McMaster University

DeGroote SCHOOL OF BUSINESS

BUSINESS F743 – Big Data in Finance
Summer 2020 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: **Online on Thursdays**
All times referenced in this document are Eastern.

Adeel Mahmood
Instructor
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 with an institutional user such as a financial services firm. 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 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

EVALUATION

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

Course Engagement	Participation in online activities (individual)	5%
Presentations	Virtual presentation in the last set of classes (group)	15%
Three (3) Mini Tests	Written online (individual)	50%
Big Data Project	Due at the end of the term (group)	30%
Total		**100%**

Course Engagement

Course Engagement marks are based on the level of student engagement in the online content items and activities contained in the course. Each item or activity is tracked online (in Avenue) to keep a record of its completion by the student. Student attendance in, or subsequent viewing of, online lectures is also recorded.

A list of the content items and activities requiring student completion is provided in Avenue to view at the beginning of the term. Refer to the *Course Engagement* section of the course for more details on the evaluation aspects of this component.

Presentations

Students will form groups for this component. Each group will be assigned a *Presentation* 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.

Mini Tests

Three (3) mini tests will be **written online 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's grade for the test will be calculated based on the scheme outlined in the *Missed Tests* document available online.

The three mini tests combine to form 50% of the final grade. If a student *writes all three mini tests*, the two tests in which the student's percentage marks are the highest will each form 20% of the student's final grade with the third test forming the remaining 10%.

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.

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 *Presentations*. *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.

ONLINE COURSE COMPONENTS

In this course we will be using Avenue to Learn (A2L). Students should be aware that when they access the electronic components of this course, private information such as first and last names, user names 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 this course will be deemed consent to this disclosure.

If you have any questions or concerns about such disclosure, please discuss it with the instructor.

McMaster University values integrity, inclusiveness and teamwork, and strives to support the personal and collective growth of the McMaster Student Community. As a McMaster student you are a member of a community that values taking responsibility for oneself and towards your fellow community members, creating an environment conducive to the intellectual and personal growth of all who study, work, and live within the McMaster Community - both in-person and online. These expectations are embodied in our Code of Student Rights & Responsibilities ('the Code,' https://secretariat.mcmaster.ca/app/uploads/Code-of-Student-Rights-and-Responsibilities.pdf).

The Code remains in effect in a virtual learning environment. It applies to any interactions that adversely affect, disrupt, or interfere with reasonable participation in University activities. Given that most virtual University functions at this time rely heavily on online platforms (e.g. use of WebEx for delivery), disruptions will continue to be addressed and could result in outcomes that restrict a student's access to these platforms. It is essential that students be mindful of their Rights and Responsibilities in their interactions online.

Some helpful information can be found here: https://sscm.mcmaster.ca/

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.

ACADEMIC DISHONESTY

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.

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, e.g. 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.

It is your responsibility to understand what constitutes academic dishonesty. For information on the various types of academic dishonesty please refer to the Academic Integrity Policy at:
www.mcmaster.ca/academicintegrity

The following illustrates only three forms of academic dishonesty:

1. Plagiarism, e.g. the 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

MISSED ACADEMIC WORK

Missed Mid-Term Examinations / Tests / Class Participation

Where students miss a regularly scheduled mid-term or class participation for legitimate reasons as determined by the Student Experience – Academic (MBA) office, the weight for that test/participation will be distributed across other evaluative components of the course at the discretion of the instructor. Documentation explaining such an absence must be provided to the Student Experience – Academic (MBA) office within five (5) working days upon returning to school.

To document absences for health related reasons, please provide to Student Experience – Academic (MBA) office the Petition for Relief for MBA Missed Term Work and the McMaster University Student Health Certificate which can be found on the DeGroote website at http://mbastudent.degroote.mcmaster.ca/forms-and-applications/. Please do not use the online McMaster Student Absence Form as this is for Undergraduate students only. University policy states that a student may submit a maximum of three (3) medical certificates per year after which the student must meet with the Director of the program.

To document absences for reasons other than health related, please provide Student Experience – Academic (MBA) office the Petition for Relief for MBA Missed Term Work and documentation supporting the reason for the absence.

Students unable to write a mid-term at the posted exam time due to the following reasons: religious; work-related (for part-time students only); representing university at an academic or varsity athletic event; conflicts between two overlapping scheduled mid-term exams; or other extenuating circumstances, have the option of applying for special exam arrangements. Such requests must be made to the Student Experience – Academic (MBA) office at least ten (10) working days before the scheduled exam along with acceptable documentation. Instructors cannot themselves allow students to unofficially write make-up exams/tests. Adjudication of the request must be handled by Student Experience – Academic (MBA).

If a mid-term exam is missed without a valid reason, students will receive a grade of zero (0) for that component.

Missed Final Examinations

A student who misses a final examination without good reason will receive a mark of 0 on the examination.

