Climate Change

Group E5- Sage Myung, Brian Lukas, Emeka Emeche, Sweena Kalamkar, Vanshika Goel

Section 1

Our group wanted to choose the topic of climate changes and the ways it has been a growing problem over the years. We will be gathering and comparing two sets of data from rstudio package “dslabs” and the World Bank websites. We wanted to use the dslabs because it gave us insight of the temperature changes and the carbon emission which helps to understand the raise of climate changes. The second source allowed us to view different factors such as greenhouse gas emission that further provides information about human factors with climate change. In order to better understands changes in our climate, we need to further be able to look at factors that contribute to this. Climate change is something our society needs to look and understand as it has huge factors of our future. Even today, changes in our world have negatively impacted thousands of people and impacting their daily life.

Section 2

The data used in the following assignment talks about the values concerning the damage caused by the emission of greenhouse gases such as carbon dioxide and more. The two datasets used primarily represent the climate data and how the earth is getting impacted by industrial activities.

WORLD BANK - https://data.worldbank.org/indicator?tab=all (https://data.worldbank.org/indicator?tab=all)

The World Bank comprises 189 member countries or shareholders who are represented by a board of governors, who are the ultimate policymakers at the World Bank. Generally, the governors are member countries’ ministers of finance or ministers of development. At the World Bank, the Development Data Group coordinates statistical and data work and maintains a number of macro, financial and sector databases. Working closely with the Bank’s regions and Global Practices, the group is guided by professional standards in the collection, compilation and dissemination of data to ensure that all data users can have confidence in the quality and integrity of the data produced. Much of the data comes from the statistical systems of member countries, and the quality of global data depends on how well these national systems perform. The World Bank works to help developing countries improve the capacity, efficiency and effectiveness of national statistical systems. Without better and more comprehensive national data, it is impossible to develop effective policies, monitor the implementation of poverty reduction strategies, or monitor progress towards global goals.

dslabs - https://cran.r-project.org/web/packages/dslabs/dslabs.pdf (https://cran.r-project.org/web/packages/dslabs/dslabs.pdf)

The package “dslabs” is a data set from the Rstudio’s source from Office of Scientific and Technical Information of the U.S. Department of Energy, that we accessed using library. This shows from years 1781 to 2018 on annual mean global temperature anomaly on land, sea and together. It also including the annual carbon emissions in millions of metric tons of carbon and gathered by the National Centers for Environmental Information, which was gathered by the National Centers for Environmental Information.

Question 1: How does the recent output of the three major greenhouse gasses in the US compare to the historical change in greenhouse gasses worldwide?

To answer this question we used the data from the World Bank regarding emissions of the three major greenhouse gases (CO2, Methane, and Nitrous Oxide). We decided to stratify the data down to just the United States to make the results of the data more immediate to our audience, students living in the United States. We also used a dslabs data frame provided in R that contains data regarding historical emissions of these gases.

Data Wrangling

We used the dslabs and World Bank data frames to make line graphs illustrating the change in these emissions over time, however, the difference in units (kt vs. ppb/ppm) required us to create two separate graphs of the data. In each graph we included a geom_smooth line to illustrate the trend of all three gases among the data. To make the data more digestible for the graphs, we gathered just the data regarding the output of the gasses to be mapped to the points on the graph.

