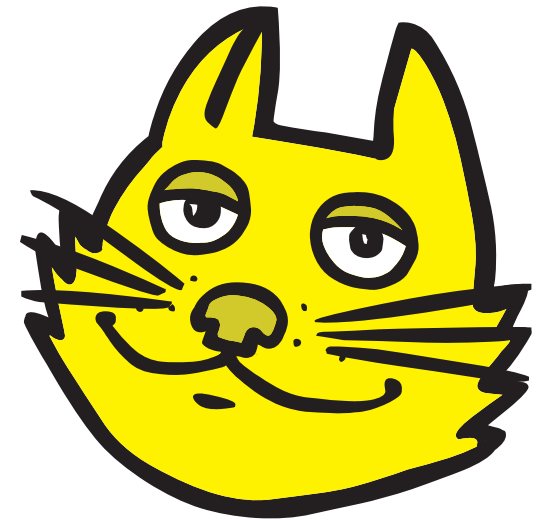


# Introduction to TOPCAT

Mark Taylor (University of Bristol)

*Shristi Astronomy from Archival Data*  
Online seminar

22 August 2020



University of  
**BRISTOL**

\$Id: topcat.tex,v 1.5 2020/08/21 13:31:00 mbt Exp \$

# Outline

## TOPCAT

- What is it?
- What can it do?

## Demo

- Pleiades in Gaia DR2 and 2MASS
- Hyades in 3D phase space using TAP

## Q+A

## Hands on

- Gaia DR2

# Overview

*TOPCAT = Tool for OPerations on Catalogues And Tables*

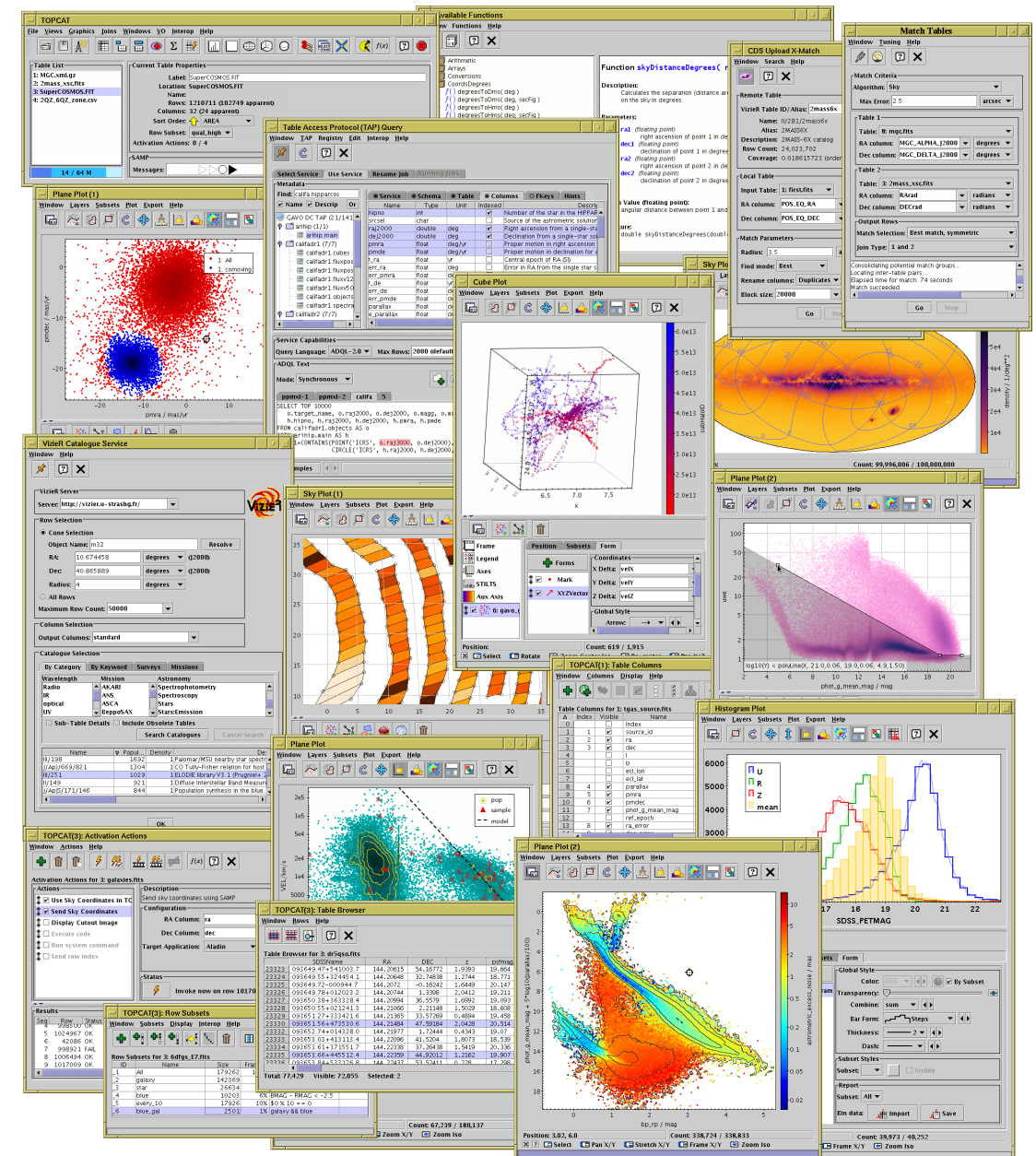
*“Does what you want with tables”*

Suitable for:

- Interactive exploration
- Quick look at unfamiliar data
- In-depth analysis

## Overall aim:

- Makes table manipulation easy, so users can concentrate on doing science



# Characteristics

## Aims:

- User-friendly
  - ▷ Easy to install and run (pure Java — one download file, no library issues)
  - ▷ Easy to get started
  - ▷ Simple things fairly obvious
  - ▷ Complicated things at least well-documented
  - ... this does get harder as more functionality is added
- High Performance
  - ▷ Most things are fast
  - ▷ Handles fairly large tables: millions of rows, hundreds of columns easily (can be much more)
  - ▷ ... even on modest hardware
- Does the things that astronomers need
  - ▷ Development is led by community input (mailing list, personal emails, tutorials, feature requests, bug reports...)
  - ▷ Feedback always welcome!

# Capabilities

## It can do:

- Read/write tables in various formats (FITS, VOTable, CSV, ...)
- View data
- View metadata
- Calculations and simple statistics (expression language)
- Visualisation (many options, interactive)
- Make/combine/display row selections in various ways (linked views)
- Crossmatching (many options)
- Access external data services (VO and others)
- Talk to other astro tools (SAMP)
- Trigger some event when a row is selected

## It can't do:

- Images, spectra (it's only for tables)
- Scripting (but see STILTS)
- *Very* large tables (but see STILTS)
- Every ASCII-like format known to man
- Write your papers for you

# Table Data and Metadata

TOPCAT(1): Table Browser

Window Subsets Help

Table Browser for 1: dr5qso.fits

	SDSSName	RA	DEC	z	psfmag_u	psfmagerr
21412	092322.64+020135.5	140.84436	2.02655			
21413	092322.67+282526.5	140.84449	28.42405			
21414	092322.86+033821.5	140.84526	3.63933			
21415	092323.01+461835.3	140.84588	46.30982			
21416	092323.65+580256.0	140.84855	58.0489			
21417	092323.92+610154.0	140.84969	61.03167			
21418	092324.25+382812.8	140.85104	38.47024			
21419	092324.47+533005.4	140.85197	53.50152			
21420	092324.49+034901.7	140.85207	3.81716			
21421	092325.25+453222.1	140.85521	45.5395			
21422	092326.45+254023.6	140.86021	25.67324			
21423	092326.53+264223.3	140.86055	26.7065			
21424	092326.86+543824.7	140.86192	54.64021			
21425	092326.88+794641.1	140.86201	79.77809			

