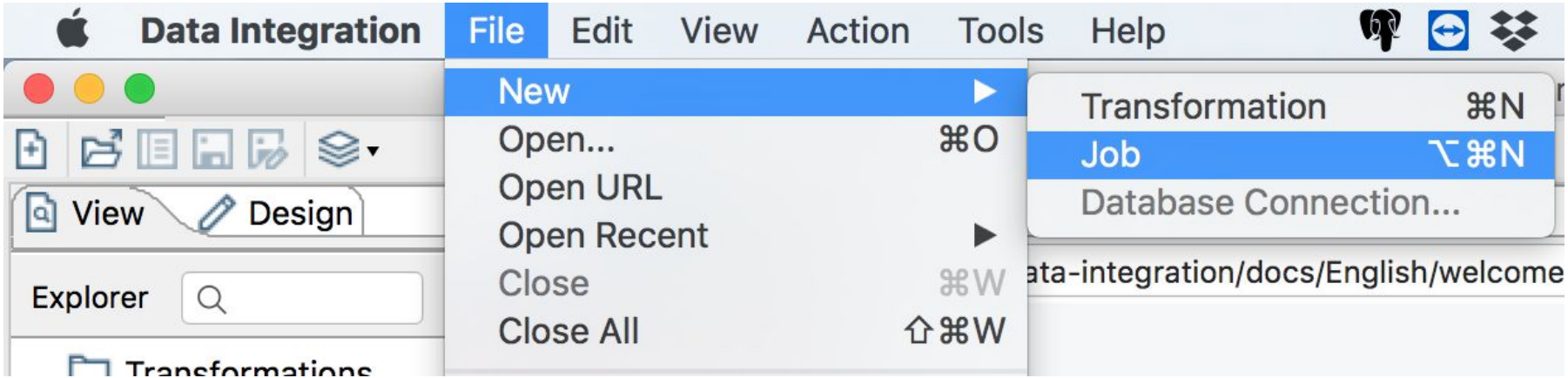
Building Sales Data mart Using Pentaho Part 3

