

Universidad Mayor de San Andrés  
Facultad de Ciencias Puras y Naturales

# Evaluación de participación 3

DAT 251 - Base de datos 3

## **Docente:**

Lic. Celia Elena Tarquino Peralta

## **Estudiante:**

Gabriel Muñoz Marcelo Callisaya

## **Fecha de entrega:**

6 de marzo del 2025

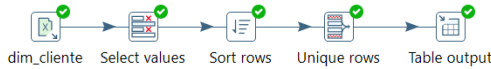
**La Paz - Bolivia**

# Hechos y dimensiones.

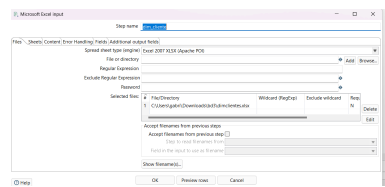
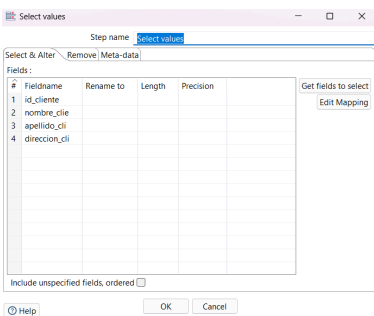
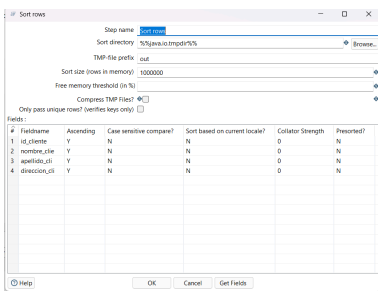
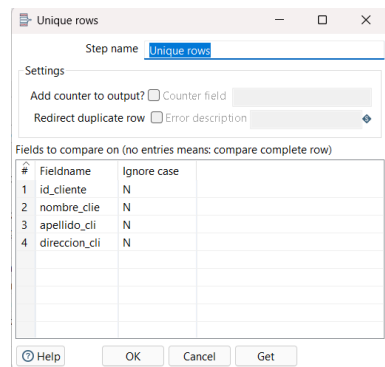
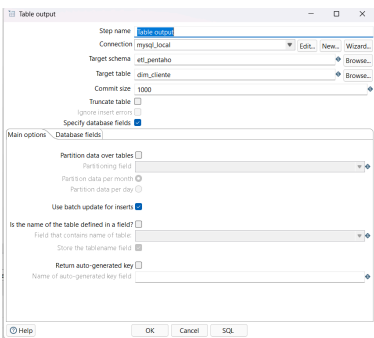
Las transformaciones para los hechos y dimensiones de dim\_cliente, dim\_producto, dim\_tiempo y hec\_venta:

## dim\_cliente

Su archivo es Transformation5-cliente.ktr

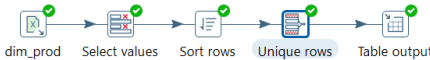
diagrama	salida																																																																	
<div></div> <p>dim_cliente    Select values    Sort rows    Unique rows    Table output</p>	<div><h3>Execution Results</h3><div><div>Logging</div><div>Execution History</div><div>Step Metrics</div><div>Performance Graph</div><div>Metrics</div><div>Preview data</div></div><div><div>First rows</div><div>Last rows</div><div>Off</div></div><table><thead><tr><th>#</th><th>id_cliente</th><th>nombre_cliente</th><th>apellido_cliente</th><th>direccion_cliente</th></tr></thead><tbody><tr><td>1</td><td>1.0</td><td>Juan</td><td>Pérez</td><td>Calle 123 #45-67</td></tr><tr><td>2</td><td>2.0</td><td>Maria</td><td>Gómez</td><td>Av. Central 89</td></tr><tr><td>3</td><td>3.0</td><td>Pedro</td><td>Rodríguez</td><td>Cra 10 #22-33</td></tr><tr><td>4</td><td>4.0</td><td>Carlos</td><td>López</td><td>Calle 89 #90-12</td></tr><tr><td>5</td><td>5.0</td><td>Ana</td><td>Martínez</td><td>Av. 7 #56-78</td></tr><tr><td>6</td><td>6.0</td><td>Luis</td><td>Hernández</td><td>Cra 15 #67-89</td></tr><tr><td>7</td><td>7.0</td><td>Sofia</td><td>García</td><td>Av. 20 #34-56</td></tr><tr><td>8</td><td>8.0</td><td>Juan</td><td>Gómez</td><td>Calle 45 #12-34</td></tr><tr><td>9</td><td>9.0</td><td>Maria</td><td>Pérez</td><td>Av. 9 #45-67</td></tr><tr><td>1.</td><td>10.0</td><td>Pedro</td><td>López</td><td>Cra 30 #12-34</td></tr><tr><td>1.</td><td>11.0</td><td>Ana</td><td>Rodríguez</td><td>Calle 50 #78-90</td></tr><tr><td>1.</td><td>12.0</td><td>Carlos</td><td>Martínez</td><td>Av. 15 #67-89</td></tr></tbody></table></div>	#	id_cliente	nombre_cliente	apellido_cliente	direccion_cliente	1	1.0	Juan	Pérez	Calle 123 #45-67	2	2.0	Maria	Gómez	Av. Central 89	3	3.0	Pedro	Rodríguez	Cra 10 #22-33	4	4.0	Carlos	López	Calle 89 #90-12	5	5.0	Ana	Martínez	Av. 7 #56-78	6	6.0	Luis	Hernández	Cra 15 #67-89	7	7.0	Sofia	García	Av. 20 #34-56	8	8.0	Juan	Gómez	Calle 45 #12-34	9	9.0	Maria	Pérez	Av. 9 #45-67	1.	10.0	Pedro	López	Cra 30 #12-34	1.	11.0	Ana	Rodríguez	Calle 50 #78-90	1.	12.0	Carlos	Martínez	Av. 15 #67-89
#	id_cliente	nombre_cliente	apellido_cliente	direccion_cliente																																																														
1	1.0	Juan	Pérez	Calle 123 #45-67																																																														
2	2.0	Maria	Gómez	Av. Central 89																																																														
3	3.0	Pedro	Rodríguez	Cra 10 #22-33																																																														
4	4.0	Carlos	López	Calle 89 #90-12																																																														
5	5.0	Ana	Martínez	Av. 7 #56-78																																																														
6	6.0	Luis	Hernández	Cra 15 #67-89																																																														
7	7.0	Sofia	García	Av. 20 #34-56																																																														
8	8.0	Juan	Gómez	Calle 45 #12-34																																																														
9	9.0	Maria	Pérez	Av. 9 #45-67																																																														
1.	10.0	Pedro	López	Cra 30 #12-34																																																														
1.	11.0	Ana	Rodríguez	Calle 50 #78-90																																																														
1.	12.0	Carlos	Martínez	Av. 15 #67-89																																																														

Y las configuraciones de cada transformación:

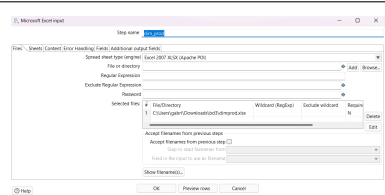
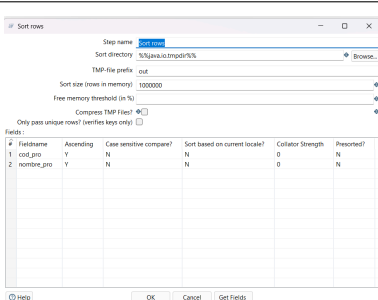
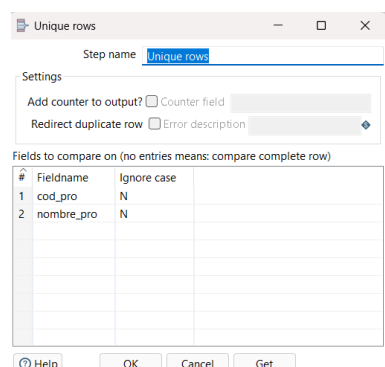
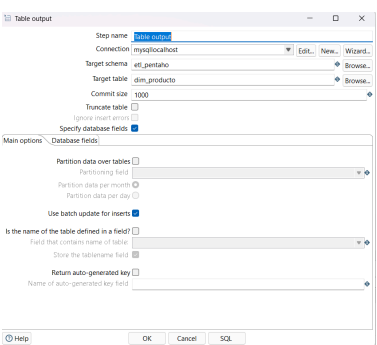
		
