

# Projects in Big Data Analytics

ISYE 4961/6961

TF 12:00-1:20 Lally 02

# Tom Morgan

- Cogswell 312
- Hours: 10-noon TF
- (office) 518-276-3887
- (cell) 281-753-1383
- Email: [morgat5@rpi.edu](mailto:morgat5@rpi.edu)

# John Erickson

- CBIS 3153, Winslow or skype
- Hours: TBD
- (office) 518-276-4384
- Email: [erickj4@rpi.edu](mailto:erickj4@rpi.edu)
- Skype: olyerickson

# Course Timeline

- Introduction and Organization
- Lectures: Data Science and Data Analytics Background
- Project 1
- Midterm Exam
- Lectures, Project 2
- Lectures, Project 3
- Final Exam

# Course Content

- Introduction
- Analytics Methods:
  - Single Channel – Univariate
  - Noise
  - Missing data
  - Multi-channel – Multivariate
  - Modeling and prediction
- Projects:
  - Acquisition and Preparation Focus
  - Univariate Focus
  - Multivariate Focus
  - Text Data Focus

# Class Period Structure

- Questions, issues, answers
- Lecture (usually, but not always)
- Hands-on project work

# Data Analytics Process

- Acquisition
- Preparation
- Analysis
- Interpretation
- Validation

# What's Missing?



What's Missing?

Problem Formulation!

# Course Objectives and Focus

- Identify and formulate data analytics problems and solution methods.
- Evaluate data sets for completeness, utility for a given purpose, and presence of errors, ambiguities and other inconsistent content.
- Identify and correct data set flaws and deficiencies, and otherwise prepare data sets for use in various analytical routines.
- Apply quantitative analyses, interpret and validate the results.

# Projects

- A means rather than an end, no specific end point, learning experiences will be the focus
- Each one will be used to present lessons in one or more of the process areas
- Will be done in teams, but individual reports will be submitted and presented
- Reports will be PDF slides

# Student Background Poll

- Computer Science / Math
- Domain experts
- Team selection

# R web site

<https://www.r-project.org/>

- Download and install
- Download manuals
  - Introduction
  - Data
  - Admin
  - Extensions