Table data view

TOPCAT(1): Table Parameters

Window Parameters Display Help

Table Parameters for 1: 6dfgs\_mini.xml.bz2

Name	Value	Units	UCD	Description
Name	6dfgs_E7_subset			Table name
URL	jar.file:/data/andromeda1/starjava/java/li...			URL of original table
Column Count	17			Number of columns
Row Count	875			Number of rows
Description	6dFGS master config file (version E7 March...			
Original Source	http://www-wfau.roe.ac.uk/6dFGS/6dfgs_...			
Credits	Column explanations provided by Mike Re...			
Conversion	Converted from 6dfgs_E7.fld.gz by Mark T...			
RESOLUTION	15	arcsec	stat.error;pos.eq.ra	

Name: Description

Class: String

Shape:

Units:

Description:

UCD:

Value: 6dFGS master config file (version E7 March 2004) - DEMO SUBSET. These data are taken from the 6dF Galaxy Redshift Survey Database see astro-ph/0505068. Kindly provided by Mike Read, ROE. These data are for EXAMPLE PURPOSES ONLY, intended for demonstrations of some of TOPCAT's properties. For science use, please consult the

Table Metadata view

TOPCAT(1): Table Columns

Window Columns Display Help

Table Columns for 1: tgas\_source.fits

Δ	Index	Visible	Name	\$ID	Class	Units	Description
0		<input type="checkbox"/>	Index	\$0	Long		Table row index
1	1	<input checked="" type="checkbox"/>	source_id	\$1	Long		Unique source id
2	2	<input checked="" type="checkbox"/>	ra	\$2	Double	deg	Right ascension
3	3	<input checked="" type="checkbox"/>	dec	\$3	Double	deg	Declination
4		<input type="checkbox"/>	l	\$4	Double	deg	Galactic longitude
5		<input type="checkbox"/>	b	\$5	Double	deg	Galactic latitude
6		<input type="checkbox"/>	ecl_lon	\$6	Double	deg	Ecliptic longitude
7		<input type="checkbox"/>	ecl_lat	\$7	Double	deg	Ecliptic latitude
8	4	<input checked="" type="checkbox"/>	parallax	\$8	Double	mas	Parallax
9	5	<input checked="" type="checkbox"/>	pmra	\$9	Double	mas/yr	Proper motion in
10	6	<input checked="" type="checkbox"/>	pmdec	\$10	Double	mas/yr	Proper motion in
11	7	<input checked="" type="checkbox"/>	phot_g_mean_mag	\$11	Double	mag	G-band mean m
12		<input type="checkbox"/>	ref_epoch	\$12	Double	yr	Reference epoch
13	8	<input checked="" type="checkbox"/>	ra_error	\$13	Double	mas	Standard error o
14	9	<input type="checkbox"/>	dec_error	\$14	Double	mas	Standard error o

Column Metadata view

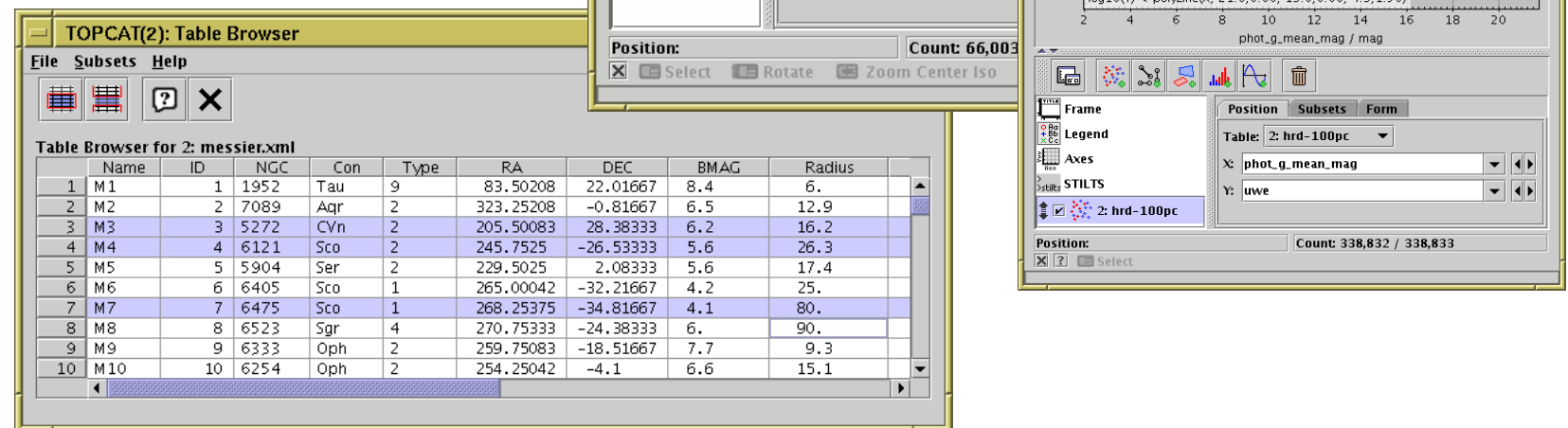
# Row Selections

Different ways to make single or multiple row selections:

- Select points graphically from a plot (freehand or polygon)
- Select rows from the table view
- Use an algebraic expression
- Combine existing subsets
- Receive from an external application (SAMP)

**Linked views** mean a selection made one way is visible in other ways

- Perform crossmatch only on items in red giant branch
- Where on the sky is this colour cut?
- Spot outliers
- Identify objects on ds9 image display

