

HA 8.2

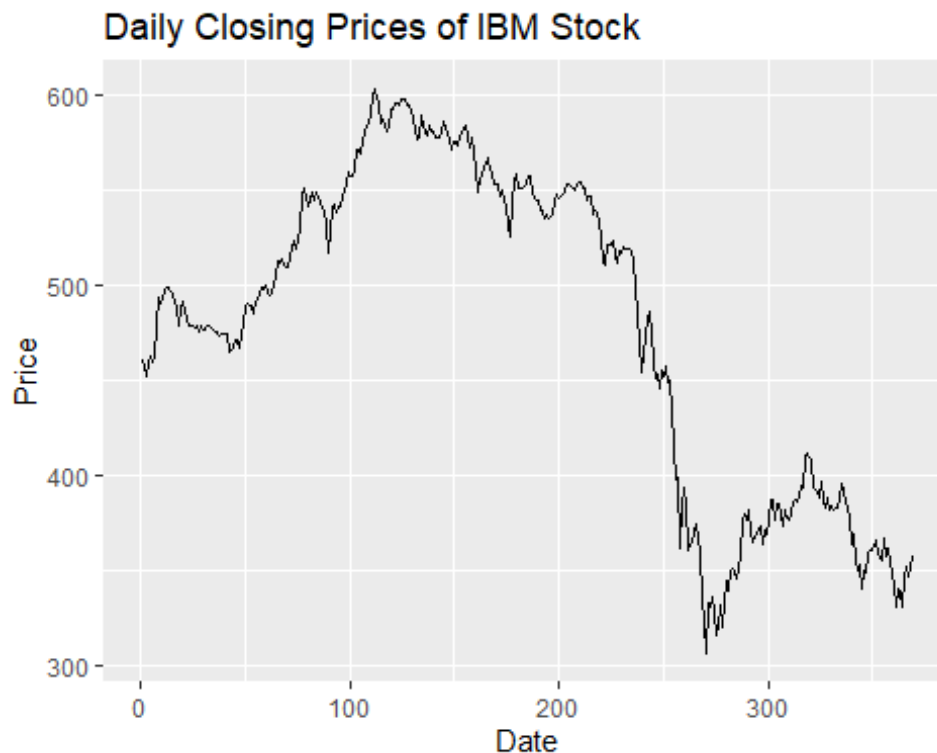
Jean Jimenez

HA 8.2

A classic example of a non-stationary series is the daily closing IBM stock price series (data set `ibmclose`). Use R to plot the daily closing prices for IBM stock and the ACF and PACF. Explain how each plot shows that the series is non-stationary and should be differenced.

Plotting Daily Closing Prices

The data set `ibmclose` is already nicely cleaned and organized after looking at it. It is time series data of the IBM stock at closing after each day. To plot the daily closing prices, I used `autoplot()` and added labels using `ggtitle()` and `labs()`.



For the span of the data set, IBM stock experienced an initial sharp period of growth where the stock price increased ~\$50 in a less than 25 day time period. Afterwards, there was a slow decline; followed by an upward trend that lasted about 75 days (at around day 110). At this point, the stock price reached its peak price of ~\$600 and then the price continued to decline until it reached its minimum of ~\$300 160 days later (at around day 275). After this period there was a short time frame of growth, followed by decreasing.

I can't really tell seasonality from this just because it is data from one year. Maybe there is some cyclical effects that come from the day of the week (maybe different prices at closing Monday vs. closing Friday) but its difficult to tell with this plot. The decline in stock prices possibly correlate to some negative news related to the company around that time period.

Plotting ACF

The Autocorrelation Function (ACF) plot is a visualization that shows how the values of a time series are related to each other at different time lags. To plot, I use the `ACF()` function.

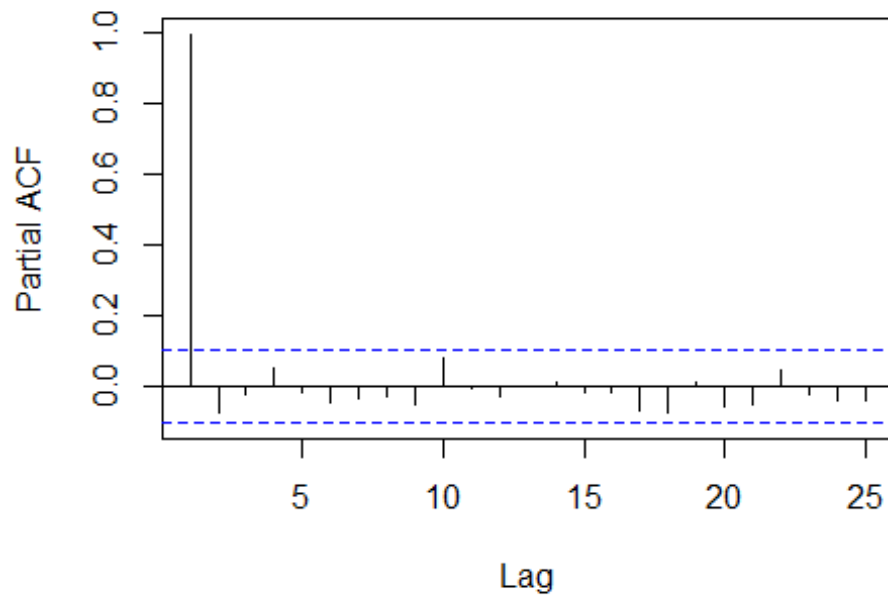


The plot shows that the autocorrelation starts very high (close to 1) at lag 1 and gradually decreases as the lag increases. This means the daily closing prices are highly correlated with their recent past values and the influence of past prices diminishes over time. The lags are not statistically significant because it exceeds the bounds of the confidence interval (dotted blue line).

Plotting PACF

Partial Autocorrelation Function (PACF) measures the correlation between a time series and its lagged values, but this time accounting for the relationships of all shorter lags. Unlike the ACF, the PACF isolates the direct effect of a specific lag. That way, we can tell which specific lag is significant or not. To plot this, I use the `pacf()` function.

PACF of IBM Closing Prices



The first lag has a significant partial autocorrelation (close to 1), which supports a strong direct relationship with the previous day's price. Lags more than 1 day have partial autocorrelations that are much lower and within the confidence intervals (blue dashed lines), meaning these lags do not have significant direct effects. IBMs stock price is really really influenced by the previous day's stock price