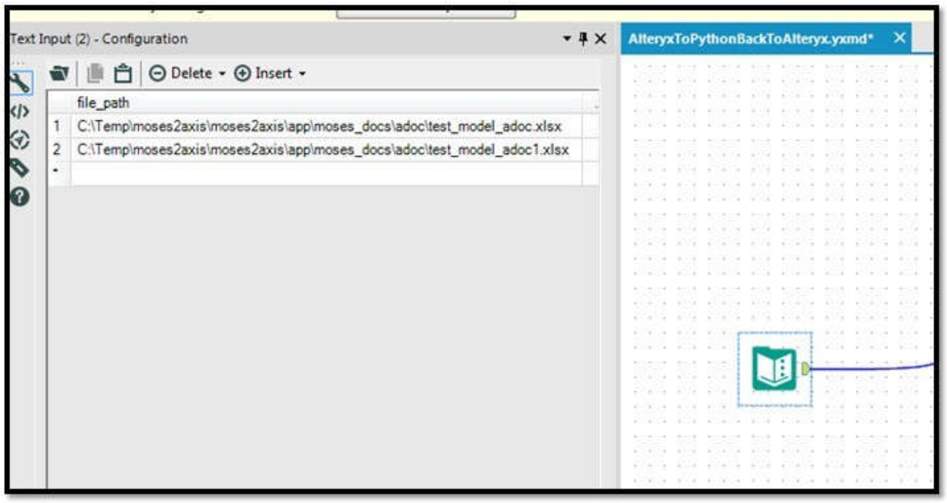
alteryx2python

Send data from Alteryx to Python, execute script and send Python output back to Alteryx workflow.

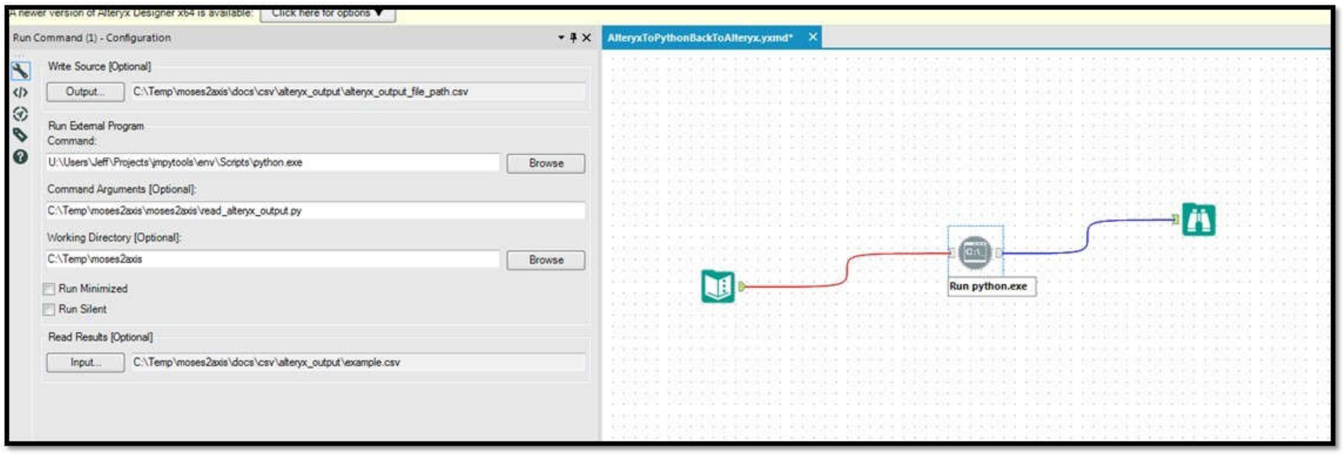
Step 1: Source of data from Alteryx.

For this example, I’ve just used a simple text input with some file paths, but this can be anything generated from Alteryx.


Text Input (2) - Configuration
AlteryxToPythonBackToAlteryx.yxmd*
Delete
Insert

	file_path
1	C:\Temp\moses2axis\moses2axis\app\moses_docs\adoc\test_model_adoc.xlsx
2	C:\Temp\moses2axis\moses2axis\app\moses_docs\adoc\test_model_adoc1.xlsx

Step 2: Connect to Run Command Tool.


A newer version of Alteryx Designer x64 is available:
Click here for options
Run Command (1) - Configuration
AlteryxToPythonBackToAlteryx.yxmd*
Run python.exe

Write Source [Optional]
Output... C:\Temp\moses2axis\docs\csv\alteryx_output\alteryx_output_file_path.csv

Run External Program
Command:
U:\Users\Jeff\Projects\jmpytools\env\Scripts\python.exe

Command Arguments [Optional]:
C:\Temp\moses2axis\moses2axis\read_alteryx_output.py

Working Directory [Optional]:
C:\Temp\moses2axis

☐ Run Minimized
☐ Run Silent

Read Results [Optional]
Input... C:\Temp\moses2axis\docs\csv\alteryx_output\example.csv

As for the options:

1. Write source will take whatever data is coming IN to the tool, that is the data stream that is attached to the input of the Run Command, and write it to the file or connection specified.
   a. The point of this is to dump the data that is currently in Alteryx to a file so that your script can then read the file.
2. Under 'Run External Program Command:' for a Python script, your command should be Python.exe. If the directory where Python exists is in your system path variable, you can simply type Python.exe, otherwise you will have to give it the full path keeping in mind to quote the string if there are spaces (e.g. "Program Files"). I however believe it is best practice to always include the entire file paths.

