



Introducing Time Series & Warp 10

Horacio Gonzalez
@LostInBrittany

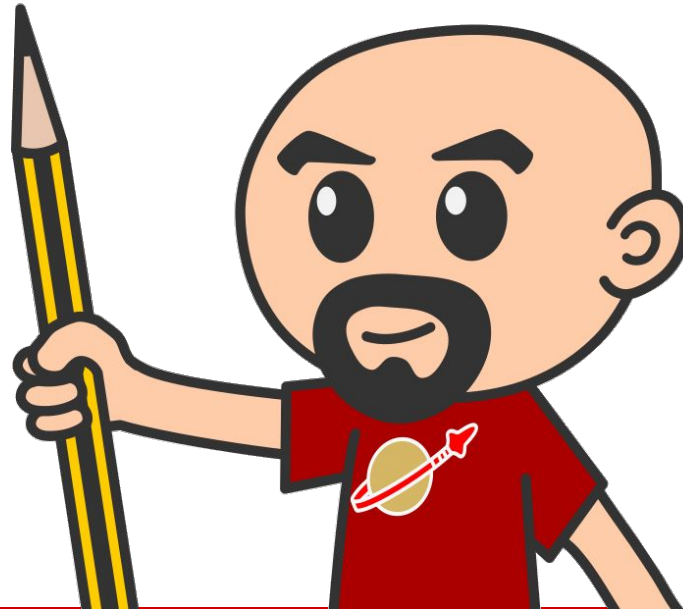


Horacio Gonzalez



@LostInBrittany

Spaniard lost in Brittany, developer,
dreamer and all-around geek





Time Series

What are they?

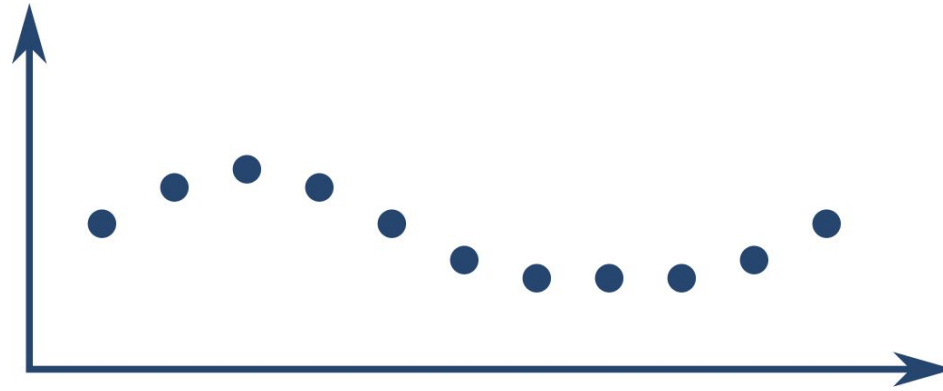


Time Series



Definition of Time Series:

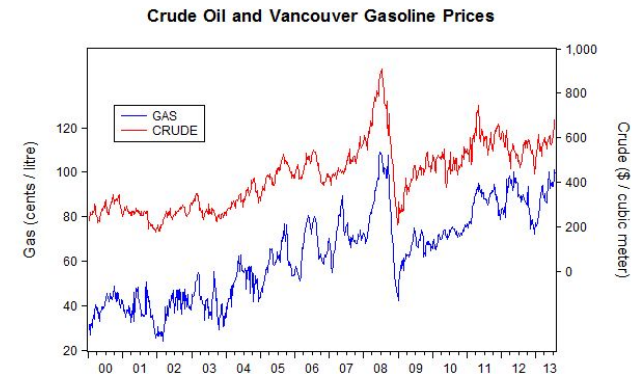
An ordered sequence of values of a variable at equally spaced time intervals.



Time Series



- Stock Market Analysis
- Economic Forecasting
- Budgetary Analysis
- Process and Quality Control
- Workload Projections
- Census Analysis
- ...

