Time Series Data

extraction, visualisation and analysis of securities

Luca Gennari

Elasticsearch Italy User Group Milan, Giugno 05, 2020

Original author: Luca Gennari - *Education Engineer EMEA*

Time Series Data

extraction, visualisation and analysis of stock market securities.

A time series is a series of data points indexed (or listed or graphed) in time order.

Most commonly, a time series is a sequence taken at successive equally spaced points in time.

Wikipedia

elastic

Time Series Data - Introduction

- Data associated with events where information refers to a moment in time..
  - logs, catalogs, metrics, stock market securities, production data, etc
  - they usually grow very fast as long as the event exists
- They must be *"denormalised"* and structured in the correct manner before an analysis can be performed.
  - denormalising is the standard way to indexing and storing data in a search engine or in any NoSQL database.

Source: elastic.co

Examples of information of type time series


```
{
  "runtime_ms" : 99,
  "method" : "GET",
  "host" : "server1",
  "geoip" : {
    "location" : {
      "lon" : -122.4128,
      "lat" : 37.7758
    },
    "country_name" : "United States",
    "country_code2" : "US",
    "continent_code" : "NA",
    "country_code3" : "US",
    "region_name" : "California",
    "city_name" : "San Francisco"
  },
  "level" : "info",
  "http_version" : "1.1",
  "@timestamp" : "2017-04-11T09:04:34.588Z",
  "status_code" : 200,
  "response_size" : 54975,
  "user_agent" : "Google Chrome"
}
```


Server Logs


```
{
  "date" : "2018-10-09",
  "high" : 68.0,
  "low" : 62.56999969482422,
  "open" : 65.58999633789061,
  "close" : 66.44999694824219,
  "volume" : 370400.0,
  "adj_close" : 66.44999694824219,
  "moving_average" : 67.46333312988281,
  "symbol" : "ESTC",
  "name" : "Elastic NV",
  "description" : "Elasticsearch is a search engine based on the
    Lucene library. It provides distributed, multitenant-capable
    full-text search engine with an HTTP web interface and schema
    -free JSON documents. Elasticsearch is developed in Java."
}
```


