

ASTR4004/ASTR8004

Astronomical Computing

Assignment 3

Sven Buder

due Tuesday, October 7, 2025

1 git in Practice

1.1 Task

Whether you are working in astronomical research or the data science industry, being able to share code and collaborate on projects is essential. This part of the assignment aims to help you build your *muscle memory* for using `git`, a common version control tool.

For this task, you will create a `git` repository to track and update your progress on the assignment, including your solutions to other tasks. Treat this repository as if you are working collaboratively with others. This typically involves the following workflow: (1) Start by creating your own branch off of the main branch. (2) Do your work (coding, updating files) in this new branch. (3) Regularly upload (commit and push) your changes to your branch, including meaningful commit messages. At least once, you should:

- Initialise and push some content (e.g., code and a README file) to the *main* branch of a new public repository.
- Create a branch with a descriptive name and upload new changes to this branch.
- Merge this branch back into the *main* branch.

The choice of whether you work mainly in the *main* branch or an additional branch is up to you, but you must demonstrate the ability to work within both.

Ensure you maintain a clear directory structure, with your data in `data/` and figures saved in `figures/`.

1.2 Marking (up to 3 points)

Your progress will be assessed based on your use of `git`, as described below. Ensure that you provide a link to your GitHub repository and specify the name of your main working branch if it differs from the default *main* branch.

Marks will be awarded based on successful completion of these milestones (3 in total), the quality of your commit messages, and your demonstrated understanding of basic `git` practices.

2 Using ADQL to Search for Bright Stars Around the Open Cluster Ruprecht 147

2.1 Task

A colleague is interested in the open cluster Ruprecht 147 (RA = 289.074 deg, Dec = -16.323 deg) and is considering an observation proposal using the 2dF fibre positioner and HERMES spectrograph (effective for Gaia G band magnitudes < 14). They need to know if there are enough bright stars in this region for observation, and how bright these stars would be in the wavelength region of the APOGEE spectrograph (which operates in the H-band). Your task is to assist by querying data from Gaia DR3 and performing some basic analysis.

- Download all stars within 1 degree of the center of Messier 67 that are brighter than $G = 14$ in Gaia DR3 (the table is called `gaiadr.gaia_source`) and include a crossmatch these stars with the 2MASS catalog and report your ADQL query text (use the `"""query"""` notation)
- Determine how many stars are returned from the initial query.
- Identify any stars with bad 2MASS photometry, where `ph_qual` is not 'AAA'.
- Identify any stars with negative (or non-positive) parallaxes in the Gaia data.
- Apply these two quality cuts (removing stars with bad 2MASS photometry and non-positive parallaxes). After applying the cuts, determine how many stars remain.
- Using the remaining stars, generate a figure with two panels (1 point per panel):
 - (a) A color-magnitude diagram (CMD) of Gaia BP-RP vs. absolute G magnitude.
 - (b) A 2MASS J-Ks vs. apparent H magnitude diagram.
- Save the figure as `figures/cmds_R147.png` with a resolution of 200 dots per inch.
- Give your colleague a recommendation for the potential proposal when only judging the fibre usage, that is, how many fibres of 2dF are available and would be used.

2.2 Marking (up to 8 points)

Marks will be awarded based on the tasks above. Marks will be deducted for improperly labeled/constructed plots.

3 The radial metallicity relation in simulated data

3.1 Background

The radial metallicity relation is a function that describes the change of metallicity - here the gas phase metallicity $A(O) = \log_{10}(N_O/N_H) + 12$ - along the galactocentric radius $R_{\text{Gal.}}$. Understanding the radial metallicity gradient in galaxies provides critical insights into their formation and evolutionary processes, such as inside-out formation, gas accretion, outflows, and radial migration. A lot of work has been done through observational studies (e.g. Ho et al., 2017, ApJ, 846, 39) and a few simulations (e.g. Grand et al., 2016, MNRAS, 460, 94), but more work needs to be done to understand the radial metallicity gradient!

Your colleague has just finished an exciting cosmological simulation that traces the gas phase metallicity for a Milky Way analogue, that is, a spiral galaxy. They have limited the simulation data to the positions (x , y , z) of the innermost gas particles ($R_{\text{Gal.}} < 25 \text{ kpc}$) and their gas phase metallicity $A(O)$ and uploaded them as a FITS file for you here: https://github.com/svenbuder/astr4004_2025_week8/blob/main/data/nihao_uhd_simulation_g8.26e11_xyz_positions_and_oxygen_ao.fits

3.2 Task

Download the file from the link above into `data/`. Load the file with `python` and then perform the following tasks to create figures that are saved in `figures/`:

- Plot a 2-panel figure:
 - (a) Logarithmic density plot of $R_{\text{Gal.}}$ vs. $A(O)$, with a linear fit and legend.
 - (b) Residuals of the fit, $R_{\text{Gal.}}$ vs. $\Delta A(O)$.
- Use a `python` fitting tool to fit a linear function to the data, reporting the intercept and slope with uncertainties. Include any hyperparameters used.
- Discuss where the linear model fits well and where it does not. Use statistical metrics, such as the root mean squares or other goodness-of-fit indicators, to quantify the performance of your linear fit in general and regions with larger residuals.
- Plot a 3-panel figure for the x vs. y plane using the same bins and sensible colormaps:
 - (a) 2D-histogram of the median simulated $A(O)$.
 - (b) 2D-histogram of the median fitted $A(O)$
 - (c) 2D-histogram of the median residuals $\Delta A(O)$.
- Describe your choice of 2D bins. Discuss what details would be missed with fewer bins or problems encountered with more bins.
- Analyze the residuals in more detail and propose an explanation for any patterns you observe.

3.3 Marking Criteria (up to 9 points)

Each sub-task is worth a point. Points will be deducted for poorly labeled or constructed plots.