Climate Change in India: A study on the rising average temperatures and the role of carbon emissions and deforestation.

Shashwat Aggarwal

1004893519

Department of Economics, University of Toronto

ECO225: Data Tools for Economists

Professor Farhoodi

Introduction

In recent times climate change and the climate emergency has become a thing of urgent attention. With the rising average global temperatures, economies, and countries all over the world are headed towards ruin. Additionally, increase in average temperature causes more frequent droughts, floods, heatwaves, and also harms ecosystems along the way (Tol, 2010). This in turn harms the local economies of the countries and livelihoods (Tol, 2010). Rising sea levels also impact coastal areas and people are often forced to relocate (Tol, 2010).

The climate crisis in India is not any better. The climate crisis in India is predicted to have severe impacts on the already depleted water reserves, immense droughts and heatwaves, negative impacts on agriculture, constrained use of sources of energy, and many other problems (NIC, 2009) . All these problems will have huge economic impacts and cause harm to the GDP and the livelihood of the citizens. For example, GDP is anticipated to drop by 2-5% in drought seasons and with intense droughts predicted with the climate crises, we are bound to see a huge fall in the GDP for the country (Dubash et al., 2018).

In the Carbon Brief Profile: India, we see several key highlighters about India's role in the wider global climate change talk (Timperley, 2020). The report describes how India is the world's third largest emitter of greenhouse gases with blame mostly falling on large dependence on coal as source of energy and large methane emissions coming from livestock and rick cultivation. Another concerning thing that the report points out is that even if India meets its pledge to reduce emissions by 33-35% by 2030, India's emissions could still increase by 90% between the years 2014 and 2030. This shows evidence by independent analysis that the climate urgency and crisis must be addressed now and seriously. Intense reforms must be made to ensure survival of the future generations.

The current literature, while does provide information about the climate crises and impacts of global warming, there is very little information providing pure data for future analysis to be conducted. This paper aims to tackle this problem by conducting independent analysis for observing the trends of climate change in India. The goal of this paper is to see if the average temperatures in the country of India are increasing are not and study potential factors for this increase. While all this is being done, one thing we need to keep in mind is that the causes of climate change are not localized and go beyond the local ecosystem. The factors this paper will discuss are carbon emissions (metric tons per capita) and the forest cover in the country.

This paper is divided into four different projects. Project One aimed at understanding and determining if the average temperature in India had increased over time using simple graphs such as bar graphs, line trends, and summary statistics. Project two aimed at answering this question further through dynamic mapping and choropleth (hover) maps. Project three aims at looking other datasets available online and using the tools of web scraping to study the impact of some of the causes of global warming - in particular, I studied deforestation and carbon emissions. Lastly, the final project looks at carbon emissions and loss in forest cover as variables and conducts different regressions to see if these variables are significant in predicting the outcome variable - average temperature in the country. We also see a machine learning regression tree for better understanding of these regression variables.

Body

When it comes to answering the question of whether the average temperatures are increasing or not, the very first thing we must do is find the data for the temperatures for a specific time period and then perform analysis on it to observe the different trends.

Data

The data that we use in this paper comes from many different sources. To begin with, the data containing the temperatures comes from Kaggle from a dataset titled 'Climate Change: Earth Surface Temperature Data' (link - https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data) . The paper uses the raw temperatures from 1796 to 2013 for its analysis. This dataset contains multiple files containing the global land temperatures since 1750 – divided by Cities, States, Countries, Major Cities, and global average land temperatures. The data comes from Berkeley Earth Data page. We use the file that has the data divided by States so that we can conduct our analysis on a State level as well where needed. This data contains about 81620 datapoints collected daily from 1796 to 2013. This is then cleaned and formatted for our use to be averaged annually.

Secondly, the dataset about the deforestation trends in India comes from Mongabay (https://rainforests.mongabay.com/deforestation/archive/India.htm). Because this data is from the years 2001 to 2018, we use the intersection of this data and the temperature data from Kaggle, i.e., 2001 to 2013. We get 17 data points on which we can conduct our analyses.

