**Stat 3220**

**Writing a statistical report**

For your Mini Projects, you will be required to submit a semi-formal statistical report of your analyses. This handout will help guide you to writing reports in statistics. Although there is no one “correct” way to write a statistical report, there are many ways to make your report more readable to the average reader, while conveying your statistical analyses. These guidelines should be used to structure your Mini Project reports, but can be enhanced to be more formal beyond this course. As you advance to further courses or even consulting out in the real world, your reports will become more formal and require more in-depth details.

**Why am I doing this?**

You may be thinking to yourself, “I can do the code in SAS (or R or even by hand), why do I need to take the time to write up everything I am doing?” Well, as a statistician, you are required to work with other researchers who have varying levels of statistical backgrounds, simply giving them raw computer output will not be sufficient when you want to explain your findings. You must learn how to communicate your findings effectively to those people. By practicing this process in this course, you will be better prepared for collaborating in your future careers.

**What do I say?**

No matter the level of statistical knowledge of your collaborators (or to whomever you are reporting), reports should not over-explain statistical techniques that were used with rigorous mathematical formulas, nor include too much output, nor a step-by-step explanations of your procedure. Your reports should contain just enough information that whoever picks it up can read it and understand exactly the purpose, analyses, and conclusions. For the purpose of this class, you should write your reports assuming you are talking to a student who is taking this course.

**How do I do this?**

The following is a list of sections that should be included in your report and what should be included in each of those sections.

1. Introduction:

In this section you will explain the purpose of your report. Here you will assume that the reader has no idea what the report is about. You should answer the following questions, while giving any pertinent background information: What is the research question you are trying to answer? How was the data collected? What is the purpose of the analysis? In larger research projects, this section will often include a literature review with relevant material about the subject. This section should not include statistical references, except the general scope of statistical analyses that will be used to answer the research question. For the purpose of this class, this section should be about one paragraph.

1. Data Summary

This section consists of data summary in two parts. The first part should be a literal summary of the data. That includes what are the variables and how were they measured. With many variables, this is best to list in a table instead of text. If your variables have a long name that you will refer to with a coded name throughout the report, mention that here. This section would also explain any data manipulation that was used, including missing data, removing data, and data transformations. The second part of this section should include exploratory data analysis. Remember the first rule of statistics: MAKE A PICTURE MAKE A PICTURE MAKE A PICTURE. The remainder of this section should be used to do just that by including all relevant numerical and graphical summaries. This section typically ends with where you will begin building your model.

1. Analysis

This is generally the most varied part of the report, and often requires several sub-sections, since there may be multiple analyses performed. You might begin by explaining the types of analyses that will be done along with their assumptions, and then perform these in some logical order, usually from simplest to most complex, although that doesn’t necessarily make sense in all analyses. In some reports, all of the analyses will have the same general format, so one explains the first analysis in explicit detail, while producing relevant summaries for the other analyses. This section should not be used to make contextual conclusions, but you should offer comment on the interpretation of all output that you include in your report. You should simply explain the process you are going to use then perform the analysis, only including and labeling relevant output that will be used in your results section.

1. Results/Conclusion

This section should synthesize your results and give more colloquial descriptions of your analyses. Here you avoid using referring to statistical modeling and concepts. For example instead of saying “the results are statistically significant,” you should explain what statistical significance means in the context of your research question. This section should clearly answer your research question in language that is understood by your collaborators. Additionally, this section should include areas for future research; any cautions that the researcher should be aware of, such as violations of statistical assumptions; and how to possibly improve the research design.

1. Appendix

For the scope of this course your appendix should include your SAS code. It would be especially useful if the code were annotated so a reader knew why you chose to use certain commands. In more formal reports this section could also include tables and figures that are relevant, but not imperative to answering the research question.

**General Comments**

Make your report readable and organized by including page numbers, headings, references table/figure by numbers.

Round off digits so that they make sense in interpretation (four decimal places aren’t often useful and can clutter a report).

Before you copy output from SAS directly in your report, be sure it is labeled properly and makes sense. SAS gives output that is not relevant for every project.