# Factor Analysis of Cereal Brand

The data file labelled [Cereal](https://olympus.greatlearning.in/courses/6116/files/423880/download?verifier=0qdOR4RYLeG742psAyajQJhhC84WOkJHqP8I9Twq&wrap=1) has the following information.

As part of a study of consumer consideration of ready-to-eat cereals sponsored by Kellogg Australia, Roberts and Lattin (1991) surveyed consumers regarding their perceptions of their favourite brands of cereals. Each respondent was asked to evaluate three preferred brands on each of 25 different attributes. Respondents used a five-point Likert scale to indicate the extent to which each brand possessed the given attribute.

For the purpose of this assignment, a subset of the data collected by Roberts and Lattin, reflecting the evaluations of the 12 most frequently cited cereal brands in the sample (in the original study, a total of 40 different brands were evaluated by 121 respondents, but the majority of brands were rated by only a small number of consumers). The 25 attributes and 12 brands are listed below

|  |  |  |
| --- | --- | --- |
| Cereal Brand | Attributes 1-12 | Attributes 13-25 |
| All Bran | Filling | Family |
| Cerola Muesli | Natural | Calories |
| Just Right | Fibre | Plain |
| Kellogg’s corn flakes | Sweet | Crisp |
| Komplete | Easy | Regular |
| Nutrigrain | Salt | Sugar |
| Purina Muesli | Satisfying | Fruit |
| Rice Bubbles | Energy | Process |
| Special K | Fun | Quality |
| Sustain | Kids | Treat |
| Vitabrit | Soggy | Boring |
| Weetbix | Economical | Nutritious |
|  | Health |  |

In total 116 respondents provided 235 observations of the 12 selected brands. How do you characterize the consideration behaviour of the 12 selected brands? Analyze and interpret your results using factor analysis.

The final deliverable should contain:

1. Exploratory Data Analysis report of the dataset
2. Factor Analysis and interpretation of the analysis which states the final set of variables along with their importance

Solution 1:

