DOCUMANTATION:

# AIM:

Explore the given dataset and analyze the number of deaths ,CFR of malaria increases

every year, from scratch.

DATASET:

The **estimated\_numbers.csv** datasetgives the Estimated no of cases across the world

**reported\_numbers.csv** dataset gives the reported no of cases across the world

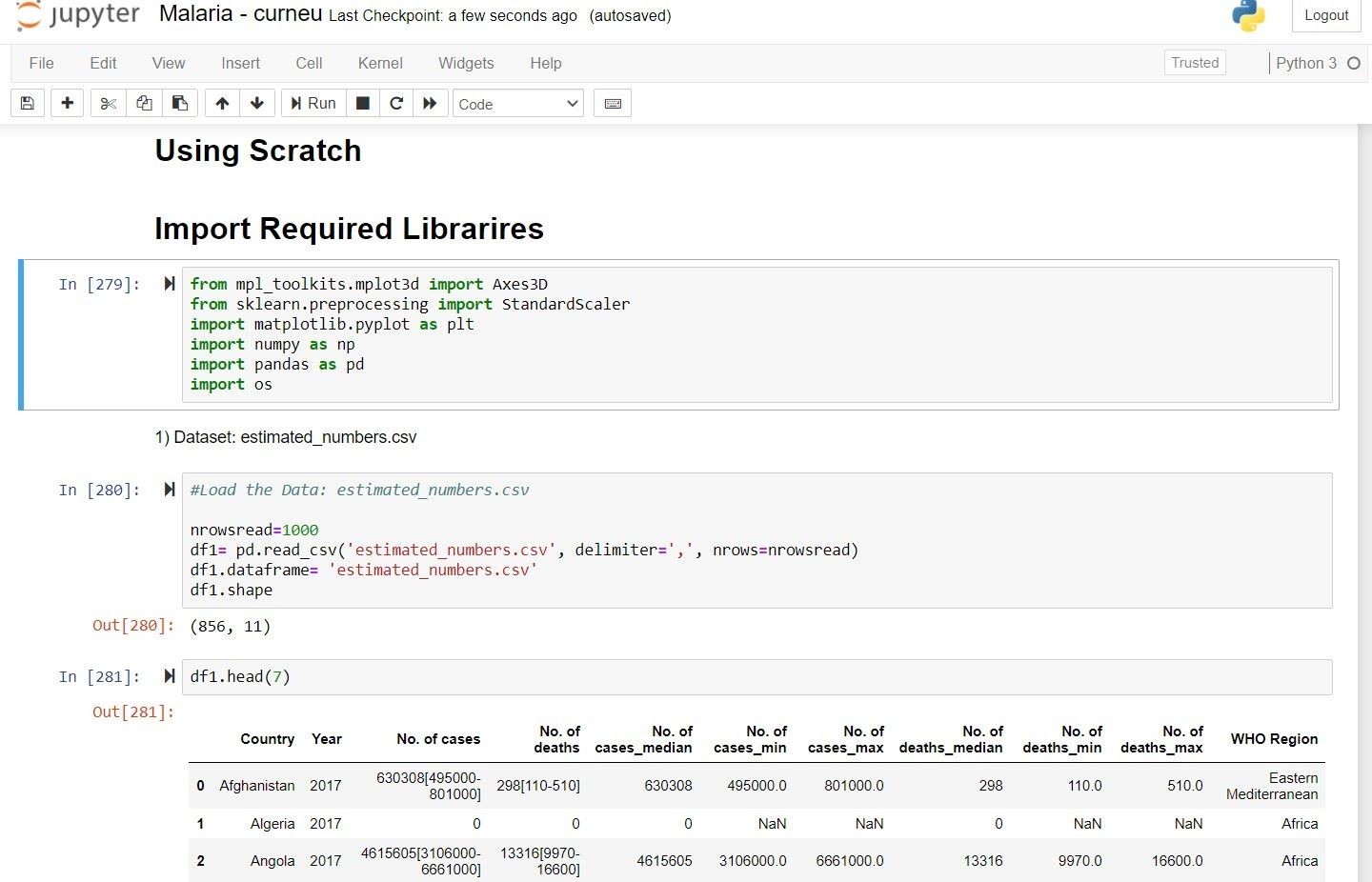
**incidenceper1000popat\_risk.csv** dataset gives Incidence per 1000 people at risk area

