



*influx*/days

# Data-driven Decision Making

---

Emanuele Della Valle

Prof. @ Politecnico di Milano & Partner  
@ Quantia Consulting

Marco Balduini

Founder & CEO @ Quantia Consulting

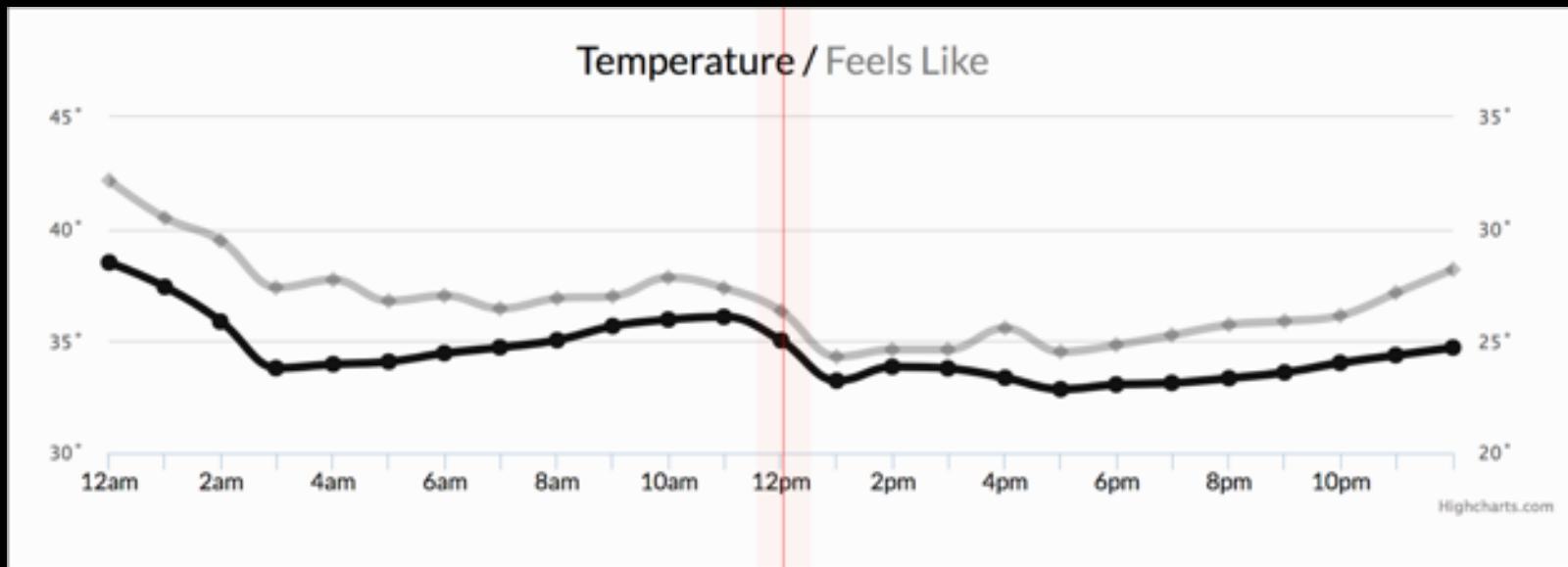


# What are they?

## Time Series

# Let's start by example

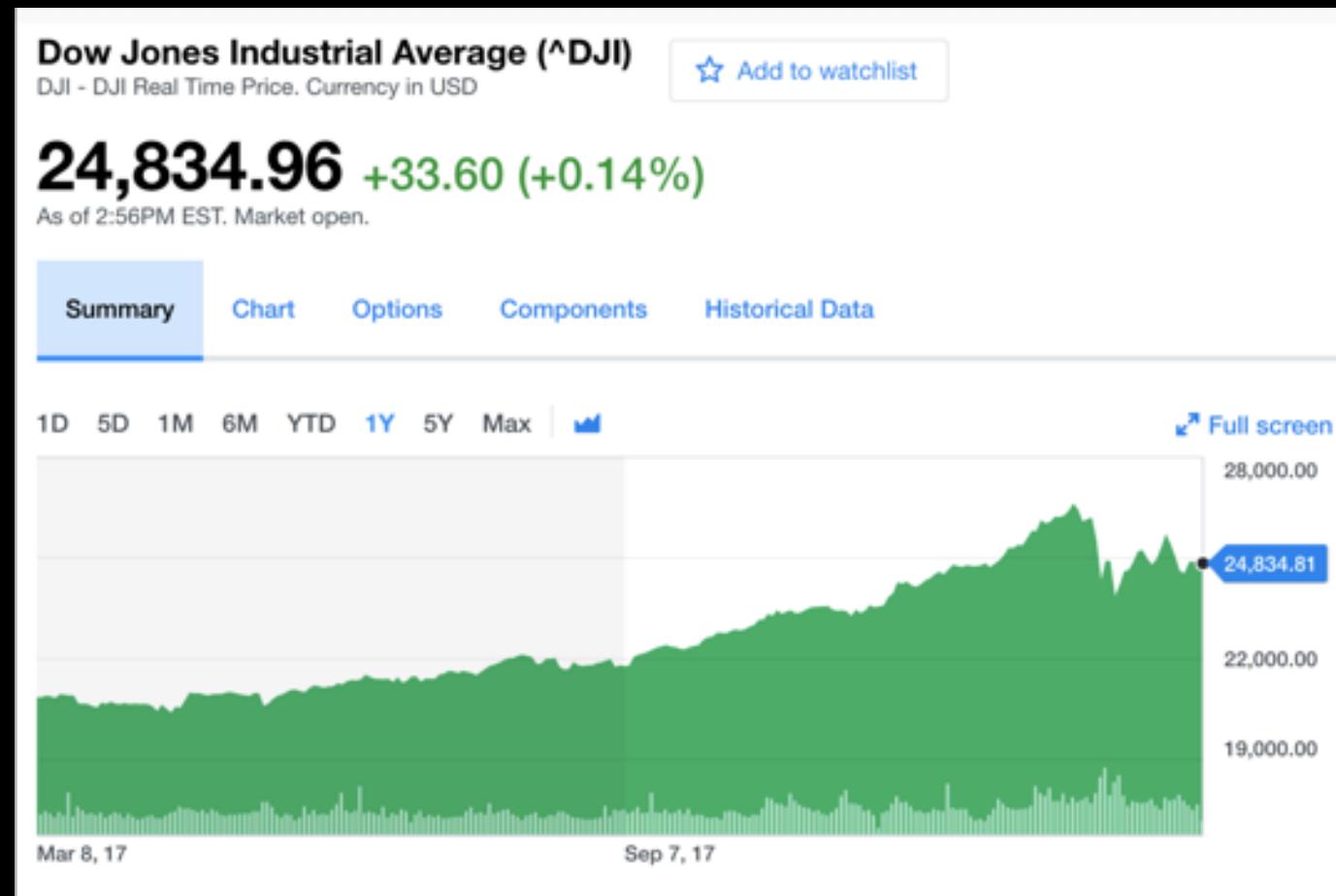
Weather conditions



# Let's start by example

## Weather conditions

# Stock exchange

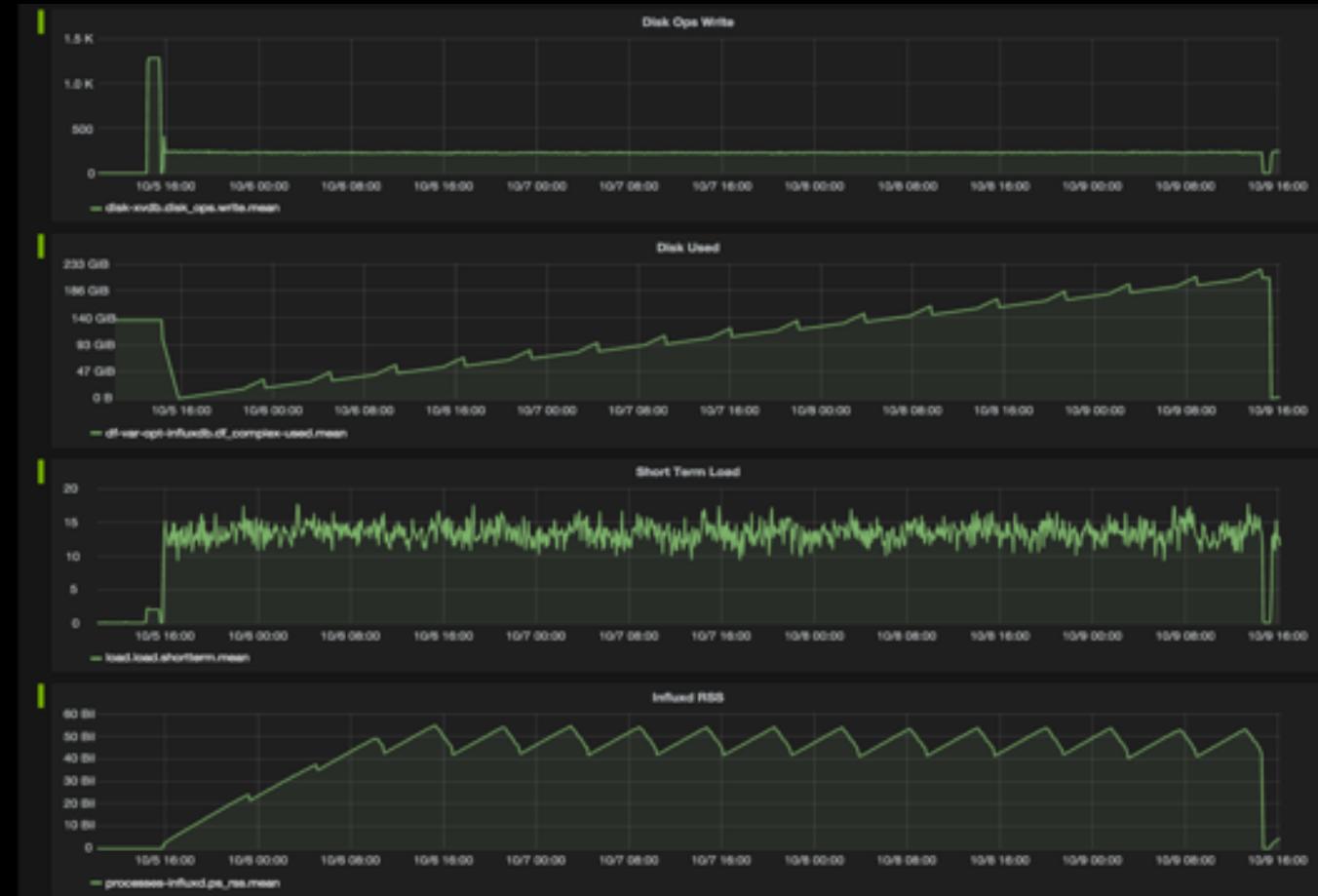


# Let's start by example

Weather conditions

Stock exchange

Cluster monitoring



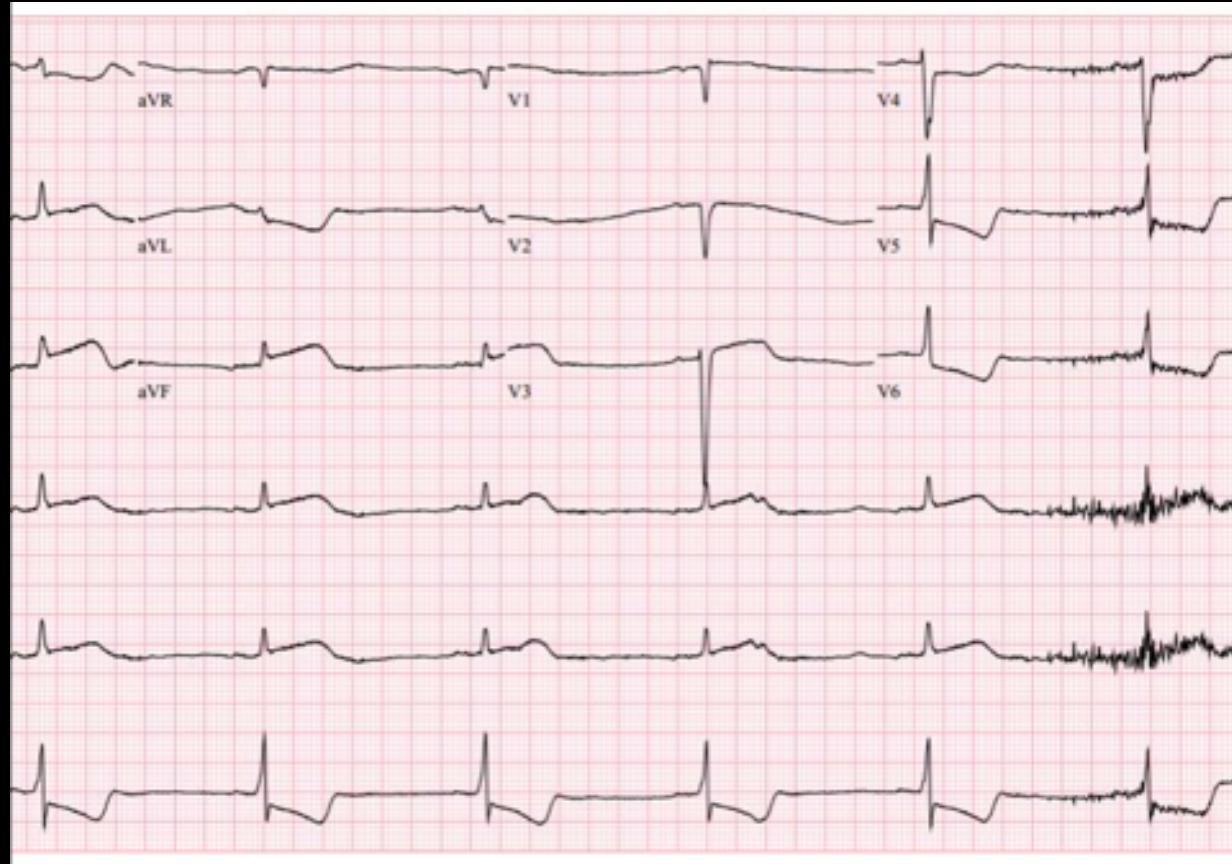
# Let's start by example

Weather conditions

Stock exchange

Cluster monitoring

Healthcare



# Now something different

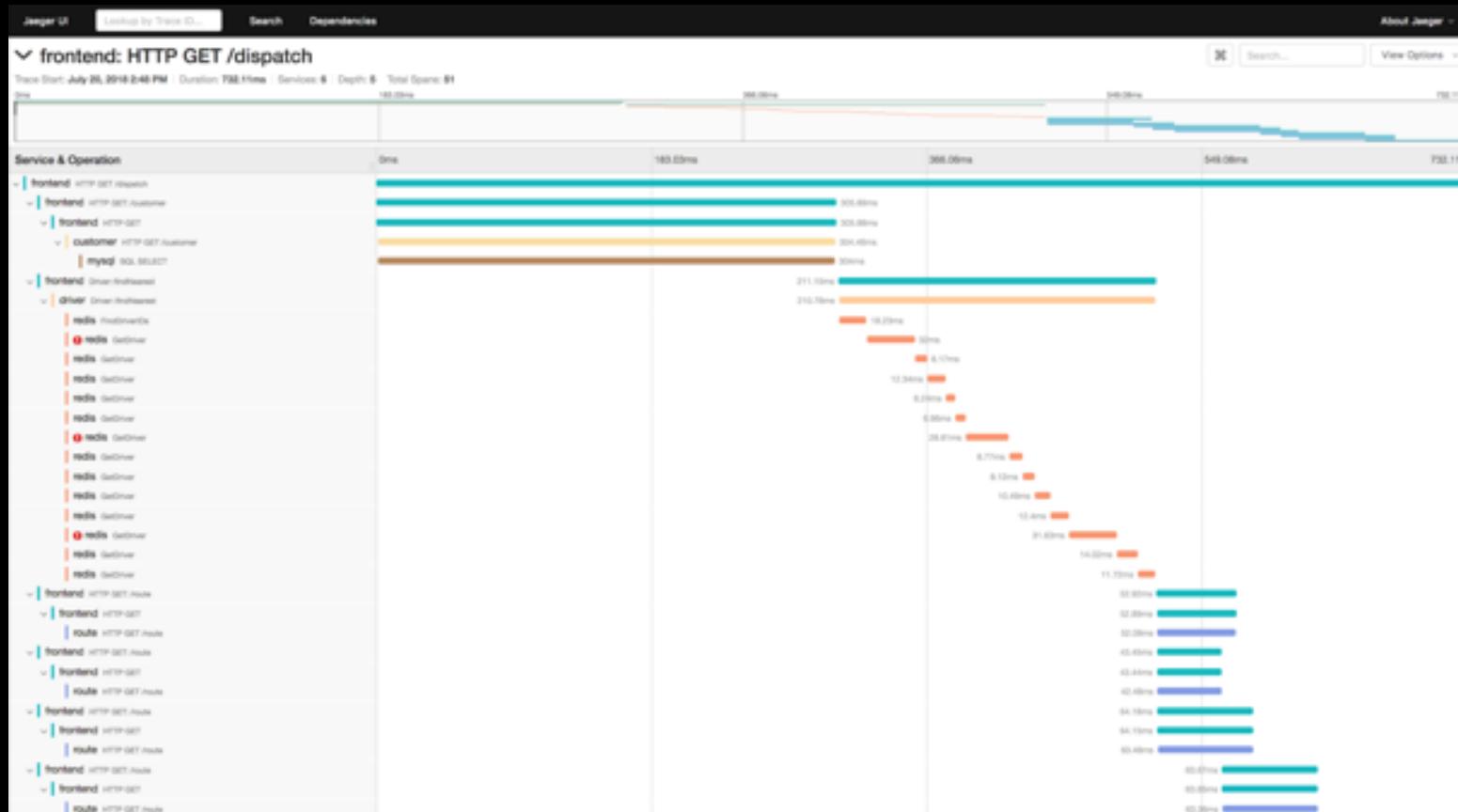
## Logs

```
Jun 24 13:45:36 haproxy epo-http.txt: dd14-034.compuserve.com [30:01:46:50] "GET /logos/small_gopher.gif HTTP/1.0" 200 935
Jun 24 13:45:38 haproxy epo-http.txt: dd14-034.compuserve.com [30:01:46:54] "GET /logos/small_ftp.gif HTTP/1.0" 200 124
Jun 24 13:45:40 haproxy epo-http.txt: ix-eve-wa2-02.ix.netcom.com [30:01:46:55] "GET /docs/EPA-WASTE/1994/October/Day-05 HTTP/1.0" 302 -
Jun 24 13:45:40 haproxy epo-http.txt: dd14-034.compuserve.com [30:01:46:56] "GET /icons/book.gif HTTP/1.0" 200 156
