Lecture 1 -Data Driven Astronomy

- **Observatory Mode (e.g., James Webb)**: Data available to selected researchers initially, then public
- **Survey Mode (e.g., Gaia)**: Continuous sky scanning, massive public datasets
- **Archival Data**: Freely available; includes the world's best datasets from past missions

Types of Data in Astronomy

- **Images**: FITS format, digital pixel arrays (like matrices)
- **Spectra**: Intensity vs. wavelength ⟶ used to derive composition, redshift, temperature
- **Time Series**: For variable stars, exoplanet transits
- **Catalogs**: Tabular data (e.g., Gaia: RA, Dec, motion, parallax)
- **Data Cubes**: Multispectral layers (like a loaf of bread)

Can be from pointed sources like stars(except the Sun) or extended sources(like galaxies)

IN-DETAIL

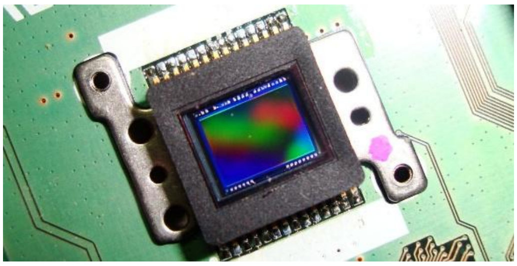
1. IMAGES->

    2D array, where each pixel corresponds to a brightness level

    This is used to create RGB images.

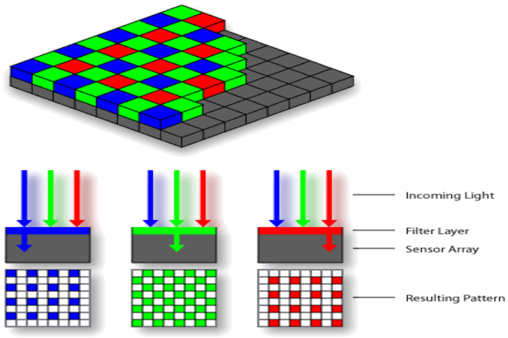
    Multiple images with different filters can be overlayed to create artificial images.

    ➔ CCD-> charge coupled devices

    (In cameras nowadays, CMOS sensors are used but for space photography and other applications CCD are still preferred due to their efficiency)

    Resource: https://www.makeuseof.com/what-is-a-ccd-charge-coupled-device/



(Bayer Filter used on CCD- 50% green, 25% red, 25% blue)

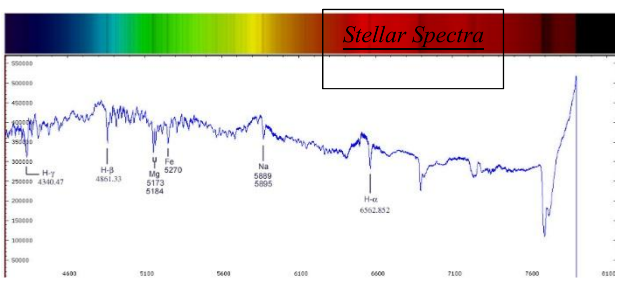
➔ FITS : Flexible Image Transport System

Analogy: astronomy's JPEG but loaded with metadata

- most commonly used digital file format in astronomy.
- **What's Inside a FITS File?**
  - **The Image Data**
    Like a grid of numbers (brightness of stars, galaxies, etc.)
  - **Header Info**
    Telescope name, time of capture, wavelength, coordinates — all saved as text
    Helps scientists know exactly what the image shows
- It can store
  - Single images-> star field from a telescope camera
  - Stacked images->different wavelengths (like X Ray + visible light)
  - Tables-> star positions, brightness, temperature
  - Spectra->Light spread out into wavelengths
- Scientists use **Python** (with astropy) or apps like **DS9, or a FITS viewer** to open and analyze it.

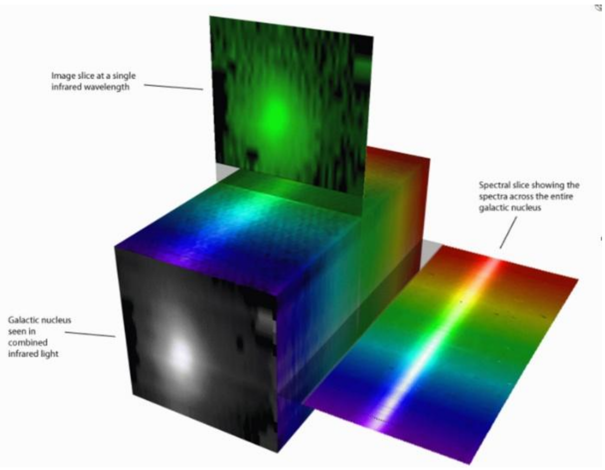
2. SPECTRA

- 1D data
- Like *'fingerprints'*, using them we can find out about :
  - Chemical composition
  - Velocity of source(doppler shift)
  - Thermal Broadening(temperature of source)



3. DATA CUBES – 3D Data

Analogy: A data cube is like a 3D photo of space, showing what the sky looks like in different kinds of light.



