Worcester Polytechnic Institute

Instructor: Narahara Chari Dingari, PhD
Office: Office hours available on-line on request
Cell: (401) 767-7769
E-mail: dnchari@gmail.com or ncdingari@wpi.edu
Skype: dnchari

Course Description

The course on Introduction to Data Science provides an overview of Data Science, covering a broad selection of key challenges in and methodologies for working with big data. Topics to be covered include data collection, integration, management, modeling, analysis, visualization, prediction and informed decision making, as well as data security and data privacy. This introductory course is integrative across the core disciplines of Data Science, including databases, data warehousing, statistics, data mining, data visualization, high performance computing, cloud computing, and business intelligence. Professional skills, such as communication, presentation, and storytelling with data, will be fostered. Students will acquire a working knowledge of data science through hands-on projects and case studies in a variety of business, engineering, social sciences, or life sciences domains. Issues of ethics, leadership, and teamwork are highlighted.

Course Objectives

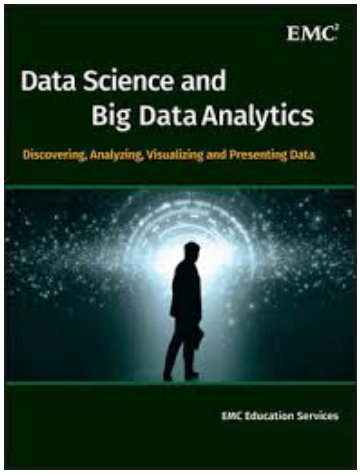
After taking this course, the student will be able to:

1) Write programs in R
2) Understand problems solvable with data science and able to attack those problems from a statistical perspective
3) Collect, Manipulate, Blend Data from Different Data Sources
4) Visualize Data and Perform Exploratory Data Analysis
5) Understand Data Science Project Lifecycle
6) Create Data Products for Business Applications
7) Get the basics of Unsupervised Machine Learning Techniques
8) Understand the basics of Supervised Regression Techniques
9) Understand the basics of Supervised Classification Techniques
10) Get introduced to Big Data Architecture

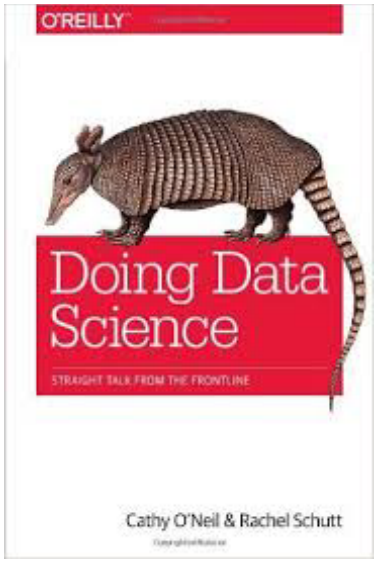
Worcester Polytechnic Institute

Required Textbook and Other Resources

- (Optional) Data Science and Big Data Analytics by Wiley



- (Optional) Doing Data Science by Oreilly



- (Mandatory) R and R Studio
  - https://cloud.r-project.org
  - https://www.rstudio.com/products/rstudio/download/

- Guidance for avoiding plagiarism, see: https://www.wpi.edu/academics/library/help/tutorials.html > Introduction to Research Tutorials > Citing Information

Worcester Polytechnic Institute

Course Approach

The online week starts on Thursday, and ends on Wednesday. The course runs for 10 full weeks.

This course is divided into ten modules that are completed over the same number of weeks; each week starts on Monday. There are no required class meeting times for the course; however, you can reach out to me if you have any questions during the office hours mentioned in announcement section and/or schedule an appointment.

To convey the basic content of the course, the instructor:

1) Posts course content in the form of PowerPoint slides and recorded lectures

2) Assigns weekly reading,

3) Posts supplemental material when appropriate,

4) Facilitates weekly online discussions with active student and faculty participation,

5) Provides timely feedback on assignments/home works/case studies.

Each week, students are required to complete:

1) One individual assignment for every week or two weeks depending on the difficulty.
2) All home works and case studies are detailed in Course Schedule Section, all are due EST Thursday.
3) Students demonstrate mastery of the course through individual assignments.

Course Communication

We will use email/online discussion for most class communication. I will be active on Canvas most weekdays and will typically respond to questions and emails within 24hours. Please post your course related questions to the Q and A Discussions so the entire class can benefit from the exchange. To discuss personal or sensitive concerns, please email me at **ndingari@wpi.edu**

There are no set office hours. However I am open to scheduling a virtual live meeting with individuals or teams, by request, at pre-agreed to times. I encourage students to be proactive in getting questions answered through the many available methods.

Individual Assignments

Please note: I try to include important content in the recorded lectures that may not be covered in the written content of the textbooks or the slides. Please be sure to listen to the recorded lectures, and take notes, before you attempt the assignment each week. Assignments range from easy to very difficult. Individual assignments successively build capability and confidence in the course material. Students hand in an individual assignment, every week/two weeks of the course. Individual assignments are due by midnight EST Thursday of the week assigned (more details in Course Schedule section below). Instructor feedback and grades are posted by midnight up to one/two weeks later.

Grading

No credit will be earned on late work unless the student has arranged an extension ahead of time with me (and that is quite possible, I am flexible with everyone's challenging circumstance and time constraints); with rare exceptions based on individual circumstances (inability to communicate with me ahead of time based on an emergency, for example).

