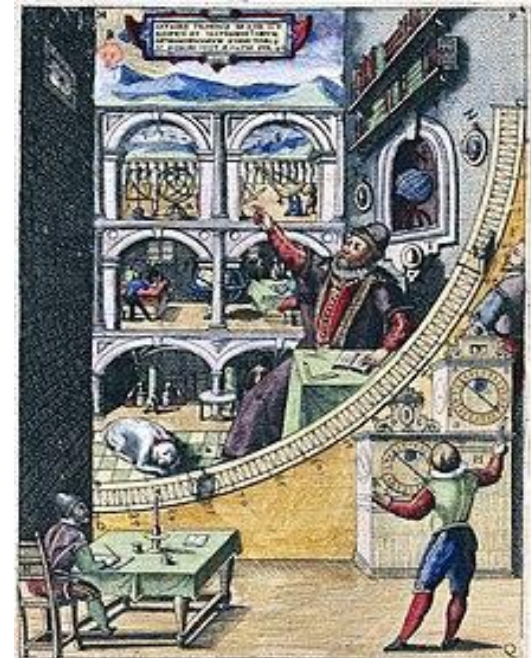




Statistical Techniques for Astronomy

Priya Hasan & Najam Hasan
MANUU, Hyderabad, India
priya.hasan@gmail.com
hasan.najam@gmail.com



Changing perspective

If your experiment needs statistics, you ought to have done a better experiment.

- Ernest Rutherford

God does not play dice with the Universe.

- Albert Einstein

Lies, damned lies and statistics.

- Benjamin Disraeli

We have come to the realisation that in the present Era of rich datasets it is **essential** to have sophisticated **Statistical techniques and methodologies** to analyse and interpret data.

Statistics has changed meaning over centuries:

- Originally it referred to collection & compilation of data.
- In 19th century it accrued the goal of mathematical interpretation of data.
- Contemporary statistics is an amalgam of science, technology and art.

- C.R. Rao (Statistics & Truth)

Astronomy too has changed over the years:

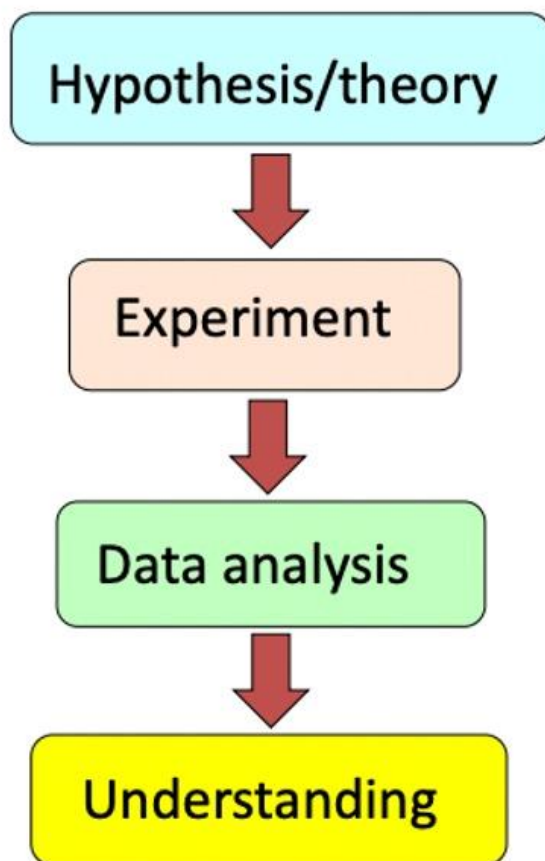
From data poor to data rich science.

We live in era of big astronomical survey:

- **2dF Galaxy Redshift Survey (2dfGRS) – redshift survey conducted by the Anglo-Australian Observatory between 1997 - 2002**
- **Sloan Digital Sky Survey (SDSS) – an optical and spectroscopic survey, 2000–2006**
- **DEEP2 Redshift Survey (DEEP2) – Keck Telescopes to measure redshift of 50,000 galaxies**
- **SAGES Legacy Unifying Globulars and GalaxieS (SAGES Legacy Unifying Globulars and GalaxieS Survey (SLUGGS) survey[8] – a near-infrared spectrophotometric survey of 25 nearby early-type galaxies (2014)**
- **Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST) – an extra-galactic and stellar spectroscopic survey**
- **Optical Gravitational Lensing Experiment (OGLE) – large-scale variability sky survey (in I and V bands), 1992-present**
- **2MASS**
- **GAIA (designed for astrometry: positions, distances and motions of stars with unprecedented precision – 1 billion stars)**

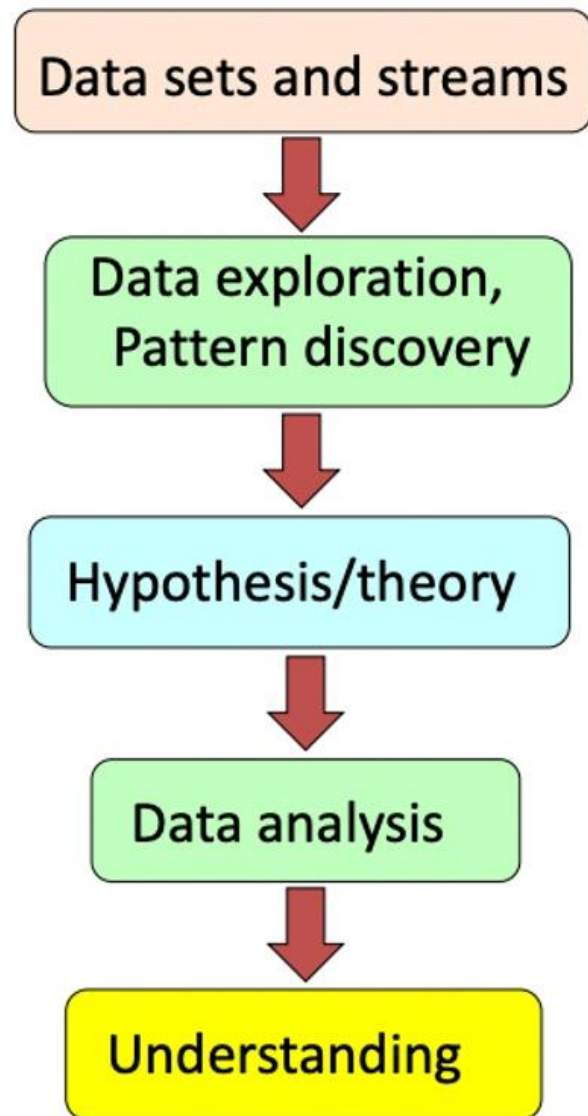
It has become essential for astronomers to handle the big datasets and interpret the complex astrophysical problems by using sophisticated statistical techniques and technology.

