Autocorrelation of Stock Returns

Company : Ultratech Cement Limited

Investors and traders make buying and selling decisions of stocks based on the current and past data. There are many approaches to formalize a strategy and make investment decision, but one of the popular approach is doing a technical analysis. Technical analyst observes the historic patterns of the stock market to make predictions about its future performance.

One of the most controversial issue in finance is possibly whether the market is efficient or not. The efficient market hypothesis (EMH) states that share prices reflect all information. If the stock prices, correctly and fully reflect all the relevant information, then no arbitrage opportunities would exist.

One of the first intuitive steps to take when studying market efficiency is to look at serial autocorrelation in the returns. Autocorrelation is a characteristic of data which refers to the degree of correlation of the same variables between two successive time intervals. The predictability of stock returns based on past returns, is essential to examine the stock market efficiency hypothesis. We can find whether the past prices are fully reflected in the current stock prices with the help of autocorrelation, also known as serial correlation.

Process of data analysis :

(I)Data Preparation:

- Get the data : The data is taken from Yahoo! Finance, it has weekly stock price data of Ultratech Cement Limited company for the period of three years from October 1, 2017 to September 30, 2020.
- Key variables :
    - Adjusted close is the closing price after adjustments for all applicable splits and dividend distributions.
    - Log return for a time period is the sum of the log returns of partitions of the time period.
- Import the required libraries
- Load the data into a pandas DataFrame
- Preprocess the data
    - Extract only the columns that is required for the analysis. Here, 'Adj Close' and 'Date' are the desired column.
    - Detect and treat the null values
    - Calculate log returns

$$\text{Log returns : } R_t = \ln(P_t / P_{t-1})$$

(II)Data Analysis:

- Summarize the data
- Lagged scatter plot
- ACF plot
- Autocorrelation test of independence

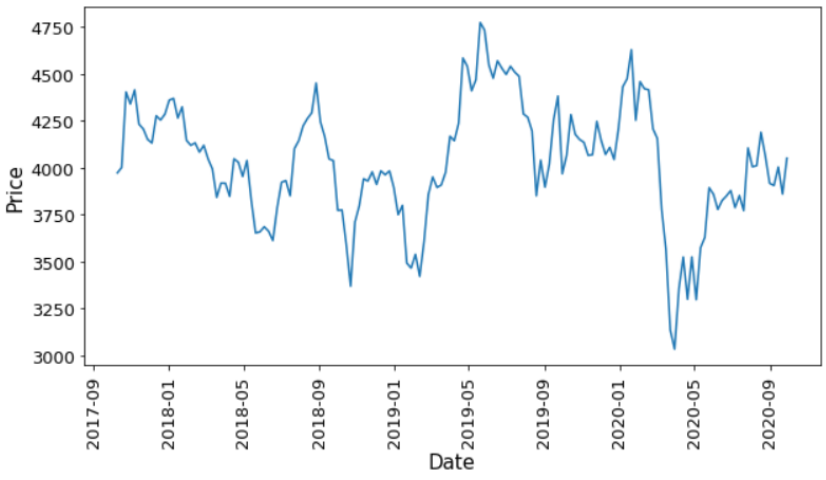
Results of data analysis:

	Adj Close	Log Returns
count	156.000000	156.000000
mean	4030.754330	0.000244
std	319.660626	0.039359
min	3032.040771	-0.128275
25%	3849.664856	-0.019458
50%	4041.183960	-0.002673
75%	4247.246948	0.020667
max	4772.781738	0.101356

