JOURNAL OF
EMERGING INVESTIGATORS

Article

Exploration of the density–size correlation of celestial objects on various scales

Thomas Thomas*[,1], Nipun Nagendra*[,2], Sreekar Garimella*[,1], Shyamal Mitra[3]

[1] Westlake High School, Austin, Texas

[2] Vandegrift High School, Austin, Texas

[3] The University of Texas at Austin, Austin, Texas

* These authors contributed equally to this work

SUMMARY

In the 1930s, Edwin Francis Carpenter calculated a size density restriction and correlation for galaxy clusters. This was reworked by Gerard de Vaucouleurs in the 1970s, proving an improved and expanded size density correlation. However, these papers still use legacy datasets, some of which do not contain data of high precision and accuracy collected by modern observational tools. To address this limitation, we used modern datasets to reinvestigate and redefine the correlation between the density and size of celestial objects in the universe. Specifically, we sought to analyze the density and size data of various celestial objects, such as neutron stars, galaxies, and galaxy clusters, and determine if a linear relationship exists between their densities and sizes. We hypothesized a negative linear correlation exists on all celestial object scales. Through logarithmically graphing data, we found that this negative correlation of luminous matter exists on such scales. We postulate that this negative correlation implies a hierarchical view of the universe, and the repeating pattern on multiple scales suggests a fractal cosmology of the universe.

INTRODUCTION

The exploration of the universe's structure has remained a sought-after area of research in astronomy, offering explanations for the underlying structures and mechanisms governing the universe. A consistent model of the universe's cosmology could serve to explain many problems that are still outstanding in astrophysics today, such as Olbers' Paradox, which asks why the night sky is not bright if it is filled with a practically infinite number of stars (1). Regardless, the two most common frameworks in modern cosmology are homogeneity and hierarchy. Homogeneity describes an overall uniformity within the universe, while hierarchy depicts a degree of tiered organization of the components of the universe. Recently, a new model has started to gain traction in the hierarchical community. This is the model of fractal cosmology, or the idea that the universe has self-similar structures on differing scales once you zoom out enough (2). One way to prove this theory is by examining the density-size correlation of varying objects.

Astronomers and cosmologists have previously analyzed the concept of a density-size correlation, attempting to decipher the relationship between an object's mass, volume, and other cosmic factors. Density, a standard variable in astrophysics, refers to the amount of mass contained in the given volume of a celestial object. Meanwhile, size is a fundamental characteristic that influences various aspects of celestial bodies, including their gravitational interaction, luminosity, and evolutionary history. Examining a correlation between the density and size of celestial bodies would allow the discernment of patterns across the universe. Particularly, patterns between these two variables across the universe suggest a cosmological principle governing the organization of mass in the universe.

Edwin Carpenter's paper on the characteristics of galaxies deeply explored the relationship between the density and size of galaxy clusters (3). This paper remains a benchmark in studying these celestial bodies and has greatly influenced ongoing research. Carpenter began his study by outlining his method for calculating the density of galaxy clusters, which he encompassed by defining the number of nebulae in thousands per cubic megaparsec (3). His data collection was based on information obtained from the Mount Wilson and Harvard databases, testifying to the credibility of his inputs (3). Carpenter plotted density against size derived from his observations of galaxy clusters, applied a linear regression analysis to his results, and revealed the significant trend that galaxy clusters with smaller radii were denser than their larger counterparts (3). Rather than view these galaxies as isolated entities, Carpenter concluded that they all form an interconnected network, further positing that these clusters represent extremes within a nonuniform space distribution limited by a certain density restriction (3).

After several years, de Vaucouleurs reworked Carpenter's study with bigger and better data sets (4). Building on the conclusions of papers written only a few years before, de Vaucouleurs constructed a table of eight "groups and clusters of galaxies," including their catalog numbers, type, mass, velocity, distance, and mass-luminosity ratios, among other measures (4). Plotting density against the radius, de Vaucouleurs found a close correlation similar to Carpenter's (4). De Vaucouleurs used these analyses to raise questions about the stability of galaxy clusters.

