Marketing Analytics Toolkit

Module Code: BEMM463J

Module Name: Marketing Analytics

Candidate Number: 211433

Word Count: 1471 words

Cluster Analysis is a multivariate method of identifying and grouping customers from a dataset based on homogeneity among customers who share similar traits/characteristics by using statistical models. "The goal of cluster analysis in marketing is to accurately segment customers in order to achieve more effective customer marketing via personalization." These homogeneous segments are known as 'personas'. **Key Characteristics: 1.** Practicality: Most accurate method to divide customers into segments. 2. Homogeneity: Helps to achieve high degree for homogeneity 3. Dynamicity: Provides dynamic results for dynamic dataset.

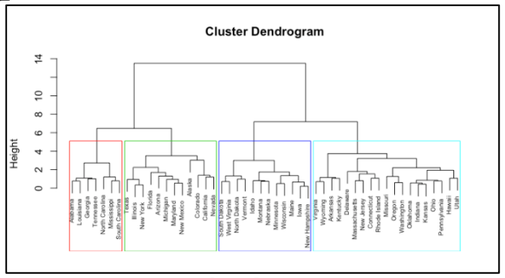

Figure no.(1) *Cluster Dendrogram*

How it works depends on the methodology used for analysis:
1.Hierarchical Method: This method starts by treating each datapoint as a cluster and then identifies closest two clusters, later combines these homogeneous clusters, the process is repeated until all clusters merged under one cluster.
Figure no. (1) shows the example of dendrogram based on similarities between various cities in the US.
2. Non-hierarchical method: This method basically creates new clusters by either merging or splitting the clusters. The data is combined to maximize or minimize based on decided evaluation criteria. **K**-mean cluster analysis is widely used mathematical algorithm, which is follows non-hierarchical approach for creating clusters. Steps of **k-mean cluster analysis**:

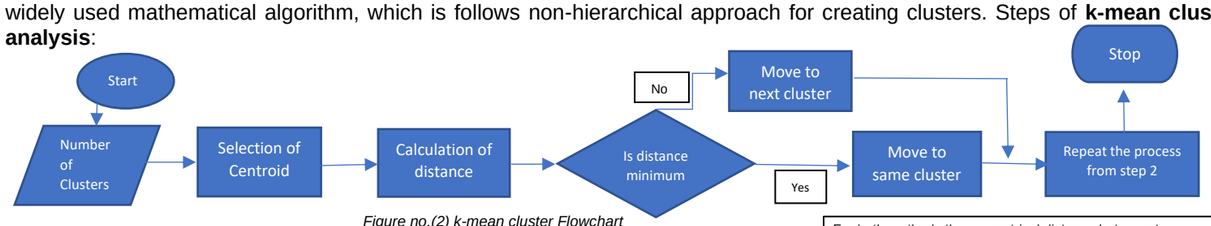

Figure no.(2) k-mean cluster Flowchart

Usability: Cluster analysis can be primarily used for market segmentation process. Segmentation helps in identifying groups of people/users, markets, communities who share common characteristics, this can further help to make decisions for target audience of a new product launch or development of existing product and its features that specific clusters finds valuable and relevant.

For both methods the geometrical distance between two datapoints or clusters is considered as **Euclidean distance**, to calculate **similarity** between these points it has following formula:

$$d(x,y) = \sqrt{\sum_{k=1}^{n}(x_k - y_k)^2}$$

Figure no.(3) Euclidean Distance Formula

Choice model: Choices/preferences of customers keep changing and evolving even if they fall under same homogeneous segment. In order to manage such customer dynamics, the **Choice Model** helps to predict the likelihood of customers' product preferences over its alternatives. It helps marketers to determine which marketing tactics(product pricing, discounts, offers) will attract most customers and/or impact their decisions.
Key Characteristics: 1. Power of prediction : model can help to predict customer preferences of a product/service and can also determine what could have influenced customers' choice. 2. Market segmentation: the segments can be identified and described based on survey data 3. Useful for experimenting new features and measuring the value of existing product in the market.
How does it work?
With the help of relevant data the Choice model can be used at each stage of AER strategies (Acquisition, Expansion, Retention). And by using GLM(generalised linear model) class of regression models in R, which supports non-normalised distribution. In R programming. glm() function uses logistic regression to provide a binary output by using set of predictor variables, basic syntax of this function is as follows: **glm**(formula, **family**=familytype(**link**=linkfunction), **data**=)

```
# Logistic Regression
# where F is a binary factor and
# x1-x3 are continuous predictors
fit <- glm(F~x1+x2+x3,data=mydata,family=binomial())
summary(fit) # display results
```