The third dataset that we use is the metric tons per capita carbon emissions in India. This dataset comes from the Millennium Development Goals Indicators website of the United Nations (http://mdgs.un.org/unsd/mdg/Data.aspx). This dataset is from 1990 – 2011 so we use it with the

Kaggle temperatures data for those specific years for conducting our analysis. We get 22 datapoints that are used for conducting analyses.

Summary Statistics

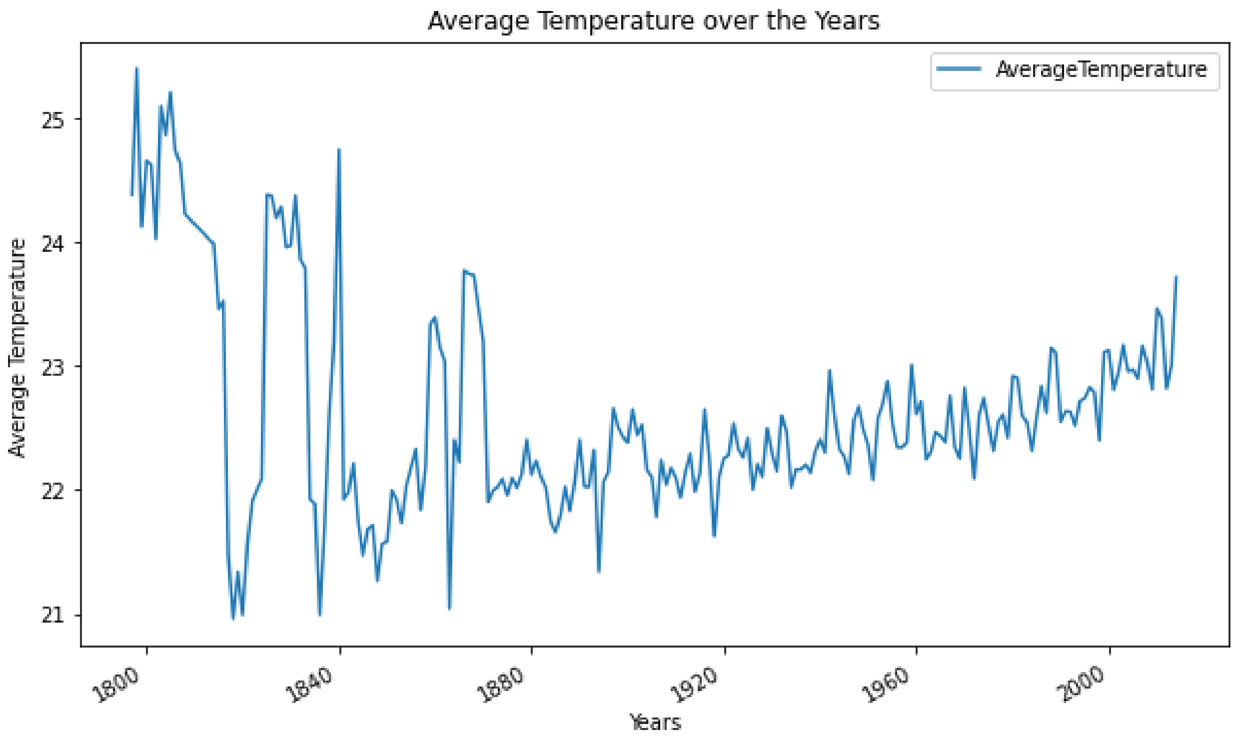
From the data we collected as described above, we first calculate the summary statistics and then perform different analysis on it. This data along with their explanations is given below.

