

# **PI World 2020 Lab**

PI DataLink with Excel's Advanced Features

---

OSIsoft, LLC  
1600 Alvarado Street  
San Leandro, CA 94577 USA  
Tel: (01) 510-297-5800  
Web: <http://www.osisoft.com>

© 2020 by OSIsoft, LLC. All rights reserved.

OSIsoft, the OSIsoft logo and logotype, Analytics, PI ProcessBook, PI DataLink, ProcessPoint, Asset Framework (AF), IT Monitor, MCN Health Monitor, PI System, PI ActiveView, PI ACE, PI AlarmView, PI BatchView, PI Vision, PI Data Services, Event Frames, PI Manual Logger, PI ProfileView, PI WebParts, ProTRAQ, RLINK, RtAnalytics, RtBaseline, RtPortal, RtPM, RtReports and RtWebParts are all trademarks of OSIsoft, LLC. All other trademarks or trade names used herein are the property of their respective owners.

#### U.S. GOVERNMENT RIGHTS

Use, duplication or disclosure by the U.S. Government is subject to restrictions set forth in the OSIsoft, LLC license agreement and as provided in DFARS 227.7202, DFARS 252.227-7013, FAR 12.212, FAR 52.227, as applicable. OSIsoft, LLC.

Published: May 27, 2020

# Table of Contents

## Contents

Table of Contents .....	3
1. INTRODUCTION .....	6
Workspace & File Locations .....	6
The Fictitious Plant Used in This Lab .....	6
How to use this Workbook .....	9
1.1 Useful Resources .....	9
1.2	
2.3 EXERCISE: Create a Report using PI DataLink .....	10
1.4 Objectives .....	10
Background .....	10
2.1	
2.2 Calculation Functions in PI DataLink .....	11
2.3 Solutions .....	13
2.4	
3. EXERCISE: Analyze Downtime Events using PivotCharts and PivotTables .....	15
3.1 Objectives .....	15
3.2 Background .....	15
3.3 Directed Activity: Retrieve Event Frames in Excel using PI DataLink .....	15
3.4 Directed Activity: Create PivotTables and PivotCharts .....	17
3.4.1 Comparing Downtime Events Based on Reason Code .....	17
3.5	
3.4.2 Comparing Downtime Events Based on Tank .....	21
3.6	
3.7 Analyze Downtime Events Using PivotTables and PivotCharts .....	23
Bonus: Making Reports More User-Friendly .....	24
4.1 Solutions .....	25
4.2	
4. EXERCISE: Improve Performance of Large Workbooks .....	26
4.3	
4.4 Objectives .....	26
4.5 Background .....	26
4.6 Directed Activity: Measure AF SDK calls made in the background .....	27
5.1 Directed Activity: Gather information about the workbook .....	31
Improve performance by reducing the number of background calls .....	33
Solutions .....	35
5. EXERCISE: Excel’s New Dynamic Array Functions .....	36
Objectives .....	36

---

	Background .....	36
	Learn about Excel’s new dynamic array functions.....	37
	Use dynamic array functions to complement PI DataLink functions.....	38
	Solutions.....	39
6.2	EXERCISE: Introduction to PI DataLink Documentation.....	41
5.3	Objective .....	41
5.4	Background .....	41
5.5	Access PI DataLink’s user guide in context .....	42
6.1	Directed Activity: Access the Customer Portal .....	43
6.2	Save the Date! .....	45
6.3		
6.4		



# 1. INTRODUCTION

This lab is designed for users who are familiar with Microsoft Excel and have already used PI DataLink.

Learn how to combine the power of PI Datalink with Excel's advanced features such as sparklines, Pivot Tables, Pivot Charts, and dynamic arrays. Using these tools, we will show you best practices for creating reports and examples to prepare PI System data and event frames for root cause analysis.

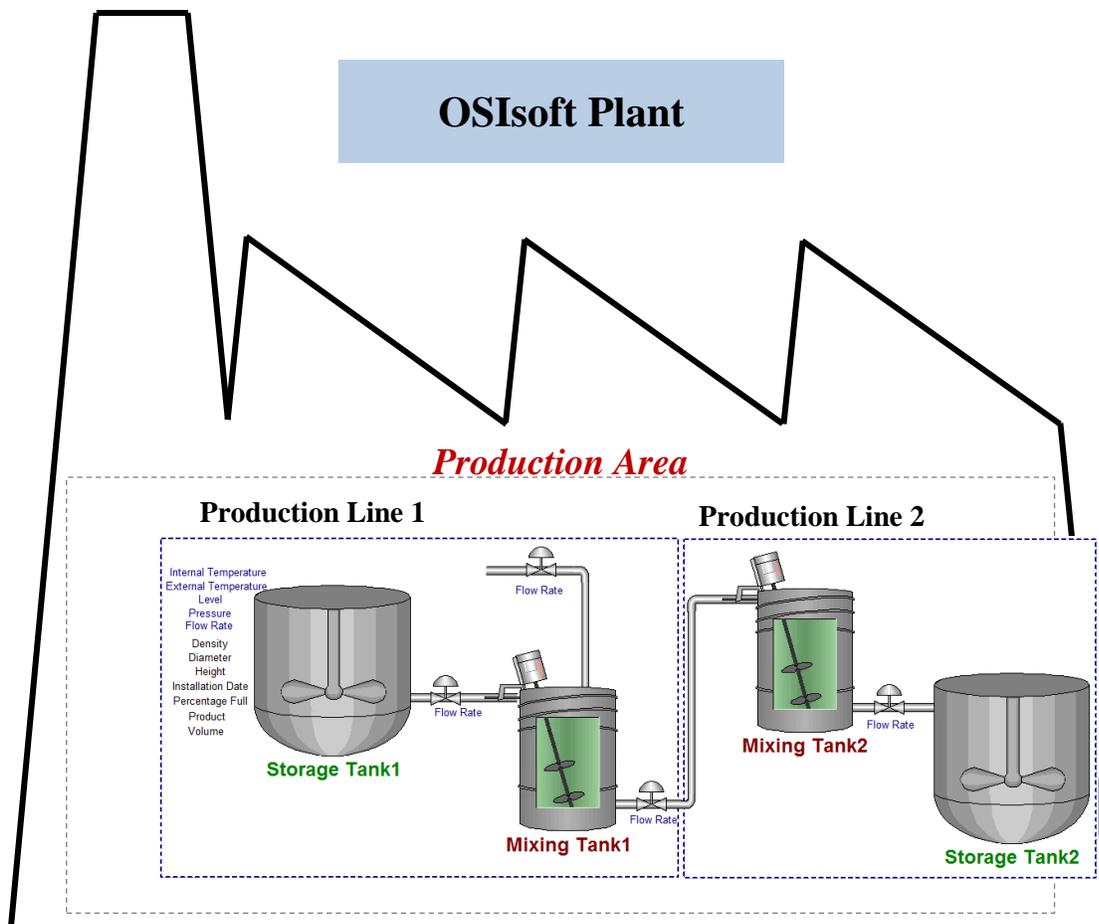
## Workspace & File Locations

- 1.1 Each lab attendee has their own client machine but will be using the same PI Data Archive and AF Server, PISRV1. All machines are VMs hosted in an Azure Virtual Learning Environment (VLE).

A soft copy of this workbook is available on the desktop of your client machine. As it contains several hyperlinks, you are encouraged to copy the workbook from your lab client machine to your laptop. (The clipboard is enabled through Remote Desktop connections)

## 1.2 The Fictitious Plant Used in This Lab

In this lab, we will have a fictitious plant named OSIsoft Plant. This simple plant has two production lines, where each has a combination of one mixing tank and one storage tank. This plant could be schematically shown as:



As shown here, each tank has different process variables such as Internal and External Temperatures, Flow Rate, Pressure and Level whose values are continuously collected from devices on the Plant. In the early days of PI System, these process variables were the only data items whose historical data could be stored in Data Archive.

There are some other data associated with each of these tanks such as the manufacturer, model and the installation date which are stored in the maintenance sheets available on tables in SQL Server. Moreover, all the information related to the material flowing in these tanks is kept in tables on the Plant's SQL Servers.

Even though these tables are available on the SQL server, their information could not easily be integrated with the historical data stored in Data Archive. Hence, using AF and hierarchy becomes critical in bringing all the important data and information in one place: the PI System.

At the OSisoft Plant, predictions on the level of each mixing tank is critical in running a smooth production. This data, Level\_Forecast, is stored in a "Future" point on the Data Archive and could be viewed on PI System displays or be compared to the actual value of level in any PI Application.

