CWRU DSCI453: Final Report

Ray Garner, CWRU Graduate Student

01 May, 2019

Contents

1 Abstract **1**

2 Introduction and Background **1**
2.1 Hertzsprung-Russell Diagrams 2
2.2 Gaia 4
2.3 Data Science Question 4

3 Data Science Methods **4**

4 Data Analysis and Discussion **10**
4.1 Cluster HR Diagrams 11
4.2 Cluster Half-Number Radii 15
4.3 Cluster Surface Density 17
4.4 Cluster Inner/Outer Separation 20
4.5 Cluster Radial Velocities 23

5 Conclusion **24**

6 Acknowledgements **24**

7 Extra Data **24**

8 References **25**

1 Abstract

In 2013, the European Space Agency launched the *Gaia* spacecraft. *Gaia* is a probe designed for astrometry: measuring the positions, distances, and motions of stars with unprecedented precision. In April 2017, the second data set was released (DR2), including positions, parallaxes, and proper motions for about 1.3 billion stars, as well as a multitude of other data. Using *Gaia* DR2, seven open clusters and seven globular clusters have been analyzed. Key results of the analysis include: open clusters are significantly younger than globular clusters, open clusters are much larger on the sky, and open clusters have lower radial velocity distributions than globular clusters.

2 Introduction and Background

Star clusters are groups of stars, usually distinguished into two types: globular and open. Globular clusters (GCs) are compact groups of hundreds or thousands of very old stars. Open clusters (OCs) are more loosely clustered and contain only a few hundred member stars. Open clusters become disrupted over time, compared to globular clusters tend to remain gravitationally bound unless the host galaxy disrupts them.

Star clusters are important objects in astronomy. Because all of the stars were born at nearly the same time, the different properties of the stars in a cluster are a function of only mass, so stellar evolution theories rely on open and globular clusters. A few of the nearest clusters are close engouh for their distances to be measured using parallax, allowing a Hertzsprung-Russell diagram to be plotted in absolute values.

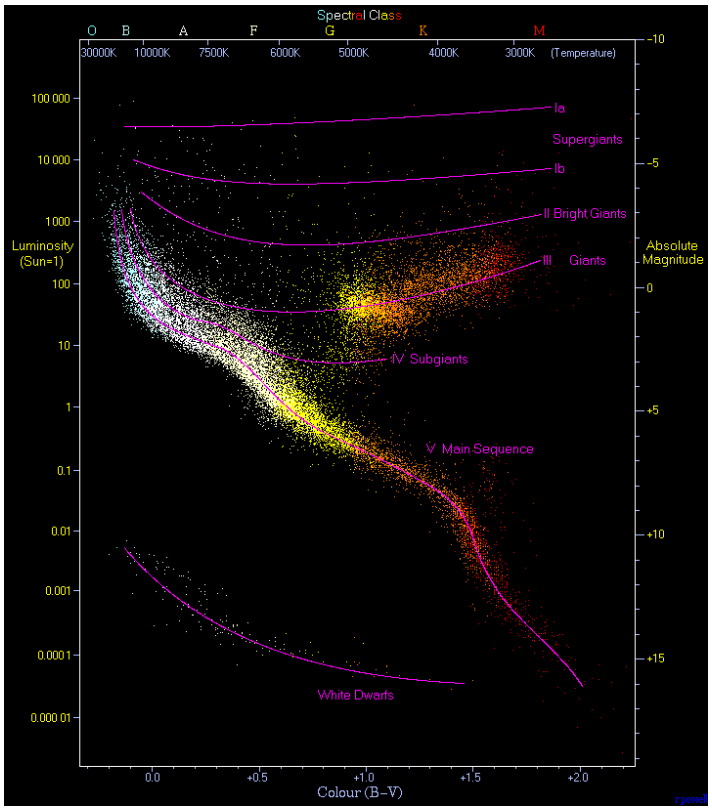
2.1 Hertzsprung-Russell Diagrams

These Hertzsprung-Russell (HR) diagrams require more explanation. An HR diagram is a plot showing the relationship between stars' absolute magnitudes or luminosities versus their stellar classifications or effective temperatures. More simply, it plots the star's brightness against its temperature.

There are two forms of HR diagrams. The original form designed by Hertzsprung (1908) and Russell (1913) displayed the spectral type of stars on the horizontal axis and absolute magnitude on the vertical axis. Although spectral type (O, B, A, F, G, K, M) is not quantitative, it is an effective measure of a star's effective temperature. Most modern diagrams replace the spectral type by the color index. This type of a diagram is commonly called an observational HR diagram, or a color-magnitude diagram. The other form of the diagram plots the effective surface temperature of the star on the horizontal axis and the luminosity of the star on the vertical axis. This is the theoretical HR diagram.

Although these two types of diagrams are similar, there is a distinction: the exact transformation from one form to the other is not trivial. To change between effective temperature and color index requires a color-temperature relation, and constructing that is difficult, being a factor of stellar composition and stellar rotation. When converting between luminosity and absolute magnitude, a bolometric correction is required, as well as the distance to the objects and the effect of interstellar extinction and reddening.

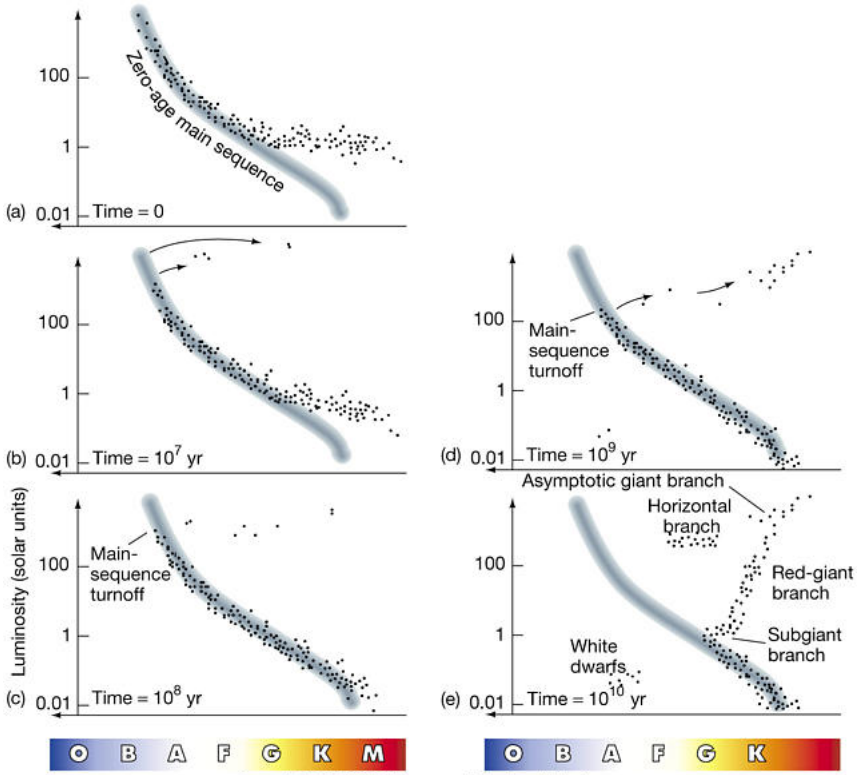
The usefulness in HR diagrams lie in studying stellar evolution since there are well-determined paths along which stars should travel as they evolve. In the figure below is an HR diagram using 22000 stars from the Hipparcos Catalogue and 1000 from the Gliese Catalogue. The Sun is located at one solar luminosity and color +0.63.



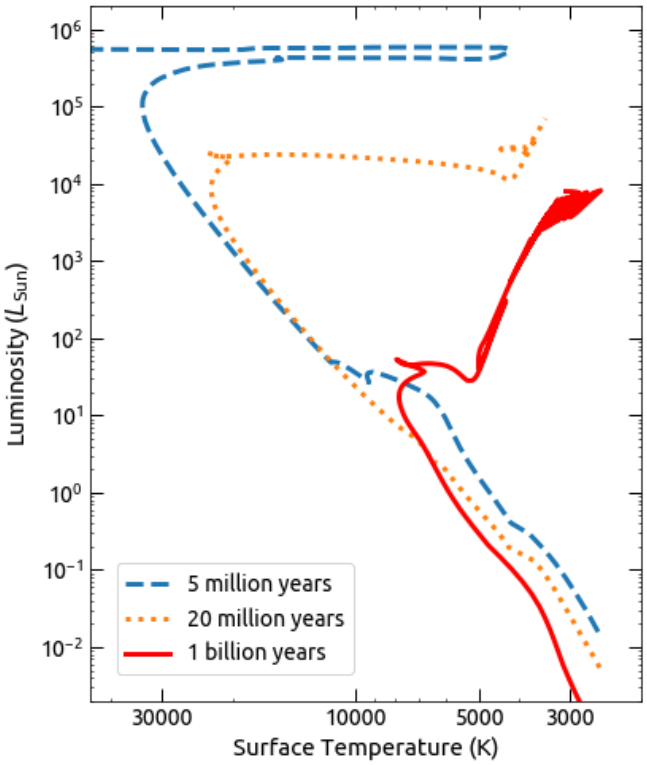
A large portion of these occupy a region in the diagram along the line called the main sequence. This is where stars begin their lives and star while burning hydrogen in their core. Stars found immediately above the main sequence are red giants (helium fusion in their core and hydrogen burning in a shell surrounding the core), whereas those below are white dwarfs.

The paths depend mostly on stellar mass; during evolution, both the effective temperature and radius (and thus luminosity $L = \pi R^2 \sigma T^4$) change. The highest mass stars, around 120 times that of the Sun, are at the upper left, while the lowest masses are at the bottom right of the main sequence. When the core hydrogen fuel is depleted, main sequence stars expand to become red giants or supergiants. The lower a star's birth mass, the longer it will remain on the main sequence. At the high-mass end, time spent on the main sequence is as short as two million years; lower mass stars remain on the main sequence for 12-13 billion years.

This time-dependent nature of an HR diagram is important for determining the ages of clusters. The stars in open and globular clusters are assumed to have been created all at the same time. Therefore, the presence of stars off the main sequence indicate a cluster that is old enough to have stars evolve off the main sequence. As time goes on, the main sequence turnoff point, where stars transition from main sequence stars to red giants, moves lower and lower along the main sequence. See the figure below.



Theorists can use a stellar evolution code to compute evolutionary tracks of stars for a number of different masses, connecting those models at the same age. These lines are called isochrones. By comparing these lines to observational HR diagrams, astronomers can estimate the ages of clusters. In the figure below are a few generated isochrones for stars ranging from five million years to one billion years old.



