Data Analyst Nanodegree Project 1

Justin Reid

7/13/2018

Reading in Temperature Data

The Temperature Data for the Globe and for the city location, *Charlotte, NC*, were loaded in from Udacity's servers via the following SQL code:

```
SELECT *
FROM city_data
WHERE city = 'Charlotte';

SELECT *
FROM global_data;
```

This data was saved in CSV format with the global data being in `global_data.csv` and Charlotte, NC data stored in `city_temp_clt.csv`.

Load in the tidyverse and read in data

```
library(tidyverse)
```

```
global_data <- read_csv('data/global_data.csv')
city_data <- read_csv('data/city_temp_clt.csv')
```

Combine Data Into One Table to Create One Plot for Both Time Series

```
complete_data <- left_join(global_data, city_data, by="year") %>%
    select(year, avg_temp.x, avg_temp.y)
colnames(complete_data) <- c("year", "global_avg_temp", "clt_avg_temp")
```

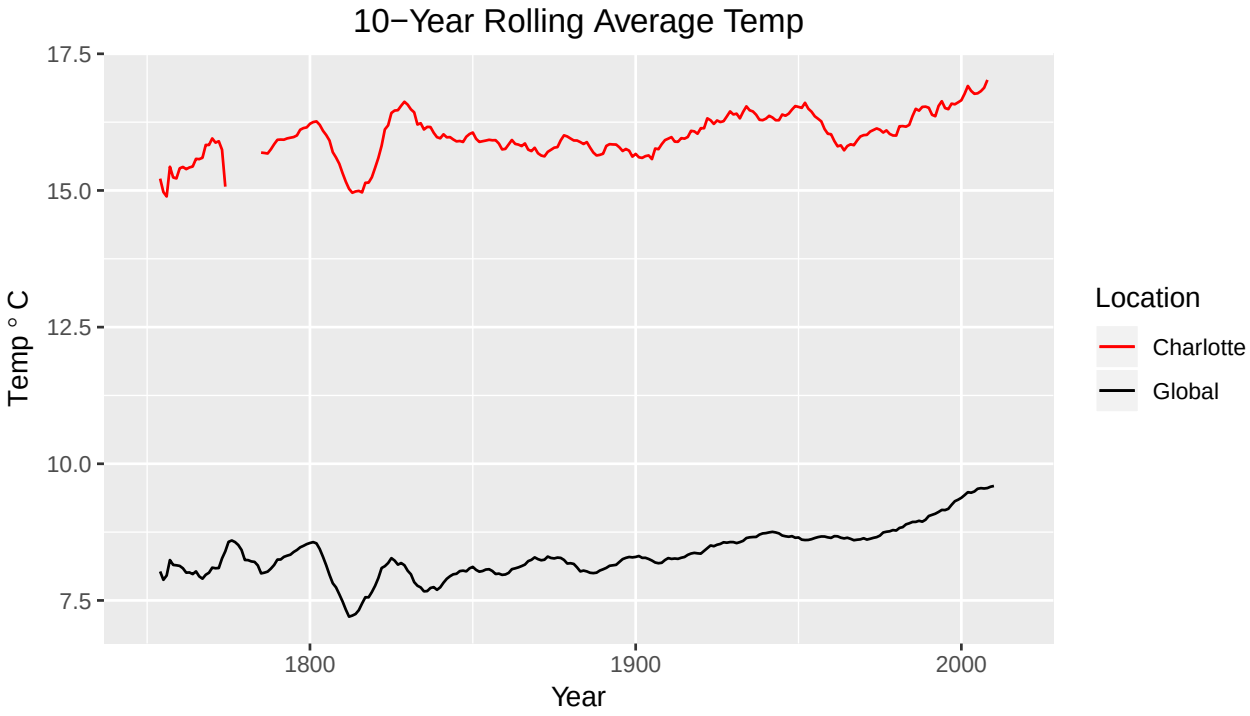
Import the Zoo Package for Running Average Calculations

```
library(zoo)
```

Create running average columns in order to create smooth plots (3 year avg)

```
complete_data <- mutate(
    complete_data,
    global_roll_avg = rollmean(global_avg_temp, 10, fill = NA)) %>%
    mutate(clt_roll_avg = rollmean(clt_avg_temp, 10, fill = NA))
```

Plot 10-Year Running Average for Both Charlotte and the World



Notable Observations

1. Charlotte is warmer location on average than the Global Mean.
2. Charlotte has more variablility (noise) in its data compared the entire globe.
3. There are some missing data before 1800 for Charlotte, more than likely due to a lack of observations so early in the record.
4. Both Charlotte's and Global data suggest a general upward trend, with Charlotte's beginning after 1900 and the World's earlier before 1900 but after 1850.