# FACTORS ASSOCIATED TO HOTSPOT OF DENGUE

**BY**

**NOR FARISHA BINTI MUHAMAD KRISHNAN**
**(2015283646)**
**NUR AZERA BINTI AHMAD**
**(2015430712)**
**NUR SHAHIRAH BINTI DOLLAH**
**(2015238474)**

# UNIVERSITI TEKNOLOGI MARA

# FACTORS ASSOCIATED TO HOTSPOT OF DENGUE

**NOR FARISHA BINTI MUHAMAD KRISHNAN**

**NUR AZERA BINTI AHMAD**

**NUR SHAHIRAH BINTI DOLLAH**

Dissertation submitted in fulfilment

of the requirements for the degree of

**Bachelor of Science (Hons) Statistics**

**Faculty of Computer and Mathematical Sciences**

**June 2018**

**APPROVED BY:**

**(PUAN NOR AZIMA ISMAIL)**

**Supervisor**

**Faculty of Computer and Mathematical Sciences**

# ABSTRACT

Dengue is an illness transmitted by mosquitoes, causing sudden high fever, severe headache, rashes, muscle and joint pain, pain behind the eyes and bleeding in some cases. Generally there are two classification of dengue fever which are dengue fever (DF) and dengue haemorrhagic fever (DHF). In Malaysia, dengue has become a major problem to the country as it lead to the serious dengue fever. The dengue fever is listed as the most prevalent disease in the country with a ratio of 328.3 cases per 100,000 population The drastically increase of dengue cases happened proved in 2014 make the researchers have more interesting to do a research about the main factors related to dengue cases in 2014. From that, the researcher can identify significant relationship between the available factors like ages, gender, occupation and locality toward hotspot of dengue that can categorize from dengue cases. The locality which the dengue occur even after 30 days of the first dengue case is called as hotspot. Logistic regression was used in order to determine whether there is significant relationship between ages, gender, occupation, locality toward hotspot of dengue (YES, NO). Then the chi square test of independence used to know whether there is an association between gender and hotspot of dengue and also identifying if there is association between age category and hotspot of dengue. The result of the logistic regression showed that only gender does not give significant effect to hotspot of dengue with the percentage of correct classification is 78.7% the variable while the value for error rate of the model is 21.3%. Based on chi square test of independence, there is association between gender and hotspot of dengue, as well as age category and hotspot of dengue. However both gender and age category showed weak association to hotspot of dengue.



***Keywords*** – Chi Square, Dengue Fever, Dengue Cases, Hotspot, Logistic Regression

# ACKNOWLEDGEMENT

In The Name Of Allah, the Most Gracious the Most Merciful

We would like to express our deepest appreciation to all those who provided us the possibility to complete this report. A special gratitude we give to our supervisor, Puan Azima binti Ismail, whose contribution in stimulating suggestions and encouragement, helped us to coordinate our project especially in witting this report. Furthermore we would like also to acknowledge with much appreciation the crucial role of the staff of Vector Borne Disease Control (VBDC) Kota Bharu, Kelantan who gave the permissions to use secondary data and gave the required information to complete the project 'Factor Associated to Hotspot of Dengue'. We are also very thankful to our parents who give fully support including financial, motivation and so on for us to complete this project. Without them maybe we could not complete this project. A special thanks to our classmates that help us to assemble the parts and gave suggestion about the project. We have appreciated the guidance given by other supervisor as well as the panels especially in our project presentation that has improved our presentation skills thanks to their comment and advices.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF APPENDICES

# CHAPTER ONE
# INTRODUCTION

## 1.1 INTRODUCTION

Dengue is an illness transmitted by mosquitoes, causing sudden high fever, severe headache, rashes, muscle and joint pain, pain behind the eyes and bleeding in some cases. Generally there are two classification of dengue fever which are dengue fever (DF) and dengue haemorrhagic fever (DHF). In many parts of the world, dengue is fast emerging pandemic-prone viral disease. Usually dengue affects more affluent neighbourhoods in tropical and subtropical countries and flourishes in urban poor areas, suburbs and the countryside. Dengue also is a mosquito-borne viral infection that sometimes causing a potentially lethal complication called severe dengue and usually causing a dangerous flu-like illness. The incidence of dengue has grown dramatically around the world in recent decades. According to World Health Organization, dengue cases has increased 30-fold over the last 50 years. Up to 50-100 million infections are now estimated to occur annually in over 100 endemic countries, putting almost half of the world's population at risk.

In Malaysia, dengue has become a major problem to the country as it lead to the serious dengue fever. The dengue fever is listed as the most prevalent disease in the country with a ratio of 328.3 cases per 100,000 population. "Six states have reported a significant increase in the number of dengue cases in the third week of January. The states are Selangor, Perak, Kedah, Pahang, Penang and Kelantan" (Bernama, 2017).

Dengue study raises concern in Malaysia as the various phases of dengue attack that could lead to fatalities. Based on Academy of Medicine and Ministry of Health Malaysia (2003), Most of the cases are reported among the urban population (70– 80%) with the highest incidence in the working and school going age group, which related to the relatively high Aedes Index in construction sites, schools and factories. It's been reported that during the month of July, August and September is the maximum number of cases occurs throughout the year.

Dengue is a virus that caused by the mosquitoes which contribute to the increasing of morbidity and mortality in Malaysia as well as other countries. However recently dengue fever has become a major issues in Malaysia as it can cause death. This kind of fever does not choose the community either kids or adults. This kind of fever does not choose the community either kids or adults. In other countries this cases are never fade of. It is increasing from year to year. This might because of several factors which can contribute to the increasing of DF and DHF.

In the article "Dengue cases in Kelantan increase 68 percent" (Thestar.com.my,2013) it was reported that in Kelantan, 68 percent increase in dengue cases as 1,245 dengue cases were reported in 2012, compared with in 2011 with 743 cases. In 2014, Kelantan has recorded the highest number of dengue cases with 14312 ("Laman Utama iDengue", 2018). Hence, with the increasing and higher dengue cases in Kelantan, the study of factors influence the hotspot of dengue is very significant as it can give an early warning and awareness to the people after identifying the hotspot of dengue.

In order to make the fight against dengue, hotspots of dengue need to be identified. "They need to identify the 'real' hotspots where the wild mosquito population is transmitting the disease away from residential homes after acquiring the infection from people who have the virus circulating in their blood stream before the onset of the disease, which is often mild in such carriers", said Malaysia Association of Environmental Health (MAEH) expert, Veeramohan Supramaniam. He continued, "The dengue vector mosquitoes especially the outdoor breeding *Stegomyia albopicta* (Aedes) mosquitoes will continue to multiply as long as those hotspots are not identified and destroyed", based on The Malay Mail Online. He also suggested, natural breeding places that posing high risks of man-mosquito contact can be cleared in a sustainable manner in a way that local authorities helping urban communities by informing and educating residents in identifying outdoor hotpots. (Dzukifly, 2014)

## 1.2 PROBLEM STATEMENT

There are many research done on factors that contribute in increasing of dengue cases in Malaysia. Sometime different places have different factors of dengue because it is depend on their lifestyle, environment and others. Some of people are ignore about dengue cases either the cases are increase or decrease neither in their place and also other places because it is not happened to them. The drastically increase of dengue cases happened proved in 2014 make the researchers have more interesting to do a research about the main factors related to dengue cases in 2014.

The locality which the dengue occur even after 30 days of the first dengue case is called as hotspot. Not only that, since the dengue cases in 2014 is the highest, the mortality rate on that year also increase make all Malaysian more worried and did not know the real reasons of this incidence. The increasing incidence of dengue actually has a significant impact on global healthcare services and other people. It is because dengue is easily spread by the Aedes Aegypti, a domestic, day-biting mosquito that prefers to bite humans. The escalation of the dengue cases give public uneasiness since dengue may cause death. Just similar to scenario that happen in 2014 which the cases is the highest. Total of the cases in that year is around 14000 and above. Dengue fever does not choose person, it can be to males or females, children or adults and so on.

One of the reasons increasing of the dengue cases is may be lacking of community support in abolishing Aedes mosquito breeding grounds whether in their house or surrounding like anti- dengue *gotong – royong* activities. Recently, it can be seen that mounds of garbage were left unattended and the blocked drains which then resulted in pools of stagnant water. This will rise the breeding of the Aedes mosquitoes. This research also may be will give an early warning and awareness to the people after identifying the hotspot of dengue. Thus the public awareness can be inculcate among the community in order to reduce the dengue fever and dengue haemorrhagic fever.

## 1.3 RESEARCH OBJECTIVES

In order to succeed this study there are some objectives to be achieved.

1. To determine whether there is significant relationship between ages, gender, occupation, locality toward hotspot of dengue (YES, NO).
2. To know whether there is an association between gender and hotspot of dengue.
3. To identify if there is association between age category and hotspot of dengue

## 1.4 RESEARCH QUESTIONS

There are several research questions that might be exist based on this study.

1. How ages, gender, occupation, locality will influenced into categorized as a hotspot of dengue?
2. Does gender and hotspot of dengue associated?
3. Is there an association between age category and hotspot of dengue?

