ORIGINAL ARTICLE

# Population Density as a Factor in the Spread of Covid-19 Cases in Peninsular Malaysia

Afifa Syuhada Binti Abd Rasid [1], Nazri Che Dom[1,2], Samsuri Abdullah[3], Hasber Salim[4]

[1] Centre of Environmental Health and Safety, Faculty of Health Sciences, Universiti Teknologi MARA Selangor, Malaysia
[2] Integrated of Mosquito Research Group (I-Merge), Faculty of Health Sciences, Universiti Teknologi MARA, Selangor, Malaysia
[3] Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia.
[4] School of Biological Sciences, Universiti Sains Malaysia, Penang, Malaysia

## ABSTRACT

**Introduction:** Coronavirus disease also known as COVID-19 in Malaysia were reported on the 25th January 2020 until present. There were several factors that influence the distribution of COVID-19 events. The objectives of this study are to explore the association between population density and the spread on early wave of COVID-19 in Peninsular Malaysia. **Methods:** The clusters of districts with the largest numbers of COVID-19 infected cases and population densities were described by using cluster analysis. Then, correlation analysis where calculated to define the strength between two parameters. **Results:** Findings of this study showed, there was a clear positive association between population density and COVID-19 infections in Peninsular Malaysia. During the study period, it was estimated that population density has a positive impact on the spread of early stage COVID-19 in Peninsular Malaysia (r= 0.752). Findings also showed there were a weak correlation between population density and COVID-19 cases in Southern region (r=0.370), Northern region (r=0.264) and East Coast region (r=0.248) as compared to Central region (r=0.917) where it have strong correlation between two variables. **Conclusion:** : This study concluded the spread of COVID-19 in Peninsular Malaysia is increasing as the population density increases.

**Keywords:** COVID-19, Population, Infectivity, Peninsular Malaysia

**Corresponding Author:**
Nazri Che Dom, PhD
Email: nazricd@uitm.edu.my
Tel: +603-32584447

## INTRODUCTION

The World Health Organization (WHO) has declared COVID-19, a deadly global pandemic caused by SARS-CoV-2, as the new coronavirus infectious disease (1,2). SARS-CoV-2 transmission modes are currently classified as viral inhalation, viral deposition on exposed mucous membranes, and virus-contaminated hands contacting mucous membranes (3). In Malaysia, the first case of COVID-19 was detected on the 25$^{th}$ January 2020 that is caused by Chinese nationals who previously had close contact with an infected person in Singapore. Day after day, the COVID-19 cases increased significantly where in March 2020, Malaysia reported the highest number of COVID-19 positive cases in South East Asia (4).

There are many factors that influence the spread of COVID-19 cases. This may be linked to the country's geographical location, socioeconomic status, level of knowledge, and health-care system in the fight against emerging and re-emerging diseases (5). Other factors such as population density, hygienic conditions of masses, length of strict lock-down and transportation of infected persons across borders influence the intensity and expansion of the COVID-19 pandemic (6,7). Population density can be defined as the number of individuals or inhabitants occupying an area of 1 km$^2$ (inhabitants/ km$^2$). It has been found out in previous study that, the spread of COVID-19 is increasing as the population density increases (8). Other than that, another study done by Carozzi, (2020) (9), stated that population density can influence the timing of outbreaks through denser location connectivity. However, on the other hand, an analysis assumed that population density is undisputed and that it cannot be a deciding factor in the spread of COVID-19. The example to support that statement are, although Macao (China) have the highest

population density (21657.70 inhabitants/km$^2$) in the world, but USA (33.86 inhabitants/km$^2$) (10) recorded the highest COVID-19 cases in the world. This is due to the effective control measures imposed by the government of China, such as they enforced the total lockdown at the beginning of the outbreak to control the transmission of the infected virus.

Thus, the aim of this research is to determine the relationship between the population density and the spread on the first wave of COVID-19 cases in Peninsular Malaysia. The results obtained from this study are vital to investigate the impacts of the higher population density to the increase of COVID-19 cases specifically in Peninsular Malaysia. Besides that, the data collection such as the total population, area of land and COVID-19 cases that gained are important to determine which location has the strong correlation between the population density and COVID-19 cases. With the entire information gathered, it can be used to predict the COVID-19 trends in specific location based on the higher population density. The outcome of this study may provide base-line information that is essential for wider studies towards better understanding of population density. This input is also useful for designing suitable strategies to control the vigorous spread of COVID-19 in location that has higher population density so, it can minimize the risk of infection.

