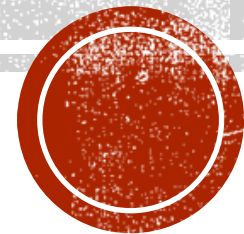


# FOOTBALL TRANSFERS

Presented by:

Kasra Ghorbani

Mohamad Mahdi Samadi



# AN INTRO TO OUR DATASET

- We found a dataset through Kaggle containing records on football players and some of their respective information such as salary, latest transfer fee, age, current league, total apps and total international apps, etc.
- Records started at 2018 and ended at 2023.
- Players' age in range of 15 to 41.
- 40,768 contracts were signed during these years.
- About 5,000 of them are signed in clubs from top five leagues.



# FEATURE ENGINEERING AND PREPROCESSING STEPS

- We added some features ourselves including:
  - duration of the contract.
  - age when signing the contract.
  - total number of matches played by the player.
  - Whether the player was under 30 at the time he signed his current contract or not.
- We refined and preprocessed the data by following these steps:
  - removing irrelevant columns and columns with too much missing values.
  - removing outlier players. Mostly players of Saudi Pro League.
  - fixing representation of salaries and transfer fees.
  - splitting date into year, month, day.



# RAW DATASET

	Name	Club	Division	Based	Nat	EU National	Caps	AT Apps	Position	Age	CR	Begins	Expires	Last Club	Last Trans. Fee	Salary
0	Cristiano Ronaldo - Portuguese	Al-Nassr (KSA) - Roshn Saudi League	Roshn Saudi League ...	Saudi Arabia (Saudi Pro League) ...	POR	Yes	200.0	673	ST (C)	38.0	7,629	31/12/2022	30/6/2025		-	€203,478,000 p/a
1	Karim Benzema - French	Al-Ittihad - Roshn Saudi League	Roshn Saudi League ...	Saudi Arabia (Saudi Pro League) ...	FRA	Yes	97.0	555	ST (C)	35.0	8,250	1/7/2023	30/6/2025		-	€199,452,000 p/a
2	Neymar - Brazilian	Al-Hilal (KSA) - Roshn Saudi League	Roshn Saudi League ...	Saudi Arabia (Saudi Pro League) ...	BRA	Yes	124.0	340	AM (LC), ST (C)	31.0	9,000	15/8/2023	30/6/2025	Paris Saint-Germain	€90M	€149,589,000 p/a
4	Kylian Mbappé - French	Paris Saint-Germain - Ligue 1 Uber Eats	Ligue 1 Uber Eats ...	France (Ligue 1) ...	FRA	Yes	70.0	221	AM (RL), ST (C)	24.0	9,415	22/7/2023	30/6/2028	Monaco	€186M	€48,467,000 p/a
5	Sadio Mané - Senegalese	Al-Nassr (KSA) - Roshn Saudi League	Roshn Saudi League ...	Saudi Arabia (Saudi Pro League) ...	SEN	No	93.0	377	AM (RLC), ST (C)	31.0	7,488	1/8/2023	30/6/2027	Bayern München	€28M	€40,977,000 p/a
7	Kalidou Koulibaly - Senegalese	Al-Hilal (KSA) - Roshn Saudi League	Roshn Saudi League ...	Saudi Arabia (Saudi Pro League) ...	SEN	Yes	70.0	367	D (C)	32.0	7,500	1/7/2023	30/6/2026	Chelsea	€25M	€29,918,000 p/a

Index: 40768 entries, 0 to 40790

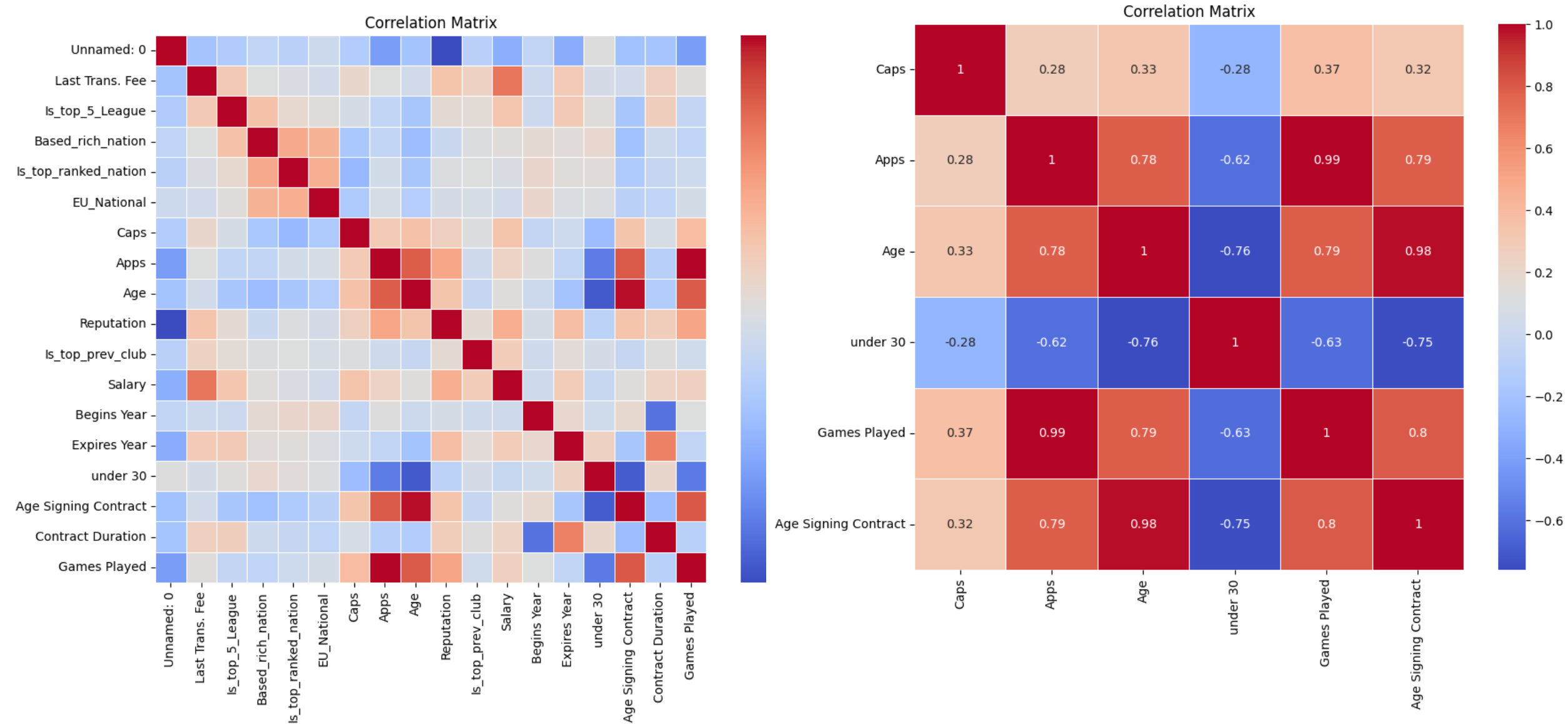
Data columns (total 21 columns):

#	Column	Non-Null Count	Dtype
---	-----	-----	-----
0	Division	40768 non-null	object
1	Based	40768 non-null	object
2	Nat	40768 non-null	object
3	Last Club	40768 non-null	object
4	Last Trans. Fee	40768 non-null	float64
5	Is_top_5_League	40768 non-null	bool
6	Based_rich_nation	40768 non-null	bool
7	Is_top_ranked_nation	40768 non-null	bool
8	EU_National	40768 non-null	bool
9	Caps	40768 non-null	int32
10	Apps	40768 non-null	int64
11	Age	40768 non-null	float64
12	Reputation	40768 non-null	float64
13	Is_top_prev_club	40768 non-null	bool
14	Salary	40768 non-null	float64
15	Begins Year	40768 non-null	int64
16	Expires Year	40768 non-null	int64
17	under 30	40768 non-null	bool
18	Age Signing Contract	40768 non-null	int32
19	Contract Duration	40768 non-null	int64
20	Games Played	40768 non-null	int64
dtypes: bool(6), float64(4), int32(2), int64(5), object(4)			

