ENTHUSIASTIC
INTERNATIONAL JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 2, Issue 2, October 2022, pp. 110-121

Multiple Correspondence Analysis towards the Change of Income and Sociodemographic of Citizens due to COVID-19 Pandemic in Malaysia

Ong Wen Xuan [a,1], Lim Chui Ting [a,2], Nurulain Nabilah Binti Muhammad Aris Fadzilah [a,3], Nora Binti Muda [a,4,*]

[a] Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.
[1] a170191@siswa.ukm.edu.my; [2] a171964@siswa.ukm.edu.my; [3] a170186@siswa.ukm.edu.my; [4] noramuda@siswa.ukm.edu.my*
* corresponding author

ARTICLE INFO

Article history
Received August 10, 2022
Revised September 10, 2022
Accepted October 12, 2022

Keywords
Coronavirus
Spending patterns
COVID-19
Movement control order
Multiple correspondence

ABSTRACT

This paper intends to observe the economic impact and spending patterns of citizens when the spread of coronavirus occurred uncontrollably in Malaysia. The data used in this paper is the result of a first-round Special Survey 'Effect of COVID-19 on Economy and Individual' conducted by the Department of Statistics Malaysia. Based on the paired-*t* test, all the aspects tested are found to have significant differences before and during the pandemic. This is likely because citizens comply with the rules set during Movement Control Order. Next, the Chi-square test between the changes in citizens' monthly income and sociodemographic factors are significantly associated. Therefore, factors state, gender, ethnicity, and age group of citizens are used in further analysis of multiple correspondence (MCA) to study the relationship between several categorical variables. Hence, citizens of age group 35–44 years old from Central region and Chinese citizens have no changes in their income before and during the pandemic. Citizens of age group 15–34 years old and Northern region have both increment or reduction income whereas men citizens received lower-income payments. Indian women and citizens from Eastern region tend to not work during the pandemic. This study can be a guide to the government in overcoming social problems.

1. Introduction

Coronavirus, known as COVID-19 is an infectious disease caused by 2019-nCoV from Wuhan, China in late December 2019. After the source of this virus is confirmed, that is from a seafood market, many operations, transportation and trading have been closed immediately in China. Unfortunately, the virus spread widely in many cities in China and expanded abroad to Malaysia, Singapore, Japan, Korea, France and many more [1].

Generally, the virus can be easily transmitted through droplets of cough or sneeze. The basic symptoms if infected with this disease are fever, cough, loss of smell or taste, sore throat and so on. Older adults with underlying medical conditions like diabetes and heart disease are at higher risk of getting this disease which causes trouble breathing or even death [2]. The challenge of dealing with

p-ISSN 2798-253X
e-ISSN 2798-3153

large clusters of coronavirus transmission is not easy when there are insufficient resources, capacity and knowledge in many countries. Hence, World Health Organization (WHO) has been concerning the spread and severity levels of the disease, announcing COVID-19 as a pandemic on 11th March 2020 [3].

The announcement of the lockdown from 18th March 2020 by Deputy Prime Minister of Malaysia, Tan Sri Muhyiddin bin Mohd Yassin has shocked and alarmed all the Malaysians [4]. The aim of implementing this policy is to reduce the spread of COVID-19. It is a big challenge for all citizens when health and economic crises occurred simultaneously in Malaysia [5]. COVID-19 has altered the physical activity, food consumption and working pattern of individuals. These behavior and lifestyle changes have created stressful life events for the majority of individuals and impacted the whole system and community in our country [6]. Many companies and factories are not allowed to operate or limit the number of employees to work physically during this period. This has caused many people to suffer pay cuts or lay-offs in a short period due to the cost-cutting measures of employers [7]. At the same time, the economy has been affected as most of the sectors could not generate their working operations as usual. Hence, the government has also introduced several financial aids to help those who are struggling during difficult times [8].

This study is related to the economic and spending patterns of citizens in Malaysia. We obtained the expenditure changes in several aspects before and after the pandemic and discusses the changes in monthly income due to the pandemic based on the sociodemographic of respondents. The purpose of this research is to compare the spending patterns of citizens before and during the pandemic COVID-19. Besides, the relationship between changes in citizens' monthly income and sociodemographic factors is determined using multiple correspondence analysis.

2. Literature Review

The spread of coronavirus outbreaks that occurs around the world has become a norm. There are many studies related to this issue in respective countries to understand the new lifestyle and spending patterns of citizens.

In Malaysia, it is not usual to use a face mask even when falling ill. After the emergence of this virus, the phenomenon of lacking face masks and cleaning accessories poses and leads to various problems in our country. A face mask is necessary for everyone whenever going out as the coronavirus emerges uncontrollably in our country. Hence, all the citizens must obey the instructions set by the government such as avoiding mass gatherings and keeping hygienic all the time. Besides, awareness programs are also necessary to raise the consciousness of this disease [9].

