Georgia Institute of Technology

Spring 2024 MGT 6655 Business Data Preparation and Visualization Syllabus

Program/School: Business Analytics Graduate Certificate/Scheller College of Business
Delivery: 100% Web-Based, Asynchronous
Dates course will run: January 8th - February 25th, 2024

Instructor Information

Mingfeng Lin, PhD.

- Please use private messaging on Ed Discussions to contact me during the semester
- Weekly office hours via Zoom. Please see Canvas for dates and times.
  NOTE: Office hours of this class will NOT be automatically recorded because office hours are optional for students. Absolutely no new content will be introduced, and no discussions will be initiated by the instructor or TA during office hours. It is only intended to answer questions students may have. If the answers to questions raised may be of use or of interest to the rest of the class, the instructor/TA will post them as additional Q&As, or record additional videos, and share them on Canvas. You will not be at any disadvantage if you do not see the need to join the office hour live.

Teaching Assistants: Mr. Jae Sang Rhee, PhD Student, Scheller College of Business; and ____ (TBD)

General Course Information

Description

Visualizing data is an important step in understanding data, exploring relationships, and "making a case." With the abundance and relevance of data in almost any type of work, the ability to understand and interpret data has become an indispensable business skill. Data visualization has become a fascinating industry, but the ease of creating visualizations also means that there are many charts out there that are wrong, ineffective, or misleading. The goal of this class is to introduce principles and tools of data visualizations, and to help you create visualizations for two different but related purposes: (1) exploration; and (2) explanation / presentation. "Exploration" is about helping you or other stakeholders understand data and form an opinion, by creating simple charts or building a dashboard. "Explanation," in contrast, is when you try to use visualizations to convince others of your opinion.

Prerequisites

Students should have basic understandings of undergraduate statistics such as statistical distributions, aggregations, and regression analyses; and be familiar with the basic operations of Microsoft Excel such as data filtering, sorting, and the use of formulas.

Course Learning Outcomes

By the end of the course, you will be able to:

(1) Identify the proper visualization for different use cases.

(2) Prepare, clean, and transform data for visualization purposes.
(3) Create visualizations using the tools covered in this class.
(4) Use visualization tools effectively to explore data, such as creating a dashboard, and interpret what you see.
(5) Use visualizations effectively to "make a case" or explain findings; and
(6) Critique visualizations presented by others (pros, cons, and suggestions for improvement).

Course Materials

Course Texts (Recommended)

1. The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios. ISBN-13: 978-1119282716
   - Available online through Georgia Tech library, https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA51310105940002947
   - Companion site (Data, workbook etc. to download): http://www.bigbookofdashboards.com/dashboards.html
2. Better Data Visualizations. ISBN 978-0231193115
   - Available online through Georgia Tech library, https://galileo-gatech.primo.exlibrisgroup.com/permalink/01GALI_GIT/1vc2ggp/alma9914891182802947

Additional Materials/Resources (Optional)

The following are the books that I recommend if you would like to pursue Tableau beyond this class:

1. Visualize this! By Nathan Yau. Available online through Georgia Tech library (https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA51158412810002947)
2. Tableau Your Data! By Daniel G. Murray. Available online through Georgia Tech library (https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA51300450890002947)
3. Practical Tableau, by Ryan Sleeper. Available online through Georgia Tech library (https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA51351229560002947)
4. Information Dashboard Design: Displaying Data for At-a-Glance Monitoring, by Stephen Few. Available online through Georgia Tech library (https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA21153492250002947).

Course Website and Other Classroom Management Tools

This class will use Canvas to deliver course materials to online students. ALL course materials and activities will take place on Canvas.

Course Requirements, Assignments & Grading

Assignment Distribution and Grading Scale

Assignment	Weight
Week 1 Quiz	**10%**
Week 2 Quiz	**10%**
Week 3 Quiz	**10%**
Week 4 Quiz	**10%**
Week 5 Quiz	**10%**
Week 5 Exercise #1 (visualization)	**5%**
Week 6 Quiz	**10%**
Week 6 Exercise #2 (dashboard)	**8%**
Visualization appreciation/critique	**2%**
Evidence of class participation	**5%**
Final Exam (week 7)	**20%**

Grading Scale

Your final grade will be assigned as a letter grade according to the following scale:

A 90-100%
B 80-89%
C 70-79%
D 60-69%
F 0-59%

Description of Graded Components

- **Quizzes:** There will be six quizzes in this course, one in each of the first six weeks. These quizzes are either multiple choices questions, or fill-in-the-blank questions. Time limit may apply.
- **Exercises:** There are two exercises in the class that you need to work on and submit to the TA for grading. It is very important that for Tableau work, you must submit TWBX files instead of the TWB because the TWB does not have data embedded in it. Failure to submit the TWBX files will result in a loss of at least 1/3 of the credit because the TA will have to request you to send the right file. Partial credits will be provided, so do your best.
- **Final Exam:** The questions on the final exams are similar to the quizzes (multiple choices or fill-in-the-blank). The exam will be digitally proctored with Honorlock. More information on digital proctoring is available in Canvas.
- **Class Participation:**
    - **Visualization appreciation/critique (2%):** You must post twice during the first six weeks to the “visualization critique and appreciation” thread on the online discussion board by sharing a visualization (cite its source). One of your posts should be for a visualization that you like; the other should be for a visualization that

you think is poorly done. For the first one, describe why you think it is great and what we can learn from it. For the second one, discuss why you think it's bad – and how we can make it better. You need to apply what you learned about visualization principles (slides & first chapter of textbook) and chart type choices (see slides as well as the textbook chapter).

- **Evidence of class participation (5%).** Throughout the course, it is very important for you to participate in the class discussions on Ed Discussions. The instructor may also occasionally post questions there for you to answer. For those questions, you're expected to work on all of them because they are designed to enhance your learning. You are not expected to write answer on Ed Discussions for all of them, but you should post an answer to at least half of those questions. It's okay if you post your message anonymously. Please note that the instructor may not be able to reply to every message but your participation will be recorded.

Extra Credit Opportunities for CIOS Participation

The CIOS (Course Instructor Opinion Survey) that you will fill out at the end of the semester is an important channel for instructors to receive feedback and suggestions about their class. Some of the contents you see in this class were added because of past student feedback. This extra credit is intended to encourage you to respond to the CIOS survey request so that we have a complete and unbiased picture of how the course is doing. A few important points to note:

- I will not see any CIOS results until 5 days after the deadline for posting the final grades.
- I DO NOT ask for individual proof that you submitted the survey (please see below the computation scheme for extra-credit)
- All CIOS answers are anonymized when shared with instructors
- **IMPORTANT**: I DO NOT offer this extra-credit opportunity in hopes of receiving a certain type of review. The extra-credit is only offered as an incentive to **participate**. I am looking for honest constructive feedback - what worked? what could use some improvement (and how can things be improved)? Future students will benefit from this feedback.
- Depending on the proportion of students who **fill out the CIOS by 2/25/2024**, all students in your section will receive up to 1 point (1% of the final grade) based on the following table. You can still answer CIOS after that date and I will still read your feedback, but it will not count toward the extra credit.

If this proportion of students in your section answer the CIOS survey **by 2/25/2024**…	**Everyone** will receive this amount of extra credit…
0-4.99%	0
5-14.99%	0.1
15-24.99%	0.2
25-34.99%	0.3
35-44.99%	0.4
45-54.99%	0.5
55-64.99%	0.6
65-74.99%	0.7
75-84.99%	0.8
85-100%	1.0

Submitting Assignments

The exercises and quizzes must be completed and submitted in Canvas by the due date stated in the course schedule. The final exam needs to be taken in Canvas during the exam window.

Assignments and Tests: Independence & Time Limits

The assignments and tests in this class are all open-everything (book, materials, internet) with one important caveat: You must complete the assignments independently. You must not ask someone else (any live person) for help to complete your assignments or tests. You can ask the instructor/TAs for clarifications about the questions or requirements, but you cannot ask them to complete the tasks for you. In addition, the students must note that the assignments and tests are timed. While the time limit is set generously and should be more than sufficient, we assume that you have thoroughly gone through the class materials (recordings and demos) *before* starting to work on the assignments and tests. If you are only watching the demos for the first time when taking the assignments and tests, you will likely run out of time.

Generative AI Policy

Generative AI such as ChatGPT is a powerful tool for learning purposes if used properly. You can certainly use it to find answers to general questions, with the following boundaries: (1) You are explicitly forbidden from sending anything you see in the assignments/quizzes/tests to ChatGPT because they are copyrighted, and they should not be shared with any third party, which includes ChatGPT or other similar tools. This is the same policy if you work for a company; you would be liable to share company info with ChatGPT. (2) You should do your best to work independently on the assignments first, before turning to ChatGPT for help. You do not need to use it at all, but if you want to use it, use it after you have at least tried. (3) ChatGPT and other generative AI tools may not always give you the correct answer.