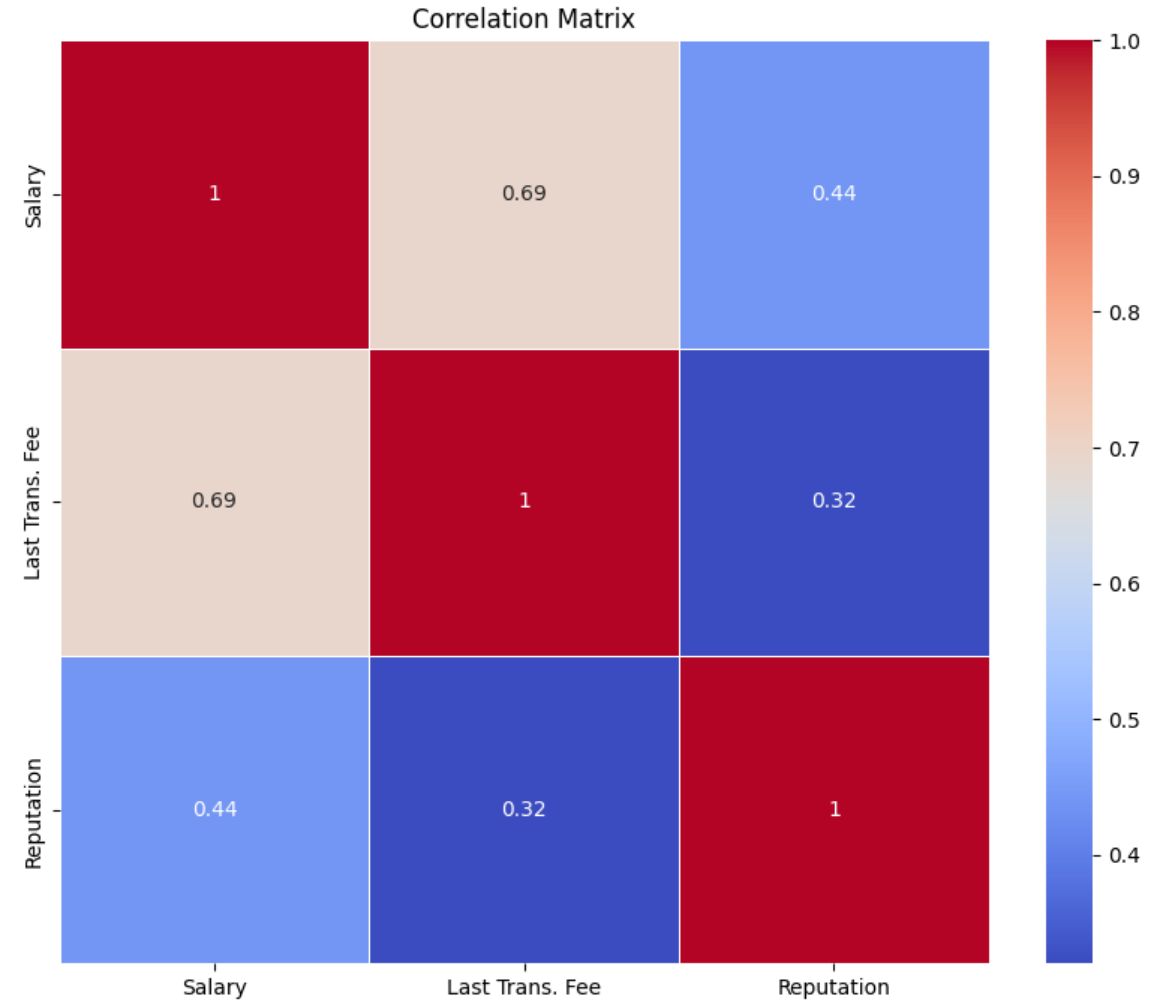
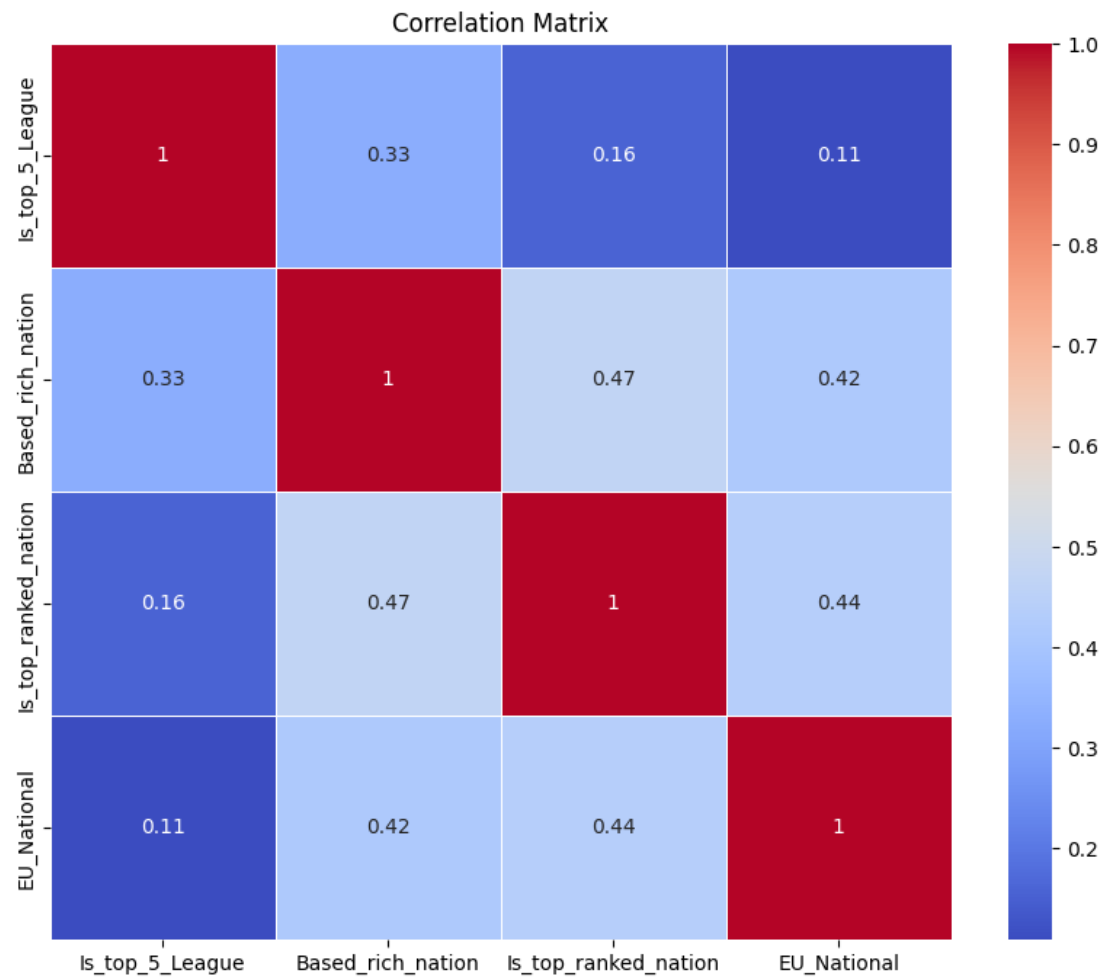
**CLEANED  
DATASET**