Jun 24 13:45:41 haproxy epo-http.txt: ix-eve-wa2-02.ix.netcom.com [30:01:46:56] "GET /EPA-WASTE/1994/October/Day-05/ HTTP/1.0" 200 623
Jun 24 13:45:42 haproxy epo-http.txt: dd14-034.compuserve.com [30:01:46:58] "GET /logos/us-flag.gif HTTP/1.0" 200 2788
Jun 24 13:45:43 haproxy epo-http.txt: ix-eve-wa2-02.ix.netcom.com [30:01:47:12] "GET /docs/EPA-WASTE/1994/October/Day-03 HTTP/1.0" 302 -
Jun 24 13:45:45 haproxy epo-http.txt: ix-eve-wa2-02.ix.netcom.com [30:01:47:14] "GET /EPA-WASTE/1994/October/Day-03/ HTTP/1.0" 200 785
Jun 24 13:45:46 haproxy epo-http.txt: dd14-034.compuserve.com [30:01:47:19] "GET /icons/ok2-0.gif HTTP/1.0" 200 231
Jun 24 13:45:48 haproxy epo-http.txt: bettong.client.uq.oz.au [30:01:47:24] "GET /enviro/html/emci/emci_overview.html HTTP/1.0" 200 2352
Jun 24 13:45:49 haproxy epo-http.txt: bettong.client.uq.oz.au [30:01:47:31] "GET /enviro/gif/efacts.gif HTTP/1.0" 200 1367
Jun 24 13:45:50 haproxy epo-http.txt: 202.96.29.111 [30:01:47:34] "GET /PressReleases/ HTTP/1.0" 200 1241
Jun 24 13:45:51 haproxy epo-http.txt: bettong.client.uq.oz.au [30:01:47:37] "GET /enviro/gif/blueball.gif HTTP/1.0" 200 903
Jun 24 13:45:53 haproxy epo-http.txt: ix-eve-wa2-02.ix.netcom.com [30:01:47:37] "GET /Rules.html HTTP/1.0" 200 3273
