

Flight Duration - Converting to Float

Question and Screenshot:

Learn / Courses / Exploratory Data Analysis In Pyt...

← Course Outline →

Daily XP 1351

Exercise

Flight duration

You would like to analyze the duration of flights, but unfortunately, the "Duration" column in the `planes` DataFrame currently contains string values.

You'll need to clean the column and convert it to the correct data type for analysis. `seaborn` has been imported as `sns`.

Instructions 3/4 25 XP

- Convert the column to `float` data type.

Take Hint (-7 XP)

script.py

```
1 # Preview the column
2 print(planes["Duration"].head())
3
4 # Remove the string character
5 planes["Duration"] = planes["Duration"].str.replace("h",
6
7 # Convert to float data type
8 planes["Duration"] = ____
```

Run Code Submit Answer

IPython Shell

```
<script.py> output:
0          19.0h
1  5.416666666666667h
2          4.75h
3  2.4166666666666665h
4          15.5h
Name: Duration, dtype: object

In [1]:
```

Question Explanation:

This task involves converting the 'Duration' column from a string data type to a float data type. This step is crucial for enabling numerical operations and analysis on the column values.

Code Solution:

```
# Preview the column
print(planes["Duration"].head())

# Remove the string character "h"
planes["Duration"] = planes["Duration"].str.replace("h", "", regex=False)

# Convert to float data type
planes["Duration"] = planes["Duration"].astype(float)
```

Solution Explanation:

1. The `str.replace()` method removes the 'h' character, preparing the values for conversion.
2. The `astype(float)` method converts the cleaned string values into a float data type.
3. This ensures the column is ready for numerical operations such as aggregations or visualizations.