2/12/23, 8:21 PM Lab 1: Intro to R

Exercise 1 Exercise 2 Exercise 3 Exercise 4 Exercise 5 Exercise 6 Exercise 7

# Lab 1: Intro to R

Code ▼

Yina Qiao 2023-02-02

```
Hide
library(tidyverse)
library(openintro)
                                                                                     Hide
```

arbuthnot

```
## # A tibble: 82 × 3
##
    year boys girls
     <int> <int> <int>
##
   1 1629 5218 4683
##
   2 1630 4858 4457
   3 1631 4422 4102
##
   4 1632 4994 4590
   5 1633 5158
   6 1634 5035 4820
##
   7 1635 5106 4928
##
   8 1636 4917 4605
  9 1637 4703 4457
## 10 1638 5359 4952
## # ... with 72 more rows
```

Hide

```
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...
```

Hide

```
arbuthnot$boys
```

```
## [1] 5218 4858 4422 4994 5158 5035 5106 4917 4703 5359 5366 5518 5470 5460 4793
## [16] 4107 4047 3768 3796 3363 3079 2890 3231 3220 3196 3441 3655 3668 3396 3157
## [31] 3209 3724 4748 5216 5411 6041 5114 4678 5616 6073 6506 6278 6449 6443 6073
## [46] 6113 6058 6552 6423 6568 6247 6548 6822 6909 7577 7575 7484 7575 7737 7487
## [61] 7604 7909 7662 7602 7676 6985 7263 7632 8062 8426 7911 7578 8102 8031 7765
## [76] 6113 8366 7952 8379 8239 7840 7640
```

# Exercise 1

Hide

```
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

Insert any text here.

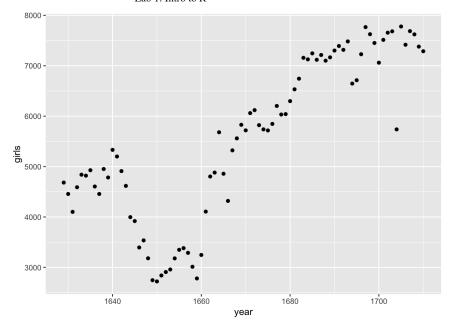
geom point()

```
Hide
ggplot(data = arbuthnot, aes(x = year, y = girls)) +
```

2/12/23, 8:21 PM

Exercise 1
Exercise 2
Exercise 3
Exercise 4
Exercise 5
Exercise 6
Exercise 7

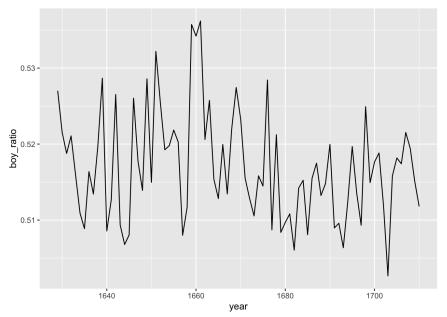




#### the plot above shown there is an increase in the number of girls being baptized from 1629 to 1710

# Exercise 3





#### it looks like propotions of boys baptized are dropping gradually overtime.

# Exercise 4

Hide data('present', package='openintro') glimpse(present)

2/12/23, 8:21 PM

```
Exercise 1
Exercise 2
Exercise 3
Exercise 4
Exercise 5
Exercise 6
Exercise 7
```

#### Lab 1: Intro to R

```
## Rows: 63
## Columns: 3
## $ year <dbl> 1940, 1941, 1942, 1943, 1944, 1945, 1946, 1947, 1948, 1949, 1950...
## $ boys <dbl> 1211684, 1289734, 1444365, 1508959, 1435301, 1404587, 1691220, 1...
## $ girls <dbl> 1148715, 1223693, 1364631, 1427901, 1359499, 1330869, 1597452, 1...
```

# Exercise 5

```
Hide
present<-present %>%
 {\tt mutate(total=boys+girls)}
head(present)
## # A tibble: 6 × 4
##
  year boys girls total
   <dbl> <dbl> <dbl> <dbl>
## 1 1940 1211684 1148715 2360399
## 2 1941 1289734 1223693 2513427
## 3 1942 1444365 1364631 2808996
## 4 1943 1508959 1427901 2936860
## 5 1944 1435301 1359499 2794800
## 6 1945 1404587 1330869 2735456
                                                                             Hide
head(arbuthnot)
## # A tibble: 6 × 6
   year boys girls total boys_to_girls_ratio boy_ratio
##
   <int> <int> <int> <int>
                                      <dbl>
                                                 <dbl>
## 1 1629 5218 4683 9901
                                       0.527
                                                 0.527
## 2 1630 4858 4457 9315
                                       0.522
                                                 0.522
## 3 1631 4422 4102 8524
                                       0.519
                                                 0.519
## 4 1632 4994 4590 9584
                                       0.521
                                                 0.521
## 5 1633 5158 4839 9997
                                        0.516
                                                 0.516
```

The values in present dataset are much bigger than arbuthnot by 3 orders of magnitude

0.511

0.511

# Exercise 6

## 6 1634 5035 4820 9855

```
Hide
present<-present %>%
 mutate(total=boys+girls)
present<-present %>%
 mutate(boys_to_girls_ratio=boys/girls)
present<-present %>%
 mutate(boys_ratio=boys/total)
ggplot(data = present, aes(x=year,y=boys_ratio))+geom_line()
```

2/12/23, 8:21 PM

Exercise 7

Exercise 1

Exercise 2

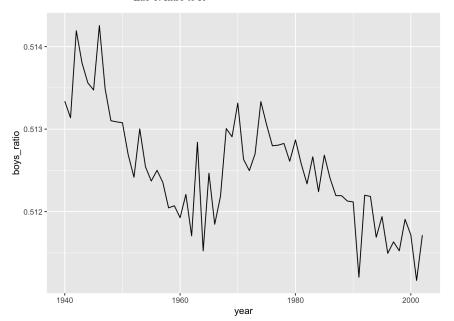
Exercise 3

Exercise 4

Exercise 5

Exercise 6

### Lab 1: Intro to R



#### the boy\_ratio is decreasing over time for the present data, different from Arbuthnot's observation.

# Exercise 7

Hide present %>% arrange(desc(total)) ## # A tibble: 63 × 6 ## year boys girls total boys\_to\_girls\_ratio boys\_ratio <dbl> <dbl> ## <dbl> <dbl> <dbl><dbl>1 1961 2186274 2082052 4268326 1.05 0.512 ## 1960 2179708 2078142 4257850 1.05 0.512 1957 2179960 2074824 4254784 1.05 0.512 1959 2173638 2071158 4244796 1.05 0.512 1958 2152546 2051266 4203812 1.05 0.512 1962 2132466 2034896 4167362 1.05 0.512 1956 2133588 2029502 4163090 1.05 0.513 1990 2129495 2028717 4158212 1.05 0.512 9 1991 2101518 2009389 4110907 1.05 0.511 ## 10 1963 2101632 1996388 4098020 1.05 0.513 ## # ... with 53 more rows