Summary statistics of Average Temperature

Count	81620
Mean	22.534576
Standard Deviation	7.749321
Min	-11.984000
25%	19.063000
50%	25.094500
75%	27.628000
Max	36.339000

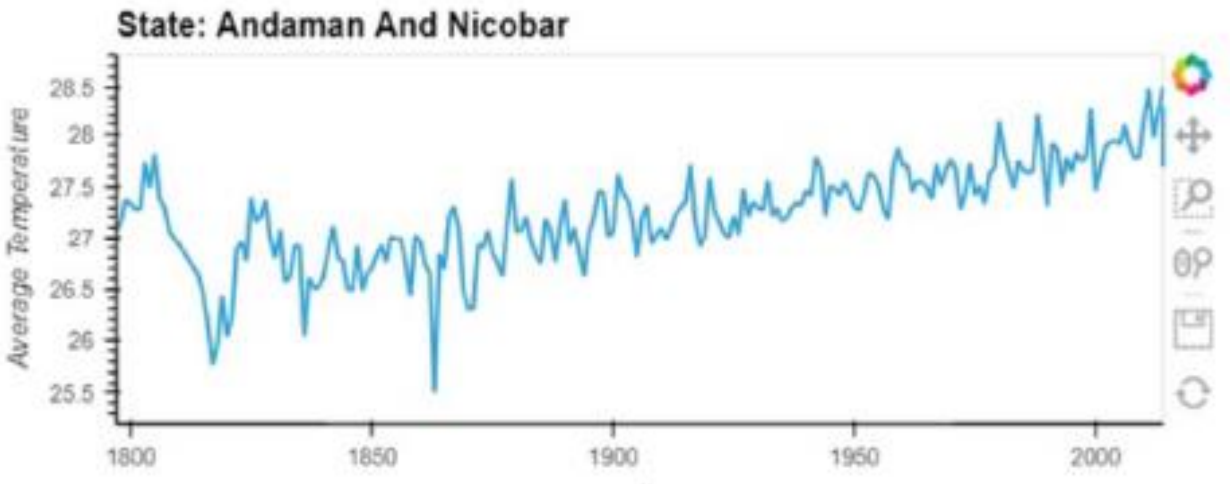
The above data shows us that there were 81620 observations in total for the average temperature data collected from Kaggle after it filtered for including only data about States in India. The mean average temperature is seen to be 22.53°C with the standard deviation being 7.749. The minimum temperature ever recorded was -11.984 and the maximum was 36.339.

Average Temperature from 1796 to 2013



The above graphs show us the trend in average temperatures from the years 1796 to 2013. The large variability pre 1900s could be because of the lack of good tools to record temperatures.

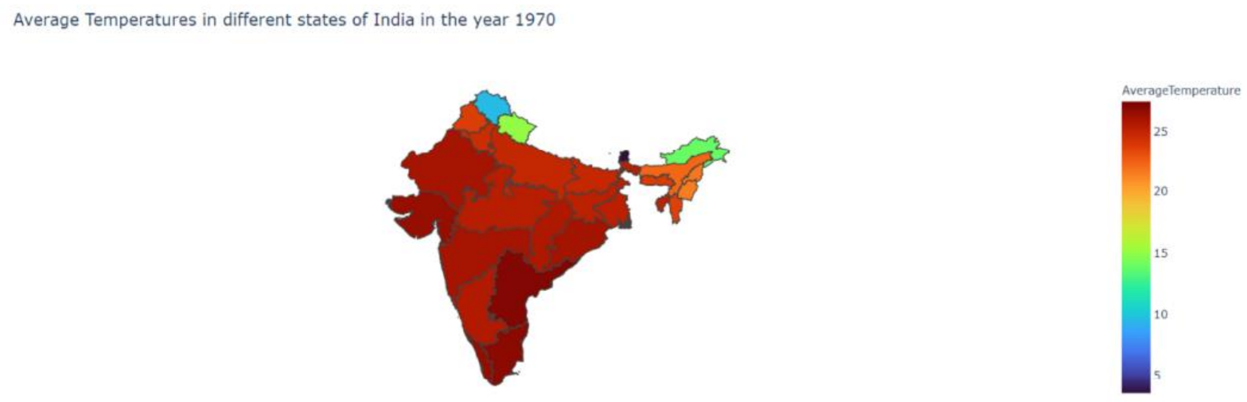
State wise average temperatures from 1800 - 2013

