

## BUSINESS F743 – Big Data in Finance Fall 2023 Course Outline

Finance and Business Economics Area  
DeGroote School of Business  
McMaster University

### COURSE DESCRIPTION

This course introduces business students to the applications of data science, its key statistical tools, and the underlying technology in the finance area. Students gain a deeper understanding of how finance, technology, and statistics intersect in an applied setting to solve tough problems in innovative ways. The course is intended to equip students with skills in solving problems requiring acquisition, management, and analysis of very large datasets.

### INSTRUCTOR AND CONTACT INFORMATION

Class Timing: **Tuesdays at 7.00 PM**

*All times referenced in this document are Eastern.*

**Adeel Mahmood**

Instructor

Tel: 905-525-9140

*Office Hours:* Before or after class, or by appointment

### COURSE ELEMENTS

Credit Value: 3	Leadership: Yes	IT skills: Yes	Global view: Yes
Avenue: Yes	Ethics: No	Numeracy: Yes	Written skills: Yes
Participation: Yes	Innovation: Yes	Group Work: Yes	Oral skills: Yes

### COURSE PREREQUISITES AND COMPLEMENTS

Students should have the academic credit of F600 or F650 – or equivalent preparation – prior to the start of this course.

## COURSE OVERVIEW

*“The world is swimming in data.”*

According to an IBM estimate, the world is generating more than 2.5 quintillion bytes of data every day, and the pace continues to rise as more technology reaches farther corners of the planet. Financial services data is of particular interest and value because of the high quality quotient of its content. Deeper and more effective analysis of the financial services data can result in improvement in such high-value decisions as economic policy, capital markets investments, credit scores and lending, and financial fraud, among others.

The course is divided into two segments: (1) skills and (2) applications.

The **skills** segment focuses on lectures, lab work, out-of-class research, and self-directed content to deliver the requisite learning. This section is designed to equip a business student with key data science skills. We begin with framing the problem and soon move into acquiring and managing the underlying data. The statistical tools and the technology infrastructure needed to analyse this data are discussed next. The higher cognitive business (esp. finance) insights are then applied to the problem.

The **applications** segment is centred around a real-life big data project that students complete in teams. The project allows students to work on an actual problem being faced by an institution in North America and apply the skills learned in the first segment to help solve the problem. This segment uses group work, instructor-to-group meetings, institution-to-group meetings, continued lab work, and periodic progress reports with feedback to deliver the learning components.

## LEARNING OUTCOMES

Upon completion of the skills segment of this course, students will be able to:

- Identify and frame a big data problem in a finance function or in the financial services industry;
- Acquire, manage, and analyze very large datasets associated with the problem or potential solutions to the problem;
- Understand and use modern applied statistics in analyzing and solving the problem systematically, including classification, clustering, regression, dimension reduction, modelling, and estimation;
- Understand the use of neural networks and apply principles of deep networks in solving business and finance problems;
- Understand and use the necessary technology infrastructure needed to solve the problem with very large datasets, including basic business programming, analytics infrastructure, operational infrastructure, and data management infrastructure.

- Learn how to apply higher cognitive skills in business and economics (esp. finance) to the solution and, in the process, add substantial value to any organization facing such a problem.

Upon completion of the applications segment, students will be able to apply the learning outcomes of the skills section in an applied institutional setting. Some examples of potential applications include:

- Predictive analytics in investments and trading
- Economic policy making, including interest rate decisions
- Credit scores and ratings used in lending decisions
- Financial fraud (incl. cybercrime)
- Customer segmentation and targeting

## REQUIRED COURSE MATERIALS AND READINGS

Course content and class communication available on Avenue:

- <http://avenue.mcmaster.ca>

## OPTIONAL COURSE MATERIALS AND READINGS

McKinney; Python for Data Analysis, First Edition; O'Reilly Media, 2012:

- ISBN: 978-1449319793

Richert and Coelho; Building Machine Learning Systems with Python, First Edition; Packt Publishing, 2013:

- ISBN: 978-1782161400

Mayer-Schonberger and Cukier; Big Data: A Revolution That Will Transform How We Live, Work, and Think, First Edition; Eamon Dolan/Mariner Books, 2014:

- ISBN: 978-0544227750

Provost and Fawcett; Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking, First Edition; O'Reilly Media, 2013:

- ISBN: 978-1449361327

<b>EVALUATION</b>
-------------------

This course will be delivered in a virtual classroom environment, comprising pre-recorded and live lectures, virtual office hours, group work, and online activities. The final student grade will be calculated as follows:

### ***Components and Weights***

Practice Assignments	Completed online (individual)	5%
Two Term Tests	Written in the class hours (individual)	2 x 25%
Demonstration Assignment	Due in the 10 <sup>th</sup> week of classes (group)	15%
Big Data Project	Due at the end of the term (group)	30%
<b>Total</b>		<b>100%</b>

### ***Practice Assignments***

The purpose of the ***Practice Assignments*** is to provide students an opportunity to learn basic programming concepts on a self-guided basis. A bit like tutorials, these assignments are designed to help learn by 'doing' it - i.e. by writing actual code.

