OSIsoft.

Regional Summit 2017

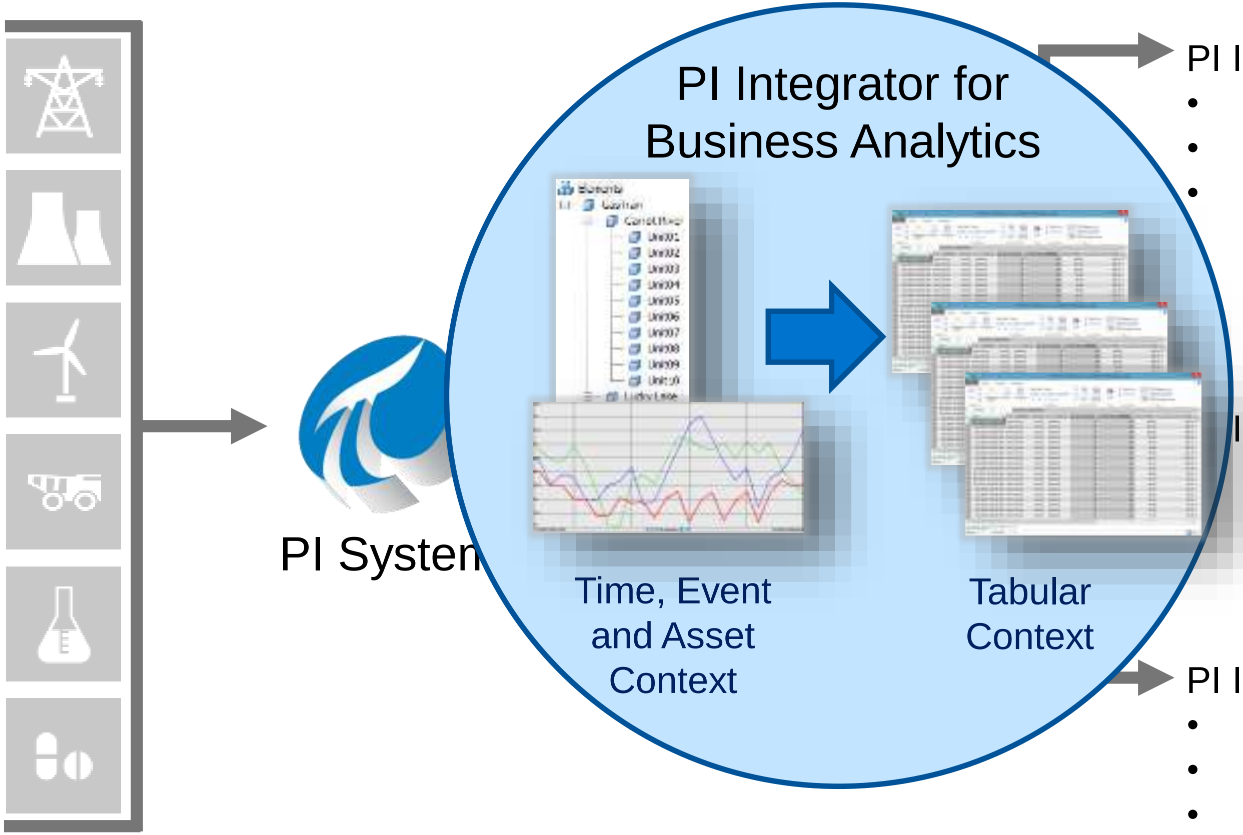
May 2-4, 2017 | West Palm Beach, FL

Your Big Data to Big Data tools using the family of PI Integrators

Presented by **Martin Bryant**

Field Service Engineer

PI Integrators



PI System

PI Integrator for Business Analytics

Time, Event and Asset Context

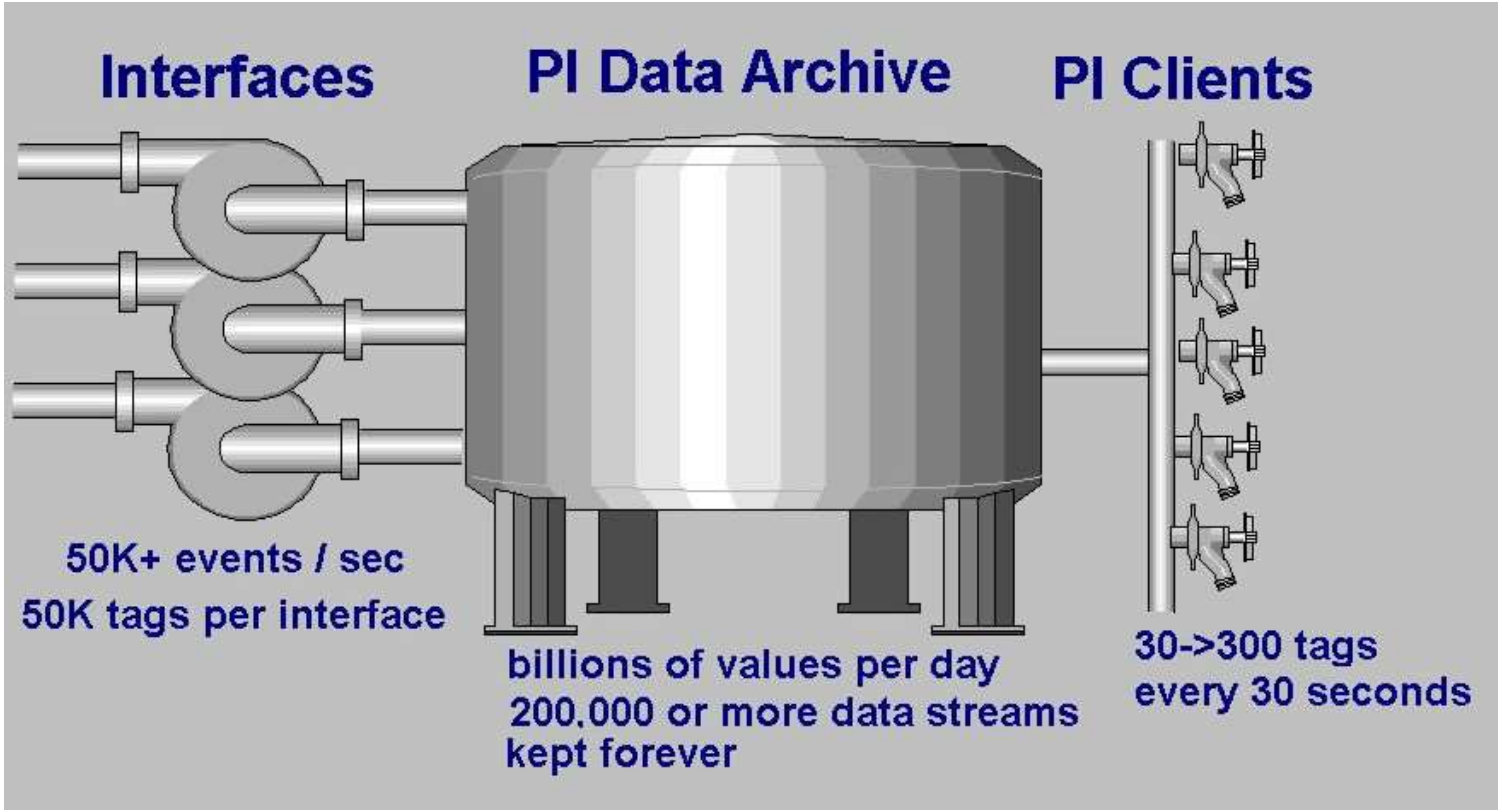
Tabular Context

PI Integrator for Business Analytics
- SQL Server
- Hadoop (HIVE & HDFS)
- Oracle RDBMS
- File
- ODBC client

Integrator for SAP HANA
- SAP HANA
- Smart Data Integration (SDI)

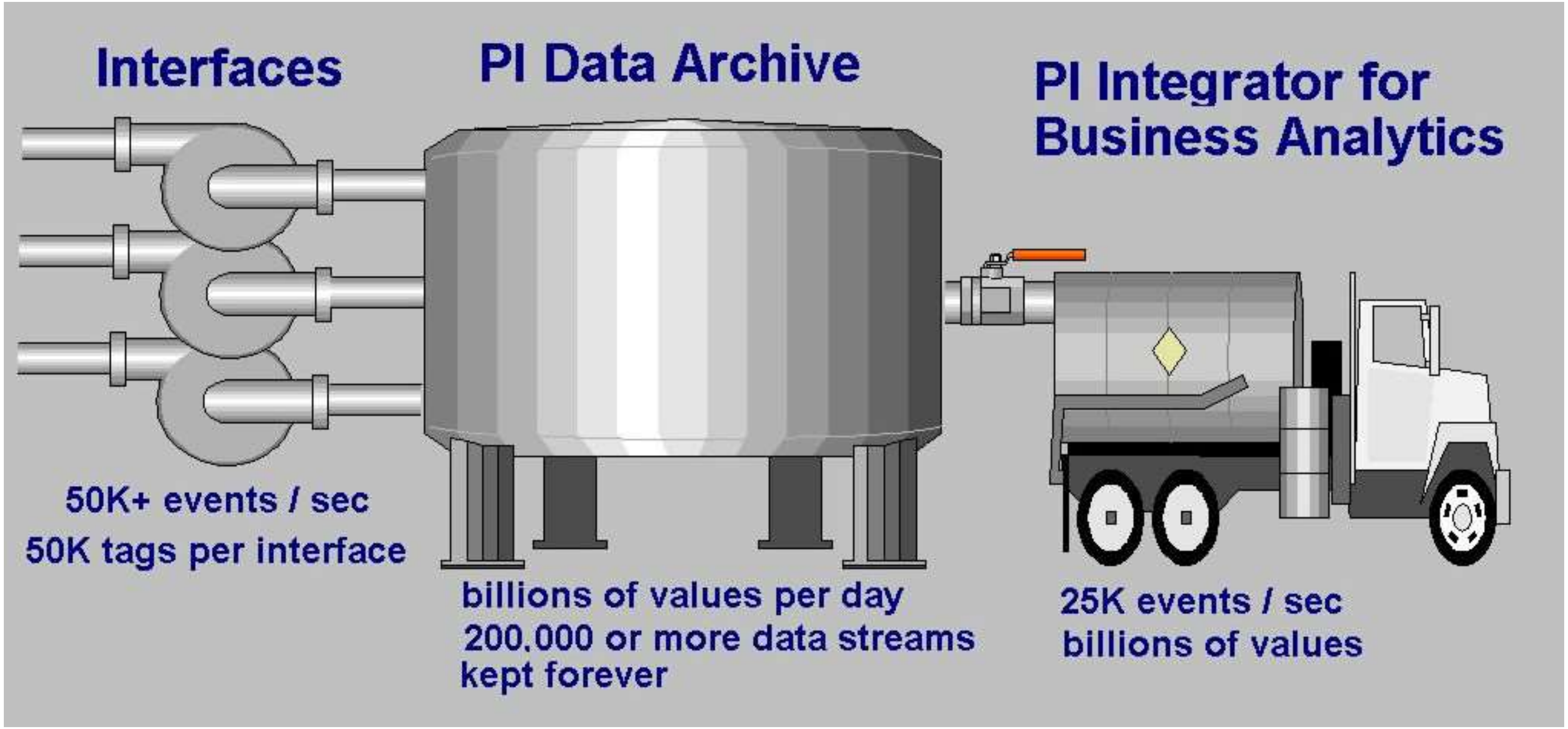
PI Integrator for Microsoft Azure
- SQL Data Warehouse
- Azure SQL
- Azure Data Lake

