Stat 110 Lab #4

Finding Life in the North Carolina Birth Data

Joe Roith

September 27, 2019

In 2004, the state of North Carolina released to the public a large data set containing information on births recorded in the state. Medical researchers who are studying the relation between habits and practices of expectant mothers and th birth of their children were keenly interested in this data. The SPSS data set `ncbirth145types.sav` contains a random ample of 1450 cases and a subset of variables; you can find this file on Moodle. Today in lab we are going to use graphs and summary statistics discussed in Chapters 7-9 to explore the North Carolina birth data.

By the end of this lab you should be able to generate and describe relevant features of the following plots:

- Histogram (for on measurement variable)
- Barplot (for one categorical variable)
- Conditional barplot (for two categorical variables)
- Boxplots (for one categorical and one measurement variable)
- Scatterplot (for two measurement variables)

Consider the following plots and sets of summary statistics (note that I picked sample variables for illustration purposes only). Remember - in SPSS, you have tons of flexibility to odify the graphs to look exactly like you want. You must first enter `Chart Editor` by double clicking on an *existing* graph in the Output Window. The plots for today's lab should have clear and meaningful axis labels and titles.

Histogram with superimposed normal curve (for one measurement variable)

- Once in Chart Builder (under the `Graph` drop-down menu), choose `Histogram` under `Gallery`
- Create a histogram as in Lab #1
- Double click on chart in Output Window to enter the Chart Editor
- Under `Elements`, choose `Show Distribution Curve` or click the icon in the upper right that is shaped like a bell curve  .
- Normal is the default, so click close

1. Pick one measurement variable of interest to you

- Generate a histogram with a superimposed normal curve. Is the normal model a good one for the variable?
- Write one sentance commenting on trends you observe from your plot (CUSS)

Summary statistics by group (for one measurement and one categorical variable)

- Click `Analyze` > `Descriptive Statistics` > `Explore`
- Place *tounces* in the `Dependent List` and *marital* in the `Factor List`
- Under `Statistics` click `Despriptives` and `Percentiles`
- Under `Plots` you'll want `Boxplots` (`Factor levels together`); you may or may not want other plots

2. Pick a combination of one measrement and one categorical variable

- Which would you consider *explanatory* and which would you consider *response* (or are they interchangeable)?
- Generate a boxplot with groups side-by-side
- Report summary statistics by group
  - Mean and standard deviation
  - Median and IQR
  - Five number summary
- Write one sentance comparing the two groups, incorporating some of your summary statistics