Udacity Project 1: Exploring Weather Trends

Josh Goldberg

12/29/2017

Load Libraries

```
library(tidyverse)
library(knitr)
```

Read Data

```
city <- read_csv(
  'results-2.csv',
  col_type = cols(
  year = col_integer(),
  city = col_character(),
  country = col_character(),
  avg_temp = col_double()
  )
  )

cities <- read_csv(
  'results-3.csv',
  col_type = cols(
  year = col_integer(),
  city = col_character(),
  country = col_character(),
  avg_temp = col_double()
  )
  )

global <- read_csv(
  'results.csv',
  col_type = cols(
  year = col_integer(),
  avg_temp = col_double()
  )
  )
```

Function: Calculate Moving Average

```
movavg <- function(x, interval = 5) {
  j <- interval
  moving_avg <- vector(mode = 'double', length = length(x))
  for(i in seq_along(x)) {
    moving_avg[j] <- mean(x[i:interval])
    if (interval >= length(x)) {
      return(moving_avg)
    } else {
      interval <- interval + 1
    }
```

```
    j <- j + 1
  }
}
```

Transform Data

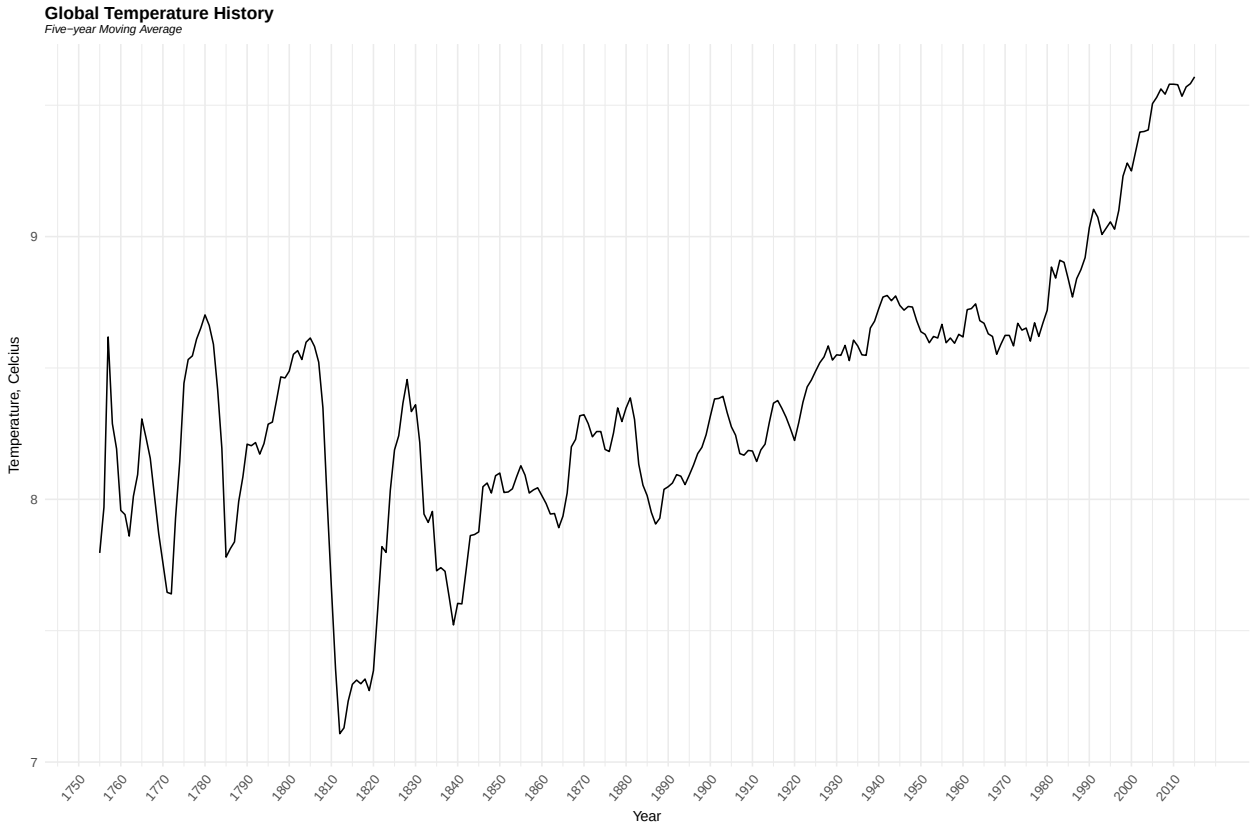
```
global$city <- 'global'
global$country <- 'global'

result_all <- global %>%
  select(year, city, country, avg_temp) %>%
  bind_rows(cities) %>%
  group_by(city) %>%
  mutate(moving_avg = movavg(avg_temp, interval = 5))

result_all_country <- result_all %>%
  filter(!is.na(avg_temp)) %>%
  group_by(country, year) %>%
  summarise(avg_temp = mean(avg_temp, na.rm = TRUE)) %>%
  mutate(moving_avg = movavg(avg_temp, interval = 5))

result_all_country$city <- result_all_country$country

result_all <- result_all_country %>%
  select(year, city, country, avg_temp, moving_avg) %>%
  filter(country == 'United States') %>%
  bind_rows(result_all)

cold_countries <- filter(result_all, moving_avg < 3, moving_avg > 0) %>%
  arrange(moving_avg, country) %>%
  distinct(country)
```

