



Journal of Advanced Research in Social and Behavioural Sciences

Journal homepage:
<https://karyailham.com.my/index.php/jarsbs/index>
ISSN: 2462-1951



A Survey on Factors Influencing Coffee Shops Preferences Among the Universiti Tun Hussein Onn Malaysia (UTHM) Pagoh Students

Armin Hafiz Rosromaizi¹, Norhaidah Mohd Asrah^{1,*}, Rohayu Mohd Salleh¹, Kiaw Tong Yun¹

¹ Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, UTHM Kampus Pagoh, Pagoh Higher Education Hub, KM 1, Jalan Panchor, 84600 Panchor, Johor, Malaysia

ARTICLE INFO

Article history:

Received 14 July 2026

Received in revised form 7 September 2026

Accepted 7 September 2026

Available online 8 September 2026

Keywords:

Coffee shop; Word cloud analysis; Factor analysis; Correspondence Analysis; Product Value

ABSTRACT

The culture of coffee shops or cafés has become very popular in Malaysia, and the trend seems to have changed coffee drinking into a lifestyle among university students who use coffee shops for studying, relaxing and socializing. This study aims to analyse data collected via a survey administered to 264 Universiti Tun Hussein Onn Malaysia (UTHM), Pagoh students from different faculties via Google Forms. This study was designed to gain insight into local coffee shops through word cloud analysis, determine the factors that influence the choice of coffee shops through factor analysis, and establish the relationships between factors and perceptions of two brands of coffee shops, A and B, through correspondence analysis. Word cloud analysis provided additional information beyond others, mainly highlighting location, high prices, poor service, and overcrowding as the main determinants of respondents' choice of coffee shops. Four main factors were derived from factor analysis: Brand Performance, Product Value, Service Quality, and Customer Engagement; however, Product Value and Service Quality had the highest loadings. From the correspondence analysis, coffee shop A was found to have a stronger association with tastes, promotions, product quality, popularity, and brand performance, while coffee shop B was found to have a stronger association with service quality and convenience, with overall perceptions showing only slight differences between the two brands.

1. Introduction

The coffee shop scene in Malaysia has seen significant changes in recent times, with coffee drinking becoming a way of life for many university students. Today, coffee shops have become popular hangout spots for studying, socializing, and relaxing, among other activities, reflecting a trend towards experiential and social consumption. Although there has been tremendous growth in this industry, many Small and Medium-sized Enterprises (SMEs), such as coffee shops, lack adequate knowledge of what drives consumer preferences, especially among university students.

* Corresponding author.

E-mail address: norhaida@uthm.edu.my

<https://doi.org/10.37934/arsbs.42.1.417431>

In Malaysia, SMEs such as coffee shops experience high business failure rates because they fail to understand the factors that affect consumer behaviour and preferences [2]. As such, marketing becomes difficult due to poor decision-making in critical areas, such as pricing and service quality. It is thus important to identify the major factors affecting customer choice and preference in coffee shop selection to enhance business success.

This study focused on UTHM Pagoh students and aimed to explore their preferences for coffee shops. The purpose of this study is to explore deeper insights into coffee shop preferences through word cloud analysis, to determine the determinants of coffee shop selection through factor analysis, and to examine the relationships between the brand factors mentioned above and brand perception through correspondence analysis. The data were collected through an online survey among some undergraduate students.

1.1 Literature Review

Previous studies have identified several important factors that influence customers' choice of coffee shops. Taste has been recognised as an important determinant of customer satisfaction and loyalty, as customers' preferences for coffee shops are closely associated with the perceived quality and sensory experience of coffee [9]. Price and perceived value are also important considerations that may influence purchasing decisions and visit frequency [10]. Other factors commonly associated with coffee shop selection include product quality, service quality, convenience, ambience, promotional activities and brand popularity. Collectively, these factors suggest that coffee shop preferences are multidimensional and may reflect a combination of product-related, economic, service-related and experiential considerations.

Methodologically, three analytical techniques are employed in this study to provide complementary perspectives on students' coffee shop preferences. First, word cloud analysis is used to identify prominent terms and emerging themes from responses to an open-ended question. Word clouds provide a visual representation of frequently occurring words within qualitative data, facilitating an initial understanding of dominant themes and patterns [4]. Second, factor analysis is employed to reduce a larger set of observed variables into a smaller number of underlying latent factors, thereby identifying the principal dimensions influencing students' decisions to visit coffee shops [11]. This technique is particularly useful for examining complex consumer preference data in which multiple observed variables may reflect common underlying constructs. Third, correspondence analysis is used to visually examine relationships among categorical variables. This technique simplifies complex categorical data and facilitates the identification and interpretation of associations that may not be immediately apparent through conventional cross-tabulation or frequency analysis [5].

The significance of this study lies in providing useful insights that can assist coffee shop owners and managers in better understanding students' preferences and the factors influencing their choice of coffee shops. By integrating word cloud analysis, factor analysis and correspondence analysis, the study provides a more comprehensive understanding of students' coffee shop preferences in Malaysia from both exploratory and multivariate perspectives. The findings may support more informed decision-making in areas such as product offerings, pricing, service quality, customer experience and marketing strategies. To maintain confidentiality, preserve research neutrality and protect the reputations of the participating businesses, the actual names of the coffee shops are anonymised throughout the study. Accordingly, the two participating establishments are referred to as Coffee Shop A and Coffee Shop B in all analyses, results and discussions.

