

Data Analysis in Contour

Course Introduction

About Contour

Contour is a powerful data analysis and visualization tool on the Palantir Foundry platform. It is designed to handle complex data transformations and visualizations with ease, providing a user-friendly interface for a broad variety of analyses.

Contour is structured to provide a linear sequence of dependent blocks of logic that operate on data. These data operations can include filtering, transforming, and visualizing data. The linear sequence of data operations can be organized into multiple paths. By structuring your analysis into paths, your analysis is easily interpretable and manageable.

Contour helps you perform various data analysis tasks, including:

- Data cleaning and transformation
- Data visualization
- Save analysis results as a new dataset for use in other Foundry tools.
- Leverage the Contour expression language for more advanced transformations and aggregations.
- Creating and managing dashboards

When to Use Contour vs. Other Tools

- **Contour:** Use Contour when you need to perform data cleaning steps, generate transformations, and build visualizations on datasets in the Foundry platform within a single tool. It is ideal for creating ad-hoc analyses and for generating regular reporting materials such as dashboards and embedding visualizations in Notepad.
- **Quiver:** Use Quiver for advanced data exploration and analysis on Ontology objects, including for time series data. It is suitable for users who are looking to embed analysis in Workshop applications.
- **Pipeline Builder:** Use Pipeline Builder for building and managing production pipelines in a low- to no-code interface that provides real time feedback. Users can integrate data, create business logic transformations, and define a rigorous release process for production pipelines.
- **Code Repositories:** Use Code Repositories to write and collaborate on production-ready code in Foundry in a web-based integrated development environment (IDE).

The application provides a user-friendly way to interact with the underlying Git repository.

The Use Case

Welcome to your training as a data analyst for the notional TitaniumWorks Manufacturing company! In this course, you'll step into the role of a data analyst tasked with prioritizing which equipment should be inspected next. As a data analyst, you'll have to report out analytic insights often. You'll use Contour to:

- **Save time and manual effort:** Automatically run the same analysis each time data is updated, eliminating the need to manually update analyses.
- **Easily generate analyses:** With point-and-click functionality, you can manage and execute complex analyses for many user types.
- **Easily interpret and present analysis:** Visualizations and automated reports make it easy to collaborate with for non-technical stakeholders to understand your findings and make informed decisions.

Your Role

As a data analyst, your primary goal is to ensure the efficiency and reliability of the manufacturing process. By analyzing the parts purity and the date of the last inspection of each piece of equipment, you will generate a list of equipment that needs inspection.

Your Goal is the Following:

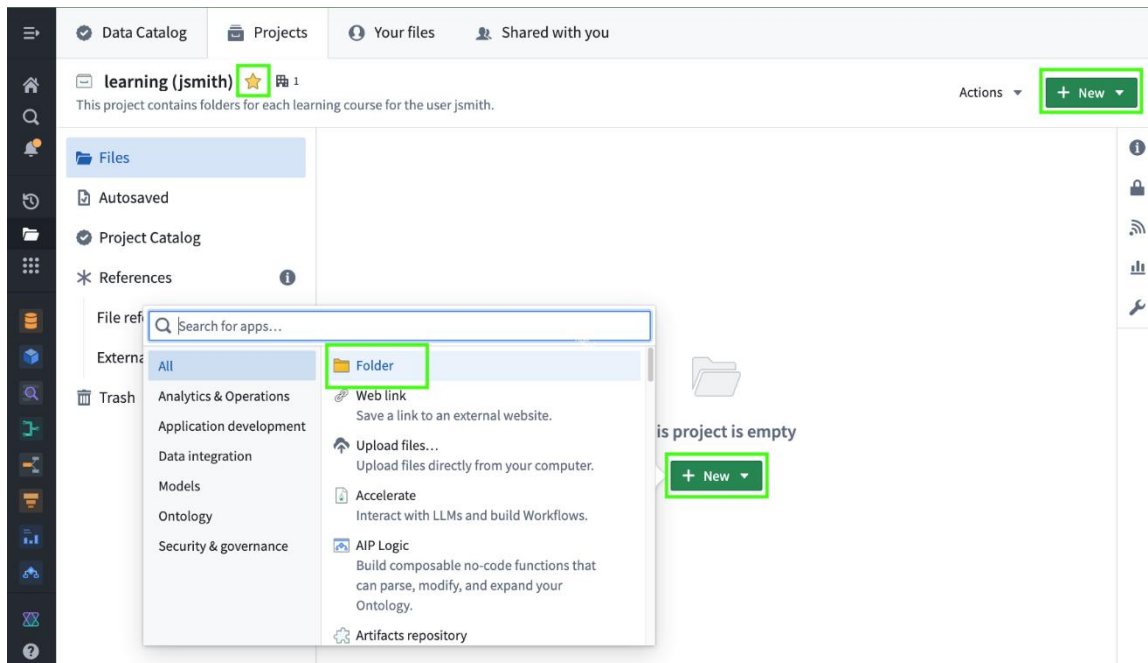
- Clean and prepare the datasets for analysis.
- Join the datasets to create a comprehensive view of the data.
- Create derived columns to gain insights into the maintenance schedule and identify high-risk machinery.
- Visualize the data and create a dashboard for easy interpretation and decision-making.
- Embed the visualizations in Notepad for collaborative analysis.

Create a Course-Specific Training Folder

Step 1: Create a folder

1. In the top left, click on the star to favorite your project
 - This will allow you to find it quickly later on

2. Click **New** in the center of the screen or in the top-right corner
3. Select **Folder**



4. Name the new folder after the current learning course
 - For example: **Deep Dive: Data Analysis with Contour**

Upload Data and Create the Analysis

Download the Data Files

In this exercise, you will upload two datasets to the Foundry platform and create a Contour resource for drill-down analysis, all stored within your own training project and folder for this course.

For the purpose of this training, you'll be working with two datasets - one that stores equipment data and one that stores the production data of parts that are produced by the equipment.

Step 1: Download two files

1. Click on the names of the two datasets below to download them to your local device.
 - You'll then upload these to your instance of Foundry in the next step.

If you know you are unable to manually upload datasets to your Foundry instance due to lack of permissions, you can skip this and the next step and add the datasets to your Foundry instance via Marketplace with the instructions in the following step.

Downloads

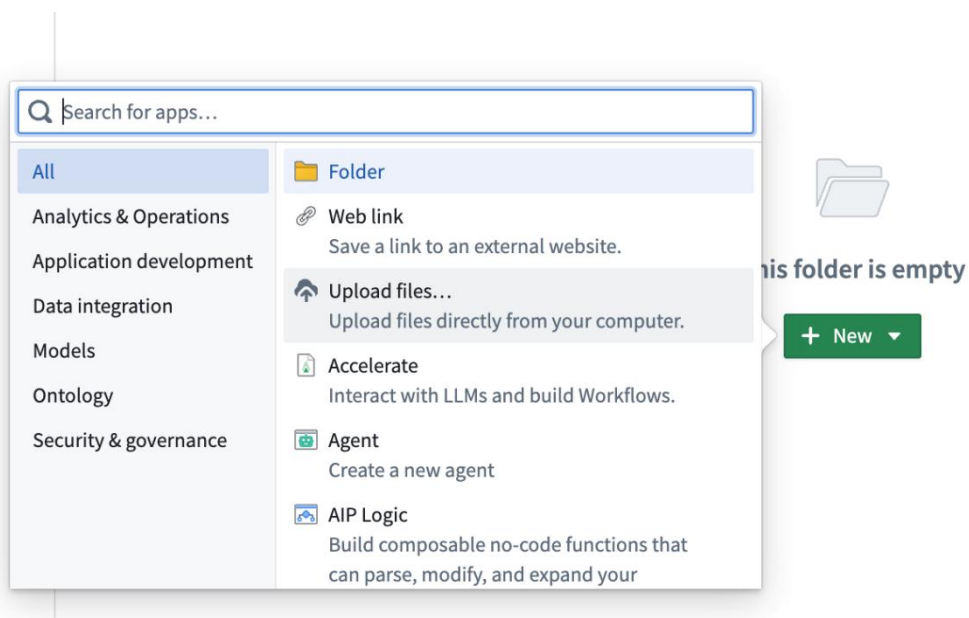
- [equipment.csv](#)
- [parts.csv](#)

Upload the Data Files

Now that you've downloaded the data files for this course, you'll next upload them as datasets to your instance of Foundry so they can be accessed within Contour.


Step 1: Upload files

1. Navigate to the course-specific folder within your training project
2. Click on **New** within the folder
 - Alternatively, you can import files by dragging them directly into your folder.
3. Click **Upload files...**







4. Click **Choose from your computer** or drag the files onto the upload interface

- Select the files downloaded earlier
- 5. Select the first option for **Upload as structured datasets**
- 6. Click **Upload**
- 7. Click **Done**



Drop files here or [choose from your computer](#)

 equipment.csv	619 B	
 parts.csv	275.22 KB	

☒ **Upload as individual structured datasets (recommended)**
Datasets are the most basic representation of tabular data. They can be used and transformed by many different applications. [Documentation](#)

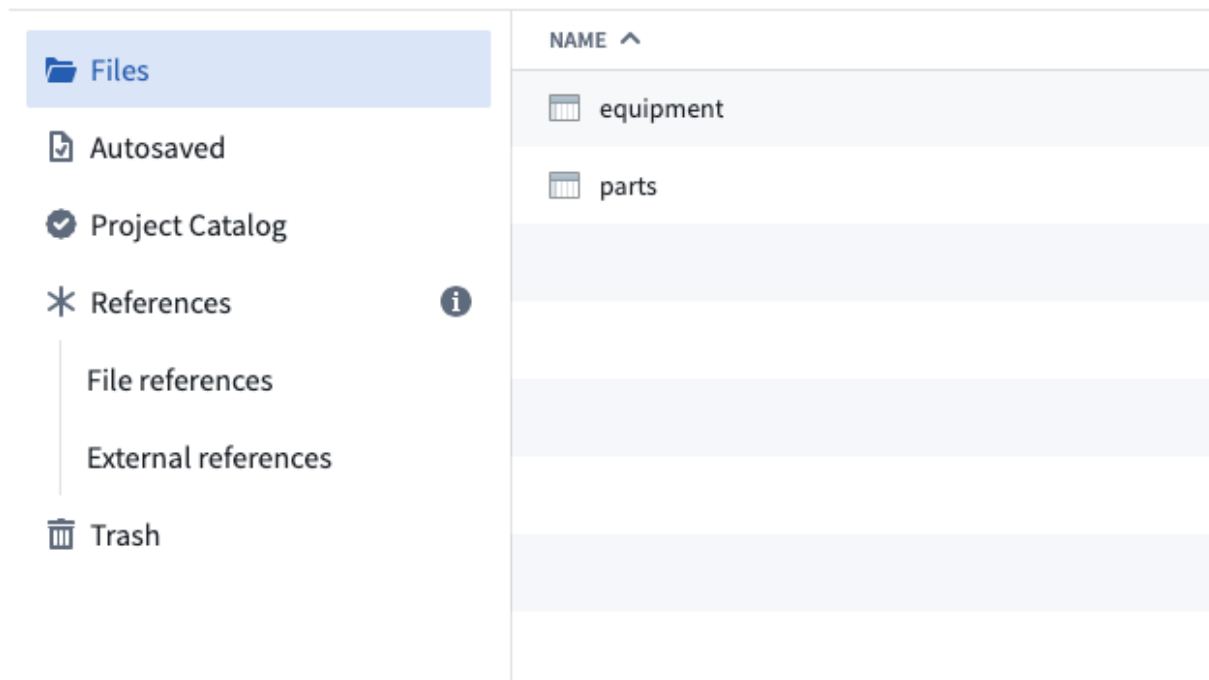
☐ **Upload to a new media set**
Media sets enable media-specific capabilities for media files (e.g. audio, imagery, video, and documents). [Documentation](#)

☐ **Upload to a new unstructured dataset**
Unstructured datasets can store arbitrary files for processing and analysis. Structured data can be extracted from unstructured datasets using Pipeline Builder or Transforms. [Documentation](#)

☐ **Upload as individual raw files**
Raw files cannot be used in data pipelines, analyses, or models.

Upload

You should now see two new datasets stored in your folder for this course

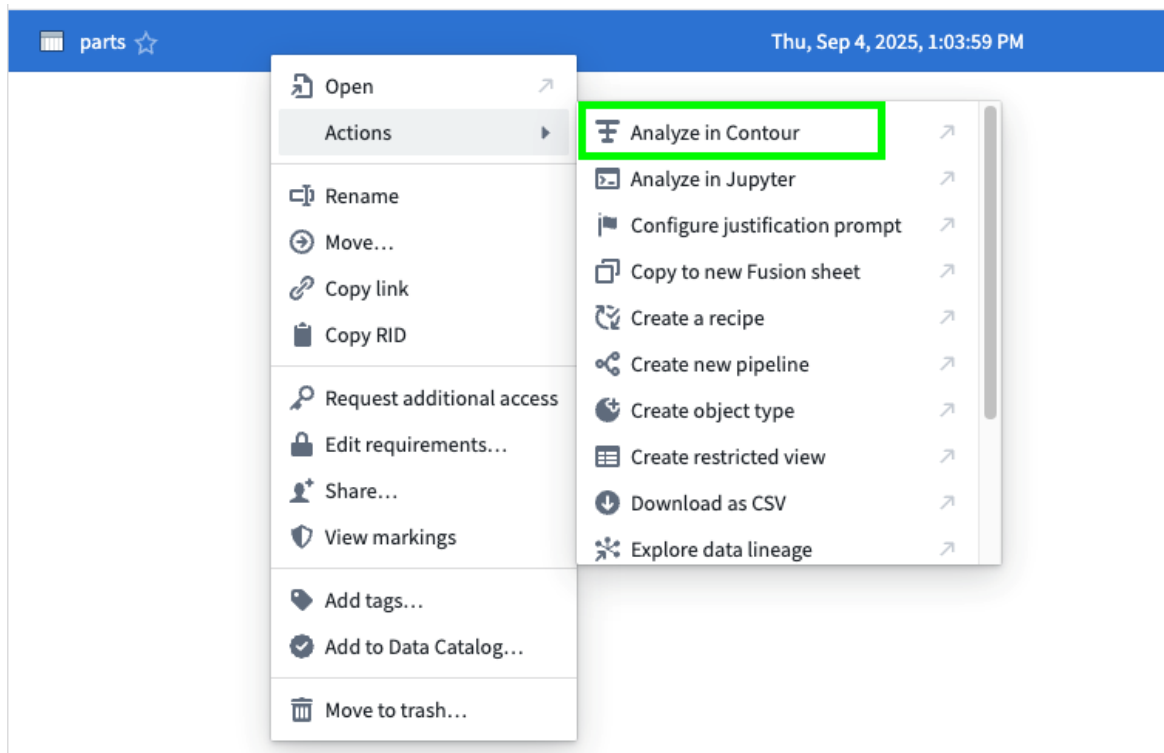


Create Contour Analysis

Next, you'll create your Contour Analysis within your folder for this course and confirm its name.

