Stellar science with the S-PLUS survey

Using giant stars to look for Galactic substructures

A IX-LaPIS Lecture

La Plata, Argentina

INSTITUTO DE ASTRONOMIA, GEOFÍSICA E CIÊNCIAS ATMOSFÉRICAS

Felipe Almeida-Fernandes

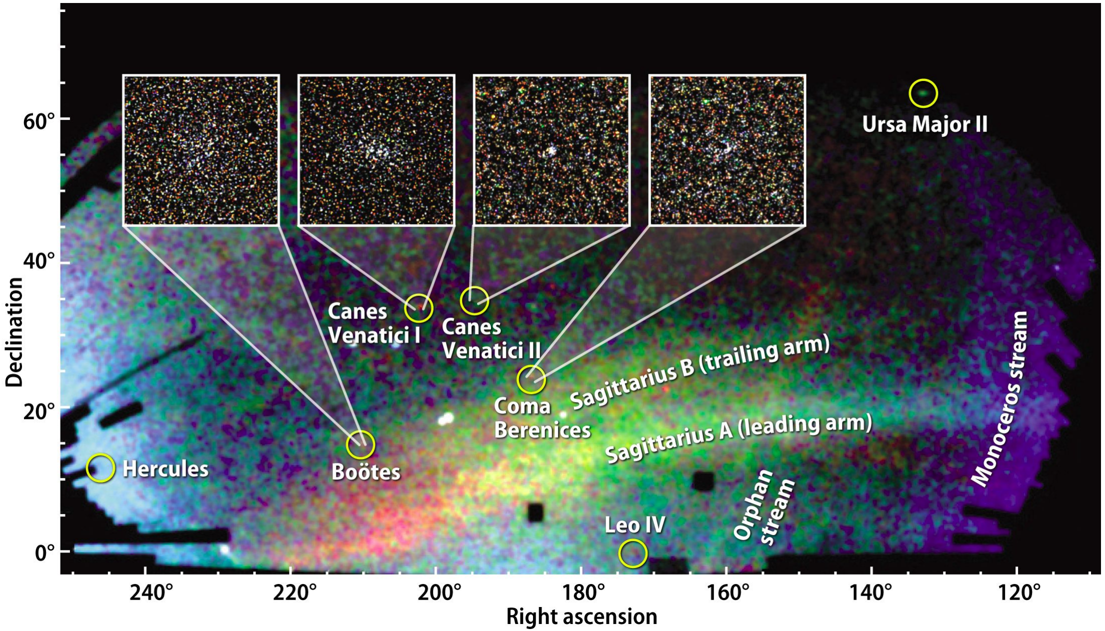
Post Doc at IAG/USP, São Paulo

Motivation

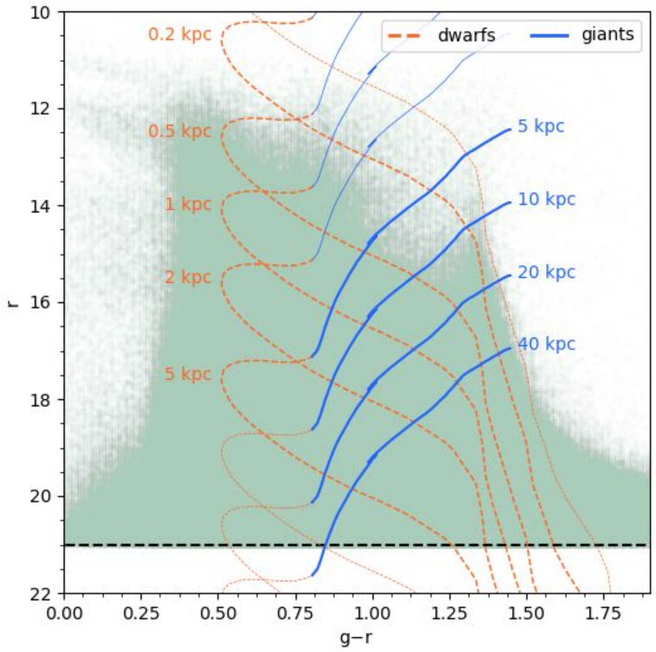
Our Galaxy is full of substructures and their nature and distribution are key for us to understand the Galaxy's evolution.

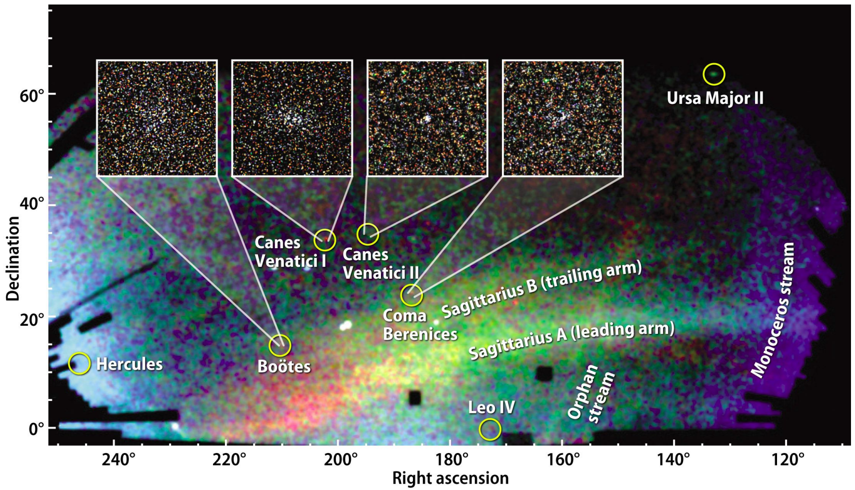
They can be very small and faint or even extend around most of the disk (stellar streams).

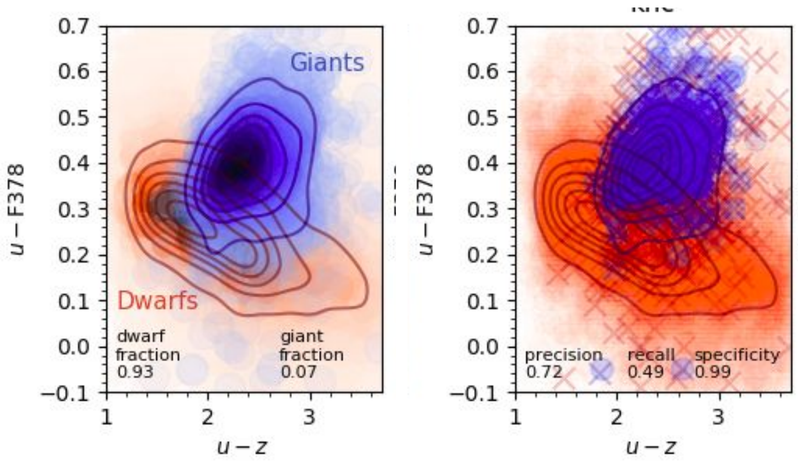
Credit: Vasily Belokurov, SDSS-II Collaboration



Spoiler Alert!







Summary

Lecture outline

A photometric search technique to identify giant stars in the context of S-PLUS.

- **The scientific background**
- **The photometric search technique**
- **An application of the technique**

Hands-on

1) Use Topcat to crossmatch S-PLUS DR1 and SDSS SSPP for the stripe-82 region.

2) Use the catalog we made, and a python script, to test different classification algorithms.

For this you will need:

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

Scientific Background

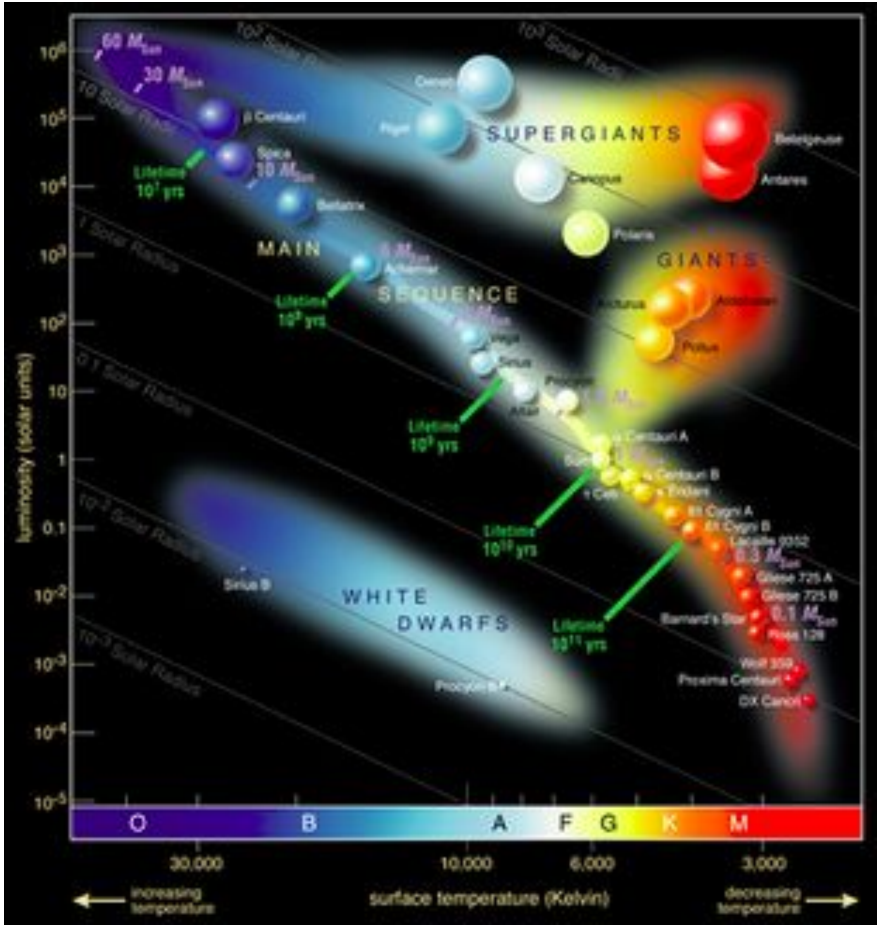
| Stellar evolution

Stars can have vastly different sizes, colors, and luminosities.

The differences come from:

- the initial stellar mass.
- the initial stellar composition.
- the stage of evolution (energy source).

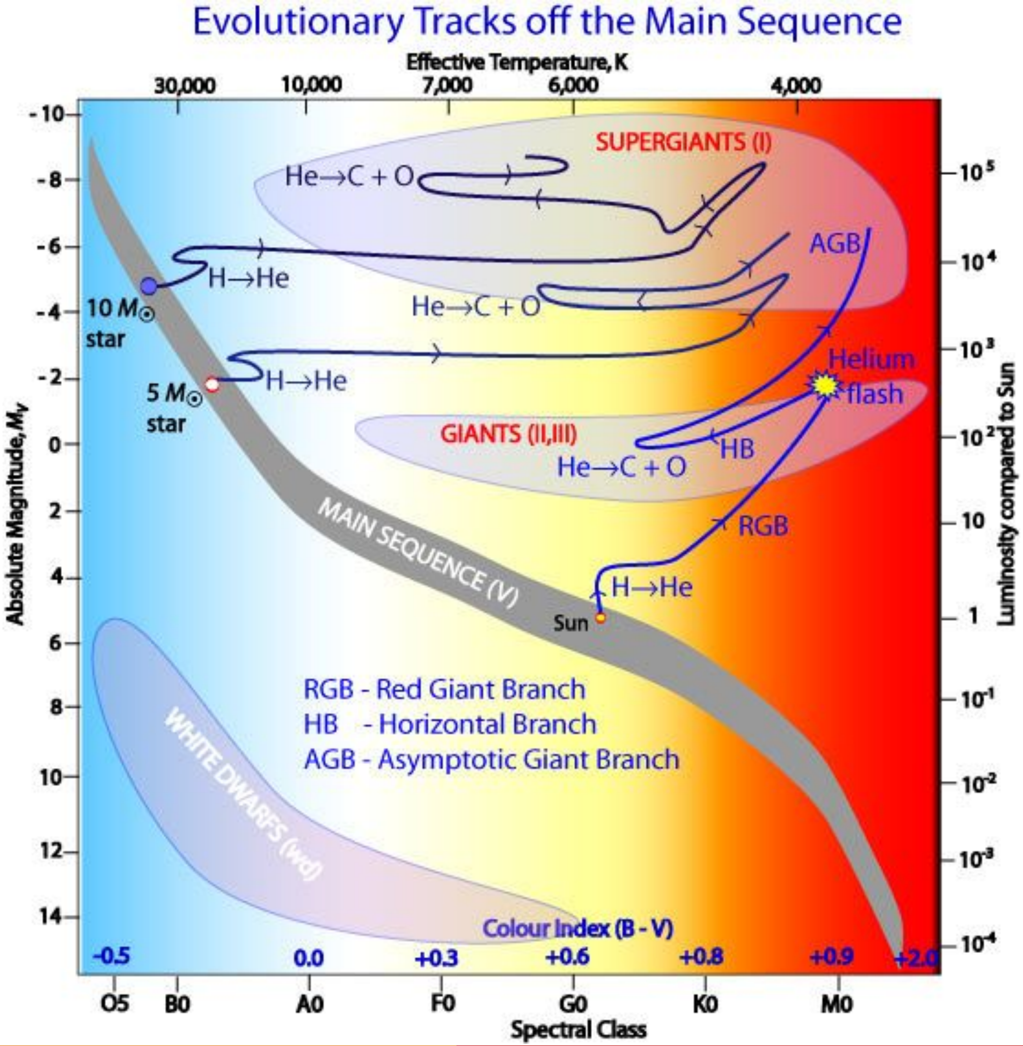
A lot can be known of a star by the region it occupies in a HR diagram



| Stellar evolution

The initial mass of the star is the main factor that dictates the evolutionary stages it will go through.

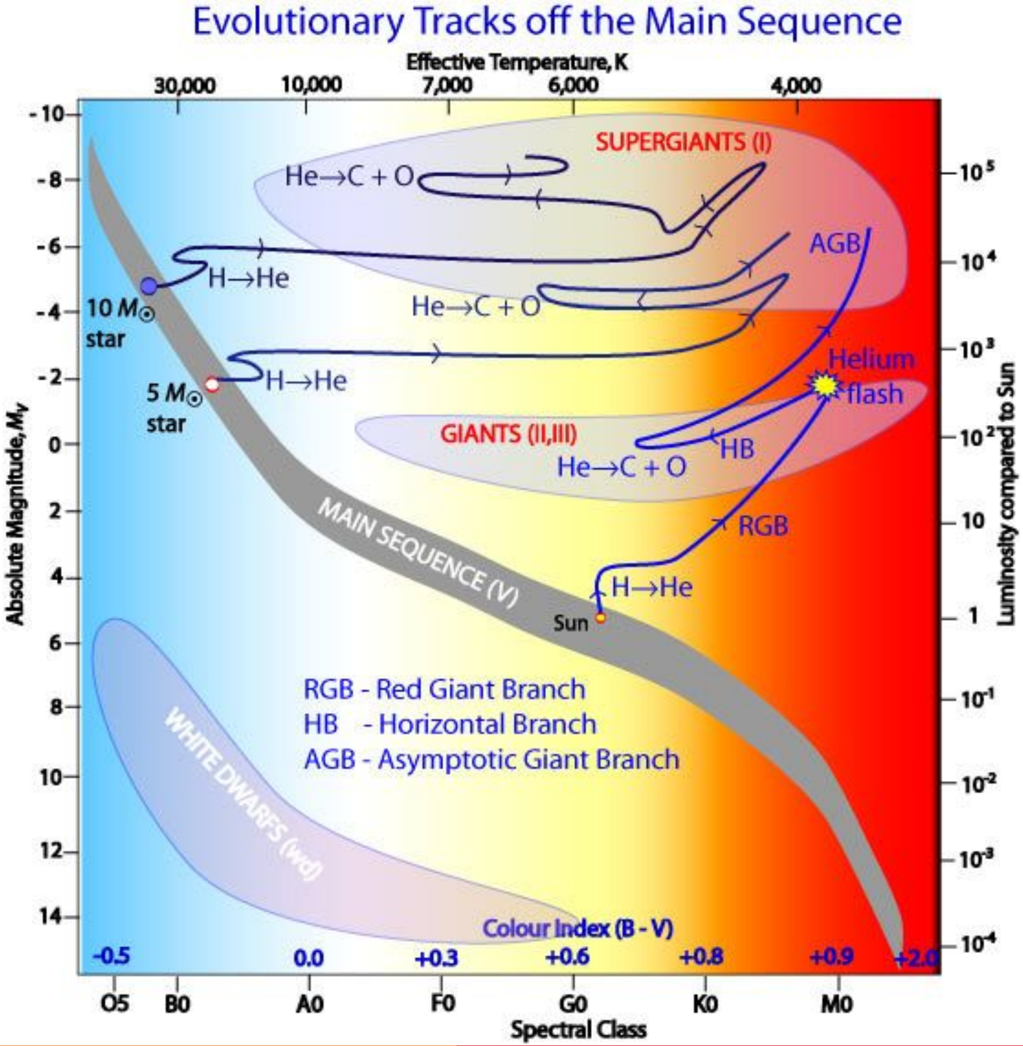
In general, for stars in similar evolutionary phases, the more massive stars are hotter and brighter.