The Traditional PI System



Interfaces
PI Data Archive
PI Clients
50K+ events / sec
50K tags per interface
billions of values per day
200,000 or more data streams
kept forever
30->300 tags
every 30 seconds

An Alternative Data Use Model



Interfaces

PI Data Archive

PI Integrator for Business Analytics

50K+ events / sec
50K tags per interface

billions of values per day
200,000 or more data streams
kept forever

25K events / sec
billions of values

The world of Big Data tools…

Visual Analytics

Excel

Spotfire
TIBCO Software

SAP BusinessObjects

COGNOS

esri

ORACLE
BUSINESS INTELLIGENCE

SAP

tableau
SOFTWARE

Power BI

SAP Lumira

QlikView

Data Warehouse / Data Lake

SQL

SQL

Microsoft
SQL Server

ORACLE
DATABASE
12c

SAP HANA

BIG DATA

Streaming Analytics – 2017

IoT Hub

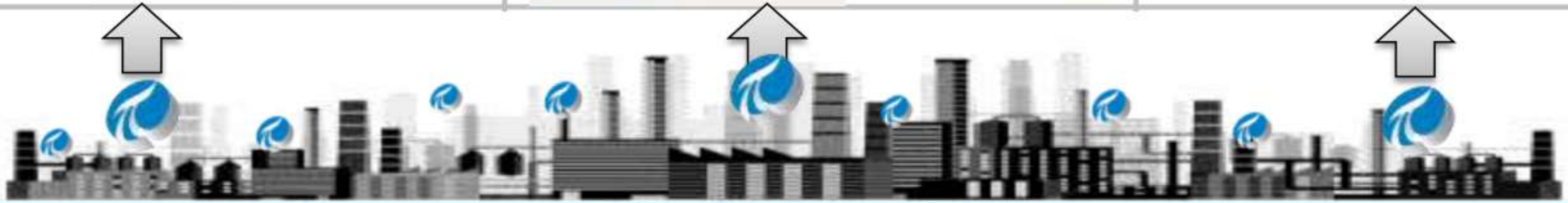
Machine Learning

Stream Analytics

Spark

kafka

SAP



Getting Value…

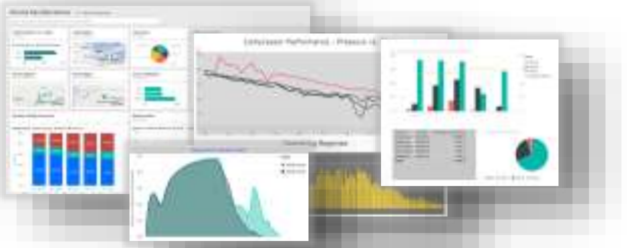
Solving complex problems for a fleet	**Dashboarding – visual reporting – real time & mobile**	**Integration to new I.T. projects and databases**
Multivariate, other statistics & machine learning resources One time answers or Running models 	Drill drown, rollup Anywhere anytime TIBCO Spotfire Power BI tableau	The right way to bring operational data to I.T.'s Big Data party… Microsoft SQL Server Reporting Services hadoop ORACLE BUSINESS INTELLIGENCE SAP

Why? … complex problem solving

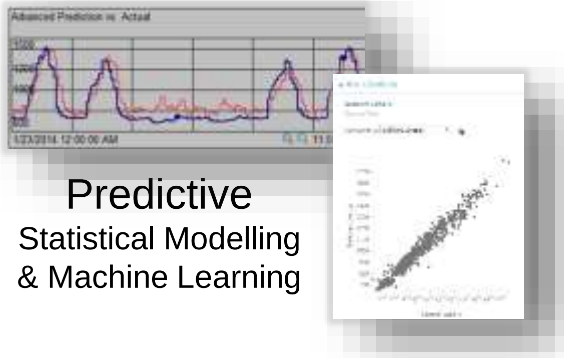
Statistical Analytics

- Identifying **patterns** and discovering problems through statistical methods that require **large** and **diverse** datasets

All PI Customers have a great deal of data.. Often too much for effective analysis in a spreadsheet. This data can be complex with many variables. Multivariable analysis and machine learning can provide interesting and invaluable answers to complex, big questions…



Multidimensional
Business Intelligence
& Dashboards



Predictive
Statistical Modelling
& Machine Learning

PI data is very large and complex.

This is about finding the answer to large, complex questions.

Complex Analyses Increase the Need for Deeper Integration

Complexity

Disparate assets
Interacting with assets on an **individual** basis

Interacting with **common** assets as a **fleet**

Monitoring

Real-time visibility

- Traditional HMI

Process Optimization

Real-time & historical views across any asset

- PI ProcessBook
- PI Coresight
- PI Datalink

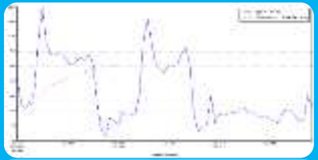
Benchmarking

Fleet-wide performance comparisons

- BI app (e.g. Tableau, Spotfire, Lumira)
- PI Integrator for Business Analytics
- PI Integrator for SAP HANA

System Optimization

Large scale multi-variate analysis



- Machine learning (Azure ML, R)
- PI Integrator for Business Analytics
- PI Integrator for SAP HANA

Applying Data to Maintenance…

Disparate assets
Interacting with assets on an **individual** basis

Interacting with **common** assets as a **fleet**

Complexity

Monitoring

Wait for failure

faster break & fix

Data aware
Usaged based

Using run times and not clock hours to schedule

Preventive Maintenance

and inspections

Performance

Compare current performance metrics, including vibration, temperature, & other KPIs to expected values

Predictive

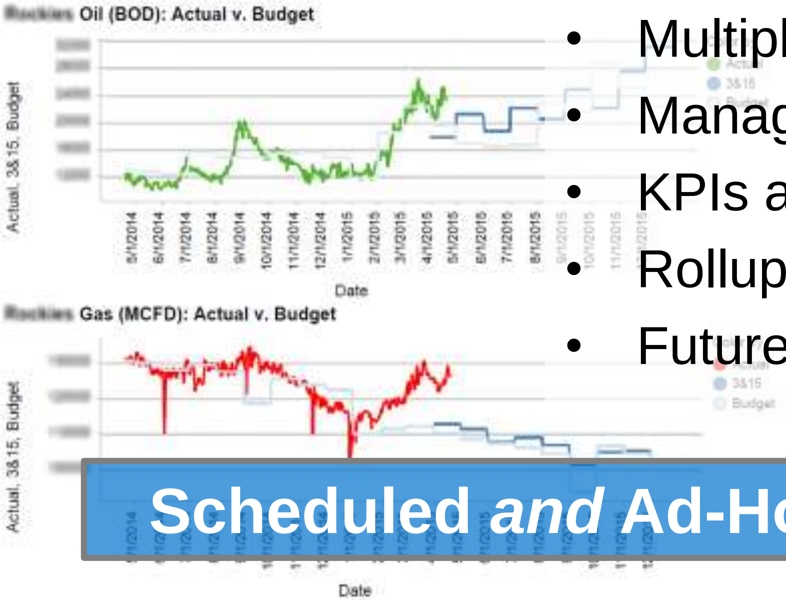
Based on the history of like equipment and circumstances – build a model that predicts failure and triages maintenance

Why? Dashboarding

- Automated daily reports
- Multiple data source views
- Management overviews
- KPIs and metrics
- Rollups by business unit
- Future Data (forecast)

Visual Analytics

- **Visualizing diverse** sets of data sources to gain insights, create reports, and improve operations



Scheduled *and* Ad-Hoc Reporting

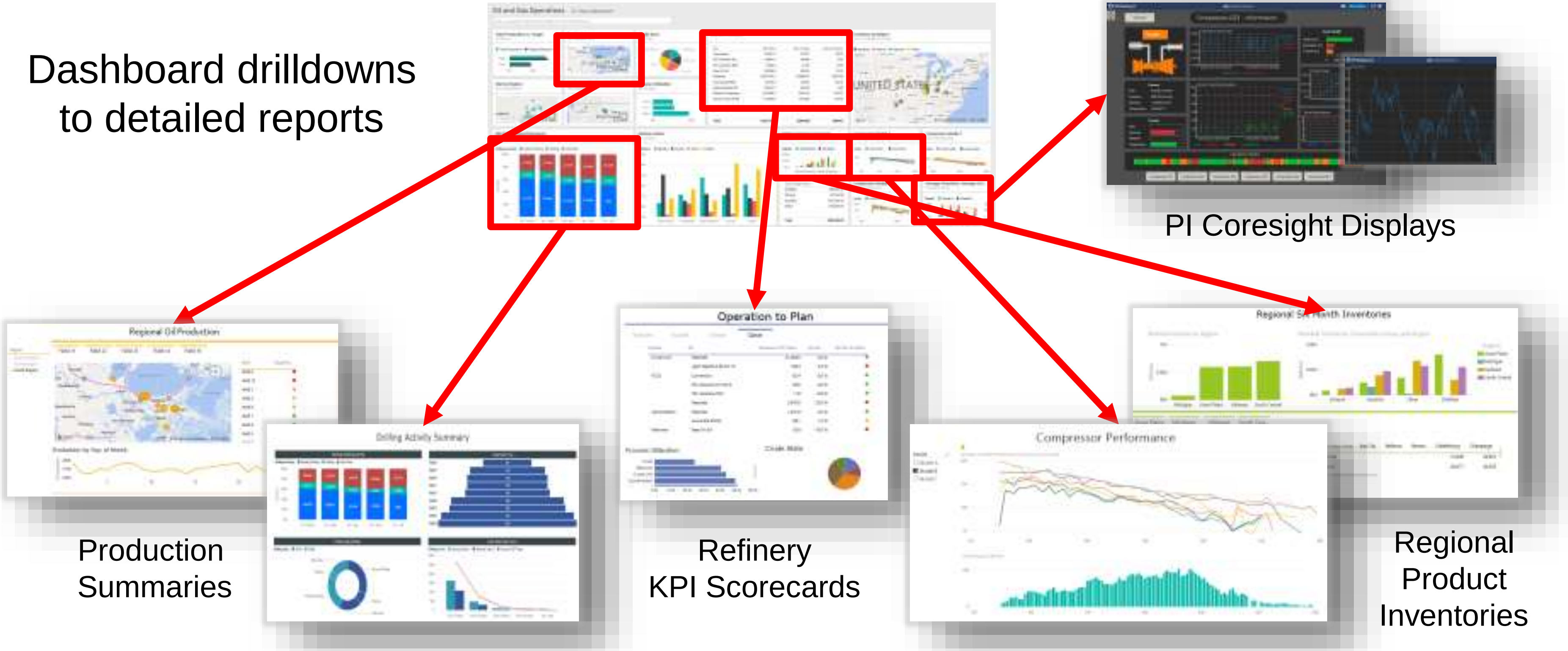
Don't wait for tomorrow to explore performance today

your operational data is a click or swipe away.

