

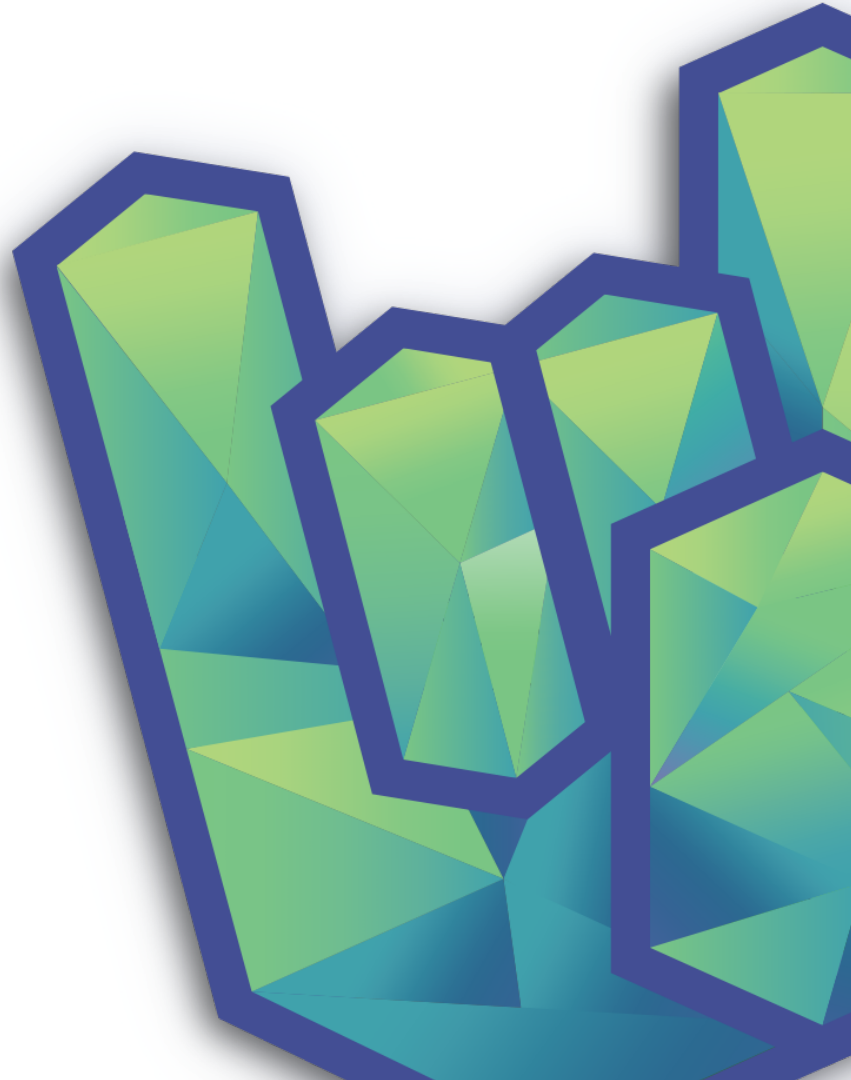
Aggregations



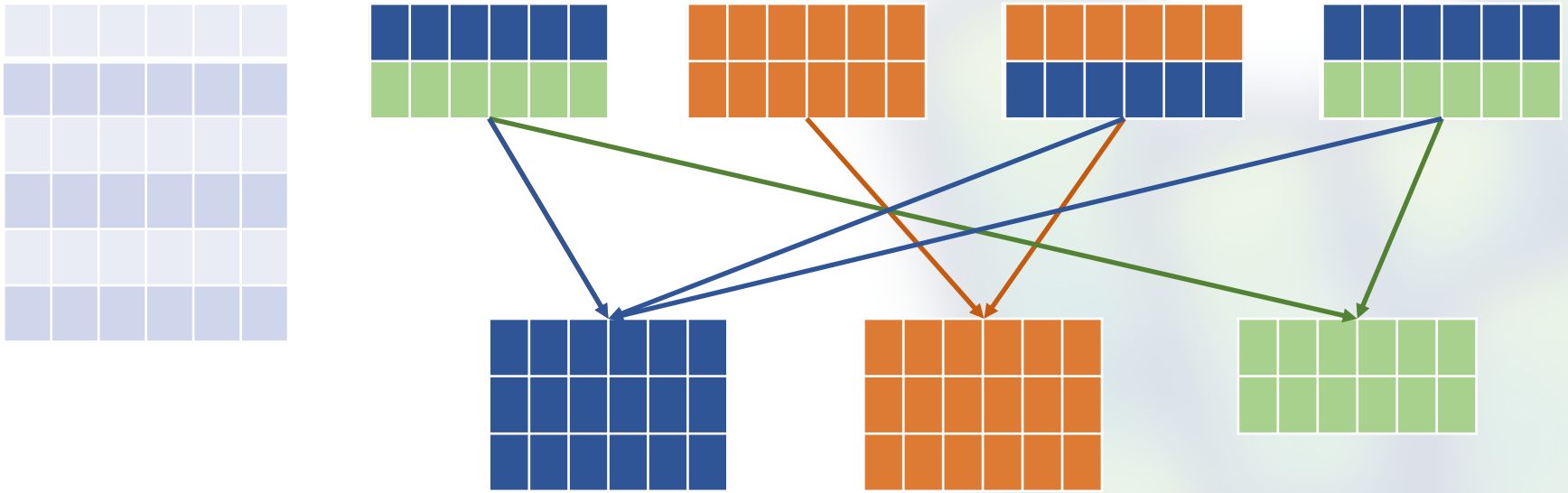
Objective

Aggregations and statistics

Grouping



Wide Transformations



One/more input partitions => One/more output partitions

Shuffle

Big performance impact

Spark rocks