## 1.5 RESEARCH HYPOTHESIS

$H_1$: There is a significant relationship between the ages, gender, occupation, locality and hotspot of dengue.

$H_1$: There is an association between gender and hotspot of dengue.

$H_1$: There is an association between age category and hotspot of dengue.

## 1.6 SCOPE OF STUDY

This study focus on dengue cases in Kelantan in order to know whether there is significant relationship between age, gender, occupation and locality towards hotspot of dengue. There are too many district in Kelantan for instance, Kota Bharu, Kubang Kerian, Pasir Mas, Pasir Puteh and so on. The scope of this study cover on every district in Kelantan to determine whether the locality will influenced the hotspot of dengue. In this case locality include either urban or rural area.

## 1.7 LIMITATION OF STUDY

There are some limitations in the study when completing this research. First and foremost, limitation in scope of study make it more complicated to understand it. This research is about factors that associates with dengue case which is usually reported by a medical researcher. There are a lot of terms that we have to find and also understand deeply since it were not in our recently study terms or familiar using that terms.

Not only that, in this study the researchers just focus a few of factors only which are dengue cases, occupation, mortality, gender and age towards hotspot of dengue. But, another factors also maybe will affect to this result either directly or indirectly. It will be uncontrolled variables which appear in this research.

Lastly, the secondary data used need to be understand more details. Since, the data given is raw data, the researchers have to clean the data to make it more clearly and related to this study. Hence it need to find as much info from previous research study in order to relate with this study. Finding of the previous study that had been done by the researcher is quite difficult due to the method used is Logistic Regression to obtain the hotspot of dengue in Kelantan.

However these of all the limitations have taught to never give up and try as harder as can to achieve the target and complete this study before submission date.

## 1.8 SIGNIFICANCE OF STUDY

The significance of this study is an attempt to identify the variables which are ages, gender, occupation, and locality influence the hotspot of dengue. This study is done in Kelantan state. Hotspot can be defined as the locality which the dengue occur even after 30 days of the first dengue case. As the hotspot can be detected, the preventive measure can be taken to overcome DF and DHF such as get rid of the places where mosquitoes breed likes plastic container, tyres, vase and so on, having *gotong – royong* and the government should give penalty to those who are not aware about cleanliness.

Next the most important is must be able to determine the place with the highest cases of dengue in Kelantan by using logistic regression method. Generally dengue fever can be divided into two which are dengue fever (DF) and dengue haemorrhagic fever (DHF). Both type of dengue fever will cause death as it is virus from Aedes sp. Mosquitoes. The place with the highest dengue cases may also high in mortality or maybe not. Thus the citizen should more aware about this dengue fever as in can be prevent from occur more rapidly.

This study will also give important information and results to the government especially for Vector Borne Disease Control (VBDC) from this research. From the result of analysis based on this study VBDC may use those information to take preventive measure to reduce dengue cases. Besides those info will aid VBDC to identify factors that contribute most to hotspot of dengue.

## 1.9 RESEARCH SCHEDULE

Table 1.1: Time Schedule of the Research

| No | Research Activities | Year 2017 | | | | Year 2018 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Sep | Oct | Nov | Dec | Jan | Feb | Mac | Apr | May | Jun |
| 1 | Chapter 1 (Introduction) | | | | | | | | | | |
| 2 | Chapter 2 (Literature Review) | | | | | | | | | | |
| 3 | Chapter 3 (Methodology) | | | | | | | | | | |
| 4 | Chapter 4 (Analysis and Finding) | | | | | | | | | | |
| 5 | Chapter 5 (Conclusion and Recommendation | | | | | | | | | | |

***subject to change**

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 INTRODUCTION

Dengue is a mosquito-borne contagion which becoming a main global public health concern and significance in a few current era Er, Rosli, Asmahani, Mohamad and Harsuzilawati (2010). There is no specific treatment for dengue, however, with frequent and appropriate medical care; the lives of patients with the more serious dengue haemorrhagic fever (DHF) could be saved. The only way to prevent the transmission of dengue virus is to combat the carrier. Currently, dengue haemorrhagic fever (DHF), sometimes a potentially lethal complication, affects most Asian countries and has become a leading cause of serious illness, hospitalization and death among children in the region.

Dengue cases has become a major problem for this few decades as the number of patients who are prone to dengue are never fade off. According to Statistics of Department, it has been increased from 357.49 per 100,000 population in 2014 to 392.96 per 100,000 population in 2015 the incidence rate for dengue fever. There are three states which are Selangor (1065.93 per 100 000 population), W.P. Kuala Lumpur (443.25 per 100 000 population) and also Johor that exceed the national level of incidence rate for dengue fever record.

A research done by Liew, Khoo, Ho, et al (2016) explained the population used in this research is patients in the Malaysian National Dengue Registry of 2013. Variables that focus on this article are associations between sociodemographic and clinical variables with the outcome. Sociodemographic such as age, gender, nationality, ethnicity and places of dwelling which are rural or urban.

## 2.2 DENGUE CASES

Mallhi et al. (2017) point out about positive reverse transcriptase PCR result, presence of dengue immunoglobulin M and G antibodies in acute-phase serum by ELISA (Pan Bio Dengue IgM ELISA, Dengue IgM Dot Enzyme Immunoassay, SD Dengue IgM and IgG capture ELISA Kits; Standard Diagnostics, Korea) and at least fourfold increase of dengue-specific haemagglutination inhibition titers in convalescent serum compared with acute-phase serum. The serum samples were also tested for dengue-specific NS1 (pan-E Early dengue ELISA kit by Panbio, Australia, and Platelia dengue NS1Ag assay by Bio-Rad Laboratories, USA) either it is dengue cases or not. The suspected cases were confirmed when at least one of the criteria above is achieved.

Ho and Ting (2015) stated that the data is provided by the Ministry of Health Malaysia, starting from the week beginning 22 February 2010 up to the week beginning 26 Jun 2015. The data collected represents the accumulated number of dengue cases reported in Malaysia that collected for a longer period of time (over 5 years).The variable that used in this article are very contribute to our research too. Total cases in dengue spread in all state in Malaysia from 2010 until 2015 are the main point of this research. In fact, by using time series analysis to predict dengue fever occurrence will be useful in the long period arrangement of dengue fever control and prevention plan (Zainudin & Shamsuddin, 2016).

## 2.3 HOTSPOT

One of the definition of hotspot based on Oxford English Dictionary is a place of significant activity, danger or violence. In context of hotspots of dengue, according to Ministry of Health in Malaysia, the dengue hotspot is a localized epidemic that lasting more than 30 days from the date the pandemic started. In simply words, hotspot of dengue means particular area has from the date the first case started. It also can be proved through I dengue website which explained more about dengue.

Based on study conducted by Jeefoo, Tripathi, and Souris (2010), state that hotspot generally defined as a condition which specify some form of clustering in spatial distribution. In fact the clusteres of cases which happen randomly may also affect on the dispersion an contagious disease. They also stated in their study that the detection

of hotspot may be useful though the global pattern is not clustered. There is also a similar study conducted by Nash and Tong (2014), which stated that mean of lifespan of female mosquito is between 14 to 30 days and can fly for average of 30 – 50m a day. This indicate that mosquitoes move around about 240 – 500m in range for the lifetime. They also said that this flight range will give to limited local spread of disease. Hence, this why the scale is lied between that range of days is the scale at which hotspot of dengue need to be identified, so that overal transmission pattern are able to be determined.

The increasing in frequency of the occurrences, and involvement of new are that were formerly unaffected are being spot recently causing dengue epidemiology to evolve rapidly. The new areas stated are small towns, land schemes and villages. The dengue cases in cities and big towns tend to be out of control and becoming the hotspot. It is a term that used to determine the outbreak happening for more than 30 days (Mohamed, Rahman, & Shaari, 2015).

## 2.4 OCCUPATION

Chen et al, 2016 has led a study to identify whether people's socio-economic give impacts to the dengue cases. At the end of his study it can conclude that the proportion of occucpations have give significant differences towards dengue fever cases and control groups.. From his study the result showed that merchants (19.9%), office workers (17.1%), and retired persons (16.4%), compared to unemployed (25.2%), retirees (21.8%), and office workers (21.4%) for community controls which means that unemployed has the highest rate of getting dengue fever. Ramzan, Ansar, & Nadeem (2015) have study the correlation of Socio-demographic charcteristics. It showed that the employed person without specific qualifiacation has the highest percentage for getting dengue fever with 42.2% compared to unemployed (18.4%) and employed with specific qualification (36.1%).