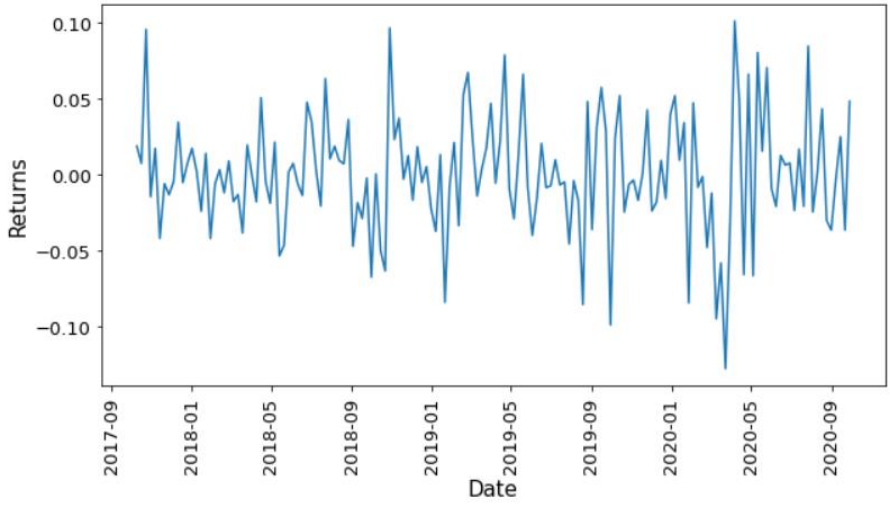
	Skewness	Kurtosis
Adj Close	-0.353034	0.272447
Log Returns	-0.089847	0.745447

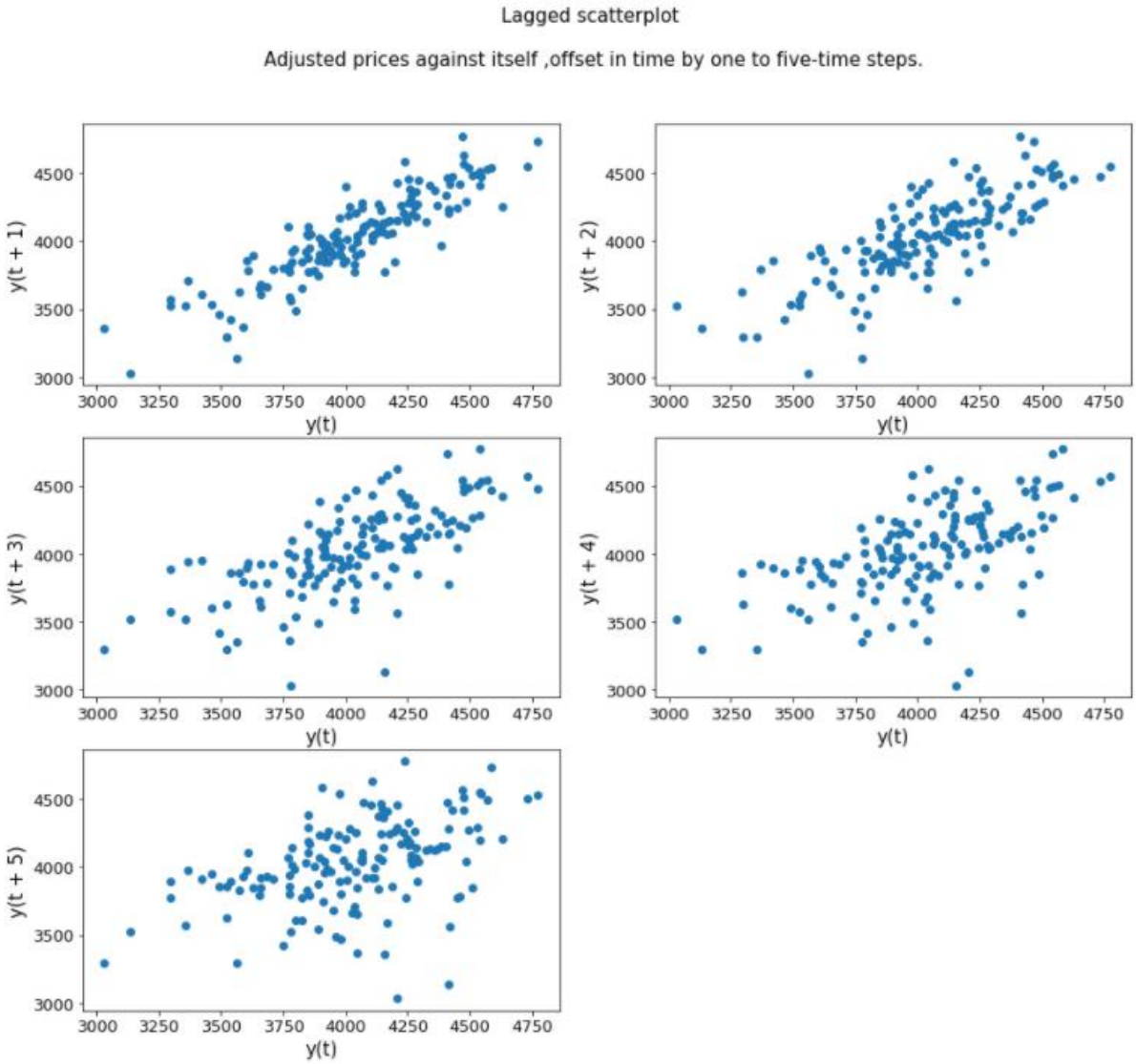
Prices and log returns are both moderately skewed. Prices are platykurtic while the returns are leptokurtic in nature.

