**IS CLIMATE CHANGE A MYTH?**

**Problem:**

There are strong claims that global warming is happening rapidly and is already having significant and harmful effects on our communities, our health, and our climate. According to the scientific community, the impacts are numerous. Some of the major impacts include:

* Extreme heat
* Rising seas and increased coastal flooding
* Longer and more damaging wildfire seasons
* An increase in extreme weather events
* Flooding
* Destruction of marine ecosystems
* Severe droughts in some areas

But there are other arguments that are completely against this claim. To find out if global warming is a real phenomenon, it is important to analyze the Earth surface temperature data over considerable time. In this study, I am attempting to find out answers to the following questions along with forecasting temperature for the coming years:

1. If the temperature change over the years is a natural process or an unexpected trend
2. If there is an unusual temperature pattern for one or more countries that can provide us with more insights about human intervention

**Dataset and source:**

The dataset that was downloaded from Kaggle.com consists of extensive Earth surface temperature information from the year 1750 to year 2013.

**Dataset Link:**

<https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data#GlobalLandTemperaturesByCountry.csv>

More information about the data is also available in <http://berkeleyearth.org/data/>

The dataset consists of 5 CSV files – Temperatures by city, temperatures by country, temperatures by major city, temperatures by state and average temperatures of the world.

**Implementation:**

This project is a time series related problem and I would be implementing the techniques learned from Time Series specialization. Since the dataset holds centuries old data, it requires thorough cleaning and validation. After cleaning, the data will be checked for trends and/or seasonality that would better assist us with modeling.

**Challenges:**

I foresee a number of challenges before coming up with a feasible solution. They might include:

* Handling missing data and validating information since the data is available from 1750
* Data is available in multiple levels – city, state and country. Each level can possibly have a different pattern and it might not be a one solution fits all approach.