Hypothesis-driven science



The two approaches are
complementary

Data-driven science



Virtual Observatory

- A virtual observatory (VO) is a collection of **interoperating data archives and software tools** which utilize the **internet** to form a scientific research environment in which astronomical research programs can be conducted.
- The **main goal** is to allow transparent and distributed **access to data** available worldwide. This allows scientists to **discover, access, analyze**, and combine nature and lab data from heterogeneous data collections in a **user-friendly manner**.
- The IVOA (International Virtual Observatory Alliance) is a standards body created by the VO projects to develop and agree the **vital interoperability standards** upon which the VO implementations are constructed.

Statistics in Astronomy

- Probability theory
- Time Series analysis
- ANOVA
- Correlation & Regression
- Principal Component analysis
- Clustering
- Hypothesis tests
- Point estimation for use in the analysis of modern astronomical data.
- Least squares
- Maximum likelihood, and Bayesian approaches to statistical inference
- Resampling methods - bootstrap
- Goodness of fit

Probability:



Question 1:

We flip an unbiased coin 3 times

All 3 times we get **heads**

What would you expect in the next flip?



Question 2:

These rocks have balanced for about 2,500 million years.

Is it safe to stand under them?

Is it safe to stand under them?

Statistics is **divided into two main branches:**

- **Descriptive statistics**

analysis of data that helps to describe, show and summarize data in a meaningful way.

- **Inferential statistics**

is used to make predictions by taking any group of data in which you are interested.

Statistics – Bayesian vs. Non-Bayesian

Similarities:

- Both forms of statistics use the likelihood function.
- Both forms of statistics are used for statistical inference.
- Both forms of statistics can be used for data analysis and visualization.

Bayesian refers to a statistical approach named after Reverend Thomas Bayes,

He formulated a theorem that describes how to update probabilities based on new data.

In a Bayesian framework:

1. Prior beliefs are represented as a probability distribution (the prior).
2. New data is observed.
3. Bayes' theorem is applied to update the prior beliefs, yielding a posterior distribution.

This approach is used in various fields, including machine learning, data science, and statistics,

Aim:

- Update beliefs based on new data
- Make predictions
- Estimate model parameters
- Perform model selection

Probability in Astronomy:

Before 1987, four naked-eye supernovae had been recorded in ten centuries.

What, before 1987, was the probability of supernova happening in the twentieth century?

Answers:

i. *Supernovae are physical events, in principle, can be accurately determined.*

They are not Random events. Probability is meaningless!

ii. *From frequentist point of view: 4/10*

(assumption: supernovae were equally likely to be reported in the past 10 centuries)

iii. *Apriori argument: In principle we might know the stellar mass function, life-time of a star as a function of mass and star formation rate in the galaxy and the detection efficiency – further take into account metallicity, obscuration due to dust cloud etc.*

Now if we sight supernova in 1987

Approach i: One supernova does not affect another (Can't predict)

Approach ii: Revise the probability to 5/10

Approach iii: Will improve models with fresh data.

Revisiting the Supernova problem:

Data: 4 Supernovae in 10 centuries

Our prior on ρ is uniform between 0 to 1

Suitable Model: Binomial distribution
(occurs or doesn't occur)

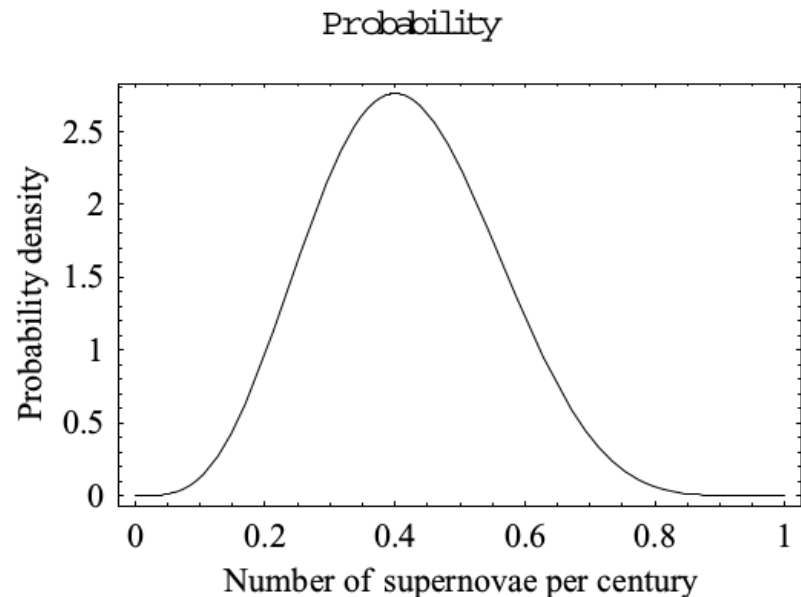
Prosterior probability is

$$\text{prob}(\rho \mid \text{data}) \propto \binom{10}{4} \rho^4 (1 - \rho)^6 \times \text{prior on } \rho.$$

Normalise

$$\begin{aligned} \int_0^1 \text{prob}(\rho \mid \text{data}) d\rho &= 1 \\ \int_0^1 \binom{10}{4} \rho^4 (1 - \rho)^6 d\rho \\ &= \frac{\Gamma(10)\Gamma(4)}{\Gamma(14)} = B[5, 7] \end{aligned}$$

In general $\text{prob}(\rho \mid \text{data}) = \frac{\rho^n (1 - \rho)^{m-n}}{B[n+1, m-n+1]}$



(for $n = 4$, $m = 5$)

Problem: Before 1987, four naked-eye supernovae had been recorded in ten centuries. In 1987 one more supernova was observed. Compare by non-Bayesian and Bayesian methods calculate the probability of supernova happening in the twentieth century?

Results Comparison

Method	Before 1987	After 1987
Frequentist	32.97%	36.5%
Bayesian	37.91%	40.7%

Key Observations

- The Bayesian approach gives higher probabilities as it accounts for uncertainty in λ
- Both methods show increased probability after the 1987 observation
- The frequentist method relies solely on observed frequencies

Principal component analysis (PCA)

When many variables are present – PCA the best correlation searcher

In a sample of ' N ' objects with ' n ' parameters measured for them

PCA is used to find

- what is correlated with what
- which variables produce primary correlation, which secondary....

