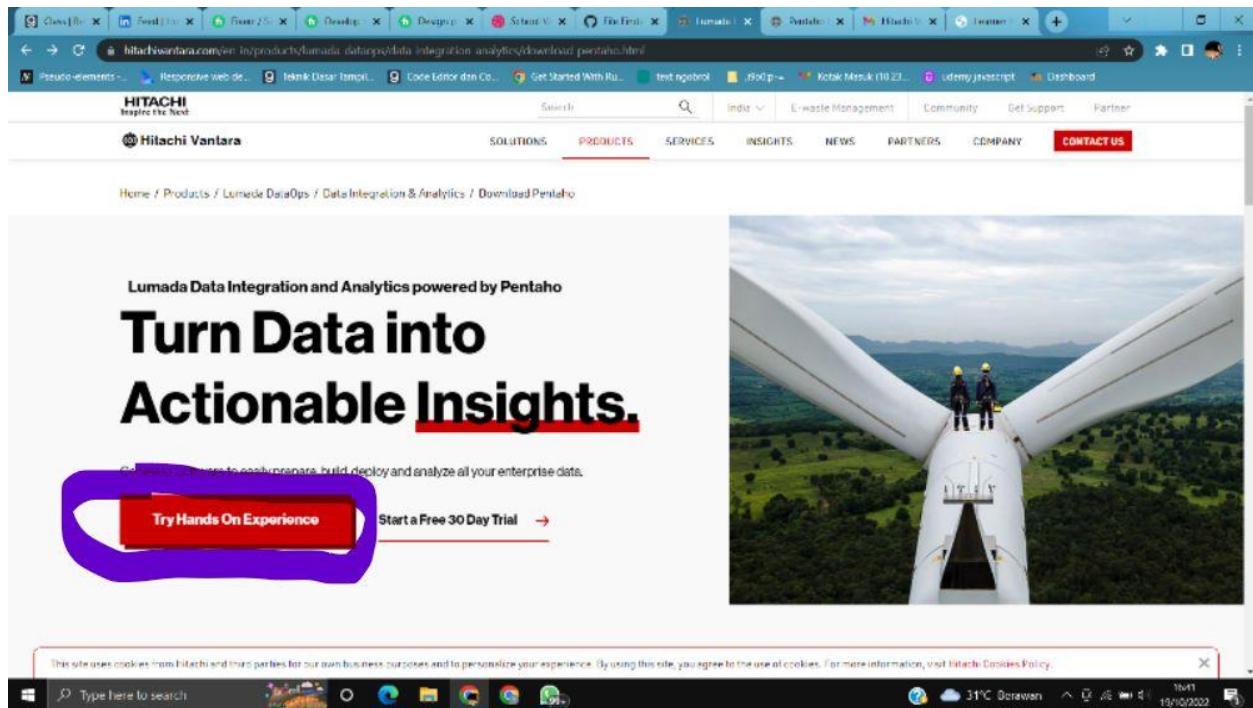


Nama : Yesaya Alehandro Silalahi
NIM : 191402096
KOM : C

Penggunaan Pentaho Melalui Layanan Hitachi

Akses Pentaho melalui URL ini

<https://www.hitachivantara.com/en-in/products/lumada-dataops/data-integration-analytics/download-pentaho.html>



Klik Try Hands On Experience , nanti akan di redirect menuju link
<https://learning.lumada.hitachivantara.com/learning-paths/evaluate-pentaho>

Class | Br... Feed | Un... Fiverr | S... Develop | Design p... School V... File Find... Pentaho | Pentaho | Hitachi | Learner | +

learning.lumada.hitachiavantara.com/courses/pentaho-data-integration-evaluation-pdf1028s

LUMADA

MY DASHBOARD CART (0)

Pentaho Data Integration: Evaluation (PDI1028S)

(60 mins) Self-paced, interactive online training with virtual lab environment for hands-on practice

Start Self-Paced Course with Lab

Description

Get started with Pentaho Data Integration (PDI)!

Read and watch the demonstrations in this course to learn about PDI. Gain hands-on experience in a virtual lab environment and build a data pipeline for a data preparation scenario.

Type here to search 31°C Berawan 16:44 19/10/2022

MILESTONE

Pentaho Evaluation

Pentaho Business Analytics: Evaluation (PBA1007S) ELECTIVE

Self-Paced Course with Lab

(60 mins) Self-paced, interactive online training with virtual lab environment for hands-on practice.

View Details ↗

Pentaho Data Integration: Evaluation (PDI1028S) ELECTIVE

Self-Paced Course with Lab

(60 mins) Self-paced, interactive online training with virtual lab environment for hands-on practice

View Details ↗

PDI Evaluation: Metadata Injection (PDI1030S) ELECTIVE

Self-Paced Course with Lab

(60 mins) Self-paced, interactive online training with virtual lab environment for hands-on practice

View Details ↗

Pilih bagian **Pentaho Data Integration: Evaluation**.

- 1) Dengan mendaftarkan diri pada pembelajaran tersebut, kita akan dialihkan ke halaman pembelajaran. Tersedia juga Lab pembelajaran yang dapat diakses via website untuk penerapan pembelajaran yang ada tanpa melakukan instalasi.

The image contains two side-by-side screenshots of a web-based learning platform. Both screenshots show a header bar with browser tabs and a navigation bar on the left.

Screenshot 1 (Top): The title bar says "Pentaho Data Integration: Evaluation (PDI1028S)". The left sidebar shows a tree view of course content under "01. COURSE": "PDI: Evaluation" (selected), "Introduction", "LAB_PDI1028S" (with a green circular icon), "PDI: Evaluation", and "Completion and Survey". The main content area displays the following text:

Course / PDI: Evaluation

Get started with Pentaho Data Integration (PDI)!

Read and watch the demonstrations in this course to learn about PDI. Gain hands-on experience in a virtual lab environment and build a data pipeline for a data preparation scenario.

This course covers Pentaho v9.2.

Screenshot 2 (Bottom): The title bar says "PDI: Evaluation" and "191403021_DWB/Langkah-Langkah". The left sidebar shows the same course structure as the first screenshot. The main content area displays the following text:

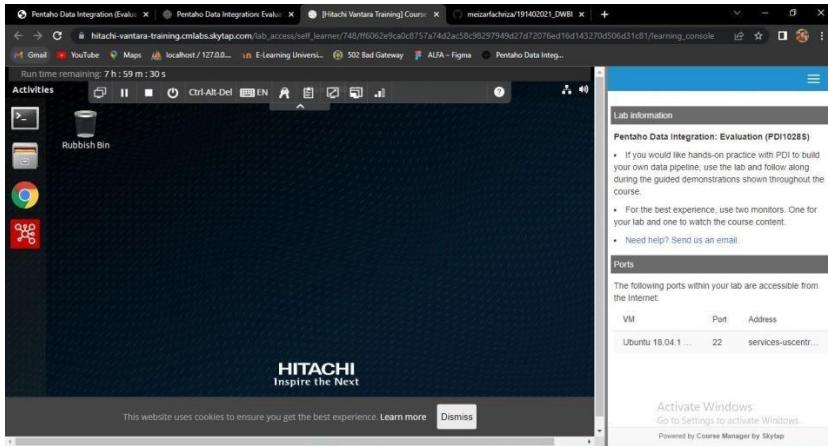
Use the course content and lab to get hands-on experience.

Important information for using the lab

- To access your lab's console, click the **Launch** button, wait up to **5 minutes** for it to be ready, then click **Open Lab**.
- The lab will launch in a new browser tab. Return to the original tab to view the course content and use the lab tab to interact with the lab.
- After you launch the lab for the first time, it will be available for **7 days**.
- Need help? Send us an email.

A modal window titled "Your Lab is Ready" is displayed in the center of the screen, containing the text: "The environment is being prepared for you. Your Lab is Ready. To access your Lab, click the button below." Below the text is a blue "Open Lab" button. At the bottom right of the main content area, there is a "Activate Windows" link.

- 2) Setelah melakukan permintaan terhadap akses pada lab, kita dapat mengakses virtual machine pada tab browser baru.



- 3) Selanjutnya pada bagian PDI: Evaluation untuk memulai pembelajaran dan window browser berisikan materi pembelajaran akan muncul.

The screenshot shows a web browser window with multiple tabs open. The active tab is titled "PDI: Evaluation". The page content includes a sidebar with a "LESSONS" section containing "01. COURSE" and "PDI: Evaluation" sections, and a "NOTES" section. The main content area displays a message: "Please complete this eLearning before continuing to the next page. (will change to after completion.)". Below this is a large red banner with the text "PDI: Evaluation" and a blue button labeled "Open Content In New Window". At the bottom right, there is a "Activate Windows" message.

