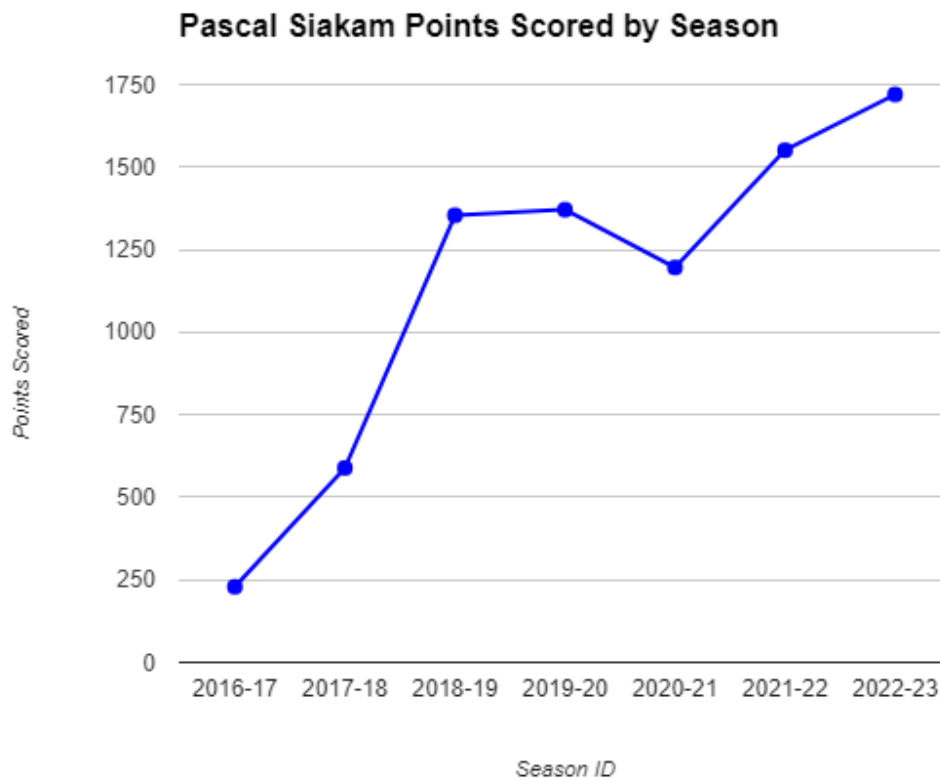


Playing with Columns (Teacher Guide)

This guide is intended to accompany the [01-01-columns](#) demo and exercise. The purpose of the lesson is to learn how to import a "Comma Separated Values" (CSV) file of Pascal Siakam career statistics and perform some first looks at the data.

Eventually this information will get used to create a plot similar to this:



Contents

Introduction.....	3
Data Fields	4
Let's Get Our Data	6
Code Explanation.....	6
The Raw Data	6
Looking at the Columns.....	8
Exercise.....	9

Introduction

The world of data science relies heavily on tools like Python for its flexibility and power. In particular, the Pandas library is essential for working with data in easy-to-understand structures.

Data Fields

To understand the data we're working with, here's a breakdown of what each column represents:

Field Name	Definition	Field Name	Definition
AST	The total number of assists a player has made.	FTM	The total number of free throws the player has made.
BLK	The total number of opponent shots a player has deflected or prevented.	GP	The number of games in which the player has appeared.
DREB	The total number of rebounds a player has grabbed on the defensive end.	GS	The number of games in which the player was in the starting lineup.
FG_PCT	The percentage of field goal attempts that are successful.	MIN	The total number of minutes the player has played.
FG2_PCT	The percentage of two-point field goal attempts that are successful.	OREB	The total number of rebounds a player has grabbed on the offensive end.
FG2A	The total number of two-point field goal attempts by the player.	PF	The total number of personal fouls committed by the player.
FG2M	The total number of two-point field goals a player has made.	PLAYER_AGE	The age of the player.
FG3_PCT	The percentage of three-point field goal attempts that are successful.	PTS	The total number of points a player has scored.
FG3A	The total number of three-point field goal attempts by the player.	REB	The total number of rebounds (offensive + defensive) a player has collected.

