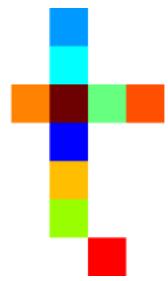




# **LEVEL<sup>UP</sup> WORKSHOP I**

## **Part A**

# 0101 - How to Think



An introduction to the basic data science concepts called Observation, Variable, and Measurement followed by a lesson on the difference between Wide and Long data.

1

There are some key concepts which make understanding data easier no matter in what format that data might appear. In this lesson we will define the concepts **Observation**, **Variable** and **Measurement** and show how to spot them in different data structures.

2

## Think like a Data Scientist



- 3** We will use this cytokine dataset during the lessons ahead.

It is a standard spreadsheet that most people will recognise.

A data scientist will look for three core concepts in this data

Observation, Variable and Measurement

## Cytokine Data

Sample ID	IFN-gamma	IL-12p70	IL-13	IL-1beta	IL-2	IL-4	IL-5	IL-6	TNF-alpha	GM-CSF	IL-18	IL-10	IL-17A	IL-21	IL-22	IL-23	IL-27	IL-9
3130035837	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0
3130035967	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452578	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452632	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452658	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452662	30.06	0	0	0	0	0	0	0	0	0	17.78	0	0	0	0	0	0	0
3370452888	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452917	6.39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452941	0	0	0	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0	0
3370456329	0	0	0	0	0	0	0	0	0.61	0	5.34	0	0	0	0	0	0	0
3400334878	27.22	0	0	0	0	0	0	0	0.75	0	33.43	0	0	0	0	0	0	0
3400334880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400334974	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400335283	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400335823	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.83	0	0	0
3400336107	3.44	0	0	0	0	0	0	0	0.75	0	21.12	0	0	0	0	0	0	0
3400336110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336138	26.15	5.12	0	0	10.7	0	2.55	0	13.47	0	20.66	1.98	16.13	0	34.93	0	0	0
3400336234	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336277	0	0	0	0	0	0	0	0	0	0	9.24	0	0	0	0	0	0	0
3400337948	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

4

**Observation** - Is the fundamental group of interest in the dataset. It is the "thing" we are trying to examine with our experiment.

It may be defined by a patient ID. (e.g. all results for each patient)

It may be defined by a batch ID. (e.g. all results for experiments with each reagent batch)

Or it might be a combination of designators (e.g. All results for each patient for each reagent batch).

In our cytokine dataset the **Observation** is defined by the Sample ID. All of the results associated with a SampleID together form the observation.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Sample ID	IFN-gamma	IL-12p70	IL-13	IL-1beta	IL-2	IL-4	IL-5	IL-6	TNF-alpha	GM-CSF	IL-18	IL-10	IL-17A	IL-21	IL-22	IL-23	IL-27	IL-9
3130035837	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0
3130035967	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452578	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452632	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452658	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452662	30.06	0	0	0	0	0	0	0	0	0	17.78	0	0	0	0	0	0	0
3370452888	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452917	6.39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452941	0	0	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0	0	0
3370456329	0	0	0	0	0	0	0	0.61	0	5.34	0	0	0	0	0	0	0	0
3400334878	27.22	0	0	0	0	0	0	0.75	0	33.43	0	0	0	0	0	0	0	0
3400334880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400334974	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400335283	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400335823	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.83	0	0	0
3400336107	3.44	0	0	0	0	0	0	0	0.75	0	21.12	0	0	0	0	0	0	0
3400336110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336138	26.15	5.12	0	0	10.7	0	2.55	0	13.47	0	20.66	1.98	16.13	0	34.93	0	0	0
3400336234	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336277	0	0	0	0	0	0	0	0	0	0	9.24	0	0	0	0	0	0	0

5

**Variable** - is the fundamental attribute we are comparing across our experiment.  
It is the "what" of an experiment.

It can be defined by the instrument (e.g. signal from the CD8+ channel)

It can be defined by the subject (e.g. Patient Arm length, number of fingers etc.)

It typically is a single entity and not a combination.

In our cytokine data the **Variable** is defined by the Cytokine being measured.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	Sample ID	IFN-gamma	IL-12p70	IL-13	IL-1beta	IL-2	IL-4	IL-5	IL-6	TNF-alpha	GM-CSF	IL-18	IL-10	IL-17A	IL-21	IL-22	IL-23	IL-27	IL-9
2	3130035837	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0
3	3130035967	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3370452578	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	3370452632	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	3370452658	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	3370452662	30.06	0	0	0	0	0	0	0	0	0	17.78	0	0	0	0	0	0	0
8	3370452888	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	3370452917	6.39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	3370452941	0	0	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0	0	0
11	3370456329	0	0	0	0	0	0	0	0.61	0	5.34	0	0	0	0	0	0	0	0
12	3400334878	27.22	0	0	0	0	0	0	0.75	0	33.43	0	0	0	0	0	0	0	0
13	3400334880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	3400334974	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	3400335283	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	3400335823	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.83	0	0	0
17	3400336107	3.44	0	0	0	0	0	0	0.75	0	21.12	0	0	0	0	0	0	0	0
18	3400336110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	3400336121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	3400336138	26.15	5.12	0	0	10.7	0	2.55	0	13.47	0	20.66	1.98	16.13	0	34.93	0	0	0
21	3400336234	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	3400336277	0	0	0	0	0	0	0	0	0	0	9.24	0	0	0	0	0	0	0

## 6 Measurement

is the individual piece of data being recorded.

It is the most important unit (the actual knowledge) but needs the context of the Observation and Variable to give it meaning.

Measurements don't have to be numbers but in biological data they usually are.

In our cytokine data there are **Measurements** in the central cells of the spreadsheet.

Sample ID	IFN-gamma	IL-12p70	IL-13	IL-1beta	IL-2	IL-4	IL-5	IL-6	TNF-alpha	GM-CSF	IL-18	IL-10	IL-17A	IL-21	IL-22	IL-23	IL-27	IL-9
3130035837	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0
3130035967	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452578	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452632	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452658	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452662	30.06	0	0	0	0	0	0	0	0	0	17.78	0	0	0	0	0	0	0
3370452888	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452917	6.39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3370452941	0	0	0	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0	0
3370456329	0	0	0	0	0	0	0	0	0.61	0	5.34	0	0	0	0	0	0	0
3400334878	27.22	0	0	0	0	0	0	0.75	0	33.43	0	0	0	0	0	0	0	0
3400334880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400334974	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400335283	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400335823	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1.83	0	0	0
3400336107	3.44	0	0	0	0	0	0	0	0.75	0	21.12	0	0	0	0	0	0	0
3400336110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336138	26.15	5.12	0	0	10.7	0	2.55	0	13.47	0	20.66	1.98	16.13	0	34.93	0	0	0
3400336234	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3400336277	0	0	0	0	0	0	0	0	0	0	9.24	0	0	0	0	0	0	0



If a biologist learns to think about their data as Observation, Variable, Measurement it becomes much easier for them to...

- Handle any data no matter which instrument it came from.
- Set up understandable views and graphs.
- Merge datasets together.
- Use powerful algorithms.
- Communicate effectively with their Bioinformatics colleagues



Datasets can look very different to a spreadsheet and scientific instruments can present the biologist with a variety of file types and data structures to try to interpret.

But the fundamental concepts of Observation, Variable and Measure can be found.

Here are some examples.

8 Here is a botany flower dataset called Iris in a **json** file format.

```
[{"sepalLength": 5.1, "sepalWidth": 3.5, "petalLength": 1.4, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.6, "sepalWidth": 3.4, "petalLength": 1.4, "petalWidth": 0.3, "species": "setosa"}, {"sepalLength": 5.0, "sepalWidth": 3.4, "petalLength": 1.5, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.4, "sepalWidth": 2.9, "petalLength": 1.4, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.9, "sepalWidth": 3.1, "petalLength": 1.5, "petalWidth": 0.1, "species": "setosa"}, {"sepalLength": 5.4, "sepalWidth": 3.7, "petalLength": 1.5, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.8, "sepalWidth": 3.4, "petalLength": 1.6, "petalWidth": 0.4, "species": "setosa"}, {"sepalLength": 4.8, "sepalWidth": 3.0, "petalLength": 1.4, "petalWidth": 0.1, "species": "setosa"}, {"sepalLength": 4.3, "sepalWidth": 3.0, "petalLength": 1.1, "petalWidth": 0.1, "species": "setosa"}, {"sepalLength": 5.8, "sepalWidth": 4.0, "petalLength": 1.2, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.7, "sepalWidth": 4.4, "petalLength": 1.5, "petalWidth": 0.4, "species": "setosa"}, {"sepalLength": 5.4, "sepalWidth": 3.9, "petalLength": 1.3, "petalWidth": 0.4, "species": "setosa"}, {"sepalLength": 5.1, "sepalWidth": 3.5, "petalLength": 1.4, "petalWidth": 0.3, "species": "setosa"}, {"sepalLength": 5.7, "sepalWidth": 3.8, "petalLength": 1.7, "petalWidth": 0.3, "species": "setosa"}, {"sepalLength": 5.1, "sepalWidth": 3.8, "petalLength": 1.5, "petalWidth": 0.3, "species": "setosa"}, {"sepalLength": 5.4, "sepalWidth": 3.4, "petalLength": 1.7, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.1, "sepalWidth": 3.7, "petalLength": 1.5, "petalWidth": 0.4, "species": "setosa"}, {"sepalLength": 4.6, "sepalWidth": 3.6, "petalLength": 1.0, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.1, "sepalWidth": 3.3, "petalLength": 1.7, "petalWidth": 0.5, "species": "setosa"}, {"sepalLength": 4.8, "sepalWidth": 3.4, "petalLength": 1.9, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.0, "sepalWidth": 3.0, "petalLength": 1.6, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.0, "sepalWidth": 3.4, "petalLength": 1.6, "petalWidth": 0.4, "species": "setosa"}, {"sepalLength": 5.2, "sepalWidth": 3.5, "petalLength": 1.5, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.2, "sepalWidth": 3.4, "petalLength": 1.4, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.7, "sepalWidth": 3.2, "petalLength": 1.6, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.8, "sepalWidth": 3.1, "petalLength": 1.6, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.4, "sepalWidth": 3.4, "petalLength": 1.5, "petalWidth": 0.4, "species": "setosa"}, {"sepalLength": 5.2, "sepalWidth": 4.1, "petalLength": 1.5, "petalWidth": 0.1, "species": "setosa"}, {"sepalLength": 5.5, "sepalWidth": 4.2, "petalLength": 1.4, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 4.9, "sepalWidth": 3.1, "petalLength": 1.5, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.0, "sepalWidth": 3.2, "petalLength": 1.2, "petalWidth": 0.2, "species": "setosa"}, {"sepalLength": 5.5, "sepalWidth": 3.5, "petalLength": 1.3, "petalWidth": 0.2, "species": "setosa"}]
```

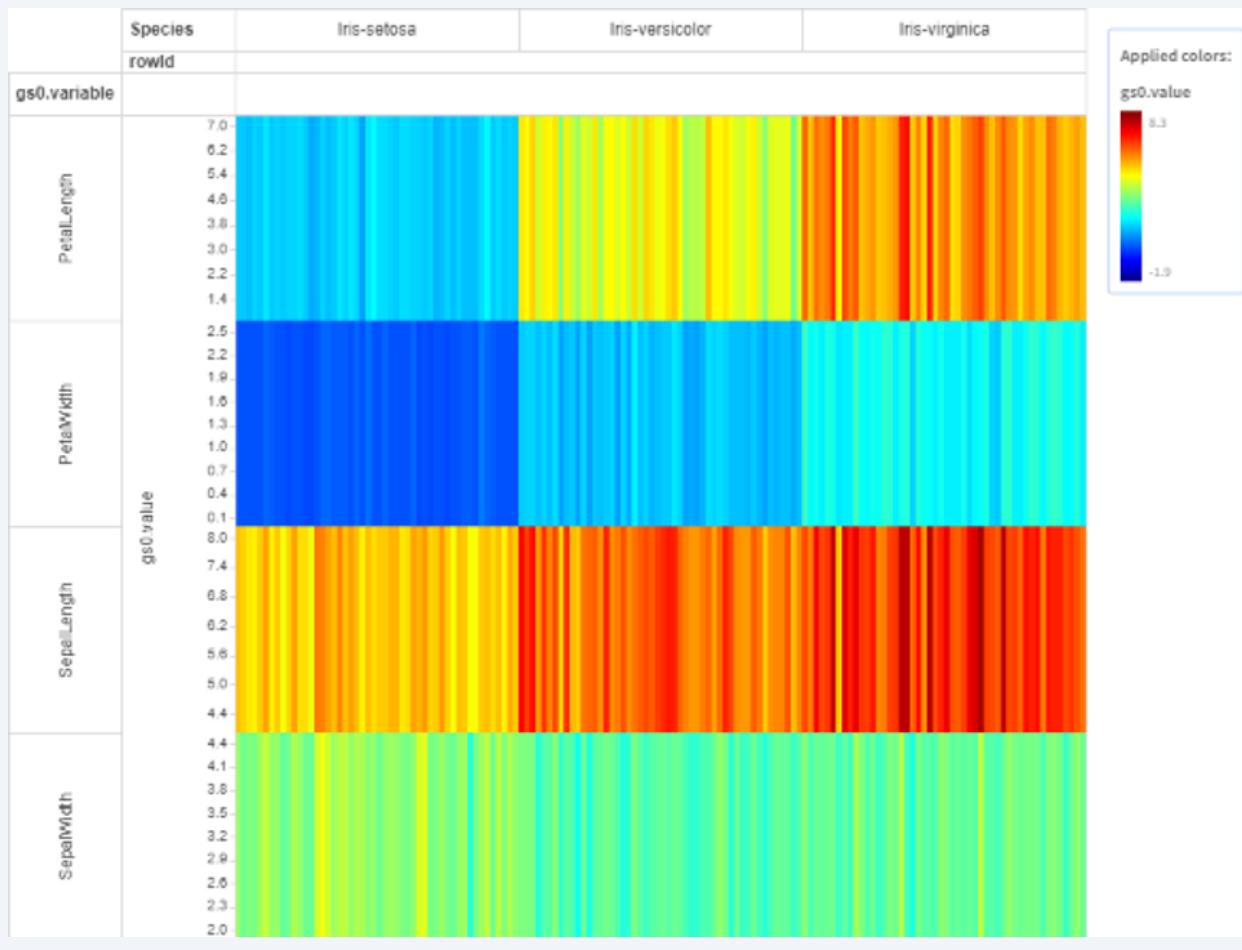
9 Here is a covid vaccination dataset in **xml** format.

```
<response>
<row>
<row_id>row-z6nm.hd7h-tzj5</row_id>
<outcome>Deaths</outcome>
<week_end>2022-07-09T00:00:00</week_end>
<age_group>0-4</age_group>
<unvaccinated_rate>0.0000000000000000</unvaccinated_rate>
<population_unvaccinated>162642</population_unvaccinated>
<outcome_unvaccinated>0</outcome_unvaccinated>
<age_group_min>0</age_group_min>
<age_group_max>4</age_group_max>
</row>
<row_id>row-t2uq-q25r_c6q2</row_id>
<outcome>Cases</outcome>
<week_end>2022-11-12T00:00:00</week_end>
<age_group>0-4</age_group>
<unvaccinated_rate>0.0000000000000000</unvaccinated_rate>
<vaccinated_rate>95.5</vaccinated_rate>
<crude_vaccinated_ratio>14.981818182</crude_vaccinated_ratio>
<population_unvaccinated>162642</population_unvaccinated>
<outcome_unvaccinated>134</outcome_unvaccinated>
<outcome_vaccinated>9</outcome_vaccinated>
<age_group_min>0</age_group_min>
<age_group_max>4</age_group_max>
</row>
<row_id>row-akdf_ec4e_jf29</row_id>
<outcome>Cases</outcome>
<week_end>2022-02-26T00:00:00</week_end>
<age_group>0-4</age_group>
<unvaccinated_rate>0.0000000000000000</unvaccinated_rate>
<population_unvaccinated>162642</population_unvaccinated>
<outcome_unvaccinated>88</outcome_unvaccinated>
<age_group_min>0</age_group_min>
<age_group_max>4</age_group_max>
</row>
<row_id>row-mdgd-hbz2.qqyh</row_id>
<outcome>Hospitalizations</outcome>
<week_end>2021-12-11T00:00:00</week_end>
<age_group>0-4</age_group>
<unvaccinated_rate>0.0000000000000000</unvaccinated_rate>
<population_unvaccinated>162642</population_unvaccinated>
<outcome_unvaccinated>5</outcome_unvaccinated>
```

10

Once Tercen understands what you call the Observation, Variable and Measurement it can reorganize the structure of the original instrument file and present your data so it's easy to read and work with.

For example, here is the botany flower dataset example laid out like a simple spreadsheet. Observations on the rows, Variables on the columns and measurements in each heatmap cell.



11

During the lessons ahead you will learn how to make these visualisations and more.

In the next part of this lesson we will discuss how data that is structured to be easy for humans to read, is not necessarily structured for optimal computer calculations.

By understanding the differences between **Wide** and **Long** data, we can know when to transform our data to get the best use from it, depending on whether a human or an algorithm needs to read it.

## 12 Here is our familiar cytokine data again.

We see how all the Measurements are in a row defined by the Observation and we can use the Variable to cross-reference and find each one.

This is called **Wide** format.

I used the word "see" because Wide format is great for humans. We have eyes and our brains are wired by evolution to find patterns.

We can do cross-references visually in milliseconds.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	Sample ID	IFN-gamma	IL-12p70	IL-13	IL-1beta	IL-2	IL-4	IL-5	IL-6	TNF-alpha	GM-CSF	IL-18	IL-10	IL-17A	IL-21	IL-22	IL-23	IL-27	IL-9
2	3130035837	0	0	0	0	0	0	0	0	0	0	0	0.43	0	0	0	0	0	0
3	3130035967	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	3370452578	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	3370452632	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	3370452658	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	3370452662	30.06	0	0	0	0	0	0	0	0	0	17.78	0	0	0	0	0	0	0
8	3370452888	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	3370452917	6.39	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	3370452941	0	0	0	0	0	0	0	0	0.15	0	0	0	0	0	0	0	0	0
11	3370456329	0	0	0	0	0	0	0	0	0.61	0	5.34	0	0	0	0	0	0	0
12	3400334878	27.22	0	0	0	0	0	0	0.75	0	33.43	0	0	0	0	0	0	0	0
13	3400334880	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	3400334974	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	3400335283	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
16	3400335823	0	0	0	0	0	0	0	0	0	0	0	0	0	1.83	0	0	0	0
17	3400336107	3.44	0	0	0	0	0	0	0	0.75	0	21.12	0	0	0	0	0	0	0
18	3400336110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
19	3400336121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
20	3400336138	26.15	5.12	0	0	10.7	0	2.55	0	13.47	0	20.66	1.98	16.13	0	34.93	0	0	0
21	3400336234	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	3400336277	0	0	0	0	0	0	0	0	0	0	9.24	0	0	0	0	0	0	0

13

Computers, on the other hand, do not have eyes. They find pattern recognition difficult and intensive on their hardware and software resources.

What computers are good at, is doing repetitive tasks extremely quickly and they never get bored, even if they get the same information millions of times in a row.

This is why algorithms are usually coded to work from lists, often called **Long Data**.

Long data has a single Measurement on a line and repeats the Observation and the Variable in each line to give the context to it.

This is what a single line from our cytokine data looks like when it is converted to **Long** format.

	OBSERVATION	VARIABLE	MEASUREMENT
1	Sample ID	Cytokine Name	Percentage
2	3400334878	IFN-gamma	27.22
3	3400334878	IL-12p70	0
4	3400334878	IL-13	0
5	3400334878	IL-1beta	0
6	3400334878	IL-2	0
7	3400334878	IL-4	0
8	3400334878	IL-5	0
9	3400334878	IL-6	0
10	3400334878	TNF-alpha	0.75
11	3400334878	GM-CSF	0
12	3400334878	IL-18	33.43
13	3400334878	IL-10	0
14	3400334878	IL-17A	0
15	3400334878	IL-21	0
16	3400334878	IL-22	0
17	3400334878	IL-23	0
18	3400334878	IL-27	0
19	3400334878	IL-9	0
20			
21			
22			
23			

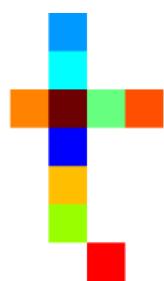
14

As data passes through an analysis pipeline it, at times, will need to appear in wide format for humans to use or, at other times, be structured as Long format for algorithms to process.

This is not something you have to do yourself. There are many computer functions to convert from Wide to Long format.

In Tercen we call this process a **Gather** and in later lessons you will learn how to use it (and our crosstab screen) to structure your data in wide or long format whenever you need it.

# 0102 - Set Up The Workshop



Welcome to the Tercen data workshop.

## Cloning Projects and Downloading Files

- 1 In this section, we will clone a copy of the training materials and download the data file to practice uploading it later.

- 2 Navigate to <https://tercen.com>

- 3 Click "Teams"

The screenshot shows the Tercen software interface. At the top, there is a navigation bar with a home icon, the text 'MartinE', and a search icon. Below the navigation bar, there is a header with the name 'MartinE'. Underneath the header, there is a menu bar with four items: 'Projects' (selected), 'Teams' (highlighted with an orange circle), 'Library', and 'Activities'. Below the menu bar, there is a button for 'New project' and a section titled 'Recent' which lists two projects: 'LevelUpWorkshopsTeam / Workshop 1' and 'HowToTeam / PREC'.

4

Click this icon.

The screenshot shows a digital workspace interface with a header "Martine". Below the header, there are four navigation tabs: "Projects", "Teams" (which is underlined), "Library", and "Activities". A "New team" button is located below the tabs. The main area displays a grid of team icons and names. The teams listed are:

Team Name	Owner
LevelUpWorkshopsTeam	fnaji
AltraBio	fnaji
Licence_Testing	Martine
Roche_Tercen_Workspace	agouy
MarketingTeam	MartinE
MyTransferTeam	MartinE
Tour	fnaji
GerTeam	fnaji
Immunology_blog_Maho	MahoTercen

5

Click "Workshop 1"

The screenshot shows a digital workspace interface with a header. Below the header, there are four navigation tabs: "Projects" (underlined), "Members", "Library", and "Activities". A "New project" button is located below the tabs. The main area displays a list of projects. The projects listed are:

Project Name	Actions
Workshop 1 -Scribe	lock icon
Workshop 1	checkbox, lock icon, trash icon

## 6 Click "files"

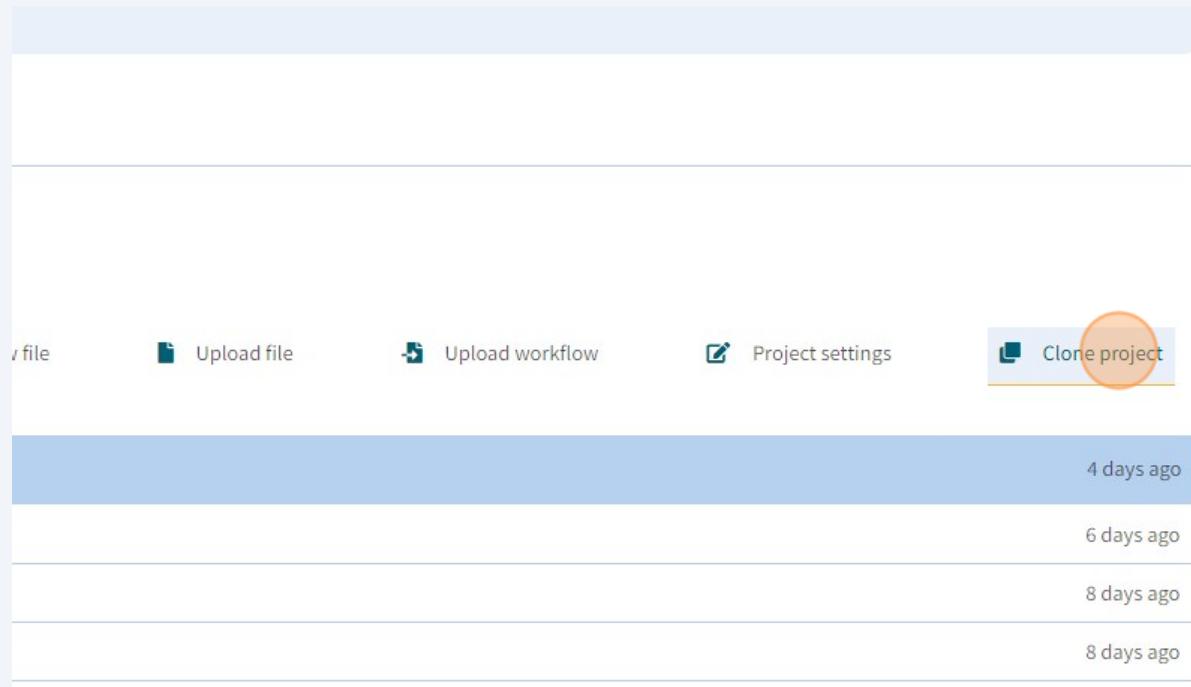
The screenshot shows a project interface with a header "Workshop 1" and a note "No description provided." Below the header are several action buttons: "New data set", "New workflow", "New file", "Upload file", and "Upload workflow". A blue banner at the top indicates "MartinE updated project Workshop 1". The main area displays a file tree with "README.md", "Part-B", "Part-A", and a folder named "files" which is highlighted with a red circle. At the bottom, there is a preview pane showing "README.md".

## 7 Click "Export"

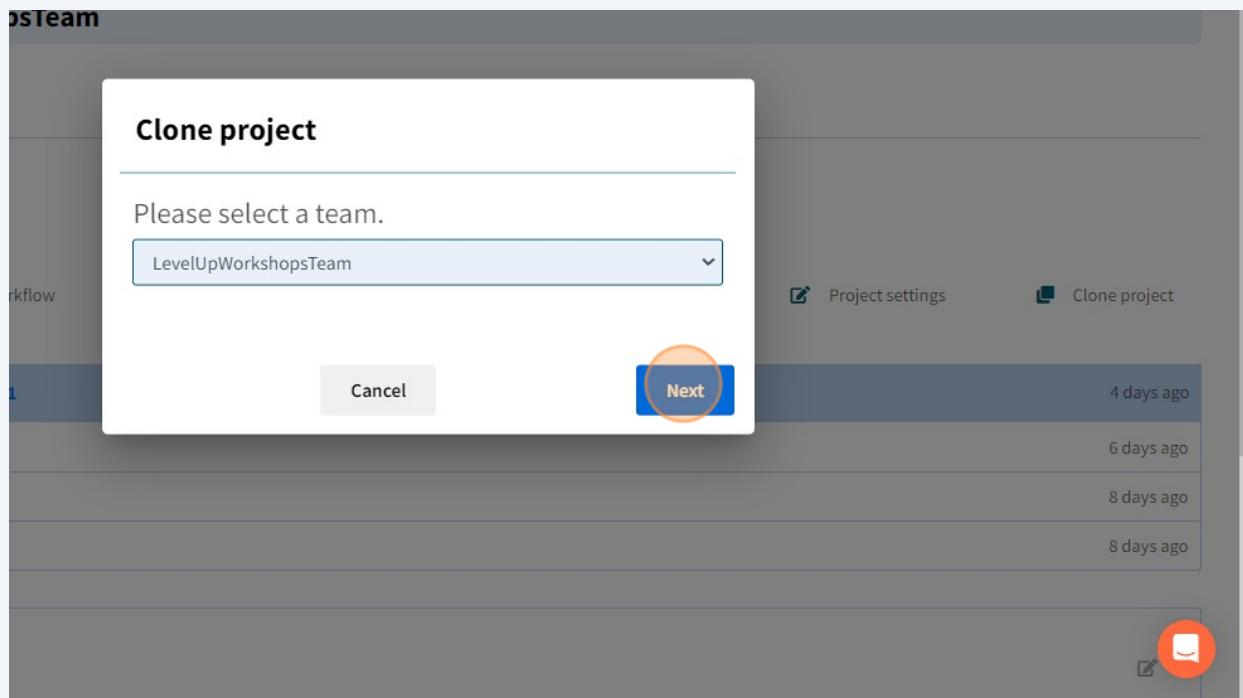
The screenshot shows a project interface for "LevelUpWorkshopsTeam" with a header "Workshop 1" and a note "No description provided." Below the header are action buttons: "New data set", "New workflow", "New file", "Upload file", "Upload workflow", "Project settings", and "Clone project". The main area displays a file list with "README.md", "grifols\_data\_final.csv", and "grifols\_cytokine\_data.csv". A red circle highlights the "Export" button, which is located next to the file list. At the bottom, there is a preview pane showing "README.md".

**8** The file will download to your local drive.

**9** Click "Clone project"

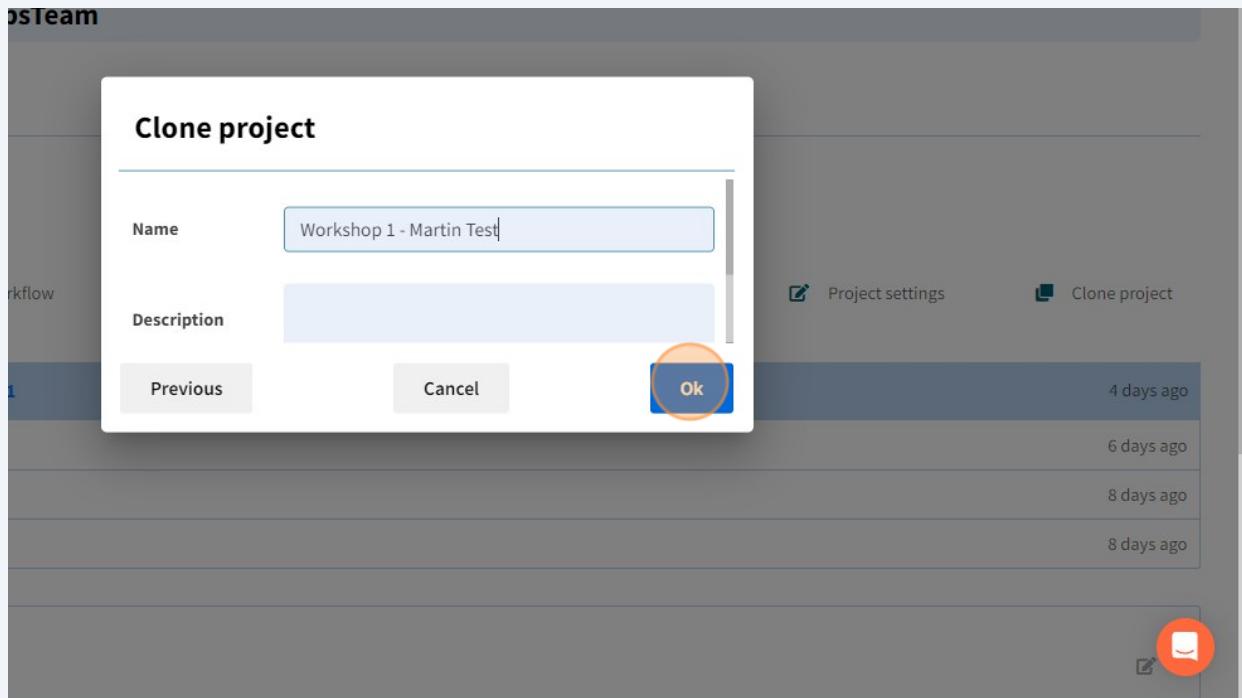


**10** Click "Next"



**11** Add your own name to the end of the cloned project name.

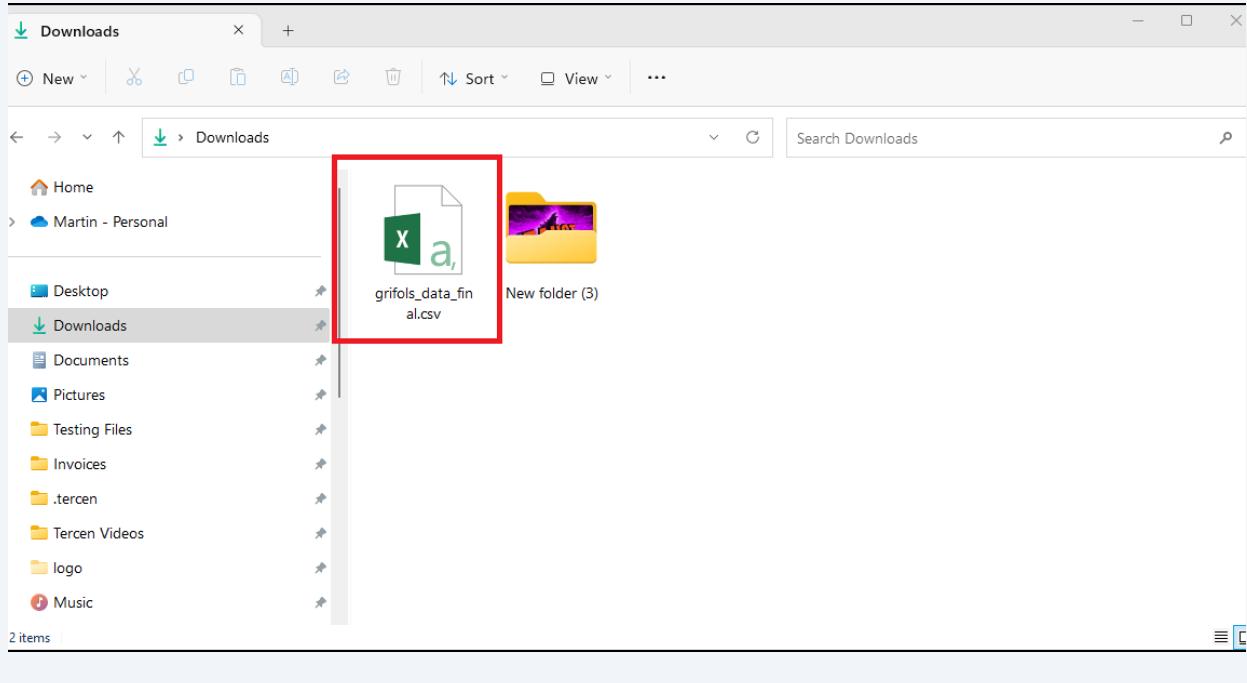
**12** Click "Ok"



This will make it easy for the instructors to help you should you need it.

## 13 Now, let's get familiar with the data.

Leave Tercen for a moment and browse to your download folder.



## 14 Open the data file in your spreadsheet program (such as MS Excel)

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	SUBJECT ID	AGE	AGEGR1N	AGEGR1C	RACEGRP	SEXN	FASFL	SEXC	Batch date	PANEL	CELL SUBSET	Unit	Parent	VALUE	URI
2	3130035837	23	1	18-29	Caucasian	1	Y	Female	02-Jul-15	Tcell	Effector Memory CD4+ T cells	%	CD4+	7.15	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
3	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	DCMNK	NK cells	%	Lineage-	8.07	
4	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	DCMNK	Myeloid DC (Live Scatter)	%	Live Scatter	0.683	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
5	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	DCMNK	Monocytes	%	Singlets	9.219	
6	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated Naive CD4+ T cells	%	CD4+ Naive	0.819	
7	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Effector CD8+ T cells	%	CD8+	14.2	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
8	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated CD8+ T cells	%	CD8+	2.69	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
9	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Central Memory CD4+ T cells	%	CD4+	31.6	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
10	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated Central Memory CD8+ T cells	%	CD8+ Central Memory	12.7	
11	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated Effector CD8+ T cells	%	CD8+ Effector	5.66	
12	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated Effector Memory CD8+ T cells	%	CD8+ Effector Memory	13.4	
13	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated Naive CD8+ T cells	%	CD8+ Naive	1.03	
14	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	CD8+ T cells	%	Lymphocytes	36.3	
15	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Central Memory CD8+ T cells	%	CD8+	4.06	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
16	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Effector CD4+ T cells	%	CD4+	1.48	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
17	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Effector Memory CD4+ T cells	%	CD4+	9.69	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
18	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Activated Effector Memory CD4+ T cells	%	CD4+ Effector Memory	5.88	
19	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Tcell	Effector Memory CD8+ T cells	%	CD8+	4.42	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
20	3400338745	22	1	18-29	African-American	0	Y	Male	24-Jan-17	Tfh	T follicular helper type 2	%	CD4+CXCR5+	28.2	
21	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Treg	Treg cells	%	CD4+	10.5	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>
22	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Treg	HLA-DR- Memory Treg cells	%	CD194+	36.9	
23	3400338745	22	1	18-29	African-American	0	Y	Male	01-Jul-15	Treg	Effector Treg cells	%	CD25+CD127lo	46.4	<a href="http://purl.obolibrary.org">http://purl.obolibrary.org</a>

15

Each row is a single measurement. There are 127 unique patients but each one has had multiple measurements taken.

Note the header names of the columns. These will be important later when we use the Tercen program.

**SUBJECT ID - The code identifying an individual donor.**

**AGE - Patient Information**

**RACEGRP - Patient Information**

**SEXC - Patient Information**

**CELL SUBSET - The immunophenotype name**

**VALUE - This is the measurement (% of population)**

16

When you have finished reviewing the data, return to Tercen.

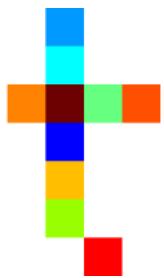
You are now ready to start the tutorial.

The screenshot shows the Tercen software interface. At the top, there is a navigation bar with a search bar labeled "Search Tercen", "Learn", "Explore", and a user icon. Below the navigation bar, the title "LevelUpWorkshopsTeam" is displayed, followed by a sub-project title "Workshop 1 - Martin Test". A message "No description provided." is shown next to a large orange circular placeholder. Below this, there are several file and workflow management buttons: "New data set", "New workflow", "New file", "Upload file", "Upload workflow", "Project settings", and "Clone project". A list of files is shown, including "README.md", "Part-B", "Part-A", and "files", all uploaded 10 seconds ago. At the bottom right of the interface, there is a red speech bubble icon.

17

Well done!

# 0103 - Upload Data



We will begin Tutorial A by learning how to upload data files to Tercen.

- Starting from your cloned **Workshop I** project.

The screenshot shows the Tercen platform interface. At the top, there is a navigation bar with the Tercen logo, a search bar labeled "Search Tercen", and links for "Learn" and "Explore". Below the navigation bar, the URL "LevelUpWorkshopsTeam > Workshop 1 - Martin" is visible. The main area displays the "LevelUpWorkshopsTeam" project. The "Project" tab is selected, showing a summary of the project: "Workshop 1 - Martin" with no description provided. Below this, there are several action buttons: "New data set", "New workflow", "New file", "Upload file", "Upload workflow", "Project settings", and "Clone project". A recent activity feed is displayed, showing a message from "MartinE" updating a workflow named "cytokine\_analysis" 22 minutes ago, along with four file attachments: "README.md", "Part-B", "Part-A", and "files", all uploaded 2 days ago. Below the activity feed, a file list shows "README.md". A sidebar on the right contains the text "Workshop I".

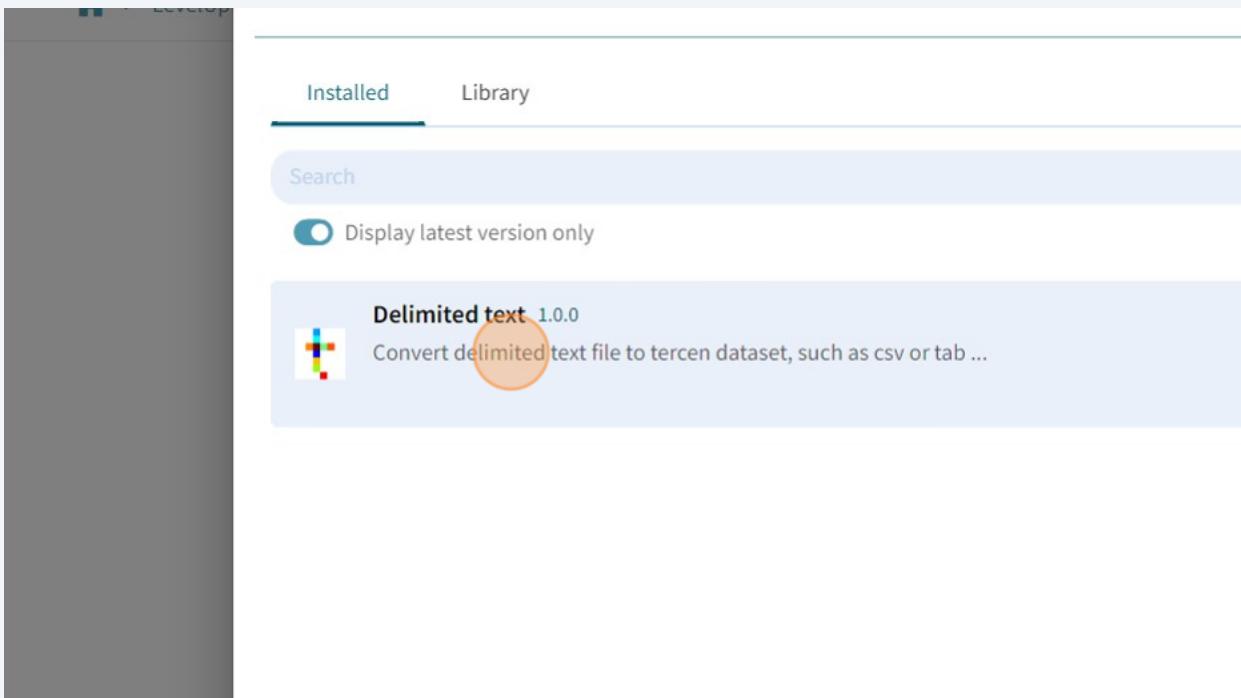
- 2 Open the "Part-A" folder.