Our study of the density-size correlation builds upon the pioneering work of Carpenter and de Vaucouleurs. Their findings laid the groundwork for this astrophysical phenomenon and set the stage for subsequent researchers to dig deeper. However, given that Carpenter's and de Vaucouleurs' research dates back to the 1930s and 1960s, respectively, it lacks the breadth and depth of modern datasets, statistical methods, and astrophysical theories. This limitation arises from the significant advancements in observational astronomy and computational power over the past few decades. Since these studies were conducted, we have developed new tools and telescopes such as the James Webb Space Telescope that allow us to observe celestial objects better than ever before. Moreover, the creation of large databases such as the Sloan Digital Sky Survey or the NASA catalogs has enabled a more comprehensive analysis across a larger dataset. Specifically, Carpenter used 42 metagalactic clusters and de Vaucouleurs used a collection of 30 celestial objects, while our study looked at 6192 celestial objects (3, 4). Carpenter and de Vaucouleurs were prevented from exploring this correlation in different types of celestial bodies beyond galaxy clusters because they were limited in terms of the size of objects measurable by their instruments. The need for a renewed examination of the density-size correlation becomes clear given these limitations. Continued research in this field will utilize newer, more accurate data to expand on Carpenter's and de Vaucouleurs' conclusions and offer a more comprehensive understanding of celestial objects across various scales, from neutron stars to clusters of galaxies.

We explored the complex relationship between the densities and sizes of celestial bodies, including neutron stars, main sequence stars, star clusters, galaxies, galaxy clusters, and superclusters. By examining this correlation, we determined whether a coherent and systematic connection exists between these variables, providing insights into the fundamental structure and evolution of the universe. Moreover, our investigation could theoretically be used in future research to help predict higher-order classifications and mathematically determine the boundary of the observable universe. However, these postulations are beyond the scope of this research. We hypothesized that the density decreases as the scale of celestial objects increases. We used recently created datasets to plot a comprehensive graph highlighting the universe's relationship between density and size.

Our data showed a clear negative correlation between density and size. In other words, the larger the scale of the celestial body, the lower the overall density. Within the scope of the present research, this pattern holds true on all scales we investigated, from neutron stars through planets to superclusters of galaxies. This could be an explanation for Olbers' paradox, because if we hold our relation to be true across all ranges of radii, the density of the universe is near zero at the universal scale, meaning that not a lot of light reaches us (1). This also suggests a fractal cosmology of the universe, as this correlation is followed across all scales (2). Carpenter's and de Vaucouleurs' findings about interconnected galaxy clusters and the idea of a "nonuniform space distribution" deserve more in-depth research. Examining this aspect of the density-size correlation could unveil further mysteries of our universe, such as the existence and formation of massive entities like superclusters. Therefore, revisiting Carpenter's and de Vaucouleurs' research with fresh perspectives and modern data offers transformative insights into the cosmos.

Type	Data Entries	Mean Mass (kg)	Mean Radius (m)
Neutron Stars	6	7.89×10^{29}	1.08×10^{4}
Planets	2,417	2.55×10^{27}	6.00×10^{7}
Main Sequence Stars	648	4.97×10^{30}	3.91×10^{9}
Star Clusters	1,000	2.71×10^{32}	1.94×10^{17}
Galaxies	800	1.91×10^{40}	9.25×10^{19}
Galaxy Clusters	8	3.28×10^{45}	2.03×10^{24}
Galaxy Superclusters	1,313	2.37×10^{42}	4.75×10^{23}

Table 1: Number of entries, mean mass, and mean radius for each type of celestial object in our dataset. Neutron stars and galaxy superclusters have a low number of entries because there has been relatively little data collected on them.

