**MALARIA DATASET**

***INTRODUCTION:***

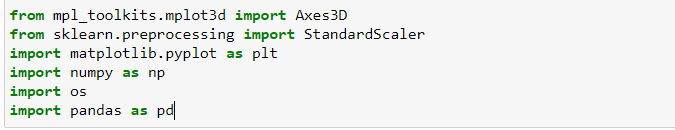
* Explore the given dataset and analyze the number of deaths ,CFR caused by malarial diseases across the word and visualize the result .we have three datasets such as
* reported\_numbers.csv - Reported no. of cases across the world
* estimated\_numbers.csv - Estimated no of cases across the world
* incidenceper1000popat\_risk.csv - Incidence per 1000 people at risk area.

***DATASET:***

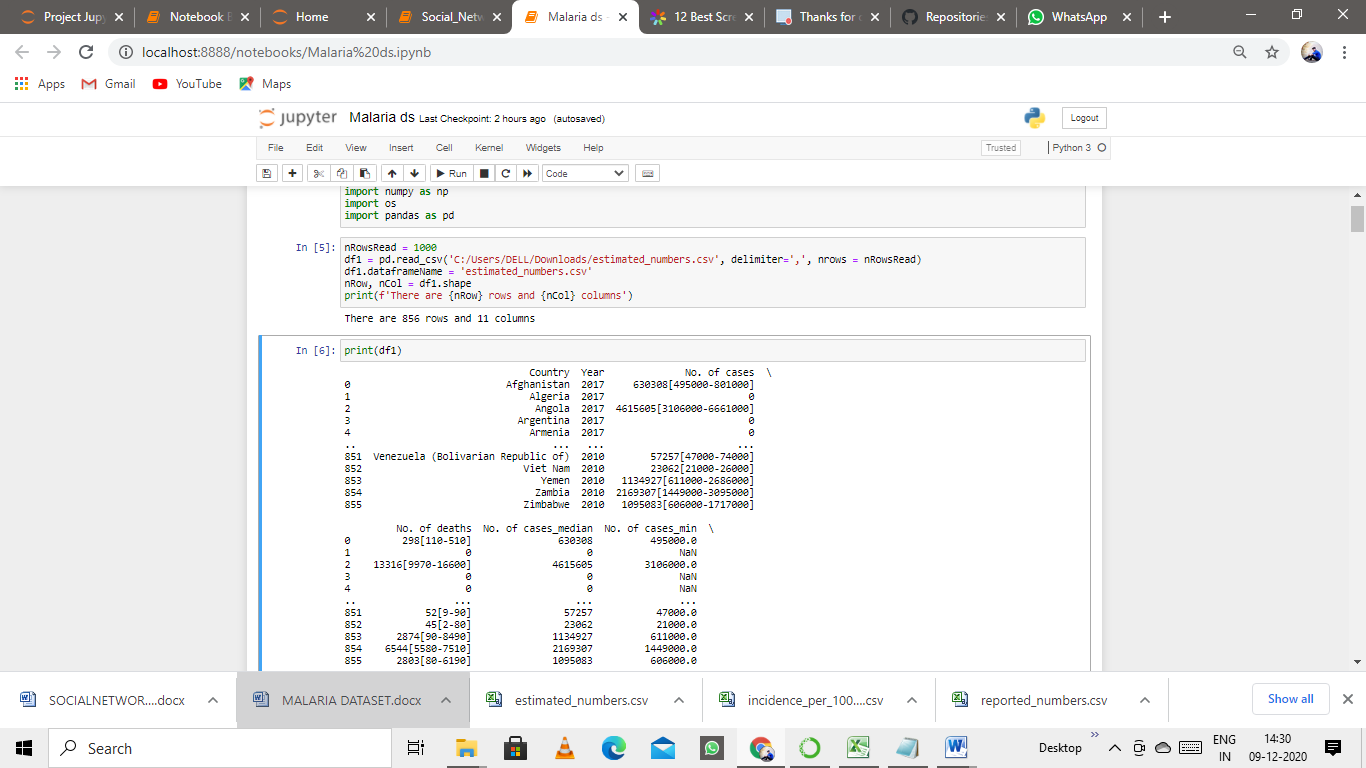
* Malaria is a life-threatening disease caused by parasites that are transmitted to people through the bites of infected female Anopheles mosquitoes.
* It is preventable and curable.
* In 2018, there were an estimated 228 million cases of malaria worldwide.
* The estimated number of malaria deaths stood at 405 000 in 2018.
* Children aged under 5 years are the most vulnerable group affected by malaria; in 2018, they accounted for 67% (272 000) of all malaria deaths worldwide.
* The WHO African Region carries a disproportionately high share of the global malaria burden.
* In 2018, the region was home to 93% of malaria cases and 94% of malaria deaths
* The **estimated\_numbers.csv** dataset gives the Estimated no of cases across the world contains the following country,region,year,no of deaths,no of cases ,WHO region etc.
* **reported\_numbers.csv** datasetgives the reported no of cases across the world has contains 5 columns and they are Country,year,no of cases,no of deaths and WHO region.
* **incidenceper1000popat\_risk.csv** dataset gives Incidence per 1000 people at risk area has 4 columns and they are Country, year,no of cases and WHO region .

