Semester 1 2024

Astroinformatics I

Tutorial 3: Astronomical Software Intro

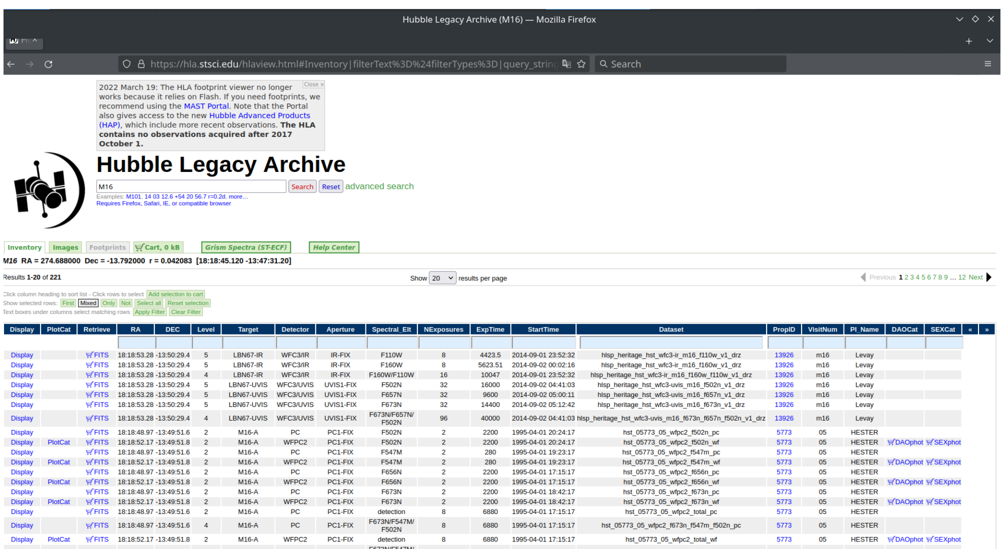
In this tutorial session, we will work with astronomical survey data which we process using SAO DS9 and Topcat.

1 Loading a Hubble image

We start with loading astronomical imaga files. We get our file from the Hubble legacy archive at the Space Telescope Science Institute (STScI), which operates the Hubble Space Telescopes and other space telescopes. The legacy archive is where all the older Hubble images are stored.

Go to `https://hla.stsci.edu/hlaview.html` and use the search field to search for M 16, the Eagle Nebula, by entering "M16" in the search field and pressing "Search".

You should see the following:



Now download the following files:

```
hst_05773_05_wfpc2_f502n_wf.fits
hst_05773_05_wfpc2_f656n_wf.fits
hst_05773_05_wfpc2_f673n_wf.fits
```

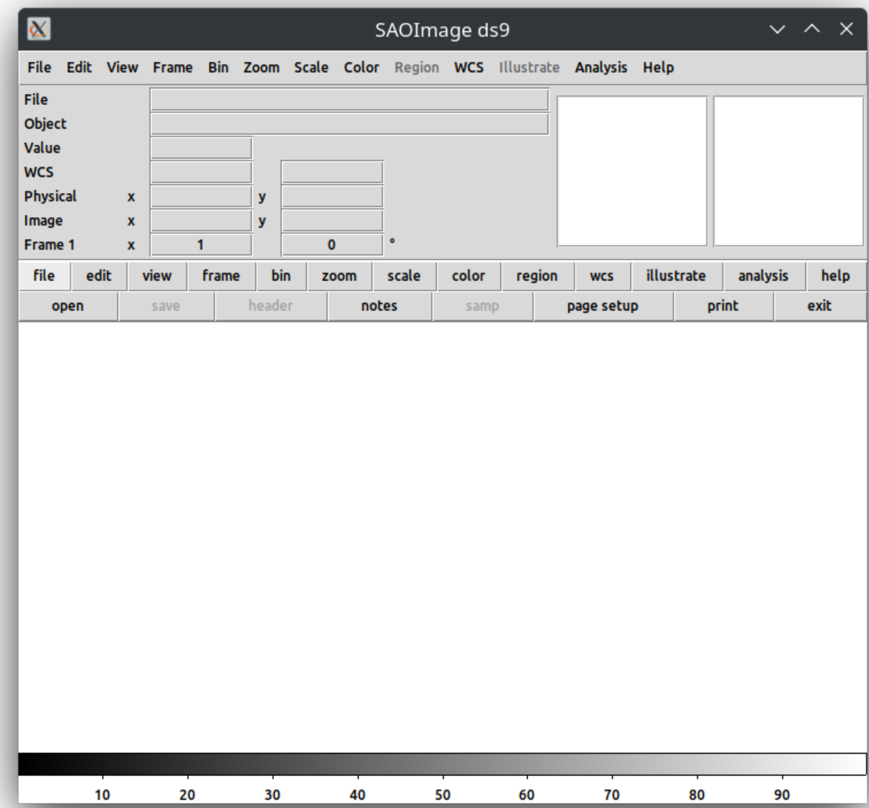
A short explanation on what these file names tell us: `hst` stands for "Hubble Space Telescope". `WFPC2` is the "Wide-Field and Planetary Camera 2" on that telescope, `wf` says that we are downloading the wide-field camera images. The `f502`, `f656` and `f673` denote different filters which have been placed before the camera for taking the respective images.