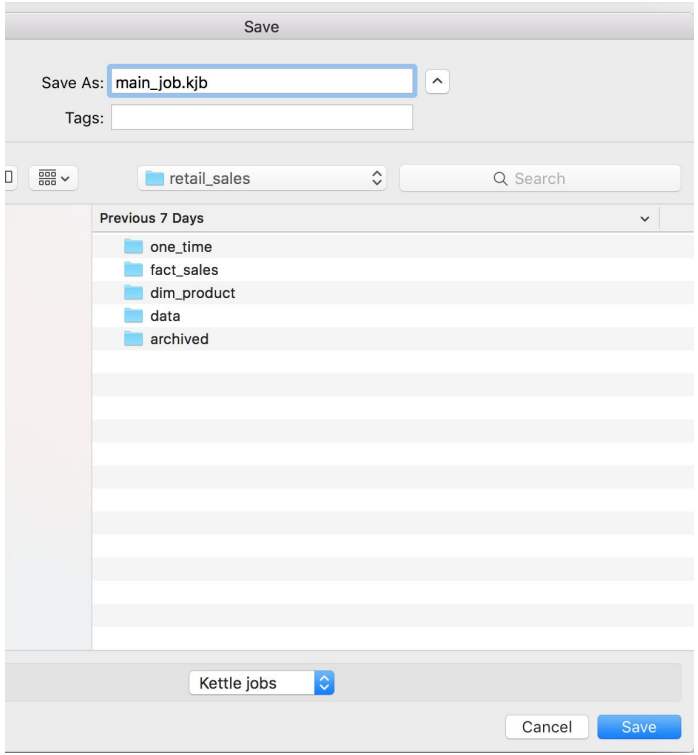
By Naheed Anjum Arafat

Create Main Job to run all Tasks

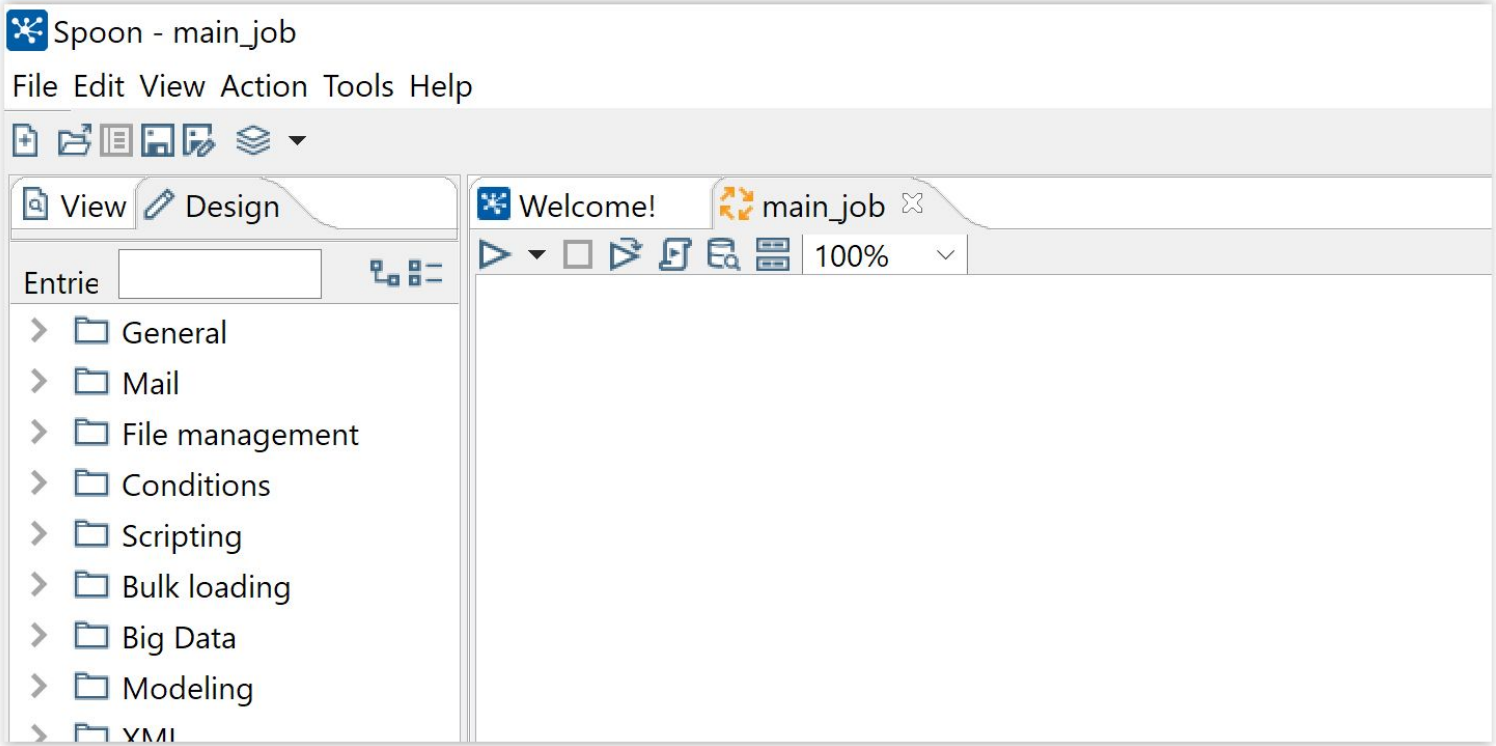
Main job


Data Integration
File
Edit
View
Action
Tools
Help
New
Open...
⌘O
Open URL
Open Recent
Close
⌘W
Close All
⇧⌘W
Transformation
⌘N
Job
⌥⌘N
Database Connection...
View
Design
Explorer
ata-integration/docs/English/welcome

