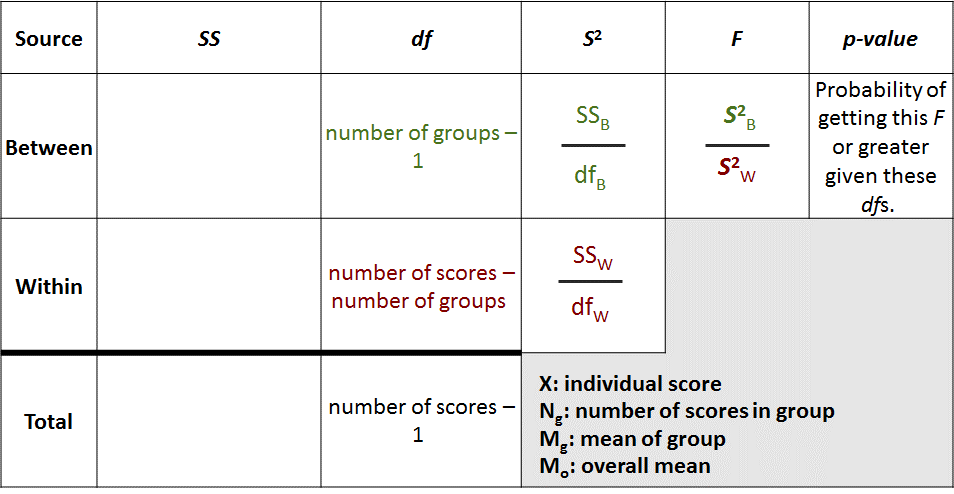
**Application Problem #3**

**Instructions**:

* Submit a typed, neat and organized final paper that shows each solution/answer in order. This should look like a professional report in quality of presentation. Please do not include a cover page. To ensure blind review, do not include your name. Use just your student ID number, along with the course and section number, as a header at the top of the first page. Please open this document in Word and use it as the template for your report. Prior to submission on Blackboard, however, you should save the document as a PDF (or print to PDF).
* Parts of problems (1.a-f and 2.a-d) and steps within hypothesis tests should be clearly labeled and highlighted on your submitted assignment. These labels should be at the edge of the left margin, with everything else indented, so as to visually set apart the labels (as in an outline). This Word document is already set up that way; please use it as a template in which to complete the assignment. Do not attach appendices showing your calculations, but rather show your calculations/demonstrations ahead of each solution, in order.
* Keep at least two decimal places during calculations.
* Show your work. When using formulas, copy the formula, substitute the values, show samples of any individual calculations you did, and then provide the answer. The formulas included in this document are editable. You can copy and paste them where you need them, or move them, then use the equation editor to add onto them. When a formula requires repetitive calculations, first provide a demonstration of the use of the formula. In this, you may show the first of the calculations by hand, write “…”, and then show the last of the calculations as your demonstration for use of formula. *In addition* present a table or copy-paste a spreadsheet to show all of the individual calculations. For a model of this, please see worksheet solutions documents posted in Blackboard under the relevant topics.
* Your grade on the assignment *will* reflect whether you followed these instructions. See rubric at end.

