

UNSUPERVISED MACHINE LEARNING: CLUSTERING PLAYERS IN FIFA 15

Elias Olcina





DATA SET

15465 Rows

104 Columns, 28 numerical

Trial and Error:

- With and without PCA
- With all/subset of variables

Removed Goalkeepers

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 15465 entries, 0 to 15464  
Columns: 104 entries, sofifa_id to rb  
dtypes: float64(17), int64(11), object(76)  
memory usage: 12.3+ MB
```



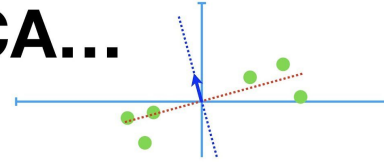
FINAL MODEL: K-MEANS

Settled for a K-Means model, with 6 variables:

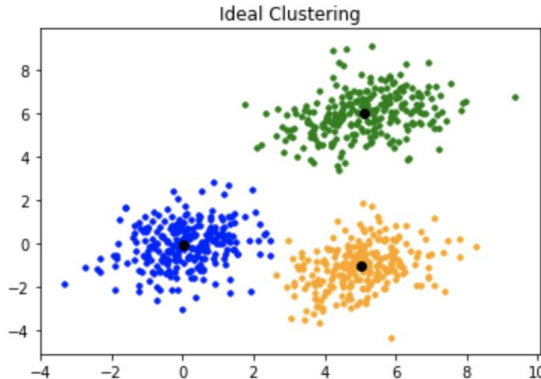
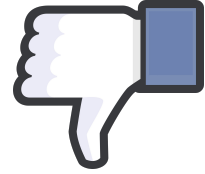
- Pace
- Shooting
- Passing
- Dribbling
- Defending
- Physics

K-Means with 5 clusters using Elbow

PCA...

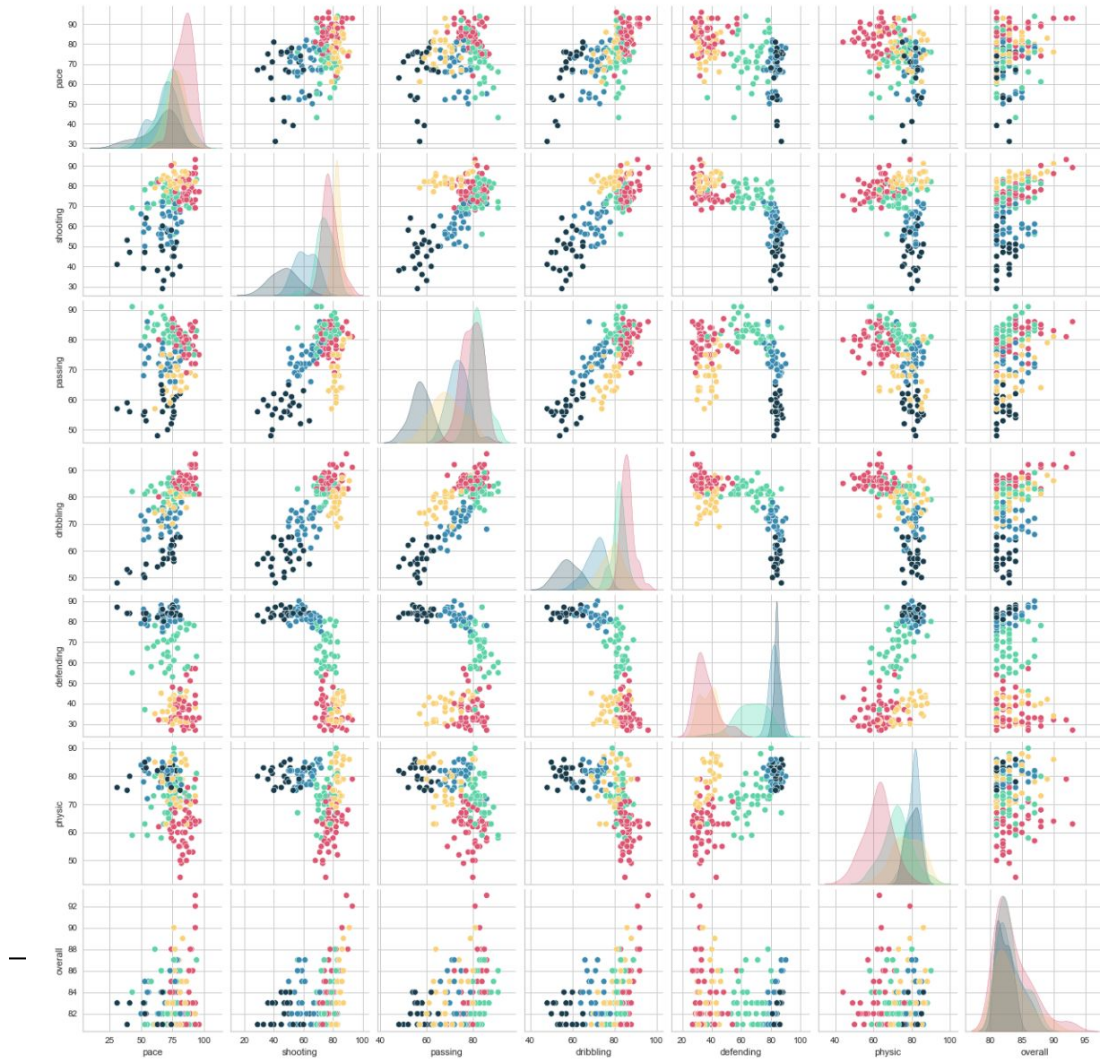


Step-by-Step!!!



VISUALIZE CLUSTERS

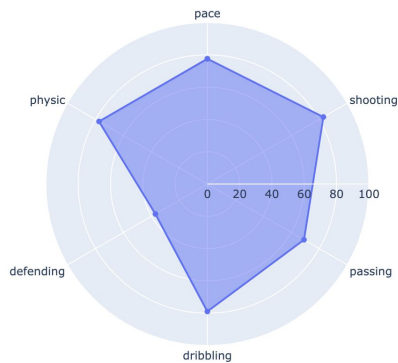
- Finesse Offense
- Goalscorer
