



# *Assignment 1*

## *Stock Price Forecast*

Meng-Chien Lin

Po-Chih Kuo

# Goal

- To make stock price forecasting by using the stock price of a company to predict the other in the same industry
- To implement the regression model to achieve the forecasting
- Preprocess the data for model training/testing

## Basic (100%)

- Use MTK's stock prices to predict TSMC's stock prices of 20 days followed by the due day (10/14) of the homework.
- That is, to predict the prices from 10/15 to 11/11 (without holidays, totally 20 days)

2021年10月						
星期天	星期一	星期二	星期三	星期四	星期五	星期六
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

Michel Zbinden / 4255

2021年11月						
星期天	星期一	星期二	星期三	星期四	星期五	星期六
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30				

Michel Zbinden / 4255

# Data

- Download link:

[https://www.twse.com.tw/zh/page/trading/exchange/STOCK\\_DAY\\_AVG.html](https://www.twse.com.tw/zh/page/trading/exchange/STOCK_DAY_AVG.html)

## MediaTek Inc. (MTK)

聯發科

957.00 TWD

+347.00 (56.89%) ↑ 過去 1 年

9月28日 下午1:30 [GMT+8] · 免責聲明

TPE: 2454

+ 追蹤

1 天 | 5 天 | 1 個月 | 6 個月 | 本年迄今 | 1 年 | 5 年 | 最久



開盤	959.00	市值	1.52兆	52 週高點	1,185.00
最高	969.00	本益比	18.74	52 週低點	581.00
最低	945.00	殖利率	3.87%		

## Taiwan Semiconductor Manufacturing Co., Ltd (TSMC)

台灣積體電路製造

594.00 TWD

+163.00 (37.82%) ↑ 過去 1 年

9月28日 下午1:30 [GMT+8] · 免責聲明

TPE: 2330

+ 追蹤

1 天 | 5 天 | 1 個月 | 6 個月 | 本年迄今 | 1 年 | 5 年 | 最久



開盤	595.00	市值	15.40兆	GDP 得分	A-
最高	596.00	本益比	27.80	52 週高點	679.00
最低	592.00	殖利率	1.77%	52 週低點	424.50

# How To Download

**臺灣證券交易所**

字體大小 [A](#) [A](#) | [會員專區](#) | [報表索引](#) | [日本語ホームページ](#) | [English Home](#) | [Q](#)

[公開資訊觀測站](#) | [基本市況報導](#) | [TWSE 網站](#) | [國內業務宣導網站](#) ▼

[關於證交所](#) | [交易資訊](#) | [指數資訊](#) | [上市公司](#) | [產品與服務](#) | [結算服務](#) | [市場公告](#) | [法令規章](#)

交易資訊

盤後資訊

- 每日收盤行情
- 每日市場成交資訊
- 每日第一上市外國股票成交量值
- 每日創新板股票成交量值
- 每日成交量前二十名證券
- 每5秒委託成交統計
- 各類指數日成交量值
- 個股日成交資訊
- 當日融券賣出與債券賣出成交量值
- **個股日收盤價及月平均價**
- 個股月成交資訊
- 個股年成交資訊
- 盤後定價交易
- 盤中零股交易行情單
- 盤後零股交易行情單
- 個股日本益比、殖利率及股價淨值比 (依日期查詢)
- 個股日本益比、殖利率及股價淨值比 (依代碼查詢)
- 暫停交易證券

升降幅度/首五日無漲跌幅

變更交易

當日沖銷交易標的

融資融券與可借券賣出額度

標借

◀ 首頁 ▶ 交易資訊 ▶ 盤後資訊 ▶ 個股日收盤價及月平均價


[Change to english version](#) [English](#)

資料日期: 民國 110 年 09 月

股票代碼 (瀏覽):

※ 本資訊自民國88年1月5日起開始提供

# How To Download



Font Size [A](#) [A](#) [A](#) | [Members](#) | [Report Index](#) | [中文首頁](#) | [日本語ホームページ](#) | [Q](#)