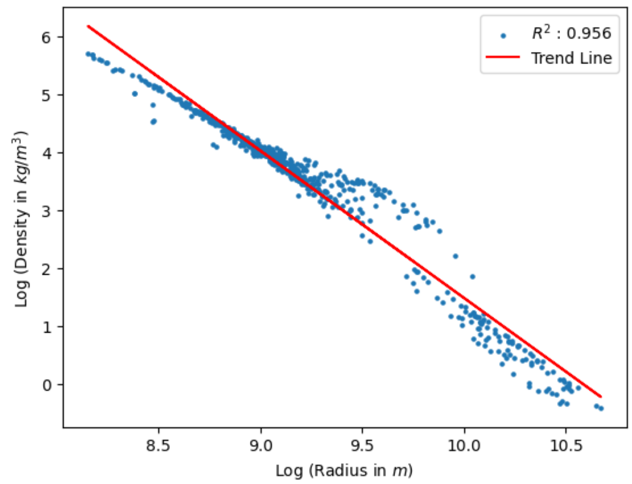

Figure 1: Negative correlation between the density and radius of main sequence stars (n=648). A logarithmic scale was applied to both axes, and a linear trend line was fitted to the graph.

RESULTS

We utilized extensive astronomical databases that compile observations from various sources encompassing ground-based telescopes and space-based missions, and a Python library called Pandas to create a dataset and organize the data (5). In total, our study examined 6 neutron stars, 2,417 planets, 648 main sequence stars, 1,000 star clusters, 800 galaxies, 8 galaxy clusters, and 1,313 galaxy superclusters (4, 6-11). We normalized the data to a standard scale and format to ensure consistency and comparability. This allowed for the effective comparison of density and size measurements across different types of celestial objects (**Table 1**). The data incorporated parameters like size, mass, volume, density, and location in the cosmos for each species of stellar object. In particular, we expressed the data in logarithmic units. Doing so allowed for alignment with the standard astronomical practice, facilitating the comparison of objects with measurements differing by multiple orders of magnitude.

Before the linear regression analysis, a preliminary analysis was conducted using only main sequence stars to detect a trend in the dataset. A base ten logarithm was added to both variables to show a clear correlation between density and radius. After fitting a trendline, a clear negative slope appeared (**Figure 1**). The observed negative correlation

JOURNAL OF
EMERGING INVESTIGATORS

https://doi.org/10.59720/23-240

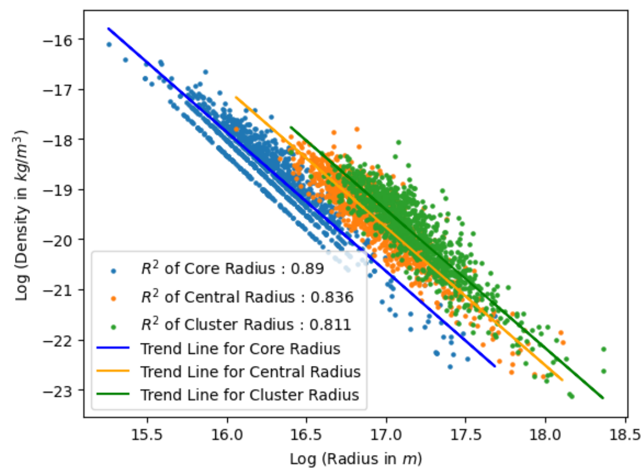

Figure 2: Negative correlations between density and radius for star clusters (n=1,000). Three measurements of radii were provided due to the obscure nature of the edge of star clusters: core radii (blue), central radii (yellow), and cluster radii (green). A logarithmic scale was applied to both axes, and trend lines were fitted to the graph's radii measurements.

between density and size aligned with Carpenter's finding. We then performed a linear regression on the main sequence stars. Linear regression, log(Density) = -2.54 * log(radius), demonstrated a statistically significant correlation (F(1, 646) = 14,000; $p < 0.213$).

