
MISSING DATA AND ITS IMPACT ON STATISTICAL ANALYSIS

Saint Michael's College
Professor Michael Larsen
Conner Lorenz
Daniel Fortier

+

o

•

+

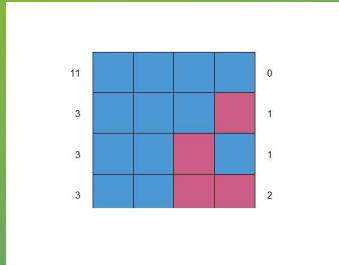
•

○

Missing Data

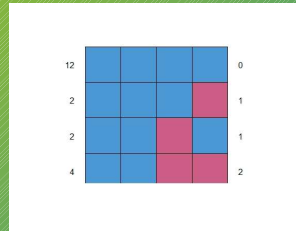
- Missing Data impacts
 - Bias
 - Standard Error
 - MSE
 - Coverage

Patterns of Missingness

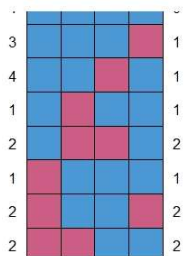


MNAR

+



MAR



MCAR

○

●



Methods of Dealing with Missingness

Complete Case Analysis

Mean Imputation

Hotdeck Imputation

Regression Imputation

Stochastic Regression Imputation

Simulation Results

Effect of missingness pattern is noticeable

Regression imputation has least bias, lowest MSE

Mean/regression imputations have least standard error

Complete case analysis has best coverage