BUSMGT 7331: Descriptive Analytics and Visualization

Spring 2019, Term 1

January 7, 2019

Instructor: Hyunwoo Park (https://fisher.osu.edu/people/park.2706)

Office: Fisher Hall 632

Email: park.2706@osu.edu

Phone: 614-292-3166

Class Schedule: 8:30am-12:pm Saturday 1/19, 2/2, 2/16

Recitation/Office Hours: Thursdays 6:30pm-8pm at Gerlach 203 (also via WebEx)

1 Course Description

Businesses and organizations today collect and store unprecedented quantities of data. In order to make informed decisions with such a massive amount of the accumulated data, organizations seek to adopt and utilize data mining and machine learning techniques. Applying advanced techniques must be preceded by a careful examination of the raw data. This step becomes increasingly important and also easily overlooked as the amount of data increases because human examination is prone to fail without adequate tools to describe a large dataset. Another growing challenge is to communicate a large dataset and complicated models with human decision makers. Descriptive analytics, and visualizations in particular, helps find patterns in the data and communicate the insights in an effective manner. This course aims to equip students with methods and techniques to summarize and communicate the underlying patterns of different types of data. This course serves as a stepping stone for further predictive and prescriptive analytics.

2 Course Learning Outcomes

By the end of this course, students should successfully be able to:

- Identify various distributions that frequently occur in observational data.
- Compute key descriptive statistics from data.
- Measure and interpret relationships among variables.
- Transform unstructured data such as texts into structured format.
- Communicate various descriptive statistics with visualization.
- Design advanced visualizations curating visual elements.
- Explain pros and cons of various visual representations depending on the context and form of data.
- Choose appropriate visual representations for special forms of data such as geographic, textual, and network data.
- Compose a visual dashboard composed of interactive visual artifacts.

3 Prerequisite

- Enrollment in Fisher's Specialized Master's Degree Program in Business Analytics (SMB-A)
- Courses
  - BUSMGT 6400: Statistics and Data Analysis for Managers
  - BUSMGT 7250: Data Management for Analytics
- Or, permission of the instructor

4 Course Materials

- Required Textbooks
  - *R for Data Science* by Hadley Wickham & Garrett Grolemund Free access online
  - *Text Mining with R* by Julia Silge & David Robinson Free access online
- Required DataCamp Courses
  - Data Visualization with R: 3 courses
    - Data Visualization with ggplot2 (Part 1)
    - Data Visualization with ggplot2 (Part 2)
    - Visualization Best Practices in R
  - Text Mining with R: 3 courses
    - String Manipulation in R with stringr
    - Sentiment Analysis in R: The Tidy Way
    - Sentiment Analysis in R
  - Spatial Data with R: 2 courses
    - Working with Geospatial Data in R
    - Interactive Maps with leaflet in R
  - Shiny Fundamentals: 2 courses
    - Building Dashboards with shinydashboard
    - Building Dashboards with flexdashboard
  - Individual Courses
    - Categorical Data in Tidyverse
    - Working with Dates and Times in R
    - Correlation and Regression
    - Cluster Analysis in R
- Additional Materials or URLs posted on Carmen

5 Weekly Schedule

- Week 1 (1/7-1/11): Data Wrangling and Descriptive Statistics
- Week 2 (1/14-1/18): More Data Wrangling, Covariation, and Customizing Visualization
- In-class Meeting #1 (1/19)
- Week 3 (1/21-1/25): Describing Textual Data
- Week 4 (1/28-2/1): Geospatial Visualization
- In-class Meeting #2 (2/2)
- Week 5 (2/4-2/8): Network Visualization
- Week 6 (2/11-2/15): Interactivity and Dashboard Design
- In-class Meeting #3 (2/16)
- Week 7 (2/18-2/22): Visualizing Models, Dimensionality Reduction, and Clustering

6 Course Components and Grading Breakdown

Assignment Category	Points
DataCamp Courses	21 (3 pts/course for 7 completed courses out of 14 assigned courses)
Concept Checker Quizzes	7 (1 pt/wk for 7 weeks)
Weekly Problem Sets	18 (3 pts/wk for 6 weeks)
Bi-weekly Assignments	18 (6 pts/wk for 3 weeks)
In-class Activities	18 (6 pts/wk for 3 weeks)
Final Exam	18
Total	**100**

Late submissions will not be accepted. Please refer to Carmen for due dates.

6.1 DataCamp Courses

In most weeks, students will be assigned to complete a set of DataCamp courses. Each course completed in time is 3 points. Even if students complete them after the due, students receive 2 point per course. Only up to seven completed DataCamp courses will be counted towards the final course grade. Note that due dates vary across DataCamp courses.

6.2 Concept Checker Quizzes

After each module of videos, there will be a very brief Carmen quiz to cover the content of the module. Students have infinite attempts on the quiz and are not penalized by multiple attempts. The purpose is simply to keep on track and self diagnose problems along the way thus generating more insightful questions.

6.3 Weekly Problem Sets

Each week, a problem set will be delivered via a Carmen quiz. These problem sets will be untimed, open material, but must be answered by individual work. Unlike concept checker quizzes, students have only one attempt on the questions. These questions can include multiple choice, true/false, numeric answers, and potentially small image uploads. The answers to the problem sets must be submitted by 11:59 PM on Fridays (to allow for discussion on the in-class weeks). Late submissions will not be accepted.

6.4 Bi-weekly Assignments

On weeks in which we don't meet in person, a small delivery will be required. This will essentially be a mini-project to show the skills required from the week. The hope is to replicate some hands-on experiential learning in the off weeks and have more 'practice' than simple problems. Bi-weekly assignments will be given in Weeks 1, 3, and 5 when no class meeting is scheduled, so each assignment is worth 6% of the grade. The instructions for submission can be found here.

