EMSE 6574: Programming for Analytics - Course Syllabus

Course Information

Name:	EMSE 6574: Programming for Analytics, Section 10
Semester:	Fall 2019
Meeting Time:	TBD
Location:	TBD

Instructor Information

Name:	John Paul Helveston
Campus Address:	Science & Engineering Hall, Office 2830
Campus Phone:	+1 (202) 994-7173
Email:	jph@gwu.edu
Office Hours:	TBD

Course Description

Official GW Bulletin Description

Introduction to programming for data analytics using the R computing language; topics include R, RMarkdown, functions, conditionals, loops, strings, file input/output, data visualization, coding style, efficiency, Monte Carlo methods, webscraping data, and package development.

Unofficial Description

This course provides students with a foundation in programming for data analytics using the R computing language. Emphasis will be on learning how to read, write, test, and debug code to produce comprehensible and reproducible analyses. Throughout the course, students will primarily work on individual programming assignments to help practice coding skills. Students will be assessed through quizzes, exams, and a final project. Teaching will involve interactive lectures and recitations with a lot time spent live coding. At the end of the semester, students will demonstrate mastery of the course's topics by working on a 3-week term project. This course will prepare students for higher level courses in data analytics.

Prerequisites

There are no prerequisites, and no prior programming experience is necessary to succeed in this class.

Learning objectives

Having successfully completed this course, students will be able to:

- Read, write, test, and debug code to produce comprehensible and reproducible analyses.
- Use computational problem solving to solve problems.
- Create visualizations of data.
- Explain the efficiency of algorithms by predicting the Big-O run time of code.
- Design and write an `R` package and deploy it on github.
- Generate fully reproducible reports in `RMarkdown` that contain contain code, equations, visualizations, and narrative text.

Texts / References

All textbooks and references are available for free on the web, though physical copies can be purchased if desired:

- Data Wrangling in R:
  - Wickham, Hadley. "R for Data Science" [free online], [buy on amazon]
- Data Visualization:
  - Healy, Kieran. "Data Visualization: A practical introduction" [free online], [buy on amazon]
  - Wilke, Claus O. "Fundamentals of Data Visualization" [free online], [buy on amazon]
- Handy R cheatsheets:
  - RMarkdown
  - Data wrangling with the `dplyr` library
  - Data visualization with the `ggplot2` library

Schedule

Week	Meeting	Date	Topics	Due
1	1	08/27	Intro to course and software	
1	2	08/29	RMarkdown & Github	
2	3	09/03	Getting familiar with R	Data Camp 1
2	4	09/05	Getting familiar with RMarkdown	
3	5	09/10	Getting familiar with data frames	Data Camp 2
3	6	09/12	Data Wrangling I: The 'verbs' of the `dplyr` library	
4	7	09/17	Data Wrangling II: Grouping, summarizing, & pipes	Data Camp 3
4	8	09/19	Data Wrangling III: Joins, reshaping for tidy data	
5	9	09/24	Fundamentals of Information Visualization	Project Proposals
5	10	09/26	Visualizations I: The grammar of graphics, the `ggplot2` library, plotting points	Data Camp 4
6	11	10/01	Visualizations II: Bars, histograms, & boxplots	Data Camp 5
6	12	10/03	Visualizations III: Facets, labeling & themes	
7	13	10/08	Data input / output & the `here` library	
7	14	10/10	In-class workshop: Getting data from the wild	
8	15	10/15	Dealing with strings: the `stringr` library	
8	16	10/17	Dealing with factors: the `forcats` library	
9	17	10/22	R as a programming language: writing your own functions	
9	18	10/24	R as a programming language: iteration with loops	
10	19	10/29	Spring Break: No Class	
10	20	10/31	Spring Break: No Class	
11	21	11/05		
11	22	11/07		
12	23	11/12		

Week	Meeting	Date	Topics	Due
12	24	11/14		
13	25	11/19		
13	26	11/21		
14	27	11/26		
14	28	11/28		
15	29	12/03		
15	30	12/05		
16	–	12/10	Reading Week Begins	
17	–	12/12	Final Exams Begin	