ANALYZING GALAXIES: STAR FORMATION AND MORPHOLOGY

Jana Bogdanoska
Assistant Professor of Astronomy and Astrophysics
Ss Cyril and Methodius University in Skopje, N. Macedonia

janabogdanoska@pmf.ukim.mk

EXPLORING THE JWST DEEP FIELD

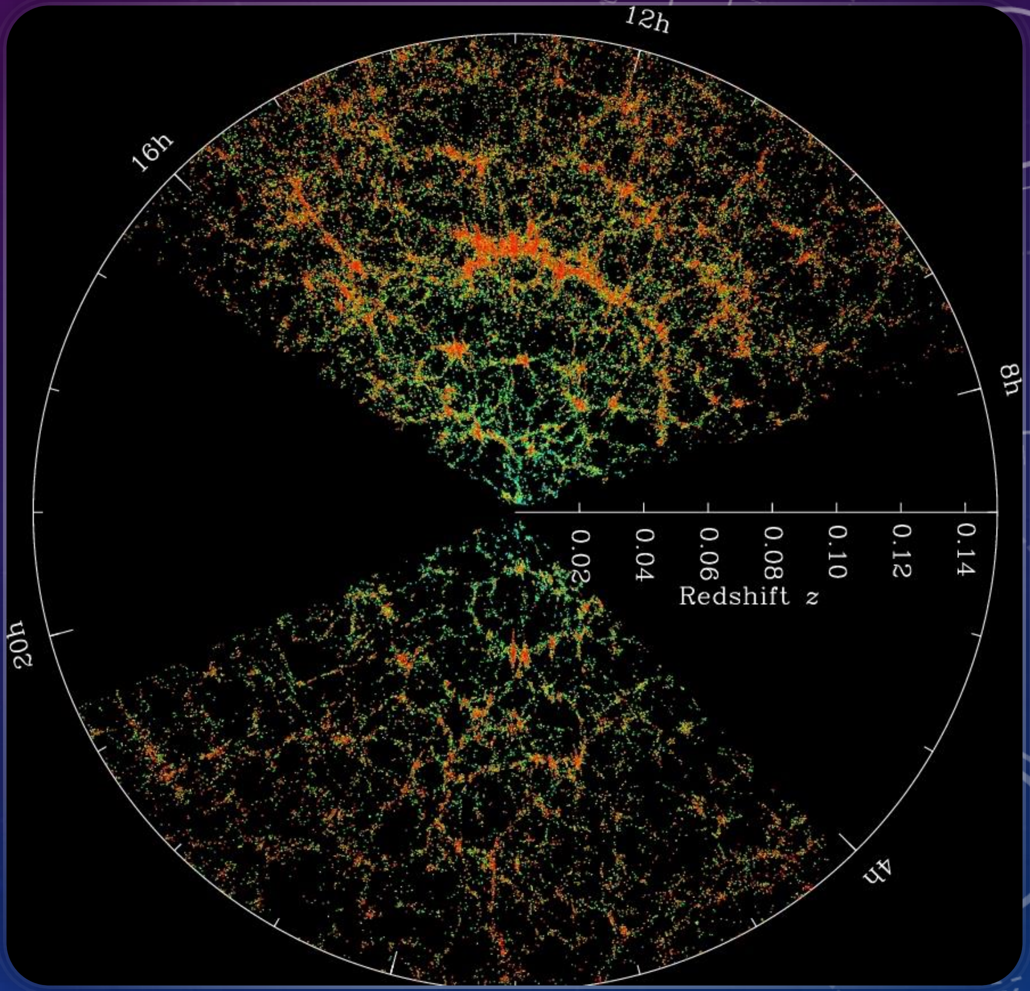
- A tiny patch of sky (size of a grain of sand at arm's length) reveals thousands of galaxies
- They stretch across cosmic time — from nearby to some of the earliest galaxies ever formed.
- The bright points with diffraction spikes are stars in our Milky Way (foreground).

Discussion Prompts:

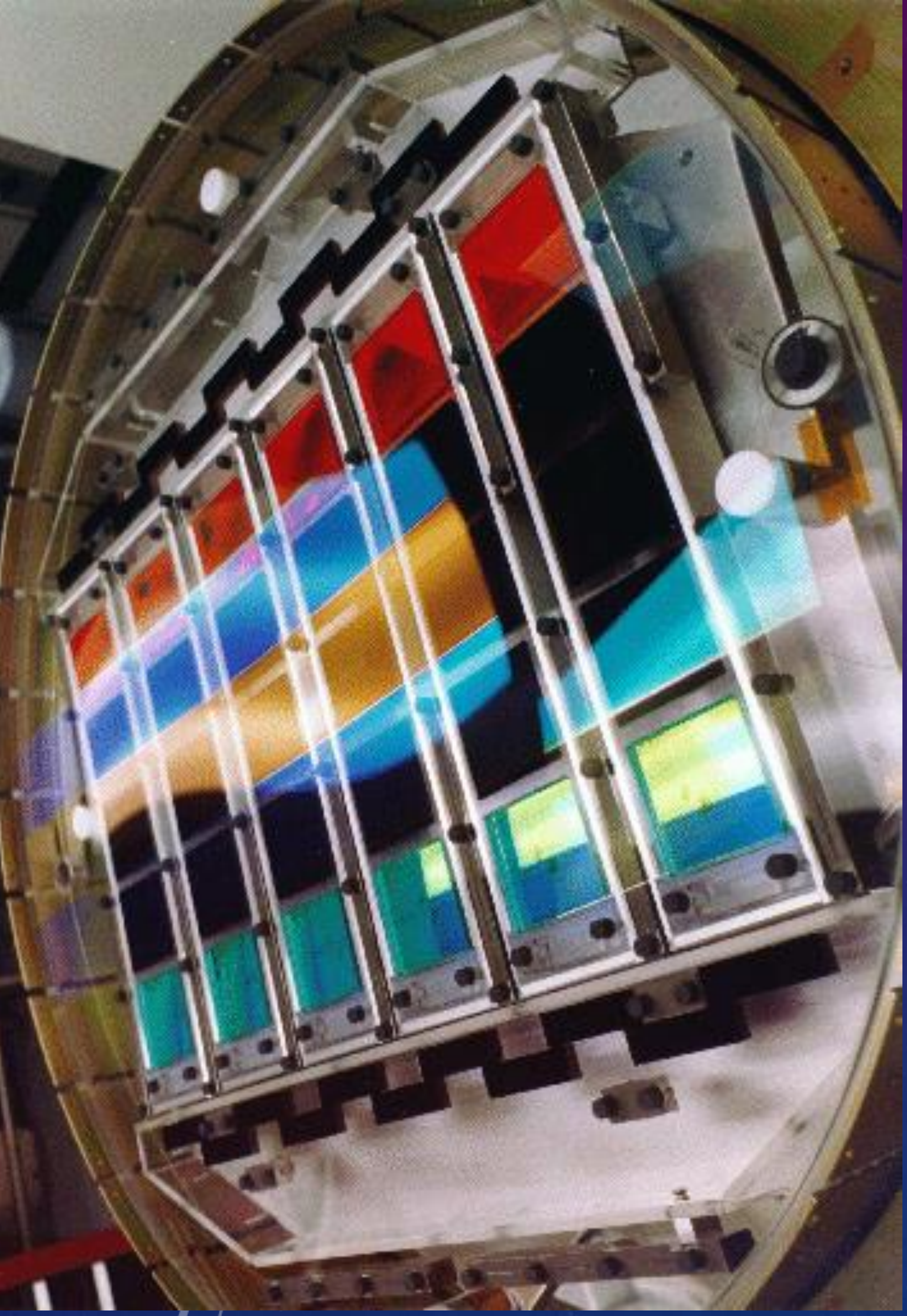
1. What different shapes and patterns do you notice?
2. How might we group or classify them?
3. Why are some galaxies redder or bluer?
4. Which ones seem older or younger?
5. What other data would astronomers want beyond an image?