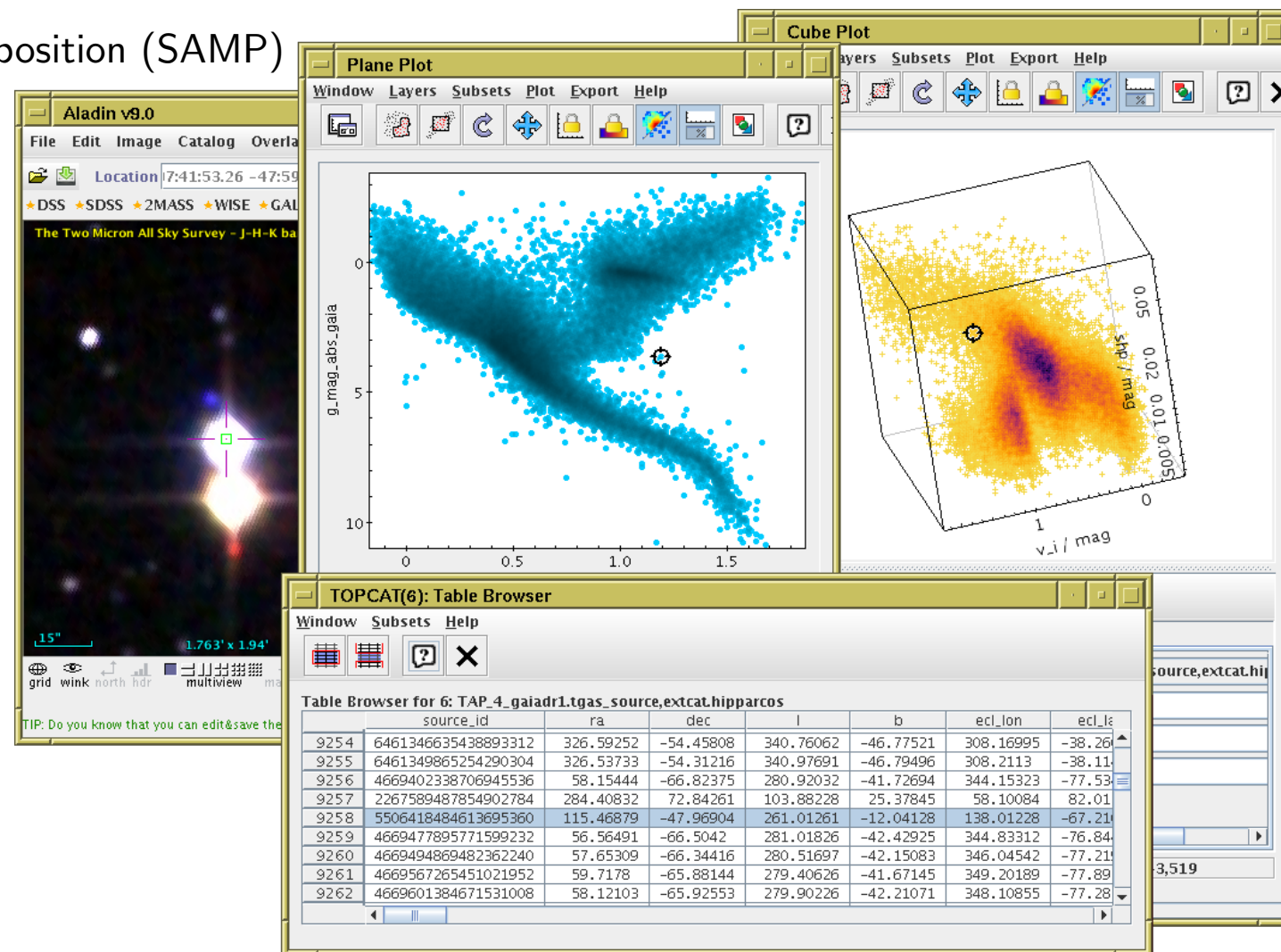




# Row Highlighting

Row selection is coordinated between linked views:

- Click on row in table browser or plot
- Same row is highlighted in other plots & table browser
- Can configure external tools to highlight same object/position (SAMP)





# Calculations

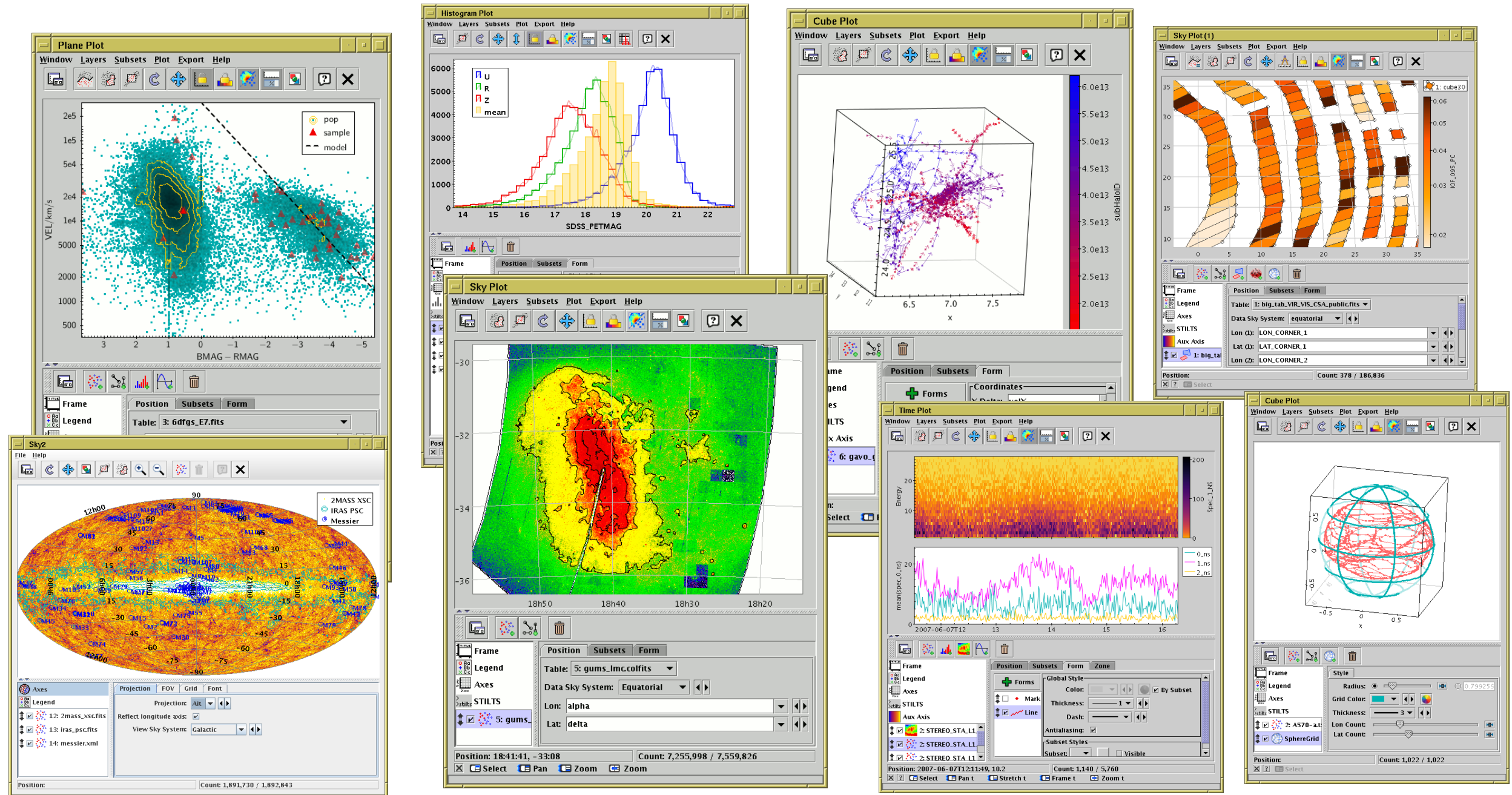
- Expression language used for creating columns, defining selections, making plots etc:
  - Straightforward arithmetic syntax (C-like)
  - Use column names as variables
  - Standard arithmetic operators (+, -, /, \*)
  - Conditional expressions (q?a:b)
  - Standard mathematical functions (abs, max, round, sin, cos, pow, ...)
  - Sky coordinates (degrees, sexagesimal, sky distances)
  - Astrometry (epoch propagation with/without errors, ...)
  - Cosmological distances (redshift, luminosity dist, lookback time, ...)
  - Fluxes (Johnson AB Magnitudes, Jansky)
  - Time conversions (ISO8601, MJD, Julian, Besselian)
  - ... and more (and it's extensible)
- Examples:
  - `mag_u - mag_g`
  - `janskyToAb(flux)`
  - `skyDistanceDegrees(ra, dec, 14.1, -72.9) < 1.2`