Imagine This:

- You're looking at the same patch of sky 🌌
- But instead of just one image, you have **many layers** stacked up
- Each layer = the sky seen with a different filter (like X-ray, infrared, visible)

So, it's like a **3D cube**:

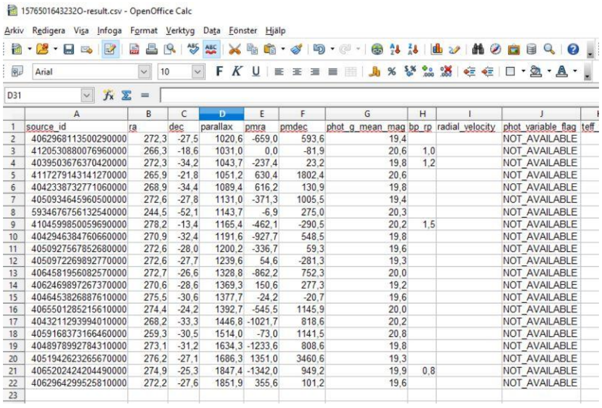
- **Width** × **Height** = the sky area (like a normal image)
- **Depth** = different types of light (or time)

Dimension	Meaning
X	Left to right on the sky
Y	Up and down on the sky
Z	Wavelength or frequency

You can find hidden things that only show up in **certain wavelengths**-> gas clouds, galaxy structures, chemical composition.

4. CATALOG DATA

Tabular data (e.g., Gaia: RA, Dec, motion, parallax)

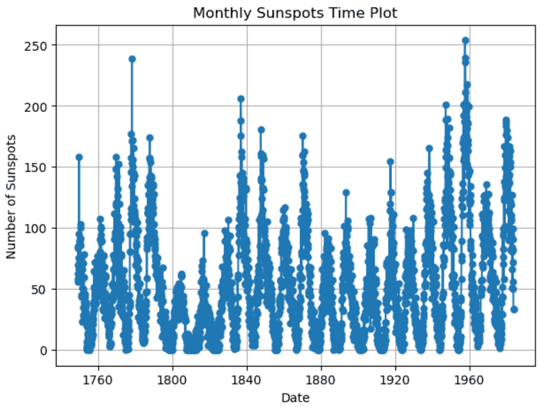


15765016432320-result.csv - OpenOffice Calc

Arkiv Redigera Visa Infoga Format Verktyg Data Fönster Hjälp

	source_id	ra	dec	parallax	pmra	pmdec	phot_g_mean_mag	bp_rp	radial_velocity	phot_variable_flag
2	4062968113500290000	272,3	-27,5	1020,6	-659,0	593,6	19,4			NOT_AVAILABLE
3	4120530880076960000	266,3	-18,6	1031,0	0,0	-81,9	20,6	1,0		NOT_AVAILABLE
4	4039503676370420000	272,3	-34,2	1043,7	-237,4	23,2	19,8	1,2		NOT_AVAILABLE
5	4117279143141270000	265,9	-21,8	1051,2	630,4	1802,4	20,6			NOT_AVAILABLE
6	4042338732771060000	268,9	-34,4	1089,4	616,2	130,9	19,8			NOT_AVAILABLE
7	4050934645960500000	272,6	-27,8	1131,0	-371,3	1005,5	19,4			NOT_AVAILABLE
8	5934676756132540000	244,5	-52,1	1143,7	-6,9	275,0	20,3			NOT_AVAILABLE
9	4104599850059690000	278,2	-13,4	1165,4	-462,1	-290,5	20,2	1,5		NOT_AVAILABLE
10	4042946384760660000	270,9	-32,4	1191,6	-927,7	548,5	19,8			NOT_AVAILABLE
11	4050927567852680000	272,6	-28,0	1200,2	-336,7	59,3	19,6			NOT_AVAILABLE
12	4050972269892770000	272,6	-27,7	1239,6	54,6	-281,3	19,3			NOT_AVAILABLE
13	4064581956082570000	272,7	-26,6	1328,8	-862,2	752,3	20,0			NOT_AVAILABLE
14	4062469897267370000	270,6	-28,6	1369,3	150,6	277,3	19,2			NOT_AVAILABLE
15	4046453826887610000	275,5	-30,6	1377,7	-24,2	-20,7	19,6			NOT_AVAILABLE
16	4065501285215610000	274,4	-24,2	1392,7	-545,5	1145,9	20,0			NOT_AVAILABLE
17	4043211293994010000	268,2	-33,3	1446,8	-1021,7	818,6	20,2			NOT_AVAILABLE
18	4059168373166460000	259,3	-30,5	1514,0	-73,0	1141,5	20,8			NOT_AVAILABLE
19	4048978992784310000	273,1	-31,2	1634,3	-1233,6	808,6	19,8			NOT_AVAILABLE
20	4051942623265670000	276,2	-27,1	1686,3	1351,0	3460,6	19,3			NOT_AVAILABLE
21	4065202424204490000	274,9	-25,3	1847,4	-1342,0	949,2	19,9	0,8		NOT_AVAILABLE
22	4062964299525810000	272,2	-27,6	1851,9	355,6	101,2	19,6			NOT_AVAILABLE

