Context

This dataset is a playground for fundamental and technical analysis. It is said that 30% of traffic on stocks is already generated by machines, can trading be fully automated? If not, there is still a lot to learn from historical data.

Content

Dataset consists of following files:

* prices.csv: raw, as-is daily prices. Most of data spans from 2010 to the end 2016, for companies new on stock market date range is shorter. There have been approx. 140 stock splits in that time, this set doesn't account for that.
* prices-split-adjusted.csv: same as prices, but there have been added adjustments for splits.
* securities.csv: general description of each company with division on sectors
* fundamentals.csv: metrics extracted from annual SEC 10K fillings (2012-2016), should be enough to derive most of popular fundamental indicators.

Acknowledgements

Prices were fetched from Yahoo Finance, fundamentals are from Nasdaq Financials, extended by some fields from EDGAR SEC databases.

Inspiration

Here is couple of things one could try out with this data:

* One day ahead prediction: Linear Regression
* Momentum/Mean-Reversion Strategies
* Security clustering, portfolio construction/hedging

Which company has biggest chance of being bankrupt? Which one is undervalued (how prices behaved afterwards), what is Return on Investment?