# Visualisation

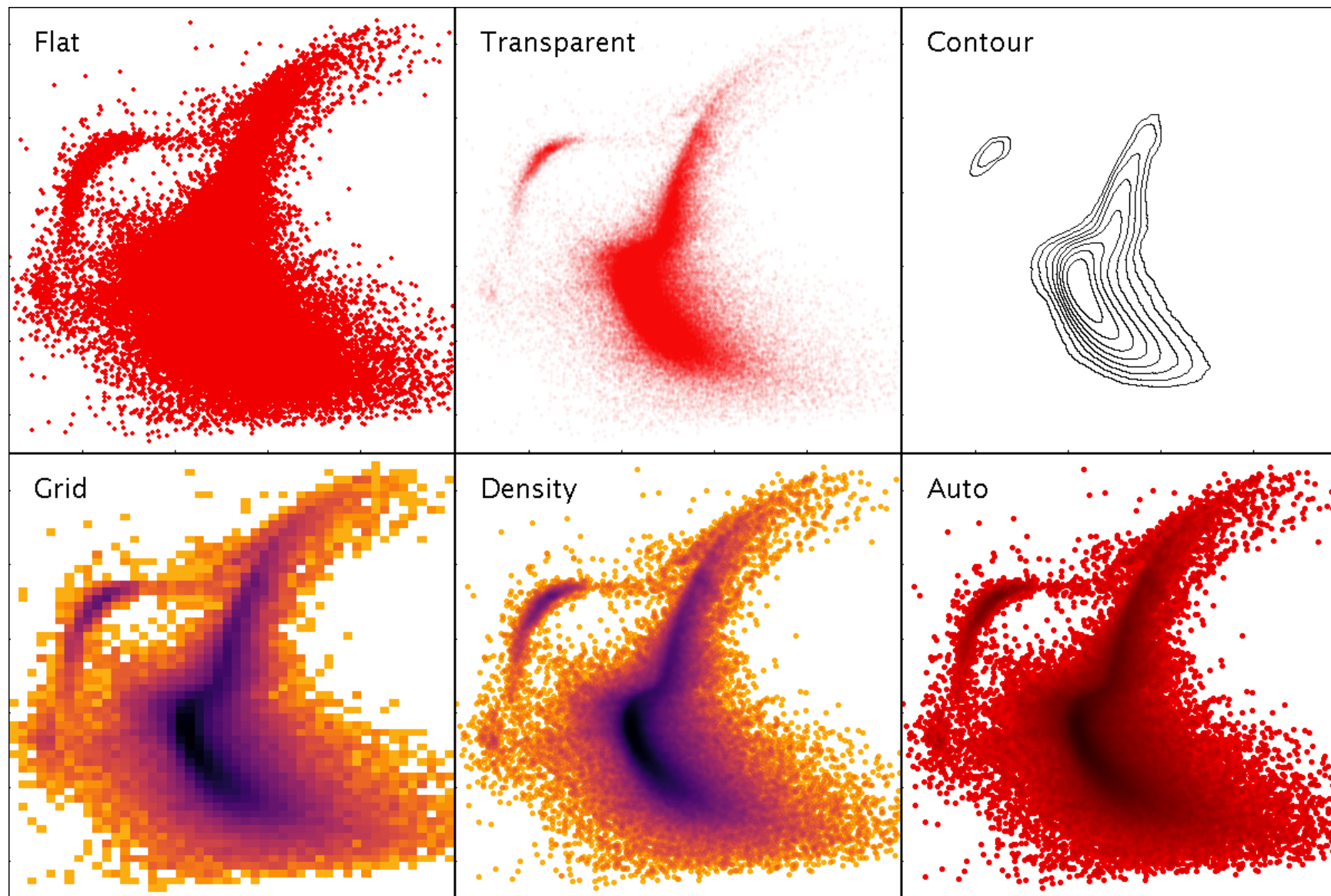
Very good for interactive exploration of large (or small) datasets:

- Many plot types!
  - ▷ 2d/3d scatter plots, histograms, HEALPix, density maps, error bars/ellipses, vectors, lines, quantiles, text labels, contours, KDEs, analytic functions, spectrograms, ...
- Many options!
  - ▷ Colour, colour maps, shading mode, weighting, marker shape/size line style, sky projection, sky system, coordinate grid, axis labelling, smoothing, binning, ...
- Highly responsive
  - ▷ Interactive changes to options update plot immediately
- Special attention to large data sets
  - ▷ Plot arbitrarily large datasets in fixed memory
  - ▷ Represent very dense plots in comprehensible ways
  - ▷ Many options for high-dimensional visualisation
- Publication-quality output?
  - ▷ Export to PDF, EPS, PNG, SVG (coming soon), ...
  - ▷ Optional LaTeX annotation
  - ▷ Script output (STILTS) for reproducibility
  - ▷ ... but not quite as good as Matplotlib/IDL/R

# Visualisation: Plot Types



## Visualisation: Dense plots



Different options for shading scatter-plot data.

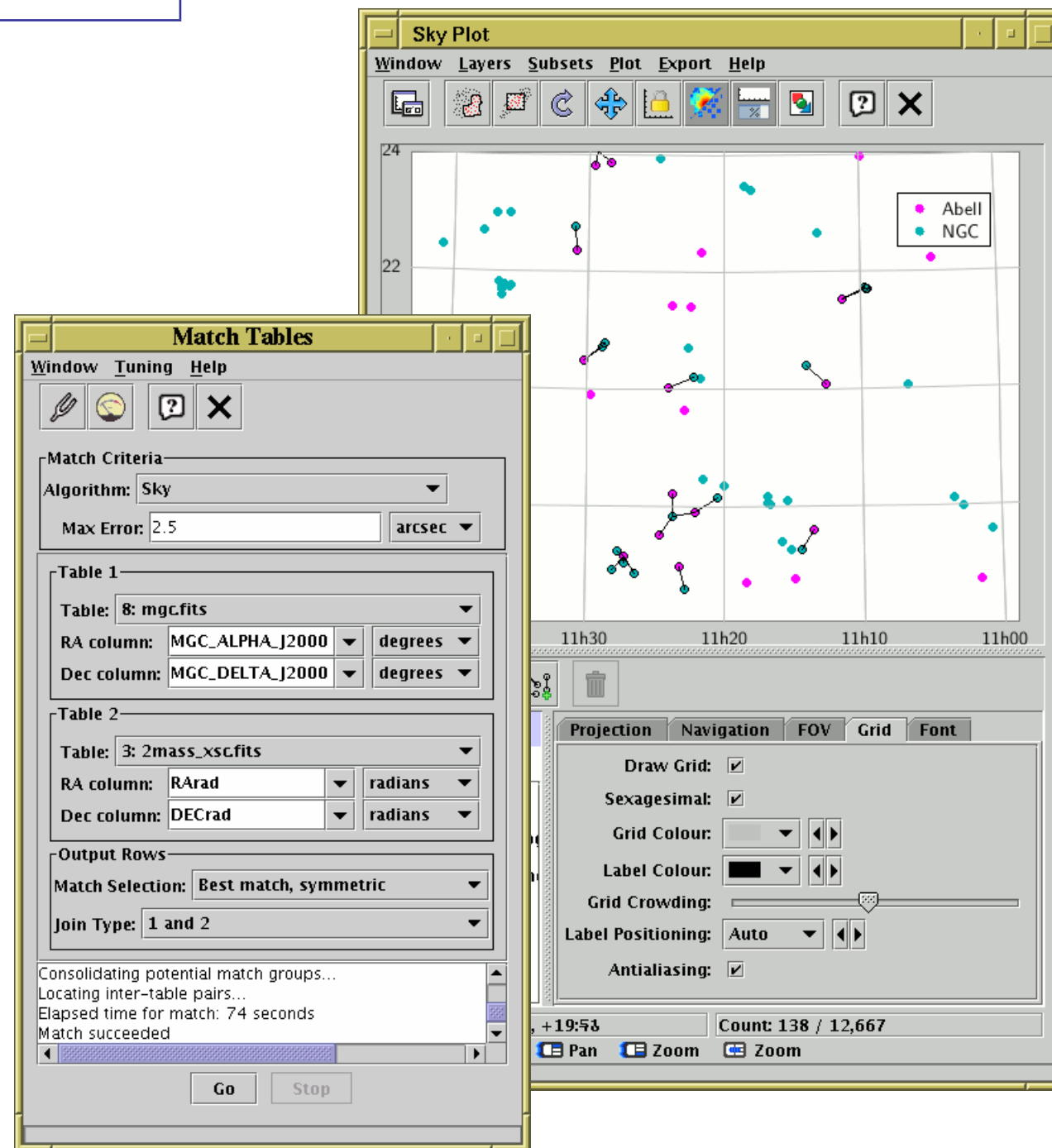
# Crossmatching

## Internal

- Both/all files loaded into TOPCAT
- Works well up to  $\sim 1$  million rows each
- Pretty fast ( $\leq$  couple of minutes)
- Very flexible (sky, Cartesian, exact, 3D, ellipses, errors, combinations...)

