

## Capstone Project One Data Story

With the help of exploratory data analysis, I tried to answer some of the questions which are listed below:

- What is the total number of records in the dataset?
- What were the causes of death in this data set?
- What is the number of deaths per each year from 1999 to 2017?
- What were the top causes of deaths in the United States during this period?
- 10806 no of cases were recorded.
- Accidents (unintentional injuries)', 'Alzheimer's disease', 'Cerebrovascular diseases', 'Chronic lower respiratory diseases', 'Diabetes mellitus', 'Diseases of heart', 'Influenza and pneumonia', 'Intentional self-harm (suicide)', 'Malignant neoplasms', 'Nephritis, nephrotic syndrome and nephrosis'. Unique cases responsible for number of deaths over a period of year 1999- 2017 is depicted on figure 1.

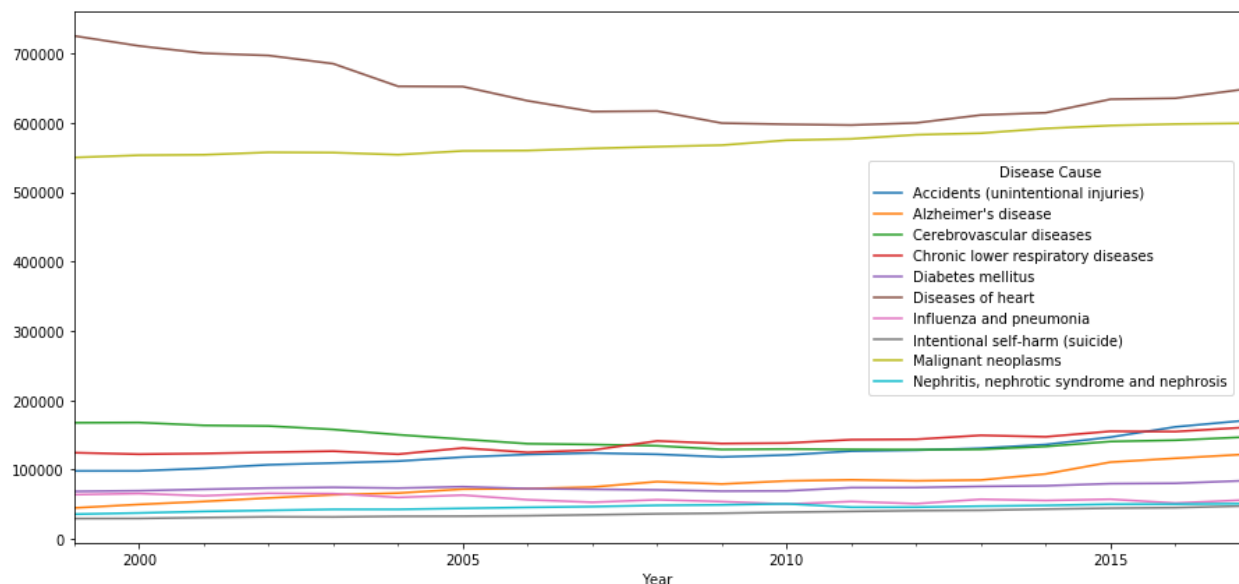


Figure 1:-Causes of Deaths Over a Period of 1999-2017

- The number of deaths declined between 2002 and 2009. Then there was a continuous growth in the number of deaths from 2010 to 2013. Finally, there was a sharp increase in the number of deaths in 2013 and 2014 which can be explained by figure 2.

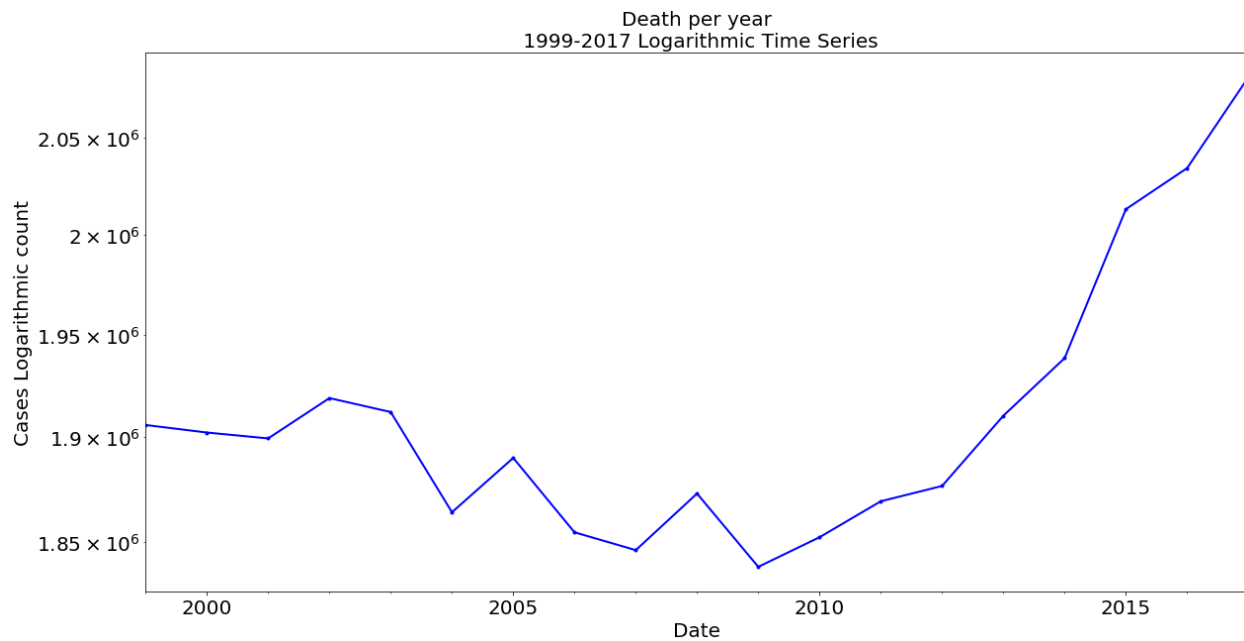
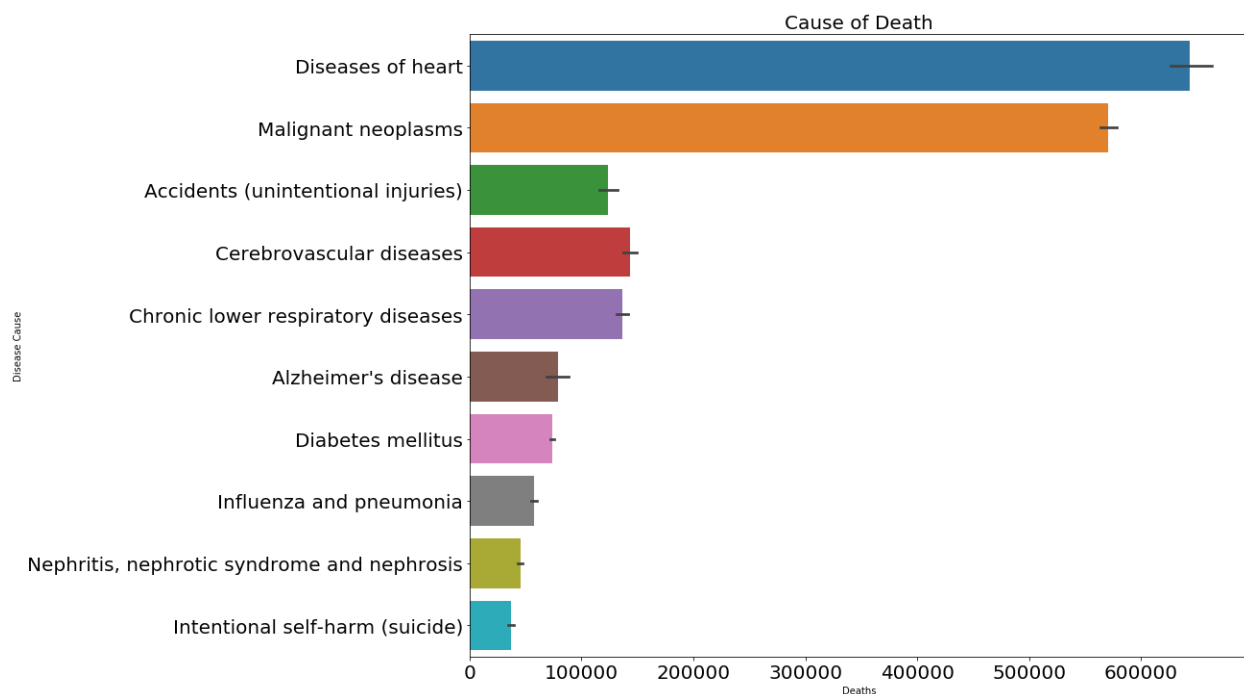


Figure 2 Number of Deaths over the period of 1999-2017

- Diseases of the heart represent the highest causes of death followed by cancer which is visualized on figure 3.



Exploratory data analysis of our data shows no of deaths are increasing and the major contributing causes belongs to non-communicable disease like heart disease and cancer.