

Introduction to InfluxDB

Michael Desa & Jack Zampolin



By the end of this section, participants will be able to...

1. Describe what time series data is and recognize its use cases.
2. Explain the differences between InfluxDB and other TSDBs.
3. Describe the InfluxDB Data model.
4. Write data into InfluxDB using the InfluxDB CLI.

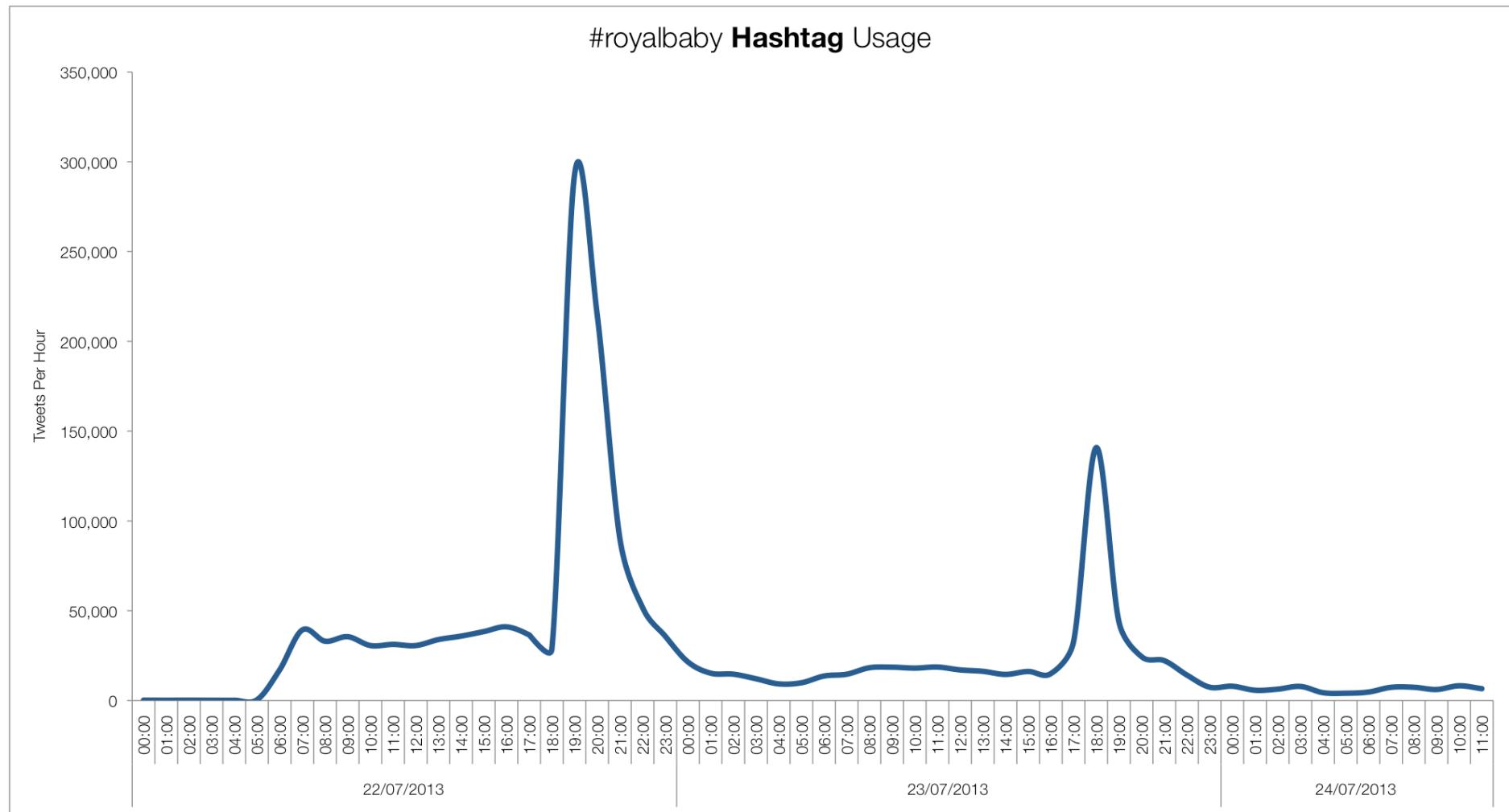
What is time series data?

Time series data is...

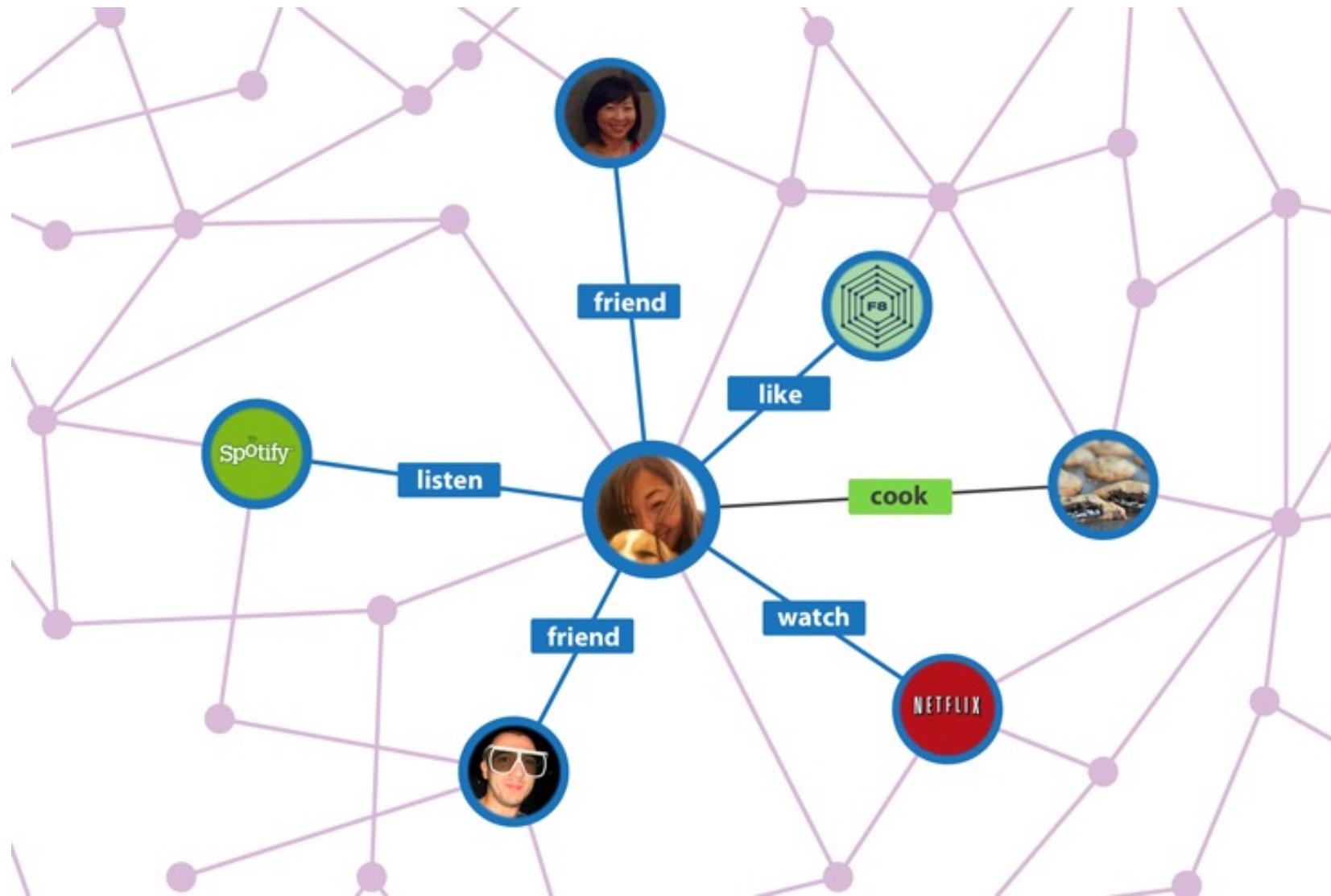
A time series is a sequence of data points, typically consisting of successive measurements made from the same source over a time interval.