What we will do in this tutorial is to combine the three images into a color image.

It is important to highlight that the resulting image will be a false-color image as the filters used are not corresponding to red, green, blue.

2 SAOImage DS9

For processing the files we downloaded, we will use a software called SAOImage DS9. It was developed at the Smithsonian Astrophysical Observatory (SAO) and is available for download from `http://ds9.si.edu/`.
When you start DS9, you should see the following:

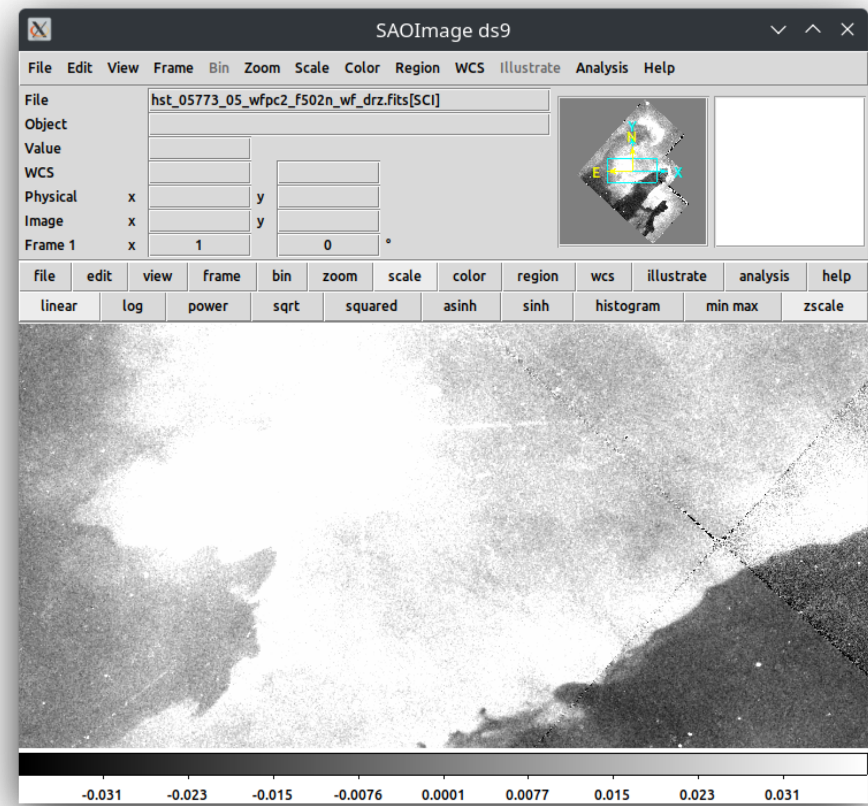


While one of the menus is where you would expect it, the other one is a bit hidden. It is a horizontal menu with two rows, starting with "file". Whenever you click on a field in the top row (the "main menu"), the bottom row will display a set of associated commands (the "secondary menu").

Below the secondary menu is the main image window, with a color bar (currently greyscale) below that and currently being empty.

We load the file `hst_05773_05_wfpc2_f502n_wf_drz.fit`.
The image should appear just as a black rectangle. This is the case as the brightness scale is not matching what our display can show:
Astronomical FITS images typically capture a very wide *dynamic range*, that is, an very wide range of different brightness values for each pixel. We have here 65536 different brightness values per pixel.
To be able to see the image, we need to find a better compressed brightness scale.