During the pandemic, the daily routines of citizens have been affected regardless of their demographic background, financial status, and country. Hence, understanding consumers' spending behavior is beneficial to reflect the changes on the business end before and during the restriction of lockdown. The effect of COVID-19 lockdown has affected many sectors like indoor entertainment, social media and life insurance schemes that have gained much profit while the citizens restricted their spending on restaurant dine-in, traveling for outdoor activity and real estate [10]. As a result, this research is focused on Malaysia in different aspects of events before and during COVID-19 through a paired-*t* test.

Multiple correspondence analysis (MCA) is an extension of correspondence analysis (CA) that studies the association between two or more categorical dependent variables. This method is used to detect the underlying structures in the data set [11]. It is useful in this study to examine how sociodemographic affect the monthly income during the pandemic in Malaysia using MCA. However, the perfect association of diagonal matrices in Burt matrix inflates the total inertia leading to a lower percentage of explained inertia in the biplot. According to Greenacre, there are two alternative approaches to overcome this problem, namely adjustment of inertia and joint correspondence analysis (JCA). The adjustment of inertia is rescaling the coordinates of the solution to best fit the pairwise cross-tabulations of Burt matrix whereas JCA is used to find the optimal

p-ISSN 2798-253X
e-ISSN 2798-3153

weighted least-squares that fit the off-diagonal tables [12]. Thus, adjustment of inertia is used in this study to optimize the solution of off-diagonal tables in the matrix.

3. Material & Methodology

4.1. Data

The data used in this study is secondary data from the first round of COVID-19 Special Impact Survey on the Economy and Individuals conducted by the Department of Statistics Malaysia from the 23rd to 31st of March 2020. This survey is not an official statistic but to get an idea of the situation during the first lockdown in our country.

There are 168145 respondents from different sectors non-government employees aged 15 years old and above involved in this survey throughout Malaysia. The questionnaire is divided into three modules, including general information and employment information from respondents to understand their sociodemographic, family members and working backgrounds. Lastly, the third module is aimed to determine their spending changes behavior before and during COVID-19 according to the Likert scale. The aspects of events before and during COVID-19 in this survey as shown in Table 1.

Table 1. Aspects of Events of Spending Behavior Before and During COVID-19

Statement
Traveling for business purposes
Traveling for personal purposes
Eat at the restaurant
Eat at fast food restaurants
Buy cooking materials at market, supermarket or grocery store
Buy ready-made products at the supermarket, grocery store or supermarket
Buy health products at the pharmacy
Using taxi services or e-hailing services such as myCar, Grabcar, maxim
Watching movies in the cinema
Online shopping
Buy non-food products in hypermarkets or stores

According to Table 2, the 5-level Likert scale is used to determine the frequency of aspects events that happened before and during COVID-19.

Table 2. Scale for the Frequency of Aspect of Events Before and During COVID-19

Scale	Frequency of Events
1	Never
2	Rarely
3	Sometimes
4	Often
5	Very Often

4.2. Data Cleaning

Data cleaning is a process of detecting and fixing incomplete, duplicate or missing data in a set of data [13]. This process is important to improve data quality as well as provide accurate and reliable information for decision-making. In a huge data set, data cleaning needs to be performed as it is likely to influence the results [14]. Due to the huge amount of data in this study, the null values are dropped from the dataset.

4.3. Paired-*t* Test

Paired-*t* test is a method of hypothesis testing against two groups of data that are dependent and interconnected with each other for a same subject [15]. Based on the survey, there are 11 different aspects before and during COVID-19 pandemic.

p-ISSN 2798-253X
e-ISSN 2798-3153

As the result, based on the aspects, the null hypothesis and alternative hypothesis are H_0: No changes in spending behavior for aspect Q_i before and during COVID-19 pandemic and H_1: Change in spending behavior for aspect Q_i before and during COVID-19 pandemic where i = 1, 2, …, 11.

The formula of test statistics for the Paired-t test, t as below:

$$t = \frac{\sum D}{\sqrt{\frac{N(\sum D^2)-(\sum D)^2}{(N-1)}}} \quad (1)$$

where $\sum D$ is the summation of differences between two sample means, $\sum D^2$ is the summation of differences between two sum of squares means and N indicates the number of paired samples.

The test statistic is compared to the critical value, $t_{N-1,\alpha}$ with degree of freedom, d.f. = n-1 from the t distribution table for a chosen confidence level, α. When the test statistic is greater than the critical value, then the null hypothesis is rejected.

4.4. Chi-square Test

Chi-square test is a non-parametric analysis used to determine the association between nominal or categorical data. It is a test to measure whether the observed data are significantly different from the expected data [16]. The association between sociodemographics of respondents (X_i) and changes of their monthly income (Y) is determined using Chi-sqaure test. The factors included are state (X_1), gender (X_2), ethnicity (X_3) and age group (X_4). The changes of respondents' monthly income during the lockdown are received income as usual, reduction, increment, no information and not working. Assuming independence, there are four hypothesis tests respectively and the null hypothesis, H_0: There is no association between X_i and Y whereas alternative hypothesis, H_1: There is an association between X_i and Y where i = 1, 2, 3, 4.

