KRITTIKA - THE ASTRONOMY CLUB OF IIT BOMBAY

A guide to Star Hopping

Akshay E, Liza Dahiya, Sahyadri Krishna,
Vivek Reddy Pininti, Yasha Kaushal,
Harshda Saxena, Chaithanya Naik, Pratham Patil,
and Vedant Shenoy

KRITTIKA - THE ASTRONOMY CLUB OF IIT BOMBAY

A guide to Star Hopping

Akshay E[1], Liza Dahiya[2], Sahyadri Krishna[3], Vivek Reddy Pininti[4], Yasha Kaushal[5], Harshda Saxena[2,6], Chaithanya Naik[2,6], Pratham Patil[2,6], and Vedant Shenoy[2,6]

[1]Central University of Karnataka, Kadaganchi 585367, Kalaburagi, India

[2]Indian Institute of Technology Bombay, Powai, Mumbai 400076, India

[3]Indian Institute of Science Education and Research, Tirupati 517507, India

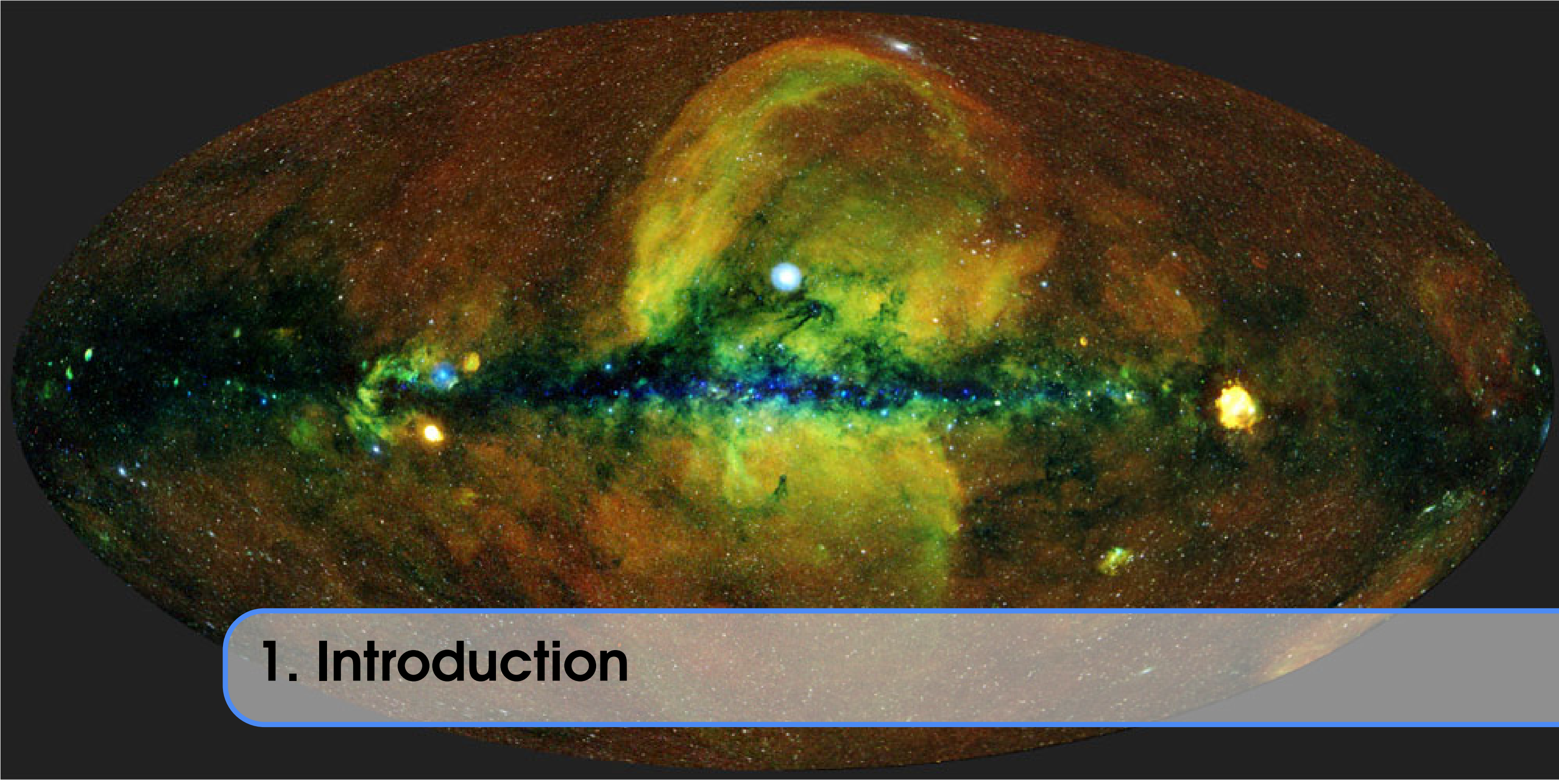
[4]Osmania University, Hyderabad 500007, India