Field Name	Definition	Field Name	Definition
FG3M	The total number of three-point field goals a player has made.	SEASON_ID	The identifier for the basketball season.
FGA	The total number of field goal attempts by the player.	STL	The total number of times a player has successfully taken the ball away from an opponent.
FGM	The total number of field goals a player has made.	TEAM_ABBREVIATION	The abbreviated name of the team.
FT_PCT	The percentage of free throw attempts that are successful.	TEAM_ID	A unique identifier for the team.
FTA	The total number of free throw attempts by the player.	TOV	The total number of times a player loses possession of the ball.

Let's Get Our Data

In this lesson, we'll use Pandas to analyze the career statistics of basketball player Pascal Siakam. Here's a breakdown of the Python code we'll use to load his data:

```
import pandas as pd

url = 'https://raw.githubusercontent.com/freeCodeCamp/boilerplate-pandas-
nba/master/nba.csv'
df = pd.read_csv(url)
display(df)
```

Code Explanation

1. We import the Pandas library and give it the shorter name 'pd'.
2. We provide a web address (URL) where the data is stored in a CSV file.
3. Pandas reads the data from the CSV and stores it in an object called 'df' (short for DataFrame).
4. Finally, we use `display()` to show the DataFrame's contents for analysis.

The Raw Data

The raw CSV file looks like this:

```
PLAYER_ID,SEASON_ID,LEAGUE_ID,TEAM_ID,TEAM_ABBREVIATION,PLAYER_AGE,GP,GS,MIN,
FGM,FGA,FG_PCT,FG3M,FG3A,FG3_PCT,FTM,FTA,FT_PCT,OREB,DREB,REB,AST,STL,BLK,TOV
,PF,PTS,FG2M,FG2A,FG2_PCT
1627783,2016-
17,00,1610612761,TOR,23.0,55,38,859.0,103,205,0.502,1,7,0.143,22,32,0.688,64,
121,185,17,26,45,33,109,229,102,198,0.515
<snip>
1627783,2022-
23,00,1610612761,TOR,29.0,71,71,2652.0,630,1313,0.48,93,287,0.324,367,474,0.7
74,131,425,556,415,65,36,169,228,1720,537,1026,0.523
1627783,2023-
24,00,1610612761,TOR,29.0,39,39,1354.0,325,623,0.522,46,145,0.317,169,223,0.7
58,54,192,246,190,32,10,83,87,865,279,478,0.584
1627783,2023-
24,00,1610612754,IND,29.0,24,24,786.0,203,370,0.549,28,73,0.384,60,86,0.698,4
6,122,168,102,17,8,34,62,494,175,297,0.589
```

1627783,2023-
24,00,0,TOT,29.0,63,63,2140.0,528,993,0.532,74,218,0.339,229,309,0.741,100,31
4,414,292,49,18,117,149,1359,454,775,0.586

This raw data presents Pascal Siakam's career statistics across various seasons. The last row summarizes his performance, potentially using sums or averages. Be aware that the two preceding rows indicate his trade to Indiana, which could impact data analysis.

Now that we understand the basic structure of the data, let's explore the columns in more detail.

Looking at the Columns

The next part of the lesson delves into understanding and manipulating the data columns within our DataFrame.

We start by listing all columns using the `df.columns` attribute, providing a clear view of the dataset's structure. This step is crucial for identifying the data available for analysis.

```
display(df.columns)
```

We also introduce looping through columns with a `for` loop, a foundational programming concept that enhances our data exploration capabilities.

```
for column in df.columns:  
    print(column)
```

Further, the lesson explains the significance of each column through a descriptive table, enhancing our understanding of the dataset's context.

Lastly, we focus on accessing specific columns, such as 'FGM' for field goals made, and illustrate how to select multiple columns, like 'FGM' and 'FGA', to analyze field goal attempts and successes.

```
display(df['FGM'])
```

```
display(df[['FGM', 'FGA']])
```


Exercise

This section is an exercise focused on refining data selection skills. The task involves modifying a given code snippet to specifically display columns related to shooting percentages: Field Goal Percentage, 2-Point Field Goal Percentage, 3-Point Field Goal Percentage, and Free Throw Percentage.

This exercise encourages students to apply their knowledge of selecting DataFrame columns, reinforcing their understanding of how to extract and analyze specific aspects of data effectively.

Here is the code snippet the student will modify:

```
import pandas as pd

# URL of the CSV file containing data for Pascal Siakam
url = 'https://raw.githubusercontent.com/pbeens/Data-Dunkers/main/Data/Pascal_Siakam.csv'

# Read the CSV file into a pandas DataFrame
df = pd.read_csv(url)

# ... (rest of your code)

# Display the DataFrame
display(df)
```