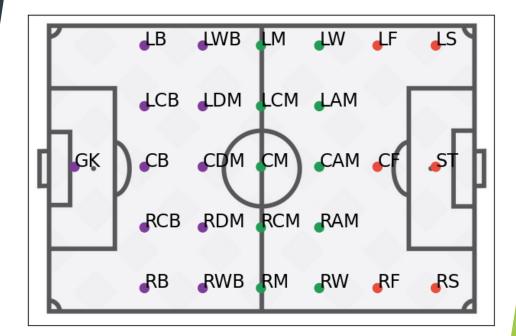
Put on your General Manager Cap: Analyzing FIFA Player Data

ECE 143 WI 2021 Group 26 Anirudh Swaminathan Hejin Liu David Liau Zhifeng Li



Background

- Goals of Soccer
- Clubs
- Players
- Positions



Background of Data

- ▶ 25 Producers & 400 outside data contributors
- ▶ 6,000 Data Reviewers worldwide
- Thorough data verification process
- ▶ Players: >30 specific attributes
 - Overall rating: Pace, Shooting, Passing, Dribbling, Defending, and Physical

Source: https://www.goal.com/en-ae/news/fifa-player-ratings-explained-how-are-the-card-number-stats/1hszd2fgr7wgf1n2b2yjdpgynu



Dataset Display

Shape: (18207, 89)

ID	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Special	Preferred Foot	International Reputation	Weak Foot	Skill Moves	Work Rate	Body Type	Real Face	
158023	L. Messi	31	Argentina	94	94	FC Barcelona	€110.5M	€565K	2202	Left	5.0	4.0	4.0	Medium/ Medium	Messi	Yes	
20801	Cristiano Ronaldo	33	Portugal	94	94	Juventus	€77M	€405K	2228	Right	5.0	4.0	5.0	High/ Low	C. Ronaldo	Yes	
190871	Neymar Jr	26	Brazil	92	93	Paris Saint- Germain	€118.5M	€290K	2143	Right	5.0	5.0	5.0	High/ Medium	Neymar	Yes	
193080	De Gea	27	Spain	91	93	Manchester United	€72M	€260K	1471	Right	4.0	3.0	1.0	Medium/ Medium	Lean	Yes	
192985	K. De Bruyne	27	Belgium	91	92	Manchester City	€102M	€355K	2281	Right	4.0	5.0	4.0	High/ High	Normal	Yes	
4																	

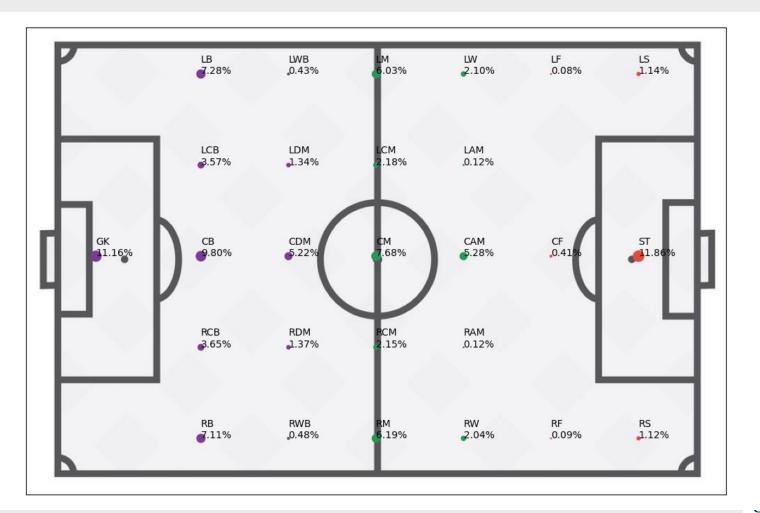
UC San Diego

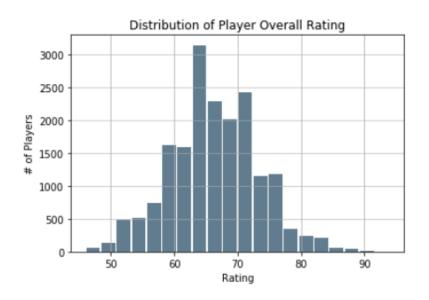
Data Preparation

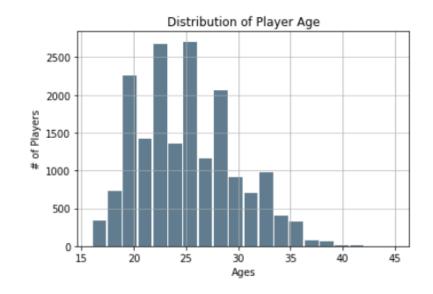
- Removing URLs
 - Dead links
- Formatting
 - Wage/Value/Release Clause
- Skill Attributes
 - Crossing, Dribbling, Ball Control, Agility, Balance

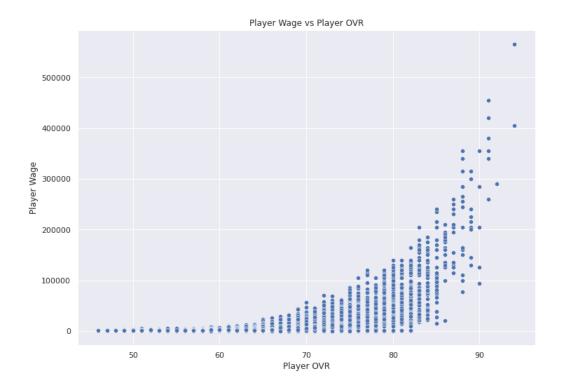
Exploratory Data Analysis











Scenario-Based Data Analysis



What playstyles do each team reflect strongly?



