**General view of Covid-19 cases and Covid-19 clusters in Hong Kong.**

1. A description of the problem and a discussion of the background. (**15 marks**)

Covid-19 is now spreading all around the world. We all know that the virus do not stop spreading. In fact, it becomes more infectious over time. All government tried their best to control the spreading of Covid-19 in their own country. As a citizen in Hong Kong, it is particular interesting for me to have a general view on the covid-19 cases in Hong Kong. Also, currently there is an outbreak in Hong Kong, which we called the third wave of covid-19 outbreak. This outbreak is the most intense than the previous two wave happened in Hong Kong. This report would like to analyze the Covid-19 cases in Hong Kong to provide a sense of idea about the current situation in Hong Kong, and also use some clustering method such as k-means and DBSCAN method to cluster the Covid cases in Hong Kong.

1. A description of the data and how it will be used to solve the problem. (**15 marks)**

The dataset are publicly available from government website. The first dataset is is the Details of propable/confirmed cases of COVID-19 in Hong Kong. (https://www.chp.gov.hk/files/misc/enhanced\_sur\_covid\_19\_eng.csv) This dataset provides the general information of every confirmed or probable cases up to date, and this report will use this dataset to look at the current Covid-19 situation in Hong Kong.

The second dataset is the Residential buildings in which probable/confirmed cases have resided in the past 14 days or non-residential building with 2 or more probable/confirmed cases in the past 14 days. (http://www.chp.gov.hk/files/misc/building\_list\_eng.csv) This dataset provides the complete building list or locations where the confirmed cases resided or gathered before. The dataset is updated daily and only shows the results over the past 14 days. This data would be useful for finding the Covid-19 cluster in Hong Kong.