Lecture 12

Table Examples
Announcements
Combining Table Methods
Important Table Methods

t.select(column, ...) or t.drop(column, ...)
t.take([row, ...]) or t.exclude([row, ...])
t.sort(column, descending=False, distinct=False)
t.where(column, are.condition(...))
t.apply(function, column, ...)
t.group(column) or t.group(column, function)
t.group([column, ...]) or t.group([column, ...], function)
t.pivot(cols, rows) or t.pivot(cols, rows, vals, function)
t.join(column, other_table, other_table_column)

http://data8.org/datascience/tables.html
### Discussion Question

Generate a table with one row per cafe that has the name and discounted price of its cheapest discounted drink

<table>
<thead>
<tr>
<th>Drinks</th>
<th>Cafe</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milk Tea</td>
<td>Tea One</td>
<td>4</td>
</tr>
<tr>
<td>Espresso</td>
<td>Nefeli</td>
<td>2</td>
</tr>
<tr>
<td>Coffee</td>
<td>Nefeli</td>
<td>3</td>
</tr>
<tr>
<td>Espresso</td>
<td>Abe's</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discounts</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>Tea One</td>
</tr>
<tr>
<td>50%</td>
<td>Nefeli</td>
</tr>
<tr>
<td>25%</td>
<td>Tea One</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cheapest</th>
<th>Cafe</th>
<th>Drink</th>
<th>Discounted Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nefeli</td>
<td>Espresso</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Tea One</td>
<td>Milk Tea</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

(Demo)
(b) (8 pt) Each row of the *trip* table from lecture describes a single bicycle rental in the San Francisco area. Durations are integers representing times in seconds. The first three rows out of 338343 appear below.

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferry Building</td>
<td>SF Caltrain</td>
<td>765</td>
</tr>
<tr>
<td>San Antonio Shopping Center</td>
<td>Mountain View City Hall</td>
<td>1036</td>
</tr>
<tr>
<td>Post at Kearny</td>
<td>2nd at South Park</td>
<td>307</td>
</tr>
</tbody>
</table>

Write a Python expression below each of the following descriptions that computes its value. The first one is provided for you. You *may* use up to two lines and introduce variables.

- The average duration of a rental.
  ```python
total_duration = sum(trip.column(2))
total_duration / trip.num_rows
  ```
- The name of the station where the most rentals ended (assume no ties).
- The number of stations for which the average duration ending at that station was more than 10 minutes.
Advanced Where
Comparison Operators

The result of a comparison expression is a `bool` value

\[
\begin{align*}
x & = 2 \\
y & = 3
\end{align*}
\]

Assignment statements

\[
\begin{align*}
x & > 1 \\
x & > y \\
y & >= 3
\end{align*}
\]

Comparison expressions

\[
\begin{align*}
x & == y \\
x & != 2 \\
2 & < x < 5
\end{align*}
\]

t.where(array_of_bool_values) returns a table with only the rows of t for which the corresponding bool is True.

(Demo)
ZIP Codes

(Demo)