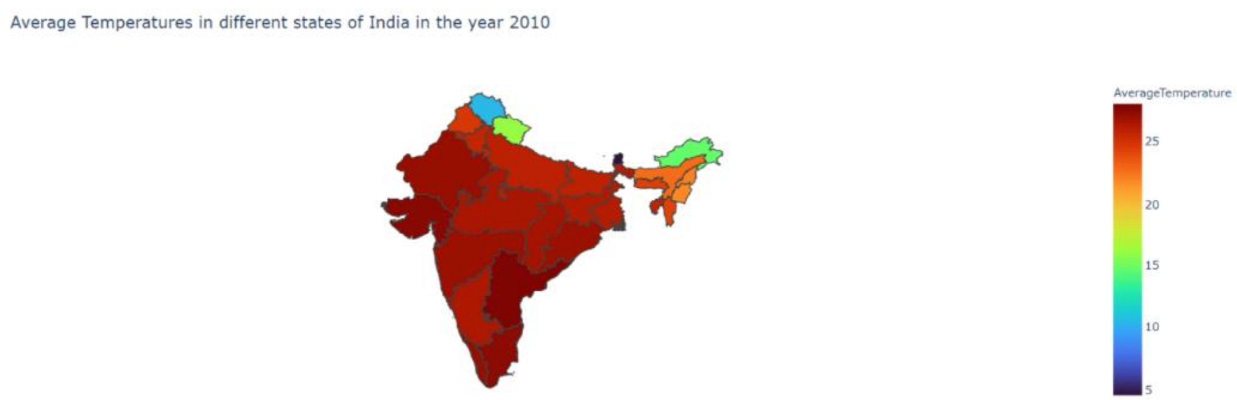


The above graph is an example of a further analysis our project conducted. We were able to derive individual graphs for each state showing their average temperature trends from 1800 to 2013. The above graph is for the state Andaman and Nicobar in particular. Due to the country having 29 states, the graphs are omitted from the paper but can be found in the final project. One key observation that was made was that all the states showed a small, but stead rise in the average temperatures since the early 1900s. This steady rise in average temperatures corresponds to the increasing global temperatures since the industrial revolution timeline. We see climate getting warmer every year in each state.

Heat Maps comparing 1976 and 2013.

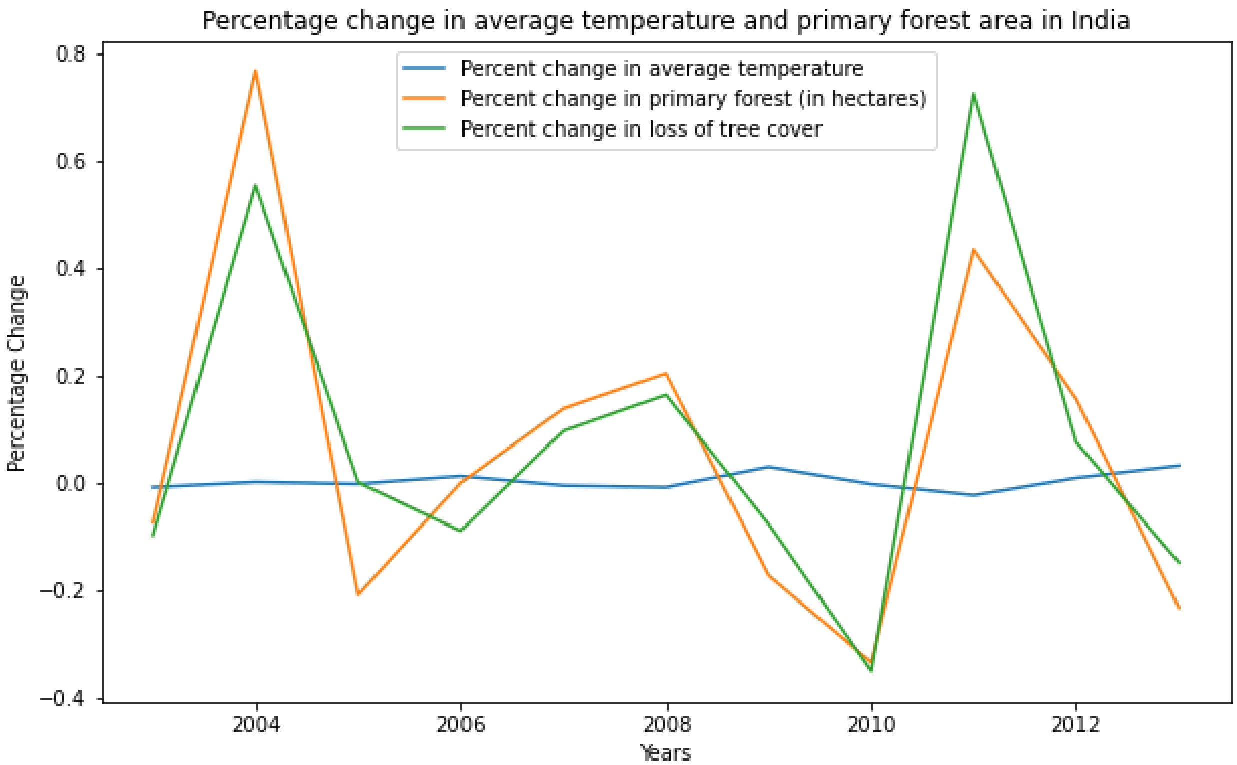
Two heat maps were created to see if there was any difference observed in the average temperatures in each state from 1970 to 2010. These maps are displayed below.