The task for PCA:

Given a sample of N objects with n measured variables x_n for each, find a new set of ξ_n variables that are orthogonal.

Each one a linear combination of the original variables:

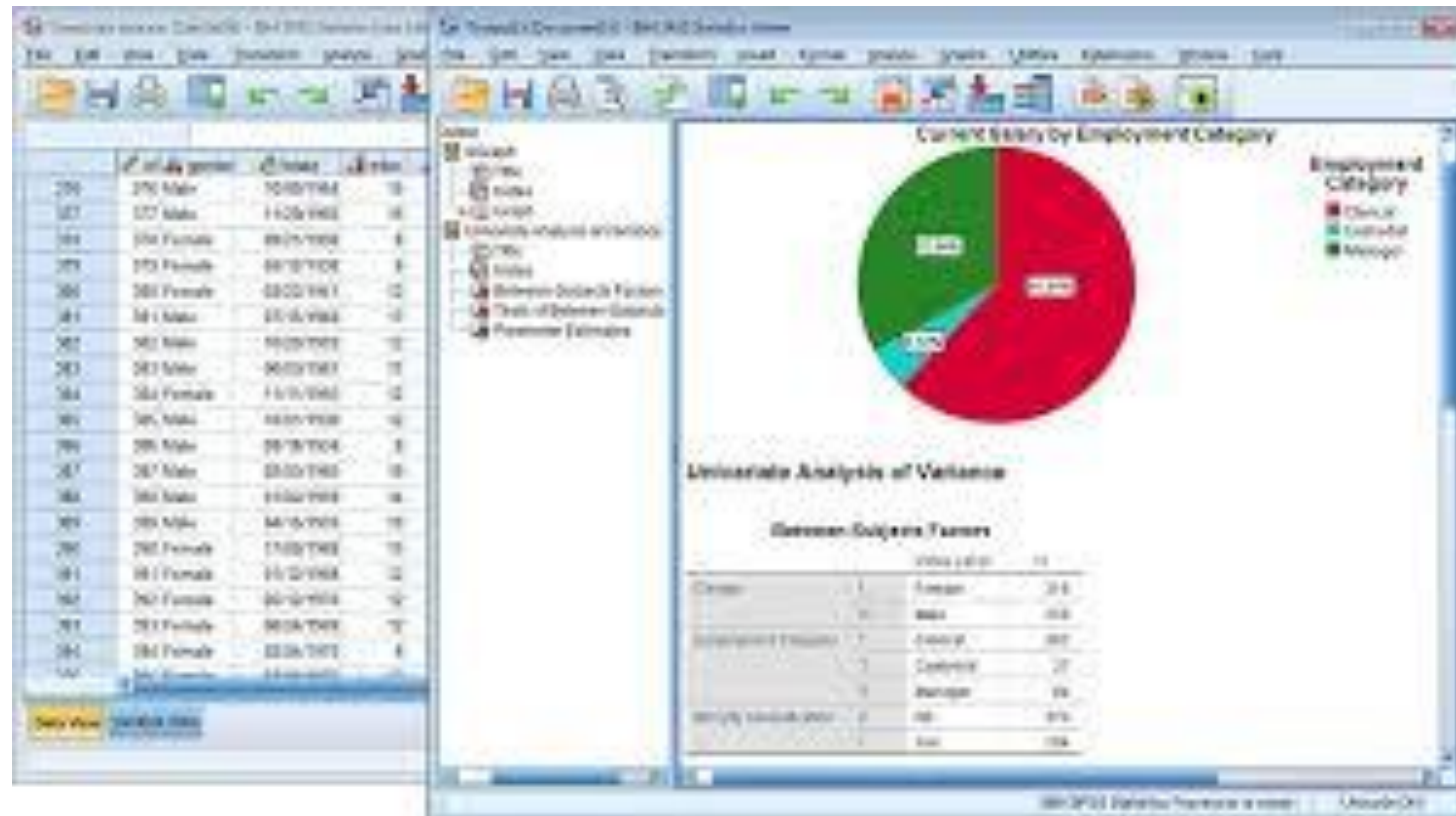
$$\xi_i = \sum_{j=1}^n a_{ij} x_j$$

with values of a_{ij} such that the smallest number of new variables accounts for as much of the variance as possible.

The ξ_i are the principal components.

Statistical Data Analysis – Tools:

- SPSS
- R
- VO tools
- Python



VO Tools

VOTable display and analysis

VOPlot, Topcat

Image display and analysis

Aladin, ESASky

Other standard display tools for downloaded data

Spectrum display and analysis

VOSpec, SpecView

Aladin
is an interactive
software sky atlas
allowing the user
to visualize
digitized
astronomical
images,
superimpose
entries
from astronomical
catalogues or
databases.



New: Aladin release 6 - April 2009

[Measurement browser by interactive histogram, Outreach mode, Full screen, SAMP compatible, RICE compression support, etc...](#)

New: The Aladin manual - April 2009 - The full user manual in English and French...

Description

Aladin is an interactive software sky atlas allowing the user to visualize digitized astronomical images, superimpose entries from astronomical catalogues or databases, and interactively access related data and information from the [Simbad](#) database, the [VizieR](#) service and other archives for all known sources in the field ([see available data](#)). Created in 1999, Aladin has become a widely-used *VO portal* capable of addressing challenges such as locating data of interest, accessing and exploring distributed datasets, visualizing multi-wavelength data. Compliance with existing or emerging VO standards, interconnection with other visualisation or analysis tools, ability to easily compare heterogeneous data are key topics allowing Aladin to be a powerful data exploration and integration tool as well as a science enabler. The *Aladin sky atlas* is available in three modes: a Java Standalone application, a Java applet interface and a simple previewer.