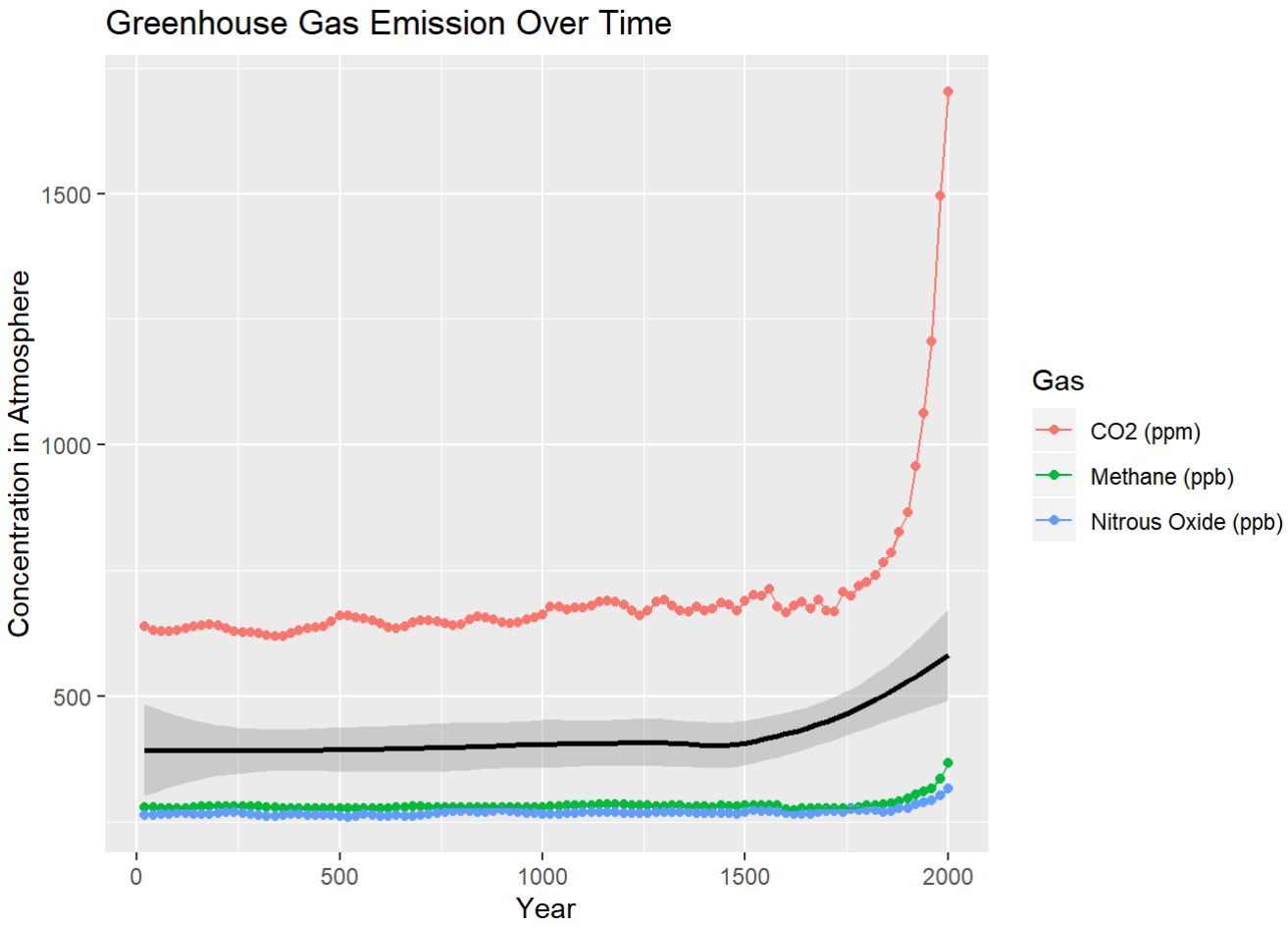
Descriptive Statistics

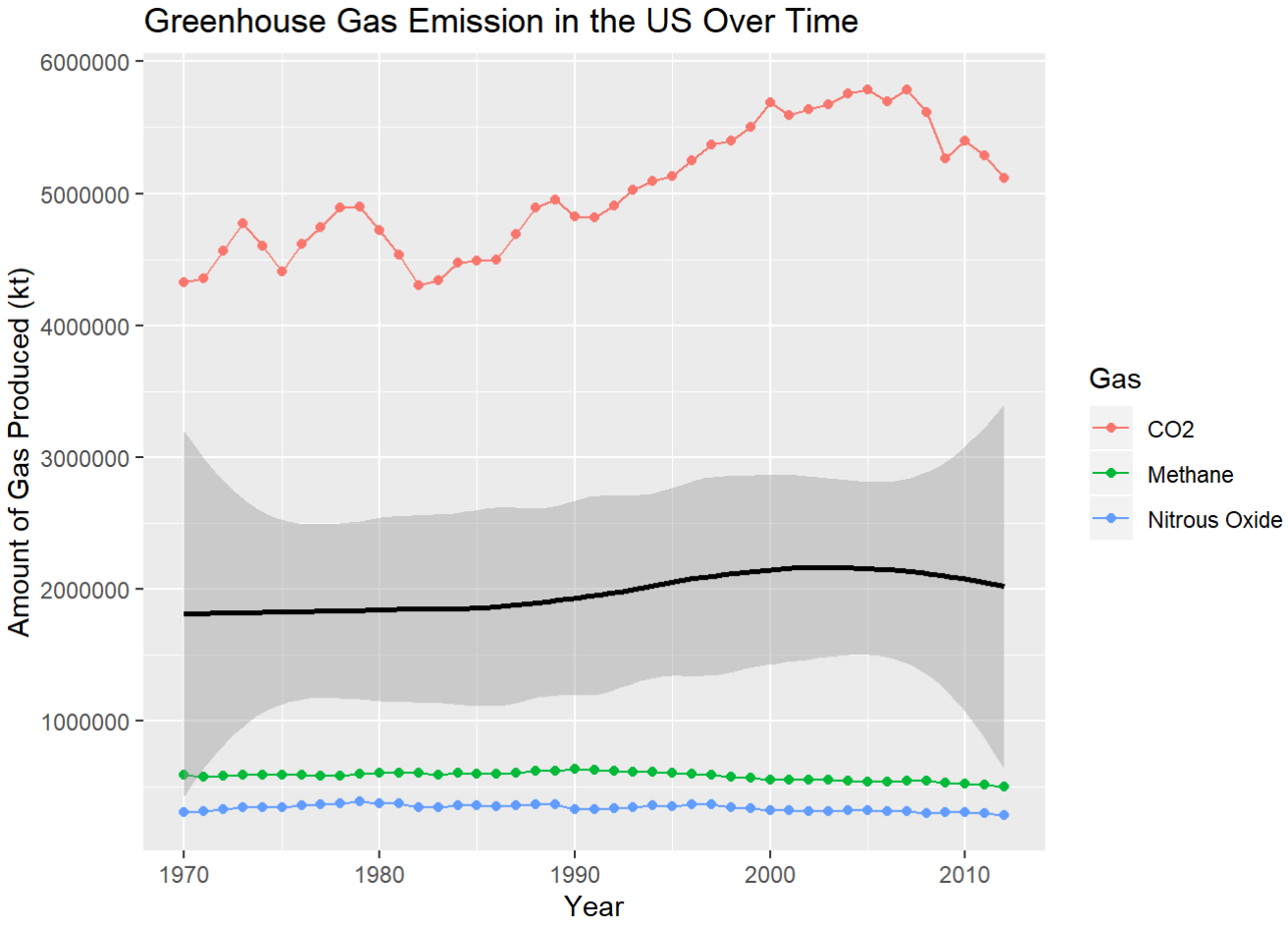
Statistics	Greenhouse Gas Emission in ppb (dslabs)	Greenhouse Gas Emission in kt (World Bank US Data)
Mean	416.2	1979434

