Star Formation Processes Visualized Using Object Oriented Databases

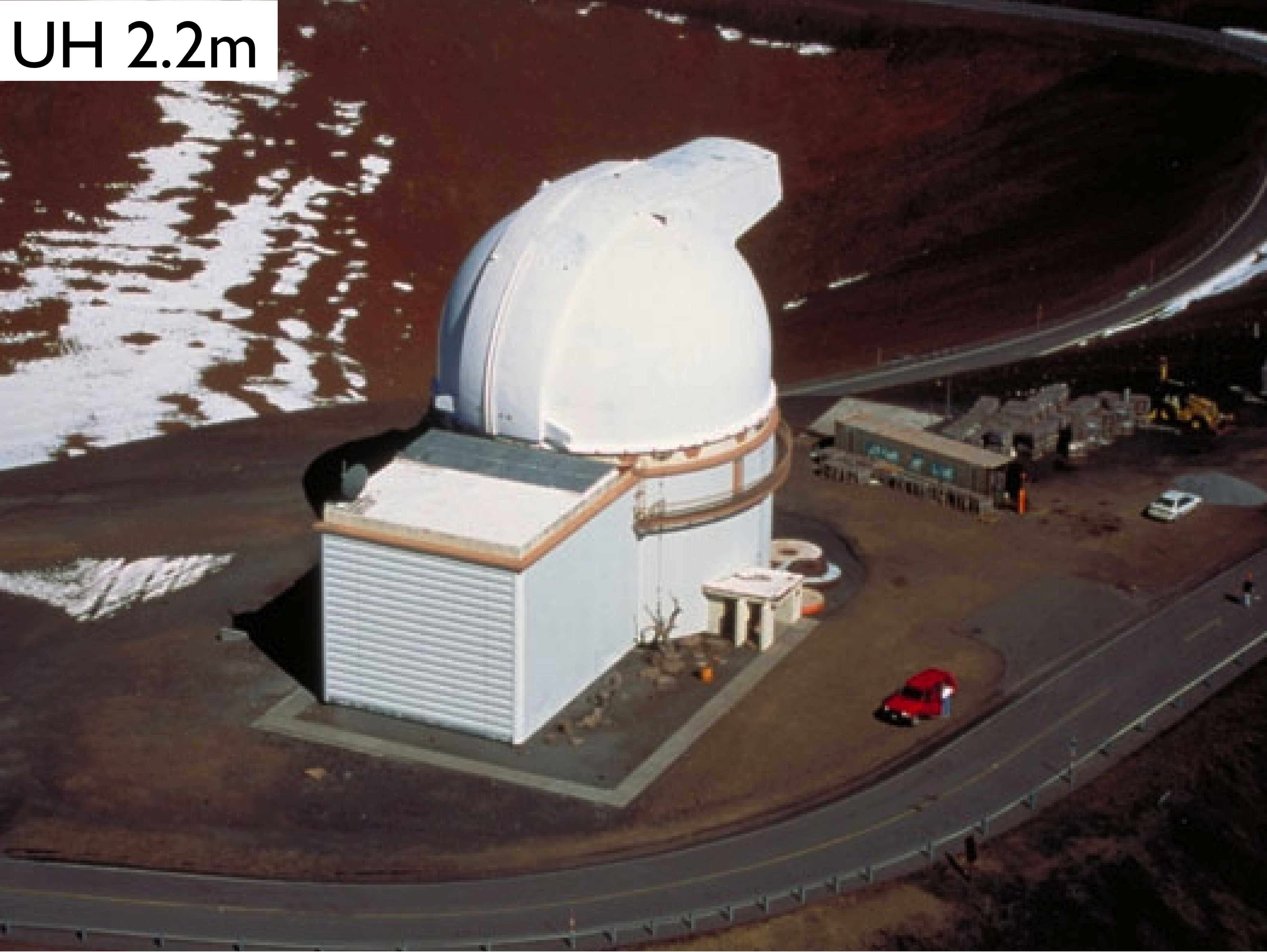
Kevin Hu
University of Hawai'i at Manoa

Mentor: Marianne Takamiya
University of Hawai'i at Hilo

AWI Akamai Workforce Initiative
BUILDING HAWAII'S SCIENTIFIC & TECHNICAL WORKFORCE

