Lab 2: Intro to R

Eric Fischer

8/8/2024

```
# Load the tidyverse and openintro packages, or libraries.

# Insert code for Exercise 0 here (already done for you)
library(tidyverse)
library(openintro)
```

Exercise 1

```
# Print the arbuthnot dataframe, available to us from the openintro package.
# Take a glimpse() of the arbuthnot dataframe.
# Print the girls column/feature/attribute of the arbuthnot dataframe.

# Insert code for Exercise 1 here
data("arbuthnot")
glimpse(arbuthnot)
```

```
## Rows: 82
## Columns: 3
## $ year  <int> 1629, 1630, 1631, 1632, 1633, 1634, 1635, 1636, 1637, 1638, 1639~
## $ boys  <int> 5218, 4858, 4422, 4994, 5158, 5035, 5106, 4917, 4703, 5359, 5366~
## $ girls <int> 4683, 4457, 4102, 4590, 4839, 4820, 4928, 4605, 4457, 4952, 4784~
```

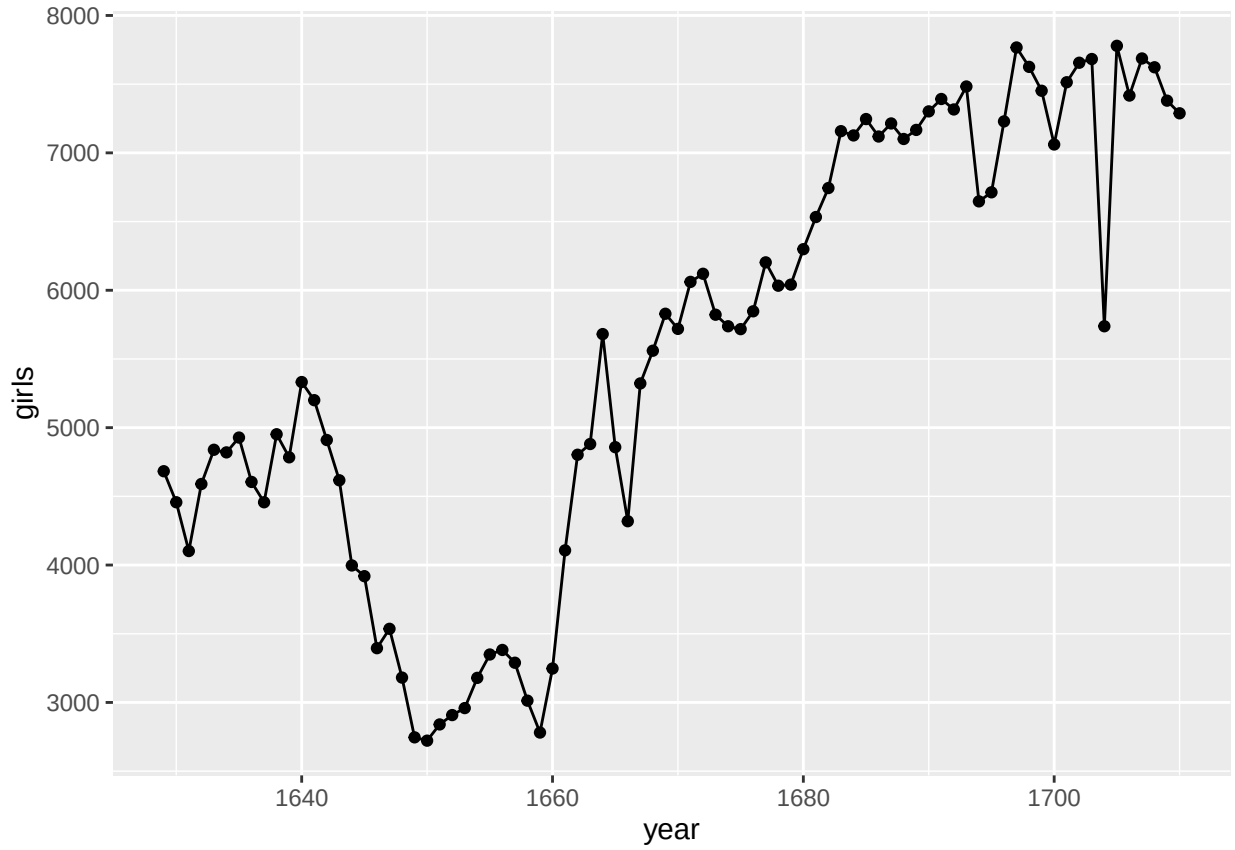
```
arbuthnot$girls
```

```
##  [1] 4683 4457 4102 4590 4839 4820 4928 4605 4457 4952 4784 5332 5200 4910 4617
## [16] 3997 3919 3395 3536 3181 2746 2722 2840 2908 2959 3179 3349 3382 3289 3013
## [31] 2781 3247 4107 4803 4881 5681 4858 4319 5322 5560 5829 5719 6061 6120 5822
## [46] 5738 5717 5847 6203 6033 6041 6299 6533 6744 7158 7127 7246 7119 7214 7101
## [61] 7167 7302 7392 7316 7483 6647 6713 7229 7767 7626 7452 7061 7514 7656 7683
## [76] 5738 7779 7417 7687 7623 7380 7288
```