**SOLUTION:**

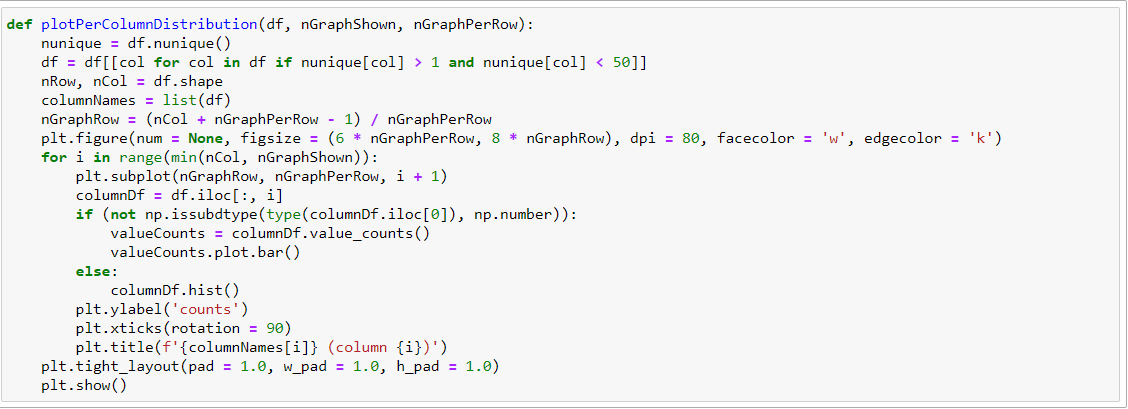
* **Import the packages**



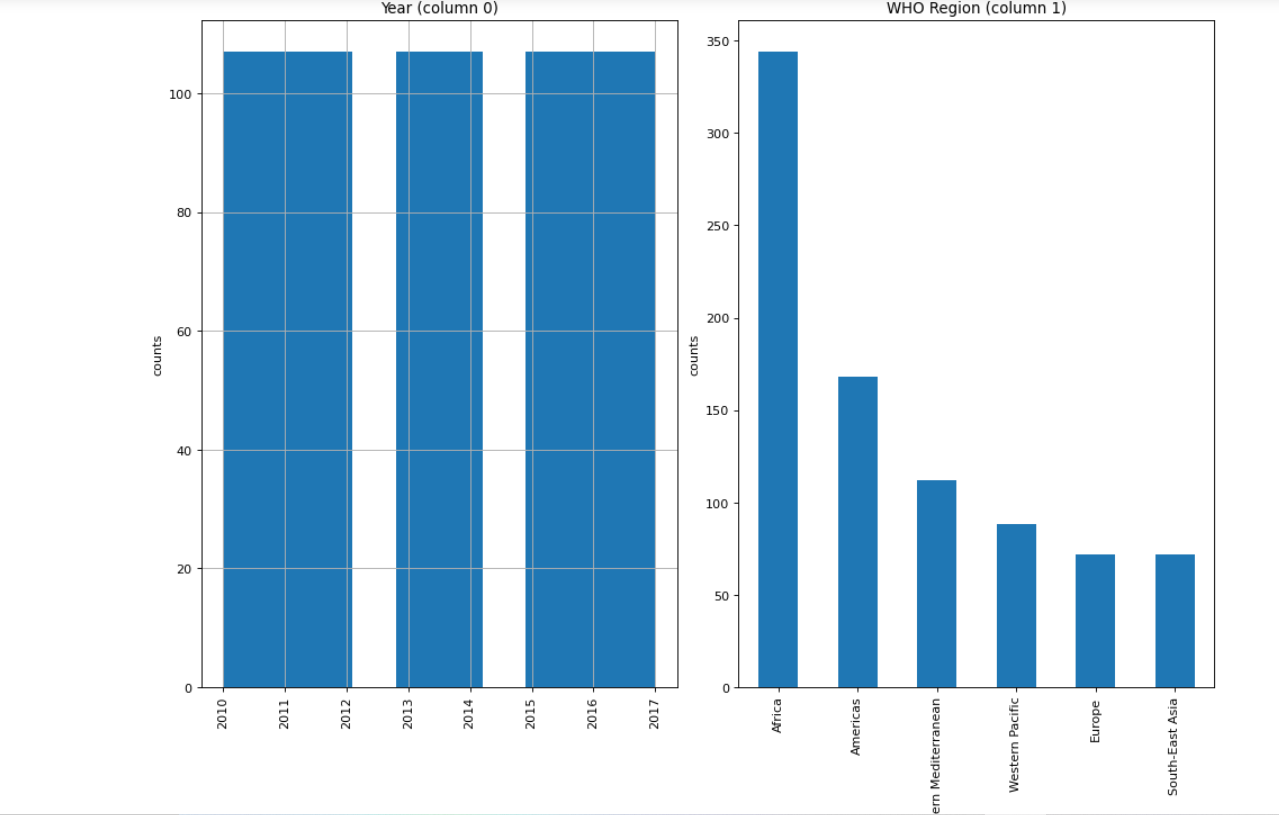
* **Load the dataset**



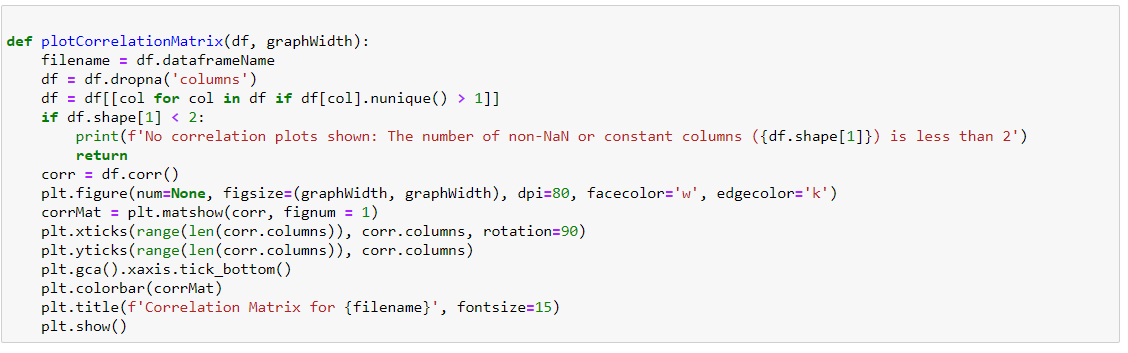
* **PLOTTING**
* **Column distribution plot**



