

The University of Western Ontario
London, Ontario, Canada
Department of Computer Science

COMPSCI 2034B / DIGIHUM 2144B

Data Analytics: Principles and Tools
Winter 2020 (0.5 Credit Course)

Instructor:	Chris Brogly	Time:	Tuesday 2:30PM – 4:30PM
Email:	cbrogly@uwo.ca	Place:	SSC 2050
Office:	MC 365, see below for office hours		

1 Course Description

A comprehensive and interdisciplinary introduction to data analytics using modern computing systems, with equal attention to fundamentals and practical aspects. Topics include sources of data, data formats and transformation, usage of spreadsheets and databases, statistical analysis, pattern recognition, data mining, big data, and methods for data presentation and visualization.

2 Learning Outcomes

By the end of this course, students will be able to:

- Describe types, sources, and facets of data.
- Use spreadsheets, macros and VBA to analyze data.
- Utilize programming constructs to process data.
- Explain relevant statistical analysis and machine learning techniques.
- Evaluate different visual artifacts for data presentation.
- Choose data analytics tools appropriate for data problems.

3 Textbook

This course requires all students to subscribe to an electronic zyBook that will be used for assigned readings and pre-class activities. A subscription may be purchased on-line via zybooks.com or through the university book store (bookstore.uwo.ca).

Instructions for Subscribing to the zyBook

1. Sign in or create an account at learn.zybooks.com. Ensure that you use your UWO e-mail for your account so your participation can be properly recorded.
2. Enter zyBook code: **TO BE POSTED ON OWL**
3. Subscribe.

A subscription costs \$65 USD (approximately \$88.33 CAD at present).

Please contact support@zybooks.com for details on refunds (e.g. if you drop the course), options to extend your subscription beyond the end of the course or other help with using the zyBook software.

An alternative format (accessibility mode) is available for students using assistive technologies. Please contact accessibility@zyBooks.com for information about alternative formats or to request access to an accessible version of the zyBook.

4 Instructor

Chris Brogly MSc

E-Mail: cbrogly@uwo.ca

Office: Middlesex College, Room 365 (MC365)

Office Hours:

- Tuesday 4:30PM to 5:30PM
- Thursday 4:30PM to 5:30PM

If you can't see me during these times, then e-mail me to setup an appointment as I am usually around campus.

5 Teaching Assistants

Teaching Assistant (TA) contact information and consulting hour details will be posted on OWL when available.

Teaching assistants will hold an open consulting hour each week. TA consulting hours start January 21st. No consulting hours will be held on the week of February 18th (reading week) or after April 6th.

6 Lectures

There will be one lecture held each week in SSC 2050 at the following time:

- Tuesday 2:30PM - 4:30PM

Students are expected to bring blank paper and writing implements to each lecture. A laptop, tablet, phone or other electronic device capable of connecting to the internet and viewing web pages is also strongly recommended (required for participation in some in-class activities).

7 Course Website & E-Mail Communication

This course uses the Online Western's Learning (OWL) system (<https://owl.uwo.ca>).

Announcements, assignments, labs, lecture notes, and other course-related information will be posted on this website. It is the responsibility of the student to ensure their UWO account is operational for both E-Mail and using the OWL course site and routinely check both for updates

Before sending questions to the course instructor, please first check the course syllabus (this document), the OWL course site and the OWL course forums before sending your inquiry. **General questions that may be helpful to other students (and do not contain personal information) should be posted to the OWL course forums so that others may benefit from the answer.** E-mail messages should be sent from your UWO e-mail account to the instructor and include "CS2034" in the subject line to avoid any spam filters.

You can normally expect a response with-in 48 hours during the week, providing it is not on weekends and holidays. Please ensure a reasonable amount of time is available for a response when asking questions regarding evaluations that are nearing a deadline.

8 Evaluation

Element	Weight
Assignments (3) A1, A2 – 5% A3 – 8%	18%
Labs – 11 10 required	10%
Midterm	32%
Exam	40%
Total	100%

If an assignment or lab has to be cancelled for any reason, the remaining assignment/quiz/lab weights will be prorated (scaled) to add up to the total weight for that element.