2. Methodology

2.1 Data Collection

Data collection is the systematic process of acquiring information for research purposes. The data in this study were collected from the Universiti Tun Hussein Onn Malaysia (UTHM) Pagoh campus. The three faculties at the UTHM Pagoh campus are the Centre for Diploma Studies (CeDS), the Faculty of Applied Sciences and Technology (FAST), and the Faculty of Engineering Technology (FTK). There are 5,627 students at UTHM Pagoh, as per the records of the Academic Management Office (AMO) for 2025. According to Krejcie & Morgan [12], the minimum sample size should be 357.

Data collection was conducted by administering an online questionnaire via WhatsApp, Telegram, Instagram, and Facebook from March to April 2026. This questionnaire had sections on demographics, coffee consumption habits, factors affecting coffee consumption, and perceptions of the coffee shop A and coffee shop B brands.

2.2 Reliability Test

A reliability test was done to determine whether the questionnaire reliably measures the construct that is expected to be measured. In this study, Cronbach's Alpha was used to evaluate the consistency of items that are measured for coffee shop preference factors. The higher the Alpha value, the more reliable the items become. Usually, an Alpha value greater than 0.7 is considered acceptable [15]. The formula for calculating Cronbach's Alpha, α is given as in Eq. (1):

$$\alpha = \frac{k}{k-1} \left(1 - \frac{\sum v_i}{v_t} \right) \quad (1)$$

where k denotes the number of questions, v_i denotes the variance of each separate question, and v_t denotes the variance of the total score. This equation calculates the correlation between the set of questions, thereby measuring the degree of internal consistency among them.

2.3 Descriptive Analysis

Descriptive analysis is used to summarize and describe the fundamental features of the gathered data. In the current case, it is utilized for establishing the consumption pattern of respondents related to their choice of coffee shops. The approach provides information about fundamental distributions, including coffee consumers and non-consumers, as well as major factors determining their preferences.

2.4 Word Cloud Analysis

A word cloud is one of the text visualization methods used to analyse the most frequent terms derived from open-ended answers. In this case, WordCloud.com was chosen as the visual software used in the analysis of the important themes in the students' responses. This method enables vast amounts of text to be visually summarized, with the most frequently mentioned terms standing out. It helps to identify the important themes that cannot be derived from closed-ended questions [7].

2.5 Factor Analysis

Factor analysis is a multivariate statistical tool employed to simplify the data structure by clustering the related variables in a few latent factors. In this research, the EFA tool is used to examine the relationships among various latent factors and variables that determine students' selection of a coffee shop. This involves evaluating the data suitability through KMO and Bartlett's test, factor extraction, varimax rotation, and interpretation of the factor loadings [14]. The formula for the factor analysis model is given as in Eq. (2):

$$y_j = \lambda_{j1}f_1 + \lambda_{j2}f_2 \dots + \lambda_{jm}f_m + \varepsilon_m \quad (2)$$

where y_j represents the observed variable, λ_{jm} is the factor loading, f_m represents the latent factors, and ε_j is the error term. This model explains how underlying latent factors and an associated error term influence observed variables.

2.6 Correspondence Analysis

Correspondence analysis is one of the methods that can be used in exploratory research with categorical variables to detect their relations and to provide a visual representation. The method assumes that the input data are organized in a contingency table, representing the frequency or percentage distributions of categories for which associations need to be determined. It simplifies the association visualization process by decomposing high-dimensional space into a lower dimensionality, where relations become clearer. The algorithm starts by conducting the Chi-square test to detect the relation between the analysed factors. If there is no relationship, expected frequencies are calculated under the assumption of independence between variables. Then row and column profiles based on the marginal total are calculated to describe the distribution of categories.

Subsequently, Singular Value Decomposition (SVD) is performed to decompose the contingency table and derive the dimensions used to represent the relationships among categories. The resulting structure can then be visualised using a biplot, which provides a graphical representation of the relationships among variables and categories. In interpreting a biplot, the relative positions, distances and directions of points provide useful information about patterns of association. Points positioned relatively close to one another generally indicate greater similarity or association, whereas points located farther apart suggest weaker similarity or different response patterns [13].

In the biplot analysis, the four quadrants provide a useful visual guide for examining the distribution of coffee shops and their associated attributes. However, the quadrants themselves should not be interpreted as direct measures of the strength of association. Instead, interpretation should primarily consider the relative positions and distances among points in the multidimensional representation [3]. Points located close to one another indicate similar profiles or stronger associations, whereas those positioned farther apart indicate greater differences in their profiles. Therefore, although the quadrants can facilitate the visual identification of clusters and patterns, the interpretation of the correspondence analysis should be based principally on the relative proximity and spatial configuration of the points rather than their quadrant locations alone [3].

3. Result and Discussion

3.1 Preliminary Analysis

A pilot test was conducted to test the reliability of the questionnaire using statistical software, SPSS. The questionnaires were distributed to 30 respondents to obtain responses from different faculties: FAST, FTK, and CeDS. The data collection was done through a questionnaire, which was sent to respondents via WhatsApp. Based on Table 1, the Cronbach's alpha for the reliability test obtained a value of 0.938. This shows that the questionnaire is highly reliable for this study and can proceed with the actual study and further analysis.