Ultratech cement limited's weekly stock price data for the period of three years from Oct 1, 2017 to Sept 30, 2020

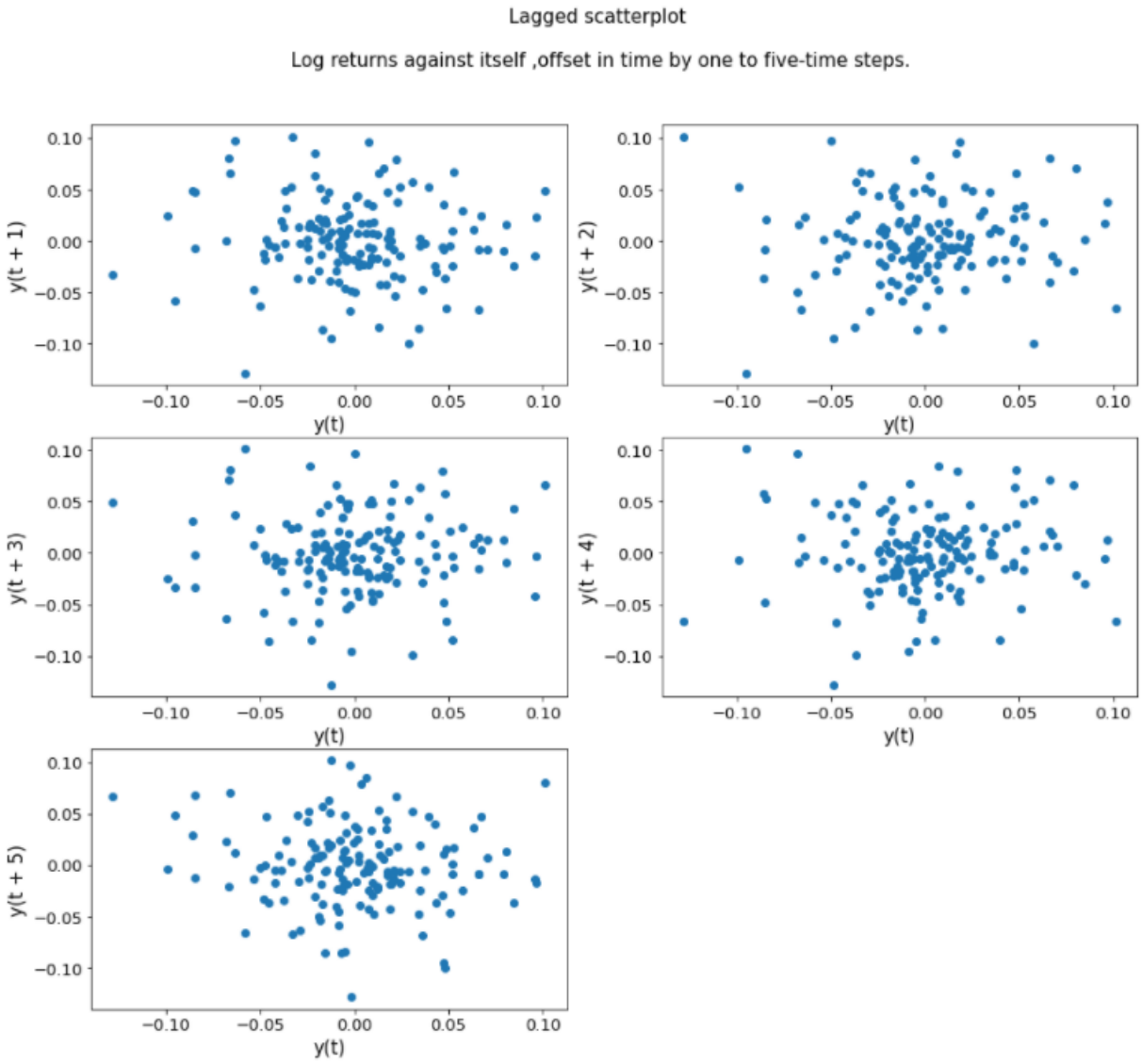


Ultratech cement limited's weekly stock price returns for the period of three years from Oct 1, 2017 to Sept 30, 2020

