San José State University
Department of Applied Data Science

DATA 200
Computational Programming for Analytics

Section 21
Spring 2024

Course and Contact Information

Instructor: Ron Mak
Office location: Clark Hall CL 325 (mostly working from home)
Email: ron.mak@sjsu.edu
Website: http://www.cs.sjsu.edu/~mak/
Office hours: TuTh 4:30 – 5:30 PM in person, Clark Hall CL 325
Class days/time: Th: 6:00 – 8:45 PM
Classroom: TBA
Prerequisites: Admission into MS in Data Analytics and instructor consent.

Course Catalog Description

"Computational thinking and problem solving; data structures, algorithms, iterations, functions, and other core concepts; standard libraries; functional programming and object-oriented programming; applications to data processing, data analytics, data visualization, and database manipulation."

Course Format

This class will meet in person in the classroom. Exams will be given in the classroom.

Faculty Web Page and Canvas

Course materials, syllabus, assignments, grading criteria, exams, and other information will be posted at my faculty website at http://www.cs.sjsu.edu/~mak and on the Canvas Learning Management System course login website at http://sjsu.instructure.com. You are responsible for regularly checking these websites to learn of any updates. You can find Canvas video tutorials and documentations at http://ges.sjsu.edu/canvas-students

Course Goals

This course emphasizes learning the Python programming language and its libraries well enough to develop **data analytics notebooks and applications**. Class meetings will encourage high class participation. Mini-labs during each class meeting will give you opportunities to try what you just learned.

Course Learning Outcomes (CLO)

Upon successful completion of this course, students will be able to:

CLO 1: Understand the role of Python in data analytics.
CLO 2: Program in Python proficiently to develop analytics notebooks and applications.
CLO 3: Apply appropriate data structures to implement application data.
CLO 4: Use Python features to implement algorithms to solve analytics problems.
CLO 5: Choose correct Python libraries and call their functions appropriately.
CLO 6: Design and develop user-friendly interactive GUI-based analytics applications.

Recommended Books

Title:	**Python for Programmers with Introductory AI Case Studies**
Authors:	Paul Deitel and Harvey Deitel
Publisher:	Pearson Education, 2019
ISBN:	978-0135224335
	The first part of the class will be based on the first ten chapters of this book.
Title:	**Hands-On Data Analysis with Pandas, second edition**
Authors:	Stephanie Molin
Publisher:	Packt Publishing, 2021
ISBN:	978-1-80056-345-2
	Some of last part of the class will be based on the first two sections of this book.
Title:	**Python for Everyone, Edition 2.0**
Authors:	Cay Horstmann and Rance Necaise
Publisher:	Wiley, 2019
ISBN:	978-1-943153-67-1
	An excellent book for beginning programmers. Dr. Horstmann was a professor in the Computer Science Department at SJSU.

Software to Install

You will need **Python**.

A good way to install Python is via **Anaconda**: https://www.anaconda.com. This will install the Python interpreter, Jupyter notebook, and several key data science libraries.

After installing Anaconda, you can execute the following commands if you need to update the installed packages to their latest versions:

```
conda update conda
conda update --all
```

[Optional] To create animations with **`matplotlib`** inside a Jupyter notebook, you must install the **`ipympl`** tool with these commands:

```
conda install -c conda-forge ipympl
conda install nodejs
jupyter labextension install @jupyter-widgets/jupyterlab-manager
jupyter labextension install jupyter-matplotlib
```

To create Python programs with a standalone GUI (graphical user interface), install the **PyQt5 modules** with this command:

```
pip3 install PyQt5
```

Install **Qt Designer** which will enable you to design applications with windows containing labels, text boxes, buttons, menus, etc.: https://build-system.fman.io/qt-designer-download

Useful Websites
- Python Syntax
  https://www.w3schools.com/python/python_syntax.asp
- **PyQt5 Tutorial**
  https://www.tutorialspoint.com/pyqt5/index.htm

Course Requirements and Assignments
Weekly assignments will provide practice with Python programming and give you experience developing code for analytical applications. Each assignment will be worth up to 100 points. Late assignments will lose 20 points and an additional 20 points for each 24 hours after the due date.

The university's syllabus policies:

- University Syllabus Policy S16-9 at http://www.sjsu.edu/senate/docs/S16-9.pdf.

- Office of Graduate and Undergraduate Program's Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/

 "Success in this course is based on the expectation that students will spend, for each unit of credit, a minimum of 45 hours over the length of the course (normally three hours per unit per week) for instruction, preparation/studying, or course related activities, including but not limited to internships, labs, and clinical practica. Other course structures will have equivalent workload expectations as described in the syllabus."

Technology Requirements
Students are required to have an electronic device (laptop, desktop, or tablet) with a camera and microphone. SJSU has a free equipment loan program available for students: https://www.sjsu.edu/learnanywhere/equipment/index.php

Students are responsible for ensuring that they have access to reliable Wi-Fi during tests. If students are unable to have reliable Wi-Fi, they must inform the instructor, as soon as possible or at the latest one week before the test date to determine an alternative. See Learn Anywhere website for current Wi-Fi options on campus.

Quizzes

Daily online quizzes will test understanding (not memorization) of the material taught during each lecture. Instant messaging, e-mails, texting, tweeting, file sharing, or any other forms of communication with anyone else during the quizzes will be strictly forbidden.

The quizzes take the place of a midterm and final exam (i.e., no midterm or final exams).

Grading Information

Individual total scores will be computed with these weights:

> 60%   Assignments
> 40%   Quizzes

This class is not graded but instead is credit/no credit (CR/NC). Programming assignments will be scored not only on correct output, but also how well the programs were written based on rubrics to be posted to Canvas for each assignment. Your total weighted score at the end of the semester must be at least 70 out of 100 to earn a CR.

University Policies

Per University Policy S16-9, university-wide policy information relevant to all courses, such as academic integrity, accommodations, etc. will be available on Office of Graduate and Undergraduate Program's Syllabus Information web page at http://www.sjsu.edu/gup/syllabusinfo/.

DATA 200
Computational Programming for Analytics

Section 21
Spring 2024
Instructor: Ron Mak

Course Schedule (subject to change with fair notice)

Week	Date	Topics
1	Jan 25	Overview of the course Waiver test
2	Feb 1	**Python fundamentals** The role of Python in data analytics Introduction to Python programming Reading user input
3	Feb 8	Python control statements
4	Feb 15	Functions
5	Feb 22	Sequences: lists and tuples
6	Feb 29	Dictionaries and sets
7	Mar 7	Array-oriented programming with `numpy`
8	Mar 14	Strings: a deeper look Files and exceptions
9	Mar 21	Files and exceptions, *cont'd* Object-oriented programming
10	Mar 28	Object-oriented programming, *cont'd* Duck typing and operator overloading Introduction to Python collections Unit testing
		Spring break, April 1 – April 5
11	Apr 11	**Python for data analytics applications** Advanced `numpy` `pandas` series and dataframes
12	Apr 18	Data visualization with `matplotlib`
13	Apr 25	Plotting with `pandas` and `seaborn`
14	May 2	Recursion Data structures and more collections Sorting and searching algorithms
15	May 9	Algorithms, *cont'd* Lambda and generator expressions GUI programming with `PyQt`