To experiment a bit to find a brightness scale that shows us the most information about the image, go to the main menu and choose "scale". From the secondary menu that will appear, choose "zscale".
Previously, we have been in "minmax" mode where the minimum pixel brightness was set to black, and the maximum pixel brightness to white. With "zscale", the greyscale colors are now mapped to values closer to the median pixel brightness. With that, the image should become visible in the main window.



The main window below only shows part of the image. In the overview window (second to right, on top) you can see the whole of the image. The little cyan frame in the overview window marks the part of the image that is visible in the main window. You can drag it around (left-click the mouse and drag) to explore other parts of the image.

With the help of the main window, try to place your cursor on a star. The detailed image (inset image top right) will shows the star's little disk, with the cursor on top, and the "Value" (top left) will give you the value of that particular pixel.

What is also shown is the sky position information for the current cursor position. "FK5" tells you that the sky coordinate system is defined using the reference stars of the *Fifth Fundamental Catalogue* (*Fundamentalkatalog 5*), which was published in 1988 by Astronomisches Recheninstitut Heidelberg, Germany. The coordinates themselves are those of the equatorial system, which is analogous to latitude and longitude on Earth: its longitude corresponds to the right ascension α (sometimes abbreviated to RA, ra, or R.A.), latitude corresponds to the declination δ (sometimes abbreviated to Dec, dec, or DE).

Both α and δ are given here in the standard sexagesimal notation: Right ascension is given in hours, minutes and seconds, declination is given in degrees northwards (positive sign) or southwards (negative sign) from the celestial equator, which itself is at $\delta = 0°$. You can change the notation from sexagesimal to degrees if you go to "WCS" in the top menu.

If you wonder how does DS9 know which sky coordinates to map to the pixel coordinates from the image: That (meta-)information comes from the FITS file header.

2.1 Meta-Information in the FITS Header

For working with astronomical images, it is extremely important to know the conditions under which they were taken. Such information is for example the filter, the exposure time, the time and location the image was taken.

FITS files store this so-called *meta-information* in a dedicated section which is called the *header*. In DS9, we can inspect the FITS header of the file we loaded by going to "file" in the main menu, and then choosing "header".

Try to find the following information: What filter was used? What is the exposure time? When was the image taken?

2.2 Creating a False-Color Image

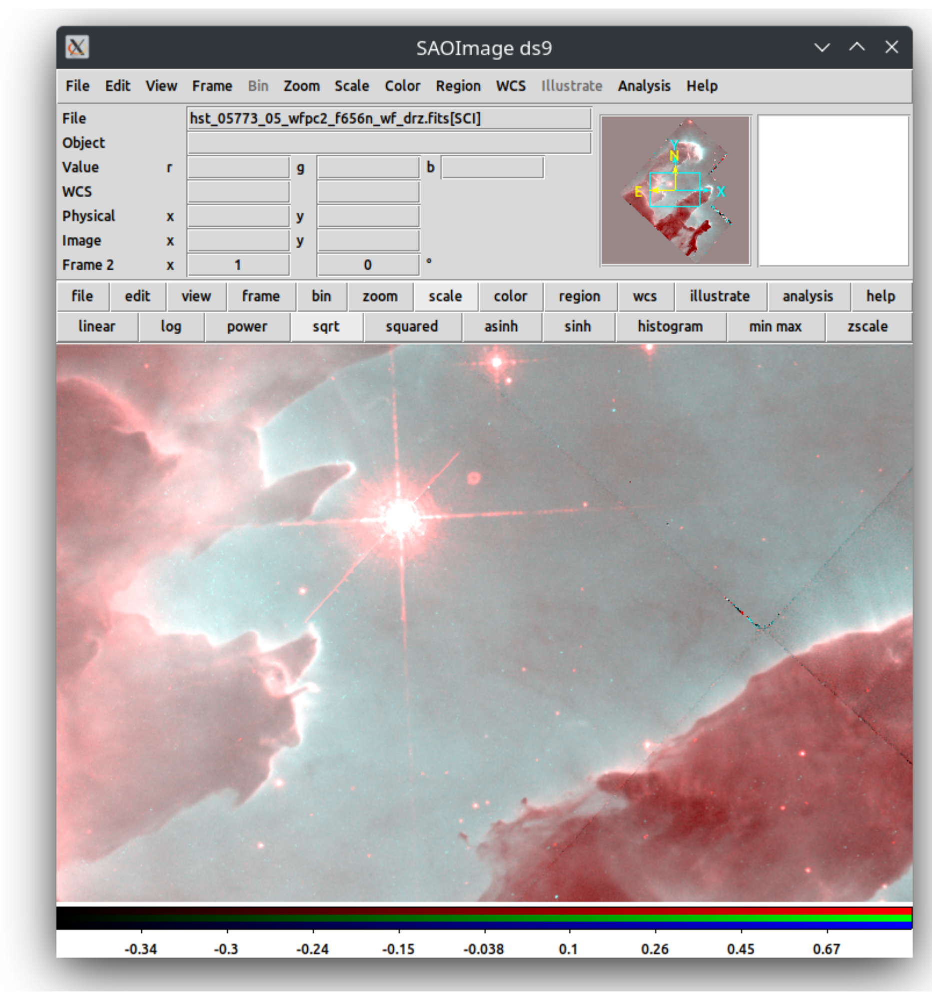
We are now going to combine partial images into an RGB color image. This will also help us understanding what the colorful astronomical images we commonly see really are.

