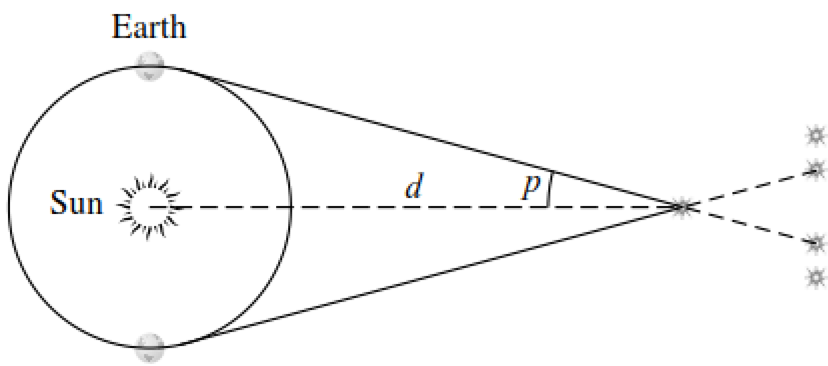
In this mini-paper, I'm going to explain different parts of TOPCAT and reproduce some famous diagrams.

For the sake of completeness, I will give some mathematical definitions and proofs which are essential for my examples.

1. Introduction to some astronomical concepts:

1.1. parallax:

One of the most important methods in astronomy for measuring stellar object's distance. It is based on the periodic motion of the stars which is caused by the Earth's motion around the Sun[1]:



As a result, the distance to a star can be calculated easily:

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

In this formula, "p" is measured in arcseconds and "d" in pc (parsecs).

Notice that the more distant an object is, the smaller parallax it has.

1.2. apparent and absolute magnitude:

Apparent magnitude (m) describes how bright an object appears to an observer on Earth. This quantity is another scale of measuring a star's brightness.

Contrary to public instinct, larger amounts of apparent magnitude mean the star is dimmer.

Absolute magnitude (M) is the apparent magnitude of a star if our detector is located exactly in 10 pc. The relationship between these two quantities can be written as:

$$M - m = 5\log\left(\frac{10}{d}\right) \xRightarrow{d=1/p} M - m = 5\log(10p)$$

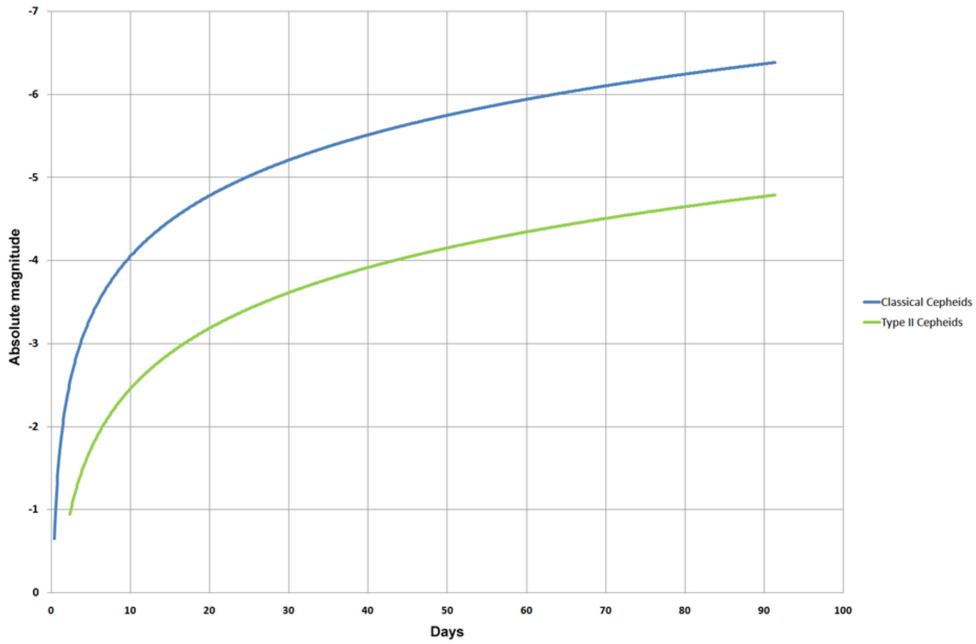
This expression is too crucial for us.

1.3. period-luminosity relation:

The relation between the pulsation period of variable stars and their luminosity. We are mainly interested in the classical Cepheid relation which is:

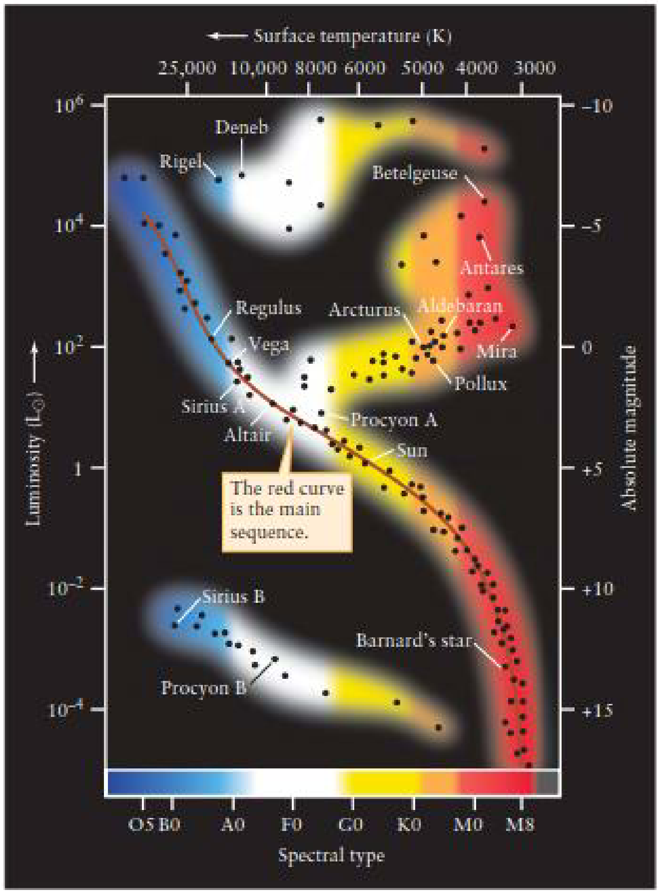
$$-2.78(\log(p)-1)-4$$

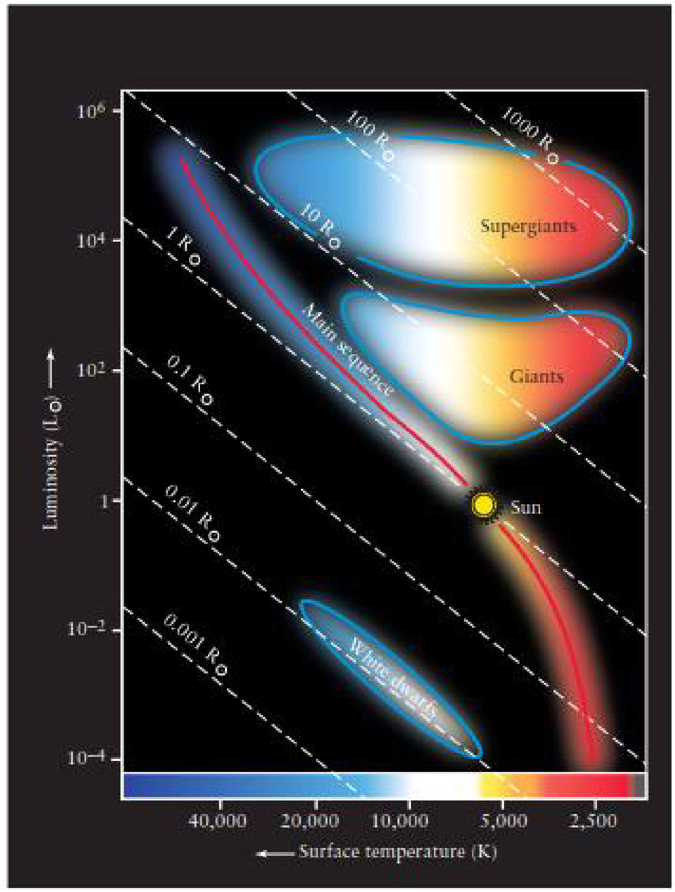
The diagram of these variables is:



1.4. HR diagrams:

In simple words, a graph of absolute magnitude versus spectral type (temperature scale) is called HR diagram [2]:





