Football Manager: Player Recommendations

What is Football Manager?

Soccer Manager Game

- Big database of players
- Manage a team
- Wage Budget

The Problem

- Scouting can be Expensive
  - Have to pay for packages
- Scouting can be Unreliable
  - Wasted wages
- Implement a Query Tool
  - Enter position
  - Specific attributes
  - Filters
  - Analysis

The packages owned determine the number of players available to search through as well as increasing the knowledge level and player role visibility of players covered. Youth packages cover players aged 21 and under.

SENIOR PACKAGES

WORLD
Best package available. Covers all players in the world.
€279K (€23K p/m)

EUROPE
Covers players in that continent.
€209K (€17.5K p/...

UK & IRELAND
Covers players in that region.
€139K (€11.5K p/...

ENGLAND
Covers players in that nation.
€59K (€4.9K p/m)

PREMIER LEAGUE, EFL CHAMPIONSHIP, EFL LEAGUE ONE
Covers players in your division and surrounding divisions.
€35.5K (€3K p/m)

EFL CHAMPIONSHIP
Lowest package available. Covers players in your division.
€11.75K (€975 p/...

NO PACKAGE
Covers only players known to the club's staff.
€0 (€0 p/m)

YOUTH PACKAGES

WORLD
Best package available. Covers all players in the world.
€557K (€46.5K p...

EUROPE
Covers players in that continent.
€418K (€35K p/m)

UK & IRELAND
Covers players in that region.
€279K (€23K p/m)

ENGLAND
Covers players in that nation.
€118K (€9.75K p/...

PREMIER LEAGUE, EFL CHAMPIONSHIP, EFL LEAGUE ONE
Covers players in your division and surrounding divisions.
€71K (€6K p/m)

EFL CHAMPIONSHIP
Lowest package available. Covers players in your division.
€23.5K (€2K p/m)

NO PACKAGE
Covers only players known to the club's staff.
€0 (€0 p/m)

Setting Up Data

1. Tough to find datasets
   https://www.kaggle.com/ajinkyablaze/football-manager-data
2. Separate dataset (160k x 89) into 15 smaller datasets (position >= 12). Each set stored in dictionary with an index. (O(n^2))
3. For each dataset, drop unnecessary columns and fit attributes into a weighted function $w\_1 * a\_1 + w\_2*a\_2 + ... + w\_m*a\_m$.
4. For every player predict a value based on the function (O(m*n)) .
5. Sort based on that value
6. Huge runtime, but only need to do once.

```
Index(['UID', 'Name', 'NationID', 'Born', 'Age', 'IntCaps', 'IntGoals',
       'U21Caps', 'U21Goals', 'Height', 'Weight', 'AerialAbility',
       'CommandOfArea', 'Communication', 'Eccentricity', 'Handling', 'Kicking',
       'OneOnOnes', 'Reflexes', 'RushingOut', 'TendencyToPunch', 'Throwing',
       'Corners', 'Crossing', 'Dribbling', 'Finishing', 'FirstTouch',
       'Freekicks', 'Heading', 'LongShots', 'Longthrows', 'Marking', 'Passing',
       'PenaltyTaking', 'Tackling', 'Technique', 'Aggression', 'Anticipation',
       'Bravery', 'Composure', 'Concentration', 'Vision', 'Decisions',
       'Determination', 'Flair', 'Leadership', 'OffTheBall', 'Positioning',
       'Teamwork', 'Workrate', 'Acceleration', 'Agility', 'Balance', 'Jumping',
       'LeftFoot', 'NaturalFitness', 'Pace', 'RightFoot', 'Stamina',
       'Strength', 'Consistency', 'Dirtiness', 'ImportantMatches',
       'InjuryProness', 'Versatility', 'Adaptability', 'Ambition', 'Loyalty',
       'Pressure', 'Professional', 'Sportsmanship', 'Temperament',
       'Controversy', 'PositionsDesc', 'Goalkeeper', 'Sweeper', 'Striker',
       'AttackingMidCentral', 'AttackingMidLeft', 'AttackingMidRight',
       'DefenderCentral', 'DefenderLeft', 'DefenderRight',
       'DefensiveMidfielder', 'MidfielderCentral', 'MidfielderLeft',
       'MidfielderRight', 'WingBackLeft', 'WingBackRight'],
      dtype='object')
```

Querying Data

1. Input Position, k, attributes, filter(height, age, weight etc.).
2. Filter rows.
3. Return top k rows.
4. If attributes defined, Give every player a value based on attributes inputted: newValue = value + avg(attributes). Then sort. (O(nlogn)).

