DATA 650 Syllabus

Communication and Collaboration for Data Science

1 Instructor Information

Instructor:	Dr. Charles Redmond
Office:	Library 408A
Phone:	824-2355
Email:	credmond@mercyhurst.edu

2 Office Hours

MWF	11 AM - 12 AM
TTh	8:15 PM - 9:45 PM

3 Learning Differences

Any student who feels s/he may need an accommodation based on the impact of a disability should contact me privately to discuss your specific needs. Please know that it is the policy of Mercyhurst University that it is the students' responsibility to provide documentation of his/her disability to the director of the Learning Differences Program. Please call the Learning Differences office at 814-824-3017 to coordinate needed accommodations.

4 Course Description

Communication and collaboration are extremely important to the data scientist. In this course, we will be learning tools that will help us make presentations, write reports and even books, and create web pages and dashboards. We will also be learning how to version control our projects and how to collaborate while writing software and documents. For collaboration and version control, we will be using Git and Github. For presentation and writing, we will be using Latex, R markdown, flexdashboard, and htmlwidgets.

During the course, you will have to produce at least one of all of the following: an article, a slide presentation, a multiple chapter book, a webpage, a blog, and a dashboard. You will not, however, need to research for material. I will be giving lessons on various programming and data science topics, and I will then ask you to turn your notes into one of the presentation types mentioned above. As you write, I will be asking you to version control with Git and Github. On

at least one of these projects you will work on a team, and you will use Git and Github for collaboration.

We should also have an exam or two, so we will have a midterm and a final, both of which will be open note.

5 Text and Software

There is no required text for the course, but I highly recommend buying Text Mining with R, by Julia Silge and David Robinson, and also R for Data Science, by Hadley Wickham and Garrett Grolemund. Information about these books and others you can find in the references section at the end of the syllabus.

The Software we need will be available in the lab, and you will be able to download it for free for your own computers. I will explain in detail how to install everything. Here is a list of the software and where it can be found, but don't worry about downloading anything until we come to that point in the course.

R	`https://www.r-project.org/`
RStudio	`https://www.rstudio.com/`
Git	`https://git-scm.com/`
MiKTeX (for Windows users)	`https://miktex.org/`
MacTeX (for Mac users)	`http://www.tug.org/mactex/`
Slack	`https://join.slack.com/t/data650/signup`

6 Course Objectives

After completing this course, students should be able to:

- communicate with Slack;
- write an article in Latex;
- write a slide presentation in Beamer;
- write a book in Latex;
- knit R code into a Latex document;
- knit R code into an R markdown document;
- create a webpage with R markdown;
- blog with Hugo, R markdown, and Github;
- publish a website on Github;
- version control projects with Git and Github;
- collaborate on projects with Git and Github;

- produce dashboards with flexdashboard;
- employ htmlwidgets into webpages and dashboards

7 Tentative Schedule

Below is our tentative schedule. Anything inside square brackets indicates a content topic, and anything outside of brackets indicates an communication and collaboration topic.

Week	**Topic**
Week 1	Slack, Git, Github, and Simple Websites
Week 2	Git, Github, and Blogging [dplyr]
Week 3	Latex Articles and Presentations[word counts]
Week 4	Latex Books [sentiment analysis]
Week 5	[sentiment analysis]
Week 6	[sentiment analysis]
Week 7	htmlwidgets[lubridate]
Week 8	Midterm Practice and Midterm (Midterm is Oct. 16)
Week 9	htmlwidgets[ggplot]
Week 10	htmlwidgets[mapping]
Week 11	htmlwidgets[networks]
Week 12	Collaboration with Github
Week 13	Collaboration with Github[sentiment analysis]
Week 14	Collaboration with Github[sentiment analysis]

8 Grading

We will have a midterm(worth 15% of your grade), and final(worth 15% of your grade), and class participation (worth 70% of your grade). Your class participation grade will be based on your timely completion of Latex, Git, and Github assignments. The midterm and final will be open note. Grades will be determined according to the following scale:

A	90%-100%
B+	85%-89%
B	80%-85%
C+	75%-79%
C	70%-74%
D+	65%-69%
D	60%-64%
F	0%-59%

References

Chen, D. (2016). *Git Essentials LiveLessons*. Addison-Wesley Professional.

Gandrud, C. (2015). *Reproducible Research with R and RStudio*. CRC Press.

Goossens, M., Mittelbach, F., Rowley, C., Carlisle, D., and Braams, J. (2004). *The Latex Companion, Second Edition*. Addison-Wesley Professional.

Silge, J. and Robinson, D. (2017). *Text Mining with R: A Tidy Approach*. O'Reilly Media.

Wickham, H. and Grolemund, G. (2016). *R for Data Science*. O'Reilly Media, Inc.