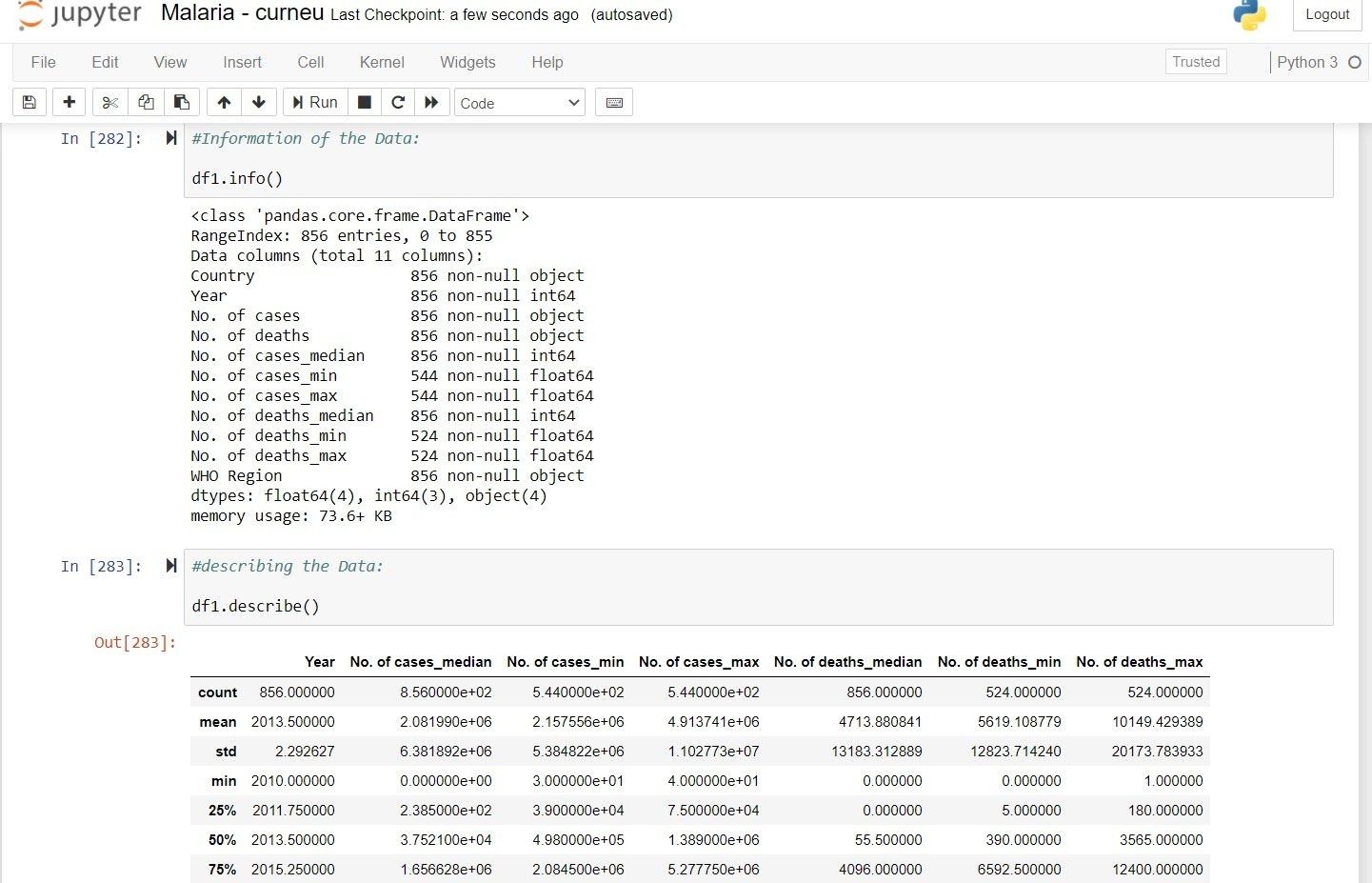
**SOLUTION:**

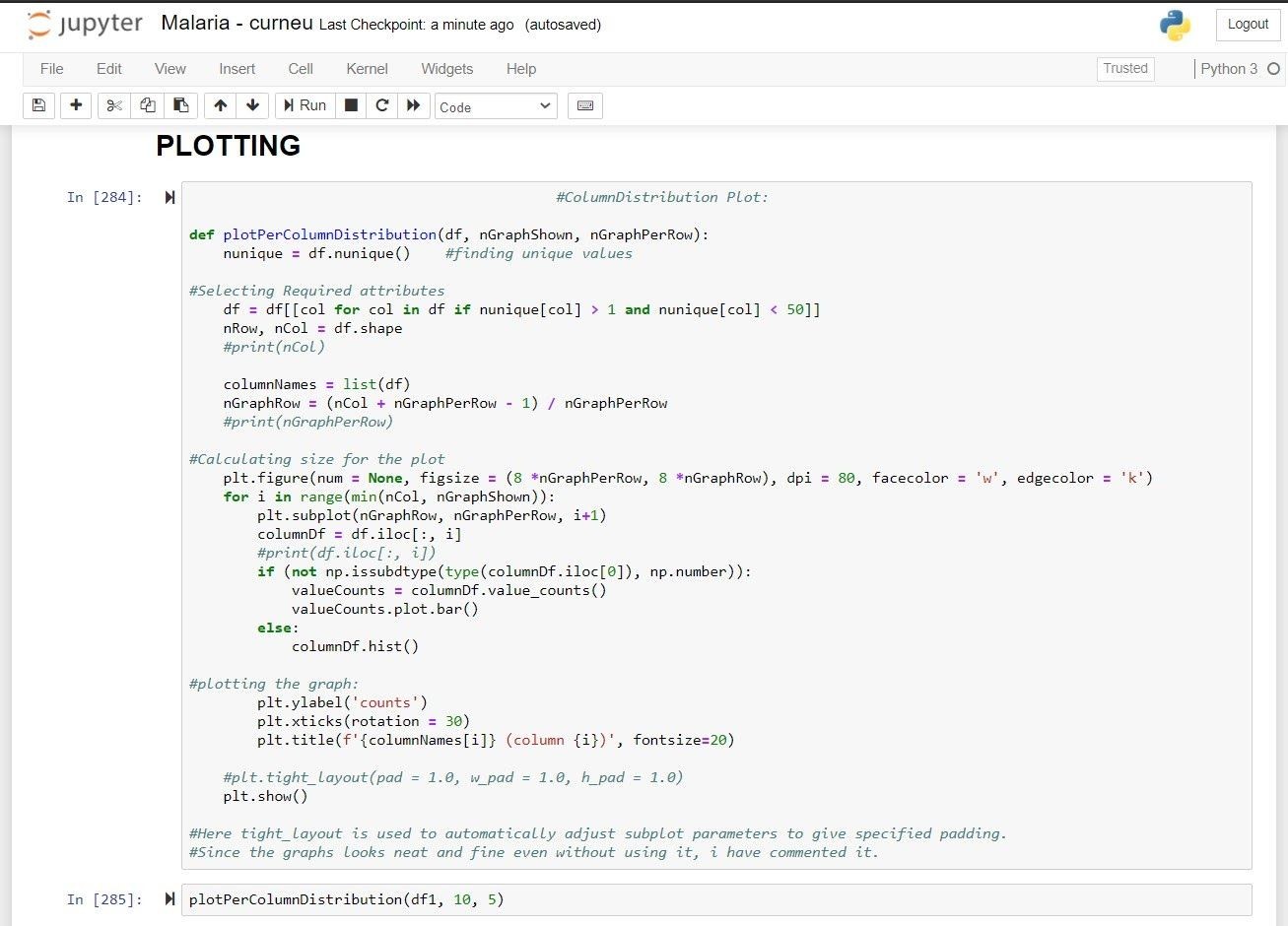
I have done Exploratory data analysis from scratch by creating user defined functions for plots,correlation matric ect. Also I have just explored the data given using the *Column Distribution plot, Scatter Plot, and Correlation Matrix* for all the given three datasets.