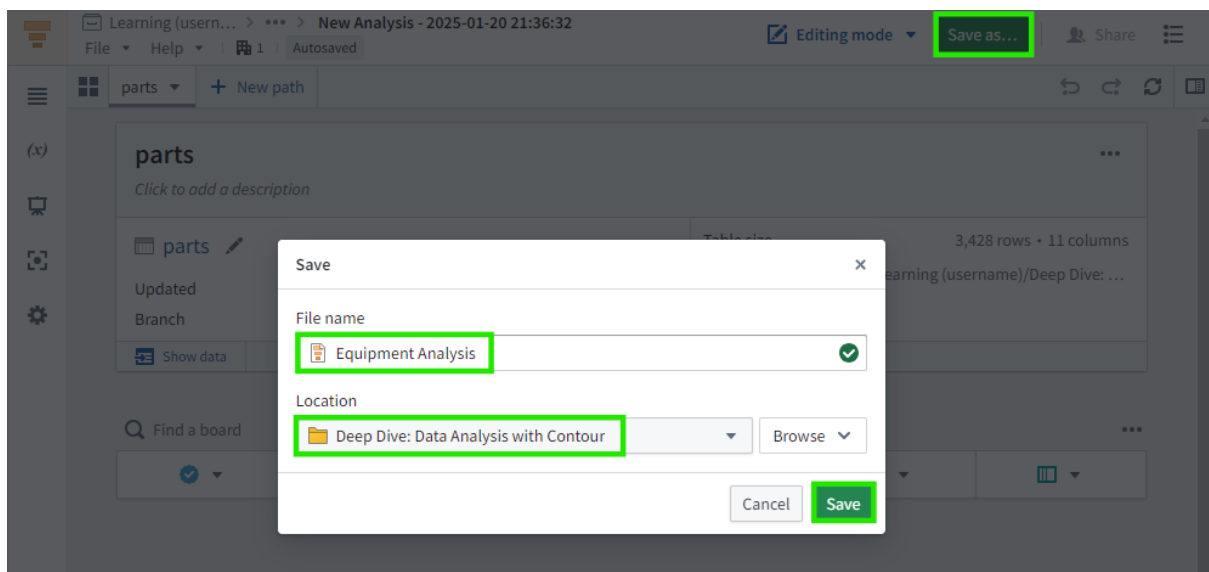
Step 1: Analyze the parts dataset in Contour

1. Right-click on the parts dataset to pull up the task options
 - Or you can use CTRL + click on a Mac
2. Select **Actions > Analyze in Contour**
 - We'll use the equipment dataset later in the training.



Step 2: Save your analysis

1. In the top right of your screen, click on the green **Save as...** button
2. Rename your analysis: Equipment Analysis
3. Confirm that you are saving into your course folder
4. Click **Save**

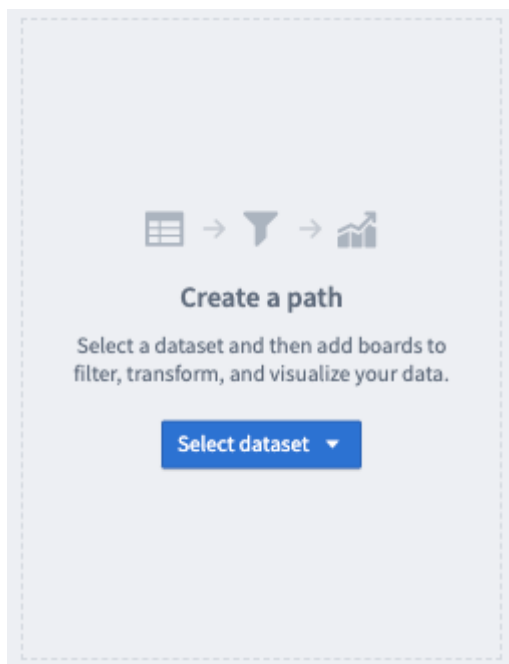


[Optional] Common Settings Configurations

This lesson covers some settings to confirm in Contour that can assist with troubleshooting common issues. You do not need to complete these steps now, but reading through them now will help you recognize common issues that may occur later.

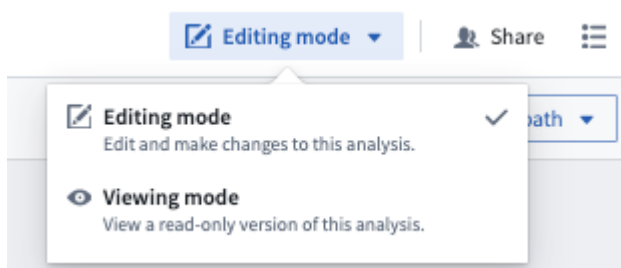
Selecting a Dataset Path

If your screen resembles the picture below at this stage, please click **Select dataset** and then navigate to and select the parts dataset as described in the previous task.



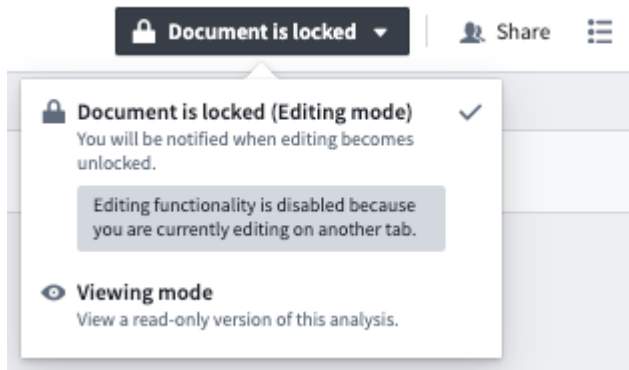
Editing Mode

Before proceeding with the following steps, please check that you're set to **Editing mode** near the top-right corner of your screen.



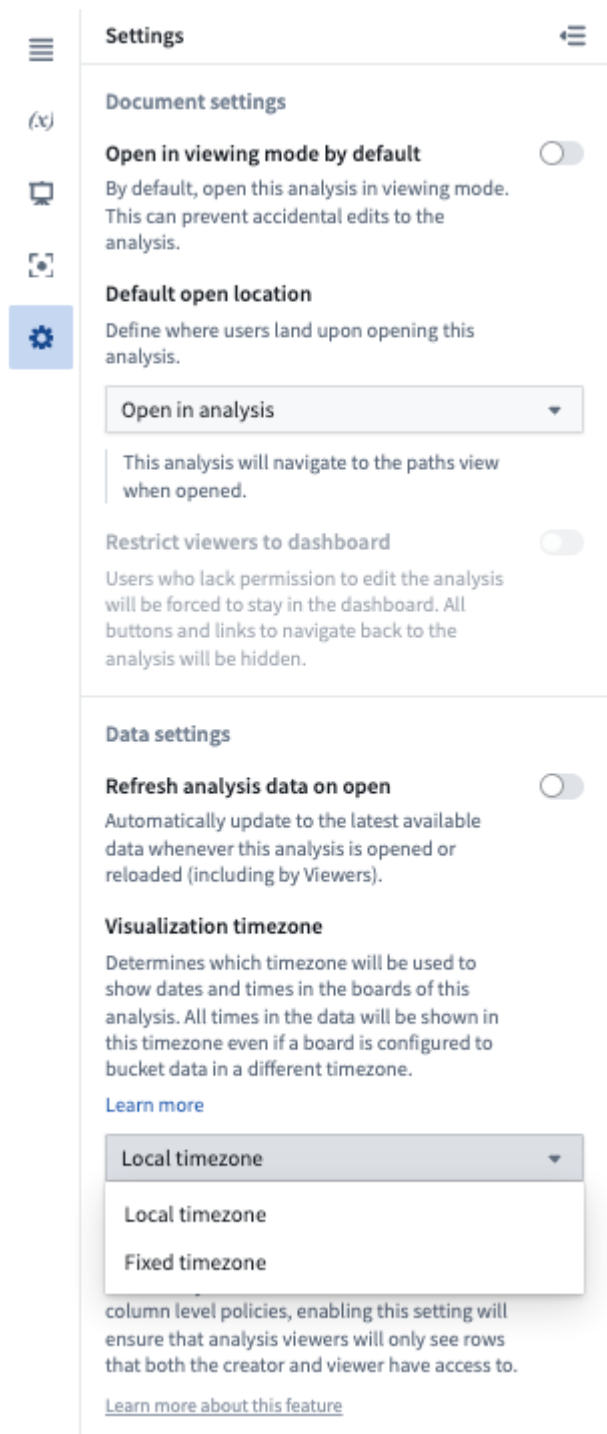
Locked Analysis

If you see an indicator that your Contour analysis is locked as pictured below, please close any other browser tabs opened to the same analysis to proceed.



Timezone Settings

- To see a few options you have for configuring your analysis, click on the Settings button located in the left-hand sidebar (cogwheel icon).
- One setting to note is **Visualization timezone**. Contour implicitly expects timestamps to be in UTC, and you can configure to display all boards to render this timestamp in the local timezone (your current timezone) or a fixed timezone.



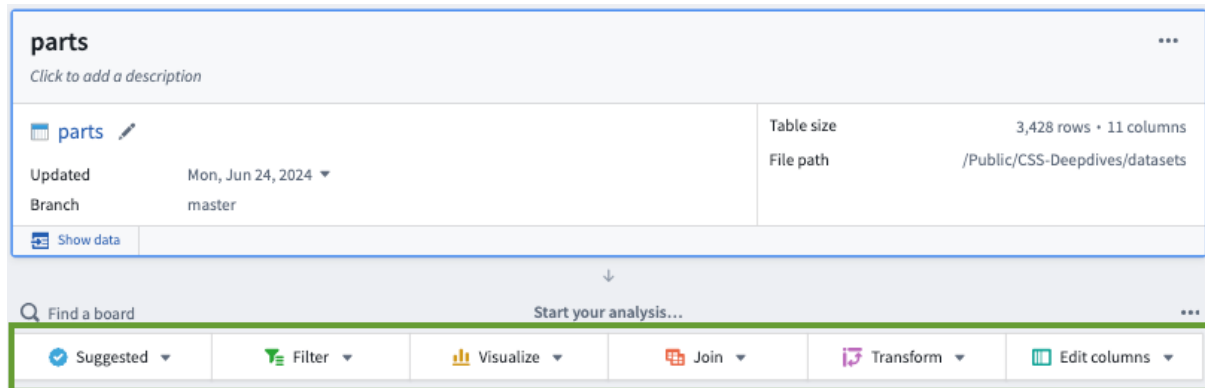
Data Cleaning

Introduction to Boards

In Contour, you build your analysis using a sequence of **boards**. The boards are used to analyze and operate on your data. Each board provides either a visualization or data transformation or both. There are boards to create charts, perform calculations, remove columns, and more.

At the top of the screen, in the first non-removable board, is your input dataset. Data always flows from the top of the page down to the bottom through the boards you have defined. This sequence of boards is called a **path**.

The **board toolbar** is always located under your last board to perform the next logical step of your analysis. Adding another board extends the current path.



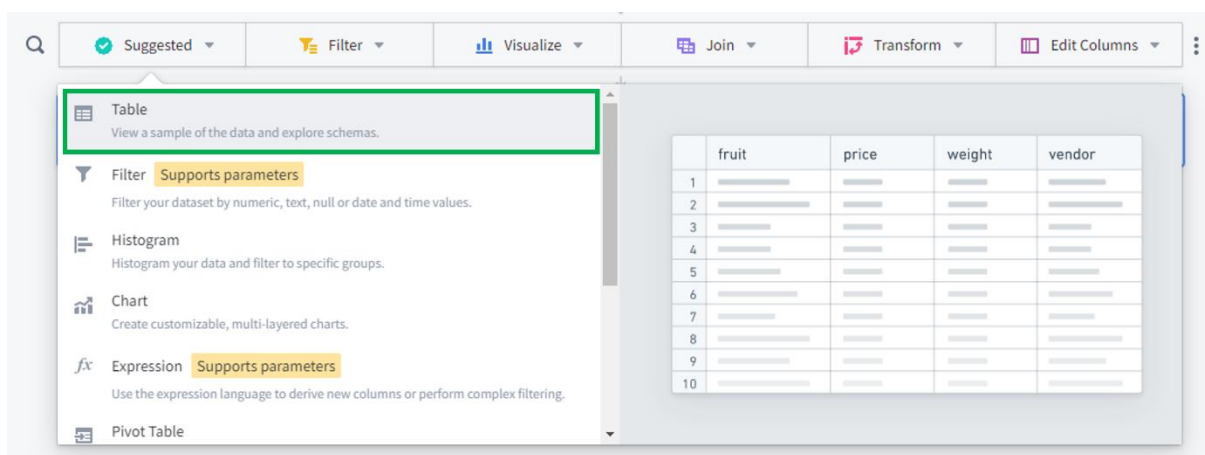
A few tips as you navigate Contour:

- At smaller screen resolutions, you can temporarily zoom out in your web browser to navigate these categories for the steps below. You can also collapse Foundry's gray navigation bar on the left to get more screen space. The category names will display when enough space is available.
- Your default view of the board toolbar may be set to an **Actions** view, which will show a list of many boards. For this training, set your board toolbar to **Display as categories** by selecting the ... icon above the the Board Toolbar.

Insert a Table Board

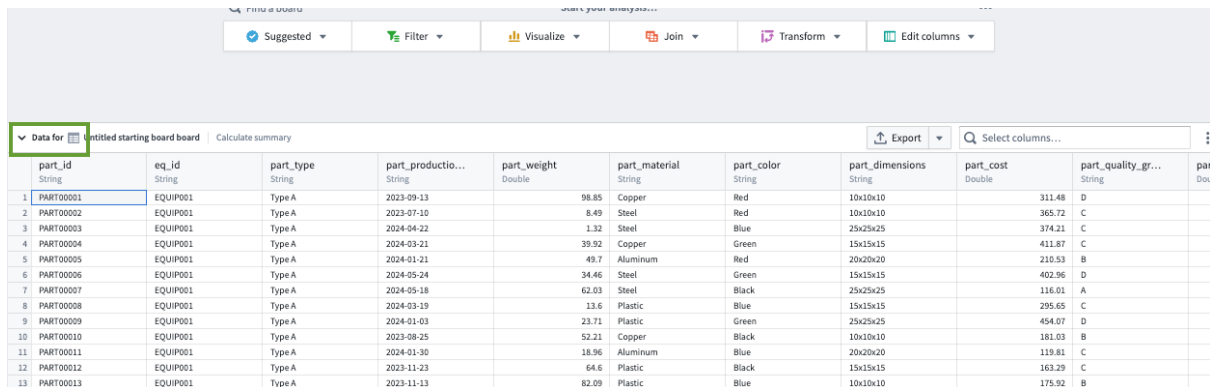
Step 1: Add a Table board

1. In the board toolbar, click **Suggested > Table**



The added Table board provides a view of the columns, their data types, and a preview of 1,000 rows.

