Climate Change Analysis

Problem

Climate change has become a huge problem that affects the very existence of life on the planet and a majority of climate change in the recent years in our industrial age is caused due to human activity. This caused a dramatic increase of temperatures on both land and water surfaces, meltdown of ice caps, rise in sea levels, increase in the magnitude and frequency of hurricanes, and extinction of many species.

This adverse effects can be slowed down / stopped / reversed with changes in human activity and proper education. Even in this science era, many people - often educated, and in powerful influential positions - disregard the climate change phenomenon as a hoax and refuse to take any action.

This project is to show the correlation between the rise in temperatures due to rise in atmospheric greenhouse gases (CO2) caused by human activity and to conclude that controlling the CO2 emissions can regulate climate change phenomenon.

Client Use Case

This solution is applicable to the general public as well as many companies in various industries like automobile, fossil fuel extraction, food processing, chemicals, etc. that cause an increase in atmospheric greenhouse gases.

The solution provides the client an education of how rise in CO2 levels is causing a rise in average temperatures, which is the reason for all the adverse effects on various life forms including humans on Earth.

This study would encourage companies to design their processes to be more efficient and build clean energy facilities, and encourage people to be cautious about their actions to reduce the carbon footprint.

Data Source

1. *Temperature* - All the temperature data necessary for the study is downloaded from National Centers for Environmental Information data repository of NOAA.
    a. The dataset is a global daily temperature readings from Jan 1975 to Oct 2018. The data set includes daily data for - Average temperature, Minimum temperature, Maximum temperature, Precipitation, details regarding thunderstorms, hail, and tornado occurrence for the day
2. *CO2 Levels* - All the CO2 levels data is downloaded from Earth System Research Laboratory - Global Monitoring Division of NOAA
    a. The dataset includes monthly data of concentration of CO2 measured at various locations in the US (different states). Even though the data is a global measure, data collected in various states is assumed to be state level CO2 data for our analysis
3. *Land Cover* - The land cover data is downloaded from the National Land Cover Database of USGS.

Methodology

1. Problem approach
    a. The aim of the project is to show the trends in temperature increase and the prime factor is the increase in CO2 levels due to human activity in the industrial age.
    b. Since the cause has many factors and the data is available to show the correlation between factors and effects, this study will be a **supervised** learning.
    c. This supervised learning model is going to be **regression and time series** forecast model to show the correlation, recent trends and future trends
    d. The variable this project is aiming to predict is temperature - average, min, and max by state in USA
    e. This prediction models are based on variables - CO2 levels, and land cover data variables (tree cover data specifically, which contributes to the atmospheric CO2 levels)
    f. The training data will be the raise of CO2 levels. The idea is to build a forecast models with varying levels of CO2 data once the relationship is established and show the sensitivity analysis by changing the trend in factor variables - increasing trend, prediction trend, decreasing trend in CO2, and also show how CO2 is changed in the atmosphere by varying the tree cover predictors

Output

The deliverable will be a presentation with associated visualizations of sensitivity analysis, and an explanation of the approach.