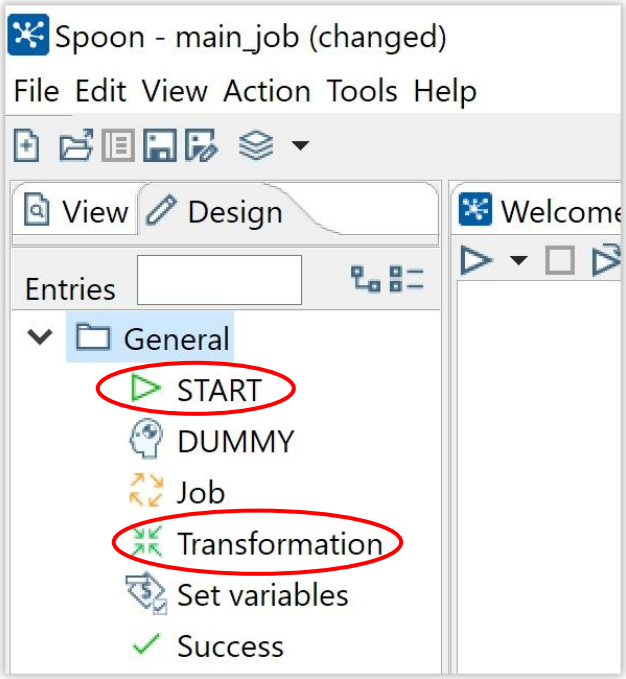
Save

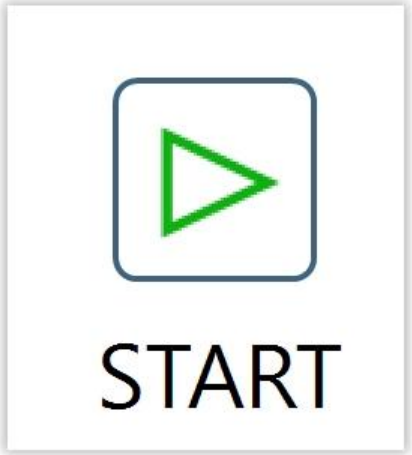

Save
Save As: main_job.kjb
Tags:
retail_sales
Search
Previous 7 Days
one_time
fact_sales
dim_product
data
archived
Kettle jobs
Cancel
Save

Main job

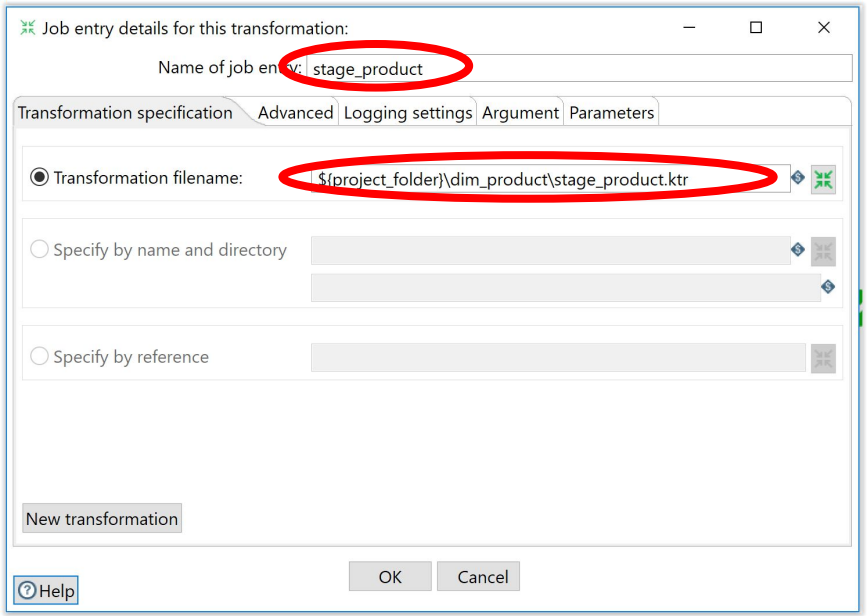

Spoon - main_job
File Edit View Action Tools Help
View
Design
Entrie
General
Mail
File management
Conditions
Scripting
Bulk loading
Big Data
Modeling
Welcome!
main_job
100%

