دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)

نام و نام خانوادگی: پیمان هاشمی

شماره دانشجویی: 400131032

درس: مبانی یادگیری آماری

پروژه نهایی

توضیحات مقدماتی

در این گزارش، کد پروژه در یک فایل در فولدر کد قرار داده شده است. کد های مربوط به هر بخش قسمت بندی شده و در حد نیاز کامنت گذاری انجام شده است.

در ابتدا داده های مربوط به شرکت را وارد کرده و سپس داده های مربوط به داده را نیز وارد می‌کنیم.

خاک (iran.china.clay)، فولاد (s*mobarakeh.steel)، شستا (social.dec.inv)، خودرو (iran.khodro)، صندوق اطلس (toseAtlasMofid.ETF)

هر کدام از داده ها شامل تعدادی ستون و سطر می‌شوند.

نمونه از داده های مربوط به صندوق اطلس مفید در زیر آورده شده است:

	<TICKER>	<DTYYYYMMDD>	<FIRST>	<HIGH>	<LOW>	<CLOSE>	<VALUE>	<VOL>	<OPENINT>	<PER>	<OPEN>	<LAST>
0	ToseAtlasMofid.ETF	20230201	308066.0	308800.0	302005.0	303822.0	72035700602	237098	480	D	305377.0	304001.0
1	ToseAtlasMofid.ETF	20230131	303002.0	308372.0	301505.0	305377.0	154805740401	506933	821	D	298771.0	308359.0
2	ToseAtlasMofid.ETF	20230130	298200.0	303998.0	294501.0	298771.0	185411002893	620579	1018	D	297877.0	303998.0
3	ToseAtlasMofid.ETF	20230129	307990.0	307990.0	293002.0	297877.0	314370975426	1055372	1078	D	307081.0	298102.0
4	ToseAtlasMofid.ETF	20230128	314950.0	314950.0	305002.0	307081.0	221471022111	721213	818	D	309176.0	307988.0
...	...	...	...	...	...	...	...	...	...	...	...	...
2082	ToseAtlasMofid.ETF	20131222	10000.0	10000.0	10000.0	10000.0	201668040000	20166804	2739	D	10000.0	10000.0
2083	ToseAtlasMofid.ETF	20131221	10000.0	10000.0	10000.0	10000.0	59571970000	5957197	1162	D	10000.0	10000.0
2084	ToseAtlasMofid.ETF	20131218	10000.0	10000.0	10000.0	10000.0	46226010000	4622601	1131	D	10000.0	10000.0
2085	ToseAtlasMofid.ETF	20131217	10000.0	10000.0	10000.0	10000.0	64871920000	6487192	1585	D	10000.0	10000.0
2086	ToseAtlasMofid.ETF	20131216	10000.0	10000.0	10000.0	10000.0	167801300000	16780130	3508	D	10000.0	10000.0

شکل 1. مجموعه داده صندوق اطلس مفید

پس از وارد کردن داده شاخص کل، تاریخ های آن بر خلاف 5 دیتاست دیگر به شمسی می‌باشد. پس با استفاده از کتابخانه persiantools داده های شمسی را به میلادی تبدیل میکنیم و داریم:

	<DTYYYYMMDD>	<VALUE>
0	20230201	1557244.00
1	20230131	1556551.90
2	20230130	1539679.61
3	20230129	1542190.69
4	20230128	1600083.66

شکل 2. مجموعه داده شاخص کل

سپس بنا به گفته سوال، داده های مربوط به دو سال گذشته را جدا میکنیم. که برای این کار از ستون تاریخ استفاده کرده و داده ها را بین بازه اخرین داده و (20000 - تاریخ) جدا می‌کنیم.

```
iran_china_clay shape :  (462, 12)
iran_khodro shape :  (449, 12)
s_mobarakeh_steel shape :  (454, 12)
social_sec_inv shape :  (410, 12)
tose_atlas_mofid shape :  (478, 12)
total_indices shape :  (480, 2)
```

شکل 3. مقدار هر مجموعه داده

عملیات پیش پردازش

در ابتدا برای اطمینان از عدم وجود داده های Nan، از عملیات dropna استفاده می‌کنیم.

```
<TICKER>        0
<DTYYYYMMDD>    0
<FIRST>         0
<HIGH>          0
<LOW>           0
<CLOSE>         0
<VALUE>         0
<VOL>           0
<OPENINT>       0
<PER>           0
<OPEN>          0
<LAST>          0
dtype: int64
```

شکل 4. تعداد داده های Nan در یکی از مجموعه داده ها

شرکت های که در بازار بورس فعالیت دارند، ممکن است در طول سال در بسیاری از روز ها به دلیل برگزاری مجمع یا افزایش سرمایه بسته باشند. در این روز ها سهم این شرکت ها معامله نمی‌شود و در حقیقت در جدول نیز، در این تاریخ ها اطلاعاتی ثبت نشده است. این در حالی است که هر روزی که بازار بورس فعالیت داشته باشد، شاخص کل نیز دچار تغییراتی می‌شود و طبیعی است که

تعداد داده های سهام شرکت ها در طول سال با تعداد داده های شاخص کل برابر نباشد. به همین برای انجام عملیات هایی در آینده و برای برابر سازی تعداد سهم ها یک عملات پیش پردازش را انجام می‌دهیم. در این عملیات در ابتدا تاریخ هایی که در مجموعه داده شاخص کل وجود دارد ولی در مجموعه داده سهم ها نیست را به آن ها اضافه می‌کنیم. در مرحله بعدی میدانیم که در روز هایی که سهم شرکت ها بسته است، سهم آن ها معامله نمی‌شود، پس مقدار Closed برابر آن روز ها برابر مقدار اخرین روز فعالیت آن ها است. پس مقدار Closed را در این روز ها برابر اخرین مقدار موجود در روز های قبل می‌کنیم و بقیه ستون ها نیز برابر صفر قرار می‌دهیم. به این صورت تعداد داده در همه مجموعه داده ها برابر یکدیگر می‌شود و میتوان به نتایج معتبری در ادامه دست یافت.

```
Iran.China.Clay shape : (479, 12)
Iran.Khodro shape : (479, 12)
S*Mobarakeh.Steel shape : (479, 12)
Social.Sec.Inv shape : (479, 12)
ToseAtlasMofid.ETF shape : (479, 12)
```

شکل 5. مقدار هر مجموعه داده بعد از پیش پردازش

سوال ۱)

الف)

در این قسمت با توجه به فرمول داده شده مقدار بازده را برای شاخص کل با توجه به ستون VALUE محاسبه می‌کنیم:

نکته ایی که باید لحاظ شود این است که چون مجموعه داده ما در ردیف اول اخرین تاریخ را نشان میدهد و در ردیف اخر اولین تاریخ را پس فرمول داده شده کمی به تغییر نیاز دارد و باید به روش زیر محاسبه شود:

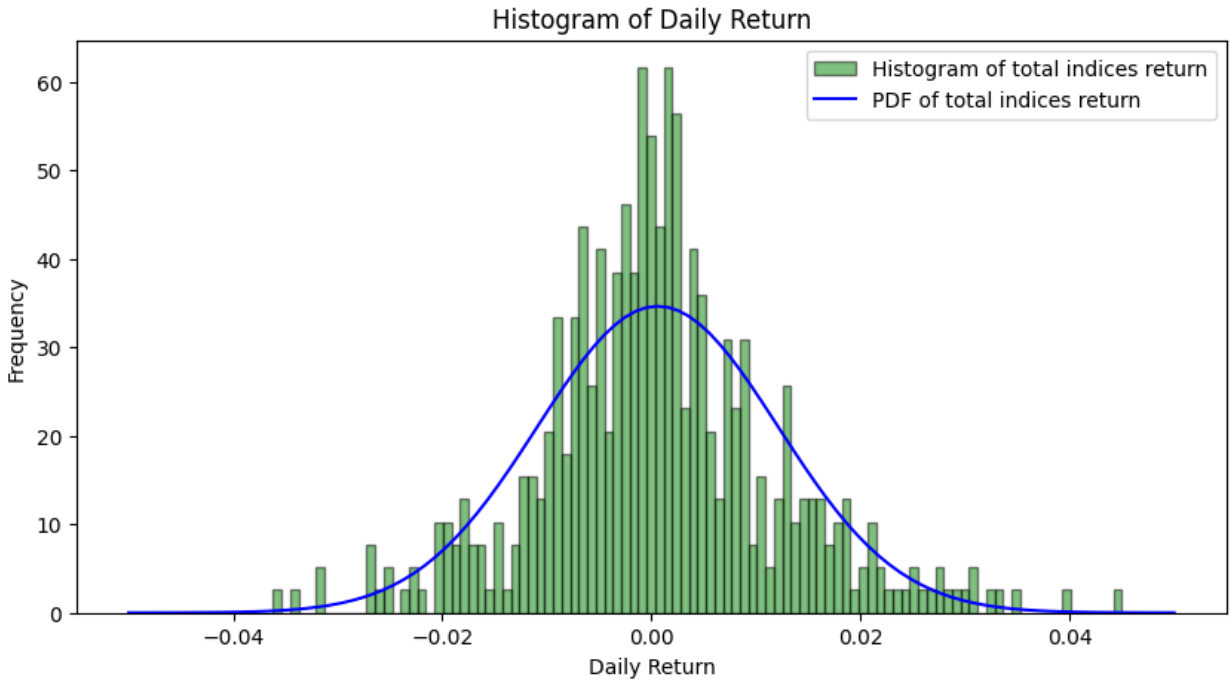
```
((closed_values[i] - closed_values[i+1])/closed_values[i+1])
```

که در این صورت اگر شاخص کل افزایش پیدا کند مقدار آن مثبت و اگر کاهش پیدا کند مقدار بازده منفی می‌شود.

	<DTYYYYMMDD>	<VALUE>	return
0	20230201	1557244.00	0.000445
1	20230131	1556551.90	0.010958
2	20230130	1539679.61	-0.001628
3	20230129	1542190.69	-0.036181
4	20230128	1600083.66	-0.007365
...	...	...	...
475	20210208	1192744.36	0.018976
476	20210207	1170532.85	0.030993
477	20210206	1135345.48	-0.032110
478	20210203	1173010.67	-0.026924
479	20210202	1205467.10	0.000000

شکل 6. مجموعه داده شاخص کل و مقدار بازده

توزیع احتمالاتی مقدار بازده شاخص کل به روش هیستوگرام و توزیع نرمال به صورت زیر است:



شکل 7. توزیع احتمالاتی مقادیر بازده شاخص کل

ب)

برای اثبات متقارن یا نامتقارن بودن یک توزیع، تست های مختلفی وجود دارد که یکی از آن ها skewness test می‌باشد.

و همچنین تست هایی مانند نرمال تست یا kstest نیز وجود دارد. همچنین میتوان از روی میانگین و انحراف معیار نیز تحلیل هایی انجام داد.

اگر نتیجه بدست آمده از skewness test نزدیک به صفر باشد و در بازی (0.05 و 0.05-) قرار گیرد میتوان توزیع را متقارن دانست و یا اگر میانگین صفر باشد با واریانس مثبت میتوان توزیع را متقارن دانست.

نتیج بدست آمده نشان میدهد که توزیع احتمالاتی متقارن نمی‌باشد.

```
Skewness test of total indices return distribution:  SkewtestResult(statistic=2.255719696414678, pvalue=0.024088191245966156)

Skewness of total indices return distribution:  0.25292167599694254

Kolmogorov-Smirnov test of total indices return distribution:  KstestResult(statistic=0.48556893244935373, pvalue=9.353798977288398e-105)

DAgostinos K^2 test of total indices return distribution:  NormaltestResult(statistic=18.34810107812115, pvalue=0.00010369563470227627)

mean of total indices return distribution:  0.0005996449201354784
standard deviation of total indices return distribution:  0.01151990441035642

Result:
The distribution of total indices return is asymmetrical
```

شکل 8. نتایج بدست آمده برای متقارن یا نامتقارن بودن توزیع احتمالاتی مقدار بازده شاخص کل

ج)

نتایج بدست آمده برای میانگین، انحراف معیار، و واریانس هر یک از سهم ها برابر زیر است:

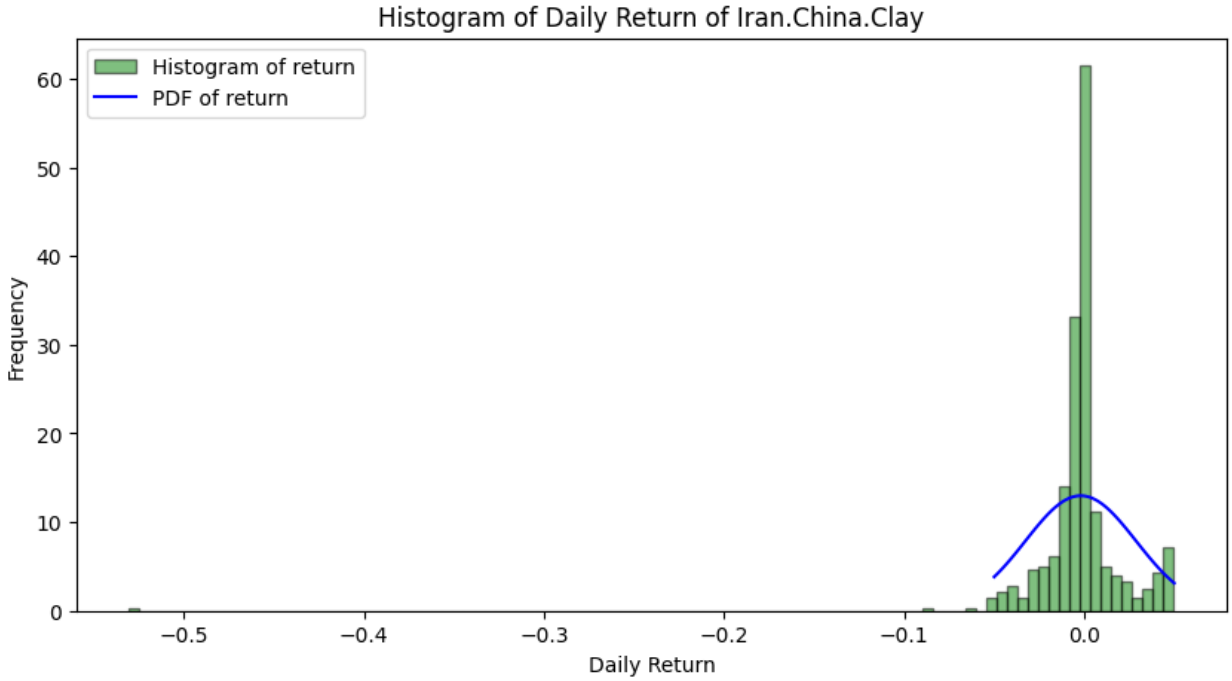
```
mean of return of  Iran.China.Clay  is:  -0.0019307187846566566
variance of return of  Iran.China.Clay  is:  0.0009453416169190381
std of return of  Iran.China.Clay  is:  0.030746408195414274
--------------------------------
mean of return of  Iran.Khodro  is:  0.0005386224251596697
variance of return of  Iran.Khodro  is:  0.0006655977602160631
std of return of  Iran.Khodro  is:  0.02579918138654913
--------------------------------
mean of return of  S*Mobarakeh.Steel  is:  -0.0005736916280916049
variance of return of  S*Mobarakeh.Steel  is:  0.0011333485089722584
std of return of  S*Mobarakeh.Steel  is:  0.03366524185227634
--------------------------------
mean of return of  Social.Sec.Inv  is:  -0.0029195098141805227
variance of return of  Social.Sec.Inv  is:  0.0025026653245125503
std of return of  Social.Sec.Inv  is:  0.05002664614495509
--------------------------------
mean of return of  ToseAtlasMofid.ETF  is:  0.0009880772053727838
variance of return of  ToseAtlasMofid.ETF  is:  0.00021951300862082138
std of return of  ToseAtlasMofid.ETF  is:  0.014815971403212865
```