Generalized Linear model

y=g(f(x)), g here is a linker function for the output of f(x)

y = dependent variable, which provides a binary result

x = set of independent variables

Figure no.(4) : R code for Choice model

Implementation example: Consider a business trying to predict churn rate of their customers after introducing a new product feature in their existing product. The dependent variables could be obtained from data of a survey, which will consist of satisfaction scores for existing features along with rating for new feature. The output will be an independent variable which will provide answer if the new feature should be added or not.
Usability: Choice model is a powerful tool when a company wishes to:

1. Design a new product or service
2. Re-design an existing product to increase profitability
3. Determine viability of a new concept before deciding on what features it may include
4. Establish better awareness what impact customers' precedence in the market share
5. Setting right prices of products/services for maximum profitability

Conjoint analysis is a statistical method based on survey data that is employed in market research to ascertain how consumers value various characteristics (features, functions, and advantages) that make up a specific good or service. The main aim of conjoint analysis is to figure out which exact combination of set of attributes of a product/service is most influenced by choice of respondents and decision making. It is a powerful tool for businesses/companies who wishes to create **Sustainable competitive advantage(SCA)** for their new offerings.
Key Characteristics: 1. Power of prediction: just like Choice model, Conjoint analysis is powerful method to predict market share a new product and its value worth for customers.
2. Freedom of innovation: By leveraging functionality of conjoint analysis, companies can invest and implement innovations with confidence.
3. Tool of SCA: The analysis is most helpful for creating SCA where **all competitors react** to new innovations or modifications in an existing product in the market.

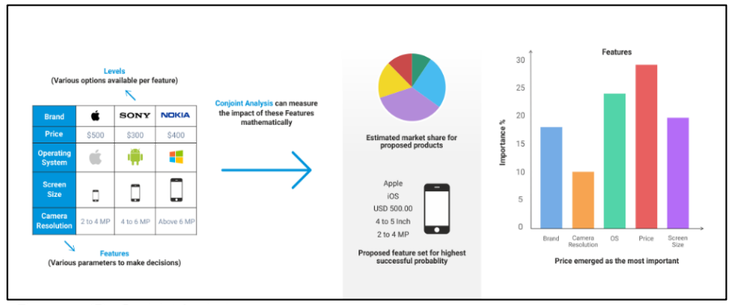

Figure no. (5) : Conjoint Analysis process

How does it work?
1. Break down product/services into features/attributes and levels, for example as presented in the Figure no. (5)
2. Identifying customer preferences by set of combinations of the features. And create limited and feasible number of product concepts using combination of features and levels. **3.** Design a product survey to ask each respondent to choose the best suitable product concept based on their choice.
4. Using all respondent data, conjoint analysis can provide numerous insights about preferences of consumers and market share of a specific product. From Figure no.(3) we can see that iphone use gives highest preference, then the OS, after that the brand itself.

Conjoint Analysis

- **Estimation of the utility values**
  - The additive model of the Conjoint Analysis is defined as:

$$y_k = \sum_{j=1}^{J}\sum_{m=1}^{M_j}\beta_{jm}\cdot x_{jm}$$

with

y_k : estimated total utility for incentive k
β_{jm} : partial utility for value (category) m of factor j

$$x_{jm} = \begin{cases} 1 & \text{if incentive k has value m of factor j} \\ 0 & \text{else} \end{cases}$$

Figure no.(6) : Conjoint Analysis Formula

Usability: **1.** Identify customer decision making patterns when choosing a product. **2.** Managers can determine how does the price of product changes preferences of the customers. **3.** Which trade-off customer is likely to make among set of given functionalities of the products.
4. Can help to find which exact improvements need to be made in existing product to retain customers and attract new ones.
5. Identify position of our product compared to other competitors in the market.

Response Models : Instead of being dependent on heuristic methods for **managing limited resources** of a business or an organization**,** the response model provides an effective way of managing marketing resources and predicting marketing outcomes by following data-driven decision making. In each stage of marketing strategy while following core marketing principles, it is vital to pay attention to resources allocation for implementing those stages. Hence even though this model is present at the end of this toolkit, in practicality budget planning of all three models is very much dependent on current one.
Key Characteristics: 1. Precision in allocation: Since the model follows a data-driven approach, the allocation suggestions produced with accuracy to maximize the profit and/or minimize the loss. 2. Highly adaptable: Even in very dynamic circumstances where customers, competitors and company's internal processes keep changing, the model can adapt or remain unaffected by such situations.

