

Predictive Analysis of Malaysian Dengue Haemorrhagic Fever data from 2010 - 2017

Foong M. Wong

November 4, 2019

Introduction

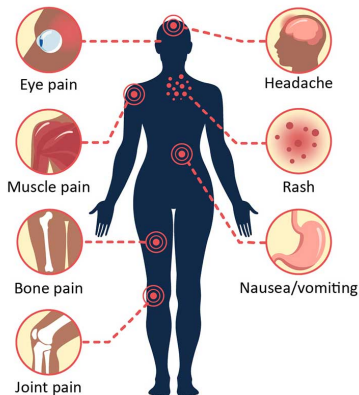
- Dengue Fever (DF)



Infected Aedes mosquitoes

Dengue Symptoms

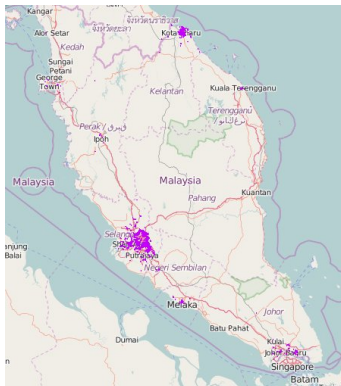
- "Break-bone fever"
- Severe joint and muscle pain, rash, fever



Symptoms of Dengue Fever

Dengue Fever in Malaysia

- Tropical weather
- Selangor - Highest number of dengue fever cases

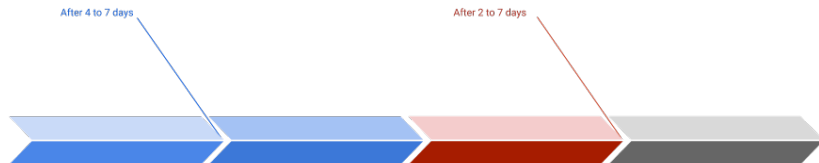


Hotspot for Dengue in Selangor based on Website IDengue for community (Government Portal, 2016)

Difference between DF and DHF

- Dengue Fever (DF) V.S. Dengue Hemorrhagic Fever (DHF)
 - ▶ There is a slight difference.
 - ▶ DF \rightarrow DHF

Evolution of DF to DHF



Aedes mosquitoes bite

Lorem ipsum dolor sit amet, consectetur adipiscing.

Dengue Fever Symptoms

- Sudden high fever
- Pain behind the eyes
- Muscle, joint, bone pain
- Severe headache
- Skin rash

Dengue Hemorrhagic Fever

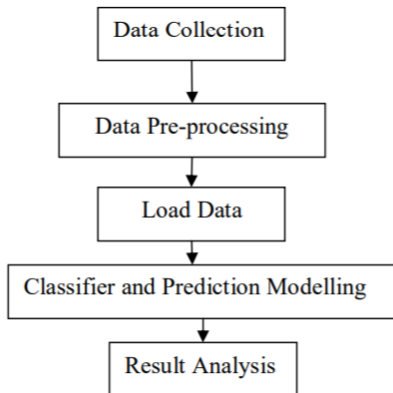
- Severe abdominal pain
- Vomiting, diarrhea
- Convulsions, bruising
- Uncontrolled bleeding

Dengue Shock Syndrome

- Circulatory system failure, shock
- Can be fatal

Dengue Fever to Dengue Hemorrhagic Fever

- Develop a model to predict the:
 - ① Infected areas/states
 - ② Age groups of dengue hemorrhagic fever
- Classical analysis to describe:
 - ① Epidemic pattern
 - ② Future outbreaks



Flow chart of research methodologies process

- Malaysian Open Data Portal

The screenshot displays the Malaysian Open Data Portal interface. At the top, a breadcrumb trail reads: Home / Data Set Provider / Ministry of Health / Number of Dengue Disease Cases ...

On the left sidebar, there is a search bar with the text 'Carian Terperinci' and a 'LOOK FOR' button. Below this is a 'Data Set Provider' section featuring the Malaysian coat of arms and the text 'health Ministry'. Further down are sections for 'Openness' (with a star rating of 3 stars) and 'Social' (with a Google+ icon).

The main content area has a search bar with the text 'Cari' and a 'LOOK FOR' button. Below the search bar are tabs for 'Data Set' and 'Clusters'. The selected 'Data Set' tab shows the title 'Number of Annual Dengue Haemorrhagic Fever (DHF) Cases by State and Age'. The description states: 'This dataset shows the number of cases of Dengue Haemorrhagic Fever (DHF) cases annually by state and age in Malaysia. Data will be provided annually once the data is finalized.' It also shows 'Total Views: 0' and 'Status: Dataset is Published'.

