Final Project  
Luke, Brett, Bill  
First Segment ToDo

* Decide on a topic for the project—think of a question that can be answered using data.  
  Notes:

What scores would separate the average player compared to a top scoring player?  
Would a machine learning module be able to predict who is a top player based on how many flawless matches they achieve?   
Compare players by region and consider tournament style?   
Use machine learning to guess live tournament victor based on current online tournament data.

* ~~Create a repository for the project and invite the other team members to join.  
  Notes: Luke created GitHub Repo and shared during lecture 8/16  
  COMPLETE~~
* Source a dataset that will suit your needs (you can even use multiple datasets if applicable).  
  Notes:  
  List of at least 4 data sources. Contemplating video game tournament data.   
  <https://developer.start.gg/docs/intro>  
  <https://api.challonge.com/v1>  
  <https://developer.start.gg/reference/query.doc.html>
* Begin to clean, organize, and perform exploratory data analysis on your datasets so that they're ready for analysis.  
  Notes:

Panda to clean once sourced.   
Need to filter by player and a win/loss ratio  
Perform statistical analysis to observe a bell curve and consider the outliers as low performers vrs average vrs high performers

* Include mockups of a machine learning model and a database.  
  Notes:

Possibly feed the machine learning one tournament data and also on randomly generated tournament data set to add to model.

## Random forest as a possible algorithm Preprocess the Data Split into the training and testing sets