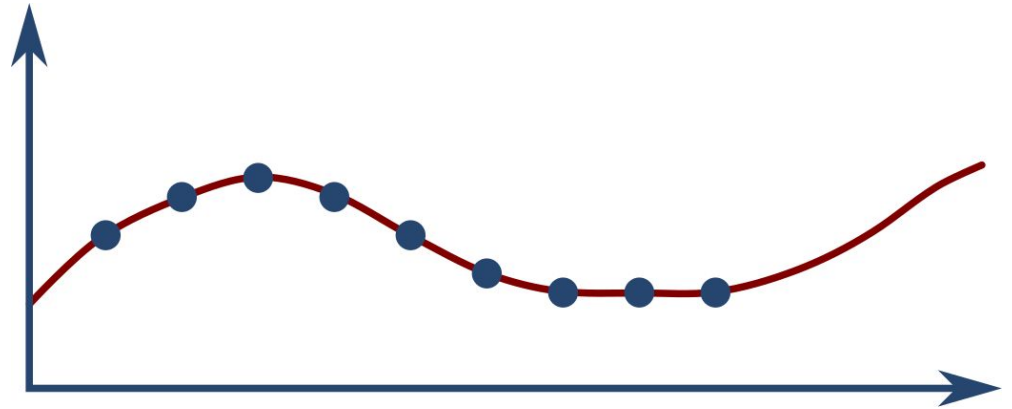


Time Series



Applications:

- Understanding the data
- Fit a model
 - Monitoring
 - Forecasting



Time Series



Stock market Analytics
Economic Forecasting



\$\$\$



Study & Research



Time Series



Many specific analytical tools:

- Moving average
- ARMA (AutoRegressive Moving Average)
- Multivariate ARMA models
- ARCH (AutoRegressive Conditional Heteroscedasticity)
- Dynamic time warping
- ...



Time Series



Specific application of general tools

- Artificial neural networks
- Hidden Markov model
- Fourier & Wavelets transforms
- Entropy encoding
- ...





Time Series Databases

What tools do I use?



Time Series Databases



General purpose data analytics tools:

- Matlab
- Python
- R
- ...



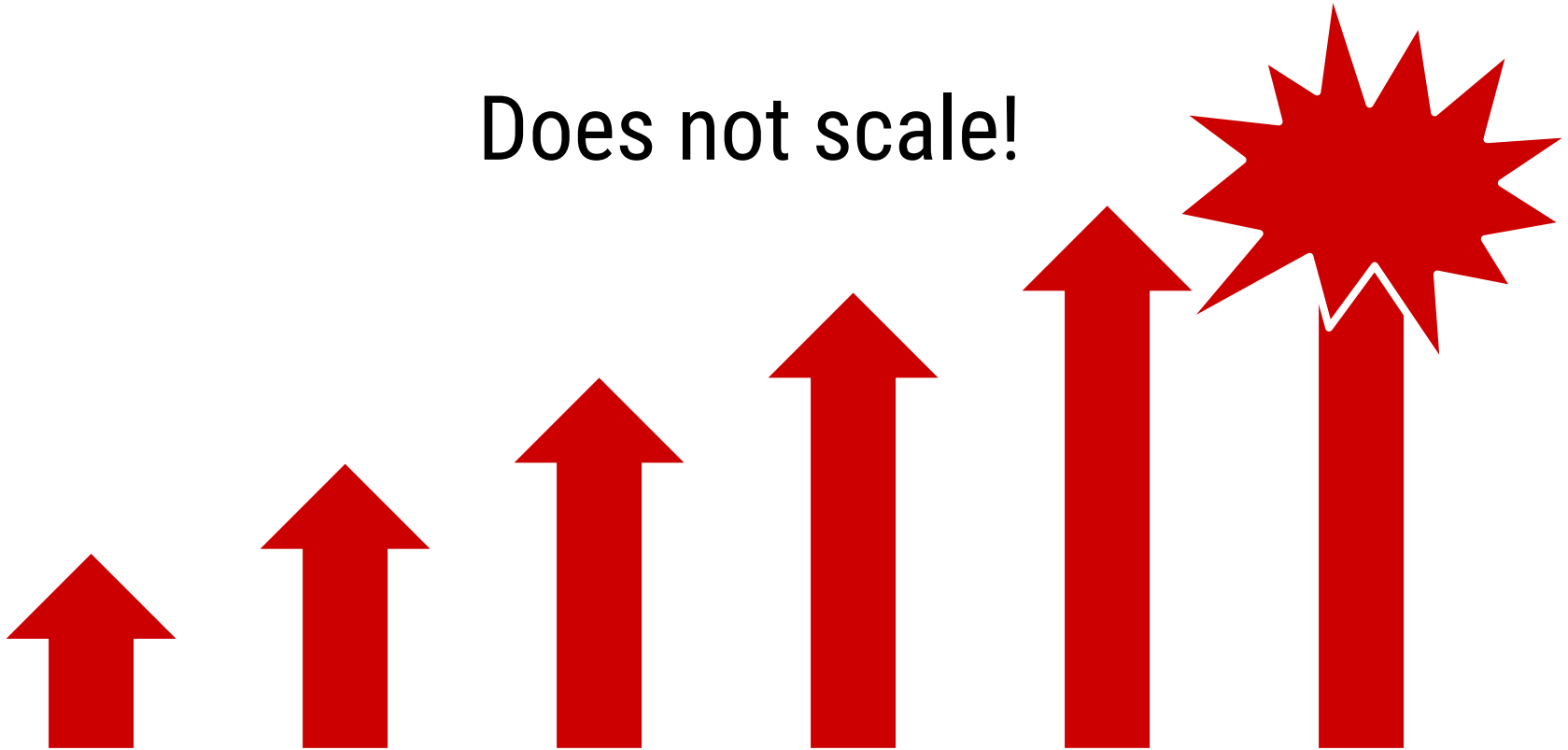
General purpose relational database engine



