

Topcat introduction + examples of classification algorithms

Hands-on session

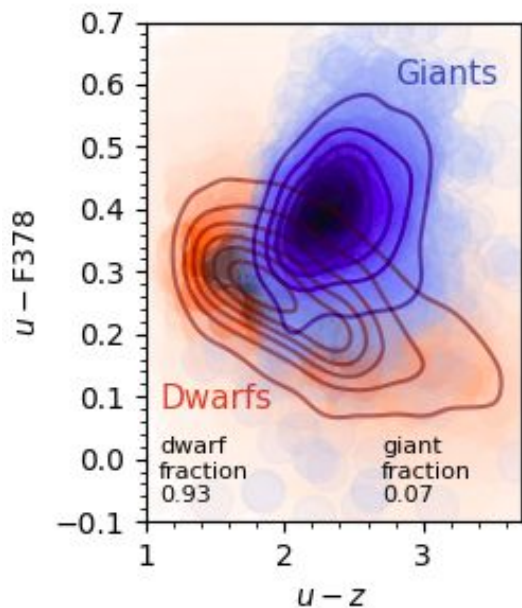
Felipe Almeida-Fernandes

Post Doc at IAG/USP, São Paulo

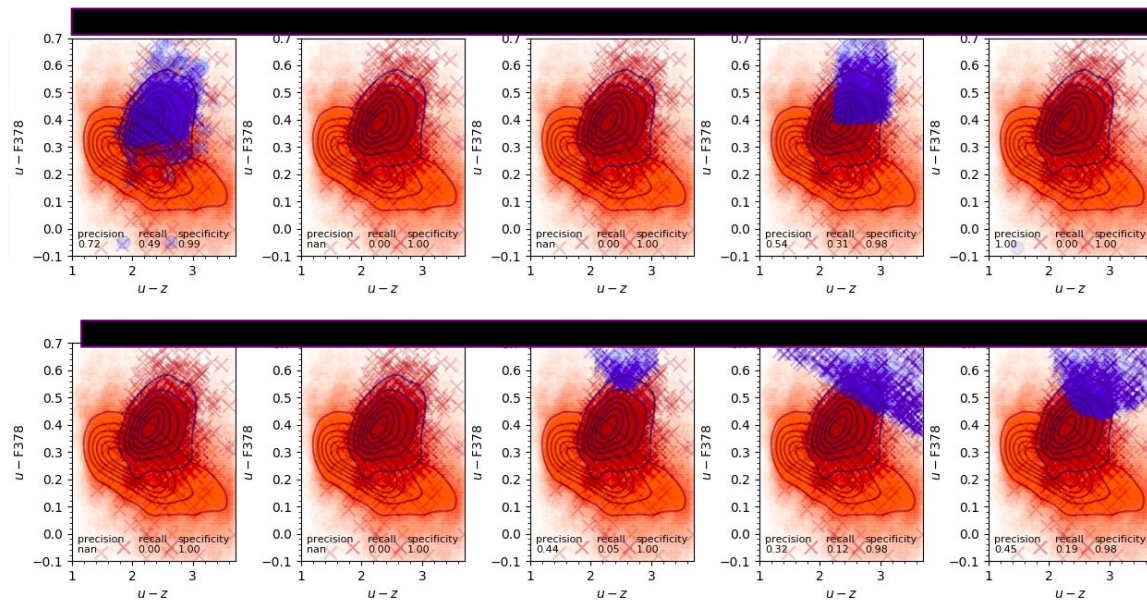
For this you will need:

- Topcat
- Python, to run a python.py script
 - scikit-learn
 - numpy
 - pandas
 - (optional) matplotlib
 - (optional) scipy

We will use topcat to crossmatch S-PLUS and SSPP samples. And we will apply some selections to (kind off) reproduce this plot:



Then, we will use the data we produced, together with a python script, to test different classification algorithm and compare their results:



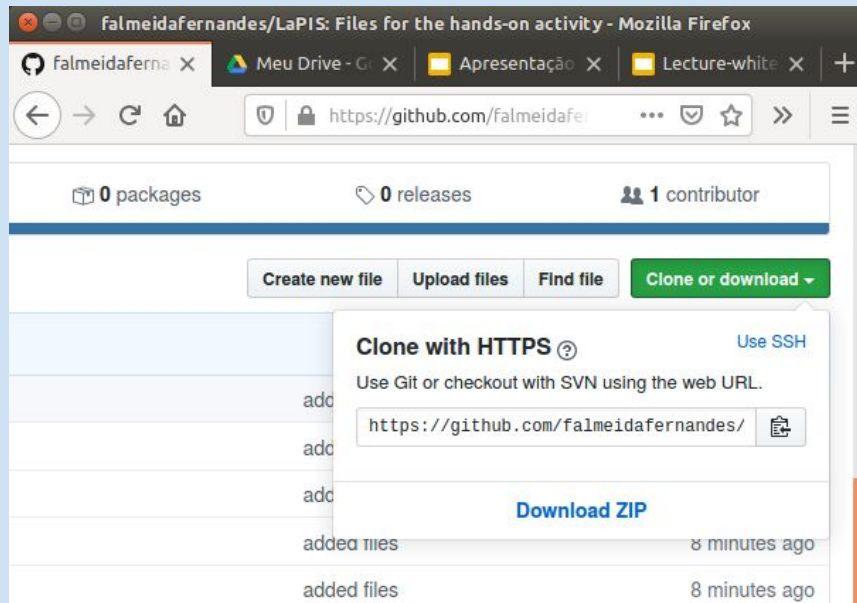
Getting the data

Option 1

```
.../your-directory$ git clone https://github.com/falmeidafernandes/LaPIS.git
```