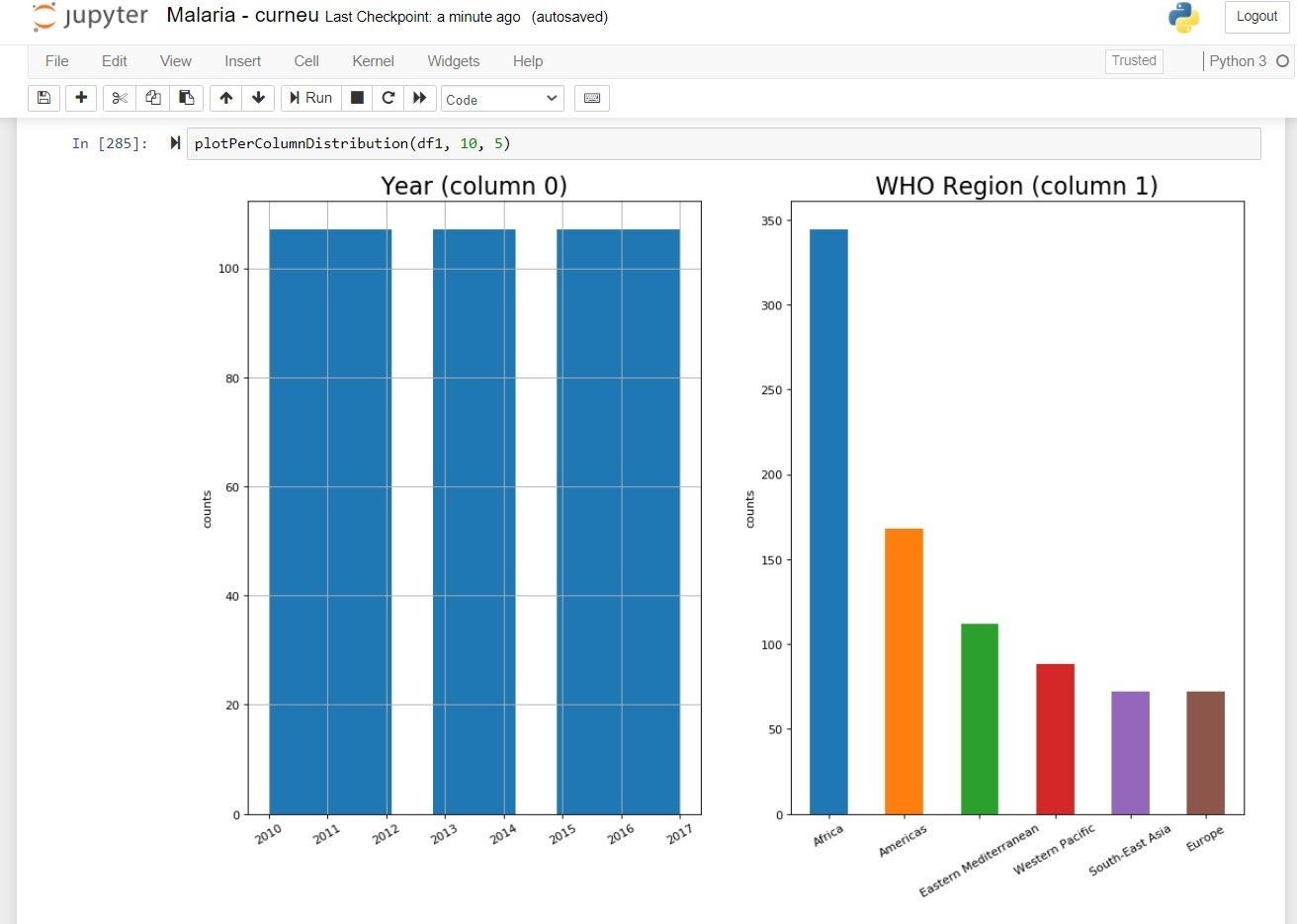
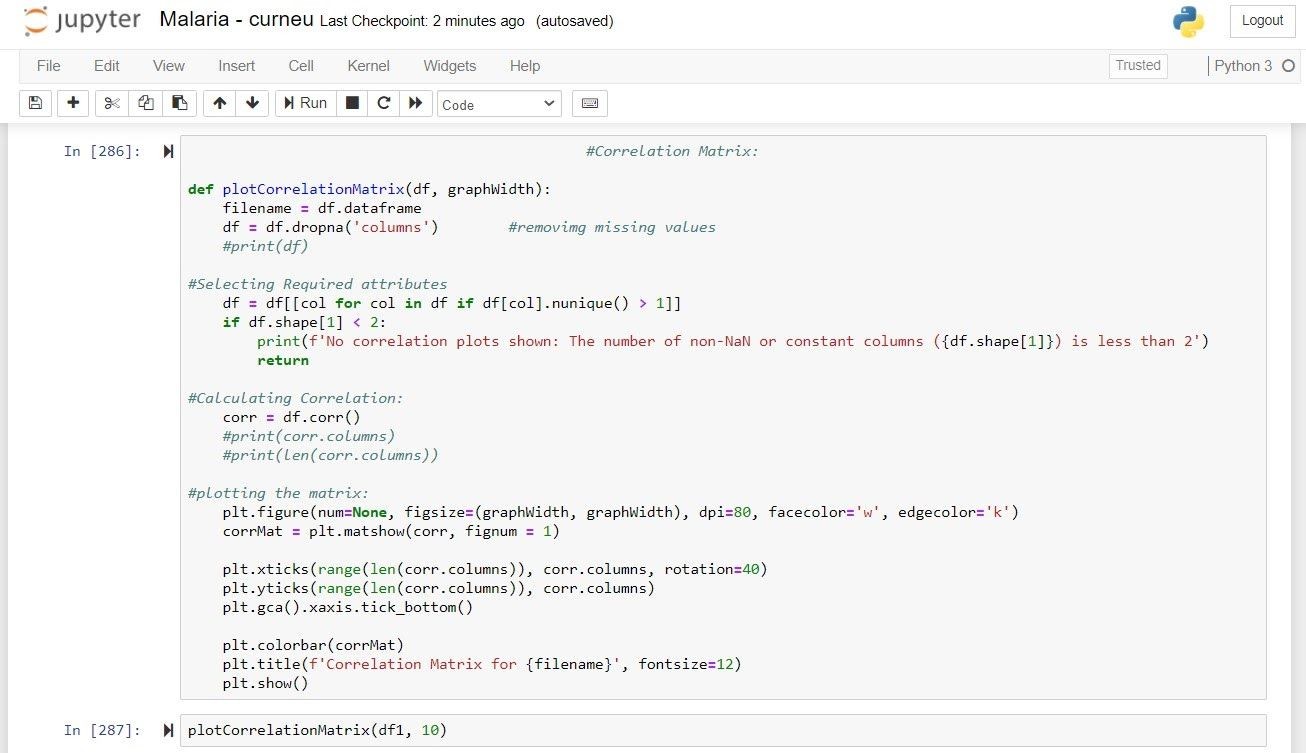
The explanation is inserted along with the code and I have inserted the screenshot of them below,