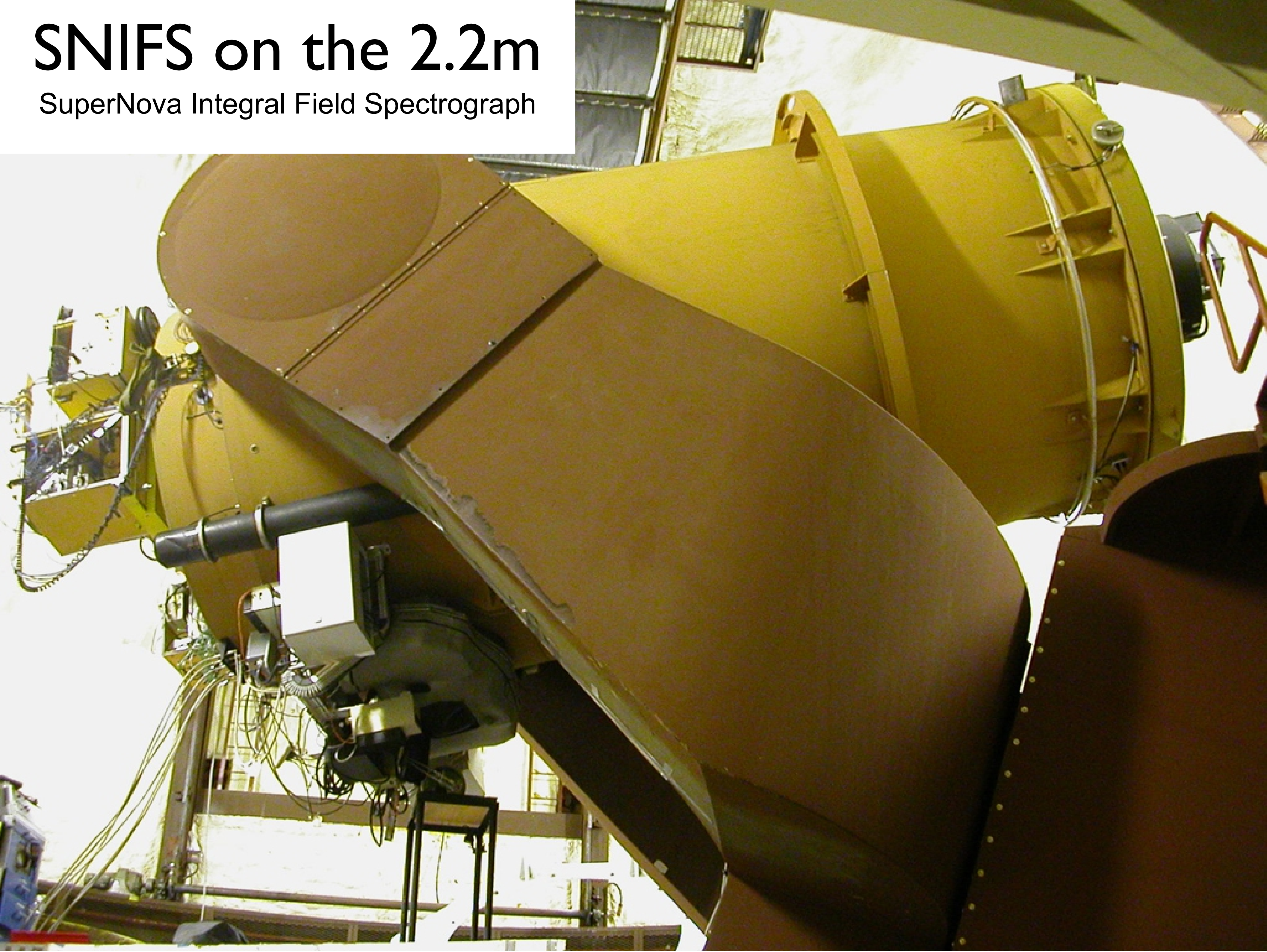
Summit of Mauna Kea

UH 2.2m



SNIFS on the 2.2m

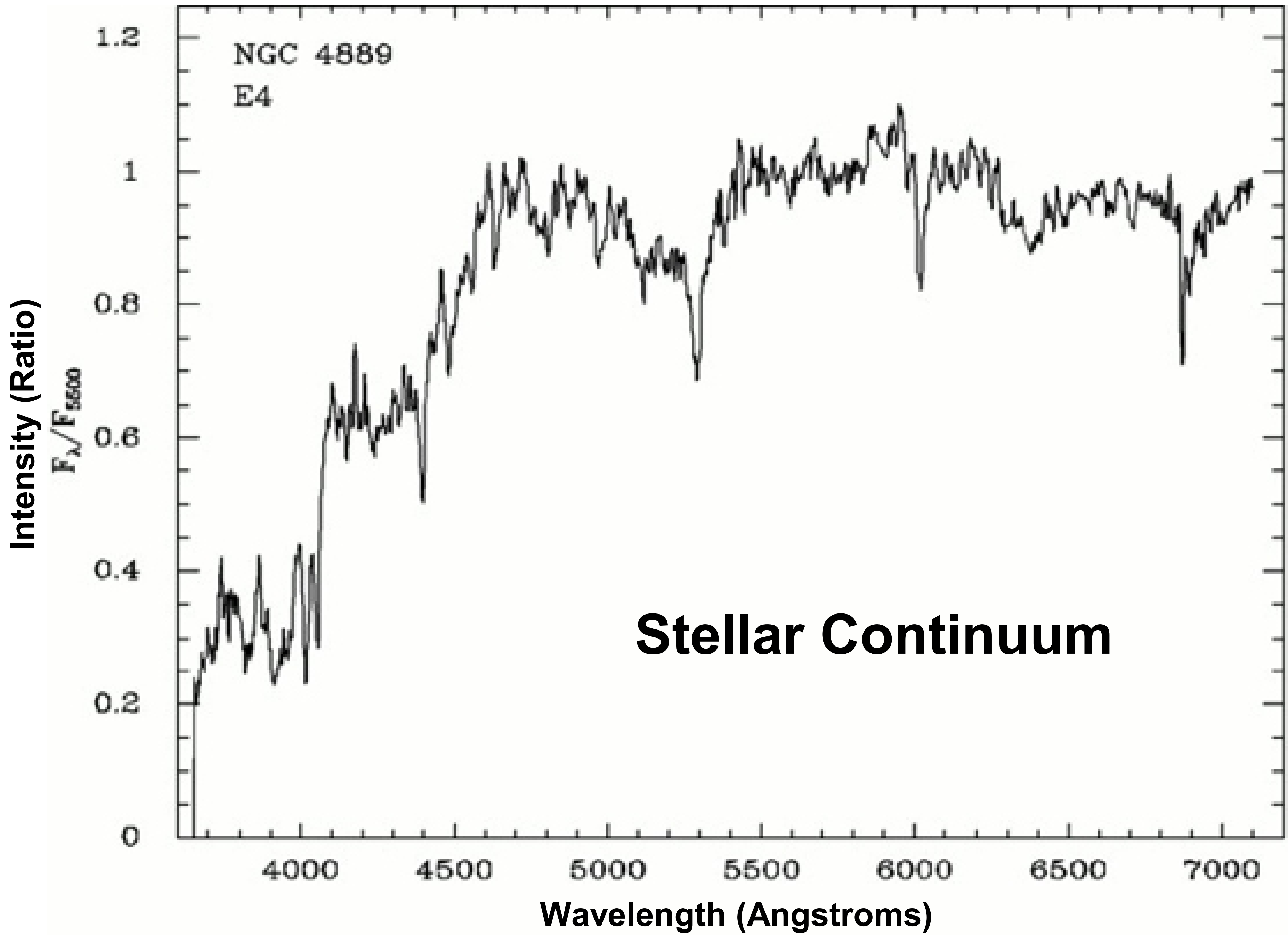
SuperNova Integral Field Spectrograph

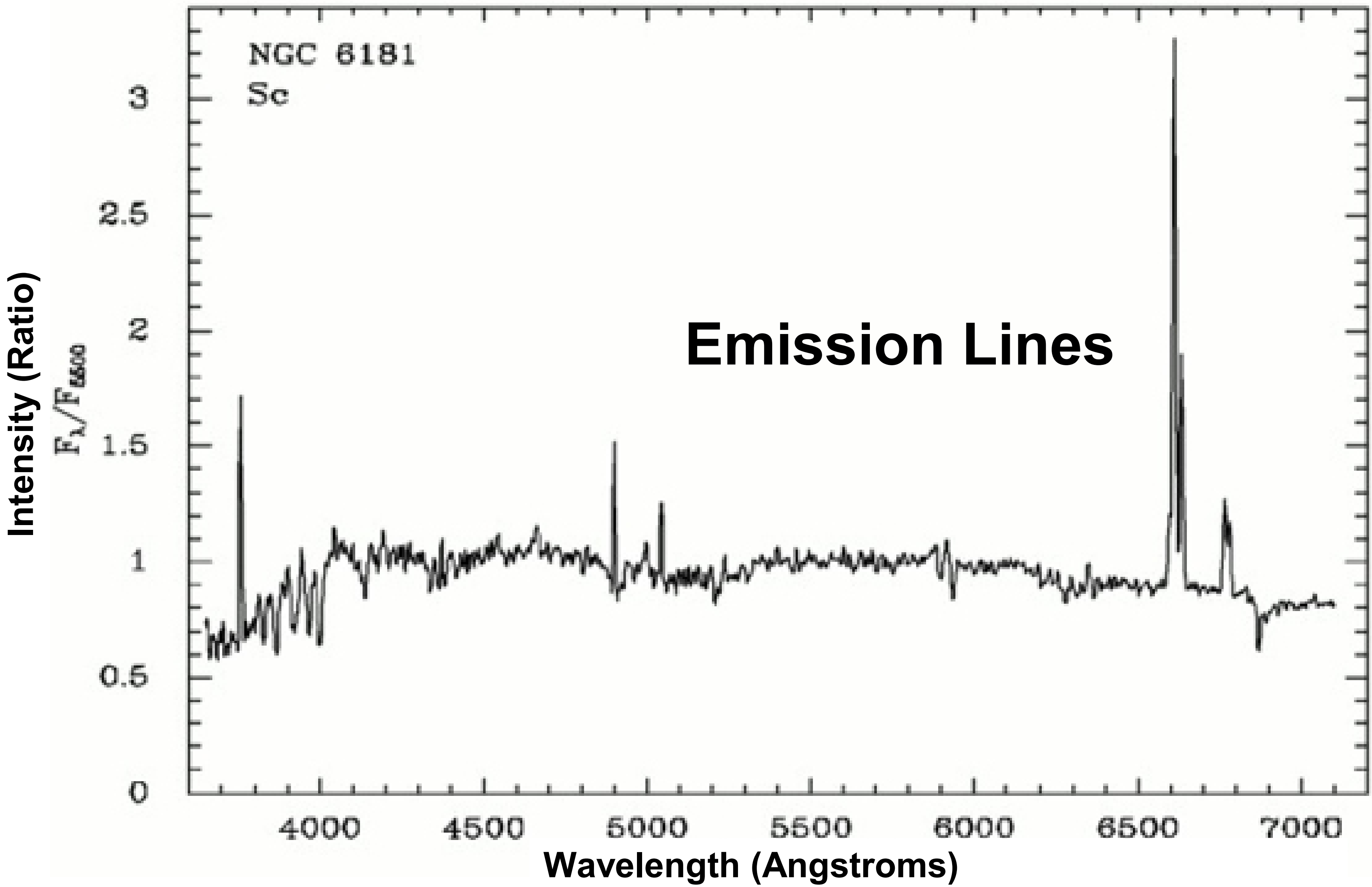


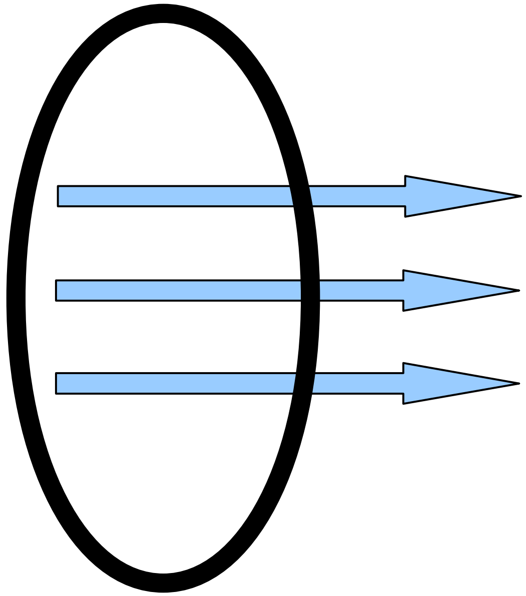
```
num obj cubeName    edshift year    day month    RA   DEC x:2 y:2 expTime airMass Rflux    Rfwhm    Bflux    Bfwhm    flux6548     eqw6548 gfwhm6548    flux6563
eqw6563 gfwhm6563    flux6583     eqw6583 gfwhm6583    flux6716     eqw6716 gfwhm6716    flux6731     eqw6731 gfwhm6731    flux3728     eqw3728 gfwhm3728    flux4363
eqw4363 gfwhm4363    flux4861     eqw4861 gfwhm4861    flux4959     eqw4959 gfwhm4959    flux5007     eqw5007 gfwhm5007    bgFile  bgSkies timeOffset  distOffset
distance    distStdDev  user_count  R_avg_vote  B_avg_vote  Classification
1   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   1   1220.0  1.126   1.039e-16   9.21    5.932e-17   11.82   -1.33E-19
0.02504 10.79   1.354E-16   -24.28  10.71   -4.02E-17   7.2 14.59   2.018E-17   -3.307  5.406   7.139E-17   -11.7   49.54   7.895e-17   4.021   -27.59
2.777e-16   64.11   -144.6  -2.08e-15   259.2   248.4   -5.93e-17   6.097   17.83   2.446e-16   33.74   -241.4  TC10_270_064    144 1845    0.0 20.300  0.872
2   1.0 -1.0    Undecided
2   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   2   1220.0  1.126   1.101e-16   8.13    0   14.37   -1.28E-17   3.408
3.094   1.089E-16   -29.73  7.645   2.352E-17   -6.562  5.885   1.972E-17   -6.781  5.124   6.016E-17   -21.08  16.2    2.513e-16   38.42   -343.9  2.329e-12
