

BDA

↳ summarizing the characteristics of data

statistical analysis + visualization

↙ help us see underlying structure
(ex. diff dataset has same mean value)

↓
but when we plot data, we will see that
data totally has diff pattern /
structure.

Univariate analysis

- mean & sd → visualize in histogram/pdf
- min, max, median, percentile, quartile
- * visualize → box plot *

Note: terminology

↳ proxy variable = variable that
we want to study instead

bivariate analysis

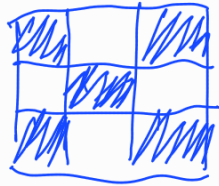
- correlation → visualization: scatter plot

ESDA (S = spatial)

- consider time & location (space)
- spatial autocorrelation

↳ relationship to neighbour

- positive value → clustered
- zero → random
- negative value → dispersed



Moran's I