You can also see a view of the dataset by clicking **Show data** at the bottom left of the input dataset board or any later board in your analysis path. This will present a preview of data as computed at the selected board, which is useful to quickly double check your logic at any point of the analysis.



	part_id String	eq_id String	part_type String	part_production_date String	part_weight Double	part_material String	part_color String	part_dimensions String	part_cost Double	part_quality_grade String	part_status String
1	PART00001	EQUIP001	Type A	2023-09-13	98.85	Copper	Red	10x10x10	311.48	D	
2	PART00002	EQUIP001	Type A	2023-07-10	8.49	Steel	Red	10x10x10	365.72	C	
3	PART00003	EQUIP001	Type A	2024-04-22	1.32	Steel	Blue	25x25x25	374.21	C	
4	PART00004	EQUIP001	Type A	2024-03-21	39.92	Copper	Green	15x15x15	411.87	C	
5	PART00005	EQUIP001	Type A	2024-01-21	49.7	Aluminum	Red	20x20x20	210.53	B	
6	PART00006	EQUIP001	Type A	2024-05-24	34.46	Steel	Green	15x15x15	402.96	D	
7	PART00007	EQUIP001	Type A	2024-05-18	62.03	Steel	Black	25x25x25	116.01	A	
8	PART00008	EQUIP001	Type A	2024-03-19	13.6	Plastic	Blue	15x15x15	295.65	C	
9	PART00009	EQUIP001	Type A	2024-01-03	23.71	Plastic	Green	25x25x25	454.07	D	
10	PART00010	EQUIP001	Type A	2023-08-25	52.21	Copper	Black	10x10x10	181.03	B	
11	PART00011	EQUIP001	Type A	2024-01-30	18.96	Aluminum	Blue	20x20x20	119.81	C	
12	PART00012	EQUIP001	Type A	2023-11-23	64.6	Plastic	Black	15x15x15	163.29	C	
13	PART00013	EQUIP001	Type A	2023-11-13	82.09	Plastic	Blue	10x10x10	175.92	B	

Generate Summary Statistics

Step 1: Add a Calculation board of unique parts

1. In the board toolbar, click **Visualize > Calculation**
2. Calculate the number of unique parts
 1. Column: **part_id**
 2. Aggregate: **Unique count**

Step 2: Add a calculation of unique equipment to the board

1. Scroll to the top of the **Data** panel
2. Click **Add calculation**
3. Calculate the number of unique equipment IDs
 1. Column: **eq_id**
 2. Aggregate: **Unique count**
4. Click **Compute**

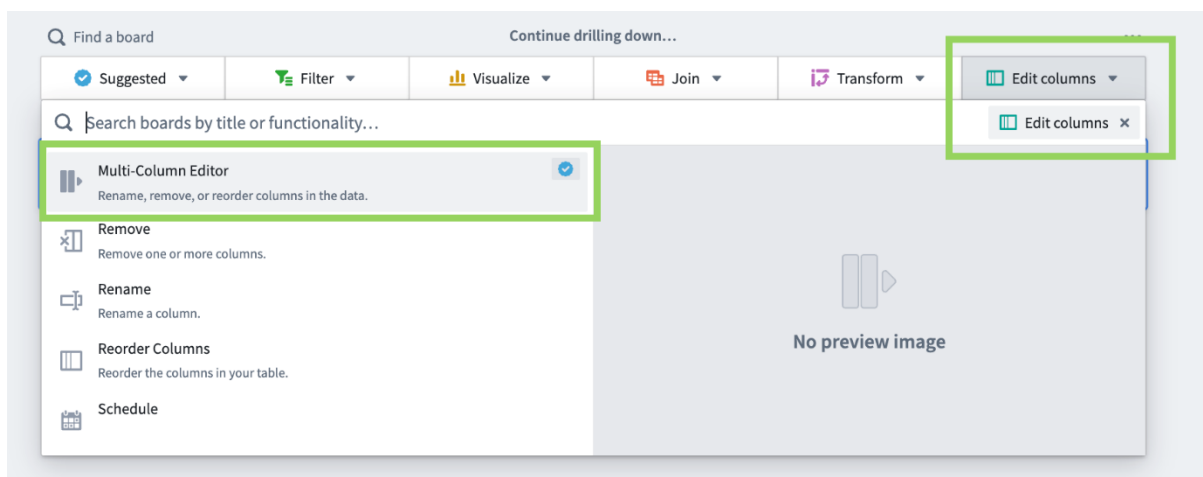


The Calculation board allows you to compute summary statistics and understand the shape of the dataset. This is useful to validate your logic when performing joins, window functions, or complex data transformations. In this case, we see that there 3,428 unique parts and 5 unique equipment. We can compare to these counts as we progress through the analysis.

Rename and Drop Columns

Step 1: Add a Multi-Column Editor board

1. In the board toolbar, **Edit columns > Multi-Column Editor**



Step 2: Rename eq_id to equipment_id

1. In the right **Kept columns** panel, click on the eq_id column
2. Rename the column to: equipment_id
 - Press the Return / Enter key to save you changes. Note, your changes will not be made to the column if you click out of the text field without tapping the Return / Enter key.

After your change, the left **All Columns** panel for should likewise update to equipment_id with a note that the old name was eq_id. In future steps, we'll use this column to join to the equipment dataset. Consistent naming conventions for columns help maintain readability for your analyses and pipelines.

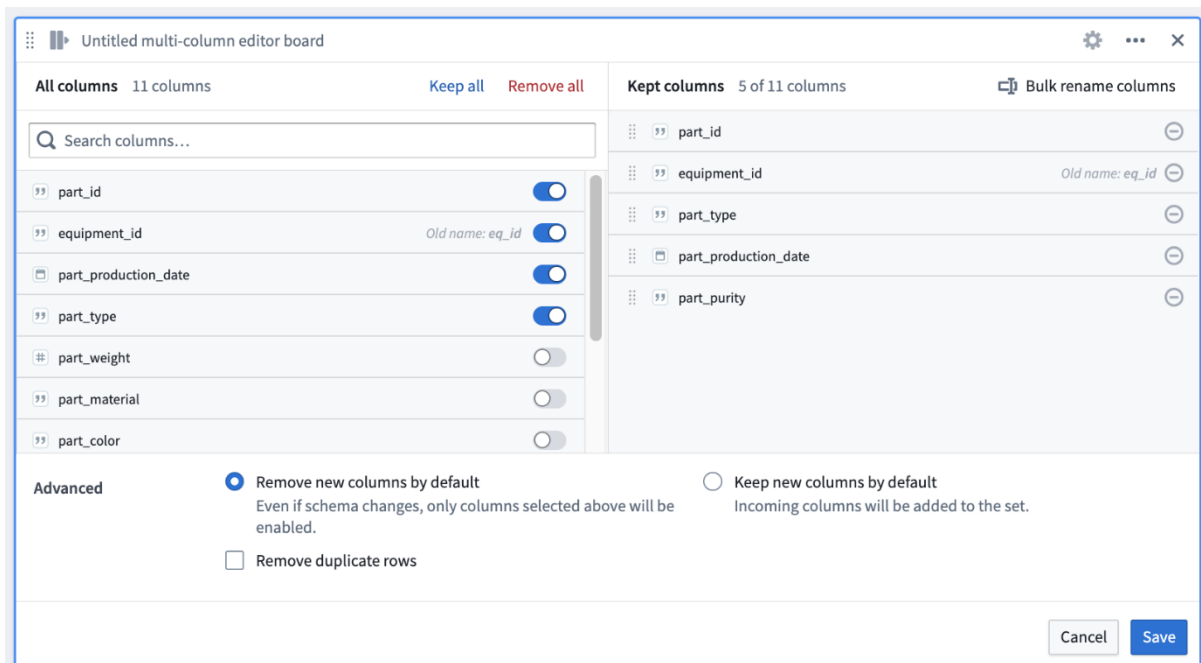
Step 3: Drop some columns

1. In **All Columns**, toggle off these columns:

1. part_weight
2. part_material
3. part_color
4. part_dimensions
5. part_cost
6. part_quality_grade

2. Click **Apply**

This will remove these columns from your dataset in any later parts of this analysis, but it has no effect above this board or on the original Foundry dataset.



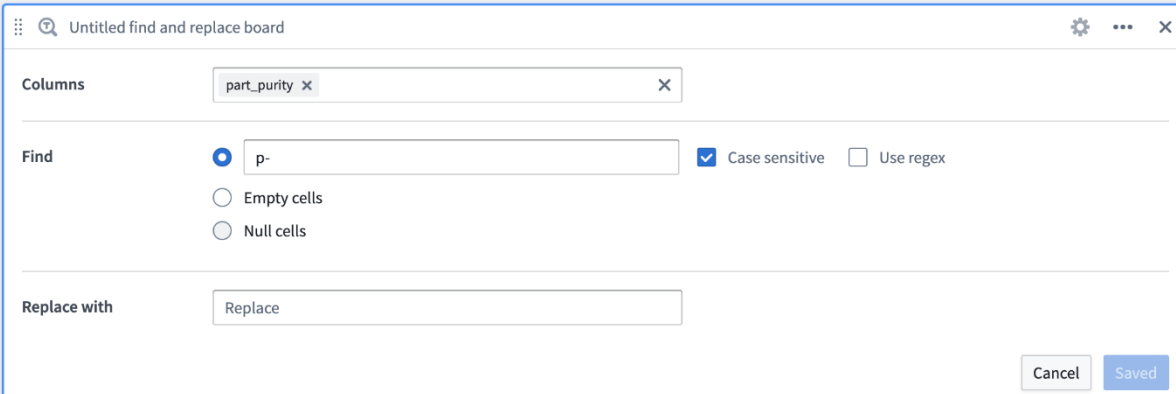
When using Multi-Column Editor in paths that follow joins or expression boards, you can toggle on **Keep new columns by default** to ensure that any additional columns are automatically included in subsequent boards.

Find and Replace Column Values

When performing analysis, it is often essential to update column values to ensure operations can be performed efficiently and consistently across multiple datasets. If you view the underlying data using the Table or **Show data**, you'll see that the values in the parts_purity column are all preceded by a p- string. Over the next two lessons, you will remove this prefix and then convert the remaining value into a Double data type. This is a numerical data type suitable for numerical operations in future steps.

Step 1: Add a Find and Replace board

1. In the board toolbar, **Transforms > Find and Replace**
 - Alternatively, if you don't know what category a board is in, you can use the **Find a board** option in the top left of the board toolbar to search for it.
2. Remove the p- prefix from part_purity
 1. Columns: part_purity
 2. Find: p-
 3. Replace: *(leave this empty)*
3. Click **Apply**



The screenshot shows a window titled 'Untitled find and replace board'. It has a search icon and a close button in the top right. The 'Columns' section has a dropdown menu with 'part_purity' selected. The 'Find' section has a radio button selected for 'p-' in the search field, with options for 'Empty cells' and 'Null cells' below it. There are checkboxes for 'Case sensitive' (checked) and 'Use regex' (unchecked). The 'Replace with' section has a text field with the word 'Replace' inside. At the bottom right, there are 'Cancel' and 'Saved' buttons.

This step will remove the letter and dash without replacing them, leaving only the numerical values. In other analyses, you may also use this board to remove null values or replace values in order to effectively join to other datasets.

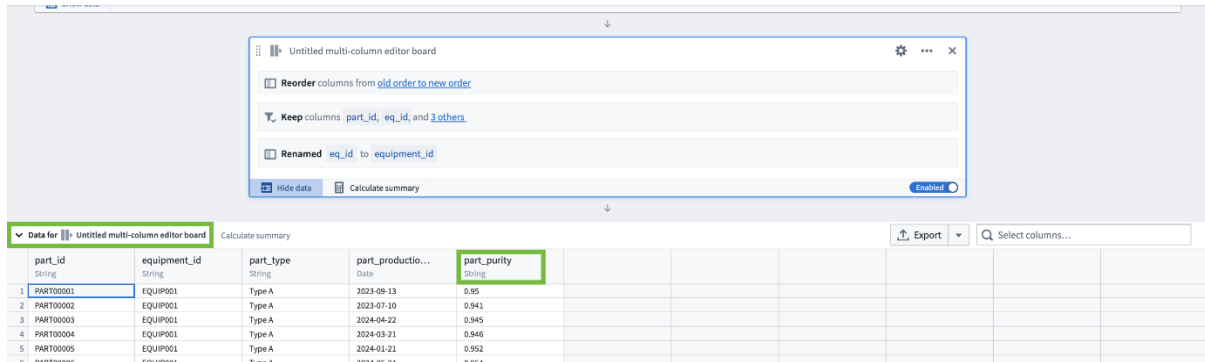
Step 2: Confirm the data change

Before performing further steps in our analysis, let's confirm that last change worked.

1. Select the **Find and Replace** board so that it is surrounded by a blue highlight

2. Expand **Show Data** drawer at the bottom of the screen.

- Alternatively, if you close the board's configuration, you can instead click **Show data** in the lower left corner of the board



Looking at the `part_purity` column, you can confirm that you've successfully removed the 'p-' from all values. However, under the column name, you can also see that `part_purity` is of String data type. In the next steps, we'll convert the column from String to Double so that we may perform mathematical operations on its values

While you are viewing the data, you can also confirm that the name of the `equipment_id` column changed thanks your earlier board.

Convert Data Types

Converting data types or casting columns is crucial in data processing as it ensures that the data is correctly interpreted and manipulated according, enabling accurate analytics, comparisons, and computations.

Step 1: Add a Convert Types board

1. Collapse the data view if it is still open
2. In the board toolbar, **Transform > Convert types**
3. Convert `part_purity` to Double
 1. Type conversions: **Select column > part_purity**
 2. On the right, change **String** to **Double**
4. Click **Apply**

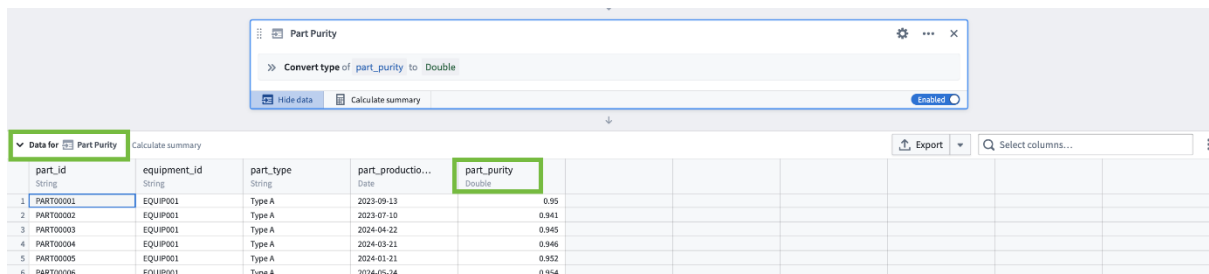
Step 2: Rename the board to Parts Purity

Sometimes it can help your rename your boards as kind of documentation to note what you're doing.

1. Click on the title of this board: "**Untitled convert types board**"
2. Enter a new title: Part Purity
3. Hit the Return / Enter key or click out of the text field to save the change

Step 3: Confirm the data

1. Click **Show data** in the bottom left corner of the board
2. Review the columns to validate that part_purity is now Double, not String
3. Hide the data view again



The screenshot shows a data board titled "Part Purity". At the top, there is a toolbar with a settings icon, a close icon, and a text field containing "Convert type of part_purity to Double". Below this, there are buttons for "Hide data" and "Calculate summary". The main area of the board displays a data table. The table has five columns: "part_id" (String), "equipment_id" (String), "part_type" (String), "part_production_date" (Date), and "part_purity" (Double). The "part_purity" column is highlighted with a green box. The table contains six rows of data.