2.2 Gaia

The *Gaia* space observatory was launched by the European Space Agency in December 2013. The spacecraft is designed for astrometry: measuring the positions, distances, and motions of stars with unprecedented precision. Since its launch, the *Gaia* team has released two datasets, Data Release 1 (DR1) and 2 (DR2). DR1 is the culmination of fourteen months of observations and contains the "positions and ... magnitudes for 1.1 billion stars using only *Gaia* data; positions, parallaxes and proper motions for more than 2 million stars" based on a combination of *Gaia* and Tycho-2 data (Gaia Collaboration, 2016). DR2, which this project uses, is based on twenty-two months of observations and contains more data at a higher precision with more predictors (Lindegren et al. 2018).

2.3 Data Science Question

The goal of this project is to confirm and challenge results commonly found in the literature about OCs and GCs in general, and about these clusters in particular.

3 Data Science Methods

This entire project was completed in the coding language Python. Python is the most used coding language in the astronomy community. Out of a survey of research astronomers at all levels of academia, Momcheva & Tollerud (2015) found that 67% use Python, compared to the 6% that use R. For that reason, this project is being completed exclusively in Python.

The data from *Gaia* DR2 can be accessed through the *Gaia* Archive. The Archive utilizes Astronomical Data Query Language (ADQL), which requires a position on the sky (right ascension and declination in degrees)

and a circular area around the sky defined by some radius. The data was downloaded using Python code provided by CWRU Professor Chris Mihos. Radii were chosen so as to grab a majority of the stars in the clusters while minimizing the presence of field stars.

The ADQL query allowed us the choice of which parameters to download for each star. We chose the following parameters:

- Source ID, the ID given by the *Gaia* catalog;
- Right ascension (RA), the angular distance eastward along the celestial equator, in degrees;
- Right ascension error, the error in the right ascension, in milliarcseconds;
- Declination (Dec), the angular distance north or south of the celestial equator, in degrees;
- Declination error, the error in the declination, in milliarcseconds;
- Parallax, the displacement in apparent position, in milliarcseconds;
- Parallax error, the error in the parallax, in milliarcseconds;
- Proper motion in RA, the apparent motion of the stars in the RA direction, in milliarcseconds per year;
- Proper motion in Dec, the apparent motion of the stars in the Dec direction, in milliarcseconds per year;
- Photometric G mean magnitude, the brightness of the star through the G filter, in magnitudes;
- $B_P - R_P$, the color of the star, in magnitudes;
- Radial velocity, the velocity of a star in the direction of the Sun, in kilometers per second;
- Radial velocity error, the error in the radial velocity, in kilometers per second;
- Effective temperature, the temperature of a black body that would emit the same total amount of electromagnetic radiation as the star, in kelvins; and
- A_G, the line-of-sight extinction in the G-band, in magnitudes.

The following table indicates which clusters were chosen, as well as the right ascension, declination, and radius used in the ADQL query.

Cluster	Type	RA..Deg.	Dec..Deg.	Radius..Deg.
Pleiades (M45)	OC	56.850000	24.116700	4
Hyades	OC	66.750000	15.866000	10
Coma	OC	185.625000	25.850000	14
IC 2602	OC	160.739583	-64.394167	4
IC 2391	OC	130.150000	-53.033000	4
M44	OC	130.100000	19.983000	4
NGC 752	OC	29.420830	37.785000	4
47 Tuc	GC	6.023625	-72.081278	2
Omega Centauri	GC	201.697000	-47.479472	1
M3	GC	205.548417	28.377278	1
M4	GC	245.896750	-26.525750	1
M5	GC	229.638417	2.081028	1
M9	GC	259.799083	-18.516250	1
NGC 1851	GC	78.528167	-40.046560	1

Once the data has been downloaded, it is saved in a CSV file for later use. For the remainder of this report, we will follow two clusters when illustrating methodology: the Pleiades and Omega Centauri. The full datasets for each cluster contains tens of thousands, often hundreds of thousands of stars. A subsection of the data can be seen below.

source_id	ra	ra_error	dec	dec_error	parallax	parallax_error	pmra
6.956253e+16	53.19046	0.0492852	25.76943	0.0302695	1.840969	0.0550270	12.188403
6.936026e+16	53.72950	0.0450155	25.61766	0.0301324	1.465799	0.0492010	2.841123
6.933115e+16	52.88175	0.1176859	25.16132	0.0747173	4.496478	0.1207077	4.950795
6.925522e+16	53.63708	0.0915370	24.94171	0.0609059	2.082618	0.0947316	23.084884
6.946165e+16	53.74380	0.0378004	25.82328	0.0274956	1.830493	0.0416930	20.320559
6.933789e+16	52.89909	0.1182796	25.34185	0.0719294	8.004872	0.1233882	-1.777644

pmdec	phot_g_mean_mag	bp_rp	radial_velocity	radial_velocity_error	teff_val	a_g_val
-8.638178	15.51943	1.555731	NA	NA	4378.333	0.4885
-4.581655	15.66517	1.369809	NA	NA	4726.910	0.3695
-19.808026	16.74960	2.560807	NA	NA	3316.000	0.3883
-25.707813	16.62549	2.126810	NA	NA	3600.000	NA
-21.236705	14.96781	1.419835	NA	NA	4737.750	0.8430
-42.028849	16.08086	2.837419	NA	NA	3921.735	0.5170

source_id	ra	ra_error	dec	dec_error	parallax	parallax_error	pmra
6.086788e+18	200.7005	0.7982284	-46.74754	0.4262193	0.9399475	0.6995934	-7.735789
6.086757e+18	200.5528	0.1747085	-47.22038	0.1082294	0.2608397	0.1688660	-12.058067
6.086767e+18	200.9685	0.7492705	-47.00994	0.5289368	0.9465377	0.6942569	-7.349937
6.086792e+18	201.1340	0.4989173	-46.87464	0.3825473	1.3928425	0.5609263	-4.649787
6.086788e+18	200.7979	0.0500064	-46.72924	0.0297321	0.3407293	0.0454473	-6.748000
6.086786e+18	200.9545	0.1822438	-46.73352	0.1325751	0.5105053	0.1947614	-7.059769

pmdec	phot_g_mean_mag	bp_rp	radial_velocity	radial_velocity_error	teff_val	a_g_val
-1.4892270	19.92348	0.9486084	NA	NA	NA	NA
-0.5858103	17.84809	1.2045288	NA	NA	NA	NA
1.0378815	20.20513	1.8943806	NA	NA	NA	NA
-2.4325116	19.66760	1.9292316	NA	NA	NA	NA
2.9269670	15.07769	1.1335726	NA	NA	4953	NA
-2.8582041	17.97304	1.2437038	NA	NA	NA	NA

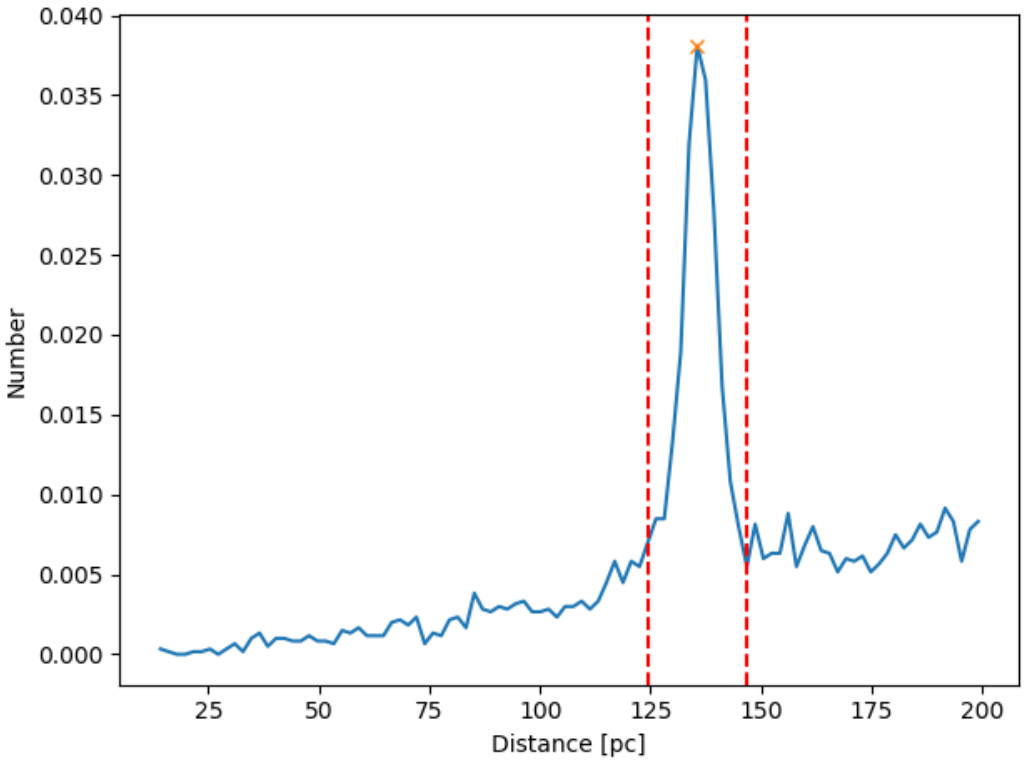
First, we have to clean the data. Many of the clusters contain stars with a negative parallax. This is physically unreasonable, and results as a fit of noisy observations to a model that connects the motion of the star through space and the motion of the observer around the Sun. Technically, to completely account for these negative parallaxes, we should apply a Bayesian fit to the data (Luri et al. 2018), but that is out of the scope of this project. Upon further inspection, those stars with negative parallaxes are only a small subset of the data. We choose to ignore these and work only with stars with reported parallaxes greater than zero.

We convert the remaining stars' parallaxes to a distance using the relation

$$d = \frac{1}{p},$$

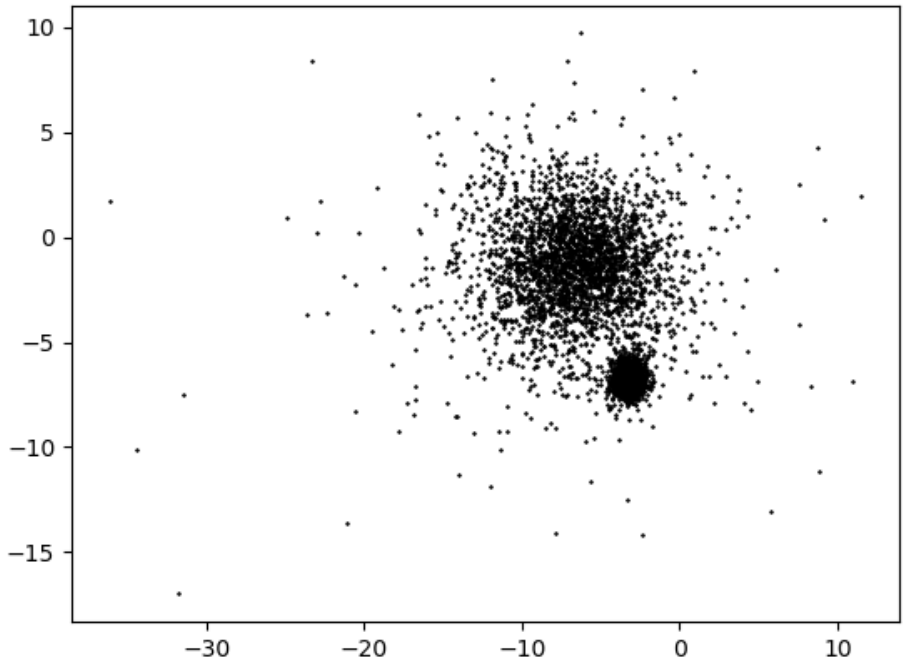
where d is the distance to the star in parsecs and p is the parallax of the star in arcseconds. The final step in cleaning the data is to remove stars that are definitely not in the cluster. Although the literature disagrees as to the exact distance value, generally-speaking, there is a distance above which none of the stars in the cluster. For the Pleiades that cut off is at 500 parsecs; for Omega Centauri that cut off is between 4000 and 6000 parsecs.

