PSYCHOLOGY 203

INTRODUCTION TO R

Before doing any analyses, we need to get some data read into a file that we can use. For the next 2 lab classes, we will be working on entering data and performing some simple operations on the variables and file.

1. Create a file on your computer for this course and use the setwd() function to set that file as the working directory. This is where all of the files you want to use or save for this course will be stored and drawn from.

2. Use the data we collected during the first day of class to create a data frame in R using the read.csv() function or by going to the Import Dataset drop down menu in the Environment window and selecting the "From Text (base)" option. The raw data file is available on Google Classroom as a CSV file. You should use this file to create a data frame that contains all of the responses for all of the items.

3. Recode the data for the MOVIE TRIVIA items (items A1 through A15) so that each response is either 0 (incorrect answer) or 1 (correct response) using the following answer key:

    - Correct Answer A: Items = A4, A7, A11, A13
    - Correct Answer B: Items = A3, A12, A14
    - Correct Answer C: Items = A1, A5, A6, A8, A10, A15
    - Correct Answer D: Items = A2, A9

4. Create a new variable that is the sum of all the MOVIE TRIVIA items. This is the MOVIE TRIVIA scale score variable.

5. Recode the following items from SECTION B (the STATS EFFICACY items) such that (1=5) (2=4) (4=2) (5=1):

    - Items B3, B4, B8, B10, B12

6. Create a new variable that is the average of all of the STATS EFFICACY items (items B1 through B13). Be sure to include the recoded items, NOT the original items, when computing this variable. This is the STATS EFFICACY scale score variable.

7. Save the modified data frame with the new variables to your working directory under a new file name using the write.csv() function so you can easily access it later. This function will save it to the working directory you set earlier. Be sure to give it a unique name as it will overwrite a file if the name is already in use in your working directory.

PSYCHOLOGY 203
Project 1

DUE: February 13, 2019 @ 10 AM
Points (20)

You should use the data file we have been working with for the last couple of weeks.

Before you do anything for this project, predict the results you will obtain.

1. What do you think the MOVIE TRIVIA scale scores and the STATS EFFICACY scale scores will look like? Specifically, do you think the distribution for each variable will be normal, skewed, etc.? Why? BE SURE YOU DO THIS FOR BOTH SCALE SCORE VARIABLES.

Now, use R to create a grouped frequency distribution (histogram) for the MOVIE TRIVIA scale scores and for the STATS EFFICACY scale scores.

2. Provide a detailed interpretation of each of the grouped frequency distributions.

3. Did the results of the analysis confirm your predictions?

4. What raw score on each variable corresponds to the 50^{th} percentile?

Create an ungrouped frequency distribution (histogram) for the MOVIE TRIVIA scale scores and for the STATS EFFICACY scale scores.

5. Are these distributions easier or more difficult to interpret than the grouped distribution? What would you recommend to someone who wanted to understand the distribution of scores? Why?

PSYCHOLOGY 203
Project 2

DUE: February 20, 2019 @ 10 AM
Points (20)

You should use the data file you used for Project 1.

Compute descriptive statistics for the MOVIE TRIVIA scale scores and for the STATS EFFICACY scale scores.

1. What is the mean, median, and mode for each of the variables?

2. What is the range, standard deviation, and variance for each of the variables?

3. If you were only able to report one measure of central tendency and one measure of variability for these variables, what would you use? Why?

4. If you were to **randomly** drop one case from your data, what do you think would happen to the mean and standard deviation for the MOVIE TRIVIA variable?

5. Choose one case (at random) and drop it. Recompute the mean and standard deviation for the remaining cases. What do you observe? Was your prediction in #4 correct? Explain what you see.

PSYCHOLOGY 203
Project 3

DUE: February 27, 2019 @ 10 AM
Points (20)

Part 1

You should use the data file you used for Projects 1 & 2.

1. Recode the month in which each individual was born so that those born in January through April are coded 1, those born between May and August are coded 2, and those born in September through December as 3.

2. For both the MOVIE TRIVIA scale scores and STATS EFFICACY scale scores variables, illustrate whether there are differences across the recoded birth month variable. You can illustrate this in any way you like, but you must efficiently and effectively communicate what you observe along with an interpretation of what the data tell you.

Part 2

Use the datafile labeled PROJ3.csv available on Blackboard.

These data include information from 15 hypothetical students as well as hypothetical quiz scores across two different sections of a class.

3. Calculate the average Exam score for PSYC students and the average Exam score for NONPSYC students. Create a figure that clearly and accurately depicts these scores. In addition, create a table that clearly and accurately communicates what you observed.

4. Create a visual representation of how the average quiz score changed across time for each of the two hypothetical class sections.

PSYCHOLOGY 203
Project 4

DUE: March 6, 2019 @ 10 AM
Points (30)

For this project, you should answer the odd numbered questions at the end of Chapter 7 in the course textbook. These questions are on pp. 255-256.

PSYCHOLOGY 203
Project 5

DUE: March 13, 2019 @ 10 AM
Points (30)

You should use the data file you used for Projects 1 & 2.

1. What will happen to the mean and standard deviation if you add 3 to every person's MOVIE TRIVIA scale score? What will happen to the mean and standard deviation if you multiple every person's MOVIE TRIVIA scale score by 3?

2. Add 3 to every person's MOVIE TRIVIA scale score and compute the mean and standard deviation. Do the results conform to what you expected?

3. Multiply every person's MOVIE TRIVIA scale score by 3 and compute the mean and standard deviation. Do the results conform to what you expected?