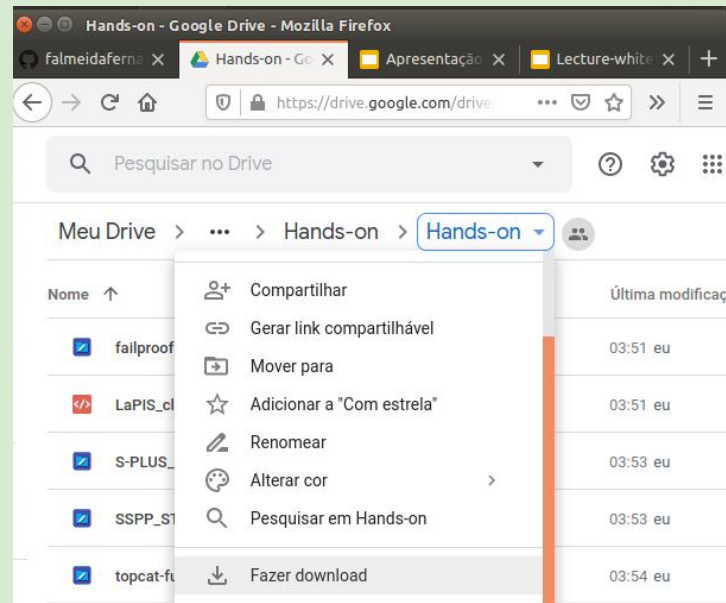
Option 2 - from github page

<https://github.com/falmeidafernandes/LaPIS>

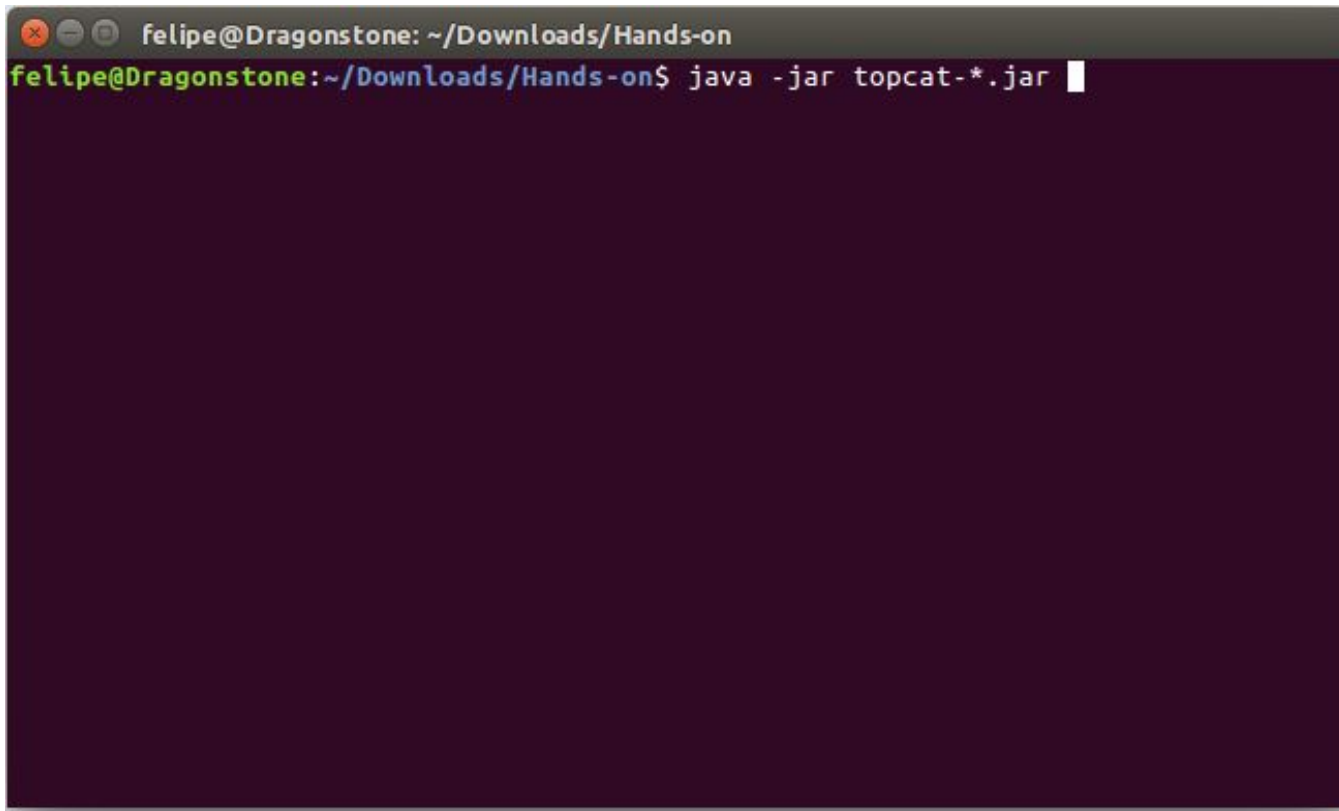


Option 3 - google drive

<http://tiny.cc/zmr9jz>



1) Open Topcat

A terminal window with a dark purple background. The title bar shows 'felipe@Dragonstone: ~/Downloads/Hands-on'. The prompt is 'felipe@Dragonstone:~/Downloads/Hands-on\$' and the command entered is 'java -jar topcat-*.jar' with a white cursor at the end.

```
felipe@Dragonstone: ~/Downloads/Hands-on
felipe@Dragonstone:~/Downloads/Hands-on$ java -jar topcat-*.jar
```

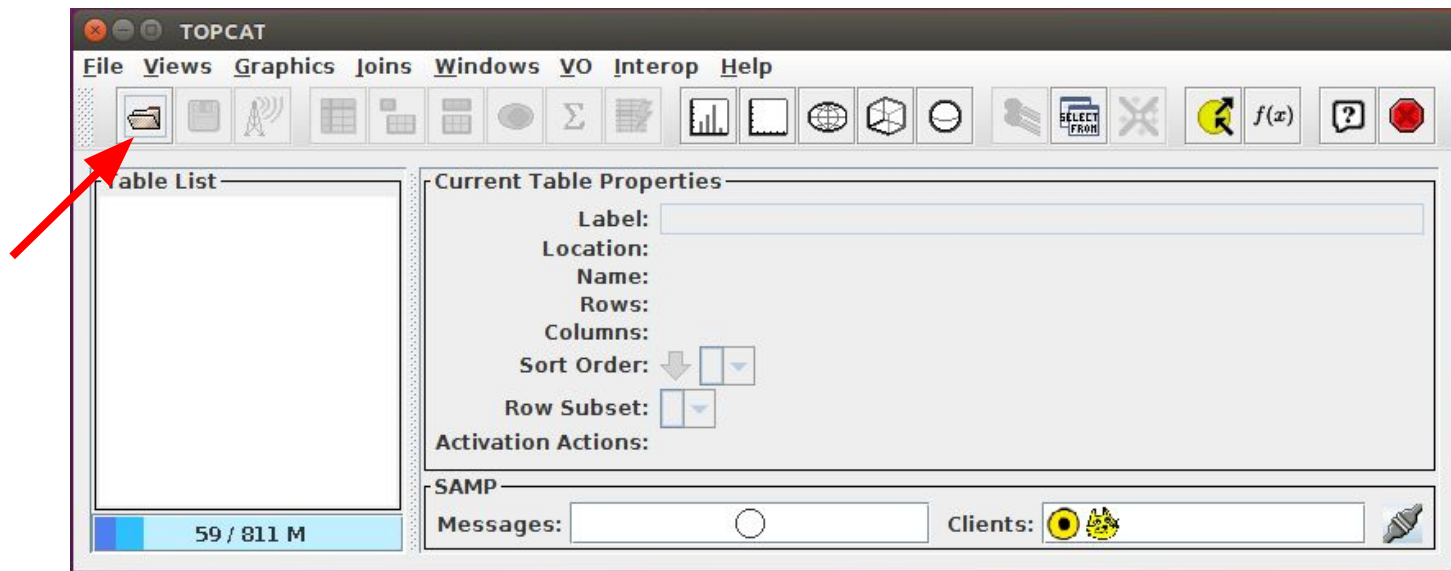
On unix systems:

```
.../Hands-on$ chmod +x topcat
.../Hands-on$ topcat
```

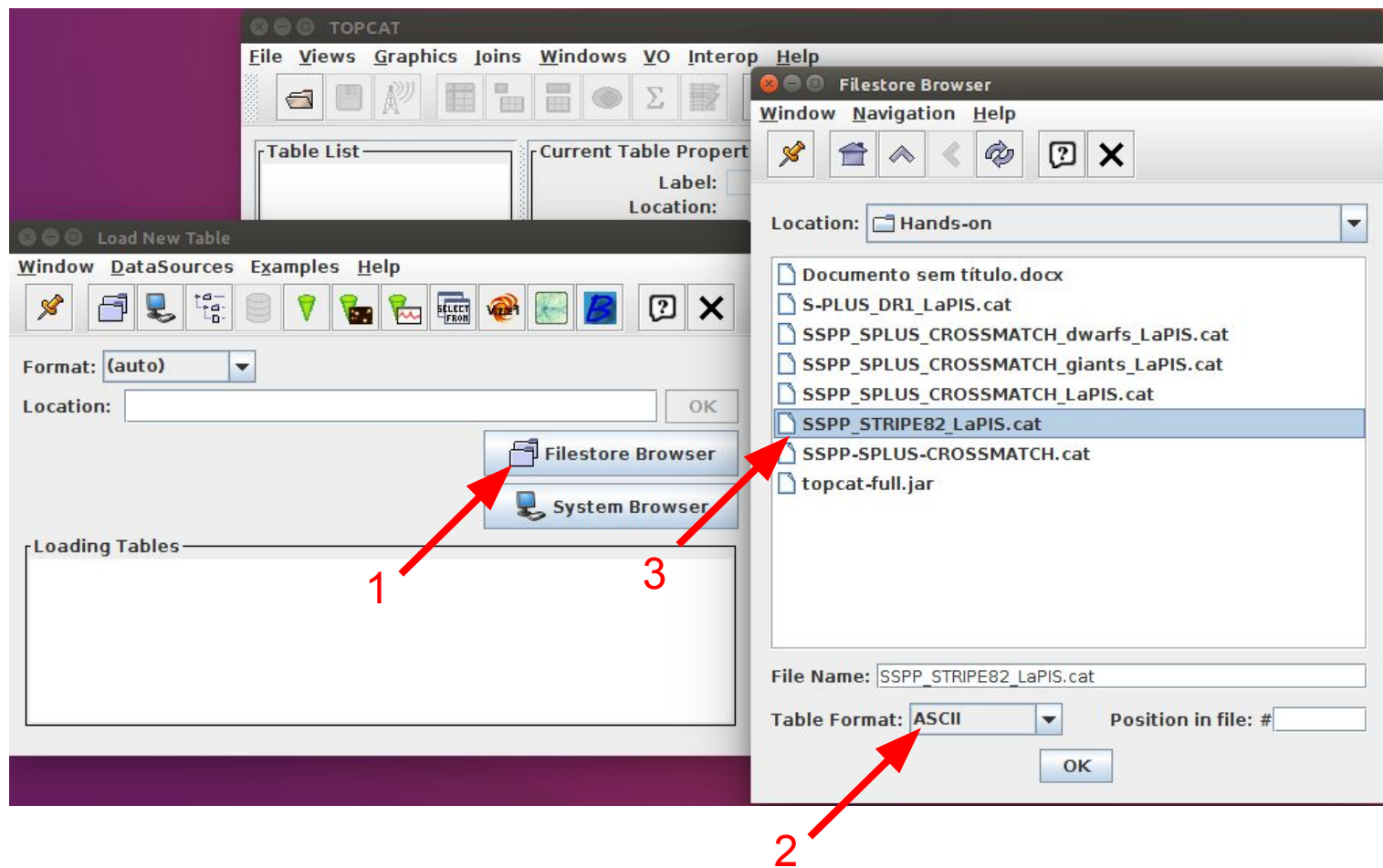
On non-unix systems:

```
.../Hands-on$ java -jar topcat-*.jar
```

