

# **Lapres Praktikum 1 ETL**

## **Data Warehouse**



Disusun Oleh :

Rifda Qurrotul 'Ain (3321600012)

Dosen Pengampu :

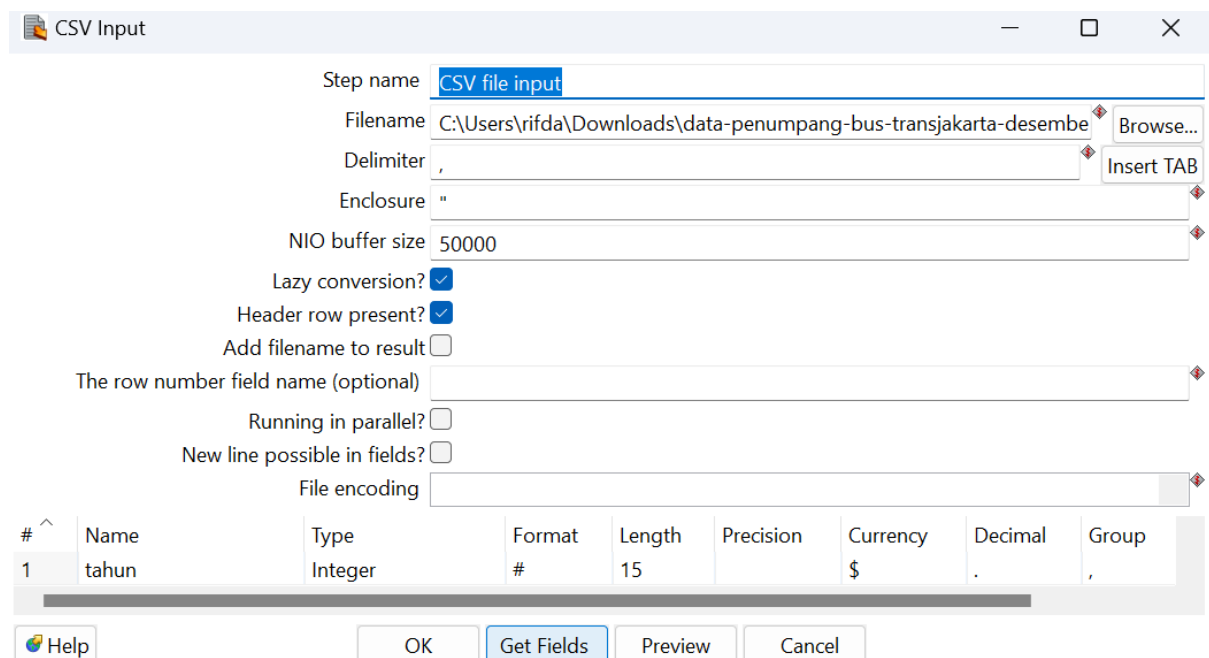
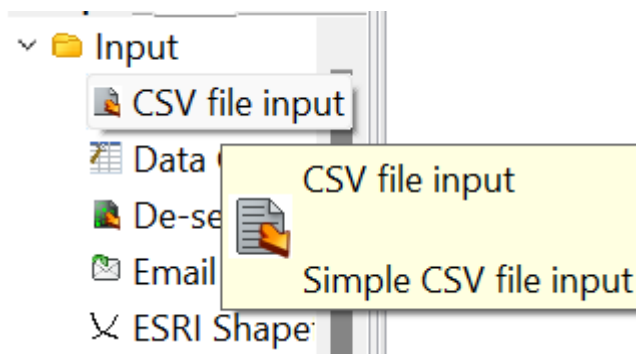
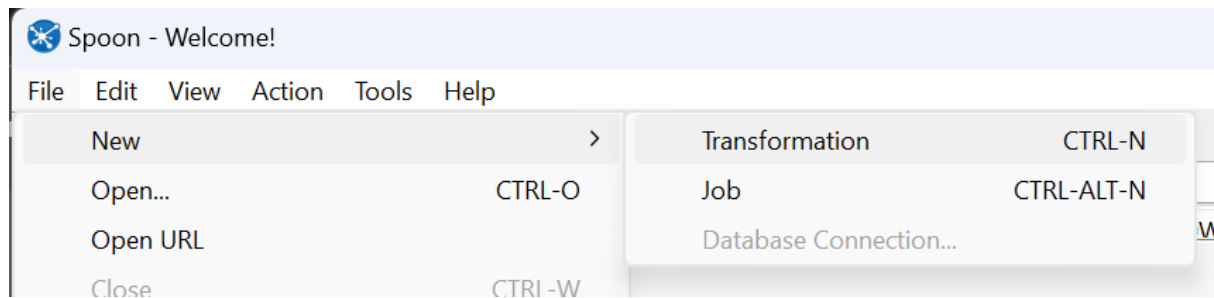
Edi Satriyanto, S.Si., M.Si.

