

PAI 730: Data-Driven Management in Public Organizations

Fall Term, 2015

Tuesday / Thursday, 2:00 pm – 3:20 pm

Room 017, Crouse Hinds

Instructor: Jesse D. Lecy (jdlecy@syr.edu)

Office Hours: Monday 2pm – 3pm

Office #320, Eggers Hall

Prerequisites: None

Philosophy of Learning

“With very few exceptions, there is no shortcut between not knowing something and knowing it. There is a beauty to awkwardness, a wisdom in the wobble.”

~Maya Stein

“Nobody tells this to people who are beginners, I wish someone told me. All of us who do creative work, we get into it because we have good taste. But there is this gap. For the first couple years you make stuff, it’s just not that good. It’s trying to be good, it has potential, but it’s not. But your taste, the thing that got you into the game, is still killer. And your taste is why your work disappoints you. A lot of people never get past this phase, they quit. Most people I know who do interesting, creative work went through years of this. We know our work doesn’t have this special thing that we want it to have. We all go through this. And if you are just starting out or you are still in this phase, you gotta know its normal and the most important thing you can do is do a lot of work. Put yourself on a deadline so that every week you will finish one story. It is only by going through a volume of work that you will close that gap, and your work will be as good as your ambitions. And I took longer to figure out how to do this than anyone I’ve ever met. It’s gonna take awhile. It’s normal to take awhile. You’ve just gotta fight your way through.”

~ Ira Glass

Course Description

This course introduces students to the field of performance management and the practice of data programming. It is a practical, tools-based course designed to build the foundations for strong data analysis as part of the performance management process. We will cover basic data operations, graphics, programming fundamentals, text analysis, and the creation of user interface. As part of the class students will work in a team to build a model performance management system that includes a data collection process, an analytical framework, and a dashboard to report key performance indicators to stakeholders.

Learning Objectives

This course builds capacity:

- 1) To understand the importance of key performance indicators:** Students will discuss several case studies that demonstrate the creation and utilization of key performance indicators and how they are used to improve organizational performance.
- 2) To become proficient in the R data programming language:** R is a popular engine for data analytics used by academics, government, and corporations. Students will leave the course with a solid foundation in data programming.
- 3) To gain experience with technical project management tools:** Managers will often have to collaborate with a technical team to do things like build a database, create a mobile apps, create a website, or build a performance management system. You will have exposure to common tools used for social coding, versioning, and project workflow in order to gain some exposure to project management in the technology field.

Course Materials

There are **two required texts** for class. The first text, *R Cookbook*, is required. The second, *The Art of R Programming*, is highly-recommended but will be treated as supplementary material.

Teetor, P. (2011). *R Cookbook*. O'Reilly Media, Inc. \$27 on Amazon

Matloff, N. (2011). *The Art of R Programming*. No Starch Press. \$24 on Amazon

Additional readings will be posted through Blackboard or provided in class.

Assignments and Tests

The class is an elective, and as such it is designed to give students a gentle introduction to data programming. We will move quickly and cover a lot of material in the class, but assignments are designed to be incremental with the knowledge that most students have not had experience programming, nor do

they intend to be a computer scientists. Learning a computer language is like learning a natural language or a musical instrument – it is awkward at first and it takes time to master the fundamentals. This class aims to open the door into an exciting new world of data science and remove the initial barriers that stand in the way of self-learning in an unfamiliar field. The jargon and syntax will be awkward at first, but once you become comfortable with the conventions of R you will find that it is feasible to teach yourself many tools (there are currently over 5,000 free packages available). Understanding the jargon allows you to access hugely helpful online resources like Stack Overflow, and familiarity with basic conventions will allow you to work with new programs in unfamiliar domains.

Labs

Each week you will be assigned a lab to give you practice on the material that we cover in the readings and lectures. They will include toy programming problems that provide practical examples of the topic for the week. **Labs are due at the beginning of class on Tuesday.** Since we will go through solutions in class, late labs will not be accepted. The labs are graded pass / fail.

You will generate your solutions for the labs using the R Markdown tools available in R Studio. You need to submit files in either a webpage complete HTML format, or as a PDF. Be sure to properly separate and format code and text in the files using the “fences” convention presented in the tutorial.

Quizzes

There will be three short quizzes covering the material from lectures and labs. Each quiz is worth the same number of points as a lab, so they are not weighted heavily in the grade. They are designed as an opportunity to review and consolidate the material that you have learned as part of the labs. Quizzes are administered in-class on the following dates:

Tuesday, October 6

Tuesday, November 3

Thursday, November 19

Case Studies

Each week you will be responsible for submitting a short 1-2 page summary of a case study. The questions for each case study are listed on the schedule portion of the syllabus. These summaries are graded pass / fail.

Final Project

For the final project students will work in teams to build a dashboard using the R Shiny web application framework. Each team will present their dashboard during the final exam period. The team will be required to use a project workflow process (Trello) and a versioning system (GitHub) while building the deliverable in order to gain familiarity with these common tools that are used by most software teams in industry.

Grades

Grades will be comprised of points from labs, quizzes, case studies, and the final project weighted as follows:

Labs (10 total):	50%
Quizzes (3 total):	15%
Case Studies (13 total):	15%
Final Project:	20%

This course uses plus-minus grading. Grades will be assigned roughly as follows with room for discretion by the instructor:

A	=	95-100
A-	=	90-94
B+	=	86-89
B	=	83-85
B-	=	80-82
C+	=	77-79
C	=	73-76
C-	=	70-72
D	=	60-69
F	=	0-59

Late assignment policy:

No late assignments will be accepted. I will drop your lowest homework or quiz grade.

Course Policies

Academic Integrity and Plagiarism

Syracuse University takes academic integrity very seriously. If you are unfamiliar with the policies they can be found at: http://supolicies.syr.edu/ethics/acad_integrity.htm

Plagiarism will result in an “F” for the course and students may be sent before the academic integrity board. In short, it is not worth it.

Disability Statement

If you believe that you need accommodations for a disability, please contact the Office of Disability Services (ODS), <http://disabilityservices.syr.edu>, located in Room 309 of 804 University Avenue, or call (315) 443-4498 for an appointment to discuss your needs and the process for requesting accommodations. ODS is responsible for coordinating disability-related accommodations and will issue

students with documented disabilities Accommodation Authorization Letters, as appropriate. Since accommodations may require early planning and generally are not provided retroactively, please contact ODS as soon as possible.

Syracuse University and I are committed to your success and to supporting Section 504 of the Rehabilitation Act of 1973. This means that in general no individual who is otherwise qualified shall be excluded from participation in, be denied benefits of, or be subjected to discrimination under any program or activity, solely by reason of having a disability.

You are also welcome to contact me privately to discuss your academic needs. Although I cannot arrange for disability-related accommodations myself, I will help you work with the Office of Disability Services at Syracuse University to support your needs. Please contact the ODS by email in advance: odssched@syr.edu

Attendance

Attendance in the course is required but is not taken and there is no grade assigned for attendance. Lack of attendance, however, can affect your final grade.

Computers and Mobile Devices in Class

The policy is to be determined by the class on the first day.

Classroom Etiquette

Treat each other with respect. Act in a way that creates an intellectually curious environment. Honor intellectual integrity of your classmates, but also be sensitive to views and opinions that are offensive. Minimize disruptions in class, especially tardiness, talking, eating, and text messaging. I tend to enforce the spirit of the law, not the letter.

Office Hours

Office hours will be negotiated during the first week of class.

Semester Schedule

Week 1 (Sept 1 & 3) – Introduction to Data-Driven Management.....Install RStudio

TOPICS: functions, packages, help files, R Studio

Course Background:

https://www.ted.com/talks/anne_milgram_why_smart_statistics_are_the_key_to_fighting_crime

US Digital Service: <https://www.whitehouse.gov/digital/united-states-digital-service>

“Data Analysts Captivated by R’s Power” New York Times, January 6, 2009.

Required Reading:

R Cookbook, CH1 Getting Started and Getting Help

The Art of R Programming: CH1 Getting Started

RMarkdown Cheatsheet

Optional:

“Data Crunchers Now the Cool Kids on Campus”, Wall Street Journal

<https://blog.javascripting.com/2014/07/28/fastest-growing-new-languages-on-github-are-r-rust-and-typescript-and-swift/>

<http://www.sitepoint.com/best-programming-language-learn-2015-job-demand-salaries/>

“Why I Joined the Digital Services”

“2016 Budget Will Fund 25 More Digital Services Teams”

Week 2 (Sept 8 & 10) – The Structure of Data.....Lab #1 + Summary Due

TOPICS: data types, classes, logical operators, subsets

Required Reading:

R Cookbook: CH5 (5.1-5.5, 5.20-5.28)

The Art of R Programming: 1.4 Preview Important Data Structures, CH2 Vectors, CH5 Data Frames, 6.1 Factors and Levels

R Style Guide

Case Study:

What is the purpose of a data-driven performance review? What are three important considerations for implementing an effective system?

Hatry, H., and Davies, E. (2011). A Guide to Using Data-Driven Performance Reviews. IBM Center for the Business of Government. Overview of the framework: pp 8-32.

