MGMT 4655 (4803)
Business Data Preparation and Visualization

Instructor

Prof. Mingfeng Lin, PhD.

Email : mingfeng.lin@scheller.gatech.edu

- *This is by far the best way to reach me. Typical response time 24 hours, often much sooner*

Class Location: SCOB4167
Office hour: 10-11am Tuesdays via Zoom at https://bit.ly/mlinzoom (Backup: MS Teams), or by appointment.

Teaching Assistant: Jiayu Yao (ITM PhD student), jyao@gatech.edu, 10-11am Thursdays via Zoom at https://gatech.zoom.us/j/95547537829, or by appointment.

Course 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 fast-growing 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 for this class are expected to have a computer with internet access. See details in "Other required resources" in Course Materials below.
- Students should have basic understanding of statistics such as statistical distributions and be familiar with the basic operations of Microsoft Excel.

Learning Objectives

Upon successful completion of the class, you should 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

Primary Materials (no need to purchase):

- **The Big Book of Dashboards: Visualizing Your Data Using Real-World Business Scenarios**. By Steve Wexler, Jeffrey Shaffer, and Andy Cotgreave. 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
- **Better Data Visualizations**, by Jonathan Schwabish, ISBN 978-0231193115
    - Available online through Georgia Tech library, https://galileo-gatech.primo.exlibrisgroup.com/permalink/01GALI_GIT/1vc2ggp/alma9914891182802947

Optional Readings / References:

- Visualize this! By Nathan Yau. Available online through Georgia Tech library (https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA51158412810002947)
- 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)
- Practical Tableau, by Ryan Sleeper. Available online through Georgia Tech library (https://gatech-primo.hosted.exlibrisgroup.com/permalink/f/1vgrnp4/01GALI_GIT_ALMA51351229560002947)

Software (all free for the class)

- **Zoom,** https://gatech.zoom.us/ -- this is our default software for virtual guest speakers etc.
- **Microsoft Teams.** https://faq.oit.gatech.edu/content/how-do-i-access-microsoft-teams
    - This is our backup virtual meeting software
    - It offers a chat function; you can reach the instructor if needed during designated hours
- **Microsoft Excel**, preferably the Windows version but the Mac version should also work. You can also access it through Office365 (http://office.com) or the VLab by visiting https://mycloud.gatech.edu. Log on with your GT credentials. Click on **CoB-Labs** in VLab.
- **Tableau Desktop and Tableau Prep Builder** (available for both Mac and PC)
    - Go to https://www.tableau.com/tft/activation and download Tableau Desktop (visualization software) and Tableau Prep Builder (data preparation tool). You can start using both in trial mode right away, but we won't start talking about Tableau until later (see course schedule).
    - You may request your student license directly from Tableau
        - https://www.tableau.com/academic/students
    - Contact the instructor if you have trouble receiving the license.
    - 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.
    - It is strongly recommended that you install Tableau Desktop on your own PC or Mac, rather than on lab computers.
- **Microsoft PowerBI:** Microsoft PowerBI is another very popular data visualization tool that you might want to practice. It is easy to use but with less flexibility than Tableau. Its desktop version is free. https://powerbi.microsoft.com/en-us/desktop/.
- **NodeXL Basic**: You don't need to worry about this until much later in the semester.
    - 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
    - This is also available on the computers in the trading floor classroom, and VLAB.

- **R (open source) and R studio (free version):** Don't worry, you will only do minimal programming for this class. You also don't need to worry about these installations until much later in the semester.
    - http://archive.linux.duke.edu/cran/ or any other mirror site listed on https://cran.r-project.org/mirrors.html
    - https://rstudio.com/products/rstudio/download/#download
        - Choose the correct version for your operating system
    - We will walk you through the installation process during class.

