Data Science Program Columbian College of Arts and Sciences	Data Science Program **DATS 2103, Spring 2026** **Data Mining in Data Science** **3 credits** **Undergraduate course**

Instructor
Max Alekseyev, Ph.D.
Associate Professor of Mathematics and Computational Biology
Office: SEH 7685
Phone: 571-553-0259
E-mail: maxal@gwu.edu
Office hours: by appointment

Course Description
This course is an introductory undergraduate course on data mining at the GW Data Science Program. It introduces the basic concepts, principles, methods, and applications of data mining, with a focus on Python programming language for implementation of the corresponding algorithms. The objective of the course is to give students an overview of data mining techniques and available machinery as well as skills to explore, analyze, and leverage data. At the beginning of the course we will cover basics of Python to perform pre-processing and data wrangling, and then we will review 'core' data mining topics, such as regression and classification techniques. Students will use Python to complete the homework assignments through the course.

Course Prerequisite(s)
CSCI1012, DATS1001, MATH1232, STAT 1051/1053/1111/1127

Learning Outcomes. Students will become familiar with
1. basic data preprocessing and analysis techniques;
2. applications of Python programming language in data science;
3. basic data mining concepts and algorithms;
4. core concepts of regression and classification;
5. visualization techniques for variety of applications.

Materials
- Python for Scientists
- Think Python 2e https://greenteapress.com/wp/think-python-2e/
- An Introduction to Statistical Learning https://statlearning.com/

Methods of Evaluation	**Percent of Grade**
6 take-home assignments	90
Participation	10

Grading Scale and Standards

A:	93-100%	C:	73-76%
A-:	90-92%	C-:	70-72%
B+:	87-89%	D+:	67-69%
B:	83-86%	D:	63-66%
B-:	80-82%	D-:	60-62%
C+:	77-79%	F:	59% or below

Workload:

This 3-credit course includes 2.5 hours of direct instruction and 5-7 hours of independent learning per week.

Class Policy: Homework Assignments

Homework assignments will be composed of a set of problems. Students are given at least a week to solve each assignment. To solve a problem students will typically need to write a code in Python, process given data, and submit the result along with the program code. Although students may discuss homework assignments with other students, program codes must be their own work. Copying another student's solution or letting someone else copy own code represent cheating and violation of academic integrity for both parties.

Class Policy: Late Work

Under extenuating circumstances a student may ask the instructor for extended time to complete the assignment. It is the instructor's choice to grant an extension or not. No late assignments will be accepted without advance permission.

Class Policy: Use of Generative Artificial Intelligence

By submitting work for evaluation in this course, you represent it as your own intellectual product. You may not submit for evaluation any content (e.g., ideas, text, code, images) that was generated, in whole or in part,by Generative Artificial Intelligence tools (including, but not limited to, ChatGPT and other large language models). Doing so in this course constitutes *cheating* under the George Washington University *Code of Academic Integrity*.

Tentative Outline

Week 1: Python programming (basic)
Week 2: Python programming (intermediate)
Week 3: Python programming (advanced)
Week 4: NumPy module
Week 5: Visualization in Python (Matplotlib, Seaborn)
Week 6: Data Science in Python (Pandas)
Week 7: Review of Linear Algebra
Week 8: Applied math in Python (SciPy)
Week 9: Data Preprocessing
Week 10: Basics of Machine Learning
Week 11: Linear Regression
Week 12: Decision Trees and Random Forests
Week 13: Support Vector Machines and Bayesian Inference
Week 14: Clustering methods

University policies:

University policy on observance of religious holidays

In accordance with University policy, students should notify faculty during the first week of the semester of their intention to be absent from class on their day(s) of religious observance. For details and policy, see: http://students.gwu.edu/accommodations-religious-holidays

Academic integrity code

Academic dishonesty is defined as cheating of any kind, including misrepresenting one's own work, taking credit for the work of others without crediting them and without appropriate authorization, and the fabrication of information. For details and complete code, see: http://studentconduct.gwu.edu/code-academic-integrity

Safety and security

In the case of an emergency, if at all possible, the class should shelter in place. If the building that the class is in is affected, follow the evacuation procedures for the building. After evacuation, seek shelter at a predetermined rendezvous location.

Support for students outside the classroom

Disability Support Services (DSS)

Any student who may need an accommodation based on the potential impact of a disability should contact the Disability Support Services office at 202-994-8250 in the Rome Hall, Suite 102, to establish eligibility and to coordinate reasonable accommodations. For additional information see: http://disabilitysupport.gwu.edu

Mental Health Services 202-994-5300

The University's Mental Health Services offers 24/7 assistance and referral to address students' personal, social, career, and study skills problems. Services for students include: crisis and emergency mental health consultations confidential assessment, counseling services (individual and small group), and referrals. For additional information see: http://counselingcenter.gwu.edu