# Best Practices for Making Advanced Analytics Relevant

Curt Hertler

Principal Pre Sales Engineer
OSIsoft



# Workbench for Relevant Operational Analytics

### Data Engineering and Preparation

PI System offers distinctive features for preparing time-series data for advanced analytics, e.g. asset context, process context and feature generation.

### Access, Analysis and Model Enablement

PI System provides multiple data access methods, meeting needs of data engineers or scientists.

### Testing, Evaluation and Operationalization

- Asset Analytics plays an essential role in testing and evaluating developed models.
- PI Vision and Future Data support model integration and socialization for gaining relevance within Operations.

### Accountability

PI System provides verification of improvement benefits



# Data Engineering and Preparation



## Real Time Data is Different

- Transactional data is recorded in a tabular format with values associated by columns in each row.
- Real-time data is recorded with only time context, i.e. value and timestamp.
- Full-featured "observations" are required for analytics.



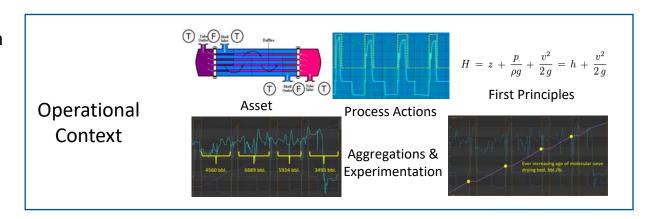
56.902 03-SEP-2016 11:23 AM

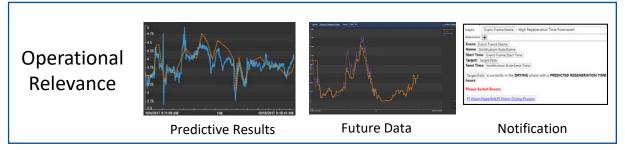


# Prepare Operational Data for Experimentation

- Sensor data unevenly timestamped
- Domain experts supply known relationships and impacting features
- Features and labeled events are added by backfilling in Asset Analytics
- Benefit for Operations and advanced analytics
- Prepared features persist for model operationalization and notification

63.781 3/9/16 11:19 AM 56.902 3/9/16 11:23 AM 58.341 3/9/16 11:41 AM

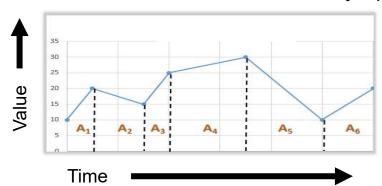






# Aggregations and Units of Measure Observance

### **Unevenly Spaced Events**



**Correct: Time-weighted** Average = 20.208

**Incorrect**: Arithmetic <u>10 + 20 + 15 + 25 + 30 + 10 + 20</u> = 18.571

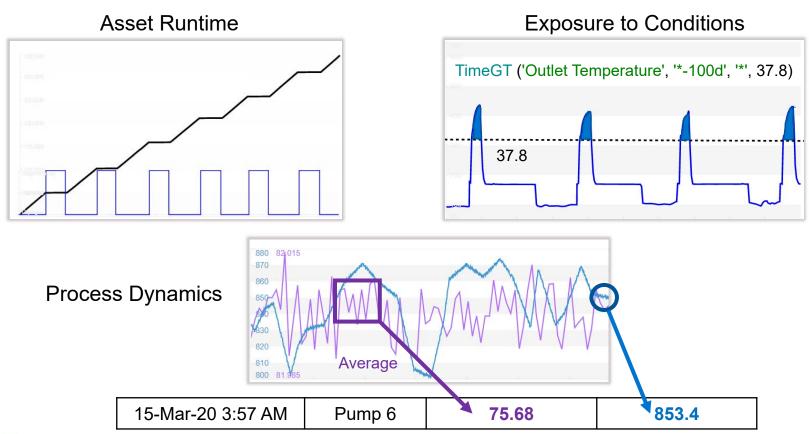
### Measured Rates Converted to Totals

Time Stamp	Value	Units	Gallons	
7/5/2017 8:00	8,828.5	gal/d	0.0	
7/5/2017 8:20	8,845.1	gal/d	122.8	
7/5/2017 8:40	8,861.6	gal/d	123.1	
7/5/2017 9:00	8,894.8	gal/d	123.5	
7/5/2017 9:20	9,045.2	gal/d	125.6	
7/5/2017 9:40	9,171.3	gal/d	127.4	
7/5/2017 10:00	9,199.9	gal/d	127.8	
	62,846.4		750.2	gal

**Correct: Unit Conversion** Total = 750.2

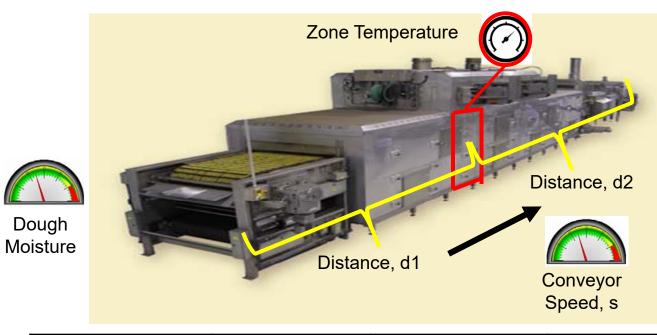
**Incorrect**: Arithmetic Sum = 62,846.4

