

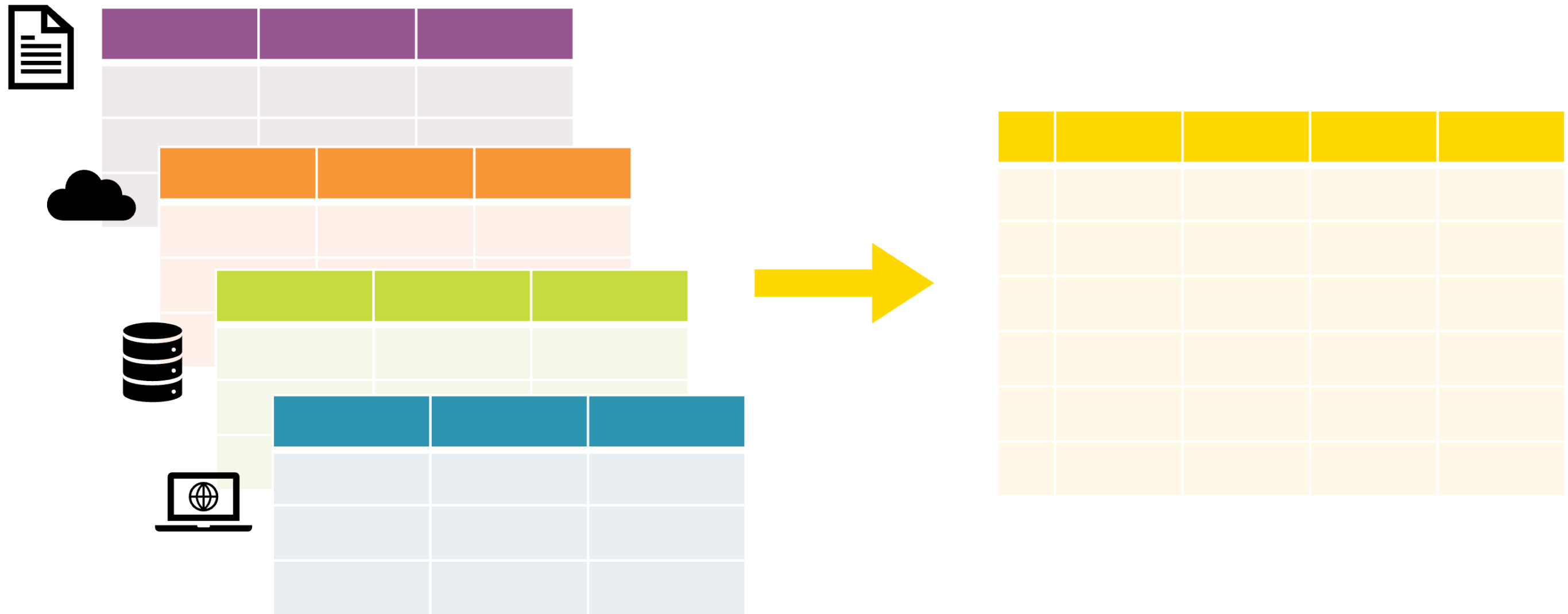
Mix and squeeze data

INTRODUCTION TO KNIME



Emilio Silvestri
Data Scientist

Data merging



Value lookup

#	EmployeeID	Name	Department	Favorite food
1	EMP-43	Jenny	Sales	Pizza
2	EMP-29	Alex	Finance	Sushi
3	EMP-11	Taylor	R&D	Tacos
4	EMP-29	Alex	Finance	Sushi
5	EMP-66	Sam	R&D	Salad

Value lookup

#	EmployeeID	Name	Department	Favorite food
1	EMP-43	Jenny	Sales	Pizza
2	EMP-29	Alex	Finance	Sushi
3	EMP-11	Taylor	R&D	Tacos
4	EMP-29	Alex	Finance	Sushi
5	EMP-66	Sam	R&D	Salad

#	ID	Salary
1	EMP-11	25000
2	EMP-29	22000
3	EMP-43	28000
4	EMP-88	22000

Value lookup

#	EmployeeID	Name	Department	Favorite food
1	EMP-43	Jenny	Sales	Pizza
2	EMP-29	Alex	Finance	Sushi
3	EMP-11	Taylor	R&D	Tacos
4	EMP-29	Alex	Finance	Sushi
5	EMP-66	Sam	R&D	Salad