Jun 24 13:45:53 haproxy epo-http.txt: 202.96.29.111 [30:01:47:38] "GET /icons/circle_logo_small.gif HTTP/1.0" 200 2624
Jun 24 13:45:54 haproxy epo-http.txt: 202.96.29.111 [30:01:48:04] "POST /cgi-
bin/waisgate/134.67.99.11=earth1.epa.gov=210=/usr1/comwais/indexes/PressReleases=gopher%40earth1=0.00=:free HTTP/1.0" 200 3993
Jun 24 13:45:54 haproxy epo-http.txt: 202.96.29.111 [30:01:48:16] "GET /waisicons/text.xbm HTTP/1.0" 200 527
Jun 24 13:45:55 haproxy epo-http.txt: dd14-034.compuserve.com [30:01:48:22] "GET /Rules.html HTTP/1.0" 200 3273
Jun 24 13:45:57 haproxy epo-http.txt: www-c8.proxy.ool.com [30:01:48:23] "GET /docs/Searchable.html HTTP/1.0" 200 765
Jun 24 13:45:58 haproxy epo-http.txt: bettong.client.uq.oz.au [30:01:48:25] "GET /enviro/gif/banner.gif HTTP/1.0" 200 14887
Jun 24 13:54:14 form-trivia-72 app/web.1: User Load (1.2ms) SELECT "users".* FROM "users" WHERE "users"."id" = $1 ORDER BY "users"."id" ASC
LIMIT 1 [{"id": 1}]
Jun 24 13:54:14 form-trivia-72 app/web.1: (1.3ms) SELECT COUNT(*) FROM "products"
Jun 24 13:54:14 form-trivia-72 heroku/router: at=info method=GET path="/a" host=form-trivia-72.herokuapp.com request_id=3a095914-887a-4b7a-9f88-81d6e2ba7771
 fwd="23.252.53.179" dyno=web.1 connect=1ms service=44ms status=200 bytes=6407
