Advances in Brain-Computer Interfaces and Neural Integration ISSN: 3142-9319

PRIME OPEN ACCESS

**Volume 2, Issue 2**

**Research Article**

**Date of Submission:** 10 July, 2026
**Date of Acceptance:** 05 August, 2026
**Date of Publication:** 12 August, 2026

# A Comparative Study of the Application of Artificial Intelligence in the Target Market: An Interdisciplinary Analysis

**Seyed Mahdi Fareghi***

Department of Sports Management, University of Tehran, Iran

***Corresponding Author:**
Seyyed Mehdi Fareghi, Department of Sports Management, University of Tehran, Iran.

**Citation:** Fareghi, S. M. (2026). A Comparative Study of the Application of Artificial Intelligence in the Target Market: An Interdisciplinary Analysis. *Adv Brain-Computer Interfaces Neural Integr*, 2(2), 01-13.

## Abstract
The purpose of this study is to examine Likert scale data with the aim of grouping and identifying the most flawless grouping through artificial intelligence. This type of data is widely used in various articles and research in various social sciences and humanities around the world. One of these valuable applications is in marketing. Data of this study were collected in the style of humanities research and their reliability and validity were examined and confirmed with common statistical methods worldwide so that their accuracy can be relied upon for generalization to the statistical community on a large scale. 460 sample responses from different individuals were collected using questionnaires from articles published in reputable journals. To analyze the success of clustering and detect similarities between clusters, artificial intelligence criteria and inferential statistics were used together to obtain stronger results. Results,Evaluated through effect size statistical measures, it was found that except for the farthest-first. Others did not show significant performance across the variables. The Findings suggest that clustering from the out toward the center is the most effective approach. This is An effective approach for categorizing when handling Likert-scale data that does not follow a normal distribution and construction.

**Keywords:** Machine Learning, Comparison Clustering, Dwass-Steel-Critchlow-Fligner, Nonparametric Data

### Introduction
The clustering technique has been around for over half a century, but there is still capacity for improvement. A frequently asked question is: what type of clustering is appropriate to use? Studies have compared various clustering methods; however, when dealing with multiple heterogeneous variables and different types of data—such as binary, nominal, ordinal, and continuous variables—direct comparisons cannot be made across varying sample sizes without considering effect sizes and error values [1]. Therefore, the aim to identify the best type of clustering for Likert scale data by utilizing data mining and statistical methods, alongside real data collected from male massage consumers. One application of clustering is in marketing.

effective segmentation through appropriate feature selection is crucial for market success. Accurate segmentation enables a company to gain strategic advantages, while irrelevant variables disrupt the process, wasting resources and leading to divergence from goals [2]. The application of clustering in marketing was first introduced by Saunders in 1980. His findings indicate that cluster analysis is valuable for segmenting consumer markets using variables such as needs, attitudes, and lifestyle. Other approaches, including benefit bundles and psychological variables, have also been employed. Recent advancements in clustering methodologies, such as flexible and componential segmentation, are under investigation. Cluster analysis has also proven useful in experimental settings and product positioning, with many researchers implementing clustering techniques in their publications.

While neural networks and decision -making trees are usually market segmentation, as defined by, "market division into distinct groups from buyers who have different needs, properties, or behaviors and who may need separate marketing

strategies" this is an important targeting [3,4]. Maintain a significant customer value. Division may include four variables: geographical factors (countries, regions, states, cities, neighborhoods, population and climate density), demographic factors (age, life cycle, gender, income, job, job, religion, ethnicity, and generation), psychological factors (lifestyle and personality), Remains. This study aims to address this issue by identifying the most appropriate method based on empirical evidence, specifically through male massage consumer feedback and the use of clustering techniques on Likert spectrum data.

What distinguishes this research is the use of multiple algorithms for comparative analysis, providing a comprehensive perspective, and results derived from scientifically validated methods based on accepted metrics, enhancing reliability. Unlike other investigations, this research incorporates innovative methodologies and emphasizes valid data grounded in standard statistical indicators, a facet often underrepresented in other studies. The credibility of the data facilitates a more robust generalization of findings to the broader population. This research focuses on comparing clustering methods to identify the most suitable approaches for Likert scale scientific data, combining data mining techniques from computer science with quantitative research methods in the humanities and sciences to address these ambiguities.

## Related Works

The First Comparison of Clustering Methods. Under the Title Comparison of Some Clus-ter Analysis Methods Done by Gower (1967), after Him, with the Increasing Growth of Algorithms, Some Scattered Comparisons Were Made, the Most Remarkable of Which Will be Introduced in the Following.

Believe That the Spectral Approach Outperforms Other Techniques with Exceptional Performance [5]. However, the Hierarchical Method is Highly Sensitive to the Number of Features, Which Affects Its Effectiveness. Algorithms Show Different Levels of Per-formance Depending on the Number of Features Used.

Researches Outcomes Concerning the Portrayal of the Data Through Centroids Reveal That K-Means Surpasses Clara When Examining Squared Euclidean Distances Relative to the Centroid [6]. Conversely, in Terms of Unsquared Distances, Clara Demonstrates Su-periority.

In their study found out that the Combination of Internal and External Evaluation Measures Can Lead to Different Cluster Results [7]. The K-Means (OS; K = 5) Performs Well with External Evaluation Measures, while K-Means (PCA; K = 7) is Rated Third Best by Cohen's Kappa and F-Score, But Lacks Internal validation. Some Methods, Like X-means and DBSCAN (OS), Show Mediocre Results. It's Important To consider Both Internal and External Evaluations to Ensure Accurate Cluster Results.

In Comparing Clustering Methods, examine the Performance of Distance-Based Parti-tioning Methods on Various Simulated Data Sets [8]. Eight Methods Were Evaluated, Including Approaches for Constructing Dissimilarity Matrices, Adapting K-Means to Mixed Data, and Reducing Variables. The Benchmark highlighted similarities and differ-ences in algorithm performance, with KAMILA, FAMD/K-Means, and K-Prototypes Standing Out as Top Performers.