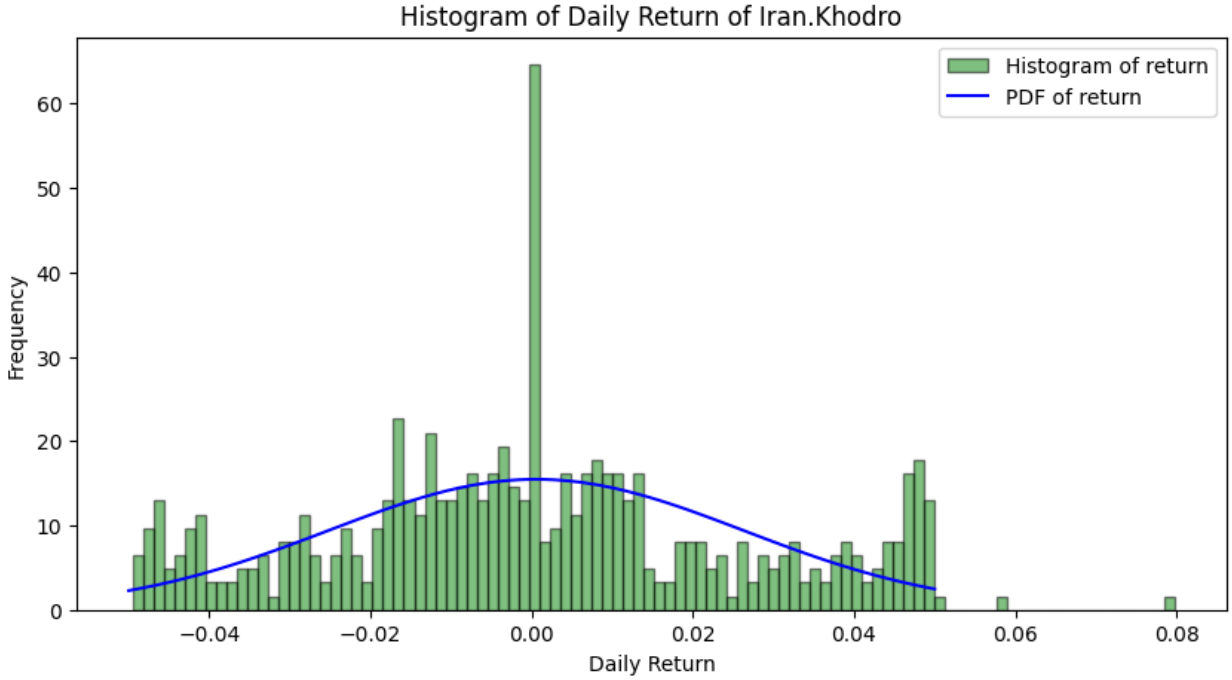
شکل 9. مانگین، انحراف معیار، و واریانس هر یک از سهم ها

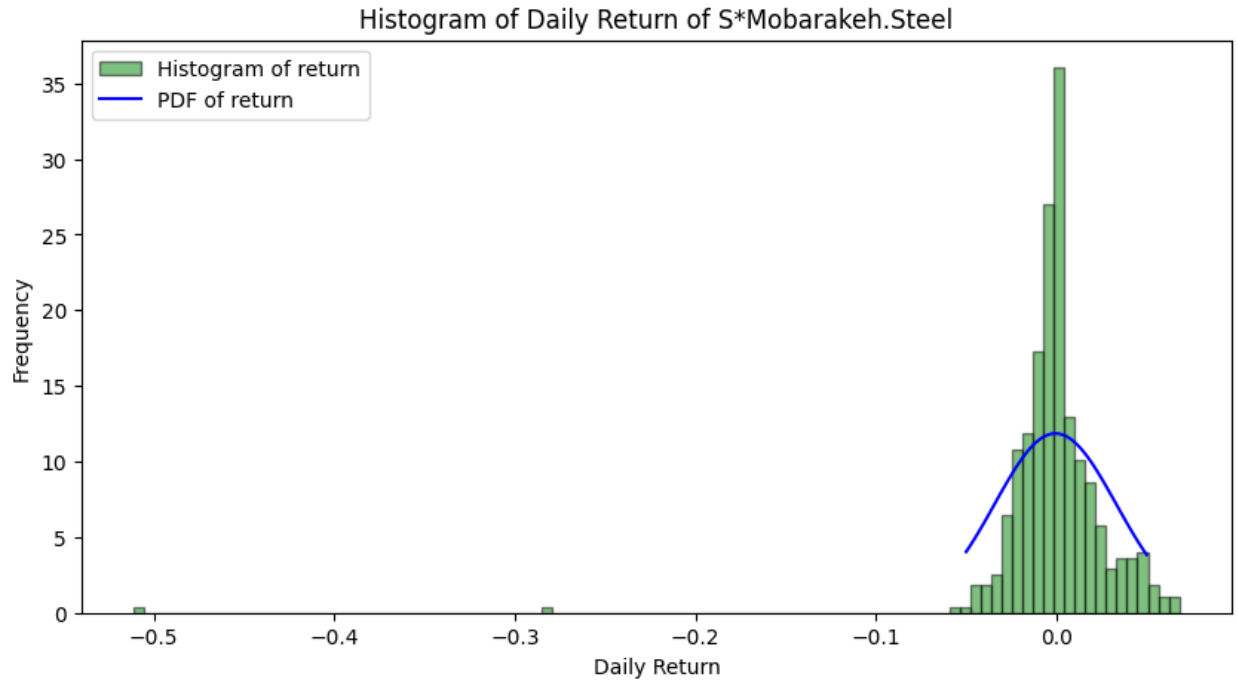
نتایج بدست آمده نشان می‌دهد میانگین سهم های کخاک، فولاد و شستا منفی بوده است. یعنی در طول 2 سال اخیر به طور کلی زیان ده بوده اند در حالی که صندوق اطلس و خودرو یا میانگین مثبت سود آور بوده اند.

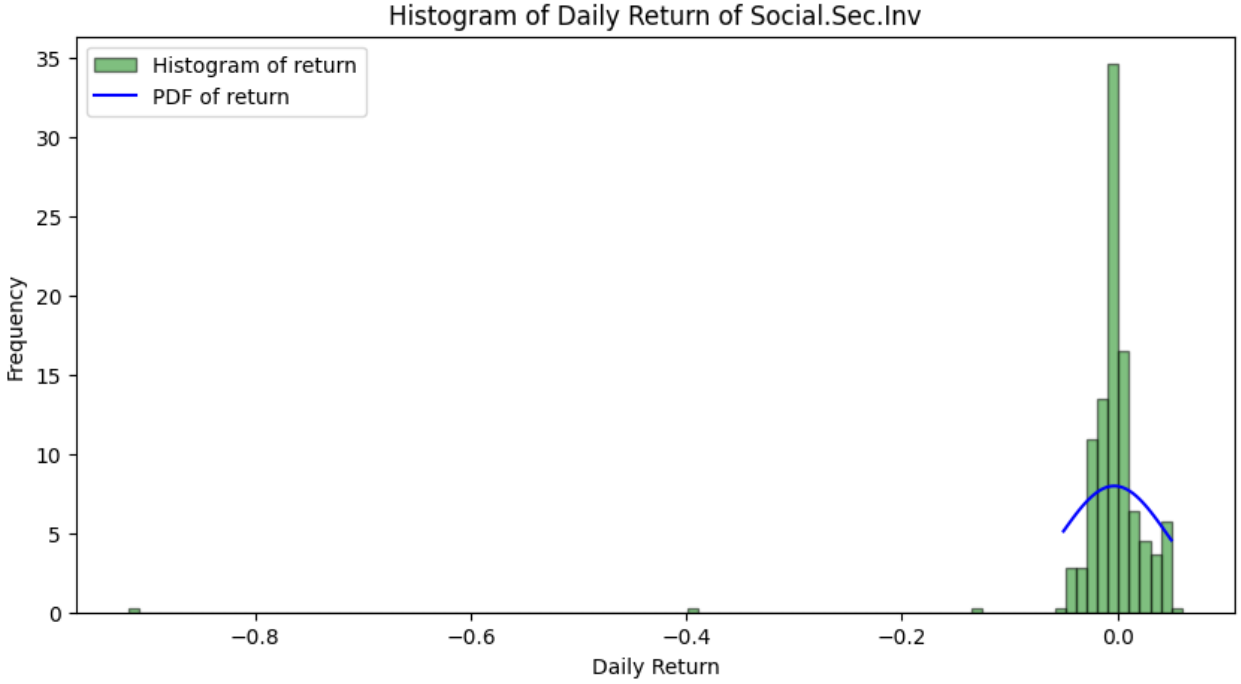
همچنین از روی واریانس میتوان تغییرات قیمتی سهم را در طول سال فهمید که به ترتیب سهم های شستا، فولاد، کخاک،خودرو و صندوق اطلس، تغییرات قیمتی از زیاد به کم داشته اند. که سهم شستا با فاصله از دیگر سهم ها دچار تغییرات قیمتی و نوسانی بودن قیمت شده است و سهم صندوق اطلس با فاصله از دیگر سهم ها دارای ثبات قیمتی بالاتری بوده است و بازه نوسان کمتری داشته است.

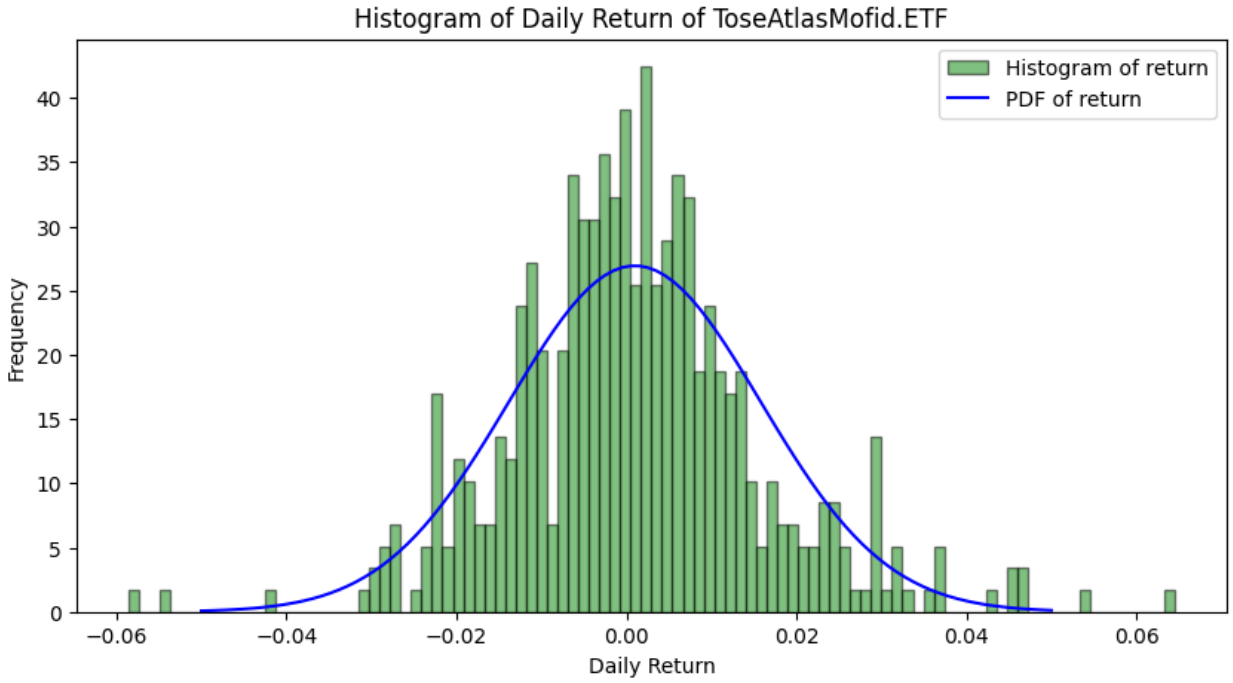
همچنین توزیع احتمالاتی هر یک سهم ها به شکل زیر است:


Histogram of Daily Return of Iran.China.Clay
Histogram of return
PDF of return
Frequency
Daily Return
0
10
20
30
40
50
60
−0.5
−0.4
−0.3
−0.2
−0.1
0.0


Histogram of Daily Return of Iran.Khodro
Histogram of return
PDF of return
Frequency
Daily Return
0
10
20
30
40
50
60
−0.04
−0.02
0.00
0.02
0.04
0.06
0.08


Histogram of Daily Return of S*Mobarakeh.Steel
Histogram of return
PDF of return
Frequency
Daily Return
0
5
10
15
20
25
30
35
−0.5
−0.4
−0.3
−0.2
−0.1
0.0


Histogram of Daily Return of Social.Sec.Inv
Histogram of return
PDF of return
Frequency
Daily Return
0
5
10
15
20
25
30
35
−0.8
−0.6
−0.4
−0.2
0.0



شکل 10. توزیع احتمالاتی سهم ها

همانطور که مشخص است، توزیع های مربوط به صندوق اطلس و خودرو به شاخص کل نزدیک تر می‌باشند.

د)

برای این بخش نیاز است تا تغییرات میانگین را در بازه های مختلف محاسبه کنیم برای همین از roll() استفاده می‌کنیم تا بازه ها مشخص شده را محاسبه کنیم و سپس میانگین را محاسبه می‌کنیم و در نهایت diff() می‌گیریم. این بازه به این صورت که اگر هر 10 مد نظر باشد داریم:

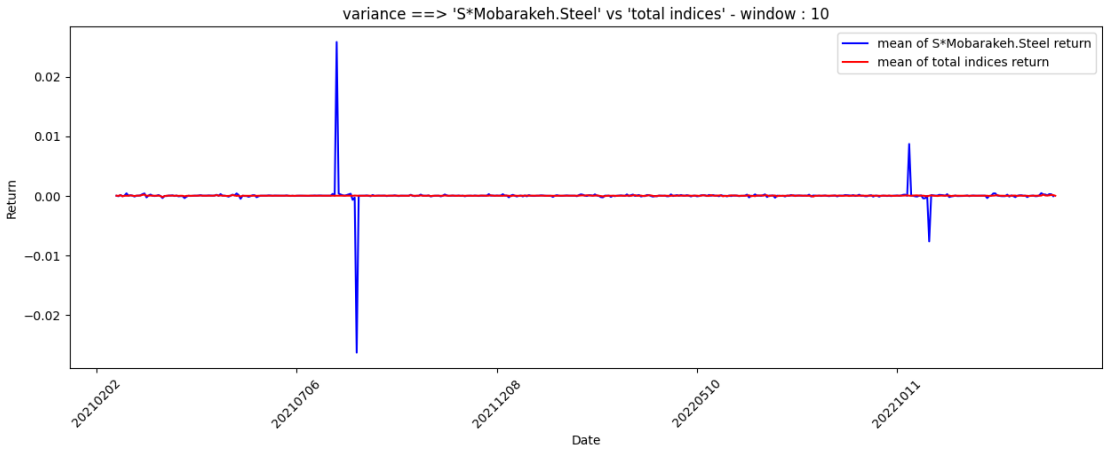
تغییرات میانگین 1-10 روز، تغییرات میانگین 2-12 روز، تغییرات میانگین 3-13 روز، و....

سپس تغییرات میانگین و واریانس هر یک سهم ها را نسبت به شاخص کل به صورت جداگانه به تصویر می‌کشیم:

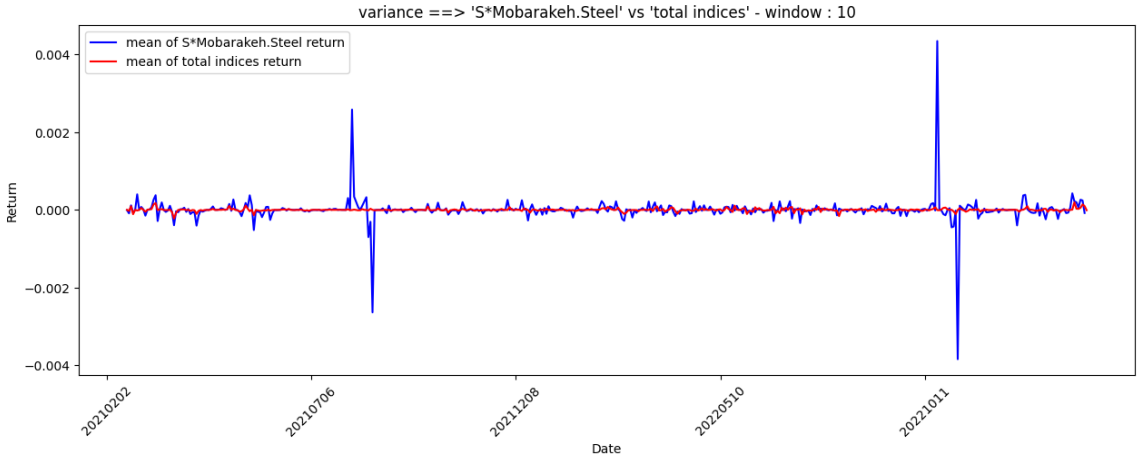
(در تیتر هر جدول میانگین یا واریانس بودن آن مشخص شده است)

(میتوان تاریخ را به 5 بخش تقسیم کرد که بخش اول از تاریخ 20210202 تا 20210706 است و بخش 5 از تاریخ 20221011 تا اخر است)

(در محاسبه واریانس و میانگین، تغییراتی که خیلی بیشتر از حد نرمال بوده اند و باعث میشد تا نتوان تحلیل درستی ارائه داد را نرمال کرده و بازه آن ها کمتر کردیم. این تغییرات تنها برای مواردی که مقدار بالایی داشتند اعمال شده است. یک نمونه از آن در زیر اورده شده است که مربوط تغییرات واریانس سهم فولاد قبل و بعد از اعمال نرمال سازی می‌باشد)