[5]University of Pittsburgh, Pittsburgh, PA 15260, USA

[6]Krittika-The Astronomy Club of IIT Bombay, Powai, Mumbai 400076, India

PUBLISHED BY KRITTIKA: THE ASTRONOMY CLUB OF IIT BOMBAY
GITHUB.COM/KRITTIKAIITB
First Release, August 2020

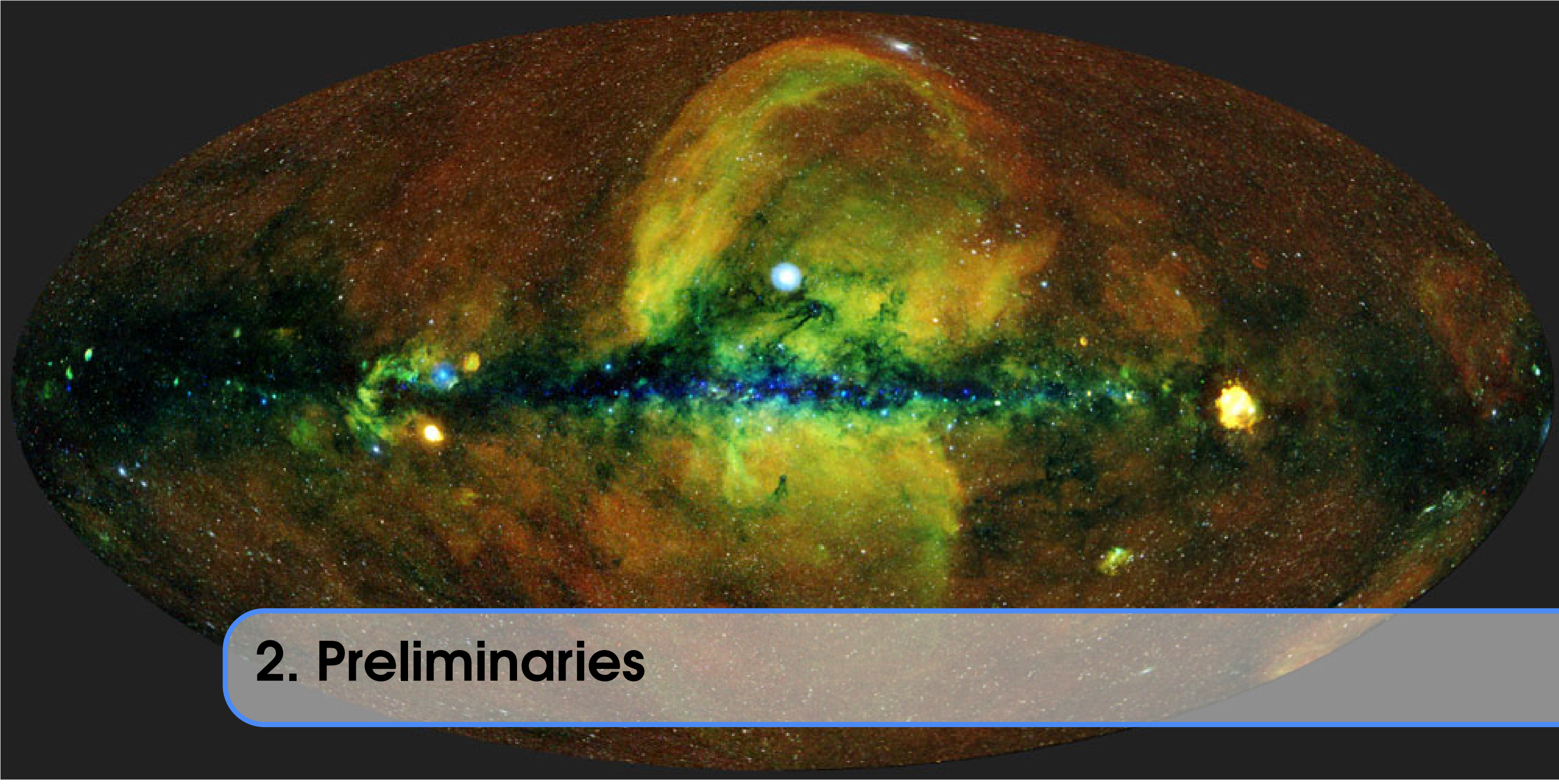
Contents



1. Introduction

Am I supposed to write this or reach out to people for this

Either is fine. If you can sketch out a structure that will also be useful and we can flesh it out later. But not now.



2. Preliminaries

2.1 Telescope parameters

2.1.1 Equatorial Coordinate System

The Equatorial coordinate system can be thought of as a projection of the latitude-longitude coordinate system we use on Earth, onto the Celestial Sphere. Lines of latitude on Earth become lines of **Declination** (Dec; measured in degrees, arcminutes and arcsecondsmaybe remove this?) and indicate how far North or South of the Celestial Equator (defined by projecting the Earth's Equator onto the Celestial Sphere) the object lies (see Figure 2.1. The Celestial Sphere equivalent to lines of longitude is **Right Ascension** (RA). Just as longitudes require reference direction (since all longitudes are equivalent, unlike latitudes which differ in size) of the Greenwich Meridian, the Right Ascension is measured from the Vernal Equinox (Figure 2.2)

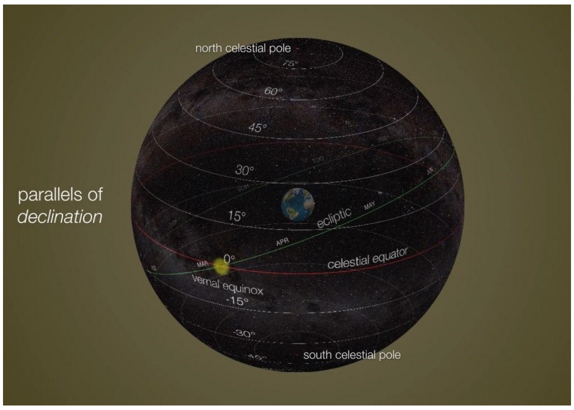


Figure 2.1: Parallel lines of declination with 0 as the celestial equator projected on the Celestial sphere

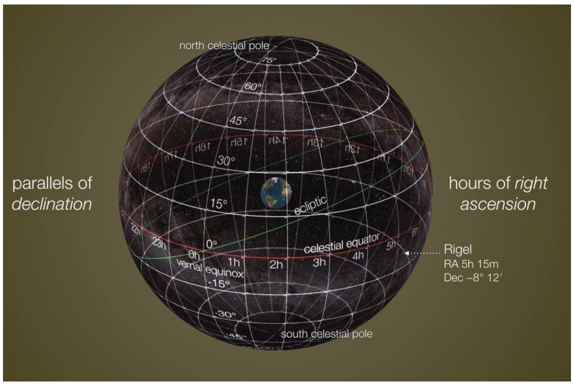
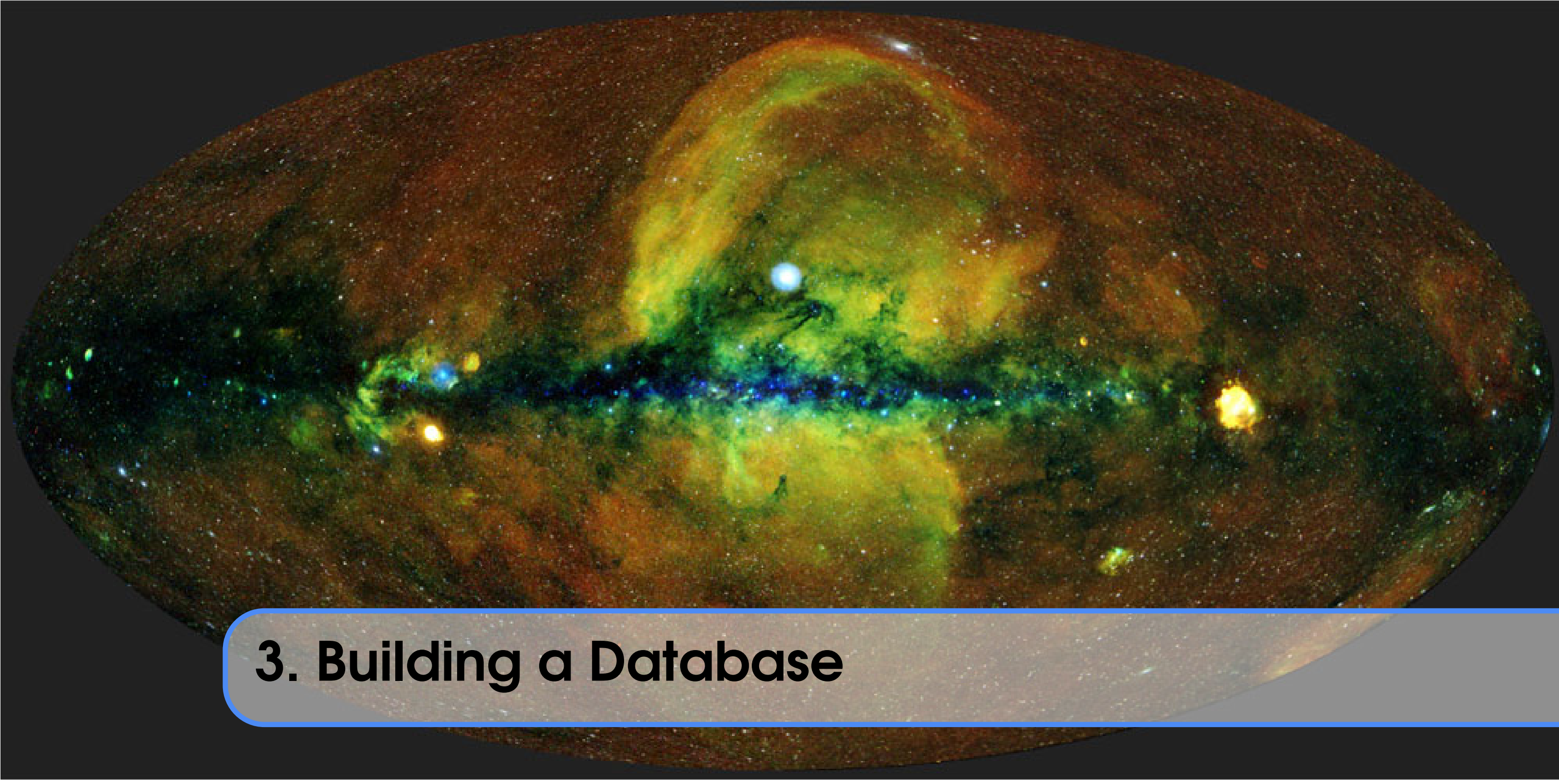


Figure 2.2: Hours of right ascension starting from 0 at vernal equinox

Conversion between Coordinates

The Celestial Equatorial Coordinate System is based on the concept of the celestial sphere. The celestial sphere is an imaginary sphere of infinite radius surrounding the earthEither say this earlier or remove. The celestial sphere is fixed with respect to the universe. Its orientation does not change. However, because the earth rotates from west to east (counterclockwise from the perspective of looking down at the north pole), an observer standing on the earth will see the celestial sphere rotate from east to west (or clockwise when looking up at sky). This cause a subtle difference in the way we perceive the celestial latitude(Declination) and celestial longitude(Right Ascension). Declination remains the same going up that is towards North is taken as positive and the other direction is negative. So the declination ranges from -90 and +90. In the case of right ascension the convention differs as it rotates from East to West. In terms of degrees of right ascension from the vernal equinox measurements towards right that is towards east is taken as negative and those towards left that is towards west is taken as positive. No nebulae, clusters, magnitude, telescope parameters, WHAT is hopping, anything about GUI???! Maybe we could cover in Intro but some should be prereqs too.

vedant: there is no coordinate conversion here, what was this section supposed to be? i would say ditch this entire chapter, but there is stuff to write here.



