**Prompt 1**

Discuss on 7/10 at 6PM EST in more detail

**Prompt 2**

Des - Getting power ranking

Wesley - Consolidate data, run [XGBoost](https://xgboost.readthedocs.io/en/latest/model.html),lightGBM

(with all features, with only test set features)

1.Combining Data:

Time: weekday, month

Location:country

Top 5 country and others

Away\_team:

Home\_team:

Num\_away\_team\_all\_star\_player(not yet)

Num\_home\_team\_all\_star\_player(not yet)

2.Feature engineering

2.1 one\_hot\_encoding(done)

2.2 missing value imputation

2.3 outlier detection

3.Grid search and other hyper-parameter tuning (k-fold cross validation)(done)

*Will add in power-ranking, count of all-star*

*Square of team-power ranking?*

Other Data Sources

<https://www.basketball-reference.com/>

<http://www.espn.com/>