Below the description is a section titled 'Data Sets and Sources' which includes a small icon of a data set and the title 'Number of cases of Dengue Haemorrhagic Fever Disease ...'. The description for this dataset is: 'This dataset shows the number of cases of Dengue Haemorrhagic Fever (DHF) ...'. It also shows 'Downloads: 6' and an 'Explore' button.

Dengue Hemorrhagic Fever data from 2010 - 2017
(Source: <http://www.data.gov.my/>)

Data Pre-processing

- Use tool for data cleaning and formatting

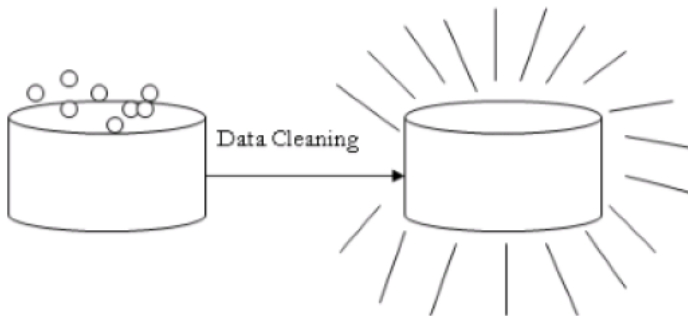


Illustration on Data Cleaning process

Data Pre-processing (cont.)

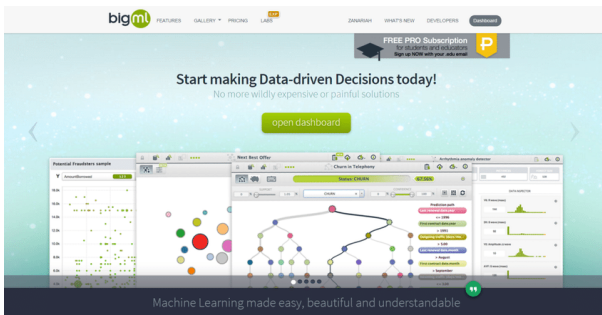
- Translate the words into English
- Get the right data format!

BILANGAN KES PENYAKIT DENGUE HAEMORRHAGIC FEVER TAHUNAN MENGIKUT NEGERI DAN KUMPULAN UMUR, TAHUN 2010																
KUMPULAN UMUR	JOHOR	KEDAH	KELANTAN	MELAKA	NEGERI SEMBILAN	PAHANG	PERAK	PERLIS	PULAU PINANG	SABAH	SARAWAK	SELANGOR	TERENGGANU	WP KUALA LUMPUR	WP LABUAN	MALAYSIA
0-4	10	0	4	1	0	0	0	0	3	1	5	32	2	5	0	63
5-9	16	0	11	2	2	4	4	0	1	3	12	63	4	21	0	143
10-14	63	0	25	29	9	12	8	0	5	2	10	182	9	37	0	391
15-19	67	0	55	38	15	16	4	0	1	2	12	243	14	38	0	505
20-24	102	0	32	52	18	10	2	0	0	0	22	348	13	44	0	643
25-29	115	0	31	41	18	13	2	0	3	2	21	330	14	48	0	638
30-34	78	0	20	28	14	13	4	0	2	2	12	246	5	33	0	457
35-39	63	0	23	28	13	10	1	0	4	1	17	166	6	20	0	352
40-44	35	1	14	27	9	7	2	0	1	3	20	132	5	15	0	271
45-49	29	1	22	18	13	12	0	0	1	1	8	83	3	6	0	197
50-54	18	0	21	15	9	11	1	0	0	0	11	61	3	11	0	161
55-59	12	1	6	16	6	4	3	0	0	0	4	34	4	10	0	100
60-64	12	0	5	5	2	2	0	0	0	0	3	16	0	5	0	50
65-69	7	0	8	3	0	1	0	0	0	0	3	10	0	2	0	34
70-74	3	0	0	4	0	0	1	0	0	0	0	7	0	2	0	17
>75	0	0	2	1	0	1	0	0	0	0	2	2	1	0	0	9
Grand Total	630	3	279	308	128	116	32	0	21	17	162	1955	83	297	0	4031

Dengue Hemorrhagic Fever Dataset in 2010

Classifier and Prediction Modeling

- Method: TBD, some data mining/ machine learning tools (e.g:scikit-learn, Weka, etc.)



Example of big data tool: BigML

- ✓ Identified research goals and collected data
- ✗ Pre-process data to get into the right format
- ✗ Decide on a method/tool to conduct dengue hemorrhagic fever predictive analysis
- ✗ Obtain research results! (infected DHF states and age groups)