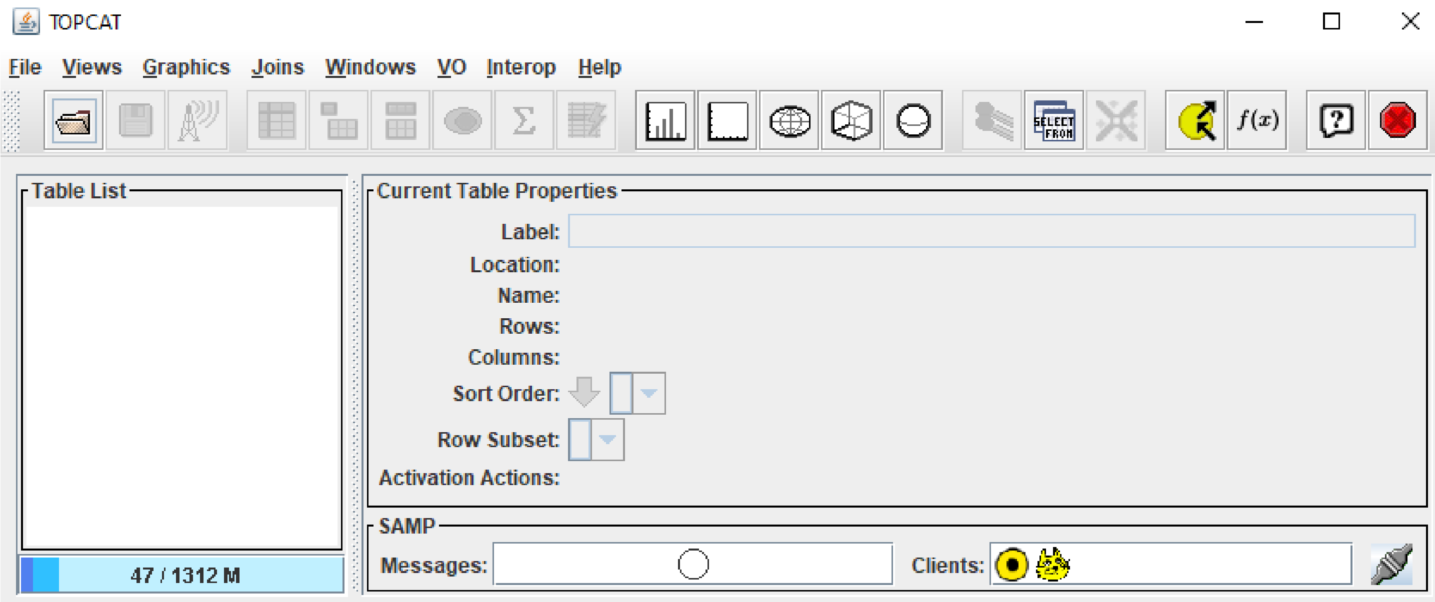
As you can see, people can infer the type a star according to its temperature and absolute magnitude. In addition, most stars are placed in the main sequence (Giants and Supergiants are located upper than main sequence since they are so bright!).

2. Introduction to TOPCAT:

If you want to access different data sources and graphing tools, TOPCAT is what you want!

2.1. some basic feature:

First of all, this is the main window of the software:

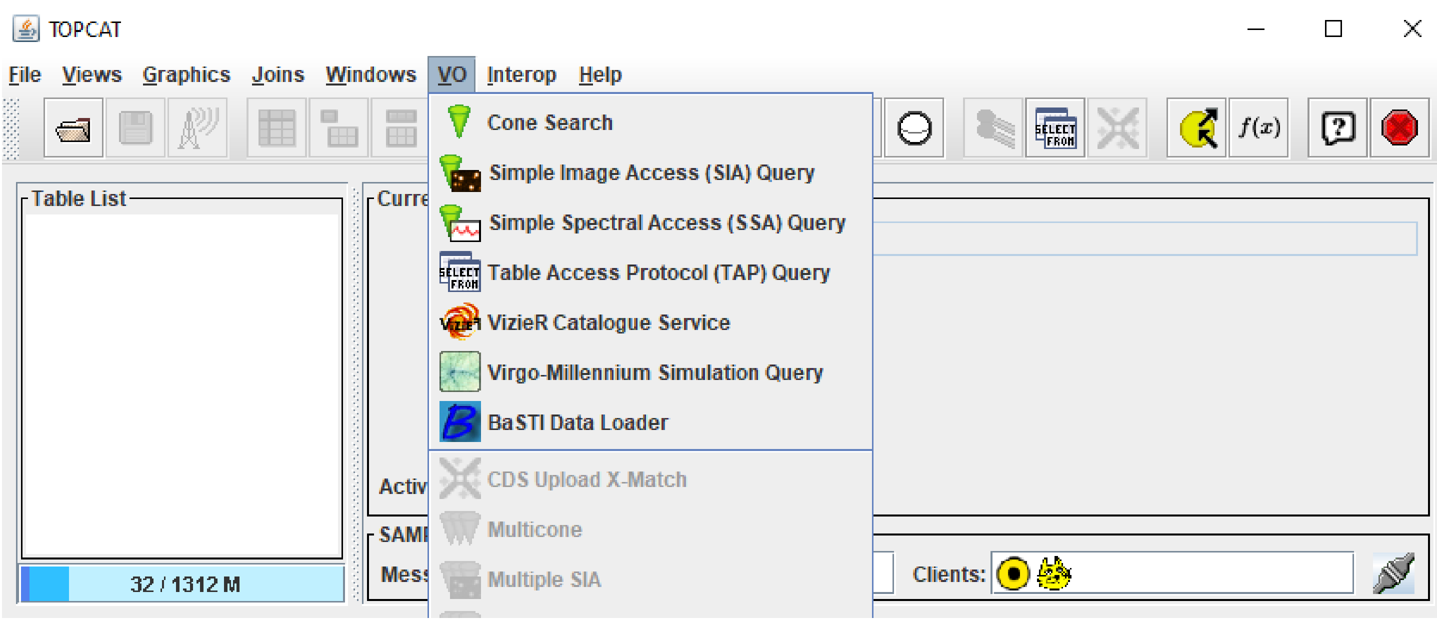


As you can see, there are several icons near the top which I will work with them soon.

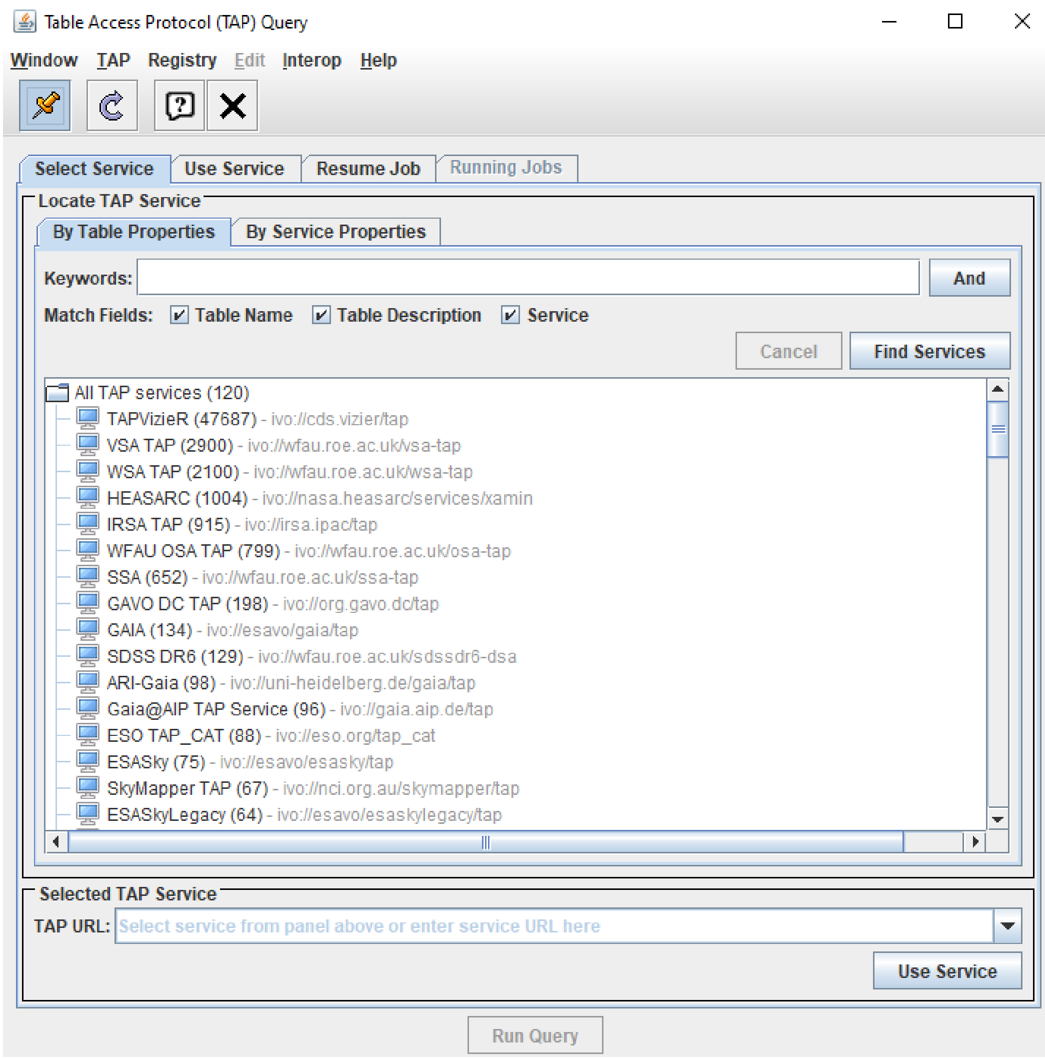
2.2. fetching data:

You can either download your target data set and import them from your local computer or use TOPCAT tools. My personal preference has been always the second option. So I will continue in this approach.