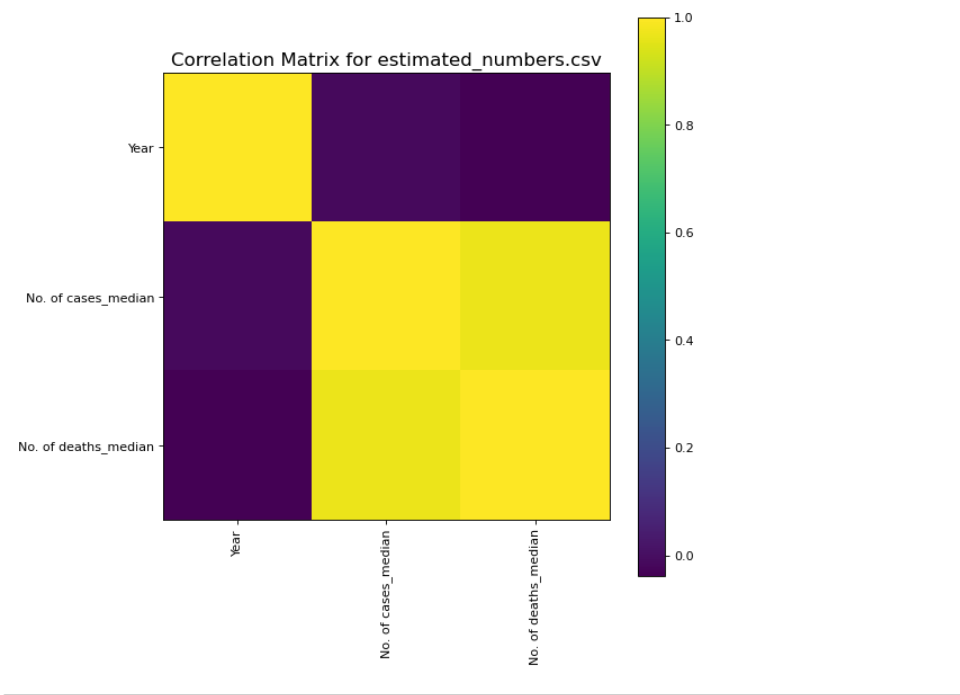




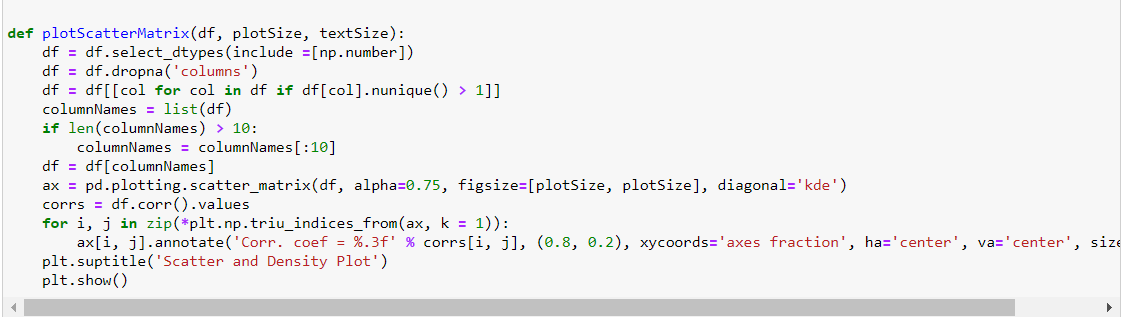
* **CorrelationMatrix**



