IS 465 Advanced Information Systems

Fall 2024

Instructor: Mark Cartwright (<u>mark.cartwright@njit.edu</u>) TA: Azadeh Naderi (<u>an57@njit.edu</u>) Format: In-person In-person Location: Central King Building 320 Class Hours: Tuesday and Friday 2:30 - 3:50 pm Office Hours: Wednesday 10–11 am, Friday 1:15-2:15 pm in GITC 3902 or by appointment

Course Description

This course introduces students to the world of data analytics from an information systems perspective, focusing on the application of various data analysis techniques in business practices. We cover a wide spectrum of topics ranging from fundamental statistics to databases, data warehouses, data visualization, and data mining. Being an introductory course, our approach is "shallow and wide," emphasizing giving students a complete view of the data analytics profession, covering as many different sub-areas as time allows while not diving too deep into any one specific domain. The goal is to serve as a "guided tour" for students to gain knowledge about the different sub-areas of data analytics and understanding of which area is a best fit for their personal development. More in-depth materials and discussion for each sub-area will be provided upon students' requests. Course topics include the rudiments of probability and random variables, visualization, data warehousing and OLAP analysis, dashboards, scorecards, data mining algorithms, optimization techniques, DSS, and knowledge systems.

Course Goals

At the end of this course, the student should be able to:

- 1. Build a foundation of data analysis such as statistics, probability theories, data structure and algorithms, database and data warehouse, data visualization, basic data mining techniques (e.g. decision trees, clustering, etc), and Python programming.
- 2. Apply them to real-world data sets for data analysis.
- 3. Communicate the results of data analysis.

The students will learn to work with the following tools:

- Excel
- Tableau

- RapidMiner
- Python

Textbook and readings

Considering the wide range of topics covered in this class, we will use PowerPoint slides and a collection of papers/articles/chapters for our class.

Assignments

We will have the following assignments. Details on each assignment will be posted on Canvas.

Labs

There will be several labs IN CLASS. Your participation is expected for these labs, and absence during the lab results in 0 points for that lab. *Laptops will be needed for labs.*

Reading Assignments

There will be several individual assignments over the semester. Some of these are individual assignments, e.g., students will write a short reflection on the reading, and some of these are group assignments, e.g., groups will develop and present a short (10-minute) PowerPoint presentation to summarize the assignment paper.

Technical Assignments

There will also be technical homework assignments in this course. Some of them are individual assignments, and some will be team-based. More details will be posted on Canvas regarding these assignments.

In-class Quizzes

There will be several (approx. 6-8) small in-class quizzes. When there is a quiz, it will be announced in advance.

Projects

Objective: To demonstrate the ability to apply data analytics techniques to solve real-world problems.

Summary: Two projects will be assigned to teams throughout the semester.

Project 1: Data Visualization Group Project

Teams are expected to find an interesting data set and visualize it using Tableau. Each team will then post the visualization model around midterm. Since we have no midterm exam in this class, this will count as the midterm milestone.

Project 2: Data Mining Group Project

Teams are expected to work with a real-world organization to gather a data set, analyze it, and try to extract insightful information/knowledge using RapidMiner and/or Python.

Grading policy

The class will be graded on a 100-point scale.

А	90 and up
B+	85 to 89
В	80 to 84
C+	75 to 79
С	70 to 74
D	65 to 69
F	Less than 65

Point distribution:

- 10 pts. Labs
- 10 pts. Reading assignments
- 10 pts. Technical assignments
- 5 pts. Quizzes
- 40 pts. Group projects
- 25 pts. Final exam

Excellent participation demonstrated by preparation for discussion and thoughtful contributions (online and in class) will have the effect of raising a final letter grade by one value (e.g. B to B+, or B+ to A). Likewise, poor participation demonstrated by consistent lack of preparation for discussion and little or no thoughtful contributions (online and in class) will have the effect of lowering a final letter grade by one value (e.g. A to B+, B to C+).

Late Assignments Policy

Unexcused late submission of homework receives a 20% penalty. This means that you start with 8 out of 10 points as the maximum. Assignments submitted after graded assignments are returned or reviewed in class receive no credit.

