

Syllabus

This is a single, concatenated file, suitable for printing or saving as a PDF for offline viewing. Please note that some animations or images may not work.

Description

This [module](#) is also available as a concatenated page, suitable for printing or saving as a PDF for offline viewing.

MET CS 689

Designing and Implementing a Data Warehouse

This course provides students with the engineering skills required to evaluate, implement, and scale data warehousing, with a focus on dimensional modeling, ETL/ELT (Extract, Transform, Load) processes, and data visualizations.

Students will learn how to design and implement a data warehouse using dimensional modeling techniques, which involve organizing data into "facts" and "dimensions" to create a structure that is optimized for querying and analysis. Students will deep dive into ETL/ELT processes to extract data from various sources via Python and SQL, transform it into a consistent format, and load it into a data warehouse, including how to handle data changes from deltas by loading data into slowly changing dimensions.

In addition, the course will cover data visualization techniques to help students present data in a clear and meaningful way. Students will learn how to use tools such as Tableau or Power BI to create visualizations that can be used to explore and analyze data from a data warehouse. The course will also explore cloud based scaled out data warehousing via the Spark architecture focusing on semi-structured data.

Overall, this course will provide students with a foundational understanding of data warehousing concepts and skills that are essential for careers in data analysis, business intelligence, and other data-related fields.

Technical Notes

The table of contents expands and contracts (+/- sign) and may conceal some pages. To avoid missing content pages, you are advised to use the next/previous page icons in the top right corner of

the learning modules.

This course requires you to access files such as word documents, PDFs, and/or media files. These files may open in your browser or be downloaded as files, depending on the settings of your browser.

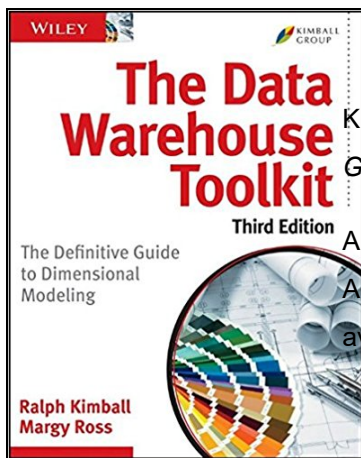
Learning Objectives

By successfully completing this course you will be able to:

- Understand the purpose of data warehousing
- Apply standard SQL Analytical functions for business intelligence reporting
- Perform dimensional data modeling
- Perform Extract, Transform, and Load (ETL and ELT) for varied data using Python and SQL techniques
- Understand, implement, and load Slowly Changing Dimensions (SCD) and Fact tables of various grains accounting for data changes (deltas).
- Understand and Implement business intelligence reporting via visualizations (dashboards)
- Introduction to working with semi-structured (Big Data) datasets and analysis with cloud-based data warehousing (Spark)
- Understand and apply performance tuning focusing specifically on the ETL and ELT process

Materials

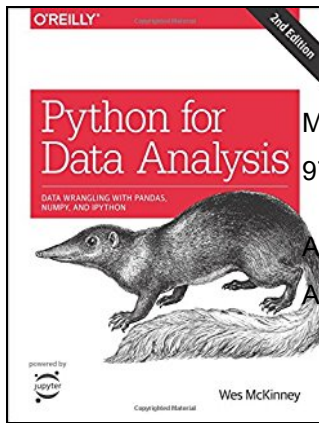
Recommended Book



Kimball, R. & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3rd ed.). Wiley & Sons. ISBN 9781118530801.

An e-book is available at [Vitalsource.com](https://vitalsource.com). An e-book is available through Amazon. An e-book is available for free through the BU library. An e-book is available through the [BU bookstore](#).

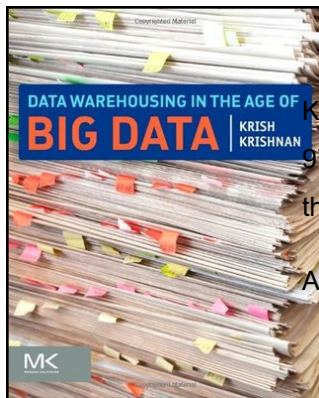
Optional Books



McKinney, W. (2017). *Python for Data Analysis* (2nd ed.). O'Reilly Media. ISBN 9781491957660.