|  |
| --- |
| This criterion is linked to a Learning OutcomeProblem 1 - EDA a) Basic data summary, Univariate, Bivariate analysis, graphs  **Exploratory data Analysis(EDA)**   * **Basic Summary** : In R function Summary(cereal)            * Rating for all obs. Min 1 and Max 6        * Plot Zoom   (Easy has Median at 5)  Mean  Median  Range  Interquartile Range  Kurtosis  Histogram  Frequency Distribution table  Box Plot   * **Univariate Analysis** : In R on single variable |
| This criterion is linked to a Learning OutcomeProblem 1 -PCA/FA a) Perform PCA/FA and Interpret the Eigen Values (apply Kaiser Normalization Rule)   * Factor Analysis: * 1. Determining significant correlation b/w the var Bartlett test of sphericity     As p-value is 0 : there is significant correlation b/w 25 variables   * KMO test:     As MSA is 0.85 sample size is sufficient enough.   * Scree plot       No. of Factors : 4 (eigen values around 1 and above 1)   * Factor analysis using principal axis factoring method.       * Communality check       Easy and Process showing least communality     * Eigen Values      * % impact of these 4 factors      * Iteration 2      * Iteration 3      * Iteration 4      * Iteration 5      * Iteration 6      * Iteration 7     So, we stop at this point as cortest Bartlett test of sphericity confirms that there is no enough correlation in the data.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Iteration** | **Cortest Bartlett test p value** | **KMO Test - MSA Value Ideally above .70** | **Scree test - Factors  Eigen value above or around 1** | **Communality - Least impacting variables** | **Eigen Values %** | | 1 | 0 | 0.84 | 4 | Process and Easy | 57.507 | | 2 | 0 | 0.86 | 4 | Soggy and Economical | 61.21 | | 3 | 0 | 0.86 | 3 | Plain and Boring | 57.44 | | 4 | 0 | 0.86 | 3 | Crish and Fruit | 61.16 | | 5 | 0 | 0.87 | 3 | Fun and Salt | 64.64 | | 6 | 0 | 0.87 | 3 | Treat and Regular | 67.34 | | 7 | 1.54E-306 | Iteration Stopped |  |  |  |   **Conclusion: Out of the 25 variable, 17 variables are not significant factors.  With the 11 variable we can effectively determining a cereal is 67.34%** |
| This criterion is linked to a Learning OutcomeProblem 1 - PCA/FA b) Output Interpretation Tell which all factors needs to be shortlisted along with their importance and which ones needs to ignored. Name the factors with correct explanations.  Output Interpretation: Significant variable in Cereal data after conducting factor analysis in each iteration:   |  |  |  |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **Iteration** | 1 | **Iteration** | 2 | **Iteration** | 3 | **Iteration** | 4 | **Iteration** | 5 | **Iteration** | 6 | | **Factors** | 25 | **Factors** | 23 | **Factors** | 21 | **Factors** | 19 | **Factors** | 17 | **Factors** | 15 | | **Removed** | 0 | **Removed** | 2 | **Removed** | 4 | **Removed** | 6 | **Removed** | 8 | **Removed** | 10 | | Filling | 2 | Filling | 2 | Filling | 2 | Filling | 2 | Filling | 2 | Filling | 2 | | Natural | 3 | Natural | 3 | Natural | 3 | Natural | 3 | Natural | 3 | Natural | 3 | | Fibre | 4 | Fibre | 4 | Fibre | 4 | Fibre | 4 | Fibre | 4 | Fibre | 4 | | Sweet | 5 | Sweet | 5 | Sweet | 5 | Sweet | 5 | Sweet | 5 | Sweet | 5 | | Easy | 6 | Easy | 6 | Easy | 6 | Easy | 6 | Easy | 6 | Easy | 6 | | Salt | 7 | Salt | 7 | Salt | 7 | Salt | 7 | Salt | 7 | Salt | 7 | | Satisfying | 8 | Satisfying | 8 | Satisfying | 8 | Satisfying | 8 | Satisfying | 8 | Satisfying | 8 | | Energy | 9 | Energy | 9 | Energy | 9 | Energy | 9 | Energy | 9 | Energy | 9 | | Fun | 10 | Fun | 10 | Fun | 10 | Fun | 10 | Fun | 10 | Fun | 10 | | Kids | 11 | Kids | 11 | Kids | 11 | Kids | 11 | Kids | 11 | Kids | 11 | | Soggy | 12 | Soggy | 12 | Soggy | 12 | Soggy | 12 | Soggy | 12 | Soggy | 12 | | Economical | 13 | Economical | 13 | Economical | 13 | Economical | 13 | Economical | 13 | Economical | 13 | | Health | 14 | Health | 14 | Health | 14 | Health | 14 | Health | 14 | Health | 14 | | Family | 15 | Family | 15 | Family | 15 | Family | 15 | Family | 15 | Family | 15 | | Calories | 16 | Calories | 16 | Calories | 16 | Calories | 16 | Calories | 16 | Calories | 16 | | Plain | 17 | Plain | 17 | Plain | 17 | Plain | 17 | Plain | 17 | Plain | 17 | | Crisp | 18 | Crisp | 18 | Crisp | 18 | Crisp | 18 | Crisp | 18 | Crisp | 18 | | Regular | 19 | Regular | 19 | Regular | 19 | Regular | 19 | Regular | 19 | Regular | 19 | | Sugar | 20 | Sugar | 20 | Sugar | 20 | Sugar | 20 | Sugar | 20 | Sugar | 20 | | Fruit | 21 | Fruit | 21 | Fruit | 21 | Fruit | 21 | Fruit | 21 | Fruit | 21 | | Process | 22 | Process | 22 | Process | 22 | Process | 22 | Process | 22 | Process | 22 | | Quality | 23 | Quality | 23 | Quality | 23 | Quality | 23 | Quality | 23 | Quality | 23 | | Treat | 24 | Treat | 24 | Treat | 24 | Treat | 24 | Treat | 24 | Treat | 24 | | Boring | 25 | Boring | 25 | Boring | 25 | Boring | 25 | Boring | 25 | Boring | 25 | | Nutritious | 26 | Nutritious | 26 | Nutritious | 26 | Nutritious | 26 | Nutritious | 26 | Nutritious | 26 |  |  |  | | --- | --- | | **Iteration** | 7 | | **Cortest test** | Failed | | **Conclusion** | **Not enough corealtion among data** | | **Final Factors** | 17 | | **Total Removed** | 10 | |

s