VO (Virtual Observatory) provides researchers with a useful portal to analyze or download different data sets. To begin, let's move our mouse to the icon and see what happens!:

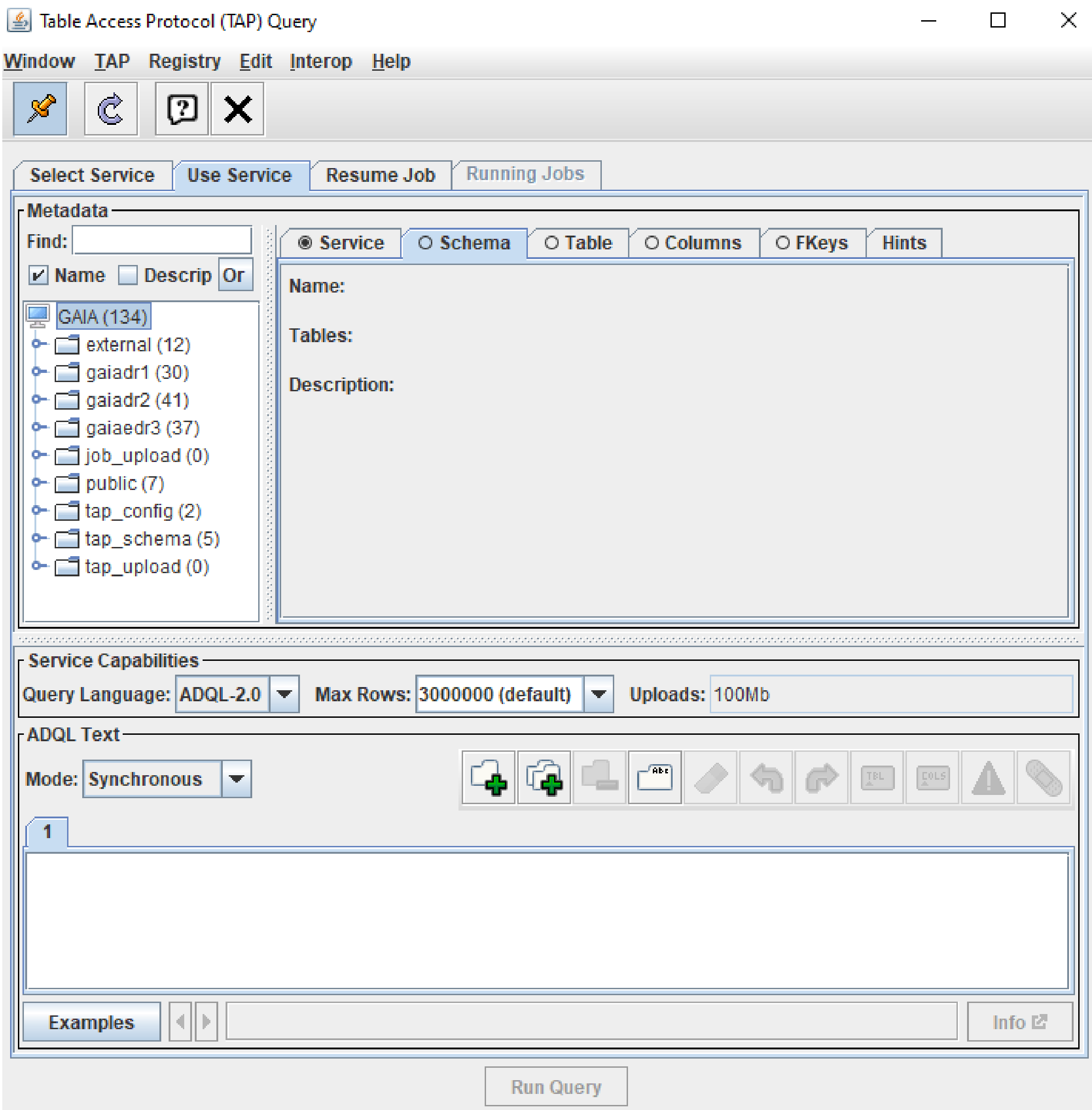


As it is obvious, you can fetch data through different methods. What I want to use here, is the "Table Access Protocol" which allows us to specify our desired features in data. After clicking on this icon, this window appears:

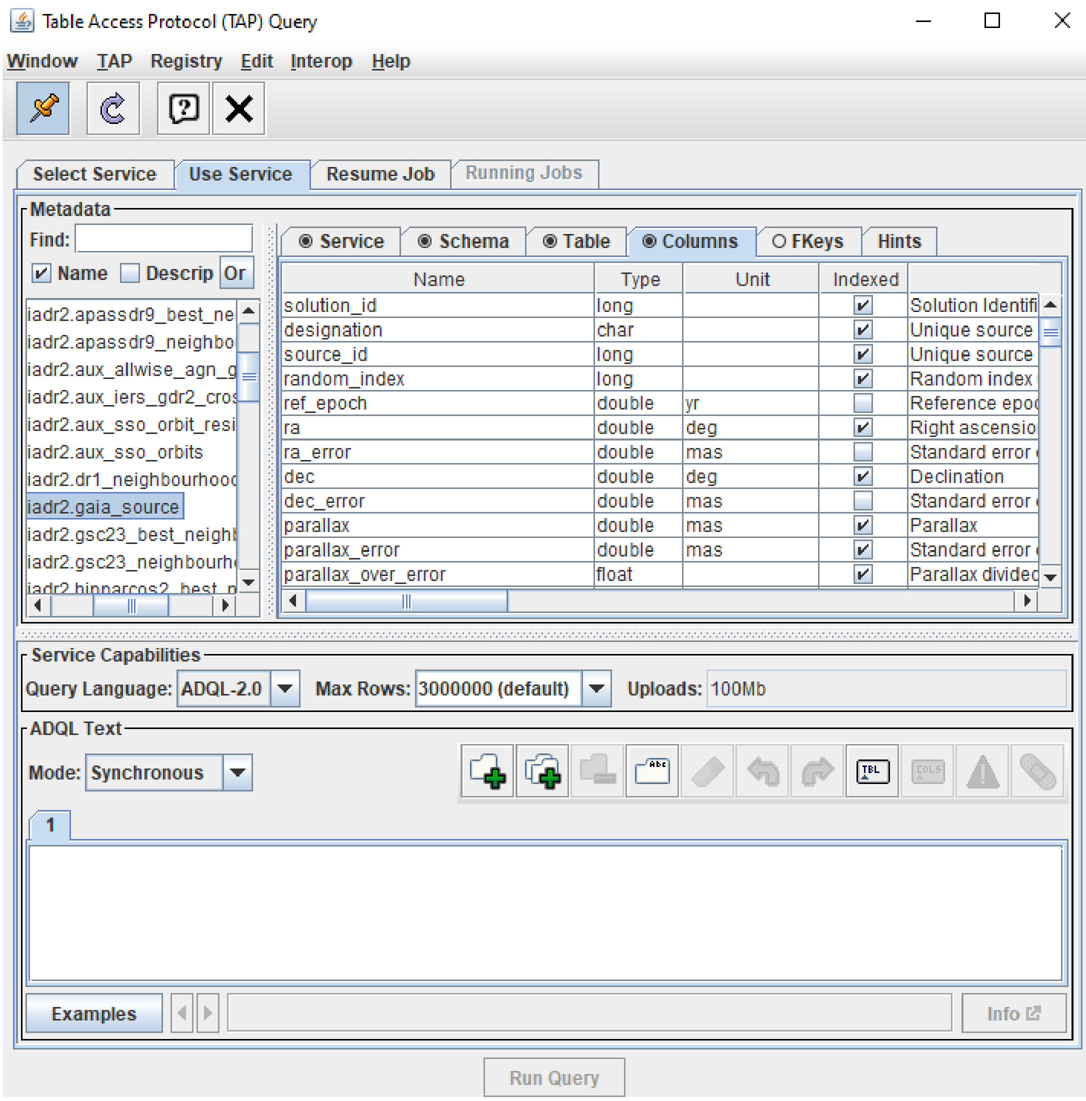


All services which is available in TOPCAT' server can be seen here. You can search the name of your data set and even your data's URL. As I mentioned earlier, I will choose GAIA (notice that there are a number of GAIA data from different centers. In this research, I chose ESA).