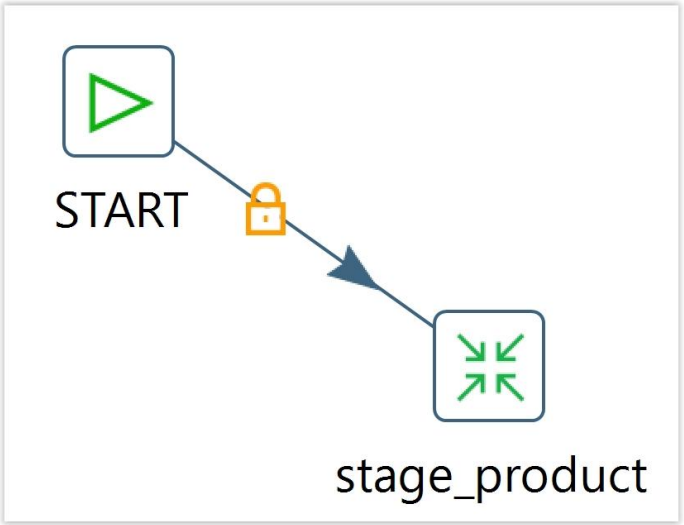
Main Job


Spoon - main_job (changed)
File Edit View Action Tools Help
View
Design
Welcome
Entries
General
START
DUMMY
Job
Transformation
Set variables
Success