[MOPS](#) | [MIS](#) | [TWSE Sites: Domestic Service Portal](#) ▼

[About the TWSE](#) | [Market Info](#) | [Indices Information](#) | [Listed Companies](#) | [Products & Services](#) | [Clearing Services](#) | [Announcements](#) | [Regulations](#)

**Market Info**

[Historical Trading Info/Data](#)

- Daily Quotes
- Highlights of Daily Trading
- Daily Quotes(Primary listed foreign companies)
- Daily Quotes(Taiwan Innovation Board)
- Top 20 Securities by Volume
- Order book & Trade per 5 seconds
- Daily Trade Summary of all Index
- Daily Trading Value/Volume of Individual Securities
- Short Sales Volume and Value
- **Daily Closing Price / Monthly Closing Average Price of Individual Securities**
- Monthly Trading Value / Volume of Individual Securities
- Yearly Trading Value / Volume of Individual Securities
- After-hour Fixed Price Trading
- Intraday Odd lot Trade
- After-hour Odd-lot Trade
- Stock P/E ratio, dividend yield and P/B ratio (Inquiry by date)
- Daily P/E ratio, dividend yield and P/B ratio (Inquiry by stock code)
- Trading Halt of Securities

[Price Variation Limit & Full Delivery](#)

[Full Delivery Securities](#)

[Objects and Statistics for Day Trading](#)

[Margin Trading & Short Sales](#)

[Securities Borrowed through](#)

Home ► Market Info ► Historical Trading Info/Data ► Daily Closing Price / Monthly Closing Average Price of Individual Securities [中文](#)

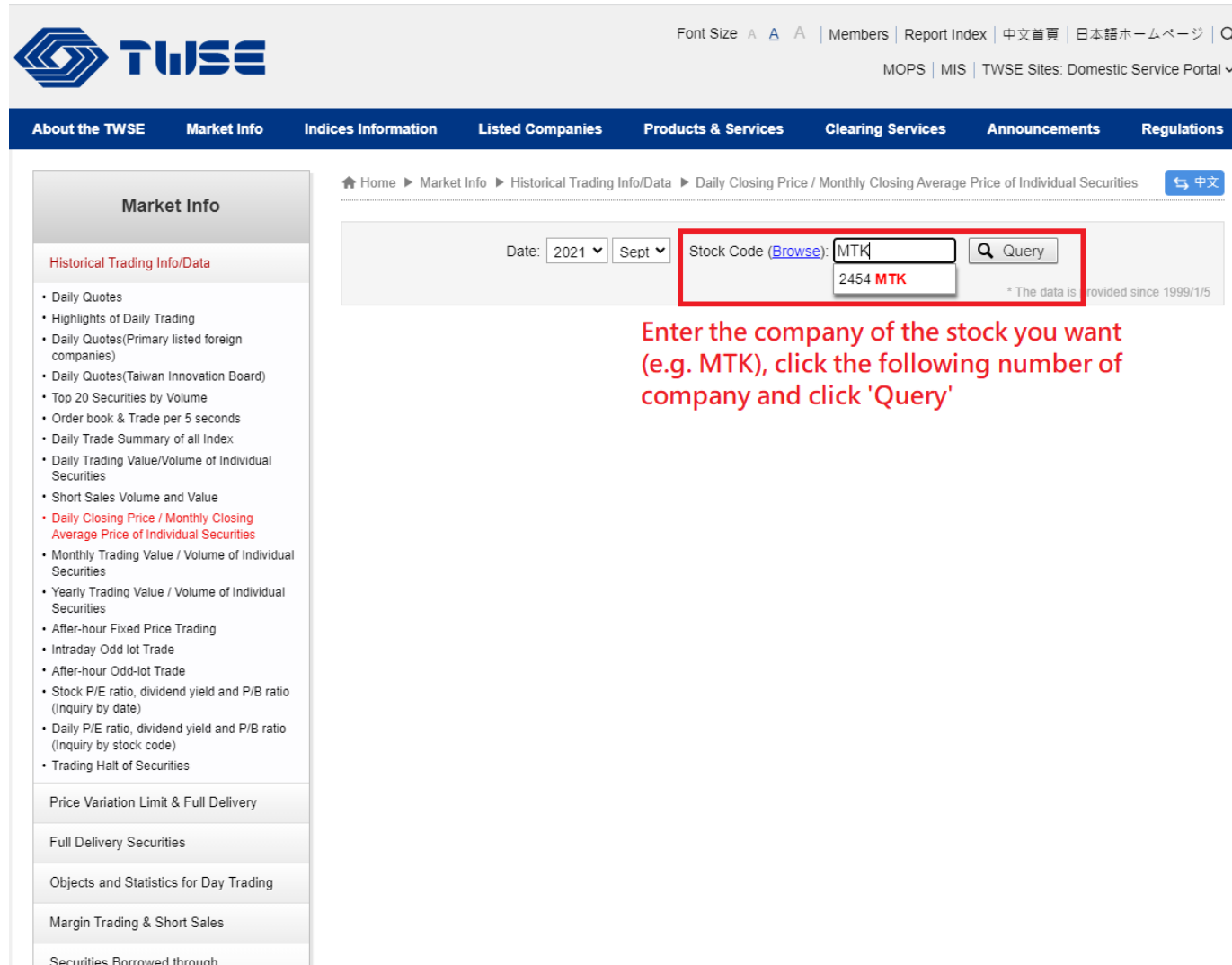
Date:   Stock Code ([Browse](#)):

