FLORIDA POLYTECHNIC UNIVERSITY

Syllabus: CAP 5320 – Data Wrangling and Exploratory Data Analysis

Spring semester 2026

Course Information

- **Course Number and Title: CAP 5320 – Data Wrangling and Exploratory Data Analysis**
- **Credit Hours:** 3 credits
- **Academic Term: Spring 2026**
- Class Meeting: Wednesdays 5:30 PM – 8:15 PM BARC-1158

Instructor Information

- **Instructor:** Abdulaziz Alhamadani, Ph.D.
- **Office Location:** ARC-1108
- **Office Hours**:
  - Tuesday 2:00 PM - 3:00 PM,
  - Wednesday 2:00 PM - 3:00 PM,
  - Thursday 1:30 PM - 3:00 PM
  - Or by appointment
- **Office Phone:** 863-874-8656
- E-mail: aalhamadani@floridapoly.edu

Course Delivery and Course Description

- **Delivery Mode**: face-to-face learning experience with class meetings twice a week. Please check the Canvas course website for all information, including announcements, discussions, and any supplementary material for topics covered in this course.

- **Classroom time/room**: **Wed 5:30 PM – 8:15 PM, BARC-1159**

- **Course Website**: https://floridapolytechnic.instructure.com/courses/8798

- **Official Catalog Course Description**:
  Preprocessing tasks often consume a large fraction of time in computational projects, and all downstream analyses depend on them. In this course, students will develop practical skills for working with large datasets. Topics will include common methods for gathering, organizing, and reshaping structured and unstructured data. We will also cover methods of exploratory data analysis that are useful to guide more focused questions and models. These include principles of information display, simple model forms and data reduction, common visualization methods, and reporting tools.

- **Communication/Computation Skills Requirement (6A-10.030)**: No

- Required Texts and Materials: None

All required classroom material will be provided in class or online.

Supplementary Textbooks and Materials:
Some supplementary resources are listed below.
"Data Visualization: A Practical Introduction" by K. Healy. Princeton University Press, 2018.
ISBN-13: 978-069118162. Available at: ***https://socviz.co/***

(R4DS) "R for Data Science" by Hadley Wickham, Mine Cetinkaya-Rundel, and Garrett Grolemund. Available online at: https://r4ds.hadley.nz/
ISBN-13: 978-1492097402, 2nd Edition. O'Reilly Media.

(PyDA) *"Python for Data Analysis: Data Wrangling with pandas, NumPy & Jupyter"* by Wes McKinney. Available online: https://wesmckinney.com/book/
ISBN-13: 978-1098104030, 3rd Edition. O'Reilly Media.

(Data Carpentry) *"R for Geospatial Data".*
Available at: https://datacarpentry.org/r-intro-geospatial/

(Data Carpentry) *"Geospatial Raster and Vector Data".*
Available at: https://datacarpentry.org/r-raster-vector-geospatial/

Other recommended materials:

"Pandas for Everyone: Python Data Analysis" by Daniel Y. Chen
ISBN-13: 978-0134546933. Addison-Wesley Data & Analytics Series (2018)

"The Art of Data Science" by R. D. Peng, E. Matsui
Available online: https://leanpub.com/artofdatascience

Equipment and Materials: R, Python, and other specialized analysis toolkits to synthesize concepts from data analytics and data science as applied to relevant projects. The format of the course will include lectures by the instructor, class discussions, directed readings, and students' presentations.

Course Objectives and Outcomes

- **Course Objectives**:
  Students will learn the fundamental skills required to acquire, munge, transform, manipulate, and visualize data in a computing environment that fosters reproducibility. Students will learn to import and manage structured and unstructured data; manipulate, transform, and summarize data, join disparate data sources, and learn best practices for exploratory data analysis.

- **Course Learning Outcomes**:

  1. **Examine** methods for data collection and aggregation. (Analyze)
  2. **Use** common techniques for preparing and processing data. (Apply)
  3. **Demonstrate** working knowledge of methods for data analysis and data management. (Apply)

- Alignment with Program Outcomes:

	Course Learning Outcome and Learning Level*		
MS in Data Science PLOs	1	2	3
(1) Demonstrate mastery in analyzing complex problems and applying knowledge of data science to formulate solutions.	Analyze		Apply
(2) Communicate data science information clearly and effectively through presentations and technical writings to both expert and non-expert audiences.		Apply	Apply
(3) Demonstrate critical evaluation of recent research literature.			
(4) Identify a novel relevant research problem in a chosen data science research field, perform the literature survey for the problem, create a plan to solve the problem, carry on the plan, and defend the research.			
(5) Recognize appropriate practices in the field of data science and their ethical implications.		Apply	

*: learning level as described in Bloom's taxonomy and Anderson and Krathwohl's taxonomy.