## External

- One or both tables too big to download
- Several options, with different pros and cons:
  - ▷ CDS X-Match  
(any VizieR table, sky match, fast, easy)
  - ▷ Multiple cone search  
(many tables available, sky match, slow)
  - ▷ TAP  
(few tables available, flexible, tricky)





# Virtual Observatory



## What is the Virtual Observatory (VO)?

- *"All astro archives in your computer"*
- A set of protocols that allows software clients to talk to external data services in a uniform way
- In most cases you (the software user) don't need to understand the details, but it's under the hood making data access work



## External data access from TOPCAT:

- Cone Search: positional query of remote catalogue
- Table Access Protocol (TAP): SQL-like queries against remote databases
- Simple Image Access/Simple Spectral Access: positional query of image/spectrum archives
- CDS services: Simbad, VizieR cone/all-sky, X-Match, Hips2fits
- Registry: service discovery
- SAMP: communication with other desktop/web applications


# STILTS

## STIL Tool Set (STIL = Starlink Tables Infrastructure Library)

- Has pretty much the same capabilities as TOPCAT
- but works from the command line (also [JyStilts](#) from Jython)



TOPCAT	STILTS
	
GUI	Command line
Interactive	Scriptable
Easy to use	Reproducible
Good for data exploration	Good for batch/programmed use
Exploratory phase	Production phase
few $\times 10^6$ rows	Unlimited size (for most things)

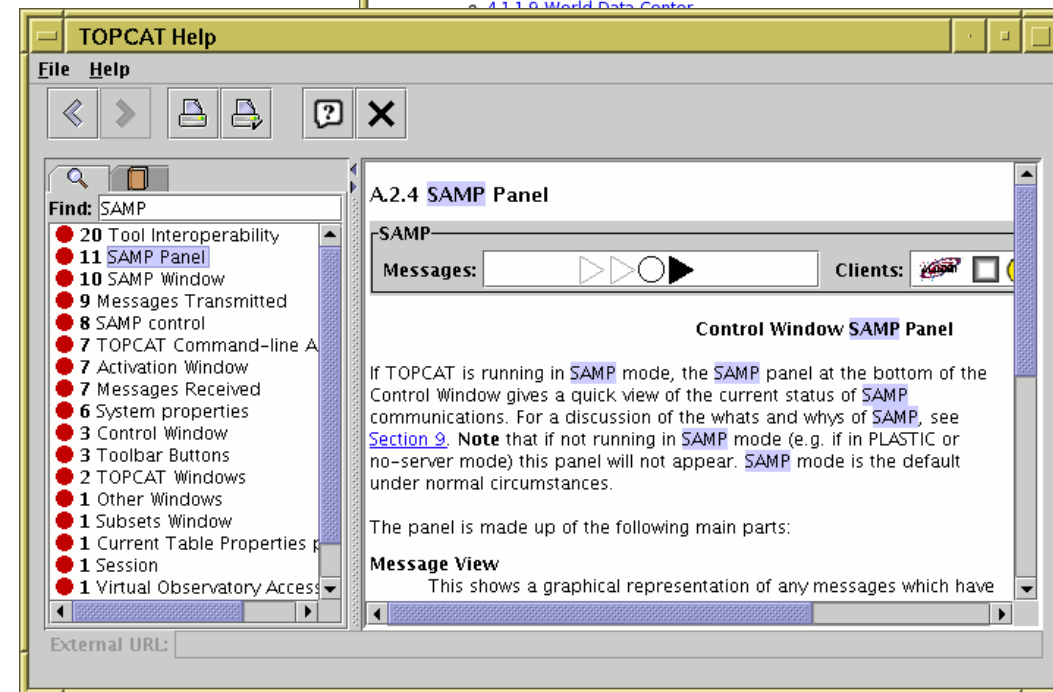
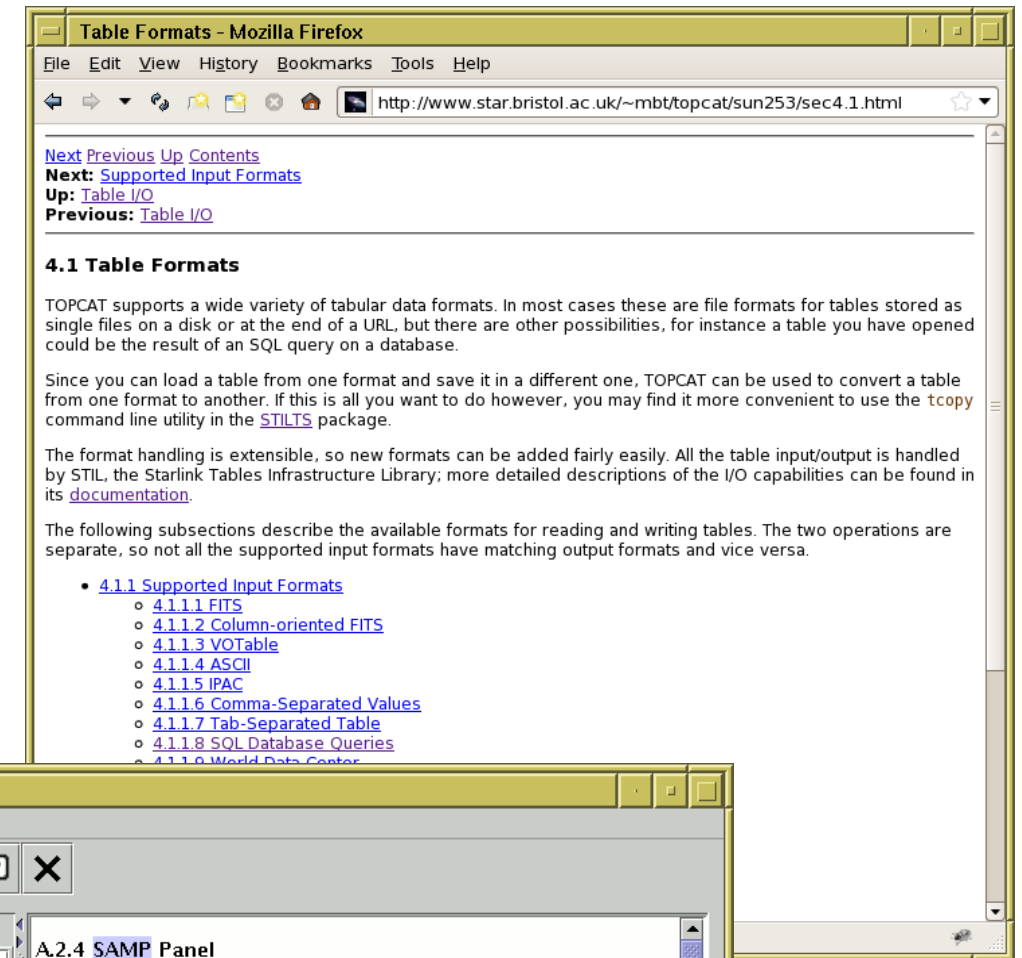
### Typical usage:

- start off with TOPCAT
- maybe move on to STILTS for more specialised requirements
- TOPCAT  STILTS control helps constructing plot commands



## Further Information

- There are things I haven't mentioned!
- Full tutorial and reference documentation:
  - ▷ HTML/PDF manual on web page  
<http://www.starlink.ac.uk/topcat/> (or Google it)
  - ▷ **Help for Window** button  on every window
  - ▷ Help browser includes search tool
  - ▷ More options in Help Menu (including **Help for Window in Browser** item )
  - ▷ Or print out the 500-page manual
- Support by email:
  - ▷ on list: [topcat-user@jiscmail.ac.uk](mailto:topcat-user@jiscmail.ac.uk)
  - ▷ in person: [m.b.taylor@bristol.ac.uk](mailto:m.b.taylor@bristol.ac.uk)
  - ▷ All feedback and questions (even dumb questions) welcome!