MEETING THE DATA – SLOAN DIGITAL SKY SURVEY (SDSS)

- SDSS is **giant sky-mapping project** with a telescope in New Mexico, USA.
- Provides **images + measurements** of millions of galaxies, stars, and quasars.
- One of the **most important tools** for studying galaxy formation and change.
- **For each galaxy, SDSS gives us:**
  - Brightness in different **colors (filters)**
  - **Distance** (redshift)
  - **Shape** (spiral or elliptical)



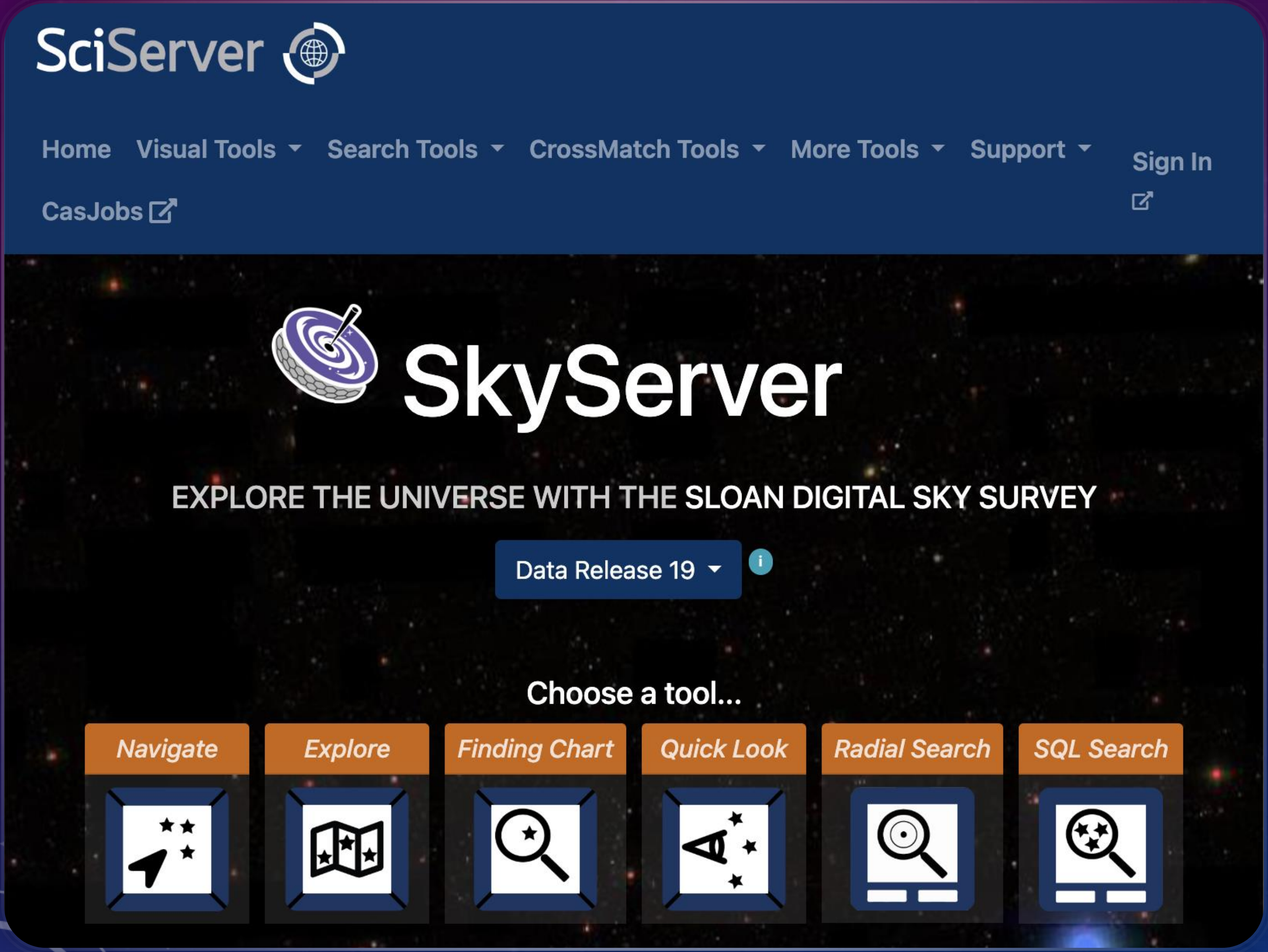
In this workshop, we'll use a **small sample** to see how color & shape link to star formation.



SDSS FILTERS AND IMAGING

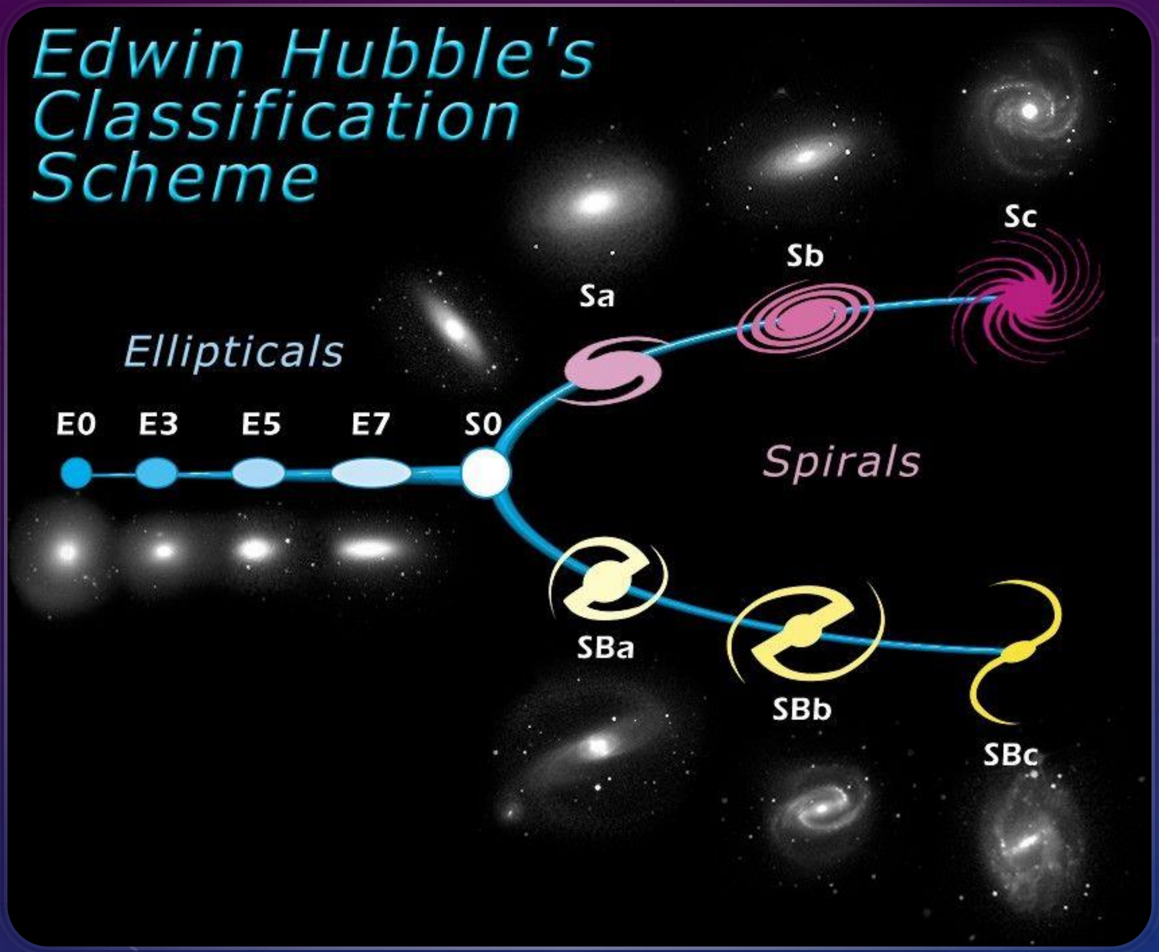
- SDSS takes not just one image → five in different filters:
  - u → ultraviolet (invisible)
  - g → green
  - r → red
  - i, z → infrared (beyond our eyes)
- Combining filters = more information + colorful galaxy images
- We'll use the u–r color (blue vs. red) to group galaxies
- We will use a sample of SDSS galaxy data, which includes:
  - Galaxy ID (name)
  - Morphology (spiral/elliptical)
  - Redshift (distance)
  - Brightness in different filters
  - Galaxy radius

Image: An illustration of SDSS camera and filters.