	Course Learning Outcome and Learning Level*		
MS in Engineering Management PLOs	1	2	3
(1) Apply business fundamentals and develop strategies to address complex challenges in engineering, applied science, and related industries.			
(2) Identify, formulate, and solve engineering problems of single or multidisciplinary nature by applying principles of engineering, science, mathematics, and analytics.	Analyze		
(3) Apply data science concepts, tools, and develop analytical insights to solve business and engineering problems.		Apply	Apply
(4) Communicate complex business and engineering problems to diverse audiences.		Apply	Apply

*: learning level as described in Bloom's taxonomy and Anderson and Krathwohl's taxonomy.

	Course Learning Outcome and Learning Level*		
MS in Computer Science PLOs	1	2	3
(1) Demonstrate mastery in analyzing complex problems and applying knowledge of computer and/or data science to formulate solutions.	Analyze		Apply
(2) Communicate computer and/or data science information clearly and effectively through presentations and technical writings to both expert and non-expert audiences.		Apply	Apply
(3) Demonstrate critical evaluation of recent research literature.			
(4) Identify a novel relevant research problem in a chosen computer and/or data science research field, perform the literature survey for the problem, create a plan to solve the problem, carry on the plan, and defend the research.			
(5) Recognize appropriate practices in the different fields of computer science and their ethical implications		Apply	

*: learning level as described in Bloom's taxonomy and Anderson and Krathwohl's taxonomy.

Academic Support Resources

- **Library**: Students can access the Florida Polytechnic University Library through the University website and Canvas, on and off campus. Students may direct questions to library@floridapoly.edu.
- **Peer Learning Strategists (PLS)**: Are specially trained student leaders who help their peers strategize approaches to course content and work through solution methods. PLS work in collaboration with the courses they support so the content and methods are aligned with your

instructors' expectations. Students can meet with a PLS in The Learning Center, which is located on the first floor of the Innovation, Science and Technology (IST) building in room 1019.

- **Academic Success Coaches**: All students at Florida Poly are assigned an Academic Success Coach. Your Academic Success Coach can assist you with academic success strategies. Please visit the Student Success Center on the second floor of the IST building to meet with an Academic Success Coach.
- **Writing Center**: Located on the second floor of the IST (2059/2061), the Writing Center helps students to develop their writing and presentation skills. Consultations are available in person and virtually. For more detail, visit https://floridapoly.edu/writingcenter.

Civility and Collegiality

Faculty and students come to the university for the same reason, which is to participate in a highly professional educational environment. To that end, both students and faculty are expected to treat each other with mutual regard and civility. Communication, written, oral and behavioral, between faculty and students must remain respectful. Within and outside of the classroom, students must refrain from derogatory comments toward the faculty member and their fellow students, and faculty as well must refrain from derogatory comments toward their students. Civility and collegiality also include respecting each other's time: for example, neither students nor faculty should arrive late to class (unless unforeseen, pressing circumstances prevail); faculty should be present at the posted office hours; and students and faculty should be punctual when meeting times are scheduled. In more general terms, collegiality means respecting the right of both faculty and students to participate fully and fairly in the educational enterprise.

Course Policies

Attendance

- Students in face-to-face (this includes labs and C-courses) courses are expected "to attend all of their scheduled University classes and to satisfy all academic objectives as defined by the instructor" (University Policy, FPU-5.0010AP).
- If you know that you will miss a class for any reason discuss the situation with your instructor in a timely manner. Exceptions to any attendance requirements may be made on a case-by-case basis.

Late Work/Make-up work

Each student must keep current on assignments. Late assignments are not graded unless *permission has been obtained from the instructor for an extension prior to the due date*. In case of a medical emergency, please notify your instructor as soon as possible who will evaluate any exceptions on a case-by-case basis.