What this essentially means is that if you were to plot the points on a graph, one of your axes would always be time.

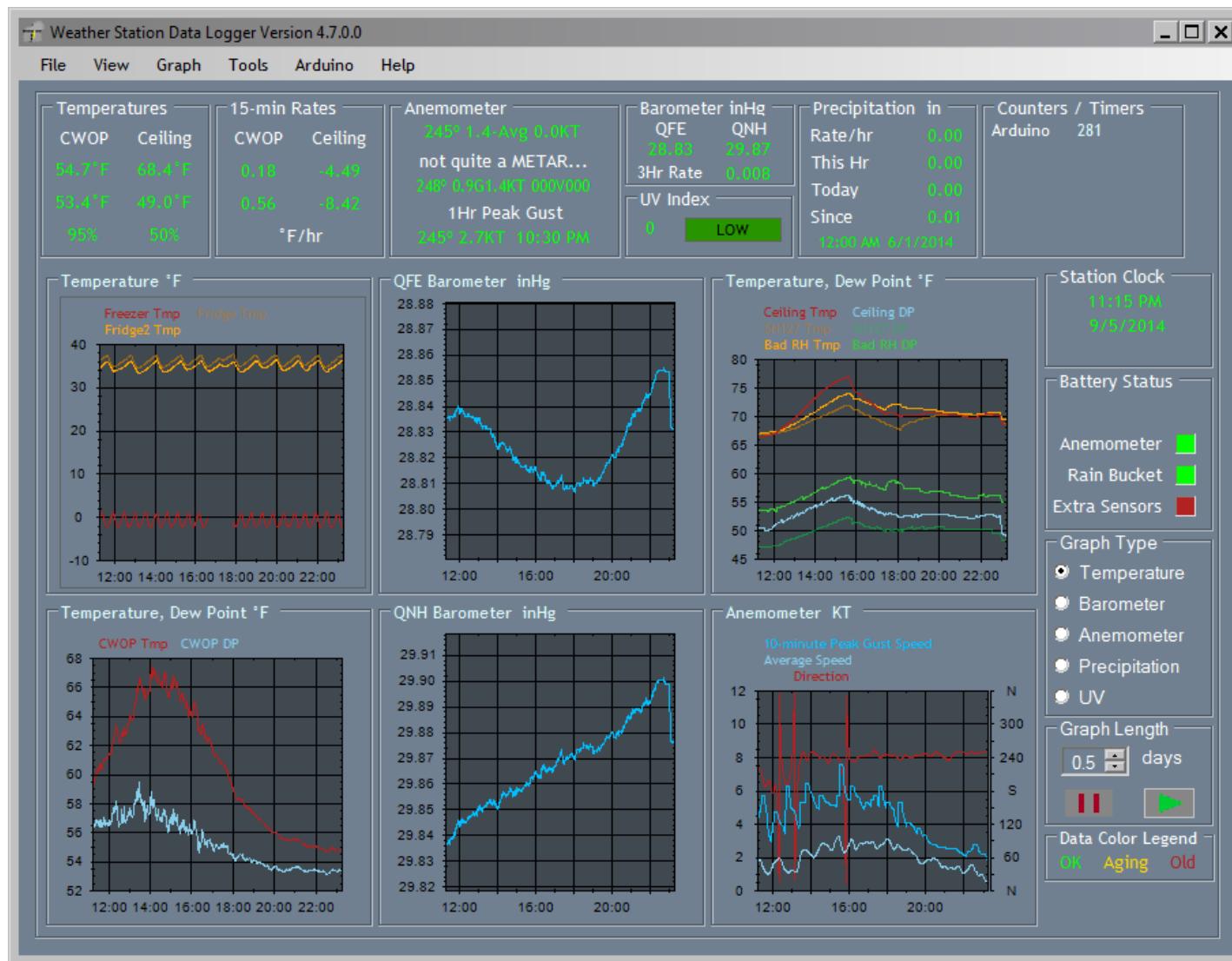
Time series data