Even from your mobile devices

Result: Improved, Detailed Reporting and Analytics

Dashboard drilldowns to detailed reports


Oil and Gas Operations
Regional Oil Production
Drilling Activity Summary
Operation to Plan
Crude Slate
Compressor Performance
Regional Six Month Inventories

PI Coresight Displays

Production Summaries

Drilling Activity Reports

Refinery KPI Scorecards

Asset Performance Displays

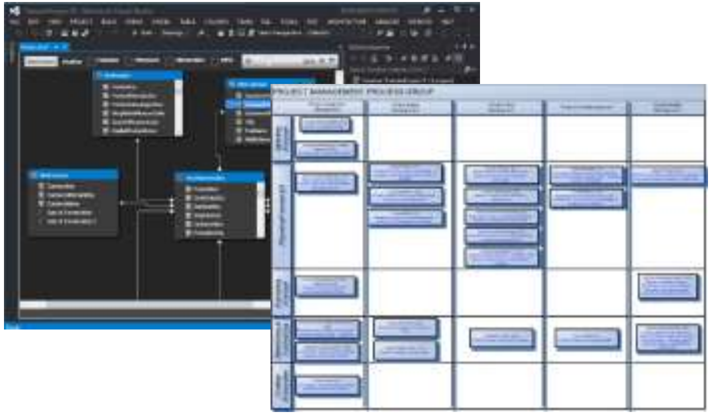
Regional Product Inventories

Why? I.T. Integration projects

Information systems have projects that provide enormous value – and those can benefit from the real-time process awareness of PI Data. But PI Data hasn't been easy to integrate – until now.

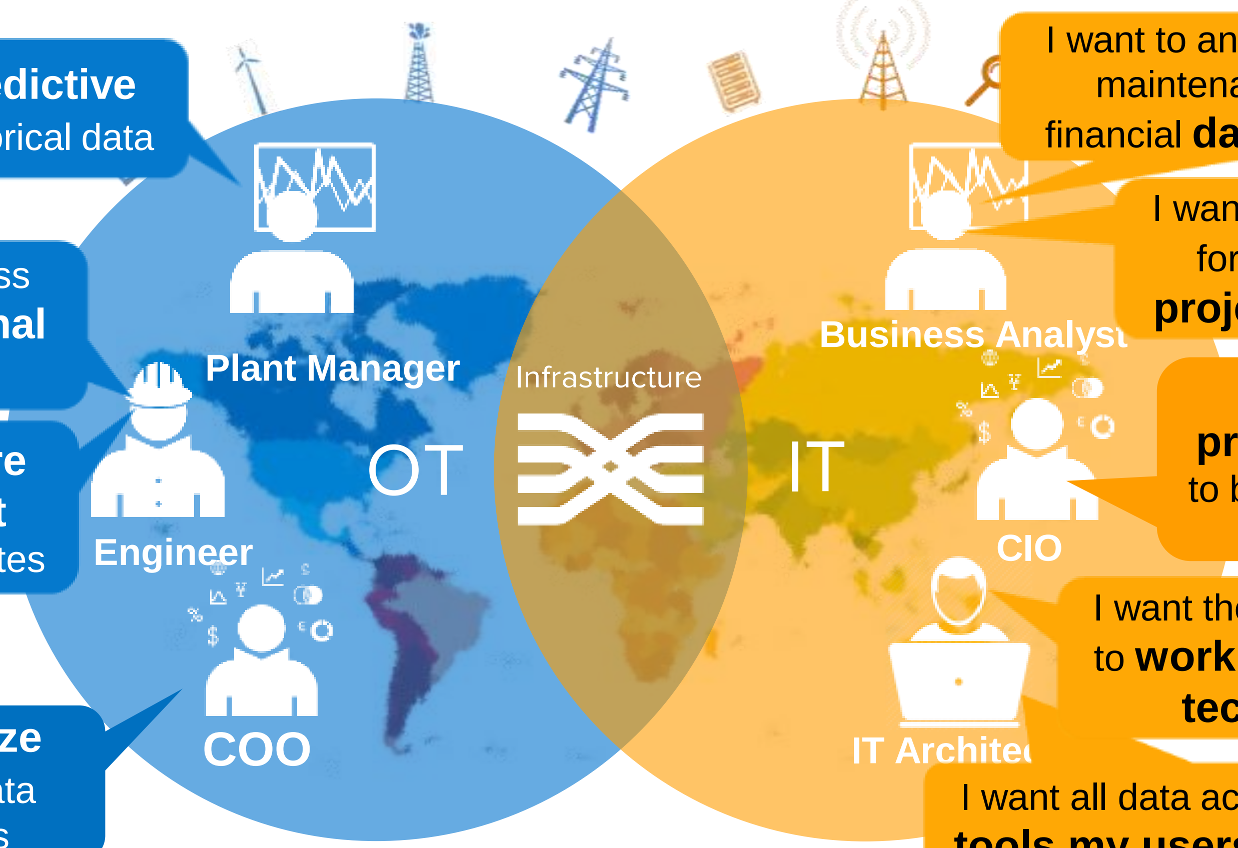
Data Warehousing

- **Centralizing data** from **different** business systems
- More effectively analyzing and reporting on business and building LOB applications



Structured PI data in a format readily consumable by the latest I.T. tools.

New Technology is bringing the IT and OT Worlds Together


Plant Manager
Engineer
COO
OT
Infrastructure
IT
Business Analyst
CIO
IT Architec

I want to build **predictive models** from historical data

I want to spend less time on **operational reports**

I want to **compare my equipment** against our other sites

I want to **minimize risks** through data driven decisions

I want to analyze production, maintenance logs, and financial **data all together**

I want operational data for the **Big Data project** we're starting

I want **trusted production data**, to be confident in our decisions

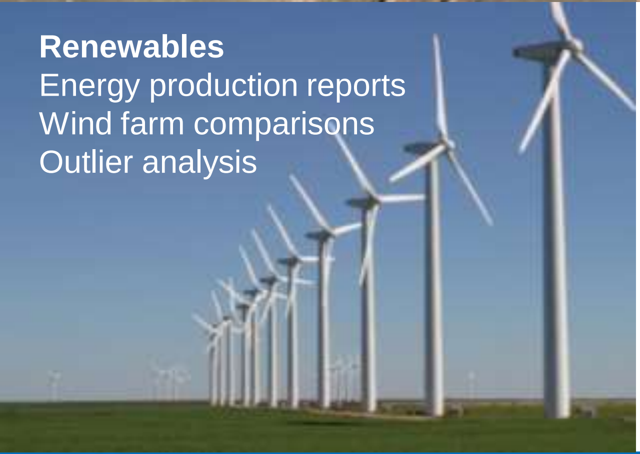
I want the operational data to **work with our other technologies**

I want all data accessible by the BI **tools my users already know**

Oil and Gas
Drilling and production comparisons
Information distribution

Mining
Route optimization
Energy reduction
Across 300 haul trucks

Life Sciences
Reactor comparisons
Process scale-ups
(1L, 3L, 10L, 1kL, 10kL)

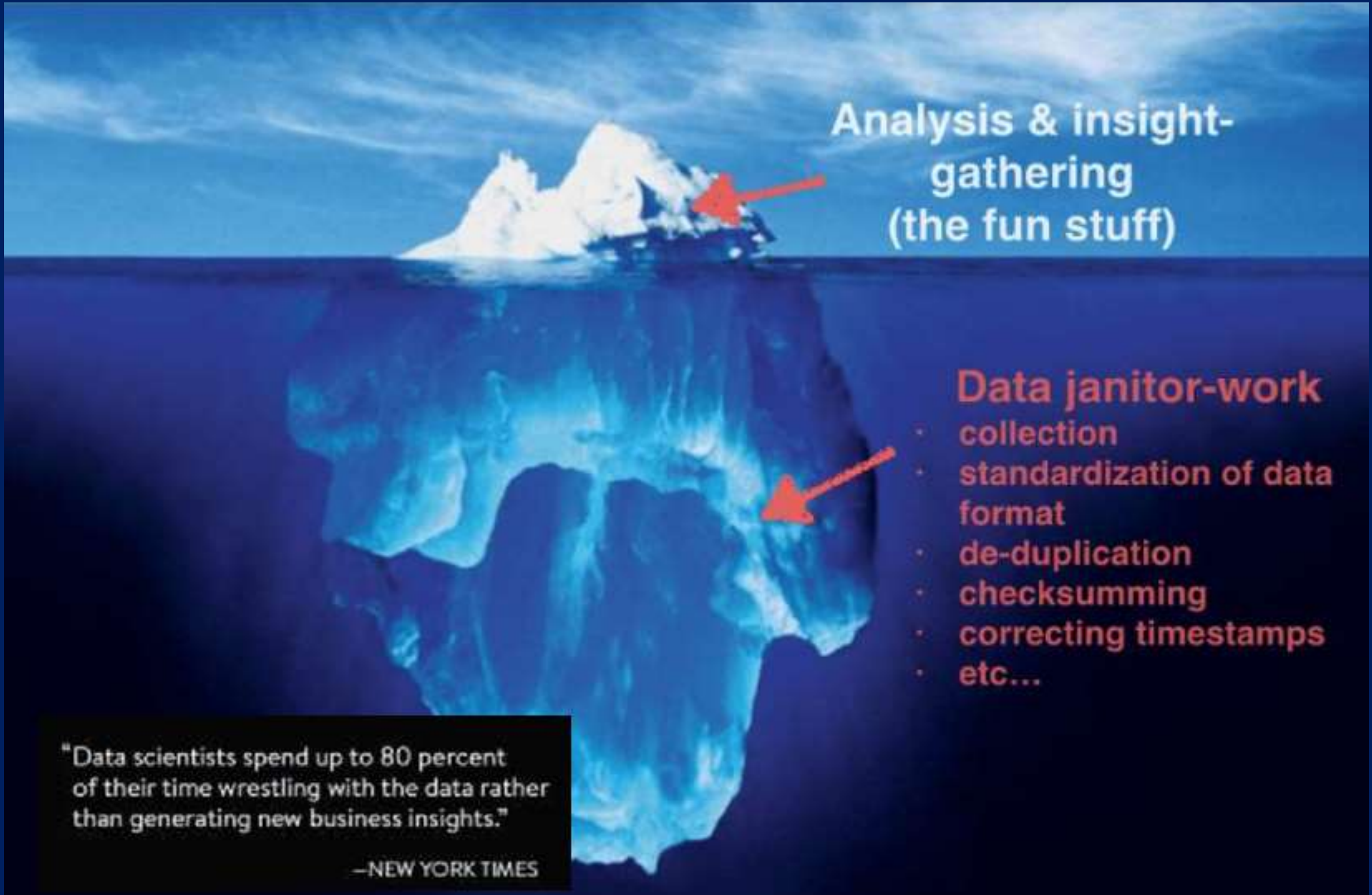


Renewables
Energy production reports
Wind farm comparisons
Outlier analysis

PI Integrator for Business Analytics is in use today!