An e-book is available at [Vitalsource.com](https://vitalsource.com). An e-book is available through Amazon.

An e-book is available for free through the BU library.



Krishnan, K. (2013). *Data Warehousing in the Age of Big Data*. ELSEVIER. ISBN 9780124058910. An e-book is available at [Vitalsource.com](https://vitalsource.com). An e-book is available through Amazon.

An e-book is available for free through the BU library.

Boston University Library Information

Boston University has created a set of videos to help orient you to the online resources at your disposal. An introduction to the series is below:

met_ode_library_14_sp1_00_intro video cannot be displayed here

All of the videos in the series are available on the [Online Library Resources](#) page, which is also accessible from the Campus Bookmarks section of your Online Campus Dashboard. Please feel free to make use of them.

As Boston University students, you have full access to the BU Library. From any computer, you can gain access to anything at the library that is electronically formatted. To connect to the library, use the link <http://www.bu.edu/library>. You may use the library's content whether you are connected through your online course or not, by confirming your status as a BU community member using your Kerberos password.

Once in the library system, you can use the links under "Resources" and "Collections" to find databases, eJournals, and eBooks, as well as search the library by subject. Some other useful links follow:

Go to <http://www.bu.edu/library/research/collections> to access eBooks and eJournals directly.

If you have questions about library resources, go to <http://www.bu.edu/library/help/ask-a-librarian> to email the library or use the live-chat feature.

To locate course eReserves, go to <http://www.bu.edu/library/services/reserves>.

Please note that you are not to post attachments of the required or other readings in the water cooler or other areas of the course, as it is an infringement on copyright laws and department policy. All students have access to the library system and will need to develop research skills that include how to find articles through library systems and databases.

Study Guide

This course meets on Mondays.

CS689 Advanced Database Management A1 Fall 2023 Schedule

	Date	Topic	What's Due

1	09/11/2023	Module 1 Course Intro, Term Project, Data Warehouse Intro, DB recovery overview	
2	09/18/2022	Module 1 Analytical SQL Functions Intro to Python	Crediting Sources Assessment Assignment 1A Installations
3	09/25/2023	Module 1 Analytical SQL Functions - continued Module 2 Dimensional Data Modeling	(Work on Assignment 1B)
4	10/02/2023	Module 2 Dimensional Data Modeling- In class Activity	Assignment 1B – SQL and Python
	10/09/2023	Indigenous People's Day Holiday	
5	10/10/2023 (Tuesday)	Module 2 Dimensional Data Modeling- In class Activity Module 3 ETL Deep dive	Term Project Update #1
6	10/16/2023	Module 3 ETL Deep dive	Assignment 2 – DW design
7	10/23/2022	Module 3 ETL Deep dive	Assignment 3A – ETL Python
8	10/30/2023	Module 3 ETL Deep dive Module 4 Reporting, Analysis and Visualization	Term Project Update #2

9	11/6/2023	Module 4 Reporting, Analysis and Visualization continued	Assignment 3B – ETL SQL
10	11/13/2023	Module 4 Reporting, Analysis and Visualization continued	Assignment 4 – Presentation
11	11/20/2023	Module 5 -Big Data Sign-up for term project presentations (link will become editable at start of class)	Term Project Update #3
12	11/27/2023	Module 6 Architecture and Tuning	Assignment 5 – Big Data Term Project Update #4
13	12/04/2023	Term Project Presentations Additional Online Presentations between 12/6 and 12/11	
14	12/11/2023	Term Project Presentations	Term Project Report & presentation
	12/14/2023	Study Period	
FE	12/18/2022	Final Exam	

Module 1 Study Guide and Deliverables

Readings: • Module 1 online content

Assignments: Assignment 1A

Assignment 1B

Module 2 Study Guide and Deliverables

- Readings:**
- Module 2 online content
 - Kimball/Ross ch 2, 18

Assignments: Assignment 2

Module 3 Study Guide and Deliverables

- Readings:**
- Module 3 online content
 - Kimball/Ross Ch 19, 20

Assignments: Assignment 3A
Assignment 3B

Module 4 Study Guide and Deliverables

- Readings:**
- Module 4 online content

Assignments: Assignment 4

Module 5 Study Guide and Deliverables

- Readings:**
- Module 5 online content

Assignments: Assignment 5

Module 6 Study Guide and Deliverables

- Readings:**
- Module 6 online content

- Assignments:**
- Final Term Project
 - Final Project Presentation

Course Please complete the course
Evaluation: evaluation. Your feedback is important to MET, as it helps us make improvements to the program and the course for future students.