Documentation

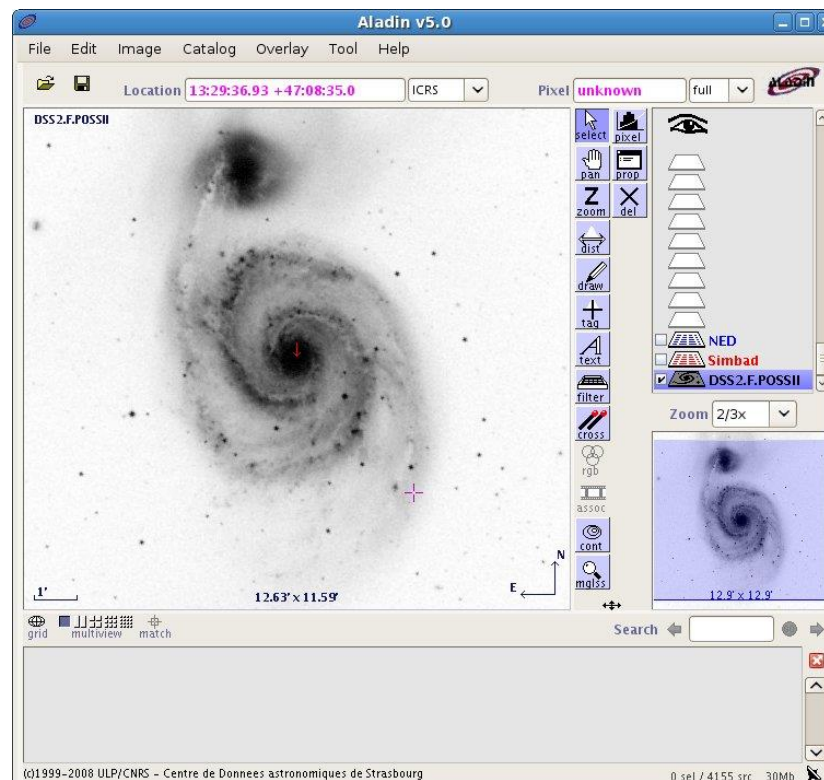
[The Aladin FAQ](#)

[The Aladin user manual](#) (2Mb)

Provide [my data in Aladin](#) ([help form](#))

[The Aladin science case tutorial](#)

[The Aladin filter manual](#)



[filter usage](#)

[y mode](#)

[y side comparison](#)

[lines](#)

from ~10 000 nodes.

developed in the framework of the EuroVO AIDA project (2008-2010).

on, Xmatcher by ellipses, Plastic integration) have been developed in the framework of the

e been developed in the framework of the Astrophysical Virtual Observatory (AVO), an EC

the IDHA project (ACI GRID of the French Ministère de la Recherche).

work, the following acknowledgment would be appreciated: "This research has [100A&AS.143...33B](#).

anbourg - France), from the [CEA](#) (Harvard - USA), from the [ADAC](#) (Tokyo - Japan), from the [OJ](#), or from the [CADC](#) (Victoria - Canada).

CDS/CNRS Contact: [\[icon\]](#)

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

© 2006 The Authors
Journal compilation © 2006 Blackwell Publishing Ltd

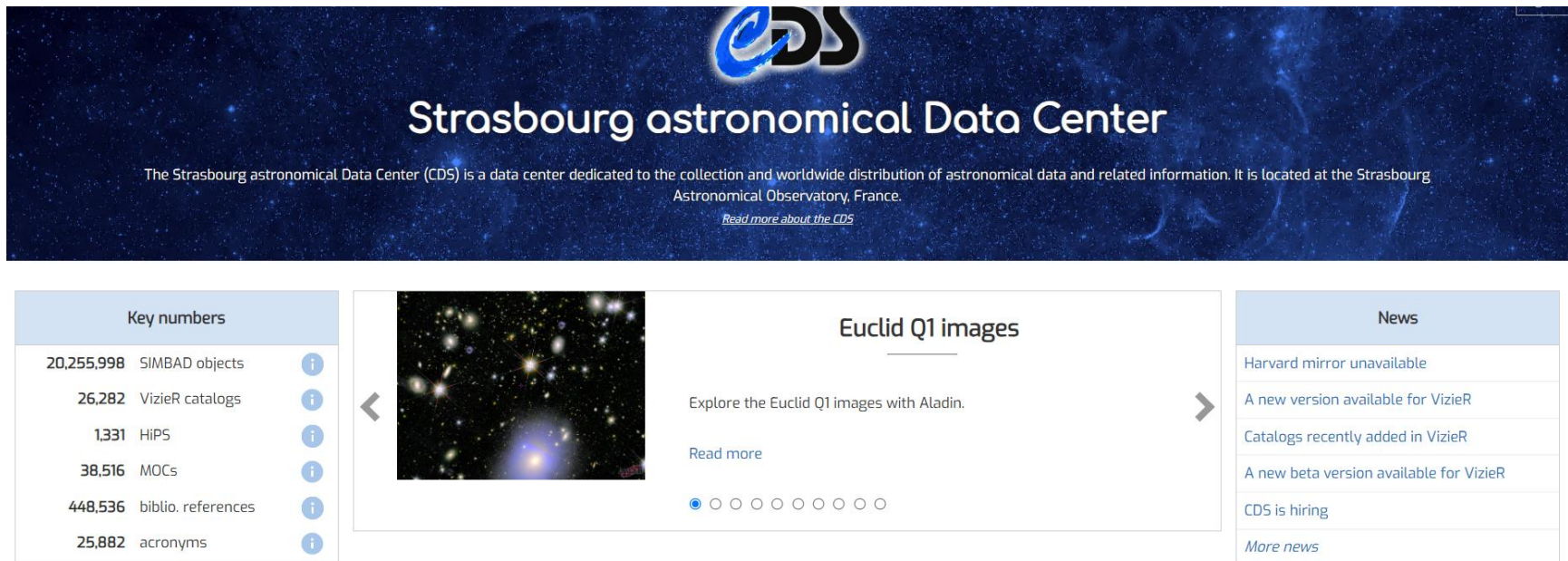
UNHCR on the Web is the WWW interface to the UNHCR database. It offers the following functionalities:

- Information on the value of the test for the null hypothesis of no effect, when the observed test statistic is greater or smaller than 1. The test statistic is used to test the null hypothesis of no effect.

Catalogues

*The **Strasbourg astronomical Data Center (CDS)***

collects and distributes astronomical data catalogues, related to observations of stars and galaxies, other galactic and extragalactic objects, solar system bodies and atomic data.



Strasbourg astronomical Data Center

The Strasbourg astronomical Data Center (CDS) is a data center dedicated to the collection and worldwide distribution of astronomical data and related information. It is located at the Strasbourg Astronomical Observatory, France.
[Read more about the CDS](#)

Key numbers	
20,255,998	SIMBAD objects
26,282	VizieR catalogs
1,331	HiPS
38,516	MOCs
448,536	biblio. references
25,882	acronyms

Euclid Q1 images

Explore the Euclid Q1 images with Aladin.