\* The data is provided since 1999/1/5

Select specific year and month  
(e.g. Sept 2021)



# How To Download



Font Size A A | Members | Report Index | 中文首頁 | 日本語ホームページ | Q

MOPS | MIS | TWSE Sites: Domestic Service Portal v

About the TWSE Market Info Indices Information Listed Companies Products & Services Clearing Services Announcements Regulations

Market Info

Historical Trading Info/Data

- Daily Quotes
- Highlights of Daily Trading
- Daily Quotes(Primary listed foreign companies)
- Daily Quotes(Taiwan Innovation Board)
- Top 20 Securities by Volume
- Order book & Trade per 5 seconds
- Daily Trade Summary of all Index
- Daily Trading Value/Volume of Individual Securities
- Short Sales Volume and Value
- **Daily Closing Price / Monthly Closing Average Price of Individual Securities**
- Monthly Trading Value / Volume of Individual Securities
- Yearly Trading Value / Volume of Individual Securities
- After-hour Fixed Price Trading
- Intraday Odd lot Trade
- After-hour Odd-lot Trade
- Stock P/E ratio, dividend yield and P/B ratio (Inquiry by date)
- Daily P/E ratio, dividend yield and P/B ratio (Inquiry by stock code)
- Trading Halt of Securities

Price Variation Limit & Full Delivery

Full Delivery Securities

Objects and Statistics for Day Trading

Margin Trading & Short Sales

Securities Borrowed through

Home ► Market Info ► Historical Trading Info/Data ► Daily Closing Price / Monthly Closing Average Price of Individual Securities 中文


Date: 2021 ▼ Sept ▼ Stock Code (Browse): MTK 2454 MTK Query

\* The data is provided since 1999/1/5

Enter the company of the stock you want (e.g. MTK), click the following number of company and click 'Query'



# How To Download



Font Size [A](#) [A](#) | [Members](#) | [Report Index](#) | [中文首頁](#) | [日本語ホームページ](#) | [Q](#)

MOPS | MIS | TWSE Sites: Domestic Service Portal [v](#)

About the TWSEMarket InfoIndices InformationListed CompaniesProducts & ServicesClearing ServicesAnnouncementsRegulations