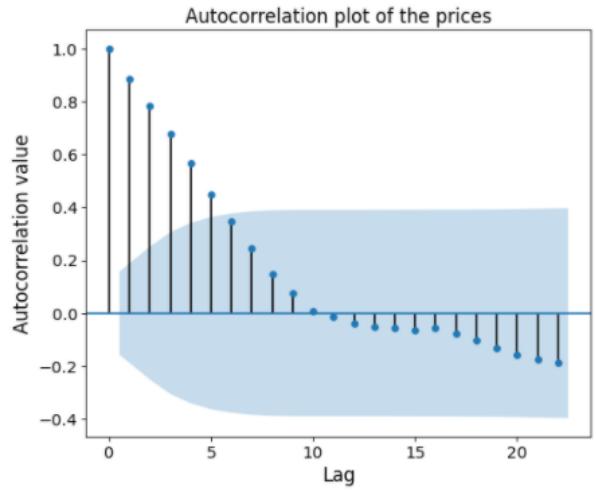


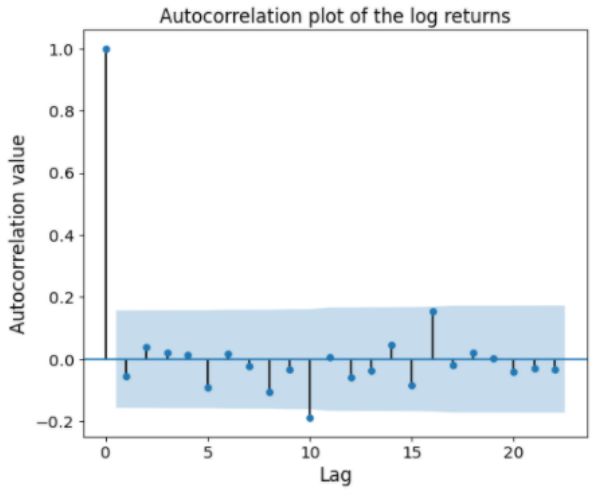


There is a pattern and the prices are positively correlated. As the lag increases, the randomness in the graph increases and the correlation decreases.

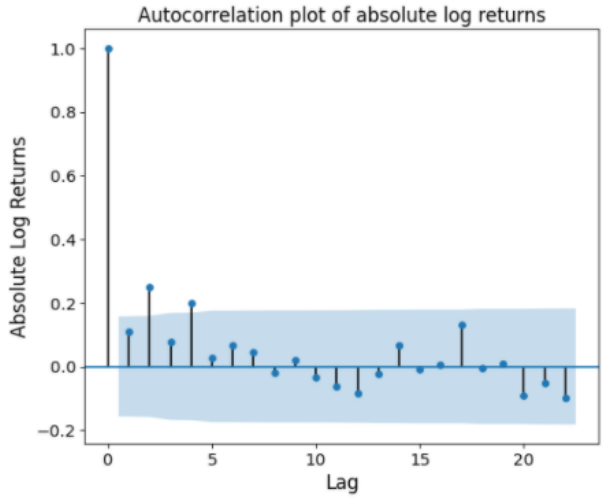


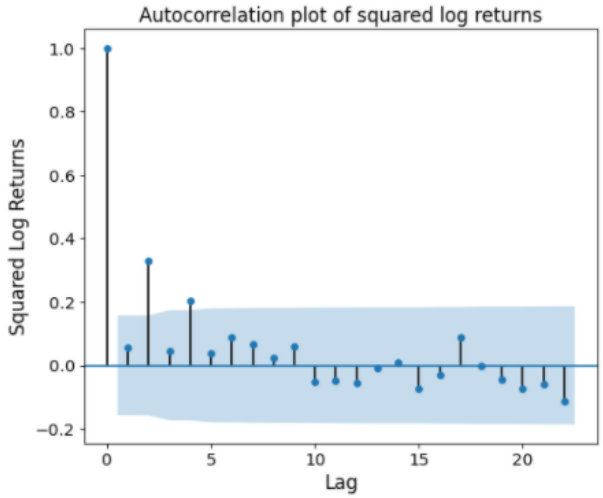
There are no peculiar patterns observed in the lagged scatter plot. The returns are most probably random and have weak or no correlation.