8.1 Assignments

- There will be 3 assignments.
- Tentative assignment schedule (dates subject to change) and weights are as follows:

Assignment #	Tentative Posting	Due On	Weight
1	January 22nd	February 8th	5%
2	February 16th	March 12th	5%
3	March 18th	April 5th	8%

- Assignments are due 5 minutes before midnight (23:55) of the due date.
- All submissions will be submitted electronically. Details will be given in the assignment descriptions.
- Late assignments are strongly discouraged:
 - 15% will be deducted from a late assignment up to 24 hours after the due date/time.
 - 30% will be deducted from a late assignment 24 to 48 hours after the due date/time. – After 48 hours from the due date/time, late assignments will receive a zero grade.
- Any changes, updates, and clarifications to assignments will also be posted on the website. It is your responsibility to monitor these pages closely.

- A program that produces the correct output is not necessarily a “working” program; it must also satisfy the specifications given in the assignment description.
- Your assignment solutions are expected to be your own individual work, not the products of group effort. **You may not share your assignment solutions** with another student for any reason nor are you to request solutions from another student or make use of solutions available on-line. Such sharing of solutions is considered a scholastic offence and will be penalized as such.
- It is your responsibility to keep up-to-date off-site backups (e.g. on OneDrive or Dropbox) of assignment files. Retain copies of all material handed in, as well as the graded assignment, to guard against the possibility of lost assignments or errors in recording marks. You should keep these materials until you are satisfied that your final mark for the course has been computed properly.
- Assignments will be marked by the TAs, who follow marking schemes provided by the instructor.
- Every effort will be made to have assignments marked and handed back within 3 weeks after the due date, preferably sooner. When assignment marking has been completed, you will be informed via the course website and/or e-mail.
- You should direct any questions or appeals about marking to your TA. If your discussion with the TA is not satisfactory, you may further discuss the issue with the course instructor.
- A request for an adjustment in an assignment mark must be made within 1 week following the date the assignment is returned. All assignment marks are considered to be final after that date.
- The course instructor reserves the right to completely remark an appealed assignment. This may raise or lower the grade compared to that given by the TA.
- Extensions are only granted for approved accommodations from the Dean’s Office, Academic Counselling or SSD (see Sections [10.5](#) and [10.6](#)).

8.2 Labs

- **Labs start the week of January 13th and run to April 3rd.**
- There will be 10 equally weighted two-hour labs, worth 1 % each.
- See the [Western Timetable](#) for location and time information.
- Labs will be run by TAs and act as practical tutorial sessions that may cover new material or review concepts discussed in lectures. Collaboration on lab problems with other students is allowed and encouraged.
- Lab descriptions will be posted on the OWL course site before the dates listed above. Any changes, updates, or clarifications to labs will also be posted on the website. It is your responsibility to monitor these pages closely.
- It is expected that you read over lab materials beforehand and come prepared to lab sessions with any required resources or reference material.
- As seating in labs is limited, **you are required to attend the lab section you are registered in.**
- To obtain full marks for a given lab session, you must attend the lab in person and complete the lab to the TA’s satisfaction (see the Lab Rubric at the end of this document for details on how lab marks are evaluated).
- TAs may give a zero or reduced lab mark to students that are late to or do not work on lab material during the lab session. It is expected that students abstain from working on assignments during lab time until they have completed and received a mark for the lab.
- **Labs can not be made up for in future lab sessions or submitted via e-mail.**

- Labs missed as part of an approved accommodation will result in the student's lab mark being re-weighted to not include the missed lab. Please contact the course instructor as soon as possible if you have an accommodated lab absence.

8.3 Midterm & Final Exam

- Tentative exam dates are as follows:

Midterm Exam

Date: Saturday February 29th or Saturday March 7th (tentative)

Time: To be determined (2 hours) (tentative)

Location: To be determined

Format: Mixed

Final Exam

Date: To be determined

Time: To be determined (3 hours)

Location: To be determined

Format: Mixed

- Both exams will be comprehensive and may include anything covered in lectures, labs, assignments, and assigned readings up until that point in the course.
- No electronic devices of any kind, including but not limited to calculators, phones, laptops and smart watches, will be allowed during the examinations.
- No additional materials will be permitted during exams.
- There will be no make-up midterm exam. Students absent due to an approved accommodation (see Section 10.6) will have their final exam mark reweighted to include the weight of the midterm exam.

9 Scholastic Offences & Academic Dishonesty

To preserve a fair and honest learning environment for all, scholastic offences and any breach of ethical conduct will not be tolerated and punished according to department policy. In addition, to the offences outlined in the department and university policy (see Section 10.1), the following will be considered academic misconduct:

- Sharing assignment solutions with other students or posting them on-line in a manner that is accessible to other students (e.g. Course Hero, in Facebook groups, via Discord, etc.).
- Using any part of another student's assignment solution including but not limited to formatting and templates.
- Sharing your account or using another student's UWO, OWL or zyBook accounts in any way.
- Submitting a Group Work Code for an activity you did not participate in or were absent for.
- Altering a lab, assignment, midterm or exam in any way after it has been graded.
- Misrepresenting work done during labs with the intent to receive a higher lab grade.
- Taking the place of another student in labs sessions, exams, quizzes, etc.
- Any other offences listed in a department or university policy (see Section 10.1).