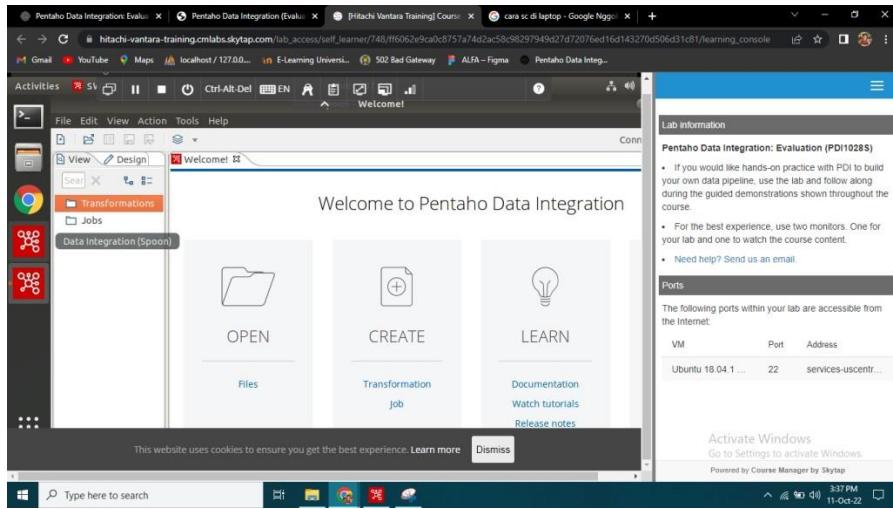
The screenshot shows a course landing page for "PDI Evaluation: Getting Started". It features a red background with a diagonal hatching pattern. A profile picture of "Joshua Morrison" is at the top left. The title "PDI Evaluation: Getting Started" is centered in large white font. Below the title are two buttons: "START COURSE" and "DETAILS". At the bottom, there is a "Activate Windows" message.

The screenshot shows the "Course Introduction" page of the course. It features a red background with a diagonal hatching pattern. The title "Course Introduction" is at the top. Below it is a profile picture of "Joshua Morrison". A descriptive text block says: "A brief look at what this course is all about." At the bottom, there is a "IN THIS COURSE" section with a red header and a black background, listing course modules like "Data Preparation Scenario", "Your First Look at PDI", etc. There is also a "Activate Windows" message.

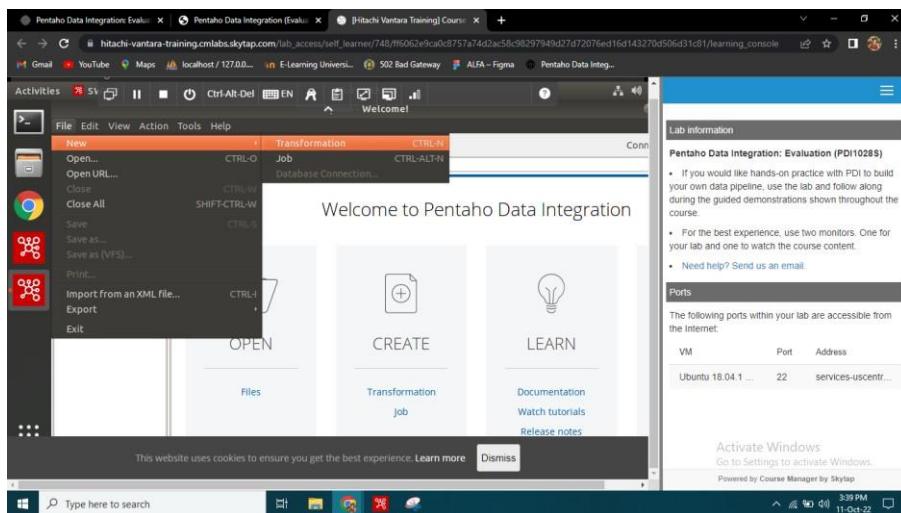
The screenshot shows the "Data Preparation Scenario" module. It features a red background with a diagonal hatching pattern. The title "Data Preparation Scenario" is at the top. Below it is a profile picture of "Joshua Morrison". A descriptive text block says: "A brief look at what this course is all about." At the bottom, there is a "IN THIS COURSE" section with a red header and a black background, listing course modules like "Data Preparation Scenario", "Your First Look at PDI", etc. There is also a "Activate Windows" message.



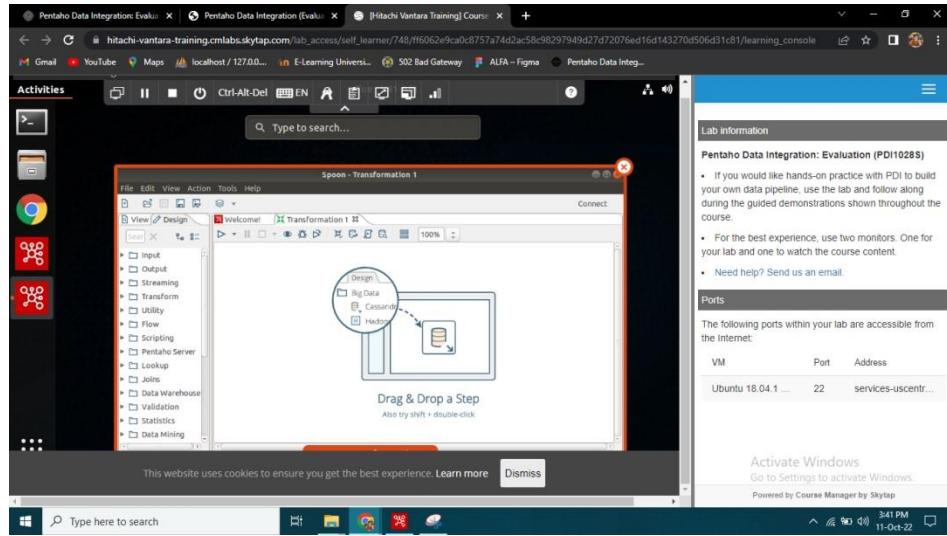
- 1) *Langkah pertama*, kita akan membaca inputan data terlebih dahulu. Kita akan menggunakan data daftar customer. Sebelumnya kita akses terlebih dahulu Spoon pada virtual machine yang sudah tersedia. Spoon merupakan produk dari Hitachi yang bertujuan untuk menangani integrasi data, dikenal sebagai Pentaho Data Integration.



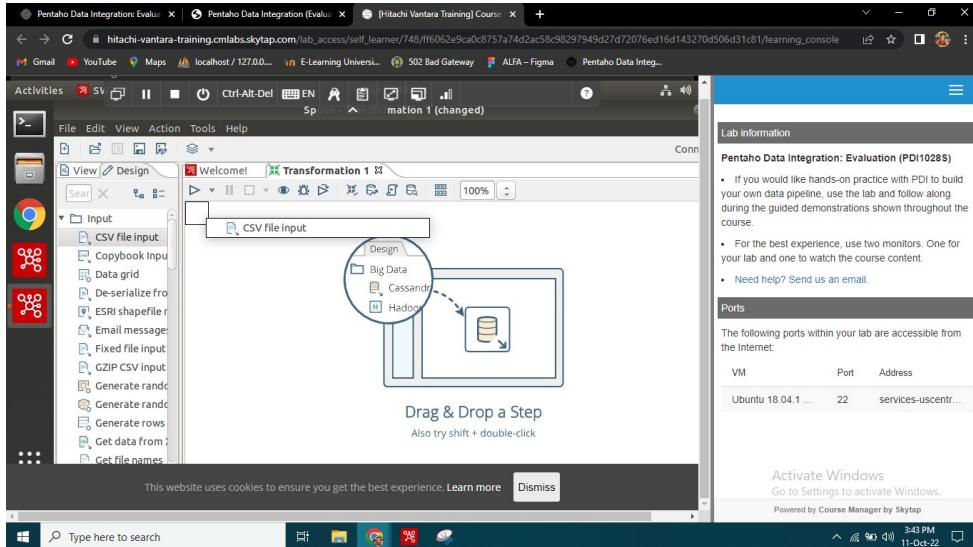
- 2) Pada penerapan kali ini, kita akan mencoba transformasi data pelanggan dari data yang belum bersih menjadi data yang bersih dan siap dipakai dalam database. Pada bagian kiri atas, pilih new kemudian transformation.



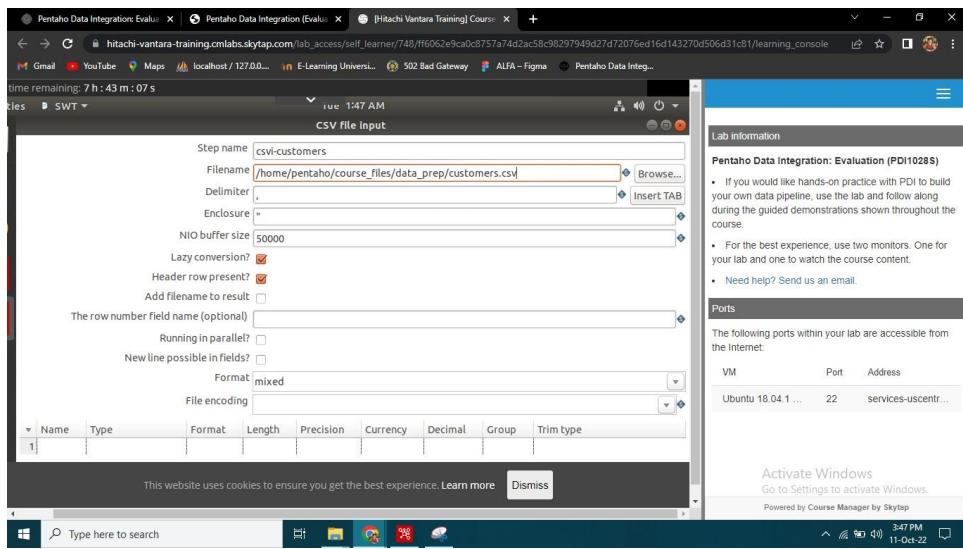
- 3) Kemudian akan muncul lembar kerja baru untuk melakukan transformation data. Lembar kerja baru ini (bagian Drag & Drop a Step pada gambar) dikenal juga sebagai canvas.