Historical Moving-average Temperatures: Chicago, Illinois vs. Global

Barring extraordinary events, global temperatures have been climbing over the past couple hundred years. The chart below could be mistaken for the stock market without aiding details, as occasional dips and a consistent upward trend is commonplace in the equity indexes.

```
ggplot(data = filter(
  result_all,
  year > 1754,
  city %in% c('global')
  ),
  aes(year, moving_avg)) +
  geom_line() +
  scale_color_brewer(
  type = 'qual',
  palette = 2
  ) +
  scale_x_continuous(breaks = seq(1750, 2010, 10)) +
  labs(title = 'Global Temperature History', subtitle = 'Five-year Moving Average', x = 'Year', y = 'Te
  theme_minimal() +
  theme(
```

```
  axis.text.x = element_text(angle = 50, hjust = 1),
  plot.title = element_text(
  size = 12,
  face = 'bold',
  color = 'black',
  vjust = -1
  ),
  plot.subtitle = element_text(size = 8, face = 'italic',
  color = 'black'),
  axis.title.x = element_text(size = 10, vjust = -0.2),
  axis.title.y = element_text(size = 10),
  legend.position = 'top'
  )
```

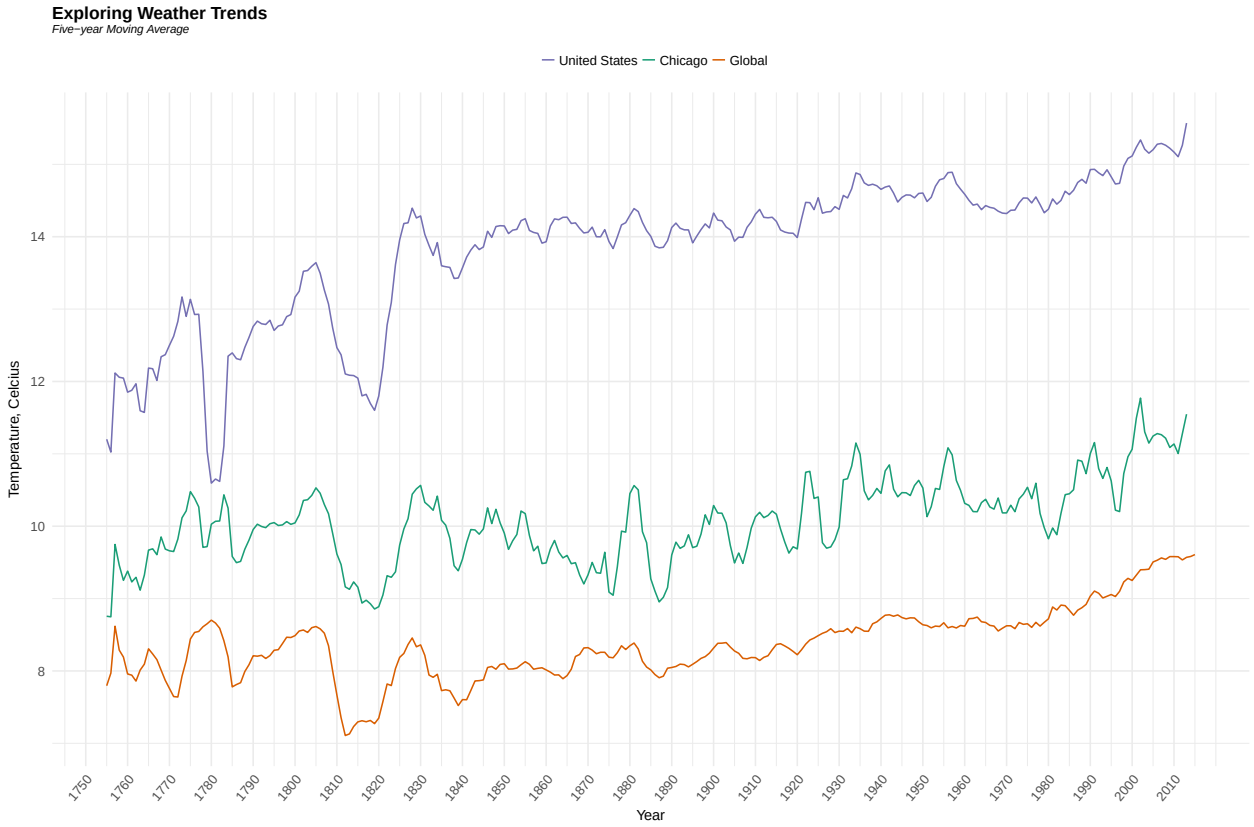


Chicago, Illinois has been warmer over time compared to global moving average temperatures (five-year period). Global temperatures have less volatility than Chicago due to the amount of data consolidated (Chicago is one city, global includes many cities; in other words: Law of Small Numbers.

There are some exceptions in Chicago and global temperature trends: * In 1770, the dip in global temperatures is not as pronounced in Chicago. * In ~1779, the dip in temperature in Chicago is not visible in global temperatures. * In 1995, the severe dip in Chicago temperatures is not pronounced in global temperatures. There are more trivial differences; however, generally speaking, the plot illustrates tandem trends between Chicago and global temperatures (with Chicago being consistently hotter).

```
ggplot(data = filter(
  result_all,
  year > 1754,
  city %in% c('Chicago', 'United States', 'global')