Aside from the trends of these data shown on the following line graphs, these means are nearly useless to compare as they are measured in different units and the dslabs data relies heavily on speculation of past conditions of the Earth's atmosphere.

Outliers

The dslabs greenhouse gases data and World Bank greenhouse gas data for the US feature no outliers, most likely because the data is illustrating climate trends, and climate is already measured on a large scale (years). Any significant outliers have most likely been quelled by previous data analysis by dslabs and the World Bank.


Greenhouse Gas Emission Over Time
Concentration in Atmosphere
1500
1000
500
0
500
1000
1500
2000
Year
Gas
CO2 (ppm)
Methane (ppb)
Nitrous Oxide (ppb)


Greenhouse Gas Emission in the US Over Time
Amount of Gas Produced (kt)
6000000
5000000
4000000
3000000
2000000
1000000
1970
1980
1990
2000
2010
Year
Gas
CO2
Methane
Nitrous Oxide

Visualization

The geom_smooth line on the Greenhouse Gas Emission Over Time graph exhibits a strong upward slope between the years 1750-2000, similarly, the Greenhouse Gas Emission Over Time in the US illustrates a similar trend between the years 1990-2010. This provides evidence of a trend that requires a large scale view to really see, while the US specific graph shows a weaker slope than the world history graph, the enlightened reader will observe that these trends when compiled among many countries will provide the dramatic results shown on the world graph. Thus we have the conclusion that the climate crisis will require a cumulative effort on a world scale to combat, and no one is less responsible than anyone else.

Question 2: Whats the impact of carbon dioxide emission on the economic growth over the past 20 years?

In order to understand more about climate change, we need to understand the impact of one of the most common greenhouse gases (CO2) on the economic growth of all the nations worldwide. In today's age, the main topic of interest is the economic growth and the climate often tends to impact the economy significantly and hence, it is essential to study this consequence of climate change.

Data Wrangling

First, the data of interest was selected (CO2 damage and including particulate damage on our economy) for all countries over the past 20 years and assigned these values to a data frame called economic_damage. The date type was then managed by using as.numeric(date). The data was then organized by using the gather function and put into long format. The columns and raw data were then renamed in order to make more sense of the data on the graphs.

