<u>The Process</u>:

- Identify a project manager (they will lead meetings, help keep things moving forward, provide accountability for all, they will lead, but not "be in charge") and determine a "meeting minutes" taker (they will document the discussion in the meeting, and take good notes on who is doing what by when).
- Get the data set and "figure out" what you have (mostly this is play around with stuff and make sure you understand what the variables you are working with mean).
- Get a firm grasp on the client's point of view. What will they be trying to accomplish with your results? What type of results are more useful?
- Look for variables that are more closely connected with the piece you are trying to predict/understand, either because they logically have a strong connection or because there is high correlation in the data.
- Figure out how different values in the variables impact the outcome (differences in averages/proportions between groups).
- Consider a predictive analysis. Can a reasonable regression model be built?
- Write a coherent, integrated, concise report for the decision makers. Use the following general structure when writing:
 - Restate the general problem, and then give a very brief overview of how you address the problem (this means tell them what to expect in the paper in a sentence or two - do NOT tell them all the things you did to get your results, they don't care).
 - Give pertinent descriptives, starting with single variable and then moving into grouped stats. (Graphs may often be important here - see Resources for Writing/Communicating Data.) Be sure that your variables are clearly defined to the reader.
 - Use the descriptives to point out the "question" that the inferential stats will answer, and use the HT/CI to answer it. (It should be a smooth ride from sample data all the way through to conclusions and how to act going forward.)
 - Include smooth transitions, recaps, and conclusions as appropriate to make the document more understandable to your audience. Assume that they have zero insight into what the data says.
 - Proof-read/revise (multiple times). Each member of the team should proof-read and revise the entire report twice for both paper submissions (everyone revise once, then start again). Collaboration on revisions is acceptable. This should be documented in meeting minutes.

Presentation and style matter, so consider formatting (for clear communication and ease of reading), proofread and revise (at least 2-3 revisions for each draft), avoid egocentric presentation (the report is not for you, you already know what it says), and say only exactly what needs to be said, nothing extra, but nothing less.

At no point should there be a play-by-play approach telling what you did. This is not a lab report; it is summary of key findings for decision makers to use to improve the quality of their decisions. Avoid first person (it just reads awkwardly; using phrases like "we find that ..." is generally okay) and use complete sentences/paragraphs with introductions, conclusions, and transitions (do not use an outline format, although judicious use of section headers is acceptable, as long as intros/conclusions/transitions are used).

• Review these instructions regularly so you maintain a big picture view of what you are trying to accomplish and don't get lost in the weeds of the nitty-gritty.

Goals of this assignment:

- 1. to help you apply statistical analysis and understand what all the theory is designed to do (<u>nothing</u> builds experience and confidence like actually doing an analysis project from start to finish)
- 2. to help you learn to manage larger scale projects and less than detailed instructions as part of a team
- 3. to give valuable insight to our clients in how they can improve their business

Signed Confidentiality \sim Due August 31 \sim 0 points

Weekly Group Update \sim Due each Monday in class, Sept 7 - Nov 2 \sim 5 points each

Descriptives Plan \sim Due as part of Update 1

Analysis Plan \sim Due as part of Update 4

Descriptives Rough Draft ~ Due September 30 in class ~ 5 points Inferential Rough Draft ~ Due October 14 in class ~ 10 points Revised Rough Draft ~ Due November 4 in class ~ 25 points Final Draft ~ Due November 18 by close of building ~ 150 points

Weekly Update Grading Guidelines:

Grades will be based on the quantity and quality of the work presented (and may not be the same for all group members)

Descriptives and Inferentials Rough Draft Grading Guidelines:

This is progress checkpoint, intended to provide feedback to help you improve your presentation. Peer evaluations are not submitted at this point. However, the grade will be awarded based on quality using the guidelines for the final draft. Write the best draft you can; be sure to revise and proofread before submitting. Revisions from previous submissions should be incorporated in the Inferentials rough draft.

Revised Rough Draft Grading Guidelines:

Revisions from previous submissions should be incorporated. See Final Draft guidelines (including initialing). Peer evaluations will not be submitted at this time. Your grade will reflect my estimate of your final grade if you revise and continue on in the same manner.

Final Draft Grading Guidelines:

- C Papers: will have identified a couple of valuable connections/relationships along with additional less significant or trivial relationships and will communicate these adequately with reasonable clarity A/B Papers: will have identified more valuable connections/relationships and will communicate these with great clarity in a <u>unified</u> paper
- Both drafts will be submitted in paper due by close of building on the day assigned. The paper copy should include handwritten initials by each paragraph/section indicting the major contributor(s) to that portion (I'm expecting that there won't be more than two major contributors to a single section).
- Evaluation of your participation on the team by your peers (10%) excessive failure to participate may result in additional penalties on your grade
- Presentation/style/formatting/professionalism/readability/clarity/conciseness (20%) including appropriate APA presentation of statistical data (integrated with English) and effective use of the introduction and conclusion towards the readability of the paper
- Quality/communication/insight of analysis (70% total)
 - Sufficient and thorough descriptive support Are your hypotheses a clear/immediate consequence of your discussion of the descriptive?
 - Appropriate discussion of the descriptives Did you explain what they mean or did you just state them and assume the reader understood why they were significant?
 - Tests which tests were run, were they appropriate, what were their results
 - Smooth integration of inferentials and descriptives Was a continuous, smooth argument presented that obviously led to the conclusion?
 - Conclusions from individual tests clearly stated/appropriate based on tests run/data presented

- Were there sufficient inferential stats run to justify a robust analysis which tests were run, were they appropriate, what were their results? was the analysis thorough? did it investigate a breadth of relationships? (extra credit may be possible here for superb analyses)
- Overall conclusion includes a summary, not a restatement, of all tests/conclusions run will the conclusion be useful to decision makers? does it highlight the important provable facts determined in your analysis, in a succinct manner?
- Executive summary includes actionable statements based on the analysis? were actions recommended that are supported by the analysis?