A study was conducted in Thailand to know about risk factor of primary school students towards dengue infection. Based on the result of the study, it showed that students has low basic knowledge on dengue infection and their household environments which are actually very important, Suwanbamrung,C. et al, (2012). Besides that, Teurlai,.M et al,(2015) had studied about socio-economic factors

associated with dengue fever spatial heterogeneity and the results showed the mean percentage of unemployed people is a variable that highly correlated with people's way of life. The recent study that aim to identify environmental and socioeconomic factors associated to dengue incidence was done by Zellwege,M.R,(2017), found that "dengue incidence rates were higher in neighbourhoods with higher unemployment (IRR = 1.25, 95%CI: 1.14–1.36), higher vegetation coverage (IRR = 1.14, 95%CI: 1.04–1.24), higher percentage of old houses (IRR = 1.12, 95%CI: 1.03–1.21), and lower percentage of apartments (IRR = 0.91, 95%CI: 0.84–0.98). For 2012–2013, higher dengue incidence rates were associated with lower revenue (IRR = 0.88, 95%CI: 0.82–0.95), lower percentage of apartments (IRR = 0.91, 95% CI: 0.84–0.97) and higher percentage of cement lodgings (IRR = 1.13, 95% CI: 1.04–1.21)". Multivariate analysis was used in their study to described closely the dengue incidence rates.

## 2.5 MORTALITY

Based on study by Zainudin and Shamsuddin (2016), mentioned that dengue cases is high in Selangor especially in Petaling Jaya and Hulu Langat in year 2010 until 2015. As both district are near to each other, the spreading of dengue becoming quite fast. It can be supported by the study by Cheong et al. (2014) explained that the dengue outbreaks covered a wide area from the north-eastern and south-eastern regions to the central. Besides based on their findings, the dengue outbreaks illustrate a significant spatio-temporal pattern at the address level. Pinto et al, 2016 has identify the factors associated with death among patients with severe dengue, in Amazonas from 2001 to 2013 in their study. They stated that patient care has not reached the level expected in any of the evaluated services, and that the recommendations from the Ministry of Health for the management of dengue cases are not being followed.

Mortality is a death within 14 days after admission from hospital (Mallhi et al, 2017). The study conducted by Liew et al. (2016), the result shows that growing age, female gender, nausea or vomiting, bleeding, lethargy or restlessness, severe plasma leakage and shock were associated with mortality. All those symptoms are related with dengue-associate mortality in the overall population. They also used multivariate analysis method to measure associations between socio-demographic and clinical variables with the outcome. The outcome measure was dengue-related mortality.

Pooransingh, Teelucksingh and Dialsingh (2016) have studied about the characteristics of the patients who died from dengue in Trinidad to determine if there were any identifiable factors that were associated with death from dengue.

## 2.6 GENDER

Gender and age have to examine separately since the effect of both factors for infectious disease are different. In relation the study is also interested in gender which one is high infection either in male or female in Kelantan. Before the study have done, the researchers have found a few of articles about gender that related to dengue cases. Chew et al. (2016) had presented data male and female were 63% (68) and 37% (40) respectively. It is because they have high probability to expose by Aedes mosquitoes. Due to male have the tendency to travel and do more outside works than the female favours.

Based on study conducted by Rafique, Saqib, Munir, and Qureshi (2017) also stated that majority of males were prone to asympotic dengue infection compared to females. They also give the prove that dengue is higher in males than females because, most females covers most of their body and rather stay indoor than outdoor. Their study can be supported by Anker and Arima (2011). They indicated the males with age between 5 – 14 year old is exceed than males who are above 15 years old. High cases in males for the reason that men are exposed to the dengue carrying mosquitoes in daytime hours whether at the workplace or going to and home from work.

## 2.7 LOCALITY

Several studies shows that the dengue has spread to the rural areas because dengue haemorrhagic fever affects the population in urban areas (Siregar, Djadja, & Arminsih, 2018). Based on their study, also explained that the dengue has spread from big cities, which the urban areas act as pools of the virus to the area with small societies. This showed that the occurrence of disease is contributed by that movement. The study conducted by Sattar, Lukose, Nellis, Pham and Sadanand (2016) stated that as the density of Aedes Aegypti mosquito population increase, the human populations of human would also increase which then contribute to the transmission of the dengue in urban area.

Based on this study, the Kelantan state will be observed. According to (Brinkhoff, 2017), the Kelantan state has 10 districts with the total population of 1,539,601 from the census in year 2017. The capital city of Kelantan is Kota Bharu with the highest population 491,237. The districts in Kelantan are Kota Bharu, Pasir Mas, Tumpat, Tanah Merah, Machang, Jeli, Pasir Puteh, Bachok, Kuala Krai and Gua Musang. Based on study conducted by Salleh and Mohd Nasir (2014), they detailed that Kota Bharu which is an islamic city, represent a symbolic icon of urbanisation of the islamic state. They also state that Kota Bharu is state capital and has become the dominant point of Kelantan's business and administrative actions. Since Kota Bharu is the capital city, this study used Kota Bharu as the reference.

## 2.8 AGE

The researchers are also interested to study about age because there are a lot of assumptions and studies about dengue related to factor of age. This disease are will affected to all without any level since either a baby or any group of age. From that, the researchers want to determine which group of age (in range) are the most contribute into dengue cases. The research had done by Chew et al. (2016) in article an epidemiological perspective study that the higher prevalence of dengue in the age group of 20-29 years might be due to more outdoor activities of this young to adult group which allowed them to be exposed to Aedes mosquitoes. In the years of 1997 until 2008, the majority of the reported dengue cases occur to both males and females with age of greater than 15 years old were 76% to 82% (Anker & Arima, 2011).

Not only that, the different studied had found in Thailand from 2005 to 2013 by Tiawilai et al. (2015) , there were 15 dengue patients, aged over 60 years old, who were admitted to Photharam Hospital, Ratchaburi, Thailand. The age range was between 60-to-87 years old, with a mean age of 68.7 years and a median age of 66 years. Older age has previously been reported to be a risk factor for mortality in patients with dengue fever (DF) or dengue hemorrhagic fever (DHF) because the co-morbidities associated with ageing and waning immunity pose a substantial risk for fatality in elderly patients with active infection. But no deaths were seen in these elderly patients with dengue disease, indicating early recognition and effective management of these dengue patients. The trend towards higher age in dengue patients is a problem of

concern, which needs further elaboration. From all of articles, it can give an ideas to other readers which population are the most affected in overall.

According to StatCan (2018) age is categorized into 4 groups which are children between 0 – 14 years old. For youth groups, age is between 15 until 24 years old followed by adults between 25 – 64 years old. The age between 65 until 90 years old and over is categorized as seniors.

# CHAPTER THREE

# METHODOLOGY

## 3.1 INTRODUCTION

Research design is important and requires evaluation independently of research results. This chapter is therefore concerned with methodological choice and the impact of this on the processes and outcome of the research. This chapter gives an outline of research methods that were followed in the study. First, it should provide the information about population, sample size, sampling method, data collection method and data analysis technique but the researcher used a secondary data in this study. So, it can only provide the information about the sources of data and explanation about the method of each variables which suitable for this study and related to all of the objectives. For this study the researcher had used multiple linear regression and logistic regression to analyses the data and choose the best method that give more accurate. Lastly, the ethical issues that were followed in the process are also discussed.

## 3.2 SOURCES OF DATA

The sources of data in research study can be divided into two which is primary and secondary data. For this study, the data used is secondary data which is obtain from Department of Health in Kelantan. Generally secondary data can be defined as data that has been collected by another researcher or organization. For instance this study used dengue data in Kelantan from Department of Health. Thus both qualitative and quantitative data will be obtained from the department. The data obtain is from 2012 until 2016. However this study just used data from 2014 as it is the highest cases among the others. Then the data will be process to get the clean data which then will be proceed for method analysis.

### 3.2.1 Type of Data and Scale of Measurement

Table 3.1 Description of Variables Used in Study

| Variables | Description | Label | Measurement Scale |
|---|---|---|---|
| Age | Age of the dengue victims | Youth=1<br>Adult=2<br>Children = 3<br>Seniors=4 | Nominal |
| Gender | The gender of dengue victims | Male=1<br>Female=2 | Nominal |
| Locality | Locality of dengue occur | Student=1<br>Field Worker=2<br>Unemployed=3<br>Officer=4<br>Housewife=5<br>Undergraduate=6<br>Private=7<br>Retirees=8<br>Ministry of Education=9<br>Non-Government=10<br>Infant=11 | Nominal |
| Occupation | The occupation of the victims | Kota Bharu=1<br>Pasir Puteh=2<br>Tanah Merah=3<br>Kuala Krai=4<br>Pasir Mas=5<br>Machang=6<br>Bachok=7<br>Tumpat=8<br>Jeli=9<br>Gua Musang=10 | Nominal |
| Hotspot (YES , NO) | Hotspot of dengue | Yes=1<br>No=0 | Binary |

## 3.3 THEORETICAL FRAMEWORK

The relationship between dependent and independent variables are shown in the theoretical framework below. This study have four independent variable and a dependent variable.

