Docker Workflow

This file talks about using docker image for the automated process of loading and validating Lending Club data.

Steps to Run the Docker Image:

Docker Image Name: sweta/luigiscript

DockerHub Path: <https://hub.docker.com/r/sweta/luigiscript/>

Note: Docker was run on AWS Cloud Instance

Docker Image Flow:

Build Docker Image on Local with the required files copied to the Image

Run your python/Luigi files in the container and check the output files created on S3

Create a docker container to run your python/luigi scripts

Pull the Latest Docker Image and check the docker Images

Login to Docker and Push the Image from Local machine to DockerHub

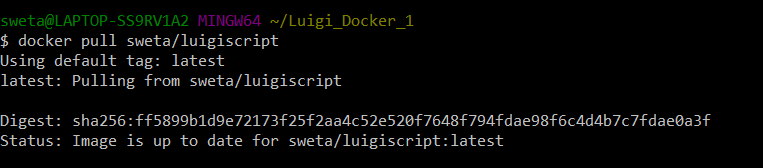
Docker Image on DockerHub:



Docker Commands to run the Docker Image

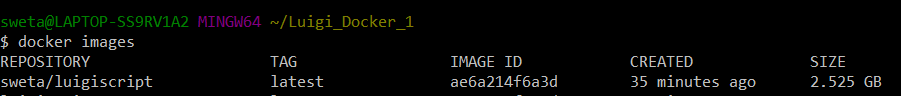
1. Pull the Docker Image

docker pull sweta/luigiscript



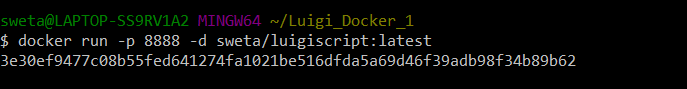
1. Check if the Image is pulled correctly

docker images

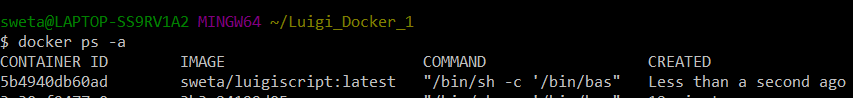


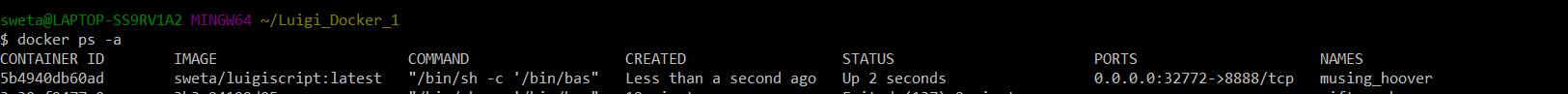
1. Now create the container to Run the Python/Luigi Files present in the images

docker run -p 8888 -d sweta/luigiscript:latest

  
4) Check the create container, with the name sweta/luigiscript

docker ps -a

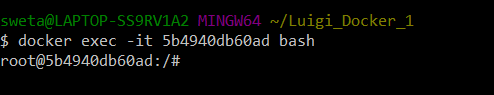




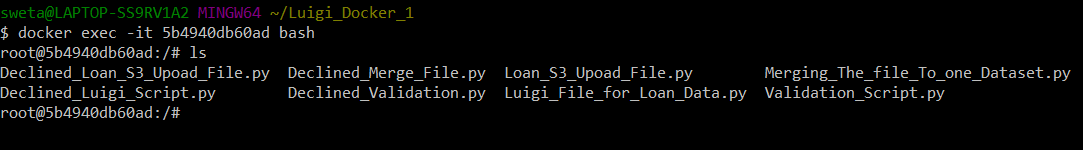
1. Once a container is created, Let’s get into the container and run the Luigi

docker exec -it 5b4940db60ad bash

# docker exec -it <present container id> bash



1. Let’s check the files in the container by using the ‘ls’ command in the container



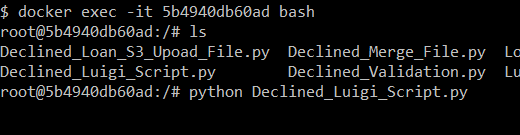
1. Now, we can run the python files for the Luigi Pipelines that we have created for Loan Data and the Declined Loan Data.

File names for each pipeline:

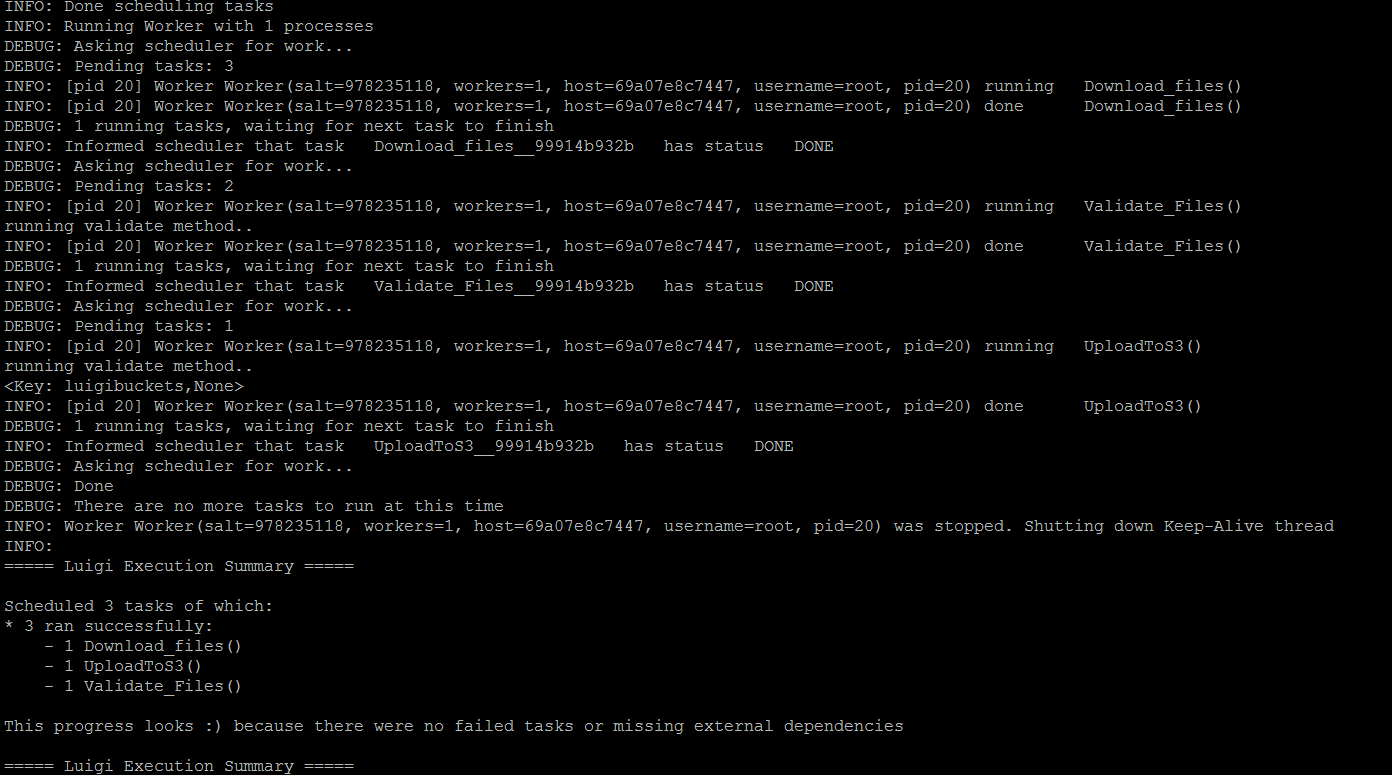
1. Loan Data - Luigi\_File\_for\_Loan\_Data.py
2. Declined Loan Data – Declined\_Luigi\_Script.py

Run the following command:

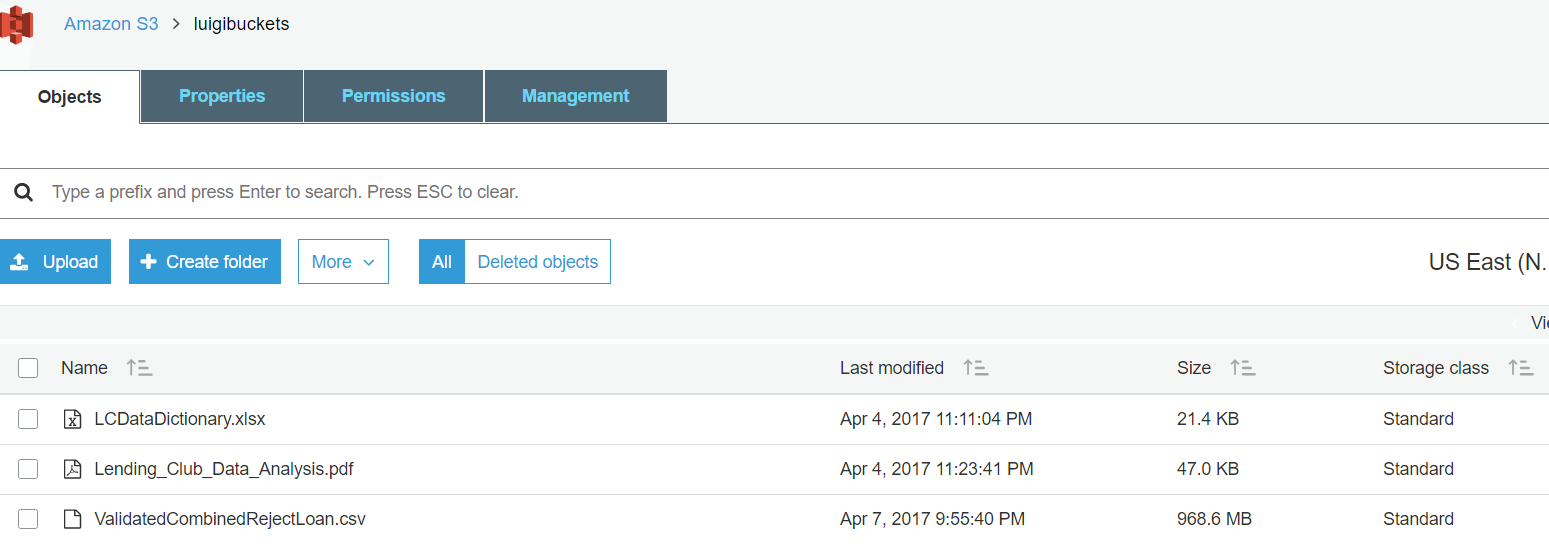
python Declined\_Luigi\_Script.py



Check the success status



Enter the AWS when the program asks for and check the output in the Amazon S3 Bucket



1. Repeat the Same commands for the Loan Data using the command and check the output in S3

python Luigi\_File\_for\_Loan\_Data.py