# Time at State and at Conditions, Dynamics



#PIWorld

# Dynamics: The Way the Cookie Crumbles



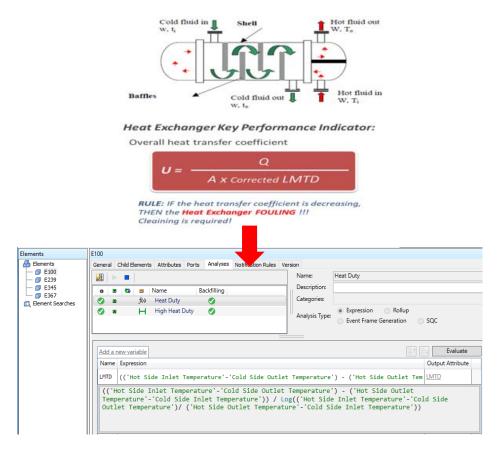
Coo	kie
Qua	lity

Modelling Phase	Moisture	Temperature	Quality
Training Observations	* - (d1 + d2) /s	* – d2 / s	*
Prediction	*	* + d1 / s	* + (d1 + d2) /s



# First Principles Analytics - Asset Analytics

- Configure calculations for transparency and scale
- Math, statistical, logical and steam table functions
- Predictive algebraic analytics
- Future data for forecasting
- Backfill! Backfill! Backfill!



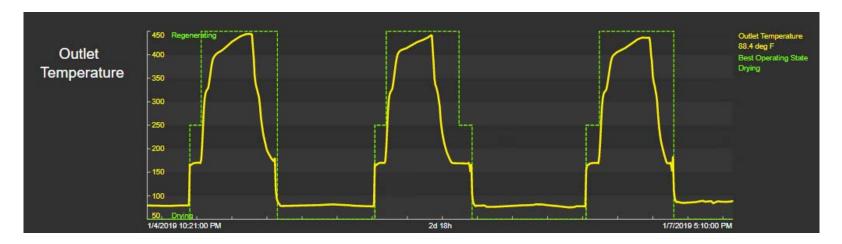


# Label Time Ranges of Interest – *Event Frames*

Filter				
□ 🖥 Name	Duration	Start Time	End Time	4
■ 🖈   Dryer A Regeneration Cycle 01-04-19 03:06	9:54:00	1/4/2019 3:06:00 AM	1/4/2019 1:00:00 PM	
🗷 🖈 💳 Dryer A Regeneration Cycle 01-05-19 04:30	8:42:00	1/5/2019 4:30:00 AM	1/5/2019 1:12:00 PM	
■ 🖈 Impryer A Regeneration Cycle 01-06-19 01:36	8:18:00	1/6/2019 1:36:00 AM	1/6/2019 9:54:00 AM	

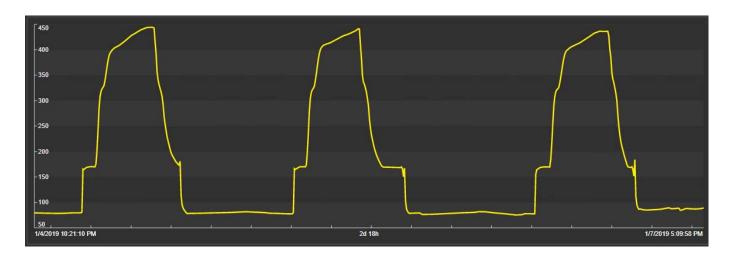








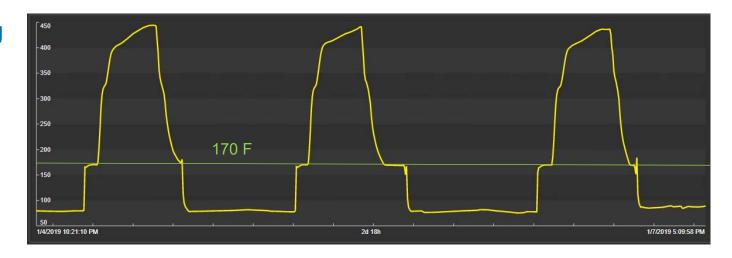
Data Engineering Process





Data Engineering Process

**First Backfill** 



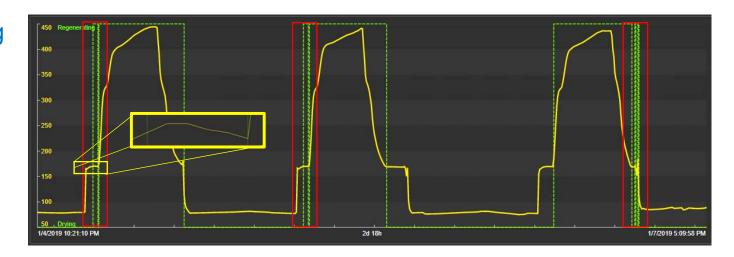
Asset Analytics If 'Outlet Temperature' >= 170
Then "Regenerating"
Else "Process"



# Data Engineering Process

### First Backfill

"False Starts"



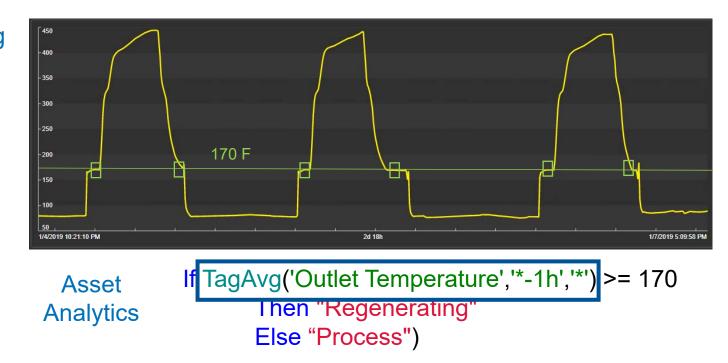


# Data Engineering Process

### First Backfill

"False Starts"

