

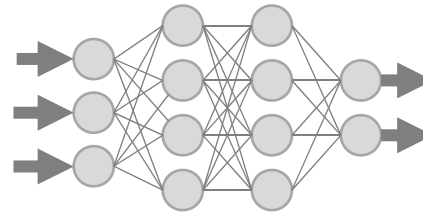
Preparing Real-time Data for Data Science

Brandon Perry

real-time data
a.k.a
real-world data



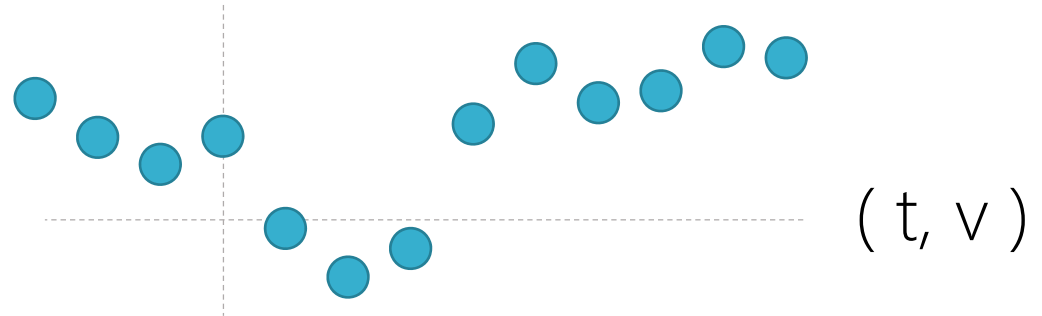
data science



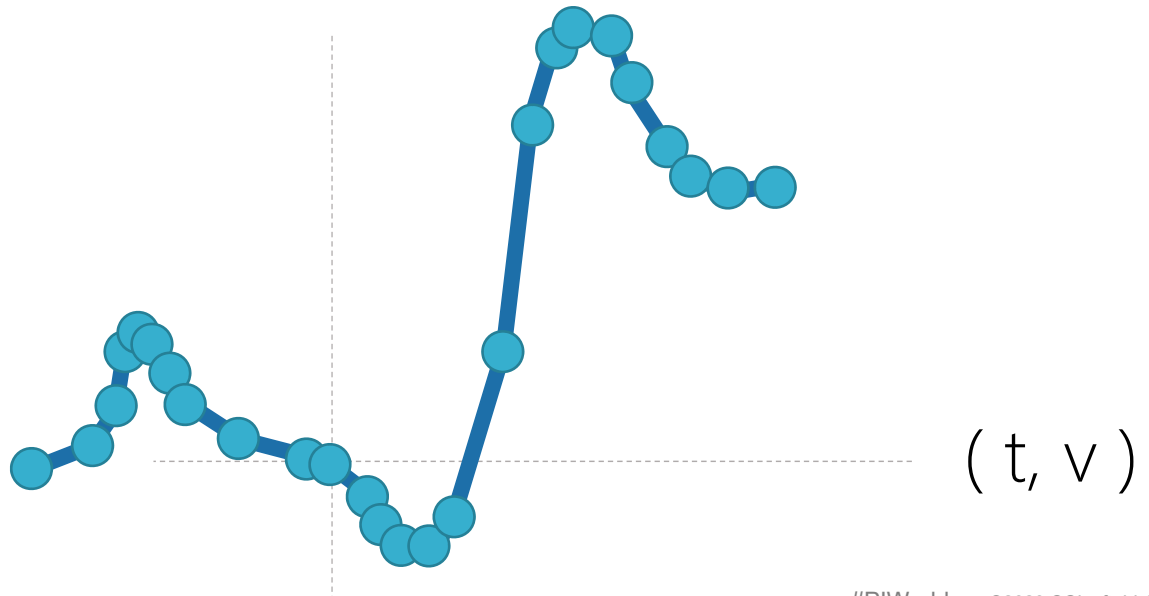
“time series”
“features”



Time Series "Samples"



