

Distributing Data in Hive

SORT BY Clause

("M", 1000)

("F", 1500)

("M", 500)

Mapper

("F", 2500)

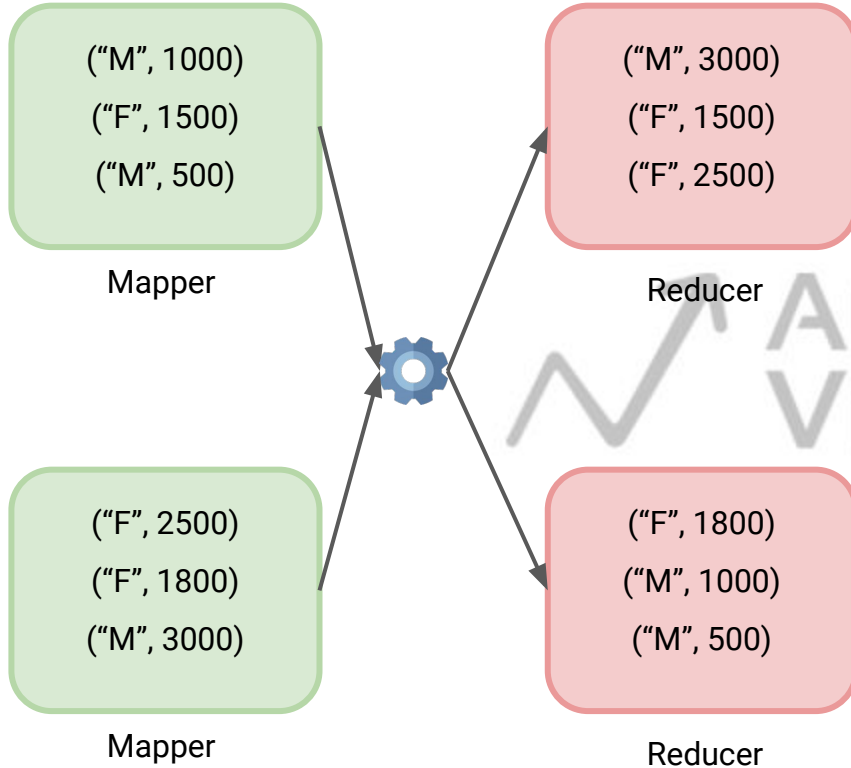
("F", 1800)

("M", 3000)

