IST 719 – Information Visualization

Tuesday, 9:30 AM to 12:15 PM

Instructor	Jeff Hemsley	Phone	TBA
Office	HH 311	E-mail	jjhemsle@syr.edu
Office Hours	TBA		

Course Description:

A broad introduction to data visualization for information professionals. Students will develop a portfolio of resources, demonstrations, recipes, and examples of various data visualization techniques.

Additional Course Description

Introduction to skills and techniques related to information visualization, through the R programming language, Adobe illustrator. These skills include data cleaning techniques, control of the R graphics environment, develop custom plots, visually explore data, use design concepts to visually communicate the story in the data, and discuss issues related to the ethics of data visualization. Conceptual themes will be presented alongside technical aspects of data visualization. Additional work and higher grading expected of graduate students.

Prerequisite / Co-requisite:

IST 687 or equivalent programming courses (Python, Java, SQL, C, C++, etc.)

Audience:

Students interested in data analytics and data science, with a focus on data/information visualization.

Credits:

3

Course Fees and/or Costs

Only costs are for the books listed below and for printing out the final poster project.

Learning Objectives:

1. Perform basic data cleaning and preparation on a wide range of data sets using R
2. Identify stories in data sets through visual data exploration
3. Create rich visual artefacts that communicate data stories

After taking this course, students will be able to:

1. Have basic proficiency with R programming language and Adobe Illustrator
2. Use functions and plots in R to explore a dataset and find simple relationships, distributions, and exceptional cases
3. Use R and Adobe Illustrator to create publish-quality data visualizations that communicate the story of the data

Required Texts / Supplies:

Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics, by Nathan Yau. Wiley Publishing, 2011. [VT in schedule]

Data Points: Visualization That Means Something, by Nathan Yau. Wiley Publishing, 2013 [DP in schedule]

Texts / Supplies – Additional:

Additional readings will be supplied as PDFs

Course Requirements and Expectations:

In order to meet the goals of the class (see above), we will use a combination of lectures to introduce topics and concepts, hands-on-labs to introduce skills, group exercises and student presentations to enable peer-to-peer learning, and homework assignments to practice skills and gain deeper knowledge of course content. These are detailed below.

Grading:

Because of the dynamic nature of the labs, time available, and any opportunities that might arise, the total points possible may vary significantly. But this should be nailed down within the first week or two of class. Thus, the numbers are approximate.

Assignment	Point Range
In-Class Labs	10 – 12
Quizzes	30 – 40
Homework	10 – 15
Group Reports and Feedback	5 – 10
Viz-a-thon*	5 - 10
Final Project	25
Total Range	100 - 110

Grading Table

Grades	Grade Points /Credit	Percentage Range
A	**4.000**	96%-100%
A-	**3.66**	93%-95.9%
B+	**3.33**	90%-92.9%
B	**3.00**	87%-89.9%
B-	**2.66**	84%-86.9%
C+	**2.33**	81%-83.9%
C	**2.00**	78%-80.9%
C-	**1.66**	75%-77.9%
F	**0**	

* source: http://www.syr.edu/registrar/students/grades/faq.html

[1]Grades of D and D- may not be assigned to graduate students.
Use Graduate syllabus template with the appropriate grading table

Course Specific Policies on attendance, late work, make up work, examinations if outside normal class time, etc.:

In Class Labs: In the real world of data visualization people work both alone and in teams to meet near and far deadlines. In this class we work in a lab setting to learn R and Adobe Illustrator skills and students are encouraged to work together to solve problems. Labs are between 1 and 3 points. Labs are also dynamic. The nature of what we turn in at the end of the lab will depend on how fast we go and what questions students ask. Because of the nature of this work, I do not allow for makeups of labs.

Lab Quizzes: These are typically one or two question quizzes that are intended to reinforce something I talked about in the last class. These won't be announced in advance.

Homework & Quizzes: These will be extensions of what we did in class or assignments out of the book. Homework may be in the form of quizzes on Blackboard, visualizations you create or some of the other assignments listed below. Quizzes often lean heavily on the reading and students who do not keep up with the readings, often do not do well on the quizzes. Homework and quizzes are usually due by 9:00am at the start of the next class. You can turn in assignments late, but there are consequences. First, 1 point will be automatically deducted. Second, late assignments may not be graded till the end of the semester.

Important: you may not receive credit if you do not follow the file naming convention specified on the assignment sheet. You may not receive credit if your file is of the wrong type. Unless otherwise specified, you will always turn in plots as .pdf files and R scripts as .R files.

~~*Advanced Topic Presentations:* While this class is focused on creating visualizations using R, many other tools exist. Examples include D3, Tableau, and Gephi. In order to give students exposure to these and many other options, students will select, research and present on an "advanced topic". Graduate students will work individually, while undergraduates may work in groups of two.~~

Final Project: The final project, including poster session, is worth a large percentage of your final grade. The final poster project leverages skills developed throughout the semester, including cleaning data, exploring data with visualization techniques, data aggregation, simple design and information organization skills, and quality graphic presentation of data visualizations. Key deliverables leading up to the final project will be due throughout the semester to help the students stay on track for this major deliverable.

Poster Session: You must be present at the poster session at the end of the semester, with your poster, in order to get credit for your poster. The exact time and date of the poster session will be announced within the first few classes. Note that the requirements for posters change each semester. Past examples on Blackboard are provided for your reference.