<http://www.businessofgovernment.org/sites/default/files/A%20Guide%20to%20Data-Driven%20Performance%20Reviews.pdf>

Kemenski, J. "Unleashing Evidence-Based Management in Government." Governing: March 23, 2015.

<http://www.governing.com/blogs/bfc/col-new-jersey-child-welfare-data-management.html>

Week 3 (Sept 15 & 17) – Merging Data.....Lab #2 + Summary Due

TOPICS: relational databases, sets (union, intersection, difference), merge, match

Required Reading:

R Cookbook: 5.29-5.30

The Art of R Programming: 5.3 Merging Data Frames

Relational Databases: <http://computer.howstuffworks.com/question599.htm>

Case Study:

How did Rudder use data to change our understanding (or maybe confirm our hunches) about dating behaviors?

Matlin, C. "Matchmaker, Matchmaker, Make Me a Spreadsheet: How Data Saved OK Cupid." FiveThirtyEight: Sept 9, 2014. <http://fivethirtyeight.com/features/christian-rudder-dataclysm-okcupid/>

Bridgeland, J. and Orszag, P. "Can Government Play Moneyball?" The Atlantic: July/August 2013.

<http://blog.ted.com/the-moneyball-effect-how-smart-data-is-transforming-criminal-justice-healthcare-music-and-even-government-spending/>

Week 4 (Sept 22 & 24) – Describing Your Data.....Lab #3 + Summary Due

TOPICS: descriptive statistics, tables, summary by group (tapply)

Required Readings:

R Cookbook: CH6, skim CH9

The Art of R Programming: 5.4 Applying Functions to Data Frames, 6.2-6.2 Common Factor Functions

Case Study:

Describe the Apgar score. How has it changed the practice of childbirth in hospitals?

Gawande, A. "The Score." The New Yorker: October 9, 2006.

Week 5 (Sept 29 & Oct 1) – Data Input.....Lab #4 + Summary Due

Required Reading:

R Cookbook: CH4 Input and Output

The Art of R Programming: CH10 Input / Output

Case Study:

Summarize one example of how Ben Wellington has used open data to change public policy:

What is the Data Act? <http://www.datacoalition.org/what-is-data-transparency/data-act/>

Ben Wellington's Blog: I Quant NY: <http://iquantny.tumblr.com/>

https://www.ted.com/talks/ben_wellington_how_we_found_the_worst_place_to_park_in_new_york_city_using_big_data?

<https://nycopendata.socrata.com/>

<http://labs.data.gov/dashboard/offices>

<http://www.forbes.com/sites/teconomy/2014/09/12/how-open-data-is-transforming-city-life/>

Week 6 (Oct 6 & 8) – Principles of Visualization.....Quiz #1 + Summary Due

TOPICS: visual hierarchy and narrative **Tuesday 10/06 – QUIZ**

Required Content:

<https://www.youtube.com/watch?v=AdSZJzb-aX8>

<http://ed.ted.com/lessons/david-mccandless-the-beauty-of-data-visualization>

<https://www.youtube.com/watch?v=vc1bq0qIKoA>

<http://ed.ted.com/lessons/the-good-news-of-the-decade-hans-rosling>

<http://drewconway.com/zia/2013/3/26/what-data-visualization-should-do-simple-small-truth>

<http://chartsnthings.tumblr.com/>

<http://www.r-bloggers.com/amanda-cox-on-how-the-new-york-times-graphics-department-uses-r/>

Case Study:

How did Paul O'Neill select a key performance indicator for his organization? Although it was not an obvious choice, why was it a good choice? What systems (technological and managerial) were put in place to collect and utilize the performance data? What factors were responsible for its success?

Duhigg, C. (2012). *The power of habit: Why we do what we do in life and business* (Vol. 34, No. 10). Random House.
CH 4: The Ballad of Paul O'Neill.

Week 7 (Oct 13 & 15) – Simple Graphics.....Lab #5 + Summary Due

TOPICS: plot function, points(), lines(), text(), ablines()

Required Reading:

R Cookbook: 10.1-10.24 Graphing Functions

The Art of R Programming: 12.1 Creating Graphs

Case Study:

What key performance indicators have been used by the examples cited in the case studies?

Hatry, H., and Davies, E. (2011). A Guide to Using Data-Driven Performance Reviews. IBM Center for the Business of Government. Case studies: pp 33-40.

<http://www.businessofgovernment.org/sites/default/files/A%20Guide%20to%20Data-Driven%20Performance%20Reviews.pdf>

Week 8 (Oct 20 & 22) – Advanced Graphics.....Lab #6 + Summary Due

TOPICS: colors, axes, labels, context, gifs

Required Reading:

R Cookbook: 10.25-10.29 Graphing Functions

The Art of R Programming: 12.2 Customizing Graphs

Case Studies:

Is measuring performance on a specific set of indicators the best way to induce productive behaviors of people in organizations?