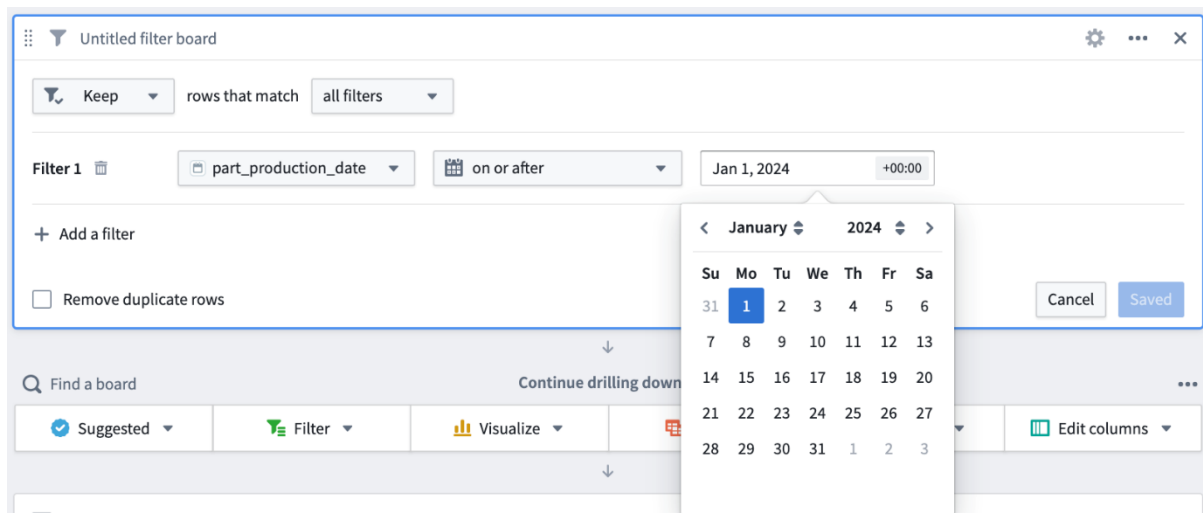
part_id	equipment_id	part_type	part_production_date	part_purity
1 PART00001	EQUIP001	Type A	2023-09-13	0.95
2 PART00002	EQUIP001	Type A	2023-07-10	0.941
3 PART00003	EQUIP001	Type A	2024-04-22	0.945
4 PART00004	EQUIP001	Type A	2024-03-21	0.946
5 PART00005	EQUIP001	Type A	2024-01-21	0.952
6 PART00006	EQUIP001	Type A	2024-05-24	0.954

Filtering Data

In our analyses, we often need to filter the data to rows that meet a specific criteria, resulting in a specific subset of all the data we have. For this analysis, we're going to exclude historical data to ensure we're using only recently manufactured parts as a basis for our decision-making.

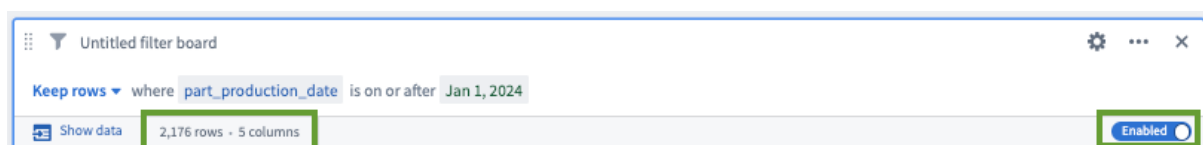
Step 1: Add a Filter board

1. In the board toolbar, **Filter > Filter**
2. Set Filter 1 to keep only production dates in 2024
 1. Select column: part_production_date
 2. Comparison: **on or after**
 3. Date: Select **Jan 1, 2024**
3. Click **Apply**



Step 2: Confirm the data change

1. Click **Calculate summary** in the lower left of your new filter board
 - After calculation, should see a row count of 2,176 rows. All subsequent boards will use this filtered data.
 - Sometimes you want to disable a board temporarily so that it does not affect the flow of data through the path.
2. Toggle **Enabled** to **Disabled** in the lower right of the board.
 - This will leave the board visible in your path, but the filter won't be applied to the data.
3. Click **Calculate summary** again.
1. The row count should be 3,428 rows again.
4. Toggle **Disabled** to **Enabled**



Contour's flexible filtering configuration can be used to combine multiple *any* or *all* conditions to filter a range of data types and columns. You can keep or remove rows and filter on multiple conditions and wildcards. The data resulting from a filter board will be used in all subsequent boards. However, you may enable or disable any board as you proceed with your analysis to review filtering impact. It's good practice to confirm the row count and validate your analysis as you go.

You should now have a path with six boards, resembling the screenshot below.

The screenshot displays a data analysis interface with the following components:

- Table:** A table with 11 columns and 4 rows of data. The columns are: part_id, equip_id, production_date, type, weight, material, color, and part_name. The rows show parts with IDs 14, 15, 16, and 17.
- Calculation Board:** Titled "Untitled calculation board", it contains two visualizations: "Unique count of part_id" with a value of 3,428 and "Unique count of eq_id" with a value of 5. A "Compute" button is at the bottom right.
- Multi-column Editor Board:** Titled "Untitled multi-column editor board", it shows actions: "Remove columns part_weight, part_material, and 4 others" and "Renamed eq_id to equipment_id".
- Find and Replace Board:** Titled "Untitled find and replace board", it shows a search for "p-" in "part_purity" with the option to replace with an empty string.
- Part Purity Board:** Titled "Part Purity", it shows a conversion action: "Convert type of part_purity to Double".
- Filter Board:** Titled "Untitled filter board", it shows a filter: "Keep rows where part_production_date is on or after Jan 1, 2024".

The bottom status bar indicates "Data for Untitled starting board board | 3,428 rows • 11 columns".

Summary

So far, we've reviewed datasets using the **Table** board and the bottom **Data** drawer to inspect columns and a data sample. We employed the **Calculation** board to review summary statistics. These are easy steps to help us understand the data we're handling.

We then removed unnecessary columns for our analysis, improved readability by renaming a column, and converted a column from a string data type to a double type to make our analysis more manageable. We've also applied a filter directly on a single column to limit the

scope of our analysis.

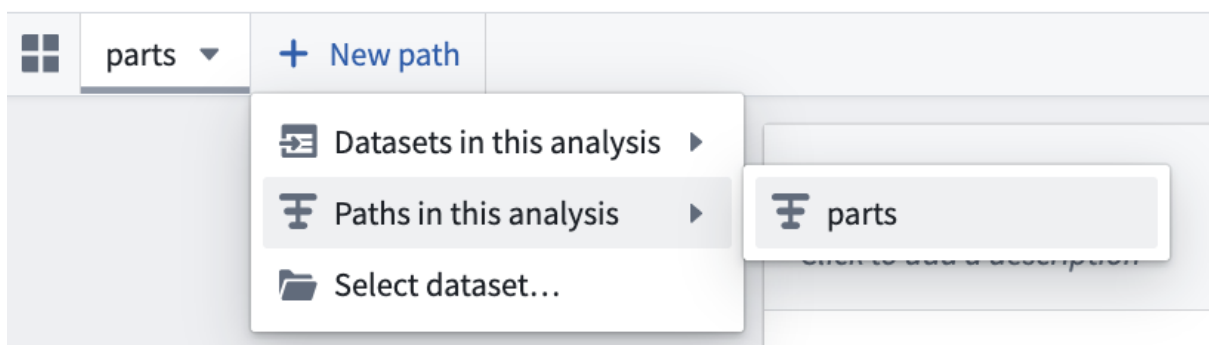
We've now applied all of the necessary data cleaning steps for the parts dataset in this path. As we'll see in the next section, keeping data cleaning steps for each dataset all within dataset-specific paths allows our analysis to be more easily managed and interpretable, and it limits potential redundant operations. Next, we'll join to the cleaned equipment dataset and generate additional columns to aid in our decision-making.

Joining Datasets

Add a New Path to the Analysis

Step 1: Add a new path

1. Click on **New path** at the top of the analysis
2. Select **Paths in this analysis > parts**

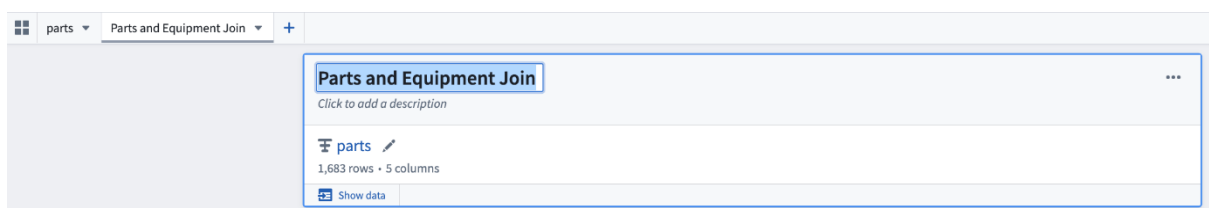


In the new tab, the first input board will display **parts (resulting set)** as the title. This means the output of the **parts** path will flow as input into this new path.

Step 2: Rename the path

1. Double-click on the title of the input board
2. Rename it: Parts and Equipment Join

After the rename, you'll notice the title of the tab at the top of the analysis has also updated to reflect the new name.



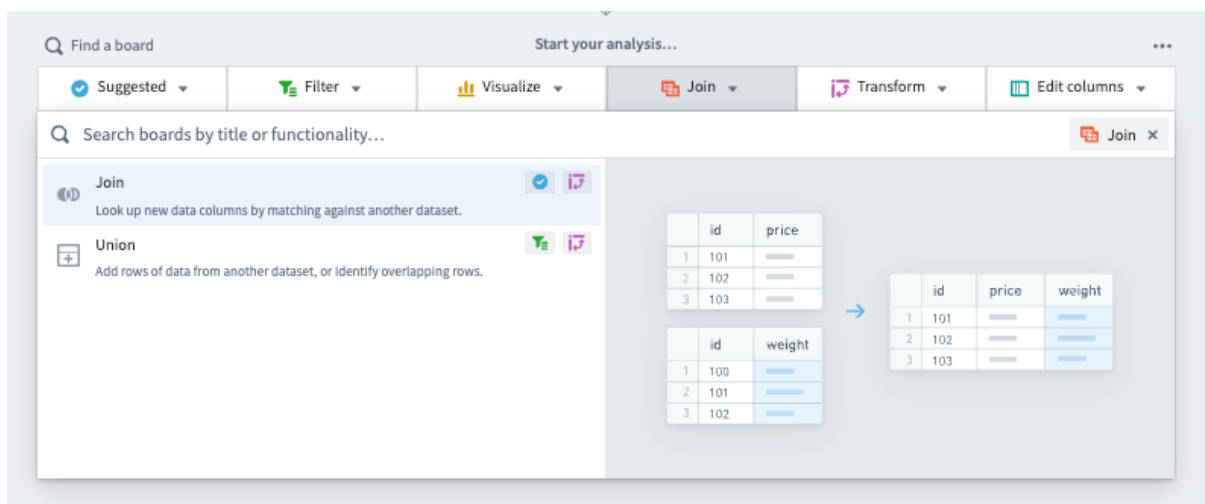
When creating a new path, Contour allows you to select any dataset on your Foundry stack, any dataset already referenced in your Contour analysis, or the result of any previous path to continue building your analysis.

To ensure a well-organized and easily comprehensible analysis, it's recommended to manage the cleaning of disparate datasets in distinct paths. For instance, you may want to join the cleaned parts data to an additional finance dataset to allow for further maintenance prioritization. For optimal performance and clarity in your analysis, it is advisable to confine your work to no more than five paths, each with a maximum of 20 boards.

Joining Datasets

Step 1: Add a Join board

1. In the board toolbar, **Join > Join**

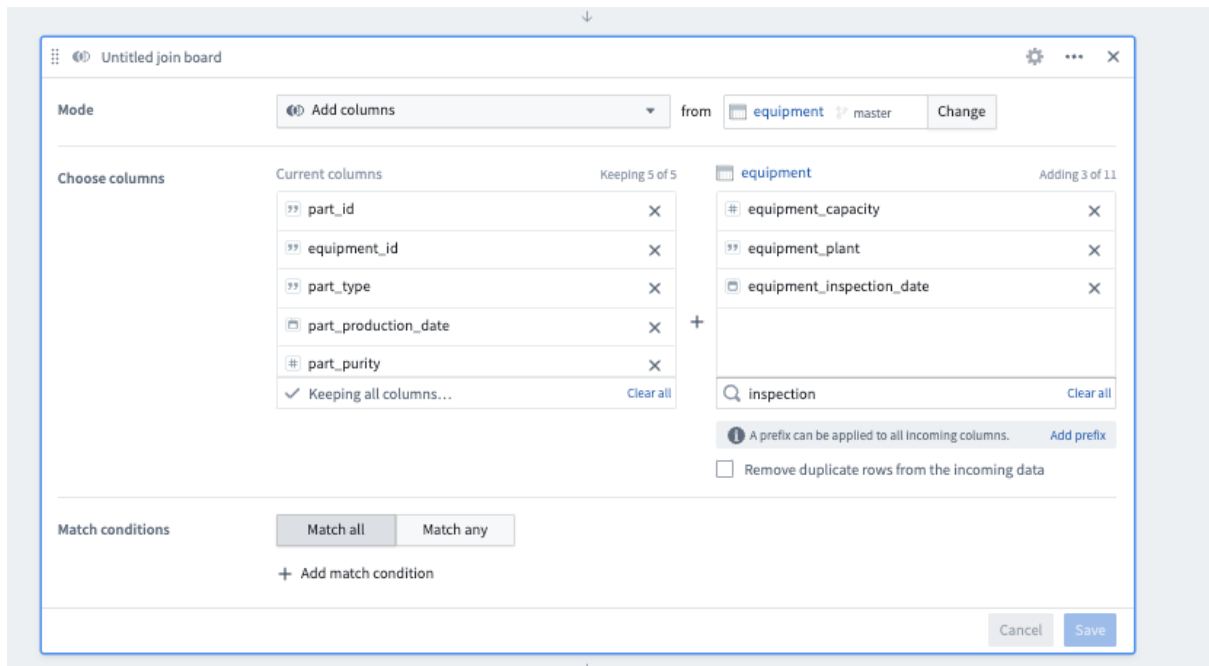


Step 2: Add columns from the equipment dataset

1. Leave the **Mode** set to **Add columns**
2. Click **Choose dataset... > Select dataset**
3. Search for and select the equipment dataset that you previously uploaded to your course folder
4. In the **equipment** field on the right side of the window, select **Add columns** to add
 1. equipment_capacity
 2. equipment_plant

3. equipment_inspection_date

- Warning: There is both an *installation* and an *inspection* date.



We've selected only the columns relevant to our analysis here, but Contour gives you the option to select all columns and to prefix all incoming columns with a unique identifier.

Step 3: Define the match condition

1. In the **Match condition** section, leave **Match all** selected
2. Click **Add match condition**
3. In the left dropdown: select: `equipment_id`
4. In the right dropdown, select: `equipment_id`.
5. Click **Apply**



If you now **Show data**, you should see the 3 new columns there.

Contour equips you with the capability to join and union datasets, with the option to match one or more columns (typically IDs) across the chosen tables. You've joined the equipment dataset to the cleaned parts dataset.

In this case, all equipment have the `equipment_active` column set as `true`. You can imagine if we had to filter out inactive equipment or perform other equipment data-cleaning steps, these would be best performed first in a distinct path for the equipment dataset.

Creating Derived Columns

Calculate Average Purity

Now that we can see which equipment produced each part, it might be useful to know whether some equipment produces parts with higher purity than other equipment. In this step, you'll use Contour's Expression board to create a new column for part purity via a code snippet.