The screenshot shows a user interface for managing files and workflows. At the top, there are four buttons: "New data set", "New workflow", "New file", and "Upload file". Below this is a header bar with the text "MartinE updated workflow cytokine\_analysis". Underneath the header, there is a list of items: "README.md", "Part-B", "Part-A" (which is circled in orange), and "files". In the bottom right corner of the main area, the text "Workshop 1" is visible.

- 3 Click "New data set"

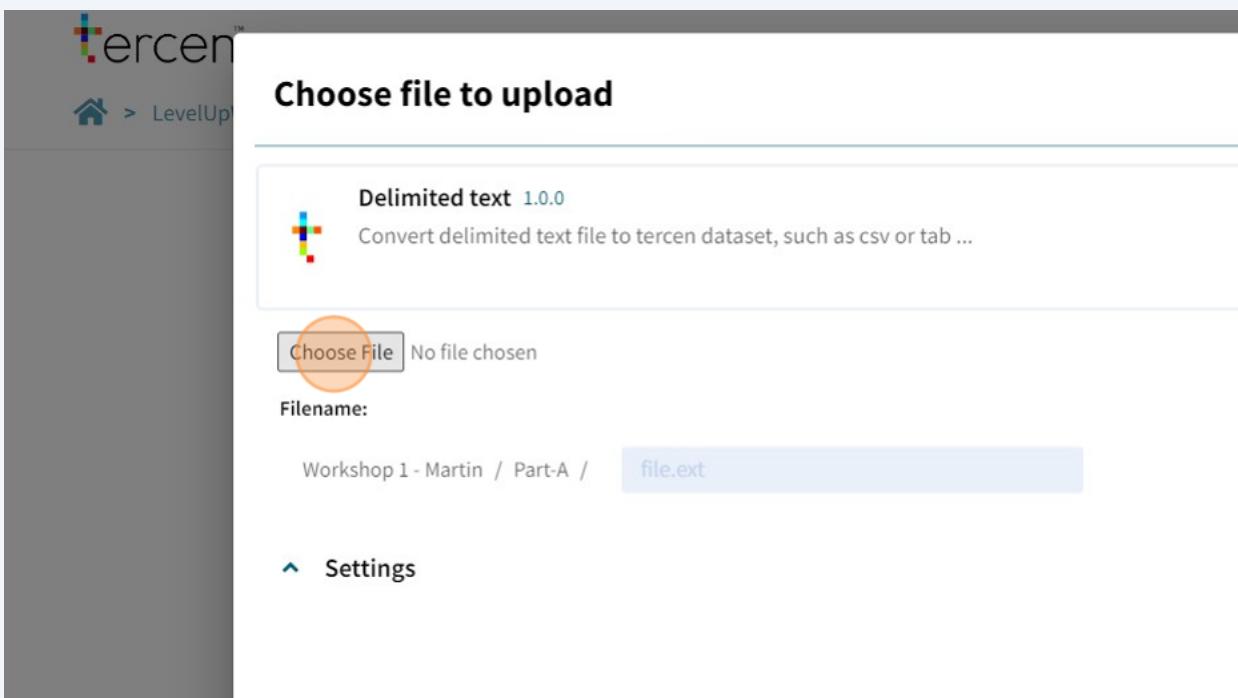
The screenshot shows a project management interface. At the top, there is a logo and the text "LevelUpWorkshopsTeam". Below this, there are two tabs: "Project" (which is selected) and "Activities". Under the "Project" tab, there is a section titled "Workshop 1 - Martin" with the sub-section "Part-A". The text "No description provided." is displayed. At the top of this section, there are four buttons: "New data set" (which is circled in orange), "New workflow", "New file", and "Upload file". Below this, there is a header bar with the text "MartinE updated workflow cytokine\_analysis" and a list of items: "README.md". In the bottom right corner of the main area, the text "Workshop 1" is visible.

**4** Click the "Delimited text" importer

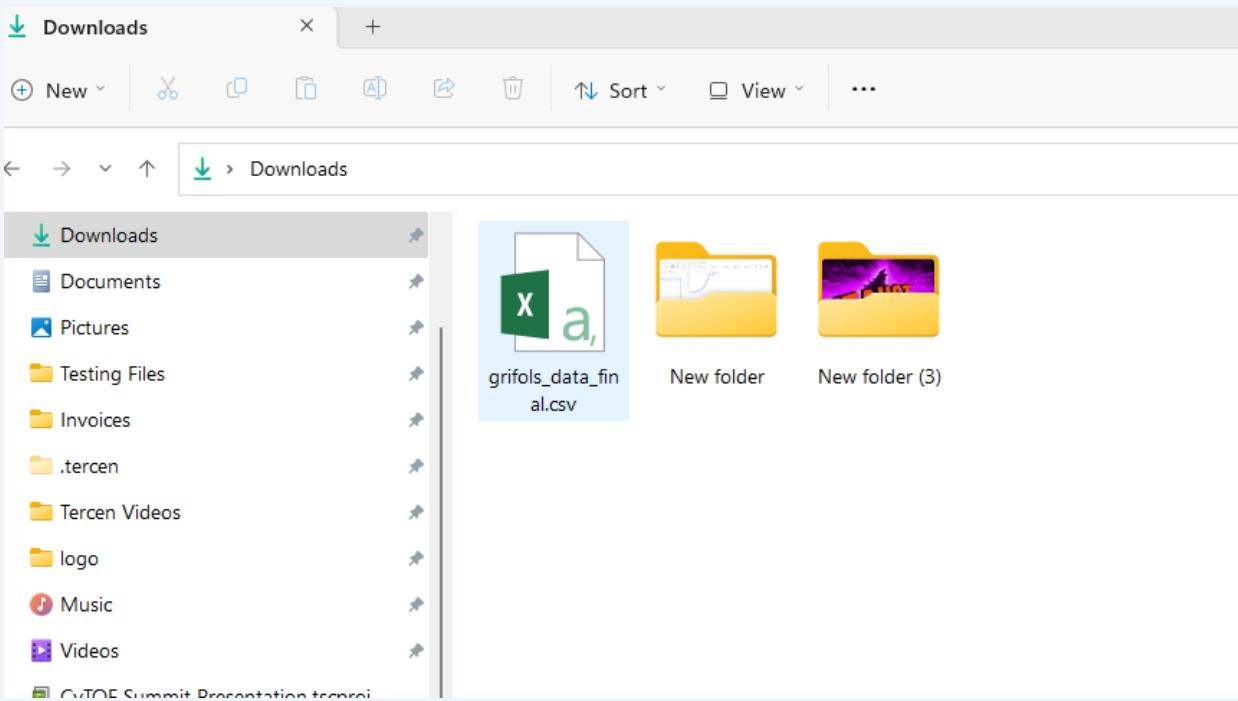


There are Importers for specific file types (such as FCS files) that will appear here when they are installed.

- 5 Press the **Choose File** button and browse to the data file.

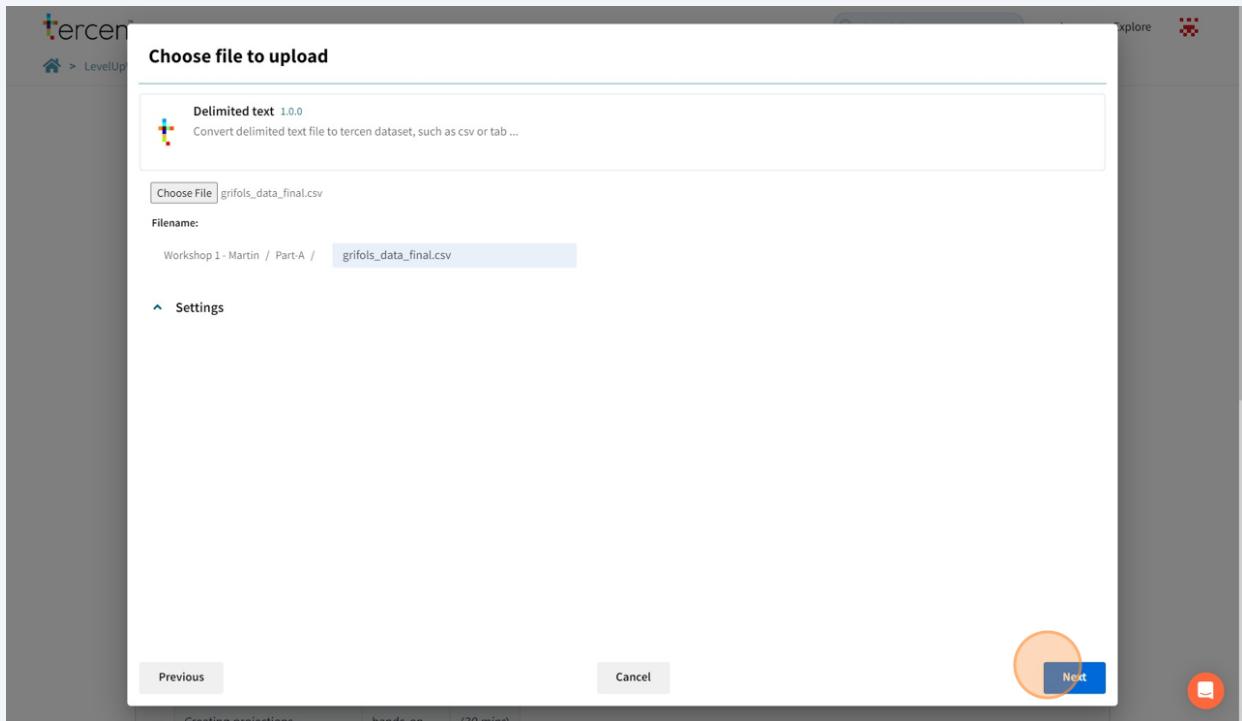


- 6 Select the file and press ok.



7

Click "Next"



8

Some importers have a review stage where you can check if your data is okay and make adjustments before it uploads.

For example

SUBJECT ID is a number in our data file, but we know it is not one we will use to do mathematical calculations.

We are more likely to use it as a label on a graph.

Change the data type from **numeric** to **character**.

Check column names and types						Advanced
	#	SUBJECT ID	AGE	AGEGRIN	AGEGRIC	RACEGRP
1		<input type="button" value="numeric ▾"/> <b>character</b> numeric ignore	<input type="button" value="numeric ▾"/>	<input type="button" value="numeric ▾"/>	<input type="button" value="character ▾"/>	<input type="button" value="character ▾"/>
2	3400338745.0	23.0	1.0	18-29	Caucasian	
3	3400338745.0	22.0	1.0	18-29	African-American	
4	3400338745.0	22.0	1.0	18-29	African-American	
5	3400338745.0	22.0	1.0	18-29	African-American	
6	3400338745.0	22.0	1.0	18-29	African-American	



Any columns you set to "ignore" will not upload.

**9** Click "Upload"

The screenshot shows the Tercen platform interface. At the top, there's a navigation bar with the Tercen logo and a 'LevelUp' link. Below it, a section titled 'Check column names and types' displays a table of data. The table has columns labeled '#', 'SUBJECT ID', 'AGE', 'AGEGR1N', 'AGEGRIC', 'RACEGRP', 'SEXN', and 'F1'. The data rows show various values for these columns. At the bottom of the table, there are 'Previous' and 'Cancel' buttons, and a prominent 'Upload' button which is circled in red.

**10** The data is uploaded and converted to a dynamic table, which **Tercen** can manipulate in visualisations and calculations.

You can click on the Table to review the data inside.

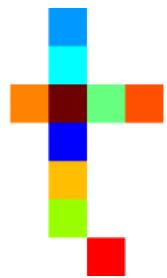
The screenshot shows the Tercen platform interface after the upload. At the top, there are buttons for 'New data set', 'New workflow', 'New file', and 'Upload file'. Below that, a navigation bar says 'Workshop 1 - Martin / Part-A'. Underneath, a blue header bar indicates 'MartinE updated workflow cytokine\_analysis'. A list of files is shown, with 'grifols\_data\_final.csv' highlighted by a red circle. Below this, another list shows 'README.md'. At the bottom, the text 'Part-A Workshop I' is displayed.

## 11 You will see a Tercen data table.

Click on the Project name in the breadcrumb to begin the next lesson.

#	SUBJECT ID	AGE	AGEGR1N	AGEGR1C	RACEGRP	SEXN	FASFL	SEXC	Batch date	PANEL	CELL SUBSET	Unit	Parent	VALUE	URI
1	3130035837	23.0	1.0	18-29	Caucasian	1.0	Y	Female	02JUL2015	Tcell	Effector Memory CD4+ T cells	%	CD4+	7.15	<a href="http://purl.obolibrary.org/obc">http://purl.obolibrary.org/obc</a>
2	3400338745	22.0	1.0	18-29	African-American	0	Y	Male	01JUL2015	DCMNK	NK cells	%	Linage-	8.07	
3	3400338745	22.0	1.0	18-29	African-American	0	Y	Male	01JUL2015	DCMNK	Myeloid DC (Live Scatter)	%	Live Scatter	0.68	<a href="http://purl.obolibrary.org/obc">http://purl.obolibrary.org/obc</a>
4	3400338745	22.0	1.0	18-29	African-American	0	Y	Male	01JUL2015	DCMNK	Monocytes	%	Singlets	9.22	
5	3400338745	22.0	1.0	18-29	African-American	0	Y	Male	01JUL2015	Tcell	Activated Naive CD4+ T cells	%	CD4+ Naive	0.82	
6	3400338745	22.0	1.0	18-29	African-American	0	Y	Male	01JUL2015	Tcell	Effector CD8+ T cells	%	CD8+	14.2	<a href="http://purl.obolibrary.org/obc">http://purl.obolibrary.org/obc</a>
7	3400338745	22.0	1.0	18-29	African-	0	Y	Male	01JUL2015	Tcell	Activated	%	CD8+	2.69	<a href="http://purl.obolibrary.org/obc">http://purl.obolibrary.org/obc</a>

# 0104 - Crosstab Basics



In this tutorial you will learn the basic concepts of the Tercen crosstab grid and make some data visualisations.

- 1 Starting from your Workshop project.

The screenshot shows the Tercen software interface. At the top, there is a navigation bar with the Tercen logo and the path: Home > LevelUpWorkshopsTeam > Workshop 1 - Martin. The 'Workshop 1 - Martin' link is highlighted with a light orange circle. Below the navigation bar, there is a sidebar with a purple square icon labeled 'Project' and a grey square icon labeled 'Activities'. The main content area is titled 'LevelUpWorkshopsTeam' and shows a section for 'Workshop 1 - Martin'. It says 'No description provided.' and features four buttons: 'New data set' (with a blue square icon), 'New workflow' (with a green square icon), 'New file' (with a yellow square icon), and 'Upload file' (with a red square icon). A blue banner at the bottom indicates that 'MartinE updated workflow cytokine\_analysis'.

- 2 In Tercen a data analysis pipeline is called a **Workflow**.

Click "New workflow"

The screenshot shows the Tercen platform interface. At the top, there's a navigation bar with a purple square icon and the text "LevelUpWorkshopsTeam". Below it, a sub-navigation bar has "Project" underlined and "Activities". Under "Project", there's a section titled "Workshop 1 - Martin" with a lock icon. A note says "No description provided.". Below this, there are four buttons: "New data set", "New workflow" (which is highlighted with an orange circle), "New file", and "Upload file". A blue header bar below these buttons displays "MartinE updated workflow cytokine\_analysis". Underneath, a sidebar lists "README.md", "Part-B", "Part-A", and "files".

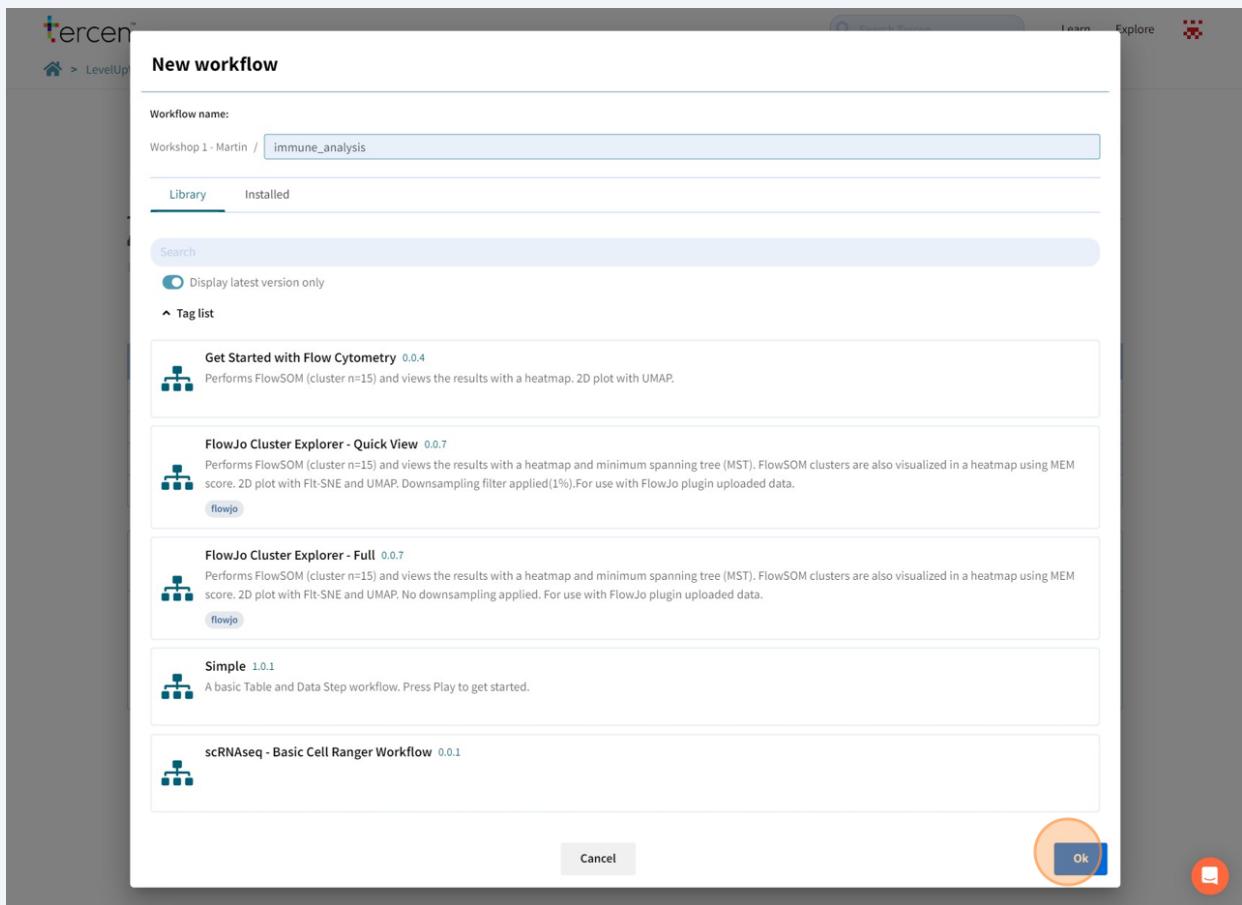
- 3 Click the "*Name field*"

Name your workflow "immune\_analysis"

The screenshot shows the "New workflow" creation dialog. At the top, it says "Workflow name:" followed by "Workshop 1 - Martin / Name your workflow ...". This input field is also highlighted with an orange circle. Below it, there are tabs for "Library" (underlined) and "Installed". There's a "Search" bar and a checked checkbox for "Display latest version only". A "Tag list" section is partially visible. At the bottom, there's a card for "Get Started with Flow Cytometry 0.0.4" which describes it as performing FlowSOM (cluster n=15) and viewing results with a heatmap, 2D plot with tSNE.

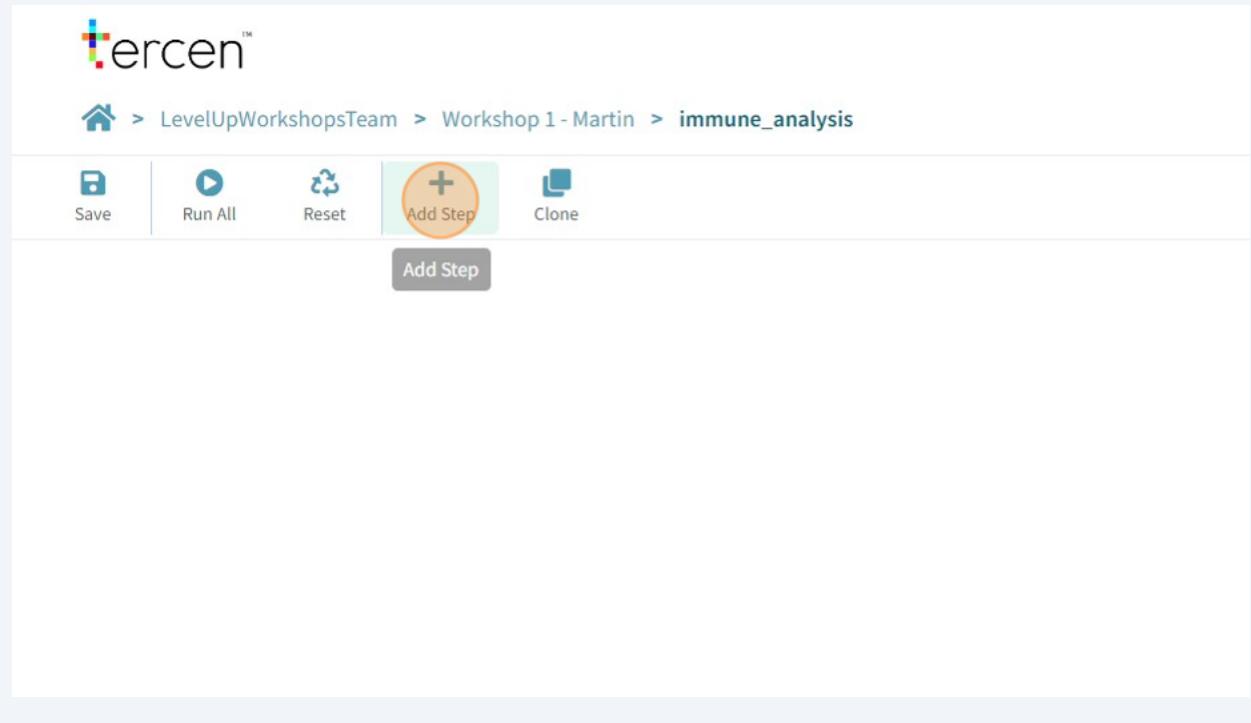
4

Click "Ok"

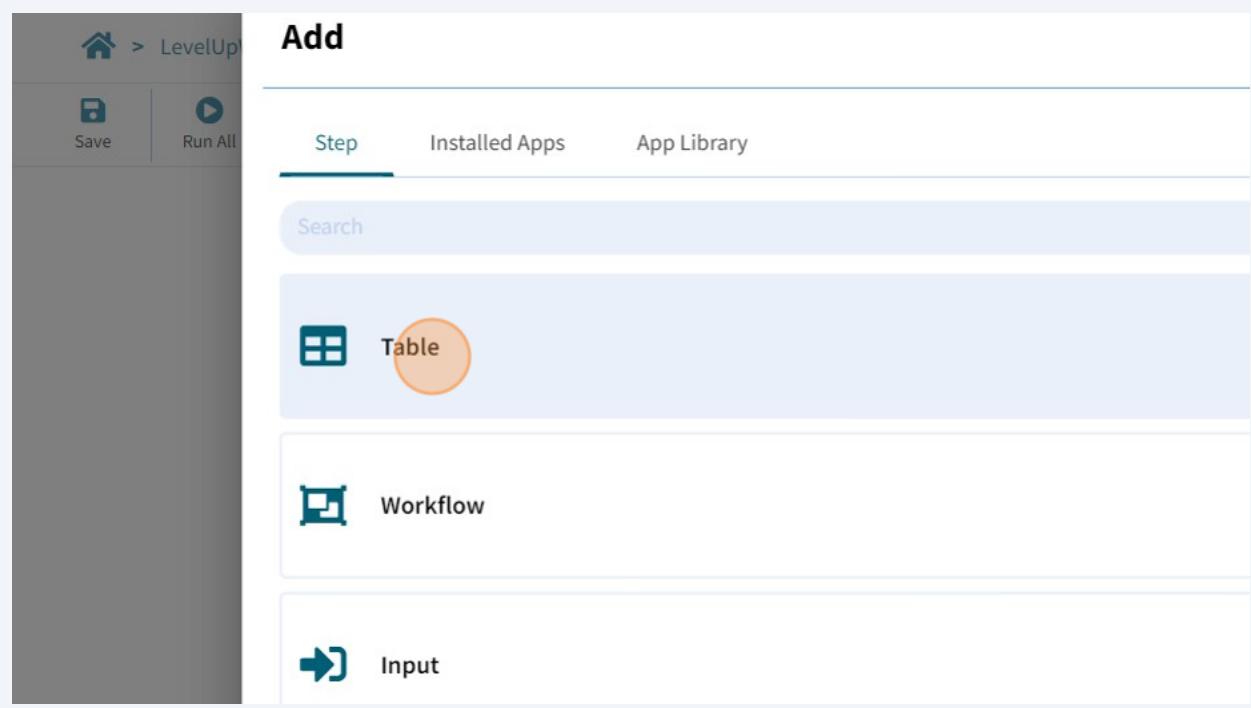


**5** The first part of making a workflow is to add the data.

Click "Add Step" in the **Global Toolbar**.



**6** Choose "Table" from the list.



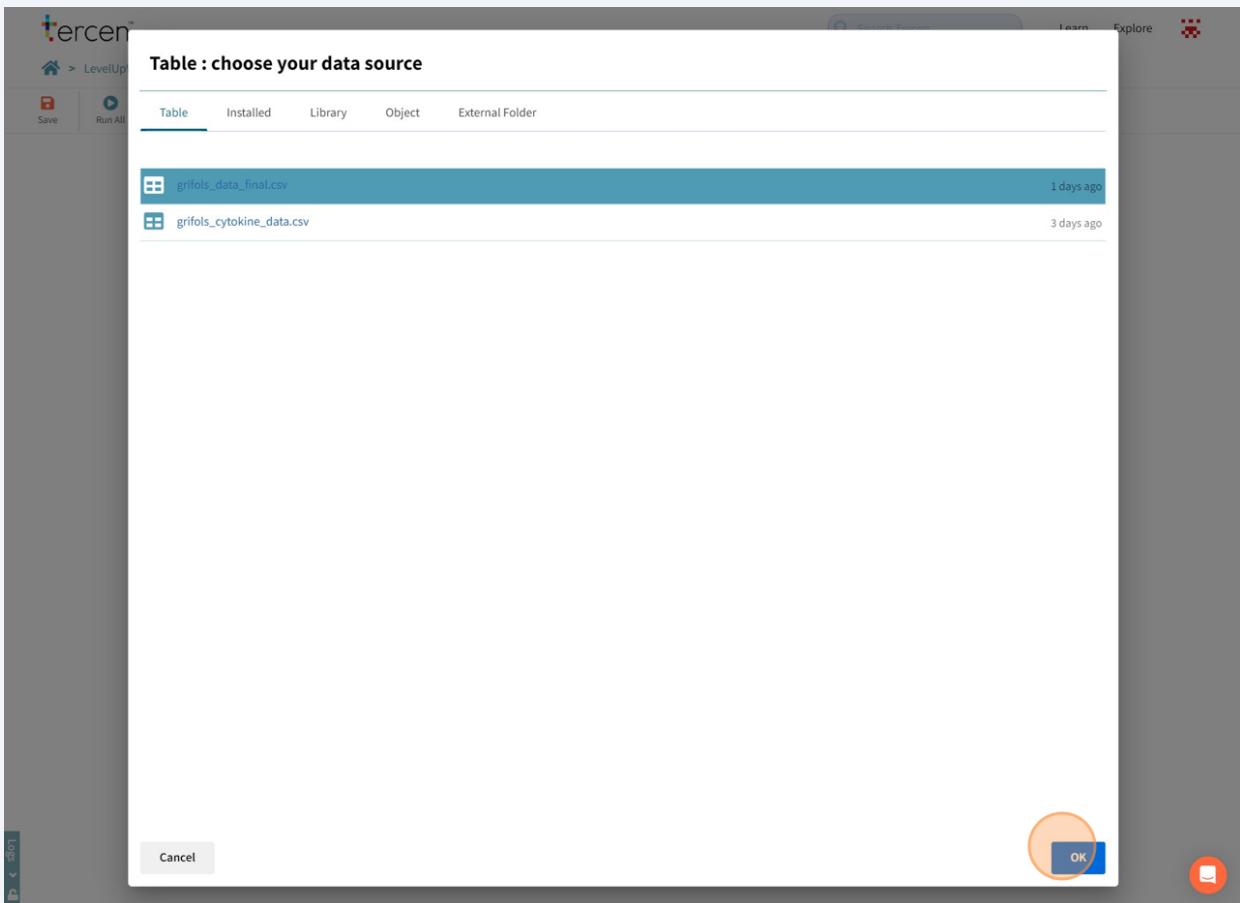
- 7 Click **grifols\_data\_final.csv** which you uploaded earlier.

The screenshot shows the tercen™ software interface. At the top, there is a navigation bar with the tercen logo, a home icon, and the text "LevelUp". Below the navigation bar, there are two buttons: "Save" and "Run All". The main area is titled "Table : choose your data source". There are five tabs at the top of this section: "Table" (which is underlined in blue), "Installed", "Library", "Object", and "External Folder". Below the tabs, there is a list of files. The first file, "grifols\_data\_final.csv", has a small orange circle drawn around it. The second file, "grifols\_cytokine\_data.csv", is also listed. Both files have a small icon next to them.

8

Click "OK"

A Table will be added to your workflow canvas.



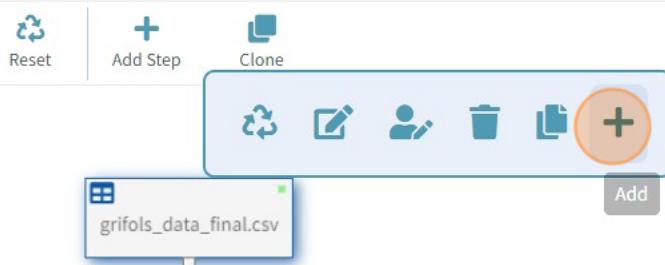
## 9 Click the Table

A **Local Toolbar** will appear with options to choose from.



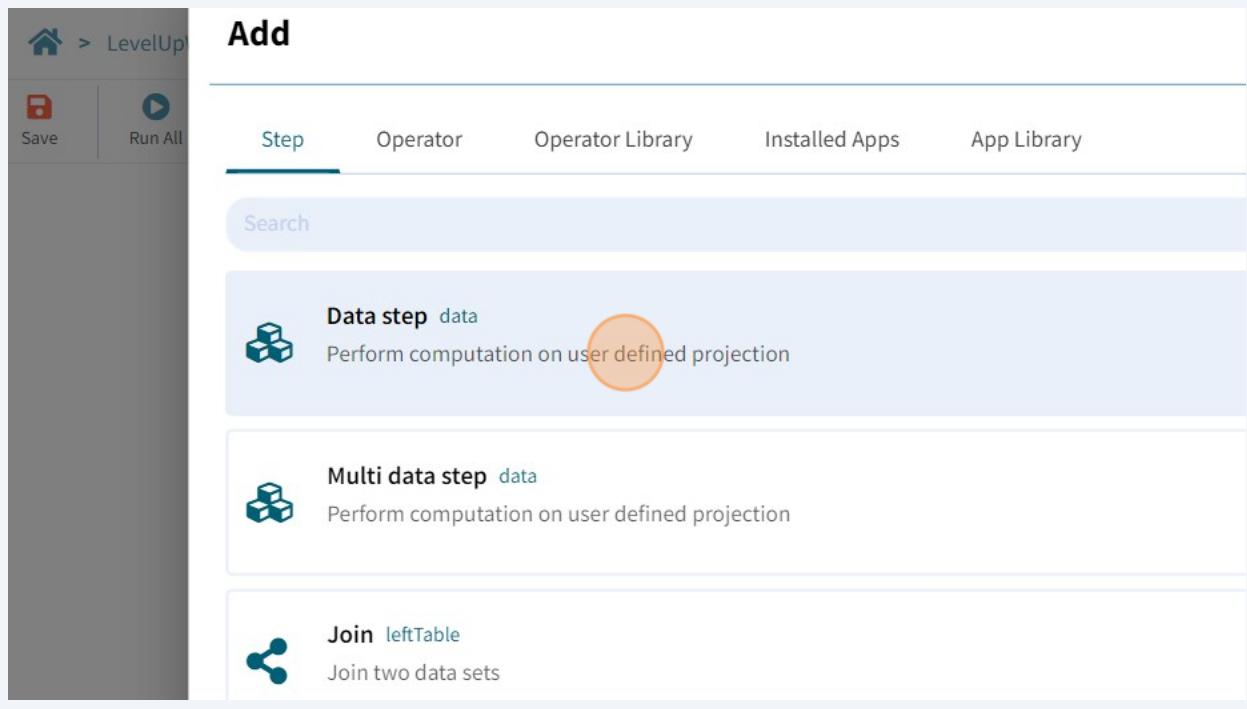
## 10 Click Add

shopsTeam > Workshop 1 - Martin > immune\_analysis



- 11** Choose **Data step** from the list.

A **Data step** is the main engine-block that performs the computations and visualisations in **Tercen**.



The screenshot shows the Tercen interface with the 'Add' screen open. The top navigation bar includes a home icon, a 'LevelUp' button, a 'Save' button, and a 'Run All' button. Below the navigation is a search bar labeled 'Search'. The main content area is titled 'Add' and has tabs for 'Step', 'Operator', 'Operator Library', 'Installed Apps', and 'App Library'. The 'Step' tab is selected. Three data step options are listed:

- Data step data**: Description: Perform computation on user defined projection. This option is highlighted with an orange circle around its icon and text.
- Multi data step data**: Description: Perform computation on user defined projection.
- Join leftTable**: Description: Join two data sets.

## 12 The **Data step** screen will open automatically.

In the left panel, you will see the column names from your data file.

These have been converted to **Factors** in Tercen

Factors are used to make projections.

Those projections are then used to submit data for computation or visualisation.

Click **VALUE** and drag it out onto the screen.

The screenshot shows the Tercen software interface in the 'Data step' mode. The top navigation bar includes 'Search Tercen', 'Learn', 'Explore', and a red user icon. Below the navigation is a toolbar with icons for Save, Add Operator, Crosstab, Tables, Layer 1, Point, Transform..., Filters, Colors, Labels, Error Bar, Clear, Legend, and Download. A sub-toolbar below the main toolbar has buttons for Factors, Environment, and Settings. On the left, a 'Factors' panel lists columns from a CSV file named 'grifols\_data\_final.csv'. The listed columns include SUBJECT ID, AGE, AGEGR1N, AGEGR1C, RACEGRP, SEXN, FASFL, SEXC, Batch date, PANEL, CELL SUBSET, Unit, Parent, and VALUE. The 'VALUE' column is highlighted with a green background and circled with a red oval. An orange arrow points from the 'VALUE' entry in the Factors panel towards the main workspace area. The main workspace is currently empty.

**13** You will see areas go green to indicate where the factor can be dropped.

The screenshot shows a software interface for data analysis. At the top, there's a navigation bar with 'Search Tercen' and a breadcrumb trail: 'n > Workshop 1 - Martin > cytokine\_analysis > Data step'. Below the navigation is a toolbar with various icons: 'Rosstab' (selected), 'Tables', 'Layer 1', 'Point', '4', 'Transform...', 'Filters', 'Colors', 'Labels', 'Error Bar', 'Clear', and 'Legend'. A large central area contains a grid with three columns and two rows. The first column is labeled 'Row', the second is 'Y-Axis', and the third is 'X-Axis'. The top row is labeled 'Column' and contains a single cell with the text 'Column'. The bottom row is labeled 'X-Axis' and also contains a single cell. The entire grid has a light green background, indicating it is a drop zone for data factors. On the far left, there's a sidebar with some text and a small upward arrow icon.

## 14 This is the **Crosstab Grid**

Factors are dragged and dropped here to make projections.

There is a standard **X-Axis** and **Y-Axis** for graphing numbers.

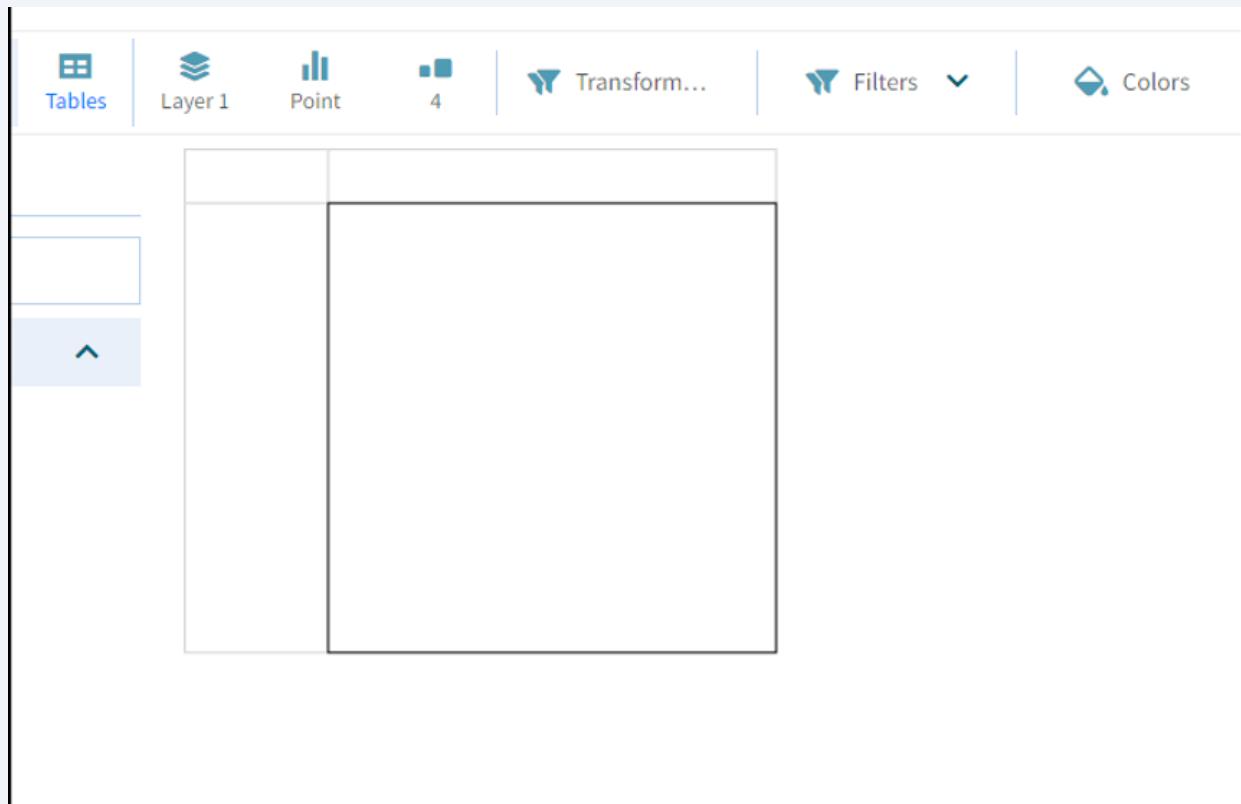
And there are **Rows** and **Columns** to make groupings of the data.

