

## **Stat 544 Mini-Project (and Project) "Pointers"/Questions/etc.**

The following are some points that I made in Stat 544 after marking the first Mini-Project in Spring 2006 and some others that I routinely make when marking projects in other courses. You may take these as advance warnings of things that will cause me to use "red ink" on your papers. ;+}

1. Use paragraph and complete sentence form! These reports are NOT meant to be sets of PowerPoint slides!
2. Use LaTeX or WORD's equation editor (or something else equivalent) to properly display mathematics. (Math should be in "italic," functions names are not set in "italic," etc. Look carefully at how a good statistics text is typeset and learn to make your stuff look similar.)
3. DO NOT write in passive voice. Take responsibility for what you are saying! Do not write "It is believed." Instead write "I believe."
4. Be consistent with singulars and plurals, especially with possessives. Do not write "The company contacted their customers" or "The customer first checks their order." Write "The company contacted its customers" or "The customer first checks his or her order." (Yes, I know that "his or her" is clumsy. If you don't like it, avoid the possessive construction completely, but DO NOT use the plural possessive "their" when you are speaking of an individual.)
5. Pick a tense and stick with it unless you really do need to contrast something that was done in the past with something in the present or the future. Don't jump back and forth between tenses without reason.
6. Be consistent in font, style, spelling, abbreviation, format, etc. There is nothing less professional/more juvenile than a document that jumps around in any of these regards.
7. Do NOT use the preposition "within" in place of "in." "Within" is commonly much over-used and is both pretentious and inaccurate when used in place of the simpler "in." "Within" is a more restrictive word than "in." It means "in the inner parts of; inside; inside the limits or extent of in time or distance; inside the fixed limits of; not beyond; in the scope or sphere of; inside a specified amount or degree." Students almost always do NOT mean "within" when they use it in place of the much simpler "in." "Within" is NOT simply a formal version of "in."
8. If you use a comparative like "sooner" you are obliged to make it clear to what you are comparing. That is, you must explicitly answer the question "Sooner than what?"
9. Explain fully what you've done in all analyses. Don't leave your reader guessing.
10. Read every sentence aloud by itself. Does it make sense on its own? If it does, does it move the whole discussion forward? Can you say the same thing just as clearly in half as many words? Then read the whole document aloud. Does it tell a story?
11. Equations are English sentences. Read them aloud. Disconnected scraps of mathematical symbols are sentence fragments and should either be incorporated into an appropriate equation or deleted.
12. What are the practical/subject matter implications of your statistical analyses?

13. You will always do far more work than will appear in a final statistical report. DO NOT waste space in a final report on failed attempts, or what in retrospect are silly analyses.
14. EVERY figure and table that you include in a report needs to be numbered, captioned, and referred to in the text. (If it is not included to illustrate some point you've made in the exposition, it has no reason to be in the document.)
15. Just cutting and pasting output without annotation and explanation is not helpful. You must tell a reader in English why it is there and what it says.
16. Include some kind of model checking for  $f(\mathbf{y} | \boldsymbol{\theta})$ . Generate  $\mathbf{Y}_{\text{new}}$  's from  $f(\mathbf{y} | \hat{\boldsymbol{\theta}})$  and/or from the posterior predictive distribution.
17. In talking about a model, the adjective "good" should mean that it does a nice job fitting the data. Whether it is easy to use in WinBUGS or the Gibbs sampler behaves decently does not make it "good" or "bad." These are matters of technical detail and implementation, not matters of whether the model (if properly used) is effective.