Grading Scale

(See also University Grading Policy).
Grades will be determined according to the following scale:

A	93% – 100%	**B**	83% – 85%	**C**	73% – 75%	**D**	63% – 65%
A–	90% – 92%	**B–**	80% – 82%	**C–**	70% – 72%	**D–**	60% – 62%
B+	86% – 89%	**C+**	76% – 79%	**D+**	66% – 69%	**F**	0% – 59%

Grading Information Specifically for Graduate Students

The grades of "A" through "C," and "SR" are passing grades. The grades of "B-," "C+," and "C" are considered passing for graduate students but indicate weak performance for a graduate student and may not be accepted for some programs.

Note: The grades of "C-," "D+," "D," "D-," "F," and "UR" are failing grades.

Assignment/Evaluation Methods

Assignment	*Percentage*	*Points*
Attendance and Participation	5%	5
Discussion	5%	5
Midterm Exam	20%	20
Quizzes & In-class Assignments	15%	15
Homework Assignments	15%	15
Exam 2 (Final)	20%	20
Final Project	20%	20
Total	100%	100

Participation in all course activities is a very important element of this course, is a basic expectation, and counts for part of your grade. Course participation consists of active and respectful involvement in class discussions, presentations, peer feedback, postings, replies, projects, and other interactions. The course grade takes into account quality, quantity, and timeliness of student participation.

Your instructor will provide you with specific guidelines for the final project report and final project presentation shortly after the first few weeks of classes (format and length, call for proposals, reference materials, presentation guidelines and logistics, rubric, etc.)

University Policies

Reasonable Accommodations

The University is committed to ensuring equal access to all educational opportunities. The University, through the Office of Disability Services (ODS), facilitates reasonable accommodations for students with disabilities and documented eligibility. It is the student's responsibility to self-identify as a student with disabilities and register with ODS to request accommodations.

If you have already registered with ODS, please ensure that you have requested an accommodation letter for this course through the ODS student portal and communicate with your instructor about your approved accommodations as soon as possible. Arrangements for testing accommodations must be made in advance. Accommodations are not retroactive.
If you are not registered with ODS but believe you have a temporary health condition or permanent disability requiring an accommodation, please contact ODS as soon as possible.

The Office of Disability Services (ODS):
DisabilityServices@floridapoly.edu
(863) 874-8770
The Access Point
ODS website: www.floridapoly.edy/disability

Accommodations for Religious Observances, Practices and Beliefs

The University will reasonably accommodate the religious observances, practices, and beliefs of individuals in regard to admissions, class attendance, and the scheduling of examinations and work assignments. (See University Policy.)

Title IX

Florida Polytechnic University is committed to ensuring a safe, productive learning environment on our campus that prohibits sex discrimination and sexual misconduct, including sexual harassment, sexual assault, dating violence, domestic violence and stalking. Resources are available if you or someone you know needs assistance. Any faculty or staff member you speak to is required to report the incident to the Title IX Coordinator. Please know, however, that your information will be kept private to the

greatest extent possible. You will not be required to share your experience. If you want to speak to someone who is permitted to keep your disclosure confidential, please seek assistance from the Florida Polytechnic University Ombuds Office, BayCare's Student Assistance Program, 1-800-878-5470 and locally within the community at Peace River Center, 863-413-2707 (24-hour hotline) or 863-413-2708 to schedule an appointment. The Title IX Coordinator is available for any questions to discussion resources and options available.

Academic Integrity

The faculty and administration take academic integrity very seriously. Violations of academic integrity regulation include actions such as cheating, plagiarism, use of unauthorized resources (including but not limited to use of Artificial Intelligence tools), illegal use of intellectual property, and inappropriately aiding other students. Such actions undermine the central mission of the university and negatively impact the value of your Florida Poly degree. Suspected violations will be fully investigated, possibly resulting in an academic integrity hearing and sanctions against the accused student if found in violation. Sanctions range from receiving a zero on the exam or assignment, to expulsion from the university. Repeat offenders are subject to more severe sanctions and penalties. Do not compromise your integrity for a perceived short-term gain.

Recording Lectures

Students may, without prior notice, record video or audio of a class lecture for a class in which the student is enrolled for their own personal educational use. Recordings may not be used as a substitute for class participation or class attendance. Recordings may not be published or shared in any way, either intentionally or accidently, without the written consent of the faculty member. Failure to adhere to these requirements is a violation of state law (subject to civil penalty) and the student code of conduct (subject to disciplinary action).

Recording class activities other than class lectures, including but not limited to lab sessions, student presentations (whether individually or part of a group), class discussion (except when incidental to and incorporated within a class lecture), and invited guest speakers is ***prohibited****.*

Course Schedule

A **tentative course calendar** is included below.

Week	*Date*	*Topic*
1	01/14	*Overview* of data science process and importance of data wrangling and EDA Tools setup (RStudio, packages, git, GitHub, etc.) Working with different data structures
2	01/21	*Data Import* Using Markdown for Documentation Documenting, commenting, and sharing code
3	01/28	*Review of Foundational Programming Skills* Tidy data principles. Creating efficient code Control statements, iterations, writing custom functions
4	02/04	*Understanding Data* Types of data, the data generation process Finding and interpreting data
5	02/11	*Data transformation* Reshaping data. Changing schemas as needed Summarizing, sorting, subsetting, merging multiple datasets

Week	*Date*	*Topic*
6	02/18	*Data transformation* Analyzing data frames by group, joining data frames. Working with missing data. Strings as factors.
7	02/25	*Exploratory Data Analysis* Generating questions about data. Measuring variation, exploring covariation. Checking missing values
8	03/04	*Data Visualization for Exploration* Selecting visual layouts. The grammar of graphics. Choosing effective graphical encodings. **MIDTERM (Wednesday March 4th)**
9	03/11	*Dates, times, and other data types* Regular Expressions Working with Spatial data
Spring Break: Monday 03/16 – Friday 03/20 (No classes)		
10	03/25	*Exploratory Data Analysis* Using data to answer questions Case studies
11	04/1	*Relational Data* Sourcing data from REST APIs Processing JSON Data, XML, and others
12	04/8	*Web scraping and dealing with text data* Discussion on other data science tools Spreadsheet Munging Strategies
13	04/15	*Model Basics* Visualizing models, model building *Storytelling with data* Dynamic reports Building interactive web applications
14	04/22	**Final Project Presentations**
Final Exam (Exam 2)		

I reserve the right to modify this schedule as required by the progression of the class.

Important Dates: https://floridapoly.edu/academics/academic-calendar/index.php

January 12 – 16	M-F	Drop/Add Week
January 19	M	Martin Luther King Jr. Holiday – *No Classes*
February 10	T	Career Day – *No Classes*
Mar 3	M	Midterm grades due
March 16 – 20	M – F	Spring Break - *No Classes*
April 17	F	Withdrawal Without Academic Penalty Deadline (W assigned)
April 28	W	Last Day of Classes
April 29 – May 1	W-F	Reading Days - *No Classes*
May 4 – 8	M – F	Final Exams
May 11	W	Final Grades Available Online

Sample Rubric for Report and Presentations

The final presentations and reports will be evaluated using rubrics similar to the ones included below.

Sample Data Analysis Rubric:

Technical requirements	Possible points	Excellent	Good	Needs work
Code and output are included and properly formatted	10			
No error messages are included (or left unexplained)	4			
Generated data visualizations are appropriate, compelling, and well-formatted	15			
Use of tools covered in class is present (other tools may be explored)	10			
Summary statistics, comparison, or any other type of quantitative analysis is present and supports the analysis	15			
Dataset description is included (or referenced) and comments on the attributes are present	10			
Data analysis is fully reproducible (or special requirements for reproducibility are detailed)	6			
Narrative	**Possible points**	**Excellent**	**Good**	**Needs work**
Discussion on findings is sound	5			
The analysis is carefully thought-out	10			
Possible future work is proposed (and/or commentary of limitations is included)	5			
Results are clear and a convincing argument is present	10			

Sample Report Rubric

Objective	Category	Below Expectations	Weak	Average	Good	Excellent
	Score	1	2	3	4	5
Students can write professional quality documents	Introduction	Opening is off-topic and inappropriate to the purpose, not concise and no clarity	Opening is somewhat related to the topic and appropriate to the purpose but is not concise and clear	Opening is related to the topic and appropriate to the purpose. Somewhat clear and concise	Opening is related to the topic and appropriate to the purpose. Clear and concise	Strong opening that is clear and concise
	Organization	Disorganized; incorrect format; unclear direction	Somewhat organized; incorrect format; unclear direction	Organized; correct format; unclear direction	Organized; correct format; clear direction	Correct formatting, strong clarity and organization

