

Soccer Analytics: Player Similarity Analysis

Team 1: Sean Jung, Stephen Thomas Jr,
Justin Wood, Cassio Salge



Our Research Question is as follows:

We want to find out which players are most similar to these following players that could be a potential business investment for the teams success when these players are retired, injured or transferred.

Darlington Nagbe



Jeff Larentowicz



Miguel Almirón



Eric Remedi



The players **we are analyzing** are as follows:

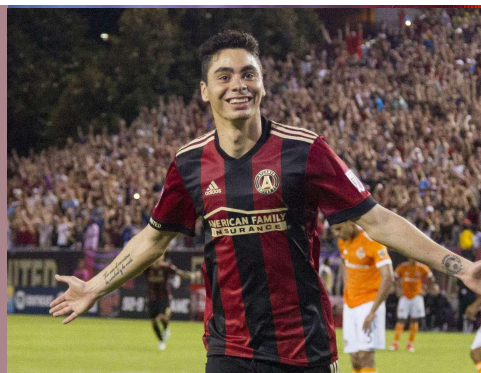


Jeff Larentowicz

AU: 2017 - 2020

Defending Midfielder

Avg 98 mins per goal (2019)

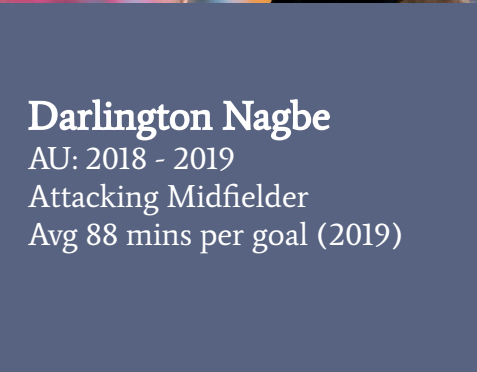


Eric Remedi

AU: 2018 - 2020

Defending Midfielder

Avg 82 mins per goal (2019)



Darlington Nagbe

AU: 2018 - 2019

Attacking Midfielder

Avg 88 mins per goal (2019)



