

Nama : Putri Laura Latersia Br.L. Tobing

NIM : 191402042

KOM : C

Hitachi Vantara & Pentaho

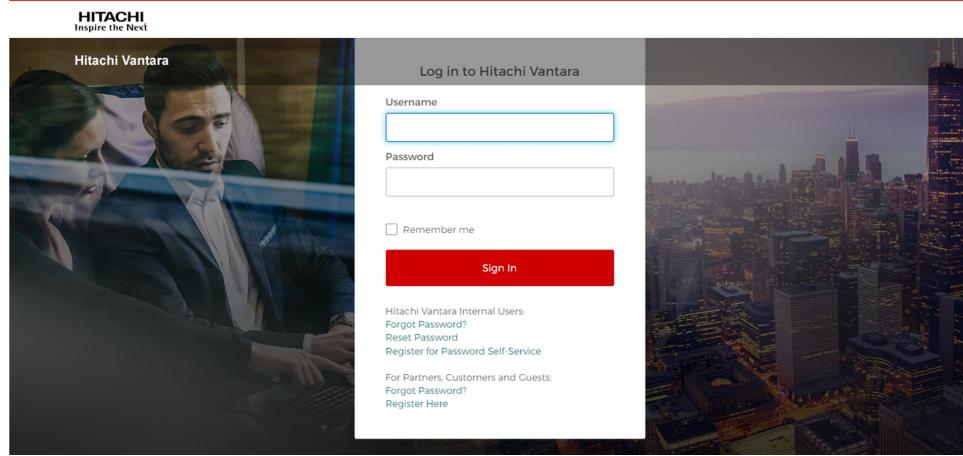
Pertama, membuka situs <https://www.hitachivantara.com/en-in/products/lumada-dataops/data-integration-analytics/download-pentaho.html> yang akan membawa kita ke beranda dari Hitachi Vantara

The screenshot shows the Hitachi Vantara website with a navigation bar at the top. The main headline reads "Lumada Data Integration and Analytics powered by Pentaho" and "Turn Data into Actionable Insights.". Below the headline, there are two buttons: "Try Hands On Experience" (highlighted with a red box) and "Start a Free 30 Day Trial". To the right of the headline is a large image of two workers standing on top of a wind turbine. A cookie consent banner at the bottom states: "This site uses cookies from Hitachi and third parties for our own business purposes and to personalize your experience. By using this site, you agree to the use of cookies. For more information, visit [Hitachi Cookies Policy](#). X".

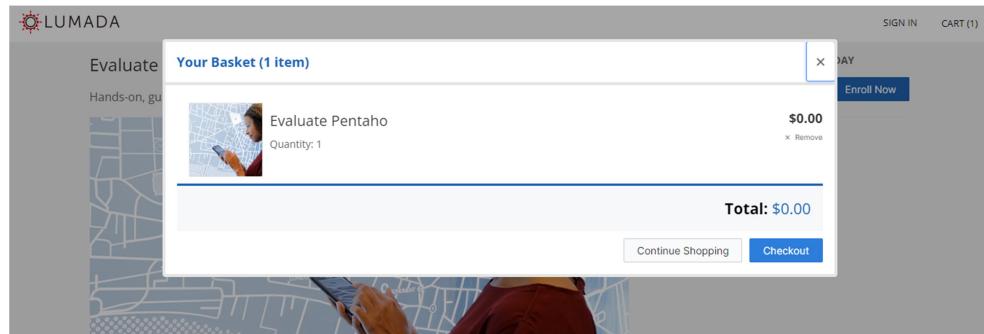
Setelah klik “TryHands On Experience” akan diarahkan untuk enroll course

The screenshot shows the Lumada website with a navigation bar at the top. The main headline reads "Evaluate Pentaho" and "Hands-on, guided evaluation, no installation required.". Below the headline is a large image of a woman looking at a smartphone screen, which displays a map. To the right, there is a "Free" offer and a "Enroll Now" button (highlighted with a red box). The text below the headline states: "Pentaho is an end-to-end, data integration and analytics platform. Here, we provide a guided tour of three aspects of the platform, complete with hands-on access to a full working environment." A bulleted list follows: • Pentaho Business Analytics • Pentaho Data Integration • Pentaho Metadata Injection (an advanced use case for Pentaho Data Integration)

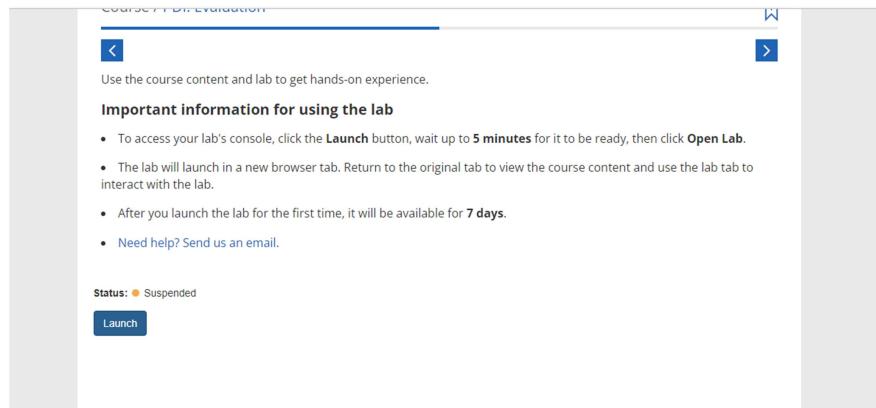
Saat klik tombol “Enroll Now” maka akan diminta login dengan username dan password



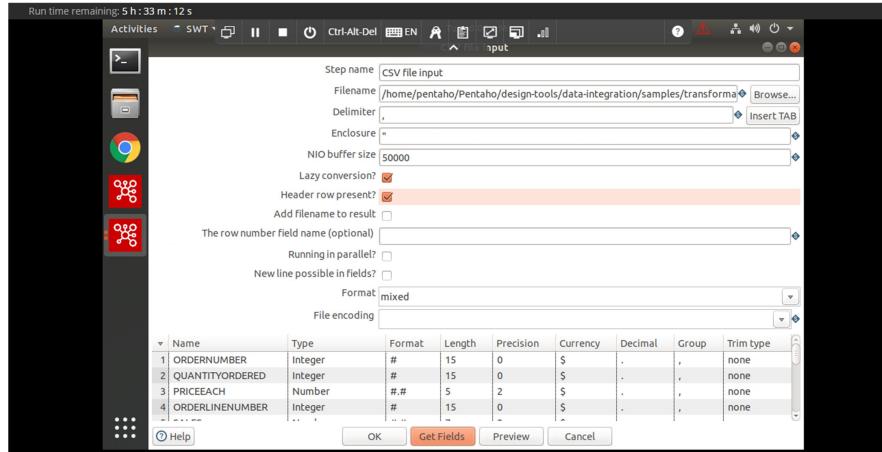
Setelah login kita akan bisa checkout course Evaluate Pentaho dan akan memulai tutorial penggunaan Pentaho.



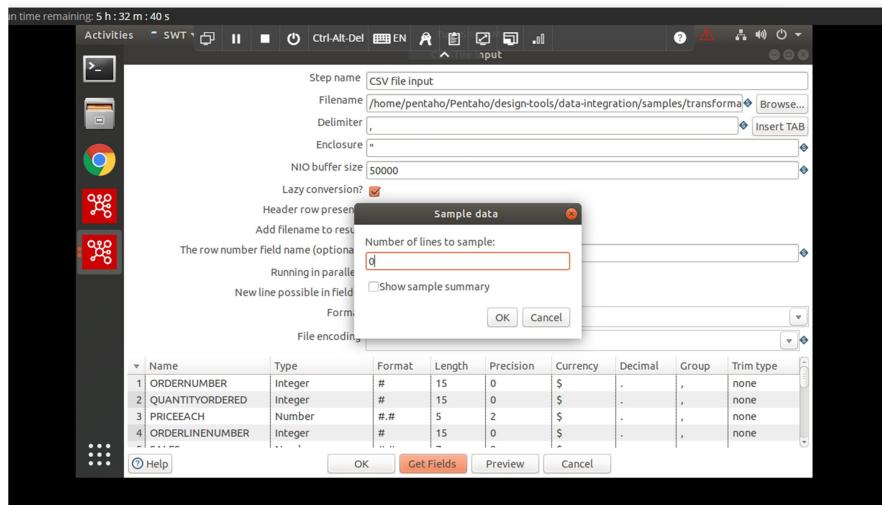
Pada bagian ini kita diminta untuk mengaktifkan mesin learning agar bisa menjalankan pentaho.