```

```
),
aes(year, moving_avg)) +
geom_line(aes(color = city)) +
scale_color_brewer(
type = 'qual',
palette = 2,
guide = guide_legend(
title = NULL,
keywidth = .75,
keyheight = .75
),
breaks = c('United States', 'Chicago', 'global'),
labels = c('United States', 'Chicago', 'Global')
) +
scale_x_continuous(breaks = seq(1750, 2010, 10)) +
labs(title = 'Exploring Weather Trends', x = 'Year', subtitle = 'Five-year Moving Average', y = 'Temp
theme_minimal() +
theme(
axis.text.x = element_text(angle = 50, hjust = 1),
plot.title = element_text(
size = 12,
face = 'bold',
color = 'black',
vjust = -1
),
plot.subtitle = element_text(size = 8, face = 'italic',
color = 'black'),
axis.title.x = element_text(size = 10, vjust = -0.2),
axis.title.y = element_text(size = 10),
legend.position = 'top'
)
```



Chicago's overall warmer temperature can be explained by the fact that its moving average temperatures are not impacted by the extremely cold moving average temperatures included in global calculations. Here are some of the coldest countries included in global temperature calculations:

```
kable(cold_countries)
```

country
Russia
Kazakhstan
Norway
China
Canada
United States

```
ggplot(data = filter(
  result_all_country,
  year > 1754,
  country %in% (cold_countries$country)
  ),
  aes(year, moving_avg)) +
  geom_line(aes(color = country)) +
  scale_color_brewer(
  type = 'qual',
  palette = 2,
  guide = guide_legend(
  title = NULL,
  keywidth = .75,
```

```
keyheight = .75
)
) +
scale_x_continuous(breaks = seq(1750, 2010, 10)) +
labs(title = 'Exploring Weather Trends', x = 'Year', subtitle = 'Five-year Moving Average', y = 'Tempe
theme_minimal() +
theme(
axis.text.x = element_text(angle = 50, hjust = 1),
plot.title = element_text(
size = 12,
face = 'bold',
color = 'black',
vjust = -1
),
plot.subtitle = element_text(size = 8, face = 'italic',
color = 'black'),
axis.title.x = element_text(size = 10, vjust = -0.2),
axis.title.y = element_text(size = 10),
legend.position = 'top'
)
```