		

## dim\_producto

Su archivo es Transformation5-producto.ktr

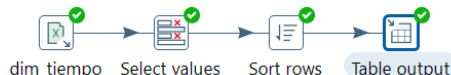
diagrama	salida																																	
 <pre>graph LR; A[dim_pro] --&gt; B[Select values]; B --&gt; C[Sort rows]; C --&gt; D[Unique rows]; D --&gt; E[Table output];</pre>	<div><b>Execution Results</b></div> <div><div><input checked="" type="radio"/> Logging</div><div><input type="radio"/> Execution History</div><div><input type="radio"/> Step Metrics</div><div><input type="radio"/> Performance Graph</div><div><input type="radio"/> Metrics</div><div><input type="radio"/> Preview data</div></div> <div><div><input checked="" type="radio"/> First rows</div><div><input type="radio"/> Last rows</div><div><input type="radio"/> Off</div></div> <table><tr><th>#</th><th>cod_pro</th><th>nombre_pro</th></tr><tr><td>1</td><td>PROD001</td><td>Camisa Azul</td></tr><tr><td>2</td><td>PROD002</td><td>Pantalón Negro</td></tr><tr><td>3</td><td>PROD003</td><td>Chaqueta Gris</td></tr><tr><td>4</td><td>PROD004</td><td>Sudadera Roja</td></tr><tr><td>5</td><td>PROD005</td><td>Vestido Floral</td></tr><tr><td>6</td><td>PROD006</td><td>Zapatos Negros</td></tr><tr><td>7</td><td>PROD007</td><td>Blusa Blanca</td></tr><tr><td>8</td><td>PROD008</td><td>Falda Negra</td></tr><tr><td>9</td><td>PROD009</td><td>Bufanda Gris</td></tr><tr><td>1..</td><td>PROD010</td><td>Chaleco Marrón</td></tr></table>	#	cod_pro	nombre_pro	1	PROD001	Camisa Azul	2	PROD002	Pantalón Negro	3	PROD003	Chaqueta Gris	4	PROD004	Sudadera Roja	5	PROD005	Vestido Floral	6	PROD006	Zapatos Negros	7	PROD007	Blusa Blanca	8	PROD008	Falda Negra	9	PROD009	Bufanda Gris	1..	PROD010	Chaleco Marrón
#	cod_pro	nombre_pro																																
1	PROD001	Camisa Azul																																
2	PROD002	Pantalón Negro																																
3	PROD003	Chaqueta Gris																																
4	PROD004	Sudadera Roja																																
5	PROD005	Vestido Floral																																
6	PROD006	Zapatos Negros																																
7	PROD007	Blusa Blanca																																
8	PROD008	Falda Negra																																
9	PROD009	Bufanda Gris																																
1..	PROD010	Chaleco Marrón																																

Y las configuraciones de cada transformación:

## dim\_tiempo

Su archivo es Transformation5-tiempo.ktr

diagrama	salida																																																																														
<div><p>dim_tiempo    Select values    Sort rows    Table output</p></div>	<div><h3>Execution Results</h3><div>Logging   Execution History   Step Metrics   Performance Graph   Metrics   Preview data</div><div><input checked="" type="radio"/> First rows   <input type="radio"/> Last rows   <input type="radio"/> Off</div><table><thead><tr><th>#</th><th>idfecha</th><th>fecha</th><th>año</th><th>mes</th><th>día</th></tr></thead><tbody><tr><td>1</td><td>20240101.0</td><td>2024/01/01 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>1.0</td></tr><tr><td>2</td><td>20240102.0</td><td>2024/01/02 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>2.0</td></tr><tr><td>3</td><td>20240115.0</td><td>2024/01/15 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>15.0</td></tr><tr><td>4</td><td>20240116.0</td><td>2024/01/16 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>16.0</td></tr><tr><td>5</td><td>20240117.0</td><td>2024/01/17 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>17.0</td></tr><tr><td>6</td><td>20240118.0</td><td>2024/01/18 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>18.0</td></tr><tr><td>7</td><td>20240119.0</td><td>2024/01/19 00:00:00.000</td><td>2024.0</td><td>1.0</td><td>19.0</td></tr><tr><td>8</td><td>20240201.0</td><td>2024/02/01 00:00:00.000</td><td>2024.0</td><td>2.0</td><td>1.0</td></tr><tr><td>9</td><td>20240202.0</td><td>2024/02/02 00:00:00.000</td><td>2024.0</td><td>2.0</td><td>2.0</td></tr><tr><td>1..</td><td>20240203.0</td><td>2024/02/03 00:00:00.000</td><td>2024.0</td><td>2.0</td><td>3.0</td></tr><tr><td>1..</td><td>20240204.0</td><td>2024/02/04 00:00:00.000</td><td>2024.0</td><td>2.0</td><td>4.0</td></tr><tr><td>1..</td><td>20240205.0</td><td>2024/02/05 00:00:00.000</td><td>2024.0</td><td>2.0</td><td>5.0</td></tr></tbody></table></div>	#	idfecha	fecha	año	mes	día	1	20240101.0	2024/01/01 00:00:00.000	2024.0	1.0	1.0	2	20240102.0	2024/01/02 00:00:00.000	2024.0	1.0	2.0	3	20240115.0	2024/01/15 00:00:00.000	2024.0	1.0	15.0	4	20240116.0	2024/01/16 00:00:00.000	2024.0	1.0	16.0	5	20240117.0	2024/01/17 00:00:00.000	2024.0	1.0	17.0	6	20240118.0	2024/01/18 00:00:00.000	2024.0	1.0	18.0	7	20240119.0	2024/01/19 00:00:00.000	2024.0	1.0	19.0	8	20240201.0	2024/02/01 00:00:00.000	2024.0	2.0	1.0	9	20240202.0	2024/02/02 00:00:00.000	2024.0	2.0	2.0	1..	20240203.0	2024/02/03 00:00:00.000	2024.0	2.0	3.0	1..	20240204.0	2024/02/04 00:00:00.000	2024.0	2.0	4.0	1..	20240205.0	2024/02/05 00:00:00.000	2024.0	2.0	5.0
#	idfecha	fecha	año	mes	día																																																																										
1	20240101.0	2024/01/01 00:00:00.000	2024.0	1.0	1.0																																																																										
2	20240102.0	2024/01/02 00:00:00.000	2024.0	1.0	2.0																																																																										
3	20240115.0	2024/01/15 00:00:00.000	2024.0	1.0	15.0																																																																										
4	20240116.0	2024/01/16 00:00:00.000	2024.0	1.0	16.0																																																																										
5	20240117.0	2024/01/17 00:00:00.000	2024.0	1.0	17.0																																																																										
6	20240118.0	2024/01/18 00:00:00.000	2024.0	1.0	18.0																																																																										
7	20240119.0	2024/01/19 00:00:00.000	2024.0	1.0	19.0																																																																										
8	20240201.0	2024/02/01 00:00:00.000	2024.0	2.0	1.0																																																																										
9	20240202.0	2024/02/02 00:00:00.000	2024.0	2.0	2.0																																																																										
1..	20240203.0	2024/02/03 00:00:00.000	2024.0	2.0	3.0																																																																										
1..	20240204.0	2024/02/04 00:00:00.000	2024.0	2.0	4.0																																																																										
1..	20240205.0	2024/02/05 00:00:00.000	2024.0	2.0	5.0																																																																										

The image displays four screenshots of SQL Server Enterprise Manager wizards, arranged in a 2x2 grid. Each screenshot shows a different step in a data processing workflow.

- Top Left: Microsoft Local Data**
  - Step name: **selectvalue**
  - Task: **Stages | Control | Error Handling | Fields | Additional Input Fields**
  - Input: **Input data from (single)** | Local 2017-10-01 (single MDS)
  - File or directory: **File or directory** | Add | Browse...
  - Regular Expression: **Regular Expression** | Add | Browse...
  - Exclude Regular Expression: **Exclude Regular Expression** | Add | Browse...
  - Parent: **Parent** | Add | Browse...
  - Source Step: **File or Directory** | Wizard | Properties | New | Delete
  - Columns: **Columns** | Columns | Add | Remove | Up | Down | OK | Cancel
  - Accept Rowsets from previous step: **Accept Rowsets from previous step** | OK | Cancel
  - Map to table: **Map to table** | Map to table | OK | Cancel
  - Search for the input to use as a function: **Search for the input to use as a function** | OK | Cancel
  - Store Rowsets: **Store Rowsets** | OK | Cancel
- Top Right: Select values**
  - Step name: **select value**
  - Select & Alter: **Select & Alter** | Remove | Meta-data
  - Fields: **Fields** | Edit Mapping