1976.   INDEF   -1.34e-15   226.2   189.8   2.617e-17   8.516   -10.01  1.098e-16   11.6    -36.45  TC10_270_064    144 1845    0.0 20.300  0.872   2   1.0
-1.0    Undecided
3   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   3   1220.0  1.126   1.306e-16   8.667   0   7.171   4.643E-17   -17.
19.31   1.331E-16   -47.26  8.594   3.318E-17   -11.28  7.66    2.400E-17   -8.151  4.656   2.920E-17   -10.13  9.415   1.549e-16   7.718   -89.92  -4.01e-16
93.86   90.2    -1.7e-16    57.52   80.71   4.219e-17   10.33   -36.45  8.962e-17   4.928   -64.62  TC10_270_064    144 1845    0.0 20.300  0.872   2   1.0
-1.0    Undecided
4   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   4   1220.0  1.126   1.503e-16   9.211   2.233e-17   3.69    1.032E-17
-3.839  2.35    1.411E-16   -52.69  8.25    8.951E-17   -33.62  31.3    4.265E-17   -13.94  11.73   3.365E-17   -10.86  10.91   2.143e-16   7.105   -519.1
4.143e-17   8.329   -22.38  4.724e-17   9.242   -31.15  1.947e-17   1.939   -7.566  6.261e-17   5.153   -37.11  TC10_270_064    144 1845    0.0 20.300  0.872
2   1.0 -1.0    Undecided
5   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   5   1220.0  1.126   1.945e-16   9.027   2.391e-17   3.54    1.198E-17
-5.007  3.068   1.945E-16   -76.76  8.781   1.983E-17   -7.217  7.089   4.570E-17   -15.07  7.273   3.857E-17   -12.81  8.874   2.436e-16   6.271   INDEF
-2.77e-17   5.084   13.34   3.95e-17    4.157   -183.8  7.784e-17   11.93   -69.25  1.685e-16   9.801   -150.2  TC10_270_064    144 1845    0.0 20.300  0.872