Time Series Databases



Does not scale!



Time Series Databases



Many type of database engine

- Relational databases
- Key-value databases
- Document databases
- Graph databases
- ...



Time Series Databases



What about Time Series?

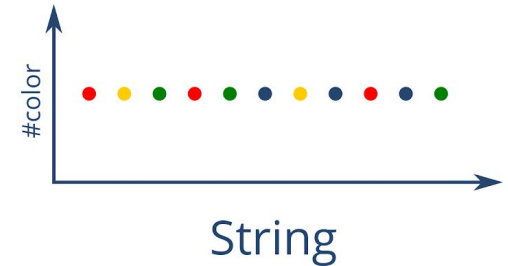
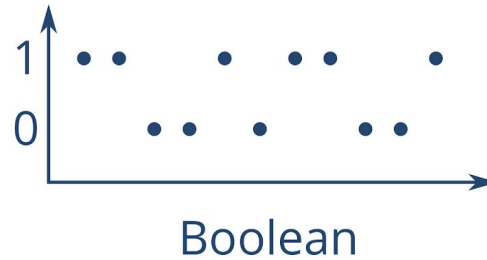
Time Series databases!



Time Series Databases



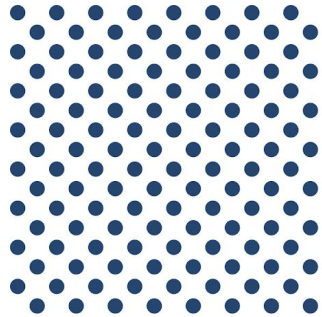
Data model: time series



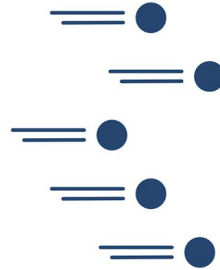
Time Series Databases



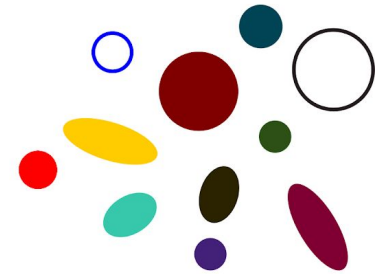
The 3 'v'



Volume



Velocity



Variety



Time Series Databases



Many options



Time Series Databases



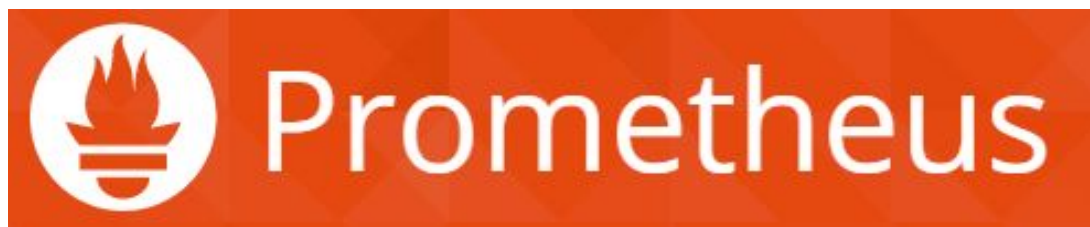
Time Series Databases



OPENTSDDB



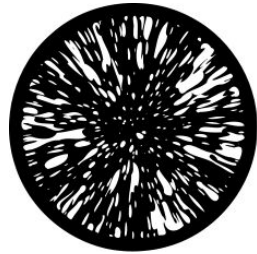
Time Series Databases



Time Series Databases



Time Series Databases



WARP 10





OVH Metrics

What did we choose?