Suggestions you may find helpful:

- 1. Make a Descriptives Plan:
 - Review all of the available variables, and plan how you are going to evaluate the data to best address the goals of the client. Doing "everything" is probably not a good idea (it's overwhelming).
 - Identify the most significant dependent variable to the decision maker (what do we want to know about/predict?)
 - Create a list/table of the relationships that you think will be most useful. You will want to consider single-variable stats in moderation and focus on multi-variable stats.
- 2. Descriptives:
 - Using the data provided, run a thorough descriptive analysis. Consider how various variables influence your dependent variable.
 - Your analysis should focus around the general research question, and the analysis should clearly lead the reader up to specific hypotheses, which be addressed in a later draft.
 - You will summarize <u>only</u> the pertinent descriptives and include the following (in this order, this is the clearest way to communicate your ideas):
 - present your general research question
 - the pertinent descriptives (there will likely be several) that led to your hypotheses and state the null and alternative hypotheses that you are considering (for each hypotheses) (but don't write these in your report)
 - Be sure to incorporate the following guidelines for statistical writing
 - don't assume your audience knows what your variables mean (however, you can assume that they are familiar with them, just maybe they have forgotten the specifics), and use English names
 - define any variable groupings that you have created and included in your final presentation
 - discuss your graphs before you present them (tell us what information the graphs convey and what conclusions we should be drawing from the graph): for example, "We can see in Figure 2 that the average grade by classification in Ex 101 is higher for freshman than any other classification."
 - be sure that all graphs have titles, captions, and labels on each axis (or clear legends depending on the style of graph)
 - Null and alternative hypotheses should not be stated explicitly, but should be clear from the narrative in your report.
- 3. Analysis Plan: determine a detailed analysis plan (HT needed?, CI needed?, run post-hoc tests?, how will predictive analytics be used?, etc.).
 - type of hypothesis test (HT) needed, run post-hoc tests?, how will predictive analytics be used?, will confidence intervals (CI) increase the value of your analysis to the decision maker?, etc.
 - Hypothesis that can only be addressed by %'s (Y/N's) can be paired with a related means question to add depth.
 - Don't forget your general question. This ties the paper together.

4. Inferentials:

• Implement your analysis plan (incorporating any suggestions received).

- You should explicitly state the name of the test/analysis tool you used, the pertinent results from that test/tool, and what we can conclude from the results in English.
- A thorough analysis can include developing a regression model to address your question.
 - regression can include demographic variables as well as appropriately handled nominal variables
 - there should be only one (maybe two depending ...) regression model developed since regression allows us to integrate multiple independent variables into the predictive model (note that HT only allows the comparison of one independent variable to the dependent variable, so multiple HT are needed even when only one regression model is useful).
- Use appropriate APA formatting to concisely present the results of your analysis (not every scenario is handled on the course webpage, you are welcome to use Internet or library resources to help with appropriate formatting). APA formatting is <u>required</u> for the presentation of the statistics (it's the most concise, easiest to read format we've seen); however, you do NOT have to use APA style for the rest of your paper (headings, paragraph/section formatting, citation, etc.). See http://math.bju.edu/ma404/ for more resources on APA.
- Point estimates (such as the sample mean or proportion) can be poor approximations of the true values. CIs along with point estimates can efficiently provide better understanding to the decision maker. See the APA guidelines for presenting CIs concisely.
- You should present content in the following order for each null hypothesis:
 - general description (and possibly single variable stats) of the question being studied (usually we look to the alternative hypothesis for this)
 - key elements of the descriptive analysis that make the alternative hypothesis seem like an important question to ask
 - the HT run and the results of the test and any other pertinent inferentials such as CIs
 - a summary/conclusion of what we know from a practical/decision-making perspective
 - you may also need an "overall" summary if several hypotheses speak to the same general question
- Use simple/natural language when possible (as opposed to formal-sounding or "mathy" language), i.e. "shows a correlation" vs. "is correlated"
- The English should hold me by the hand through your argument to the "obvious" conclusion. Your audience is not stupid, but they are also not informed about the data or the research, so you need to lead them carefully through the logic of your thinking.

5. Final Draft

- Revise your descriptives/inferentials based on personal proofreading and feedback from previous submissions.
- Write an executive summary and a conclusion.
 - The conclusion should be the final paragraph of your paper and ensure that it incorporates the breadth of your analysis into a summary that can be used to make decisions (1/2-1 page in length). Actionable recommendations should be included either immediately following the presentation of the analysis or incorporated into the conclusion (whichever makes more sense based on the recommendations being made).
 - The executive summary (1/2-3/4 page in length) should be at the beginning of your paper (but is not the introduction). It summarizes the conclusions of your analysis targeting actionable statements, not simply a restatement of the results of the inferential stats, for the appropriate decision maker.

Note that an executive summary should include all the conclusions of the paper, but NONE of the supporting evidence, and your target audience is someone who needs to know what the bottom line is and has 3-5 minutes to spend on it.

 Your paper should also include a proper introduction after the executive summary to communicate the general topic of the paper to anyone who is reading the entire paper.