2   1.0 -1.0    Undecided
6   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   6   1220.0  1.126   2.123e-16   9.346   0   4.618   1.846E-17   -5.941
6.252   2.092E-16   -64.51  9.024   1.117E-17   -3.227  4.896   7.078E-17   -26.6   14.02   1.518E-17   -5.508  6.431   2.739e-16   6.324   -477.   4.287e-16
99.72   INDEF   -4.41e-16   185.1   152.2   4.532e-17   3.72    -15.1   1.921e-16   7.655   -107.4  TC10_270_064    144 1845    0.0 20.300  0.872   2   1.0
-1.0    Undecided
7   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   7   1220.0  1.126   3.011e-16   10.68   4.23e-17    5.259   3.114E-17
-10.48  18.86   2.915E-16   -96.48  9.981   5.413E-17   -17.67  10.4    4.829E-17   -16.77  7.861   5.360E-17   -17.34  9.982   2.789e-16   7.584   -163.2
-7.05e-17   14.54   18.16   2.346e-19   0.07051 -0.0839 5.416e-17   1.175   -31.42  1.416e-16   5.18    -162.1  TC10_270_064    144 1845    0.0 20.300  0.872
2   1.0 -1.0    Undecided
8   NGC0337 TC10_270_062    0.005497    2010    27  09  00:59:50.73 -07:34:58.59    1   8   1220.0  1.126   2.98e-16    8.882   8.798e-17   8.014   9.855E-18
-3.059  1.681   2.940E-16   -91.01  8.462   3.685E-17   -11.36  12.42   6.472E-17   -22.01  8.376   3.696E-17   -12.46  6.846   4.273e-16   9.257   INDEF
```

display all data within 0h1m11s by 60° and 0h;1m11s,60°

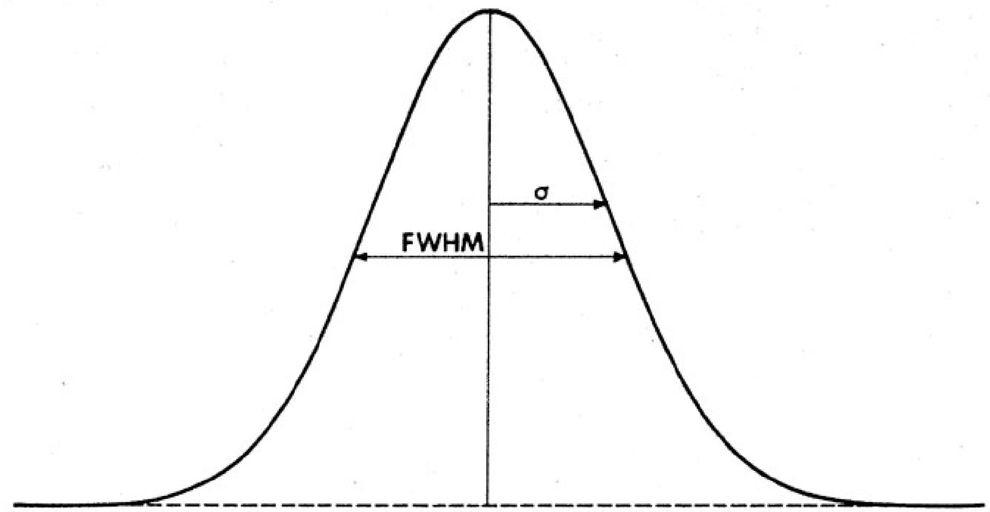
NGC 6946
NGC 02
NGC 03
NGC 04
NGC 05
NGC 06
NGC 07
NGC 08
NGC 09
NGC 10
NGC 11
NGC 12
NGC 13
NGC 14
NGC 15
NGC 16