The OCs require another distance selection. Given that the data can be visualized as a cylinder of arbitrary length with a radius as given in the table above, a histogram of the data should include a peak at the distance of the cluster. This is exactly what we observe. The histogram of the Pleiades dataset distances is below.



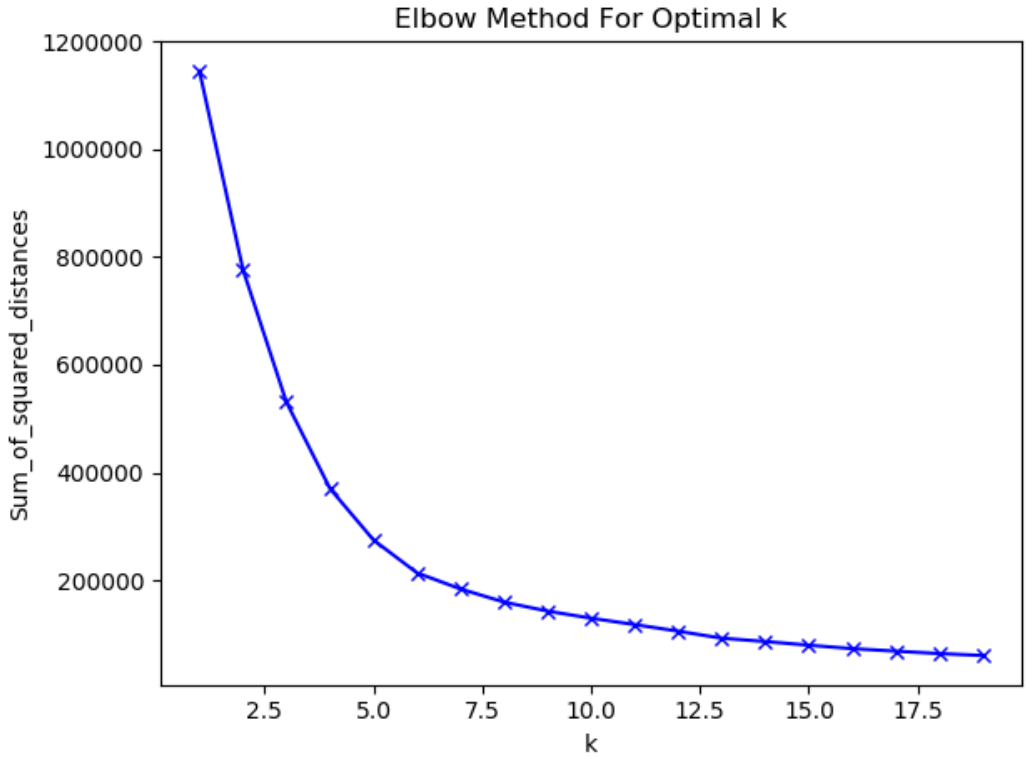
Note the peak at about 130 parsecs, roughly the distance to the Pleiades. If we approximate this peak as a Gaussian distribution, we can select the stars that fall entirely within the peak, $\pm 6\sigma$. This same procedure was performed on each of the OCs.

The GCs are not as simple. Since they are so far away, $\geq$ 4000 parsecs, a histogram of their distances does not result in a sharply defined peak as there are more stars far away than near. Following the procedure lined out in Gaia Collaboration (2018), we investigate the proper motion plots and cut. These stars are all in a comoving cluster, therefore they should all have the same proper motion and the cluster should appear in proper motion space.

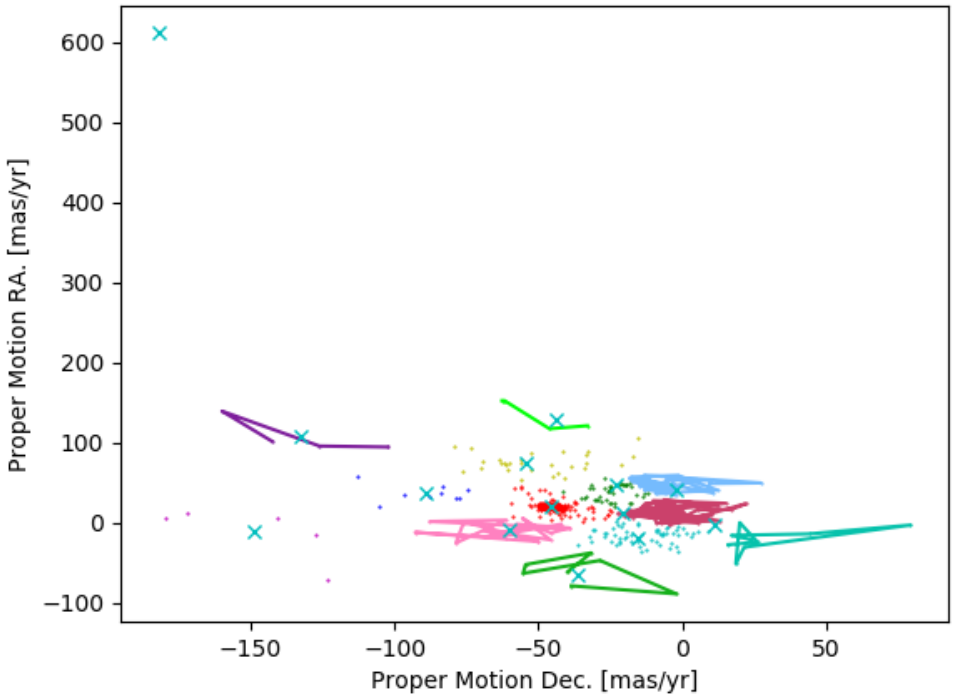


This is a plot of the proper motion in both RA and Dec for the Omega Centauri dataset, pre-selected for stars between 4000 and 6000 parsecs. Note the tight cluster centered around (-4,-7); this is Omega Centauri. All of the GCs were selected similarly.

The OCs also underwent a proper motion cut. Their proper motions were much more scattered than the GCs, so a k-means algorithm was used. First, an elbow diagram was created to aid in selecting the correct number of clusters.



There was no "elbow" in the diagram, so k was chosen semi-randomly and modified to fit the cluster that appears in the proper motion plot. The Pleiades utilized $k = 14$ clusters; the other OCs utilized between 6-8 clusters. The results of the k-means plot can be seen below for the Pleiades.



Each cluster is color-coded differently, with the cyan crosshairs marking the cluster centroids. The red cluster at the bottom center is the Pleiades. Of the other OCs, only NGC 752 required more detail. That dataset required two rounds of clustering: once to get the center of the dataset and another to isolate the specific cluster corresponding to NGC 752. Both algorithms used $k = 8$. At the end of all of these cuts, the datasets have been reduced to approximately 4% of their original length.

At this point, both OCs and GCs have been thoroughly cleaned and are now ready for data analysis. The next steps are to plot HR diagrams and fit isochrones so as to determine the ages of the clusters. This also serves as a good check as to whether or not our data matches the data gathered previously. Half-number radii will also be calculated for each of the clusters. This is another check, but will also be used in the analysis of their surface densities. Finally, by splitting the clusters by their half-number radii, we can explore the differences in the inner and outer parts of each cluster.

4 Data Analysis and Discussion

First, we fit a normal distribution to the datasets' distance to get a distance estimate. This is summarized in the table below, with references to the literature.

Cluster	Distance..pc.	Distance.Error..pc.	Lit.Distance..pc.	Lit.Error..pc.	Reference	Percent.Error
Pleiades (M45)	135.874	4.337	134.400	2.9	Galli et al. (2017)	1.097
Hyades	47.340	3.147	47.000	2.0	Majaess et al. (2011)	0.723
Coma	85.181	2.180	85.000	6.0	Majaess et al. (2011)	0.213
IC 2602	152.073	3.465	166.725	NA	Pandey et al. (2010)	8.788
IC 2391	152.334	4.919	176.000	NA	Kharchenko et al. (2005)	13.447
M44	186.937	5.464	187.000	NA	Kharchenko et al. (2005)	0.034
NGC 752	441.351	7.223	458.000	NA	Kharchenko et al. (2005)	3.635
47 Tuc	3981.377	287.668	4690.000	136.0	Woodley et al. (2012)	15.109
Omega Centauri	4880.538	590.520	4840.000	340.0	van de Ven et al. (2006)	0.838
M3	10431.545	891.313	10000.000	NA	Forbes et al. (2008)	4.315

M4	2006.598	325.388	2200.000	NA	Forbes et al. (2008)	8.791
M5	7467.658	582.828	7500.000	NA	Paust et al. (2010)	0.431
M9	7901.217	545.156	7900.000	NA	Boyles et al. (2011)	0.015
NGC 1851	20152.744	8886.815	12000.000	NA	Forbes et al. (2008)	67.940

The distances calculated via the normal distribution reasonably match the distances calculated in the literature. NGC 1851 has the largest percent error with the literature value at 68%. However, the standard deviation gives a distance that lies near the literature value as reported by Forbes et al. (2008). Some of the stars might not be tied to this cluster.

4.1 Cluster HR Diagrams

To create HR diagrams for each of the clusters, the datasets' apparent G-band magnitude, m_G, had to be converted to an absolute G-band magnitude, M_G:

$$M_G = m_G - 5(\log(d) - 5),$$

where d is the distance to the cluster in parsecs. This allows the color index to be plotted against the absolute magnitude to create the HR diagram. Isochrones were also fitted to the data. Isochrones were generated by the Padova and Trieste Stellar Evolution (PARSEC) code. The only parameters changed from the default were the photometric system ("Gaia's DR2 G, G_BP and G_RP (Vegamags, Gaia passbands from Weiler 2018)"), the ages, and, for the GCs, the metallicity. Data files were generated for isochrones corresponding to stellar populations 1 Myr, 10 Myr, 100 Myr, and 1 Gyr to be used for all clusters, as well as isochrones corresponding to the ages of the clusters as reported in the literature. The table below reports the literature ages for each cluster.

