Fall Data Competition at Virginia Tech
Hosted by CMDA Club
Sponsored by General Dynamics Mission Systems
In Collaboration with Computer Science
Monday, October 30, 2017 – Monday, November 6, 2017
Signup by Friday, October 13

Overview

The fall data competition is hosted by the CMDA Club at Virginia Tech. It is a week-long event from Monday, October 30, 2017 to Monday, November 6, 2017. Teams will receive the dataset on October 30, have 1 week to work on the analysis and presentation, and will present their conclusions November 6. Teams will consist of 2-3 members (undergraduate students) to ensure meaningful participation from every member.

Fall Data Competition vs. DataFest

The fall data competition was created to give students an opportunity to be hands-on in a data set. DataFest allows students to have fun, network with companies, peers, and faculty, and get experience with data analysis. The Fall Data Competition will also do this. What is different is that this competition emphasizes training and mentorship. There is a beginner track so that less experienced students have the ability to explore data analysis without the discouragement of being uncompetitive to the more advanced teams. More advanced students still have the ability to join a more competitive environment. The Fall Data Competition is around the time of internship and full time recruiting and interviews. Students will have this experience to talk about when looking for jobs.

Dataset

One or two datasets will be chosen from the VT library database. Depending on the difficulty of the dataset, two datasets can be used (a harder one for the advanced track and an easier one for the beginner track).

Tracks

For this competition, two tracks will be available for each team. The first track is the beginner track. The beginner track will receive an easier dataset and more direct analysis questions and goals. This will promote participation from students who are new to CMDA and out-of-major students.

The second track is the advanced track. The advanced track will receive a more complicated dataset and broader analysis questions and goals. This track is intended for students who have had experience in data analysis and competitions. The cutoff for the advanced track is students who are in or have taken CMDA 3654, Intro to Data Analysis, or equivalent.

Every member of each team must agree on one track for the team to take. Students who qualify for the advanced track will not be able to compete on a team in the beginner track.

Teams

Both the beginner and advanced track teams will have 2-3 members per team. All participants in the competition must be undergraduate students at Virginia Tech. Participation is open to students from any major. Students can either sign up as a team or individually. Individual signups will be randomly placed on a team with other individuals on the same track. Teams with members in both the advanced and beginner tracks are not encouraged. However, if it does happen, the team must register for the advanced track.
Judges

Depending on the turnout of the event, there will either be 3 or 5 judges per track (6-10 total). This event is expected to be as large as ASA Datafest. Since General Dynamics Mission Systems is sponsoring the event, they will have precedence in providing judges. If any more judges are needed, spaces will be filled with unbiased professors and grad students (those who did not help out teams during the competition).

Introduction Night

The introduction night will be on Monday, October 30. This will start at 6:30pm, since that is when most students are out of classes, and end around 9pm. Dinner and networking with GDMS will be from 6:30-7. A presentation will then be given by GDMS. After, CMDA club will explain the details and rules of the data competition. Both the beginner and advanced track teams will receive their dataset and instructions at this time. Poster boards and materials will be distributed.

Office Hours

We want to provide help during this week for teams. It is intended that a few professors and upper level students to hold 30 minute – 1 hour office hours at various times throughout the week for teams to come in and share their thoughts/problems. Specific information about office hours will come later.

Presentations

Presentations will be on Monday, November 6. We ask that students arrive around 5:30pm or earlier. There will be preliminary presentations to choose the top three teams of each track in no specific order. After, there will be final presentations of these finalist teams. The preliminary presentations will be in the form of a poster board session. Each team will have been given a trifold to fill with its slides/visualizations/analysis during the introduction night. The night of the presentations, the judges will walk around and preliminarily judge teams (who will be with their trifolds in a classroom) from 6:45-7:45. Teams will give 2-3 minute presentations about their approach and analysis to judges that come to their tables. There will be separate judges for the beginners track and the advanced track. They will judge on visualizations, analysis, and presentation. The judges can then decide on the top 3 teams in no specific order from each track.

From 8-8:45 there will be dinner. During this time, students and recruiters can network. We will announce the 3 top teams from each track. After dinner, each of the 6 teams (top three from each track) will give a 5 minute PowerPoint presentations for all of the judges (this will take about an hour). These will be back-to-back presentations. The judges can then break, choose the first, second, and third place team for each track and we can give awards.

This means that all teams must prepare both a trifold with analyses and visualizations and a five-minute PowerPoint presentation for Monday, November 6.
Prizes

We know that incentive drives turnout to competitions. We want to have monetary awards for each member of the first, second, and third place teams for both of the beginner and advanced tracks. We also want to make the awards slightly more for the advanced track for even more incentive. The following are the awards:

- Beginner Track:
  - $90 for each member of the 1st place team
  - $75 for each member of the 2nd place team
  - $50 for each member of the 3rd place team
- Advanced Track:
  - $120 for each member of the 1st place team
  - $90 for each member of the 2nd place team
  - $75 for each member of the 3rd place team

Calendar of Events

Friday, October 13
- Last Day to Sign Up - Team and individual sign ups due for fall data competition

Monday, October 30
- 6:30pm-9pm: Introduction meeting – teams will receive datasets, GDMS will give a presentation and network.
- Dinner will be provided (Location: NCB 260)

Monday, November 6
- 5:30-6:45: Setup and introduction – Teams will setup trifold stations, GDMS and CMDA club will give opening remarks (Location: NCB 160)
- 6:45-7:45: Preliminary poster board session – Teams will give 2-3 minute presentations with their trifolds to judges (Location: NCB 210, NCB 270)
- 7:45-8 Preliminary judging – Judges will decide top 3 teams (in no specific order) from each track
- 8-8:45: Dinner and networking (Location: NCB 160)
- 8:30: Announce Finalists - Judges will announce top 3 teams from each category
- 8:45-9:30: Final presentations – Top 3 teams from each track will give PowerPoint presentations to all judges
- 9:30-9:45: Final Judging -- Judges will determine first, second, and third place from each track
- 9:45-10: Closing - Awards and closing remarks