3. In command argument, you will type the location of your Python script your_python_script.py (Alteryx's default working directory is the directory of the running module so if is easiest to keep your script nearby and simply type "your_python_script.py" instead of the full path) and any - or -- options necessary. Remember to quote this string.
4. The default working directory for Alteryx is the directory of the running module unless set otherwise.
   a. It is possible to find the working directory of your python script by running the following in python:

```
import os
os.getcwd()
# 'C:\\temp\\moses2axis'
```

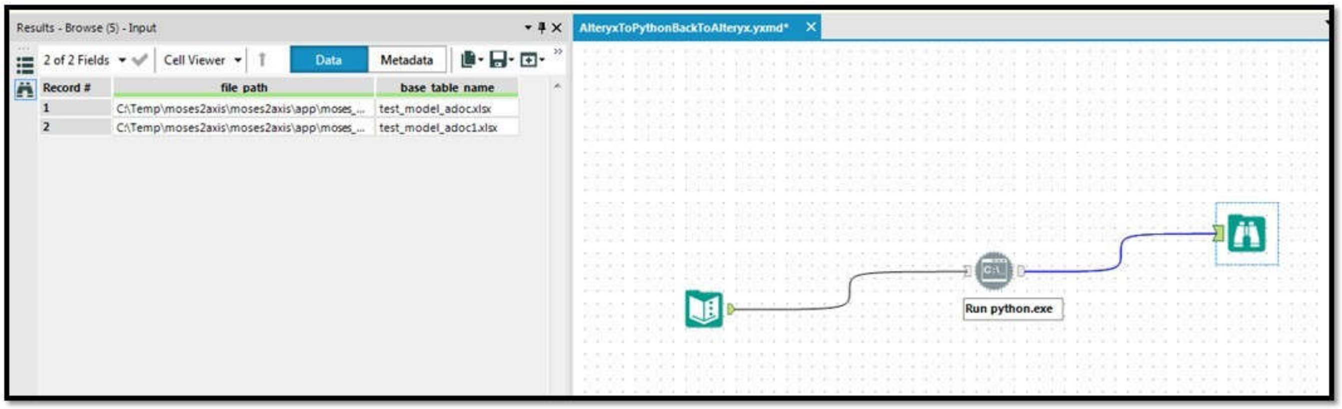
5. Your Python script should then write to a file at the end of its processing.
6. The Read Results option should point to the file written by your Python script. This way, Alteryx can read back in the data that was processed by the external command.

Step 3: Write Python Script That Reads Alteryx Output

```
#!/usr/bin/env python
# -*- coding: utf-8 -*-
"""
author: jeff maxey
date: 2018.05.15
purpose: example demonstrating integration of python and alteryx
"""

# // import packages
import os
import subprocess

def install(name):
    subprocess.call(['pip', '--proxy=cr-proxy.us.aegon.com:9090', 'install', name])
install('pandas')

import pandas as pd

# // define a function that will list all file paths in specified root directory with specified file extension
def list_file_paths(root_path, ext):
    """
    Purpose: Search all folders in root_path for specific file type.
    :param root_path: The parent directory for traversal search
    :param ext: Specified file extension desired to be searched for.
    :return: A list of files located within root_path that possess specified file extension
    """
    # // create an empty list for storing all file paths
    all_files = []

    # // walk entire directory of root_path; all sub-folders and files will be checked.
    for root, dirs, files in os.walk(root_path):
        for filename in files:
            # // check if file name ends with specified extension
            if filename.lower().endswith(ext):
                # // if it does, generate the full path using os.path.join() and append it to all_files list
                all_files.append(os.path.join(root, filename))

    # // return the list of files
    return all_files


# // specify the directory containing output from alteryx workflow
dir_alteryx_output = r"C:\Temp\moses2axis\docs\csv\alteryx_output"

# // call the function you created above to collect a list of files in the folder; returns a list
# // returns: ['C:\\Temp\\moses2axis\\docs\\csv\\alteryx_output\\alteryx_output_file_path.csv']
file_list = list_file_paths(root_path=dir_alteryx_output, ext='.csv')

# // select the file from the list using indexing; returns a string
# // returns: 'C:\\Temp\\moses2axis\\docs\\csv\\alteryx_output\\alteryx_output_file_path.csv'
desired_file = file_list[0]

# // easiest way to read a file is using pandas; you can even do pd.read_excel, pd.read_sql and many more:
df = pd.read_csv(desired_file, sep=',', header=0)

# // just to show an example of python performing a procedure, lets create an additional column and add the basename
# // of the file path.
# // e.g. 'C:\\Temp\\moses2axis\\docs\\csv\\alteryx_output\\alteryx_output_file_path.csv' --> 'alteryx_output_file_path.csv'
df['base_table_name'] = df['file_path'].apply(lambda x: os.path.basename(x))


# // generate a csv file path and output data frame; e.g. 'C:\\Temp\\moses2axis\\moses2axis\\app\\project_directory\\docs\\csv\\xref.cs
filepathCSV = os.path.join(dir_alteryx_output, "example.csv")
df.to_csv(filepathCSV, sep=",", header=True, index=False)
```

Step 4: Stream Python Output Back Into Alteryx Workflow

The Read Result Option path specified in the Run Command Tool in Alteryx matches the output file path generated from the python script, and the result will be read back into Alteryx after the python script has been completed (e.g. base_table_name column has been added).


Results - Browse (5) - Input
2 of 2 Fields
Cell Viewer
Data
Metadata
AlteryxToPythonBackToAlteryx.yxmd*
Run python.exe

Record #	file_path	base_table_name
1	C:\Temp\moses2axis\moses2axis\app\moses_...	test_model_adoc.xlsx
2	C:\Temp\moses2axis\moses2axis\app\moses_...	test_model_adoc1.xlsx