In study, K-means clustering, OPTICS, and Gaussian mixture model (GMM) clustering were compared on vibration datasets, examining the effect of feature combinations, PCA feature selection, and the number of clusters [9]. The results showed that averag-ing and variance-based features were more effective than shape-based features, and clustering with K-means was slightly better than GMM MoreTechnical Details of These Papers are Presented in the Table.1

| Authors | Sample Size | Subject Used for Clustering | Compare Method | Compared Algorithm | Result |
|---|---|---|---|---|---|
| Rodriguez et al. 2019 | 400 | Artificial datasets | distances between classes and correlations between features and average of the best accuracies | hierarchical, Clara, k-means, spectral, hc model, subspace, optics, DBSCAN, EM | Spectral algorithm consistently provided best performance for datasets with 10+ feature |
| Hening 2021 | 35043 sample | Different categories (images, Texts, digits) there is large variation between the data sets | Just compare output of software | K-means, (Clara), (m-clust), (em-skew), (teigen), Single linkage hierarchical clustering, Average linkage hierarchical clustering, Complete linkage hierarchical clustering, Spectral clustering | The Gaussian mixture is the best for the largest amount of data. Differences between the other methods are not that pronounced, and all of them did best in some data sets. |

| | | | | | |
|---|---|---|---|---|---|
| Kaya And Schoop 2022 | Ten negotiation experiments of several hundred participants with a total of 7026 exchanged negotiation interactions | Negotiation Support Systems (text) | Principal Component Analysis | K-means, X-means, DBSCAN agglomerative | With internal index k-Means and with external index k-Means or DBSCAN, performance is better. |
| Costa et al 2023 | n = 100 low, 600 moderates and 1000 | Normal Mixed-type data | ANOVA-ARI-AMI-Effect size η2 | KAMILA, FAMD/K-Means, K-Prototypes, M-S K-Means, Mixed RKM, HL/PAM, Mixed K-Means, Gower/PAM | KAMILA, K-Prototypes and sequential Factor Analysis and K-Means clustering typically performed better than other methods |
| Sepin et al 2024 | data1:512 data2: ambiguous data3: ambiguous | vibration data | 1 Arithmetic mean of absolute values (Abs Mean), 2Median of absolute values (Abs Median), 3 Standard deviation (Std), 4 Interquartile range (IQR), 5 Skewness of absolute values, 6 Kurtosis of absolute values | K-means, Gaussian Mixture Model, OPTICS | K-means outperformed GMM slightly, whereas OPTICS performed significantly worse. |

**Table 1: Description of Literature**

## Metodology

The paradigm of this study is positivism, and its approach is quantitative. The target population consists of male massage service consumers in Iran. The Steps of Data Mining Include Collecting Data, selecting a Suitable Attribute for the Algorithm, running it and receiving output, finding the best Cluster and finally mining additional Data from the Output. Comparing Methods Add to the Chart According to the Figure.1


**Figure 1: Steps Path Way**

A cluster sampling method was employed to obtain a representative sample. To com-pare cluster variances, a maximum of 20 clusters was anticipated, and the sample size was determined using formula-based methods with Gpower software Figure 2 [10,11]. with a maximum of Twenty number of clusters and a minimum of two clusters, along with seven variables having an effect size of 0.1, a confidence level of 95%, and a pow-er of 0.95, the required sample size was calculated to be 460 members.

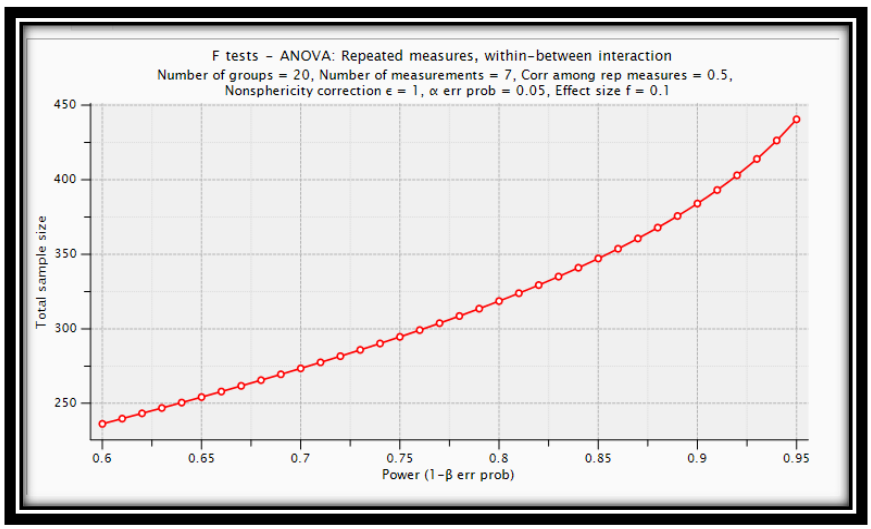


**Figure 2: Gpower Software Output**

**Instruments**

For Measuring the Current Variables, Data Collection Instruments Developed in Recent Years Were Selected, Such as the Developed Scales [12-14]. These Instruments Underwent Face Validity and Content Validity tests. Subsequently, the Final Questionnaire Was Developed Based on the Reflective Nature of the Items, Which Included Demographic Questions and Items With a 5-Point Likert Scale [15]. in Pre-Processing and Primary Reliability Measurement and For Post-Hoc Tests Were Used to Calculate and Compare Clustering Methods, JAMOVI 2.5 Used to Calculate and Compare Clustering Methods and Weka 3.9.6 With Its Plugins and RapidMiner 9.1.0 With its Plugins as Data Mining Platform's. The Reason of Using These Software's, Primarily Their Free Availability, Which Makes Them Accessible to Everyone Additionally, both are User-Friendly Thirdly the Open-Source Library of All Methods (Table2) and Forth in the Clustering Section are Different and Has Overlap With Each Other، and Use Them Helps to Spread Research in More Fields.

| Data Mining Platform | Rapid Miner | Weka |
|---|---|---|
| Algorithms Compared | K-Mean, K-Mean (H2O), K-Mean Kernal, K-Mean fast, X-Mean, K-Medoids, Fuzzy C-mean, Agglomerative Clustering, DBSCAN, Top-Down Clustering, Flatten Clustering, G-Means, Random | BANG-File, Canopy, Cascade Simple K-Mean, Cobweb, farthest first, Gen-Clust plus plus, Hierarchical Clustering, LVQ, EM, make density-based Clustering, OPTICS, SOM, sIB, , X-means |

**Table 2: Software Libraries**

## Performance

**First Step**

After collecting and preparing Data, Clustering Operations Were Performed for Each Algorithm with Varying Quantity until the Initial Clustering Threshold Without Any Specific Membership Was Determined.

**Second Step**

The Quality of the Generated Clusters Was Evaluated in Terms of the Number of Members and Recommend That Clusters With 2% Or Less Compared to the Total Sample Population, be Considered Poor Clustering, and Clusters With 5% Until 2% be Considered Near to Poor. After Identifying This Range, it Should be Further Examined Using the ANOVA Test to Determine Which Clusters are Similar, This Test is Most Known for Data Miners as F Test [16].

$$SST = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(\chi_{ij-\bar{\chi}}\right)^2 \quad (1)$$

$$SSB = \sum_{j=1}^{k}\left(\overline{\chi_j} - \bar{x}\right)^2 \quad (2)$$

$$SSw = \sum_{j=1}^{k}\sum_{i=1}^{n_j}\left(\chi_{ij-\overline{\chi_j}}\right)^2 \quad (3)$$

This test could be run if data known normal by one of D'Agostino, Anderson-Darling, Shapiro-Wilk, or Kolmogorov-Smirnov tests [17]. results of ANOVA Test Determine Which Clusters are Similar, Indicating Poor Clustering Quality [16]. If Data Known as un-normal and There is More Than Two Groups to Compare and Many Variables; the Non-Parametric Equivalent Test Means Kruskal–Wallis Test That Must be Implemented [18]. If the Results Show that Significant, it Only Proves that There are at Least two groups that are not Similar Together, But the Task is Unclear and the Necessity for a Post-hoc Test will be Disclosed and the Next Step is Following that [19].