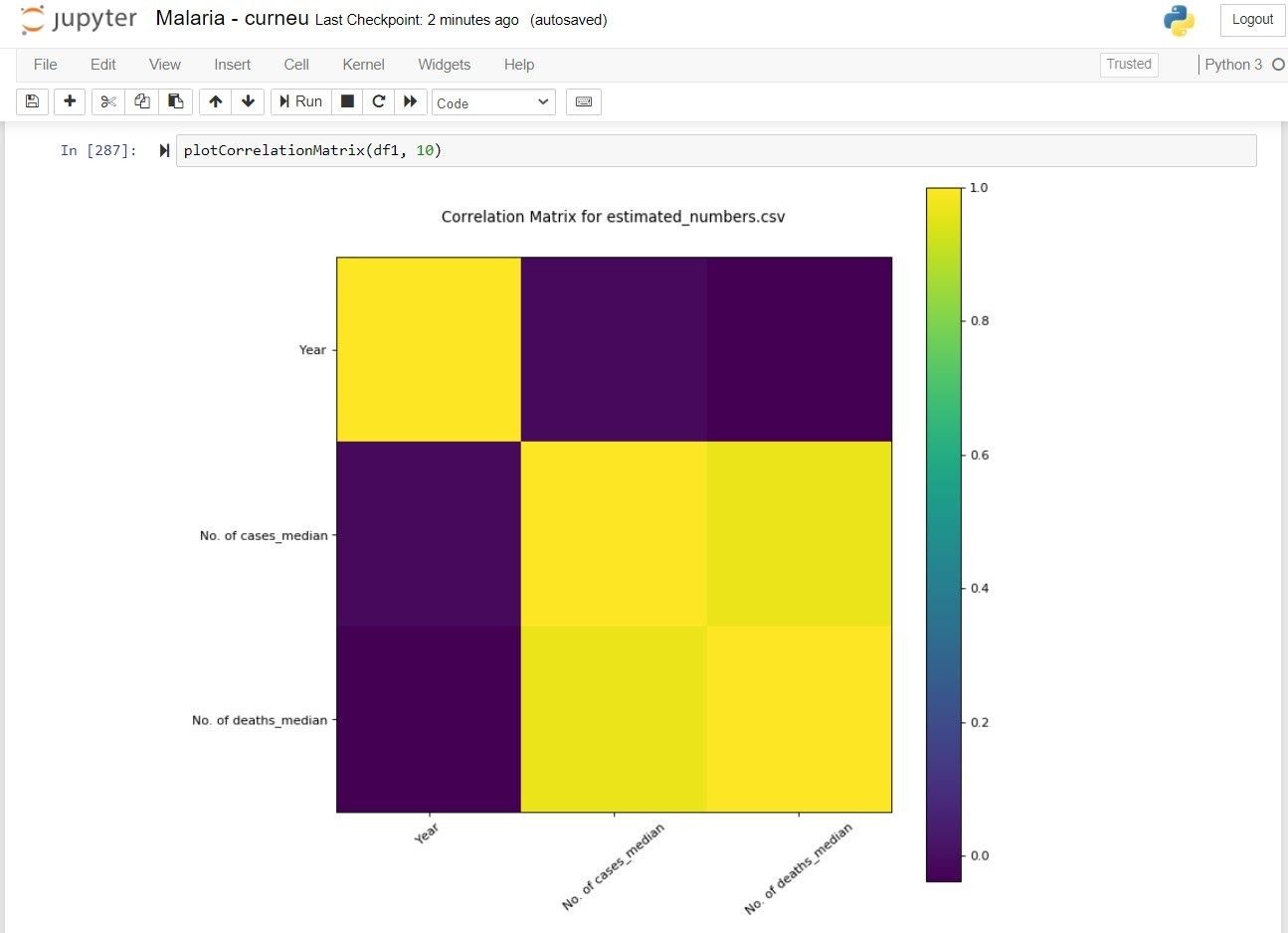
1. *Dataset: ‘estimated\_numbers.csv’*