## METHODS AND MATERIALS

This management of the observational study involves of three-stages which include; (i) Data collection; (ii) Data processing and (iii) Data Analysis). After identification of the study population, the data were collected from reliable sources and were processed before being analysis. In this study, Statistical Package for Social Sciences (SPSS) software for Windows, version 21.0 was used to apply the statistical methods analysis. These studies were adapted from Kadi, & Khelfaoui (2020) (8).

### Study population

This research aim to determine the relationship between population densities at early wave of COVID-19 cases beginning from 25th January 2020 until 31st May 2020. Peninsular Malaysia comprises of four main geographical regions: Northern region, Southern region, East Coast region and Central region. Northern area comprises of Perak, Kedah, Perlis and Penang, which have total of 30 districts. Southern region comprises three states, which are Johor, Melaka and Negeri Sembilan consist of total 20 districts. Next, East Coast region consist of Terengganu, Kelantan and Pahang with total of 29 districts. Lastly, Central area that consists of Selangor, Wilayah Persekutuan Kuala Lumpur (WPKL) and WP Putrajaya where contains of 11 districts in total. There are 3 states that do not have districts: WPKL, WP Putrajaya and Perlis. Thus, in this analysis, each of it were considered to have 1 districts. To summarise, this research observed total of 90 districts for evaluation of the relationship that consists of total between population density and COVID-19 case in Peninsular Malaysia.

### Study design

Correlational study between population density and number of cases of COVID-19 was done through analysis of secondary data. Population density and COVID-19 cases data were operated as secondary data for this study. The data were collected at reliable sources that were easily accessible (Table I). The study time was selected between 25th January and 31st May 2020 in order to determine if population density played a role in the early wave of the COVID-19 outbreak in Peninsular Malaysia.

**Table I: Datasets of correlation and cluster analysis between population density and COVID-19 cases data in Peninsular Malaysia**

| Datasets | Sources | Specifications | Attributes | Value | Being pre-processed |
|---|---|---|---|---|---|
| COVID-19 cases data in Malaysia | Official website for updated information on COVID-19 21 January 2020 to 31st May 2020 at http://covid-19.moh.gov.my/ | Cumulative reported positive COVID-19 cases | Data time interval<br>Day<br>Districts<br>Coronavirus disease | January 25, 2020 – May 31, 2020<br>1-311<br>90 districts<br>Positive local cases | Data tabulation into cumulative cases for each districts in Peninsular Malaysia. |
| Population density<br>-Number of People<br>-Total Land Area | Open Data of Population by states, administrative districts 2018 at https://www.dosm.gov.my<br>District Land Office Official Website and State Government Official Portal. | Total population and total land area each districts in Peninsular Malaysia | Data time<br>Districts<br>Unit | Total population each districts, year 2018 and latest updated data at relevant source for Total Land Area<br>90 districts<br>Inhabitants/km | Data tabulation of population density by districts |

**Data Collection**

Quantitative data gathered for this study consists of population density and cumulative COVID-19 from 25th January 2020 to 31st May 2020 in each 90 districts in Peninsular Malaysia. Population density can be defined as the number of individuals or inhabitants occupying an area of 1 $km^2$ (inhabitants/ $km^2$). Data on number of people and total land area of each districts were collected and were derived by using the formula of number of people (inhabitants) divided by total land area ($km^2$) (8). The data of number of people in each districts were generated from Department of Statistics Malaysia Official Portal-Open Data of Population by states, administrative districts 2018 (https://www.dosm.gov.my) while the total land area ($km^2$) data of each 90 districts were derived from various sources for example District Land Office Official Website and State Government Official Portal. The data information for cumulative cases of COVID-19 were generated by Ministry of Health (MOH) on official COVID-19 website on updated information of states until 31st May 2020 (http://covid-19.moh.gov.my). The summary of sources of data, specifications, attributes, value and pre-processed method best described in Table 1.

