Astro 17: Extragalactic Astronomy and Cosmology

Instructor's Guide to Lab Notebooks

Christopher Stubbs with his friend Claude

Contents

1 Introduction

This guide accompanies the Google Colab notebooks for Astro 17, Harvard's introductory course in extragalactic astronomy and cosmology. The notebooks provide hands-on computational exercises covering topics from stellar populations to cosmological distance measurements.

All notebooks are designed to run on Google Colab, a cloud-based Python notebook environment that requires no local software installation.

2 Google Colab: Benefits and Considerations

2.1 Benefits of Using Colab

- **No local setup required** – Students can begin working immediately without installing Python, Jupyter, or any packages on their personal computers.
- **Platform independence** – Works on any computer with a web browser (Windows, Mac, Linux, Chromebook).
- **Pre-installed scientific libraries** – NumPy, Matplotlib, SciPy, Astropy, and Pandas are available by default.
- **GPU/TPU access** – For advanced exercises, Colab provides free access to accelerated computing resources.
- **Easy sharing** – Notebooks can be shared via Google Drive links for collaboration or grading.
- **Built-in AI assistant** – The Gemini AI integration helps students understand code and debug issues.

2.2 Important Considerations

- **Google Drive mounting required** – To access data files, students must mount their Google Drive and navigate to the folder containing the data files. This requires a Google account.
- **Data files must be uploaded** – Students need to copy the data files from the course repository to their own Google Drive's `Colab Notebooks` folder (or another folder of their choice, adjusting the path in the notebook accordingly).
- **pip installations do not persist** – Any packages installed with `pip install` during a session are lost when the runtime disconnects. Students must re-run pip install cells each time they open a notebook. For example, the `LineProfileGalaxies` notebook requires `pip install ipympl` at the start of each session.
- **Runtime timeouts** – Colab disconnects after periods of inactivity (typically 90 minutes). Students should save work frequently.
- **Internet required** – Colab is entirely cloud-based and requires an internet connection.

2.3 Student Setup Instructions

Students should follow these steps before their first lab session:

1. Create a Google account if they don't have one.
2. In their individual Google Drive, if it does not exist already each student should create a folder called `Colab Notebooks`.
3. Download all data files from the course repository's `data/` folder.
4. Upload the data files to their `Colab Notebooks` folder on Google Drive.
5. Open each notebook from the course repository in Colab.
6. When prompted, authorize Colab to access Google Drive.

3 Individual Notebook Descriptions

3.1 Gaussians

Topic: Statistical fitting and Gaussian distributions

Learning Objectives: Understanding probability distributions, curve fitting, chi-squared analysis, parameter estimation with uncertainties

Required Data: None (synthetic data generated in notebook)

Prerequisites: Basic Python, NumPy arrays

Common Issues: None – self-contained notebook

3.2 Linear_Regression_Correlation

Topic: Linear regression and parameter correlation

Learning Objectives: Understanding how slope and intercept become correlated when fitting data far from origin; mean subtraction to improve numerical stability

Required Data: None (synthetic data generated in notebook)

Prerequisites: Basic statistics, curve fitting concepts

Common Issues: None – self-contained notebook

3.3 FIRASfit

Topic: CMB spectral fit to Planck function

Learning Objectives: Fitting a model to data, consistency with Big Bang

Required Data: None (included in notebook)

Prerequisites: Basic cosmology concepts, basic statistics and fitting

Common Issues: None – self-contained notebook

3.4 Distances

Topic: Cosmological distance measures

Learning Objectives: Understanding luminosity distance, angular diameter distance, comoving distance; cosmological redshift; distance modulus

Required Data: None (calculations only)

Prerequisites: Basic cosmology concepts

Common Issues: None – self-contained notebook

3.5 gaia

Topic: Hertzsprung-Russell diagram using Gaia data

Learning Objectives: Stellar parallax and distance; absolute vs. apparent magnitude; color-magnitude diagrams; main sequence, giant branch, white dwarfs

Required Data: `gaia_200pc_sample.fits`

Prerequisites: Magnitude system, stellar classification basics

Common Issues: Students must correctly calculate absolute magnitude from parallax

3.6 Galaxy_Rotation_Curve_Fitting

Topic: Galaxy rotation curves and dark matter

Learning Objectives: Rotation curve analysis; evidence for dark matter; mass modeling (disk, bulge, halo components)

Required Data: None (NGC 7331 data embedded in notebook)

Prerequisites: Keplerian dynamics, mass estimation

Common Issues: Understanding the disk/halo decomposition

3.7 LineProfileGalaxies

Topic: Galaxy morphology through brightness profiles

Learning Objectives: FITS image analysis; surface brightness profiles; exponential vs. de Vaucouleurs profiles; morphological classification; data visualization

Required Data: Multiple FITS images: `M74DSS2red.fits`, `M87SDSSr.fits`, `M101DSS2red.fits`, `M84SDSSr.fits`, `M89SDSSr.fits`, `M105SDSSr.fits`, `NGC4874DSS2red.fits`, and others

Prerequisites: Galaxy classification, magnitude systems

Special Setup: Requires `pip install ipympl` at start of each session, followed by kernel restart

Common Issues: Interactive widget may not work if ipympl not installed; students must restart runtime after pip install

3.8 NumberCounts

Topic: Galaxy number counts and star-galaxy separation, inconsistent with static Euclidean cosmology.