We have seen in the lecture that astronomical cameras take greyscale images, each through a previously chosen filter. We now start creating the false-color image by combining our images in DS9. It is best to first restart DS9. Then create a color frame by clicking "frame" in the main menu, and then "rgb" in the secondary menu.

You will see a little menu popping up having the option "red" checked. While keeping this pop-up open, go to "file" in the main menu, "open" in the submenu.

For "red", open the file `hst_05773_05_wfpc2_f673n_wf_drz.fits`. Next, while still keeping the pop-up open, check "green" and load `hst_05773_05_wfpc2_f656n_wf_drz.fits`. Finally, check "blue" and load `hst_05773_05_wfpc2_f502_wf_drz.fits`. With this, the composite color image was already produced, but it is very dark. We can change this, still keeping the pop-up open: Click "Lock", "scale" in the pop-up. This will make your changes affect all three images (channels) equally.

We now set the scale to "sqrt" and the scale limits to "99%" from the "scale" top menu. The resulting image looks similar to the one you should be familiar with:



2.3 Searching SDSS

Nowadays, research in astronomy often combines information from present and past surveys, such as using coordinates from previous surveys to find objects in new survey data.

We will work with an image from the Sloan Digital Sky Survey (SDSS) data release 9 (DR9).

The original SDSS was the first large-scale digital survey, immediately making data from more than 900,000 galaxies available. SDSS DR9 is part of the third version of the survey, SDSS-III.

You can chose an image by going to the website `https://dr9.sdss.org/fields/` were you enter 20.0 both in the RA and Dec fields of "Search by Object Coordinates" and click "Submit". A result page should appear, where you click on the link "g-band FITS" to download the image taken through the SDSS *g* filter. Unzipping the file will give you a 12.4 MB FITS file named `frame-g-007923-5-0307.fits`. As previously done with the Hubble image, open it with DS9 and chose the zscale.

DS9 is able to show the positions of known stars using the positioning data from the FITS header. To display them, chose "analysis" from the main window's top menu. Chose the submenu "Catalogues", go to "Optical" and select "SDSS Release 9".

The image and the surrounding area should now be covered in small green circles. Each circle marks a star (or possibly other object) that is listed in the SDSS DR9 catalogue. In the catalog window that opens, you can see that in addition to an object's position (in RA and Dec), information on the object's brightness is available. SDSS provides magnitudes measured through the filter set *ugriz* which goes from the near-UV to the infrared.

The filter field in the catalog window allows you to filter on a variety of criteria. To filter the *g* band magnitude, try entering `$gmag < 17`, where `gmag` is the name of the column, to which you have to add a dollar sign $ to denote it is a variable. Then click on the "Filter" button near the bottom of the window. With doing so, you keep only those objects who are brighter than 17 mag in the *g* band.

Try out connecting several such conditions with `&&` for a logical AND, or `||` for logical OR, e.g. `$gmag < 17 || $rmag < 16`.