Time Sequence "Signal"



values



Time



Transport



what | select, project



when | Gaps, interpolation, summaries, events



HOW | push, pull, cloudify



what | select, project (& know what you're looking at)

what | (be able to) select your data



OCS
Metadata Rules

OCS Stream
Metadata

```
{  
  "site": "Yukon 1",  
  "Asset Type": "Pump",  
  "Asset Id": "314",  
}
```



PI Asset
Framework

PI AF Builder

& know your data source

Regenerant Temperature	172.9 °F
Process maximum	220 °F
Process minimum	150 °F
Sensor accuracy	2 delta °F
Sensor maximum	480 °F
Sensor minimum	0 °F
Sensor type	Type K Thermocouple

what | Project a.k.a. shape, associate, align


Stream
Metadata ▶

OCS Data Views


PI Assets ▶

PI Integrator

\$self

via PI AF SDK, PI SQL, PI Web API, etc.

Configure Data View

Use the configuration pane on the left to manage the data view. Then click the **Apply** button to generate a preview of the current configuration.

Index Configuration ⓘ

Field Management ⓘ

Index Field

Timestamp
Index

Grouping Fields

{IdentifyingValue} Location
Metadata

{IdentifyingValue} Device Number
Metadata

Data Fields (Query1) [Manage Queries](#)

✓ {IdentifyingValue} Value
Property Id

■ {IdentifyingValue} Id
Id

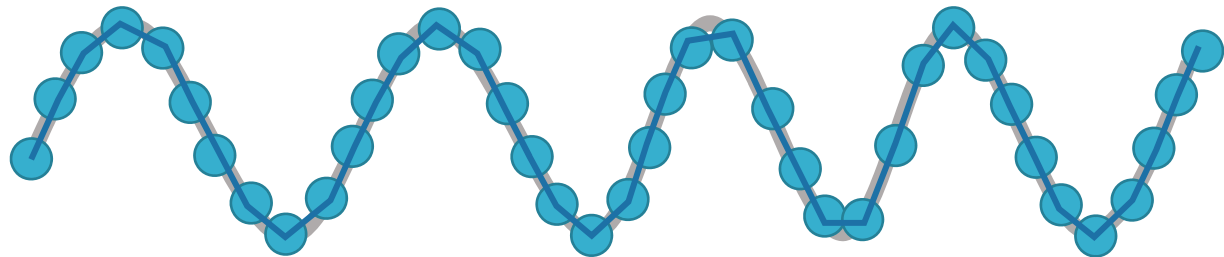
Name Tank_Inflow_Outflow | **Description** Description...

Timestamp	Location	Dev
Jun 7, 2020, 12:00:00 AM	Chicago	516
Jun 7, 2020, 1:00:00 AM	Chicago	516
Jun 7, 2020, 2:00:00 AM	Chicago	516
Jun 7, 2020, 3:00:00 AM	Chicago	516
Jun 7, 2020, 4:00:00 AM	Chicago	516
Jun 7, 2020, 5:00:00 AM	Chicago	516
Jun 7, 2020, 6:00:00 AM	Chicago	516
Jun 7, 2020, 7:00:00 AM	Chicago	516
Jun 7, 2020, 8:00:00 AM	Chicago	516
Jun 7, 2020, 9:00:00 AM	Chicago	516
Jun 7, 2020, 10:00:00 AM	Chicago	516
Jun 7, 2020, 11:00:00 AM	Chicago	516
Jun 7, 2020, 12:00:00 PM	Chicago	516
Jun 7, 2020, 1:00:00 PM	Chicago	516