What's OVH Metrics



Managed Cloud Platform for Geo Time Series®



What's a metric?



[**me**-trik] : the science of measuring



What's a metric?



“To measure is to know”

William “Lord Kelvin” Thomson



What's a metric?



Actionable metric \neq vanity metric



What is a metric?



Metrics are Time Series!



How do we deal with metrics?



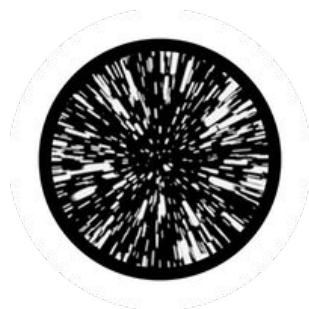
Using a Time Series Database!

But... which one?

Why choose? Let's support all of them!



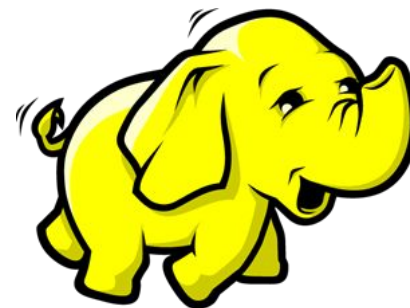
Metrics Data Platform



+



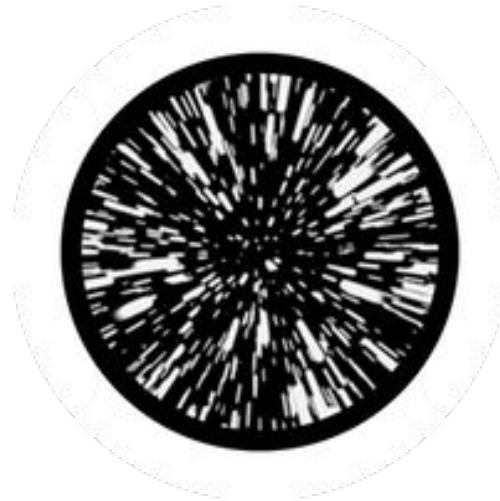
+



Metrics Data Platform



And why Warp 10?



Metrics Data Platform



Some Metrics's metrics:

- 1.5M datapoints/s, 24h/7
- Peaks at ~10M datapoints/s
- 500M unique series
- Deleted (varying) up to 8M datapoints/s
- 2 Trillion deletes capacity per week





Warp 10

Open-source Time Series Database



More than a Time Series DB



Warp 10 is a software platform that

- Ingests and stores time series
- Manipulates and analyzes time series

The screenshot shows the Warp 10 Platform website. The browser address bar displays 'www.warp10.io'. The page has a blue header with the 'Warp 10' logo and a search bar. A left sidebar contains a navigation menu with items: Home, Overview, Concepts, History, Team, Versions, How-to, APIs, WarpScript, Tools, and Support. The main content area features the title 'The Warp 10 Platform' followed by a paragraph: 'The Warp 10 Platform is designed to collect, store and manipulate sensor data. Sensor data are ingested as sequences of measurements (also called time series). The Warp 10 Platform offers the possibility for each measurement to also have spatial metadata specifying the geographic coordinates and/or the elevation of the sensor at the time of the reading. Those augmented measurements form what we call Geo Time Series® (GTS)'. Below this is the 'Geo Time Series®' section, which states: 'The first differentiating factor of Warp 10 is that both space (location) and time are considered first class citizens. Working with Geo Time Series® (GTS) allows you to have geo-located readings without having to use four separate series and having to keep track of the reading context.' and 'Complex searches like "find all the sensors active during last Monday in the perimeter delimited by this geo-fencing polygon" can be done without involving expensive joins between separate time series for the same source.' The 'Platform components' section explains: 'Readings are pushed into the Warp 10 platform via a HTTP call to a component called Ingress. In the distributed version of the platform, the Ingress component can be instantiated multiple times to support very high ingestion rates. Once data is dealt with by Ingress it is considered persisted by the Warp 10 storage layer called Continuum (for the spacetime continuum).' and 'Warp 10 also offers streaming endpoints to push data using a WebSocket or to consume data as it enters the platform to build dynamic dashboards and integrate Warp 10 into a more elaborate stream processing workflow.' The final paragraph notes: 'Security and privacy have also been addressed by Warp 10 since its very inception, this includes fine grain access control mechanisms, encryption capabilities and throttling management to enable full multi-tenancy of the platform.'





