Pivot

- Cross-classifies according to two categorical variables
- Produces a grid of counts or aggregated values
- Two required arguments:
  - First: variable that forms column labels of grid
  - Second: variable that forms row labels of grid
- Two optional arguments (include both or neither)
  - values='column_label_to_aggregate'
  - collect=function_with_which_to_aggregate

(Demo)
## Joining Two Tables

Keep all rows in the table that have a match ...

... for the value in this column ...

... somewhere in this other table's ...

... column that contains matching values.

```python
drinks = {
    'Milk Tea': {'Cafe': 'Tea One', 'Price': 4},
    'Espresso': {'Cafe': 'Nefeli', 'Price': 2},
    'Latte': {'Cafe': 'Nefeli', 'Price': 3},
    'Espresso': {'Cafe': 'Abe''s', 'Price': 2}
}

discounts = {
    'Tea One': {'Coupon': 25, 'Location': 'Tea One'},
    'Nefeli': {'Coupon': 50, 'Location': 'Nefeli'},
    'Tea One': {'Coupon': 5, 'Location': 'Tea One'}
}

drinks['Milk Tea']['Coupon'] = discounts['Tea One']['Coupon']

The joined column is sorted automatically.
```
Monty Hall Problem
Random Selection

\texttt{np.random.choice}

- Selects at random
- with replacement
- from an array
- a specified number of times

\texttt{np.random.choice(some_array, sample_size)}
Discussion Question

```python
d = np.arange(6) + 1
```

What results from evaluating the following 2 expressions? Are they the same? Do they describe the same process?

1. `np.random.choice(d, 1000) + np.random.choice(d, 1000)`
2. `2 * np.random.choice(d, 1000)`
Probability

- Lowest value: 0
  - Chance of event that is impossible
- Highest value: 1 (or 100%)
  - Chance of event that is certain

- If an event has chance 70%, then the chance that it doesn’t happen is
  - 100% - 70% = 30%
  - 1 - 0.7 = 0.3
Equally Likely Outcomes

Assuming all outcomes are equally likely, the chance of an event A is:

\[ P(A) = \frac{\text{number of outcomes that make A happen}}{\text{total number of outcomes}} \]
Simulating Monty Hall

x 1,000
Multiplication Rule

Chance that two events $A$ and $B$ both happen

$$= P(A \text{ happens}) \times P(B \text{ happens given that } A \text{ has happened})$$

- The answer is less than or equal to each of the two chances being multiplied
- The more conditions you have to satisfy, the less likely you are to satisfy them all
Addition Rule

If event $A$ can happen in exactly one of two ways, then

$$P(A) = P(\text{first way}) + P(\text{second way})$$

- The answer is \textit{greater than or equal to} the chance of each individual way
Example: At Least One Head

- In 3 tosses:
  - Any outcome except TTT
  - \( P(TTT) = \left(\frac{1}{2}\right) \times \left(\frac{1}{2}\right) \times \left(\frac{1}{2}\right) = \frac{1}{8} \)
  - \( P(\text{at least one head}) = 1 - P(TTT) = \frac{7}{8} = 87.5\% \)

- In 10 tosses:
  - \( 1 - \left(\frac{1}{2}\right)^{10} \)
  - 99.9%

Test 1

- **Topics**
  - Causality
  - Python
    - Data (names, values & types)
    - Expressions (numbers, strings, arrays, & tables)
    - Functions
  - Probability
  - Visualization
    - Charts & Histograms

- Do the review questions. Post questions to Piazza!
- Bring 2 sheets of notes