Jun 24 13:54:14 form-trivia-72 app/web.1: Product Load (1.4ms) SELECT "products".* FROM "products" ORDER BY products.updated_at desc LIMIT 1
Jun 24 13:54:14 form-trivia-72 app/web.1: User Load (1.4ms) SELECT "users".* FROM "users" ORDER BY users.updated_at desc LIMIT 1
Jun 24 13:54:14 form-trivia-72 app/web.1: (1.2ms) SELECT COUNT(*) FROM "users"
Jun 24 13:54:14 form-trivia-72 app/web.1: method=GET path=/a/ format=html controller=rails_admin/main action=dashboard status=200 duration=35.71 view=20.85 db=6.39
 remote_ip=23.252.53.179 user_id=1 params={}
Jun 24 13:54:16 form-trivia-72 heroku/router: at=info method=GET path="/a/product?_pjax=0&data=pjax-container%5D" host=form-trivia-72.herokuapp.com
 request_id=4e7f806e-63b2-493a-88d4-ec8ebab5f0a6 fwd="23.252.53.179" dyno=web.1 connect=3ms service=102ms status=200 bytes=17350
Jun 24 13:54:16 form-trivia-72 app/web.1: Product Load (1.7ms) SELECT "products".* FROM "products" ORDER BY products.id desc LIMIT 20 OFFSET 0
Jun 24 13:54:16 form-trivia-72 app/web.1: User Load (1.2ms) SELECT "users".* FROM "users" WHERE "users"."id" = $1 ORDER BY "users"."id" ASC LIMIT 1
 [{"id": 1}]
Jun 24 13:54:16 form-trivia-72 app/web.1: (1.3ms) SELECT COUNT(*) FROM "products"
```