[Read more](#)

◉ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○

News
Harvard mirror unavailable
A new version available for VizieR
Catalogs recently added in VizieR
A new beta version available for VizieR
CDS is hiring
More news

Simbad contains on 2025.05.28

20,255,998 objects

68,571,124 identifiers

448,536 bibliographic references

46,118,463 citations of objects in papers

15,697 acronyms described for
Simbad

The screenshot shows the SIMBAD Astronomical Database website. At the top, there is a navigation bar with icons for various services: Search, Visual, Table, Alerts, Catalogs, Dictionary, Basic, Tables, and Downloads. Below this, a banner reads "SIMBAD Astronomical Database" and "18-Dec-2009: The archival links are now available for several objects. See details [here](#)." The main content area is divided into three columns of links. The first column, titled "Query", includes links for "basic search", "by identifier", "by coordinates", "by criteria", "reference query", and "scripts". The second column, titled "Documentation", includes links for "User's guide", "Query by id", "Simbad's Dictionary", "Object name", "List of names", "Measurement description", and "Special type coding". The third column, titled "Information", includes links for "Presentation", "Release history", and "Acknowledgment". Below these columns, there are two boxes. The left box, titled "CANONI", provides a description of the database and its capabilities. The right box, titled "Statistics", lists the number of objects, identifiers, bibliographic references, and citations of objects in papers as of 2010.01.01. At the bottom, there is an "Acknowledgment" section with a template text and a "Basic search" section with a search form and links.

Query	Documentation	Information
basic search	User's guide	Presentation
by identifier	Query by id	Release history
by coordinates	Simbad's Dictionary	Acknowledgment
by criteria	Object name	
reference query	List of names	
scripts	Measurement description	
	Special type coding	

CANONI	Statistics
The SIMBAD astronomical database provides basic data, cross-identifications, bibliography and measurements for astronomical objects outside the solar system. SIMBAD can be queried by object name, coordinates and various criteria. Lists of objects and scripts can be submitted. Links to some other on-line services are also provided.	Simbad contains on 2010.01.01
	4,714,823 objects
	18,477,280 identifiers
	237,781 bibliographic references
	6,822,781 citations of objects in papers

Acknowledgment

If the Simbad database was helpful for your research work, the following acknowledgment would be appreciated.

This research has made use of the SIMBAD database, operated at CDS, Strasbourg, France

Basic search

identifier, coordinate (ra/dec= hh:mm:ss, or hh:mm) or hllcode

[SIMBAD search](#) [user](#) [help](#)

[Insert the Simbad basic search in your tool bar](#)

SIMBAD on the Web is the WWW interface to the SIMBAD database. It offers the following functionalities:

- Query by identifier and named identifier
- Query by coordinates, specifying the radius and the equinox
- Query by hllcode and partial hllcode
- Sampling with a set of physical criteria
- Query by lists of objects, coordinates or hllcodes
- Display charts for list of objects resulting from coordinate query


Moreover, the interface provides links with many other data services:

- Links to the other CDS services: Tables, in Visual, giving access to the whole catalogued data, links to Aladin images, surveys and observation logs.
- Through the coordinate to basic data, you can query around an object, using a provided radius.
- Identifiers are linked to the crossmatch dictionary, providing full description of original list, and, when available, offering direct access to the corresponding catalogue in the CDS catalogue service (VizieR and Bessier).
- Every hllcode is a link to the underlying bibliographic information, either at CDS, at ADS, or at the journal site when available. Links to the full text of paper are available in most cases.
- For articles containing tables stored at CDS, the reference provides also a link to the table or catalogue.
- A link to the Messier database (NASA GPC) is proposed when an object has identifier in high energy catalogue.
- FIT measurements contain an anchor which points to the spectra stored in the IRIS database.
- The maps produced by coordinate queries are clickable and return the object information from SIMBAD.


SIMBAD on the Web has all the functionalities provided by XSimbad, after the addition of list queries in March 2001. XSimbad software is then not maintained any more.


VizieR Services


Query service for catalogues, bibliography, cross-matching across catalogues, NED objects, Simbad objects.





EDS
EUROPEAN DATA SERVICE



Simbad



VizieR



Aladin


Catalogs


Dictionary


Biblio


Tutorials


Developers

VizieR Service

JavaScript is used, and should be enabled to get full functionality.

[Browsing through Catalogues - Output Preferences](#)
[FAQ - More about VizieR](#)

Direct access to Catalogues from Name or Designation [\(tips and examples\)](#)
Find Catalogue

Clear

Find catalogues or Data [\(tips and examples\)](#)

Find catalogues among 7880 available

Words matching author's name, word(s) from title, description, etc.

Select from **Wavelength**, **Mission**, and controlled **Astronomical keywords**:

Radio	ANS	AGN
IR	ASCA	Abundances
optical	BeppoSAX	Ages
UV	OGRO	Associations
EUV	COBE	Atomic_Data
X-ray	Chandra	BL_Lac_objects
Gamma-ray	CoRoT	Binaries:cataclysmic

Target Name (resolved by [Simbad](#)) or Position:

J2000

Position in ☒ Sexagesimal, or ☐ Decimal °

Target radius:
 arcmin

☒ Radius or ☐ Box size

Find Catalogues

☐ Select from UCDS

☐ Use [LISTs of Targets](#)

☒ Show footprints
☒ Show all columns
☒ Show column UCDS

Clear

Find Data around Target

Search by Position across 8129 tables

Output preferences [\(usage\)](#)
 Maximum Entries per table:

Output layout:
☒ HTML Table

ALL columns
☐

Reset All

	r	x,y	Position	Galactic	J2000	B1950	EclJ2000	none
Compute	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sort by	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

r and xy are the distance to the Target; Position is in the same coordinate system as Target.

This Bookmark Button will help you for bookmarking: by clicking on this button, the current page, completed with your input, will be reloaded to be safely included into your bookmark or favorite list



Browsing through Catalogues
 Browsing modes via: [Designations](#) · [Acronyms](#) · [Favorites](#) · [Data](#) · [Images/Spectra](#)

This Kohonen Self-Organizing Map is based on a neural network analysis of the keywords associated to the catalogues (see Poinçot et al., [1998A&AS...130..183P](#); and Lesteven et al., [1996VA.....40..393L](#))

Each dot marks a map area; colour denotes the *density* or the *clustering tendency* of the documents; deep blue areas have the lowest density. Just click any area on the map to get the corresponding list of catalogues found in that area.



