

# SYLLABUS

**Technology Fundamentals for Business Analytics**

**MGMT6963**

**3 Credits**

**Fall/2015 Monday 6:00 – 8:50**

**Room Location: Pittsburgh Building 4114**

**Websites:** [**https://jkuruzovich.github.io/tech-fundamentals-analytics/**](https://jkuruzovich.github.io/tech-fundamentals-analytics/)

**& Blackboard**

**Prerequisites or Other Requirements**: None.

***INSTRUCTOR***

**Instructor Name:** Jason Kuruzovich **Office Location:** Pitt. 4th Floor (Sev. Center)

**Tel. No.:** 518-698-9910 **Email Address:** kuruzj@rpi.edu

**Office Hours:** Thusday 10:00 AM-12:00 PM

***TEACHING ASSISTANT***

**TA Name(s):** TBD **TA Office Location:**

**TA Office Hours:** **TA(s) Email Address:**

## 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.

## COURSE GOALS/OBJECTIVES

The goal of this course will be to provide the technical foundation to enable students to become data scientists.

## STUDENT LEARNING OUTCOMES

Through this course you are expected to:

1. Demonstrate an understanding of analytics based problem solving and analytics thinking.
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 REFERENCE MATERIALS

There is no text book but students may find that the following freely available texts are useful:

**R for Data Science** by Garrett Grolemund & Hadley Wickham

**(**<http://r4ds.had.co.nz)>

**Think Python 2e** by Allen B. Downey

(<http://greenteapress.com/thinkpython2/thinkpython2.pdf>)

## COURSE ASSESSMENT MEASURES

**Class Participation and In-Class Assignments**

Participation in class is absolutely necessary. I expect this to be a hands on class where you will be actively working on problems, demonstrating solutions, and helping your colleagues. This may also include in-class assignments designed to push absorption of the material.

**Assignments (Weekly)**

You will have assignments associated with many classes. Be sure to check the course website and submit assignments via 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/Organizational Assignment (Due 14th Class)**

There will be an individual data assignment to be completed as part of the class. The overall goal is to undergo all stages of the data understanding, data munging, modelling, and further development of understanding.

**For complete details, see the course web page.**

**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 and Quizzes | 30% |
| Kaggle/Organizational Assignment (Due 14th Class) | 30% |
| Midterm (7th Class) | 20% |
| Final Exam | 20% |

## ATTENDANCE POLICY

A maximum of 2 unexcused absences are allowed. After that 10% of the overall participation grade will be removed for each missed class.

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 behaviours 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.)

## OTHER COURSE POLICIES

After the add deadline, 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.

## 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 that violate this trust undermine the educational process. The Rensselaer Handbook of Student Rights and Responsibilities and the Graduate Student Supplement (For Graduate courses) 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. Submission of any assignment that is in violation of this policy will result in a ***grade of zero will be given on the first assignment where a violation is detected, and the infraction will be reported to the Associate Dean for Academic Affairs. If there is a subsequent infraction the student will receive a grade of F for the course*.**

If you have any question concerning this policy before submitting an assignment, please ask for clarification.

## COURSE CALENDAR

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Date** | **Topic** | **Overview** |
| 1 | 29-Aug | The Rise of Analytics | [Click Here](https://gist.github.com/jkuruzovich/672945f488de09dee47010a6f3f343ba) |
|  | 5-Sep | Labor Day |  |
| 2 | 12-Sep | Intro to Python | [Click Here](https://gist.github.com/jkuruzovich/6cc69cda7778fa58c66ac3da0bd2ce39) |
| 3 | 19-Sep | Intro to Python (continued) | [Click Here](https://gist.github.com/jkuruzovich/f55f10c4e8dee3e137fec4be06c58e91) |
| 4 | 26-Sep | Intro to R | [Click Here](https://gist.github.com/jkuruzovich/8190532832731f26252a43abd4a43d7d) |
| 5 | 3-Oct | Intro to R (continued) | [Click Here](https://gist.github.com/jkuruzovich/aa5c5dbd1a8030f0795c6155427447a6) |
|  | 10-Oct | Columbus Day - No Class |  |
| 6 | 11-Oct | Intro to Visualization | [Click Here](https://gist.github.com/jkuruzovich/6072f9b59cda03d5303e31b47553f360) |
| 7 | 17-Oct | Data Munging | [Click Here](https://gist.github.com/jkuruzovich/c2a8e97cd697abb0c61f34862a9c0a27) |
| 8 | 24-Oct | Midterm | [Click Here](https://gist.github.com/jkuruzovich/ef2f0c9f56087d13803a6b64f97dcbbe) |
| 9 | 31-Oct | Introduction to Modeling I | [Click Here](https://gist.github.com/jkuruzovich/bedf36300cec584aec47d3873ec361f8) |
| 10 | 7-Nov | Introduction to Modeling II | [Click Here](https://gist.github.com/jkuruzovich/65268fc7d45519c828ec4a5ae0ae2a73) |
| 11 | 14-Nov | Text Mining and Unstructured Data | [Click Here](https://gist.github.com/jkuruzovich/a78ca939dea78fe0893e258178f6d96b) |
| 12 | 21-Nov | Image Processing | [Click Here](https://gist.github.com/jkuruzovich/9b6a0d89d9edb7bc0f8b243cb5143b84) |
| 13 | 28-Nov | Introduction to Big Data with Spark | [Click Here](https://gist.github.com/jkuruzovich/936d214689c6077bc11bb35e775afca5) |
| 14 | 5-Dec | Presentations |  |

