
COURSE: MATH 50 DARTMOUTH COLLEGE (MWF 11:15 AM-12:20 PM), FALL 2015
INSTRUCTOR: NISHANT MALIK
HOMEWORK SHEET NUMBER: 1
POSTED ON: 09/18/2015
DUE ON: 09/25/2015

Directions: Any problem marked with asterisk (*) should be completed using IPython Notebook (Jupyter) and can be uploaded at <https://dropitto.me/m50f15> .

1*. Samples of Roman-British pottery were collected from four different sites in the United Kingdom, a chemical analysis was done to find out the chemical (metal oxides) composition of these samples. Create a IPython notebook (Jupyter) with a box plot for the percentage of aluminum oxide and sodium oxide in the sample.

Data source and description: <http://lib.stat.cmu.edu/DASL/Datafiles/Pottery.html>

Data abstract and reference: <http://lib.stat.cmu.edu/DASL/Stories/Pottery.html>

2*. Use the dataset from the "rdatasets" package with the following description:

Package	Item	Title
robustbase	delivery	Delivery Time Data

to plot a scatter plot of delivery time versus distance in IPython notebook (Jupyter). Size of the marker should represent the number of products.

3*. The **central limit theorem** states that the distribution of the sum (or mean) of a large number of independent, identically distributed variables will be approximately normal, regardless of the underlying distribution. Visualize the central limit theorem in IPython notebook (Jupyter). Following steps will be involved in this visualization:

1. Generate uniformly distributed random numbers.
2. Plot their distribution.
3. Calculate means of each of the multiple sequences of uniformly distributed random variables of length n i.e., every sequence consists of n random variables.
4. Plot distributions for different n , as n is increased distributions are expected to resemble a normal distribution. One can also plot all the distributions for different n in a single figure by using "subplot" function from matplotlib.