Course: **Analytics Programming (DAV 5400)**
Credits: 3 Credits / Graduate
Pre/Coreqs: N/A
Instructor: James Topor
Instructor Contact: james.topor@yu.edu

COURSE OVERVIEW

Code-based solutions can be richer, more accurate, and more flexible than those that rely on off-the-shelf software and analytic packages. This course teaches the programming skills that data analysts need to prepare structured and unstructured data for downstream analysis. Students will learn to use high-level programming languages to create rich data analysis workflows.

COURSE LEARNING OUTCOMES

By the end of this course, students will be able:

- Write programs to clean, filter, aggregate, restructure, and combine data.
- Load data from both structured and unstructured sources.
- Perform basic exploratory data analysis.
- Apply principles of scientific method and reproducible reporting to the development and presentation of code-based solutions.
- Create high quality explanatory narratives and visualizations in support of reproducible analytical work

REQUIRED MATERIALS

- Wes McKinney, *Python for Data Analysis*, 2nd edition, O'Reilly. 2018.
- Mark Pilgrim, *Dive Into Python 3*, freely available web-based content: http://diveintopython3.problemsolving.io/table-of-contents.html

Web-based readings and videos on related topics will also be assigned.

Relevant Software, Hardware, or Other Tools:
We will make use of Python via the freely available **Anaconda** environment, including Jupyter Notebooks and the Spyder IDE. Students are also welcome to use Google's **Colab** platform where feasible. Details for obtaining and installing the appropriate software will be provided in the course materials. All of the software will work on (or from) both PCs and Macs.

ASSIGNMENTS & GRADING

Approach to Assignments. All Python-based projects and assignments are to be written in IPython (Jupyter) notebooks. All assignments and projects will be submitted directly within the DAV 5400 Canvas portal.

Evaluation Criteria. All course assignments and projects will be evaluated like work assignments from a demanding employer. The primary evaluation basis is adherence to the deliverables stated in each assignment's functional requirements. To achieve a top grade, students must also adhere to best practices for software engineering principles, including reproducibility; following appropriate coding guidelines; and DRY. Furthermore, assignments must be clearly and

concisely written using proper English language grammar and should present relevant supporting text in a logical flow. Presentations should include an appropriate level of detail for the intended audience.

Assignments	Grading
Discussions / Response Assignments (8 x 10pts) The eight discussions will focus primarily on use cases related to the Modules to which they apply. Students will prepare short responses to these discussion questions, which will be used to prompt group discussion.	8%
Assignments (6 x 100pts) On most weeks when projects are not due, there will be short-form assignments to help reinforce the current learning material. These assignments may include completing tasks using course analytical tools. Some assignments may require working in small groups.	25%
Projects (4 x 100pts) Students will work individually and in teams on four analytical programming projects. At the end of the course, each student will have a portfolio of increasingly complex projects ready to show an employer.	32%
Final Project (150pts) and Presentation (50pts) Working individually or as part of a small team, students will create a requirements document that outlines a *useful* analytical application that will be applied to available source data. They will then implement the application described in their requirements document. Students will present their final projects to their peers for feedback.	*20%*
Final Exam (100 Pts) Students will address a series of practical challenges derived from the content of DAV 5400 Course Modules 1 through 14.	*15%*

- All projects and assignments, unless otherwise noted, are due end of day on Sundays.

- Each week's materials will be made available via Canvas no later than the previous Friday at 6:00 a.m. ET.

- **Course Completion Requirements**: As a prerequisite to passing this course, you must complete all five projects (including the final), and make the final presentation during the final class session. Failure to either submit any one of the five projects or present your final project will preclude you from achieving a passing grade in this course. Please note that completion of the five projects is not the sole determinant of whether you will receive a passing grade: however, failure to submit any one of the five will prevent you from achieving a passing grade.

- **Discussions / Response Assignments**: While this material is important, please note that this work only makes up 8% of your grade. Please do the readings, and participate in the discussions and any discussion-related group assignments. If you have limited time for the course, please remember to invest the majority of your efforts in completing the projects and assignments. The assignments merit close attention because they will help you to be successful on the projects.