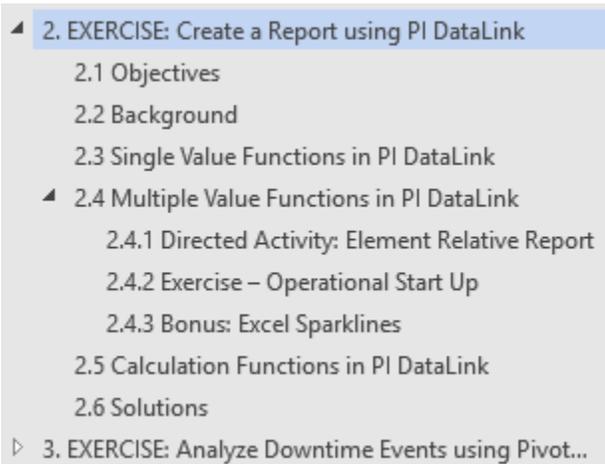
A collection of PI Points are built on Data Archive for storing the values of process variables. There is also a hierarchy built in AF for this Plant, bringing all the important information and data, including the process variable time series data, to one place.

Name	Value
Asset Location	Production Line1
Asset Name	Mixing Tank1
Density	4321 g/L
Diameter	5 ft
External Temperature	223.497 °F
Average	229.7221 °F
Flow Rate	152.0182 US gal/min
Height	10 ft
Installation Date	3/2/2016 3:00:00 PM
Manufacturer	ACME
Serial Number	8T498-C54
Internal Temperature	104.3854 °F
Average	135.7157 °F
Level	4.597035 ft
Maximum	10 ft
Minimum	0 ft
Target	5.099405 ft
Level_Forecast	4.971141 ft
Percentage Full	45.9703493118286 %
Pressure	88.58934 psi
Product	BCS1717
RandomSeed	0.047563022490387324
Status	Emptying
Tank Status	3
Tank Volume	1468.797 US gal

## How to use this Workbook

This workbook is structured around exercises. Each exercise will begin with objectives and background sections, then proceed through a variety of individual exercises, directed activities, and questions. Solutions for exercises and questions are available at the end of the overall exercise; directed activities guide you through the solution already. For example, this is the structure for the first exercise:

1.3



- ▲ 2. EXERCISE: Create a Report using PI DataLink
  - 2.1 Objectives
  - 2.2 Background
  - 2.3 Single Value Functions in PI DataLink
  - ▲ 2.4 Multiple Value Functions in PI DataLink
    - 2.4.1 Directed Activity: Element Relative Report
    - 2.4.2 Exercise – Operational Start Up
    - 2.4.3 Bonus: Excel Sparklines
  - 2.5 Calculation Functions in PI DataLink
  - 2.6 Solutions
- ▷ 3. EXERCISE: Analyze Downtime Events using Pivot...

To more easily browse this workbook, go to **View > Show** in the ribbon and check **Navigation Pane**.

1.4

## Useful Resources

- PI DataLink User Guide, available
  - via the F1 key or question mark on a PI DataLink task pane
  - on [LiveLibrary](#)
  - as a PDF download from the [OSIsoft Customer Portal](#)
- [PI DataLink Playbook](#), which contains
  - Version history
  - Common data gather steps for troubleshooting issues
  - Architecture and data flow overview
  - Links to articles addressing common issues, sorted by issue type (installation, runtime, admin tasks, etc.)

## 2. EXERCISE: Create a Report using PI DataLink

### Objectives

- Re-familiarize yourself with PI Datalink
- Create a report using the **Calculated Data** function.

2.1

### Background

2.2

PI DataLink is an add-in for Microsoft Excel that enables you to retrieve information from your PI System directly into a worksheet. Combined with the computational, graphic, and formatting capabilities of Microsoft Excel, PI DataLink offers powerful tools for gathering, monitoring, analyzing, and reporting PI System data.

PI DataLink sorts its functions into groups:



- Single-value functions retrieve the value of a data item at a specific time. They return exactly one value per data item.
- Multiple-value functions associate a PI point or PI AF attribute with a time period over which there can be one or many corresponding values.
- Calculation functions compute values from PI point values, PI AF attributes, or performance equation evaluations during a specific time period.
- Events functions return events that meet specified criteria in a PI AF database. Use the Explore Events function to view and explore events in a hierarchical format, or the Compare Events function for a flat format. These functions return one event per row.
- The Search functions return data items that meet specified criteria.
- The Properties function returns the property value of a specified data item.

## Directed Activity: Calculation Functions in PI Datalink

### Problem Description

2.3 As a production manager, you want to create a report showing last week's production statistics. You want to display the following for the production from each day of the past week:

- Total
- Average
- Maximum

You also want to do the same calculations for the entire week.

### Approach

Step 1 : On your PI Server, the production is the sum of the productions from the two production lines and is stored on your AF Server as an attribute named **Production** under the element of **Production Area**, as shown below:

The screenshot displays the PI Datalink configuration interface. On the left, the 'Elements' tree shows a hierarchy: 'Production Area' (circled in red), which contains 'Production Line1' and 'Production Line2'. 'Production Line1' includes 'Mixing Tank1' and 'Storage Tank1'. 'Production Line2' includes 'Mixing Tank2' and 'Storage Tank2'. The main window shows the 'Production Area' configuration. The 'Attributes' tab is active, displaying a table with the following data:

Name	Value
Asset Name	Production Area
Production	5797.61 US gal/min

The 'Production' attribute is highlighted with a red dashed box. On the right, the 'Settings...' section contains a formula for the 'Production' attribute:

```
A=.\Production Line1\Mixing Tank1|Flow Rate;B=.\Production Line2\Mixing Tank2|Flow Rate;[A+B]
```

**Note:** Use the PI Point **CDT158** if you do not have access to the AF Database.

Step 2 : Spend a few minutes and fill out the following table:

<b>Root Path</b>	
<b>Data item</b>	
<b>Start time:</b>	
<b>End time:</b>	
<b>Time interval:</b>	
<b>UOM</b>	

Step 3 : You will use the template provided in sheet *Production Summaries* of the file *PI\_DataLink-Exercises.xlsx*. Use the values of your table in the provided template.

**Important Notes:**

When working with the weekly total, do not use the time interval. Only use it for the daily total (hint).

Be sure to show the Start Time for the daily Calculations.

The percent good field is always located to the right of an aggregate calculation, so use

Show percent good with the **Maximum**.

## Solutions

### Calculation Functions (Production Summaries)

2.4

1. The table could be filled out as:

<b>Root Path</b>	\\PISRV1\OSIsoft Plant\Production Area
<b>Data item</b>	Production
<b>Start time:</b>	04-Jan-2020 00:00:00
<b>End time:</b>	11-Jan-2020 00:00:00
<b>Time interval:</b>	1d
<b>UOM</b>	US gal/min

2. Open the file *PI\_DataLink-Exercises.xlsx*, then work on sheet *Production Summaries*.
3. Select cell B11.
4. In the PI DataLink ribbon, select the Properties function.
5. Select cell B5 as the root path, B6 as the data item, and UOM as the property.
6. Click OK.
7. Select cell A21.
8. In the Ribbon, select **Calculated Data**.
9. Make sure the **Data Item** radio button is selected near the top of the task pane.
10. In the task pane, click in the **Root path (optional)** field.
11. Click on cell B5.
12. In the task pane, click in the **Data item(s)** field.
13. Click on cell B6.
14. In the task pane, click in the **Start time** field.
15. Click on cell B7.
16. In the task pane, click in the **End time** field.
17. Click on cell B8.
18. In the task pane, click in the **Time Interval (optional)** field.
19. Click on cell B10.
20. In the **Calculation mode** drop-down, select **total**.
21. Since the UOM of the production is **per min**, select minute from the conversion factor drop down.
22. Click OK to have *1440* entered in the **Conversion Factor** field for the total calculation mode only. For all other calculation modes, leave this field at *1*.
23. For column A only, check the **show end time** check box to fill the **Time stamp** column.
24. Click on the **OK** button.
25. Repeat steps 5-22 for columns C through D using the appropriate selection for **Calculation Mode** (e.g., select *Average* in column C). When you get to column D, make sure the **show percent good** check box is checked.
26. Select cell B32.
27. In the Ribbon, select **Calculated Data**.
28. In the task pane, make sure the **Data Item** radio button is selected.

29. In the task pane, click in the **Root path (optional)** field.
30. Click on cell B5.
31. In the task pane click in the **Data item(s)** field.
32. Click on cell B6.
33. In the task pane click in the **Start time** field.
34. Click on cell B7.
35. In the task pane click in the **End time** field.
36. Click on cell B8.
37. In the **Calculation mode** drop-down, select **total**.
38. Since the UOM of the production is **per min**, select minute from the conversion factor drop down.
39. Click OK to have *1440* entered in the **Conversion Factor** field for the total calculation mode only. For all other calculation modes, leave this field at *1*.
40. Click on the **OK** button.
41. Repeat steps 24-38 for columns C through D using the appropriate selection for **Calculation Mode** (e.g., select *Average* in column C). When you get to column D, check the **show percent good** check box to fill the **Percent Good** column.

