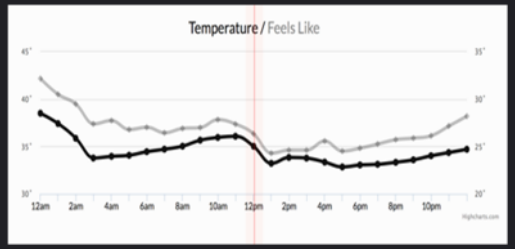
Time series

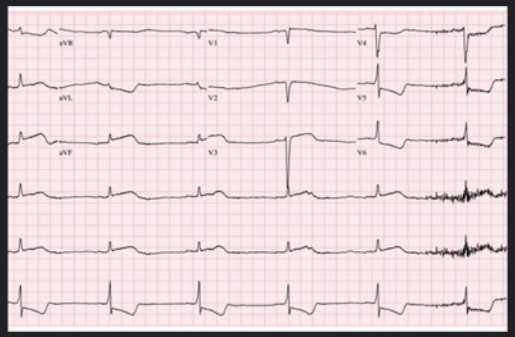
Emanuele Della Valle
Prof. @ Politecnico di Milano
Founder & Partner @ Quantia Consulting
Marco Balduini
Founder & CEO @ Quantia Consulting
Riccardo Tommasini
Prof. @ University of Tartu

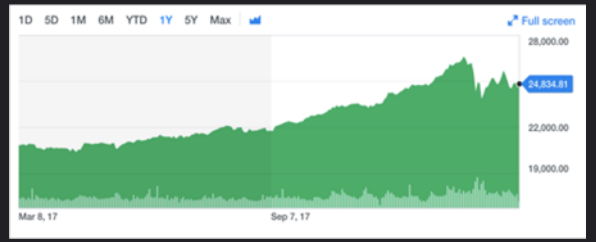
What are time series?