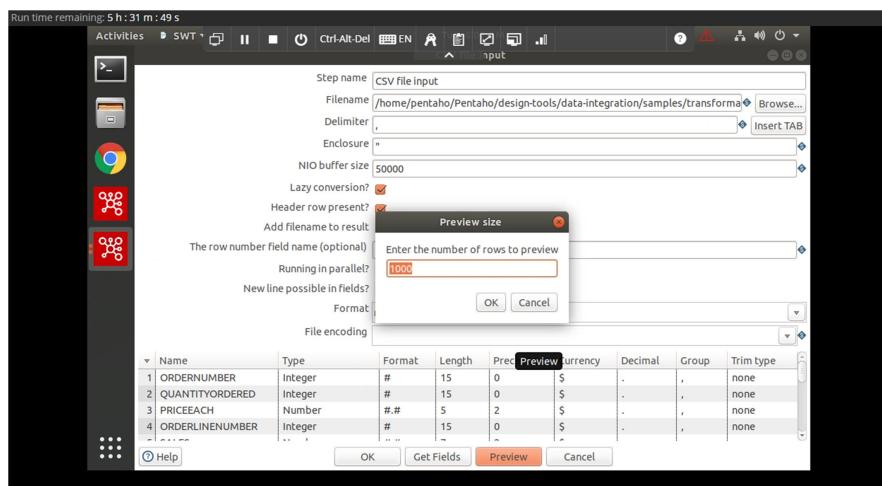
Input CSV pada canvas, kemudian double klik untuk browse file data sample



Klik Get Fields untuk memunculkan data yang telah diinput



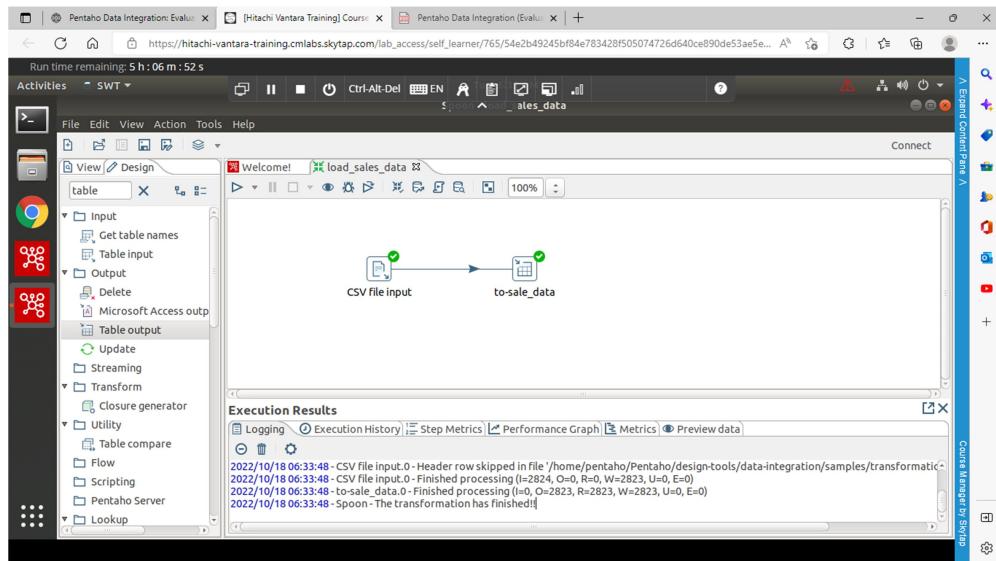
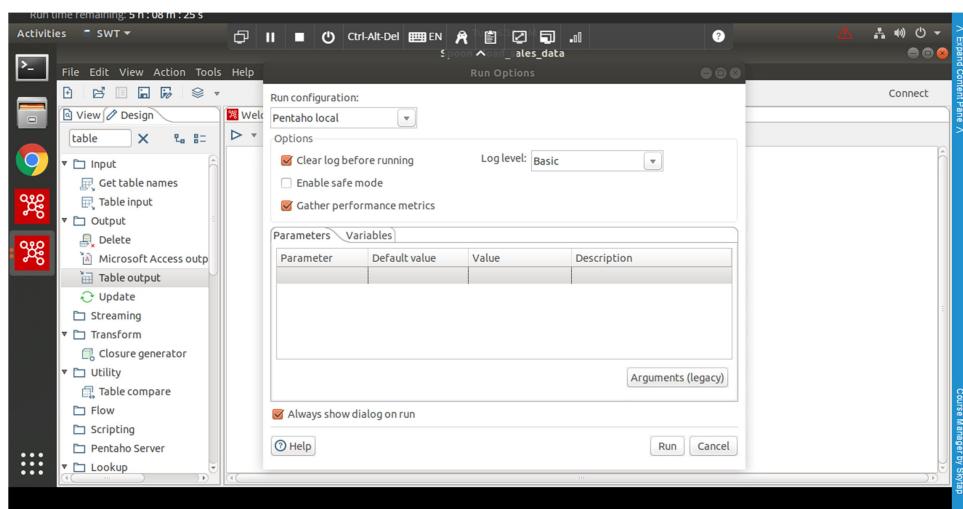
Preview untuk menampilkan 1000 data



Run time remaining: 5 h : 31 m : 43 s

	ORDERNUMBER	QUANTITYORDERED	PRICEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS
1	10107	30	95.7	2	2871.0	2/24/2003 0:00	Shipped
2	10121	34	81.3	5	2765.9	5/7/2003 0:00	Shipped
3	10134	41	94.7	2	3884.3	7/1/2003 0:00	Shipped
4	10145	45	83.3	6	3746.7	8/25/2003 0:00	Shipped
5	10159	49	100	14	5205.3	10/10/2003 0:00	Shipped
6	10168	36	96.7	1	3479.8	10/28/2003 0:00	Shipped
7	10180	29	86.1	9	2497.8	11/11/2003 0:00	Shipped
8	10188	48	100	1	5512.3	11/18/2003 0:00	Shipped
9	10201	22	98.6	2	2168.5	12/1/2003 0:00	Shipped
10	10211	41	100	14	4708.4	1/15/2004 0:00	Shipped
11	10223	37	100	1	3965.7	2/20/2004 0:00	Shipped
12	10237	23	100	7	2333.1	4/5/2004 0:00	Shipped
13	10251	28	100	2	3188.6	5/18/2004 0:00	Shipped
14	10263	34	100	2	3676.8	6/28/2004 0:00	Shipped
15	10275	45	92.8	1	4177.4	7/23/2004 0:00	Shipped
16	10285	36	100	6	4099.7	8/27/2004 0:00	Shipped
17	10299	23	100	9	2597.4	9/30/2004 0:00	Shipped
18	10309	41	100	5	4394.4	10/15/2004 0:00	Shipped
19	10318	46	94.7	1	4358.1	11/2/2004 0:00	Shipped
20	10320	42	100	4	4200.4	11/15/2004 0:00	Shipped

Untuk Run data yang telah dibuat



Menampilkan data dari DATA_SALE dalam bentuk table

The screenshot shows the Kettle Data Integration Platform interface. The 'Database Explorer' window is open, displaying a tree view of database objects. Under the 'Tables' node, the 'DATA_SALE' table is selected and highlighted in orange. A preview window titled 'Preview first 100' is overlaid on the main interface, showing the first 100 rows of the DATA_SALE table.