**Data Processing**

Data processing is the conversion of information into a form that is useful and desired. This conversion is using a predefined series of operations such as using computers and other data processing devices (10). After the both data were collected, the data then transferred to spread sheet; Microsoft Excel. Then, the data were arranged accordingly to districts. In order to determine the population density, the total land area divided the numbers of people by using the formula derived from Carozzi (2020) (8).

**Data Analysis**

Statistical Package for Social Sciences (SPSS) software for Windows, version 21.0 was used to analyse data collected from this study. To categorize districts based on the number of cases of COVID-19 infection during the study period, hierarchical cluster method was created and it was interpreted in the form of tree diagram (dendrogram). Then, Pearson correlation coefficient, *(r)* and the coefficient of determination, $R^2$ were determined to indicate the strength of the relationship between population density and the number of infected COVID-19 cases through data of 90 districts. Simple linear models were constructed to demonstrate the effect of population density on the number of cases of COVID-19 infection over a study period. The linear model were constructed into two models: *Model 1*: Data from 90 districts to evaluate the relationship and degree of impact of population density on the early stages of spread of COVID-19 in Peninsular Malaysia. *Model 2:* Data from the Central region that consist of 11 districts. Central region generated the highest population density as compared to other geographical region by according to the $R^2$ value.

**RESULTS**

**Classification of districts according to early wave cases of COVID-19 infections in Peninsular Malaysia.**

This study was undertaken in thirteen states comprise of 90 districts in Peninsular Malaysia. A cluster hierarchy is constructed by the agglomerative hierarchical clustering algorithms, which is usually seen as a tree diagram called a dendrogram. Dendrogram created cluster of early stages of COVID-19 cases by districts in Peninsular Malaysia. The dendrogram consists of a horizontal axis and a vertical axis. Horizontal axis indicated COVID-19 cases and vertical axis indicated 90 districts of 13 states in Peninsular Malaysia. In data analysis, the use of the classifications method were allowed to classify districts according to the degree of cases infected with COVID-19 within the study period. The results obtained are shown in Figure 1.

Figure 1 indicates dendrogram of cluster analysis of COVID-19 cases by districts in Peninsular Malaysia. By referring to the dendrogram present in Figure 1, the study population of 90 districts were divided into three groups. The first group contains 86 districts starting from of Kota Setar to Klang with an average of 38 infected cases. Next, the second group includes 3 districts which are Seremban, Petaling and Hulu Langat, with an average of 449 infected cases. Lastly, the third group contains only 1 district of WPKL, with an average of 2028 infected cases.

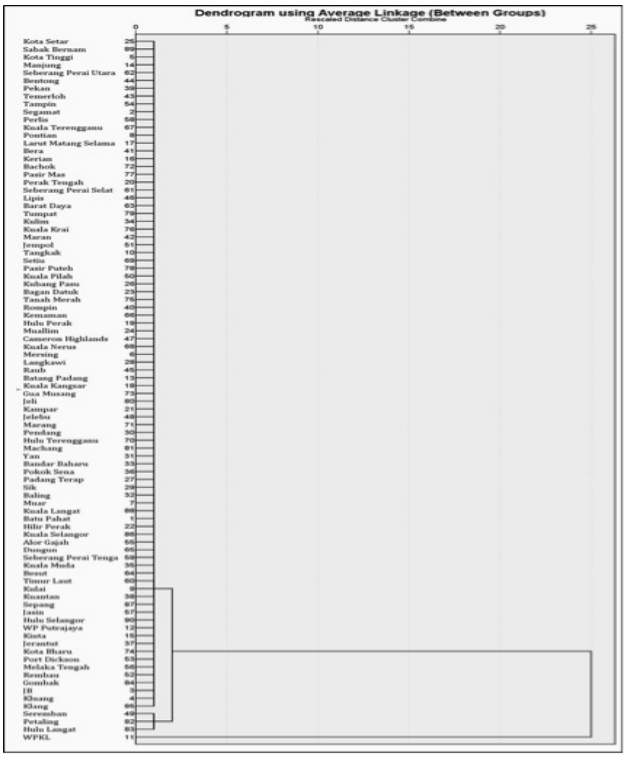


**Figure 1: Cluster analysis of early wave of COVID-19 cases by districts in Peninsular Malaysia**

**Classification of districts according to early stages cases of COVID-19 infection and population density in Peninsular Malaysia.**

