

 <p>2023-2024</p>	<p><b>DATA105N</b>  <b>Data Mining</b>  <b>Section ES</b></p>	<p><b>Fall 2023</b></p>
<p><b>Department</b></p>	<p>Department of Mathematics &amp; Science</p>	
<p><b>Instructor</b></p>	<p>Dr. K Seefeld, EdD</p>	
<p><b>NCC Email</b></p>	<p>kseefeld@ccsnh.edu</p>	
<p><b>Telephone Number</b></p>	<p>(603) 578-6815 ext. 1554</p>	
<p><b>Office Hours</b></p>	<p>Thurs 4-6 on tutoring or by appointment</p>	
<p><b>Office Location</b></p>	<p>167 Streeter Hall</p>	
<p><b>Class Days/Meeting Time</b></p>	<p>Online, asynchronous with a weekly deadline structure</p>	
<p><b>Class Location</b></p>	<p>Canvas</p>	

---

### Rationale

This course teaches how to find patterns in data which is the essence of data mining. Topics covered include association analysis, cluster analysis, classification and prediction and how to work with large datasets.

---

### Course Description

Students will learn how to consolidate data from multiple sources, mine relevant information from source data, and display and summarize information effectively. An important component of this course is the completion of an applied project utilizing a current business intelligent tool software such as Microsoft PowerBI. Prerequisite: DATA101N.

---

### Course Competencies

1. Describe data mining in the context of big data.
2. Describe and apply the data mining methodology.
3. Identify inputs to the data mining process.
4. Produce outputs from the data mining process.
5. Describe and apply basic algorithms.
6. Describe and apply advanced algorithms.
7. Analyze real-world problems based on data mining techniques.
8. Explain solutions.

### Essential Questions

1. Why is it important to understand the origin of data being mined?
2. How can programming languages help with the mining of data?
3. Why is the data mining process important in solving real life problems?
4. Why are ethical considerations important when mining data

## Required Materials

### Textbook(s):

There is no required text for this course.

### Software (no cost to students, course videos will explain how to obtain/install)

Python (will install as part of Orange, free)

Orange data mining tool (free) which uses Python via visual programming (no coding necessary)

### Course Structure and Expectations:

This course is 100% online with no live meetings. It requires students to have the ability to work independently. There are however course activities designed to foster a sense of community in the class. Although we will not be together at the same time you will interact with each other and get to know your classmates during this course.

**Canvas is the home base for the class – all work and course activities occur on Canvas.** This course will follow a structured weekly pattern where weeks (except for week 1) begin on Friday 8 am and ends on the following Friday at 4 pm (there is an overlap on Fridays, and I am generally more available to assist online students on Fridays). All work for that week must be done when expected – this is not a self-paced course. Many of the course components build upon each other and several involve timely interaction with classmates.

Every week I will post a video for you as a weekly overview and lecture. This will cover the material the week's work and discussing any issues or special things for that week. This will be approximately one hour long and will be the first thing you will do each week.

Weekly activities will always involve some analytical lab activities (for which I will post guided activities and videos demonstrating), quizzes on what is covered in videos, and associated homework (for which you will do on your own and submit) using Orange software. Generally, there will be a discussion board to post on as well. It is very important you participate in discussions in this online class to foster community and create a learning environment.

This course will culminate in students doing a final applied project incorporating data mining technologies on a large dataset of their choice. Students will present (via video) this analysis the final week of the course and it serves as the final exam for the course.

### Grading:

Final project	25%
Labs/HW	30%
Quizzes	15%
Discussions	30%

The time to negotiate a disappointing grade is immediately after the grade is awarded (within 72 hours of posting the grade) and not the end of the semester. I am generally very good about promptly posting grades.

Final course grades are assigned based on the standard scale below.

		<b>B+</b>	87-89	<b>C+</b>	77-79	<b>D+</b>	67-69		
<b>A</b>	93-100	<b>B</b>	83-86	<b>C</b>	73-76	<b>D</b>	63-66	<b>F</b>	< 60
<b>A-</b>	90-92	<b>B-</b>	80-82	<b>C-</b>	70-72	<b>D-</b>	60-62		

## Available Support Services

### NCC Library:

Get help with research for your papers and assignments from the NCC librarian and staff. Email [ncclibrary@ccsnh.edu](mailto:ncclibrary@ccsnh.edu)

### Tutoring and Writing Center:

The Tutoring and Writing Center, located in the library, offers free academic support services to all NCC students. <https://library.nashuacc.edu/home/tutoring>

### Accessibility:

Students who think they may have or have a documented disability (physical, learning, or mental health) and/or had an IEP or 504 Plan in high school who may need classroom accommodations must meet with the Accessibility Services Coordinator to set up an NCC Reasonable Accommodation Plan (RAP). For more information view the Accessibility page: <https://www.nashuacc.edu/student-services/disability-accessibility-information> and/or contact Jodi Quinn, Accessibility Services Coordinator (office located in the library), at [jquinn@ccnsh.edu](mailto:jquinn@ccnsh.edu) or (603) 578-8996.

### CARE Team:

Students in need of assistance in the areas of mental health or basic support services (such as food, clothing, shelter, substance abuse or others) should contact the [NCC CARES Team](#).

### Student Assistance Program:

Students in need of confidential counseling services should visit the [NCC Student Assistance Program](#) page on the NCC website

### English for Speakers of Other Languages:

The ESOL Program provides a variety of services for non-native speakers of English. <https://nashuacc.edu/programs/esol-esl/>

## Department Policies

The data analytics program at NCC is part of the NCC Mathematics and Science department.

