VU VRIJE UNIVERSITEIT AMSTERDAM | Universiteit van Amsterdam

Data Processing

Market Data

Market Data is temporal, it's a series of trading events that happen in a moment, each moment is called a tick. A tick contains the time of the event (Data, Time), stock's name (Ticker), trade data (Volume, Last Trade), and quote data (price and size of bid and ask).

In modern marketplaces, there are tons of trade data per second. To reduce the volume of data, we usually bucket these data into equally space-time integrals like minutes or days.

Market Data

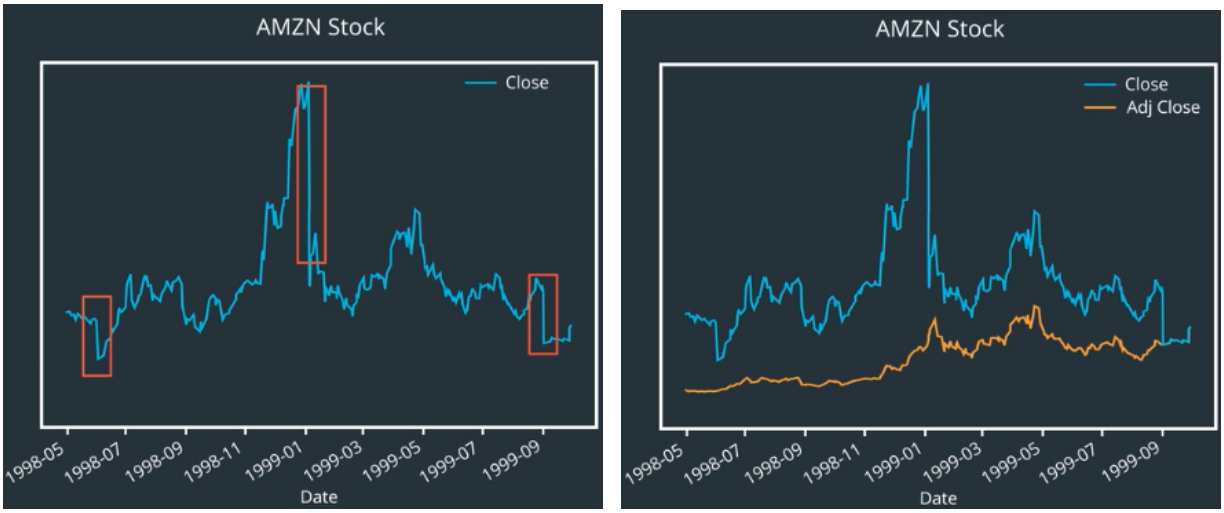
Date	Time	Ticker	Bid Size	Bid	Ask	Ask Size	Last Trade	Volume
07/19	09:59.2	MSFT	400	94.08	94.09	1000	94.06	100
07/19	09:59.3	MSFT	300	94.06	94.08	600	94.05	200
07/19	09:59.5	IBM	400	159.63	160.39	400	160	500
07/19	09:59.9	MSFT	700	94.06	94.07	1800	94.01	100
07/19	10:00.1	IBM	700	159.50	160.00	500	160.00	200
07/19	10:00.3	GOOG	300	1150.02	1152.00	100	1150.00	50
07/19	10:00.7	MSFT	500	94.08	94.09	700	94.10	200
07/19	10:00.8	GOOG	300	1151.00	1152.50	500	1150.20	20

Corporate Actions

Stock Splits

A company split its stock, after the split, the number of its stock will increase and the price per stock will decrease, and the total market capitalization has not changed. After the split, the expense of the stock decrease, the liquid will increase, and current shareholders and potential investors can buy and sell more granularity.

But stock price graph will distortion like fig (a) below, it looks like the total market capitalization decreased after the split because the figure doesn't consider the volume of stocks. To fix this problem, we usually compress the price before the split, for example, turn the price before the tree-to-one split into thirds. the price graph after adjustment as shown in fig (b).



(a) (b)

Dividend

Dividends are partial cash distribution of company earns, which means companies share some fraction of their profits with their shareholders.

In the figure below, it looks like the shareholder lost \$0.5 on both stocks, but in reality, the person made \$0.5 on company A because of the dividend of \$1. So we should adjust prices based on dividends to reflect this. We can normalize the prices, first, calculate the adjusted price factor $1 + D/S$, where D means dividend and S means the Stock price at the ex-dividend date. To normalize the price, we should divide the historical price by the Adjusted Price Factor up until the day before the ex-dividend date.

Liquidity