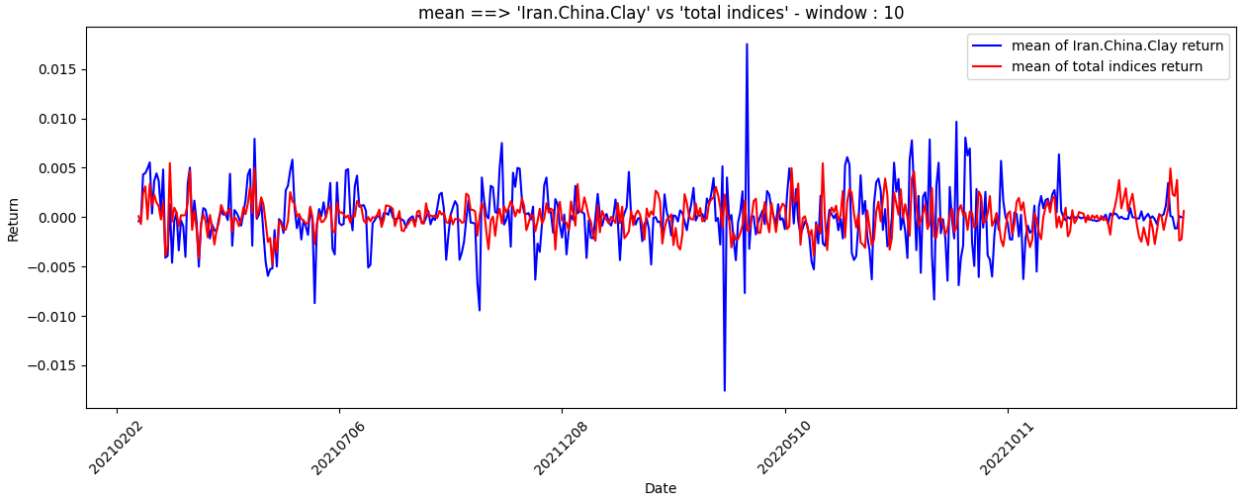


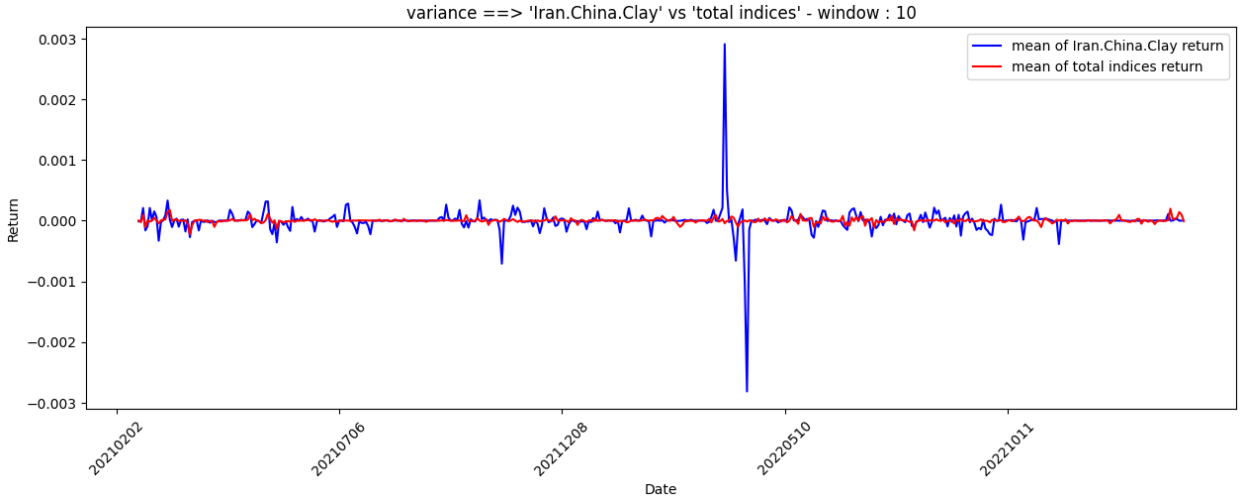
قبل از اعمال نرمال سازی

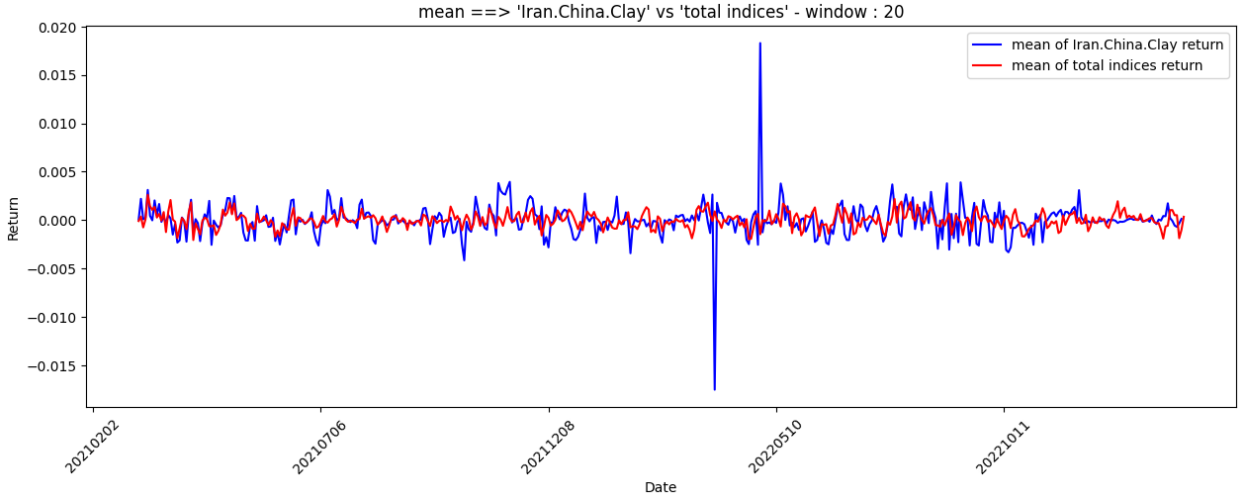


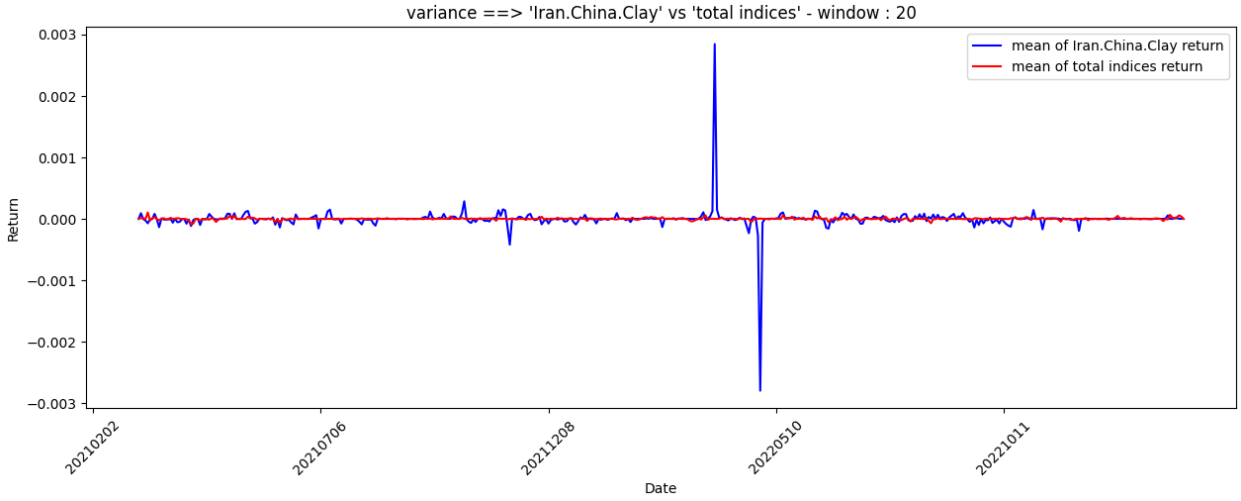
بعد از اعمال نرمال سازی

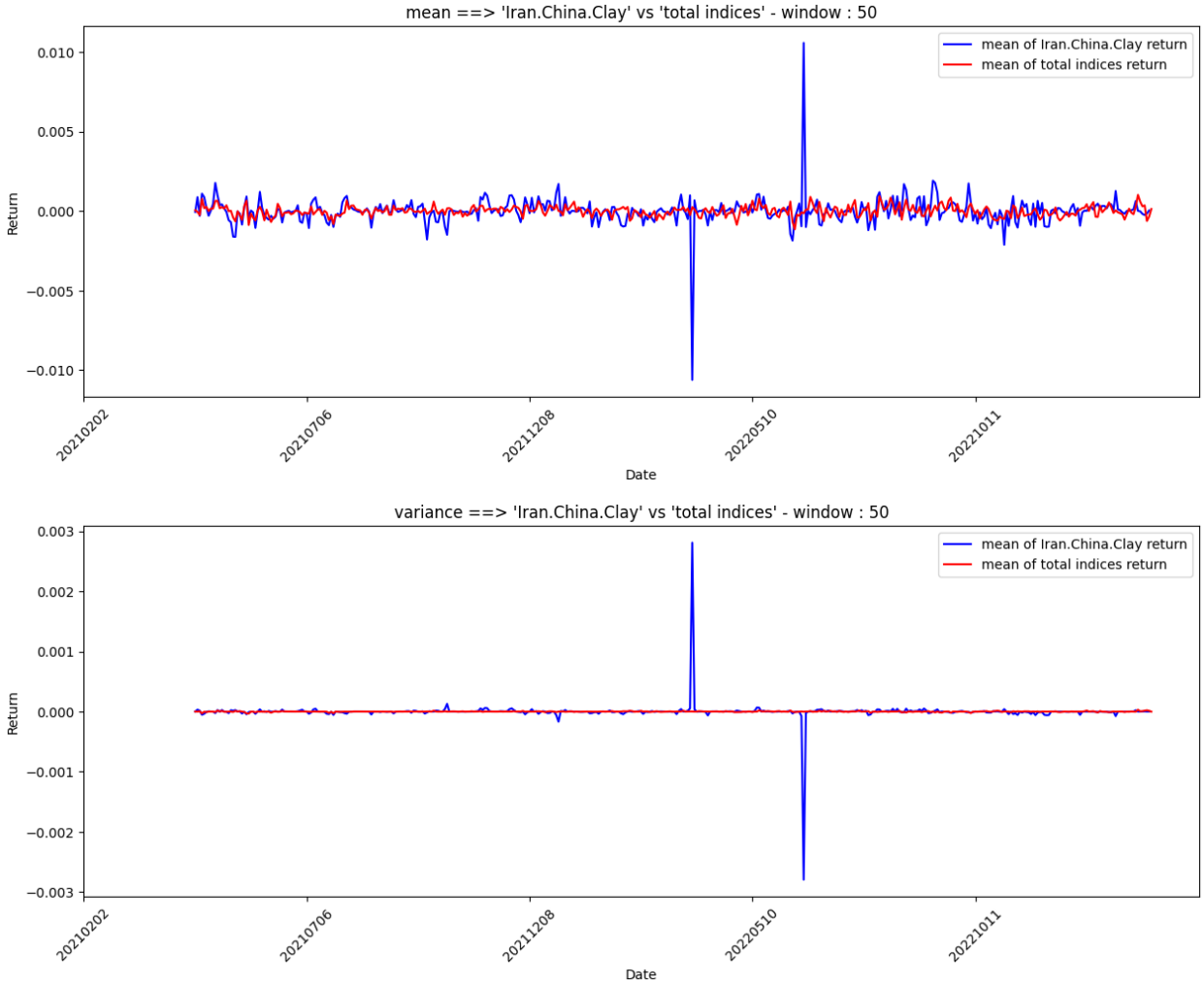
میتوان دید که تغییرات با جزئیات بهتری قابل مشاهده است.




variance ==> 'Iran.China.Clay' vs 'total indices' - window : 10
mean of Iran.China.Clay return
mean of total indices return
Return
Date
0.003
0.002
0.001
0.000
−0.001
−0.002
−0.003
20210202
20210706
20211208
20220510
20221011


mean ==> 'Iran.China.Clay' vs 'total indices' - window : 20
mean of Iran.China.Clay return
mean of total indices return
Return
Date
0.020
0.015
0.010
0.005
0.000
−0.005
−0.010
−0.015
20210202
20210706
20211208
20220510
20221011


variance ==> 'Iran.China.Clay' vs 'total indices' - window : 20
mean of Iran.China.Clay return
mean of total indices return
Return
Date
0.003
0.002
0.001
0.000
−0.001
−0.002
−0.003
20210202
20210706
20211208
20220510
20221011

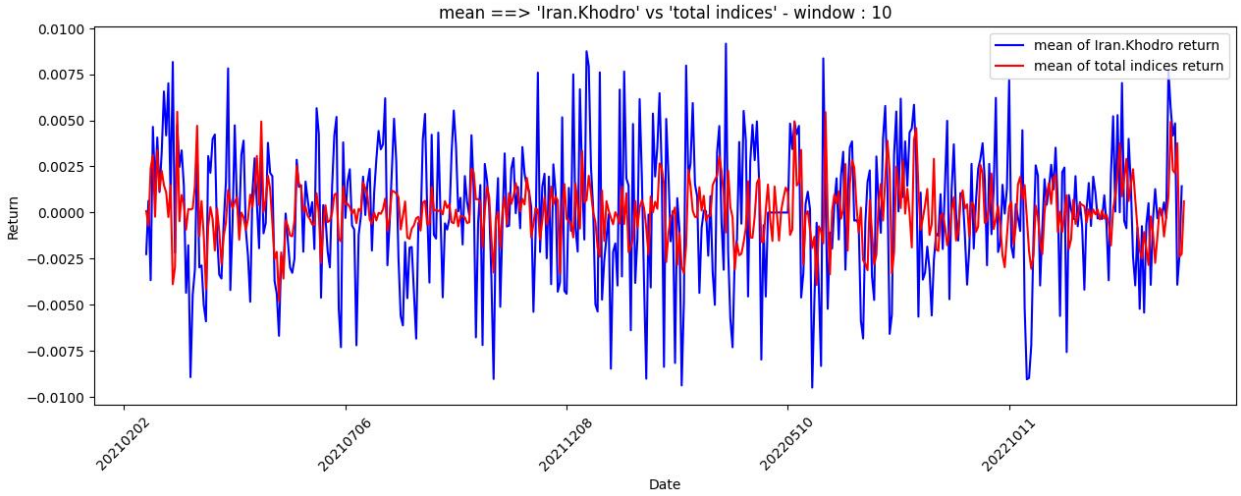


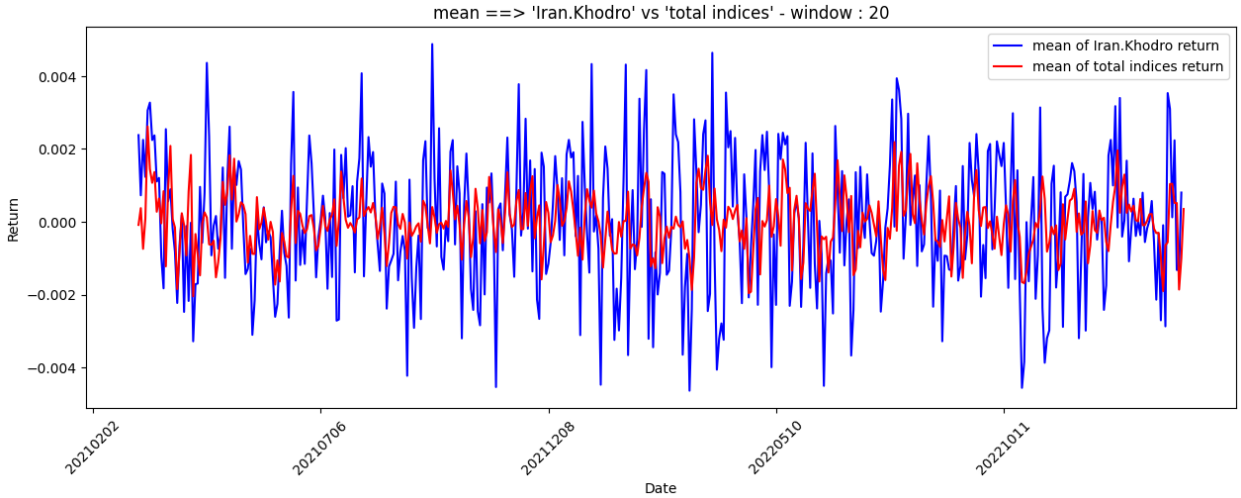
شکل 11. میانگین و تغییرات کخاک با شاخص کل در بازه های 10، 20، 50 روزه

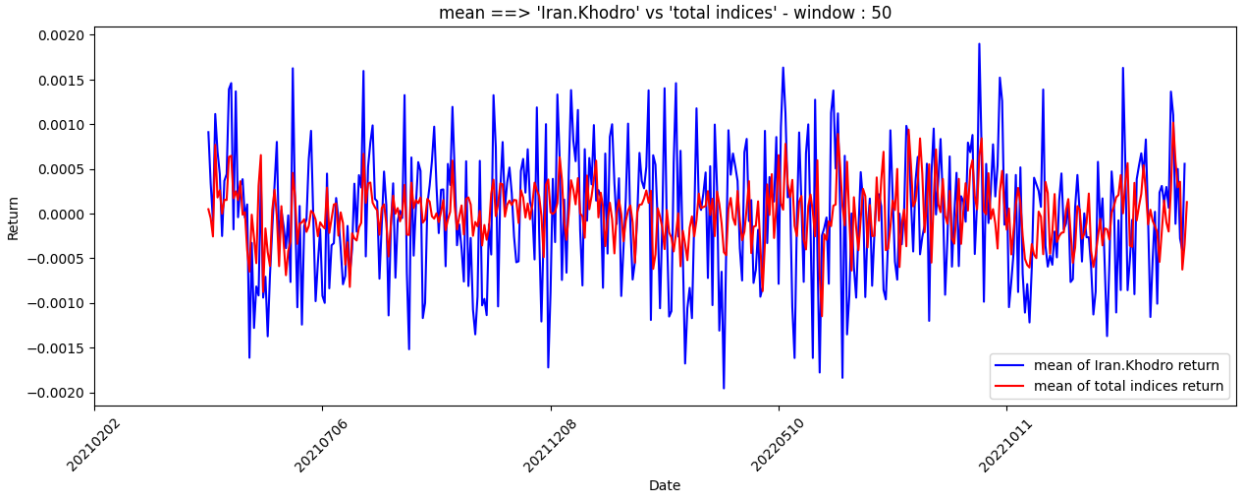
می‌توان دید که در بازه 10 روزه به طور کل با شباهتی نزدیک به شاخص کل تغییر میکند. برای میانگین در بخش 4 بسیار نوسانی است و مطابق با شاخص کل نیست در حالی که در بخش 5 شاخص کل نوسان بیشتری دارد. تغییرات واریانس کخاک بسیار بیشتر از شاخص بوده است و شباهت زیادی با آن ندارد.