Many time-series solutions





But they are only stores...



Fetching data is only the tip of the iceberg





Analysing the data

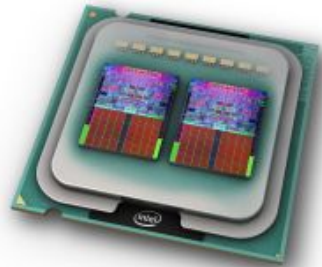


High level analysis must be done elsewhere





Algorithms are resource hungry



Your computer is not a datacenter



A few days?...
but I want it now!





Manipulating Time Series



To be scalable, analysis must be done in the TSDB engine, not in user's computer





Manipulating Time Series with Warp 10

A true Time Series analysis toolbox

- Hundreds of functions
- Manipulation frameworks
- Analysis workflow





Manipulating Time Series with Warp 10

Why not a simple REST API?

- One endpoint by function?
- How to chain a workflow analysis?



REST API not suitable for
complex manipulations





Manipulating Time Series with Warp 10

Why not a SQL dialect?

- How do you do a simple moving average in SQL?
- How do you geo-time fencing in SQL?



SQL is not adapted to TS analysis!



Manipulating Time Series with Warp 10



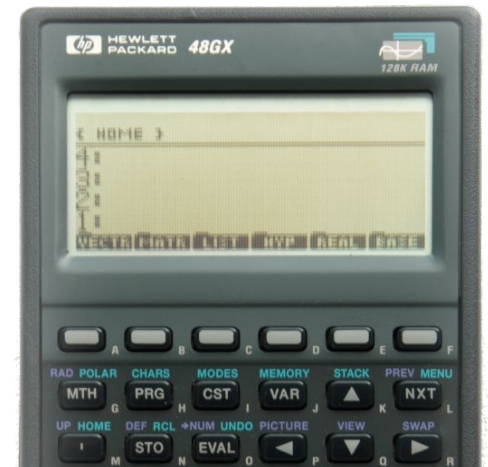
Solution: a Time Series manipulation language