After double clicking on GAIA, the next stage is to have a look at the data:



You can see that several data set are listed here that each of them has numerous subsets (the number of them is written in front of them). The best choice for our purpose is gaiadr2 (which refers to the second data release). You can see the detailed information for any subset. For example, in gaiadr2, “gaia_source” presents these features (you have to click on columns icon):

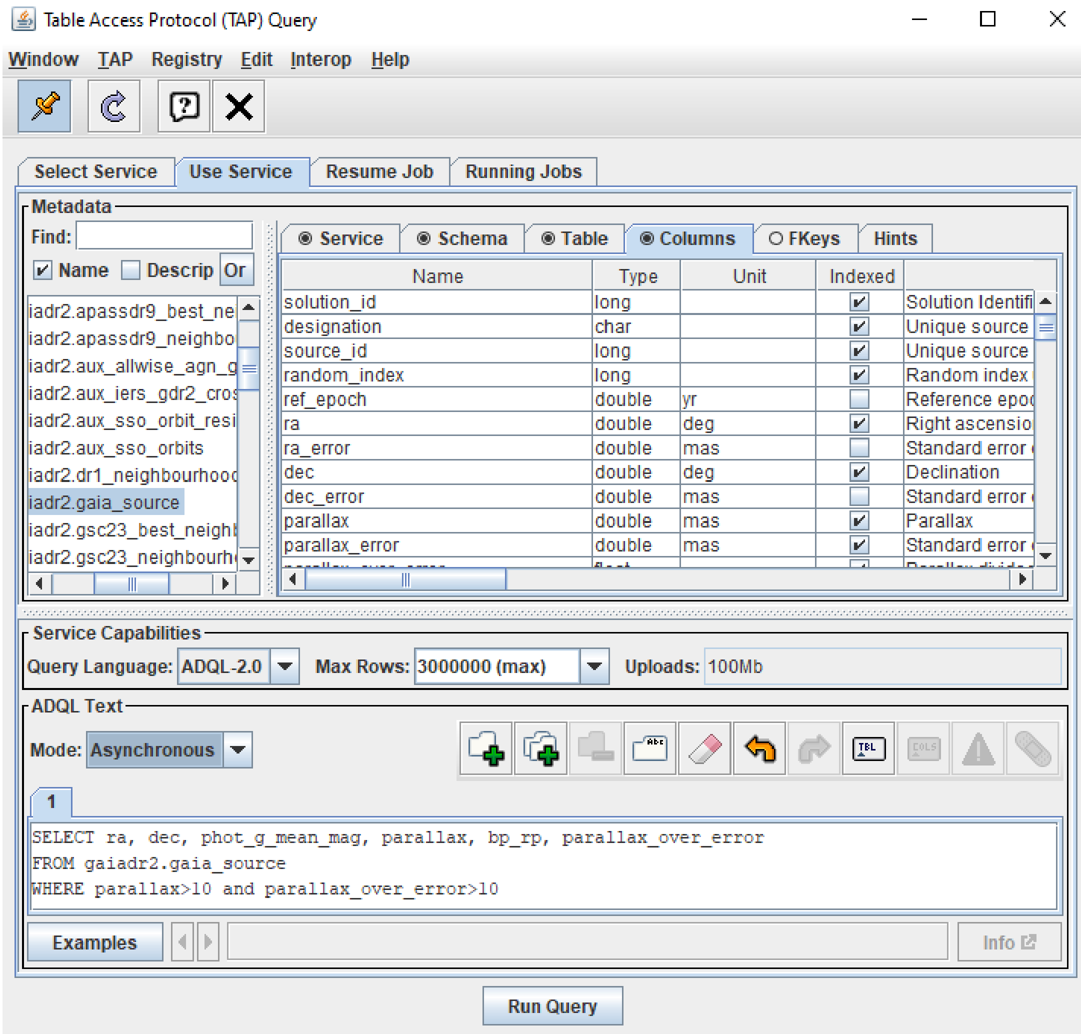


You can see the names, types, units and other additional information for every subset.

What is crucial here is that we have to search data set using ADQL language. In the following I will generate three distinct set and depict them in various plots:

2.2.1 stars which are close to us:

What is the most relevant quantity to distance?? Without a doubt, it is parallax!. So, to draw a sample of stars which are not too far, we must use their parallax measurements:



In the first line, I specified quantities which I want (they are abbreviated. For example, ra means right ascension). In the second line, I gave the target data set and in the last one, I wrote some conditions to fetch only near stars and low-error measurements. If you click on "Run Query", the program will provide you with your data set.

2.2.2 most bright stars:

In this case, I want to collect information of top 500000 stars which are brighter than other ones on sky. Again, which quantity is most relevant to this? According to previous chapter, it is apparent magnitude. So, I have to employ that these lines of codes:

```
SELECT TOP 500000 ra, dec, bp_rp, phot_g_mean_mag, parallax
FROM gaiadr2.gaia_source
WHERE parallax>0
ORDER BY phot_g_mean_mag ASC
```

In the last line, I told TOPCAT to arrange data according to their "phot_g_mean_mag" number (ASC stands for ascending).

Note that due to some errors, a number of parallaxes are negative! For this reason, I imposed positive amounts by the third line

2.2.3 Cepheids:

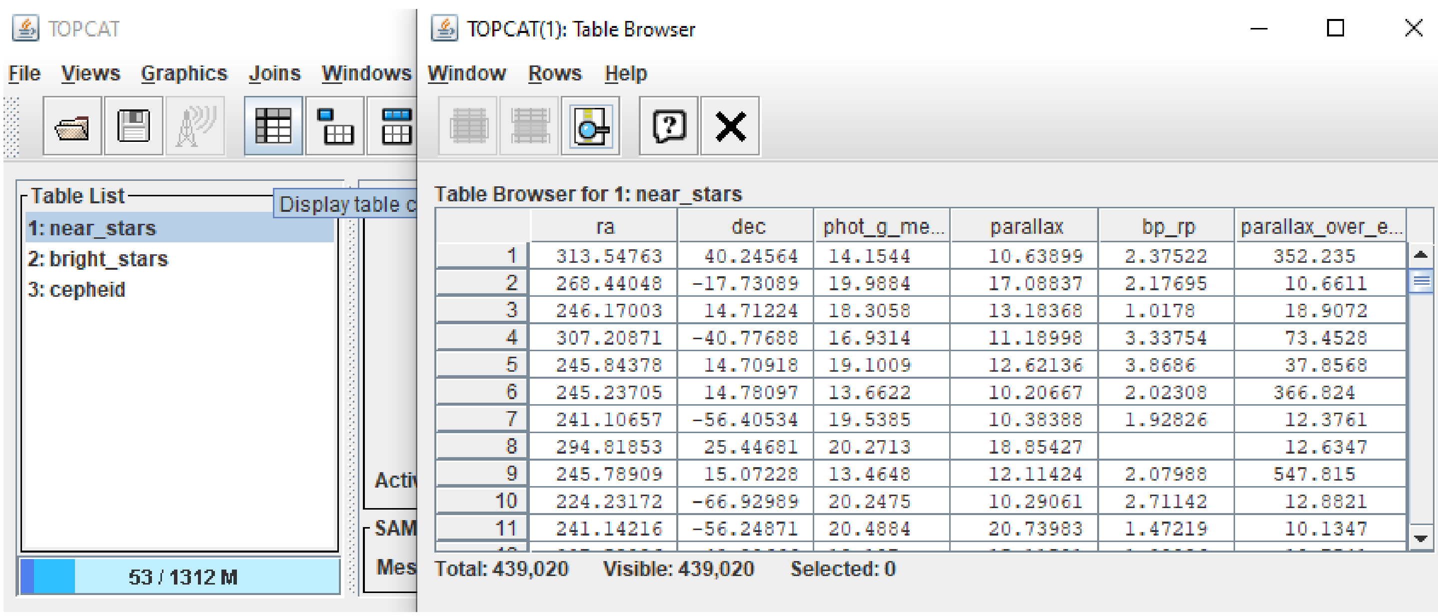
Cepheids play a crucial role in astronomy and GAIA has collected information about them too. Those data are accessible in "gaiadr2.vari_cepheid". If you take a look at this subset, you can see that it does not include the star's ra,dec,parallax, etc. What can we do about this? Luckily, TOPCAT can get data for a specified range of stars from two distinct subsets:

```
SELECT s.ra, s.dec, s.parallax, s.phot_g_mean_mag, c.pf
FROM gaiadr2.gaia_source AS s
JOIN gaiadr2.vari_cepheid AS c
USING (source_id)
WHERE parallax>0
```

As it is indicated, in the third line, using JOIN we linked two different data subsets. Most importantly, we said TOPCAT to identify stars by means of their "source_id" (which is common in each subset).