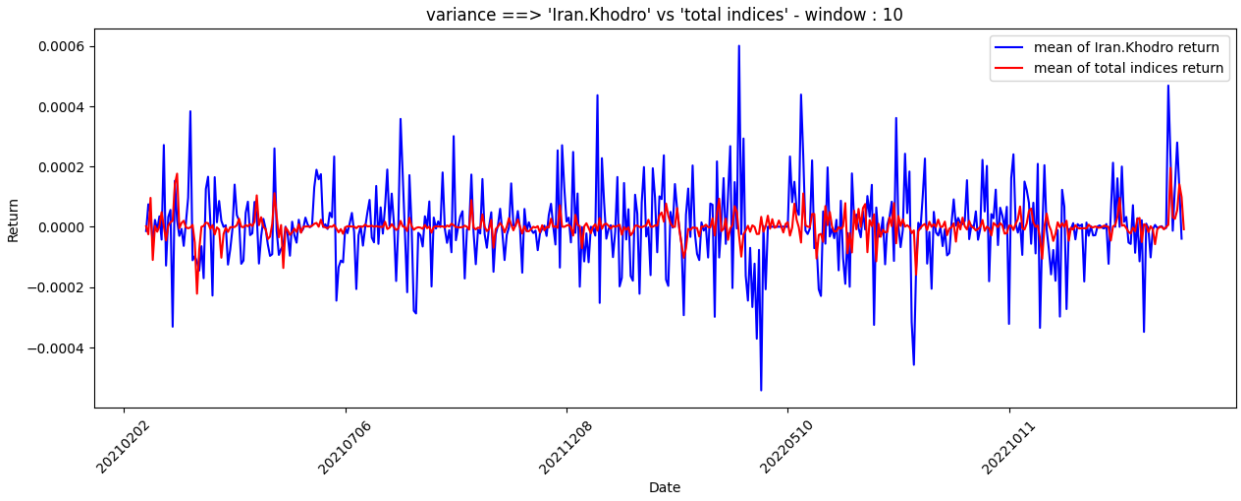
در بازه 20 روزه شباهت تغییرات کخاک با شاخص کل بیشتر شده و نوسان های موجود با نوسان های شاخص کل تغییر میکند.

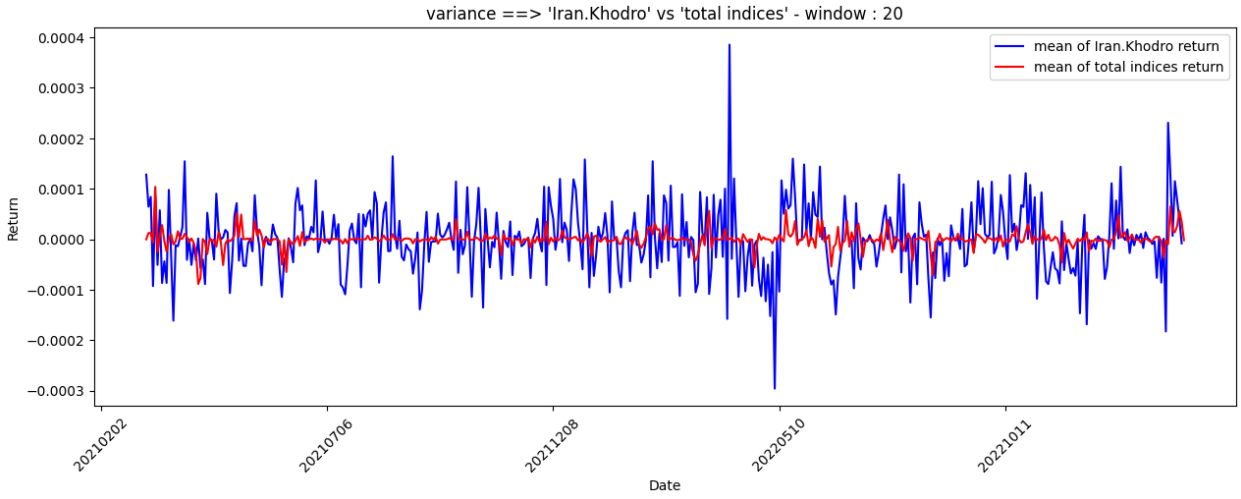
همچنین در بازه 50 روزه نیز این شباهت در بخش هایی زیاد تر و در بخش های مانند بخش 5 ام کمتر می‌شود. به طور کلی می‌توان دید که این تغییرات میانگین این سهم بر روی شاخص تاثیر گذار بوده است ولی تغییرات واریانس آن شباهت چندانی به شاخص کل ندارد.

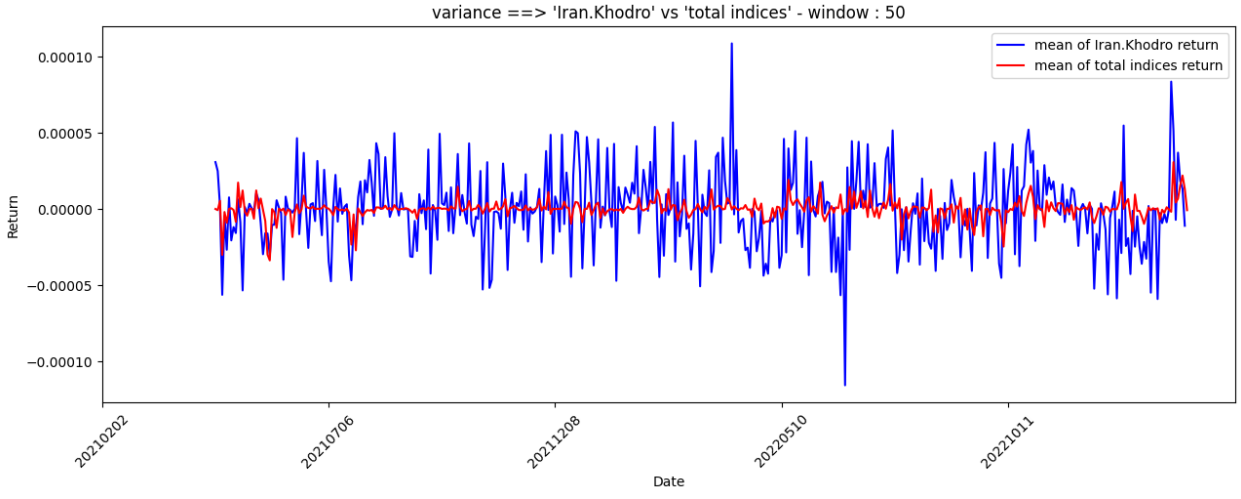

mean ==> 'Iran.Khodro' vs 'total indices' - window : 10
mean of Iran.Khodro return
mean of total indices return
Return
Date
0.0100
0.0075
0.0050
0.0025
0.0000
−0.0025
−0.0050
−0.0075
−0.0100
20210202
20210706
20211208
20220510
20221011


mean ==> 'Iran.Khodro' vs 'total indices' - window : 20
mean of Iran.Khodro return
mean of total indices return
Return
Date
0.004
0.002
0.000
−0.002
−0.004
20210202
20210706
20211208
20220510
20221011


mean ==> 'Iran.Khodro' vs 'total indices' - window : 50
mean of Iran.Khodro return
mean of total indices return
Return
Date
0.0020
0.0015
0.0010
0.0005
0.0000
−0.0005
−0.0010
−0.0015
−0.0020
20210202
20210706
20211208
20220510
20221011







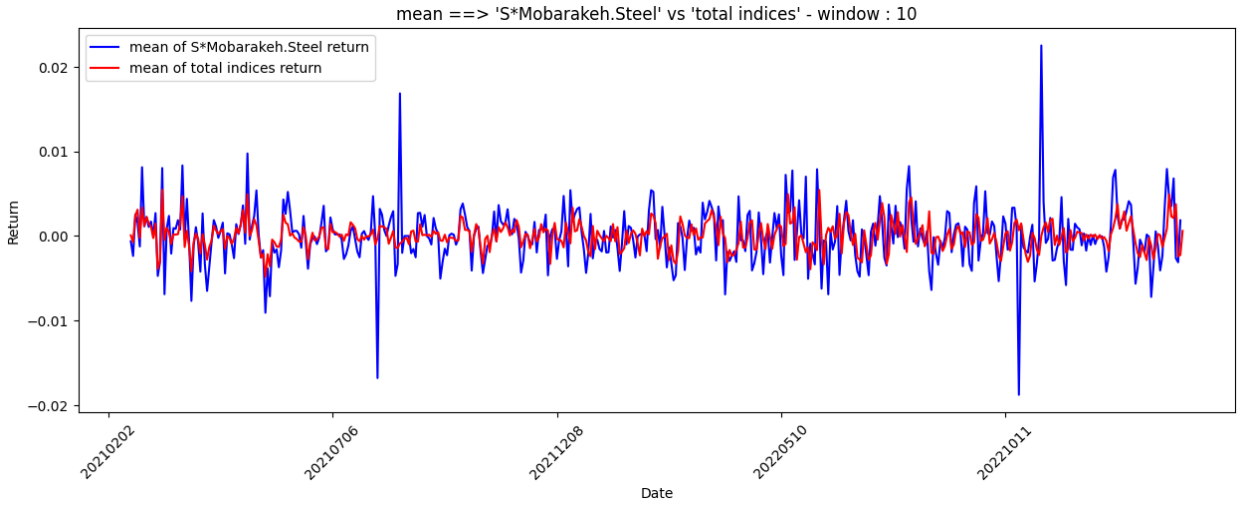
شکل 12. تغییرات میانگین و واریانس خودرو با شاخص کل در بازه های 10، 20، 50 روزه

میتوان دید که در بازه 10 روزه میانگین سهم خودرو نسبتا شباهت بالایی با شاخص کل دارد و فقط نوسان بیشتری داشته است و به نوعی تغییرات آن نسبتا بر روی شاخص کل اثر گذار بوده است. ولی واریانس آن خیلی با شاخص کل شباهت ندارد.

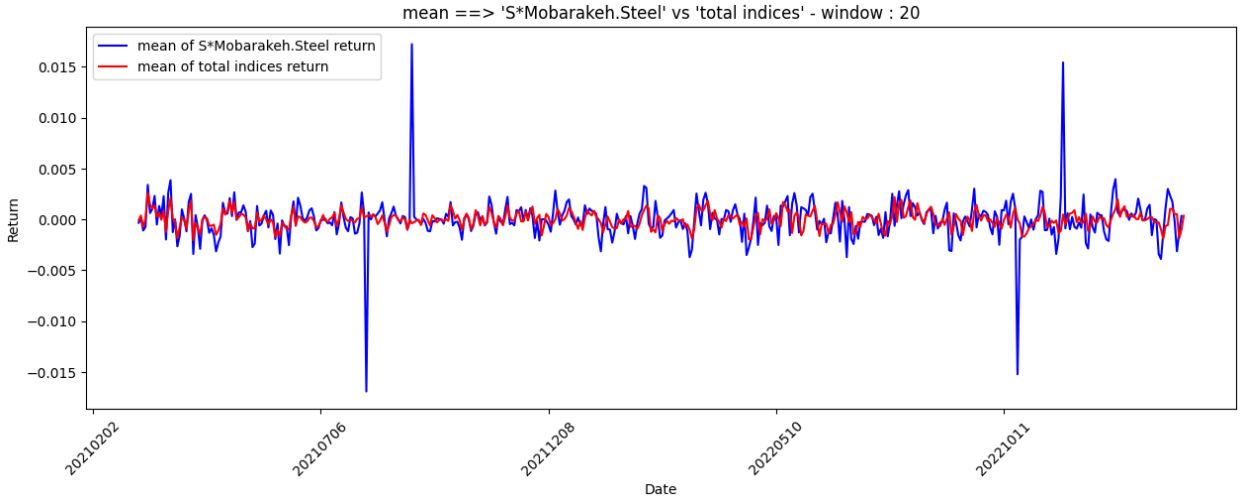
در بازه 20 روزه این تغییرات شباهت بیشتری به شاخص کل پیدا میکند و در بازه 50 روزه نیز بیشتر می‌شود. این بدین معنی است که سهم خودرو بسیار بر روی شاخص تاثیر گذار بوده است و یکی از سهم های اصلی در بورس است.

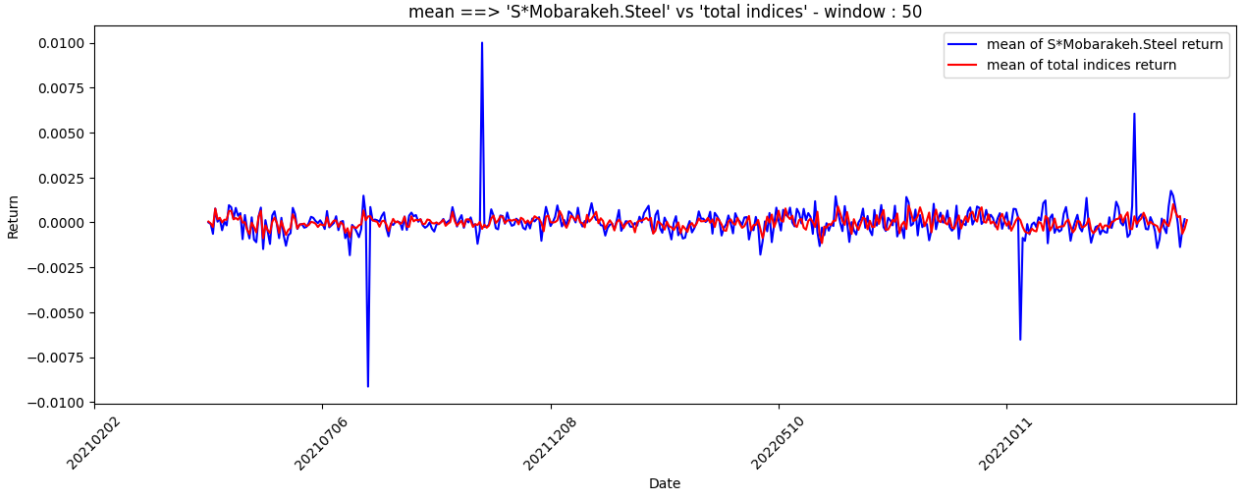
در بازه 10 روزه میتوان دید که واریانس بسیار نوسانی بوده است ولی در زمانی که تغییرات واریانس به شدت زیاد میشود به همان اندازه نیز کاهش پیدا میکند مانند انتهای بخش سوم یا انتهای بخش پنجم. این تغییرات با زیاد شدن بازه نیز به نوعی تاثیر گذار تر می‌شود. به این معنی که در بخش هایی که شباهت وجود دارد این شباهت بیشتر شده و در بخش هایی که تفاوت وجود دارد این تفاوت بیشتر می‌شود. مثلا در بازه 50 روزه در اوایل بخش سوم میتوان دید که تغییرات واریانس خودرو کاملا خلاف جهت شاخص بوده است. ولی در اواسط بخش اول این تغییرات شباهت بیشتری نسبت بهم دارند تا بازه 10 روزه.

