Lab3 – read me

1. Installation

Since we both had problems installing phpmyadmin, we decided to directly use mysql.

```
cindy@cindy-QEMU-Virtual-Machine:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 8
Server version: 8.0.34-0ubuntu0.22.04.1 (Ubuntu)

Copyright (c) 2000, 2023, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql>
```

2. Data collection/ storage & preprocessing

2.1 Symbol check

We first downloaded stock symbol online and traversed to get initial respond for each of the symbol using yfinance for future reference. This was done in **getnifo.py** with **stocksymbol.xlsx**, and **symbol_response.xlsx** is the new file with stock history response:

	0	1
0	AA	Date2023-03-09 47.6238862023-03-10 44.6511192023-03-13 44.45227420
1	AAC	Date2023-03-09 10.2802023-03-10 10.2002023-03-13 10.2702023-03-14 1
2	AACG	Date2023-03-09 2.472023-03-10 2.002023-03-13 1.992023-03-14 1.92202
3	AACI	Date2023-03-09 10.042023-03-10 10.042023-03-13 10.042023-03-14 10.0
4	AACIU	Date2023-03-09 10.052023-03-10 10.052023-03-13 10.052023-03-14 10.0
5	AACT	Date2023-06-13 10.112023-06-14 10.102023-06-15 10.112023-06-16 10.1
6	AADI	Date2023-03-09 8.222023-03-10 7.402023-03-13 8.002023-03-14 8.38202
7	AAIC	Date2023-03-09 2.932023-03-10 2.822023-03-13 2.702023-03-14 2.81202
8	AAIN	Date2023-03-09 22.2658732023-03-10 22.2658732023-03-13 22.29491620
9	AAL	Date2023-03-09 15.882023-03-10 15.462023-03-13 14.852023-03-14 14.6
10	AAM	Series([], Name: Adj Close, dtype: float64)
11	AAMC	Date2023-03-09 58.9800002023-03-10 59.8600012023-03-13 53.50600120
12	AAME	Date2023-03-09 2.3931112023-03-10 2.3436672023-03-13 2.4227782023-0
13	AAN	Date2023-03-09 11.2732842023-03-10 10.9304812023-03-13 10.24487820
14	AAOI	Date2023-03-09 2.812023-03-10 2.662023-03-13 2.672023-03-14 2.6620
15	AAON	Date2023-03-09 61.1361392023-03-10 60.1325952023-03-13 58.43787420

2.2 Portfolio creation & data preprocessing

Then we let user create portfolio and store the data using **data_storage.py** script. Inside the function, first we check the response of user input stock symbol using the last generated symbol_response.xlsx. If all symbols are valid then we get the adjusted close data for the past 2 years. Then calculate metrics based on the data fetched, and finally store the portfolio stock data, the history data and the portfolio metric data separately for each portfolio. There will be null analysis and calculated metrics shown on the terminal to help users know their results.

```
cindy@cindy-QEMU-Virtual-Machine:~/Desktop/XinyiZhang_9328705976/lab03$ python3
data_storage.py
Please enter name of your portfolio: portfolio1
Please enter name of your intended stocks (Use comma and space to seperate your
symbols. Sample input: AAPL, WMT, IBM, META, AMZN, GOOG):
AAPL, WMT, IBM, META, AMZN, GOOG
Please type in the weight for renewed each stock (Use comma and space to separat
e. Sample input: 0.3, 0.2, 0.2, 0.1, 0.2):
0.3, 0.2, 0.2, 0.1, 0.1, 0.1
Downloading...
[**********************100%**********************]  6 of 6 completed
Null analysis of close stock price: AAPL    0
AMZN    0
GOOG    0
IBM     0
META    0
WMT     0
dtype: int64
        metrics     value
0    return_std  0.016913
1   return_mean -0.009749
2  sharpe_ratio -9.150498
```

```
mysql> show tables;
+-------------------------+
| Tables_in_yahoo_finance |
+-------------------------+
| portfolio1_content      |
| portfolio1_history      |
| portfolio1_metrics      |
| portfolio2_content      |
| portfolio2_history      |
| portfolio2_metrics      |
+-------------------------+
6 rows in set (0.00 sec)
```

From the results, this API is very robust, no missing values are found.

2.3 Portfolio management

Then we created a portfolio management system using the **portfolio_management.py** file. It's a loop so users could continuously update their portfolio. The system will display current portfolio names and creation time, current stocks inside the portfolio in interest, adding and deleting new stock in the portfolio (importing function from data_storage.py).