# CORRELATION BETWEEN FEATURES



# CORRELATION BETWEEN FEATURES

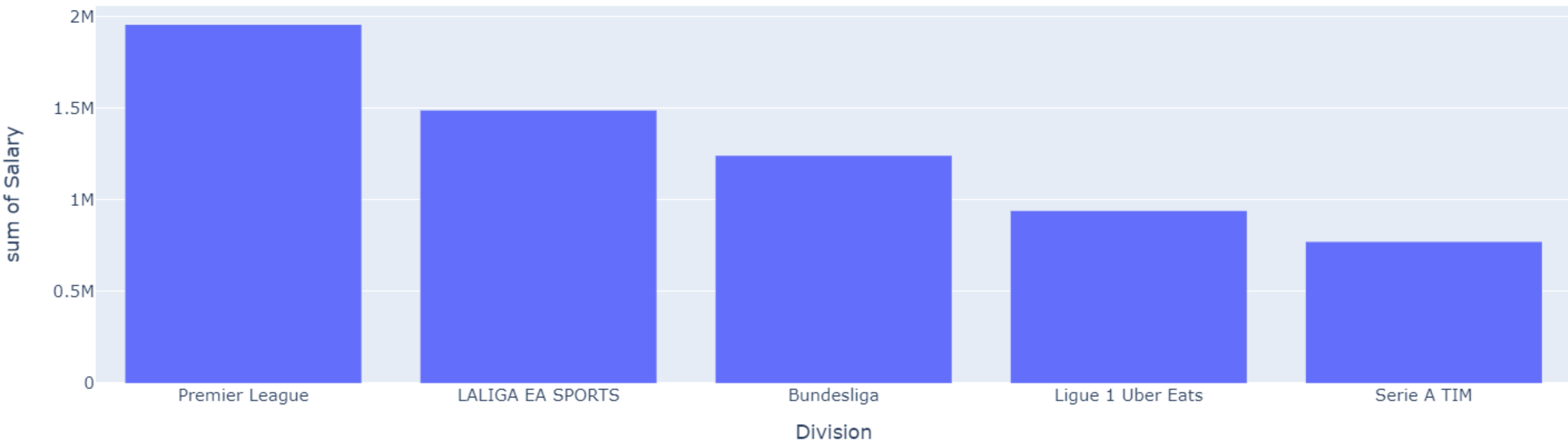


# SOME INSIGHTS INTO OUR DATASET



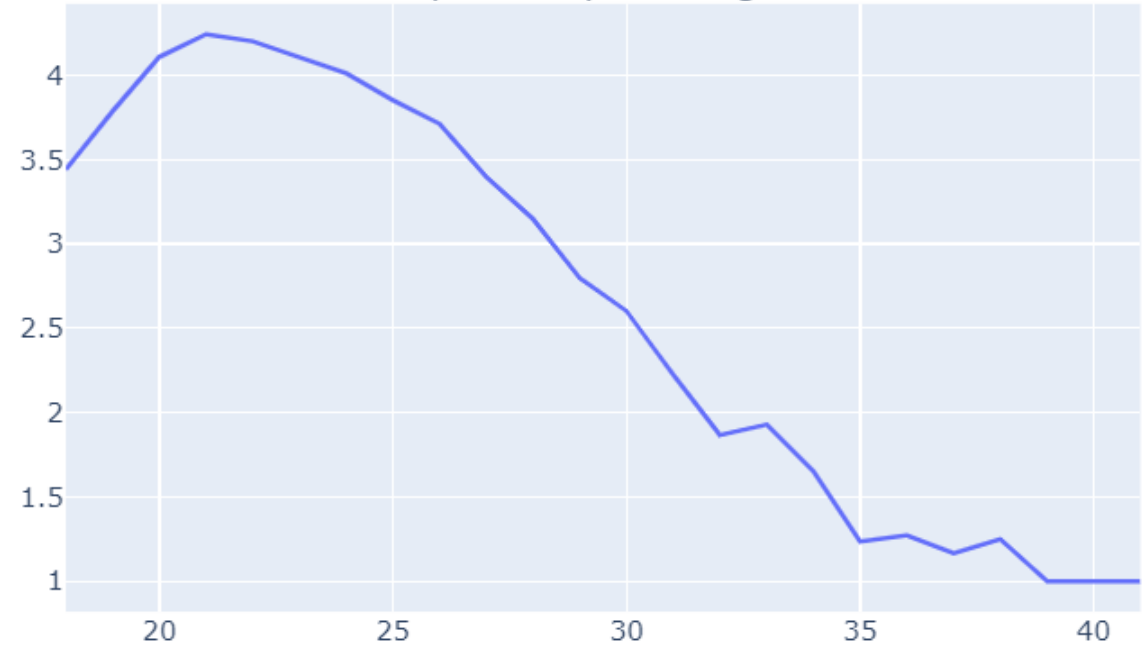


# AVERAGE SALARIES PAID IN TOP EUROPEAN LEAGUES

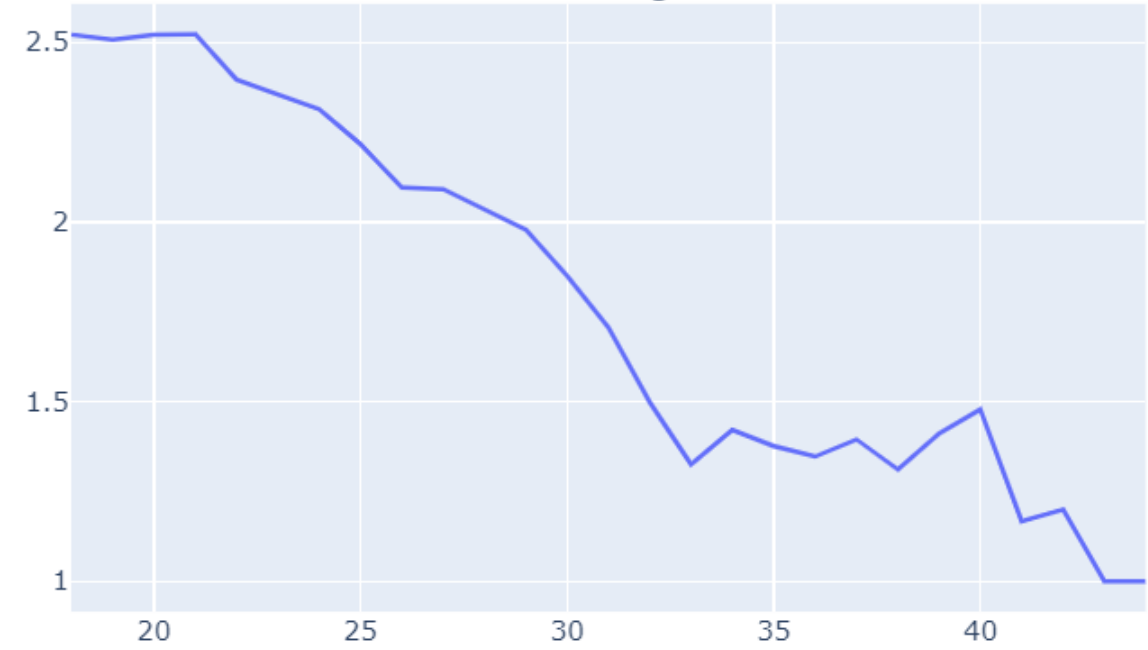


# AVERAGE CONTRACT DURATION PER AGE WHEN SIGNING THE CONTRACT

Top 5 European Leagues

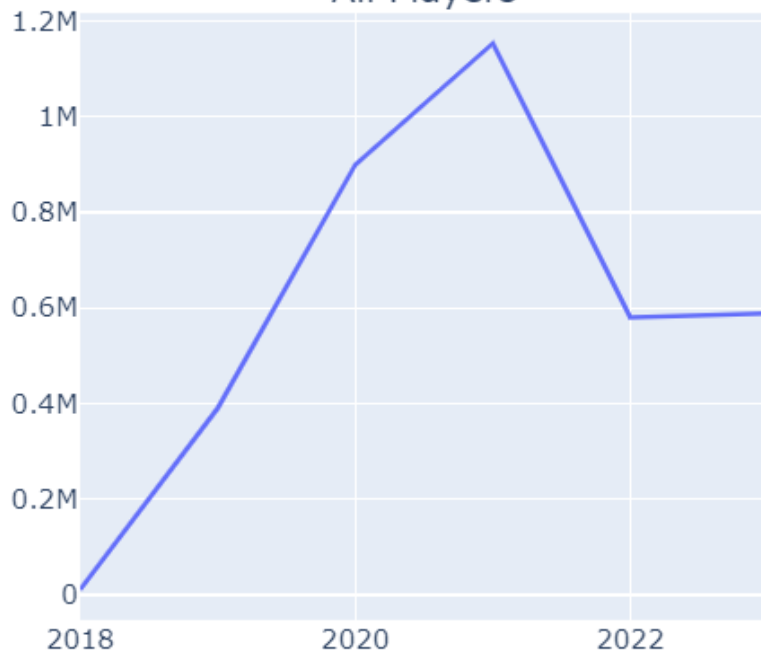


Other leagues

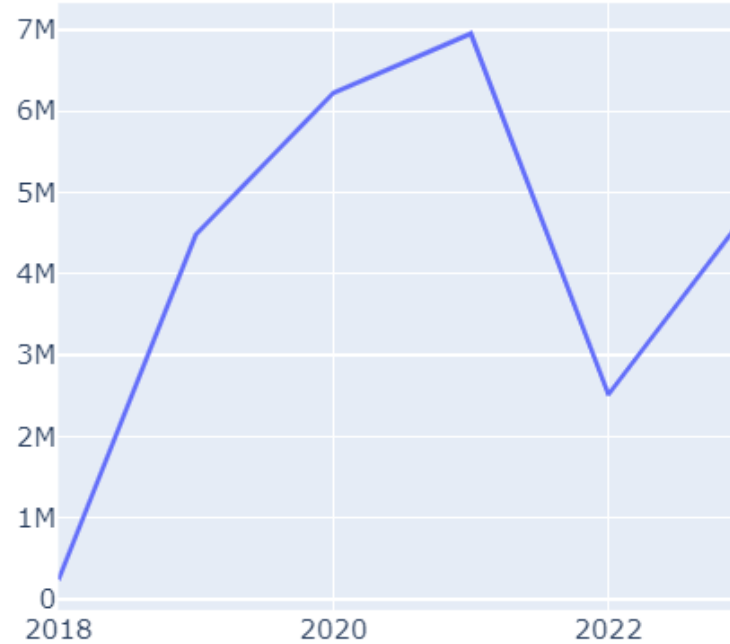


# AVERAGE SALARIES PER YEAR OF SIGNING CONTRACT

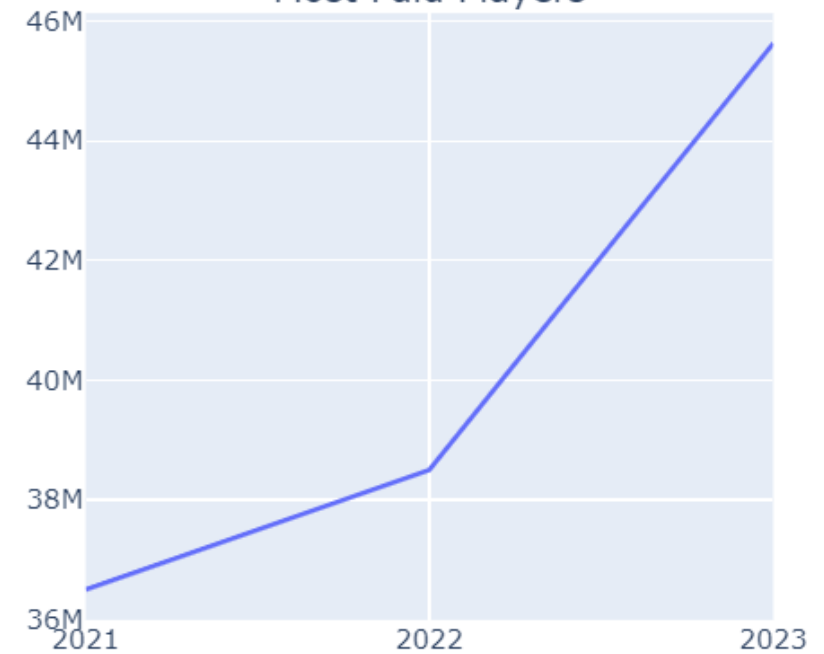
All Players



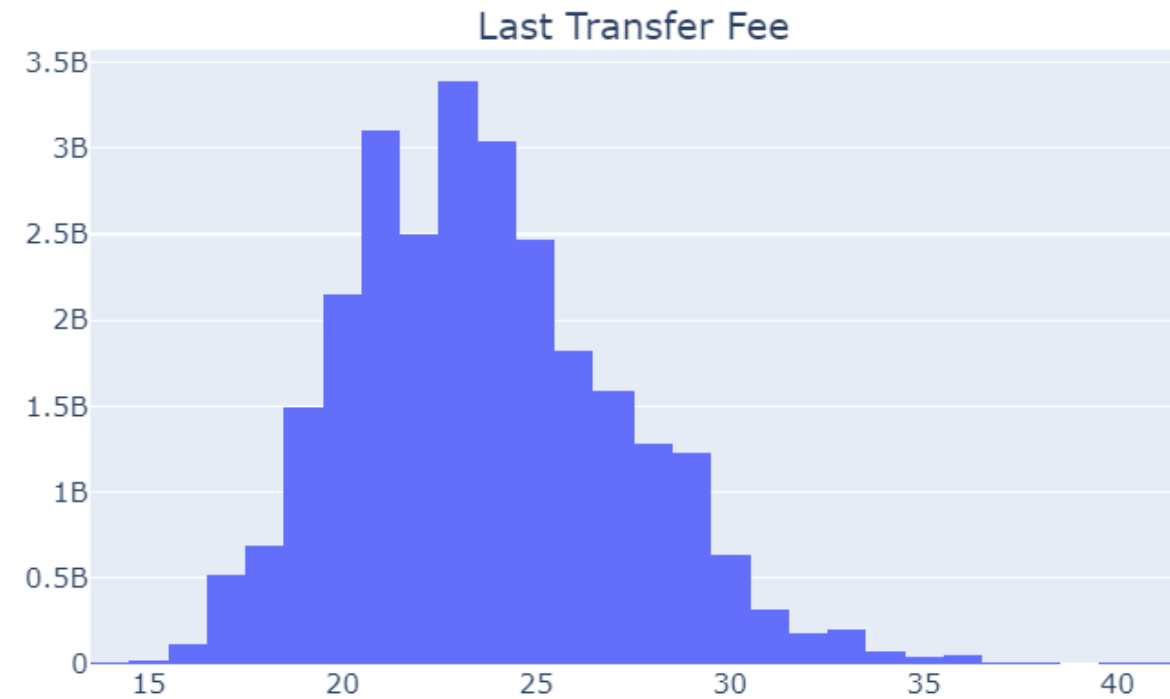
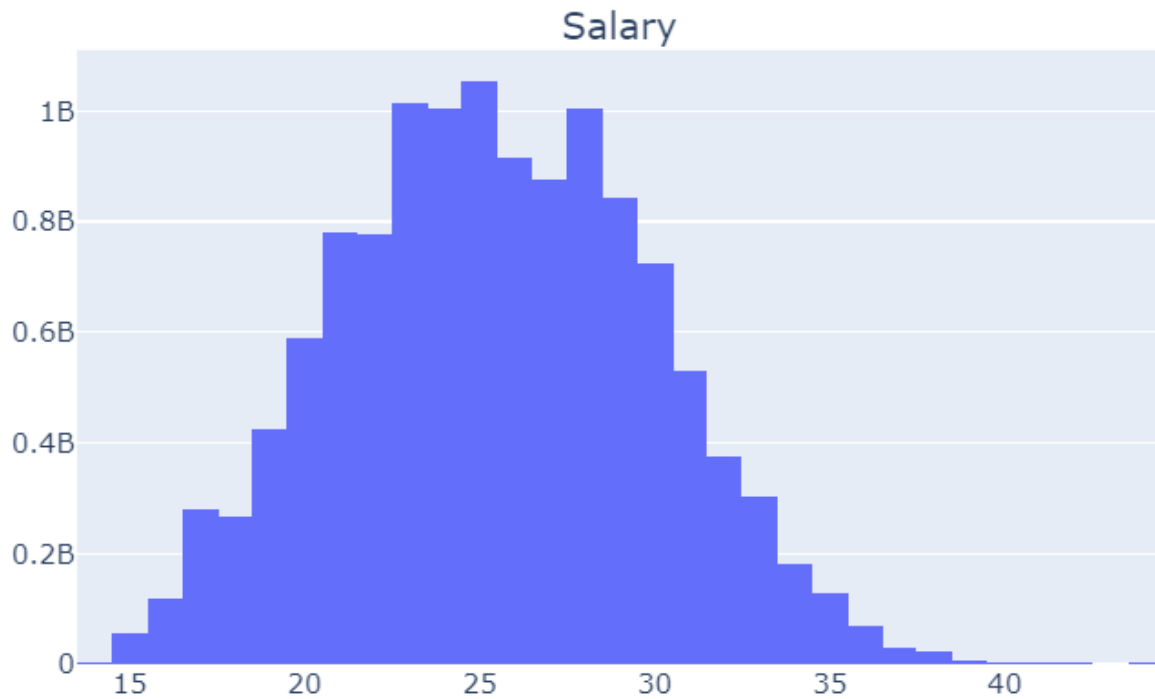
Top European Leagues Players