## 3. Directed Activity: Analyze Downtime Events using PivotCharts and PivotTables

### Objectives

- Retrieve Event Frames in PI DataLink
  - Create a PivotTable and PivotChart from Event Frame Data
- 3.1
- Analyze Event Frames retrieved from PI DataLink using PivotCharts and PivotTables

### Background

- 3.2
- PivotTables and PivotCharts are powerful tools to calculate, summarize, and analyze data that let you see comparisons, patterns, and trends in your PI Event Frame data.

### Directed Activity: Retrieve Event Frames in Excel using PI DataLink

3.3

#### Objectives

- Retrieve Event Frames in PI DataLink

#### Problem Description

The operations manager needs a report that shows which downtime reason is most prevalent and a comparison showing which tanks are the most problematic. He also would like to see information about the total production loss from the tanks.

In this exercise, you will import event data into Excel. Later, you will create PivotTables and PivotCharts to analyze the data.

#### Approach

Step 1 : Open the *Downtime Raw Data* sheet and fill in the values for cells B1:B3 as:

- a. **Database:** '\\PISRV1\OS\soft Plant  
\*\*Note: use the single quote ' to format the cell as text and not a function
- b. **Search Start:** t-7d
- c. **Search End:** \*

Step 2 : Select cell A7 as your Output Cell

Step 3 : Select **Compare** from the Events group

Step 4 : Make the following selections:

- a. **Database:**  
\\PISRV1\OS\soft Plant
- b. **Event name:** B1
- c. **Search start:** B2
- d. **Search end:** B3
- e. **Event template:**  
Downtime
- f. **Element name:** \*

Compare Events

Database	Event name
Downtime Raw Data!\$B\$1	*
Search start	Event template
Downtime Raw Data!\$B\$2	Downtime
Search end	Element name
Downtime Raw Data!\$B\$3	*

Step 5 : From Columns to display select Event name, Start time, End time, Duration, Event Template, Primary element, Event Duration (min), Maximum External Temperature, Maximum Internal Temperature, Reason Code, Lost Production (gal), and Temperature Difference.

<input checked="" type="checkbox"/> .IEvent template	↑
<input checked="" type="checkbox"/> .IPrimary element	↓
<input checked="" type="checkbox"/> .ITemperature Difference	✖
<input checked="" type="checkbox"/> .IEvent Duration (minutes)	
<input checked="" type="checkbox"/> .IMaximum External Temperature	
<input checked="" type="checkbox"/> .IMaximum Internal Temperature	
<input checked="" type="checkbox"/> .ILost Production (gal)	
<input checked="" type="checkbox"/> .IReason Code	+

Step 6 : Click OK.

## Directed Activity: Create PivotTables and PivotCharts

### Objectives

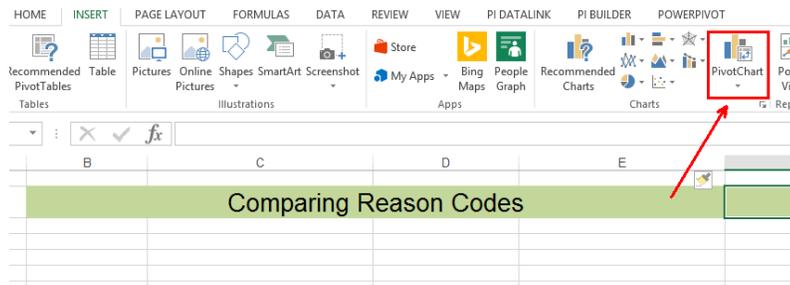
- Create a PivotTable and PivotChart from Event Frame Data

### 3.4 3.4.1 Comparing Downtime Events Based on Reason Code

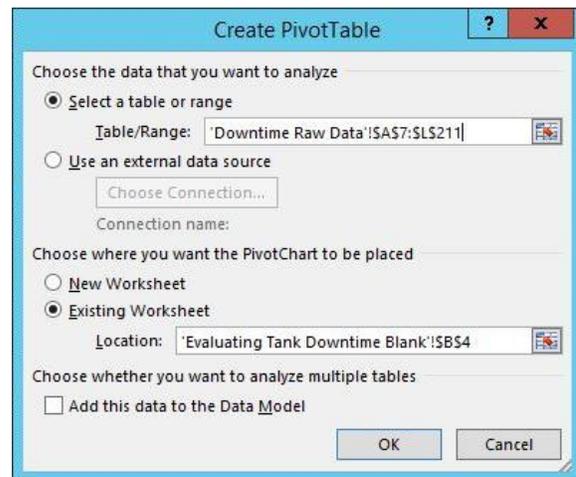
We would like to determine which causes of events are the most frequent and costly. To do this, we will create a PivotChart report from Event Frames generated in AF.

After completing the Event Frame search in the previous section, use the following instructions to create the Pivot Chart and Pivot Table.

- Step 1 : Go to the *Evaluating Tank Downtime Blank* sheet, select the *Insert* ribbon and select the *PivotChart* option. Select the columns and rows of interest from the Downtime Raw Data tab. This will create a PivotTable and a PivotChart.

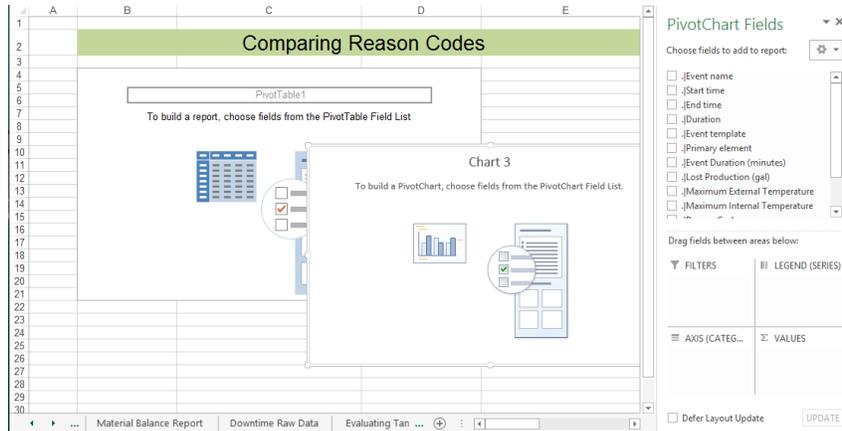


- Step 2 : As input for the PivotTable select the cell range in the *Downtime Raw Data* sheet where the CompareEvents function has returned the data (including the header line). Then choose to place the PivotTable and PivotChart in the *Evaluating Tank Downtime Blank* sheet.



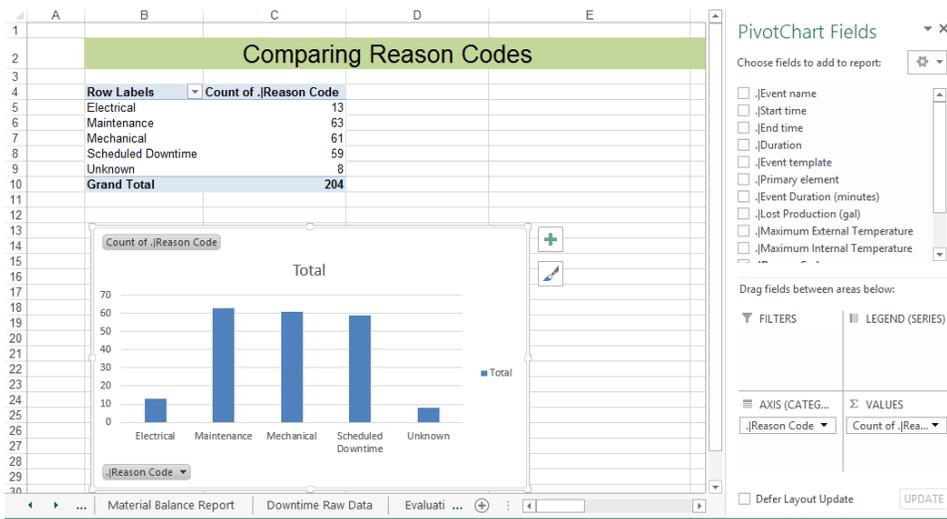
Tip: if you want to correct the source area later in time, select all cells of your PivotTable (or choose the Analyze ribbon), then from the Analyze Ribbon, select **Change Data Source**.

Step 3 : The PivotChart field list should now be shown in your Excel worksheet and a range of the worksheet should be designated where the pivot table will be located, as shown below.



