IELE756: Tools Workshop - Google Colab and Data Science Workflow

2-hour session structure

Leo Ferres, PhD

September 5, 2025

Contents

1 Part 1: Setup and Google Colab Fundamentals (45 minutes) 2
- 1.1 Opening (10 minutes) . . . 2
- 1.2 Google Colab Deep Dive (35 minutes) . . . 2
  - 1.2.1 Account setup and access (5 min) . . . 2
  - 1.2.2 Interface and navigation (10 min) . . . 2
  - 1.2.3 Collaborative features (10 min) . . . 3
  - 1.2.4 Runtime management (10 min) . . . 3

2 Part 2: Python Environment and Key Libraries (40 minutes) 3
- 2.1 Environment Setup in Colab (15 minutes) . . . 3
- 2.2 Essential Libraries Walkthrough (25 minutes) . . . 3
  - 2.2.1 pandas (8 min) . . . 3
  - 2.2.2 geopandas (7 min) . . . 3
  - 2.2.3 xarray & rioxarray (5 min) . . . 3
  - 2.2.4 rasterio & dask (5 min) . . . 3

3 Break (10 minutes) 4

4 Part 3: Workflow Integration and Project-Specific Applications (20 minutes) 4
- 4.1 Documentation and Version Control (10 minutes) . . . 4
  - 4.1.1 Markdown in Colab . . . 4
  - 4.1.2 GitHub integration . . . 4
  - 4.1.3 Obsidian . . . 4

4.2 Data Pipeline Preview (10 minutes) 4

5 Part 4: Hands-On Team Exercise (25 minutes) 4
5.1 Mini-Project Simulation (20 minutes) 4
5.2 Showcase and Troubleshooting (5 minutes) 5

6 Key Takeaways and Next Steps 5

7 Pre-Class Preparation (*Inverted Learning Approach*) 5

8 Post-Class Assignment 5

1 Part 1: Setup and Google Colab Fundamentals (45 minutes)

1.1 Opening (10 minutes)

- Quick overview of the project-driven approach
- Why these tools matter for your Chile disease modeling project
- Team formation logistics (2-3 people per team)

1.2 Google Colab Deep Dive (35 minutes)

1.2.1 Account setup and access (5 min)

- Google account requirements
- Accessing Colab through different entry points

1.2.2 Interface and navigation (10 min)

- Notebook structure: cells, markdown, code
- Toolbar and menu exploration
- Keyboard shortcuts essential for efficiency

1.2.3 Collaborative features (10 min)

- Real-time collaboration (crucial for team assignments)
- Sharing and permissions
- Comments and suggestions

1.2.4 Runtime management (10 min)

- GPU/TPU access and when to use them
- Session limits and best practices
- Connecting to Google Drive

2 Part 2: Python Environment and Key Libraries (40 minutes)

2.1 Environment Setup in Colab (15 minutes)

- Installing packages with `!pip install`
- Managing dependencies for your project
- Persistent storage strategies

2.2 Essential Libraries Walkthrough (25 minutes)

2.2.1 pandas (8 min)

Basic data manipulation for your datasets

2.2.2 geopandas (7 min)

Spatial data handling for Chilean administrative boundaries

2.2.3 xarray & rioxarray (5 min)

Climate and satellite data processing

2.2.4 rasterio & dask (5 min)

Large-scale geospatial data handling

Each library section includes a mini hands-on example with sample data

3 Break (10 minutes)

4 Part 3: Workflow Integration and Project-Specific Applications (20 minutes)

4.1 Documentation and Version Control (10 minutes)

4.1.1 Markdown in Colab

Creating professional reports

4.1.2 GitHub integration

Version control for team projects

4.1.3 Obsidian

Note-taking and knowledge management

4.2 Data Pipeline Preview (10 minutes)

- Connecting to external data sources (MINSAL, INE, ERA5)
- File organization strategies for your multi-dataset project
- Preview of typical data workflow for the course

5 Part 4: Hands-On Team Exercise (25 minutes)

5.1 Mini-Project Simulation (20 minutes)

Teams work together to:

1. Create a shared Colab notebook
2. Load sample Chilean administrative boundary data
3. Perform basic data exploration using the tools introduced
4. Create a simple visualization
5. Document their process using markdown

5.2 Showcase and Troubleshooting (5 minutes)

- Quick team presentations of their mini-projects
- Address common issues encountered
- Best practices recap

6 Key Takeaways and Next Steps

- Assignment 0 expectations and timeline
- Resources for continued learning
- Preview of next week's content
- Q&A session

7 Pre-Class Preparation (*Inverted Learning Approach*)

Students should review before class:

- Google Colab official documentation
- Basic Python refresher (pandas basics)
- Create Google account if needed

8 Post-Class Assignment

Complete **Assignment 0**: Set up all tools and submit a working Colab notebook demonstrating basic functionality.

This structure aligns with your course's collaborative, hands-on approach while ensuring students are prepared for the data-intensive, team-based projects ahead.