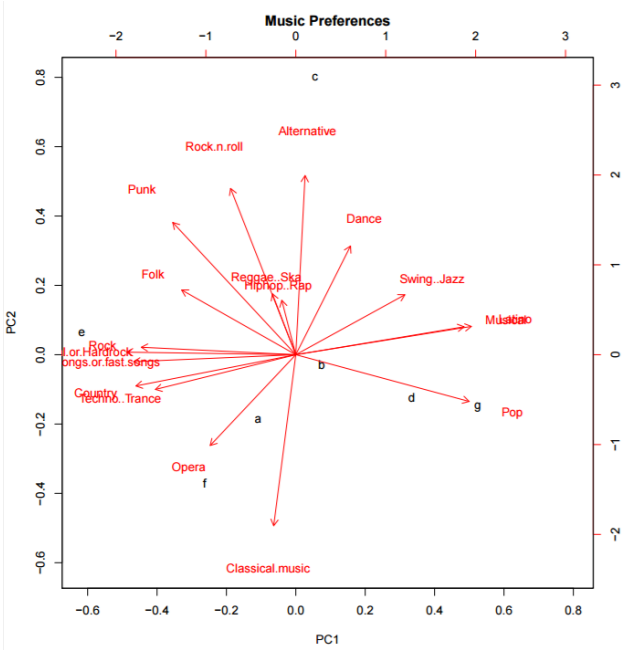
Summary, Learnings and Implications

Implications

Bernie: Marketing Director of XYZ Advertising Company | **Team B-05: Analytics Consultants**

Bernie: *But wait, I have to also devise an advertising strategy to target youth users who stream movies from Fandango Slovakia. Can your analysis help me do that?*

Team B-05: *Certainly, our analysis can help you identify music choices, personality traits, phobias and spending habits of the youth users based on their demographics and movie genre likes.*



Vďaka!