3. Building a Database

3.1 Data Extraction

For the extraction of data from the various online databases and archives, three tools were used. All the tools were used through Python scripts. There were other tools available, but these three were used for their applicability to Python. yes we getit, python. redundant sentences

3.1.1 Astroquery

Astroquery is a Python package of the Astropy module. It mimics the function of web interfaces which were created to access astronomical data. Astroquery is used to construct queries to access remotely-stored data. It combines the functionalities of the Astropy and Requests modules (Ginsburg (2019)). Astroquery allows, among other things, for cone searching, region searching, catalog requests, and identifier searches (Ginsburg (2020)). Astroquery also allows for application of search criteria similar to the capabilities offered by ADQL. ew green refrence color change

While astroquery allows for parsing/searching through specific databases/websites, we used astroquery to write queries to VizieR and SIMBAD, both maintained by the CDS. This is because Vizier is itself a database of various catalogues (Ochsenbein et al. (2000)).

3.1.2 Web Scrapping

For certain databases, where astroquery could not be used, we used the HTML parsing capabilities of BeautifulSoup to extract information.

Web scraping using python is a useful tool for extracting data from web pages, and was extremely useful in situations where the data obtained from astroquery needed to be supplemented with additional data.

3.1.3 ADQL/SQL queries

ADQL is a variant of SQL - a language which allows for constructing queries for online databases. Being a standardised form of SQL, ADQL allows communication with a host of astronomical databases. (ESA (2013)).

Like most SQL queries, ADQL queries consist of three main fields: SELECT, FROM and WHERE. Through the capability of joining multiple data tables and using mathematical operations in queries, ADQL allows for the construction of complex queries (ESA (2013)).

3.2 Catalogues

3.2.1 Star Catalogues:-

Several catalogues were considered for this project. These catalogues were mainly accessed through Vizier or Simbad. The choice of catalogues used was determined by the need for complete data on positions and brightness of all the bright stars in our neighbourhood. Since hopping requires prominent stars as guides, we deemed any catalog of stars complete at magnitudes beyond 12 to be superfluous.

The catalogues considered were:

- Hipparcos catalogue
- Tycho-1 and Tycho-2 catalogues
- Guide Source Catalogue of HST (GSC)
- NOMAD catalogue
- Gaia
- UCAC-4/5 catalogues
- Bright Star Catalog (BSC5P)

A short description of each is provided below.

Hipparcos

The Hipparcos catalogue is a catalogue documenting the positions and motions of the brightest stars in our vicinity. The Hipparcos mission, which was completed in 1997, catalogued details of over 100,000 stars. The catalogue is complete in the range of 7.3 to 9 magnitudes (Perryman et al. (1997)). The Hipparcos catalogue was not chosen due to the greater extent and completeness of the Tycho catalogues.

Tycho-1

The Tycho-1 catalogue is the successor to the Hipparcos catalogue (Egret & Fabricius (1997)). It contains around 1 million stars and is 99% complete down to 10.5 mag and 66% between 10.5 and 11.5 - the limit of its observations (Hog et al. (1998)). The catalogue contains both astrometric and tycho-band photometric data (Egret & Fabricius (1997))

The original Tycho catalogue omitted/failed to record 6301 stars present in the Hipparcos catalogue. This includes very bright stars like Sirius and stars beyond Tycho's observation limit. To ensure the completeness of the Tyhco-1 catalogue, these missing stars were added from the Hipparcos catalogue (Hog et al. (1998)).

The size and completeness of the Tycho database, along with photometric data, which is converted to Johnson's V band, makes Tycho-1 a suitable catalogue for this project.

Data from this catalogue was extracted using Vizier and integrated into python using astroquery. The data extracted includes J2000 RA and Dec, V band magnitudes, B-V colour index, and reference numbers for other catalogues.(ESA (1997)).

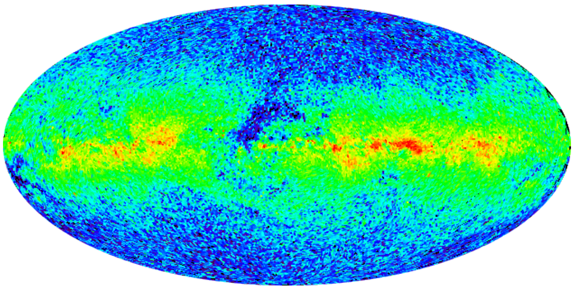


Figure 3.1: Representation of the coverage of the Tycho-1 stellar catalogue. This was taken from Vizier. The interactive plot, provided by Aladin, is accessible here.

Tycho-2

The Tycho-2 catalogue is a catalogue containing data on roughly 2.5 million bright stars. Observations used to create Tycho-1 were also used to create Tycho-2. Tycho-2 data however is more accurate. It boasts a 90% completeness upto 11.5 magnitudes. (Høg et al. (2000)).

While the Tycho-2 catalogue is much larger than Tycho-1, it is not complemented by Hipparcos data for the brightest of stars. As a result, it is missing several bright stars present in the Hipparcos and Tycho-1 catalogues (Høg et al. (2000)). This motivated us to stick with the Tycho-1 catalogue to cover stars upto its 90% completion limit, and insert Tycho-2 catalogue data beyond this limit up to 12 mag.

Data extracted from the Tycho-2 catalogue includes J2000 RA and Dec, and Tycho band photometry (Hog et al. (2000)) The Tycho-2 photometry data available in Vizier has not been converted to Johnson's filter. Hence a simple conversion was used to obtain the Johnson V band magnitudes, maintaining consistency with the Tycho-1 data table (Egret & Fabricius (1997)).

$$V = V_T - 0.090 * (B_T - V_T)$$
$$where, -0.2 < (B_T - V_T) < 1.8$$

NOMAD

The Naval Observatory Merged Astrometric Dataset (NOMAD) is a dataset composed of data from the Hipparcos, Tycho-2, UCAC-2, and USNO-B catalogues. The data for each star is the best data out of these datasets. The NOMAD catalogue comprises around 1 billion stars. Along with photometric data from its component surveys, it also contains infrared photometric data from the 2MASS survey (Zacharias et al. (2004)).

We were directed to NOMAD while researching the catalogues used by Stellarium. While NOMAD offers a greater flexibility in terms of magnitude limit, since the need for stars up to mag 12 was already being met by Tycho, this catalogue was deemed redundant for our purposes.

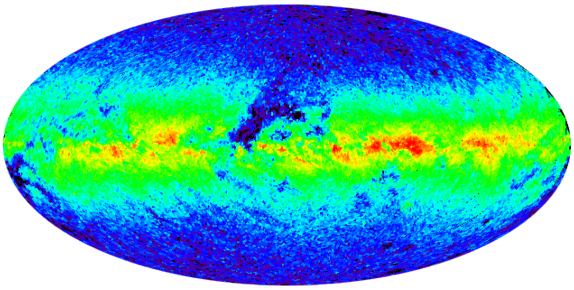


Figure 3.2: Representation of the coverage of the Tycho-2 stellar catalogue. This was taken from Vizier. The interactive plot, provided by Aladin, is accessible through https://cdsarc.u-strasbg.fr//vizier/cat/aladinLite.html?I/259.