Grading Information

Please check the **Study Guide** in the syllabus for Live Classroom dates and specific due dates for assignments and assessments.

The course is conducted by means of a sequence of readings from the textbooks, lectures in text and graphic form, and exercises. There are one or more lectures per module.

Graded Items:

- **Assignments:** There are assignments due each module. You submit the assignment in the "Assignments" area.
- **Term Project:** There is a term project that you will define and develop with the help of your facilitators and myself. There are weekly milestones to help you stay on schedule and to help your facilitator and professor guide you as you work on your project. You submit each piece of the term project in the "Assignments" area.
- **Final Exam:** There will be a proctored Final Exam in this course using a proctor service called Examity. Detailed instructions regarding your proctored exam will be forthcoming from the Assessment Administrator. You will be responsible for scheduling your own appointment.

Grading Policy

All students will be expected to demonstrate database knowledge and techniques. Your professor may in exceptional circumstances, such as disabilities, modify these distributions to more accurately reflect a student's performance in the course.

Grading Scheme	
Assignments	35%
Term Project Updates	5%

Term Project	30%
Final Exam	30%

Grade	Numeric Grade Range	Grade Points
A	≥ 95	4.0
A-	$\geq 90 < 95$	3.7
B+	$\geq 87 < 90$	3.3
B	$\geq 83 < 87$	3.0
B-	$\geq 80 < 83$	2.7
C+	$\geq 77 < 80$	2.3
C	$\geq 73 < 77$	2.0
C-	$\geq 70 < 73$	1.7
D	$\geq 60 < 70$	1.0
F	< 60	0

The Term Project Update grades are Acceptably on track (1), Not yet acceptably on track (0). Updated summatively account for 5% of grade and are not an average. Otherwise:

≥ 2.5 "Acceptably...": A

≥ 2 "Acceptably...": B

≥ 1 "Acceptably...": C

< 1 "Acceptably...": D

None "meets ...": F

Expectations

You are expected to communicate proactively with your facilitator and team members. Especially for group projects, delays will impact other team members' work, so extra sensitivity towards your virtual teammates is appreciated.

Delays

All assignments must be completed. We understand that it is sometimes not possible for students to submit their assignments by the deadline, and we make every effort to accommodate our typically very busy students' schedules. We recognize that students with crunch times at work or other things that keep them from submitting their work on time are already at a disadvantage, so we allow late submissions without penalty, provided that you communicate your situation to us and your facilitator can grade them on time. In practice, your facilitator will be very busy grading your term projects near the end of the term, so it may not be possible to grade assignments submitted more than one week late near the end of the term; these may count as a zero in your course grade. For this reason, be sure to communicate with your facilitator, preferably in advance, if you need to submit your assignments late.

On-Campus Late Penalty

- One week late - no penalty
- Past one week late (more than 7 calendar days from original due date) 50% credit

Near the end of the term, so it may not be possible to grade assignments submitted more than one week late; these may count as a zero in your course grade. For this reason, be sure to communicate with your facilitator, preferably in advance, if you need to submit your assignments late.

Academic Conduct Policy

Please visit Metropolitan College's website for the full text of the department's [*Academic Conduct Code*](#).

A Definition of Plagiarism

“The academic counterpart of the bank embezzler and of the manufacturer who mislabels products is the plagiarist: the student or scholar who leads readers to believe that what they are reading is the original work of the writer when it is not. If it could be assumed that the distinction between plagiarism and honest use of sources is perfectly clear in everyone’s mind, there would be no need for the explanation that follows; merely the warning with which this definition concludes would be enough. But it is apparent that sometimes people of goodwill draw the suspicion of guilt upon themselves (and, indeed, are guilty) simply because they are not aware of the illegitimacy of certain kinds of “borrowing” and of the procedures for correct identification of materials other than those gained through independent research and reflection.”