Although star clusters are a common association of stars, the nature of the body leads to a loose classification of their size. Thus, the dataset of star clusters provided multiple values for the radius of a cluster. However, all three categories of radii showed evidence of Carpenter's negative density-size relation (**Figure 2**). We performed the regression analysis using the cluster radius for star clusters. Linear regression, log(Density) = -2.77 * log(radius), demonstrated a statistically significant correlation (F(1, 998) = 4,500; $p < 0.283$).

A similar trend was found in the galaxy dataset (**Figure 3**). We then analyzed this galaxy dataset trend. Linear regression, log(Density) = -0.48 * log(radius), showed a statistically significant correlation (F(1, 798) = 120; $p < 0.393$).

The data for superclusters are not as indicative of a trend as we would like them to be. This can likely be attributed to the looser definition for superclusters compared to other cosmic structures. Superclusters are vast regions containing multiple galaxy clusters, and their boundaries are often not well-defined. This leads to variability in density and radius measurements, making it harder to identify a clear trend. However, a negative trend can still be found (**Figure 4**).

Nevertheless, we performed the regression analysis on the galactic superclusters. Linear regression, log(Density) = -1.4 * log(radius), demonstrated a statistically significant correlation (F(1, 1,311) = 400, p < 0.0872). Repeating this analysis and combining all other datasets, the observable universe follows this trend (**Figure 5**). Regression analysis on this combined, unified dataset showed a statistically significant correlation (F(1, 6,190) = 420,000; $p < 0.235$):

$$\log(\text{density}) = -2.1 \log(\text{radius}) + 19.34 \quad \text{(Eqn 1)}$$

Fractal cosmological theory was considered during our study, although it would only hold true on the scale of star clusters and larger objects in our data (2). The data were replotted with only star clusters, galaxies, galaxy clusters, and galaxy superclusters to emphasize this trend (**Figure 6**). The linear regression analysis, log(Density) = -1.54 *log(radius), of the dataset of star clusters through galactic superclusters revealed a highly significant correlation (F(1, 3,119) = 42,000, $p < 0.167$).

DISCUSSION

Our research focused on exploring the relationship between cosmic density and size, extending the work of Carpenter and de Vaucouleurs (3, 4). We sought to understand this relationship across a broader range of celestial structures using modern, high-precision data. Our findings show a consistent negative correlation between density and size, with the regression line described by:

$$\log(\text{density}) = -2.1 \log(\text{radius}) + 19.34 \quad \text{(Eqn 1)}$$

Both Carpenter and de Vaucouleurs found negative relationships between density and size (3,4). Unfortunately, Carpenter does not provide a specific equation for the density-size relationship, though his work does include relations between volume and density (3). This means that while we can refer to his findings regarding how density changes with volume, we lack a direct formula to compare the density-size relationship quantitatively with our results. This limitation affects our ability to make precise comparisons but highlights the need for further research to refine these relationships and develop more explicit models for cosmic density and size correlations (12). Our finding is more consistent with de Vaucouleurs' second reworking, finding that such a relation does exist and tying the physics of star clusters to galactic superclusters together through fractal cosmology (4):

$$\log(\text{density}) = -2.1 \log(\text{radius}) + 19.34 \quad \text{(Eqn 1)}$$

as compared to de Vaucouleurs' line defined for stellar bodies:

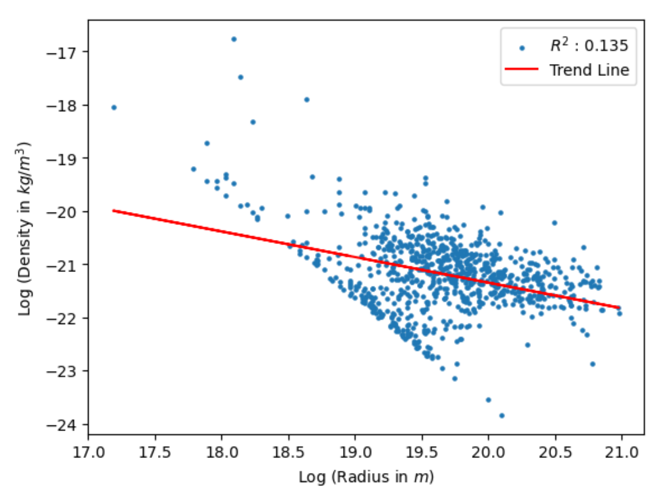

Figure 3: Loose negative correlation between density and radius for galaxies (n=800). A logarithmic scale was applied to both axes. Visual confirmation of a trend could not be obtained, but a linear trend line shows a slight negative correlation when fitted to the graph.

