````
---
title: "Climate Change"
author: "Stephen Smitshoek"
date: "24/05/2022"
output: pdf_document
---

# Introduction

  Climate change is an issue that is impacting our world and is directly caused by humans.  Some issues created by climate change are, more common extreme weather events, polar ice caps melting, rising sea levels and a rise in the global average temperatures.  If climate change is not stopped in the near future Earth could be permanently impacted and may become uninhabitable.

There are a number of people that do not believe in climate change, using data science it should be possible to prove that climate change is a fact and show the impacts that it is having on our world.

# Problem Statement

  To start it is required to prove that climate change does in fact exist. To do this global temperatures and sea levels will be plotted over time.  If global temperatures and sea level are both rising then it can be assumed that the climate is changing, and the next step is to see why.  Because scientists often say the change is caused by CO2 emissions, atmospheric CO2 levels over time will be plotted and a linear model will be fitted between CO2 levels and global temperatures. If there is a good fit between the two then a second model will be done checking how global CO2 emissions effect the atmospheric CO2 levels.  If this second model has a good fit than it is would be strong evidence in favor of humans being responsible for the climate change phenomenon.

# Analysis
## Import and Clean Data
```{r libraries, echo=FALSE, warning=FALSE, message=FALSE}
library(dplyr)
library(ggplot2)
```

```{r setup, echo=FALSE}
setwd("C:\\Users\\sksmi\\PeytoAccess\\Personal\\Bellevue\\DSC520\\dsc520")
global_temp <- read.csv("final_project\\data\\GlobalTemperatures.csv")
country_temp <- read.csv("final_project\\data\
\GlobalLandTemperaturesByCountry.csv")
co2_atmo <- read.csv("final_project\\data\\co2_atmo.csv")
co2_emiss_country <- read.csv("final_project\\data\\co2_emission.csv")
sea_level <- read.csv("final_project\\data\\sea_levels_2015.csv")
```

```{r global_temp Clean Up, echo=FALSE}
global_temp$dt <- as.Date(global_temp$dt, "%Y-%m-%d")
global_temp <- global_temp[global_temp$dt >= as.Date("1753-01-01", "%Y-%m-
%d"),]
global_temp <- subset(global_temp, select = c(dt,
````
````

```
                                                LandAverageTemperature,
                                                LandAverageTemperatureUncertainty))
global_temp <- aggregate(global_temp[,c(2,3)], by=list(format(global_temp$dt,
format="%Y")), FUN=mean)
colnames(global_temp) <- c("year", "AvgTemp", "AvgTempUncer")
```

```{r country_temp Clean Up, echo=FALSE}
country_temp$dt <- as.Date(country_temp$dt, "%Y-%m-%d")
country_temp <- aggregate(cbind(AverageTemperature,
                                AverageTemperatureUncertainty) ~
                            format(country_temp$dt, format="%Y") + Country,
                          data = country_temp,
                          FUN=mean)
colnames(country_temp)[1] <- "year"
```

```{r co2_atmo Clean Up, echo=FALSE}
co2_atmo$date <- as.Date(paste0(co2_atmo$Year, "-", co2_atmo$Month, "-",
"01"),
                         "%Y-%m-%d")
co2_atmo <- subset(co2_atmo, select = c(date,
                                        Carbon.Dioxide..ppm.,
                                        Seasonally.Adjusted.CO2..ppm.))
co2_atmo <- aggregate(co2_atmo[,c(2,3)], by=list(format(co2_atmo$date,
format="%Y")), FUN=mean, na.action=na.omit)
colnames(co2_atmo) <- c("year","CO2_ppm", "CO2_ppm_Seas_Adj")
```

```{r co2_emiss Clean Up, echo=FALSE}
colnames(co2_emiss_country)[4] <- "co2_tonnes"
co2_emiss_global <- aggregate(co2_emiss_country$co2_tonnes,
                              by=list(co2_emiss_country$Year),
                              FUN=sum)
colnames(co2_emiss_global) <- c("year", "co2_tonnes")
co2_emiss_global$year <- as.character(co2_emiss_global$year)
```

```{r sea_level Clean Up, echo=FALSE}
sea_level$date <- as.Date(paste0(substring(sea_level$Time, 1, 4), "-",
                                 "01", "-", "01"), "%Y-%m-%d")
sea_level <- aggregate(sea_level[,c(2,3)], by=list(format(sea_level$date,
format="%Y")), FUN=mean)
colnames(sea_level)[1] <- "year"
```

```{r climate_change_df Creation, echo=FALSE}
climate_change_df <- left_join(global_temp, co2_atmo, by="year")
climate_change_df <- left_join(climate_change_df, co2_emiss_global,
by="year")
climate_change_df <- left_join(climate_change_df, sea_level, by="year")
climate_change_df$year <- as.integer(climate_change_df$year)
```

