BUS 240 f1: Information Visualization

Course Syllabus
Spring 2020

Bharatendra Rai
brai@brandeis.edu

Class Hours: Th 6:30 PM – 9:20 PM.
Location: Golding Judaica Center101

Office Hours:
Thursdays  5:00 – 6:30 PM
and by appointment

Office:     tbd

TA:        tbd
Overview

Catalog Description

Modern computer graphics provide many ways to tame "big data," allowing users not only to view multidimensional information, but to interactively explore, combine, and interpret massive volumes of information using software tools including R and Tableau.

Meets for one-half semester and yields half-course credit.

Course Description

Among the most promising developments in data analytics is the growth in Information visualization capabilities. Across numerous disciplines, tools and techniques are emerging that help people interactively analyze and understand the flood of data now available. Not only is there is staggering growth not only in the availability of tools, but also in the domains in which these techniques are deployed.

Visualization can tap into the ways in which humans rapidly and intuitively process information, taking advantage of IT tools to move beyond tabular displays and simple graphs. The theory and techniques of data visualization not only make information accessible, but can open paths to exploration of cause and effect.

This module provides an overview of the field of Data Visualization, presenting current theory and best practices to students. We will rely heavily on hands-on learning, interspersed with readings, cases, lectures and occasional guest speakers. Although we naturally will use particular software, the lessons are aimed at principles so that students can continue to refine their skills as new tools emerge. Students will learn to evaluate and assess existing visualizations as well as develop their own information-rich interactive displays.

Learning Goals and Objectives

Upon successful completion of this module, students will be able to:

• Create complex data visualizations the address the needs of business users;
• Understand and apply strategies of analytical design;
• Use visualization packages and original coding to produce effective visualization products
• Acquire, parse, and analyze abstract data sets
• Make specific recommendations for improvement and enhancement of visualizations created by others;
• Demonstrate agility in rapidly creating prototype visualizations

Intended Audience

Graduate students at IBS in various degree programs, of different nationalities and different work experience.

Delivery Method and Instructional Approach

Course will take place in a classroom and will include online elements for accessing materials and submitting work on the Brandeis learning management system (LATTE).

This course will use both standard classroom lecture-structure and a flipped classroom. For certain class days, the students will access lecture-style materials prior to class where they will then apply the concepts and topics while their work is facilitated by the instructor and TAs. On other days, the instructor will present his lecture within the designated class-time while students are expected to complete any assignments outside of class.
Recommended Readings


Other required readings as posted on LATTE site.

Assessments

1. 6 individual on-line chapters at DataCamp.com (Note: 5 assignments, consisting of 6 chapters). Scoring explained on p. 4
2. 2 individual peer-reviewed visualization tasks (LATTE Workshops)
3. 1 pair visualization assignment (done with 1 partner)
4. 2 team-based assignments
5. several short Forum entries in response to reading-related prompts
6. “Speed” presentations (team-based) instead of final exam

Course grade weightings

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course engagement (includes Forum posts and in-class participation)</td>
<td>10%</td>
</tr>
<tr>
<td>Peer-reviewed visualizations (2)</td>
<td>20%</td>
</tr>
<tr>
<td>Individual DataCamp assigned chapters (6)</td>
<td>20%</td>
</tr>
<tr>
<td>Pair visualization assignment</td>
<td>10%</td>
</tr>
<tr>
<td>Team-based longer assignments (average of 2)</td>
<td>40%</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

Prerequisites

BUS 211f or permission of instructor. You must have familiarity with R and RStudio.

Other Course Technology

All of the software we will use in this course can be accessed on the public computer clusters at IBS and/or on your personal laptops. If you do use a laptop, the class schedule below indicates dates when it will be useful to have it with you.

Most of the data we will analyze will come from a variety of open-access and IBS-supported website and databases. Additionally, we will use the following tools:

- **R**: R is a free software environment for statistical computing and graphics, and is widely used by both academia and industry. In this module we will use R mainly for data visualization.

  R Software: [http://www.r-project.org/index.html](http://www.r-project.org/index.html)

In addition to aspects of R and RStudio that you know well, we’ll also use Brandeis’ RStudio Connect server. This allows us to post Markdown and Shiny products to share with the class.