#	ID	Salary
1	EMP-11	25000
2	EMP-29	22000
3	EMP-43	28000
4	EMP-88	22000

Value lookup

#	EmployeeID	Name	Department	Favorite food	Salary
1	EMP-43	Jenny	Sales	Pizza	
2	EMP-29	Alex	Finance	Sushi	
3	EMP-11	Taylor	R&D	Tacos	
4	EMP-29	Alex	Finance	Sushi	
5	EMP-66	Sam	R&D	Salad	

#	ID	Salary
1	EMP-11	25000
2	EMP-29	22000
3	EMP-43	28000
4	EMP-88	22000

Value lookup

#	EmployeeID	Name	Department	Favorite food	Salary
1	EMP-43	Jenny	Sales	Pizza	
2	EMP-29	Alex	Finance	Sushi	
3	EMP-11	Taylor	R&D	Tacos	
4	EMP-29	Alex	Finance	Sushi	
5	EMP-66	Sam	R&D	Salad	

#	ID	Salary
1	EMP-11	25000
2	EMP-29	22000
3	EMP-43	28000
4	EMP-88	22000

Value lookup

#	EmployeeID	Name	Department	Favorite food	Salary
1	EMP-43	Jenny	Sales	Pizza	28000
2	EMP-29	Alex	Finance	Sushi	
3	EMP-11	Taylor	R&D	Tacos	
4	EMP-29	Alex	Finance	Sushi	
5	EMP-66	Sam	R&D	Salad	

#	ID	Salary
1	EMP-11	25000
2	EMP-29	22000
3	EMP-43	28000
4	EMP-88	22000

Value lookup



Value lookup

#	EmployeeID	Name	Department	Favorite food	Salary
1	EMP-43	Jenny	Sales	Pizza	28000
2	EMP-29	Alex	Finance	Sushi	22000
3	EMP-11	Taylor	R&D	Tacos	25000
4	EMP-29	Alex	Finance	Sushi	22000
5	EMP-66	Sam	R&D	Salad	

#	ID	Salary
1	EMP-11	25000
2	EMP-29	22000
3	EMP-43	28000
4	EMP-88	22000

Dictionary

Value lookup

#	EmployeeID	Name	Department	Favorite food	Salary
1	EMP-43	Jenny	Sales	Pizza	28000
2	EMP-29	Alex	Finance	Sushi	22000
3	EMP-11	Taylor	R&D	Tacos	25000
4	EMP-29	Alex	Finance	Sushi	22000
5	EMP-66	Sam	R&D	Salad	

Lookup column	#	ID	Salary	Dictionary
	1	EMP-11	25000	
	2	EMP-29	22000	
	3	EMP-43	28000	
Key column	4	EMP-88	22000	

Value lookup

#	EmployeeID	Name	Department	Favorite food	Salary
1	EMP-43	Jenny	Sales	Pizza	28000
2	EMP-29	Alex	Finance	Sushi	22000
3	EMP-11	Taylor	R&D	Tacos	25000
4	EMP-29	Alex	Finance	Sushi	22000
5	EMP-66	Sam	R&D	Salad	?

Lookup column	#	ID	Salary	Dictionary
	1	EMP-11	25000	
	2	EMP-29	22000	
	3	EMP-43	28000	
	4	EMP-88	22000	