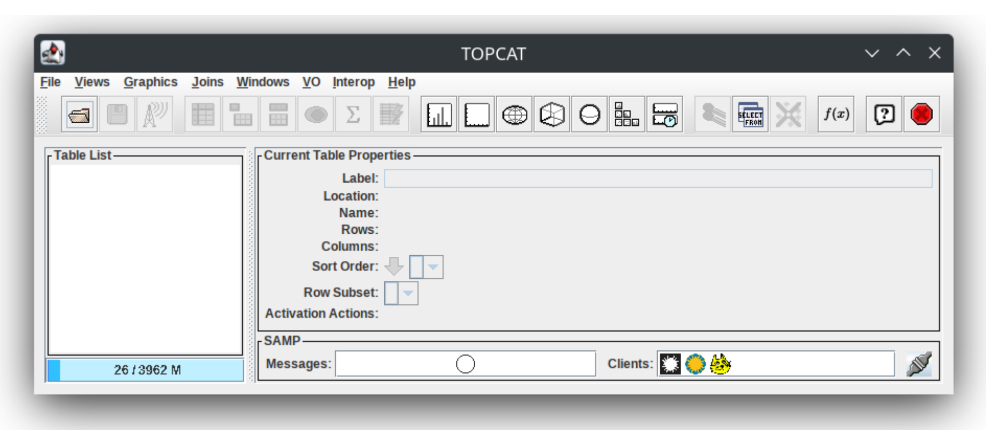
3 TOPCAT and Table Data

After we have looked at how to retrieve astronomical data (on the examples of the Hubble legacy achive and SDSS) and how to process astronomical images with SAO DS9, we will take a look at a software for dealing with astronomical tables. We will use here TOPCAT (Tool for Operations on Catalogues And Tables). TOPCAT covers a wide range of functionalities for table data, including creating plots, histograms, filtering data.

TOPCAT can be downloaded from `https://www.star.bris.ac.uk/~mbt/topcat/`. Further installation instructions can be found here: `https://www.star.bris.ac.uk/~mbt/topcat/#install`. TOPCAT is based on Java, so if you haven't a recent version of the Java Runtime Environment (JRE) installed, you will need to install it first.

Please make sure you can install and run topcat before you proceed.

When you start TOPCAT, you will see the base window from which other windows will open.



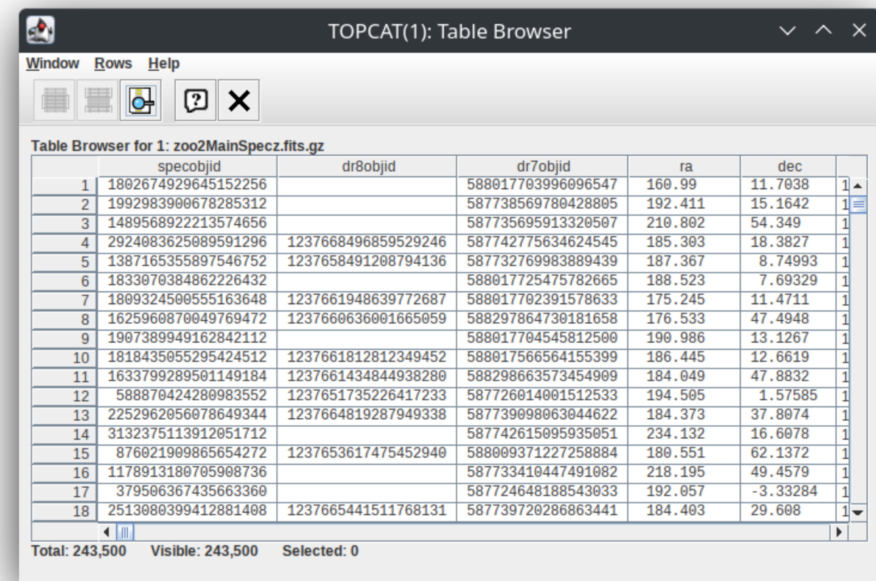
3.1 Galaxy Zoo data

To have a table to work with, we go to the Galaxy Zoo project at `https://data.galaxyzoo.org/`. From there, we will go to Galaxy Zoo 2 (`https://data.galaxyzoo.org/#section-7`) and download table 5 as a FITS file. This table contains the classifications of the 243,500 galaxies in the main sample with spectroscopic redshifts. This is the primary GZ2 data release, containing the largest number of galaxies and the most reliable morphologies.

