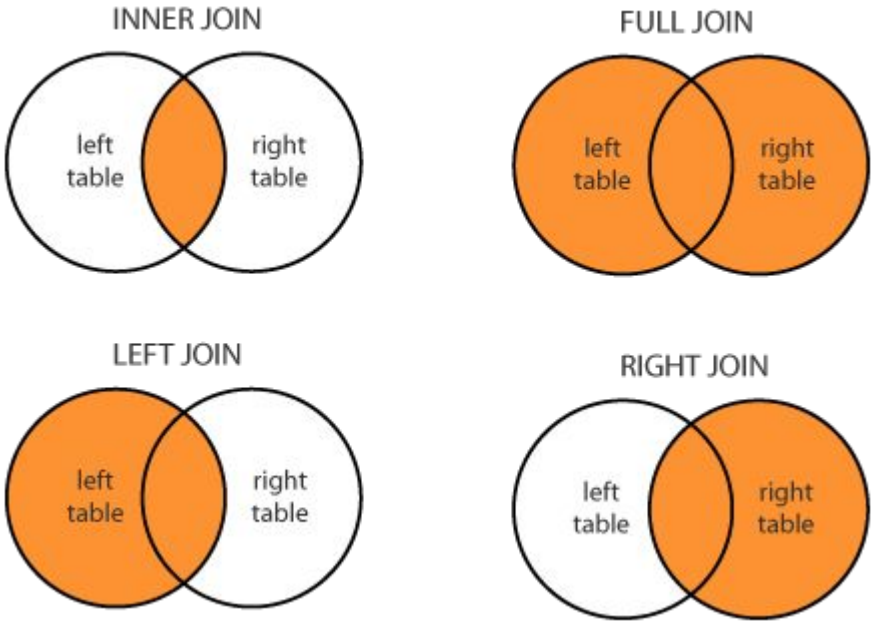
Combining and Merging Data Sets.

McKinney Chapter 8. Section 2.

Agenda

1. **Pandas.merge()** is similar to relational database *join* operations.
2. **Pandas.concatenate()** glues or stacks together objects along axis.
3. **Pandas.DataFrame.combine_first()** fills in missing values in one object with values from another.

Different types of joins / mergers:



Different types of joins / mergers cont.

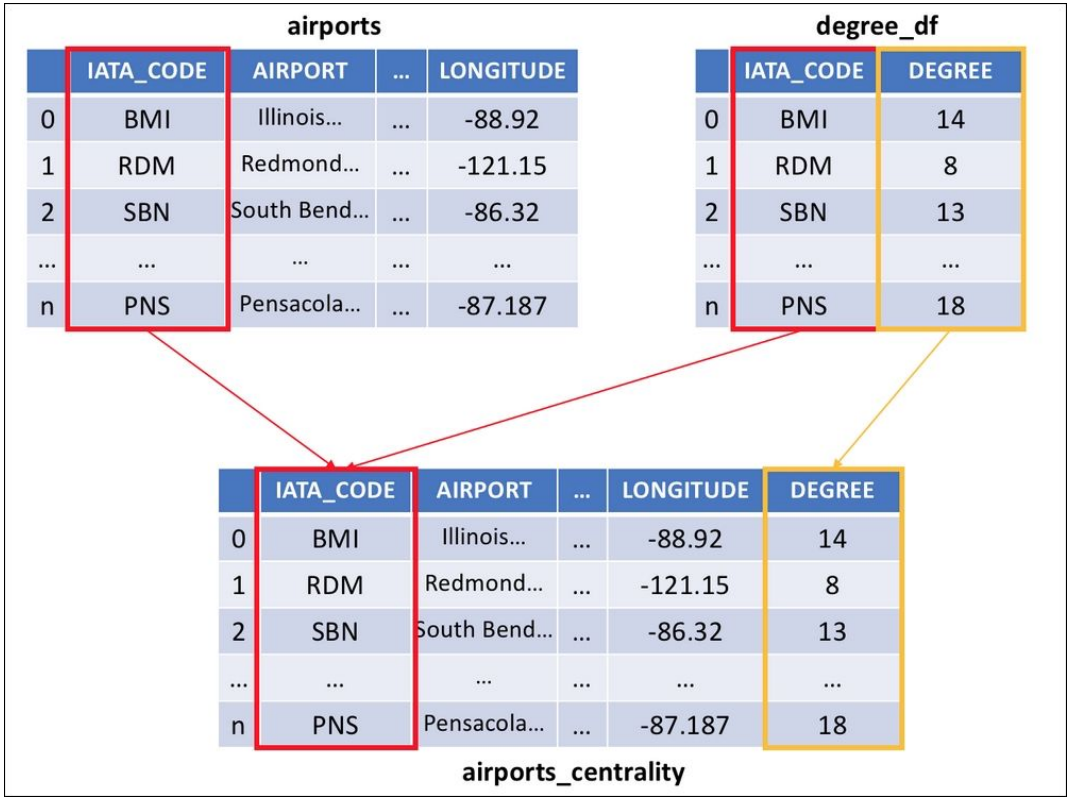
INNER JOIN: Select records that have matching values in both DataFrames.

LEFT OUTER JOIN: Select all of the records from the left-most DataFrame with matching right DataFrame records.

RIGHT OUTER JOIN: Select all of the records from the right-most DataFrame with matching left DataFrame records.

FULL OUTER JOIN: Selects all records that match either left or right DataFrame records.

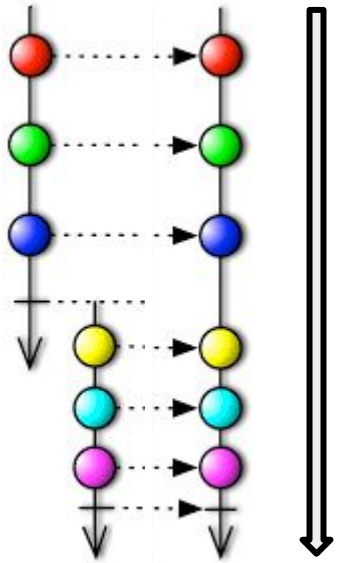
Join example



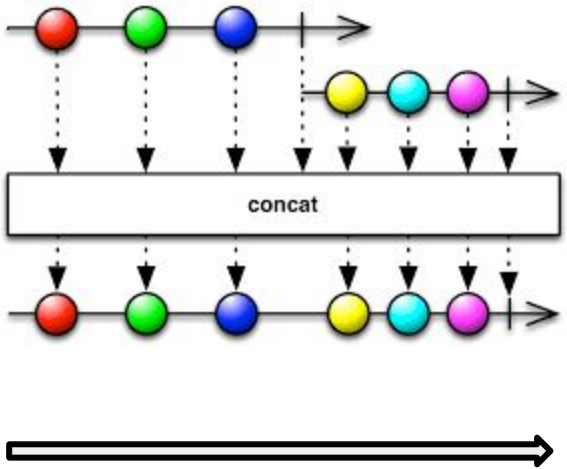
Pandas.concatenate() along axis 0

Pandas Concatenating Objects

Binding or stacking objects.



Along axis 0



Along axis 1

pandas.DataFrame.combine_first()

You can think of it as “**patching**” missing data in the calling object with the data from the object you pass.