Table 1
Output Cronbach Alpha Value

Cronbach's Alpha	Cronbach's Alpha based on standardized items	N of items
0.938	0.940	40

3.2 Descriptive Analysis

The descriptive analysis was used to summarize and analyse demographic information of 264 respondents. Moreover, some visualization methods were used to present the gathered information in charts and graphs. Microsoft Excel was used to create a graph for better understanding and analysis of information.

3.3 Demographic Profile

Table 2 shows the demographic profile of the respondents. Most respondents were females (65%), while males constituted 35% of the total participants. Most respondents were Malay students (63%), followed by Chinese (24%) and Indians (11%). Regarding age, most respondents were in the 21-22 years age group (46%), while 23-24 years accounted for 36%. Most respondents came from FAST (64%) and were degree students (83%). Year 3 was the leading year group among respondents (38%), followed by Year 1 (29%).

Table 2
Demographic profile

No.	Demographic	Frequency
1	Gender	Female: (65%), Male: (35%)
2	Race	Malay: (63%), Chinese: (24%), Indian: (11%), Sikh: (1%), Bumiputera Sabah: (1%), Bidayuh: (0.38%)
3	Age	21-22: (46%), 23-24: (36%), 19-20: (9%), above 25: (9%)
4	Faculty	FAST: (64%), FTK: (19%), CeDS: (17%)
5	Level of Study	Degree: (83%), Diploma: (17%)
6	Year of Study	Year 3: (38%), Year 1: (29%), Year 2: (23%), Year 4: (10%)

3.4 Coffee Drinker Status

The number of coffee drinkers and non-drinkers is presented in Figure 1(a). Of the 264 respondents, 164 indicated they are coffee consumers, while 100 indicated they are non-consumers, suggesting that more than half are coffee drinkers. These findings indicate that the practice of coffee drinking is widespread among the students of UTHM Pagoh, and hence, the collected information is appropriate for further study regarding the choice of coffee shops. On the other hand, Figure 1(b) depicts the distribution of coffee consumers by gender. Among the female respondents, 100 were coffee drinkers, while 72 were non-drinkers. Among male respondents, 64 were coffee consumers, while 28 were non-consumers. These findings indicate that coffee drinking behaviour is prevalent among both genders.

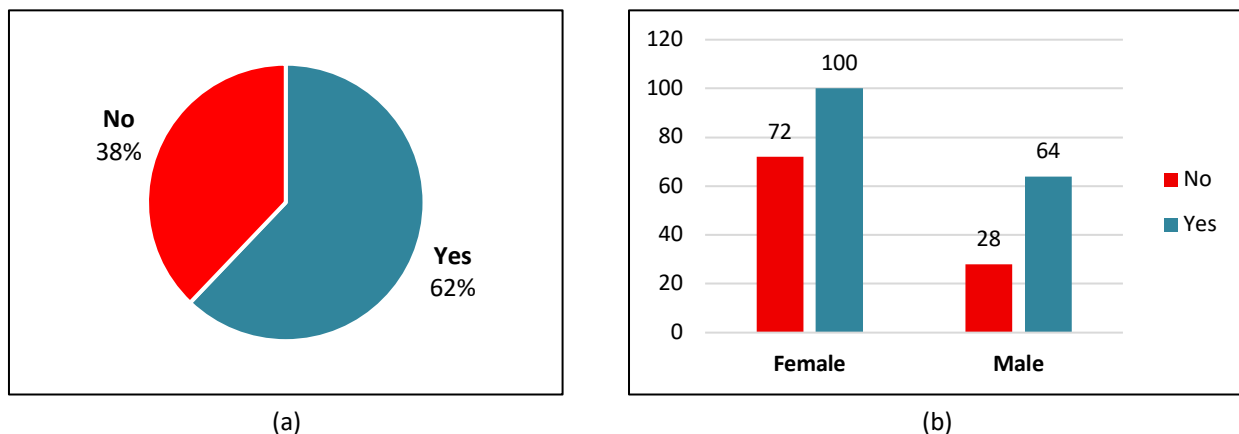
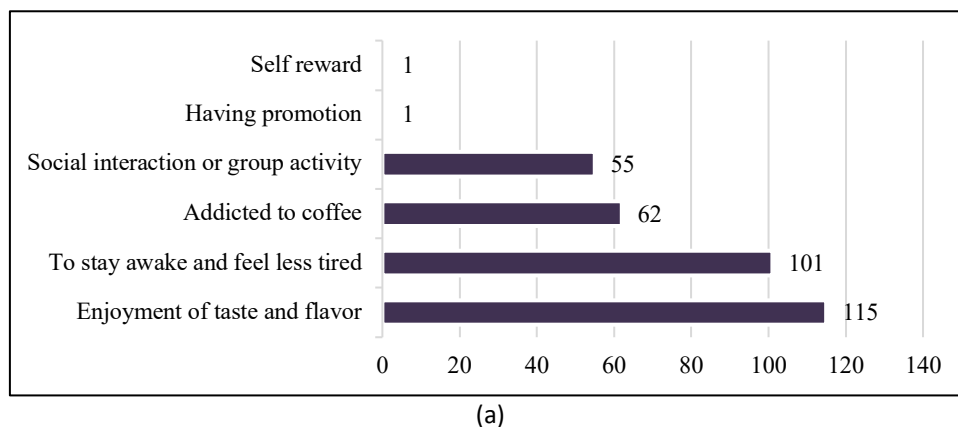