https://doi.org/10.59720/23-240

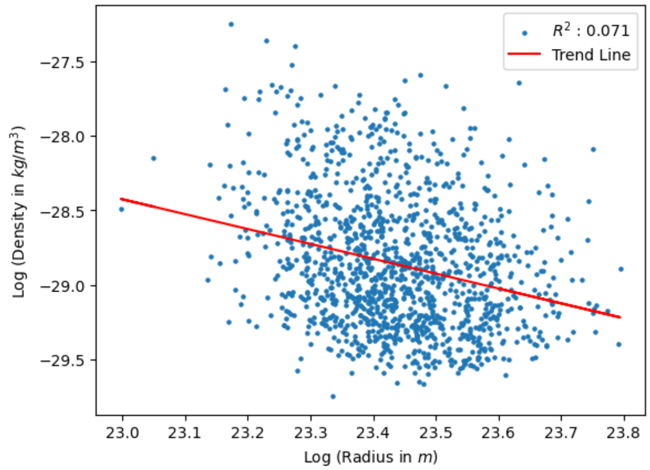

Figure 4: Loose negative correlation between density and radius for superclusters (n=1,313). A logarithmic scale was applied to both axes. Visual confirmation of a trend could not be obtained, but a linear trend line shows a slight negative correlation when fitted to the graph.

log(density) = -2.7 log(radius) + 29.7 (Eqn 2)

The differences in the slope and intercept between our findings and de Vaucouleurs' results can be primarily attributed to the scale and diversity of the data sets used. Our study incorporates modern, high-precision data covering a broader range of galactic structures, from star clusters to superclusters, whereas de Vaucouleurs' work was limited to the data available in his time. The inclusion of more massive and diverse galactic structures in our analysis leads to a different density-radius relationship, reflected in a less steep slope (-2.1 compared to de Vaucouleurs' -2.7) and a lower intercept (19.34 compared to de Vaucouleurs' 29.7).

Based on the data, there are two populations of main sequence stars in separate parts of the graph (**Figure 1**). The lower right category is likely supermassive stars, which are known to be a lot bigger and slightly more massive than the average main sequence star. This observation is complemented by the population density of each group, which shows that supermassive stars are much smaller in number than moderately sized stars. Meanwhile, the upper left cluster likely represents the more numerous, moderately sized main-sequence stars. These stars are smaller in size and mass compared to the supermassive stars found in the lower right category.

The original data collection may have excluded a range of values due to observational constraints, which could be why the data appears truncated to the lower left (**Figure 2, Figure 3**). Ultimately, these data sets still retain accuracy.

As previously mentioned, the idea that "as the size of a celestial object increases its density decreases" serves as a solution to Olbers' paradox, showing that the night sky must have a density so low that it cannot be illuminated and bright. For this same reason, this negative density-size correlation supports "empty universe" hypotheses on larger scales. On these larger scales, dark matter could also factor into the negative correlation, providing data on the effects of the theoretical particle. Another theory this correlation could support, which discounts any gravitational effects on the data, is the fractal cosmological theory, which as mentioned before only applies to star clusters and above. The new regression line that we plotted to emphasize fractal cosmological theory (**Figure 6**) had the equation:

log(density) = -1.54 log(radius) + 7.407 (Eqn 3)

This was very close to other work done on this range of objects, such as de Vaucouleurs' 1970 paper, where the slope for star and galaxy systems fell within the range -1.5 to -1.9 in his calculations (12).