- ✓ IT/ OT integration
- ✓ Business intelligence and reporting
- ✓ Data warehouse integration
- ✓ Supporting cross-platform projects

Food and Beverage
Utility usages
Process analytics



Analysis & insight-gathering (the fun stuff)

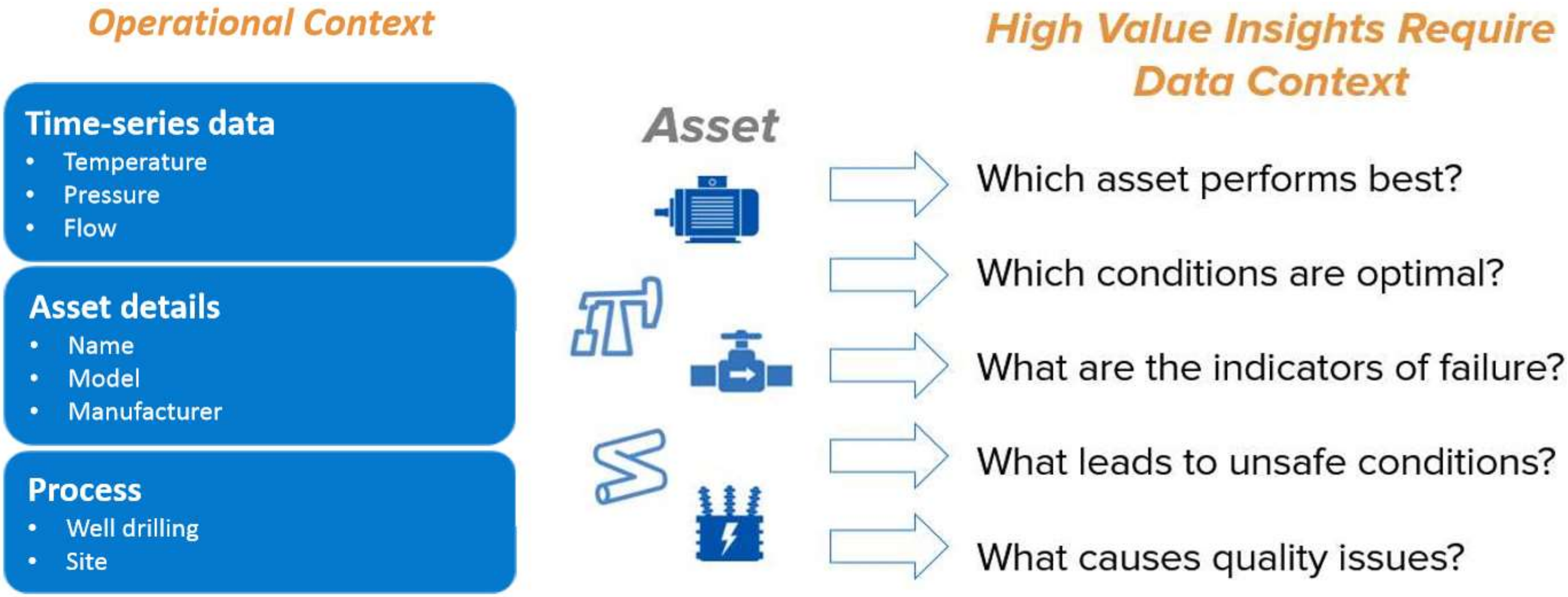
Data janitor-work

- collection
- standardization of data format
- de-duplication
- checksumming
- correcting timestamps
- etc…

> "Data scientists spend up to 80 percent of their time wrestling with the data rather than generating new business insights."
>
> —NEW YORK TIMES

Source: https://hbr.org/2014/04/the-sexiest-job-of-the-21st-century-is-tedious-and-that-needs-to-change/

PI AF can provide context and structure



Operational Context

Time-series data
- Temperature
- Pressure
- Flow

Asset details
- Name
- Model
- Manufacturer

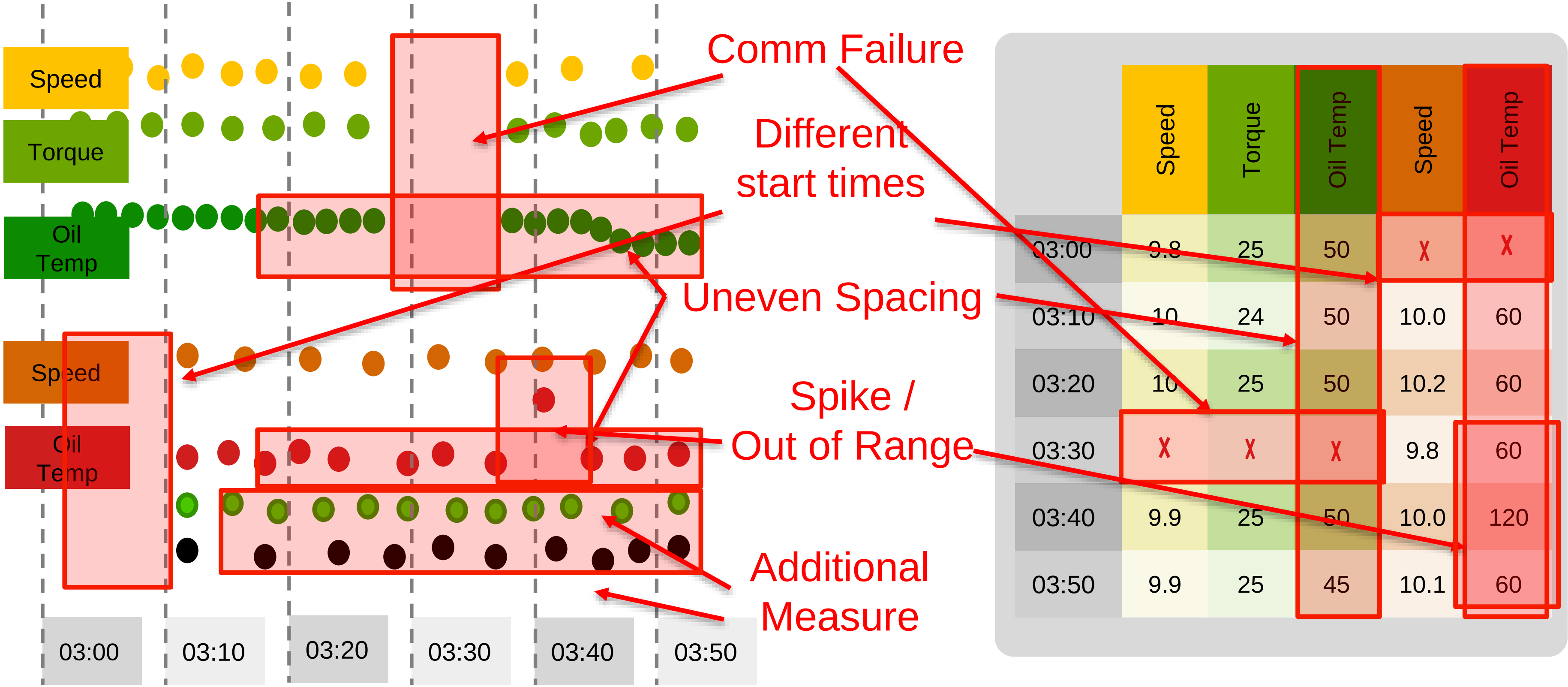
Process
- Well drilling
- Site

Asset

High Value Insights Require Data Context

- Which asset performs best?
- Which conditions are optimal?
- What are the indicators of failure?
- What leads to unsafe conditions?
- What causes quality issues?

Cleaning & Preparing Sensor Data: It’s Challenging


Speed
Torque
Oil Temp
Speed
Oil Temp
03:00
03:10
03:20
03:30
03:40
03:50

Comm Failure

Different start times

Uneven Spacing

Spike / Out of Range

Additional Measure

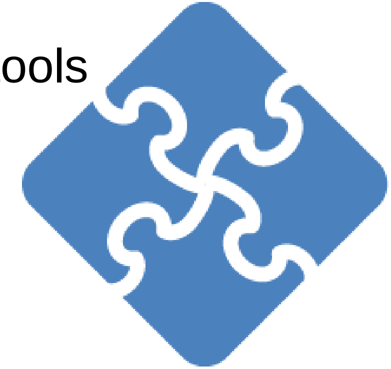
	Speed	Torque	Oil Temp	Speed	Oil Temp
03:00	9.8	25	50	✗	✗
03:10	10	24	50	10.0	60
03:20	10	25	50	10.2	60
03:30	✗	✗	✗	9.8	60
03:40	9.9	25	50	10.0	120
03:50	9.9	25	45	10.1	60

Three steps from PI System (AF) data to Big Data tools.. .