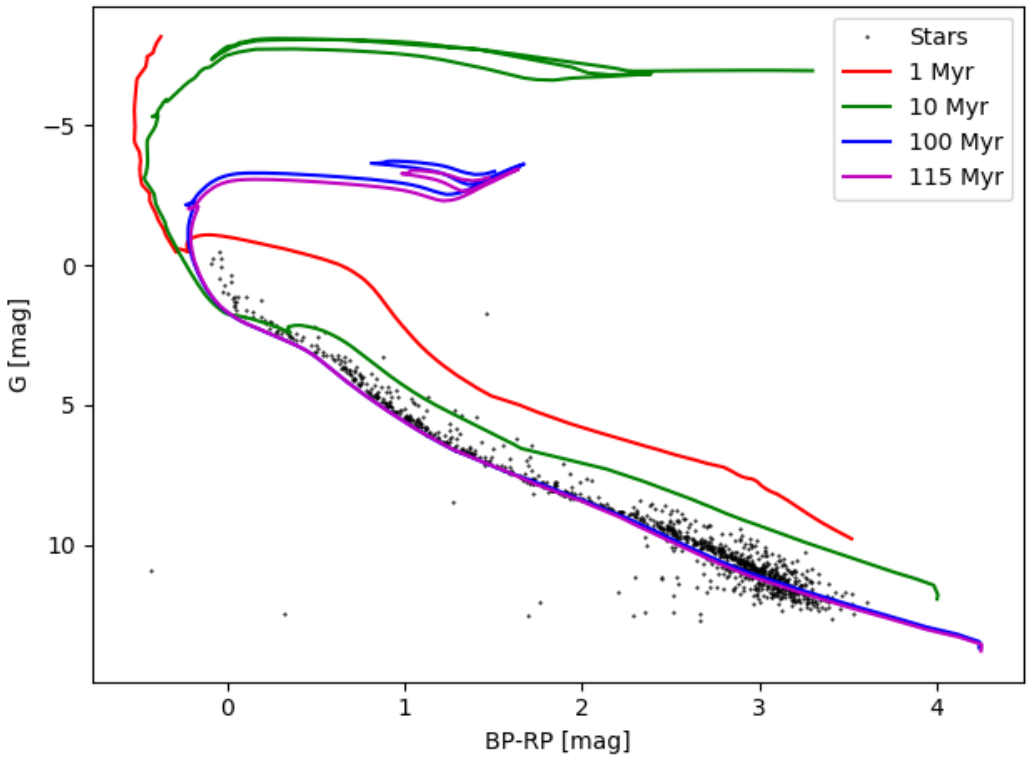
Cluster	Age..Gyr.	Reference
Pleiades	0.115	Basri et al. (1996)
Hyades	0.625	Perryman et al. (1998)
Coma	0.450	Odenkirchen et al. (1998)
IC 2602	0.050	Barrado y Navascués et al. (2004)
IC 2391	0.050	Barrado y Navascués et al. (2004)
M44	0.600	Kraus & Hillenbrand (2007)
NGC 752	1.340	Agüeros et al. (2018)
47 Tuc	13.060	Forbes & Bridges (2010)
Omega Centauri	11.520	Forbes & Bridges (2010)
M3	11.390	Forbes & Bridges (2010)
M4	12.200	Caputo et al. (1985)
M5	10.620	Forbes & Bridges (2010)
M9	12.000	Koleva et al. (2008)
NGC 1851	9.200	Koleva et al. (2008)

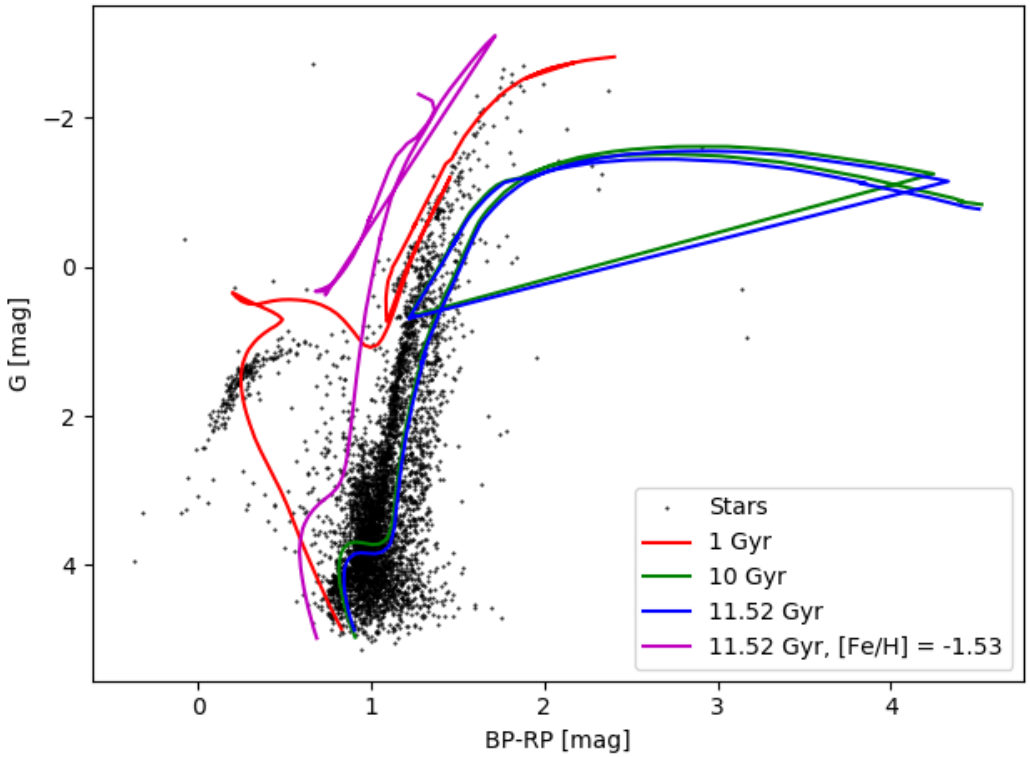
Metallicity was only taken into account for the GCs. GCs, being much older than OCs, have significantly older stellar populations. These stars were created at a time when the universe was not as enriched by metals, resulting in a lower metallicity. This has the effect of making the stars appear redder, which the isochrones need to take into account to give a good fit. OCs have near-zero metallicities and are thus considered negligible. The reported metallicity values for each GC are in the table below. All values are from Harris (1996; 2010 edition).

Cluster	Metallicity
47 Tuc	-0.72
Omega Centauri	-1.53
M3	-1.50
M4	-1.16
M5	-1.29

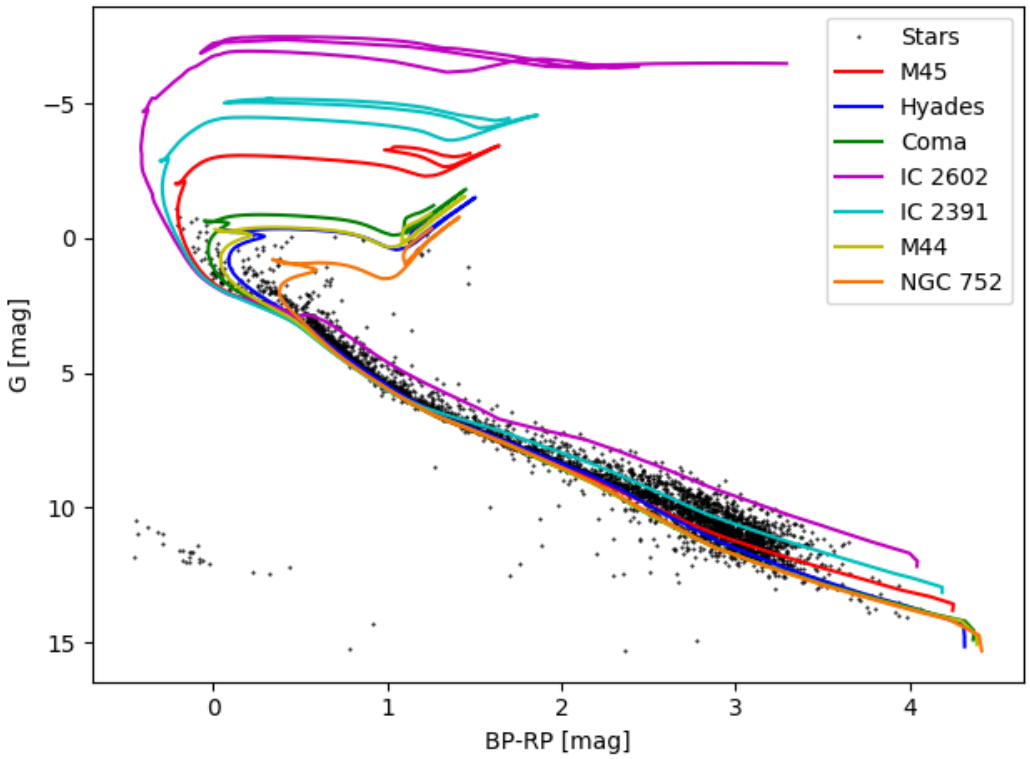
M9	-1.77
NGC 1851	-1.18

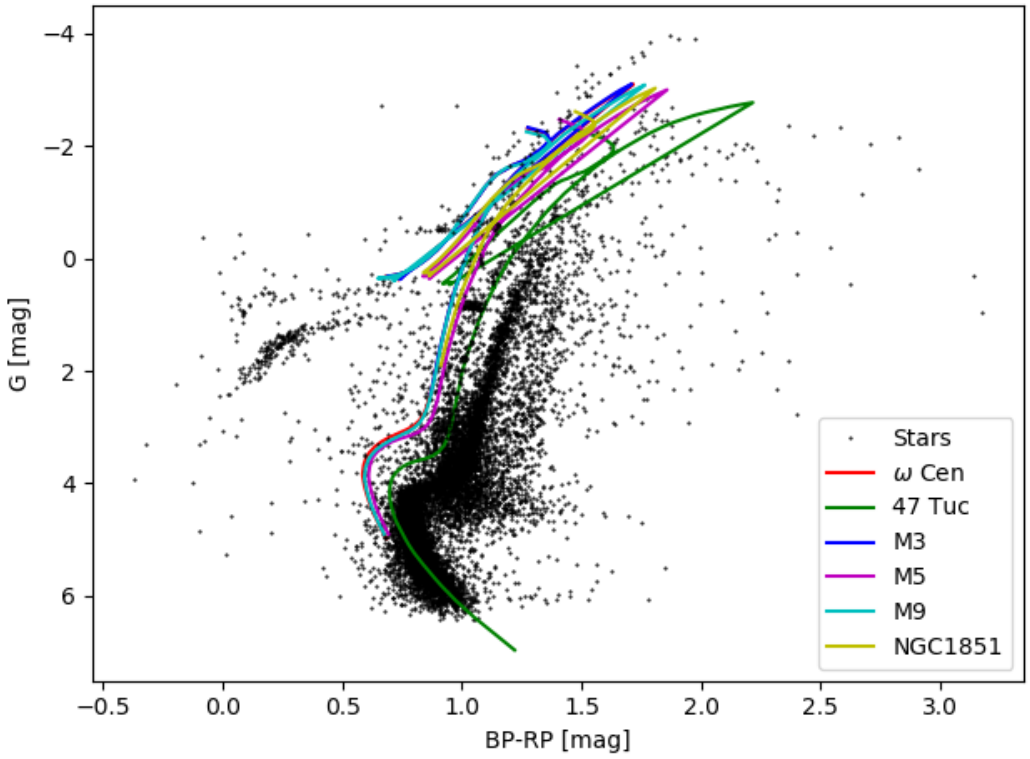
The completed HR diagram for the Pleiades and Omega Centauri is below.