The test statistic for Chi-square test is based on the formula below:

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{p}\frac{(O_{ij}-E_{ij})^2}{E_{ij}} \quad (2)$$

where O_{ij} is the observed value for i-th row and j-th column and E_{ij} is the expected value for i-th row and j-th column.

$$E_{ij} = \frac{O_{i.}-O_{.j}}{O_{..}} \quad (3)$$

where $O_{i.}$ is the frequency for column by i-th row, $O_{.j}$ is the frequency for column by j-th column and $O_{..}$ is the overall frequency.

The test statistic is compared to the critical value, $\chi^2_{0.05(\,(I-1)*(J-1)}$ with degree of freedom d.f.=(I-1)*(J-1) from the Chi-square table for a chosen confidence level, α. If the test statistic is greater than critical value, then the null hypothesis can be rejected.

4.5. Multiple Correspondence Analysis

Multiple correspondence analysis (MCA) is a multivariate technique that explains the relationships between several categorical variables. It shows the relationships between rows and columns simultaneously from two-way contingency tables in a low-dimensional vector space. The results can be interpreted according to the position and distribution of the respective dimensions through a biplot graph [17]. Based on the result of Chi-square test, the variables that are significant can be taken in this analysis. MCA is obtained by using standard correspondence analysis on an indicator matrix $\boldsymbol{X}$ or its cross tabulation which is known as Burt matrix $\boldsymbol{B}$ [18].

Indicator matrix $\boldsymbol{X}$ is formed to indicate the presence of elements of each respondent. The elements of the matrix are recorded as binary numbers, that is the selected element denoted as 1 and the unselected element denoted as 0.

When there is a huge amount of data and will be time-consuming, Burt matrix $\boldsymbol{B}$ is built from the matrix $\boldsymbol{I}$ as below:

$$\boldsymbol{B} = \{\boldsymbol{b_{ij}}\} = \boldsymbol{X^T X} = \begin{bmatrix} \boldsymbol{X_1^T X_1} & \boldsymbol{X_1^T X_2} & \cdots & \boldsymbol{X_1^T X_w} \\ \boldsymbol{X_2^T X_1} & \boldsymbol{X_2^T X_2} & \ldots & \boldsymbol{X_2^T X_w} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X_w^T X_1} & \boldsymbol{X_w^T X_2} & \ldots & \boldsymbol{X_w^T X_w} \end{bmatrix} \quad (4)$$

Correspondence matrix, $\boldsymbol{P}$ is calculated by dividing the matrix $\boldsymbol{B}$ element, b_{ij} with the total elements, n of matrix $\boldsymbol{B}$, hence the formula as below:

$$\boldsymbol{P}_{ij} = \frac{b_{ij}}{n} \quad (5)$$

Each element of the variables has its own mass of overall data. The total row mass can be calculated from the matrix $\boldsymbol{P}$ by $r_i = P \cdot 1_{J\times 1}$ while the mass of the column will obtain the same value because matrix $\boldsymbol{B}$ is a symmetry matrix [19].

Further the formation of a standard residual matrix $\boldsymbol{S}$ with equation as following:

$$\boldsymbol{S} = \boldsymbol{D_r^{-\frac{1}{2}}(P - rr^T)D_c^{-\frac{1}{2}}} \quad (6)$$

Next, singular value decomposition (SVD) from matrix $\boldsymbol{S}$ with $\boldsymbol{S = UD_\lambda V^T}$ where $\boldsymbol{U^T U = V^T V = I \cdot D_\lambda}$ is formed. $\boldsymbol{D_\lambda}$ is n × n diagonal matrix from SVD in the ascending order, $\boldsymbol{a_1 \geq a_2 \geq a_3 \geq \cdots}$, $\boldsymbol{V}$ is matrix of vector eigen and $\boldsymbol{I}$ is identity matrix. The results of vector eigen and the eigen value λs with s dimension also can be obtained.

In MCA, each principal inertia values indicated as a proportion of total variance explained by the corresponding principal dimension. Hence, the total inertia is found by calculating the trace from the matrix $\boldsymbol{S}$ and its transposition matrix [17] as follows:

$$trace(\boldsymbol{SS^T}) = \sum_{i=1}^{I}\sum_{j=1}^{J}\frac{(p_{ij}-r_i c_i)^2}{r_i c_i} \quad (7)$$

The coordinates of each *i*-th row for *s*-th dimension are found according to the following equation:

$$x_{is} = \frac{v_{ij}}{\sqrt{r_i}} \quad (8)$$

where v_{ij} is the eigen vector and $\sqrt{r_i}$ is the row mass.

3.5.1 Adjusted Inertia

The Burt matrix $\boldsymbol{B}$ obtained from the MCA is a square symmetric block matrix that causing an overestimation of its total inertia. Hence, the aim of inertia adjustment is to omit the cross-tabulation of perfect association in order to obtain a higher proportion of inertia and reduce loses information from the original data [20].