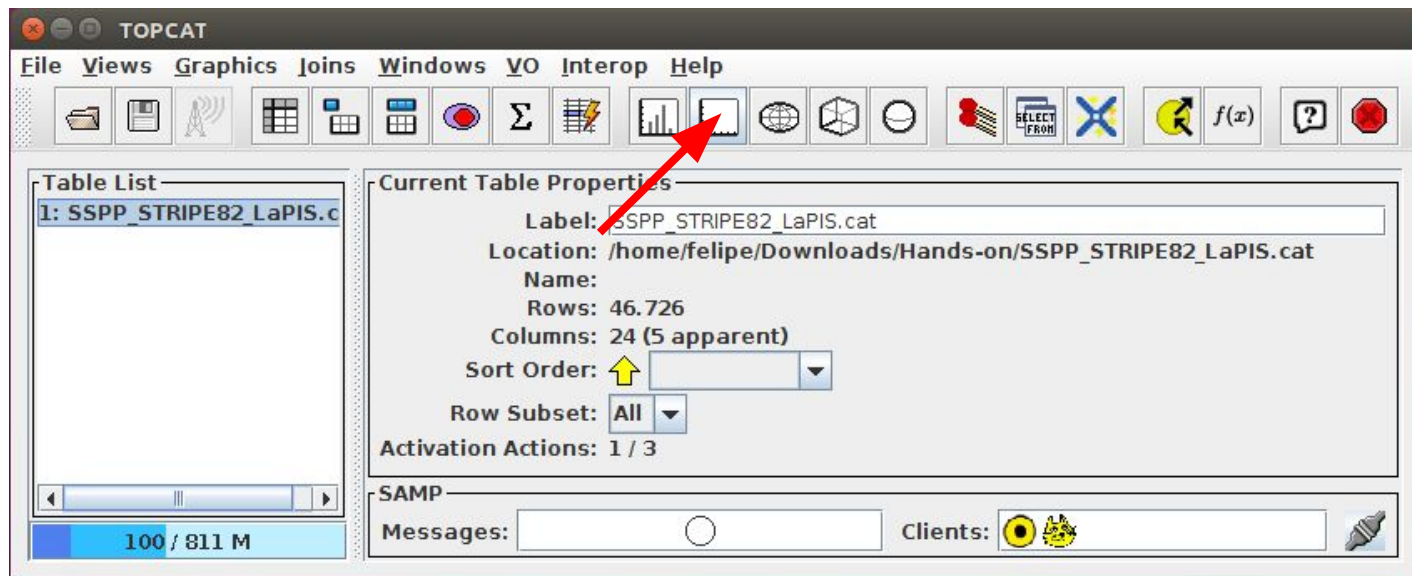
- 1) Open Topcat
- 2) Load the SSPP sample



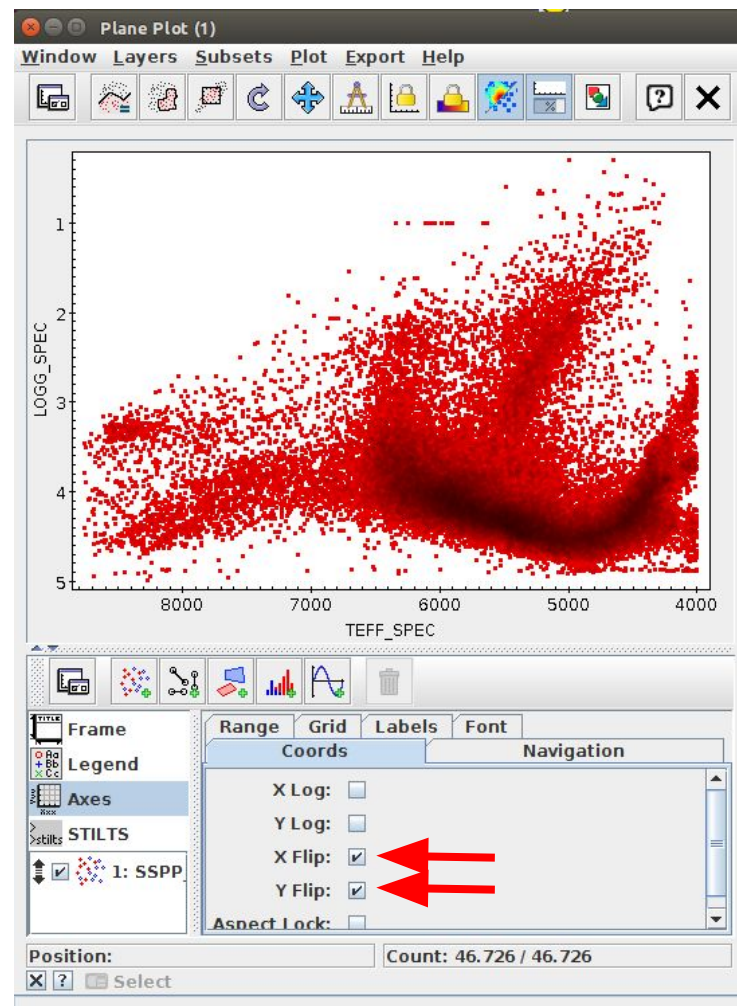
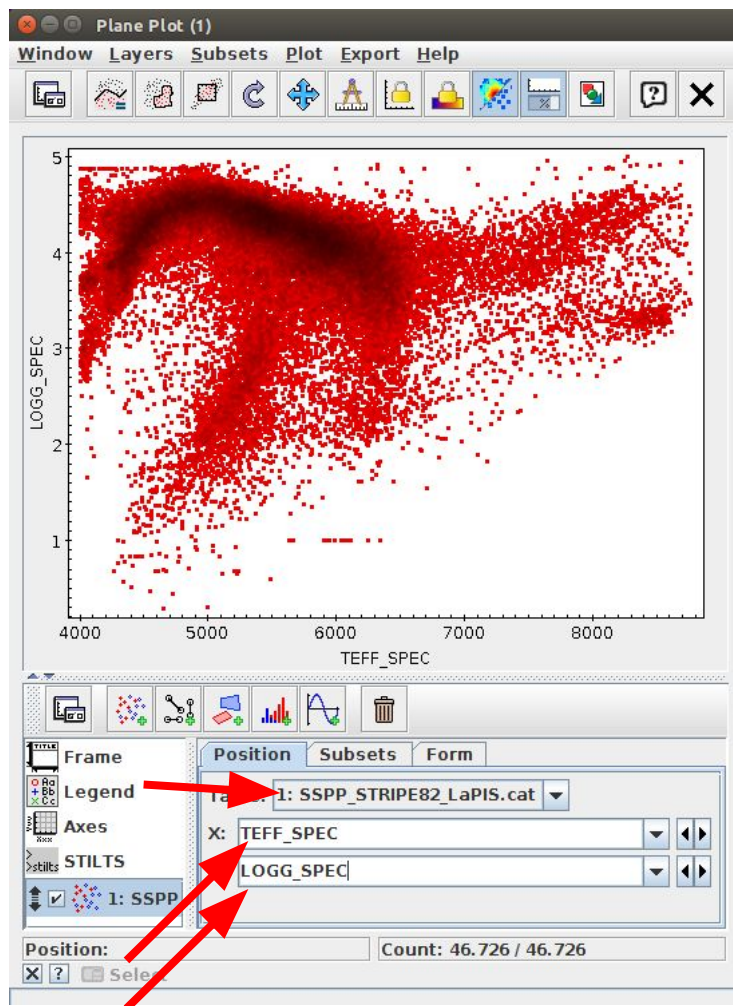
- 1) Open Topcat
- 2) Load the SSPP sample



- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data



- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data



1) Open Topcat

2) Load the SSPP sample

3) Plot spec data

4) Let's get rid of hot and cold dwarfs

The image shows the TOPCAT software interface with two windows. The main window, 'TOPCAT', has a menu bar (File, Views, Graphics, Joins, Windows, VO, Interop, Help) and a toolbar. The 'Table List' on the left shows '1: SSPP_STRIPE82_LaPIS.c'. The 'Current Table Properties' on the right show details for this table, including its location, name, row count (46,726), column count (24, with 5 apparent), and sort order. A red arrow labeled '1' points to the 'Current Table Properties' window.

