
Sociology 350

Your Name: _____

Lab# 1

Subject: Interpreting Correlations and partial-order correlations

1. Enter the **SPSS** program
2. Load the **GSS93subset** file from the list of data files contained within the SPSS program
3. Next, left click the "*Variable View*" menu option from the bottom left-hand corner of your screen.
4. After you have done this, left click the far right-hand side of the "*Values*" column that corresponds to the variable name you want to work with. For example, look at the variable "sex". You should see three consecutive dots [...] in this location of the box. Once you have done this, you will have gained access to the dialog box that will allow you to make any changes in *Value labels* that you desire to make.
5. Today we will only "**Recode**" a few variables. Under the "**Transform**" main menu header of SPSS select "**Recode**" first, then select "**Into Different Variables**". Next, choose the variable named "**educ**" to recode. Then type in the new variable name as "**neweduc**". Then select "**Old and New Variables**" as follows--- **1 through 95 = same; 96 through 99 = Sysmis**. End all SPSS commands with "**OK**".
6. Next "**Recode**" the variable named "**race**" with "**Into Different Variables**" named "**Black**". Now select "**Old and New Values**" so that **1=1, 2=2, and All Else = Sysmis**.
7. By reading these "**Value labels**" and only by reading these will you be able to determine how to interpret a particular correlation coefficient. For example, the variable named "**attsprts**" measures whether or not the respondent attends sporting events but you cannot tell whether the low number code (usually "1") represents "no" or whether the high number code (usually "2") represents "no" unless you inspect the "*Value labels*".

Do not forget this invaluable piece of information; you will have frequent occasion to use both today and throughout the semester!

8. Now select the **Analyze** function off the SPSS main menu. Next, select the **correlate** option off the statistics menu. Next, select the **bivariate** option off the correlation menu. Next select the variable names you desire and move them into the dialog box with the **Variables** heading. Now have the computer generate correlation coefficients for the variable named “**income91**” and **each** of the other variables listed below. Then interpret the meaning of **each** of these correlations **in a single sentence** in the space provided.

Be sure to include the sign (+ or -) of the correlation coefficient and be sure to consider this sign as you proceed to interpret your correlations.

<i>r</i> <u>correlation coefficient</u>	<i>Significance Level</i> <u>p value</u>	<u>Write Meaning of “r”</u>
---	--	--

rincom91

educ

sex

age

Which of the above variables is the *strongest predictor* of a person’s family income?

Which of the above variable is the *weakest predictor* of a person’s family income?

9. Once again select the **Analyze** function off the SPSS main menu. And select the **correlate** option off the **statistics** menu. Next, select the **bivariate** option off the correlation menu. Next select the variable names you desire and move them into the dialog box with the **Variables** heading. Now make the computer generate correlation coefficients for the variable named “**childs**” with **each** of the other variables listed below and then interpret the meaning of **each** of these correlations in a single sentence in the space provided.

Be sure to include the sign (+ or -) of the correlation coefficient and be sure to consider this sign as your proceed to interpret your correlations.

<i>r</i> <u>correlation coefficient</u>	<i>Significance Level</i> <u>p value</u>	<u>Write Meaning of “r”</u>
---	--	--

age

degree

Income91

sex

Which of the above variables is the strongest predictor of how many children a person has?

Which of the above variable is the weakest predictor of how many children a person has?

10. Again select the **Analyze** function off the SPSS main menu. And select the **correlate** option off the **statistics** menu. Next, select the **bivariate** option off the correlation menu. Next select the variable names you desire and move them into the dialog box with the **Variables** heading. Now instruct the computer to generate **correlation coefficients** for the variable named “**neweduc**” (the dependent variable) and **each** of the other variables listed below and then interpret the meaning of **each** of these correlations in a single sentence in the space provided

Be sure to include the sign (+ or -) of the correlation coefficient and be sure to consider this sign as you proceed to interpret your correlations.

<i>r</i> <u>correlation coefficient</u>	<i>Significance Level</i> <u>p value</u>	<u>Write Meaning of “r”</u>
---	--	--

agewed

attsprts

tvnews

tvpbs

Which of the above variables is the **strongest predictor** of how much education a person has?

Which of the above variable is the **weakest predictor** of how much education a person has?

