

# Data

David Robinson

## Data Processing

1. Remove features that are irrelevant.
2. Identify records with missing or outlier data and either drop the records or replace the data with the estimated value based on other information. For example, a human age of 10000 should be dropped or replaced with the mean age.
3. Engineer new features based on the existing features. For example, a boolean feature, `isAdult`, can be marked true for records with an age above 17.