ACCESSING THE DATA

- Needs SQL code (easily generated with AI chatbots)
- Here, data provided as a spreadsheet
- For other interesting activities, check https://voyages.sdss.org/



GALAXY MORPHOLOGY – THE SHAPES OF THE UNIVERSE

Galaxy morphology = the shape and structure of a galaxy

- **Spiral galaxies**: flat disks with spiral arms
- **Elliptical galaxies**: round, smooth systems
- **Irregular galaxies**: chaotic, no clear shape

These shapes are not just appearances — they tell us about how stars formed, how galaxies evolved, and whether they have merged or interacted.

Shown here: Hubble's "Tuning Fork" diagram — a classic way of organizing galaxy shapes.

ACTIVITY 1: GALAXY GALLERY

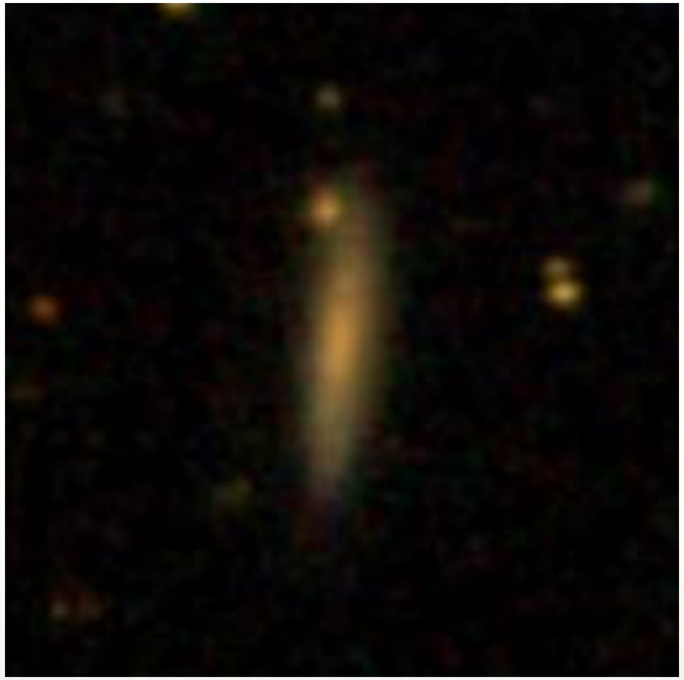
Goal: Decide if each galaxy is a **Spiral (S)** or **Elliptical (E)**, then compare with the computer's simple rule.

Instructions

1. **Look carefully** at the 12 galaxies (3 per slide).
2. **Individually** decide for each: Spiral (S) or Elliptical (E).
3. **Discussion in pairs**: Compare your answers when all 12 are shown together.
4. **Share and reveal**: The computer's "auto" label will be revealed as each group shares their best estimate.

