FIFA data analytics and visualization (December 2021)

Lamiss Gargouri

Master student in Bahcesehir University

Department of Big Data Analytics and Management.

ABSTRACT The act of adding order and organization to acquired data is known as data analysis. It transforms data into actionable information for teams. The process of turning data into a chart, graph, or other visual format that aids analysis and understanding is known as data visualization. Furthermore, data analysis is the act of analyzing, cleaning, manipulating, and modeling data in order to identify usable information, generate conclusions, and help decision-making. So, what is the significance of data visualization in data analysis?By placing data in a visual context, such as maps or graphs, data visualization helps us understand what it means. This makes the data more natural to understand for the human mind, making it simpler to see trends, patterns, and outliers in vast data sets.

INDEX TERMS Data analysis, Data visualization.

# I. INTRODUCTION

FIFA 21 is a FIFA series association football simulation video game produced by Electronic Arts. It is the 28th entry in the FIFA series, and it was launched on Microsoft Windows, Nintendo Switch, PlayStation 4, and Xbox One on October 9, 2020. On December 3, 2020, enhanced versions for the PlayStation 5 and Xbox Series X and Series S were published, as well as a Stadia version in March 2021.

This data visualization will employ statistics information from football players who appear in the game FIFA 21, which has been updated for the 2020-2021 season. We will utilize the dataset provided on Kaggle for our data visualization, which comprises information on 18+ players with 100+ attributes taken from the most recent edition of FIFA.

Using Python, this data visualization will provide a quick exploratory examination of the FIFA 21 dataset.

# II. BACKGROUND

## A. BASIC CONSEPTS

* Big data is a combination of technologies for storing, analyzing, and managing large amounts of data, as well as a macro-tool for seeing patterns in

the chaos of this information explosion in order to construct smart solutions.

* The science of studying raw data in order to draw conclusions about it is known as data analytics.
* Data analytics techniques and procedures have been turned into mechanical processes and algorithms that operate on raw data for human consumption. A company's performance may be improved by using data analytics.
* By visualizing data in the form of maps or graphs, we can have a better understanding of what it implies. This makes the data easier to interpret for the human mind, making it easier to see trends, patterns, and outliers in vast data sets.
* Data visualization software may be utilized in a number of different ways. The most prevalent application nowadays is as a reporting tool for business intelligence (BI).

# Graphical user interface, application, Teams Description automatically generated

# III. Dataset

1. Context

The player data for Career Mode from FIFA 15 to FIFA 21 is included in the databases. The data enables for repeated comparisons of the same gamers over the videogame's last seven versions.

Here are some examples of potential analyses:

* The ideal budget for building a competitive team (at the level of Europe's top n teams) in which the budget does not allow for the purchase of much superior players for the 11-man lineup. An extra is a comparison of the lineup's Potential attribute rather than the Overall attribute.
* Comparison between two players (In comparison to real-life numbers, which skill qualities changed the most over time).
* Sample analysis of the top n percent players to evaluate if some key traits like agility, ball control, and strength have remained popular or not over FIFA editions. For example, in FIFA 20, the top 5% of players are faster (better Acceleration and Agility) than in FIFA 15. The trend of qualities is also a good indicator of how crucial specific attributes are for players to win games.

1. Content

* Every player in FIFA 15, 16, 17, 18, 19, 20, and FIFA 21 is accessible.
* 100+ attributes
* The scraped players' URLs
* Chart, bar chart, histogram

  Description automatically generatedPositions of players, as well as their roles in the club and the national team.
* Attacking, Skills, Defense, Mentality, GK Skills, and so on are all player traits with statistics.
* Nationality, Club, DateOfBirth, Wage, Salary, and other personal information about players are collected.

1. Acknowledgements

The data was collected from a publicly accessible website: [https://sofifa.com](https://sofifa.com/).

# IV. Importing Libraries & Data Exploration

To begin the project, import libraries like NumPy for data visualization and Matplotlib for data visualization, and then load the data using Pd. Read csv and save it to the Fifa object. Data is entered by typing. The  Data.Head() function displays the first five records of the data set, providing a quick overview of the data frame's rows and columns.