NGC 4889
E4
Intensity (Ratio)
F_λ/F_{5500}
Stellar Continuum
Wavelength (Angstroms)
1.2
1
0.8
0.6
0.4
0.2
0
4000
4500
5000
5500
6000
6500
7000

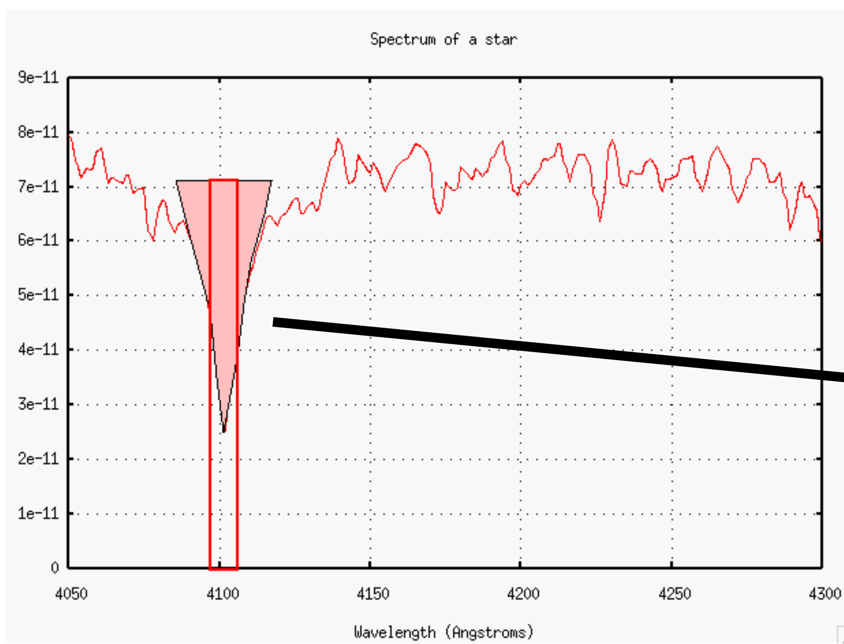

NGC 6181
Sc
Emission Lines
Intensity (Ratio)
F_λ/F_{5500}
3
2.5
2
1.5
1
0.5
0
4000
4500
5000
5500
6000
6500
7000
Wavelength (Angstroms)