The concept of fractals proved central to our study (12). Fractals, complex structures exhibiting self-similarity at varying scales, have been suggested to represent the structure of the universe, namely an overarching universal pattern repeating at every scale, from mere star clusters to colossal superclusters (2).

A limitation of this study is the lack of white dwarfs within the overall dataset and analysis. Since a mass-luminosity relationship does not hold for white dwarfs since they cool over time, we could not find any white dwarf catalogs that included both mass and radius. However, it can be predicted that white dwarfs would still show a negative correlation between density and size since all white dwarfs were once main sequence stars and maintain many of the same properties, including mass.

Another limitation of our study is the predefined nature of certain structures, such as star clusters and galactic clusters, which are specifically defined based on their density. This inherent classification can introduce a bias in our analysis, as these structures are selected and categorized due to their density characteristics. Consequently, our findings might reflect these classification criteria rather than a more general cosmic density-size relationship at larger scale structures where limits and ranges are defined empirically. Future research should aim to include a broader range of cosmic structures with varied definitions and characteristics to ensure a more comprehensive and unbiased analysis of the density-size correlation.

Future work could be done using and expanding upon these data and correlations, including the addition of a white

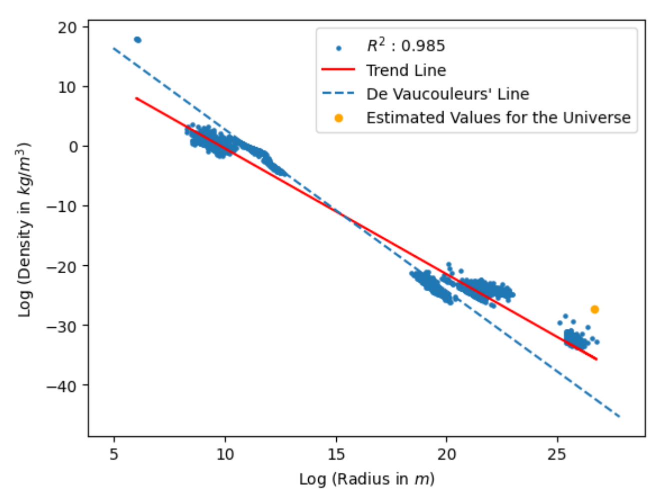

Figure 5: Negative correlation between density and radius for all categories (n=6,192). A logarithmic scale was applied to both axes, and a linear trend line was fitted to the graph.

https://doi.org/10.59720/23-240

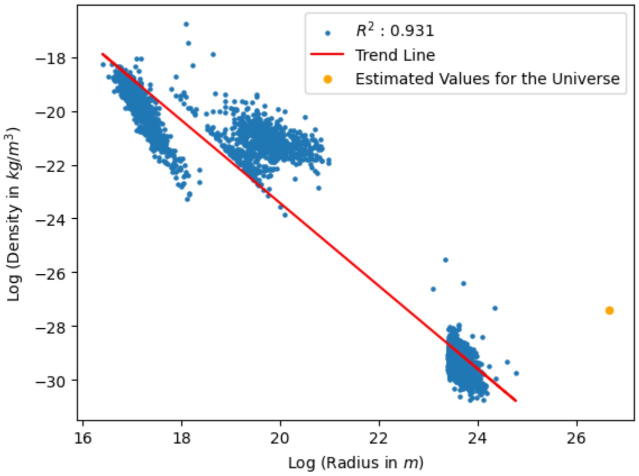

Figure 6: Negative correlation between density and radius for star clusters, galaxies, and galactic superclusters (n=3,113). A logarithmic scale was applied to both axes, and a linear trend line was fitted to the graph.

dwarf catalog to confirm the correlation exists for white dwarf stars. Alternatively, one could investigate the impact of dark matter on the correlation, which would provide insight as to the nature of dark matter in and around larger celestial objects such as galaxies, galaxy clusters, and superclusters. Extending the density-size correlation and relating other variables, one could analyze the density present in the estimated universe and compare that figure with critical densities extrapolated from different expected shapes of our universe. This could help predict the shape of the universe. If one could find distance data along with mass and radius, the 'age' of each data point could be calculated with the Hubble constant and a density-size-time plot could be made. This could demonstrate and assist in measuring the Hubble constant if a fluctuation occurs.

