

# SOP 85: Using NOAA Climate Records

Marc Los Huertos

July 4, 2017

## 1 Introduction

## 2 Preparing CSV files

### 2.1 Preprocessing CSV files

In most cases, we don't need to preprocess the csv files.

### 2.2 Upload CSV Files into Appropriate Rstudio Directory

## 3 Reading CSV Files into R

Although the csv file might be present in an Rstudio directory, it is not in the R environment. One way to confirm this is to look at the Rstudio window tab 'Environment', where the file listed.

To have the file in the R environment, we read the file using the `read.csv()` function. The function is expecting the path and name of the file as an argument inside the parentecies, which is tough to type without making errors.

### 3.1 Assigning the File Path and Name

Okay, now we know what the file name and path look like from the eyes of R. In this example, the path and filename is:

```
"/home/CAMPUS/mwl04747/github/Climate_Change_Narratives/Data/Los Angeles_878024.csv"
\end{verbatim}
```

We could paste the whole mess into the `read.csv()` function. Okay, let's keep moving

First, we will import the CSV file. In most cases, we don't need to preprocess the csv files

We will use a pop-up window to select the file the first time and then we'll assign the file

```
\begin{verbatim}
> choose.file()
```

Need an image of the pop-up window...

## 3.2 Reading a CSV File into a Dataframe

```
maunaloa <- read.csv(filename)
```

## 3.3 Importing CSV Files

Importing Scripts

```
import = read.csv(file)
```

```
# Read CSV
```

```
file = "/home/CAMPUS/kaj41925/Climate_Change_Narratives/Khalil/NewOrleansNOAAdata.csv"
```

```
import = read.csv(file)
```

## 3.4 Confirming the Dataframe

## 3.5 Checking the Data

Now we will check the data by plotting it.

## 3.6 Checking the Data

What's going on? What is the deal with the -9999? These are used for missing data. We need to remove them!

Okay, now we'll check again, for dates less than Dec. 31, 1913:

Yikes! What's wrong? As it turns out the problem is that with how the dates are specified. In particular, the Dec 31 to Jan 1 transition:

19131231 -> 19140101, if you use these as numbers it should be 19131232, 19131232, etc. You see we are missing lots numbers!

## 3.7 Fixing the Dates

First, we convert the date to a string of character values. Next, we'll convert the strings to a data format.

## 3.8 Checking the New Dates

## 3.9 Subset Sites

Olivia has two sites in here data, so, we need to subset it.

Let's choose the FAIRPLEX NY US because the record is longer than the airport.

Okay, you know have created a data frame. To confirm this, type `str(maunaloa)` and you should see some strange text that describes the data frame. This function allow you to peer into the data frame structure. You you see it is a data frame and it has several variables and each one has certain characteristics and R even shows you some of the observations. This is a good thing to get into the habitat of check, for you want to ensure the data have been imported in a way that you expect.

Remember, a data frame is a list of vectors. To access the data inside the data frame, you can use the following command

```
#maunaloa$average
```

to dump the average CO<sub>2</sub> concentrations readings onto your screen as a vector. You should see some 627 observations, depending on how recent the data have been uploaded. So, the dollar symbol is used to drill into the data frame vectors. And when you look at the `str()` function again, you will see these dollar signs again.

## 4 Preparing Records for Analysis

### 4.1 Converting Dates

### 4.2 Re-assigning Missing Values