The cluster method also allowed us to classify districts by cases of COVID-19 and population density for the study period. Figure 2 highlights the cluster analysis of COVID-19 cases and population density by districts in Peninsular Malaysia that are divided into four groups. The first group contains 74 districts (from Batang Padang to Seremban) with an average population density of 181.30 (inhabitants/km$^2$) and an average of 36 infected cases. While, the second group includes 13 districts (from Melaka Tengah to Hulu Langat) with an average population density of 1360.95 (inhabitants/km$^2$) and an average of 129 infected cases. The third group contains only 2 districts of Timur Laut and Petaling, with an average population density of 4160.19 (inhabitants/km$^2$) and an average of 233 infected cases. The fourth group consists of 1 district only which is WPKL with an average population density of 7088.45 (inhabitants/km$^2$) and an average of 2028 infected cases. From the findings, the fourth group contributed the highest average of population density and COVID-19 cases although it contains only 1 district of WPKL.

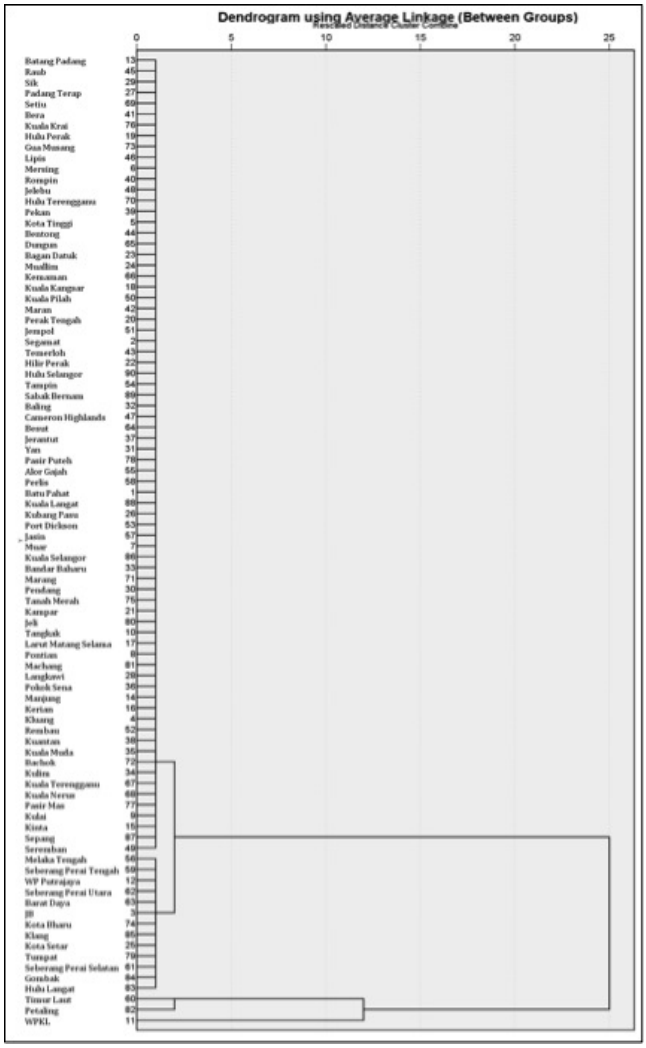


**Figure 2: Cluster analysis of early wave of COVID-19 cases and population density by districts in Peninsular Malaysia**

**Correlation between population density and COVID-19 cases of districts in Peninsular Malaysia.**

To further explore and understanding on the relationship between population density and COVID-19 cases, Pearson correlation coefficient (*r*) and the coefficient of determination, $R^2$ were determined to know the strength and the nature of the correlation between population density and the number of cases infected. Generally, the positive sign of *r* indicates that the positive association between population density and the number of infected cases, while the negative sign of R indicates the negative association between population density and the number of infected cases of COVID-19. According to (11), when $r > 0.7$, the relationship between two variables is considered strong, while if $0.5 < r < 0.7$, the strength of relationship is moderate. Weak strength of relationship when $0.3 < r < 0.5$ and very weak or no strength of relationship when $r < 0.3$.