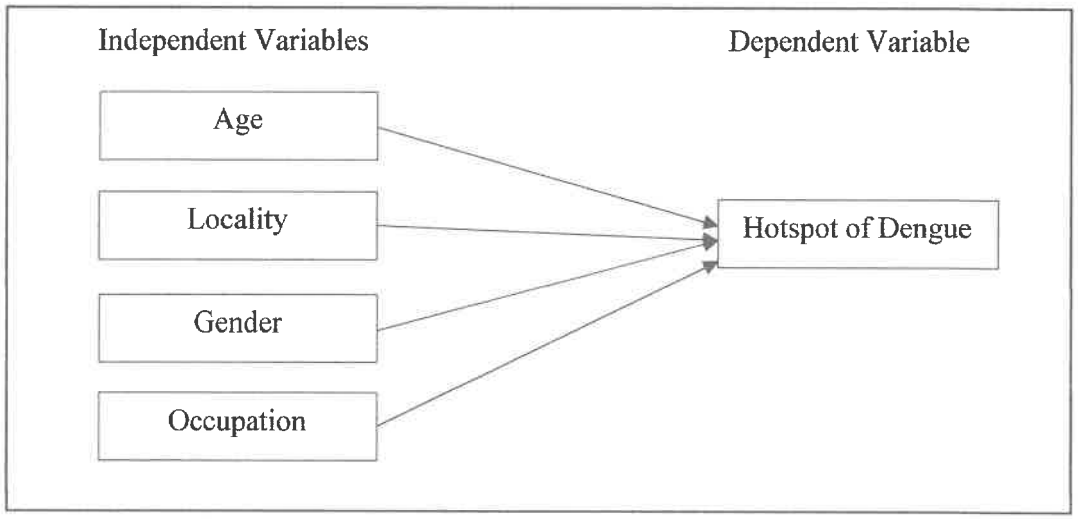


Figure 3.1 Theoretical framework of the Hotspot of Dengue

## 3.4 METHOD OF ANALYSIS

### 3.4.1 Method of Data Analysis

For this study, the data will be analysed using IBM SPSS 21.0 (Statistical Package for the Social Sciences). In order to interpreting the data, identifying the correlation among the variables, checking for logical inconsistencies, checking for ROC curve in dataset, this software would help a lot.

### 3.4.2 Logistic Regression

One of the method use in this study is the logistic regression. Logistic regression model is often referred to a special case of a generalized linear model and analyses models where the outcome is a nominal variable. Analysis for the logistic regression model assumes the outcome variable is a categorical variable which is in this study is hotspots of the dengue. It is common practice to assume that the outcome variable (Y), is a dichotomous variable having either a success or failure as the outcome. This study use yes or no as the outcome variable. The logistic regression model use in this study is to determine the best model of hotspot of dengue.

In order to generalized the linear model and also analyses the models which the outcome is a nominal variable, the logistic regression model or logit model is used. The analysis for the logistic regression model is assumed that the outcome variable is a categorical variable. Commonly the practice to assume that the outcome variable is denoted as Y is a dichotomous variable. It means that it just have either success or failure, good or bad, high or low and others outcomes.

The odds ratio (OR) which is meant to indicate whether the odds of a success (case) are equally likely to the odds of failure is given by $oddsratio = \frac{odds\ of\ cases}{odds\ of\ non\ cases}$

An odds ratio of one is an indication that the odds of a success (case) outcome are equally likely for to the odds of a failure (non-case).

For logistic regression analysis, the model parameter estimates (β1, β2, β3, β4) should be obtained and it should be determined how well the model fits the data. The potential explanatory variables were examined to determine whether or not they are significant enough to be used in this study models. The complete model contained all the explanatory variables believed to influence hotspots of dengue. Factor analysis was used to determine the significant combination of 4 factors in our model which are age, locality, gender and occupation.

### 3.4.2(a) Odds and Odds Ratios

To appreciate the logit model, it's helpful to have an understanding of odds and odds ratios. Most people regard probability as the "natural" way to quantify the chances that an event will occur. We automatically think in terms of numbers ranging from 0 to 1, with a 0 meaning that the event will certainly not occur, and a 1 meaning that the event certainly will occur.2 Probability, can be computed as follows: p=(number of outcomes)/(number of all possible outcomes)

### 3.4.2(b) The Logit Model

Now we're ready to introduce the logit model, otherwise known as the **logistic regression** model. For $k$ explanatory variables and $i = 1, \dots, n$ individuals, the model is

$$\log(\frac{p}{1-p}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n x_n \qquad (1)$$

Where $p_i$ is the probability that $P(Y_i) = 1$. The expression on the left-hand side is usually referred to as the logit or log-odds.Logit, is the log of the odds, is not only linear in X, but also linear in the parameters.

- The positive logit values indicate that the odds are in favour of an event happening, while
- Negative logit values indicate that the odds are against the occurrence of an event. We can solve the logit equation for $p_i$ to obtain

$$p = \frac{e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}}{1+ e^{(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \qquad (2)$$

We can simplify further by dividing both numerator and denominator by the numerator itself:

$$p = \frac{1}{1+ e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n)}} \qquad (3)$$

In mathematical expression, this formula is called the logistic function and can be written as:

$$p(x) = \frac{1}{1+e^{-z}}, z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n x_n \qquad (4)$$

### 3.4.2(c) Assessing the Goodness of Fit of the Model

i. **Omnibus Tests of Model Coefficients** gives us an overall indication of how well the model performs compared with model with none of the predictors entered into the model. For this results, we want a highly significant value (p-value must be less than 0.05)

ii. **Hosmer-Lemeshow test** will be used to test the goodness-of-fit of the model

iii. **Cox & Snell R-square and Nagelkerke R-square values** provide an indication of the amount of variation in the dependent variable explained by the model.

iv. **The Classification table** was also used in this study to know how well the model is able to predict the correct category. This table also provides the sensitivity and specificity of the model. **Sensitivity** measures the proportion of actual positives which are correctly identified, whereas **Specificity** measures the proportion of negative which are correctly identified. A model with high percentage of sensitivity and low in specificity are good and can be used for prediction.

v. **ROC Curve** was used in this study in order to provide the means comparison between classification models. It is a graph of True Positive Rate (Sensitivity) versus False Positive Rate (1 – Specificity). On X – axis show the proportion of positive targets that labelled correctly while on Y – axis is the proportion of negative targets that are mislabel as positive.

### 3.4.3 Chi Square Test of Independence

The Chi Squared Test of Independence is used to decide whether two categorical variables are independence. The data will be displayed in the contingency table which each row and column represent the categorical variables need to be tested.

In order to know how strength the association between the variables of categorical, Phi and Cramer's V value is used. Even the variables are significant, it does not mean that they are strongly associated. The value of Phi and Cramer's V lied between 0 and 1. If value closed to 1, it means that there is a strong association between the variables.

Assumptions:

- Random samples
- All expected frequencies are greater than equal to 1
- At most 20% of the expected frequencies are less than 5

1. Testing Hypothesis :

   $H_0$ : The two variables are independence

   $H_1$ : The two variables are not independence

2. Level of Significance, α
3. Test Statistic

   $\chi^2 = \sum \frac{(0_i - e_i)^2}{e_i}$ where

   $0_i$ = observed value

   $e_i = \frac{(row\ total)(column\ total)}{(grand\ total)}$

4 Decision Rule

   Reject $H_0$ if $\chi^2 > \chi^2\ \alpha$ , (r-1)(c-1)

5 Conclusion

# CHAPTER FOUR
# DATA ANALYSIS AND RESULT

## 4.1 INTRODUCTION

In this chapter presents the variables used in the binary logistic as well as discussing about model estimation and validation. The data was analysed using IBM SPSS which is called Statistical Product and Service Solution.

## 4.2 DESCRIPTIVE ANALYSIS

### 4.2.1 Age Category of the Victims

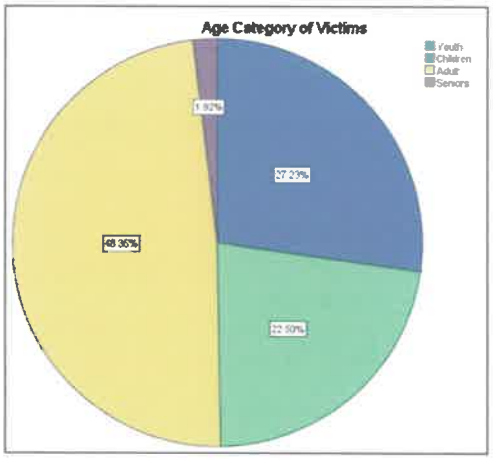


Figure 4.1 Pie Chart of Age_Category

Based on the pie chart Figure 4.1, it can be seen that the major age_category for the victims are adults with 48.35%. The remaining are youth, children and seniors with the percentage of 27.23%, 22.5% and 1.92% respectively.

