Python Project

1. Download the zip file to your VM ,extract the zip file content to folder name **/stocks**
2. The zip file contain the following files
   1. **Symbols\_valid\_meta.csv** - CSV with the Stock Symbol and its full name   
      For example **A** - **Agilent Technologies, Inc. Common Stock**
   2. Under folder **stocks** - there is CSV for each company symbol and its rates
3. Develop Python Pipeline **load\_stock\_history.py** that reads the company's CSV files from **/stocks** folder 1 by 1 from the file system and writes the data to the **stock\_rates** table in MySQL DB as follows:a. add the **company\_name** from the CSV **Symbols\_valid\_meta.csv** (don't load the **Symbols\_valid\_meta.csv** to the database)

b. create table according to the CSV columns + **company\_name** field

c. Use Pandas to read the CSV files and write the database in chunks of 500 rows. The process should enrich the company name before writing to the database .

1. Create a python module **stock\_details** that retrieves company stock details by the following filters . The maximum range should be 30 days.

Incase filters are not accurate raise exception

* 1. Inputs
     1. Company/stock symbol
     2. Start\_date
     3. End\_date
  2. Json array of the results - verify that the json format are valid

1. Create a python module **stocks\_stats** that retrieves companןקד stock details by the following filters
   1. Input
      1. Json - List of stock symbols
      2. Start\_date
      3. end\_date
   2. Output
      1. Json array with the following fields
         1. Stock Symbol
         2. Company full name
         3. MAX Rate
         4. MIN Rate
         5. AVG Rate
         6. Yield

Sort by Yield descending

1. General Instructions
   1. Use the Virtualbox for development and running pipelines on several companies + testing
   2. You will have a strong VM to test whole company data
   3. Add module which handle the logging mechanism and use it to log important steps
   4. Use a table called ETL\_PROCESS\_LOGS to write the progress and logs of each company . The table should contain the following fields
      1. ID - auto increment - PK
      2. ETL\_PROCESS - for now **stocks\_data**
      3. Log\_time - default courant time
      4. Log\_level - INFO Default , ERROR in case of error
      5. ERROR\_DESC - Error Description