Go to the class website and download this file: SDSS_DR7.dat
It contains data from the final (Data Release 7) SDSS galaxy redshift survey. Specifically, the file contains 550,166 lines, where each line has the following info for a single galaxy:

RA – right ascension
DEC – declination
z – redshift
M_g – absolute *g*-band magnitude
M_r – absolute *r*-band magnitude

a) Familiarize yourself with the data: Make a plot of *DEC* (y-axis) vs. *RA* (x-axis) showing the positions of the galaxies on the sky. Make a plot of M_r (y-axis) vs. *z* (x-axis) also showing all the galaxies (invert the y-axis so that more luminous galaxies are up).

b) Make a plot of the *(g-r)* color distribution of galaxies. Using a cut between red and blue galaxies of *(g-r)*=0.75, calculate the fraction of blue galaxies in the sample.

c) Plot the *r*-band luminosity function of galaxies: a histogram of $\log(dn/dM_r)$ as a function of M_r, where d*n* is the number density of galaxies (in units of $h^3\mathrm{Mpc}^{-3}$) in bins of M_r (use bins of width 0.1). To compute the density, you need to estimate a volume. This is difficult for a flux-limited sample since it has no clear redshift boundary, so just use the median redshift survey depth of $z<0.1$. You will also need the area on the sky covered by the SDSS DR7, which is 7,675.2 deg^2 (2.295 steradians).

d) Construct three volume-limited sub-samples of the full dataset that contain galaxies more luminous than -20, -19, and -18, respectively. List the redshift bounds, the volume, and the number of galaxies for each sample. Calculate new blue galaxy fractions for these samples.

e) Make a single plot of the luminosity functions measured from these three volume-limited samples. How do these look different from the luminosity function made using the flux-limited sample?

f) Make a new plot of the galaxy luminosity function using the full flux-limited sample, but this time applying a $1/V_{max}$ weighting. How does the weighting change the luminosity function?