

Stat 110 Lab #2

Introduction to SPSS

Joe Roith

September 13, 2019

Today we are going to explore data collected from 173 college students at UC Davis in two different statistics classes. One class was for students not in the liberal arts (n=148), while the other class was for students in the liberal arts (n=25). The data were collected in the Spring quarter of 2000. (Source: Utts and Heckard). We will be using graphical and numerical summaries in the software package PASW Statistics (formerly and commonly called SPSS). The instructions should work for either Windows or Mac with minor modification. Let me know if you see anything confusing or get stuck at any point.

Name	Description
Sex	Male or Female
TV	hours spent watching tv each week
computer	hours spent on a computer each week
Sleep	hours of sleep previous night
Seat	typical classroom seat (front, middle, back)
alcohol	number of alcoholic beverages consumed each week
Height	self-reported height (inches)
momheight	mother's height (inches)
dadheight	father's height (inches)
exercise	hours spent exercising each week
GPA	student GPA
class	liberal arts or not liberal arts

Reading in the Data

First some preliminaries. we will read in the data, look it over, and make sure it is usable by SPSS.

1. Download the data from Moodle. Pay careful attention to where the file is saved!
2. Open SPSS Statistics. If you cancel the initial dialog box, you will see the Data Editor. This should look like an (empty) excel spreadsheet.
3. Bring in our data using **File > Open > Data > UCDavis.xlsx** (after changing **File Type** to **.xlsx**). Click OK to the dialog box so that SPSS uses its default choices for importing Excel data. Note that the spreadsheet has a **Data View** (where you can see the raw data), and a **Variable View** (where you can see properties of each variable in the data set).
4. In **Variable View**, we must make sure to define the Measure of each variable. **Scale** for a measurement/quantitative variable, **Nominal** for a categorical variable, and **Ordinal** if the categories have a natural order to them. For this data, the measures of all the variables are already correct.
5. After cleaning and defining your data, it can be saved as an SPSS data set for future use. Go to **File > Save As** and type in a filename (which will be given a **.sav** extension).

Now it's time to examine this data! In statistics, we always begin by visualizing our data using appropriate and insightful graphs. The best graph for a situation depends on the number of variables being examined (1 or 2 for now) and the type of variables (measurement/quantitative or categorical).

For each situation below, begin by choosing **Chart Builder** under **Graphs**.

Graphs that explore a single variable

1. *Histogram*: One measurement/quantitative variable
 - Choose **Histogram** under Gallery.
 - Drag the first (left) image into the Chart Preview box above.
 - Drag **Height** from the Variables list onto the x-axis (horizontal). The the y-axis (vertical) be the default of counts per interval and hit OK.
 - In the Output Window , you will get a histogram of **Height** with a few summary statistics.
2. *Bar plot*: One categorical variable.
 - Choose **Bar** under Gallery.
 - Drag the first (upper left) image into the Chart Preview box above.
 - Drag **Seat** from the Variables list onto the x-axis. The the y-axis be the default of counts per category and hit OK.
 - In the Output Window , you will get a barplot of **Seat**.

Graphs that explore the relationship between a pair of variables

1. *Scatterplot*: Two measurement variables.
 - Choose **Scatter/Dot** under Gallery.
 - Drag the first (upper left) image into the Preview box above.
 - Drag **dadheight** from the Variables list onto the x-axis and **Height** onto the y-axis.
 - In the Output Window, you will get a scatterplot of **Height** vs **dadheight**.
2. *Conditional barplot*: Two categorical variables
 - Choose **Bar** under Gallery.
 - Drag the 3rd image from the left in the top row of the Chart Preview box above.
 - Drag **Class** from the Variables list into the Stack box and **Sex** onto the x-axis.
 - Under Element Properties, set **Statistic** = **Percentage**, and under Set Parameters choose **Denominator** = **total** for each category.
 - Hit Apply.
 - In the Output Window, you will get a conditional barplot of **Class** by **Sex**.
3. *Side by side boxplots*: One measurement variable and one categorical variable.
 - Choose **Boxplot** under Gallery.
 - drag the first (left) image into the Chart Preview box above.
 - Drag **Height** from the Variables list onto the y-axis and **Seat** onto the x-axis.
 - Take note of which variable type goes where.
 - In the Output Window, you will get a boxplot of **Height** by **Seat**.

In SPSS, you have tons of flexibility to modify the graphs to look exactly like you want. You must first enter Chart Editor by double clicking on your graph in the Output Window. For instance, the boxplot of **Height** by **Seat** could use some work...

- To change the range of the y-axis: Double click on the y-axis, choose Scale, and set the Maximum to something reasonable (you can play with a few options here).
- To change a label: After double clicking on the y-axis, click on the label **Height** and you can type in your changes, say “Respondent Height in inches”.
- To change your bar labels on the x-axis: Double click on the x-axis, click **Seat** and type in your new label, say “Seat in Class”. this can also be done by going to the data itself, and using the Value column in Variable View. You can also re-order your categories!
- Add titles, textboxes, etc.: Click the appropriate button in Chart Editor.
- Add a fitted line to a scatterplot: Click on the appropriate button in Chart Editor, or select Add Fit Line under Elements.

Lab Assignment

1. Come up with at least two interesting questions that you can investigate with variables from this data.
2. For each question use one of the plots explored in lab to graph the appropriate pair of variables to address the question.

As one example, you could use boxplots motivated by the question “Do taller students prefer to sit in the back?” Each person should include two-nice-looking, nicely labeled SPSS plots with a corresponding question, followed by a couple questions stating what you learned about this class at UC Davis in their Homework submission due **Monday, Sept 16, 5:00 PM**. You can also include any comments you have on data collection.

Be creative! Be artistic! Have fun!

Note that you can save your output as a .spv file, but it is often easier to cut and paste graphs into Word documents. You can also take a screenshot of your visualizations (snipping tools in Windows and CMD + SHIFT + 4 on Mac). SPSS does not play well with Google. But once you have a Word document you should be able to upload that to Google and protect the figures.