

BDA400 – Lecture 0

Course Logistics

Les King

les.king@senecacollege.ca

ca.linkedin.com/pub/les-king/10/a68/426



- University of Toronto graduate in Mathematics & Computer Science
- Started at IBM at beginning of DB2 for Linux, Unix & Windows in 1992
- 25 years of Information Management and Database experience
- Previous experience as an instructor at University of Toronto in Mathematics
 - Calculus
 - Algebra
- Currently a professor at Seneca College covering:
 - Introduction to Data Warehousing
 - Introduction to Big Data
 - Introduction to DB2

All course material that you will be tested on will come from the lecture notes.

I will post the lecture notes to the Blackboard prior to the class

I will post the assignments to the Blackboard within the first two weeks

There are no mandatory text books for this class.

All assignments are due the **week before** the final exam – **no exceptions**

I will communicate through Blackboard. Announcements. Course Information. Course Documents. Discussion Room. Grade Center. Keep an eye on email.

The hands-on portion of the course (assignments) will be taught using a cloud-based Hadoop environment

Best way to reach me

- les.king@senecacollege.ca
- Office Hours

Tuesday Nights

- 6:30pm – 7:00pm – Office Hours, Extra Help
- 7:00pm – 9:00pm – Class Time
- 9:00pm – 10:00pm – Office Hours, Extra Help

Course Outline

- Course Addendum
- Purpose of assignments
- Purpose of lectures
- Assignments
- Exams

Skills Needed

- Database skills – you must know SQL for this course
- Warehousing concepts – DBW624 is a suggested pre-requisite for this course

Marking Scheme

- Assignment #1 – 6% (Hadoop)
- Assignment #2 – 6% (Spark)
- Assignment #3 – 6% (Big SQL)
- Assignment #4 – 6% (Programming in R)
- Assignment #5 – 6% (Machine Learning)

- Midterm – 30% (Lecture 1-5)

- Final Exam – 40% (Lecture 1-10)

Passing Requirements

- Minimum of 50% is required to pass the course
- Minimum of 50% on the final exam is required to pass the course

Important Dates

- Mid-Term Due: **Tuesday, February 26**
 - Lectures 1-5
 - Will be posted by **Thursday, February 21**
- Assignments Due: **Tuesday, April 9**
- Final Exam Due: **Tuesday, April 16**
 - Lectures 1-5 : ~30%-35%
 - Lectures 6-10 : ~65%-70%
 - Will be posted by **Thursday, April 11**
- Mid-Term and Final Exam are take-home exams. You will have 4-5 days to complete them. They are, obviously, open book.
- I will communicate class schedule out on the BlackBoard so there's clarity on what nights we are meeting and what nights we are not meeting.

Environment for Assignments

Access to Big Data (Hadoop) environment

<https://bigdatauniversity.com/learn/big-data/>

Videos are available to assist you with setting up the environment

Videos are available to assist with how to go about assignments

To receive marks for your assignments, use the Big Data University facility above and provide me with your personalized certificate for the specific module referenced in the assignment

WARNING: Once you fail the test, you cannot redo it. Make sure you watch the videos and do the hands-on portions before taking the test.