						in the development of main points
	Literature Review	Does not present information from any source	Presents information from irrelevant sources representing limited points of view/approaches	Presents information from relevant sources representing limited points of view/approaches	Presents in-depth information from relevant sources representing limited points of view/approaches	Synthesizes in-depth information from relevant sources representing limited points of view/approaches
	Research Design (weighted twice)	Does not provide information on research design	Inquiry design demonstrates misunderstanding of the methodology or theoretical framework	Critical elements of the methodology or theoretical framework are missing, incorrectly developed or unfocused	Critical elements of the methodology or theoretical framework are appropriately developed however, more subtle elements are ignored or unaccounted for	All elements of the methodology or theoretical framework are skillfully developed and may be synthesized from across disciplines or relevant subdisciplines
	Analysis (weighted twice)	Incorrect, Irrelevant, no supporting evidence	Correct, irrelevant, no supporting evidence	Correct, relevant, no supporting evidence	Relevant and correct with supporting evidence	Relevant, correct, complete, incorporates innovative insights
	Next Steps	Missing or content does not support conclusion	Conclusion irrelevant to the findings	Conclusion somewhat relevant to the findings	Conclusion relevant to the findings	Strong conclusion that is clear, complete and compelling
	Grammar & Spelling	Uses language that often impedes meaning due to errors	Uses language that often sometimes meaning due to errors	Uses language that generally conveys meaning to readers with clarity, although writing includes some errors	Uses straightforward language that conveys meaning to readers. Language has few errors	Uses graceful language that communicates meaning to readers with clarity and fluency and is virtually error free
	Reference Style (APA)	Did not follow APA style	Numerous errors in APA style, did not cite sources correctly, formatting issues	Some errors in APA style, cited correctly but formatting issues persist	Minimum errors in style and formatting but does not detract from readability	No errors in APA style
Total points for Report = 50						

Sample Presentation Rubric

Objective	Category	Below Expectations	Weak	Average	Good	Excellent
	Score	1	2	3	4	5
Students can demonstrate mastery of communication technology	Use of Media	Lack of media detracts from the presentation objective	Misuse of media that detracts from the presentation objective	Use of media barely supports and contributes to the presentation objective	Use of media supports and contributes to the presentation objective	Use of media supports, clarifies and reinforces the presentation objective
	Quality of Slides	Very poor quality. Not enough or too much colors, fonts and animations that detract from project objective	Poor quality. Not enough or too much colors, fonts and animations that detract from project objective	Fonts, colors and animations barely support the presentation objective	Fonts, colors and animations support the presentation objective	Fonts, colors and animations support, clarify and reinforce the presentation objective

Students can develop and deliver a compelling oral talk with relevant facts and information	Opening statement	Opening is off-topic and inappropriate to the purpose, not concise and no clarity	Opening is somewhat related to the topic and appropriate to the purpose but is not concise and clear	Opening is related to the topic and appropriate to the purpose. Somewhat clear and concise	Opening is related to the topic and appropriate to the purpose. Clear and concise	Strong opening that is clear and concise
	Organization	Disorganized; incorrect format; unclear direction	Somewhat organized; incorrect format; unclear direction	Organized; correct format; unclear direction	Organized; correct format; clear direction	Correct formatting, strong clarity and organization in the development of main points
	Literature Review	Does not present information from any source	Presents information from irrelevant sources representing limited points of view/approaches	Presents information from relevant sources representing limited points of view/approaches	Presents in-depth information from relevant sources representing limited points of view/approaches	Synthesizes in-depth information from relevant sources representing limited points of view/approaches
	Analysis	Incorrect, Irrelevant, no supporting evidence	Correct, irrelevant, no supporting evidence	Correct, relevant, no supporting evidence	Relevant and correct with supporting evidence	Relevant, correct, complete, incorporates innovative insights
	Next Steps	Missing or content does not support conclusion	Conclusion irrelevant to the findings	Conclusion somewhat relevant to the findings	Conclusion relevant to the findings	Strong conclusion that is clear, complete and compelling
	Timing	Presentation is too short, insufficient coverage of material	Presentation is too long. Unable to cover all the material	Able to cover all the material within five extra minutes	Utilizes allotted time to provide sufficient coverage of material	Well-paced coverage of material within the allotted time
Students can deliver an oral talk with clarity and appropriate poise	Delivery Techniques	Does not participate in the oral presentation	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) detract from the understandability of the presentation, and speaker appears uncomfortable.	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) make the presentation understandable, and speaker appears tentative.	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) make the presentation interesting, and speaker appears comfortable.	Delivery techniques (posture, gesture, eye contact, and vocal expressiveness) make the presentation compelling, and speaker appears polished and confident.
	Peer Evaluation	5 points				
Total Points = 50						