Let’s start by example

- Weather conditions
- Stock exchange
- Cluster monitoring
- Healthcare

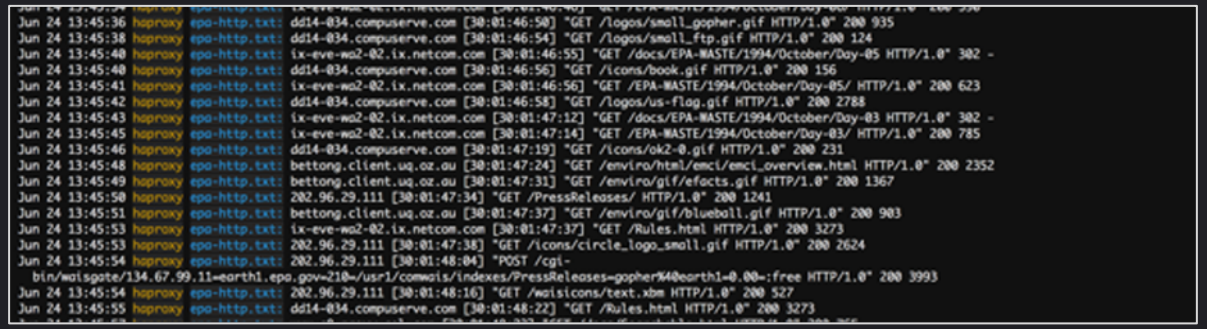


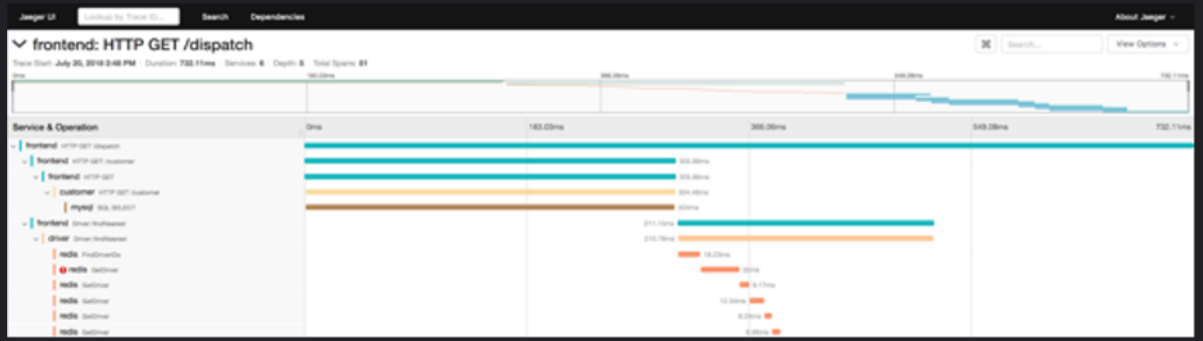




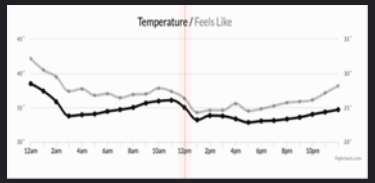
Now some different type of time series

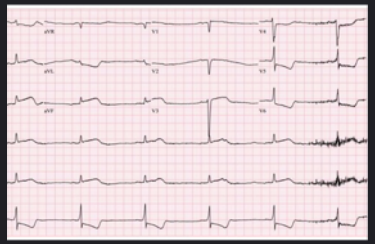
- Logs
- Traces



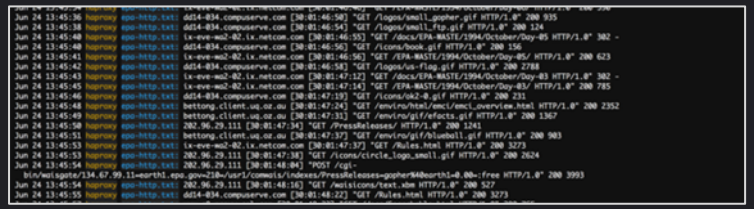


What's the difference?





VS.



What's the difference?

both of them are time series, but …

We monitor the phenomena

regular

metrics

VS.

The phenomena happen and we observe them

irregular

events

Metrics	Events
Regular Time Series	**Irregular Time Series**
Measurements **gathered** at *regular* time intervals	Measurements **observed** at *irregular* time intervals

Summarization of Events

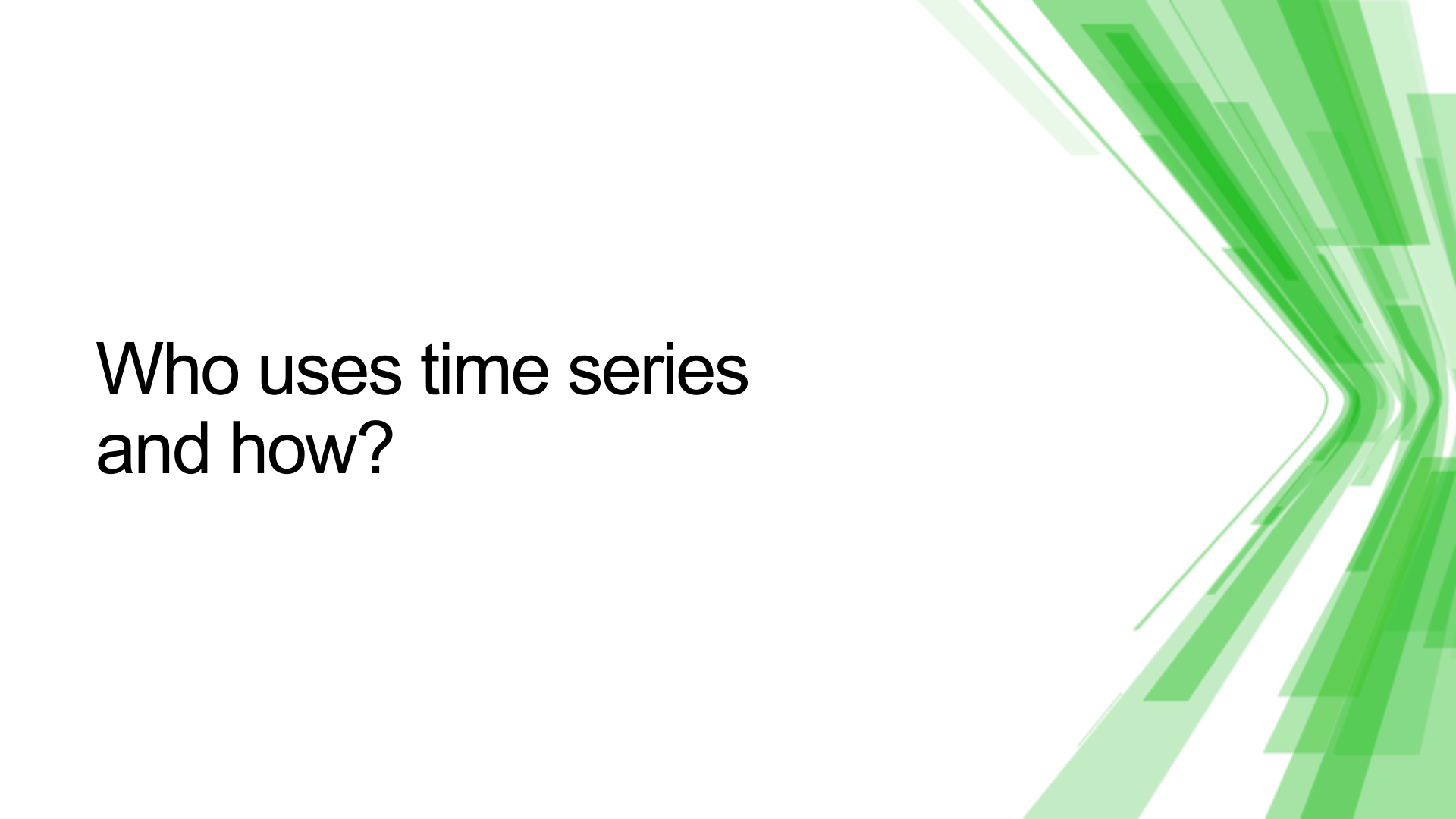
Events become regular time intervals, for example

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

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

Who uses time series and how?