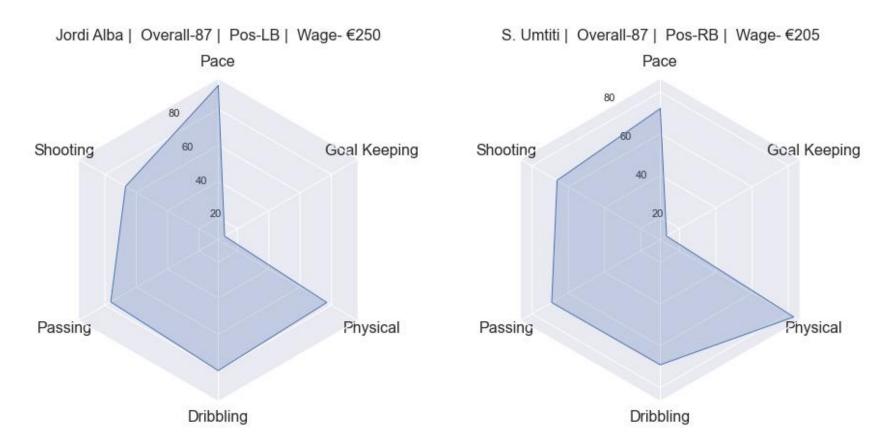
Which players are best fit for a given team?



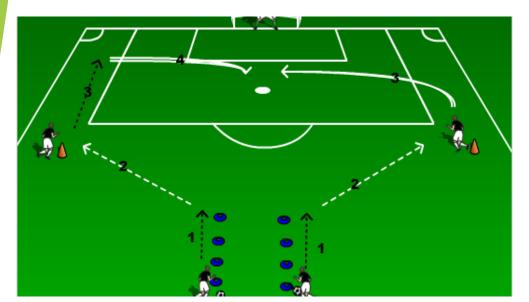
Scenario 1

- Long Passing, FC Barcelona
 - ►[ST, LB, RB, CB]





Scenario 2

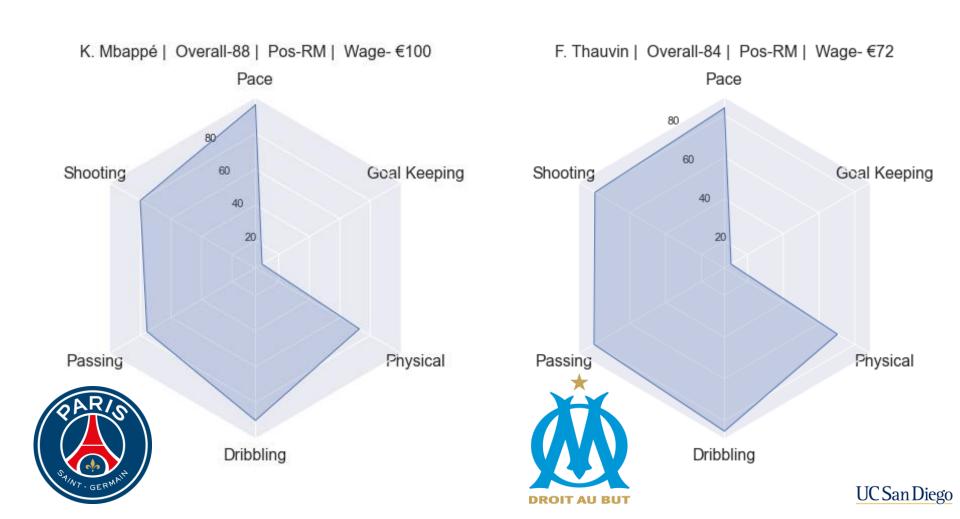


Crossing, Juventus[ST, LM, RM, LW, RW]

JUVENTUS

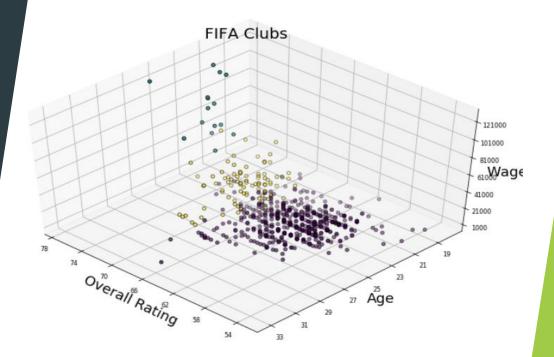
ST, LW, RW are good enough, need LM and RM





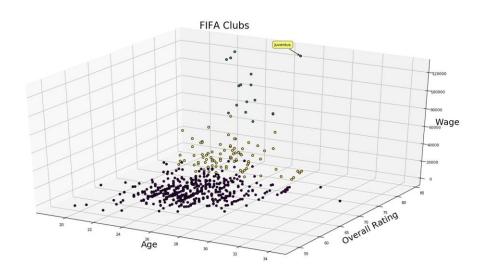
ML Model - Team Clustering

- What kind of team to play?
- 3 Clusters
 - Championship Contenders
 - Middle of the Pack
 - Developmental Rosters



* + > + Q = 2 8

3



- FIFA dataset is comprehensive
- More data comes out every year (2k)
 - Player rating comparison to real life
 - Contract movement over years

Takeaways