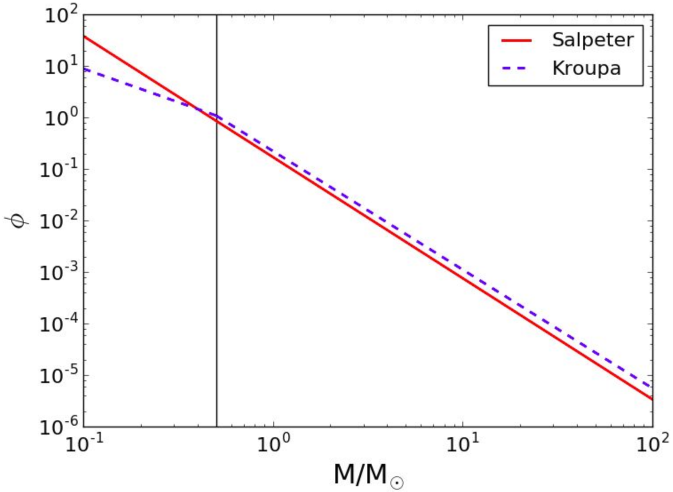
Stellar evolution

Three things determine how populated each region is:

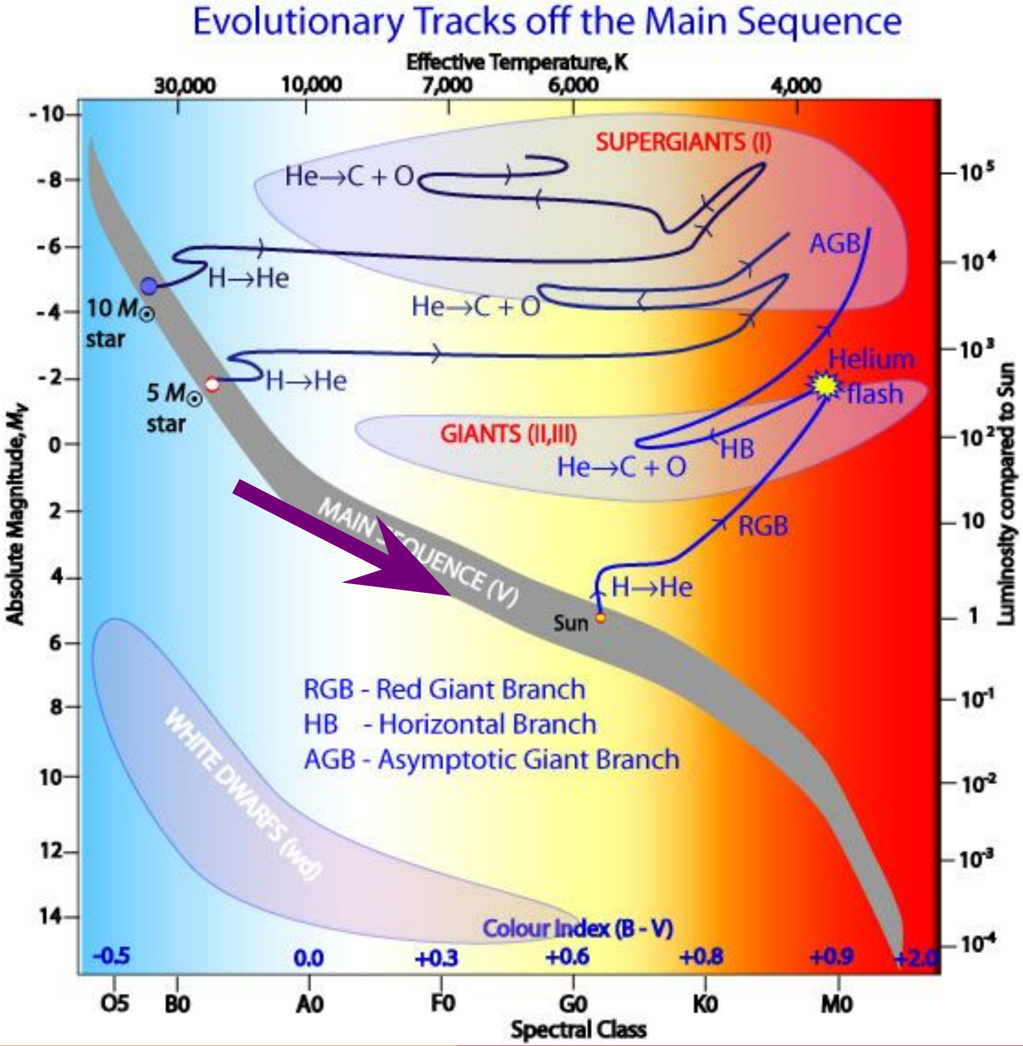
- The initial mass function
- The mass-life time relation
- The time scale of each evolutionary phase for each stellar mass



The initial mass function



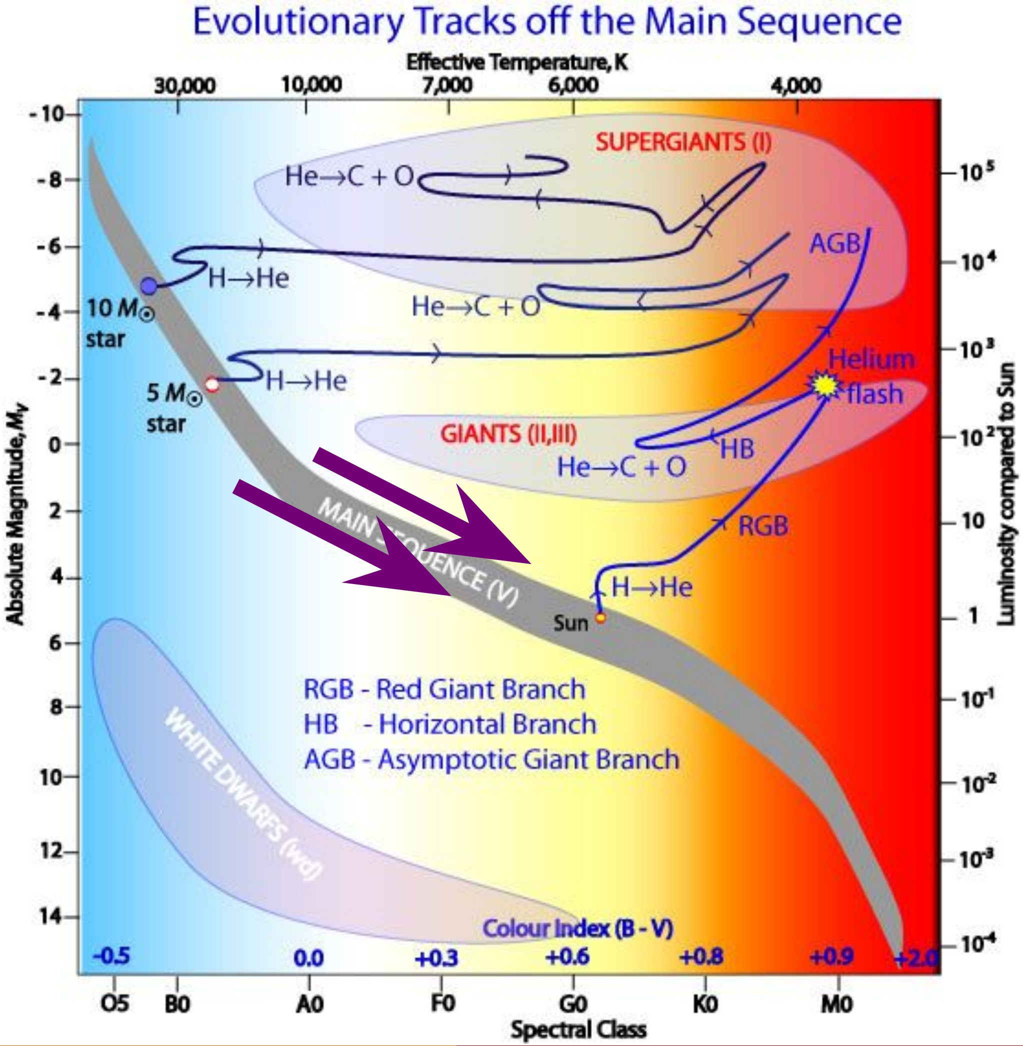
The process of star formation forms more low mass than high mass stars.



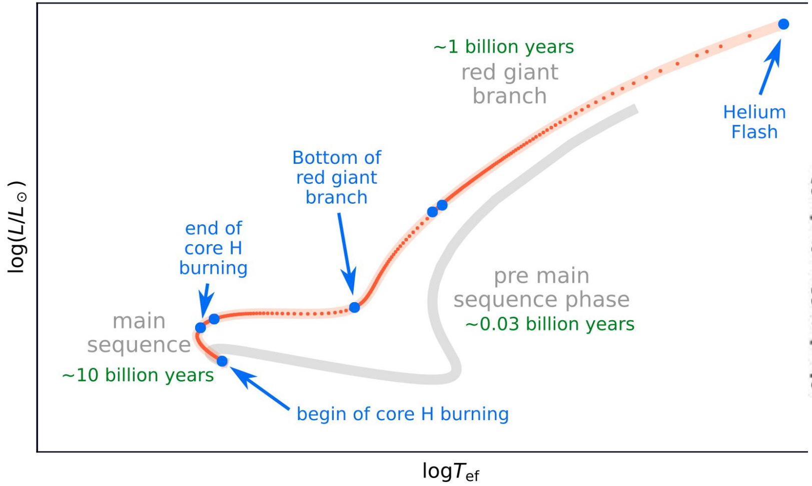
The mass-life time relation

Mass (solar masses)	Time (years)
60	3 million
30	11 million
10	32 million
3	370 million
1.5	3 billion
1	10 billion
0.1	1000s billions

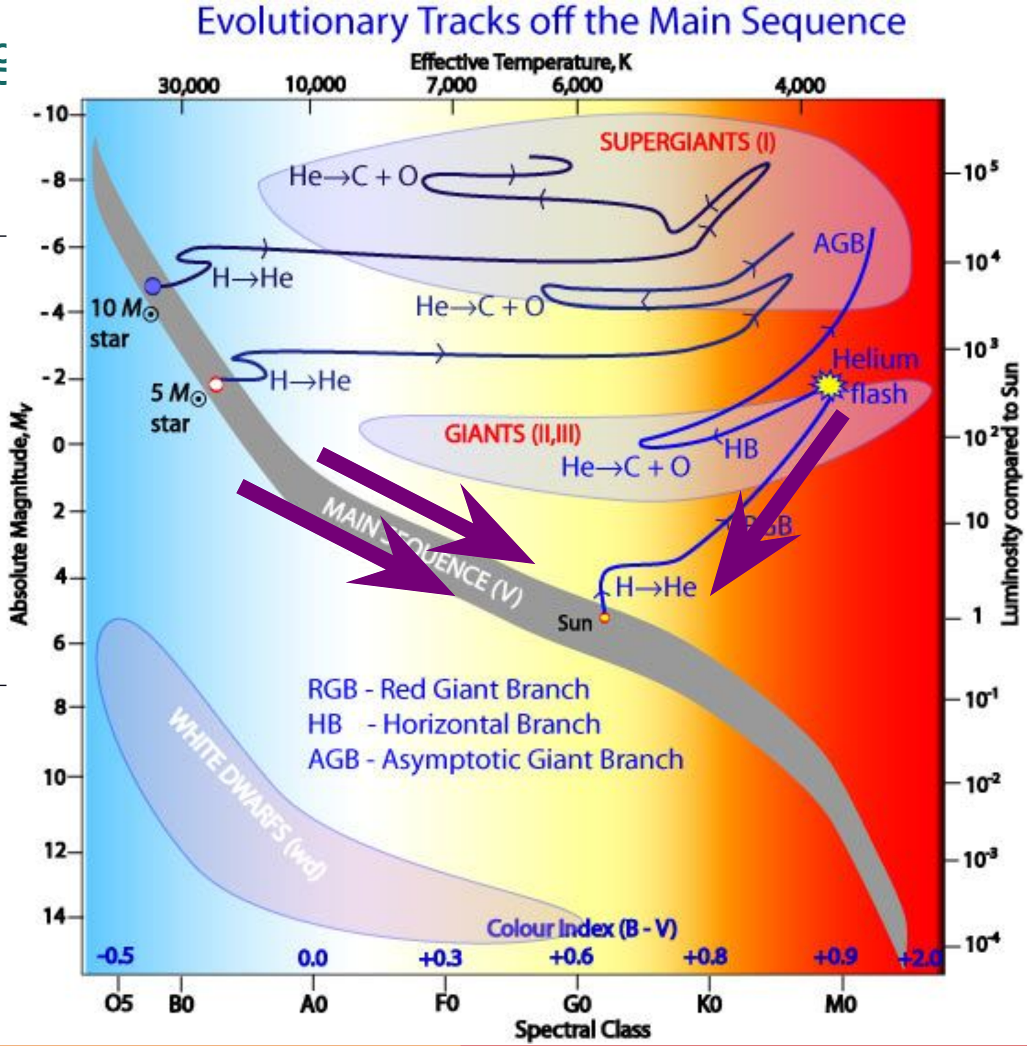
High mass stars evolve much faster than low mass stars.



Time scale of evolutionary stag



Stars spend most of its life in the Main Sequence (burning Hydrogen in the core)

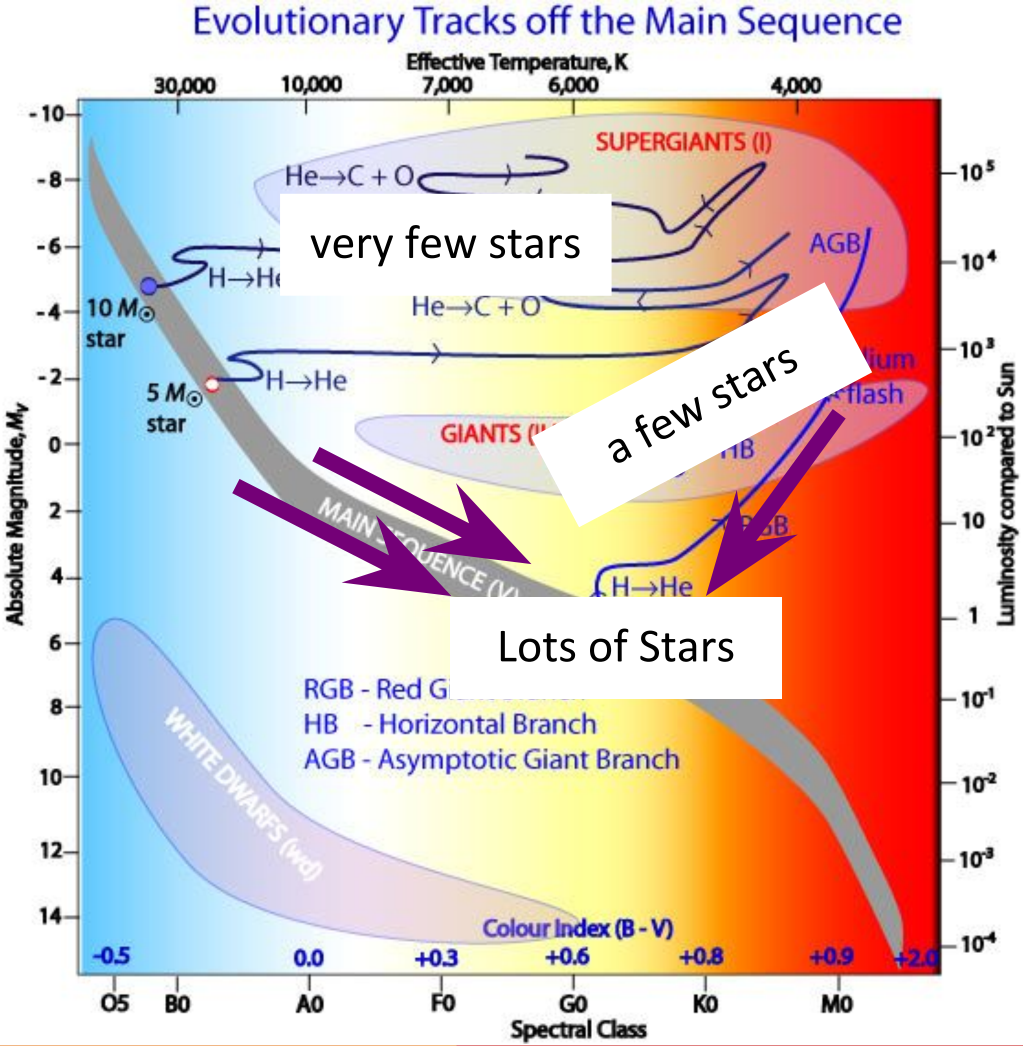


| Stellar evolution

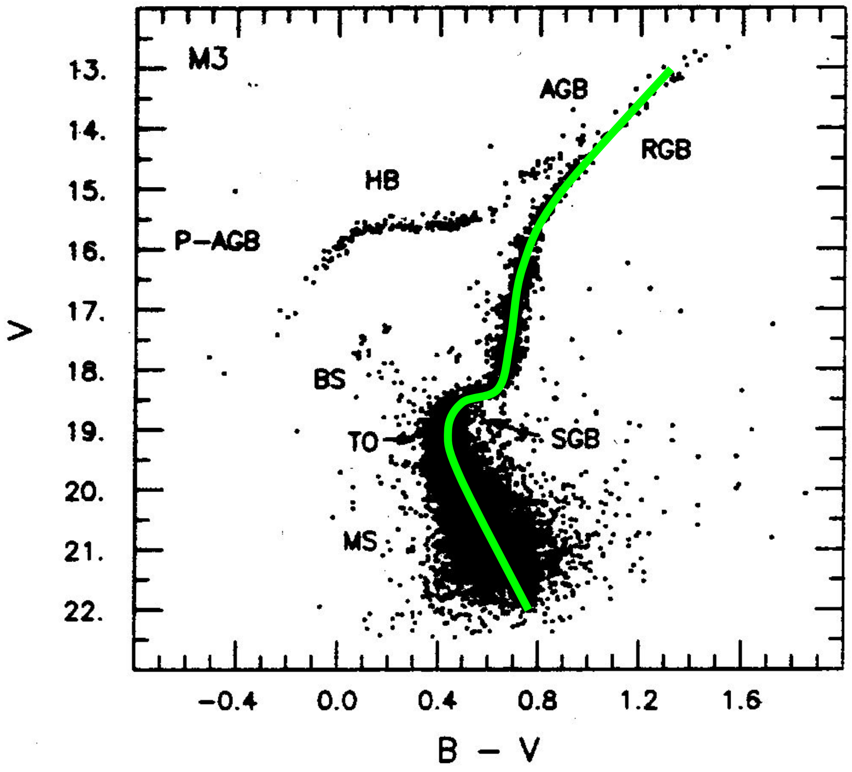
Three things determine how populated each region is:

- The initial mass function
- The mass-life time relation
- The time scale of each evolutionary phase for each stellar mass

In a given stellar population, the amount of dwarf stars (main sequence stars) is greater than the amount of giant stars.



