CS 6240: Project Presentation

**Goals**: (1) Report the results of your project. (2) Practice an important job skill: presenting your work to others in oral form.

Each team has to create all deliverables from scratch. In particular, it is not allowed to copy another team’s code or text and modify it. If you use publicly available code or text, you need to **cite the source** in your report!

All deliverables for this HW are due as stated on Blackboard. (To ensure a smooth transition between presentations, we will make a dedicated presentation laptop available and upload all presentations to that laptop before the presentation.) For late submissions you will lose one percentage point per hour after the deadline. This HW is worth 100 points and, together with the oral presentation, accounts for 10% of your overall homework score. Please submit your solution through the Blackboard assignment submission system and make sure your presentation slides are stored as a single **PDF** file. Do not submit anything else (source code, result files, log files).

**File naming convention**: use the last names of all team members, sorted alphabetically, as the file name. For instance, the two-person team Joe Smith and Mary Miller would use file name MillerSmith.pdf.

# Project Presentation

Each team will present their work for about 7 minutes, followed by 1-2 minutes of open discussion and questions. (Precise time limits will be announced once teams are finalized.) Consider the following recommendations when preparing your presentation:

* Think of this as a hiring talk. You are interviewing for a high-paying job with your favorite Big Data company and the class is a group of experts from this company who will decide if they should hire you or not. How do you show this technically versed audience in a few minutes that you are capable of solving a Big Data problem?
* Do not waste time by explaining small details about input data or all the different tasks you worked on. Just give a 1-minute overview of the project.
* Then focus on a few carefully selected results that best showcase your expertise as a MapReduce programmer. All other tasks can be summarized on a single slide.
  + E.g., if you wrote 20 different group-by/count queries and a decision tree algorithm, do not waste time on the simple jobs. Instead, focus on the decision tree challenges: discuss the main idea how you solved it and present some concrete results such as running time numbers for different problem sizes and/or numbers of machines used. For the 20 simple tasks, just enumerate them on a slide that says what else you computed, maybe including visualizations of selected interesting results.
* When selecting graphs or tables showing concrete measurements, think about which ones you need to make a case that you (1) indeed solved the problem and (2) your code is scalable or at least fast enough for the given problem.

In addition to presenting your own work, every student in class will also evaluate the other teams’ presentations. The main evaluation criteria are:

* Would you hire this team for your Big Data company?
* Did the presentation give you a good idea about what the team was working on and how they solved the problem?
* Did you see enough concrete results to be confident that the team solved the problem well and thought carefully about their solution and how to present it?

# Deliverables

1. Presentation slides. (1 PDF file)