Chosen Admission and Fatalities CSV data sets Plan of Action.

From the studied observation of two Kaggle.com **manufactured** data sets, a presence of two preattentional values columns can be observed. One table should be formed and a relationship between the data sets provided two value columns can be attempted to originate. We are attempting to track All Hospital Admissions and All Hospital Mortalities on an annual basis.

From the admissions.csv column:

1. Year
2. ICD10 Code
3. Value

From the admissions.csv column:

1. Year
2. ICD10 Code
3. Value

From the admissions.csv:

1. data table column 1 named Year, cell A:2 thru A:1750
2. Modify the entire column cell element to display a year in a single YYYY-format
3. Combine the column data elements to display one value per each year 2004-15
4. Goal is to contain a combined number annual amount, total of 11 years.
5. data table column 2 named ICD10 code, cell B:2 thru B:1750

a. Two data tables are merged on ICD10 key

b. Combine the column data elements to display one value per each year 2004-15 and ICD10 code row data containing All Codes cell value only.

c. Goal is to combine All codes cell values on an annual basis

3) data table column 7 named Value, cell G:2 thru G:1750

a. A column should be renamed from Value to Total Admissions

b. Perform a mathematical cell value summation from the above combined Annual and All Codes data.

c. Goal is to combine All codes cell values on an annual basis and draw a Total Admission amount per year for 11 years.

From the fatalities.csv:

1) data table column 1 named Year, cell A:2 thru A:1750

1. Modify the entire column cell element to display a year in a single YYYY-format
2. Combine the column data elements to display one value per each year 2004-15
3. Goal is to contain a combined number annual amount, total of 11 years.

2) data table column 2 named ICD10 code, cell B:2 thru B:1750

a. Two data tables are merged on ICD10 key

b. Combine the column data elements to display one value per each year 2004-15 and ICD10 code row data containing All Codes cell value only.

c. Goal is to combine All codes cell values on an annual basis

3) data table column 7 named Value, cell G:2 thru G:1750

a. A column should be renamed from Value to Total Fatalities

b. Perform a mathematical cell value summation from the above combined Annual and All Codes data.

c. Goal is to combine All codes cell values on an annual basis and draw a Total Fatalities amount per year for 11 years.

Following the data sets merger, newly formed table should display:

* Year-cleaned combined annual
* ICD10-KEY-All Codes(data duplicates may not be necessary to display)
* Total Fatalities-annual sum
* Total Admissions-annual sum

Plot charts should display relationships:

* Year vs Total Admissions
* Year vs Total Fatalities
* Total Fatalities vs Total Admission

Steps to complete the project:

1. Data sets cleanup.
2. Post-cleanup new Data array formation and merge.
3. Form a relation and display visualization via plotting
4. Load the formed html converted data table onto Mongo DB and Flask
5. Display the data table and plot chart Html output

Preclusion:

**The provided data sets fail to rectify the observed mortalities are directly related to smoking tobacco products.** Furthermore, the diagnosis column data statements in the cancer description are speculative by nature at best. Example, “Cancers which can be caused by smoking” is not a medical diagnosis and cannot be excepted in our data science study.

However, based on the initial inspection, possible trends between Hospital admissions and Mortality rates may be discovered in the final data analysis.