Market Info

Historical Trading Info/Data

- Daily Quotes
- Highlights of Daily Trading
- Daily Quotes(Primary listed foreign companies)
- Daily Quotes(Taiwan Innovation Board)
- Top 20 Securities by Volume
- Order book & Trade per 5 seconds
- Daily Trade Summary of all Index
- Daily Trading Value/Volume of Individual Securities
- Short Sales Volume and Value
- Daily Closing Price / Monthly Closing Average Price of Individual Securities
- Monthly Trading Value / Volume of Individual Securities
- Yearly Trading Value / Volume of Individual Securities
- After-hour Fixed Price Trading
- Intraday Odd lot Trade
- After-hour Odd-lot Trade
- Stock P/E ratio, dividend yield and P/B ratio (Inquiry by date)
- Daily P/E ratio, dividend yield and P/B ratio (Inquiry by stock code)
- Trading Halt of Securities

Price Variation Limit & Full Delivery

Full Delivery Securities

Objects and Statistics for Day Trading

Margin Trading & Short Sales

Securities Borrowed through Competitive Bid Initiated by Securities Finance Company

Foreign & Other Investors

Home > Market Info > Historical Trading Info/Data > Daily Closing Price / Monthly Closing Average Price of Individual Securities [中文](#)

Date: [2021](#) [Sept](#) Stock Code ([Browse](#)):  [Q](#) Query

\* The data is provided since 1999/1/5

[Print / HTML](#) [Download CSV](#)

Click here to download csv file

2021/09 Monthly Closing Average Price of 2454

(NT\$)

Date	Closing Price
2021/09/01	927.00
2021/09/02	933.00
2021/09/03	919.00
2021/09/06	896.00
2021/09/07	903.00
2021/09/08	912.00
2021/09/09	901.00
2021/09/10	925.00
2021/09/13	927.00
2021/09/14	947.00
2021/09/15	935.00
2021/09/16	936.00
2021/09/17	940.00
2021/09/22	911.00
2021/09/23	915.00
2021/09/24	920.00
Monthly Closing Average Price	921.69



# How To Download

- Then you can get one month's stock data !
- Combine several months of data to create the final input file

2021/09 Monthly Closing Average Price of 2454	
Date	Closing Price
2021/9/1	927
2021/9/2	933
2021/9/3	919
2021/9/6	896
2021/9/7	903
2021/9/8	912
2021/9/9	901
2021/9/10	925
2021/9/13	927
2021/9/14	947
2021/9/15	935
2021/9/16	936
2021/9/17	940
2021/9/22	911
2021/9/23	915
2021/9/24	920
Monthly CI	921.69
Remarks:	
If the last digit of the 6 digit alphanumeric ETF security code is K, M, S or C, it is traded in foreign currency.	