Key
column

Data aggregation



Data aggregation



Data aggregation



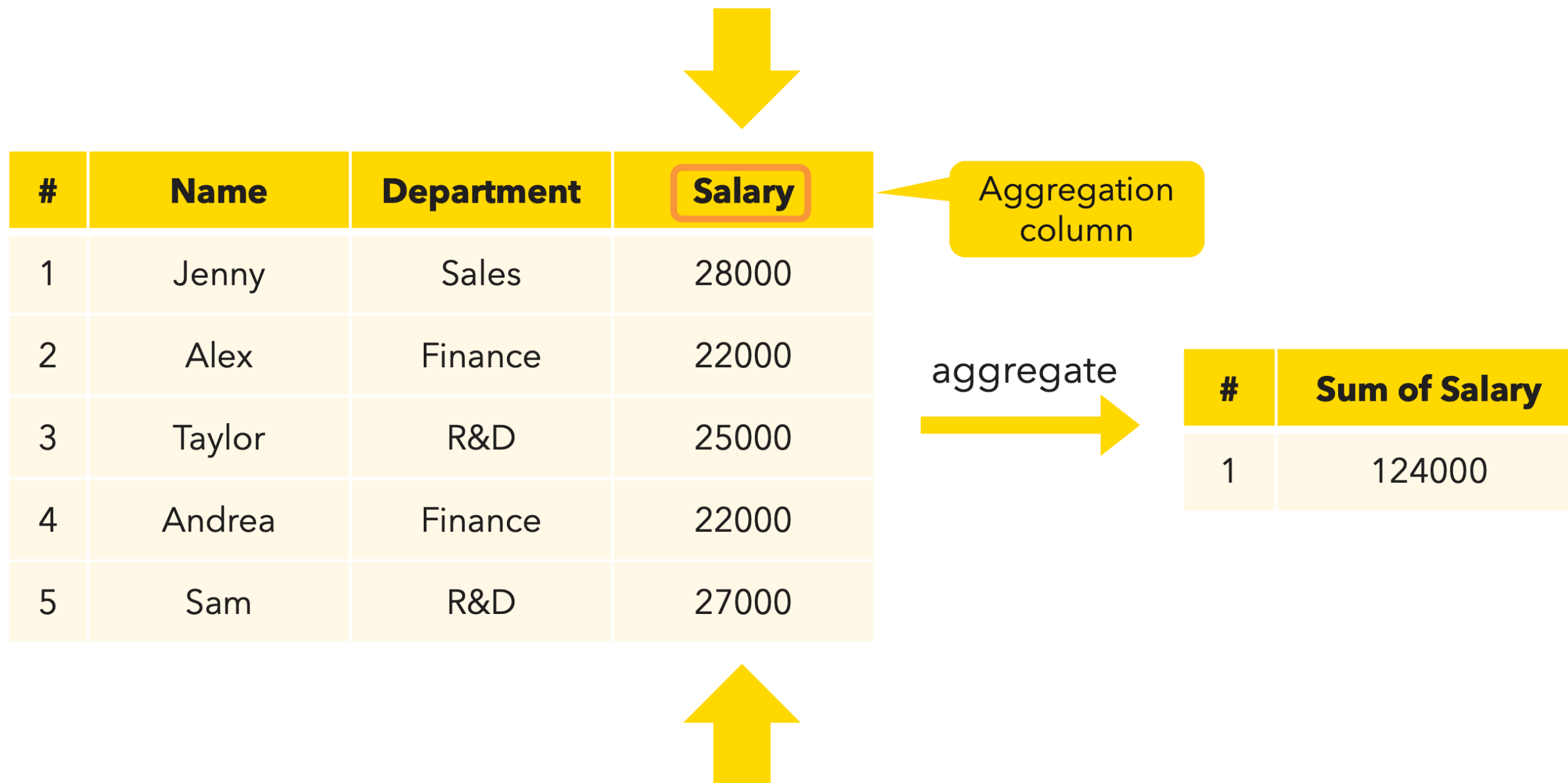
#	Name	Department	Salary
1	Jenny	Sales	28000
2	Alex	Finance	22000
3	Taylor	R&D	25000
4	Andrea	Finance	22000
5	Sam	R&D	27000