2.2.4 conclusion:

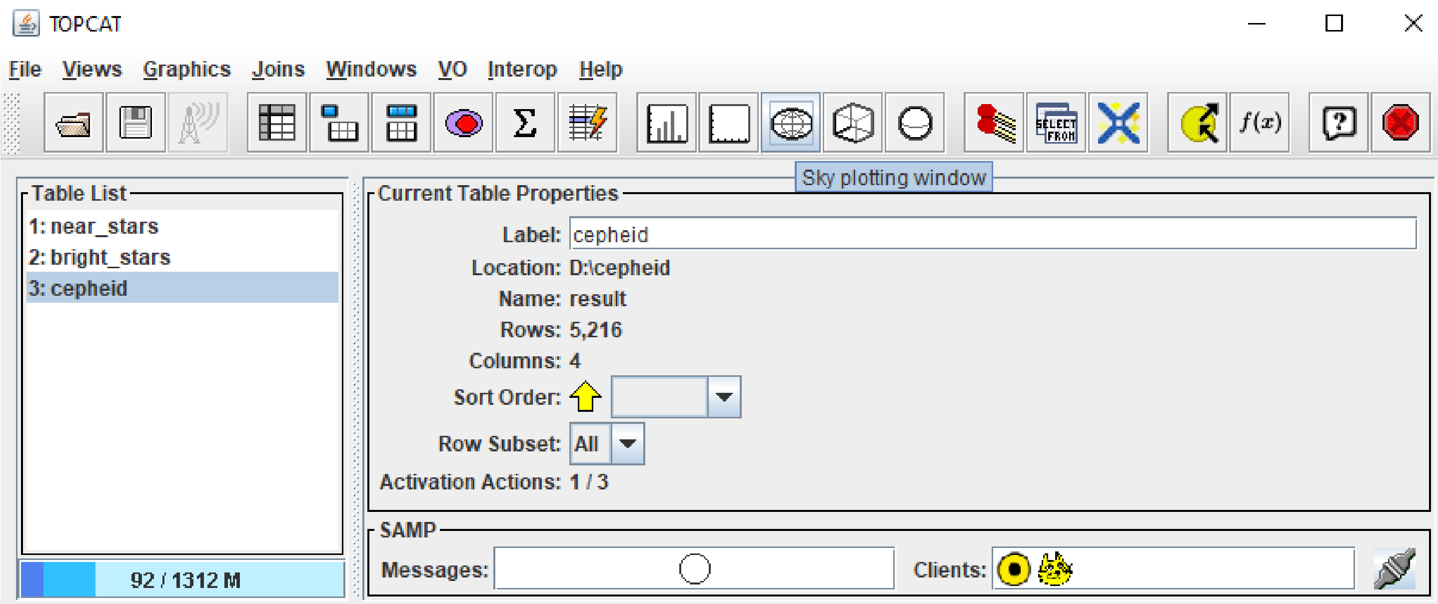
We gathered information from gaiadr2 and its subsets. Whenever you run a query, the gathered data will be added to your main window in TOPCAT and you can see its data using "table browser" icon:

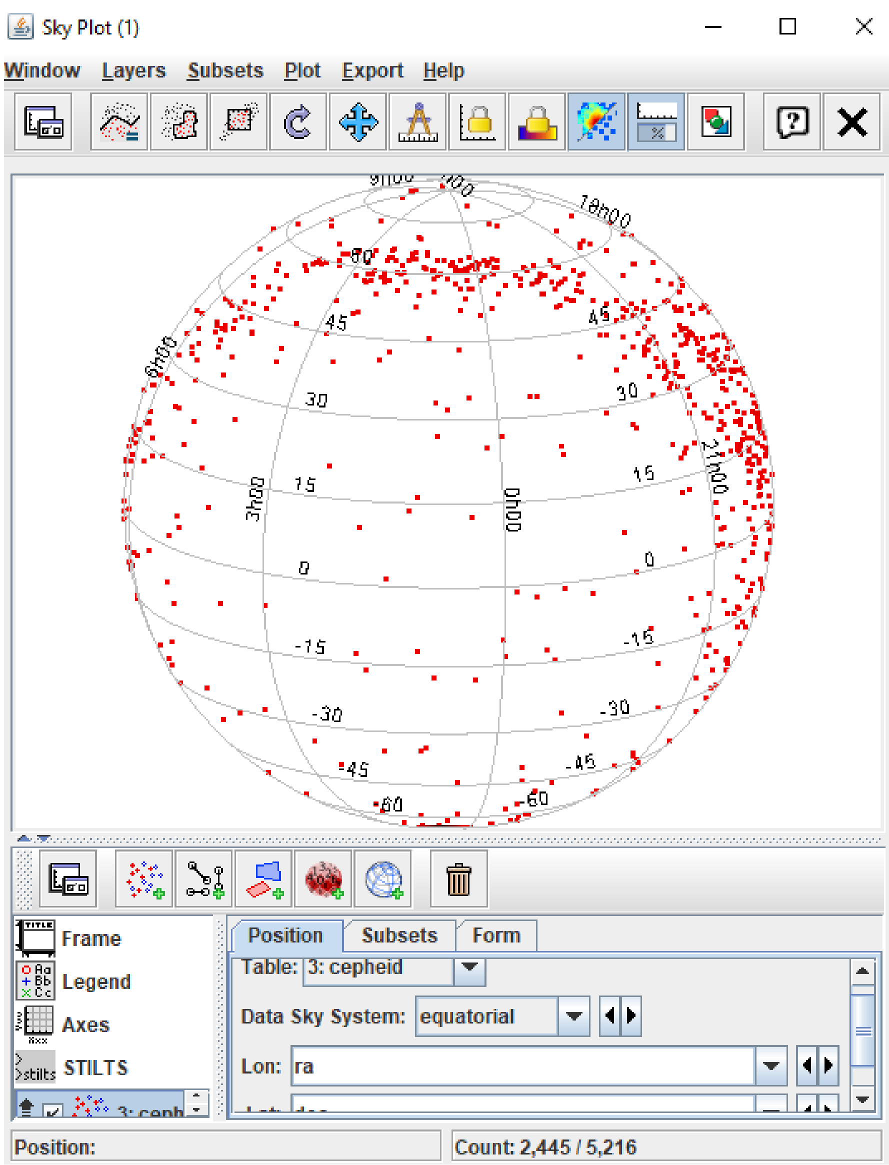


2.3. Visualizing data:

2.3.1 sky plot:

Sometimes we are interested to know where is our celestial object in the sky?? To achieve this, we can use the "sky ploy" module in TOPCAT. For example, consider cepheids, click on the related data and then plot it, we will have:

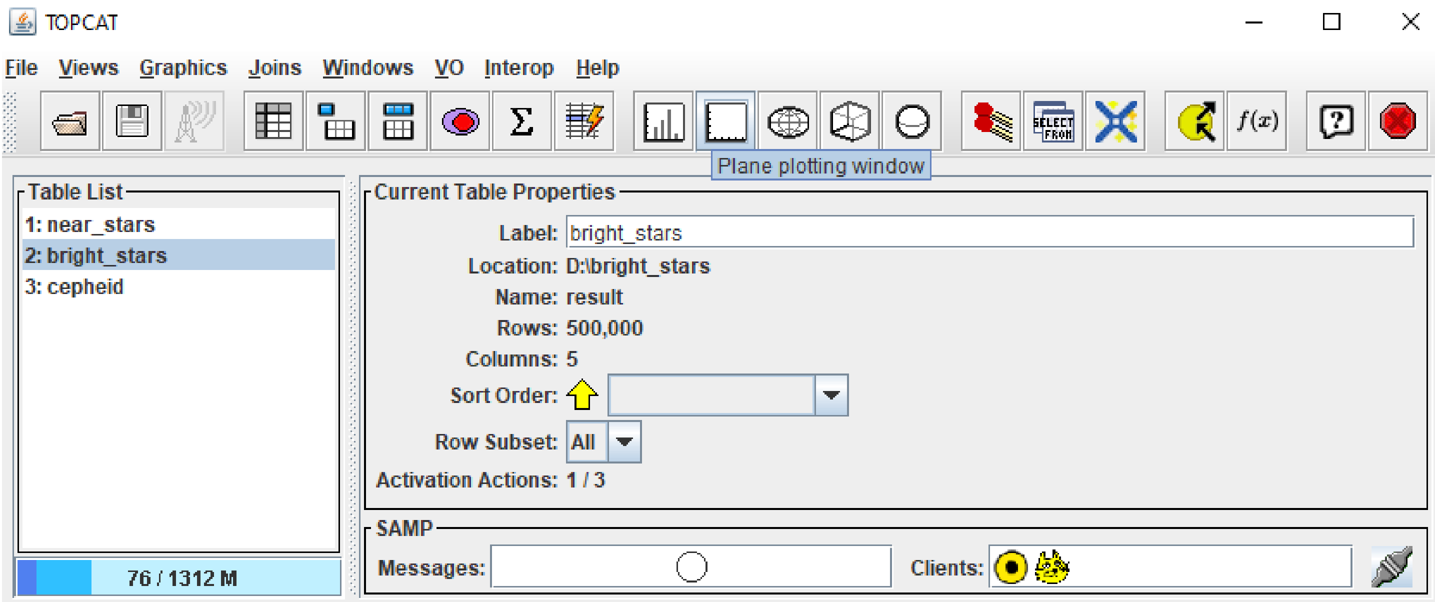


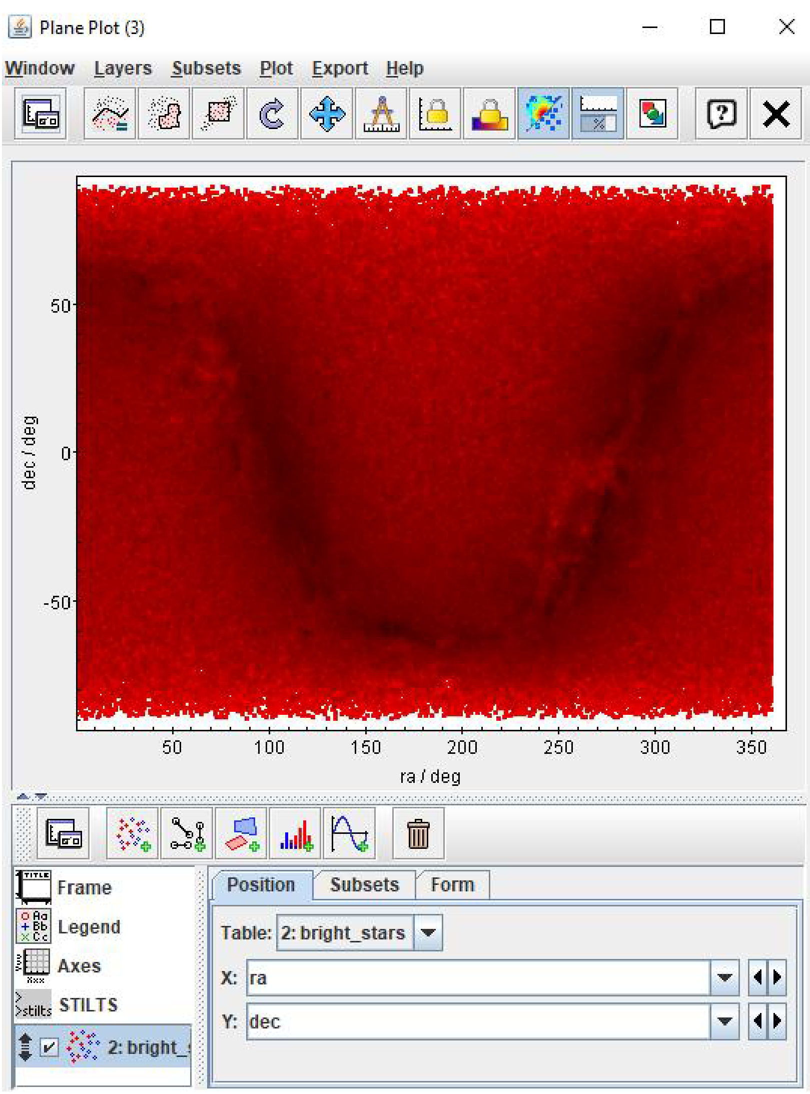


You can identify the location of each Cepheid star easily.

2.3.2. plane plot

Sometimes it is helpful to look at data in a plane plot. In this way, you will be able to understand your data better. For example, for bright stars, we have:

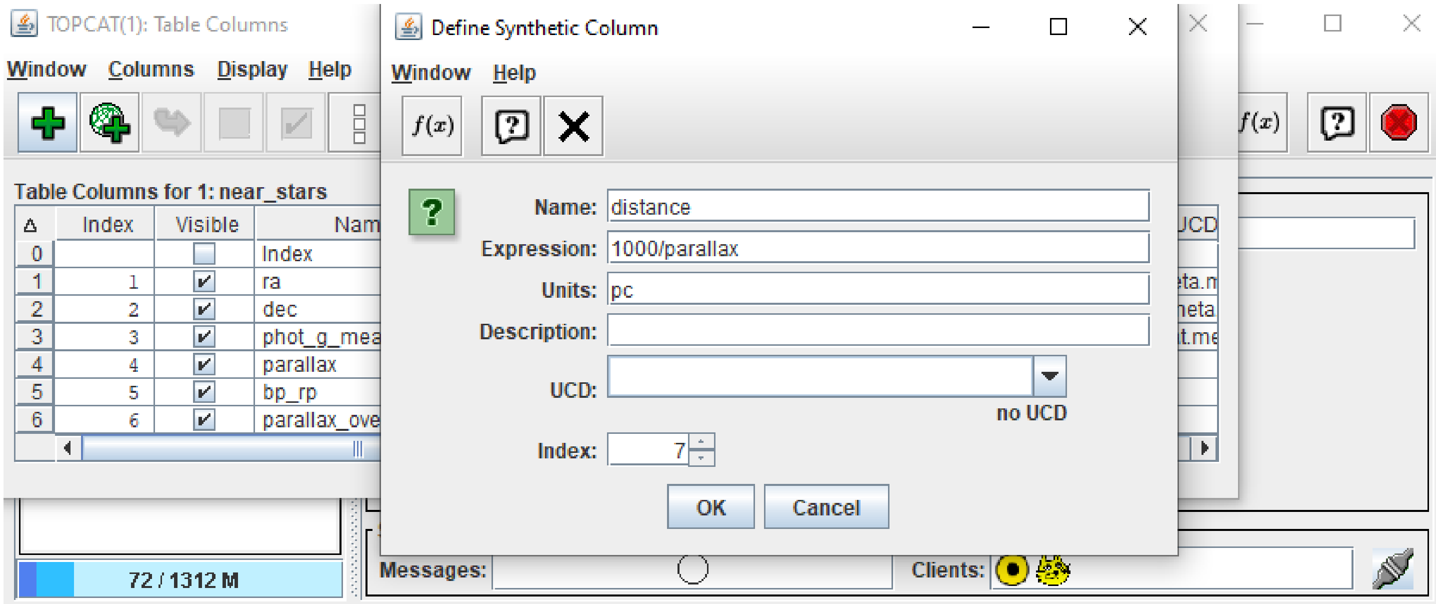




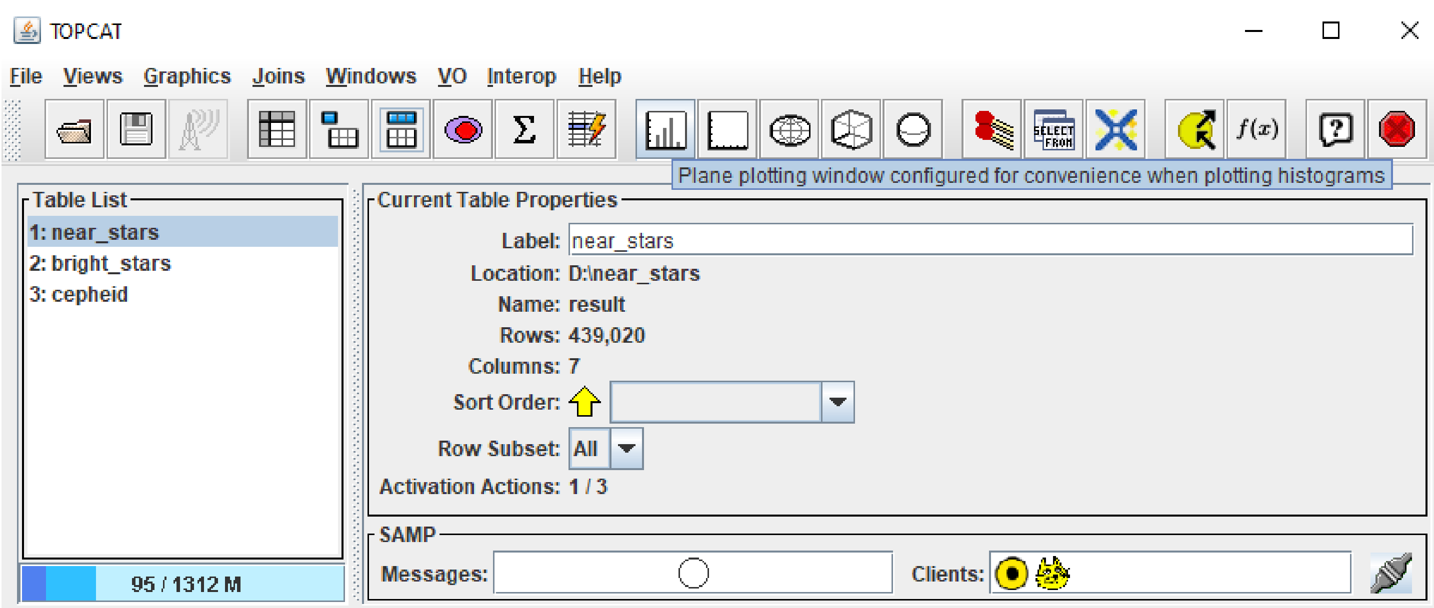
As you can see, the footprint of Milky Way is obvious in the middle (which means that a large percent of bright stars is in the Milky Way).