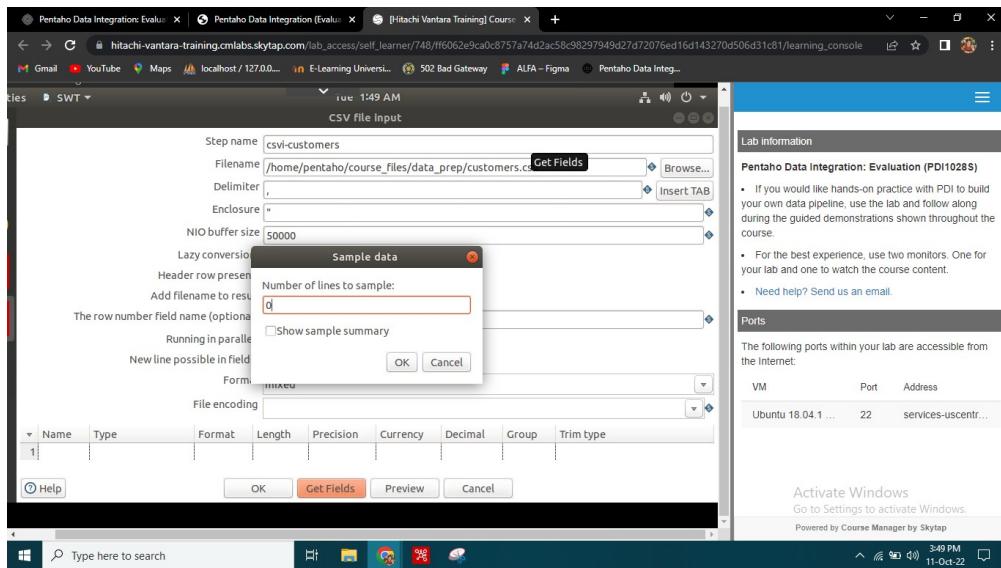
- 4) Kita mulai dengan membuat step untuk input customer data dengan mendrag CSV file input pada tab Design ke dalam canvas. Diikuti dengan double-click pada step tersebut untuk melakukan konfigurasi.



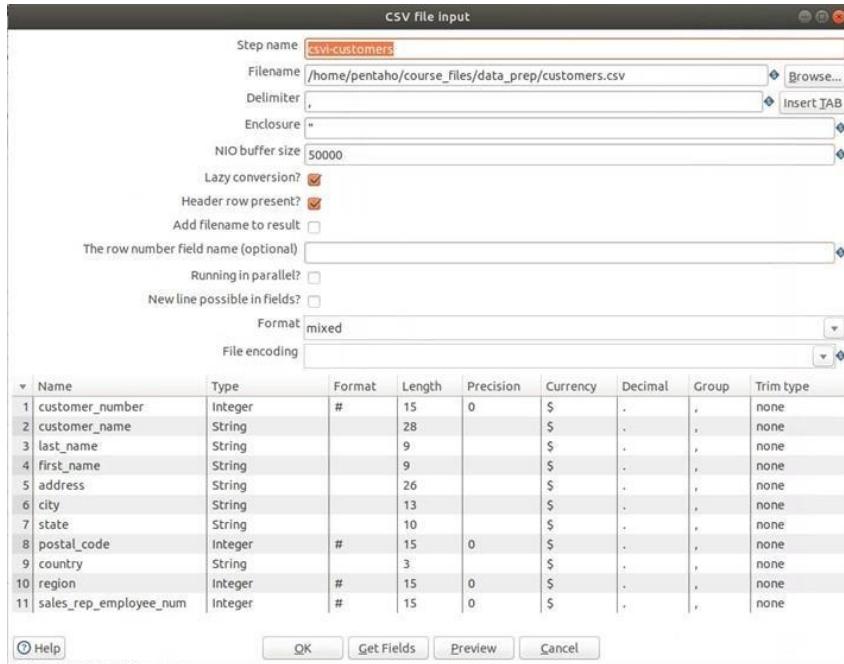
- 5) Konfigurasi yang dilakukan dapat berupa step name, filename sebagai sumber data, delimiter, enclosure, dan sebagainya sebagaimana yang sesuai untuk menginput data.



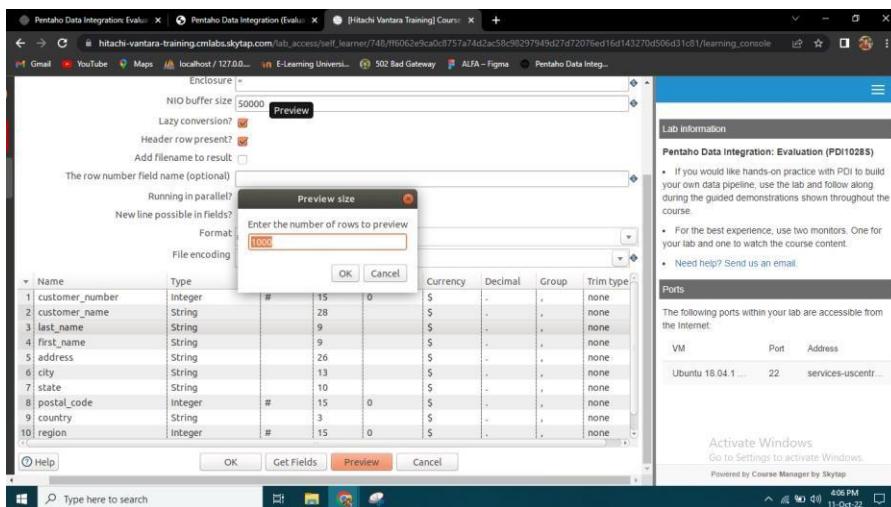
- 6) Kita lanjutkan dengan click pada Get Fields untuk mendapatkan nama kolom yang ada pada input data. Tetapkan 0 pada bagian Number of lines sample untuk mendapatkan nama kolom data secara keseluruhan.



- 7) Maka nama kolom beserta propertinya akan muncul pada bagian bawah dialog konfigurasi step.



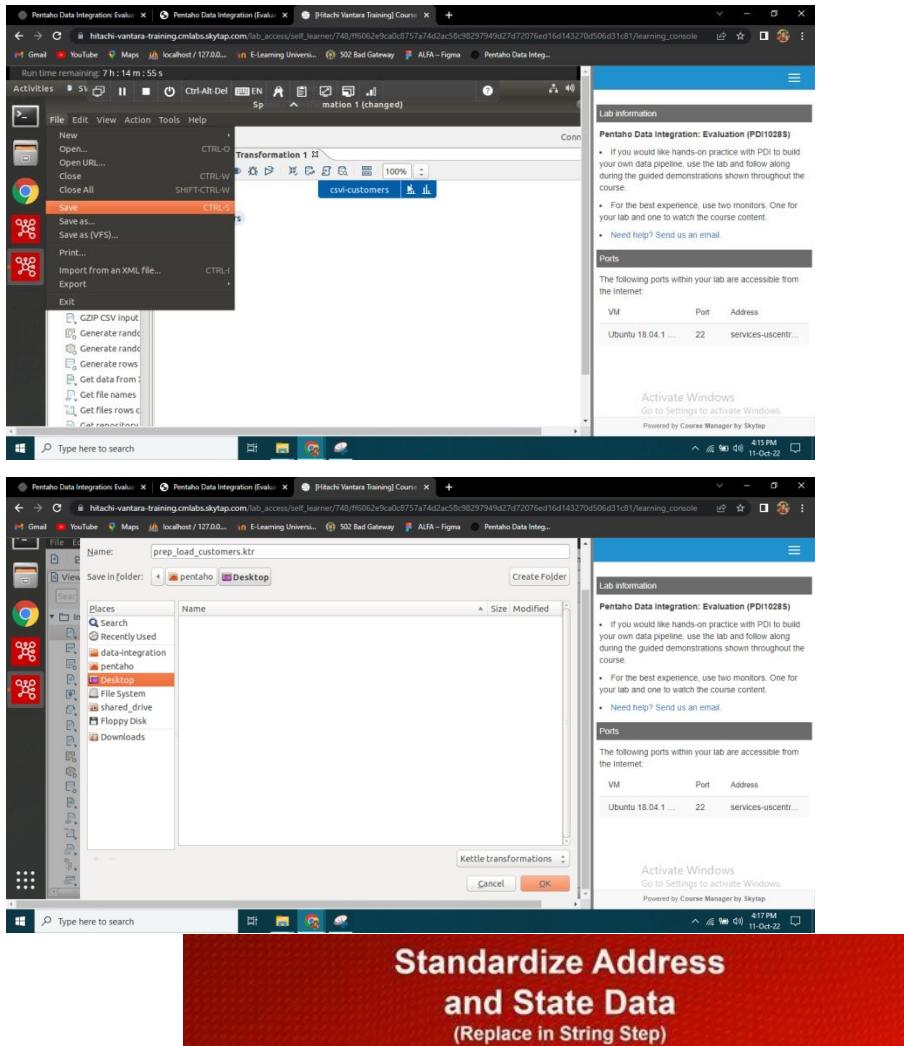
- 8) Untuk melihat contoh data yang tersedia dari file tersebut, klik tombol preview pada bagian bawah dialog dan masukkan preview size sesuai keinginan banyak data yang ditampilkan.