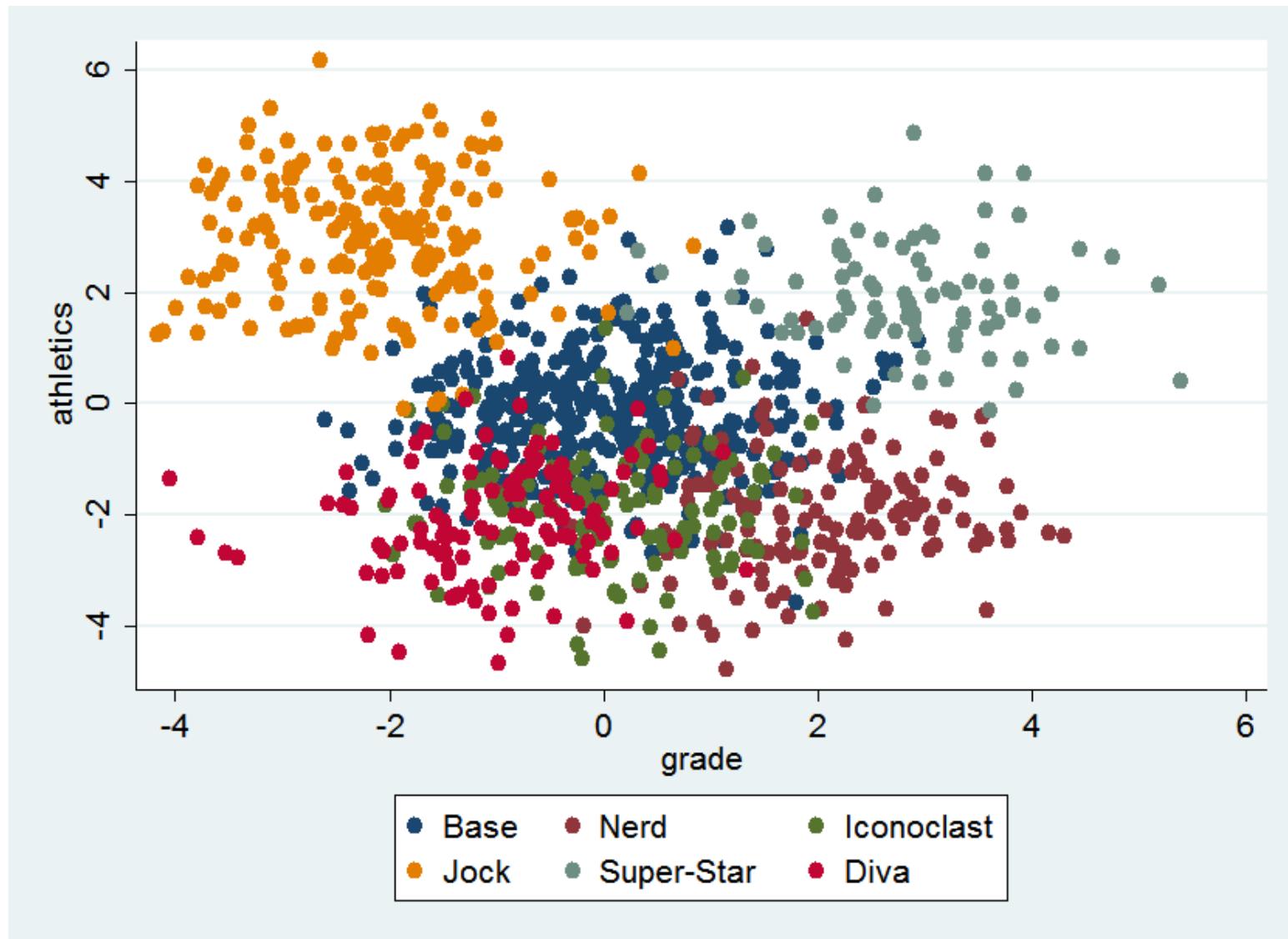
Not time series data



Time series data



Not time series data

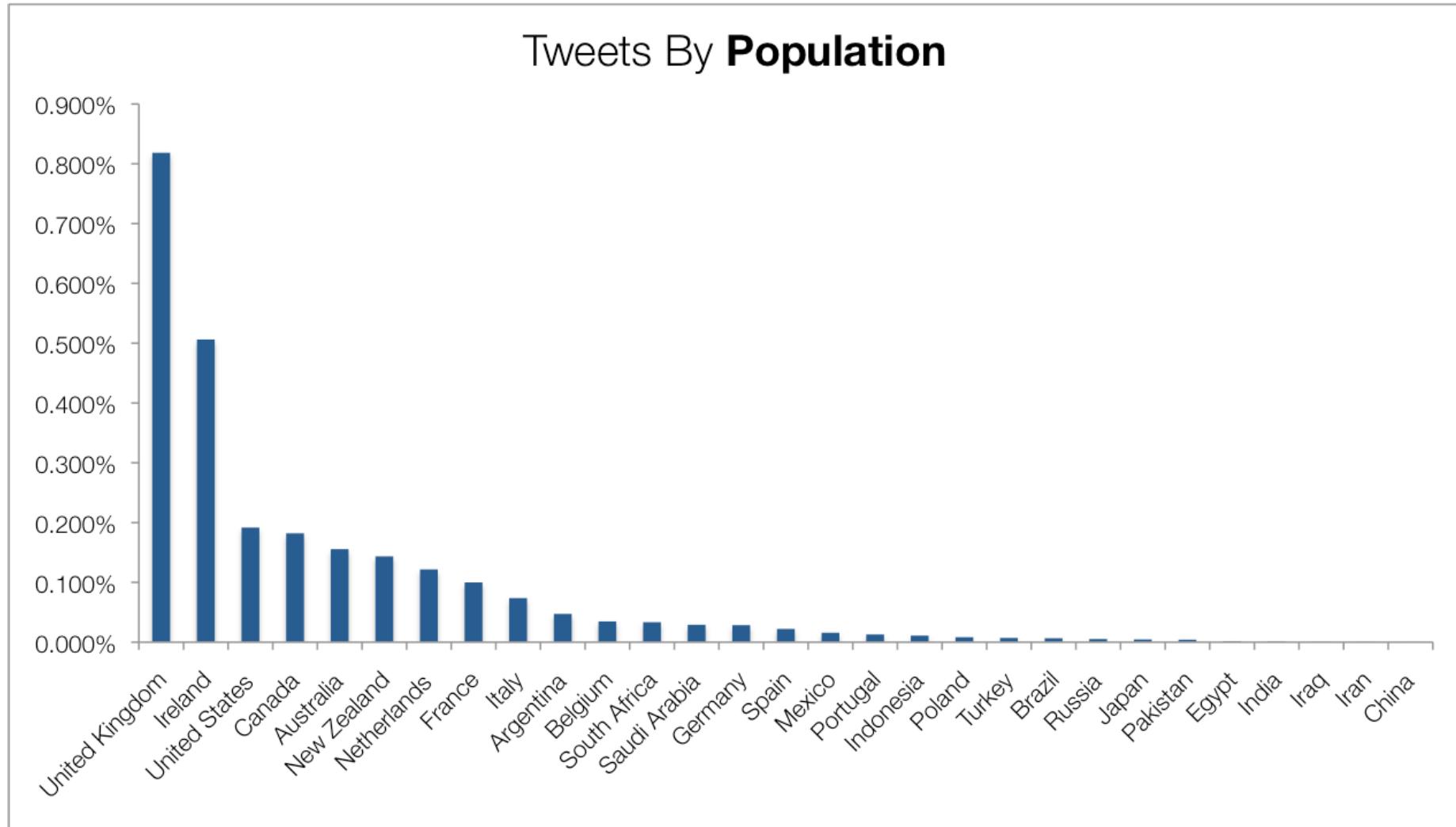


Time series data

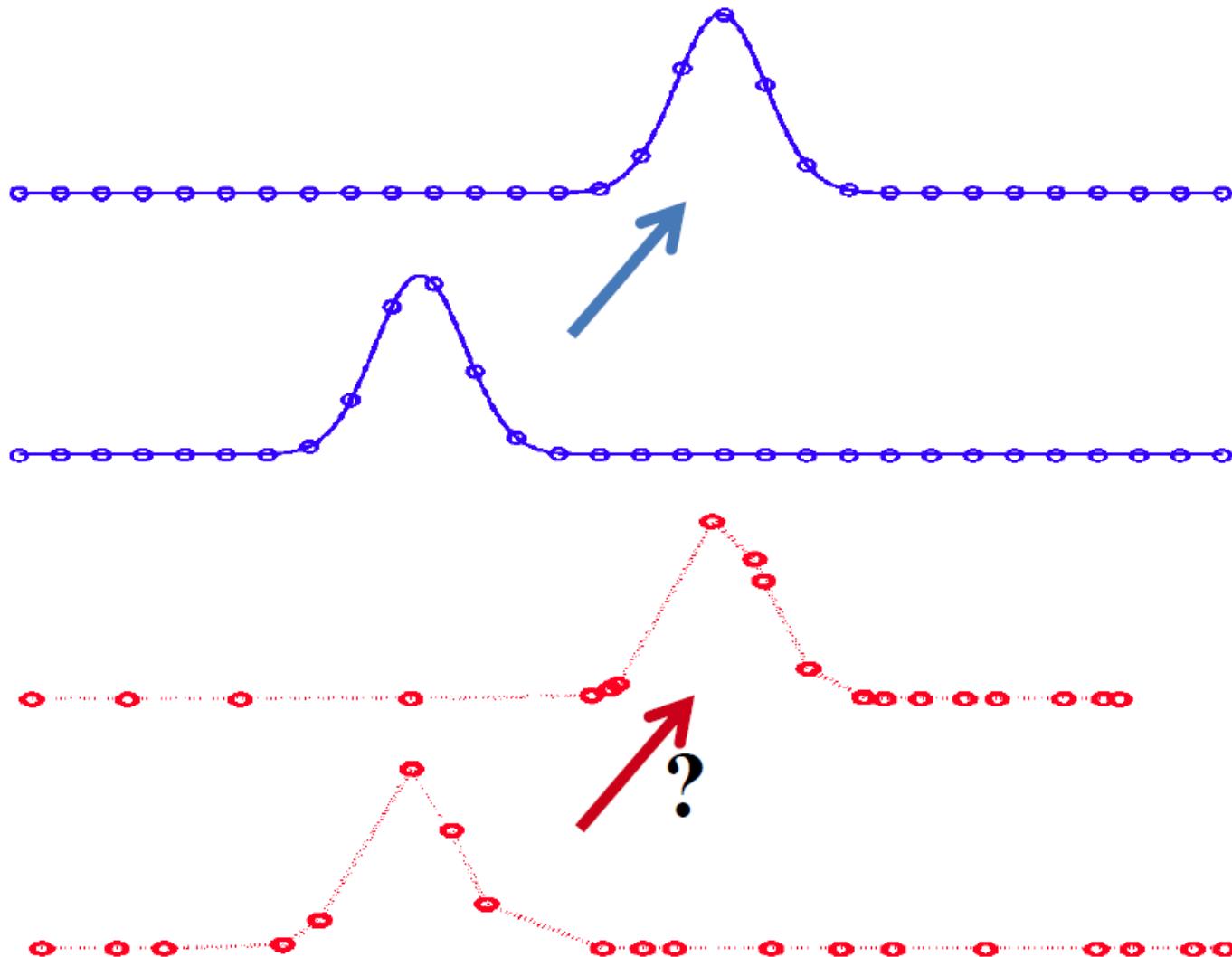