Assignment Due Dates and Time

All deadlines above are in the US Eastern Time Zone. Be sure to refer to the table header for the specific time of the day. These times are subject to change so please check back often. Please convert from UTC to your local time zone using a Time Zone Converter.

Late Submission Policy

Submissions that are late but turned in within 24 hours of deadline will be penalized at 10% of the grades received for that assignment. Submissions that are late between 24 hours after the deadline but before 72 hours of the deadline will be penalized at 20%. Those that are late more than 72 hours will receive no credit. No late submission will be accepted on the Final Exam.

Timing Policy

- The Modules follow a logical sequence that includes knowledge-building and experience-building. It is strongly recommended that you take the class according to this schedule.
- Assignments should be completed by their due dates.
- You will have access to the course content for the scheduled duration of the course.

Grading and Feedback

Results of your quizzes will be released the day after the deadline. Results of the two exercises will be released once the TA has finished grading, which may vary based on class size.

Technology Requirements and Skills

Network Connectivity

High-speed Internet connection

Computer Hardware

Laptop or desktop computer with a **minimum** of a 2 GHz processor and 2 GB of RAM, and a minimum of 2GB hard drive space. Please note: Chromebooks will *not* be sufficient.
While both PC and Mac should work, windows PC is strongly preferred. Mac users will have to use GT's VLAB when they need to use software that cannot be installed on Mac. Since VLAB is web-based, please be sure to have a **stable** high speed internet connection, and also take into account the fact that it will be slower to respond when many other users log in at the same time. You can access the VLAB by visiting https://mycloud.gatech.edu. Log on with your GT credentials. Click on **CoB-Labs**.

IMPORTANT: The instructor and the TA do not host or manage VLAB. If you have trouble with VLAB, please contact Digital Learning Support (https://b.gatech.edu/digitallearningsupport) for assistance. The instructor and TA cannot help. Again, a PC will be a better option.

Software

1. **Microsoft Excel**, preferably the Windows version but the Mac version should also work. If you don't have this on your computer, you can access it through Office365 (available to all GT students) or you can use the VLAB.
2. **Tableau Desktop and Tableau Prep** (available for both Mac and PC)
    - Request your student license for Tableau and Tableau Prep at https://www.tableau.com/academic/students
    - Go to https://www.tableau.com/tft/activation and download both Tableau Desktop (visualization software) and Tableau Prep (data preparation tool). Use the license you obtained from the previous step.
    - Please note that Tableau Desktop is different from the free Tableau Public (available on VLAB at the time of writing). Tableau Public does not require a license, but you must know that *Tableau Public offers no privacy for data and is intended for use on public datasets* --- if you use your own data you must make sure it does not contain any private or confidential information.
3. **Microsoft PowerBI:** Microsoft PowerBI is another very popular data visualization tool that you may want to practice and we will talk about it later in the class. Its desktop version is free. The link to download is at https://powerbi.microsoft.com/en-us/desktop/. While there is an online version, we will use the free desktop version for the purpose of the class.
4. **NodeXL Basic** for network data visualization.
    - You only need the basic version for this class. Register using the link below and you will receive a link to download the basic version: https://www.nodexlgraphgallery.org/Pages/RegistrationBasic.aspx
    - NodeXL Basic is for PC only. If you have a Mac, you can install it on VLAB.
5. **R (open source) and R studio (free version):**
    - R: http://archive.linux.duke.edu/cran/ or any other mirror site listed on https://cran.r-project.org/mirrors.html

- R Studio: https://rstudio.com/products/rstudio/download/#download

6. Mozilla Firefox, Chrome and/or Safari browsers. **Please note that** if you run into trouble in a browser, a simple fix is to clear your browser's cache, or use "private browsing" ("incognito" in Chrome) which creates a sandbox where everything is fresh. Try this first before requesting assistance.
7. Microsoft Office software (recommended: Office 365 provided by GT)
8. Adobe Acrobat Reader (for PDFs)

Technology Help Guidelines

30-Minute Rule: When you encounter struggles with technology, give yourself 30 minutes to 'figure it out.' If you cannot, then post a message to the discussion board; your peers may have suggestions to assist you. You are also directed to contact the Helpdesk 24/7.