All applications for deferred and special examination arrangements must be made to the Student Experience – Academic (MBA) office. Failure to meet the stated deadlines may result in the denial of these arrangements. Deferred examination privileges, if granted, must be satisfied during the examination period at the end of the following term. There will be one common sitting for all deferred exams.

Failure to write an approved deferred examination at the pre-scheduled time will result in a failure for that examination, except in the case of exceptional circumstances where documentation has been provided and approved. Upon approval, no credit will be given for the course, and the notation N.C. (no credit) will be placed on the student's transcript. Students receiving no credit for a required course must repeat the course. Optional or elective courses for which no credit is given may be repeated or replaced with another course of equal credit value.

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

Any student who is unable to write a final examination because of illness is required to submit the Application for Deferred MBA Final Examination and a statement from a doctor certifying illness on the date of the examination. The Application for Deferred MBA Final Examination and the McMaster University Student Health Certificate can be found on the DeGroote website at http://mbastudent.degroote.mcmaster.ca/forms-and-applications/ Please do not use the online McMaster Student Absence Form as this is for Undergraduate students only. Students who write examinations while ill will not be given special consideration after the fact.

In such cases, the request for a deferred examination privilege must be made in writing to the Student Experience – Academic (MBA) office within five business days of the missed examination.

Special examination arrangements may be made for students unable to write at the posted exam time due to compelling reasons (for example religious, or for part-time students only, work-related reasons):

Students who have religious obligations which make it impossible to write examinations at the times posted are required to produce a letter from their religious leader stating that they are unable to be present owing to a religious obligation.

Part-time students who have business commitments which make it impossible to write examinations at the times posted are required to produce a letter on company letterhead from the student's immediate supervisor stating that they are unable to be present owing to a specific job commitment.

In such cases, applications must be made in writing to the Student Experience – Academic (MBA) office at least ten business days before the scheduled examination date and acceptable documentation must be supplied.

If a student is representing the University at an academic or athletic event and is available at an overlapping scheduled time of the test/examination, the student may write the test/examination at an approved location with an approved invigilator, as determined by the Student Experience – Academic (MBA) office.

In such cases, the request for a deferred examination privilege must be made in writing to the Student Experience – Academic (MBA) office within ten business days of the end of the examination period.

Note: A fee of $50 will be charged for a deferred exam written on campus and a fee of $100 for deferred exams written elsewhere. In cases where the student's standing is in doubt, the Graduate Admissions and Study Committee may require that the student with one or more deferred examination privileges refrain from re-registering until the examination(s) have been cleared.

STUDENT ACCESSIBILITY SERVICES

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

RESEARCH DATA

Any research data obtained by a student from publicly available sources for the purposes of completing term work (including data used in assignments, projects, and tests) can be used by the student, the instructor, and the university. All parties are free to use the research data subject to the original copyright of such data. Students hereby agree not to use in their term work any proprietary data or data subject to copyright protection without the prior written approvals of the instructor and the owner of any such copyright.

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.

POTENTIAL MODIFICATIONS TO THE COURSE

The instructor and university reserve the right to modify elements of the course during the term. The university may change the dates and deadlines for any or all courses in extreme circumstances. If either type of modification becomes necessary, reasonable notice and communication with the students will be given with explanation and the opportunity to comment on changes. It is the responsibility of the student to check their McMaster email and course websites weekly during the term and to note any changes.

COURSE SCHEDULE

CL.	DATE	CONTENT	DUE
1	Thu. Jun 4	**Introduction** Intro to Data Science, Example Applications	-
2	Thu. Jun 11	**Programming for Business Students I** Introduction to Python and Object-Oriented Languages	-
3	Thu. Jun 18	**Programming for Business Students II** Basic Programming Principles	• *Group information*
4	Thu. Jun 25	**Programming for Business Students III** Advanced Programming Techniques	• **Mini Test 1** *(Written at 12.00 Noon)*
5	Thu. Jul 2	**Programming for Business Students IV** Using Libraries such as Numpy, Pandas	-
6	Thu. Jul 9	**Applied Machine Learning I** Statistical background, clustering, data modelling, estimation/prediction, data visualization	• **Mini Test 2** *(Written at 12.00 Noon)*
7	Thu. Jul 16	**Applied Machine Learning II** Regression Algorithms	-
8	Thu. Jul 23	**Applied Machine Learning III** Classification Algorithms	• **Mini Test 3** *(Written at 12.00 Noon)*
9	Thu. Jul 30	**Big Data Ecosystem I** Data Collection, Cleaning, and Preparation	• *Presentation 1* • *Presentation 2*
10	Thu. Aug 6	**Big Data Ecosystem II** Hadoop, MapReduce, Scala, and Spark	• *Presentation 3* • *Presentation 4*
-	Fri. Aug 7	**Big Data Project** – Due by 11.59 PM	