Project Code:
Github link: https://github.com/christalwang35/si507_final_project/tree/main

README:

When running the program:

First Step: Enter the stock code of Chinese A-stock, for example 000001 or 601111.

Second Step: Enter a start year, for example 2020 or 2021.

Third Step: Enter a end year, which is later than the start year, for example 2021 or 2022.

Forth Step: Enter a number representing the principal money amount. This number cannot be too small, preventing the amount from being too small to complete the simulated transaction. A number like 100000 or 1000000 is suitable.

Fifth Step: After completing login, logout and simulating. The program will show the loss and win rate, average PE, average PB, average PS and total profit. There will be a tree method questioning. We can answer “yes” or “no” to get the investment advice.

Sixth Step: There will be two graphs shown.

Python Packs:

```
import sys
listone=sys.path
listone.append('C:\\Users\\user\\AppData\\Local\\Programs\\Python\\Python37\\Lib\\site-packages')
sys.path = listone
from mplfinance.original_flavor import candlestick_ohlc
from matplotlib.dates import date2num
import matplotlib.pyplot as plt
import requests
from dateutil import parser
from time import sleep
from datetime import datetime,time,timedelta
import pandas as pd
import os
import numpy as np
import baostock as bs
import pysnooper
```

Data Sources:

Origins: www.baostock.com

Format: CSV

How to access:

Use the function from baostock pack to get the PE, PB and time-series data. The code uses caching to store data as DataFrame form and then saves data as csv files to local.

Summary of data:

There are more than 20000 records available each year, a new record is generated every 5 minutes. I will mainly use five data in the record: datetime, open, high, low, close, which respectively represent the time node, opening price, highest price, lowest price and closing price of the stock price. Based on the above five data, simulate the change of the stock price in the selected period.

Data Structure:

README:

Run the function run_backtesting_dual_thrust to perform the following.

The function stock_PEPB gets the stock' s PE and PB from previously saved csv files. The function get_ticks_for_backtesting gets the stock' s price, which is given every 5 minutes, from previously saved csv files. Then it estimates price changes within the 5-minute intervals, saving the results in attribute as a list.

According to financial knowledge, the function strategy_dual_thrust imitates trading based on the function Buy, and the function Sell. Then, it saves the date purchased and sold, price, and number of stocks as dictionary into the attribute.

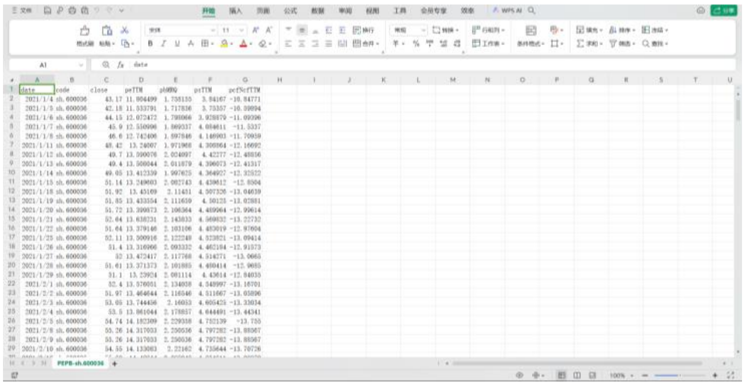
using stock prices generated by the function get_ticks_for_backtesting, the function candlestick_ohlc from the package mplfinance.original_flavor generates a graph that shows the stock price k-line. The function ax.plot uses information from the function strategy_dual_thrust, and generates blue dots which represent price and date of purchase and sell.

Based on the difference between prices of purchase and sell, the function plt.subplots generates a bar chart that shows loss and gain of each trading.

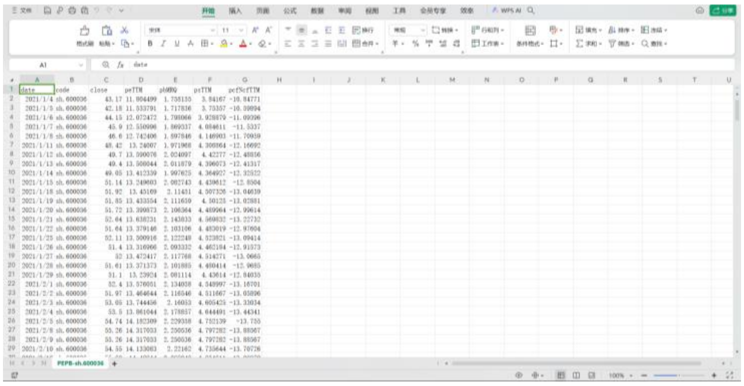
According to PE and PB obtained by the function stock_PEPB, a tree method is implemented, which allows users to answer questions based on the average PE of their ideal stocks to determine whether this stock is the user's ideal stock.

Screenshots and description:

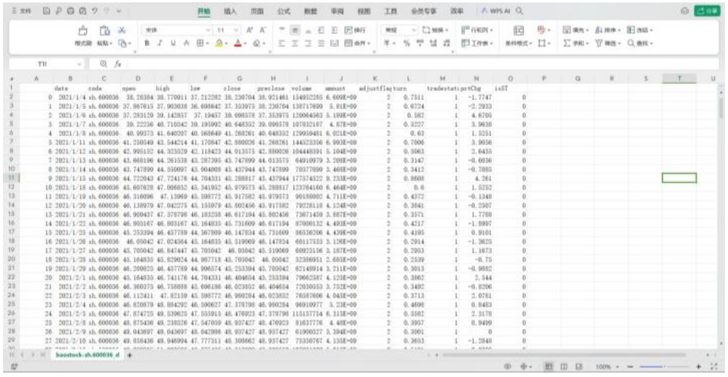
Data:



date	code	close	peTTM	pbMRQ	psTTM	pcfNcfTTM
2021/1/4	sh.600036	43.17	11.804499	1.758135	3.84167	-10.84771
2021/1/5	sh.600036	42.18	11.533791	1.717836	3.75357	-10.59894
2021/1/6	sh.600036	44.15	12.072472	1.798066	3.928879	-11.09396
2021/1/7	sh.600036	45.9	12.550998	1.869337	4.084611	-11.5337
2021/1/8	sh.600036	46.6	12.742406	1.897846	4.146903	-11.70959
2021/1/11	sh.600036	48.42	13.24007	1.971968	4.308864	-12.16692
2021/1/12	sh.600036	49.7	13.590076	2.024097	4.42277	-12.48856
2021/1/13	sh.600036	49.4	13.508044	2.011879	4.396073	-12.41317
2021/1/14	sh.600036	49.05	13.412339	1.997625	4.364927	-12.32522
2021/1/15	sh.600036	51.14	13.249603	2.082743	4.439612	-12.8504
2021/1/18	sh.600036	51.92	13.45169	2.11451	4.507326	-13.04639
2021/1/19	sh.600036	51.85	13.433554	2.111659	4.50125	-13.02881
2021/1/20	sh.600036	51.72	13.399873	2.106364	4.489964	-12.99614
2021/1/21	sh.600036	52.64	13.638231	2.143833	4.569832	-13.22732
2021/1/22	sh.600036	51.64	13.379146	2.103106	4.483019	-12.97604
2021/1/25	sh.600036	52.11	13.500916	2.122248	4.523821	-13.09414
2021/1/26	sh.600036	51.4	13.316966	2.093332	4.462184	-12.91573
2021/1/27	sh.600036	52	13.472417	2.117768	4.514271	-13.0665
2021/1/28	sh.600036	51.61	13.371373	2.101885	4.480414	-12.9685
2021/1/29	sh.600036	51.1	13.23924	2.081114	4.43614	-12.84035
2021/2/1	sh.600036	52.4	13.576051	2.134058	4.548997	-13.16701
2021/2/2	sh.600036	51.97	13.464644	2.116546	4.511667	-13.05896
2021/2/3	sh.600036	53.05	13.744456	2.16053	4.605425	-13.33034
2021/2/4	sh.600036	53.5	13.861044	2.178857	4.644491	-13.44341
2021/2/5	sh.600036	54.74	14.182309	2.229358	4.752139	-13.755
2021/2/8	sh.600036	55.26	14.317033	2.250536	4.797282	-13.88567
2021/2/9	sh.600036	55.26	14.317033	2.250536	4.797282	-13.88567
2021/2/10	sh.600036	54.55	14.133083	2.22162	4.735644	-13.70726



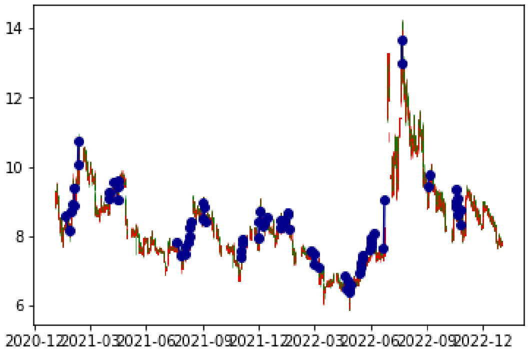
date	code	close	peTTM	pbMRQ	psTTM	pcfNcfTTM
2021/1/4	sh.600036	43.17	11.804499	1.758135	3.84167	-10.84771
2021/1/5	sh.600036	42.18	11.533791	1.717836	3.75357	-10.59894
2021/1/6	sh.600036	44.15	12.072472	1.798066	3.928879	-11.09396
2021/1/7	sh.600036	45.9	12.550998	1.869337	4.084611	-11.5337
2021/1/8	sh.600036	46.6	12.742406	1.897846	4.146903	-11.70959
2021/1/11	sh.600036	48.42	13.24007	1.971968	4.308864	-12.16692
2021/1/12	sh.600036	49.7	13.590076	2.024097	4.42277	-12.48856
2021/1/13	sh.600036	49.4	13.508044	2.011879	4.396073	-12.41317
2021/1/14	sh.600036	49.05	13.412339	1.997625	4.364927	-12.32522
2021/1/15	sh.600036	51.14	13.249603	2.082743	4.439612	-12.8504
2021/1/18	sh.600036	51.92	13.45169	2.11451	4.507326	-13.04639
2021/1/19	sh.600036	51.85	13.433554	2.111659	4.50125	-13.02881
2021/1/20	sh.600036	51.72	13.399873	2.106364	4.489964	-12.99614
2021/1/21	sh.600036	52.64	13.638231	2.143833	4.569832	-13.22732
2021/1/22	sh.600036	51.64	13.379146	2.103106	4.483019	-12.97604
2021/1/25	sh.600036	52.11	13.500916	2.122248	4.523821	-13.09414
2021/1/26	sh.600036	51.4	13.316966	2.093332	4.462184	-12.91573
2021/1/27	sh.600036	52	13.472417	2.117768	4.514271	-13.0665
2021/1/28	sh.600036	51.61	13.371373	2.101885	4.480414	-12.9685
2021/1/29	sh.600036	51.1	13.23924	2.081114	4.43614	-12.84035
2021/2/1	sh.600036	52.4	13.576051	2.134058	4.548997	-13.16701
2021/2/2	sh.600036	51.97	13.464644	2.116546	4.511667	-13.05896
2021/2/3	sh.600036	53.05	13.744456	2.16053	4.605425	-13.33034
2021/2/4	sh.600036	53.5	13.861044	2.178857	4.644491	-13.44341
2021/2/5	sh.600036	54.74	14.182309	2.229358	4.752139	-13.755
2021/2/8	sh.600036	55.26	14.317033	2.250536	4.797282	-13.88567
2021/2/9	sh.600036	55.26	14.317033	2.250536	4.797282	-13.88567
2021/2/10	sh.600036	54.55	14.133083	2.22162	4.735644	-13.70726



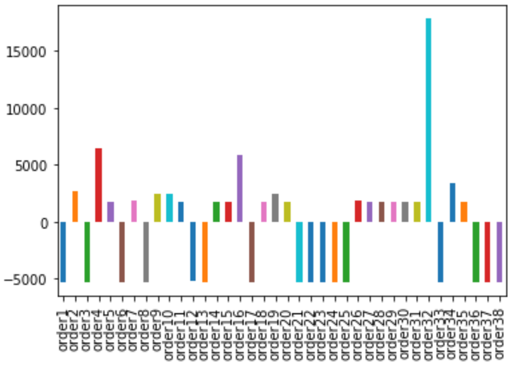
	date	code	open	high	low	close	preclose	volume	amount	adjustflag	turn	tradestat	pctChg	isST
0	2021/1/4	sh.600036	38.28384	38.770911	37.212282	38.230704	38.921461	134952285	6.609E+09	2	0.7511	1	-1.7747	0
1	2021/1/5	sh.600036	37.867615	37.903038	36.698642	37.353973	38.230704	138717609	5.81E+09	2	0.6724	1	-2.2933	0
2	2021/1/6	sh.600036	37.283129	39.142857	37.19457	39.098578	37.353973	120064563	5.189E+09	2	0.582	1	4.6706	0
3	2021/1/7	sh.600036	39.22256	40.710342	39.195992	40.648352	39.098578	107832187	4.87E+09	2	0.5227	1	3.9638	0
4	2021/1/8	sh.600036	40.99373	41.640207	40.568649	41.268261	40.648352	129959481	6.021E+09	2	0.63	1	1.5251	0
5	2021/1/11	sh.600036	41.230549	43.544214	41.170847	42.880026	41.268261	144523336	6.993E+09	2	0.7006	1	3.9056	0
6	2021/1/12	sh.600036	42.995152	44.323529	42.118423	44.013575	42.880026	104448591	5.104E+09	2	0.5063	1	2.6435	0
7	2021/1/13	sh.600036	43.668196	44.261538	43.287395	43.747899	44.013575	64910979	3.208E+09	2	0.3147	1	-0.6036	0
8	2021/1/14	sh.600036	43.747899	44.350097	43.004008	43.437944	43.747899	70377099	3.468E+09	2	0.3412	1	-0.7085	0
9	2021/1/15	sh.600036	44.722043	47.724176	44.704331	45.288817	43.437944	177574522	9.233E+09	2	0.8608	1	4.261	0
10	2021/1/18	sh.600036	45.607628	47.006852	45.341952	45.979573	45.288817	123764160	6.464E+09	2	0.6	1	1.5252	0
11	2021/1/19	sh.600036	46.316096	47.13969	45.598772	45.917582	45.979573	90188002	4.711E+09	2	0.4372	1	-0.1348	0
12	2021/1/20	sh.600036	46.138979	47.042275	45.155979	45.802456	45.917582	79228118	4.124E+09	2	0.3841	1	-0.2507	0
13	2021/1/21	sh.600036	46.909437	47.378798	46.183258	46.617194	45.802456	73671459	3.887E+09	2	0.3571	1	1.7788	0
14	2021/1/22	sh.600036	46.903167	46.903167	45.164835	45.731609	46.617194	87000132	4.493E+09	2	0.4217	1	-1.8997	0
15	2021/1/25	sh.600036	45.253394	46.457789	44.367809	46.147834	45.731609	86536206	4.439E+09	2	0.4195	1	0.9101	0
16	2021/1/26	sh.600036	46.05042	47.024564	45.164835	45.519069	46.147834	60117533	3.126E+09	2	0.2914	1	-1.3625	0
17	2021/1/27	sh.600036	45.705042	46.847447	45.705042	46.05042	45.519069	60923136	3.187E+09	2	0.2953	1	1.1673	0
18	2021/1/28	sh.600036	45.164835	45.829024	44.987718	45.705042	46.05042	52386951	2.685E+09	2	0.2539	1	-0.75	0
19	2021/1/29	sh.600036	46.209825	46.437789	44.996574	45.253394	45.705042	62148914	3.211E+09	2	0.3013	1	-0.9882	0
20	2021/2/1	sh.600036	45.164835	46.741176	44.704331	46.404654	45.253394	79662587	4.123E+09	2	0.3862	1	2.544	0
21	2021/2/2	sh.600036	46.360375	46.738888	45.696186	46.023852	46.404654	72030353	3.752E+09	2	0.3492	1	-0.8206	0
22	2021/2/3	sh.600036	46.112411	47.82159	45.598772	46.980284	46.023852	76587806	4.045E+09	2	0.3713	1	2.0781	0
23	2021/2/4	sh.600036	46.820879	48.884292	46.590627	47.378798	46.980284	96910977	5.23E+09	2	0.4698	1	0.8483	0
24	2021/2/5	sh.600036	47.874725	49.539625	47.553915	48.476923	47.378798	115157734	6.315E+09	2	0.5582	1	2.3178	0
25	2021/2/8	sh.600036	48.875436	49.238526	47.547039	48.937427	48.476923	81637776	4.48E+09	2	0.3957	1	0.9499	0
26	2021/2/9	sh.600036	49.043697	49.043697	48.042986	48.937427	48.937427	61900327	3.394E+09	2	0.3001	1	0	0
27	2021/2/10	sh.600036	49.858436	49.946994	47.777311	48.308662	48.937427	75330767	4.133E+09	2	0.3653	1	-1.2848	0

Graphs:

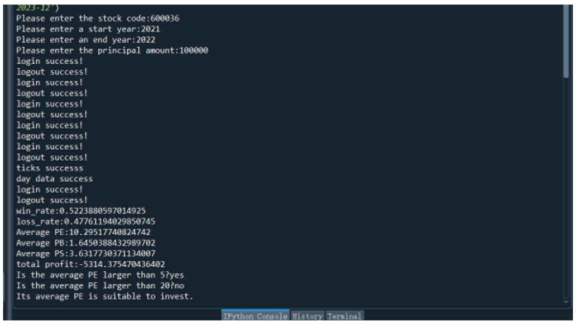
The following graph is created by using the above data records and stock prices within the 5-minute intervals estimated based on financial and statistical knowledge. It shows the stock price K-line according to the stock code and time interval entered by the user.The blue lines connect the buying points and selling points of the simulated transactions on the K-line chart.



The following graph is a histogram drawn according to the return of simulated trading in the backtesting.



Tree:



```
2023-12')
Please enter the stock code:600036
Please enter a start year:2021
Please enter an end year:2022
Please enter the principal amount:100000
login success!
logout success!
login success!
logout success!
login success!
logout success!
login success!
logout success!
login success!
logout success!
ticks successs
day data success
login success!
logout success!
win_rate:0.5223880597014925
loss_rate:0.47761194029850745
Average PE:10.29517740824742
Average PB:1.6450388432989702
Average PS:3.6317730371134007
total profit:-5314.375470436402
Is the average PE larger than 5?yes
Is the average PE larger than 20?no
Its average PE is suitable to invest.
```

Interaction and Presentation Options

Users can enter any existing Chinese A-stock stock code and any time span during the listing period of the stock.

Users can answer the questions based on the performance of their ideal stock, and let the program judge whether the stock is worth investing.

The interactive technologies used are command line prompts.

Video Link:

https://youtu.be/99jX-wdwSCE?feature=shared