Attribute Analysis

1. A player may not know exactly what attributes they need.
2. Input position
3. Query for top 200 players in the position. (O(nlogn))
4. Plot attributes (48 plots)
5. Skew left more important

DC Attributes
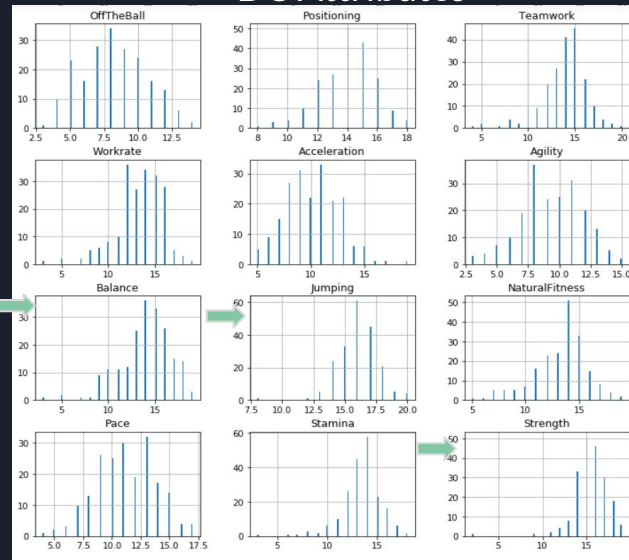
Query : New Striker

GetTopK(Position = 'ST', k = 10)

	UID	Name	NationID	Born	Age	IntCaps	IntGoals	U21Caps	U21Goals	Height	...	Adaptability	Ambition	Loyalty	Pressure	Profess
125352	735216	Cristiano Ronaldo	788	05-02-1985	31	133	61	10	3	186	...	18	20	13	19	
90196	5132312	Gareth Bale	801	16-07-1989	26	61	22	13	3	186	...	8	14	11	14	
119672	678445	Theofanis Gekas	772	23-05-1980	36	78	24	0	0	179	...	14	17	8	16	
114345	661804	Mario Mandzukic	761	21-05-1986	30	69	24	9	1	190	...	15	15	6	18	
115242	6705233	Aritz Aduriz	796	11-02-1981	35	9	1	0	0	182	...	14	14	18	14	
60570	35006608	Kostas Mitroglou	772	12-03-1988	28	47	9	17	4	188	...	14	16	14	11	
154906	950975	Radamel Falcao	1653	10-02-1986	30	62	25	13	8	177	...	13	18	14	19	
132367	762557	Øyvind Gausdal	786	14-03-1986	30	0	0	0	0	192	...	18	13	13	12	
27884	18009139	Christian Benteke	757	03-12-1990	25	29	7	9	4	191	...	18	16	8	16	
126257	7458323	Imanol Agirretxe	796	24-02-1987	29	0	0	0	0	187	...	13	11	17	12	

More Speed

GetTopK(Position = 'ST', k = 10, Attributes = 'Acceleration, Pace')