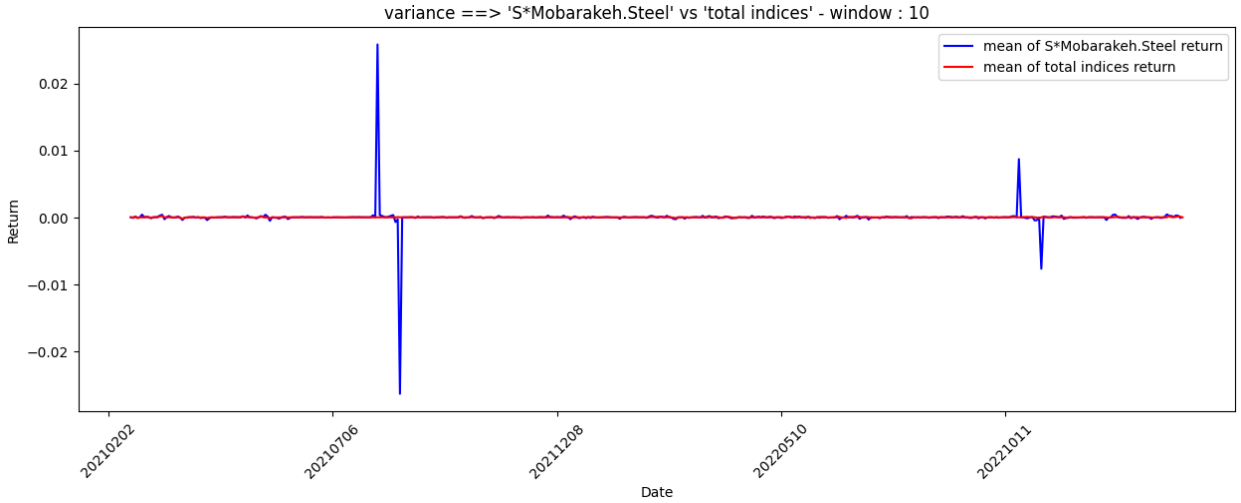
قبل از اعمال نرمال سازی



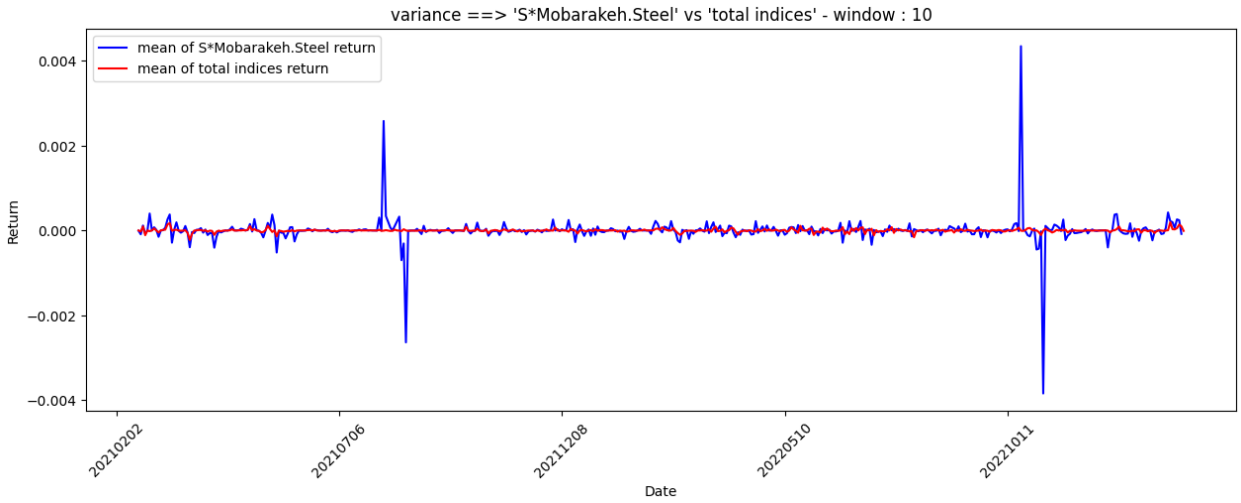
بعد از اعمال نرمال سازی



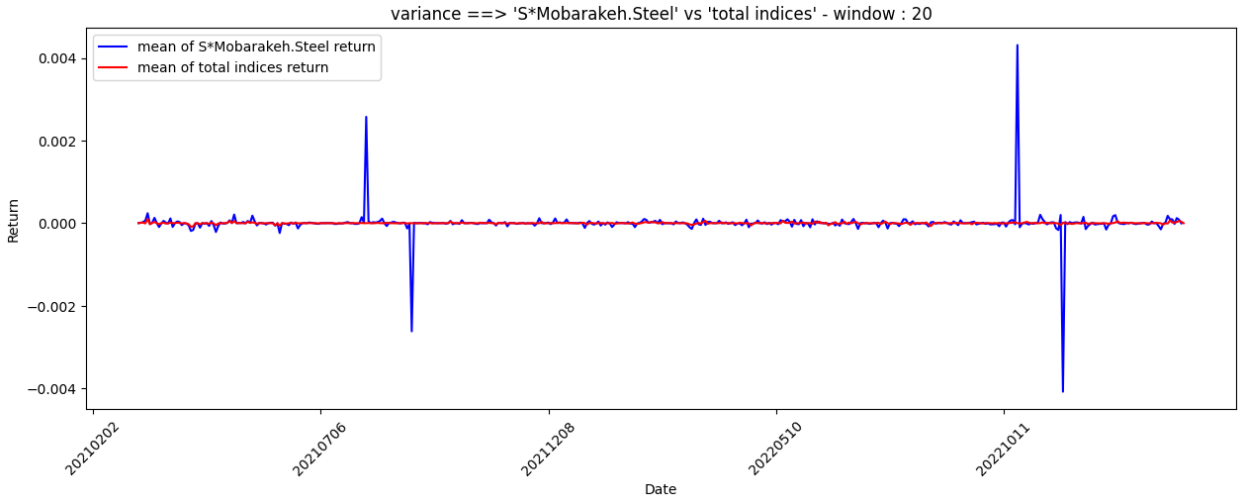


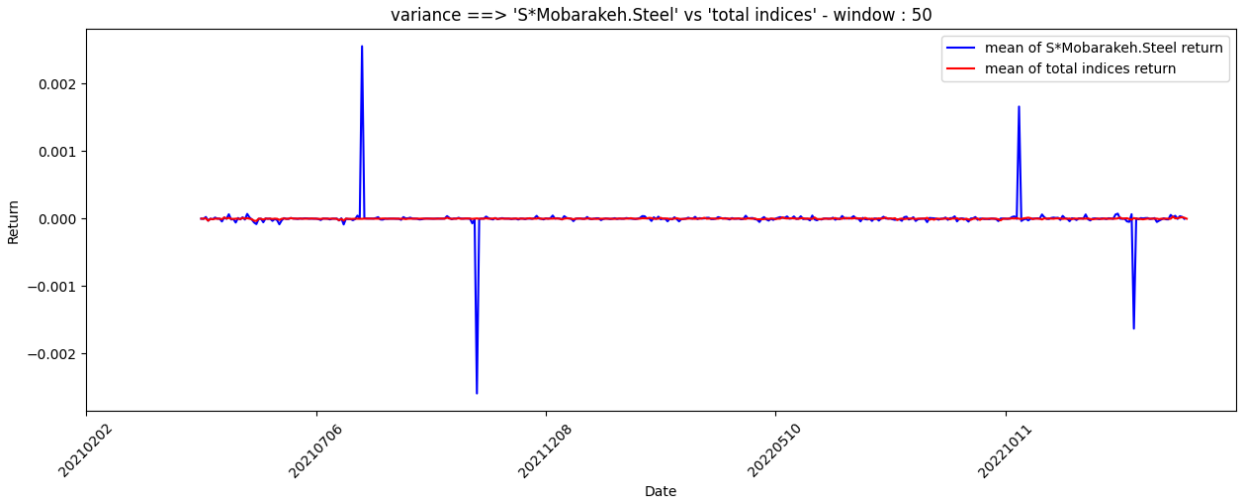


قبل از اعمال نرمال سازی



بعد از اعمال نرمال سازی

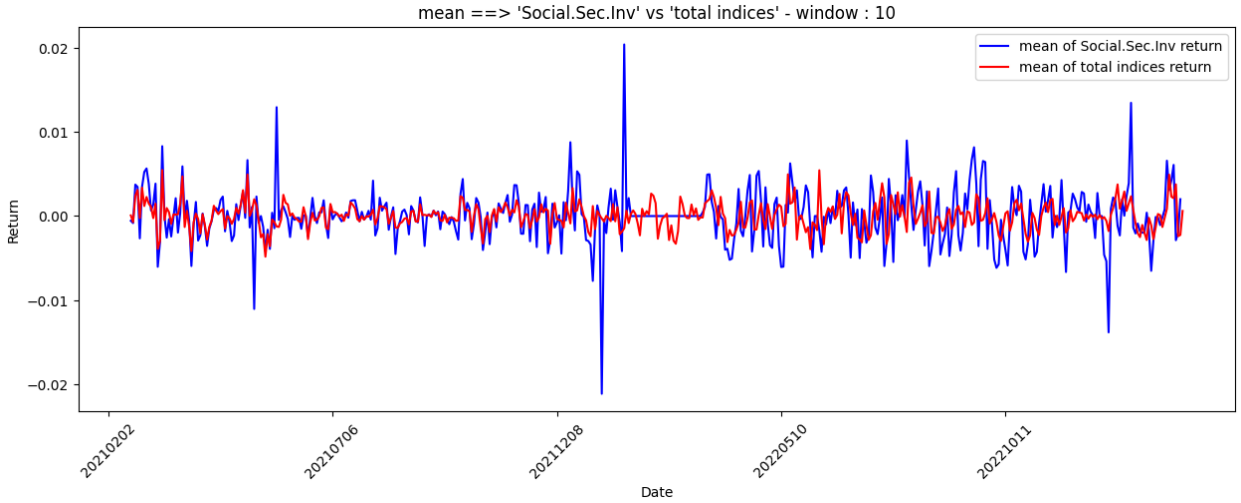


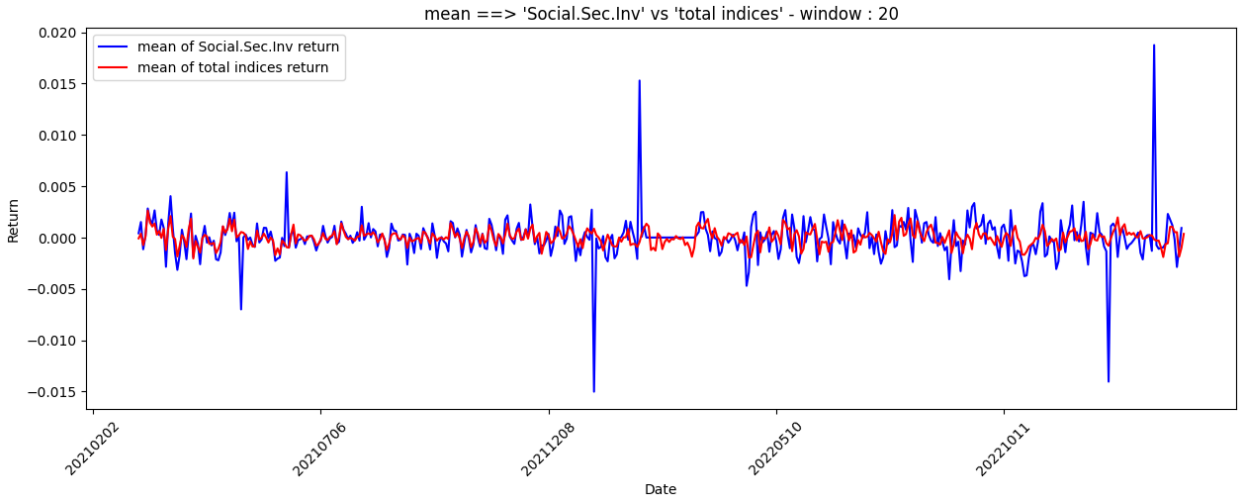


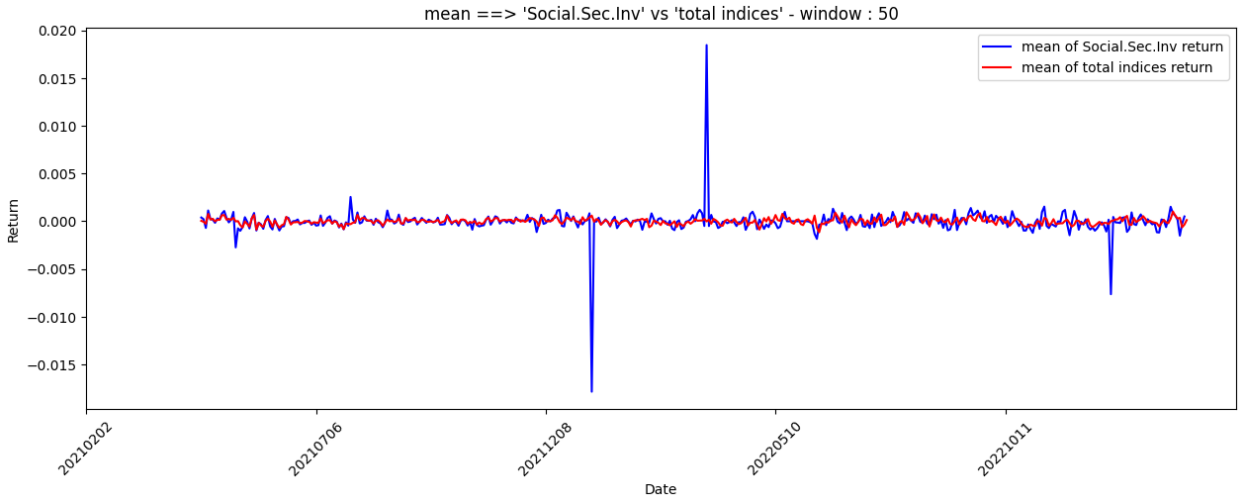
شکل 13. تغییرات میانگین و واریانس فولاد با شاخص کل در بازه 10، 20، 50 روزه

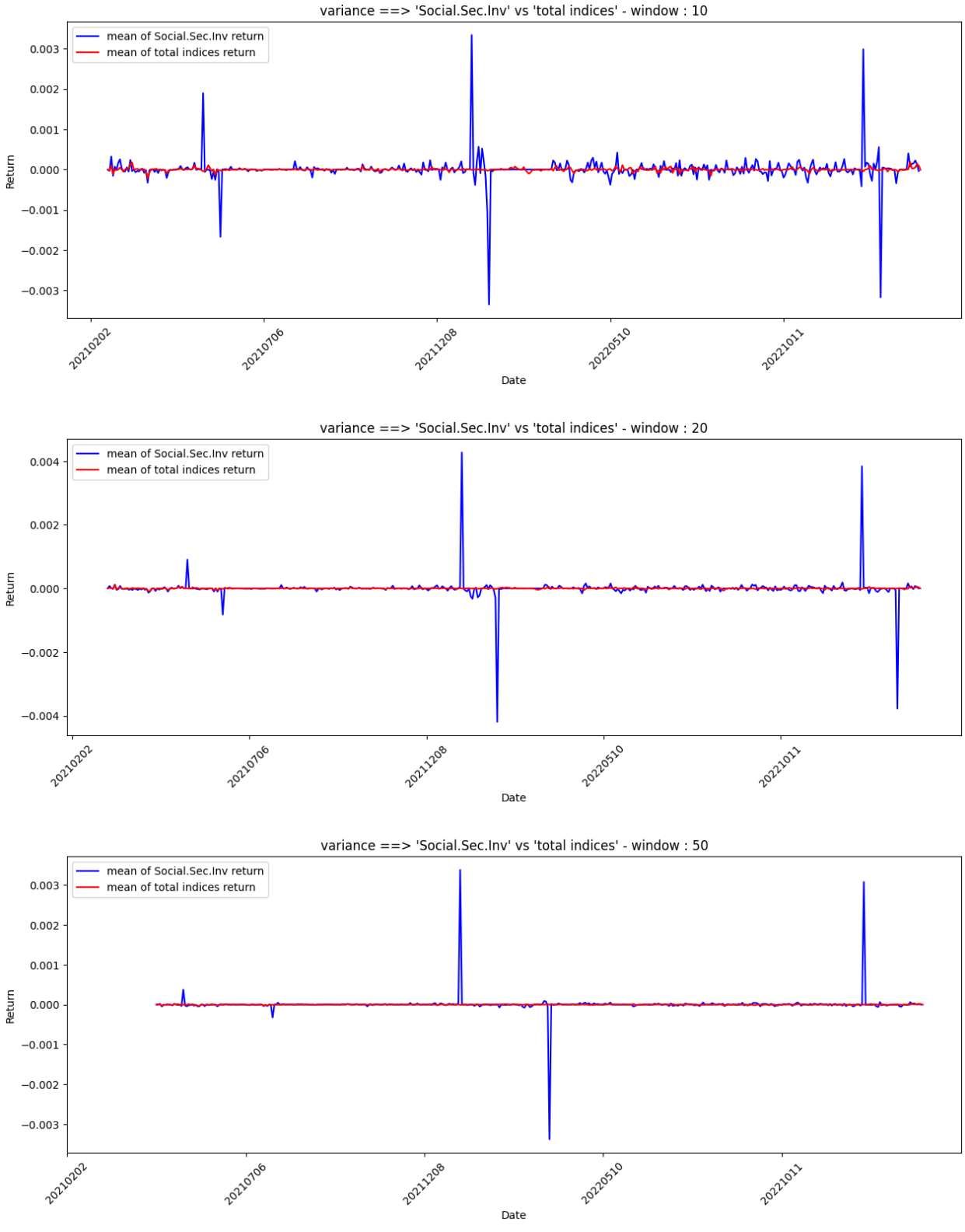
این نمودار ها نشان می‌دهد که در بازه 10 روزه، تغییرات میانگین فولاد و شاخص بسیار نزدیک بهم است و سهم فولاد نیز بر روی شاخص اثر گذار می‌باشد. با افزایش بازه، میتوان مشاهده کرد که نوسانات فولاد و شاخص کمتر شده و به یکدیگر نزدیک تر می‌شوند. میتوان گفت شاخص کل به نوعی پیرو و تابع سهم فولاد است و نوسانات فولاد بر روی شاخص کل بسیار تاثیر گذار است.