Gaia

Gaia is the successor to the Hipparcos mission launched by the ESA. It aims to measure parallaxes, coordinates and proper motions with greater accuracy. Gaia started operations in 2013 and aims to provide a better understanding of the Milky Way. Gaia has also recorded the photometric data of stars in the Milky Way upto 20 magnitudes in the Gaia G-band (Gaia Collaboration (2016)).

Access to Gaia was provided by Gaia TAP+ feature of astroquery (Carlos Segovia (2016)). Queries in TAP were written in ADQL. Gaia's coverage for some of the brighter stars is poorer than that offered by other catalogues (ESA (????)). It also offers more data in the fainter end than is required by this project. As a result we chose not to use the Gaia catalogue.

Bright Star Catalogue BSC5-P

The Bright Star Catalogue contains data on stars whose magnitudes are lower than 6.5. It provides the IDs, classifications, spectral types, proper motions, coordinates, UBVRI photometry and other information on their subject stars (Hoffleit (1964) Hoffleit D. (1991) HEASARC (2014)).

3.2.2 Extended objects catalogue

The objective of this project is to provide hopping sequences from prominent stars and constellations, to various objects with distinct features that are easily visible using a telescope. Objects which are not point sources in the night sky, are called extended objects. These include galaxies, nebulae, and clusters. Several catalogues of extended objects have been compiled over the years. Two of the most famous extended object catalogues are the Messier and the NGC catalogues.

Messier Catalogue

The Messier catalogue is a catalogue of celestial objects compiled by Charles Messier in 1781. The Messier catalogue consists of 110 objects known as deep sky objects. This catalogue forms the central theme of this project (Rob Garner (2014)). Data on the messier objects was extracted from SIMBAD using its catalog search feature. SIMBAD sources photometric and astrometric data on the messier objects from recent surveys such as GALEX. The dataset this project uses contains the ID, J2000 coordinates, object type, distance, and B, V magnitudes of the messier objects. V band magnitudes of the messier objects extracted from astropixels (api

(2012)) were added to the project dataset. The data available on this website was sourced from SEDS (Frommert et al. (1994-2013a)). This data was added owing to the incompleteness of photometric data from SIMBAD.(NOTE: NEED TO INFORM THE TWO PAGES THAT WE HAVE USED THEIR STUFF FOR NON-PROFIT PURPOSES)

NGC Catalogue

The NGC 2000.0 catalogue is composed of the original NGC, IC and Second IC catalogues written by JLE Dreyer (HEASARC (2015)). Dreyer was driven to compile the original NGC in 1871 due to discrepancies in the earlier data compiled by John Herschel in the "General Catalogue of Nebulae" as well as the discovery of additional "nebulae" (ngc (1878)). Since then, many astronomers have added to and made corrections to the data. This led to the creation of the NGC 2000.0 catalogue (ngc (1998)). The NGC 2000.0 dataset obtained from VizieR contains the IDs, RA and Dec (J2000), constellations and the V band magnitudes of several types of objects.

3.2.3 Cross Reference Catalogues

While the catalogues mentioned above had astrometric and photometric data, they were missing a few details, such as the names of stars and extended objects, and Bayer and Flamsteed designations. Addition of these details would fine-tune the available data tables for the project's needs. Since this data wasn't always available with the larger catalogues, other sources of data on Vizier and elsewhere had to be used to add this information. A brief description of the supplementary data and their sources is given below.

Star Names

During a hop, several bright stars are required as reference points. These stars are best known by their common name, such as Sirius or Aldebaran. To make it easier to guide the user, common names of stars were required.

The IAU's Working Group on Star Names(WGSN) has catalogued a list of star names that are based on the most popular historical names for stars (IAU WGSN (2018)). The catalogue used by this project is an extension of that provided by the IAU. Along with the names and HR numbers of stars, it contains the HIP and HD numbers. Since these are present in the Tycho catalogue, they are ideal for cross-referencing. The catalogue was downloaded off Kaggle (kag (2018)).

Bayer Designations

As per the IAU WGSN, there are only 330 official star names. Since hopping may require stars apart from these stars, an additional naming scheme was required.

The Bayer scheme uses a greek/roman letter plus the name of the star's constellation to reference a star. The brightest stars are usually designated by the letter α. An example of this is the star Aldebaran, which has the Bayer designation α Tauri (Blackman (????)). The Bayer designations for stars were obtained from the HD-DM-GC-HR-HIP-Bayer-Flamsteed Cross Index catalogue available in Vizier. It is constructed using several catalogues, thereby serving as a cross referencing catalogue (Kostjuk (2002)). The HIP number provided in the catalogue was used to cross match the Bayer designations to their corresponding stars.

Flamsteed Designations

The Flamsteed naming scheme involves naming all stars with a number, indicating their proximity to the western edge of the star's constellation, and the name of the constellation. An example of this would be the star 70 UMa in Ursa Major (Blackman (????)). We have assigned stars a Flamsteed Designation only when the Bayer designation is missing. The Flamsteed number is obtained from the same catalogue as above.

Messier common names

To make it easier for the user to access hops, the common names of Messier objects were added. The Wikipedia page on messier objects was scraped to obtain the common names, although the source for this is the SEDS webpage for Messier objetcs (Frommert et al. (1994-2013b)).

NGC common names

Like the Messier common names, names were added to the NGC catalogue in order to make it easier to access objects. The common names were added to the NGC catalogue using an appendix to the NGC 2000.0 catalogue available through Vizier (Sinnott (1988)).

3.2.4 Miscellaneous Catalogues

In addition to all of the catalogues mentioned above, a catalogue of IAU-approved constellation boundaries was used for the project (Davenhall & Leggett (1989)).

3.3 Explanation of Code

In order to build a database for the project, a program was constructed to download data from the various online catalogues mentioned above. This program's code can be divided into three broad sections:

1. Download and manipulation of the extended object catalogues.
2. Download and manipulation of Tycho-1 catalogue.
3. Download of Tycho-2 catalogue.

All the code was written in Python 3.7, using either Spyder or Jupyter notebooks. Table manipulations are performed using Pandas and Astropy Tables. The code documentation will cover the details of the code in greater depth. which code documentation, link to github?

3.3.1 Extended Objects

Constellation Borders

To obtain the constellation borders, the Vizier class of astroquery is used. The get_catalogs method not graying this? of Vizier is fed the argument "VI/49" corresponding to the vizier catalogue one wants to download. The `get_constellation` function of astropy coordinates is used to add the full name of each coordinate (short forms are already given).

A small change that is made by the program immediately after each call of `get_constellation` is changing the name of the constellation Boötes. Since Boötes contains a utf-8 character, it isn't displayable unless encoding is changed to utf-8. Thus to make the data more universal, the program changes it to Bootes.

Messier Objects

Before downloading the messier objects, the program scrapes the astropixels web page on Messiers in order to get the visual magnitudes.

Simbad was used to download the entire messier catalogue. To do this, the Simbad class function `get_catalogs` is used with the argument "Messier". The output is a table of the VOTable format. The columns in this VOTable can be selected from a host of available options.

As is the case with many online catalogues, an internal identification is maintained for easy navigation and access to each entry in the catalogue. The internal IDs used here consist of the object type, constellation, and the object ID, separated by underscores. The catalogue download is followed by the addition of common names to the catalogues. These common names are scraped off Wikipedia.

NGC Catalogue

The NGC 2000.0 catalogue is downloaded from Vizier (VII/118). The `Vizier` class was used to choose only relevant columns from this catalogue. In order to make the names and types more consistent with the Messier data, as well as more uniform, the program uses list comprehension to make modifications to the names and types columns.

After adding constellation names and internal ID number, the program moves onto adding common names to the NGC catalogue. The common names are extracted from an annex to the NGC 2000.0 catalogue on Vizier. The downloaded file's NGC ID column is then changed to be consistent with the previously downloaded catalogue. The NGC ID's are then used as an index to cross match the common names onto the main NGC 2000.0 catalogue.

Plotting both the Messier and NGC catalogues, we noticed a lack of coincidence between the Messier objects from both the catalogues. To remove this, the RA and Dec of the Messier objects alone in the NGC catalogue are replaced by the RA and Dec values from SIMBAD, the source for the downloaded Messier catalogue. This was done by using the common names (which contained Messier IDs) as an index.