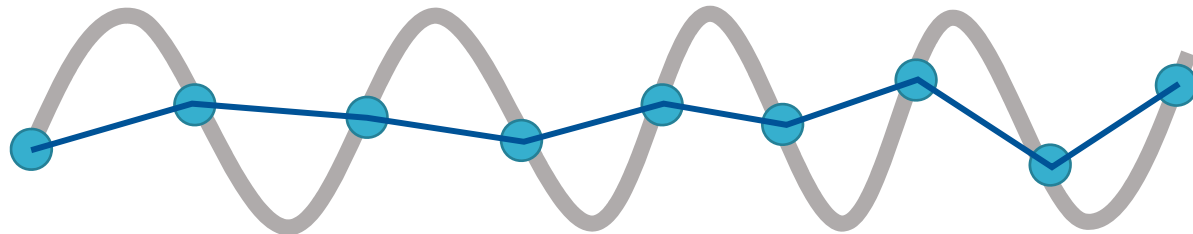
when | Gaps, interpolation, summaries, events

when | continuity

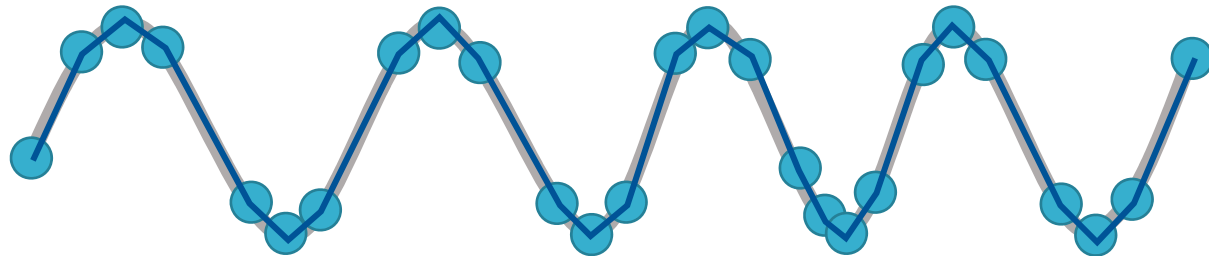
Well-sampled



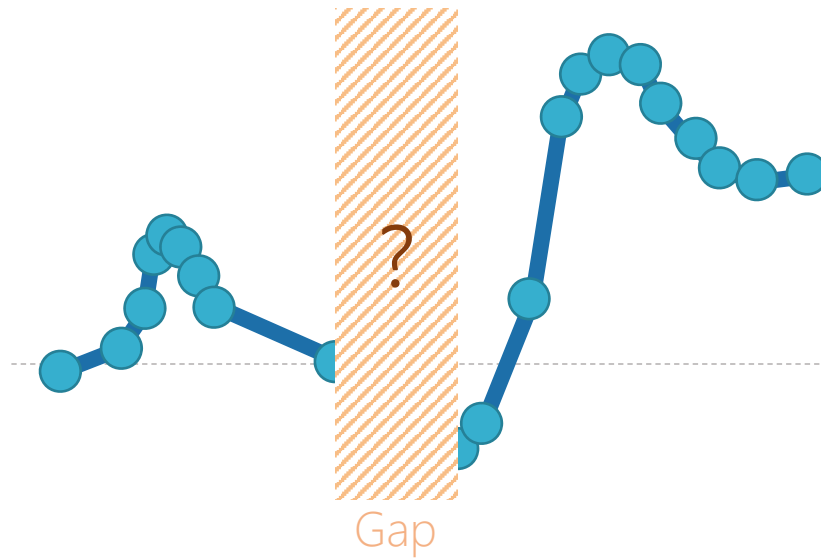
Under-sampled



Well-sampled,
compressed



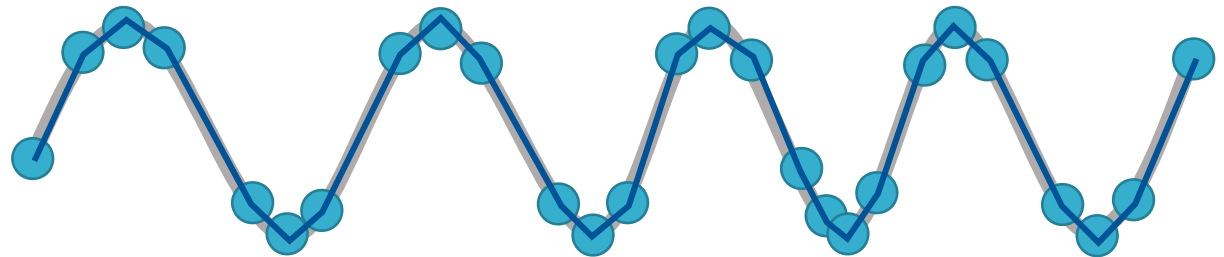
when | Gaps



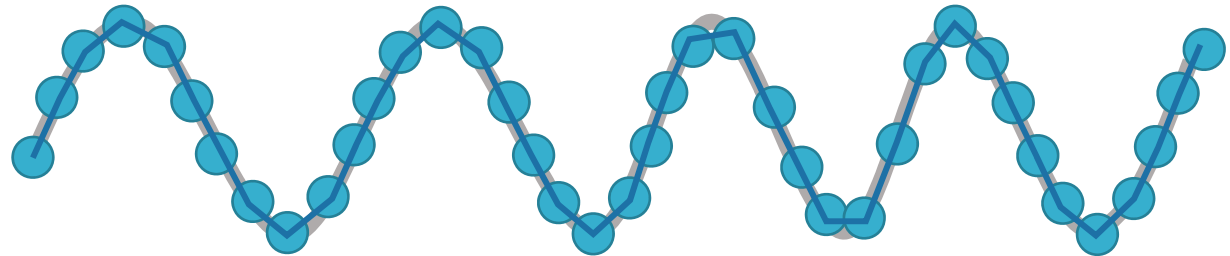
9/5/18 21:35	989.96
9/5/18 21:36	SOURCE OFFLINE
9/5/18 22:01	986.92