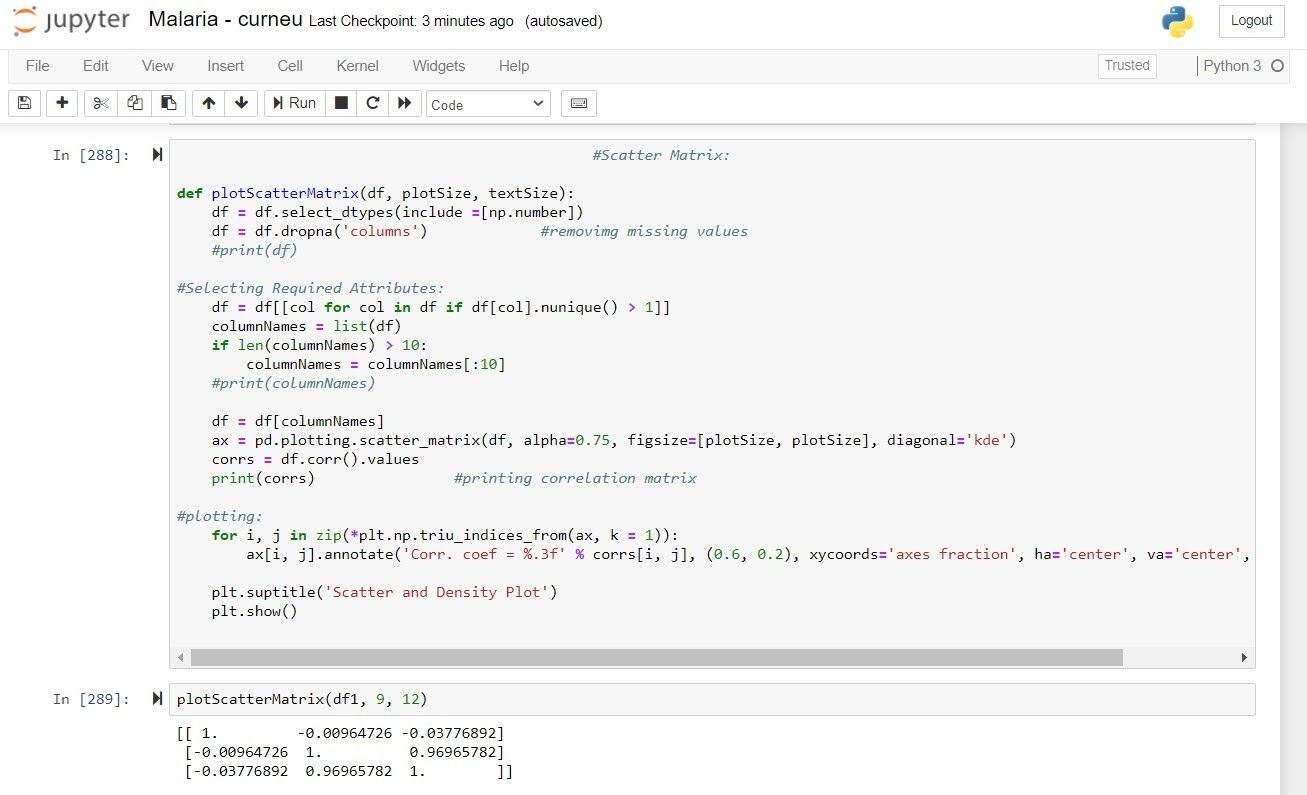


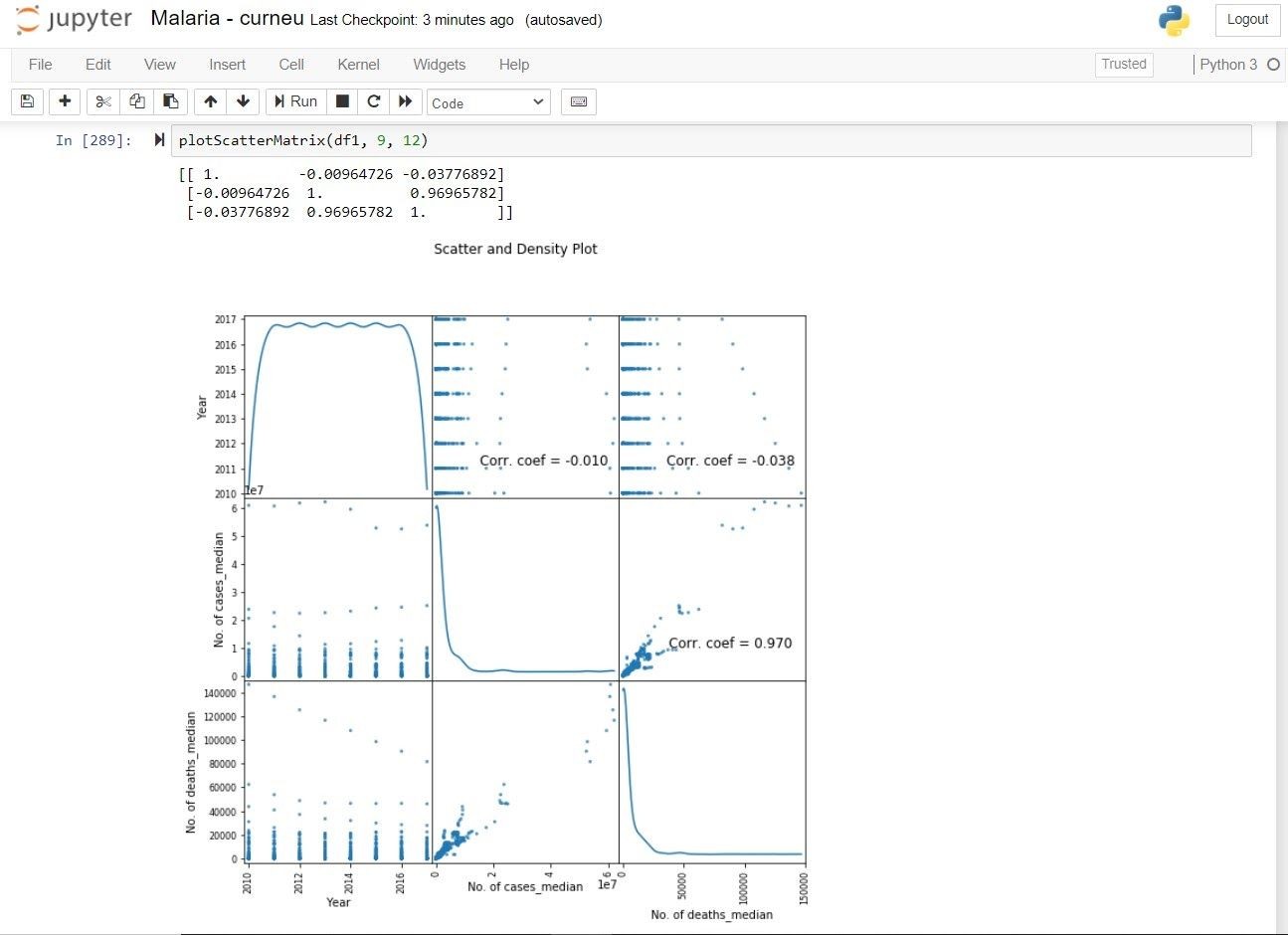




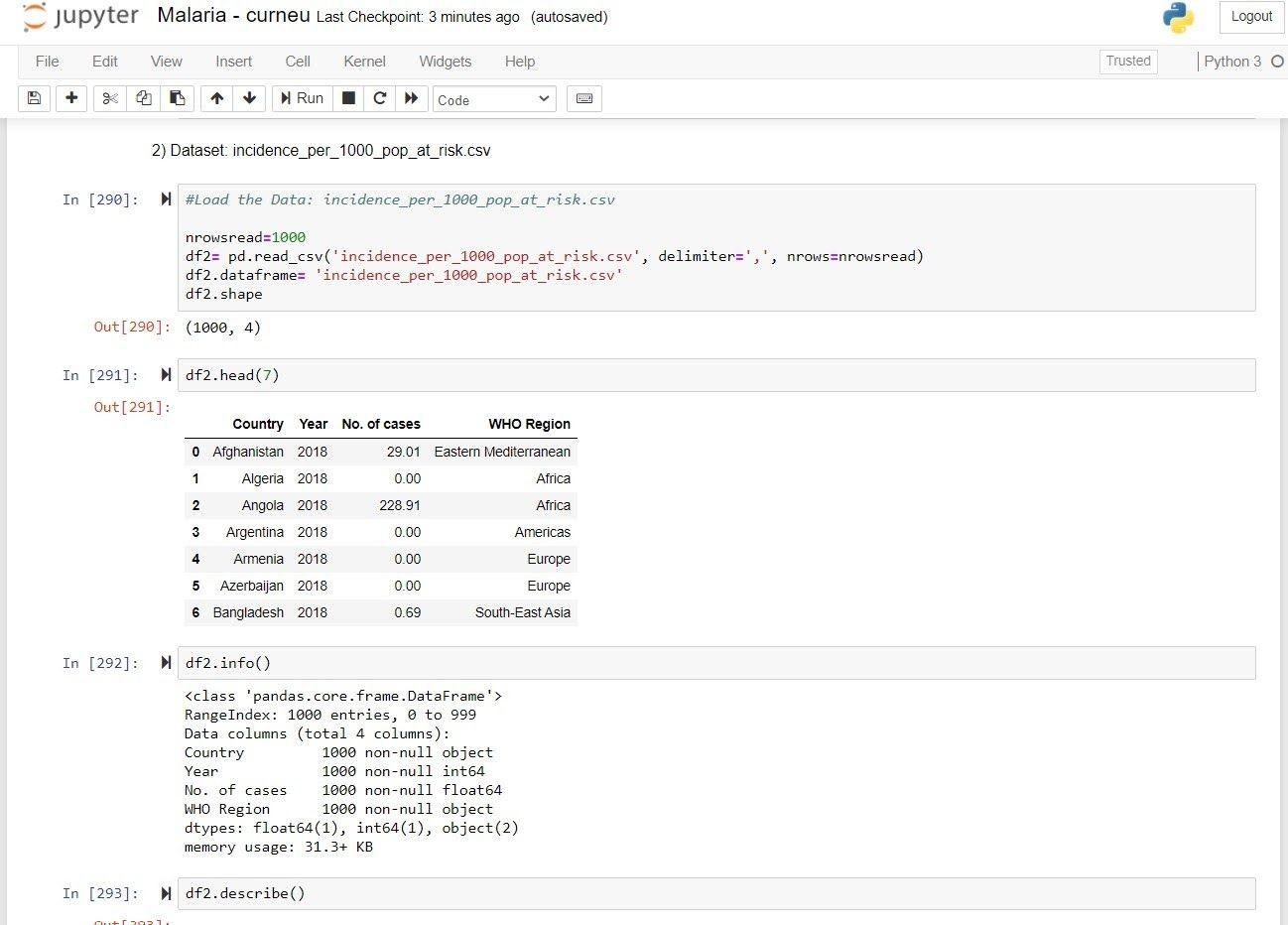


