Adapted from Mario Juric's Galactic Structure Course.

For this project, create a jupyter notebook with all your work answering the questions. There are many citations that are contained here which will help guide you through understanding the "why" for the calculations I'm asking. You will also need to do background literature searches in addition to the refs provided through this document.

The data file you need is Astr511HW1data.dat — this contains SDSS measurements for about 600,000 stars with $b>80^0$ and $14<r<21$. The data are listed as one line per star, with each line containing the following quantities:

- **ra dec:** right ascension and declination (J2000.0) in decimal degrees
- **run:** SDSS observing night identifier
- **Ar:** the value of the r band ISM extinction used to correct photometry (adopted from the SFD maps; for bands other than r standard SDSS coefficients are used)
- **u g r i z:** SDSS photometry (corrected for the ISM extinction)
- **uErr gErr rErr iErr zErr:** photometric errors
- **pmL pmB:** proper motion vector components in the longitudinal and latitudinal directions (mas/yr); set to 999.99 when no measurement is available
- **pmErr:** mean proper motion error (mas/yr); set to 999.99 when no measurement is available

For stars from this file, compute absolute magnitude using a photometric parallax relation, $M_r(g-i,[Fe/H])$, given by eqs. A2, A3 and A7 from Ivezić et al. 2008 (ApJ, 684, 287). For computing metallicity, $[Fe/H]$, instead of their eq. 4, use an updated expression from Bond et al. 2010 (ApJ, 716, 1):

$$[Fe/H] = A + Bx + Cy + Dxy + Ex^2 + Fy^2 + Gx^2y + Hxy^2 + Ix^3 + Jy^3, \qquad (1)$$

with $x = (u-g)$ and $y = (g-r)$, and the best-fit coefficients (A–J) = (−13.13, 14.09, 28.04, −5.51, −5.90, −58.68, 9.14, −20.61, 0.0, 58.20). This expression if valid only for $g-r < 0.6$; for redder stars use $[Fe/H] = -0.6$.

Since $b > 80°$, for these stars the distance from the galactic plane, Z, and the distance from us, D, are approximately the same. Using $Z = D$, where D is computed from $r - M_r = 5 * \log(D/(10\text{pc}))$, do the following:

1. For stars with $0.2 < g - r < 0.4$, plot $\ln(\rho)$ vs. Z, where ρ is the stellar number density in a given bin (e.g. look at Figs. 5 and 15 in Jurić et al. 2008, ApJ, 673, 864 for similar examples). You can approximate $\rho(Z) = N(Z)/V(Z)$, where $N(Z)$ is the number of stars in a given bin, and $V(Z)$ is the bin volume (note that the solid angle is $\Delta\Omega \sim 314$ deg°). What is the Z range where you believe the results, and why?

2. Add $\ln(\rho)$ vs. Z for stars with $0.4 < g - r < 0.6$, $0.6 < g - r < 0.8$, and $0.8 < g - r < 1.0$ (you can rescale all curves to the same value at some fiducial Z, or leave them as they are). Discuss the differences compared to the $0.2 < g - r < 0.4$ subsample. Why do we expect larger systematic errors for $0.8 < g - r < 1.0$ than for the adjacent bin with $0.4 < g - r < 0.6$?

3. For subsample with $0.2 < g - r < 0.4$, separate stars into low-metallicity sample, $[Fe/H] < -1.0$, and high-metallicity sample, $[Fe/H] > -1.0$. Compare their $\ln(\rho)$ vs. Z curves. What do you conclude?

4. For these low-metallicity and high-metallicity samples, plot and compare their differential r band magnitude distributions (i.e. the number of sources per unit magnitude, in small, say 0.1 mag wide, r bins). What do you conclude? How would you numerically describe these curves (i.e. what kind of functional form for the fitting functions would you choose)?

5. What should be the faint r band limit for a survey to be able to map the $\ln(\rho)$ vs. Z profile out to 100 kpc using main-sequence stars? Assume the same color distribution as for the SDSS sample. For a solid angle of 1 deg^2, how many stars with $0.2 < g - r < 0.4$ would you expect with distances between 90 kpc and 100 kpc? Assume whatever additional information you need to solve this problem (not all required information is provided here).