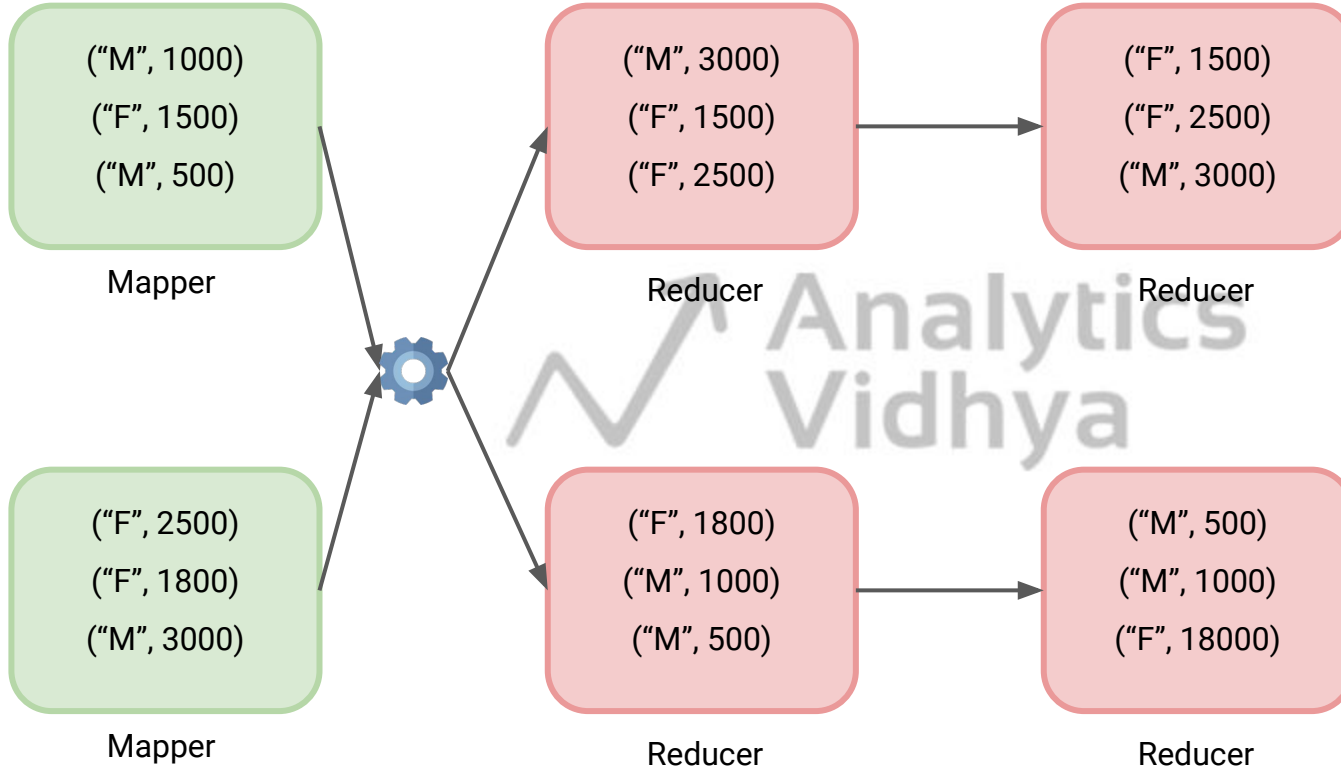
Mapper



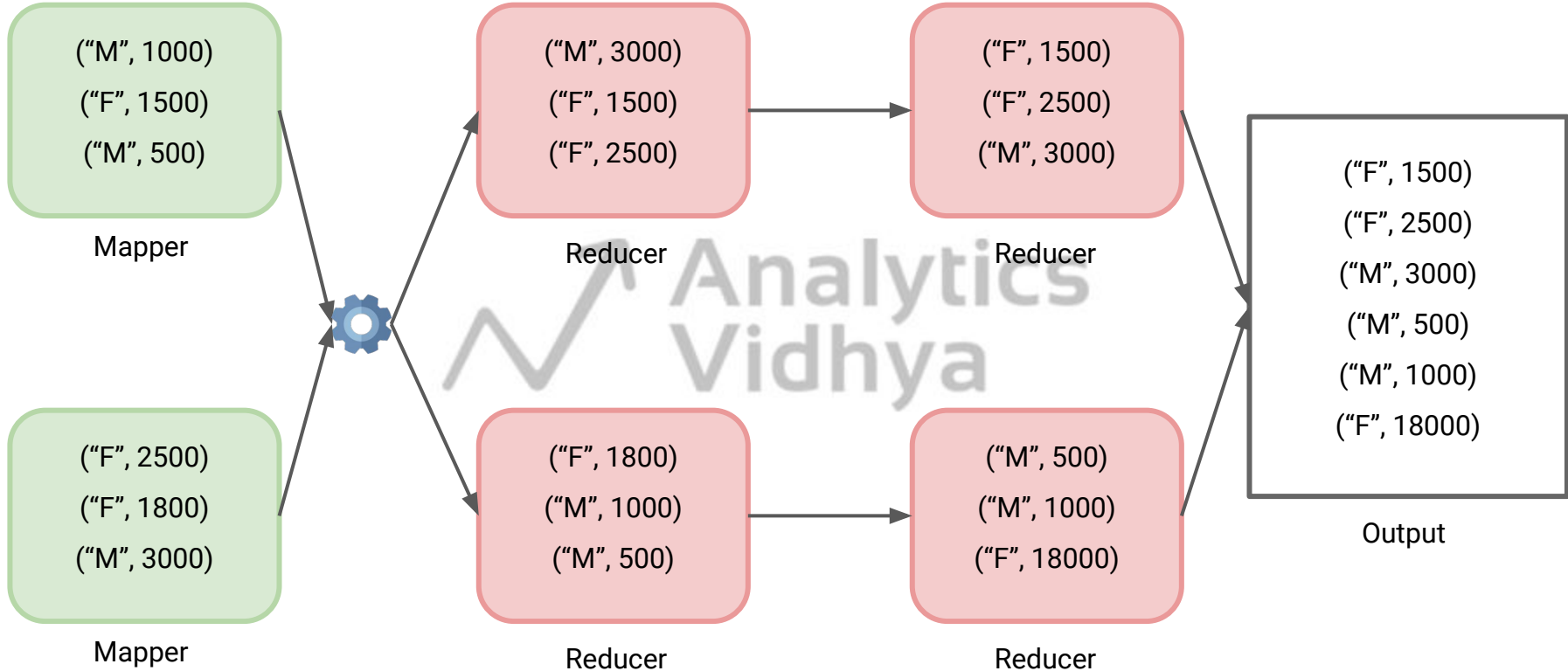
SORT BY Clause