Class Participation

1. Pre-recorded videos of all classes and demos will be available to you each week at the time of each class or soon after. If new materials arise from discussions or student questions, I will also record them and make them available to everyone. For this reason, to ensure the safety of everyone, if you feel sick in any way, please do not feel like you have to come to the class. You will not fall behind if you take advantage of these videos. Be sure to refer to GT's guidelines regarding COVID-19 isolation and quarantine at https://www.health.gatech.edu/coronavirus/isolation-quarantine. In particular (as of August 22, 2022)
    a. If you get sick and test positive: please stay home for at least 5 full days, and wear a mask after that for another 5 days when around others
    b. If you get sick and test negative: please still be careful about contact with others.
2. If you need to isolate and want to join live, I'm happy to live stream the class via Zoom.
3. Due to the technical nature of this class, pre-recorded videos greatly complement live demos because each student learns differently. The live demos can be too fast for some students and too slow for others. These videos allow you to replay on demand, and contribute to better learning outcome for you. They are therefore an integral part of the class.
4. Please note that these pre-recorded videos are also our "plan B" if there is any unanticipated class disruption during the fall semester.
5. You are welcome to use these videos for your personally learning purposes. You must not, however, post these videos online or share them with anyone outside the class.
6. It is very important that you work closely with your teammates and fully participate**. I understand that you are all very busy, so I strongly recommend that you set up group meeting schedule at the beginning of the semester and put it on your calendars.** Otherwise it will be difficult for you to find a good time to meet later in the semester. **Please also refer to the later section on "Peer Evaluation" so you understand the grading aspect clearly.**
7. **Technical Difficulties or Emergencies:** I will provide alternative links to recorded contents in the rare event that Canvas is down – in that case, please send an email to the instructor and the TAs and we will share the backup URLs with you. If interruptions occur during timed events such as tests, please document as much as you can (take screenshots), and I will do my best to accommodate.

Course Grading

Course grades will be determined based on 100 points from the following items:

Item	Weight
Individual effort: Assignments (6)*	42%
Group project: milestone 1 (topic and data)	5%
Group project: milestone 2 (draft TWBX)	5%
Group project: final deliverable	12%
Individual evaluation of other teams' reports	6%
Group Project: Peer evaluation (of your teammates)	5%

Midterm	10%
Final (take home)	15%
TOTAL	**100%**

* You will see 7 assignments listed on the course schedule, but you can skip one without penalty.

Special note about Canvas: The table above reflects the "official" way of calculating your grade in this class. Do not rely on the total score that you see on Canvas.

The final course grade breakdown is as follows:

Points	Percentage	Letter Grade
90 and above	90% and above	**A**
80-89	80%-89%	**B**
70-79	70%-79%	**C**
60-69	60%-69%	**D**
Lower than 60	Lower than 60%	**F**

Assignments and Assessments

Group Project

Students will self-organize into **teams of 5-7 people**[1]. If you cannot form such teams on your own, we will assign you randomly. You will work with teammates to (1) create a dashboard for a stakeholder of your choice (where you build the dashboard as a tool, and you have not formed opinions of your own from the data), and (2) create a "data story" or make a case using a series of charts (where you have formed an opinion from your knowledge of the data and are now trying to persuade others), from the **datasets** that the instructor makes available to you or from your own dataset. **These are two separate, unrelated tasks.** Additional details will be provided for the group project in a **separate document**. Since this class is for MBA students, I expect that you use some regression analyses and/or predictive analytics if you have done so in previous classes.

Your final deliverables will be evaluated by both the instructor and some of your classmates in other teams (after removing outliers).

Assignments (Individual Effort)

There will be six individual-effort assignments throughout the semester (such as working on creating a specific visualization that the instructor asks you to), and you must know how to complete them correctly by yourself. These are typically extensions of the course demos. You are allowed to discuss with your classmates, but **each student should still write their assignment independently**. If you discuss an assignment with a classmate before you submit it, you must clearly mention their names in your assignment. **Identical submissions are not allowed**.