Since matrix $\boldsymbol{B}$ is a symmetry matrix, adjusted inertia can be calculated from the average of upper or lower submatrix. As the total inertia of matrix $\boldsymbol{B}$ is available, the constant amounts of problematic diagonal matrices can be subtracted, and the adjusted inertia is as follow:

$$Adjusted\ inertia\ of\ matrix\ \boldsymbol{B} = \frac{Q}{Q-1}\left(inertia(\boldsymbol{B}) - \frac{J-Q}{Q^2}\right) \quad (9)$$

where Q is the number of variables and J is the total number of elements in each variable.

Let $\sqrt{\lambda_k}$ denote the *k*-th principal inertia of matrix $\boldsymbol{B}$, only $\sqrt{\lambda_k}$ which is greater than $\frac{1}{Q}$ retained. Hence, the adjustment can be shown as below:

$$\frac{Q}{Q-1}\left(\sqrt{\lambda_k} - \frac{1}{Q}\right), k = 1,2,\ldots, for\ \sqrt{\lambda_k} > \frac{1}{Q} \quad (10)$$

p-ISSN 2798-253X
e-ISSN 2798-3153

Based on this equation, an effective MCA is produced with new singular and adjusted inertia. In short, biplot graph can be explained the proportion of inertia that is the higher variance of the original data.

4. Results and Discussion

4.1. Data Cleaning

Out of 168145 respondents, a total of 6365 incomplete data were removed from this study. After the data cleaning process, a total of 161780 respondents' data is used to conduct the next analysis.

4.2. Descriptive Analysis

Out of the 161780 respondents, there are 43560 respondents (26.93%) from Selangor, followed by 33341 respondents (20.61%) from Johor, 13787 respondents (8.52%) from Kuala Lumpur and 11359 respondents (7.02%) from Sabah while the other states are only between 0.49% to 4.85%.

In terms of changes in monthly income due to COVID-19, 47173 respondents (29.16%) who received income as usual while 30903 respondents (19.10%) experienced a reduction of income and only 346 respondents (0.21%) who received increment of monthly income. In fact, there are 56 459 respondents (34.90%) who had no information while 26899 respondents (16.63%) who is in not working status when the survey conducted.

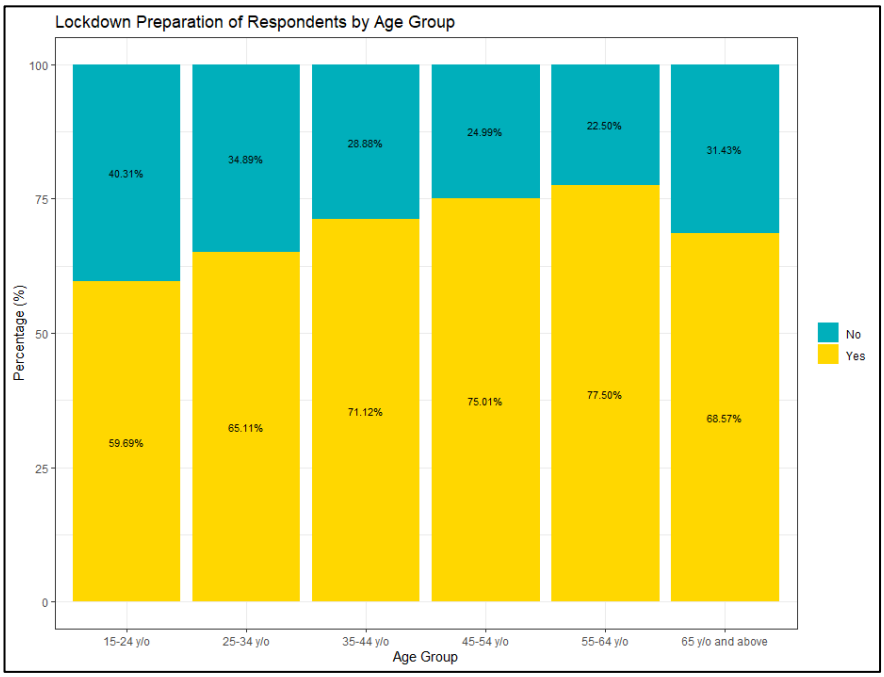

Fig. 1. Lockdown preparation of respondents by age group

Fig. 1 shows respondents' preparation towards lockdown in our country according to the age group. Respondents from aged group 55–64 years are most (77.50%) well prepared to face this situation, followed by respondents aged 45–54 years (75.01%), 35–44 years (71.12%), 65 years and above (68.57%), 25–34 years (65.11%) and finally 15 - 24 years (59.69%). Therefore, it can be concluded that most of the respondents in the group of 15 - 24 years old are not ready for the lockdown compared to other age groups.

4.3. Paired-*t* Test

Pair-*t* test was used because each aspect have two related groups and significant differences in means. There are 11 aspect variables in this study used to perform the hypothesis testing for two sample means between before and during COVID-19. The test statistics and results of the 11 aspects at 0.05 significance level in this study as shown in Table 3.