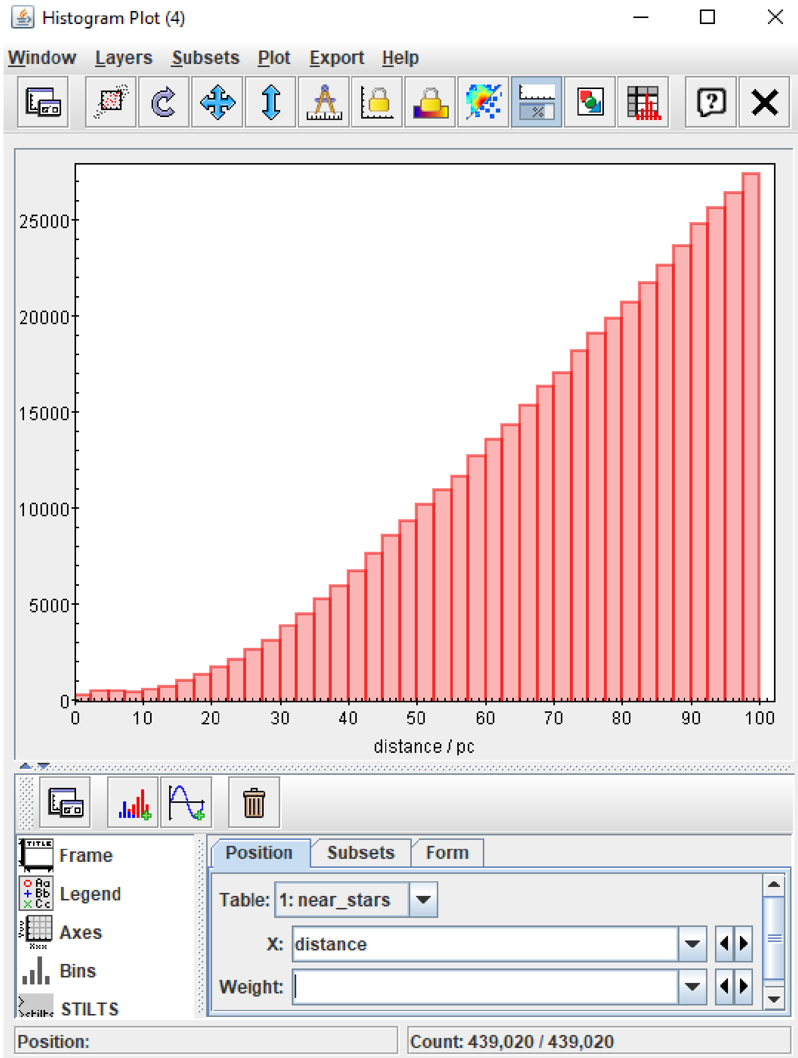
2.3.3 distances:

Distance is a key factor in many astrophysical calculations. But the amount of this quantity is not directly mentioned in data (we have to use parallax). So, I am going to calculate distance for all three data sets and add it to the table. For this, I click on "Display column metadata", then plus icon (on the top left side corner) and finally, write distance's name and expression in the boxes:



Notice that units of parallax measurements were mas, so we have to transform it. After this, the column of distances will be added to your data set. To visualize this, the best choice would be histogram:

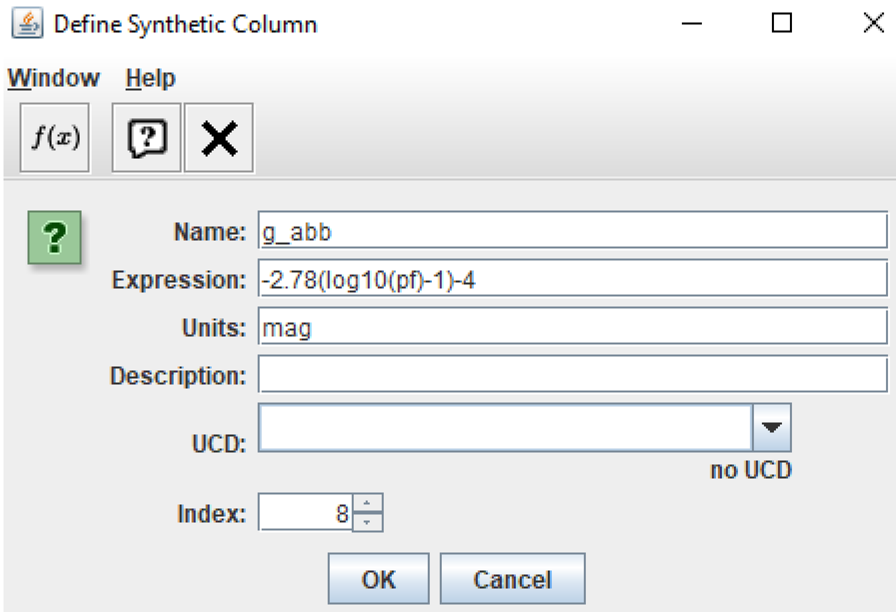


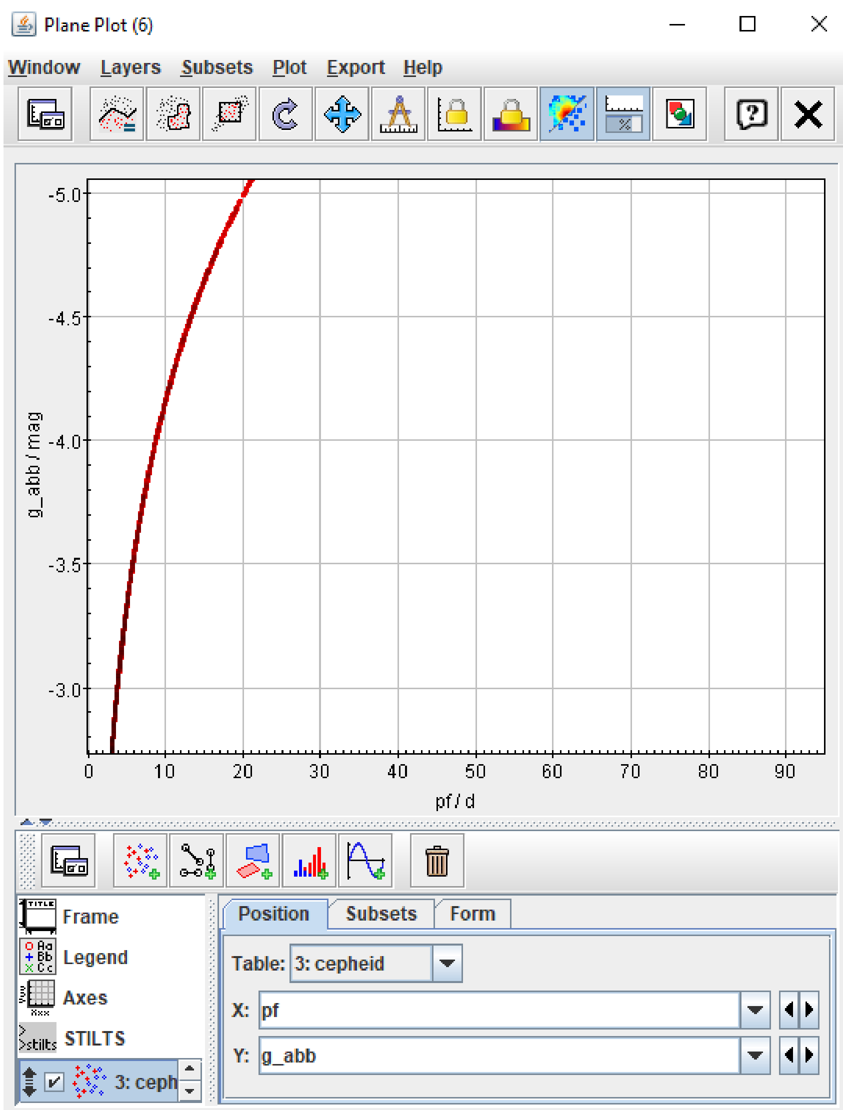


As it is clear, the distance ranges from approximately 1pc to 100pc (remember that these are nearest stars to Earth).