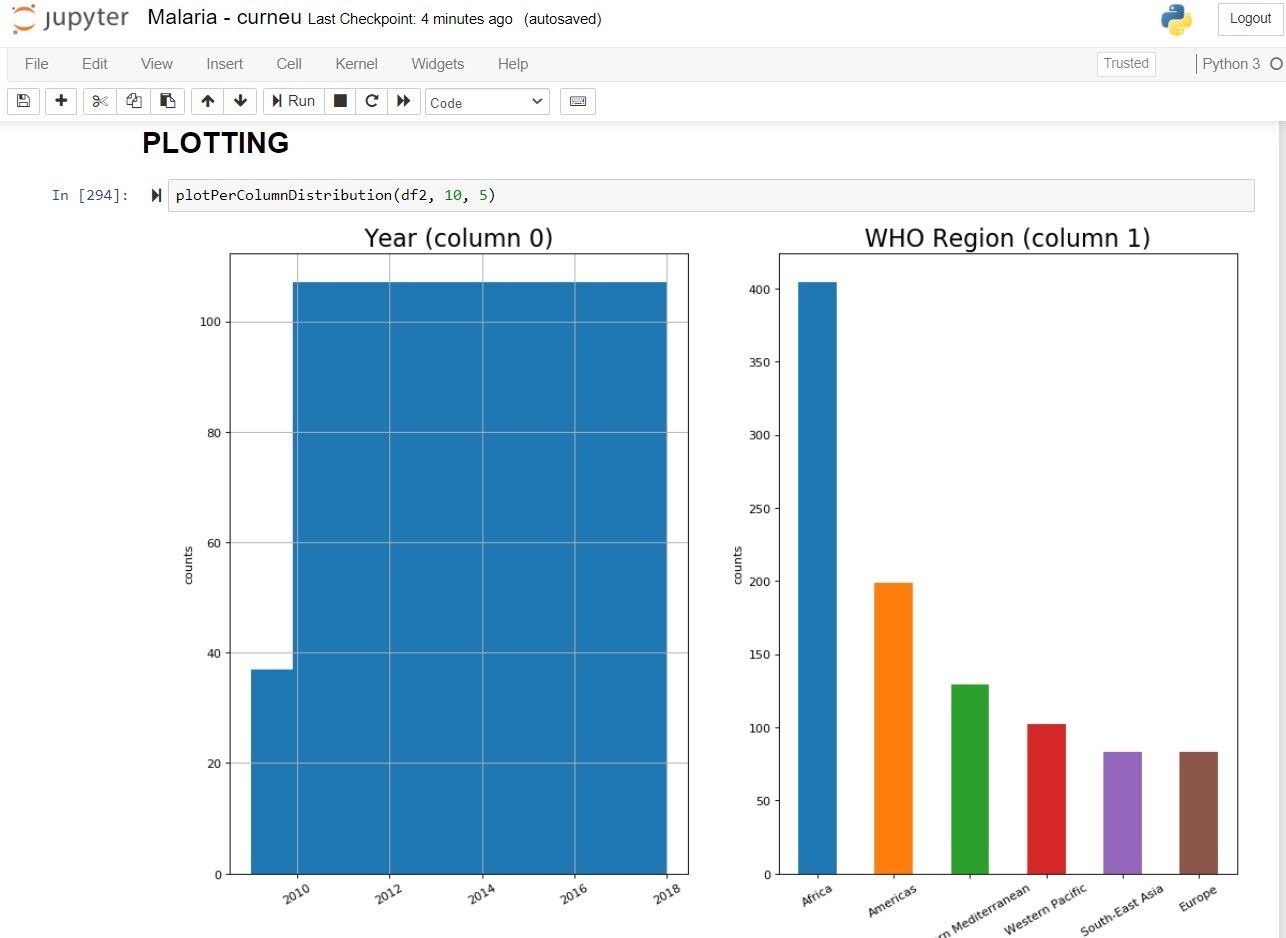


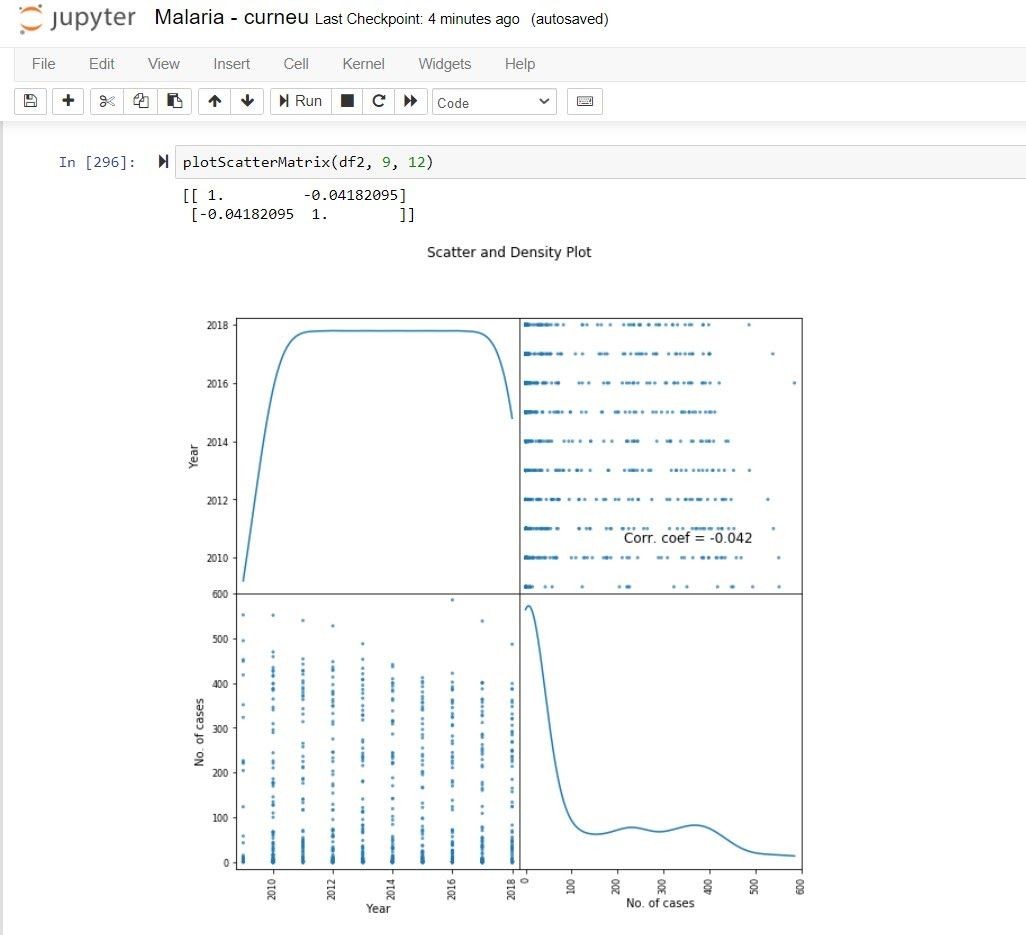