Other Installations of VizieR
 Some other installation of VizieR could be closer to you, and answer faster:
[Tokyo, Japan](#) · [IUCAA, India](#) · [CADC, Canada](#) · [Cambridge, UK](#) · [CfA/Harvard, USA](#) · [UKIRT-Hawaii, USA](#) · [INASAN, Russia](#) · [Beijing Obs., China](#)

Aladin
is an interactive
software sky atlas
allowing the user
to visualize
digitized
astronomical
images,
superimpose
entries
from astronomical
catalogues or
databases.



New: Aladin release 6 - April 2009

[Measurement browser by interactive histogram, Outreach mode, Full screen, SAMP compatible, RICE compression support, etc...](#)

New: The Aladin manual - April 2009 - The full user manual in English and French...

Description

Aladin is an interactive software sky atlas allowing the user to visualize digitized astronomical images, superimpose entries from astronomical catalogues or databases, and interactively access related data and information from the [Simbad](#) database, the [VizieR](#) service and other archives for all known sources in the field ([see available data](#)). Created in 1999, Aladin has become a widely-used *VO* portal capable of addressing challenges such as locating data of interest, accessing and exploring distributed datasets, visualizing multi-wavelength data. Compliance with existing or emerging VO standards, interconnection with other visualisation or analysis tools, ability to easily compare heterogeneous data are key topics allowing Aladin to be a powerful data exploration and integration tool as well as a science enabler. The *Aladin* sky atlas is available in three modes: a Java Standalone application, a Java applet interface and a simple previewer.

Documentation

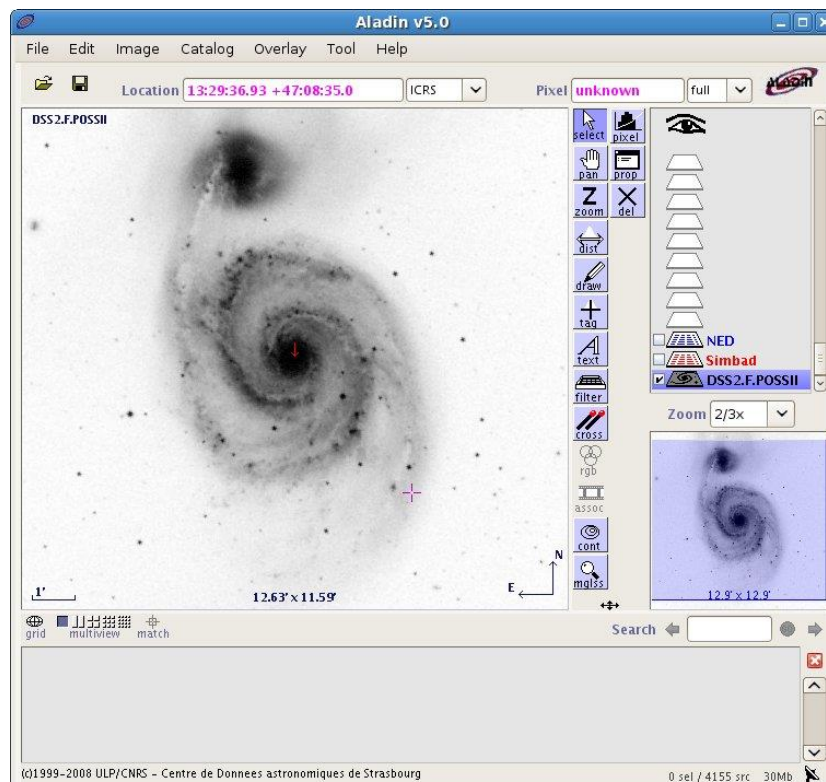
[The Aladin FAQ](#)

[The Aladin user manual](#) (2Mb)

Provide [my data in Aladin](#) ([help form](#))

[The Aladin science case tutorial](#)

[The Aladin filter manual](#)



[filter usage](#)

[y mode](#)

[y side comparison](#)

[lines](#)

from ~10 000 nodes.

developed in the framework of the EuroVO AIDA project (2008-2010).

on, Xmatcher by ellipses, Plastic integration) have been developed in the framework of the

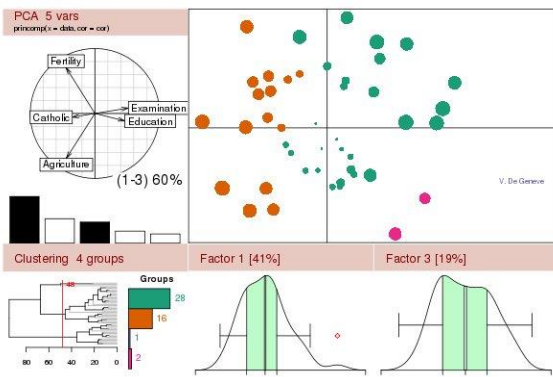
have been developed in the framework of the Astrophysical Virtual Observatory (AVO), an EC

the IDHA project (ACI GRID of the French Ministère de la Recherche).

work, the following acknowledgment would be appreciated: "This research has [100A&AS.143...33B](#).

anbourg - France), from the [CEA](#) (Harvard - USA), from the [ADAC](#) (Tokyo - Japan), from the [CADC](#) (Victoria - Canada).

CDS/CNRS Contact: [\[icon\]](#)



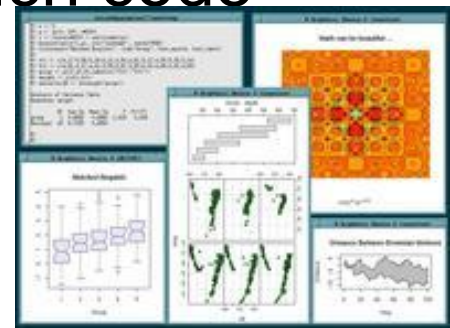
R

R is a free software environment for statistical computing and graphics.

It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS.

It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues.

R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.



