

Math Q114
Collecting, Representing and Analyzing Data

Project 1
Due, February 12

Instructions: Each group will study a column of data from the student data set that we collected during the first week of class. Your goal is to use statistics to analyze this data and to prepare a written analysis of your results. This handout summarizes the project.

Introduce yourself to other people in the group. Get telephone numbers and e-mail addresses so that you can communicate outside of class. We will do some work in class, but you are expected to complete the assignment and outside of class. Please plan accordingly!

Step one: Take your column of data and copy it into a new spreadsheet. **Calculate mean and median.** What is the **range** of your data (take the difference between the highest and lowest value)?

Step two: **Make a histogram.** Discuss what interval size you should use and whether you want to make a frequency histogram or a relative frequency histogram. Once you have decided these questions as a group, each of you **individually** should create your own version of the histogram based on what you decided. When completed, save your work and print out a copy.

Step three: **Discuss together** your group's data, including the histogram, the distribution of the data and the mean and median. Be analytical and jot down patterns that emerge from the data. What conclusions can you draw? Are there any limitations to the data? What other questions are raised and how might they be resolved? If you can, go on the internet and search for information about the variable you are studying.

Consider the following points as you discuss your data. This will help you write an analysis of your data:

- What is the range of values for your data set? (For example: Class ages go from 17 to 59 years, a span of 42 years. Are there any values that don't fit a pattern (ie., are there outliers)?
- What is the shape of distribution of your data set in your histogram? Is it mostly grouped around the mean or median? Is it skewed left or skewed right? Is it bimodal? Is it spread out more or less evenly across the whole range?
- Look closely at the mean and median. Are they close in value or are they very different? Does the shape of the distribution explain why? If they are very different, which is the better measure of the class "average"?
 - Having answered all the above, how would you summarize your conclusions about your data set? What is the main point you want to make?

(more on back)



Step four: **Individually write** an analysis of your data taking into account all the points you discussed above. It should be a paragraph in length at minimum.

Guidelines for written work:

- It must be typed and carefully edited (use spellchecker and grammar check).
- Make sure your name is on the final report and that your work is **your own**.
- Attach a copy of your spreadsheet with mean and median calculations, plus the graphs you produced. (You may want to cut and paste your Excel graphs into the Word document. If not, then label your graphs so that you can clearly refer to them in your written summary.)
- Your written summary should include:
 - Title that clearly states the main point you discovered about your data set.
 - Topic sentence that reiterates the main point and serves as an introduction to the paragraph

A paragraph about your data that is **analytic not just descriptive**. You should cite specific things such as mean, median and shape of the distribution to back up your analysis. Your analysis should draw a conclusion about our class with respect to the data you were assigned to investigate. You should put the assigned data about our class into perspective by comparing it to similar data about all UMass students, or typical college students, or adults in the USA or world. You will need to consult the mass database (where you found Fast facts) or do a Google search.

Due February 8:

- Spreadsheet with data, Intervals and frequency and relative frequency counts.
- Histogram.
- Draft written analysis of your assigned data set.

Due February 12:

- Revised spreadsheet with data, Intervals and frequency and relative frequency counts.
- Revised histogram.
- Final written analysis of your assigned data set.