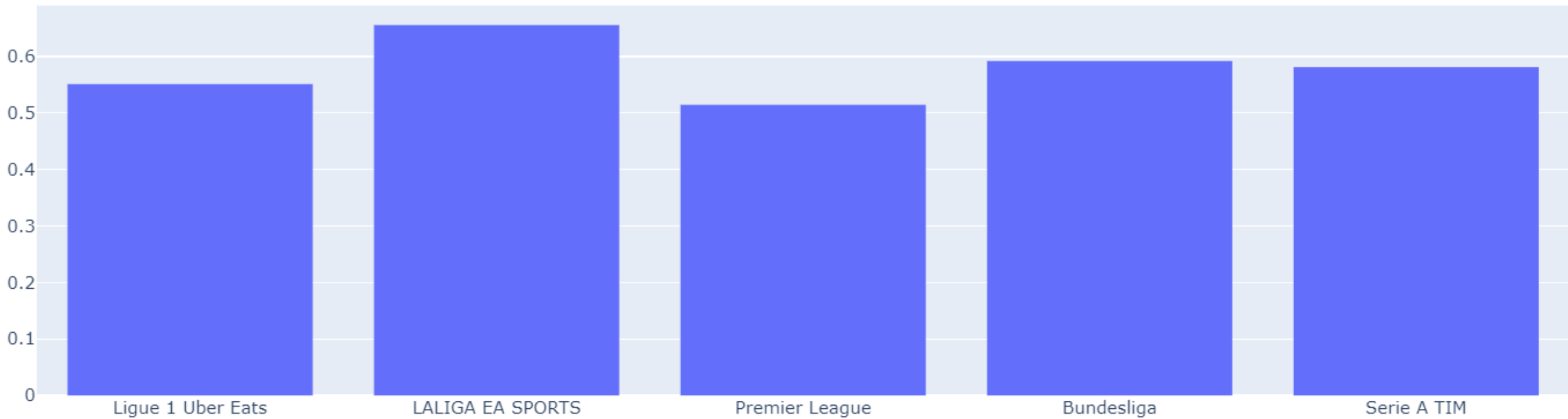
Most Paid Players



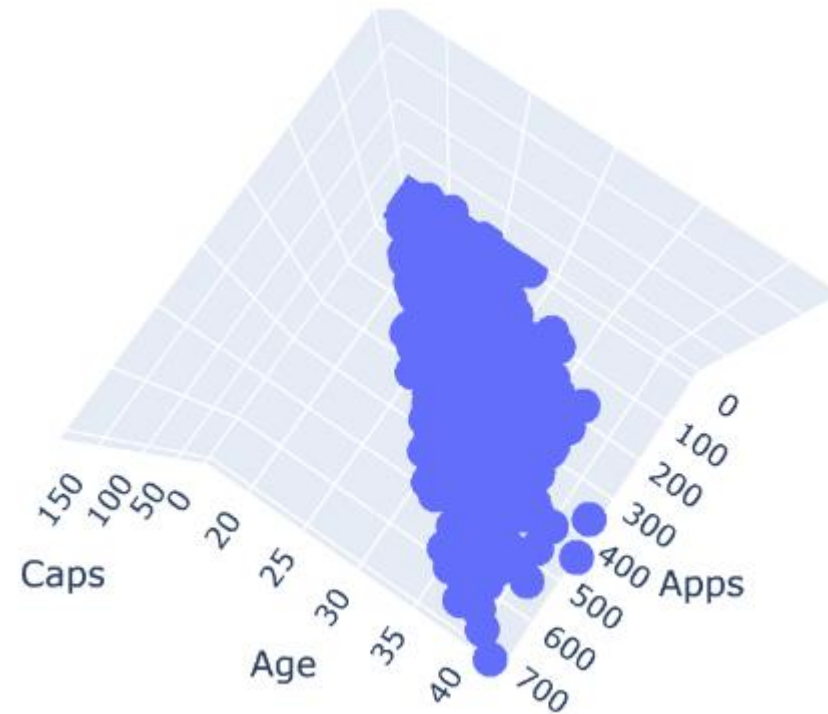
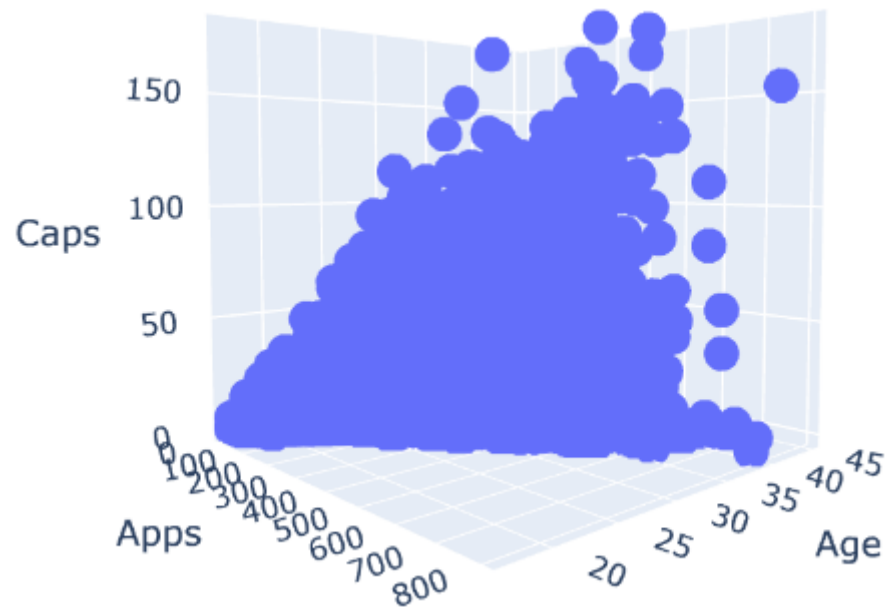
# SALARIES AND LAST TRANSFER FEES PER AGE



