

Statistics for Engineers, 4C3/6C3

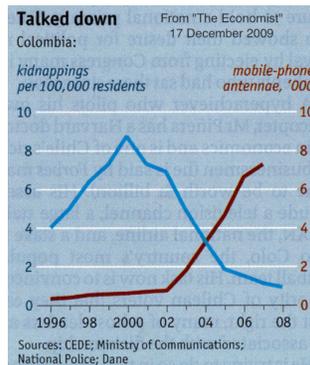
Assignment 1

Kevin Dunn, dunnkg@mcmaster.ca

Due date: 12 January 2011

Note: 600-level students must complete all the question; 400-level students may attempt the 600 level question for extra credit.

Question 1 [1.5]



1. What type of plot is this?
2. Describe the phenomenon displayed.
3. Which plot type asks you to draw a cause and effect relationship?
4. Use rough values from the given plot to construct an approximate example of the plot you proposed in part 3.
5. What advantage is there to the plot given here, over the type in your answer to part 3.

Question 2 [1.5]

Load the [room temperature](#) dataset from the general [Datasets website](#) into R, Python or MATLAB.

1. Plot the 4 trajectories, `FrontLeft`, `FrontRight`, `BackLeft` and `BackRight` on the same plot using the default settings in the software.
2. Comment on any features you observe in your plot.
3. Be specific and describe how sparklines of these same data would improve the message the data is showing.

Question 3 [1]

Load the [six point board thickness](#) dataset, available from datasets website.

1. Plot a boxplot of the first 100 rows of data to match the figure in the course notes (page 9 in the PDF version).
2. Explain why the thick center line in the box plot is not symmetrical with the outer edges of the box.

Question 4 [1]

Pie charts are widely criticized in the technical literature as being inappropriate - there is almost never a case where it is suitable - yet we see them in the media all the time. Next time you open a daily newspaper or magazine count how many times you see this type of plot.

Read the article by Stephen Few, “[Save the pies for dessert](#)” and explain in your own words the shortcomings of the pie chart. Which is an appropriate alternative?

Question 5 [1]

Using the [Website traffic data set](#):

1. Create a chart that shows the *variability* in website traffic for each day of the week.
2. Use the same data set to describe any time-based trends that are apparent.

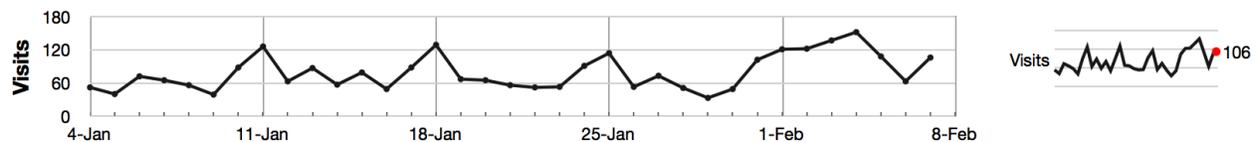
Question 6 [1] (600 level)

Copy a plot from any recently-graduated PhD student in your department. Include the plot in your assignment and comment on any shortcomings in the plot: how would you improve it and if necessary, reproduce your version of the improved plot.

Non-credit question

Note: *Question 1* from the course notes was a question from the 2010 midterm. Repeated below.

The data shown here are the number of visits to a university website for the 4C3/6C3 statistics course in 2010. There were 90 students in the course, however the site is also publicly available.



1. What are the names (type) of the 2 plots shown?
2. List any 2 interesting features in these data.

END