A screen shot of a computer

Description automatically generated with low confidence

# V. Analysis and Visualization

Chart, bar chart, histogram

Description automatically generated

## Conclusion: Dribbling is not affected by height.

Chart, bar chart, histogram

Description automatically generated

* **Conclusion: Height does affect movement agibility.**

# Chart, bar chart, histogram Description automatically generated

* **Conclusion: Weight does not affect dribbling.**

**Chart, bar chart, histogram

Description automatically generated**

* **Conclusion: Weight does affect acceleration.**

**Chart, bar chart

Description automatically generated**

# VI. Comparing Stats of Messi, Lewandowski & CR7

To compare these three players I have chosen these attributes:

* Shooting
* Passing
* Dribbling
* Defending
* Physic
* Player traits
* Attacking crossing
* Attacking finishing
* Attacking heading accuracy
* Attacking short passing
* Attacking volleys
* Skill dribbling
* Skill curve
* Skill fk accuracy
* Skill long passing
* Skill ball control
* Movement acceleration
* Movement sprint speed
* Movement agility
* Movement reactions
* Movement balance
* Power shot power
* Power jumping
* Power stamina
* Power strength
* Power long shots

Chart, line chart

Description automatically generated

# VII. Top 10 players based on overall skill

Code:

display(

    HTML(data.sort\_values('overall', ascending=False)[['short\_name', 'overall']][:10].to\_html(index=False)

))

A picture containing text, scoreboard, several

Description automatically generated

# VIII. Age distribution of players in the clubs

# Code :

top\_club\_names = ('FC Barcelona','Real Madrid', 'Juventus', 'Paris Saint-Germain', 'Chelsea', 'Manchester City', 'Manchester United')

clubs = data.loc[data['club\_name'].isin(top\_club\_names) & data['age']]

fig, ax = plt.subplots()

fig.set\_size\_inches(20, 10)

ax = sns.boxenplot(x="club\_name", y="age", data=clubs)

ax.set\_title(label='Age distribution in the top clubs', fontsize=25)

plt.xlabel('Clubs', fontsize=20)

plt.ylabel('Age', fontsize=20)

plt.grid()

Chart, box and whisker chart

Description automatically generated

* Conclusion: we can deduce the strategy of each club through this chart. For instance PSG is clearly the one who is investing the most in its academy while Manchester City is clearlytaking only initiated player from the age of 22.5 to 25 (at their prime).

# IX. Age distribution of players in countries

Code:

countries\_names = ('France', 'Brazil', 'Germany', 'Belgium', 'Spain', 'Netherlands', 'Argentina', 'Portugal', 'Chile', 'Colombia')

countries = data.loc[data['nationality'].isin(countries\_names) & data['age']]

fig, ax = plt.subplots()

fig.set\_size\_inches(20, 10)

ax = sns.boxenplot(x="nationality", y="age", data=countries)

ax.set\_title(label='Age distribution in countries', fontsize=25)

plt.xlabel('countries', fontsize=20)

plt.ylabel('age', fontsize=20)

plt.grid()

Chart, box and whisker chart

Description automatically generated

* Conclusion: As expected, the age distribution between countries is similar.

# X. Conclusion

Using The FIFA 2021 Dataset Data, I Was Able To Explore And Extract The Data Of The Players Concerning His Nationality, Wages, Etc. In addition, the player's various skills, such as defending and shooting, are sorted in order of highest to lowest. We learn about the players' nationalities, income, and skills through this exploratory analysis and visualization, which helps them become world-renowned players.

# REFERENCES

Sandeepanmukherjee. (2021, March 15). *FIFA 21 data analysis and visualisation*. Kaggle. Retrieved

Leone, S. (2020, October 8). *FIFA 21 complete player dataset*. Kaggle. Retrieved December 28, 2021, from https://www.kaggle.com/stefanoleone992/fifa-21-complete-player-dataset

Corporativa, I. (n.d.). *Big data: Main uses and applications*. Iberdrola. Retrieved December 28, 2021, from <https://www.iberdrola.com/innovation/what-is-big-data-what-is-it-used-for>

Borjigin, K. (n.d.). *Players*. SoFIFA. Retrieved December 28, 2021, from https://sofifa.com/

Frankenfield, J. (2021, December 7). *What is data analytics?* Investopedia. Retrieved December 28, 2021, from <https://www.investopedia.com/terms/d/data-analytics.asp>

*Why data visualization is important*. Analytiks. (2021, September 11). Retrieved December 28, 2021, from https://analytiks.co/importance-of-data-visualization/