Table 3. Test Statistics and Results of 11 Aspects Using Paired-*T* Test

	Aspect	Test Statistics	Results
Q1	Traveling for business purposes	374.21	Rejected
Q2	Traveling for personal purposes	412.55	Rejected
Q3	Eat at the restaurant	465.22	Rejected
Q4	Eat at fast food restaurants	423.36	Rejected
Q5	Buy cooking materials at market, supermarket or grocery store	248.37	Rejected
Q6	Buy ready-made products at the supermarket, grocery store or supermarket	265.97	Rejected
Q7	Buy health products at the pharmacy	257.61	Rejected
Q8	Using taxi services or e-hailing services such as myCar, Grabcar, maxim	257.45	Rejected
Q9	Watching movies in the cinema	327.22	Rejected
Q10	Online shopping	303.26	Rejected
Q11	Buy non-food products in hypermarkets or stores	367.01	Rejected

Test statistics with degree of freedom 161779 at 0.05 significance level for the above 11 statements is 1.960. Based on the Table 3, test statistics for all the aspects are greater than critical value, hence the null hypothesis were rejected at 95% confidence level. There was sufficient evidence to conclude that the spending behavior for all the aspects changed before and during COVID-19 pandemic.

4.4. Chi-square Test

The assocation between sociodemographic of respondents (X_i) and changes of their monthly income was determined through Chi-sqaure test. There were four factors including the state (X_1), gender (X_2), ethnic (X_3) and age group (X_4) performed in this test. The test statistics and results at 0.05 significance level in this study as shown in Table 4.

Table 4. Results and Test Statistics of Four Testing Using Chi-Square Test

Test, i	Degree of Freedom	Chi-Square Value	Test Statistics	Results
1	20	31.41	3132.5	Rejected
2	4	9.49	2023.5	Rejected
3	12	21.03	332.57	Rejected
4	8	15.51	5872.9	Rejected

Based on the Table 4, the test statistics with degree of freedom at 0.05 level of significance for all the testing above are greater than the Chi-square values respectively. Hence, the null hypotheses are rejected at 95% confidence level. There are sufficient evidence to conclude that the association between the sociodemographic factors and their monthly income are significant. The factors state, gender, ethnic and age group give an impact on citizens' monthly income when the first lockdown happened in Malaysia.

4.5. Multiple Correspondence Analysis

Based on the result of Chi-square test, the factors state and age group are significantly associated with the changes in monthly income at 5% significance level. Hence, they can be used in the following multiple correspondence analysis (MCA).

Indicator matrix $\boldsymbol{X}$ shows the variables changes of monthly income, state and age group of respondents. The elements of the variables represent in table matrix $\boldsymbol{X}$ and each element chosen notated as 1 while 0 for not chosen as shown in the Table 5.

Table 5. Tabulation of Matrix $\boldsymbol{X}$ for State, Age Group and Monthly Income

Response	State	Age Group	Monthly Income

p-ISSN 2798-253X
e-ISSN 2798-3153

	Southern Region	*Northern Region*	*Eastern Region*	*Sabah*	*Sarawak*	*Central Region*	*15-34 y/o*	*35-44 y/o*	*45+ y/o*	*Normal*	*Reduction*	*Increment*	*No Information*	*Not Working*
1	0	0	0	0	0	1	0	1	0	1	0	0	0	0
2	0	0	0	0	0	1	0	1	0	1	0	0	0	0
3	0	0	0	0	0	1	1	0	0	0	0	0	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
161780	0	1	0	0	0	0	1	0	0	1	0	0	0	0

Burt matrix ***B*** obtained from matrix ***I*** shows the factors state, age group and monthly income is built. It is used to compare columns and rows of combined variables from the contingency table simultaneously. The proportion of variance for the first dimension is 0.4131 while the second dimension is 0.3739. The cumulative inertia of two dimensions is 21.46%. Hence, there is 78.54% (100% - 21.46%) of information might loss from the raw data as the value of inertia affect the result of which biplot to describe the actual data. Inertia, proportion of inertia and cumulative proportion of inertia (%) are described in Table 6.

Table 6. Inertia, Proportion of Inertia and Cumulative Proportion of Inertia (%)

Dimension	Inertia	Proportion of Inertia (%)	Cumulative Proportion of Inertia (%)
1	0.4131	11.27	11.27
2	0.3739	10.20	21.46
3	0.3429	9.35	30.82
4	0.3353	9.14	39.96
5	0.3337	9.10	49.06
6	0.3322	9.06	58.12
7	0.3288	8.97	67.09
8	0.3224	8.79	75.88
9	0.3131	8.54	84.42
10	0.3003	8.19	92.61
11	0.2710	7.38	100.00
Total	3.667	100.00	-

Since the cumulative proportion of inertia for two dimensions are low, an adjustment of the singular values is done so that MCA solution best fits the off-diagonal tables. Hence, the adjusted inertia of matrix ***B*** is as follows:

$$Inertia(\boldsymbol{B}) = \frac{3}{2}\left(1.2359 - \frac{14-3}{9}\right) = 0.0205$$