1. A statistics instructor is curious to find out if an assignment designed to help students apply concepts ahead of an exam is effective. As a first pass at testing this question, the researcher tested the relationship between assignment and exam performance by collecting data from 30 students on assignment (AP) and exam scores over the course of a semester. The data are included in an Excel file called “AP3 dataset”.
2. (2 pt) Draw (or generate in Excel) a scatterplot of the data with labeled axes. The plot should be large and precise, so every individual point is easily discernable.
3. (7 pt) Conduct a hypothesis test (significance level 0.05) to see if there is a significant positive relationship between AP scores and exam scores. A list of useful formulas is provided below.
4. (2 pt) Write up the results of the study in one sentence as it would appear in a publication (being sure your conclusion respects the limitations of the methodology used). At the end of the sentence, in brackets, report your test statistics and p-value.
5. (1 pt) Calculate and report (in a sentence) the proportion of variance in these two measures you have explained through their relationship with each other.
6. (2 pt) Calculate and report the equation for the best fit regression line for these data. Also, reproduce the scatterplot you created for part a, and plot the regression line precisely on this new copy of the scatterplot.
7. (2 pt) Predict the exam score for an individual with an AP score of 50. Report the actual exam scores(s) for individual(s) with an AP score of 50. Comment on why the actual exam score(s) is/are not the same as your prediction.
8. The same investigator decided to analyze the data differently, dividing up the participants into Low, Medium, and High AP-score groups. In other words, you will convert AP score from a numeric to a nominal variable, leaving the exam score as the numeric variable for their analysis. *(The dataset is already sorted by AP score, so the Low-AP group will include the first 10 participants, the Medium-AP group the next 10, and the High-AP group the last 10 participants listed.)* You should now conduct statistical analyses to see if the mean **exam score** differs among these three groups.
9. (2 pt) Draw (or generate in Excel) a bar graph of the group exam-score means with labeled axes. Error bars are nice, but not necessary.
10. (7 pt) Conduct a hypothesis test that corresponds to this new analysis (significance level 0.05). Step 4 should conclude with a fully-filled-out summary table.
11. (2 pt) Write up the results of the study in one sentence as it would appear in a publication (being sure your conclusion respects the limitations of the methodology used). At the end of the sentence, in brackets, report your test statistic and p-value.
12. (4 pt) This new statistical design resembles one that might be used to analyze the results of an experiment. List the required features of experimental design that allow for causal conclusions. Then comment on whether it is appropriate to draw cause-effect conclusions from this analysis and why (or why not).

USEFUL FORMULAS:



**Grading Rubric for Application Problem 3**

For each part, how well did the student do on this criterion? Was it on target, or did it need improvement? The sum of the number of on-target criteria is the score for that part.

1a. Scatter Plot

* Is the plot precise and accurate, with the correct variables on the correct axes?
* Are there accurate axis labels, and is the plot large enough so that every individual point is easily discernable?

1b. Hypothesis test on relationship

* Are the populations and hypotheses accurately identified as per in-class model?
* Are the mean and standard deviation of the comparison population accurately stated?
* Is the appropriate coefficient accurately calculated, with all calculation work and formulas shown?
* Is the appropriate standard deviation accurately calculated, with all calculation work and formulas shown?
* Is the cutoff score accurately identified along with its determinants with no evidence of logical inconsistencies?
* Is the test score accurately calculated, with all calculation work and formulas shown?
* Is the correct decision stated, with no evidence of logical inconsistencies?

1c. Results sentence on relationship

* Are the results stated with the variables clearly identified, and are causal inferences and the word significant appropriately used?
* Are the statistics represented with appropriate punctuation and with no evidence of logical inconsistencies?

1d. Proportion of variance explained

* Is the appropriate measure accurately calculated, and its meaning interpreted in a clear sentence?

1e. Best fit regression line

* Are the slope and intercept of the regression line accurately calculated, with all calculation work and formulas shown?
* Is the line represented accurately on the scatter plot, with all calculation work and plotting points shown?

1f. Exam score prediction

* Is the exam score predicted accurately, with all calculation work and formulas shown?
* Is the actual exam score(s) identified and discrepancy with prediction adequately explained as modelled in class?

2a. Group means graph

* Is the graph precise and accurate, with the correct variables on the correct axes?
* Are there accurate axis labels?

2b. Hypothesis test on mean exam score differences

* Are the populations and hypotheses accurately identified as per in-class model?
* Are the characteristics of the comparison population accurately stated?
* Is the cutoff score accurately identified along with its determinants?
* Are the sums of squares accurately calculated, with all calculation work and formulas shown?
* Are the variances accurately calculated, with all calculation work and formulas shown?
* Is the test score accurately calculated, with all calculation work and formulas shown?
* Is the correct decision stated, with no evidence of logical inconsistencies?

2c. Results sentence on mean exam score differences

* Are the results stated with the variables clearly identified, and are causal inferences and the word significant appropriately used?
* Are the statistics represented with appropriate punctuation and with no evidence of logical inconsistencies?

2d. Causal conclusions essay

* Are the required features of experimental design that allow for causal conclusions accurately and clearly identified?
* Is the first design feature accurately analyzed with reference to the current study’s features?
* Is the second design feature accurately analyzed with reference to the current study’s features?
* Is the correct conclusion drawn regarding appropriateness of cause-effect conclusions in the current study?

Global criteria:

* Was the assignment organized and formatted according to instructions?
* Was the writing clear and grammatically sound?