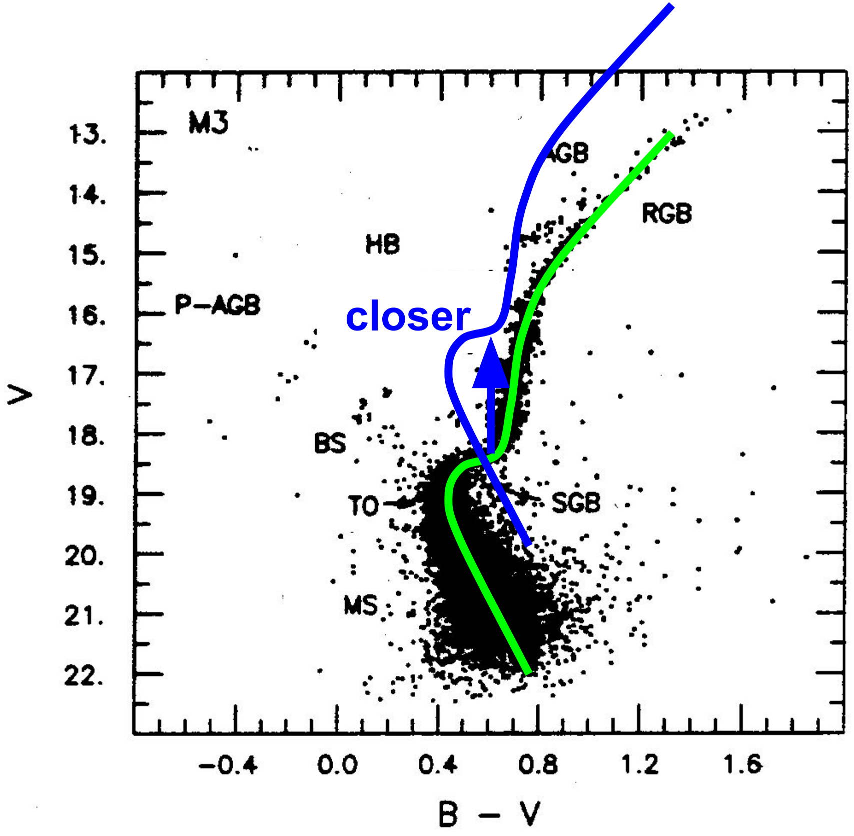
Example: the M3 globular cluster



The role of distance

The brightness of a star is determined not only by its own properties, but also by their distance from the Sun.

If a stellar population is closer to the Sun, all of its stars will be brighter.

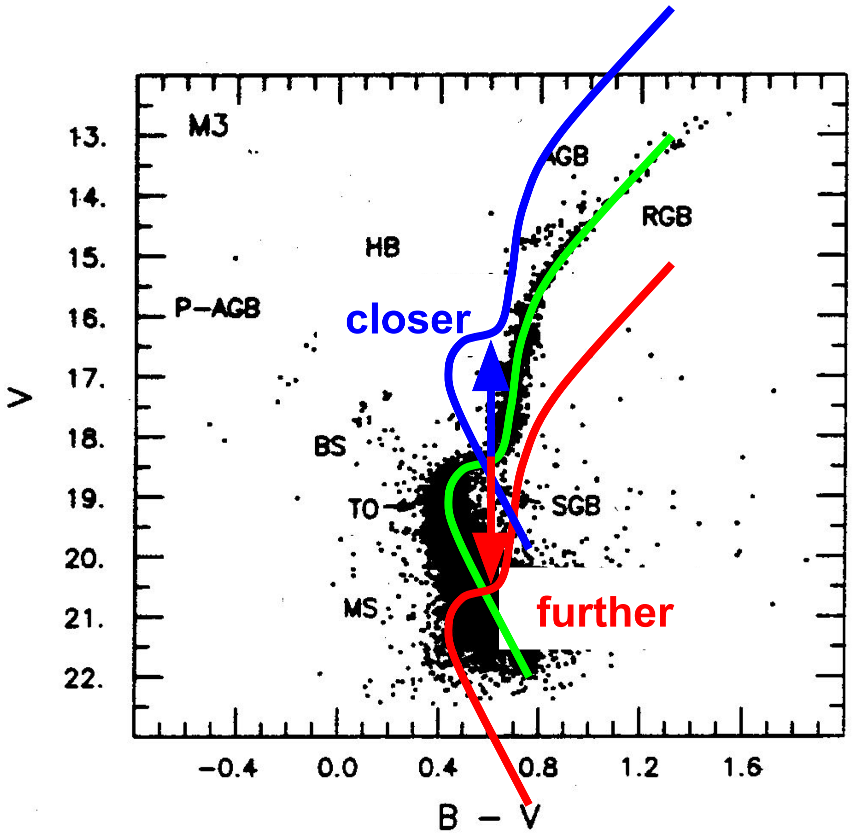


The role of distance

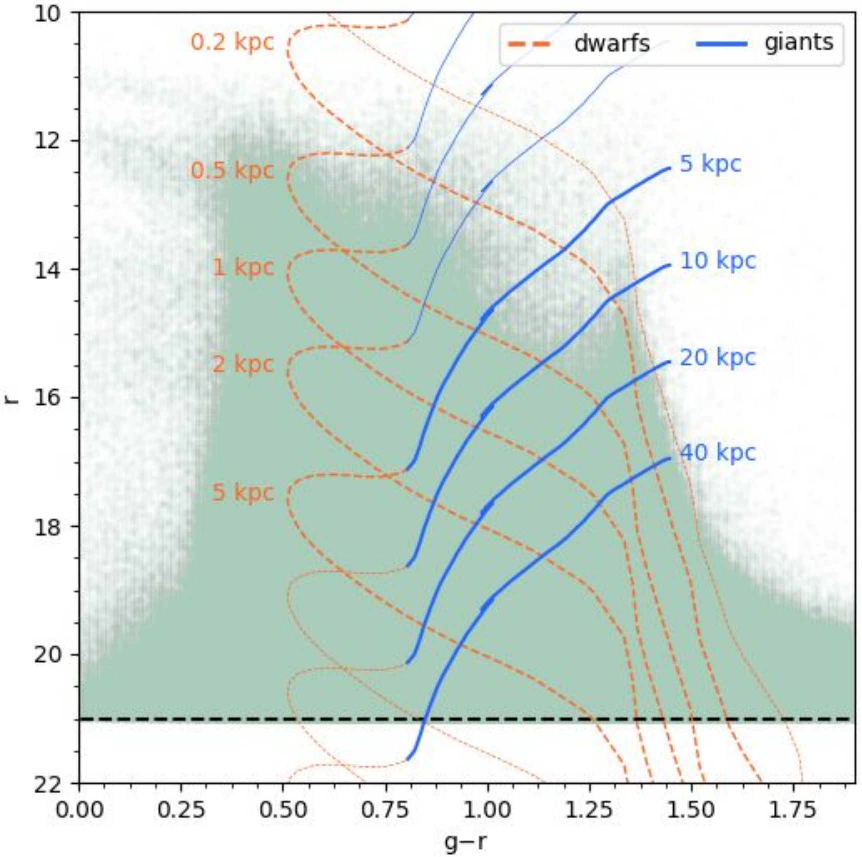
The brightness of a star is determined not only by its own properties, but also by their distance from the Sun.

If a stellar population is closer to the Sun, all of its stars will be brighter.

And if its further away, all of its stars will be fainter.



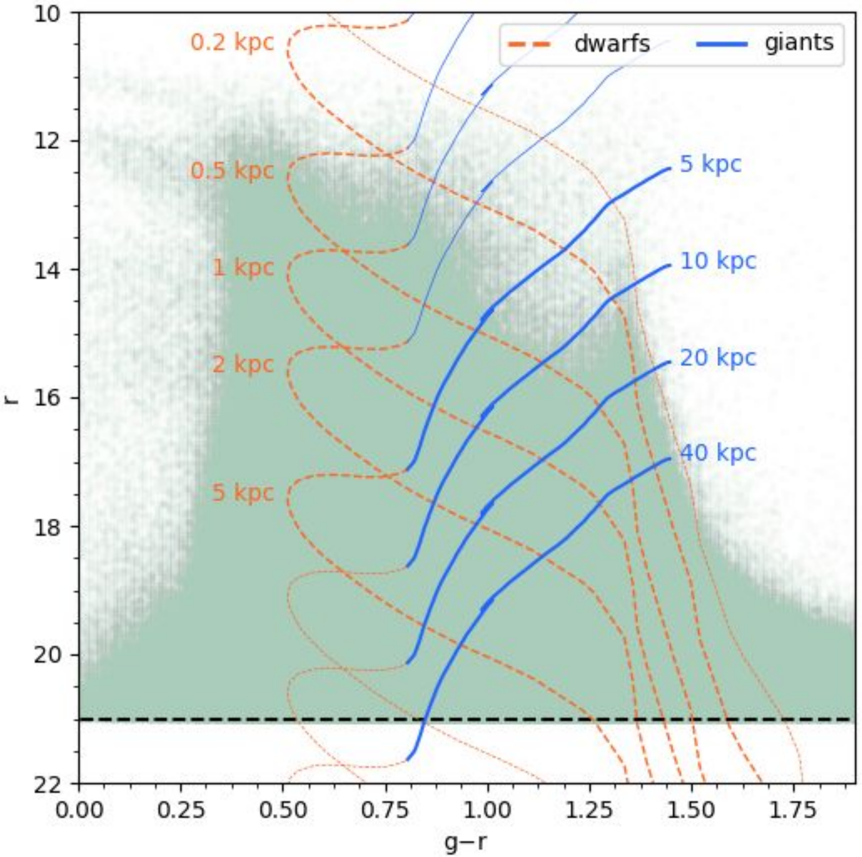
Population Tracer



A color-magnitude diagram is a combination of nearby dwarf stars and far away giants.

Due to the initial mass function, and the lifetime spent in each phase of stellar evolution, most of the stars in a cmd (~>95%) are dwarfs.

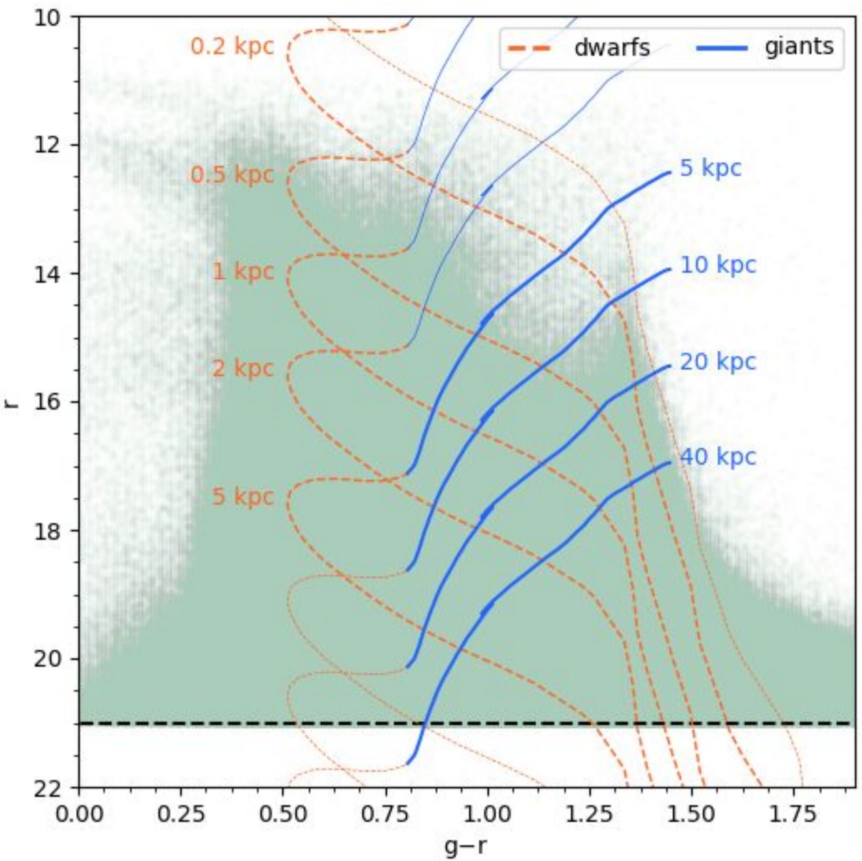
| Population Tracer



The majority of the stellar overdensities are located more than 10 kpc away from the Sun.

Only the giant stars of distant stellar populations can be observed by photometric surveys with depths <~20.

| Population Tracer



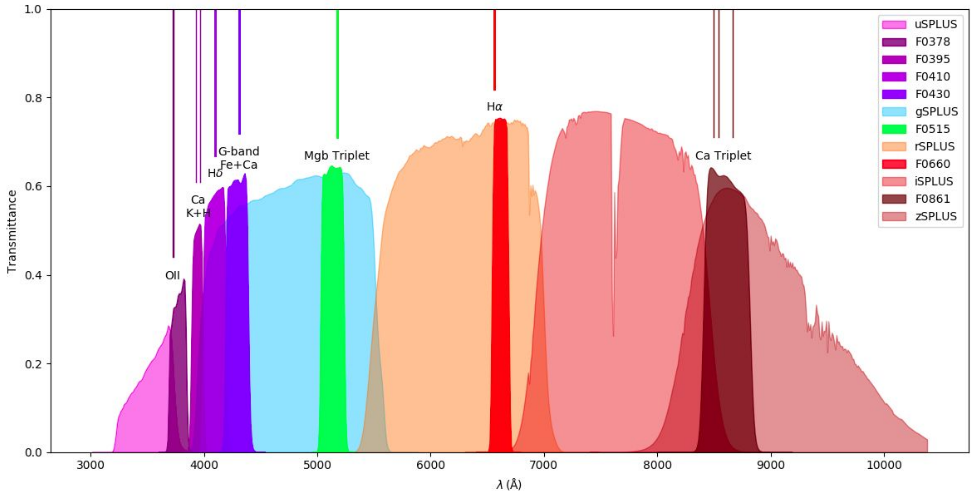
In a given stellar population, the amount of dwarf stars (main sequence stars) is greater than the amount of giant stars.

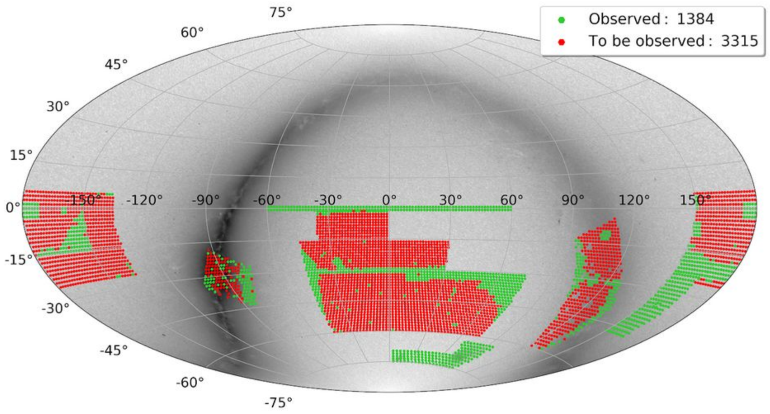
Only the giant stars of distant stellar populations can be observed by photometric surveys with depths <~20.

In order to use this data to look for stellar populations at these greater distances, we must be able to distinguish the giant stars.

Advantages of using S-PLUS

The S-PLUS narrow band filters are designed to measure the fluxes in some key stellar spectral features like the Mgb triplet and Halpha.