Learning Objectives: Photometric catalogs; PSF vs. Petrosian magnitudes; star-galaxy separation; number counts as cosmological probe

Required Data: `A1942nohead.tsv`

Prerequisites: Magnitude systems, basic statistics

Common Issues: Understanding the magnitude uncertainty as inverse SNR

3.9 ClusterLRG_Rubin

Topic: Galaxy cluster analysis with Rubin Observatory data

Learning Objectives: cModel photometry; color-magnitude diagrams; red sequence selection; photometric redshifts

Required Data: `abell360_rubin_dp1_303_1_galaxy_photometry_v6.fits`

Prerequisites: NumberCounts notebook concepts

Common Issues: Large file size may slow upload

3.10 Zshell

Topic: Galaxy clustering as compared to random Poisson statistics.

Learning Objectives: Redshift shells; K-corrections; luminosity functions; Schechter function fitting

Required Data: `sdss_photometry.csv`

Prerequisites: Redshift, magnitude systems

Common Issues: Understanding the narrow redshift slice selection

3.11 SN1a_Hubble_Fit

Topic: Type Ia supernovae and the Hubble diagram, evidence for dark energy

Learning Objectives: Standard candles; distance modulus; Hubble diagram; cosmological parameter estimation

Required Data: None (supernova data embedded in notebook)

Prerequisites: Distances notebook concepts

Common Issues: Understanding distance modulus vs. redshift relationship

3.12 Diffraction_Grating_Lab

Topic: Spectroscopy with diffraction gratings

Learning Objectives: Diffraction grating equation; wavelength calibration; spectral line identification

Required Data: Student-provided spectrum images (taken in lab)

Prerequisites: Basic optics, wavelength concepts

Common Issues: Image quality affects measurements; students upload their own images via `files.upload()`

4 Required Python Packages

All notebooks use packages that are pre-installed on Google Colab:

- `numpy` – Numerical computing
- `matplotlib` – Plotting
- `scipy` – Scientific computing (curve fitting, optimization)
- `astropy` – Astronomical utilities (FITS, units, coordinates)
- `pandas` – Data manipulation

One notebook requires an additional package:

- `ipympl` – Required by `LineProfileGalaxies.ipynb` for interactive plotting. Must be installed with `pip install ipympl` at the start of each session.

5 Data Files Summary

File	Size	Used By
`A1942nohead.tsv`	2.9 MB	NumberCounts
`abell360_rubin_dp1_...fits`	1.7 MB	ClusterLRG_Rubin
`gaia_200pc_sample.fits`	360 KB	gaia
`sdss_photometry.csv`	220 KB	Zshell
`M74DSS2red.fits`	4.2 MB	LineProfileGalaxies
`M87SDSSr.fits`	4.2 MB	LineProfileGalaxies
`M87SDSSu.fits`	4.2 MB	LineProfileGalaxies
`M101DSS2red.fits`	4.2 MB	LineProfileGalaxies
`M101WISE3.5.fits`	4.2 MB	LineProfileGalaxies
`M84SDSSr.fits`	4.2 MB	LineProfileGalaxies
`M89SDSSr.fits`	4.2 MB	LineProfileGalaxies
`M105SDSSr.fits`	4.2 MB	LineProfileGalaxies
`NGC4874DSS2red.fits`	4.2 MB	LineProfileGalaxies

File	Size	Used By
DECAMr.fits	33.5 MB	LineProfileGalaxies

Total data size: approximately 75 MB.

6 Suggested Notebook Sequence

A recommended order for working through the notebooks:

1. **Gaussians** – Foundational statistics
2. **Linear_Regression_Correlation** – Fitting techniques
3. **FIRASfit.ipynb** – Fitting a Planck function to CMB spectrum, single fit parameter (Temp.)
4. **Diffraction_Grating_Lab** – Spectroscopy (lab-based)
5. **gaia** – Stellar HR diagram
6. **Distances** – Cosmological distances
7. **Galaxy_Rotation_Curve_Fitting** – Dark matter evidence
8. **LineProfileGalaxies** – Galaxy morphology
9. **NumberCounts** – Galaxy number counts compared to static Euclidean cosmology.
10. **ClusterLRG_Rubin** – Galaxy clusters
11. **Zshell** – 2-d Spatial distribution of galaxies in redshift shell
12. **SN1a_Hubble_Fit** – Cosmological measurements

7 Technical Contact

For technical questions about these notebooks, contact:

Christopher Stubbs
Harvard University
[stubbs@g.harvard.edu]

8 Acknowledgements

This material was developed by the Astronomy 17 instructional staff and members of the Bok Center for Teaching and Learning. Contributors include Kaylee de Soto, Dongpeng Huang, Jungyoon Koh, Marlon Kuzmick, Andrés Alejandro Plazas Malagón, Theo O'Neill, and Madeleine Woods. The notebooks and this manual were developed with assistance from AI tools including Google's Gemini, OpenAI's GPT, and Anthropic's Claude Code.