Oct 8

Data Science Notes

Missing Data

* MCAR (Missing Completely at Random):
  + Missing value doesn’t depend on X or itself.
  + Unbiased estimates
* MAR (Missing at Random)
  + Missing value depend on X but not on itself.
  + Biased estimates
* MNAR (Missing not at Random)
  + Missing value depends on itself

Techniques to deal with missing data

* Imputation or Partial Imputation (give some values)
* Deletion or Partial Deletion
* Analysis
* Interpolation