# What's the difference?

Logs

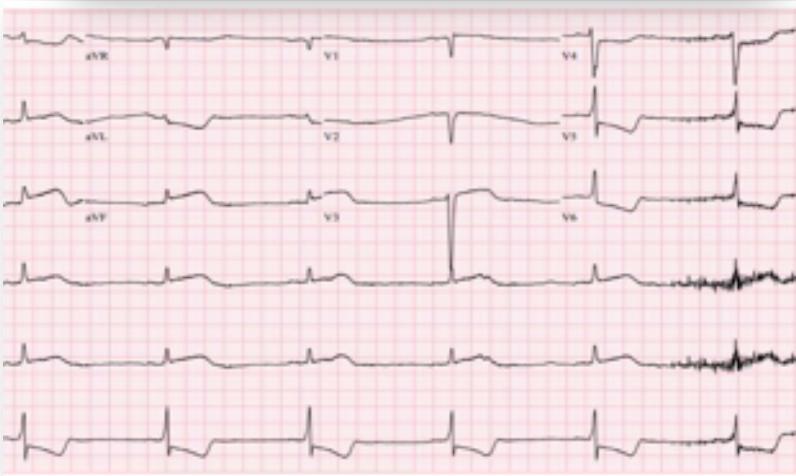
Traces



# What's the difference?



vs.

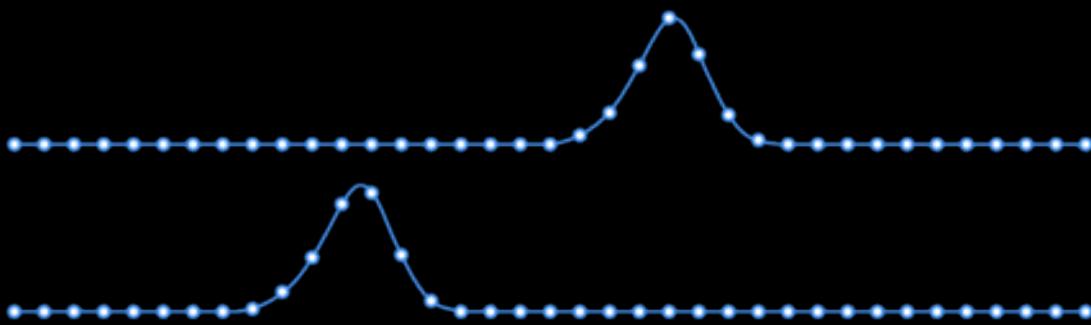


# What's the difference?

both of them are time series, but ...

We monitor the  
phenomena

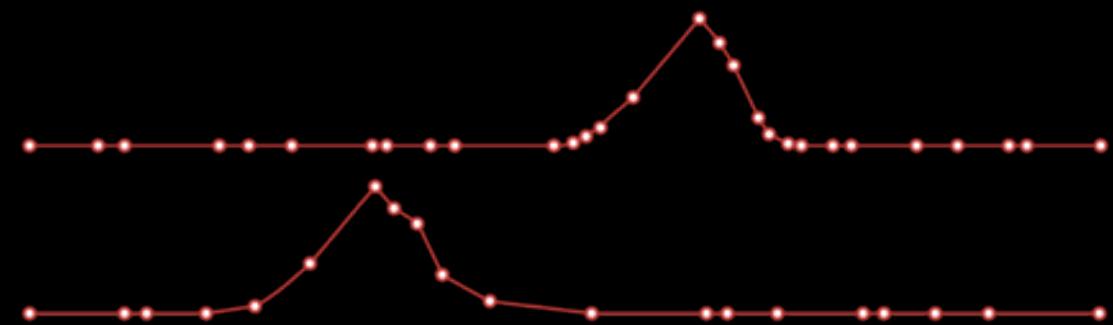
regular



metrics

The phenomena happen  
and we observe them

irregular



VS.

