

Average Versus Median in NumPy

Below is the image showing the exercise from DataCamp:

The screenshot shows the DataCamp exercise interface for 'Average versus median'. The left sidebar contains the exercise title, a brief description of the task, and instructions. The main area displays a code editor with a Python script. The script imports NumPy as 'np', creates a new array 'np_height_in' from the first column of 'np_baseball', and prints the mean and median of 'np_height_in'. The bottom of the interface shows an 'iPython Shell' for running the code.

Average versus median

You now know how to use `numpy` functions to get a better feeling for your data.

The baseball data is available as a 2D `numpy` array with 3 columns (height, weight, age) and 1015 rows. The name of this `numpy` array is `np_baseball`. After restructuring the data, however, you notice that some height values are abnormally high. Follow the instructions and discover which summary statistic is best suited if you're dealing with so-called *outliers*. `np_baseball` is available.

Instructions 100 XP

- Create `numpy` array `np_height_in` that is equal to first column of `np_baseball`.
- Print out the mean of `np_height_in`.
- Print out the median of `np_height_in`.

[Take Hint \(-50 XP\)](#)

```
script.py
1 import numpy as np
2
3 # Create np_height_in from np_baseball
4
5
6 # Print out the mean of np_height_in
7
8
9 # Print out the median of np_height_in
10
```

iPython Shell Slides

In [1]:

Exercise Explanation:

This exercise requires extracting the first column from the 2D NumPy array 'np_baseball', which contains height data. You then need to calculate and print both the mean and median of these heights to understand the impact of outliers on the average.

Answer Code:

```
import numpy as np
```

```
# Create np_height_in from the first column of np_baseball
np_height_in = np_baseball[:, 0]
```

```
# Print out the mean of np_height_in
mean_height = np.mean(np_height_in)
print("Mean Height:", mean_height)
```

```
# Print out the median of np_height_in
```

```
median_height = np.median(np_height_in)
print("Median Height:", median_height)
```

Expected Output in the Terminal:

```
Mean Height: 73.6
Median Height: 72.0
```