Humboldt University Berlin

Institute of Marketing

Dr. Narine Yegoryan & Mareike Sachse

Customer Analytics and Customer Insights

WS 2022/23

Assignment 3

Perceptual and Preference Mapping

Data for the assignment

QuestionaireData_CityTrips.csv

Page Limit

The assignment should not be longer than 5 pages. The suggested length for single tasks is just that – a suggestion.

What to submit

The report in PDF format and an R script (.R file) that reproduces the analysis. Include meaningful comments, so we understand which part of the code is related to which task.

Background

To investigate young adults' perception and evaluation of various European cities as weekend trip destinations, a survey was designed and distributed among students at Humboldt University Berlin (convenience sample). In total, 266 repondents completed the questionnaire.

A list of 20 different European cities was selected. To keep the questionnaire manageable for the respondents (possible to answer within 10-15 minutes), each respondent saw only 6 out of 20 cities. Ten subsets of 6 cities were created a priori. Each city was included in 3 different subsets (see Table 1 for details). Respondents were then randomly assigned to one of these subsets. The overall structure of the questionnaire is summarized in Figure 1. An example of the questionnaire for Subset 1 of the cities in the file SAMPLE Questionaire.pdf.

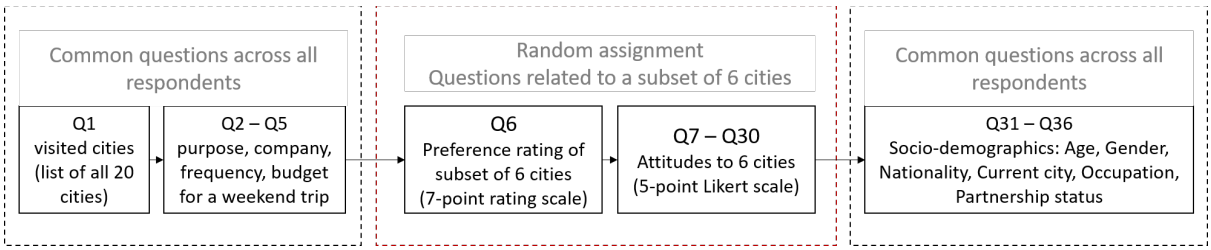

Figure 1: Questionnaire structure

	Subset										
City	1	2	3	4	5	6	7	8	9	10	Sum
Amsterdam	0	1	1	0	0	0	0	1	0	0	3
Athens	0	1	1	0	0	1	0	0	0	0	3
Barcelona	0	0	0	0	0	1	1	0	0	1	3
Berlin	0	0	0	0	0	1	1	1	0	0	3
Brussels	1	0	0	0	1	1	0	0	0	0	3
Budapest	0	0	0	1	0	0	0	0	1	1	3
Dublin	0	0	0	1	0	0	1	0	0	1	3
Geneva	1	1	1	0	0	0	0	0	0	0	3
Istanbul	0	0	0	0	1	0	0	1	0	1	3
Krakow	0	0	0	0	0	1	1	0	1	0	3
Lisbon	0	0	0	1	0	0	1	0	1	0	3
London	1	1	1	0	0	0	0	0	0	0	3
Madrid	0	0	0	0	0	1	0	1	1	0	3
Paris	1	0	1	1	0	0	0	0	0	0	3
Prague	1	0	0	1	1	0	0	0	0	0	3
Riga	0	1	0	1	1	0	0	0	0	0	3
Rome	0	0	0	0	1	0	1	0	1	0	3
St. Petersburg	0	0	0	0	0	0	0	1	1	1	3
Stockholm	1	1	1	0	0	0	0	0	0	0	3
Vienna	0	0	0	0	1	0	0	1	0	1	3
Sum	6	6	6	6	6	6	6	6	6	6	

Table 1: Constructed ten subsets of cities

Data Description and Codebook

Column #	Variable name	Description
1	Sample	Index for the subset of cities
2	ID	Respondent index within a subset of cities (is not a unique index across the whole sample)
3	Timestamp	A character variable indicating the date and time of the respondent accessing the questionnaire
4-23	[city] (e.g., Berlin)	A dummy variable indicating whether the respondent has visited the respective city (= 1) or not (= 0) (More than one option could have been chosen for this question, i.e., could have visited multiple cities)
24-28	Purpose[i] (i = 1, …, 5)	A dummy variable indicating whether the respondent usually goes on weekend trips for a specific purpose: 1. Visit family 2. Visit friends 3. Partying 4. Exploring a new city 5. Visiting various cultural event (More than one option could have been chosen for this question)
29	Purpose_other	A character variable indicating the respondent specified other purposes of a weekend trip. An empty cell (which will show as “” in R) indicates no other purpose was specified.