Assignment will be subject to metadata analysis to detect sharing of files and solutions. For first offences, misconduct on assignments, quizzes and labs will result in a 0% grade for that assignment/lab/quiz in

addition to a penalty equal to the weight of the assignment/quiz/lab. First offence relating to misconduct on any participation component will result in a 0% overall participation grade and an additional penalty equal to 10%. Second offences on any course component will result in a 0% overall course grade.

10 University & Department Policies

10.1 Ethical Conduct

Scholastic offences are taken seriously and students are directed to read the appropriate policy, specifically, the definition of what constitutes a Scholastic Offence, at the following Web site: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/scholastic_discipline_undergrad.pdf.

Students must write their essays and assignments in their own words. Whenever students take an idea, or a passage from another author, they must acknowledge their debt both by using quotation marks where appropriate and by proper referencing such as footnotes or citations. Plagiarism is a major academic offence.

Assignments that are judged to be the result of academic dishonesty will, for the student's first offence, be given a mark of zero with an additional penalty equal to the weight of the assignment also being applied.

Computer-marked multiple-choice tests and/or exams may be subject to submission for similarity review by software that will check for unusual coincidences in answer patterns that may indicate cheating.

All required papers and assignments may be subject to submission for textual and metadata similarity review to commercial or custom plagiarism detection software under license to the University for the detection of plagiarism. All papers submitted for such checking via Turnitin.com will be included as source documents in the reference database for the purpose of detecting plagiarism of papers subsequently submitted to the system. Use of this service is subject to the licensing agreement, currently between The University of Western Ontario and Turnitin.com (<http://www.turnitin.com>).

Students are responsible for reading and respecting the Computer Science Department's policy on [Scholastic Offences](#) and [Rules of Ethical Conduct](#).

10.2 Tutoring

The role of tutoring is to help students understand course material. Tutors should not write assignments or tests for the students who hire them. Submitting an assignment that contains material written by a tutor is an academic offence.

Having employed the same tutor as another student is not a legitimate defence against an accusation of collusion, should two students hand in assignments judged similar beyond the possibility of coincidence.

10.3 Email Contact

We occasionally need to send email messages to the class or to students individually. Such emails are sent to the UWO email address as assigned to you by Information Technology Services (ITS), i.e. your email address @uwo.ca. It is your responsibility to read your email account on a frequent and regular basis, or to have it forwarded to an alternative email address if preferred. See the ITS website for directions on forwarding email.

However, note that the email at ITS (your UWO account) and other email providers may have quotas or limits on the amount of space they dedicate to each account. Unchecked emails may accumulate beyond those limits and you may be unable to retrieve important messages from your instructors. Losing emails is not an acceptable excuse for not knowing about the information that was sent.

Students are encouraged to contact their course instructor via e-mail with brief, e-mail appropriate questions regarding lecture materials or clarification of assignments. However, before sending email to the instructor, the student should check the course website to see if the requested information is already there and post on the OWL forums if appropriate. Students must send emails from their UWO ITS account and include “CS2034” in the subject line of the email.

10.4 Computing Facilities & OWL Forums

Each student will be given an account on the Computer Science Department undergraduate computing facility, GAUL. In accepting the GAUL account, a student agrees to abide by the department’s [Rules of Ethical Conduct](#).

Students are expected to act professionally when making posts to the OWL course forums. All submissions should be relevant to the course, made in the correct subforum and not contain any content that would violate department or university policies including the Student Code of Conduct. Assignment solutions (including partial or incomplete answers) should never be posted to the OWL Forums. When seeking help with an assignment or answering questions, only small code segments or specific technical questions should be posted.

10.5 Accessibility Statement

Please contact the course instructor if you require material in an alternate format or if you require any other arrangements to make this course more accessible to you. You may also wish to contact Services for Students with Disabilities (SSD) at 661-2111 x82147 for any specific question regarding an accommodation.