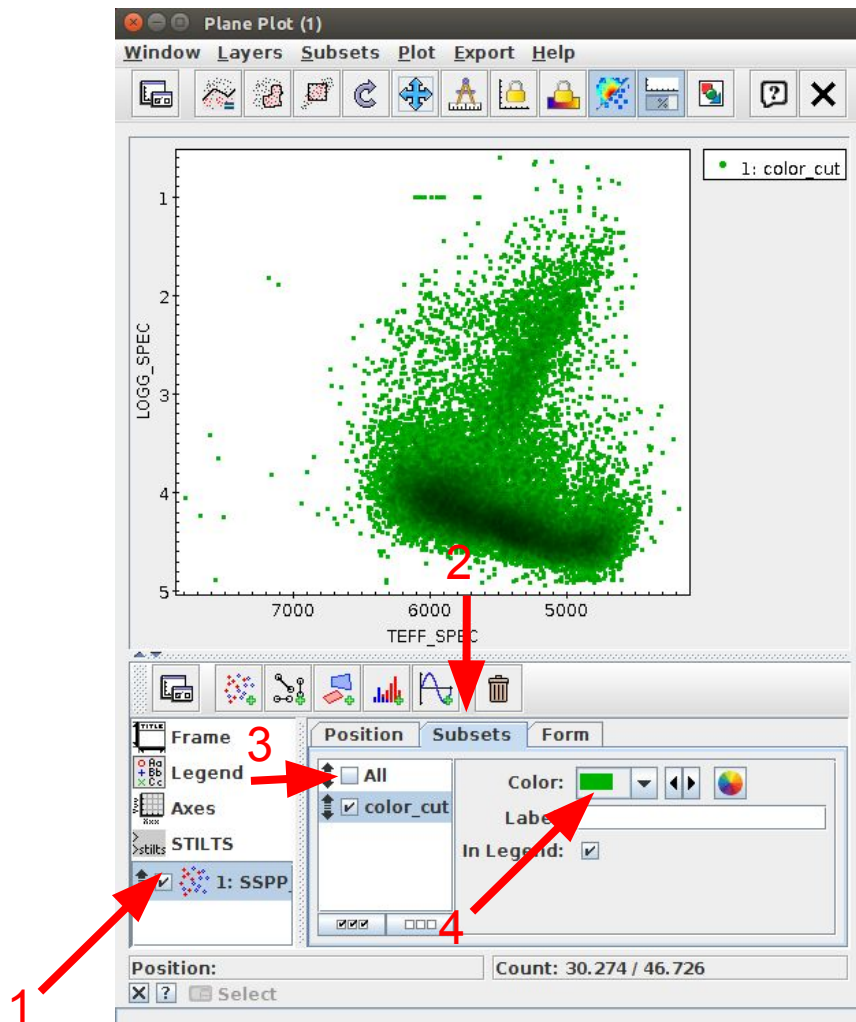
The second window, 'TOPCAT(1): Row Subsets', is open in the foreground. It has a menu bar (Window, Subsets, Display, Interop, Help) and a toolbar. The 'Row Subsets for 1: SSPP_STRIPE82_LaPIS.cat' table is shown with a single subset named 'All'. A red arrow labeled '2' points to the 'Add' button (a green plus sign) in the toolbar.

A 'Define Row Subset' dialog box is open over the 'Row Subsets' window. It has a menu bar (Window, Help) and a toolbar. The 'Subset Name' is set to 'color_cut' and the 'Expression' is 'GR > 0.3 & GR < 0.85'. A red arrow labeled '3' points to the 'OK' button in the dialog box.

ID	Name	Size
1	All	4

Subset Name	Expression
color_cut	GR > 0.3 & GR < 0.85

- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs



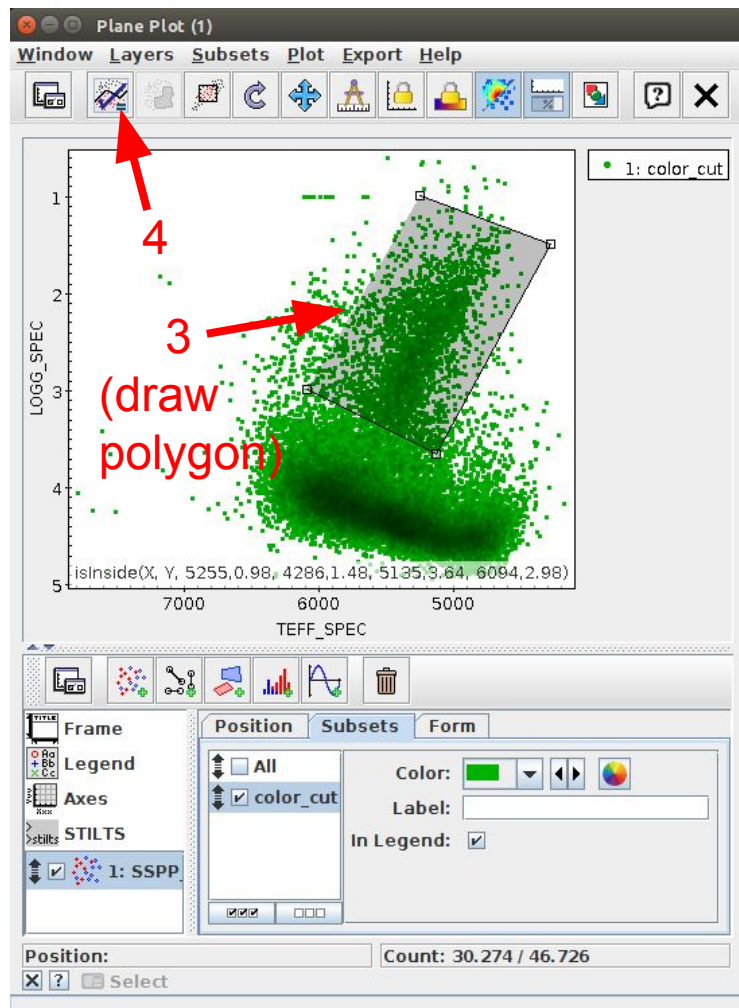
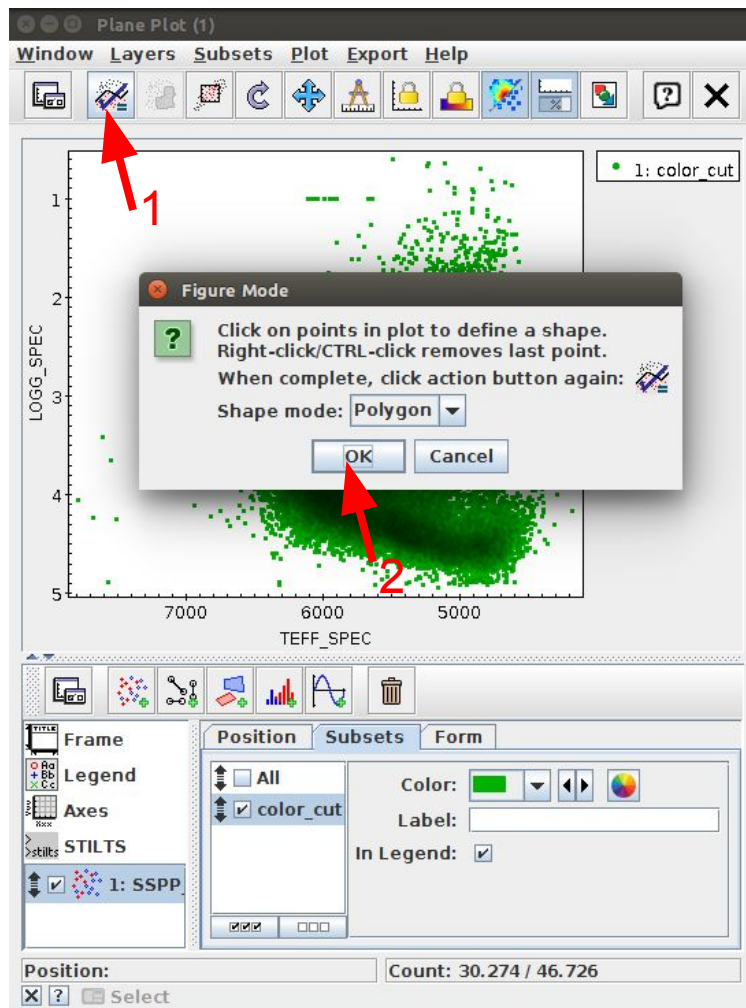
1) Open Topcat