In conclusion, our study verifies the hypothesis that as the size of a celestial object increases, its density decreases. Our contributions to the field include a broader and more precise analysis of the density-size relationship across various galactic structures, supporting fractal cosmological theory and providing a potential solution to Olbers' paradox. Future research could expand upon these findings to explore the impact of dark matter and further refine our understanding of the universe's structure and evolution.

MATERIALS AND METHODS

Our study preliminarily identified celestial bodies of various scales in the universe (listed below). The choice of these celestial bodies was an extension of the bodies chosen by de Vaucouleurs' research. Then, through our research, we found databases that contained specific required parameters such as mass and size (4, 6-11). Our next step was a rigorous data cleaning procedure through custom code to ensure data quality, including the removal of data points with missing values. After this step, we performed data conversions to ensure consistency in the units used across all the datasets. Lastly, after the consolidation of each dataset into a comprehensive one, calculations were performed to derive certain values, such as density.

A logarithmic scale proved instrumental when graphically representing the correlation between density and size. This scale's unique attribute of denoting each increment as a multiplication rather than a linear addition allowed for visual confirmation of a correlation between the density and radius of a range of astronomical objects. Also, the logarithmic scale ensured that all data entries were neatly visible on a single graph. The data were plotted, and a regression line was fitted using Matplotlib, another Python module (13). Code for the data analysis and regression can be found here: https://github.com/NipunNagendra/DensitySizeDataAnalysis.

Neutron Stars

Due to the difficulty of recognizing neutron stars in space, catalogs of the objects lack data about their physical properties. However, a range of radii and masses were derived in a 2016 paper (6). The physical properties of six neutron stars were estimated by finding the median of these ranges. Of the six data entries for neutron stars, the mean radius was determined to be 1.08×10^4 m, and the mean mass was determined to be 7.89×10^{29} kg.

Planets

We used NASA's Exoplanet Archive, a constantly growing catalog of exoplanets, accessed 30 June 2023 (7). Out of the catalog, only the exoplanets with a non-null radius and mass were selected from the composite data. Of the 2,417 data entries for planets, the mean radius was determined to be 6.00×10^7 m, and the mean mass was determined to be 2.55×10^{27} kg.

Main Sequence Stars

Data on the mass and radius of main sequence stars were found by focusing on binary stars compiled in the "DEBCat" catalog (8). Of the 648 data entries for main sequence stars, the mean radius was determined to be 3.91×10^9 m, and the mean mass was determined to be 4.97×10^{30} kg.

Star Clusters

Since there are many star clusters observable in the Milky Way, we were able to find a thousand data entries for star clusters in NASA's Milky Way Star Clusters Catalog (9). Of the 1000 data entries for star clusters, the mean radius was determined to be 1.94×10^{17} m, and the mean mass was determined to be 2.71×10^{32} kg.

Galaxies

The galaxy dataset contained the Holmberg mass and diameter of 800 galaxies (10). Of these data entries, the mean radius was determined to be 9.25×10^{19} m, and the mean mass was determined to be 1.91×10^{40} kg.

Galaxy Clusters

We found our data on galaxy clusters in de Vaucouleurs' paper (4). Although there are more recent data for galaxy clusters, the intended purpose of including it within our research was to link our work to de Vaucouleurs' study, as it is an extension of his research. Furthermore, data on galaxy superclusters has been collected below from the Sloan Digital Sky Survey (SDSS). Since the classifications for both celestial structures are similar, as reflected by the similar masses and radii, we pair modern datasets alongside the parent study's data. De Vaucouleurs' data entries (in CGS) were converted into the SI system used in our analysis. Of the eight data entries for galaxy clusters, the mean radius was determined to be 2.03×10^{24} m, and the mean mass was determined to be 3.28×10^{45} kg.

