Programming for Data Science

Rafael C. Alvarado

5/8/21

Table of contents

1 Preface

Welcome to Programming for Data Science, a collection of materials designed to support the course DS 5100 in the Data Science curriculum at UVA.

In this course, you will develop skills in Python and R Programming, as well as how to use the command line and GitHub.

The objective of this course is to introduce essential programming concepts, structures, and techniques.

You will gain confidence in not only reading code, but learning what it means to write good quality code.

Additionally, essential and complementary topics are taught, such as testing and debugging, exception handling, and an introduction to visualization.

Part I

Getting Started

During the technical orientation, you set up a GitHub account.

You also spent a little time browsing a sample repository, which you may wish to revisit:

https://github.com/UVADS/orientation-technical

You also should be able check off the following items:

- Understand the difference between Git and GitHub.
- Understand the purpose of Git and Github for data science work.
- Ensure Git is installed on your computer.
- Understand how to find a repository on GitHub.