2) Load the SSPP sample

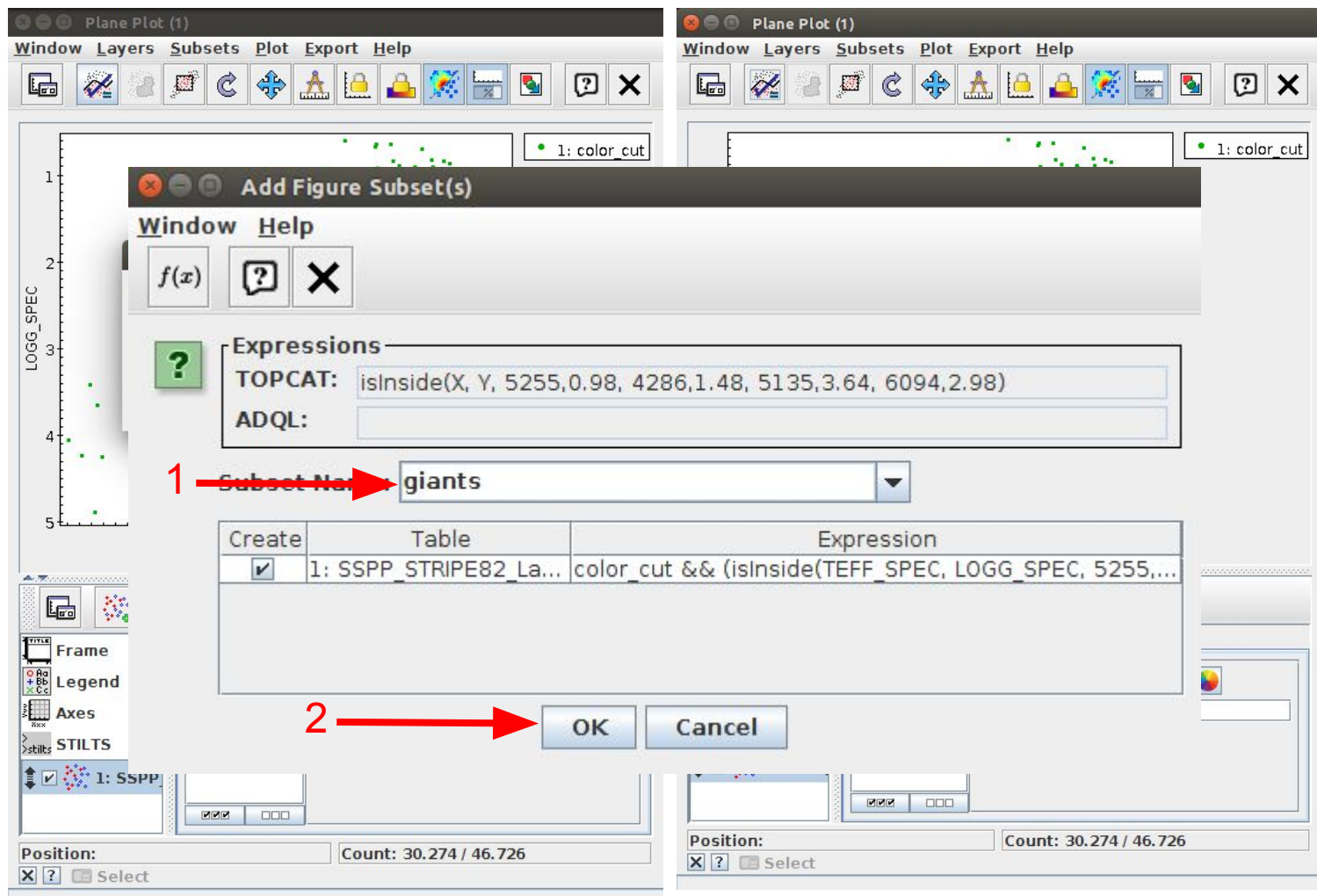
3) Plot spec data

4) Let's get rid of hot and cold dwarfs

5) Select the giants



- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants

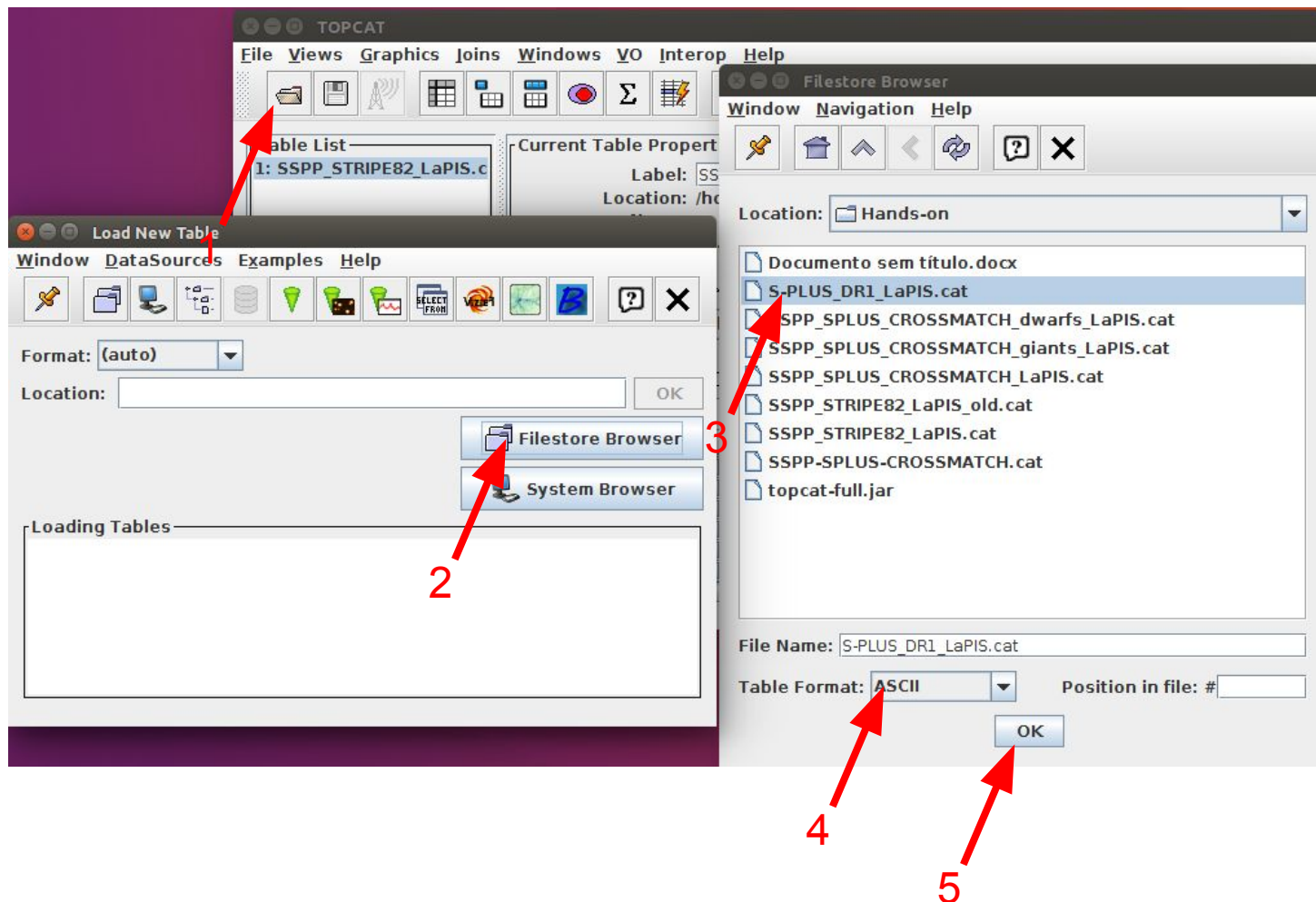


- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column

The screenshot displays the TOPCAT software interface with several windows open. A red arrow labeled '1' points to the 'Windows' menu. Another red arrow labeled '2' points to the 'Define Synthetic Column' dialog box. A third red arrow labeled '3' points to the 'Name' field, which contains the text 'giant'. A fourth red arrow labeled '4' points to the 'Expression' field, which contains the text '3'. A fifth red arrow labeled '5' points to the 'OK' button. The 'Row Subsets' window is also visible, showing a table with columns ID, Name, Size, Fraction, and Expression.