Syracuse University Policies

Syracuse University has a variety of policies designed to guarantee that students live and study in a community respectful of their needs and those of fellow students. **The policies and services are listed on the new Syracuse University Senate approved syllabus appendix titled, '*Syracuse University Student Policies and Services*'. These statements are an official part of this course syllabus.**

Course Schedule:

Week	Topic	Required Reading and Assignment
1	**What is data visualization? What is R?** ***Learning Outcomes*** - Students will be able to: • Differentiate between Information Visualization and other prominent forms of visualization • Describe the two main purposes of Information Visualization: exploration and communication • Describe the 7 basic steps of visualization • Describe the purpose of the 4 windows of RStudio • Create variables in R • Create simple single variable plots in R, such as pie and bar charts, histograms Use R's help system to lookup the available parameters for plotting functions	***Lab 1***: R and Basic Plots ***Readings***: • VT: Chapters 1 and 2 • DS: Chapters 1 and 2 • Ben Fry, Visualizing Data ***Quiz 1***: covers readings and lecture
2	**Data and R** ***Learning Outcomes*** – Students will be able to: • List sources of freely available data • Describe the process of data exploration • Discuss the role of context markers in visualization • Differentiate between common data types • Open data files • Use R functions to explore and clean data • Use R to retype, suset and filter data Create rough data exploration plots	***Lab 2***: Exploring Data in R ***Readings***: • DP: Chapter 1 • DS: Chapter 3 • VT: Chapters 3 and 4 ***Homework 1***: Reproduce Figures in VT Chapter 4
3	**Using Data Libraries and visualizing multi-dimensional data** ***Learning Outcomes*** – Students will be able to: • Find datasets from online data libraries such as Data Planet • Describe ways to make comparisons with visualizations • Describe ways to identify and show relationships in data • Differentiate between single and multi dimension plots • Use R to make simple multi-dimensional plots Identify the appropriate plot type for a given set of data	Lab 3: Visualizing multi-dimensional data Assignment: Find a dataset and be prepared to describe/discuss it in class
4	Covering GGPlot in Depth Learning Outcomes – Students will be able to: • Describe the elements of the grammar of graphics needed to build a ggplot plot	Lab 4: ggplot

Week	Topic	Required Reading and Assignment
	• Build simple and complicated plots using ggplot • Describe the pros and cons of ggplot and when not to use it	
5	**Beginning to find and tell the story in the data & Illustrator Intro** ***Learning Outcomes*** - Students will be able to: • Describe a dataset • Identify questions that might be answered with the data • Identify elements of a visual artifact that make it compelling • Interpret the meaning(s) of a data visualization • Use illustrator to modify R plots • Add context elements to a data visualization • Use Illustrator to modify plot colors, type face and layout Explain the difference between raster and vector graphics	Lab 5: Illustrator Intro Quiz 2: Optimal visual encoding of data Readings: • DP: Chapters 2 and 3 • Few: Chapter 3 • VT: Chapter 4 (Illustrator parts) Homework 2: Illustrator portions of VT Ch 4 plots
6	**Graphic Design Principles: Color, tools and R color functions** ***Learning Outcomes*** – Students will be able to: • Describe the use of contrasting and harmonious color in visualization • Describe how hue, saturation and value combine to make a color • Use online tools to choose and create color schemes • Use R's color setting and transformation functions Use color to provide visual cues in visualizations	***Lab 6***: Working with color ***Reading***: • DP: Chapter 4 ***Homework 3***: Visualize your data: creating a visual report of your dataset
7	**Graphic Design Principles: Type Face & Layout, R plot area control** ***Learning Outcomes*** – Students will be able to: • Describe how type face and layout work together to create a visual hierarchy • Describe how visual hierarchies direct viewers attention • Explain how lines, gutters, grids and colors can be used to highlight visual elements Critically assess example posters and discuss useful and detracting design elements	***Lab 7*** ***Reading***s: • DS: Chapter 5 • VT: Chapter 6 • Few: Chapter 4 ***Homework 4***: Create plots in VT, chapter 6 ***Quiz 3***: Design supported comparisons
8	**Advanced Plots & Posters in Illustrator** ***Learning Outcomes*** – Students will be able to: • Prepare geographic based data for plotting	***Lab 8***: Maps and posters ***Readings***: • DS: Chapter 7

Week	Topic	Required Reading and Assignment
	• Create national, regional and world map plots in R • Enhance cognitive apprehension of complex maps through the use of design elements • Use Illustrator to merge 2 or more plots and incorporate context text and design elements Create poster sized, high quality vector graphic ready for printing	• VT: Chapter 7 ***Quiz 4***: The right plot of the data
9	Data scaling, word clouds and social network analysis	**Lab 9**
10	**Viz-a-thon** *Students are given a new dataset and must work together in small groups to explore the data to find the story, then create a mini-poster using R and Illustrator* ***Learning Outcomes*** – Students will be able to: • Demonstrate knowledge of the process of creating a visual artifact Work in a group to create a mini-poster from an unknown dataset	
11	Ethics and Project Group Feedback ***Learning Outcomes*** – Students will be able to: • Critically assess visualizations • Identify the audience for a visual artifact • Identify visual credibility markers Discuss the ethical concerns around visual artifacts	Readings: • DS: Chapter 8 • VT: Chapter 8 Quiz 6: The right plot for the data
12	Advanced ggplot	**Lab 12**
13	**Interactivity in R Plotting: Shiny** ***Learning Outcomes*** – Students will be able to: Create an online interactive visualization portfolio	Lab: Shiny
14	**Final Project Poster Session** *Students present a poster they created themselves at the all-iSchool poster session*	