

Comparing Arrest Rates by Violation and Weather Rating

The screenshot shows a web browser window displaying a DataCamp exercise titled "Comparing arrest rates by weather rating". The exercise instructions are as follows:

- Do police officers arrest drivers more often when the weather is bad? Let's find out!
- First, you'll calculate the overall arrest rate.
- Then, you'll calculate the arrest rate for each of the weather ratings you previously assigned.
- Finally, you'll add violation type as a second factor in the analysis, to see if that accounts for any differences in the arrest rate.

Since you previously defined a logical order for the weather categories, `good < bad < worse`, they will be sorted that way in the results.

The exercise is worth 30 XP. The instructions are 3/3. The first instruction is checked. The second instruction is checked. The third instruction is highlighted. A "Take Hint (4 XP)" button is visible.

The code editor shows the following code:

```
1 # Calculate the arrest rate for each 'violation' and 'rating'
2 print(ri_weather. ....)
```

The output shows the following data:

violation	rating	is_arrested
Equipment	good	0.059
Equipment	bad	0.066
Equipment	worse	0.097
Moving violation	good	0.056
Moving violation	bad	0.058
Moving violation	worse	0.066
Other	good	0.077
Other	bad	0.067
Other	worse	0.065
Registration/plates	good	0.082
Registration/plates	bad	0.068
Registration/plates	worse	0.116
Seat belt	good	0.029
Seat belt	bad	0.022
Seat belt	worse	0.066
Speeding	good	0.013
Speeding	bad	0.023
Speeding	worse	0.027

Question

Do police officers arrest drivers more often when the weather is bad? Let's find out!

1. First, you'll calculate the overall arrest rate.
2. Then, you'll calculate the arrest rate for each of the weather ratings group you previously assigned.
3. Finally, you'll calculate the arrest rate for each combination of violation and rating. How do the arrest rates differ by group?

Solution

```
# Import pandas library
```

```
import pandas as pd
```

```
# Calculate the arrest rate for each combination of violation and rating
```

```
arrest_rate_by_violation_and_rating = ri_weather.groupby(['violation',  
'rating']).is_arrested.mean()  
print(arrest_rate_by_violation_and_rating)
```

Explanation

1. The first step imports pandas for data manipulation.
2. We use `groupby()` to group the data by both 'violation' and 'rating'.
3. The mean of 'is_arrested' is calculated for each group, providing insights into how arrest rates differ across combinations of violation type and weather rating.