### **Second Backfill**





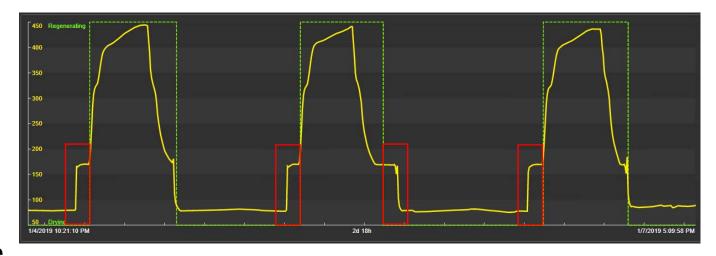
# Data Engineering Process

### First Backfill

"False Starts"

### Second Backfill

Missing third "Standby" State





# Data Engineering Process

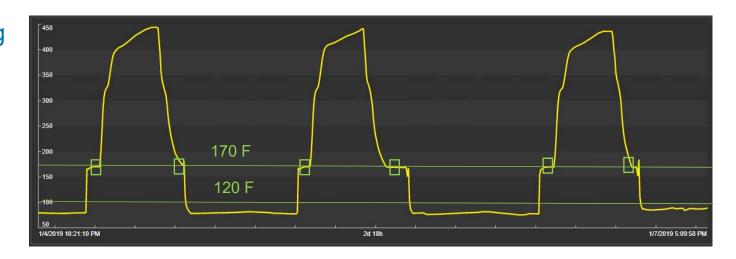
### First Backfill

"False Starts"

### Second Backfill

Missing "Standby" State

### **Third Backfill**



```
Asset Analytics
```

```
If TagAvg('Outlet Temperature','*-1h','*') >= 170
Then "Regenerating"

Else (If 'Outlet Temperature' >= 120
Then "Standby"
Else "Process")
```



# Data Engineering Process

### First Backfill

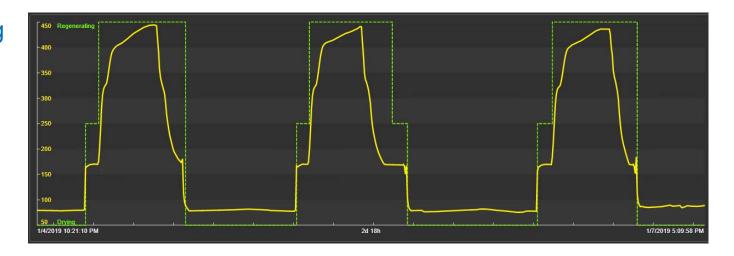
"False Starts"

### Second Backfill

Missing "Standby" State



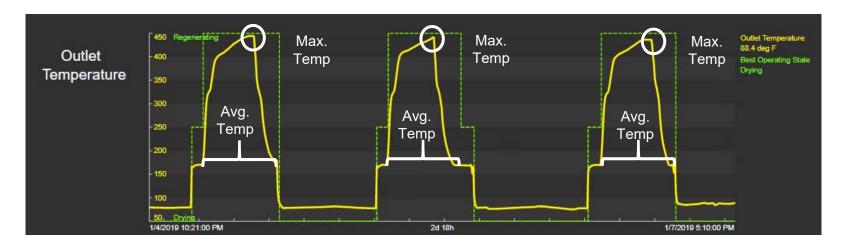
Use it!





# Complete Observation with Event Frame Attributes

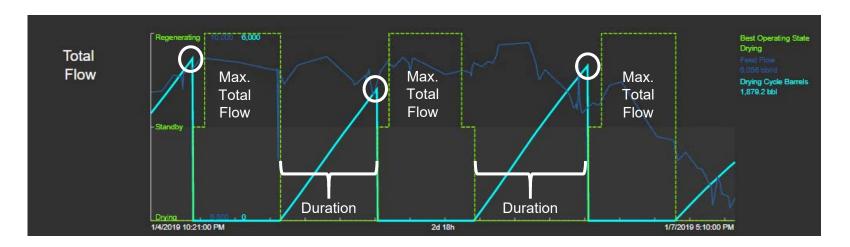
Filter					
■ Name	Duration	Start Time	End Time	Avg Outlet Temp	Max Outlet Temp
🗷 🖈   Dryer A Regeneration Cycle 01-04-19 03:06	9:54:00	1/4/2019 3:06:00 AM	1/4/2019 1:00:00 PM	324.4 deg F	444.3 deg F
🗷 🖈   Dryer A Regeneration Cycle 01-05-19 04:30	8:42:00	1/5/2019 4:30:00 AM	1/5/2019 1:12:00 PM	346.9 deg F	444.3 deg F
■ 🖈 ├─ Dryer A Regeneration Cycle 01-06-19 01:36	8:18:00	1/6/2019 1:36:00 AM	1/6/2019 9:54:00 AM	339.0 deg F	441.3 deg F





# Complete Observation with Event Frame Attributes

Filter							
□ 🗟 Name	Duration	Start Time	End Time	△   Avg Outlet Temp	Max Outlet Temp	Drying Cycle Duration	Drying Cycle Barrel
🗷 🖈 💳 Dryer A Regeneration Cycle 01-04-19 03:06	9:54:00	1/4/2019 3:06:00 AM	1/4/2019 1:00:00 PM	324.4 deg F	444.3 deg F	13.8 h	5259.6 bbl
🗷 🖈 💳 Dryer A Regeneration Cycle 01-05-19 04:30	8:42:00	1/5/2019 4:30:00 AM	1/5/2019 1:12:00 PM	346.9 deg F	444.3 deg F	14.0 h	5189.7 bbl
■ 🖈 Impryer A Regeneration Cycle 01-06-19 01:36	8:18:00	1/6/2019 1:36:00 AM	1/6/2019 9:54:00 AM	339.0 deg F	441.3 deg F	11.1h	4207.6 bbl