5. TIME SERIES



TOOLS

1. ds9 – image visualization
   - *https://sites.google.com/cfa.harvard.edu/saoimageds9/download?authuser=0*
2. TOPCAT - tool that helps astronomers explore and analyze huge star catalogs using tables, plots, and filters—without needing to code.
3. Programming languages like Python, modules, libraries, etc.

METADATA

Metadata is **data about data**—it describes what's inside a file, like the time, location, telescope, or settings used to capture an image or measurement.

ASTRONOMY ARCHIVE

- Organized, systematic information about the sky above us
- The amount of data is huge…. In terabytes and petabytes

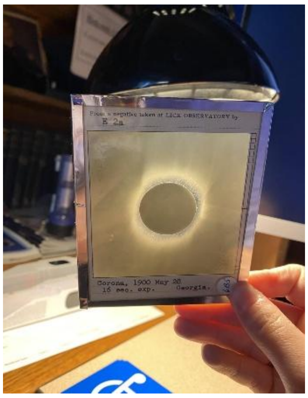
PALOMAR **Observatory Sky Survey (POSS)**

The **Palomar Sky Survey** was a major project started in the 1950s to **photograph the entire night sky** from the northern hemisphere using **large photographic plates** at the **Palomar Observatory** in California.

- It used a 48-inch Schmidt telescope.
- Thousands of glass plates were taken over years, each covering a small patch of the sky.
- The survey created one of the **first complete visual maps of the sky**, which astronomers studied for decades.

Later, these plates were scanned and **digitized into DPOSS**, making the data easier to analyze using computers.

In this survey, photographic plates were used.

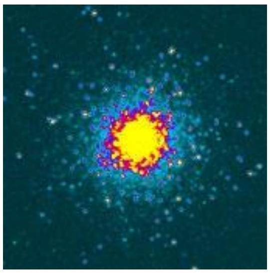


CCD is much better, photographic plates are like old films,

More exposure in both cases :

- Photographic images flatten out
- Since CCD use more exposure to create electrical signals, it only increases their efficiency

DPOSS image of **NGC 6229:**



GAIA

Gaia is a space telescope launched by the European Space Agency (ESA) to map the positions, distances, and motions of over 1 billion stars in our galaxy.