#1

objid: 1237648675068445173

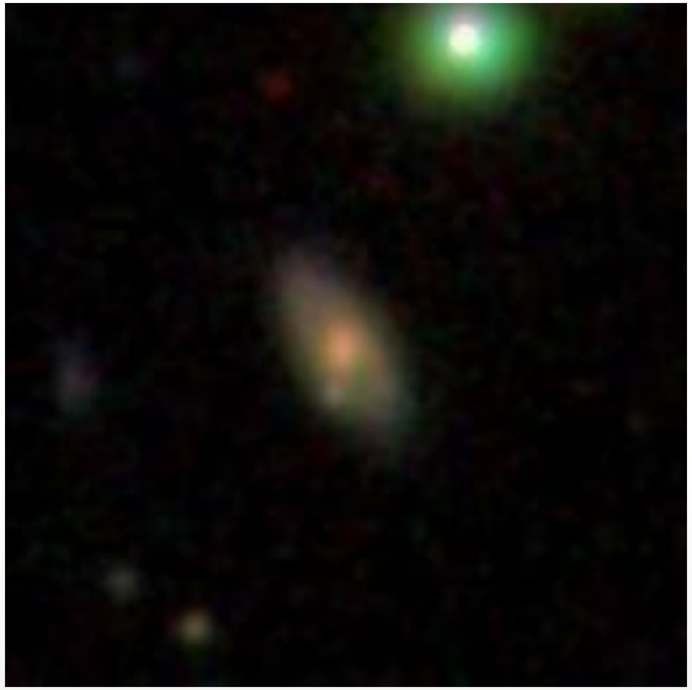


#2

objid: 1237648704602243453

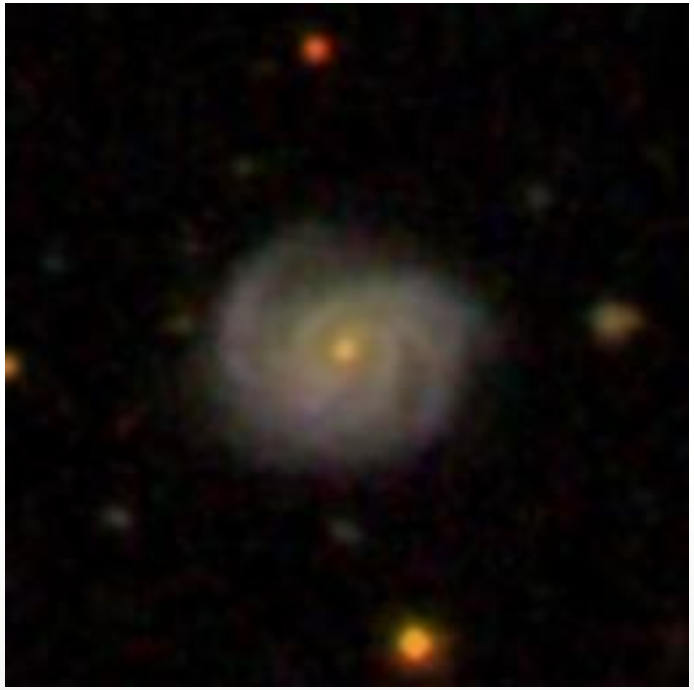
#3

objid: 1237648704044007728



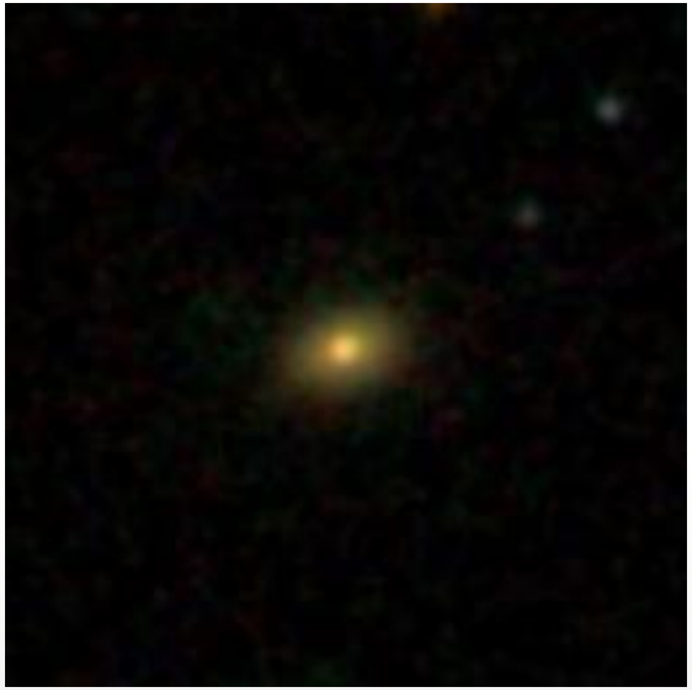
#4

objid: 1237648704054493316



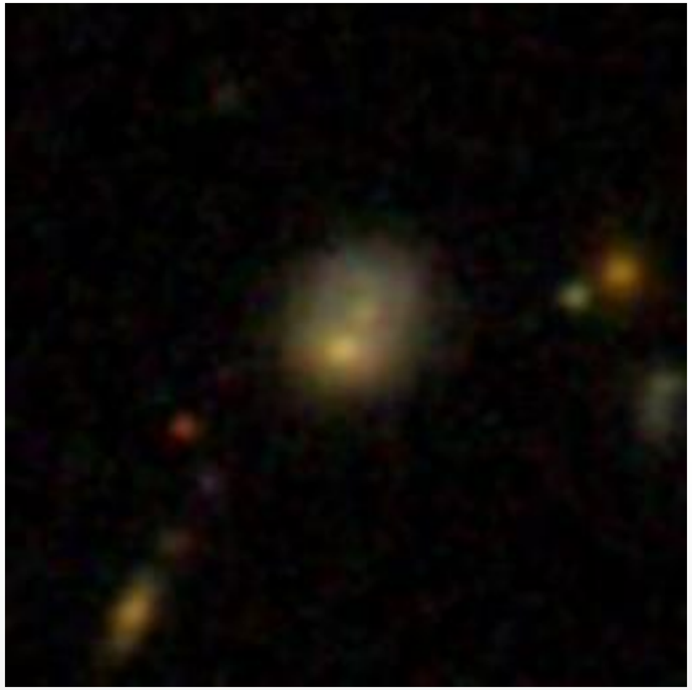
#5

objid: 1237648705123844289



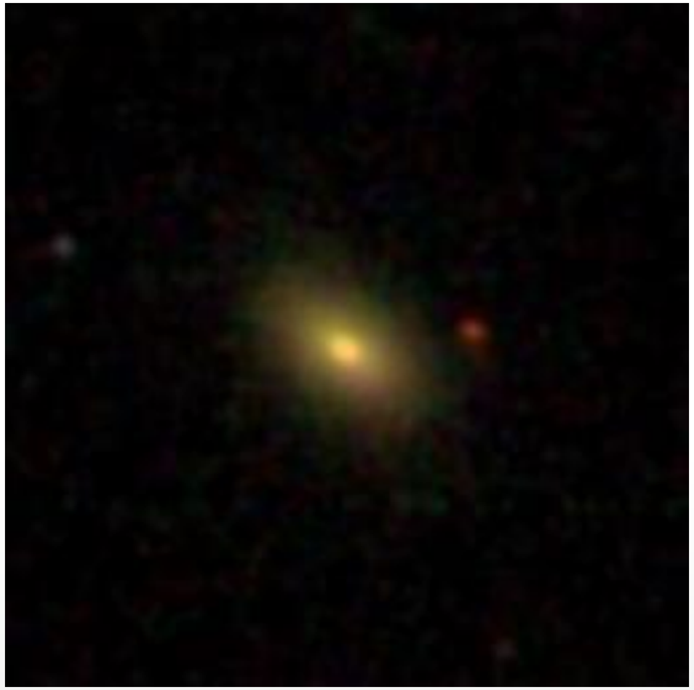
#6

objid: 1237648703508840617



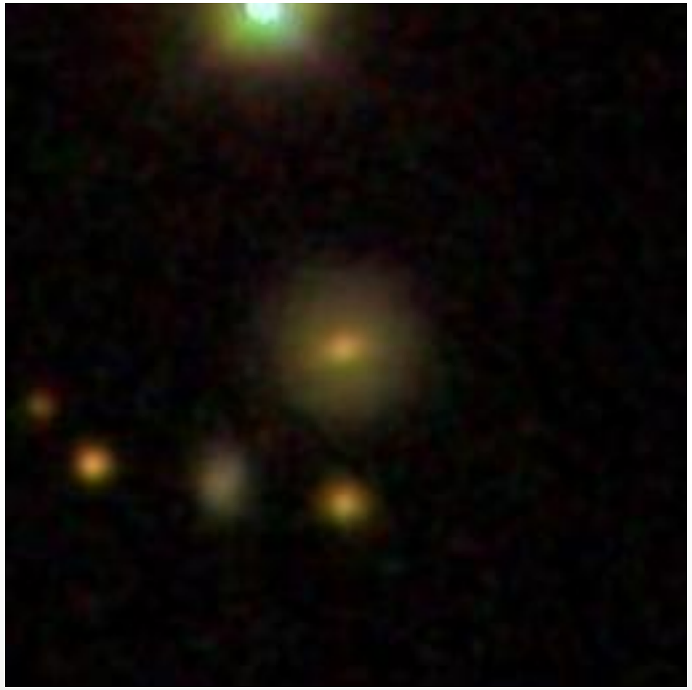
#7

objid: 1237648702973870224



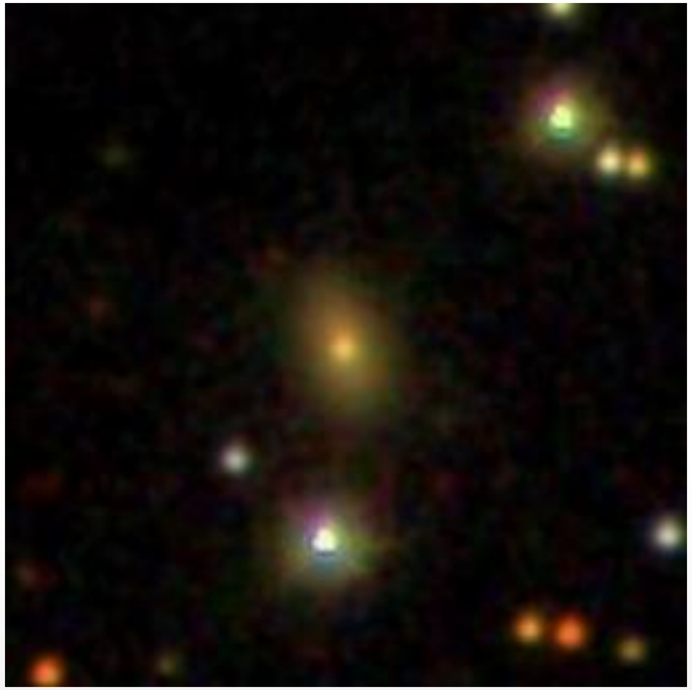
#8

objid: 1237648704585007316



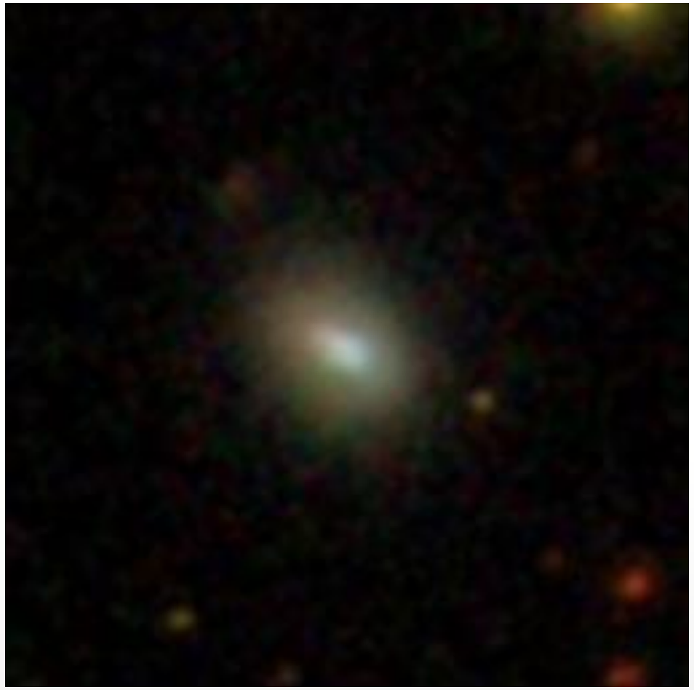
#9

objid: 1237648704054689846