The combined HR diagram for the OCs and GCs is below as well.





It is important to note that the data gathered from the *Gaia* Archive has not been corrected for interstellar reddening. Correcting for this could shift the stars towards the bluer colors (to the left), allowing the isochrones and main sequence to correlate stronger. The data has also not taken into account any binary stars, which would have an artificially increased luminosity.

Some general properties we note are that the OCs have a wider range of ages and are much younger than the GCs. The ages of the OCs range from the youngest, IC 2602 and IC 2391 at 13.7 and 50 Myr, respectively, to the oldest, NGC 752 at 1.34 Gyr.

The GCs are much less precise. There are a few properties of GCs that affect the ability of the isochrones to accurately fit them. GCs are much further away than open clusters, allowing more reddening dust to scatter blue light, shifting the stars to the right on the HR diagram. This is especially true of Messier 4 (M4), where the stars have been shifted so far to the right that the isochrones do not fit the stars at all, resulting in the omission from the total GC HR diagram. They are also more populous than OCs, overloading the instruments on *Gaia* when measuring stars at the centers of the clusters, thus reducing the number of stars in the data. Perhaps the worst offender is Omega Centauri.

However, the HR diagrams for the GCs still shows generally what we would expect. Since there is a very small main-sequence and a large RGB, asymptotic giant branch (AGB), and small horizontal branch (HB), this confirms the old age of the GCs. In fact, the overlapping isochrones indicates that they are all of roughly the same age, 10-13 Gyr, about the age of the universe.

Although many of the clusters fit their isochrones well with the reported ages, there are a few that are noticeably bad fits. IC 2602 has a reported age of 50 Myr, but the best-fit isochrone is at 13.7. This is a discrepancy that is not explained in the literature. Another method that measures the lithium depletion in the lowest-mass stars in the cluster gives the reported age. It is unsure what is causing this discrepancy. Perhaps the presence of or lack of lithium does not affect the star's luminosity as much as the theory suggests.

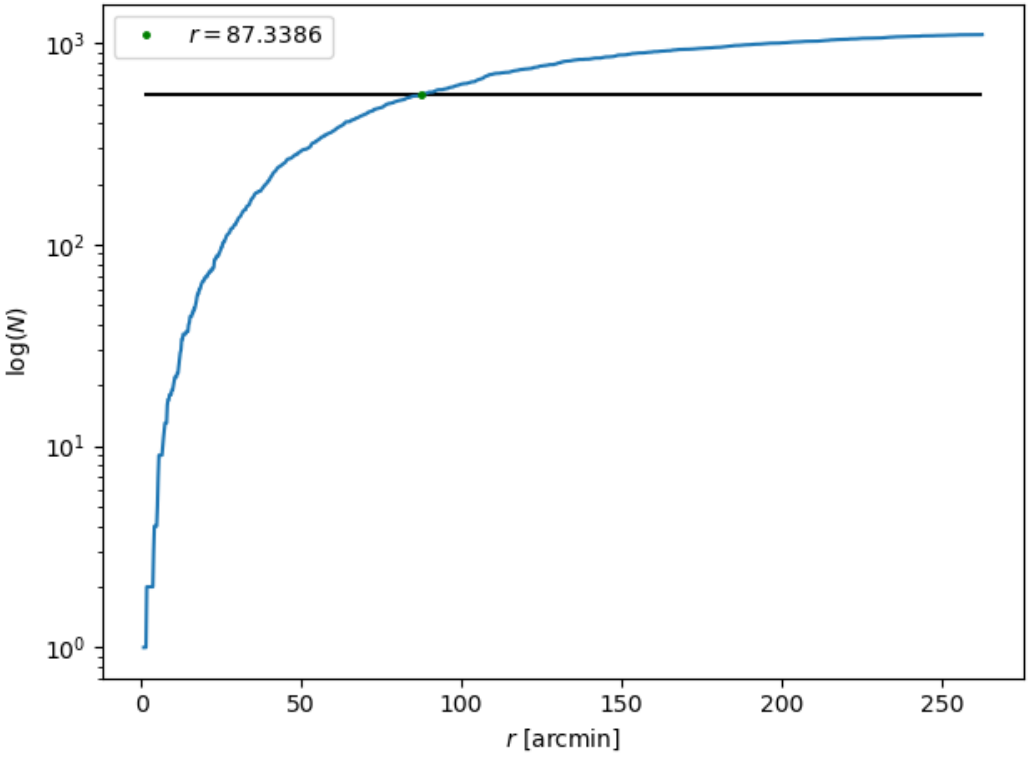
Messier 4 is not the only GC that has a bad isochrone fit. The HR diagram of Messier 9 (M9) does not seem to have any definitive tracks of stars. There might be a horizontal branch and perhaps a main sequence, but

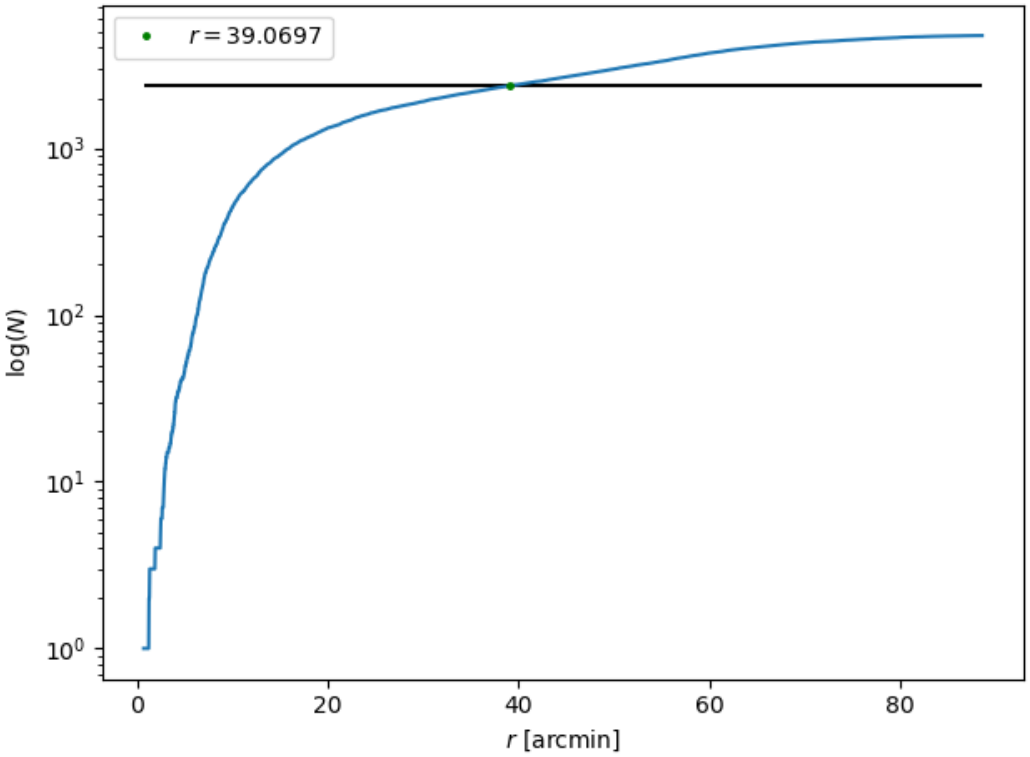
neither are tight fits. These might be field stars, but there is no reason why they should be, considering the proper motion and distance cuts are correct.

4.2 Cluster Half-Number Radii

The half-number radius is defined as the radius at which half of the stars in the cluster are contained. First, we need to calculate the distance from the center of the cluster for each star. This requires converting the RA and Dec in degrees into an RA and Dec in arcminutes. The center of each cluster is the RA and Dec positions defined in the table above. The distance from the center is a simple Euclidean distance using the Pythagorean Theorem.

Sorting the datasets by the distance from the center allows us to make a cumulative histogram. The half-number is half of the total number of stars in the cluster. For the Pleiades, the total is 1103 stars; for Omega Centauri it is 4379. The half-number radius was found by inspecting the plots of distance from the center vs number. The plots for the Pleiades and Omega Centauri are shown below.





This is yet another check of the data. The table below tabulates the sizes of the clusters as derived by the half-number radius and the apparent dimensions as measured on Earth.

Cluster	Half.Number.Radius..arcminute.	Apparent.Size..arcminute.	Percent.Error
Pleiades	87.33860	110.00	20.601
Hyades	280.98800	330.00	14.852
Coma	256.27800	225.00	13.901
IC 2602	161.41200	25.00	545.648
IC 2391	85.98320	25.00	243.933
M44	76.42580	47.50	60.896
NGC 752	42.86470	37.50	14.306
47 Tuc	22.95750	30.90	25.704
Omega Centauri	39.06970	36.30	7.630
M3	3.62129	9.00	59.763
M4	6.33025	13.00	51.306
M5	3.70101	11.50	67.817
M9	16.50800	4.65	255.011
NGC 1851	2.39802	6.50	63.107

Generally, the OCs have low percent errors with the observed apparent sizes. The two notable exceptions are IC 2602 and IC 2391. It is likely that the data still has several field stars, which would make the cluster appear larger than it is. The GCs have reasonable percent errors, the errors being caused by, once again, the instruments on *Gaia* getting overloaded at the central bright portions of the clusters. There are many more stars there, which, if included, would lower the half-number radius. Perhaps the apparent size and half-number radius are measurements of different concepts.

4.3 Cluster Surface Density

The surface density (in units of number per square arcminute) measures the "concentration" of stars in the cluster. We calculate the surface density by computing a histogram of the datasets' distances from the cluster center and dividing that by the circular area at that bin of some radius (in units of arcminutes). The surface area histogram is then the distance histogram divided by the binned circular area.