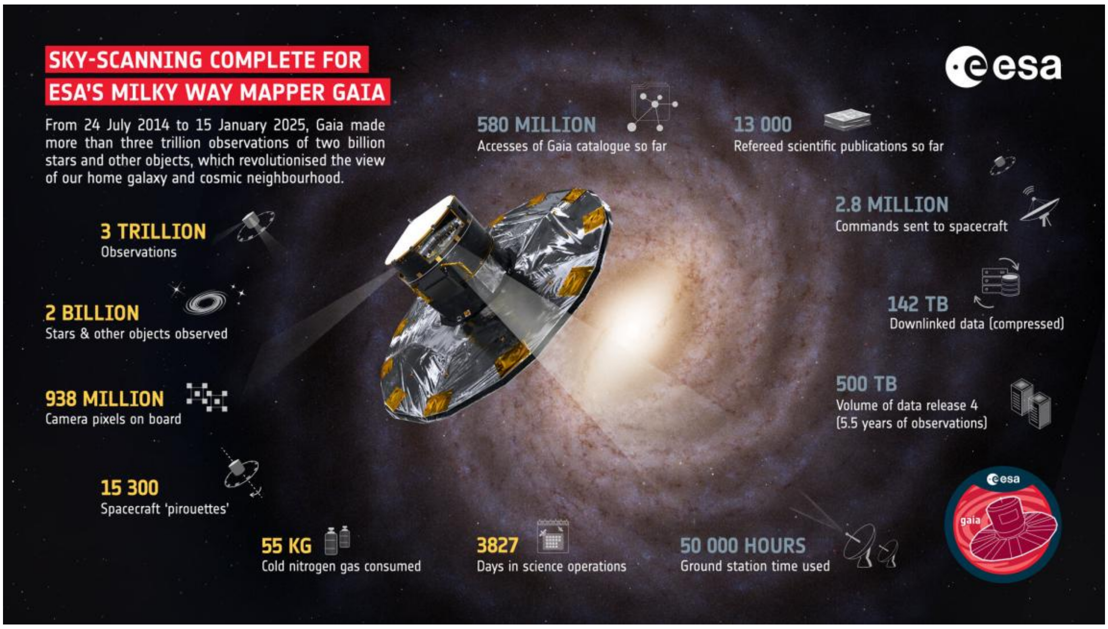
- It creates a 3D map of the Milky Way
- Measures stars’ brightness, color, and motion with extreme accuracy
- Helps scientists understand the structure and history of our galaxy

Gaia produces one of the largest and most precise star catalogs ever made, used widely in astronomy today.

6 D Revolution:

Gaia doesn’t just take pictures of stars—it tells us their full position and motion in 6 dimensions:

Dimension	What it means
X, Y, Z	Where the star is in space (3D position)
Vx, Vy, Vz	How the star is moving (3D velocity)



GMRT (Giant Metrewave Radio Telescope), India

The GMRT is one of the world's largest radio telescope arrays, located near Pune, India.
It studies the universe using low-frequency radio waves, helping us understand galaxies, black holes, and cosmic signals.

Chandra X-ray Observatory

Chandra is a NASA space telescope that observes the universe in X-rays, especially from black holes, neutron stars, and hot gas clouds.
It gives us deep insight into high-energy events invisible to normal telescopes.

SDSS (Sloan Digital Sky Survey)

SDSS is a project that has mapped millions of stars and galaxies using optical and spectroscopic data.
It created some of the most detailed 3D maps of the universe ever made, revolutionizing modern astronomy.

Uses ground-based telescopes.

VIRTUAL OBSERVATORY

Virtual Observatory (VO)

A Virtual Observatory is an online system that lets scientists access and analyze astronomy data from many telescopes around the world as if it's all in one place.

IVOA (International Virtual Observatory Alliance)

IVOA is the global group that sets standards so different observatories and databases can work together in the virtual observatory system.

IN VO-India

VO-India is India's virtual observatory project that helps Indian scientists access, analyze, and contribute to global astronomy data using VO tools and platforms. All individual VOAs are self-funded.

CITIZEN SCIENCE PROJECTS

1. Asteroid Search (*astrometrica*)
2. Planet Hunting
3. Galaxy Cruise
4. Exoplanets(using Kepler/TESS)

1- 🌎 **Project: Using TESS Data to Find Exoplanets (Light Curve Analysis)**

This project uses data from **TESS** (Transiting Exoplanet Survey Satellite), which observes thousands of stars to detect dips in brightness.
These tiny dips, seen in the **light curves**, may indicate a planet passing in front of the star—called a **transit**.

By analyzing these curves using Python (e.g., with lightkurve or AstroPy), you can:

- Detect periodic dimming
- Estimate the planet's size and orbit
- Confirm potential **exoplanet candidates**

In short:

The project finds new worlds by looking for tiny shadows in star brightness data collected by TESS.

2- 🌠 **Project: Finding Exoplanets with Gravitational Microlensing**

Gravity is like a lens- it bends light

In this project, we simulate how gravitational microlensing can be used to detect hidden exoplanets.

- When a massive object (like a star) passes in front of another star, its gravity bends and magnifies the light—this creates a brightening curve called a microlensing event.
- If there's a planet orbiting the lens star, it causes a small, sharp distortion in the curve.

By plotting and analyzing this light curve, we can:

- Understand how magnification changes over time
- Recognize signs of a planet through tiny "blips" in the smooth curve