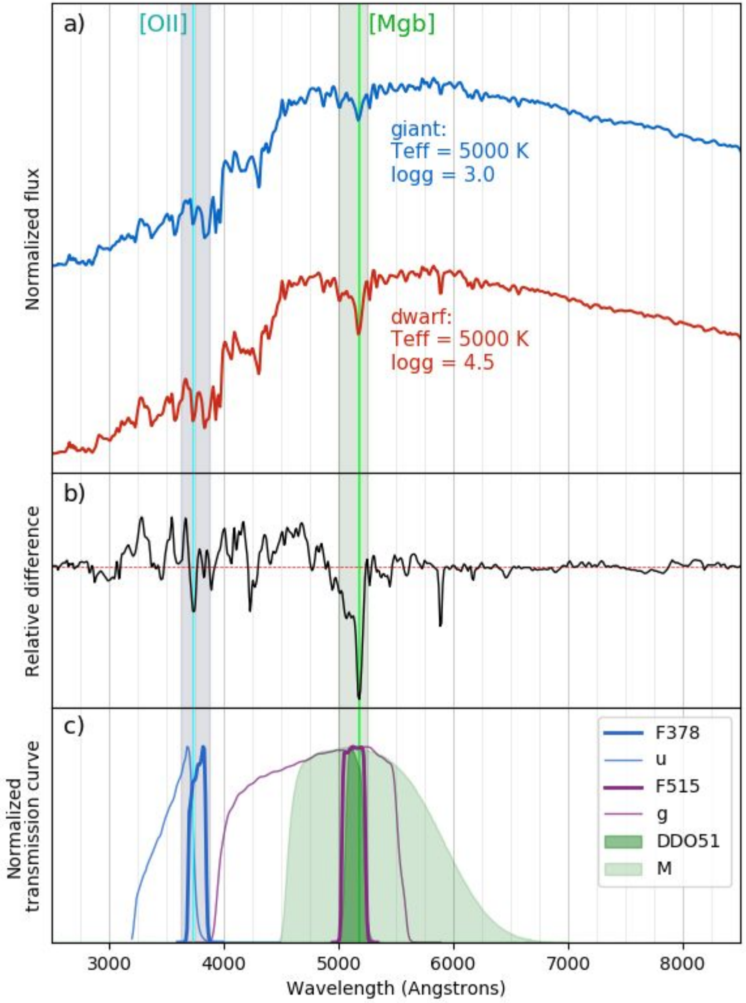
The very great sky coverage will allow us to look for overdensities in regions not yet explored. It also allows us to search and study extended structures, like streams.

Photometric Search Technique

| The photometric search technique

This technique relies on a spectral feature that is sensitive to logg, and strong enough to be detected by a narrow band filter.

It was used by APOGEE to select the Giant stars for spectral observation. But in that case, they used the Washington filter system + the DDO51 narrow band.

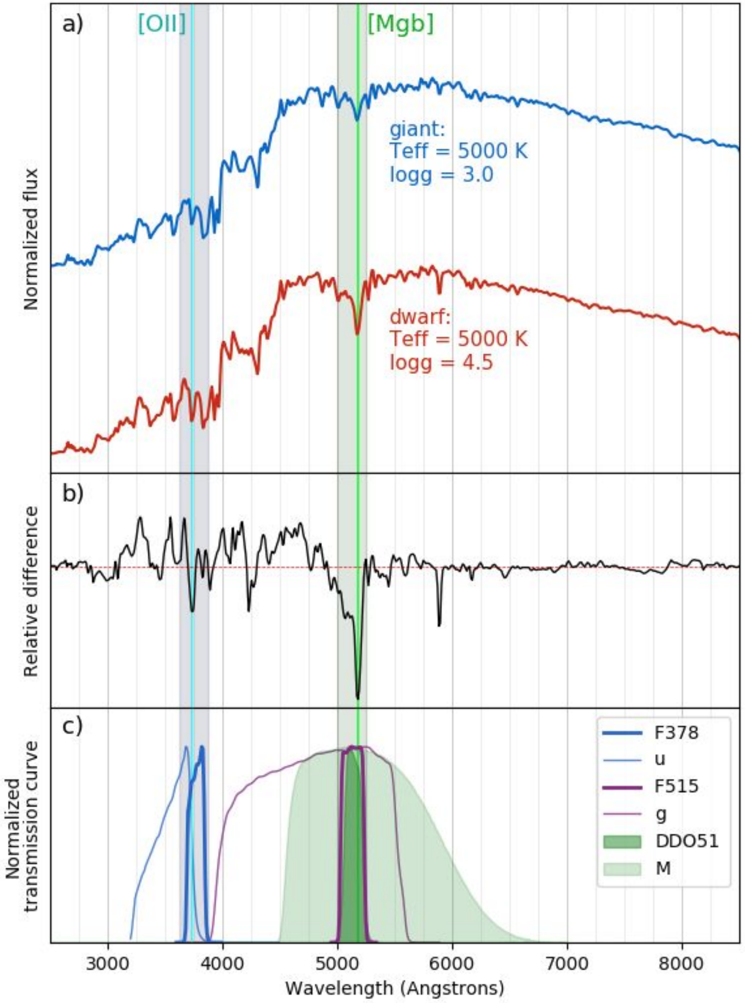


| The photometric search technique

On the upper panel of the plot, we see the synthetic spectral model of a giant and a dwarf star (Coelho, 2014)

Overall, the spectra seems to be identical.

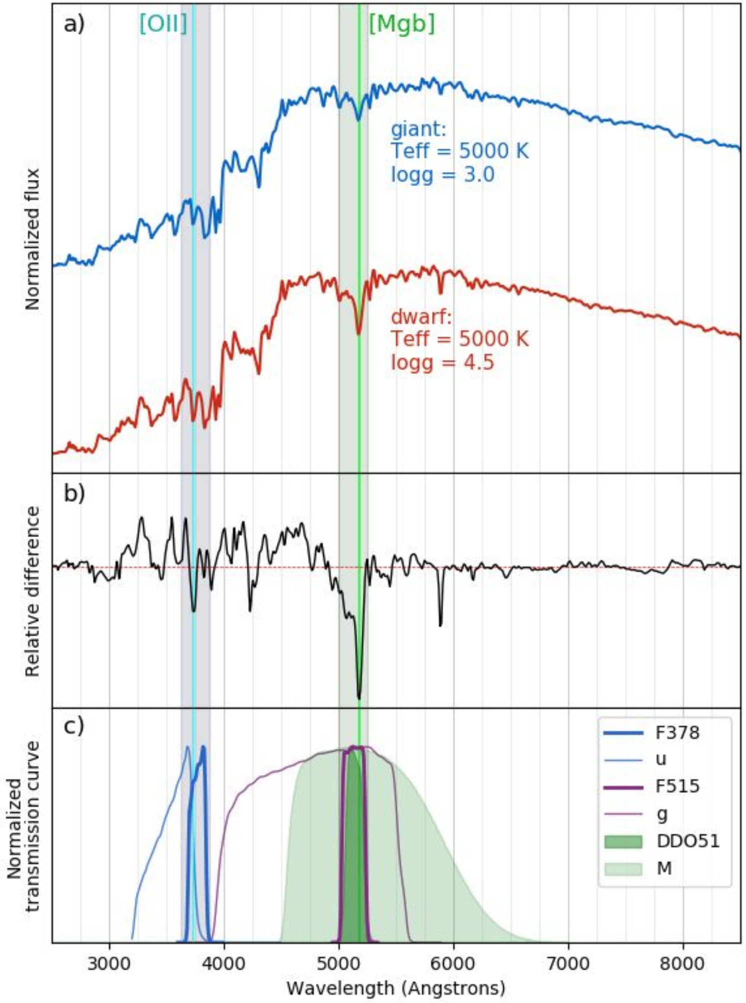
But looking at the relative difference (b), we can see how the Mgb triplet is sensitive to logg.



| The photometric search technique

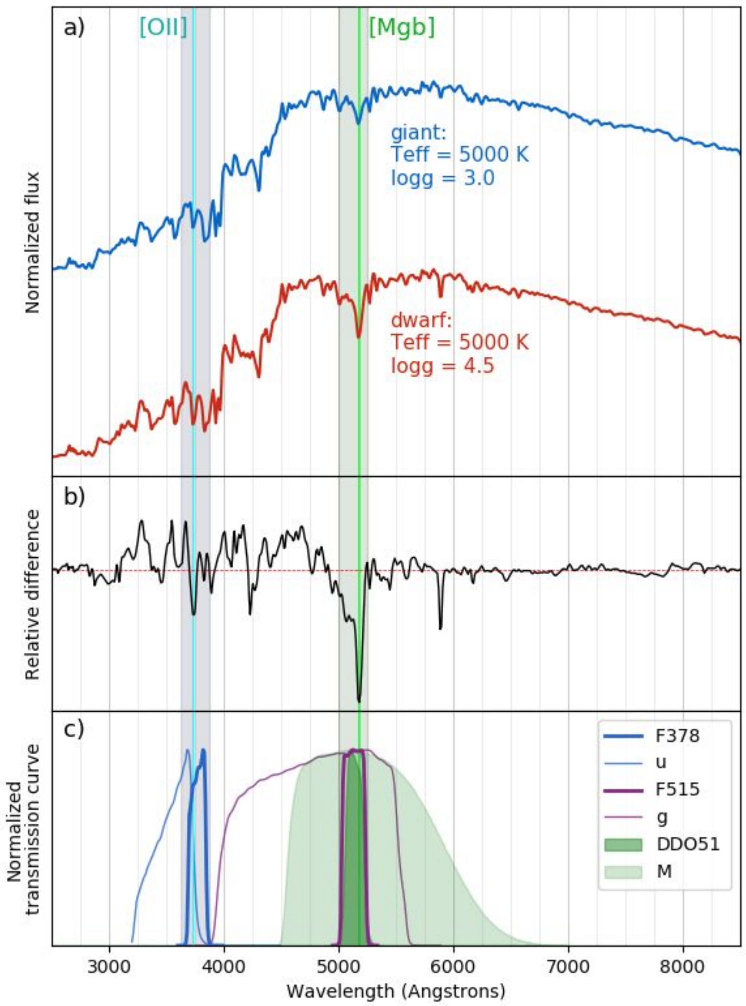
In S-PLUS, the Mgb line strength is measured by the g − F515 color.

The spectral region observed by the u and F378 filters is also sensitive to logg, but the relation is not so straightforward.



| The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.



$$m_x = -2.5 \log_{10}\left(\frac{F_x}{F_{x,0}}\right)$$

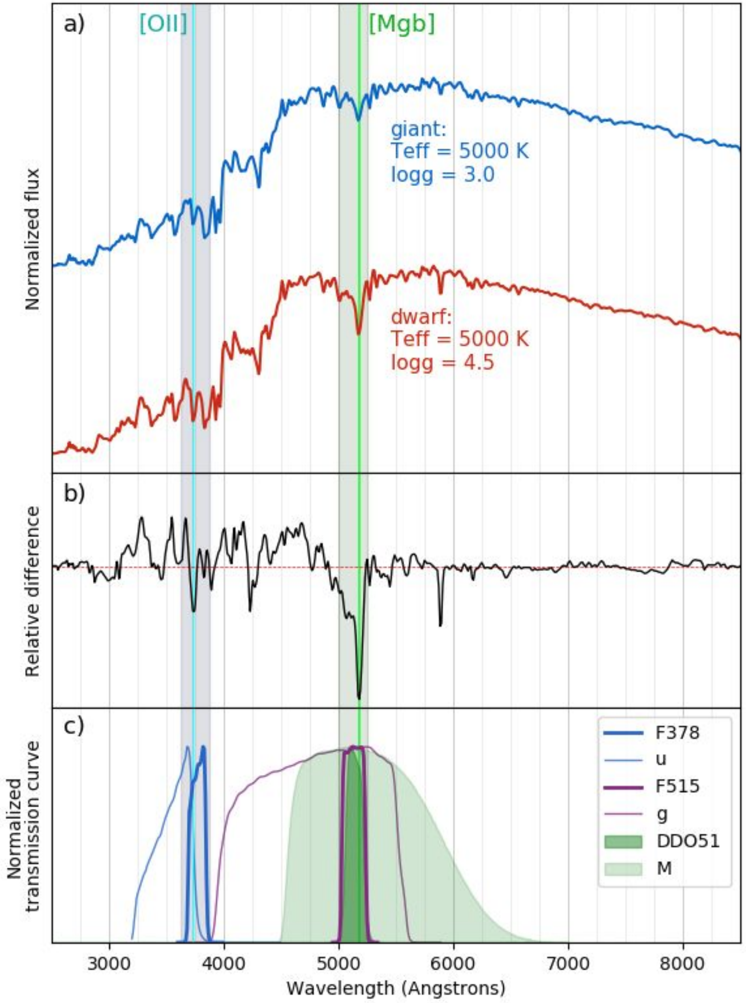
| The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

Giants

less depth of Mgb

$$m_x = -2.5 \log_{10}\left(\frac{F_x}{F_{x,0}}\right)$$

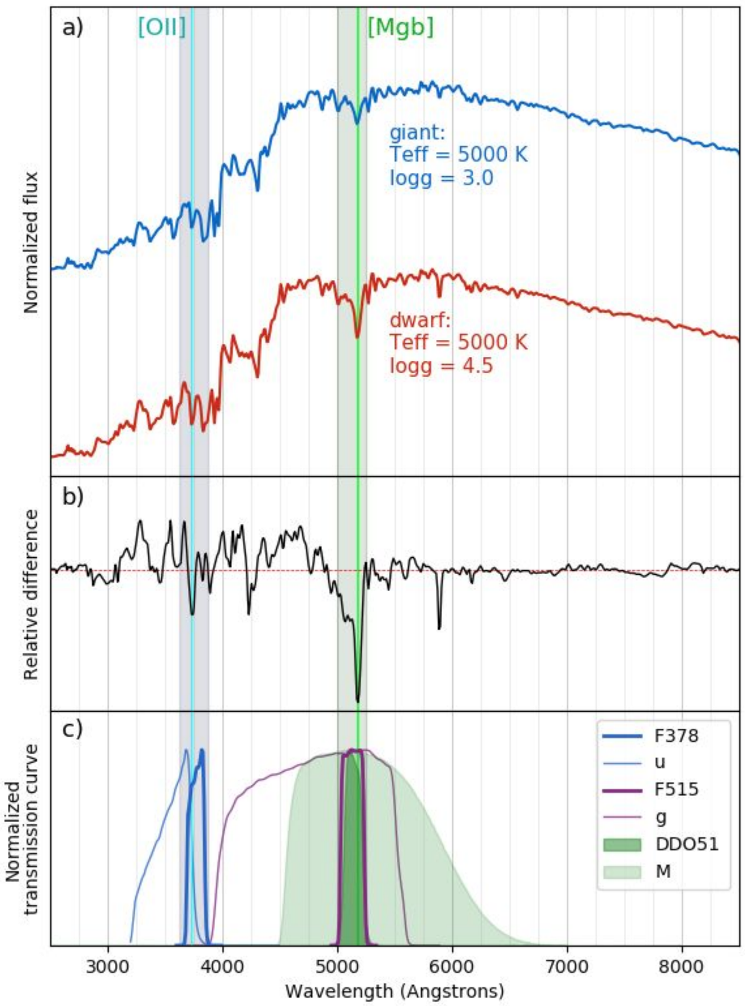


| The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

Giants | less depth of Mgb | higher flux in F515 filter

$$m_x = -2.5\log_{10}\left(\frac{F_x}{F_{x,0}}\right)$$

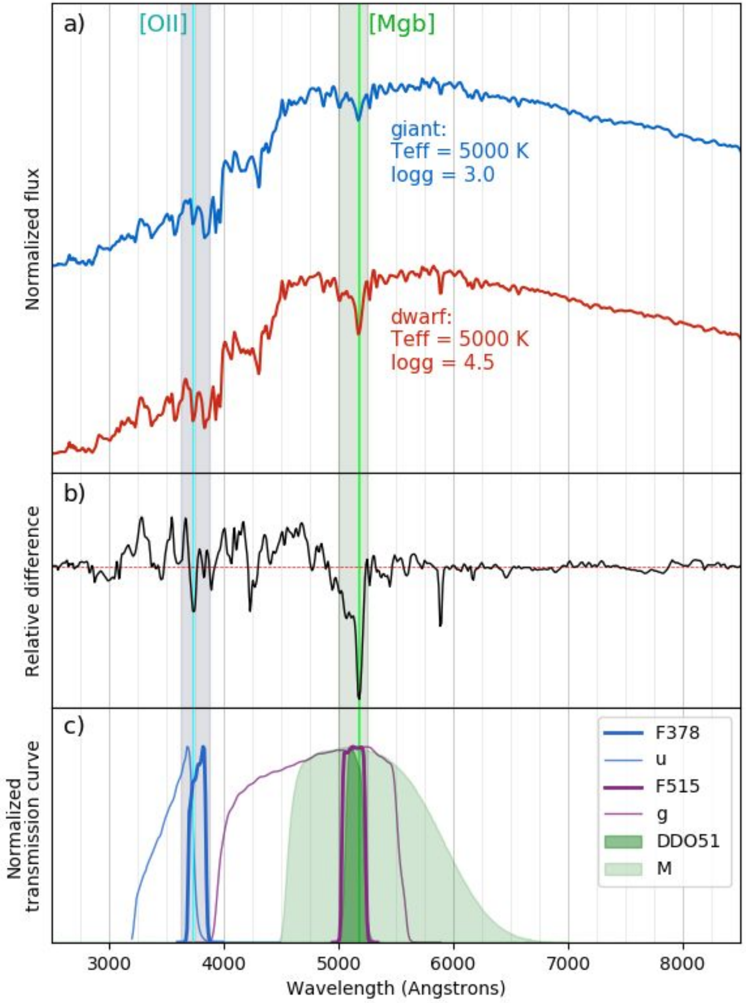


| The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

Giants	less depth of Mgb	higher flux in F515 filter	lower F515 magnitude

$$m_x = -2.5 \log_{10}\left(\frac{F_x}{F_{x,0}}\right)$$



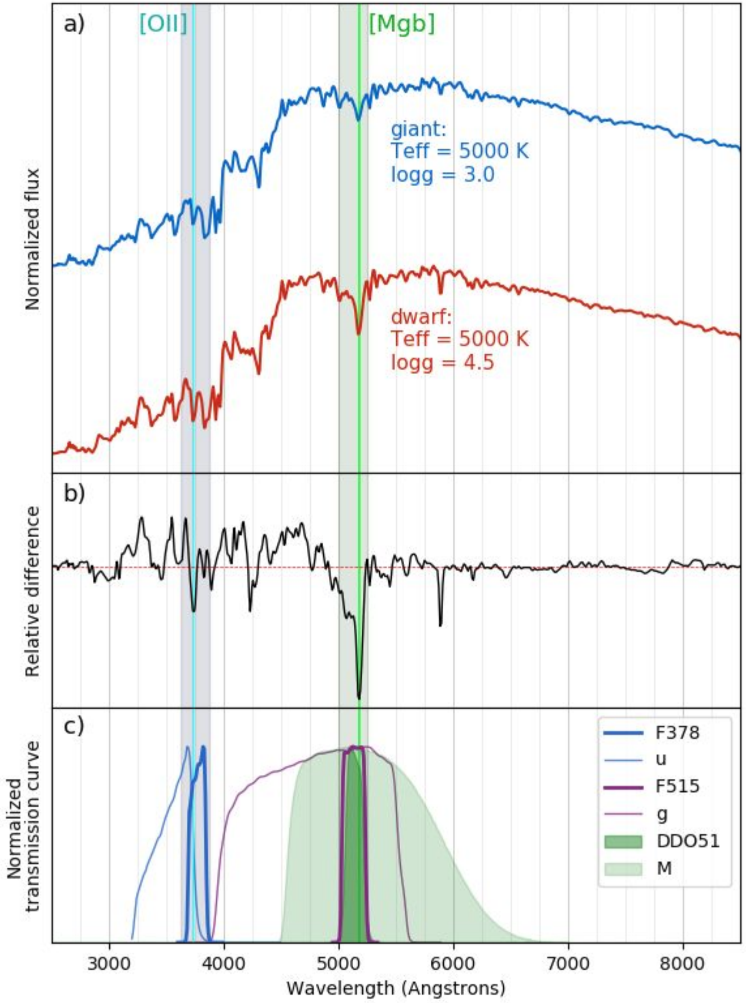
The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

