

## Cross-tabulation of Job\_Category and Company\_Size

Question:

Perform a cross-tabulation to analyze how observations in Job\_Category and Company\_Size relate to each other. Use pandas crosstab to group by Job\_Category and use Company\_Size as columns.

Explanation of the Question (40 words):

The task involves creating a cross-tabulation using pandas to compare Job\_Category as rows and Company\_Size as columns. This identifies trends, patterns, or relationships in the data, particularly to observe mean Salary\_USD for each combination.

Full Answer:

```
import pandas as pd

# Sample dataset
data = {
    "Job_Category": ["EN", "EX", "EN", "MI", "SE", "MI", "SE", "EN"],
    "Company_Size": ["L", "M", "L", "S", "M", "L", "S", "M"],
    "Salary_USD": [70000, 85000, 72000, 40000, 91000, 63000, 41000,
95000],
}

# Create a DataFrame
df = pd.DataFrame(data)

# Perform cross-tabulation
cross_tab = pd.crosstab(index=df["Job_Category"],
                        columns=df["Company_Size"],
                        values=df["Salary_USD"],
                        aggfunc='mean')

# Print the result
print(cross_tab)
```

Explanation of the Answer (50 words):

The script loads a dataset into a pandas DataFrame. Using pandas.crosstab(), it groups data by Job\_Category and Company\_Size,

calculating the mean Salary\_USD for each combination. The crosstab provides a summarized view of average salaries based on job categories and company sizes.