$$H = (N-1)\frac{\sum_{i=1}^{g}(\overline{r_i} - \bar{r})^2}{\sum_{i=1}^{g}\sum_{j=1}^{n_i}\left(r_{ij} - \bar{r}\right)^2} \quad (4)$$

N; is the total number of observations across all groups

g; is the number of groups, $n_i$ is the number of observations in group i, $r_{ij}$; is the rank (among all observations) of observation j from group i,

$$\overline{r_i} = \frac{\sum_{j=1}^{n_i} r_{ji}}{n_i} \quad (5)$$

is the average rank of all observations in group i, $\bar{r} = \frac{1}{2}(N+1)$ is the average of all the $r_{ij}$

### Third Step

After Narrowing Down the Remaining Choices, the Post-Hoc Dwass-Steel-Critchlow-Fligner (DSCF) Test Identifies Which Clusters Have the Least Similarity Between Them, as it Compares Each Cluster Pair [20]. it Necessitates the Recalculation of Ranks for Every Treatment Combination. the $w_{ij}$ Statistic is Computed for Each of These Combinations. This Test All-Treatment Multiple Comparison Procedure is Based on the Maximum of the Studentized Two-Sample Wilcoxon Statistics, Computed Over All Pairs of Samples. for Each Pair $i<j$ let

$$S_{ij} = \sum_{a=1}^{n_i} \sum_{b=1}^{n_j} \psi(X_{ia} - X_{jb}) + \frac{n_i(n_{i+1})}{2} \qquad (6)$$

Where ψ(t)=1 if t>0 and 0 otherwise, be the rank sum associated with the i^thSample When Ranked with Sample j. in Addition, let

$$W_{ij} = \frac{S_{ij} - \frac{n_i(n_i+n_j+1)}{2}}{\left\{\frac{n_i n_{ij}(n_i+n_j+1)}{24}\right\}^{\frac{1}{2}}} \qquad (7)$$

denote the studentized Wilcoxon statistic (multiplied by √2). for Each Pair i, j with $1 \le i < j \le k$, the Steel-Dwass procedure declares that $\theta_i \neq \theta_j$ whenever $|W_{ij}| \ge w_\propto$ Where $w_\propto$ Is Chosen so That

$$P_0\left(max_{i<j}|W_{ij}| \ge w_\propto\right) \qquad (8)$$

and the Probability $P_0$(-) is Computed Under the Null Hypothesis Assumption $\theta_1 = \cdots = \theta_k$ an Expression Analogous to (8) is Given for the Case When Not All the $\theta_i$ are Equal, That Utilizes the Same value of $w_\propto$. it is Used to Derive the Simultaneous Confidence Intervals for the Differences

$\Delta_{ij} = \theta_i - \theta_j \; 1 \le i < j \le k$, and to show that the procedure satisfies property (Control of the maximum type I error rate). Also, for a large Sample Approximation for $w_a$ is Provided Below, Which Ensures That the Probability in (8) Has Limit ∞ for Equal $n_i$, and limit ≤∝ when the sample sizes are unequal.

$$\text{Let } D_{ij} = \frac{\sqrt{2}(\bar{X}_i - \bar{X}_j)}{S_p\sqrt{\frac{1}{n_i}+\frac{1}{n_j}}} \qquad (9)$$

where $\bar{X}_i = \sum_{m=1}^{n_i} X_{im}/n_i$

And $S_P^2 = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(X_{ij}-\bar{X}_i)^2}{(N-k)}$. note that under the null hypothesis, the $\binom{n}{2}$ component vectors

$$D = \left(D_{12}, D_{13,\ldots,} D_{k-1,k}\right) W = \left(W_{12}, W_{13,\ldots,} W_{k-1,k}\right)$$

both have limiting multivariate normal distributions with mean vector 0 and the same covariance matrix, when $n_i \to \infty$ and $\frac{n_i}{N} \to \lambda_i, 0 < \lambda_i < 1$ for the covariance structure). Thus, for large sample sizes, we may take

$W_\propto = q_{k,\infty}^{a}$ thereby ensuring that the probability in (8) has limit a for equal sample sizes, and limit < a for unequal sample sizes. This gives the standard approximation for large and equal sample sizes. the conservative nature of this result for unequal n_i is new. as with Tukey-Kramer, the probability will generally be quite close to □ [21].

### Forth Step

The Optimal Clustering is Determined Among the Remaining Options Using the Silhouette Test who offered by Rousseeuw1987 [22].

$$S(i) = \frac{b(i)-a(i)}{max\{a(i),b(i)\}} \qquad (10)$$

### Fifth step

And for Scientific Result's Comparison, The Effect-Size Index Used at the End. it is Expressed to Highlight the Practical Significance and Implications for Information Systems Research, it is Essential for Statistical Studies to Consistently Report Effect Sizes.

$$\eta_p^2 = \frac{SS_{ef}}{SS_{ef}+SS_{er}} SS_{ef} \quad (11)$$

; sum of squares for the effect $SS_{er}$ ;sum of squared errors

$$\omega^2 = \frac{df_{ef}(MS_{ef}-MS_{er})}{SS_t+MS_{er}} \quad (12)$$

$MS_{ef}$ mean square of the effect, $MS_{ef}$; mean square error, $SS_t$; the total sum of squares, $DF_{ef}$; degrees of freedom for the effect by Detailing Effect Sizes, Researchers Not Only Demonstrate the Applicability of Their Findings But Also Facilitate Evidence-Based Practitioners in Comparing the Effects of Various Interventions Across Different Studies [8,23]. About Importance of Effect Sizes express that Effect sizes serve several critical functions: They help determine the practical or theoretical importance of an effect [24]. They allow for the comparison of effects across different studies. They assist in power analysis, which is crucial for determining appropriate sample sizes in future research [24].