Step 1: Add an Expression board

1. In the board toolbar, **Transform > Expression**
2. Click the **Editor** tab

Step 2: Add an avg_purity column

1. Leave **Add new column** selected in the top bar
2. In the **New column name** field, enter: `avg_purity`
3. In the subsequent code text box, input the code below to calculate the average purity by equipment:

```
format_number(avg("part_purity") OVER (  
PARTITION BY "equipment_id"  
) , 2)
```

4. Click **Apply**

Add new column	Replace column	Filter	Aggregate
avg_purity			
<pre>1 format_number(avg("part_purity") OVER (2 PARTITION BY "equipment_id" 3), 2)</pre>			

This code partitions the data by `equipment_id` and then computes the average `part_purity` for each of those partitions. It then rounds each average purity result to 2 decimal places.

Contour's **Expression** board lets you use an expression language to update existing columns, derive new columns, perform sophisticated filtering, use CAST and CASE statements, window functions, obtain complex aggregations, and other functions. With an Expression board, you can perform SQL-like operations on refreshed data, without having to re-run queries and then transferring data extracts to Excel for further analysis each time a report is due.

When editing in the Expression board, click the **Help** icon for a quick reference of the expression syntax. Click the **Documentation** link for the full list of supported functions. As you type, suggested functions appear in a dropdown and suggested columns appear to help autofill your expressions. Click or use the Enter key to select desired functions.

You've now created a column to calculate the average `part_purity` per equipment since the start of 2024. You could alternatively evaluate the minimum or maximum value.

Following this, we'll implement simple case logic to assess whether a piece of equipment requires an inspection.

Generate Inspection Alerts

Step 1: Use another Expression board to generate inspection alerts

1. Add another **Expression** board (**Transform > Expression**)
2. Switch to the **Editor** tab
3. Select **Add new column**
4. Name the new column: `inspection_alert`
5. Paste in the following code block:

```

CASE
WHEN date_diff(current_date(), "equipment_inspection_date") > 365
AND "avg_purity" < .90 THEN 'high'
WHEN date_diff(current_date(), "equipment_inspection_date") > 365
AND "avg_purity" >= .90 AND "avg_purity" <.95 THEN 'medium'
WHEN date_diff(current_date(), "equipment_inspection_date") <= 365
AND "avg_purity" < .90 THEN 'medium'
ELSE 'low'
END

```

6. Click **Apply**

 The function `current_date` is only evaluated when the path is recalculated.

Add new column
Replace column
Filter
Aggregate

Documentation Help AIP Assist

inspection_alert

```

1 CASE
2 WHEN date_diff(current_date(), "equipment_inspection_date") > 365
3 AND "avg_purity" < .90 THEN 'high'
4 WHEN date_diff(current_date(), "equipment_inspection_date") > 365
5 AND "avg_purity" >= .90 AND "avg_purity" <.95 THEN 'medium'
6 WHEN date_diff(current_date(), "equipment_inspection_date") <= 365
7 AND "avg_purity" < .90 THEN 'medium'
8 ELSE 'low'
9 END

```

This code computes three levels of inspection priority based on the equipment's average output purity and whether that equipment has been inspected within the past year.

You may see a pop-up warning that “The function `current_date` is only evaluated when the path is recalculated.” This flags that the current date value will only be calculated when the analysis is run. For a production use case, we’d likely then want this analysis to be running on a daily basis so the `current_date` value is accurate. We can ignore the warning for the purpose of this training.

Step 2: Confirm the data changes

1. Click **Show data** in the lower left corner of the Expression board you just added
2. Verify that the two columns that were added in the previous two steps are present: `inspection_alert` and `avg_purity`

	inspection_alert String	avg_purity String	part_id String	equipment_id String	part_type String	part_productio... Date	part_purity Double
1	low	0.95	PART00003	EQUIP001	Type A	2024-04-22	0.945
2	low	0.95	PART00004	EQUIP001	Type A	2024-03-21	0.946
3	low	0.95	PART00005	EQUIP001	Type A	2024-01-21	0.952
4	low	0.95	PART00006	EQUIP001	Type A	2024-05-24	0.954
5	low	0.95	PART00007	EQUIP001	Type A	2024-05-18	0.943
6	low	0.95	PART00008	EQUIP001	Type A	2024-03-19	0.947
7	low	0.95	PART00009	EQUIP001	Type A	2024-01-03	0.949
8	low	0.95	PART00011	EQUIP001	Type A	2024-01-30	0.956
9	low	0.95	PART00015	EQUIP001	Type A	2024-05-07	0.954
10	low	0.95	PART00022	EQUIP001	Type A	2024-05-03	0.959
11	low	0.95	PART00023	EQUIP001	Type A	2024-02-14	0.948
12	low	0.95	PART00024	EQUIP001	Type A	2024-01-18	0.95
13	low	0.95	PART00025	EQUIP001	Type A	2024-01-20	0.949
14	low	0.95	PART00027	EQUIP001	Type A	2024-02-12	0.942
15	low	0.95	PART00028	EQUIP001	Type A	2024-05-09	0.95
16	low	0.95	PART00029	EQUIP001	Type A	2024-04-08	0.959
17	low	0.95	PART00030	EQUIP001	Type A	2024-03-05	0.956
18	low	0.95	PART00033	EQUIP001	Type A	2024-04-21	0.943
19	low	0.95	PART00039	EQUIP001	Type A	2024-01-09	0.959
20	low	0.95	PART00040	EQUIP001	Type A	2024-01-08	0.949
21	low	0.95	PART00041	EQUIP001	Type A	2024-01-26	0.945
22	low	0.95	PART00042	EQUIP001	Type A	2024-06-02	0.948
23	low	0.95	PART00044	EQUIP001	Type A	2024-06-05	0.959
24	low	0.95	PART00046	EQUIP001	Type A	2024-04-12	0.947
25	low	0.95	PART00049	EQUIP001	Type A	2024-03-04	0.94
26	low	0.95	PART00050	EQUIP001	Type A	2024-05-15	0.942
27	low	0.95	PART00051	EQUIP001	Type A	2024-06-10	0.942
28	low	0.95	PART00055	EQUIP001	Type A	2024-03-08	0.951
29	low	0.95	PART00056	EQUIP001	Type A	2024-05-09	0.96

You've now derived an inspection priority flag generated based data from both the parts and equipment datasets in a unified path. Next, we'll create a parameter to better suit our dashboard for plant managers.

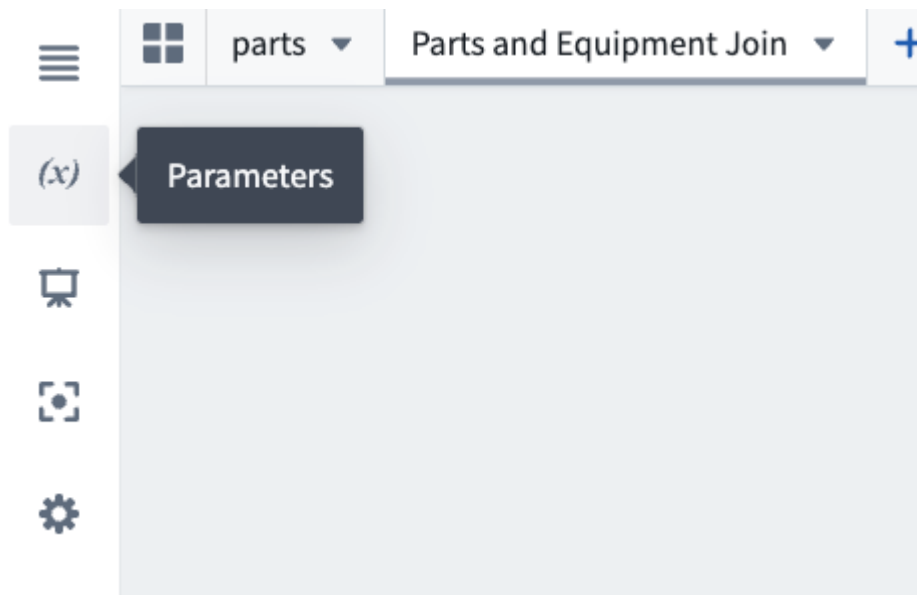
Parameters

Create a Parameter for Plant

Looking ahead, we may want the view of the data and the inspection alerts you've just created to vary based on inputs from the viewing user. For example, that user might only care about data associated with a given manufacturing plant. This can be done through the use of parameters.

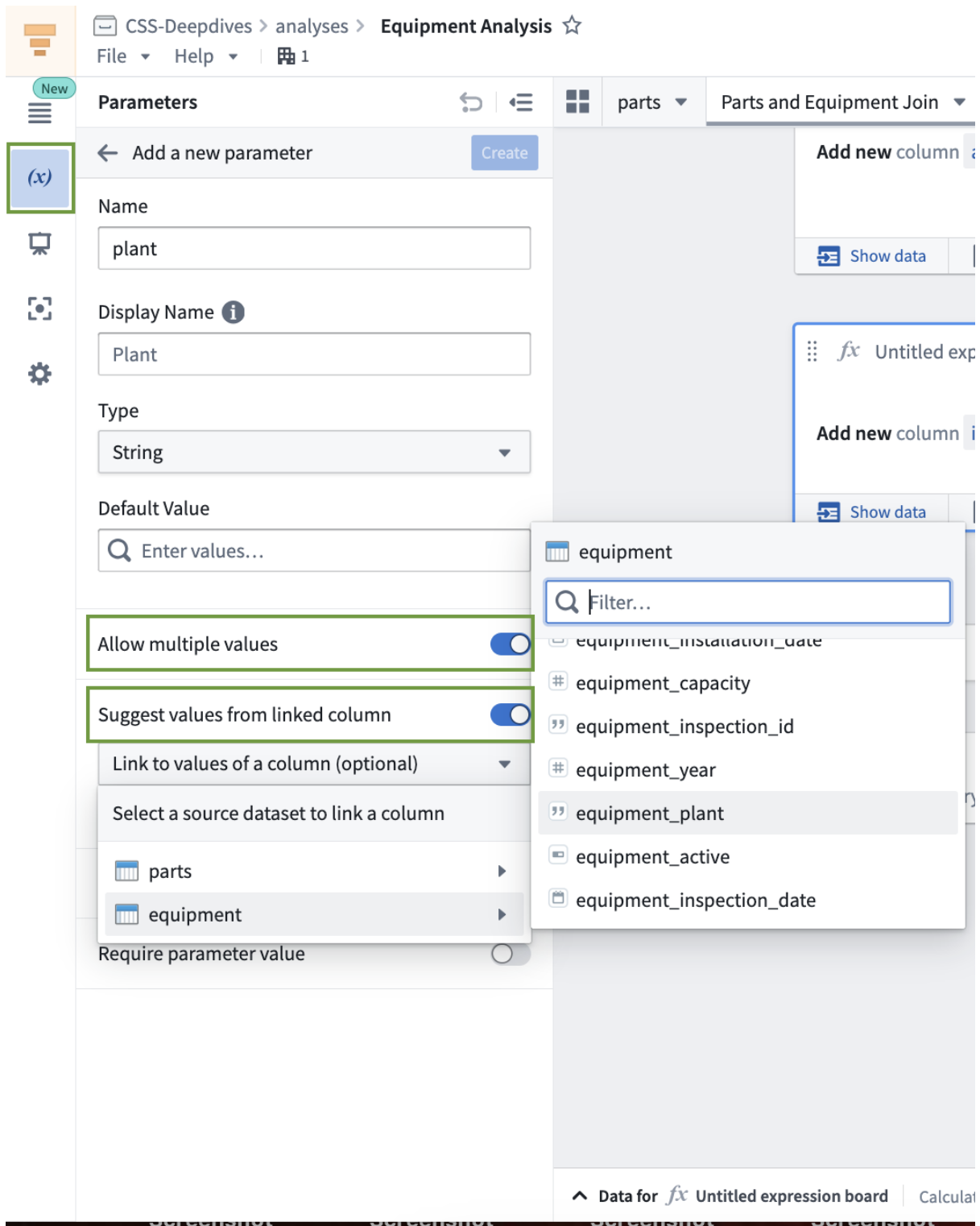
Step 1: Create a new parameter

1. Click on the **Parameters** icon (**x**) from the left-hand sidebar
2. Click the **+ Create parameter** button at the center of the left panel.



Step 2: Configure the new parameter

1. Name: plant
 - This will also set the Display Name
2. Toggle on **Allow multiple values**
3. Toggle on **Suggest values from linked column**
4. In the **Link to values of a column** dropdown, choose **equipment > equipment_plant**.
5. Click the blue **Create** button at the top right of the panel
6. Close the **Parameters** sidebar by clicking on the **(x)** tab again



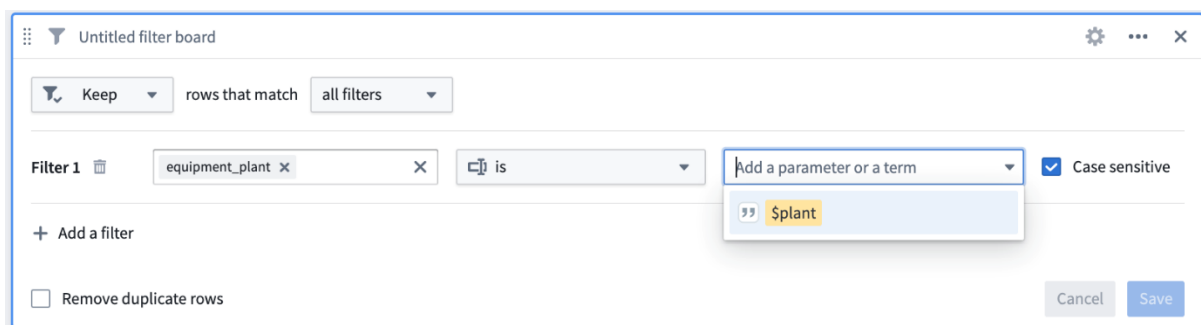
Apply Parameter as Filter

For the new parameter to affect your analysis, it must be used in a board somewhere. In this case, we want to filter down to the selected plants, so we can use a new Filter board to do that.

Step 1: Add a Filter based on the plant parameter

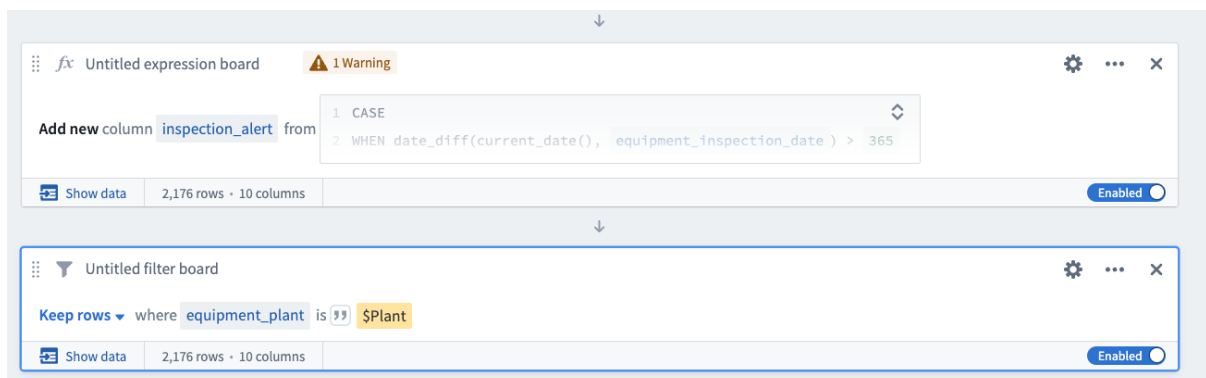
1. From the board toolbar, **Filter > Filter**
2. Define the filter
 1. Select column: equipment_plant.
 2. Leave the comparison set to: **is**
 3. Add a parameter or a term: \$plant
 - This should appear as a yellow option

3. Apply



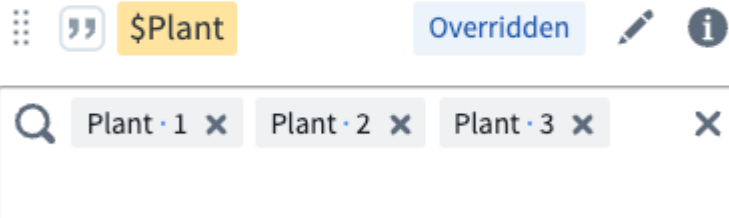
Step 2: Confirm the effects of parameter selection

1. Click **Calculate summary** in the lower left corner of the two most recent expression boards
 - Notice that the datasets are the same at this point. You have not defined a value for the \$plant parameter, so no filtering is being done.