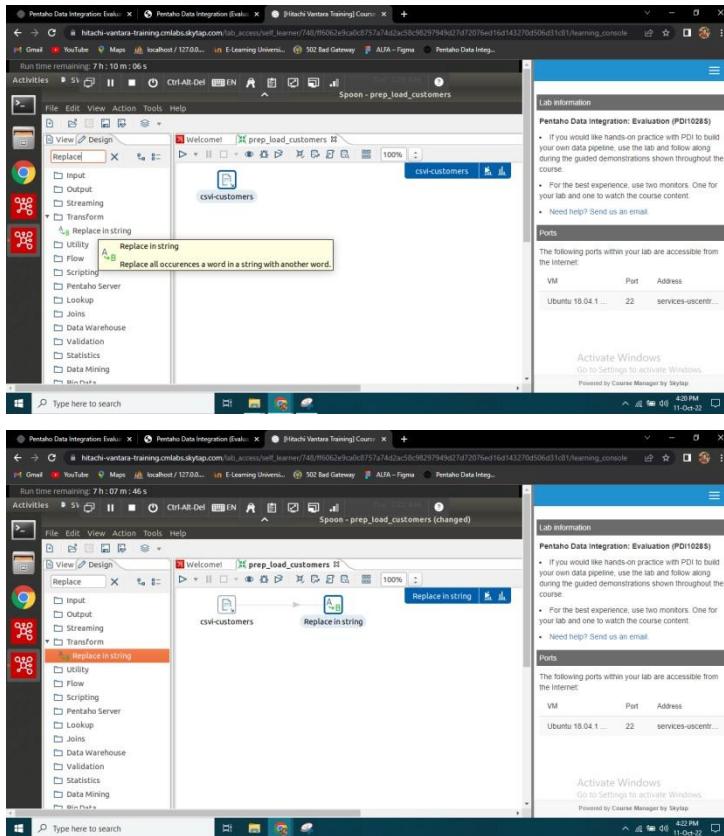
- 9) Akan muncul kotak dialog baru yang memuat data preview dari file input yang ditentukan. Dari preview ini kita dapat mengamati bagian data yang belum bersih seperti masih terdapatnya null pada kolom sales_rep_employee_num, penulisan state dan address yang tidak konsisten dalam penggunaan tanda titik, hingga tidak konsistennya penggunaan singkatan pada kolom address. Jika kita click pada salah satu data kolom state, ternyata masih terdapat white-space yang tidak diperlukan. Untuk itu kita harus membersihkan data agar dapat lebih efisien untuk digunakan.

Examine preview data										
number	customer_name	last_name	first_name	address	city	state	postal_code	country	region	sales_rep_employee_num
112	Signal Gift Stores	King	Sue	245 Edgefield St.	Las Vegas	NV.	89030	USA	14	1166
124	Mini Gifts Distributors Ltd.	Nelson	Valarie	982 Lilac Street	San Rafael	CA	94901	USA	15	<null>
129	Gifts Galore Co.	Murphy	Julie	1324 West Plaza St.	San Francisco	CA	94117	USA	15	<null>
131	Land of Toys Inc.	Yu	Kwei	919 South Main Road	NYC	NY	10021	USA	10	1323
151	Muscle Machine Inc.	Young	Jeff	9788 Union Avenue	NYC	NY	10022	USA	10	<null>
157	Diecast Classics Inc.	Yu	Kyung	907 Windsor St.	Allentown	PA	70267	USA	4	1216
161	Technics Stores Inc.	Hirano	Juri	7473 Bishop Ave.	Burlingame	CA	94117	USA	16	1165
168	American Souvenirs Inc.	Franco	Sue	541 East Halifax Street	New Haven	CT	97823	USA	13	1286
173	Cambridge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51347	USA	5	1188
175	Gift Depot Inc.	King	Julie	9820 Pennsylvania Lane	Bridgewater	CT	97562	USA	3	1323
181	Vitachrome Inc.	Frick	Michael	48 Tower Lane	NYC	NY	10022	USA	10	<null>
198	Auto-Moto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	USA	2	1216
204	Online Mini Collectables	Barajas	Miguel	8093 Archer St.	Brickhaven	MA	58339	USA	2	1188
205	ToysAGoWards.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	USA	16	1166
219	Boards & Toys Co.	Young	Leslie	377 Java Ave.	Glendale	CA	92561	USA	16	<null>
239	Collectable Mini Designs Co.	Thompson	Valarie	341 South Lumber St.	San Diego	CA	91217	USA	17	<null>
286	Marta's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	USA	5	1216
319	Mini Classics	Frick	Steve	429 Anchorage St.	White Plains	NY.	24067	USA	11	1323
320	Mini Creations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA	50553	USA	7	1188
321	Corporate Gift Ideas Co.	Brown	Julie	908 Gray St.	San Francisco	CA	94217	USA	15	1165
328	Tekno Collectables Inc.	Brown	William	266 Court Drive	Newark	NJ	94019	USA	9	1323
339	Classic Gift Ideas, Inc	Cervantes	Francisca	798 Pinnacle Ave.	Philadelphia	PA	71270	USA	12	1188
347	Men 'R' US Retailers, Ltd.	Chandler	Michael	529 Carrige St.	Los Angeles	CA	91003	USA	16	1166
362	Gifts4AllAges.com	Yoshido	Juri	222 Green Lake St.	Boston	MA	51003	USA	1	1216
363	Online Diecast Creations Co.	Young	Valarie	7987 Manchester St.	Nashua	NH	62005	USA	8	1216
379	Collectables For Less Inc.	Nelson	Allen	26 Corporation St.	Brickhaven	MA	58339	USA	2	<null>
424	Classic Legends Inc.	Hernandez	Maria	8847 West Oceanview St.	NYC	NY	10022	USA	10	1286
447	Gift Ideas Corp.	Lewis	Dan	655 Wentworth Street	Glendale	CT	97561	USA	6	1323
450	The Sharp Gifts Warehouse	Frick	Sue	817 Pin Oak St.	San Jose	CA	94217	USA	15	1165
455	Super Scale Inc.	Murphy	Leslie	834 Bank Lane	New Haven	CT	97823	USA	13	1286
456	Microscale Inc.	Kuo	Kee	49 Fairfield Court	NYC	NY	10022	USA	10	1286

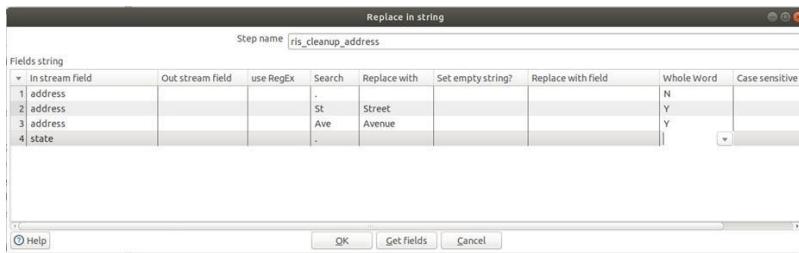
- 10) Setelah menyelesaikan konfigurasi terhadap file input, click OK. Lanjutkan dengan menyimpan perubahan yang kita lakukan pada canvas transformation terlebih dahulu. Pada tab file pada pojok kiri atas, pilih save dan simpan pada lokasi yang diinginkan.



11. Selanjutnya kita akan melakukan standarisasi format untuk beberapa kolom, dalam contoh penerapan ini kolom address dan state. Sebelumnya, cari terlebih dahulu step Replace in string pada tab design. Untuk menghubungkan kedua step, click terlebih dahulu step csvi-customers dan double-click pada step Replace in string, secara otomatis penghubung kedua step akan terbentuk.

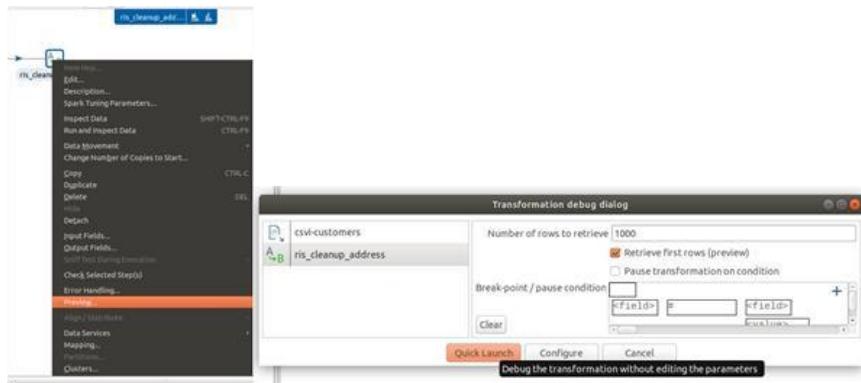


12. Kemudian kita lanjutkan konfigurasi step Replace in string dengan double-click. Standarisasi yang diterapkan berupa menghapus tanda titik pada kolom address dan state, juga mengubah singkatan St dan Ave menjadi Street dan Avenue. Setelah selesai, click OK.



13. Right-click pada step tersebut, pilih preview untuk menampilkan data hasil standarisasi.

Sebelum mengakhiri preview, kita dapat melakukan konfigurasi pada transformation debug dialog. Bila konfigurasi selesai dilakukan, click pada Quick Launch.