Based on the Table 7, there are only five principal inertias λ_k that satisfy $\sqrt{\lambda_k} > \frac{1}{3} = 0.3333$ to be replaced by the singular values in the MCA. The singular value of matrix ***B*** for first dimension is 0.1197 and its cumulative proportion of adjusted inertia is 70.02%. Therefore, cumulative proportion of adjusted inertia for two dimensions shows 88.09%

Table 7. Singular Value, Proportion of Adjusted Inertia and Cumulative Proportion of Adjusted Inertia (%)

Dimension	Singular Value	Proportion of Adjusted Inertia (%)	Cumulative Proportion of Adjusted Inertia (%)
1	0.1197	70.02	70.02
2	0.0608	18.07	88.09
3	0.0143	1.01	89.10
4	0.0029	0.04	89.14
5	0.0005	0.001	89.14
Total	0.1983	89.14	-

A two-dimensional graph known as biplot is built. The dimensions are ordered decreasingly based on the amount of variance explained in the solution. Dimension 1 which is x-axis explains the most

variance in the solution followed by Dimension 2 (y-axis) [21]. The coordinates of biplot are according to their dimensional values based on mass and inertia respectively as shows in the Table 8.

Table 8. Coordinates of Element of Each Variables for Dimension 1 and Dimension 2

Variable	Element	Dimension 1	Dimension 2	Mass	Inertia
State	Southern region	-1.232	0.1552	0.0928	0.0814
	Northern region	0.6163	0.9004	0.0429	0.0970
	Eastern region	0.4612	1.6679	0.0322	0.1007
	Sabah	-0.3418	0.5646	0.0250	0.1028
	Sarawak	-0.229	0.2736	0.0123	0.1071
	Central region	0.6593	-0.9705	0.1281	0.0693
Age group	15 – 34 y/o	1.1655	0.5657	0.1381	0.0673
	35 – 44 y/o	-0.5124	-1.2147	0.1244	0.0706
	45+ y/o	-1.3721	1.0308	0.0708	0.0891
Monthly income	Normal	0.5441	-1.4395	0.0972	0.0797
	Reduction	1.2551	0.1147	0.0637	0.0910
	Increment	1.8839	-0.1658	0.0007	0.1109
	No information	-1.5005	0.1588	0.1163	0.0753
	Not working	0.7290	2.0616	0.05542	0.0937

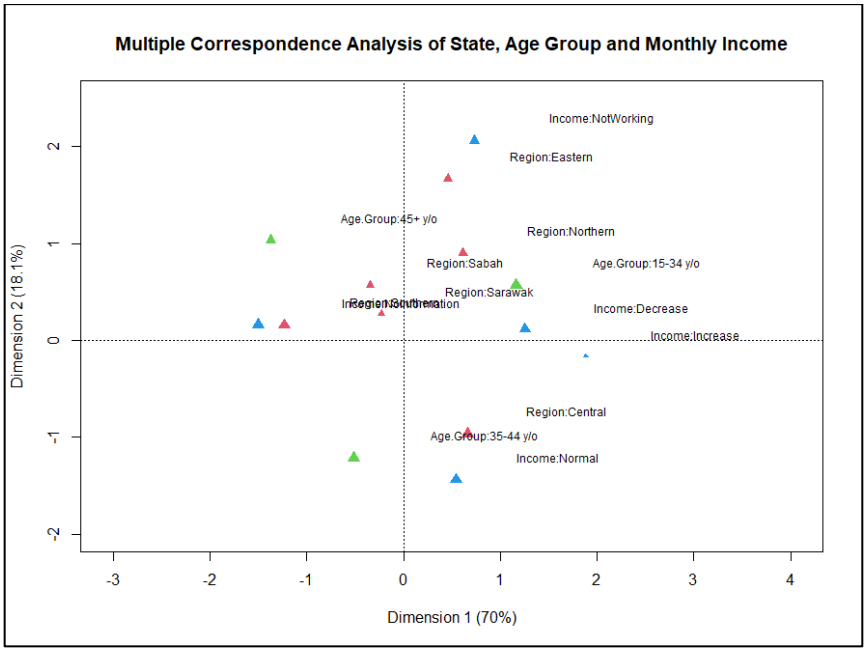

Fig. 2. Biplot for correspondence analysis of state, age group, and monthly income

Based on the Fig. 2, no changes of income and income reduction before and during the lockdown period have significant difference in x-axis as they are far apart from each other, whereas not working and receive income as usual in y-axis. Respondents who receive income as usual are from Central region and age group 35–44 years old. Respondents from Eastern region tend to not working whereas respondents received increment or reduction of income are from Northern region and age of 15–34 years old. However, respondents who do not have information on their income changes during the lockdown in our country are 45 years old and above from Southern region, Sabah and Sarawak.