Miguel Almirón

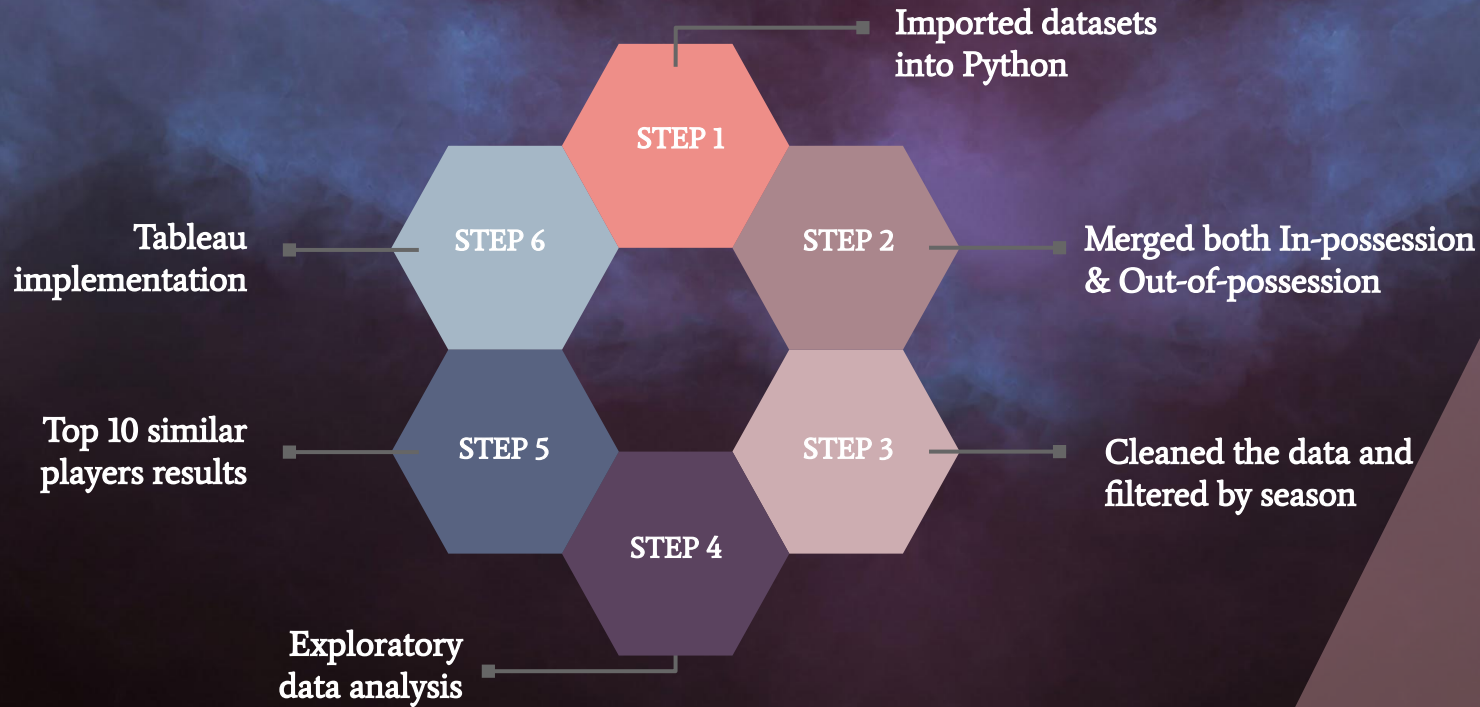
AU: 2017 - 2018

Midfielder

Avg 86 mins per goal (2018)



Our methodology is as follows:



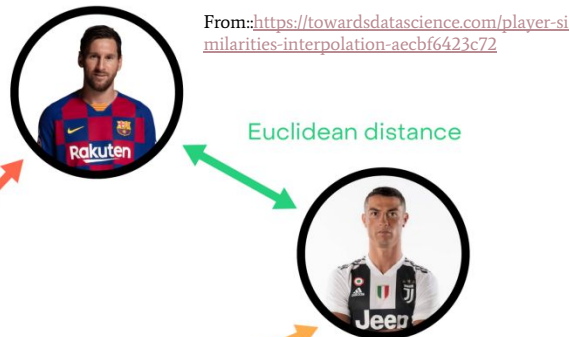
We Considered Different Similarity Measures:

Euclidean Distance

Cosine Similarity

Pearson Correlation

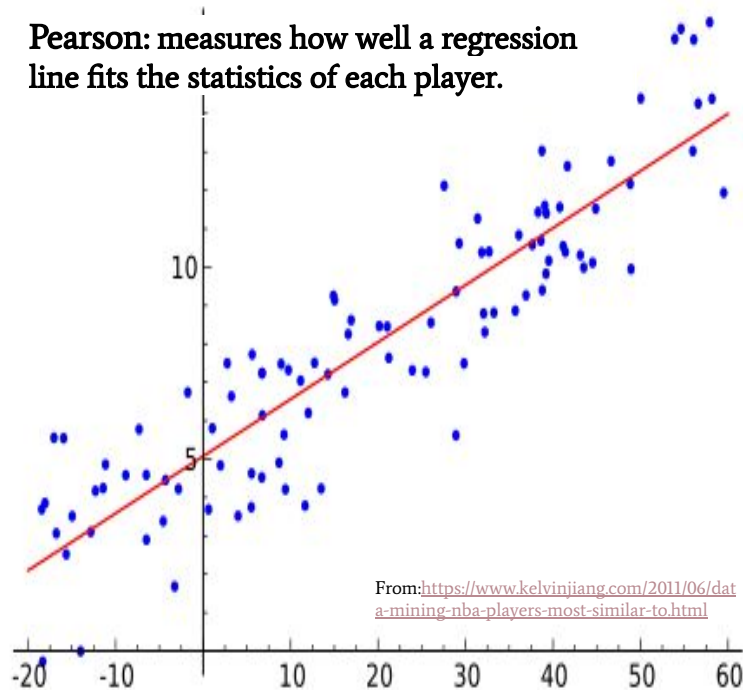
Euclidean: measure the distance between two players.



Cosine: measures how much the ratio of a player's statistics differs from the ratio of another player.