Statistics

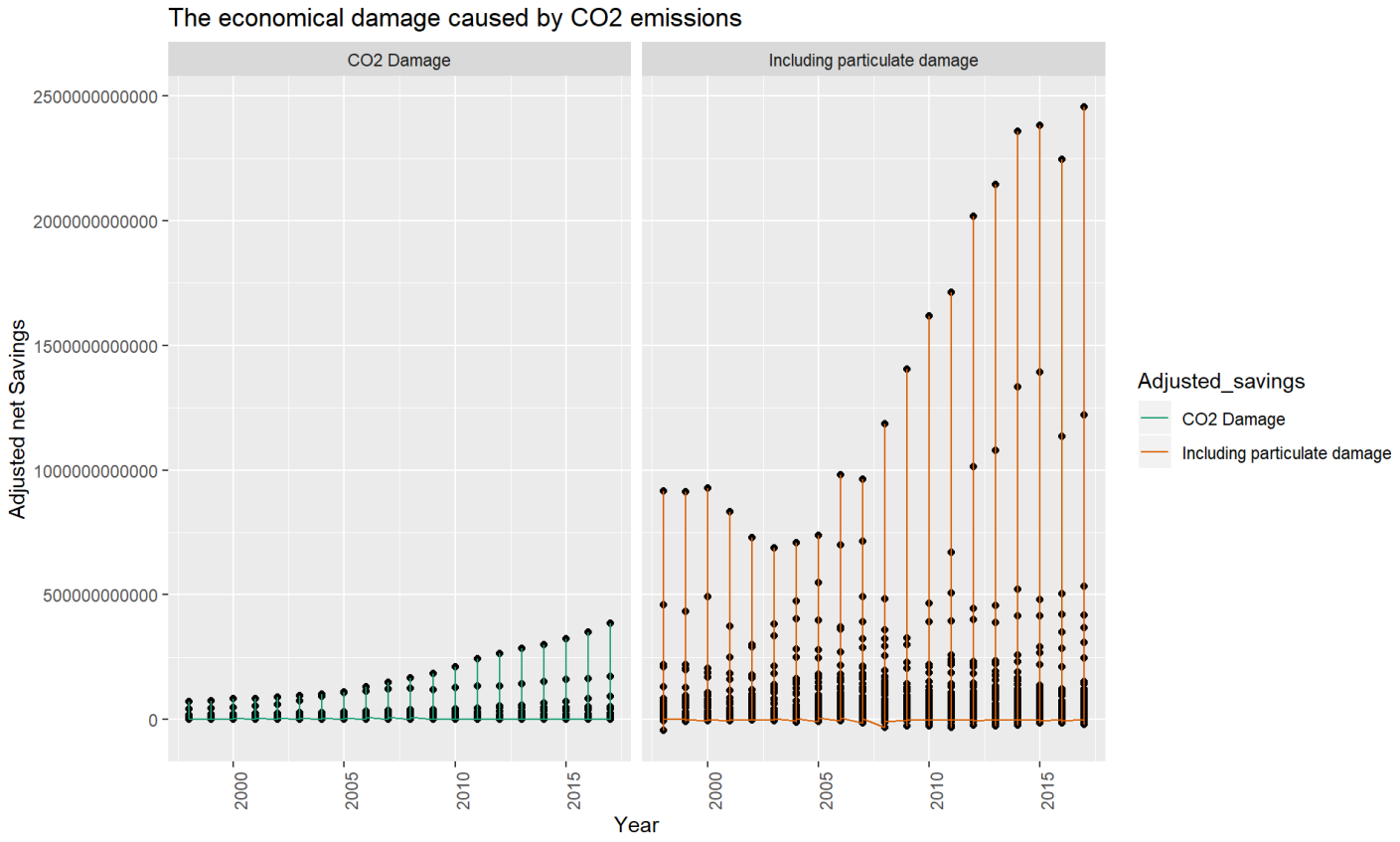
Statistics	CO2 Damage	Including particulate damage
Mean	4703041305	40863431705
Outliers	384735482377	2455370116969

Descriptive Statistics

The summary of the following raw data reflects the mean of the total CO2 emissions to be 4703041305 while the mean of adjusted net savings including particulate emission damage to be 40863431705. This shows that on average the CO2 emissions damage is lower as compared particulate damage even if its including particulate emission damage. This is the same case with regard to the maximum values of the CO2 emissions and the adjusted net savings including particulate emission damage. The CO2 emissions data did not include any missing values however, the other column did so we used na.omit in order to remove all the NA values so that the descriptive statistics are not altered greatly by the missing values.

Outliers

The outliers package allows to figure out the outlying values which might skew with the mean or the statistical analysis. The following section utilizes the outlier function to find the values which is the furthest away from the mean value. These are the values that might obscure the trend of the following data. As calculated, the value of for China in 2017 had the highest values with regards to CO2 emissions and adjusted net savings including as compared to the mean.



Visualization

As shown by the visualization, the Carbon dioxide emissions for the past 20 years and the economic damage caused by it have been on a steeping increasing trend. Similarly, we can also observe the increasing trend of the particulate damage caused to our economy including CO2 and particulate damage. This shows the interdependence and contribution of the CO2 damage on our economy. The increase of CO2 levels in the past 20 years is truly alarming and now, observing its impact on our economy should set us to be more precautious about our environment so that our future generations are also able to experience the world we live in.

Question 3: What is the impact of carbon dioxide emissions on the land and ocean abnormalities?

To study climate change, we also need to understand the direct impact of CO2 emissions on land and ocean anomalies as this is the deciding factor for weather and natural disasters. These factors account for the major changes in day-to-day human activities hence, climate change can be accounted for by better understanding CO2 emissions and its impact on our Earth.

