Using Transformers

Overview

In this exercise, we will cover the basic use of Formats and Transformers using the Microsoft Excel Transformer. This exercise expands on content and concepts covered in the “Lab 1: Creating a Simple Route” tutorial. This exercise will create a Route, which picks up a Microsoft Excel file from a directory, transforms it into XML, transforms it from XML back to Microsoft Excel, and delivers it to another directory.

Steps

1. Create a new Route called “Using Transformers”.

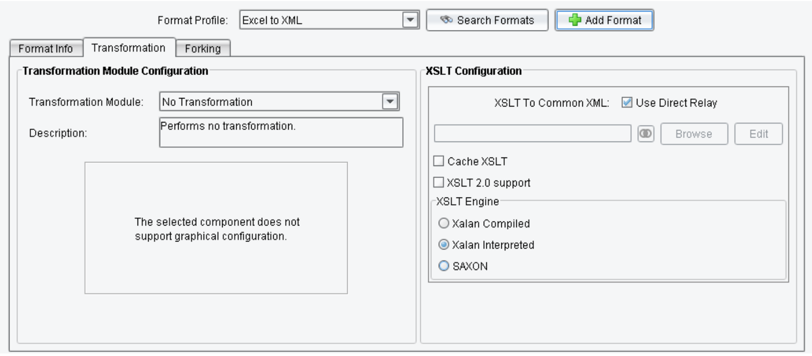
2. Configure a “Directory“ Listener restricting the extension of files we pick up to “xls,” (Microsoft Excel only)

3. Configure the Transport to create files with an “xls” extension, creating Microsoft Excel files as output.

4. Create a new Format by selecting our Source Transform stage and clicking “Add Format” in the “Format Profile” on the bottom half of the screen.

5. Name the format “Excel to XML” and click OK.

   The Format panel should now look like this:



Format Profile: Excel to XML

Transformation Module: No Transformation

Description: Performs no transformation.

The selected component does not support graphical configuration.

XSLT To Common XML: ☑ Use Direct Relay

□ Cache XSLT

□ XSLT 2.0 support

○ Xalan Compiled

● Xalan Interpreted

○ SAXON

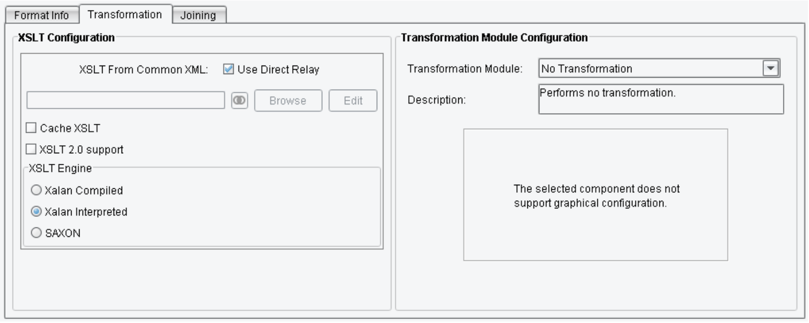
6. Select the Microsoft Excel Transformer from the Transformation Module dropdown.

For our Source Transform, no other configuration is necessary. This Format will take a Microsoft Excel workbook and convert it into an intermediary XML format.

With the Target Transform we will be reversing direction, going from our intermediate XML back to Microsoft Excel. From a functional perspective, this is actually quite useless, as the output will be more-or-less identical to the input. However, once we begin testing the Route, the value will become somewhat more apparent.

7. Select the Target Transform stage and click “Add Format”. You could select the previously created Format from our drop-down, but we'll go ahead and create a new Format called “XML to Excel”:

   With the Format added (and selected), you'll notice that the Target Transform configuration looks slightly different than the Source Transform:



XSLT Configuration

XSLT From Common XML: ☑ Use Direct Relay

☐ Cache XSLT

☐ XSLT 2.0 support

XSLT Engine

○ Xalan Compiled

● Xalan Interpreted

○ SAXON

Transformation Module Configuration

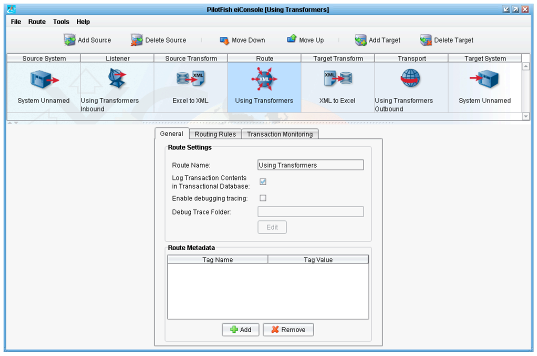
Transformation Module: No Transformation

Description: Performs no transformation.