Select based on template (schema), intelligent filters, and time

Prepare clean, filter, & synchronize the data for tabular tools

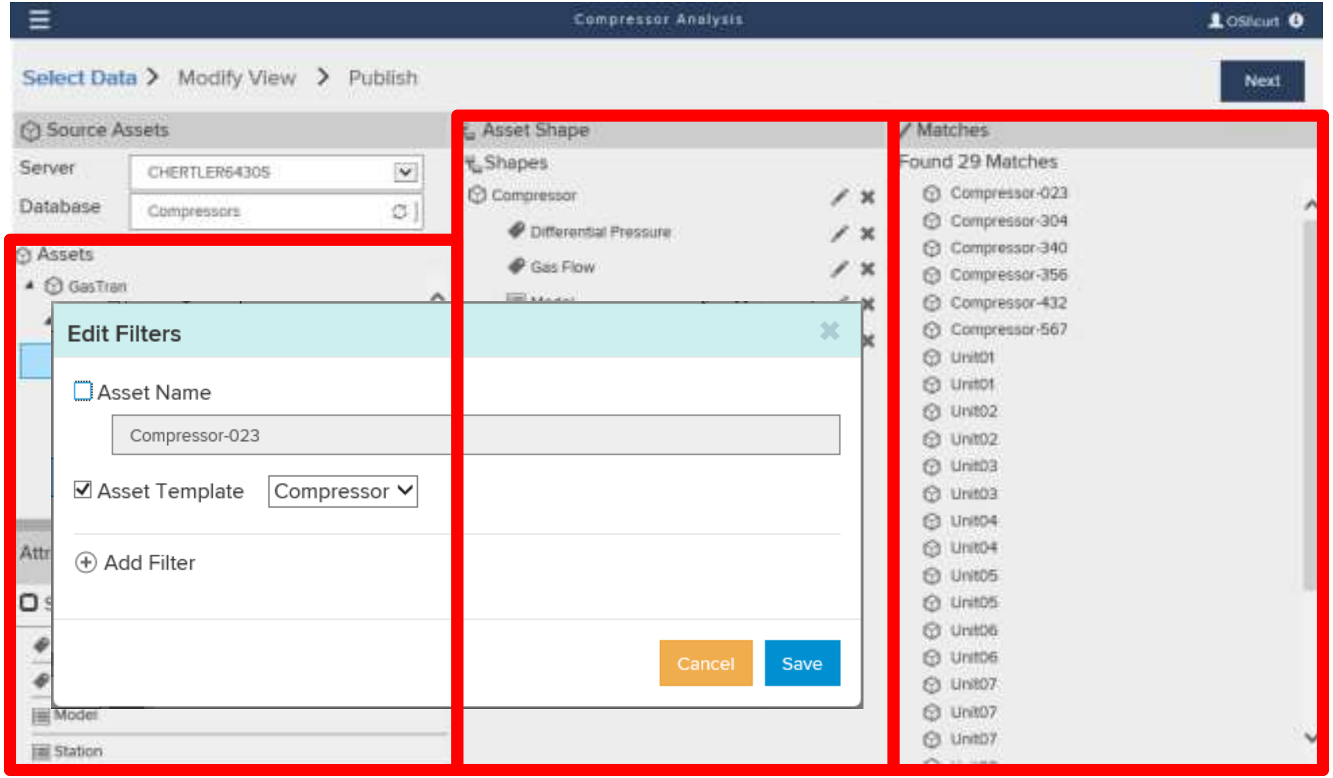
Publish and update the data on a schedule



A configured tool (no code) for PI System (AF) data and PI Event Frames
PI is still the source, rebuild your big data cache on your rules at any time.

PI Integrator for Business Analytics - *“Select Data”*

- Intuitive way to create tabular content in “PI Views”
- Requires AF Hierarchy
- Select AF Elements and Attributes
- or Select Event Frames
- Scale up leveraging name, hierarchy, or category



Compressor Analysis
OSIcuut
Select Data
Modify View
Publish
Next
Source Assets
Assets
GasTran
Attr
Model
Station
Asset Shape
Shapes
Compressor
Differential Pressure
Gas Flow
Model
Edit Filters
Add Filter
Cancel
Save
Matches

Server: CHERTLER64305

Database: Compressors

☐ Asset Name: Compressor-023

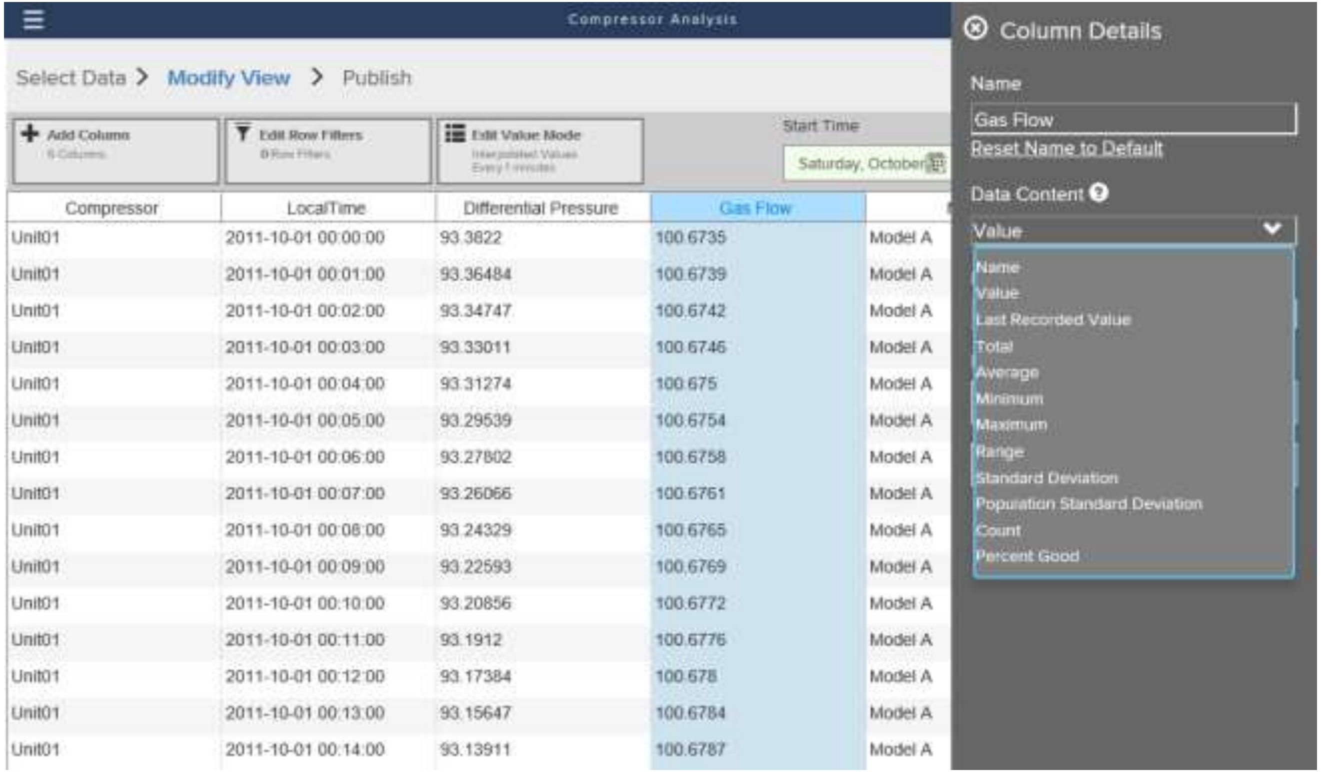
☑ Asset Template: Compressor

Found 29 Matches

- Compressor-023
- Compressor-304
- Compressor-340
- Compressor-356
- Compressor-432
- Compressor-567
- Unit01
- Unit01
- Unit02
- Unit02
- Unit03
- Unit03
- Unit04
- Unit04
- Unit05
- Unit05
- Unit06
- Unit06
- Unit07
- Unit07
- Unit07

PI Integrator for Business Analytics - *“Modify View”*

- Select time range and interval
- Add columns for proper aggregating & synchronization of PI System data
- Filter / clean data for null or bad values, missing data, out of range data or based on process conditions


Compressor Analysis
Select Data
Modify View
Publish
Add Column
Edit Row Filters
Edit Value Mode
Column Details
Reset Name to Default
Name
Value
Last Recorded Value
Total
Average
Minimum
Maximum
Range
Standard Deviation
Population Standard Deviation
Count
Percent Good

Start Time: Saturday, October

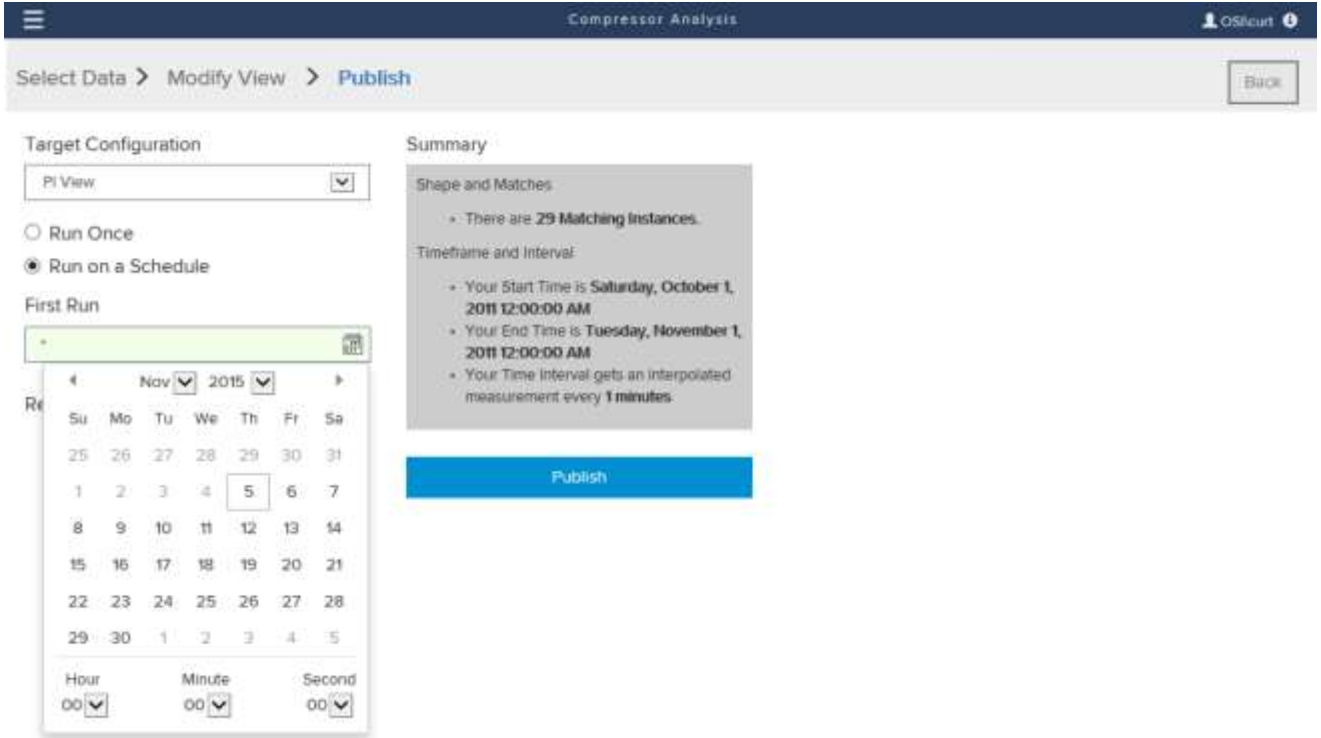
Name: Gas Flow

Data Content: Value

Compressor	LocalTime	Differential Pressure	Gas Flow	
Unit01	2011-10-01 00:00:00	93.3822	100.6735	Model A
Unit01	2011-10-01 00:01:00	93.36484	100.6739	Model A
Unit01	2011-10-01 00:02:00	93.34747	100.6742	Model A
Unit01	2011-10-01 00:03:00	93.33011	100.6746	Model A
Unit01	2011-10-01 00:04:00	93.31274	100.675	Model A
Unit01	2011-10-01 00:05:00	93.29539	100.6754	Model A
Unit01	2011-10-01 00:06:00	93.27802	100.6758	Model A
Unit01	2011-10-01 00:07:00	93.26066	100.6761	Model A
Unit01	2011-10-01 00:08:00	93.24329	100.6765	Model A
Unit01	2011-10-01 00:09:00	93.22593	100.6769	Model A
Unit01	2011-10-01 00:10:00	93.20856	100.6772	Model A
Unit01	2011-10-01 00:11:00	93.1912	100.6776	Model A
Unit01	2011-10-01 00:12:00	93.17384	100.678	Model A
Unit01	2011-10-01 00:13:00	93.15647	100.6784	Model A
Unit01	2011-10-01 00:14:00	93.13911	100.6787	Model A

