

Final Project

For the purpose of my business plan, I conducted a survey with 235 participants. The survey was designed to gain insight on coffee habits and purchasing decisions of coffee consumers. The aim was to discover patterns in attitude and preferences to create segments of consumers through clustering analysis for marketing purposes. The most promising segment will be targeted in the positioning and marketing strategy.

Moreover, I want to learn whether there is a variance in attitudes within demographic groups and what variables mostly drive the variance between consumers. As previously agreed, I will be doing a Clustering & Correspondence analysis. This will include several steps:

1. Preparing the data. Re-scaling & selecting the variables
2. Clustering analysis, excluding demographics
3. Correspondence analysis, bring in demographics
4. Individual cluster analysis - sub sampling for additional findings and profiling

Naturally this additionally entails, visualization, hypothesis testing & validation testing.

Data set

For the clustering I will most likely be using columns (from excel): B-D, F-U. Only excluding the demographic variables.

As you can see there are many different types of variables included in the data set. The nominal values will be taken as 0 = negative, 1 = positive (one-hot-encoding). And all others will be standardized 0-1.

Field	Description	Scales
AmountWeek	How many cups of coffee do you typically consume weekly?	Ratio, Continuous
AmountOutMonth	How frequently do you drink out-of-home per month on average?	Ratio, Continuous
MoneyCoffee	How much money on average do you estimate you spend on coffee per month?	Ratio, Continuous
MoneyGroceries	How much on average do you spend on general groceries per month?	Ratio, Continuous
Machine	How do you brew your coffee at home?	Nominal
Brand change	How often do you switch between coffee brands?	Nominal
Purchase location	Where do you usually purchase your coffee?	Nominal
Supermarket_Positive_Reasons	When you purchase coffee from the supermarket what are your main reasons for doing so?	Nominal
Supermarket_Negative_Reasons	What would be reasons why you would not purchase coffee from the supermarket?	Nominal
Criteria_Type_Coffee	What are your main criteria's or evaluation points for choosing the type of coffee?	Nominal
KnowledgeCoffee	How would you describe your knowledge level regarding coffee in general?	Ordinal. 0-10, Discrete
Purchase_Price	I believe that the ____ is important to my decision on which coffee to purchase.	Ordinal, likert 0-5
Purchase_Sustainability	I believe that the ____ is important to my decision on which coffee to purchase.	Ordinal, likert 0-5
Purchase_Sustainability	I believe that the ____ is important to my decision on which coffee to purchase.	Ordinal, likert 0-5

coffee to purchase.	Ordinal, likert 0-5	
Purchase_Packaging	I believe that the ____ is important to my decision on which coffee to purchase.	Ordinal, likert 0-5
Frequency_Specialty	How often do you drink specialty coffee?	Ordinal
Subscription_Likely	How likely are you to have an online subscription for (specialty) coffee?	Ordinal 0-10, Discrete
Subscription_Not_Likely	What is the number one reasons why you would be hesitant?	Nominal
App_Likely	How likely are you to value and use an app for your online subscription?	Ordinal, 0-10, Discrete
Gender	What is your gender?	Nominal
AgeCategory	What is your age category?	Ordinal
Occupation	What is your occupational status?	Nominal
Education	What level of education have you completed?	Ordinal
Home	How would you describe the place you currently live in?	Nominal

Below are the heads of the data set:

AmountWeek	AmountOutMonth	MoneyCoffee	MoneyGroceries	Machine	BrandChange
3	5	15	320	Filter machine	Sometimes
21	4	10	125	Espresso machine	Sometimes
22	8	30	350	CupMachine	Sometimes

PurchaseLocation	Supermarket_Positive_Reasons	Supermarket_Negative_Reasons	Criteria_Type_Coffee	KnowledgeCoffee
The supermarket	Time-saving	Not wanting to support big cooperations	Price	4
Specialty stores or cafés	I do not purchase coffee from the supermarket	It is not fresh, Better quality elsewhere	Origin, Flavour profile	7
The supermarket	Price, Time-saving	No reason	Roast level, Flavour profile	5

Purchase_Price	Purchase_Sustainability	Purchase_Certificate	Purchase_Fairtrade	Purchase_Packaging
2	5	1	5	3
2	4	1	5	3
3	3	3	3	3

Frequency_Specialty	Subscription_Likely	Subscription_Not_Likely	App_Likely
Only in cafes	3	The price	1
Always	10	No reason	9
Never	3	I am happy with my coffee now, I do not like being stuck with subscriptions	2

App_Likely	Gender	AgeCategory	Occupation	Education	Home
1	Male	18-25	Student	Bachelor's degree	Urban (City)
9	Female	18-25	Student	Bachelor's degree	Urban (City)
2	Male	18-25	Student	Bachelor's degree	Suburbs

I am also attaching my github for further reference with the project file called: Thesis Cluster Survey
<https://github.com/danielle-kotter/Master-Thesis>

Prepared by: Danielle Kotter