INFO659: INTRODUCTION TO DATA ANALYTICS

Fall 2024 Course Syllabus

AT A GLANCE

Course

Course Number	INFO 659
Course Title	Introduction to Data Analytics
Credits	3
Schedule	Monday 6:00-8:50 PM (in-person)
	Asynchronous (online)
Web site	Blackboard
Discussion	Blackboard Discussion Boards
Prerequisites	None

Instructor

Name	Michael Ekstrand
Email	mdekstrand@drexel.edu
Web site	https://md.ekstrandom.net
Office hours	by appointment (see link in Blackboard)

COURSE INFORMATION

Course Description

Provides an overview of data analytics foundations and techniques for information professionals. Introduces fundamental concepts and theories of data and data science. Discusses methods and techniques of data representation, data analysis, and data visualization. Covers a range of tools and systems that can be used to discover insight and derive values from data.

Course Purpose within a Program of Study

Currently, this course is counted as a free elective for CCI Master programs in Information Systems, Health Informatics, and Library and Information Science.

Course Learning Outcomes

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

- Understand the background of data analytics and Big Data in the context of current and future needs of information institutions.

- Evaluate the scope and functionality of several state-of-the-art data analytics tools.
- Apply selected data analytics and visualization tools and methods to explore and gain a broad understanding of a large dataset.
- Report and communicate data analytics results to a large audience and decision makers.

COURSE MATERIALS

Required

Brian Godsey. (2017). *Think Like a Data Scientist*. Manning. [Available free from the library](#).

Recommended

Hadley Wickham, Mine Çetinkaya-Rundel, and Garrett Grolemund. (2015). *R for Data Science*, 2nd Edition. Available free at https://r4ds.hadley.nz/.

Online Readings

I will also be posting online articles on Blackboard throughout the quarter to provide information on specific topics or to supplement your learning.

ASSIGNMENTS AND ASSESSMENT

Grade Components

40%	**Assignments**	4 assignments focused on understanding domain data, data transformation, methodology and modeling, evaluation, and final analysis/report. Emphasis on practice and application of theories, with R programming.
20%	**Final exam**	Focused on theoretical and conceptual understanding.
30%	**Final project**	A group project on data analytics in a domain of your interest.
10%	**Class participation**	Online participation, posting relevant questions and responses, exchanges of ideas, with instructor and other students.

Grade Scale

The following scale will be used to convert points to letter grades:

Points	Grade	Points	Grade	Points	Grade
97-100	A+	82-86.99	B	70-71.99	C-
92-96.99	A	80-81.99	B-	67-69.99	D+
90-91.99	A-	77-79.99	C+	60-66.99	D
87-89.99	B+	72-76.99	C	0-59.99	F

Assignments
- Assignment 1: Loading and examining data
- Assignment 2: Data integration, visualization, and correlation
- Assignment 3: Building and evaluating models
- Assignment 4: Data analysis and visualization
- Final group project: Domain data analysis & presentation
- Final Exam

COURSE POLICIES

Announcements
I will be using Blackboard for course announcements, including any changes to the course. You are responsible for making sure that you receive course announcements in a timely fashion.

Course Communications and Support
We will be using Blackboard discussions for course discussions and support. For any course questions that do not involve your grades or direct information about your solutions to the assignment, I encourage you to ask questions in class or in the Blackboard discussion channels so other students can see the answer.

I am also available by appointment for one-on-one or small group meetings, either in-person or via Zoom. Blackboard has a link to schedule meetings with me.

I would prefer to receive questions about the course over Blackboard instead of e-mail to keep things in one place.

Course Conduct

I expect you to behave in a civil, respectful manner in all class interactions, both in official meetings such as lectures and out-of-classroom activities such as project group meetings and study sessions, and to contribute to a constructive learning environment.

The [Recurse Center Social Rules](#) set a good standard constructive and educational learning environments. The [Drexel Student Code of Conduct](#) provides additional information about expectations for all Drexel students.

Academic Integrity and External Resources

As an information professional, data analyst, and student, you are expected to do your own work, attribute sources, and respect the legal and moral rights of others with respect to their work; as a student, you are also required to abide by the university policies regarding academic integrity. While I aim to allow you to make reasonable use of resources, cheating (including copying code, using unauthorized resources during tests, etc.) is not acceptable. If you are found to be cheating, the penalty may range from an F on the assignment to an F on the course, and will also be reported to the university.

You may consult external resources such as other books and web sites for understanding how to solve assignments or portions of the project. In your assignment submissions, list all external resources you used; if they are available online, provide the URL. You do not need to cite the textbooks or the official documentation for the software we are using.

Besides the course forum, you may ask questions related to the course material and concepts required to complete the work on publicly accessible discussion forums such as Stack Overflow, newsgroups, or publicly-archived mailing lists. To qualify as publicly-accessible, a site must provide access to complete discussions without requiring payment or registration. Provide URLs to the forum discussion on the relevant web site or archive (Google Groups works well for newsgroup archives) with your assignment submission. When you ask a question for one of the assignments, mention that it is for a course project and that your instructor permits you to make reasonable use of discussion forums.

A good question will ask about how to go about a particular sub-portion of the problem, how something works, why something you are trying doesn't work, or other specific difficulties. Do not ask "how do I solve <the problem description>?", or similarly direct translations of the project requirements, or for specific code.

Questions should be written to fill in a gap in your understanding that will then enable you to continue your work, not to get a solution to the assignment.

This course is also subject to the [Drexel Academic Integrity, Plagiarism, Dishonesty, and Cheating Policy](#).

Accessibility and Disability Accommodations
If you need accommodations to be able to fully participate in this course, please talk with me as soon as possible. If you have documentation from Disability Resources for specific accommodations, please bring it, but I am happy to discuss anything you need to fully participate in the class.

See also [Drexel Disability Resources](#) for support available from the university.

Adds, Drops and Withdrawls
- [Course Add/Drop Policy](#)
- [Course Withdrawal Policy](#)

Course Changes
I may need to make changes to the course as the term progresses to better support your learning and the logistics of delivering the course. Such changes will be announced through Blackboard Announcements, as well as mentioned in lecture when the timing of the change permits.

COURSE SCHEDULE

The tentative course schedule is as follows (date column is lecture day, Tuesday):

Week	Date	Topic	Due
1	9/23	Introduction, Context, and Data Problems	
2	9/30	Sourcing, Loading, and Processing Data	A0
3	10/7	Exploratory Analytics with R	
4	10/14	*Mon. University Holiday*	A1
5	10/21	Visualization	
6	10/28	Probability and Statistics	A2
7	11/4	Modeling and Prediction	
8	11/11	Evaluation and Decision-Making	A3
9	11/18	Text and Clustering	
10	11/25	Data Management, Storage, & Curation	
11	12/2	Becoming a Data Scientist	A4, FP
F	12/9	*Final Exam*	

Assignments are due **11:59 PM on Sunday** at the end of the week in which they are listed.