PI Integrator for Business Analytics - *“Publish”*

- Select targeted endpoint from CSV ASCII files to the Azure cloud
- Publish once or schedule updates
- Manage your Integrator jobs


Compressor Analysis
OSIcurt
Select Data
Modify View
Publish
Back
Nov
2015
Su Mo Tu We Th Fr Sa
Publish

Target Configuration

PI View

○ Run Once

● Run on a Schedule

First Run

Hour 00

Minute 00

Second 00

Summary

Shape and Matches

- There are **29 Matching Instances.**

Timeframe and Interval

- Your Start Time is **Saturday, October 1, 2011 12:00:00 AM**
- Your End Time is **Tuesday, November 1, 2011 12:00:00 AM**
- Your Time Interval gets an interpolated measurement every **1 minutes**

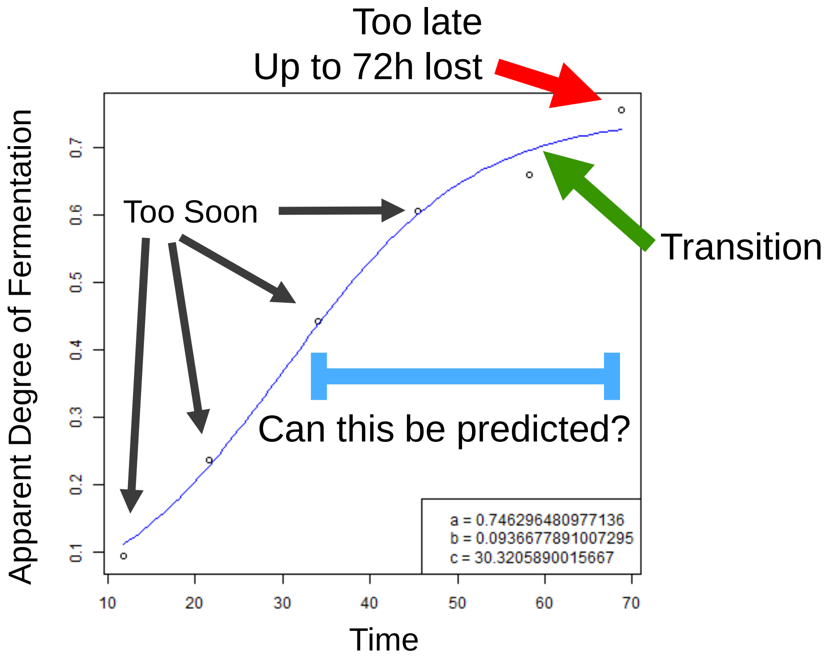
Deschutes Brewery

- Located in Bend, OR
- Founded in 1988
- Pub opened in Portland, OR in 2007
- 2 brewhouses
- 50+ vessels
- Bottling and kegging
- 7th largest craft brewer in the U.S.

DESCHUTES
BLACK BUTTE PORTER
DESCHUTES
MIRROR POND PALE ALE
DESCHUTES
OBSIDIAN STOUT
DESCHUTES
CHAINBREAKER WHITE IPA
DESCHUTES
DESCHUTES RIVER ALE
DESCHUTES
BIG RIG
DESCHUTES
DOWN 'N DIRTY IPA

Production Challenges

Filling ▶ Fermentation ▶ Free Rise ▶ ▪ ▪ ▪



Too late
Up to 72h lost

Too Soon

Transition

Can this be predicted?

Options?

Invest $750K (or more) in inline density meters

or

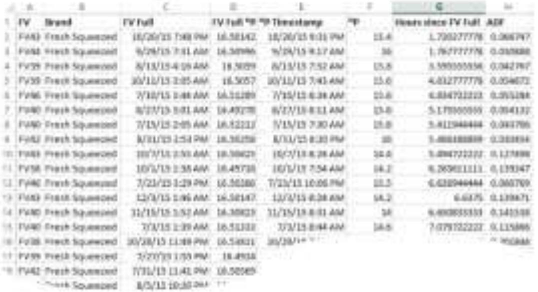
Continue predicting transitions with spreadsheets and manual readings every 8-10 hours

More Challenges

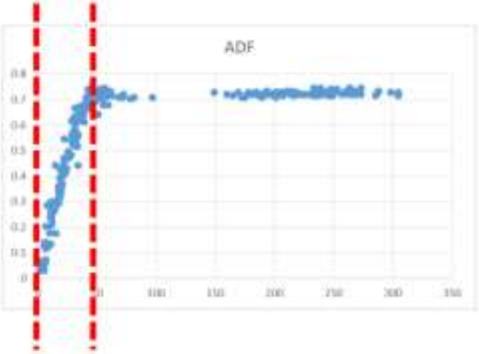
- Transition occurs between manual density measurements
- Beer brands vary in transition

Predictive Analytics in a Spreadsheet

Bring Raw Data In



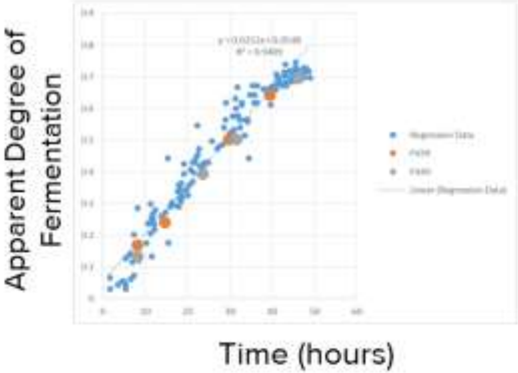
Clean it Up



New Challenges

- How can the **data preparation** be **automated**?
- How can the **predictions** be **operationalized**?
- How can the **predictions** become **more accurate** over time?

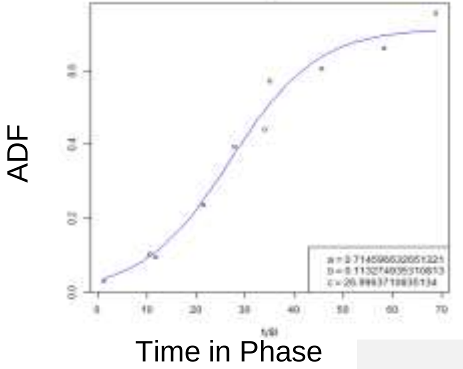
Fit to a Line



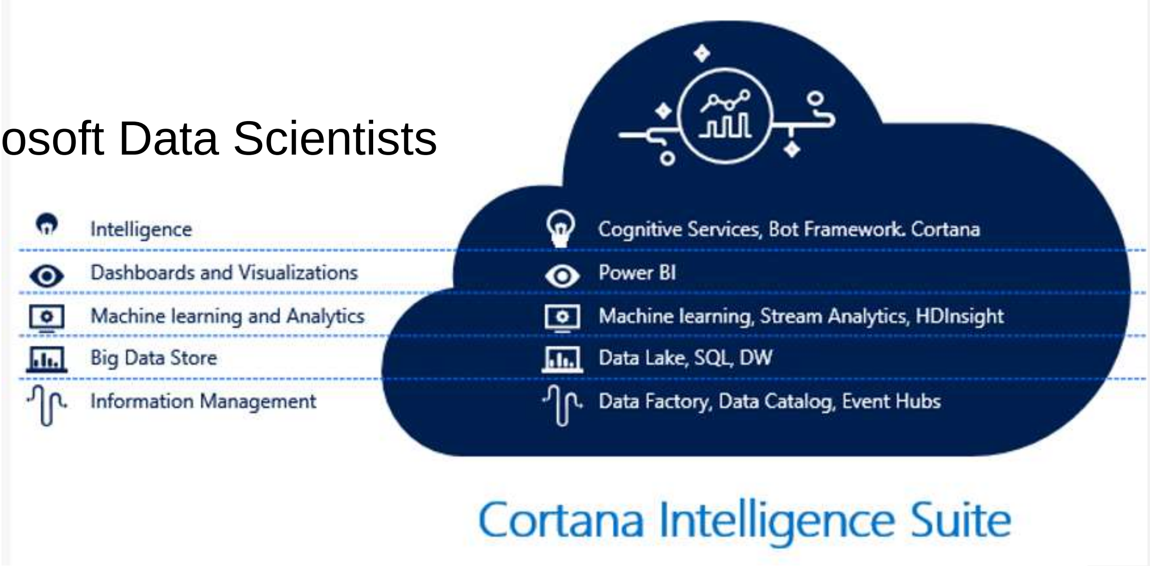
Hypothesis: Use early density readings to predict Free Rise Transition Time

influenced by:

- Brand of Beer
- Fermentation dynamics (Temperatures, pressures,..)
- Vessel’s dimensions & volume

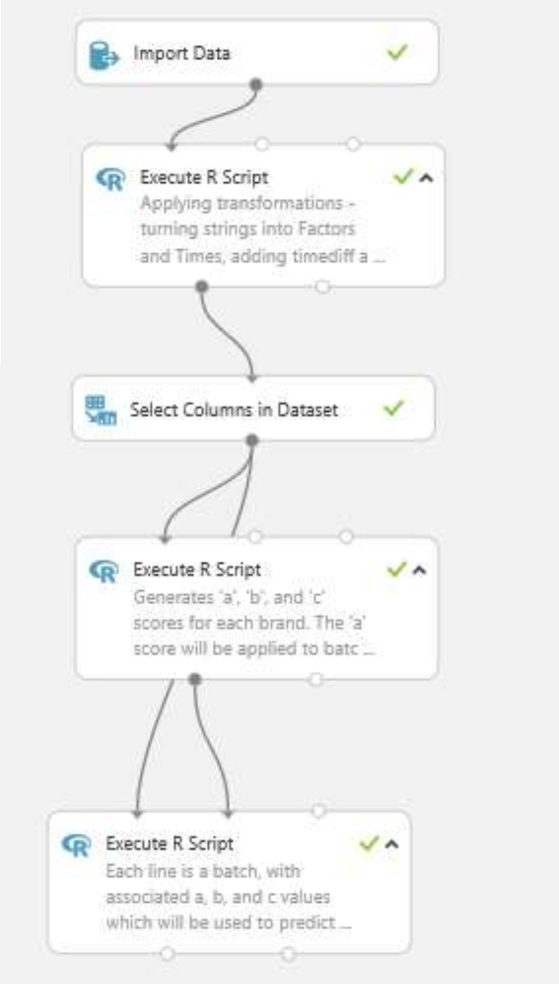


Microsoft Data Scientists



Intelligence	Cognitive Services, Bot Framework. Cortana
Dashboards and Visualizations	Power BI
Machine learning and Analytics	Machine learning, Stream Analytics, HDInsight
Big Data Store	Data Lake, SQL, DW
Information Management	Data Factory, Data Catalog, Event Hubs