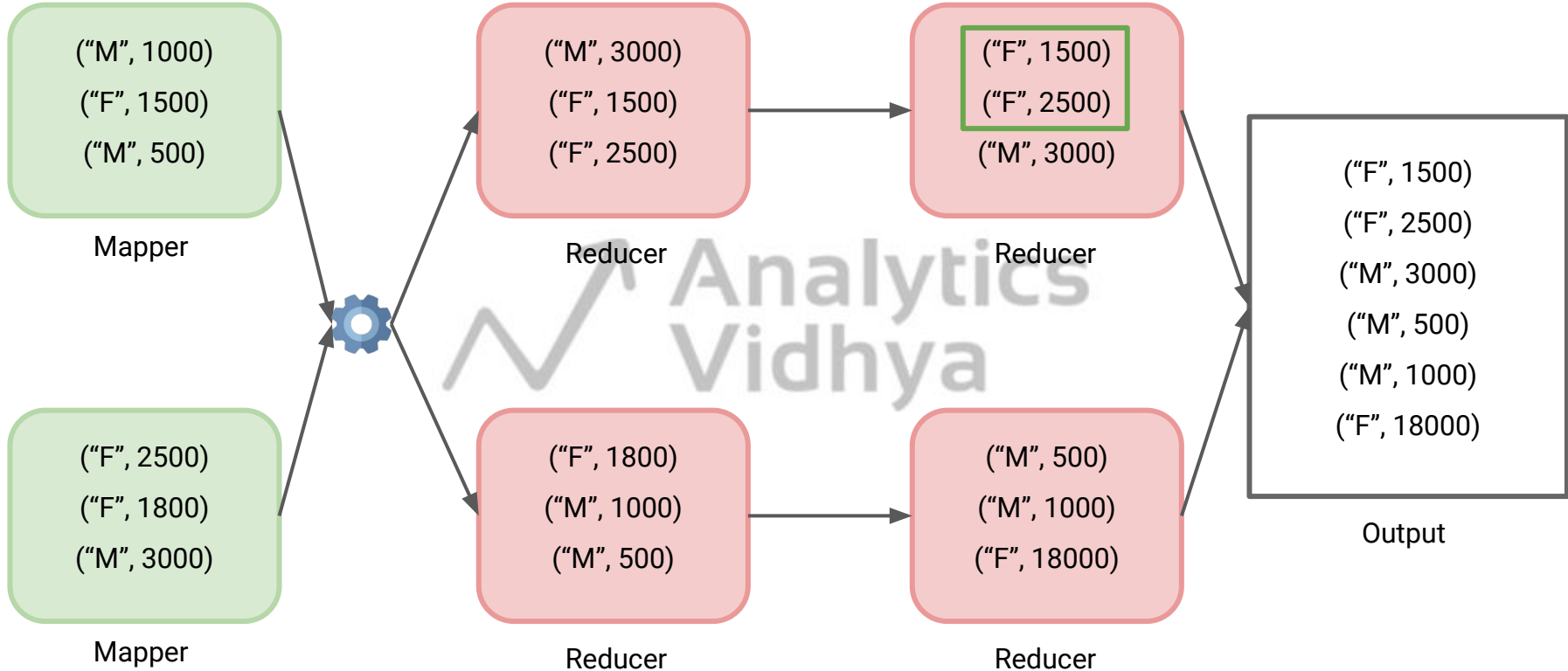
SORT BY Clause



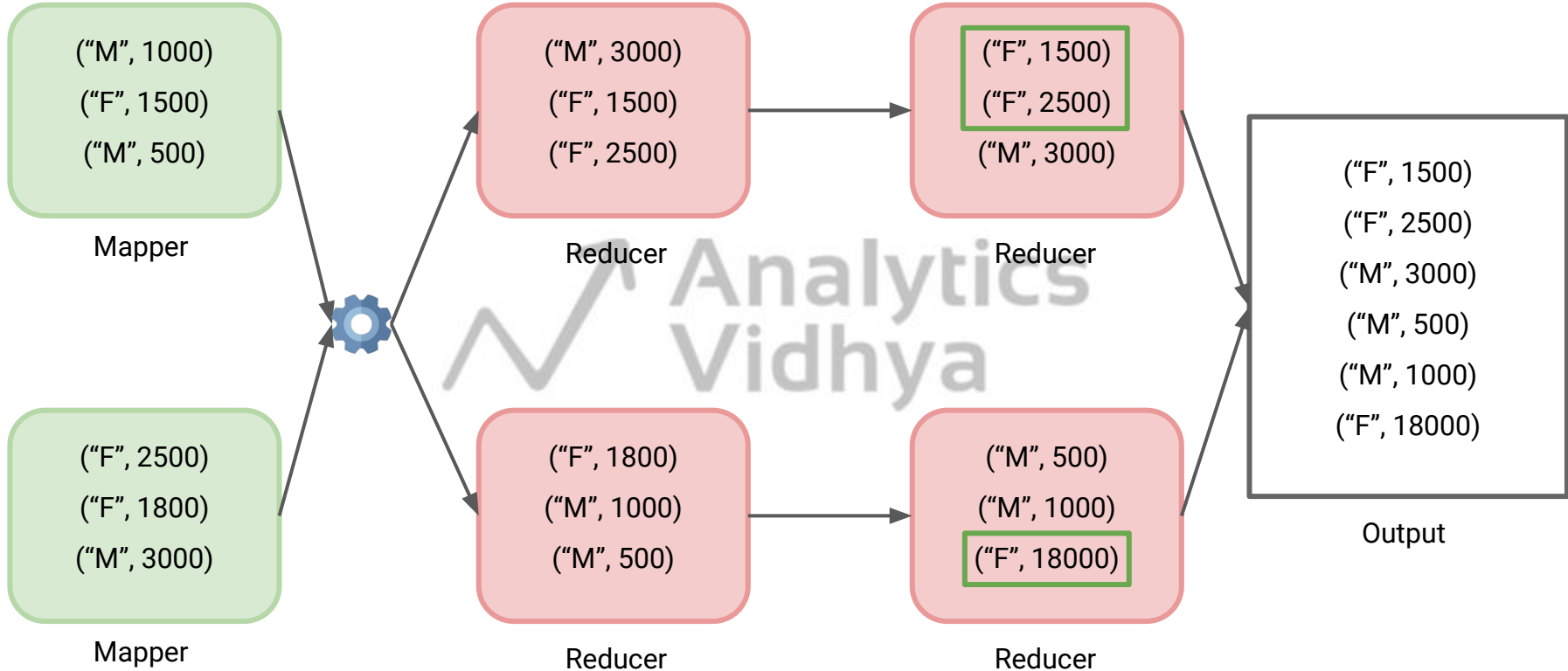
SORT BY Clause



SORT BY Clause



SORT BY Clause



DISTRIBUTE BY
Clause



DISTRIBUTE BY
Clause

("M", 1000)
("F", 1500)
("M", 500)

Mapper

("F", 2500)
("F", 1800)
("M", 3000)