# *Report*

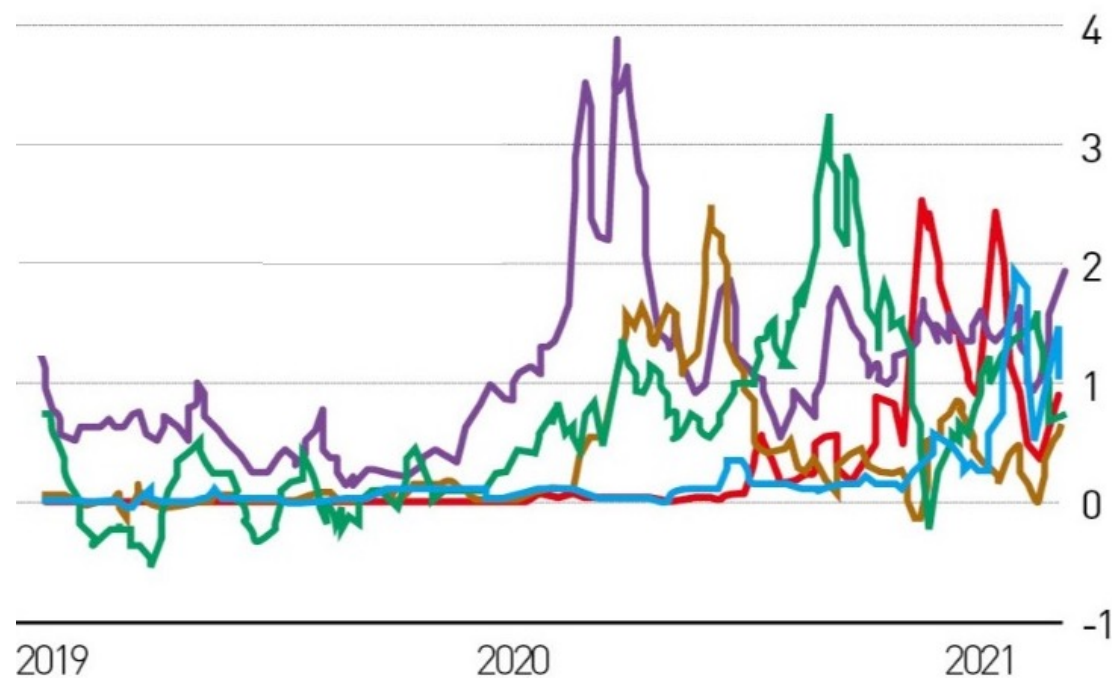
- Briefly describe how to implement regression
- Summarize your work
- No more than two pages

# *Basic Grading Policy*

- Model – 50%
- Performance – 40%
- Report – 10%

## *Bonus (extra 20%)*

- Use more than two companies to predict the stock price of TSMC in the next month (10/15~11/11 without holidays, totally 20 days)



# *Bonus Grading Policy*

- Model -- 5%
- Performance – 10%
- Report – 5% (You should write basic and bonus part of the report together in three pages)

# Template

- Basic part must follow this format (There is no restriction for the format of bonus part)
- You must use the given file “template.py” to build the model
- Except for the imported packages in the template, you cannot use any other packages

```
template.py X
1 #!/usr/bin/env python
2 # coding: utf-8
3
4 #import package
5 #Note: You cannot import any other package!
6 import numpy as np
7 import matplotlib.pyplot as plt
8 import csv
9 import math
10 import random
11
12
13
14 # Global attributes
15 StudentID = '' # TODO 1 : Fill your student ID here
16 input_dataroot = 'input.csv' # Please name your input csv file as 'input.csv'
17 output_dataroot = StudentID + '_basic_prediction.csv' # Output file will be named as '[StudentID]_basic_prediction.csv'
18
19 input_datalist = [] # Initial datalist, saved as numpy array
20 output_datalist = [] # Your Prediction, should be 20 * 2 matrix and saved as numpy array
21 # The format of each row should be [Date, TSMC_Price_Prediction]
22 # e.g. ['2021/10/15', 512]
23
24 # You can add your own global attributes here
25
26
27
28 # Read input csv to datalist
29 with open(input_dataroot, newline='') as csvfile:
30     input_datalist = np.array(list(csv.reader(csvfile)))
31
32 # From TODO 2 to TODO 6, you can declare your own input parameters, local attributes and return parameters
33
34 def SplitData():
35     # TODO 2: Split data, 2021/10/15 ~ 2021/11/11 for testing data, and the other for training data and validation data
36
37
38 def PreprocessData():
39     # TODO 3: Preprocess your data e.g. split datalist to x_datalist and y_datalist
40
41
```

# Input File Format

- Basic part must follow this format (There is no restriction for the format of bonus part)
- Named as “input.csv” and contains a  $n * 3$  matrix,  $n$  means number of dates
- Each row represents “Date, MTK Price, TSMC Price”
- The part to be predicted (TSMC Price of 10/15 ~ 11/11 ) is filled with 0
- You can refer to the given file “sample\_input.csv”
- We will use this format of csv file to test your model with  $n = 209$  (1/4 ~ 11/11)
- Please make sure your model can be correctly input into this format of csv file