Primary Use Cases

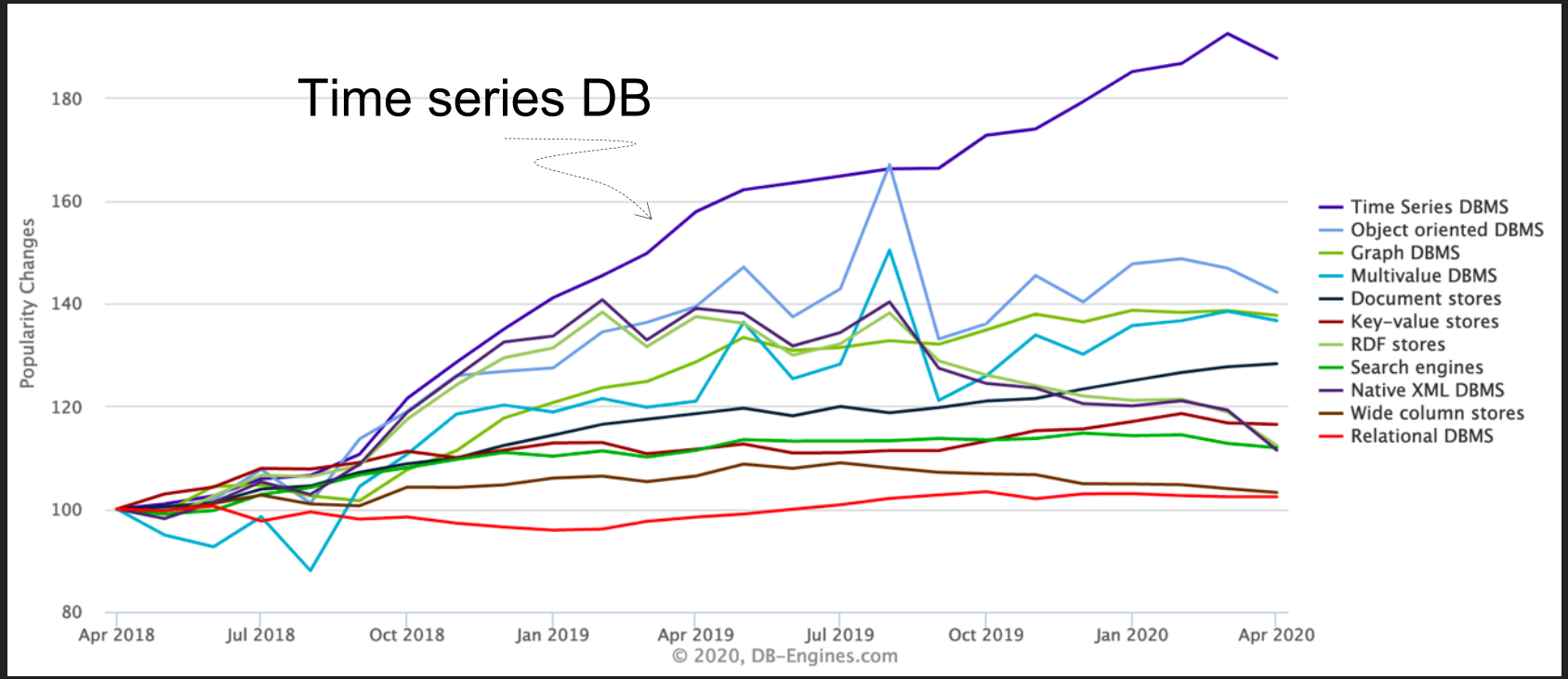
IoT	DevOps	Real-Time Analytics
Industrial settings: factories, oil & gas, agriculture, smart roads & infrastructure **Consumer:** wearables, consumer devices & trackers	**Custom monitoring solutions to track** servers, VMs, applications, users or events	**Apps that instrument** business, social or development metrics in real-time

So what’s a time series DB?

Time series DB	Other DBs
optimized for • collecting • storing • retrieving • processing (historical and real-time) timestamped & semi-structured data	**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



influx/days

Time series

Emanuele Della Valle Prof. @ Politecnico di Milano & Partner @ Quantia Consulting
Marco Balduini Founder & CEO @ Quantia Consulting
Riccardo Tommasini Prof. @ University of Tartu