**PRCP-1004-Fifa20**

**Problem Statement**

Task 1:-Prepare a complete data analysis report on the given data.

Task 2:- Explore football skills and cluster football players based on their attributes.

Task3:- Explore the data and attempt all the below asked questions in a

step by step manner:

● Prepare a rank ordered list of top 10 countries with most players. Which countries are producing the most footballers that play at this level?

● Plot the distribution of overall rating vs. age of players. Interpret what is the age after which a player stops improving?

● Which type of offensive players tends to get paid the most: the striker, the right-winger, or the left-winger?

**Dataset Link:**

FIFA 20 Football is arguably the most popular sport in the world and FIFA is the most popular football (soccer) simulation game by Electronic Arts (EA Sports).

The dataset provided includes the players data for the Career Mode from FIFA 15 to FIFA 20 ("players\_20.csv"). The data allows multiple comparisons of the same players across the last 6 versions of the videogame.

Some ideas of possible analysis:

* Historical comparison between Messi and Ronaldo (what skill attributes changed the most during time - compared to real-life stats);
* Ideal budget to create a competitive team (at the level of top n teams in Europe) and at which point the budget does not allow to buy significantly better players for the 11-men lineup. An extra is the same comparison with the Potential attribute for the lineup instead of the Overall attribute;
* Sample analysis of top n% players (e.g. top 5% of the player) to see if some important attributes such as Agility or BallControl or Strength have been popular or not across the FIFA versions. An example would be seeing that the top 5% players of FIFA 20 are faster (higher Acceleration and Agility) compared to FIFA 15. The trend of attributes is also an important indication of how some attributes are necessary for players to win games (a version with more top 5% players with high BallControl stats would indicate that the game is more focused on the technique rather than the physical aspect).

**Domain:** Sports

Link : <https://d3ilbtxij3aepc.cloudfront.net/projects/CDS-Capstone-Projects/PRCP-1004-Fifa20.zip>

The attributes used in the project are:

* Name: Name of the player.
* Age: Age of the player.
* Height: Height of the player in inches (transformed to centimeters in preprocessing).
* Overall: General performance quality and value of the player representing the key positional skills and international reputation rated between 1-99. Overall attribute is used only in preprocessing and discussion stages because using it in modelling could lead to domination by this feature. The aim of the project is not basically sort and categorize the players using their overall talent and international reputation, but to cluster them based on using their whole skillset.
* Potential: Maximum Overall rating expected to be reached by a player in the top of his career rated between 1-99.
* PreferredFoot: Right or Left. Label encoder is applied as 0 for left and 1 for right.
* WeakFoot: Represents how well a player uses his weak foot (e.g. left for righties) rated between 1 to 5.
* WorkRate: Degree of the effort the player puts in terms of attack and defense rated as low, medium and high. This feature is divided into two new features as AttackWorkRate and DefenseWorkRate. Besides, label encoder is applied as 0 for low, 0.5 for medium and 1 for high.
* Position: Position of the players on the pitch which determines their roles and responsibilities in the team. Forward positions in the football and FIFA 19 can be grouped as striker (ST: center striker, RS: right striker, LS: left striker), forward (CF: center forward, RF: right forward, LF: left forward) and winger (RW: right winger, LW: left winger). The word, forward, is used both as a general term and a special position. Strikers are positioned in front of forwards and wingers and very closed to the opposing goal. Their main responsibilities are attacking and scoring goals, that’s why their ball control, shooting and finishing skills are expected to be well. Center forwards are positioned right behind the strikers. They are expected to receive balls from the others and score assists to the others or goals. In addition to the skills expected from strikers, they have to be good at passing. Right forwards and left forwards are positioned at the right and left of the center forwards with the same expectations. Wingers are positioned near the touchlines to create chances for strikers and forwards from the right and left side of the field by breakthrough and crosses and to score goals. They are expected to be good at dribbling, acceleration, passing and crossing. Positions are used only in preprocessing and discussion stages.
* ST: Positional skill. Player’s general ability when playing in ST position rated between 1-99.
* RS: Positional skill. Player’s general ability when playing in in RS position rated between 1-99.
* LS: Positional skill. Player’s general ability when playing in in LS position rated between 1-99.
* CF: Positional skill. Player’s general ability when playing in in CF position rated between 1-99.
* RF: Positional skill. Player’s general ability when playing in in RF position rated between 1-99.
* LF: Positional skill. Player’s general ability when playing in in LF position rated between 1-99.
* RW: Positional skill. Player’s general ability when playing in in RW position rated between 1-99.
* LW: Positional skill. Player’s general ability when playing in in LW position rated between 1-99.
* Crossing: Crossing skill of the player rated between 1-99. Cross is a long-range pass from wings to center.
* Finishing: Finishing skill of the player rated between 1-99. Finishing in football refers to finish an attack by scoring a goal.
* HeadingAccuracy: Player’s accuracy to pass or shoot by using his head rated between 1-99.
* ShortPassing: Player’s accuracy for short passes rated between 1-99.
* LongPassing: Player’s accuracy for long passes rated between 1-99.
* Dribbling: Dribbling skill of the player rated between 1-99. Dribbling is carrying the ball without losing while moving in one particular direction.
* SprintSpeed: Speed rate of the player rated between 1-99.
* Acceleration: Shows how fast a player can reach his maximum sprint speed rated between 1-99.
* FKAccuracy: Player’s accuracy to score free kick goals rated between 1-99.
* BallControl: Player’s ability to control the ball rated between 1-99.
* Balance: Player’s ability to remain steady while running, carrying and controlling the ball rated between 1-99.
* ShotPower: Player’s strength level of shooting the ball rated between 1-99.
* Jumping: Player’s jumping skill rated between 1-99.
* Penalties: Player’s accuracy to score goals from penalty rated between 1-99.
* Strength: Physical strength of the player rated between 1-99.
* Agility: Gracefulness and quickness of the player while controlling the ball rated between 1-99.
* Reactions: Acting speed of the player to what happens in his environment rated between 1-99.
* Aggression: Aggression level of the player while pushing, pulling and tackling rated between 1-99.
* Positioning: Player’s ability to place himself in the right position to receive the ball or score goals rated between 1-99.
* Vision: Player’s mental awareness about the other players in the team for passing rated between 1-99.
* Volleys: Player’s ability to perform volleys rated between 1-99.
* LongShots: Player’s accuracy of shoots from long distances rated between 1-99.
* Stamina: Player’s ability to sustain his stamina level during the match rated between 1-99. Players with lower stamina get tired fast.
* Composure: Player’s ability to control his calmness and frustration during the match rated between 1-99.
* Curve: Player’s ability to curve the ball while passing or shooting rated between 1-99.
* Interceptions: Player’s ability to intercept the ball while opposite team’s players are passing rated between 1-99. It is a defensive skill.
* StandingTackle: Player’s ability to perform tackle (take the ball from the opposite player) while standing rated between 1-99. It is a defensive skill.
* SlidingTackle: Player’s ability to perform tackle by sliding rated between 1-99. It is a defensive skill.
* Marking: Player’s ability to apply strategies to prevent opposing team from taking the ball rated between 1-99. It is a defensive skill.

**Model Comparison Report**

Create a report stating the performance of multiple models on this data and suggest the best model for production.

**Report on Challenges faced**

Create a report which should include challenges you faced on data and what technique used with proper reason.

Note:-All above tasks have to be done on a single jupyter notebook and share the same for the final submission of the project.