3.3.2 Star Catalogues

Star Common Names

The file on common names of stars was downloaded off Kaggle. The draft version of the database construction program included the python API for Kaggle which allows for downloads of files from Kaggle. This API, however, is complicated to use. Thus, the final program doesn't incorporate the Kaggle API. Rather the program is shipped with the Kaggle file.

Bayer and Flamsteed Designations

The database construction program downloads the Bayer and Flamsteed designations from Vizier (IV/27A). The output file contains not just the Bayer and Flamsteed designations, but also the HD, HIP and Flamsteed numbers. It also contains a column of the constellations of each star in the file. Since a Bayer or Flamsteed designation carries meaning only when paired with the name of the constellation, the program appends the constellation name to the end of the designation. The program then attempts to fill any missing HIP values in the Bayer designation catalogue by feeding the corresponding HD values into SIMBAD. This is done solely because the Tycho-1 catalogue to be download has a more complete record of HIP numbers.

Tycho-1

The download of the Tycho-1 catalogue is performed through Vizier(I/239) and executed in three runs. These runs are based on filters of V band magnitude applied through the Vizier class column filter feature. These three runs are from brightest to 6 mag, 6 to 8 mag and 8 to 10.5, which marks the limit of maximum completeness of Tycho-1. The objective behind this scheme was to offer the user flexibility to choose which parts of the database they desired. The initial draft of the program was designed to take V mag inputs from the user and apply those as column filters. This is, to a degree, a false sense of flexibility, as the user is anyways constrained to a limit of 10.5 mag. The user input feature was removed and replaced with fixed runs.

After the download of a file during a run, the constellations and internal IDs are added in a manner similar to previous sections of the program, but with one small change - the internal ID contains an 8 digit number. The final number in each run is stored as a variable and used for the succeeding run. This means the internal ID numbering continues for each run where the last run left off.

The common names and Bayer/Flamsteed designations are added to the Tycho-1 table by using the HIP catalogue numbers as a cross matching index. In addition to the HIP numbers, the HD numbers are also used to fill up missing Bayer/Flamsteed designations.

After each run, the program increases the magnitude limits by 3, till a maximum of 10.5, and executes another run. After each run, the resultant data table is appended to the table generated in the previous run. When the database is complete till 10.5, the Tycho-1 download loop automatically terminates.

Tycho-2

Before commencing the Tycho-2 download, the program will ask the user if they wish to continue. Since this file is large and covers the fainter end of our database, the download of this file was left as optional to the user.

The Tycho-2 data file is downloaded through Vizier (I/259). Unlike previous downloads, however, Johnson's V band magnitude is not available as a filter. The magnitude filter available is through the Tycho band system. It was discovered that filtering from Tycho mag 10.5 to 12, and then converting to Johnson's V band resulted in Johnson's V band magnitude values beyond the limits. This also raised concerns regarding Tycho band values below the cutoffs that, upon conversion, lie within the Johnson's band cutoff. Thus, to make sure all stars within 10.5 to 12 range in Johnson's V band are selected, the program filters for a larger range in the Tycho band, i.e. 10.2 to 12.3, and then converts to Johnson's system. This large interval provides confidence that all stars between 10.5 and 12 have been selected. After applying the transformation, the table is filtered and only stars between 10.5 and 12 are kept. The Tycho B-V difference is also converted to Johnson's B-V colour index.

Constellations names and internal ID are added as before, but neither the common names nor Bayer/Flamsteed designations are added as the hops will be confined to stars below 6 magnitudes.

3.4 Caveats and Issues

- The mapping of Bayer and Flamsteed designations is not perfect - missing HD/HIP values in either the Tycho-1 or cross reference catalogue make it hard to ensure all the values from the cross catalogue are matched onto the Tycho-1 table.
- During the replacement of RA and Dec of NGC catalogue Messiers, the program searches through common names for words starting with M. Thus, the Ra and Dec has not only been replaced through SIMBAD for Messiers, but also other objects such as (name). These outliers, however, are few.
- Some of the large catalogues such as GSC, NOMAD and UCAC 5 were discounted partially because they could be used only with region search capabilities. A potential workaround we explored for this was to build up several region searches covering the entire area of the night sky, for a given magnitude range.
- V band magnitudes for several Messier objects is missing from the SIMBAD database. To complement this data, V band magnitudes from SEDS was added onto the downloaded catalogue. In many cases the magnitudes from these two sources disagree. The SIMBAD photometric sources for Messiers (such as the GALEX survey) are more recent than those of SEDS. However, since the plotting requires completeness in magnitudes, the SEDS magnitudes were kept.
- After download, the Tycho-2 table is kept separate from the the combined Tycho-1 table - the two files are not outer joined. Owing to storing a large table (~150 MB) onto memory before joining, the joining of Tycho-2 table onto Tycho-1 results in the program crashing due to memory overuse. To avoid such issues on other computers, the tables are kept separate.
- The conversion formula for Tycho band to Johnson's band magnitude was applied rather indiscriminately - the BT-VT values often lie beyond the limits for the application of this formula. Nonetheless, the formula was applied, keeping in consideration the errors in Vmag that will occur as a consequence.

weird empty page in middle

4. Interface

4.1 Editor Interface

The very first and crucial step towards building Star Hopping tool is creating a catalog file of hops for all the extended source in the sky (like Messiers, NGCs etc.) using the database we have created and an interactive interface. This makes it natural to have two interfaces in the project- one for the editor and other for the end user. Editor interface will aid the editor in selection and storing of hop-sequences in a convenient and logical manner. The idea is implemented in two parts:

- The Interactive Plot
- The Hopping Algorithm

4.1.1 The Interactive Plot

is there supposed to be something here

4.1.2 The Hopping Algorithm

Motivation and Introduction

Star hopping is quite interesting and a great way to learn the night sky that a lot of amateur astronomers often use. But sometimes it can become too tiresome and since we had to prepare an entire hopping catalogue, it would have been difficult to do all the hops, manually. So, to ease this work of manually locating stars in the sky plot, we devised an algorithm that would do the hopping for them.

We needed an algorithm to return a hop whenever an editor would click at a point in the interactive star plot. Not to confuse hop with something special, it would just be a star in the sky. The algorithm that is used in this project is a rather simple one which owes the simple problem that it had to cater for finding hops. very weird sentence - need to frame better

Problem Statement Breakdown

Whenever designing any algorithm, the easiest way to go about it is to break down the problem mathematically and then start thinking of what can be done to solve it not needed thx for gyaan. A mathematical breakdown of the problem would be: Consider a 2-D graph with some fixed points and an editor can click at any arbitrary point within the graph. For each click, the algorithm must return the closest fixed point to that clicked point. Here, the fixed points refer to the **stars** no need to bold?? in the sky, the clicked point would be a rough estimate of the editor trying to hop the nearest star and the returned fixed point by the algorithm would be the **hopped star**. redundant

Further, the stars would be plotted considering their RA values as x-coordinates and DEC values as y-coordinates. Hence, they could be thought of as 2-D points only with `(ra, dec)` as coordinates. The clicked point by the editor would also be just a pair of (x,y). The problem just demands the returning of the closest `(ra, dec)` coordinate to the clicked (x, y).

First Outlook

After the mathematical breakdown, the problem seemed quite easy - After this, we calculate, which is just calculating the Euclidean distance between the clicked point and all the fixed points and choosing the minimum value of those. Such a solution would also be logically pleasing because of its time complexity being O(N). But the catch here is that N would be the number of stars in the database which is quite huge, around 1 million! In such a case, a simple for loop used for calculating the distance would also become tremendously slow.

To fix this up, the approach followed overall is reducing the value of N, in some way. Reducing the value of N doesn't require manipulation with data, it just demands taking some obvious steps to cut down unnecessary calculations which obvious steps - assuming this is the square cutoff explained below, so this sentence redundant. A new temporary dataset is created as the part of the algorithm and is used for the distance calculations.