WarpScript

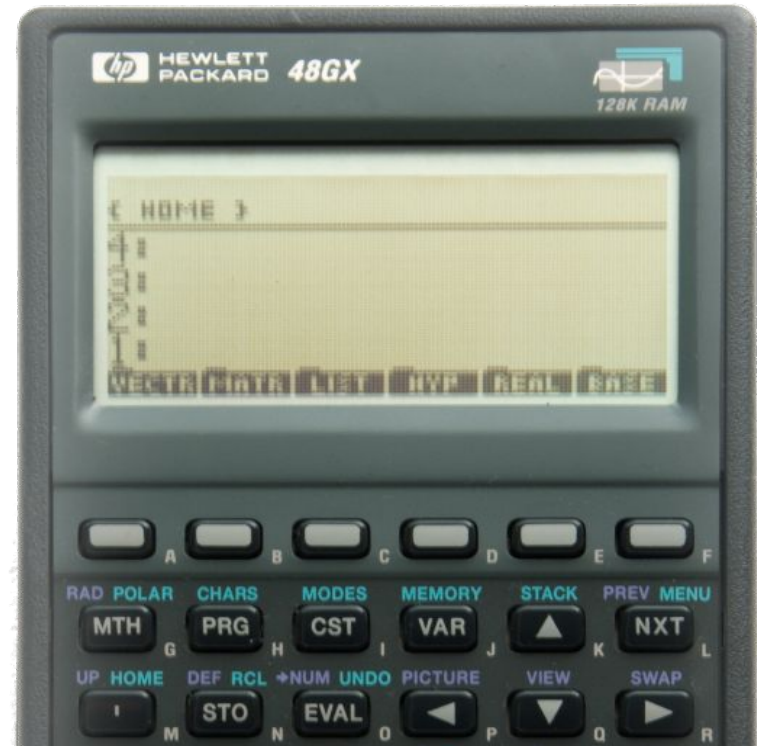
Time Series manipulation language



A stack based language



Input	2	3	add	11	mul	1	add
Stack	2	3		11		1	
	2	2	5	5	55	55	56



Basic operations



```
// This is a commentary  
'a' // string value  
true // boolean value  
42 // long value  
3.14159 // double value  
+ // operations  
  

```



Five frameworks



- BUCKETIZE
- MAP
- REDUCE
- FILTER
- APPLY



More than 800 functions



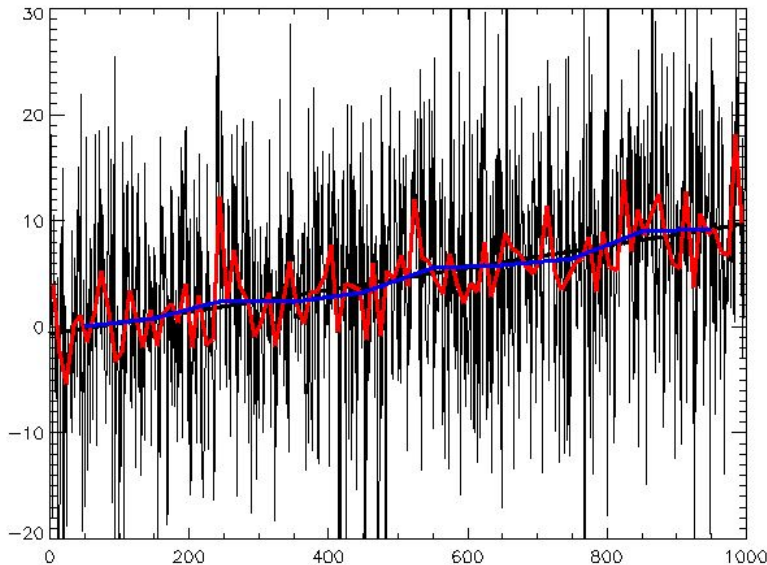
Trigonometry

Data & Time

Macros

List & Maps

Logic
Structures



Strings

Maths &
statistics

Loop
Structures



Time series functions



TIMECLIP

TIMESPLIT

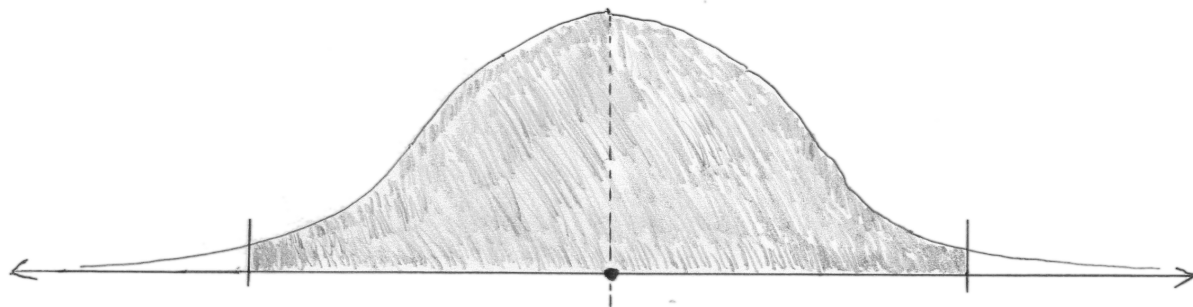
SHRINK

MERGE

...



Time series functions



MUSIGMA

NORMALIZE

NSUMSUMQ

STANDARDISE

ZSCORE



Geo-Time Series functions



Geo mapping (WKT)

Horizontal & vertical speed

Horizontal & vertical distance

Haversine

...



