**Math 80 – Group Project**

For this project, you will pose a problem, determine its underlying statistical structure, gather data, organize and describe the data, draw conclusions from the data, and give a critique of what you have done. You will work in groups of 4-5 students and each of you will be responsible for collecting at least 25 responses to a survey that you create to answer the question that you pose.

**MAKE A PLAN:** Email me an outline of your plan and a copy of the survey you plan on using no later than **Thursday, April 21st**; the sooner, the better. I will make sure your plans are reasonable and will suggest changes if necessary. ***If you don’t turn in an outline by this date, you will lose points for each day it is late!***

**WRITE AN ESSAY:** Hand in the write-up of the completed project no later than **Tuesday, May 24** (the Tuesday before the final exam)**.** If you hand it in sooner, that will be fine, but you will probably need to wait until you have learned some of the techniques in the later chapters and you might want to keep it until you complete your group presentation to the class.

Your outline and write up must be typed, or printed using a word processor. They must be clear and easy to read. I would like the final projects to be submitted both via email and 1 printed copy handed in to me.

**PRESENT YOUR FINDING TO CLASS:** Each group will give a presentation to the class on their findings at the end of the semester. We will have project presentations in class on **Thursday, May 19th and Tuesday, May 24th**.

**Computing**: There are hand-held calculators (and computer programs like Excel) with statistical functions available for computing means, standard deviations, and correlation coefficients. Ask me for help if you need it. In class, we will learn how to use Excel/Google Spreadsheets and you can always ask me how to use your calculator. I just ask that you show your data in your project, but you do not have to show all of the steps in calculating things like mean, standard deviation, etc. You will have to interpret these values and discuss their meaning in your write up. With your completed project, I ask that you also send me your data in Excel (or whatever program you use – google sheets is fine).

**The Outline (due by 4/21)**:

Your outline should give the following:

\* The names of everyone in your group

\* Statement of the problem/question to investigate. You should be looking at a problem/question that can be answered by a survey that you can administer and that will involve both categorical and numerical data. (What would you like to learn about?)

\* Motivation. (Why do you want to learn it?)

\* Statistical context. (Subjects/Units, population(s), variables)

\* How do you plan to collect the data? What sample size do you expect? (Remember that each person in the group is responsible for collecting at least 25 responses!)

\* How do you expect to organize and describe the data?

\* What tools do you expect to use to draw conclusions from the data?

\* Critique. (What difficulties do you anticipate that you will encounter in collecting, organizing, describing, and drawing conclusions from the data?)

**The Survey (due by 4/21)**:

The survey you turn in to me should have at least 25 questions. We won’t need them all, but I can help you narrow down the questions to create a survey that will help answer the question your group poses. Sometimes we don’t get exactly what we think we will get from the questions we ask and I want to be SURE that your project will gather the right types of data (categorical and numerical) as well as give you the data you need to answer the question you pose.

**The Write-Up of the Project (due by 5/24)**:

The completed write-up should contain:

0. **The names of everyone in the group and the date**.

1. **Statement of the problem**. What do you want to find out? Be as clear and specific as possible.

2. **Statistical context**. State the unit, population(s), population size(s), sample(s), sample size(s), variables, variable types and their intrinsic nature.

3. **Sampling method**. How did you gather the data? Describe in detail the actual process used.

4. **The raw data**. Present the raw data. (This can be in a separate Excel file that you send to me)

5. **The data organized and described**. Use whatever *univariate* or *bivariate* tools are appropriate; for example, frequency or relative frequency tables, histograms, bar graphs, circle graphs, contingency tables, etc. Also, for example, mean, median, mode, standard deviation, range, shape of the distribution, percentiles, percentages, independence or association, correlation, regression line, etc.

6. **Inferences about the population**. What can you conclude about the population and its parameters using confidence intervals and hypothesis tests? [In order to demonstrate that you understand these tools, use them despite difficulties such as the sampling being non-random. Discuss in the critique the reasons why any of these tools were used inappropriately.]

7. **A critique of your work**.

(1) Did you answer the question you posed?

(2) If someone were to criticize your work, what would they object to? Would their objections be justified? How could the project be improved?

(3) What, if anything, surprised you about the data, the procedures you used, or the overall task of doing the project?

(4) What did you learn in doing this project?

**The Presentation of the Project (5/19 and 5/24)**:

You will present your findings to the class. These are the requirements for the presentation:

0. Introduce yourselves

1. State the problem in a clear and concise way

2. Each group member must speak for at least 4 minutes (The overall presentation should be 16-20 minutes). Each group member must introduce findings in the data, a concept used from class to analyze the data and draw a conclusion from it.

3. Visual Displays must be used (graphs, tables, …)

4. Overall conclusions you drew

5. Any problems you ran into or parts of your process that could be improved

**GRADING**

**I will grade your projects on the following criteria:**

(1) is neat, clear, well-organized, well-written, and correct;

(2) is appropriate to the question you are posing;

(3) includes all 7 required sections, with well written descriptions and information;

(4) includes both univariate and bivariate descriptions of the data, investigates relationships involving 2 or more variables that include (at least) the comparison of:

2 numerical variables,

2 categorical variables

and

1 numerical and 1 categorical variable

(5) includes a variety of descriptive methods and a more than one inferential method (in other words, both confidence intervals and hypothesis testing).

**Basically, your project should include many, if not all, types of evaluation we have learned in class, from beginning to end. I will create a rubric for how I will grade your projects and your presentation later this semester and I will share it with you once it is complete. After reading your outline, I will make sure that you are posing a wide enough question to target all course concepts and that your survey could adequately answer the question posed, as well as enable you to gather enough data to implement all concepts from class.**

Occasionally an innovative project does not fit neatly into all of the above categories but is still outstanding. It is possible that such a project might still receive an A. Hopefully I can determine this with your outline.

**Note (caution)**: If you do nothing more than present descriptive statistics on a single variable, or just present a single hypothesis test or a single confidence interval with no descriptive statistics, or just do some other “minimal” project, you will not get a high grade. If you have questions, please, always ask me. I can tell you if your project is lacking.

**Note (good news)**: You are **not** the Gallup Poll or a research institute. Your sample size might be too small; your sampling method might not be random enough; there might be extraneous variables. ***Discuss any of these difficulties in the critique.*** A project can have all of these difficulties (within reason) and still earn an A.