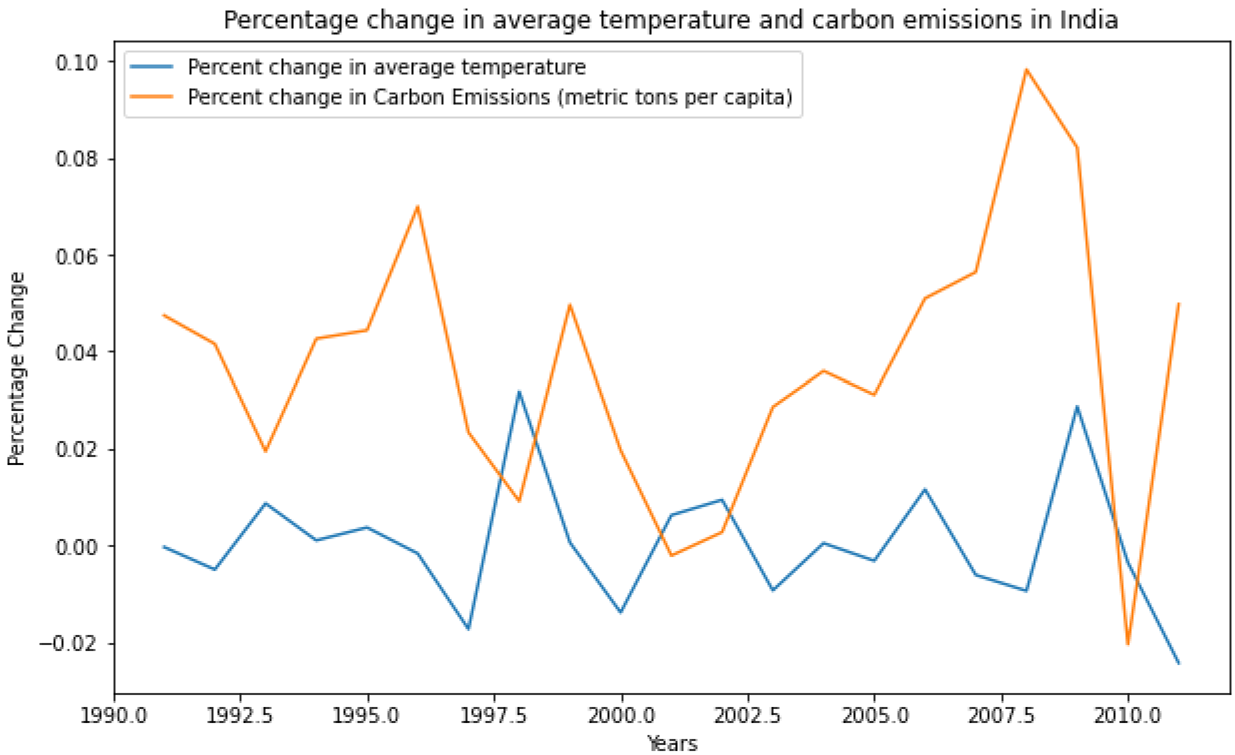
While the observation and the colors may not be very obvious, when observed carefully we can see slightly darker shades in the map for 2010 indicating the average temperatures have indeed increased. While this increase is not drastic, it is still very much present. This is alarming cause a small increase in average temperatures has been very harmful for a lot of economies around the world, including India’s. If this gradual increase takes place every year, we could soon see a major human survivability crisis.