The policy on Accommodation for Students with Disabilities can be found here: https://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_disabilities.pdf

10.6 Academic Accommodation

If you are unable to meet a course requirement due to illness or other serious circumstances, you must provide valid medical or other supporting documentation, depending on the missed requirement, as soon as possible and contact your instructor immediately. It is the student’s responsibility to make alternative arrangements with their instructor once the accommodation has been approved and the instructor has been informed. In the event of a missed final exam, a “Recommendation of Special Examination” form must be obtained from the Dean’s office immediately. For further information, please see: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_medical.pdf.

A student requiring academic accommodation due to illness should use the Student Medical Certificate when visiting an off-campus medical facility or request a Record’s Release Form (located in the Dean’s office) for visits to Student Health Services. The form can be found here: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/medicalform.pdf.

10.7 Support Services

Learning-skills counsellors at the Student Development Centre (<http://www.sdc.uwo.ca>) are ready to help you improve your learning skills. They offer presentations on strategies for improving time management, multiple-choice exam preparation/writing, textbook reading, and more. Individual support is offered throughout the Fall/Winter terms in the drop-in Learning Help Centre, and year-round through individual counselling.

Students who are in emotional/mental distress should refer to Mental Health at Western (<http://www.uwo.ca/uwocom/mentalhealth>) for a complete list of options about how to obtain help.

The policy on Accommodation for Religious Holidays can be found here: http://www.uwo.ca/univsec/pdf/academic_policies/appeals/accommodation_religious.pdf

Students may refer to <https://westernusc.ca/your-services> for services provided by the USC.

Students may refer to <http://www.registrar.uwo.ca> for the Registrarial Services.

11 Tentative Course Outline (Subject to Change)

Week of	Tentative Lecture Topics	Assigned Readings	Lab Topic	Important Dates
Week 1 January 7 th	<ul style="list-style-type: none"> • Introduction & Syllabus • Data, Information 	<ul style="list-style-type: none"> • Course Syllabus (this document) • CH1: Computing History • CH2: Data 	No Labs This Week	Lectures start this week
Week 2 January 14 th	<ul style="list-style-type: none"> • Data, Information • Computing History • Hardware Basics • Spread Sheet Basics 	<ul style="list-style-type: none"> • CH3: The Basics: Hardware and Software • CH7: Spreadsheet Basics • Optional Readings on the History of Spreadsheets: http://www.dssresources.com/history/sshistory.html & http://www.dssresources.com/history/mattessichspreadsheet.htm 	Lab 1: Spreadsheet Basics	<ul style="list-style-type: none"> • Labs start this day to week
Week 3 January 21 st	<ul style="list-style-type: none"> • Spreadsheet Basics • Statistics: Basics • Data Prep & Transformation • Regular Expressions (RegEx) 	<ul style="list-style-type: none"> • CH10: Statistics: Basics • RegexOne: https://regexone.com 	Lab 2: Correlating Data Sources	Assignment 1 Posted: January 22 nd
Week 4 January 28 th	<ul style="list-style-type: none"> • Regular Expressions (RegEx) • Introduction to Computation • Getting Started With VBA • Excel Macros 	<ul style="list-style-type: none"> • CH4: Computation, Algorithms and Programming • Home & Learn: Getting Started (http://www.homeandlearn.org) 	Lab 3: Rank Analysis	

Week 5 February 4 th	<ul style="list-style-type: none"> • VBA Programming Part 1 	<ul style="list-style-type: none"> • CH5: Advanced Algorithms • Home & Learn: VBA Programming Variables • Home & Learn: Conditional Logic • Home & Learn: Strings and String Functions 	Lab 4: Benford's Law	Assignment 1 Due: February 9 th
Week 6 February 11 th	<ul style="list-style-type: none"> • VBA Programming Part 2 • Debugging 	<ul style="list-style-type: none"> • CH6: Troubleshooting and Debugging • Home & Learn: Programming Loops • Home & Learn: Programming Arrays • Home & Learn: Subs and Functions 	Lab 5: Debugging	Assignment 2 Posted: February 12 th

Week of	Tentative Lecture Topics	Assigned Readings	Lab Topic	Important Dates
February 18 th	Reading Week			
Week 7 February 25 th	<ul style="list-style-type: none"> • VBA Programming Part 3 • Textual and Social Media Analytics • Midterm Review 	<ul style="list-style-type: none"> • VBA Functions StrComp, Split and Replace from https://www.excelfunctions.net • CH17.3 and CH17.8 (only if you are unfamiliar with these concepts) • Midterm Overview • Midterm Tutorial 	Lab 6: Course Calendar Search Engine	Midterm: Saturday Feb 29 th or Mar 7 th

<p>Week 8 March 3</p>	<ul style="list-style-type: none"> • Intro to Visualization 	<ul style="list-style-type: none"> • CH9: Data Visualization • The Data Visualization Catalogue (https://datavizcatalogue.com/) • Interaction Design Foundation: What are Gestalt Principles? (https://www.interaction-design.org/literature/topics/gestalt-principles) 	<p>Lab 7: Visualization</p>	<ul style="list-style-type: none"> • Assignment 2 • Due: March 8th • March 7: Last day to drop.
<p>Week 9 March 10</p>	<ul style="list-style-type: none"> • Intro to AI and Machine Learning, Images • Intro to Feature Engineering • Classification 	<ul style="list-style-type: none"> • CH13: Data Mining and Machine Learning • Turing Test: https://www.youtube.com/watch?v=sXx-PpEBR7k • A Brief History of AI: https://www.youtube.com/watch?v=056v4OxKwll • Machine Learning & AI: https://www.youtube.com/watch?v=z-EtmaFJieY 	<p>Lab 8: Entropy and Decision Trees</p>	<p>Assignment 3 Posted: March 10th</p>
<p>Week 10 March 17</p>	<p>Tentatively:</p> <ul style="list-style-type: none"> • Intro to Privacy, Ethics, Issues in Machine Learning <p>Potentially:</p> <ul style="list-style-type: none"> • Intro to AI and Machine Learning 2, Natural Language Processing <p>OR</p> <p>Relational Databases</p>	<ul style="list-style-type: none"> • Tentatively: CH16: Privacy and Ethics <p>Possibly: Relational DB notes</p>	<p>Lab 9: TBD</p>	
<p>Week 11 March 24</p>	<p>Selected Topic From</p> <ul style="list-style-type: none"> • Modeling and Simulation <p>OR</p> <p>- Relational Databases</p>	<ul style="list-style-type: none"> • CH12: Modeling and Simulation <p>Possibly: Relational DB notes</p>	<p>Lab 10: TBD</p>	

Week of	Lecture Topics	Assigned Readings	Lab Topic	Important Dates
<p>Week 12 March 31</p> <p>April 3rd, last day of classes</p>	<p>Time permitting, brief selected topic from:</p> <ul style="list-style-type: none"> • Big Data, Cloud Computing • Other Data Analytics Tools • Relational Databases, Continued <p>And</p> <p>Final Exam Review</p>	<ul style="list-style-type: none"> • CH14: Big Data • CH15: Cloud Computing • CH18: Python (optional reading) <p>Possibly: Relational DB notes</p>	<p>Lab 11: TBD</p>	<ul style="list-style-type: none"> • Assignment 3 Due: April 3 <p>Last lab this week!</p> <ul style="list-style-type: none"> • Last lecture this week!
<p>Final Exam (Date, time and location to be announced)</p>				

CS2034/DH2144 Lab Rubric

Name: _____

Total: ____ / 10 marks*

	Exemplary (10 marks)	Above Average (9 - 8 marks)	Satisfactory (7 - 6 marks)	Needs Improvement (5 - 0 marks[†])
	<ul style="list-style-type: none"> • All required components of lab completed and correct. • Demonstrates understanding of lab content and correctly answers questions about their work. • Arrives on time and prepared for the lab with all required prelab readings and activities completed. 	<ul style="list-style-type: none"> • Most required components of lab completed. • Few or insignificant errors in work. • Likely that lab could be completed correctly if more time was given. • Demonstrates understanding of completed lab content and correctly answers questions about their work. • Arrives on time and prepared for the lab with all required prelab readings and activities completed. 	<ul style="list-style-type: none"> • At least half of lab content is completed. • Several or significant errors in work. • Likely that errors could be corrected if more time was given. • Likely that most of lab could be completed with more time. • Demonstrates some understanding of completed lab content and able to explain most completed work. • Arrives on time and with most prelab readings and activities completed. 	<ul style="list-style-type: none"> • Majority of lab tasks are incomplete. • Many significant errors in work. • Unlikely that errors could be corrected if more time was given. • Unlikely that most of lab could be completed with more time. • Fails to demonstrate understanding of lab content. • Arrives late, without prelab readings and activities completed.

* Some labs may have a bonus/optional part that will allow for a grade over 10. Bonus marks can make up for labs with less than a 10/10 mark but not raise the overall lab grade above 15% of your final grade.

† Lab absences will result in a grade of 0/10 for the missed lab.