2.3.4. period-luminosity graph:

As described earlier, by adding another column to Cepheid data, we can plot this:

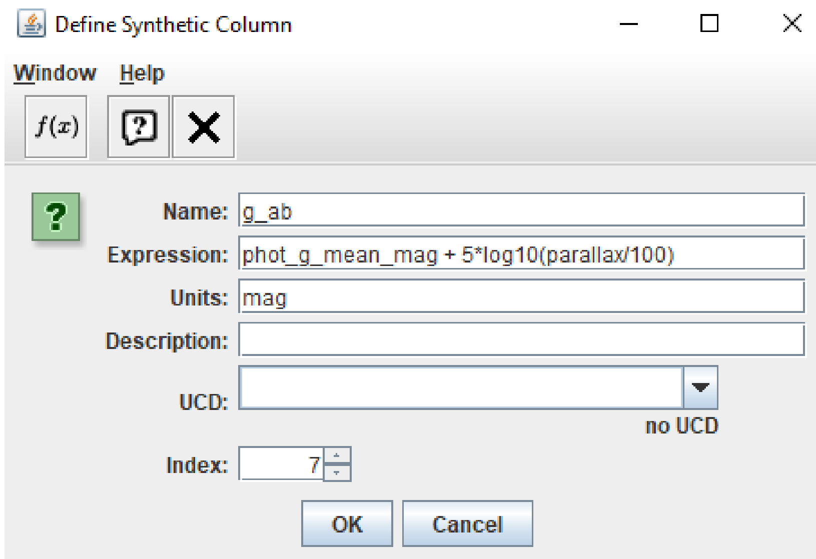


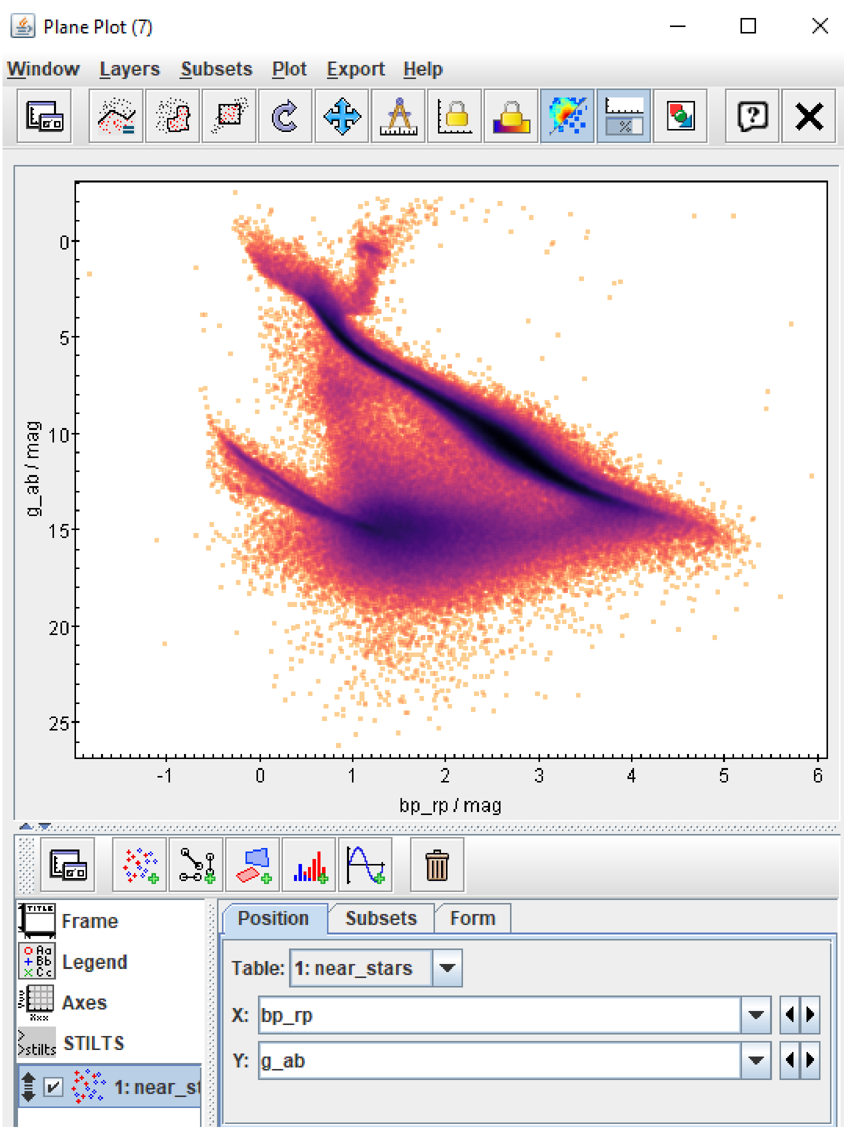


This graph implies that cepheids are classical.

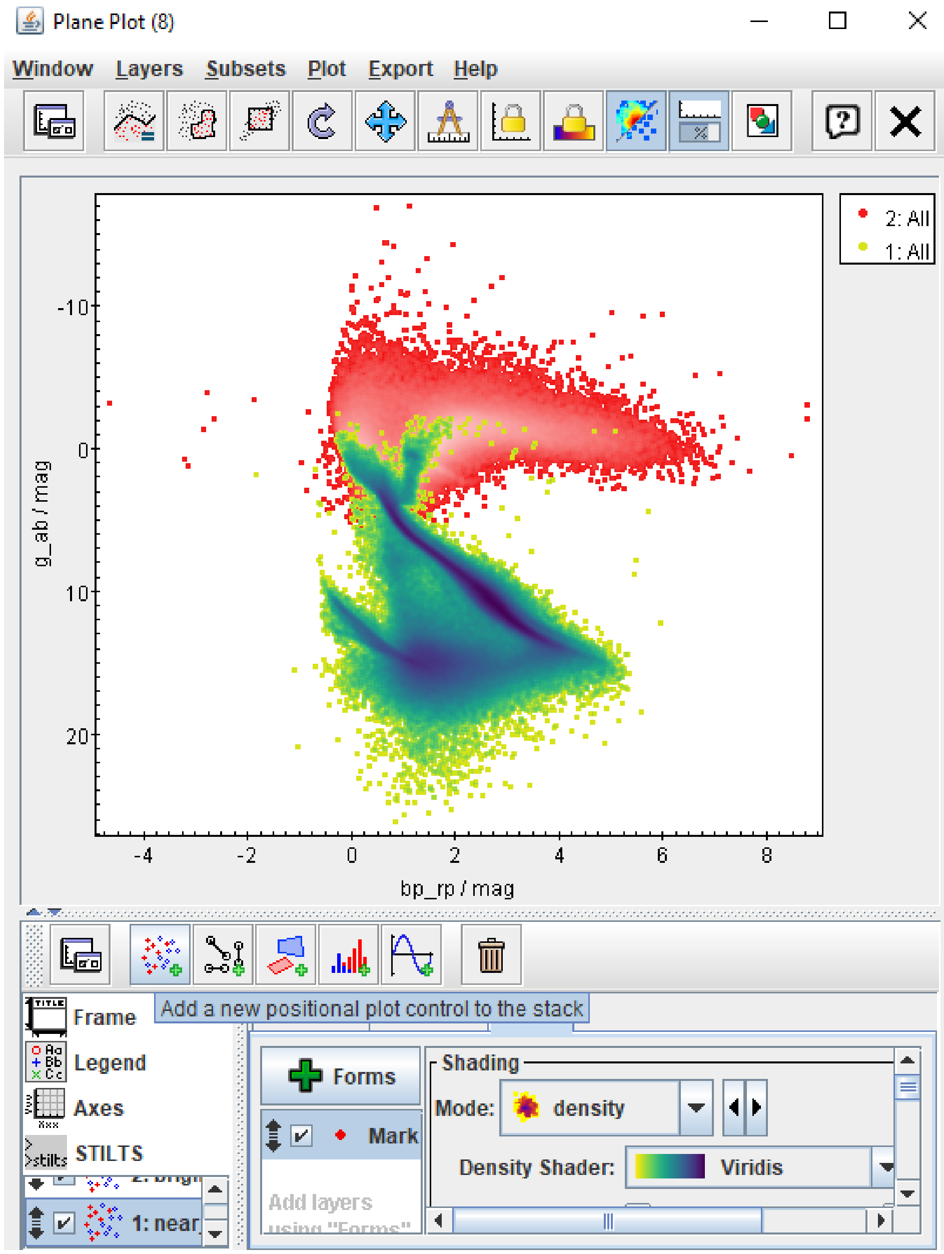
2.3.5. HR diagram:

One of the most famous diagrams in astronomy is HR diagram. For plotting this, first we must calculate the absolute magnitude of the near stars according to chapter one's formula:





But we can move further and combine the HR diagram of near stars and bright stars. To do this, we first absolute magnitude for bright stars, then combine them in a single graph:



As it is clear, bright stars are more likely to be supergiants.