Lab 0

Max Wagner

August 30, 2015

---

1. What years are included in this data set? What are the dimensions of the data frame and what are the variable or column names?

    - The dimensions of the data set are **63, 3**. The column names are **year, boys, girls**, and the years included are:

```
##  [1] 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953
## [15] 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967
## [29] 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981
## [43] 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995
## [57] 1996 1997 1998 1999 2000 2001 2002
```

---

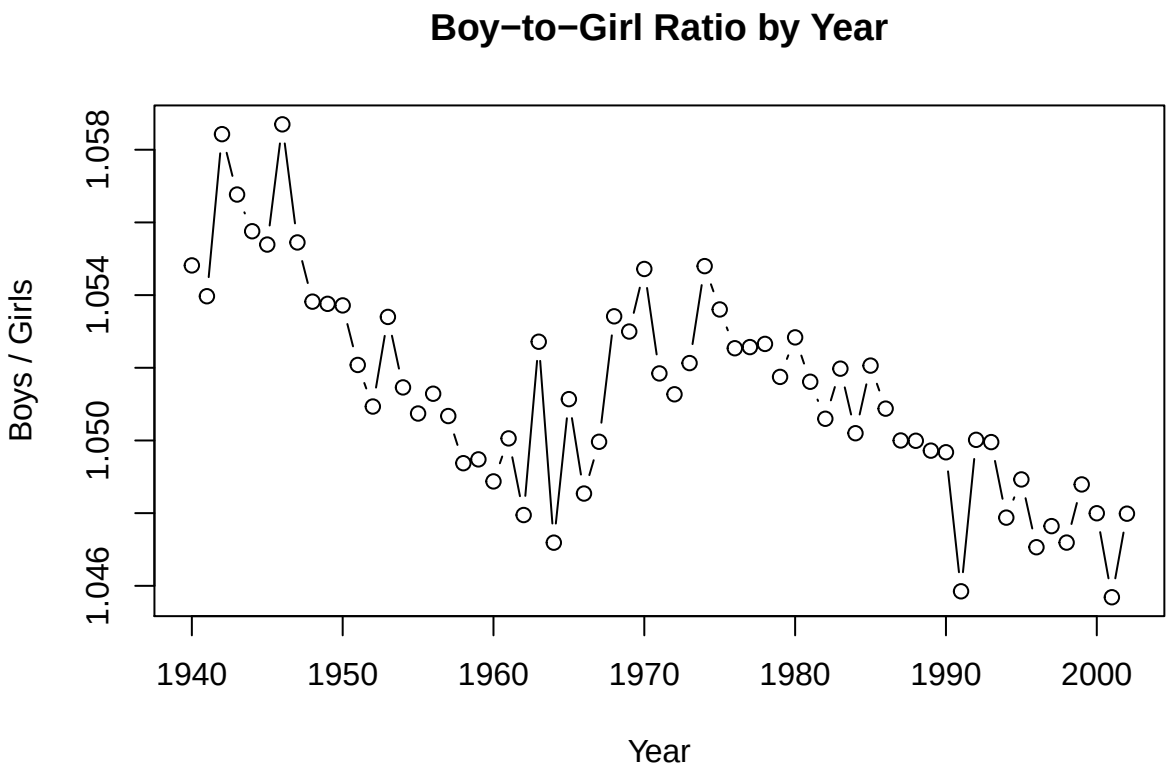
2. How do these counts compare to Arbuthnot's? Are they on a similar scale?

    - The count of years in present is **63** while the number of years in the arbuthnot set is **82**, which makes the present set **0.77%** of the arbuthnot set's size. The column names are the same in both, **year, boys, girls**. The dimesions of present are **63, 3** versus **82, 3** for arbuthnot, which is caused by the difference in years included.

---

3. Make a plot that displays the boy-to-girl ratio for every year in the data set. What do you see? Does Arbuthnot's observation about boys being born in greater proportion than girls hold up in the U.S.? Include the plot in your response.

    - The average boy-to-girl ratio in the present set is **1.05** compared to **1.07** in the arbuthnot set. The observation seems to still be valid.



---

4. In what year did we see the most total number of births in the U.S.? You can refer to the help files or the R reference card http://cran.r-project.org/doc/contrib/Short-refcard.pdf to find helpful commands.

   - The highest number of births occured in **1961** with a total of **4,268,326**.