aggregate

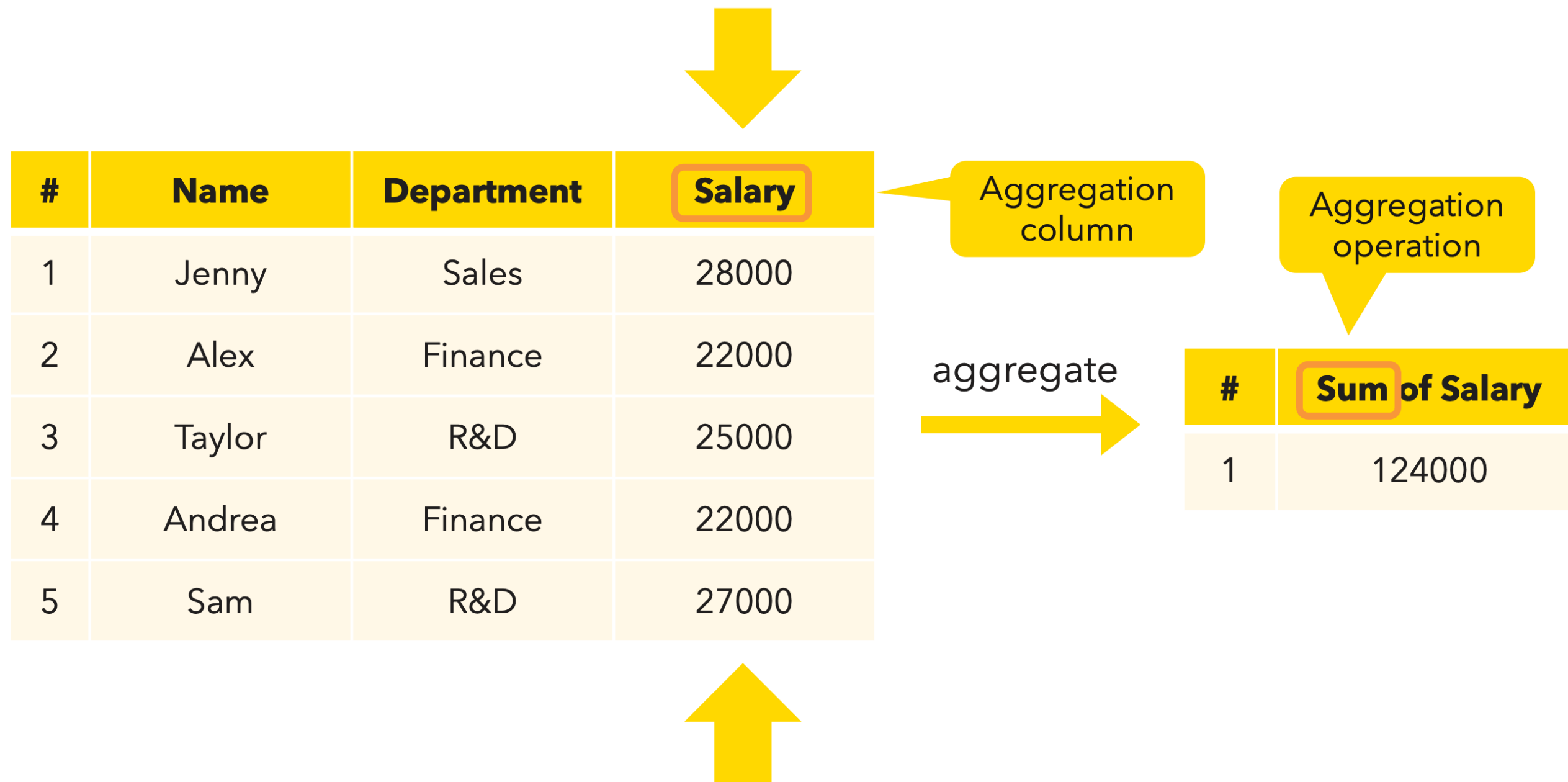
#	Sum of Salary
1	124000



Data aggregation



Data aggregation



Data aggregation

Category
column


#	Name	Department	Salary
1	Jenny	Sales	28000
2	Alex	Finance	22000
3	Taylor	R&D	25000
4	Andrea	Finance	22000
5	Sam	R&D	27000

aggregate

#	Department	Sum of Salary
1	Sales	28000
2	Finance	44000
3	R&D	52000

Data aggregation


#	Name	Department	Salary
1	Jenny	Sales	28000
2	Alex	Finance	22000
3	Taylor	R&D	25000
4	Andrea	Finance	22000
5	Sam	R&D	27000

