# Assignment Summary

In this assignment, we have to find the mortality rates according to the given data set. We use different python libraries to calculate or find the given problem. So first we import three different libraries to read the data set we use pandas function read_csv to read the whole dataset. We use the head function to get the first few values of our data set and then the info function is used to check the column values and the datatypes that are available in our data set. As there is a lot of null values in our main data set that we use the is_null function to get the null values in our dataset after that we sum all the values to get the total number of null values in our dataset. According to the given conditions we separate the condition from our dataset and make a new_dataframe and apply condition according to that for this data we use the head function to get the values but there are still a lot of nulls in this data set again we use is_null function to get the null values and then get the sum of these columns to get the total number of null values in our new data frame. To remove the Null value, we use the panda's function known as interpolation Data frames and series can be filled with NA values using this function. The function is powerful, but it can be overused. Interpolation is used instead of hard-coding the value to fill in the missing values. Linear interpolation is used to fill in the missing data. By default, the Linear method ignores the index and considers the values to be evenly spaced. Now the second condition is to get the excess death rate according to that we use another library one is matplotlib and the other is seaborn these are used to display the data In the form of the graph according to the user. Then the third condition is to get the highest death rate according to a year for that we have to convert the data into year format for that we first convert data into string format using Str and then use the split function to split the data into year format after that again convert the string into integer form for this use NumPy fuctionto_numeric and pass our dataset column to get the year in Integer format. Now we use group by function Using the Pandas group, data are grouped according to categories, and then a function is applied to the categories. The data can also be aggregated more efficiently. Data can be divided into groups based on some criteria using the Pandas group by function. You can split Panda objects along any axis. As a mapping of labels to groups, grouping defines what it means. So we group by the dataframe first for location and high mortality column and use sum function to get the sum of all the values now we have to show the highest mortality for that data frames can be sorted by passing a column to Pandas' sort_values() function. Because it cannot sort a data frame and particular columns cannot be selected, it is different from the sorted Python function.