influxdata

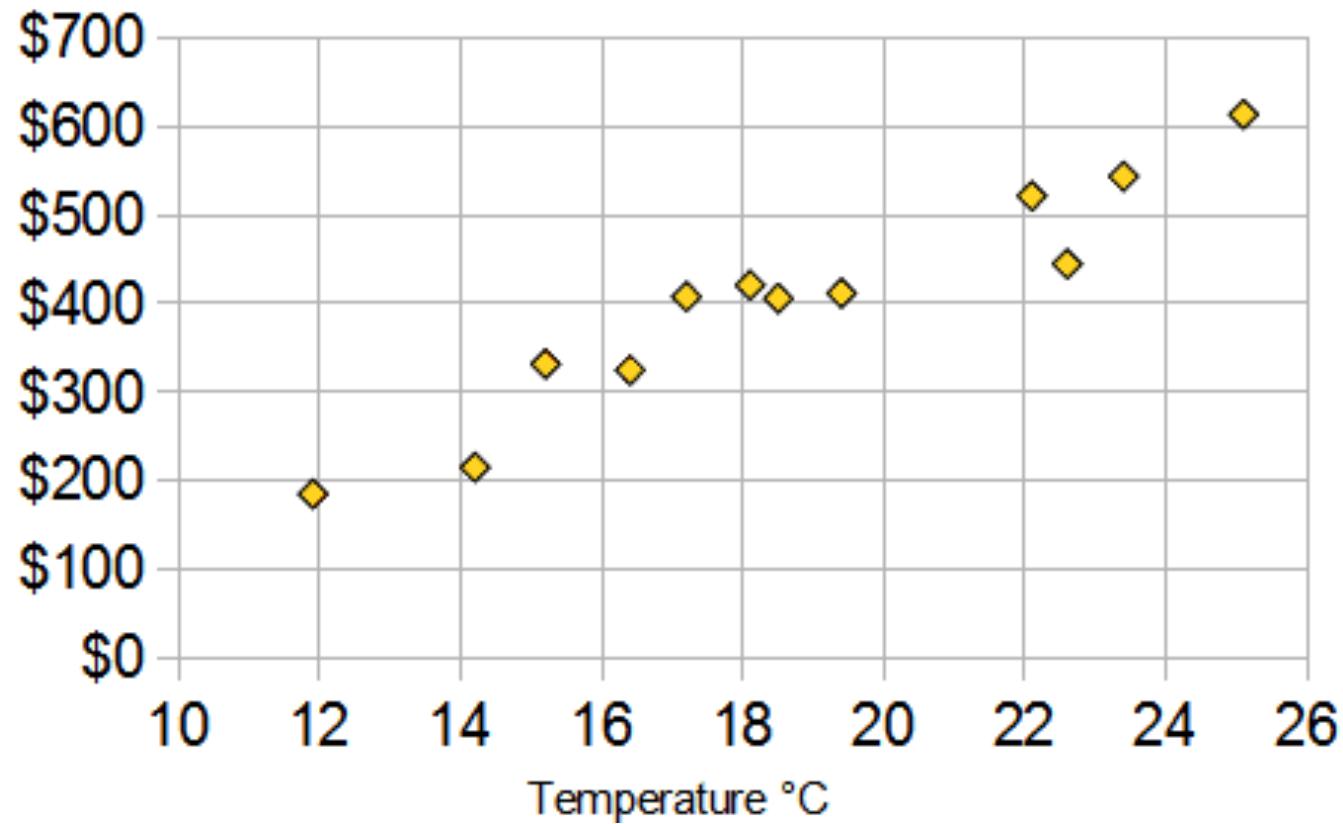
Not time series data



Regular vs Irregular Timeseries



Exercise



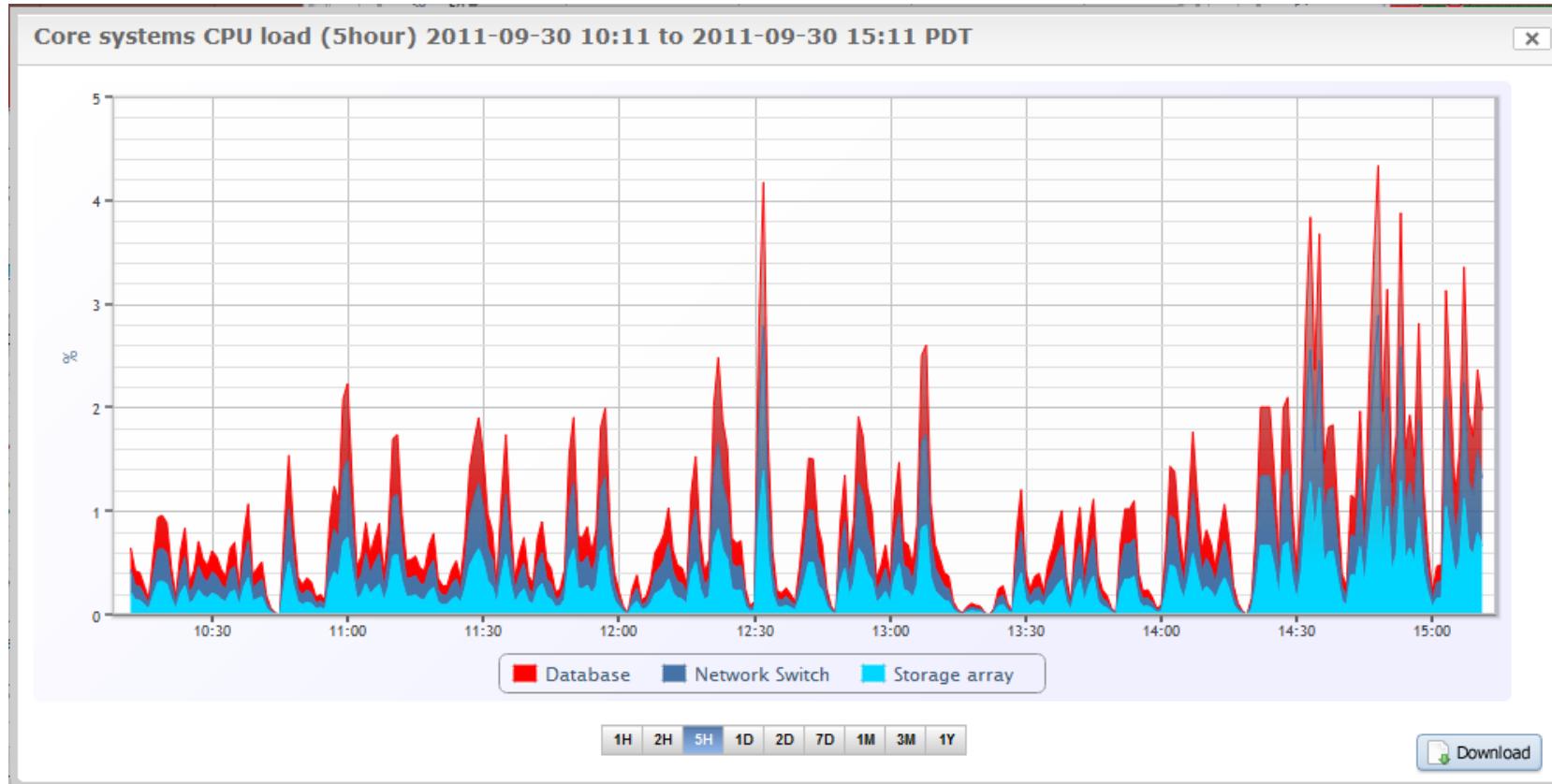
Is it time series?