Fig. 1. (a) Coffee Drinker Status (b) Coffee Drinker Status between Males and Females

3.5 Reasons for Coffee Consumption and Non-Consumption

The reasons for respondents' decisions to consume and not consume coffee are shown in Figure 2(a) and (b), respectively. Among the 164 coffee drinkers, 115 respondents chose to drink coffee for its taste and flavour. Apart from that, 110 respondents chose coffee to stay awake and overcome tiredness, and 62 respondents claimed that their coffee consumption was due to addiction. Social activity has also been listed as a reason by 55 respondents. Promotion and self-reward reasons were provided by a few respondents who drink coffee.



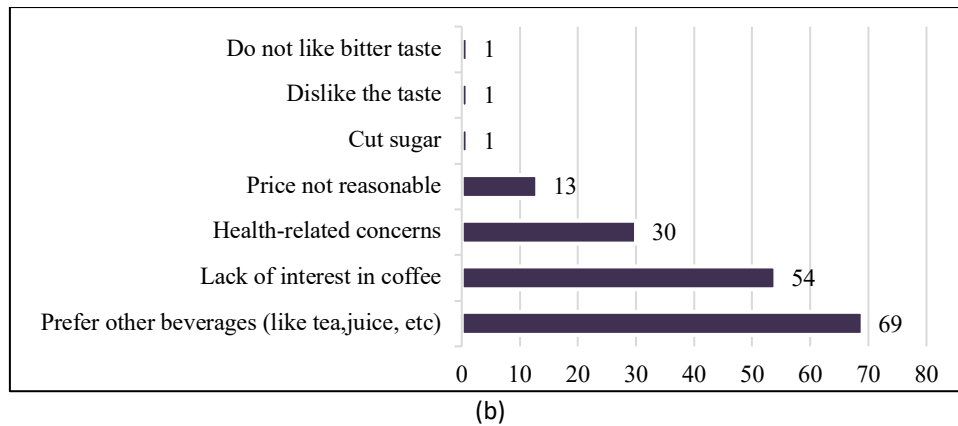


Fig. 2. (a) Reasons Respondents Drink Coffee (b) Reasons Respondents Do Not Drink Coffee

As for the 100 non-coffee consumers, 69 preferred other beverages such as tea, juice, and soft drinks, while 54 claimed they did not like coffee. Moreover, the health concern was answered by 30 respondents, among whom 13 thought coffee prices were not reasonable.

3.6 Respondents' Preferred Coffee Menus

The types of coffee preferred by the 164 respondents are presented in Figure 3, where respondents were permitted to select more than one option. Latte is the most selected type of coffee with 108 responses, followed by caramel latte (85), frappe and mocha (77). It implies that the respondents like their coffee to be sweetened and flavoured. On the contrary, conventional coffee types, such as cappuccino (59), Americano (57), and macchiato (56), have gained only average popularity. The flat white coffee type has received 31 responses. In addition, several other preferences were expressed by a few respondents, such as Spanish latte, coconut latte, and iced black coffee.

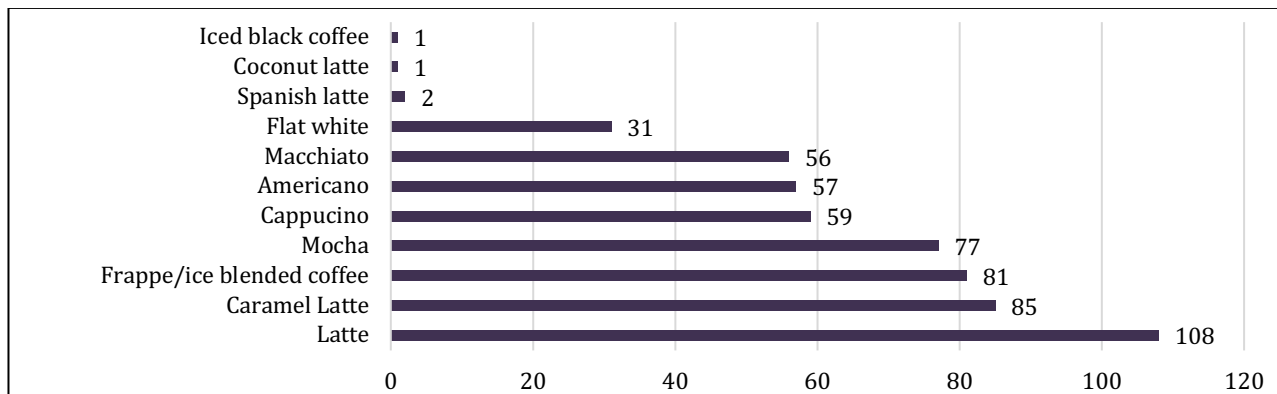


Fig. 3. Type of Coffee Menus Respondents Like

3.7 Frequency of Coffee Consumption

The frequency of coffee consumption among the respondents is shown in Figure 4 (a). Most of the respondents were found to have coffee a few times a week (64 respondents), followed by a few times a month (47 respondents), and once a week (26 respondents). Meanwhile, only 16 respondents had coffee every day, while 11 had it only a few times a year. This reveals that most respondents consume moderate amounts of coffee and are not heavy coffee drinkers.

