

Task 1

Summary of Data Exploration

Dataset 1: Player Statistics (all_seasons.csv)

- **Shape:** 12,844 rows and 22 columns.
- **Columns:** Includes details like player_id, player_name, team_abbreviation, age, player_height, player_weight, college, country, and performance metrics (pts, reb, ast, etc.).
- **Data Types:**
 - Numeric: Columns like age, player_height, player_weight, and performance metrics.
 - Categorical: Columns such as player_name, team_abbreviation, college, country.
- **Missing Values:** No missing values in any column.
- **Summary Statistics:**
 - Average player age is 27 years.
 - Most players are from the USA (10,721 entries).
 - Average points (PTS) per game: 8.21.

Dataset 2: Bank Marketing Data (bank.csv)

- **Shape:** 4,521 rows and 17 columns.
 - **Columns:** Includes features like:
 - **Demographics:** age, job, marital, education.
 - **Financial Information:** default, balance, housing, loan.
 - **Campaign Details:** contact, month, day, duration, campaign, pdays, previous, poutcome.
 - **Target Variable:** y (indicating whether a client subscribed to a term deposit).
 - **Data Types:**
 - Numeric: age, balance, day, duration, campaign, pdays, previous.
 - Categorical: job, marital, education, default, housing, loan, contact, month, poutcome, y.
 - **Missing Values:** No missing values in any column.
 - **Summary Statistics:**
 - Average client age: ~41 years.
 - Most common job: management.
 - Average balance: €1,422.66.
 - Campaign success rate: 11.5% (based on target variable y)
-