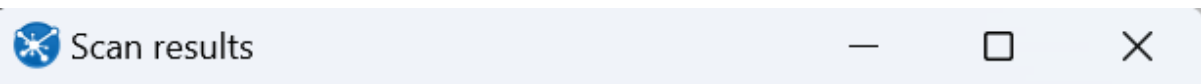
PROGRAM STUDI SARJANA TERAPAN SAINS DATA TERAPAN

POLITEKNIK ELEKTRONIKA NEGERI SURABAYA

2023

## Percobaan 1 : Ekspor CSV ke XML





Here are the results of the document scan:

```
Result after scanning 100 lines.
-----
Field nr. 1 :
  Field name      : tahun
  Field type      : Integer

Field nr. 2 :
  Field name      : bulan
  Field type      : Integer


Field nr. 3 :
  Field name      : jenis
  Field type      : String
  Maximum length  : 23
  Minimum value   : Angkutan Umum Integrasi
  Maximum value   : Mikrotrans
  Nr of null values : 0


Field nr. 4 :
  Field name      : kode_trayek
  Field type      : String
  Maximum length  : 7
  Minimum value   : 1
  Maximum value   : JAK.88
  Nr of null values : 0


Field nr. 5 :
  Field name      : trayek
  Field type      : String
  Maximum length  : 46
  Minimum value   : Andara - Stasiun Universitas Pancasila
  Maximum value   : Tongtek - Cililitan
  Nr of null values : 1

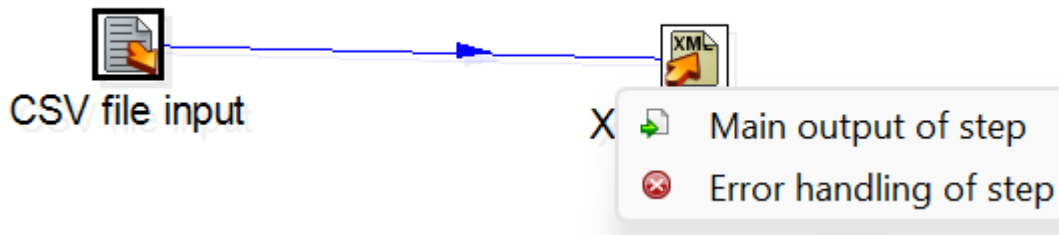
Field nr. 6 :
  Field name      : jumlah_penumpang
  Field type      : Integer
```

Close

 Text file output

 Update

 XML Output



XML output

Step name XML Output

File Content Fields

#	Fieldname	Element name	Content type	Type	Format	Length	Precision	Currency	Decimal	Group	Num
1	tahun		Element	Integer			0				
2	bulan		Element	Integer			0				
3	jenis		Element	String							
4	kode_trayek		Element	String							
5	trayek		Element	String							
6	jumlah_penumpang		Element	Integer			0				
7											
8											

Get Fields Minimal width

Help OK Cancel

Execute a transformation

Local, remote or clustered execution

☒ Local execution ☐ Execute remotely ☐ Execute clustered

Remote host

☐ Pass export to remote server ☒ Post transformation ☒ Prepare execution ☒ Start execution ☐ Show transformations

Details

☐ Enable safe mode ☒ Gather performance metrics ☐ Clear the log before execution

Log level

Replay date (yyyy/MM/dd HH:mm:ss)

