I continued my work from last quarter by switching to data from the Gaia data release 2, refining my query with cuts made by Trick et al. in their 2018 paper, *"The Galactic Disc in Action Space as seen by Gaia DR2."* The plots in this paper were used as a reference to my own. I also continued to check my coordinate transformations and was able to successfully implement a transformation into galactocentric coordinates. I researched several methods of doing this, including the Python library "GALPY", which is meant for use with galactic dynamics. I ended up using the coordinates package from Astropy, as it was easier for me to understand and use, rather than attempting to use an entirely new library.

While I was working on figuring out the coordinate transformations, I noticed the query I was using for the Gaia database was "parallax > 300", which meant I was querying for objects beyond 300 pc, instead of within. I fixed this error, which increased the length of the data table I was working with from ~39,000 rows to ~420,000 rows. I proceeded to create plots to visualize the data, including 2D and 3D plots of velocity space, kernel density estimate plots, and k-means.

Given that the data table I was now working with was close to half a million rows, I implemented a random subsample of 3000 rows to make computations quicker. I then proceeded to apply Gaussian mixture models to the data, using the Python library Scikit-Learn. I made plots of both the full and random subsample of the data for a Gaussian mixture model, and a Bayesian Gaussian mixture model. These plots reveal some substructure within the data, but not clearly enough to identify particular clusters (as it is in Trick et al.). Further investigation is required.