# Joins

Say that you have a dataset with airline tickets. An observation is a ticket with an airline and route.

Say that you have a dataset with each airline’s market share of a route. An observation is an airline-route combination.

Say that you want a dataset where each observation is an airline ticket, but you also have the airline’s market share.

The way to achieve this is called a join. In Stata it is done via the merge command.

First, make sure that both of the datasets that you want to join are in Stata format.

Second, load one of the datasets into memory using the “use” command.

Third, figure out what the common variable is between the two datasets. For example, in the above the example it would be the airline-route combination.

Fourth, figure out how many copies of that key are present in each dataset. In the above example, the airline shows up on many rows in the ticket data, but only once in the dataset about market shares.

Write the appropriate command:

* Use /path/to/market\_share.dta, clear
* Merge 1:m airline\_id route\_id using /path/to/ticket.dta