Data Wrangling

The required data needed for this section (CO2 emissions, land and ocean anomaly data) was selected from the temp_carbon section and assigned to a new data frame called emissions_anomaly. Then the data was organized by using the gather function and put into long format.

Statistics

Statistics	Carbon emissions	Land anomaly	ocean anomaly
Mean	2942	0.03259	0.03252

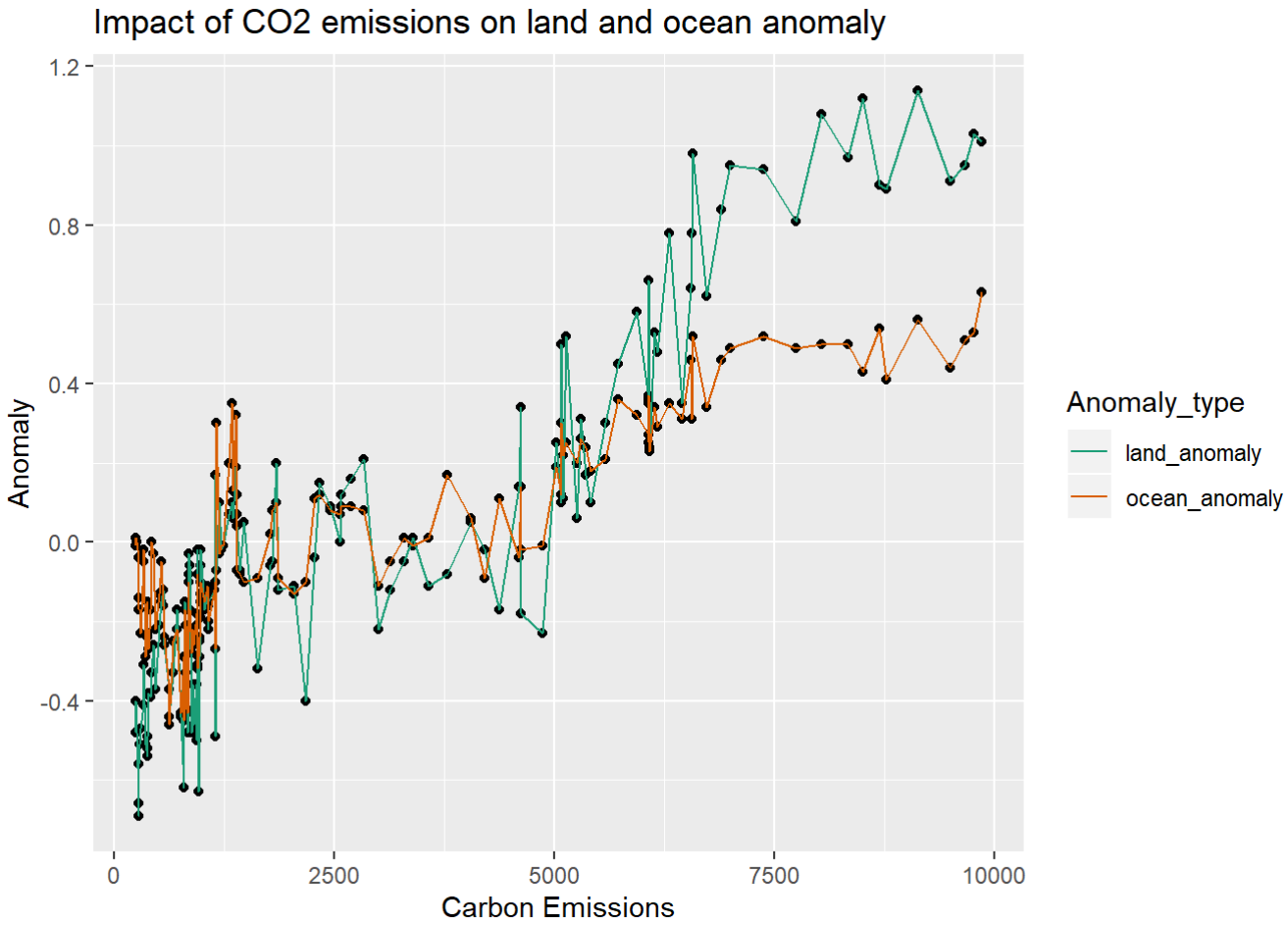
Statistics	Carbon emissions	Land anomaly	ocean anomaly
Outliers	9855	1.14	0.63

Descriptive Statistics

The descriptive statistics of the following raw data highlights the average CO2 emissions to be 2942 while the average effect upon land and ocean anomaly appears to be 0.03259 and 0.03252 respectively. This shows the mean correlation between the emissions and the anomalies that have been caused. The main columns of interest in the following section do not have any missing values that might derail the overall mean or statistical analysis of the data.