## College Policies

### Non-Discrimination:

The NCC Equity Committee is designated to coordinate compliance with the Non-Discrimination Policy and handles all concerns of discrimination that occur on campus that are not covered under Title IX. View more information here [Non-Discrimination Policy and Grievance Procedure](#)

### Title IX:

Classroom instructors at Nashua Community College are encouraged to create and maintain a safe learning environment in which students feel able to share opinions and related life experiences in classroom discussions, in written work, and in meetings with professors. To the greatest extent possible, this information will be kept private. However, staff and faculty have a legal obligation to report information concerning sexual misconduct, violence and exploitation of individuals per federal statute and in compliance with established policies and procedures at Nashua Community College. If you have been subjected to sexual misconduct, violence or exploitation, we encourage you to contact your NCC Title IX Coordinator, Vice President of Student and Community Affairs, Lizbeth Gonzalez, for support and assistance. She can be reached at Nashua Community College, 505 Amherst Street, Nashua, NH, 603-578-8928 or [lgonzalez@ccsnh.edu](mailto:lgonzalez@ccsnh.edu). You may also contact Bridges, the local crisis center, for free and confidential services at their 24hr support line 603-883-3044. You do not need to be in crisis to call. Nashua Community College encourages you to download uSafeUS® <https://usafeus.org/>, a free and confidential app that helps protect yourself and your community. Additional information including resources can be found here [Title IX Policy and Grievance Procedure](#).

### Attendance:

Regular attendance and active participation are essential for academic success and NCC takes student attendance very seriously.

Every student should carefully review the [NCC attendance policies](#).

### **NCC Academic Failure Policy**

Students with excessive absences or who display disruptive classroom behavior may be subject to the college's Academic Failure policy. All students should carefully review the policy found here: [AF Policy | Nashua Community College](#)

### **Classroom Audio Recording Policy**

[Classroom Audio Recording Policy | Nashua Community College](#)

### **Canvas:**

Canvas is the online learning management system used by instructors and learners at Nashua Community College. Information regarding the specific uses of Canvas and Canvas help for students can be found here: [Canvas | Nashua Community College](#)

### **Forgotten Login Information:**

If you have forgotten your Easy Login Password information, please visit the following link: <https://password.ccsnh.edu/accounts/Reset>

### **EAB Alerts:**

Faculty can issue "Alerts" for students at any time. The administration at NCC encourages professors to issue alerts for any student who may be having difficulty in the course. When a faculty member raises an alert, both the student and student's advisor receive an automated email based on the alert raised. Faculty can issue an alert through EAB Navigate on the college website or directly using the following link: <https://nashuacc.campus.eab.com>

### **College Email System:**

Nashua Community College utilizes a college email system as a means of the College sending official information to enrolled students, and for students to send communication to their instructors and College personnel. Specific details regarding the use of the college's email system can be found here [College Email System | Nashua Community College](#)

### **Sensitive Materials Policy:**

During the semester, in order to cover certain academic topics, there may be occasions to view or discuss material which may not meet the student's own personal definition of appropriateness. In those cases, students should refer to the [NCC Sensitive Materials Policy](#)

### **Academic Conduct:**

All students are expected maintain high standards of academic integrity as they complete their coursework at NCC. Cheating and plagiarism are violations of the Student

Code of Conduct and all instances of academic dishonesty are taken very seriously by the College.

[Cheating](#)

[NCC Plagiarism Policy](#)

**Diversity, Equity, and Inclusion Statement:**

[Diversity, Equity, and Inclusion Statement | Nashua Community College](#)

**NCC Credit Hour Guidelines:**

[Credit Hour Guidelines | Nashua Community College](#)

## Course Calendar

Date start and end	Week	Theme	Coverage	Targeted Competencies
10/23 to 10/27	1	<b>Intro</b>	Overview of Data Mining  Definitions History <b>Setup Orange Software</b>	<i>1. Describe data mining in the context of big data. 2. Describe and apply the data mining methodology.</i>
10/27 to 11/3	2	<b>Data exploration</b>	Uploading data EDA and basic statistics Issues with large datasets <b>Applied exercises - EDA</b>	<i>3. Identify inputs to the data mining process.</i>
11/3 to 11/10	3	<b>Preparing for mining</b>	Date wrangling Dimension reduction Normalization <b>Applied exercises - data prep</b>	<i>4. Produce outputs from the data mining process.</i>
11/10 to 11/17	4	<b>Cluster Analysis</b>	Intro to clustering K Means clustering Hierarchical clustering <b>Applied Exercises - clustering</b>	<i>5. Describe and apply basic algorithms.</i>
11/17 to 11/24	5	<b>Classification part 1</b>	Intro to classification Linear and logistic regression <b>Applied Exercises - regression</b>	<i>6. Describe and apply advanced algorithms.</i>
11/24 to 12/1	6	<b>Classification part 2</b>	Decision Trees K Nearest neighbors <b>Applied Exercises - classification</b>	

12/1 to 12/8	7	<b>Market basket analysis</b>	Probability background Definition of Association Association metrics <b>Applied Exercises - association</b>
12/8 to 12/16	8	<b>End of course</b>	Project due and presentation <i>7. Analyze real-world problems based on data mining techniques.</i> <i>8. Explain solutions.</i>

**This schedule is subject to change**