

Overview of data manipulation techniques

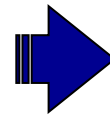
Data manipulation is the most intensive part of a project

- Data manipulation usually takes 80% of the time of your entire project.
- So, avoid inefficient code as it...
 - ...will take longer to execute.
 - ...is more likely to have bugs or mistakes.

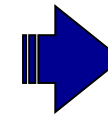
Three basic data operations

Select

Customer	TransDate
149332	2005-11-15
172951	2008-08-29



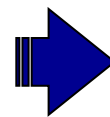
Select rows where
TransDate < 2007



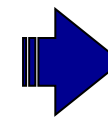
Customer	TransDate
149332	2005-11-15

Aggregate

Customer	Quantity
149332	1
172951	1
149332	1



Sum up Quantity
for each customer



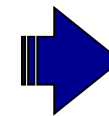
Customer	Quantity
149332	2
172951	1

Merge

Customer	TransDate
149332	2005-11-15
172951	2008-08-29

+

Customer	Quantity
149332	1
172951	1

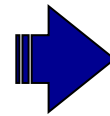


Customer	TransDate	Quantity
149332	2005-11-15	1
172951	2008-08-29	1

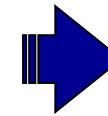
Three basic data operations

Select

Customer	TransDate
149332	2005-11-15
172951	2008-08-29



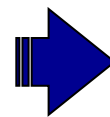
Select rows where
TransDate < 2007



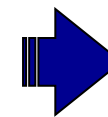
Customer	TransDate
149332	2005-11-15

Aggregate

Customer	Quantity
149332	1
172951	1
149332	1



Sum up Quantity
for each customer



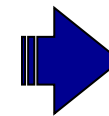
Customer	Quantity
149332	2
172951	1

Merge

Customer	TransDate
149332	2005-11-15
172951	2008-08-29

+

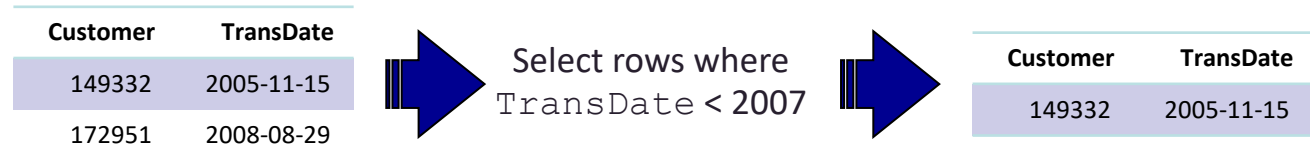
Customer	Quantity
149332	1
172951	1



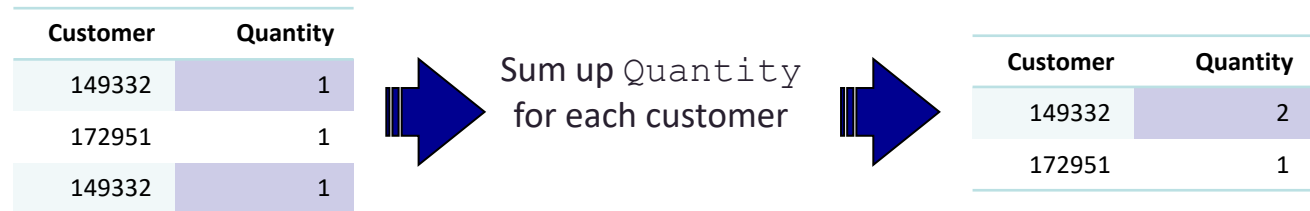
Customer	TransDate	Quantity
149332	2005-11-15	1
172951	2008-08-29	1

Three basic data operations

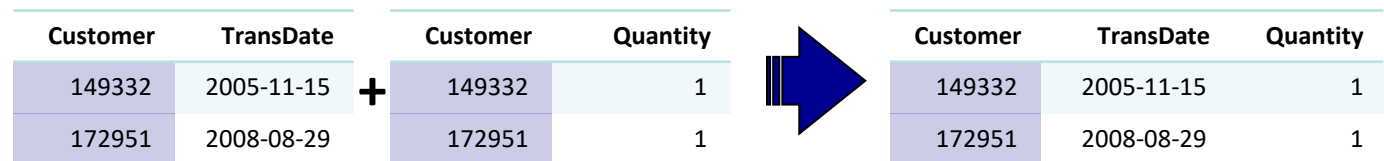
■ Select



■ Aggregate

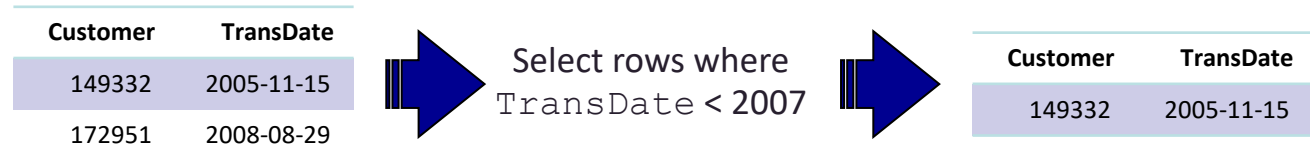


■ Merge

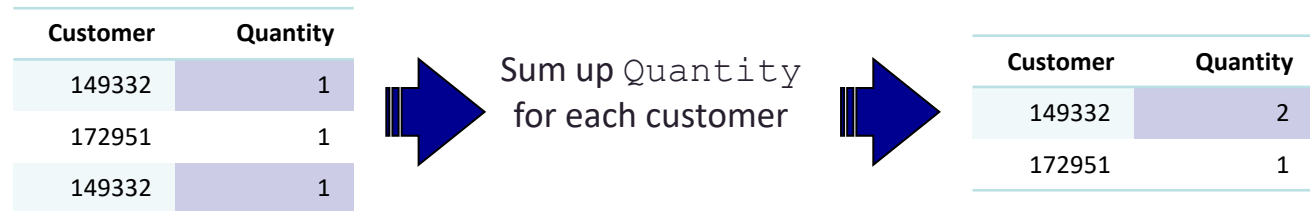


Three basic data operations

■ Select



■ Aggregate



■ Merge