You should be able to complete the assignments on your own computer. Alternatively, you may also use the MyCloud.gatech.edu virtual computers. **If you use the VLAB to complete your assignments, please make sure that you save your progress frequently, or send to your own email inbox for backup purposes.**

Midterm and Final

There are two tests in the semester, one mid-term and one final. Questions will typically include (but are not limited to) creating and interpreting charts; critiquing an existing visualization and providing suggestions for improvement, based on the principles discussed in class; and envisioning / planning a visualization to either explore the data, answer a question, or present a relationship. These tests will be on Canvas.

[1] We ask that you form teams of at least 5 to ensure our capacity of giving you meaningful feedback. If you have special circumstances that require smaller teams please check with the instructor.

Peer Evaluation (of Your Teammates)
While most of our MBA students are great with teamwork, it is important that we provide explicit incentive for collaboration. This is the rationale for the peer evaluations. I encourage you to share your feedback about your teammates as soon as possible after issues arise. You can do that within your team or involve me if needed. The earlier we identify issues the more likely that we can resolve the issues and ensure a great team environment.

When you evaluate your teammate, you must list relevant facts (with date and time etc.) that show violations of the team contract, especially when you rate someone lower than 80% (out of 100%). If multiple teammates list the same facts and they all rate one student lower than 80%, and the facts are verified, then that student will lose points proportionally on the peer evaluation part. Points awarded to individual students will be discounted proportionally if a member does not substantially contribute to the group product. In addition, **you must submit your evaluations of your peers before you can receive your grade on this item**.

Individual Evaluation of Other Teams' Reports
We often learn a great deal from our classmates. To that end, once the submission deadline for the group project's final deliverable has passed, the TA will collect the submissions and share that with the rest of the class. You (not as a team) will be asked to evaluate their work and provide some constructive feedback. This part of the grade reflects whether you take this seriously – how high (or low) you rate the other teams' presentations will not affect this part of the grade, as long as you provide meaningful critiques.

As an extra bonus, if you take your own critique/suggestion one step further and execute it, and show meaningful improvement over a team's presentation, it counts as extra credit for your own grade (up to 2 percentage points). See course schedule for details.

Late Assignment Policy

Due to the hands-on nature of this class, to earn a passing grade in the course you will need to complete and submit every assigned assessment before the final exam period (except the one that you choose to skip), **even if** this implies submitting an assignment late and earning reduced points toward the final grade. If you expect to be absent or unavailable around the due dates, please do your best to finish the assignments beforehand. Late submissions will receive penalty based on the extent of your delay, up to the full credit of that assignment.

Meanwhile, the instructor understands the challenges presented by the ongoing pandemic and will make every effort to accommodate reasonable requests. Please note that for **medical excuses, do NOT submit your doctor's notes to the instructor or the TA.** You should submit those notes to the Dean of Students, and we will be notified. Out of respect for your privacy, **we do not need to know, and will never ask about, your medical conditions.**

It is not just medical issues. If you have other challenging circumstances that prevent you from completing assignments on time, **please reach out to the instructor as early as you can**. We will help you in any way we can while preserving the integrity of the class.

Policies and Procedures

Academic Honesty / Honor Code: The Georgia Tech Honor Code is in force. The complete text is on the Honor Advisory Council website: http://www.honor.gatech.edu. **You are expected to be aware of actions that constitute cheating, fabrication or falsifications, multiple submissions of essentially the same work for different classes, plagiarism, and complicity in academic misconduct.** Note that academic misconduct includes "Submission of material that is wholly or substantially identical to that created or published by another person or persons, without adequate credit notations indicating authorship," (plagiarism). Suspected cases of academic misconduct are investigated by the Office of Dean of Students. The Honor Code is an important mechanism to protect the value of the degree that you are working hard to earn.