“The spectrum is a wide one. At one end there is a word-for-word copying of another’s writing without enclosing the copied passage in quotation marks and identifying it in a footnote, both of which are necessary. (This includes, of course, the copying of all or any part of another student’s paper.) It hardly seems possible that anyone of college age or more could do that without clear intent to deceive. At the other end there is the almost casual slipping in of a particularly apt term which one has come across in reading and which so aptly expresses one’s opinion that one is tempted to make it personal property.”

“Between these poles there are degrees and degrees, but they may be roughly placed in two groups. Close to outright and blatant deceit-but more the result, perhaps, of laziness than of bad intent-is the patching together of random jottings made in the course of reading, generally without careful identification of their source, and then woven into the text, so that the result is a mosaic of other people’s ideas and words, the writer’s sole contribution being the cement to hold the pieces together. Indicative of more effort and, for that reason, somewhat closer to honest, though still dishonest, is the paraphrase, and abbreviated (and often skillfully prepared) restatement of someone else’s analysis or conclusion, without acknowledgment that another person’s text has been the basis for the recapitulation.”

The paragraphs above are from H. Martin and R. Ohmann, *The Logic and Rhetoric of Exposition, Revised Edition*. Copyright 1963, Holt, Rinehart and Winston.

Academic Conduct Code

I. Philosophy of Discipline

The objective of Boston University in enforcing academic rules is to promote a community atmosphere in which learning can best take place. Such an atmosphere can be maintained only so long as every student believes that his or her academic competence is being judged fairly and that he or she will not be put at a disadvantage because of someone else’s dishonesty. Penalties should be carefully determined so as to be no more and no less than required to maintain the desired atmosphere. In defining violations of this code, the intent is to protect the integrity of the educational process.

II. Academic Misconduct

Academic misconduct is conduct by which a student misrepresents his or her academic accomplishments, or impedes other students’ opportunities of being judged fairly for their academic work. Knowingly allowing others to represent your work as their own is as serious an offense as submitting another’s work as your own.

III. Violations of this Code

Violations of this code comprise attempts to be dishonest or deceptive in the performance of academic work in or out of the classroom, alterations of academic records, alterations of official data on paper or

electronic resumes, or unauthorized collaboration with another student or students. Violations include, but are not limited to:

- A. **Cheating on examination.** Any attempt by a student to alter his or her performance on an examination in violation of that examination's stated or commonly understood ground rules.
- B. **Plagiarism.** Representing the work of another as one's own. Plagiarism includes but is not limited to the following: copying the answers of another student on an examination, copying or restating the work or ideas of another person or persons in any oral or written work (printed or electronic) without citing the appropriate source, and collaborating with someone else in an academic endeavor without acknowledging his or her contribution. Plagiarism can consist of acts of commission-appropriating the words or ideas of another-or omission failing to acknowledge/document/credit the source or creator of words or ideas (see below for a detailed definition of plagiarism). It also includes colluding with someone else in an academic endeavor without acknowledging his or her contribution, using audio or video footage that comes from another source (including work done by another student) without permission and acknowledgement of that source.
- C. **Misrepresentation or falsification of data** presented for surveys, experiments, reports, etc., which includes but is not limited to: citing authors that do not exist; citing interviews that never took place, or field work that was not completed.
- D. **Theft of an examination.** Stealing or otherwise discovering and/or making known to others the contents of an examination that has not yet been administered.
- E. **Unauthorized communication during examinations.** Any unauthorized communication may be considered prima facie evidence of cheating.
- F. **Knowingly allowing another student to represent your work as his or her own.** This includes providing a copy of your paper or laboratory report to another student without the explicit permission of the instructor(s).
- G. **Forgery, alteration, or knowing misuse of graded examinations, quizzes, grade lists, or official records of documents,** including but not limited to transcripts from any institution, letters of recommendation, degree certificates, examinations, quizzes, or other work after submission.
- H. **Theft or destruction of examinations or papers** after submission.
- I. **Submitting the same work in more than one course** without the consent of instructors.
- J. **Altering or destroying another student's work or records,** altering records of any kind, removing materials from libraries or offices without consent, or in any way interfering with the work of others so as to impede their academic performance.
- K. **Violation of the rules governing teamwork.** Unless the instructor of a course otherwise specifically provides instructions to the contrary, the following rules apply to teamwork: 1. No team member shall intentionally restrict or inhibit another team member's access to team meetings, team work-in-progress, or other team activities without the express authorization of the instructor. 2. All team members shall be held responsible for the content of all teamwork submitted for evaluation as if each team member had individually submitted the entire work product of their team as their own work.

