Lab 1: Intro to R

Code

Kevin

1/19/2023

Hide

```
data(arbuthnot)
```

Exercise 1:

Hide

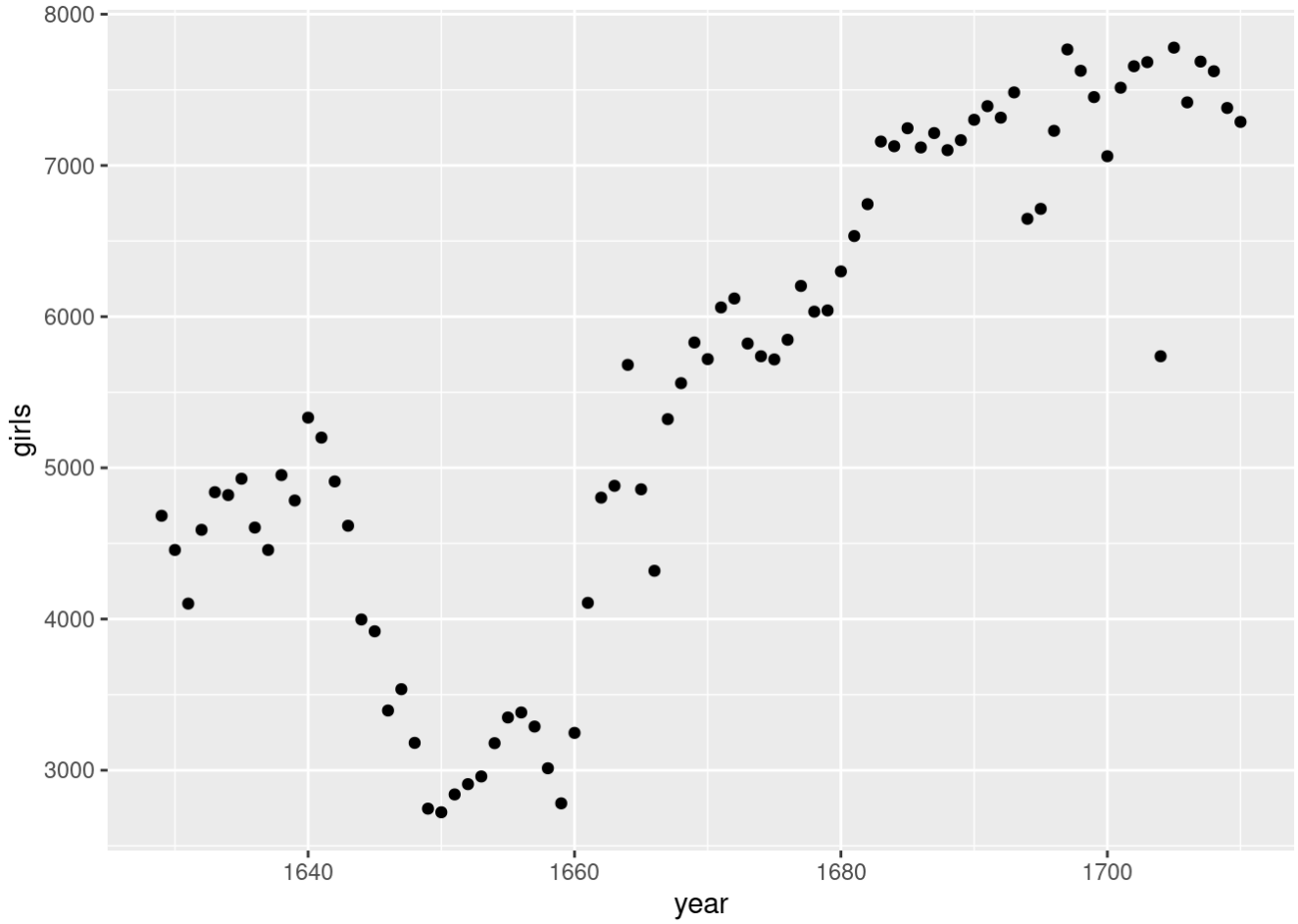
```
arbuthnot$girls
```

Exercise 2:

Hide

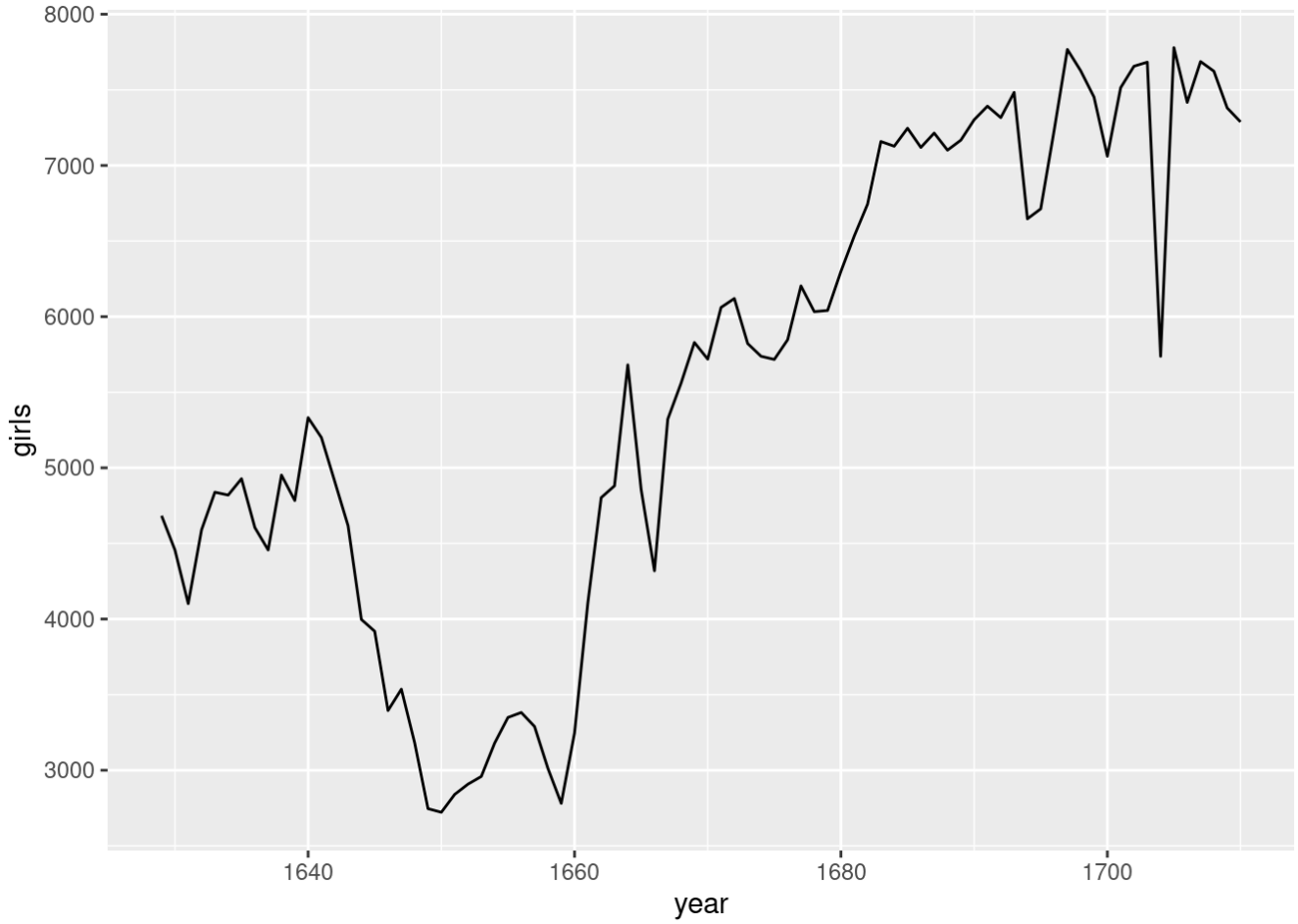
```
qplot(x = year, y = girls, data = arbuthnot)
```

```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
```



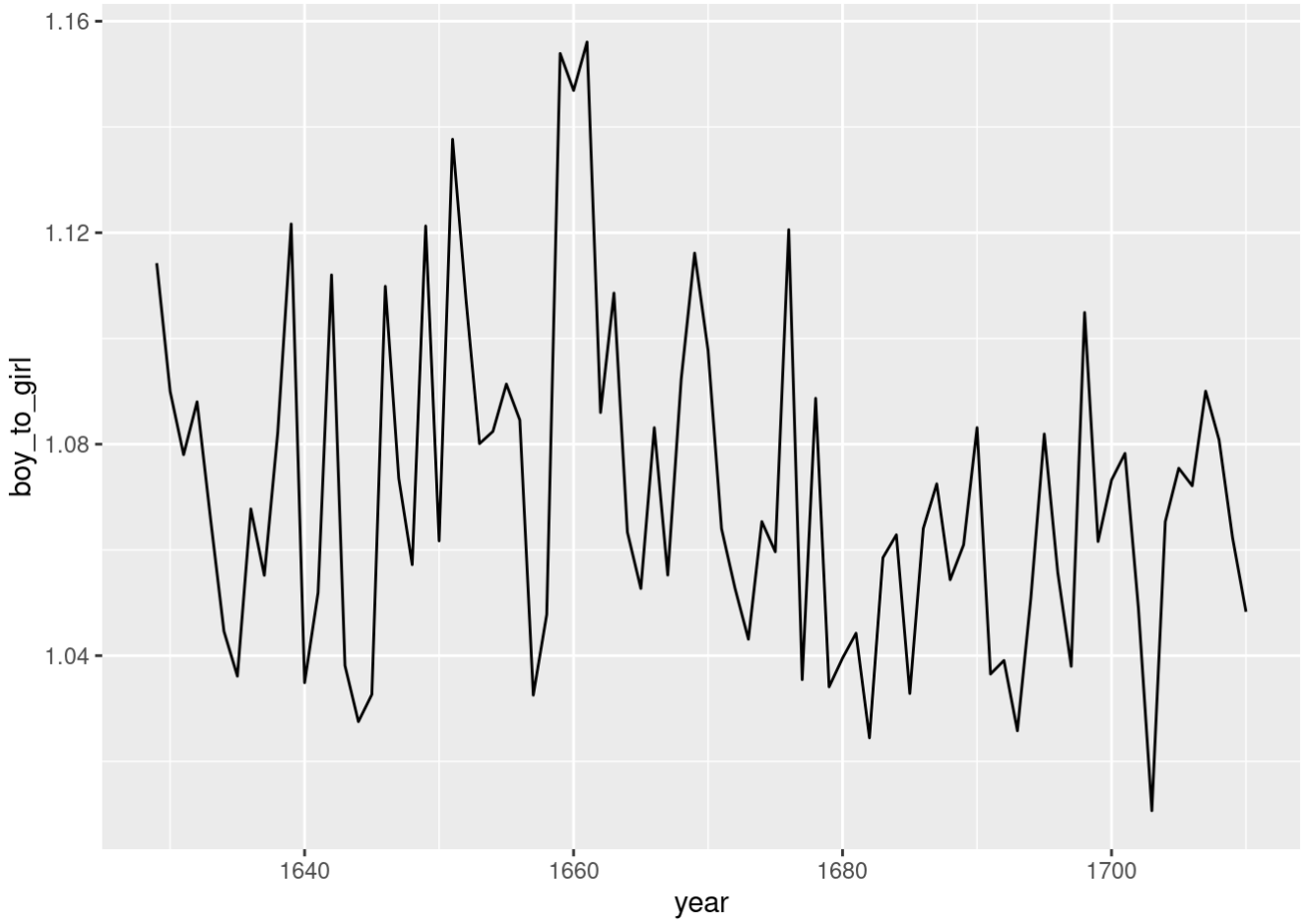
Hide

```
qplot(x = year, y = girls, data = arbuthnot, geom = "line")
```



The trend is going up, with a dip in the 40’s through 60’s

Exercise 3:



Exercise 4:

Hide

```
data(present)
range(present$year)
```

```
## [1] 1940 2013
```

Hide

```
dim(present)
```

```
## [1] 74  3
```

Hide

```
names(present)
```

```
## [1] "year"  "boys"  "girls"
```

Exercise 5:

They are not similar in scale, the present data has significantly more births than the data from arbuthnot