events

## Metrics

### Regular Time Series

Measurements  
**gathered at *regular***  
time intervals

## Events

### Irregular Time Series

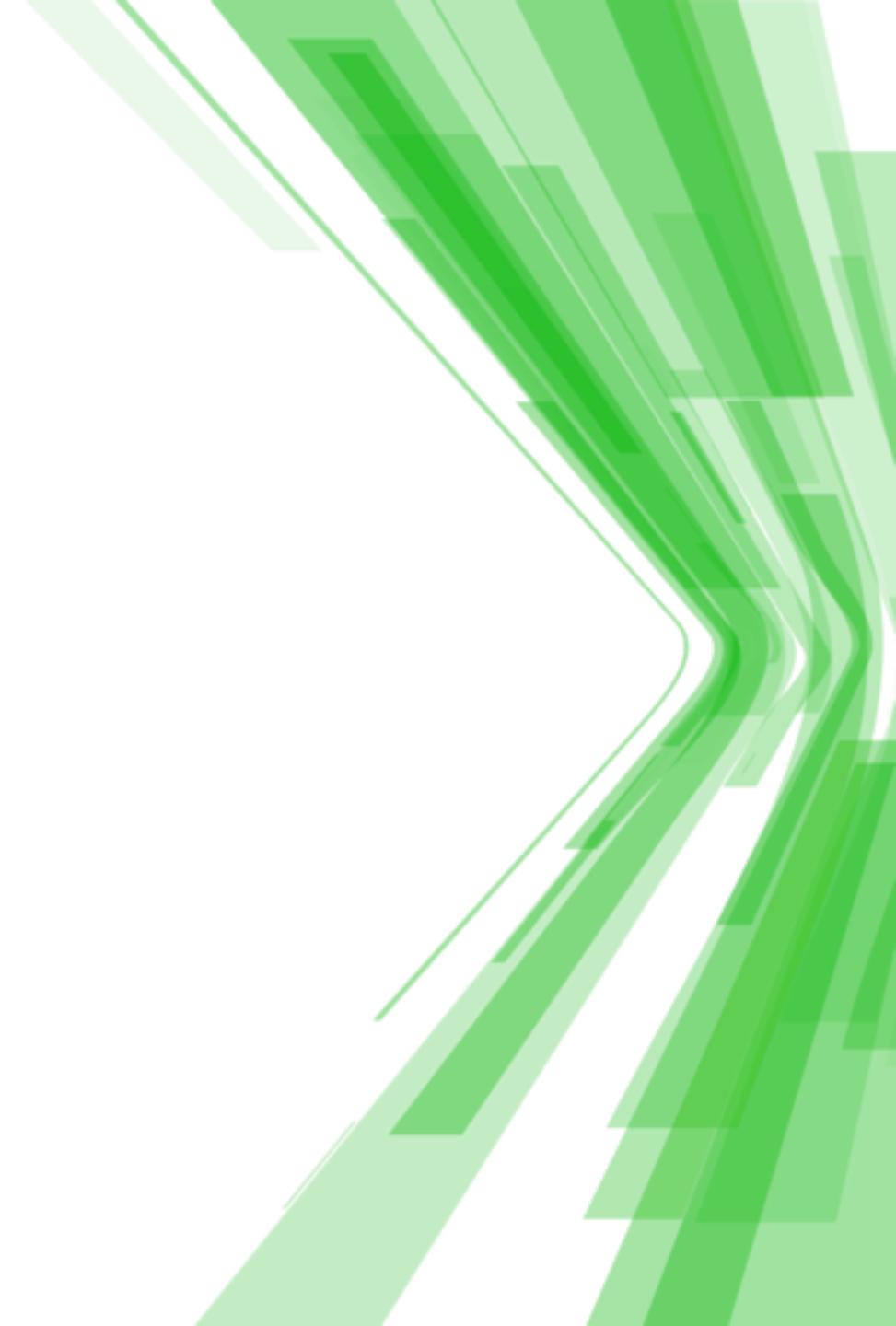
Measurements  
**observed at *irregular***  
time intervals

## Summarization of Events

Events become regular time intervals, for example

Summarizing the average trade price of Apple stock every 10 minutes over the course of a day

Summarizing the average response time for requests in an application over 1 minute intervals