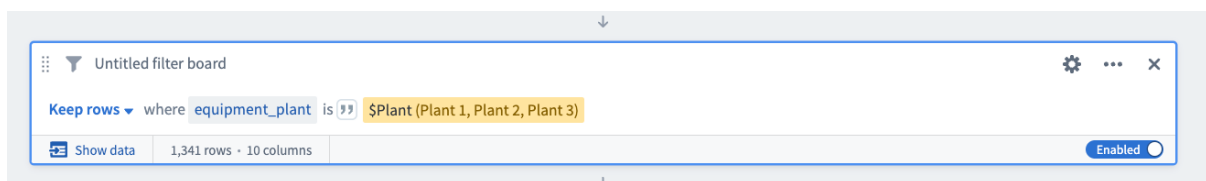
2. Open the left-side **Parameters** panel
3. Enter values for **\$plant**: Plant 1, Plant 2, and Plant 3
4. Click **Apply** at the bottom right of the Parameters panel

 Temporary override applied to 1 parameter



5. Now click **Calculate summary** on your latest filter board

- The row count should update to 1,341 as you are now filtering only to rows containing data for Plant 1, Plant 2, and Plant 3.



6. Clear the parameter

1. Back in the left-side **Parameters** panel, clear the selection for Plant 1, Plant 2, and Plant 3
2. Click **Apply**
3. Close the **Parameters** panel

In analyses, you may input a specific value(s) into the left parameter panel to filter consistently across multiple paths where the parameter is used, whereby reducing redundancy and chance for human error.

Parameters are also accessible in dashboards to allow multiple users at a time to select and interact with different values per their particular needs. In the context of our analysis, we've created a parameter that will enable users, like plant managers, to efficiently filter equipment based on their plants of interest. In future steps, we'll create a dashboard where users can apply this filter.

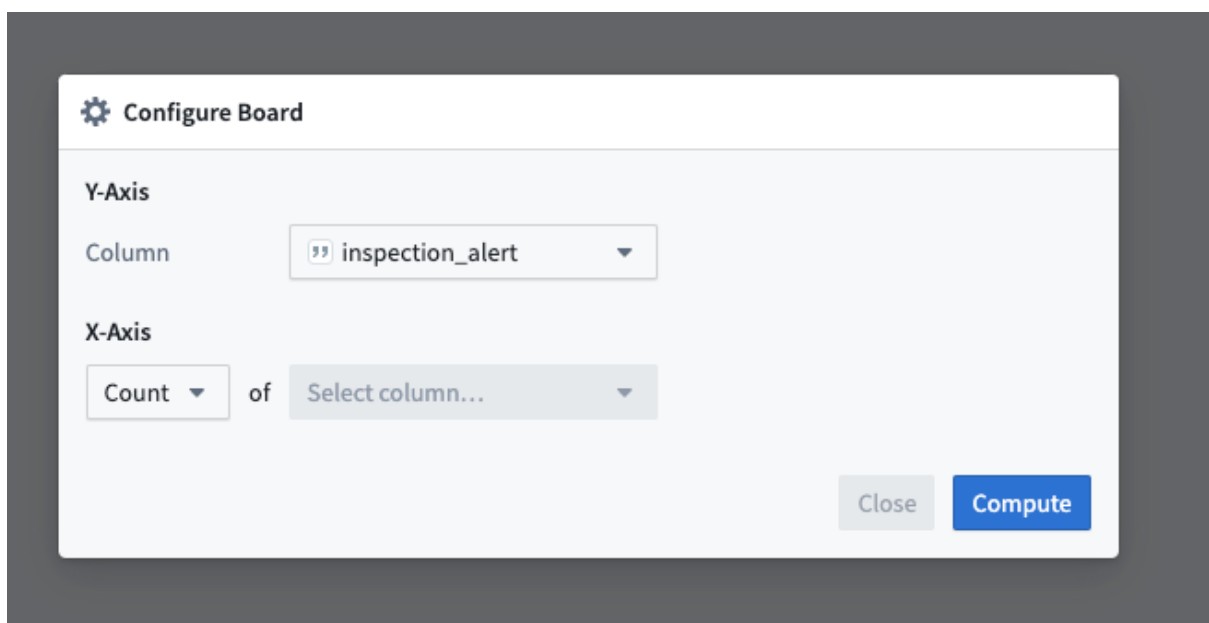
Create Visualizations

Create Histogram for Inspection Alerts

In this section, we'll explore how patterns and trends in the data can be made more apparent using visualizations.

Step 1: Add a Histogram of inspection alerts

1. In the board toolbar, **Visualize > Histogram**
2. Configure the board to show a count of inspection alerts
 1. Under Y-Axis, Column: inspection_alert
 2. Under X-Axis, select **Count**
3. Click **Compute**



You can modify the X-Axis options to perform different operations on your columns. For this analysis, however, we'll leave **Count** as the selected option. This will tally the number of rows (parts) associated with equipment bucketed into low, medium, and high inspection alerts.

Step 2: Use the Histogram as a filter

1. Click on the bar representing **high** values

A grey bar at the bottom of the board will appear displaying, "Keep rows where inspection_alert is high." This will filter the data in all subsequent boards.



Histogram boards are one of the most commonly used boards in Contour. Users will often implement histogram and other visualization boards to perform filtering steps in lieu of filter boards because they offer an easy way to both view your data and apply filtering functionality simultaneously.

Generate a Pivot Table to Identify High-Capacity Plants

Another useful view would be a summary of which equipment in which plants are producing the greatest variety of parts. One way to display this is as a pivot table.

Step 1: Add a Pivot table

1. In the board toolbar, **Transform > Pivot Table**
2. Configure the pivot table
 1. Rows: Add equipment_plant and equipment_id
 2. Columns: Leave empty
 - While you have the option to select a column for a deeper analysis of each aggregated row, we'll leave it blank for this simple analysis.
 3. Aggregates: Select part_id and ensure that **Unique count** is selected.
3. Click **Compute**

Untitled pivot table board

Displaying 3 rows

equipment_plant	equipment_id	count_distinct_of_part
Plant 2	EQUIP002	500
Plant 3	EQUIP003	442
Plant 5	EQUIP005	389

Reorder columns Format cells

ROWS

Display row subtotals

Press enter to add...

equipment_plant

equipment_id

COLUMNS

Press enter to add...

AGGREGATES

Press enter to add...

part_id Unique count

Display results as: Compute

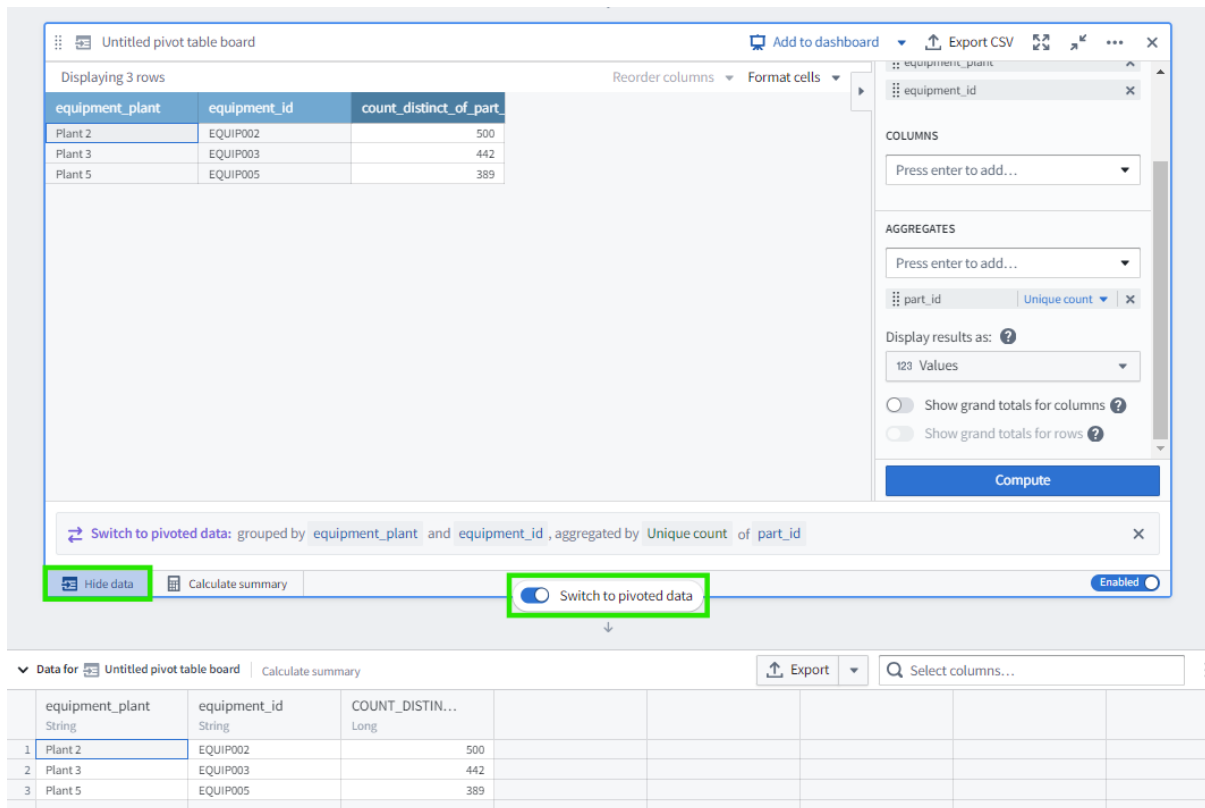
Show data Switch to pivoted data

This data contains only five plants and equipment. Currently you see only three in this pivot table because of the filter applied in the histogram board above to only equipment with a high inspection alert level. Given the small data range, it's easy to pick out which equipment to perform inspections on. But for users like plant managers that may need to allocate resource while wrangling many thousands of rows, this repeatable drill-down to the aggregated pivot table serves as a quick tool for analysis.

Step 2: Toggle pivoted data on and off

In addition to data visualization, a pivot table can be used to transform the data for later boards.

1. Click **Show data**
2. Toggle on **Switch to pivoted data**
3. Toggle off **Switch to pivoted data**
4. Click **Hide data**



Other boards, such as this Histogram board, will also let you pivot to aggregated data like this.

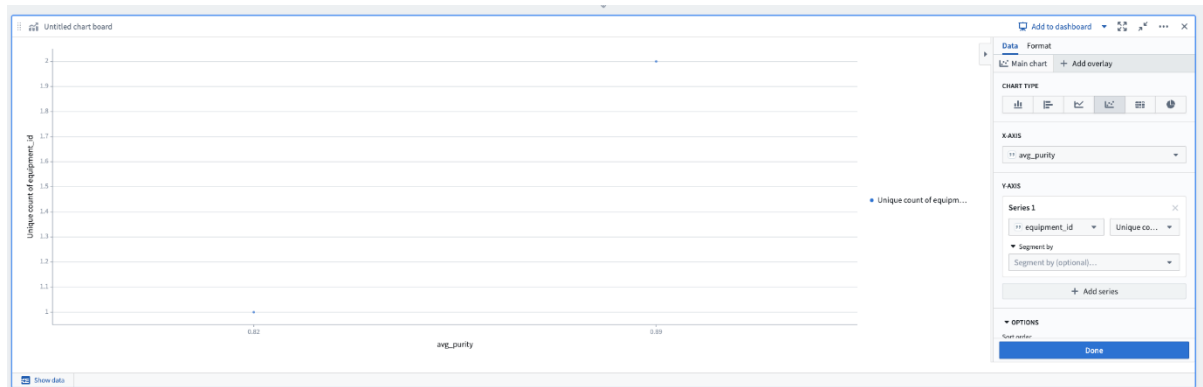
Create Chart for Average Purity by Count of Parts

In this lesson, you will build a visualization of average purity across different equipment.

Step 1: Add a Chart of equipment count by average purity

1. In the board toolbar, **Visualize > Chart**
2. Configure the chart
 1. Chart Type: **Scatter plot**
 - If you hover your cursor over each option, the tooltip will tell you the title of each one.
 - Note that the same Chart board can produce many different kinds of charts, such as bar, line, scatter, heat grid and pie charts.
 2. X-Axis: avg_purity.
 3. Y-Axis: Series 1: Select equipment_id and **Unique count**
 4. Options: Set **Sort order** to **X axis ascending**

1. Again, you can hover over each option to reveal its title.
3. Click **Compute and save**

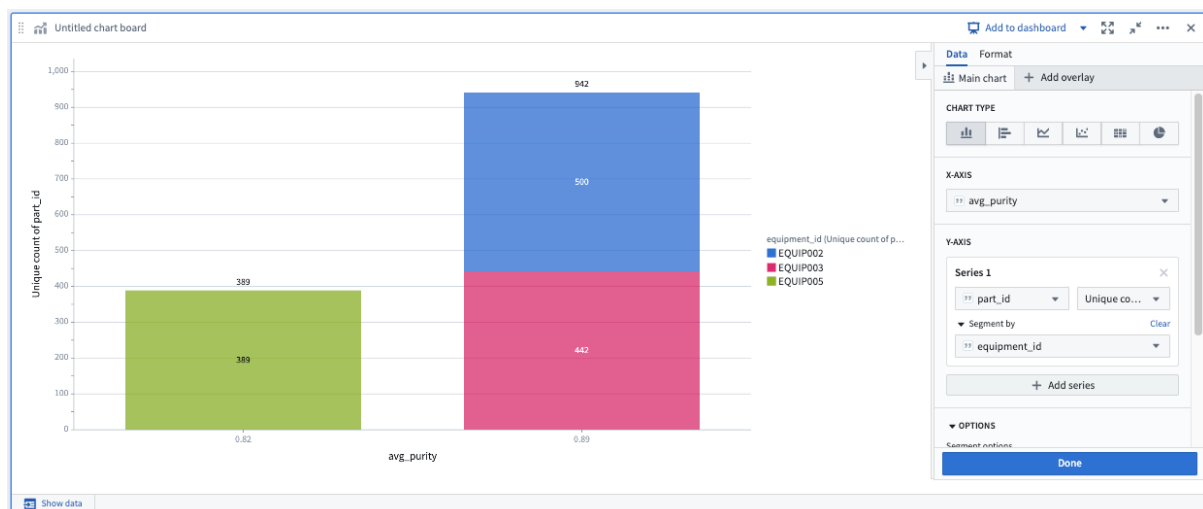


While this chart tells us the count of equipment producing at each purity level, we want to try a different visualization to also see a count of part types produced. We'll switch to the bar chart in the steps below.