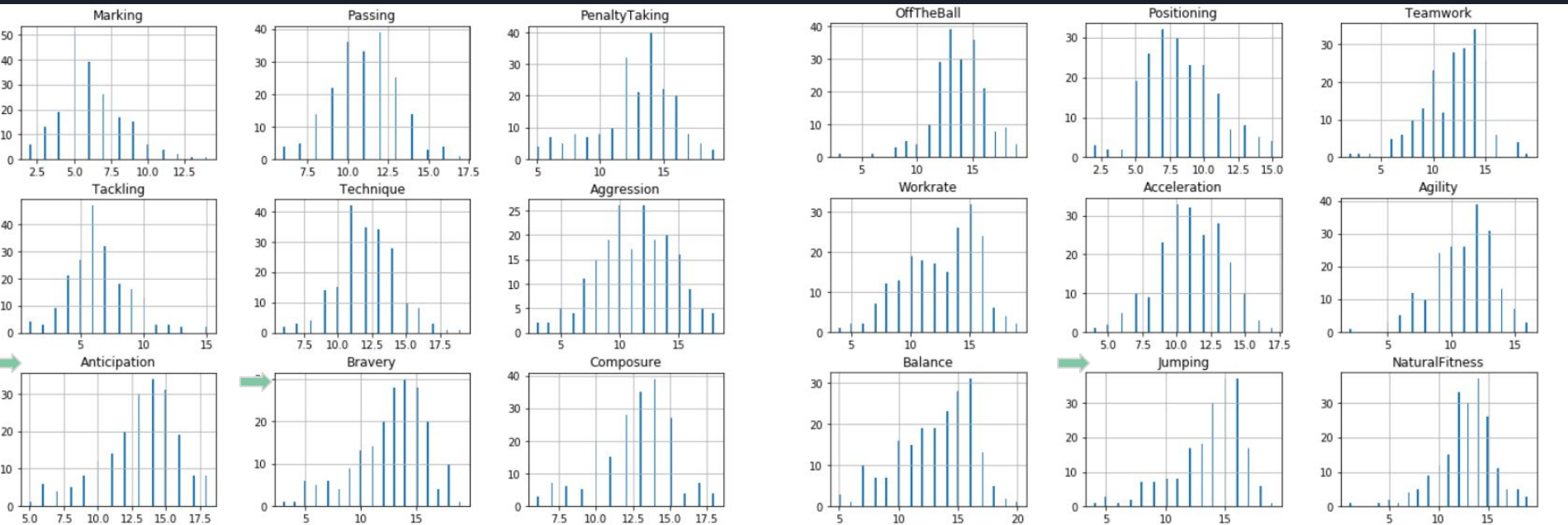| UID		Name	NationID	Born	Age	IntCaps	IntGoals	U21Caps	U21Goals	Height	...	Adaptability	Ambition	Loyalty	Pressure	Profe
74688	43017197	Pierre-Emerick Aubameyang	19	18-06-1989	27	49	20	0	0	187	...	13	15	10	12	
90196	5132312	Gareth Bale	801	16-07-1989	26	61	22	13	3	186	...	8	14	11	14	
1420	12021802	Emma Nwachi	38	12-10-1986	29	0	0	0	0	180	...	17	10	11	10	
99074	5680375	David Radebe	45	27-09-1979	36	1	0	0	0	189	...	17	7	14	5	
1562	12036758	Kofi Boakye	21	05-04-1987	29	1	0	0	0	168	...	20	14	9	11	
148619	90008752	Jamie Vardy	765	11-01-1987	29	11	4	0	0	178	...	8	16	13	15	
87115	51001380	Javier Hernández	379	01-06-1988	28	86	45	5	1	175	...	14	14	13	16	
125352	735216	Cristiano Ronaldo	788	05-02-1985	31	133	61	10	3	186	...	18	20	13	19	
126266	7458500	Lionel Messi	1649	24-06-1987	29	113	55	18	14	170	...	17	20	13	19	
51425	28010491	Danny Welbeck	765	26-11-1990	25	34	14	14	5	185	...	9	15	16	14	

Better Shooting

GetTopK(Position = 'ST', k = 10, Attributes = 'LongShots, Finishing')