Table II represents the correlation coefficient, (*r*) and coefficient of determination, ($R^2$) between population density and COVID-19 cases in all districts and four geographical region Peninsular Malaysia: Southern region, Northern region, East Coast region and Central region. Southern region consists of 20 districts in 3 states of Johor, Negeri Sembilan and Melaka. Northern region consists of 30 districts in 4 states of Perak, Kedah, Pulau Pinang and Perlis. Next, East Coast region contains 29 districts in 3 states which are Kelantan, Terengganu and Pahang. Lastly, Central region that consist of 11 districts in 3 states of Selangor, WPKL and WP Putrajaya. Based on the Table II, it is noted that the correlation coefficient between population density and COVID-19 cases are differ from one region to another. The highest and strongest correlation is in the Central region ($r = 0.917$, $R^2 = 0.841$), followed by in Southern region ($r = 0.370$, $R^2 = 0.137$), marginally less strong in the Northern region ($r = 0.264$, $R^2 = 0.070$) and lastly in the East Coast region ($r = 0.248$, $R^2 = 0.062$). Central region that consist of 3 districts of WPKL, WP Putrajaya and Selangor contributed the highest total average number of COVID-19 cases and as compared to other region.

**Table II: Correlation coefficient, R and coefficient of determination, $R^2$ between population density and COVID-19 cases.**

| Geographical regions | COVID-19 cases | |
|---|---|---|
| | **r** | **$R^2$** |
| All the districts | 0.752 | 0.565 |
| Southern region | 0.370 | 0.137 |
| Northern region | 0.264 | 0.070 |
| East Coast region | 0.248 | 0.062 |
| Central region | 0.917 | 0.841 |

**The degree of influence of population density on the propagation of the COVID-19 cases in Peninsular Malaysia.**

A scatter plot were used to show the relationship between two quantitative variables. Simple linear model were constructed to illustrate the impact of population density on the number of cases of COVID-19 infection. According to (11), the slope, provides information on the strength of the relationship. The strongest linear relationship occur when the slope, m is 1. Two clouds of points were generated to illustrate the degree of influence of population density on the spread of COVID-19 in Peninsular Malaysia. Figure 3 shows the population density and cases of COVID-19 for all districts in Peninsular Malaysia started from 25 January 2020 until 31 May 2020. The Central region was chosen as it has strong correlation between population density and COVID-19 cases which it has the highest R values as compared to other geographical region (Figure 4). By referring to the figure, an upward slope on the scatterplot were formed as it represented the positive correlation between two variables. It is noted that there was a significant effect of the population density on the increase in infections of COVID-19. It is because each time the population density increases by 1 individual/ $km^2$, there is an increase in the number of cases of COVID-19 estimated at 90 districts of Peninsular Malaysia.

