**Manual File Loads - Running Python Programs (With or Without Optional SSIS Pkg)**

Contents

[Summary 2](#_Toc83645302)

[Jupyter Notebook Processes 2](#_Toc83645303)

[Software Installs 2](#_Toc83645304)

[GitHub 2](#_Toc83645305)

[Local Folders 2](#_Toc83645306)

[Customer Data Table 2](#_Toc83645307)

[Execution 3](#_Toc83645308)

[Using SSIS Package 3](#_Toc83645309)

[Not Using SSIS Package (i.e. Just Running Python Prog) 4](#_Toc83645310)

[Special Customer Processing 4](#_Toc83645311)

[Amazon Fresh PN & BM 4](#_Toc83645312)

[Albertsons 4](#_Toc83645313)

[Meijer 4](#_Toc83645314)

[Walmart OPD 5](#_Toc83645315)

## Summary

* This document outlines how to prepare for and execute manual file loads using Python program suite.
* An optional **SSIS** pkg may also be used as a wrapper. (**ExecManualLoad.dtsx**.)

## Jupyter Notebook Processes

1. Files processed via **Jupyter Notebook** are handled separately. (See relevant documentation.)
2. Customer data processed via Jupyter Notebook are as follows:

|  |  |
| --- | --- |
| **Customer** | **Comment** |
| Walmart OPD | * Jupyter Notebook used for **initial processing** of file. * SSIS pkg may optionally be used for **subsequent** SS Agent job execution & recon report. |
| Amazon BM | * Jupyter Notebook used for **initial processing** of file. * **Subsequent processing** of file is by regular Amazon FreshPN process as outlined in this document. |

## Software Installs

1. **Python**: Install Python. (Programs have been run using Ver 3.7 which may be superseded by now.)
2. **IDE**: Install Visual Studio Code or other IDE (e.g. Pycharm).
3. **Python libraries**: You’ll also probably also need to install several of these. Pls do so as necessary dependent on specific Python program/customer requirements. (Using **pip.exe**.)

## GitHub

1. **GitHub**: Clone/download Manual Load programs/files from GitHub. Location**: https://github.com/pepsico-ecommerce/py-ETL-V2.**
2. **KT Documentation:** Located on GitHub in [**https://github.com/pepsico-ecommerce/py-ETL-V2/KT**](https://github.com/pepsico-ecommerce/py-ETL-V2/KT)**.**
3. **SSIS package**: Located on GitHub in [**https://github.com/pepsico-ecommerce/py-ETL-V2/SSIS**](https://github.com/pepsico-ecommerce/py-ETL-V2/SSIS)**.**

## Local Folders

1. Pls create following folder on C: drive.

**C:\CompaniesSourceFilesReceived**

1. Pls create folders within above folder using table in next section.

## Customer Data Table

|  |  |  |
| --- | --- | --- |
| **Customer (Option for Python Prog)** | **Folder** | **Usable w/ SSIS pkg? (Yes unless otherwise indicated)** |
| kroger\_ship | KrogerShip |  |
| sams | SamsDotCom |  |
| sams\_pickup | SamsPickup |  |
| albertsons | Albertsons | No (has multiple input files) |
| amazon\_freshpn | AmazonFreshPN | No (has multiple input files) |
| meijer | Meijer |  |
| hyvee | HyVee |  |
| walmart\_og | NA | No (for initial process of file via Jupyter Notebook) |