## Results

After performing clustering for each algorithm and identifying the optimal cluster, the final results for the best cluster in each algorithm are presented in Table 3

| Cluster Metod | Cluster quantity | Seeding | With Out Membership | Poor Cluster | Kruskal Wallis | DSCF |
|---|---|---|---|---|---|---|
| **K-Mean** | 2 | - | no | no | Significant For Courtesy, Perceived Value, Aesthetics | Significant Courtesy-Perceived Value-Aesthetics |
| **K-Mean Fast** | 2 | - | no | no | Significant For Courtesy, Aesthetics, Perceived value | Significant For Courtesy, Aesthetics, Perceived value |
| **K-Mean H2O** | 2 | - | no | no | Unsignificant For at attitude | Significant For Access Aesthetics, Perceived Value, Courtesy, Loyalty |
| **K-Mean Kernel** | 2 | - | no | no | Unsignificant For Access, | Unsignificant For Access, |
| **K-Medoid** | 4 | - | no | no | Significant All Except Loyalty | Significant For at attitude |
| **Fuzzy C-Mean** | 2 | - | no | no | Significant For Aesthetics, Courtesy, Access, at attitude, Perceived value | Significant For Aesthetics, Courtesy, Access, at attitude, Perceived value - |
| **Canopy** | 2 | 5 | no | no | Significant For Access, Courtesy, Perceived value, Quality | Significant For Access, Courtesy, Perceived value, Quality |
| **Cascade k-Mean** | 2 | - | no | no | Significant For Aesthetics, Courtesy, Perceived value | Significant For Aesthetics, Courtesy, Perceived value |
| **EM** | 2 | - | no | no | Significant For Access, Courtesy, Perceived value | Significant For Access, Courtesy, Perceived value |
| **Farthest First** | 2 | 2 | no | no | Significant All Index | Significant All Index |
| **Gen clust plus Plus** | 2 | 10 | no | no | Significant For Courtesy, at attitude | Significant For Courtesy, at attitude |
| **LVQ** | 2 | - | no | no | Significant Except Loyalty | Significant Except Loyalty |
| **Make Density Based** | 2 | - | no | no | Significant Aesthetics, Perceived value, Courtesy | Significant Aesthetics, Perceived value, Courtesy |
| **sIB** | 2 | 1 | no | no | Significant Except Loyalty | Significant Except Loyalty |
| **X-Mean** | 2 | 10 | no | no | Significant For Aesthetics, Courtesy, Perceived value | Significant For Aesthetics, Courtesy, Perceived value |

**Table 3: Optimum Clustering for Each Method**

Researches comparison Thirteen normality tests concluded that for large sample sizes, the Shapiro-Wilk and D'Agostino-Pearson tests are the most powerful, Also Shapiro-Wilk Test Is The Most Powerful Test Followed By The Anderson-Darling Test [17]. According to these tests, Data does not follow a normal distribution; however, all indexes Table 4 yielded valuable results.

| **KMO** | **$\chi^2$** | **df** | **p-value** | **CFI** | **TLI** | **SRMR** | **RMSEA** |
|---|---|---|---|---|---|---|---|
| 0.780 | 7160.6 | 329 | < .001 | 0.922 | 0.911 | 0.054 | 0.049 |
| **veriable** | **access** | **Aesthetic** | **Attitude** | **Courtesy** | **Loyalty** | **quality** | **PrecievedValue** |
| McDonald's ω | 0.849 | 0.848 | 0.867 | 0.832 | 0.826 | 0.822 | 0.844 |

**Table 4: Kaiser-Mayer-Olkin, Bartlets, Validity and reliability Indexes**

After conducting the KMO and Bartlett's test to assess sample adequacy and sphericity, a KMO(Kaiser-Meyer-Olkin Measure of Sampling Adequacy) value of 0.784 was obtained, exceeding the threshold of 0.6. According to the Researchers the number of observations is sufficient for factor analysis [25]. The sphericity of relationships among indicators has been validated as crucial for factor analysis due to the significance of the Bartlett test.

The Shapiro-Wilk results indicated abnormality in the data, as shown by high skewness and kurtosis. Consequently, nonparametric methods were employed. The equivalent nonparametric test for normal analysis of variance is the Kruskal-Wallis test, which assesses the null hypothesis that the clusters (groups) are similar. However, Kruskal-Wallis test only indicates whether there is a difference among at least one of the clusters, necessitating a post-hoc test to compare the clusters individually.The Dwass-Steel-Critchlow-Fligner post-hoc test can determine which groups are similar or different by compairing each pair of groups [26].This test is available in NCSS, SAS,Stats Direct, Jamovi, JASP, R packages, and Excel macros. it is noted to be more robust and effective for large datasets compared to the Conover-Iman [27]. Among nonparametric tests, Mood's test performs best in controlling Type I error, and in Duncan's test results, both the Type I error rate and the power of the tests are significantly improved . It has also been noticed that a changed version of the Steel-Dwass method, called the Steel-Dwass-Critchlow-Fligner method, along with its different step-by-step versions that use their own rankings, works better than the method that uses a combined ranking [21,26,28]. Once the optimal cluster is identified, the next step is to calculate the effect size of the variables in clustering, as shown in Table 5.

| **Cluster Method** | **Quality** | **Aesthetics** | **Perceived value** | **Attitude** | **Access** | **Courtesy** | **Loyalty** | **Silhouette Index** |
|---|---|---|---|---|---|---|---|---|
| **K-Mean E-S** | 0.00 | 0.04 | 0.33 | 0.00 | 0.00 | 0.57 | 0.00 | 0.209 |
| **P Value** | unsignificant | significant | significant | unsignificant | unsignificant | significant | unsignificant | |
| **K-Mean Fast E-S** | 0.00 | 0.04 | 0.33 | 0.00 | 0.00 | 0.57 | 0.00 | 0.210 |
| **P Value** | unsignificant | significant | significant | unsignificant | unsignificant | significant | unsignificant | |
| **X-Mean E-S** | 0.00 | 0.05 | 0.34 | 0.00 | 0.00 | 0.54 | 0.00 | 0.210 |
| **P Value** | unsignificant | significant | significant | unsignificant | unsignificant | significant | unsignificant | |
| **Cascade K-Mean E-S** | 0.00 | 0.06 | 0.34 | 0.00 | 0.00 | 0.54 | 0.00 | 0.209 |
| **P Value** | unsignificant | significant | significant | unsignificant | unsignificant | significant | unsignificant | |
| **Make-Density Based E-S** | 0.00 | 0.04 | 0.29 | 0.00 | 0.00 | 0.60 | 0.00 | 0.209 |
| **P Value** | unsignificant | significant | significant | unsignificant | unsignificant | significant | unsignificant | |
| **K-Mean-H2o E-S** | 0.01 | 0.04 | 0.33 | 0.00 | 0.02 | 0.55 | 0.01 | 0.207 |
| **P Value** | unsignificant | significant | significant | unsignificant | significant | significant | significant | |
| **Fuzzy C-Mean E-S** | 0.003 | 0.012 | 0.390 | 0.044 | 0.009 | 0.462 | 0.004 | 0.201 |
| **P Value** | unsignificant | significant | significant | significant | significant | significant | unsignificant | |
| **sIB E-S** | 0.01 | 0.02 | 0.27 | 0.01 | 0.02 | 0.54 | 0.00 | 0.198 |
| **P Value** | significant | significant | significant | significant | significant | significant | unsignificant | |
| **Em E-S** | 0.00 | 0.01 | 0.28 | 0.00 | 0.01 | 0.57 | 0.00 | 0.200 |
| **P Value** | unsignificant | unsignificant | significant | unsignificant | significant | significant | unsignificant | |
| **Canopy E-S** | 0.01 | 0.00 | 0.43 | 0.00 | 0.04 | 0.35 | 0.00 | 0.196 |
| **P Value** | significant | unsignificant | significant | unsignificant | significant | significant | unsignificant | |
| **LVQ E-S** | 0.08 | 0.01 | 0.49 | 0.04 | 0.06 | 0.02 | 0.00 | 0.166 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **P Value** | significant | significant | significant | significant | significant | significant | unsignificant | |
| **Gen clust Plus Plus E-S** | 0.00 | 0.00 | 0.01 | 0.68 | 0.00 | 0.01 | 0.00 | 0.157 |
| **P Value** | unsignificant | unsignificant | unsignificant | significant | unsignificant | significant | unsignificant | |
| **Farthest First E-S** | 0.11 | 0.04 | 0.05 | 0.04 | 0.01 | 0.05 | 0.02 | 0.175 |
| **P Value** | significant | significant | significant | significant | significant | significant | significant | |
| **K-Mean-Kernel E-S** | 0.03 | 0.02 | 0.01 | 0.12 | 0.00 | 0.05 | 0.06 | 0.047 |
| **P Value** | significant | significant | significant | significant | unsignificant | significant | significant | |
| **Kmedoid E-S** | 0.16 | 0.03 | 0.13 | 0.24 | 0.05 | 0.39 | 0.04 | 0.167 |
| **P Value** | unsignificant | unsignificant | unsignificant | significant | unsignificant | unsignificant | unsignificant | |
| **Random, Clope, SOM, Cobweb, DBSCAN** | Unsignificant all | | | | | | | |

