**Scenario 1: Data Cleaning and Preparation**

**Objective:** Clean and prepare the dataset for analysis.

**Steps:**

1. Load the dataset.
2. Identify and handle missing values.
3. Convert data types to appropriate formats.
4. Remove duplicates.
5. Standardize units for physical attributes like height and weight.

**Tasks:**

* Load the FIFA 21 dataset into a DataFrame.
* Check for missing values and decide on a strategy to handle them.
* Ensure columns have appropriate data types.
* Remove any duplicate rows.
* Standardize the height (convert to centimeters) and weight (convert to kilograms).

**Scenario 2: Descriptive Statistics and Basic Visualization**

**Objective:** Perform descriptive statistics and create basic visualizations to understand the dataset.

**Steps:**

1. Calculate summary statistics for numerical columns.
2. Visualize the distribution of player ratings.
3. Analyze the frequency of categorical attributes (e.g., Preferred Foot, Position).
4. Identify top clubs and nationalities based on average player ratings.

**Tasks:**

* Calculate mean, median, and standard deviation for key attributes like Overall, Potential, Value, and Wage.
* Create histograms for the Overall and Potential ratings.
* Create bar plots for Preferred Foot and Position.
* Identify and visualize the top 10 clubs and nationalities based on average Overall rating.

**Scenario 3: Correlation and Feature Engineering**

**Objective:** Analyze correlations and create new features for deeper insights.

**Steps:**

1. Calculate and visualize the correlation matrix.
2. Identify highly correlated attributes.
3. Create new features (e.g., Total Physical Score, Skill Index).
4. Analyze the impact of these new features on player performance.

**Tasks:**

* Calculate the correlation matrix and create a heatmap.
* Identify pairs of attributes with high correlation.
* Create a Total Physical Score as the sum of physical attributes (e.g., Strength, Stamina).
* Create a Skill Index combining key skill attributes (e.g., Dribbling, Ball Control).
* Analyze how Total Physical Score and Skill Index relate to Overall rating.

**Scenario 4: Grouping and Aggregation**

**Objective:** Group and aggregate data to gain insights on different player segments.

**Steps:**

1. Group players by Club and calculate average ratings and values.
2. Analyze player attributes based on positions.
3. Compare average ratings for different age groups.
4. Investigate the relationship between wage and value for top players.

**Tasks:**

* Group players by Club and calculate the average Overall, Potential, and Value.
* Compare the average Overall rating for different positions (e.g., Forwards, Midfielders, Defenders, Goalkeepers).
* Create age groups (e.g., <20, 20-25, 26-30, 31+) and compare average ratings for each group.
* Visualize the relationship between Wage and Value for the top 100 players by Overall rating.

**Scenario 5: Predictive Analysis and Modeling**

**Objective:** Use machine learning to predict player ratings based on attributes.

**Steps:**

1. Prepare the dataset for modeling (feature selection, encoding, splitting).
2. Choose a machine learning algorithm (e.g., Linear Regression, Decision Tree).
3. Train and evaluate the model.
4. Interpret the model results and identify important features.

**Tasks:**

* Select relevant features for predicting Overall rating.
* Encode categorical variables (e.g., Position, Preferred Foot).
* Split the dataset into training and testing sets.
* Train a regression model to predict Overall rating based on player attributes.
* Evaluate the model performance using metrics like R-squared and Mean Absolute Error (MAE).
* Identify the most important features influencing the Overall rating.