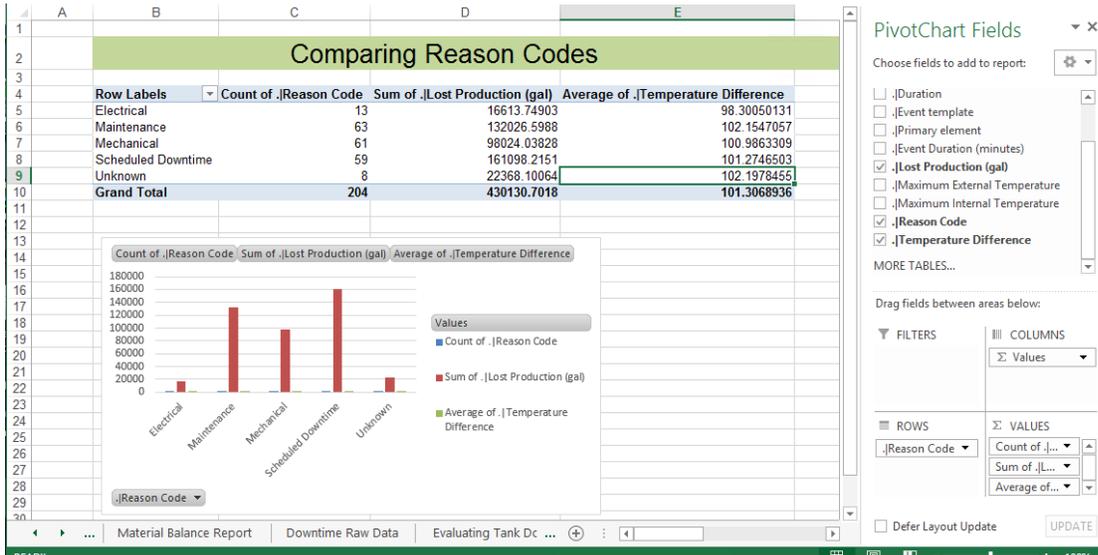
Step 4 : Select the PivotTable, and review the PivotTable Field list. These fields come from the column names of the *Downtime Raw Data* sheet.

Step 5 : To perform a downtime analysis for our Event Frames based on the corresponding reason code, select the **.|Reason Code** line and drag into the *Values* area. The applied Aggregation for the reason codes is **COUNT**, because these are non-numeric values. Select the **.|Reason Code** line again and drag into the *Rows* area:



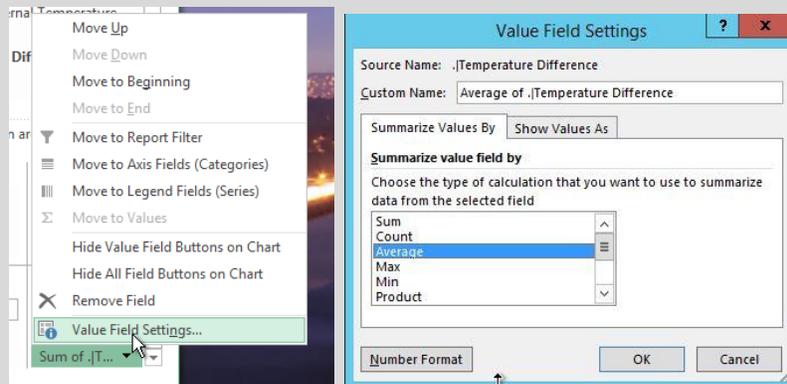
Step 6 : Select the **.|Lost Production (gal)** line and drag into the *Values* area. The aggregation applied for these numeric values is SUM. Select the **.|Temperature Difference** line and drag into the *Values* area. Change the aggregation type to Average.

Your PIVOT table is extended 2 additional columns, which summarize corresponding production losses and temperature differences, based on the reason codes:



**Tip1:** if the PivotTable Fields pane was closed and you want to have it available again, select a cell of your PivotTable. From the right-mouse button menu, select **Show Field List**.

**Tip2:** to change the aggregation that is applied to your data, select the dropdown icon on the field, and choose Value Field Settings... to select another aggregation type.



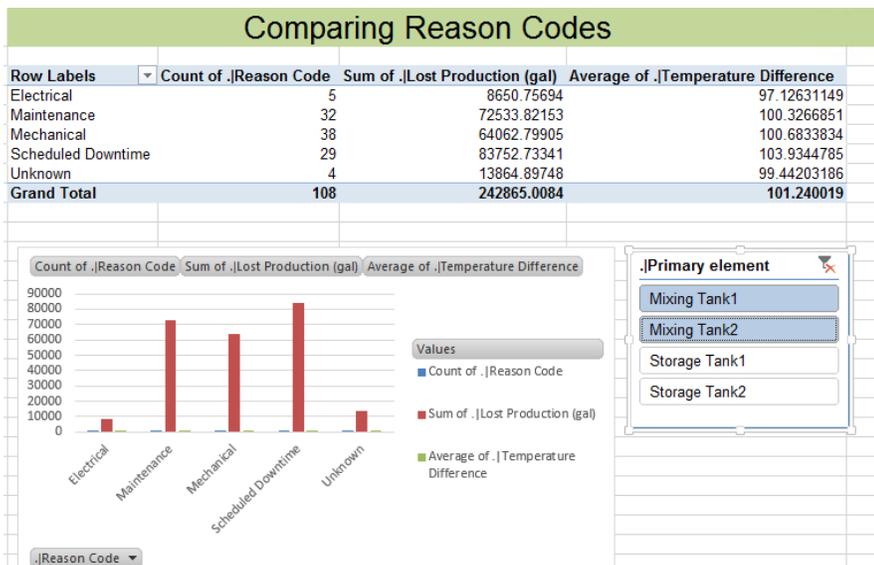
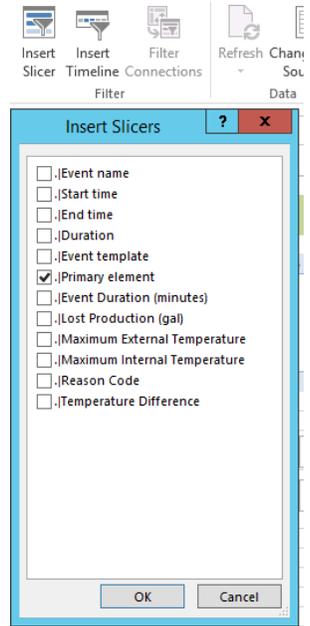
Step 7 : Let's enhance our Pivot table for analysis depending on individual tank selections.

Which column of our data represents a tank?

Step 8 : Select a cell in the Pivot Table, and select the Analyze ribbon from the Pivot Table tools. Click on Insert Slicer, select **.|Primary Element** and click on OK.

Step 9 : Add a slicer for the primary element, which allows you to select any combination of one or more tanks for our analysis. Check various combinations (use Shift- and Ctrl-key for selections in the slicer):

- all tanks
- Mixing Tank1 only
- all Mixing Tanks



The PivotTable and the PivotChart will update to show you what reason code is causing most of the downtime events. In the screenshot above, during the observed period of time, Electrical and Unknown events were the least common sources of downtime.

## EXERCISE : Comparing Downtime Events Based on Tank

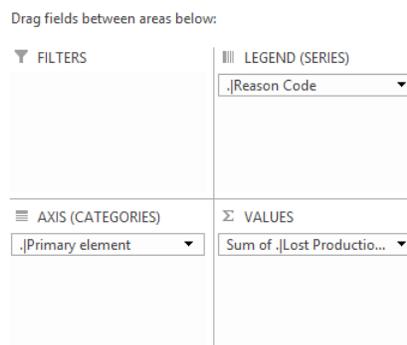
Now that we have determined which downtime causes are most frequent, we need to determine which tank should get our attention first.

3.5

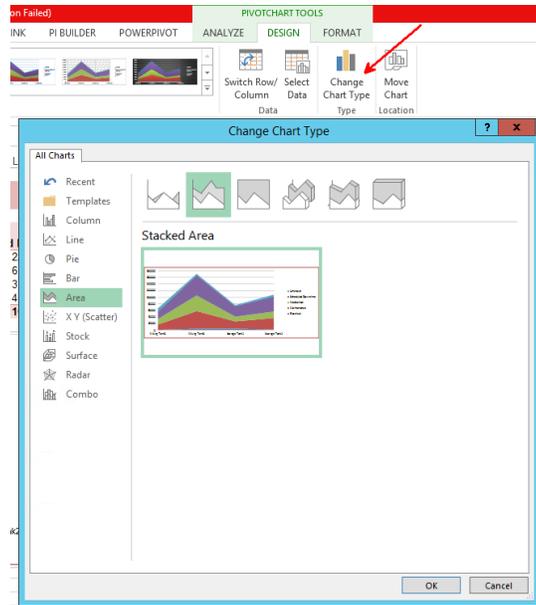
Step 1 : To create a report to compare downtime events based on tanks, we will follow a similar procedure as above. Repeat the procedure outlined in Steps 1-4 above to create a new PivotChart and PivotTable with the same Event Frame data.