Galaxy Superclusters

We found our supercluster data in a database stemming from the Sloan Digital Sky Survey Data Release 7 (11). The data on superclusters were hard to find because measuring the volume and mass of superclusters is a relatively new field with few papers currently published. Of the 1,313 data entries for galaxy superclusters, the mean radius was determined to be 4.75×10^{23} m, and the mean mass was determined to be 2.37×10^{42} kg.

Regression

Each plot with a regression line was created using the same regression algorithm, a standard linear regression applied to the graph after the conversion to a logarithmic scale. This allowed for easy comparison between graphs and, ultimately, easy extrapolation of conclusions from the data.

ACKNOWLEDGMENTS

The authors would like to thank the High School Research Academy program at The University of Texas at Austin and its staff without whom this research would not be possible.

Received: September 24, 2023
Accepted: February 23, 2024
Published: March 24, 2025

REFERENCES

1. "Olbers' Paradox." *Olbers' Paradox | Astronomy 801: Planets, Stars, Galaxies, and the Universe*, *Penn State University*. https://e-education.psu.edu/astro801/content/l10_p2.html. Accessed 15 Aug. 2023.
2. Pietronero, L. et al. "On the Fractal Structure of the Visible Universe." *arXiv.* 25 November 1996. https://doi.org/10.48550/arXiv.astro-ph/9611197.
3. Carpenter, E. F. "Some Characteristics of Associated Galaxies. I. a. Density Restriction in the Metagalaxy." *The Astrophysical Journal*, vol. 88, October 1938, pp. 344-355. https://doi.org/10.1086/143987.
4. De Vaucouleurs, G. "The Apparent Density of Matter in Groups and Clusters of Galaxies." *Astrophysical Journal*, vol. 131, May 1960, pp. 585-597. https://doi.org/10.1086/146871.
5. McKinney, W. "Data structures for statistical computing in python". *Proceedings of the 9th Python in Science Conference,* vol. 445, 2010, pp. 56-61. https://doi.org/10.25080/Majora-92bf1922-00a.
6. Ozel, F. et al. "The Dense Matter Equation of State from Neutron Star Radius and Mass Measurements." *The American Astronomical Society,* vol. 820, no. 1, March 2016, p. 28. https://doi.org/10.3847/0004-637X/820/1/28.
7. NASA Exoplanet Archive. "Planetary Systems." *NExScI-Caltech/IPAC*, 29 June 2023, https://doi.org/10.26133/NEA12. Accessed 30 June 2023.
8. Southworth, J. "The DEBCat detached eclipsing binary catalog." *arXiv*. 5 November 2014. https://doi.org/10.48550/arXiv.1411.1219.
9. Karchenko, N. V. "Global survey of star clusters in the Milky Way." *Astronomy & Astrophysics*, vol. 543, July 2012, p. A53. https://doi.org/10.1051/0004-6361/201118708.
10. Kaisina, E. I. et al. "Catalog & atlas of the LV Galaxies." *Proceedings of the International Astronomical Union*, vol. 344, 30 October 2019, pp. 377-380. https://doi.org/10.1017/S1743921318005987.
11. Liivamagi, L. J. et al. "SDSS DR7 superclusters - The catalogs." *Astronomy & Astrophysics*, vol. 539, 2012, p. A80. https://doi.org/10.1051/0004-6361/201016288.
12. De Vaucouleurs, G. "The Case for a Hierarchical Cosmology." *Science*, vol. 167, no. 3922, 27 February 1970, pp. 1203-1213. https://doi.org/10.1126/science.167.3922.1203.
13. Hunter, J. "Matplotlib: A 2D Graphics Environment", *Computing in Science & Engineering*, vol. 9, no. 3, 2007, pp. 90-95. https://doi.org/10.1109/MCSE.2007.55.