## 15 The central part of the Crosstab Grid is called the **Data Cell**

The Data Cell will split or merge based on how you make projections with factors.

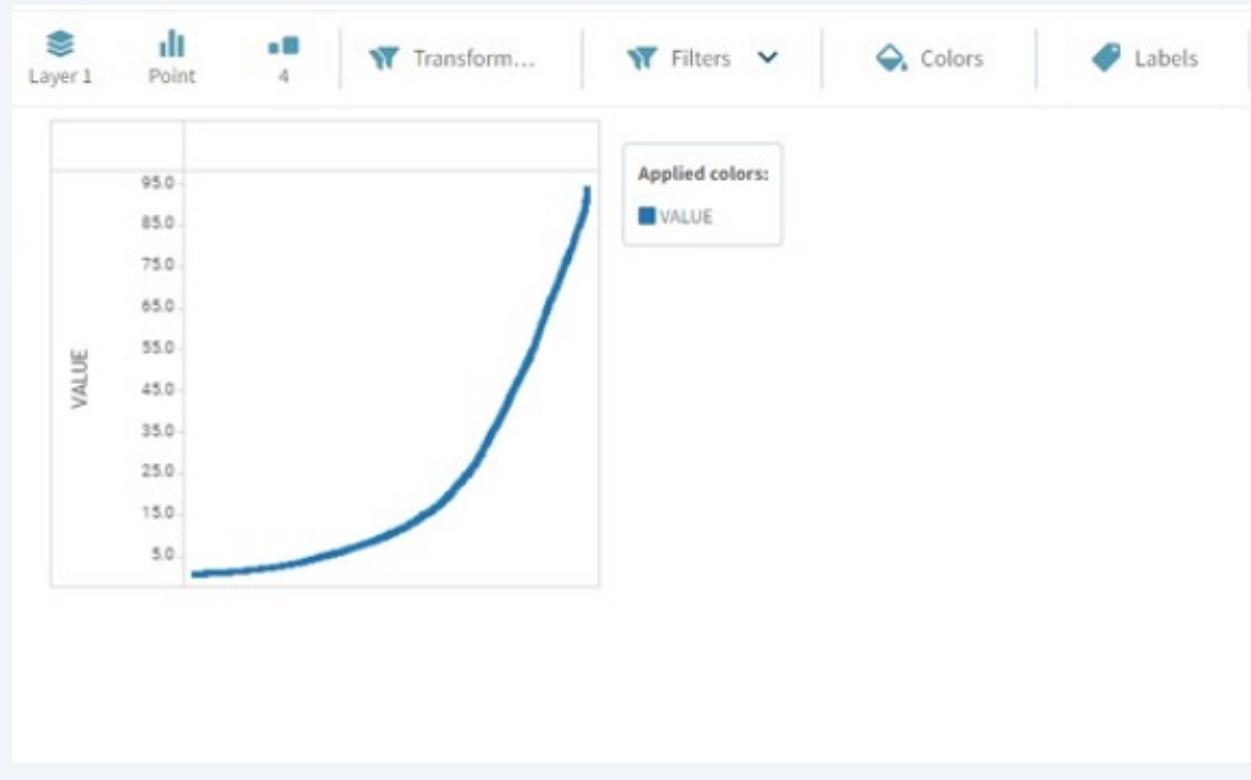
A Data Cell can hold one or more numerical values depending on how the projection is designed.

This is important because holding multiple data points or a single data point impacts how a visualisation looks or a calculation is performed.



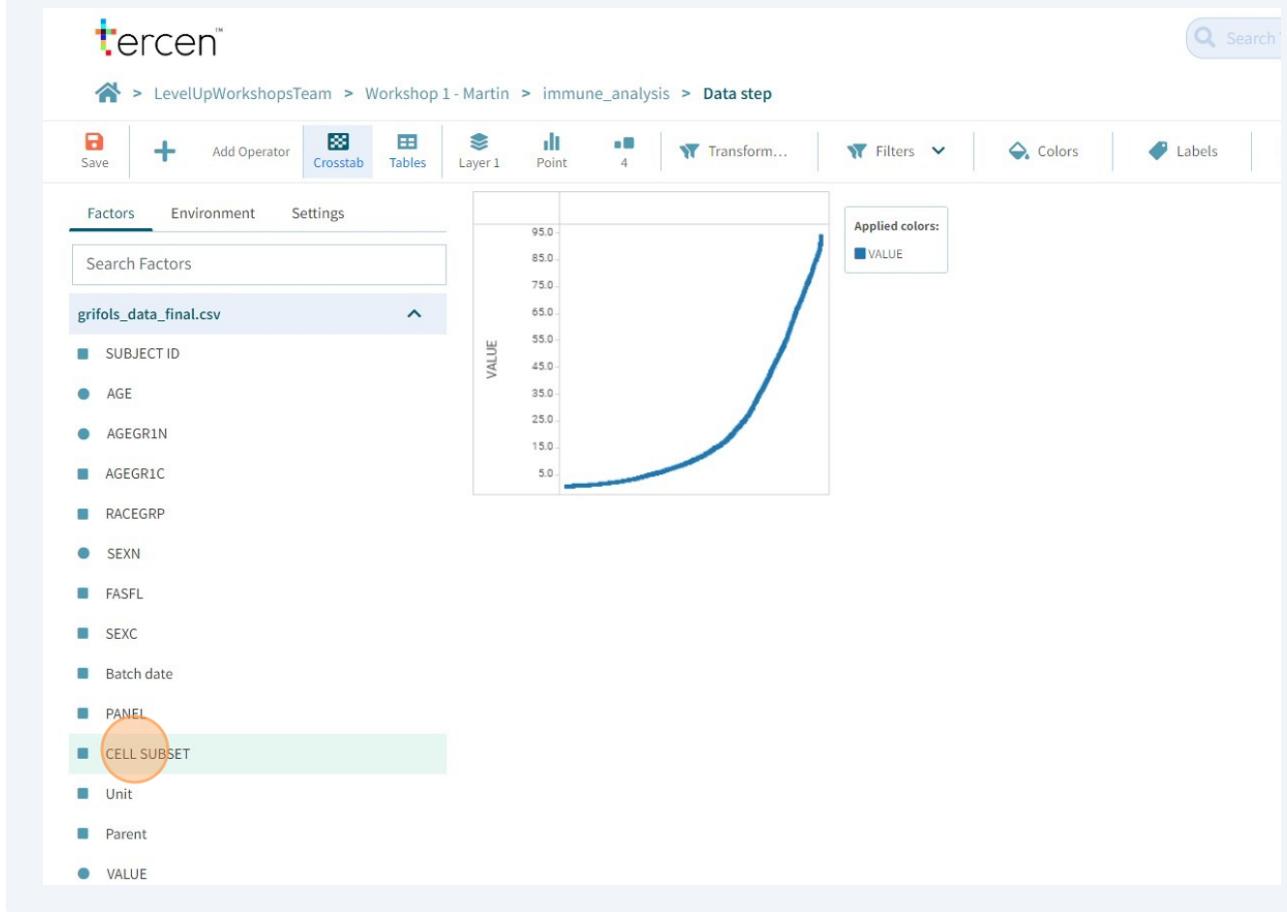
**16** Drop VALUE onto the Y-Axis.

The Data Cell will automatically order the measurements from lowest to highest.



We recommend you start every projection by dropping your primary measurement value onto the Y-Axis.

**17** Drag "CELL SUBSET" and drop it onto the X-Axis.



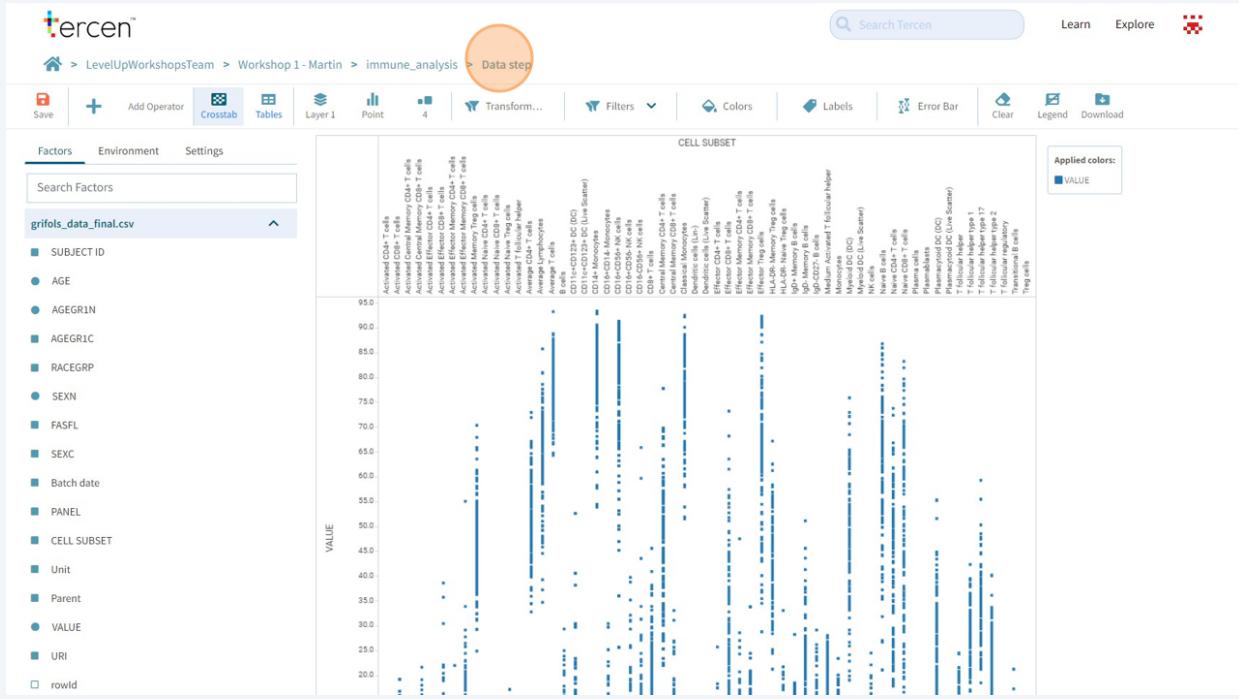
18

You have made a simple projection on the X and Y Axes.

The Data Cell holds all measurement values and lists them by their Cell Type.

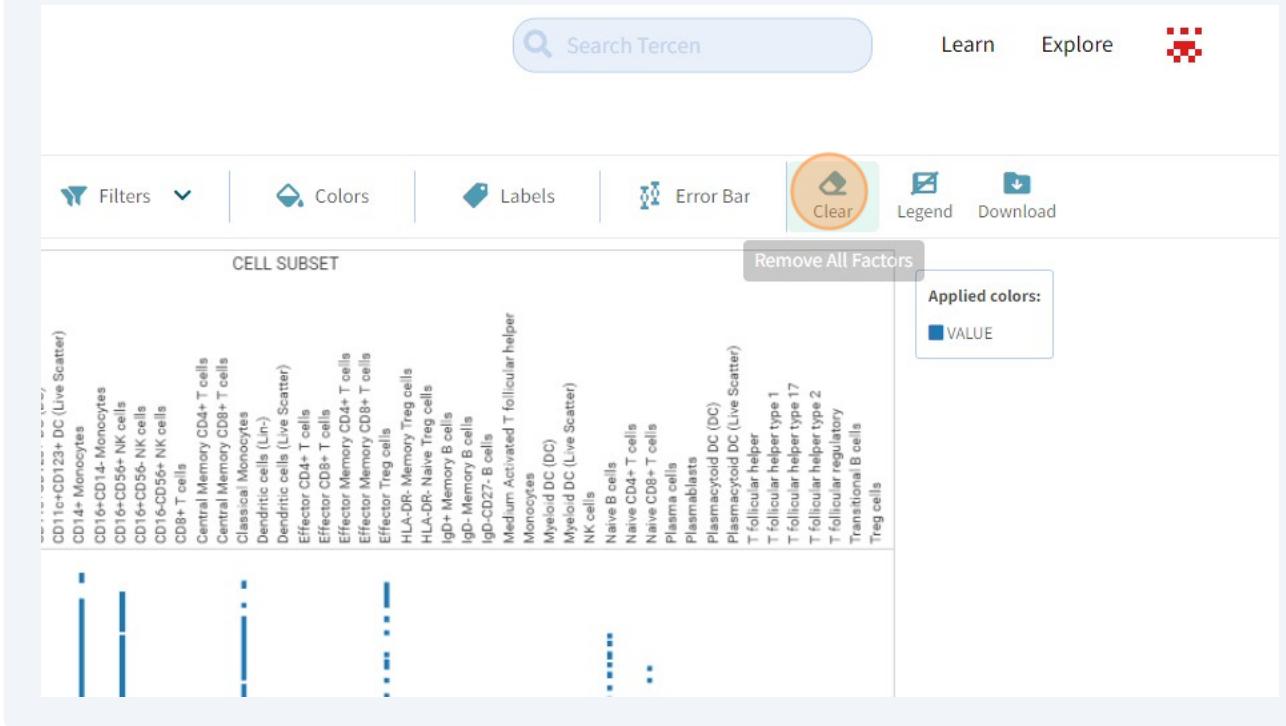
Click on the black lines of the Crosstab Grid and drag them out or in to adjust how the graph looks.

Hold down the CTRL on your keyboard and scroll with your mouse wheel to Zoom in and out of the Crosstab Grid.

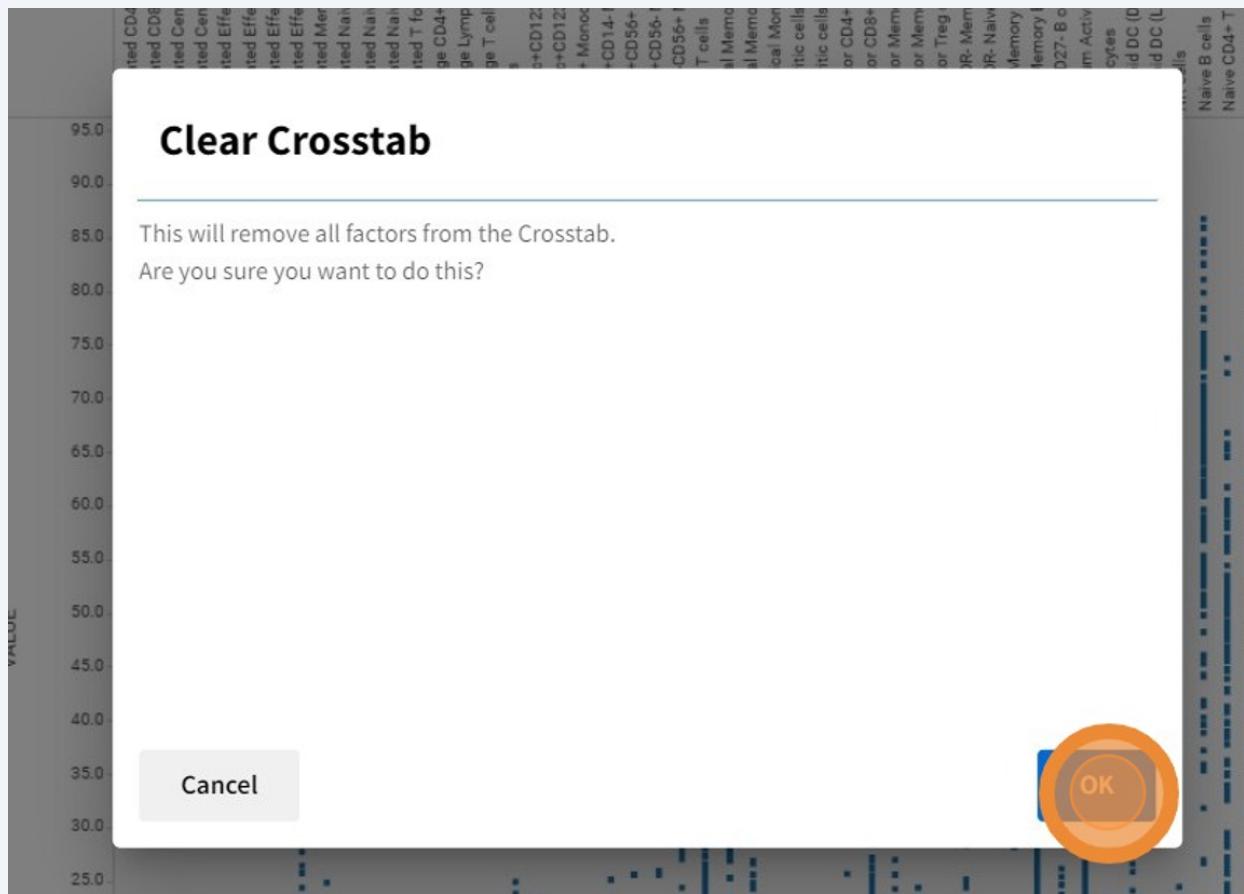


## 19 Now, let's make a more complex projection

Click the **Clear** button.



20 Click "OK"

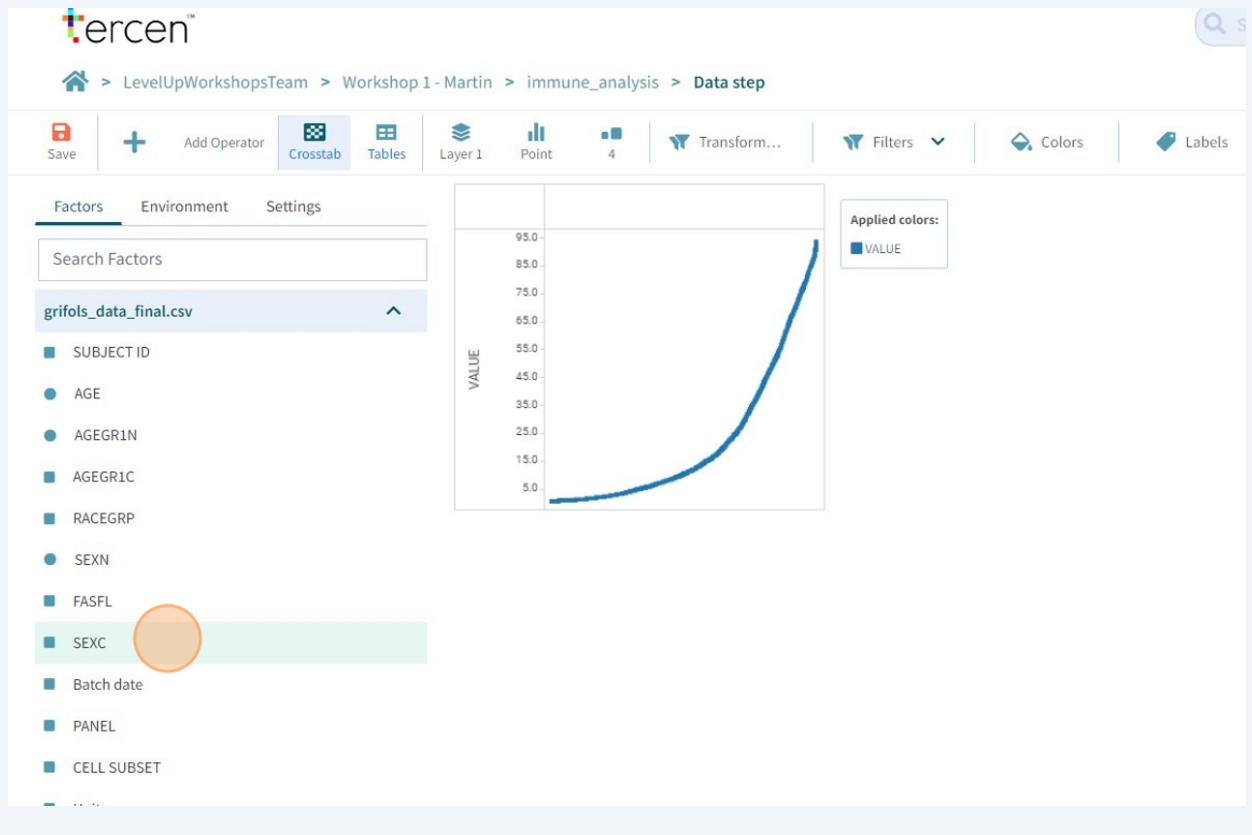


## 21 Drag VALUE to Y-Axis

The screenshot shows the Tercen software interface with the following details:

- Header:** Tercen™, Search Tercen, Learn, Explore, Help.
- Breadcrumbs:** Home > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Data step.
- Toolbar:** Save, Add Operator (+), Crosstab, Tables, Layer 1, Point, 4, Transform..., Filters, Colors, Labels, Error Bar, Clear, Legend, Download.
- Left Panel (Factors):** Factors, Environment, Settings. Search Factors: grifols\_data\_final.csv. Variable list:
  - SUBJECTID
  - AGE
  - AGEGRIN
  - AGEGRIC
  - RACEGRP
  - SEXN
  - FASFL
  - SEXC
  - Batch date
  - PANEL
  - CELL SUBSET
  - Unit
  - Parent
  - VALUE** (highlighted and circled in orange)
  - URI
  - rowId
  - tableId
- Right Panel:** A large empty rectangular area representing the target for dragging the 'VALUE' variable.
- Bottom:** Width: 330px (8.73cm) | Height: 290px (7.67cm) | dpi: 96.

## 22 Drag SEXC to Column



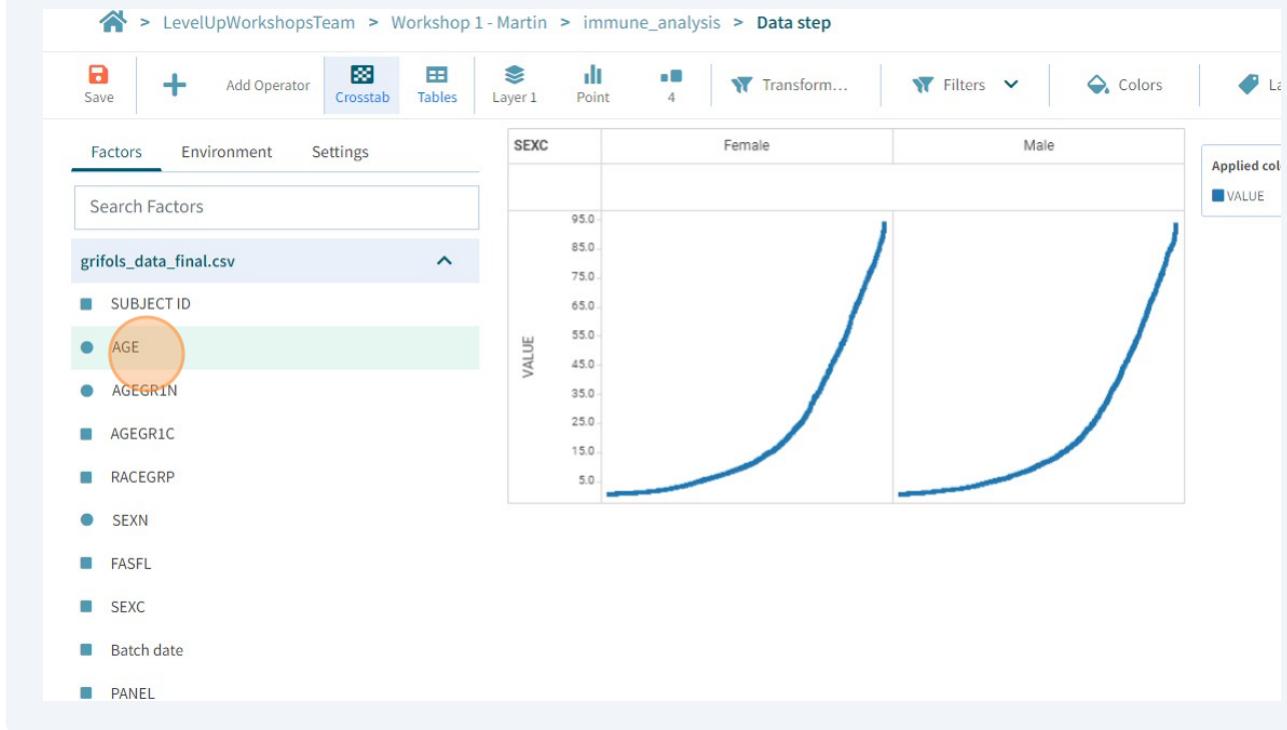
23

The Crosstab Grid has grouped the VALUE into two Data Cells with a separated Y-Axis.

Male and Female

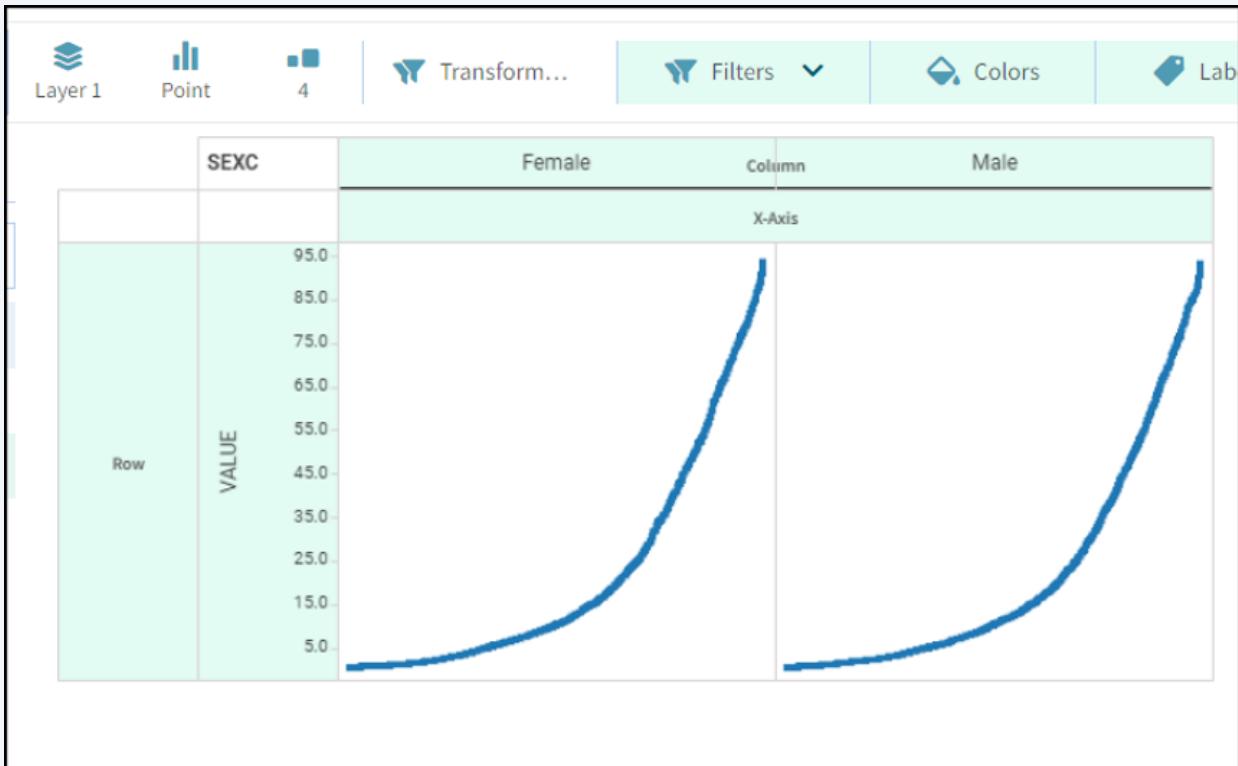
Drag "AGE" out to the Crosstab Grid, but don't drop it yet.

Instead, hover above and below the SEXC Column on the Crosstab Grid.



**24** A black line indicates that you can drop AGE above or below the existing column.

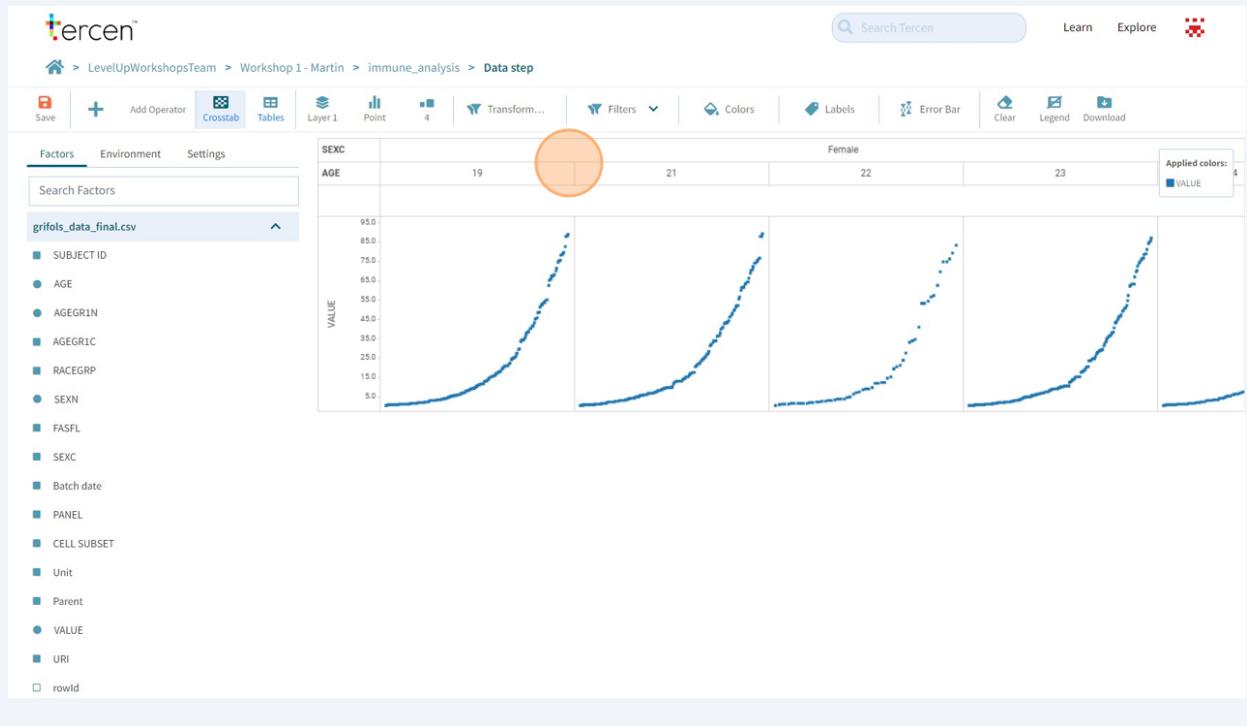
Drop AGE below the SEXC column.



**25** The Crosstab Grid has now made a sub-grouping.

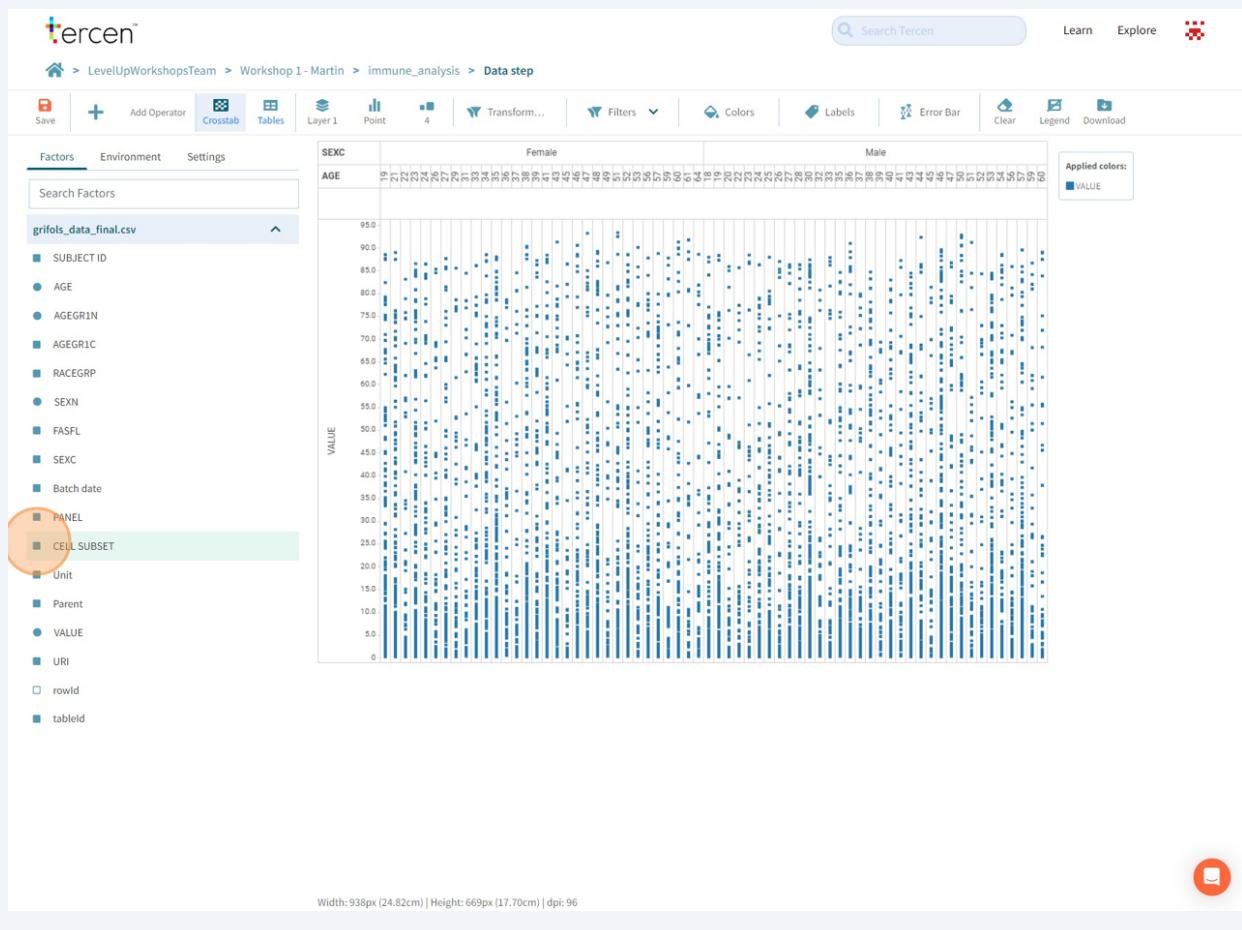
It has separated the Male and Female data by age.

Use the black lines to drag the data close together.

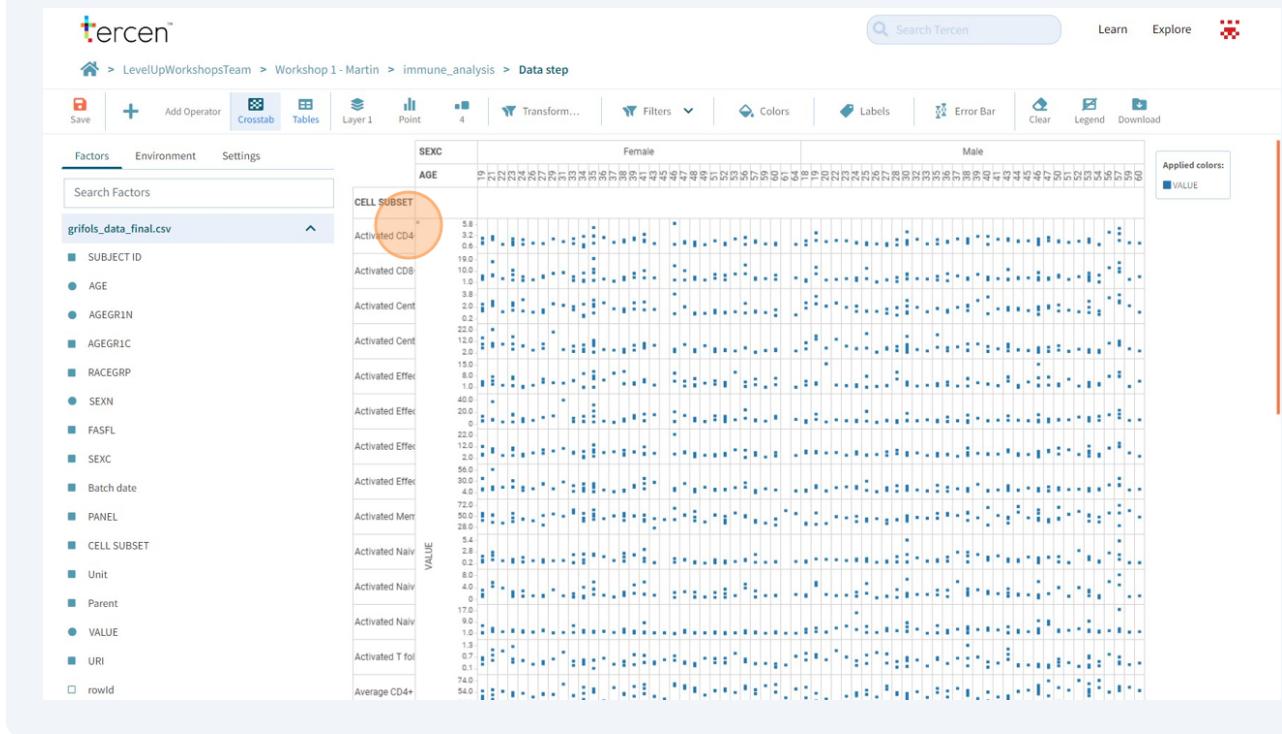


## 26 Up to 10 Columns and Rows can be stacked in a projection.

Drag "CELL SUBSET" to Row



**27** The Crosstab is now a multivariate grid with many Data Cells.

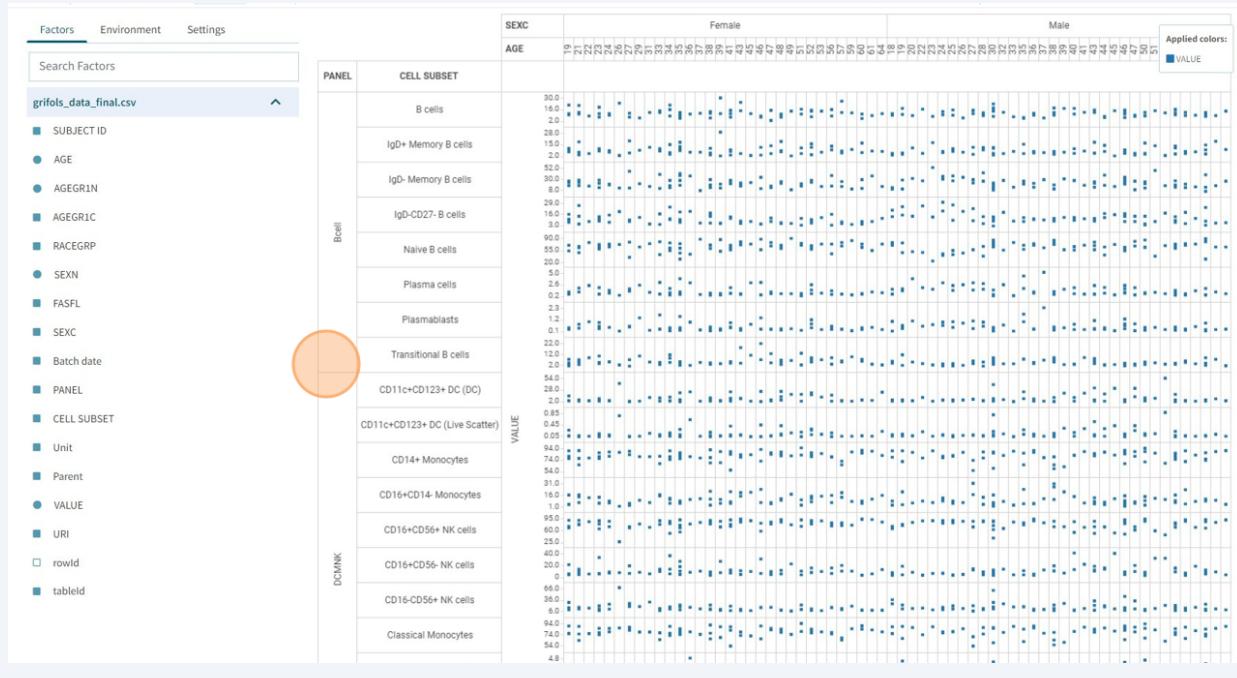


**28** Drag "PANEL" to the black line before CELL SUBSET (on the left).

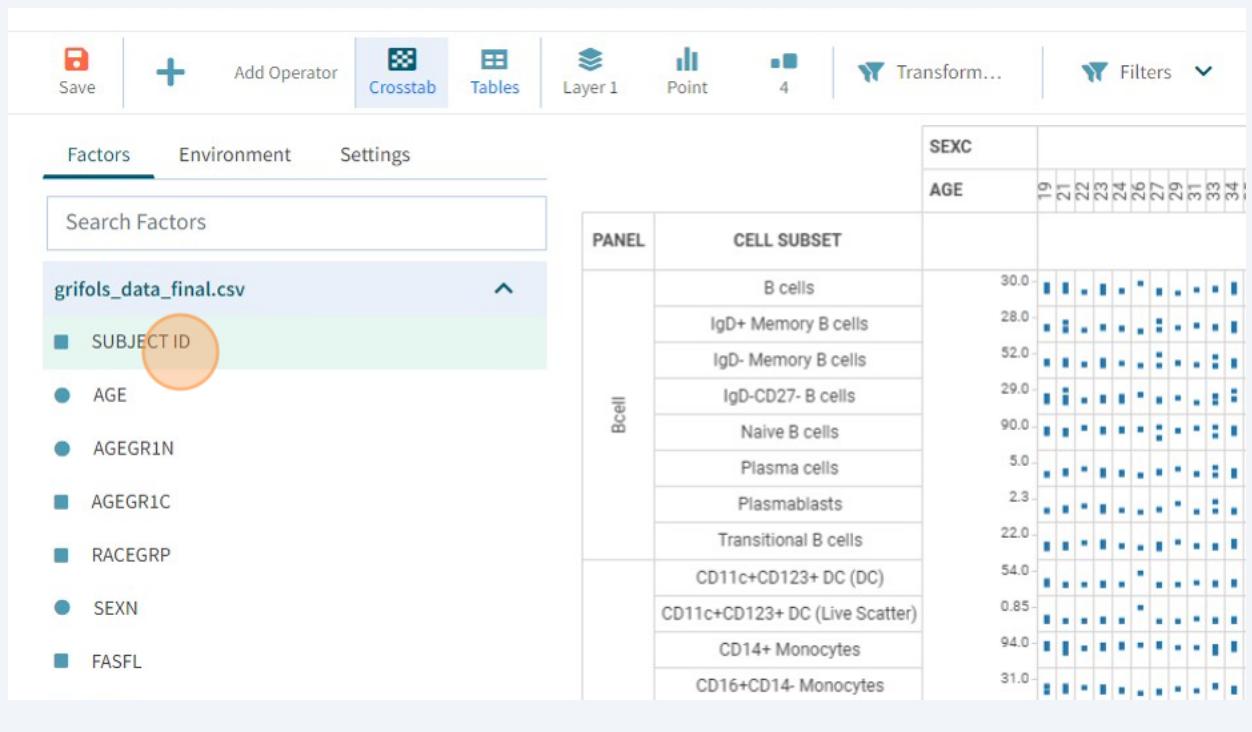
Groupings on a Crosstab projection work from **Outside to In**

For Columns, the topmost factor will be grouped first, and then sub-groups are made in order of what is below it.