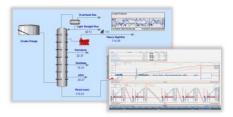
# Access, Analysis and Model Enablement



# **Enabling Analytics for Operational Intelligence**

Real-Time Decision Analysis

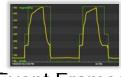
Retrospective & Predictive Analysis



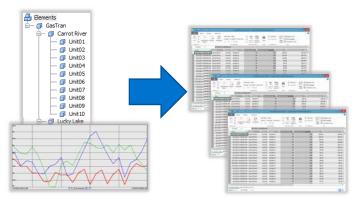
Diagnostic
Trending & Awareness

$$Q=rac{\Delta P_{DD}*kh}{141.2\mu B_0\Big\{lnrac{r_e}{r_w}-rac{3}{4}+S\Big\}}$$

Descriptive
Condition & Performance



**Event Frames** 



Time, Event and Asset Context

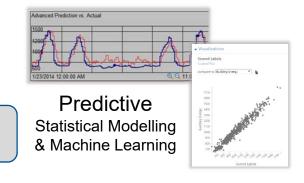
Tabular Context

Common Ground between Technological Contexts



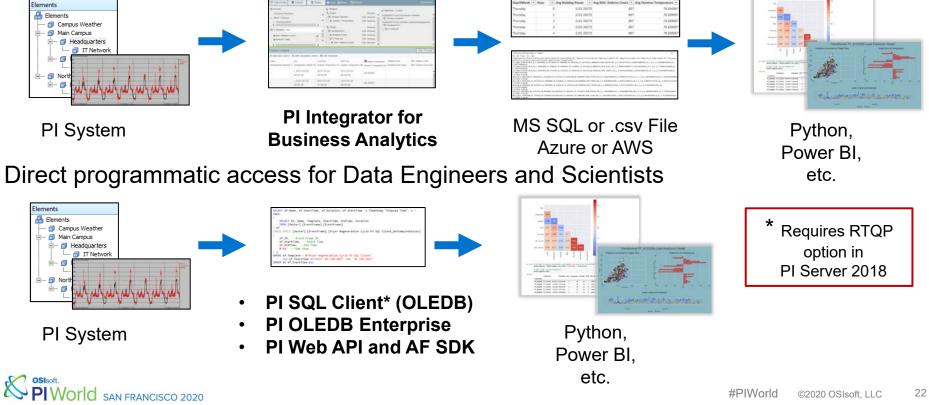
Visual
Dashboards &
Multidimensional Assessment

#PIWorld



# Open Data Access for Advanced Analysis Tools

### Self-service access for Everyone



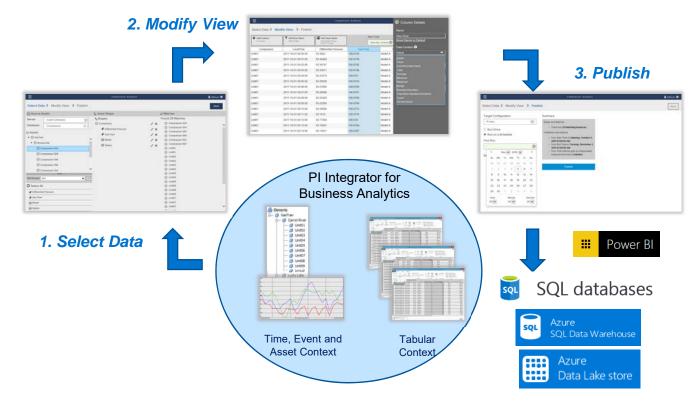
#PIWorld

Easy, scalable way for users to create contextualized views of operational data.

- Select assets and their attributes from an AF hierarchy.
- Modify view by setting time range, row interval, and column aggregations.

**Add** filtering rules to "cleanse" data.

 Publish once or on a scheduled bases.



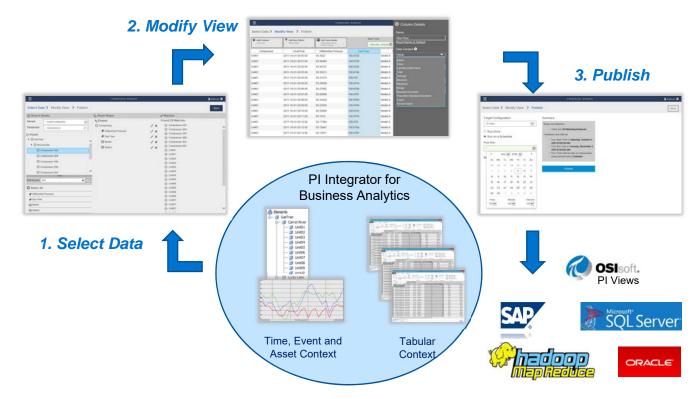


Easy, scalable way for users to create contextualized views of operational data.

- Select assets and their attributes from an AF hierarchy.
- Modify view by setting time range, row interval, and column aggregations.

**Add** filtering rules to "cleanse" data.

 Publish once or on a scheduled bases.



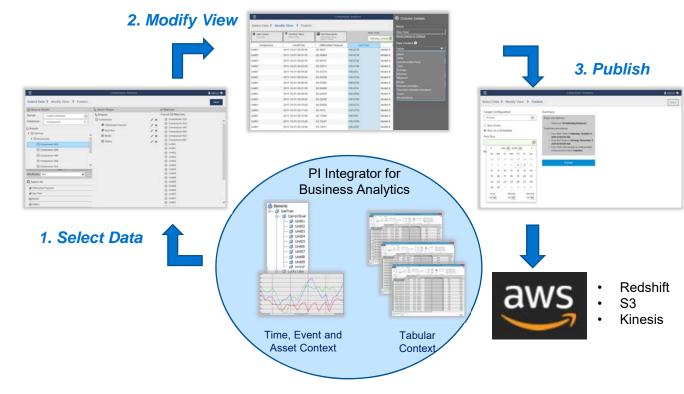


Easy, scalable way for users to create contextualized views of operational data.

