Introduction to programming for data science

STAT 201

2025-01-07

Table of contents

Preface

This book serves as the course notes for STAT201 Winter 2025, and it is an evolving resource developed to support the learning objectives of the course. It builds upon the foundational work of the original iteration, authored and maintained by Professor Arvind Krishna. We are deeply grateful for Professor Krishna's contributions, as his work has provided a robust framework and valuable content upon which this version of the book is based.

As the course progresses during this quarter, the notes will be continually updated and refined to reflect the content taught in real time. The modifications aim to enhance the clarity, depth, and relevance of the material to better align with the current teaching objectives and methodologies.

This book is a living document, and we welcome feedback, suggestions, and contributions from students, instructors, and the broader academic community to help improve its quality and utility.

Thank you for being part of this journey, and we hope this resource serves as a helpful guide throughout the course.

Part I

Getting started: Coding environment

1 Setting up your environment with VS Code

```
<IPython.core.display.Image object>
```

1.1 Learning Objectives

By completing this lecture, you will be able to:

- Set up your Python coding environment with VS Code.
- Create and manage Python virtual environments using both `pip` and `conda`.
- Install and verify packages within these environments.
- Export and recreate environments using environment files.
- Use Jupyter Notebook for data science tasks in VS Code.

1.2 Introduction to Visual Studio Code (VS Code)

Visual Studio Code (VS Code) is a free, open-source, and lightweight code editor developed by Microsoft. It's widely used for coding, debugging, and working with **various programming languages** and frameworks. Here's an overview of its key features and functionalities:

1.2.1 Core Features

- **Multi-language Support:** VS Code supports a wide range of programming languages out of the box, including Python, JavaScript, TypeScript, HTML, CSS, and more. Additional language support can be added via extensions.
- **Extensibility:** The editor has a rich ecosystem of extensions available through the Visual Studio Code Marketplace. These extensions add support for additional programming languages, themes, debuggers, and tools like Git integration.