Percent change in average temperatures and primary forest areas in India



The above analysis was conducted to see if the change in forest cover and the change in tree cover was associated with the change in average temperatures. What we observed that whenever the primary forest cover declined, the average temperature rose in the subsequent period. This might point towards the longer-term impact of deforestation as opposed to immediate effects.

Percent change in average temperature and carbon emissions in India



A similar analysis was conducted for observing the percent change in average temperature and the percent change in carbon emissions (metric tons per capita). A similar pattern to the deforestation graphs was observed. Whenever the carbon emissions were high in one period, the percent increase in average temperatures was positive in the next. This indicates that the carbon takes time to accumulate in the atmosphere and has effects in the longer run. Additionally, we need to be careful of the fact that percent change only shows the amount of change of carbon emissions and not how much already exists. A negative percent change still indicates carbon being emitted – only smaller in amount as compared to the previous period.

Results

Lastly, we conducted four different regressions analyses to find if the different factors were, individually and together, a significant explanatory variable for the average temperatures in the country. The results showed that only carbon emissions, when compared individually with average temperatures, were a significant predictor. The results for all four regressions is shown below.

For the first regression - Average Temperature and Carbon Emissions: we considered carbon emissions and average temperatures in the country. The model used was:

$$avgtempi=\beta 0+\beta 1carboni+ui$$

- $\beta 0$ is the intercept of the linear trend line on the y-axis.
- $\beta 1$ is the slope of the linear trend line, representing the *marginal effect* of carbon emissions on the average temperatures.
- ui is a random error term (deviations of observations from the linear trend due to factors not included in the model)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     AverageTemperature   R-squared:                       0.395
Model:                            OLS   Adj. R-squared:                  0.365
Method:                 Least Squares   F-statistic:                     13.05
Date:                Thu, 22 Apr 2021   Prob (F-statistic):            0.00174
Time:                        05:32:41   Log-Likelihood:                 4.1434
No. Observations:                  22   AIC:                            -4.287
Df Residuals:                      20   BIC:                            -2.105
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================================================
                                               coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------------------------------
const                                       22.1812      0.206    107.679      0.000      21.751      22.611
Carbon Emissions (metric tons per capita)    0.6198      0.172      3.612      0.002       0.262       0.978
==============================================================================
Omnibus:                        1.697   Durbin-Watson:                   2.106
Prob(Omnibus):                  0.428   Jarque-Bera (JB):                1.186
Skew:                          -0.560   Prob(JB):                        0.553
Kurtosis:                       2.800   Cond. No.                         9.24
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

To assess the performance of the regression, we can look at the coefficient, t-value, the p-value, and the confidence interval of the slope or the X variable's coefficient, and the R-squared value. In our regression we see that the above values for $\beta 1$ or the coefficient for carbon emissions as a predictor for average temperatures in the country are given below -

- coefficient = 0.6198
- t-value = 3.612
- p-value = 0.002
- 95% confidence interval = [0.262, 0.978]
- R-squared = 0.395

With a p-value of 0.002, we see it is significant at the 5% and the 1% level. This means we have very strong evidence that carbon emissions correlate well with the average temperatures in the country and are a good predictor of the Y variable. Since the coefficient is significant, we can interpret it as a unit increase in metric tons per capita of carbon emissions was associated with increase in average temperatures in the country by 0.6198°C. Additionally, we are 95% confidence that our average value falls between 0.262 and 0.978.

R-squared explains the amount of variability explained by the model and therefore, higher the R-squared value means a better model. In our case, R-squared is only 0.392 which is not that high meaning only 39.2% of the variation in our model is explained by the independent variable.

