Pipeline and instructions on how to run it

To create a database you must first launch the SQL Shell psql and enter the proper credentials. To make the actual database you must type in

CREATE DATABASE databasename

To connect to the new database you must type in

\c databasename

To add tables to this database you must specify the number of variables, as well as the types of variables and also the primary key of the table. For example:

CREATE TABLE tablename(

id INTEGER,

name VARCHAR(100),

age INTEGER,

blood_type VARCHAR(10)

PRIMARY KEY (id)

);

In order to filter the necessary information from the csv files you must specify the filtering parameter of which rows we need and create a new file using the following command in a PowerShell pipeline

Import-Csv .\input.csv -Delimiter ',' |

? { $_.[Variable name] -like '*[Variable Value]*' } |

Export-Csv -NoTypeInformation output.csv

In order to populate this database with a pre-existing csv file you must specify which table you want to populate, specify the source and then add the delimiter. For example:

COPY tablename(id, name, age, blood_type)

FROM 'C:\sampledata\tablenamedata.csv'

DELIMITER ','

CSV HEADER;