Shana Green

DATA 606 - Lab 1

Due Date: 8/30/2020

```r
library(tidyverse)
library(openintro)
library(tinytex)
```

```r
#source('more/arbuthnot.r')
source('https://raw.githubusercontent.com/jbryer/DATA606/master/inst/labs/Lab1/more/arbuthnot.r')
```

```r
arbuthnot <- arbuthnot %>%
  mutate(total = boys + girls)
```
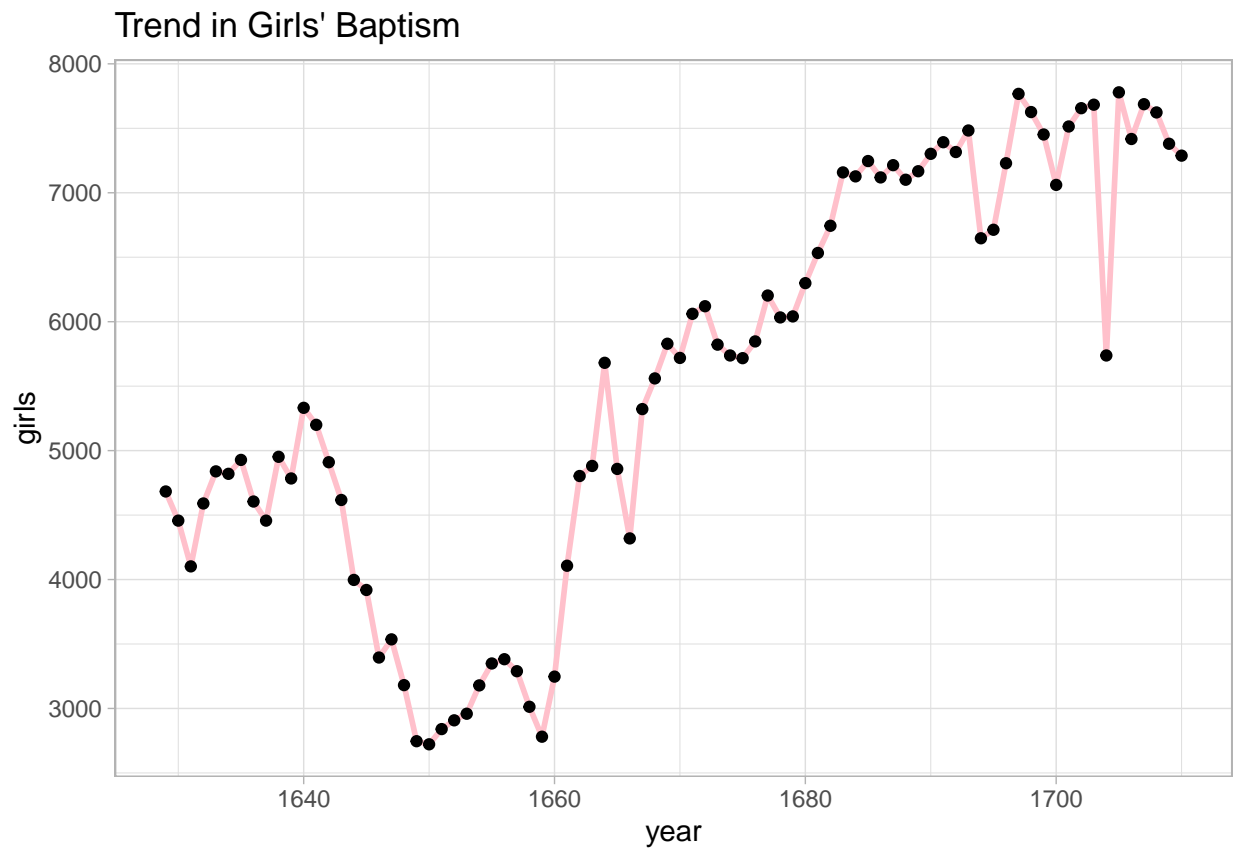
Exercise 1

```r
arbuthnot$girls
```