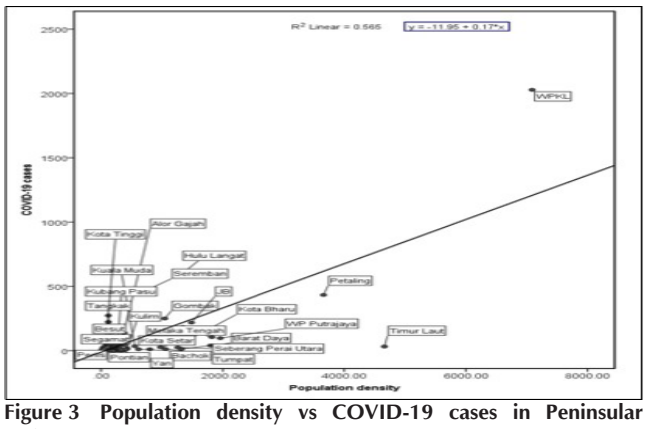


**Figure 3 Population density vs COVID-19 cases in Peninsular Malaysia**

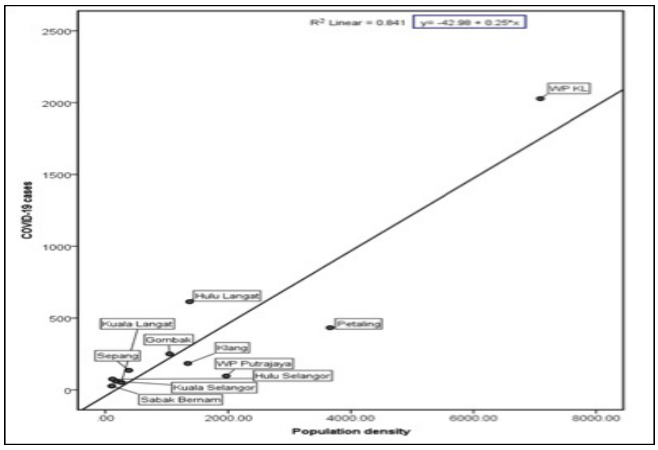


**Figure 4 Population density vs COVID-19 cases in Central region**

## DISCUSSION

Population density in one area can be affected by many factors for instance social, economic and political factors. Countries with stable governments, numerous job opportunities and areas rich in resources tend to have a high population density (12). There are advantages and disadvantages for an area that have high population density. Previous study indicated that population density had contributed to the spread of communicable diseases such as measles (9). A higher density of the population may mean more interactions among people and thus a higher risk of contagion. COVID-19 is spreading over the world including Malaysia. COVID-19 outbreak in Malaysia initially only involve 22 cases then increasing to significant number over the country (2).

In this study, a correlation analysis were conducted to confirm the association between population density and COVID-19 cases in Peninsular Malaysia. Findings of this study showed weak to strong association strength were observed between the two variables. Strong correlation was detected for the relationship between population density and COVID-19 cases in all districts in Peninsular Malaysia and specifically in Central region. Central region that consists of 11 districts of 3 main states; WKPL, WP Putrajaya and Selangor (9 districts) has contributed the highest average of COVID-19 cases at early stage of the pandemic and also the highest average population density as compared to other geographical region. Among the 11 districts, WPKL has reached four digits cumulative local COVID-19 cases, which are 2028 cases, and it also has highest population density of 7088.45 inhabitants/$km^2$.

Social distancing had an effect on the spread of COVID-19 cases especially in the area that have high population density. If the infectious disease progressed, areas with a higher population were more prone to see a higher number of cases. As a result, while population density might not be a major factor in the early phases of outbreaks, it is influential later on. The research provided here offers compelling proof to support this argument. Furthermore, our findings clearly indicated that, in addition to taking population density into account at all stages of pandemic growth, local variables such as the prevalence of some facilities in local communities could play certain roles in the early stages of the pandemic. Social distancing can be defined as the practice of purposefully reducing close contact between people. This study found that WPKL had recorded the highest number of confirmed local COVID-19 cases (2,028 cases) to date 31st May 2020 and it found to have the highest and strong correlation between population density and COVID-19 cases among all study population. In addition, the Government had announced that two parliamentary constituency in WPKL as red zones which are Lembah Pantai and Cheras on March 2020. Many research has proved that population density have a

strong relationship with the spread of COVID-19 since this virus are easily spread at crowded places with poor social distancing. Previous study conducted by Bhadra et al (11), stated that people living in areas of high population density, such as urban or metropolitan cities, are more likely to come into close contact with others, so any infectious disease is expected to spread rapidly in dense areas as high-density cities may offer more opportunities for crowding. WPKL is urban areas that often be a places with a lot of social interactions, crowded living and close contact. So, there are not impossible for the virus to widely spread to other people in very short time. This statement can be supported by the COVID-19 cases incident that happen in Sri Petaling, Kuala Lumpur where religious event participated by 16,000 people led to an exponential rise in COVID-19 cases. Bringing people together has many positive social and economic consequences, but it does have some detrimental consequences. When population density becomes too high, disease transmission can increase, sometimes more than proportionately due to population interaction. Densely inhabited regions are often ideal for the emergence and spreading of such respiratory epidemics. Frequent contacts between people whose physical presence rises non-linearly with the density of individuals in any given location result in the rapid spread of infectious diseases to a broad population.