**Table 5: Effect Size Index (E-F)**

By examining the output results of the cluster selection table for identifying optimal clusters, it is clearly observable that the k-medoids algorithm, due to the differences in the input data (selection of median values for each variable), indicates a four-cluster so-lution as optimal when applying the same criteria and indices, whereas in all other algorithms, a two-cluster solution is identified as the optimal clustering.

Another notable point that can be observed is that using the same data in such comparisons helps better reveal the performance of the algorithms. In fact, the algorithms each time use one of the seven available variables as the basis attribute for clustering, and accordingly, a different effect size is assigned to each variable. Evidence shows that K-Mean, K-Mean Fast, X-Mean, Cascade K-Mean, Make Density Based, K-Mean-H2O, k-medoid, Fuzzy C-Mean, sIB, Em algorithms indicate the courtesy index with the highest effect size; the Gen clust Plus Plus and Kernel K-Mean algorithms simultaneously show Attitude as the variable with the highest effect size; the LVQ and Canopy algorithms indicate perceived value as the variable with the highest effect size; and the Farthest First algorithm shows Quality variable as the the highest effect size.

As mentioned in literature review, clustering based on central tendencies (such as the mean) is inappropriate for binary data (such as gender or employment status) and ordinal data (such as education levels) due to the meaningless nature of calculating averages for this type of data. Therefore, these were excluded from the evaluation. author only considered the 5-point Likert scale questionnaire responses for standardization purposes, which some studies in the literature have overlooked. In the initial clustering phase, G-Means, BANG-File, and OPTICS were not executed; instead, numeric attributes were first adopted. The resulting clustering outputs for agglomerative (919 clusters), hierarchical (459 in the first step and 1 in the second), and top-down algorithms (34 clusters, with 23 lacking members) were imbalanced and unacceptable. The similarity in clustering results between the X-means and Fast K-Means algorithms in RapidMiner indicated that they differed only in name while having the same structure. Consequently, X-Means was disregarded in RapidMiner, and its output was recorded and utilized in Weka.

clustering method testing all indices, meaning that each time, one variable is consid-ered as an index to calculate the distance to other variables. For all algorithm Euclidean distance cosidered. The clustering process began with two clusters, one of which was poorly formed, ultimately resulting in 16 clusters. The poorly performing cluster began at 12 clusters, with 38 clusters showing no membership. Except for the 2 and 3 clus-ters, all remaining clusters were similar, sharing common members. The final step re-vealed that a higher silhouette index was associated with 2 clusters, with observed similarities in courtesy, aesthetics, and perceived value indices. In the K-Means H2O clustering process, a decline in clustering quality was noted starting at 11 clusters, with poor clustering observed at 17 clusters. A cluster without membership appeared at 30 clusters. Again, aside from the 2 and 3 clusters, all other clusters were similar, with a higher silhouette index associated with 2 clusters, showing similarities in access, aes-thetics, perceived value, courtesy, and loyalty indices. For K-Means Kernel, clustering quality began to deteriorate at 11 clusters, and poor clustering was observed at 13 clusters, with a cluster lacking membership appearing at 60 clusters. This evidence sug-gests that this method behaves similarly to its counterparts, except it fails to reveal clusters without membership. Except for the 2 and 3 clusters, all clusters were similar, and a higher silhouette index was linked to 2 clusters, with dissimilarities observed only for the access index.

For K-Means Fast, clustering quality decreased starting at 11 clusters, with poor cluster-ing observed at 13 clusters and clusters without membership noted at 38. This evidence suggests that this method behaves like K-Means. Except for the 2 and 3 clusters, all others were similar, with a higher silhouette index belonging to 2 clusters. Unsimilarity between

clusters was observed only for courtesy, aesthetics, and perceived value indi-ces. For the execution of the K-medoids algorithm, instead of using the average of the data, the mode was calculated and utilized for clustering. This algorithm operates based on the median index.Additionally, we observed that the arrangement or order of varia-bles does not significantly impact the clustering result, as this model performs cluster-ing based on the dataset itself. Poor clustering began at 9 clusters and continued until 16 clusters, with no membership observed at 17. Except for the 2, 3, and 4 clusters, all other clusters were similar, and the higher silhouette index belonged to the 4 clusters, with dissimilarity between clusters observed only for attitude.There are two adjustable variables in the DBSCAN model: epsilon and the minimum points. Different values for these variables were tested; however, no improvement in clustering quality was ob-served. This means that some members did not belong to any cluster and were referred to as noise. For the random algorithm, clustering quality decreased starting at 11 clus-ters, with poor clustering observed at 26 clusters and clusters without membership not-ed at 46. This method behaves similarly to K-Means and K-Means Kernel, with all clus-ters being similar, resulting in the silhouette test not being conducted.

For Fuzzy C-Means, clustering quality decreased starting at 8 clusters, with poor cluster-ing noted at 12 clusters and a cluster without membership observed at 13. This sug-gests that this method is sensitive; except for 2 and 3 clusters, all others were similar, with a higher silhouette index belonging to 2 clusters. Unsimilarity between clusters was observed only for aesthetics, courtesy, access, attitude, and perceived value. In the Flatten method, poor clustering appeared at 2 clusters, with a cluster without member-ship observed at 8, indicating that this method is unfriendly to Likert scale data. For the Canopy method, there are three adjustable variables: T1, T2, and seeding quantity. Different values for these variables were tested. The default values are seeding = 1, T1 = -1.25, and T2 = -1.0. T1 represents the distance to use; values less than 0 are taken as a positive multiplier for the T2 distance, which also represents distance. Values less than 0 should be set using a heuristic based on the attribute's standard deviation, mak-ing it effective only during batch training. Since these numbers result from scientific work by the producers, we did not alter their relationships; instead, we adjusted them by adding or subtracting from this ratio. These changes did not affect the output re-sults, but modifying the seeding made the results distinguishable. Clusters in field 10 are weak, while in field 1, they are better. Consequently, field 10 is gradually being abandoned as the number of clusters increases. Poor clustering observed at 4 clusters,. All others were similar, with a higher silhouette index belonging to 2 clusters, with 5 seeds, and unsimilarity between clusters was observed for access, courtesy, perceived value, and Quality.