when | Interpolation

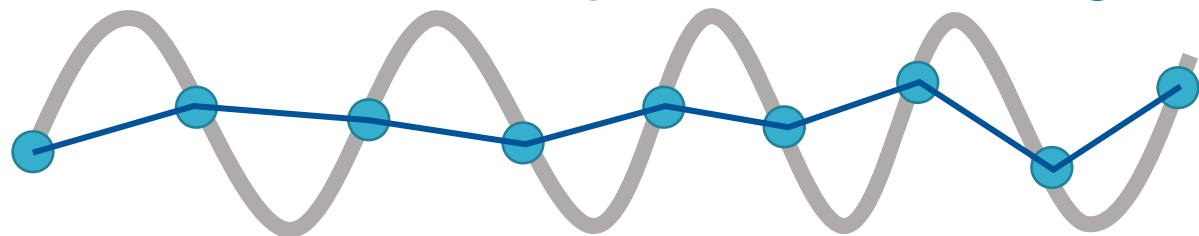
Recorded



Interpolated

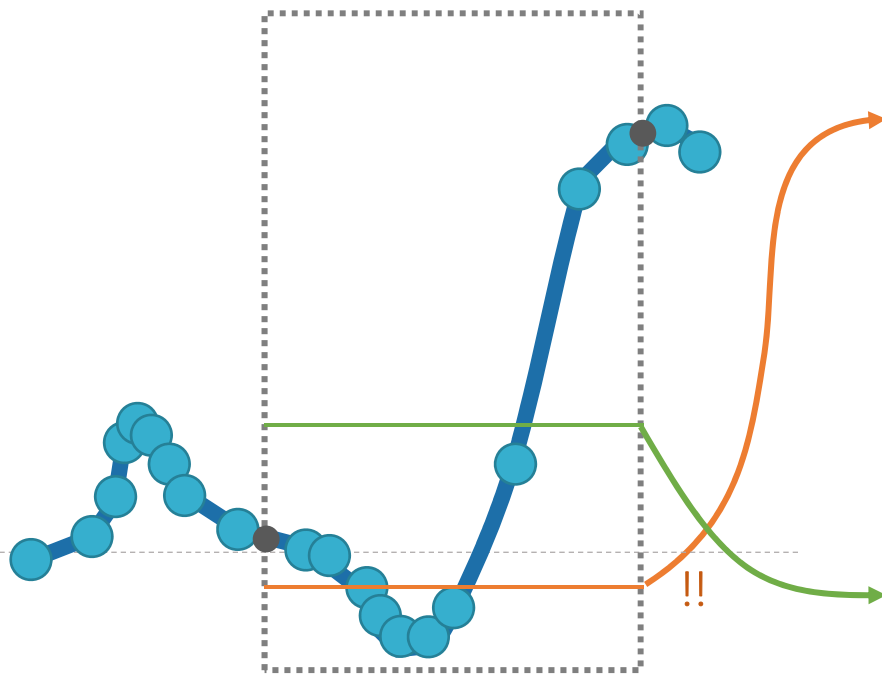


Under-interpolated



when | summaries

