

ECE 20875: Python for Data Science

Zeke Ulrich

June 10, 2024

Lecture notes for Purdue's ECE 20875.

Contents

<i>Course Introduction</i>	1
<i>Introduction to data science</i>	2

Course Introduction

This course will introduce Python programming to students through data science problems. Students will learn Python concepts as well as introductory data science topics, and will use their knowledge of Python (and prior programming experience) to implement data analyses.

Learning Objectives:

1. Understanding of regular expressions.
2. Use Python to write data analyses.
3. Explain when data analyses are appropriate.
4. Explain the results of data analyses.
5. Incorporate classes in Python.
6. Incorporate associative arrays in programs.

Introduction to data science

Data is defined as quantified information. By analyzing data we can make informed decisions and take actions likely to achieve our goals. Recently, humans have been able to record a lot more data and create machines to use it. Data science collects data from a variety of sources, puts them in a consistent format, makes observations, visualizes trends, interprets, and most importantly predicts. Data science is pervasive and examples can be found in fields spanning healthcare, education, finance, and shopping logistics.

Python is the industry standard for data science and has many useful, expansive open source libraries for data processing like pandas, numpy, and keras.