Giants

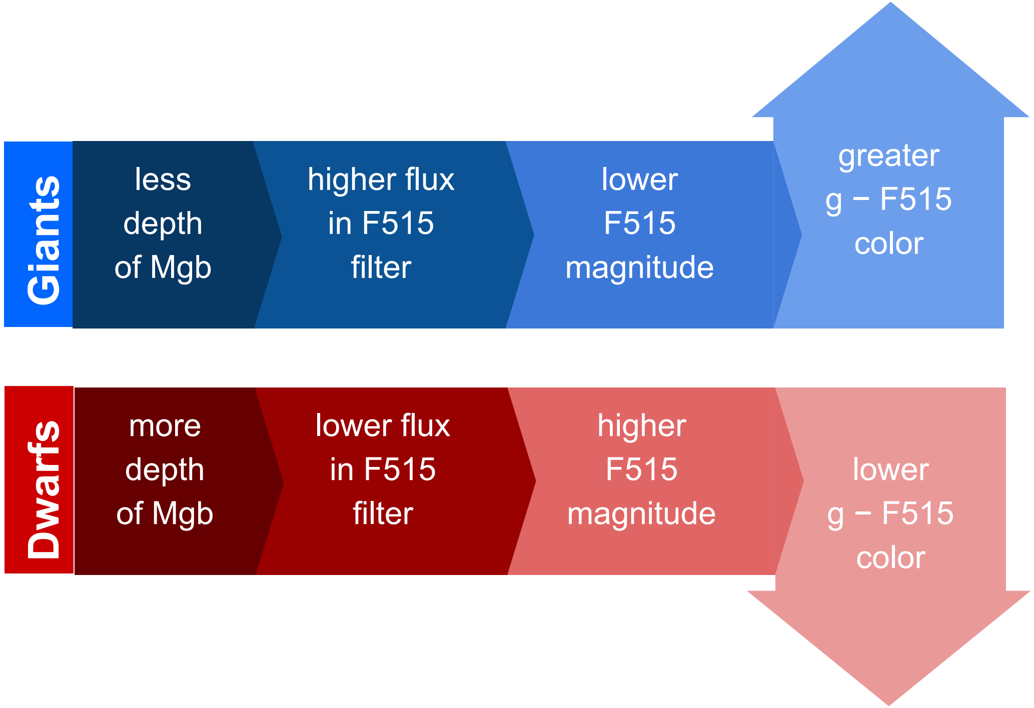
less depth of Mgb → higher flux in F515 filter → lower F515 magnitude → greater g − F515 color

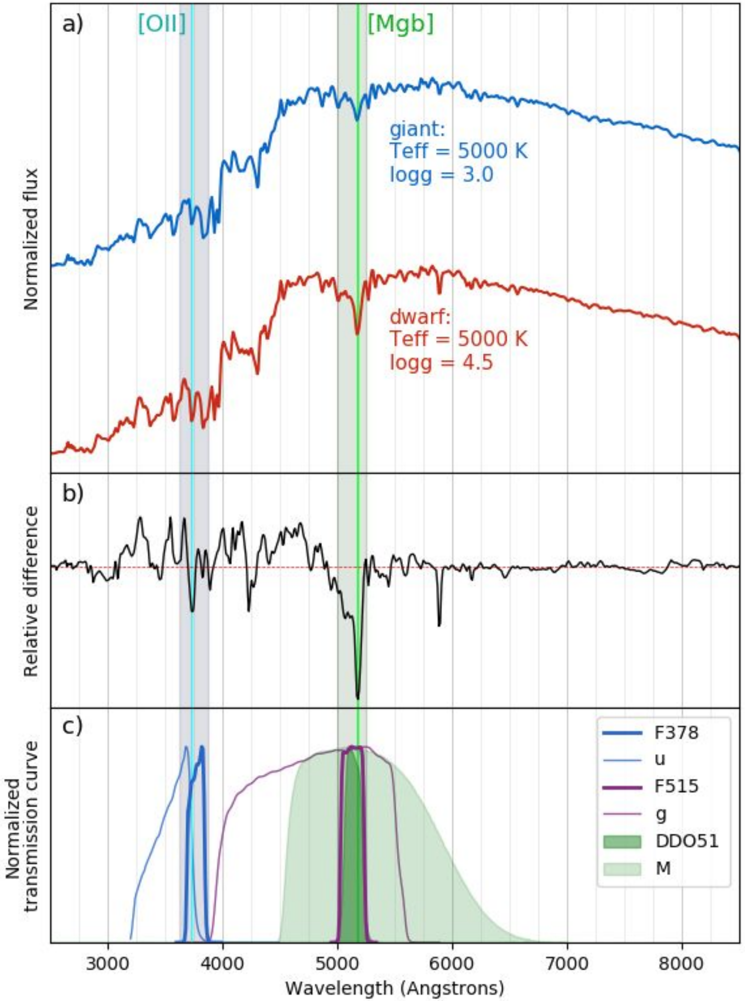
$$m_x = -2.5 \log_{10} \left(\frac{F_x}{F_{x,0}} \right)$$



| The photometric search technique

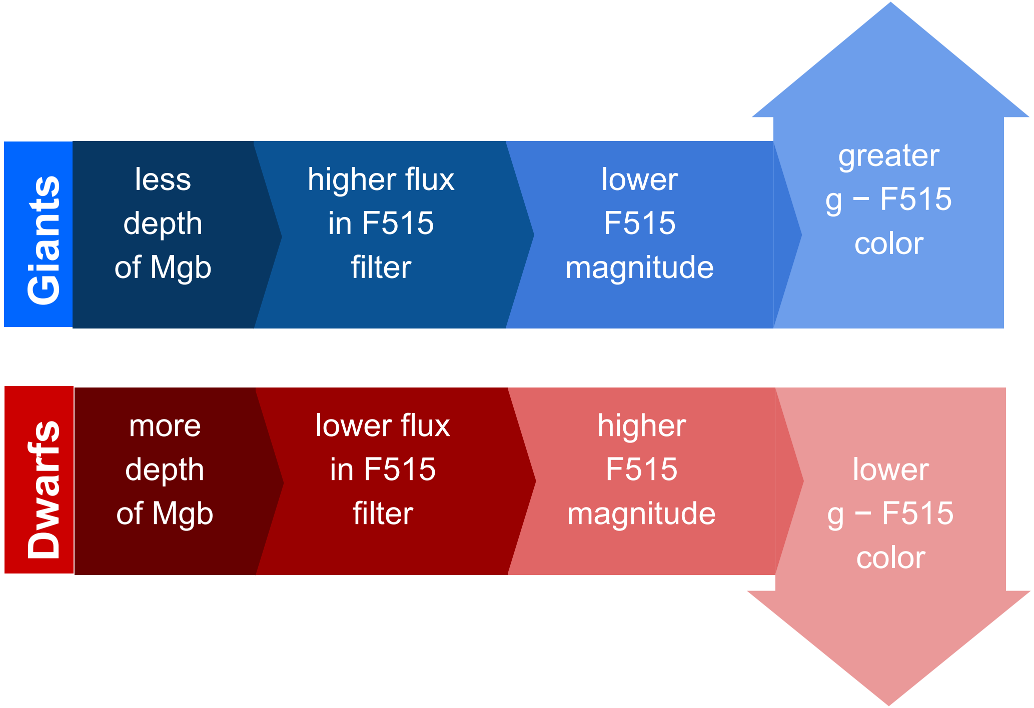
In S-PLUS, the Mgb line strength is measured by the g − F515 color.

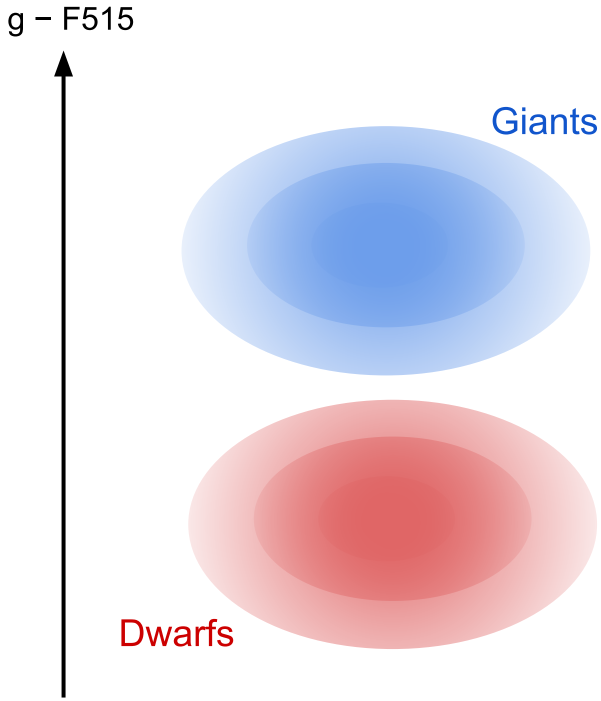




The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

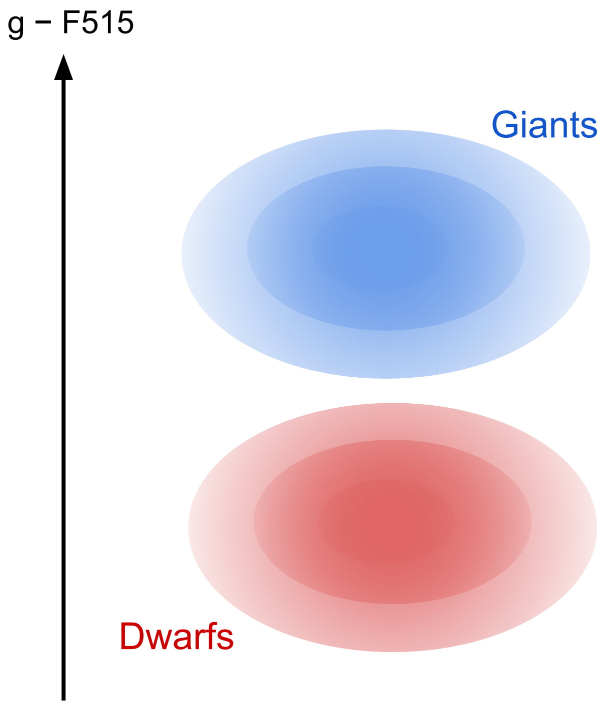




The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

But, of course, it's not so simple!

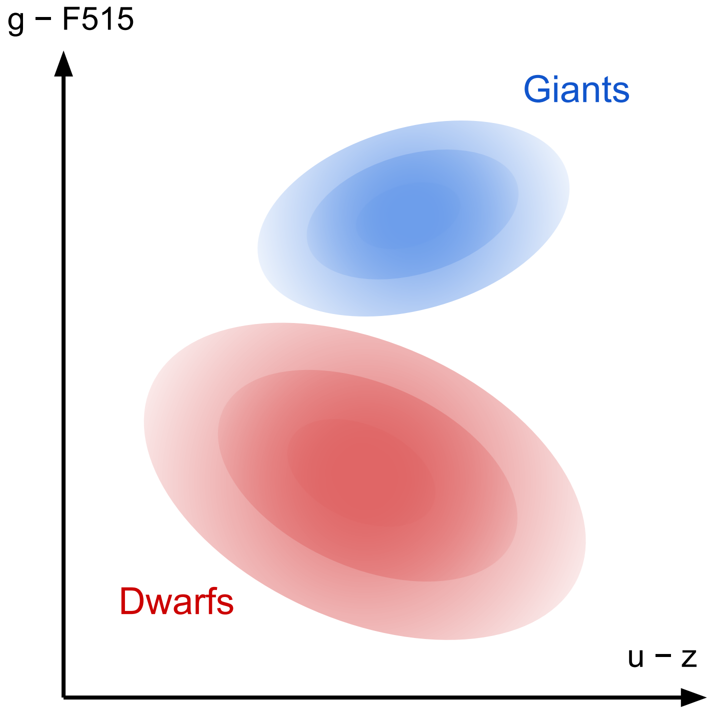


The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

But, of course, it's not so simple!

- There is also correlation with the temperature.



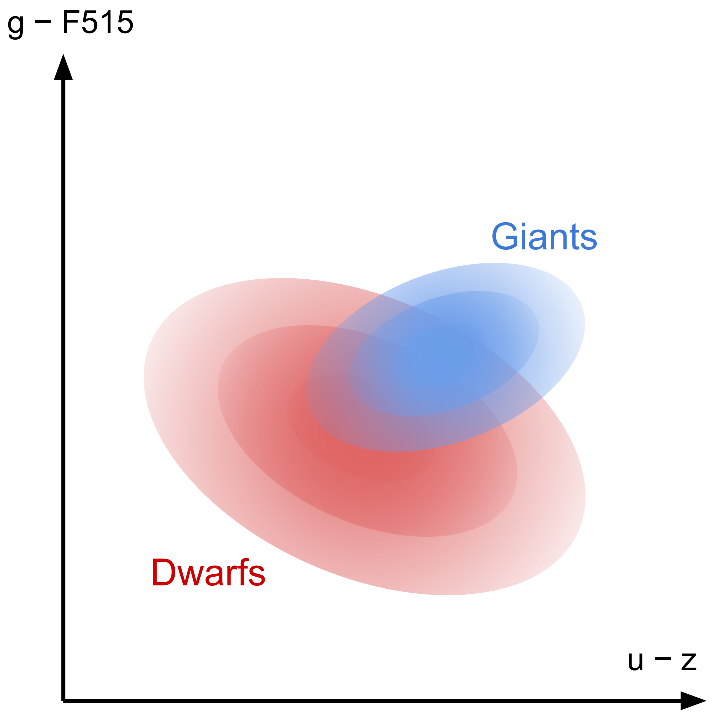
| The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

But, of course, it's not so simple!

- There is also correlation with the temperature.

- The distributions overlap (significantly)

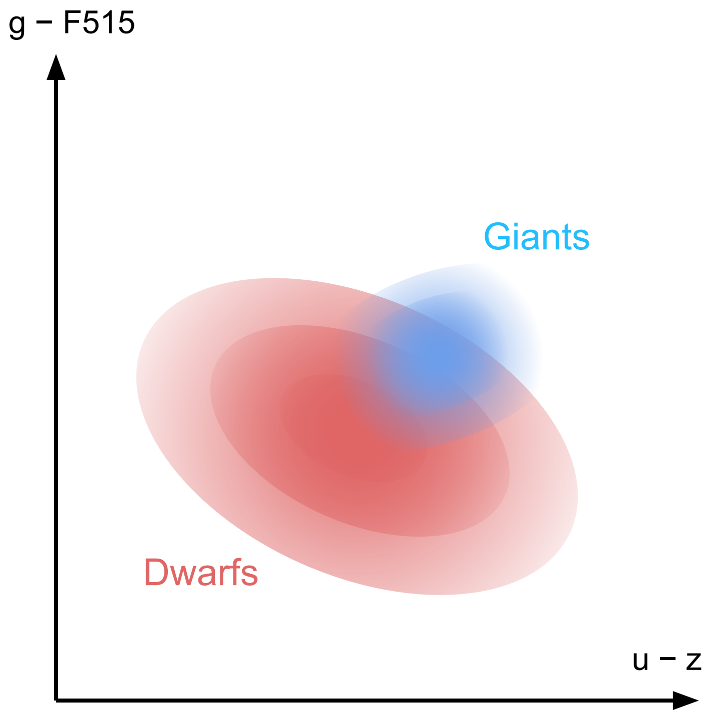


| The photometric search technique

In S-PLUS, the Mgb line strength is measured by the g − F515 color.