- Middle-Man
- Balanced Defender
- Pure Goal Stopper



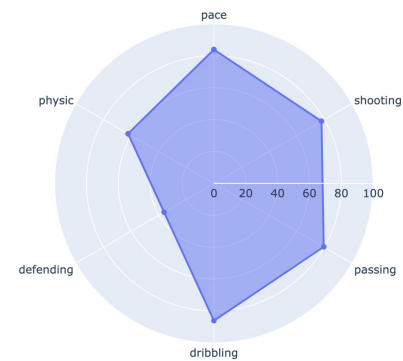
ATTRIBUTES BY CLUSTER



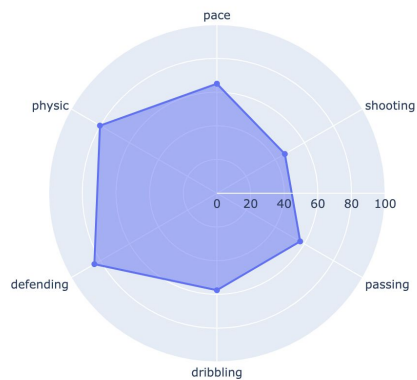
Player Type: Goalscorer



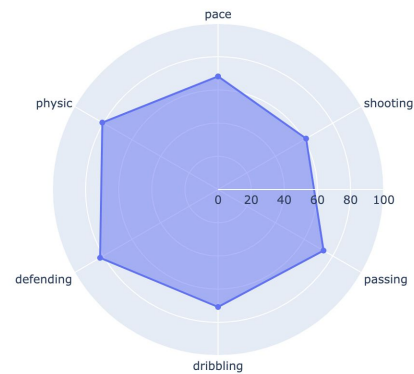
Player Type: Finesse Offense



Player Type: Pure Goal Stopper

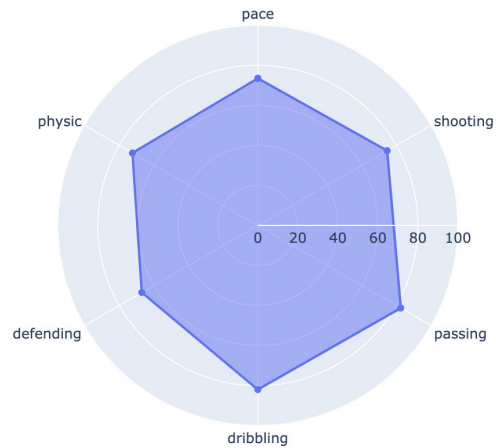


Player Type: Balanced Defender



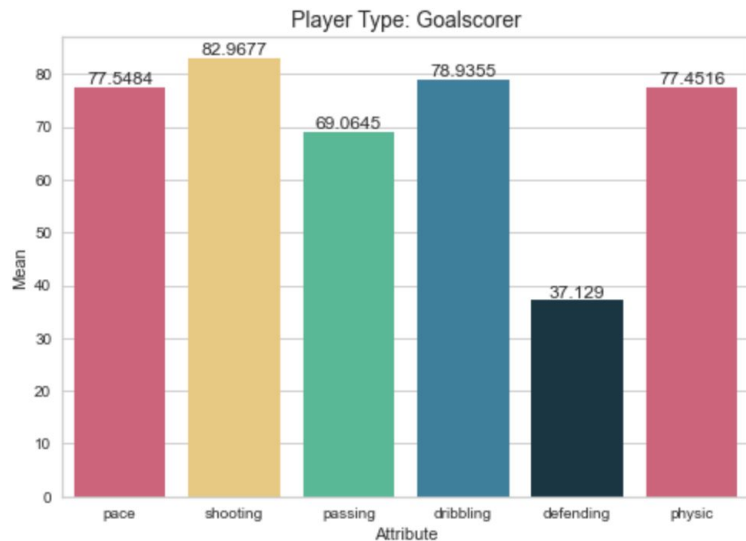
ATTRIBUTES BY CLUSTER