For the second regression - Average Temperature and loss in forest cover: forest cover and average temperatures in the country.

$$avgtempi=\beta 0+\beta 1 forestlossi+ui$$

where:

- $\beta 0$ is the intercept of the linear trend line on the y-axis.
- $\beta 1$ is the slope of the linear trend line, representing the *marginal effect* of loss of forest cover on average temperatures.
- ui is a random error term (deviations of observations from the linear trend due to factors not included in the model)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     AverageTemperature   R-squared:                       0.016
Model:                            OLS   Adj. R-squared:                 -0.073
Method:                 Least Squares   F-statistic:                    0.1821
Date:                Thu, 22 Apr 2021   Prob (F-statistic):              0.678
Time:                        05:32:41   Log-Likelihood:               -0.87105
No. Observations:                  13   AIC:                             5.742
Df Residuals:                      11   BIC:                             6.872
Df Model:                           1
Covariance Type:            nonrobust
======================================================================================
                         coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------
const                 23.2732      0.401     58.003      0.000      22.390      24.156
loss of tree cover -2.336e-06   5.47e-06     -0.427      0.678   -1.44e-05    9.71e-06
==============================================================================
Omnibus:                        6.057   Durbin-Watson:                   1.581
Prob(Omnibus):                  0.048   Jarque-Bera (JB):                3.229
Skew:                           1.200   Prob(JB):                        0.199
Kurtosis:                       3.451   Cond. No.                     3.77e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 3.77e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```

The above regression results shows that our p-value is not significant as well as R-squared is not high therefore this is not a good predictor of average temperatures.

For the third regression - Average Temperature and primary forest cover in hectares:

primary forest cover (in hectares) and average temperatures in the country.

$$avgtempi=\beta 0+\beta 1 forestcoveri+ui$$

where:

- $\beta 0$ is the intercept of the linear trend line on the y-axis.
- $\beta 1$ is the slope of the linear trend line, representing the *marginal effect* of forest cover on average temperatures.
- ui is a random error term (deviations of observations from the linear trend due to factors not included in the model).

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     AverageTemperature   R-squared:                       0.127
Model:                            OLS   Adj. R-squared:                  0.040
Method:                 Least Squares   F-statistic:                     1.456
Date:                Thu, 22 Apr 2021   Prob (F-statistic):              0.255
Time:                        05:33:44   Log-Likelihood:               -0.39763
No. Observations:                  12   AIC:                             4.795
Df Residuals:                      10   BIC:                             5.765
Df Model:                           1
Covariance Type:            nonrobust
==================================================================================================
                                  coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------------------
const                          23.6073      0.413     57.119      0.000      22.686      24.528
primary forest (in hectares) -3.126e-05   2.59e-05    -1.207      0.255    -8.9e-05    2.65e-05
==============================================================================
Omnibus:                        3.634   Durbin-Watson:                   1.588
Prob(Omnibus):                  0.163   Jarque-Bera (JB):                1.773
Skew:                           0.941   Prob(JB):                        0.412
Kurtosis:                       3.077   Cond. No.                     8.33e+04
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 8.33e+04. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Similar to the second regression, the above regression results shows that our p-value is not significant as well as R-squared is not high therefore primary forest cover is not a good predictor of average temperatures.

For the fourth regression - Average Temperature along with carbon emissions, loss in tree cover, and primary forest cover in hectares:

$$avgtempi=\beta0+\beta1carboni+\beta2forestlossi+\beta3forestcoveri+ui$$

- $\beta0$ is the intercept of the linear trend line on the y-axis.
- $\beta1$ is the slope of the linear trend line, representing the *marginal effect* of carbon emissions on average temperatures.
- $\beta2$ is the slope of the linear trend line, representing the *marginal effect* of loss of forest cover on average temperatures.
- $\beta3$ is the slope of the linear trend line, representing the *marginal effect* of forest cover on average temperatures.