But, of course, it's not so simple!

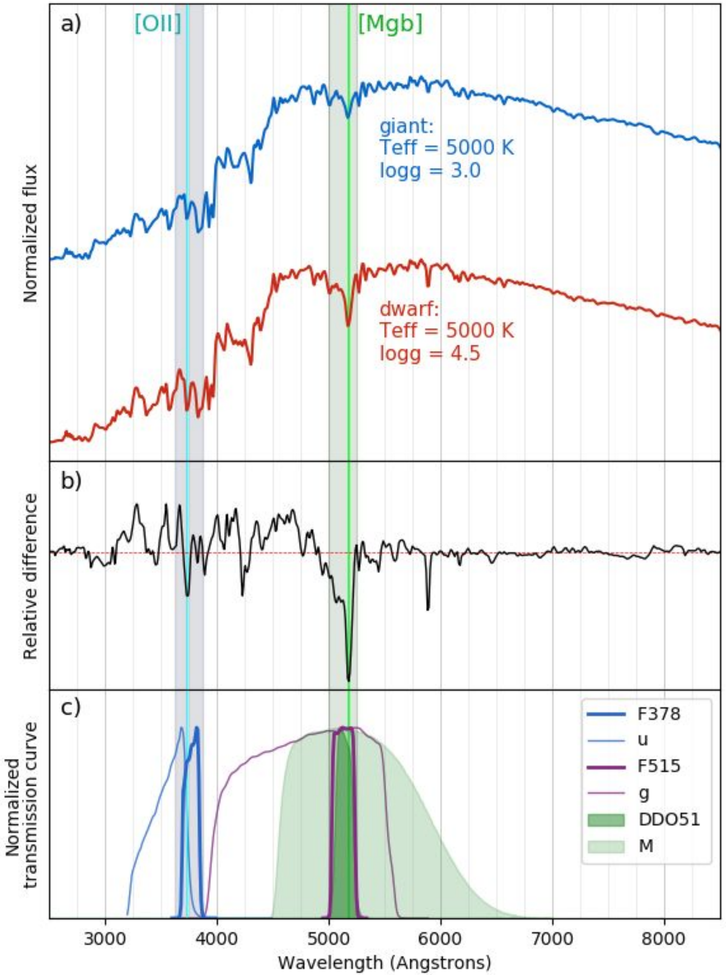
- There is also correlation with the temperature.
- The distributions overlap (significantly)
- The Density of giants is smaller than density of dwarfs (by a lot)



The photometric search technique

We saw that, in theory, the S-PLUS filter system is sensitive to the features that allow the use of the photometric search technique to select giant stars.

To do that in practice, we need a sample of known giants and dwarfs to train some classifier algorithm.

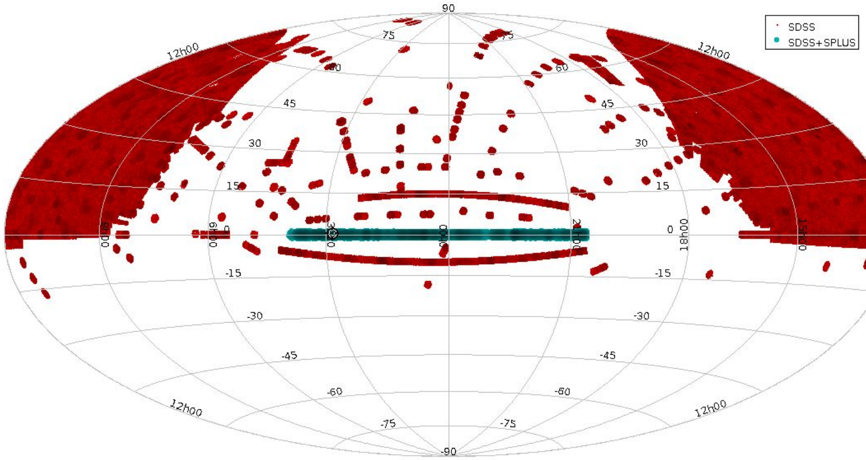


Application

| The SEGUE - SSPP sample

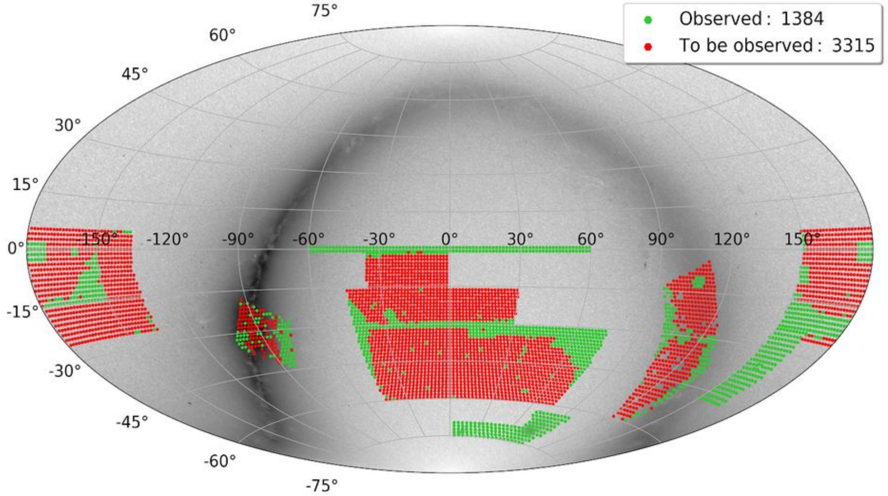
The SEGUE - SSPP has everything we need to test the method:

- Spectroscopical stellar parameters:
  - allows us to select the giant stars directly by their values of Teff and logg
- Overlap with S-PLUS DR1
  - Allows us to obtain the distributions of giants and dwarfs in the S-PLUS color-color diagrams

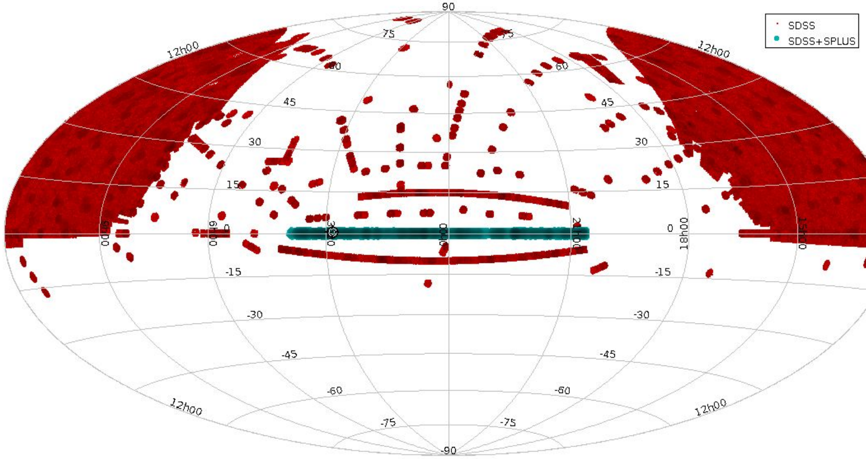


| The SEGUE - SSPP sample

While the surveys mostly cover different hemispheres, they overlap for small declinations. For this study, we used the S-PLUS DR1, covering the STRIPE-82 region.



S-PLUS

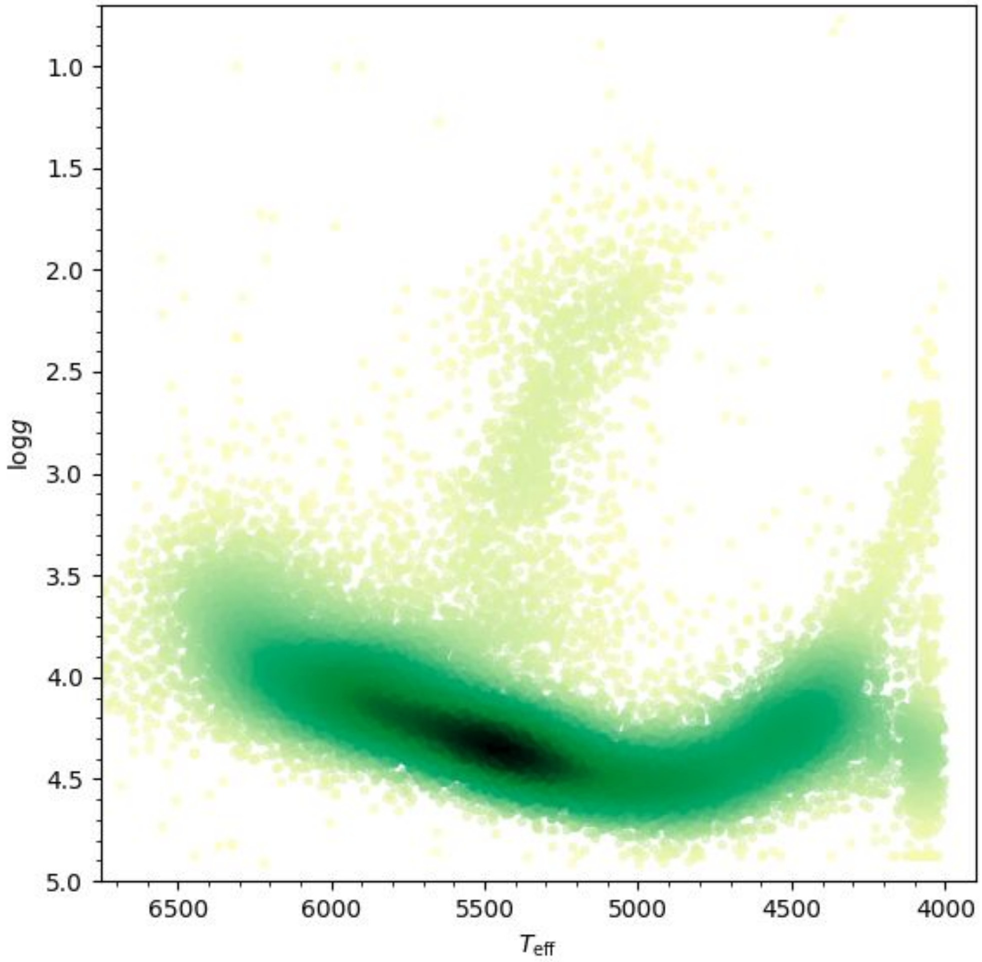


SSPP

| The SEGUE - SSPP sample

We use the stellar parameters logg and Teff to select the giant stars.

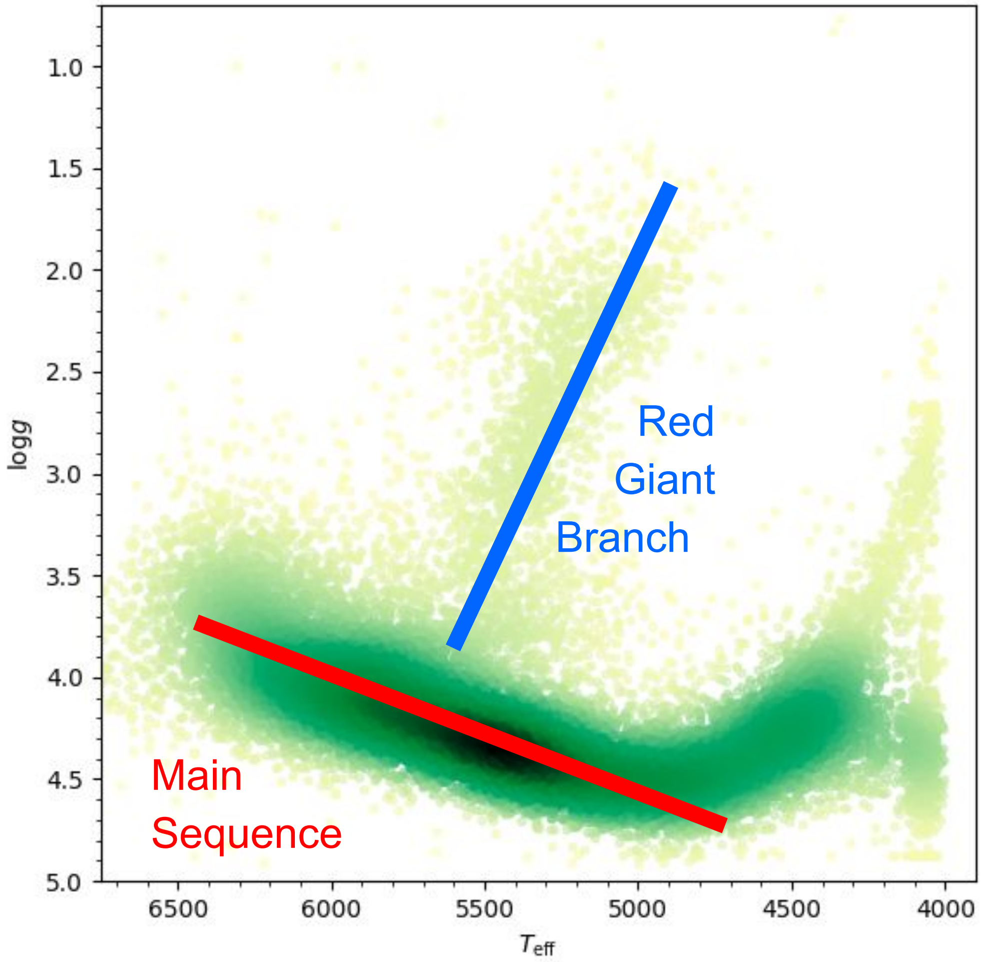
Many features discussed for the HR diagram can be seen in this plot (logg strongly correlates with luminosity).



| The SEGUE - SSPP sample

We use the stellar parameters logg and Teff to select the giant stars.

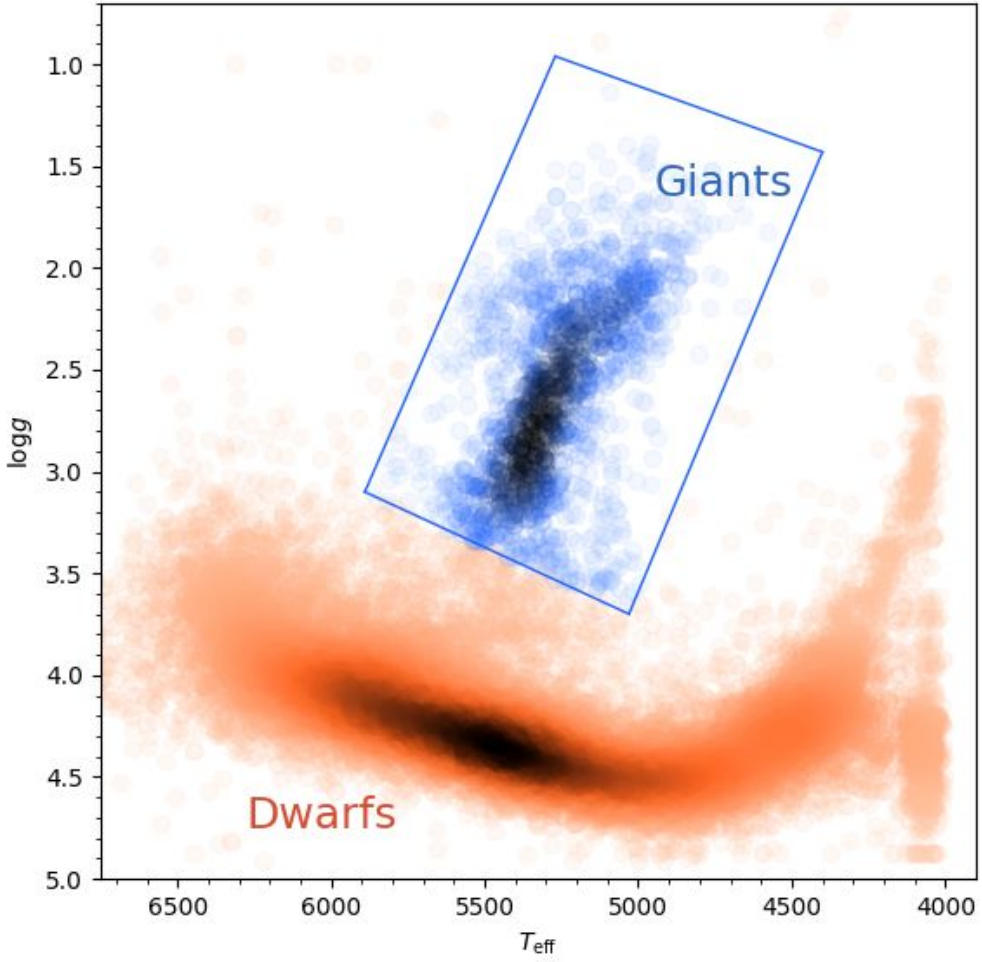
Many features discussed for the HR diagram can be seen in this plot (logg strongly correlates with luminosity).