<https://www.youtube.com/watch?v=avnHUXSVfVM>

Levitt, S. (2005). Do This, Get That: Making Sense of Incentives. *Associations Now*.

Optional

Levitt, S. D., & Neckermann, S. (2014). What field experiments have and have not taught us about managing workers. *Oxford Review of Economic Policy*, 30(4), 639-657.

Kerr, S. (1995). "On the Folly of Rewarding A, While Hoping for B." *Academy of Management Executive*, 9(1): 7-14.

Week 9 (Oct 27 & 29) – Maps and GIS.....Lab #7 + Summary Due

TOPICS: shapefiles, map APIs, choropleth maps

Required Reading:

TBA

Case Study:

How do people respond to the measurement of their productivity? Are there certain types of tasks where pay-for-performance works better than others? Are there unintended consequences or perverse incentives that result from these schemes?

Lazear, E. P. (1996). *Performance pay and productivity* (No. w5672). National bureau of economic research.

Hong, F., Hossain, T., List, J. A., & Tanaka, M. (2013). *Testing the Theory of Multitasking: Evidence from a Natural Field Experiment in Chinese Factories* (No. w19660). National Bureau of Economic Research.

Week 10 (Nov 3 & 5) – Basic Programming.....Quiz #2 + Summary Due

TOPICS: loops, simulation, debugging **Tuesday 11/03 – QUIZ**

Required Reading:

A Beginner's Guide to R: CH6 Loops and Functions

The Art of R Programming: CH7 Programming Structures, 8.6 Simulations in R, CH13 Debugging

Case Study:

We think of incentives primarily on an individual basis, even though much of the work in organizations occurs in teams. What is an example of a social incentive? Is there any evidence that they work better than individual incentives?

Pentland, A. (2014). *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. Penguin. CH 4: Engagement

Week 11 (Nov 10 & 12) – Text Analysis.....Lab #8 + Summary Due

TOPICS: letters, strings, find, replace

Required Reading:

R Cookbook: CH7 Strings and Dates

The Art of R Programming: CH11 String Manipulation

Case Study:

How do organizations change as a result of performance management systems?

Heinrich, C.H. (2010). "Incentives and Their Dynamics in Public Sector Performance Management Systems." *Journal of Policy Analysis and Management*, 29(1): 183-208.

Week 12 (Nov 17 & 19) – Text Analysis..... **Quiz #3 + Summary Due**

TOPICS: fuzzy matching, regular expressions ****Thursday 11/19 – QUIZ****

Required Reading:

The Art of R Programming: 11.2 Regular Expressions

Case Study:

Identify one data dashboard that is used by a community or an organization. What are the key performance indicators? Where does the data come from? Who are the key stakeholders? How is the dashboard used? Submit your report as PowerPoint slides.

[https://en.wikipedia.org/wiki/Dashboard_\(business\)](https://en.wikipedia.org/wiki/Dashboard_(business))

Week 13 (Nov 24 & 26) – TG Break..... **Tofurky Turducken**

Week 14 (Dec 1 & 3) – Dynamic Reporting..... **Lab #9 + Summary Due**

TOPICS: R Shiny, social coding and versioning

Required Reading:

Shiny Tutorial: <http://shiny.rstudio.com/tutorial/>

Version Control: <http://git-scm.com/video/what-is-version-control>

Skim: https://en.wikipedia.org/wiki/Revision_control#Graph_structure

Skim: <http://nvie.com/posts/a-successful-git-branching-model/>

Case Study:

Write a one-page memo to the mayor of Syracuse outlining the benefits and costs of implementing a CityStat system.

Behn, R.D. (2007). What All Mayors Would Like to Know about Baltimore's CitiStat Performance Strategy. IBM Center for the Business of Government, Washington, DC.

<http://web.pdx.edu/~stipakb/download/PerfMeasures/CitiStatPerformanceStrategy.pdf>

Supplementary:

Behn, R.D. (2006). "The Varieties of CitiStat." *Public Administration Review*, 66(3): 332-340.

Week 15 (Dec 8 & 10) – The Dashboard.....Lab #10 + Summary Due

TOPICS: R Shiny, user interface design

Case Study:

What is the GPRA Modernization Act? What are some reasons that the original Modernization Act failed to achieve adoption? What are common factors for successful implementation in the crime and homelessness examples?

Moynihan, D. (2013). *The New Federal Performance System: Implementing the GPRA Modernization Act*. IBM Center for the Business of Government, Wash., DC.

Final Exam (Presentation of Team Dashboards): Fri, Dec 18, 8:00am - 10:00am

*This date and time may change