For Cascade K-Mean, there is only one adjustable variable for seeding quantity, with a default of 1. Different values for seeding were tested, but there was no difference be-tween them, even in the silhouette output. Thus, the higher silhouette index belonged to 2 clusters. For CLOPE, one condition for implementing this method is that the data must be nominal, and there is only one adjustable variable, repulsion. Different values were tested for it. All outputs had clusters without membership, indicating that this method is not appropriate for Likert data. For Cobweb, there are two adjustable varia-bles: Acuity and Seeding. Different values were tested for them. The default value for Acuity is 1, and for Seeding, it is 42. It was observed that clusters 1, 2, 3, 32, and 33 were similar, indicating that this method is not appropriate for Likert data. For EM, alt-hough it is automated, the number of clusters is regulative, similar to K-Means. A de-crease in clustering quality started at 11 clusters, with poor clustering at 13 clusters. Clusters without membership were observed at 49. Except for clusters 2, 3, and 5, all clusters were similar, with a higher silhouette index belonging to cluster 2, and unsimi-larity between clusters was observed for Access, Courtesy, and Perceived Value indexes.

For the farthest first, the only regulative index is Seeding. The best seeding is consid-ered a rise in that seeding. Clustering started at 4 clusters, with poor clustering at 2 clusters, and clusters without membership were observed at 17. Except for 2 clusters, all clusters were similar, with a higher silhouette index belonging to it, and similarity between clusters was observed for all indices.For Gen Clust Plus Plus, there are three adjustable variables: Initial Population size, Generations, and Seeding. Different values were tested for them. The default value for Initial Population size is 30, Generations quantity is set to 60, and Seeding is 10. There are only three options that yield similar clusters and a higher silhouette index. They belong to 2 clusters with an Initial Popula-tion size of 20, a Generations quantity of 60, and a Seeding of 10. Unsimilarity be-tween clusters was observed only for Courtesy and Attitude indexes. Additionally, there was no significant difference observed when changing Generations quantity values.

For LVQ, there are two adjustable variables: Learning Rate and Cluster Quantity. Differ-ent values were tested for them. The default value for the Learning Rate is 1.0. The best output was observed at a 0.5 Learning Rate for two clusters. This rate was tested on other cluster quantities. Except for clusters 2, 3, and 4, all others were similar, with a higher silhouette index belonging to cluster 2 with a 0.5 Learning Rate, and unsimi-larity between clusters was observed, except for Loyalty.For Density-Based Clustering, poor clustering was observed at 13 clusters, with clusters without membership ob-served at 51 clusters. Except for 2 clusters, all clusters were similar, with a higher sil-houette index belonging to 2 clusters, and unsimilarity between clusters was observed for Aesthetics, Perceived Value, and Courtesy indexes. For the SOM algorithm, known for automation decisions, the only regulative index is the Learning Rate. Different val-ues were tested for it. Overall, there are only 4 cluster outputs, and all of them were similar, leading to disappointing results.

For sIB, clustering quality decreased starting at 15 clusters, with poor clustering ob-served at 20 clusters. Clusters without membership were noted at 27 clusters, except for the 2 and 3 clusters tested; others were similar and had higher silhouette outputs, with the index belonging to 2 clusters. Unsimilar clusters were disappointing, except for

loyalty. The only regulative index is seeding, with different values tested, showing the best results belong to 1 seeding. For X-Mean, the only regulative index is also seeding; different values were tested, and there were no differences between them, so the de-fault was used. Clustering quality decreased starting at 12 clusters, with poor clustering at 17 clusters. Clusters without membership were observed at 31 clusters, and except for the 2, 3, and 5 clusters, all others were similar. The higher silhouette index be-longed to 2 clusters, and unsimilarity between clusters was observed only for courtesy, aesthetics, and perceived value.

## Discussion and Conclusion

Interpretation of the results shows various points or perspectives. Firstly, K-Mean, K-Mean Fast, X-Mean, Cascade K-Mean, Make Density Based, K-Mean-H2O, k-medoid, Fuzzy C-Mean, sIB, Em, algorithms perform similarly on the basis of their influence on the clustering of results since they group the data in such a way that the politeness feature has the largest effect size. Gen clust Plus Plus and kernel k-means algorithms perform similarly since they group the data in such a way that the attitude feature has the greatest effect size. LVQ and Canopy algorithms perform similarly since they group the data in such a way that the perceived value feature with the highest effect size is selected. However, the Farthest-First algorithm is completely distinct or unique from the other algorithms since it selects the quality feature with respect to effect size. Evidence that helps support or confirms the assertion that the algorithms perform similarly is that they select features to work on data. In this context, effect size refers to the power or influence of a variable, or feature, involved in the grouping or clustering phases of an algorithm.

This group is named as the first group, consisting of ten algorithms. From Table 5, they are listed with the help of the effect size that indicates the power of each. The first algorithm is Make Density Based, the second rank is jointly held by K-Mean, K-Mean Fast, and EM algorithms, the third algorithm is K-Mean-H2O, the fourth rank is jointly held by X-Mean, Cascade K-Mean, and sIB algorithms, the fifth algorithm is Fuzzy C-Means, and finally ending the group ranking, the sixth and last rank is held by K-Medoids algorithms. The second group comprises of Gen clust Plus Plus and K-Means Kernel algorithms. Of these, Gen clust Plus Plus has the highest power in this group and among other algorithms. The third group comprises of LVQ and Canopy algorithms; of these, LVQ is more powerful. The last group comprises of only one algorithm: Farthest First. In statistics, it possesses the lowest power; it possesses the highest geographical or meaningful separation among all other algorithms in the artificial intelligence area that possesses all seven indicators. This characteristic indicates that this algorithm is significantly far away from all other algorithms with no functionality or relationship with any one of them. Though it possesses low power, it is more effective in separating clusters since it is of significant importance in utilizing all seven variables.

## Artificial Intelligence Perspective

This analysis outcome showed that the 2-cluster solution was the most stable and meaningful for this type of dataset. The results are new and have not been previously observed in this comparison in the field of machine learning and clustering. From an artificial intelligence perspective, the performance of the algorithms in producing better clusters can be examined through a comparison of the effect size and power index of each variable. One significant limitation noticed is that all variables cannot be clustered effectively by these algorithms, which results has no overlapping members and suboptimal clustering shapes. Thus only farthest first algorithm succeed to separate data correctly.

