DUKE FUQUA SCHOOL OF BUSINESS

Quantitative Management Programs

Data Competition

Background:

Dr. Shelly D. Farnham
"I do not know how you teach someone to love to learn, but being self-motivated is integral to this field. Once you have the core concepts, to be able to be really excited about and continue to seek out new information is something that I look for…when we are recruiting people."

Gregory Piatetsky-Shapiro
"You can best learn data mining and data science by doing, so start analyzing data as soon as you can!"

Claudia Perlich
"Learning how to do data science is like learning to ski. You have to do it."

As data science academic programs have continued to expand over the last 10 years, employers are becoming more selective in the candidates they recruit. More than ever, recruiters want to see evidence that candidates have inherent motivation, natural curiosity, and experience *doing* data science. As you will quickly find, every data science problem is unique and has its own set of challenges. There is an art to data science that can only be developed through trial-and-error, iteration, and collaboration with the larger data science community.

The Winter 2023 Data Competition is designed to give you the incentive and structure to begin developing a digital portfolio that can demonstrate your personal talents, curiosity, and passions to potential recruiters. The winning team will receive an award of $750 to help support its team members in their professional development.[1] The winning team will receive further instructions on award payment, including tax implications.

Project Scope\Deliverables:

The goal of this competition is to create a collection of digital projects that demonstrates your team's application of the data science techniques you are developing, broadly defined. Examples of projects that would achieve this goal can be found at:

- https://www.hannahyan.com/ (if you get a security warning you can ignore it)
- https://branwalker19.medium.com/
- https://davidventuri.com/portfolio (especially the Personal Projects section)

Please note that as you have just recently been introduced more advanced data science techniques, you should think of these projects as "bite-sized" as opposed to full analyses incorporating multiple techniques, nuanced consideration, etc. For example, projects demonstrating:

- "I learned a new skill. Here is a demonstration of what I learned."
- "I learned a new skill and wanted to apply it to something interesting."
- "I wanted to take a first pass at a question I found interesting."

Semi-Finalists will be based on your team members' digital portfolios. A team can submit a single team portfolio or links to each individual team member's portfolio. Please submit a Word document to Canvas (MQM:24 – Assignments - Winter Data Competition) with the links included and the names of everyone on your team. Teams selected for the Final round will also present their work to a panel of judges on January 30, 2024.

1 Please note this award will be subject to US Income Tax per IRS Guidelines.

DUKE FUQUA SCHOOL OF BUSINESS

Quantitative Management Programs

Data Competition

Project Constraints:

1. Teams can consist of, at most, 3 individuals. One member of each team is responsible for submitting the project deliverables to Canvas (MQM 24 – Assignments -Winter Data Competition). Please make sure each team member's name is on the front page.
2. The minimum number of unique digital projects that must be submitted is determined by your team size:
    - 1 team member – 3 projects
    - 2 team members – 4 projects
    - 3 team members – 5 projects
3. Teams having multiple members can work on each project individually, as a team, or some combination of the two. For projects in which you work with team members, make sure all contributors are cited.
4. Only 1 of your team's digital projects can have been created prior to December 1, 2023. In other words, all but 1 of your projects must be new.
5. Data used must be obtained via. the original source's Data Use Agreement, if applicable.
6. You may not scrape any data accessed using your NetID without prior approval from Carlton Brown (carlton.brown@duke.edu) in the Ford Library. To appropriately set expectations, most licenses unfortunately explicitly forbid utilizing scraping technologies.
7. As is common in industry, you may discuss specific techniques with anyone. Feel empowered to ask for help, research solutions, and discuss ideas with each other.
8. You may use any tool/software at your disposal.
9. You may design your digital portfolio using any platform you choose. While more complex than is necessary, one solution can be found at https://towardsdatascience.com/how-to-build-a-data-science-portfolio-website-335b0f253822.
10. Your submission is due on Sunday, January 21th at 11:59pm EST on the MQM:24 Resource site under Assignments, Winter Data Competition. If you are selected as a finalist, the final presentations will be on January 30, 2024. Each team will have 10 minutes to present their work, followed by judges' questions.

Competition Judging Rubric:

Criteria	Description
Variety	Your projects demonstrate a variety of approaches. At a minimum, you must demonstrate 3 different tools/techniques/etc. you have learned in your classes, on your own, etc.
Aesthetic	Your digital portfolio(s) is aesthetically pleasing and professional. It is designed with design considerations in mind, offers intuitive navigation, and is designed for the digital medium.
Originality	Your projects are original and interesting. For instance, they are not all based on Kaggle competitions (although one of them can be if it aligns with your interests). The projects reflect your individual personality and interests. **Note:** Recreating Github or Kaggle submissions will result in extremely low Originality scores.
Citations	Data sources, sources of inspiration, technical resources, etc. are all clearly cited, generally by including relevant links.
Technical Accuracy	The tools/techniques/etc. used are appropriate for your projects and utilized correctly.
Presentation	You follow data science blogging best-practices, which can be mostly found at https://towardsdatascience.com/thinking-of-blogging-about-data-science-here-are-some-tips-and-possible-benefits-680ff0e51d67. Your portfolio projects do not contain grammar mistakes/typos/etc.

Bonus points are given to projects that demonstrate a focus on business for good. Students are allowed to analyze historical, current, or potentially new ways in which businesses have contributed to or will contribute to the world to address societal issues.

Guest Judges:

Jiaming Xu, Associate Professor, Decision Sciences
Eric Weng, Analyst II, Strategy Analytics, Align Technology (MQM Alum)