Q: Average value in this window?



Naïve

e.g. **AVG()** of a simple array

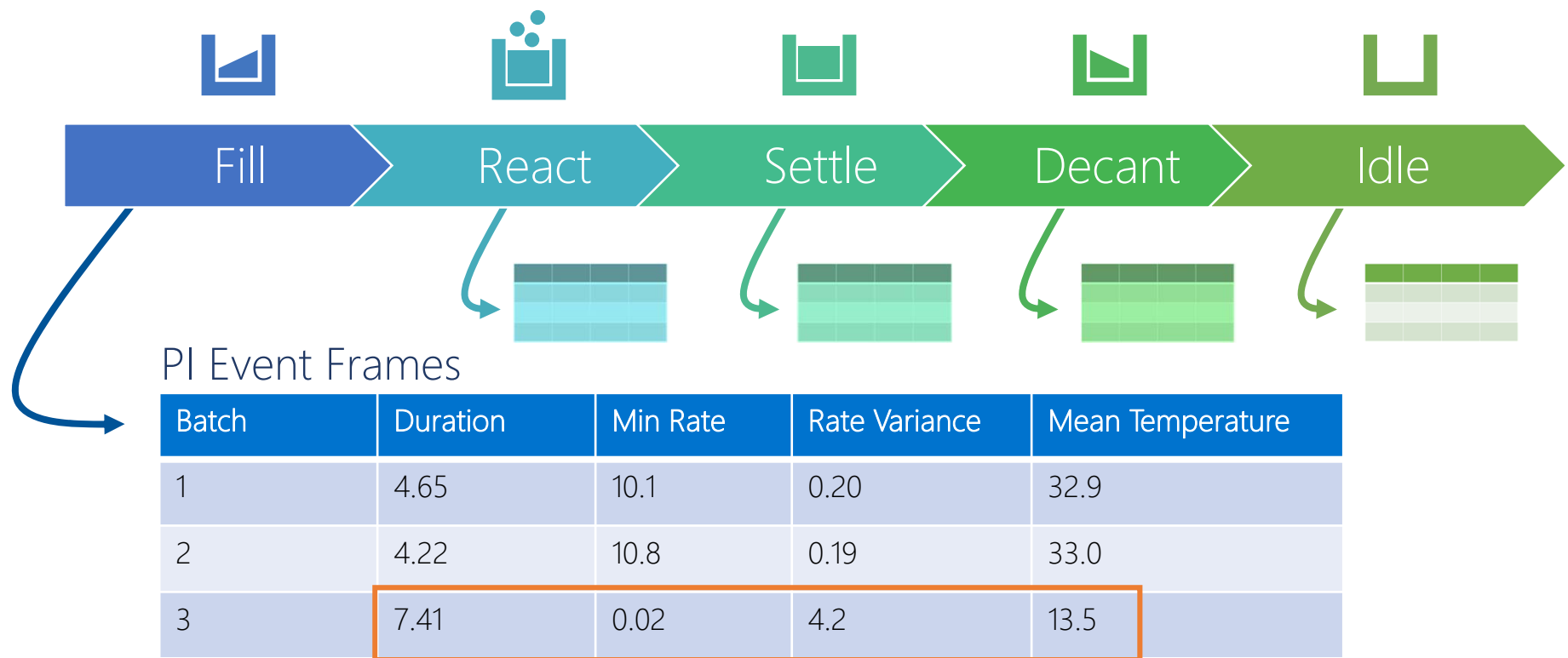
$$\frac{A + B + C + \dots}{n} = \boxed{-0.13}$$