- L. Failure to sit in a specifically assigned seat during examinations.
- M. Conduct in a professional field assignment that violates the policies and regulations of the host school or agency.
- N. Conduct in violation of public law occurring outside the University that directly affects the academic and professional status of the student, after civil authorities have imposed sanctions.
- O. Attempting improperly to influence the award of any credit, grade, or honor.
- P. Intentionally making false statements to the Academic Conduct Committee or intentionally presenting false information to the Committee.
- Q. Failure to comply with the sanctions imposed under the authority of this code.

Microsoft Azure Dev Tools for Teaching

Microsoft Azure Dev Tools for Teaching is a Microsoft program that supports technical education by providing access to Microsoft software for learning, teaching, and research purposes. Our membership allows faculty and students currently enrolled in MET courses to obtain certain Microsoft products free of charge. All MET students are granted access to download the software for the duration of their study at MET College.

FAQ and basic information are at [Microsoft Azure Dev Tools for Teaching](#), (You may have to enter your personal BU login credentials to access this page.)

Disability and Access Services

In accordance with University policy, every effort will be made to accommodate students with respect to speech, hearing, vision, or other disabilities. Any student who may need an accommodation for a documented disability should contact [Disability and Access Services](#) at 617-353-3658 or at access@bu.edu for review and approval of accommodation requests.

Once a student receives their accommodation letter, they must send it to their instructor and/or facilitator each semester. They must also send a copy to their Faculty & Student Support Administrator, who may need to update the course settings to ensure accommodations are in place. Accommodations cannot be implemented if the student does not send their letter.

Netiquette

BU Virtual has produced a netiquette guide to help you understand the potential impact of your communication style.

Before posting to any discussion forum, sending an email, or participating in any course or public area, please consider the following:

Ask Yourself...

- How would I say this in a face-to-face classroom or if writing for a newspaper, public blog, or wiki?
- How would I feel if I were the reader?
- How might my comment impact others?
- Am I being respectful?
- Is this the appropriate area or forum to post what I have to say?

Writing

When you are writing, please follow these rules:

- **Stay polite and positive in your communications.** You can and should disagree and participate in discussions with vigor; however, when able, be constructive with your comments.
- **Proofread your comments before you post them.** Remember that your comments are permanent.
- **Pay attention to your tone.** Without the benefit of facial expressions and body language, your intended tone or the meaning of the message can be misconstrued.
- **Be thoughtful and remember that classmates' experience levels may vary.** You may want to include background information that is not obvious to all readers.
- **Stay on message.** When adding to existing messages, try to maintain the theme of the comments previously posted. If you want to change the topic, simply start another thread rather than disrupt the current conversation.
- **When appropriate, cite sources.** When referencing the work or opinions of others, make sure to use correct citations.

Reading

When you are reading your peers' communication, consider the following:

- **Respect people's privacy.** Don't assume that information shared with you is public. Your peers may not want personal information shared. Please check with them before sharing their information.
- **Be forgiving of other students' and instructors' mistakes.** There are many reasons for typos and misinterpretations. Be gracious and forgive other's mistakes or point them out privately and politely.
- **If a comment upsets or offends you, reread it and/or take some time before responding.**

Important Note

Don't hesitate to let your instructor or your faculty and student support administrator know if you feel others are inappropriately commenting in any forum.

All Boston University students are required to follow academic and behavioral conduct codes. Failure to comply with these conduct codes may result in disciplinary action.