Next, as the factors gender and ethnicity group are also significantly associated with the changes in monthly income at 5% significance level, the factors can be taken for the next multiple correspondence analysis. Follow the process as in the previous, there are 4 principal inertia λ_k that

satisfy $\sqrt{\lambda_k} > \frac{1}{3} = 0.3333$ after the adjustment. Hence, the cumulative proportion of adjusted inertia for two dimensions is 76.86%.

Table 9. Inertia, Proportion of Inertia and Cumulative Proportion of Inertia (%)

Dimension	Singular Value	Proportion of Adjusted Inertia (%)	Cumulative Proportion of Adjusted Inertia (%)
1	0.0588	70.28	70.28
2	0.0180	6.58	76.86
3	0.0043	0.38	77.24
4	0.0015	0.04	77.28
Total	0.0826	77.28	-

The biplot graph is obtained according to their dimensional values based on mass and inertia respectively as shows in the Table 10.

Table 10. Coordinates of Element of the Each Variables for Dimension 1 And Dimension 2

Variable	Element	Dimension 1	Dimension 2	Mass	Inertia
Gender	Male	1.4155	0.2566	0.1373	0.0662
	Female	-0.9911	-0.1797	0.1961	0.0463
Ethnicity	Others	0.0947	-1.8021	0.0298	0.1011
	Chinese	-1.2449	3.7687	0.0244	0.1032
	Indian	-0.9238	-1.2810	0.0089	0.1082
	Malay	0.1326	-0.0990	-0.2701	0.0211
Monthly income	Normal	-0.5509	1.7995	0.0972	0.0789
	Reduction	2.2597	0.0514	0.0637	0.0909
	Increment	2.1576	4.5527	0.0007	0.1109
	No information	-0.0794	-0.9787	0.1163	0.0724
	Not working	-1.4911	-1.2192	0.0554	0.0931

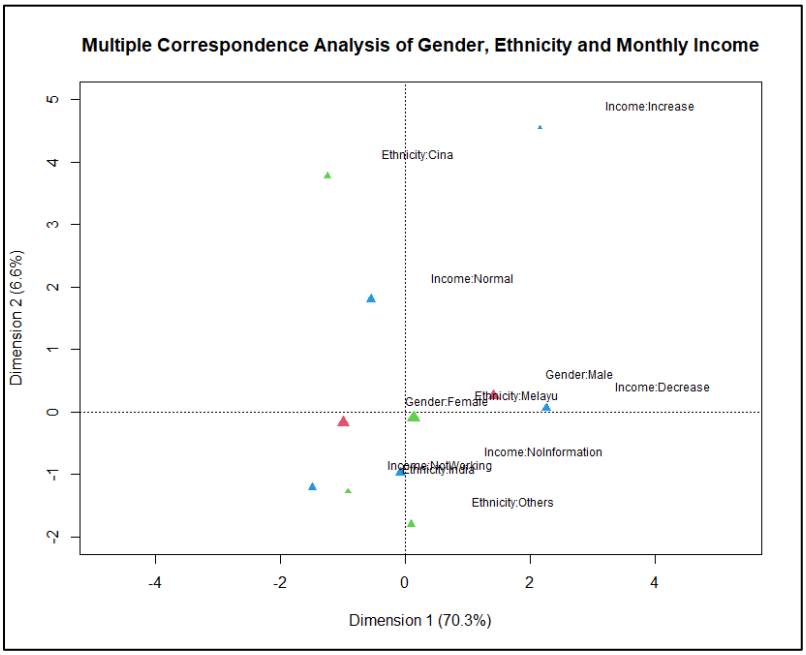

Fig. 3. Biplot in correspondence analysis of gender, ethnicity, and monthly income

Based on the Fig. 3, Chinese ethnicity and income reduction are significant difference in axis 1 whereas the increment of income and other ethnicity in axis 2. Based on this analysis, Malay ethnic respondents may play a small role because the coordinate point is located close to the center in this biplot. Chinese respondents are found to receive income as usual during the pandemic lockdown. In

addition, men have experienced a reduction of income as majority of them are working in "other services" sector and education sector. Indian women are more likely to be unemployed while the other ethnic respondents have no information on changes in their current income during the lockdown.

5. Conclusion

Pandemic of COVID-19 that happened in Malaysia give a huge impact on economy and lifestyle to the citizens and country. This research studied and identified the spending pattern and changes in daily activities that occurred during the first lockdown since March 2020.

Based on the paired-*t* test, all the aspects including eating in the restaurant, online shopping and traveling for personal and business purposes and so on have significantly difference before and during the pandemic. Chi-square test between the changes of monthly income and sociodemographic factors such as state, gender, ethnicity, and age group of citizens were all significantly associated.

From the result of Multiple correspondence analysis, citizens of age group 35–44 years old from Central region and Chinese citizens have no changes in their monthly income before and during the pandemic. Other than that, citizens of age group 15–34 years old and Northern region states have both increased or decreased income whereas men citizens received lower-income payments. Lastly, Indian women and citizens from Eastern region tend to not working during the pandemic. Citizens of 45 years old and above from Southern region, Sabah or Sarawak and citizens from other ethnicity have lack of information on their monthly income changes when the survey was conducted.

