**Data Mining and Data Warehousing**

**MASY1-GC 3510** | **100** | **Summer 2024 |** **07/03/2024 – 08/12/2024 | 3 Credit**

**Modality: In-person**

**Course Site URL:** <https://brightspace.nyu.edu/>

**General Course Information**

**Name/Title:** Kathleen Hayes, Adjunct Instructor, She/Her/Hers

**NYU Email:** kmh9021@nyu.edu

**Class Meeting Schedule:** 07/03/2024 – 08/12/2024 | Mondays & Wednesdays | 06:20pm – 08:55pm

**Class Location:** TBD

**Office Hours:** By appointment; email [kmh9021@nyu.edu](mailto:kmh9021@nyu.edu) to schedule an NYU Zoom office hour appointment

**Description**

In an increasingly competitive information age, data mining and data warehousing are essential in business decision-making. This course teaches students concepts, methods, and skills for working with data warehouses and mining data from these warehouses to optimize competitive business strategy. In this course, students develop analytical thinking skills required to identify effective data warehousing strategies such as when to use outsource or in-source data services. Students also learn to Extract, Transform and Load data into data warehouses (the ETL process) and use the CRISP approach to data mining to extract vital information for data warehouses. The course also teaches students how to secure data and covers the ethical issues associated with the uses of data and data models for business decisions.

**Prerequisites**

1210 - Quantitative Models for Decision Makers

**Learning Outcomes**

At the conclusion of this course, students will be able to:

* Translate business requirements into a well-constructed, normalized conceptual and logical data models
* Apply logical database design and the relational model
* Apply the CRISP model to conduct successful data mining
* Establish a successful ETL process to load a data warehouse
* Write basic SQL statements including some advanced SQL features
* Employ appropriate data governance principles to assure data quality and security

**Communication Methods**

Be sure to turn on your [NYU Brightspace notifications](https://www.nyu.edu/servicelink/KB0018507) and frequently check the “Announcements” section of the course site. This will be the primary method I use to communicate information critical to your success in the course. To contact me, send me an email. I will respond within 24 hours during the week and no later than 48 hours over weekends.

**Structure | Method | Modality**

There are 12 class sessions, including time for your exams and final presentations.

Active learning experiences and small group projects are key components of the course. Assignments, papers, and exams will be based on course materials (e.g., readings, videos), lectures, and class discussions. Some course sessions may be conducted synchronously on NYU Zoom, which you can access from the course site in [NYU Brightspace](https://brightspace.nyu.edu/).

This course is in-person and will meet once a week on Wednesday, with assignments, announcements and emails being sent through Brightspace. Students are expected to check email and/or Brightspace at least twice a week for announcements related to assignments, class changes or adjustments, and other important information.

**Expectations**

Learning Environment

You play an important role in creating and sustaining an intellectually rigorous and inclusive classroom culture. Respectful engagement, diverse thinking, and our lived experiences are central to this course, and enrich our learning community.

Participation

You are integral to the learning experience in this class. Be prepared to actively contribute to class activities, group discussions, and work outside of class.

Assignments and Deadlines

Students are required to complete (on their own) homework assignments following various materials used for class; applied coding and/or case problems that may be periodically assigned. Students are also required to independently complete midterm/final exams.

I expect you to complete all assignments and to come to each class fully prepared, having read the assigned readings and completed the assignment(s), as well as actively engaging in class discussions.

The details for all assignments and projects will be available on NYU Brightspace, including assignment description and grading criteria. Assignments due dates are final and not negotiable, and late submissions will be not accepted without written permission from the faculty and cannot be granted more than once throughout the semester. Failure to submit assignments on time will result in grade reductions up to 100%.

All exams will be open book, open notes exams. The exams will test students’ acquisition of topics, concepts, and competencies learned in class. The midterm will assess content learned up to mid-term, and the final exam will only cover material covered in the second half of the term.

Please submit all assignments to the appropriate section of the course site in [NYU Brightspace](https://brightspace.nyu.edu/). Assignments will be graded for content and clarity.

Course Technology Use

We will utilize multiple technologies to achieve the course goals. I expect you to use technology in ways that enhance the learning environment for all students. Class sessions may require use of Zoom, and all class sessions will require use of technology (e.g. laptop) for learning purposes.