Mapper

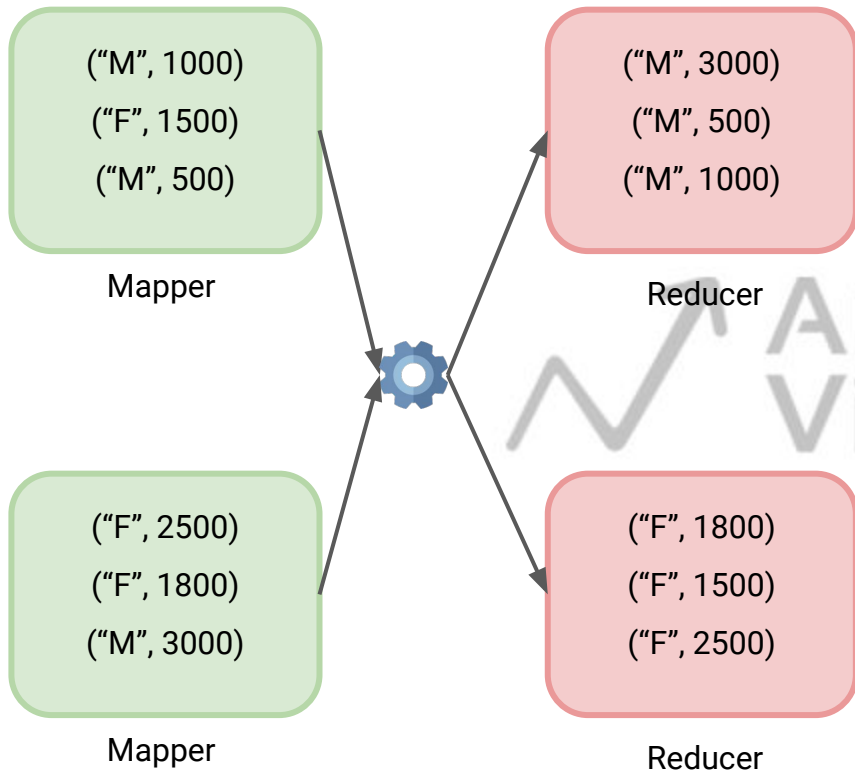


Reducer

Reducer

Analytics
Vidhya

DISTRIBUTE BY Clause



DISTRIBUTE BY
Clause

SORT BY Clause

("M", 1000)
("F", 1500)
("M", 500)

Mapper

("M", 3000)
("M", 500)
("M", 1000)

Reducer

("F", 2500)
("F", 1800)
("M", 3000)

Mapper

("F", 1800)
("F", 1500)
("F", 2500)