*there are certainly times where event weighting is the right thing, but this choice should be made deliberately

Time-weighted

$$\frac{\overline{A} \cdot \Delta T_A + \overline{AB} \cdot \Delta T_{AB} + \dots}{\Delta T} = \boxed{4.38}$$

when | Events



HOW | Push, pull, cloudify

How | Push



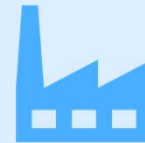
PI
Integrator



How | Pull

OCS
Data Views

REST
+ 



REST
+ 

OCS
Streams

REST
+ 

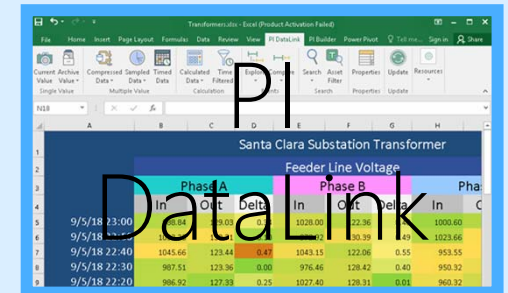
.NET

PI AF SDK

PI Web
API

*MORE IN NEXT TALK THIS TRACK!

UI



How | cCloudify



OCS

PI to OCS



PI System

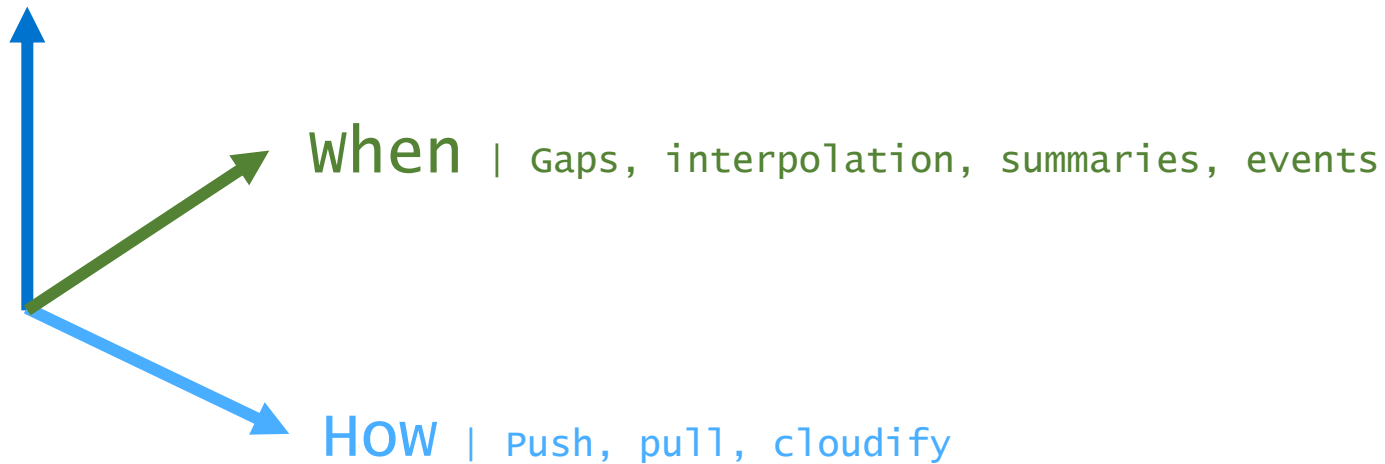
Data Views

REST + 

Streams

REST + 

what | Find the right things with metadata, assets



Takeaway | respect the axes

Contact information



- Brandon Perry
- Staff Software Developer
- OSIssoft, LLC
- bperry@osisoft.com



Questions?

Please wait for
the **microphone**

State your
name & company



Please remember to...

Complete Survey!

Navigate to this session in
mobile agenda for survey