Cortana Intelligence Suite



Import Data ✓

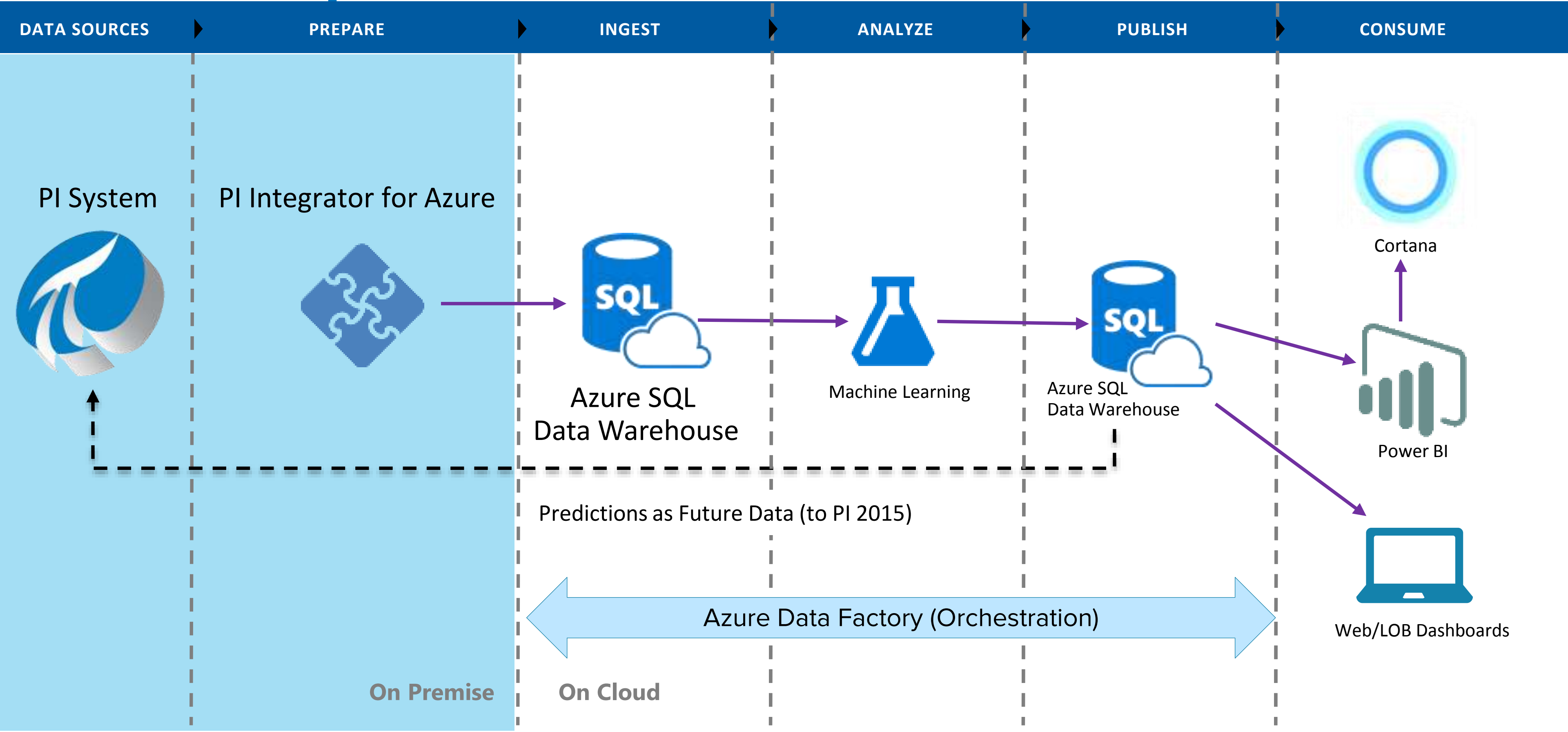
Execute R Script ✓
Applying transformations - turning strings into Factors and Times, adding timediff a ...

Select Columns in Dataset ✓

Execute R Script ✓
Generates 'a', 'b', and 'c' scores for each brand. The 'a' score will be applied to batc ...

Execute R Script ✓
Each line is a batch, with associated a, b, and c values which will be used to predict ...

How to Operationalize Predictions



DATA SOURCES | PREPARE | INGEST | ANALYZE | PUBLISH | CONSUME

PI System

PI Integrator for Azure

Azure SQL Data Warehouse

Machine Learning

Azure SQL Data Warehouse

Cortana

Power BI

Web/LOB Dashboards

Predictions as Future Data (to PI 2015)

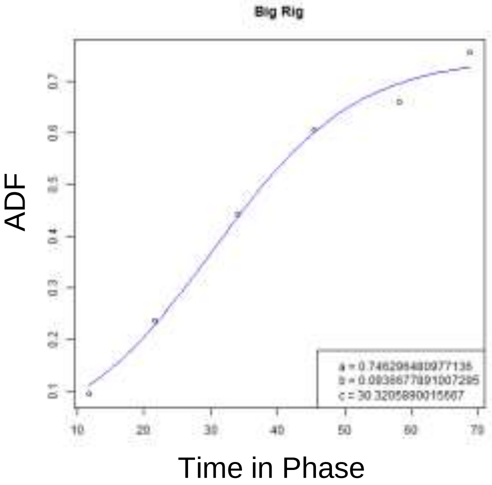
Azure Data Factory (Orchestration)

On Premise

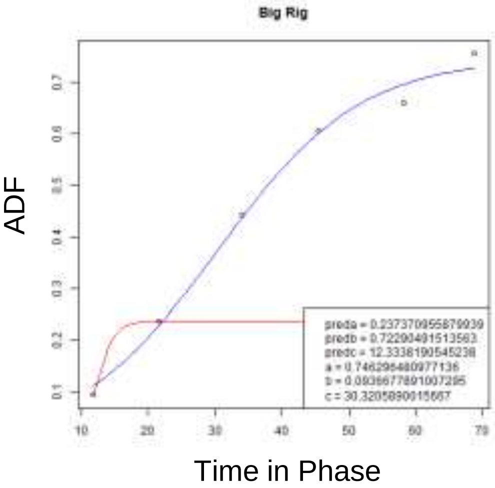
On Cloud

Azure ML Predicts Accurate Transition Time

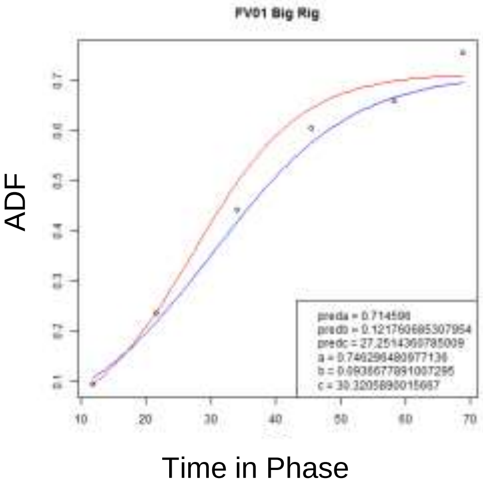
Benchmark: Measure accuracy against a standard (based on historical data)



Predict: use 2 early densities to estimate transition time



Refine: base predictions on brand for greater accuracy

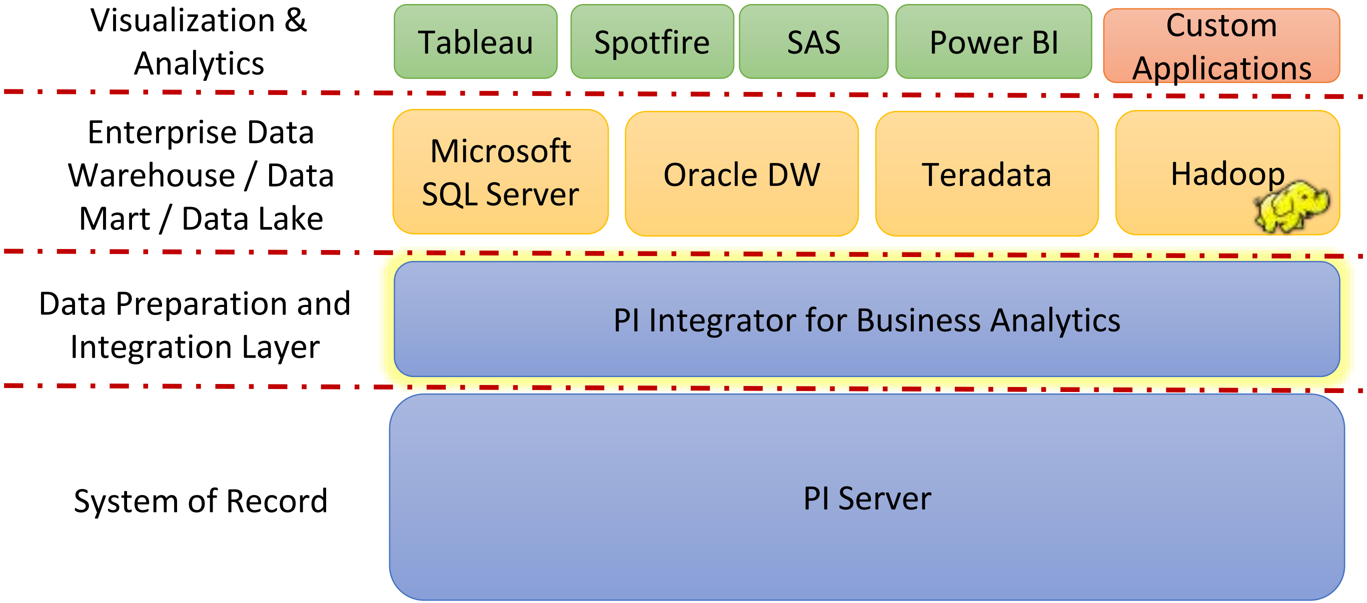


You can see Tim & Brian do this talk

- Go to OSIsoft.com
- Select Events / Past Events / Event Presentations
- Select EMEA 2016 Berlin
- Look for “Reducing Beer Production Time with Predictions”