- Select assets and their attributes from an AF hierarchy.
- Modify view by setting time range, row interval, and column aggregations.

**Add** filtering rules to "cleanse" data.

 Publish once or on a scheduled bases.



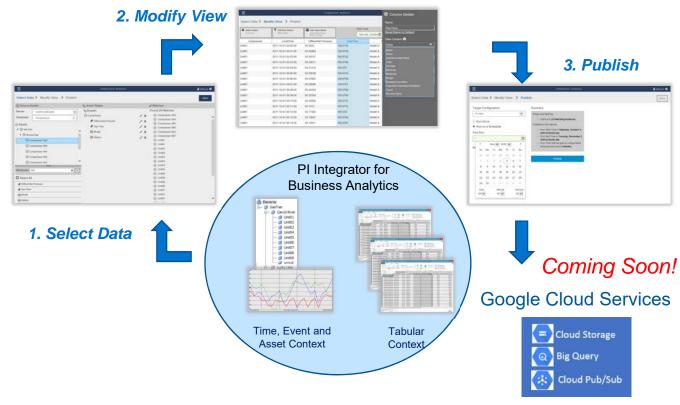


Easy, scalable way for users to create contextualized views of operational data.

- Select assets and their attributes from an AF hierarchy.
- Modify view by setting time range, row interval, and column aggregations.

**Add** filtering rules to "cleanse" data.

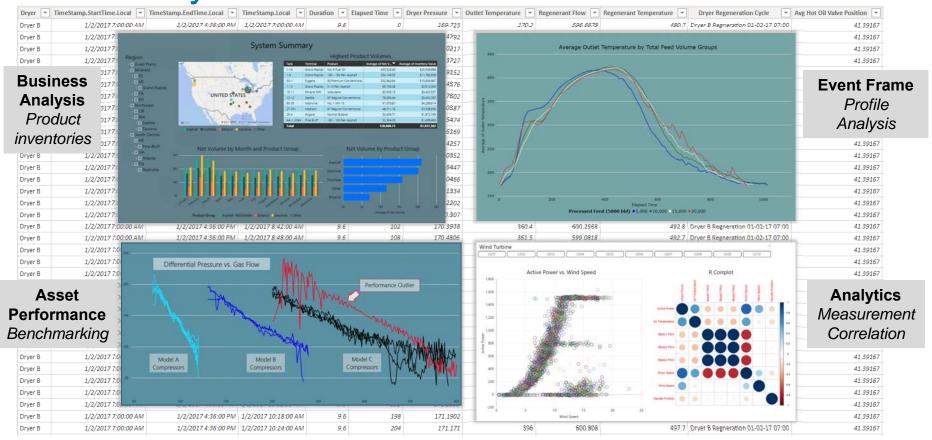
 Publish once or on a scheduled bases.





#PIWorld

# Visual Analytics – Multidimensional Tools



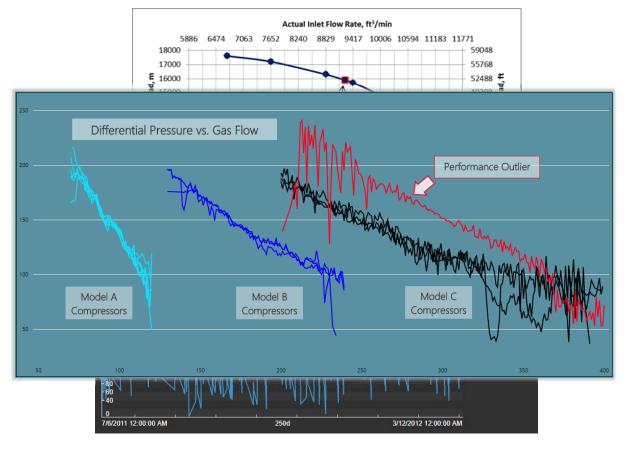


27

# Visual Analytics – Asset Benchmarking

Benchmark <u>all</u> similar assets against know performance characteristics.

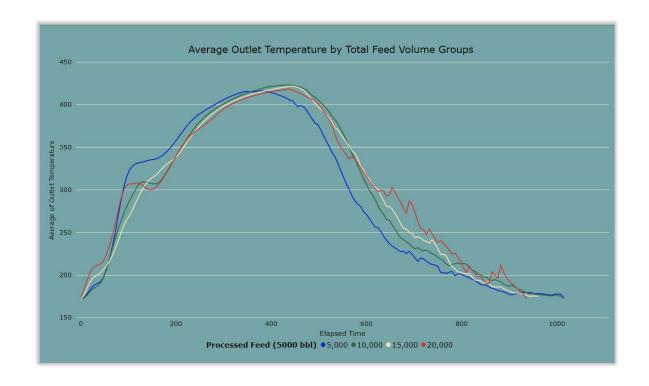
- Real-time trend of one asset is essential for current operation.
- Different tools required to analyze groups of assets.
- PI Asset Views summarize months of actual operations exposing actual performance profile.





# Visual Analytics - Event Frame Evaluation

- Sampled Event View dataset imported into Power BI.
- Shows 200 Event Frames grouped by Event Frame Attribute.