### 4.2.2 Gender

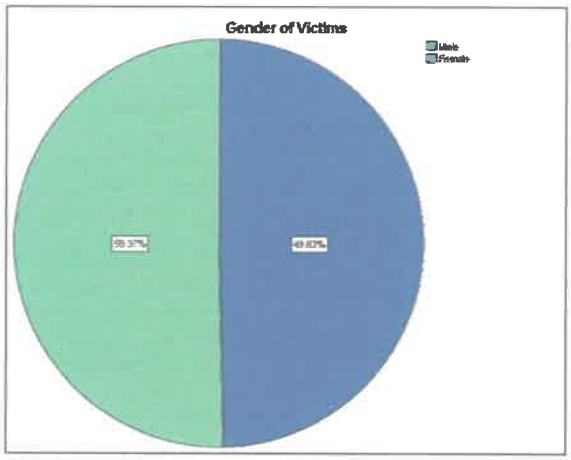


Figure 4.2 Pie Chart for Gender

From the pie chart in Figure 4.2, more victims was female with 50.37% and remaining 49.63% are male in the total population of 7899 observation

### 4.2.3 Locality

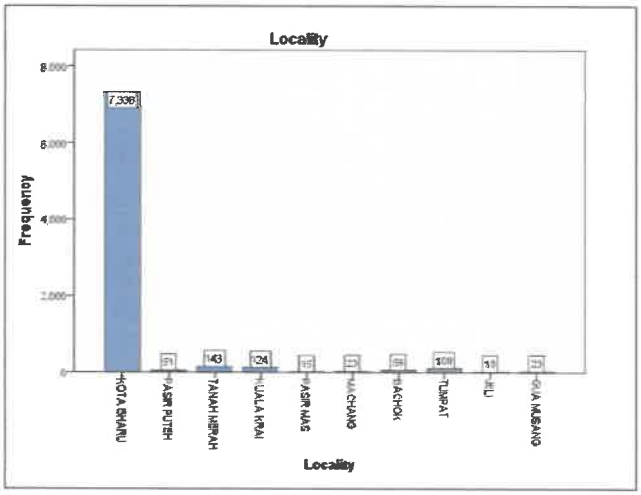


Figure 4.3 Bar Chart of the Locality

From bar chart in Figure 4.3 it is high cases in Kota Bharu with 7338 cases. Then followed by Tanah Merah, Kuala Krai and Tumpat with the cases of 143, 124 and 108 respectively. It is minor cases in district of Bachok, Pasir Puteh, Jeli and Pasir Mas. The cases are 56, 51, 18 and 15 respectively. Machang and Gua Musang was sharing the amount of cases that is 23 cases.

### 4.2.4 Occupation

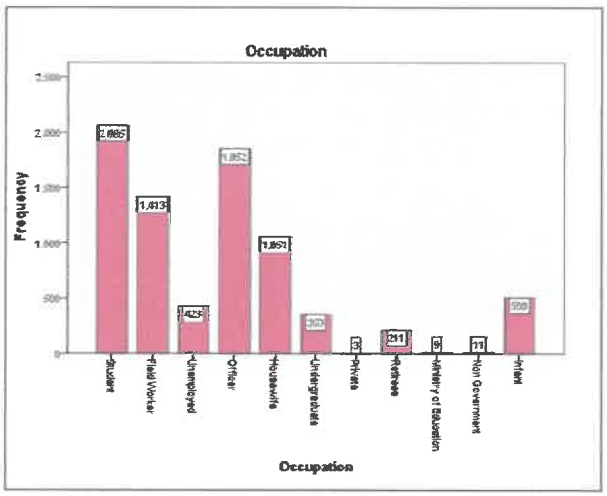


Figure 4.4 Bar Chart of the Occupation

Based on the bar chart in Figure 4.4 majority are students who prone to dengue with 2065 cases followed by officer, field worker, housewife, infant, unemployed, undergraduate and retirees with the frequency of cases 2065, 1852, 1413, 1051, 508, 423, 353 and 211 respectively. Least amount of cases for non-government, ministry of education and private with cases of 11, 9 and 3.

## 4.3 LOGISTIC REGRESSION

### 4.3.1 Omnibus Tests of Model Coefficients

Since this study used enter method of model fitting which mean entering all variables at the same step. Based on result in Table shows the model chi – square and the significance levels for test of the null hypothesis are equal to zero.

Table 4.1 Omnibus Tests of Model Coefficients

| | Chi-square | df | Sig. |
|---|---|---|---|
| Model | 1335.953 | 23 | 0.000 |

The model chi – square value in Table 4.1 is 1335.953 and the null hypothesis is rejected since the p – value (sig. value in Table 4.1) is less than 0.05 (significance level). It meant that the model is significant and one of the predictor variable is significant.

### 4.3.2 Model Summary

In order to determine how good the model fits the data, it is shown in model summary based on Table 4.2

Table 4.2 Model Summary

| -2 Log likelihood | Cox & Snell R Square | Nagelkerke R Square |
|---|---|---|
| 7933.094 | 0.156 | 0.225 |

Based on Table 4.2, Cox & Snell R Square described as pseudo R square statistics with values ranges between 0 and 1. The value of 0.156 suggests only 15.6% of the variability in the outcome is explained by the exposure. It is not a good measure of fit since the value is small and implying a poor fit. Nagelkerke R Square is another pseudo R square statistics with value ranges between 0 and 1. The value of Nagelkerke R Square is 0.225 suggests that only 22.5% of the variability in the outcome is explained by the exposure. It indicates that not a good measure of fit too.

### 4.3.3 Hosmer and Lemeshow Test

The Hosmer – Lemeshow test shown in the Table 4.3. Discovers whether the predicted probabilities are the same as the observed probabilities.

Table 4.3 Hosmer and Lemeshow Test

| Chi-square | df | Sig. |
|---|---|---|
| 13.876 | 8 | 0.085 |

Since the chi – square with 8 degree of freedom is 13.866 and p - value of the Hosmer and Lemeshow is 0.085 and greater than 0.05, the Hosmer – Lemeshow goodness of fit is significant which indicating the logistic regression model fit the data well.

### 4.3.4 Interpretation of the Model

Table 4.4 Variables in the Equation

| | B | S.E. | df | Sig. | Exp(B) |
|---|---|---|---|---|---|
| Age_Category | | | 3 | 0.009 | |
| Age_Category(1) | 0.158 | 0.110 | 1 | 0.153 | 1.171 |
| Age_Category(2) | 0.198 | 0.087 | 1 | 0.022 | 1.220 |
| Age_Category(3) | -0.258 | 0.215 | 1 | 0.230 | 0.773 |
| Gender(1) | 0.099 | 0.062 | 1 | 0.111 | 1.104 |
| Occupation | | | 10 | 0.000 | |
| Occupation(1) | -0.625 | 0.122 | 1 | 0.000 | 0.535 |
| Occupation(2) | -0.057 | 0.149 | 1 | 0.703 | 0.945 |
| Occupation(3) | -0.400 | 0.120 | 1 | 0.001 | 0.670 |
| Occupation(4) | -0.233 | 0.141 | 1 | 0.099 | 0.792 |
| Occupation(5) | 0.253 | 0.168 | 1 | 0.132 | 1.288 |
| Occupation(6) | -0.623 | 1.228 | 1 | 0.612 | 0.536 |
| Occupation(7) | -0.126 | 0.215 | 1 | 0.557 | 0.881 |
| Occupation(8) | -0.384 | 0.774 | 1 | 0.620 | 0.681 |
| Occupation(9) | -0.141 | 0.735 | 1 | 0.848 | 0.868 |
| Occupation(10) | -0.095 | 0.133 | 1 | 0.472 | 0.909 |
| Locality | | | 9 | 0.000 | |
| Locality(1) | -22.145 | 5614.349 | 1 | 0.997 | 0.000 |
| Locality(2) | -2.729 | 0.220 | 1 | 0.000 | 0.065 |
| Locality(3) | -22.422 | 3588.623 | 1 | 0.995 | 0.000 |
| Locality(4) | -22.154 | 10348.539 | 1 | 0.998 | 0.000 |
| Locality(5) | -22.394 | 8336.659 | 1 | 0.998 | 0.000 |
| Locality(6) | -22.243 | 5350.906 | 1 | 0.997 | 0.000 |
| Locality(7) | -3.143 | .289 | 1 | 0.000 | 0.043 |
| Locality(8) | -22.269 | 9433.922 | 1 | 0.998 | 0.000 |
| Locality(9) | -22.371 | 8319.837 | 1 | 0.998 | 0.000 |
| Constant | 1.316 | 0.091 | 1 | 0.000 | 3.730 |

The fitted model using the enter method is shown in Table 4.4. The final estimated logistic regression is:

$$\ln\left(\frac{p}{1-p}\right) = 1.316 + 0.158*\text{Age\_Category(1)} + 0.198*\text{Age\_Category(2)} - 0.258*\text{Age\_Category(3)} + 0.099*\text{Gender(1)} - 0.625*\text{Occupation(1)} - 0.057*\text{Occupation(2)} - 0.400*\text{Occupation(3)} - 0.233*\text{Occupation(4)} + 0.253*\text{Occupation(5)} - 0.623*\text{Occupation(6)} - 0.126*\text{Occupation(7)} - 0.384*\text{Occupation(8)} - 0.141*\text{Occupation(9)} - 0.095*\text{Occupation(10)} - 22.145*\text{Locality(1)} - 2.729*\text{Locality(2)} - 22.422*\text{Locality(3)} - 22.154*\text{Locality(4)} - 22.394*\text{Locality(5)} - 22.243*\text{Locality(6)} - 3.143*\text{Locality(7)} - 22.269*\text{Locality(8)} - 22.371*\text{Locality(9)}$$

i. Age_Category(2): OR=1.220

The person who are in category as an adult is approximately 1.22 times more likely that contribute towards hotspot of dengue compared to youth.

ii. Occupation(1): OR=0.535

The students are 1.8692 more likely to contribute towards hotspot of dengue compared to field worker.

iii. Occupation(3): OR=0.670

The students are 1.493 more likely to contribute towards hotspot of dengue compared to those who works as officer.

iv. Locality(2): OR=0.065

The locality of Kota Bharu is 15.3846 more likely to contribute towards hotspot of dengue compared to Tanah Merah.

v. Locality(7): OR=0.043

The locality of Kota Bharu is approximately 23.26 more likely to contribute towards hotspot of dengue compared to Tumpat.