From the variance index perspective, which calculates the similarity of the cluster members and the differences of members from other clusters, even one shared member between clusters is considered. Otherwise, the p-value, which indicates the measurement error, will not be significant. Thus, though they all have high effect sizes, if the measurement error for all indicators is significant and less than 5 percent, they all may be desirable indicators. The non-significance of measurement error of a variable indicates that there are members common between clusters in the clusters made for that variable, stating that the clusters have not functioned well.

All of the methods were not able to achieve a significant level of accuracy with their clustering, but as suggested by Table 5(Appendix), Farthest First has effectively separated data, though still with low effect size. Therefore, this success of the algorithm may be due to handling abnormal data. This method sorts the data from the farthest point onwards, and typically, abnormal data shows high kurtosis and skewness because of dispersion, thereby having an advantageous merit over the methods based on central tendencies.

Moreover, regarding the Silhouette index, an internal clustering metric does not entirely agree with the Dwass-Steel-Critchlow-Fligner index results and shows a high value with the datasets' differentiation. The index gives a side view of how well each point is fitted in its cluster compared to the others. The range of this index is from -1 to 1:value 1 means perfect, well-separated clusters, value of 0 means indistinct, overlapping clusters [29]. The indexes observed show that all values are below 0.25, and according to common interpretation, this indication is that clusters have weak structure and are not very well separated. However, the fact that the top algorithms [K-Means, etc.] have scores around 0.21 while the poor one are much lower (~0.047 for K-Mean-Kernel) shows that some algorithms are way better at finding the underlying weak structure than others. In the provided case study, considering that Silhouette value has not exceeded 0.25, it may indicate that first, the range of variations in the numbers is rather small, and secondly, the repetition of numbers among the clusters is high.

## Statistical Perspective

Statistically, the highest effect size comes from the group; afterward comes K-Mean, K-Mean Fast, X-Mean, K-Mean H2O, Cascade K-Mean, Making Density, and Fuzzy C-Mean. Coming next, good methods include sIB, EM, Canopy, LVQ, and Gen clust Plus Plus. Farthest First, K-Mean-Kernel, and K-Medoid fall into the moderate group.

The current study also has highlighted the importance of a nonparametric Post-hoc test of Dwass-Steel-Critchlow-Fligner because previous studies revealed that the Kruskal-Wallis test cannot alone present the differences between groups, in case more than two independent groups across seven variables are taken into consideration. In such a case, when the Kruskal-Wallis test indicates significant difference between groups, pairwise comparisons may be conducted to specify the groups with differences, similar to the method used in one-way ANOVA, because reduction of Type I error risk is highly important in this context.

With this regard and considering the configurations which have been set in GPower software, the sample size was 460, and the obtained power was 0.96. The 0.96 of power means that there is 96% probability that the test will correctly detect an effect in differences among existing clusters. The high power shows that the test is sensitive and has a low risk of missing real differences. Thus, the GPower software has considerable capabilities to provide services for solving structural equation problems. As observed, the level of power in this research is higher than the computational level which is common and usual in most international scientific journals. About the analysis of the result, if we like to interpret them based on the science of statistics, the true success belongs to the genetic algorithm with an effect size of 0.68 on the variable attitude. In the following ranks, 'volume' and then K-Mean family group occupy the mean positions. These results could not be satisfying for data analysts because, in the real world and when the volume of data and also the number of variables are big, performing the test is practically impossible and is a risk.

Quality, aesthetics, perceived value, attitude, access, courtesy, Loyalty were the variables used for clustering purposes. The corresponding values represent effect sizes (E-S), which reflect the magnitude of each variable in outlining the clusters. A higher value means that the variable was a more significant discriminator for the groups. It does appear that the pattern found is remarkably consistent across the top methods: Courtesy (0.54 - 0.60) is the most critical driver. This variable differentiates the clusters better than any other variable. Perceived Value (0.29 - 0.43) is the second most significant driver, and aesthetics (0.04 - 0.06) is a reliable but weaker third driver. These findings suggest that for massage clients, Courtesy is the most critical driver. It reveals that a friendly respectful attentive approach to service is crucial in forming the customer experience and customer interface. It suggests that if the customer service is positive, respectful, and attentive, there is a high level of likelihood that the customer is pleased and will show loyalty to the brand. The second most important driver is perceived value. The perceived value is important because customers want to evaluate the product or service received against the price paid. When customers feel they are getting good value, they are likely to be loyal. Aesthetics ranks third but is weaker, which would suggest that while looks may attract them initially, it is not the most significant driver of customer retention or loyalty. Indeed, this would point to an area of improvement because increasing aesthetics can provide a more significant and positive feel towards the brand when married with solid courtesy and perceived value. Overall, based on these dynamics, there is substantial value that this can provide to businesses as they seek to optimize their approach to retain customers and engage customers.

Also, the P-value gives the statistical significance of the effect size of each variable. By closely analyzing the p-values, one can effectively determine the correct effect sizes. Variables with large effect sizes like politeness, perceived value, and aesthetics are more important than others when assessed in the 15 algorithms; on the contrary, a variable with a relatively smaller effect size is regarded as nonsignificant. This observation indeed proves that the identified trends are not a result of sheer luck.

Furthermore, The P-value reveals the significance of the effect sizes of the variables. A careful observation of the p-values reveals a true representation of the effect sizes of the variables. Variables with large effect sizes (such as politeness, perception of value, and aesthetics) are regarded as more significant variables among the 15 algorithms, while variables with small effect sizes are regarded as non-significant variables. This statement confirms that the observed trends are not random phenomena.

Atlast Using This research verifies that data clustering, especially for customers and their opinions, is not a trivial problem, whereas it is of great significance since, using mathematical algorithms and artificial intelligence, it can obtain very useful clusters with desired qualities for marketers and researchers. Furthermore, the division of customers into clusters using demographics and geographical divisions can no longer be applicable for the complex demands of marketing science. This research clearly implies a bright future for marketing that calls for researchers to investigate whether or not the data is normally distributed before proceeding with any analysis, that in statistics basically, there are two broad streams used for the aforesaid tests with their respective formulas, implications, and name: the parametric statistics, and the non-parametric statistics. This research clearly verifies that by using the non-normal data distributions, the furthest-first algorithm provides the cleanest data for clusters that can be further used by the broad public of marketers. This is since, when a researcher is dealing with more than one hundred variables, along with millions of data about the customers, trying each of these approaches is tiresome, time-wasting, and impossible.

## Conclusion

results emerges that the effectiveness of different algorithms for data clustering as far as identifying significant features of customer perception is concerned, with the 'Farthest-First' algorithm being unique in its approach to data clustering. The variability of effects of courtesy, value, aesthetics, overlapping membership, optimum shape of clusters, and other factors pertaining to marketing data analysis reveals the importance of more sophisticated approaches in improving customer engagement and retention. Successful measurement of significant variables in marketing data analysis is critical in realising the complexities involved in such data analysis.