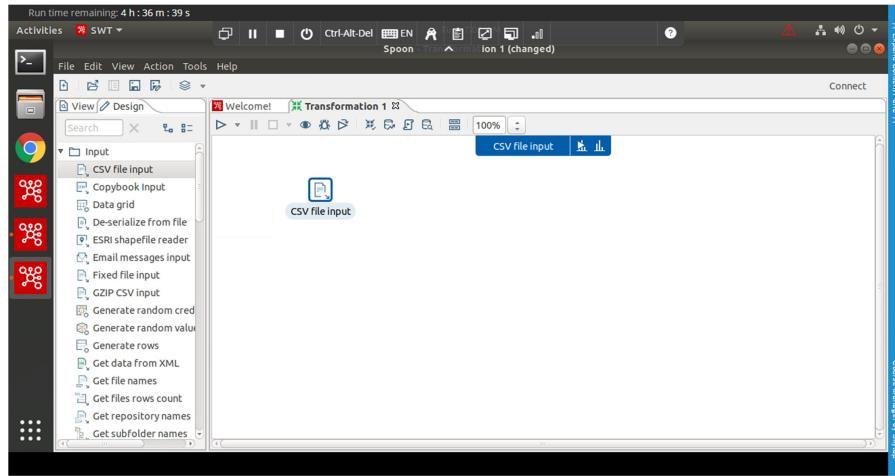
The screenshot shows the Kettle Data Integration Platform interface. The 'Database Explorer' window is open, displaying a tree view of database objects. Under the 'Tables' node, the 'DATA_SALE' table is selected. A context menu is open over the table, with the 'Preview first 100' option highlighted.

The screenshot shows the Kettle Data Integration Platform interface. The 'Database Explorer' window is open, displaying a tree view of database objects. Under the 'Tables' node, the 'DATA_SALE' table is selected. A preview window titled 'Rows of step: DATA_SALE (100 rows)' is open, showing the first 20 rows of the DATA_SALE table. The columns listed are ORDERNUMBER, QUANTITYORDERED, PRICEEACH, ORDERLINENUMBER, SALES, ORDERDATE, STATUS, QTR_ID, MONTH_ID, YEAR_ID, and P.

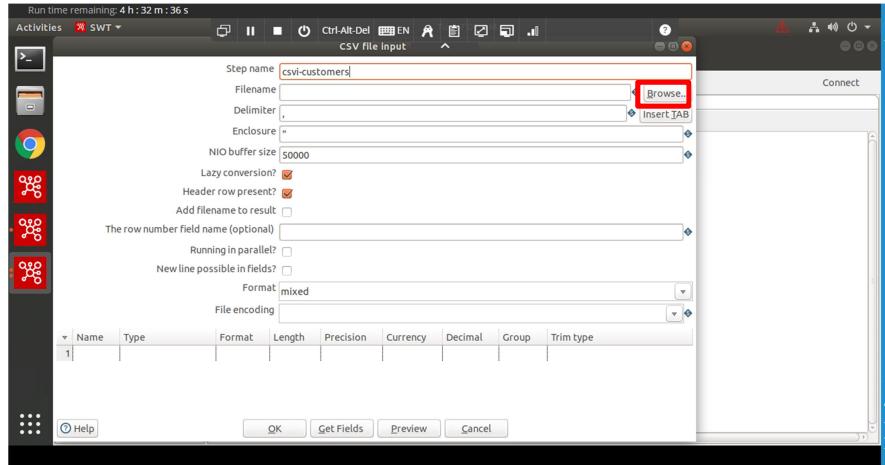
	ORDERNUMBER	QUANTITYORDERED	PRICEEACH	ORDERLINENUMBER	SALES	ORDERDATE	STATUS	QTR_ID	MONTH_ID	YEAR_ID	P
1	10107	30	95.7	2	2871.0	2/24/2003 0:00	Shipped	1	2	2003	M
2	10121	34	81.35	5	2765.9	5/7/2003 0:00	Shipped	2	5	2003	M
3	10134	41	94.74	2	3884.34	7/1/2003 0:00	Shipped	3	7	2003	M
4	10145	45	83.26	6	3746.7	8/25/2003 0:00	Shipped	3	8	2003	M
5	10159	49	100.0	14	5205.27	10/10/2003 0:00	Shipped	4	10	2003	M
6	10168	36	96.66	1	3479.76	10/28/2003 0:00	Shipped	4	10	2003	M
7	10180	29	86.13	9	2497.77	11/11/2003 0:00	Shipped	4	11	2003	M
8	10188	48	100.0	1	5512.32	11/18/2003 0:00	Shipped	4	11	2003	M
9	10201	22	98.57	2	2168.54	12/1/2003 0:00	Shipped	4	12	2003	M
10	10211	41	100.0	14	4708.44	1/15/2004 0:00	Shipped	1	1	2004	M
11	10223	37	100.0	1	3965.66	2/20/2004 0:00	Shipped	1	2	2004	M
12	10237	23	100.0	7	2333.12	4/5/2004 0:00	Shipped	2	4	2004	M
13	10251	28	100.0	2	3188.64	5/18/2004 0:00	Shipped	2	5	2004	M
14	10263	34	100.0	2	3676.76	6/28/2004 0:00	Shipped	2	6	2004	M
15	10275	45	92.83	1	4177.35	7/23/2004 0:00	Shipped	3	7	2004	M
16	10285	36	100.0	6	4099.68	8/27/2004 0:00	Shipped	3	8	2004	M
17	10299	23	100.0	9	2597.39	9/30/2004 0:00	Shipped	3	9	2004	M
18	10309	41	100.0	5	4394.38	10/15/2004 0:00	Shipped	4	10	2004	M
19	10318	46	94.74	1	4358.04	11/2/2004 0:00	Shipped	4	11	2004	M
20	10329	42	100.0	1	4396.14	11/15/2004 0:00	Shipped	4	11	2004	M

Reading Customers Data

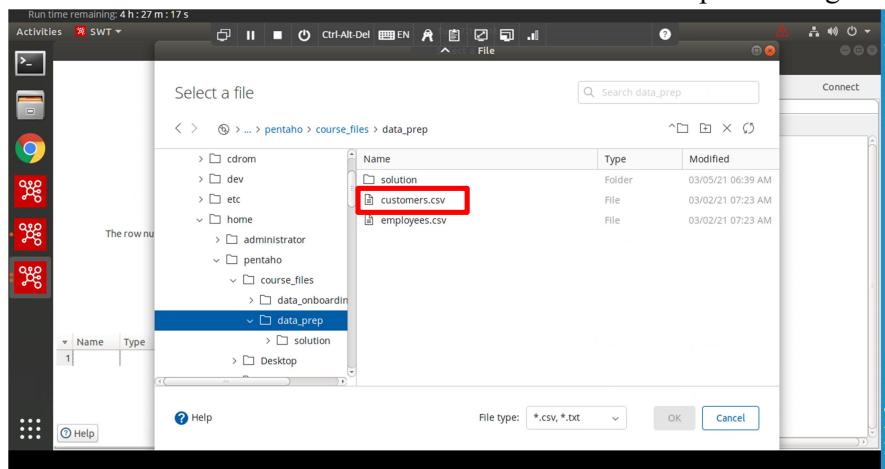
1. Menambahkan CSV file input ke dalam canvas dengan cara drag file



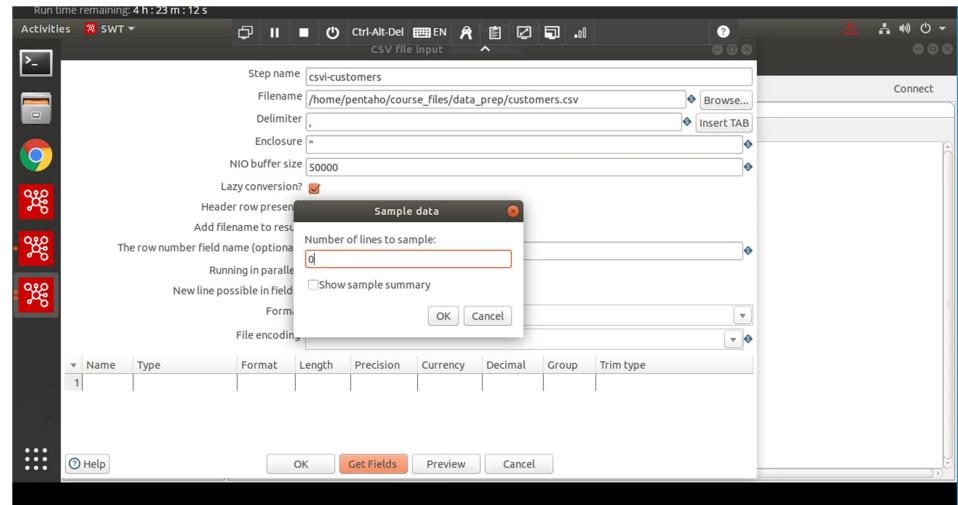
2. Double klik pada file csv dan berikan nama untuk file csv tersebut. Klik browse yang ada pada sebelah kanan untuk mencari data.



