**裰矺㡰矵䒨矹絰矵診矵꿬Ῠ**

|  |  |
| --- | --- |
| COURSE INFORMATION Course Title: Technology Fundamentals for Business Analytics  Course Number: MGMT-6963-01  Credit Hours: 3  Semester / Year: Fall 2016  Meeting Days: Monday 6:00-8:50 PM  Room Location: Pittsburgh Building 4114  Website: | INSTRUCTOR Jason Kuruzovich: Associate Professor of Business Analytics  Office location: Pittsburgh 4108  Office Telephone Number: 518.698.9910  Office Hours: Thursday 9-12 or by appointment  Email Address: kuruzj@rpi.edu |

**COURSE DESCRIPTION**

The widespread proliferation of IT-influenced economic activity leaves behind a rich trail of micro-level data, enabling organizations to use analytics and experimentation in both strategy and operations. This course provides a hands-on introduction to the concepts, methods and processes of business analytics. We will learn how to obtain data and draw business inferences from data by asking the right questions and using the appropriate tools.

**STUDENT LEARNING OUTCOMES**

Through this course you are expected to:

1. Demonstrate an understanding of analytics based problem solving.
2. Be able to extract, match, transform, and clean data from a variety of sources;
3. Conduct structured statistical analyses to answer business related questions;

**COURSE ASSESSMENT MEASURES**

**Class Participation**

Participation in class is absolutely necessary. You are expected to volunteer solutions to in class exercises.

**Assignments (Weekly)**

You will have assignments associated with many classes. Be sure to check the LMS. Preparing the case questions in advance is expected to help you participate and contribute to class discussion. Assignments up to 24 hours late will have their grade reduced by 25%; assignments up to one week late will have their grade reduced by 50%. After one week, late assignments will receive no credit. Please turn in your assignment early if there is any uncertainty about your ability to turn it in on time.

**Kaggle Competition / Organizational Assignments (2)**

There will be two individual assignments to be completed as part of the class.

*Assignment 1 – Solution Assessment and Analysis.*

[*https://www.kaggle.com/c/titanic-gettingStarted*](https://www.kaggle.com/c/titanic-gettingStarted)

The goal of the first assignment is to address an existing Kaggle sample project and identify and diagnose 3 different solutions to the project, research any methods used, assess the relative quality of each, and make a recommendation on which is the best to utilize for the context. We will go through the process of how to go through the process of identifying and analysing solutions in class.

The Titanic is one of the most well developed and documented projects. We will all go through this.

*Assignment 2 – Solution Creation.* In this assignment you will select a challenging problem from Kaggle or from our relationships with different organizations and develop a unique solution to the problem. The solution may employ methods of analysis examined in the class or others you have investigated.

**GRADING CRITERIA**

All grading is out of 100%. Grade breakdown and feedback will be given through the course learning management system.

|  |  |
| --- | --- |
| **Component** | **Weight** |
| Participation and Homework Assignments | 25% |
| Quizzes | 25% |
| Kaggle/Organizational Assignment (Due 13th Class) | 25% |
| Midterm (7th Class) | 10% |
| Final Exam (14th Class) | 15% |