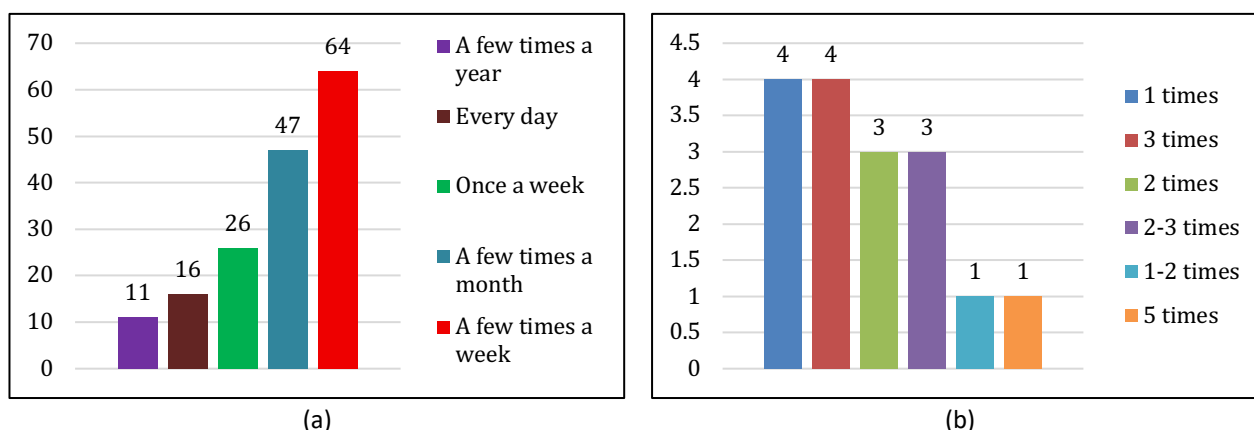


Fig. 4. (a) Coffee Consumption Frequency (b) Daily Coffee Consumption Frequency

On the other hand, the chart in Figure 4 (b) displays the daily coffee consumption frequency among the 16 respondents who have coffee every day. Most respondents were found to consume coffee one to three times per day, and one respondent had five cups of coffee per day. The consumption of 4 to 5 cups of coffee per day (approximately 400 mg of caffeine) is regarded as the limit of caffeine dependence [1]. Hence, the findings reveal that there is no excessive risk of experiencing caffeine dependence among the respondents due to their moderate consumption levels.

3.8 Factor Analysis

Before applying factor analysis, the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy and Bartlett's Test of Sphericity were conducted to assess whether the data were suitable for factor analysis. The KMO measure assesses whether there is sufficient common variance among the variables for factor analysis. It is considered appropriate when the KMO measure is greater than 0.60. According to Table 3, the KMO value in this case was 0.903.

Table 3

KMO test value

	Value
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.903

On the other hand, Bartlett's Test of Sphericity tests whether there are correlations between the variables. In this test, it is appropriate when the p -value is less than 0.05. The result from Bartlett's Test of Sphericity in this case Table 4, was 0.001. From this result, note that it is less than 0.05 and hence reject the null hypothesis that it is an identity matrix.

H_0 : The correlation matrix is an identity matrix (there are no correlations between variables).

H_1 : The correlation matrix is not an identity matrix (there are correlations between variables).

Table 4

Bartlett's test of sphericity p -value

	p -value
Bartlett's test of sphericity	0.001

Table 5 highlights the Rotated Component Matrix generated by factor analysis. Factor loading is a measure of the association between each variable and its component; values above 0.50 are considered significant. For ease of interpretation, the Varimax rotation technique was used to generate the rotated component matrix, through which high factor loadings were maximized within one component and low loadings minimized among other components. As per the results shown in Table 5, four components were effectively extracted.

Based on Table 6, the first component identified was that of "Brand Performance," which comprised variables related to brand recognition, appealing drink appearance, efficient service, cleanliness, and social media marketing. The second component, named "Product Value", had variables related to fair price, value, quality, reliability, and taste of drinks. The third component of "Service Quality", consisted of variables related to friendly employees, responsiveness, a comfortable ambiance, and convenient online ordering. Finally, Component 4 was labelled "Customer Engagement" since it included variables concerned with promotions, campaigns, Wi-Fi availability, customer interaction, and interior design. Thus, the factor analysis helped categorize the selected variables into four components that represent students' coffee shop preferences.