So the final equation is,

$$\ln\left(\frac{p}{1-p}\right) = 1.316 + 0.198*\text{Age\_Category(2} - 0.625*\text{Occupation(1)} - 0.400*\text{Occupation(3)} - 2.729*\text{Locality(2)} - 3.143*\text{Locality(7)}$$

### 4.3.5 Classification Table

The classification table specifies how well the model predicts cases to the dependent categorical for example in this study how true the prediction of the hotspot of dengue whether the prediction is true based on actual.

Table 4.5 Classification Table

| Observed | | Predicted | | |
|---|---|---|---|---|
| | | Hotspot | | Percentage Correct |
| | | No | Yes | |
| Hotspot | No | 521 | 1640 | 24.1 |
| | Yes | 40 | 5698 | 99.3 |
| Overall Percentage | | | | 78.7 |

Based on the classification table in Table 4.5, it shows the cross tabulation of the predicted response category with the actual response category and predicted. Sensitivity is 0.99 where measures the proportion of actual positives which are correctly identified and specificity is 0.241, measures the proportion of negative which are correctly identified. It shown that a model with high percentage of sensitivity and low in specificity are good and can be used for prediction. The percentage of correct classification is 78.7% the variable. Hence, the value for error rate of the model is 21.3%.

### 4.3.6 ROC Curve

Table 4.6 Area under the Curve

| Area | Std. Error | Asymptotic Sig. | Asymptotic 95% Confidence Interval | |
|---|---|---|---|---|
| | | | Lower Bound | Upper Bound |
| 0.668 | 0.007 | 0.000 | 0.654 | 0.683 |

Area under the Receiver Operating Characteristic (ROC) curve was 0.668 represents the probability that the assay result for randomly chosen positive case will exceed the result for a randomly chosen negative case. P – Value is 0.000 which is significant. It means that using the assay is better than guessing. (95% Ci= 0.654, 0.683 with standard deviation 0.007 and p – value less than 0.000.

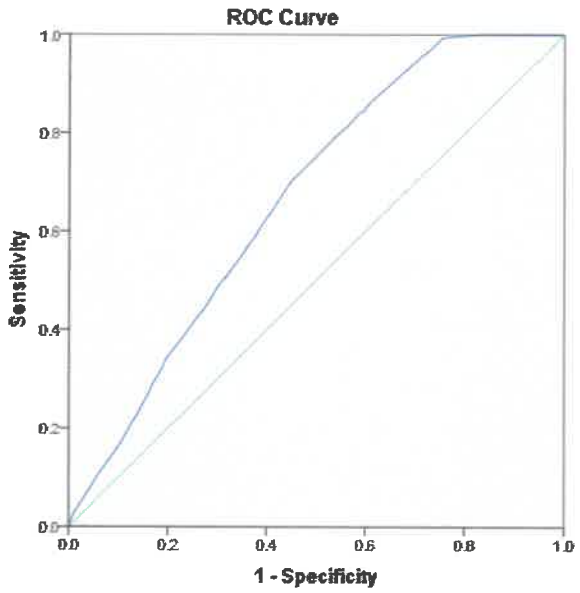


Figure 4.5 ROC Curve

According to ROC curve in Figure 4.5, it shows the best point gives 99.3% of the sensitivity and 24.11% of the specificity. This means that 99.3% of the model is able to identify that there is a hotspot of dengue.

## 4.4 CHI SQUARE TEST ON GENDER AND HOTSPOT OF DENGUE

### 4.4.1 The Cross Tabulation between Gender and Hotspot of Dengue

Table 4.7 Cross tabulation between Gender and Hotspot of Dengue

| | | | Hotspot | | Total |
|---|---|---|---|---|---|
| | | | No | Yes | |
| Gender | Male | Count | 1157 | 2763 | 3920 |
| | | Expected Count | 1072.4 | 2847.6 | 3920.0 |
| | | % within Gender | 29.5% | 70.5% | 100.0% |
| | | % within Hotspot | 53.5% | 48.2% | 49.6% |
| | | % of Total | 14.6% | 35.0% | 49.6% |
| | Female | Count | 1004 | 2975 | 3979 |
| | | Expected Count | 1088.6 | 2890.4 | 3979.0 |
| | | % within Gender | 25.2% | 74.8% | 100.0% |
| | | % within Hotspot | 46.5% | 51.8% | 50.4% |
| | | % of Total | 12.7% | 37.7% | 50.4% |
| Total | | Count | 2161 | 5738 | 7899 |
| | | Expected Count | 2161.0 | 5738.0 | 7899.0 |
| | | % within Gender | 27.4% | 72.6% | 100.0% |
| | | % within Hotspot | 100.0% | 100.0% | 100.0% |
| | | % of Total | 27.4% | 72.6% | 100.0% |

Based on Table 4.7, there is 29.5% male and 25.2% female are not in area of hotspots of dengue while the majority of male (70.5%) and female (74.8%) are in the area of hotspot of dengue. At the area that there is no hotspots of dengue, there are 53.5% male and 46.5% female while at the area of hotspot of dengue is occur, there are 48.2% of male and 74.8% of female.

### 4.4.2 Chi Square Test of Independence for Gender

Table 4.8 Chi Square Test of Gender

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 18.226 | 1 | 0.000 |

In Table 4.8, since the Pearson Chi-Square value is 18.226 (degree of freedom=1) with p-value 0.000 is less than $\alpha$=0.05, null hypotheses is rejected. Therefore, there is association between gender and hotspot of dengue.

### 4.4.3 The Strength of Relationship between Gender and Hotspot of Dengue

Table 4.9 Symmetric Measure of Gender

| | | Value | Approx. Sig. |
|---|---|---|---|
| Nominal by Nominal | Phi | 0.048 | 0.000 |
| | Cramer's V | 0.048 | 0.000 |
| N of Valid Cases | | 7899 | |

The value of Phi and Cramer's V in Table 4.9 is 0.048 which indicates there is weak association between gender and hotspot of dengue.

## 4.5 CHI SQUARE TEST ON AGE CATEGORY AND HOTSPOT OF DENGUE

### 4.5.1 The Cross Tabulation between Age Category and Hotspot of Dengue

Table 4.10 Cross tabulation between Age Category and Hotspot of Dengue

| | | | Hotspot | | Total |
|---|---|---|---|---|---|
| | | | No | Yes | |
| Age Category | Youth | Count | 597 | 1554 | 2151 |
| | | Expected Count | 588.5 | 1562.5 | 2151.0 |
| | | % within Age Category | 27.8% | 72.2% | 100.0% |
| | | % within Hotspot | 27.6% | 27.1% | 27.2% |
| | | % of Total | 7.6% | 19.7% | 27.2% |
| | Children | Count | 382 | 1395 | 1777 |
| | | Expected Count | 486.1 | 1290.9 | 1777.0 |
| | | % within Age Category | 21.5% | 78.5% | 100.0% |
| | | % within Hotspot | 17.7% | 24.3% | 22.5% |
| | | % of Total | 4.8% | 17.7% | 22.5% |
| | Adult | Count | 1131 | 2688 | 3819 |
| | | Expected Count | 1044.8 | 2774.2 | 3819.0 |
| | | % within Age Category | 29.6% | 70.4% | 100.0% |
| | | % within Hotspot | 52.3% | 46.8% | 48.3% |
| | | % of Total | 14.3% | 34.0% | 48.3% |
| | Seniors | Count | 51 | 101 | 152 |
| | | Expected Count | 41.6 | 110.4 | 152.0 |
| | | % within Age Category | 33.6% | 66.4% | 100.0% |
| | | % within Hotspot | 2.4% | 1.8% | 1.9% |
| | | % of Total | 0.6% | 1.3% | 1.9% |
| Total | | Count | 2161 | 5738 | 7899 |
| | | Expected Count | 2161.0 | 5738.0 | 7899.0 |
| | | % within Age Category | 27.4% | 72.6% | 100.0% |
| | | % within Hotspot | 100.0% | 100.0% | 100.0% |
| | | % of Total | 27.4% | 72.6% | 100.0% |

Based on Table 4.10, there are 27.8% of youth, 21.5% of children, 29.6% of adults and 33.6% of seniors who are not living in the area of hotspot of dengue occur while the highest percentage of age category that live in the area of hotspot of dengue is children with 78.5%, followed by youth 72.2%, adult 70.4% and the lowest percentage is seniors with 66.4%.