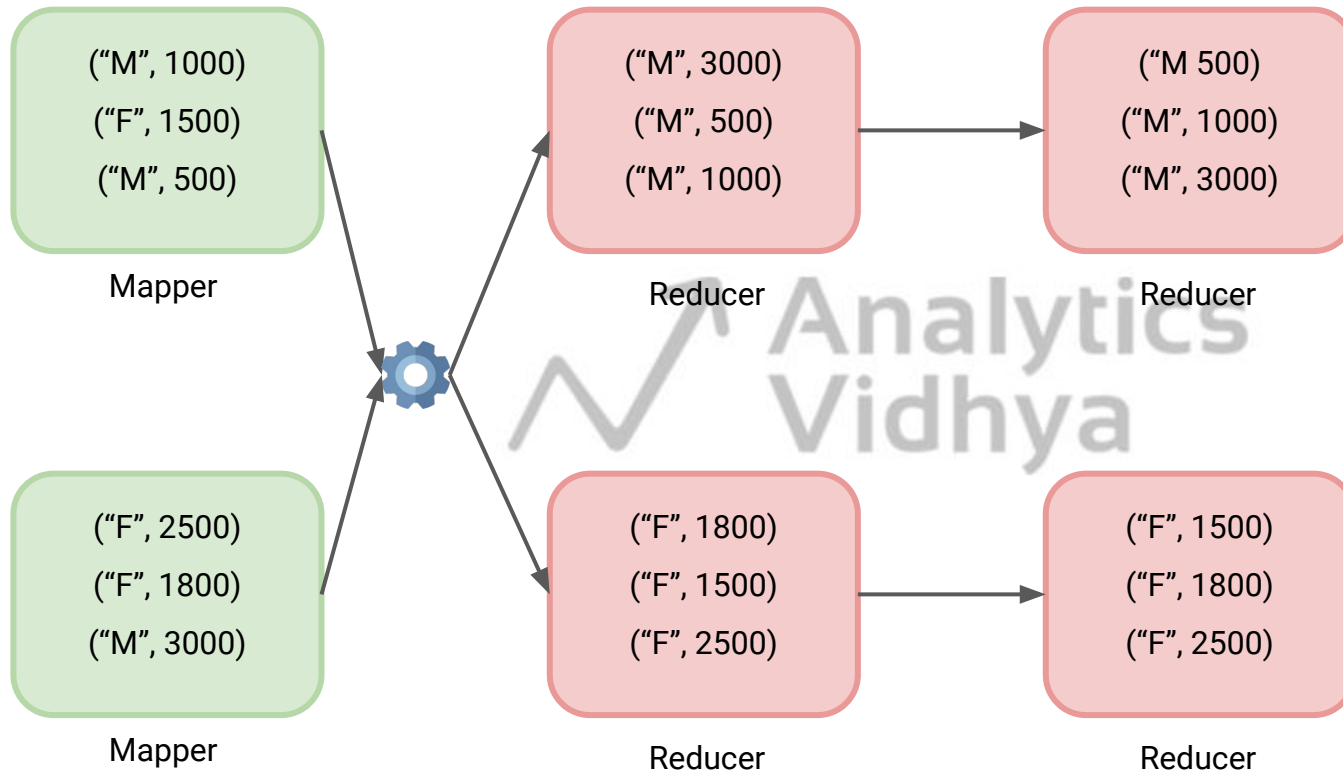
Reducer



Analytics
Vidhya

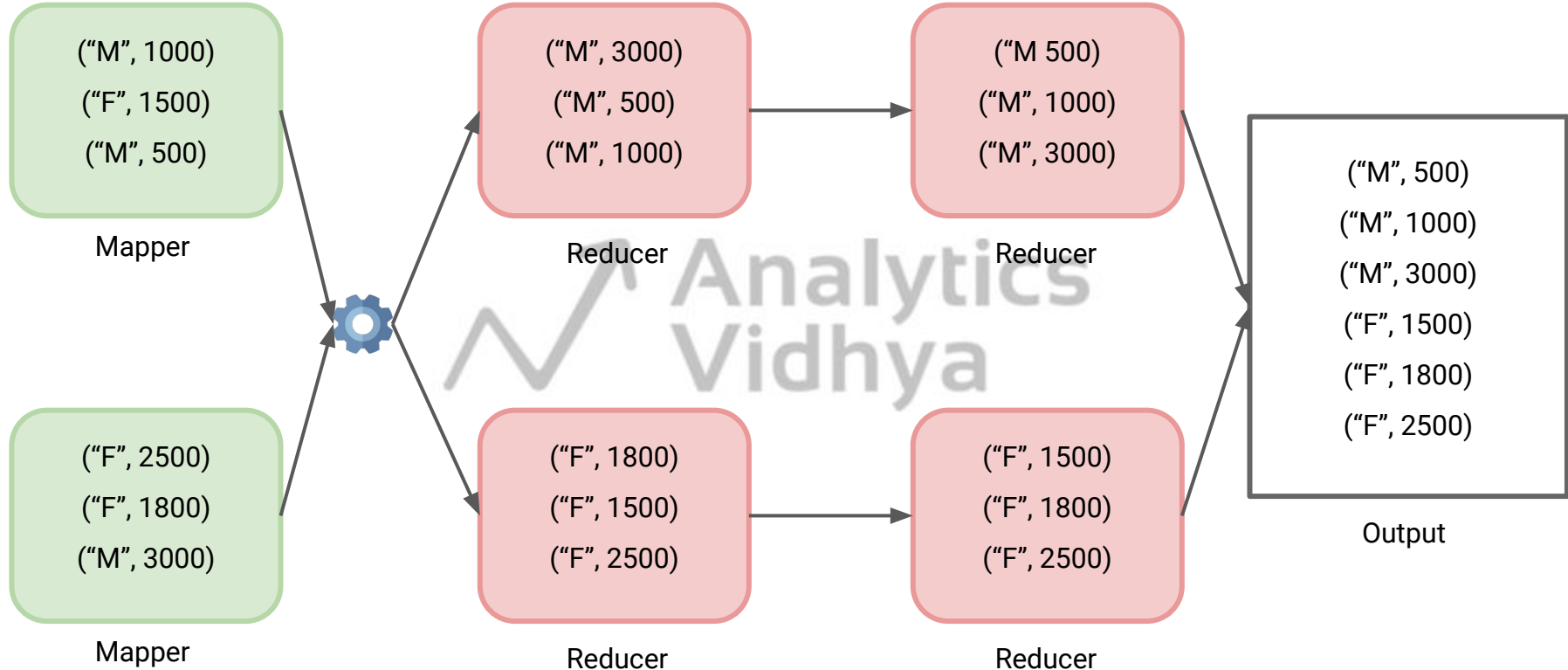
DISTRIBUTE BY
Clause

SORT BY Clause



DISTRIBUTE BY
Clause

SORT BY Clause





Thank You!