**VOLLEY BALL PLAYS**

**(**USING OF PYTHON CREATED EDA)

* I've started to collecting and sorting a dataset about volleyball players in Volleyball National League .
* the recourse I've used is [volleyballworld.com](https://www.kaggle.com/datasets/yeganehbavafa/url)
* In each factor, I have used the average of the player during a game, which was the author of errors, points and percentage of positive performance of that player in each section.
* **Here are some information about the Dataset:**
* Position Feature represent the real position of each player

1. OH : outside hitter
2. OP : Opposite hitter
3. MB : Middle blocker
4. S : Setter
5. L : Libero

* Attack : A player's overall average during each game in the offense factor
* Block : A player's overall average during each game in the defense factor.  
  Defense on the net includes direct points, errors and touching the ball without changing points.
* Serve : Except for the libero player, the rest of the players serve during each turn to start the game. Each player's service average during the match is listed here.
* Set : The setter is responsible for setting the players. But in special cases, the rest of the players cooperate in this matter, and in the data analysis, we will find out which position the players have the most participation in setting after the setter.
* Set feature represent the average of successful sets, errors and attempts for each ball during a rally.
* Dig : the average of digs, errors and receptions
* Receive : Receive feature also represents the average of successful receptions, errors and attempts per mach.

Here is the basic diagram of players in court  
A screen shot of a computer screen

AI-generated content may be incorrect.

I hope you find it useful for analyzing.

**EDA ANALYST**

* Load the CSV,
* Show the first few rows,
* Summarize structure (columns, datatypes, missing values).

Here’s what I found :

* Dataset: **131 players × 10 columns**
* Columns:
  + Player → Player name
  + Country → National team
  + Age → Player age
  + Attack, Block, Serve, Set, Dig, Receive → Performance stats
  + Position → Playing role (e.g., OH = Outside Hitter, OP = Opposite, etc.)
* No missing values — clean dataset.
* Numeric stats (float) are available for performance comparison.

**What I did (quick bullet)**

* Loaded and checked the data (131 rows × 10 columns; no missing values).
* Produced descriptive statistics for each numeric metric.
* Listed top-10 players for each metric (Attack, Block, Serve, Set, Dig, Receive).
* Calculated country-level averages and position-level stats.
* Produced boxplots for each metric and scatter plots (Attack vs Receive, Attack vs Dig).
* Computed and plotted the correlation matrix for the six metrics.

**Quick findings (high-level)**

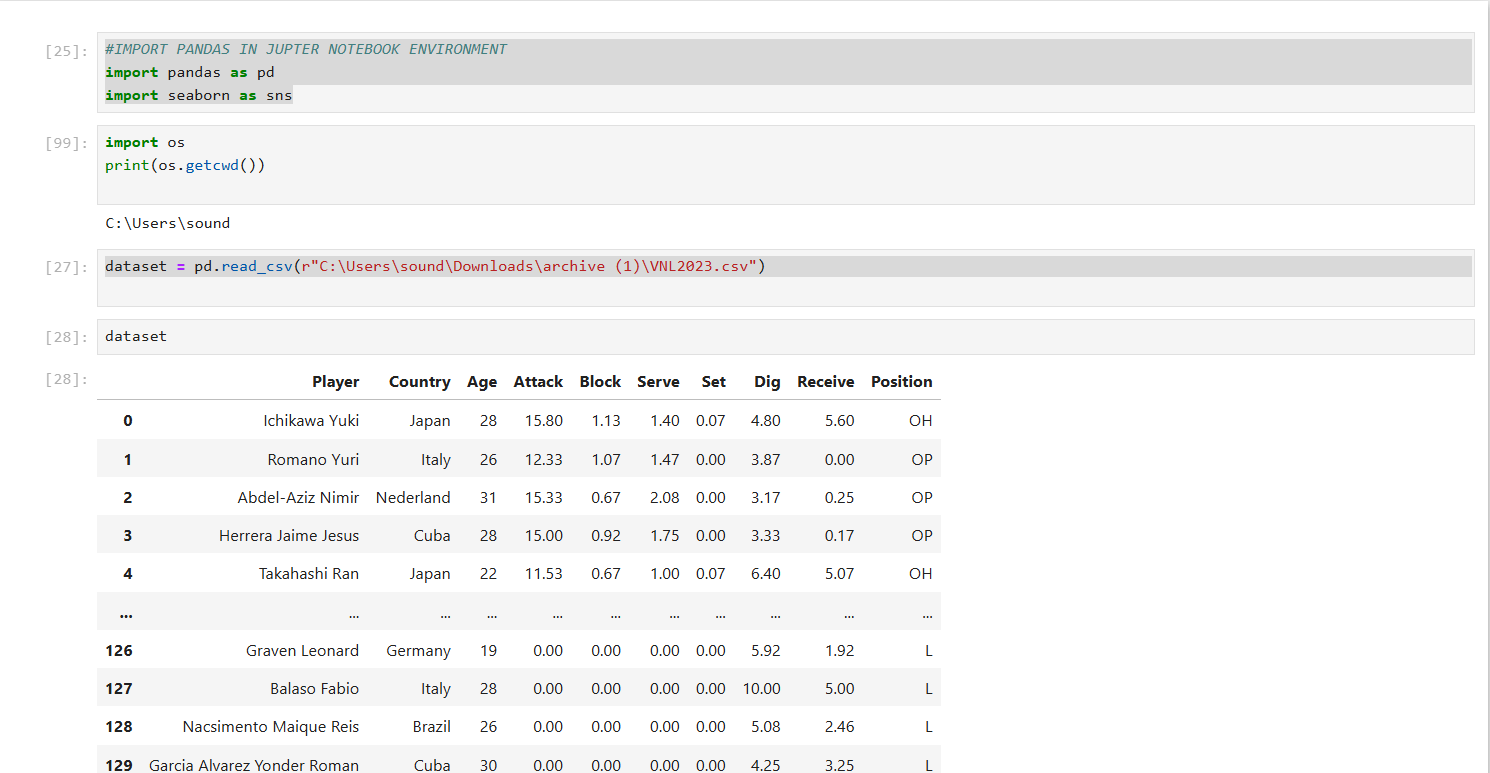
* Dataset clean — no nulls.
* Attack, Serve, and Dig show wider spread (some high outliers).
* Correlations:
  + **Attack** correlates positively with **Serve** and **Block**.
  + **Dig** and **Receive** show a positive relationship.
  + **Set** tends to be negatively correlated with Attack in this data (likely due to role differences — setters have high set counts but fewer attacks).
* Position counts and averages are available so you can compare roles (OH, OP, MB, S, L, etc.).

**Files & downloads**

You can download the outputs here from the workspace:

* Excel full report (sheets with describe / country averages / position stats / top-10s / correlation):  
  sandbox:/mnt/data/volleyball\_eda\_outputs/volleyball\_eda\_report.xlsx
* CSV summary: sandbox:/mnt/data/volleyball\_eda\_outputs/volleyball\_eda\_summary.csv
* Text summary: sandbox:/mnt/data/volleyball\_eda\_outputs/summary.txt
* Plots (boxplots, correlation heatmap, scatter plots) saved in /mnt/data/volleyball\_eda\_outputs (each metric has a box\_<metric>.png, plus correlation\_heatmap.png, scatter\_attack\_receive.png, scatter\_attack\_dig.png).

(Links above are workspace download links — click to download.)



A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

A screenshot of a graph

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A graph of a number of people

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A diagram of a scatter plot

AI-generated content may be incorrect.

A screen shot of a graph

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer screen

AI-generated content may be incorrect.

A screen shot of a graph

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.S