	UID	Name	NationID	Born	Age	IntCaps	IntGoals	U21Caps	U21Goals	Height	...	Adaptability	Ambition	Loyalty	Pressure	Professio...
125352	735216	Cristiano Ronaldo	788	05-02-1985	31	133	61	10	3	186	...	18	20	13	19	
90196	5132312	Gareth Bale	801	16-07-1989	26	61	22	13	3	186	...	8	14	11	14	
126266	7458500	Lionel Messi	1649	24-06-1987	29	113	55	18	14	170	...	17	20	13	19	
145916	8708271	Lukas Podolski	771	04-06-1985	31	129	48	5	0	182	...	14	12	17	18	
158329	962988	Gonzalo Higuaín	1649	10-12-1987	28	63	30	0	0	184	...	18	16	11	12	
147278	8828106	Oscar Cardozo	1655	20-05-1983	33	49	9	0	0	193	...	12	16	10	16	
134922	78000335	Luis Suárez	1657	24-01-1987	29	84	46	4	2	182	...	15	18	8	15	
21677	142525	Robin van Persie	784	06-08-1983	32	101	50	12	1	183	...	17	18	10	17	
32550	19189145	Talisca	1651	01-02-1994	22	0	0	2	0	188	...	12	11	13	14	
154906	950975	Radamel Falcao	1653	10-02-1986	30	62	25	13	8	177	...	13	18	14	19	

Don't Know?

getDistribution('ST')

New Query

GetTopK(Position = 'ST', k = 10, Attributes = 'Anticipation, Bravery, Jumping')

UID		Name	NationID	Born	Age	IntCaps	IntGoals	U21Caps	U21Goals	Height	...	Adaptability	Ambition	Loyalty	Pressure	Profess
125352	735216	Cristiano Ronaldo	788	05-02-1985	31	133	61	10	3	186	...	18	20	13	19	
115242	6705233	Aritz Aduriz	796	11-02-1981	35	9	1	0	0	182	...	14	14	18	14	
135037	78005742	Edinson Cavani	1657	14-02-1987	29	76	27	13	9	188	...	10	16	12	14	
87115	51001380	Javier Hernández	379	01-06-1988	28	86	45	5	1	175	...	14	14	13	16	
104209	6002115	Claudio Pizarro	1656	03-10-1978	37	85	20	0	0	187	...	16	13	10	18	
95564	55002038	Diego Costa	796	07-10-1988	27	10	1	0	0	186	...	16	16	11	15	
64420	37008621	Luuk de Jong	784	27-08-1990	25	12	3	18	5	188	...	12	16	12	10	
60030	34000647	Olivier Giroud	769	30-09-1986	29	55	20	0	0	192	...	15	14	14	15	
114345	661804	Mario Mandzukic	761	21-05-1986	30	69	24	9	1	190	...	15	15	6	18	

Age Limit

GetTopK(Position = 'ST', k = 10, Attributes = 'Anticipation, Bravery, Jumping', Age = 29)

	UID	Name	NationID	Born	Age	IntCaps	IntGoals	U21Caps	U21Goals	Height	...	Adaptability	Ambition	Loyalty	Pressure	Profess
135037	78005742	Edinson Cavani	1657	14-02-1987	29	76	27	13	9	188	...	10	16	12	14	
87115	51001380	Javier Hernández	379	01-06-1988	28	86	45	5	1	175	...	14	14	13	16	
95564	55002038	Diego Costa	796	07-10-1988	27	10	1	0	0	186	...	16	16	11	15	
64420	37008621	Luuk de Jong	784	27-08-1990	25	12	3	18	5	188	...	12	16	12	10	
60030	34000647	Olivier Giroud	769	30-09-1986	29	55	20	0	0	192	...	15	14	14	15	
64372	37005776	Bas Dost	784	31-05-1989	27	6	1	10	5	192	...	12	17	8	13	
158329	962988	Gonzalo Higuaín	1649	10-12-1987	28	63	30	0	0	184	...	18	16	11	12	
149132	91014711	Pierre-Michel Lasogga	771	15-12-1991	24	0	0	11	4	189	...	12	15	8	16	
1479	12031625	Islam Slimani	5	18-06-1988	28	44	24	0	0	188	...	12	14	14	13	
74642	43010773	Ciro Immobile	776	20-02-1990	26	15	1	16	8	185	...	7	16	10	12	

Observations/Improvements

Observations

- Basing selection off weighted functions is unreliable (GK)
- More attributes = better selection of players

Improve Runtime

- Fagin's (FA) Algorithm (O($N \cdot \sqrt[m]{\frac{k}{N}}$))
- Parallelize data

Future Work

- Filters
- Player Roles
- Regens

Questions?