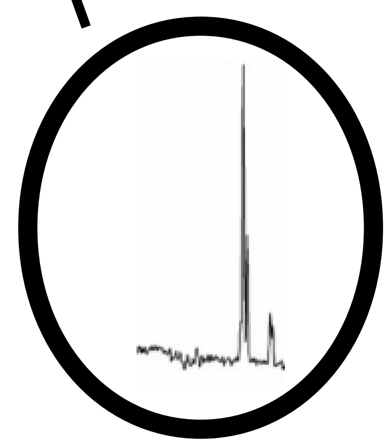


Energy Flux



(Gaussian) Full Width of Half-Maximum

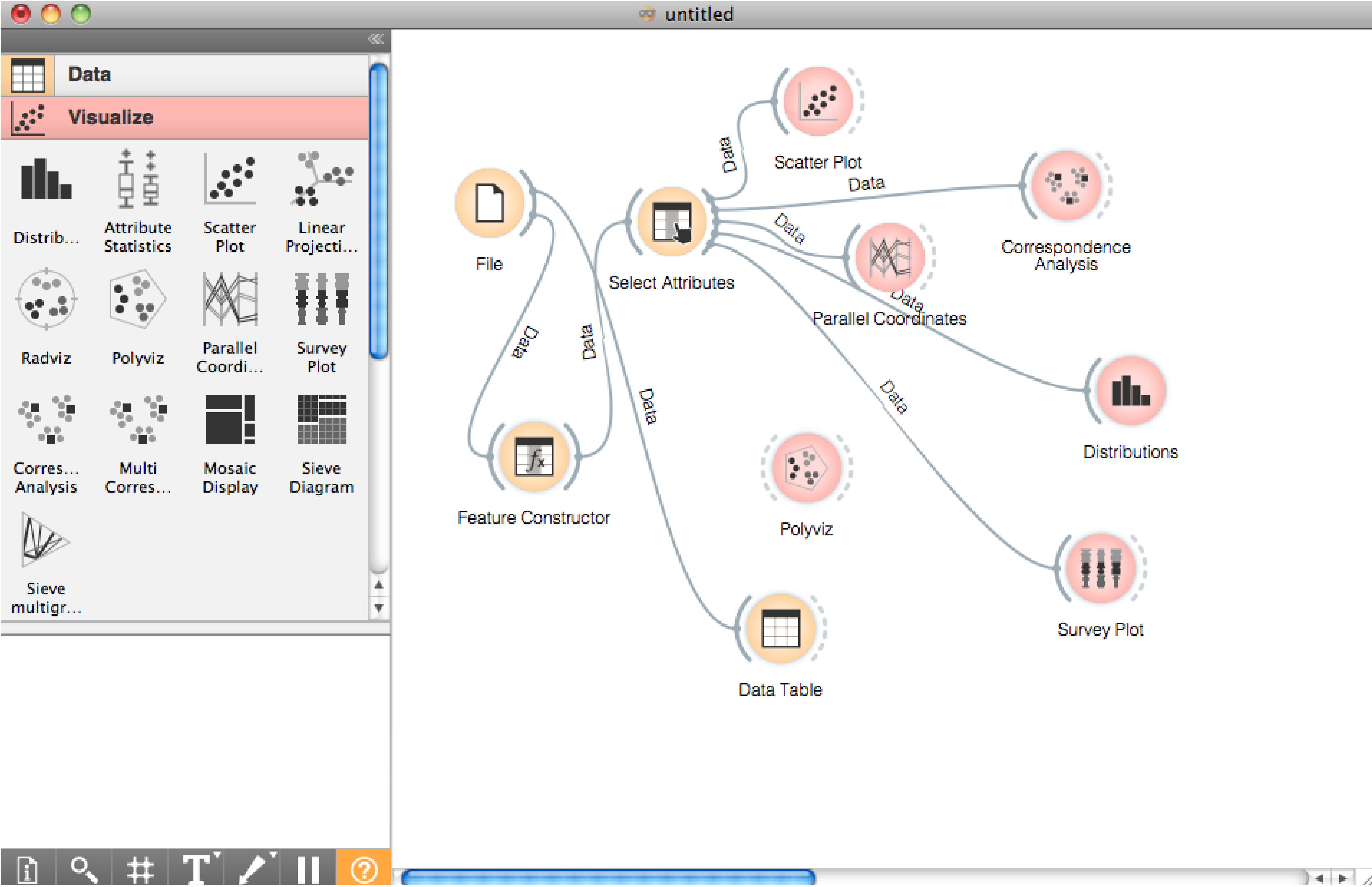


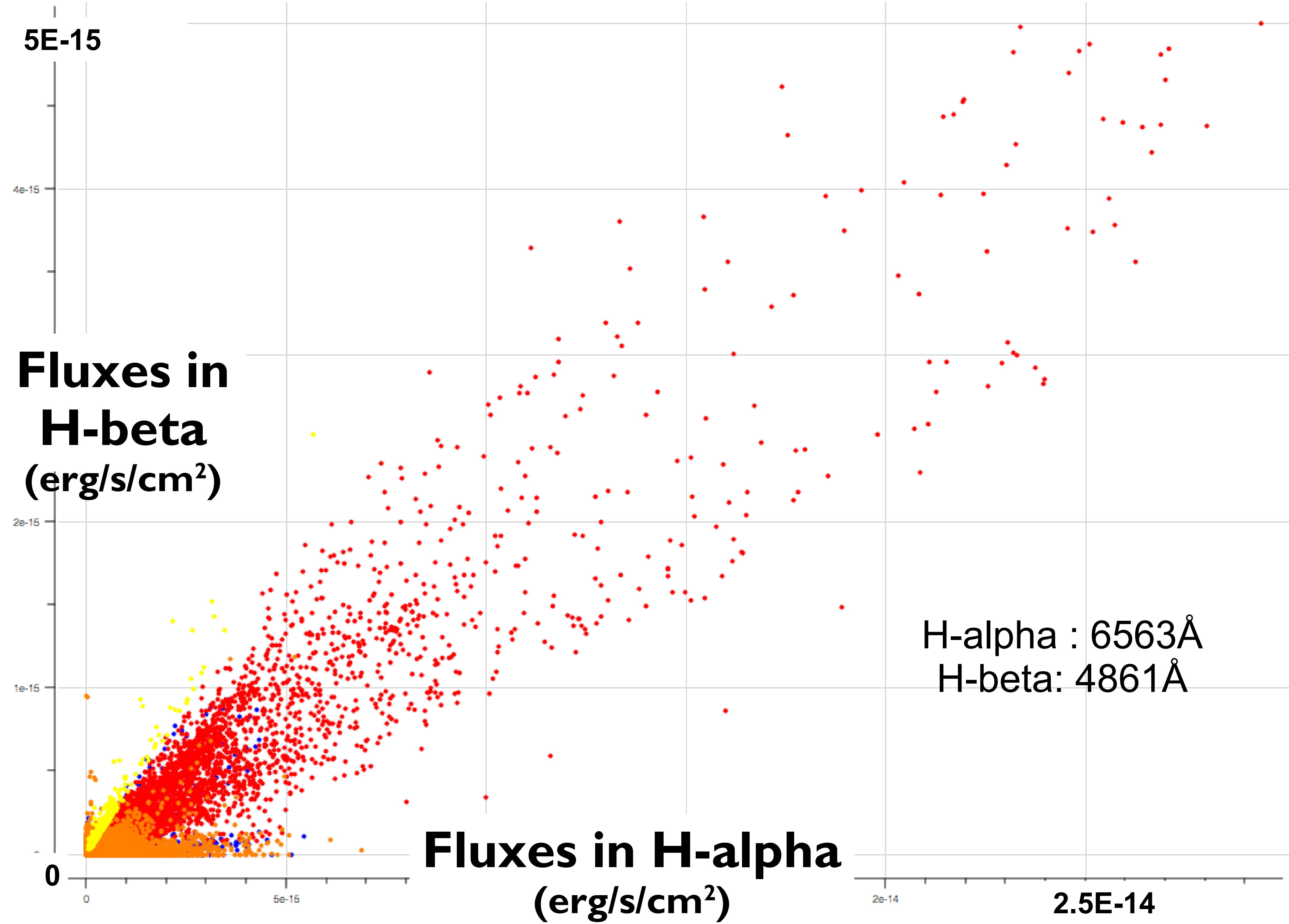


Equivalent Width

orange

DATA MINING
FRUITFUL&FUN


untitled
Data
Visualize
Distrib...
Attribute Statistics
Scatter Plot
Linear Projecti...
Radviz
Polyviz
Parallel Coordi...
Survey Plot
Corres... Analysis
Multi Corres...
Mosaic Display
Sieve Diagram
Sieve multigr...
File
Select Attributes
Scatter Plot
Correspondence Analysis
Parallel Coordinates
Distributions
Feature Constructor
Polyviz
Survey Plot
Data Table
Data


5E-15
4e-15
Fluxes in
H-beta
(erg/s/cm²)
2e-15
1e-15
0
0
5e-15
Fluxes in H-alpha
(erg/s/cm²)
2e-14
2.5E-14
H-alpha : 6563Å
H-beta: 4861Å

How To Use *Orange* Software

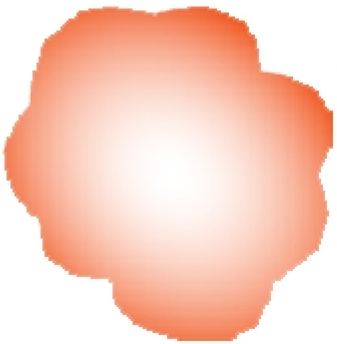
Kevin Hu, Student Intern, Akamai Initiative Workforce

Abstract – Scientists from all fields are constantly obtaining data from their research, sometimes with over thousands of pages worth of information. With this much data, manually extracting specific information is virtually impossible. *Orange: Data Mining Fruitful & Fun* is a computer software program which allows visual organization, representation, and analysis of information. This visualization allows users to understand their obtained data and possibly discover unexpected properties and trends. There are limitless ways to organize data using this program, but only the fundamental yet essential methods including two-way plotting will be explored using our astronomical database as an example.