The weekly prices are highly correlated. Overall, the auto-correlation values for the log returns go on decreasing and suddenly increases for few lags. But almost all the autocorrelation values are within the confidence interval except for the value at lag 10.





For both the above graphs, we can see that the auto-correlation values corresponding to lag values 2 and 4 are lying outside the confidence interval. There is decreasing auto-correlation for both absolute and squared log returns.

Testing the Significance of Autocorrelation

The Ljung–Box test is a way to test for the absence of serial autocorrelation, up to a specified lag k. It is a more quantitative way to test for autocorrelation at multiple lags jointly.

H_0 : The data are independently distributed.

H_A : The data are not independently distributed; they exhibit serial correlation.

If the p-value ≤ 0.05, then we reject the null hypothesis i.e. . the data are not independently distributed.

If the p-value > 0.05, then we fail to reject the null hypothesis i.e. the data are independently distributed.

---Ljung-Box test for adjusted prices---

Lags	Test statistic	p-value	Reject/Accept H0
1	124.552977	6.375410e-29	Reject
2	223.153535	3.490002e-49	Reject
3	297.422945	3.593423e-64	Reject
4	349.436579	2.320748e-74	Reject
5	381.944586	2.306817e-80	Reject
6	401.784504	1.155643e-83	Reject
7	411.859458	6.821112e-85	Reject
8	415.574346	8.712996e-85	Reject
9	416.518784	4.082179e-84	Reject
10	416.530093	2.846263e-83	Reject
11	416.565557	1.855502e-82	Reject
12	416.819607	1.034371e-81	Reject
13	417.289999	4.962459e-81	Reject
14	417.807808	2.234228e-80	Reject
15	418.545936	8.710486e-80	Reject
16	419.061108	3.656126e-79	Reject
17	420.151606	1.126093e-78	Reject
18	421.966648	2.378348e-78	Reject
19	425.062752	2.647997e-78	Reject
20	429.563639	1.472837e-78	Reject

From the Ljung-box test, we can say there is significant autocorrelation between the Adjusted prices and its lagged version. The prices are not distributed independently.

---Ljung-Box test for log returns---

Lags	Test statistic	p-value	Reject/Accept H0
1	0.492075	0.483003	Accept
2	0.758638	0.684327	Accept
3	0.831136	0.842006	Accept
4	0.859690	0.930271	Accept
5	2.235070	0.815754	Accept
6	2.277651	0.892475	Accept
7	2.372482	0.936389	Accept
8	4.185627	0.839999	Accept
9	4.375685	0.884994	Accept
10	10.439414	0.402824	Accept
11	10.450169	0.490405	Accept
12	11.003767	0.528596	Accept
13	11.244841	0.590315	Accept
14	11.611722	0.637453	Accept
15	12.834961	0.615044	Accept
16	17.067164	0.381264	Accept
17	17.137766	0.445071	Accept
18	17.228423	0.507456	Accept
19	17.229526	0.574321	Accept
20	17.519851	0.619005	Accept

From the Ljung-box test, we can say, there is no autocorrelation between the log returns and any of its lagged version. The returns are distributed independently.

Conclusion:

The randomness of the Adjusted prices and the log returns are assessed with the help of serial autocorrelation and Ljung-Box test. Weekly stock price data of Ultratech Cement Limited company for the period of 3 years from October 1, 2017 to September 30, 2020 has autocorrelation that decays very slowly. This means they have some memory of past events. But the returns behave differently. The lagged scatter plots between the returns and its lagged version are most probably random and have weak or no correlation. The autocorrelation values for the log returns go on decreasing and suddenly increases for few lags, mostly all the values are within the confidence interval except for the value at lag 10. The ljung box test helps us to conclude that all the autocorrelation coefficients of the log returns up to order 20 are equal to

zero i.e. the log returns are independently distributed. Hence, we can say that probably the market is weakly efficient and technical analysis is an ineffective way of predicting stock returns.