و همچنین واریانس آن ها نیز بسیار به یکدیگر نزدیک است.


mean ==> 'Social.Sec.Inv' vs 'total indices' - window : 10
mean of Social.Sec.Inv return
mean of total indices return
Return
Date
20210202
20210706
20211208
20220510
20221011


mean ==> 'Social.Sec.Inv' vs 'total indices' - window : 20
mean of Social.Sec.Inv return
mean of total indices return
Return
Date
20210202
20210706
20211208
20220510
20221011

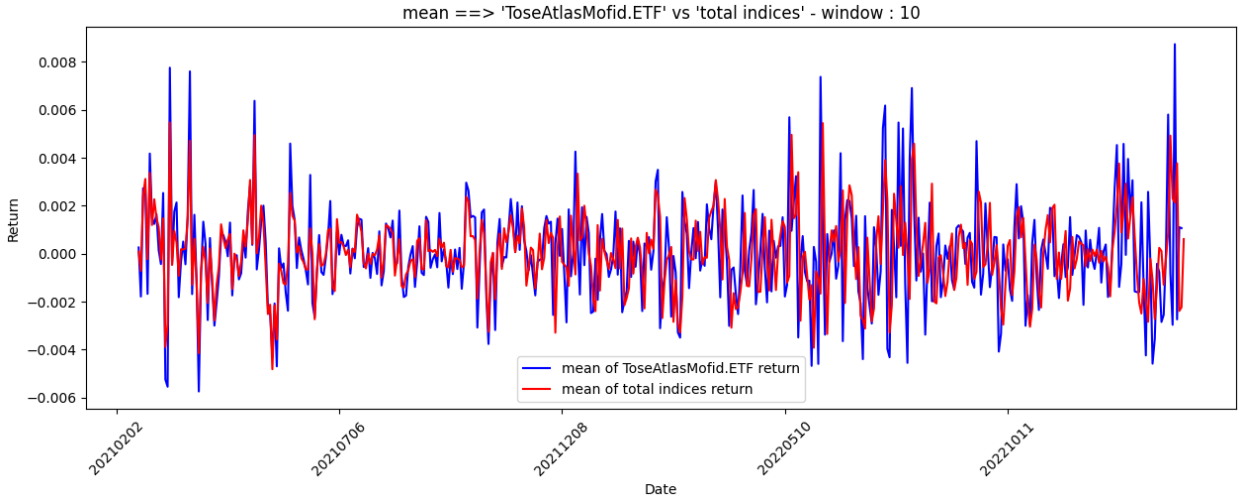

mean ==> 'Social.Sec.Inv' vs 'total indices' - window : 50
mean of Social.Sec.Inv return
mean of total indices return
Return
Date
20210202
20210706
20211208
20220510
20221011

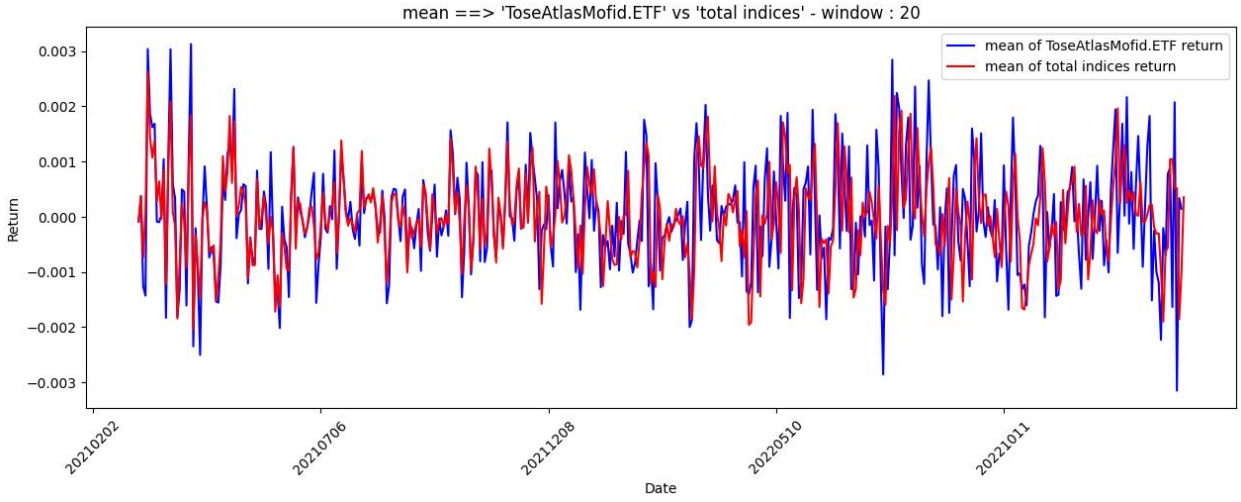


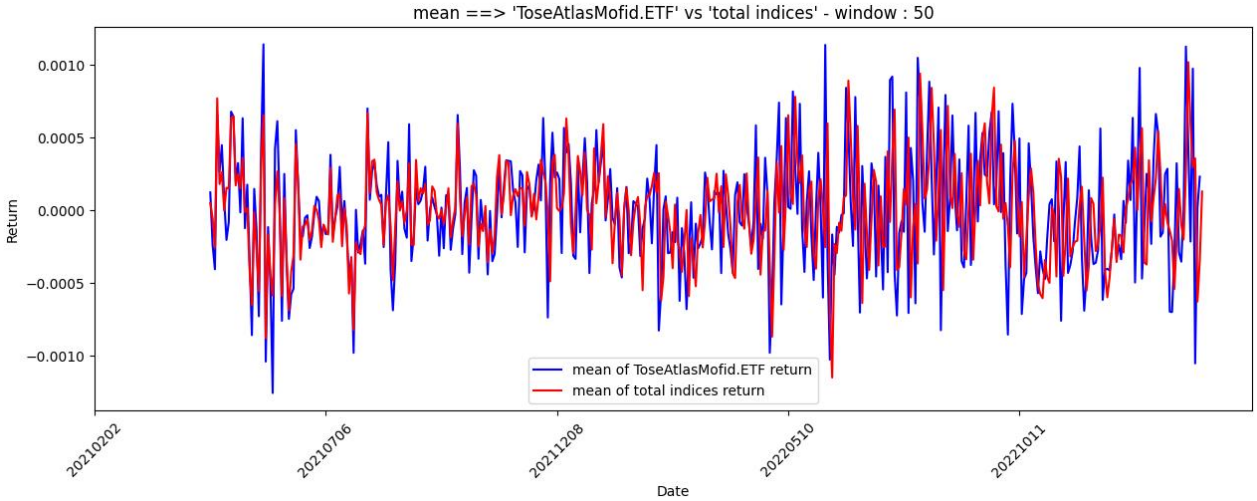
شکل 14. تغییرات میانگین و واریانس شستا با شاخص کل بازه های 10، 20، 50 روزه

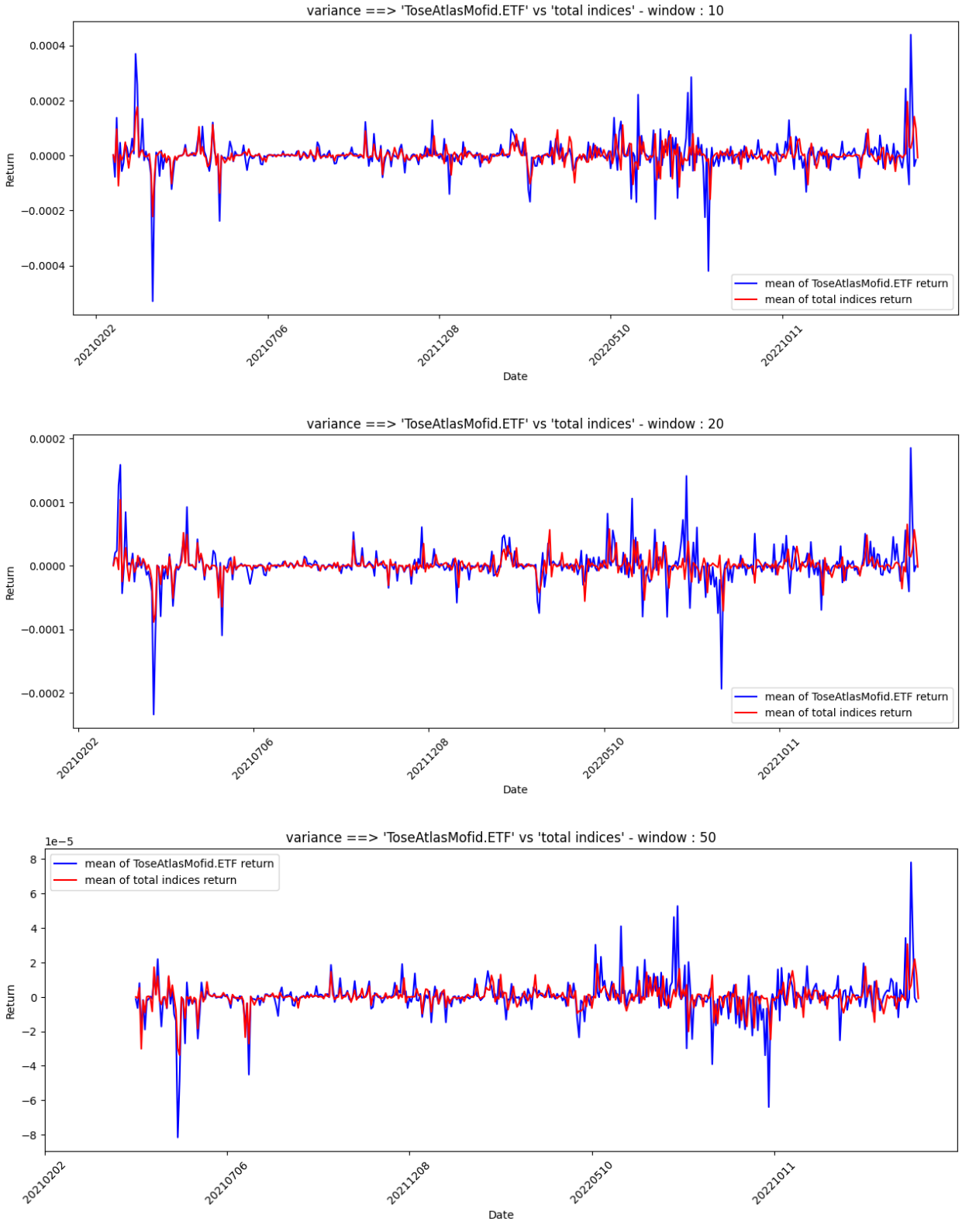
تغییرات میانگین سهم شستا در همه بازه ها نیز در بخش های 1 و 2 بسیار نزدیک به شاخص است و بر روی شاخص اثر گذار بوده است ولی در بخش سوم این اثر کمتر شده است. یکی از دلایل آن این است که میتوان مشاهده کرد که برای مدتی تغییرات میانگین برابر صفر بوده است که یعنی سهم بسته بوده و همچنین در بخش های 4 و 5 نیز این اثر کمتر شده است. دلیل این موضوع این است که در سال های اولی که سهم شستا وارد بورس شد بسیار بر روی شاخص تاثیر گذار بود و به نوعی جز سهم های شاخص ساز در بورس ایران بود و بعد از دوره افت طولانی بورس، تاثیر گذاری این سهم بر روی شاخص کمتر از قبل شد و به همین دلیل در یک سال اخیر نسبت به دو سال پیش کمتر بر روی شاخص تاثیر گذار بوده است.

تغییرات واریانس نزدیک به شاخص کل نمیباشد و بسیار نوسانی تر است. موضوع اشاره شده در بالا را نیز میتوان در واریانس هم مشاهده کرد. این بدان معنی است که در بخش های 1 و 2، در صورتی که سهم سهم شستا مثبت بود، شاخص کل نیز عمدتا مثبت بوده است و بر عکس ولی در سال اخیر این تاثیر گذاری کمتر شده است.


mean ==> 'ToseAtlasMofid.ETF' vs 'total indices' - window : 10
Return
Date
mean of ToseAtlasMofid.ETF return
mean of total indices return
0.008
0.006
0.004
0.002
0.000
−0.002
−0.004
−0.006
20210202
20210706
20211208
20220510
20221011


mean ==> 'ToseAtlasMofid.ETF' vs 'total indices' - window : 20
Return
Date
mean of ToseAtlasMofid.ETF return
mean of total indices return
0.003
0.002
0.001
0.000
−0.001
−0.002
−0.003
20210202
20210706
20211208
20220510
20221011


mean ==> 'ToseAtlasMofid.ETF' vs 'total indices' - window : 50
Return
Date
mean of ToseAtlasMofid.ETF return
mean of total indices return
0.0010
0.0005
0.0000
−0.0005
−0.0010
20210202
20210706
20211208
20220510
20221011



شکل 15. تغییرات میانگین و واریانس صندوق اطلس با شاخص کل در بازه های 10، 20، 50 روزه

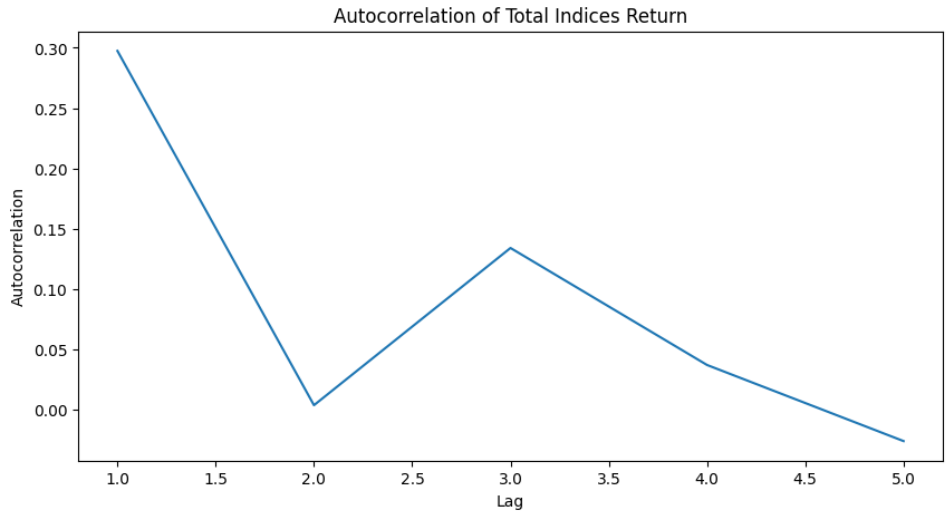
در همه بازه ها تغییرات میانگین و واریانس صندوق اطلس بسیار شبیه به شاخص کل است. دلیل این موضوع این است که صندوق های مالی موجود در بورس به نوعی تابع و پیرو شاخص کل می‌باشند و با شاخص کل تغییر میکنند. به همین دلیل با تغییرات در شاخص کل، صندوق های مالی مانند اطلس نیز تغییر می‌کنند. هر چه بازه بیشتر باشد این شباهت بیشتر می‌شود تا جایی که در بازه 50 روزه تغییرات میانگین صندوق اطلس منطبق بر شاخص کل است و تغییرات واریانس نیز بسیار نیز به تغییرات واریانس شاخص کل نزدیک است.