Exercise



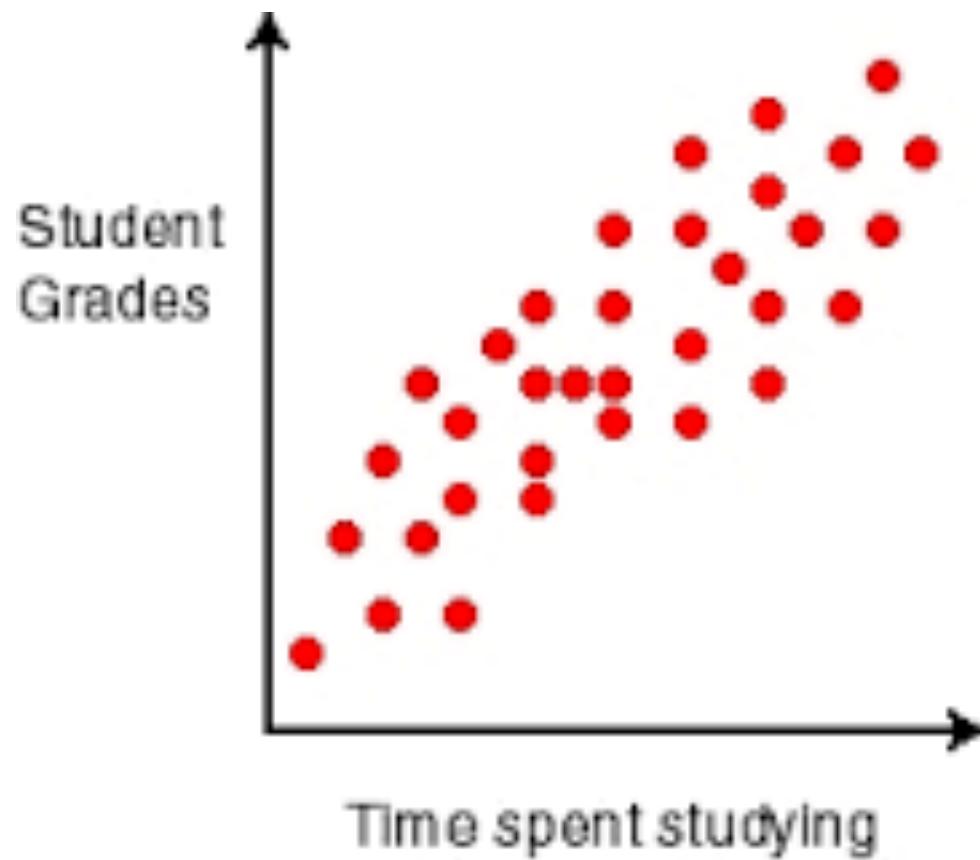
Is it time series?

Exercise



Is it time series?

Exercise



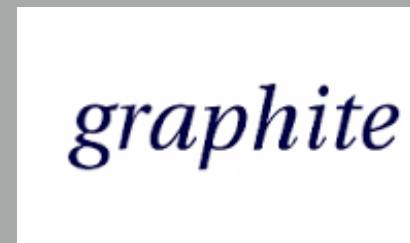
What is time series database?

Time series database is...

A database where you manage and store time series data...

Other TSDB's

- KDB+
- OpenTSDB
- Graphite
- Riak TS



Why cant I just use [your favorite db]?

You could...

...but you would be implementing a lot of the functionality that a time-series database comes with by default.

Why Choose InfluxDB?



- Easy to get started with
- Familiar query syntax
- No external dependencies
- Allows for regular and irregular time series
- Horizontally scalable
- Member of a cohesive time series platform

