



Lecture 12

Table Examples

Announcements

Table Review

Important Table Methods

`t.select(column, ...) or t.drop(column, ...)`

`t.take([row_num, ...]) or t.exclude([row_num, ...])`

`t.sort(column, descending=False)`

`t.where(column, are.condition(...))`

`t.apply(function_name, column, ...)`

`t.group(column) or t.group(column, function_name)`

`t.group([column, ...]) or t.group([column, ...], function_name)`

`t.pivot(cols, rows) or t.pivot(cols, rows, vals, function_name)`

`t.join(column, other_table, other_table_column)`

<http://data8.org/sp22/python-reference.html>

Table Practice

Join for Value Annotation

One common use of `t.join(_, u, _)`:

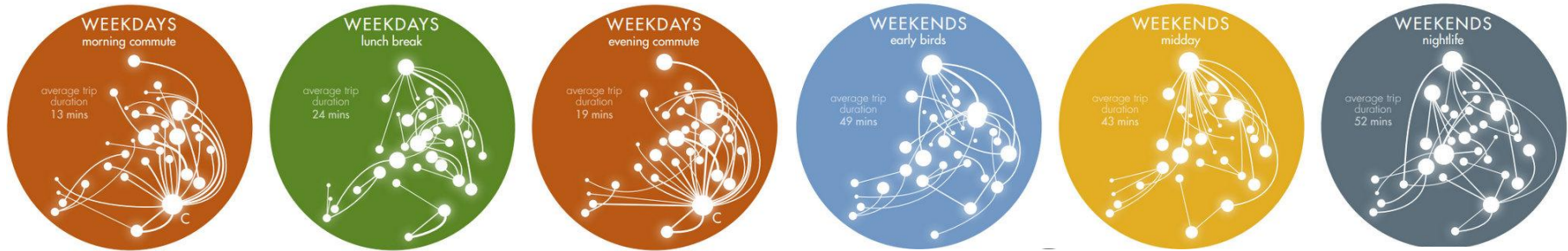
- A dataset `t` has a categorical variable `x`.
- A table `u` has one row per possible value of `x` that describes some properties of that value.
- The joined table has the same rows as `t`, but each row in `t` is now annotated with the properties of its `x` value.

(Demo)

Bike Sharing in SF Bay Area

Hourly bike sharing in the Bay Area began with a pilot program in 2014-2015 that produced a public dataset.

- The SF Metropolitan Transportation Commission organized an Open Data Challenge in which participants visualized the dataset in interesting ways.



by Bjorn Vermeersch

(Demo)