Registration Information and Important Dates

[View the drop dates for your course.](#)

[Withdraw or drop your course.](#)

- If you are dropping down to zero credits for a semester, please contact your college or academic department.
- **Nonparticipation in your online course does not constitute a withdrawal from the class.**
- If you are unable to drop yourself on MyBU Student Portal, please contact your college or academic department.
- Online courses will open to students in Blackboard on the first day of the term.
- Online courses close to students three weeks after the last day of the term. Please plan to download and save any assignments or material you'd like to keep by that date.

Technical Support

Help Desk

Boston University IT Help Desk can be reached via email (ithelp@bu.edu), phone (617-353-4357) or by filling out the [support form](#) on their website. For IT Help Desk hours of operation, visit the [contact page](#). If you are contacting IT outside of business hours, you will receive a response the following day. Visit the BU Information Services & Technology (IS&T) [news page](#) for announcements and system-wide alerts.

Technology Requirements and Resources

To successfully view all content in your course, it is important that your computer setup meets the necessary minimum technical requirements. Certain courses with specific functionality or educational tools may require additional technical requirements, these details can be found on the Course Resources or Materials page in the Syllabus.

System Requirements

- Access to reliable, high-speed internet: Check your [internet connection speeds](#)
- Learning Management System (Blackboard): [System Requirements](#)
- Synchronous live classroom sessions (Zoom): [System requirements for Windows, macOS, and Linux](#)
- Courses with proctored exams (Examity): [System requirements for Windows, macOS](#)
- Two-factor authentication service for BU applications: [Duo Security](#)

Downloads

- Recommended web browsers: [Mozilla Firefox](#) or [Google Chrome](#)
- Synchronous live classroom sessions (Zoom): [Zoom download center](#)
- Courses with proctored exams (Examity): Desktop or laptop computer with [Google Chrome](#) or [Microsoft Edge](#)
- Two-factor authentication service for BU applications (Duo Security): optional [Duo Mobile download for iOS](#) or [Duo Mobile download for Android](#)

Recommended Hardware

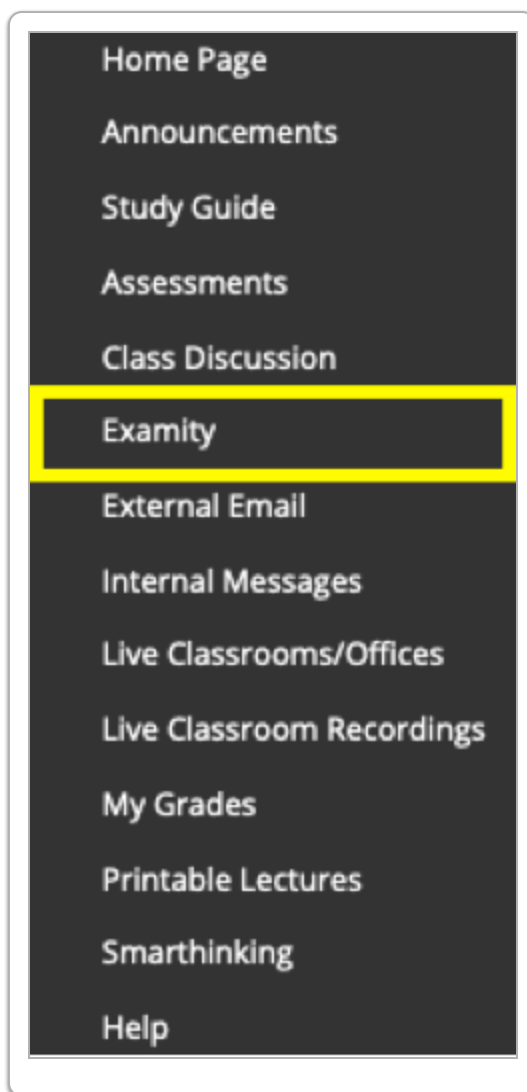
- Desktop or laptop computer recommended for best experience, some course functionality including proctored exams are not compatible with phones or tablets
- Headset with built-in microphone for high quality audio during live classroom sessions
- Webcam (required for proctored exams)
- Working computer speakers (required for proctored exams)

Clearing Your Browser Cache

It is recommended that users periodically [clear their browser cache](#) to ensure they are viewing the most current course content. Completing this step often resolves login issues and problems viewing course materials.