Quantum IDE



Quantum ☰

WarpScript
Ingress
Delete

Choose your backend:

Distributed Warp

Choose another backend

WarpScript

```
1
2 // This is a commentary
3 'a' // string value
4 true // boolean value
5 42 // long value
6 3.14159 // double value
7 + // operations
8
9 20 22 + // several items in one line
10
```

Permalink:
[Ci8vIFRoaxMgaXMgYsBjb21tZW50YXJ5CidhJyAgCS8vIHNoCmluZyB2YWx1ZQp0cnVlIAkvLyBib29s...](#)

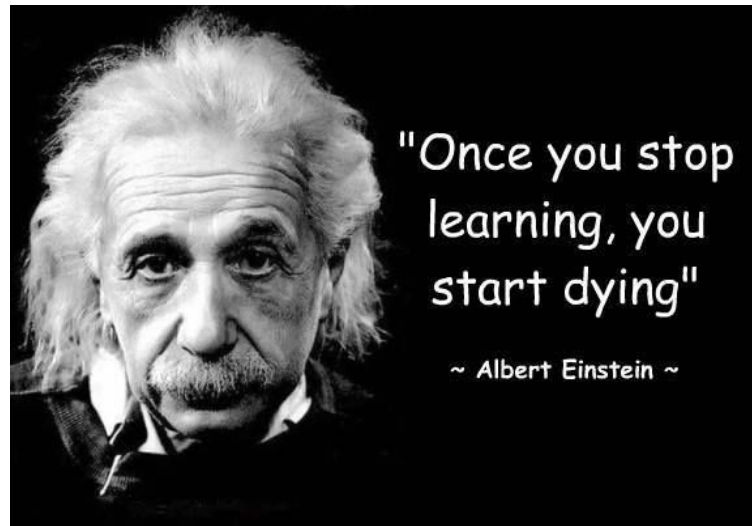
Execute!



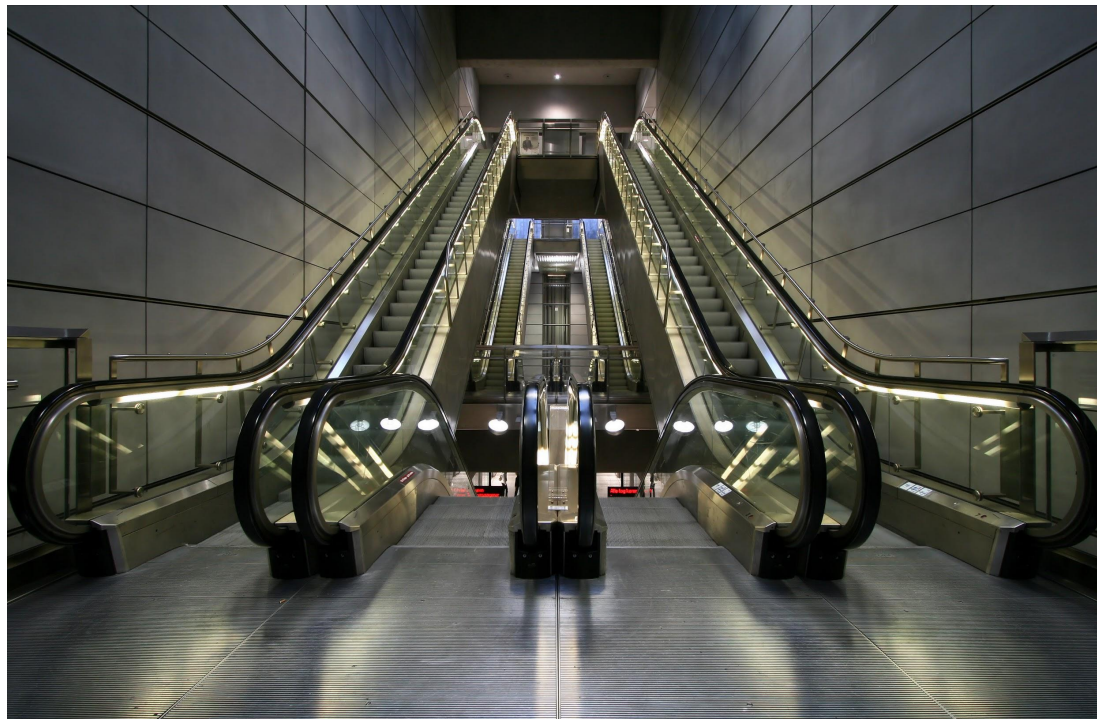


And this is only the surface

Possibilities are endless



Think differently



Time Series are everywhere