Feedback and Viewing Grades

I will provide timely, meaningful feedback on all your work via our course site in NYU Brightspace. You can access your grades on the course site Gradebook.

Attendance

I expect you to attend all class sessions. Attendance will be taken into consideration when determining your final grade.

Excused absences are granted in cases of documented serious illness, family emergency, religious observance, or civic obligation. In the case of religious observance or civic obligation, this should be reported in advance. Unexcused absences from sessions may have a negative impact on a student’s final grade. Students are still responsible for assignments given during any absence.

Each unexcused absence or being late may result in a student’s grade being lowered by a fraction of a grade. A student with three unexcused absences may earn a Fail grade.

Students who join the course during the add/drop are responsible for ensuring that they identify what assignments and preparatory work they have missed and complete and submit those per the syllabus.

Refer to the [SPS Policies and Procedures page](https://www.sps.nyu.edu/homepage/student-experience/policies-and-procedures.html) for additional information about attendance.

**Textbooks and Course Materials**

*Students can purchase these items through the NYU Bookstore.*

**Required:**

Data Mining: Concepts, Models, Methods, and Algorithms, 3rd Edition

* Author: Mehmed Kantardzic
* Publisher: Wiley-IEEE Press, 2019
* ISBN: 978-1-119-51607-1

Cloud computing / software

* Be prepared to create accounts for Mode Analytics, Google BigQuery, StrataScratch, Tableau. Where possible, we will use free / student versions of software offerings, but there may be costs associated with usage throughout the semester.
* You do not need to register for any of these ahead of time; we will do this together as a class.

**Recommended:**

* Instructor may also provide session-by-session content, which will be posted online.

**Grading | Assessment**

Your grade in this course is based on your performance on multiple activities and assignments. Since all graded assignments are related directly to course objectives and learning outcomes, failure to complete any assignment will result in an unsatisfactory course grade. All written assignments must be typed and 1.5x- or double-spaced. Grammar, punctuation, and spelling will be considered in grading. Please carefully proof-read your written assignments before submitting them for a grade. Students can have multiple submissions toward the same assignment unless otherwise noted; the latest submission will be considered for the grading. I will update the grades on the course site each time a grading session has been completed.

DESCRIPTION PERCENTAGE

Class participation 12%

Homework assignments 24%

Final project 24%

Midterm Exam 20%

Final Exam 20%

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

TOTAL 100%

See the [“Grades” section of Academic Policies](https://www.sps.nyu.edu/homepage/student-experience/policies-and-procedures.html#Graduate1) for the complete grading policy, including the letter grade conversion, and the criteria for a grade of incomplete, taking a course on a pass/fail basis, and withdrawing from a course.

**Course Outline**

**Start/End Dates:** 07/03/2024 - 08/12/2024 | Mondays &amp; Wednesdays

**Time:** 06:20 pm – 09:20 pm

**Summer Session**: 6W2

**No Class Date(s):** N/A

**Special Notes**: N/A

**Session 1** - 07/03/24

Topics: Introduction to Data Warehousing, SQL, and Tools

Assignments (due one week from class):

 HW #1

 Project milestone: Company selection

**Sessions 2 and 3** – 07/08/24 & 07/10/24

Topics: ETL/ELT, BRDs, SQL techniques

Assignments (due one week from class):

 HW #2 and HW #3

 Project milestone: Project Proposal

**Sessions 4 and 5** – 07/15/24 & 07/17/24

Topics: Data Modeling, Schemas, and Data Governance

Assignments (due one week from class):

 HW #4 and HW #5

 Project milestone: Dimension and Fact Tables

**Sessions 6 and 7** – 07/22/24 & 07/24/24

Topic: Data Visualization

Assignments (due one week from class):

 HW #6

 Reading: Chapters 3-5 (Data Mining: Concepts, Models, Methods, and Algorithms)

**Sessions 8 and 9** – 07/29/24 & 07/31/24

Topic: Introduction to Data Mining &amp; Data Science Techniques

Assignments (due one week from class):

 HW #7 and HW #8

 Reading: Chapters 6, 9 (Data Mining: Concepts, Models, Methods, and Algorithms)

**Sessions 10 and 11** – 08/05/24 & 08/07/24

Topic: Data Science Techniques - Supervised and Unsupervised Learning

Assignments (due one week from class):

 Project milestone: Final project presentation (due 8/11/24)

**Session 12** – 08/12/24

Topic: Final presentations

Assignments (due one week from class):

 N/A

**NOTES:**

The syllabus may be modified to better meet the needs of students and to achieve the learning outcomes.