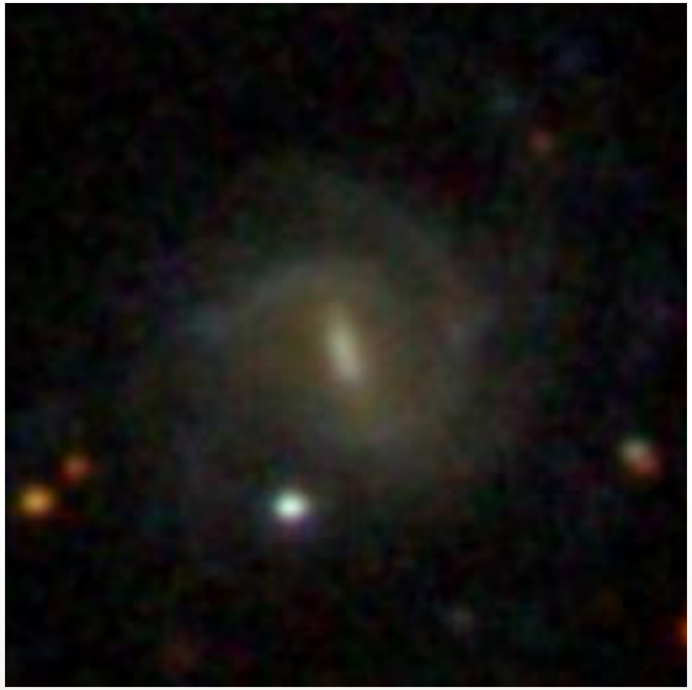
#10

objid: 1237648673460912298



#11

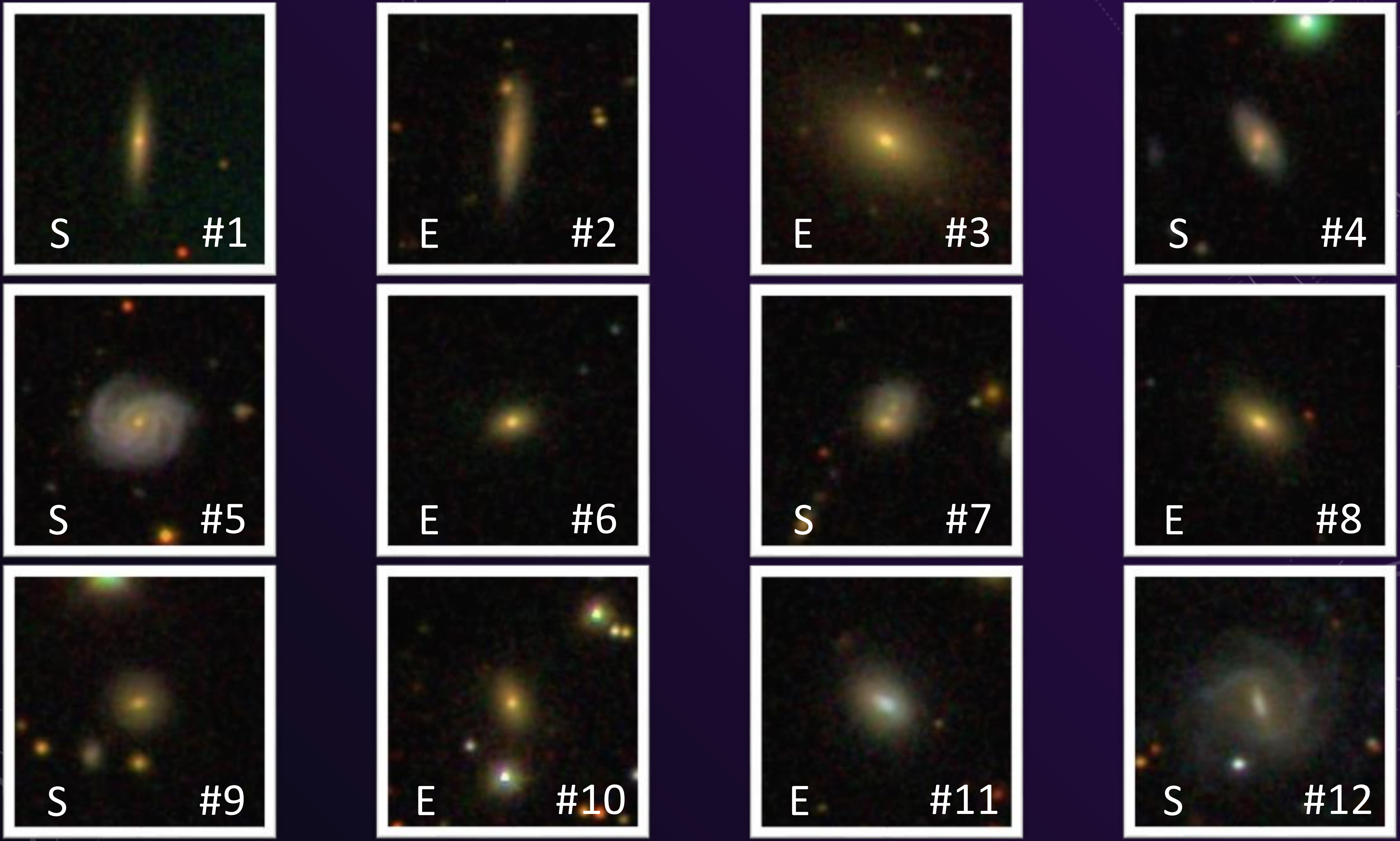
objid: 1237648703516311675



#12

objid: 1237648705137541215


#1
#2
#3
#4
#5
#6
#7
#8
#9
#10
#11
#12


S #1
E #2
E #3
S #4
S #5
E #6
S #7
E #8
S #9
E #10
E #11
S #12

DISCUSSION AND CONCUSIONS

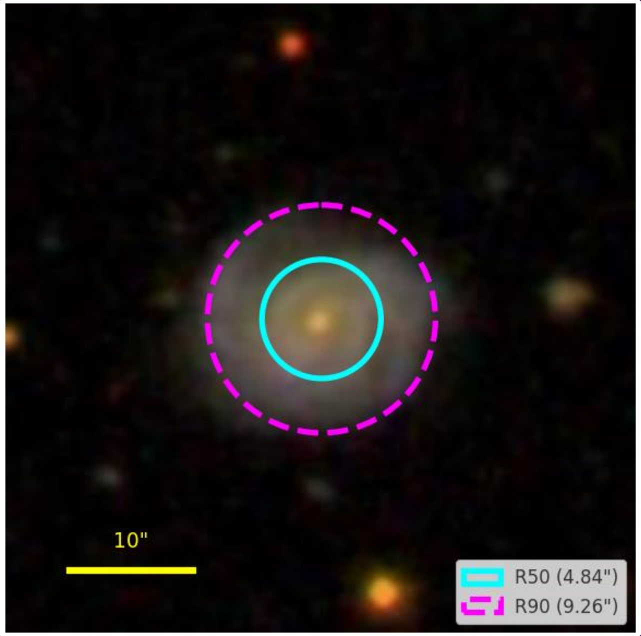
1. What patterns distinguish the galaxies you found easy to label from those you found difficult? Be specific about what you saw.

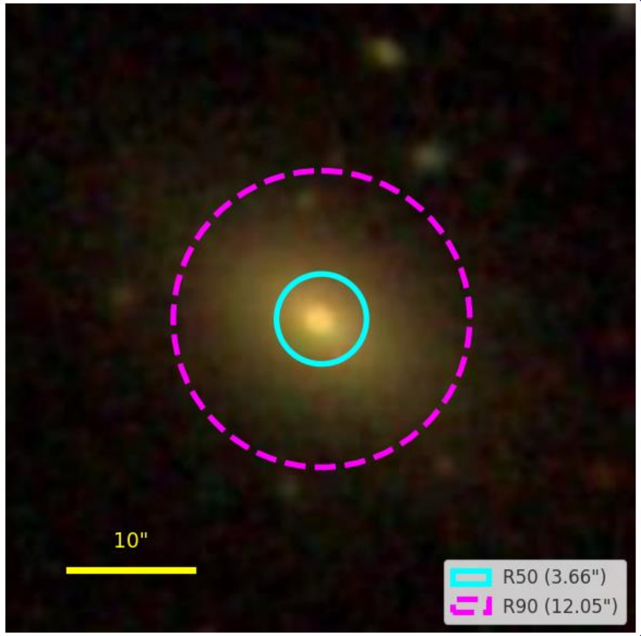
2. Which visual cues did you rely on most? Rank your top three and explain why.