Step 2: Reconfigure the chart

In the configuration of your existing chart:

1. Chart type: **Vertical bar chart**
2. X-Axis: avg_purity
3. Y-Axis: Series 1: Select part_id and **Unique count**
4. Segment by: equipment_id
5. Options: Segment options: **Stacked**
6. Click **Compute & save**



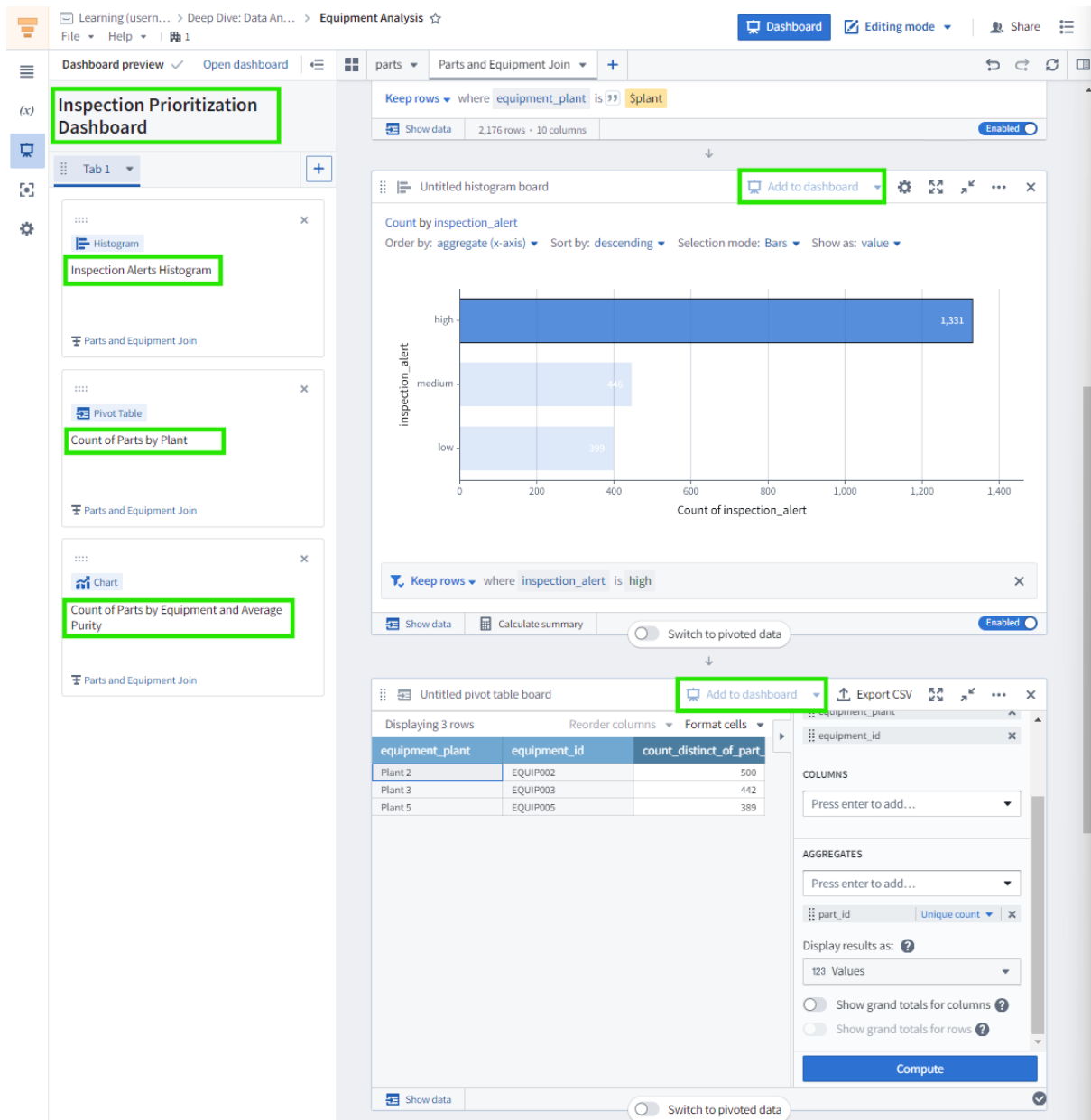
Create and Export a Dashboard

Create a Dashboard

Often you may find that you want to share the results of your data analysis with decision makers or other users. In this case, you will typically want to point them to only a subset of the boards in your analysis and potentially provide additional contextual notes or information. Building a Contour dashboard lets you do this.

Step 1: Create a dashboard of your analysis

1. Scroll to your Histogram board
2. Click **Add to dashboard** at the top of that board
 - The dashboard preview panel will appear on the left side of your screen and a blue *Dashboard* button will appear at the top.
3. Title this board: Inspection Alerts Histogram
 - The box in the center of the Histogram card in the dashboard preview allows you to title this board within the dashboard.
4. Add a title to your dashboard
 1. Click into the text box at the top of the dashboard preview, which says *Untitled dashboard*
 2. Enter: Inspection Prioritization Dashboard
5. Add your pivot table to the dashboard
 1. Scroll to your pivot table board
 2. Click **Add to dashboard**
 3. Title the board: Count of Parts by Plant
6. Add your bar chart to the dashboard
 1. Scroll to your pivot table board
 2. Click **Add to dashboard**
 3. Title the board: Count of Parts by Equipment and Average Purity

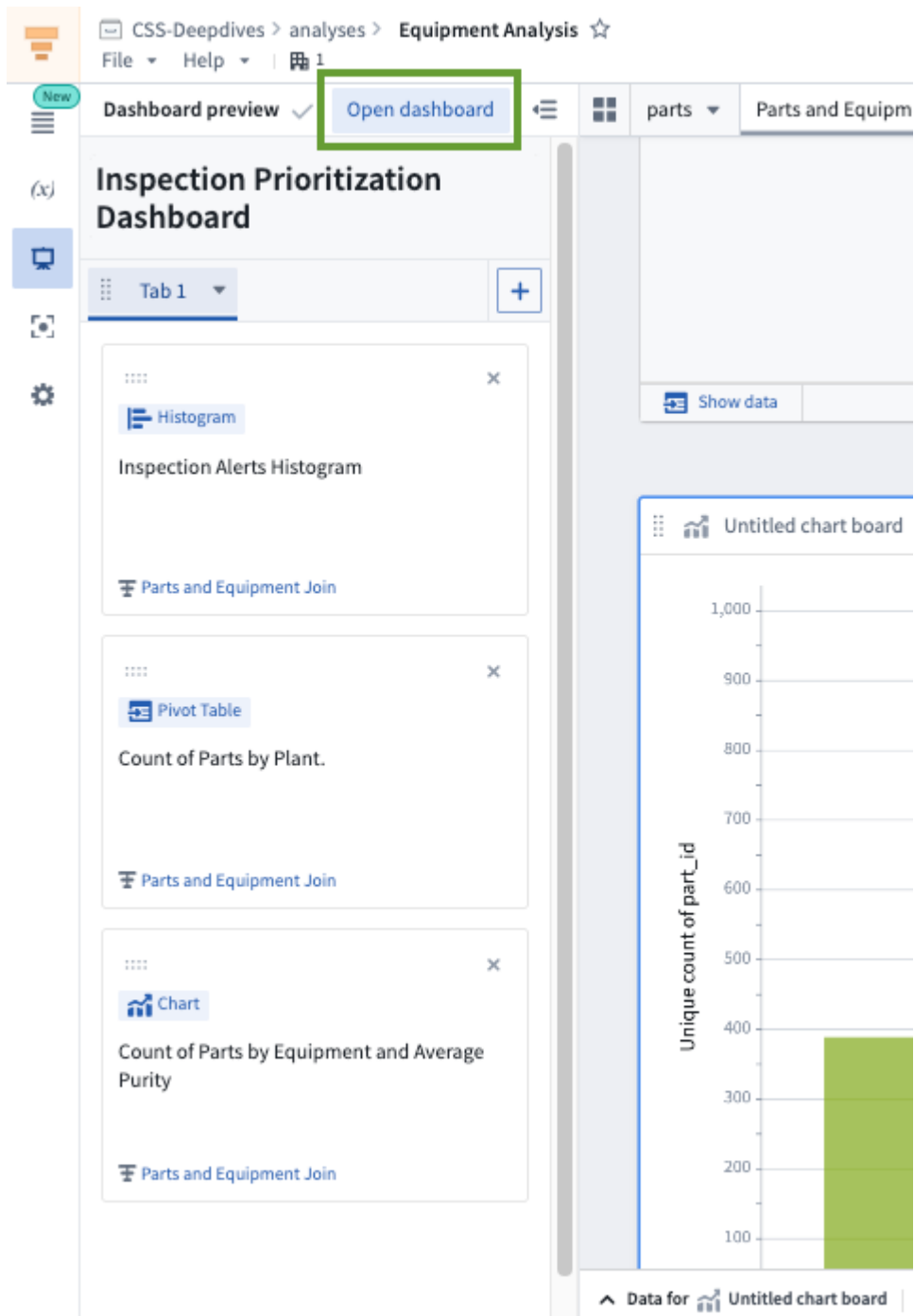


Interact with Your Dashboard

Let's see what this dashboard will look like to viewers of your analysis.

Step 1: Open your dashboard

1. Click **Open dashboard** at the top of the left dashboard panel
 - Alternatively, you can click the blue **Dashboard** button in the top right of the screen.



Step 2: Explore your dashboard filters

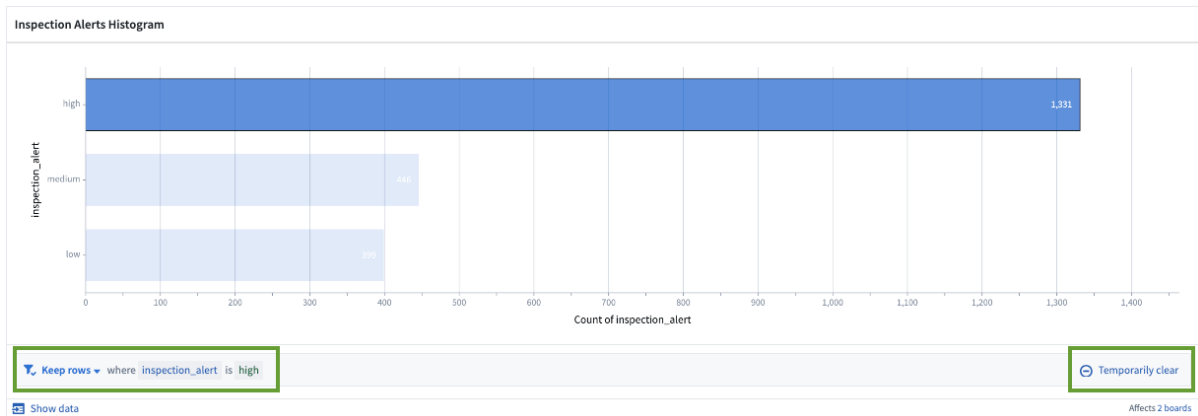
In your dashboard, you may see the filters applied to each board as you selected or defined them in your analysis. These are interactive for viewers of the dashboard.

1. Click on the **high** bar in the Histogram
 - Note how the two lower boards update accordingly.

- Also, the note at the bottom of the Histogram board indicates that the analysis's saved filters have been overridden.

2. Click **Reset**

- This is the curved arrow in the lower right of the Histogram board.



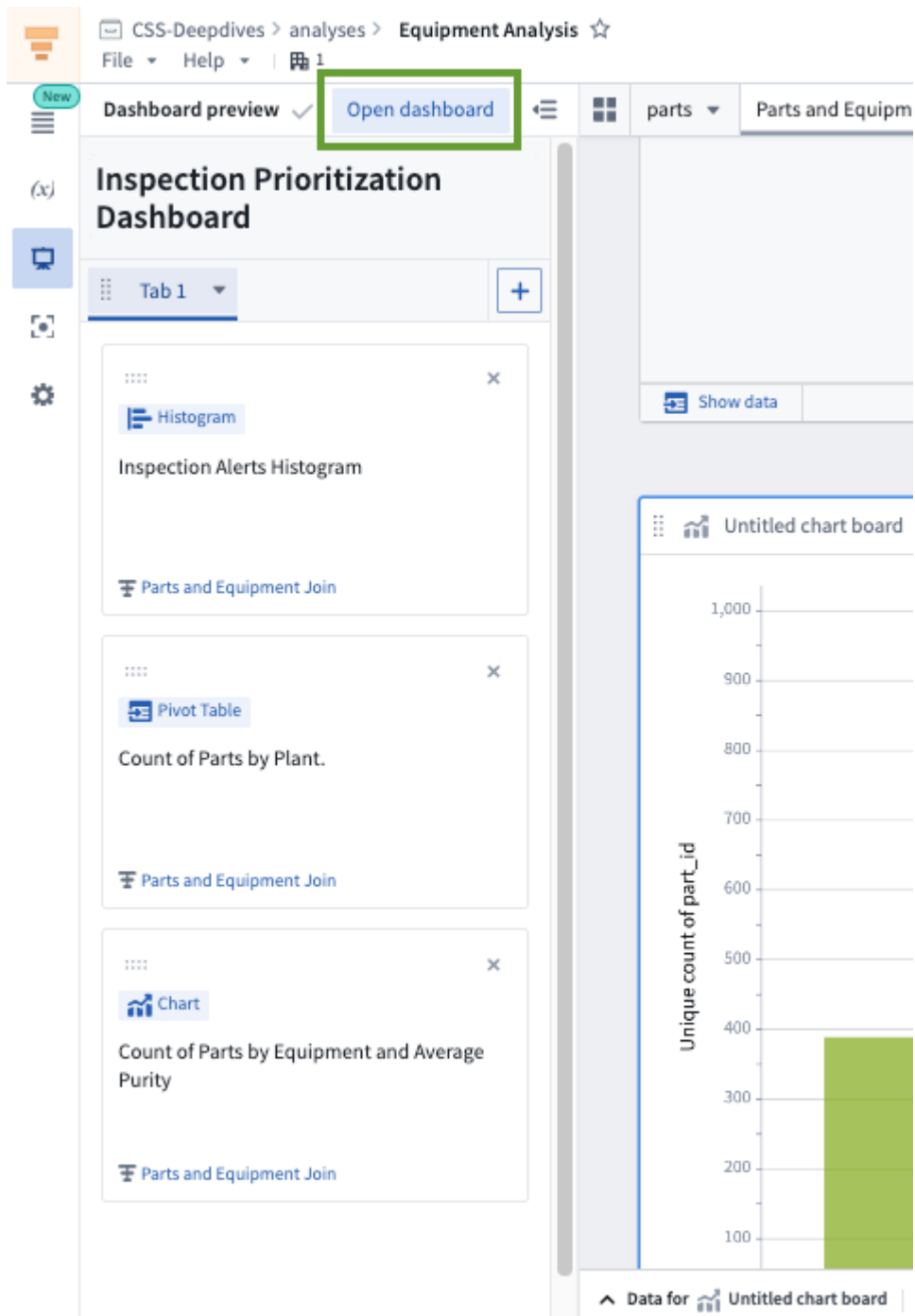
Step 3: Explore your dashboard parameters

On the left side of the page:

1. Click the **(x)** to view the **Parameters**



1. In the **Plant** filter, select **Plant 2**
2. Click **Apply** at the bottom of the Parameters panel
 1. The boards in the dashboard will dynamically update based upon your selection.
3. Clear the parameter
 1. Click the **X** in the **Plant** text field to remove the list of plants
 2. Click **Apply**



Step 4: Explore your dashboard filters

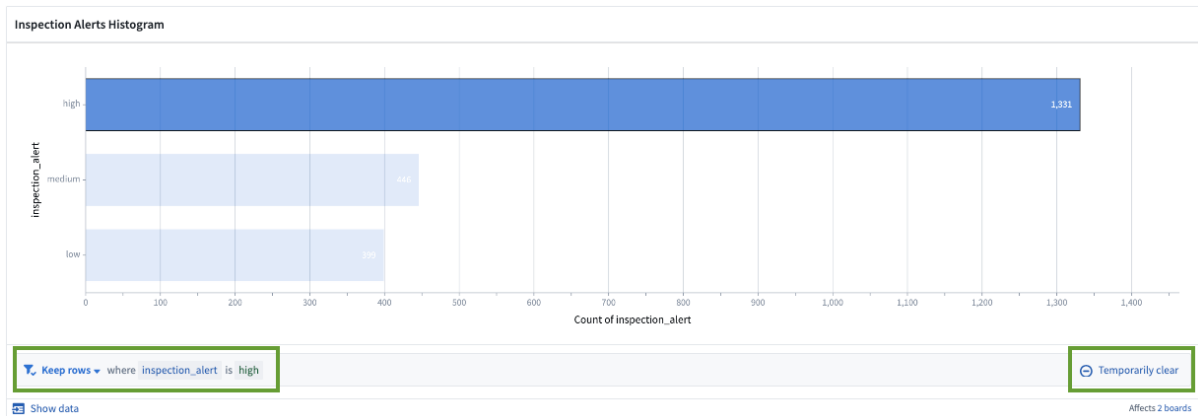
In your dashboard, you may see the filters applied to each board as you selected or defined them in your analysis. These are interactive for viewers of the dashboard.