# Sampled Event View

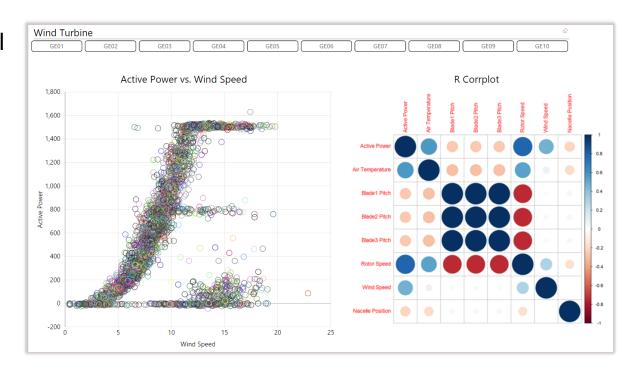
	Dryer	Time Stamp	Duration	Elapsed Time	Dryer Pressure	Outlet Temperature	e Regenerant Flow	Regenerant Tempe	Dryer Regeneration Cycle	Avg Outlet Temp	Avg Regen Temp	Dryer Processing Age	Total Pro
_	Dryer A	1/2/2017 12:00:00 AM	5.3	0	170.4697	436.9	603.2525	496.1	Dryer A Regneration 01-02-17 00:00	319.9179	229.7132	0.0005072668	4474.716
- [	Dryer A	1/2/2017 12:06:00 AM	5.3	6	170.4824	437.8	603.6448	496.2	"	"	"	"	"
	Dryer A	1/2/2017 12:12:00 AM	5.3	12	170.495	438.6	605.063	496.4		"	"	"	
	Dryer A	1/2/2017 12:18:00 AM	5.3	18	170.5076	439.5	599.6411	496.8	<u>"</u>	"	"		
				:						"		"	"
		:		:	:	•	:	:			"	"	"
١	Dryer A	1/2/2017 5:18:00 AM	5.3	318	204.1571	173.8	314.5962	287.9	"	"	"	"	"
-	Drver B	1/2/2017 7:00:00 AM	9.6	0	169.723	170.2	596.6678	480.7	Dryer B Regneration 01-02-17 07:00	357.6974	404.2589	0.05128649	2742.962
- 1	Dryer B	1/2/2017 7:06:00 AM	9.6	6	170.4792	170.3	598.8013	484.8					
	Dryer B	1/2/2017 7:12:00 AM	9.6	12	171.0217	170.4	597.2024	487.9	"	"	"	"	"
ľ									"	"	"	"	"
				•	•	•		:	"	"	"	"	"
		•	•	•	•	•	•	•	"	"	"	"	"
	D D	1/2/2017 4:36:00 PM	9.6	576	168.8051	174.9	896.8521	173.8	"	"	"	"	"
	Dryer B Dryer A	1/2/2017 4:36:00 PM	11.3	0	168.8051	174.9	596.4086	173.8 425.6	Dryer A Regneration 01-02-17 19:00	332.5292	364.0818	0.1004348	5302.277
- 1	Dryer A	1/2/2017 7:06:00 PM	11.3	6	170.3642	170.1	598.5709	442.7	Diyer A Regileration 01-02-17 13.00	002.0232	004.0010	0.1004040	0002.211
- 1	Dryer A	1/2/2017 7:12:00 PM	11.3	12	170.2456	170.2	600.7331	466.3	"	"	"	"	"
	Differ / C	172720 11 1.12.00 1 M	11.0	12	170.2400	170.2	000.7001	400.0	"	"	"	"	"
			: •	: .	: :	: :			"	"	"	"	"
		:	:	:	:	:	:	:	"	"	"	"	"
					:		:		"	"	"	"	"
_	Dryer A	1/3/2017 4:54:00 AM	11.3	594	167.147	196.4	547.8572	173.4	:	:	:	:	-
- 1	Dryer B	1/3/2017 8:48:00 AM	9.3	0	169.3378	170.2	592.8909	491.6	Dryer B Regneration 01-03-17 08:48	357.8753	399.4613	0.1699348	6347.083
- 1	Dryer B Dryer B	1/3/2017 8:54:00 AM 1/3/2017 9:00:00 AM	9.3	6 12	168.1517 167.7706	171.1 205.2	617.2303 617.502	490 488.7	"	"	"	"	"
	Diyei B	1/3/2017 9.00.00 AW	9.3	12	167.7706	205.2	617.302	400.7	"	"	"	66	"
- (				: .		: :	: .	:	"	"	"	66	"
١		•	•	•	•	•			<b>\</b>	"	"	"	"
	\							/	\	"	"		"

San Francisco 2020

#PIWorld ©2020 OSIsoft, LLC

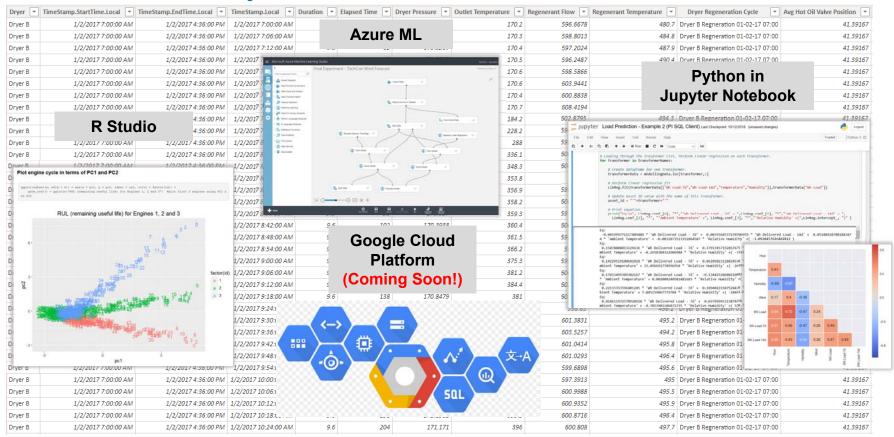
# Visual Analytics - R and Python Integration

- Ad hoc, multidimensional front end for driving R and Python scripts.
- R "corrplot" to identify correlated variables.
- Identify data to be used for predictive model training.