- ui is a random error term (deviations of observations from the linear trend due to factors not included in the model)

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     AverageTemperature   R-squared:                       0.448
Model:                            OLS   Adj. R-squared:                  0.173
Method:                 Least Squares   F-statistic:                     1.626
Date:                Thu, 22 Apr 2021   Prob (F-statistic):              0.280
Time:                        05:33:47   Log-Likelihood:                 4.1826
No. Observations:                  10   AIC:                           -0.3653
Df Residuals:                       6   BIC:                            0.8451
Df Model:                           3
Covariance Type:            nonrobust
===============================================================================================================
                                              coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------------------------------------
const                                      22.8278      0.516     44.254      0.000      21.566      24.090
Carbon Emissions (metric tons per capita)   0.7306      0.413      1.769      0.127      -0.280       1.741
loss of tree cover                      -1.736e-05    1.2e-05     -1.449      0.198   -4.67e-05     1.2e-05
primary forest (in hectares)             2.768e-05   4.64e-05      0.596      0.573    -8.6e-05       0.000
==============================================================================
Omnibus:                        1.172   Durbin-Watson:                   2.627
Prob(Omnibus):                  0.556   Jarque-Bera (JB):                0.494
Skew:                           0.526   Prob(JB):                        0.781
Kurtosis:                       2.722   Cond. No.                     6.93e+05
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.93e+05. This might indicate that there are
strong multicollinearity or other numerical problems.
```

Similar to the third regression, the above regression results shows that our p-values are not significant as well as R-squared is not high. Therefore, when taken together none of the factors are significant enough to make an impact on the average temperatures.

Discussion and Conclusion

The aim of this paper was to determine if the average temperature in the country of India have been increasing or not and understand the reality of the climate crisis better. We conducted many different types of analysis in this paper and saw many different types of results – some which help us achieve this goal and some do not. One of the main, highlighting, results we found in our analysis is that the average temperatures in each individual state have been increasing since 1800. This highlights that the climate change is, in fact real, and needs to be addressed. When it comes to studying the different factors behind climate change, it is known that carbon emissions and deforestation are two such responsible factors. Therefore, we decided to conduct four regressions to see which factors were significant predictors of average temperatures. We discovered that only carbon emission was a significant predictor of average temperatures, that too only when it was regressed in a bivariate model. It was not a significant predictor in the multivariate regression that was conducted with carbon emissions, loss in forest cover, and primary forest cover, taken together. The potential reason for that could be the lack of data points available. With only about 11 datapoints available to conduct the multivariate regression, the sample size is too small to make accurate predictions and therefore we have a poor model. The lack of data points, i.e., a small sample size stands to be the reason why R-squared is small for all the regressions conducted.

In the future more research should be conducted with these factors but taking into consideration more datapoints. This requires more datasets to be available for public use. Additionally, we are only considering two potential factors for the average temperature trends in India, but it is known that many other factors such as population, pollution, use of fossil fuels, increase in agriculture etc. are also responsible for causing climate change. For a more detailed, better regression analyses, future studies should consider these factors as well.

References

Cho, R. (2019, June 20). *How Climate Change Impacts the Economy*. State of the Planet. https://news.climate.columbia.edu/2019/06/20/climate-change-economy-impacts/.

Dubash, N. K., Khosla, R., Kelkar, U., & Lele, S. (2018). India and Climate Change: Evolving Ideas and Increasing Policy Engagement. *Annual Review of Environment and Resources*, *43*(1), 395–424. https://doi.org/10.1146/annurev-environ-102017-025809

India: Development while fighting climate change. Environmental Defense Fund. (n.d.). https://www.edf.org/climate/india-development-while-fighting-climate-change.

National Intelligence Council. (2009). (rep.). *India: The Impact of Climate Change to 2030 A Commissioned Research Report*. Retrieved from https://www.dni.gov/files/documents/climate2030_india.pdf

Timperley, J. (2020, April 6). *The Carbon Brief Profile: India*. Carbon Brief. https://www.carbonbrief.org/the-carbon-brief-profile-india.

Tol, R. S. (2010). The Economic Impact of Climate Change. *Perspektiven Der Wirtschaftspolitik*, *11*, 13–37. https://doi.org/10.1111/j.1468-2516.2010.00326.x