To describe the relationship between the surface area and the distance from the center of the cluster, we fit three profiles. The first is a logarithmic relationship between surface density, σ, and radius, r

$$\log \sigma = ar + b,$$

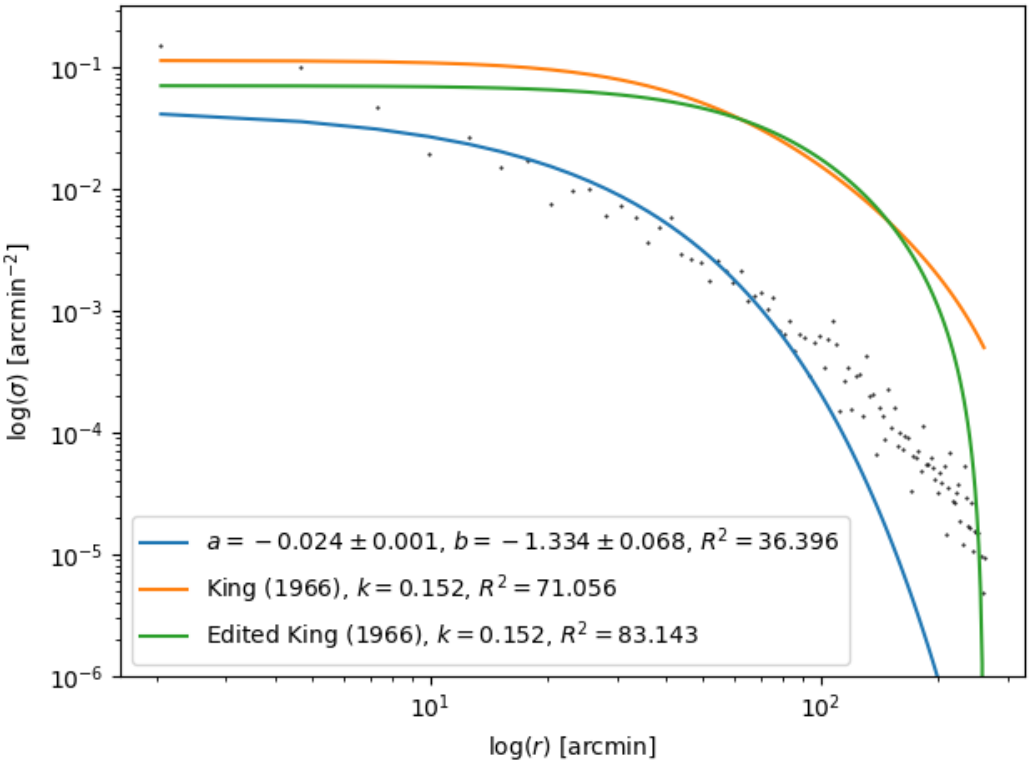
where a and b are linear coefficients. Those coefficients are found by using Numpy's PolyFit function in Python which performs a least squares polynomial fit. The second profile is the King (1966) profile,

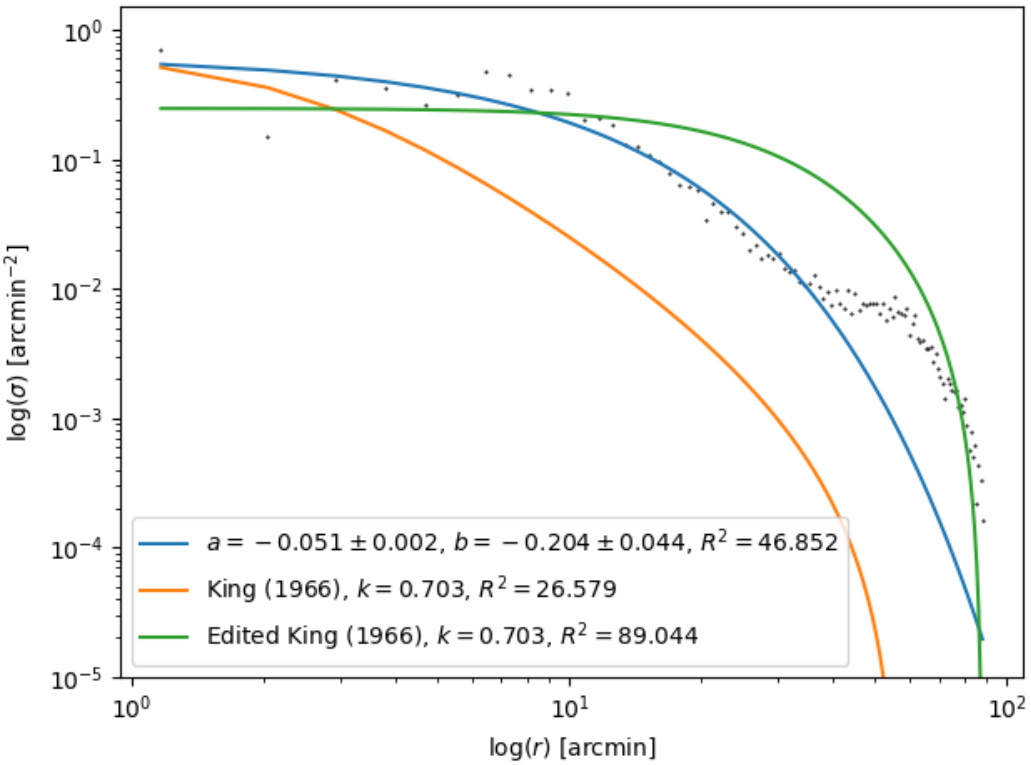
$$\sigma = k \left(\frac{1}{(1 + (r/r_c)^2)^{1/2}} - \frac{1}{(1 + (r_t/r_c)^2)^{1/2}} \right),$$

where k is the central density, r_c is the core radius, and r_t is the tidal radius. Both the core radius and tidal radius are defined by the literature. The central density is the surface density at the first binned radius. The King profile assumes that the velocity of the stars in the cluster is isotropic everywhere. This is generally true for globular clusters, but not necessarily true for open clusters. The core radius is the distance at which the apparent surface luminosity has dropped by half. The tidal radius is the distance at which the gravitation of the galaxy has more influence on the stars than the gravitation of the globular cluster. The third profile is an "edited" King profile, where r_c is replaced by the half-number radius and r_t is replaced by the radius at the total number of stars. The table below tabulates the reference data used for the core and tidal radius.

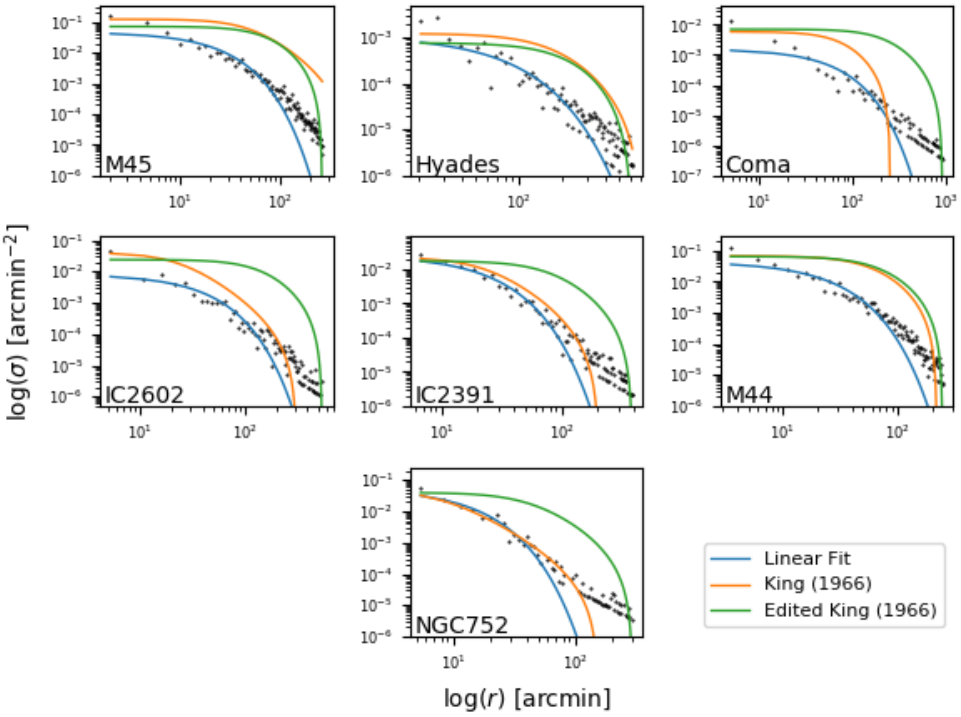
Cluster	Core.Radius..arcmin.	Tidal.Radius..arcmin.	Core.Reference	Tidal.Reference
Pleiades	50.55	378.000	Adams et al. (2001)	Converse & Stahler (2010)
Hyades	226.62	656.400	Röser et al. (2011)	Röser et al. (2011)
Coma	96.00	383.820	Odenkirchen et al. (1998)	Piskunov et al. (2008)
IC 2602	30.00	316.230	Kharchenko et al. (2005)	Piskunov et al. (2008)
IC 2391	21.00	207.114	Kharchenko et al. (2005)	Piskunov et al. (2008)
M44	36.00	343.488	Kharchenko et al. (2005)	Piskunov et al. (2008)
NGC 752	22.20	155.880	Kharchenko et al. (2005)	Piskunov et al. (2008)
47 Tuc	0.36	42.860	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)
Omega Centauri	2.37	57.030	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)
M3	0.37	38.190	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)
M4	1.16	32.490	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)
M5	0.44	28.400	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)
M9	0.45	8.160	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)
NGC 1851	0.09	11.700	Harris (1996; 2010 edition)	Harris (1996; 2010 edition)

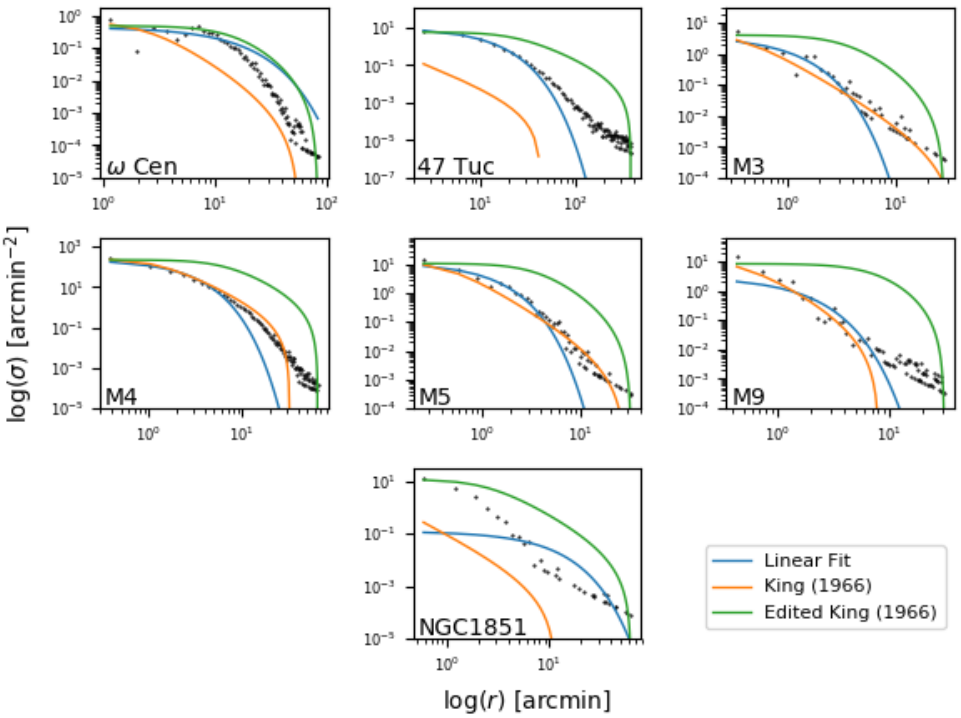
The results of these profile fits to the Pleiades and Omega Centauri are shown below, as well as the R^2 values for those fits as calculated by the Scipy Stats.LinRegress function in Python.