- **Reproducibility Requirement, Testing Requirement, But Not Perfection!** Students are responsible for providing all code and data so that your work can be reproduced by others. If you turn in code that does not run, you will not receive credit, unless you also include an explanatory note at the time of submission. At the same time, you don't need to turn in perfect code. Generous partial credit will be given for deliverables that are timely, tested, and reproducible.

- **Policy on Sharing and "Stealing" Code.** In this course, you may collaborate and you may take base code from whatever sources you wish. But **you must document what you started with, and what you added**, so you are graded on your own contributed work! Furthermore, **use of any automated/GPT tools for any aspect of your assignment, discussion, project, or exam work will be treated as CHEATING, and is strictly prohibited**. Failure to provide proper citations for any third party components of the content you submit will be treated as a violation of the Katz School's **Student Code of Conduct** and will be treated accordingly.

- **Late work policy**. Please note: **Assignments, discussion responses, exams, and projects cannot be accepted after their due dates for any reason**. Any assignment, discussion, exam, or project that is not submitted before its associated deadline will automatically be assigned a grade of **ZERO**. **Late submits will be neither reviewed nor graded**. You will enhance your chances for success in this class if you start early, and turn in your work on time (even if it's not perfect!).

- Students that complete all work in a satisfactory and timely manner will earn a maximum grade of A-. To earn a grade of A in *Analytics Programming*, you'll need to demonstrate work above and beyond what is expected.

GRADING SCALE:

ACADEMIC GRADES				ADMINISTRATIVE GRADES	
Quality of Performance	**Letter Grade**	**Range %**	**GPA/Quality Pts.**	**GRADE**	**DESCRIPTION**
Excellent - work is of exceptional quality	A	94- 100	4	G	Stopped attending without filing an official withdrawal form (counted as failure)
	A-	90 – 93.9	3.7	I	Incomplete
Good - work is above average	B+	87 - 89.9	3.3	L	Audit (no credit)
Satisfactory	B	83 - 86.9	3	W	Withdrawal without penalty or prejudice
Below Average	B-	80 - 82.9	2.7		Note that credit is given only for grades A through C and P. No credit is given for grades F, G, I, L, N, or W.
Poor	C+	77 - 79.9	2.3		
	C	70 - 76.9	2		
Failure	F	< 70	.000		

How This Course Works:

Class Sessions are held on campus every week on *Tuesdays from 5:30pm to 7:30pm ET*, **with the exception of Katz School official holidays**. You are strongly encouraged to attend these weekly classes since each will include opportunities for hands-on learning via in-class assignments and case studies as well as a presentation / demonstration of many of the concepts you will need to use for any assignment or project due that week. You are also required to bring your laptop to these Live Sessions as this will serve to facilitate the hands-on learning segments. Class dates can be found in the Course Schedule shown on the following page.

Office Hours can be scheduled by appointment. If you need extra help and are willing to invest the time and effort to be successful, the instructor will make time available to help you. ***But...***you should not be asking for extra help on a project or assignment the day before or the day it is due, since this will indicate that you are not investing the time and effort required to be successful in the course.

You are encouraged to ask questions on Canvas where other students will be able to benefit from your inquiries. For the most part, you can expect me to respond to questions asked either via email or via Canvas within one business day.

KATZ SCHOOL CLASS ATTENDANCE POLICY

For programs that meet in face-to-face format, students are expected to attend all scheduled classes in their entirety. Students who fail to fulfill this requirement will receive an academic penalty appropriate for the course work missed.

Students may not miss 20% (3 classes) or more of their scheduled class. If a student misses 20% or more of a course during the semester, they will receive a final grade of "F." This grade will be reflected on the student's official university transcript.

