

# Homework Assignment #2

ENVE 681

*Nicola Horscroft*

*October 20, 2017*

## Introduction

This problem is designed to introduce work with R and to do some basic data manipulation and simple programming. You will download and use precipitation data (see ppt.xlsx file, given in inches) for the Philadelphia Airport. The rows represent years from 1980-2014 and the columns represent months from January to December.

### 1. Read the data into R

```
library(readxl)
library(tidyverse)
library(lubridate)

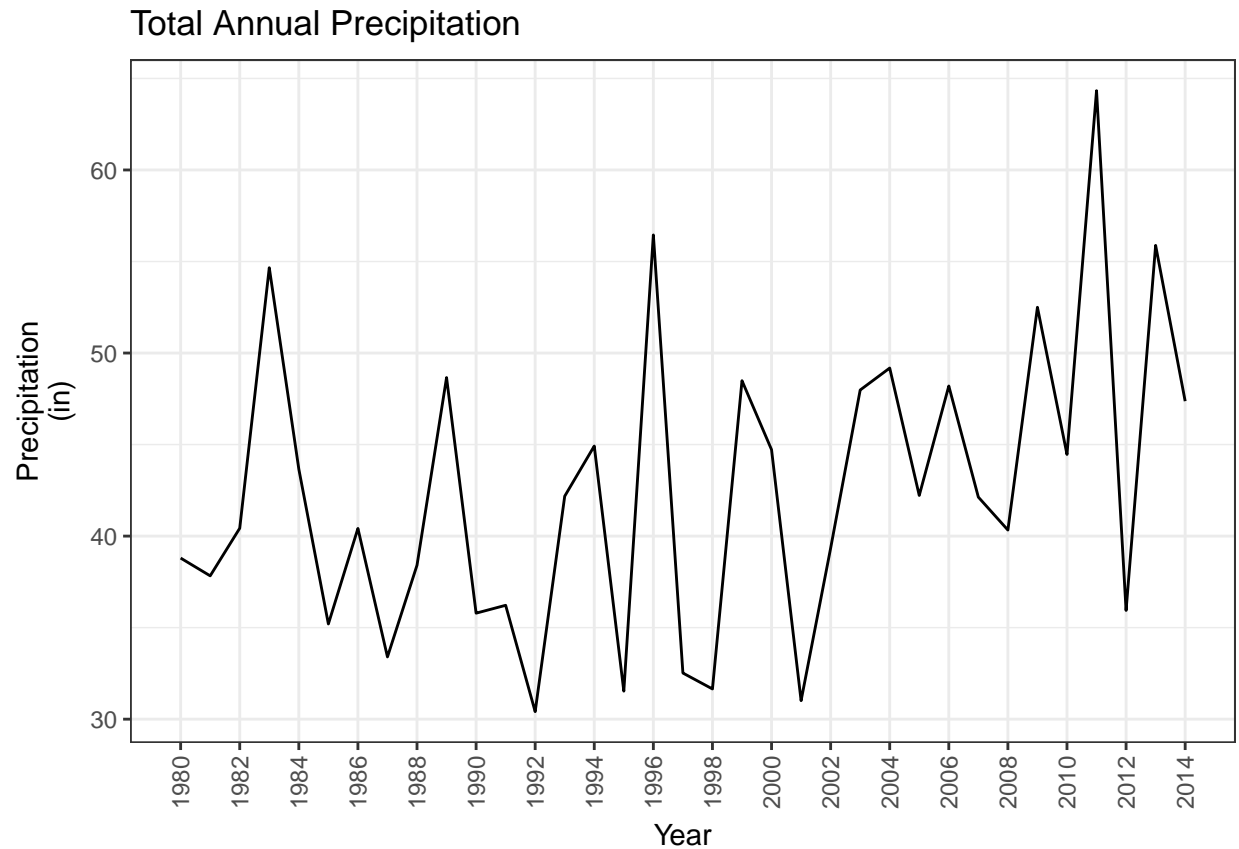
ppt <- read_excel("HW 2/ppt.xlsx")

xppt <- ppt %>%
  mutate(year = 1980:2014)
```

### 2. Calculate the total annual precipitation for each year and plot these data versus year

```
answer_2 <- xppt %>%
  gather(key = "month", value = "value", Jan:Dec) %>%
  group_by(year) %>%
  summarize(sum = sum(value))

ggplot(answer_2, aes(year, sum)) +
  geom_line() +
  scale_x_continuous(breaks = seq(1980, 2014, 2)) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 90, hjust = 1, vjust = 0.5),
        panel.grid.minor.x = element_blank()) +
  labs(x = "Year", y = "Precipitation\n(in)", title = "Total Annual Precipitation")
```



3. For a given year (you can pick), calculate and display the minimum, maximum, and mean monthly precipitation

```

xppt %>%
  gather(key = "month", value = "value", Jan:Dec) %>%
  filter(year == 1990) %>%
  group_by(year) %>%
  summarize(mean = mean(value),
            min = min(value),
            max = max(value)) %>%
  pandero::pander()