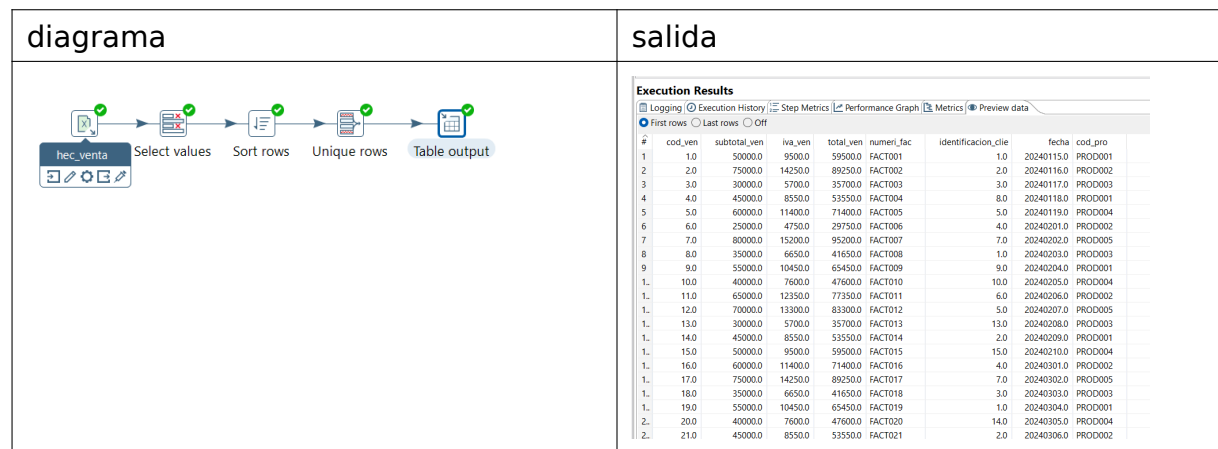
#	Fieldname	Rename to	Length	Precision
1	idfecda			
2	fecda			
3	afda			
4	mes			
5	dia			

  - Include unspecified fields, ordered: ☐
  - Buttons: **OK** | **Cancel** | **Help**
- Bottom Left: Table output**
  - Step name: **tableoutput**
  - Connection: **myhsdbcon** | Edit... | New... | Wizard...
  - Target schema: **etl\_jaradba** | Browse...
  - Target table: **dim\_itemspo** | Browse...
  - Commit size: **1000** | Browse...
  - Truncate table: ☐
  - Signal insert errors: ☐
  - Specify database fields: ☐
  - Main options: **Database fields** | Partition data over tables: ☐ | Partitioning type: **Partitioning type** | Partition data per month: ☐ | Partition data per day: ☐ | Use batch update for inserts: ☒
  - Is the name of the table defined in a field? ☐ | Field that contains name of table: **Field that contains name of table** | Store the tablename field: ☐ | Return auto-generated key: ☐ | Name of auto-generated key field: **Name of auto-generated key field**
  - Buttons: **OK** | **Cancel** | **SQL** | **Help**
- Bottom Right: Sort rows**
  - Step name: **sort rows**
  - Sort directory: **Sort directory** | Browse...
  - TMP-file prefix: **out**
  - Sort size (rows in memory): **1000000**
  - Free memory threshold (in %): **Free memory threshold (in %)**
  - Compare TMP-files: ☐
  - Only pass unique rows? (verifies keys only): ☐
  - Fields: **Fields** | Edit Mapping

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted?
1	idfecda	Y	N	N	0	N
2	fecda	Y	N	N	0	N
3	afda	Y	N	N	0	N
4	mes	Y	N	N	0	N
5	dia	Y	N	N	0	N

  - Buttons: **OK** | **Cancel** | **Get Fields** | **Help**

Su archivo es Transformation5-venta.ktr



Y las configuraciones de cada transformación:

Step name: **sort**

File: C:\Users\Control-E\OneDrive\Files\Additional output\Files\Sort-03-05-2025\Sort-03-05-2025.kjb