1. Click on the **high** bar in the Histogram
 - Note how the two lower boards update accordingly.

- Also, the note at the bottom of the Histogram board indicates that the analysis's saved filters have been overridden.

2. Click **Reset**

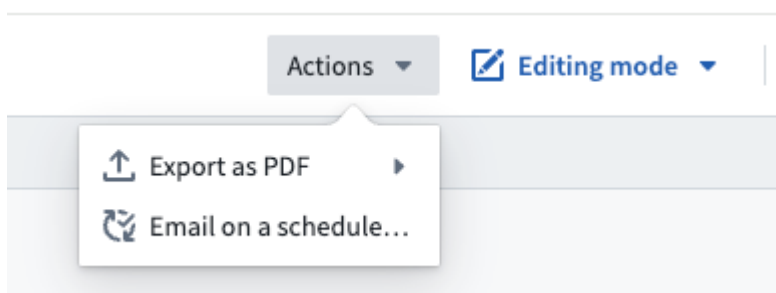
- This is the curved arrow in the lower right of the Histogram board.



Step 5: Export to PDF

Although sending people to your dashboard in Foundry gives them the ability to interact with it and explore the analysis that produced it, sometimes you need a static snapshot.

1. In the top right of the screen, click **Actions > Export as PDF > Portrait**
 - It will take Foundry a few seconds to generate the PDF.
 - Any filters and overrides applied in the dashboard are included in the PDF report.

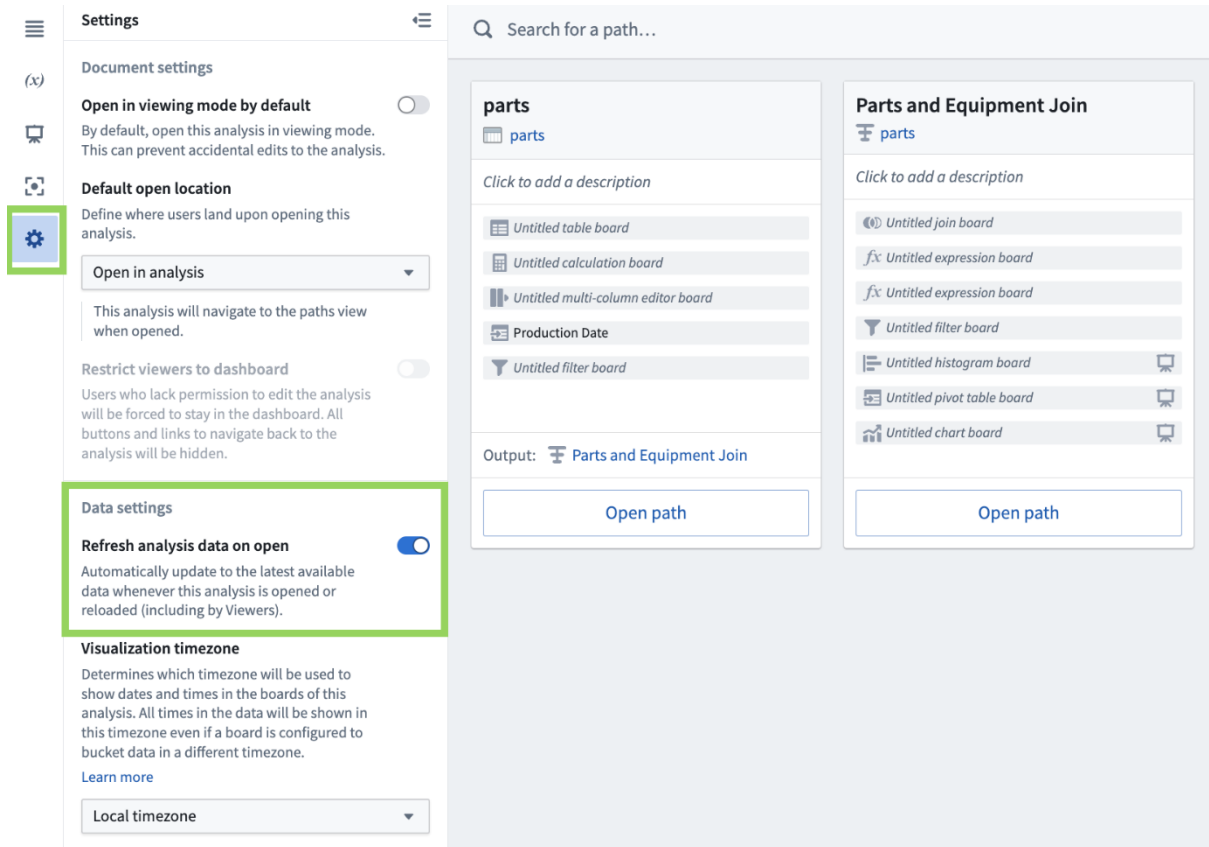


[Optional] Step 5: Email to User

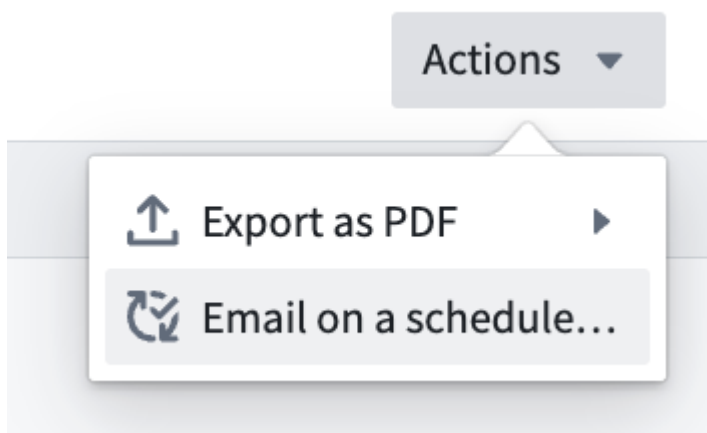
You may not have the permissions required to complete this step. This step is not necessary for completing the course, so you can skip it.

You can configure the dashboard report to be automatically emailed to a specified Foundry user, as follows:

1. Enable data refresh whenever the analysis is opened
 1. Click **Back to analysis** in the top left of the dashboard view
 2. Open **Settings** in the left panel by selecting the cogwheel icon
 3. Under **Data settings**, toggle on **Refresh analysis data on open**



2. Return the **Dashboard** view using the blue button at the top of the screen
3. Select **Actions > Email on a schedule**



4. In the right panel, click **Email this Contour dashboard on a schedule**

- This start to create a recipe. On the Foundry platform, recipes are used to monitor for conditions of interest, to automatically send notifications through Foundry or email to other users, and to deliver a preview of additional context related to the conditions of interest.

5. Configure the recipe

1. Select a recipient (such as yourself) or user group
2. Customize the message with a subject and message
3. Click **Next**
4. Configure the reporting cadence
5. Click **Next**
6. Configure the name, description, and expiration of the recipe that sends this email
7. Click **Save recipe**

An email will be sent to the specified users with a link to the Foundry resource and a PDF of the configured dashboard.

Equipment Analysis

Details > Recipes

Create a recipe

Compose your email

Recipients

Search...

Subject line

e.g. Warning: Threshold exceeded

Message

Write a message (optional)

Customize button

View in Foundry

https://papaya.yojoe.local/workspace/contour-i

Additional email content

Equipment Analysis

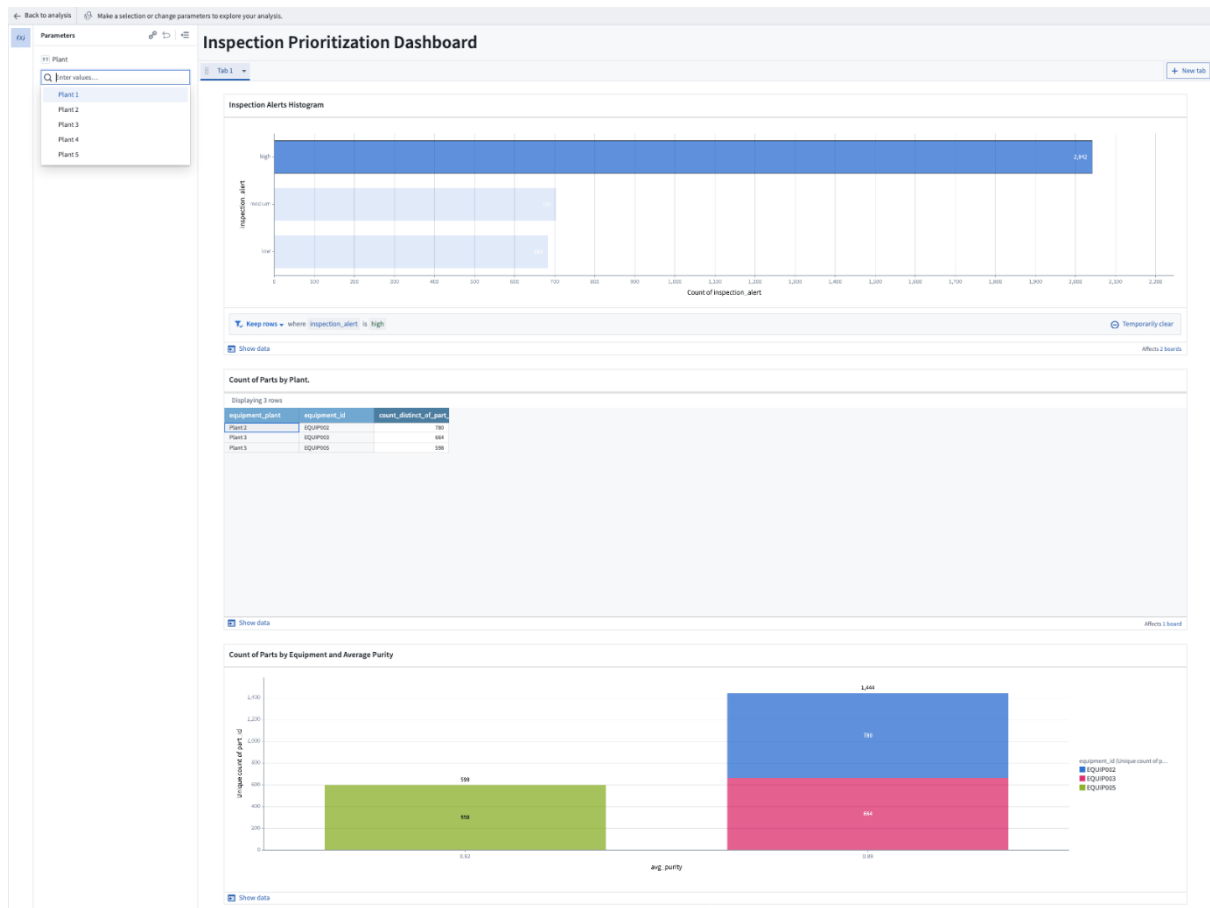
/Public/[yourUsername] Training/Deep Dive: Data Analy...

Include as

PDF attachment

Preview email

At this point, you've built a dashboard tailored for plant managers to showcase your analytical findings in an easily understandable format. The dashboard can auto-update whenever new data is available. The audience will benefit from several features, including chart-to-chart filtering, inline parameter references, a fullscreen presentation view, and the capability to export as PDF. Contour dashboards empower users to interact with live data and results, thereby facilitating the sharing of findings and insights in a well-organized and directed manner.



Recap, Resources, and Tips

Throughout this training, we've explored various features and functionalities of Contour, Palantir Foundry's analytic tool, including data cleaning, joining datasets, creating derived columns, building parameters, visualizing data, and designing an interactive dashboard.

Recap

You've learned how to:

1. Create a new Contour analysis.
2. Use **boards** to analyze and operate on data, such as creating charts, performing calculations, and removing columns. The board toolbar, located at the bottom of the last board, helps you to perform the next logical step of your analysis.
3. Insert a **Table** board to add a tabular view of the active dataset, providing a view of columns, their data types, and a preview of 1,000 rows.
4. Use **Show data** to preview the data backing any board.

5. Use the **Calculation** board to compute summary statistics and understand the dataset's shape, which is useful for validating logic when performing joins, window functions, or complex data transformations.
6. Use the **Multi-Column Editor** to rename and drop columns and the **Convert types** option to change data types of columns.
7. Implement **flexible filtering configurations** with a **Filter** board to keep or remove rows, filter on multiple conditions, any data type or column, and on wildcards.
8. **Join datasets** using Contour's capability to join and union datasets, with the option to join one or more columns across the chosen tables.
9. Derive new columns using an **Expression** board, which provides an expression language that also be used to update existing columns, perform sophisticated filtering, and more.
10. Create **Parameters** to input a specific value(s) into the left parameter panel to filter consistently across multiple paths where the filter board is enabled.
11. Create **Visualizations** such as Histograms, Pivot Tables, and Charts to help interpret and analyze data.
12. Create and interact with **Dashboards** that can be easily exported as a PDF or configured to be automatically emailed to a specified Foundry user.

You also learned best practices for enabling our analysis to power operational decision-making:

1. **Clean and prepare data in separate paths:** Manage the cleaning of disparate datasets in distinct paths to maintain consistency, minimize redundancy, and facilitate the use of results from one path as inputs for another.
2. **Check your logic and validate data at each step:** Use Calculation Boards, Table Boards, and Data bottom drawers to inspect columns, review summary statistics, and validate your logic throughout the analysis. You can always remove boards after validation!
3. **Create Parameters for consistent filtering:** Implement Parameters to input specific values for filtering across multiple paths and to allow users to interact with different values in dashboards.
4. **Optimize performance with limited paths and boards:** Aim to confine your work to no more than five paths, each with a maximum of 20 boards, for optimal performance and clarity in your analysis.