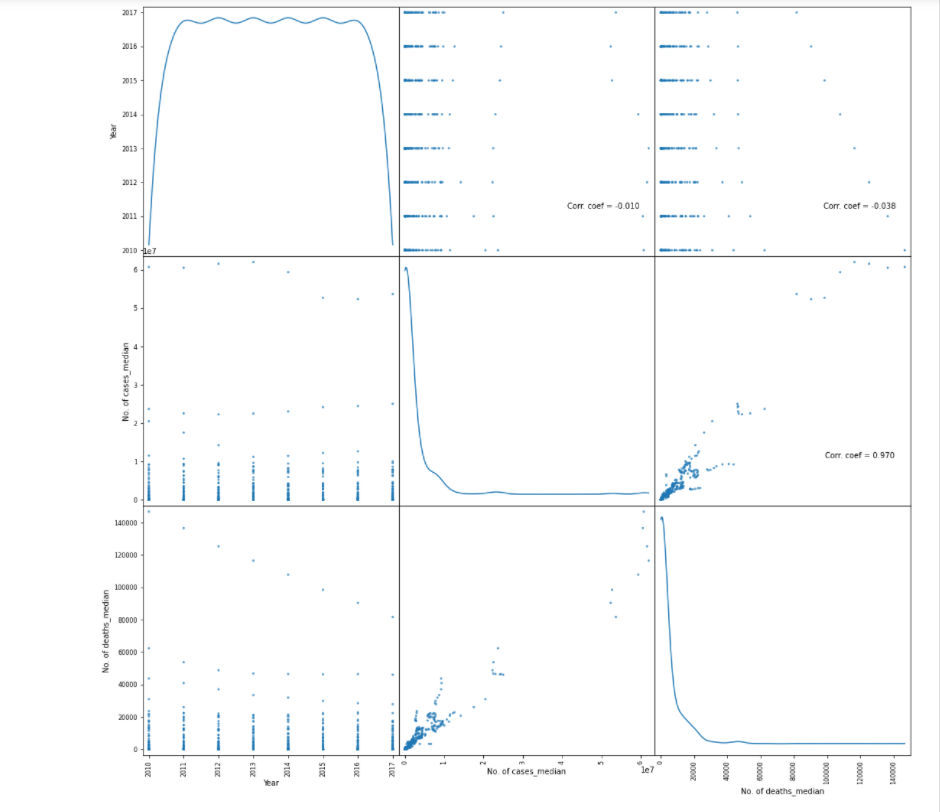




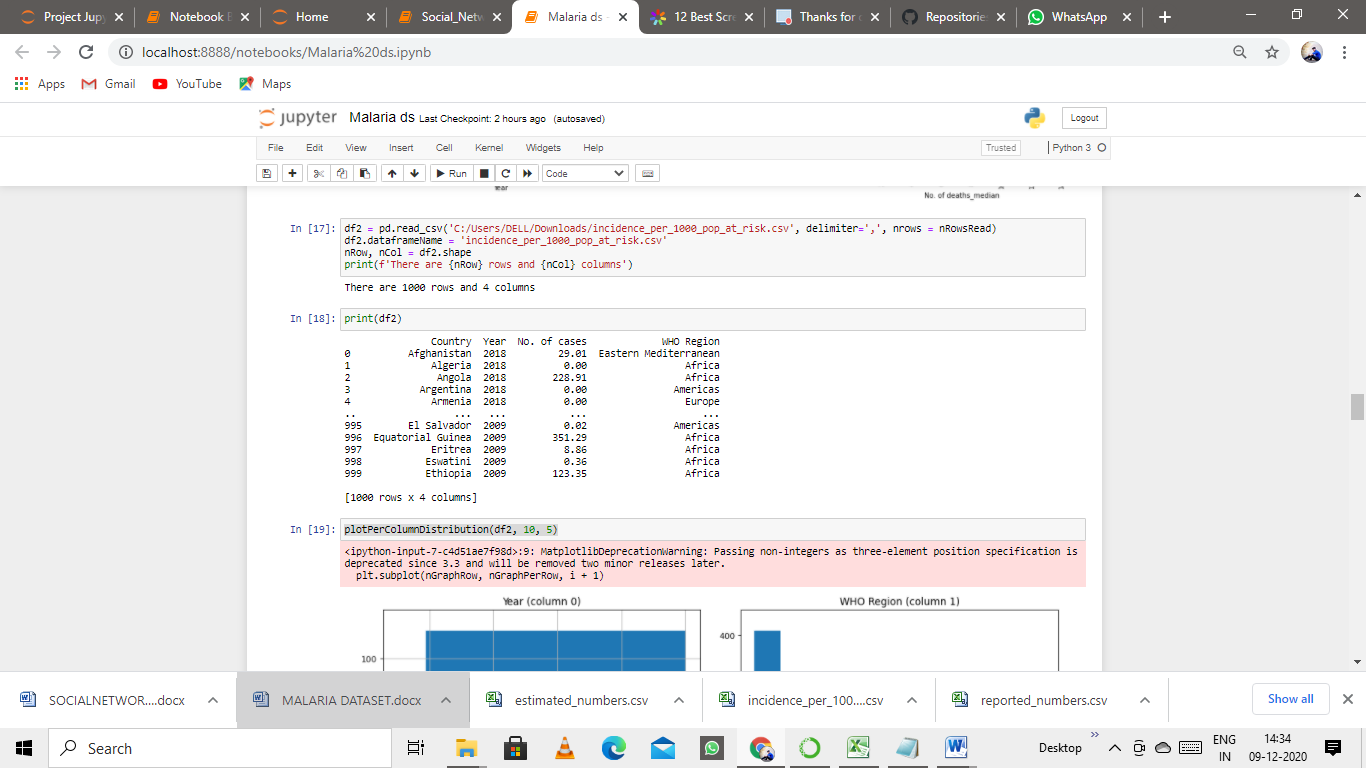
* **ScatterMatix**