Exercise 2

There is no linear, quadratic, or exponential relationship of the trend in girls' baptisms over time. The largest decrease and increase of girls' baptism occurred between 1703 and 1705.

```r
trendgirls<-ggplot(data = arbuthnot, aes(x = year, y = girls)) +
    geom_line(linetype="solid",color="pink", size=1)+
    theme_light()
print(trendgirls+ggtitle("Trend in Girls' Baptism")+geom_point())
```
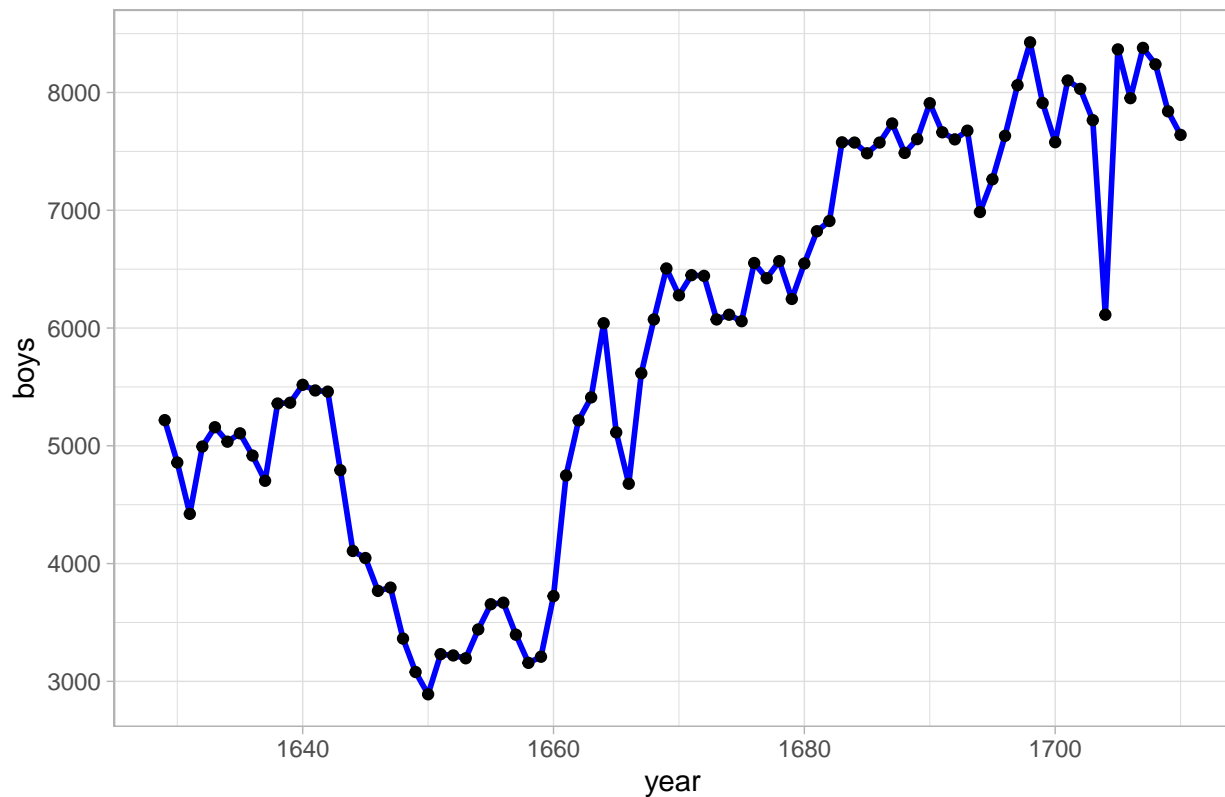
Trend in Girls' Baptism

Exercise 3

I see a graph that looks very similar to the trend in girl's baptism over time!

```
trendboys<-ggplot(data = arbuthnot, aes(x = year, y = boys)) +
    geom_line(linetype="solid",color="blue", size=1)+
    theme_light()
print(trendboys+ggtitle("Trend in Boys' Baptism")+geom_point())
```

Trend in Boys' Baptism

Exercise 4

```r
range(present$year) #Years included in the data set

dim(arbuthnot) #Dimensions of data frame

names(arbuthnot)#List of column names here
```

Exercise 5

First I used the dataframe arbuthnot to add the data for boys and girls below:

```r
arbuthnot$boys+arbuthnot$girls
```

Second, I compared it to the dataframe present to add the data for boys and girls:

```r
present$boys+present$girls
```

Upon reviewing both data sets, the number count per year of boys and girls in present is in the millions in comparison to the 10,000 range for arbuthnot. They are not similar in either dimension.