2021/1/4	861	624
2021/1/5	591	840
2021/1/6	677	748
2021/1/7	542	866
2021/1/8	512	766
2021/1/11	915	852
2021/1/12	589	997
2021/1/13	552	508
2021/1/14	947	743
2021/1/15	707	784
2021/1/18	680	539
2021/1/19	679	700
2021/1/20	669	899
2021/1/21	511	753
2021/1/22	747	840
2021/1/25	820	746
2021/1/26	823	639
2021/1/27	776	795
2021/1/28	589	705
2021/1/29	724	781
2021/2/1	927	781
2021/2/2	967	780
2021/2/3	735	545
2021/2/4	941	505
2021/2/5	838	592
2021/2/17	554	697
2021/2/18	638	512
2021/2/19	865	527
2021/2/22	705	954
2021/2/23	870	682
2021/2/24	614	959
2021/2/25	696	847
2021/2/26	670	916

2021/9/27	504	863
2021/9/28	704	713
2021/9/29	833	995
2021/9/30	929	604
2021/10/1	952	793
2021/10/4	883	809
2021/10/5	695	723
2021/10/6	703	783
2021/10/7	691	657
2021/10/8	858	652
2021/10/12	928	876
2021/10/13	536	843
2021/10/14	922	836
2021/10/15	519	0
2021/10/18	565	0
2021/10/19	625	0
2021/10/20	975	0
2021/10/21	804	0
2021/10/22	579	0
2021/10/25	822	0
2021/10/26	734	0
2021/10/27	729	0
2021/10/28	730	0
2021/10/29	906	0
2021/11/1	926	0
2021/11/2	957	0
2021/11/3	880	0
2021/11/4	712	0
2021/11/5	684	0
2021/11/8	938	0
2021/11/9	789	0
2021/11/10	929	0
2021/11/11	766	0

# Output File Format

- Both basic and bonus parts must follow this format
- The prediction you turned in must also follow this format
- Named as “[StudentID]\_basic(bonus)\_prediction.csv” and contains a 20 \* 2 matrix
- Each row represents “Date, TSMC Price(Prediction)”
- You can refer to the given file “sample\_output.csv”
- Please make sure your model can correctly output this format of csv file

2021/10/15	564
2021/10/18	423
2021/10/19	526
2021/10/20	628
2021/10/21	730
2021/10/22	799
2021/10/23	710
2021/10/26	654
2021/10/27	622
2021/10/28	543
2021/10/29	510
2021/11/1	499
2021/11/2	526
2021/11/3	589
2021/11/4	614
2021/11/5	688
2021/11/8	624
2021/11/9	635
2021/11/10	611
2021/11/11	584



# Assignment 1 Requirement

- Do it individually! Not as a team! (team is for final project)
- Announce date: 2021/9/30
- Deadline: 2021/10/14 23:59 (Late submission is not allowed!)
- Hand in your files in the following format
  - [StudentID]\_basic.py
  - [StudentID]\_basic\_prediction.csv
  - [StudentID]\_bonus.py (optional)
  - [StudentID]\_bonus\_prediction.csv (optional)
  - [StudentID]\_report.pdf
  - Compress all files into [StudentID]\_HW1.zip

# *The Evaluation Metric*

- MAPE (Mean absolute percentage error):

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

- For example:
- The value you predicted:
- $\hat{y} = [592, 486, 538, 689, 752, 841, 491]$
- Ground Truth :
- $y = [491, 584, 541, 599, 615, 741, 512]$
- $\text{MAPE} = 1/7 * 0.928 = 0.1326 = 13.26\%$ .



# Penalty

- 0 points if any of the following conditions
  - Plagiarism
  - Late submission
  - Not using template or import any other packages in basic part
  - Incorrect input/output format
  - Incorrect submission format

# Questions?

- TA: Meng-Chien Lin (charlie2039667@gapp.nthu.edu.tw)