```

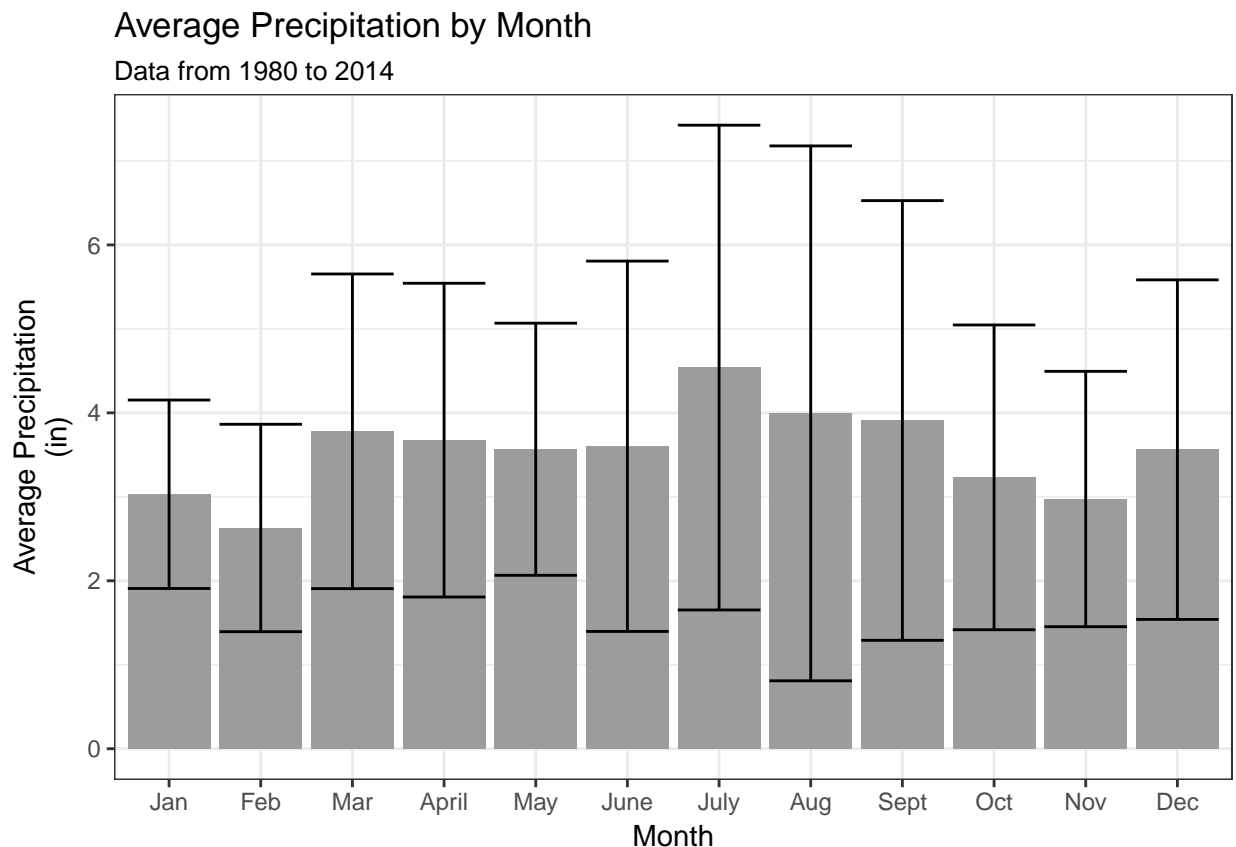
year	mean	min	max
1990	2.982	1.17	6.08

4. Calculate the mean monthly precipitation for each month and plot the values using a bar chart. Now add “error bars” showing the standard deviations of the monthly means.

```
answer_4 <- xppt %>%
  gather(key = "month", value = "value", Jan:Dec) %>%
  mutate(month = factor(month, levels = c("Jan", "Feb", "Mar",
                                           "April", "May", "June",
                                           "July", "Aug", "Sept",
                                           "Oct", "Nov", "Dec"))) %>%

  group_by(month) %>%
  summarize(mean = mean(value),
            std = sd(value))

ggplot(answer_4, aes(month, mean)) +
  geom_col(fill = "#9C9C9C") +
  geom_errorbar(aes(x = month, ymin = mean - std, ymax = mean + std), color = "black") +
  theme_bw() +
  labs(x = "Month", y = "Average Precipitation\n(in)",
       title = "Average Precipitation by Month",
       subtitle = "Data from 1980 to 2014")
```



## 5. Plot all the monthly precipitation values consecutively as a line plot

```
answer_5 <- xppt %>%
  gather(key = "month", value = "value", Jan:Dec) %>%
  mutate(month = factor(month, levels = c("Jan", "Feb", "Mar",
                                           "April", "May", "June",
                                           "July", "Aug", "Sept",
                                           "Oct", "Nov", "Dec"))) %>%

  mutate(date = ifelse(as.numeric(month) < 9,
                       paste(as.character(year), "-0",
                             as.numeric(month), "-01", sep = ""),
                       paste(as.character(year), "-",
                             as.numeric(month), "-01", sep = ""))) %>%
  mutate(date = as.POSIXct(strptime(date, format = "%Y-%m-%d")))

ggplot(answer_5, aes(date, value)) +
  geom_line() +
  scale_x_datetime(date_breaks = "5 years", date_labels = "%Y") +
  theme_bw() +
  labs(x = "Date", y = "Precipitation\n(in)",
       title = "Yearly Precipitation",
       subtitle = "Data from 1980 to 2014")
```

