

A dynamic photograph of a soccer player's lower body in mid-kick. The player is wearing blue and white striped socks and black and yellow cleats. A black and white soccer ball is positioned on the grass in front of the player's foot. The background is blurred, showing other players and the green of the soccer field.

Soccer Analysis -FIFA

Claudia Vargas,
Tianying Gong,
Qingyuan Wang,
Kexin Mao



Background Information

- Who we are?



- What are the clients need?



- Database Selection



- Authorized data from FIFA
- 18207 Rows, 89 columns
- Quantify the ability of each player
- Strength compared with other data

ID	Name	Age	Photo	Nationality	Flag	Overall	Potential	Club	Club Logo	Value	Wage	Specia!	Preferred Fo!	Internationa!	Weak Foot	Skill Move	Work Rate	Body Type	Racial Face	Position	Jersey Num	Joined
158023	L. Messi	31	https://cdn.i	Argentina		94	94	FC Barcelona		€110.5M	€65K	2202	Left	5	4	4	Medium/M. Messi	Yes	RF	10	1-Jul-04	
20801	Cristiano Ron	33	https://cdn.i	Portugal		94	94	Juventus		€77M	€405K	2228	Right	5	4	5	High/Low C. Ronaldo	Yes	ST	7	10-Jul-18	
190871	Neymar Jr	26	https://cdn.i	Brazil		92	93	Paris Saint-G		€118.5M	€290K	2143	Right	5	5	5	High/M. Neymar	Yes	LW	10	3-Aug-17	
193080	De Gea	27	https://cdn.i	Spain		91	93	Manchester		€72M	€260K	1471	Right	4	3	1	Medium/M. De Gea	Yes	GK	1	3-Jul-11	
192985	K. De Bruyne	27	https://cdn.i	Belgium		91	92	Manchester		€102M	€355K	2281	Right	4	5	4	High/High Normal	Yes	RCM	7	30-Aug-15	
183277	E. Hazard	27	https://cdn.i	Belgium		91	91	Chelsea		€93M	€340K	2142	Right	4	4	4	High/Medium Normal	Yes	LF	10	1-Aug-12	
177003	L. Modrić	32	https://cdn.i	Croatia		91	91	Real Madrid		€67M	€420K	2280	Right	4	4	4	High/High Lean	Yes	RCM	10	1-Aug-12	
176580	L. Suárez	31	https://cdn.i	Uruguay		91	91	FC Barcelona		€80M	€455K	2346	Right	5	4	3	High/Medium Normal	Yes	RC	9	11-Jul-14	
155862	Sergio Ramo	32	https://cdn.i	Spain		91	91	Real Madrid		€51M	€380K	2201	Right	4	3	3	High/Medium Normal	Yes	RCB	15	1-Aug-05	
200389	J. Oblak	25	https://cdn.i	Slovenia		90	93	Atletico Mad		€68M	€94K	1331	Right	3	3	1	Medium/M. Normal	Yes	GK	1	16-Jul-14	
188545	R. Lewandov	29	https://cdn.i	Poland		90	90	FC Bayern M		€77M	€205K	2152	Right	4	4	4	High/Medium Normal	Yes	ST	9	1-Jul-14	
182521	T. Kroos	28	https://cdn.i	Germany		90	90	Real Madrid		€76.5M	€355K	2190	Right	4	5	3	Medium/M. Normal	Yes	LCM	8	17-Jul-14	
182493	D. Godín	32	https://cdn.i	Uruguay		90	90	Atletico Mad		€44M	€125K	1946	Right	3	3	2	Medium/H. Lean	Yes	CB	10	4-Aug-10	
168542	David Silva	33	https://cdn.i	Spain		90	90	Manchester		€60M	€285K	2115	Left	4	2	4	High/Medium Normal	Yes	LCM	21	14-Jul-10	
215914	N. Kanté	27	https://cdn.i	France		89	90	Chelsea		€25K	€218K	2189	Right	3	3	2	Medium/H. Lean	Yes	LDM	13	16-Jul-10	
211110	P. Dybala	24	https://cdn.i	Argentina		89	94	Juventus		€89M	€205K	2092	Left	3	3	4	High/Medium Normal	Yes	LF	21	1-Jul-15	
202126	H. Kane	24	https://cdn.i	England		89	91	Tottenham H		€83.5M	€205K	2165	Right	3	4	3	High/High Normal	Yes	ST	9	1-Jul-10	
194765	A. Griezman	27	https://cdn.i	France		89	90	Atletico Mad		€78M	€145K	2246	Left	4	3	4	High/High Lean	Yes	CAM	7	28-Jul-14	
192448	M. ter Stege	26	https://cdn.i	Germany		89	92	FC Barcelona		€58M	€240K	1328	Right	3	4	1	Medium/M. Normal	Yes	GK	22	1-Jul-14	

Normalization

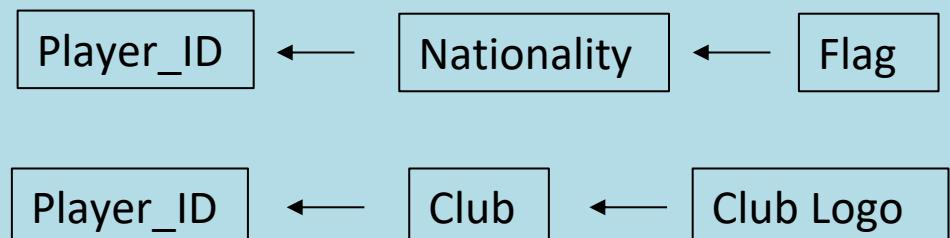
- 1 NF: 1) split work rate
2) merge adjacent position attributes

Work Rate	LAM	CAM	RAM
Medium/ Medium	93+2	93+2	93+2
High/ Low	88+3	88+3	88+3
High/ Medium	89+3	89+3	89+3
Medium/ Medium	88+3	88+3	88+3
High/ High	89+3	89+3	89+3

1 NF: *Player, player_style, value, rating, defensive, forwards, midfielders, player_position_rating, dribbling, defending, passing, shooting, pace, physical, goal_keeping, player_ability_rating*

2 NF: Add the primary key, which contains only one attribute, to all the tables.

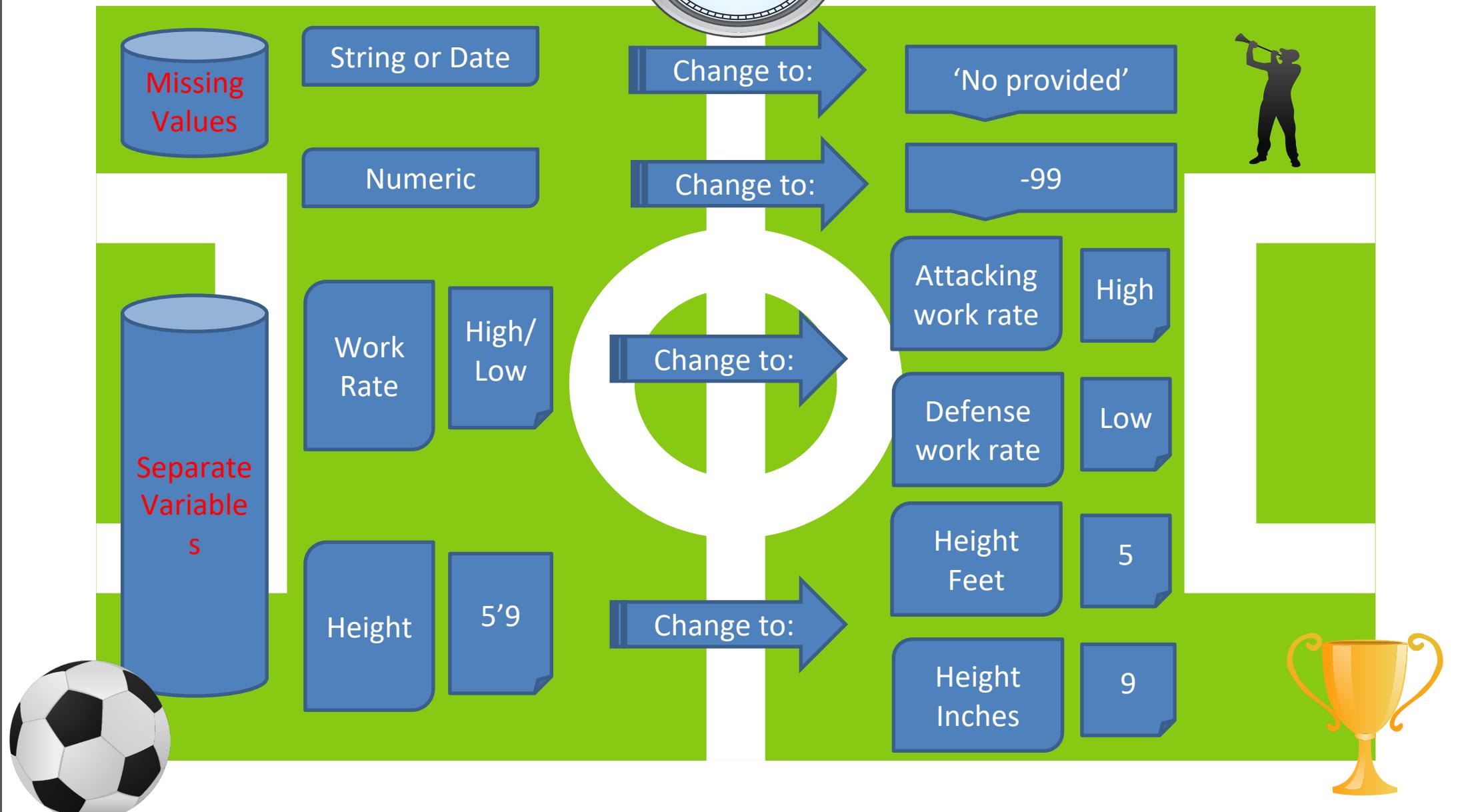
- 3 NF: Reduce transitive dependence



3 NF (22 tables in total):
Player, foot_preference, player_habits, player_style, value, player_significance, rating, country, contract, club, defensive, forwards, midfielders, player_position_rating, dribbling, defending, passing, shooting, pace, physical, goal_keeping, player_ability_rating

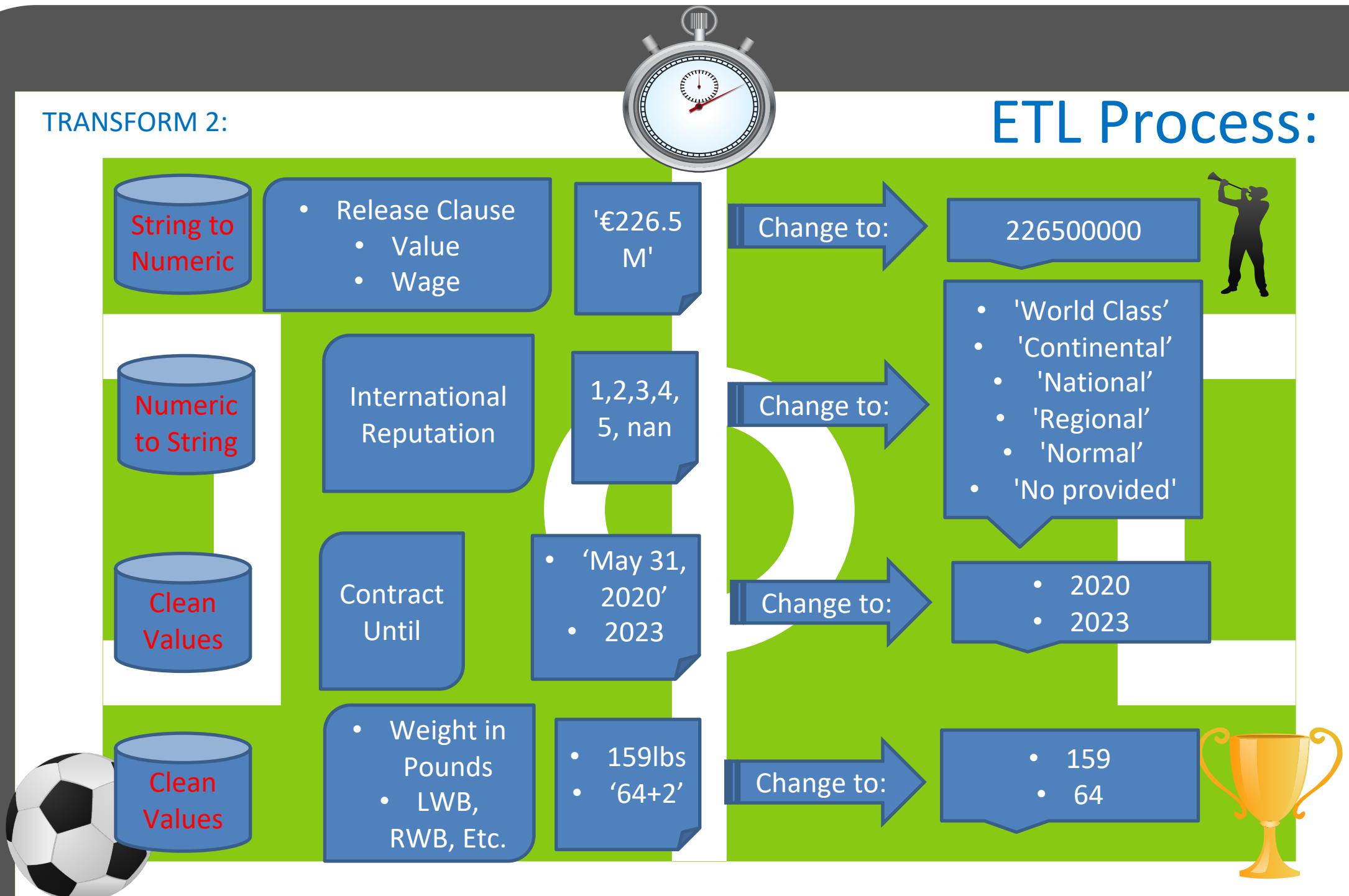
TRANSFORM 1:

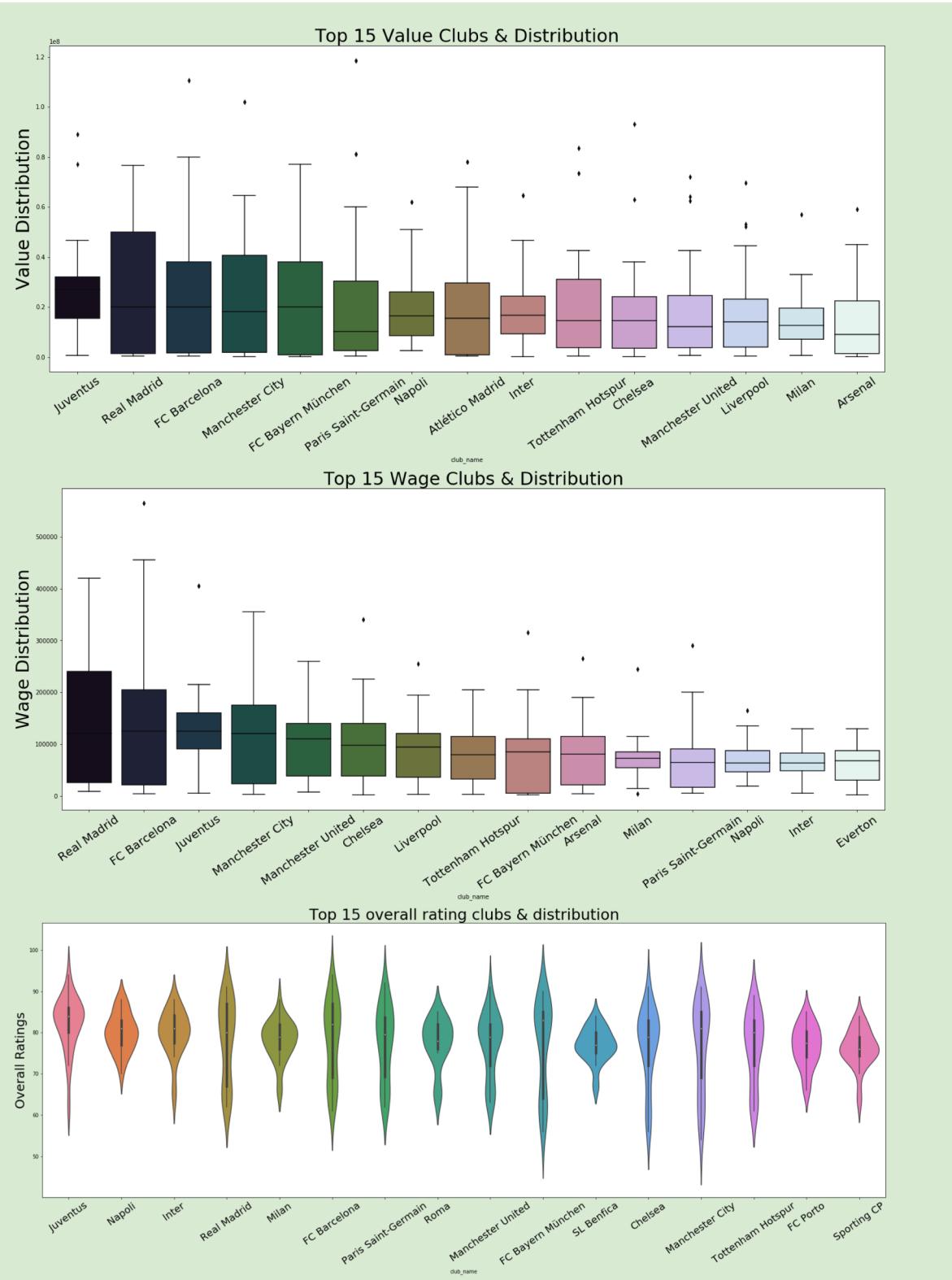
ETL Process:



ETL Process:

TRANSFORM 2:

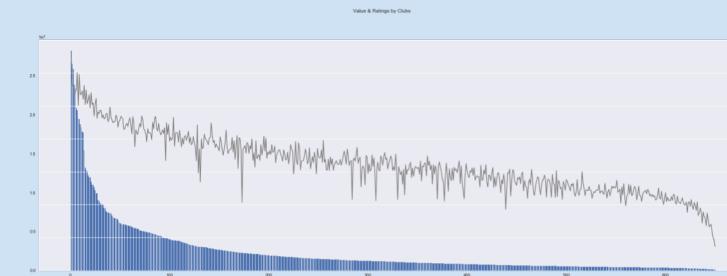




Descriptive Analysis:

1. Distribution of Value, wages, & Ratings in Top Ranking Clubs:
Italian clubs have high rankings and more centralized distribution

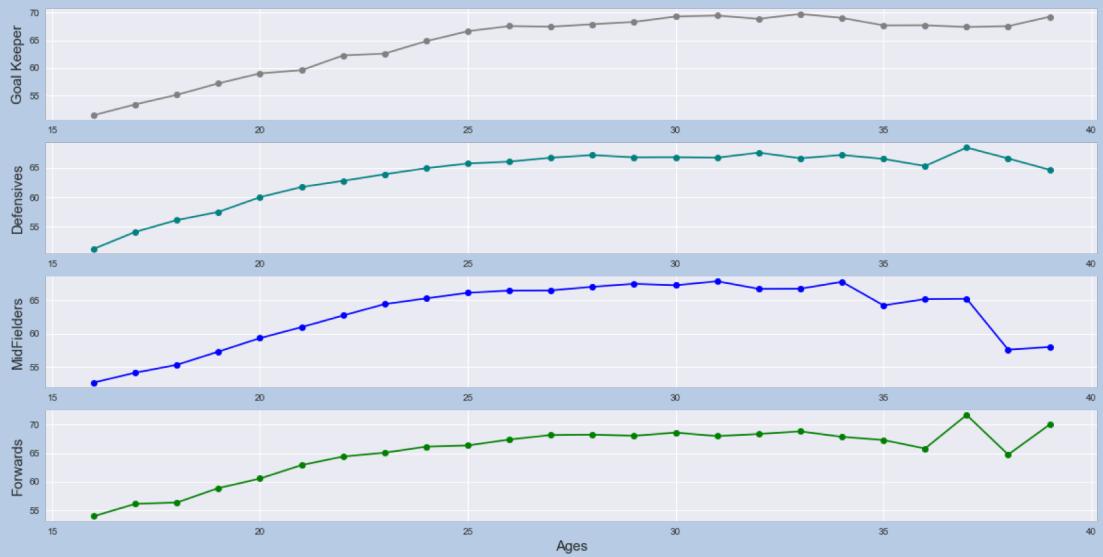
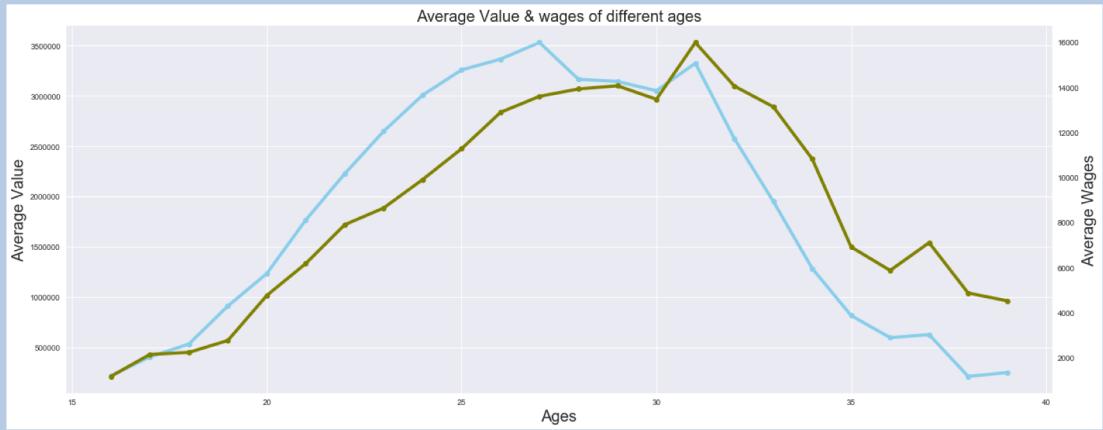
2. Correlations of Value, wages with overall ratings of clubs



