**Anomaly Detection Methods:**

* Statistical - can do experiments to test for multivariate normality
  + Andrew Ng’s videos(he left baidu a few weeks ago)
  + Gaussian mixture models (EM algorithm)
* Density-based (DBSCAN)
* Clustering-based

**Tasks:**

1. Write scraper and scrape data
2. Figure out how to split players into forwards, mids, defenders, goalkeepers (using positions list) and save data as separate CSV files
3. Figure out how to convert CSV file into Numpy format
4. Preliminary data analysis (visualizations of clusters based on certain features for different positions)

Lynn (statistical):

Sid (clustering-based):

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0152173>

http://trevorwhitney.com/data\_mining/anomaly\_detection

Chenhao (density-based DBSCAN):

**http://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf**