

LET'S GET MESSI

Predicting a Player's Salary



Jules Morris
Kayla Brock
Eriberto Contreras
Ray Cerna

AGENDA

- ⚽ Data Acquisition
- ⚽ Data Preparation
- ⚽ Explanatory Data Analysis
- ⚽ Modeling
- ⚽ Conclusion



DATA ACQUISITION AND PREPARATION

>>>



- Source - Kaggle.com
- Years 2015 - 2022 data
- 93% data retained for exploration
- 70% data retained for modeling



EXECUTIVE SUMMARY



BEST MODEL

GLM with
Tweedie Regressor



BEST FEATURES

Int'l Reputation
Reactions
Overall
Potential



RMSE BASELINE

€20154.67



TEST RESULTS

€11095.90
45% improvement



Main Model RMSE

Baseline: €20155

Model: €11096

Forwards
Model



Midfielders
Model



Defenders
Model



Goalkeepers
Model



RMSE: ???

*All models performed best using the Tweedie Regressor

OVERALL DATASET

01



Forwards

02



Midfielders

03



Defenders

04



Goalkeepers



FORWARD SKILLS



OVERALL

93%

SHOOTING

91%

BALL CONTROL

94%

#STRIKERS
#WINGERS
#FORWARDS



100

75

50

25

CROSSING
Min: 18
Avg: 51
Max: 90

100

75

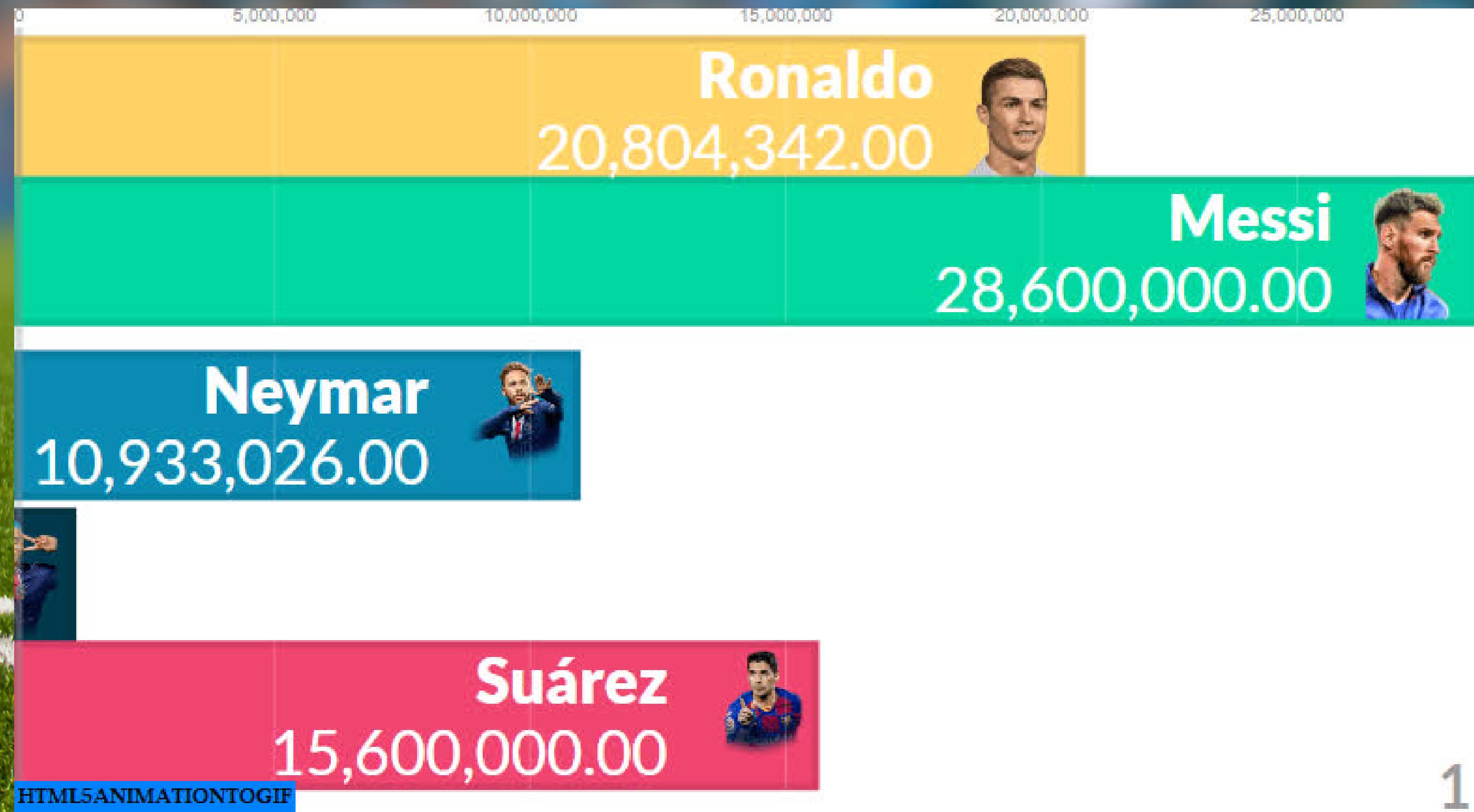
50

25

FINISHING
Min: 30
Avg: 66
Max: 95

SHORT PASS
Min: 22
Avg: 60
Max: 90

FORWARD SALARIES

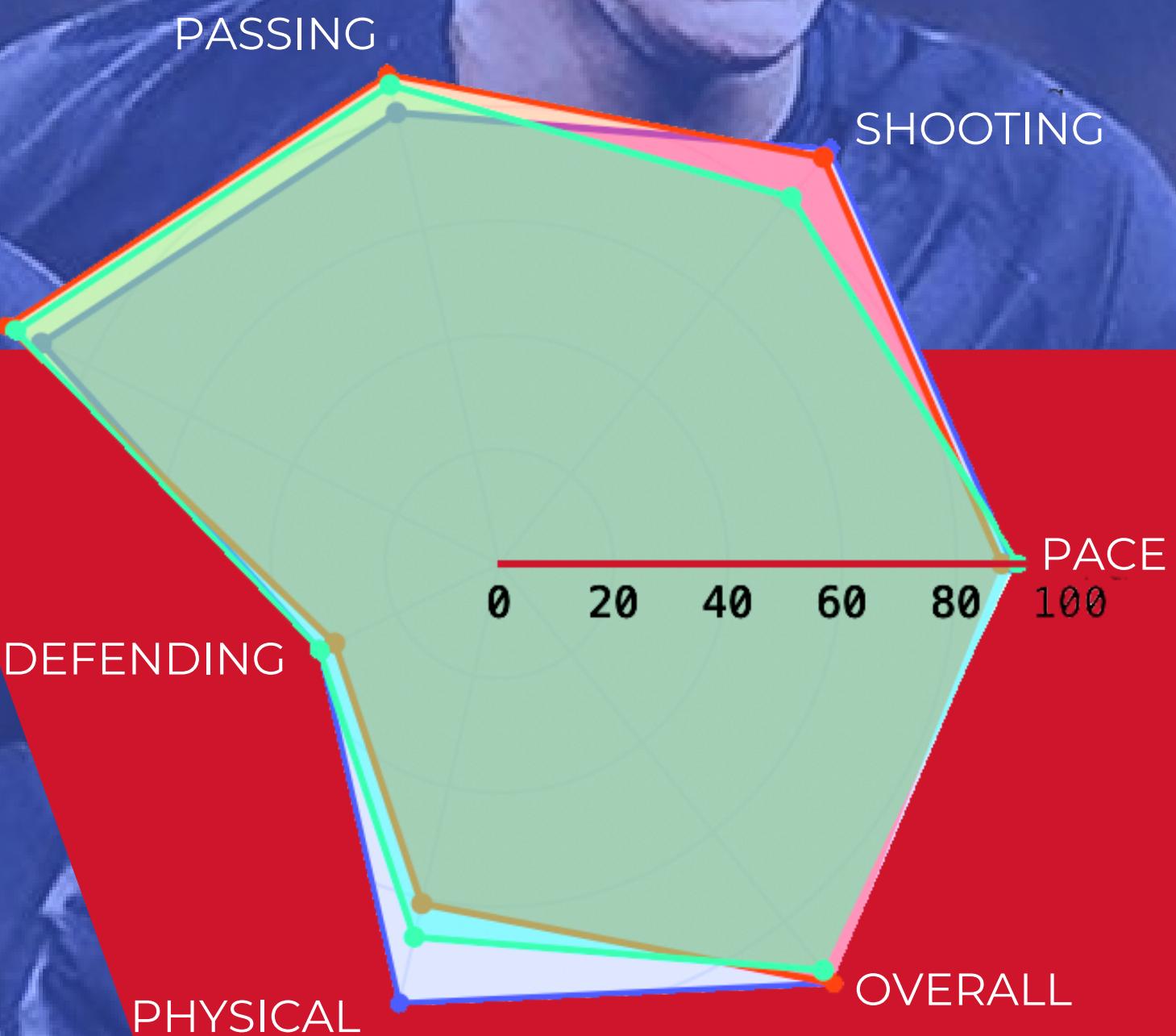


BEST BUDGET FORWARD

>>>>

WAGES

- NAME: EDEN HAZARD