**COURSE SCHEDULE**

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Date** | **Topic** | **Readings** |
| 1 | 8/25 | Overview of Analytics and Business | 1. The Rise of the Data Scientist [[link]](http://flowingdata.com/2009/06/04/rise-of-the-data-scientist/) 2. Data Scientist: The Hottest Job You Haven’t Heard Of [[link]](http://jobs.aol.com/articles/2011/08/10/data-scientist-the-hottest-job-you-havent-heard-of/) 3. Drew Conway’s Venn Diagram [[link](http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram)] 4. What is Data Science? [[link](http://radar.oreilly.com/2010/06/what-is-data-science.html)] 5. [eScience: The Fourth Paradigm](http://research.microsoft.com/en-us/collaboration/fourthparadigm/) (Foreward and Introduction, pages xi - xxxi; Gray's Laws, pages 5-12) 6. Volume, Velocity, and Variety [[link](http://blogs.gartner.com/doug-laney/deja-vvvue-others-claiming-gartners-volume-velocity-variety-construct-for-big-data/)]   [LAB 1 Install Software] |
|  | 9/1 | Labor Day – No Classes | |
| 2 | 9/8 | Relational Databases and Introduction to Relational Algebra | 1. Relational Algebra (Wikipedia Entry) [[link](https://en.wikipedia.org/wiki/Relational_algebra)] 2. SQL (Wikipedia Entry) [[link](http://en.wikipedia.org/wiki/Sql)] 3. Relational Algebra I Video [[link](http://www.youtube.com/watch?v=3Xu_LWK3SWw)] 4. Relational Algebra II Video [[link](http://www.youtube.com/watch?v=auMkHSJdW2Q)] 5. Introduction to SQL Video [[link](http://www.youtube.com/watch?v=OmYKfFELcGE)] 6. MySQL Tutorial [[link](http://dev.mysql.com/doc/refman/5.7/en/tutorial.html)] 7. Chapters 1&2 in Provost and Fawcett   [LAB 2 SQL and Movies] |
| 3 | 9/15 | API’s, Scrapers, Semantic Data, and SPARQL | 1. The semantic web and social machines [[link](http://www.stanford.edu/class/cs227/Readings/hendler-berners-lee-semantic-web.pdf)] 2. Ontologies [[link](http://bmir.stanford.edu/file_asset/index.php/108/SMI-2001-0880.pdf)] 3. Tutorial, Linked Open Data Using R [[link](http://semanticweb.cs.vu.nl/R/sparql_lop/sparql_lop.html)] 4. REST API (Wikipedia) [[link](http://en.wikipedia.org/wiki/REST_API)] 5. Twitter Rest API [[link](https://dev.twitter.com/docs/api/1.1)] 6. Tutorial, web scraping with beautiful soup [[link](http://www.pythonforbeginners.com/python-on-the-web/web-scraping-with-beautifulsoup/)] |
| 4 | 9/22 | Introduction to Predictive Modeling | 1. Google’s R Class [[link](http://www.r-bloggers.com/google-developers-r-programming-video-lectures/)] 2. Chapters 3, 4 in Provost and Fawcett |
| 5 | 9/29 | Overfitting and Cross Validation | 1. Overfitting [[link](http://blog.kaggle.com/2012/07/06/the-dangers-of-overfitting-psychopathy-post-mortem/)] 2. Chapter 5 Provost and Fawcett |
| 6 | 10/6 | Similarity, Nearest Neighbors, Clusters, and Text Mining | 1. Chapter 6 Provost and Fawcett 2. Chapter 10 Provost and Fawcett |
|  | 10/13 | Columbus Day – No Classes | |
| 7 | 10/14 | Midterm | 1. Short Midterm with project presentations.   [Kaggle 1 Assignment due]. |
| 8 | 10/20 | Data Mining and Machine Learning Based Analyses | 1. A few Useful things to know about Machine Learning [[link](http://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf)] 2. Top 10 Algorithms in Data Mining [[link](http://www.cs.uvm.edu/~icdm/algorithms/10Algorithms-08.pdf)] 3. Chapter 9 Provost and Fawcett |
| 9 | 10/27 | Intro to Big Data | 1. Planning for Big Data [O’Reilly Media]   Mining Massive Datasets (Chapters   1. “Data Jujitsu” DJ Patil 2. Ullman, Rajaraman, [Mining of Massive Datasets](http://infolab.stanford.edu/~ullman/mmds.html), Chapters 1. |
| 10 | 11/3 | Big Data With Hadoop | 1. Ullman, Rajaraman, [Mining of Massive Datasets](http://infolab.stanford.edu/~ullman/mmds.html), Chapters 2 |
| 11 | 11/10 | Visualization (using Tableau) | 1. Hans Rosling, The Joy of Stats [[link](http://www.gapminder.org/videos/the-joy-of-stats/)] 2. Pat Hanaran, Tools for Data Enthusiasts [[link](http://vimeo.com/50723101)] 3. Jeffrey Heer, Michael Bostock, Vadim Ogievetsky, A Tour through the Visualization Zoo, Communications of the ACM, Volume 53 Issue 6, June 2010 [[link](http://queue.acm.org/detail.cfm?id=1805128)] 4. Chapter 8 Provost and Fawcett |
| 12 | 11/17 | Applied Analytics in Business | 1. “Building Data Science Teams” by DJ Patil [file] 2. Case Study [TBD] 3. Chapter 7, 11, 13 Provost and Fawcett |
| 13 | 11/24 | Final Project Work | Guest speaker. Short presentation. |
| 14 | 12/1 | Final Exam |  |

**COURSE POLICIES**

These will be designated by the instructor in advance to the student and will deal with the following issues.