| The SEGUE - SSPP sample

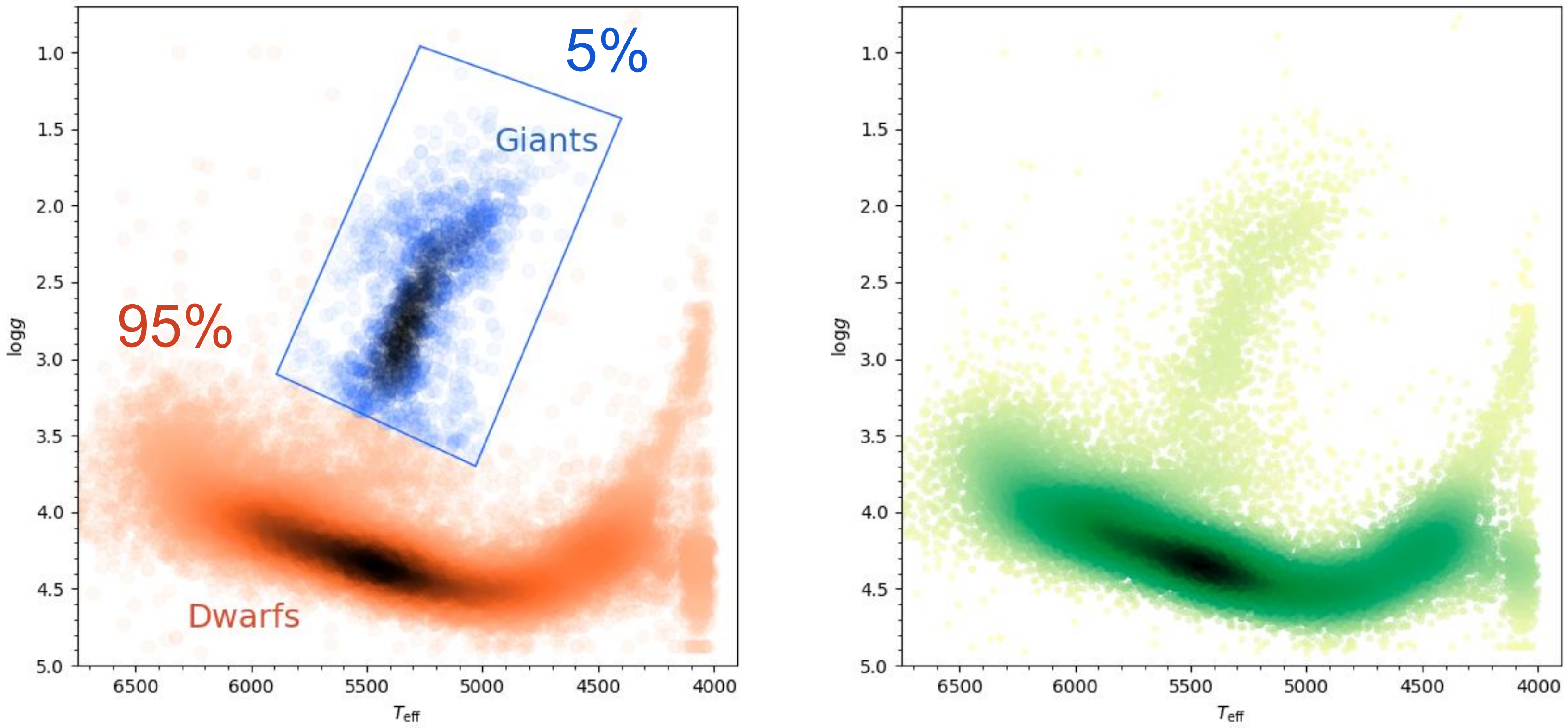
Therefore, we can use these parameters to select the giant stars of our study.

This selection was made using the topcat polygon selection tool (the same one that you will use in the Hands-on activity).

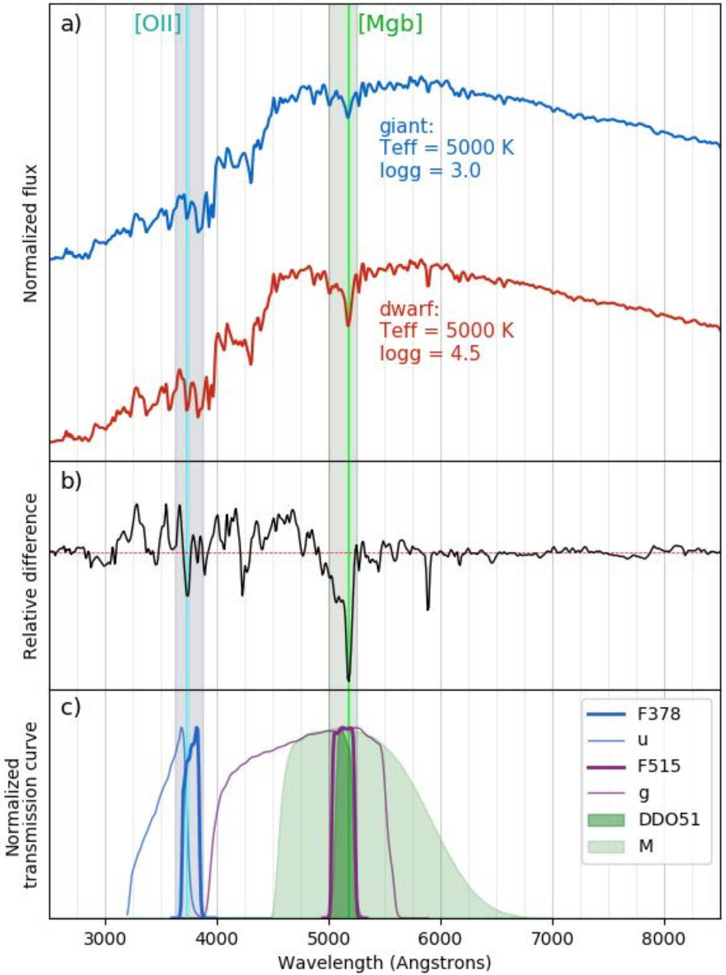


| The SEGUE - SSPP sample

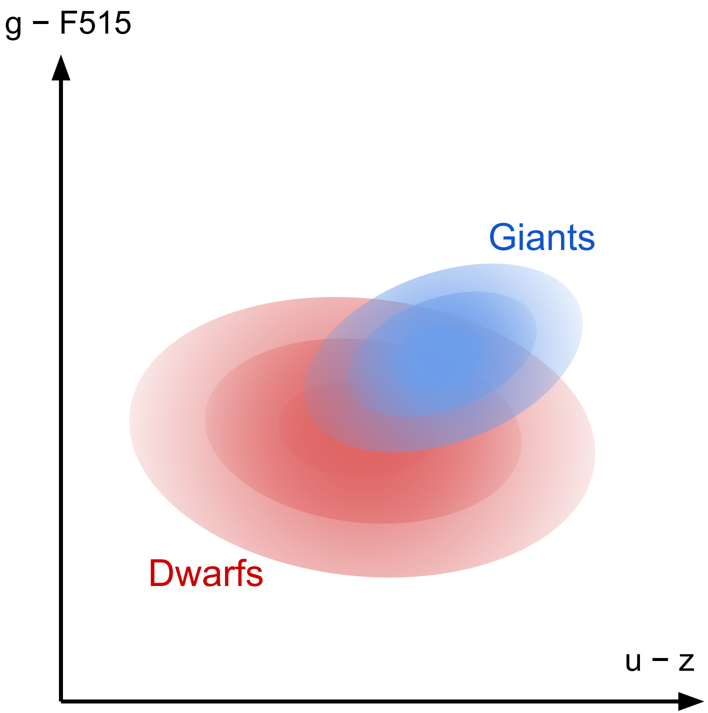
But as we discussed, the fraction of dwarfs is much greater than the fraction of giants.



| Recap



a)
[OII]
[Mgb]
giant:
Teff = 5000 K
logg = 3.0
dwarf:
Teff = 5000 K
logg = 4.5
Normalized flux
b)
Relative difference
c)
Normalized transmission curve
F378
u
F515
g
DDO51
M
3000
4000
5000
6000
7000
8000
Wavelength (Angstrons)

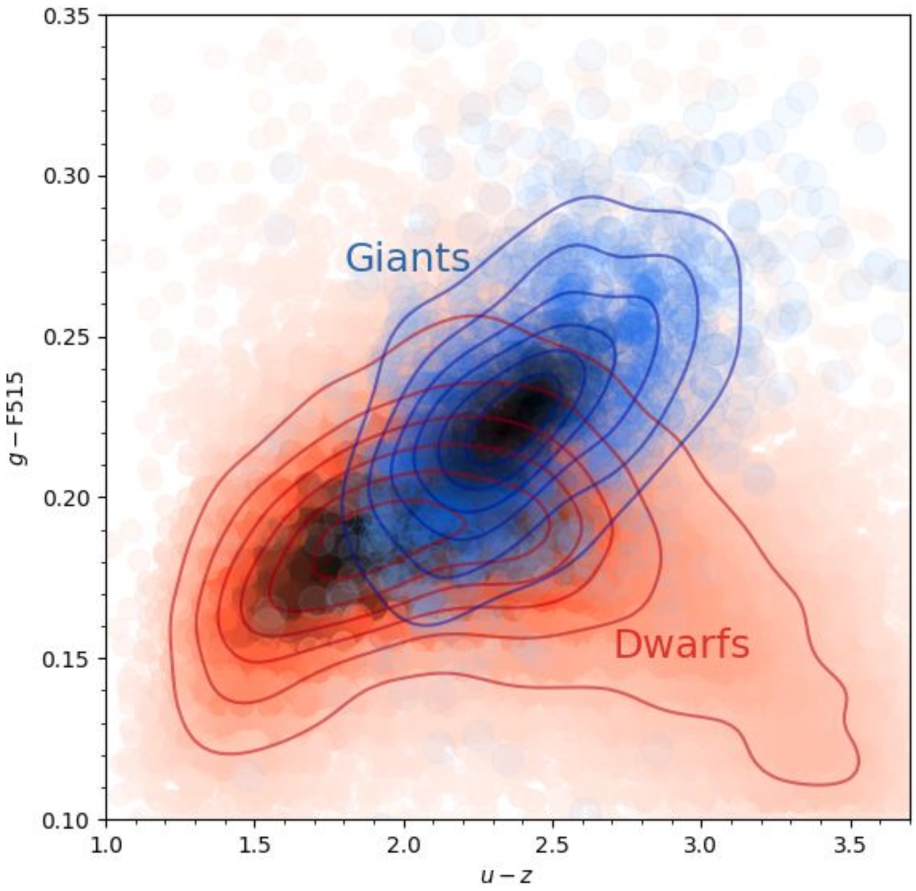


g − F515
Giants
Dwarfs
u − z

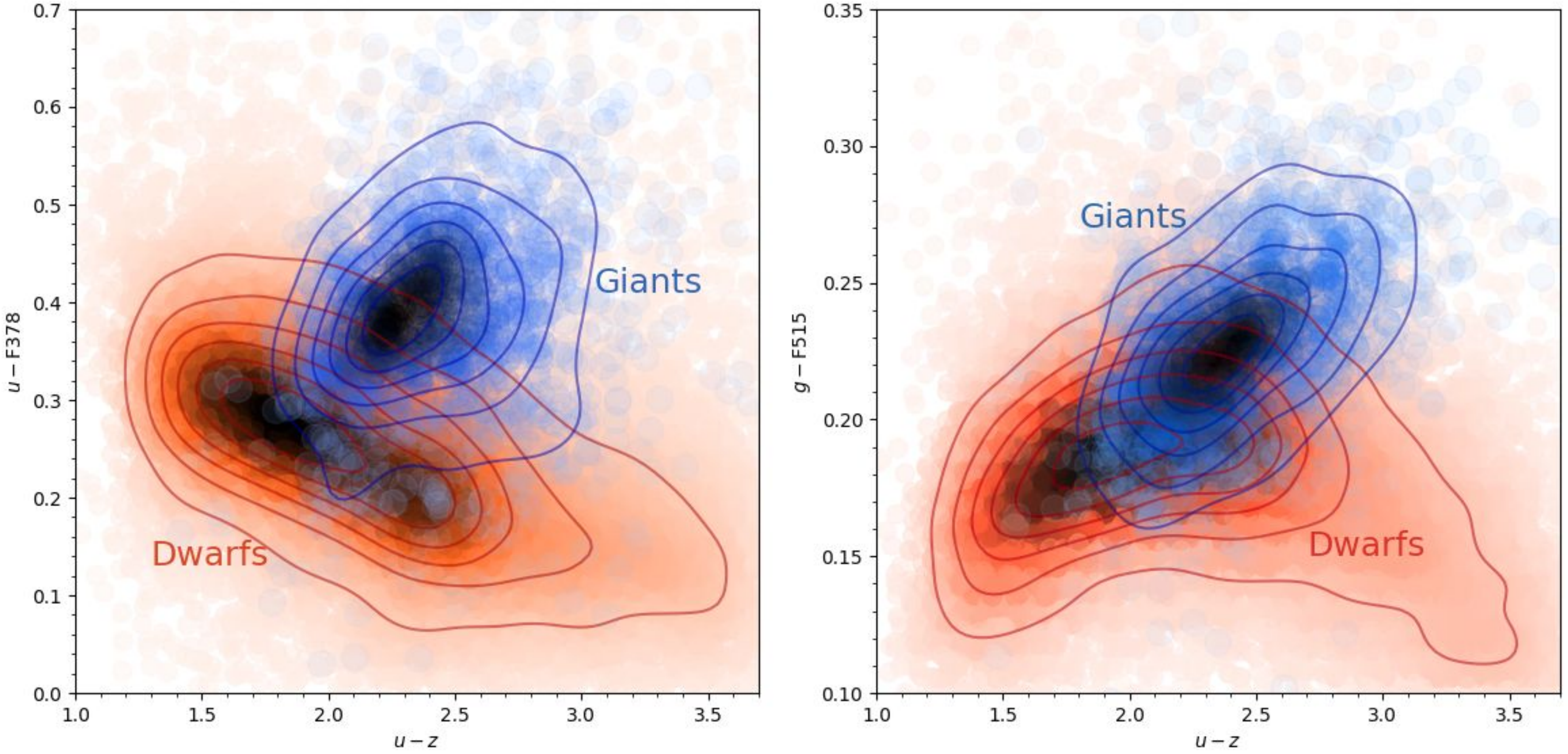
| Giants in the S-PLUS survey

As expected, the dwarfs and giants distributions are clearly distinct.

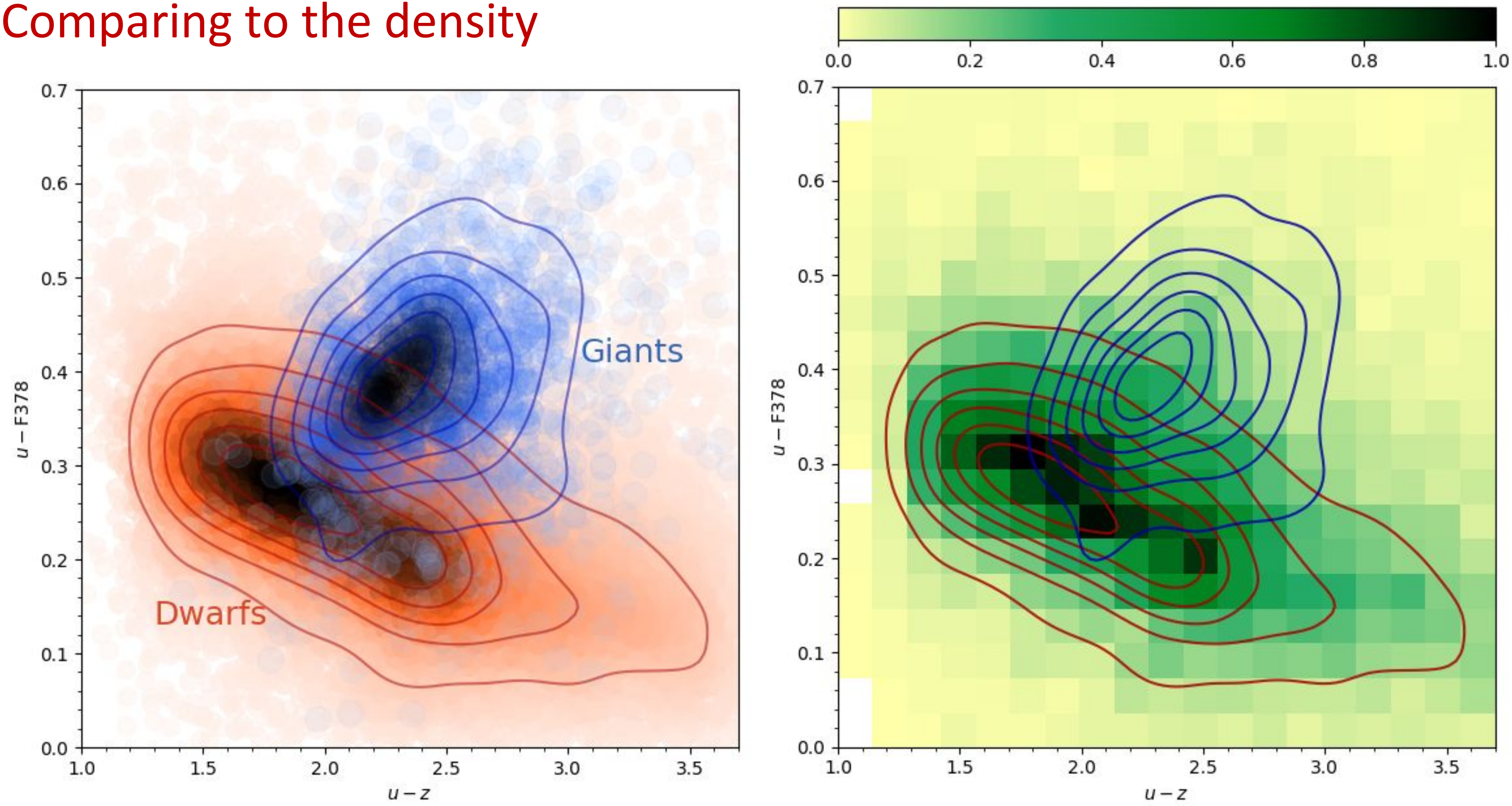
But they also significantly overlap.