Value - Overall Ratings
0.78



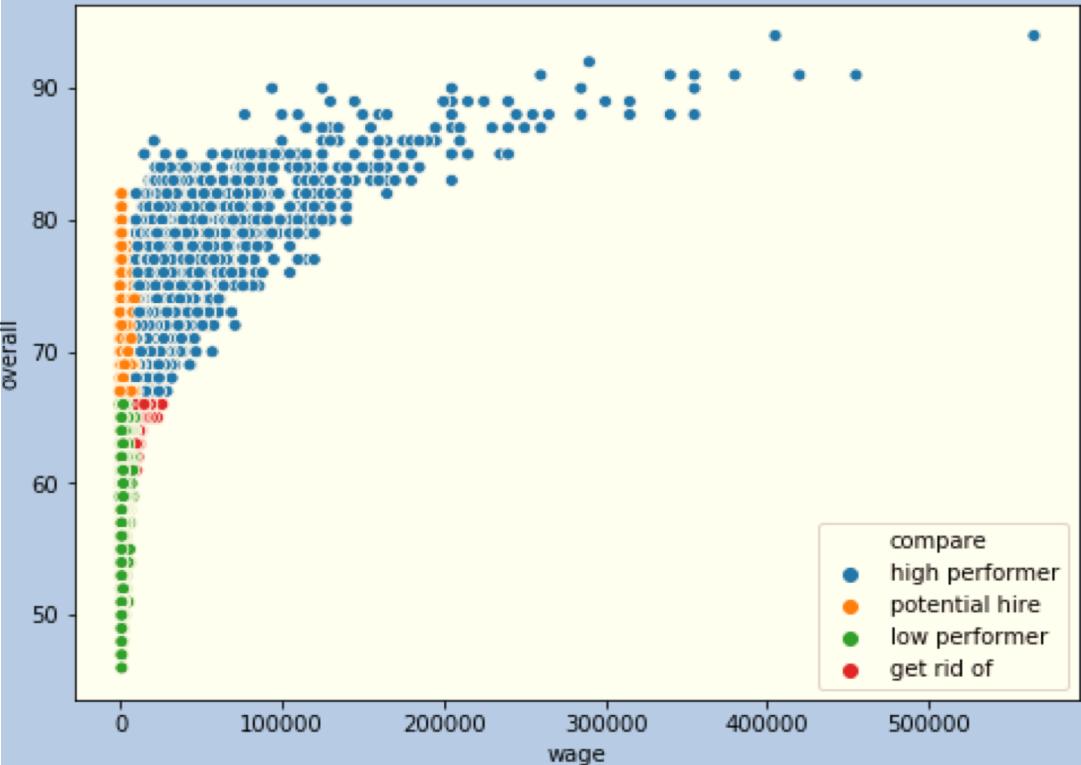
Wage - Overall Ratings
0.67



3. Age's Influences:

- **30s is the 'changing point'**
- **The changing point of player's wage comes few years later than player's value**
- **Trends in different positions varies.**
- Midfielders' performances decreases rapidly.**
- **More details in Metadata**

Who Should I Hire?



Average wage per week(Euro):
9731

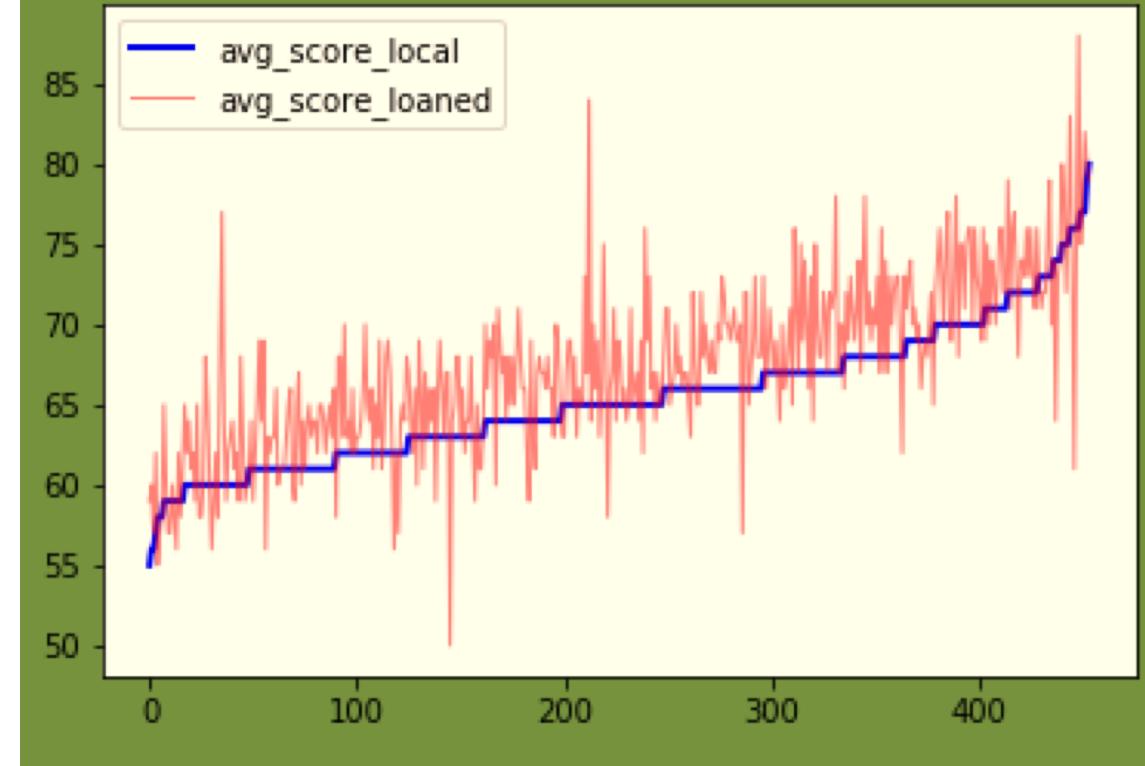
Average overall score: 66

Most needed: Attacking, shooting

454 clubs have players on loan.

