

# Forest Fire Exploratory Data Analysis and Prediction

-By Aman Kumar

I used python and imported dataset from obtained from Kaggle (Which is saved locally) using pandas, and using pandas I cleaned data, because there were some missing values some were also invalid. I also had to replace month name because there in some other language. I used Matplotlib and Seaborn for visualizations.

## Inference

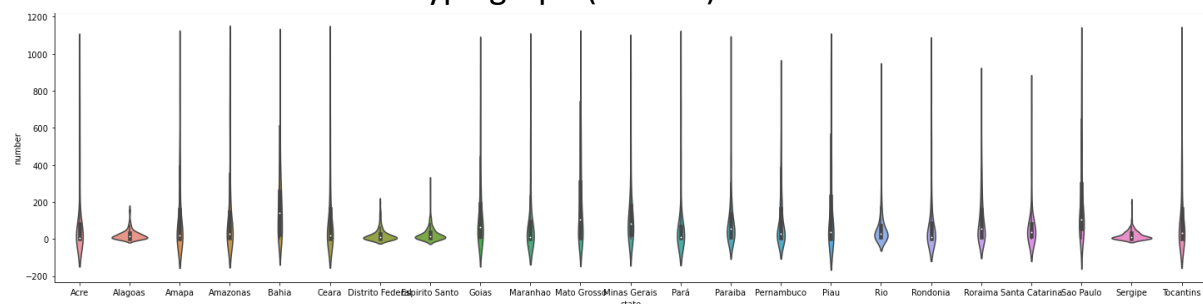
Worst hit states are:

Mato Grosso

Paraiba

Sao Paulo

As can be seen from violin type graph (Line 15)



Worst hit years are:

2003

2015

2016

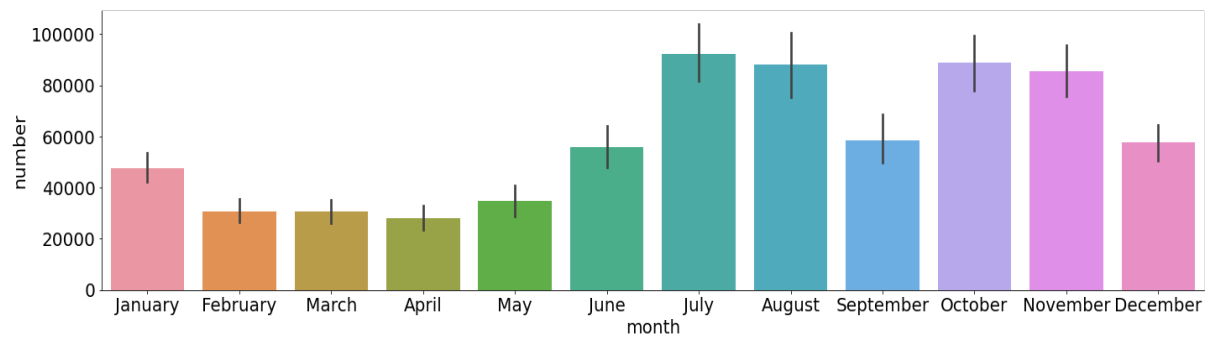
Worst hit months are:

August

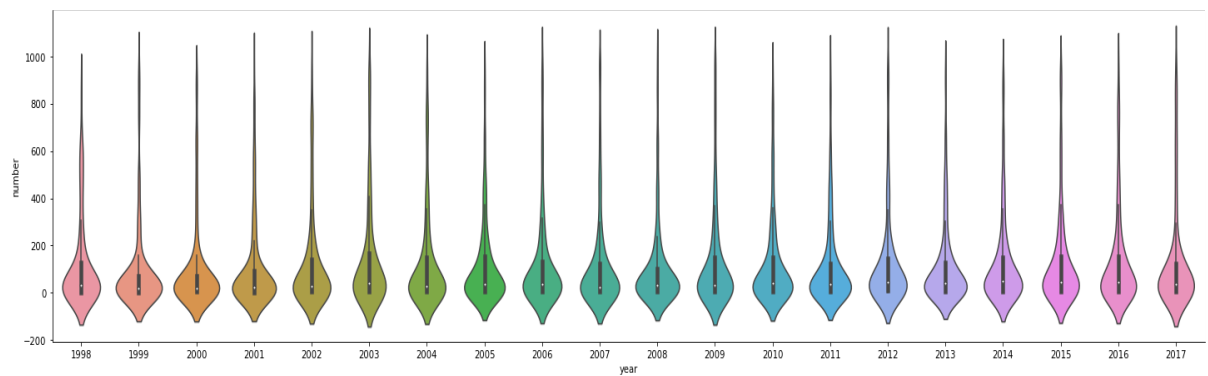
July

November

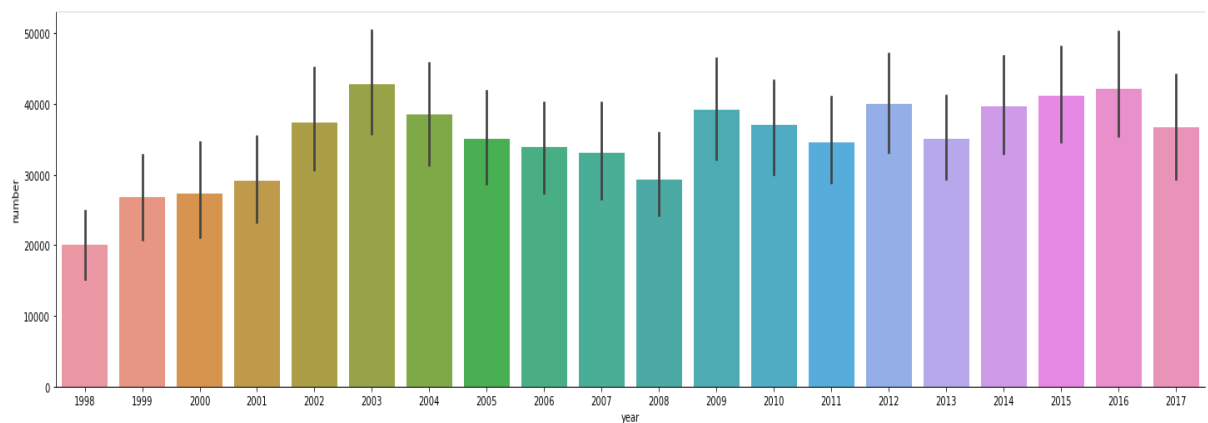
October



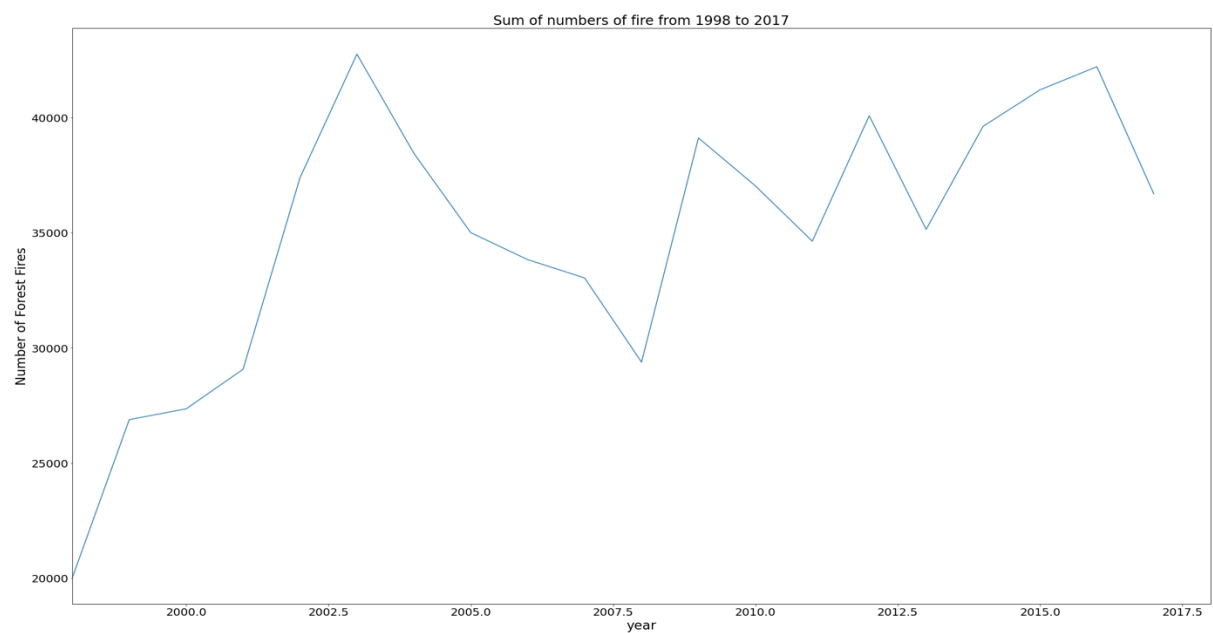
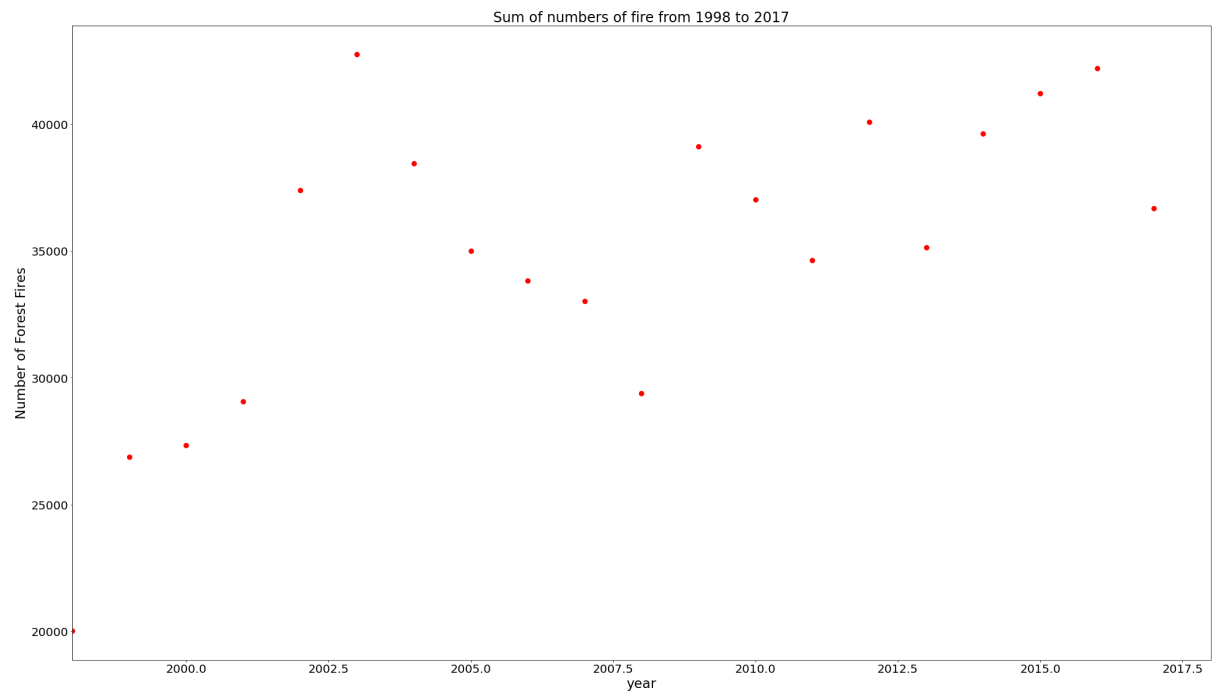
As can be seen from violin type graph, bar type graph and scatter plot graph.  
(Line 17, 18, 19 and 20)