Official Communications: Both Canvas and the class email distribution list are used to post information on class changes, upcoming assignments, grades, etc. Please make sure that you can receive notifications from Canvas. Communications will also be directly sent to your Georgia Tech student email (which ends in @gatech.edu), especially during Canvas outages. It is ***your*** responsibility to have the Georgia Tech email account active and to receive and regularly read messages written to that address. **If you do not use your Georgia Tech email address, you should have your GT email forwarded to an account that you check several times each day and make sure it is not filtered as spam.**

Accommodations for Disabilities: If you have established accommodations with the Offices of Disability Services, please communicate your approved accommodations to me at your earliest convenience so we can discuss your needs in this course.

If you have a need but have not yet established accommodations through Disability Services, you should contact Disability Services at 404.894.2563 or dsinfo@gatech.edu or http://disabilityservices.gatech.edu. Disability Services offers resources and coordinates reasonable accommodations for students with disabilities and/or temporary health conditions. Reasonable accommodations are established through an interactive process between you, your instructor(s) and Disability Services.

Should a new disability be identified during the term, the instructor will work with Disability Services to accommodate your needs from that time forward.

Other accommodations, for example those allowed under the ADA for documented conditions, might be granted in the class if they do not interfere with the normal conduct of the class, such as the posing of daily questions and quick provision of feedback, or if they do not require exemption from a crucial element of the class such as attendance, participation in a team, or making presentations.

If you have a color deficiency, please let the instructor know as well. This is a visualization class so we will use colors extensively, and I will not penalize you if this affects your choices of colors.

Changes to this document: Changes to any part of this document may be made to reflect changing situations. They will be announced in class or via email and posted online.

Campus Resources for Students

In your time at Georgia Tech, you may find yourself in need of support. Below you will find some resources to support you both as a student and as a person. ***If you believe that your performance is being impacted by events occurring outside of this class, please don't hesitate to contact me as I want to be a resource for you.***

Academic support

- Center for Academic Success http://success.gatech.edu
  - 1-to-1 tutoring http://success.gatech.edu/1-1-tutoring
  - Peer-Led Undergraduate Study (PLUS) http://success.gatech.edu/tutoring/plus
  - Academic coaching http://success.gatech.edu/coaching
- Residence Life's Learning Assistance Program https://housing.gatech.edu/learning-assistance-program
  - Drop-in tutoring for many 1000 level courses
- OMED: Educational Services (http://omed.gatech.edu/programs/academic-support)
  - Group study sessions and tutoring programs
- Communication Center (http://www.communicationcenter.gatech.edu)
  - Individualized help with writing and multimedia projects
- Academic advisors for your major: http://advising.gatech.edu/

Personal Support

- The Office of the Dean of Students: http://studentlife.gatech.edu/content/services; 404-894-6367; Smithgall Student Services Building 2nd floor
    - You also may request assistance at https://gatech-advocate.symplicity.com/care_report/index.php/pid383662?
- Counseling Center: http://counseling.gatech.edu; 404-894-2575; Smithgall Student Services Building 2nd floor
    - Services include short-term individual counseling, group counseling, couples counseling, testing and assessment, referral services, and crisis intervention. Their website also includes links to state and national resources.
    - *Students in crisis may walk in during business hours (8am-5pm, Monday through Friday) or contact the counselor on call after hours at 404-894-2204.*
- Students' Temporary Assistance and Resources (STAR): http://studentlife.gatech.edu/content/need-help
    - Can assist with interview clothing, food, and housing needs.
- Stamps Health Services: https://health.gatech.edu; 404-894-1420
    - Primary care, pharmacy, women's health, psychiatry, immunization and allergy, health promotion, and nutrition
- OMED: Educational Services: http://www.omed.gatech.edu
- Women's Resource Center: http://www.womenscenter.gatech.edu; 404-385-0230
- LGBTQIA Resource Center: http://lgbtqia.gatech.edu/; 404-385-2679
- Veteran's Resource Center: http://veterans.gatech.edu/; 404-385-2067
- Georgia Tech Police: 404-894-2500

Change log:

1. August 21st first published version