سوال 2)

الف)

میزان خود همبستگی یک سهم نشان دهنده این است که مقادیر قبلی سهم مورد نظر چقدر بر روی خودش تاثیر گذار است .
میزان خودهمبستگی بازده شاخص کل لگ های 1 تا 5 روزه به صورت زیر است:

```
autocorrelation with lag= 1 is : 0.2974189392944751
autocorrelation with lag= 2 is : 0.003669522910023966
autocorrelation with lag= 3 is : 0.13400637342822638
autocorrelation with lag= 4 is : 0.03701574877039716
autocorrelation with lag= 5 is : -0.026083400538123948
```



شکل 16. میزان خودهمبستگی بازده شاخص کل با لگ های 1 تا 5 روزه

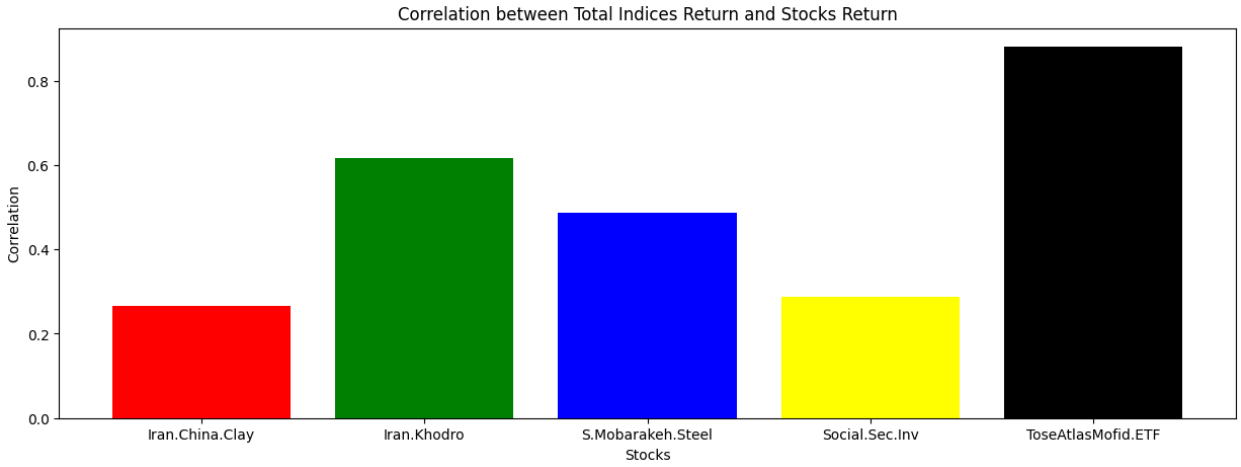
بدین معنی است که بازده شاخص کل در روز قبل میتواند بسیار بر روی بازده امروز تاثیر گذار باشد. این تاثیر گذاری به صورت نزولی به این صورت است که روز قبل بیشترین تاثیر، سپس روز سوم سپس روز چهارم، سپس روز پنجم و در اخر روز دوم بر روی شاخص کل تاثیر گذار هستند.

1 => 3 => 4 => 5 => 2

ب)

برای بدست آوردن میزان هبستگی بازده هر یک سهم ها با بازده شاخص کل، میزان بازده شاخص کل را با هر یک از سهم بر روی تاریخ merg کرده تا بتوان به طور دقیق آن را اندازه گیری کرد. میزان همبستگی هر یک سهم ها به صورت زیر است:

```
correlation between total indices return and Iran.China.Clay is : 0.2668998351007557
correlation between total indices return and Iran.Khodro is : 0.6172163845630182
correlation between total indices return and S*Mobarakeh.Steel is : 0.4875790807149877
correlation between total indices return and Social.Sec.Inv is : 0.28653975856350394
correlation between total indices return and ToseAtlasMofid.ETF is : 0.880946787387012
```



شکل 17. همبستگی بازده هر یک از سهم ها با بازده شاخص کل

این بدین معنی است که بازده صندوق اطلس بسیار با بازده شاخص کل شباهت دارد و منطبق بر آن است که از لحاظ منطقی و بر اساس مفاهیم بورسی نیز درست است. چرا که صندوق های مالی به نوعی نماینده شاخص کل می‌باشند و عمدتا با صعود یا نزول شاخص آن ها نیز صعود یا نزول می‌کنند.

سهم خودرو در جایگاه بعدی قرار دارد که به این معنی این تغییرات سهم خودرو وشاخص کل بسیار به یکدیگر نزدیک اند و شباهت بالایی دارند. همانطور که در قسمت قبل توضیح داده شد، یک سری از سهم های شرکت های بزرگ، سهم های شاخص ساز هستند که به این معنی است با توجه به حجم معاملاتی آن تغییرات این سهم ها بر روی شاخص کل بسیار تاثیر گذار بوده است و میتوانند شاخص را به طور قابل توجه تغییر دهند که شرکت خودرو یکی از آن ها می‌باشد و به همین دلیل همبستگی بالایی با شاخص کل دارد.

در جایگاه سوم سهم فولاد است که این سهم نیز جز سهم های شاخص ساز می‌باشد و همبستگی بالایی با شاخص کل دارد و تغییرات آن، بر روی شاخص کل بسیار تاثیر گذار است.

در جایگاه چهارم سهم شستا و در جایگاه پنجم سهم کخاک می‌باشد. که این سهم ها نیز جز سهم های شاخص ساز بورس هستند و بازده این دو سهم نسبت به سهم های موجود در بازار بورس (به غیر از این 5 سهم) همبستگی بالایی با بازده شاخص کل دارند. ولی نسبت به 3 هم دیگر همبستگی کمتری با شاخص کل دارند که به این معنی است که تغییرات آن ها بر روی شاخص کل تاثیر گذار هستند ولی نه به اندازه 3 سهم دیگر.

ج)

برای بدست آوردن مجموعه داده دلار و طلا از کتابخانه finpy_tse استفاده میکنیم و با دستور زیر در بازه دو ساله اخیر قیمت دلار و طلا را استخراج می‌کنیم.

```
import finpy_tse as fpy

dollar_dataset = fpy.Get_USD_RIAL(
                start_date='1399-11-14',
                end_date  ='1401-11-14',
                ignore_date=False,
                show_weekday=False,
                double_date=False).iloc[::-1]
dollar_dataset
```

J-Date	Open	High	Low	Close
1401-11-13	436050	436370	435810	436100
1401-11-12	437070	438090	425880	436440
1401-11-11	435880	438900	433310	437290
1401-11-10	439400	440900	433810	436560
1401-11-09	436480	439800	432800	438810
...	...	...	...	...
1399-11-19	237490	237550	236940	236960
1399-11-18	237520	237550	237440	237550
1399-11-16	237500	237550	237440	237470
1399-11-15	238680	238750	237640	237660
1399-11-14	238650	238750	238640	238710

486 rows × 4 columns

شکل 18. مجموعه داده دلار

اما برای طلا، یکی از سهم های مرتبط با قیمت طلا، که صندوق پشتوانه طلای لوتوس است انتخاب شده است که تغییرات آن بسیار به خود طلا نزدیک است. برای مجموعه داده طلا نیز داریم:

```
gold_dataset = fpy.Get_Price_History(
            stock='طلا',
            start_date='1399-11-14',
             end_date  ='1401-11-14',
            ignore_date=False,
            adjust_price=False,
            show_weekday=False,
            double_date=False)
gold_dataset
```

✓ 1.9s

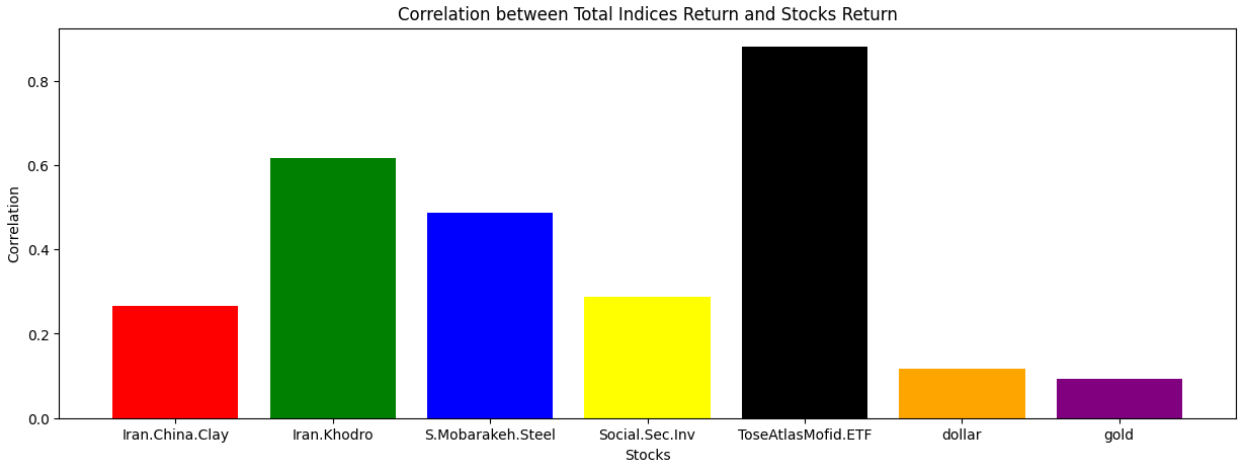
J-Date	Open	High	Low	Close	Final	Volume	Value	No	Ticker	Name	Market
1399-11-14	78000	78497	77410	78010	77816	302887	23569486634	452	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1399-11-15	76200	77200	75025	76980	76533	304792	23326504081	386	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1399-11-18	77127	77199	76660	77050	76849	1041142	80010877358	550	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1399-11-19	77400	78789	77131	78506	77896	454145	35376055884	652	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1399-11-20	81010	82202	81010	82200	81691	746437	60977179330	1034	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
...	...	...	...	...	...	...	...	...	...	...	...
1401-11-08	157200	161408	157200	157895	158732	10231523	1624074717704	8395	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1401-11-09	159050	165998	159020	161500	162270	13707035	2224240695262	9501	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1401-11-10	164490	164490	161711	162000	163088	5996543	977967134398	6295	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1401-11-11	160000	162000	157501	161900	159857	5470458	874492111895	4664	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم
1401-11-12	162500	163569	161960	163441	162887	4833843	787370720529	4486	طلا	صندوق س.پشتوانه طلای لوتوس	نامعلوم

479 rows × 11 columns

همانطور که مشاهده می‌شود، تاریخ های این دو مجموعه داده بر حسب شمسی است و بار دیگر از کتابخانه استفاده شده در بخش مقدماتی persiantools استفاده میکنیم تا این تاریخ ها را به میلادی تبدیل کنیم. سپس مقدار بازده برای هر کدام از این دو مجموعه داده بر حسب ستون close محاسبه میکنیم و سپس برای تطبیق این تاریخ ها با شاخص کل، هر یک را با شاخص کل بر حسب تاریخ merg می‌کنیم.

سپس بار دیگر همبستگی آن ها با شاخص کل را محاسبه میکنیم و داریم:

```
correlation between total indices return and Iran.China.Clay is : 0.26689983510075557
correlation between total indices return and Iran.Khodro is : 0.6172163845630182
correlation between total indices return and S*Mobarakeh.Steel is : 0.4875790807149877
correlation between total indices return and Social.Sec.Inv is : 0.28653975856350394
correlation between total indices return and ToseAtlasMofid.ETF is : 0.880946787387012
correlation between total indices return and USD is : 0.1168086433693989
correlation between total indices return and Gold is : 0.09325321654958031
```



شکل 18. همبستگی بازده دلار، سکه، و سایر سهم ها با بازده شاخص کل

میتوان دید که بازده های مربوط به دلار و طلا نیز بر روی شاخص کل تاثیر گذار است. **(شاید در این جدول کمترین مقدار را داشته باشند ولی این همبستگی در بازار بورس نیز، همبستگی نسبتا بالایی می‌باشد)**

همانطور که مشاهده می‌شود، دلار تاثیر گذار تر از طلا است که در بازار بورس نیز تاثیر دلار بیشتر از طلا است. این دو همبستگی نشان میدهد که در صورت صعود یا نزول قیمت دلار و سکه نیز شاخص سهام صعود و نزول میکند و بر روی آن تاثیر گذار است. چرا که با افزایش یا کاهش قیمت دلار و طلا، قیمت سهم ها نیز افزایش می‌یابد و در نتیجه شاخص نیز افزایش یا کاهش می‌یابد.

سوال 3)

الف)

در این قسمت از الگوریتم linear Regression برای پیشبینی استفاده می‌کنیم. در ابتدا داده ها به train و test با درصد 70 به 30 تقسیم میکنیم. این تقسیم به این صورت که از تاریخ 2 سال گذشته 70 درصد داده ها با افزایش تاریخ به عنوان آموزش و 30 درصد داده های جدید تر را به عنوان داده های تست در نظر می‌گیرد.

سپس تابعی به اسم make_window تعریف می‌کنیم خروجی آن ارایه ای از داده های است که در آن داده ها به صورت پنجره پنجره تقسیم شده اند. به عنوان مثال اگر پنجره برابر 1 باشد، به اسم صورت که داده 0 را x_train میگیرد و داده بعدی را به عنوان y_train در نظر میگیرد و تا اخر و اگر پنجره برابر 3 باشد، 3 تای اول را به عنوان ورودی 1 در نظر میگیرد و داده 4 ام را به عنوان y_train در نظر میگیرد. سپس از داده 2 تا 4 به عنوان ورودی دوم و ... .

همچنین برای تست نیز همین کار را تکرار می‌کنیم.

برای ارزیابی از MSE و MAE استفاده می‌کنیم.

برای بدست آوردن تعداد مناسب پنجره هم میتوان از مرحله قبل استفاده کرد ولی برای سرچ بهتر و مقایسه بهتر، چندین پنجره متفاوت در نظر گرفته شده است (**داده های مربوط به بخش قبل هم در آن ها وجود دارد**) و یک grid search بر روی پنجره، بهترین مقدار برای پنجره در نظر گرفته می‌شود.

اعداد در نظر گرفته شده برای پنجره برابر زیر است:

```
WINDOW_SIZE = [1, 2, 3, 4, 5, 10,15,16,17,18,19,20,21,22,23,24,25,30,35,40,45,50,55,60,65,70,75,80,85,90,95,100]
```

به دلیل حجم بالای تعداد پنجره ها، برای مشخص شدن بهرین پنجره، مقداری که کمترین MSE و MAE را داشته باشد استخراج می‌کنیم (جواب های هر پنجره در کد موجود است).

بهترین جواب های بدست آمده در این قسمت برابر زیر است:

```
Min MSE_results happen when window size is : 17
Min MSE_results is : 0.00010995653350643559
Min MAE_results happen when window size is : 18
Min MAE_results is : 0.0080093903806594 97
```

مقادیر بدست آمده برای خطا بسیار کم است که مقداری دور از واقعیت است که ممکن است این خطای کم به دلیل این باشد که مقادیر بازده بسیار کوچک هستند. میتوان برای اینکه جواب بهتر و مطابق با دنیای واقعی شود، روی بازده شاخص کل یک نرمال

سازی انجام شود و سپس وارد الگوریتم linear regression شود. که برای نرمال سازی، بازده شاخص کل را بین 1 تا 1- می‌آوریم.

نتایج بدست آمده برای این قسمت به صورت زیر است:

```
results for normalized data :
Min MSE_results happen when window size is : 17
Min MSE_results is : 0.0667584639322744
Min MAE_results happen when window size is : 18
Min MAE_results is : 0.19735226248588267
```

ب)

در این قسمت از الگوریتم های logestic regression و gradiant boosting tree استفاده می‌شود و از داده های نرمال سازی نشده استفاده می‌شود .

(به دلیل حجم بالای نتایج دیگر از داده های نرمال سازی شده استفاده نشده است ولی تنها لازم است به جای در نظر گرفتن داده های اصلی از ستون noramalized_return استفاده شود).

در این قسمت علاوه بر کار های قسمت قبلی، باید اعداد موجود در y_train و y_test را باینری کرد به ان صورت که اگر بزرگتر از صفر باشد برابر 1 و در غیر این صورت برابر صفر در نظر گفته شود که نشان دهنده این است که بازده مثبت بوده است یا خیر.

همچنین در این قسمت خواسته شده است تا چند روز آینده پیش بینی شود. به یاد داریم که در قسمت قبل y_train و y_test برابر مقدار بازده در روز بعد از آموزش میشد. برای پیشبینی چندین روز آینده، تعداد روز های مورد نیاز برای پیشبینی را وارد y_train و y_test میکنیم و به اندازه آن پیشبینی میکنیم. (**تعداد پنجره ها برابر قسمت قبل است**).

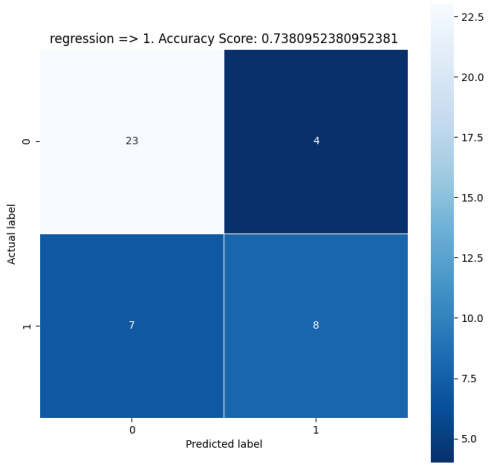
مقادیر و تعداد پیشبینی ها در ماتریس confusion متناسب با افزایش مقدار پنجره کم میشود.

همچنین نتایج برای همه پنجره ها در کد موجود است که به دلیل حجم بالای آن در گزارش آورده نشده است.