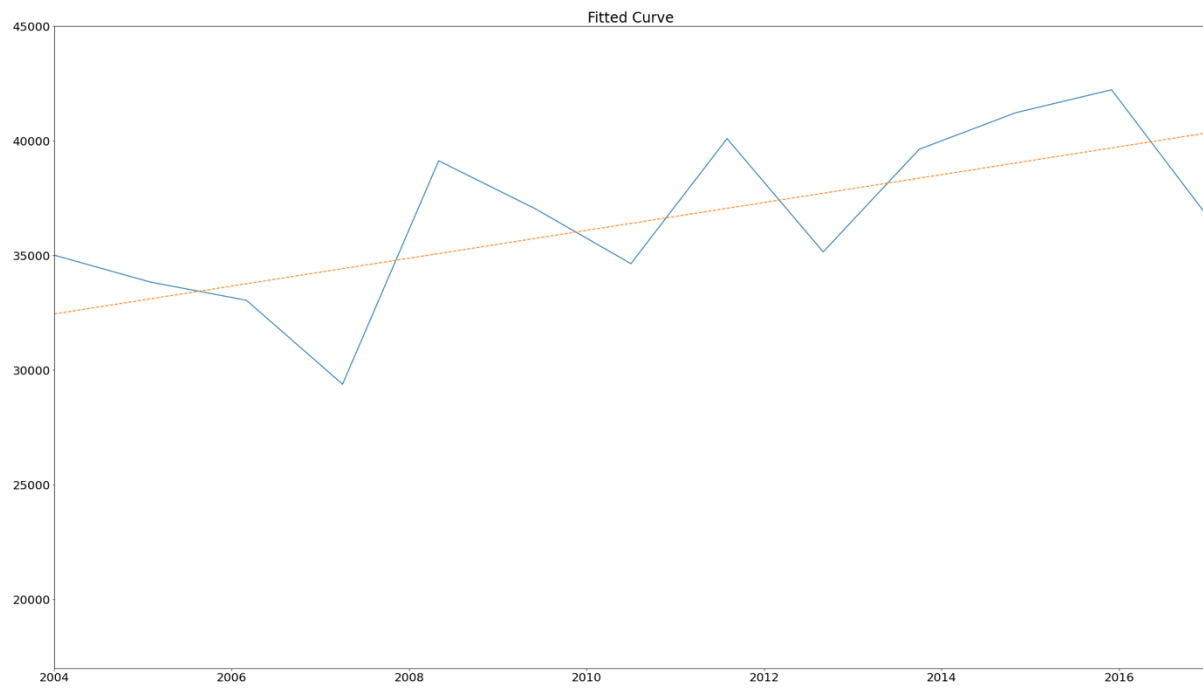
Bar graph for clear analysis



Scatter and spike type graph for visualizing trends and numbers as well



Predicting the trends in cases with 1<sup>st</sup> degree polynomial



### Final prediction results:

2018 - 40942  
2019 - 41549  
2020 - 42156  
2021 - 42763  
2022 - 43370  
2023 - 43978  
2024 - 44585  
2025 - 45192