Table 5
The rotated component matrix

		Components			
		1	2	3	4
1	Drinks are presented in an appealing way	0.727			
2	The coffee shop is well known among customers	0.805			
3	The coffee shop is clean and hygienic	0.663			
4	Orders are served fast and quickly	0.693		0.375	
5	The coffee shop provides Wi-Fi for customers				0.573
6	The price matches the quality of the products		0.752		
7	Staff help when needed			0.668	
8	The drinks provide good value for money		0.735		
9	Ingredients used in drinks are of high quality		0.556	0.562	
10	The coffee shop has a relaxing environment			0.638	0.500
11	Discounts, loyalty programs or special offers encourage customers				0.664
12	The coffee taste is consistent every time		0.458	0.542	
13	The coffee shop has good feedback from customers		0.450	0.411	0.554
14	Staff are friendly			0.704	
15	Ordering through apps or online platforms is easy	0.461		0.546	
16	The taste follows the menu description	0.439	0.479	0.494	
17	The coffee shop always advertises promotions through social media	0.674			
18	The coffee shop promotions are interesting and encouraging				0.621
19	The drinks on every menu taste good		0.566		0.543
20	The coffee shop is widely recognized and trusted	0.524			
21	Payment methods are convenient and fast			0.424	0.481
22	The coffee shop's interior design is visually appealing and suitable	0.482			0.523
23	The prices are reasonable		0.788		
24	The coffeeshop offers consistent product quality over time	0.478	0.540		

Table 6

Category of Factors in Components

Component	Factor
1	Brand Performance
2	Product Value
3	Service Quality
4	Customer Engagement

3.9 Correspondence Analysis

The correspondence analysis method (CA) was employed in investigating the relationship between attributes of coffee shops and the perceptions of respondents. Because the use of the Brand $\times$ Factor matrix with only two brands could offer limited dimensional information, the data was restructured into a 16×5 contingency table including brand-factor pairs (such as Taste A, Price B) and five levels of perceptions on a Likert scale from Strongly Disagree to Strongly Agree.

3.10 Chi Square Test

A Chi-Square test of independence was conducted before the Correspondence Analysis to determine whether there exists an association between the combinations of brand and attributes and perception levels. This is because Correspondence Analysis is based on Chi-Square statistics. The test was conducted using a contingency table consisting of frequencies of brand-attribute combinations for each of the five Likert scale perception levels, as in Table 7.

Table 7

Contingency Table Brand Factor and Respondents' Perception

Brand Factors	Perception Level				
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Taste A	11	16	19	51	67
Price A	8	15	41	61	39
Promotion A	10	12	20	58	64
Quality A	7	9	27	64	57
Service A	10	17	18	64	55
Environment A	4	10	19	68	63
Popularity A	7	10	18	46	83
Convenience A	8	11	19	51	75
Taste B	8	16	59	64	17
Price B	13	26	59	50	16
Promotion B	6	37	66	40	15
Quality B	8	19	59	56	22
Service B	13	16	49	60	26
Environment B	4	19	46	54	41
Popularity B	8	19	65	44	28
Convenience B	11	27	46	49	31
Total	136	279	630	880	699

Chi-square test of independence was done to find out if there was any statistically significant relationship between combinations of brands and factors and perception levels. The results obtained

from the Chi-Square test in Table 8 were used to justify the use of Correspondence analysis (CA). The following hypotheses were tested:

H_0 : There is no significant relationship between combinations of brands factor and perception levels.

H_1 : There is a significant relationship between combinations of brands factor and perception levels.

Table 8

Chi-Square test of independence from SPSS output

Test	Value	Degree of Freedom	<i>p</i> -value
Pearson Chi-Square	390.977	60	0.000

From the Chi-Square output analysis, the Chi-Square value equals 390.977 with 60 degrees of freedom, while the *p*-value equals 0.000. The null hypothesis is therefore rejected since the *p*-value is less than the 0.05 significance level. It can be concluded that there is enough evidence to suggest an association between the brand factor and perception exists.

3.11 Interpretation of Correspondence Analysis Biplot

Correspondence Analysis (CA) was done to see the relationship of brand factor combinations and their perceptions. CA involves breaking down the total inertia (0.149), which measures the total relationship of the contingency table. The following Table 9 gives the percentage of inertia associated with each dimension.

Based on Table 9 in terms of inertia analysis, Dimension 1 explained 85.1% of the total inertia and is the principal dimension that explained the relationship between the brand-factors associations and perceptions. The second dimension, on its part, explained 8.0% of the total inertia, thus summing up to 93.1% of the total inertia for the first two dimensions. Since the first two dimensions were responsible for explaining most of the total inertia, a two-dimensional model sufficed in analysing the category relationships. On the other hand, Dimension 3 and Dimension 4 contributed collectively 6.9% to the total inertia, showing little significance in explaining the total inertia relationship.

Table 9

Correspondence Analysis Summary

Dimension	Inertia	Proportion of Explained Variance	Cumulative Proportion
1	0.127	85.1%	85.1%
2	0.012	8.0%	93.1%
Other Dimensions (3 and 4)	0.010	6.9%	100%
Active total	0.149	100%	-

The CA biplot of brand-factor combinations and perception levels is shown in Figure 5. Points' closeness in this figure represents the association structure among categories. For coffee shop A, factors Popularity A, Convenience A, and Taste A are located close to the "Strongly Agree" category, while Quality A, Environment A, Promotion A, and Service A are located near the "Agree" category. On the other hand, Price A is located between positive and negative perception categories, showing a weak association with positive perception levels.

For coffee shop B, Price B, Popularity B, and Quality B are located close to the "Neutral" category. Furthermore, Taste B and Service B are located distant from the cluster of positive perception categories, while Environment B, Convenience B, and Promotion B are located close to the negative

perception categories. From this CA biplot, coffee shop A factors have an association with positive perception levels compared to coffee shop B factors, which have an association with neutral and even negative perception levels. There is a relation between brand-factor combinations and perception levels.