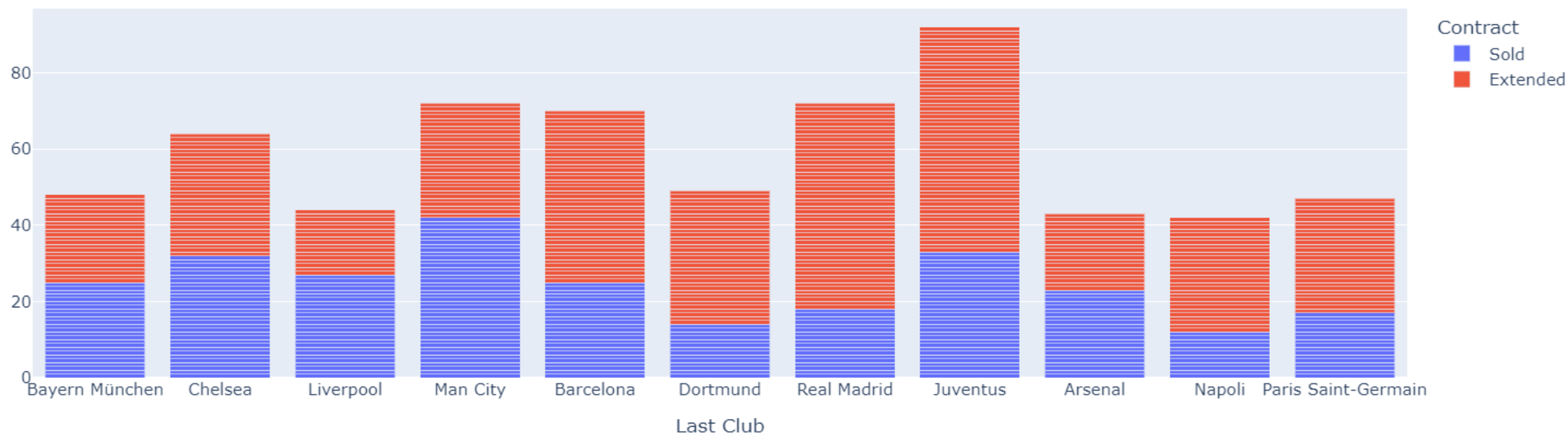
# PROPORTION OF LOCAL PLAYERS IN TOP EUROPEAN LEAGUES



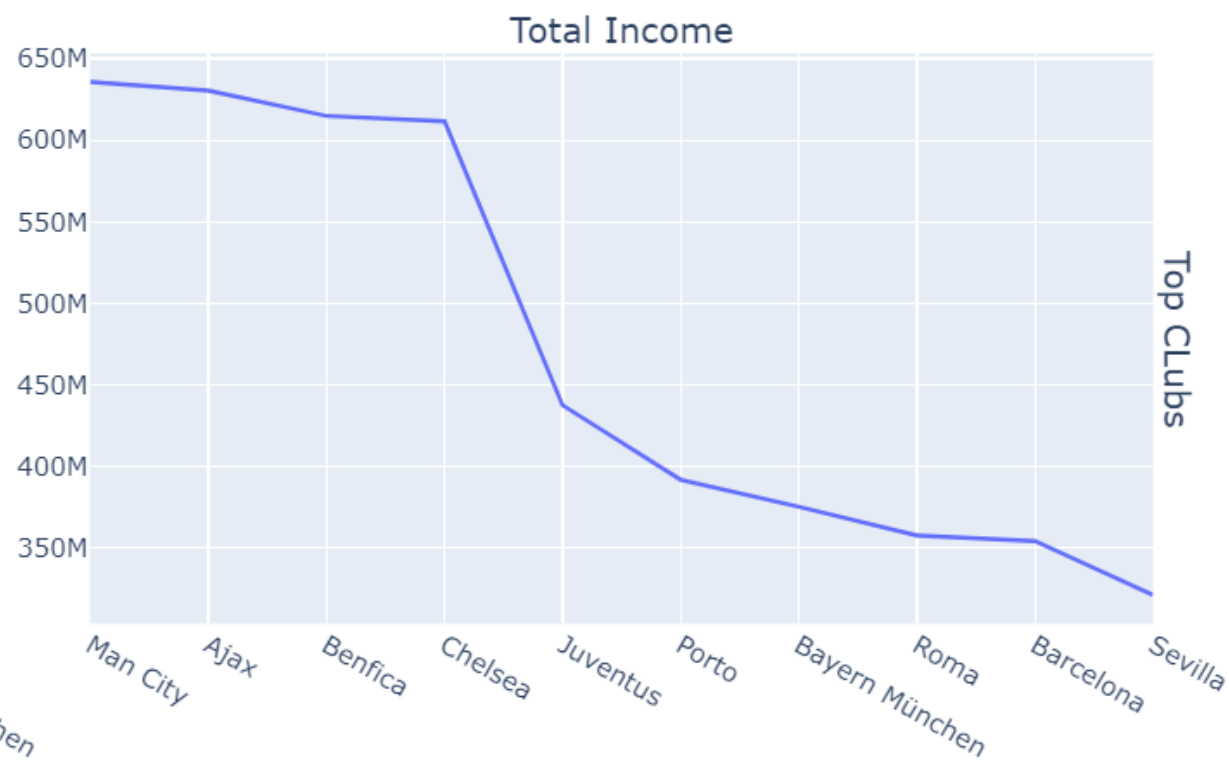
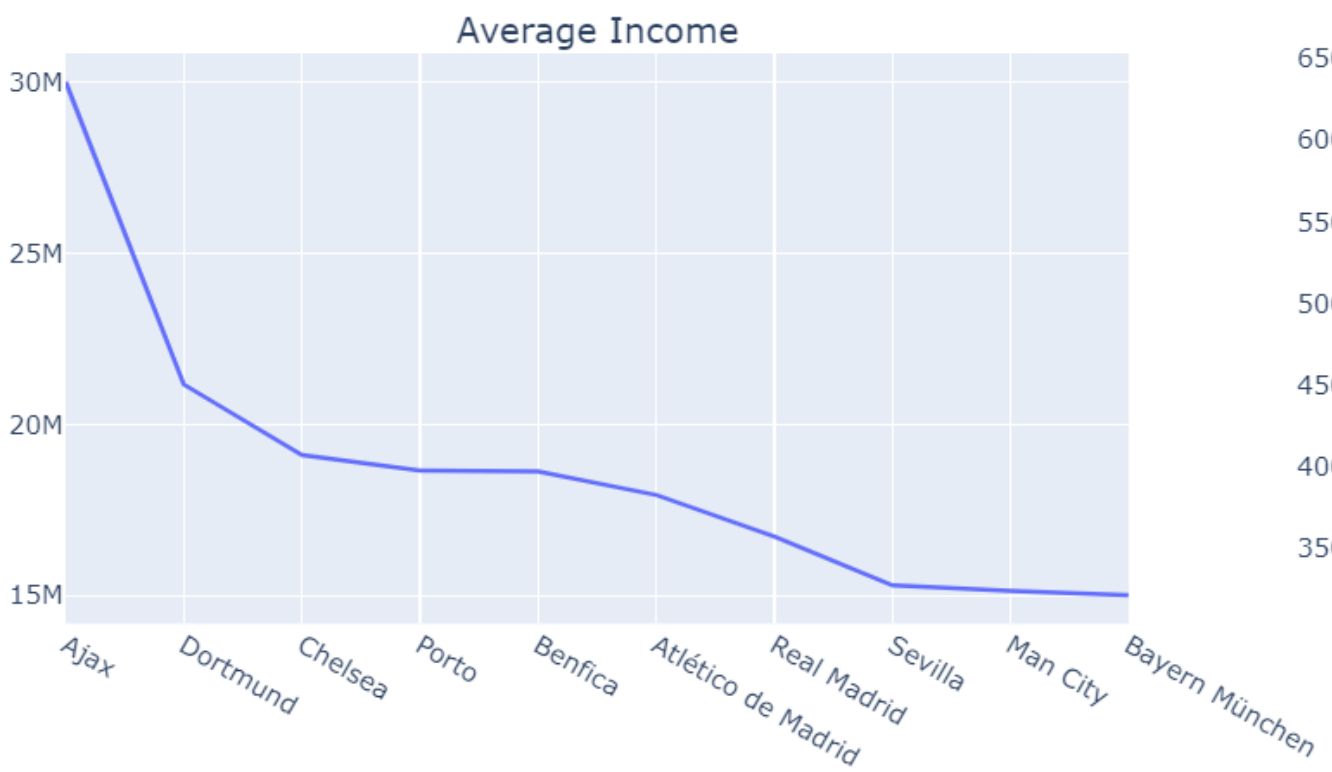
# EXPLORING RELATIONSHIP BETWEEN APPS, CAPS AND AGE