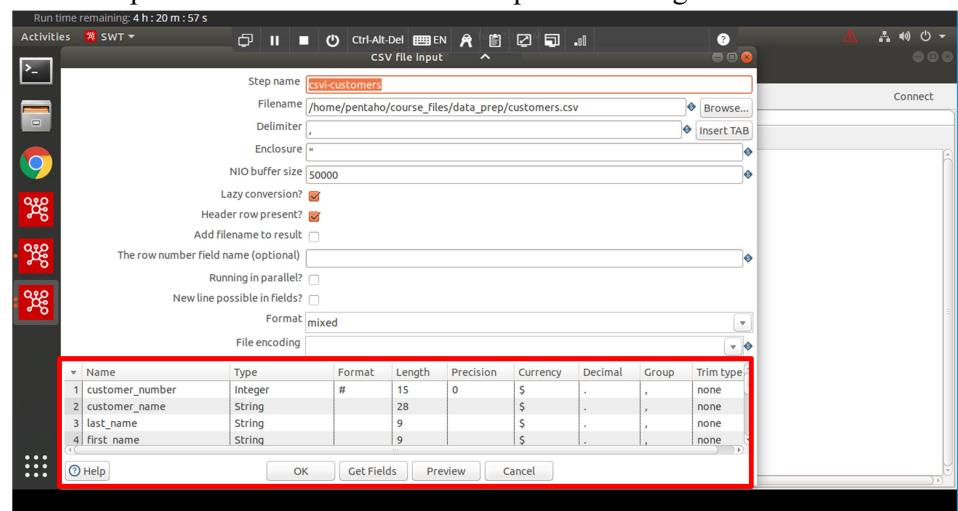
3. Mencari data dari `/home/pentaho/course_files/data_prep` dan pilih "customers.csv" kemudian klik tombol OK pada bagian bawah.



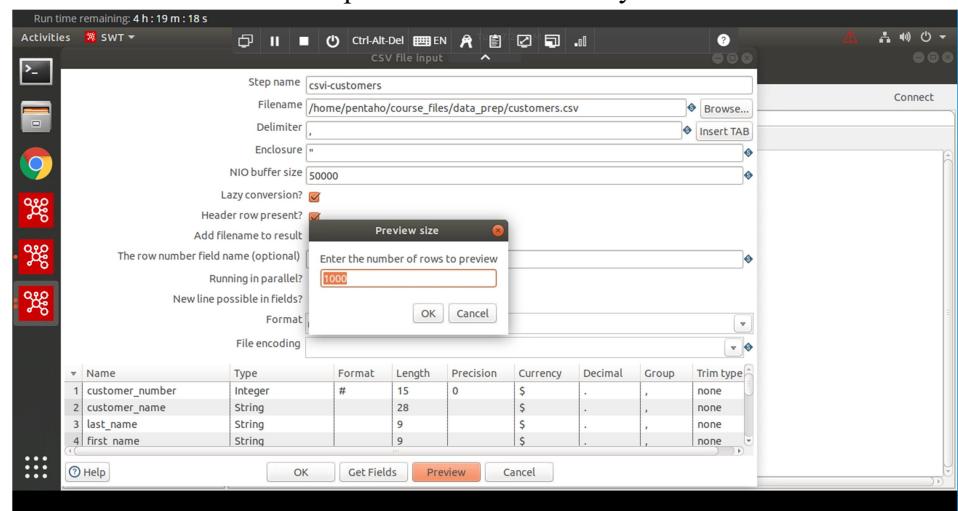
4. Klik tombol Get Fields dan masukkan number of line to data sample “0” kemudian klik OK



Data Sample customers.csv akan muncul pada table bagian bawah



5. Klik Preview untuk menampilkan data table sebanyak 1000 baris



Data akan tampil dalam bentuk table yang sudah tersusun rapi yang tersusun dari label dan contain.

Run time remaining: 4 h : 18 m : 15 s

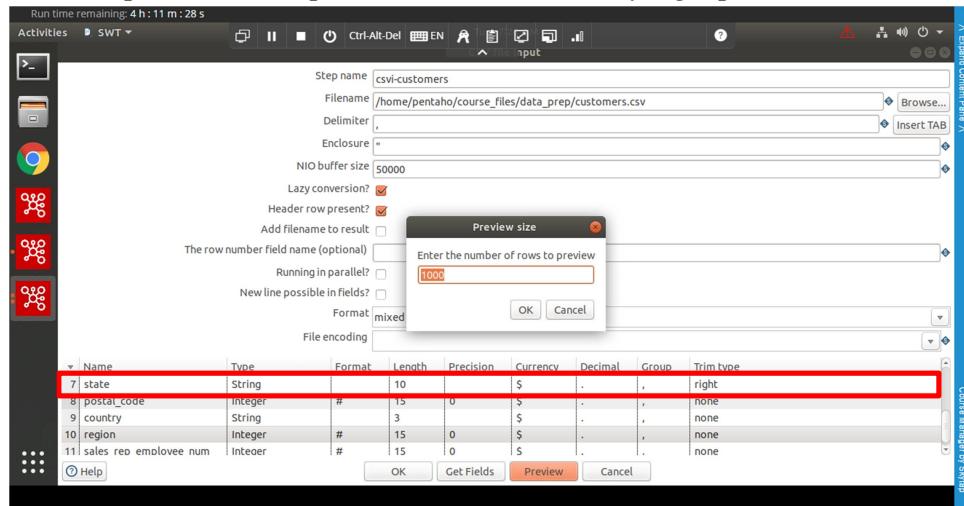
Activities SWT

Rows of step: csv-customers (36 rows)

	customer_number	customer_name	last_name	first_name	address	city	state	postal_code	country
1	112	Signal Gift Stores	King	Sue	245 Edgefield St.	Las Vegas	NV.	83030	USA
2	124	Mini Gifts Distributors Ltd.	Nelson	Valarie	982 Lilac Street	San Rafael	CA	97562	USA
3	129	Mini Wheels Co.	Murphy	Julie	125 Augusta St.	San Francisco	CA	94217	USA
4	131	Land of Toys Inc.	Yu	Kwai	9 South Summit Road	NYC	NY	10022	USA
5	151	Muscle Machine Inc	Young	Jeff	9788 Union Avenue	NYC	NY	10022	USA
6	157	Diecast Classics Inc.	Yu	Kyung	907 Windsor St.	Allentown	PA	70267	USA
7	161	Technics Stores Inc.	Hirano	Juri	7473 Bishop Ave.	Burlingame	CA.	94217	USA
8	168	American Souvenirs Inc	Franco	Sue	541 East Halifax Street	New Haven	CT.	97823	USA
9	173	Cambridge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51247	USA
10	175	Gift Depot Inc.	King	Julie	9920 Pennsylvania Lane	Bridgewater	CT	97562	USA
11	181	Vitachrome Inc.	Frick	Michael	48 Tower Lane	NYC	NY	10022	USA
12	198	Auto-Moto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	USA
13	204	Online Mini Collectables	Barajas	Miguel	8093 Archer St	Brickhaven	MA	58339	USA
14	205	Toys4GrownUps.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	USA
15	219	Boards & Toys Co.	Young	Leslie	377 Java Ave.	Glendale	CA	92561	USA
16	239	Collectable Mini Designs Co.	Thompson	Valarie	341 South Lumber St.	San Diego	CA	91217	USA
17	286	Marta's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	USA
18	319	Mini Classics	Frick	Steve	429 Anchorage St.	White Plains	NY.	24067	USA
19	320	Mini Creations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA.	50553	USA
20	321	Corporate Gift Ideas Co.	Brown	Julie	908 Gray St.	San Francisco	CA	94217	USA

Close Show Log

- Untuk mengubah state ke bagian sebelah kanan kita dapat mengubah trim type dan pilih “right” kemudian klik preview untuk melihat hasil dari 1000 baris data. Dan data pun akan ditampilkan dalam bentuk table yang rapi.