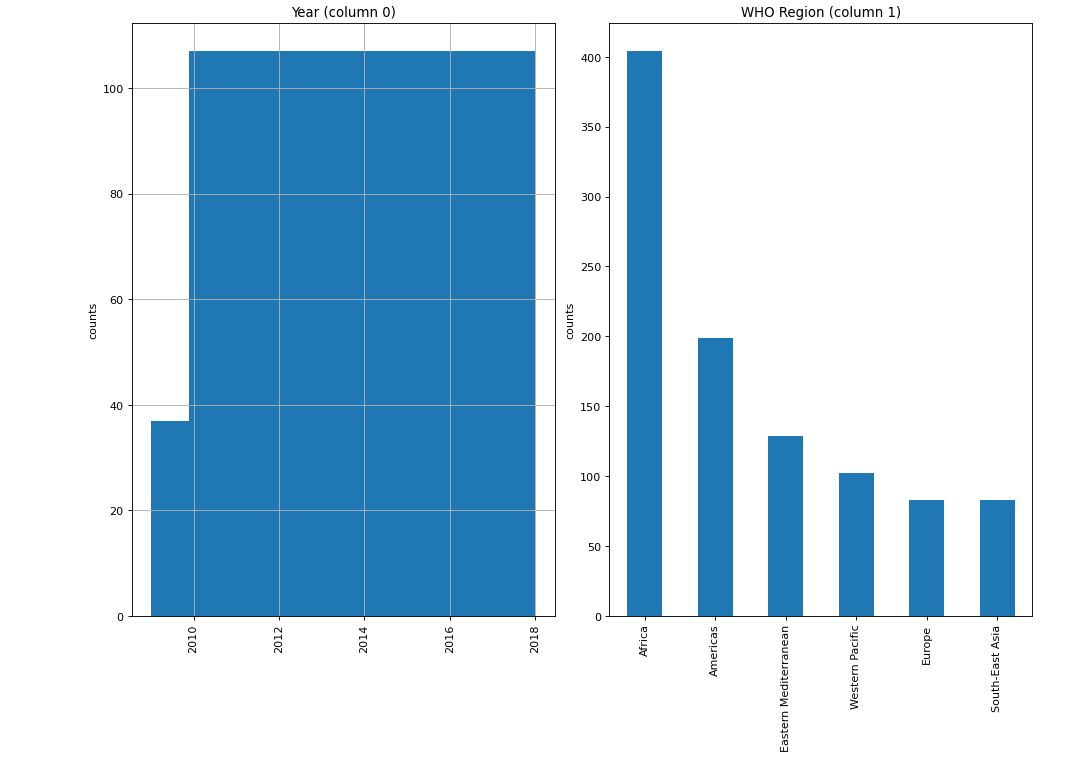


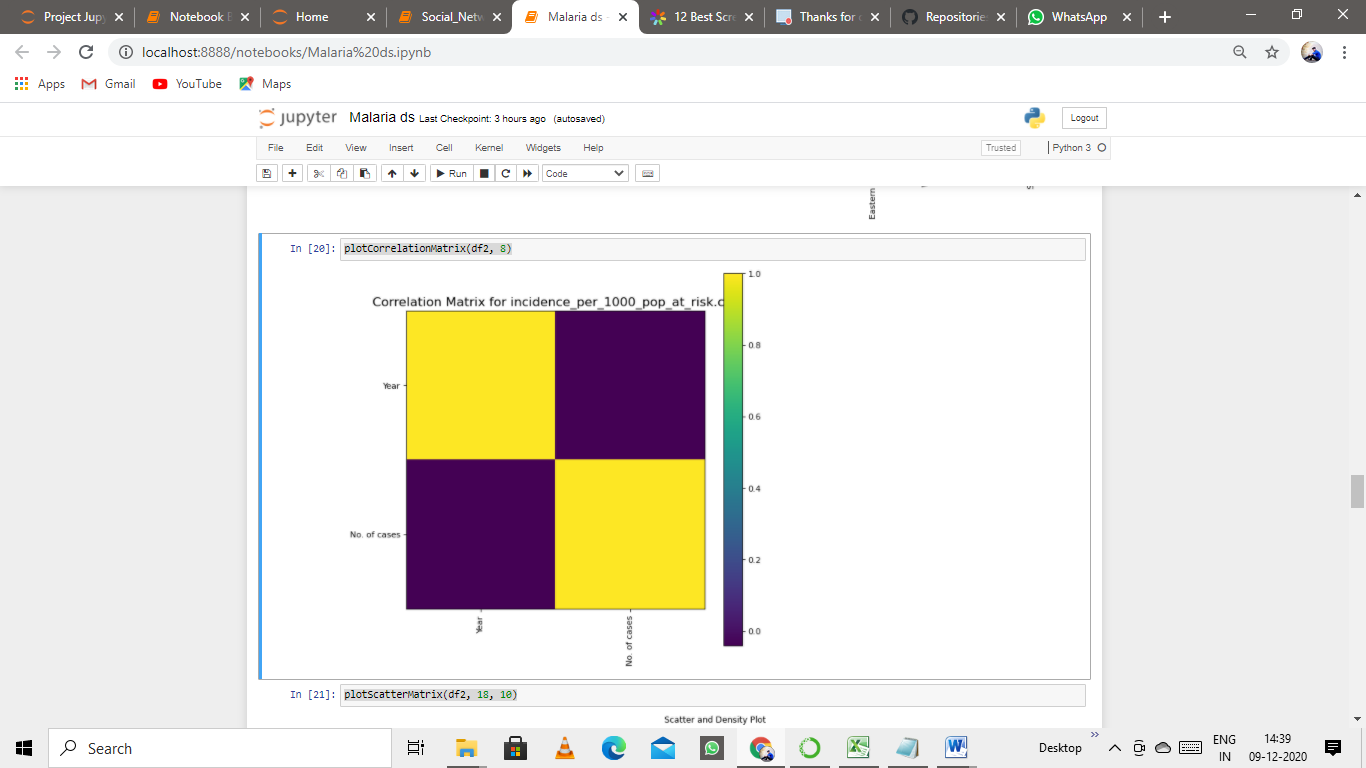
* **Load incidence per 1000 pop at risk.csv dataset**