Column #	Variable name	Description
30-34	With_Whom_[i], (i = 1, …, 5)	A dummy variable indicating with whom the respondent usually goes on a weekend trip 1. with family 2. with friends 3. with a partner 4. with colleagues 5. by yourself (More than one option could have been chosen for this question)
35	Number_of_Trips	A numerically coded variable indicating the number of trips in the respective year of data collection (2015). Values 1 to 5 represent the following categories given in the question: 1. none 2. once 3. 2-3 times 4. 4-5 times 5. more than 5 times
36	Avg_Budget	A numerically coded variable indicating the average budget for a weekend trip. Values 1 to 5 represent the following categories given in the question: 1. less than 200€ 2. 200-300€ 3. 300-400€ 4. 400-500€ 5. more than 500€
37-56	Pref_[city]	Preference rating of a city as a weekend trip destination on a 7-point scale (1 = “Not preferred at all” and 7 = “Greatly preferred”). If the city was not in the subset of cities the respondent saw, it would be recorded as an empty cell (or NA in R). However, missing values are also denoted with an empty cell (or NA in R).

Column #	Variable name	Description
57- 456	[city]_Att[i] (i = 1, …, 20)	Attribute ratings of a city on a 5-point Likert scale (1 = "Strongly Disagree" and 5 = "Strongly Agree"). If the city was not in the subset of cities the respondent saw, it would be recorded as an empty cell (or NA in R). However, missing values are also denoted with an empty cell (or NA in R). Att[i], where i = 1, …, 20 represent the following items: 1. is friendly 2. is historical 3. is affordable 4. is trendy 5. has a vibrant nightlife 6. has delicious food 7. is easy to get around in (by foot or public transport) 8. is a good place for shopping 9. full of cultural events (incl. festivals, theater, etc.) 10. has interesting museums 11. is clean 12. is green (e.g., has lots of parks) 13. is international / multicultural 14. is too touristic 15. is fun 16. is noisy 17. is romantic 18. is safe 19. is beautiful 20. is English-speaker friendly
457	Age	An integer variable indicating the age of the respondent
458	Gender	A character variable indicating the gender of the respondent
459	Nationality	A character variable indicating the nationality of the respondent
460	CurrentCity	A character variable indicating the city the respondent currently lives in
461	Occupation	A character variable indicating the current occupation of the respondent
462	PartnershipStatus	A character variable indicating the partnership status of the respondent

Tasks

1 **Investigate the Data (1.5-2 pages)**

1.1 Briefly describe the sample of respondents based on their characteristics (socio-demographics), overall travel patterns, and city preferences. Use both univariate and bivariate statistics, tables, and graphs. Report the most interesting insights from your analysis.

1.2 Check for missing values (survey or item non-response) and outliers. Hint: not all NAs in the data are actual missing values: some NAs are due to respondents seeing different subsets of cities. If any, document and explain your approach to dealing with missing values and outliers and report the resulting sample size.

2 **Perceptual and Preference Mapping (1.5-2 pages)**

Your task is to produce a joint map of perceptions and preferences for European cities as weekend trip destinations using attribute evaluations and preference ratings.

2.1 Describe and explain your approach to producing a perceptual map. For example,

- Do you opt for factor analysis or multidimensional scaling? Check for the appropriateness of the data for FA or MDS.
- What method for MDS (dissimilarity measure, transformation) or FA (factor extraction, rotation method) do you apply?
- How many dimensions do you retain?
- How do you map attributes (if applicable)?

Do not simply report but argue and justify your decisions!

2.2 Augment your perceptual map by mapping consumer preferences. Describe and explain your approach.

- Vector vs. ideal point model?
- Individual or segment-level? In the latter case, what are segmentation criteria and why?

Do not simply report but argue and justify your decisions!

2.3 Interpret the resulting map

3 **Plan of Attack: Suggestions to the Management (1-1.5 pages)**

From the point of view of an airline providing flights to various European cities, what insight can you derive?

3.1 What positioning strategy should the airline pursue, and which destinations should it focus on?

3.2 Which customers or customer segments would you suggest the airline to target and how?