# Demo 1: Pleiades in Gaia and 2MASS

**Cone Search**

Window Columns Registry Interop Help

Available Cone Services

Registry:  ☐ RegTAP

Keywords:

Match Fields: ☒ Short Name ☒ Title ☒ Subjects ☒ ID ☒ Publisher ☐ Description

☒ Accept Resource Lists

Short Name	Title	
ARI-Gaia	ARI's Cone Search Service for the last Gaia Data Release (DR2)	Gaia DR
GAIA DR2	Gaia DR2 at ESA	Gaia, DR
GAIA DR2 CS	GAIA DR2 ConeSearch	Surveys
GDR2light SCS	Gaia DR2-light Cone Search	Gaia, As
I/345	Gaia DR2 (Gaia Collaboration, 2018)	Stars:va
I/347	Distances to 1.33 billion stars in Gaia DR2 (Bailer-Jones+, 2018)	Stars, Po

AccessURL	Description	Version
<a href="http://gaia.ari.uni-heidelberg...">http://gaia.ari.uni-heidelberg...</a>		

Resource Count: 20

Cone Parameters

Cone URL:

Object Name:

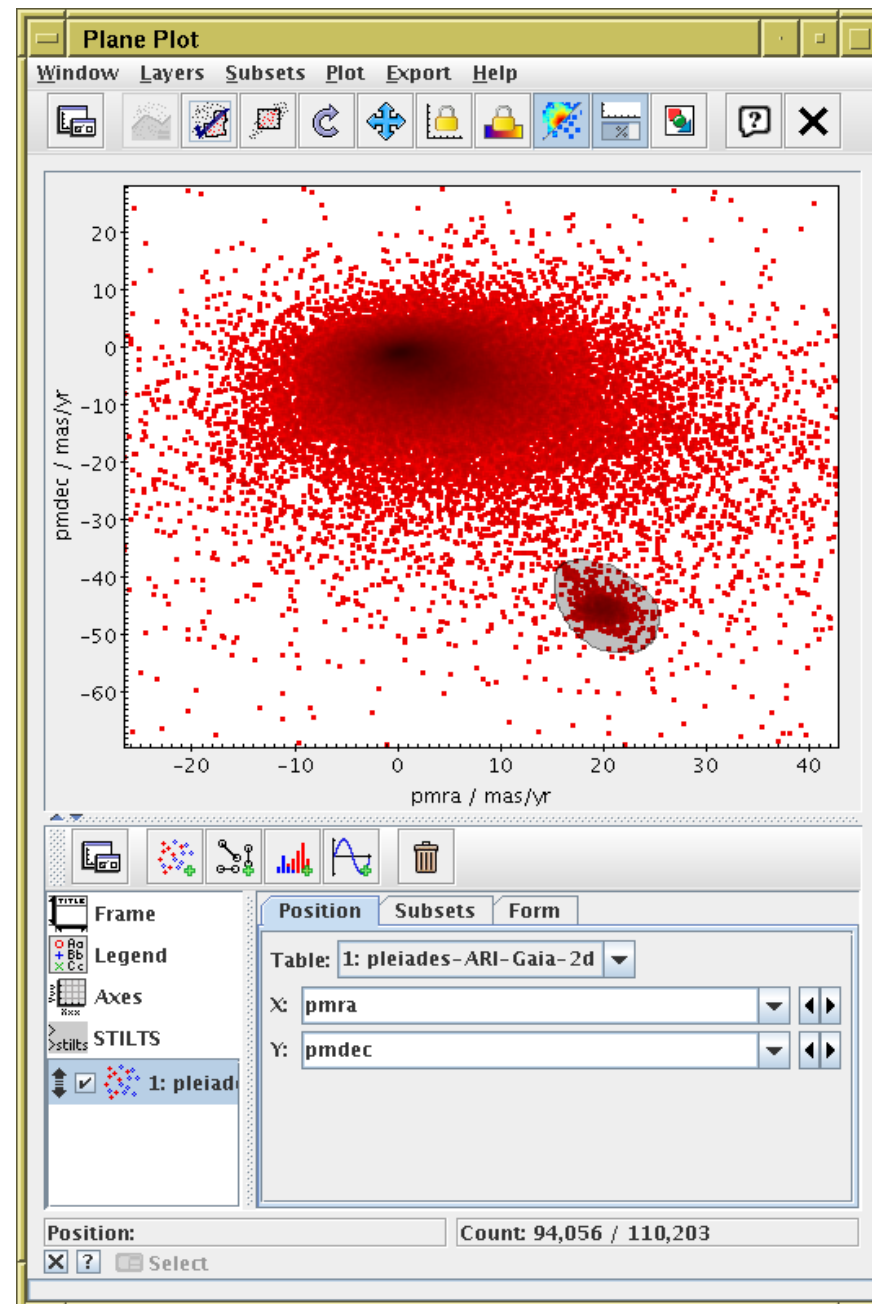
RA:  degrees (J2000) ☒ Accept Sky Positions

Dec:  degrees (J2000)

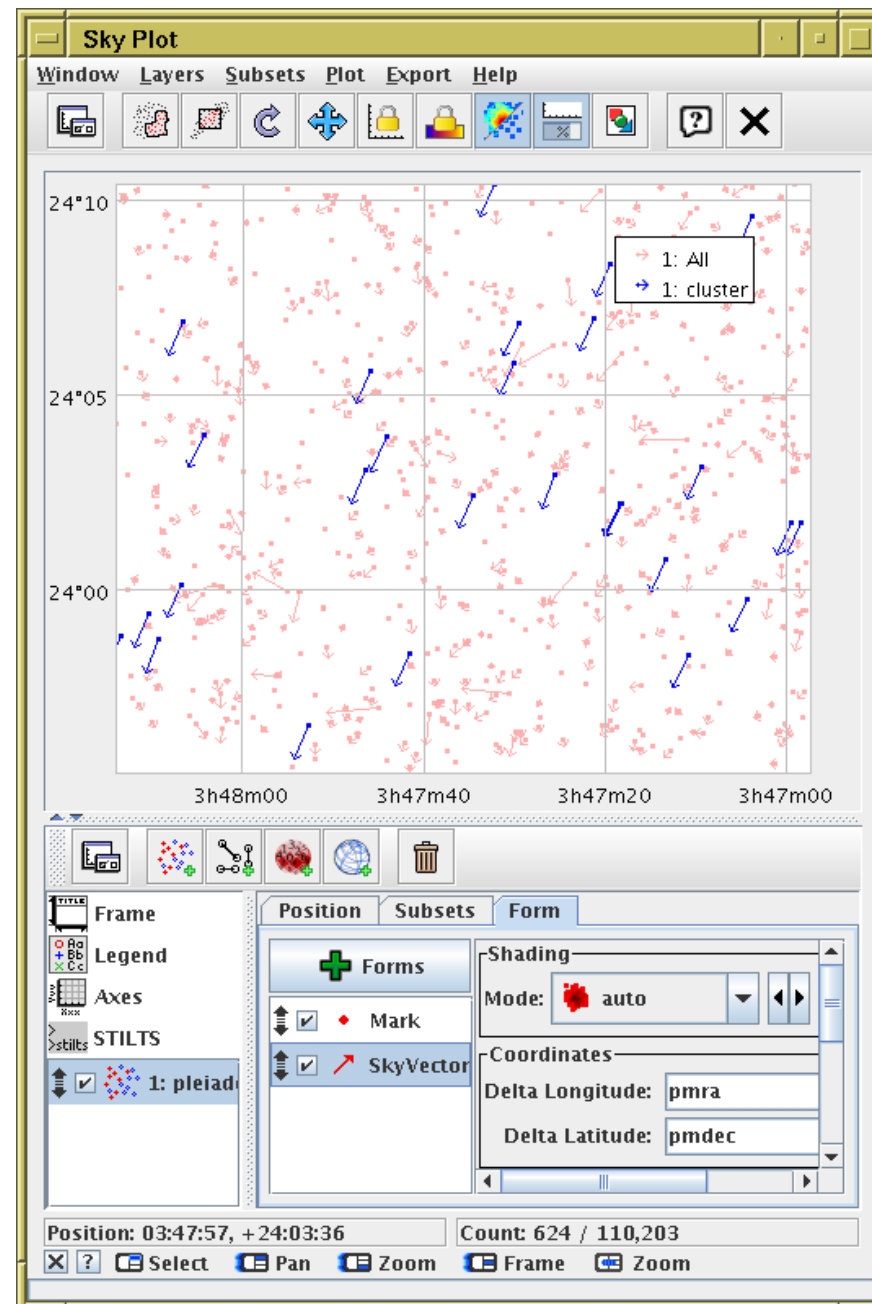
Radius:  degrees

Verbosity:

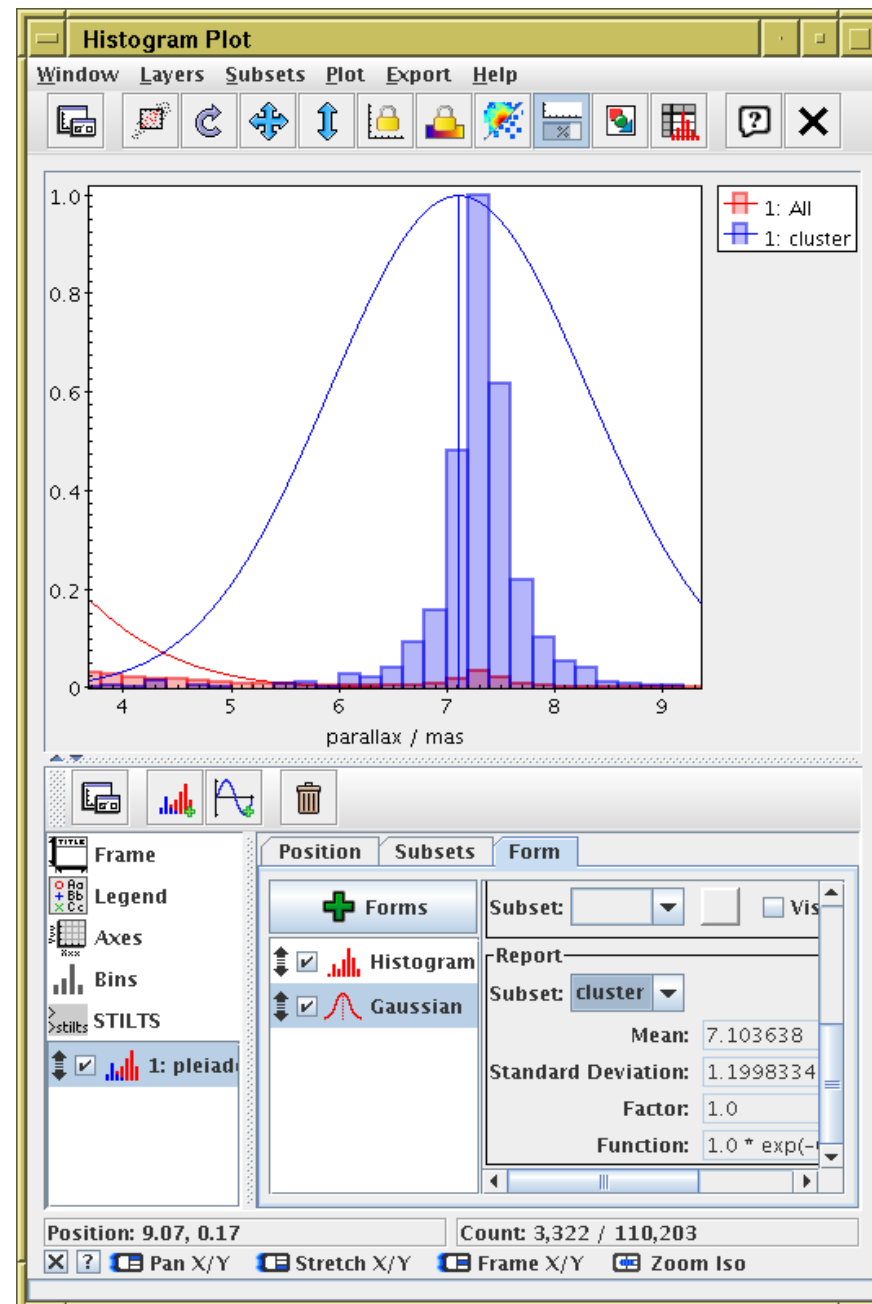
Acquire Gaia DR2 data in the region of the Pleiades using the Cone Search window



Plot the points in proper motion space,  
and select the comoving sub-population graphically to create a new subset

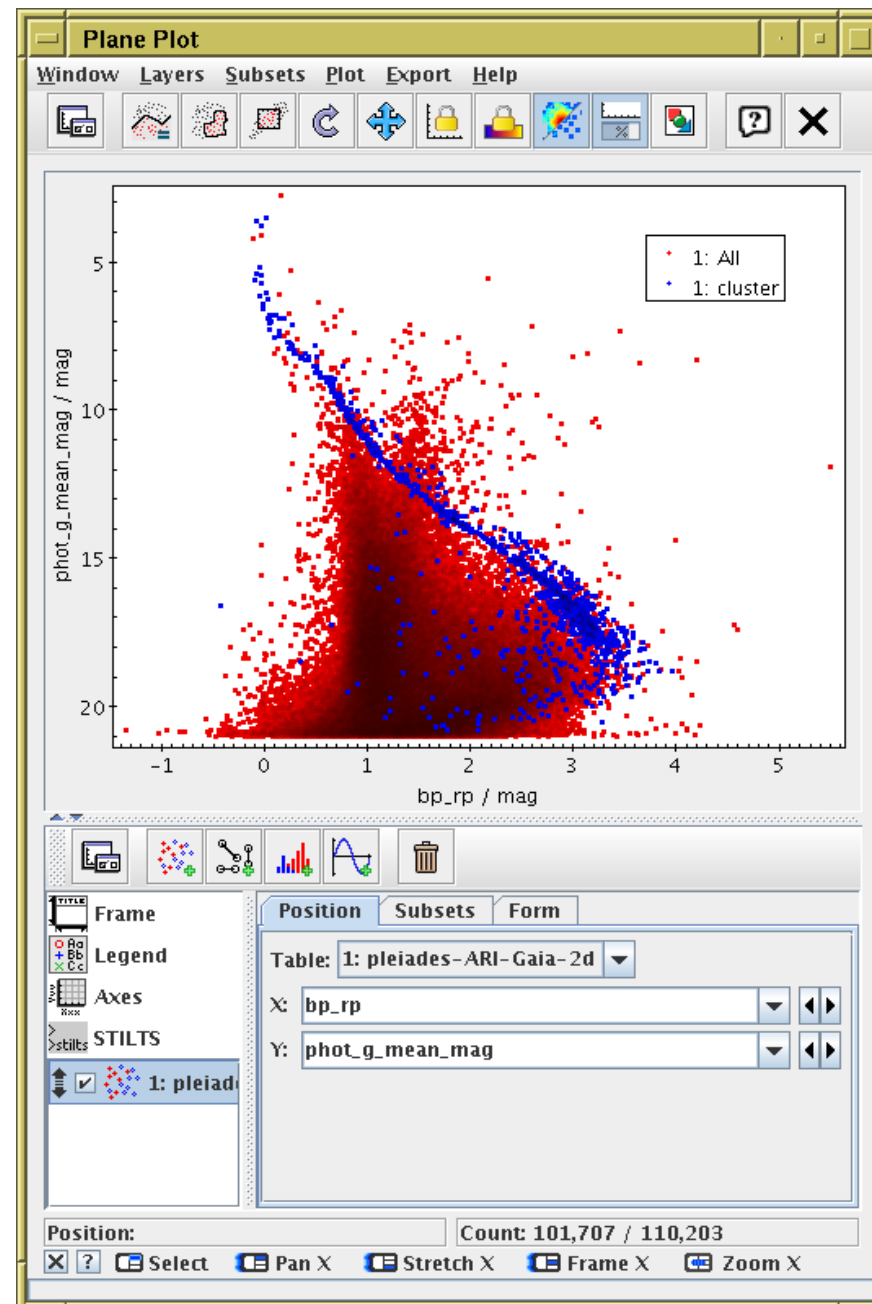


Plot the background and cluster objects on the sky with their proper motion vectors