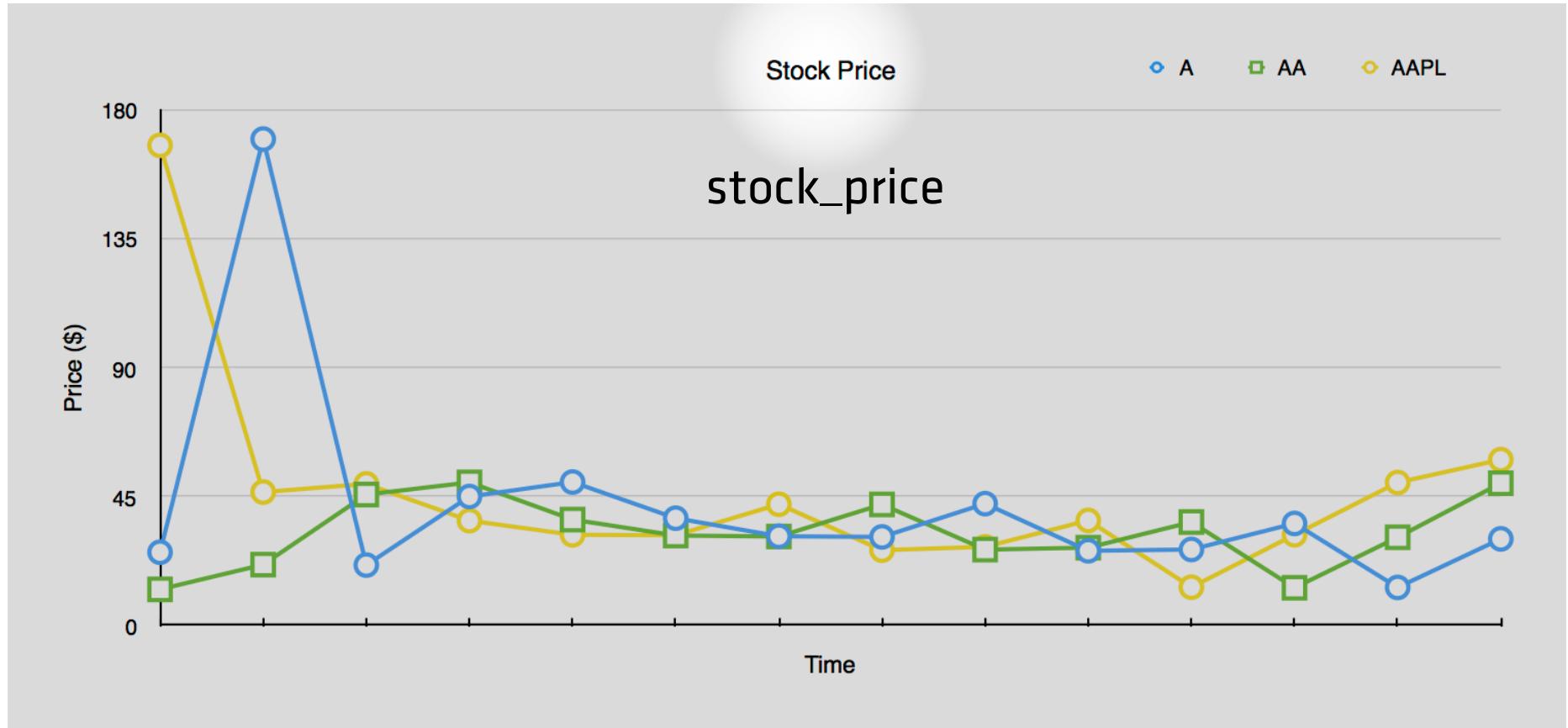
The InfluxDB Data Model

A typical time series graph



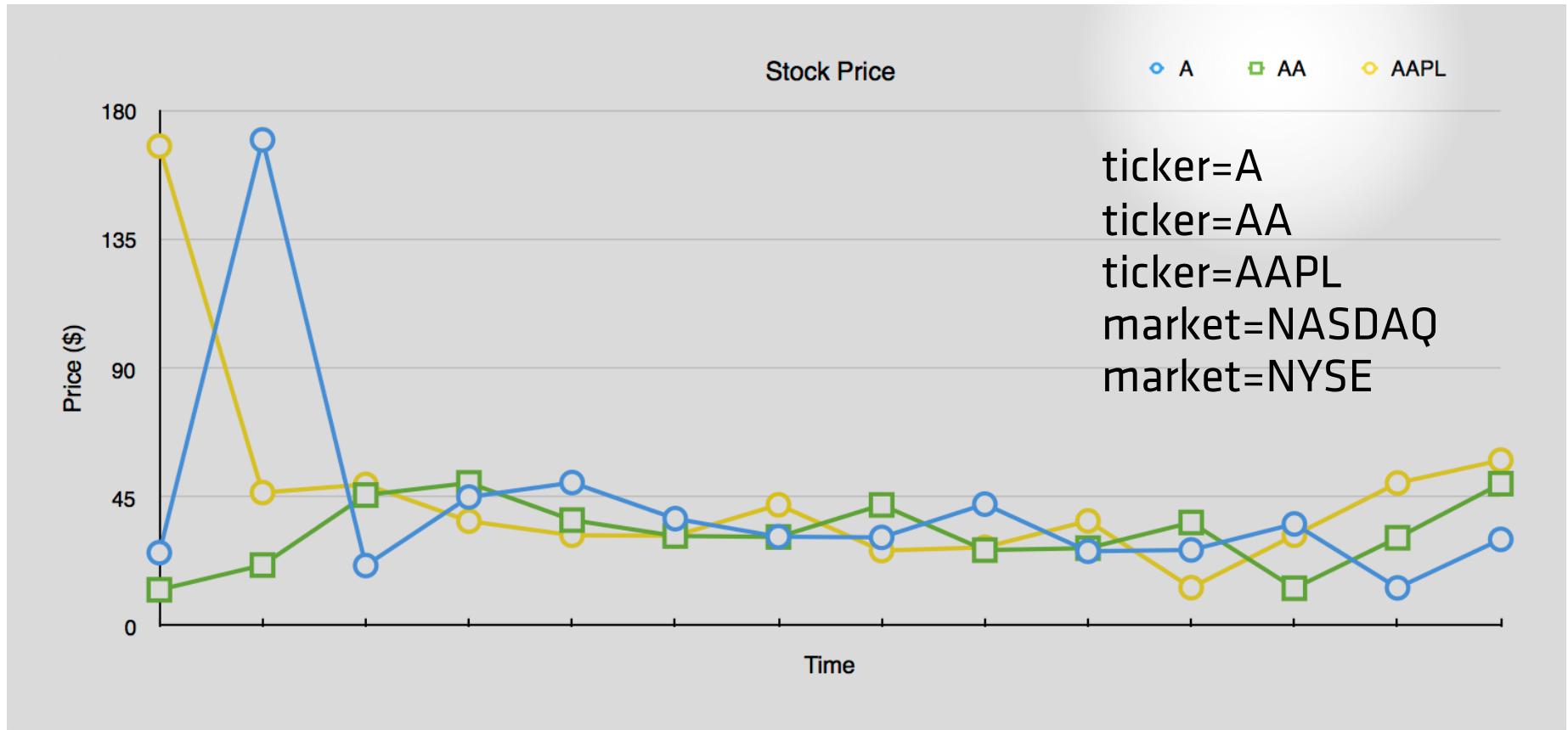
What components do we have?

The Label



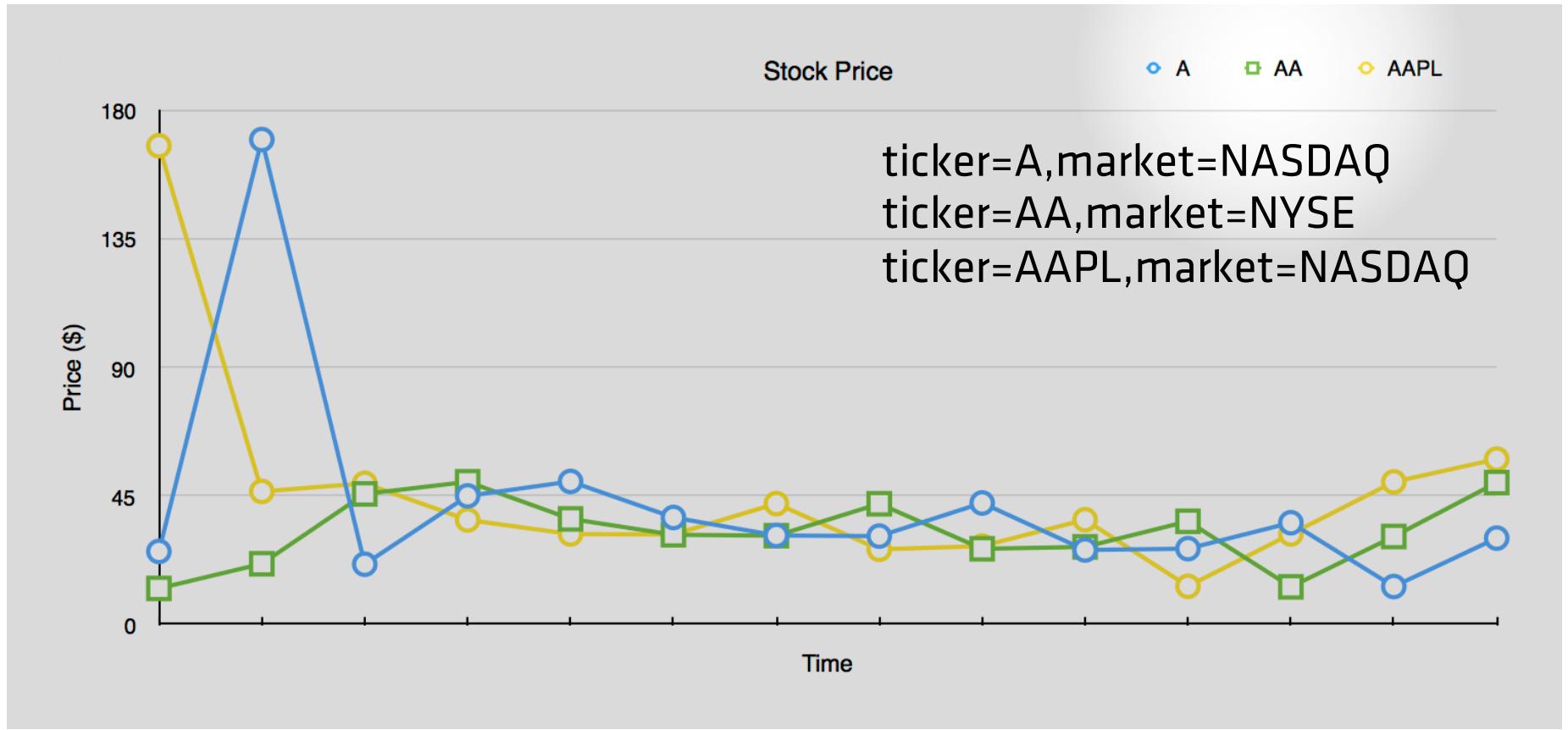
We call this the measurement

The Legend (metadata)