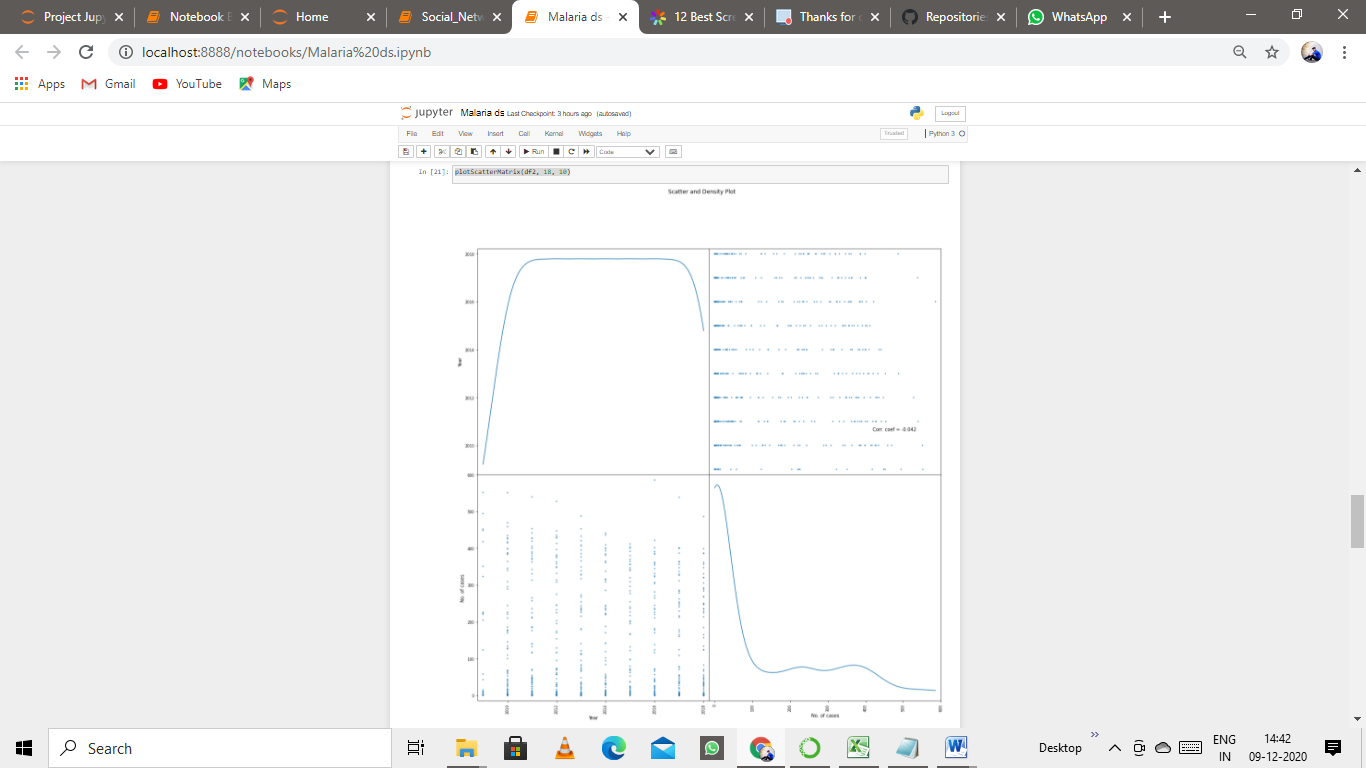


* **Column Distribution for incidenceper1000popat\_risk.csv**

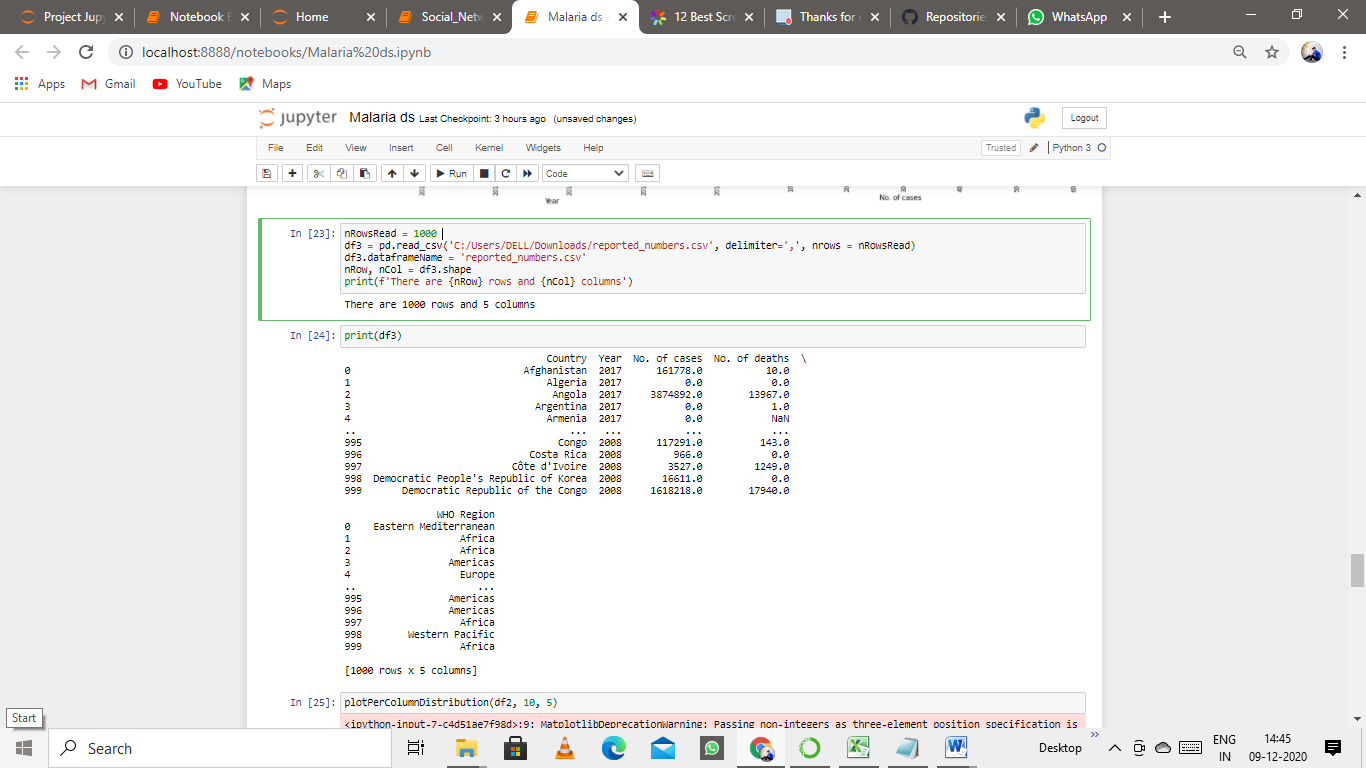
**Correlation Matrix**



* **ScatterMatix**

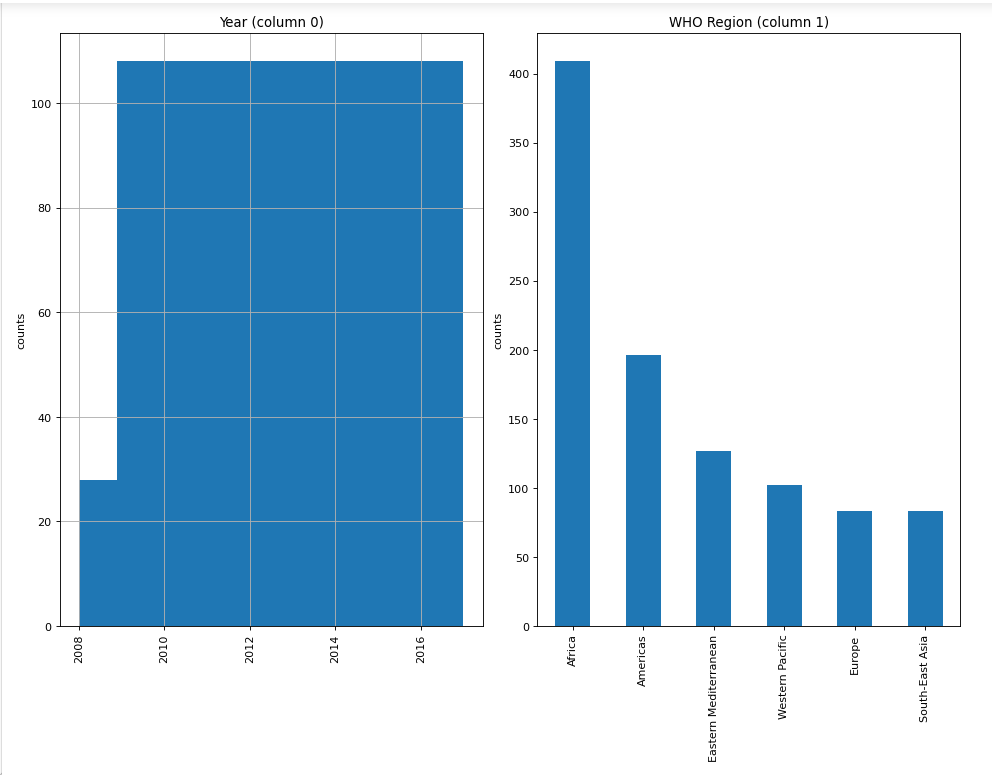


