ASTR222

SQL, Sloan, Hipparcos, and Gaia, due date: Tuesday, November 20, 2018

Part 1. The goal of part one of this assignment is to become familiar with the basic querying structure of SQL databases. This will be accomplished using the SDSS SQL interface to first run through the SQL tutorial, then perform science by exploring a common parameter relation for galaxies: The Faber-Jackson relation (a.k.a. the L-σ relation) for elliptical and S0 galaxies.

The starting point for this assignment is the SDSS SkyServer:

http://skyserver.sdss.org/dr14/en/home.aspx

A. Complete the SDSS SQL Tutorial located here:

   http://skyserver.sdss.org/dr14/en/help/howto/search/searchhowtohome.aspx

   The tutorial provides an introduction to SQL search commands and table structure, and should provide everything you need to complete part B.

B. Use the SDSS SQL interface to select a sample of elliptical galaxies which have measured velocity dispersions (veldisp) and measured photometry. Thankfully, the default query helps you out here quite a bit. You will extract values from the `PhotoObj` and `SpecObj` tables, and do a join on those tables. This should be set up in the default search, but if it's not (or just for your reference), the join should look like this:

```
<select some stuff>
FROM PhotoObj AS p
JOIN SpecObj AS s ON s.bestobjid = p.objid WHERE
<some search criteria>
```

In addition to the default fields, you'll also want to extract `petroMag_r`, `petroMagErr_r`, `extinction_r`, `petroMag_g`, `extinction_g`, `petroR50_r`, `petroR90_r`, `z` (redshift), and `velDisp`.

Be sure to `SELECT TOP 10` and return to HTML until you're confident your search works. Construct a query with the following criteria:

1. Select only elliptical and S0 galaxies. That's what petroR90_r and petroR50_r are for. These values are the radii containing 90% and 50% of the Petrosian flux. (Petrosian flux is a galaxy thing–it's a method of determining the extent of an object that does not have distinct edges.) The relation between the R90 and R50 terms tells you something about how "bulgy" a galaxy is. petroR90_r/petroR50_r > 2.6 is a reasonable rule of thumb value for selecting bulge-dominated galaxies.

2. To further ensure we have only ellipticals, we want extinction-corrected g-r > 1.0, i.e., red galaxies. The extinction correction is easy since the extinction_<FILTER> fields are in units of magnitudes, so don't overcomplicate this step.

3. Galaxies with r magnitudes (petroMag_r) in the range of 0 - 19 and r magnitude errors (petroMagErr_r) less than 0.05 mag.

4. Galaxies with measured redshift z < 0.35.

5. Galaxies with measured stellar velocity dispersion (veldisp) > 30 km/s and error (veldispErr/veldisp) < 0.2 (less than 20% error).

When you have a working query, you're ready to go! Change the number of lines returned from 10 to 10000, and change the format from html to csv. Depending on your browser, this will either automatically download the csv file or it will dump it into a browser window and you'll need to copy and paste it into a text file (don't try to use emacs or excel–trust me, they won't be happy about a 10000 line paste).

B. Now you're ready to do some science. Produce a plot of absolute R magnitude vs. stellar velocity dispersion. If you choose reasonable ranges for the X and Y axes, your plot should look like this one, only with a lot more data points!

https://en.wikipedia.org/wiki/File:Faber_Jackson.png

To calculate absolute R magnitude:

1. Correct r magnitude for extinction.

2. Since we restricted our query to low redshift galaxies, no cosmological correction is necessary–you can assume a simple Hubble law. Recall that:

   $v_{rs} = c\,z$ (where v_{rs} is redshift velocity)

   and

   $v = H_0 D$

   Assume $H_0 = 72$ km/s/Mpc

3. Also recall that $M = m - 5(\log_{10}D - 1)$ where distance is in parsecs.

Part 2. In this part of the assignment, we will resolve the Pleiades distance paradox. In order to do this, you will rely on the astroquery skills you employed in the second assignment and the SQL skills you have now mastered from part 1 of this assignment. Plus plotting.

A. Use Vizier to query the Hipparcos catalog (catalog='I/239/hip_main') to return all the stars within a 2 degree radius of the position of the Pleiades. Generate a plot of proper motion (pmRA vs. pmDE). This should allow you to easily identify members of the Pleiades vs. outliers. Try applying a cut in parallax (Plx) to see if you can limit the number of outliers, and pply any other parameter cuts you think might help limit the sample to members of the Pleiades. Look at the columns available in Hipparcos to see if any others will help.

B. Use the filtered list from part A to construct an H-R diagram (B-V vs. Absolute Vmag) for the Pleiades. You may ignore extinction for the purposes of this exercise, though note that you'll need to use parallax to determine distance and then compute absolute magnitude.

C. Now we'll do the same thing for Gaia DR2. Gaia can be queried through astroquery and accepts an SQL query (really an ADQL query) as a string. You can find an example of how to construct such a query online pretty easily. Construct the same two plots, only this time with MANY more points. You will need to choose different photometric bands for Gaia.

D. What is the average distance to the stars you identified as being in the Pleiades from Hipparcos? What is this distance from Gaia? Upon how many stars is the number based in each case? Are there any other interesting features you can identify in the Gaia HR diagram?

For part 1, please submit your SQL query and a copy of your plot. For part 2, please submit to me the 4 labeled plots from part D, any code you wrote in the process of doing this exercise, and written responses from part D (although you could also choose to just print those values on the plots). Please send all of these things in a single email.