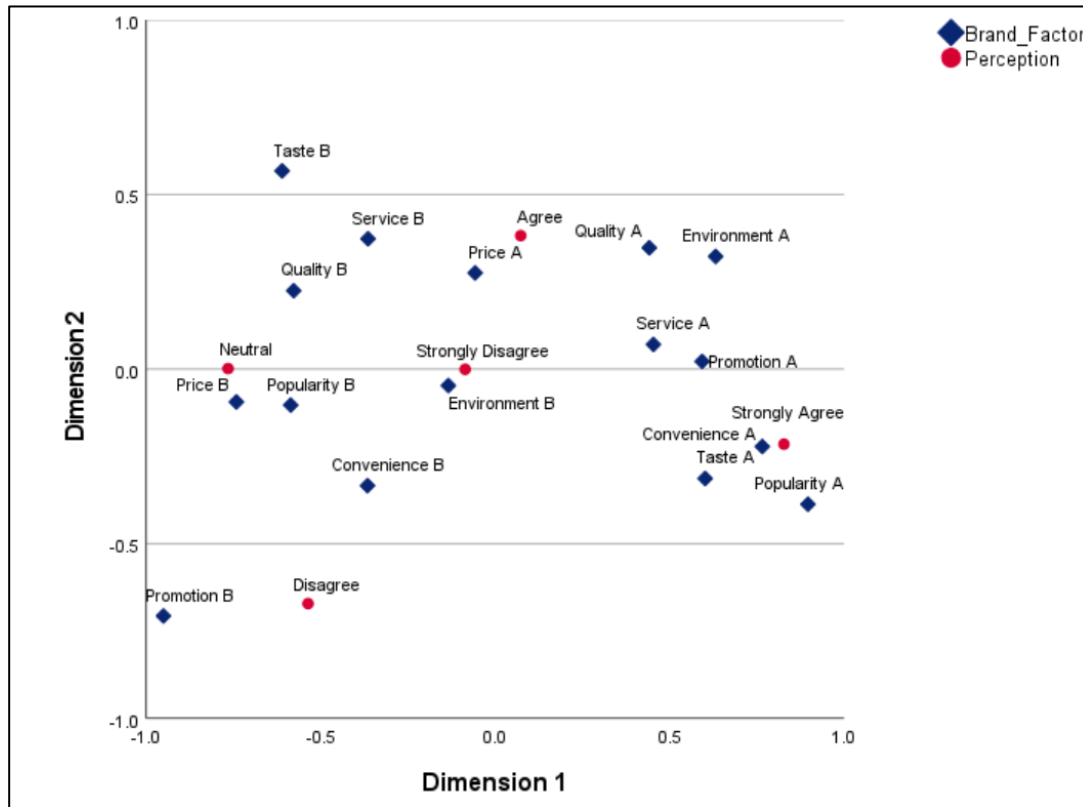


Fig. 5. Correspondence Analysis Biplot

3.12 Word Cloud Analysis

A word cloud analysis was conducted to extract additional information from the respondents' free-response answers regarding their preferences for coffee shops. In contrast to factor analysis and correspondence analysis, word cloud analysis enables respondents to express their views freely and reveals themes that may not emerge from questionnaires. The word sizes in the word cloud depict the frequency with which a particular factor or view is mentioned.

Figure 6 identifies the key factors that determine coffee shop choice, with price, taste, convenience and location among the most important criteria.



Fig. 6. Word cloud analysis of the main factor influencing respondents' choice of coffee shop

Figure 7 (a) illustrates the primary reasons customers are discouraged from visiting the coffee shop, including higher costs, poor service, and overcrowding. Also, Figure 7(b) depicts several changes recommended by the respondents to encourage more visits. The proposed changes include greater affordability, tastier, higher-quality coffee, and improved atmosphere and facilities. Thus, the word cloud analysis successfully provided additional information overlooked in the factor and correspondence analyses.



(a)



(b)

Fig. 7. (a) Word cloud analysis of the main factors that make respondents uninterested in visiting a coffee shop and (b) Word cloud analysis of respondents' suggestions on changes or additions that would encourage them to visit a coffee shop more often

4. Conclusions

This study successfully employed word cloud analysis, factor analysis and correspondence analysis to examine students' preferences for coffee shops at UTHM Pagoh. The word cloud analysis provided complementary insights beyond those obtained from the structured questionnaire, identifying crowdedness as a key deterrent and location as an important but previously unmeasured factor influencing coffee shop selection. These findings demonstrate the usefulness of word cloud analysis in summarising textual information and identifying frequently occurring terms and emerging themes [8]. Future studies should therefore consider incorporating location and crowdedness as explicit variables when examining coffee shop preferences.

Factor analysis identified four key dimensions influencing students' preferences: Brand Performance, Product Value, Service Quality and Customer Engagement. These findings indicate that coffee shop preference is a multidimensional construct shaped by several underlying factors. Factor analysis is particularly appropriate in this context because it enables a larger set of observed variables to be reduced into a smaller number of underlying dimensions, thereby facilitating the interpretation of complex patterns within the data [16].