Quiz Policy

- Please bring your laptop, as the quiz will often be given through Canvas.
- Please install the LockDown browser. Otherwise, you cannot take the quiz.
- If you miss the quiz with a valid reason (e.g., lateness or absence vilified by DOS) and in advance notice, you can schedule an in-person meeting with the lecturer or TA to take the quiz. But, you will not be able to take the quiz after the results have been announced (most likely, before the next class).

Class environment

Please do your part by seeking to promote the success of others, and by treating each other in ways that respect and celebrate the diversity of talent that is drawn to this field. The classroom is an open forum for discussion, and I encourage all students to feel free to ask questions in class. Please do not be afraid to ask any question, no matter how basic it may seem. What is basic to some of the class may be completely new to the rest.

Accessibility

If you are in need of accommodations due to a disability please contact Chantonette Lyles, Associate Director of the Office of Accessibility Resources & Services (OARS), Fenster Hall Room 260 to discuss your specific needs. A Letter of Accommodation Eligibility from the OARS authorizing your accommodations will be required.

Academic Integrity

Academic Integrity is the cornerstone of higher education and is central to the ideals of this course and the university. Cheating is strictly prohibited and devalues the degree that you are working on. As a member of the NJIT community, it is your responsibility to protect your educational investment by knowing and following the academic code of integrity policy that is found at: NJIT Academic Integrity Code.

Please note that it is my professional obligation and responsibility to report any academic misconduct to the Dean of Students Office. Any student found in violation of the code by cheating, plagiarizing or using any online software inappropriately will result in disciplinary action. This may include a failing grade of F, and/or suspension or dismissal from the university.

If you have any questions about the code of Academic Integrity, please contact the Dean of Students Office at dos@njit.edu

Generative AI Policy

Student use of artificial intelligence (AI) is permitted in this course for certain assignments and activities and is recommended to speed up the development of plotting and data-wrangling code. However, it is recommended to only use such tools as an assistant when coding, since you will be expected to fully understand any code you submit and will be assessed on exams accordingly with that expectation. Furthermore, it is not permitted to be used in reading assignments, reports, and presentations (with the exception of generated images in presentations), as this would impact the student's own efforts and understanding.

Feedback

I will solicit (anonymous) feedback from students throughout the course through anonymous surveys in Canvas, but if you have pressing or specific issues, please do not hesitate to let me know if any aspect of our course or class community can be improved.

Schedule

Tentative. Subject to modification as the semester progresses.

WEEK	ΤΟΡΙϹ
1 (9/3,9/7)	 Course Introduction syllabus, self-introduction, course overview
	Introduction and overview of data analysis
9/9	Last day to add/drop a class
2 (9/10,9/13)	 Data basics types of data, level of measurement, summary/contingency tables, stem-and-leaf plots, frequency distribution tables, etc.
	Descriptive statistics
3 (9/17,9/20)	 Data visualization: basic charts Amount, distribution, proportion, association

	Lab: Excel basics Function, formula, absolute/relative reference, basic charts, pivot
	table
4 (9/24,9/27)	Inferential statistics
	especially hypothesis testing
	Lab: Tableau
5 (10/1,10/4)	Data management (database, data warehouse, OLTP, OLAP, etc)
	Lab: Tableau
6 (10/8,10/11)	(buffer)
7 (10/15,10/18)	Microsoft HR Visualization Challenge
8 (10/22,10/25)	Project 1 Presentation
(· · · · _ · · · _ · · · _ ·)	
9 (10/29,11/1)	Data Mining - Introduction
	Lab: Python basics with NumPy, Pandas, and Matplotlib
10 (11/5, 11/8)	Data Mining - Classification
	Lab: Classification
11/11	Last day to withdraw from classes
11 (11/12, 11/15)	Data Mining - Clustering
	Lab: Clustering
12 (11/19, 11/22)	Data Mining - Association rules (also cover probability theory here)
	Lab: Association Rules
13 (11/27)	Final project check-in
11/28-12/1	Thanksgiving recess
14 (11/3, 11/6)	Course review and preparation for the final exam
	Final project presentation
15 (11/10)	Final project presentation (continued)
15 (11/15-21)	Final exam