Below is a summary of the surface density profiles for each cluster, as well as a table summarizing the R^2 values for each fit.





Cluster	Linear.R.2	King.R.2	Edited.King.R.2
Pleiades	36.396	71.056	81.143

Hyades	69.047	85.343	88.945
Coma	25.992	75.848	78.301
IC 2602	28.355	46.120	79.727
IC 2391	24.056	98.155	71.245
M44	36.078	61.159	80.851
NGC 752	18.213	57.148	58.254
47 Tuc	46.852	26.579	89.044
Omega Centauri	10.664	33.201	34.607
M3	16.711	11.719	54.969
M4	13.108	23.473	47.314
M5	16.382	12.256	52.084
M9	21.539	35.497	77.084
NGC 1851	48.501	26.256	24.545

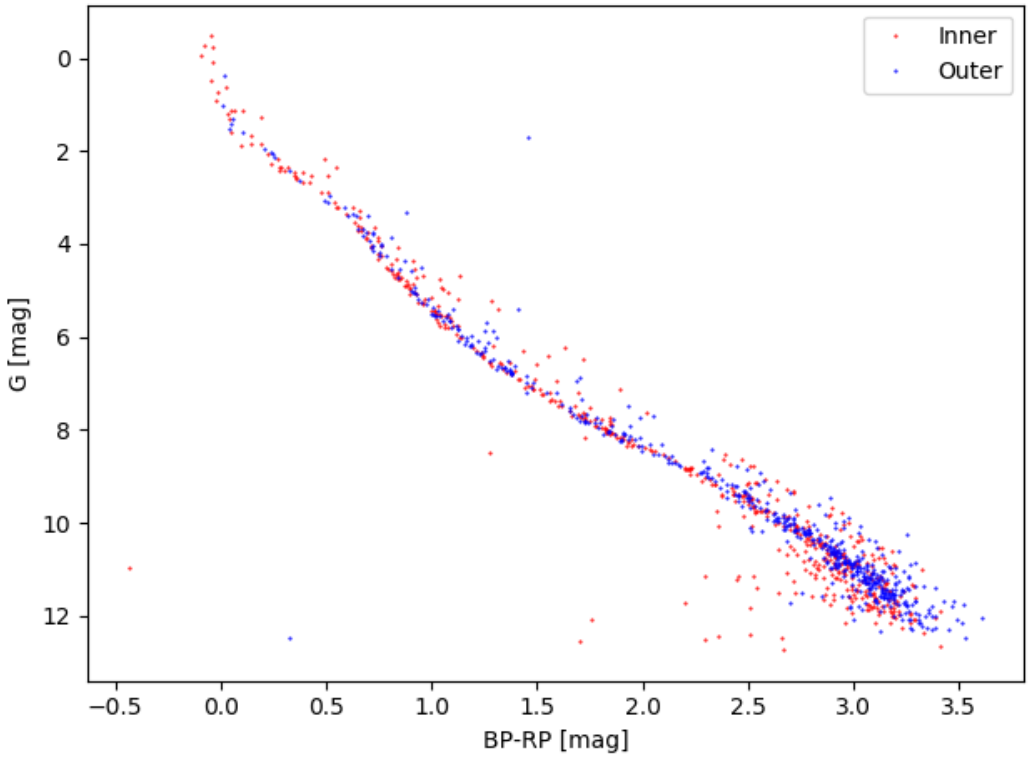
By looking at the R^2 values, it appears that the King model fits the OCs much better than the GCs. However, by investigating the plots, it appears the opposite. Furthermore, the "edited" King model accounts for the variability for all of the clusters the best. However, this should be taken warily, as the "edited" King model has parameters chosen specifically for the data; it is expected that it would explain the data well. It is more likely that the plots are correct and, on average, the King model fits the GCs best.

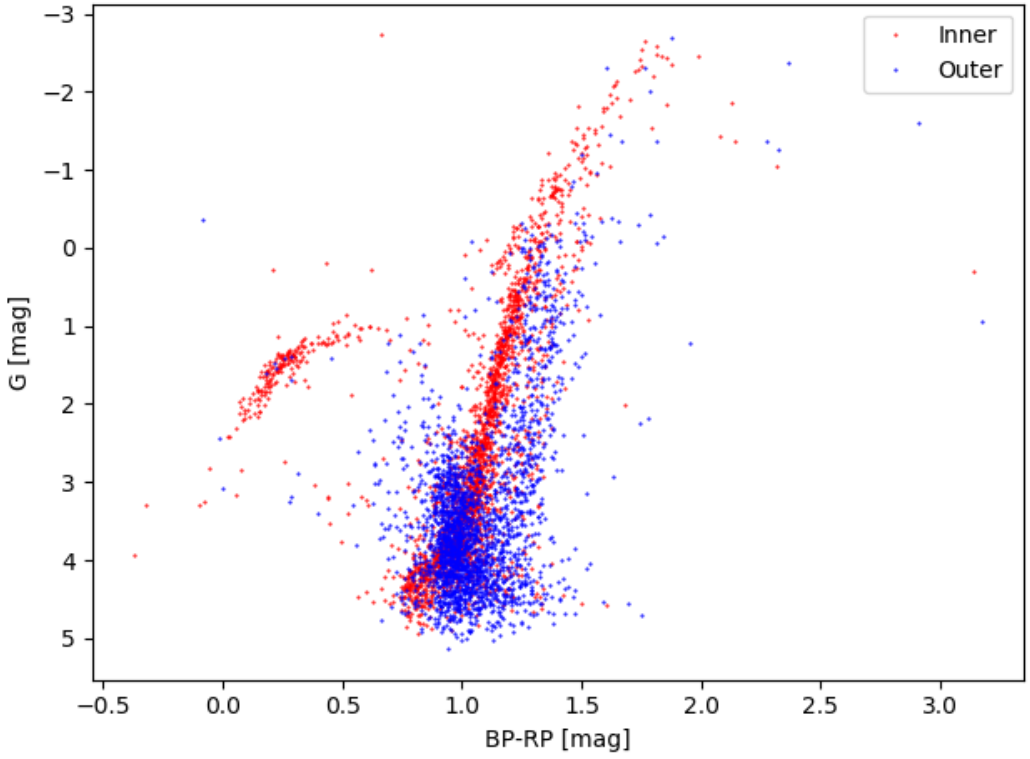
However, there are exceptions. It is likely that NGC 752 is fitted better by the King model. NGC 752 might be undergoing evaporation; the low mass stars have been tossed out of the cluster (Giardino et al. 2008). As the cluster makes more passes around the disk of the Milky Way, it will encounter spiral density waves which will cause more stars to slow down or speed up and disrupt the cluster. In the Pleiades, the King model had $R^2 = 72.393\%$, while the custom model had $R^2 = 36.396\%$. It is possible that the expansion of the Pleiades is isotropic everywhere, though more research in the literature needs to be done to confirm this.

In Omega Centauri, the custom model had $R^2 = 64.403\%$ while the King model had $R^2 = 26.470\%$. This is likely due to the overexposure problem mentioned before, as can be seen in the widely varying data towards the center of the cluster. The King model did not fit the globular cluster NGC 1851 well ($R^2 = 26.256\%$), likely due to the lack of stars in the dataset for that cluster. It is unsure why the King model did not fit 47 Tuc well.

4.4 Cluster Inner/Outer Separation

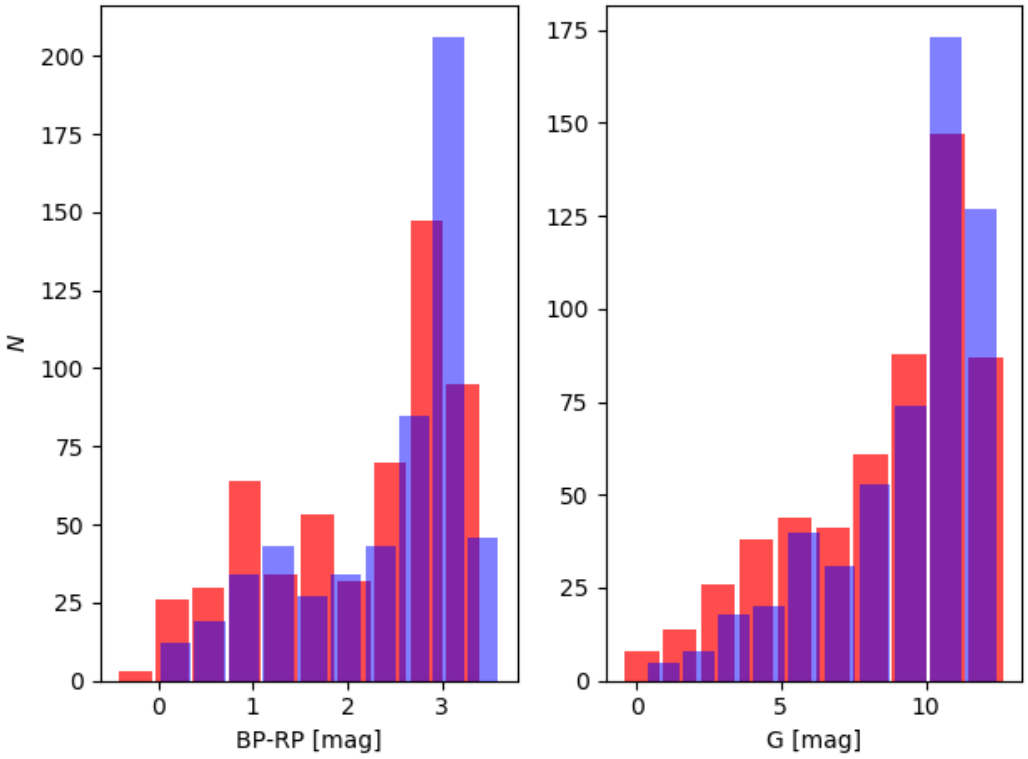
With the half-number radii for each cluster, the stars were split into an inner and outer portion of each cluster. Once again, an HR diagram was created with an emphasis on those portions. The split HR diagrams for the Pleiades and Omega Centauri can be seen below.