Correspondence analysis further revealed distinct patterns in students' perceptions of the two coffee shops. Coffee Shop A was more closely associated with positive evaluations, whereas Coffee Shop B was associated with relatively more neutral and negative perceptions. These findings illustrate the usefulness of correspondence analysis in visually examining relationships between brands and customer perceptions and in providing insights that may support stronger brand positioning [6]. From a managerial perspective, coffee shop operators can use such information to identify attributes that contribute positively or negatively to their brand image and subsequently improve their product offerings, service quality and customer engagement strategies.

Despite these contributions, the study has several limitations. First, the sample was restricted to students from a single campus, which limits the generalisability of the findings to broader student or consumer populations. Second, the use of convenience sampling may have introduced selection bias, while relatively low respondent engagement may have affected the reliability and representativeness of the data. Third, the correspondence analysis included only two coffee shop brands, limiting the diversity of brand comparisons. Future studies should therefore involve multiple campuses, employ probability-based sampling techniques, combine multiple data collection methods and include a wider range of coffee shop brands. These improvements would enhance the representativeness of the sample, strengthen the robustness of the analyses and provide a more comprehensive understanding of consumer preferences in the coffee shop market.

Acknowledgement

The authors would like to thank the Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, for its support. This research was not funded by any grant

References

- [1] Ágoston, Csilla, Róbert Urbán, Michael J. Richman, and Zsolt Demetrovics. 2018. "Caffeine Use Disorder: An Item-Response Theory Analysis of Proposed DSM-5 Criteria." *Addictive Behaviors* 81: 109–116. <https://www.sciencedirect.com/science/article/abs/pii/S0306460318300753>.
- [2] Ambad, Sylvia Nabila Azwa, Jennifer V. Andrew, and Dayang Haryani Diana Awang Amit. 2020. "Growth Challenges of SMEs: Empirical Evidence in Sabah, Malaysia." *ASEAN Entrepreneurship Journal* 6 (1): 8–14. https://www.researchgate.net/publication/343404542_Growth_Challenges_of_SMEs_Empirical_Evidence_in_Sabah_Malaysia.
- [3] Bendixen, Mike. 2003. "A Practical Guide to the Use of Correspondence Analysis in Marketing Research." Technical note. *Marketing Bulletin* 14: 1–15.
- [4] Chandrapaul, S., R. Soni, S. Sharma, H. Fagna, and S. Mittal. 2018. "News Analysis Using Word Cloud." In *Advances in Signal Processing and Communication: Select Proceedings of ICSC 2018*, 55–64. Singapore: Springer.
- [5] Greenacre, Michael. 2017. *Correspondence Analysis in Practice*. 3rd ed. Boca Raton, FL: Chapman & Hall/CRC.
- [6] Guzzetti, Andrea, Riccardo Crespi, and Valeria Belvedere. 2024. "Phygital Luxury Experiences: A Correspondence Analysis on Retail Technologies." *International Journal of Consumer Studies* 48 (2): e13008. <https://doi.org/10.1111/ijcs.13008>.
- [7] Jin, Y. 2017. "Development of Word Cloud Generator Software Based on Python." *Procedia Engineering* 174: 788–792.
- [8] Kabir, A. I., K. Ahmed, and R. Karim. 2020. "Word Cloud and Sentiment Analysis of Amazon Earphones Reviews with R Programming Language." *Informatica Economica* 24 (4): 55–71.
- [9] Karabulut, A. N. 2023. "Comparing Young People's Coffee Shop Perceptions with Their Senses of Taste." *Yönetim ve Ekonomi Dergisi* 30 (1): 1–19.

- [10] Kevin, A., and Z. Z. Simon. 2024. "The Influence of Price Perception, Service Quality, Store Atmosphere, and Word of Mouth on Purchasing Decisions." *Research of Economics and Business* 2 (2): 94–106. <https://www.researchgate.net/publication/384449897>.
- [11] Kim, Jae-On, and Charles W. Mueller. 1978. *Introduction to Factor Analysis: What It Is and How to Do It*. Beverly Hills, CA: Sage Publications.
- [12] Krejcie, Robert V., and Daryle W. Morgan. 1970. "Determining Sample Size for Research Activities." *Educational and Psychological Measurement* 30: 607–610.
- [13] Sarah, M. 2025. "Understanding Correspondence Analysis: A Comprehensive Guide for 2025." *Displayr*. <https://www.displayr.com/understanding-correspondence-analysis-a-comprehensive-guide/>.
- [14] Shrestha, N. 2021. "Factor Analysis as a Tool for Survey Research." *Journal of Education and Research* 11 (2): 1–12.
- [15] Siswaningsih, W., H. Firman, and A. Khoirunnisa. 2017. "Development of Two-Tier Diagnostic Test Pictorial-Based for Identifying High School Students' Misconceptions on the Mole Concept." *Journal of Physics: Conference Series* 812 (1): 012117. <https://doi.org/10.1088/1742-6596/812/1/012117>.
- [16] Yang, B. 2005. "Factor Analysis Methods." In *Research in Organizations: Foundations and Methods of Inquiry*, edited by F. Drasgow and N. Schmitt, 181–199. Mahwah, NJ: Lawrence Erlbaum Associates.