14. Dengan demikian, kita dapat melihat hasil standarisasi data customer pada kotak dialog baru. Tidak lupa juga untuk melakukan save progress yang kita lakukan sejauh ini dengan click icon save pada bagian kiri atas aplikasi.

	customer_number	customer_name	last_name	first_name	address	city	state	postal_code	country	region	sales
1	112	Signal Gift Stores	King	Sue	245 Edgefield Street	Las Vegas	NV	83030	USA	14	
2	124	Mini Gifts Distributors Ltd.	Nelson	Valarie	982 Lilac Street	San Rafael	CA	97562	USA	15	
3	129	Mini Wheels Co.	Murphy	Julie	125 Augusta Street	San Francisco	CA	94217	USA	15	
4	131	Land of Toys Inc.	Yu	Kwai	9 South Summit Road	NYC	NY	10022	USA	10	
5	151	Muscle Machine Inc.	Young	Jeff	9788 Union Avenue	NYC	NY	10022	USA	10	
6	157	Diecast Classics Inc.	Yu	Kyung	907 Windsor Street	Allentown	PA	70267	USA	4	
7	161	Technics Stores Inc.	Hirano	Juri	7473 Bishop Avenue	Burlingame	CA	94217	USA	16	
8	168	American Souvenirs Inc.	Franco	Sue	541 East Halifax Street	New Haven	CT	97823	USA	13	
9	173	Cambridge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51247	USA	5	
10	175	Gift Depot Inc.	King	Julie	9920 Pennsylvania Lane	Bridgewater	CT	97562	USA	3	
11	181	Vitachrome Inc.	Frick	Michael	48 Tower Lane	NYC	NY	10022	USA	10	
12	198	Auto-Moto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	USA	2	
13	204	Online Mini Collectables	Barajas	Miguel	8093 Archer Street	Brickhaven	MA	58339	USA	2	
14	205	Toys4GrownUps.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	USA	16	
15	219	Boards & Toys Co.	Young	Leslie	377 Java Avenue	Glendale	CA	92561	USA	16	
16	239	Collectable Mini Designs Co.	Thompson	Valarie	341 South Lumber Street	San Diego	CA	92127	USA	17	
17	286	Marta's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	USA	5	
18	319	Mini Classics	Frick	Steve	429 Achorage Street	White Plains	NY	24067	USA	11	
19	320	Mini Creations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA	50553	USA	7	
20	321	Corporate Gift Ideas Co.	Brown	Julie	908 Gray Street	San Francisco	CA	94217	USA	15	
21	328	Tekni Collectables Inc.	Brown	William	266 Court Drive	Newark	NJ	94019	USA	9	
22	339	Francisca	Cervantes	Francisca	798 Pinnacle Avenue	Philadelphia	PA	71270	USA	12	
23	347	Men 'R' US Retailers, Ltd.	Chandler	Michael	529 Carrige Street	Los Angeles	CA	91003	USA	16	
24	362	Gifts4Allages.com	Yoshida	Juri	222 Green Lake Street	Boston	MA	51003	USA	1	
25	363	Online Diecast Creations Co.	Young	Valarie	7987 Manchester Street	Nashua	NH	62005	USA	8	
26	379	Collectables For Less Inc.	Nelson	Allen	26 Corporation Street	Brickhaven	MA	58339	USA	2	
27	424	Classic Legends Inc.	Hernandez	Maria	8847 West Oceanview Street	NYC	NY	10022	USA	10	
28	447	Gift Ideas Corp.	Lewis	Dan	655 Wentworth Street	Glendale	CT	97561	USA	6	
29	450	The Sharp Gifts Warehouse	Frick	Sue	817 Pin Oak Street	San Jose	CA	94217	USA	15	
30	455	Super Scale Inc.	Murphy	Leslie	834 Bank Lane	New Haven	CT	97823	USA	13	
31	456	Microscale Inc.	Kuo	Kee	43 Fairfield Court	NYC	NY	10022	USA	10	

Close

File Edit View Action Tools Help

Save current file

View Design

replace

Input Output Streaming Transform Replace in string

Welcome! prep_load_customers

100%

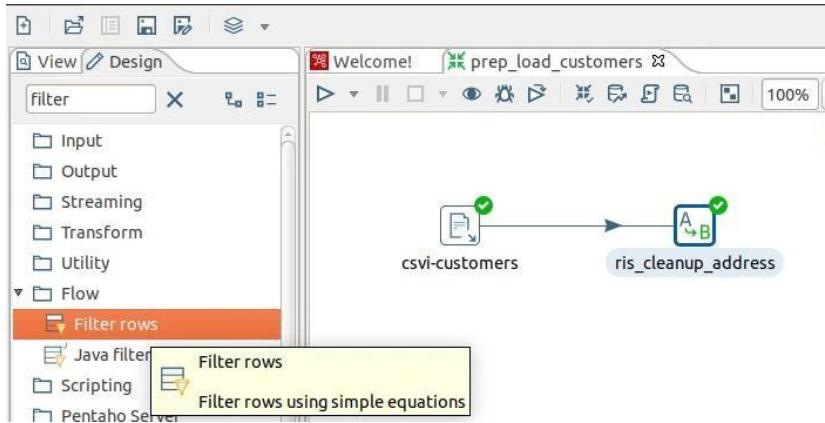
ris_cleanup_address

csv-customers

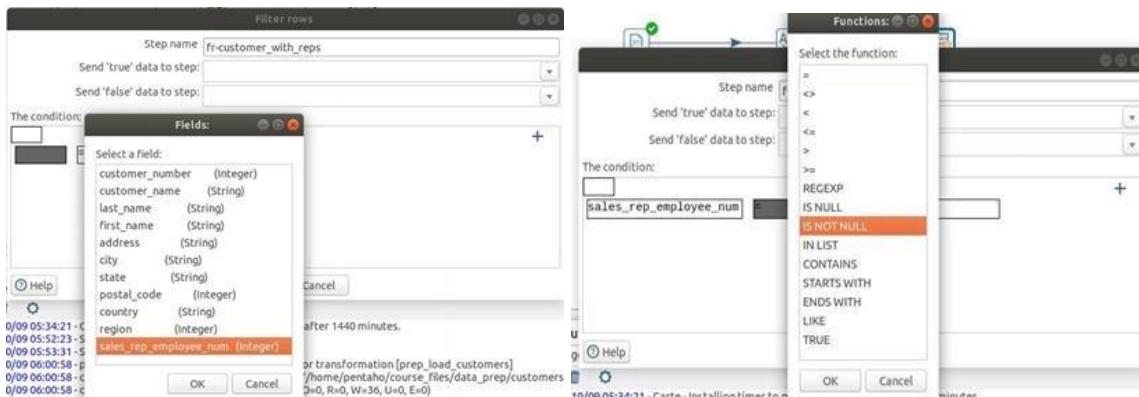
Diagram showing the flow from 'csv-customers' to 'ris_cleanup_address'.

Filter Customers (Filter Rows Step)

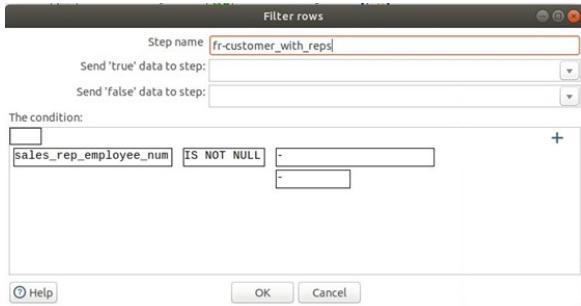
15. Selanjutnya melakukan filtrasi terhadap data terstandarisasi. Pada penerapan kali ini kita akan memfiltrasi kolom sales_rep_employee_num. Sebelumnya, cari step Filter rows pada tab Design terlebih dahulu dan double-click. Pastikan step ini terhubung dengan step ris_cleanup_address.



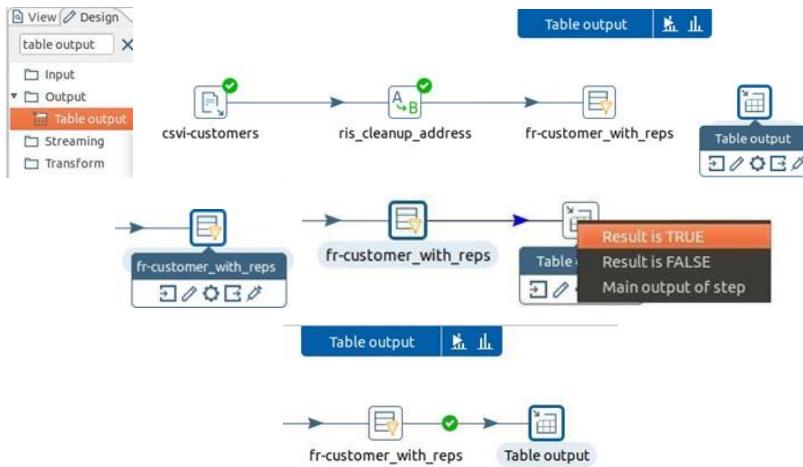
16. Double-click pada step tersebut untuk memulai konfigurasi. Kali ini kita kosongkan saja bagian send 'true' dan 'false' data to step karena kita akan menambahkan penghubung step secara manual nantinya. Beralih ke kotak pada baris kedua bagian condition dan pilih nama kolom yang akan difiltrasi, kolom sales_rep_employee_num. Beralih ke kotak di sebelah kanan, kotak ini bertujuan untuk menentukan function yang akan diterapkan pada nama kolom, pilih IS NOT NULL.