3. After reviewing the mismatches, how much do you trust the automatic label (1–5)? What evidence would raise or lower that score?

4. Did any images resist a simple spiral/elliptical label? If so, propose a third label or an “unsure” policy and when to use it.

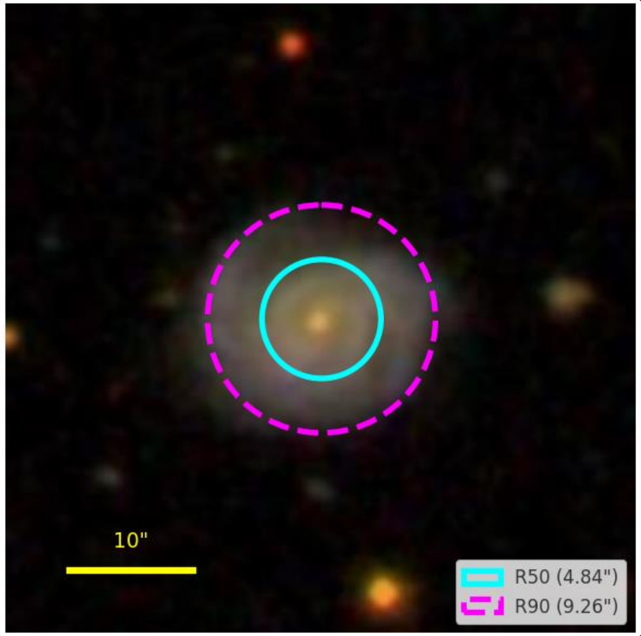
5. List two things that could make the features hard to see. How could you tell which one is at play using only the image?

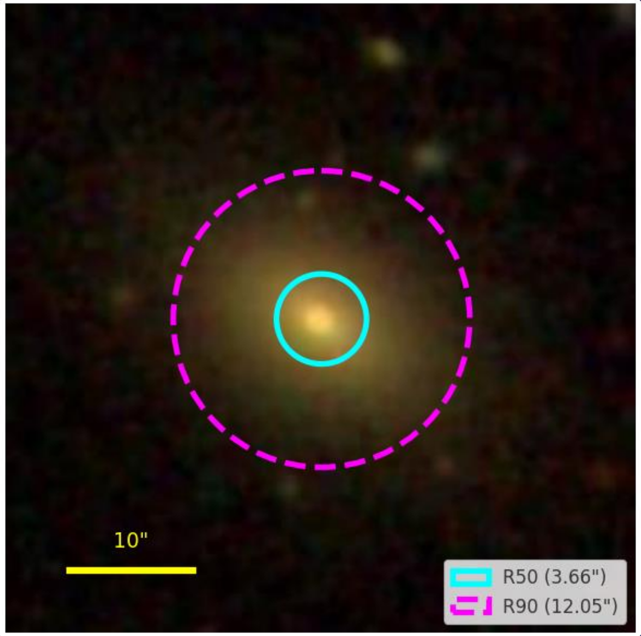




HOW THE "AUTO" CLASSIFICATION IS DONE: CONCENTRATION INDEX

- Computers don't look for spiral arms or smoothness
- They measure how light is distributed in the galaxy by finding two radii:
  - R50 → radius enclosing 50% of the light
  - R90 → radius enclosing 90% of the light
- Then the Concentration Index is calculated:
  - $C_r = R90/R50$
- Rule of thumb
  - $C_r \geq 2.6$ → elliptical-like (centrally packed)
  - $C_r < 2.6$ → spiral-like (spread out)
- Caution: Not perfect → small, faint, or edge-on galaxies can fool it
- But it is simple, consistent, and reproducible





HOW THE "AUTO" CLASSIFICATION IS DONE: CONCENTRATION INDEX

Why do spiral-like galaxies look more “spread out”?

- Spirals:
  - Flat disks with arms stretching far from the center
  - Light spread across the disk and arms
  - R50 is relatively large, and R90 isn’t much bigger → Lower concentration (R90/R50 smaller)
- Ellipticals:
  - Round, centrally packed
  - Light rises steeply in the core, fades outward
  - R50 is small, R90 much larger → Higher concentration (R90/R50 bigger)

ACTIVITY 2: CONCENTRATION LADDER — ORDER FROM MOST TO LEAST

Goal: Look at 6 real SDSS galaxy images. Rank them from most centrally packed to most spread out (by eye).

Instructions

1. **Look carefully** at the 12 galaxies
2. **Individually** decide which has the highest concentration index, and which lowest. Rank them.
3. **Discussion in pairs**: Compare your answers in pairs.
4. **Share and reveal**: The ranking based on the measured index will be revealed as each group shares their best estimate.

#1
#2
#3
#4
#5
#6

HIGHEST CONCENTRATION INDEX

LOWEST CONCENTRATION INDEX

HIGHEST CONCENTRATION INDEX

LOWEST CONCENTRATION INDEX

FROM SHAPES TO COLORS

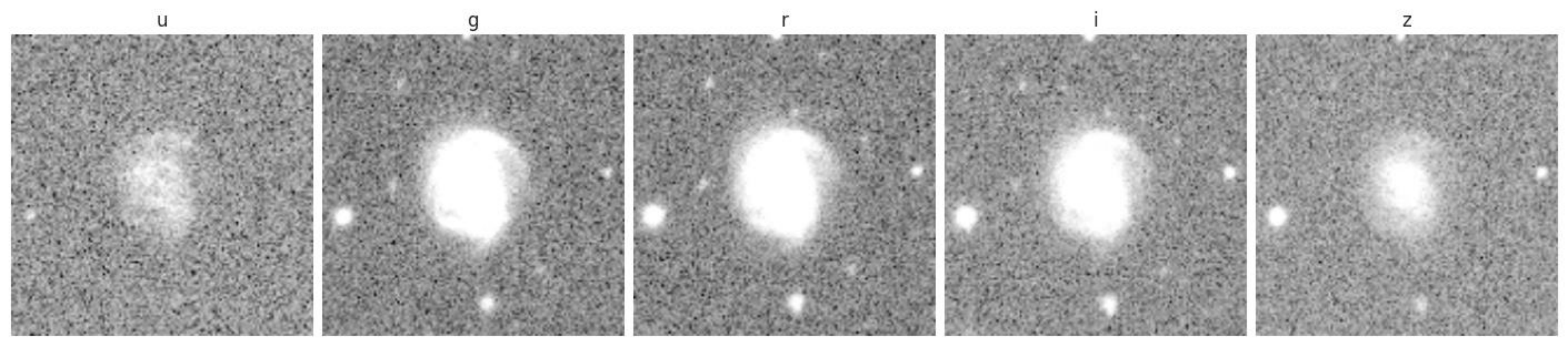
- We’ve sorted galaxies by their outer shapes using concentration
- But shapes don’t tell the whole story
- Next we “measure” the colors of galaxies

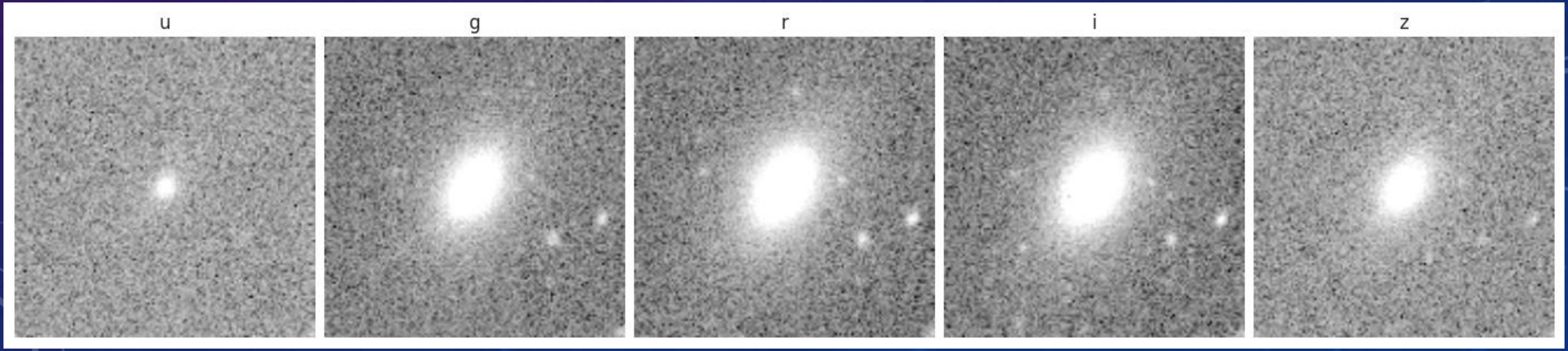
It's no accident that SDSS didn’t take just one picture of each galaxy, but five: through the u, g, r, i, z filters.