How the Climate is Effected

```{r climate_fig_1, echo=FALSE, message=FALSE, warning=FALSE, fig.width=3, fig.height=3}
climate_change <- ggplot(data=climate_change_df)

climate_change + aes(x=year, y=AvgTemp) + geom_point() +
geom_smooth(method="gam")

climate_change + aes(x=year, y=GMSL) + geom_point() + geom_smooth() +
  ggtitle("Sea Level vs Time") + theme(plot.title=element_text(hjust=0.5)) +
  xlab("Year") + ylab("Global Mean Sea Level (mm)")
```

The above graphs show a clear increase in both the global sea level and the global temperature. This shows that there is indeed a change in climate occurring. It must now be shown why this is happening.

CO2 Levels
```{r co2, warning=FALSE, message=FALSE, echo=FALSE, fig.width=3, fig.height=3}
climate_change <- ggplot(data=climate_change_df)
climate_change + aes(x=year, y=CO2_ppm) +  geom_point()
climate_change + aes(x=year, y=co2_tonnes) + geom_point()
```

These two plots clearly show that CO2 levels in the atmosphere are rising and that human CO2 emissions starting increasing around 1850 then really took off in the 1950s.

Climate Change and CO2 Correlation
```{r correlations, echo=FALSE}
temp_CO2_lm <- lm(AvgTemp ~ CO2_ppm, data=climate_change_df)
lm_summary <- rbind(c(summary(temp_CO2_lm)$coefficients[2,c(1,4)],
                      summary(temp_CO2_lm)$adj.r.squared))


ppm_emmis_lm <- lm(CO2_ppm ~ co2_tonnes, data=climate_change_df)
lm_summary <- rbind(lm_summary, c(summary(ppm_emmis_lm)
$coefficients[2,c(1,4)],
                                  summary(ppm_emmis_lm)$adj.r.squared))

temp_emiss_lm <- lm(AvgTemp ~ co2_tonnes, data=climate_change_df)
lm_summary <- rbind(lm_summary, c(summary(temp_emiss_lm)
$coefficients[2,c(1,4)],
                                  summary(temp_emiss_lm)$adj.r.squared))

row.names(lm_summary) <- c("Temp vs PPM", "PPM vs Emissions",
                           "Temp vs Emissions")
colnames(lm_summary) <- c("Coefficient", "p<", "Adj. R2")
lm_summary[,c(1,2)] <- formatC(lm_summary[,c(1,2)], format="e", digits=2)
lm_summary[,c(3)] <- round(as.numeric(lm_summary[,3]), 2)
```

```{r kable, echo=FALSE}
lm_summary %>% knitr::kable(align="c", caption="Linear Model
Characteristics")
```

```

```{r echo=FALSE, message=FALSE, warning=FALSE}
ggplot(data=climate_change_df, aes(x=co2_tonnes, y=AvgTemp)) +
  geom_point() + stat_smooth(method="lm")
```

Implications

  The first linear model between the average global temperature and the CO2 ppm shows that 81% of the temperature rise can be correlated to a rise in CO2 levels with a $p < 0.001$. This shows that while not all of the temperature rise can be explained by rising CO2 levels the vast majority of it can be, and there is almost no chance that this relationship is by chance.

  The second model shows that 96% of the rise in CO2 levels in our atmosphere can be attributed to human CO2 emissions, and again there is almost no chance that this relationship is by chance.

  The final model confirms the direct relationship between the CO2 emitted by humans each year and the rise in global temperatures. It shows that 52% of the rise in global temperatures can be explained by a rise in CO2 emissions. The figure above also shows how closely the temperature and emissions are related.

Limitations

  The data set that was used for this analysis was missing a large portion of data for both the global sea levels and CO2 levels. If the data could be filled in it would strengthen the meaning behind this analysis.

  Rises in the global temperature are having compounding effects by causing more forest fires and CO2 sequestered in the oceans to be released. This data is not taken into account in this analysis but would help to build a stronger model.

  Although it is very clear that there is a strong correlation between CO2 emissions and the climate change the globe is seeing it does not confirm a causation. Scientists are required to explain how the two events are related and prove that in this case the correlation is a causation.

Concluding Remarks

  The climate of the planet is changing and much of this can be explained by the emissions that humans are releasing into the atmosphere. By visually observing the data and further analyzing it with a linear model it can be clearly seen that there is the average global temperature is rising and with that the sea levels are rising as well. Further analysis of these trends fits a very strong relationship between the CO2 emissions that humans are causing and the changes to our climate. This data paired with scientific explanation proves that humans are having a direct and immense impact on the climate change.
```