All assignments will be anonymously peer reviewed by your classmates, however, I will over ride the comments/grades and my grade will be final. Peer reviewing gives you the opportunity to learn how other students are approaching the problem.

Specific details on the assignments are found in the course Assignments (in the online classroom) and on the Course Schedule (next section).

Type of Assignment	Number of Items	Points Per item	Total Points
Home Works	3	100	300
Case Studies	2	200	400
Final Project: Creating a Data Product	1	300	300
		Total Course Points	1000

TOTAL **100%**

Please note that assignments in this class may be submitted to a web-based anti-plagiarism system, for an evaluation of their originality.

Extra Credit

No extra credit assignments are given.

Worcester Polytechnic Institute

Course Schedule

Week	Primary Topics/ Weekly Lectures	Assignments	Max Points
1	**Lecture 1 – Introduction to Data Science**	*Home Work 1 – Introduction to Data Science*	100
2	**Lecture 2 – Introduction to R**	*Home Work 2 – Introduction to R*	100
3	**Lecture 3 – Data Collection and Data Blending**	*Home Work 3 (Case Study 1) - Collecting Data from Twitter*	200
4	**Lecture 4 – Business Intelligence and Data Warehousing (Guest Lecture by Lurdu Kunireddy, Principal Data Scientist, Deutsche Bank)**		
5	**Lecture 5 - Data Science Project Life Cycle**	*Home Work 4 - Home Work 2 – Introduction to R*	100
6	**Lecture 6 - Data Visualization**	*Home Work 5 (Case Study 2) - Analyzing data from MovieLens*	200
7	**Lecture 7 - Introduction to Machine Learning (Clustering and Principal Component Analysis)**		
8	**Lecture 8 - Introduction to Machine Learning (Association Rules)**	*Home Work 6 (Case Study 3) – Creating a Data Product using Machine Learning Algorithm*	300
9	**Lecture 9 - Introduction to Machine Learning (Regression and Classification)**		
10	**Lecture 10 - Big Data Analytics (Guest Lecture by Anil K Batchu, Big Data Engineer, Intuit)**		

Worcester Polytechnic Institute

SUPPORT AND HELP

Urgent Messages

Any urgent messages about the class will be distributed via e-mail and on the announcements area of the myWPI class site. It is expected that e-mail will be read by each student within one business day after it is posted.

Contacting Me

My availability is a high priority. I return phone calls and e-mail, typically within 24 hours, unless I'm traveling. I am available via e-mail, phone and Skype. When you leave a message, please speak slowly and clearly; make sure you leave your phone number or e-mail.

E-mail: dnchari@gmail.com or ndingari@wpi.edu
Cell: 401-767-7769

Tips for Online Learning

The most important step for being successful in an online graduate course is to create a weekly schedule that accounts for 10-15 hours per week for class work – spread over different days.

- ☐ About one-third of the time is needed for 'in class' work such as:
  - listening to/reading lectures
  - participating in class and online discussions
  - taking assignments

- ☐ The rest of the time is used for 'outside class' work such
  - completing the assigned reading
  - performing research
  - completing individual assignments

Taking the class online saves you commute time but now it's up to you to dedicate enough time, normally 100-150 hours per graduate course, to be successful.

- ☐ Other tips include:
  - logging in the course nearly every day even for a quick check
  - contributing your ideas and perspectives to the discussions
  - reading or listening to other students' ideas and perspectives – that is learning from each other, not just the instructor
  - keeping an open mind about and becoming an advocate for online learning
  - building a support structure of folks who will help you achieve your learning goals

Technical Support

For technical support, please reach out to helpdesk@wpi.edu and they will triage your request/problem to the appropriate team. The Help Desk is also open for phone calls (508-831-5888) in the evenings and weekends during the regular academic year. For more information on their hours and services, please visit: http://www.wpi.edu/+helpdesk.

Worcester Polytechnic Institute

POLICIES

Academic Honesty

Please be familiar with the *Student Guide to Academic Integrity at WPI* that is downloadable from http://www.wpi.edu/offices/policies/studentguide.html. Issues involve copying and pasting text directly from a source, not putting that text in quotes, and not giving credit to the source; collaborating on individual assignments; rephrasing another person's work without citing that work; or turning in work where a good portion of the work is someone else's work even when properly cited. Consequences for violating the Academic Honesty Policy range from earning a zero on the assignment, failing the course, or being suspended or expelled from WPI.

Library Access

Simply access a database, ejournal, or ebook from http://www.wpi.edu/academics/Library/ and login with a WPI username and password (on campus users will not be prompted to login) when prompted.

American Disability Act (ADA) Accommodations

Students who need special accommodations to participate in class should contact me, the instructor, as well as the Student Disabilities Services Office (see http://www.wpi.edu/offices/disabilities/).

Grading Policy

Final course grades are based on the student's performance as follows:

A	Excellent: earned 90 – 100 points
B	Good: earned 80 – 89 points
C	Pass: earned 70 – 79 points
D	Unacceptable for graduate work: earned 60 – 69 points
F	Fail: earned 59 points or below

Course incompletes may be granted if the major part of the course is completed; however, no additional credit can be given for missed class discussions beyond the end of the course. In addition, in the case of an incomplete, the student is responsible for handing in the final work within the WPI required timeframe (currently one year); after this time an incomplete grade changes to a Failing (F) grade.