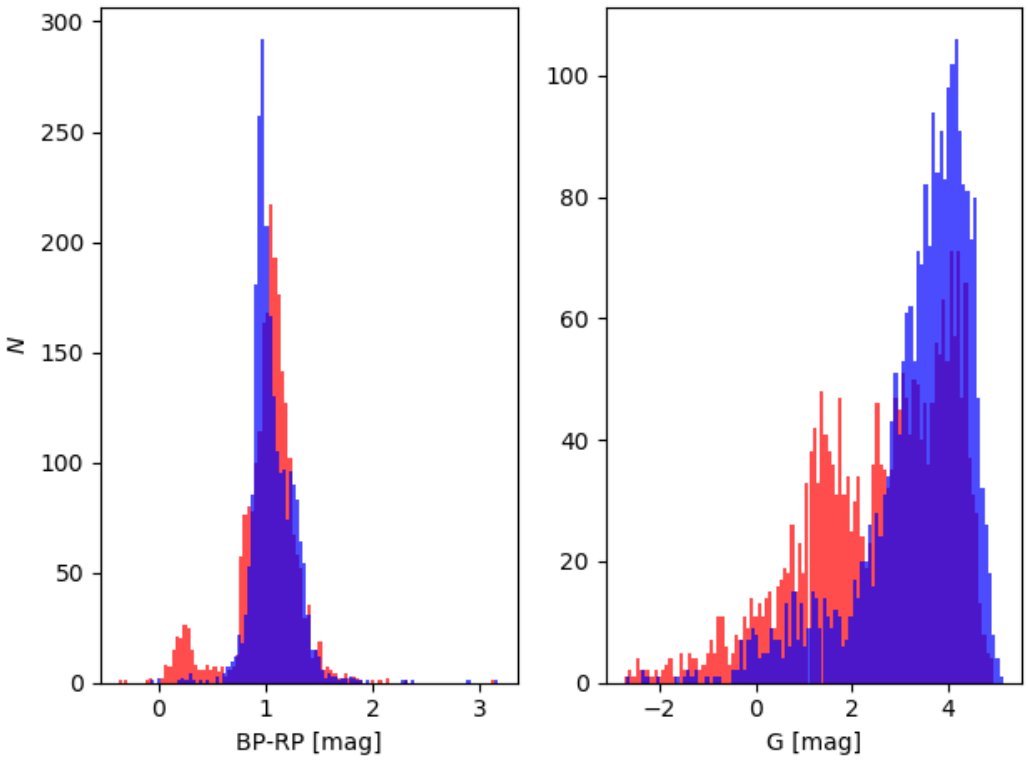




Upon first examination, the Pleiades appears well-mixed; there is not an obvious preference as to what stars occupy what region. Omega Centauri does appear to have a preference. There is an obvious trend that stars

with a higher absolute magnitude (and thus lower luminosity) are in the outer region of the cluster, while stars with a lower absolute magnitude (and thus higher luminosity) are in the inner region. It is evident that almost the entire horizontal branch of giants that are fusing helium into carbon. This only happens in the massive stars. We can investigate this further by plotting histograms, one for the color index and another for the absolute magnitude.





This confirms that the Pleiades is well-mixed, whereas Omega Centauri has obvious trends. Since luminosity roughly correlates with mass, we can deduce that the outer region of Omega Centauri contains the low-mass stars and the inner region contains the high-mass stars. This process of moving the low-mass stars outward and the high-mass star inward is called mass segregation. While this is not direct evidence of mass segregation in Omega Centauri, it might hint at it. The literature does not mention this effect, which might be a new result available to those who use the *Gaia* data. The remaining clusters, including the remaining GCs, have trends resembling the Pleiades: they all appear evenly mixed. Mass segregation is expected to happen over billions and billions of years, so many GCs don't appear segregated yet. OCs will evaporate over time, but the data indicates that none are currently doing that; more research is required on this front.

4.5 Cluster Radial Velocities

Finally, using this dataset, we can calculate the radial velocity of each cluster. Radial velocity is the rate of change of distance between the observer and the object. While this is not hard to measure as compared to the proper motion, we can use this to present new radial velocity determinations or confirm results from the literature. The radial velocity was calculated just as the average of the radial velocities of the stars, with proper error propagation.

Cluster	Radial.Velocity..km.s.	Error..km.s.	Ref.Radial.Velocity..km.s.	Ref.Error..km.s.	Reference
Pleiades	6.866	0.227	5.7	0.60	Conrad et al. (2017)
Hyades	34.659	0.085	39.1	0.02	Ciardi et al. (2018)
Coma	-2.071	0.155	-0.1	0.20	Conrad et al. (2017)
IC 2602	17.570	0.348	5.2	3.00	Conrad et al. (2017)
IC 2391	18.060	0.289	14.1	0.20	Conrad et al. (2017)
M44	27.798	0.149	34.5	0.10	Conrad et al. (2017)
NGC 752	-8.290	0.282	4.7	0.10	Conrad et al. (2017)
47 Tuc	-18.024	0.365	-18.7	0.20	Harris (1996)
Omega Centauri	233.195	0.417	232.1	0.10	Harris (1996)
M3	-142.831	0.961	-141.2	5.60	Smolinski et al. (2011)

M4	70.917	0.192	70.4	0.40	Harris (1996)
M5	54.057	0.529	51.8	0.50	Harris (1996)
M9	111.384	0.915	229.1	7.00	Harris (1996)
NGC 1851	322.956	1.296	320.9	1.00	Harris (1996)

As is evident by the above table, *Gaia* data can accurately predict the radial velocities of OCs and GCs. It should be noted that Messier 3 (M3) has only one star with radial velocity data, but it obviously confers well with the radial velocity from the literature.

5 Conclusion

As stated before, he goal of this project is to confirm and challenge results commonly found in the literature about OCs and GCs in general, and about these clusters in particular. We summarize the results of the project below:

1. GCs, in general, are about as old as the universe. This is confirmed by the HR diagrams and isochrones, showing fitted ages ranging from 10-13 Gyr. In general, they are very far away from the Sun, orbiting outside the disk of the Milky Way. This is confirmed by the computed distances. This large distance also results in small angular sizes on the sky, confirmed by the half-number radii.
2. OCs, in general, are very young. This is confirmed by the HR diagrams and isochrones, showing fitted ages ranging from 13 Myr to 1.3 Gyr. They have a large range of ages, but still much younger than the Galaxy. They orbit within the disk of the Milky Way. This is confirmed by the distances and positions on the sky. As such, they have large angular sizes, confirmed by the half-number radii.
3. NGC 1851 has a large range in distance, but this is likely to be an artifact of the data and not a challenge to the literature-derived value. We defer to the literature value, and do not challenge this result.
4. IC 2602 has a reported age of 50 Myr, but an HR diagram-derived age of 13 Myr. We challenge the results from the literature here. Perhaps the low-mass stars that are the basis for the lithium-depletion burning technique do not reside in the cluster proper or are the remainder of stars from a bygone cluster. We see no reason the isochrone fitting technique would result in an incorrect age determination.
5. Although we have used the King model to explore the surface density of each cluster in our dataset, it appears (based on the R^2 values) that it does not explain the variability in the GCs and does explain the variability in the OCs. This is incorrect according to King (1966), as only GCs and elliptical galaxies should have an isotropic velocity distribution. We challenge the literature in this regard; perhaps more OCs have isotropic velocity distributions than previously thought, or perhaps a new model is needed.
6. No new radial velocity measurements are found. All are reproduced in the literature using other methods.

6 Acknowledgements

The author would like to thank Professor Chris Mihos for the as-needed assistance on this project.

7 Extra Data

A .zip file has been prepared and attached to the submission of this report containing all the data used, as well as the plots produced and code used to analyze the data.

8 References

- Adams, J. D., Stauffer, J. R., Monet, D. G., Skrutskie, M. F., & Beichman, C. A. 2001, AJ, 121, 2053
- Agüeros, M. A., Bowsher, E. C., Bochanski, J. J., et al. 2018, ApJ, 862, 33
- Barrado y Navascués, D., Stauffer, J. R., & Jayawardhana, R. 2004, ApJ, 614, 386
- Basri, G., Marcy, G. W., & Graham, J. R. 1996, ApJ, 458, 600
- Boyles, J., Lorimer, D. R., Turk, P. J., et al. 2011, ApJ, 742, 51
- Caputo, F., Castellani, V., & Quarta, M. L. 1985, A&A, 143, 8
- Ciardi, D. R., Crossfield, I. J. M., Feinstein, A. D., et al. 2018, AJ, 155, 10
- Conrad, C., Scholz, R.-D., Kharchenko, N. V., et al. 2017, A&A, 600, A106
- Converse, J. M., & Stahler, S. W. 2010, MNRAS, 405, 666
- Forbes, D. A., Lasky, P., Graham, A. W., & Spitler, L. 2008, MNRAS, 389, 1924
- Forbes, D. A., & Bridges, T. 2010, MNRAS, 404, 1203
- Gaia Collaboration 2016, A&A, 595, A2
- Gaia Collaboration 2018, A&A, 616, A12
- Galli, P. A. B., Moraux, E., Bouy, H., et al. 2017, A&A, 598, A48
- Giardino, G., Pillitteri, I., Favata, F., & Micela, G. 2008, A&A, 490, 113
- Harris, W. E. 1996, AJ, 112, 1487
- Hertzsprung, E. 1911, POPot, 63
- Kharchenko, N. V., Piskunov, A. E., Röser, S., et al. 2005, A&A, 438, 1163
- King, I. R. 1966, AJ, 71, 64
- Koleva, M., Prugniel, Ph., Ocvirk, P., et al. 2008, MNRAS, 385, 1998
- Kraus, A. L., & Hillenbrand, L. A. 2007, AJ, 134, 2340
- Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
- Luri, X., Brown, A. G. A., Sarro, L. M., et al. 2018, A&A, 616, A9
- Majaess, D. J., Turner, D. G., Lane, D. J., & Krajci, T. 2011, JAVSO, 39, 219
- Momcheva, I., & Tollerud, E. 2015, arXiv:1507.03989
- Odenkirchen, M., Soubiran, C., & Colin, J. 1998, NewA, 3, 583
- Pandey, A. K., Sandhu, T. S., Sagar, R., & Battinelli, P. 2010, MNRAS, 403, 1491
- Paust, N. E. Q., Reid, I. N., Piotto, G., et al. 2010, AJ, 139, 476
- Perryman, M. A. C., Brown, A. G. A., Lebreton, Y., et al. 1998, A&A, 331, 81
- Piskunov, A. E., Schilbach, E., Kharchenko, N. V., Röser, S., Scholz, R.-D. 2008, A&A, 477, 165
- Röser, S., Schilbach, E., Piskunov, A. E., Kharchenko, N. V., & Scholz, R.-D. 2011, A&A, 531, A92
- Russell, H. N. 1914, PA, 22, 275
- Smolinsky, J. P., Lee, Y. S., Beers, T. C., et al. 2011, AJ, 141, 89
- van de Ven, G., van den Bosch, R. C. E., Verolme, E. K., de Zeeuw, P. T. 2006, A&A, 445, 513
- Woodley, K. A., Goldsbury, R., Kalirai, J. S., et al. 2012, AJ, 143, 50