Outliers

The outliers package allows to figure out the outlying values which might skew with the mean or the statistical analysis. The following section utilizes the outlier function to find the values which is the furthest away from the mean value. These are the values that might obscure the trend of the following data. As calculated, the value of 9855 is the greatest which is very high as compared to the mean of 2942. This is similarly observed in the land and ocean anomalies which have outliers of 1.14 and 0.63 respectively.



Visualization

As observed through the visualization, the CO2 emissions have been on a constant rise for the past few years and this has had a significant impact upon the temperatures, land and oceans of our earth. As specified by the line and point graph, we can conclude the trends that the land and ocean abnormalities follow as caused by the

increase in CO2 emissions. Overall, the land anomalies are higher as compared to the ocean anomalies as caused by the CO2 emissions.

Question 4: What are the top 12 countries that were affected and how was agriculture output affected in 2009?

For climate change, we also have to understand how the change in envirnment can affect people's lives. Drastic changes in the environment can greatly affect the populations residing in how they live their life. We also see how climate change can affect the agricultural production in how much countries output, effectively changing the value of the products shipped out in the world.

Data Wrangling

The data required here were: "Droughts, floods, extreme temperatures (% of population, average 1990-2009)" and "Agriculture, forestry, and fishing, value added (% of GDP)." The data was wrangled from the two ID codes given from the dataset and turned into 'pop_percent_affected' and 'gdp_gross_output.' Then the data was filtered to show the year of countries in 2009 and used the gather function to turn the two columns into long format. After, the slice function was used to filter out which data points had a full set of values to analyze since there were NA values included within the chart.

Statistics

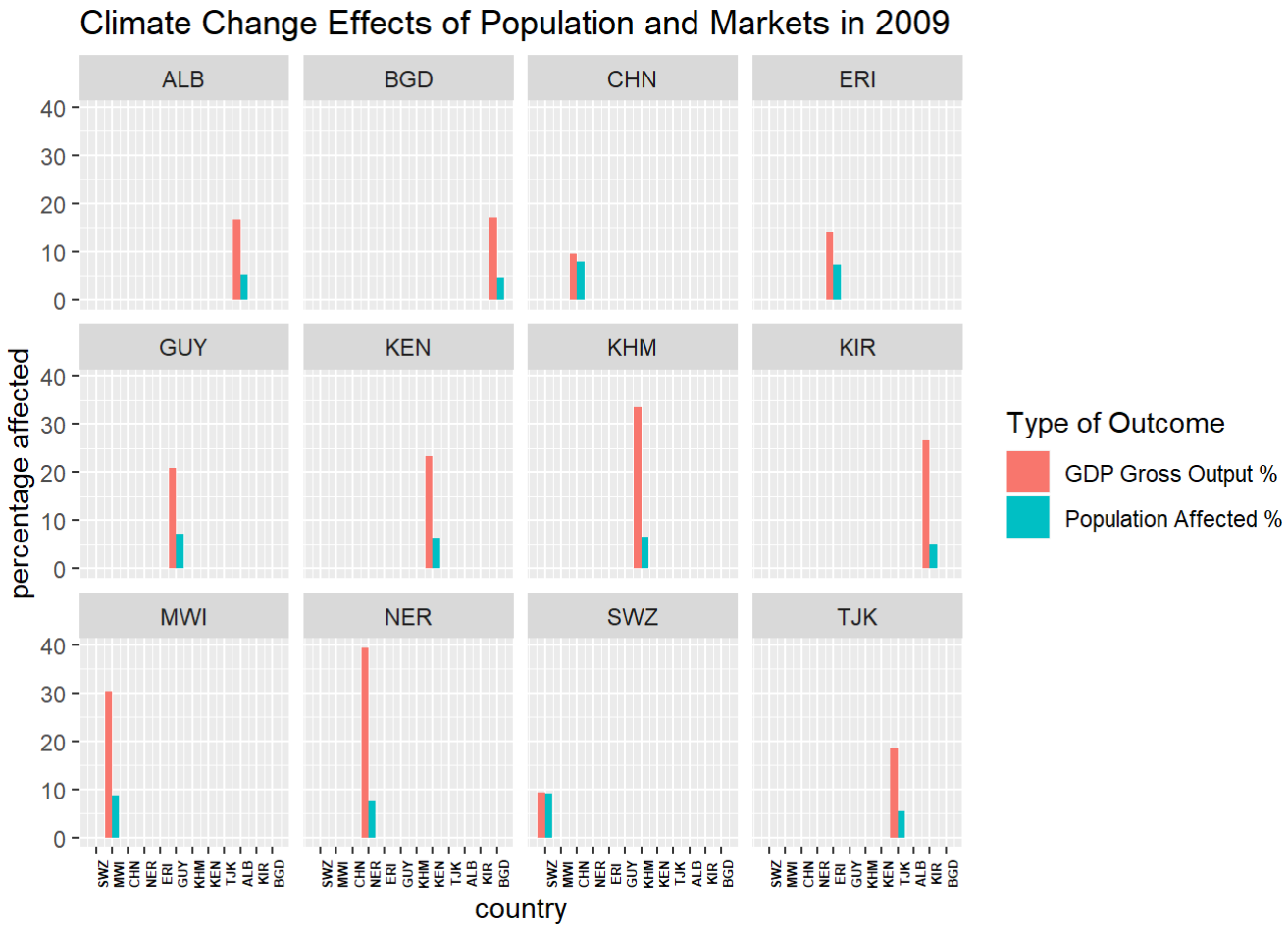
Statistics	Country Percentages	Total Population Affected	Total GDP Affected
Mean	14.207	1.16990	11.28260