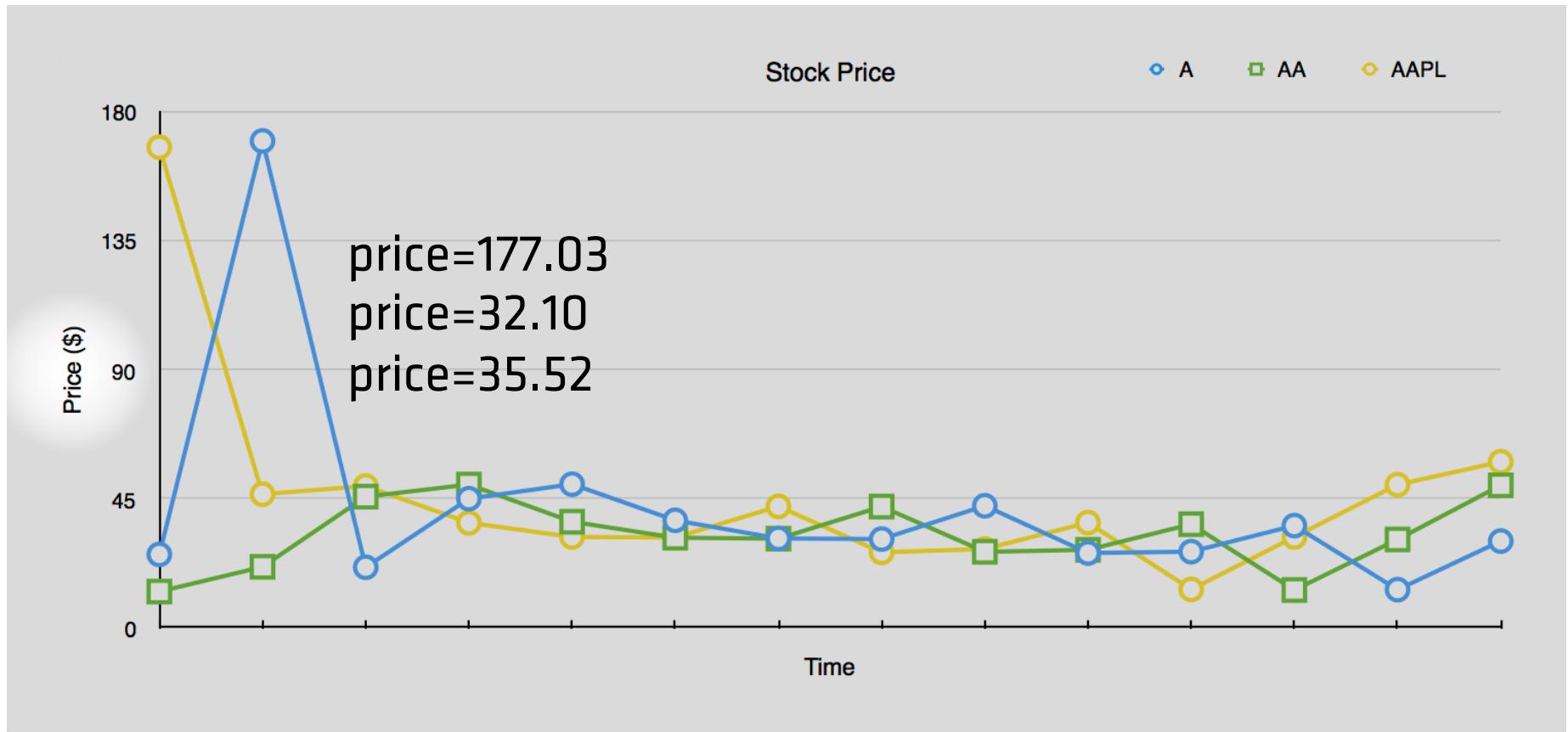
We call these tags.
Tags are indexed.

Collection of all tags



We call this the tagset.

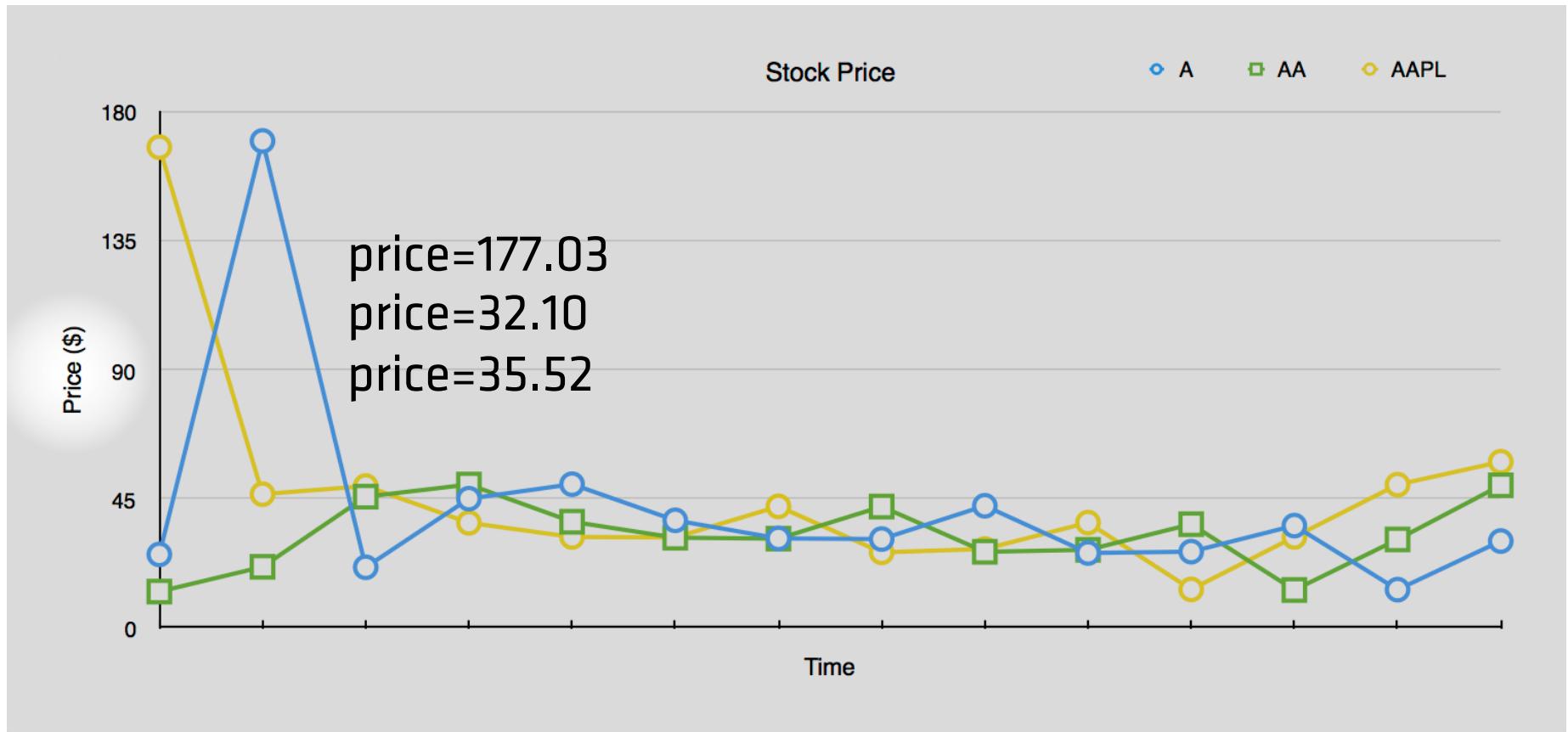
Y-Axis Values



We call these fields.

Note that the values that the field stores can be floats, ints, strings, or bools.