6.5 In-class Activities

The weeks in which we meet, we will also have a mini-project to complete worth 18% of the grade. These will involve some more focused intense data analysis individually or in groups and some data collection as well. Part of the in-class session will be devoted to questions/clarification but most of the time will be spent in a more holistic activity covering the 2 weeks of online content consumed.

6.6 Final Exam

The final exam will be opened on the first day of exams and closed the following evening. The precise format of the exam as well as time permitted is TBD.

7 Other Course Policies

7.1 Recitation/Office Hours

Each week, I will be available remotely via WebEx (link to be provided) for synchronous office hours. Office hours will be held every Thursday from 6:30pm to 8pm at Gerlach 203. Students are welcome to walk in, but remote participation will also be possible. These sessions will provide students with the ability to ask probing questions on the current material and clarify gaps in understanding. Please feel free to bring problems from the book or outside questions for practice as well. If you have a specific question that you'd like to delve into, please bring it up on the discussion board so I can adequately prepare a solution in advance if possible.

7.2 Course Technology and Help

For help with your password, university e-mail, Carmen, or any other technology issues, questions, or requests, contact the OSU IT Service Desk. Standard support hours are available at https://ocio.osu.edu/help/hours, and support for urgent issues is available 24x7.

- Self-Service and Chat support: http://ocio.osu.edu/selfservice
- Phone: 614-688-HELP (4357)
- Email: 8help@osu.edu
- TDD: 614-688-8743

7.3 Microsoft Office 365

All Ohio State students are now eligible for free Microsoft Office 365 ProPlus through Microsoft's Student Advantage program. Full instructions for downloading and installation can be found https://ocio.osu.edu/kb04733.

7.4 Faculty Feedback and Response Time

I am providing the following list to give you an idea of my intended availability throughout the course. Remember that you can call 614-688-HELP at any time if you have a technical problem.

- Grading and feedback: For large weekly assignments, you can generally expect feedback within 7 days.
- E-mail: I will reply to e-mails within 24 hours on school days.
- Discussion board: I will check and reply to messages in the discussion boards every 24 hours on school days.

7.5 Academic Integrity Policy

7.5.1 Policies for this Online Course

- Quizzes and exam: You must complete the answers to quizzes, problem sets, and exam yourself, without any external help or communication.

- Reusing past work: In general, you are prohibited in university courses from turning in work from a past class to your current class, even if you modify it. If you want to build on past research or revisit a topic you've explored in previous courses, please discuss the situation with me.
- Falsifying research or results: All research you will conduct in this course is intended to be a learning experience; you should never feel tempted to make your results or your library research look more successful than it was.
- Collaboration and informal peer-review: The course includes many opportunities for formal collaboration with your classmates. While study groups and peer-review of major written projects is encouraged, remember that comparing answers on a quiz or assignment is not permitted. If you're unsure about a particular situation, please feel free just to ask ahead of time.

7.5.2 Ohio State's Academic Integrity Policy

Academic integrity is essential to maintaining an environment that fosters excellence in teaching, research, and other educational and scholarly activities. Thus, The Ohio State University and the Committee on Academic Misconduct (COAM) expect that all students have read and understand the University's Code of Student Conduct, and that all students will complete all academic and scholarly assignments with fairness and honesty. Students must recognize that failure to follow the rules and guidelines established in the University's Code of Student Conduct and this syllabus may constitute "Academic Misconduct."

The Ohio State University's Code of Student Conduct (Section 3335-23-04) defines academic misconduct as: "Any activity that tends to compromise the academic integrity of the University, or subvert the educational process." Examples of academic misconduct include (but are not limited to) plagiarism, collusion (unauthorized collaboration), copying the work of another student, and possession of unauthorized materials during an examination. Ignorance of the University's Code of Student Conduct is never considered an "excuse" for academic misconduct, so I recommend that you review the Code of Student Conduct and, specifically, the sections dealing with academic misconduct.

If I suspect that a student has committed academic misconduct in this course, I am obligated by University Rules to report my suspicions to the Committee on Academic Misconduct. If COAM determines that you have violated the University's Code of Student Conduct (i.e., committed academic misconduct), the sanctions for the misconduct could include a failing grade in this course and suspension or dismissal from the University. If you have any questions about the above policy or what constitutes academic misconduct in this course, please contact me.

Other sources of information on academic misconduct (integrity) to which you can refer include:

- The Committee on Academic Misconduct web pages
- Ten Suggestions for Preserving Academic Integrity
- Eight Cardinal Rules of Academic Integrity

7.5.3 Copyright Disclaimer

The materials used in connection with this course may be subject to copyright protection and are only for the use of students officially enrolled in the course for the educational purposes associated with the course. Copyright law must be considered before copying, retaining, or disseminating materials outside of the course.

8 Accessibility Accommodations for Students with Disabilities

8.1 Requesting Accommodations

If you would like to request academic accommodations based on the impact of a disability qualified under the Americans with Disabilities Act and Section 504 of the Rehabilitation Act of 1973, contact your instructor privately as soon as possible to discuss your specific needs. Discussions are confidential. In addition to contacting the instructor, please contact the Student Life Disability Services at 614-292-3307 or ods@osu.edu to register for services and/or to coordinate any accommodations you might need in your courses at The Ohio State University.

Go to http://ods.osu.edu for more information.

8.2 Accessibility of Course Technology

This online course requires use of Carmen (Ohio State's learning management system) and other online communication and multimedia tools. If you need additional services to use these technologies, please request accommodations with your instructor.

- Carmen (Canvas) accessibility
- Streaming audio and video
- Synchronous course tools