

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

- a) 1710
- b) 82
- c) 2
- d) 4
- e) 3

Ans: e

Explanation: There are total three columns/variables ["year", "boys", "girls"]

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

- a) girls
- b) \$girls
- c) arbuthnot\$girls
- d) arbuthnot\$boys
- e) arbuthnot[girls]

Ans: c

Explanation: The "\$" sign fetches the column of the given dataset in R.

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

- a) 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.
- b) There is initially an increase in the number of girls baptised. This number peaks around 1640 and then after 1640 the number of girls baptised decreases.
- c) There appears to be no trend in the number of girls baptised from 1629 to 1710
- d) The number of girls baptised has decreased over time.
- e) There is an initial increase in the number of girls baptised but this number appears to level around 1680 and not change after that time point.

Ans: a

Explanation: See the file Week1\_Solutions.html which contains a picture/plot depicting the given information.

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

- a) 2013
- b) 4
- c) 3
- d) 74
- e) 2

Ans: 3

Explanation: There are total of three variables [“year”, “boys”, “girls”]

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.

- a) True
- b) False

Ans: a

Explanation: Refer to Week1\_Solutions.html

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?

- a) Half of the years there are more boys born, and the other half more girls born.
- b) Every year there are more girls born than boys.
- c) Every year there are more boys born than girls.

Ans: c

Explanation: Refer to Week1\_Solutions.html

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?

- a) The boy-to-girl ratio has increased over time.
- b) There appears to be no trend in the boy-to-girl ratio from 1940 to 2013.
- c) There is an initial decrease in the boy-to-girl ratio but this number appears to level around 1960 and remain constant since then.
- d) There is initially an increase in boy-to-girl ratio, which peaks around 1960. After 1960 there is a decrease in the boy-to-girl ratio, but the number begins to increase in the mid 1970s.
- e) There is initially a decrease in the boy-to-girl ratio, and then an increase between 1960 and 1970, followed by a decrease.

Ans: e

Explanation: Refer to [Week1\\_Solutions.html](#)

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

- a) 2007
- b) 1961
- c) 1940
- d) 1991
- e) 1957

Ans: a

Explanation: Refer to [Week1\\_Solutions.html](#)