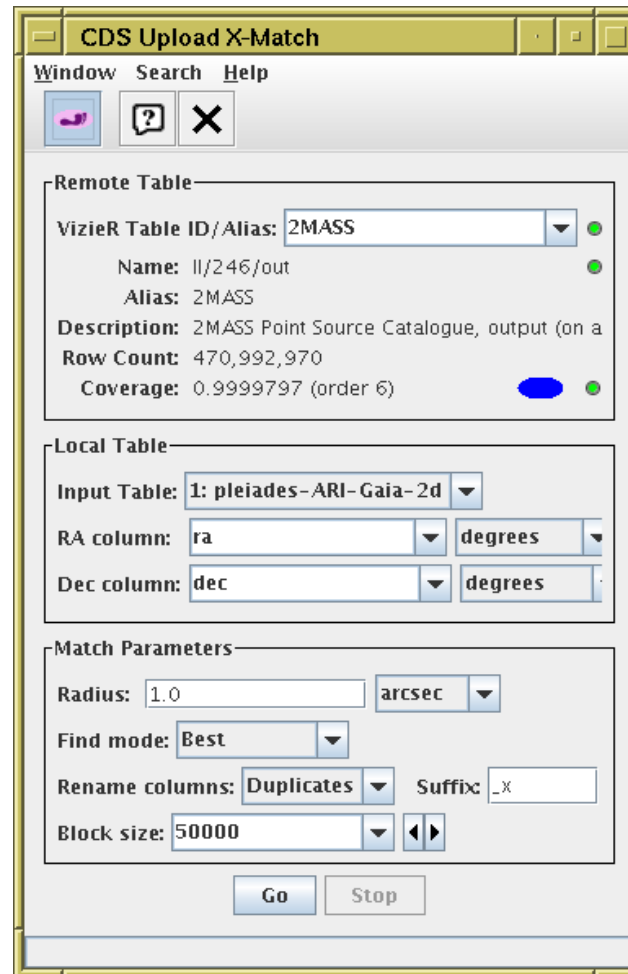
Plot a histogram of the background and cluster objects.

Fit a Gaussian to calculate mean parallax hence distance of cluster objects:  $1000/7.1 \text{ mas} \simeq 140 \text{ pc}$ .

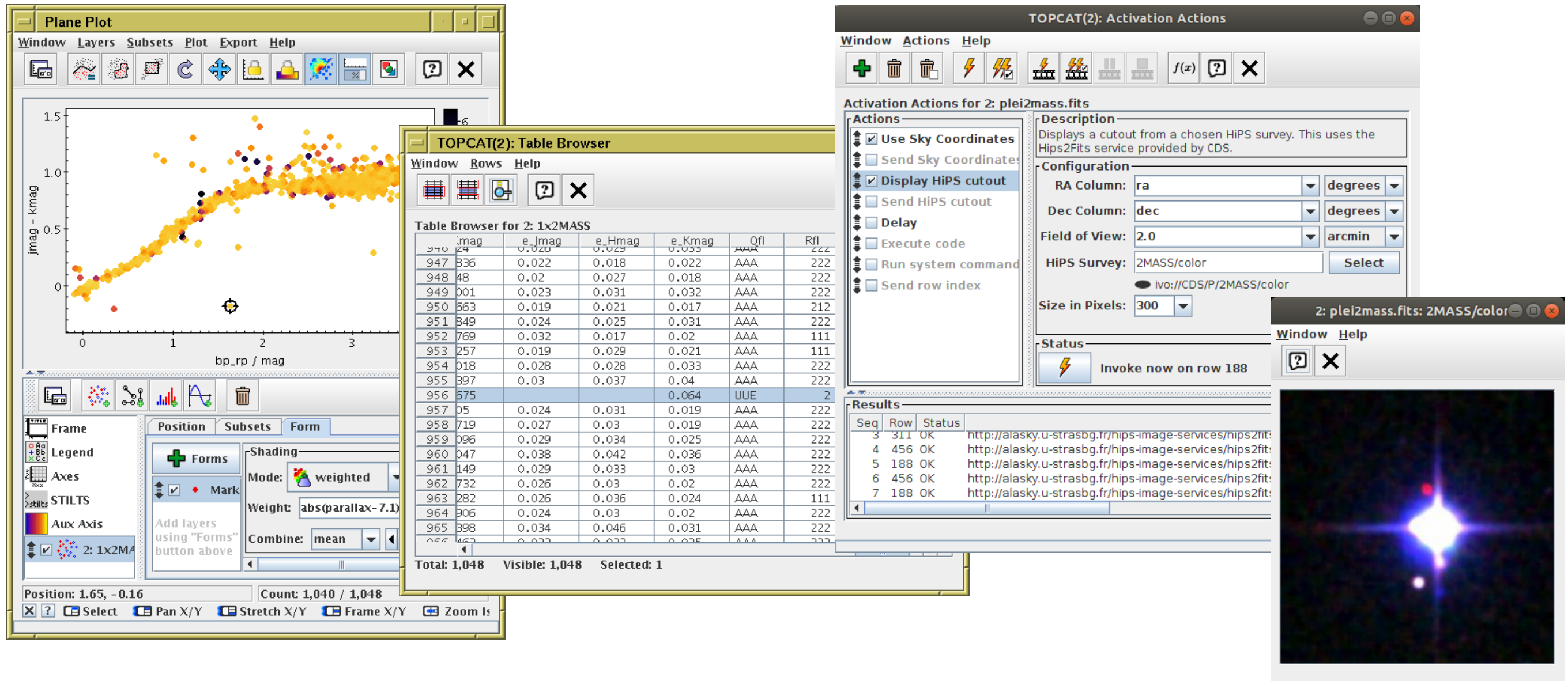


Plot colour-magnitude diagram using Gaia photometry for cluster and background objects





Find 2MASS associations for cluster objects using CDS X-Match window



Plot colour-colour diagram using Gaia and 2MASS photometry (Gaia  $bp\_rp$  vs. 2MASS  $Jmag - Kmag$ ) and investigate outliers. Colour points using distance from mean cluster parallax; darker ones are PM interlopers. Set up Activation Action **Display HiPS Cutout**; use survey **2MASS/color**. Click on point of interest for linked view: in table display and 2MASS imagery in image window.

## Demo 2: Hyades in 3D using TAP

The screenshot shows the Table Access Protocol (TAP) Query interface. The left pane displays a list of selected TAP services, with 'GAIA (71/83)' selected. The right pane shows the 'Use Service' tab, where the 'gaiadr2.gaia\_source' table is selected. The 'ADQL Text' field contains the following query:

```
SELECT ra, dec, pmra, pmdec, parallax, radial_velocity,
       phot_g_mean_mag, bp_rp
FROM gaiadr2.gaia_source
WHERE parallax > 15
      AND parallax_over_error > 5
      AND radial_velocity IS NOT NULL
```

The 'Run Query' button is visible at the bottom of the interface.

Use ESA Gaia TAP service  
Investigate metadata  
Load position and velocity information  
for nearby sources:

```
SELECT ra, dec, pmra, pmdec, parallax,
       radial_velocity, bp_rp,
       phot_g_mean_mag + 5*log10(parallax/100) as g_abs
FROM gaiadr2.gaia_source
WHERE parallax > 15
      AND parallax_over_error > 5
      AND radial_velocity IS NOT NULL
```



# Do It Yourself!

Work through the tutorial at

<https://github.com/mbtaylor/tctuto/releases/download/asterics-vo-school-4/tctuto.pdf>

- ▷ The first part is what I've demonstrated today
- ▷ Later parts contain other examples and functionality
- ▷ Do whichever parts you like!

Full documentation available online (or in topcat):

<http://www.starlink.ac.uk/topcat/>

<http://www.starlink.ac.uk/stilts/>