Making the algorithm work

For designing a new dataset, first all the farther stars and even the medium-distanced stars were removed. This was the most logical step that anyone would take while calculating the minimum distanced star - ok buddy. Pictorially, it is quite easy to eliminate the far away stars but seeing only numbers as (x, y) points, it's a bit tricky to eliminate stars without calculating the distance between them. Hence, a rather pictorially intuitive solution was thought for this.what is this sentence Considering the clicked point to be the center a square with some length would be drawn theoretically and any points lying outside that would obviously be farther than the ones lying inside. Hence, only the points inside the square would be included in the new database for calculating the Euclidean distances between the clicked point and the stars in the new database. The algorithm incorporates this by checking if both the coordinates lie within certain limits which would be the boundary lines of the "limiting square". The length of the square is left up to the editor to decide as per the ra and dec values. This is done so that an optimal solution can be thought of which doesn't lead to too many stars being inside the square and not even removing all of them.

The figure shows the easiest way of representing what has been trying to convey in previous paragraphs which figure. It can be clearly seen that the Point 1, which is inside that square is closer to the clicked point as compared to the Point 2, which is outside the square. Infact, Point 2 is farther away from all the points inside the square. ????

Even after including all the proximal stars in the database, the value of "reduced N" might be still large and that would still reduce the speed of the program. Such a case would arise when the editors would have clicked at a point with too many proximal stars. It would not be reasonably correct to assume further smaller squares, because it would involve iterations of choosing the correct square that has just one nearest star inside it.

Hence, another approach based on the visual magnitude was used to eliminate stars further. We set a limit such that if the number of stars gets greater than that limit then we would only take the brightest "M" of those. Here, M is the brightness limit that would be again left to the editor to be decided. Logically, as well it is good to assume that when an editor clicks at a point, he would not be aiming for the faintest of the star to be selected. not he, they

Error Handling

Whenever any application is built to be used for some external purpose and would be used by people who did not write the code, it is always important to take care of the corner cases since nobody would like the program to halt in the middle. - did not know that thx Hence, some error handling was done in the cases where the editor might click at a boundary point, with no proximal stars inside the 'square' and also if the clicked point has more than 1 star at the same distance from itself. This way, the editor will get an error message in the end and would have to do those hops again without worrying about them while the program runs.

Final Moment Additions

This addition was done and thought of while writing this report itself. There was one point where we were going entirely wrong. It was when we were trying to remove the stars outside the boundary lines of the square to make the new database. While coding this logic, we were using 'for' loop for comparing the x and y coordinates of stars with the limiting boundary points, which would require "N" complete iterations with the original number of stars, which would in no way mean the reduction of time complexity. Hence, the `for` loop was replaced by a `pandas` hook instead which goes something like this: why is this commented where `a` is the column name and `b` is the limit to compare with all the values in that column `a`. This way multiple filters can be added in a single line and remove the undesirable stars without using a 'for' loop. And undoubtedly, it made the code a lot faster and efficient! This was one of the takeaways from designing the algorithm to hop stars.- ok.

Scope of Improvement

The current algorithm will work best when the all limits would be correctly decided. It is advised to read the coding documentationlink? before to understand every term properly. However, the algorithm is still of time complexity **O(N)** type. If an algorithm of time complexity **O(log(N))** could be designed to achieve the same target then it would be much more time efficient.

Keyword	Description
llcrnrlon	longitude of lower left hand corner of the desired map domain (degrees).
llcrnrlat	latitude of lower left hand corner of the desired map domain (degrees).
urcrnrlon	longitude of upper right hand corner of the desired map domain (degrees).
urcrnrla t	latitude of upper right hand corner of the desired map domain (degrees).

Table 4.1: Defining map based on two corner locations

4.2 User Interface

????

4.3 User Plot

4.3.1 Introduction

The objective is to plot the database that we created. Given the right ascension and declination of a target create a function that parse through the entire database and return plots showing projection of the sky in relevant areas of observation. The final output displayed to the end user will be consisting of three plots. Two of them showing the field of view through the eye piece and the finder scope of the telescope, and the third plot displaying stereographic projection(**?**)wrong cite imo of broad area of sky centered at target which is received as input from user. The final hops to the target will be plotted on the third map.

4.3.2 Method Implemented

Basemap toolkit

The matplotlib basemap toolkit (Whitaker (2011)) is a library for plotting 2D data on maps in Python. Basemap does not do any plotting on it's own, but provides the facilities to transform coordinates to one of 25 different map projections (using the PROJ.4 C (Warmerdam et al. (2009)) library). Matplotlib is then used to plot contours, images, vectors, lines or points in the transformed coordinates.

Stereographic Projection using Basemap toolkit

In order to represent the curved surface of sphere on a two-dimensional map, a map projection is needed. Since this cannot be done without distortion, there are many map projections, each with it's own advantages and disadvantages. Basemap provides 24 different map projections. Some are global, some can only represent a portion of the globe. When a Basemap class instance is created, the desired map projection must be specified, along with information about the portion of the sphere's surface that the map projection will describe. There are two basic ways of doing this.

Table 4.1 defines the map by specifying the lower left and upper right corner latitude and longitude coordinate values respectively.Such a plot can be made use of when dealing with projections which requires specific field of view or only some relevant area of of the map need to be displayed. Table 4.2 defines the map by centering at the given latitude, longitude values and shows plot with the given width and height in meters. Such plots can be used when we want to display a broad region of the map centred at some specific points.

Keyword	Description
width	width of desired map domain in projection coordinates (meters)
height	height of desired map domain in projection coordinates (meters).
lon_0	center of desired map domain (in degrees).
lat_0	center of desired map domain (in degrees).

Table 4.2: Defining map based on width, height and centre

Parallels, Meridians and Map Coordinates

Most maps include a graticule grid, a reference network of labelled latitude and longitude lines. Basemap does this with the `drawparallels()` and `drawmerideans()` instance methods. The longitude and latitude lines can be labelled where they intersect the map projection boundary. The declination circle runs from 90 degree to -90 degree with gaps of 10 degree. We can set the maximum declination circle that would be plotted, the default set to 80 degree is changed in these plots to 90 degree. Since celestial coordinate system is used the circles of right ascension runs from 0 to -180 degrees towards East and 0 to +180 degrees towards West.

In order to plot data on a map, the coordinates of the data must be given in map projection coordinates. Calling a Basemap class instance with the arguments `Ra` , `Dec` will convert Ra/Dec (in degrees) to x/y map projection coordinates (in meters). So if Ra, Dec are the arrays of data coordinates they are converted to map coordinates by `x,y=m(RA,Dec)` where `m` is the basemap variable.

4.3.3 Telescope PLots

The aim of these two plots is to aid the navigation to target object through telescope. Centered at the target(Messier Object) two plots would be shown one through the eyepiece and one through the finder scope where the field of view will be calculated by the input parameters from user.

Setting up the Plots

The right ascension and declination of the target is stored from the Messier user enters. A basemap variable is created centered at the target but showing only the relevant field of view. This is done by giving the coordinates of two corners of the plot: the bottom left corner Ra, Dec and the top right corner Ra, Dec as mentioned in method one of section 4.3.2. The coordinates of the two corners are taken to be the coordinates of the target +/- field of view. Only the stars lying below the limiting magnitude of the telescope are shown in the plot with their sizes proportional to their apparent magnitude. The value of limiting magnitude, field of view of eye piece and the finder scope are calculated from the input parameters from user.

Scatter plotting the data

The stars from the database are scatter plotted in the projection as white dots, while the Messier objects are divided based on their type into open clusters, globular clusters, galaxies, nebula and supernova remnant and other messiers. Each of these types are then plotted using different markers. Green dots in the map represent the constellation borders.