For rows, the left factor is grouped first, and then subsequent factors work to the right.

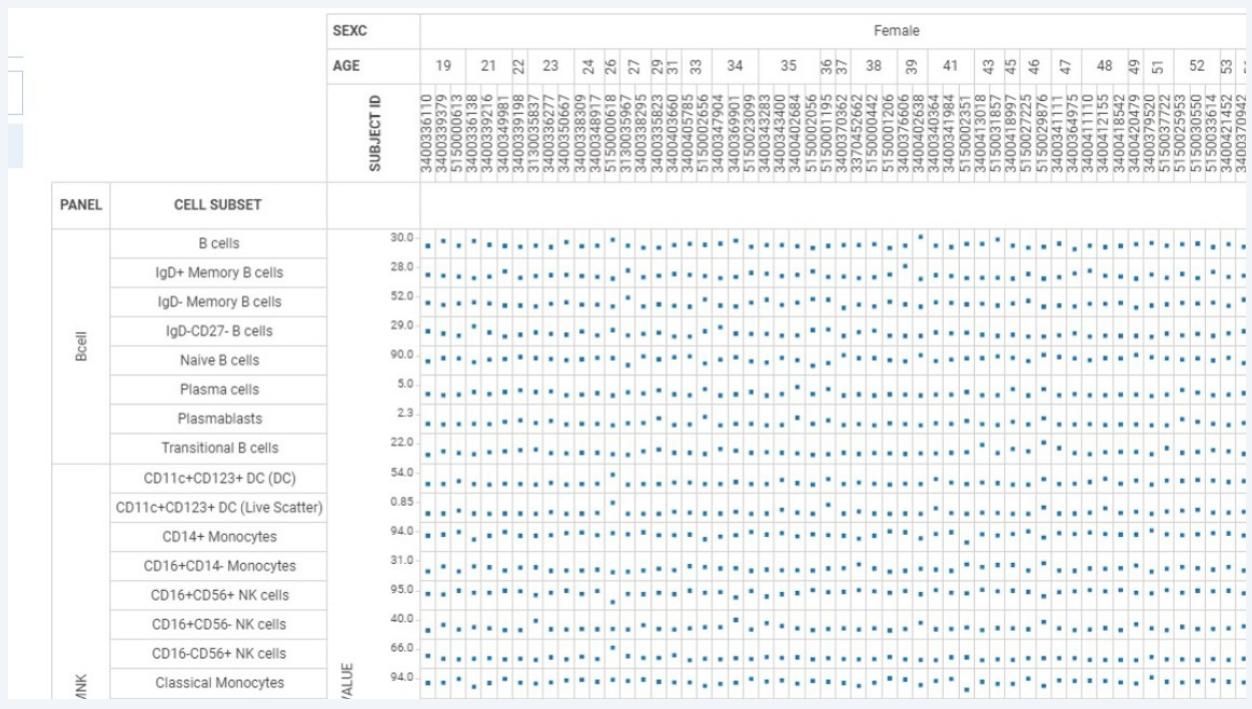


**29** Drag SUBJECT ID to the column, drop it below AGE,



**30** Now, the Crosstab Grid has a data cell for every measurement.

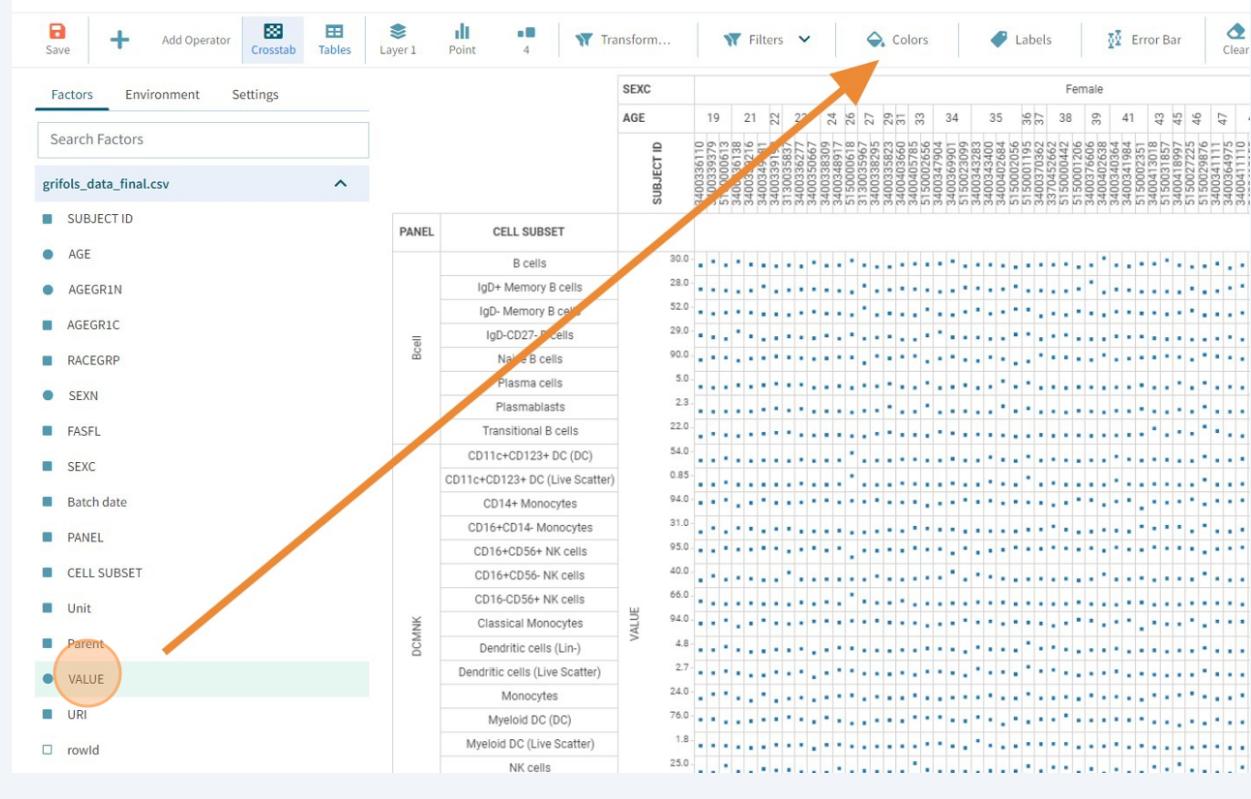
It shows a complete layout of the data.



31

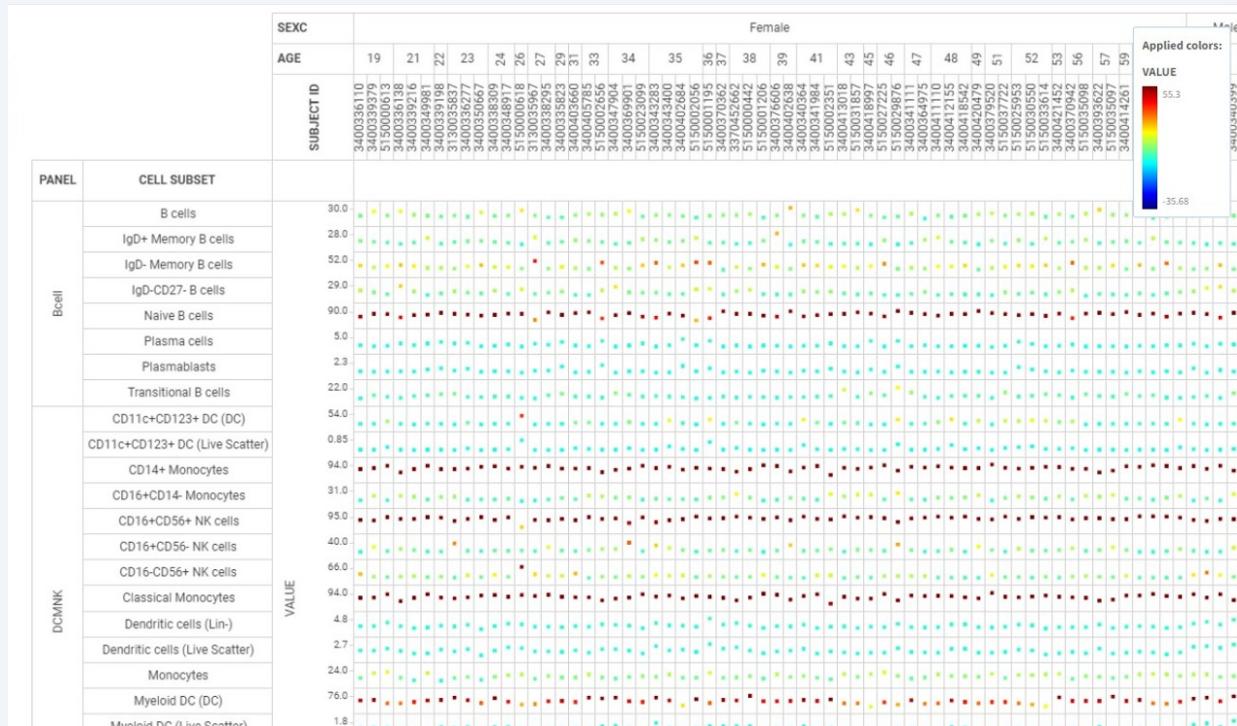
The top bar has settings that control how a visualisation looks or how a calculation is made.

Drag "VALUE" to Colors



32

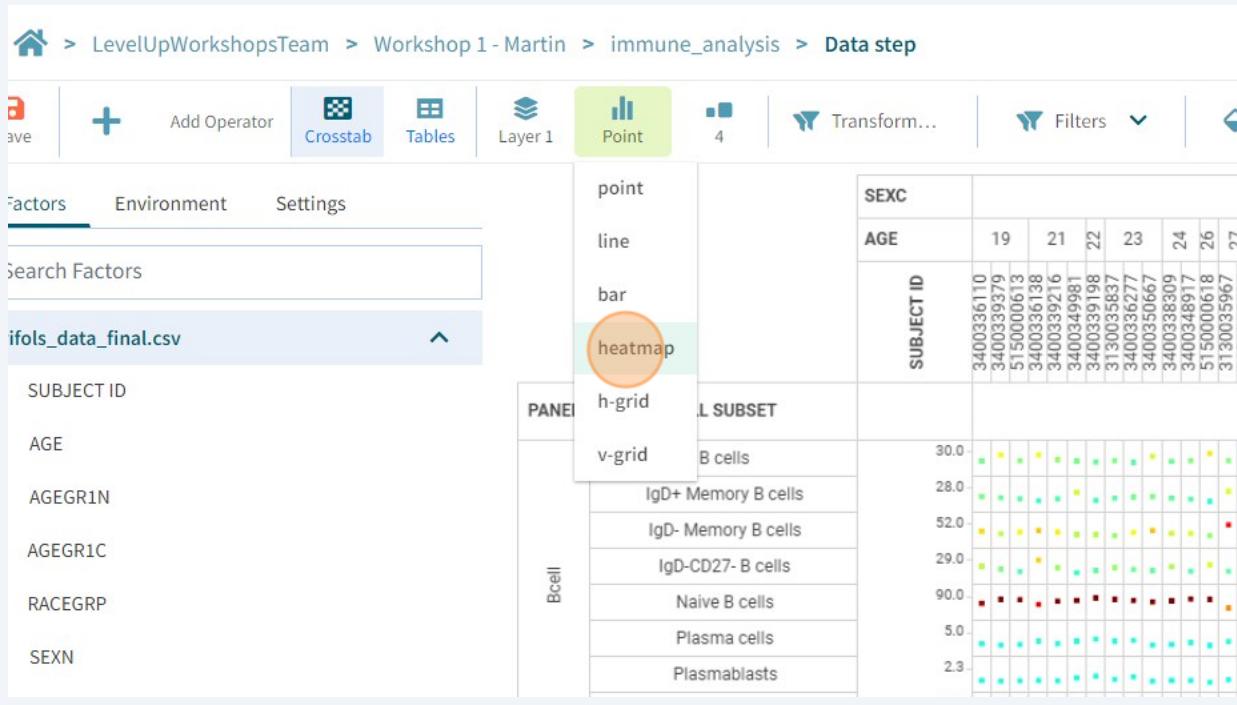
The values in the data cells now display color, with higher values trending toward red and lower values trending toward blue.



33

Click on the **Style** Button

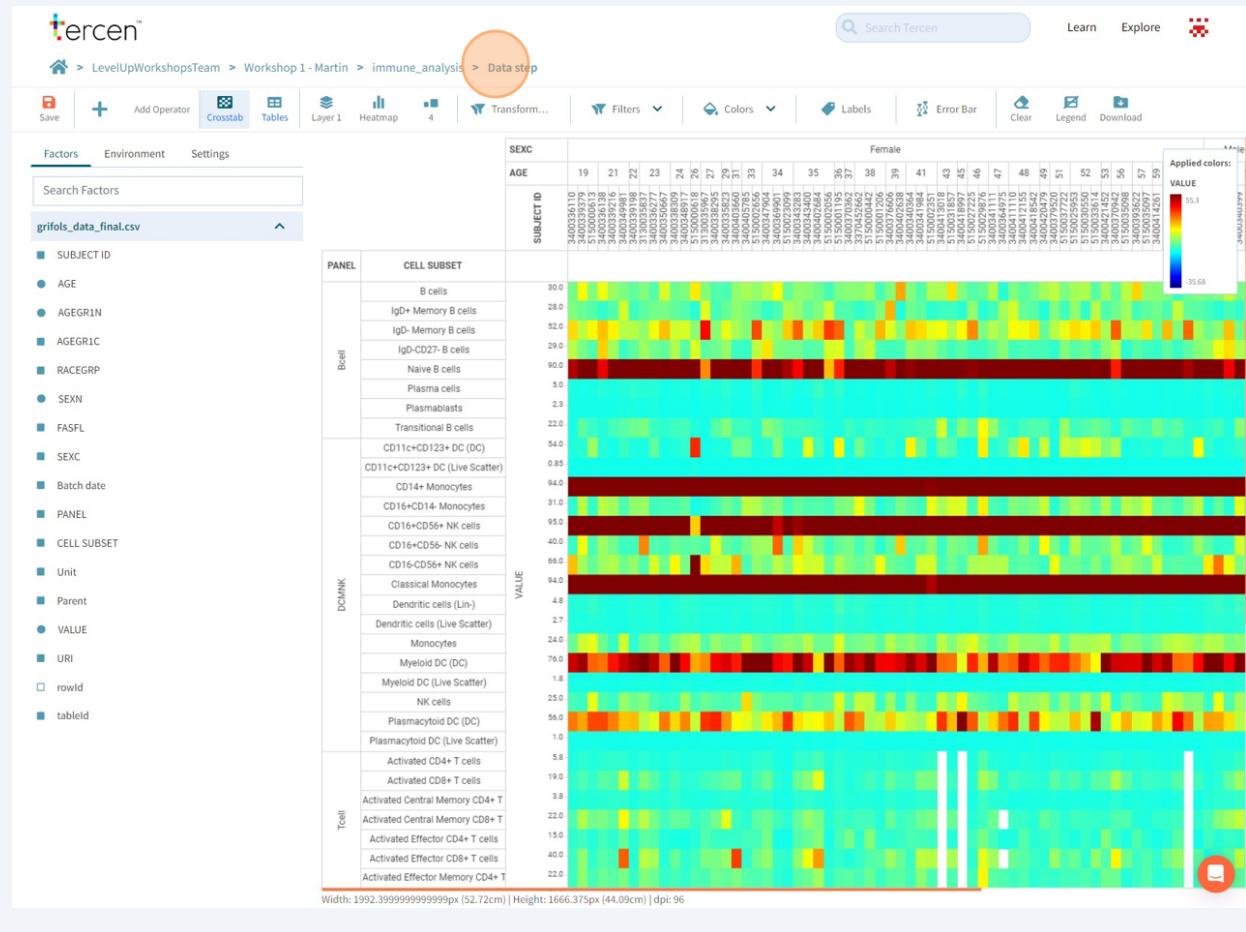
Change the graph style from **Point** to **Heatmap**



## 34 You have made a heatmap overview of the data.

Take a moment to move the black lines and zoom in to examine the data.

Try to look for immunological patterns.



35 Click "Save"

The screenshot shows the Tercen software interface. At the top, there is a navigation bar with the project path: Home > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Data step. Below the navigation bar is a toolbar with various icons: Save (highlighted with an orange circle), Add Operator, Crosstab, Tables, Layer 1, Heatmap, a numeric input field (4), Transform..., Filters, and a dropdown menu.

The main workspace has tabs for Factors, Environment, and Settings, with Factors selected. A search bar labeled "Search Factors" is present. Below the search bar is a file list containing "grifols\_data\_final.csv".

On the right side, there is a detailed view of the data. It includes a table with columns for SEXC, AGE, and SUBJECT ID, and a heatmap titled "Bcell" showing cell subsets like B cells, IgD+ Memory B cells, IgD- Memory B cells, IgD-CD27- B cells, Naive B cells, Plasma cells, Plasmablasts, Transitional B cells, CD11c+CD123+ DC (DC), and CD11c+CD123+ DC (Live Scatter). The heatmap uses a color scale from 0.85 (blue) to 30.0 (red).

36 Return to the workflow screen by clicking **immune\_analysis** in the project breadcrumb.

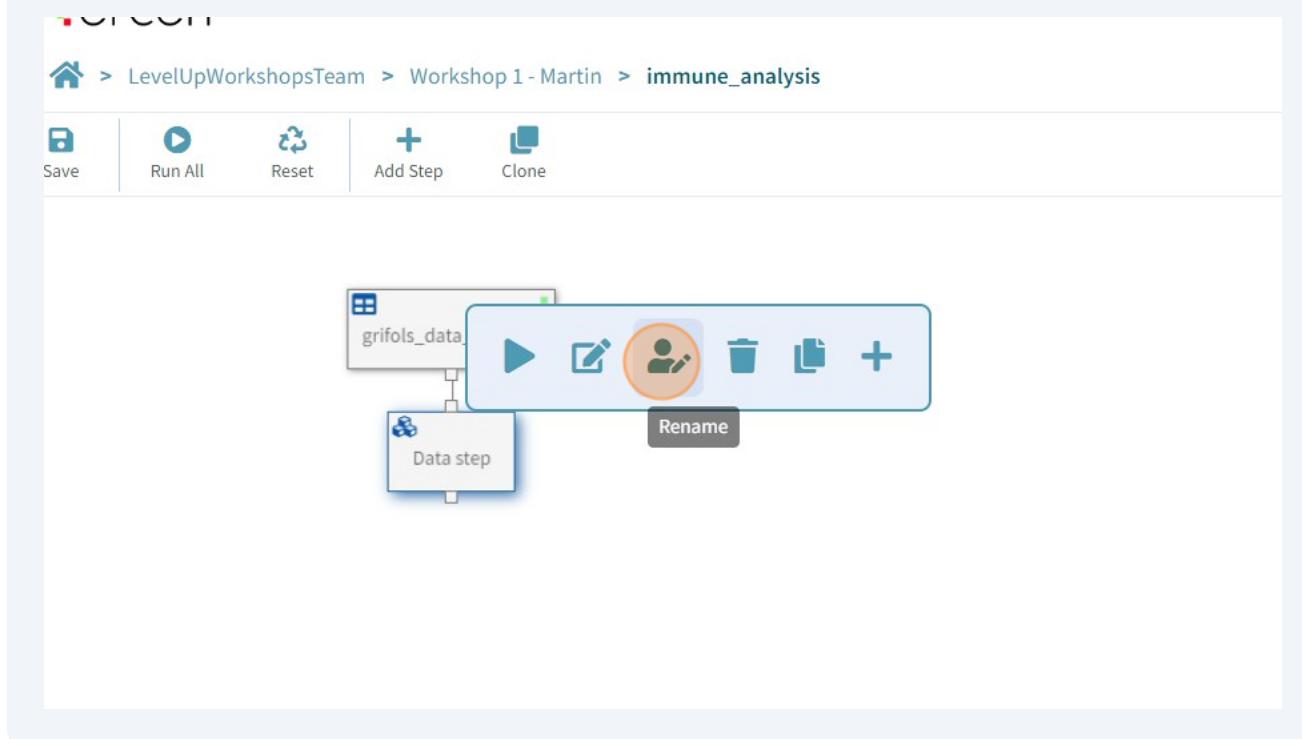
The screenshot shows the Tercen software interface on the workflow screen. The navigation bar at the top shows the project path: Home > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Data step. The "immune\_analysis" part of the path is highlighted with an orange circle.

The toolbar below the navigation bar is identical to the one in the previous screenshot, featuring icons for Add Operator, Crosstab, Tables, Layer 1, Heatmap, a numeric input field (4), Transform..., Filters, and a dropdown menu.

The main workspace shows the same "Factors" tab selected, with a search bar and a file list containing "grifols\_data\_final.csv".

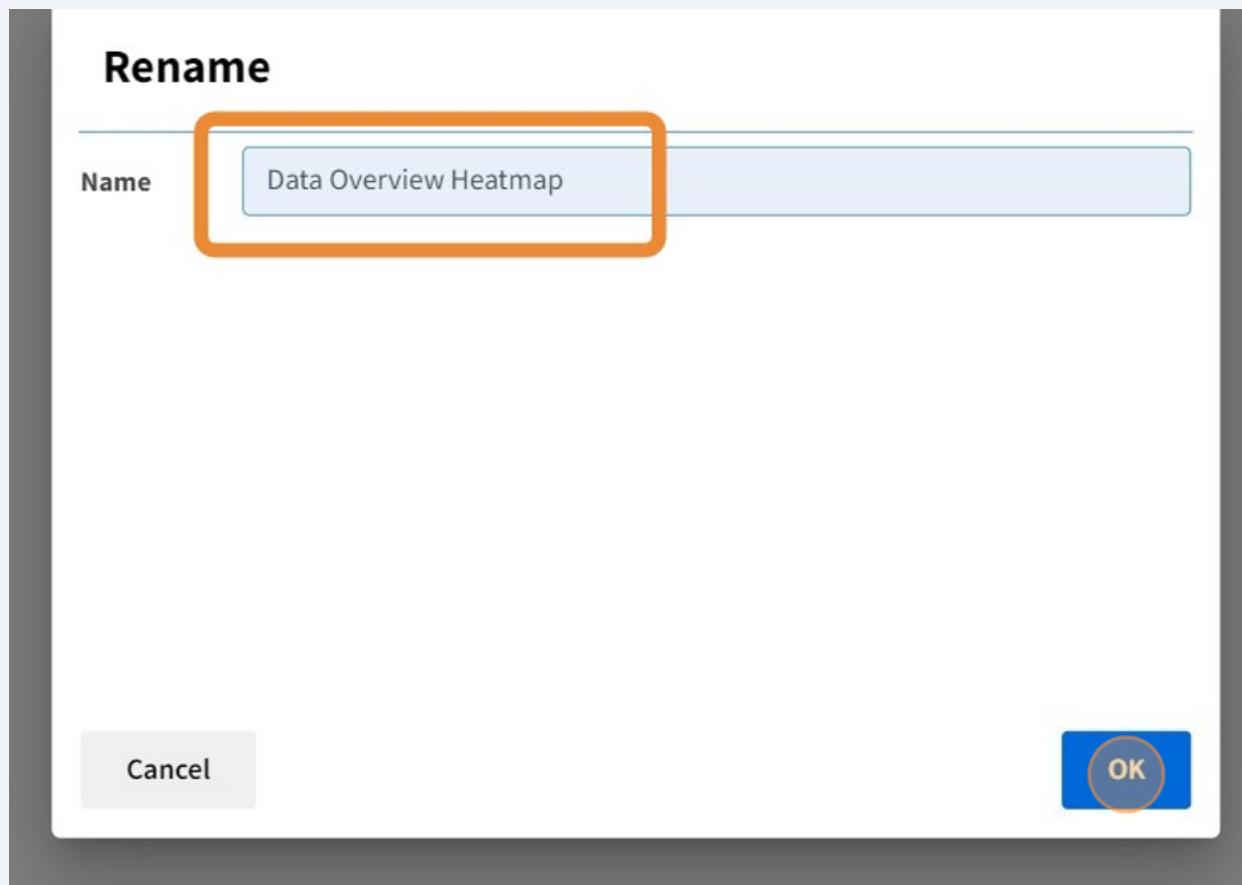
On the right side, the detailed view of the data is identical to the previous screenshot, showing the "Bcell" panel with its corresponding table and heatmap.

- 37** Click the rename button in the local toolbar.

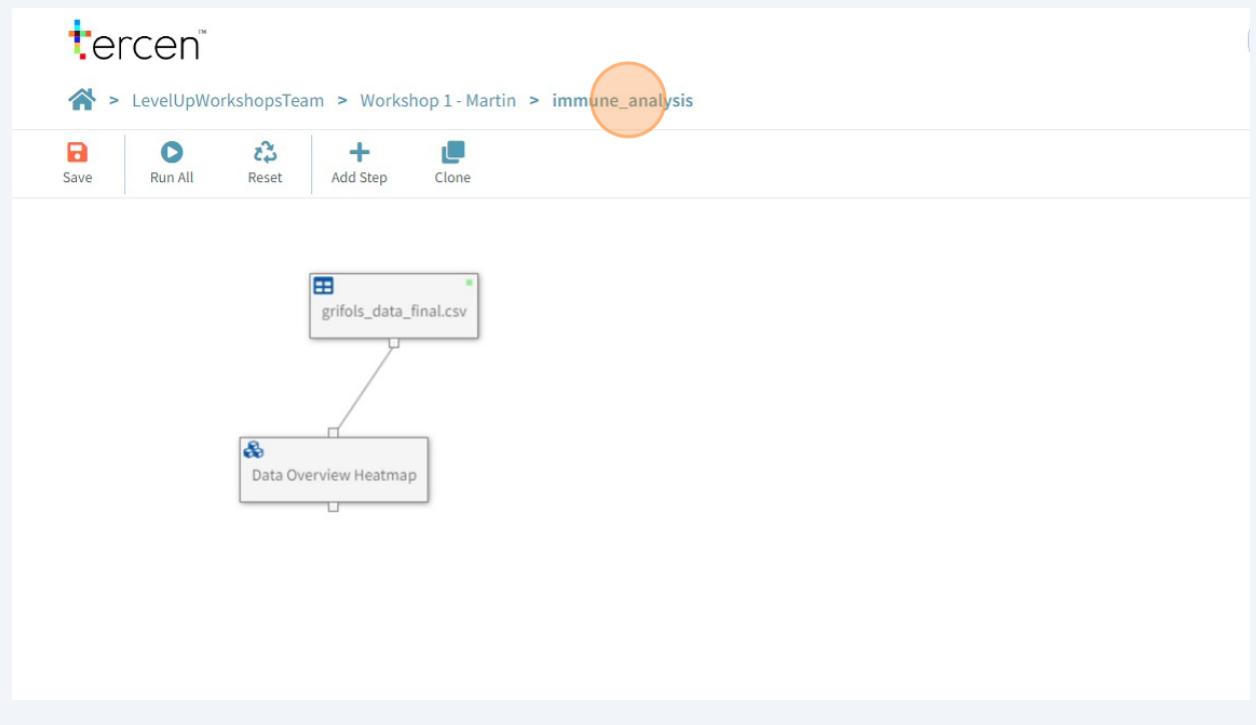


- 38** Name your Heatmap "Data Overview Heatmap"

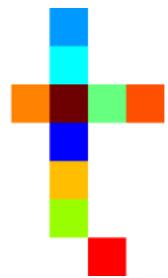
Click Ok



**39** Congratulations! You have made a heatmap visualisation in Tercen.



# 0105 - Calculate a Median



In Tercen, data manipulation can be done by an "Operator."

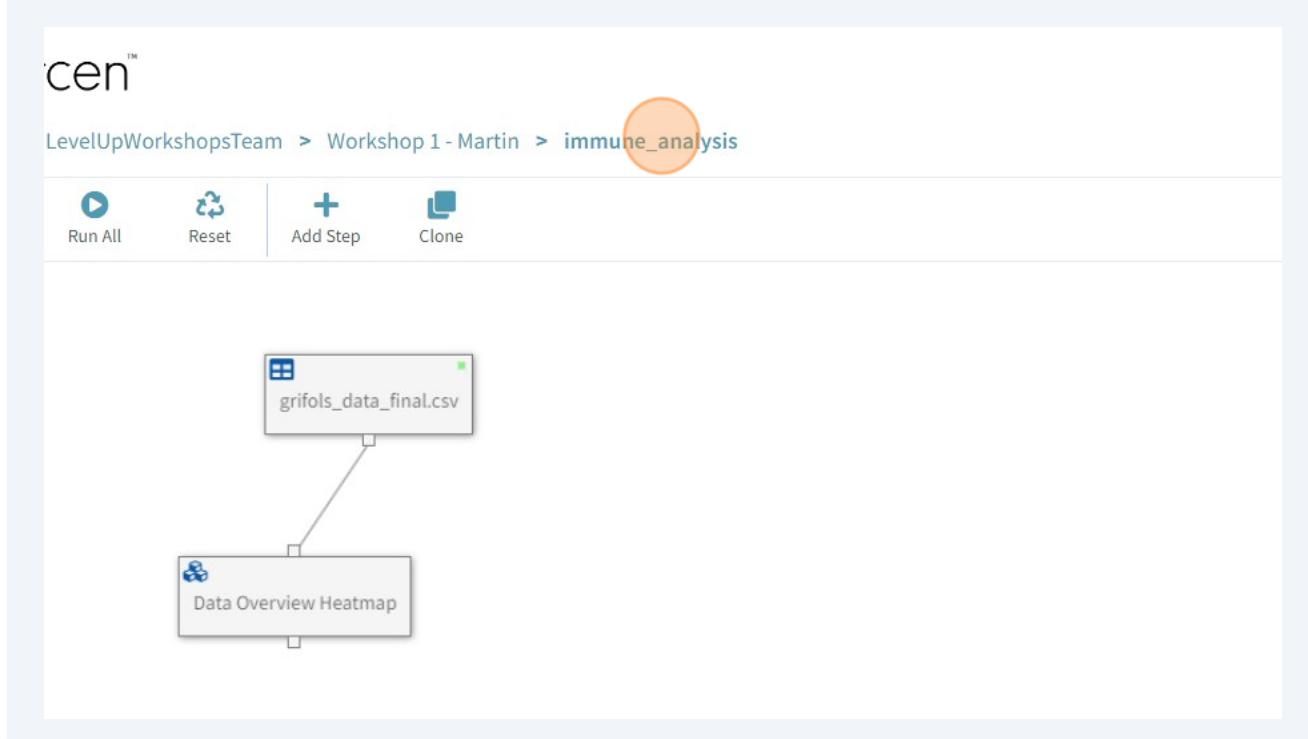
Operators are computer programs written in R or Python by a bioinformatician.

They can be scripts for simple calculations such as COUNT or MEDIAN.

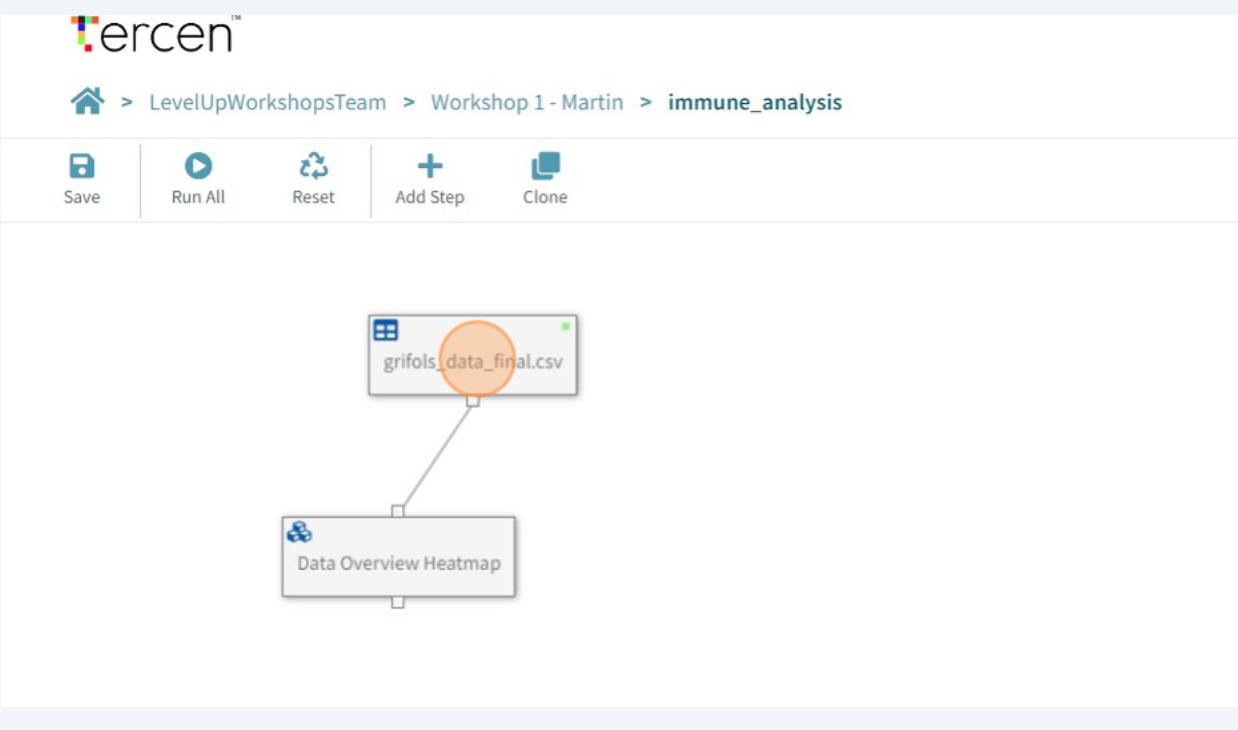
Or they can be complex programs with their control panels and specialised graphs.

In this tutorial, we will learn the basic concepts of Tercen operators and use one to make a simple calculation of the data.

- 1 Start from the **immune\_analysis** workflow canvas

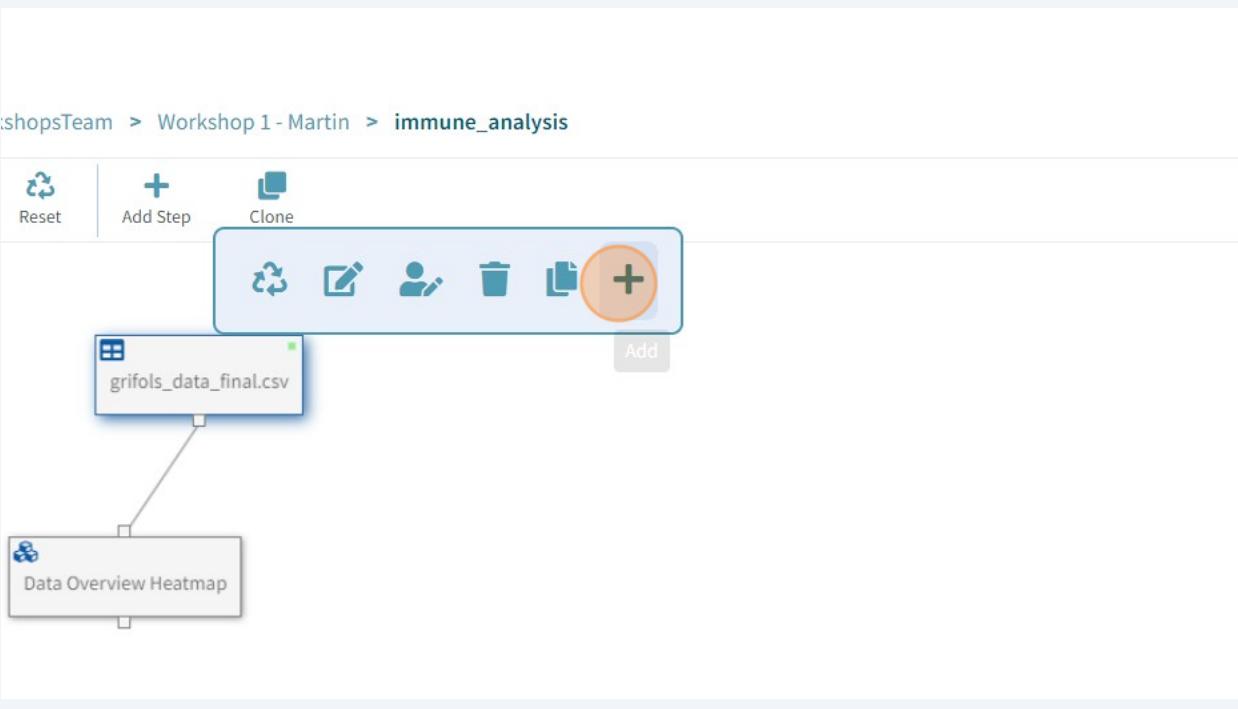


- 2 Click the **Table Step** "grifols\_data\_final.csv"



- 3 The local toolbar will open.

Click Add.



#### 4 Add a Data Step

The screenshot shows the 'Add' screen with the 'Step' tab selected. There are three items listed:

- Data step** data  
Perform computation on user defined projection
- Multi data step** data  
Perform computation on user defined projection
- Join** leftTable  
Join two data sets

#### 5 The new Data Step will open on screen.

Click the **+ button** to add an operator.

