

INCREASING COMPLEXITY

## EXPLORATORY ANALYSIS

Identify the main characteristics of the data set

## DISTRIBUTIONS & CORRELATIONS

Distributions inform subsequent statistical tests. Correlations identify relationships and confounding factors

## NULL HYPOTHESIS TESTS

Using a the Mann-Whitney test identify if there is a significant difference between metric populations

## SIMPLE LINEAR MODELLING

Test a linear model with team factors as the independent variables and CK metrics as the dependent variables

## LINEAR MIXED MODELLING

Enrich the model using the 'random intercepts' approach to accommodate for project-specific idiosyncrasy

## PROJECT ANALYSIS

Using clustering, identify two projects for deeper analysis, studying the code-level drivers that drive observed trends