Cosine similarity

Pearson: measures how well a regression line fits the statistics of each player.



Why did we choose cosine similarity?

- ◀ Measures similar player style rather than similar player statistics
- ◀ Similarities made based on skill/ability than performance and popularity
- ◀ Works best when many features are available

Best way to score/rank most similar players to our target players

| | Euclidean Distance | Cosine Similarity | Pearson Correlation |
|------|---|--|---|
| Pros | Easy to implement and intuitive | It allows better capture “style” rather than pure “statistics” attributes. | Accuracy of score increases when data is not normalized. |
| Cons | The variables which have the largest value greatly influence the result | Doesn't work efficiently with nominal data | Accuracy of score increases even when data is not normalized. |

Here are our top 10 similar players we recommend for 2020

Darlington Nagbe



1. Jose Artur de Lima Junior
2. Emerson Hyndman
3. Matheus Rossetto
4. Jackson William Yuell
5. Felipe Martins Campanholi
6. Bryan Josue Acosta Ramos
7. Thomas Liam McNamara
8. Blaise Matuidi
9. Jonathan dos Santos Ramirez
10. Anibal Cesis Godoy

Jeff Larentowicz



1. Jordi Jose Delem
2. Will Trapp
3. Scott Caldwell
4. Brian Anunga
5. Wilfried Aimeric Zahibo
6. Darwin Adeldo Cerén Delgado
7. Caleb Stanko
8. Tanner Tessmann
9. Matt LaGrassa
10. Russell Teibert

Miguel Almirón (2018)



1. Victor Rodríguez Romero
2. João Jimmy Plata Coterá
3. Ezequiel Barco
4. Ilson Pereira Dias Junior
5. Pablo Mauricio
6. Carlos Alberto Vela Garrido
7. Sebastian Marcelo Blanco
8. Borek Dockal
9. Victor Vazquez Solsona
10. Mauro Alberto Díaz