# NUMBER OF CONTRACTS SIGNED IN TOP CLUBS

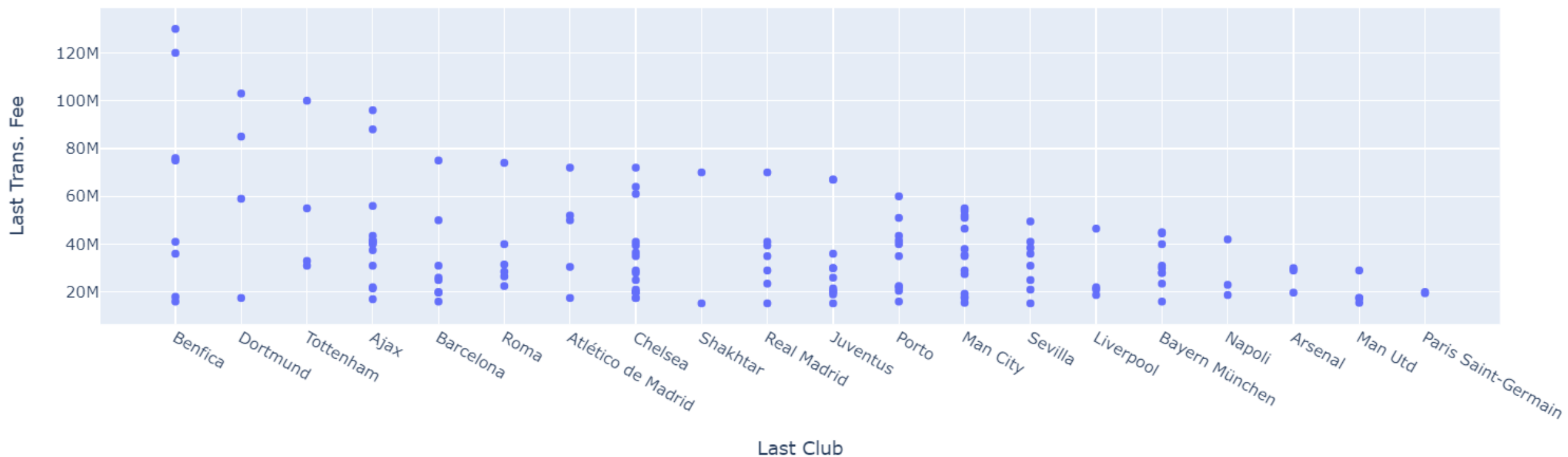


# TOP 10 CLUBS BASED ON INCOME THROUGH ONLY SELLING PLAYERS



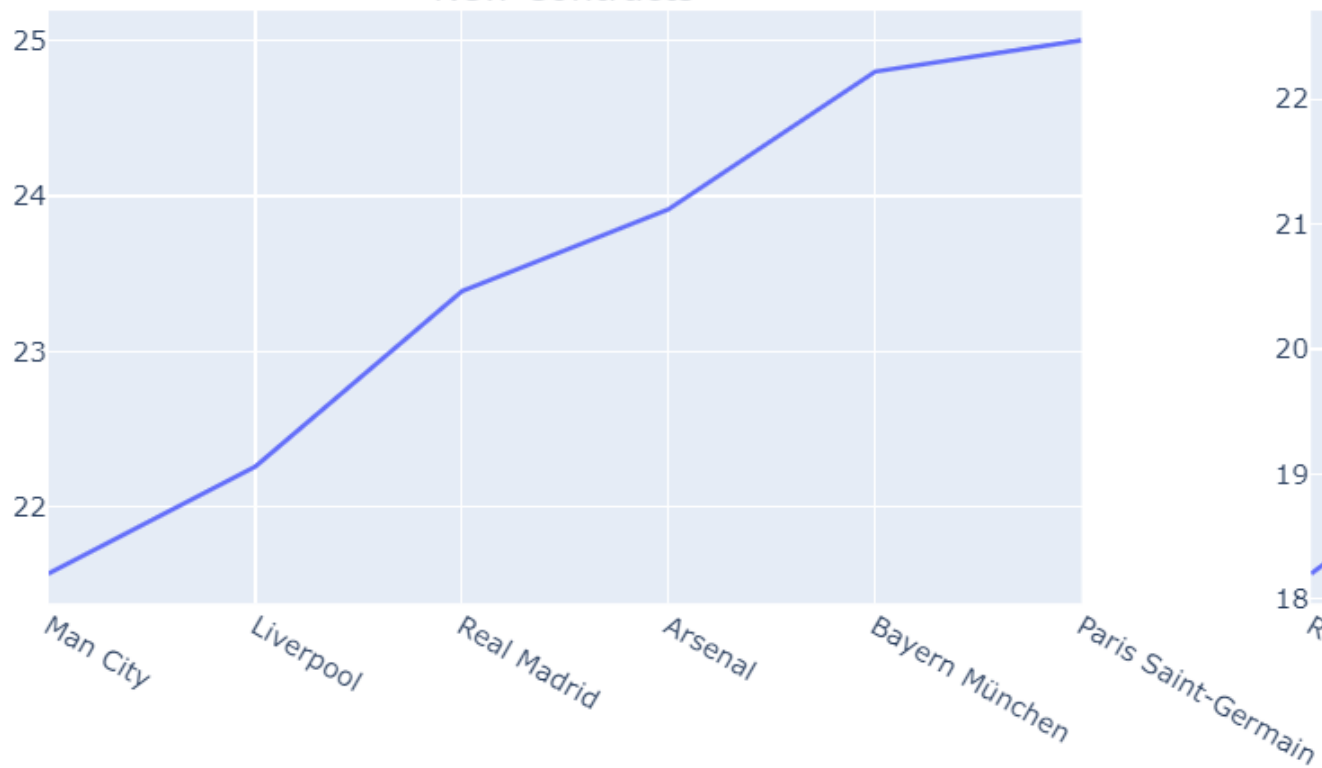


# MOST EXPENSIVE TRANSFERS FROM TOP CLUBS

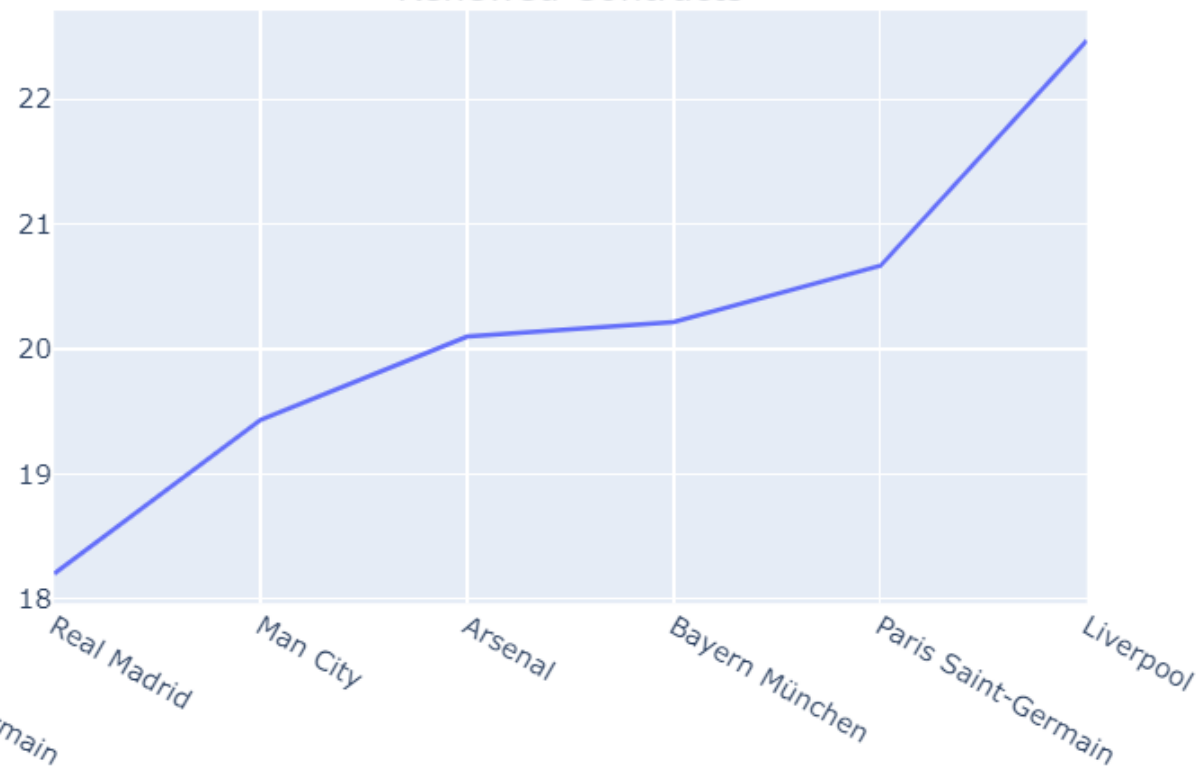


# TOP CLUBS CONTRACTS

New Contracts



Renewed Contracts



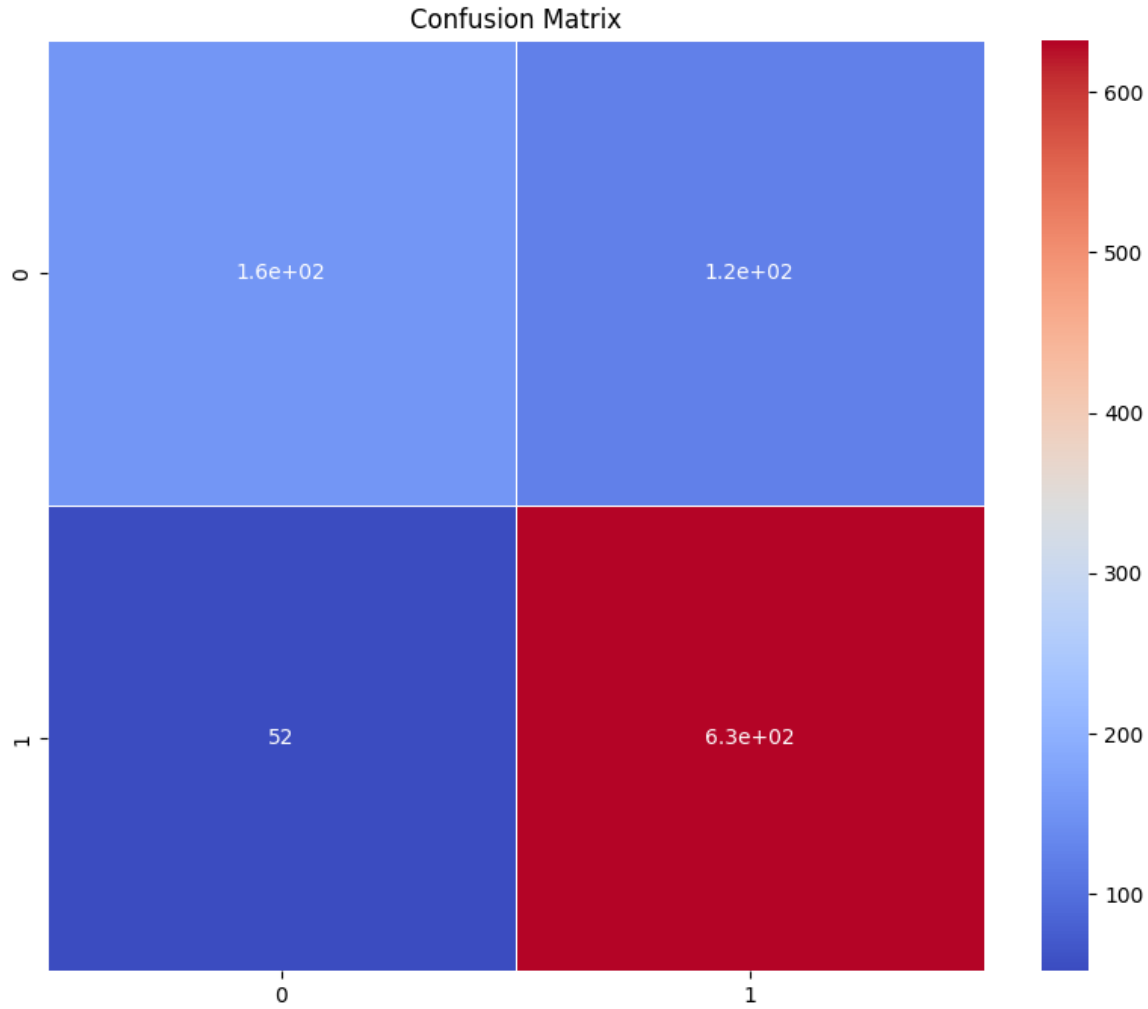
# REGRESSION AND CLASSIFICATION

- Some continuous variables we aim to predict using regression models:
  - Salary
  - Last Transfer Fee
  - Reputation Score
- discrete variable we aim to predict using classification models:
  - Whether the contract is long-term or short-term contract. Contracts longer than 3 years are considered long-term.