The School of Professional Studies (SPS) and its faculty celebrate and are committed to inclusion, diversity, belonging, equity, and accessibility (IDBEA), and seek to embody the IDBEA values. The School of Professional Studies (SPS), its faculty, staff, and students are committed to creating a mutually respectful and safe environment (*from the* [*SPS IDBEA Committee*](https://www.sps.nyu.edu/homepage/about-us/idbea/about-idbea.html)).

**New York University School of Professional Studies Policies**

1. Policies - You are responsible for reading, understanding, and complying with [University Policies and Guidelines](http://www.nyu.edu/about/policies-guidelines-compliance.html), [NYU SPS Policies and Procedures](http://sps.nyu.edu/academics/academic-policies-and-procedures.html), and [Student Affairs and Reporting](https://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/student-services.html).

2. Learning/Academic Accommodations - New York University is committed to providing equal educational opportunity and participation for students who disclose their dis/ability to the [Moses Center for Student Accessibility](https://www.nyu.edu/students/communities-and-groups/student-accessibility.html). If you are interested in applying for academic accommodations, contact the [Moses Center](https://www.nyu.edu/students/communities-and-groups/student-accessibility/academic.html) as early as possible in the semester. If you already receive accommodations through the Moses Center, request your accommodation letters through the [Moses Center Portal](https://www.nyu.edu/students/communities-and-groups/student-accessibility.html) as soon as possible ([mosescsa@nyu.edu](mailto:mosescsa@nyu.edu) | 212-998-4980).

3. Health and Wellness - To access the University's extensive health and mental health resources, contact the [NYU Wellness Exchange](https://www.nyu.edu/students/health-and-wellness/wellness-exchange.html). You can call its private hotline (212-443-9999), available 24 hours a day, seven days a week, to reach out to a professional who can help to address day-to-day challenges as well as other health-related concerns.

4. Student Support Resources - There are a range of resources at SPS and NYU to support your learning and professional growth. For a complete list of resources and services available to SPS students, visit the [NYU SPS Office of Student Affairs site](https://www.sps.nyu.edu/homepage/student-experience/resources-and-services.html).

5. Religious Observance - As a nonsectarian, inclusive institution, NYU policy permits members of any religious group to absent themselves from classes without penalty when required for compliance with their religious obligations. Refer to the [University Calendar Policy on Religious Holidays](https://www.nyu.edu/about/policies-guidelines-compliance/policies-and-guidelines/university-calendar-policy-on-religious-holidays.html) for the complete policy.

6. Academic Integrity and Plagiarism - You are expected to be honest and ethical in all academic work. Moreover, you are expected to demonstrate how what you have learned incorporates an understanding of the research and expertise of scholars and other appropriate experts; and thus recognizing others' published work or teachings—whether that of authors, lecturers, or one's peers—is a required practice in all academic projects.

Plagiarism involves borrowing or using information from other sources without proper and full credit. You are subject to disciplinary actions for the following offenses which include but are not limited to cheating, plagiarism, forgery or unauthorized use of documents, and false form of identification

[Turnitin](https://www.nyu.edu/servicelink/KB0018471), an originality detection service in NYU Brightspace, may be used in this course to check your work for plagiarism.

Read more about academic integrity policies at the NYU School of Professional Studies on the [Academic Policies for NYU SPS Students](https://www.sps.nyu.edu/homepage/student-experience/policies-and-procedures.html) page.

7. Use of Third-Party Tools - During this class, you may be required to use non-NYU apps/platforms/software as a part of course studies, and thus, will be required to agree to the “Terms of Use” (TOU) associated with such apps/platforms/software.

These services may require you to create an account but you can use a pseudonym (which may not identify you to the public community, but which may still identify you by IP address to the company and companies with whom it shares data).

You should carefully read those terms of use regarding the impact on your privacy rights and intellectual property rights. If you have any questions regarding those terms of use or the impact on the class, you are encouraged to ask the instructor prior to the add/drop deadline.