In class quiz 2

The stock price of Apple is listed as follows:

一張含有 桌 的圖片

自動產生的描述

We import it as **apple** as a dataframe.

1. If we want to extract the data from “2022-02-04 09:30:10”, which of the following codes is correct:
2. df.iloc["2022-02-04 09:30:10"]
3. df.iloc[3]
4. df.loc[3]
5. df.loc[2]
6. df.iloc[2]
7. With code:

**apple.shape**

we find that there are 14040 rows and 7 columns in our data, we then slice the data into two sets: train set and test set. If we want the ratio to be train 9: test 1, which of the following codes is incorrect:

A. train\_size= 14040\*0.9

test\_size=14040-train\_size

B. train\_size= 14040\*0.1

test\_size=14040-train\_size

C. train\_size= int(14040\*0.9)

test\_size=int(14040\*0.1)

D. train\_size= int(14040\*0.9)

test\_size=14040-train\_size

1. If we create a new column “dummy” with code:

**apple["dummy"]=[1 if x <300 else 2 if x<500 else 3 if x <700 else 4 if x<900 else 5 for x in apple["volume"]]**

what values of dummy would be for “2022-02-04 09:30:05”, “2022-02-04 09:30:10”, “2022-02-04 09:30:15”?

1. 2 4 3
2. 3 3 4
3. 4 3 3
4. 4 4 3
5. According to the lecture note, the fast signal is the rolling(30) average of stock price (close price in every 5 seconds).

We create a new column as RollSum30 with code:

**apple["RollSum30"]=apple["close"].rolling(30).sum()**

then we drop the NaN values with code:

**apple.dropna()**

how many rows do we drop?

1. 28
2. 29
3. 30
4. 31