The sizes of the stars are plotted as proportional to their magnitude, using the equation 4.1. V is the apparent magnitude of the stars and k is a constant, it's values can be changed to increase the star sizes in the plot proportionally. By trial and error for the telescope plots the constant `star_constant` is set as 500. The right ascension values in the database ranges from 0 to 360 degree use degrees?. It is converted to celestial coordinate, 0 to +180 degrees towards West and 0 to -180 degree towards East. Before plotting the dataset, Ra and Dec are then further converted to map coordinates.The parallels and meridians that is the right ascension and declination lines in the respected field of view are also added on to the plot.Fig 4.1 and Fig 4.2 shows sample telescope plots for the Messier M6, commonly known as Butterfly cluster.

$$s = \frac{k}{2.5^V} \tag{4.1}$$

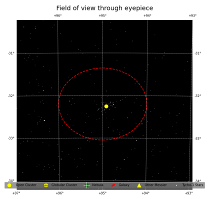


Figure 4.1: Field of view through eye piece

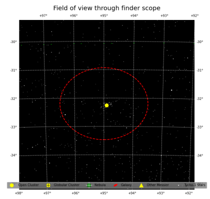


Figure 4.2: Field of view through finder scope

Adding Patch Circle

The plot displayed will be of the width twice the field of view. To display the actual field of view or exactly how the sky will be visible through eye piece and finder scope a red dotted patch circle is added with field of view in degrees as the diameter.The radius of the circle need to be calculated in the map coordinates. Since stereographic projection is used the radius of the patch circle vary based on the coordinates of the centre. The reason being that circles of right ascension aggregate together as we go towards the poles, that is distance in map coordinates between two points say 5 degrees apart is different based on their declination. So basically radius need to be computed each time based on the target coordinates. A direct approach by initialising radius of circle to field of view is not possible.

4.3.4 Hop Sequence Plot

The objective of this plot is to display hops to target object. Centered at the target messier, a broad region of sky is displayed with the hops connected by lines terminating at the Messier object. Name of some prominent stars, all constellations, constellation boundaries and Messier ID's of all objects are displayed on the plot for the ease of user to hop. The axis of right ascension is inverted to resemble the sky as seen from earth, otherwise the plot will be mirror image of the sky that we observe.

Plotting the Dataset

A basemap variable is created centered at the target object showing a broad region of sky. The area of the map can be altered by changing the value of variables `width` and `height` as discussed in the method 2 of section 4.3.2 . Only the stars lying below the visual magnitude (6) are plotted, since the brightest of the stars will be used for making hops.

The stars from the database are scatter plotted in the projection as white dots, while the Messier objects are divided based on their type into open clusters, globular clusters, galaxies, nebula and supernova remnant and other messiers. Each of these type are plotted then using different markers. Green dots in the map represent the constellation borders.

The sizes of the stars are plotted as proportional to their magnitude, using the equation 4.1. V is the apparent magnitude of the stars and k is a constant, it's values can be changed to increase the star sizes in the plot proportionally. By trial for the hop sequence plot the constant `star_constant` is set as 200. 200 or 500? this section seems similar to the last one, last one is redundant.Similar to the telescope plots the values of right ascencion are converted to celestial coordinates before plotting. Fig 4.3 shows a sample hop to the messier M6.

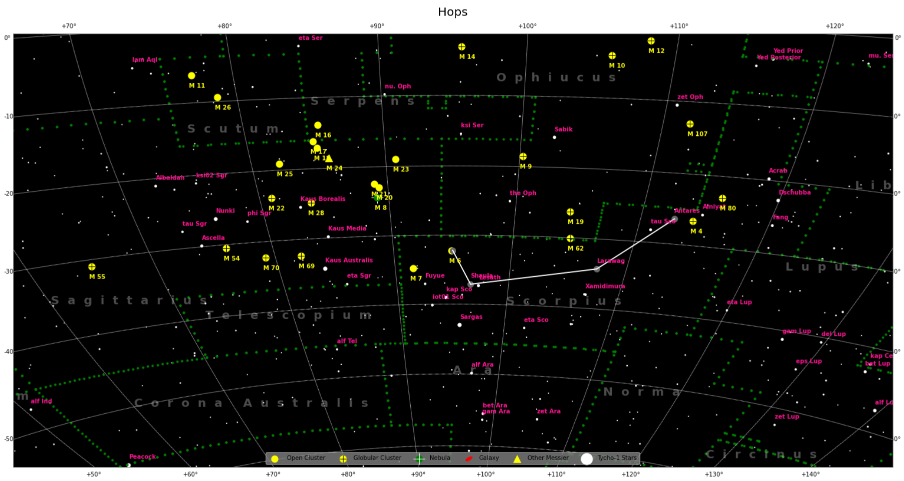


Figure 4.3: Star hops to M6:Butterfly cluster

Adding names on the Plot

The hop sequence plot will be the one used to show star hops to reach the Messier. To ease the hopping, names of prominent stars and all the constellations are added. Adding names of all the stars shown on the plot is surely not a good option since that bulks up the plot, and moreover most stars don't have names. So the plot basically shows the names of 300 brightest stars in the sky. To include more star names the value of variable `name_lim` can be changed. Common names are given for most stars and Bayer designation are used for those stars that doesn't have common names. The Messier objects are named with their Messier ID(M1 to M110).

Regarding the constellations we have the data of constellation boundaries and their names from the database. But working out a position to display the names from the data such that every constellation name was at the right coordinates was not possible. Various combination to find the coordinate for naming were tried, which are explained in section 5.2.2. The best way out of this was found to manually try out the 88 constellations and find a location best suitable for each constellation. The arrays of `ra_const` and `dec_const` gives the coordinates for naming the constellations.

4.3.5 Caveats and Scope for Improvement

The issues that raise while working with the plots are mainly due to the nature of stereographic projection. The aggregation of right ascension circles towards the poles is one major point. Following are the problems faced during plotting, these are discussed briefly in section 5.2.2

- Telescope plots nearing poles shows an error.
- Telescope plots tends to narrow down as we go from decination of 0 degrees to poles.
- The common way to label constellations was not found.
- At some targets the hop sequence plots can be found to be crowding up.
- The distance between star names and objects tend to increase as we zoom in.

All of the three plots given as output are made using the stereographic projection method of the basemap toolkit. The plots are done to aid hopping to Messier objects. The plot satisfies the idea for hopping we had at our hand but betterment of the plot could be done with future updates. Following are some points where we can improve the plots:

- The plots that currently works with 110 Messier objects could be modified to include deep sky objects in NGC catalogue.
- The plots differentiates between constellations by the constellation boundaries that are plotted on the map. Constellation sticks could be added here to aid the user.
- More efficient method to name the constellations could be arrived at rather than manually doing that task.
- Star are plotted with their sizes proportional to their magnitude, but all the messiers are plotted with the same size irrespective of their apparent magnitude.
- The plot currently displays white stars on a black background and addition to invert the color combination could be added.
- The hops are shown as lines connecting the stars, a better approach for hopping with angular separation and directions between hops could be added.

5. Replicating this work

5.1 Database

5.1.1 Database construction program

The program used to construct this project's database was written entirely in Python. A basic understanding of Python is more than sufficient to reconstruct such a program.

Astroquery can be easily picked up from the documentation that is provided on the Astroquery GitHub repository. This GitHub repository, which also contains the Astroquery API, is actively maintained by a large team of coders. This means most issues are proactively resolved, and updates rolled out quite frequently. The documentation contains very commonly used examples and thus does not require long hours of learning and practice. In addition to the documentation, the repository also has a forum for clearing any issues.

5.1.2 Catalogues

There is no dearth of astronomical data available on the internet. In fact, simply google searching astronomy catalogue should bring up many guides to selecting catalogues.

This project can be expanded to include objects from several other catalogues. Some other famous catalogues are the UGC, Caldwell, GC and LDN catalogues. The stellar database can also be expanded beyond 12 mag by using catalogs such as Gaia and UCAC-5. The object content can also be expanded upon by including objects from surveys of different wavelength regimes, although this may not be very useful for amateur astronomy.

In order to add a new catalog, simply search for it on Vizier, learn to download catalogues from Vizier and save the catalog to a csv file. At that point it is ready to use for plotting. why commented

5.2 Interface

5.2.1 User Interface

???

