# Data 607 - FinalProject - Climate Change

*Suma Gopal*

## Objective

The objective of this project is to observe how the climate and the environment have changed over time, and to prove that our planet is facing dramatic shifts in climate and the environment overall. We perform exploratory analysis on the following various climate data:
* Sea levels * CO2 emissions * Glacier Mass

```
  library(jsonlite)
  library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
  library(DT)
  library(readr)
  library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------- tidyverse
```

```
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.0     v stringr 1.3.1
## v tidyr   0.8.2     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------------------- tidyverse_confl
## x dplyr::filter()  masks stats::filter()
## x purrr::flatten() masks jsonlite::flatten()
## x dplyr::lag()     masks stats::lag()
```

## Data Acquisition

- JSON
- CSV

### JSON Data

#### Sea Levels

**Data Source**
Below, we acquire data in JSON format. The data is copyrighted by CSIRO. Here is the data source site link: https://datahub.io/core/sea-level-rise#resource-csiro_alt_gmsl_yr_2015
This data describes cumulative changes in sea level for the world's oceans since 1880. It is based on a

combination of long-term tide gauge measurements and recent satellite measurements. The average refers to the height of the ocean surface. It is said in the data source's site, that the ocean floor has been gradually sinking since the last Ice Age peak which was 20,000 years ago!

```
json_seaLevelFile <- 'https://datahub.io/core/sea-level-rise/datapackage.json'
json_seaLevelData <- fromJSON(paste(readLines(json_seaLevelFile), collapse=""))
```

```
## Warning in readLines(json_seaLevelFile): incomplete final line found on
## 'https://datahub.io/core/sea-level-rise/datapackage.json'
```

```
# get list of all resources:
print(json_seaLevelData$resources$name)
```

```
##  [1] "validation_report"              "csiro_alt_gmsl_mo_2015_csv"
##  [3] "csiro_alt_gmsl_yr_2015_csv"     "csiro_recons_gmsl_mo_2015_csv"
##  [5] "csiro_recons_gmsl_yr_2015_csv"  "epa-sea-level_csv"
##  [7] "csiro_alt_gmsl_mo_2015_json"    "csiro_alt_gmsl_yr_2015_json"
##  [9] "csiro_recons_gmsl_mo_2015_json" "csiro_recons_gmsl_yr_2015_json"
## [11] "epa-sea-level_json"             "sea-level-rise_zip"
## [13] "csiro_alt_gmsl_mo_2015"         "csiro_alt_gmsl_yr_2015"
## [15] "csiro_recons_gmsl_mo_2015"      "csiro_recons_gmsl_yr_2015"
## [17] "epa-sea-level"
```

```
# get all tabular data(if exists any)
for(i in 1:length(json_seaLevelData$resources$datahub$type)){
  if(json_seaLevelData$resources$datahub$type[i]=='derived/csv'){
    path_to_file = json_seaLevelData$resources$path[i]
    seaLevelData <- read.csv(url(path_to_file))
  }
}
```

Here is the raw data:

```
datatable(seaLevelData, options = list(pageLength = 5))
```

```
## PhantomJS not found. You can install it with webshot::install_phantomjs(). If it is installed, pleas
```

**Data Transformation**

Here we perform a data transformation by performing a select operation on the data from 2 specific columns, called: Year and CISRO Adjusted Sea Level. CISRO Adjusted Sea level is measured in units of inches, representing changes in sea level.

```
seaLevels <- seaLevelData %>%
 select("Year", "CSIRO.Adjusted.Sea.Level")

# output to data table format
datatable(seaLevels, options = list(pageLength = 5))
```

### Glacier Mass

**Data Source**

Below we acquire average cumulative mass balance of reference Glaciers worldwide, from years 1945 - 2014. Here is the data source's website: https://datahub.io/core/glacier-mass-balance#readme
Negative mass values point to a net loss of ice and snow compared with 1945. The units of mass measurement are in meters, representing changes in average glacier thickness.

```
  json_glacierFile <- 'https://datahub.io/core/glacier-mass-balance/datapackage.json'
json_glacierData <- fromJSON(paste(readLines(json_glacierFile), collapse=""))
```

```
## Warning in readLines(json_glacierFile): incomplete final line found on
## 'https://datahub.io/core/glacier-mass-balance/datapackage.json'
```

```
# get list of all resources:
print(json_glacierData$resources$name)
```

```
## [1] "validation_report"        "glaciers_csv"
## [3] "glaciers_json"            "glacier-mass-balance_zip"
## [5] "glaciers"
```

```
# print all tabular data(if exists any)
for(i in 1:length(json_glacierData$resources$datahub$type)){
  if(json_glacierData$resources$datahub$type[i]=='derived/csv'){
    path_to_file = json_glacierData$resources$path[i]
    glacierData <- read.csv(url(path_to_file))
  }
}
```

Here is the raw data:

```
  datatable(glacierData, options = list(pageLength = 5))
```

### CSV Data

### CO2 Emissions

#### Data Source

Below we acquire Carbon Dixoide emissions data, from a CSV file. Here is the data source's website: https://datahub.io/core/co2-ppm#readme
The data represents trends in atmospheric Carbon Dioxide, of a global average over marine surface sites. Here is the raw data:

```
  emissionsData <- read_csv("co2-gr-gl_csv.csv")
```

```
## Parsed with column specification:
## cols(
##   Year = col_character(),
##   `Annual Increase` = col_double(),
##   Uncertainty = col_double()
## )
```

```
  datatable(emissionsData, options = list(pageLength = 5))
```

#### Data Transformation

Next we will perform a data transformation by performing a select operation on the data from 2 specific columns, called: Year and Annual Increase. Annual Increase values are in units of ppm (parts per million), measuring annual CO2 increase.

```
  emissions <- emissionsData %>% select("Year", "Annual Increase")
  datatable(emissions, options = list(pageLength = 5))
```

### CO2 Emissions From Fossil Fuels

Below we acquire data on Carbox Dioxide emissions from fossil fuels by nation, since the year of 1751. Here is the data source's website: https://datahub.io/core/co2-fossil-by-nation#data.
Data citation: Boden, T.A., G. Marland, and R.J. Andres. 2013. Global, Regional, and National Fossil-

Fuel CO2 Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A. doi 10.3334/CDIAC/00001_V2013

Here is the raw data:

```
  nationEmissionsData <- read_csv("fossil-fuel-co2-emissions-by-nation_csv.csv")
```

```
## Parsed with column specification:
## cols(
##   Year = col_double(),
##   Country = col_character(),
##   Total = col_double(),
##   `Solid Fuel` = col_double(),
##   `Liquid Fuel` = col_double(),
##   `Gas Fuel` = col_double(),
##   Cement = col_double(),
##   `Gas Flaring` = col_double(),
##   `Per Capita` = col_double(),
##   `Bunker fuels (Not in Total)` = col_double()
## )
```

```
  datatable(nationEmissionsData, options = list(pageLength = 5))
```

```
## Warning in instance$preRenderHook(instance): It seems your data is too
## big for client-side DataTables. You may consider server-side processing:
## https://rstudio.github.io/DT/server.html
```