Proctored Exams

Courses with proctored exams will have an Examity link in the left-hand course navigation. This link will not appear until scheduling opens. The BU Virtual Assessment Administrator will notify you when it is time to schedule your exam. Details on Examity's technical requirements and how to schedule your exam are in the Proctored Exam Information module on the course homepage. The Assessment Administrator can be reached at pexams@bu.edu. Examity support is available 24/7 via phone (855-392-6489), email (support@examity.com), or 'live chat' when logged in to the Examity dashboard.



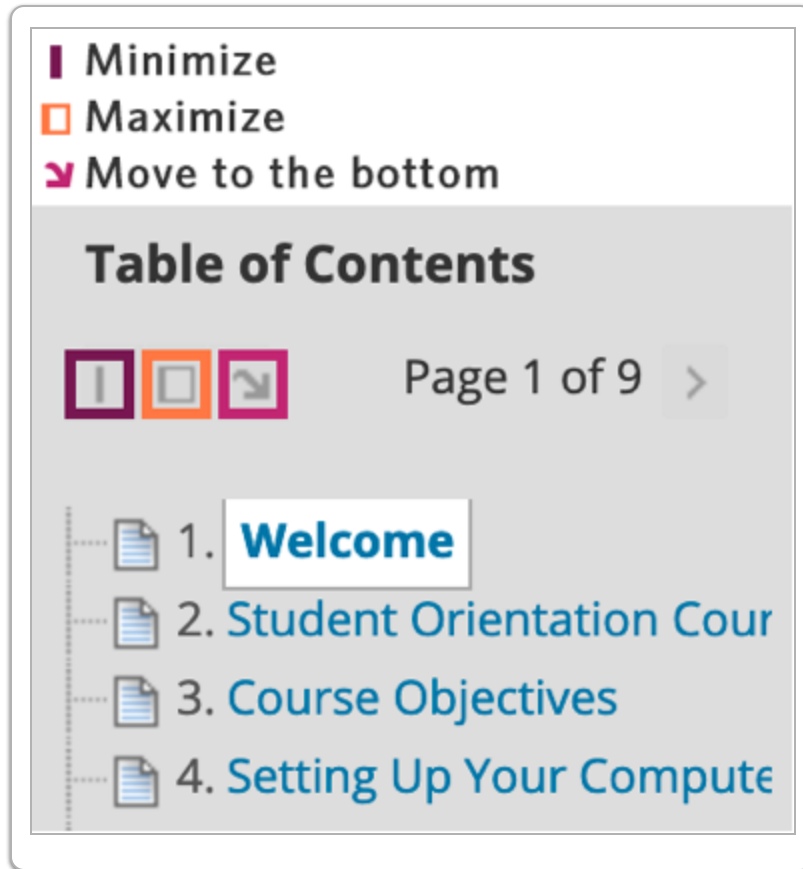
Navigating Courses

While navigating through your courses it's important to note that all hyperlinks will open in a new browser window.


The Blackboard navigation tools—shown in the images below—allow you to show and hide both the Course Menu and the Table of Contents which can free up space when moving through weekly lecture material.

The Table of Contents may contain folders that open and close (+ and – signs) and may conceal some pages. To avoid missing content pages, you are advised to use the next- and previous-page buttons (and icons) in the top-right corner of the learning content.

Navigation tools for the Table of Contents are shown in the image below:



Clicking the space between the Course Menu and the Table of Contents allows you to show or hide the Course Menu on the left:

DE Student Orientation 
(2020 Spring 2)

[Home Page](#)
[Announcements](#)
[Assessments](#)
[Assignments](#)
[Calendar](#)
[Class Discussion](#)
[External Email](#)
[Internal Messages](#)

Table of Contents
Page 1 of 9 >
Hide Course Menu

1. **Welcome**
2. Student Orientation Cou
3. Course Objectives
4. Setting Up Your Comput
5. Blackboard App for Onli
6. Disability Services
7. Netiquette
8. Registration Information
9. Technical Support

Boston University Metropolitan College