u - F378 shows an even clearer separation



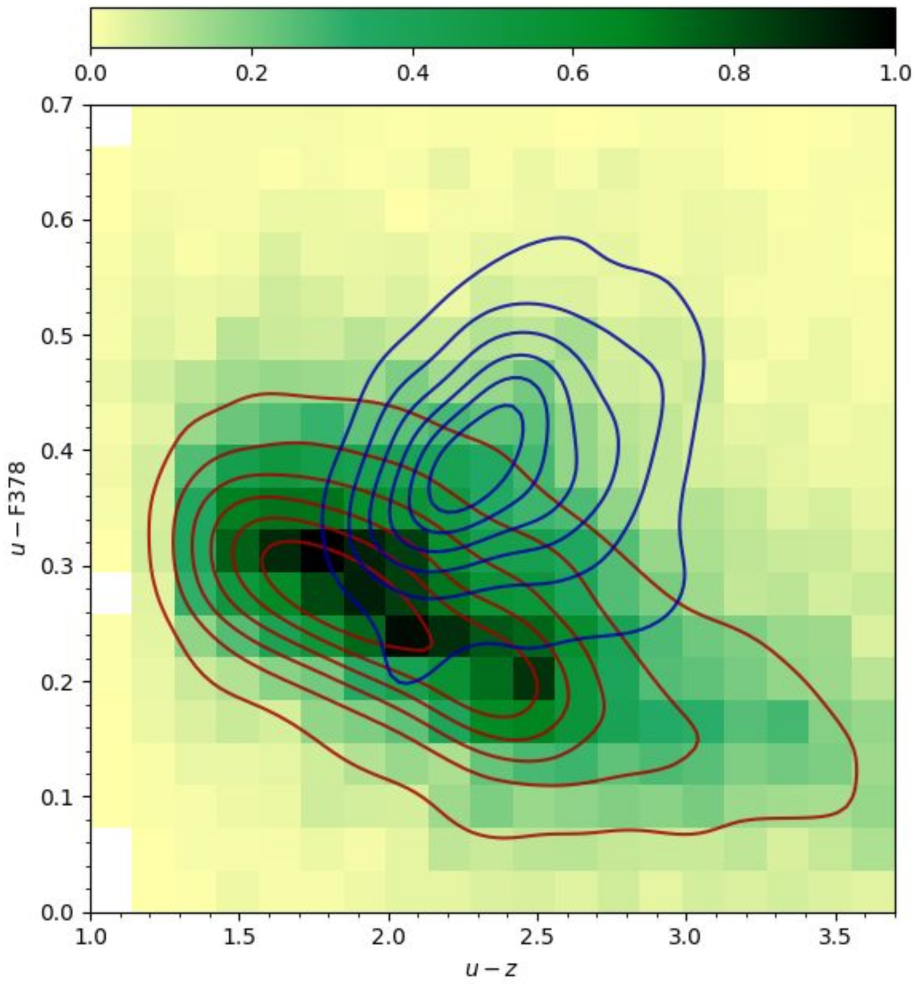
Comparing to the density



Comparing to the density

In fact, around the peak of the giants distribution, there are more dwarf stars than giants.

So the selection of giant stars (**candidates**) is not so straightforward.

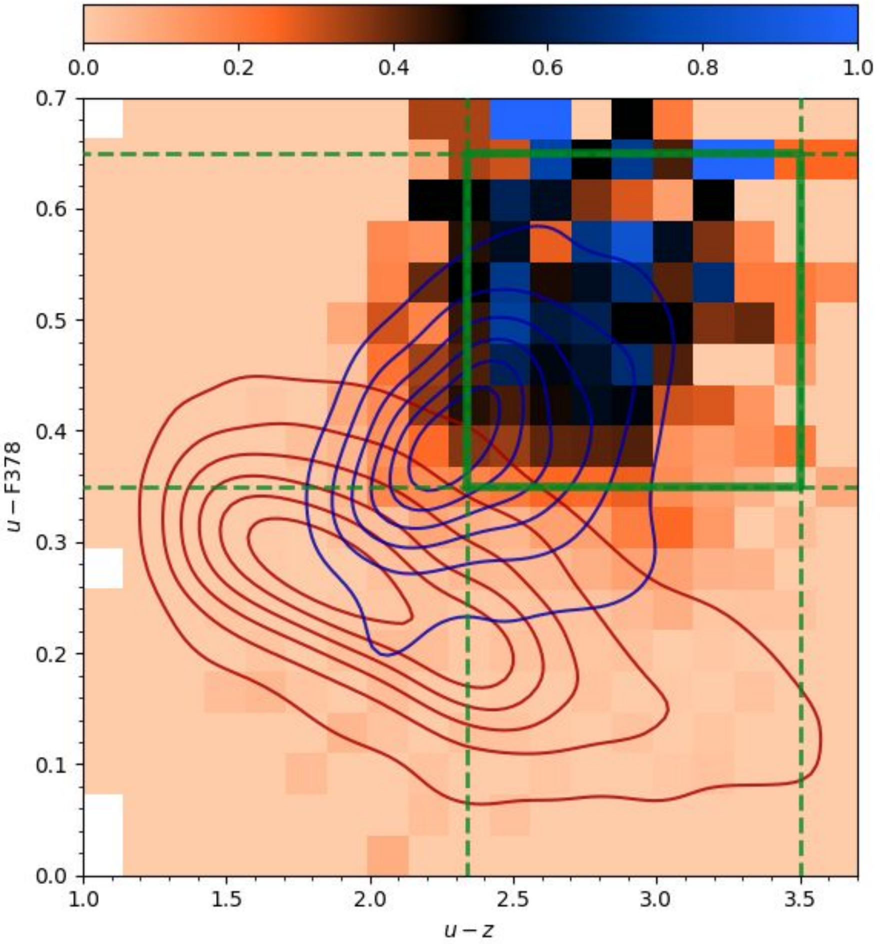


| Fraction of giants

In fact, around the peak of the giants distribution, there are more dwarf stars than giants.

So the selection of giant stars (**candidates**) is not so straightforward.

We get better results looking for the region where the fraction of giants is greater than the fraction of dwarfs



How to choose the best selection

		Predicted classification: Giant	Predicted classification: Dwarf
True classification	Giant	True Positive (TP)	False Negative (FN)
True classification	Dwarf	False Positive (FP)	True Negative (TN)

Precision

TP / (TP + FP)

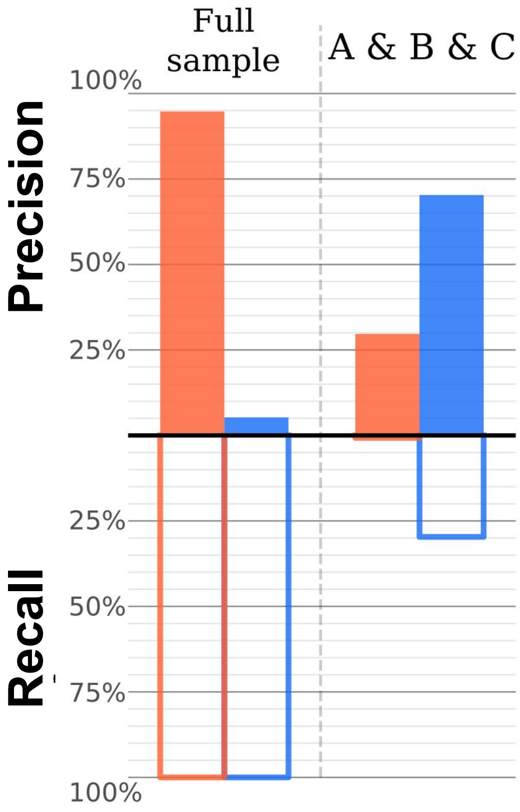
$$\frac{TP}{TP + FP}$$

Recall

TP / (TP + FN)

$$\frac{TP}{TP + FN}$$

| Results with color cuts



dwarfs

giants

$A = \{2.34 < u - z < 3.50\}$

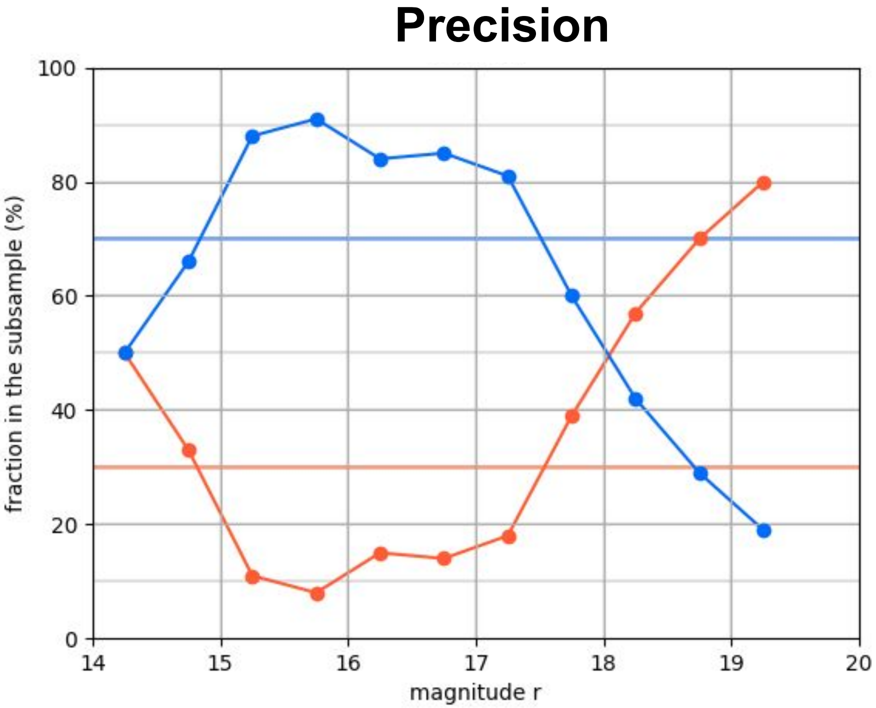
$B = \{0.22 < g - F515 < 0.33\}$

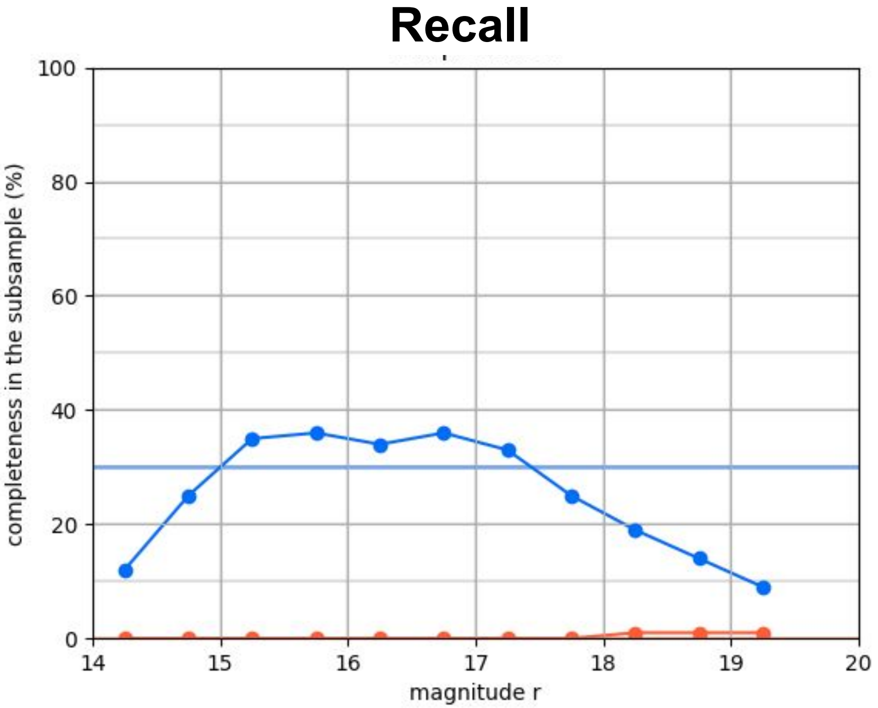
$C = \{0.35 < u - F378 < 0.65\}$

Just using simple color cuts it’s possible to reach values of **70% precision** (fraction of selected stars that are actually giants) and **30% recall** (fraction of giants selected from the whole sample of giants).

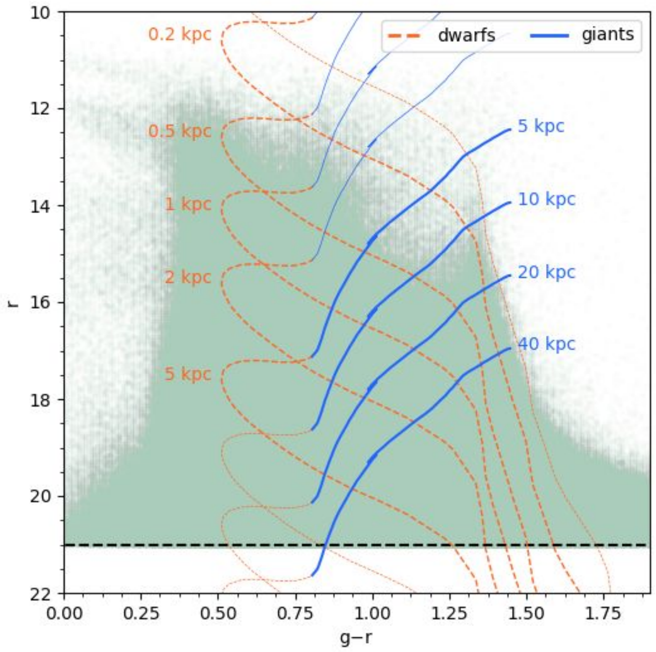
Result as a function of magnitude

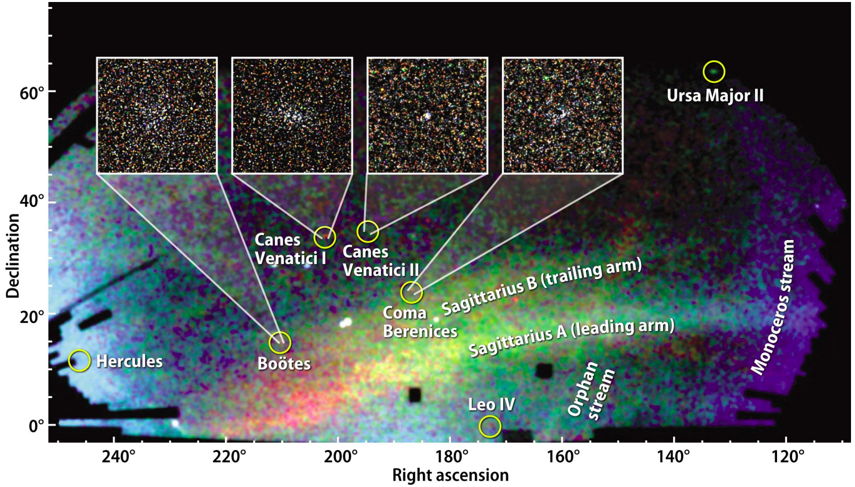
The results gets significantly worse for fainter objects

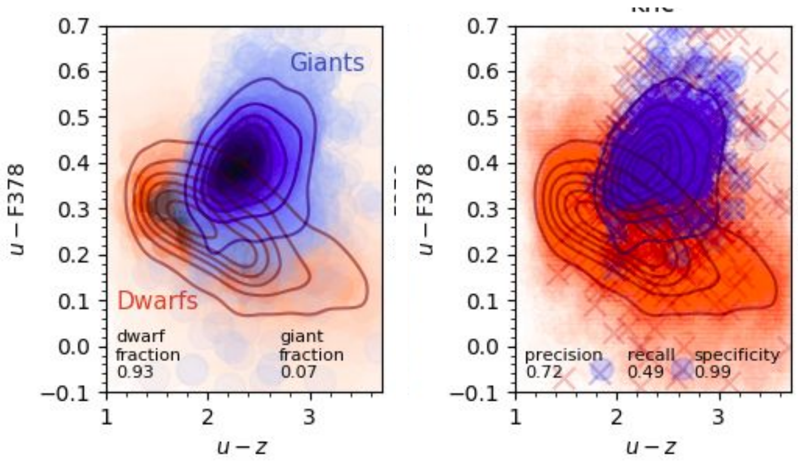




Conclusions







Hands-On Activity

Testing different classifiers

For this specific problem, one of the classifiers works better than the rest.