RIT SYLLABUS
ISTE.782, VISUAL ANALYTICS
FALL 2019 (TERM 2191), DRAFT OF AUGUST 26, 2019

DETAILS

Important note: The information presented in this syllabus is subject to expansion, contraction, change, or stasis during the semester. In case of conflict between versions, the copy on myCourses takes precedence.

Course Number. 13522

Prerequisites. None

Time. Online

Place. Online

Dates. 26 Aug 2019–11 Oct 2019

Final Exam. Online. Time to be announced.

Instructor. Dr. Michael McQuaid (call me Mick)

Email. mjmics@rit.edu

Office. 70-2675

Office Hours. M,W,F 1000–1145 or by appointment

DESCRIPTION

This course introduces students to Visual Analytics, or the science of analytical reasoning facilitated by interactive visual interfaces. Course lectures, reading assignments, and practical lab experiences will cover a mix of theoretical and technical Visual Analytics topics. Topics include analytical reasoning, human cognition and perception of visual information, visual representation and interaction technologies, data representation and transformation, production, presentation, and dissemination of analytic process results, and Visual Analytic

case studies and applications. Furthermore, students will learn relevant Visual Analytics research trends such as Space, Time, and Multivariate Analytics and Extreme Scale Visual Analytics.

LEARNING OUTCOMES

At the end of this course, a student will be able to

- Understand key elements and techniques used in visualizations
- Know how to interpret charts and visualizations
- Design various visualization projects
- Create and present projects using effective visualizations

MATERIALS

Each student will need a computing environment capable of running the R programming language. A minimal Linux, Mac, or Windows computer should suffice. R will be taught as part of this course. Students should have proficiency with some full-featured text editor. A network connection capable of streaming video is essential for this online course.

The textbook for this course is *R for Data Science*, Wickham and Grolemund (2017). This book is available free online and for purchase worldwide. We will also make use of Adler (2012) to learn R basics.

SCHEDULE

Arabic numerals refer to days. The course runs for fifteen days over seven weeks (Roman numerals).
Week I

1. Course overview — Syllabus — Tableau — Data visualization workshop — Create visualizations — Create a dashboard — Create story points

2. Bertin (2010) — Semiology of graphics — Sign systems — The invariant — Components of information — Visual variables — Retinal variables — Levels of organization — The basic graphic problem — Five aspects of image theory — Angular legibility — Retinal legibility

Week II

3. Adler (2012) — The R programming language — Installation of R and RStudio — Basic operations — Functions — Variables — Data structures — Objects and classes — Models and Formulas — Wickham and Grolemund (2017) — RMarkdown — *Homework 1 is due*

4. Wilkinson (2005) — Grammar of graphics — Data extraction — Data scales — Statistics — Geometry — Coordinate systems — Aesthetics — Rendering

Week III

5. Wickham (2010) — Layered grammar of graphics — Building a plot — Transforming input — Mapping to aesthetic space — Transforming data to a graphic — Faceting — Components of the layered grammar — Defaults — Layers — Scales — Coordinate Systems — Facets — Role of developers and users — Embedding the grammar — Implications of the grammar — Examples of the grammar — Appearance of visualizations — Limitations and future work

6. Wickham and Grolemund (2017), Chapter 1 — R — Implementation of aesthetic mappings in R — Implementation of facets in R — Implementation of geometric objects in R — Implementation of statistical transformations in R — Implementation of position adjustments in R — Implementation of coordinate systems in R — *Homework 2 is due*

Week IV

7. Wickham and Grolemund (2017), Chapter 3 — R basics — The dplyr library — The filter() function — De Morgan's laws — The arrange() function — The select() function — The mutate() function — The summarize() function — The group_by() function — Use of pipes — Missing values — Examples of data transformation

8. Wickham and Grolemund (2017), Chapter 5 — Exploratory data analysis — Variation in categorical variables — Variation in continuous variables — Exploring categorical and continuous variables together — Finding typical values — Finding unusual values — boxplots — scatterplots — tiled plots — Abbreviating ggplot() calls to speed up exploration

Week V

9. Wickham and Grolemund (2017), Chapters 7 and 8 — Unix utilities — Creating tibbles — Reading with the readr library — Dataframes and tibbles — Encodings — Numeric encodings — String encodings — Date-time encodings — Other data sources — *Homework 3 is due*

10. Wickham and Grolemund (2017), Chapter 9 — Untidy data examples — Reasons for untidyness — Tidy data

and databases — Using the gather() function — Using the spread() function — Using the separate() function — Using the unite() function — Implicit missing values — Explicit missing values

Week VI

11. Wickham and Grolemund (2017), Chapter 10 — Relational data — Entity relationships — Mutating joins — Inner joins — Left outer joins — Right outer joins — Full outer joins — Defining keys — Filtering joins — Using the semijoin() function — Using the antijoin() function — Set operations — Intersection — Union — Set difference

12. Wickham and Grolemund (2017), Chapter 11 — String examples — String functions — Vectors of strings — Locales — Regular expressions — Metacharacters — Character classes — Alternatives — Repetition — Detecting matches — Logical subsetting — Finding matches per string — Regular expression examples — Replacing matches — Backreferences — Splitting a string into pieces — *Homework 4 is due*

Week VII

13. Wickham and Grolemund (2017), Chapter 18 — Model basics — Defining models — Model aspects — Generating models — Searching among models — Comparing models — Using the optim() function — Visualizing models — Visualizing residuals — Using the model_matrix() function — Fitting a model and generating predictions — Using transformations in modeling — Modeling nonlinear functions — *Homework 5 is due*

14. Wickham and Grolemund (2017), Chapter 22 — Communicating differs from exploring — Titles, subtitles, and captions for visualizations — Annotation of visualizations — Representing scales — Using breaks — Using logarithmic scales — Using color — Colorbrewer for discrete scales — Viridis for continuous scales — Context in plotting — Themes for ggplot2

15. Tufte (2001) — Use cases of visualization — Anscombe's quartet — Scatterplots — Maps — Summarizing — Comparing information display formats — Faceting — Untruthful graphical practices — Depicted dimensions and data dimensions — Data ink ratio — Chartjunk — Grid reduction — Multi functioning elements — *Homework 6 is due*

GRADING

The grading scale used along with the grade components follow.