The screenshot shows the 'Data step' editor with the 'Factors' tab selected. The left sidebar lists variables from a file named 'grifols\_data\_final.csv':

- SUBJECT ID
- AGE
- AGEGR1N
- AGEGR1C
- RACEGRP

6 Click "Library"

The screenshot shows the tercen software interface. On the left, there's a sidebar with a 'Save' button, a plus sign button, and tabs for 'Factors' and 'Environ'. Below these are buttons for 'Search Factors' and a file named 'grifols\_data\_final.csv'. A list of variables is shown: SUBJECT ID, AGE, AGEGR1N, AGEGR1C, and RACEGRP. The main area is titled 'Choose an operator' with tabs for 'Installed', 'Library' (which is highlighted with an orange circle), and 'Log'. There's a 'Search' field with a magnifying glass icon. A checkbox for 'Display latest version only' is checked. Below this is a 'Tag list' section with two items: 'Bar plot 0.1.5' and 'Count 1.0.0'. Both items have small icons and tags below them.

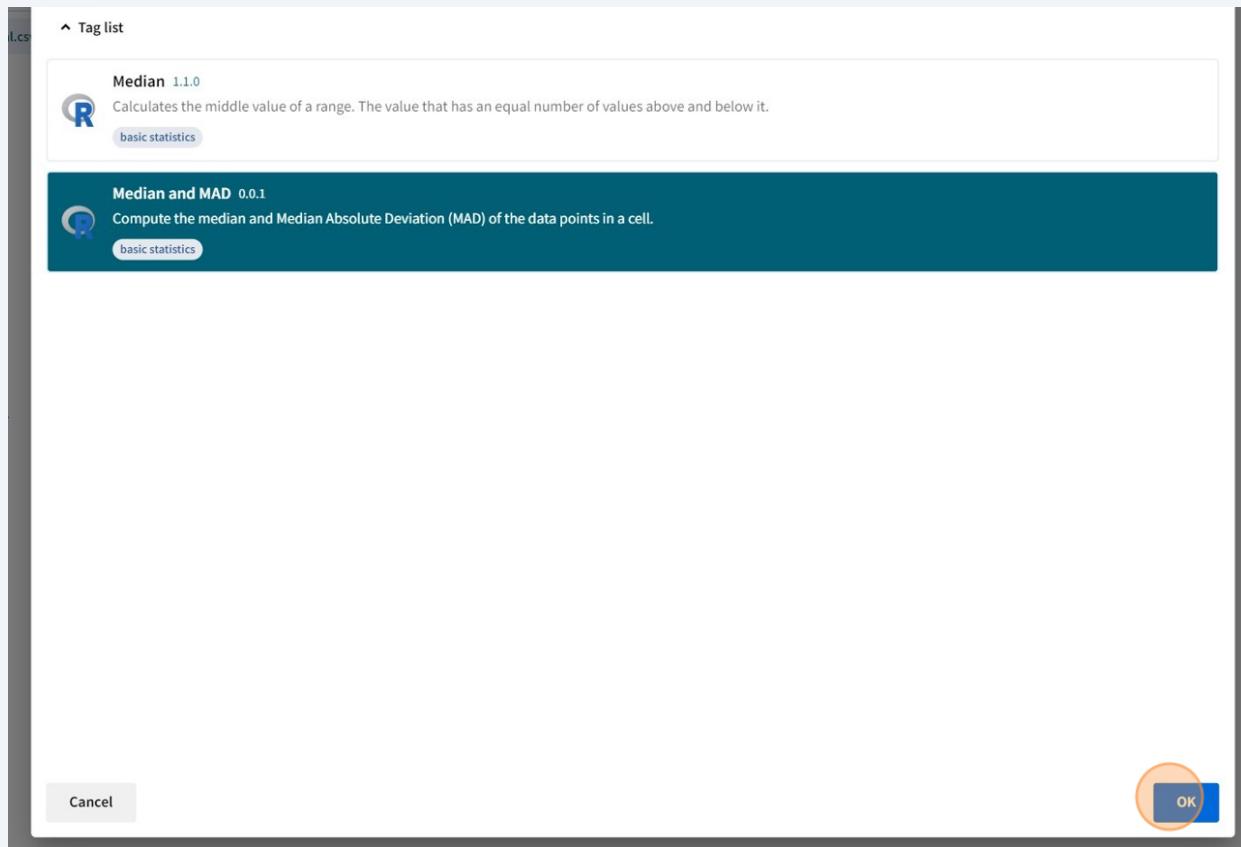
7 Click the **Search** field.

Type "Median"

The screenshot shows the same tercen software interface as the previous one, but with the 'Search' field in the top bar highlighted with an orange circle. The rest of the interface is identical to the previous screenshot.

- 8 Select the **Median and MAD** operator.

Click "OK"



9

Operators have requirements for how data is laid out in a projection before they can work properly.

To discover the "rules" an Operator expects you to follow, click on the version link under the Operator name.

Click "0.0.1"

The screenshot shows the Tercen software interface. At the top, there's a navigation bar: Home > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Data step. Below the navigation, there's a toolbar with various icons: Save, Remove, Median and MAD 0.0.1 (which is highlighted with an orange circle), Run, Crosstab (which is selected and highlighted with a blue background), Tables, Layer 1, Point, and four small square icons. To the right of the toolbar, there are Transform... and Filter icons. Below the toolbar, there are three tabs: Factors (selected), Environment, and Settings. A search bar labeled "Search Factors" is present. Under the Factors tab, a file named "grifols\_data\_final.csv" is listed. To the right of the file name is an upward-pointing arrow icon. Below the file name, there's a list of variables: SUBJECT ID, AGE, AGEGR1N, and AGEGR1C. The main workspace area is currently empty, showing a grid structure.

**10** Your browser will open the GitHub page of the operator.

This is where the computer code "lives".

The screenshot shows a GitHub repository page for the 'median mad operator'. The repository is public and was generated from [tercen/template-R-operator-lite](#). It has a version of 0.0.1, one branch, and one tag. The repository contains files for .github/workflows, renv, tests, and .Rprofile. A commit by agouy is shown with the commit ID 839b986. The repository has 839b986 commits and 12 branches.

File/Folder	Commit
.github/workflows	first version
renv	first version
tests	first version
.Rprofile	first version
...	...

**11** The programmer will have left instructions for the data projection.

This is called the **Input Projection**.

For Median and MAD, all that is required is that a numeric factor is put on the Y-Axis.

It will calculate a Median and MAD for every data cell it sees.

The last thing to note is the **Output relations**.

These are **New Factors** that the operator will create containing the results of its calculations.

The screenshot shows a portion of a README.md file. At the top, there's a navigation bar with icons for back, forward, and search. Below that is a header section with a title and a subtitle. The main content area contains two tables: one for 'Input projection' and one for 'Output relations'. The 'Input projection' table has two rows: 'Input projection' (empty) and 'y-axis' (type: numeric, measurement). The 'Output relations' table has two rows: 'Output relations' (empty) and two entries under 'median': 'median' (type: median, per cell) and 'mad' (type: median absolute deviation, per cell).

Input projection	
y-axis	numeric, measurement

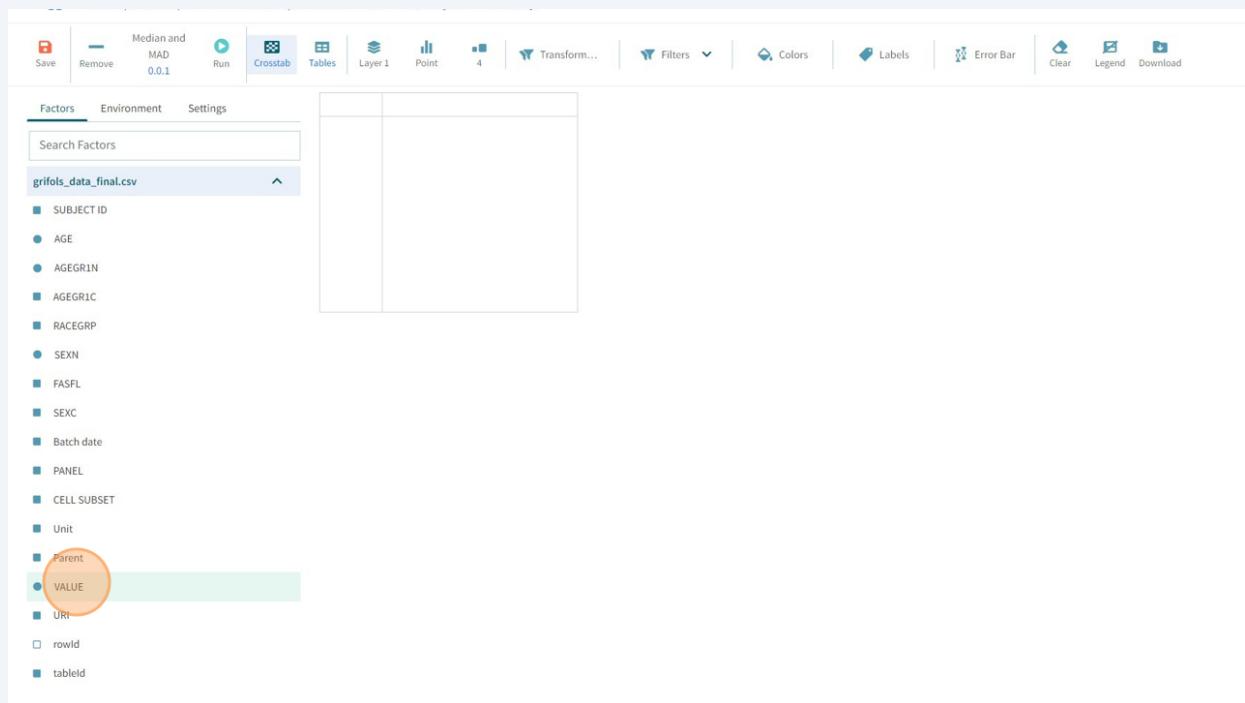
Output relations	
median	median, per cell
mad	median absolute deviation, per cell

**i** Operators can calculate from a single Data Cell, a whole row, or the entire crosstab grid, depending on their design.

You can find this information in the Description. Be careful though; when a programmer says "cell" they usually mean a "data cell" not a biological cell.

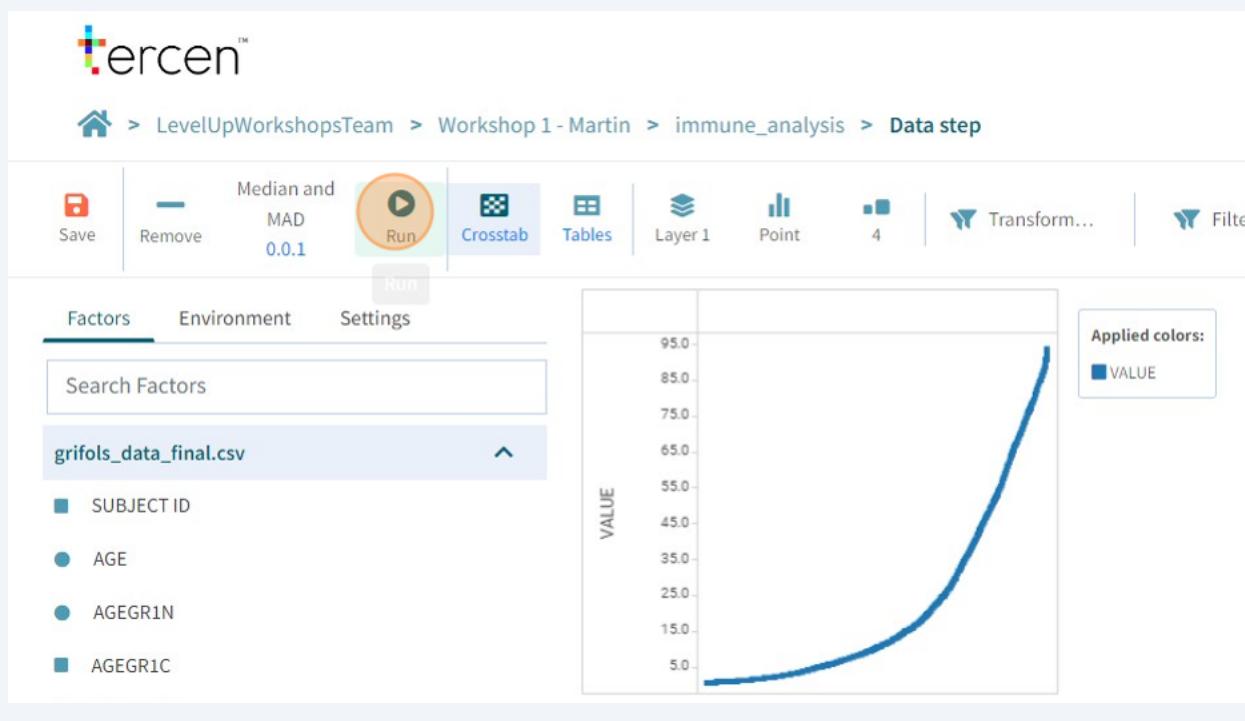
**12** On your browser tab, return to your Tercen project.

Drag VALUE to the Y-Axis.

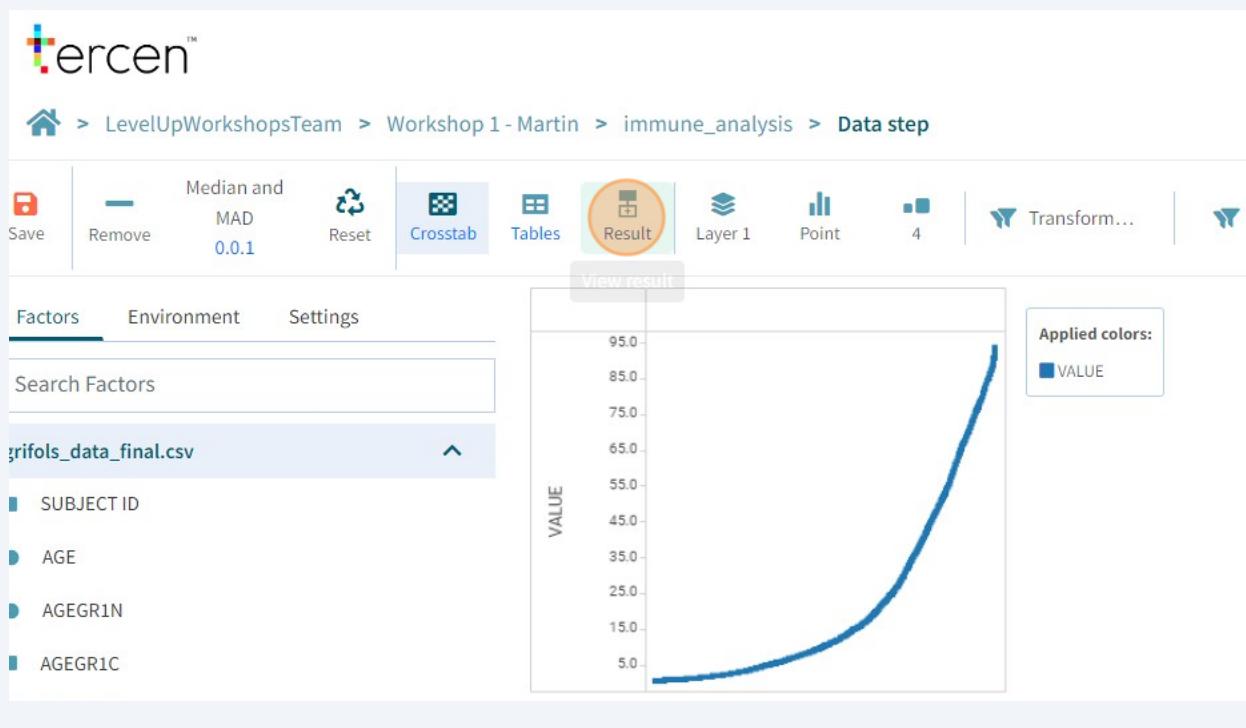


**13** Run the operator.

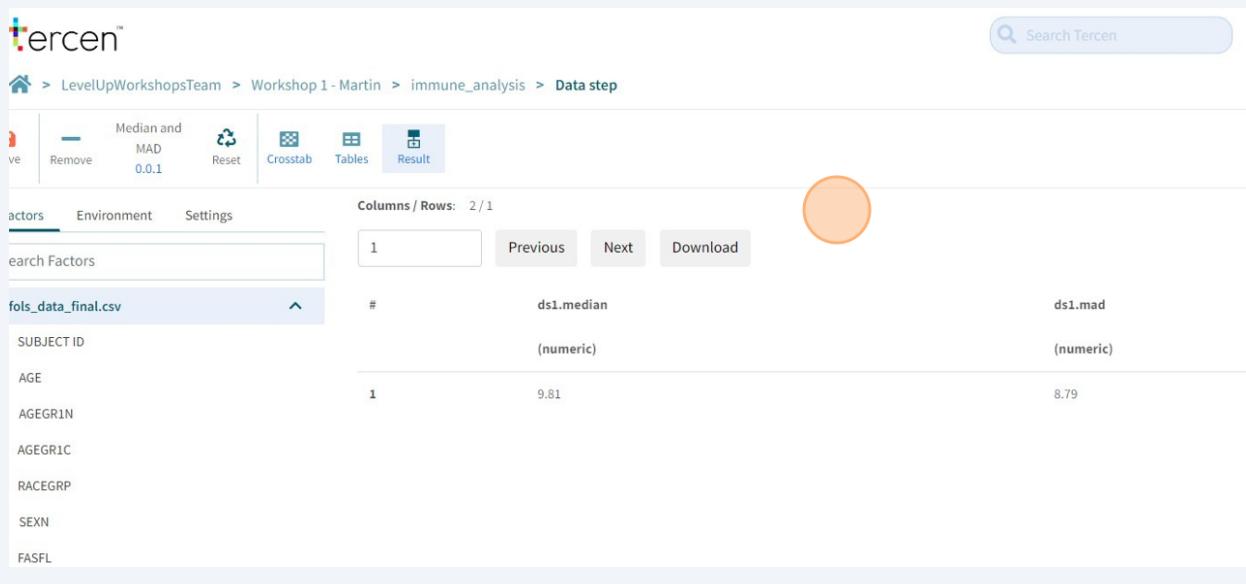
It will calculate a Median and a MAD of all the values in the data cell.



- 14** Click the Result button to see what the operator has produced.



- 15** It has calculated the the Median and the Median Absolute Deviation (MAD) of the all of the values.



**16** Click "Crosstab"

The screenshot shows the tercen™ software interface for a data analysis step. The top navigation bar shows the path: Home > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Data step. The main toolbar includes buttons for Save, Remove, Median and MAD (0.0.1), Reset, Crosstab (which is highlighted with an orange circle), Tables, and Result. Below the toolbar, there are tabs for Factors, Environment, and Settings, with Factors selected. A search bar labeled "Search Factors" contains the text "grifols\_data\_final.csv". The results table shows one row of data:

#	ds1.median
1	(numeric) 9.81

Buttons for View crosstab, Columns / Rows: 2 / 1, Previous, Next, and Download are located at the bottom of the table view.



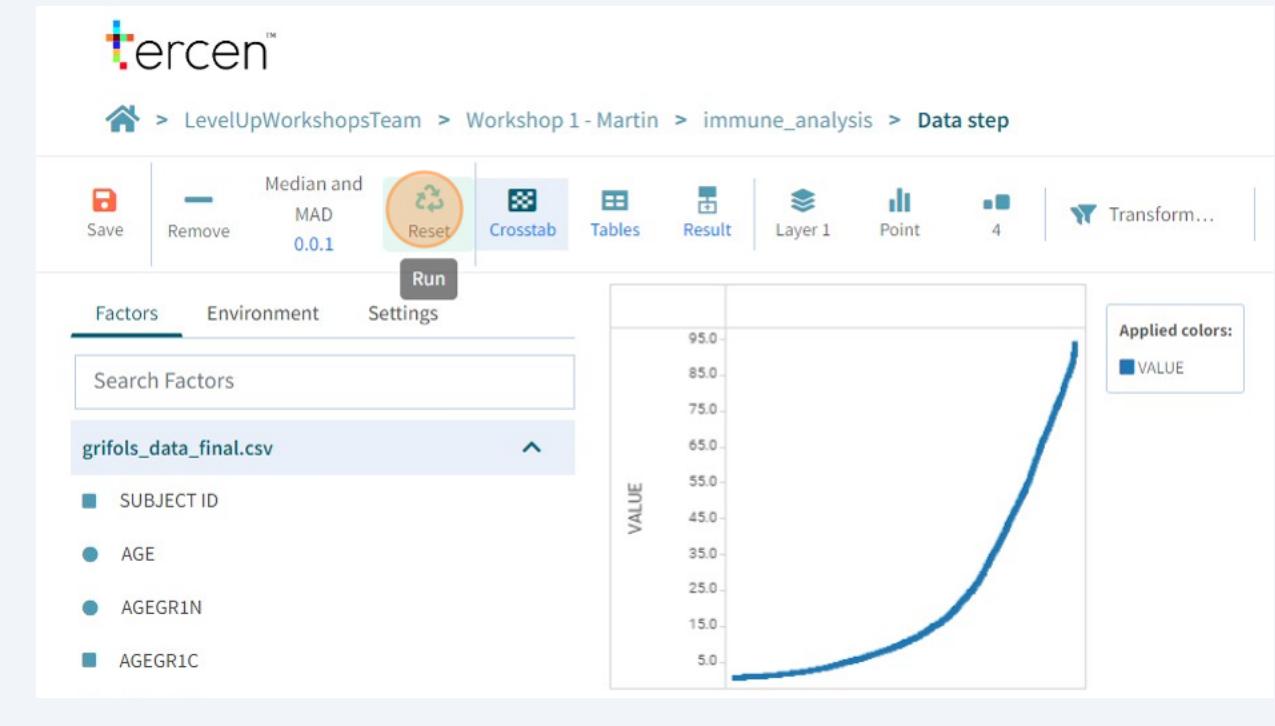
When an operator runs, it locks the Data Step to preserve its results. It must be reset before any changes are made.

17

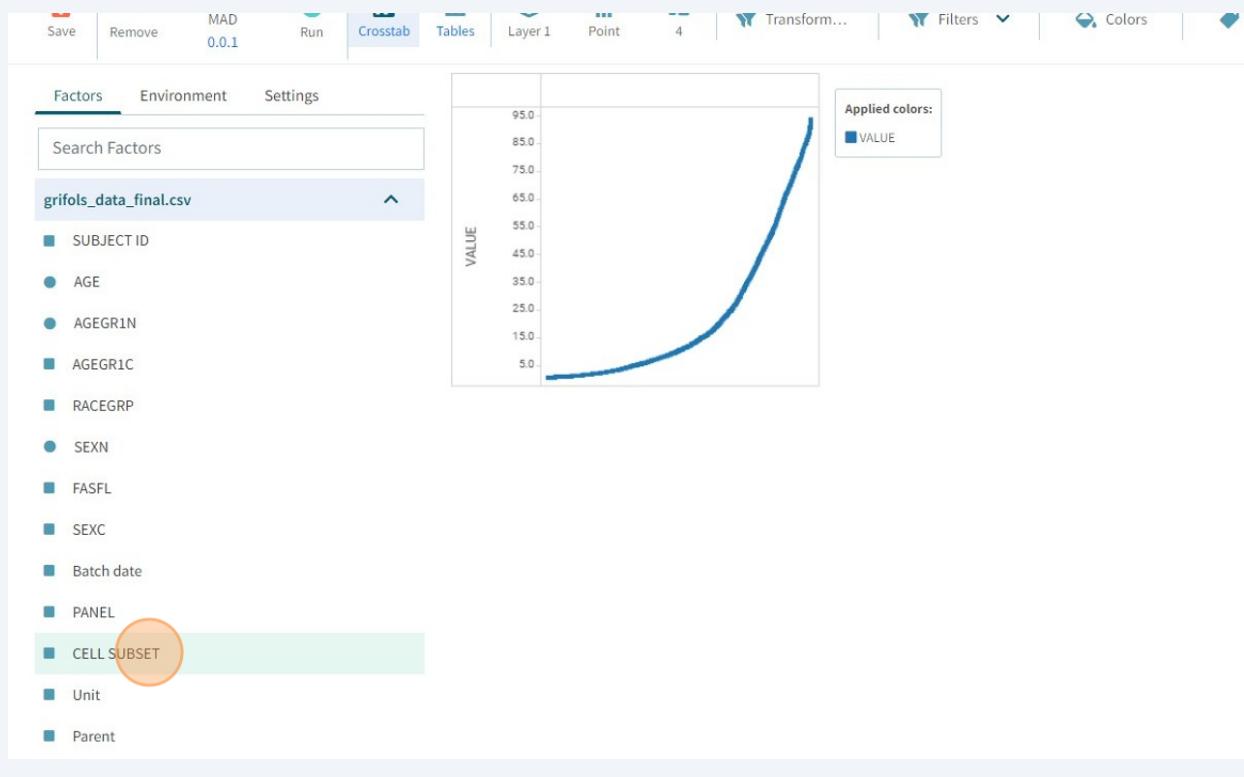
The Median and MAD operator will calculate on all the Data Cells it sees.

We will repeat the process on a multi data cell projection.

Press the **Reset** button.

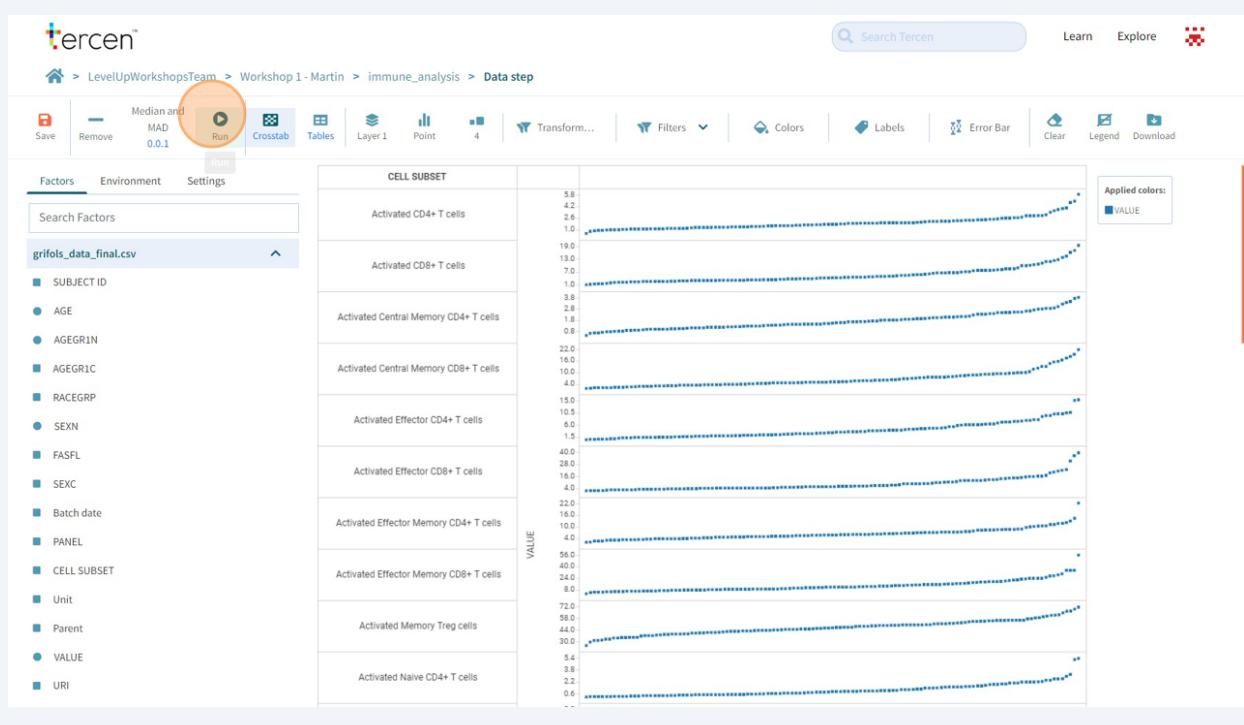


## 18 Drag CELL SUBSET to row.

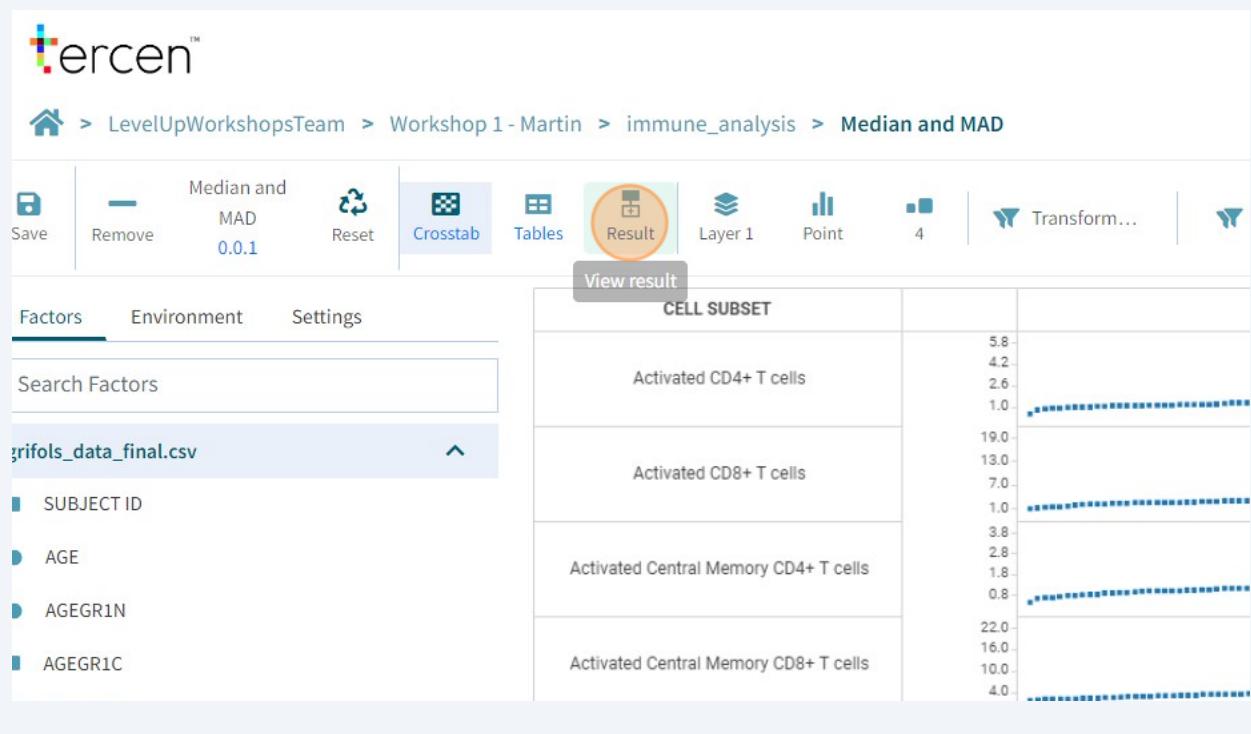


## 19 Now, the values are grouped by Cell Type.

Run the Operator again.



20 Click "Result"



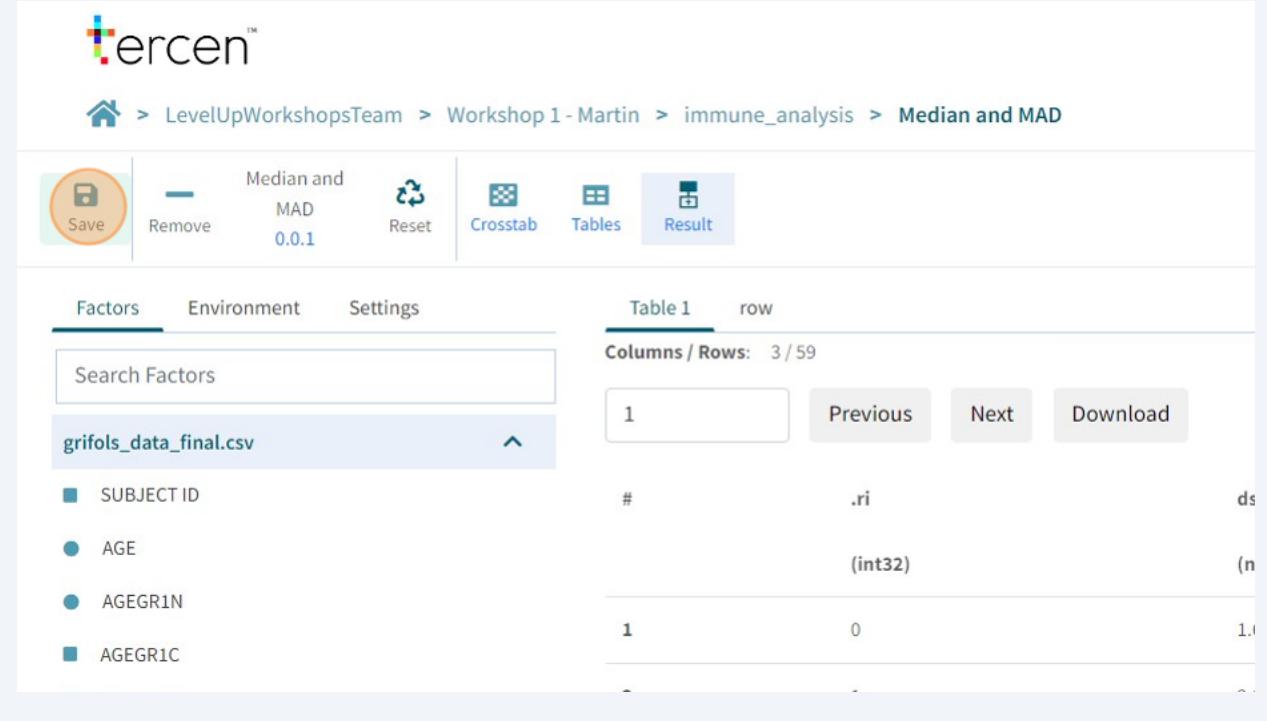
21 The operator has calculated 59 Median and 59 MAD

One for each row.

The screenshot shows the Tercen software interface. At the top, there is a search bar 'Search Tercen' and a 'Learn' button. Below the search bar is a navigation bar with the path: IUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Median and MAD. Below the navigation bar is a toolbar with 'Median and MAD' (set to 0.01), 'Reset', 'Crosstab', 'Tables', and 'Result' (highlighted with a yellow circle). Under the toolbar, there are two tabs: Environment and Settings. A table titled 'Table 1' is displayed with columns: '#', 'row', 'ds1.median', and 'ds1.mad'. The table contains 6 rows of data. The 'Download' button in the table header is also highlighted with a yellow circle.

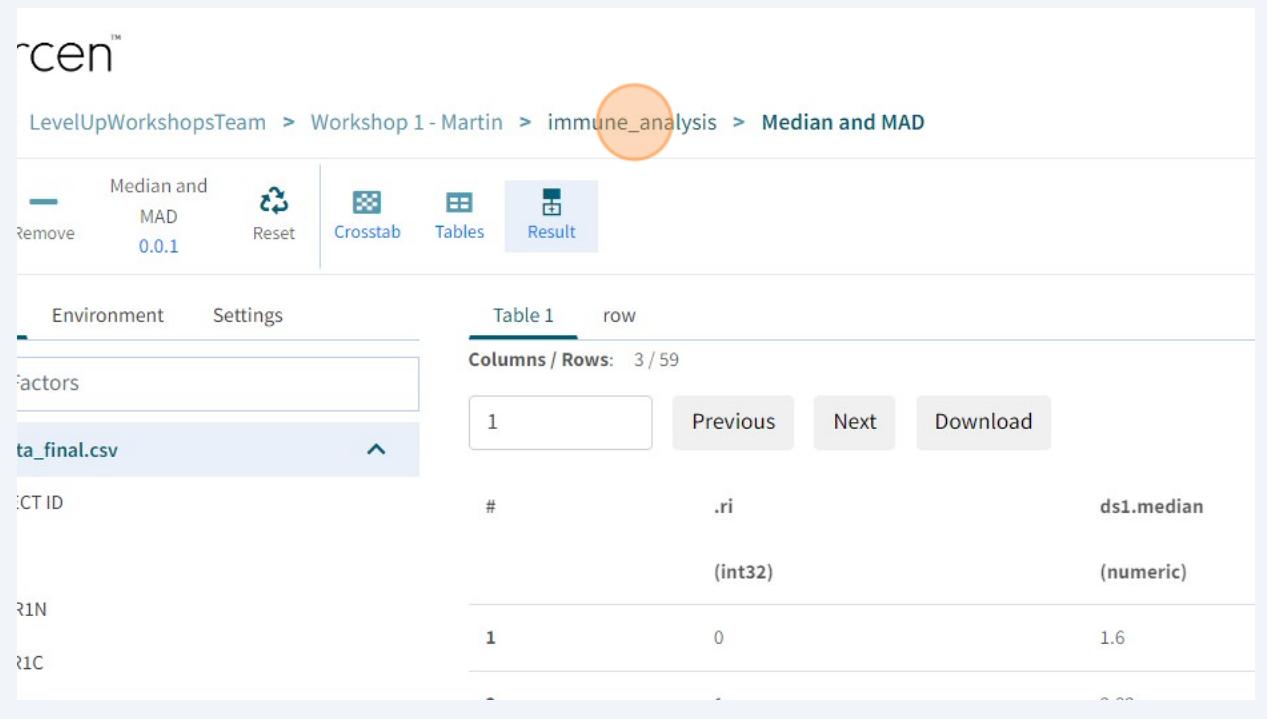
#	row	ds1.median	ds1.mad
		(int32)	(numeric)
1	0	1.6	0.43
2	1	3.82	1.53
3	2	1.47	0.47
4	3	4.63	1.97
5	4	2.84	1.52
6	5	5.13	2.57

22 Click "Save"



The screenshot shows the tercen™ software interface. At the top, there's a navigation bar: Home > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Median and MAD. Below the navigation is a toolbar with buttons for Save (highlighted with a red circle), Remove, Median and MAD (value 0.0.1), Reset, Crosstab, Tables, and Result (selected). On the left, there are tabs for Factors, Environment, and Settings, with a search bar for 'Search Factors' containing 'grifols\_data\_final.csv'. Under the Factors tab, a list of variables is shown: SUBJECT ID, AGE, AGEGR1N, and AGEGR1C. To the right, a table titled 'Table 1' shows data for SUBJECT ID. The table has columns for '#', 'ds', and '(int32)'. The first row shows '#': 1, 'ds': .ri, and '(int32)': 0. There are also 'Previous', 'Next', and 'Download' buttons.

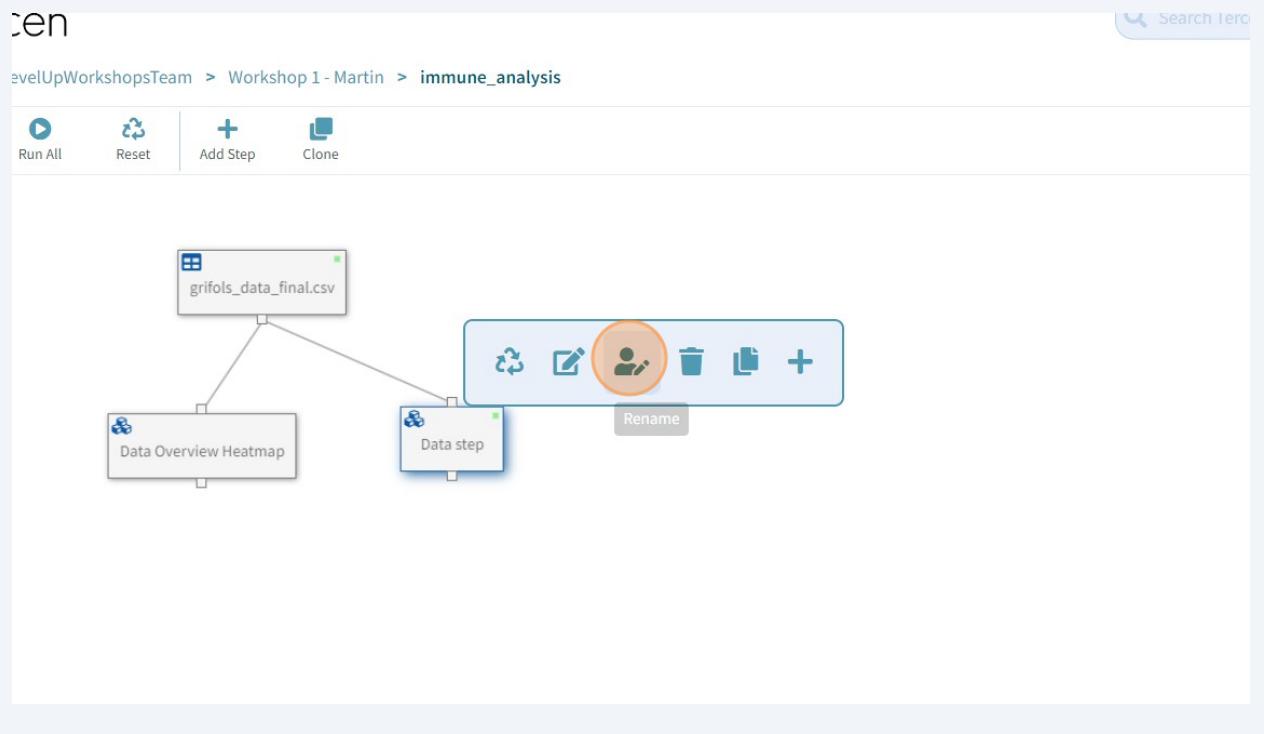
23 Return to the "**immune\_analysis**" workflow canvas.