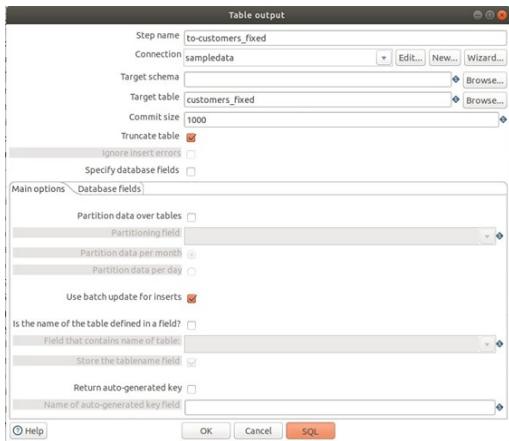
17. Aturan yang ditetapkan pada konfigurasi ini bertujuan untuk filtrasi kolom sales_rep_employee_num yang bernilai null. Melalui step ini, kita dapat mengalihkan data yang memiliki nilai pada kolom tersebut ke step selanjutnya dan data dengan nilai null pada kolom untuk diproses lebih lanjut pada step lainnya.



18. Setelah menyelesaikan konfigurasi, click OK. Kita lanjutkan untuk penanganan kondisi/aturan true untuk filtrasi. Pada tab Design, cari step Table output dan drag ke dalam canvas. Hover pada step filter dan pilih icon kedua dari kanan untuk menarik garis penghubung ke step Table output dan pilih Result is TRUE.



19. Lalu kita double-click pada Table output untuk memulai konfigurasi. Pada dialog konfigurasi ini, kita juga dapat melihat SQL yang tercipta dengan click pada tombol SQL pada bagian bawah.



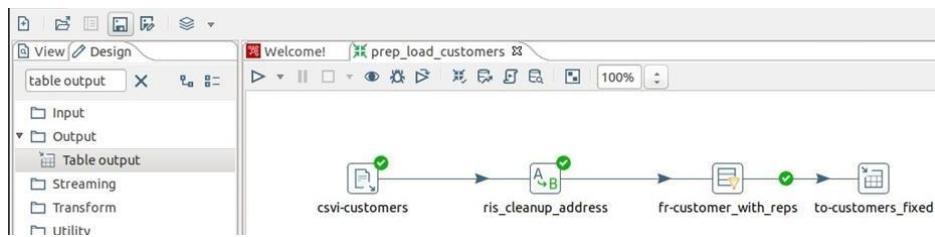
20. Melalui dialog Simple SQL Editor yang muncul setelah click SQL, kita dapat melihat hasil Query berdasarkan konfigurasi sebelumnya. Untuk menjalankan Query, click tombol Execute dan table customers_fixed pada penerapan contoh kali ini akan terbentuk pada database sampledata beserta struktur-struktur tablenya.

Two screenshots of the Oracle SQL Developer interface. On the left, the 'Simple SQL editor' window shows the SQL code for creating the 'customers_fixed' table:

```
CREATE TABLE customers_fixed
(
    customer_number BIGINT,
    customer_name VARCHAR(28),
    last_name VARCHAR(9),
    first_name VARCHAR(9),
    address VARCHAR(26),
    city VARCHAR(13),
    state VARCHAR(10),
    postal_code BIGINT,
    country VARCHAR(3),
    region BIGINT,
    sales_rep_employee_num BIGINT
);
```

 Below the code, it says 'Line 1 column 0'. At the bottom are 'Execute', 'Clear cache', and 'Close' buttons. On the right, the 'Results of the SQL statements' window shows the confirmation message: 'The SQL statements had the following results' and 'SQL executed: CREATE TABLE customers_fixed'. It also indicates '1 SQL statements executed'. At the bottom are 'OK' and 'Cancel' buttons.

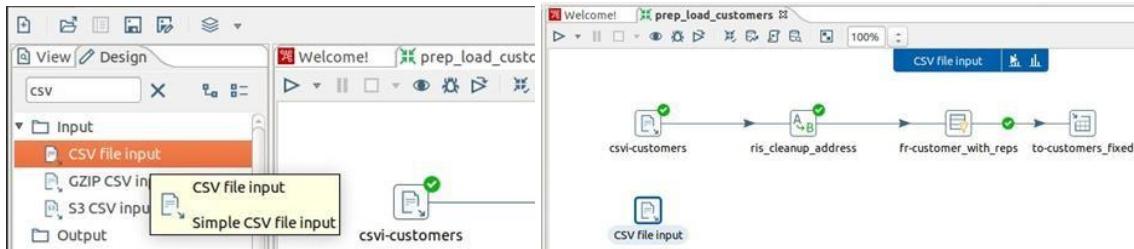
21. Table dan struktur yang terbentuk akan kita lihat pada penerapan selanjutnya. Setelah selesai konfiguras, tidak lupa untuk kita OK dan close. Tidak lupa juga untuk menyimpan progress yang telah dikerjakan dengan click icon save pada bagian kiri atas aplikasi ini.



1)

Reading Employee Data (CSV File Input Step)

Selanjutnya kita akan menginput file employee data sebagai dasar penanganan untuk bagian false pada filtrasi data bernilai null. Sama seperti cara sebelumnya, cari step CSV file input dan drag ke dalam canvas.



22. Lakukan konfigurasi terhadap file input dengan double-click pada step ini dan jangan lupa untuk Get Fields dan Preview pada data employee ini. Pada penerapan ini, file input data employee sudah bersih untuk kolom tertentu yang diperlukan saja.

Step name: csvi-employees
Filename: /home/pentaho/course_files/data_prep/employees.csv
Format: mixed
File encoding: UTF-8

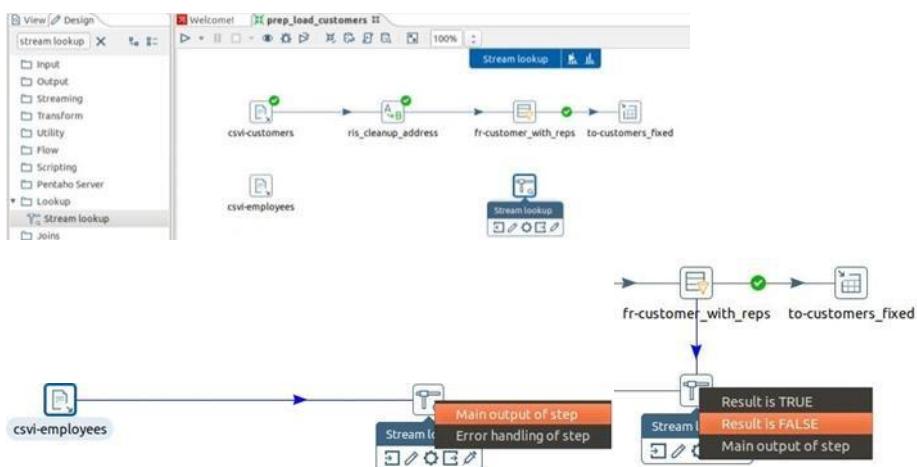
Name	Type	Format	Length	Precision	Currency	Decimal	Group	Trim type
1 employee_number	Integer	#	15	0	\$,	,	none
2 last_name	String		9		\$,	,	none
3 first_name	String		8		\$,	,	none
4 email	String		31		\$,	,	none
5 region	Integer	#	15	0	\$,	,	none
6 job_title	String		27		\$,	,	none

Examine preview data

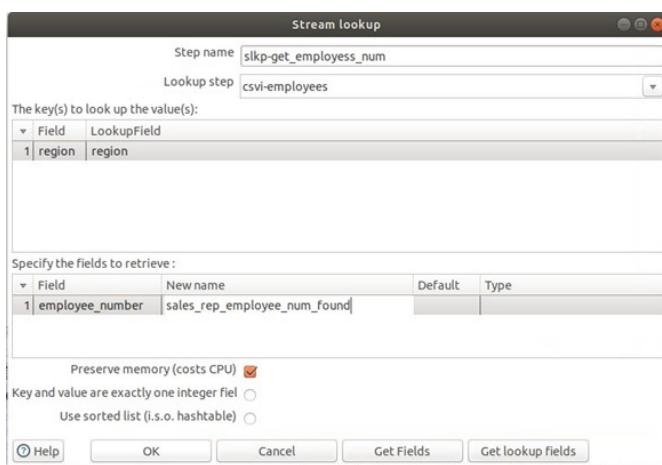
	employee_number	last_name	first_name	email	region	job_title
1	1002	Murphy	Diane	dmurphy@classicmodelcars.com	<null>	President
2	1056	Patterson	Mary	mpatterso@classicmodelcars.com	<null>	VP Sales
3	1076	Firrelli	Jeff	jfirrelli@classicmodelcars.com	<null>	VP Marketing
4	1088	Patterson	William	wpatterson@classicmodelcars.com	<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

Lookup Employee Number (Stream Lookup Step)

23. Selanjutnya, cari step Stream lookup pada tab Design dan drag ke dalam canvas. Jangan lupa untuk menghubungkan csvi-employees dengan step ini, dapat dilakukan dengan menahan Shift dan menarik garis dari step csvi menuju step ini, atur sebagai Main output of step. Hubungkan juga step filtrasi sebelumnya (fr-customer_with_reps) dengan step ini dengan cara yang sama dan pilihan Result is FALSE.



24. Double-click pada step dan lakukan beberapa konfigurasi. Pada step ini, data akan mengacu pada kolom region sebagai lookup value untuk data employee dan customer (yang tersaring filtrasi sebagai false karena nilai sales_rep_employee_num null). Juga menetapkan employee_number dari data employee sebagai sales_rep_employee_num_found yang nantinya akan menggantikan nilai null agar data customer terintegrasi. Jangan lupa untuk click OK dan save progress setelah selesai konfigurasi.

