

# FREC 408 - Fall 2003

## Assignment 1

Issued: 9/3/03

Due: End of day, 9/15/03

### Be sure to:

- Put your name and the Assignment # on the front
- Answer as completely as you can. All I can go on is what you give me, so show your work.
- Be as neat as possible. You can write it out, but please be neat.
- Staple or place in a folder

1. Go to the following web site for the **Current Population Survey** 25 pts

<http://www.bls.census.gov/cps/cpsmain.htm>

1. Read the *Overview* of the CPS on the main page - this gives an overview of this survey.
2. Go to *History and Concepts, Basic Monthly Survey*, and read the *Concepts* part. This section provides definitions for variables used in the survey, such as "Who is a Jobseeker" and "what is the definition of a family." **Take one definition and summarize and comment on it.**
3. Click on the *Publications* on the main page, and then *Index of Topics*. Pick any topic and visit it. Note: some will link you directly to data and some will go to other sites.
4. **All I want you to do is to visit the topic of your choice, view the data there, and briefly describe it.** This should be one or two paragraphs. You can describe the topic coverage, and anything that seems to stick out in your mind. For example, if you clicked on geographic mobility you would end up at a table showing males and females who lived in a different house, different county, or different state between March 1996 and March 1997. I could report the percentage of males and females who moved during the time period and how they differed.

2. Below is some data from the U.S. Bureau of Labor and Statistics for labor force participation rates in the United States from 1970 to 2006 (projected). Labor Force Participation rates reflect the percentage of the group who is either employed full or part time) or is actively seeking employment. The data are broken down by males and females. 20 pts

1. **Graph the data and show differences between males and females** over time using any method you wish, or any number of graphs.
2. **Briefly summarize the data in terms of any pattern you might observe.**

	1970	1980	1990	1998	2006
Males	79.7	77.4	76.4	74.9	73.1
Females	43.3	51.5	57.5	59.8	61.4

3. Would you expect the data sets described below to possess relative frequency distributions that are symmetrical, skewed to the right, or skewed to the left? **For each one (a to f) give your opinion and a short reason why you think it is so.** 20 pts
- The salaries of all persons employed by a large university
  - The grades on an easy test
  - The grades on a difficult test
  - The amount of time students in your class studied in a typical week
  - The age of automobiles on a used-car lot
  - The amount of time spent by students on a difficult examination (maximum time is 50 minutes)
4. On the resources page on STAT 200 Home page there is an Excel data set of the SAT scores by state. You can download this data. The data are 1999 state average scores for the Verbal part of the test (VERBAL), the math part of the test (MATH), the total scores (TOTAL) and the percent of the high school students who take the SAT test. Using Excel I want you to: 35 pts
- Sort the data by the Math Scores.** To do this you use the following commands:
    - Grab the entire data (all rows and columns for the data, including the Header Row - it should be shaded)
    - Under DATA in the Command Bar find SORT
    - In the SORT command box you want to designate MATH as the Column used to sort the data. You also want to specify that there is a Header Row
    - Then click OK and the data will be sorted by the Math Scores
  - Then create you own Stem and Leaf Plot** like we did in the group assignment.
  - Calculate the mean, median, and mode for VERBAL, MATH, TOTAL, and % TAKING** using the following Excel commands.
    - Below the last row of data for the VERBAL Column enter the following commands in successive rows
 

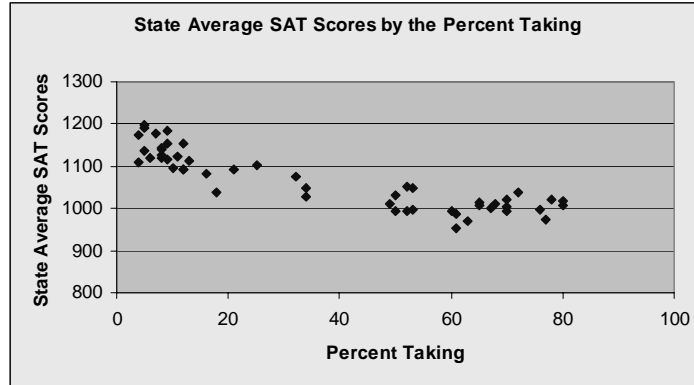
=AVERAGE(C5:C55)	this calculates the average of the column
=MEDIAN(C5:C55)	this calculates the median value
=MODE(C5:C55)	this calculates the mode or most frequent value
    - You can copy and paste the commands under the columns for MATH, TOTAL, and % TAKING
  - Construct a graph using Excel that shows the relationship between the TOTAL and % Taking.** This would be done by using the following commands:
 

INSERT  
CHART  
XY SCATTER  
Pick the first Plot Type

Once in the command box, select the data for the Y axis, and then specify the X axis by using the Series option in the command box. Don't grab the Labels when specifying the data range - graphs don't give you the option of a header row.

It might help to change the axis scale for TOTAL (Y axis) from 0 to 1400 to something like 800 to 1400. You can do this by clicking on the values for the Y axis in the graph, and once the command box appears setting the scale to go from 800 to 1400.

Give it a Title and any other window dressing. The finished product should look like this:



- d. **Briefly comment on the graph (one paragraph will do).** Does there appear to be a relationship between the average state score and the percent of high school students who take the SAT test?