This screenshot is similar to the previous one but shows a red circle highlighting the 'immune\_analysis' link in the navigation bar. The rest of the interface is identical, displaying the Median and MAD analysis results for the 'grifols\_data\_final.csv' dataset.

**24** Rename the Data Step to "Median and MAD".

Press OK.

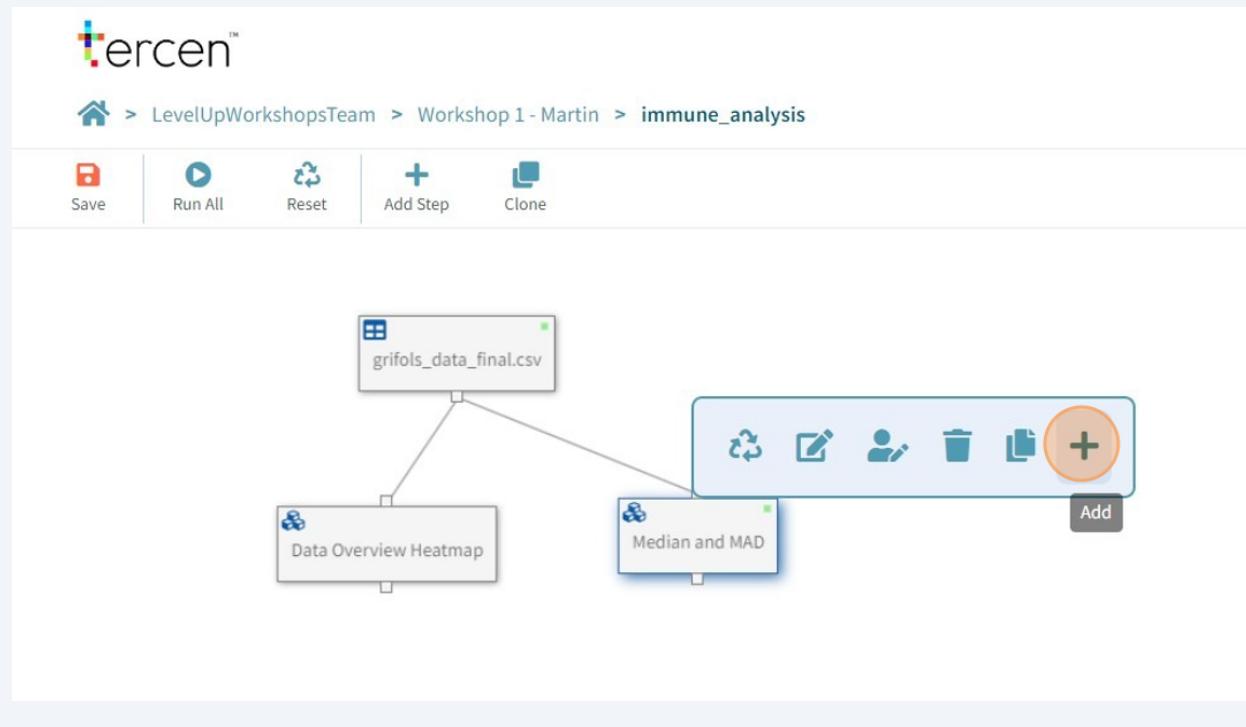


25

The results of calculations made by an operator can be used to create new projections and further calculations.

This is done by adding a new Data Step downstream from the one with the operator.

Add a new data step from the local toolbar.

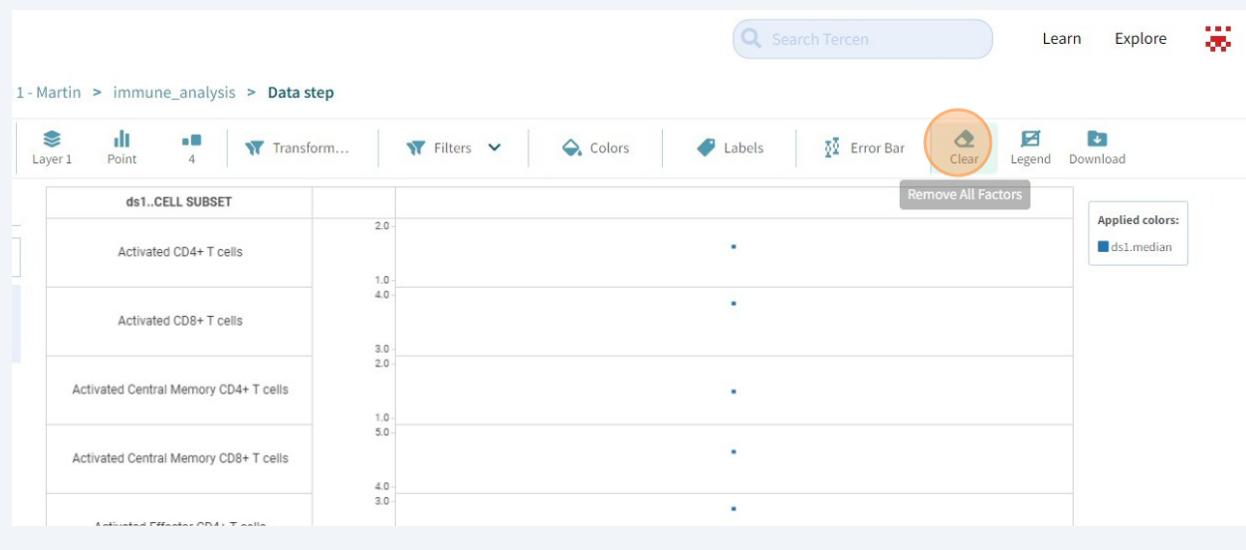


26

The Tercen AI will attempt to predict the projection you want make with the new factors.

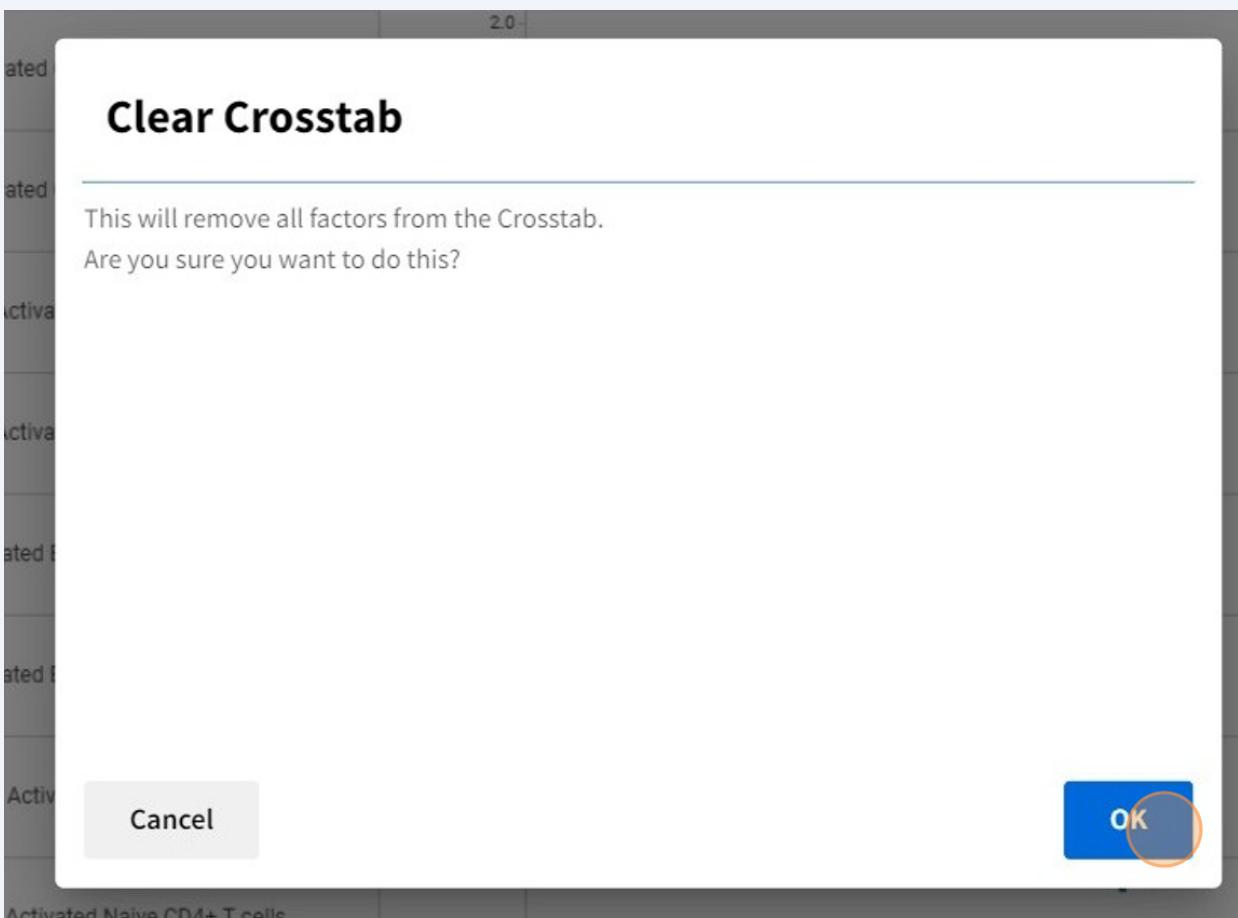
In this tutorial we will start from the beginning.

Press the Clear button to clear the crosstab grid.



27

Click "OK"



**28** Note the Factor Panel to the left has a new section.

The original factors taken from the grifols\_data\_final.csv are still available.

As are the new Factors created by the Median and MAD operator.

To make a visualisation with the newly created factors.

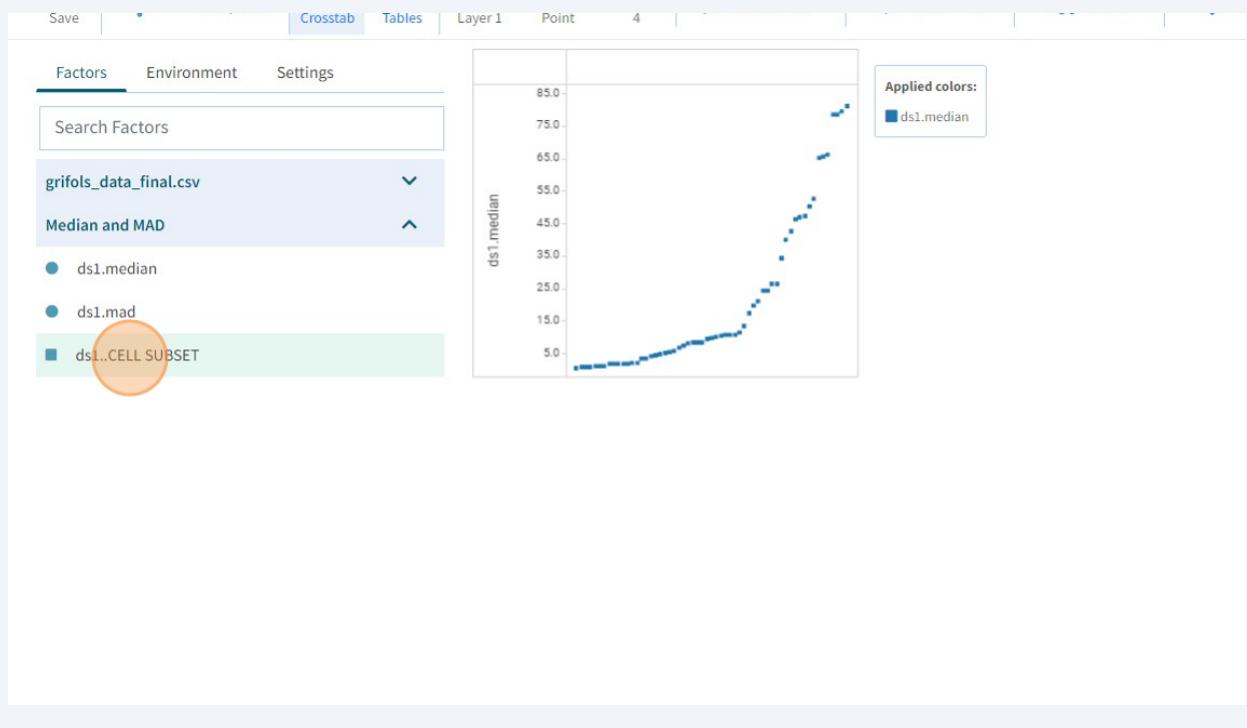
Drag **ds1.median** to the Y-Axis.

The screenshot shows the Tercen software interface. At the top, there is a toolbar with various icons: Save, Add Operator, Crosstab (selected), Tables, Layer 1, Point, Transform..., and Filters. Below the toolbar is a navigation bar with three tabs: Factors (selected), Environment, and Settings. Under the Factors tab, there is a search bar labeled "Search Factors". A dropdown menu is open, showing "grifols\_data\_final.csv" (selected) and "Median and MAD" (also selected). Under "Median and MAD", there are three items: "ds1.median" (highlighted with an orange circle), "ds1.mad", and "ds1..CELL SUBSET". To the right of the Factor Panel is a large, empty workspace area.

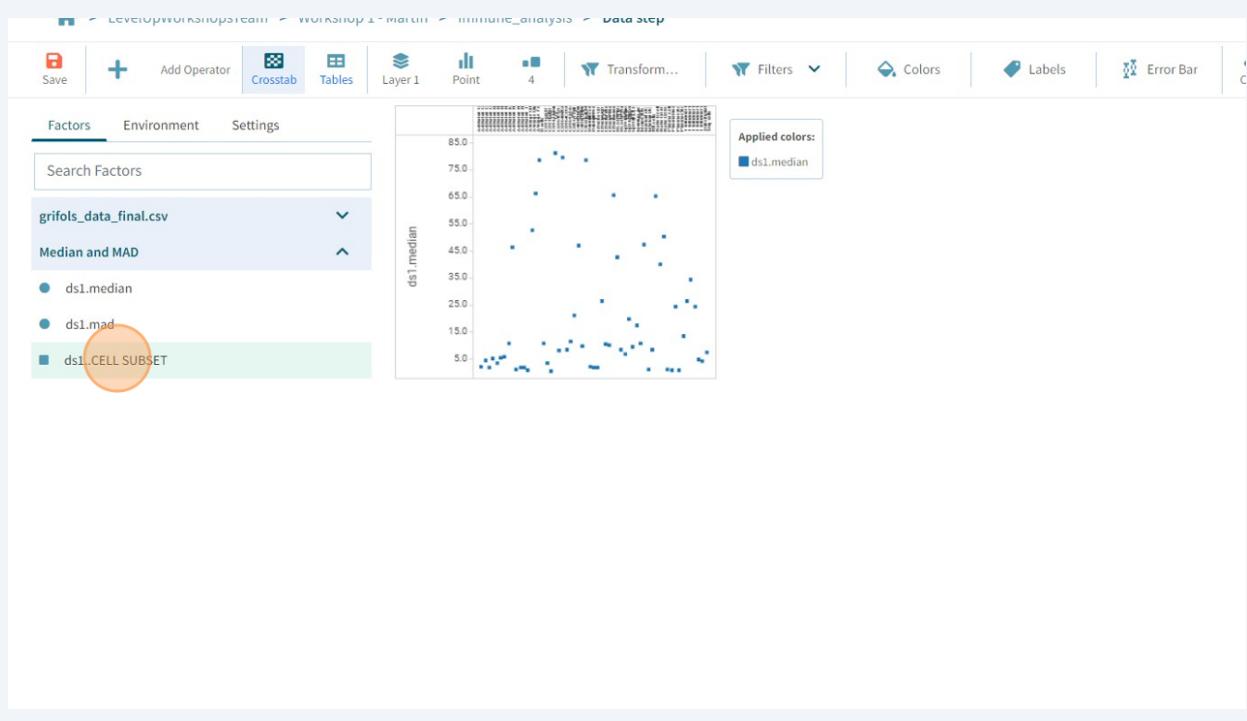


Tercen inserts a "Name-space" code to every new factor that is calculated. "ds1" comes from "Data Step 1". This is to help you differentiate which Data Step calculated which factor as there may be many factors named "value" or "result".

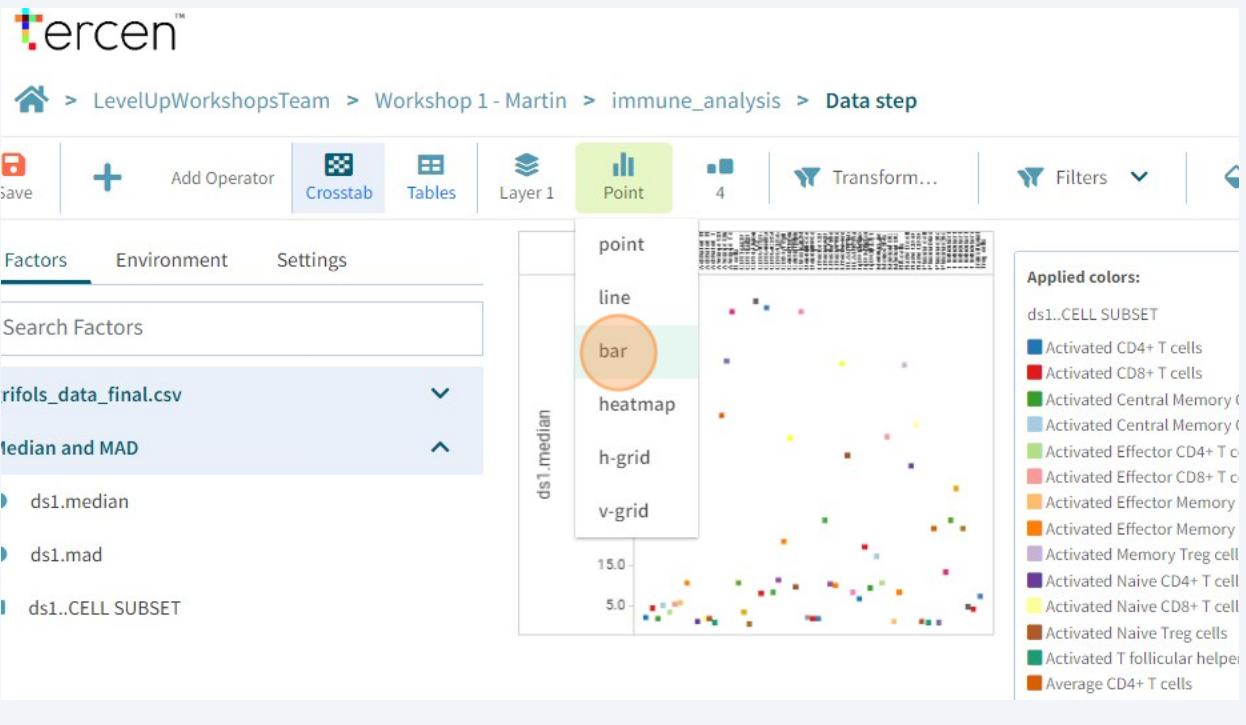
## 29 drag ds1..CELL SUBSET to X-Axis



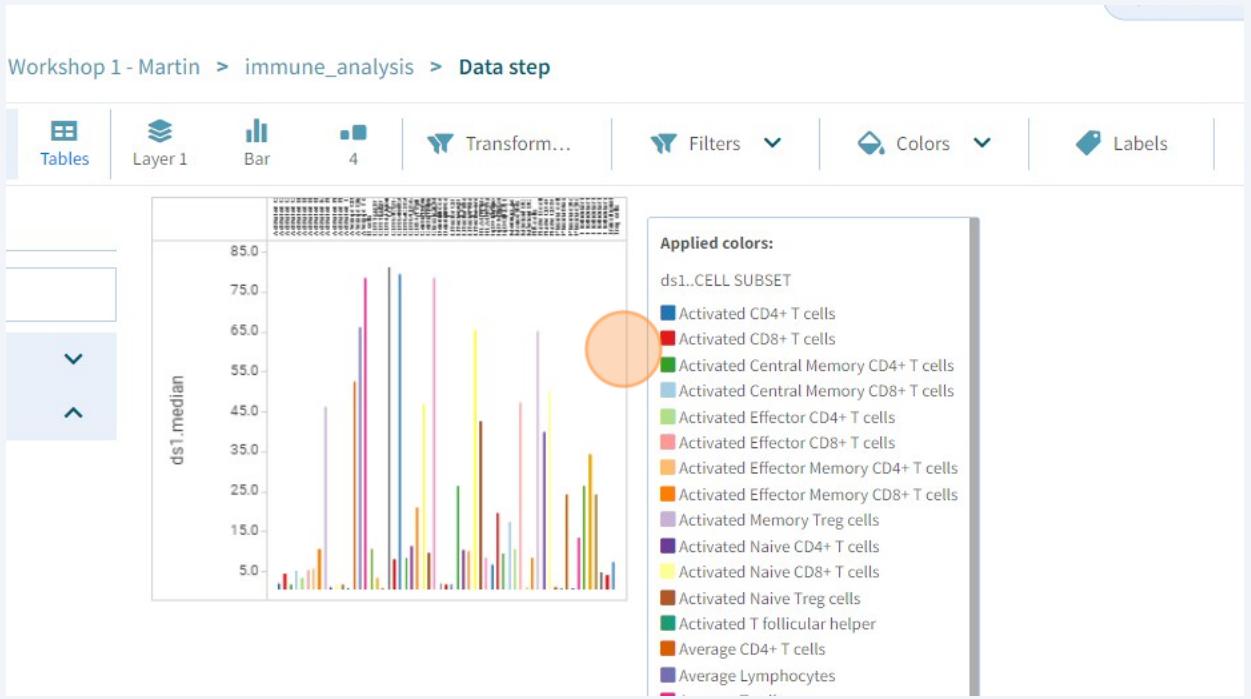
## 30 Drag ds1..CELL SUBSET a second time to Color



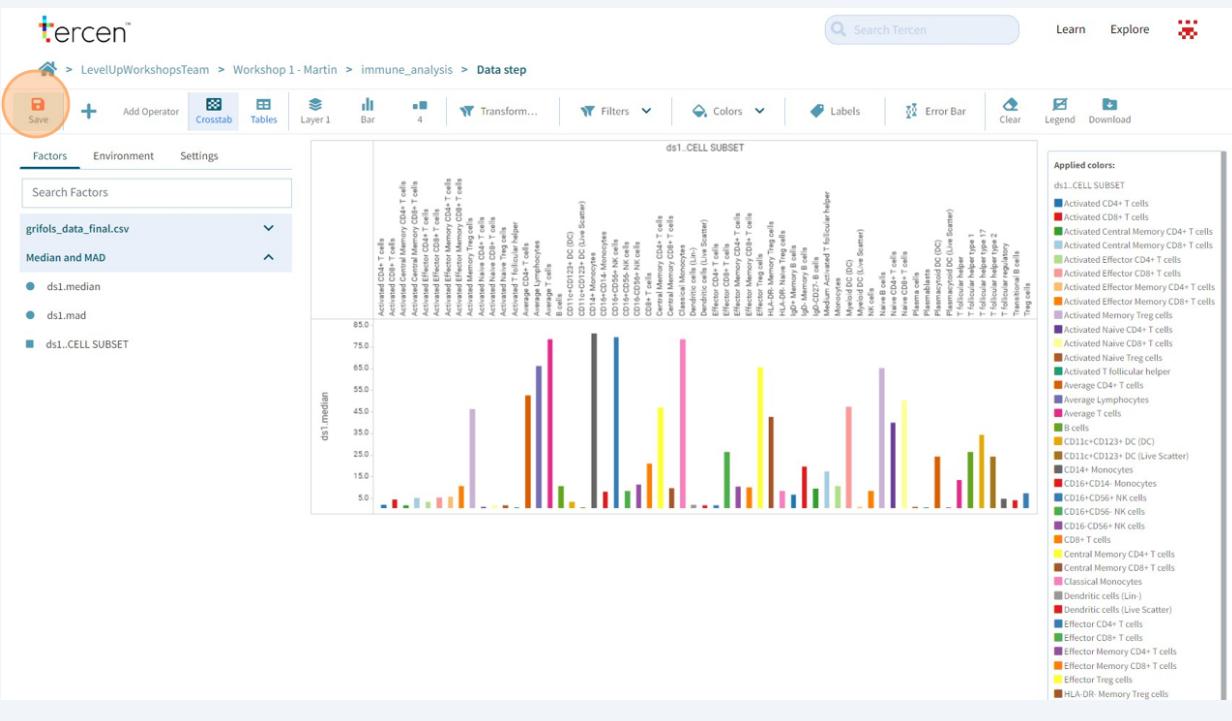
### 31 Change the Graph Style to Bar



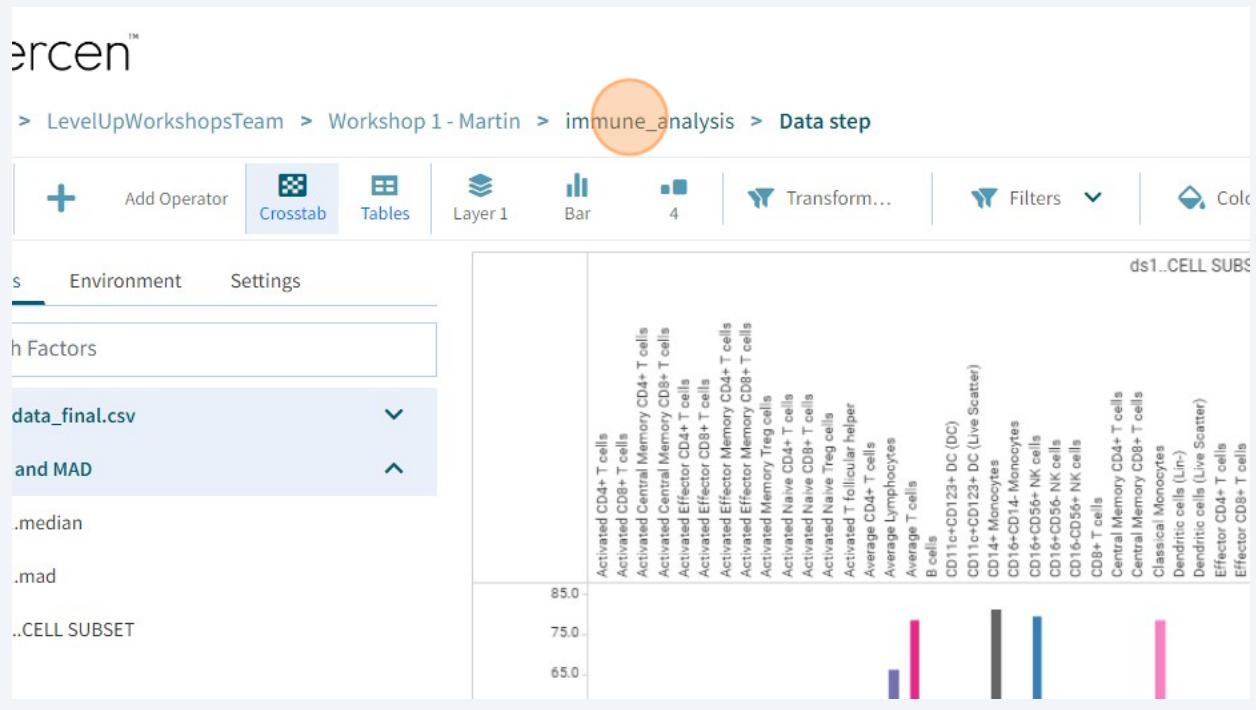
### 32 Use the black lines to adjust your visualisation



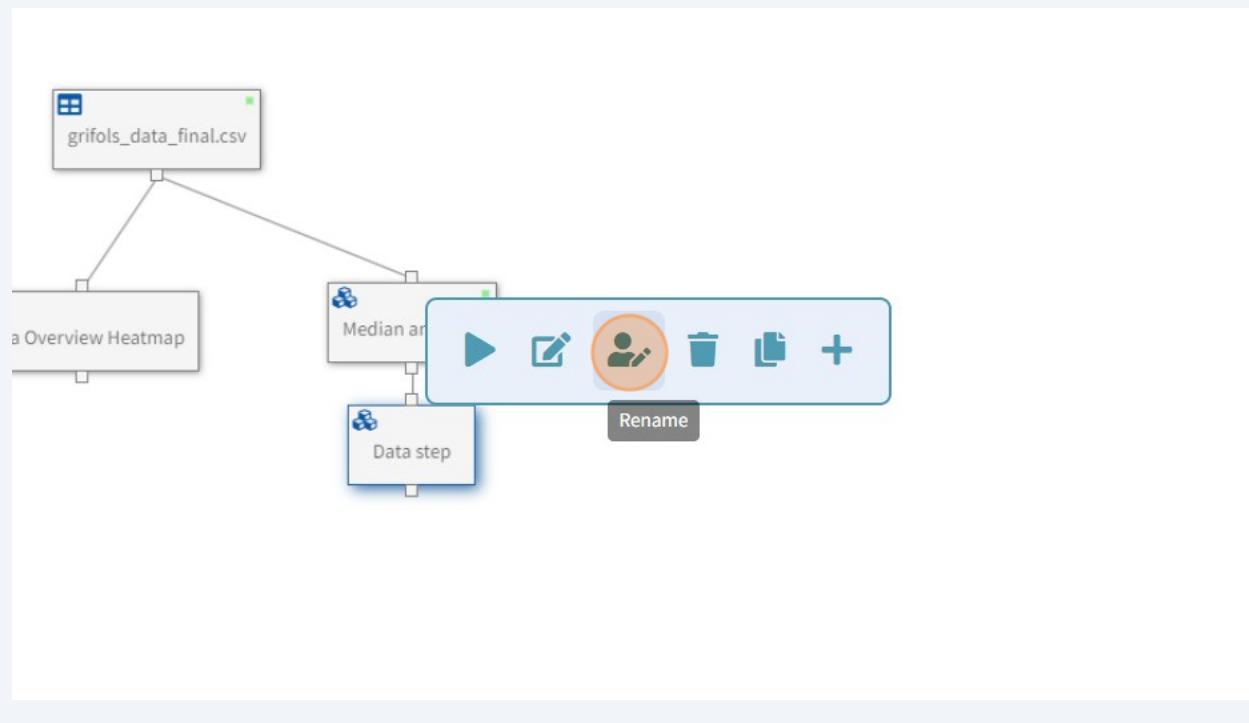
33 Click Save



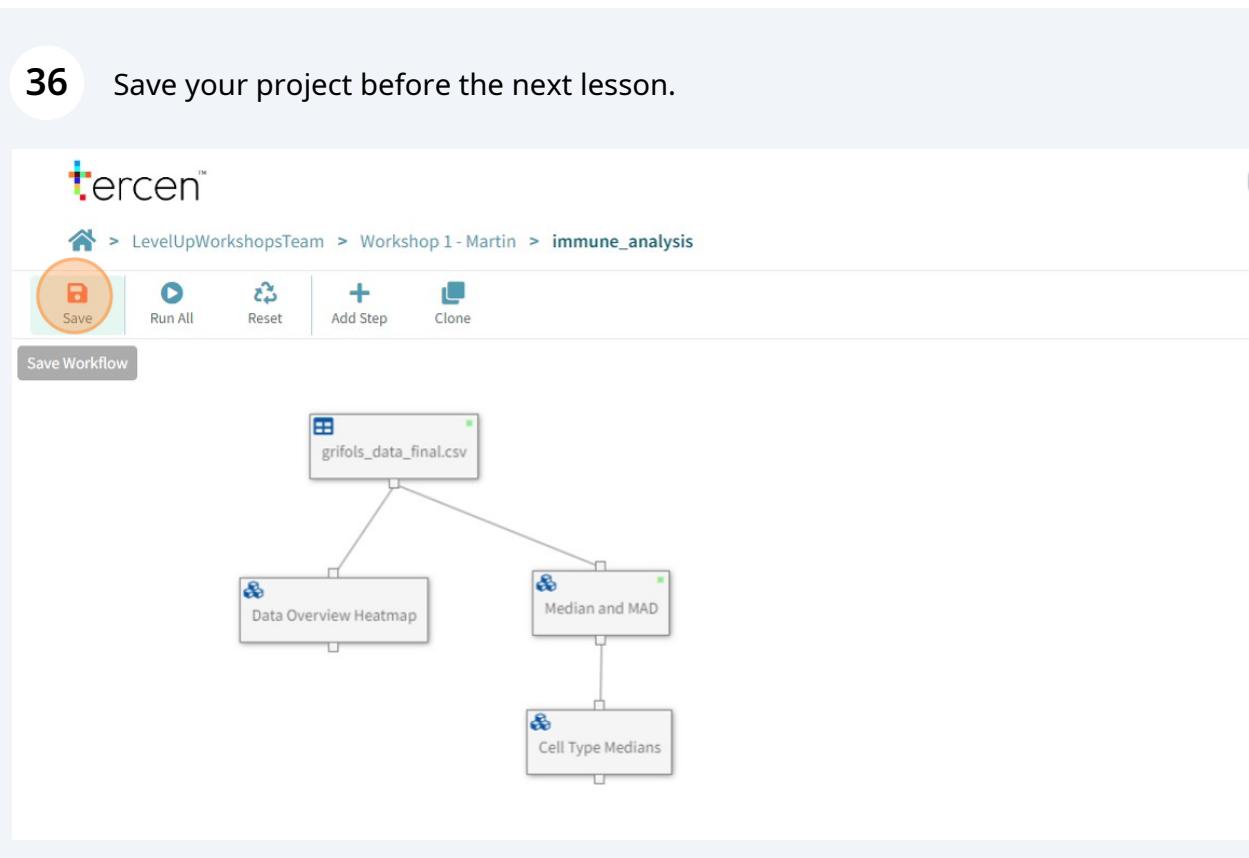
**34** Return to the **immune\_analysis** workflow canvas.



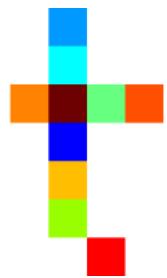
**35** Rename the data step to **Cell Type Medians**



**36** Save your project before the next lesson.



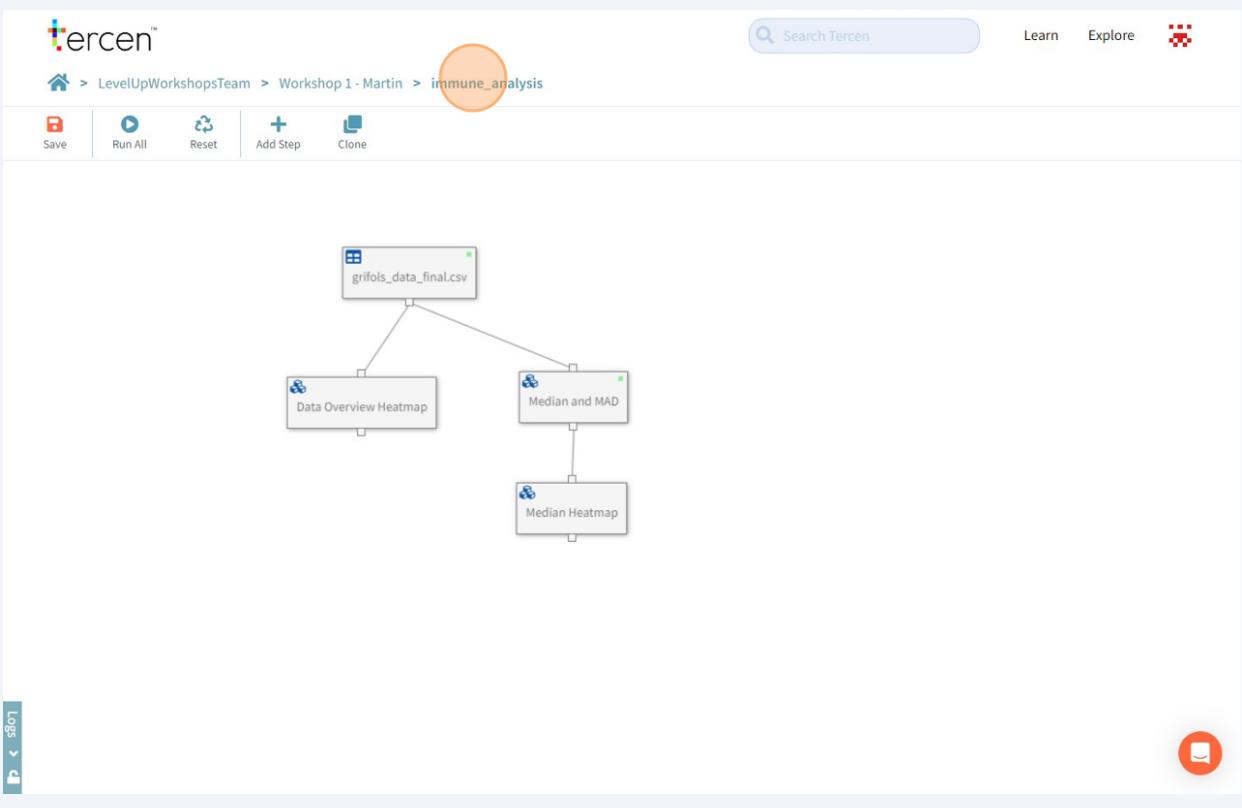
# 0106 - Export Plots and Data



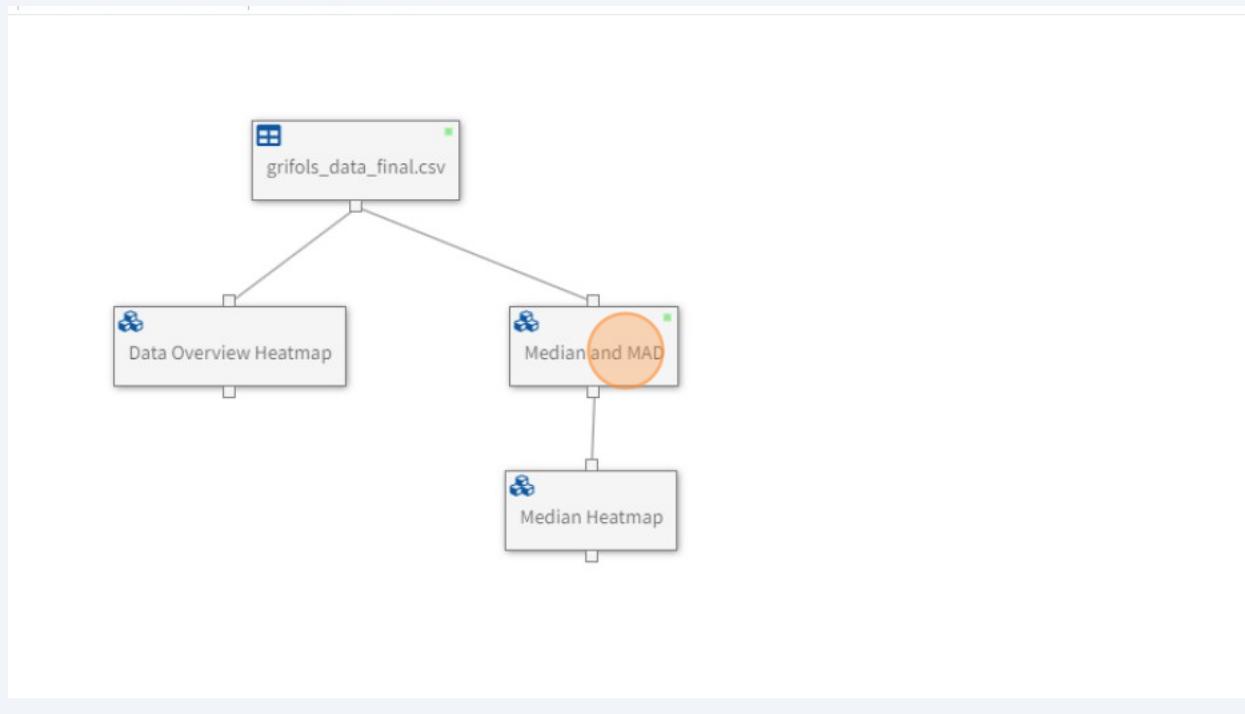
In this tutorial you will learn how Tercen can create images from the Crosstab Grid or from an Operator. You will also learn how to make a projection that allows you to export your results as a data table.

