Team MEHL - IND Discovery Spark

A	B
Berlin	Berlin
	New York

Berlin	A, B
New York	B

FlatMap: DFs -> (value, attribute) ➡️ **GrouBy: Value (-> (value, including attributes))**
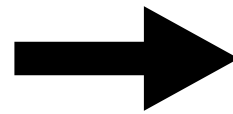
MapGroups: Including attributes -> toSet

A, B
B

A
B

FlatMap: Including attributes -> (Including attribute, excluding attributes) ⬅️ *Diff* **Broadcasted list of DF attributes**
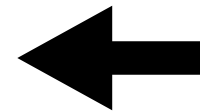
A	{}
B	A

GrouBy: -> Including attribute

Diff

MapGroups: (Including attribute, excluding attrs.) -> (column, INDs)

A	B