Stock Exchange Trading Info

Source: elastic.co

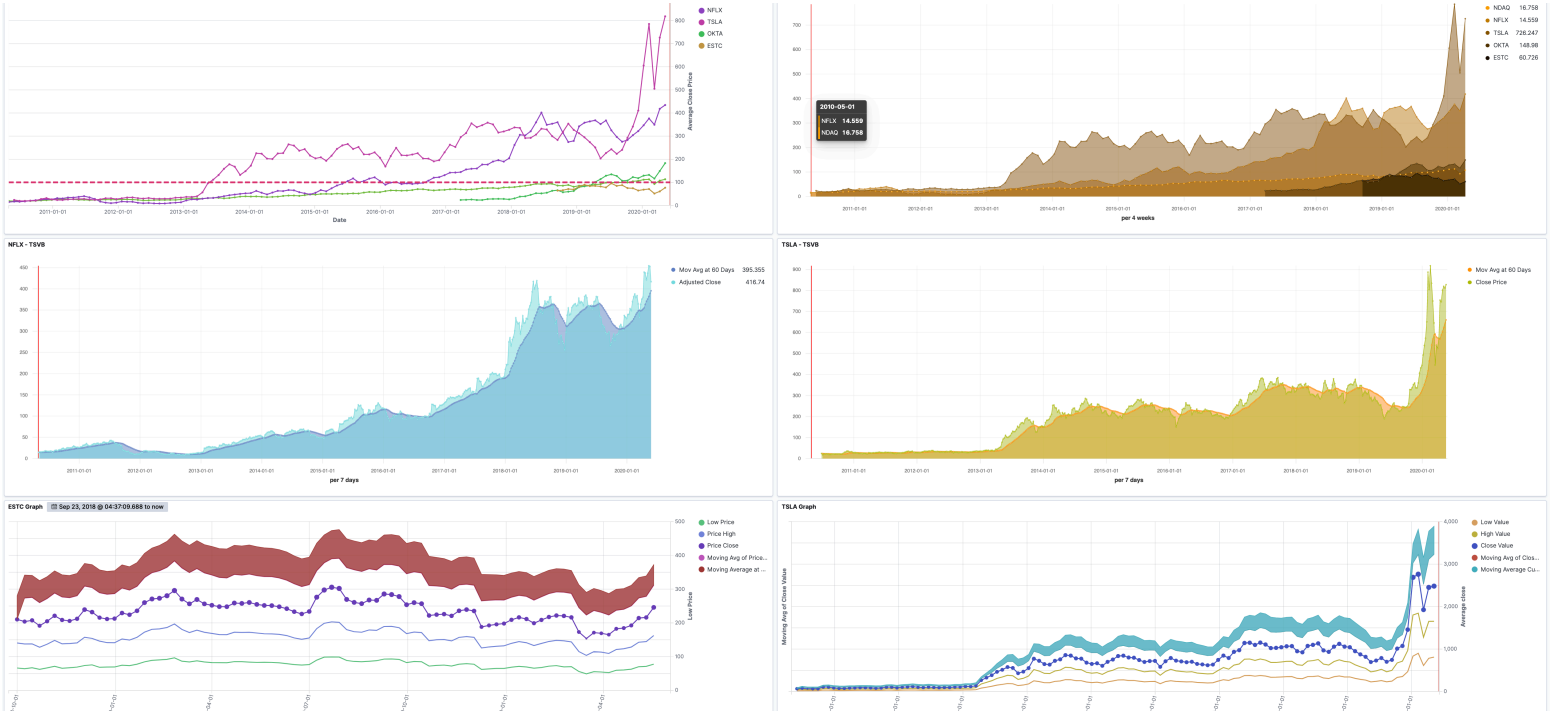
Time Series in Elasticsearch

Indexing documents that represent time series data in Elasticsearch is simple but requires that fields are correctly mapped.

Kibana natively has tools that allow you to analyse Time Series data in all forms such as TSVB (Time Series Visual Builder)

Let's see some...

Standard visualisations of time series...

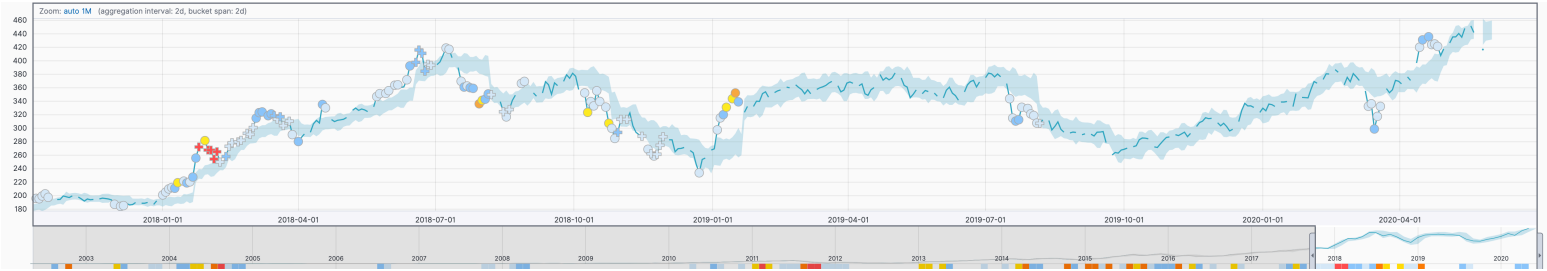


Dashboard with TSVB and Line Graphs

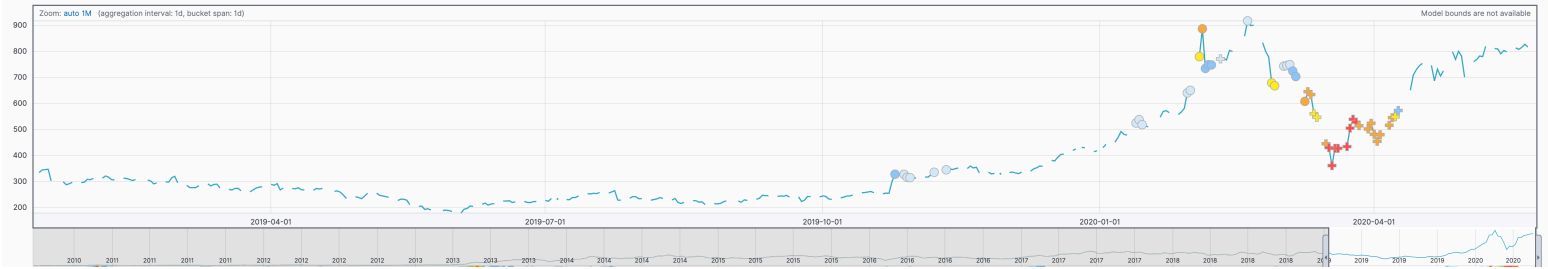
Source: elastic.co

elastic

...and with Machine Learning features.



Machine Learning Analytics



Machine Learning Anomaly Detection

Source: elastic.co

elastic

Dear audience, here you go...
Time Series Data

DEFINITION

- Time series are a collection of observations, points or data collected at specific regular intervals.
- They can be used to predict future values based on previous observations.
- Time series must have only one variable: the time.

REPRESENTATION

- Time series are very often plotted through line graphs.
- Other information can be overlapped to find correlations.
- They are indispensable in the search for patterns (graphical or numerical).

USAGE

- Statistic analysis.
- Exchange forecast.
- Signal processing.
- Models recognition.
- Weather or earthquake forecast.
- Astronomy and in any sector of applied science and engineering.
- IoT and industrial applications.

Kibana TSVB
Time Series Visual Builder

TSVB - Introduction

- ✦ It combines a wide range of aggregations, including pipeline aggregations.
  - ✓ allows to analyse data in different ways
- ✦ Supports different types of customisations on data views
  - ✓ background colour
  - ✓ information displayed on different axes
  - ✓ overlay of aggregate data
  - ✓ different chart types in one single view
- ✦ Supports multiple index patterns
  - ✓ allows to evaluate the differences between different information
- ✦ Graphical annotations can be created
  - ✓ makes it easier to spot anomalies or points of interest.