Tim Alexander – Assistant Brewmaster – Data Wrangler
Brian Faivre - Brewmaster

Enterprise Data Warehouse Architecture



Visualization & Analytics

Tableau

Spotfire

SAS

Power BI

Custom Applications

Enterprise Data Warehouse / Data Mart / Data Lake

Microsoft SQL Server

Oracle DW

Teradata

Hadoop

Data Preparation and Integration Layer

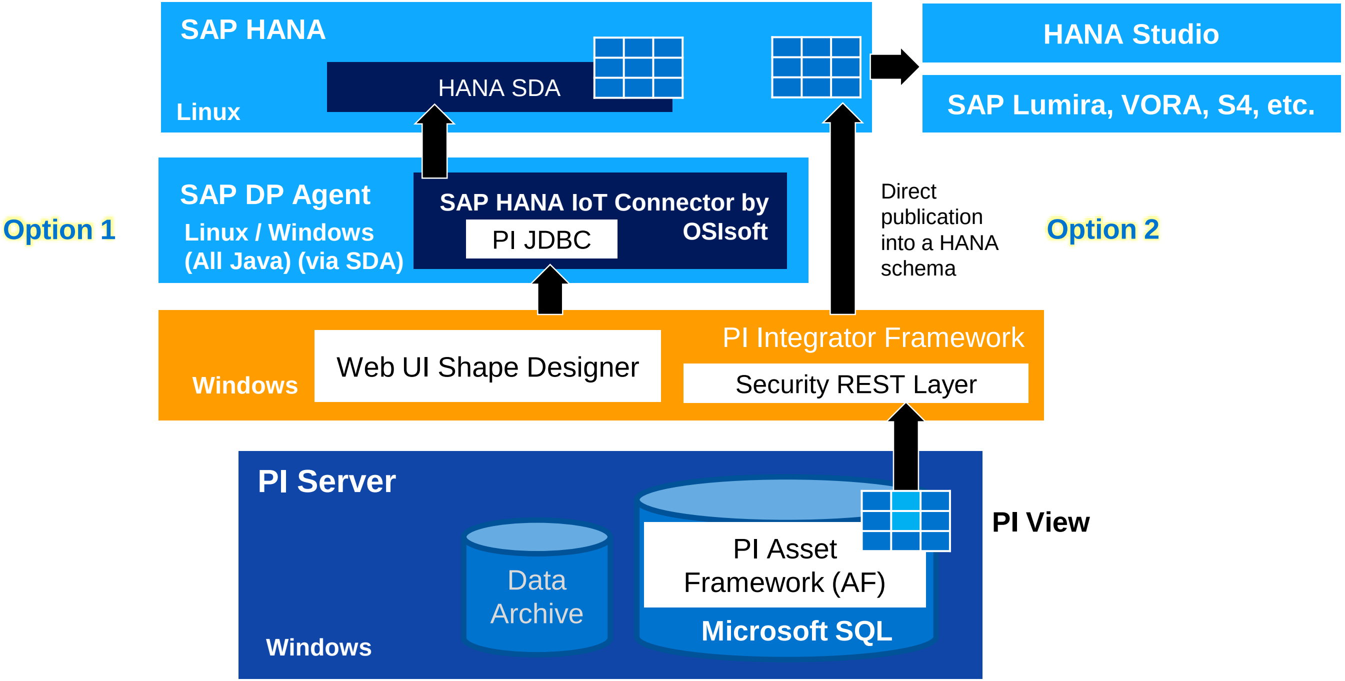
PI Integrator for Business Analytics

System of Record

PI Server

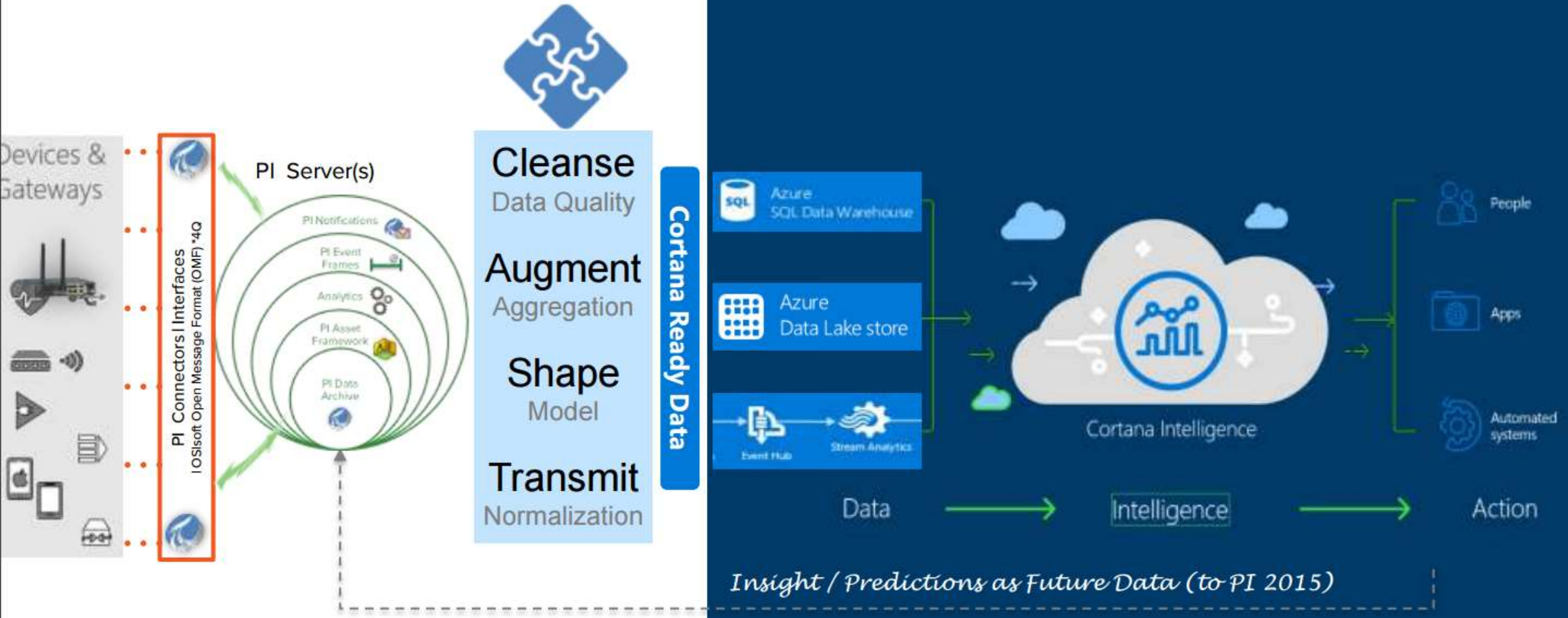
“I need to fit operational data into my existing company IT information architecture!”

PI Integrator for SAP HANA Architecture


SAP HANA
Linux
HANA SDA
HANA Studio
SAP Lumira, VORA, S4, etc.
Option 1
SAP DP Agent
Linux / Windows
(All Java) (via SDA)
SAP HANA IoT Connector by OSIsoft
PI JDBC
Direct publication into a HANA schema
Option 2
Windows
Web UI Shape Designer
PI Integrator Framework
Security REST Layer
PI Server
Data Archive
PI Asset Framework (AF)
Microsoft SQL
Windows
PI View

PI Integrator for Microsoft Azure

Cortana Intelligence



evices &
ateways
PI Connectors I Interfaces
I OSIsoft Open Message Format (OMF) *4Q
PI Server(s)
PI Notifications
PI Event Frames
Analytics
PI Asset Framework
PI Data Archive

Cleanse
Data Quality

Augment
Aggregation

Shape
Model

Transmit
Normalization

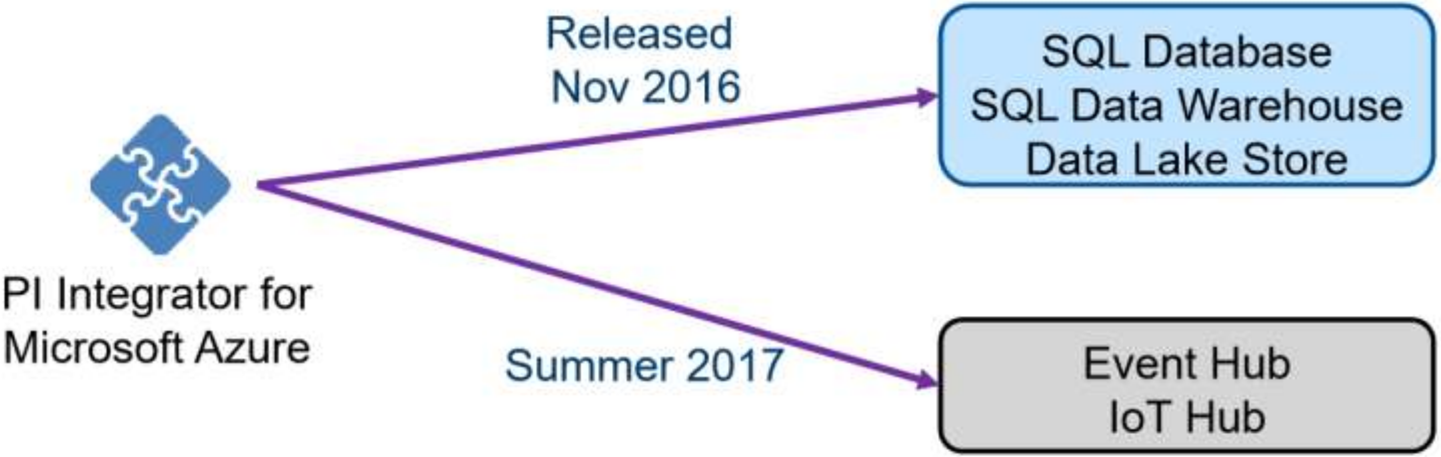
Cortana Ready Data

Azure SQL Data Warehouse
Azure Data Lake store
Event Hub
Stream Analytics
Cortana Intelligence
People
Apps
Automated systems

Data → Intelligence → Action

Insight / Predictions as Future Data (to PI 2015)

Bridging the PI System and Microsoft Azure



PI Integrator for Microsoft Azure

Released Nov 2016

SQL Database
SQL Data Warehouse
Data Lake Store

Summer 2017

Event Hub
IoT Hub

Getting Value…

Solving complex problems for a fleet	Dashboarding – visual reporting – real time & mobile	Integration to new I.T. projects and databases
Multivariate, other statistics & machine learning resources One time answers or Running models 	Drill drown, rollup Anywhere anytime 522 ($1.39M) $723.16 TIBCO Spotfire Power BI + a b l e a u	The right way to bring operational data to I.T.'s Big Data party… Microsoft SQL Server Reporting Services hadoop ORACLE BUSINESS INTELLIGENCE SAP

Contact Information

Martin Bryant

MBryant@osisoft.com

Field Service Engineer

OSIsoft, LLC

Thank You