Descriptive Statistcs

The statistics shown here show 3 different aspects of the data presented: the wrangled dataset showing the top 12 country percentages, as well as the data from the total population affected and total GDP affected. The mean from the data here is 14.207, 1.16990, and 11.28260. The country percentage mean represents the value as a whole showing a low turnout of output when different populations were affected in the top 12 countries. The mean of total population affected in the whole world is low which shows that world's populations aren't influenced as much, possibly due to adapting along with the environemnt.The mean of the total GDP is low which can show how climate change effect products due to varying places every place is in a different environment which can induce differing values of products everywhere.

Outliers

The datasets didn't provide any outliers listed. It mainly showed the percentages each country had regarding this topic. There weren't too many factors that could have drastically changed the data to skew it. Each country is fundamentally different due to living conditions and resources they have, so it's expected that most countries can be higher or lower than others in development.



Vizualization

The graph here shows the top 12 countries which were afflicted the most in percentage. Depending how badly each country was hit with climate change, it showed a correlation on how much GDP output they pushed out. From what is shown on the data, the higher the population affected usually means there is a low expectancy for GDP gross output. Eswatini had the lowest GDP value of around 9 percent with around 9 of the population affected. In Niger however, even though around 7.5 percent of the population were overcome with climate change, they had the highest GDP gross value of around 39 percent. This goes into question of what factors can cause the GDP gross value rate to grow or diminish. Niger for example could have a strong labor force, due to a massive influx of people who live there. Though the population may be affected, it wouldn't effect the GDP output too much. A low labor force, especially after effects of climate can make countries output very little products. The state of the country itself also plays a role too. Natural disasters can alter how much agricultural resources the country produces and the amount of output sent out.

Question 5: Which countries have had the highest contribution of greenhouse gas emissions from 1990 to 2018?

To study climate change, we also need to understand the direct impact of CO2 emissions on land and ocean anomalies as this is the deciding factor for weather and natural disasters. These factors account for the major changes in day-to-day human activities hence, climate change can be accounted for by better understanding CO2 emissions and its impact on our Earth.

Data Wrangling

The required data needed for this section (CO2 emissions, land and ocean anomaly data) was selected from the temp_carbon section and assigned to a new data frame called emissions_anomaly. Then the data was organized by using the gather function and put into long format.