الگوریتم logestic regression

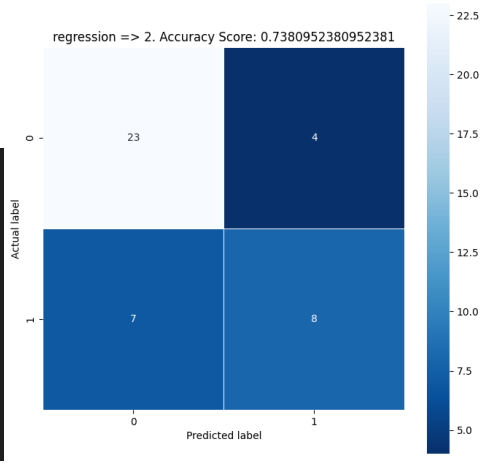
بهترین نتایج بدست آمده برای یک روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 102
Min Classifier_MSE_results is : 0.2619047619047619
Min Classifier_MAE_results happen when window size is : 102
Min Classifier_MAE_results is : 0.2619047619047619
Max Classifier_SCORE_results happen when window size is : 102
Max Classifier_SCORE_results is : 0.7380952380952381
confusion matrix for the best model is :
[[23  4]
 [ 7  8]]
```



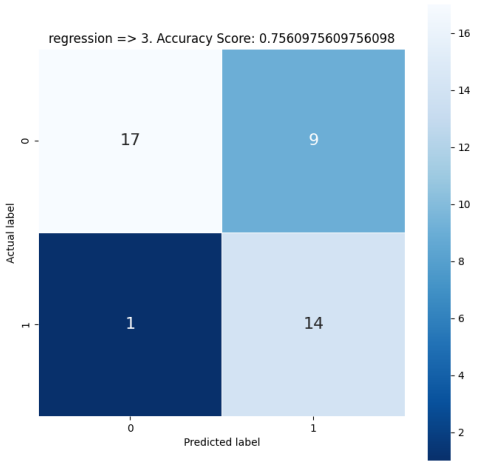
بهترین نتایج بدست آمده برای دو روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 101
Min Classifier_MSE_results is : 0.2619047619047619
Min Classifier_MAE_results happen when window size is : 101
Min Classifier_MAE_results is : 0.2619047619047619
Max Classifier_SCORE_results happen when window size is : 101
Max Classifier_SCORE_results is : 0.7380952380952381
confusion matrix for the best model is :
[[23  4]
 [ 7  8]]
```



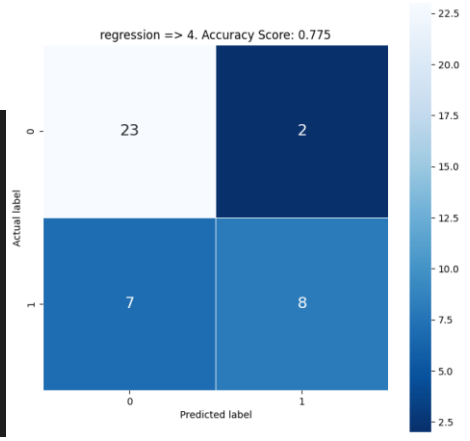
بهترین نتایج بدست آمده برای سه روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 101
Min Classifier_MSE_results is : 0.24390243902439024
Min Classifier_MAE_results happen when window size is : 101
Min Classifier_MAE_results is : 0.24390243902439024
Max Classifier_SCORE_results happen when window size is : 101
Max Classifier_SCORE_results is : 0.7560975609756098
confusion matrix for the best model is :
[[17  9]
 [ 1 14]]
```



بهترین نتایج بدست آمده برای چهار روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 101
Min Classifier_MSE_results is : 0.225
Min Classifier_MAE_results happen when window size is : 101
Min Classifier_MAE_results is : 0.225
Max Classifier_SCORE_results happen when window size is : 101
Max Classifier_SCORE_results is : 0.775
confusion matrix for the best model is :
[[23  2]
 [ 7  8]]
```

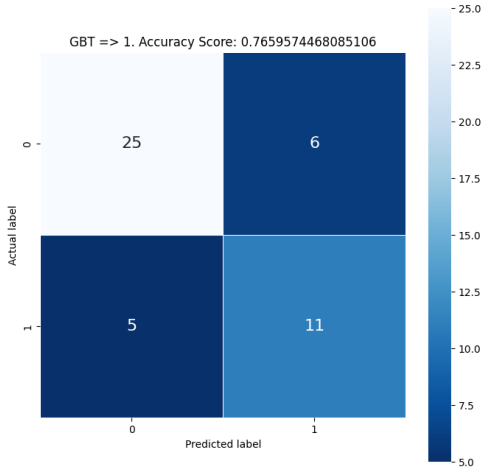


با توجه به نتایج بدست آمده میتوان دید که بهترین نتایج برای الگوریتم logestic regression برابر 0.775 درصد برای روز چهارم است. این نتایج را میتوان از روی ماتریس confusion نیز مشاهده کرد که تعداد خطا ها در روز 4 ام برابر 9 است.

الگوریتم Gradiant boosting tree

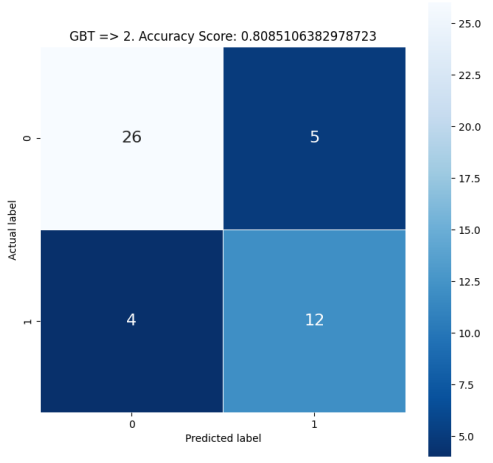
بهترین نتایج بدست آمده برای یک روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 97
Min Classifier_MSE_results is : 0.23404255319148937
Min Classifier_MAE_results happen when window size is : 97
Min Classifier_MAE_results is : 0.23404255319148937
Max Classifier_SCORE_results happen when window size is : 97
Max Classifier_SCORE_results is : 0.7659574468085106
confusion matrix for the best model is :
[[25  6]
 [ 5 11]]
```



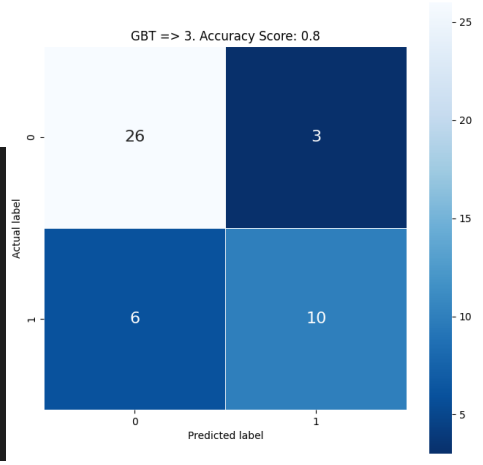
بهترین نتایج بدست آمده برای دو روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 96
Min Classifier_MSE_results is : 0.19148936170212766
Min Classifier_MAE_results happen when window size is : 96
Min Classifier_MAE_results is : 0.19148936170212766
Max Classifier_SCORE_results happen when window size is : 96
Max Classifier_SCORE_results is : 0.8085106382978723
confusion matrix for the best model is :
[[26  5]
 [ 4 12]]
```



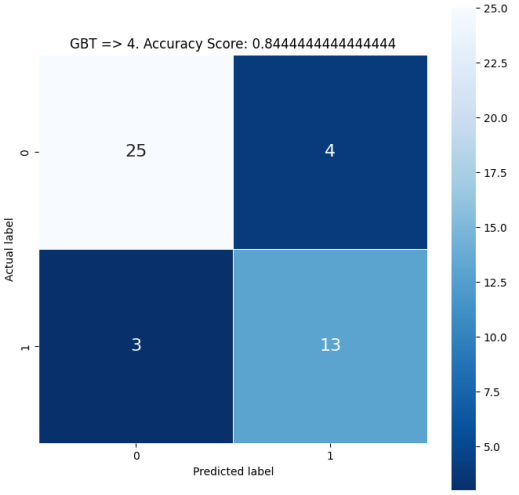
بهترین نتایج بدست آمده برای سه روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 97
Min Classifier_MSE_results is : 0.2
Min Classifier_MAE_results happen when window size is : 97
Min Classifier_MAE_results is : 0.2
Max Classifier_SCORE_results happen when window size is : 97
Max Classifier_SCORE_results is : 0.8
confusion matrix for the best model is :
[[26  3]
 [ 6 10]]
```



بهترین نتایج بدست آمده برای چهار روز پیشبینی:

```
Min Classifier_MSE_results happen when window size is : 96
Min Classifier_MSE_results is : 0.15555555555555556
Min Classifier_MAE_results happen when window size is : 96
Min Classifier_MAE_results is : 0.15555555555555556
Max Classifier_SCORE_results happen when window size is : 96
Max Classifier_SCORE_results is : 0.8444444444444444
confusion matrix for the best model is :
[[25  4]
 [ 3 13]]
```



بنابر نتایج بدست آمده میتوان بار دیگر مشاهده کرد که پیشبینی 4 روز بهترین دقت را بدست آورده و ماتریس confusion نیز تنها دارای 7 اشتباه است و در بقیه موارد پیشبینی درست بوده است.

ج)

در این قسمت یک دیتافریم جدید تشکیل میدهیم که ستون های آن برابر بازده هر یک از سهم ها است و در نهایت یک ستون دیگر برابر مثبت یا منفب بودن بازده شاخص کل تشکیل می‌دهیم.

	date	iran_china_clay	iran_khodro	s_mobarakeh_steel	social_sec_inv	tose_atlas_mofid	total_indices	P_or_N
0	20230201	-0.001497	-0.001365	0.000000	-0.006309	-0.005092	0.000445	1
1	20230131	-0.002240	0.033134	0.020443	0.007415	0.022111	0.010958	1
2	20230130	-0.014717	-0.004911	0.000000	0.007471	0.003001	-0.001628	0
3	20230129	-0.048985	-0.046488	-0.034539	-0.046796	-0.029973	-0.036181	0
4	20230128	-0.039005	-0.013527	0.000000	-0.008073	-0.006776	-0.007365	0
...	...	...	...	...	...	...	...	...
474	20210209	-0.002257	0.023346	0.049029	0.030065	-0.003525	0.018170	1
475	20210208	-0.012483	0.044715	0.049515	0.040924	0.064443	0.018976	1
476	20210207	-0.016012	-0.042802	-0.044527	-0.048557	-0.041440	0.030993	1
477	20210206	-0.007619	-0.030189	-0.040071	-0.040018	-0.010228	-0.032110	0
478	20210203	0.000000	0.000000	0.000000	0.000000	0.000000	-0.026924	0

479 rows × 8 columns

که مقادیر Nan در آن برابر صفر است.

در نهایت برای ستون های بازده کخاک تا بازده شاخص کل را به عنوان x و ستون بازده شاخص کل را به عنوان y به مدل میدهیم. همانند قسمت الف y_train و y_test برابر مقدار بازده شاخص کل در روز بعد است.

بهترین نتایج بدست آمده در این قسمت برابر زیر است:

```
Min multi_MSE_results happen when window size is : 1
Min multi_MSE_results is : 5.5499984426334 53e-05
Min multi_MAE_results happen when window size is : 1
Min multi_MAE_results is : 0.005686531082135615
```

که خطای MSE عدد بسیار کوچکی است و برابر 0.000055 است که مقداری دور از واقعیت است. دلیل این موضوع ممکن است کوچک بودن بازده سهم ها باشد.که میتوان آن را بر اساس بخش قبلی نرمال سازی کرد و داریم:

	date	iran_china_clay	iran_khodro	s_mobarakeh_steel	social_sec_inv	tose_atlas_mofid	total_indices	P_or_N
0	20230201	0.823403	-0.255454	0.765107	0.864868	-0.130765	-0.097535	1
1	20230131	0.820841	0.278640	0.835695	0.892970	0.311597	0.161523	1
2	20230130	0.777854	-0.310350	0.765107	0.893083	0.000846	-0.148611	0
3	20230129	0.659781	-0.954047	0.645844	0.781970	-0.535364	-1.000000	0
4	20230128	0.694170	-0.443746	0.765107	0.861257	-0.158150	-0.289968	0
...	...	...	...	...	...	...	...	...
474	20210209	0.820783	0.127116	0.934399	0.939345	-0.105278	0.339212	1
475	20210208	0.785549	0.457948	0.936077	0.961581	1.000000	0.359069	1
476	20210207	0.773390	-0.896969	0.611358	0.778364	-0.721848	0.655172	1
477	20210206	0.802311	-0.701700	0.626743	0.795849	-0.214280	-0.899681	0
478	20210203	0.828561	-0.234326	0.765107	0.877786	-0.047959	-0.771910	0

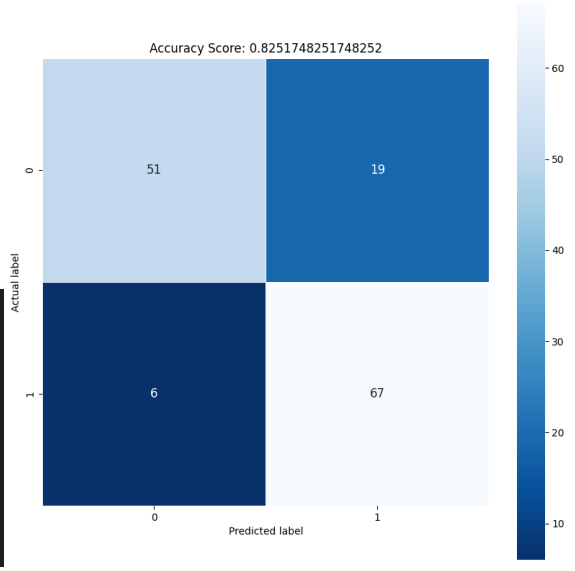
479 rows × 8 columns

که بهترین نتایج بدست آمده از داده های نرمال سازی شده برای پیشبینی بازده شاخص کل با استفاده الگوریتم linear regression برابر زیر است:

```
Min multi_MSE_results happen when window size is : 1
Min multi_MSE_results is : 0.0336959850444029 44
Min multi_MAE_results happen when window size is : 1
Min multi_MAE_results is : 0.140116753138377
```

که بهترین نتایج بدست آمده از داده های نرمال سازی شده برای پیشبینی مثبت یا منفی بودن شاخص کل با استفاده الگوریتم gradiant boosting tree برابر زیر است:

```
Min Classifier_MSE_results happen when window size is : 1
Min Classifier_MSE_results is : 0.17482517482517482
Min Classifier_MAE_results happen when window size is : 1
Min Classifier_MAE_results is : 0.17482517482517482
Max Classifier_SCORE_results happen when window size is : 1
Max Classifier_SCORE_results is : 0.8251748251748252
Max Classifier_CONF_MAT_results happen when window size is :
[[51 19]
 [ 6 67]]
```



د)

در این قسمت نیز الگوریتم lasso را پیاده سازی میکنیم. همانطور که در صورت سوال خواسته شده است، اطلاعات ده روز گذشته را گرفته و روز بعد را پیشبینی می‌کنیم. کاری که لاسو انجام میدهد این که ممکن اطلاعات ده روز قبل به یک اندازه مهم نباشند و درصد اهمیت هر روز متفاوت باشد به همین دلیل به ده روز گذشته وزن خاصی نسبت میدهد و روز های مهم وزن های بیشتری را میگیرند.

برای پیاده سازی مقدار alpha برابر چندین مقدار مختلف قرار داده شده است که در زیر نشان داده است و alpha که کمترین خطا MSE را بدست بیاورد در زیر نمایش داده است.

```
alph =[0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, 0.00005, 0.00001]
```

```
Min MSE_results happen when alpha is : 1e-05
Min MSE_results is : 0.00013134683365775 89
Min MAE_results happen when alpha is : 1e-05
Min MAE_results is : 0.008795771308151413
best coef is :
[ 0.          0.         -0.02861085 -0.         -0.         -0.
 -0.          0.06028942 -0.          0.17206386]
```

همانطور که مشاهده میشود روز های سوم و هشتم و دهم از وزن بیشتری برخوردار هستند و مهم تر از بقیه روز ها می‌باشند.

همچنین این کار را یک بار داده های نرمال شده برای بازده شاخص کل انجام میدهیم و داریم:

```
Min MSE_results happen when alpha is : 1e-05
Min MSE_results is : 0.07256000254345552
Min MAE_results happen when alpha is : 1e-05
Min MAE_results is : 0.20737903808542096
best coef is :
[-0.01092747  0.11250129 -0.12441043 -0.00919772 -0.05049381 -0.00134642
 -0.06270402  0.19167007 -0.11296993  0.29644328]
```

می‌توان دید که مقدار خطا به واقعیت نزدیک شده است و همچنین مقادیر وزن انتخابی بر 10 روز دیگر 0 نیست و هر روز از یک وزن خاص برخوردار است. بیشترین وزن به ترتیب روز دهم، هشتم، سوم و.. است.