Player Type: Middle-Man

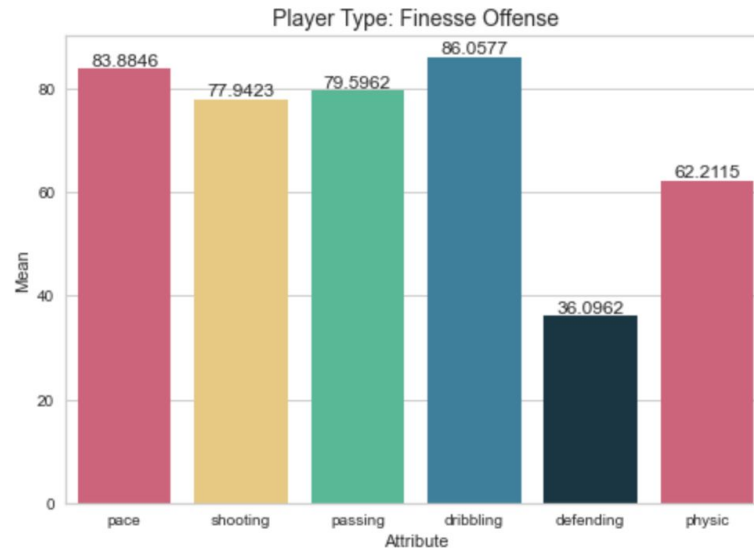




FURTHER COMPARISONS WITH EXAMPLE



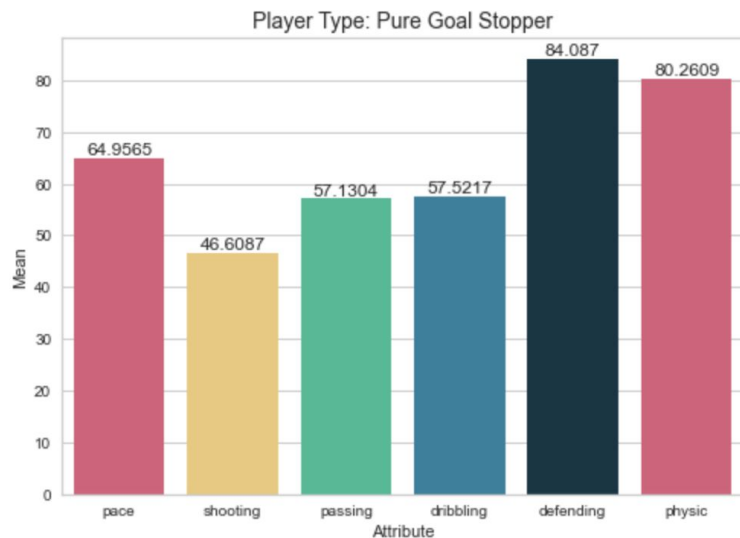
Example: Zlatan Ibrahimovic



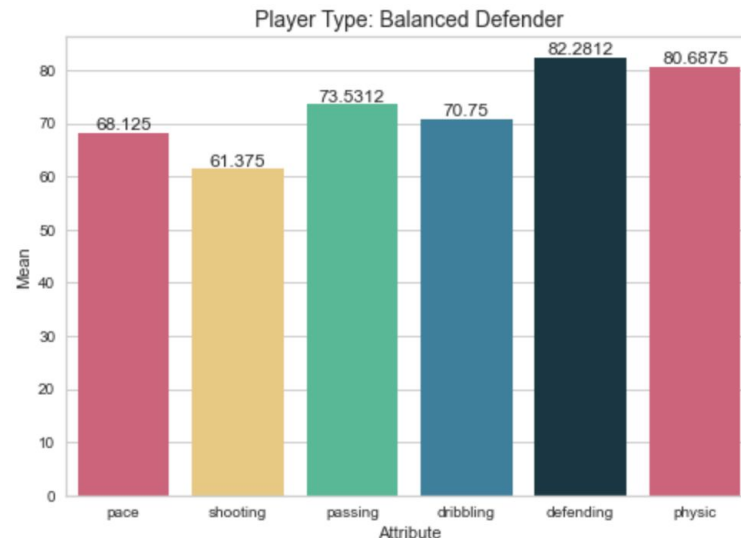
Example: Lionel Messi



COMPARISONS WITH EXAMPLE



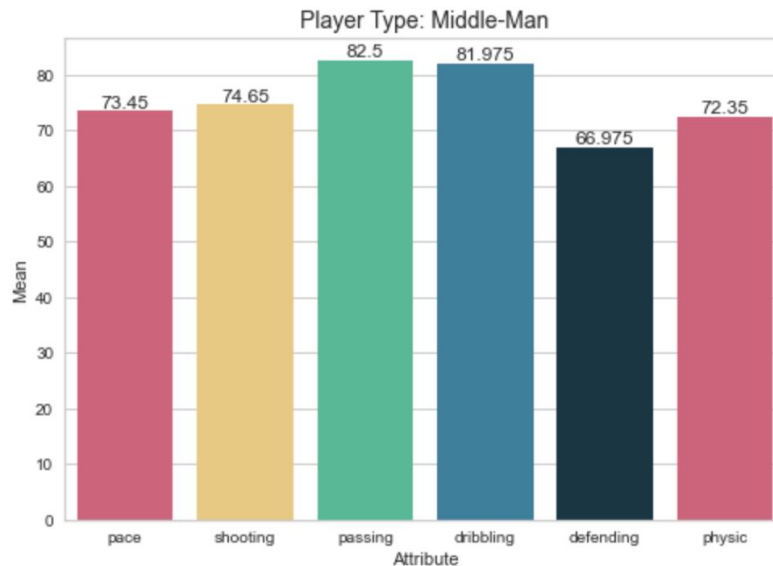
Example: Giorgio Chiellini



Example: Sergio Ramos



COMPARISONS WITH EXAMPLE



Example: Bastian Schweinsteiger



QUESTIONS?