Average score of players on loan is 2.06
higher than the score of regular players

Player on Loan





	club_name character varying (100)	club_overall numeric (4,2)	club_dribbling numeric (4,2)	club_defending numeric (4,2)	club_goalkeeping numeric (4,2)	club_passing numeric (4,2)	club_shooting numeric (4,2)	club_pace numeric (4,2)	club_physical numeric (4,2)
1	Juventus	82.28	0.20	0.31	0.02	0.23	0.20	0.12	0.16
2	Napoli	80.00	0.19	0.23	0.10	0.22	0.19	0.11	0.06
3	Inter	79.75	0.16	0.22	0.20	0.18	0.18	0.06	0.10
4	Real Madrid	78.24	0.16	0.17	0.12	0.20	0.15	0.09	0.07
5	Milan	78.07	0.17	0.19	0.17	0.19	0.15	0.06	0.07
6	FC Barcelona	78.03	0.17	0.18	0.13	0.24	0.20	0.05	0.05
7	Paris Saint-Germain	77.43	0.20	0.18	0.02	0.24	0.19	0.13	0.07
8	Roma	77.42	0.12	0.20	0.06	0.18	0.16	0.05	0.06
9	Manchester United	77.24	0.17	0.22	0.02	0.26	0.24	0.09	0.13
10	FC Bayern München	77.00	0.12	0.12	0.13	0.21	0.18	0.04	0.04

Top 10 Performing Club in the World

Below average: 0

Most good at: Attacking, shooting, defending

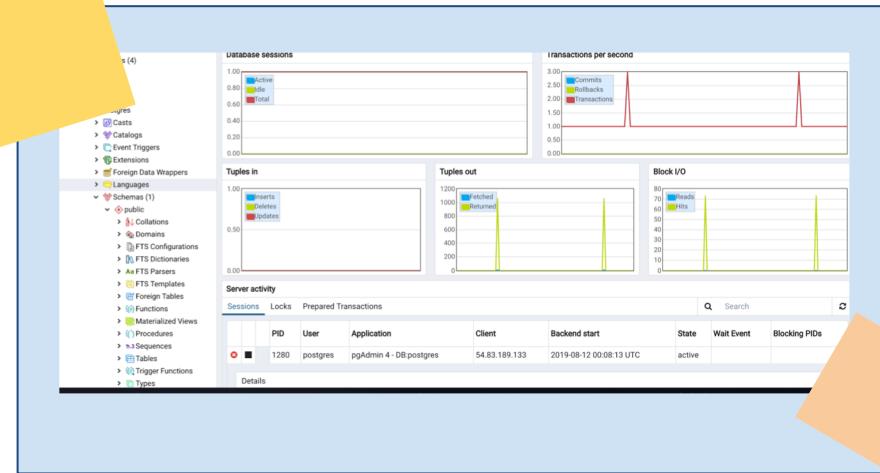
Physical score: slightly above average

Database Interaction SQL/Metadata

```
STMT_CLUBRANKS = """
DROP VIEW average_clube;
CREATE VIEW average_clube AS
SELECT cb.club_id,
       ROUND(AVG(cb.overall),2) AS avg_overall_rating,
       ROUND(AVG(v.value),2) AS avg_value,
       ROUND(AVG(v.wage),2) AS avg_wage_rating
  FROM player p
  JOIN club cb ON cb.club_id = p.club_id
  JOIN rating r ON r.player_id = p.player_id
  JOIN value v ON v.player_id = p.player_id
 GROUP BY cb.club_id
"""

CREATE VIEW club_ranks AS
SELECT
    cb.club_name,
    p.player_id,
    r.overall,
    a.avg_overall_rating,
    (DENSE_RANK() OVER(ORDER BY a.avg_overall_rating DESC)) AS overall_rating_rank,
    v.value,
    a.avg_value,
    (DENSE_RANK() OVER(ORDER BY a.avg_value DESC)) AS value_rank,
    v.wage,
    a.avg_wage_rating,
    (DENSE_RANK() OVER(ORDER BY a.avg_wage_rating DESC)) AS wage_rank
  FROM player p
  JOIN club cb ON cb.club_id = p.club_id
  Join value v ON v.player_id = p.player_id
  JOIN average a ON a.club_id = cb.club_id
  JOIN rating r ON r.player_id = p.player_id
 ORDER BY overall_rating_rank;
```

Python
Extract & Analysis
Build up Views for conveniences



Codio
Server for the database

Metadata
Visualize with table/views
Insert SQL Queries
Interact with charts



Q&A

Thanks for listening !