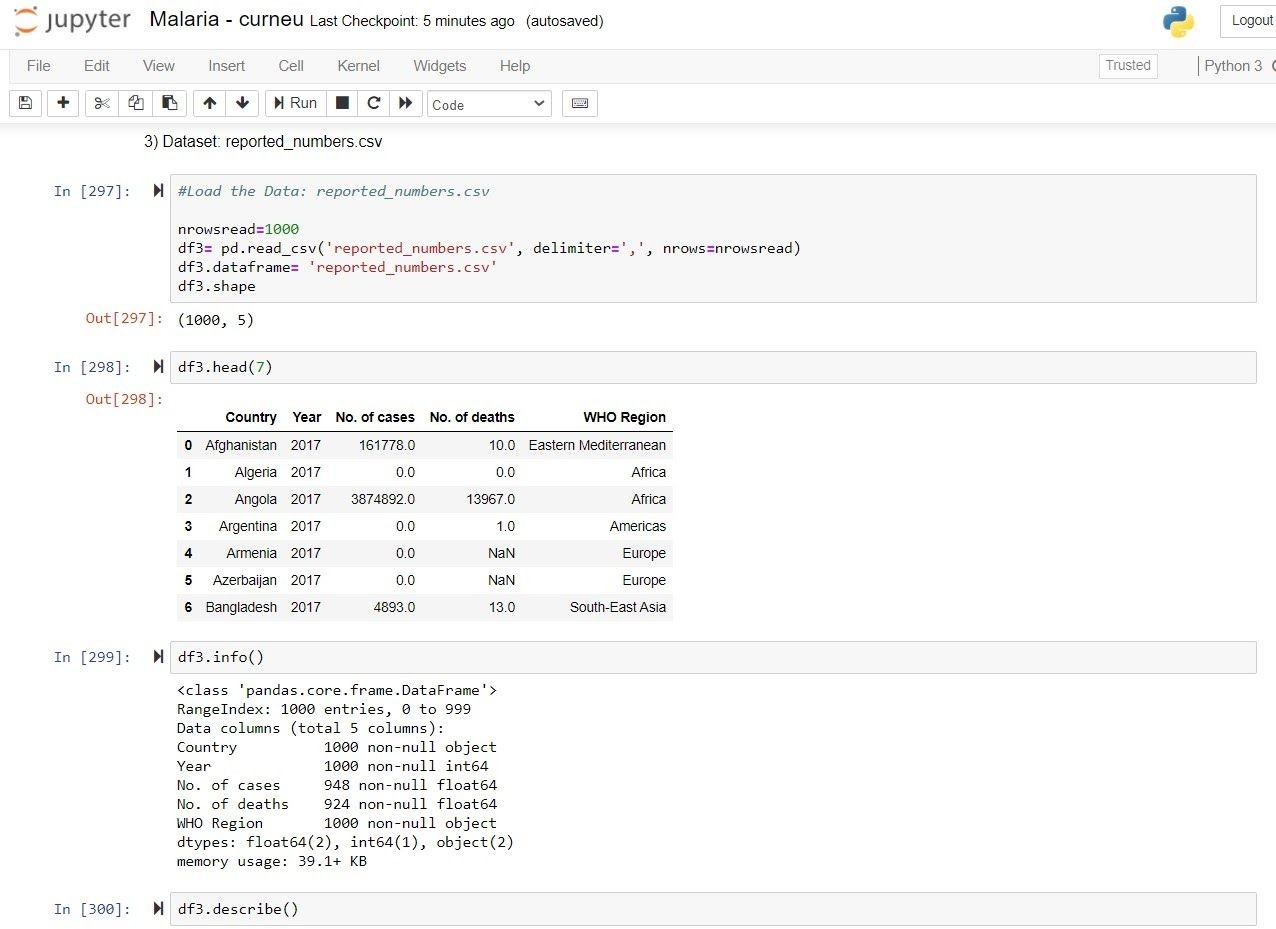
1. *Dataset: ‘incidence\_per\_1000\_pop\_at\_risk.csv’*

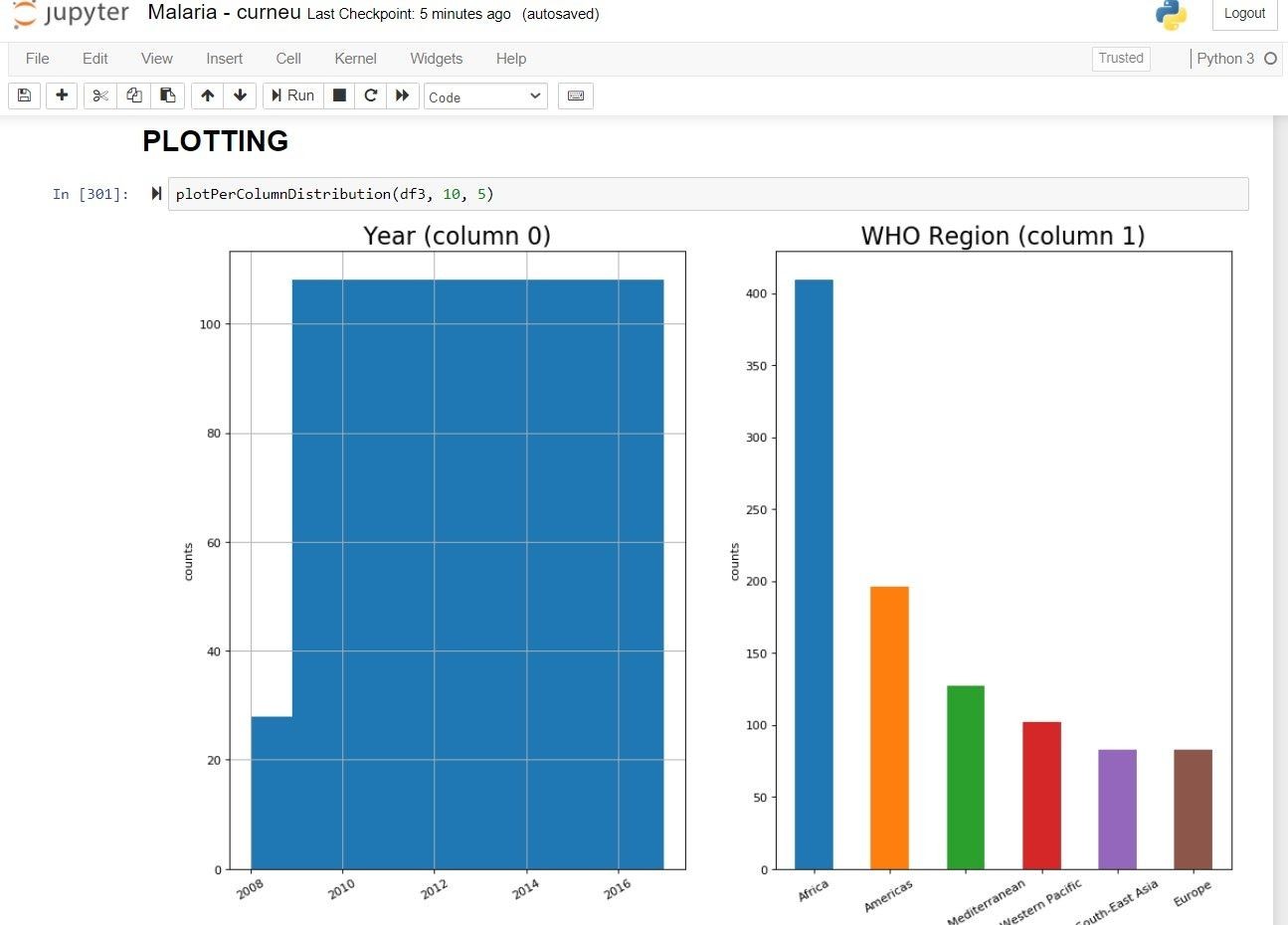
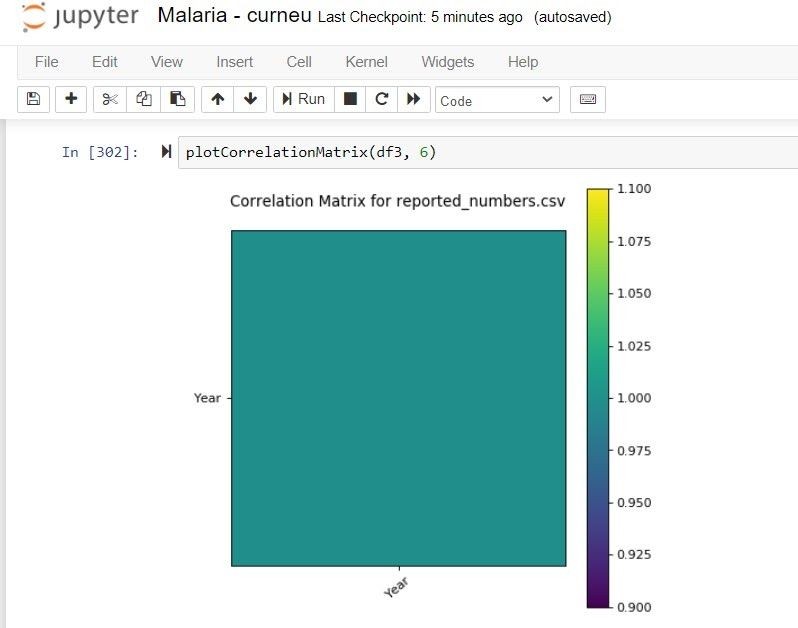


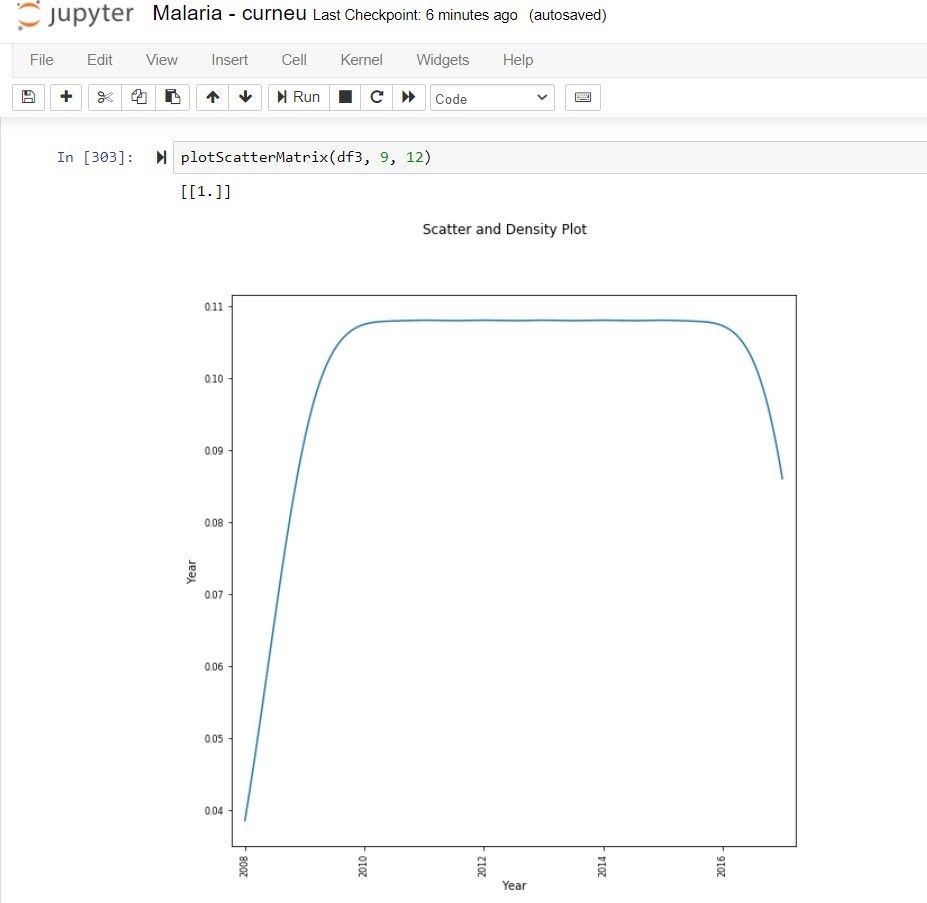




1. *Dataset: ‘reported\_numbers.csv’*







# CONCLUSION:

With help of various plots we have visualized the dataset of malaria across the world during the period of time in different years without predefined functions.

From Scatter plot we can find that from the beginning(ie. 2008) it starts increasing for certain periods and remains the same for a long period of time;