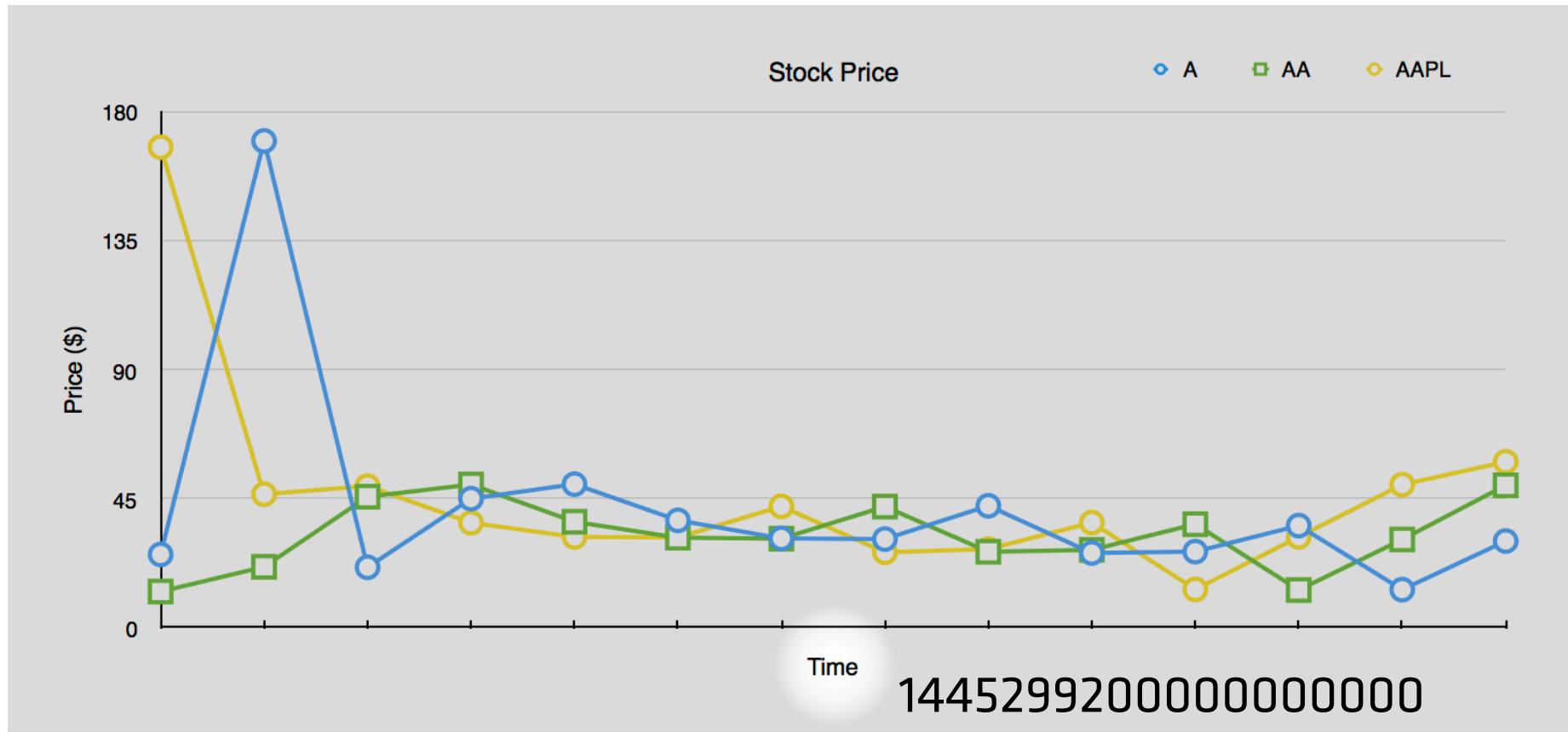
The Collection of Fields



We call the fieldset.

Note that in this case, there's only one field.
(There could be many.)

X-Axis Value



Exercise



How do we represent points textually?

The Line protocol

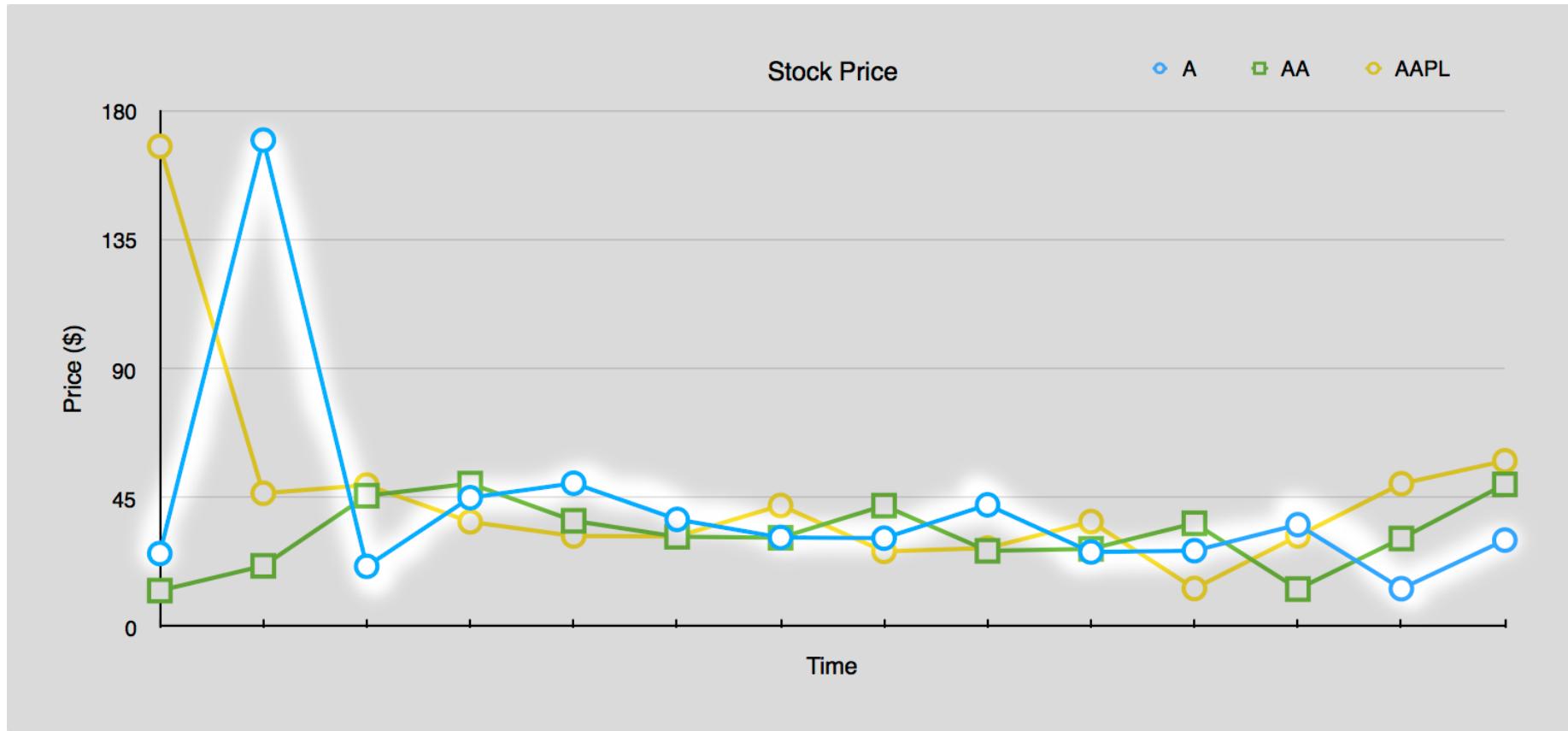
measurement, tagset fieldset timestamp

```
stock_price, ticker=A, market=NASDAQ price=177.03 14452992000000000
```

Points in InfluxDB look like...



A Series in InfluxDB



measurement + tagset = the series as a whole

measurement + tagset + timestamp = single point

Examples of points in Line Protocol

- `cpu,host=server1 value=100 14452992000000000000`
- `temperature,zipcode=94107,country=usa value=75,humidity=10`
- `response_time,method=GET,precision=ms value=12i 14452992000000000000`

Exercise

Do the following points belong to the same series?

stock_price,ticker=NN value=100 14452992000000000000

stock_price,ticker=NN other=100 14452992000000000000

stock_price,ticker=NN value=100,other=100 14452992000000000000

Exercise

Express a point with measurement

stock_price

with one tag

ticker=NN

that has a single field

value=10

without a timestamp.

Exercise

Express a point with measurement

`stock_price`

with two tags

`ticker=AA`

and

`exchange=NASDAQ`

with a single field

`value=10`

with a timestamp.

Exercise

Express a point under the same measurement as the example below, but belonging to a different series.

stock_price,ticker=AAPL,market=NASDAQ value=1032 1445299200

Writing Data into InfluxDB

Creating a database

```
CREATE DATABASE mydb
```

```
Last login: Mon Oct 19 10:50:43 on ttys006
~$ influx
Connected to http://localhost:8086 version 0.9
InfluxDB shell 0.9
> create database mydb
> █
```

Verifying that it was created

SHOW DATABASES

```
> show databases
name: databases
-----
name
_internal
mydb
```

Using the database we just created

```
USE mydb
```

```
> use mydb
Using database mydb
> |
```

Inserting data into the database

```
insert cpu,host=server1,location=us-west value=10  
insert cpu,host=server1,location=london value=11  
insert cpu,host=server2,location=us-west value=12
```

```
> insert cpu,host=server1,location=us-west value=10  
> insert cpu,host=server1,location=london value=11  
> insert cpu,host=server2,location=us-west value=12  
> |
```

Verifying that the data was written

```
SELECT * FROM cpu
```

PS. Be careful! This query can be very expensive.

```
SHOW SERIES
```

```
SHOW MEASUREMENTS
```

```
> select * from cpu
name: cpu
-----
time          host    location   value
1445282634342214207  server1 us-west   10
1445282643540349747  server1 london    11
1445282651029129406  server2 us-west   12
```