- COST: €375,000/WEEKLY
- INTERNATIONAL REPUTATION: 4



- RONALDO
- MESSI
- HAZARD

OVERALL DATASET

01



Forwards

02



Midfielders

03



Defenders

04



Goalkeepers





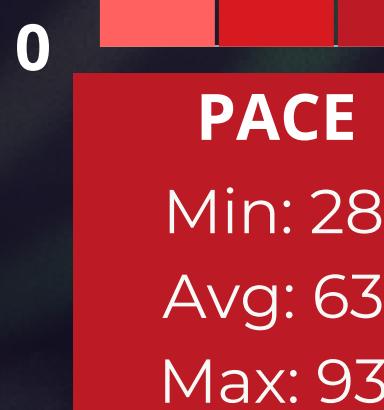
MIDFIELDER SKILLS

100

75

50

25

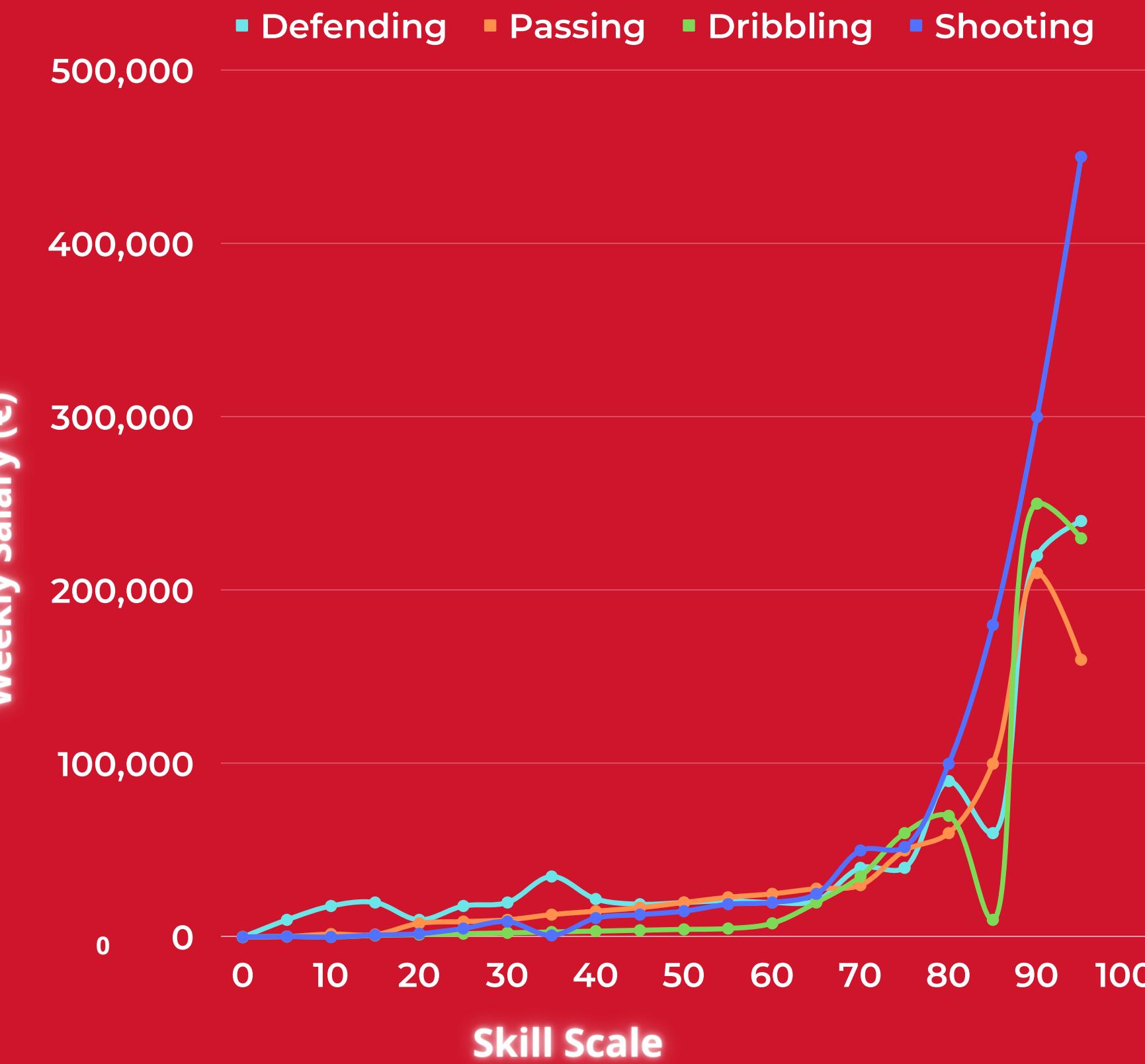




WEEKLY WAGE STATS IN EUROS

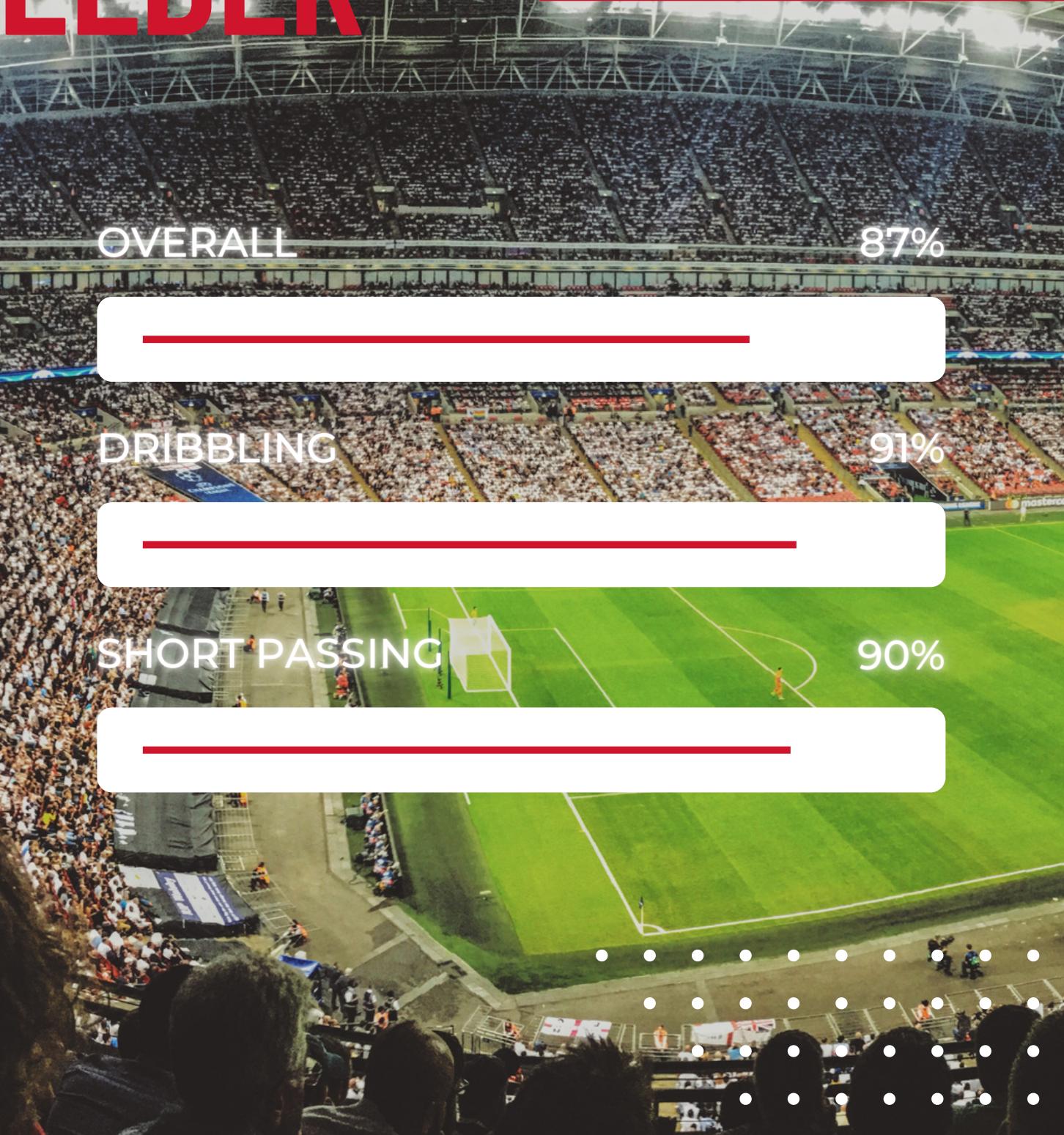
- MIN: €500
- AVG: €5000
- MAX: €475,000