When posting or sending email requesting help with technology issues, whether to the Helpdesk, or the message board, use the following guidelines:

- Include a descriptive title for the subject field that includes 1) the name of course 2) the issue. Do NOT just simply type "Help" into the subject field or leave it blank.
- List the steps or describe the circumstance that preceded the technical issue or error. Include the exact wording of the error message.
- When possible, always include a screenshot(s) demonstrating the technical issue or error message.
- Also include what you have already tried to remedy the issue (rebooting, trying a different browser, etc.).

Course Policies, Expectations & Guidelines

Communication Policy

- Do NOT submit posts of a personal nature to the discussion board unless it is a private post on Ed Discussions.
- Use private messaging on Ed Discussions to contact me during the semester. During the week, you can expect to receive a response within 24 hours; on weekends and holidays, allow up to 48 hours. If there are special circumstances that will delay the response from the TA or instructor, we will make an announcement.
- Student Forum/Q&A discussion boards will be checked twice per day Monday through Friday; Saturday and Sunday, these discussion boards will be checked at least once.
- Virtual office hours will be held using the Zoom. See my Virtual Office Hours at the beginning of this document.
- For questions related to technology, please contact: https://b.gatech.edu/digitallearningsupport

Online Student Conduct and (N)etiquette

Communicating appropriately in the online classroom can be challenging. In order to minimize this challenge, it is important to remember several points of **"internet etiquette"** that will smooth communication for both students and instructors:

1. *Read first, Write later*. Read the ENTIRE set of posts/comments on a discussion board before posting your reply, in order to prevent repeating commentary or asking questions that have already been answered.

2. *Avoid language that may come across as strong or offensive*. Language can be easily misinterpreted in written electronic communication. Review email and discussion board posts BEFORE submitting. Humor and sarcasm may be easily misinterpreted by your reader(s). Try to be as professional as possible.
3. *Follow the language rules of the Internet*. Do not write using all capital letters, because it will appear as shouting. Also, the use of emoticons can be helpful when used to convey nonverbal feelings. ☺
4. *Consider the privacy of others*. Ask permission prior to giving out a classmate's email address or other information.
5. *Keep attachments small*. If it is necessary to send pictures, change the size to an acceptable 250kb or less (one free, web-based tool to try is picresize.com). For large files such as datasets that's more than 10mb in size, consider using links (e.g., through OneDrive or Dropbox) instead of uploading actual files.
6. *No inappropriate material*. Do not forward virus warnings, chain letters, jokes, etc. to classmates or instructors. The sharing of pornographic material is forbidden.

***NOTE**: The instructor and TAs reserve the right to remove posts that are not collegial in nature and/or do not meet the Online Student Conduct and Etiquette guidelines listed above.*

University Use of Electronic Email

A university-assigned student e-mail account is the official university means of communication with all students at Georgia Institute of Technology. Students are responsible for all information sent to them via their university-assigned e-mail account. If a student chooses to forward information in their university e-mail account, he or she is responsible for all information, including attachments, sent to any other e-mail account. To stay current with university information, students are expected to check their official university e-mail account and other electronic communications on a frequent and consistent basis. Recognizing that some communications may be time-critical, the university recommends that electronic communications be checked minimally twice a week.

Plagiarism & Academic Integrity

Georgia Tech aims to cultivate a community based on trust, academic integrity, and honor. Students are expected to act according to the highest ethical standards. All students enrolled at Georgia Tech, and all its campuses, are to perform their academic work according to standards set by faculty members, departments, schools and colleges of the university; and cheating and plagiarism constitute fraudulent misrepresentation for which no credit can be given and for which appropriate sanctions are warranted and will be applied. For information on Georgia Tech's Academic Honor Code, please visit https://policylibrary.gatech.edu/student-life/academic-honor-code.

Any student suspected of cheating or plagiarizing on a quiz, exam, or assignment will be reported to the Office of Student Integrity, who will investigate the incident and identify the appropriate penalty for violations.

Copyright

You are explicitly forbidden from sharing any of the course materials including the slides, the video recordings (or their audio, screenshots, or transcripts), the samples (e.g., Tableau workbooks), the exam questions and answers, assignment questions and answers, with anyone who is not

registered in the class. This includes generative AI tools such as ChatGPT. You are not allowed to upload them to any websites, servers or social media without the explicit prior written consent of the instructor and GTPE.