## Limitations of other Approaches

The study revealed that several algorithms were fundamentally unsuitable for Likert scale data: CLOPE, SOM, Cobweb, DBSCAN, and Random methods produced non-significant results across all indices and were deemed inappropriate.

Hierarchical and agglomerative algorithms produced imbalanced and unacceptable clustering outputs.

The other limitations of this study are, first, the algorithms that were not used in this review. Second, various combined comparisons can be made by integrating different types of binary, categorical, and continuous data.

## Declarations and Confilict Statement

Author guarantee that his article is devoid of any competing interests and there is no conflict of interests.

## References


1. Rubarth, K., Pauly, M., & Konietschke, F. (2022). Ranking procedures for repeated measures designs with missing data: estimation, testing and asymptotic theory. Statistical Methods in Medical Research, 31(1), 105-118.
2. D'Urso, P., Disegna, M., Massari, R., & Prayag, G. (2015). Bagged fuzzy clustering for fuzzy data: An application to a tourism market. Knowledge-Based Systems, 73, 335-346.
3. Kotler, P. (2020) Principles of marketing. 18th edn. Harlow, England: Pearson.
4. Pitts, B.G. and Stotlar, D.K. (2013) Fundamentals of sport marketing. 4th edn. Morgantown, W.Va: Fitness Information Technology (Sport management library). Available at: (Accessed: 23 October 2024).
5. Rodriguez, M. Z., Comin, C. H., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. D. F., & Rodrigues, F. A. (2019). Clustering algorithms: A comparative approach. PloS one, 14(1), e0210236.
6. Hennig, C. (2022). An empirical comparison and characterisation of nine popular clustering methods. Advances in Data Analysis and Classification, 16(1), 201-229.
7. Kaya, M.-F. and Schoop, M. (2022) 'Analytical comparison of clustering techniques for the recognition of communication patterns', Group Decision and Negotiation, 31(3), pp. 555-589. Available at: (Accessed: 23 October 2024).
8. Costa, E., Papatsouma, I., & Markos, A. (2023). Benchmarking distance-based partitioning methods for mixed-type data: E. Costa et al. Advances in Data Analysis and Classification, 17(3), 701-724.
9. Sepin, P. et al. (2024) 'Comparison of clustering algorithms for statistical features of vibration data sets', in Data Science—Analytics and Applications. Cham: Springer Nature Switzerland, pp. 3-11. Available at: (Accessed: 23 October 2024).
10. Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods, 39(2), 175-191.
11. Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. Behavior research methods, 41(4), 1149-1160.
12. Sánchez-Fernández, R., Iniesta-Bonillo, M. A., & Holbrook, M. B. (2009). The conceptualisation and measurement of consumer value in services. International Journal of Market Research, 51(1), 1-17.
13. Diallo, M. F., & Seck, A. M. (2018). How store service quality affects attitude toward store brands in emerging countries: Effects of brand cues and the cultural context. Journal of Business Research, 86, 311-320.
14. Soares-Silva, D., de Moraes, G. H. S. M., Cappellozza, A., & Morini, C. (2020). Explaining library user loyalty through perceived service quality: What is wrong?. Journal of the Association for Information Science and Technology, 71(8), 954-967.
15. Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial least squares structural equation modeling (PLS-SEM) using R: A workbook (p. 197). Springer Nature.
16. Thorpe, M. G., Milte, C. M., Crawford, D., & McNaughton, S. A. (2016). A comparison of the dietary patterns derived by principal component analysis and cluster analysis in older Australians. International Journal of Behavioral Nutrition and Physical Activity, 13(1), 30.
17. Kamath, A., Poojari, S., & Varsha, K. (2025). Assessing the robustness of normality tests under varying skewness and kurtosis: a practical checklist for public health researchers. BMC Medical Research Methodology, 25(1), 206.
18. He, F. et al. (2017) "Non-parametric MANOVA approaches for non-normal multivariate outcomes with missing values," Communications in Statistics - Theory and Methods [Preprint]. Available at: (Accessed: November 14, 2025).
19. Elliott, A. C., & Hynan, L. S. (2011). A SAS® macro implementation of a multiple comparison post hoc test for a Kruskal–Wallis analysis. Computer methods and programs in biomedicine, 102(1), 75-80.
20. Quintavalle Pastorino, G., Preziosi, R., Faustini, M., Curone, G., Albertini, M., Nicoll, D., ... & Mazzola, S. (2019). Comparative personality traits assessment of three species of communally housed captive penguins. Animals, 9(6),

376.
21. Spurrier, J. D. (2006). Additional tables for Steel–Dwass–Critchlow–Fligner distribution-free multiple comparisons of three treatments. Communications in Statistics-Simulation and Computation, 35(2), 441-446.
22. Lengyel, A., & Botta-Dukát, Z. (2019). Silhouette width using generalized mean—A flexible method for assessing clustering efficiency. Ecology and evolution, 9(23), 13231-13243.
23. Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. Behavior research methods, 37(3), 379-384.
24. Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. Journal of experimental psychology: General, 141(1), 2.
25. Sarstedt, M., & Mooi, E. (2018). Principal component and factor analysis. In A concise guide to market research: The Process, Data, and Methods using IBM SPSS Statistics (pp. 257-299). Berlin, Heidelberg: Springer Berlin Heidelberg.
26. Dolgun, A., & Demirhan, H. (2017). Performance of nonparametric multiple comparison tests under heteroscedasticity, dependency, and skewed error distribution. Communications in Statistics-Simulation and Computation, 46(7), 5166-5183.
27. Buchan, I.E. (2025) Kruskal-Wallis Test (Nonparametric One-way ANOVA) - StatsDirect. Available at:  (Accessed: 7 November 2025).
28. Fagerland, M. W., Sandvik, L., & Mowinckel, P. (2011). Parametric methods outperformed non-parametric methods in comparisons of discrete numerical variables. BMC Medical Research Methodology, 11(1), 44.
29. Vergara, V. M., Abrol, A., Espinoza, F. A., & Calhoun, V. D. (2019, July). Selection of efficient clustering index to estimate the number of dynamic brain states from functional network connectivity. In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 632-635). IEEE.
30. Dolnicar, S., Grün, B., & Leisch, F. (2018). Market segmentation analysis. In Market segmentation analysis: Understanding it, doing it, and making it useful (pp. 11-22). Singapore: Springer Singapore.
31. Li, Y. Z., Shen, J. H., Zhang, W. X., Zhang, K. Q., Peng, Z. H., & Huang, M. (2023). Slope deformation partitioning and monitoring points optimization based on cluster analysis. Journal of Mountain Science, 20(8), 2405-2421.
32. Zangana, H. M., & Abdulazeez, A. M. (2023). Developed clustering algorithms for engineering applications: A review. International Journal of Informatics, Information System and Computer Engineering (INJIISCOM), 4(2), 160-182.