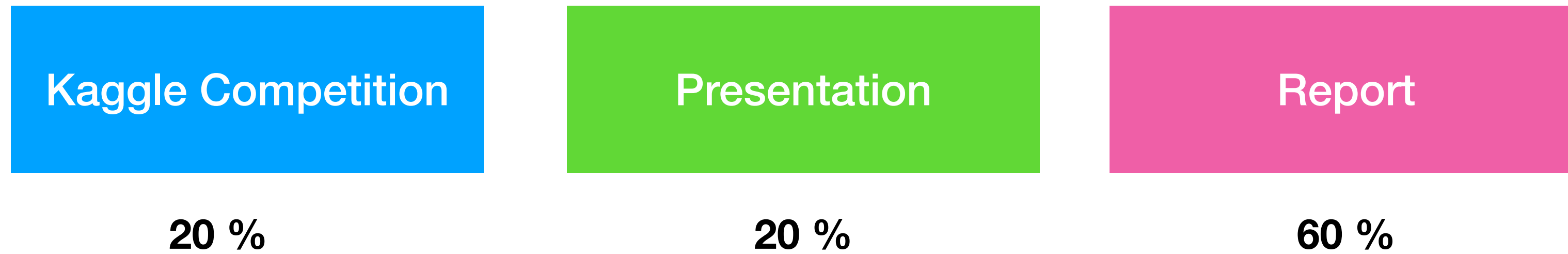


General Guidelines for Final Project

STATS101C - Introduction to Statistical Models and Data Mining

Scoring Guidelines



Teams

- At most **3 students** per team.
- Students **must not change** teams during the competition.
- Students without a team must first post an announcement on Campuswire. If this does not work, then please email me so I can post an announcement on CCLE.
- Please register your team in the corresponding google sheet:
 - Lecture 3
 - Lecture 4

Kaggle Competition

- Teams can use any method seen in class or available in the textbook.
- Teams can also use other method not seen in class. However, they **must first ask** for permission to your TA and me.
- In your email you must motivate the method:
 - Why do you think it will be useful?
 - What advantages does it offer compared to the method seen in class?
 - Brief description of the method.
 - R libraries including the implementation of this method.
- If you decide to use a method not seen in class, you must make room in your report to clearly explain what the method is doing.
- Teams do not need to do this for preprocessing methods. However, a good motivation in the report (with possibly references) is expected

Kaggle Score

- We will use five models to evaluate the performance of the teams.
 - You will get 5% if you beat Model 1
 - 10% if you beat Model 2
 - 15% if you beat Model 3
 - 20% if you beat Model 4

Presentation

- Duration: **4 min** and **2 min of questions**.
- The presentation must be **previously recorded** and all team members must participate.
- We will evaluate: Clarity, structure, motivation of methods and results, critical thinking, etc.
- We will ask quick questions to a team member at random. Examples of questions are:
 - What do you think the model works? Why does it fail?
 - What are the most important predictors? How is the most important predictor driving the predictions?
 - How would you explain your model to your boss or little brother? How confident are you that the model is working well?

Report

- Similar guidelines as midterm report. If the students decide to use an advanced analysis or preprocessing method, then references must be included.
- Statement of contribution.

Time Line

- Posting Kaggle Competition: **November 30.**
- End of Competition: **December 11.**
- Maximum number of submissions per day: 5.
- Please submit your report and presentation before **December 14** at 11.59 PM (PST). You must add a link to your presentation in the google sheet you used to register your teams.
- During the scheduled time for the final exam, we will watch the videos and ask questions.