Documentation

Tuesday, June 28, 2022

7:42 PM

Segment 1 - Class 1

* Data and Resources:
  + Main Data source: <https://www.seanlahman.com/baseball-archive/statistics/>
  + Data source: <https://www.kaggle.com/datasets/seanlahman/the-history-of-baseball?select=salary.csv>
  + Data source: <https://www.kaggle.com/datasets/open-source-sports/baseball-databank?select=Salaries.csv>
  + ML Predicting Salaries Project: <https://www.kaggle.com/code/samfenske/ml-predicting-mlb-salaries>
  + Moneyball Project: <https://www.kaggle.com/code/juder0y/moneyball-prediction>
  + Other Baseball Data Science Projects: <https://www.baseballdatascience.com/>
* Pitch
  + Our team is analyzing MLB player statistics including but not limited to: RBIs, Hits, On Base Percentage, Batting Average, Wins, Doubles, Triples, Home Runs, and position, to predict average salaries and determine if and how players are being over or under compensated within 10% of the predicted salary.