IDS 707: Data Communication, Visualization and Storytelling

Spring 2019, Tu/Th 10:05-11:20 AM

Instructor: Dr. Jana Schaich Borg

js524@duke.edu

TA: Ryan Huang, PhD (ryan.huang@duke.edu)

*Course Description*

One of the characteristics that differentiates effective data scientists from ineffective data scientists is the ability to communicate practical implications of data analyses. Even the most sophisticated statistical analyses are not useful outside of academic pursuits if they do not lead to actionable advice, or if the answers the analyses provide are not conveyed in a way that non-technical people can understand. Great data scientists can explain the significance of any statistical model to any audience member. Data visualizations are one of the most powerful tools in a data scientist’s communication toolbox. One of the main goals of this course is for you to learn how to use best visualization practices to uncover and communicate the significance of your data efficiently.

That said, creating effective visualizations is only one piece of an analysis pipeline, and the success of the visualization process is contingent on what happens before and after the visualizations are made. To ensure you have all the tools you need to succeed, this course will therefore make sure you understand and practice the role that asking good questions, listening to stakeholders, leveraging structured analysis plans, creating logically-valid arguments, and telling persuasive stories have on your visualizations and, ultimately, your influence as a data scientist. We will use Tableau to practice the visualization concepts discussed in this class, but you will be given opportunities to learn other visualization software if you wish, such as the visualization packages offered in R and Python.

*Learning Objectives*

By the end of this course you will:

• appreciate the role communicating with stakeholders has in ensuring your analysis projects yield successful results;

• practice interviewing stakeholders to determine what project directions will be considered successful;

• harness the human brain’s innate perceptual and cognitive tendencies to optimize your visualizations and other people’s visualizations;

• employ wireframing and storyboarding to receive iterative feedback on visualizations;

• create the most commonly used graphs in Tableau;

• design effective data dashboards with Tableau;

• practice creating logical, persuasive stories with data.

We will also extend what is learned in this course to the topics of making slides and giving presentations during selected Friday workshops.

*Course Structure*

Like many aspects of data science, the best way to learn how to communicate about data and use data visualization software is to practice…a lot. This course will have many assignments to help you do that. You will be given short reading assignments to complete before most classes, and we will have quizzes at the beginning of most classes on that reading material. Class time will be dedicated primarily to discussing and applying principles you read about to real scenarios and analyses. You should practice the technical aspects of implementing the principles primarily outside of lecture times, and ask questions through discussion boards and during scheduled Q/A sessions. The course has both individual assignments and group assignments. The assignments will include a diverse set of activities, including (but not limited to) watching videos I created to teach you how to use Tableau, exercises in Tableau, quizzes based on analyses you complete using real data sets in Tableau, and case studies. When you complete exercises as a group, *make sure every person in the group knows how to complete the exercise on their own*, because there will be individual assignments that build on the skills practiced in the group activities. The class will culminate with a group final project due on the day of our assigned final exam.

*Preparing for Class and Classroom Norms*

I am dedicated to providing you with opportunities to actively test out data communication and visualization concepts. To benefit from these opportunities, it is critical that you engage fully in all class discussions and activities. Please be prepared to speak up, participate, and get feedback! We will be covering a large amount of material that can only be mastered with practice, and you will only get that practice if you are committed to getting the most out of each assignment and exercise, and push yourself to take risks and learn how to improve. I expect you to make mistakes; if you don’t make mistakes, you probably are not challenging yourself enough. In order to provide a place for you to feel comfortable making mistakes, it is critical that we maintain a collegial, supportive environment in class. Feedback and comments should always be given with the intention of helping each other grow, learn, and get better. Please contact me if you struggle with giving or receiving feedback constructively.

You will be expected to arrive at class on time with all assignments completed. Due to the dynamic nature of the classroom activities, there will be no mechanism for making up classes, but you will be provided with recordings of classes if you have a reasonable justification for missing class.

One of my pet peeves is seeing students on their phones or checking social media on another device during class. It is distracting to me and your classmates, and makes it difficult to have engaged discussions. I also find it disrespectful. Please step out of the classroom if you have a matter that requires you to use your cell phone. In addition, be prepared for me to call you out if your social media activities distract me or others in class. If you have a special circumstance that requires you to check your cell phone or social media during class, please speak to me before class begins.

*Honor Code*

The Duke Community Standard applies to all aspects of this course (see https://trinity.duke.edu/undergraduate/academic-policies/community-standard-student-conduct). Although the curriculum will strongly encourage you to use outside resources to explore aspects of Tableau not explicitly covered in class or assignments, **you are not allowed to use resources that directly give you the answers to problem sets or quizzes**. If you are uncertain about the nature of the consultation you can use for any given assignment, please ask me in advance. **No quizzes, assignments, or cases are to be discussed with future years of MIDS students without my permission.**

*Accommodations*

If you need special accommodations due to disabilities, medical needs, religious practices, or other reasons, please inform me as soon as possible to make a plan to accommodate those needs.

*Schedule, Assignments, and Final Project*

We will be using the Sakai website and Piazza for all course-related communications and discussions. Instructions and due dates for all assignments will be posted on the course Sakai website.

The work schedule I suggest you follow is provided in a google doc that you can reach through Sakai. Note that I am making the official due date for assignments many days later than I think they should actually be completed to give you flexibility. This will give you wiggle room if many people on your team have to travel, but it is also very risky because you will not learn the material at the level you need to if you do not mostly follow the schedule I suggest. ***I strongly recommend that you schedule regular times to meet with your team each week*** so that you know you will be able to complete the assignments in the time frames I suggest. Some of the team assignments are time-consuming, so ***plan ahead***!

Although the lectures for this class end on Feb. 27, the Final Project will be due on Wednesday, April 29 (during our scheduled Final Exam period). The Final Project will be completed as a team, and will consist of a presentation, white paper, and visualizations. You will be given a data sets from real stakeholders to analyze. You will be asked to create visualization tools and make recommendations about what the team should do next using visualizations to support your points. The analyses are complex, **so please start working on your final projects early.**

*Grading*

Course grades will be based on the following weighting scheme:

Assignments (Individual): 25%

Assignments (Team): 30%

In-class quizzes and other class participation: 10%

Final Project: 35%

Many assignments will be graded on a 3 point scale. In these cases, the most common score is 3 points. A score of 3 points represents a submission that meets all expectations with respect to discussion, analysis, and directions. Scores below 3 points represent submissions that are in some aspects incomplete or inaccurate. Scores above 3 points are considered extra credit, and indicate your discussion and analysis exceed expectations.

The case studies will be worth a lot more points than the Tableau assignments, so take them seriously and plan on spending a significant amount of time on them (since the case studies are completed in teams, that means you will also need to schedule and coordinate carefully with your team). The case studies are meant to integrate everything you are learning in class; you will be expected to think about them deeply and provide thoughtful, thorough solutions.

*Course Materials*

There will not be a course packet or required textbook for this course. All of the assignments, materials, videos, and data sets used for the course will be available on Sakai course website. We will be using the Sakai website frequently, so please refer to the website on a regular basis. You are responsible for monitoring the Course announcements.

Although no textbooks are required for this course, the following books are useful resources about data visualization:

Storytelling with Data

*Cole Nussbaumer, Wiley (2015)*

Information Dashboard Design: Displaying Data for At-a-Glance Monitoring

*Stephen Few, O’Reilly Media (2013)*

The Visual Display of Quantitative Information

*Edward Tufte, Graphics Press, 2nd Edition (2001)*

The following blogs and websites are also helpful:

* The Information Lab Blog (https://www.theinformationlab.co.uk/category/blog/)
* Makeover Monday Project by Andy Kriebel and Eva Murray (http://www.makeovermonday.co.uk/)
* University of Washington Interactive Data Lab (http://idl.cs.washington.edu/about http://idl.cs.washington.edu/about)
* VizWiz by Andy Kriebel (http://www.vizwiz.com/)
* Storytelling with Data by Cole Nussbaumer (http://www.storytellingwithdata.com/)
* Data Revelations by Steve Wexler (http://www.datarevelations.com/)
* Perceptual Edge by Stephen Few (http://www.perceptualedge.com/blog/)
* Flowing Data by Nathan Yau (the Tutorials section; http://flowingdata.com)
* Evergreen Data by Stephanie Evergreen (http://stephanieevergreen.com)

You may enjoy this resource for using R and Python visualization tools:

“Data Visualization: A practical introduction” by Duke’s own Kieran Healy

https://socviz.co/index.html#preface

We will also have a Piazza discussion where we can collectively post other resources that are useful for learning different data visualization packages.

*Getting Help*

Please ask all technical or conceptual questions about the course content through Piazza rather than through email. The discussion forums will be checked approximately every 24 hours, but we strongly encourage you to help your classmates and respond to their questions whenever possible.

We also strongly encourage you to use outside resources to help you practice using Tableau

throughout the course. The Tableau Community Website (https://www.tableau.com/community)

provides a wealth of tutorials and is a great place to ask specific questions. If you would like to try

something that hasn’t been taught explicitly through the course materials, a quick web search will also

likely lead you to supplementary instructions. Additionally, you can make appointments with the course

TAs or Duke Library Data and Visualization Services (https://library.duke.edu/data/data-visualization) for

targeted help with specific issues related to visualizing in other platforms.