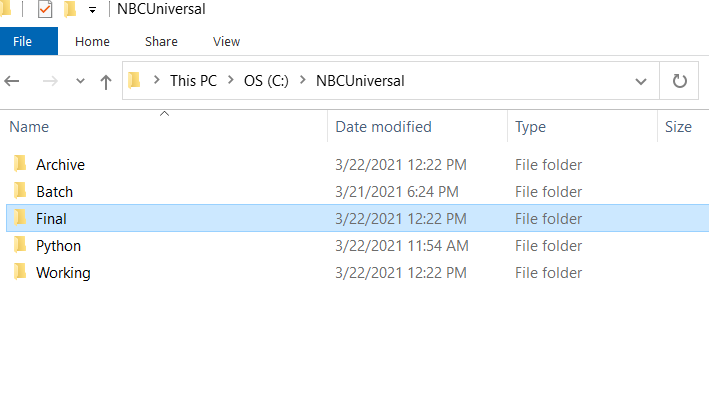
**Data Engineer Runtime Instructions**

You’re environment will require Python3.7 or more as a local install. The main python script is NBCUniversal\_Formula1.py and can be run in conjunction with the attached NBCUniversal\_challenge.bat file. Your folder structure will need to be as follows locally:



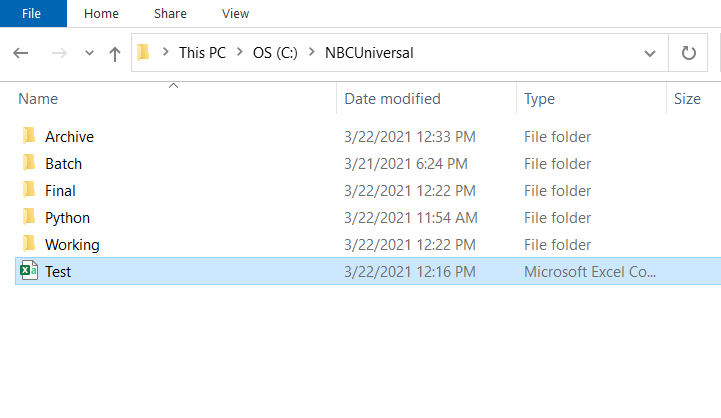
Attached files will need to be placed as follows:

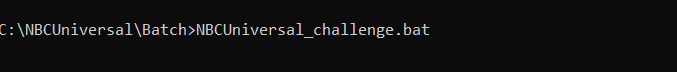
NBCUniversal\_challenge.bat -> c:\NBCUniversal\Batch

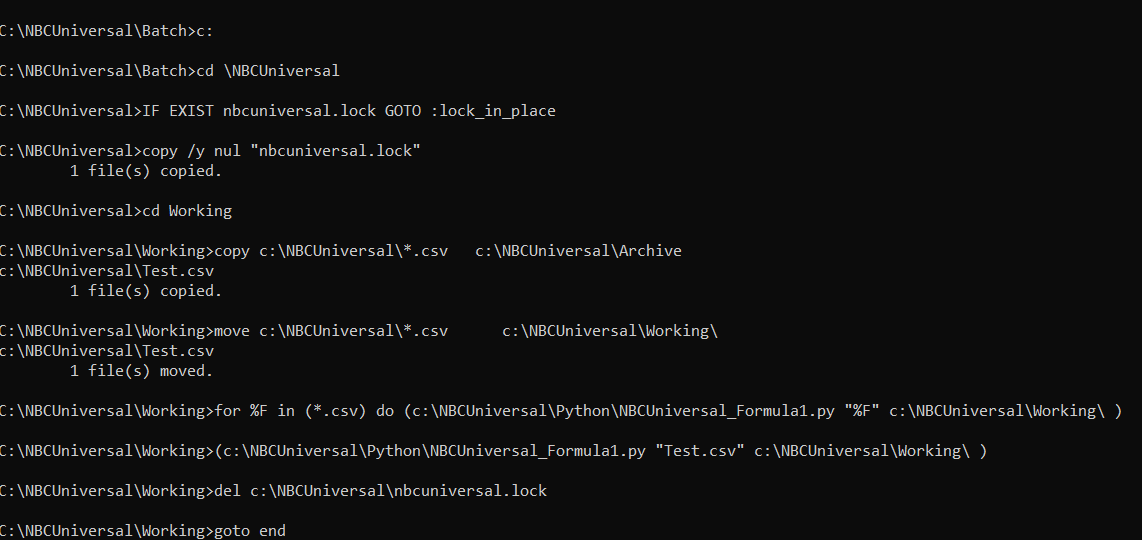
NBCUniversal\_Formula1.py -> c:\NBCUniversal\Python

Test files would need to be placed at the root: c:\NBCUniversal directory as our landing pad. In a real world environment, we would have an infrastructure set up already where the file would be ftp’d to this location, and a subsequent scheduler would execute the .bat script at certain intervals. (That was my setup at my previous shop). But that is beyond the scope of this exercise.

Once one or many .csv files are placed at the landing pad, the script can be executed by initiating the .bat script at the command line level:







More than one file can be placed at the landing pad. Each will be run individually.

If you prefer to not use the batch script, this command will also work. But since you are bypassing the .bat, please put your test file in the c:\NBCUniversal\Working folder. And please note the file will not be archived, this is handled by the .bat script:



An original copy of the file will be placed in the c:\NBCUniversal\Archive folder. A new file will be placed in the c:\NBCUniversal\Final folder with the same name and todays date attached to it.

The python program will take the file and calculate the average lap time of all drivers in the .csv file. It will then report on the top 3, lowest to highest. It will report in a dense ranking, meaning if there are ties it will report on all ties with no gaps. So while you will get the top 3, you can have more than three records reported.