Parameters

#	Parameter	Value	Default value
1			

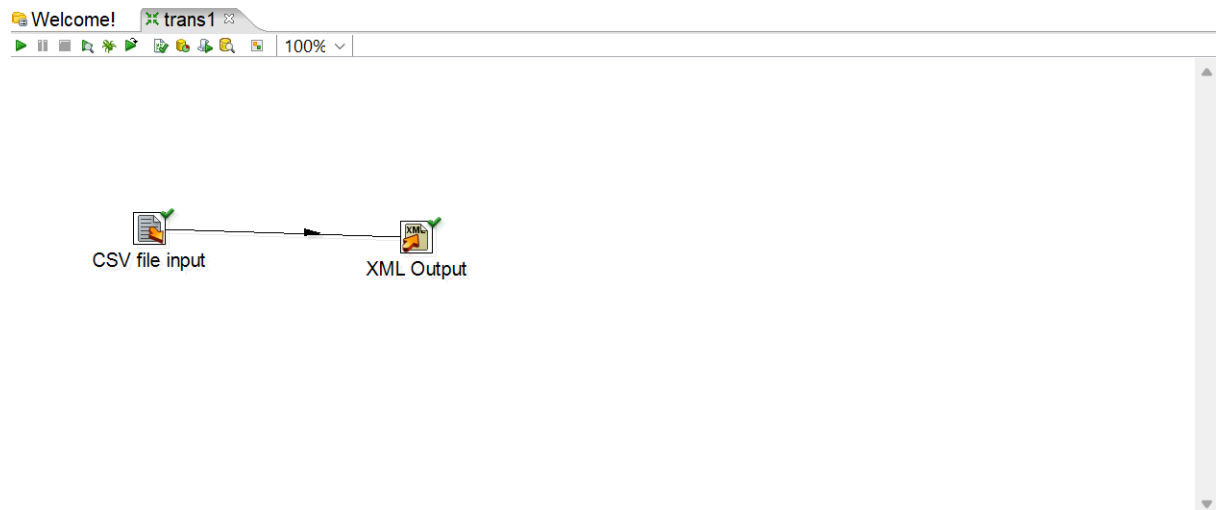
Variables

#	Variable	Value
1	Internal.Job.Filename.Directory	Parent Job File Directory
2	Internal.Job.Filename.Name	Parent Job Filename
3	Internal.Job.Name	Parent Job Name
4	Internal.Job.Repository.Directory	Parent Job Repository Directory

Arguments

#	Argument	Value
1		

Launch Cancel



## Execution Results

Execution History | Logging | Step Metrics | Performance Graph | Metrics | Preview data

#	Stepname	Copynr	Read	Written	Input	Output	Updated	Rejected	Errors	Active	Time	Speed (r/s)
1	CSV file input	0	0	122	123	0	0	0	0	Finished	0.0s	4,920
2	XML Output	0	122	122	0	122	0	0	0	Finished	0.1s	1,649