By combining and comparing the light from these filters, we can learn a lot about each galaxy.

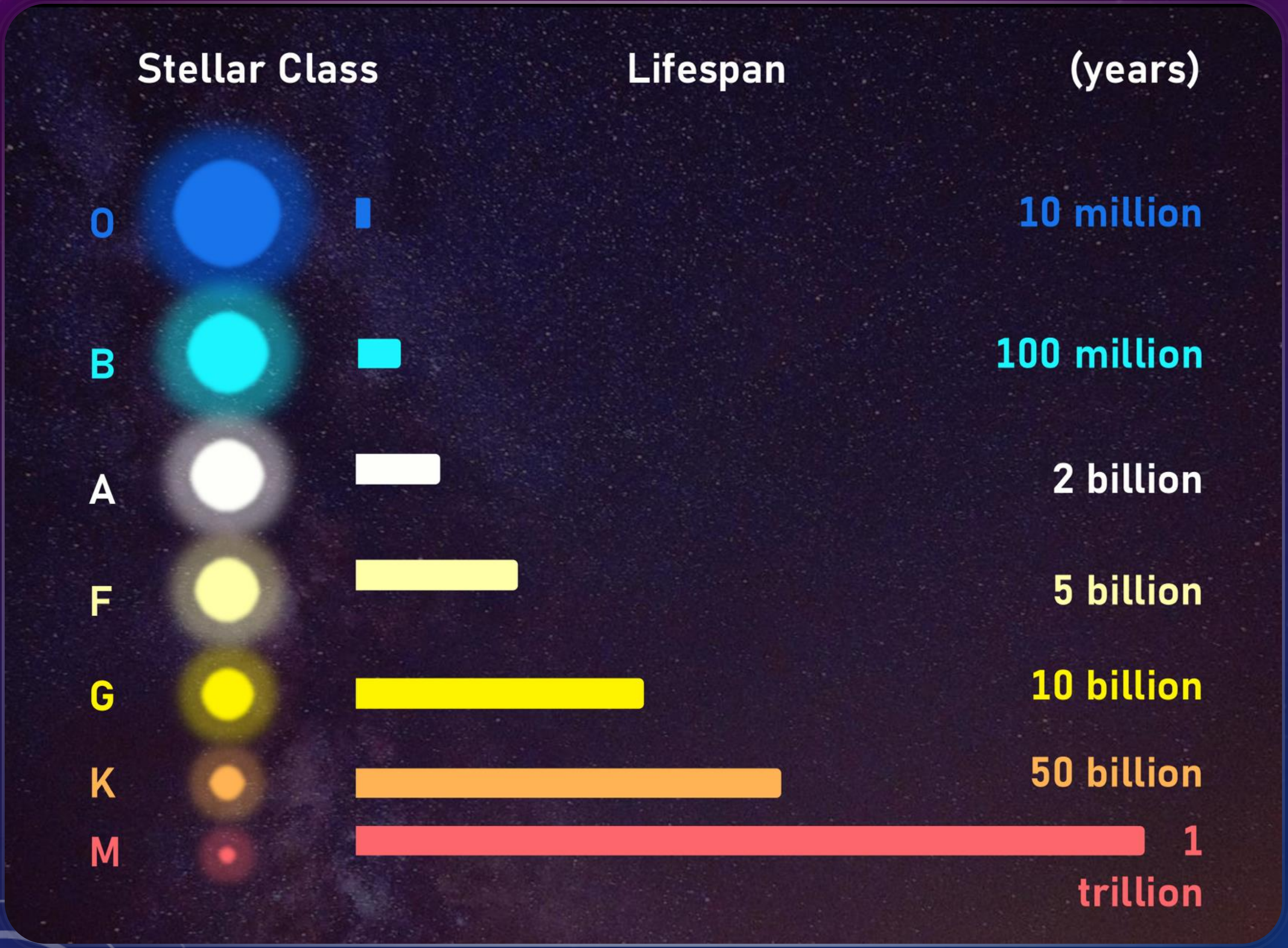
- A galaxy’s color comes mainly from its stars

WHAT GALAXY IMAGES ACTUALLY LOOK LIKE





STARS AND GALAXY COLORS



Stars are classified by their colors:

- O/B stars → hot, blue, short-lived
- K/M stars → cool, red, long-lived

Galaxy color = snapshot of its stars:

- Blue galaxy → has young, massive stars (must be forming stars recently)
- Red galaxy → dominated by older, cooler stars

Important:
Dust can absorb/scatter blue light → galaxies (especially edge-on spirals) may look redder and dimmer than they really are.

COLOR INDEX (u–r)

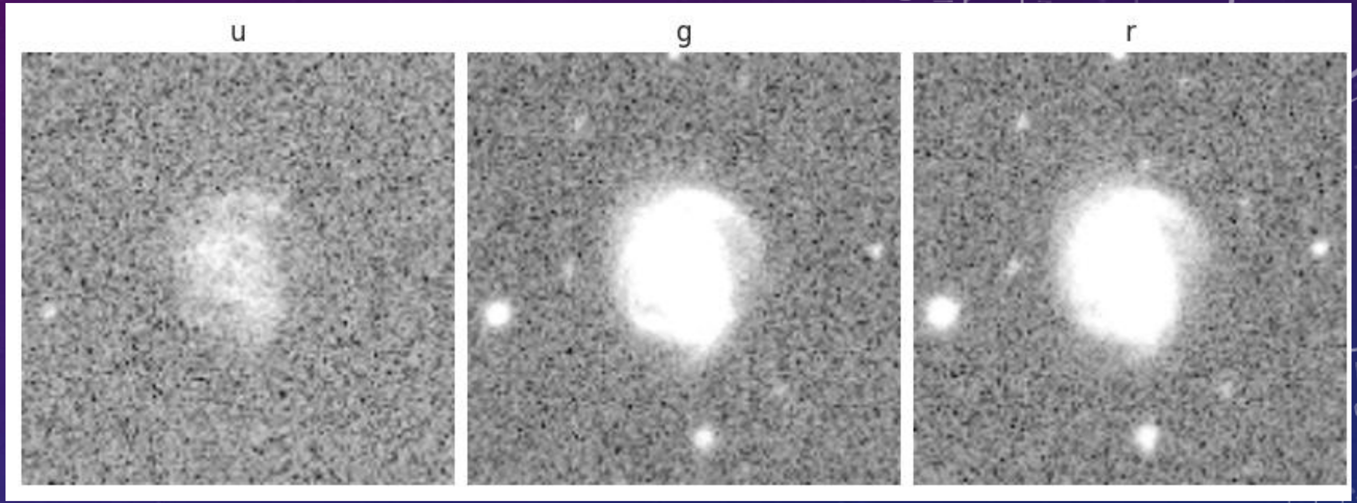
Compare brightness in **u (ultraviolet)** and **r (red)**