I. Introduction

Orange: Data Mining Fruitful & Fun, or simply *Orange*, is an open-source software computer program developed at the Bioinformatics Laboratory at the University of Ljubljana, Slovenia [1]. This software allows visual organization and representation of information and databases, essential for scientists who have gathered a plethora of information where reading through all the data is simply impossible. Other powerful data mining tools such as Interactive Data Language or IDL are available, but the one niche *Orange* provides is the targeted audience: non-programmers. Several features include visualization through various plots, classification through methods such as the naïve Bayes classifier and tree builders, evaluation including calibration plots and lift curves, and even Python scripting for elaborate or unavailable algorithms in *Orange*. Despite all the possibilities, only the basic components

Star Formation Rates



Dust Extinction

Chemical Abundance

$$\textbf{Luminosity(H}\boldsymbol{\alpha}\textbf{)} = F(H\alpha)*4\pi d^2$$
(erg/s)

$$\textbf{O3N2} = 8.73 - 0.32\log \frac{F[OIII] / F(H\beta)}{F[NII] / F(H\alpha)}$$

$$\textbf{N2} = 8.9 + 0.57\log \frac{F[NII]}{F(H\alpha)}$$

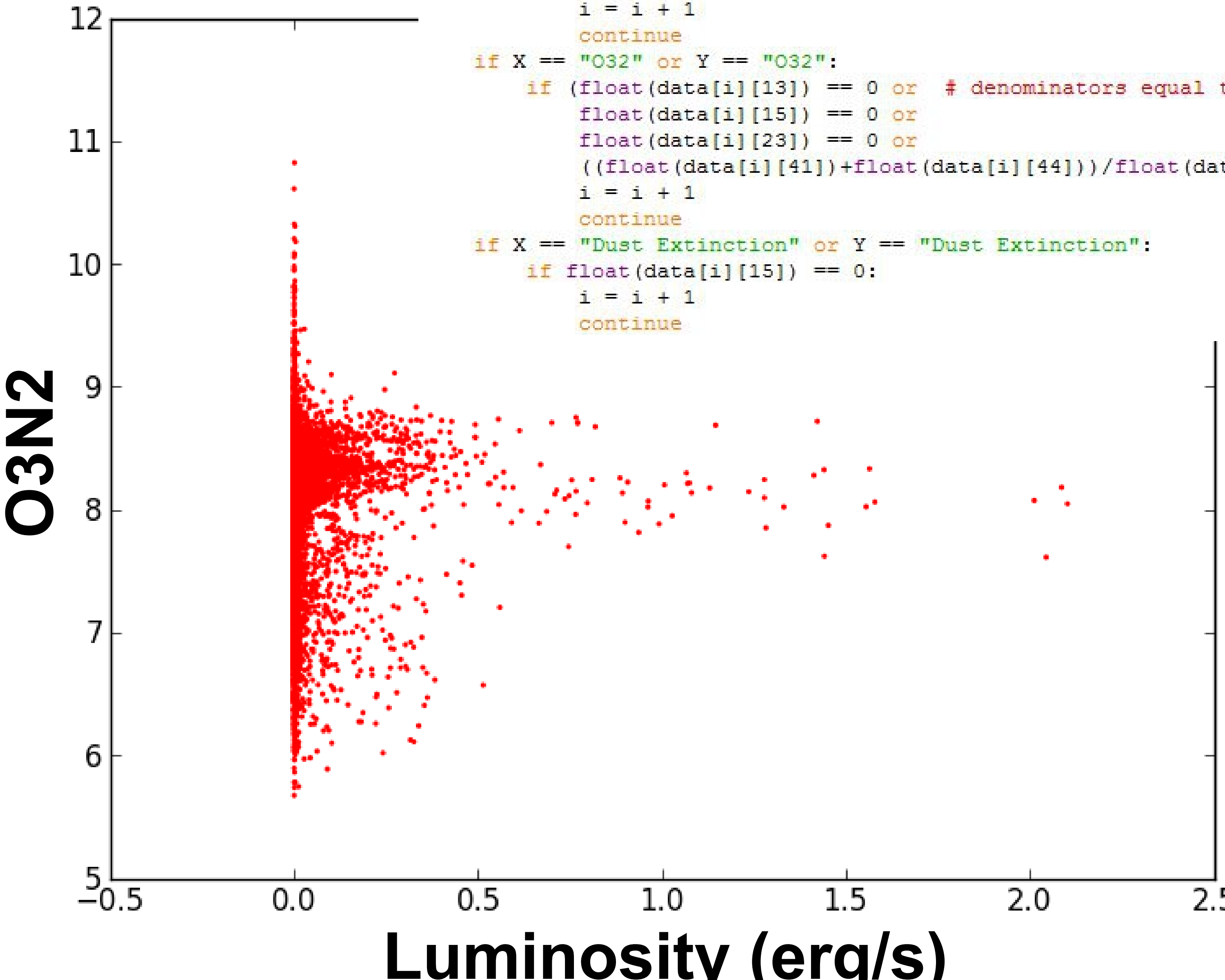
[NII] : 6583Å
[OIII] : 4959 and 5007Å
Hα : 6563Å
Hβ : 4861Å