## Execution

1. Ensure you’re connected to **VPN**.

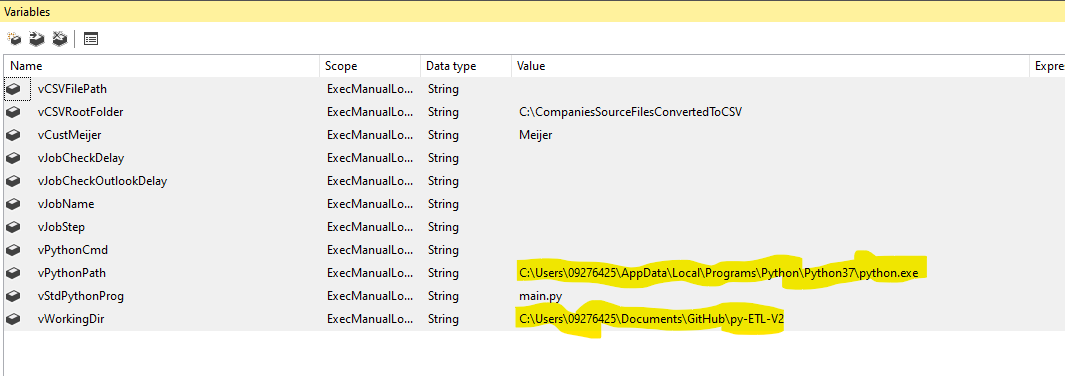
### Using SSIS Package

Note: SSIS pkg acts as a processing wrapper and will execute:

* Local Python program. (**main.py** – see below.)
* SQL Server Agent job.
* Recon report at end of processing. (Note: Recon currently mailed to [David.McIntyre.Contractor@pepsico.com](mailto:David.McIntyre.Contractor@pepsico.com). To change, edit line 122 in Python prog **get\_job\_status.py** in py-ETL-V2 folder.)

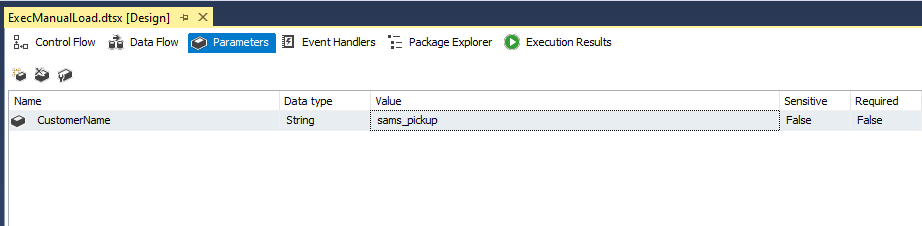
#### Set Up

1. Set package variables as approp. (Highlighted below.)



#### Execution

1. Set package parameter to relevant customer name.



1. Execute package.
2. Note that program will usually prompt for confirmation of file to be loaded in dialog window.

### Not Using SSIS Package (i.e. Just Running Python Prog)

1. Cd to local Github location of py\_etl\_v2 (e.g. Documents\GitHub\py-ETL-V2).
2. Start IDE (e.g. Visual Studio Code, Pycharm).
3. Open **main.py** in IDE.
4. Copy Excel file to be processed to relevant customer folder within **C:\CompaniesSourceFilesReceived**.
5. Execute **main.py** for relevant **customer** (per *Customer Data* *Table* above). Syntax:

main.py --customer <customer>

e.g. main.py –customer hyvee

## Special Customer Processing

### Amazon Fresh PN & BM

These files are separate but should be processed via Python **before** **executing SQL Server Agent job**. Pls see instructions below.

|  |  |
| --- | --- |
| **File** | **Instructions** |
| BM | 1. Do **initial** file processing using **Jupyter Notebook** (see related documentation). 2. Do **subsequent** file processing using Python program **main.py**. (Process as **Amazon** **Fresh PN** file.) |
| Fresh PN | Process file using Python program **main.py**. |

### Albertsons

These files are separate but should be processed via Python **before** **executing SQL Server Agent job**. Pls see instructions below.

|  |  |
| --- | --- |
| **File** | **Instructions** |
| HD | Process file using Python program **main.py**. |
| DUG | Process file using Python program **main.py**. |

### Meijer

If **not** **using SSIS** **pkg**, follow steps below.

1. Perform following **before** executing SQL Server Agent job:
   1. Load CSV created by Python program into following table:
      1. [TDWH].[DataStage].[STG\_MeijerSales]

### Walmart OPD

If **not** **using SSIS** **pkg**, follow steps below.

1. Perform following **before** executing SQL Server Agent job (and after processing file using **Jupyter Notebook)**:
   1. Connect to SQL Server and execute following commands:

Use TDWH

Exec [DataStage].[ConvertWalmartOGSalesDataFormat];

Exec [DataStage].[ProcessWalmartOGSalesData]