- A >= 90.0%
- B >= 80.0% & < 90.0%
- C >= 70.0% & < 80.0%
- D >= 60.0% & < 70.0%
- F < 60.0%

The course grade is composed of 10 percent for each of the six homeworks, and 40 percent for the take home final exam, for a total of 100 percent. Homeworks will be graded strictly and with an attention to detail. Instructions will be provided that must be followed carefully to expect a passing grade.

The following are brief statements of policy that are, in many places, expanded at the URLs provided. You are bound by these policies and any protest that you did not read the extended versions at the provided links will not be heeded. Your familiarity with the following policies, dates, and and parameters will be assumed in this course.

Last day of 7-day add/drop period. Tuesday 3 Sep 2019

Last day to withdraw with W. Friday 8 November 2019 (This is the date on the Academic Calendar although it doesn't make sense for a half-term course.)

myCourses. All project assignments, lecture notes, and other distributable course materials will be available via myCourses. Except where otherwise indicated, all student project assignments will be submitted via myCourses dropboxes.

Grade Challenges. School of Information policy states that a student has one semester to challenge any grade. After that, grades cannot be challenged.

Late Work. Any work not submitted by the final due date receives a grade of zero unless arrangements are made previous to the initial due date.

Extra Credit. No extra credit is available in this course.

Accommodations. If you have a "Notice of Accommodation", you must provide your instructor with a copy of it within 1 week of starting this course. You must follow all the rules of the relevant office.

Academic Dishonesty. The policy on dishonesty is simple: Anyone caught cheating receives an "F" as a course grade, is removed from the section and a letter detailing the incident is placed into his or her folder. Any student accused of cheating should realize that the evidence has already been verified

by other faculty members and will withstand an appeal. Additionally, please review the institute policy at http://www.rit.edu/studentaffairs/studentconduct/rr_academicdishonesty.php

Acceptable Use. We are bound by the following Acceptable Computer Use policy at http://www.rit.edu/academicaffairs/policiesmanual/sectionC/C82.html

Student Responsibilities. Please review the general student responsibilities as outlined at http://www.rit.edu/~301www/rr.php3

Policy on Reporting Incidents of Discrimination and Harassment. RIT is committed to providing a safe learning environment, free of harassment and discrimination as articulated in our university policies located on our governance website. RIT's policies ***require faculty to share information*** about incidents of gender based discrimination and harassment with RIT's Title IX coordinator or deputy coordinators, regardless whether the incidents are stated to them in person or shared by students as part of their coursework. RIT Governance website: https://www.rit.edu/academicaffairs/policiesmanual/policies/governance

If you have a concern related to gender-based discrimination and/or harassment and prefer to have a ***confidential*** discussion, assistance is available from one of RIT's confidential resources on campus:

1. The Center for Women & Gender: Campus Center Room 1760; 585-475-7464; CARES (available 24 hours/7 days a week) Call or text 585-295-3533.
2. RIT Student Health Center – August Health Center/1st floor; 585-475-2255.
3. RIT Counseling Center - August Health Center /2nd floor - 2100; 585-475-2261.

4. The Ombuds Office – Student Auxiliary Union/Room 1114; 585-475-7200 or 585-475-2876.
5. The Center for Religious Life – Schmitt Interfaith Center / Rm 1400; 585-475-2137.
6. NTID Counseling & Academic Advising Services – 2nd Floor Lyndon B. Johnson; 585-475-6468 (v), 585-286-4070 (vp).

RIT Resilience. Success in this course depends heavily on your personal health and wellbeing. Recognize that stress is an expected part of the college experience, and it often can be compounded by unexpected setbacks or life changes outside the classroom. Your other instructors and I strongly encourage you to reframe challenges as an unavoidable pathway to success. Reflect on your role in taking care of yourself throughout the term, before the demands of exams and projects reach their peak. Please feel free to reach out to me about any difficulty you may be having that may impact your performance in this course as soon as it occurs and before it becomes unmanageable. In addition to your academic advisor, I strongly encourage you to contact the many other support services on campus that stand ready to assist you.

REFERENCES

Adler, Joseph. 2012. *R in a Nutshell, Second Edition*. Sebastopol, CA: O'Reilly Media.

Bertin, Jacques. 2010. *Semiology of Graphics: Diagrams, Networks, Maps (English Translation)*. Redlands, CA: ESRI Press.

Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*. Cheshire, Conn: Graphics Press.

Wickham, Hadley. 2010. "A Layered Grammar of Graphics." *Journal of Computational and Graphical Statistics* 19

(1). Informa UK Limited: 3–28.
https://doi.org/10.1198/jcgs.2009.07098.
Wickham, Hadley, and Garrett Grolemund. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. 1st ed. O'Reilly Media, Inc.
Wilkinson, Leland. 2005. *The Grammar of Graphics (Statistics and Computing)*. Secaucus, NJ, USA: Springer-Verlag.