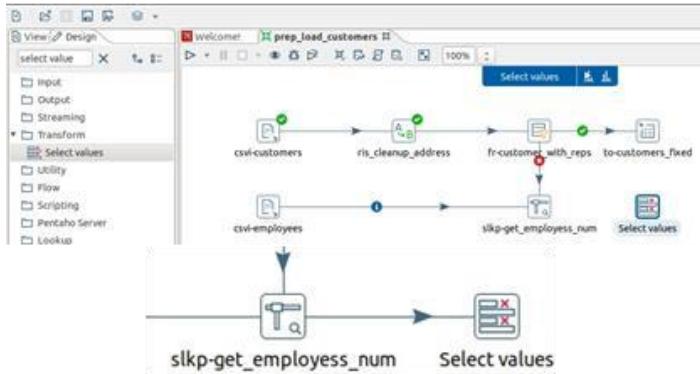


1)

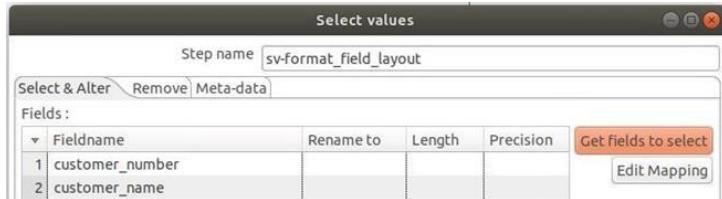
Format Field Layout

(Select Values Step)

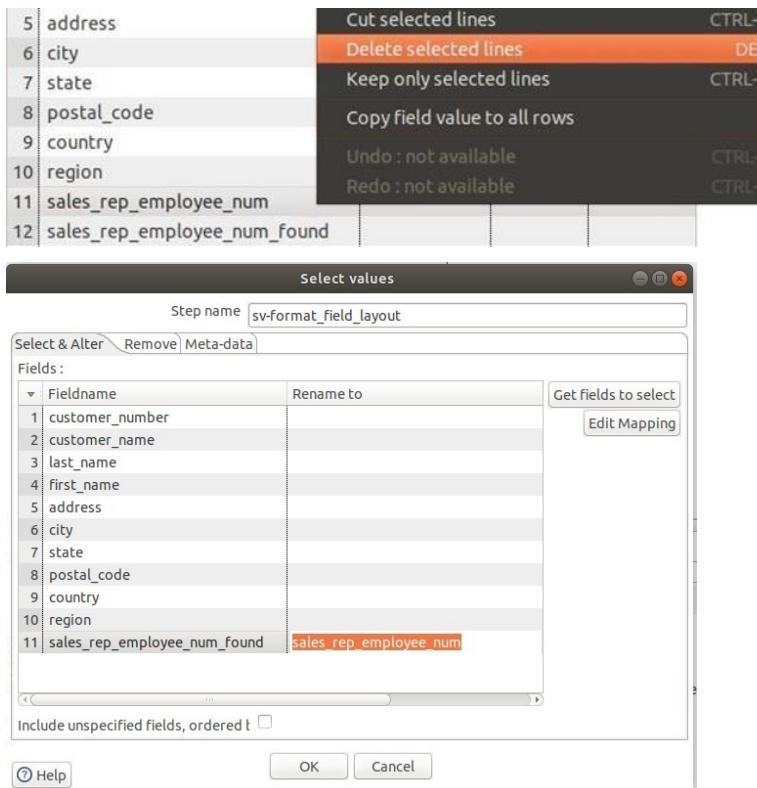
25. Selanjutnya, cari step Select Values pada tab Design dan drag pada canvas, serta hubungkan step stream lookup sebelumnya dengan step ini



26. Double-click pada step untuk memulai konfigurasi. Silahkan click tombol Get fields to select bila tidak tertera daftar nama kolom (fields) pada tampilan konfigurasi.

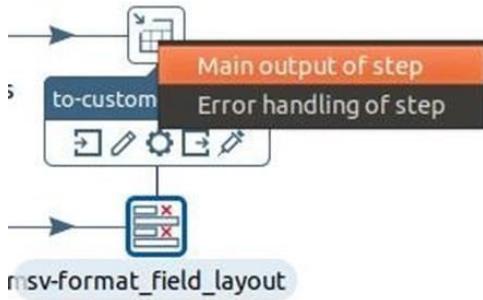


27. Cari nama kolom sales_rep_employee_num dan right-click padanya, kemudian Delete selected lines. Kolom tersebut dihapus karena memuat nilai null saja teruntuk data yang tidak lulus step filtrasi. Dan juga cari kolom sales_rep_employee_num_found dan isi pada kolom rename di sebelah kanannya dengan sales_rep_employee_num. Hal ini bertujuan untuk menetapkan employee_number berdasarkan lookup sebelumnya sebagai nilai baru untuk kolom sales_rep_employee_number dan menyesuaikan dengan kolom-kolom yang ada pada table customer_fixed.

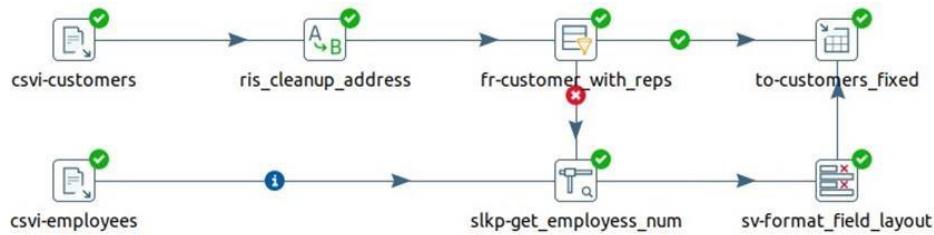


1)

28. Setelah konfigurasi selesai, click OK. Hubungkan juga step ini dengan step table output sebelumnya (to-customer_fixed) dan pilih Main output of step. Jangan lupa untuk melakukan save progress yang ada dengan click icon save pada kiri atas aplikasi ini.

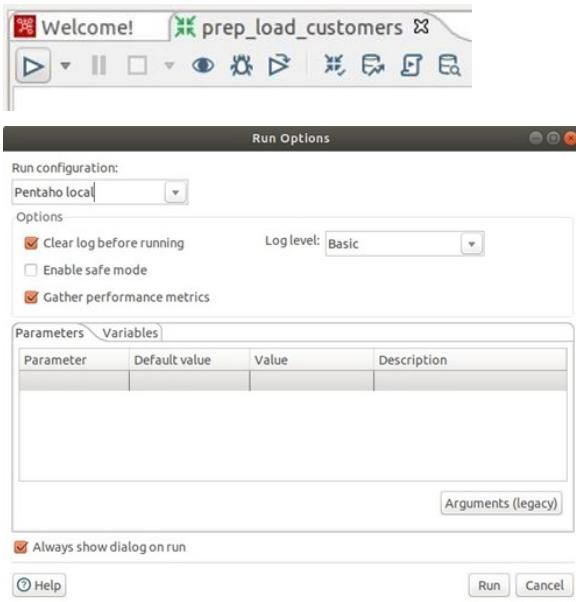


29. Susunan akhir canvas transformation:



Running Your Transformation

Setelah melalui penerapan sebelumnya, transformation yang dibentuk untuk membersihkan dan mengintgrasi data customer telah selesai. Untuk menjalankan hasil transformasi, click icon play pada bagian kiri atas canvas transformation. Akan muncul kotak dialog Run Options, click Run setelah selesai konfigurasi.



30. Proses dan hasil run transformation akan muncul pada bagian bawah aplikasi pada bagian Execution Results tab Logging. Silahkan cek lagi bila terdapat error pada logging ini dan belum mendapatkan indikasi Spoon – The transformation has finished.

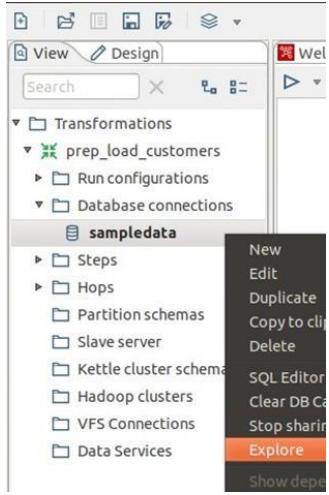
Execution Results										
	Logging	Execution History	Step Metrics	Performance Graph	Metrics	Preview data				
	<pre>2022/10/09 06:48:17 - csvi-customers.0 - Header row skipped in file '/home/pentaho/course_files/data_prep/employees.csv' 2022/10/09 06:48:17 - csvi-customers.0 - Header row skipped in file '/home/pentaho/course_files/data_prep/customers.csv' 2022/10/09 06:48:17 - csvi-customers.0 - Finished processing (I=24, O=0, R=0, W=23, U=0, E=0) 2022/10/09 06:48:17 - csvi-customers.0 - Finished processing (I=37, O=0, R=0, W=36, U=0, E=0) 2022/10/09 06:48:17 - ris_cleanup_address.0 - Finished processing (I=0, O=0, R=36, W=36, U=0, E=0) 2022/10/09 06:48:17 - fr-customer_with_reps.0 - Finished processing (I=0, O=0, R=36, W=36, U=0, E=0) 2022/10/09 06:48:17 - slkp-get_employees_num.0 - Finished processing (I=0, O=0, R=33, W=10, U=0, E=0) 2022/10/09 06:48:17 - sv-format_field_layout.0 - Finished processing (I=0, O=0, R=10, W=10, U=0, E=0) 2022/10/09 06:48:17 - to-customers_fixed.0 - Finished processing (I=0, O=36, R=36, W=36, U=0, E=0) 2022/10/09 06:48:17 - Spoon - The transformation has finished!</pre>									

31. Tab Step Metrics bertujuan untuk menunjukkan pemrosesan data pada setiap step, sebagai contoh pada step to-customers_fixed memiliki property output bernilai 36, yang artinya terdapat 36 data yang masuk ke dalam table customer_fixed pada database sampledata yang sudah dibentuk.