Step 2 : To perform a downtime analysis for our Event Frames based on the corresponding tank, select the **.|Lost Production (gal)** line and drag into the *Values* area. The applied Aggregation for the reason codes should be SUM. Select the **.|Primary element** line again and drag into the *Rows* area

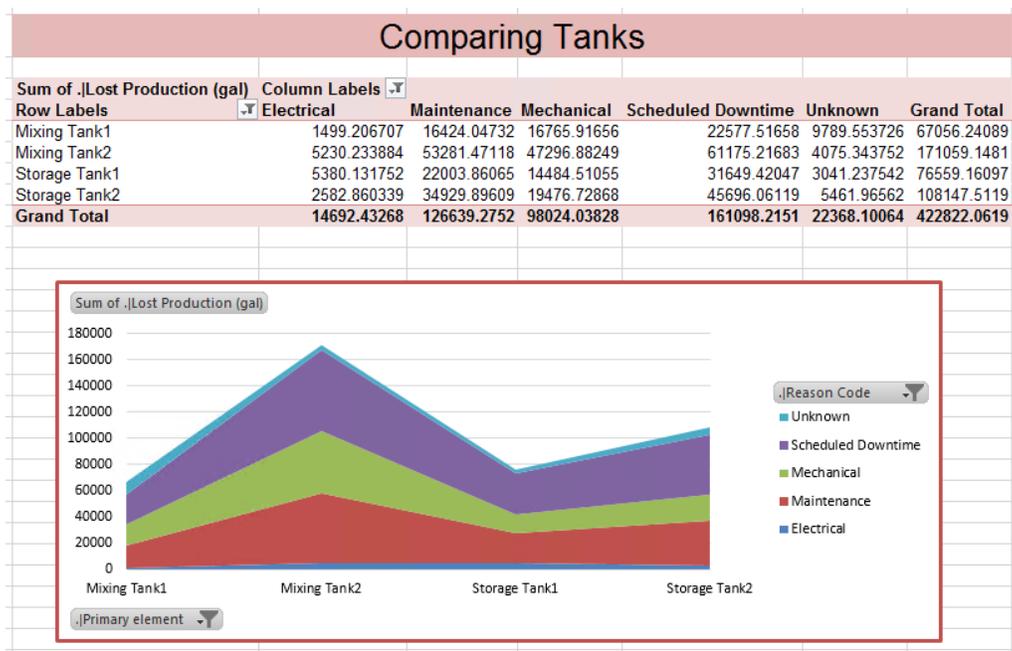
Step 3 : To add an additional visual, drag the **.|Reason Code** to the *Legend (Series)* area.



Step 4 : Select the **PivotChart**, then go to the **Design** tab and choose **Change Chart Type**. Choose the *Area* category on the left, then choose *Stacked Area* as the **Chart Type**.



The result is a report that shows us which tank has had the most production loss, along with a quick glimpse of which reason codes are causing the production loss.



# Analyze Downtime Events Using PivotTables and PivotCharts

## Objectives

- Analyze Event Frames retrieved from PI DataLink using PivotCharts and PivotTables

3.6

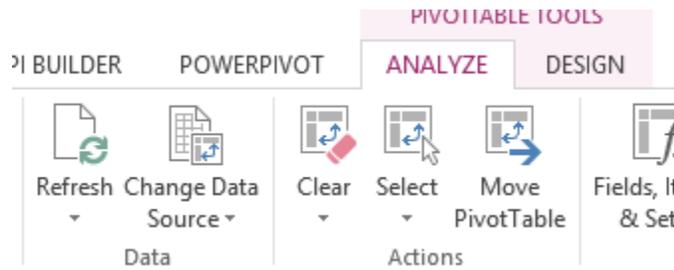
## Problem Description

The operations manager needs a report that shows which downtime reason is most prevalent and a comparison showing which tanks are the most problematic. He also would like to see information about the total production loss from the tanks.

Now that you have imported event data into Excel and created PivotTables and PivotCharts, use them to analyze the data.

## Approach

- Step 1 : You will use the template provided in sheets *Downtime Raw Data* and *Evaluating Tank Downtime* of the file *PI\_DataLink-Exercises.xlsx*.
- Step 2 : Open the *Evaluating Tank Downtime* sheet.
- Step 3 : Select the PivotTable under *Comparing Reason Codes*, then, from the Analyze tab, select **Refresh**



- Step 4 : Select the PivotTable under *Comparing Tanks*, then, from the Analyze tab, select **Refresh**
- Step 5 : Which Reason Code caused the most production loss overall?  
\_\_\_\_\_
- Step 6 : Which Reason Code caused the most production loss for the Storage tanks?  
\_\_\_\_\_
- Step 7 : Which Tank has caused the most production loss? \_\_\_\_\_

## **Bonus: Making Reports More User-Friendly**

### **Objective**

- Leverage out-of-the-box Excel functionality to make your report easier to use.

3.7

### **Problem Description**

You would like to put some finishing touches on your workbook so that it is easier for others to use.

### **Approach**

Step 1 : Hide the sheet containing raw data.

Step 2 : Configure the PivotTables to automatically refresh data when the workbook is opened. (As a result, the PivotCharts will automatically refresh as well).

Step 3 : Prevent column widths and cell formatting from adjusting on data refresh so the PivotCharts will print consistently.

## Solutions

### Analyze Downtime Events Using PivotTables and PivotCharts

- 3.8
1. **Which Reason Code caused the most production loss overall?** Use the Comparing Tanks PivotTable. Find the reason code corresponding to the highest individual value in the Grand Total row (H10:L10).
  2. **Which Reason Code caused the most production loss for the Storage tanks?** Use the Row Labels filter to choose only the storage tanks, then find the reason code corresponding to the highest individual value in the Grand Total row.
  3. **Which Tank has caused the most production loss?** Make sure all tanks are selected in the filter for the Comparing Tanks PivotTable, then look at which tank has the highest sum of .|LostProduction (gal) in the PivotChart.

### Bonus

1. Right-click the Downtime Raw Data sheet and select Hide
2. Select the PivotTable, then go to Tools > Analyze > PivotTable > Options > Data > Refresh data when opening file ([Microsoft doc](#))
  - Or VBA, although we recommend using out of the box functionality instead if possible. Example: when source data changes via [Excel Campus](#)
3. PivotTable Tools > Analyze > PivotTable > Options > Layout & Format > check **Autofit column widths on update** and **Preserve cell formatting on update** ([Microsoft doc](#))

## 4. Directed Activity: Improve Performance of Large Workbooks

### Objectives

- Prevent Excel from recalculating when you don't want it to
- Use Microsoft's built-in Inquire add-in to create a summary of what a large workbook contains
- Measure how many data calls PI DataLink makes in the background when opening a workbook or recalculating
- Use Excel's Formula Auditing tools to see relationships between formulas
- Improve performance of large workbooks by substantially reducing the number of calls PI DataLink must make

### Background

Under the hood, PI DataLink takes the information entered into the task pane for a given function and creates an AF SDK call. So the more PI DataLink functions you have, the more AF SDK calls must be made to retrieve data. Thus, it is important to keep in mind how you can minimize calls when building out or modifying large workbooks containing many hundreds of formulas (or more).

This section will teach you some best practices when troubleshooting why workbooks containing many PI DataLink functions open or recalculate slowly. You will learn how to minimize unnecessary recalculations (each of which could take minutes); quantify the number of calls made by PI DataLink to its data access layer, the AF SDK; and finally, reduce the number of calls necessary by using cell ranges as input and/or bulk data retrieval methods.

For more detailed information on architecture and data flow, see PI DataLink's [playbook](#) in the Customer Portal. For your convenience, here is the data flow section:

#### Data Flow

PI System data is retrieved by PI DataLink into Excel in the following manner:

1. User inputs parameters via the function pane
2. Recalculation is initiated via OK/Apply or right-click > Recalculate (Resize) Function
3. PI DataLink takes the input parameters and forms an AF SDK call
4. AF SDK handles reaching out to back-end resources (PI Data Archive, AF Server) and calculations
5. AF SDK returns values to PI DataLink
6. PI DataLink creates output appropriate for Excel and pastes values into the worksheet

## Directed Activity: Measure AF SDK calls made in the background

### Objective

- Measure how many data calls PI DataLink makes in the background when opening a workbook or recalculating

4.3

### Problem Description

How do you quantitatively measure the background activity to recalculate the PI DataLink functions in this workbook? Specifically, how do you log the calls PI DataLink makes to its data access layer, the AF SDK, to retrieve PI System data?

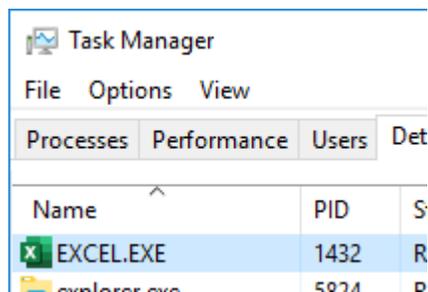
### Approach

Step 1 : Open Excel

Step 2 : Open the **Task Manager** (right-click the task bar, or Ctrl+Shift+Esc)

Step 3 : Right-click on **Microsoft Excel (32 bit)** and **Go to details**

Step 4 : What is the process ID (**PID**) for the highlighted EXCEL.EXE process? Write it down.  
Note: if you do not have a **PID** column, add it by right-clicking a column header > **Select columns**.