Statistics

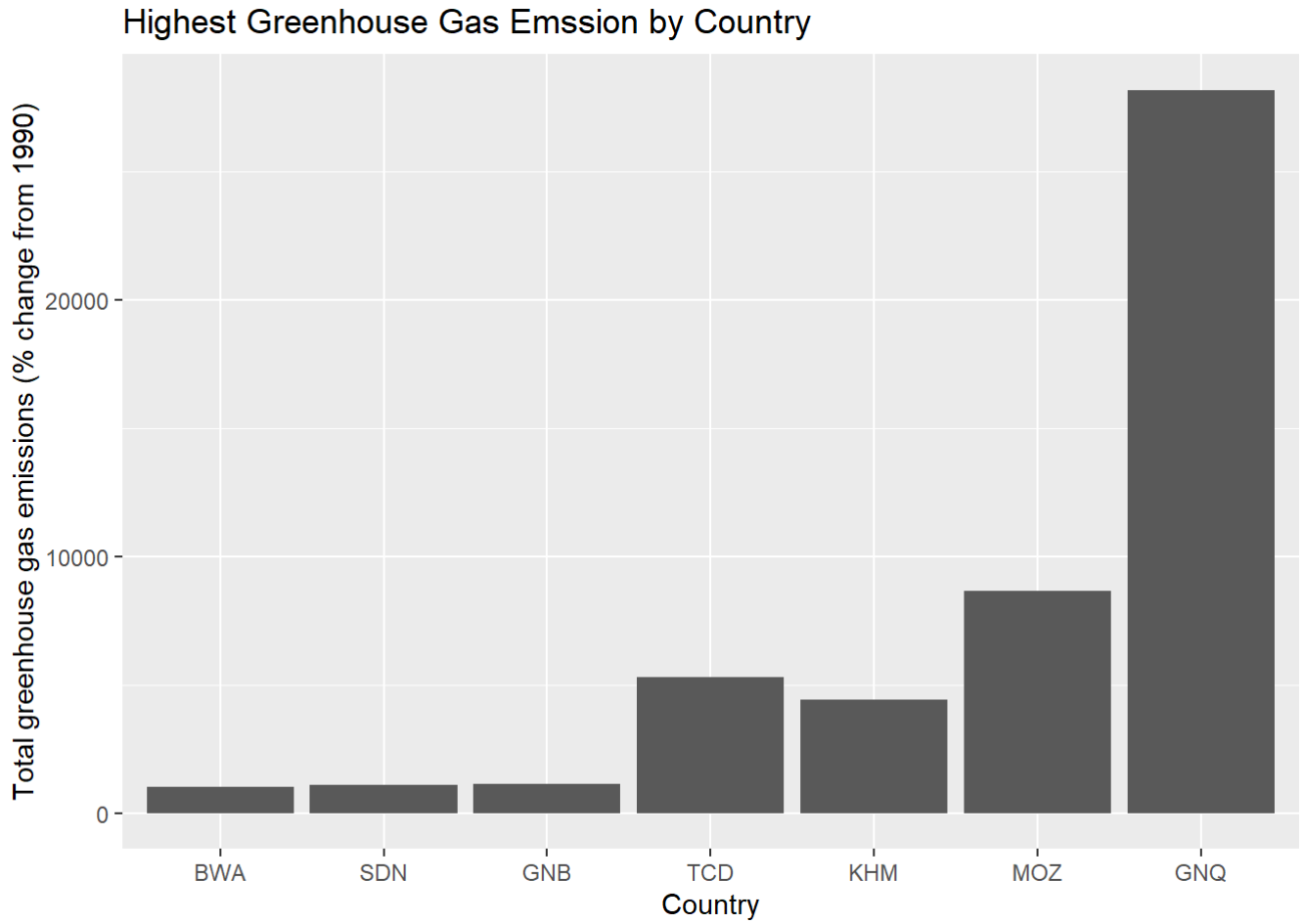
Statistics	BWA	SDN	GNB
Value	1,026.3536	1,066.2091	1,110.1458

Descriptive Statistics

To breifely summarize the statistics, it was all of the top 50 data points over the time frame from 1980 to 2018. If the country was repeated, it would add up all of the total greenhouse gas emissions.

Outliers

This data set not required information that needed the outliers package to review the data because it was only adding the top 50 data points. Therefore, there was no mean or anything factors that could have affected the overall analysis. It's also possible that due to having the statistics gathering the value a certain way, there was no significant data sets that might hinder the analysis.



Visualization

From the graph, it is easily shown that there are top 6 countries who have contributed the highest greenhouse gas emssion. Although GNQ had the higest overall, it's only with this country having the top over the past years is there a huge gap between the other countries. We can understand through this visualization how much GNQ's

emissions can lead to a huge contribution as greenhouse gas emission is one of the biggest factors with climate change.

Evaluation

Overall, as a way to understand climate change we must be able to understand the different factors that have contributed to this. First being greenhouse gas, as it contains properties absorbing infrared radiation keeping heat inside the atmosphere. This process itself contributes to climate change as the greenhouse gas is one of the biggest factors affecting it. By looking at just the US's greenhouse gas emission, the increasing trend can not be only that country's fault. It's most likely that all countries, even if a small increase together are strong enough to make huge impacts. Furthermore, another impact to climate change is carbon dioxide emission. With this increase of this, we are able to understand how it has damaged our economy. As it links into precautions we must take now, it shows us that more actions of how we handle our environment is important for economic reasons can be negatively impacted. In addition, carbon dioxide emissions also have a link to impacting land and ocean anomaly. By being able to understand environmental changes, we can better support how things on our earth are changing. Land and ocean temperatures are crucial to understand or atmosphere because they are a huge factor in predicting weather and natural disasters. Which further brings concerns to how climate change is an ongoing problem that brings many worrisome effects on humans. Although, we should argue that humans are already being impacted in different ways. One argument being how climate change has affected agricultural production in multiple countries. Our data has shown the top 12 countries that were affected and how their GDP output changed because of this. Agricultural production is important in many countries as this is a way to survive and what they can bring in and out of the country. By decreasing this, it decreases survival rate and the economy. Finally, an argument to how climate change has been impacted so much is by humans. Although not just one country is to blame, we must be able to determine what countries are contributing to most in order to find methods to help decrease climate change before it impacts us any further. Being able to determine over the past years what countries had the highest greenhouse gas emission, gives us a better understanding of how we can move forward to change our human negative impact to climate change.