There were some recommended suggestions for further studies. Analysis for spending patterns of citizens before and during the pandemic in different aspects can be further carried out to understand the sector which suffer profits or vice versa which incur losses. Joint Correspondence Analysis can be used as further analysis of MCA by rescaling the coordinates of the solution to obtain the pairwise cross-tabulation off the main diagonal of the Burt matrix.

Acknowledgment

This research is fully supported by Universiti Kebangsaan Malaysia (UKM). We greatly thank UKM and supervisor for the professional guidance and valuable support.

References

[1] Y.C. Wu, C.S. Chen, and Y.J. Chan, "The Outbreak of COVID-19: An Overview," *Journal of Chinese Medical Association* 83 (3), 217 (2020).

[2] Unknown. "Symptoms of COVID-19," https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html. Retrieved 7 January 2022.

[3] World Health Organization. "WHO Director-General's opening remarks at the media briefing on COVID-19–11 March 2020," https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020 Retrieved 7 January 2022.

[4] A. Tang, "Malaysia Announces Movement Control Order After Spike in COVID-19 Cases (Updated)," https://www.thestar.com.my/news/nation/2020/03/16/malaysia-announces-restricted-movement-measure-after-spike-in-covid-19-cases. Retrieved 15 October, 2021.

[5] I.S.Z Abidin, "Analisis Ekonomi Bukti Negara Tidak Gagal Urus Impak COVID-19," https://www.bharian.com.my/kolumnis/2021/07/841952/analisis-ekonomi-bukti-negara-tidak-gagal-urus-impak-covid-19. Retrieved 15 October 2021.

[6] A. Ahmad *et al.*, "Changes in Health-Related Lifestyles and Food Insecurity and Its Association with Quality of Life during the COVID-19 Lockdown in Malaysia." *BMC Public Health*, 22(1), 1-9 (2022).

[7] Unknown. "COVID-19 Pandemic: Frequently Asked Questions on Employment," https://www.skrine.com/insights/alerts/april-2020/covid-19-pandemic-frequently-asked-questions-o-1. Retrieved 13 January 2022.

[8] R. Raihan, "A Complete Guide to COVID-19 Financial Assistance In Malaysia," https://www.homage.com.my/resources/covid-19-financial-assistance-malaysia/. Retrieved 14 January 2022.

[9] A.A. Azlan, M.R. Hamzah, T.J. Sern, S.H. Ayub, and E. Mohamad. "Public Knowledge, Attitudes and Practices towards COVID-19: A Cross-Sectional Study in Malaysia," *Plos one* 15(5), e0233668 (2020).

[10] S. Jha, and V. Pradhan, "Factors causing change in customer behaviour during COVID-19 Pandemic," *Journal of Critical Reviews* 7(19), 3792-3804 (2020).

[11] N. Topaç, M. Bardak, S.B. Kalkan, and M. Kirisci. "Early Childhood Children in COVID-19 Quarantine Days and Multiple Correspondence Analysis," *Journal of Mathematical Sciences and Modelling* 3(3): 130-134 (2020).

[12] O. Nenadic and M. Greenacre, "Correspondence Analysis in R, with Two-and Three-Dimensional Graphics: the ca Package," *Journal of Statistical Software*, 20: 1-13 (2007).

[13] A. Crossman, "Pembersih Data," https://ms.eferrit.com/pembersihan-data/#. Retrieved 12 November 2021.

[14] K.F. Hidayati. *Tak Hanya Agar Rapi, Pahami Arti dan Pentingnya Data Cleansing, Yuk!*. Glints. https://glints.com/id/lowongan/data-cleansing-cleaning/#.YrCOz0ZBzIX. Retrieved 22 November 2021.

[15] T.K. Kim, "T Test as a Parametric Statistic," *Korean Journal of Anesthesiology*, 68(6), 540-546 (2015).

[16] M.L. McHugh, "The chi-square test of independence," *Biochemia Medica*, 23(2), 143-149 (2013).

[17] Greenacre, M., *Correspondence Analysis in Practice, edition.* Boca Raton, Chapman dan Hall/CRC, 2007.

[18] L. Lebart, A. Morineau, and K.M. Warwick, *Multivariate Descriptive Statistical Analysis; Correspondence Analysis and Related Techniques for Large Matrices*. New York US, John Wiley & Sons, Inc. (1984).

[19] O. Nenadic and M. Greenacre, *Computation of Multiple Correspondence Analysis, with code in R.* Barcelona, Universtat Pompeu Fabra, 2012.

[20] M. Greenacre, *Biplots in Practice*. Bilbao dan Madrid, Fundación BBVA, 2010.

[21] A. Kassambara, Articles–Principal component methods in R: practical guide, HCPC–Hierarchical Clustering on Principal Components: Essentials. STHDA (Statistical tools for highthroughput data analysis) (2017a), 2018.

p-ISSN 2798-253X
e-ISSN 2798-3153