Run time remaining: 4 h : 10 m : 21 s

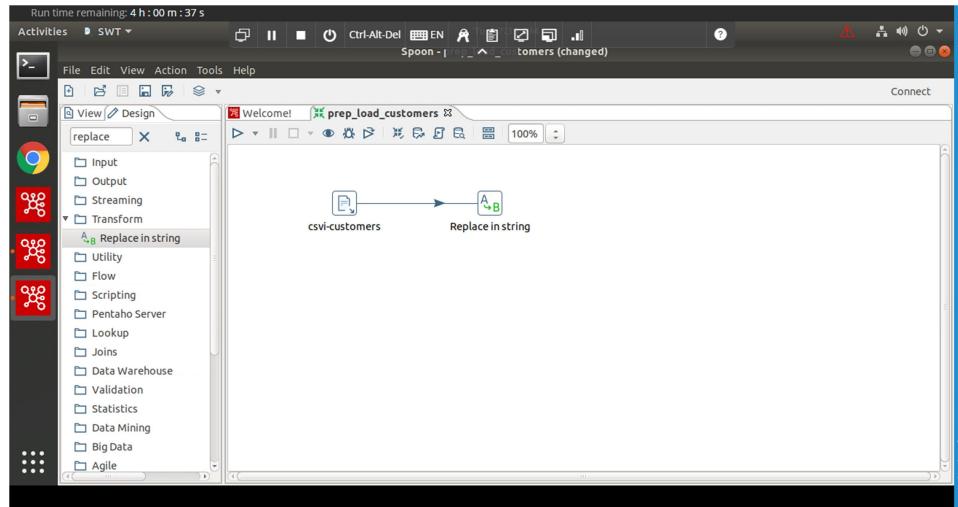
Activities SWT

Rows of step: csv-customers (36 rows)

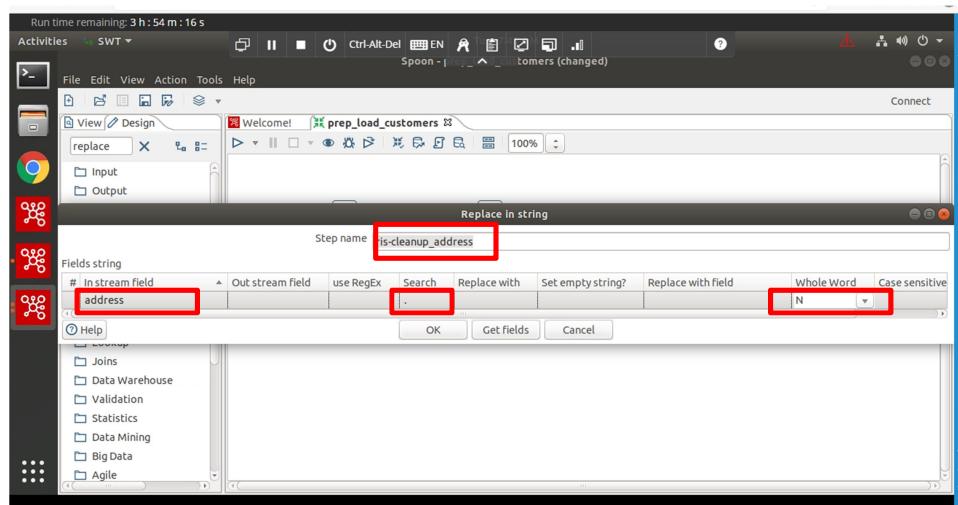
	customer_number	customer_name	last_name	first_name	address	city	state	postal_code	country
1	112	Signal Gift Stores	King	Sue	245 Edgefield St.	Las Vegas	NV.	83030	USA
2	124	Mini Gifts Distributors Ltd.	Nelson	Valarie	982 Lilac Street	San Rafael	CA	97562	USA
3	129	Mini Wheels Co.	Murphy	Julie	125 Augusta St.	San Francisco	CA	94217	USA
4	131	Land of Toys Inc.	Yu	Kwai	9 South Summit Road	NYC	NY	10022	USA
5	151	Muscle Machine Inc	Young	Jeff	9788 Union Avenue	NYC	NY	10022	USA
6	157	Diecast Classics Inc.	Yu	Kyung	907 Windsor St.	Allentown	PA	70267	USA
7	161	Technics Stores Inc.	Hirano	Juri	7473 Bishop Ave.	Burlingame	CA.	94217	USA
8	168	American Souvenirs Inc	Franco	Sue	541 East Halifax Street	New Haven	CT.	97823	USA
9	173	Cambridge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51247	USA
10	175	Gift Depot Inc.	King	Julie	9920 Pennsylvania Lane	Bridgewater	CT	97562	USA
11	181	Vitachrome Inc.	Frick	Michael	48 Tower Lane	NYC	NY	10022	USA
12	198	Auto-Moto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	USA
13	204	Online Mini Collectables	Barajas	Miguel	8093 Archer St	Brickhaven	MA	58339	USA
14	205	Toys4GrownUps.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	USA
15	219	Boards & Toys Co.	Young	Leslie	377 Java Ave.	Glendale	CA	92561	USA
16	239	Collectable Mini Designs Co.	Thompson	Valarie	341 South Lumber St.	San Diego	CA	91217	USA
17	286	Marta's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	USA
18	319	Mini Classics	Frick	Steve	429 Anchorage St.	White Plains	NY.	24067	USA
19	320	Mini Creations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA.	50553	USA
20	321	Corporate Gift Ideas Co.	Brown	Julie	908 Gray St.	San Francisco	CA	94217	USA

Standardize Address and State Data

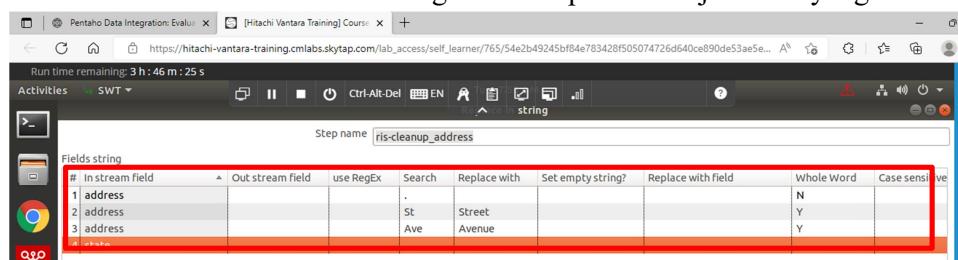
1. Replace dalam bentuk string dengan mencari “Replace” pada kolom search dan double klik pada “Replace in string” dan akan terhubung langsung dengan csvi-customers



2. Double klik pada “Replace In string” untuk melakukan konfigurasi dan berikan nama “ris-cleanup_address”. Pilih “address” pada bagian In Stream field, “.” Pada kolom search dan “N” pada whole word



Tambahkan kolom baru untuk mengubah setiap data menjadi data yang baru

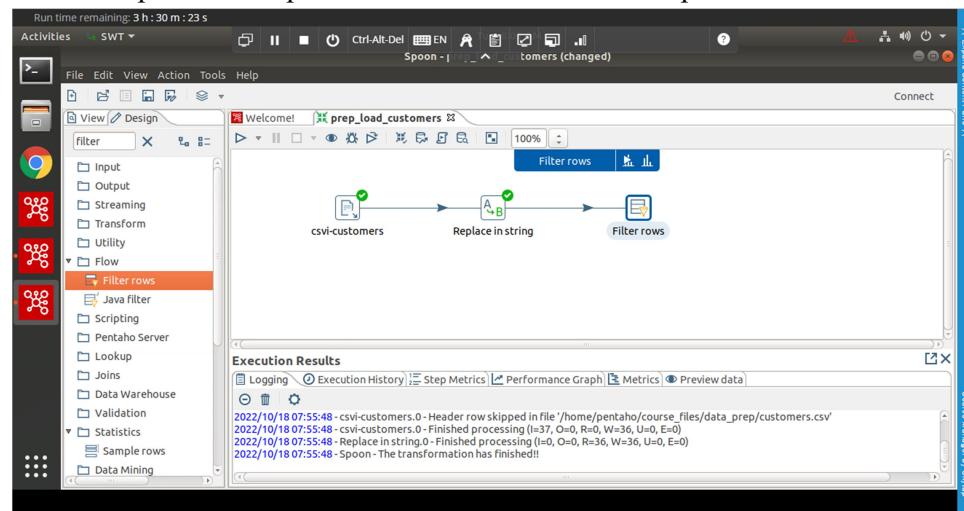


3. Untuk melihat hasilnya klik kanan pada Replace In string kemudian pilih preview. Terlihat data yang dicari sudah berubah menjadi kata baru yang telah diinput