For programs within clinical components students may not miss 20% or more of any course, clinical or not. At the Katz School, this pertains to only to students in the Speech Language Pathology program. If a student misses 20% or more of a course during the semester, they will receive a final grade of "F." This grade will be reflected on the student's official university transcript.

If the student is absent because of a disability which is documented with the Office of Disability Services at Yeshiva, falls ill or there are other extenuating circumstances, the student must inform the instructor in advance. The instructor may require appropriate documentation to make any exception to this policy.

COURSE SCHEDULE

Students should expect to spend <u>a minimum</u> of 9 hours each week outside of the classroom sessions on the materials, assignments, discussions, and projects required for this course.

Module	TOPIC	SCHEDULE OF MAJOR ASSIGNMENTS
Module 1 Jan 15 – Jan 21 Class: **T Jan 16**	Python Basics: A Refresher (syntax, data types, control flow)	M1 Assignment
Module 2 Jan 22 – Jan 28 Class: **T Jan 23**	Python Data Structures + Functions (includes list comprehensions)	M2 Assignment
Module 3 Jan 29 – Feb 4 Class: **T Jan 30**	Intro to Numpy: Fast Vectorized Operations	M3 Assignment
Module 4 Feb 5 – Feb 11 Class: **T Feb 6**	Pandas Series and Dataframe Objects	M4 Assignment
Module 5 Feb 12 – Feb 18 Class: **T Feb 13**	Exploratory Data Analysis	
Module 6 Feb 19 – Feb 25 Class: **T Feb 20**	Creating Visualizations in Python ***** Final Project Requirements Distributed *****	**Project 1 Due**
Module 7 Feb 26 – Mar 3 Class: **T Feb 27**	Text Processing: Regular Expressions	M7 Assignment
Module 8 Mar 4 – Mar 10 Class: **T Mar 5**	Working with Web Data	
Module 9 Mar 11 – Mar 17 Class: **T Mar 12**	Data Cleaning & Preparation	**Project 2 Due** **** 1st Draft of Final Project Proposal Due ****
Module 10 Mar 18 – Mar 24 Class: **T Mar 19**	Combining and Reshaping Data in Pandas	M10 Assignment
Module 11 Mar 25 – Mar 31 Class: **T Mar 26**	Data Aggregation & Grouping in Pandas	**Final Project Proposal Due** **Project 3 Due**
Module 12 Apr 1 – Apr 7 Class: **T Apr 2**	Modeling + Machine Learning with Python Libraries	
Module 13 Apr 8 – Apr 14 Class: **T Apr 9**	Text Mining	**Project 4 Due**
Module 14 Apr 15 – Apr 21 Class: **T Apr 16**	Ethics In Statistics & Data Ethics **FINAL EXAM** (Online via Canvas, Thur 3pm – Fri 3pm)	**Final Exam**
Apr 22 – Apr 30	****** NO CLASSES: PASSOVER BREAK ******	
Module 15 **W** May 1– May 8 Class: **T May 7**	Final Project Presentations + Writeups Due **** Final Project Presentations Tuesday May 7****	**** Final Project Writeups Due Monday May 6 ****

ONLINE LEARNING POLICIES

Online Learning Formats
Your course consists of two online learning formats:
- **Synchronous Learning**: While our class sessions will be held on campus, we may occasionally be required to meet via a live real time sessions using Zoom (webinar system). During these sessions, we will be able to see and talk with each other. Attendance is required.
- **Asynchronous Learning**: Pre-created content such as videos, assignments, links and articles. There will also be the use of community and collaboration tools like discussion boards and group tools. These sessions are not in real-time but rather involve engagement over the course of each week.

Online Learning Engagement Policy
A successful online class only happens when there is an active community. Students are required to attend both the weekly live synchronous sessions and participate in other community building activities such as the discussion boards.