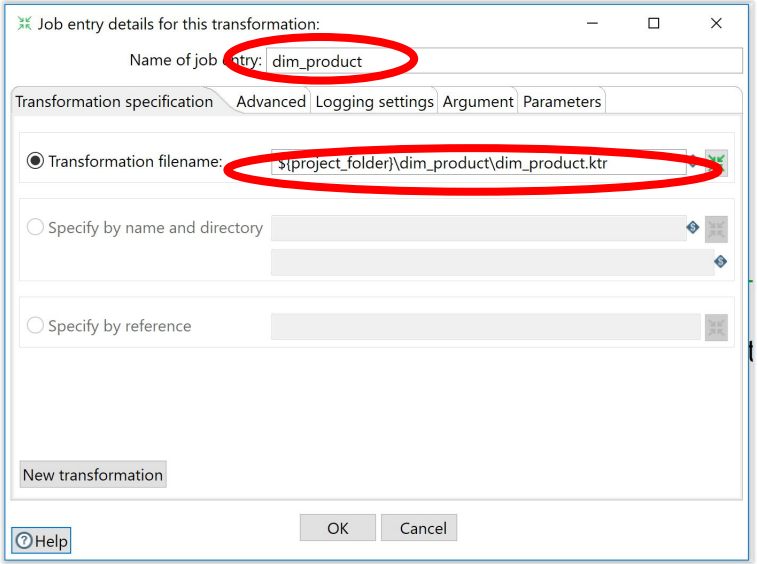

START

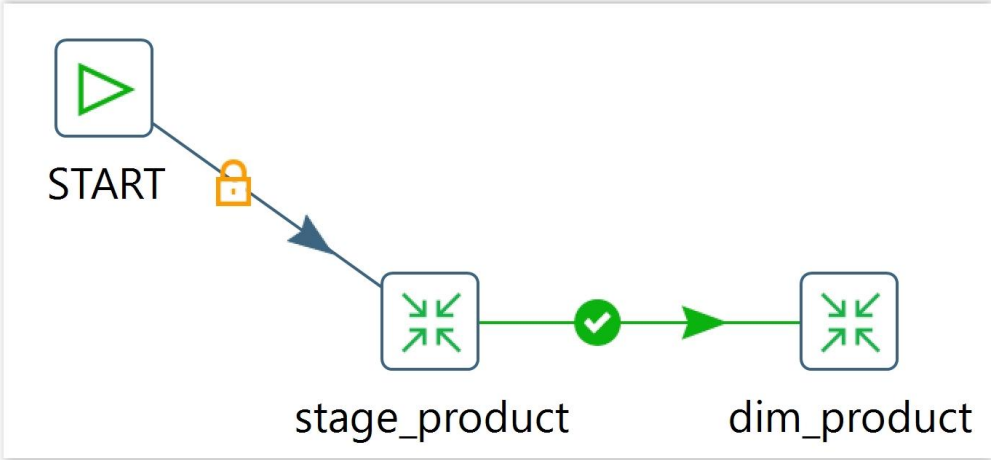
stage_product



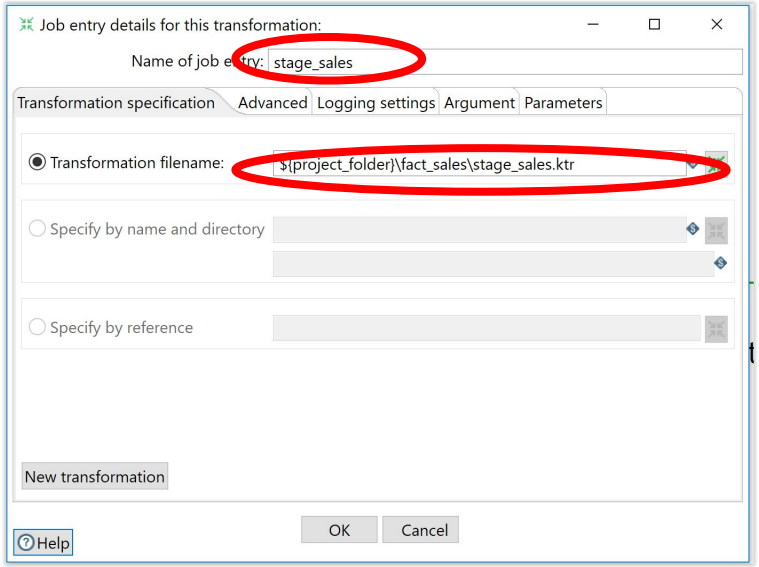


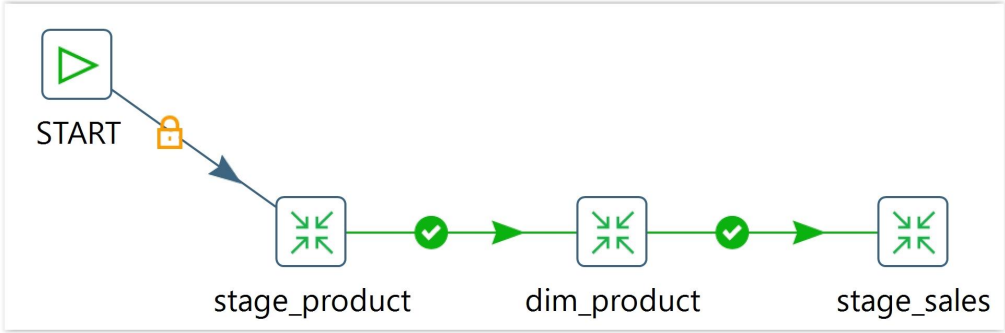
dim_product


Job entry details for this transformation:
Name of job entry: dim_product
Transformation specification
Advanced
Logging settings
Argument
Parameters
Transformation filename:
${project_folder}\dim_product\dim_product.ktr
Specify by name and directory
Specify by reference
New transformation
Help
OK
Cancel

