**TITLE: WILDFIRES AND ITS IMPACTS**

**Name:** Parikh Malav Arunkumar

**ID:** 30708974

**Tutor Name:** Fahimeh Sadat Saleh and Pratik Bhumkar

**Questions:**

1. Identify the relation between temperatures across all states in US and the number of wildfires that have taken place in that state over the years. Whether there is a direct relation between them or not indicating a rise in temperature resulting in more wildfires?
2. Identify the relation between wildfires, CO2 emissions and respiratory diseases to answer whether they are co-related or not? Does the rise in CO2 emissions levels and number of wildfires result in rise of respiratory diseases?
3. Explore how wildfires impacts the economy, business and population across different states? Does rise in wildfires impact the number of business and shut down business and impacts economy and is there a change in population due to it?

**Data Sources:**

1. **Population dataset** contains information about population by years and zip code

https://data.world/lukewhyte/us-population-by-zip-code-2010-2016

1. **Zip code dataset** contains data about every zip code. Its location, state and every related data.

https://simplemaps.com/data/us-zips

1. **Number of Wildfires** contains month, year and state wise data of number of wildfires and acres burned

<https://www.ncdc.noaa.gov/societal->impacts/wildfires/month/12?params[]=acres&params[]=fires

1. **Wildfires by state** contains data about number of wildfires and acres burned in every state every year

https://www.iii.org/table-archive/23284

1. **GDP** contains data about GDP of US, GDP per capita, inflation and unemployment rate

https://www.imf.org/external/pubs/ft/weo/2019/02/weodata/weorept.aspx?sy=1980&ey=2019&scsm=1&ssd=1&sort=country&ds=.&br=1&pr1.x=21&pr1.y=11&c=111&s=NGDP\_RPCH%2CNGDP%2CNGDPPC%2CPCPIPCH%2CLUR%2CBCA\_NGDPD&grp=0&a=#download

1. **CO2 emissions by state** contains year and state wise data of CO2 emissions

https://www.eia.gov/environment/emissions/state/

1. **Respiratory Disease** contains data about the number of deaths caused by respiratory diseases across states in different years

https://www.cdc.gov/nchs/pressroom/sosmap/lung\_disease\_mortality/lung\_disease.htm

1. **Temperature by State and Year** contains data about state wise and year wise temperatures

https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data

1. **Business Data (Table Name: State)** contains data about every state wise and year wise new business establishments, closures, job creations, job closings and all related data

<https://www.census.gov/programs-surveys/bds/data/data-tables/legacy-establishment->characteristics-tables-1977-2014.html

* Data sources 3,4,8 will help answer question 1.
* Data sources 3,4,6,7 will help answer question 2.
* Data sources 1,3,4,5,9 will help answer question 3.
* Data source 2 will be used when wanted to join zip code related data to another datasets.

**Description of data sources:**

1. **Population dataset:** Textual data as comma separated values (rows: 33121, columns: 9)
2. **Zip code dataset:** Textual data as comma separated values (rows: 33100, columns: 18)
3. **Number of wildfires:** Textual data as comma separated values (rows: 249, columns: 4)
4. **Wildfires by state:** Tabular data .xlsx format. 9 files each containing same number of rows and columns (rows: 56, columns: 3)
5. **GDP:** Tabular data .xlsx format (rows: 7, columns: 46)
6. **CO2 emissions by state:** Tabular data .xlsx format (rows: 53, columns: 33)
7. **Respiratory disease:** Textual data as comma separated values (rows: 301, columns: 5)
8. **Temperature by state and year:** Textual data as comma separated values (rows: 2296, columns: 5)
9. **Business data:** Textual data as comma separated values (rows: 256, columns: 27)