Source: elastic.co

TSVB - Features

- ✦ Supported aggregations
  - ✓ mathematical operations (max, min, sum, etc...)
  - ✓ static values
  - ✓ overall highs and lows
  - ✓ aggregations on aggregate data (pipeline aggregations)
  - ✓ moving average, standard deviation, percentile, sum of squares, variance, etc...
  - ✓ ...and much more
- ✦ Offers more flexibility than standard charts
  - ✓ time series shifting
  - ✓ cloning of existing series
  - ✓ different styles of visualisations
  - ✓ comparisons of different data from different indices

Source: elastic.co

Components of a time series analysis

- Direction in which the information is directed

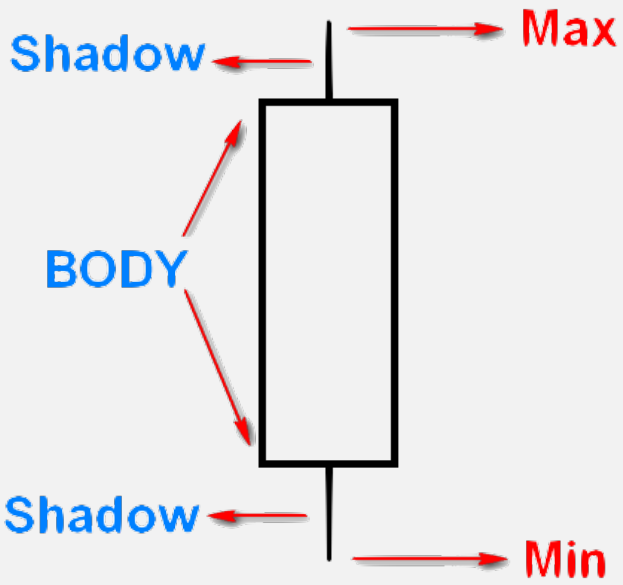
- Patterns that data can assume in particular periods

- Changes of patterns that do not follow seasonality or cyclicity

- Repetition of patterns with constant character and independent of seasonality

Data of a stock security

- Generally, for every single trading day you have five basic information:
  - open price
  - close price
  - minimum achieved by the price
  - maximum reached by the price
  - closure adjustment
- Another not basic but useful data is the volume exchanged
  - refers to the amount of exchanged contracts, both for sale and for purchase
  - provides an idea of the market interest on a particular security (high volumes, high interest and vice versa)
- Data generally not provided or which are not part of the bargaining, however can be very useful for analysis purposes:
  - Moving average. The most used:
    - *30, 60 or 90 days*
  - Earnings dates
    - *forecast and expected data from earning announcement (generally released a few weeks before the announcement date)*
    - *real data (released at the time of the announcement and generally used for the next forecast)*



Japanese candlestick

Document structure

- A date field (in this case we are only interested in the date, not the time)
- There are six numerical fields:
  - ***Low*** - minimum price reached
    - ***type : float***
  - ***High*** - maximum price reached
    - ***type : float***
  - ***Open*** - opening price
    - ***type : float***
  - ***Close*** - close price
    - ***type : float***
  - ***Adj_close*** - closing price adjustment
    - ***type : float***
  - ***Volume*** - contracts exchanged
    - ***type : Integer***
- A field generated in pre-analysis which is the 60-day moving average.
- Descriptive fields:
  - symbol of the ticker
  - name of the underlying
  - description of the underlying


```
{
  "_index" : "indeces_combined",
  "_type" : "_doc",
  "_id" : "mfgmdHIBzFEP8t6vM-TW",
  "_score" : 1.0,
  "_source" : {
    "date" : "2018-10-09",
    "high" : 68.0,
    "low" : 62.56999969482422,
    "open" : 65.58999633789061,
    "close" : 66.44999694824219,
    "volume" : 370400.0,
    "adj_close" : 66.44999694824219,
    "moving_average" : 67.46333312988281,
    "symbol" : "ESTC",
    "name" : "Elastic NV",
    "description" : "Elasticsearch is a search..."
  }
}
```


Indexed document structure

elastic

DEMO TIME

Legal Information Notice

This presentation has been prepared for ***information purposes only***.

Neither Elastic nor the author can guarantee the accuracy of the data or information provided, when applied in specific circumstances.

If any individual or entity chooses to utilise the information or data provided in this presentation, they do so at their own risk.

No liability will be accepted by Elastic or the author for any claims, losses, damages or costs (arising directly or indirectly) howsoever caused by any reliance on the data or information provided in this presentation.

Source: elastic.co

Script and data

Material used in this presentation is freely downloadable and usable without restriction, at the following github address:

https://github.com/lucagennari-es/time-series-data-meetup.git

GitHub page:

https://github.com/lucagennari-es/time-series-data-meetup

Source: elastic.co

Q&A