- 1 Starting from the "immune\_analysis" workflow canvas.

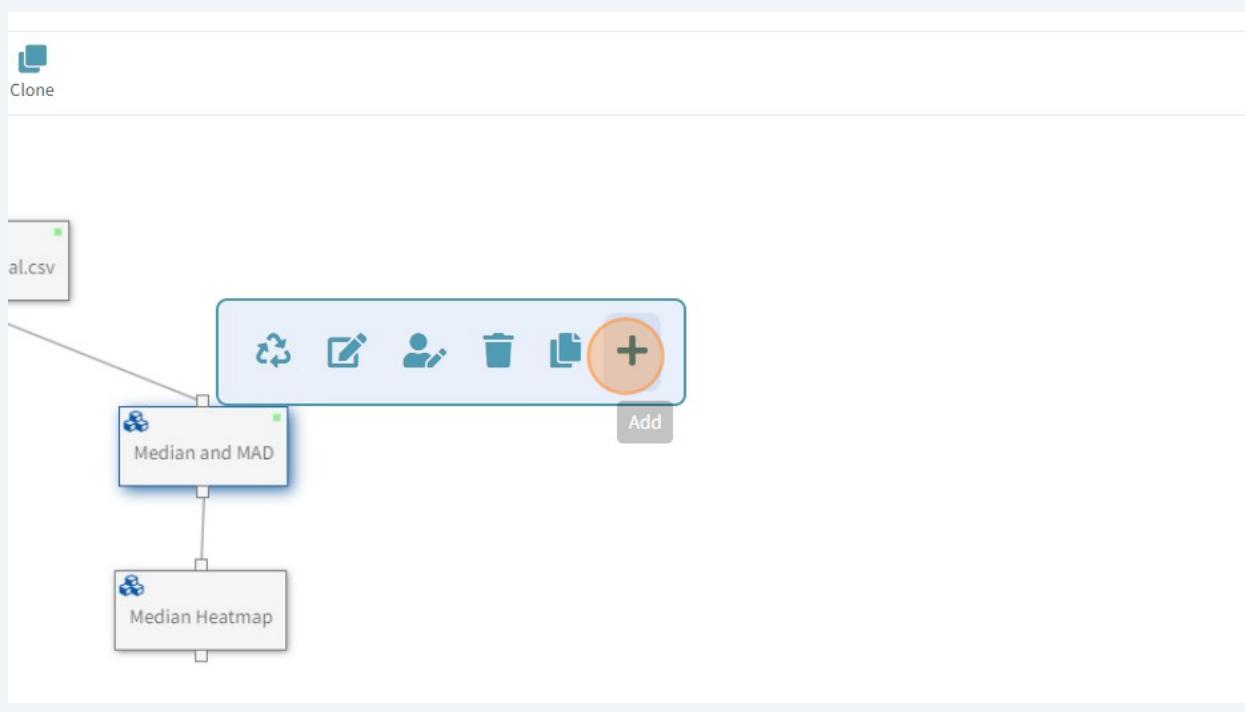
We will make a Bar Graph visualisation.



- 2 Click the Median and MAD data step



- 3 Click Add.



#### 4 Add a Data Step

Add

Step    Operator    Operator Library    Installed Apps    App Library

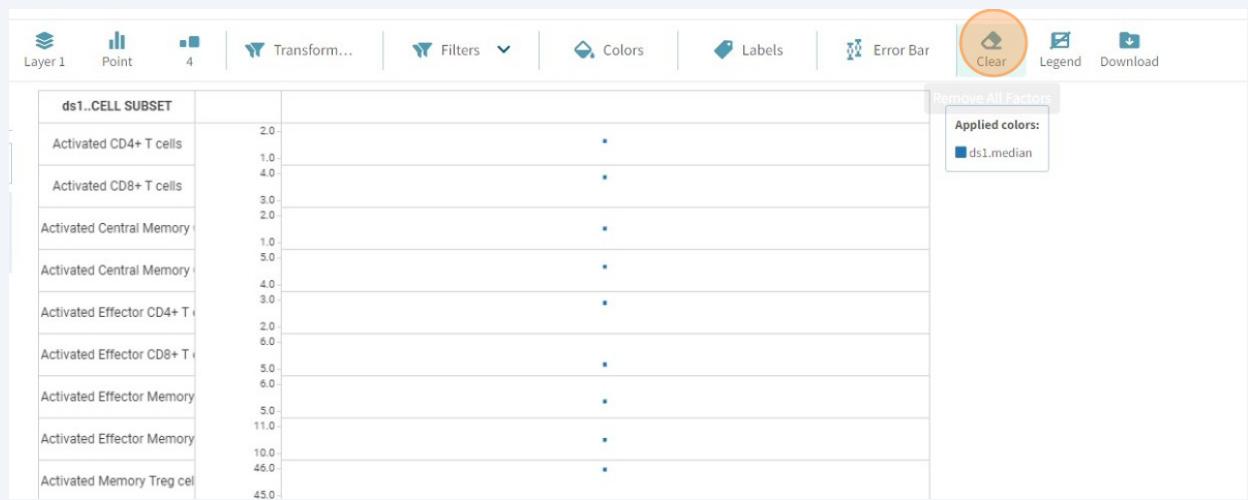
Search

**Data step** data  
Perform computation on user defined projection

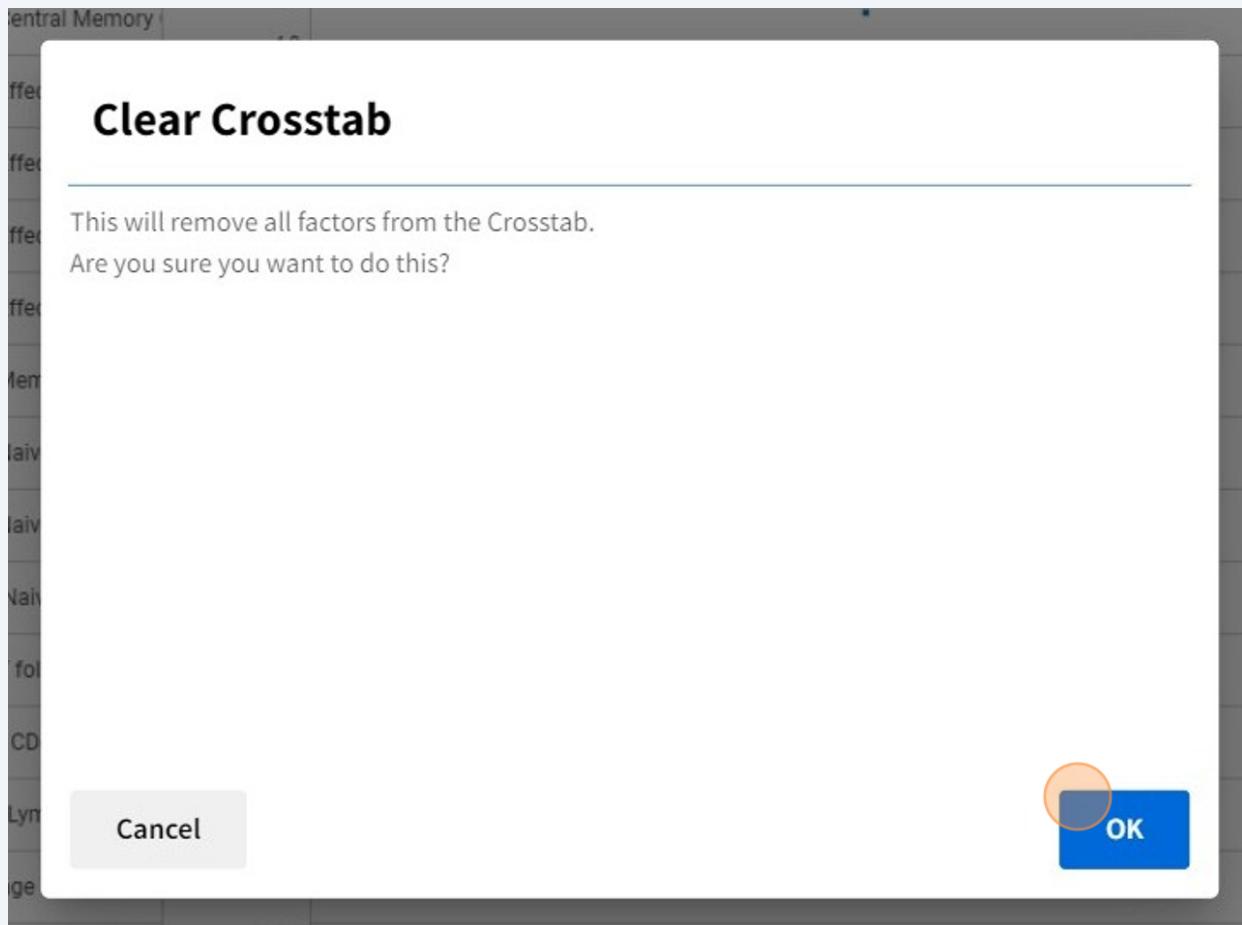
**Multi data step** data  
Perform computation on user defined projection

**Join** leftTable  
Join two data sets

#### 5 Clear the automatic projection.

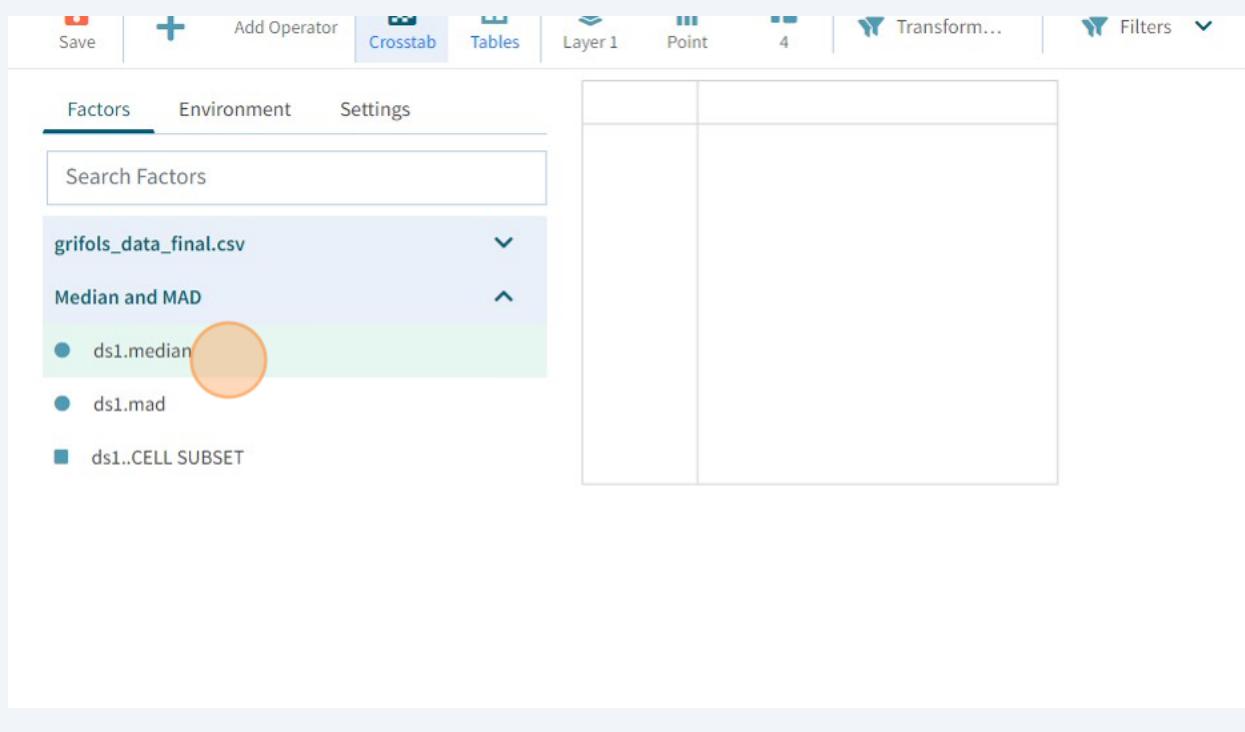


6 Click "OK"

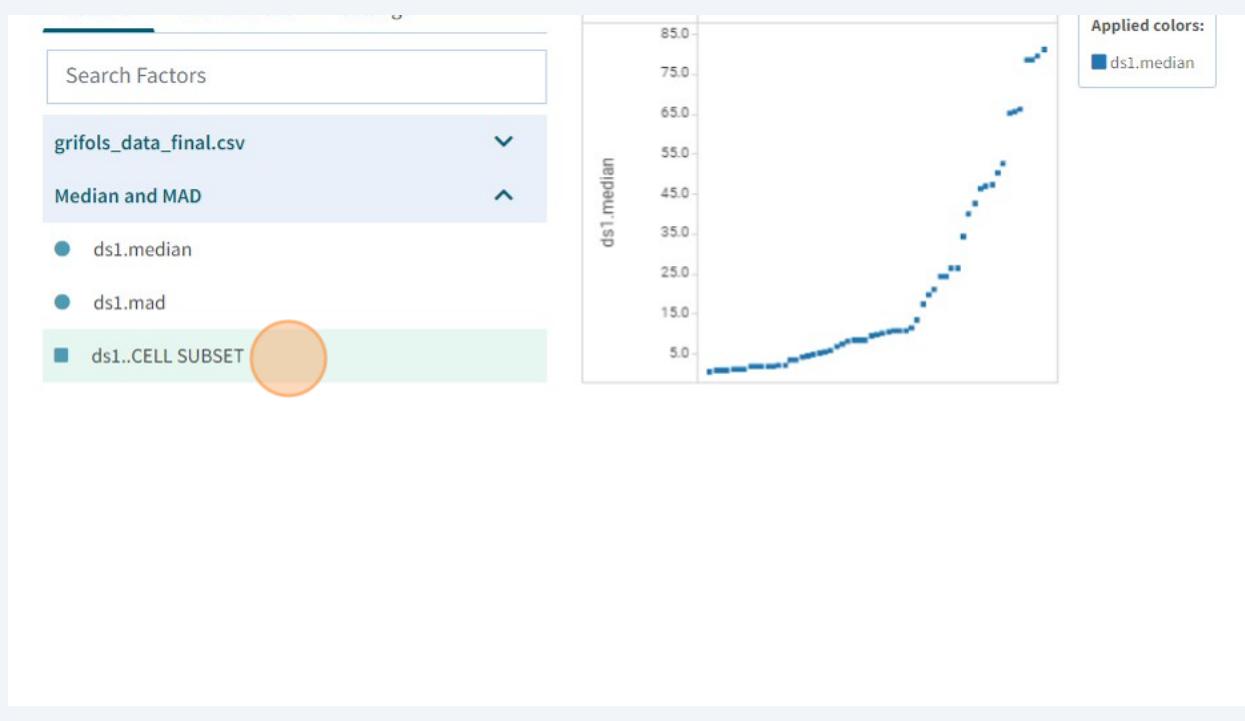


## 7 Let's make a Bar Graph projection

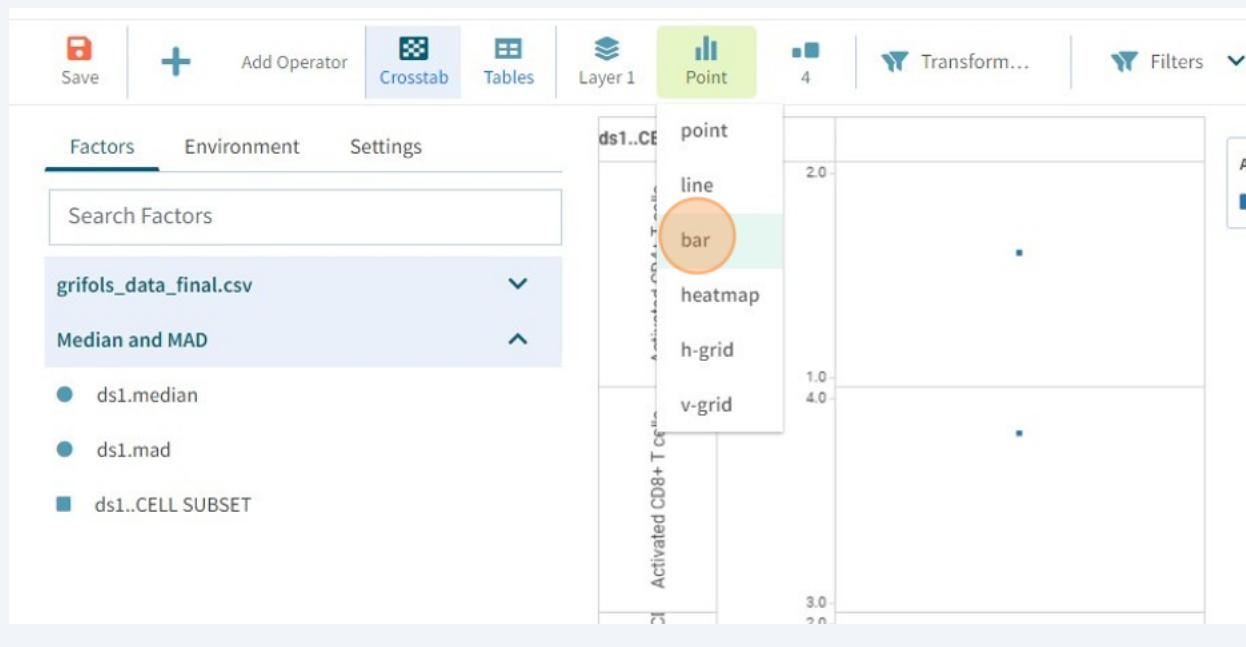
Drag **ds1.Median** to Y-Axis



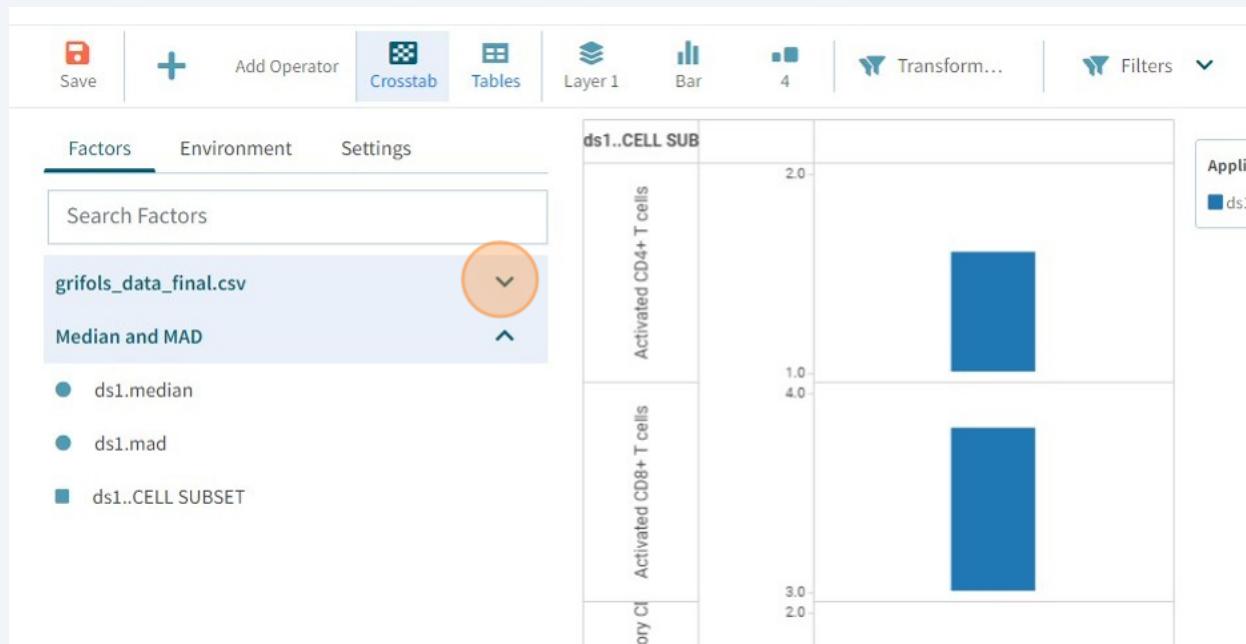
## 8 Drag **ds1.CELL SUBSET** to Row.



## 9 Change the Graph Style to Bar



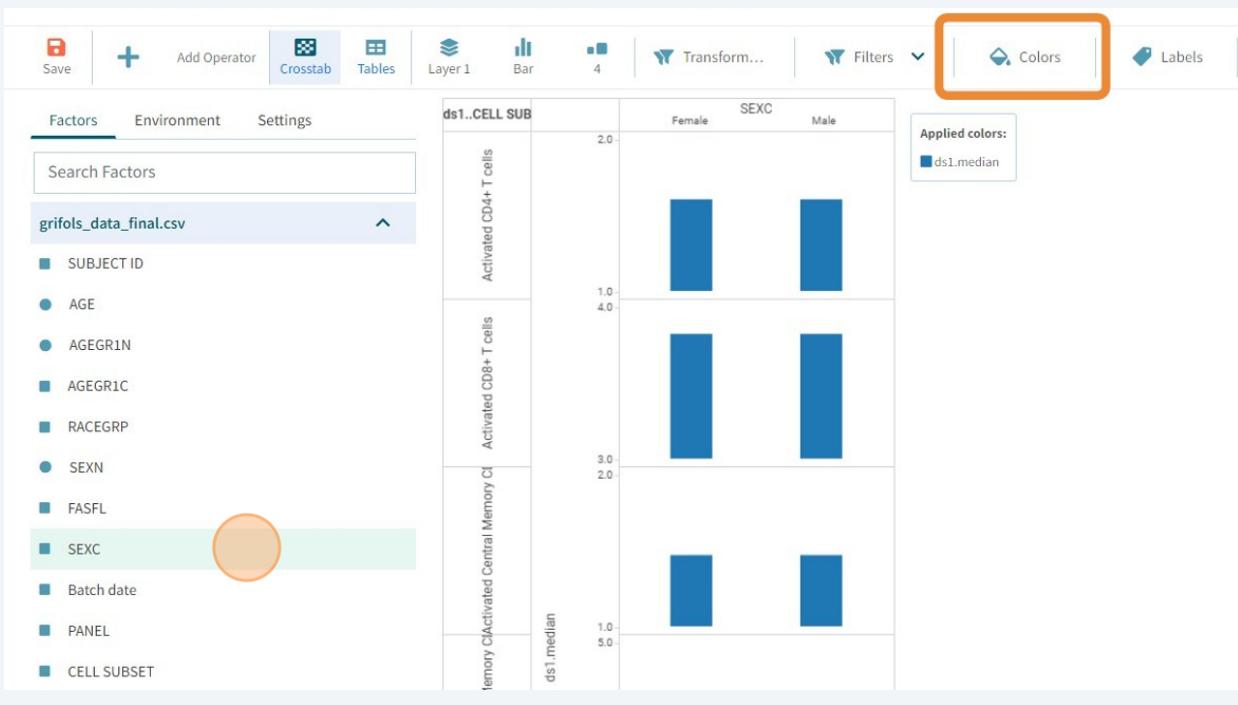
## 10 Open the original data table.



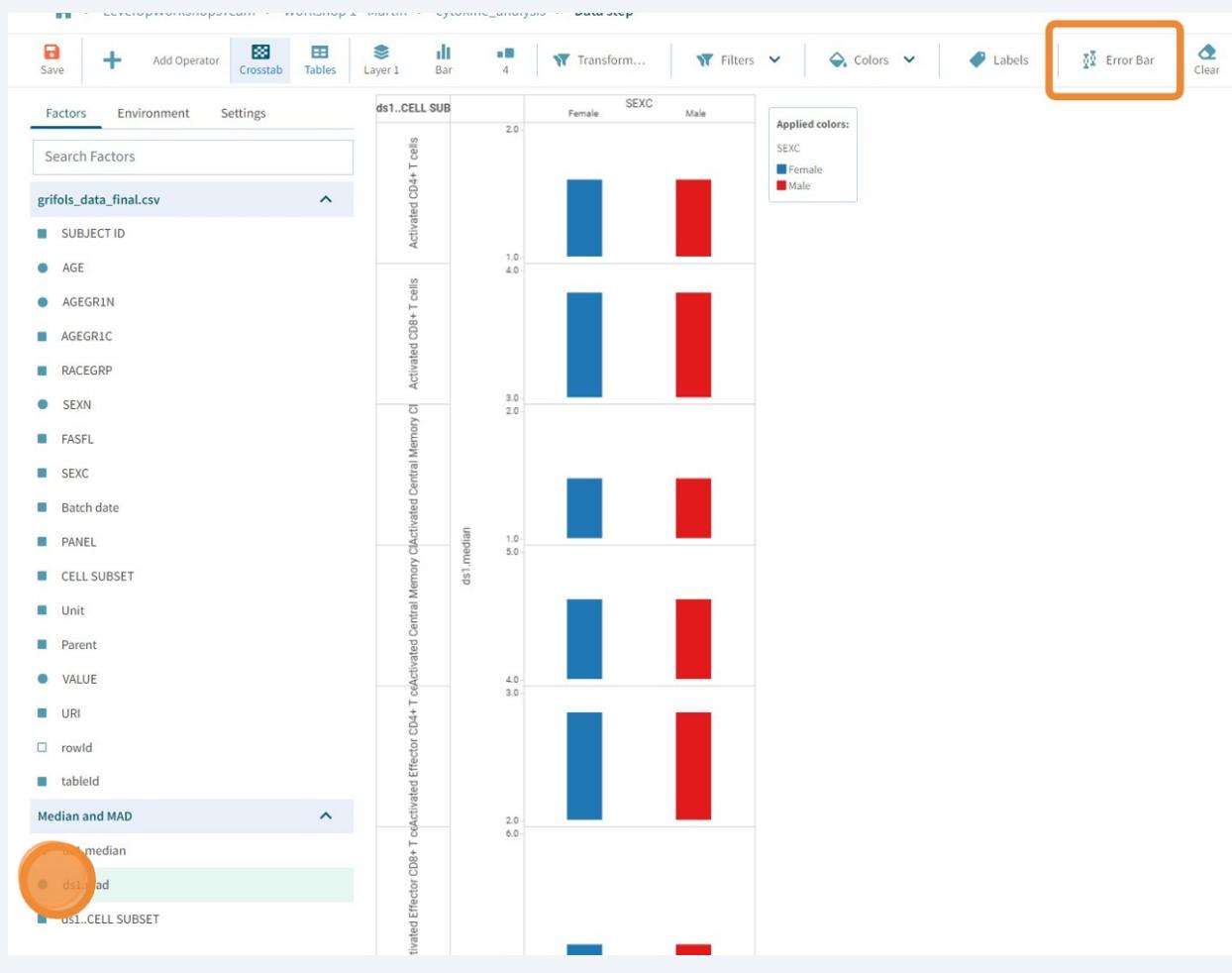
## 11 Drag **SEXC** to X-Axis



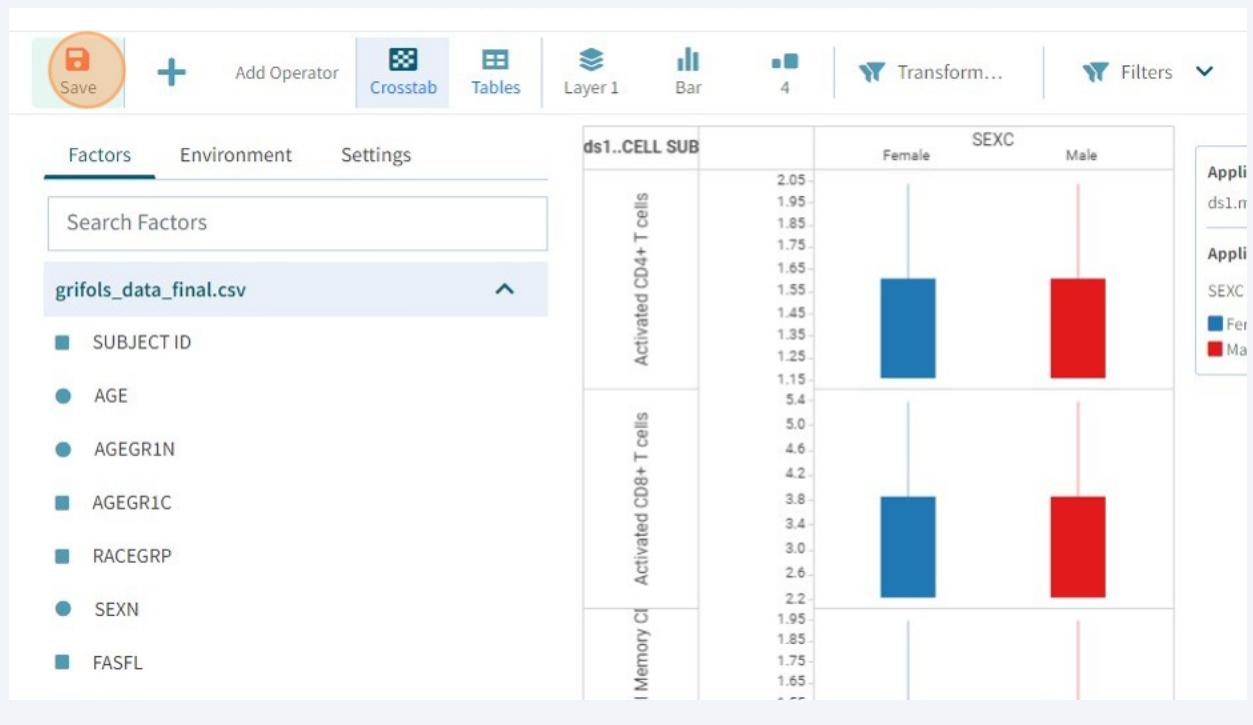
## 12 Drag **SEXC** a second time, to Color.



### 13 Drag ds1.mad to Error Bar

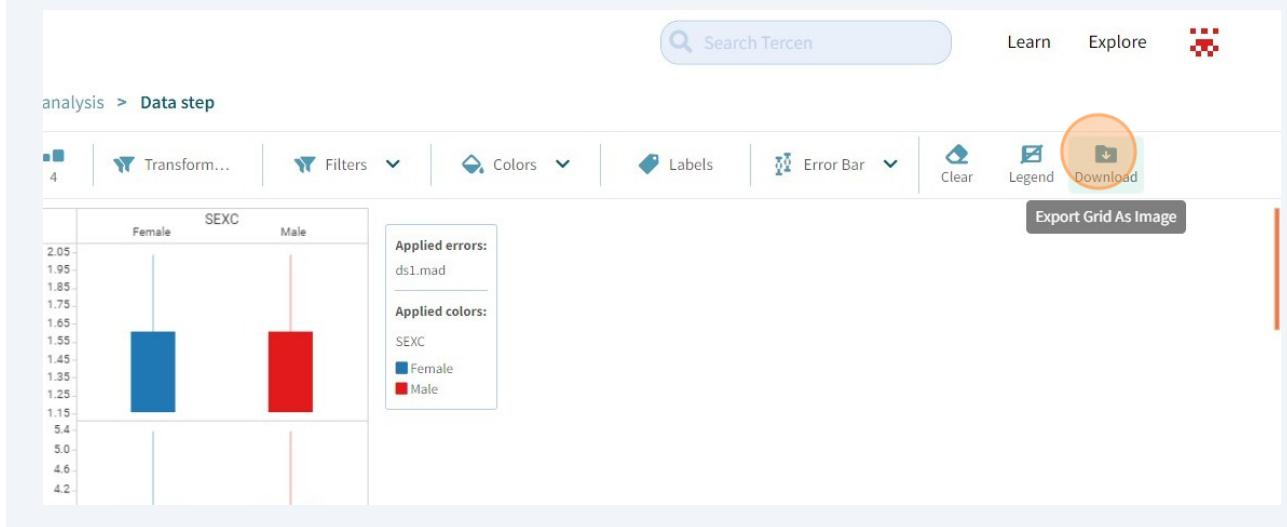


## 14 Click "Save"

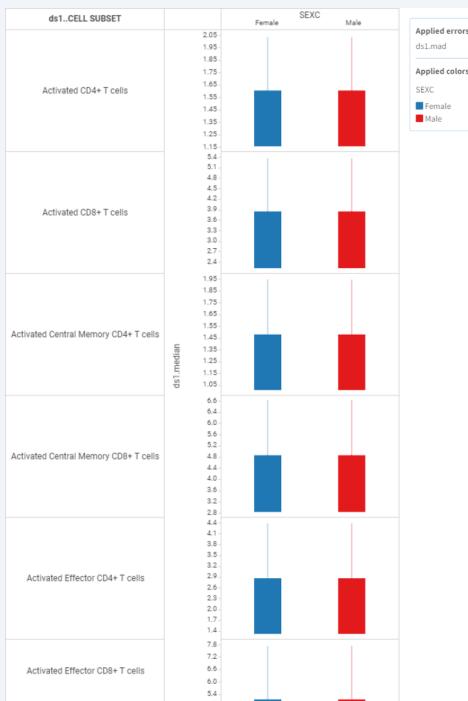


## 15 Tercen has a quick image download button.

Click **Download**



## 16 Review your image from your download area.

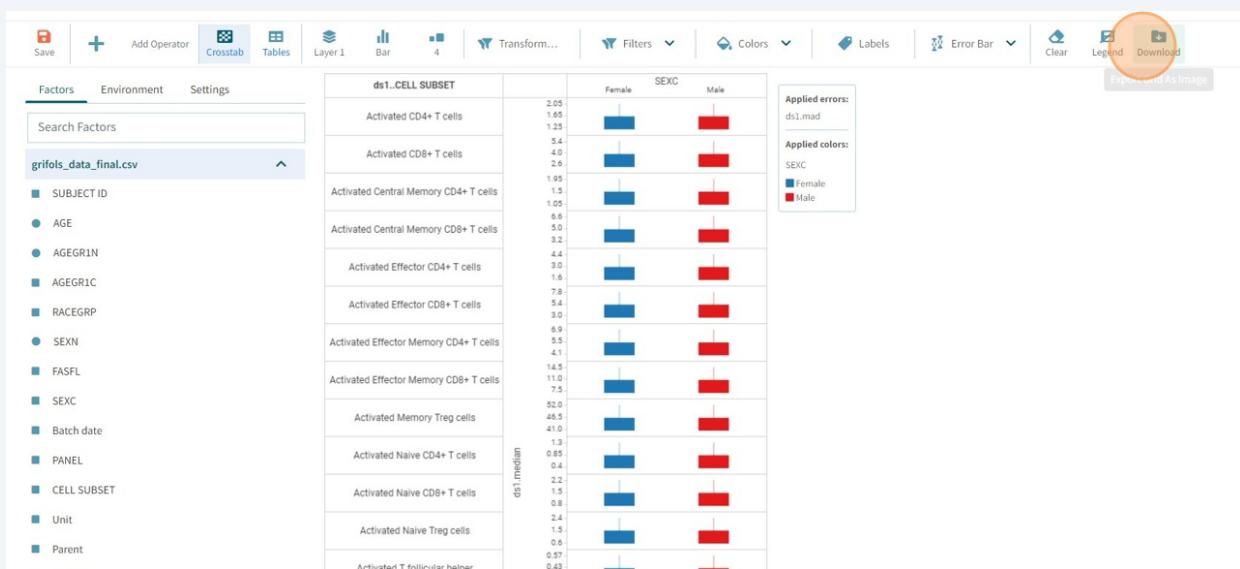


Quick Download generates a .png file of what it sees on screen.

If your projection goes off the bottom of the screen then the hidden parts will not be included in the quick image.

## 17

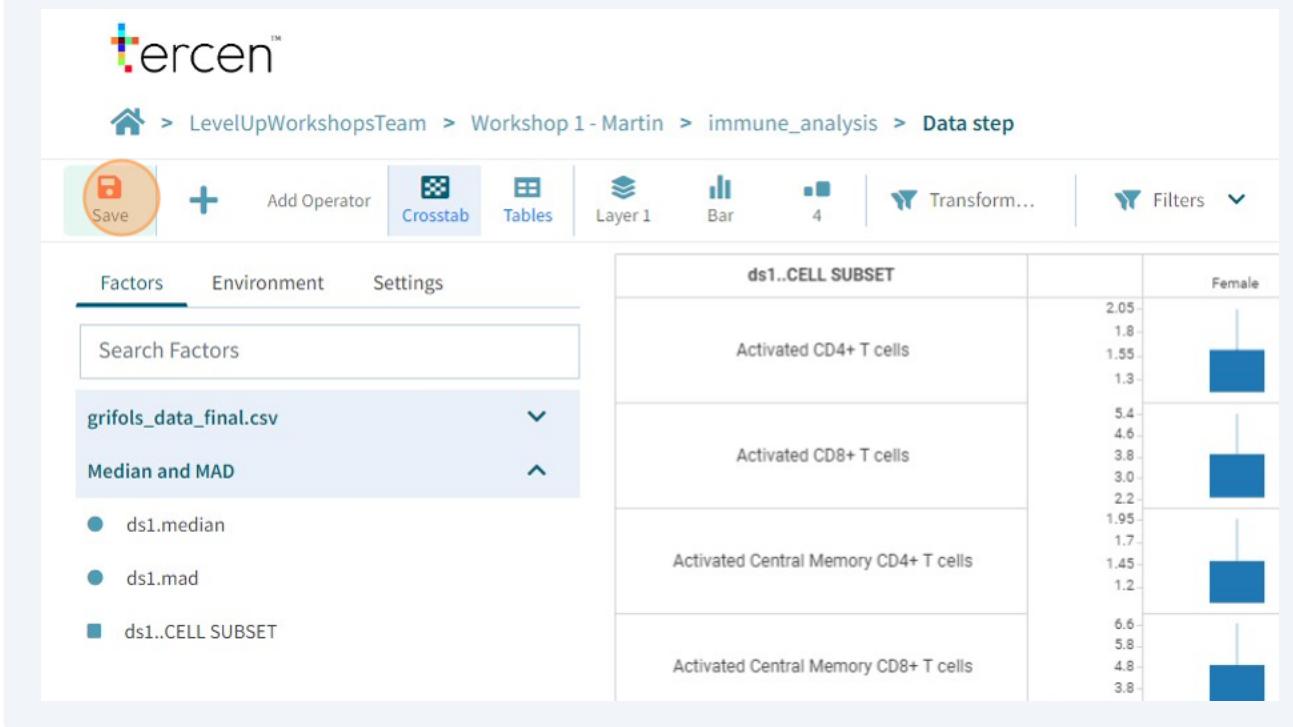
You can test this by changing the image on screen and downloading again.



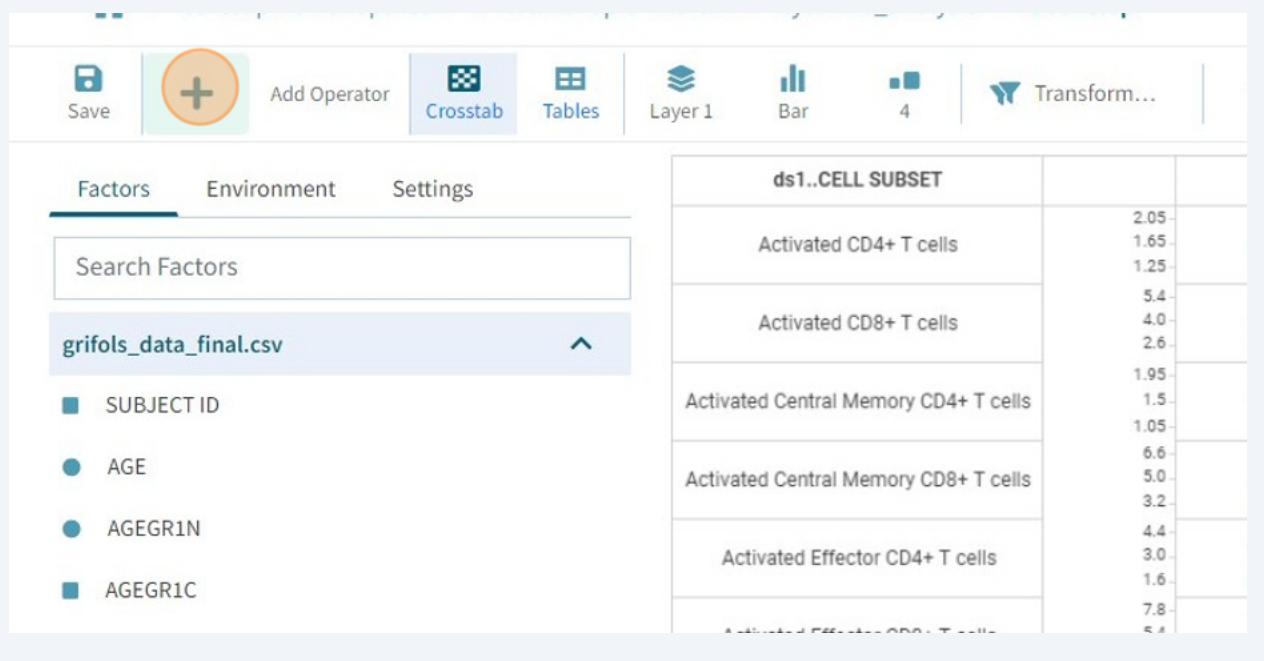
## 18 Save the Data Step.

