

# Data Processing

- Supervised ML
  - Defining “long” LOS  $> 220$  weeks
- Cross Validation
- Overfitting risk
- Forward selection
  - 90  $\rightarrow$  1, 2, 3..., 9 variables