The selected component does not support graphical configuration.

Primarily, the directionality is reversed. First we potentially go through a XSLT Transformation, then through a Transformation Module. Once again, select the Microsoft Excel Transformer.

We've now configured our Route from end-to-end. It should look something like this:



Route Settings

Route Name: Using Transformers

Log Transaction Contents in Transactional Database: ☑

Enable debugging tracing: ☐

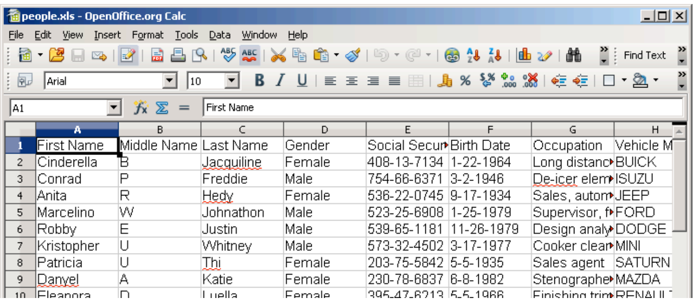
Debug Trace Folder:

Route Metadata

Tag Name	Tag Value

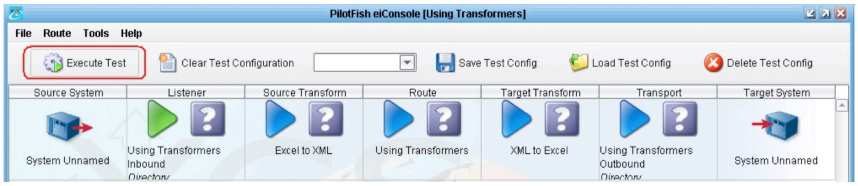
Testing

For our sample file, we've created an Excel sheet:

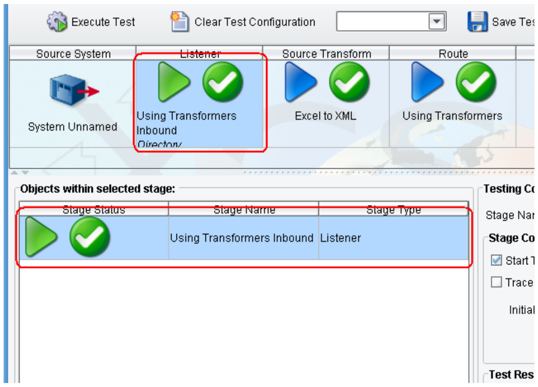

people.xls - OpenOffice.org Calc
File Edit View Insert Format Tools Data Window Help
Find Text
Arial
10
A1
First Name

	A	B	C	D	E	F	G	H
1	First Name	Middle Name	Last Name	Gender	Social Secur	Birth Date	Occupation	Vehicle M
2	Cinderella	B	Jacquiline	Female	408-13-7134	1-22-1964	Long distanc	BUICK
3	Conrad	P	Freddie	Male	754-66-6371	3-2-1946	De-icer elem	ISUZU
4	Anita	R	Hedy	Female	536-22-0745	9-17-1934	Sales, autom	JEEP
5	Marcelino	W	Johnathon	Male	523-25-6908	1-25-1979	Supervisor, f	FORD
6	Robby	E	Justin	Male	539-65-1181	11-26-1979	Design analy	DODGE
7	Kristopher	U	Whitney	Male	573-32-4502	3-17-1977	Cooker clean	MINI
8	Patricia	U	Thi	Female	203-75-5842	5-5-1935	Sales agent	SATURN
9	Danyel	A	Katie	Female	230-78-6837	6-8-1982	Stenographe	MAZDA
10	Eleanora	D	Luella	Female	395-47-6213	5-5-1966	Finishing trim	RENAULT

Copy the sample file ("people.xls") into the input directory ("C:\in"), switch to Testing Mode, and Execute Test:

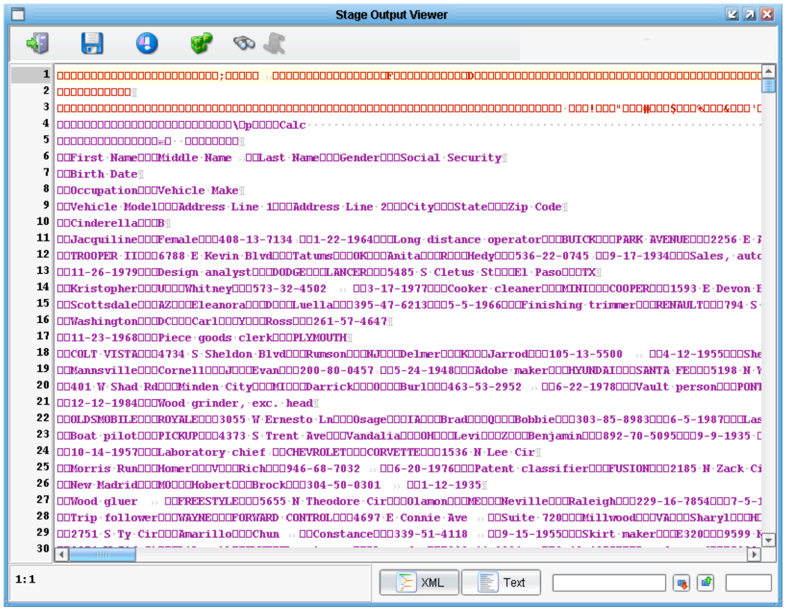


Once the processing completes, we'll view the output at three different stages. First, at the Listener, prior to going into the Source Transform:



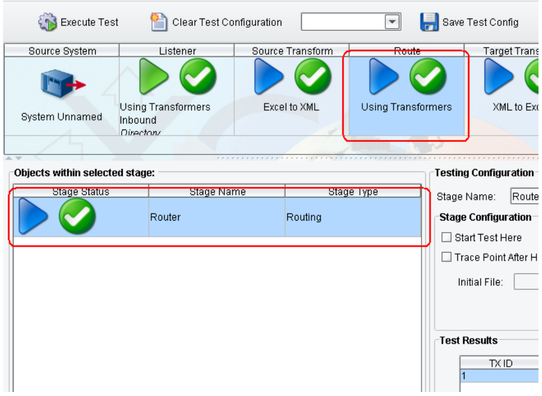
Stage Status	Stage Name	Stage Type
▶ ✓	Using Transformers Inbound	Listener

In this case, our stage output will be the binary representation of the Excel worksheet:



Stage Output Viewer

```
□□□□□□□□□□□□□□□□□□□□□□□□□;□□□□□ ››□□□□□□□□□□□□□□□□□□F□□□□□□□□□□□D□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□□□□□□□□□□¶
□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□·□□□!□□□"□□□#□□□$□□□%□□□&□□□'□
□□□□□□□□□□□□□□□□□□□□□□□□□□□□\□p□□□□Calc·····································································
□□□□□□□□□□□□□□□□□□□□□⁂□··□□□□□□□□¶
□□First·Name□□□Middle·Name ››□□Last·Name□□□Gender□□□Social·Security¶
□□Birth·Date¶
□□Occupation□□□Vehicle·Make¶
□□Vehicle·Model□□□Address·Line·1□□□Address·Line·2□□□City□□□State□□□Zip·Code¶
□□Cinderella□□□B¶
□□Jacquiline□□□Female□□□408-13-7134›□□1-22-1964□□□Long·distance·operator□□□BUICK□□□PARK·AVENUE□□□2256·E·A
□□TROOPER·II□□□6788·E·Kevin·Blvd□□□Tatums□□□□OK□□□Anita□□□R□□□Hedy□□□536-22-0745›□□9-17-1934□□□Sales,·auto
□□11-26-1979□□□Design·analyst□□□DODGE□□□LANCER□□□5485·S·Cletus·St□□□El·Paso□□□TX¶
□□Kristopher□□□U□□□Whitney□□□573-32-4502 ›› □□3-17-1977□□□Cooker·cleaner□□□MINI□□□COOPER□□□1593·E·Devon·E
□□Scottsdale□□□AZ□□□Eleanora□□□D□□□Luella□□□395-47-6213□□□5-5-1966□□□Finishing·trimmer□□□RENAULT□□□794·S·
□□Washington□□□DC□□□Carl□□□Y□□□Ross□□□261-57-4647¶
□□11-23-1968□□□Piece·goods·clerk□□□PLYMOUTH¶
□□COLT·VISTA□□□4734·S·Sheldon·Blvd□□□Rumson□□□NJ□□□Delmer□□□K□□□Jarrod□□□105-13-5500 ›› □□4-12-1955□□□She
□□Mannsville□□□Cornell□□□J□□□Evan□□□200-80-0457›□□5-24-1948□□□Adobe·maker□□□HYUNDAI□□□SANTA·FE□□□5198·N·Y
□□401·W·Shad·Rd□□□Minden·City□□□MI□□□Darrick□□□O□□□Burl□□□463-53-2952 ›› □□6-22-1978□□□Vault·person□□□PONT
□□12-12-1984□□□Wood·grinder,·exc.·head¶
□□OLDSMOBILE□□□ROYALE□□□3055·W·Ernesto·Ln□□□Osage□□□IA□□□Brad□□□Q□□□Bobbie□□□303-85-8983□□□6-5-1987□□□Las
□□Boat·pilot□□□PICKUP□□□4373·S·Trent·Ave□□□Vandalia□□□OH□□□Levi□□□Z□□□Benjamin□□□892-70-5095□□□9-9-1935·□
□□10-14-1957□□□Laboratory·chief›□□CHEVROLET□□□CORVETTE□□□1536·N·Lee·Cir¶
□□Morris·Run□□□Homer□□□V□□□Rich□□□946-68-7032 ›› □□6-20-1976□□□Patent·classifier□□□FUSION□□□2185·N·Zack·Ci
□□New·Madrid□□□MO□□□Hobert□□□Brock□□□304-50-0301 ›› □□1-12-1935¶
□□Wood·gluer ›› □□FREESTYLE□□□5655·N·Theodore·Cir□□□Olamon□□□ME□□□Neville□□□Raleigh□□□229-16-7854□□□7-5-1
□□Trip·follower□□□WAYNE□□□FORWARD·CONTROL□□□4697·E·Connie·Ave ›› □□Suite·720□□□Millwood□□□VA□□□Sharyl□□□HO
□□2751·S·Ty·Cir□□□Amarillo□□□Chun ›› □□Constance□□□339-51-4118 ›› □□9-15-1955□□□Skirt·maker□□□E320□□□9599·K
```



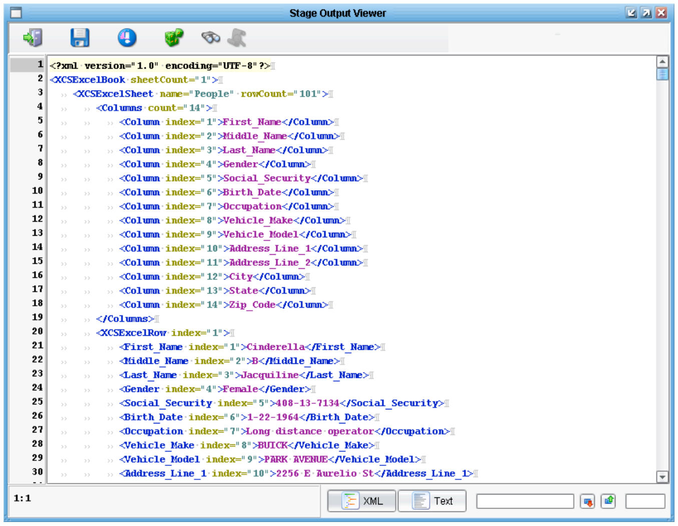
Objects within selected stage:

Stage Status	Stage Name	Stage Type
✓	Router	Routing

Stage Name: Route

□ Start Test Here
□ Trace Point After H
Initial File:

Next, we'll view it at the Routing stage, which is after it has undergone transformation to XML:

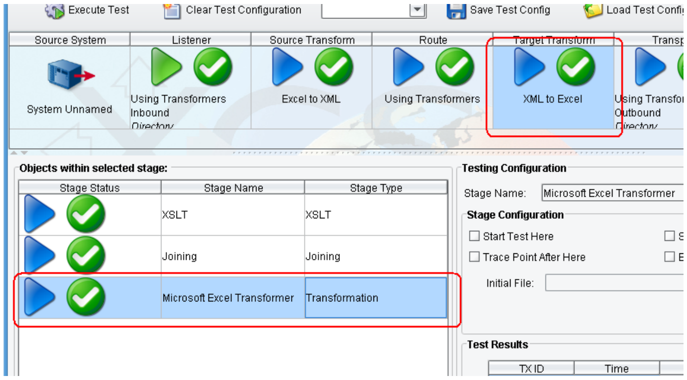


```
<?xml version="1.0" encoding="UTF-8"?>
<XCSExcelBook sheetCount="1">
  <XCSExcelSheet name="People" rowCount="101">
    <Columns count="14">
      <Column index="1">First_Name</Column>
      <Column index="2">Middle_Name</Column>
      <Column index="3">Last_Name</Column>
      <Column index="4">Gender</Column>
      <Column index="5">Social_Security</Column>
      <Column index="6">Birth_Date</Column>
      <Column index="7">Occupation</Column>
      <Column index="8">Vehicle_Make</Column>
      <Column index="9">Vehicle_Model</Column>
      <Column index="10">Address_Line_1</Column>
      <Column index="11">Address_Line_2</Column>
      <Column index="12">City</Column>
      <Column index="13">State</Column>
      <Column index="14">Zip_Code</Column>
    </Columns>
    <XCSExcelRow index="1">
      <First_Name index="1">Cinderella</First_Name>
      <Middle_Name index="2">B</Middle_Name>
      <Last_Name index="3">Jacquiline</Last_Name>
      <Gender index="4">Female</Gender>
      <Social_Security index="5">408-13-7134</Social_Security>
      <Birth_Date index="6">1-22-1964</Birth_Date>
      <Occupation index="7">Long distance operator</Occupation>
      <Vehicle_Make index="8">BUICK</Vehicle_Make>
      <Vehicle_Model index="9">PARK AVENUE</Vehicle_Model>
      <Address_Line_1 index="10">2256 E Aurelio St</Address_Line_1>
```

Our output here is XML:

Spend some time reviewing the output and comparing it to the original sample. The XML created has a root tag called XCSExcelBook, and each sheet creates an XCSExcelSheet, which in turn contains a Columns tag (detailing each column defined) and then an XCSExcelRow tag for each row. Columns are then defined and named within each XCSExcelRow tag (for example, “Middle_Name”).

Finally, view the stage output at the Target Transform's third stage:



Stage Status	Stage Name	Stage Type
▶ ✓	XSLT	XSLT
▶ ✓	Joining	Joining
▶ ✓	Microsoft Excel Transformer	Transformation

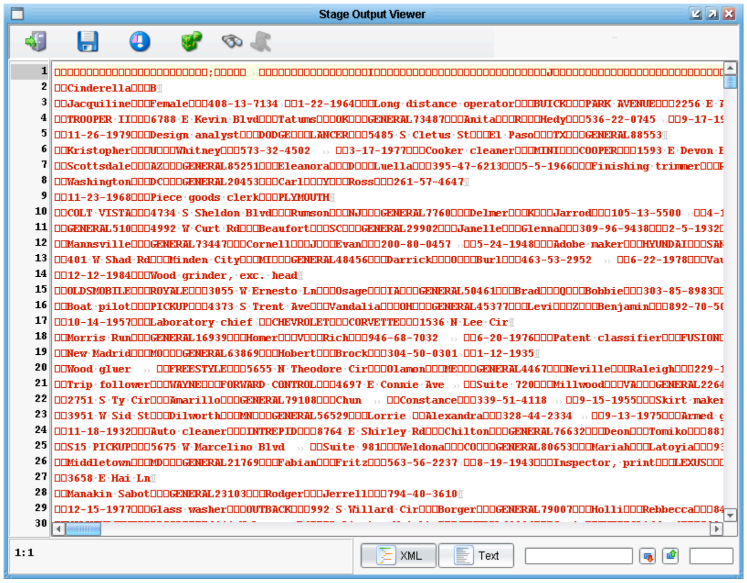
Stage Name: Microsoft Excel Transformer

□ Start Test Here

□ Trace Point After Here

Initial File:

We should see binary data again:



```
□□□□□□□□□□□□□□□□□□□□□□□□□□;□□□□□	□□□□□□□□□□□□□□□□□□I□□□□□□□□□□□□□□□□□□□□□□□□□□□J□□□□□□□□□□□□□□□□□□□□□□□□□□□
□□Cinderella□□□B
□□Jacquiline□□□Female□□□408-13-7134	□□1-22-1964□□□Long distance operator□□□BUICK□□□PARK AVENUE□□□2256 E 
□□TROOPER II□□□6788 E Kevin Blvd□□□Tatums□□□OK□□□GENERAL73487□□□Anita□□□R□□□Hedy□□□536-22-0745	□□9-17-1
□□11-26-1979□□□Design analyst□□□DODGE□□□LANCER□□□5485 S Cletus St□□□El Paso□□□TX□□□GENERAL88553
□□Kristopher□□□U□□□Whitney□□□573-32-4502	□□3-17-1977□□□Cooker cleaner□□□MINI□□□COOPER□□□1593 E Devon 
□□Scottsdale□□□AZ□□□GENERAL85251□□□Eleanora□□□D□□□Luella□□□395-47-6213□□□5-5-1966□□□Finishing trimmer□□□
□□Washington□□□DC□□□GENERAL20453□□□Carl□□□Y□□□Ross□□□261-57-4647
□□11-23-1968□□□Piece goods clerk□□□PLYMOUTH
□□COLT VISTA□□□4734 S Sheldon Blvd□□□Rumson□□□NJ□□□GENERAL7760□□□Delmer□□□K□□□Jarrod□□□105-13-5500	□□4-1
□□GENERAL510□□□4992 W Curt Rd□□□Beaufort□□□SC□□□GENERAL29902□□□Janelle□□□Glenna□□□309-96-9438□□□2-5-1932□
□□Mannsville□□□GENERAL73447□□□Cornell□□□J□□□Evan□□□200-80-0457	□□5-24-1948□□□Adobe maker□□□HYUNDAI□□□SA
□□401 W Shad Rd□□□Minden City□□□MI□□□GENERAL48456□□□Darrick□□□O□□□Burl□□□463-53-2952	□□6-22-1978□□□Va
□□12-12-1984□□□Wood grinder, exc. head
□□OLDSMOBILE□□□ROYALE□□□3055 W Ernesto Ln□□□Osage□□□IA□□□GENERAL50461□□□Brad□□□Q□□□Bobbie□□□303-85-8983□□
□□Boat pilot□□□PICKUP□□□4373 S Trent Ave□□□Vandalia□□□OH□□□GENERAL45377□□□Levi□□□Z□□□Benjamin□□□892-70-5
□□10-14-1957□□□Laboratory chief	□□CHEVROLET□□□CORVETTE□□□1536 N Lee Cir
□□Morris Run□□□GENERAL16939□□□Homer□□□V□□□Rich□□□946-68-7032	□□6-20-1976□□□Patent classifier□□□FUSION□
□□New Madrid□□□MO□□□GENERAL63869□□□Hobert□□□Brock□□□304-50-0301	□□1-12-1935
□□Wood gluer	□□FREESTYLE□□□5655 N Theodore Cir□□□Olamon□□□ME□□□GENERAL4467□□□Neville□□□Raleigh□□□229-1
□□Trip follower□□□WAYNE□□□FORWARD CONTROL□□□4697 E Connie Ave	□□Suite 720□□□Millwood□□□VA□□□GENERAL2264
□□2751 S Ty Cir□□□Amarillo□□□GENERAL79108□□□Chun	□□Constance□□□339-51-4118	□□9-15-1955□□□Skirt maker
□□3951 W Sid St□□□Dilworth□□□MN□□□GENERAL56529□□□Lorrie	□□Alexandra□□□328-44-2334	□□9-13-1975□□□Armed 
□□11-18-1932□□□Auto cleaner□□□INTREPID□□□8764 E Shirley Rd□□□Chilton□□□GENERAL76632□□□Deon□□□Tomiko□□□881
□□S15 PICKUP□□□5675 W Marcelino Blvd	□□Suite 981□□□Weldona□□□CO□□□GENERAL80653□□□Mariah□□□Latoyia□□□9
□□Middletown□□□MD□□□GENERAL21769□□□Fabian□□□Fritz□□□563-56-2237	□□8-19-1943□□□Inspector, print□□□LEXUS□□□
□□3658 E Hai Ln
□□Manakin Sabot□□□GENERAL23103□□□Rodger□□□Jerrell□□□794-40-3610
□□12-15-1977□□□Glass washer□□□OUTBACK□□□992 S Willard Cir□□□Borger□□□GENERAL79007□□□Holli□□□Rebbecca□□□84
```

Finally, check your output directory and view the resulting Microsoft Excel file.