Plotting Tools

- **Topcat:** An interactive graphical viewer and editor for tabular data.
- **Python:** Matplotlib, Seaborn, Plotly.....
- **VOPlot** A tool for visualizing astronomical data.
- **VOmegaPlot** or handling large number of points (in the range of millions), similar to VOPlot and both these tools have certain common functionality.
- **STILTS:** processing of tabular data; the package has been designed for, but is not restricted to, astronomical tables such as object catalogues.
- **VOConvert** A tool for converting files from one format to another. It supports following file format conversions: (1) ASCII to VOTable (2) FITS to VOTable and (3) VOTable to ASCII.

Using Python

```
In [1]: matplotlib inline
```

```
In [3]: import pandas as pd
from ydata_profiling import ProfileReport
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [4]: plt.style.use('ggplot')
```

```
In [5]: data=pd.read_csv('dias2018.csv')
```

```
In [6]: data
```

```
Out[6]:
```

	LII	BII	LD	RGC	LOG_AGE	PM_RA	PM_DEC	NUM_CLUSTER_STARS	RAD_VEL	RAD_VEL_ERROR	METAL
0	306.5592	-21.4418	4.012222	8262.035772	9.300	-6.30	-5.41	67	0.00	0.00	0.000
1	313.5830	-30.7574	5.752015	6566.127096	9.700	0.96	-4.82	37	0.00	0.00	0.000
2	292.4817	-21.6535	1.128074	8592.322145	6.900	-30.00	27.80	18	16.10	0.50	0.000
3	307.2214	-16.3409	3.197567	7567.324146	9.350	-2.81	-3.68	122	0.00	0.00	-0.050
4	300.2173	-15.6085	0.596600	8478.627331	6.602	-39.50	-1.00	10	13.00	5.00	0.000
5	291.1373	-35.7047	1.622333	8923.234812	9.300	4.39	3.83	22	0.00	0.00	0.000
6	317.7172	-19.3386	6.498056	9881.110039	9.315	-1.93	-7.56	238	0.00	0.00	0.000
7	305.9925	-8.6218	4.459963	10344.985938	9.100	-2.49	-0.93	77	0.00	0.00	0.000
8	283.7728	-38.2608	7.975018	8007.584768	9.480	2.86	5.71	120	4.96	3.60	0.000
9	308.6803	-8.2906	11.629630	7001.344056	8.300	-3.48	-3.06	670	0.00	0.00	0.000

In [7]: data.describe()

Out[7]:

	LII	BII	LD	RGC	LOG_AGE	PM_RA	PM_DEC	NUM_CLUSTER
count	2166.000000	2166.000000	2166.000000	2166.000000	2166.000000	2166.000000	2166.000000	2166.000000
mean	190.859480	-0.797745	4.799574	8778.562255	7.737534	-1.872461	-1.226519	288.974146
std	98.873882	8.289514	5.558965	1938.267074	2.281574	4.260588	4.251714	802.063735
min	0.009400	-79.261200	0.000000	1369.671399	0.000000	-69.590000	-45.000000	0.000000
25%	107.977725	-2.596550	1.604889	7711.184526	7.506250	-3.417500	-3.380000	23.000000
50%	204.348550	-0.403800	3.198148	8540.524642	8.450000	-1.570000	-1.340000	76.000000
75%	276.256400	1.630850	5.882848	9704.337318	8.900000	0.000000	0.677500	203.000000
max	359.994200	84.591400	75.705981	23360.246914	10.100000	36.390000	27.800000	14341.000000



```
In [8]: pp.ProfileReport(data)
```

Overview

Dataset info

Number of variables	11
Number of observations	2166
Total Missing (%)	0.0%
Total size in memory	186.2 KiB
Average record size in memory	88.0 B

Variables types

Numeric	11
Categorical	0
Boolean	0
Date	0
Text (Unique)	0
Rejected	0
Unsupported	0

Warnings

rejected	0
Unsupported	0

Warnings

- LD has 132 / 6.1% zeros Zeros
- LOG_AGE has 154 / 7.1% zeros Zeros
- METALLICITY has 1868 / 86.2% zeros Zeros
- NUM_CLUSTER_STARS has 65 / 3.0% zeros Zeros
- PM_DEC has 64 / 3.0% zeros Zeros
- PM_RA has 65 / 3.0% zeros Zeros
- RAD_VEL has 1464 / 67.6% zeros Zeros
- RAD_VEL_ERROR has 1493 / 68.9% zeros Zeros

Variables

BII

Numeric

Distinct count 2151
Unique (%) 99.3%
Missing (%) 0.0%
Missing (n) 0
Infinite (%) 0.0%
Infinite (n) 0

Mean -0.79774
Minimum -79.261
Maximum 84.591
Zeros (%) 0.0%



[Toggle details](#)

LOG_AGE

Numeric

Distinct count 587
Unique (%) 27.1%
Missing (%) 0.0%
Missing (n) 0
Infinite (%) 0.0%
Infinite (n) 0

Mean 7.7375
Minimum 0
Maximum 10.1
Zeros (%) 7.1%



[Toggle details](#)

