

CSV to DataFrame (Final Results)

After completing the data manipulations, the final step is to view the results of the analysis, which includes the total number of passengers, total bumps, and bumps per 10k passengers for each airline.

Instructions:

1. Print the `airline_totals` DataFrame to see the results of your manipulations.

Original Uploaded Image:

The screenshot shows a DataCamp exercise titled "CSV to DataFrame". The instructions state: "You work for an airline, and your manager has asked you to do a competitive analysis and see how often passengers flying on other airlines are involuntarily bumped from their flights. You got a CSV file (airline_bumping.csv) from the Department of Transportation containing data on passengers that were involuntarily denied boarding in 2016 and 2017, but it doesn't have the exact numbers you want. In order to figure this out, you'll need to get the CSV into a pandas DataFrame and do some manipulation." The instructions also note: "pandas is imported for you as pd. 'airline_bumping.csv' is in your working directory." The code editor shows the following Python code:

```
1 # From previous steps
2 airline_bumping = pd.read_csv("airline_bumping.csv")
3 print(airline_bumping.head())
4 airline_totals = airline_bumping.groupby("airline")[["nb_bumped", "total_passengers"]].sum()
5 airline_totals["bumps_per_10k"] = airline_totals["nb_bumped"] / airline_totals
6   ["total_passengers"] * 10000
7
8 # Print airline_totals
9
```

The "Python Shell" output window displays the following DataFrame:

airline	total_passengers	nb_bumped	bumps_per_10k
ALASKA AIRLINES	1392	36543121	0.361
AMERICAN AIRLINES	11115	197345225	0.563
DELTA AIR LINES	1591	197833215	0.081
EXPRESSJET AIRLINES	3326	27858678	1.194
FRONTIER AIRLINES	1228	22954995	0.535
HAWAIIAN AIRLINES	122	16577572	0.074
JETBLUE AIRWAYS	3615	53245066	0.679
SKYWEST AIRLINES	3894	47692737	0.457
SOUTHWEST AIRLINES	16585	228142636	0.815
SPIRIT AIRLINES	2920	32304571	0.904
UNITED AIRLINES	4941	134468897	0.367
VIRGIN AMERICA	242	12617967	0.201

Python Code Implementation:

```
# From previous steps
```

```
airline_bumping = pd.read_csv("airline_bumping.csv")
airline_totals = airline_bumping.groupby("airline")[["nb_bumped",
"total_passengers"]].sum()
airline_totals["bumps_per_10k"] = airline_totals["nb_bumped"] /
airline_totals["total_passengers"] * 10000
```

```
# Print airline_totals
print(airline_totals)
```

Explanation of Code:

1. ****Load and group data****: Previously, the CSV file was read and grouped by `airline` to calculate totals for `nb_bumped` and `total_passengers`.
2. ****Calculate bumps per 10k****: A column `bumps_per_10k` was added to calculate the number of bumps per 10,000 passengers.
3. ****Print the DataFrame****: Finally, `airline_totals` is printed to display the results.