At the area with hotspots of dengue is not occur, the highest percentage of age category that are not live there is adult with 52.3% and the lowest percentage is seniors with only 2.4%. While the others, youth 27.6% and children 17.7%. Besides, at the area that hotspot are exist, the highest percentage is adult with 46.8%, the second highest is youth (27.1%), followed by children (24.3%) and the lowest is seniors (1.8%).

### 4.5.2 Chi Square Test of Independence for Age_Category

Table 4.11 Chi Square Test of Age_Category

| | Value | df | Asymp. Sig. (2-sided) |
|---|---|---|---|
| Pearson Chi-Square | 43.612 | 3 | 0.000 |

The value of Pearson Chi-square in Table 4.11, with degree of freedom 3 is 43.612 with the p-value is 0.000 is less than $\alpha$=0.05. Thus there is association between age category and hotspot of dengue.

### 4.5.3 The Strength of Relationship between Age_Category and Hotspot of Dengue

Table 4.12 Symmetric Measure of Age_Category

| | | Value | Approx. Sig. |
|---|---|---|---|
| Nominal by Nominal | Phi | 0.074 | 0.000 |
| | Cramer's V | 0.074 | 0.000 |
| N of Valid Cases | | 7899 | |

The value of Phi and Cramer's V in Table 4.12, is 0.074 with p=value is 0.000 show that there is weak association between age category and hotspot of dengue.

# CHAPTER FIVE

# CONCLUSION AND RECOMMENDATION

The factors that are interested in this study are dengue cases, mortality, gender, age and occupation. The results showed that model is significant when at least one of the predictor variables is significant. Using Hosmer and Lemeshow to check goodness of fit is significant which indicating the logistic regression model fit the data well. The percentage of correct classification is 78.7% the variable while the value for error rate of the model is 21.3%. So that, the data is suitable for this analysis. Based on result of this study, the highest cases mentioned is in Kota Bharu with 7338 cases. Majority students who prone to dengue followed by officer, field worker, housewife, infant, unemployed, undergraduate and retirees. Female with 50.37% and remaining 49.63% are male in the total population of 7899 observation who are the victims of this disease at Kelantan in 2014 while for age_category adults with 48.35% has been categorized as the most victims' category.

Not only that, we also have conducted the study using chi square method to identify the relationship between independence variables. Both gender and age category have a relationship to hotspot of dengue as the p – value is less than alpha=0.05. However, there are weak association between gender and age category towards hotspot of dengue as the Cramer's V value too far from 1.

The next researchers also can use another methods to compare the result analysis using this secondary data with same factors that associated for example multiple regression or data mining. The best analysis is when all the methods are given the consistent result to this study.

Ho, C. C., & Ting, C. Y. (2015). Time Series Analysis and Forecasting of Dengue Using Open Data. *Advances in Visual Informatics: 4th International Visual Informatics Conference. 9429*, pp. 51-63. Bangi: LNCS.

*iDengue*. (2018). Retrieved from iDengue Web site: http://idengue.remotesensing.gov.my/idengue/

Jeefoo, P., Tripathi, N. K., & Souris, M. (2010). Spatio-Temporal Diffusion Pattern and Hotspot Detection of Dengue in Chachoengsao Province, Thailand. *Int J Environ Res Public Health, 8*(1), 51 -74.

Liew, Su May; Khoo, Ee Ming; Ho, Bee Kiau; Lee, Yew Kong; Omar, Mimi; Ayadurai, Vickneswari;. (2016). Factors Associated with Dengue Mortality from a National Registry. *Dengue in Malaysia*, 1-14.

Mallhi, T. H., Khan, A. H., Sarriff, A., Adnan, A. S., & Khan, Y. H. (2017). Determinants of mortality and prolonged hospital stay among dengue patients attending tertiary care hospital: a cross-sectional retrospective analysis. *US National Library of Medicine National Institutes of Health, 7*(7).

Mohamed, M. S., Rahman, A. J., & Shaari, A. H. (2015). Dengue in Malaysia - A Commentary. *The International Medical Journal Malaysia, 14*(1), 1-2.

Pinto, R. C., de Castro, D. B., de Albuquerque, B. C., Sampaio, V. d., & al, e. (2016). Mortality Predictors in Patients with Severe Dengue in the State of Amazonas, Brazil. *PLoS One, 11*(8), 1-10.

Pooransingh, S., Teelucksingh, S., & Dialsingh, I. (2016). Dengue Deaths: Associated Factors and Length of Hospital Stay. *Advances in Preventive Medicine, 2016*, 1-4.

Rafique, I., Saqib, M. A., Munir, M. A., & Qureshi, H. (2017). Asymptomatic dengue infection in adults of major cities of Pakistan. *Asian Pacific Journal of Tropical Medicine, 10*(10), 1002-1006.

Ramzan, M., Ansar, A., & Nadeem, S. (2015). Dengue Epidemics: Knowledge Perhaps is the Only Key to Success. *Journal Ayub Med Coll Abbottabad, 27*(2), 402-406.

Salleh, I. H., & Mohd Nasir, M. R. (2014). Kelantan Malay Cultural Landscape: The Concept of Kota Bharu Islamic City. *SHS Web of Conferences. 12*, pp. 2097-2102. Tanjung Malim: EDP Sciences. Retrieved from https://www.shs-conferences.org/articles/shsconf/abs/2014/09/shsconf_4ictr2014_01021/shsconf_4ictr2014_01021.html

Sattar, A., Lukose, D., Nellis, S., Pham, D. N., & Sadanand, A. A. (2016). A Literature Review of Methods for Dengue Outbreak Prediction. *The Eighth International Conference on Information, Process, and Knowledge Management* (pp. 7-13). eKNOW.

Siregar, D., Djadja, M., & Arminsih, R. (2018). Analysis of the Risk Factors of Dengue Hemorrhagic Fever (DHF) In Rural Populations in Panongan Subdistrict, Tangerang 2016. *ICGH Conference Proceedings. 2017*, pp. 119-128. Depok: Knowledge E.

StatCan. (2018, May 8). *Statistics Canada*. Retrieved from Statistics Canada Web site: https://www.statcan.gc.ca/eng/concepts/definitions/age2

Suwanbamrung, C. (2012). Children's basic knowledge and activities for dengue problem solution: an islamic religious school, Southern Thailand. *Asian Pacific Journal of Tropical Disease, 2*(6), 456-464. Retrieved from https://www.sciencedirect.com/science/article/pii/S2222180812601005

Teurlai, M., Menkes, C. M., Virgil, C., Degallier, N., Descloux, E., Grangeon, J. P., & al, e. (2015). Socio-economic and Climate Factors Associated with Dengue Fever Spatial Heterogeneity: A Worked Example in New Caledonia. *PLoS Neglected Tropical Diseases, 9*(12), 1-31.

Thestar.com.my. (2018, May 29). Dengue cases in Kelantan increase 68 percent - Nation. *The Star Online*. Retrieved from https://w7ww.thestar.com.my/news/nation/2013/01/26/dengue-cases-in-kelantan-increase-68-percent

Tong, S., & Nash, S. (2014). Hotspot Detection And SPatio-Temporal Dynamics of Dengue in Queensland, Australia. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL*(8), 197-204.

Zainudin, Z., & Shamsuddin, S. M. (2016). Predictive Analytics in Malaysian Dengue Data from 2010 until 2015 using BigML. *Int. J. Advance Soft Compu. Appl, 8*(3), 18-30.

Zellweger, R. M., Cano, J., Mangeas, M., Taglioni, F., & et al. (2017). Socioeconomic and environmental determinants of dengue transmission in an urban setting: An ecological study in Nouméa, New Caledonia. *PLoS Neglected Tropical Disease, 11*(4), 1-18.

# APPENDICES

## APPENDIX A

Letter applying for data to Department of Health in Kelantan

**Pejabat Penolong Rektor (Kampus Kota Bharu)**
*Office of Assistant Rector (Kota Bharu Campus)*

**Universiti Teknologi MARA Cawangan Kelantan**

UNIVERSITI TEKNOLOGI MARA

Surat Kami :
Tarikh :

**KEPADA SESIAPA YANG BERKENAAN**

YBhg.Datuk/Datin Paduka/Prof/Tuan/Puan

**PENGESAHAN SEBAGAI PELAJAR UNIVERSITI TEKNOLOGI MARA (UiTM) KELANTAN, KAMPUS KOTA BHARU**

Adalah dimaklumkan bahawa penama di bawah merupakan pelajar dari **Sarjana Muda Sains (Kepujian) Statistik (CS241), bahagian 05** yang ditugaskan untuk menyiapkan satu laporan projek yang berkaitan dengan matapelajaran mereka ikuti iaitu ***RESEARCH METHODOLOGY(STA 650).***

| BIL | NAMA | NO.UiTM |
|---|---|---|
| 1 | NOR FARISHA BINTI MUHAMAD KRISHNAN | 2015283646 |
| 2 | NUR SHAHIRAH BINTI DOLLAH | 2015238474 |
| 3 | NUR AZERA BINTI AHMAD | 2015430712 |

2. Sehubungan dengan itu, pihak kami berharap agar pihak tuan dapat memberi kerjasama dalam menyumbangkan sebanyak mungkin maklumat seperti yang diperlukan. Segala maklumat tersebut akan kami rahsiakan dan hanya untuk tujuan akademik sahaja.