**Who uses them and how?**

**Time Series**

# Primary Use Cases

## IoT

**Industrial settings:**  
factories, oil & gas,  
agriculture, smart roads  
& infrastructure

**Consumer:** wearables,  
consumer devices &  
trackers

## DevOps

**Custom monitoring  
solutions to track**  
servers, VMs,  
applications, users or  
events

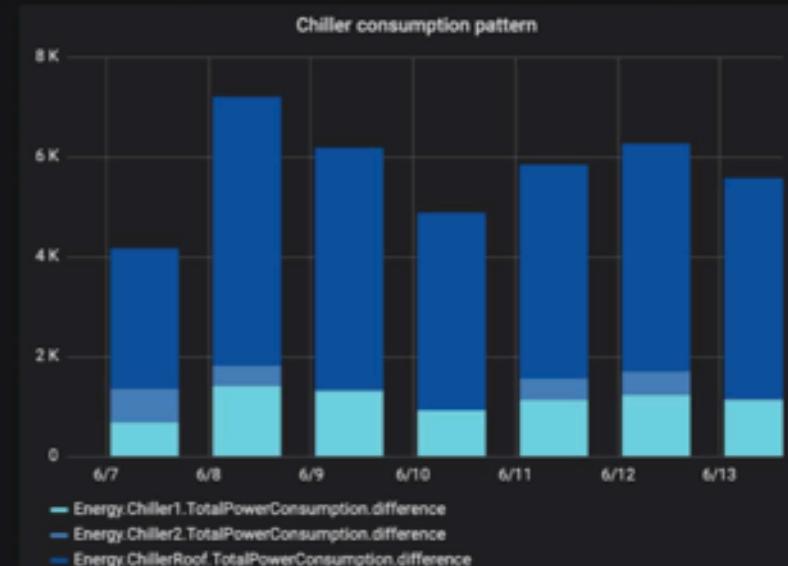
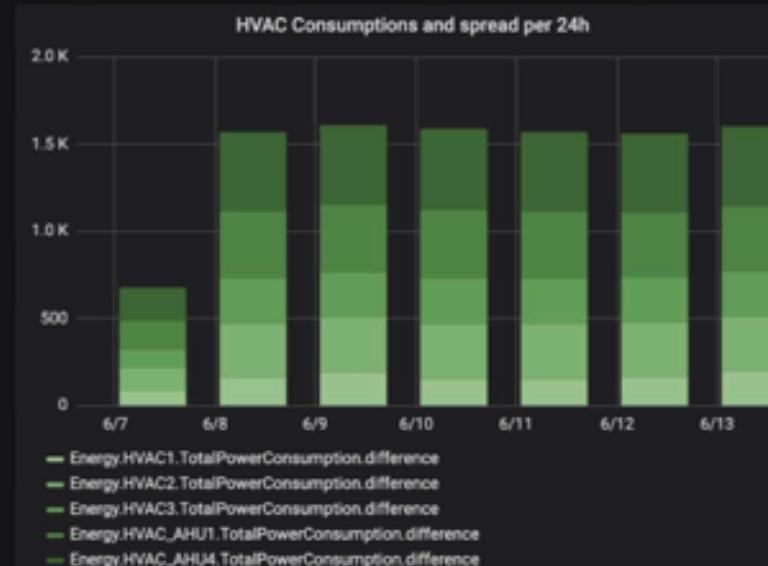
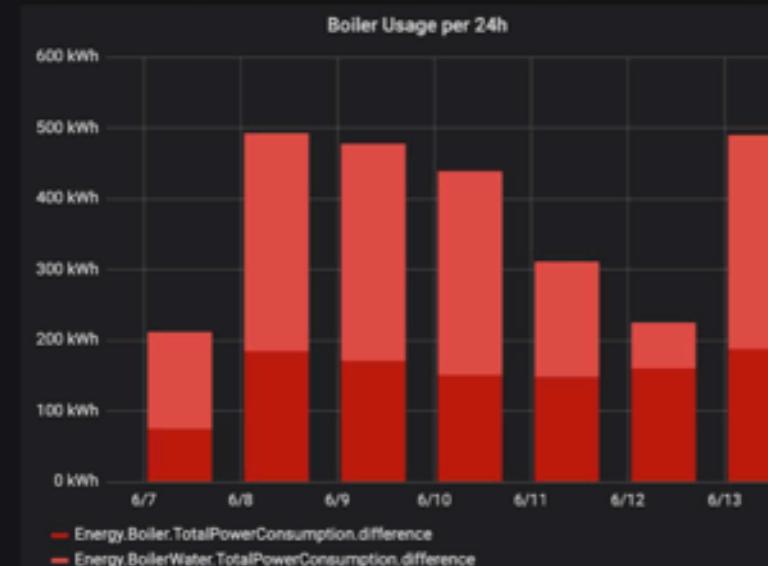
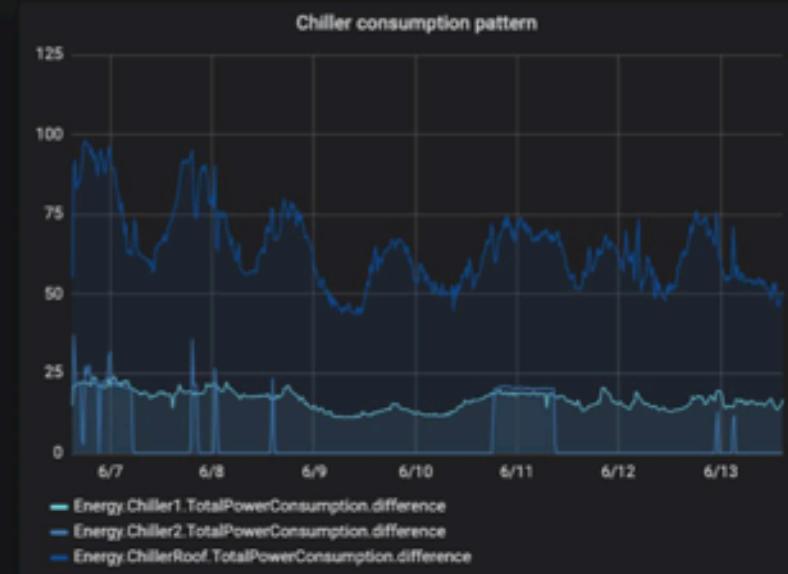
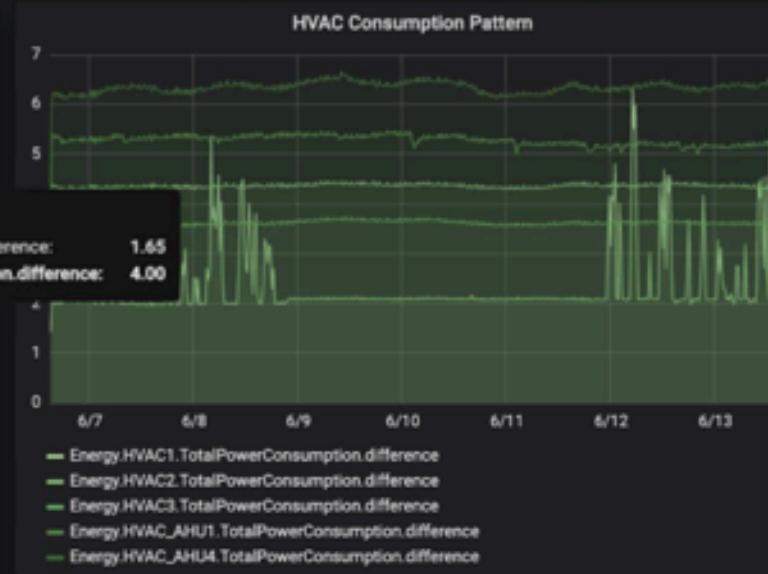
## Real-Time Analytics