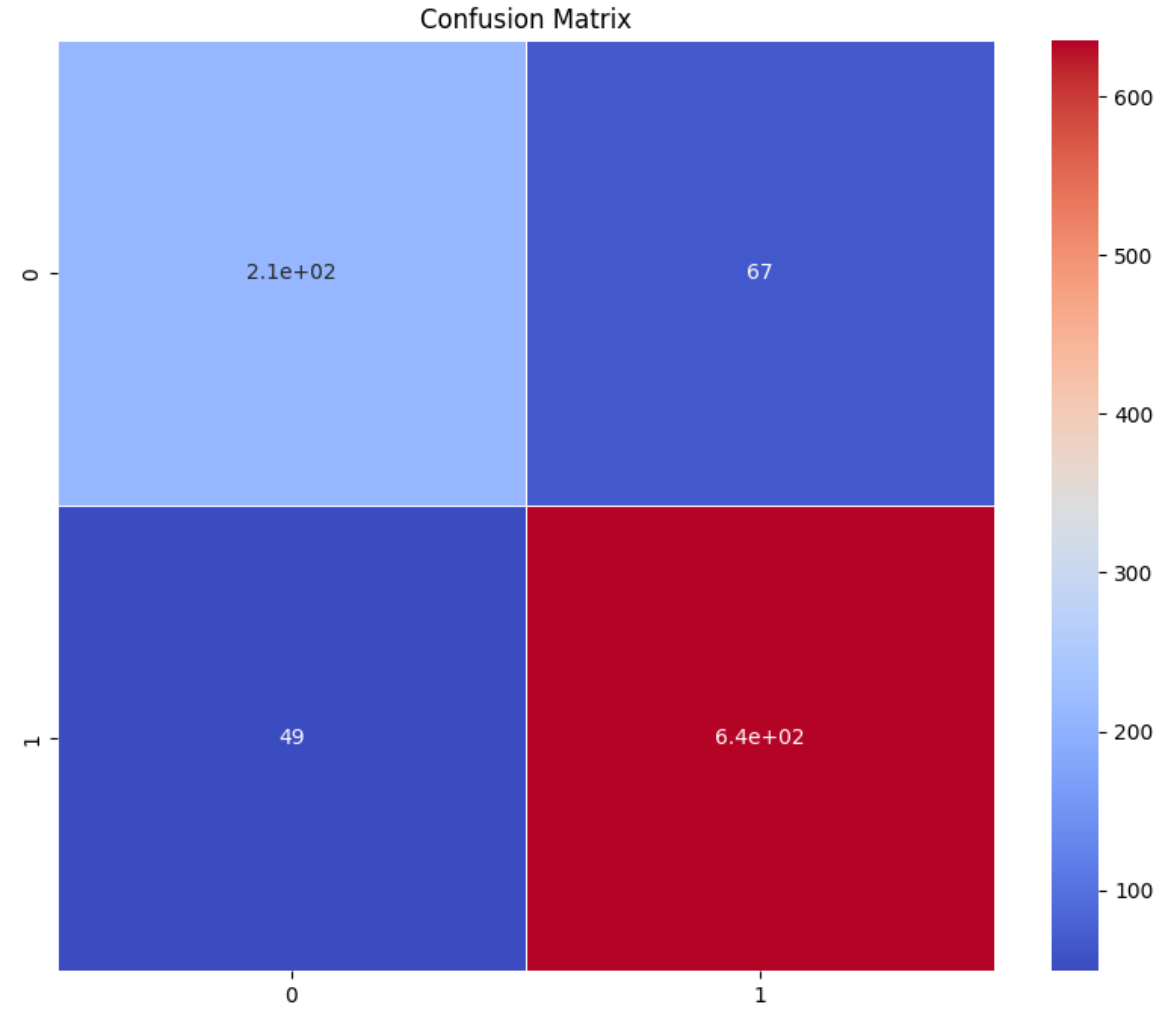


# DECISION TREE CLASSIFIER

Initial tree

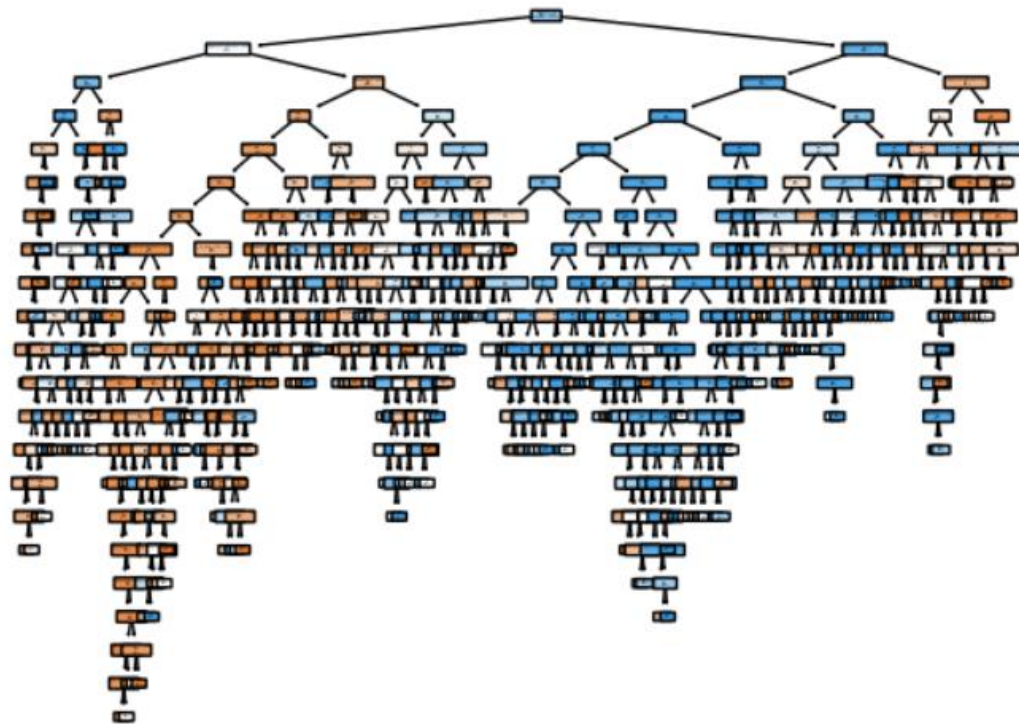


Parameter Tuning using grid search CV



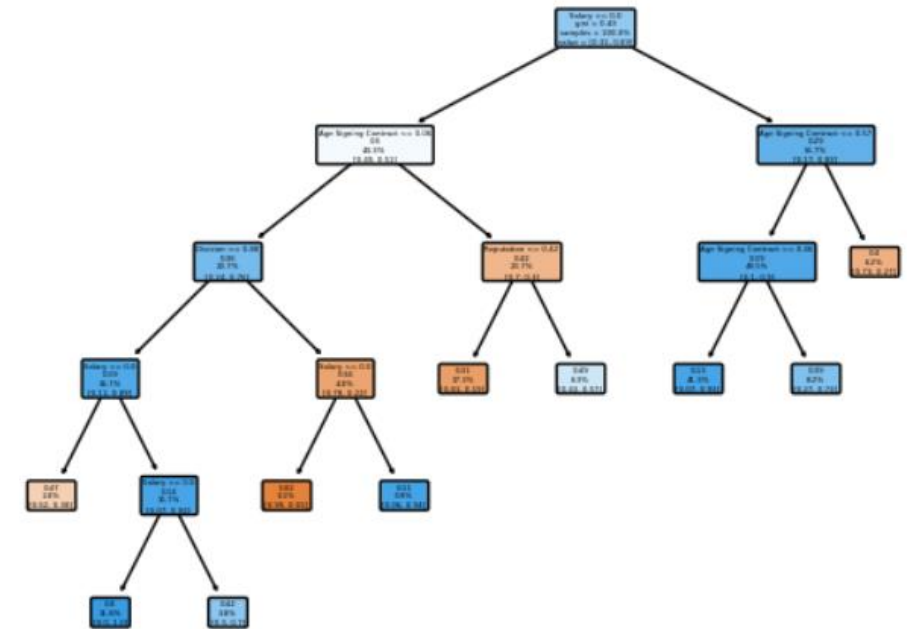
# EVALUATION METRICS BEFORE AND AFTER PARAMETER TUNING

Acc:0.77, Prec:0.87, Rec:0.84, F1:0.85



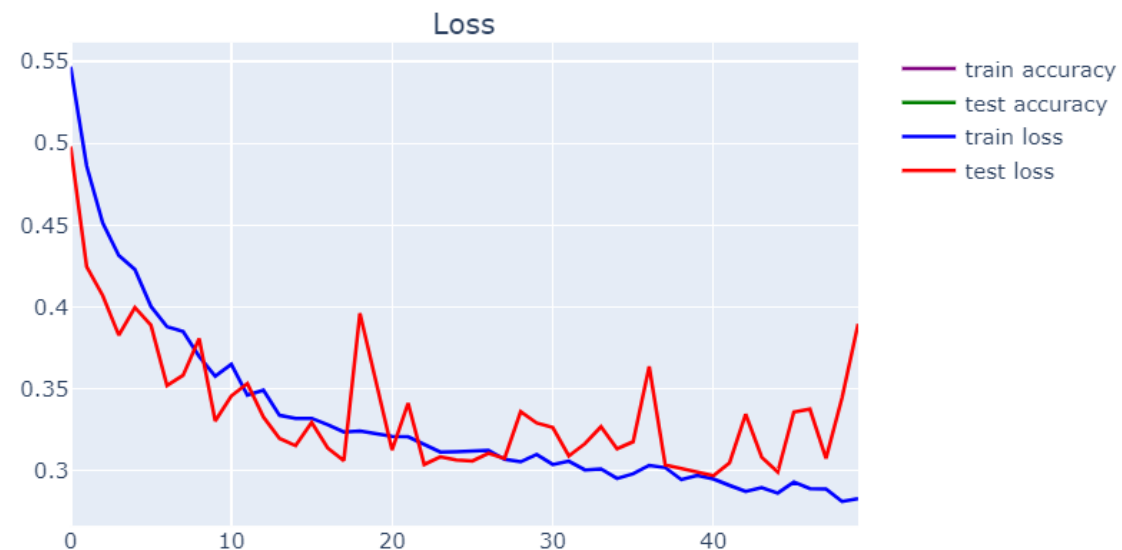
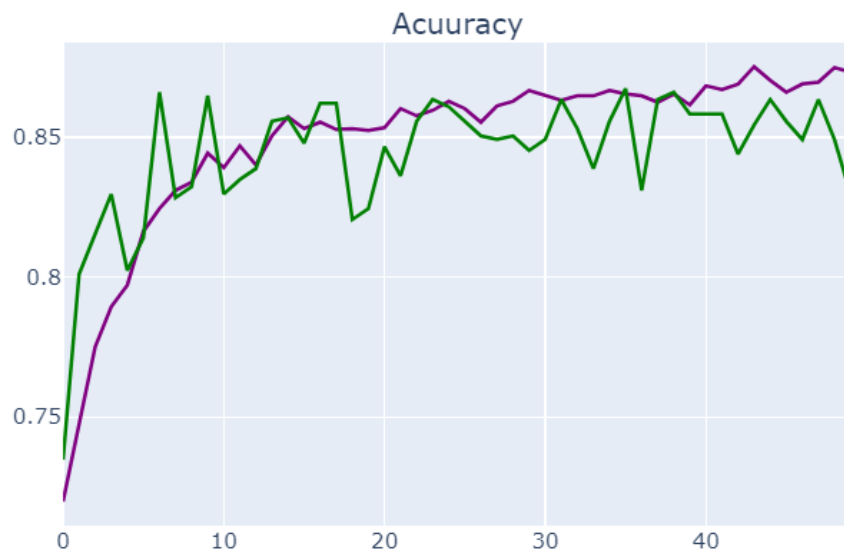
Acc:0.87, Prec:0.89, Rec:0.91, F1:0.9