Netiquette
Netiquette is a set of rules for behaving properly in an online course. Often the anonymity of online courses can cause a lapse in judgement when learners are excited or passionate about a subject. This can lead to statements that could be demeaned as offensive. You are all adults and are treated as such. However, it is still important to talk about these issues. The following bullet points cover some basics communicating in an online course:

- Be sensitive to the fact that there will people with different cultural and linguistic backgrounds, as well as different political and religious beliefs.
- Use good taste when composing your responses in Discussion Forums. Swearing and profanity is also part of being sensitive to your classmates and should be avoided.
- Don't use all capital letters when composing your responses as this is considered "shouting" on the Internet and is regarded as impolite or aggressive.
- Be respectful of your others' views and opinions. Avoid "flaming" (publicly attacking or insulting) them as this can cause hurt feelings and decrease the chances of getting all different types of points of view.
- Be careful when using acronyms. If you use an acronym it is best to spell out its meaning first, then put the acronym in parentheses afterward, for example: Frequently Asked Questions (FAQs). After that you can use the acronym freely throughout your message.
- Use good grammar and spelling (avoid using text messaging shortcuts).
- If you aren't sure what someone meant, consider asking for clarification.
- Remember that your peers are not required to respond to your specific post, so don't be offended if your question goes unanswered.

UNIVERSITY POLICIES & RESOURCES

ACCESSIBILITY AND ACCOMMODATIONS
The Office of Disability Services collaborates with students, faculty and staff to provide reasonable accommodations and services to students with disabilities. Students with disabilities who are enrolled in this course and who will be requesting documented disability-related accommodations should make an appointment with the Office of Disability Services, (646) 592-4132, rkohn1@yu.edu, during the first week of class. Once you have been approved for accommodations, please submit your accommodation letter to ensure the successful implementation of those accommodations. For more information, please visit:
http://yu.edu/Student-Life/Resources-and-Services/Disability-Services/

ACADEMIC INTEGRITY
The submission by a student of any examination, course assignment, or degree requirement is assumed to guarantee that the thoughts and expressions therein not expressly credited to another are literally the student's own. Evidence to the contrary will result in appropriate penalties.

Academic integrity is a set of responsibilities and standards to facilitate high academic quality and rigor with the purpose of clarifying expectations and student conduct. The submission by a student of any coursework, or degree requirement is assumed to guarantee that the thoughts and expressions therein not expressly credited to another are literally the student's own. Examples of violations on academic integrity are, but not limited to:

- Cheating
- Plagiarism
- Dishonesty
- Assisting or attempting to assist another student in an act of academic dishonesty
- Providing papers, essays, research, or other work to aid another student in Intentional Misrepresentation
- Engaging in unauthorized cooperation with other individuals in completing assignments or examinations
- Submitting the same assignment, in part or whole, in more than one course, whether at YU or another institution, without prior written approval from both faculty members.

These Academic Integrity principles are incorporated within the Student Code of Conduct/Code of Ethics, with which all students are required to adhere to at all times.

For more information, visit http://yu.edu/registrar/grad-catalog/

Disciplinary Actions for Violations of Academic Integrity Standards
In this course, disciplinary actions for violations of the Student Code of Conduct/Code of Ethics are as follows:

- **For the first incident**: A grade of **zero** will be assigned for the assignment, project, or discussion. **The student's final grade for the course will also be reduced by a full letter grade**, i.e., if, at the end of the semester, the student would have otherwise earned a grade of A-, that grade will be lowered to a grade of B-.

- **For the second incident**: Student will be **assigned a final grade of 'F' for the course**, and no further assignment/project/discussion work will be accepted from the student for the balance of the semester.

YU Refund Policy
You should be aware of the universities refund policy. Please review this information:
https://www.yu.edu/osf/undergraduate-accounts/withdrawal.

Academic Calendar
You should review the academic calendars, including add/drop dates. Please review this information: https://www.yu.edu/registrar/grad-calendar.

STUDENT SUPPORT SERVICES
Katz School offers academic support through the Learning Hub. This support service includes writing, academic integrity (APA format), English as a Second Language, and general academic tutoring. For more information, please contact katz@yu.edu.

If you need any additional help, please visit Student Support Services:
http://yu.edu/academics/services/