Sort directory: %Java.io.tmpdir\5%

Sort size (rows in memory): 1000000

Free memory threshold (in %): 90

Only pass unique rows? (verified keys only): ☐

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted
1	cod_yen	Y	N	N	0	N
2	subtotal_yen	Y	N	N	0	N
3	iva_yen	Y	N	N	0	N
4	total_yen	Y	N	N	0	N
5	numeri_fac	Y	N	N	0	N
6	identificacion_cliente	Y	N	N	0	N
7	fecha	Y	N	N	0	N
8	cod_pro	Y	N	N	0	N

Step name: **select values**

Select & Alter Remove Meta-data

Fields:

#	Fieldname	Rename to	Length	Precision
1	cod_yen			
2	subtotal_yen			
3	iva_yen			
4	total_yen			
5	numeri_fac			
6	identificacion_cliente			
7	fecha			
8	cod_pro			

Step name: **sort rows**

Sort directory: %Java.io.tmpdir\5%

Sort size (rows in memory): 1000000

Free memory threshold (in %): 90

Only pass unique rows? (verified keys only): ☐

Fields:

#	Fieldname	Ascending	Case sensitive compare?	Sort based on current locale?	Collator Strength	Presorted
1	cod_yen	Y	N	N	0	N
2	subtotal_yen	Y	N	N	0	N
3	iva_yen	Y	N	N	0	N
4	total_yen	Y	N	N	0	N
5	numeri_fac	Y	N	N	0	N
6	identificacion_cliente	Y	N	N	0	N
7	fecha	Y	N	N	0	N
8	cod_pro	Y	N	N	0	N

Step name: **unique rows**

Settings

Add counter to output? ☐ Counter field

Redirect duplicate row ☐ Error description

Fields to compare on (no entries means: compare complete row)

#	Fieldname	Ignore case
1	cod_yen	N
2	subtotal_yen	N
3	iva_yen	N
4	total_yen	N
5	numeri_fac	N
6	identificacion_cliente	N
7	fecha	N
8	cod_pro	N

Step name: **table output**

Connection: mymysqllocal

Target schema: gtd\_pentaho

Target table: hec\_venta

Commit size: 1000

Truncate table: ☐

Specify output schema: ☐

Main options: Database fields

Partition data over tables: ☐

Partitioning field:

Partition data per month: ☐

Partition data per day: ☐

Use batch update for inserts: ☒

Is the name of the table defined in a field? ☐

Field that contains name of table:

Store the tablename field: ☐

Return auto-generated key: ☐

Name of auto-generated key field:

# ETLs con jobs

Su archivo es Job1.kjb

diagrama

```
graph LR; Start --> dim_cliente; dim_cliente --> dim_producto; dim_producto --> dim_tiempo; dim_tiempo --> hec_ventas;
```

logs

Execution Results

Logging History Job metrics Metrics

2025-03-05 17:46:23.688 - Job1 - Starting entry [hec\_ventas]

2025-03-05 17:46:23.688 - hec\_ventas - Using run configuration [Pentaho local]

2025-03-05 17:46:23.697 - Transformation5-venta - Dispatching started for transformation [Transformation5-venta]

2025-03-05 17:46:23.697 - Table output.0 - Connected to database [mysqlocal] (commit=1000)

2025-03-05 17:46:23.721 - hec\_venta.0 - Finished processing (I=0, O=0, R=0, W=25, U=0, E=0)

2025-03-05 17:46:23.728 - Select values.0 - Finished processing (I=0, O=0, R=25, W=25, U=0, E=0)

2025-03-05 17:46:23.745 - Sort rows.0 - Finished processing (I=0, O=0, R=25, W=25, U=0, E=0)

2025-03-05 17:46:23.759 - Unique rows.0 - Finished processing (I=0, O=0, R=25, W=25, U=0, E=0)

2025-03-05 17:46:23.771 - Table output.0 - Finished processing (I=0, O=25, R=25, W=25, U=0, E=0)

2025-03-05 17:46:23.773 - Job1 - Finished job entry [hec\_ventas] (result=[true])

2025-03-05 17:46:23.773 - Job1 - Finished job entry [dim\_tiempo] (result=[true])

2025-03-05 17:46:23.773 - Job1 - Finished job entry [dim\_producto] (result=[true])

2025-03-05 17:46:23.773 - Job1 - Finished job entry [dim\_cliente] (result=[true])

2025-03-05 17:46:23.773 - Job1 - Job execution finished

2025-03-05 17:46:23.773 - Spoon - Job has ended.