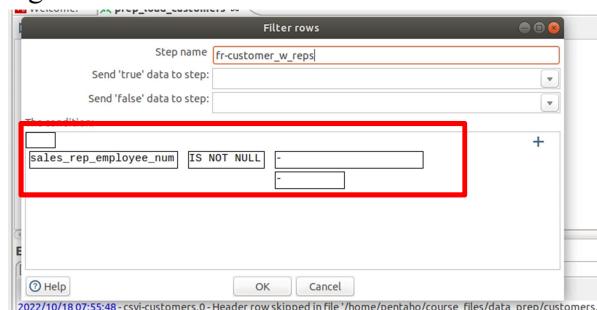
customer_number	customer_name	last_name	first_name	address	city	state	postal_code	country
1	112 Signal Gift Stores	King	Sue	245 Edgefield Street	Las Vegas	NV	83030	USA
2	124 Mini Gifts Distributors Ltd.	Nelson	Valarie	982 Lilac Street	San Rafael	CA	97562	USA
3	129 Mini Wheels Co.	Murphy	Julie	125 Augusta Street	San Francisco	CA	94217	USA
4	131 Land of Toys Inc.	Yu	Kwai	9 South Summit Road	NYC	NY	10022	USA
5	151 Muscle Machine Inc	Young	Jeff	9788 Union Avenue	NYC	NY	10022	USA
6	157 Diecast Classics Inc.	Yu	Kyung	907 Windsor Street	Allentown	PA	70267	USA
7	161 Technics Stores Inc.	Hirano	Juri	7473 Bishop Avenue	Burlingame	CA	94217	USA
8	168 American Souvenirs Inc	Franco	Sue	541 East Halifax Street	New Haven	CT	97823	USA
9	173 Cambridge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51247	USA
10	175 Gift Depot Inc.	King	Julie	9920 Pennsylvania Lane	Bridgewater	CT	97562	USA
11	181 Vitachrome Inc.	Frick	Michael	48 Tower Lane	NYC	NY	10022	USA
12	198 Auto-Moto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	USA
13	204 Online Mini Collectables	Barajas	Miguel	8093 Archer Street	Brickhaven	MA	58339	USA
14	205 Toys4GrownUps.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	USA
15	219 Boards & Toys Co.	Young	Leslie	377 Java Avenue	Glendale	CA	92561	USA
16	239 Collectable Mini Designs Co.	Thompson	Valarie	341 South Lumber Street	San Diego	CA	91217	USA
17	286 Marta's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	USA
18	319 Mini Classics	Frick	Steve	429 Anchorage Street	White Plains	NY	24067	USA
19	320 Mini Creations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA	50553	USA
20	321 Corporate Gift Ideas Co.	Brown	Julie	908 Gray Street	San Francisco	CA	94217	USA

Filters Customers

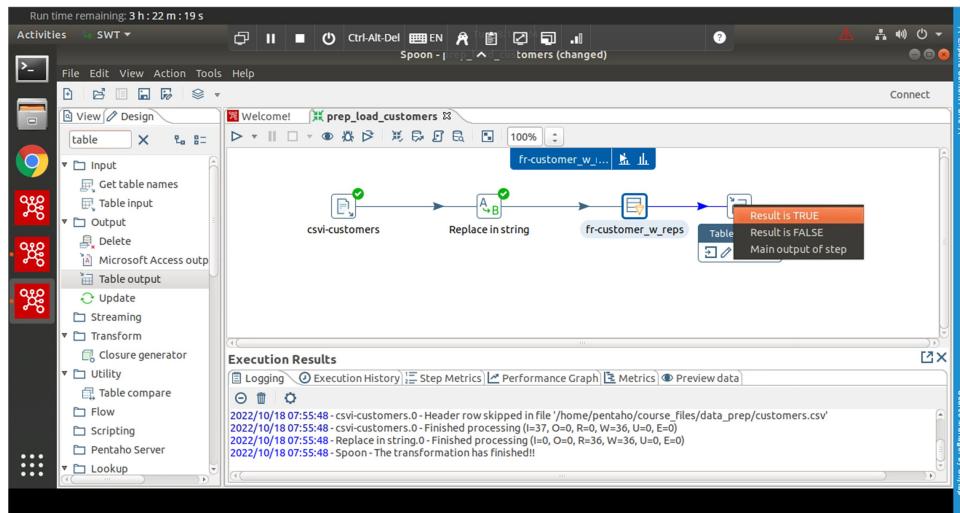
1. Cari filter pada kolom pencarian kemudian double klik pada filter rows



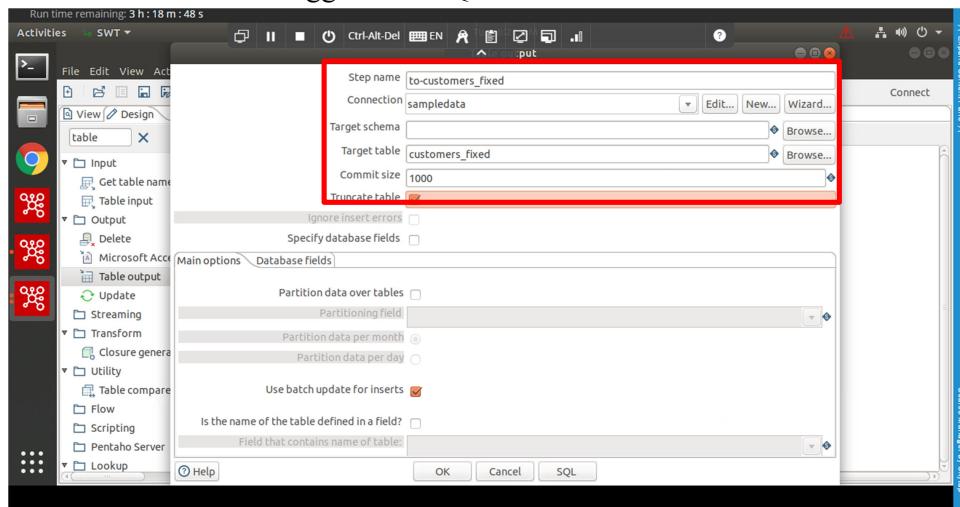
2. Double klik filter rows untuk menentukan kondisi dari data yang akan di filter dan beri nama pada bagian ini



3. Search table out pada kolom pencarian kemudian drag table output dan hubungkan fr-customer_w_reps dengan table output dengan kondisi TRUE

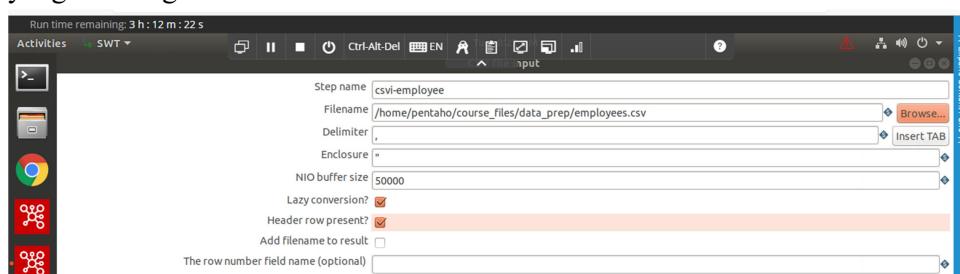


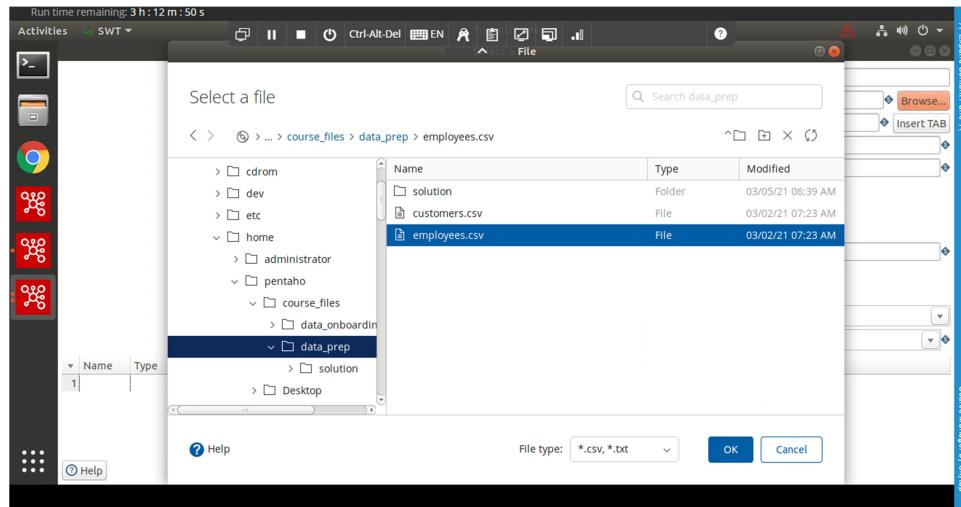
4. Untuk melihat hasilnya double klik pada table output dan isi setiap table untuk membuat table baru menggunakan SQL