Pruning →

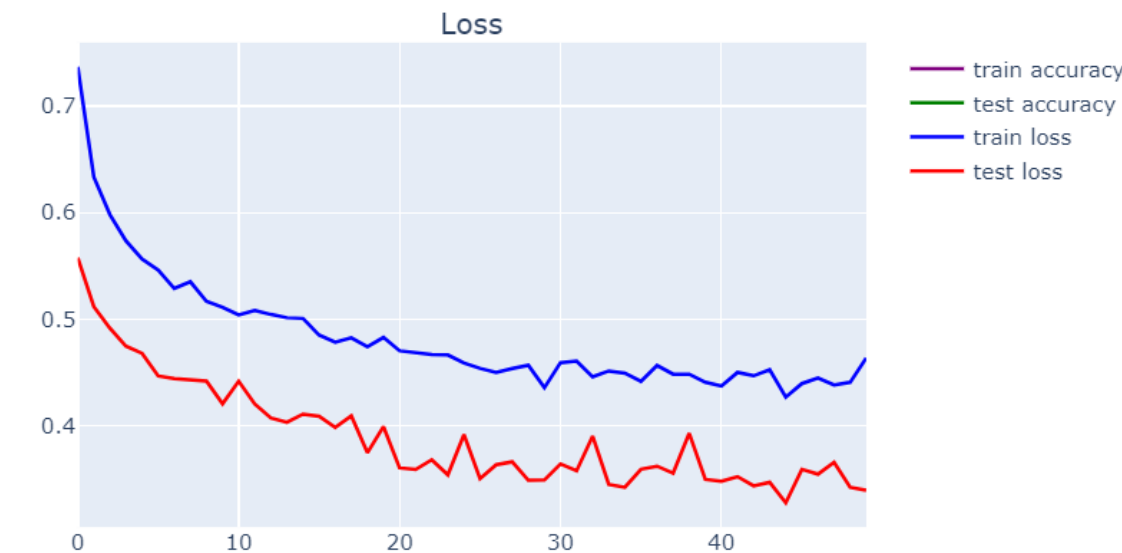
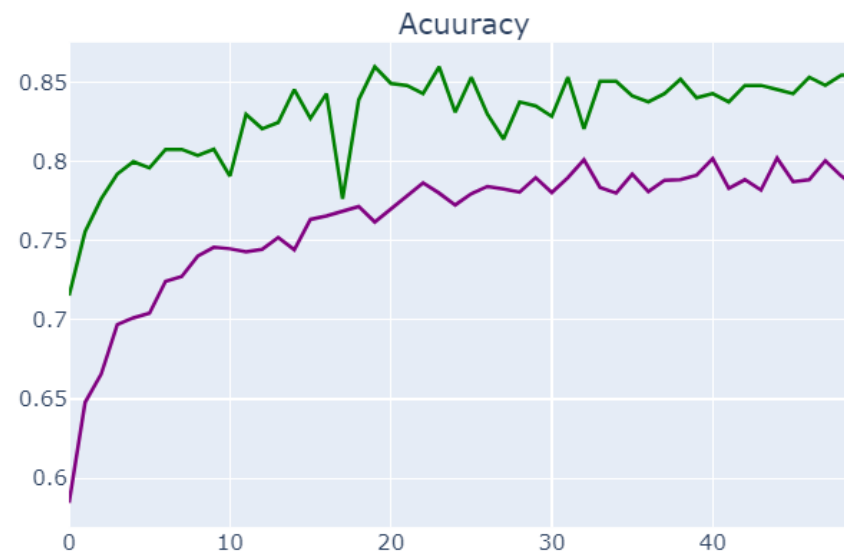


# NEURAL NETWORK CLASSIFIER

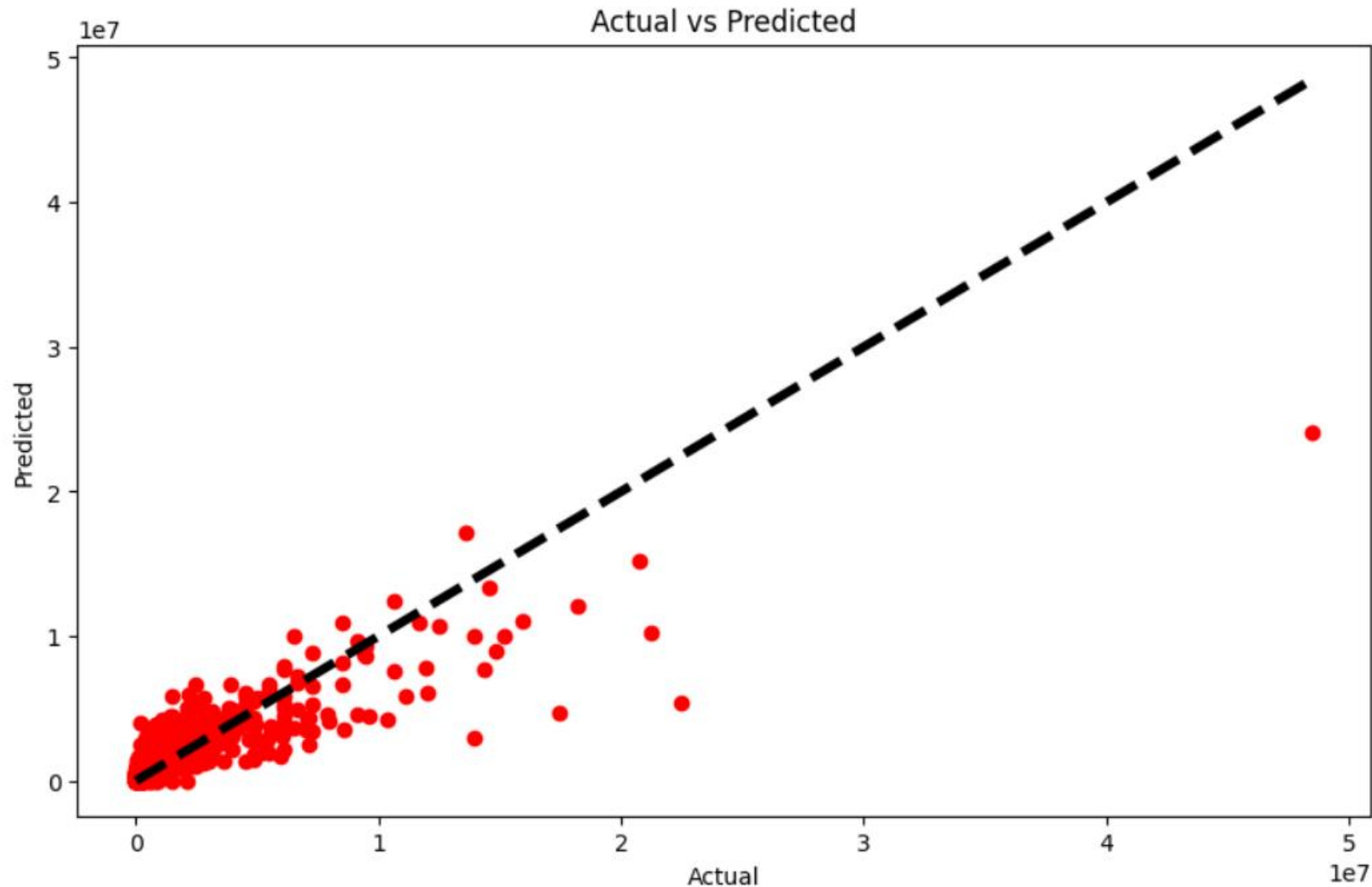
Without any  
regularization



Using  
regularization  
methods



# MULTIPLE REGRESSION

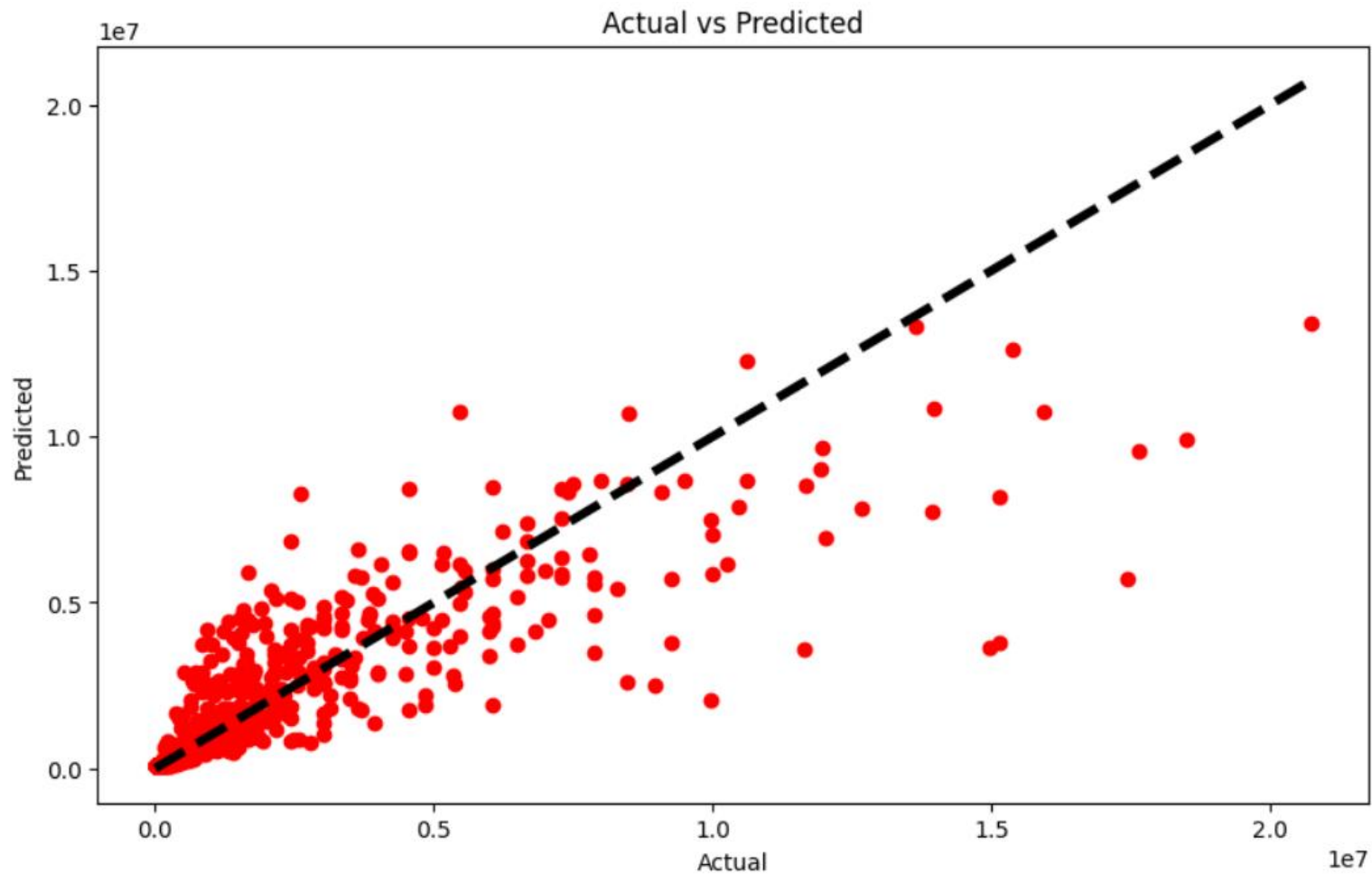


*Target variable: Salary*

$R^2 : 0.71$



# NEURAL NETWORK



*Target variable: Salary*

$R^2 : 0.79$