Exercise 6

Insert any text here.

```
boyspresent<-ggplot(data = present, aes(x = year, y = boys)) +
    geom_line(linetype="solid",color="red", size=1)+
    theme_light()
print(boyspresent+ggtitle("Boy's Present Day Birth Records")+geom_point())
```
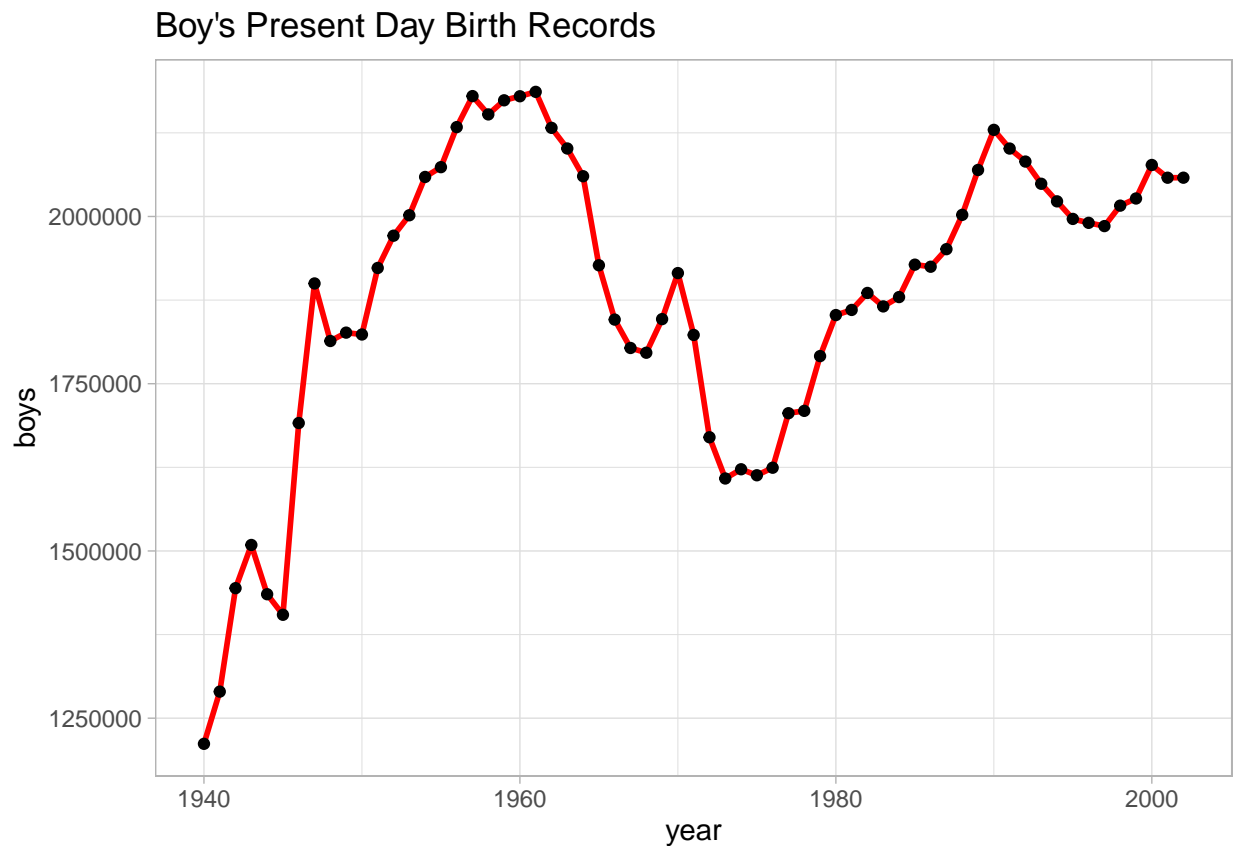

Exercise 7

According to the information provided in the lab, I used a mutate function to create and add a total column. This is the same exact method used for arbuthnot dataframe. Upon reviewing the results in descending order for total, 1961 had the most combined totals for boys and girls.

```
present <- present %>%
  mutate(total = boys + girls)

present %>%
  arrange(desc(total))
```