Webapplication-project

This is code factory of Project Phase 1 - Data Collection Module (Group Task). We implemented following programs.

- Python Web Crawler for Historical Stock Data
  - folder name: Xin_Yang_Historical_Python
- Python Realtime Stock Data Getter
  - folder name: Song_Yang_Realtime_Python
- Python Implementation with third-party API
  - folder name: Zhuohang_Li_API_Python

You will see Readme file in each folder. We highly recommend to check our project on Github. We will stop pushing after the deadline.

Github Link : https://github.com/Lynnes001/Webapplication-project

General Instruction

PLEASE MAKE SURE YOU HAVE ALREADY INSTALLED 'pip' for Python!

- Environment:
  - Python 2.7
  - Google Chrome
  - Jupyter Notebook:
    > pip install jupyter notebook
  - MacOS/Ubuntu (not required)

For realtime data program

- Dependencies:
  - MySQLdb:
    > pip install mysqldb-python
  - request 2.8
    > pip install requests

For historical data program

- Dependencies:
  - MySQLdb:
    > pip install mysqldb-python
  - numpy:
    > pip install numpy
  - Beautiful Soup:
    > pip install beautifulsoup4

- Selenium 3.9

  > pip install -U selenium or pip install selenium

- chromedriver: Download drivers for your browsers: https://pypi.python.org/pypi/selenium/

Build Database:

1. Login to mysql as root
2. Load stockDB.sql

   > Mysql>source <Fullpath.sql>

   - Example:

     > Mysql>source D:\test\ss.sql

   - If failed, you can try following steps

```
CREATE DATABASE stockDB;
USE stockDB;

CREATE TABLE Real_Time_Data(
    rid INT UNSIGNED AUTO_INCREMENT NOT NULL,
    symbol VARCHAR(4) NOT NULL,
    stock_time DATETIME,
    price FLOAT,
    volume INT,
    PRIMARY KEY(rid)
);

CREATE TABLE History_Time_Data(
    hid INT UNSIGNED AUTO_INCREMENT NOT NULL,
    symbol VARCHAR(4) NOT NULL,
    stock_date DATE,
    open_price FLOAT,
    high_price FLOAT,
    low_price FLOAT,
    close_price FLOAT,
    adj_close FLOAT,
    volume INT,
    PRIMARY KEY(hid)
);
```

Output Database to .csv File

```
USE stockDB;
SELECT * FROM Real_Time_Data
INTO OUTFILE '/var/lib/mysql-files/mytable.csv'
FIELDS TERMINATED BY ','
OPTIONALLY ENCLOSED BY '"'
LINES TERMINATED BY '\n';
```