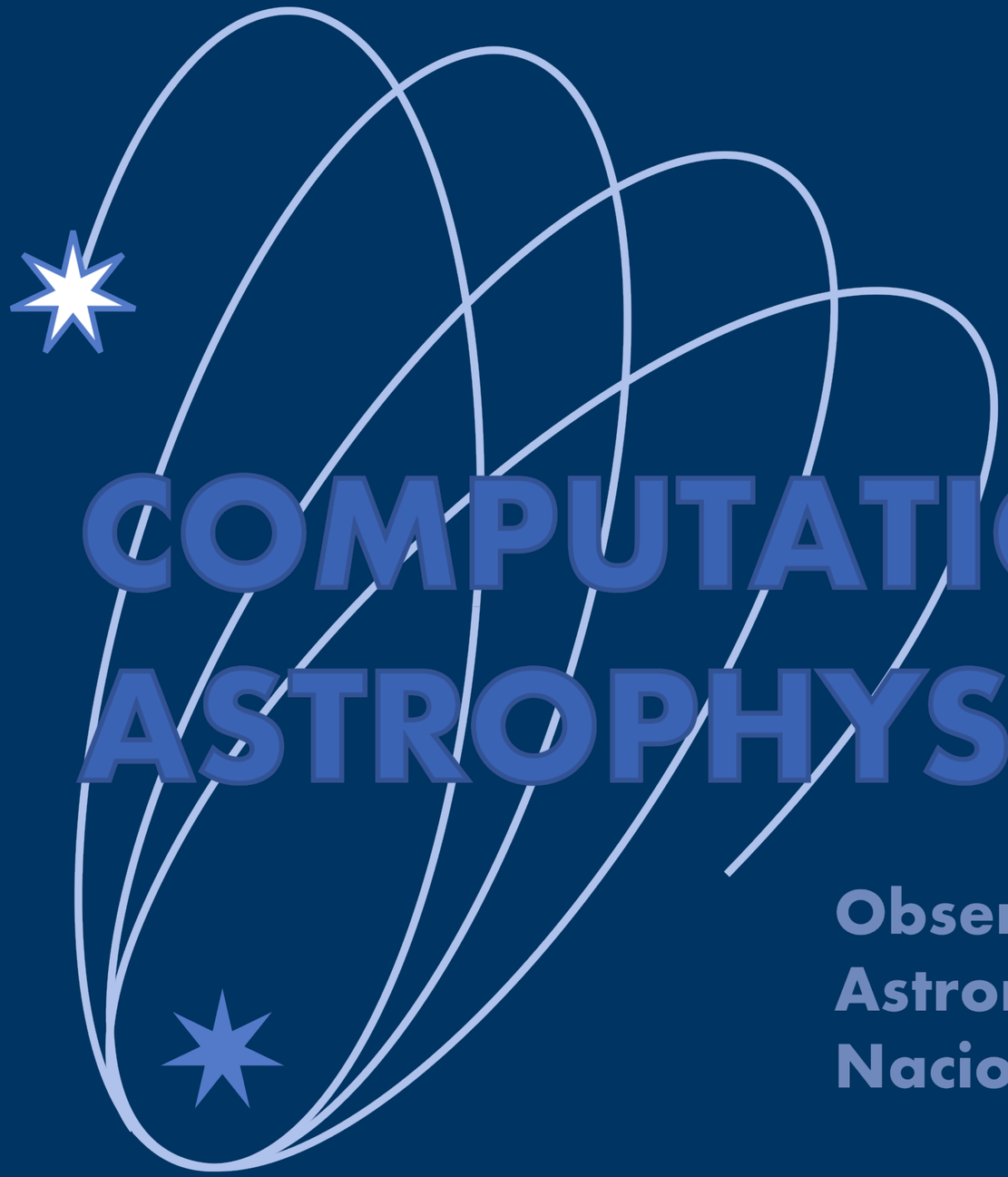

COMPUTATIONAL ASTROPHYSICS

Observatorio
Astronómico
Nacional

Computational Astrophysics 2022

02. Data in Astrophysics

Eduard Larrañaga
Observatorio Astronómico Nacional
Universidad Nacional de Colombia

Knowing the Data

Usual data formats in astrophysics

.txt (plain text)

.fits (Flexible Image Transport System)

.dat + ReadMe

VOTable

.csv

Data files

Feature

Data point or Sample

Name	z	sigma*	e_sigma*	n_sigma*	FWHM	e_FWHM	ogL	e_logL	logM	E_logM	e_logM
SDSS J000805.62+145023.4	0.0454	140.0	27.0		7610	380	41.13	0.04	7.7		0.1
SDSS J004236.86-104921.8	0.0419	78.4	10.0		1960	97	41.58	0.14	6.7		0.
SDSS J011703.58+000027.3	0.0456	98.8	16.0		2270	110	41.45	0.08	6.8		0.1
SDSS J020459.25-080816.0	0.0772	121.0	9.4	a	3720	180	41.13	0.05	7.0		0.1
SDSS J020615.99-001729.1	0.0426	216.0	30.0		3860	190	41.91	0.07	7.5		0.1

Data files

Non-numerical values

Missing values

Name	z	sigma*	e_sigma*	n_sigma*	FWHM	e_FWHM	logL	e_logL	logM	E_logM	e_logM
SDSS J000805.62+145023.4	0.0454	140.0	27.0		7610	380	41.13	0.04	7.7		0.1
SDSS J004236.86-104921.8	0.0419	78.4	10.0		1960	97	41.58	0.14	6.7		0.1
SDSS J011703.58+000027.3	0.0456	98.8	16.0		2270	110	41.45	0.08	6.8		0.1
SDSS J020459.25-080816.0	0.0772	121.0	9.4		3720	180	41.13	0.05	7.0		0.1
SDSS J020615.99-001729.1	0.0426	216.0	30.0		3860	190	41.91	0.07	7.5		0.1

DataBases

SIMBAD - VizieR

Set of
Identifications,
Measurements and
Bibliography for
Astronomical
Data.

DataBases

SIMBAD - VizieR http://simbad.u-strasbg.fr/simbad/

4’500.000 stars

3’500.000 non-stellar objects (galaxies, nebula, clusters, novae, supernovae, etc)

No data for Solar System objects

DataBases

SIMBAD - VizieR http://simbad.u-strasbg.fr/simbad/

Data is used in bibliographical references

DataBases

SIMBAD - VizieR http://simbad.u-strasbg.fr/simbad/

- Type of the object
- Coordinates
- Proper motion
- Radial velocity
- Paralax
- Spectral type
- Redshift
- Integral magnitude
- Flux
- ...

DataBases

SDSS https://www.sdss.org/

Sloan
Digital
Sky
Survey

Data Release 16 (DR16): 9 Dec 2019

DataBases

SDSS https://www.sdss.org/

3D map of the Universe with multicolor (photometric bands) images of one third of the sky and spectroscopical data of more than 3 million objects (stars, galaxies, quasars).

Synthetic Data

We will also use synthetic data in some examples, obtained from known physical models and numerical solution of some systems.

Conociendo los Datos

- Is your data enough ? Do you need more data?
- How many features do you have? Are them too many? Are them enough?
- Are there missing data values ? Is it possible to drop out those missing values?
- What is the question you want to answer from the data analysis? Is the data enough to answer that question?