5.2.2 User Plot

The three final plots displayed to the user:two of the telescope plots and one hopping sequence plot are the stereographic projection of the celestial sphere, not a simple Ra Vs Dec scatter plot. The stereographic projection proves to be a suitable choice for the plots, but it does have slight issues that though not significant enough exists nevertheless. These caveats are addressed below:

Plots nearing poles

The telescope plots which gives the field of view through the eye piece and finder errors out at poles. This happens because these plots are defined by giving the coordinates of two corners of the plot, the lower left corner and the upper right corner. Further these corners are calculated as the coordinates of the target +/- field of view. So for an input with declination of 90 or -90 the declination value of one of the corner crosses the range(-90,90). So basically the telescope plots are available for the inputs such that `dec_in+fov<90` and `dec_in−fov>−90`, where `dec_in` is the declination of target and `fov` is the field of view.

Rather than showing an error at poles this problem is overcome by displaying a plot that shows a broad region of sky centered at the target. So every time declination of one of the corner crosses the declination limit, a plot displaying broad region of sky will be given as output rather than a plot in the given field of view. This issue doesn't raise a significant problem for hopping since all the Messier objects to be hopped lies within the declination of -35 degree and 70 degree.

Narrowing of the Telescope Plots

The telescope plots are found to narrow down as the declination of the target changes from 0 degree to either poles. This happens essentially due to the crowding of the right ascension circles as we approach poles. In the stereographic projection the distance in map coordinates between two points separated by the same degree of Ra vary with their declination, that is a distance in map coordinates of 10 degrees Ra is different when declination is 0 degree and when declination is 70 degree. So because of this, area of the plot based on the coordinates of corner gets reduced and gets narrowed down.

This issue can be solved by tempering with the aspect ratio. Basemap by default fixes the aspect ratio of the plots this can be override by setting the function `fix_aspect=False`. It prevents from the plot to be narrowed down but raises another slight issue that the red dotted patch circle spreads out resembling an ellipse. This is bound to happen since stretching a plot with circle to fix the aspect ratios, will elongate the circle. Fig **??** and Fig **??** shows the plots with `fix_aspect` set as `True` and `False`

Plotting Constellation Names

The database provides with the data for constellation names and constellation boundaries but figuring out a position that best displays the constellation name was not possible. Many approaches to attain coordinates were tried, the first idea

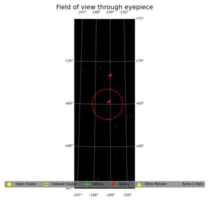


Figure 5.1: Plot with aspect fixed by basemap by default

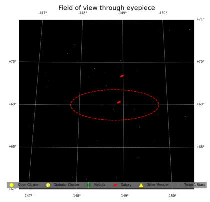


Figure 5.2: Plot without fixing the aspect which results in the broadening of the circle

was to calculate the mean of Ra and Dec values of individual constellations using `groupby()` and `aggregate()` functionalities of pandas **?**. Then the naming was tried at coordinates by calculating the mean declination of constellation and minimum right ascension. Another way by naming the constellation at or near the brightest or one of the dimmest star was also tried. Though these approaches came out to correctly name many of the constellation the entire 88 constellations could not be named by any of them.

The way this problem was tackled was to try out all of the 88 constellations manually and come up with coordinates that would be best suitable for each constellation. The arrays of `ra_const` and `dec_const` are the manually found coordinates for naming the constellation.

Crowding up of the Hopping Plot

The plot to display the hopping sequence contains many information on the plot as such, names of prominent stars, constellation names, Messier ID's etc. These details aimed at helping ease out the user to make hops sometimes. However, because of this the plot may get crowded, especially around targets which have large number of Messier Objects in it's vicinity. It doesn't cause any problem for plotting or hopping as such but can be improved on to beautify the plots in future updates. This can be done by changing the `width` and `height` which for the same input kind of acts as zooming feature. Also, the opacity of all the Messier object markers other than the target could be reduced, which may prevent the crowding of Messiers.

Resolving of Names with Zooming

In the hop sequence plot the stars and Messier objects are named at coordinates such that names doesn't overlap with the markers they represent. This is done by naming these object at `x,y + value` where `value` is a integer number in map coordinates and `x,y` is the coordinate of the object. This suitably provides names to the objects on the map with given width and height, but while zooming in on the plot because of `value` the name does not lie at the same location, that is the distance between name and the object also increases with zooming.

this became more of a problems section rather than how to replicate in the last stages?

References

1878, A supplement to Sir John Herschel's "General catalogue of nebulae and clusters of stars." (The (Royal Irish) Academy, Dublin)

1998, NGC2000.0: Complete New General Catalog and Index Catalog (Sky Publishing Corporation and Cambridge University Press)

2012, http://astropixels.com/messier/messiercat.html

2018, https://www.kaggle.com/ecotner/named-stars

Blackman, E. ????, Naming the Stars, http://www.pas.rochester.edu/~blackman/ast104/naming.html

Carlos Segovia, J. 2016, Gaia TAP+ (astroquery.gaia) https://astroquery.readthedocs.io/en/latest/gaia/gaia.html

Davenhall, A., & Leggett, S. 1989, Catalogue of Constellation Boundary Data

Egret, D., & Fabricius, C. 1997, in ESA Special Publication, Vol. 402, Hipparcos - Venice '97, ed. R. M. Bonnet, E. Høg, P. L. Bernacca, L. Emiliani, A. Blaauw, C. Turon, J. Kovalevsky, L. Lindegren, H. Hassan, M. Bouffard, B. Strim, D. Heger, M. A. C. Perryman, & L. Woltjer, 31–34

ESA. ????, Gaia FAQs, https://www.cosmos.esa.int/web/gaia/faqs#BrightStar

ESA. 1997, VizieR Online Data Catalog: The Hipparcos and Tycho Catalogues

ESA. 2013, An Astronomy Data Query Language cookbook to accompany Gaia Data Release 1

Frommert, H., Kronberg, C., & McArthur, G. 1994-2013a, Messier Object Data https://www.messier.seds.org/dataMag.html

—. 1994-2013b, The Messier Catalogue https://www.messier.seds.org/

Gaia Collaboration. 2016, Astronomy and Astrophysics, 595, A1, doi: 10.1051/0004-6361/201629272

Ginsburg, A. 2020, Astroquery Documentation https://astroquery.readthedocs.io/en/latest/

Ginsburg, A. e. 2019, The Astronomical Journal

HEASARC. 2014, https://heasarc.gsfc.nasa.gov/W3Browse/star-catalog/bsc5p.html

HEASARC. 2015, https://heasarc.gsfc.nasa.gov/W3Browse/all/ngc2000.html

Hoffleit, D. 1964, New Haven, Conn.: Yale University Observatory, 1964, 3rd rev.ed., edited by Hoffleit, Dorrit

Hoffleit D., Warren Jr, W. 1991, Bright Star Catalogue, 5th Revised Ed. : V/50 (VizieR). https://cdsarc.unistra.fr/viz-bin/cat/V/50

Hog, E., Fabricius, C., Makarov, V. V., et al. 1998, in IAU Symposium, Vol. 179, New Horizons from Multi-Wavelength Sky Surveys, ed. B. J. McLean, D. A. Golombek, J. J. E. Hayes, & H. E. Payne, 395

Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, A&A, 355, L27

Hog, E., Fabricius, C., Makarov, V. V., et al. 2000, VizieR Online Data Catalog, I/259

IAU WGSN. 2018, Naming Stars https://www.iau.org/public/themes/naming_stars/

Kostjuk, N. 2002, HD-DM-GC-HR-HIP-Bayer-Flamsteed Cross Index

Ochsenbein, F., Bauer, P., & Marcout, J. 2000, A&AS, 143, 23, doi: 10.1051/aas:2000169

Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1997, A&A, 500, 501

Rob Garner. 2014, Explore - The Night Sky; Hubble's Messier Catalog. https://www.nasa.gov/content/goddard/hubble-s-messier-catalog

Sinnott, R. 1988, NGC 2000.0

Warmerdam, F., FRASER, B., BRUCE, C., et al. 2009, GitHub repository: https://github.com/OSGeo/proj, 4

Whitaker, J. 2011, Matplotlib Basemap Toolkit documentation, February

Zacharias, N., Monet, D. G., Levine, S. E., et al. 2004, 205, 48.15