aggregate


#	Department	Number of employees
1	Sales	1
2	Finance	2
3	R&D	2

Data aggregation

#	Name	Department	Salary
1	Jenny	Sales	28000
2	Alex	Finance	22000
3	Taylor	R&D	25000
4	Andrea	Finance	22000
5	Sam	R&D	27000

aggregate


#	Department	Average Salary
1	Sales	28000
2	Finance	22000
3	R&D	26000

Documentation



Documentation

Calculate average income

Access

Read employees dataset and income

Excel Reader



Read employees dataset

SQLite Connector



hr_data.sqlite

DB Table Selector



income table

DB Reader



Execute query

Clean

Filter and remove duplicate

Row Filter



Filter employees

Duplicate Row Filter



Remove duplicate

Column Filter



Filter columns

Merge and aggregate

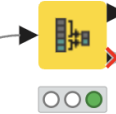
Append income and calculate average

Value Lookup



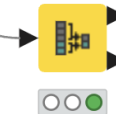
Append income

Row Aggregator



Average income in the company

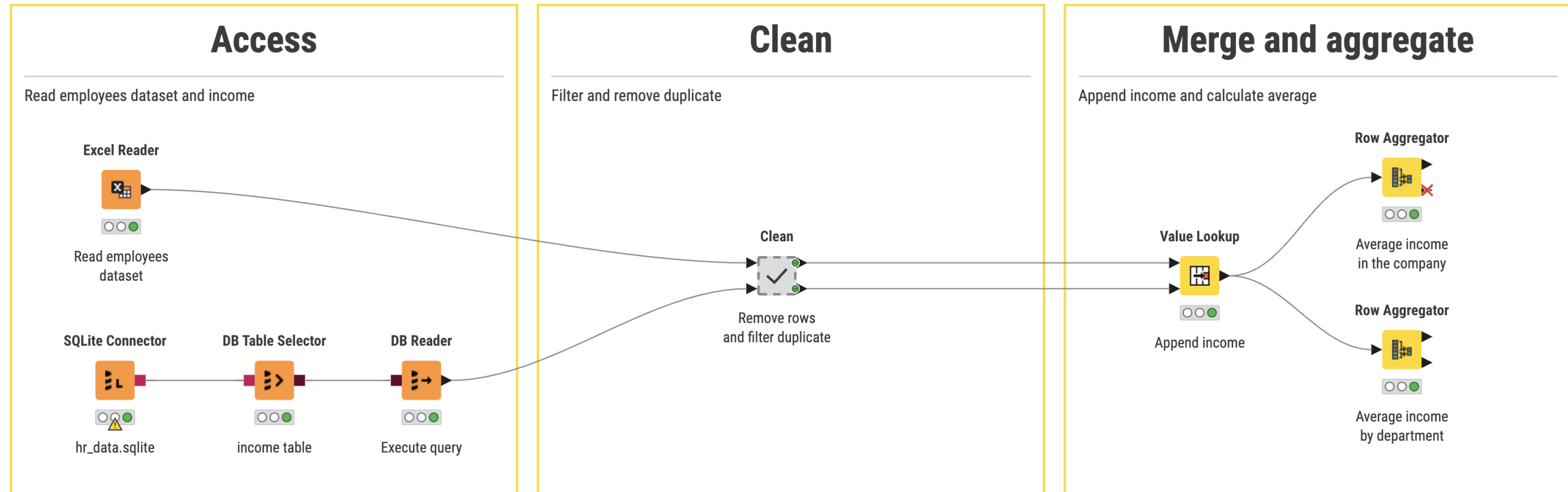
Row Aggregator



Average income by department

Metanodes

Calculate average income



Let's practice!

INTRODUCTION TO KNIME

Crunching data in KNIME Analytics Platform

INTRODUCTION TO KNIME



Emilio Silvestri
Data Scientist

Let's practice!

INTRODUCTION TO KNIME

Congratulations!