Accommodations for Students with Disabilities

If you are a student with learning needs that require special accommodation, contact the Office of Disability Services at (404)894-2563 or http://disabilityservices.gatech.edu/, as soon as possible, to make an appointment to discuss your special needs and to obtain an accommodations letter. You must also e-mail the instructor as soon as possible in order to set up a time to discuss your learning needs.

Deadline Extensions & Re-Scheduled/Missed Exams

Students with approved institute activities (http://www.catalog.gatech.edu/rules/4/) can request extensions for their deadlines. The instructor will also make an effort to accommodate other requests. However, given the compressed schedule of the class, the students should do their best to keep up with the course schedule, and make a sincere effort to catch up with missed deadlines as soon as they are able.

Student-Faculty Expectations Agreement

At Georgia Tech we believe that it is important to strive for an atmosphere of mutual respect, acknowledgement, and responsibility between faculty members and the student body. See http://www.catalog.gatech.edu/rules/22/ for an articulation of some basic expectation that you can have of me and that I have of you. In the end, simple respect for knowledge, hard work, and cordial interactions will help build the environment we seek. Therefore, I encourage you to remain committed to the ideals of Georgia Tech while in this class.

Subject to Change Statement

The syllabus and course schedule may be subject to change. Changes will be communicated via email, Ed Discussions, the Canvas announcement tool, and edX bulk email. It is the responsibility of students to check email messages and course announcements regularly to stay current in their online courses.

Course Schedule

Week	Topic	Overview	Deliverables	Due Date (by 11:59pm)	Readings
1	**Getting Started: Course Overview & Visualization Principles**	This week introduces student to the class and review basic statistics and data storage knowledge in preparation for the class, and also introduces high level data visualization principles.	Week 1 Quiz	1/14	All course videos Chapter 1 of Textbook (will be covered in the exercise) Visualization appreciation exercise
2	**Data Preparation**	Students will learn how to prepare data for visualization in various tools, particularly Excel and Tableau Prep Builder.	Week 2 Quiz	1/21	All course videos (must be able to follow and re-produce all demos)
3	**Data Visualization Types and Choices; Visualizing Data in Excel and Tableau**	Students will learn about the major types of data visualizations and when it is appropriate to use each. This is followed by creating some visualization types in Excel that illustrates Excel capabilities, and will also get started with Tableau	Week 3 Quiz	1/28	All course videos (must be able to follow and re-produce all demos) Visual Vocabulary
4	**Tableau Data Visualization (1)**	This week introduces basic visualization types in Tableau, and how to create them. Students are also expected to understand and interpret the results appropriately (i.e., what the visualization tells you).	Week 4 Quiz (Requested, not required): submit course feedback at https://bit.ly/bdpvfeedback	2/4	All course videos (must be able to follow and re-produce all demos)

Week	Topic	Overview	Deliverables	Due Date (by 11:59pm)	Readings
5	**Tableau Data Visualization (2)**	This week continues the journey of creating flexible visualizations in Tableau, and introduces more advanced visualization types and features	Week 5 Quiz Exercise #1 (visualization)	2/11	All course videos (must be able to follow and re-produce all demos)
6	**Tableau Data Visualization (3), Network Visualization, and Introduction to Programming for Visualization**	This week covers (1) the use of custom shapes in Tableau, and more importantly the creation of Tableau Dashboards; (2) the sharing of Tableau visualizations; (3) the concept of network visualization, and (4) programming for charts. The goal of introducing programming here is to give you the flexibility of creating charts that are not available off-the-shelf. You will just need to understand and tweak existing codes so that they can read your own data.	Visualization appreciation/critique Week 6 Quiz Exercise #2 (dashboard)	2/18	All course videos (must be able to follow and re-produce all demos)

Week	Topic	Overview	Deliverables	Due Date (by 11:59pm)	Readings
7	**PowerBI, Certifications, and Final Exam**	This week we will introduce PowerBI which is also used widely but generally less customizable or flexible than Tableau. Certification opportunities in Tableau will be discussed as well, and students will complete the final exam.	Re-produce the PowerBI demo Complete Final Exam Complete Course Instructor Opinion Survey (CIOS)	2/25	All course videos (must be able to follow and re-produce the PowerBI demo)