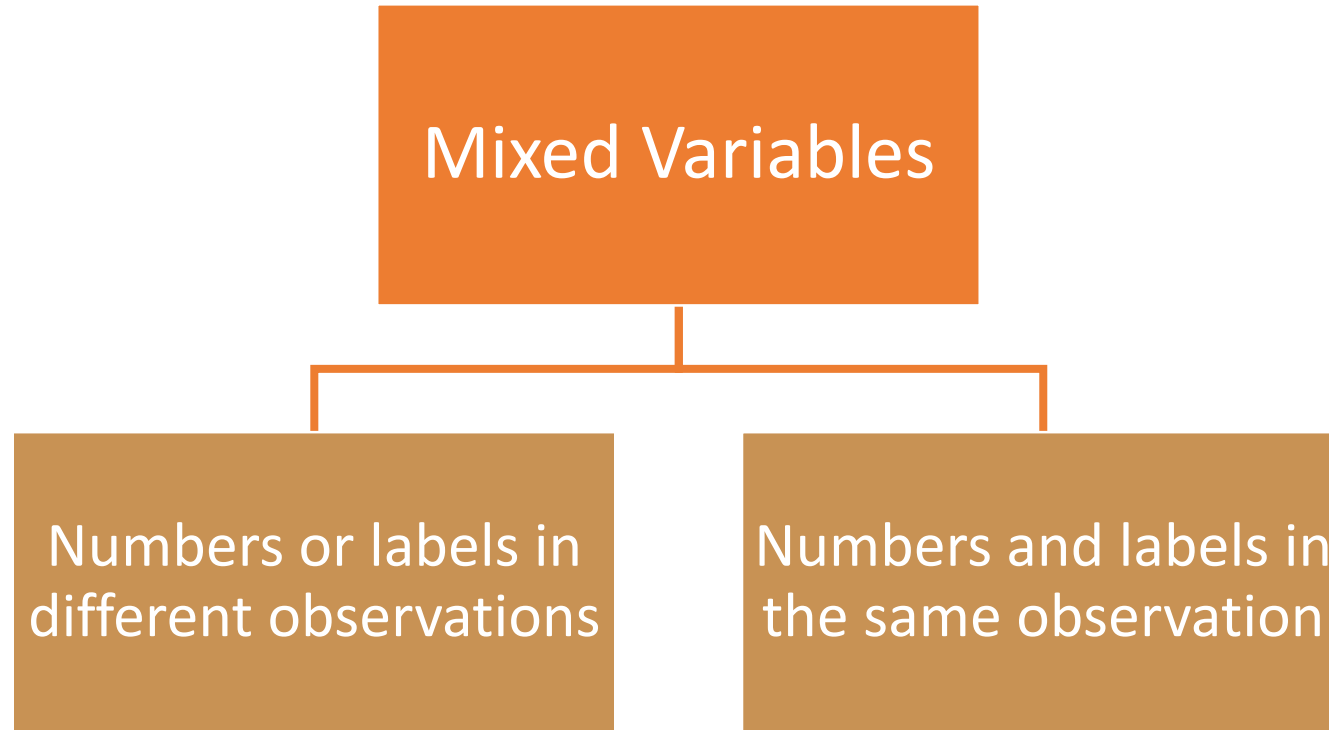




Mixed Variables

Mixed Variables



Mixed Variables

Observations show either numbers or categories among their values

- Number of credit accounts (1-100, U, T, M)
 - U = unknown, T = unverified, M = unmatched)
- Number of missed payments (1-3, D, A)
 - D = defaulted, A = arrangements

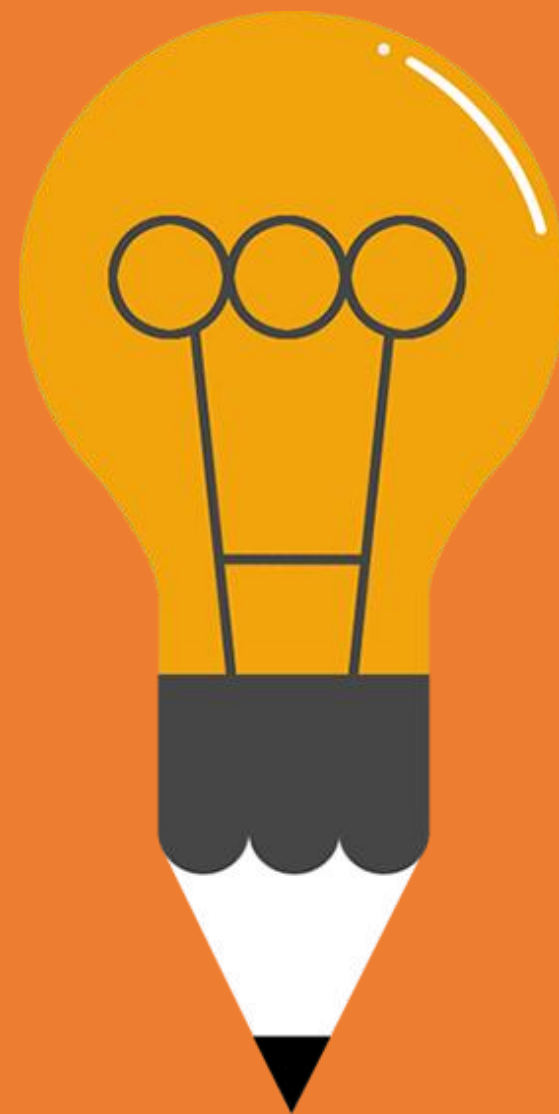
Mixed Variables

Observations show both numbers and categories in their values

- Cabin (Titanic) (A15, B18, ...)
- Ticket (Titanic) (A103349)
- Vehicle registration (AB500)
- Postcode (SE18)

Pre-processing mixed variables

- We can enrich the dataset dramatically by extracting information from the label and the number of the mixed variable.
- We will learn how in a dedicated section.



Accompanying Jupyter Notebook



- Read the accompanying Jupyter Notebook
- Download the sample dataset attached to this lecture.

THANK YOU

www.trainindata.com