Astron 98 Final Project: Analyzing the Distribution of Light Within the Galaxy

1. Introduction

This project has explored the distribution of light across 3,000,000 stars within the Milky Way Galaxy according to their position. There was usage of data analysis and generation, the application of filters, and deriving error and fitting the data with it.

2. Chosen Phenomenon and Data Source

For this project, the mean flux of certain stars was analyzed. The location of these stars were collected with data on their right ascension and declination. A total of ___ stars were randomly chosen to be studied. All data was taken from the Gaia Archive. The specific dataset used can be found within the Auxiliary data provided on Gaia Archive website linked here, We used the XP-MERGE/XP-SAMPLING Tables that took spectral data from the Gaia Data Release 3 and focused on information regarding absolute flux and the wavelength system.

3. Equation to Fit Data

The Equatorial Coordinate System:

This system was used to organize our data into a graph. It gives us the positions of stars relative to the celestial poles in a way that consistently describes the positions of objects in the sky regardless of the observer's location or the time of observation; equatorial coordinates remain relatively fixed against background stars. It uses 4 separate components: right ascension, declination, geocentricity, and epochs.

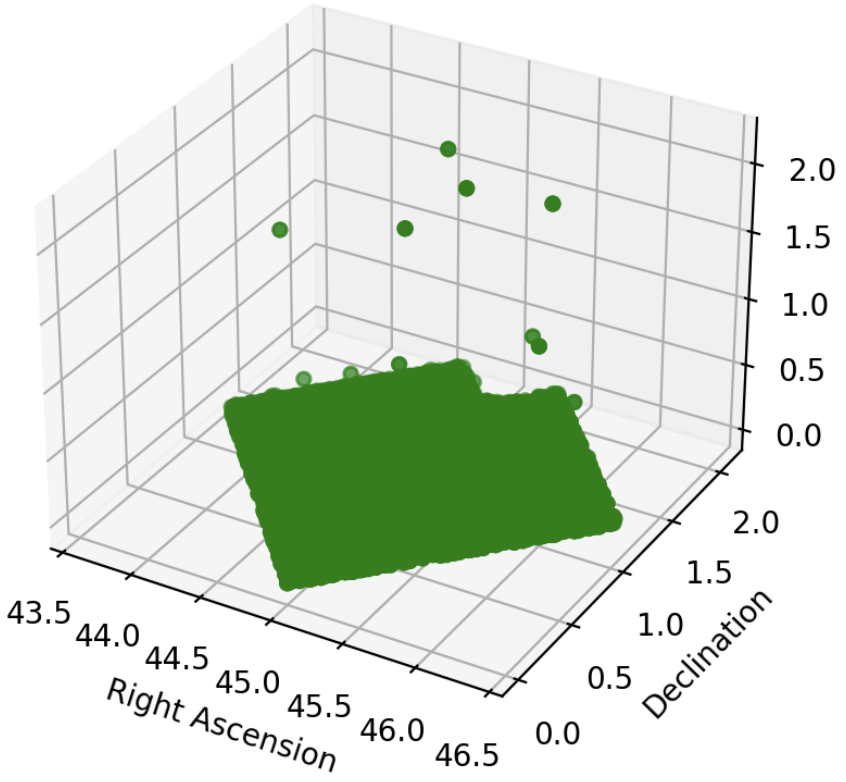
- Right Ascension (RA)
    - Measures the eastward angular distance along the celestial equator
    - Expressed in hours, minutes, and seconds
    - The primal direction is towards the Vernal Equinox
- Declination (Dec)
    - Measures the angular distance North or South of the celestial equator
    - Measured in degrees, minutes, and seconds
    - Positive values indicate positions North of the celestial equator, whilst negative values indicate Southern positions
- Geocentricity
    - Seen from the center of the Earth (as if the Earth were transparent)
    - For observers on Earth's surface, the coordinates slightly change due to the observer's location

- Epochs
    - Due to Earth's axis slowly rotating over time, the orientation of the equatorial coordinate system gets disrupted, so astronomers specify "epochs" (particular dates) when giving a position

Equation Used:

- y = ax^2 + bx + c
    - d[DEC] = a * r[RA]^2 + b * r[RA] + c

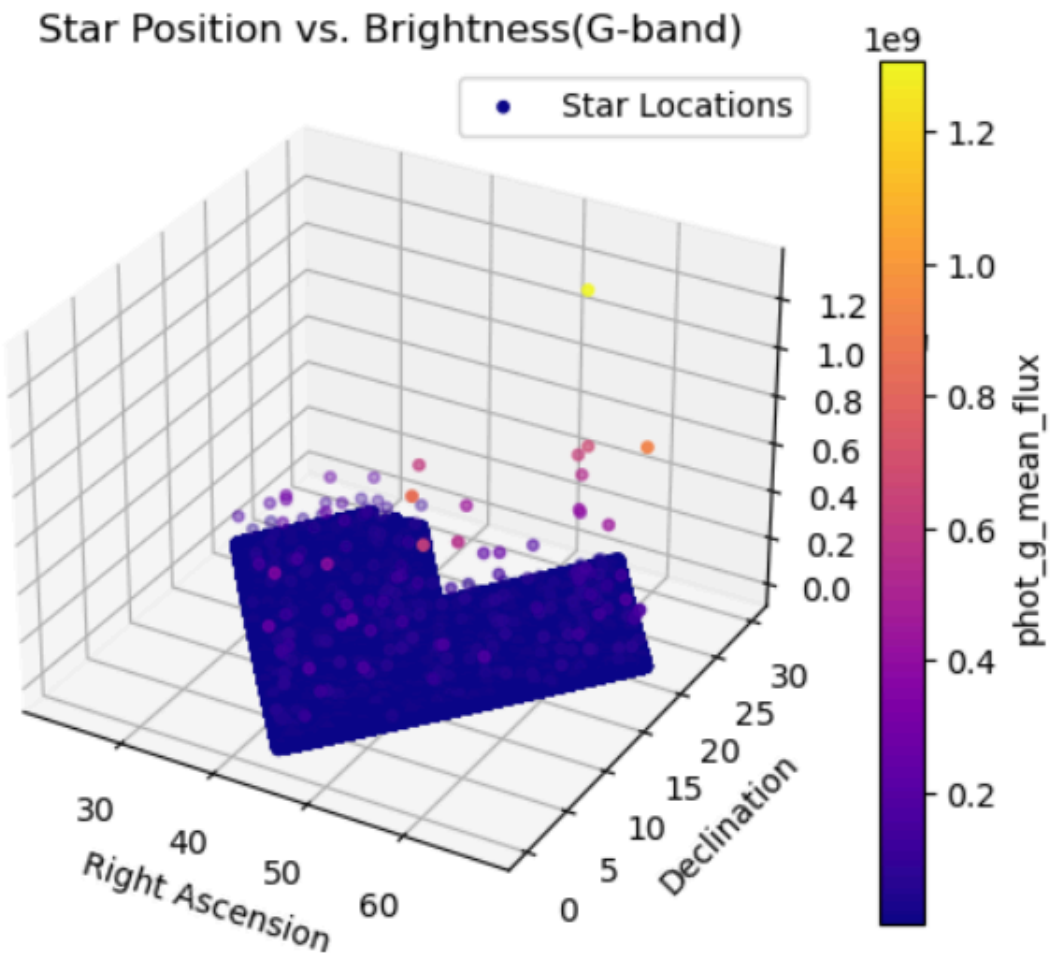
4. Data Generation for Testing



3,000,000 objects measured within the Gaia Data Release 3 with the highest measured luminous flux were chosen within the dataset. The spatial measurements (right ascension and declination with respect to geocentricity and Epochs) are plotted here without any indication of their mean flux.