**Apps that instrument**  
business, social or  
development metrics in  
real-time

## USE-CASE: BIOREACTOR



# USE-CASE: ENERGY MONITORING



[Park](#) [6HD](#) [Turbine](#) [All](#)

Dashboard Row

## USE-CASE: WIND TURBINE OPERATIONS



6HD01



6HD02



6HD03



Back to the nature  
of time series ...  
Time Series

# Question: Time-Series?



yes

no

Data items of **time series** can be complex and **semi-structured** ...

# Characteristics of the time series

All Time-stamped data

Generated in

- regular (Metric) and
- irregular (Event) time periods

Huge volumes of data

High variety of semi-structured data

Real-time

Time sensitive

# So what's a time series DB?

## Time Series

## Time series DB

optimized for

- collecting
- storing
- retrieving
- processing  
(historical and  
real-time)

## Other DBs

### **Traditional relational**

Databases optimized for storing and querying structured data

### **Document databases**

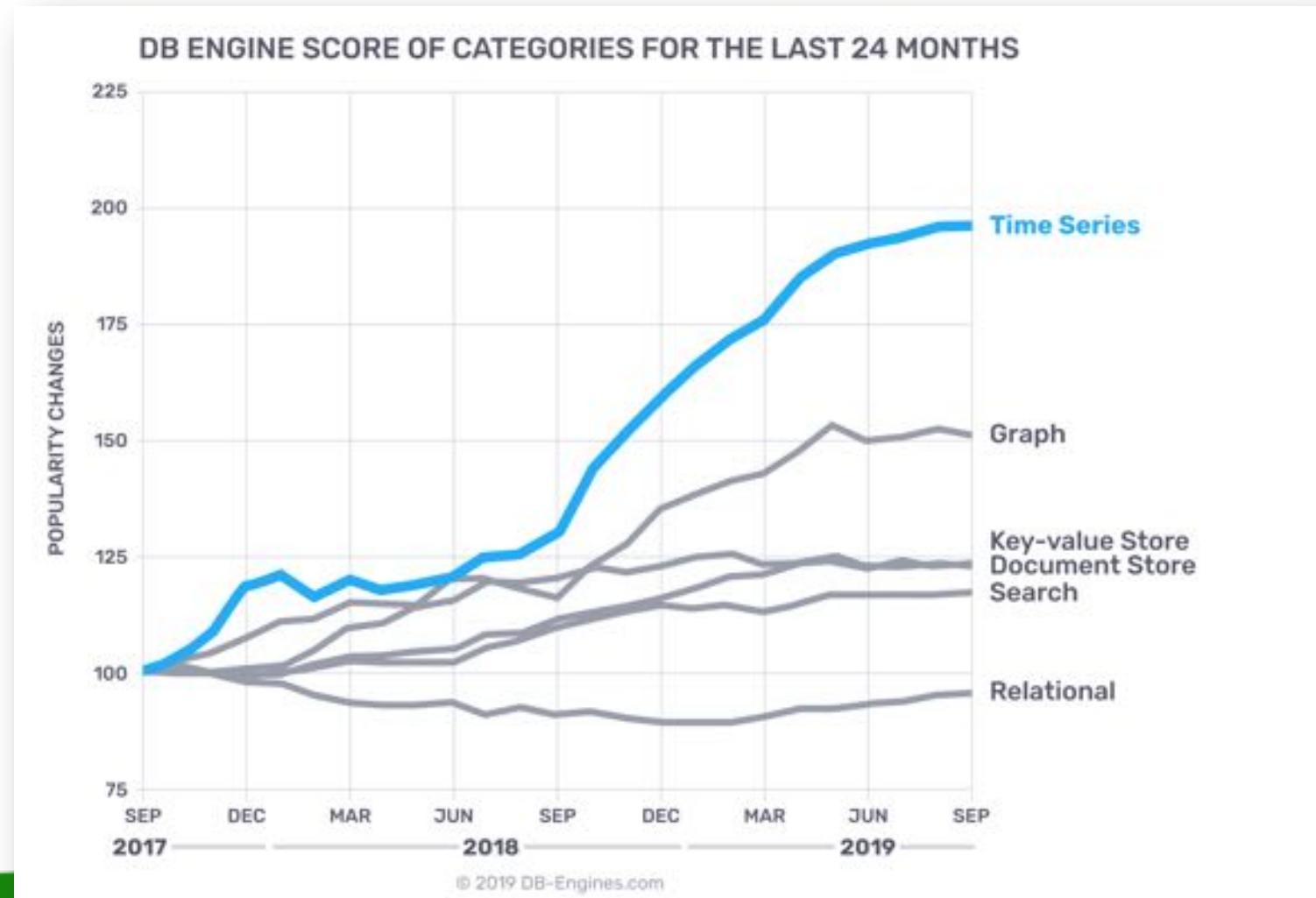
Optimized for storing and retrieving semi-structured JSON documents

### **Search databases**

Optimized for storing and retrieving unstructured data (e.g., full-text searches)

# The interest about time series data bases is growing

[https://db-engines.com/en/ranking\\_categories](https://db-engines.com/en/ranking_categories)





*influx*/days



---