

## FINAL COMPUTER ASSIGNMENT FOR STAT 1: Salary equity at SAFECORP

### 1. Overview

You have been asked to provide statistical expertise in a lawsuit alleging discrimination against minority employees at Safecorp, a major banking institution. One of the issues at stake is salary equity. Your clients, the plaintiffs' lawyers, have supplied a random sample of records from 474 employees (dataset = safecorp\_data.dta). For each sample employee, you have this information:

- Employee's current annual salary (in dollars) (**annual\_sal**)
- Number of years working for Safecorp (**yr\_work**)
- Employee's position at Safecorp (**position**)
- Employee's minority status (**minority**)
- Employee's sex (**sex**)
- Highest grade completed in formal education (**highest**)

Your task is to address the following questions in a brief memo to the legal team:

- Are salaries for minority employees lower than those for majority employees?
- What characteristics are associated with salary?
- Do minority and majority employees differ with respect to the characteristics that are associated with salary?
- Do differences in characteristics of minority and majority employees "explain" the lower salaries of minority employees?

Begin your work by thinking about why we are recommending that you use these four questions to guide your data analysis. Then, as you analyze the data, stop periodically and reflect on what you have found. Reflect (and re-reflect) on your findings until you see a pattern in the results, or a "story". Tell that story in your memo.

## 2. POINTS TO CONSIDER

### Your audience

Write at a level that this audience can appreciate. You are addressing people who are familiar with basic numerical concepts (averages, percents). However, most are probably not aware of concepts like hypothesis testing, regression coefficients, and so on – after all, you didn't know about these until you took Stat 1. Such technical terms do not belong in your memo, unless you also explain them in plain English. A big part of your assignment is to convey your findings, using words that your audience can understand.

### Tips on organizing and writing a memo of this kind

Typically, memos of this sort are organized as follows:

1. Introduction. What are you going to talk about? Why are these issues important? (Note that because this is a statistics class rather than a policy course, this will be just one short paragraph.)
2. Methods: What is your data source for addressing these issues? Very briefly, how did you analyze the data?
3. Findings: What did you find? (This section will be the bulk of your memo). Here are a few tips

#### Do

- Use this section to tell the story that was revealed in your analysis
- Highlight the most important and interesting findings. Tell important things first, and save details for last (or omit them altogether)
- Use bolding and bullets as needed to separate elements of the story, as needed
- Draw your reader's attention to important trends in the tables. For example, "As shown in Table 1, Group 1 and Group 2 differed on the following characteristics..."

#### Do not

- Bury important findings in minutiae
  - Use space in your "Findings" section to recount every statistic in your table(s). In other words, be selective!
4. Conclusions: How do your findings address the question that you posed in the introduction? What do you conclude? Briefly, what are the limitations of your analysis?

### Deliverables

Your memo should be 3 pages double-spaced, with a 12 point font. Adhere strictly to the page limit --- your audience needs to be briefed briefly!

You will submit two appendices to your memo. These won't count toward your page limit.

1. Tables and/or figures. You must include **at least one good table** that illustrates your findings. You may also include figures. Tables and figures shouldn't be STATA output --- they should be clearly constructed displays of information in plain English, accessible to your client audience. Additional guidance on creating good tables and figures will be presented in class readings, and discussed in lab sessions.

Do not go overboard on tables and figures. Do not present redundant information, or information that is better conveyed in simple text.

2. Annotated STATA output for your instructor's review. This is just like the material that you have been attaching to your computer assignments. The annotations can be machine or handwritten on the output, and should guide your instructor in finding the analyses that you describe in your memo. For example, in your memo, you might have a sentence like "Roses are exceptionally thorny. While x% of roses have thorns, only y% of ferns do". In your annotated output, jot something like: "Percent with thorns: ferns vs. roses". This will help your instructor to see where you got your information.

To avoid generating lengthy printout, you should limit the scale of your STATA output. Your lab instructor will provide instructions for doing this.

### **Additional points on style and content**

In preparing any memo that includes statistical data, keep these in mind:

1. Round numbers to a user-friendly number of digits. For example, "The average height of women is 65.3 inches", not "65.357 inches". Readers can't absorb numbers if there are too many of them.
2. Don't just talk about statistical significance. Give specific statistics that your reader can evaluate. For example if you want to highlight that you found a difference in height between men and women, don't just tell your reader "There is a statistically significant difference between men and women". Rather, say something like "Men average 68 inches in height; women average 65 inches. This difference is statistically significant."
3. Style matters, both in written presentation and in the presentation of numbers. Be thoughtful, and be prepared to revise and re-think.

**Important note.** Doing the final assignment takes planning, time and concentration. Many students report that completing the assignment is the best learning experience in the course. You should allocate significant time to it, starting in Week 10 of the course. Please note the following:

- Your lecture instructor and your lab instructor will orient you to the assignment.

- You may work with your fellow students, but this is not a group project. You must run your own analysis, write your own memo, and create your own tables.
- Final assignments written by prior students in Stat 1 \*may not\* be used as sources for writing your final assignment memo.
- Problem-solving is a big part of the exercise. Lab & lecture instructors will field \*specific\* questions that are limited in scope. "How should I do this part of the analysis?" will not be answered. "I'm thinking of using procedure X to answer question Y because I believe that the assignment is asking me to examine relationship Z. Does this make sense?" is an appropriate question.
- Working and re-working the presentation of your findings is also a big part of the exercise. However, lab & lecture instructors cannot review your written drafts.
- Please do not ask your tutors for assistance with this assignment. Their responsibility is to handle questions related to the lecture/homework.