

# Calculating Median Duration by Airline

## Question and Screenshot:

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

← Course Outline →

Daily XP 1412

Exercise

### Adding descriptive statistics

Now "Duration" and "Price" both contain numeric values in the `planes` DataFrame, you would like to calculate summary statistics for them that are conditional on values in other columns.

Instructions 2/3 35 XP

✓ Add a column to `planes` containing the standard deviation of "Price" based on "Airline".

2

Calculate the median for "Duration" by "Airline", storing it as a column called "airline\_median\_duration".

Take Hint (-10 XP)

3

Find the mean "Price" by "Destination", saving it as a column called "price\_destination\_mean".

script.py

Light Mode

```
1 # Median Duration by Airline
2 planes["airline_median_duration"] = _____._____("_____")
3   ["_____"]._____(): _____()
4 print(planes[["Airline", "airline_median_duration"]].
      value_counts())
```

↺ Run Code Submit Answer

IPython Shell

Slides

<script.py> output:

Airline	airline_price_st_dev	
Jet Airways	4230.749	3685
IndiGo	2266.754	1981
Air India	3865.872	1686
Multiple carriers	3763.675	1148
SpiceJet	1790.852	787
Vistara	2864.268	455
Air Asia	2016.739	309
GoAir	2790.815	182

dtype: int64

In [1]:

## Question Explanation:

The task involves calculating the median duration of flights for each airline in the `planes` DataFrame. The result is stored in a new column called 'airline\_median\_duration', and the frequency of unique values is displayed.

## Code Solution:

```
# Median Duration by Airline
planes["airline_median_duration"] = planes.groupby("Airline")
["Duration"].transform("median")

# Display the frequency of unique values in the new column
print(planes[["Airline", "airline_median_duration"]].value_counts())
```

## Solution Explanation:

1. The ``groupby('Airline')`` method groups the DataFrame by the 'Airline' column.
2. The ``transform('median')`` calculates the median duration of flights for each group.
3. The calculated median values are stored in a new column called 'airline\_median\_duration', and the ``value_counts()`` method displays the frequency of unique airline and median duration combinations.