$$\textbf{Dust Extinction} = \frac{F(H\alpha)}{F(H\beta)}$$

python™

matplotlib

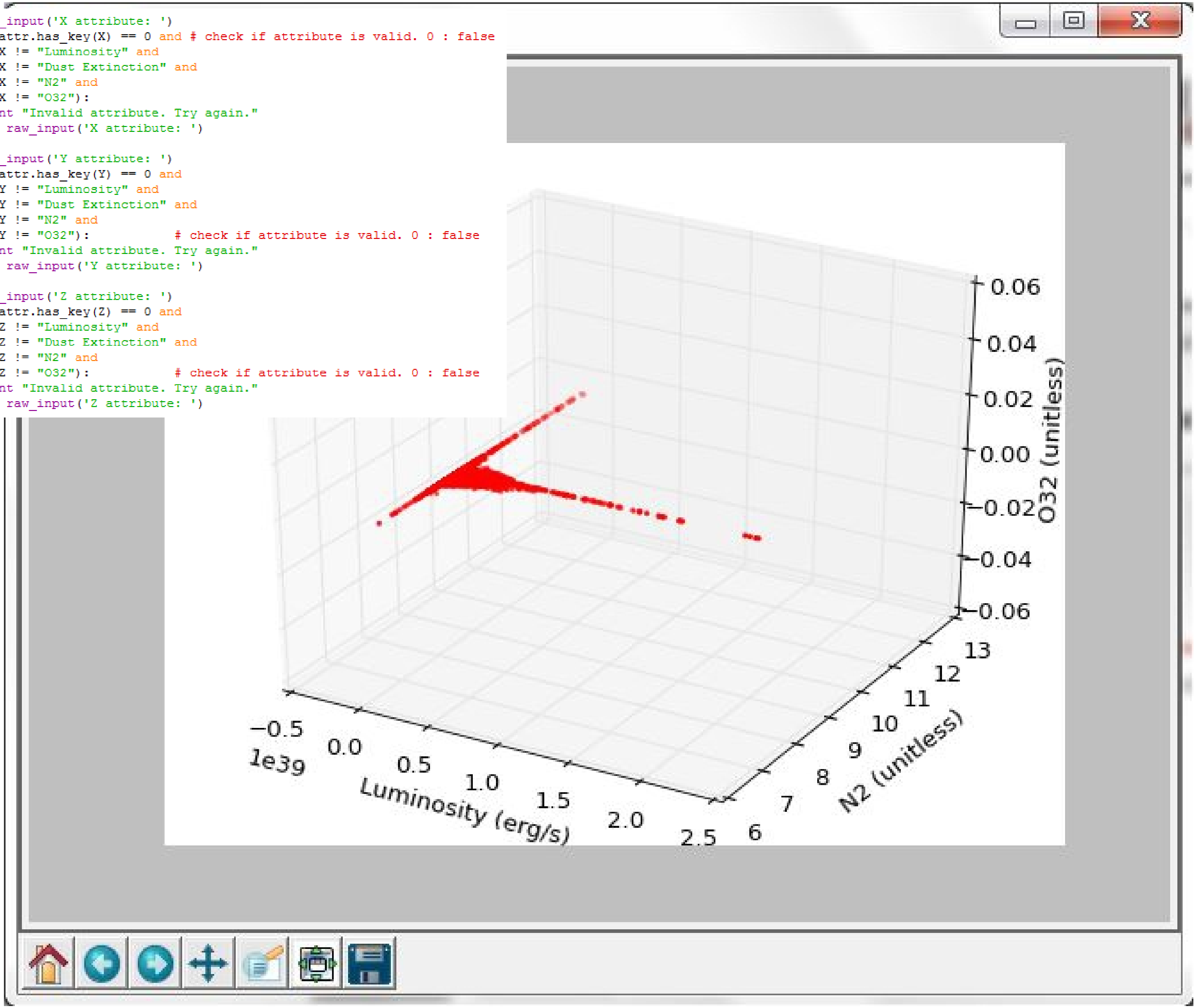
```
i = 1
while i < len(data):
    # ----------- check if value will be invalid; skip onto next -------
    if X == "N2" or Y == "N2":
        if float(data[i][13]) == 0 or float(data[i][23])/float(data[i][13]) < 0:
            i = i + 1
            continue
    if X == "O32" or Y == "O32":
        if (float(data[i][13]) == 0 or  # denominators equal to zero
            float(data[i][15]) == 0 or
            float(data[i][23]) == 0 or
            ((float(data[i][41])+float(data[i][44]))/float(data[i][15]))/(float(
            i = i + 1
            continue
    if X == "Dust Extinction" or Y == "Dust Extinction":
        if float(data[i][15]) == 0:
            i = i + 1
            continue
```



```
X = raw_input('X attribute: ')
while (attr.has_key(X) == 0 and # check if attribute is valid. 0 : false
        X != "Luminosity" and
        X != "Dust Extinction" and
        X != "N2" and
        X != "O32"):
    print "Invalid attribute. Try again."
    X = raw_input('X attribute: ')

Y = raw_input('Y attribute: ')
while (attr.has_key(Y) == 0 and
        Y != "Luminosity" and
        Y != "Dust Extinction" and
        Y != "N2" and
        Y != "O32"):          # check if attribute is valid. 0 : false
    print "Invalid attribute. Try again."
    Y = raw_input('Y attribute: ')

Z = raw_input('Z attribute: ')
while (attr.has_key(Z) == 0 and
        Z != "Luminosity" and
        Z != "Dust Extinction" and
        Z != "N2" and
        Z != "O32"):          # check if attribute is valid. 0 : false
    print "Invalid attribute. Try again."
    Z = raw_input('Z attribute: ')
```



display all data within 0h1m11s by 60° and 0h;1m11s,60°

NGC 6946
NGC 02
NGC 03
NGC 04
NGC 05
NGC 06
NGC 07
NGC 08
NGC 09
NGC 10
NGC 11
NGC 12
NGC 13
NGC 14
NGC 15
NGC 16

Acknowledgments

Work Team:

- Marianne Takamiya
- Daniel Berke
- Forest Bremer
- Aaron McDonald

Akamai Workforce Initiative

Support for the 2013 Akamai Internship Program and the Akamai Workforce Initiative is provided by: the National Science Foundation (AST#0836053); Air Force Office of Scientific Research (FA9550-10-1-0044); University of Hawaii; Thirty Meter Telescope Corp.; University of California, Santa Cruz; Kamehameha Schools.