How does it work?
From the Figure no.(5) we can see the possible Input parameters for the model, which could be product design, Price of the product, advertising budget, promotional spending, selling efforts of the teams etc. External inputs involves Competitive measures taken by the organization and uncontrollable environmental factors. If we consider expected output as profit levels, we can use formula from figure no. (6) to get allocation required for each input resource from the formula.

Usability:

1. Useful for reliable budget allocation for marketing strategy
2. Helps in understanding drivers of the sales
3. Understand and measure ROI
4. It can help to predict future performance
5. Determine optimal budget to reach a specific target

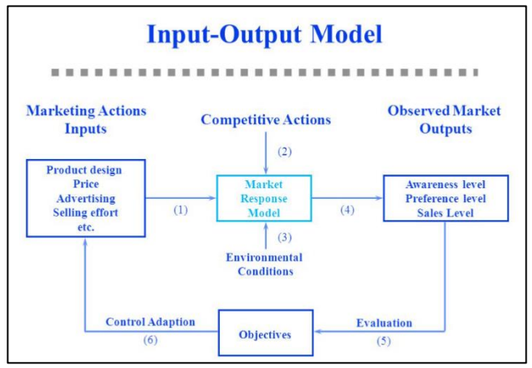

Figure no.(7) : Response model flowchart

Hanssens and Dekimpe (2017) have formulated the basic response model as below:

$$S_t = e^c M_t^{\beta} X_t^{\gamma} Z_t^{\delta} e_t^{u}$$

where S_t = Sales in time t;

M_t^{β} = Marketing support in time t,

X_t^{γ} = Firm – controlled variables,

Z_t^{δ} = Uncontrollable (Environmental) factors, and

e_t^{u} = Residual.

By taking logarithm of equation: (1), it may be expressed as:

$$\ln(S_t) = c + \beta \ln(M_t) + \gamma \ln(X_t) + \delta \ln(Z_t) + u_t$$

where u_t = Error term (Residual).

Figure no.(8) : Response model formula

References

1106518271. (n.d.). *Cluster Dendrogram*. Biostars.

https://www.biostars.org/p/326557/

https://uc-r.github.io/public/images/analytics/clustering/hierarchical/unnamed-chunk-13-1.png

Conjoint analysis white paper. (2022, May 18). *Qualtrics XM // The Leading Experience Management Software*.

https://www.qualtrics.com/support/conjoint-project/getting-started-conjoints/getting-started-choice-based/conjoint-analysis-white-paper/

Conjoint analysis. (2003, August 14). In *Wikipedia, the free encyclopedia*.

https://en.wikipedia.org/wiki/Conjoint_analysis

Conjoint Analysis. (n.d.). Institute for Statistics and Mathematics: Resources Web Page.

https://statmath.wu.ac.at/courses/as_spss/Conjoint

Conjoint.ly. (n.d.). What is conjoint analysis? *Tools and support for product and pricing research - Conjoint.ly*.

https://conjointly.com/guides/what-is-conjoint-analysis/

De Gruyter. (2020, September 8). *Sales – Response model in marketing revisited in the times of uncertainty and global turmoil*.

https://www.degruyter.com/document/doi/10.1515/roms-2020-0050/html

Girish Punj, & David W. Stewart. (1983). *Cluster Analysis in Marketing Research: Review and Suggestions for Application*. School of Business Administration.

https://journals.sagepub.com/doi/full/10.1177/002224378302000204

Heather Turner. (2008). *Introduction to Generalized Linear Models*. ESRC National Centre for Research Methods, UK and Department of Statistics, University of Warwick, UK.

https://statmath.wu.ac.at/courses/heather_turner/glmCourse_001.pdf

An Introduction to Cluster Analysis. (2018, February 15). Alchemer.

https://www.alchemer.com/resources/blog/cluster-analysis/#:~:text=Marketers%20commonly%20use%20cluster%20analysis,clusters%20find%20relevant%20and%20valuable

Nguyen, M. (2022, May 6). *19.6 market response model | Marketing research*. Bookdown.

https://bookdown.org/mike/marketing_research/market-response-model.html

Priya Pedamkar. (2021, October 22). *GLM in R*. EDUCBA.

https://www.educba.com/glm-in-r/

Rosie Cornish. (2007). *Cluster Analysis*. Statistics support for students - www.statstutor.ac.uk.

https://www.statstutor.ac.uk/resources/uploaded/clusteranalysis.pdf

Saif Ali Kheraj. (2019, September 22). *Euclidean distance formula*. Towardsdatascience.

https://towardsdatascience.com/similarity-measures-e3dbd4e58660

(n.d.). https://www.statstutor.ac.uk/resources/uploaded/clusteranalysis.pdf