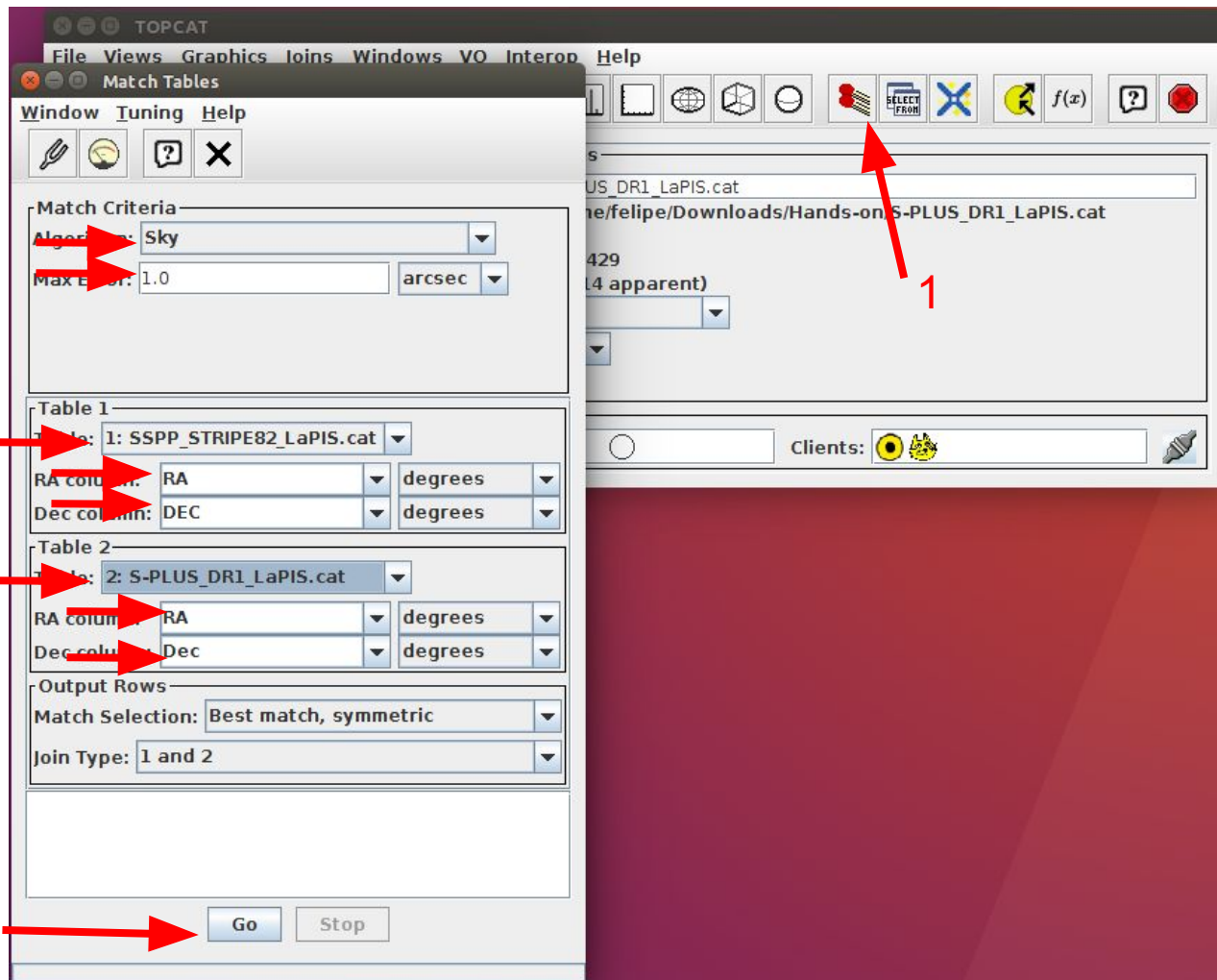
ID	Name	Size	Fraction	Expression
1	All	46726	100%	
2	color_cut	30274	65%	GR > 0.3 & GR < 0.85
3	giants	2913	6%	color_cut && (isinsid...

- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1

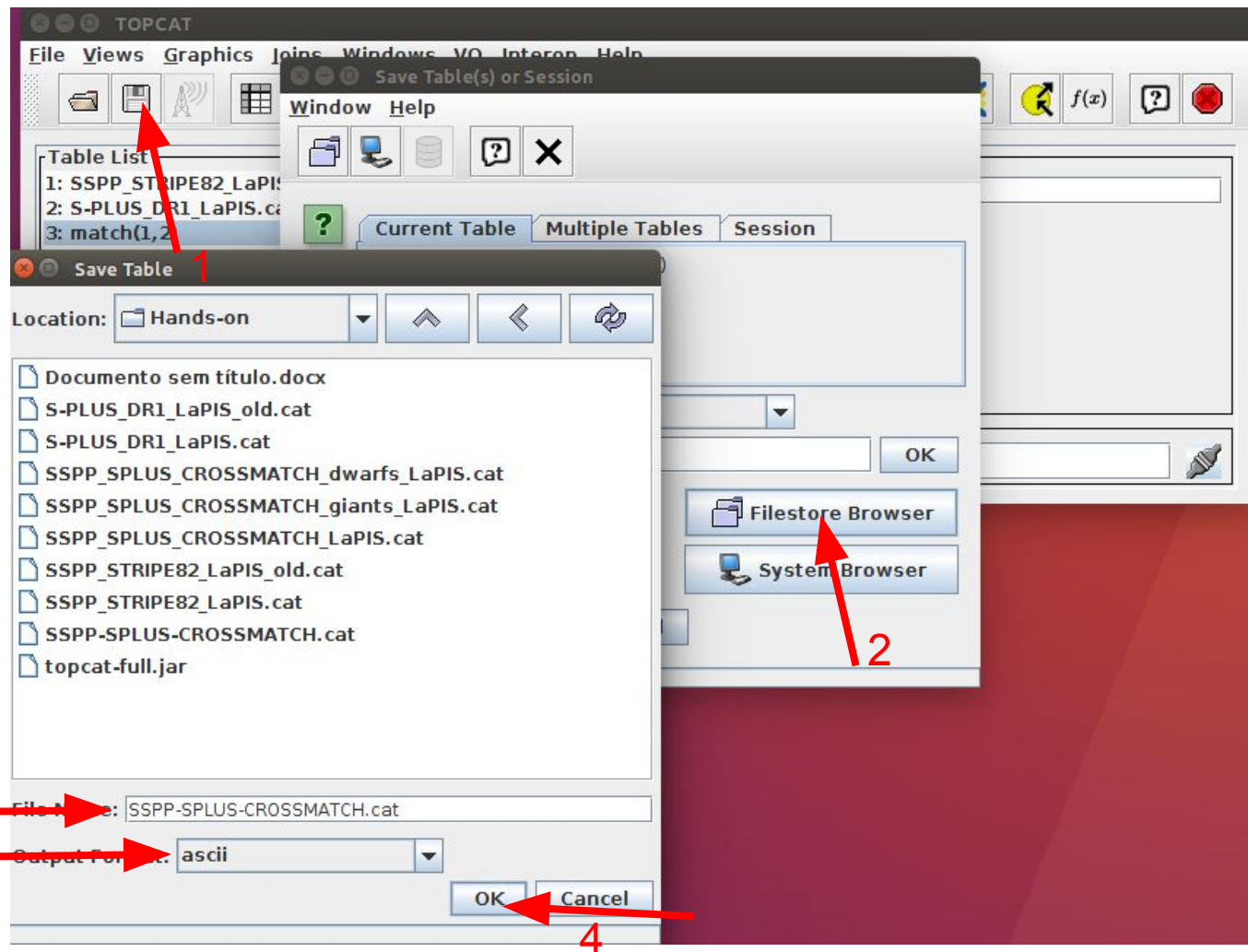


- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Crossmatch both catalogs

last

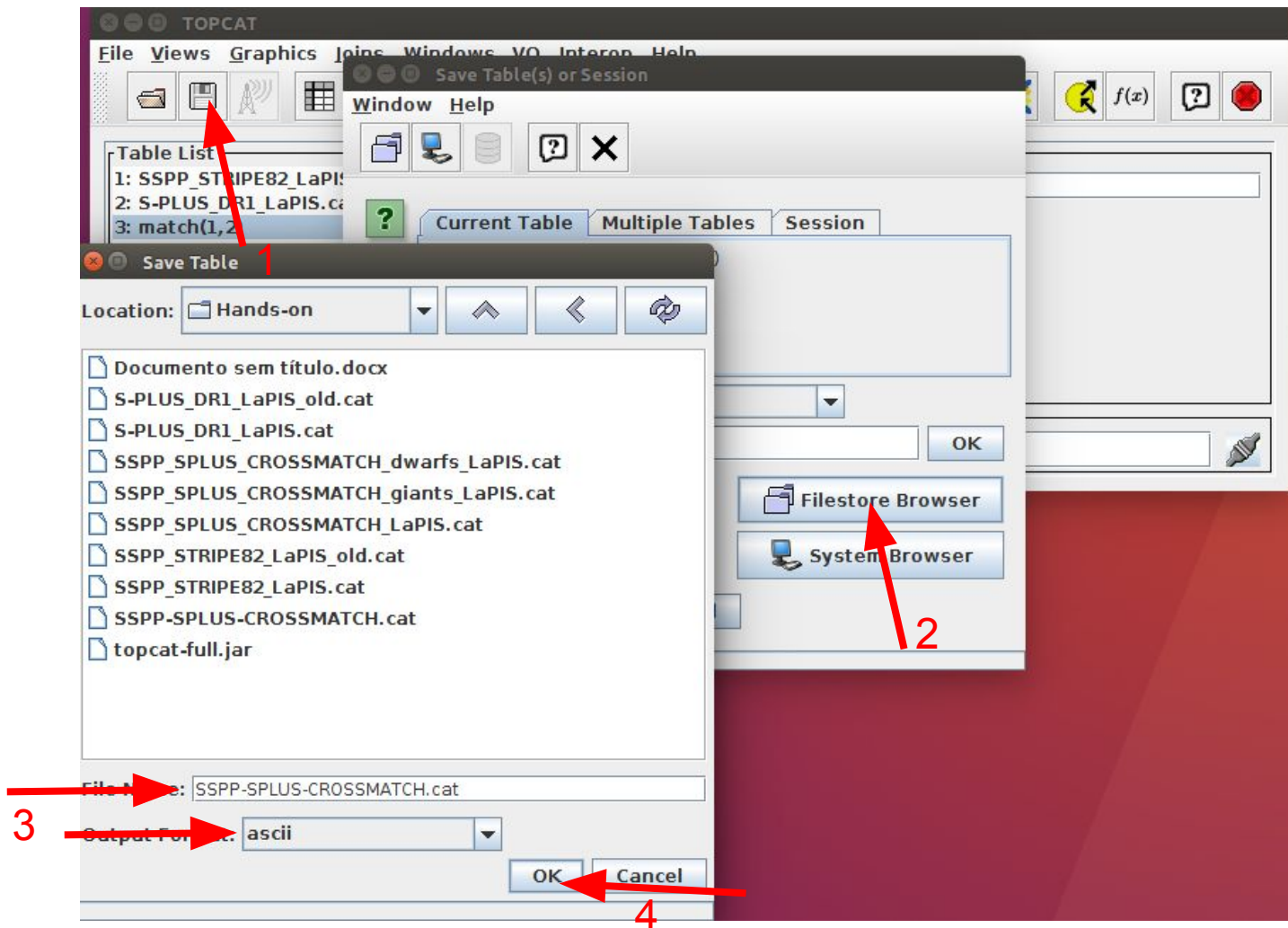


- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table



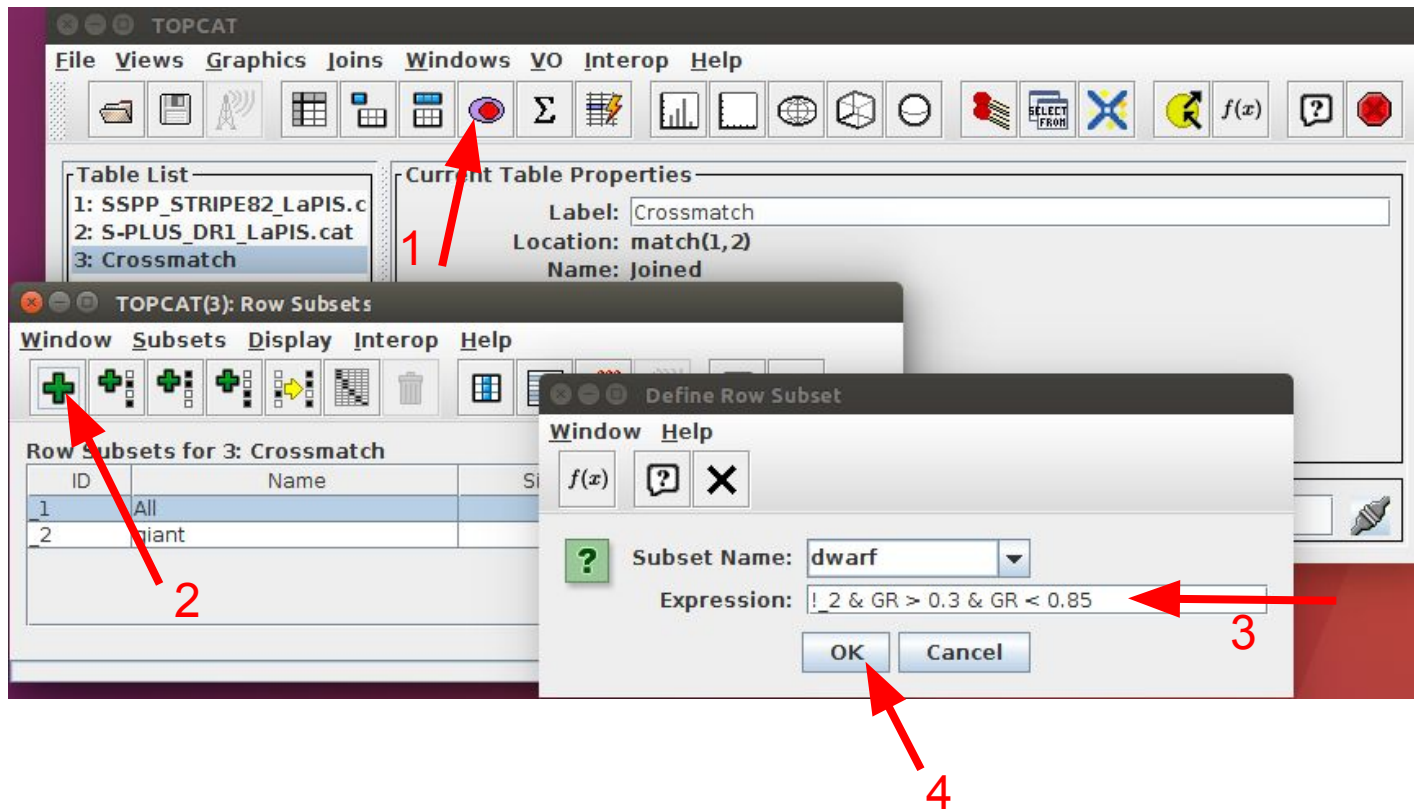
- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table

We're done. But let's make some plots.



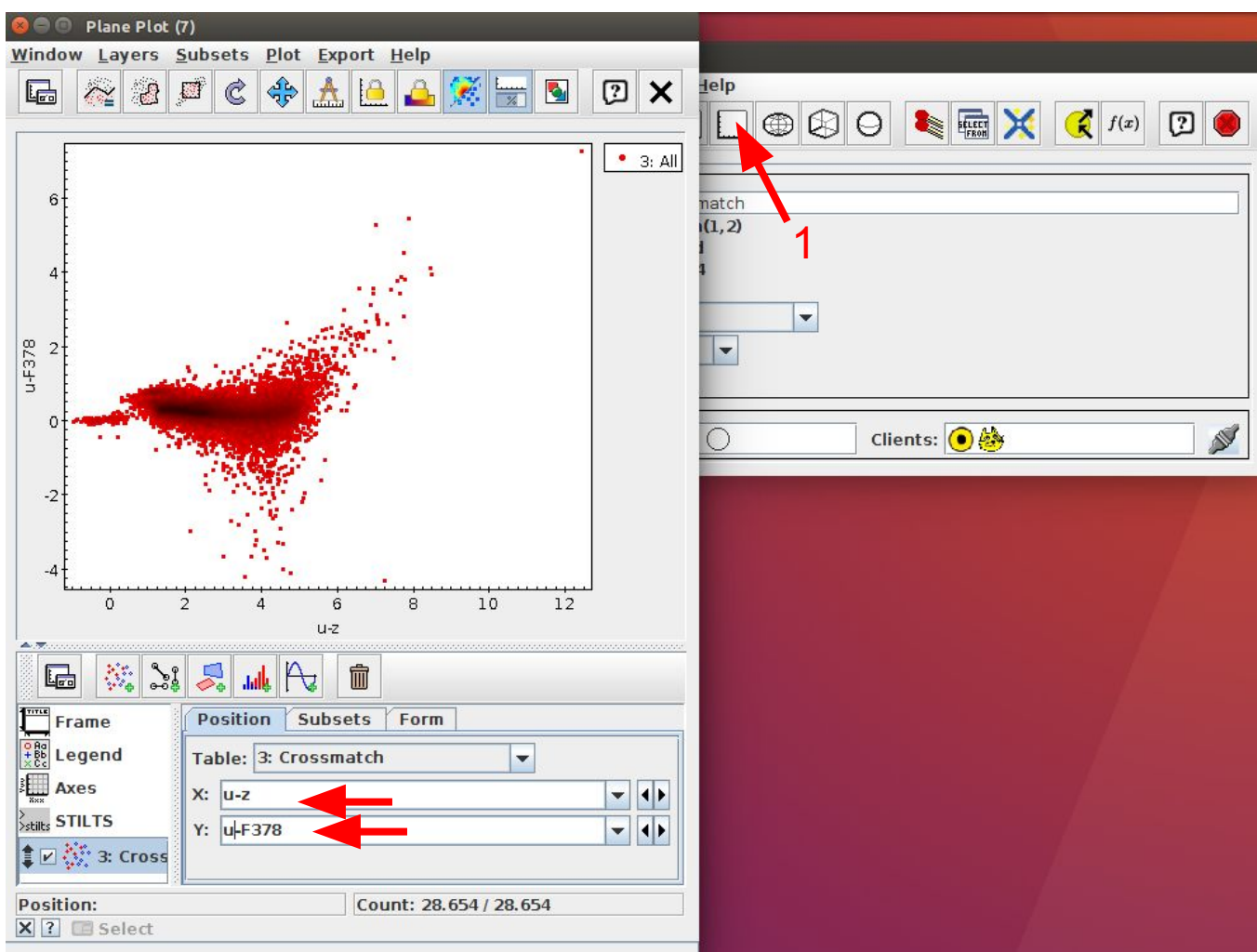
- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table

We're done. But let's make some plots.



- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table

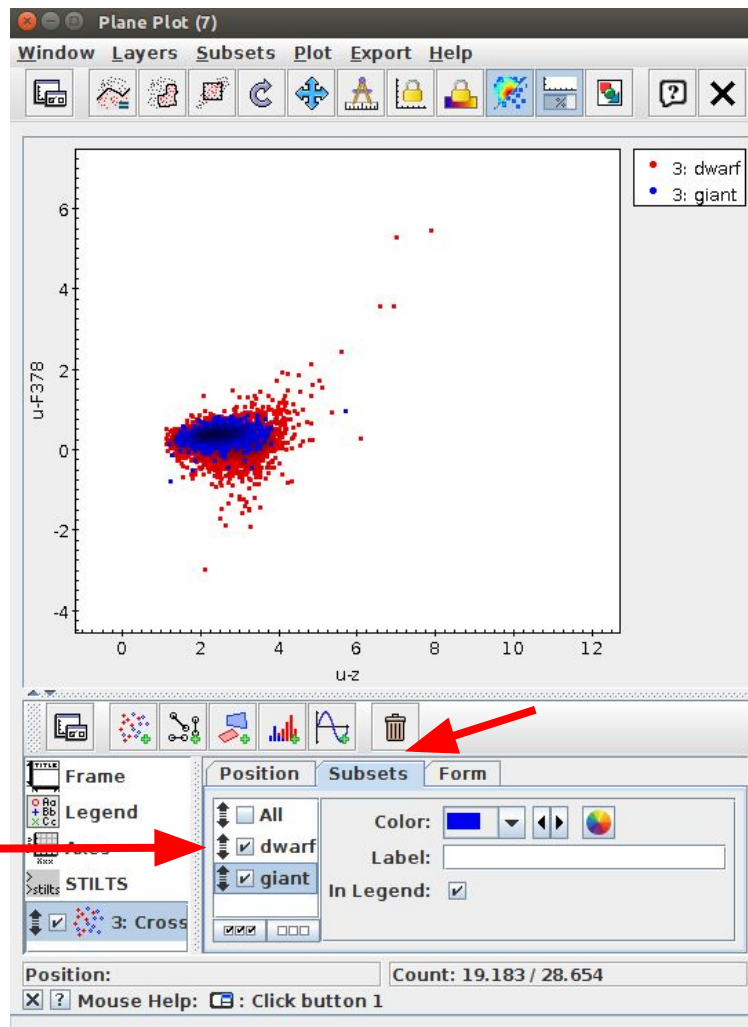
We're done. But let's make some plots.



- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table

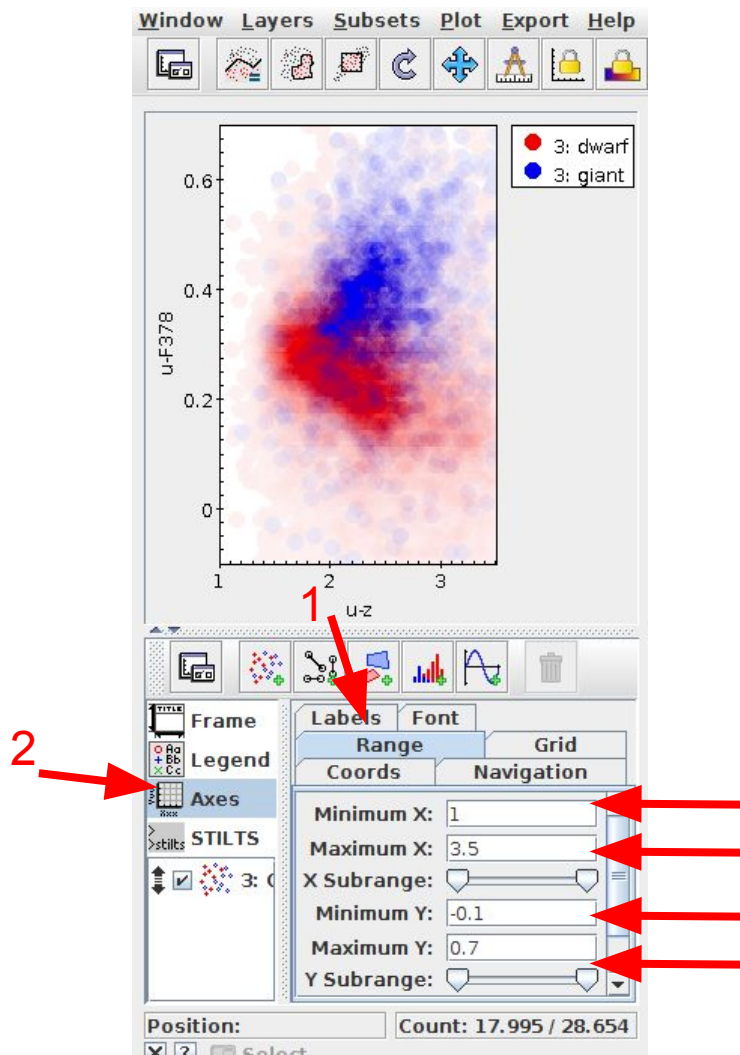
We're done. But let's make some plots.

select only dwarf
and giant. Change
their order.



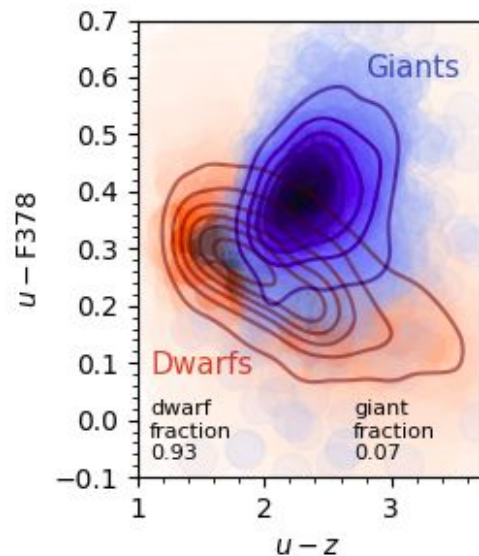
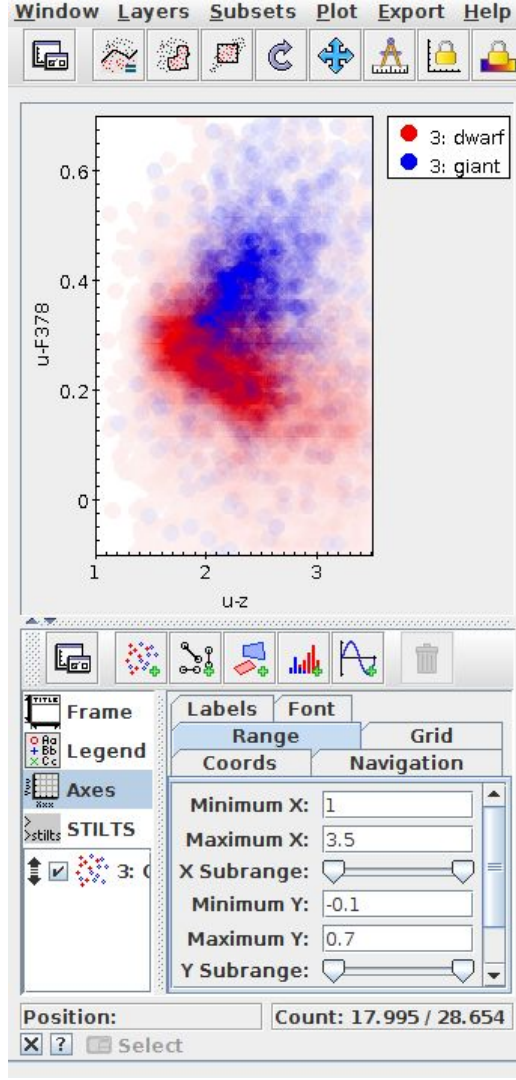
- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table

We're done. But let's make some plots.



- 1) Open Topcat
- 2) Load the SSPP sample
- 3) Plot spec data
- 4) Let's get rid of hot and cold dwarfs
- 5) Select the giants
- 6) Create a flag column
- 7) Load S-PLUS DR1
- 8) Save the table

We're done. But let's make some plots.



Activity

Use the `LaPIS_classifier_example.py` script to find out which classifier algorithm works better in this case.