METALLICITY

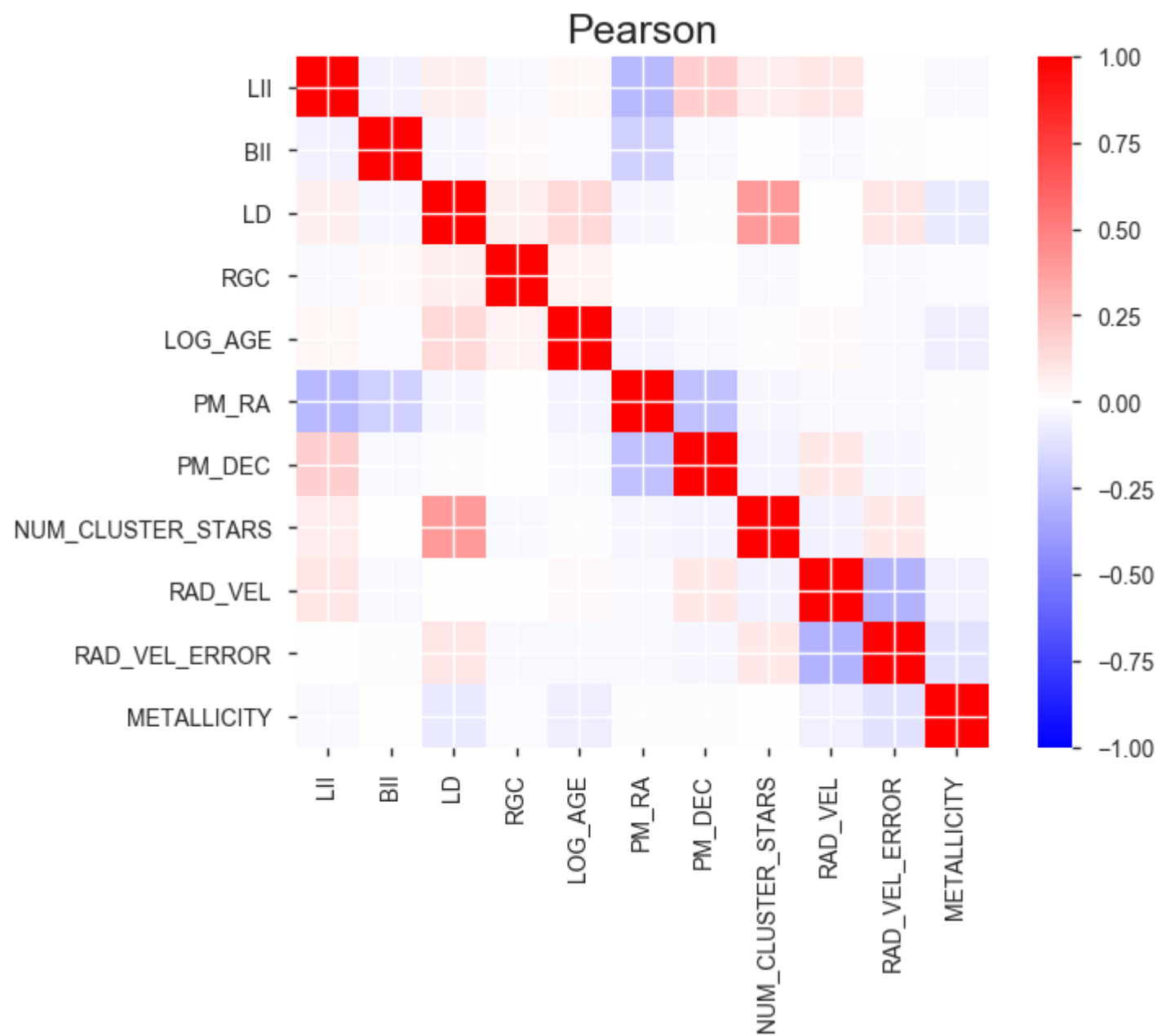
Numeric

Distinct count 128
Unique (%) 5.9%
Missing (%) 0.0%
Missing (n) 0

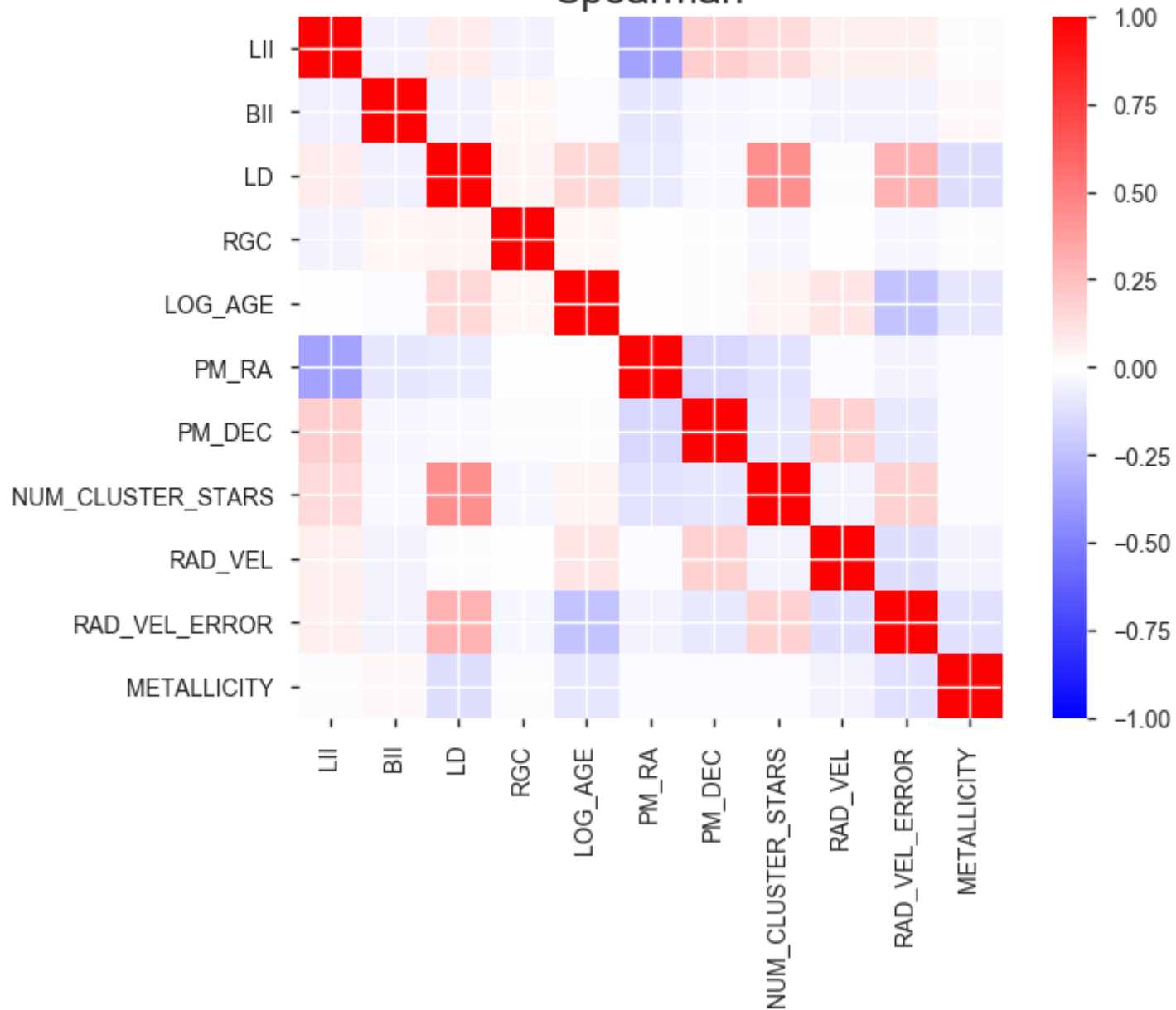
Mean -0.01663
Minimum -1.544
Maximum 0.46
Zeros (%) 86.2%



Correlations



Spearman



Session with python

Thank You