

**COLLEGE OF COMPUTING AND INFORMATION SCIENCES**

**SCHOOL OF COMPUTING AND INFORMATION SCIENCES**

**BACHELOR OF SCIENCE IN SOFTWARE ENGINEERING**

**RECESS GROUP DATA SCIENCE MINI PROJECT**

|  |  |  |
| --- | --- | --- |
| **NAME** | STUDENT NUMBER | **REGISTRATION NUMBER** |
| MUSINGUZI JOSEPH | **217015788** | 17/U/6671/PS |
| MUKWATSE COLLIN | **217012303** | 17/U/6509/PS |
| KALANZI TONNY | **217015739** | 17/U/4463/EVE |
| BETARIZA ALVIN | **217012330** | 17/U/3574/PS |

**FIFA PLAYER ANALYSIS**

The project is for people who love data science and have grown up playing football and are FIFA enthusiasts.

**THE DATASET**

The data was downloaded from <https://www.kaggle.com/karangadiya/fifa19>.

The dataset contains data from the EA Sports' game FIFA and gets updated regularly with the release of new versions of the game. The data is developed by Electronic Arts for the latest edition of their FIFA game franchise. Through several research projects done on soccer analytics, it has been established in the field of academia that the use of data from the FIFA franchise has several merits that traditional datasets based on historical data do not offer. Since 1995 the FIFA Soccer games provide an extensive and coherent scout of players worldwide.

The FIFA 19 dataset that has been used for this analysis provides statistics of about 18000 players on over 70 different attributes. These attributes are optimal indicators to determine the performance of a player at a particular playing position.

For each attribute, we have an integer from 0-100 that measures how good a player is at that attribute. Examples of the attributes include dribbling, aggression, vision, marking and ball control.

Observe that it seems to be unfeasible to accurately characterize players in these attributes automatically. Thus all of those are gathered and curated by the company whose job is to bring the gameplay closer to reality as possible, hence preserving coherence and representativeness across the dataset.

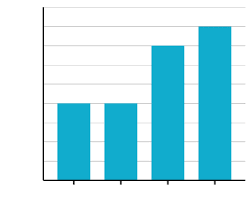
### Contents of the dataset

* Every player featuring in FIFA 19
* 70+ attributes
* Player and Flag Images
* Playing Position Data
* Attributes based on actual data of the latest EA's FIFA 19 game
* Player personal data like Nationality, Photo, Club, Age etc.

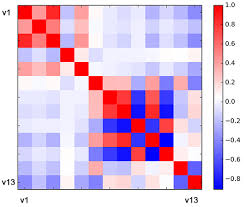
**ANALYTICS APPROACH**

**Exploratory Data Analysis to be done**:

* Show the variation of ages and nationality of the players in the game: - Bar graphs (*chart that uses****bars****to show comparisons between categories of data*) shall be plotted to display the count of players falling under a given nationality and age group.



* Predict the best player to play a given position during the game: - A player to be placed in a given position shall be determined based on his attributes. The player with the best attributes to play a given position shall be selected.
* Determine the correlation amongst certain player attributes: - We shall use a correlation matrix to determine which factors are more dependent on each other using the player skill attributes provided in the dataset.



**Project Lifecycle**

DATA PREPARATION & CLEANING

DATA ACQUISITION

HYPOTHESIS

&

MODELLING

DOCUMENTATION

IMPLEMENTATION