* **Load reported\_numbers.csv dataset**



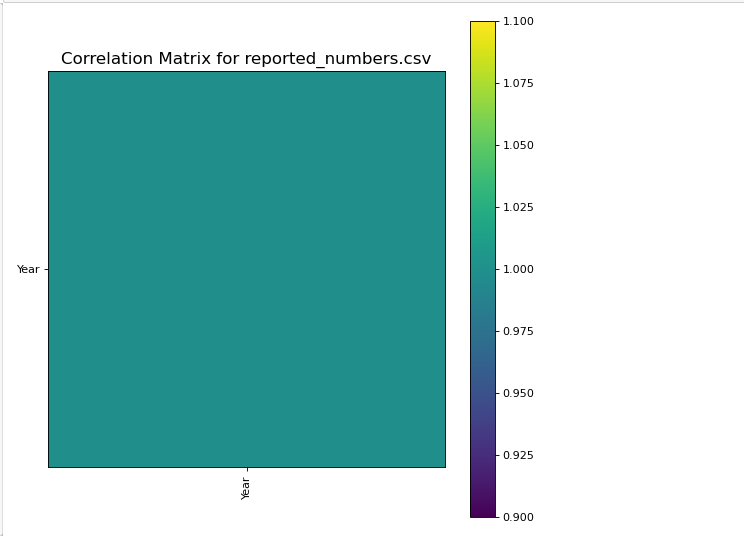
* **ColumnDistribution for reported\_numbers.csv dataset**





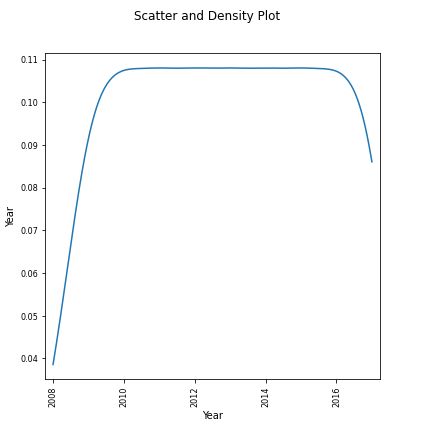
* **CorrelationMatrix for reported\_numbers.csv dataset**





* **ScatterMatrix for reported\_numbers.csv dataset**





* **CONCLUSION:**

Here, I used various plot for visualizing the no of deaths caused by malaria across the world during the period of time in different years by the given data set