INTRODUCTION TO KNIME



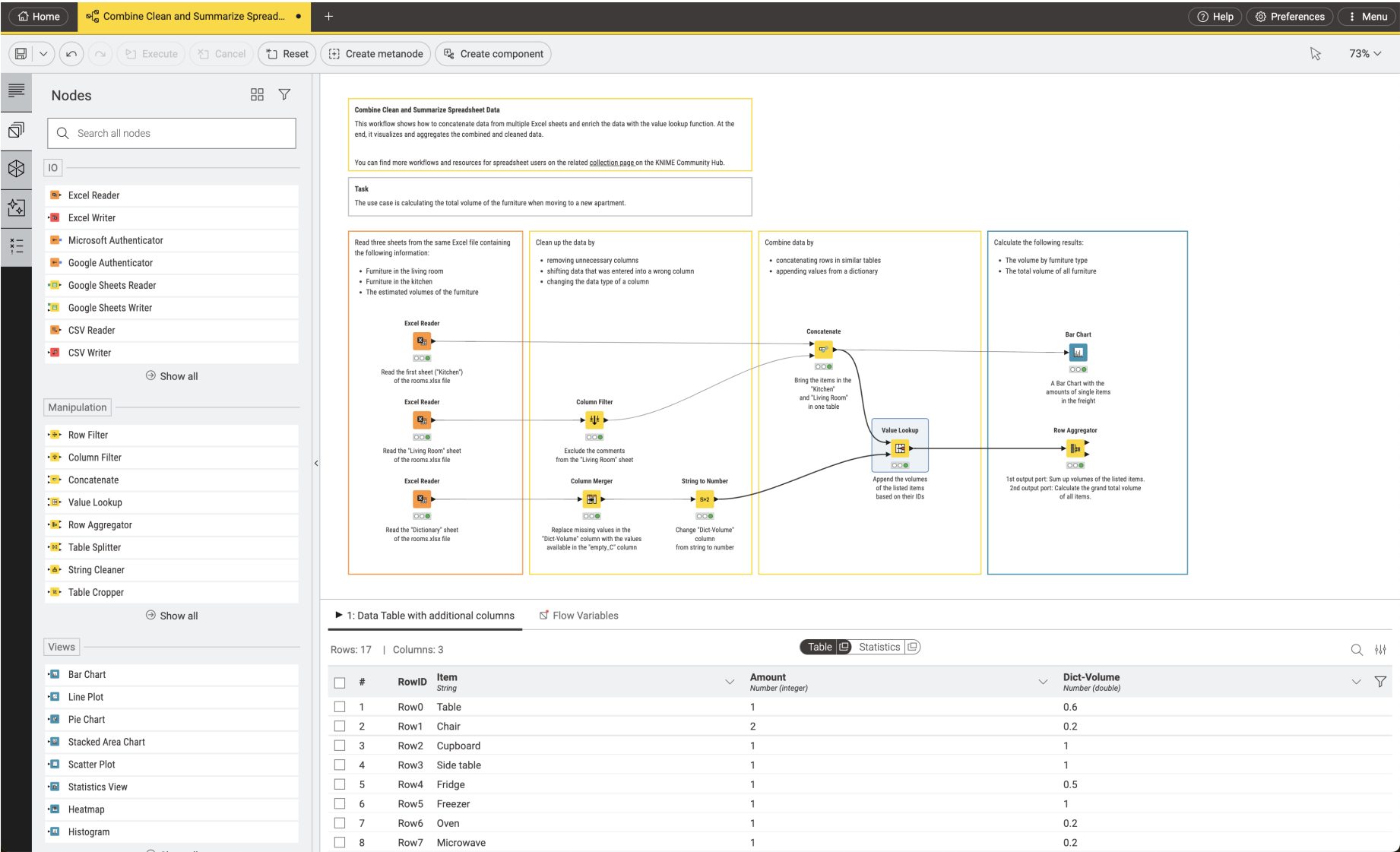
Emilio Silvestri

Data Scientist, KNIME

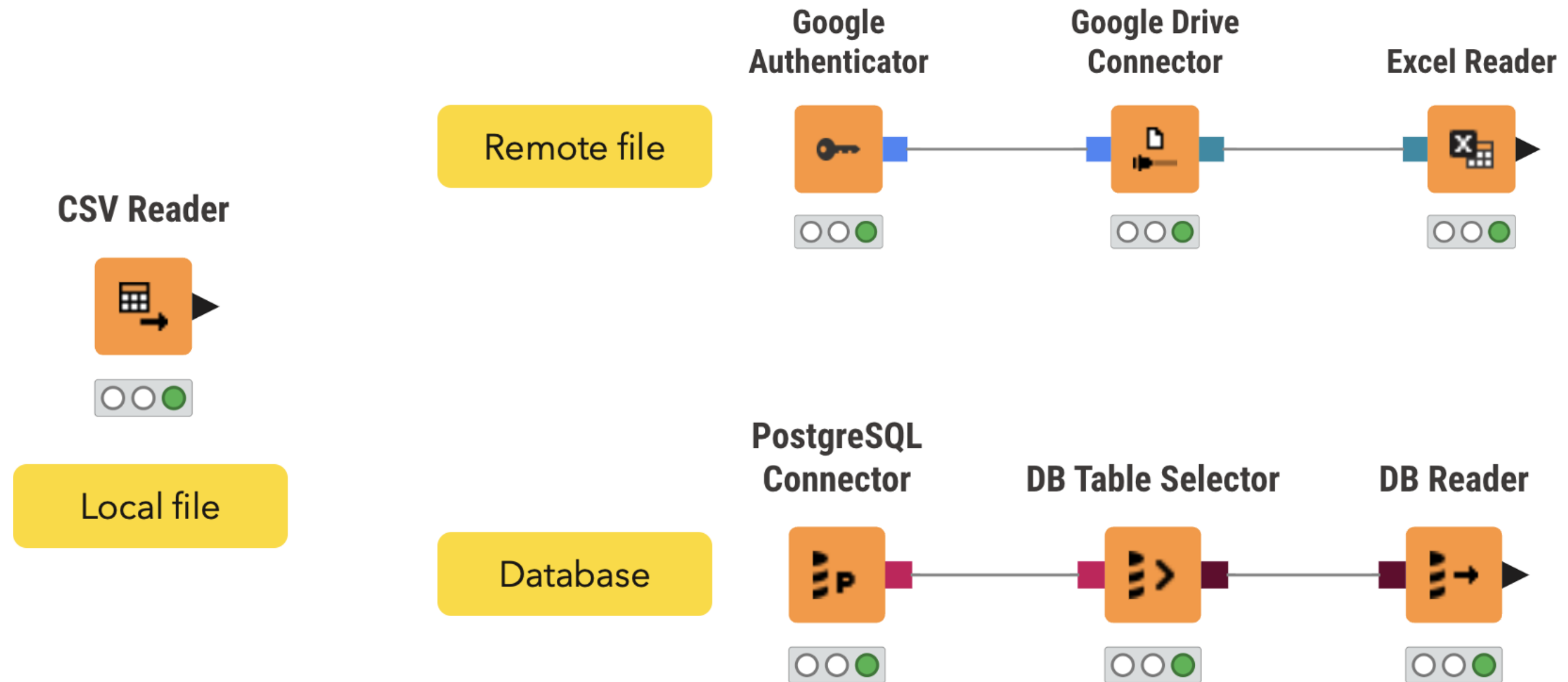