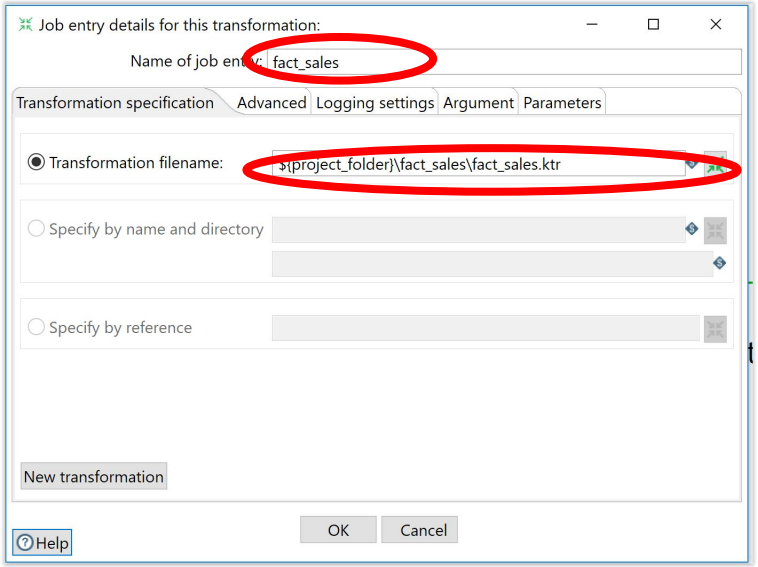

START
stage_product
dim_product

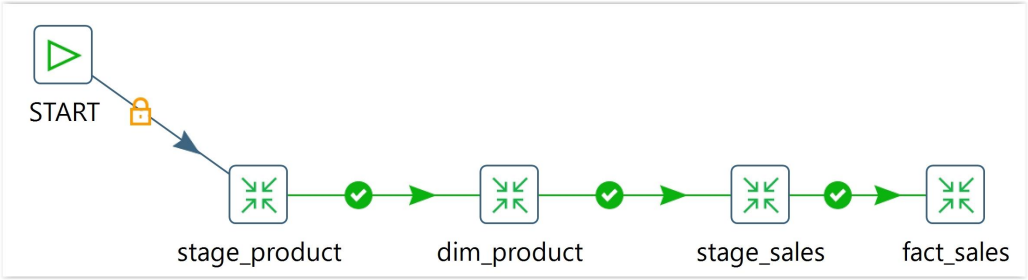
stage_sales


Job entry details for this transformation:
Name of job entry: stage_sales
Transformation specification
Advanced
Logging settings
Argument
Parameters
Transformation filename:
${project_folder}\fact_sales\stage_sales.ktr
Specify by name and directory
Specify by reference
New transformation
Help
OK
Cancel


START
stage_product
dim_product
stage_sales

fact_sales


Job entry details for this transformation:
Name of job entry: fact_sales
Transformation specification
Advanced
Logging settings
Argument
Parameters
Transformation filename:
${project_folder}\fact_sales\fact_sales.ktr
Specify by name and directory
Specify by reference
New transformation
Help
OK
Cancel