MIDFIELDER



BEST BUDGET MIDFIELDER

- NAME: MARCO VERRATTI
- COST: €155,000/WEEKLY
- INTERNATIONAL REPUTATION: 3



OVERALL DATASET

01



Forward

02



Midfielder

03



Defender

04



Goalkeeper



DEFENDER SKILLS



SLIDING TACKLE

STANDING TACKLE

DEFENDING

INTERCEPTIONS



Min: 13
Avg: 65
Max: 95

Min: 14
Avg: 67
Max: 94

Min: 18
Avg: 64
Max: 91

Min: 13
Avg: 64
Max: 93



GOALKEEPER WAGE DISTRIBUTION



Max: € 375k
Min: €500
Average:
€12.4k

BEST BUDGET DEFENDER



- NAME: DIEGO GODÍN
- COST: €135,000/WEEKLY
- INTERNATIONAL REPUTATION: 3



Overall

85%

Defender

87%

OVERALL DATASET

01



02



03



04



Forwards

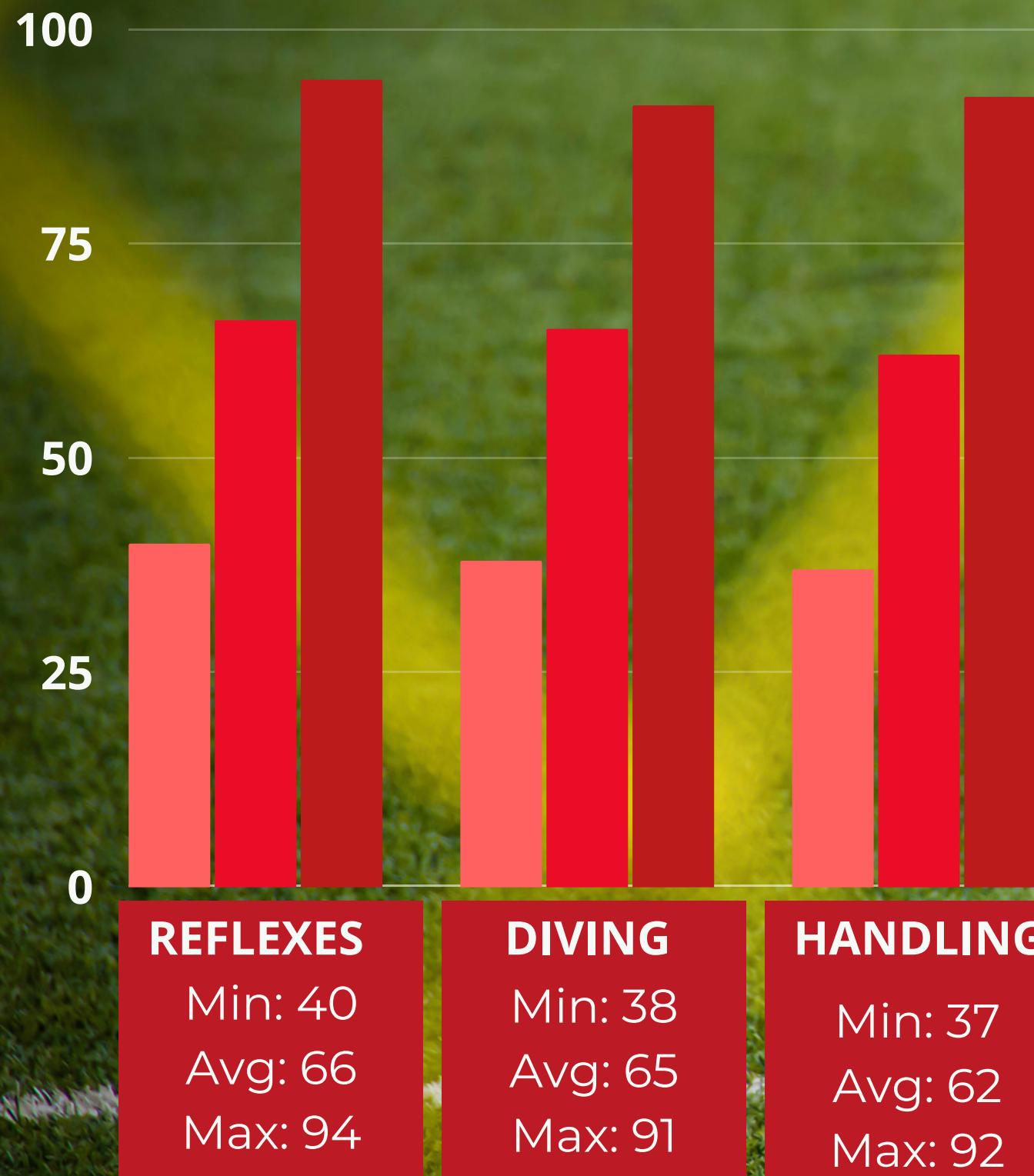
Midfielders

Defenders

Goalkeepers



GOALKEEPER SKILLS



>>>

GOALKEEPER WAGE DISTRIBUTION



Max: € 300k
Min: €500
Average:
€9.7k

BEST BUDGET GOALKEEPER

- NAME: JAN OBLAK
- COST: €130,000/WEEKLY