Recap



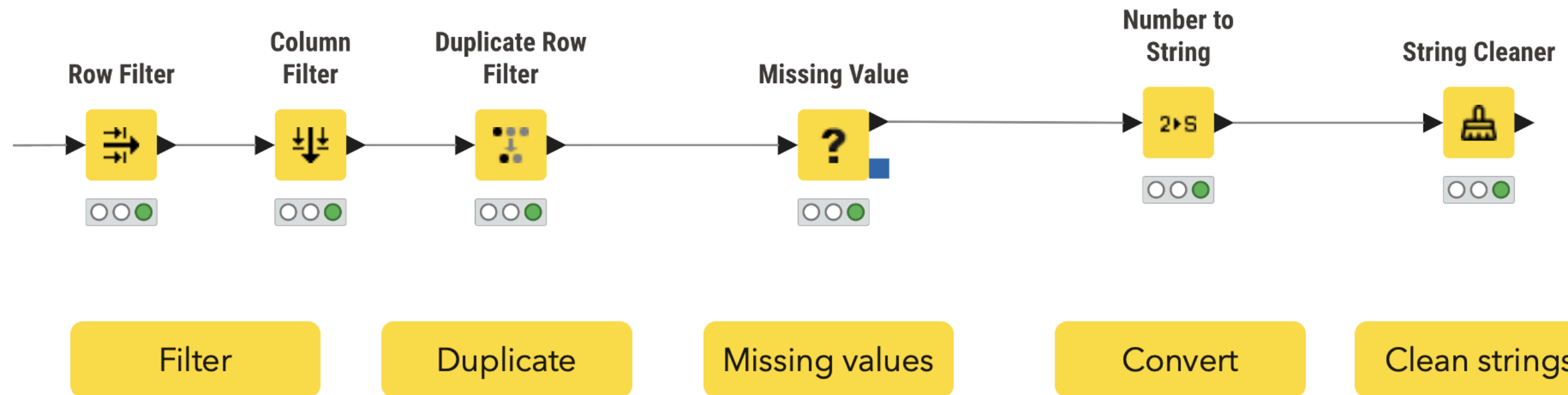
KNIME Analytics Platform



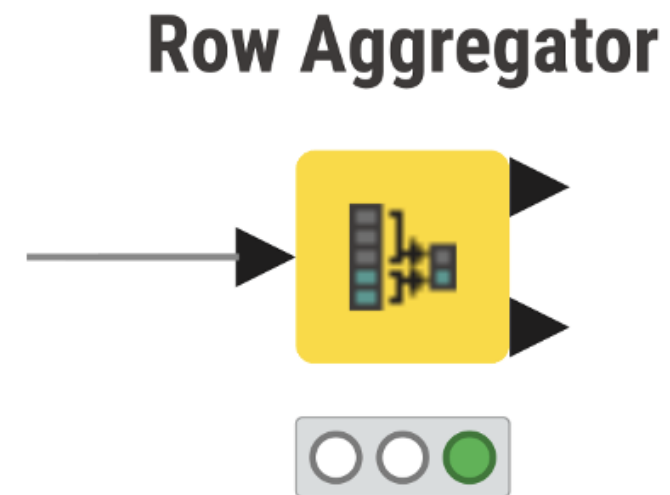