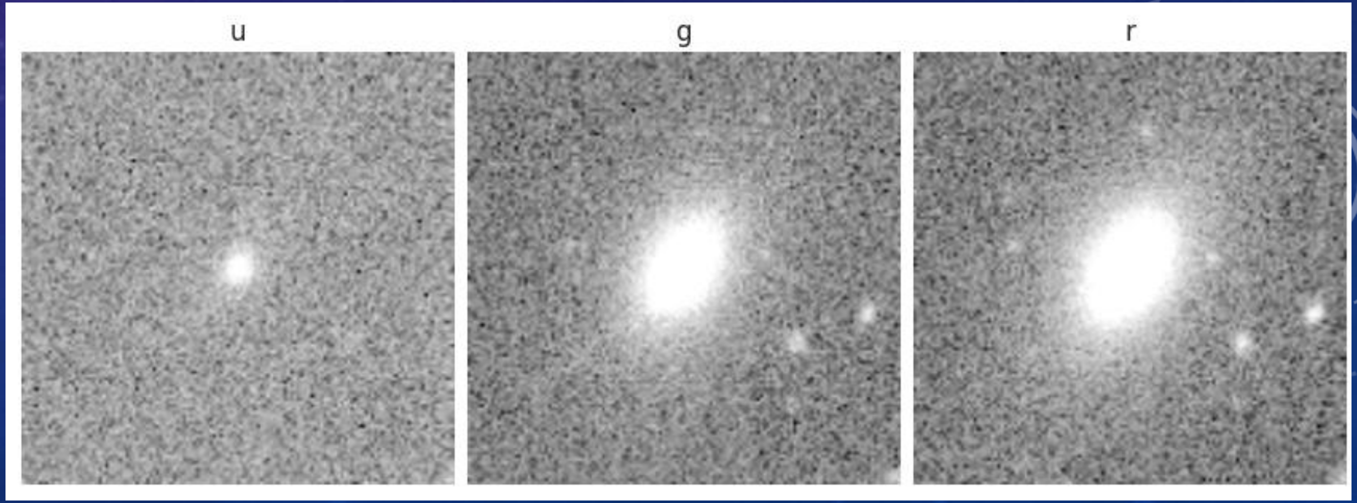
$$u - r$$
$$= brightness\ in\ u - brightness\ in\ r$$

- **Smaller u–r** → brighter in u → **bluer** galaxy
- **Larger u–r** → brighter in r → **redder** galaxy

Rule of thumb

- $u–r < 2.2$→ spiral-like (blue)
- $u–r \geq 2.2$→ elliptical-like (red)





DOES COLOR MATCH SHAPE?

- Galaxy type so far: from **concentration index**
- Now: check if **color (u–r)** tells the same story

For the next activities, we will switch onto a spreadsheet, to work with a prepared data table.

A CSV file will be provided with:

- Galaxy label from concentration index (C_r)
- Brightness values in all 5 SDSS filters (u, g, r, i, z)
- Redshift and other useful data

In your favorite spreadsheet software, you'll explore how color and shape connect by making your own plots and calculations.

Several activities will guide you step by step — no "ready answers" on the slides, you'll find them yourself!

ACTIVITY 3: WORKING WITH DATA

Task 3.1: Creating a u-r color index histogram

- A histogram is a plot that shows how many galaxies fall in each u-r value range.

Steps:

1. Open the provided CSV file.
2. Make two new tables (can be on the same sheet):
   - Table 1: copy u_r values for spiral-like galaxies
   - Table 2: copy u_r values for elliptical-like galaxies
3. Insert a chart:
   - Select both tables.
   - Choose Insert → Chart → Histogram (Google Sheets) or Insert → Histogram/Column (Excel).
4. Adjust the chart:
   - Both groups should appear in different colors automatically.
   - Add axis labels:
     - X-axis = “u–r (color index)”
     - Y-axis = “Number of galaxies”

ACTIVITY 3: WORKING WITH DATA

Link: https://github.com/JanaBog/Galaxies_workshop

To work with spreadsheets, open the data file: Galaxy Data -> galaxy_data.csv

For the python notebook, open Analyzing_Galaxies_Star_Formation_and_Morphology.ipynb, then open in Colab to be able to run it

DISCUSSION: WHAT CAN WE CONCLUDE?

- Does color mostly agree with our shape label?
- Do the two curves peak at different u-r values, or mostly sit on top of each other?
- If you had to pick one cutoff, where would you place it—and why? Is the rule of thumb at u-r=2.2 the best choice for this sample?
- In which u-r range do the two groups mix the most? What does that say about confidence in labels there?
- Can you spot bins where spiral-like galaxies appear at very red u-r or elliptical-like appear at very blue u-r? What might explain those cases?
- For which u-r ranges would you feel most confident in the label, and least confident? What visual evidence from the plot supports that?
- What extra clue could a second color (e.g., u-g or g-r) add to separate the groups?

ACTIVITY 3: WORKING WITH DATA

Task 3.2: Using Two Colors on a Color–Color Diagram, g-r vs u-g

Steps:

1. Open the CSV file and make two new tables:
   - Table 1 (spirals): copy u_g (u–g) and g_r (g–r) values for spiral-like galaxies
   - Table 2 (ellipticals): copy the same columns for elliptical-like galaxies

2. Insert a chart:
   - Select both tables
   - Choose Insert → Chart → Scatter plot

3. Adjust the chart:
   - Put g–r on the X-axis and u–g on the Y-axis
   - Each galaxy = one dot
   - Spirals and ellipticals should show up in different colors automatically
   - Label axes clearly:
     - x-axis = “g–r (color)”
     - y-axis = “u–g (color)”

ACTIVITY 3: WORKING WITH DATA

Task 3.2.1: Add the Mystery Dots

Steps:

1. Open the Mystery_Dots CSV file and add these data to your color-color diagram
2. Determine the type of galaxy the Mysery Dots represent, based on the position
3. Check your answers
4. Which did you get wrong? Why do you think that is?

The answer is:
SSSEEESE

DISCUSSION: WHAT CAN WE CONCLUDE?

- Better separation: Blue, star-forming systems tend to land in a different area than red, older systems — clearer than with one color alone.

- Dust vs. stars (which way points move): Dust dims blue/UV more than red, so both u−g and g−r get larger—points slide up-right along a “reddening” arrow. Changing the stellar mix mostly changes the UV relative to blue, so points move mainly up/down (u−g changes more than g−r). Using two colors lets you tell these different directions apart.

- Find outliers: Very unusual objects (e.g., super-blue starbursts, active nuclei) stick out from the main clumps.

- Cross-check labels: If our shape label (from concentration) and color disagree, the diagram shows where and how they disagree.

- Quality sanity check: If your sample is reasonable, you’ll see recognizable clusters/“sequences.” Bad data points often sit in odd places.

ACTIVITY 3: WORKING WITH DATA

Task 3.3: Color vs. Redshift
See how galaxy color (u–r) changes with distance (redshift).

Steps:
1. From the CSV, make two tables:
- Table 1: redshift and color (u_r) for spiral-like galaxies
- Table 2: redshift and color (u_r) for elliptical-like galaxies

2. Insert a chart:
- Select both tables
- Choose Insert → Chart → Scatter plot

3. Adjust the chart:
- x-axis = redshift
- y-axis = u–r (color index)
- Two groups should appear in different colors automatically
- Add axis titles:
    - x = "Redshift "
    - y = "u–r (color index)"

DISCUSSION: WHAT CAN WE CONCLUDE?

- Do spiral-like and elliptical-like galaxies keep their relative ordering in u–r across z?
- Is there any overall drift in u–r with redshift for either group, or is it roughly flat here?
- Where are the bands widest (most spread in colors)? What could cause that spread?
- If we extended to higher z, what new challenges might appear (fainter galaxies, stronger color shifts, more dust effects)?

WORKSHOP RECAP: OUR JOURNEY THROUGH GALAXIES

- Learned how astronomers collect data with SDSS
- Explored colors & magnitudes as clues to stellar populations
- Practiced classifying galaxies: spirals vs ellipticals using concentration & color
- Investigated color vs. redshift → how galaxies change over cosmic time

Takeaway:

Astronomy connects images, measurements, and theory to reveal the life stories of galaxies and their place in the universe.

THANK YOU!

QUESTIONS OR FEEDBACK?