Displaying all portfolios and asking user to explore one of them:

```
cindy@cindy-QEMU-Virtual-Machine:~/Desktop/XinyiZhang_9328705976/lab03$ python3
portfolio_management.py
Please indicate which type of operation you want to implement (Choose from add,
delete, display):
display
Creation time of all portfolios:
[('portfolio1', '09/09/2023, 11:38:27'), ('portfolio2', '09/09/2023, 11:39:00')]
Please enter the name of the portfolio you want to explore/edit: portfolio1
Current stocks inside this portfolio:
AAPL
AMZN
GOOG
IBM
META
WMT
Still want to continue?(y/n)
```

Allowing users to add stock to their portfolios:

```
Please indicate which type of operation you want to implement (Choose from add, de
lete, display):
delete
Please enter the name of the portfolio you want to explore/edit: portfolio2
Current stocks inside this portfolio:
AAPL
IBM
META
WMT
Please type in the symbol of stock which you want to delete in this portfolio: WMT
Please type in the weight for renewed each stock (Use comma and space to separate.
 Sample input: 0.3, 0.2, 0.2, 0.1, 0.2):
0.3, 0.3, 0.4
Downloading...
[**********************100%%**********************]  3 of 3 completed
Null analysis of close stock price:
 AAPL    0
IBM     0
META    0
dtype: int64
        metrics     value
0    return_std  0.018822
1   return_mean -0.009650
2  sharpe_ratio -8.138825
Still want to continue?(y/n)
```

Allowing users to delete stocks from their portfolios:

```
Please indicate which type of operation you want to implement (Choose from add,
delete, display):
add
Please enter the name of the portfolio you want to explore/edit: portfolio2
Current stocks inside this portfolio:
AAPL
IBM
WMT
Please type in the symbol of stock which you want to add in this portfolio: META
Please type in the weight for renewed each stock (Use comma and space to separat
e. Sample input: 0.3, 0.2, 0.2, 0.1, 0.2):
0.2, 0.2, 0.3, 0.3
Downloading...
[**********************100%%**********************]  4 of 4 completed
Null analysis of close stock price:
 AAPL    0
IBM     0
META    0
WMT     0
dtype: int64
        metrics      value
0    return_std   0.015163
1   return_mean  -0.009649
2  sharpe_ratio -10.101805
Still want to continue?(y/n)
```

There's a validation check for wrongly spelled portfolios:

```
Please indicate which type of operation you want to implement (Choose from add, de
lete, display):
delete
Please enter the name of the portfolio you want to explore/edit: portfolio
Wrong portfolio name! Please try again.
```

Also validation checks for stock symbol names:

```
Please indicate which type of operation you want to implement (Choose from add, de
lete, display):
add
Please enter the name of the portfolio you want to explore/edit: portfolio2
Current stocks inside this portfolio:
AAPL
IBM
META
Please type in the symbol of stock which you want to add in this portfolio: AKO
Please type in the weight for renewed each stock (Use comma and space to separate.
 Sample input: 0.3, 0.2, 0.2, 0.1, 0.2):
0.3, 0.3, 0.3, 0.3
Stock symbol AKO not found, please try again.
```

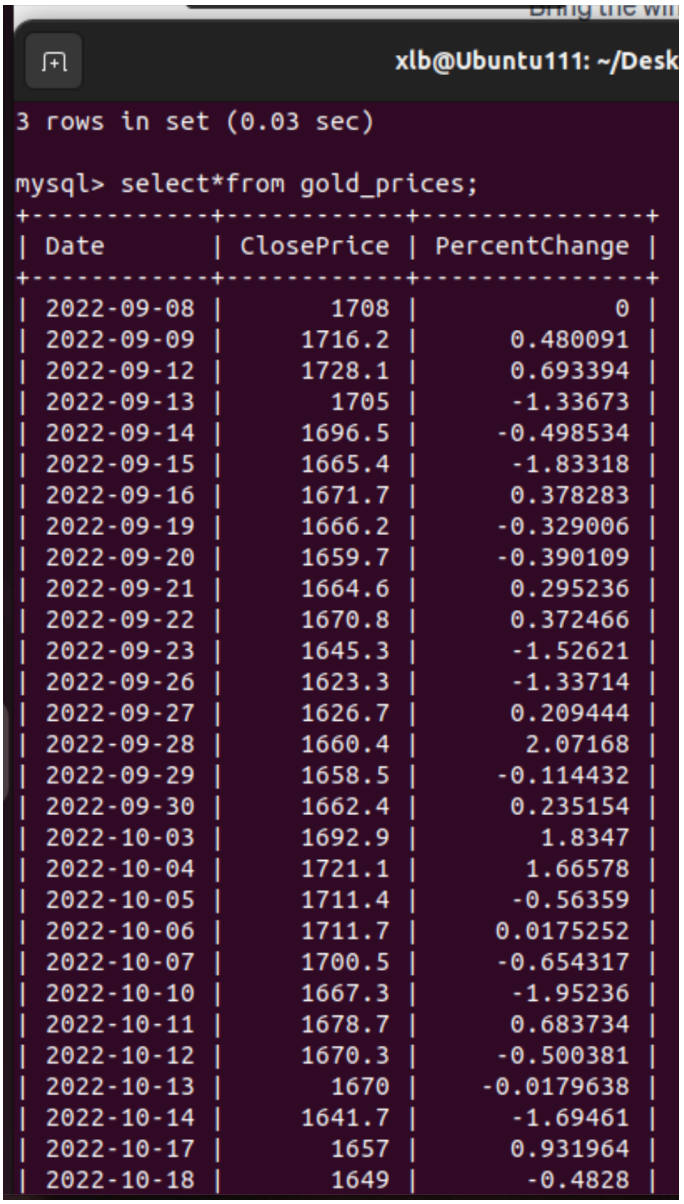
3. Other data gathering (intended for Lab 3 Part 2)

We aim to investigate the relationships between the price of gold, weather data in major cities, and stock prices to ascertain whether significant correlations exist between them. Our hypothesis posits that gold and stocks exhibit a negative correlation: as stock prices decrease, gold prices increase, and vice versa. To validate this hypothesis, we will employ the following methodologies:

- Exploratory Data Analysis (EDA): Initial examination of the data will provide insights into trends, patterns, and potential anomalies.
- Regression Analysis: This will help determine the strength and nature of the relationship between gold and stock prices.

- Time Series Models: Given the temporal nature of our data, models like ARIMA and VAR will be particularly useful in discerning any causative influences and forecasting future trends.

Furthermore, we will delve into machine learning to possibly uncover nuanced patterns or relationships in the data, providing a deeper understanding than traditional statistical methods might offer. The weather data were directly downloaded from www.weather.gov and https://www.visualcrossing.com/weather/weather-data-services/San%20Francisco,USA/metric/2021-09-01/2023-09-08, while gold data acquired using yfinance in **goad.py** and uploaded using **Mysql_cone.py**.
Example of external data:




```
3 rows in set (0.03 sec)

mysql> select*from gold_prices;
+------------+------------+---------------+
| Date       | ClosePrice | PercentChange |
+------------+------------+---------------+
| 2022-09-08 |       1708 |             0 |
| 2022-09-09 |     1716.2 |      0.480091 |
| 2022-09-12 |     1728.1 |      0.693394 |
| 2022-09-13 |       1705 |      -1.33673 |
| 2022-09-14 |     1696.5 |     -0.498534 |
| 2022-09-15 |     1665.4 |      -1.83318 |
| 2022-09-16 |     1671.7 |      0.378283 |
| 2022-09-19 |     1666.2 |     -0.329006 |
| 2022-09-20 |     1659.7 |     -0.390109 |
| 2022-09-21 |     1664.6 |      0.295236 |
| 2022-09-22 |     1670.8 |      0.372466 |
| 2022-09-23 |     1645.3 |      -1.52621 |
| 2022-09-26 |     1623.3 |      -1.33714 |
| 2022-09-27 |     1626.7 |      0.209444 |
| 2022-09-28 |     1660.4 |       2.07168 |
| 2022-09-29 |     1658.5 |     -0.114432 |
| 2022-09-30 |     1662.4 |      0.235154 |
| 2022-10-03 |     1692.9 |        1.8347 |
| 2022-10-04 |     1721.1 |       1.66578 |
| 2022-10-05 |     1711.4 |      -0.56359 |
| 2022-10-06 |     1711.7 |     0.0175252 |
| 2022-10-07 |     1700.5 |     -0.654317 |
| 2022-10-10 |     1667.3 |      -1.95236 |
| 2022-10-11 |     1678.7 |      0.683734 |
| 2022-10-12 |     1670.3 |     -0.500381 |
| 2022-10-13 |       1670 |    -0.0179638 |
| 2022-10-14 |     1641.7 |      -1.69461 |
| 2022-10-17 |       1657 |      0.931964 |
| 2022-10-18 |       1649 |       -0.4828 |
```

Note that if you try to run the codes, mysql may cause error because of localhost. Please refer to our report and video for more information or change the path of MySQL engine to your own path (in both **data_storage.py, portfolio_management.py and Mysql_cone.py**). Thank you!

Reporter: New Bee team
Zegang Cui 6293177952

Libang Xia 7450994620
Xinyi Zhang 9328705976