Exercise 2

```
# Make a ggplot() using the arbuthnot dataframe, with year on the x-axis and girls on the y-axis.
# (Ensure the plot is a combined scatterplot and line graph.)
```

```r
# Insert code for Exercise 2 here
# geom_point() -> scatterplots
# geom_line() -> line graph
ggplot(data = arbuthnot, mapping = aes(x = year, y = girls)) +
  geom_point() +
  geom_line()
```



Exercise 3

```r
# Mutate() the arbuthnot dataframe in memory such that it has 2 new columns/features/attributes,
# total (boys + girls) and boy_ratio (boys / total).
# Make a line graph plot using the arbuthnot dataframe, with year on the x-axis and total on the y-axis
# Make a line graph plot using the arbuthnot dataframe, with year on the x-axis and boy_ratio on the y-

# Insert code for Exercise 3 here
arbuthnot = arbuthnot %>%
  mutate(total = boys + girls) %>%
  mutate(boy_ratio = boys / total)
```

Exercise 4

```r
# Print the unique() values of the year column/feature/attribute of the present dataframe (e.g., presen
# available to us from the openintro package.
# Print the dimensions of the present dataframe.
# Print the column names of the present dataframe.

# Insert code for Exercise 4 here
data("present")
unique(present$year)
```

```
##  [1] 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954
## [16] 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969
## [31] 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984
## [46] 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999
## [61] 2000 2001 2002
```

```r
dim(present)
```

```
## [1] 63  3
```

```r
colnames(present)
```

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

Exercise 5

```r
# Print the mean of the boys column/feature/attribute of the present dataframe, divided by
# the mean of the boys column/feature/attribute of the arbuthnot dataframe.
# Print the mean of the girls column/feature/attribute of the present dataframe, divided by
# the mean of the girls column/feature/attribute of the arbuthnot dataframe.

# Insert code for Exercise 5 here
```

Exercise 6

```r
# Mutate() the present dataframe in memory such that it has 2 new columns/features/attributes,
# total (boys + girls) and boy_ratio (boys / total).
# Make a line graph plot using the present dataframe, with year on the x-axis and total on the y-axis.
# Make a line graph plot using the present dataframe, with year on the x-axis and boy_ratio on the y-ax

# Insert code for Exercise 6 here
present = present %>% mutate(total = boys + girls)
```

Exercise 7

```r
# Arrange() in descending order using desc() the total column/feature/attribute of the present datafram
# and print the result.

# Insert code for Exercise 7 here
present %>% arrange(desc(total))
```

```
## # A tibble: 63 x 4
##     year    boys   girls   total
##    <dbl>   <dbl>   <dbl>   <dbl>
##  1  1961 2186274 2082052 4268326
##  2  1960 2179708 2078142 4257850
##  3  1957 2179960 2074824 4254784
##  4  1959 2173638 2071158 4244796
##  5  1958 2152546 2051266 4203812
##  6  1962 2132466 2034896 4167362
##  7  1956 2133588 2029502 4163090
##  8  1990 2129495 2028717 4158212
##  9  1991 2101518 2009389 4110907
## 10  1963 2101632 1996388 4098020
## # i 53 more rows
```

```r
# Knit (or generate) the R Markdown file into a PDF and submit both this .Rmd file and the PDF.
```