5. Data Filtering

[3D blue graph with gradient and colorbar]

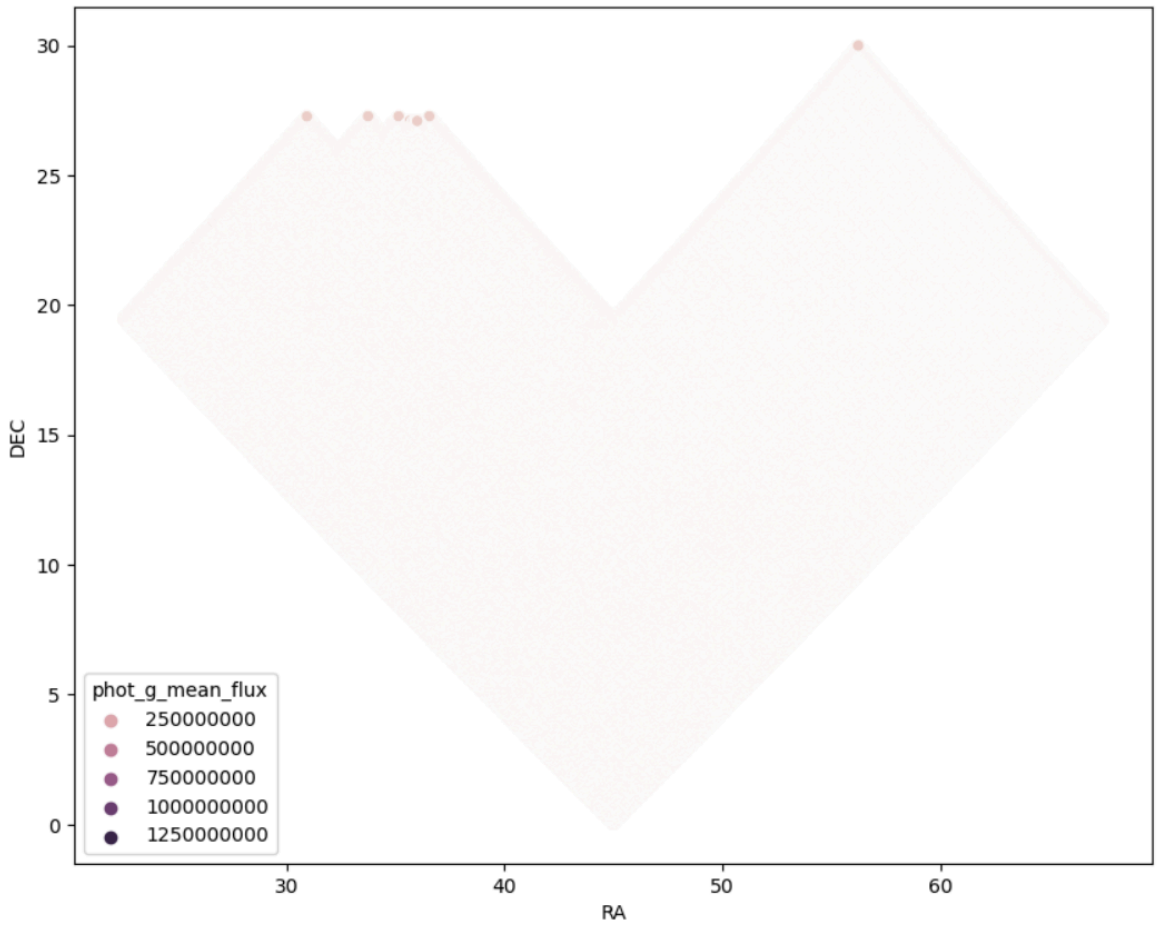


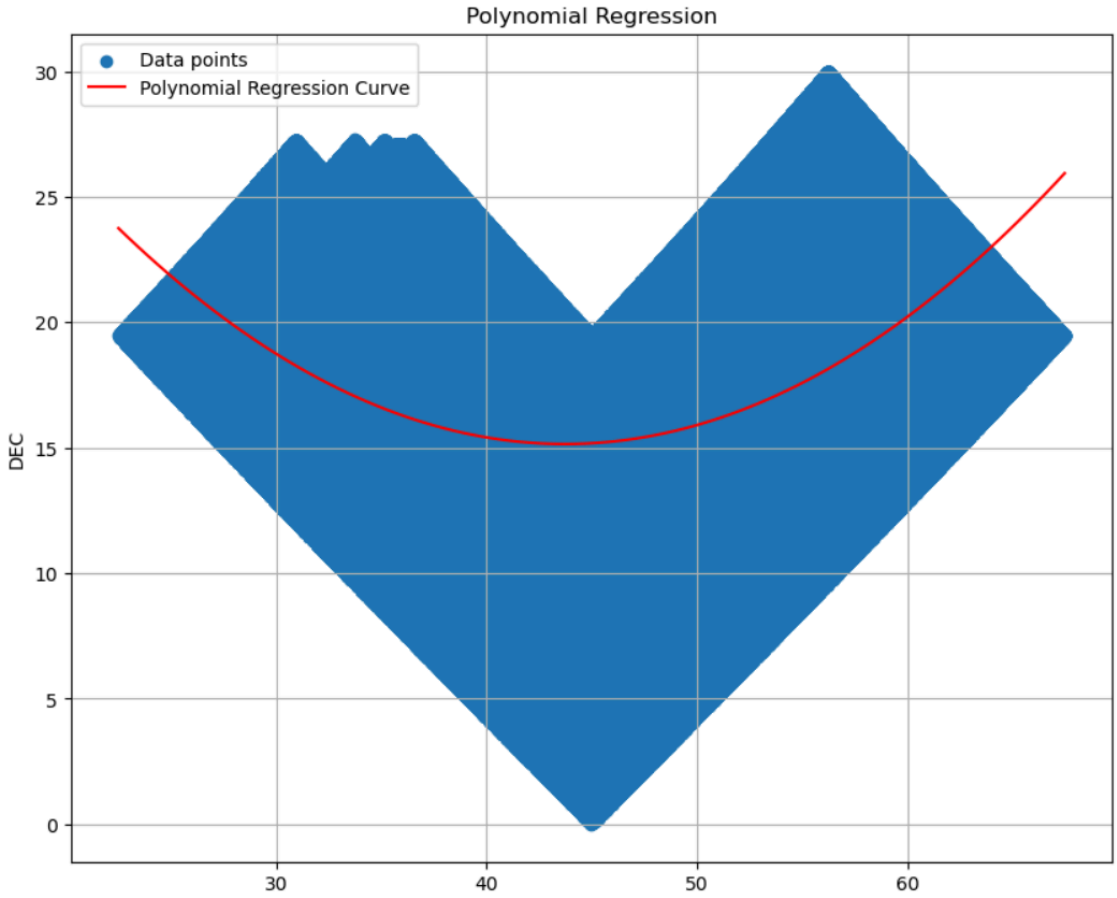
In order to condense our data and ensure its relevance, we will filter it according to the following methods:

A. Data selection within Gaia Data Release 3: choose to only include data relevant to the equatorial coordinate system.
B. Data selection within the datasets: The 3,000,000 astronomical bodies with the largest flux were chosen because they are the most significant in luminosity.

6. Data Fitting

In choosing different methods to analyze our findings, we decided to try plotting the data on 2D graphs to visualize the data differently.


30
25
20
15
10
5
0
DEC
30
40
50
60
RA
phot_g_mean_flux
250000000
500000000
750000000
1000000000
1250000000



We then created a summary of statistics from the dataset that describe the mean flux of the measured coordinates derived from the data collected within the Gaia Data Release 3

```
Summary Statistics for the mean flux of measured coordinates
The mean value is : 32334.876202135187
The standard deviation is : 1969839.0916212068
The minimum value is : 19.68103712262059
The maximum value is : 835597662.8828965
```

7. Explanation of Model Fit

This model represents the individual placements of 3 million stars recorded in Gaia data with the regression curve being the expected placement of a star based on its inherent brightness. The error in relation to this graph are all of the data points surrounding the regression line and represent the fact that there are other factors that go into calculating a star's brightness, like their specific size and type.

8: Conclusion

Organizing this data into these graphs using the Equatorial Coordinate System allows us to see where the highest rates of luminosity can be found within the Milky Way Galaxy. Since we

only took in the top 3,000,000 points, we are able to see where the most luminous bodies are concentrated. In order to more comprehensively analyze the distribution of these luminous bodies, it may be necessary for future studies to intake additional information about what exactly these bodies are and their sizes.

Throughout this project, the difficulty of simply finding usable and relevant data files as well as the difficulty of fitting graphs in a way that effectively displays intended results were emphasized. We realized the necessity of contextual research on the data within databases like the Gaia archive and adaptability in projects such as this one.