* + Lateness (and penalties, if any, for lateness/absences)

• Class participation: If participation will be graded, explain how it will be measured/graded and how participation will be graded

• Attendance

• Missed exams or assignments – the syllabus should inform the student if assignments and exams can be made up

• Extra credit policies (if any)

• Lab safety/health - This should include reference to safety manuals and procedures that are available to all research personnel at Rensselaer.

* + Incomplete grades, Withdrawal from course

• Use of Wireless Devices- This should include a statement regarding use of electronic devices (cell phones, laptops, iPads, etc.) during class, or during exams (e.g. “any interaction with a wireless device (cell phone, iPad, etc. ) during an exam (except when explicitly indicated otherwise by the instructor) will be considered an illicit data exchange an will result in a zero for the entire exam”)

**ATTENDANCE POLICY**

This course employs case evaluations along with a textbook that includes “labs” or mini-cases. The success of case learning experiences depends upon the engagement of everyone involved. Active, goal-oriented contributions to discussions are important! Labs are designed to be worked on during class time in a collaborative environment in which you each help one another. We have very different levels of technical expertise in the class, and it is important to work together on the labs to help one another in a classroom environment. **While you may be able to complete the work at home by yourself, in doing so you will be robbing your classmates of your expertise. I will consider helping behaviors during the class as part of the participation component of the class.**

Therefore, the imperative clearly stated: each participant attends class fully prepared, willing and able to offer constructive criticism, provide goal-oriented analytic and synthetic insights, and encourage investigative dialectic. You earn your grade on participation through consistent, daily contribution. Merely "COMING TO CLASS” is not sufficient, but is necessary.

Simply put: Do not miss class hours or group meetings! Understandably, there are circumstances (e.g., job interviews, family matters, extracurricular activity, etc.) that may cause you to miss class; nevertheless, excessive absences will reduce your class participation grade. Notify the instructor and group IN ADVANCE of any planned absences (especially students who participate in extracurricular activities as representatives of RPI.)

As a matter of common courtesy, arrive to class on time (If you are not early, you are late!).

The quality of class participation is far more important than the quantity, and the following points tend to characterize effective ‘attending participation’; they give you guidance to the subjective nature of these evaluation criteria.

• Do comments make substantive and relevant points to the discussion? Do they link to the comments of others?

• Do comments show the participant has been listening?

• Do comments clarify and highlight the important aspects of earlier comments and lead to a clearer statement of the concept under consideration?

• Is the participant willing to interact with other class members?

• Do comments show evidence of analysis, synthesis, or integration?

• Do comments add to our understanding of the situation?

• Does the participant distinguish among different kinds of data (i.e., facts, opinions, beliefs, etc.)?

**ACADEMIC INTEGRITY**

Student-teacher relationships are built on trust. For example, students must trust that teachers have made appropriate decisions about the structure and content of the courses they teach, and teachers must trust that the assignments that students turn in are their own. Acts, which violate this trust, undermine the educational process. The Rensselaer Handbook of Student Rights and Responsibilities and the Graduate Student Supplement define various forms of Academic Dishonesty and you should make yourself familiar with these. In this class, all assignments that are turned in for a grade must represent the student’s own work. In cases where help was received, or teamwork was allowed, a notation on the assignment should indicate your collaboration.  A grade of zero will be given on the first assignment where a violation is detected, and the Associate Dean of Academic Affairs will be notified, in accordance with Lally’s ‘three strikes’ policy on Academic Dishonesty. If there is a subsequent infraction the student will receive a grade of F for the course.

**Grading :**

Student tests and individual exams are graded by specific points and then weighted according to the overall percentages indicated earlier. Students will have access to all grades through the learning management system. While the final grades will be given by points and percentages, the instructor reserves the right to curve final overall grades as appropriate.

**Grade Percentage GPA**

A 93-100 4.00

A− 90-92 3.67

B+ 87-89 3.33

B 83-86 3.0

B− 80-82 2.67

C+ 77-79 2.33

C 70-76 2.0

D 60-69 1.0

F 0-59 0.0