- **Tableau**: Tableau provides an environment for data management and visualization, as well as some basic statistical functions. It is one of the fastest-growing tools available for laptops and desktops, and there are free Academic versions for Windows and Mac-OS.
• **DataCamp:** [www.datacamp.com](http://www.datacamp.com) offers an excellent free (to us) instructional platform. For some technical skills, you will complete some independent individual assignments at your own pace via DataCamp. By moving these skill lessons out of the classroom, we will have more in-class time to discuss complex topics and solve problems together.

Class participation is important in this course both as a means of developing understanding and as an indicator of student progress. Participation can take many forms, and each student is expected to contribute actively, freely, and effectively to the classroom experience by raising questions, demonstrating preparedness and proficiency in the analysis of assignments and cases, and explaining the implications of analyses in context. Homework-based discussion and presentations are an important part of participation. **To this end, regular class attendance is required, and students should use name cards.** We meet only twelve times, so absence can become a serious problem. Even if you must arrive late or leave early, be here.

With assistance from the TA, I will evaluate the quality of your contributions in class each evening, as well as the quality of your contributions via email, LATTE discussion, etc. These will all be factored together in determining your ultimate Engagement grade. In general, absence from class reduces your contribution grade.

Your engagement in the course will also take the form of a few short Forum posts in response to assigned reading and to other student comments.

Early in the term, you’ll receive a link to join our DataCamp course group. **To access and receive credit for the assignments, you need to be in the group.**

We’ll use DataCamp to introduce several important technical details of the r packages `ggplot2`, `shiny`, and `leaflet`. There are four on-line assignments to complete comprising 8 chapters in all. Each chapter should take 30 to 60 minutes. You must complete these by yourself prior to the deadline.

**DataCamp assignments**

<table>
<thead>
<tr>
<th>DataCamp score achieved</th>
<th>Points earned</th>
</tr>
</thead>
<tbody>
<tr>
<td>80 – 100</td>
<td>3</td>
</tr>
<tr>
<td>60 - &lt; 80</td>
<td>2</td>
</tr>
<tr>
<td>40 - &lt; 60</td>
<td>1</td>
</tr>
<tr>
<td>&lt; 40 or missed deadline</td>
<td>0</td>
</tr>
</tbody>
</table>

**Other Assignments and Projects**

a) There will be two “Peer-Reviewed” LATTE WORKSHOP assignments in which you (a) prepare some visualizations by yourself and then (b) review and comment on a classmate’s work.

b) One assignment will be done in pairs – you and one partner will collaborate on designing and creating a visualization.

c) Two other assignments will be “Projects” requiring more significant time and analysis. The projects will be prepared in teams of four to five students (I will assign you to a team).

a. The 2nd project also includes a “Speed Session”. In place of a final exam, we will spend 80 minutes during which every team will have 5 minutes to present and demonstrate their interactive dashboard.
All assignments should be submitted via LATTE upload prior to the start of class. Papers should be professional in appearance and use clear, grammatically correct business English. Consistent with the orientation of the module, all work should reflect an understanding of the impact of visual appeal.

**Workload Expectations**

Success in this two-credit course is based on the expectation that students will spend a minimum of 9 hours of study time per week for six weeks in preparation for class (readings, papers, discussion sections, preparation for exams, etc.)."

**Academic Integrity**

You are expected to follow the University’s policies on academic integrity (see [http://www.brandeis.edu/studentaffairs/srcs/ai/index.html](http://www.brandeis.edu/studentaffairs/srcs/ai/index.html)). Instances of alleged dishonesty will be forwarded to the Office of Campus Life for possible referral to the Student Judicial System. Potential sanctions include failure in the course and suspension from the University.

**Disabilities**

If you are a student with a documented disability on record at Brandeis and wish to have a reasonable accommodation made for you in this class, please see me immediately.

**Course Teams**

Working with partners is an excellent way to gain understanding of this subject. Most of the assignments in the course must be done as individuals, but you will be assigned to a team for the two project assignments. As you work in your teams to complete assignments, with a few caveats:

- Be sure that you are neither carrying nor being carried by the group; each member of the group is entitled to learn and is expected to contribute.
- Even in the context of group work, each student is responsible for the quality and timeliness of the submitted work.
- If your team is experiencing performance problems, please speak with me at once.
**Course Outline**

Note: for each session, you should complete the assigned reading *before* coming to class. Detailed assignments will be posted on LATTE, and all assignments and handouts will also be available on our LATTE site. Watch LATTE closely for updates.

<table>
<thead>
<tr>
<th>Session # &amp; Date</th>
<th>Topics and Readings to do <em>before</em> class</th>
<th>DataCamp assignment due by class</th>
<th>Required Deliverable Due by class time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theme 1: Foundations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| #1 Thurs Mar 14  | Business Uses of Visualization & Available Technologies  
• **WATCH** on LATTE  
  o Carver video #1 on LATTE  
  o Hans Rosling (Gapminder.org) TED talk  
• Yau: Intro & Chap 1  
• Cairo Chap 1 |  | Forum post responding to Rosling video;  
  *Complete your “Data Science Profile”* |
| #2 Tues Mar 19   | Human Perception & Cognition: Visual Thinking Process  
• Ware, Chapter 11 (on LATTE)  
• **WATCH**: Cairo: The Functional Art Intro (LATTE)  
• Cairo, Chap 5  
• Yau, Chap 3 (to p. 75, skim the rest)  
• **INSTALL** Tableau desktop  
• **WATCH**: Intro to Tableau (link on LATTE)  
  *Introduction to the tool kit: out-of-the-box software (Tableau) and custom programming (R)* |  | Forum Post: See prompt on LATTE |
| **Theme 2: Exploration and Discovery** |
| #3 Thurs Mar 21  | Methods for Detecting and Displaying Patterns  
• **WATCH** Cairo, Chap 2 video  
• Skim: Cairo Chap 2  
| #4 Tues Mar 26   | Design choices: Matching Data Types, Plots and Objectives  
• **WATCH**: Carver video #2 on LATTE  
• Yau, Ch. 4 & 5 | DC1: Data Visualization with ggplot2 (Part 1) Ch 1 Introduction. |  |
| # 5 Thurs Mar 28 | **Making Complex Relationships Visible**  
• Yau Ch. 6 & 7  
• Cairo, Ch. 6  
• B. Schneiderman, “The Eyes Have it” (1996) |  | Visualization 2 in pairs |
<table>
<thead>
<tr>
<th>Session # &amp; Date</th>
<th>Topics and Readings to do before class</th>
<th>DataCamp assignment due by class</th>
<th>Required Deliverable Due by class time</th>
</tr>
</thead>
</table>
| #6 Tues April 2  | Special considerations for time-based and geographic data  
  - Yau, Ch. 8  
  - *(optional)* Basic Leaflet Maps in R  
  https://rpubs.com/mattdray/basic-leaflet-maps | DC2: Interactive Maps with Leaflet in R, Ch 1: Setting up Interactive Web Maps | Forum Post: See prompt on LATTE |
|                  | Theme 3: Presentation Graphics                                                                        |                                  |                                       |
| #7 Thurs April 4 | Static Visuals that Inform and Persuade  
  - Cairo, Ch. 8  
  - WATCH Carver Video #3 (LATTE) | DC3: Building Web Applications in R with Shiny, Ch 1: Intro & Shiny Basics |                                       |
|                  | Automating and Animating  
  - Yau, Ch. 2 (pp. 22-38)  
  - Package `rvest` intro (see links on LATTE) | Project 1                        |                                       |
|                  | Theme 4: Interactive Story-Telling                                                                    |                                  |                                       |
| #9 Thurs April 11| Planning Ahead: Prototyping and Story-Boarding  
  - Yau, Ch. 9  
  - Cairo, Ch. 8 (revisited)  
  - Berinato, *Good Charts*, Chap 4 (LATTE) | DC4: Interactive Data Visualization with Plotly in R, Ch 1: Introducting to plotly |                                       |
| #10 Tues April 16| Incorporating User Controls  
  - WATCH: Carver video #4, LATTE  
  - Cairo, Ch. 9  
  - Kirk, *Data Visualization*, Chap7 (LATTE)  
  - Skim: Herr & Shneiderman 2012 (LATTE) | DC5 & 6: Building Dashboards with Shinydashboard, Ch 1: Building Static Dashboards AND Ch 2: Building Dynamic Dashboards (2 chaps) |                                       |
| # 11 Thurs April 18| Dashboards for Business Intelligence and Decision Support  
| Apr 19 – Apr 26  | Passover and Spring Break – no classes                                                                |                                  |                                       |
| #12 Tues April 30| Project “Speed Presentations” instead of final exam                                                    |                                  |                                       |
| Fri May 3        | • Submit final project before midnight                                                               | Project 2                        |                                       |