- INTERNATIONAL REPUTATION: 5



Overall



Reflexes



Main Model RMSE

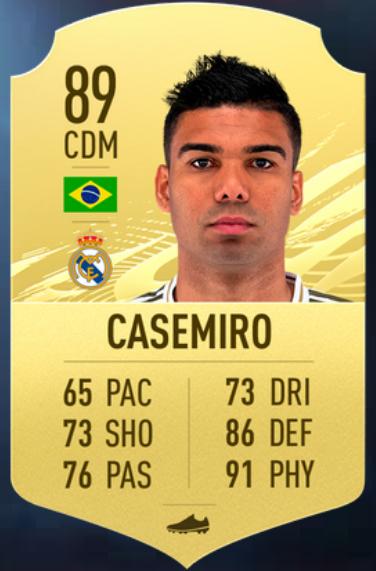
Baseline: €20155
Model: €11096

Forwards
Model



RMSE: €12580

Midfielders
Model



RMSE: €11015

Defenders
Model



RMSE: €10524

Goalkeepers
Model



RMSE: €9127

*All models performed best using the Tweedie Regressor

CONCLUSION

-  Highest rated attributes by field position
-  Improvement was best for midfielders, defenders, and goalkeepers
-  Best performing model - Tweedie Regressor

RECOMMENDATIONS

-  Separate data by player position
-  Separate by league and club due to budget variation
-  Utilize dimensionality reduction techniques to improve R2 score



DATA SCIENCE TEAM



**Eriberto
Contreras**
[Github/
EribertoContreras](#)



**Jules
Morris**
[Github/
JulesMorris](#)



**Ray
Cerna**
[Github/
RayCerna](#)



**Kayla
Brock**
[Github/
Kayfj41901](#)