BALL STATE
UNIVERSITY

Practical Assignment I

Analysis of Variance

Quinton Quagliano, M.S., C.S.P

Department of Educational Psychology

Table of Contents

1 Pre-assignment Reminders

- Prior to attempting this homework, I strongly recommend that you have completed all other work through module 3, including:
  - Likely the most useful resource will be the video I specifically record for this assignment (near the assignment portal on Canvas)
  - Having watched all lectures thoroughly and taken notes (optionally with the guided notes)
  - Reading all assigned sections of the book, and completing practice examples as helpful and useful
  - Completing relevant lecture check-ins, and using answer key for review
  - Completing all quizzes, and reviewing correct answers at the end of the week
- This practical assignment will be graded for accuracy, please give yourself enough time to complete your best work
  - You may use your notes, book, SPSS software, provided data, and lectures to aid you in completing this
  - Where possible, include more detail to ensure you fully explain your rationale for a question
- All syllabus and university policies on academic integrity, plagiarism, and other forms of misconduct apply to this assignment. Please review them if you are unfamiliar

2 Context

The following information will be useful for answer some questions on the homework, please ensure you read and understand the following context

You are working at an education research firm that is trying to summarize data about enrolled college students; this data been gathered as part of a partnership with all public universities in Indiana. For the purpose of this project your team is only interested in extrapolating conclusions from the data to students at public colleges in Indiana. At the start of the project, it is decided that getting all data from all public schools is too burdensome, so your team decides to list all the public schools and then use simple random selection to choose 4 of them. During this random selection process, one university is selected at a time, and removed from the "drawing" on future iterations of the random selection. Thus, the odds of any given university being chosen for the sample increases slightly until all 4 are chosen. Then, all students at those 4 randomly chosen universities have their data used for the study.

At this stage in the project, your team is mostly interested in ensuring that the data is described accurately, prior to any sort of inferential analysis. The partnered universities are asking for some basic aggregate data to use of their own use cases.

Your goal is to use your skills in SPSS and growing statistical knowledge to create a preliminary report showing relevant descriptive statistics and graphs. Secondarily, you

want to accurately describe the nature of the data and variables that are present so that such characteristics are apparent during inferential analysis (as that information will be especially important then!).

Please use the provided dataset on Canvas, edu_data.sav, to complete this assignment.

3 Instructions

3.1 Knowledge Checks (5 pts)

1. Identify the population of interest in this study (1pt)
2. Identify the sample in this study (1pt)
3. Explain whether the sampling technique used will produce a sample representative population or not (2pt)
4. The following descriptive sample statistics are meant as estimates of the population _____ (fill-in-the-blank; 1pt)

3.2 SPSS Applied Work (15 pts)

1. Determine how many variables and how many constants are in the dataset (1pt)
2. How many total records/rows are there in the data, i.e. the n of the sample? (1pt)
3. For each of the variables, identify their level/scale-of-measurement (2pts)
4. For each of the nominal/ordinal scale variables, determine the frequency of each level (2pts)
5. For each of the ratio/interval scale variables, determine the mean, median, variance, standard deviation (2pts)
6. For each of the nominal scale variables, display a bar chart (2pts)
7. For each of the ratio/interval scale variables, display a histogram with a normal curve overlaying it (2pts)
8. Compute a new variable called 20_over from the age variable and use the new variable to determine the number of individual in the sample over 20 (2pts)
9. Provide the SPSS syntax for generating a frequency table for ‘Enrolled’ column (1pts)