Step 5 : Open Windows Explorer and go to %pihome64%\AF. (On the lab machine, this is equivalent to **C:\Program Files\PIPC\AF**)

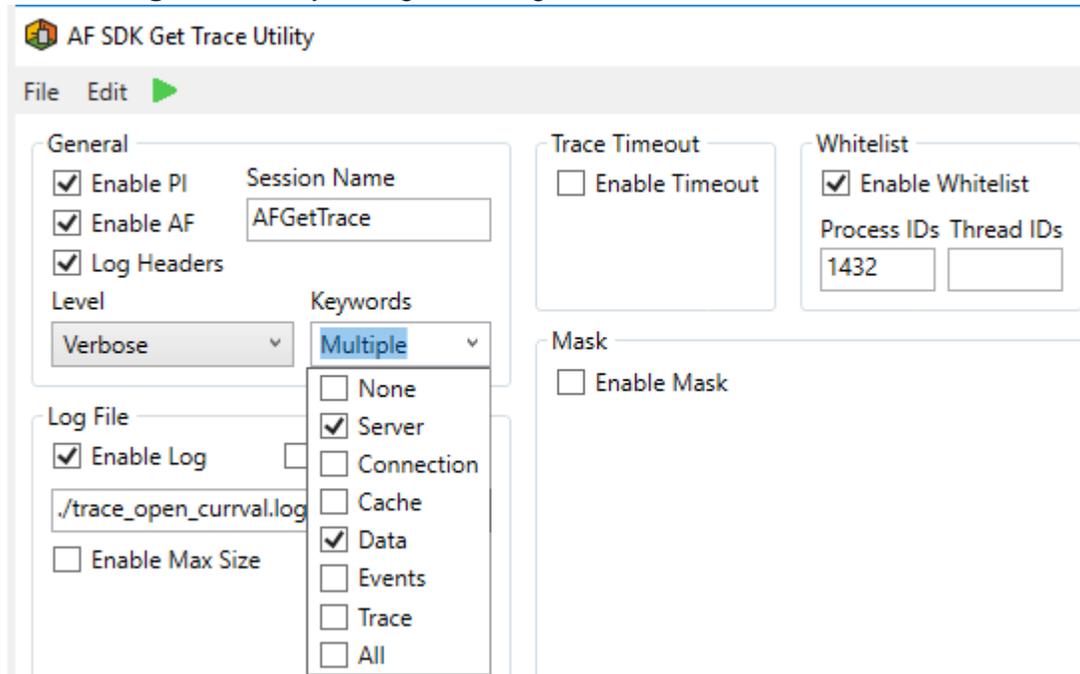
Step 6 : Launch **AFGetTrace.exe**.

Step 7 : Configure the tracing settings: **Enable Whitelist** and enter Excel's PID in the **Process IDs** field

Step 8 : Verify **Enable PI**, **Enable AF**, and **Log Headers** are all checked

Step 9 : Select **Server** and **Data** keywords

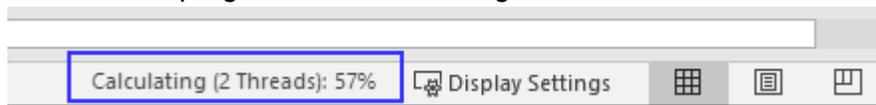
Step 10 : **Enable Log** and name your log/select log location



Step 11 : Press the green, triangular **Play** button

Step 12 : Open **Performance Exercise.xlsx**

Step 13 : Notice the delay in opening as the workbook recalculates. Excel estimates its recalculation progress in the bottom right:



Step 14 : Save the trace

Step 15 : Close the workbook while keeping Excel open (e.g. Book1 is still open)

Step 16 : Go to **File > Options > Formulas > Calculation Options** and select **Manual**, then uncheck **Recalculate workbook before saving**. Alternatively, in the Formulas tab, go to **Calculation > Calculation Options** and select **Manual**. Note this second method does not give you an option to control the save behavior.

Step 17 : Change the name of your log and start a new trace

Step 18 : Open **Performance Exercise.xlsx** again, and notice the difference in open speed because a recalculation was not triggered.

Step 19 : Save the trace

Step 20 : Change the name of your log and start a new trace

Step 21 : Switch the calculation mode back to **Automatic**.

Step 22 : Notice how any change you make to the workbook results in a lengthy recalculation. For example, type "test" into cell N3, then delete it.

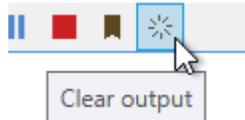
Step 23 : Review the three traces you just generated:

- when opening the workbook in the default Automatic calculation mode
- when opening the workbook in Manual calculation mode
- and when making any change to the workbook

Step 24 : Which trace is empty?

Step 25 : To determine the exact number of AF SDK calls made for a particular action, use the AFGetTrace log and Notepad++'s Find All in Current Document functionality.

Step 26 : For example, clear the AFGettrace log output



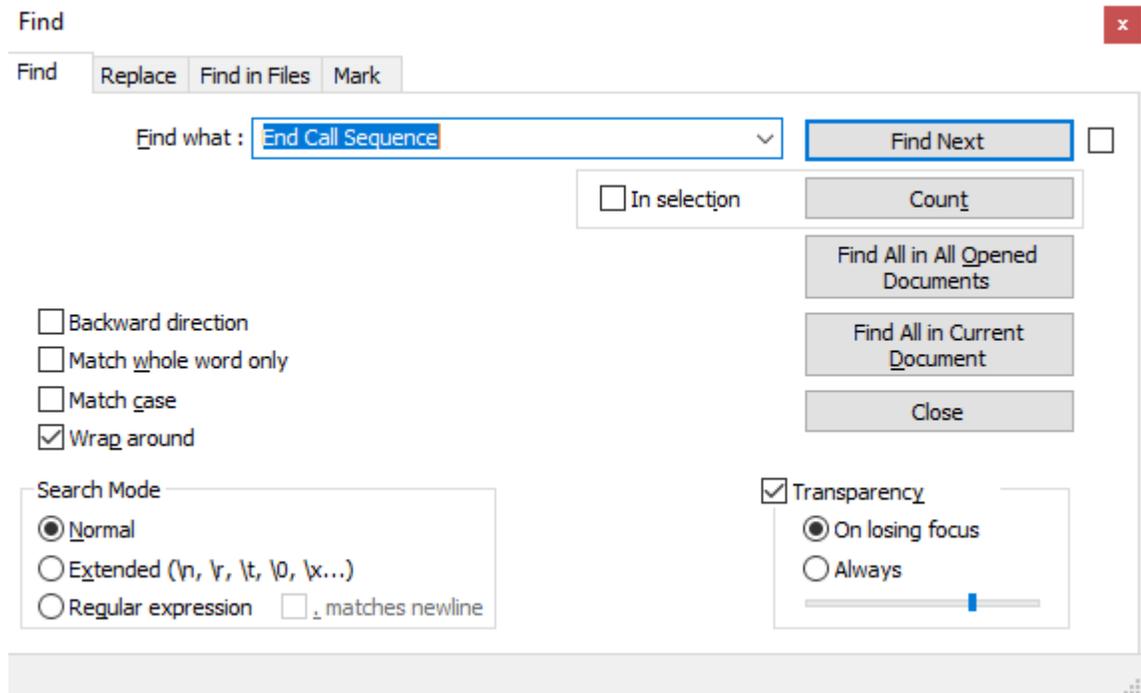
Step 27 : Perform an action that you saw results in AF SDK calls, such as loading the workbook or making a change while in Automatic mode.

Step 28 : Stop the trace (red square).

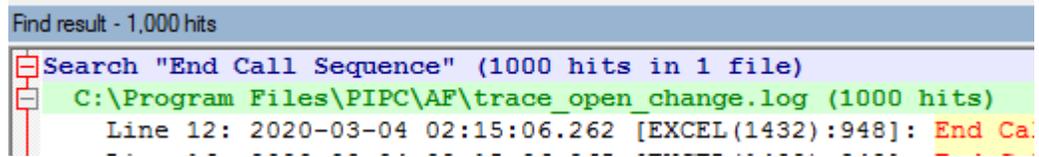
Step 29 : Open the log in Notepad++

Step 30 : Highlight "End Call Sequence" and press Ctrl+F to open the Find window with those search terms auto-filled

Step 31 : Select Find All in Current Document



Step 32 : See how many call sequences exist in the log



Step 33 : Calculate the total recalculation time by subtracting the first **Begin Call Sequence** time from the last **End Call Sequence** time

Step 34 : Calculate the average call sequence time. A relatively fast baseline is about 3-4 milliseconds per call.