Y las configuraciones de cada transformación:

<p><b>Start</b></p> <p>Job entry name: <input type="text" value="Start"/></p> <p>Repeat: <input type="checkbox"/></p> <p>Type: <b>No Scheduling</b></p> <p>Interval in seconds: <input type="text" value="0"/></p> <p>Interval in minutes: <input type="text" value="60"/></p> <p>Time of day: <input type="text" value="12"/> : <input type="text" value="0"/></p> <p>Day of week: <b>Monday</b></p> <p><input type="checkbox"/> Help f month: <input type="text" value="1"/></p>	<p><b>Transformation</b></p> <p>Entry Name: <input type="text" value="sim_cliente"/></p> <p>Transformation: <input type="text" value="C:\bd3transformaciones\Transformation5-cliente.ktr"/> Browse...</p> <p>Options \ Logging \ Arguments \ Parameters</p> <p>Run configuration: <b>Pentaho local</b></p> <p>Execution</p> <p><input type="checkbox"/> Execute every input row</p> <p><input type="checkbox"/> Clear results rows before execution</p> <p><input type="checkbox"/> Clear results files before execution</p> <p><input checked="" type="checkbox"/> Wait for remote transformation to complete</p> <p><input type="checkbox"/> Follow local abort to remote transformation</p> <p><input type="checkbox"/> Suppress result data from remote transformation</p> <p><input type="button" value="Help"/> <input type="button" value="OK"/> <input type="button" value="Cancel"/></p>	<p><b>Transformation</b></p> <p>Entry Name: <input type="text" value="sim_producto"/></p> <p>Transformation: <input type="text" value="C:\bd3transformaciones\Transformation5-producto.ktr"/> Browse...</p> <p>Options \ Logging \ Arguments \ Parameters</p> <p>Run configuration: <b>Pentaho local</b></p> <p>Execution</p> <p><input type="checkbox"/> Execute every input row</p> <p><input type="checkbox"/> Clear results rows before execution</p> <p><input type="checkbox"/> Clear results files before execution</p> <p><input checked="" type="checkbox"/> Wait for remote transformation to complete</p> <p><input type="checkbox"/> Follow local abort to remote transformation</p> <p><input type="checkbox"/> Suppress result data from remote transformation</p> <p><input type="button" value="Help"/> <input type="button" value="OK"/> <input type="button" value="Cancel"/></p>
<p><b>Transformation</b></p> <p>Entry Name: <input type="text" value="sim_cliente"/></p> <p>Transformation: <input type="text" value="C:\bd3transformaciones\Transformation5-tempo.ktr"/> Browse...</p> <p>Options \ Logging \ Arguments \ Parameters</p> <p>Run configuration: <b>Pentaho local</b></p> <p>Execution</p> <p><input type="checkbox"/> Execute every input row</p> <p><input type="checkbox"/> Clear results rows before execution</p> <p><input type="checkbox"/> Clear results files before execution</p> <p><input checked="" type="checkbox"/> Wait for remote transformation to complete</p> <p><input type="checkbox"/> Follow local abort to remote transformation</p> <p><input type="checkbox"/> Suppress result data from remote transformation</p> <p><input type="button" value="Help"/> <input type="button" value="OK"/> <input type="button" value="Cancel"/></p>	<p><b>Transformation</b></p> <p>Entry Name: <input type="text" value="sim_cliente"/></p> <p>Transformation: <input type="text" value="C:\bd3transformaciones\Transformation5-venta.ktr"/> Browse...</p> <p>Options \ Logging \ Arguments \ Parameters</p> <p>Run configuration: <b>Pentaho local</b></p> <p>Execution</p> <p><input type="checkbox"/> Execute every input row</p> <p><input type="checkbox"/> Clear results rows before execution</p> <p><input type="checkbox"/> Clear results files before execution</p> <p><input checked="" type="checkbox"/> Wait for remote transformation to complete</p> <p><input type="checkbox"/> Follow local abort to remote transformation</p> <p><input type="checkbox"/> Suppress result data from remote transformation</p> <p><input type="button" value="Help"/> <input type="button" value="OK"/> <input type="button" value="Cancel"/></p>	