3.2 Making a sky plot

Open the file you just downloaded with TOPCAT. The description of the columns can be found here: `https://data.galaxyzoo.org/data/gz2/zoo2MainSpecz.txt`.

A double click on the file name in the left part of the window opens the Table Browser:



	specobjid	dr8objid	dr7objid	ra	dec
1	1802674929645152256		588017703996096547	160.99	11.7038
2	1992983900678285312		587738569780428805	192.411	15.1642
3	1489568922213574656		587735695913320507	210.802	54.349
4	2924083625089591296	1237668496859529246	587742775634624545	185.303	18.3827
5	1387165355897546752	1237658491208794136	587732769983889439	187.367	8.74993
6	1833070384862226432		588017725475782665	188.523	7.69329
7	1809324500555163648	1237661948639772687	588017702391578633	175.245	11.4711
8	1625960870049769472	1237660636001665059	588297864730181658	176.533	47.4948
9	1907389949162842112		588017704545812500	190.986	13.1267
10	1818435055295424512	1237661812812349452	588017566564155399	186.445	12.6619
11	1633799289501149184	1237661434844938280	588298663573454909	184.049	47.8832
12	588870424280983552	1237651735226417233	587726014001512533	194.505	1.57585
13	2252962056078649344	1237664819287949338	587739098063044622	184.373	37.8074
14	3132375113912051712		587742615095935051	234.132	16.6078
15	876021909865654272	1237653617475452940	588009371227258884	180.551	62.1372
16	1178913180705908736		587733410447491082	218.195	49.4579
17	379506367435663360		587724648188543033	192.057	-3.33284
18	2513080399412881408	1237665441511768131	587739720286863441	184.403	29.608

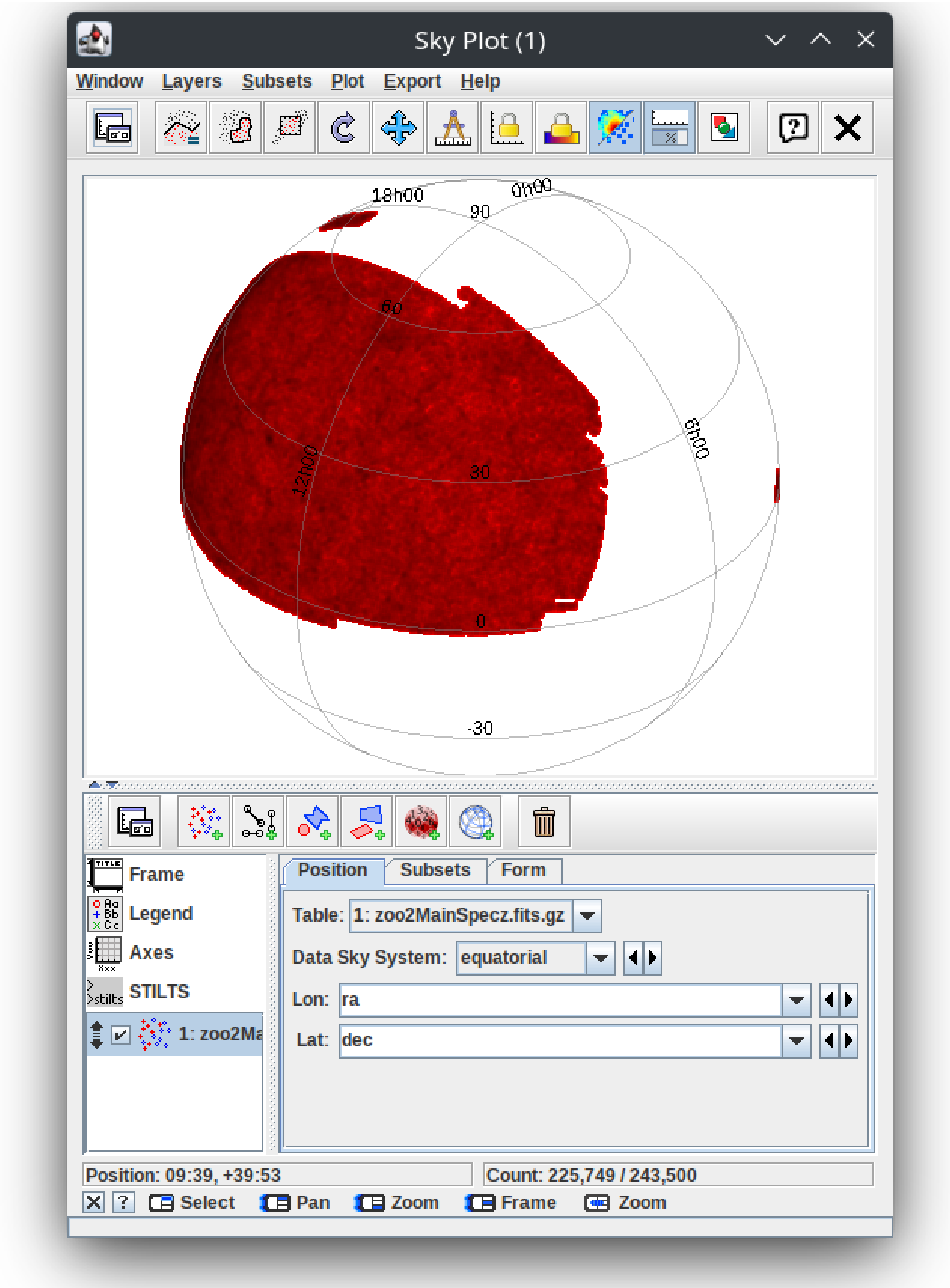
TOPCAT provides several possibilities of viewing your data.