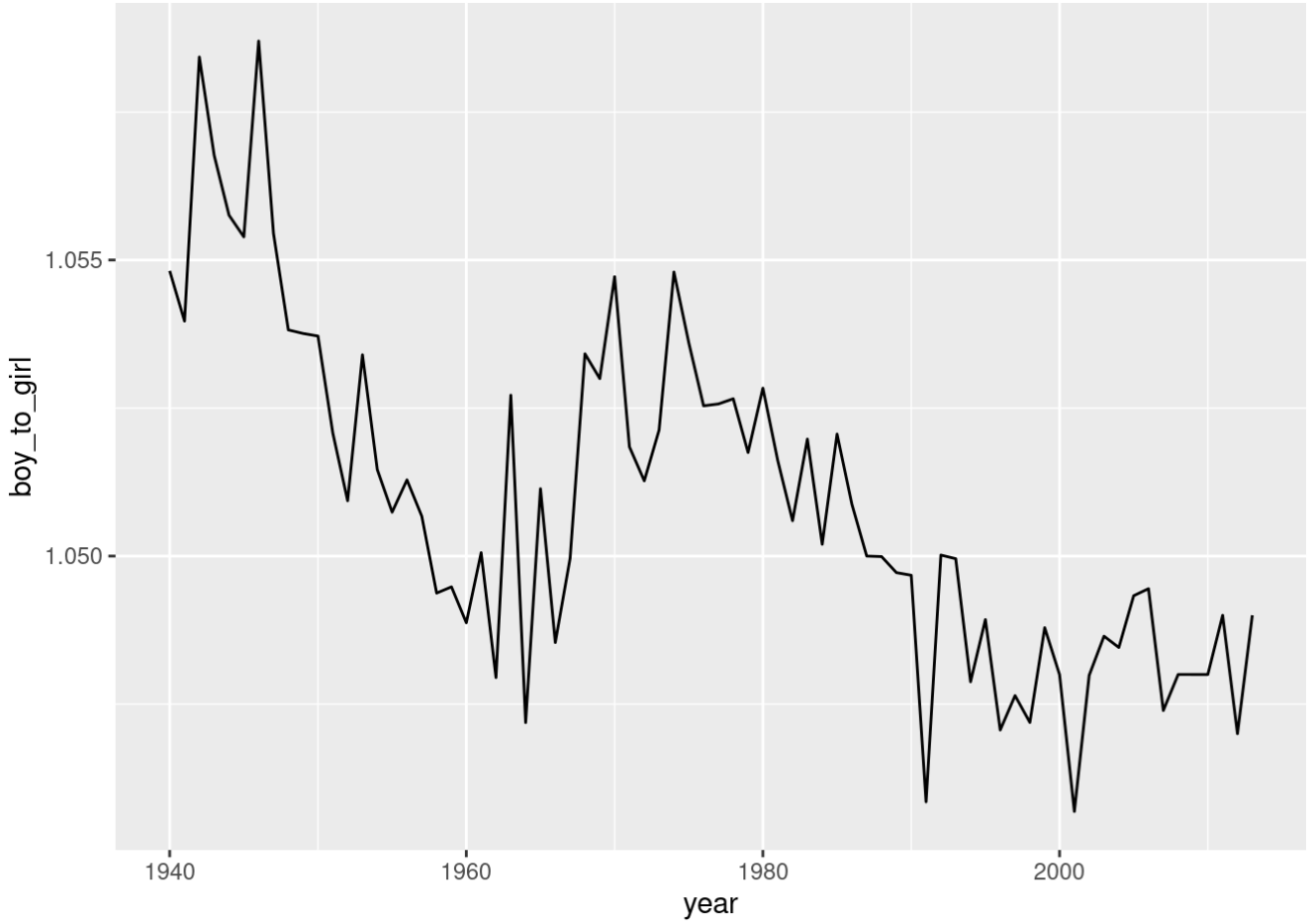
Exercise 6:

Hide

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

present <- present %>%
  mutate(boy_to_girl = boys / girls)

qplot(x = year, y = boy_to_girl, data = present, geom = "line")
```



In the present data we can see that the ratio of boys is greater than girls by a small amount, only a few present higher than girls

Exercise 7:

Hide

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

```
## # A tibble: 74 × 5
##     year    boys   girls   total boy_to_girl
##    <dbl>   <dbl>   <dbl>   <dbl>       <dbl>
##  1  2007 2208071 2108162 4316233        1.05
##  2  1961 2186274 2082052 4268326        1.05
##  3  2006 2184237 2081318 4265555        1.05
##  4  1960 2179708 2078142 4257850        1.05
##  5  1957 2179960 2074824 4254784        1.05
##  6  2008 2173625 2074069 4247694        1.05
##  7  1959 2173638 2071158 4244796        1.05
##  8  1958 2152546 2051266 4203812        1.05
##  9  1962 2132466 2034896 4167362        1.05
## 10  1956 2133588 2029502 4163090        1.05
## # … with 64 more rows
```

2007 had the most total number of births at 4,316,233 total births