The assignments are graded for attempt and submission – but not accuracy. There are 16 such assignments available, of which any ten (10) can be attempted and submitted to receive the full grade allocated for this component. Submission timelines are provided in the ***Course Schedule***.

### ***Demonstration Assignment***

Students will form groups for this component. Each group will be assigned a ***Demonstration*** topic to present virtually in one of the weeks. Refer to the course website for a list of past topics.

The group will also complete the ***Big Data Project*** referred to in this document. *The group members will be assigned individual grades relative to the group grade based on the peer assessments completed towards the end of the course.* More details of the format, structure, and length of the project will be provided during the term.

## **Term Tests**

Two (2) term tests will be written on specific dates and times. More details of the format, structure, and content coverage will be provided in the first week of classes.

A student *missing a test* is required to contact the 'Student Experience – Academic (MBA) office' and obtain an official approval of relief if he or she wishes to avoid getting a zero (0) grade for the test. If the 'Student Experience – Academic (MBA) office' adjudicates that relief be provided, the student will be able to write an alternate exam, in lieu of the missed test, during the final exam period at the end of the term.

## **Big Data Project**

The project allows students to work on a data analytics problem faced by an institution and apply key skills such as the analysis of large datasets, application of modern machine learning, and use of the necessary technology infrastructure. The project will culminate in presenting a management-level solution to the proposed problem.

The machine learning aspect of the project has two components:

1. Traditional learning
2. Deep learning

Examples of projects include:

- *Predictive analytics in ecommerce and retail*: Which products or categories to market to a customer given the customer profile?
- *Predictive analytics in investments and trading*: Which stocks or securities to purchase following a sequence of events?
- *Pricing for new technology products or services*: How to price a new tech product given competitive, customer, and transactional data?
- *Credit scores and ratings*: How to assess the credit risk of a borrower given the borrower profile and meta data?
- *Financial fraud*: What is the likelihood of fraud for a user attempting to access your personal finance solution?
- *Customer segmentation and targeting*: What value to assign to a customer based on the past purchase and/or transactional data and customer profile?

Students will work in the same groups for this component as for the ***Demonstration Assignment***. *The group members will be assigned individual grades relative to the group grade based on the peer assessments completed towards the end of the course.* More details of the format, structure, and length of the project will be provided during the term.

Any student missing the submission deadline will be subject to a marks deduction equal to 25% of the project grade for each day rounded **up** that the submission is late.

## COMMUNICATION AND FEEDBACK

Students that are uncomfortable in directly approaching an instructor regarding a course concern may send a confidential and anonymous email to the respective Area Chair or Associate Dean:

<http://mbastudent.degroote.mcmaster.ca/contact/anonymous/>

Students who wish to correspond with instructors or TAs directly via email must send messages that originate from their official McMaster University email account. This protects the confidentiality and sensitivity of information as well as confirms the identity of the student. Emails regarding course issues should NOT be sent to the Administrative Assistant.

Instructors should conduct an informal course review with students to allow time for modifications in curriculum delivery. Instructors should provide evaluation feedback for at least 10% of the final grade to students prior to Week #8 in the term.

## COURSES WITH AN ONLINE ELEMENT

Our course may use on-line elements (e.g. e-mail, Avenue to Learn, etc.). Students should be aware that, when they access the electronic components of a course using these elements, private information such as first and last names, usernames for the McMaster e-mail accounts, and program affiliation may become apparent to all other students in the same course. The available information is dependent on the technology used. Continuation in a course that uses on-line elements will be deemed consent to this disclosure. If you have any questions or concerns about such disclosure, please discuss this with the course instructor.

## CONDUCT EXPECTATIONS

As a McMaster student, you have the right to experience, and the responsibility to demonstrate, respectful and dignified interactions within all of our living, learning and working communities. These expectations are described in the *Code of Student Rights & Responsibilities* (the “Code”). All students share the responsibility of maintaining a positive environment for the academic and personal growth of all McMaster community members, **whether in person or online**.