# Advanced Analytics - R Studio, Python, Azure ML, GCP





#PIWorld ©2020 OSIsoft, LLC

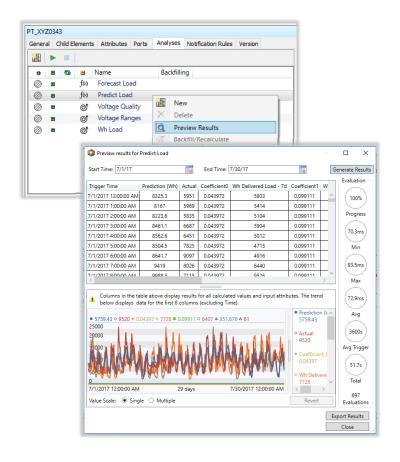
# Testing, Evaluation and Operationalization



# Testing and Evaluation: Asset Analytics

### **Preview Results**

feature of Asset Analytics generates model results without posting values to PI Points



1	Α	В	С	D
1	Trigger Time	Prediction (Wh)	Actual	
2	7/1/2017 0:00	8325.3	5951	
3	7/1/2017 1:00	8167	5969	
4	7/1/2017 2:00	8223.6	5835	
5	7/1/2017 3:00	8461.1	6687	
6	7/1/2017 4:00	8562.6	6451	
7	7/1/2017 5:00	8504.5	7825	
8	7/1/2017 6:00	8641.7	9097	
9	7/1/2017 7:00	9419	8026	
10	7/1/2017 8:00	9988.5	7115	
11	7/1/2017 9:00	10458	11052	
12	7/1/2017 10:00	11710	13563	
13	7/1/2017 11:00	11126	11933	
14	7/1/2017 12:00	11666	11642	
15	7/1/2017 13:00	12935	10199	

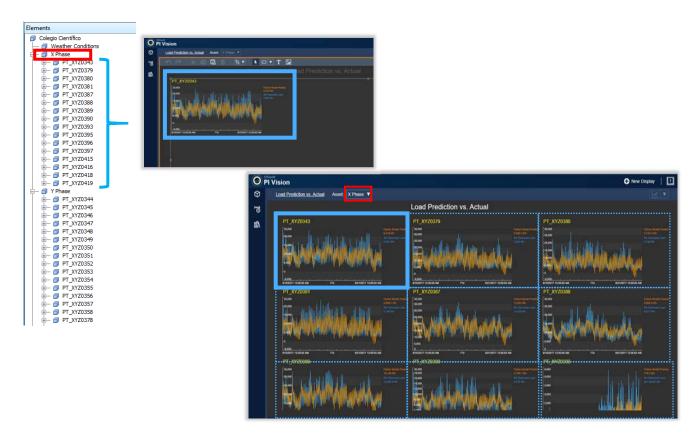
Export results for analysis in Excel.