Eric Remedi

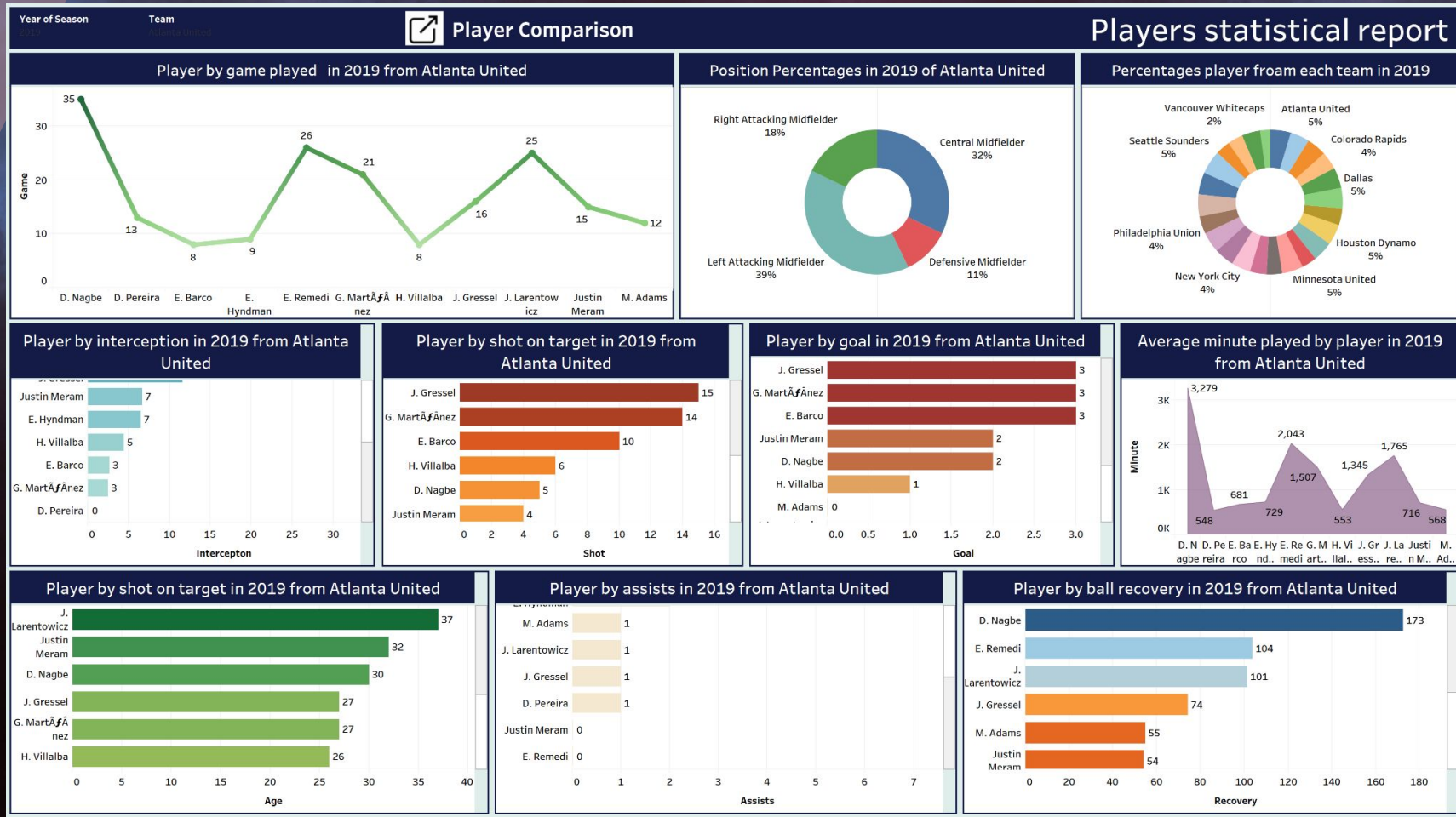


1. Frankie Amaya
2. Victor Wanyama
3. Mohammed Adams
4. Tanner Tessmann
5. Scott Caldwell
6. Matthew Ryan Polster
7. Russell Canouse
8. Matías Gabriel Vera
9. Emanuel Fernando Maciel
10. Dax McCarty

We have used Tableau for Implementation

- Our dynamic visualization can select players and get high level results:
 - MLS official website
 - Player analysis
 - Skill sets and history
- Able to make quick analysis and decisions on potential players through our interactive model
- Quick demonstration on Tableau

We created interactive visualization to compare players



We have also implemented player similarity analysis

Year

2018

Team

(All)


Position

(All)

Main Player


Miguel Ángel Almirón...

Miguel Ángel Almirón Rejala





Player Statistics


| | | | | |
|-----------------------------|--------------------------|----------------|---------------------------|---------------------------------|
| Víctor Rodríguez Romero | João Jimmy Plata Cotera | Ezequiel Barco | Ilson Pereira Dias Júnior | Pablo Mauricio Aránguiz Salazar |
| 0.9981 | 0.9979 | 0.9974 | 0.9974 | 0.9960 |
| Carlos Alberto Vela Garrido | Sebastián Marcelo Blanco | Borek Dockal | Víctor Vázquez Solsona | Mauro Alberto Díaz |
| 0.9957 | 0.9955 | 0.9949 | 0.9947 | 0.9944 |




NewsVideoStandingsScheduleStatsClubsCompetitionsRosters...

🔍 TicketsStoreEspañolAccount

CHANGE YOUR GAME.  



SHOP NOW






Miguel Almiron
Midfielder

ProfileCareer StatsMatch LogNewsVideo

Player Details

| | | | |
|------------|--------------------|---------------|----------------|
| Name | Miguel Almiron | Height | 5' 9" |
| Weight | 141 lbs | Date of Birth | 2.10.1994 (27) |
| Birthplace | Asuncion, Paraguay | Position | Midfielder |

SCORE **FREE** MLS GEAR



There are some limitations in our approach

FEATURE SELECTION

Selecting certain features can
change outcomes in player
similarity

INJURIES

Player injuries could
distort stats



COVID-19

Covid altered certain lengths of
seasons on average were lower
than previous seasons and
player eligibility

A black and white photograph of a crowd of people cheering with their arms raised in a stadium. The image is overlaid with a semi-transparent red banner on the right side, which contains the text "Thank You!".

Thank You!