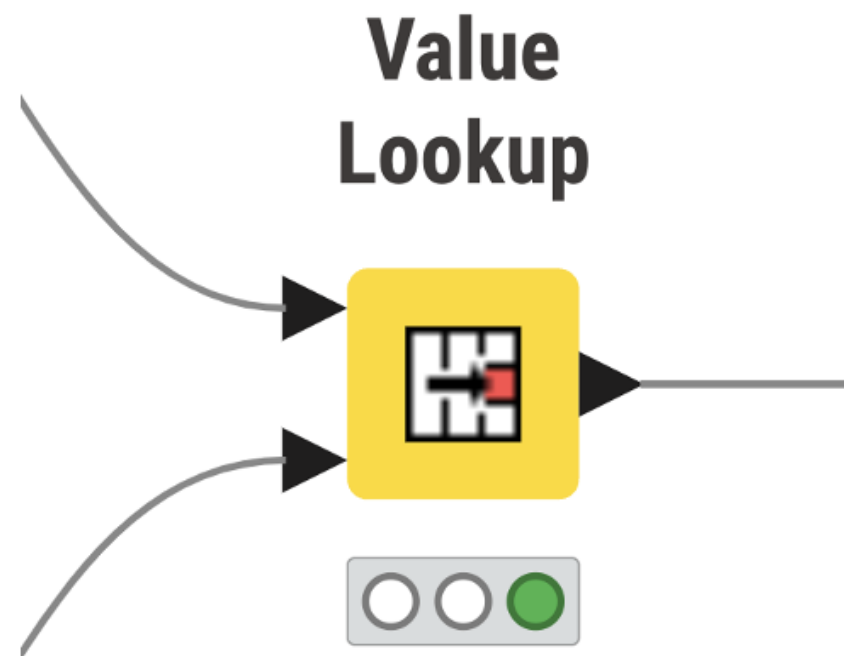
Data access



Data cleaning



Data merging and aggregation



More to come



Congratulations!

INTRODUCTION TO KNIME