Lab 1: Intro to R

VIctor H Torres

r Sys.Date()

{r load-packages, message=FALSE} library(tidyverse) library(openintro) data('arbuthnot',
package='openintro')

Exercise 1

{r view-girls-counts} arbuthnot\$girls

Exercise 2

There is a significant decrease of girls baptized from 1620 to 1660. After that year, the amount of girls baptized increase drastically until the 1700's, where we can some reduction of the number of baptisms. The graphic below shows the changes on the amount of females baptisms over the years.

```
{r trend-girls} ggplot(data = arbuthnot, aes(x = year, y = girls)) + geom_line() +
geom_point( colour = 'red' )
```

Exercise 3

With the new variable created(total) I am able to generate a linear chart, showing the proportion of boys born over time displaying a increase pattern over the years.

```
{r plot-prop-boys-arbuthnot} arbuthnot <- arbuthnot %>% mutate(total = boys + girls)
ggplot(data = arbuthnot, aes(x = boys , y = total)) + geom_line()
```

Exercise 4

The years included in this data ranges from 1940 to 2002. The dimensions of the data frame are 63 observations and 3 variables There are three variables: Year, Boys, and Girls

```
{r dim-present} data('present', package='openintro') present %>% summarize(min = min(boys),
max = max(boys))
```

Exercise 5

The range of years are bigger on the Arbuthnot dataframe, also the dimmensions are bigger, it contains more rows,

```
{r count-compare} arbuthnot <- present %>% summarize(min = min(boys), max = max(boys))
```

Exercise 6

The Proportion of boys being born in greater proportion than girls in the U.S changes from the first data frame to the second one, an increase of girls is evident in the charts.

```
{r plot-prop-boys-present} present <- present %>% mutate(total = boys + girls) ggplot(data = present, aes(x = boys , y = total)) + geom_line()
```

Exercise 7

The most total number of births in the U.S was in 1961, with a total of 4268326.

```
{r find-max-total} present %>% arrange(desc(total))
```