4. Redo items 1-3 for the STATS EFFICACY scale score variable.

5. Provide a visual representation of the original variables and the transformed values.

PSYCHOLOGY 203
Project 6

DUE: March 27, 2019 @ 10 AM
Points (30)

For this project, you should answer the odd numbered questions at the end of Chapter 10 in the course textbook. These questions are on pp. 358-359.

PSYCHOLOGY 203
Project 7

DUE: April 10, 2019 @ 10 AM
Points (30)

For this project, you should answer the odd numbered questions at the end of Chapter 12 in the course textbook. These questions are on pp. 426-427.

PSYCHOLOGY 203
Project 8

DUE: April 17, 2019 @ 10 AM
Points (35)

You should use the data file labeled PROJ8 on Blackboard.

There should be 40 cases and 3 variables: *chapters*, *attend*, and *grade*. *Chapters* represents the number of chapters read by students taking a particular statistics course. *Attend* is the number of lectures each student attended. *Grade* reflects each student's final grade in the course.

1. For each of the three possible relations between these variables, predict whether the correlations will be strong and positive, strong and negative, weak and positive, weak and negative, or near zero. **Provide a short explanation to support your predictions.**

Produce and report the three scatterplots for each of the potential relations.

2. Based on the scatterplots, is there support for your three predictions? Explain your answer.

Compute the correlations between these three variables.

3. Report the correlations in a **correlation matrix formatted in APA style**.

4. Based on the computed correlations, evaluate your predictions. Were they supported? Why? Justify your answer.

5. How much variance in grades is explained by course attendance? Provide an explanation of this value for someone who is not familiar with statistics.

Part 2

Find some data (with at least 100 cases) that permits computing a correlation.

6. Describe the variables. What are they? How are they coded?

7. What is your prediction of the relationship between these variables? What do you expect to see?

8. Compute and report the correlation you observe in these data. Was your prediction supported? How much variance was explained?

9. What are the practical implications of this relationship? If you collected additional data, do you think the relationship would replicate? Why or why not?

PSYCHOLOGY 203
Project 9

DUE: April 24, 2019 @ 10 AM
Points (35)

You should use the data file labeled PROJ9 on Blackboard.

There should be 130 cases and 3 variables: *bodytemp*, *zombie*, and *heartrate*. *Bodytemp* represents the individual's temperature (in °F). *Zombie* is coded as 1 = non-zombie and 2 = zombie. *Heartrate* reflects the resting HR (in beats per minute).

Imagine that you are interested in whether zombie status or body temperature can predict someone's resting heart rate.

1. Do you think body temperature will predict heart rate? Why or why not? If you think there will be a relation, will it be positive or negative? Do you think zombie status will predict heart rate? Why or why not? If you think there will be a relation, will it be positive or negative?

Compute a regression equation in which you predict heart rate from body temperature.

2. What is the value of the slope of the line (B_1)? Is it statistically significant? What does this value tell you?

3. What is the value of the intercept (B_0)? Is it statistically significant? What does this value tell you?

4. What is the predicted heart rate for someone with a body temperature of 98.6?

5. Are the results consistent with your predictions?

Compute a regression equation in which you predict heart rate from zombie status.

6. What is the value of the slope of the line (B_1)? Is it statistically significant? What does this value tell you?

7. What is the value of the intercept (B_0)? Is it statistically significant? What does this value tell you?

8. What is the predicted heart rate for someone who is a zombie?

9. Are the results consistent with your predictions?

PSYCHOLOGY 203
Project 10

DUE: May 1, 2019 @ 10 AM
Points (50)

You should use the data file labeled PROJ10A on Blackboard.

These data reflect heights of randomly selected NFL and NBA players and are expressed in feet (i.e., a value of 6.5 indicates the player is 6 feet, 6 inches).

You are interested in evaluating whether NBA and NFL players are, on average, the same height.

1. What are your null and alternative hypotheses? Provide a brief statement in support of your prediction.

Run an independent samples *t*-test.

2. Report the mean and standard deviation for each group.

3. What is the observed *t* for this analysis? Is the difference between NFL and NBA players statistically significant? What information did you use to determine this?

4. Provide a **graphical depiction** of the observed difference in means along with an **explanation**.

You should use the data file labeled PROJ10B on Blackboard.

These data reflect test scores for a group of students enrolled in a tutoring program. Values reflect the percentage of questions answered correctly on a *pretest* and *posttest* competed at the conclusion of the course.

You are interested in evaluating whether individuals, on average, saw an improvement in test scores as a function of the tutoring program.

5. What are your null and alternative hypotheses? Provide a brief statement in support of your prediction.

Run a paired samples *t*-test.

6. Report the mean and standard deviation for the *pretest* and *posttest* scores.

7. What is the observed *t* for this analysis? Is there a statistically significant difference between *pretest* and *posttest* scores? What information did you use to determine this?

8. Provide a graphical depiction of the observed difference in means along with an explanation.

9. Can you conclude that (a) all students saw their scores increase after the program, and (b) that the tutoring program caused any changes in scores? Defend your answers to both parts of the question.

You should use the data file labeled PROJ10C on Blackboard.

These data reflect student scores on the first quiz in a biology class. Students were categorized in terms of where they sat during class sessions prior to the quiz. Students sitting in the first two rows were coded as GROUP A, students in rows 3 and 4 were coded as GROUP B, and students in rows 5 and 6 were coded as GROUP C.

10. What are your null and alternative hypotheses? Provide a brief statement in support of your prediction.

11. Before running the one-way ANOVA, report the means and standard deviations for each of the groups.

Conduct a one-way ANOVA

12. Report the **source table following APA style**.

13. Is there evidence for a significant difference between these groups? What information did you use to answer this question?

14. If there is a significant effect, report the results of any post hoc tests that you did. What groups are significantly from one another?