# Testing and Evaluation: Model Prediction vs. Actual

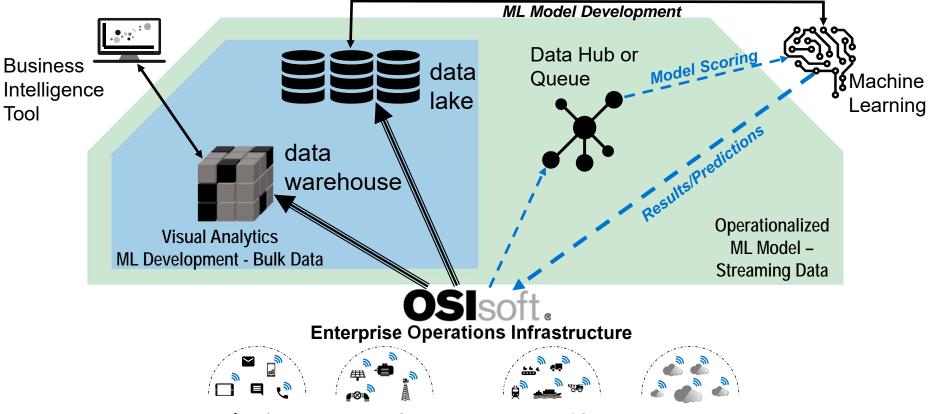
Backfill model results into PI Point

PI Vision
Collections
to inspect
model results





# Operationalize - Advanced Analytics Patterns



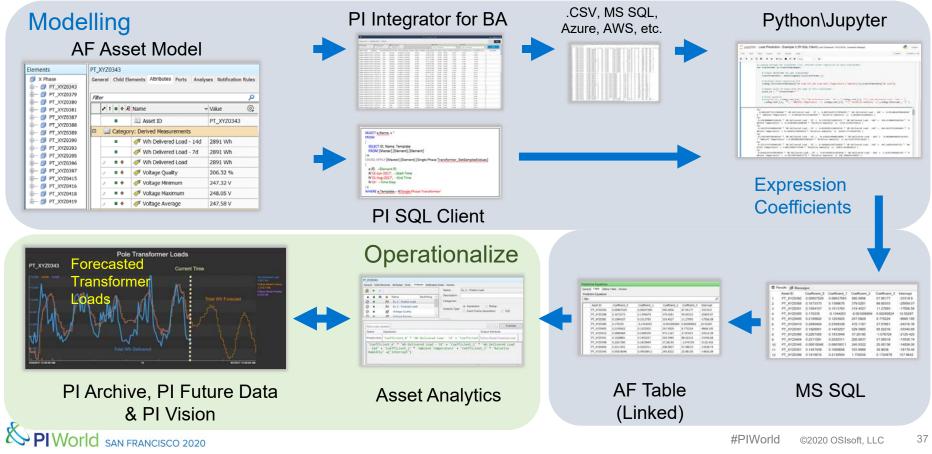
**Assets** 

Automation Systems Edge Devices / Sensors IoT solutions



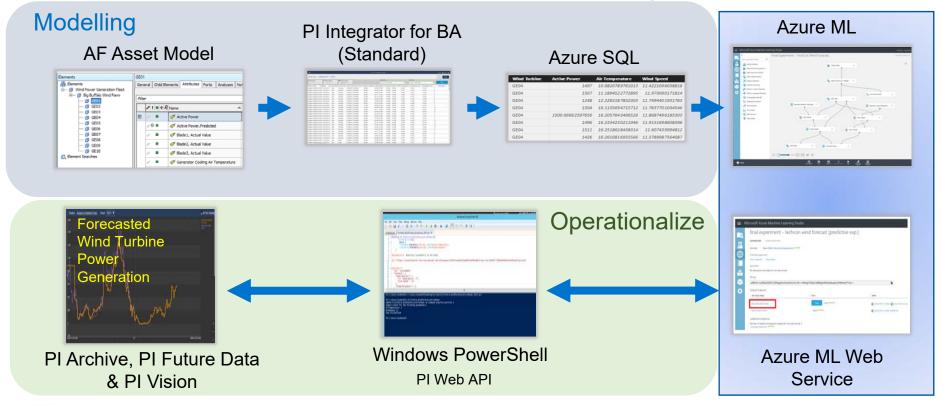
## Example 1 – Expression-based Model

PI World 2019 - "Exploring AF Analytics for Advanced Analysis and Prediction"



## Example 2 – Web Service Endpoint Model

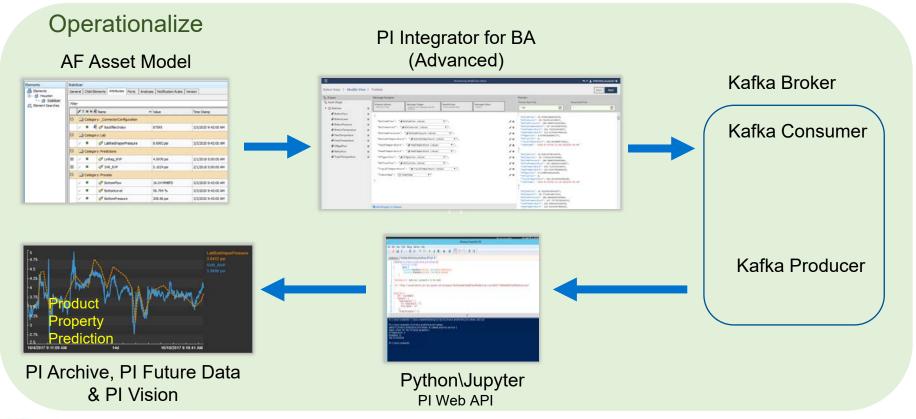
PI World 2017 - "Create and Operationalize Forecasting Models with the PI Infrastructure and Azure Machine Learning"





# Example 3 – Streaming Analytic

PI World 2018 - "Apply Predictive Machine Learning Models to Operations"



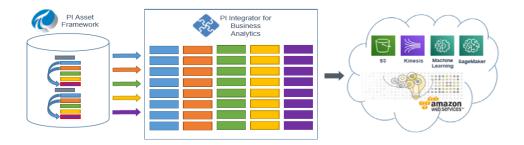


# TransCanada – Gas Pipeline Demand Forecasting

### Navigating the Sea of Data



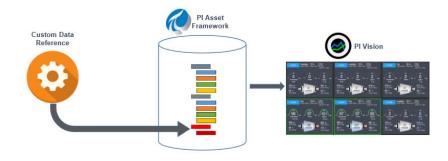
### Feeding the Machine



### Automating the Demand Forecast



### Consuming the Results

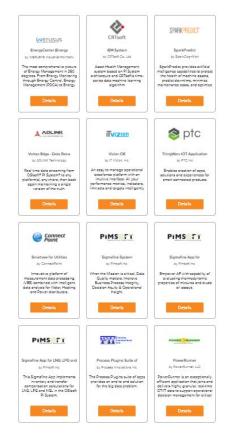


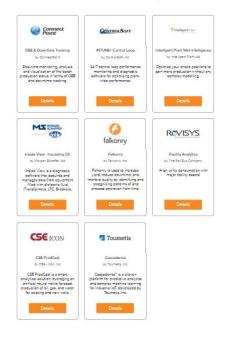
https://www.osisoft.com/Presentations/TransCanada-s-Journey-to-Advanced-Analytics---Integrating-TransCanada-s-PI-AF-with-AWS-Machine-Learning/



# OSIsoft Partner Ecosphere – "Advanced Analytics"







https://www.osisoft.com/marketplace/



# Accountability



# "PI Don't Lie"





# Workbench for Relevant Operational Analytics

### Data Engineering and Preparation

PI System offers distinctive features for preparing time-series data for advanced analytics, e.g. asset context, process context and feature generation.

### Access, Analysis and Model Enablement

PI System provides multiple data access methods, meeting needs of data engineers or scientists.

### Testing, Evaluation and Operationalization

- Asset Analytics plays an essential role in testing and evaluating developed models.
- PI Vision and Future Data support model integration and socialization for gaining relevance within Operations

### Accountability

"PI don't lie."



# **Contact Info**



- Curt Hertler
- Principal Pre-Sales Engineer
- OSIsoft, LLC
- Curt@osisoft.com



# Questions?

Please wait for the **microphone** 

State your name & company

### Save the Date...



AMSTERDAM October 26-29, 2020





# KÖSZÖNÖM MULŢUMESC GO RAIBH MAITH AGAT NATION OF THE STATE OF ДЗЯКУЙ TAKK SKAL DU HA **MERC RAHMAT** MATUR NUWUN CẨM ƠN BẠN **UATSAUG RAU KOJ**