Step 35 : How many calls are made when opening the workbook in the default Automatic mode? How much total time does this data access layer activity take?

Step 36 : How many calls are made when making any change to the workbook? How much total time does it take?

## Directed Activity: Gather information about the workbook

### Objective

- Prevent Excel from recalculating when you don't want it to
- Use Excel's Formula Auditing tools to see relationships between formulas
- Use Microsoft's built-in Inquire add-in to create a summary of what a large workbook contains

4.4

### Problem Description

The operators for a plant your company recently acquired have been using a workbook for a long time, designed by someone who left or retired years ago. It has been kind of slow for a long time, but it's finally gotten so bad they are now asking you to fix it. You've never seen this workbook before. How do you analyze the workbook to find out a likely root cause?

### Approach

Step 1 : Open **Performance Exercise.xlsx**

Step 2 : Select the upper left cell in the block of **PIAdvCalcVal** functions (C2).

Step 3 : In the Ribbon, select **Formulas > Formula Auditing > Trace Precedents**

Step 4 : Arrows point to the referenced cells containing the data item (B2), start time (C2), and end time (D2):

	B	C	D
Data Item	2/23/2020 0:00	2/24/2020 0:00	2/24/2020 0:00
TestTag1_1m	43.37546399	43.92727177	

Step 5 : Notice each date cell in row 1 depends on the cell to the right, such as D1 referencing E1.

Step 6 : **Remove Arrows** and select the right-most date cell (M1).

Step 7 : Select **Trace Dependents**. Notice the huge number of arrows generated:

	L	M
30	3/3/2020 0:00	3/4/2020 0:00
37	45.68610187	
36	59.54379406	
46	77.69614246	
33	85.26468703	
33	15.48875883	
59	53.40881068	

Clicking **Trace Dependents** additional times will show successive dependents, eventually going all the way to column C. This is a visual reference of how the entire

block of PI DataLink calls is influenced by this one date cell with a formula of TODAY(). We will discuss the performance impact of this function in the next exercise.

Step 8 : Switching gears, read Microsoft's [article](#) for more broad coverage of their Inquire add-in than will be presented in this lab.

Step 9 : Go to **File > Options > Add-ins > Manage: COM Add-ins** and press **Go...**

Step 10 : Check **Inquire** and press **OK**. An Inquire tab should load in the Ribbon.

Step 11 : Select **Workbook Analysis**. If prompted, do not save changes.

Step 12 : In the **Items** pane, check **Summary**, **Workbook**, and **Formulas**, then press **Excel Export...**

Step 13 : Save the analysis, then press **Load Export File**.

Step 14 : How many formulas, number of formulas with errors, and hidden/very hidden sheets are in this workbook?

## Improve performance by reducing the number of background calls

### Objective

- Improve performance of large workbooks by reducing the number of calls PI DataLink must make
- 4.5
- Learn about volatile functions in Excel recalculation
  - Learn which Excel and PI DataLink functions are volatile

### Problem Description

Now that you know this workbook is making a large amount of relatively fast data calls, how can you help performance by reducing the number of calls?

Read the [Volatile and Non-Volatile Functions](#) section of Microsoft's Excel Recalculation article. This is also discussed in certain OSIssoft KB articles, such as [After upgrading Microsoft Office, opening complex spreadsheets using PI DataLink is much slower](#).

After reading, you should be aware of which Excel functions are volatile. The only PI DataLink function that is volatile by default is Current Value. Why?

Insert a column into the workbook that contains the current value for each data item. Do so in such a way that only one call sequence is made to populate the current values.

Bonus: verify only one call sequence is made using AFGetTrace.

### Approach

Step 1 : Open **Performance Exercise.xlsx**

Step 2 : Select cell M1 containing the volatile formula, **TODAY()**

Step 3 : Type today's date or press **Ctrl+;**

Step 4 : Using AFGetTrace, observe that AF SDK calls no longer occur when making an unrelated change to the workbook, such as typing "test" in cell N3

Step 5 : To recalculate the entire block of PI DataLink formulas, go to **PI DataLink > Resources > Settings** and select **Full calculate, 0s Interval** for the **Automatic Update**. Then click the green **Update** button.

Step 6 : Using AFGetTrace, observe 1000 AF SDK calls are made for each recalculation cycle.

Step 7 : Click **Update** again to turn auto-update off.

Step 8 : Select cell C2 and notice the formula references a single data item, **\$B2: {=PIAdvCalcVal('Raw Data!\$B2,'Raw Data!C\$1,'Raw Data!D\$1,"average (time-weighted)","time-weighted",0,1,0,"")}**

Step 9 : Right-click the cell and select **Recalculate (Resize) Function** (or use the **task pane > OK**)

- Step 10 : Notice how *the entire row* recalculated, not just that single cell, and the formula now shows the contiguous range of data items (\$B\$2:\$B\$101) instead of the single data item.
- Step 11 : Repeat for each column so that there are ten formula arrays, each referencing 100 data items, instead of 1000 individual formula arrays.
- Step 12 : Turn **Update** back on
- Step 13 : Using AFGetTrace, observe the number of AF SDK calls has been reduced by an order of magnitude. This is because these formulas now meet the conditions under which PI DataLink is able to take advantage of bulk AF SDK calls, as documented in the user guide's [Retrieval of large amounts of data](#) topic.

## Solutions

### Directed Activity: Measure AF SDK calls made in the background

- 4.6
1. **Calculate the average call sequence time.** Total time / # calls = time/call. So 3.8/1000 = 0.0038s, or 3.8 milliseconds.
  2. **How many calls are made when opening the workbook in the default Automatic mode?** Should be 1000.
  3. **How much total time does this data access layer activity take?** Should be around 3-4 seconds. During testing, it took 3.8s.
  4. **How many calls are made when making any change to the workbook? How much total time does it take?** # of calls and time should be very similar to the open, as the functions all being volatile results in all of them recalculating on change, just like they do on workbook open.
  5. **Which trace is empty?** The one covering workbook open when in Manual calculation mode should be empty.

### Directed Activity: Gather information about the workbook

1. **How many formulas, number of formulas with errors, and hidden/very hidden sheets are in this workbook?** Use the Inquire report you just generated. Example:

	A	B	
1	<b>Summary</b>		
2	<a href="#">https://github.com/microsoft/officeapps-sdk/blob/master/docs/insights/PI World 2020</a>		2020
3			
4	<b>Item</b>		<b>Value</b> <b>Rev</b>
5	Creation Date	Monday, March 2, 2020 8:03:24 PM	
6	Modified Date	Wednesday, March 4, 2020 1:23:52 AM	
7	File Size (bytes)	48,088	
8	Title		
9	Author	Jeremy Roberts	
10	Linked Workbooks	0	
11	DDE Links	0	
12	Data Connections	0	
13	Visible Sheets	1	
14	Hidden Sheets	0	
15	Very Hidden Sheets	0	
16	Formulas	1,111	
17	Array Formulas	1,000	
18	Formulas With Errors	0	
19	Formulas With Logical Values	0	
20	Formulas With Numeric Values	1,000	

## 5. EXERCISE: Excel's New Dynamic Array Functions

### Objectives

- Learn about Excel's new dynamic array functions
  - Create a report using PI DataLink and these new Excel functions
- 5.1
- Brainstorm other ways these functions could prove useful to you

### Background

5.2 Microsoft has recently made changes to Excel's calculation engine that enable "dynamic arrays", which allow you to write one formula and return many values. [As of January 17<sup>th</sup>](#), dynamic arrays and the new functions that leverage them are available in all Office 365 channels except Semi-Annual.

In this exercise, we will go through Microsoft's preview of dynamic arrays. Then, we will see a few ways these new functions could be used with PI DataLink functions.

Note: In most cases, there are other ways to accomplish the same tasks besides these new functions, such as VLOOKUP or INDEX(MATCH()). However, the new functions will generally be simpler and easier to use.

# Learn about Excel's new dynamic array functions

## Objective

- Get hands-on experience with dynamic array formulas

### 5.3 Problem Description

Microsoft's [Preview of Dynamic Arrays in Excel](#) describes several aspects of dynamic arrays and lists multiple dynamic array functions. Read through this article and the articles on specific functions, such as **SORT** and **FILTER**. In addition, read the [article](#) on the **XLOOKUP** function. Then follow along in the workbook.

## Approach

Step 1 : Open *PI\_DataLink-Exercises.xlsx* to the *Dynamic Functions* sheet.

Step 2 : **Sort** the names in column A with a single dynamic function.

Step 3 : Notice a couple of new behaviors for dynamic functions—first, when you select any cell in the resulting array, the entire array is outlined in blue. In addition, only the top left cell in the dynamic array is editable; all other cells in the array have greyed-out formulas, and if you attempt to edit the cell, the formula disappears.