**[ALWAYS CHECK BLACKBOARD FOR LATEST ASSIGNMENTS]**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Date** | **Topic** | **Readings/Notes** |
| 1 | 8/29 | The Rise of Analytics | **Modules (**   1. Join Analytics Dojo 2. Analytics Platform 3. Rise of Analytics |
|  | 9/5 | Labor Day |  |
| 2 | 9/12 | Intro to Python | **Readings**  **Modules (During Class)**   1. Concepts – Intro to Programming 2. Python – Overview and Packages 3. Python – Data structures Intro 4. Python – Numpy 5. Python - Pandas |
| 3 | 9/19 | Intro to Python | 1. Python – Accessing Files 2. Python – Logic and Loops 3. Python – Functions and Classes 4. Python – Filtering, Merging, Splitting, Aggregating Data 5. Python – Basic Visualizations |
| 4 | 9/26 | Intro to R | 1. R – Overview and Packages 2. R – Data Structures |
| 5 | 10/3 | Intro to R | 1. R - Accessing Files 2. R – Logic and Loops   R – Functions   1. R – Filtering, Merging, Splitting, Aggregating Data |
|  | 10/10 | Columbus Day - No Class | Data Science for Business Chapters 2-3  Columbus Day/Class Tuesday Rather than Monday |
| 6 | 10/11 | Intro to Visualization with R and Python | 1. Introduction to Visualization with R and Python Tableau |
| 7 | 10/17 | DataMunging | 1. Python – Missing Data 2. Python – Feature Creation 3. Python – Cross Validation 4. R – Missing Data   R – Feature Creation |
| 8 | 10/24 | Midterm | 1. Intro to Model Evaluation |
| 9 | 10/31 | Introduction to Modeling I | 1. R – Classification and Regression 2. R – Cross Validation |
| 10 | 11/7 | Introduction to Modeling II | 1. Python – Classification and Regression 2. Python – Cross Validation |
| 11 | 11/14 | Text Mining and Unstructured Data | R – Introduction to Text Mining |
| 12 | 11/21 | Image Processing |  |
| 13 | 11/28 | Introduction to Big Data with Spark | 1. R – Introduction to Spark |
| 14 | 12/5 | Presentations |  |

<http://www.nytimes.com/interactive/2016/upshot/presidential-polls-forecast.html?_r=1>

<http://www.scipy-lectures.org/intro/numpy/array_object.html>

Books:

**R for Data Science**

**Garrett Grolemund**

**Hadley Wickham**

<http://r4ds.had.co.nz>

Open Tech School

<http://opentechschool.github.io/python-data-intro/core/data.html>

[Advanced R](http://adv-r.had.co.nz/) by Hadley Wickham

<http://adv-r.had.co.nz>

Think Python

<http://greenteapress.com/thinkpython2/thinkpython2.pdf>

Conceptual

Tidy Data

https://github.com/hadley/tidy-data

Cases:

Classification and the Federalist Papers.

<https://github.com/akavka/federalistPapers>

Titanic

http://trevorstephens.com/kaggle-titanic-tutorial/

https://github.com/trevorstephens/titanic

Random Acts of Pizza

Rossman Store Sales.

Apply lambda function

<https://www.kaggle.com/omarelgabry/rossmann-store-sales/a-journey-through-rossmann-stores>

Expedia

Kaggle Pandas Example.

https://www.dataquest.io/blog/kaggle-tutorial/