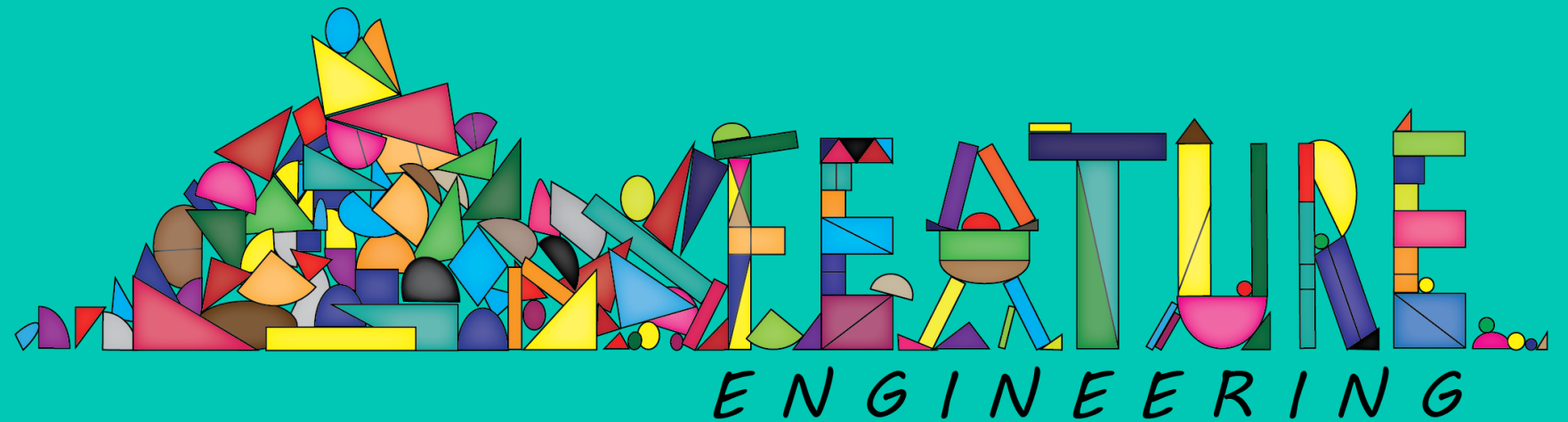
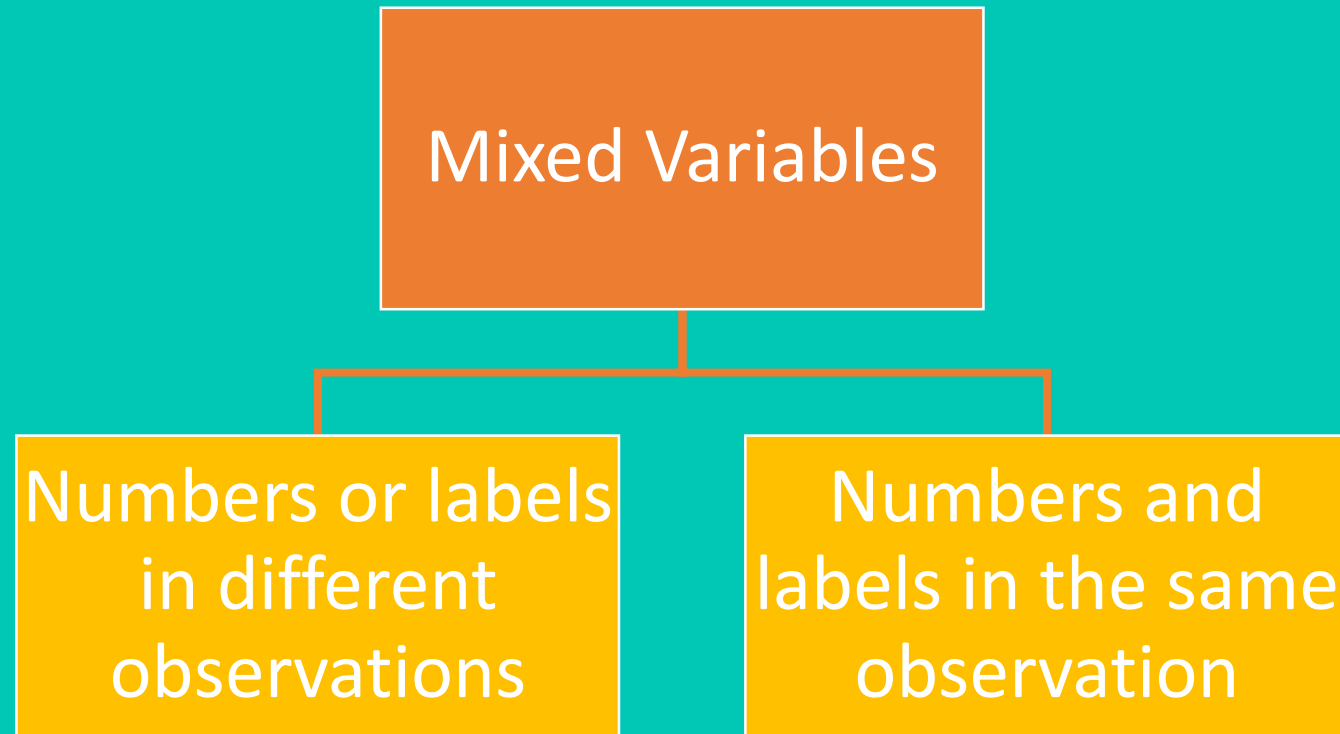


# Mixed variables



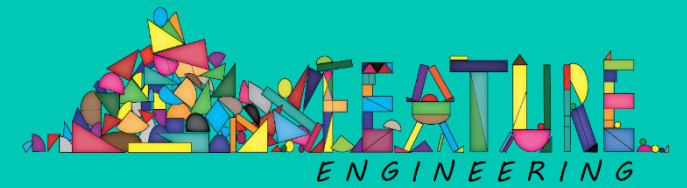
# Mixed variables



# Mixed variables

Observations present either numbers or categories among their values

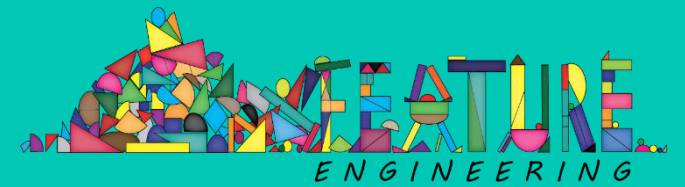
- Number of credit accounts (1-100, U, T, M)
  - U = unknown, T = unverified, M = unmatched)
- Number of missed payments (1-3, D, A)
  - D = defaulted, A = arrangements



# Mixed variables

Observations show both numbers and categories in their values

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



# Mixed variables

## Pre-processing

We can enrich the dataset dramatically by extracting information from the label and the number of the mixed variable

We will see how in a dedicated section of this course!