It is essential that students be mindful of their interactions online, as the Code remains in effect in virtual learning environments. The Code applies to any interactions that adversely affect, disrupt, or interfere with reasonable participation in University activities. Student disruptions or behaviours that interfere with university functions on online platforms (e.g. use of A2L, WebEx or Zoom for delivery), will be taken very seriously and will be investigated. Outcomes may include restriction or removal of the involved students’ access to these platforms.

## EXTREME CIRCUMSTANCES

The University reserves the right to change the dates and deadlines for any or all courses in extreme circumstances (e.g., severe weather, labour disruptions, etc.). Changes will be communicated through regular McMaster communication channels, such as McMaster Daily News, A2L and/or McMaster email.

## COPYRIGHT AND RECORDING

Students are advised that lectures, demonstrations, performances, and any other course material provided by an instructor include copyright protected works. The Copyright Act and copyright law protect every original literary, dramatic, musical and artistic work, **including lectures** by University instructors.

The recording of lectures, tutorials, or other methods of instruction may occur during a course. Recording may be done by either the instructor for the purpose of authorized distribution, or by a student for the purpose of personal study. Students should be aware that their voice and/or image may be recorded by others during the class. Speak with the instructor if this is a concern for you.

## ACADEMIC INTEGRITY

You are expected to exhibit honesty and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity. **It is your responsibility to understand what constitutes academic dishonesty.**

Academic dishonesty is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage. This behaviour can result in serious consequences, such as the grade of zero on an assignment, loss of credit with a notation on the transcript (notation reads: “Grade of F assigned for academic dishonesty”), and/or suspension or expulsion from the university. For information on the various types of academic dishonesty, refer to the policy: [www.mcmaster.ca/academicintegrity](http://www.mcmaster.ca/academicintegrity)

The following illustrates only three forms of academic dishonesty:

1. Plagiarism, e.g. submission of work that is not one’s own or for which other credit has been obtained.
2. Improper collaboration in group work.
3. Copying or using unauthorized aids in tests and examinations

## GENERATIVE AI

Students are **not permitted** to use generative AI to complete any graded assessments in this course. In alignment with McMaster academic integrity policy, it “shall be an offence knowingly to ... submit academic work for assessment that was purchased or acquired from another source”. This includes work created by generative AI tools. Also stated in the policy is the following: “Contract Cheating is the act of outsourcing of student work to third parties with or without payment.” Using generative AI tools is a form of contract cheating. Charges of academic dishonesty will be brought forward to the Office of Academic Integrity.

## COURSE MODIFICATION

From time to time, there may be a need to add and/or remove topics or to change the schedule or delivery format. If these are necessary, you will be given as much advance notice as possible.

## **MISSED ACADEMIC WORK**

### ***Missed Mid-Term Examinations / Tests / Class Participation***

Please do not use the online McMaster Student Absence Form (MSAF) as this is for Undergraduate students only. The MBA program will not accept an MSAF.

When students miss regularly scheduled term work which contributes 10% or more to the final grade, for legitimate reasons as determined by the Student Experience – Academic Office (SEAO), the activity necessary to compensate for the missed work will be determined by the course instructor. The compensatory activities assigned will vary with the nature of the course and the missed requirement. They include, but are not restricted to, an alternative assignment, a rescheduled midterm exam, or re-weighting the marks for the missed component to other mark components. Documentation explaining such missed work must be provided to the SEAO within five (5) working days of the scheduled date for completion of the work.

Acceptable reasons for missed work, along with the Petition for Missed Term Work and the MBA Student McMaster University Student Health Certificate, can be found on the DeGroot MBA Student website ([mbastudent.degrootemcmaster.ca](http://mbastudent.degrootemcmaster.ca)). Please direct any questions about acceptable documentation to the MBA Academic Advisors ([askmba@mcmaster.ca](mailto:askmba@mcmaster.ca)).

University policy states that a student may submit a maximum of three (3) Petition for Missed Term Work per academic year, after which the student must meet with the Director of the program.

If term work is missed without an approved reason, students will receive a grade of zero (0) for that component.

### ***Missed Final Examinations***

Students must be available for the duration of the posted exam period regardless of their personal exam schedule. This is to ensure student availability throughout the entire exam period in the event that an exam must be rescheduled due to unforeseen circumstances (university closure, power outage, storm policy, etc.). A student who misses a final examination without valid reason will receive a mark of 0 on the examination.

Students who have missed a final exam for a valid reason can apply to the SEAO to write a deferred examination by submitting an Application for Deferring a Final Exam with supporting documentation. The application must be made within five days of the scheduled exam.

The Application for Deferring a Final Exam and the MBA Student McMaster University Student Health Certificate can be found on the DeGroot MBA Current Student website ([mbastudent.degrootemcmaster.ca](http://mbastudent.degrootemcmaster.ca)).