11. Once again select the **Analyze** function off the SPSS main menu. And select the **correlate** option off the **statistics** menu. Next, select the **bivariate** option off the correlation menu. Next select the variable names you desire and move them into the dialog box with the **Variables** heading. Now instruct the computer to generate **correlation coefficients** for the variable named “**partners**” and **each** of the other variables listed below and then interpret the meaning of **each** of these correlations in a single sentence in the space provided.

Be sure to include the sign (+ or -) of the correlation coefficient and be sure to consider this sign as you proceed to interpret your correlations.

<i>r</i> <u>correlation coefficient</u>	<i>Significance Level</i> <u>p value</u>	<u>Write Meaning of “r”</u>
---	--	--

age

sex

Income91

sexfreq

Which of the above variables is the **strongest predictor** of how many sex partners a person has?

Which of the above variable is the **weakest predictor** of how many sex partners a person has?

12. Now you are asked to calculate a **bivariate correlation** between the recoded race variable (**black**) and the person's "highest year of school completed" recoded as "**neweduc**". Once you have completed this task write your findings below:

<i>r</i> correlation coefficient	<i>Significance Level</i> p value	Write Meaning of "r"
--	---	---------------------------------

Next you are going to introduce a sequence of control variables to see **how low** you can reduce the statistically significant correlation that you just calculated between race and educational attainment. Introduce each of the independent variables listed below as control variables beginning with a single control variable, "sibs". Then add a second control variable "faminc" in addition to "sibs". Then add a third control variable, father's highest degree" in addition to the first two control variables. And keep adding additional control variables as they are listed below without removing any until you have completed all the control variable groups. After each set of control variables write the resulting "partial-order correlation" coefficient and its significance level in the space provided under it.

-
- a) **sibs (number of brothers and sisters)---** only
- b) **faminc (total family income)---** plus a)
- c) **padeg (father's highest degree) ---** plus a) and b)
- d) **madeg (mother's highest degree)---** plus a), b) and c)
- e) **classic3 (classical music) ---** plus a), b), c), and d)

What conclusion can you draw about the correlation between race and the highest year of school completed based on this set of control variables?

13. Now you are asked to calculate the bivariate correlation between the recoded race variable (**black**) and “liking Rap music” (**rap**). Once you have completed this task write your findings below:

r
correlation
coefficient

Significance
Level
p value

Write
Meaning of “r”

As before you are going to introduce a sequence of control variables to assess how much of the correlation that you just calculated between race and liking rap music can be eliminated by adding these control variables. Introduce each of the independent variables listed below as control variables beginning with a single control variable, “sibs”. Then add a second control variable “faminc” in addition to “sibs”. Then add a third control variable, father’s highest degree” in addition to the first two control variables. And keep adding additional control variables as they are listed below without removing any until you have completed all the control variable groups. After each set of control variables write the resulting “partial-order correlation” coefficient and its significance level in the space provided under it.

a) **sibs (number of brothers and sisters) --- only**

b) **news (how often does R read newspapers) --- plus a)**

c) **polviews (thinks of self as liberal or conservative --- plus a) and b)**

d) **sex (respondent’s sex) --- plus a), b), and c)**

e) **rincom91 (respondent’s income) --- plus a), b), c), and d)**

What conclusion can you draw about the correlation between race and the person’s likelihood of liking “rap music” based on this set of control variables?

14. Finally, your job is to determine exactly how it is that “**age**” relates to personal preferences in “**musical styles**”, specifically, you want to determine if older or younger persons have a relatively higher preference for “**rap**”, for “**big band**”, for “**country**”, for “**blues**”, for “**musicals**”, for “**classical**”, for “**folk**”, for “**jazz**”, for “**opera**”, and for “**heavy metal**”. To test this you will have to run bivariate correlations with the variable of “respondent’s age” (**age**). Run bivariate correlations between age and the variety of musical styles. Then answer the following questions:
- a. Which three styles of music are most popular with the **older persons** in the US population? List them below in their relative order popularity starting with the #1 most popular.
- 1.
 - 2.
 - 3.
- b. Which three styles of music are most popular with the **younger persons** in the US population? List them below in their relative order of most popular starting with the #1 most popular.
- 1.
 - 2.
 - 3.
- c. And for which style of music is there absolutely **no significant relationship** between the “respondent’s age” and their preference for that style of music? **What explanation do you have for why this style of music bears no relation to age?**

Congratulations! You have just cleared another pivotal hurdle in your life. You may now proceed to the next one!