output.xml.xml X

D: > Applied Data Science > SEMESTER 4 > DATA WAREHOUSE > output.xml.xml

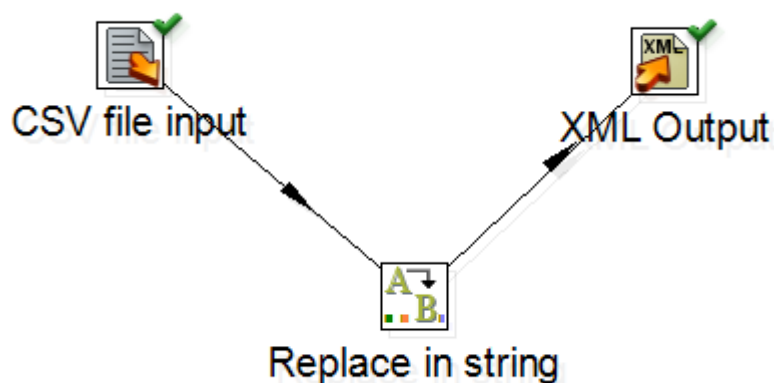
```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <Rows>
3 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.88</kode_trayek> <trayek>T
4 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.85</kode_trayek> <trayek>B
5 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.84</kode_trayek> <trayek>T
6 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.80</kode_trayek> <trayek>R
7 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.77</kode_trayek> <trayek>T
8 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.75</kode_trayek> <trayek>C
9 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.74</kode_trayek> <trayek>T
10 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.73</kode_trayek> <trayek>J
11 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.72</kode_trayek> <trayek>K
12 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.71</kode_trayek> <trayek>K
13 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.64</kode_trayek> <trayek>L
14 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.61</kode_trayek> <trayek>C
15 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.60</kode_trayek> <trayek>K
16 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.59</kode_trayek> <trayek>R
17 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.58</kode_trayek> <trayek>C
18 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.56</kode_trayek> <trayek>G
19 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.54</kode_trayek> <trayek>G
20 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.53</kode_trayek> <trayek>G
21 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.52</kode_trayek> <trayek>K
22 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.51</kode_trayek> <trayek>T
23 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.50</kode_trayek> <trayek>K
24 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.49</kode_trayek> <trayek>L
25 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.47</kode_trayek> <trayek>P
26 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.46</kode_trayek> <trayek>P
27 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.45</kode_trayek> <trayek>L
28 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.44</kode_trayek> <trayek>A
29 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.43</kode_trayek> <trayek>T
30 <Row><tahun>0000000002021</tahun> <bulan>0000000000012</bulan> <jenis>Mikrotrans</jenis> <kode_trayek>JAK.42</kode_trayek> <trayek>Y
```

## Percobaan Kedua: Filter (Ganti Kata)

tahun	bulan	jenis	kode_tray
2021	12	Mikrotran	JAK.88
2021	12	MKTR	JAK.85
2021	12	MKTR	JAK.84
2021	12	Mikrotran	JAK.80
2021	12	Mikrotran	JAK.77
2021	12	MKTR	JAK.75

### Transform

- Add a checksum
- Add constants
- Add sequence
- Add value fields changing sequ
- Add XML
- Calculator
- Closure Generator
- Concat Fields
- Example plugin
- Get ID from slave server
- Number range
- Replace in string



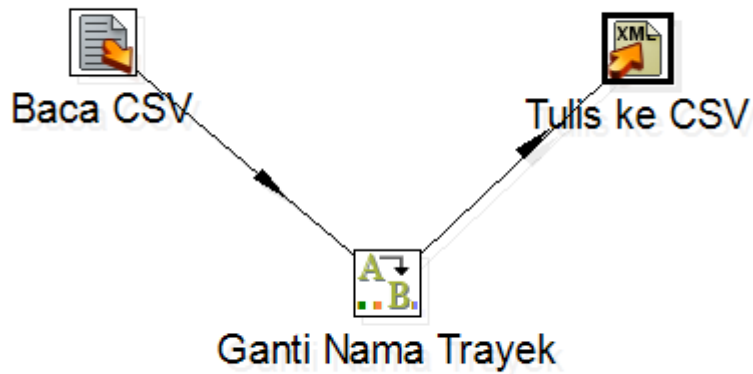
Replace in string

Step name

Ganti Nama Trayek

Fields string

#	^	In stream field	Out stream field	use RegEx	Search	Replace with	Set empty string?	Replace with field	Whole Word	Case sensitive
1		jenis		N	MKTR	Mikrotrans	N		N	N
2										



```

output.xml.xml X
D: > Applied Data Science > SEMESTER 4 > DATA WAREHOUSE > output.xml.xml
1  <?xml version="1.0" encoding="UTF-8"?>
2  <Rows>
3  | <Row><trayek>Terminal Tanjung Priok - Ancol Barat</trayek> <bulan> 00000000000012</bulan> </Row>
4  |

```

## Task Ketiga: Output ke database

### Databases


Create database


Trasjakarta


utf8mb4\_general\_ci


Create


```
1 CREATE TABLE alamat (  
2   tahun varchar(50) not null  
   auto_increment primary key,  
3   bulan varchar(50) not null,  
4   jenis varchar(50) not null,  
5   kode_trayek varchar(50),  
6   trayek varchar(50),  
7   jumlah_penumpang int  
8 )
```


 Synchronize after merge

 Table output

 Text file output

 Update

 XML Output

 Table output  
Write information to a database table