Reading Employee Data

1. Tambahkan csv baru ke canvas dan beri nama csv tersebut. Lalu browse file data yang akan digunakan





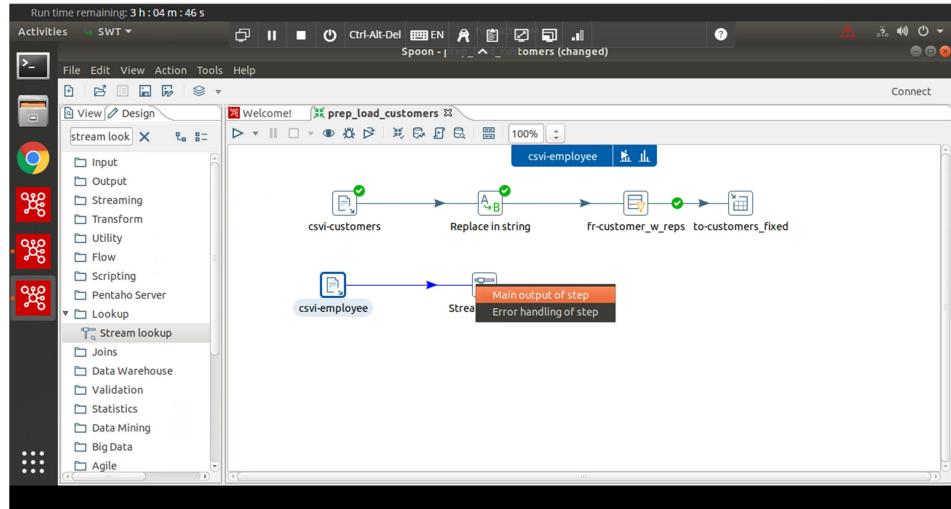
2. Get Fields untuk menampilkan label data dan preview untuk menampilkan table data yang sudah diinput. Transformasi employee data selesai

Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1 employee_number	Integer	#	15	0	\$.	,	none
2 last_name	String		9		\$,	,	none
3 first_name	String		8		\$.	,	none
4 email	String		31		\$.	,	none
5 region	Integer	#	15	0	\$.	,	none

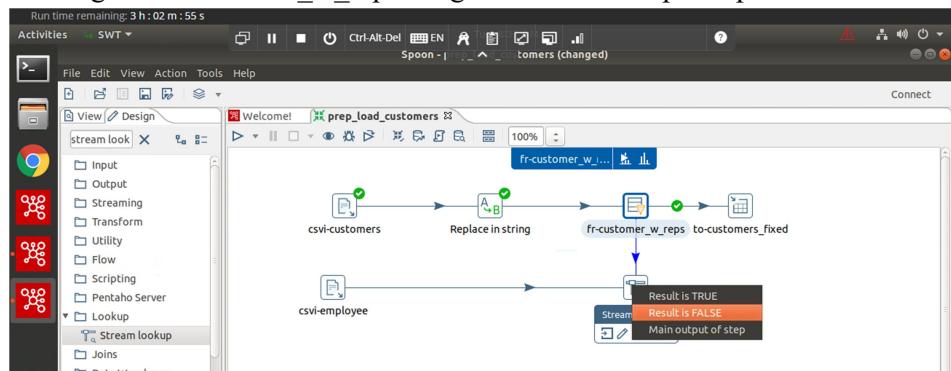
Rows of step: csvi-employee (23 rows)	employee_number	last_name	first_name	email	region	job_title
1	1002	Murphy	Diane	dmurphy@classicmodelcars.com	<null>	President
2	1056	Patterson	Mary	mpatterson@classicmodelcars.com	<null>	VP Sales
3	1076	Firrelli	Jeff	jfirrelli@classicmodelcars.com	<null>	VP Marketing
4	1088	Patterson	William	wpatterson@classicmodelcars.com	<null>	Sales Manager (JAPAN, APAC)
5	1102	Bondur	Gerard	athompson@classicmodelcars.com	<null>	Sale Manager (EMEA)
6	1143	Bow	Anthony	bhoward@classicmodelcars.com	<null>	Sales Manager (NA)
7	1165	Jennings	Leslie	ljennings@classicmodelcars.com	2	Sales Rep
8	1166	Thompson	Leslie	lthompson@classicmodelcars.com	1	Sales Rep
9	1188	Firrelli	Julie	jfirrelli@classicmodelcars.com	4	Sales Rep
10	1216	Patterson	Steve	spatterson@classicmodelcars.com	8	Sales Rep
11	1286	Tseng	Foon Yue	ftseng@classicmodelcars.com	14	Sales Rep
12	1323	Vanauf	George	gvanauf@classicmodelcars.com	3	Sales Rep
13	1337	Bondur	Loui	lbondur@classicmodelcars.com	17	Sales Rep
14	1370	Hernandez	Gerard	ghernande@classicmodelcars.com	9	Sales Rep
15	1401	Castillo	Pamela	pcastillo@classicmodelcars.com	13	Sales Rep
16	1501	Bott	Larry	lbott@classicmodelcars.com	16	Sales Rep
17	1504	Jones	Barry	bjones@classicmodelcars.com	7	Sales Rep
18	1611	Fixter	Andy	afixter@classicmodelcars.com	11	Sales Rep
19	1612	Marsh	Peter	pmarsht@classicmodelcars.com	10	Sales Rep
20	1619	King	Tom	tking@classicmodelcars.com	15	Sales Rep

Lookup Employee Number

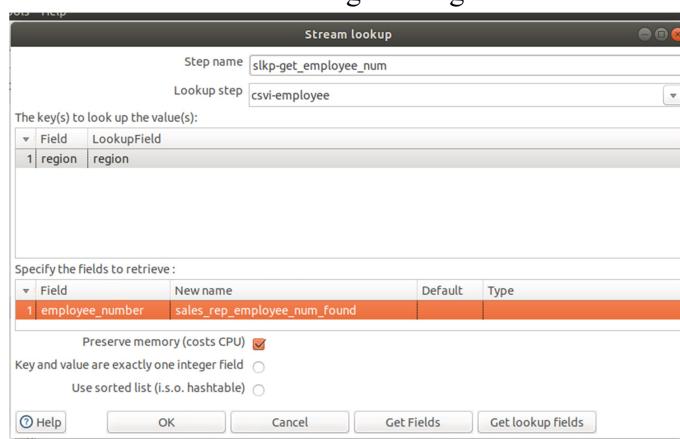
1. Search “Stream Lookup” kemudian drag ke canvas untuk dihubungkan dengan file csvi-employee lalu pilih “main output of step”



2. Hubungkan fr-customer_w_reps dengan Stream lookup dan pilih “Result FALSE”

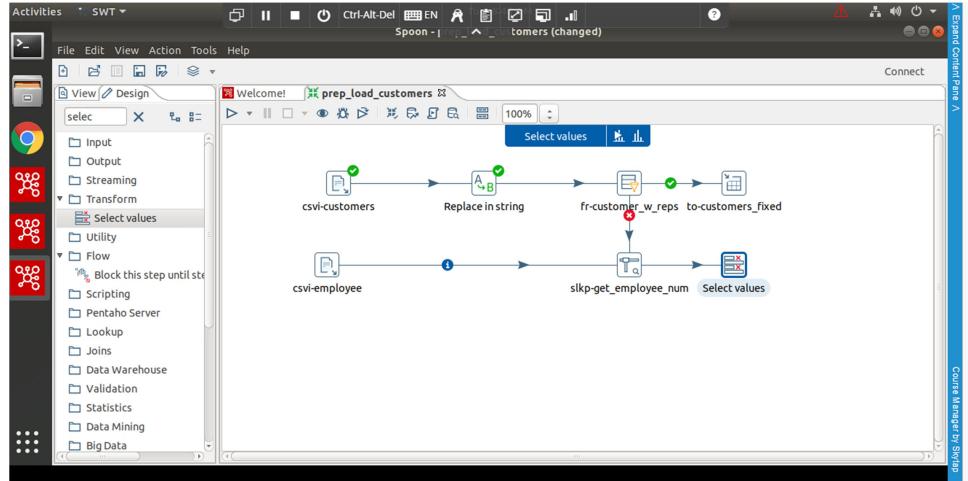


3. Untuk memberitahukan kepada PDI apa yang harus dilakukan kita terlebih dahulu harus menentukan rules dengan mengisi table dari stream lookup.