3. Pihak kami mengucapkan terima kasih di atas kerjasama dan maklumat yang diberikan.

Sekian.

Yan

**IRO AFFIZAN MOHAMED**
Pen ksekutif
Bah al Ehwal Akademik
b.p. **OLONG REKTOR (KAMPUS KOTA BHARU)**

Kepada:

Pengarah Jabatan Kesihatan Kelantan

Wisma Persekutuan

Jalan Bayam

15590 Kota Bharu

KELANTAN DARUL NAIM

Berikut adalah data yang dimohon:

| Bil | Jenis Data | Lokasi | Durasi | Format Data |
|---|---|---|---|---|
| 1. | Jumlah kes denggi | Seluruh Kelantan (mengikut daerah) | 2005 – 2017 (mengikut minggu) | Excel |
| 2. | Jumlah kematian disebabkan denggi | Seluruh Kelantan (mengikut daerah) | 2005 – 2017 (mengikut minggu) | Excel |
| 3. | Kategori mangsa denggi (Umur dan Jantina) | Seluruh Kelantan (mengikut daerah) | 2005 – 2017 (mengikut minggu) | Excel |
| 4. | Data cuaca mengikut lokaliti | Seluruh Kelantan (mengikut daerah) | 2005 – 2017 (mengikut minggu) | Excel |
| 5. | Senarai lokaliti berlaku denggi | Seluruh Kelantan (mengikut daerah) | 2005 – 2017 (mengikut minggu) | Excel |
| 6. | Skop pekerjaan | Seluruh Kelantan (mengikut daerah) | 2005 – 2017 (mengikut minggu) | Excel |

Untuk makluman pihak Dato' kami pelajar tahun akhir Ijazah Sarjana Muda Kepujian Statistik Universiti Teknologi Mara Cawangan Kelantan Kampus Kota Bharu. Kami dikehendaki menyiapkan *Final Year Project* bagi melengkapkan pengajian kami. Tajuk kajian kami ialah **'Factors Associated to Hotspot of Dengue'** dan kami memerlukan beberapa data berkaitan denggi dari pihak Jabatan Kesihatan Negeri Kelantan

Sekian, terima kasih.

## APPENDIX B

## OUTPUT LOGISTIC REGRESSION USING SPSS

| Case Processing Summary | | | |
|---|---|---|---|
| Unweighted Cases[a] | | N | Percent |
| Selected Cases | Included in Analysis | 7899 | 100.0 |
| | Missing Cases | 0 | .0 |
| | Total | 7899 | 100.0 |
| Unselected Cases | | 0 | .0 |
| Total | | 7899 | 100.0 |
| a. If weight is in effect, see classification table for the total number of cases. | | | |

| Dependent Variable Encoding | |
|---|---|
| Original Value | Internal Value |
| No | 0 |
| Yes | 1 |

There are two coding used which are 0 and 1, according to dependent variable encoding. For hotspot of dengue, no is coded with 0 and yes is coded with 1. Since this study is interested to investigate on hotspot of dengue, coding 1 is used.

**Block 0: Beginning Block**

| Classification Table[a,b] | | | | | |
|---|---|---|---|---|---|
| | Observed | | Predicted | | |
| | | | Hotspot | | Percentage Correct |
| | | | No | Yes | |
| Step 0 | Hotspot | No | 0 | 2161 | .0 |
| | | Yes | 0 | 5738 | 100.0 |
| | Overall Percentage | | | | 72.6 |
| a. Constant is included in the model. | | | | | |
| b. The cut value is .500 | | | | | |

| Variables in the Equation | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | B | S.E. | Wald | df | Sig. | Exp(B) |
| Step 0 | Constant | .977 | .025 | 1497.004 | 1 | .000 | 2.655 |

| Variables not in the Equation | | | | | |
|---|---|---|---|---|---|
| | | | Score | df | Sig. |
| Step 0 | Variables | Age_Category | 43.612 | 3 | .000 |
| | | Age_Category(1) | 39.631 | 1 | .000 |
| | | Age_Category(2) | 18.955 | 1 | .000 |
| | | Age_Category(3) | 2.993 | 1 | .084 |
| | | Gender(1) | 18.226 | 1 | .000 |
| | | Occupation | 159.941 | 10 | .000 |
| | | Occupation(1) | 130.450 | 1 | .000 |
| | | Occupation(2) | 2.367 | 1 | .124 |
| | | Occupation(3) | 1.467 | 1 | .226 |
| | | Occupation(4) | 1.011 | 1 | .315 |
| | | Occupation(5) | 14.875 | 1 | .000 |
| | | Occupation(6) | .054 | 1 | .816 |
| | | Occupation(7) | 1.462 | 1 | .227 |
| | | Occupation(8) | .162 | 1 | .687 |
| | | Occupation(9) | .000 | 1 | .995 |
| | | Occupation(10) | 11.513 | 1 | .001 |
| | | Locality | 1322.718 | 9 | .000 |
| | | Locality(1) | 136.298 | 1 | .000 |
| | | Locality(2) | 217.350 | 1 | .000 |
| | | Locality(3) | 334.502 | 1 | .000 |
| | | Locality(4) | 39.905 | 1 | .000 |
| | | Locality(5) | 61.249 | 1 | .000 |
| | | Locality(6) | 149.756 | 1 | .000 |
| | | Locality(7) | 196.235 | 1 | .000 |
| | | Locality(8) | 47.904 | 1 | .000 |
| | | Locality(9) | 61.249 | 1 | .000 |
| | Overall Statistics | | 1392.564 | 23 | .000 |

## APPENDIX C

## Output of Chi Square Independent Test

### Between Gender and Hotspot of Dengue

| Case Processing Summary | | | | | | |
|---|---|---|---|---|---|---|
| | Cases | | | | | |
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Gender * Hotspot | 7899 | 100.0% | 0 | 0.0% | 7899 | 100.0% |

The percentage of valid cases is 100% with non-missing values for both gender and hotspot of dengue. Hence, gender and hotspot of dengue can be used in this test.

### Between Age_Category and Hotspot of Dengue

| Case Processing Summary | | | | | | |
|---|---|---|---|---|---|---|
| | Cases | | | | | |
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| Age Category * Hotspot | 7899 | 100.0% | 0 | 0.0% | 7899 | 100.0% |

The valid cases percentage of age category and hotspot dengue is 100.0% with no missing value thus, the test can be done between age category and hotspot of dengue.

# REFERENCES


Anker, M., & Arima, Y. (2011). Male–female differences in the number of reported incident dengue fever cases in six Asian countries. *Emerging Diseases Surveillance and Reponse, Division of Health Security amd Emergencies, 2*(2), 1-7.

Bernama. (2017, January 30). Concerns rise as dengue cases surge in six states this month. *New Straits Times*. Retrieved from https://www.nst.com.my/news/2017/01/208174/concerns-rise-dengue-cases-surge-six-states-month

Brinkhoff, T. (2017, December 30). *City Population*. Retrieved from citypopulation.de website: https://www.citypopulation.de/php/malaysia-admin.php?adm1id=03

Chen, B., Yang, J., Luo , L., Yang , Z., & Liu, Q. (2016). Who Is Vulnerable to Dengue Fever? *International Journal of Environmental Research and Public Health, 13*(712), 1-15.

Cheong, Y. L., & Leitão, J. P. (2014). Assessment of land use factors associated with dengue cases in Malaysia using Boosted Regression Trees. *Spatial and Spatio-temporal Epidemiology, 10*, 75-84.

Chew, C.H., Woon, Y.L. , Amin, F., Adnan, T.H., Wahab, A.H.A., Ahmad, Z.E., Bujang, M.A., Hamid, A.M.A., Jamal, R., Chen, W.S., Hor, C.P., Yeap, L., Hoo, L.P., & Goh, P.P. (2016). Rural-urban comparisons of dengue seroprevalence in Malaysia. *BMC Public Health, 16*(824), 1-9.

Dzukifly, D. (2014, July 16). *Malay Mail Online*. Retrieved from Malay Mail Online Web site: http://www.themalaymailonline.com/malaysia/article/identify-real-hotspots-to-fight-dengue-effectively-association-tells-govern#7jhHYoWeJaUOuw5v.97

Er, A.C., Rosli, M.H., Asmahani A., Mohamad Naim M.R., Harsuzilawati M. (2010). Spatial Mapping of Dengue Incidence: A Case Study in Hulu Langat District, Selangor,Malaysia. *International Journal of Geological and Environmental Engineering, 4*(7), 251-255.