In order to make the code production ready:

* The functions currently written in one .py files will be packaged using classes .The pandas dataframes are passed as argument , can be inherited will provide faster response from python.
* Error handling is very basic need to add logic for :
  + Data cleaning and handling data related errors:
    - Null values can be handled in load process
    - Type cast errors
    - Encoding errors
  + Handle issues when worksheet names are different or is not available

From infrastructure perspective:

* Current deployment is single threaded and memory utilization will be based on the volume of data .The 3 requirements are independent of each other and can be run in parallel using multiple threads or using pyspark where spark engine handles parallelism.
* An increase in data volume can be handled using scaling horizontally .Again spark engines provide the functionality which is embedded in them.
* Better handling of OS errors
* Operating system portability can be built in. Currently tested in windows laptop.