Deferred examination privileges, if granted, are normally satisfied during the examination period at the end of the following semester. In select cases, the deferred examination may be written at a time facilitated by the SEAO and agreed to by the course instructor.

Requests for a second deferral or rescheduling of a deferred examination will not be considered.

### **ACADEMIC ACCOMMODATION FOR STUDENTS WITH DISABILITIES**

Student Accessibility Services (SAS) offers various support services for students with disabilities. Students are required to inform SAS of accommodation needs for course work at the outset of term. Students must forward a copy of such SAS accommodation to the instructor normally, within the first three (3) weeks of classes by setting up an appointment with the instructor. If a student with a disability chooses NOT to take advantage of an SAS accommodation and chooses to sit for a regular exam, a petition for relief may not be filed after the examination is complete. The SAS website is:

<http://sas.mcmaster.ca>

Students who are leveraging accommodation for tests and exams are supported by the SEAO. These exams are written at the Ron Joyce Centre and do not take place in the Tim Nolan Testing Centre. Correspondence for accommodations is managed via the DSBSAS@mcmaster.ca email address. Students must communicate their intent to leverage accommodations on a test or exam a minimum of 10 business days prior to the assessment.

### **ACADEMIC ACCOMMODATION FOR RELIGIOUS, INDIGENOUS OR SPIRITUAL OBSERVANCES (RISO)**

Students requiring academic accommodation based on religious, indigenous or spiritual observances should follow the procedures set out in the RISO policy. Students should submit their request to their Faculty Office **normally within 10 working days** of the beginning of term in which they anticipate a need for accommodation or to the Registrar's Office prior to their examinations. Students should also contact their instructors as soon as possible to make alternative arrangements for classes, assignments, and tests.

### **ACKNOWLEDGEMENT OF COURSE POLICIES**

Your registration and continuous participation (e.g. on A2L, in the classroom, etc.) to the various learning activities of this course will be considered to be an implicit acknowledgement of the course policies outlined above, or of any other that may be announced during lecture and/or on A2L. **It is your responsibility to read this course outline, to familiarize yourself with the course policies and to act accordingly.**

Lack of awareness of the course policies **cannot be invoked** at any point during this course for failure to meet them. It is your responsibility to ask for clarification on any policies that you do not understand.

<b>COURSE SCHEDULE</b>			
<b>CL.</b>	<b>WEEK OF</b>	<b>CONTENT</b>	<b>DUE*</b>
1	Tue. Sep 12	<b>Introduction</b> Intro to data science, example applications	-
2	Tue. Sep 19	<b>Programming for Business Students</b> Review of Python and object-oriented languages	<b>P-1 to P-5</b>
3	Tue. Sep 26	<b>Programming for Big Data</b> Big data programming principles, big data libraries	<b>P-6 to P-10</b> <i>Group information</i>
4	Tue. Oct 3	<b>Term Test 1 (Written in Class Hours)</b>	
5	<i>Virtual</i>	<b>Machine Learning for Business Students</b> Statistical background, regression/classification algorithms	<b>P-11, P-12</b>
6	Tue. Oct 17	<b>Machine Learning for Big Data I</b> Clustering and dimension reduction for big datasets	<b>M-1, M-2</b>
7	Tue. Oct 24	<b>Machine Learning for Big Data II</b> Boosting and pipelines for big datasets	<b>M-3, M-4</b>
8	Tue. Oct 31	<b>Term Test 2 (Written in Class Hours)</b>	
9	Tue. Nov 7	<b>Big Data Ecosystem</b> Hadoop/Spark, Scala, Databricks, cloud-based platforms	-
10	Tue. Nov 14	<b>Deep Learning with Big Data I</b> Neural networks, optimizations, TensorFlow basics	-
-	Mon. Nov 20	<b>Demonstration Assignment – Due online by 10.00 AM</b>	
11	Tue. Nov 21	<b>Deep Learning with Big Data II</b> Deep networks, CNNs, RNNs, TensorFlow extended	-
12	Tue. Nov 28	<b>Deep Learning Applications</b> AI in finance, applications in business and investments	-
13	<i>Virtual</i>	<b>Big Data Project</b> Progress Check, Advice, and Feedback	-
-	Fri. Dec 8	<b>Big Data Project – Due online by 11.59 PM</b>	

\* *Practice Assignments* are due on a weekly basis and are numbered **P-1 to P-12** and **M-1 to M-4**. Each assignment is due online by 10.00 AM on the Monday following the listed date of the respective week.