

# Math 113 - Semester Project - 100 points

---

There are many claims that are made which are suspect. Have you ever wondered if really "four out of five dentists surveyed recommend sugarless gum to their patients who chew gum" or if "people prefers Hardees chicken 2 to 1 over KFC's chicken"?

Your task is to perform some real-world inferential statistics. Try to come up with something of interest to you instead of some contrived situation. I have had people ask managers something about their business and then testing to see if the manager knew what s/he was talking about (ex: KFC's average serving time is 30 seconds or 57% of all drinks sold at Carlos O'Kelly's are Marguerites). Some have challenged a commercial advertisement (ex: people prefer Hardees chicken 2:1 over KFC's). Still others have tested the claim that local data is different from the national data (ex: The percentage of blood types in McLean county is different from the national percentages). One group tested to see whether the percentage of in town drivers wearing their seat belts was different from the percentage of out of town drivers wearing their seat belts.

The project will involve forming a hypothesis, then collecting the data necessary to conduct the test of the hypothesis, and finally analyzing and interpreting the results. Projects which involve using pre-existing data, rather than the collecting of data through sampling will require more analysis since the work of data collection was less involved.

You may work in groups of up to four (4) persons. Pick people you can work with; part of the grade will be assigned by the people in the group as to the work you contributed. Do not necessarily pick your friends.

You need to submit a proposal defining what it is that you wish to test, and how you wish to go about testing it (think back to the types of sampling). The instructor will peruse these proposals, make suggestions and give it back to you. The proposal will be included as part of the final project. If your group can't decide on a project or need help defining it, see the instructor.

Be sure to define clearly what the population being studied is (Richland Students, People listed in the Decatur phone book, People driving a car, etc).

Some of your projects will require information from sections in the book not yet covered. See the instructor if you have trouble identifying these areas. If you read the section(s) and don't understand the material, see the instructor for an explanation. Don't wait for the class to cover the material, it may be too late.

Make sure you get the project cleared with the instructor before you go collect the data. One person wanted to telephone survey some people and was talking in the range of \$100 phone bill if she called everyone she said she was going to. The project should not cost you very much money to implement. It will take some time, however, and you should not wait until it's due to get started on it.

While you are not precluded from doing any of the given examples, it is certainly better if you can come up with something original which has an interest to you. Things like the KFC/Hardees debate, while acceptable, are really poor choices.

The instructor will keep a copy of your final project.

The project will be comprised of several parts, due at different times during the remainder of the semester.

It is highly recommended the student read Chapter 14 before deciding upon their topic.

# Project Components

## Pre-project proposal (10 points)

This is to make sure you're on the correct tract before wasting lots of time collecting useless information. These points are awarded for having it done. Must be typed.

## Rough Draft (10 points)

This is a rough draft of the final report so the instructor can suggest corrections. These points are awarded for having it done. Must be typed.

## Final report (60 points)

The final report will include a description of the problem, and why you think it is important, or what you hope to gain from testing the hypothesis. It should also include all data collected, the statistical formulas used to perform the hypothesis test (you may write the formulas by hand) and the values generated by Statdisk or the calculator. A decision and conclusion should be stated. An analysis should follow with what the conclusion means in terms of the original problem. Must be typed.

## Presentation (10 points)

Classroom presentation of 4-8 minutes on why you picked the project you did, and what your results were. The class and/or instructor may ask questions on why you did something the way you did. These points will be assigned by the other class members, not by the instructor. Evaluations by the students will be handwritten.

## Self-evaluation (10 points)

Evaluate yourself and the other people in your group as to their ability to perform within a group and their ability to do assigned work. Can be handwritten.

## Some previous projects:

- Are the rates paid by the insurance company for dental cleaning in line with the rates charged by the dentists? A student called 30 dentists to find out the rates.
- Are serving times at Kentucky Fried Chicken less than the thirty seconds claimed? A student used the computer records during different times of the days. In the same vein, other students have sat in McDonald's for an hour and timed the delivery of the orders.
- Does the blood type ratios in McClean county agree with the national percentages as published by the American Red Cross? Students went through Red Cross records using stratified sampling until there were over two hundred people in the sample.
- Do people prefer Pepsi over Coke? People's preference was asked and then they were given a taste test.
- Are the men and women's shoe prices at Foot Locker, MC Sports, and Finish Line the same?
- A company claimed that the machine produced parts that were acceptable 90% of the time. They collected printouts of runs from the machines and tested the claim.
- A machine used as the Clinton Power Plant for measurements was tested daily to see if it was accurate. The student collected data from the machine to see if it was within the specified tolerance at least 95% of the time.
- Cubs, Krogers, and Eagles' prices were compared in several different categories to see if Cubs' claim of lowest prices was correct. The same thing has been done with WalMart, KMart, and Ventures. That group went and interviewed the managers about pricing strategies.
- Does Firestone/Bridgestone produce splices with the mean size claimed?
- Is the GPA of smokers lower than the GPA of non-smokers?
- Do higher priced bullets have a smaller shot pattern?
- Do Chex potato chips have 60% less fat than their competitors?
- Does the day of the week a test is given on affect the scores?

# Sample Proposal

Title: Smokers and GPA

Group Members: John Smith, Jan Doe, Jane Public

Claim: GPA is related to smoking

Population: Richland Students

We would like to test the claim that Grade Point Average is related to the number of cigarettes smoked per day. We will use convenience sampling to ask Richland students the following questions.

1. How many cigarettes do you smoke per day
2. What is your Grade Point Average

We will also record the gender of each person for possible further analysis.

Once we have collected the data, we will perform a hypothesis test to see whether there is any linear correlation between the number of cigarettes smoked per day and the GPA.

## What can we test?

Some things are easier to test than other things. The purpose of this project is not to do a full-scale PhD level research project, it is to expose you to the process of hypothesis testing in a real-world application.

You may test means, proportions, standard deviations, or linear correlation. It is very unlikely that you will test a standard deviation. You may have one or more samples. You may categorize your variables in several ways.

If you are dealing with one sample, then you will need some numerical value to test against. The claim “more people prefer Pepsi than Coke” becomes a claim that the proportion of Pepsi drinkers is greater than 0.5. There are not two independent samples (Pepsi drinkers / Coke drinkers), just one sample categorized differently. A problem with the Pepsi / Coke thing is that it omits other possibilities and becomes more difficult to do.

Again, read through chapter 14 for some suggestions. Realize that some of their topics are really lame and you should, if at all possible, come up with a claim that you have heard or that interests you rather than one out of the book.