**Business Requirements:**

**1. Project Overview**

* **Project Name**: FIFA 21 Player Performance Analysis
* **Objective**: To analyze the FIFA 21 player data to uncover key trends, insights, and potential opportunities related to player performance, market value, and attributes across different positions and age groups.

**2. Business Requirements**

**2.1 Key Analysis Goals**:

* **Player Performance Insights**: Provide insights on player performance metrics such as overall ratings, potential, and attributes.
* **Market Value Analysis**: Understand trends in player market values, especially in relation to positions and age groups.
* **Potential Growth**: Identify players with the highest growth potential based on the difference between their current overall ratings and future potential.
* **Age and Position-Based Analysis**: Explore how age and player positions impact overall performance, potential, and market value.

**2.2 Deliverables**:

* Clear insights on player performance distribution across various attributes (e.g., overall, potential, age, market value).
* Visualizations of key performance metrics such as player value trends, position-based performance, and correlation between player attributes.
* List of high-potential players with low current performance (valuable for recruitment and scouting).
* A final report summarizing all findings, trends, and actionable insights.

**Task 1: Initial Data Exploration & Cleaning** ✓

**Objective**: Prepare the dataset for analysis by cleaning and processing the raw data.

**Actions**:

1. Check for missing values across all columns.
2. Identify columns that are not relevant for our analysis and drop them (e.g., sofifa\_id, dob, player\_url).
3. Fix any issues with player names (e.g., different languages, special characters).
4. Ensure that all columns have the correct data types (e.g., value\_eur, wage\_eur as numeric).
5. Address any inconsistencies with player positions and other categorical columns.

**Task 2: Handling Missing Values & Imputation** ✓

**Objective**: Handle missing data appropriately to avoid bias or inaccuracies in the analysis.

**Actions**:

1. Identify columns with missing data (e.g., nation\_position, release\_clause\_eur, goalkeeper stats).
2. Decide how to handle missing values:
   * Drop irrelevant columns like player\_tags, loaned\_from, nation\_position.
   * Impute missing values for relevant columns:
     + Goalkeeper stats should be filled with 0 for non-goalkeepers.
     + Free agents' missing club and league details should be filled with "Free Agent".
     + For missing numeric values (e.g., pace, shooting for goalkeepers), impute 0 for non-applicable players.
3. Drop columns that are not necessary for the analysis (e.g., contract\_valid\_until, release\_clause\_eur).

**Task 3: Analyzing Player Distributions and Summary Statistics** ✓

**Objective**: Understand the distribution of key metrics such as overall, potential, age, value\_eur, and wage\_eur.

**Actions**:

1. Use .describe() to get summary statistics for key metrics.
2. Create histograms or box plots for overall, potential, age, value\_eur, and wage\_eur to explore their distributions.
3. Analyze the distributions:
   * Identify if the distributions are normal, skewed, or have outliers.
   * Provide commentary on the central tendencies, spread, and any noticeable trends (e.g., age distribution skewed toward younger players).

**Task 4: Advanced Insights and Trend Analysis**

**Objective**: Perform a deeper analysis to uncover key trends and insights regarding player growth, performance across age groups, and market value.

**Actions**:

1. **Player Growth Potential**:
   * Create a new column called potential\_growth by calculating the difference between potential and overall.
   * Identify the top 10 players with the highest potential growth and analyze their positions and attributes.
2. **Age vs. Performance**:
   * Divide players into age groups (e.g., 18-22, 23-26, 27-30, 31+).
   * Calculate the average overall, potential, and value\_eur for each age group.
   * Visualize the performance and value of players across these age groups.
3. **Market Value Trend by Position**:
   * Group players by position and calculate the average value\_eur for each group.
   * Create bar or line plots to show how player market value changes by position (e.g., attackers, midfielders, defenders, goalkeepers).
4. **High-Value, Low-Overall Players**:
   * Identify players with high market value but low overall ratings (e.g., overall < 70 but value\_eur > 5 million).
   * Analyze their traits and what contributes to their high value despite lower overall performance.

**Task 5: Feature Engineering**

**Objective**: Create new features from the existing dataset to derive additional insights.

**Actions**:

1. **Performance per Value**:
   * Create a feature that represents the player's performance per value (e.g., overall divided by value\_eur).
   * Analyze the most cost-efficient players based on this new metric.
2. **Future Potential Growth**:
   * Create a feature comparing potential and overall to identify players with the largest room for growth.
   * Analyze players who are under-performing but have high future potential.
3. **Position-Specific Analysis**:
   * Create position-specific features to understand how different positions impact performance and value.
   * For example, assess which attributes (e.g., pace, passing) contribute the most to a player’s overall performance depending on their position (e.g., forwards, defenders).

**Task 6: Visual Storytelling and Final Summary**

**Objective**: Create a final report that summarizes all the findings and insights from the analysis and present it in a visually compelling way.

**Actions**:

1. **Create a Summary of Key Findings**:
   * Summarize the most important insights from the previous tasks (e.g., trends in player performance, market value, potential growth).
2. **Generate a Visual Dashboard**:
   * Create a dashboard with key visualizations (e.g., distribution of player metrics, top players, age vs. performance trends).
   * Ensure that the visualizations clearly communicate the findings to stakeholders (e.g., recruitment teams, analysts).
3. **Highlight Key Player Groups**:
   * Emphasize any standout players or groups (e.g., top performers, players with high growth potential).
   * Provide actionable insights that stakeholders can use for player scouting or strategic decisions.