Plotting the positions of astronomical objects is one of the most common. When you are given a catalog of astronomical objects you might be interested in their distribution on the sky, which also gives you an idea about the coverage of the survey: Is it an all-sky survey, or is it a targeted survey that only covers a small region?

Here, we want to plot where on the celestial sphere the galaxies from the Galaxy Zoo DR2 list are.

Among the icons in the main window, locate the "sky plot" icon. Alternatively, you can chose in the menu Graphis → Sky Plot.

TOPCAT automatically identifies the RA and Dec columns, and plots the object's positions in the sky.
You can rotate the sphere in the plot by click and hold with the left mouse button.



The Sky Plot window comes with its own options for customisation. For instance, you can easily vary at least some aspects of the visualisation. If you click on "Axes", the options field directly to the right will change. In the tab "Projection" you can change the projection from the default "sin" (which shows you a projection of the celestial sphere) to the world-map like "aitoff" or to the rectilinear coordinate plot "car". On the tab "grid", you can configure the density of the grid coordinates less or more dense by using the slider "Grid Crowding".

3.3 Virtual Observatory (VO) services

With astronomical data bases getting getting larger and larger, downloading all the data before selecting the data we really want to use for our science becomes more and more cumbersome. The alternative, which is widely available noway, is searching online data bases for exactly the data you need, and then only download these data.

For table data, this search is commonly done by using the *table access protocol* (TAP) which uses the *Astronomical Database Query Language* (ADQL), which is very similar to SQL.

We will see here how we can do so directly from within TOPCAT.

In the top menu, go to the tab "VO" (which stands for "Virtual Observatory"). Chose the TAP service, and then select the ESA database for the Gaia satellite, which is simply listed as GAIA in the TAP window. Double-click on GAIA and you will get additional information.

In order to access data from a VO table, we need to execute a query, which asks the database to return a specific set of data according to what we want to select. Such queries are written in ADQL. We enter our queries into the bottom entry field in the TAP window ("ADQL text").

As an example, we want to get the RA and Dec cordinates of objects close to the position of the Andromeda galaxy, and those objects should come from the `gaiadr2.gaia` table.

This is the query:

```
SELECT ra, dec
FROM gaiadr2.gaia_source
WHERE 1=CONTAINS(POINT('ICRS',ra,dec),
CIRCLE('ICRS',10.684708,41.268750,
3.2))
```

We enter this and then press the button "Run Query" to execute the query. In case there was a syntax error, it will not be executed but an error message will show up. During the download, there will be a temporary download window with an animated progress bar.

The plot we see is not an image (i.e.: photography) of the Andromeda galaxy. It is a density plot, a diagram showing the density of individual point sources (i.e.: stars) identified by Gaia DR2 to trace Andromeda by individual stars identified within.

3.4 More on TOPCAT

This was an introduction to TOPCAT. You can find more tutorials online, e.g. on the TOPCAT website: `http://www.star.bristol.ac.uk/~mbt/topcat/#further`.