START
stage_product
dim_product
stage_sales
fact_sales

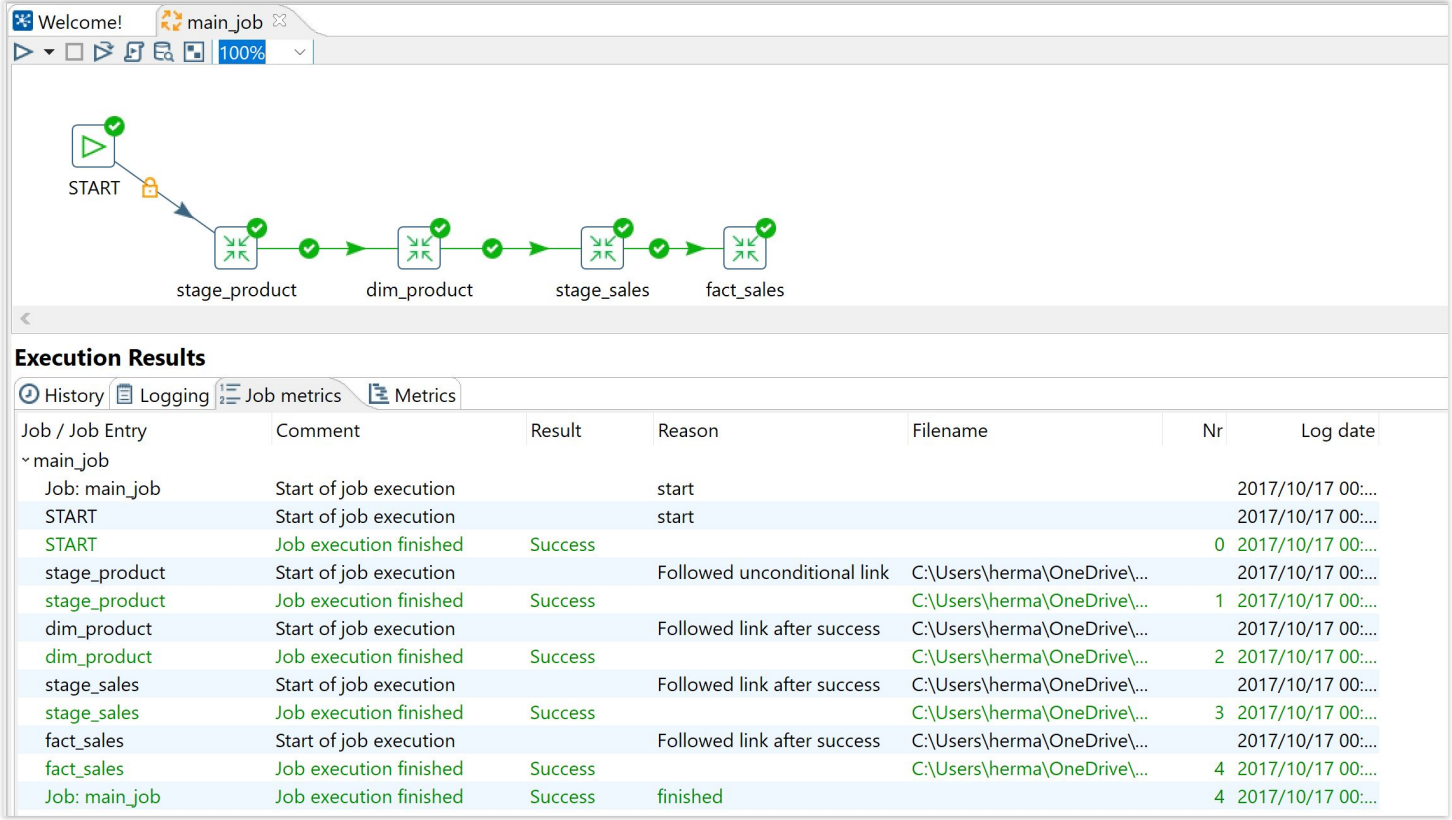
Before you run Main job

- Before you create and run a job to streamline the dimension, recall that, running step 2.2 twice created duplicate entries

- In practice, we make sure that, we never process same pos files twice.
  - How? By moving the processed files to an archive directory.
  - More on this topic : https://wiki.pentaho.com/display/EAI/Move+files

For now, Truncate the fact_sales manually (using pgadmin / spoon database explorer), before you run the job.

Run main_job.kjb



Execution Results

History | Logging | Job metrics | Metrics

Job / Job Entry	Comment	Result	Reason	Filename	Nr	Log date
main_job						
Job: main_job	Start of job execution		start			2017/10/17 00:...
START	Start of job execution		start			2017/10/17 00:...
START	Job execution finished	Success			0	2017/10/17 00:...
stage_product	Start of job execution		Followed unconditional link	C:\Users\herma\OneDrive\...		2017/10/17 00:...
stage_product	Job execution finished	Success		C:\Users\herma\OneDrive\...	1	2017/10/17 00:...
dim_product	Start of job execution		Followed link after success	C:\Users\herma\OneDrive\...		2017/10/17 00:...
dim_product	Job execution finished	Success		C:\Users\herma\OneDrive\...	2	2017/10/17 00:...
stage_sales	Start of job execution		Followed link after success	C:\Users\herma\OneDrive\...		2017/10/17 00:...
stage_sales	Job execution finished	Success		C:\Users\herma\OneDrive\...	3	2017/10/17 00:...
fact_sales	Start of job execution		Followed link after success	C:\Users\herma\OneDrive\...		2017/10/17 00:...
fact_sales	Job execution finished	Success		C:\Users\herma\OneDrive\...	4	2017/10/17 00:...
Job: main_job	Job execution finished	Success	finished		4	2017/10/17 00:...

END

Acknowledgement:- Slides adapted from Herman Tan, guest lecturer, AY 2017/18