Other than that, when individuals are in close proximity to one another, the underlying mechanism for the associations of social distancing and population density for COVID-19 is possibly associated with increased droplet transmission and potentially airborne transmission, as spread occurs when an infected person coughs, sneezes, or communicates, and droplets from their mouth or nose are released into the air and to people nearby (11). As mentioned by (10), throughout the pandemic, communities consisting of more essential workers who have sustained food service, public transport and health care systems are understandably less interested in social distancing and are at greater risk of having infection (12-14).

Population density contributed the positive effect on increasing the number of cases of COVID-19 virus at early stage in Peninsular Malaysia. As population density increases, the number of infections increases, especially in Central region. Social distancing is the important factor related with population density. It can be concluded that, social distancing measures had an impact on reducing the transmission of COVID-19. When population density increases, the social distancing will be poor (12, 13).

## CONCLUSION

Overall correlation of population density and early stages of COVID-19 cases in Peninsular Malaysia have strong positive association. While, Central region contributed the strong association between two variables by according to geographic region. Higher the population density, the closer people to public areas, thus the higher the risk of virus transmission. As a result, some approaches to control and prevent the spread of the virus, especially in relation to the population density factor, must be considered. It is critical to raise public awareness of social distancing, especially in high-density areas like the Central region in order to cut the spread of the virus.

## ACKNOWLEDGEMENTS

First and foremost, the authors would like to express deep and sincere gratitude to all organizations were involved in this project. We also would like to offer thanks to the Faculty of Health Sciences, Universiti Teknologi MARA (UiTM) for the technical assistances rendered.

## REFERENCES

1 Salim N, Chan, WH, Mansor S, Bazin NEN, Amaran S, Faudzi AAM, Shithil SM COVID-19 epidemic in Malaysia: Impact of lock-down on infection dynamics. medRxiv. 2020;
2 Elengoe A. COVID-19 outbreak in Malaysia. Osong public health and research perspectives. 2020; 11(3): 93-7
3. Scientific Brief: SARS-CoV-2 Transmission, (2021, May 7). Center for Disease Control and Prevention. https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/sars-cov-2-transmission.html
4. Kejela T. Probable Factors Contributing to the Fast Spread of the Novel Coronavirus (COVID-19) in Ethiopia. Journal of Infectios Disease Epidemiol. 2020; 6, 169.
5. Collivignarelli MC, Abba A, Bertanza G, Pedrazzani R, Ricciardi P, Miino MC. Lockdown for CoViD-2019 in Milan: What are the effects on air quality?. Science of the Total Environment. 2020;732:139280.
6. Sharif N, Dey SK. Impact of population density and weather on COVID-19 pandemic and SARS-CoV-2 mutation frequency in Bangladesh. Epidemiology & Infection. 2020; 1-34.
7. Kadi N, Khelfaoui M. Population density, a factor in the spread of COVID-19 in Algeria: statistic study. Bulletin of the National Research Centre. 2020; 44(1): 1-7.
8. Carozzi F. Urban density and COVID-19. Centre of Economic Performance. 2020
9. Anser MK, Yousaf Z, Khan MA, Sheikh AZ, Nassani AA, Abro MMQ, Zaman K. Communicable Diseases (Including COVID-19)- Induced Global Depression: Caused by Inadequate Healthcare

Expenditures, Population Density, and Mass Panic. Frontiers in public health. 2020; 8: 398
10. Tarwater PM, Martin CF. Effects of population density on the spread of disease. Complexity. 2001; 6(6): 29-36.
11. Bhadra A, Mukherjee A, Sarkar K. Impact of population density on Covid-19 infected and mortality rate in India. Modeling Earth Systems and Environment. 2021; 7(1): 623-629.
12. Garnier R, Benetka JR, Kraemer J, Bansal S. Socio-economic disparities in social distancing during the COVID-19 pandemic in the United States. 2020; medRxiv
13. Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non- pharmaceutical interventions on COVID-19 in Europe. Nature 2020: 1–8.
14. Tammes P. Social distancing, population density, and spread of COVID-19 in England: a longitudinal study. BJGP open. 2020; 4(3).