=SORT(E5:E11)	
E	F
<b>Dynamic Arrays</b>	Dt
<b>Sort your data</b>	
<u>Names</u>	<u>Sorted</u>
Mary	Alison
Peter	Charles
John	Jane
Simon	John
Charles	Mary
Jane	Peter
Alison	Simon

Step 4 : **Cleanse** the list of names in column D of **duplicates** with a single dynamic function.

Step 5 : **Filter** the list of names in column G by the diet preferences in column H with a single dynamic function.

Step 6 : With a single function, **look up** the country code for Brazil from the data in columns P and R.

Step 7 : Bonus: Which aspects of this new functionality will work in earlier versions of Excel? What will not work?

## Use dynamic array functions to complement PI DataLink functions

### Objective

- Use dynamic array functions with PI DataLink functions to derive more value out of your data
- 5.4
- Brainstorm other ways these functions could prove useful to you

### Problem Description

You want to enhance your production summaries by sorting production days below a certain threshold by their total production. After hearing about new dynamic functions, you think they could be used to quickly accomplish your goals.

### Approach

Step 1 : Go to the *Production Summaries* sheet

Step 2 : Use the **SORT** and **FILTER** functions to display a list of the days below 200,000 gallons of total production from the *Daily Calculations* table, in ascending order.

Step 3 : Go to the *Downtime Raw Data* sheet

Step 4 : In column T, use **Compressed Data** to retrieve values from yesterday for the tag CDM158.

Note: you will need to enter '**CDM158** for the data item; otherwise, it will be recognized as a reference to the *cell* CDM158.

Step 5 : Use the **UNIQUE** function to retrieve (1) the unique digital states in yesterday's data for CDM158 and (2) the unique reason codes in the event frame data.

Bonus: Sort the digital states and reason codes alphabetically.

## Solutions

### Learn about Excel's new dynamic array functions

5.5

1. Open *PI\_DataLink-Exercises.xlsx* to the *Dynamic Functions* sheet.
2. In cell B5, use the **SORT** function and select the range of names: **=SORT(A5:A11)**
3. In cell E5, use the **UNIQUE** function and select the range of names: **=UNIQUE(D5:D14)**
4. Cell E9 contains a period, which prevents the dynamic array from spilling into the full cell range it requires. This results in a **#SPILL** error. Delete this period, and the **UNIQUE** formula will auto-populate.
5. In cell U4, use the **XLOOKUP** function with a **lookup value** of T4, a **lookup array** of P4:P13, and a **return array** of R4:R13. If it's hard to keep track of the different fields, use **Formulas > Insert Function**. **=XLOOKUP(T4,P4:P13,R4:R13)**
6. **Bonus:** It's probably wise to avoid using dynamic array functionality until your IT department rolls out an O365 update containing dynamic arrays to you/your users. Dynamic arrays will be converted to legacy Ctrl+Shift+Enter (CSE) formulas in older versions of Excel and still work, but dynamic array functions will not be recognized, resulting in **#NAME?** errors.

Remember that if you need to verify compatibility across multiple Office versions, you can use **File > Info > Check for Issues > Check Compatibility**. For more information, read [Dynamic array formulas in non-dynamic aware Excel](#) and [Dynamic array formulas vs. legacy CSE array formulas](#).

### Use dynamic array functions to complement PI DataLink functions

1. On the *Production Summaries* sheet, in cells G20:I20, label columns **Criterion**, **Day**, and **Production Total**.
2. Enter **200,000** into cell G21
3. In cell H21, filter the **Daily Calculations** range containing the **start time** and **Production Total (gal)** by the **criterion**: **=FILTER(A21:B27,B21:B27<G21,"")**
4. Now wrap the filtered range in a **SORT** to specify **ascending order**: **=SORT(FILTER(A21:B27,B21:B27<G21,""),2,1)**  
Note: This is an application of the example given in the **SORT** function's [documentation](#).
5. On the *Downtime Raw Data* sheet, in cell T7, use **Compressed Data** to retrieve yesterday's values for the tag CDM158: **{=PICompDat("cdm158","y","t",1,"","inside")}**
6. In cell W8, use the **UNIQUE** function with the values from the **Compressed Data** function, and use an ascending **SORT** for alphabetical order: **=SORT(UNIQUE(U8:U1934),,1)**

7. Similarly, in cell Q8, use the **UNIQUE** function with the values in the **Reason Code** column: **=SORT(UNIQUE(L8:L36),,1)**

## 6. EXERCISE: Introduction to PI DataLink Documentation

### Objective

- Learn about key articles in OSIsoft Technical Support's Knowledge Base for PI DataLink

6.1

### Background

6.2 PI DataLink has been and continues to be one of OSIsoft's most-used client tools. Over the past two decades, OSIsoft Technical Support has answered well over 100,000 calls concerning PI DataLink. As you might expect, we have created documentation along the way to help us efficiently answer these calls and deflect some questions outright as customers can find what they need in our Customer Portal without our assistance.

This section will give you an overview of the most-used articles for PI DataLink so that the next time you have a question about PI DataLink the user guide doesn't answer, you will know how to find the same resources we use!

## Access PI DataLink's user guide in context

### Objective

- Access the PI DataLink user guide in the context of your current task
- Understand the structure of the PI DataLink user guide

6.3

### Problem Description

PI DataLink does not appear in the Ribbon, and you wish to search and view OSIssoft Knowledge Base articles to help you troubleshoot.

### Approach

- Step 1 : You will use the template provided in sheet *Operational Start Up* of the file *PI\_DataLink-Exercises.xlsx*.
- Step 2 : Select a cell containing a PI DataLink function, such as cell A11, and the task pane will appear.
- Step 3 : Select any field in the Compressed Data task pane and press F1
- Step 4 : Alternatively, in the top right of the task pane, click the question mark.
- Step 5 : This brings up the **Compressed Data function** topic in the local PI DataLink User Guide, which describes each possible input for the function
- Step 6 : Navigate down a level to the **Compressed Data example** topic. (Each PI DataLink function topic has an example.)
- Step 7 : You can also use the function reference for more formula-focused help. Return to cell A11 and view the formula.
- Step 8 : In the user guide, navigate to the **Function reference** topic, select the group matching the Ribbon group for your function, then select the exact function. For example, Function reference > Multiple-value functions > PISCompDat() for Compressed Data in cell A11
- Step 9 : Go to the **Downtime Raw Data** sheet and view the documentation for the Compare Events function

## Directed Activity: Access the Customer Portal

### Objective

- Configure your account (if necessary)
- Access the Customer Portal

6.4

### Problem Description

PI DataLink does not appear in the Ribbon, and you wish to search and view OSIsoft Knowledge Base articles to help you troubleshoot.

### Approach

Step 1 : Go to <https://customers.osisoft.com> and log in with your OSIsoft single sign-on (SSO) account

 <p>Tip</p>	<p>If you do not have an SSO account or have not been added to a site, at your convenience go to <a href="https://explore.osisoft.com/myosisoft-customer-portal/how-to-get-a-login">https://explore.osisoft.com/myosisoft-customer-portal/how-to-get-a-login</a> and follow the step-by-step instructions there.</p>
--	--

Step 2 : In the top search bar, search for “datalink playbook”

Step 3 : Click on Playbook – PI DataLink

Step 4 : In the Troubleshooting section under the Runtime list, select “PI DataLink or PI Builder does not appear in the Ribbon in Excel”

Step 5 : Instead of using the playbook, now access the same article via direct search. Search for your issue to find the article, filtering by the **KB Articles** content type if necessary.

Step 6 : Pretend two months have passed, and you vaguely remember something about an AF tracing utility you used in a PI World lab. You think it could be useful to troubleshoot your current situation but can’t recall exactly how to use it. Search to find an article describing how to use the utility.



Have an idea how to  
improve our products?  
**OSISOFT wants to hear  
from you!**

<https://feedback.osisoft.com/>





REGISTER YOUR INTEREST

## Save the Date!

OSISOFT PI World Users Conference in Amsterdam; October 26-29, 2020.

Register your interest now to receive updates and notification early bird registration opening.

[https://pages.osisoft.com/UC-EMEA-2020-Q4-10-26-PI-World-AMS-Register-Your-Interest\\_Register-Your-Interest.html](https://pages.osisoft.com/UC-EMEA-2020-Q4-10-26-PI-World-AMS-Register-Your-Interest_Register-Your-Interest.html)



## CONTINUE YOUR PI SYSTEM LEARNING



After the conference, the **PI SYSTEM LEARNING** does not have to stop. All registered attendees for **PI World SFO 2020** will have access to all PI World Hands-on Lab cloud environments for 21 days using the discount code below. You will receive detailed instructions via email after the conference.

**Discount Code: 2020PIW SF-LAB-100**

Offer expires July 3, 2020