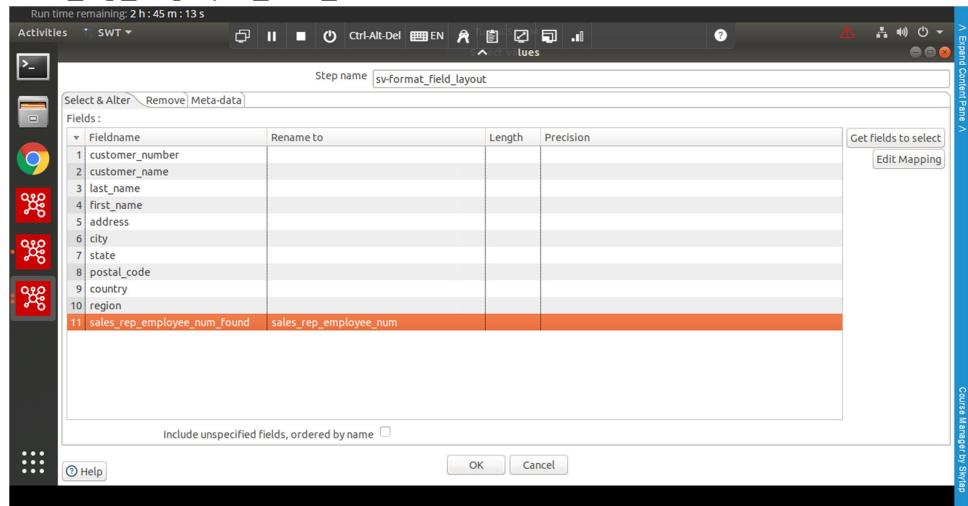


Format Field Layout

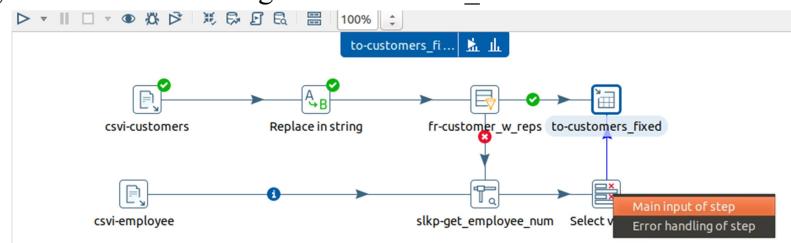
- Untuk memastikan bahwa layout yang ada pada employee_num akan sama dengan to-customer_fixed kita perlu membuat format. Pertama tambahkan “Select Values” pada canvas dan hubungkan dengan slkp-get_employee_num



- Kemudian beri nama pada langkah ini dan klik tombol “Get field to select” hapus table yang tidak diperlukan dan rename pada table sales_rap_employee_num_found

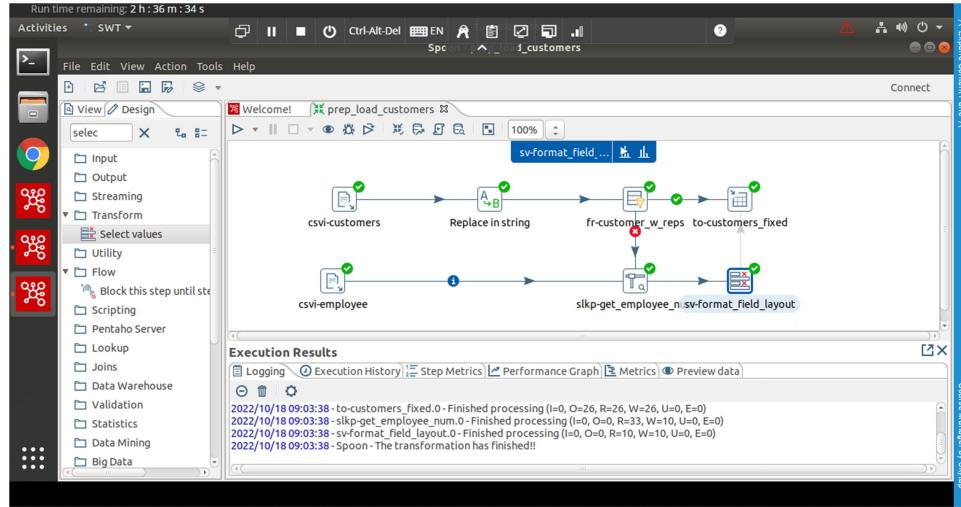


- Hubungkan select values dengan to-customers_fixed



RUNNING TRANSFORMATION

1. Klik button run pada toolbar dan transformation akan berjalan



Terdapat 26 table to-customers_fixed

Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time
skip-get_employee_num	0	26	26	0	26	0	0	0	0	0.0
6 to-customers fixed	0	26	26	0	26	0	0	0	0	Finished 0.0
7 sv-format_field_layout	0	10	10	0	0	0	0	0	0	0.2

Hasil dari table CUSTOMERS_FIXED terdapat 26 rows

	CUSTOMER_NUMBER	CUSTOMER_NAME	LAST_NAME	FIRST_NAME	ADDRESS	CITY	STATE	POSTAL_CODE	CC
1	112	Signal Gift Stores	King	Sue	245 Edgefield Street	Las Vegas	NV	83030	US
2	131	Land of Toys Inc.	Yu	Kwai	9 South Summit Road	NYC	NY	10022	US
3	157	Diecast Classics Inc.	Yu	Kyung	907 Windsor Street	Allentown	PA	70267	US
4	161	Technics Stores Inc.	Hirano	Juri	7473 Bishop Avenue	Burlingame	CA	94217	US
5	168	American Souvenirs Inc	Franco	Sue	541 East Halifax Street	New Haven	CT	97823	US
6	173	Cambridge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51247	US
7	175	Gift Depot Inc.	King	Julie	9920 Pennsylvania Lane	Bridgewater	CT	97562	US
8	198	Auto-Moto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	US
9	204	Online Mini Collectables	Barajas	Miguel	8093 Archer Street	Brickhaven	MA	58339	US
10	205	Toys4GrownUps.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	US
11	286	Marta's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	US
12	319	Mini Classics	Frick	Steve	429 Anchorage Street	White Plains	NY	24067	US
13	320	Mini Creations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA	50553	US
14	321	Corporate Gift Ideas Co.	Brown	Julie	908 Gray Street	San Francisco	CA	94217	US
15	328	Teknik Collectables Inc.	Brown	William	266 Court Drive	Newark	NJ	94019	US
16	339	Classic Gift Ideas, Inc.	Cervantes	Francisica	798 Pinnacle Avenue	Philadelphia	PA	71270	US
17	347	Men R' US Retailers, Ltd.	Chandler	Michael	529 Carriage Street	Los Angeles	CA	91003	US
18	362	Gifts4AllAges.com	Yoshido	Juri	222 Green Lake Street	Boston	MA	51003	US
19	363	Online Diecast Creations Co.	Young	Valarie	7987 Manchester Street	Nashua	NH	62005	US
20	424	Classic Legends Inc.	Hernandez	Maria	8847 West Oceanview Street	NYC	NY	10022	US

CONFIGURATION SUCCESSFULLY