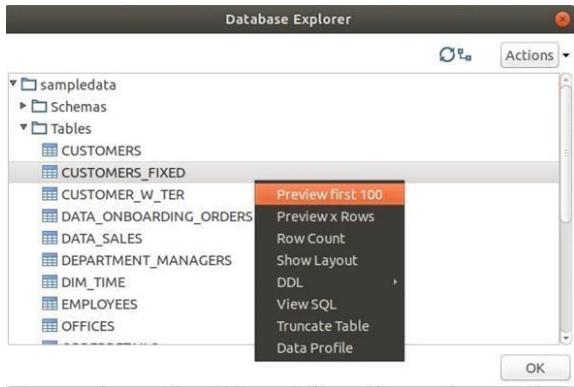
Execution Results																																																																																																									
	Logging	Execution History	Step Metrics	Performance Graph	Metrics	Preview data																																																																																																			
			<table border="1"> <thead> <tr> <th>Stepname</th> <th>Copynr</th> <th>Read</th> <th>Written</th> <th>Input</th> <th>Output</th> <th>Updated</th> <th>Rejected</th> <th>Errors</th> <th>Active</th> <th>Time</th> <th>Speed (r/s)</th> </tr> </thead> <tbody> <tr> <td>1 csvi-customers</td> <td>0</td> <td>0</td> <td>36</td> <td>37</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.05</td> <td>5,286</td> </tr> <tr> <td>2 ris_cleanup_address</td> <td>0</td> <td>36</td> <td>36</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.05</td> <td>2,118</td> </tr> <tr> <td>3 csvi-employees</td> <td>0</td> <td>0</td> <td>23</td> <td>24</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.05</td> <td>4,000</td> </tr> <tr> <td>4 fr-customer_with_reps</td> <td>0</td> <td>36</td> <td>36</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.05</td> <td>1,800</td> </tr> <tr> <td>5 slkp-get_employees_num</td> <td>0</td> <td>33</td> <td>10</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.25</td> <td>205</td> </tr> <tr> <td>6 sv-format_field_layout</td> <td>0</td> <td>10</td> <td>10</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.25</td> <td>61</td> </tr> <tr> <td>7 to-customers_fixed</td> <td>0</td> <td>36</td> <td>36</td> <td>0</td> <td>36</td> <td>0</td> <td>0</td> <td>0</td> <td>Finished</td> <td>0.25</td> <td>199</td> </tr> </tbody> </table>	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)	1 csvi-customers	0	0	36	37	0	0	0	0	Finished	0.05	5,286	2 ris_cleanup_address	0	36	36	0	0	0	0	0	Finished	0.05	2,118	3 csvi-employees	0	0	23	24	0	0	0	0	Finished	0.05	4,000	4 fr-customer_with_reps	0	36	36	0	0	0	0	0	Finished	0.05	1,800	5 slkp-get_employees_num	0	33	10	0	0	0	0	0	Finished	0.25	205	6 sv-format_field_layout	0	10	10	0	0	0	0	0	Finished	0.25	61	7 to-customers_fixed	0	36	36	0	36	0	0	0	Finished	0.25	199						
Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)																																																																																														
1 csvi-customers	0	0	36	37	0	0	0	0	Finished	0.05	5,286																																																																																														
2 ris_cleanup_address	0	36	36	0	0	0	0	0	Finished	0.05	2,118																																																																																														
3 csvi-employees	0	0	23	24	0	0	0	0	Finished	0.05	4,000																																																																																														
4 fr-customer_with_reps	0	36	36	0	0	0	0	0	Finished	0.05	1,800																																																																																														
5 slkp-get_employees_num	0	33	10	0	0	0	0	0	Finished	0.25	205																																																																																														
6 sv-format_field_layout	0	10	10	0	0	0	0	0	Finished	0.25	61																																																																																														
7 to-customers_fixed	0	36	36	0	36	0	0	0	Finished	0.25	199																																																																																														

1)

32. Untuk melihat hasil transformation pada table, pilih tab View dan expand folder Transformations, prep_load_customers (nama transformation), Database connections. Pilih database sampledata dan right-click serta pilih explore.



33. Setelah muncul kotak dialog Database Explorer, right-click pada CUSTOMER_FIXED pada bagian Tables and pilih Preview first 100. Kemudian akan tampil data hasil transformation yang sudah masuk ke dalam table tersebut.



Examine preview data

Rows of step: CUSTOMERS_FIXED (36 rows)

IMER_NAME	LAST_NAME	FIRST_NAME	ADDRESS	CITY	STATE	POSTAL_CODE	COUNTRY	REGION	SALES_REP_EMPLOYEE_NUM
Gift Stores	King	Sue	245 Edgefield Street	Las Vegas	NV	83030	USA	14	1166
If Toys Inc.	Yu	Kwai	9 South Summit Road	NYC	NY	10022	USA	10	1323
t Classics Inc.	Yu	Kyung	907 Windsor Street	Allentown	PA	70267	USA	4	1216
cs Stores Inc.	Hirano	Juri	7473 Bishop Avenue	Burlingame	CA	94217	USA	16	1165
can Souvenirs Inc	Franco	Sue	541 East Halifax Street	New Haven	CT	97823	USA	13	1286
idge Collectables Co.	Tseng	Kyung	271 South Moon Drive	Cambridge	MA	51247	USA	5	1188
spot Inc.	King	Julie	9920 Pennsylvania Lane	Bridgewater	CT	97562	USA	3	1323
Auto Classics Inc.	Taylor	Leslie	8132 East Riverside Street	Brickhaven	MA	58339	USA	2	1216
Mini Collectables	Barajas	Miguel	8093 Archer Street	Brickhaven	MA	58339	USA	2	1188
GrownUps.com	Young	Julie	663 Blue Spring Lane	Pasadena	CA	90003	USA	16	1166
's Replicas Co.	Hernandez	Marta	904 Meadowbrook Street	Cambridge	MA	51247	USA	5	1216
lassics	Frick	Steve	429 Anchorage Street	White Plains	NY	24067	USA	11	1323
reations Ltd.	Tam	Wing C	357 John Drive	New Bedford	MA	50553	USA	7	1188
rate Gift Ideas Co.	Brown	Julie	908 Gray Street	San Francisco	CA	94217	USA	15	1165
Collectables Inc.	Brown	William	266 Court Drive	Newark	NJ	94019	USA	9	1323
Gift Ideas, Inc.	Cervantes	Francisca	798 Pinnacle Avenue	Philadelphia	PA	71270	USA	12	1188
' US Retailers, Ltd.	Chandler	Michael	529 Carriage Street	Los Angeles	CA	91003	USA	16	1166
AllAges.com	Yoshido	Juri	222 Green Lake Street	Boston	MA	51003	USA	1	1216
Diecast Creations Co.	Young	Valarie	7987 Manchester Street	Nashua	NH	62005	USA	8	1216
Legends Inc.	Hernandez	Maria	8847 West Oceanview Street	NYC	NY	10022	USA	10	1286
eas Corp.	Lewis	Dan	655 Wentworth Street	Glendale	CT	97561	USA	6	1323
arp Gifts Warehouse	Frick	Sue	817 Pin Oak Street	San Jose	CA	94217	USA	15	1165
Scale Inc.	Murphy	Leslie	834 Bank Lane	New Haven	CT	97823	USA	13	1286
cale Inc.	Kuo	Kee	43 Fairfield Court	NYC	NY	10022	USA	10	1286
Coast Collectables Co.	Thompson	Steve	385 Windsor Drive	Burbank	CA	94019	USA	16	1166
Mint Distributors Inc.	Hernandez	Rosa	9296 South Wharf Avenue	Philadelphia	PA	71270	USA	12	1323
Ifts Distributors Ltd.	Nelson	Valarie	982 Lila Street	San Rafael	CA	97562	USA	15	1619
heels Co.	Murphy	Julie	125 Augusta Street	San Francisco	CA	94217	USA	15	1619
Machine Inc	Young	Jeff	9788 Union Avenue	NYC	NY	10022	USA	10	1612
rome Inc.	Frick	Michael	48 Tower Lane	NYC	NY	10022	USA	10	1612
s & Toys Co.	Young	Leslie	377 Java Avenue	Glendale	CA	92561	USA	16	1501

Close

34. Dengan demikian, Hands On Experience untuk Pentaho Data Integration: Evaluation bagian Getting Started telah selesai

PDI Evaluation: Getting Started

100% COMPLETE

Course Introduction

Data Preparation Scenario

Your First Look at PDI

Take a Peak at the Solution

Hands-On Building Your Transformation

OVERACHIEVER

If you want to do even more on your own, consider using your lab environment to enhance your transformation. Here are some things you could do:

- Add notes on the canvas to tell others what the transformation does and how it works
- Change the transformation so that it outputs the fixed customers to a text file instead of a database table
- Browse the list of available steps and maybe even try out a few. Some good ones are Sort Rows, Group By, and Calculator. If you know how to code, check out JavaScript, Java Expression, or even the R Script Executor steps.
- Create a job that will run the transformation

Congratulations!

You've completed PDI Evaluation: Getting Started

Click Save & Return in the upper-right of the window.