**Week 1 Lab: Introduction to R and RStudio**

1.How many variables are included in this data set (data set: arbuthnot)?

3

2.What command would you use to extract just the counts of girls born?

arbuthnot$girls

3.Which of the following best describes the number of girls baptised over the years included in this dataset?

There is initially an increase in the number of girls baptised, which peaks around 1640. After 1640 there is a decrease in the number of girls baptised, but the number begins to increase again in 1660. Overall the trend is an increase in the number of girls baptised.

4.How many variables are included in this data set (data set: present)?

3

5.Calculate the total number of births for each year and store these values in a new variable called total in the present dataset. Then, calculate the proportion of boys born each year and store these values in a new variable called prop\_boys in the same dataset. Plot these values over time and based on the plot determine if the following statement is true or false: The proportion of boys born in the US has decreased over time.

True

6.Create a new variable called more\_boys which contains the value of either TRUE if that year had more boys than girls, or FALSE if that year did not. Based on this variable which of the following statements is true?

Every year there are more boys born than girls.

7.Calculate the boy-to-girl ratio each year, and store these values in a new variable called prop\_boy\_girl in the present dataset. Plot these values over time. Which of the following best describes the trend?

There is initially a decrease in the boy-to-girl ratio, and then an increase between 1960 and 1970, followed by a decrease.

8.In what year did we see the most total number of births in the U.S.?

2007