**US Wildfires Dataset**

**Content:**

This data publication contains a spatial database of 1.88 million wildfires that occurred in the United States between 1992 and 2015 from the reporting systems of federal, state and local fire organisations. The dataset is in the form of an SQLite database, which contains geo spatial data along with the main wildfires dataset. Records include information on dates and times, burn time duration, size, State of origin, Lat and Long coordinates and cause of fire.

<https://www.kaggle.com/rtatman/188-million-us-wildfires?>

**Questions:**

From an initial review of the data available I would like to answer the following questions:

1. Which states have the highest likelihood of a wildfire occurring.
2. Have the occurrences of wildfires increased or decreased over the recording period.
3. What are the main causes of wildfires within the US.
4. Is there a significant difference in causation of fires between varying geographical areas within US.
5. Has there been a significant change in the main causes of fires over the reporting period.
6. Is there any correlation between average size of fires and the State of origin.
7. Are any particular States more susceptible to wildfires at certain times of the year compared to others.

If time I would like to try:

1. Produce a model to predict the likelihood of a wildfire occurring at a particular day of the year.
2. Produce a model to predict the likely cause of a wildfire dependant on state of origin.