Tercen uses operators to generate the more comprehensive or complex images for reports.

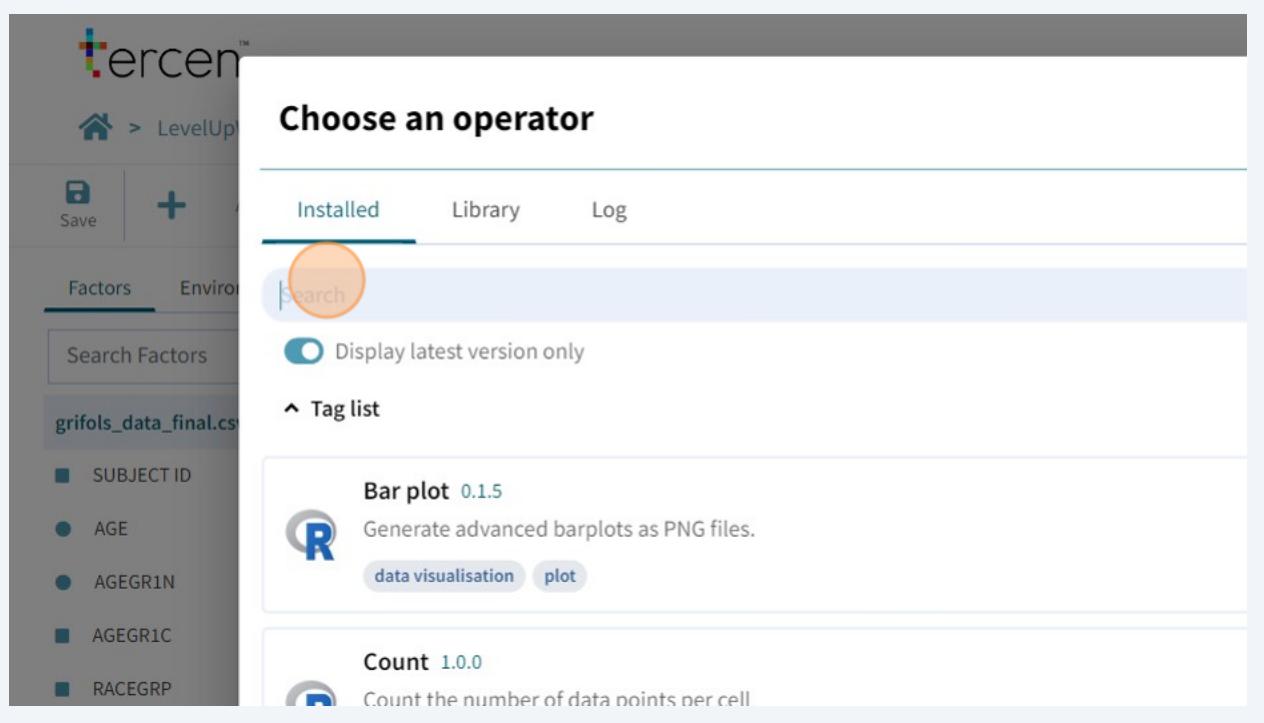
There are specialised operators for specialised graphs but the standard graph operator is called **Plot**.



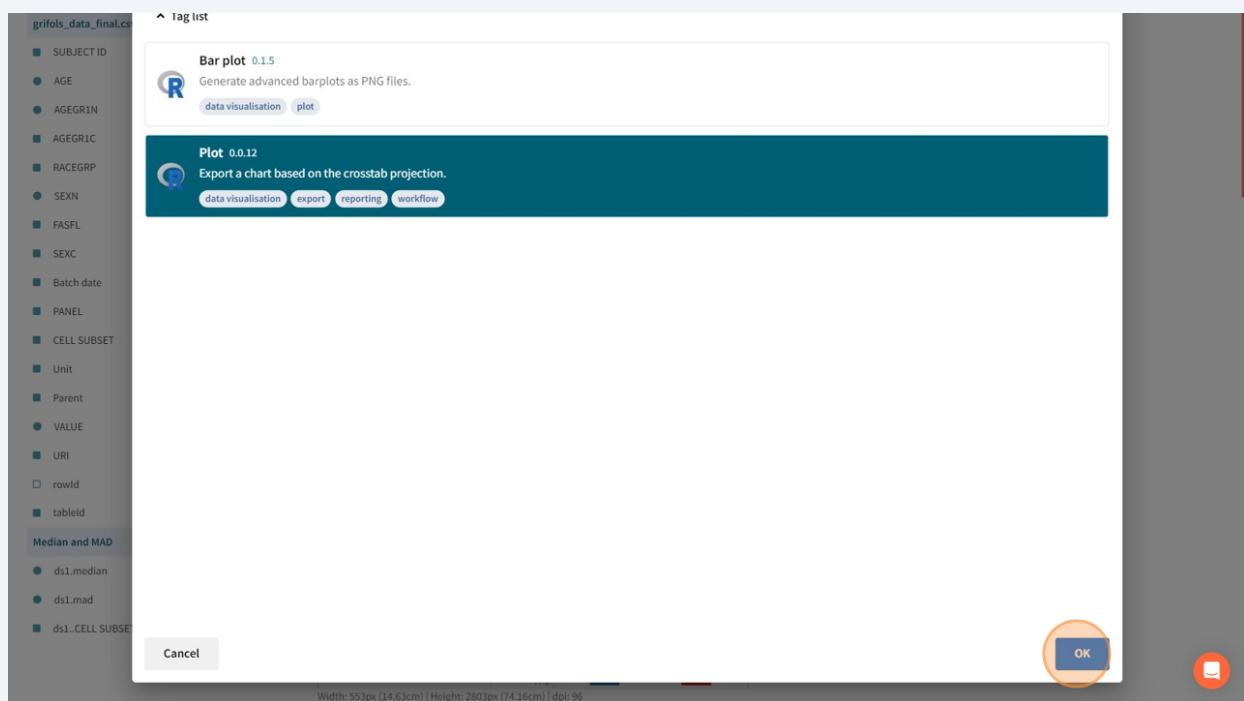
19 Click to add an operator,



20 Search for the "Plot" operator

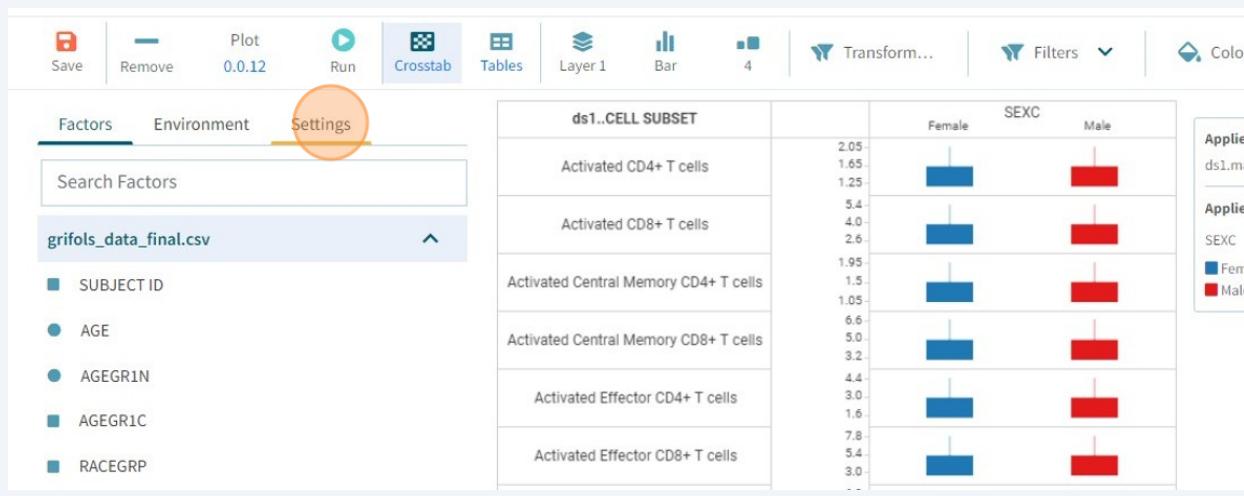


## 21 Click "OK"



22 Operators have settings to control how they generate a visualisation or make a calculation.

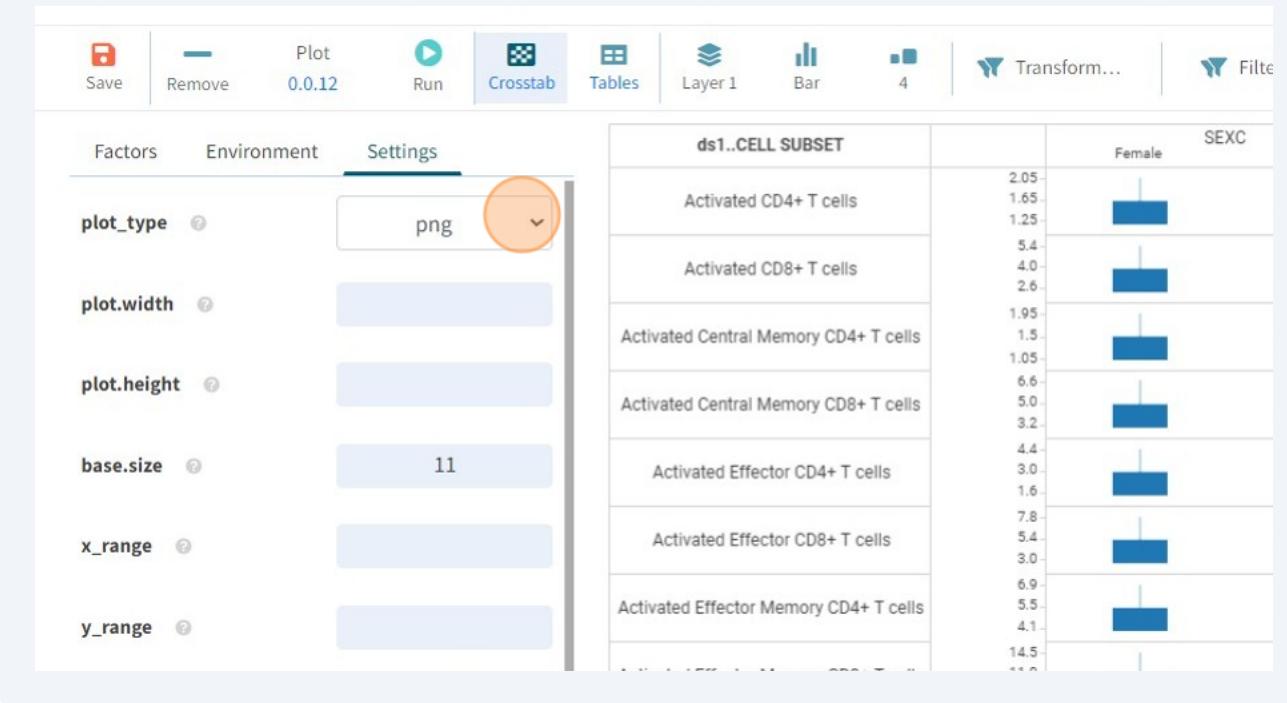
Click on the **Settings** tab.



**23** Click plot\_type to see the output file formats.

Plot outputs to PNG SVG and PDF files.

We will stay on .png for this example.



Hovering over the ? icon on a setting will give additional information about what it controls.

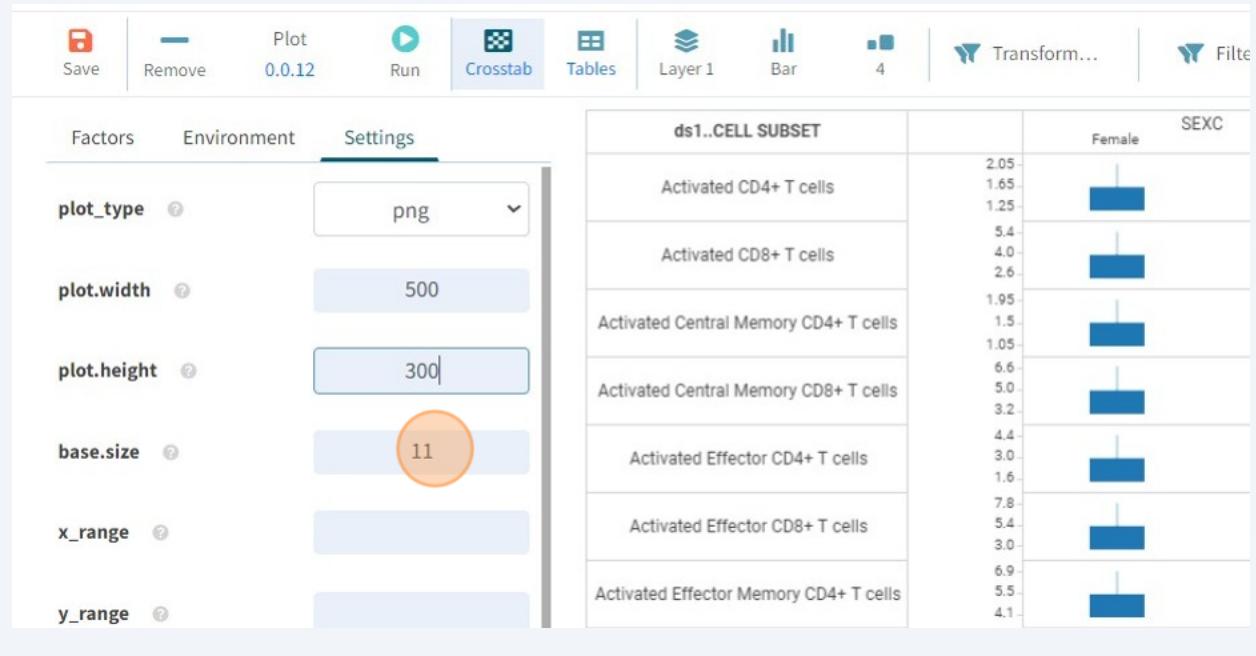
**24** **plot.width** and **plot.height** set the size of the downloaded image.

The size is calculated in **pixels**

Set our example to be 500 pixels wide and 300 pixels high.

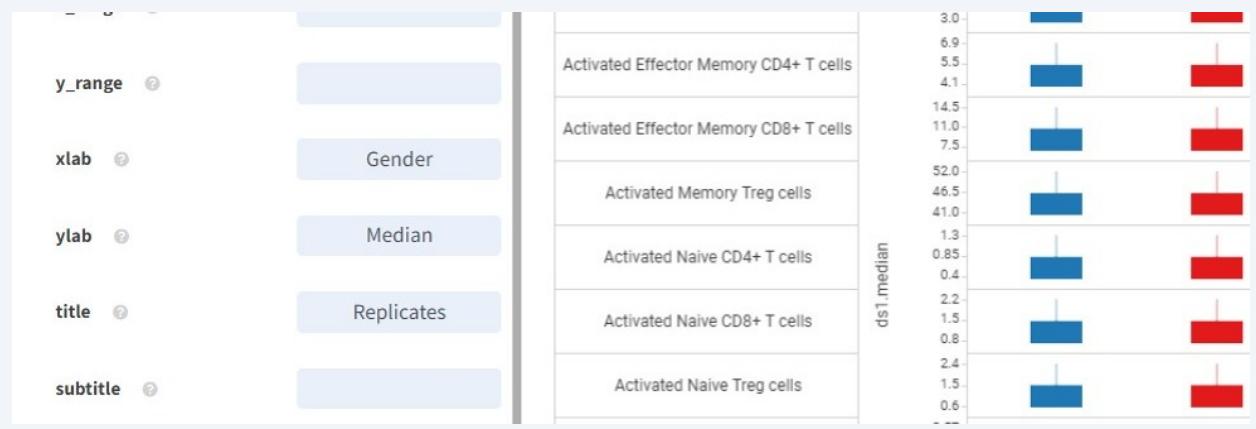
**base.size** is the font size for the text.

Change it to 10



**25** **xlab** and **ylab** will override the labels on the crostab projection with labels of your choosing

Set the **title** to be "Replicates"



**26** Run the operator.

It will generate a separate image for every **Data Cell** it sees on the Crosstab Grid.

Save Remove Plot 0.0.12 Run Crosstab Tables Layer 1 Bar

Factors Environment Settings

plot\_type ? png

plot.width ? 500

plot.height ? 300

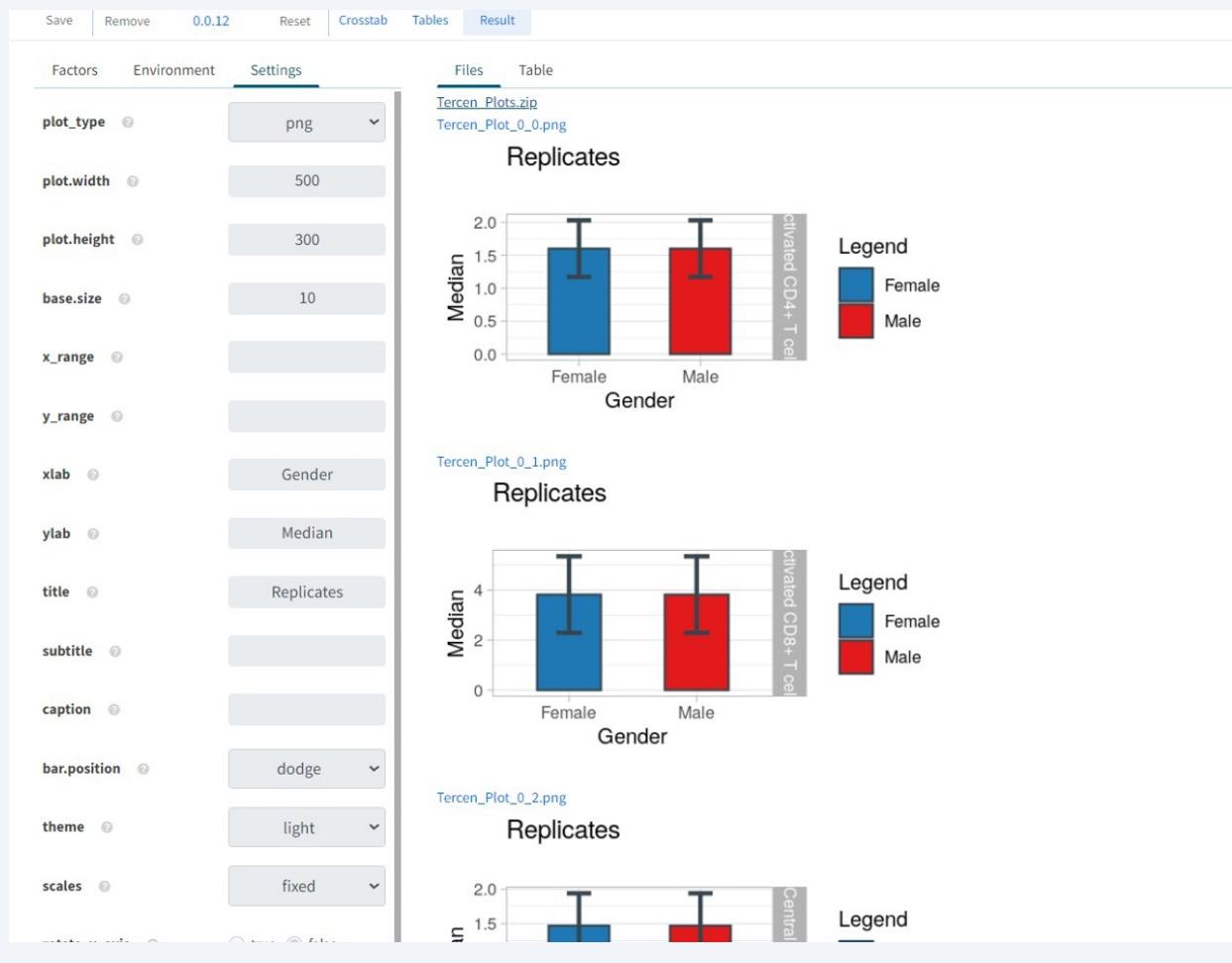
ds1..CELL SUBSET

- Activated CD4+ T cells
- Activated CD8+ T cells
- Activated Central Memory CD4+
- Activated Central Memory CD8+

**27** The images can be downloaded as a batch

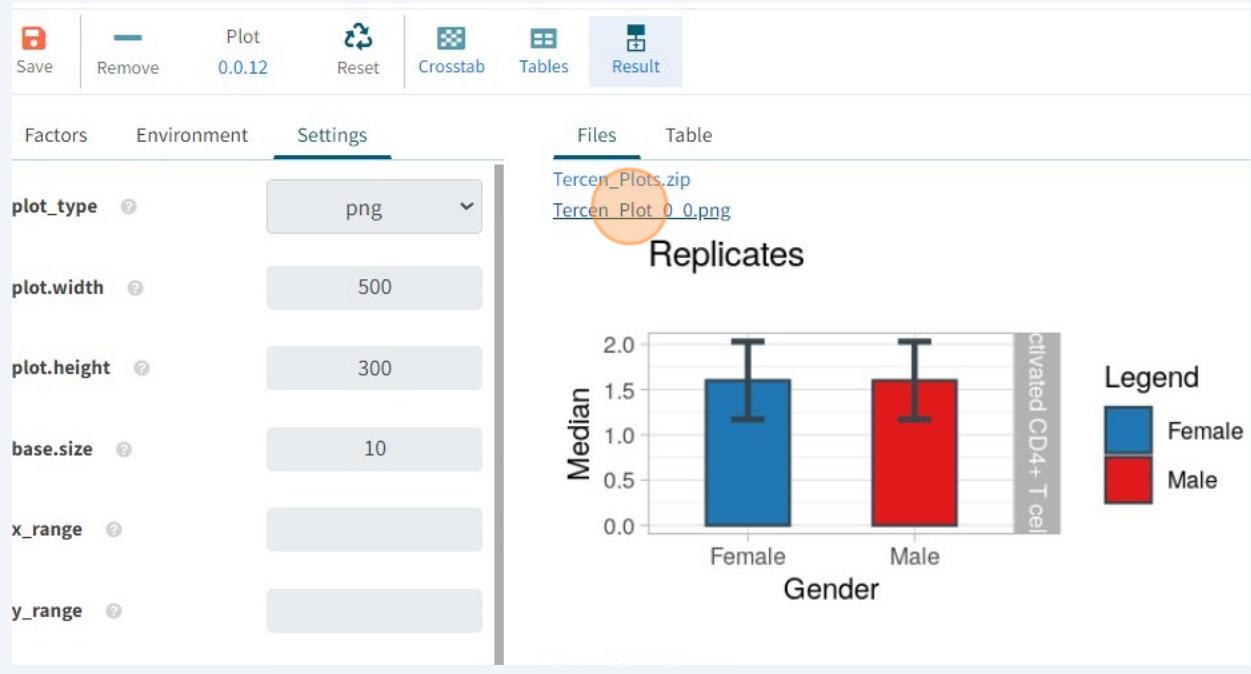
Click "Tercen\_Plots.zip"

Or

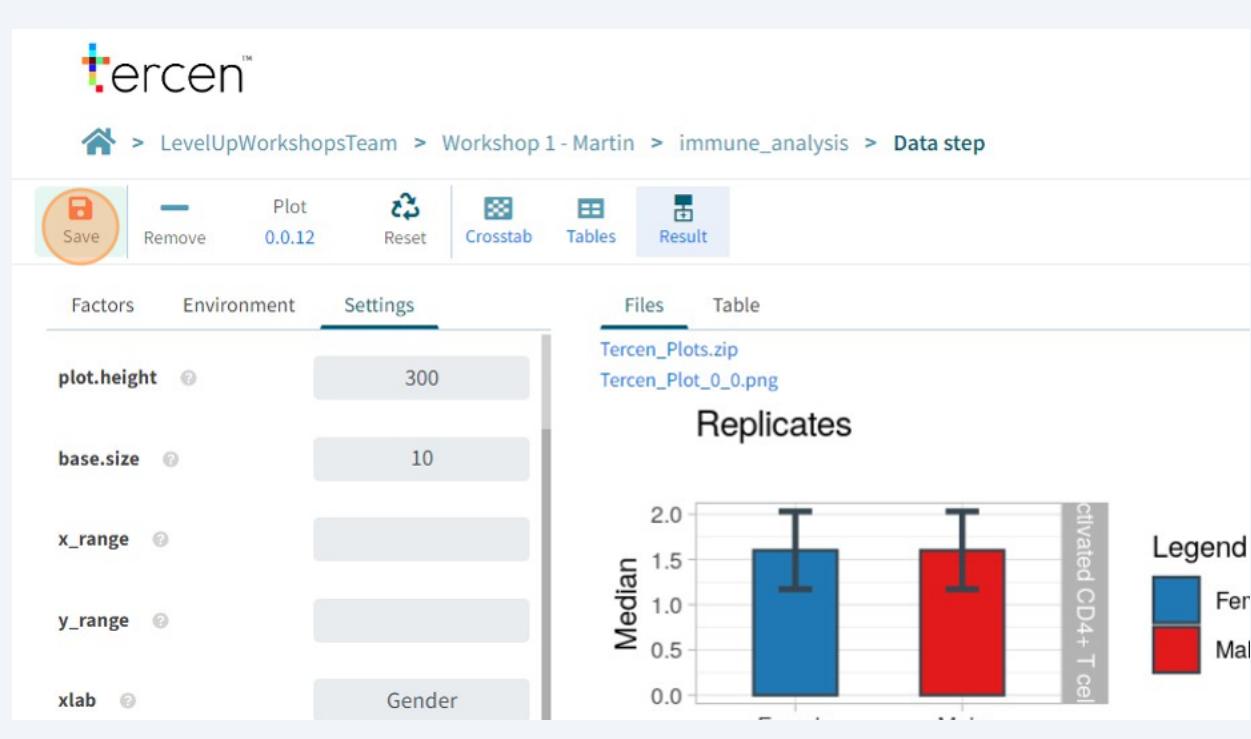


**28** Or be downloaded individually

Click "Tercen\_Plot\_0\_0.png"



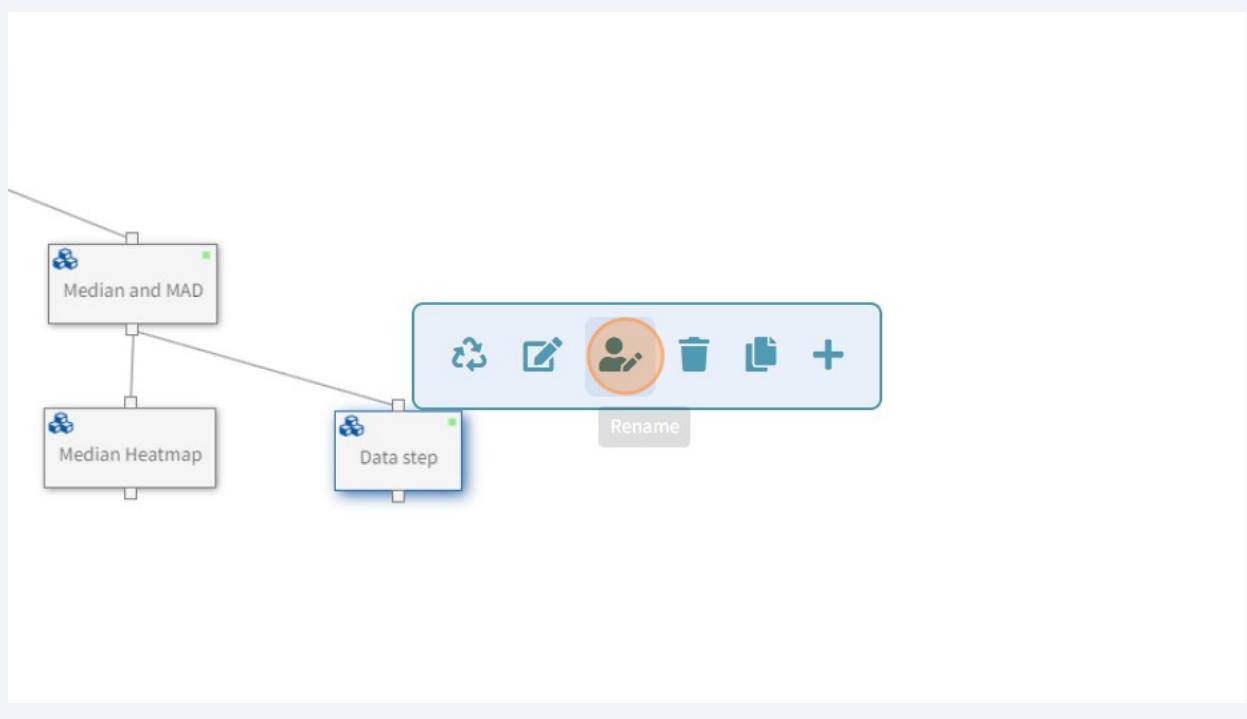
**29** Click "Save"



**30** Return to the "immune\_analysis" workflow canvas.



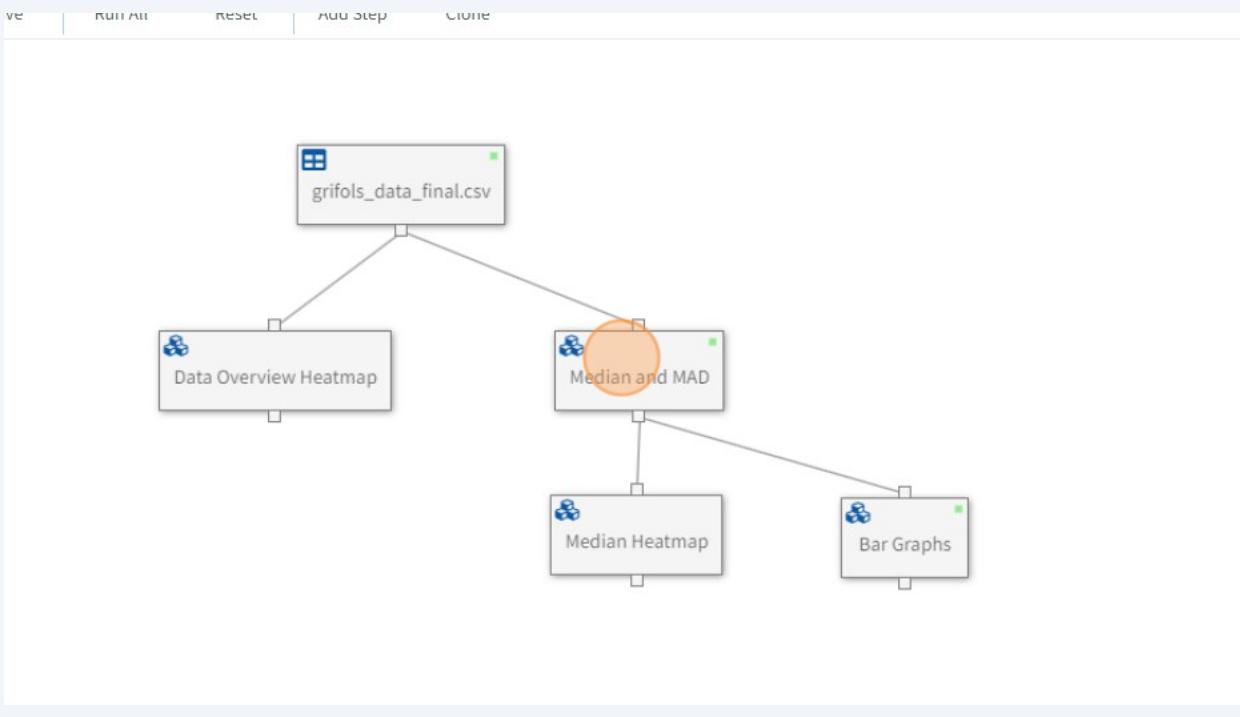
**31** Rename the Data Step to "Bar Graphs"



32

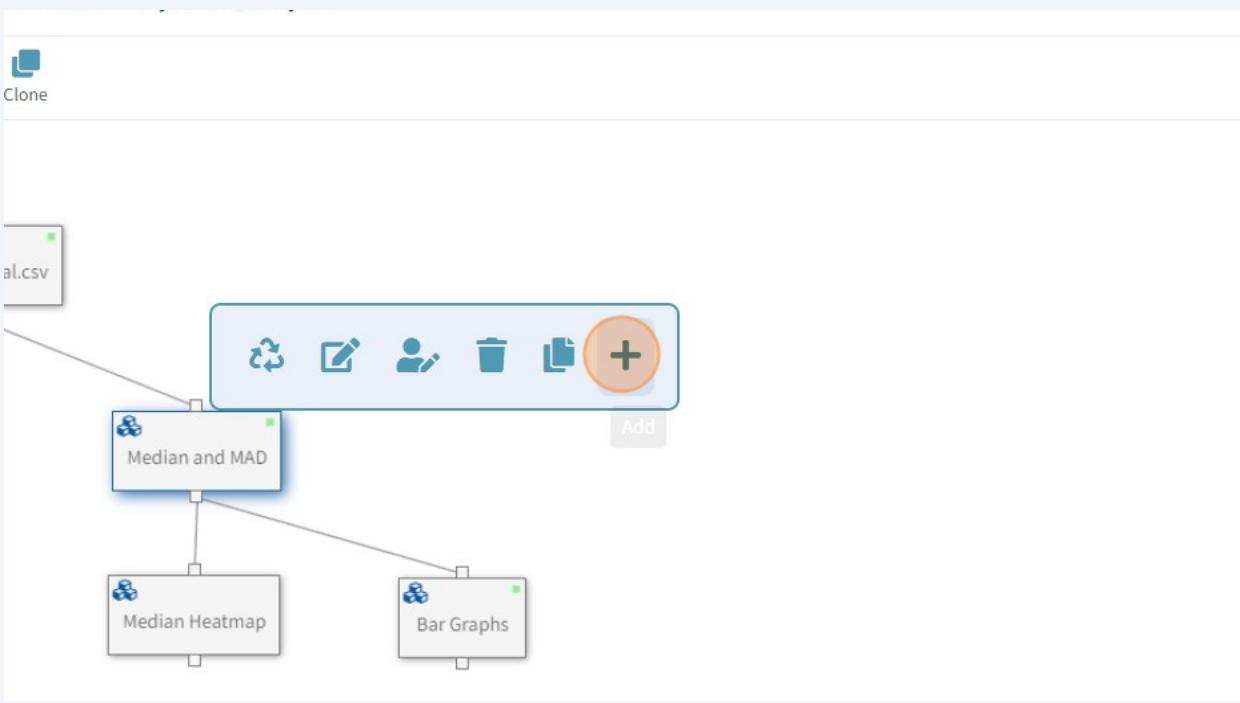
You have exported Tercen visualisations from the crosstab grid to be used in reports.

Next we will export a table with the new data generated by the Median and MAD operator.



33

Click Add on the local toolbar of the Median and MAD data step.



**34** Add a new data step.

The screenshot shows a software interface with a navigation bar at the top containing tabs: Step (which is selected), Operator, Operator Library, Installed Apps, and App Library. Below the navigation bar is a search bar labeled "Search". The main content area displays three data step options:

- Data step** data  
Perform computation on user defined projection
- Multi data step** data  
Perform computation on user defined projection
- Join** leftTable  
Join two data sets

An orange circle highlights the "Data step" option.

**35** Clear the system generate projection.

Press OK

The screenshot shows a data visualization interface with a toolbar at the top featuring various icons and buttons. One button, labeled "Clear", is circled in orange. The main area displays a scatter plot with data points. A legend on the right side indicates that the data is colored by "ds1.median".

**36** Make a projection of the data we want to export.

Drag ds1.CELL SUBSET to row

The screenshot shows a software interface for data analysis. On the left, there's a sidebar titled 'Factors' with tabs for 'Environment' and 'Settings'. Below the tabs is a search bar labeled 'Search Factors'. Underneath the search bar is a dropdown menu set to 'grifols\_data\_final.csv'. Within this dropdown, there's a section titled 'Median and MAD' with two items: 'ds1.median' and 'ds1.mad'. At the bottom of the dropdown, there's a blue square icon followed by the text 'ds1..CELL SUBSET'. A large orange circle highlights this 'ds1..CELL SUBSET' item. To the right of the sidebar is a large, empty grid area with three columns and four rows, intended for data export.



When building data table exports drag all factors to the **Row**.

This is a case where that Y-Axis is not used.

### 37 Drag ds1.Median to row

The screenshot shows a software interface for data analysis. At the top, there are several tabs: 'Save', 'Add Operator', 'Crosstab' (which is selected), 'Tables', 'Layer 1', 'Point', a numerical field '4', 'Transform...', and 'Filters'. Below these are three tabs: 'Factors' (selected), 'Environment', and 'Settings'. A search bar labeled 'Search Factors' contains the text 'grifols\_data\_final.csv'. Under the 'Median and MAD' section, there are three items: 'ds1.median' (highlighted with an orange circle), 'ds1.mad', and 'ds1..CELL SUBSET'. To the right, a table titled 'ds1..CELL SUB' has two rows: 'Activated CD4+ T cells' and 'Activated CD8+ T cells'. The table has four columns, with the first column header being 'ds1..CELL SUB'.

### 38 Drag ds1.mad to row.

This screenshot shows the same software interface as the previous one. The 'Median and MAD' section now has 'ds1.mad' selected (highlighted with an orange circle), while 'ds1.median' is no longer highlighted. The rest of the interface, including the table on the right, remains the same.

## 39 A simple table is built on the crosstab grid.

Click the Tables button

The screenshot shows the Tercen interface with the 'Tables' button highlighted by a yellow circle. The interface includes a toolbar with 'Save', 'Add Operator', 'Crosstab', 'Tables', 'Layer 1', 'Point', '4', 'Transform...', 'Filters', 'Colors', 'Labels', and 'Error Bar'. Below the toolbar, there are tabs for 'Factors', 'Environment', and 'Settings', with 'Factors' selected. A search bar for 'Search Factors' is present. A dropdown menu for 'grifols\_data\_final.csv' is open, showing 'Median and MAD' with three items: 'ds1.median', 'ds1.mad', and 'ds1..CELL SUBSET'. The main area displays a table with three columns: 'ds1..CELL SUBSET', 'ds1.median', and 'ds1.mad'. The data rows are: Activated CD4+ T cells (1.6, 0.43), Activated CD8+ T cells (3.82, 1.53), Activated Central Memory CD4+ T cells (1.47, 0.47), Activated Central Memory CD8+ T cells (4.63, 1.97), Activated Effector CD4+ T cells (2.84, 1.52), Activated Effector CD8+ T cells (5.13, 2.57), Activated Effector Memory CD4+ T cells (5.32, 1.5), Activated Effector Memory CD8+ T cells (10.45, 3.84), and Activated Memory Treg cells (45.9, 5.8).

ds1..CELL SUBSET	ds1.median	ds1.mad
Activated CD4+ T cells	1.6	0.43
Activated CD8+ T cells	3.82	1.53
Activated Central Memory CD4+ T cells	1.47	0.47
Activated Central Memory CD8+ T cells	4.63	1.97
Activated Effector CD4+ T cells	2.84	1.52
Activated Effector CD8+ T cells	5.13	2.57
Activated Effector Memory CD4+ T cells	5.32	1.5
Activated Effector Memory CD8+ T cells	10.45	3.84
Activated Memory Treg cells	45.9	5.8

## 40 Click the "Row" tab.

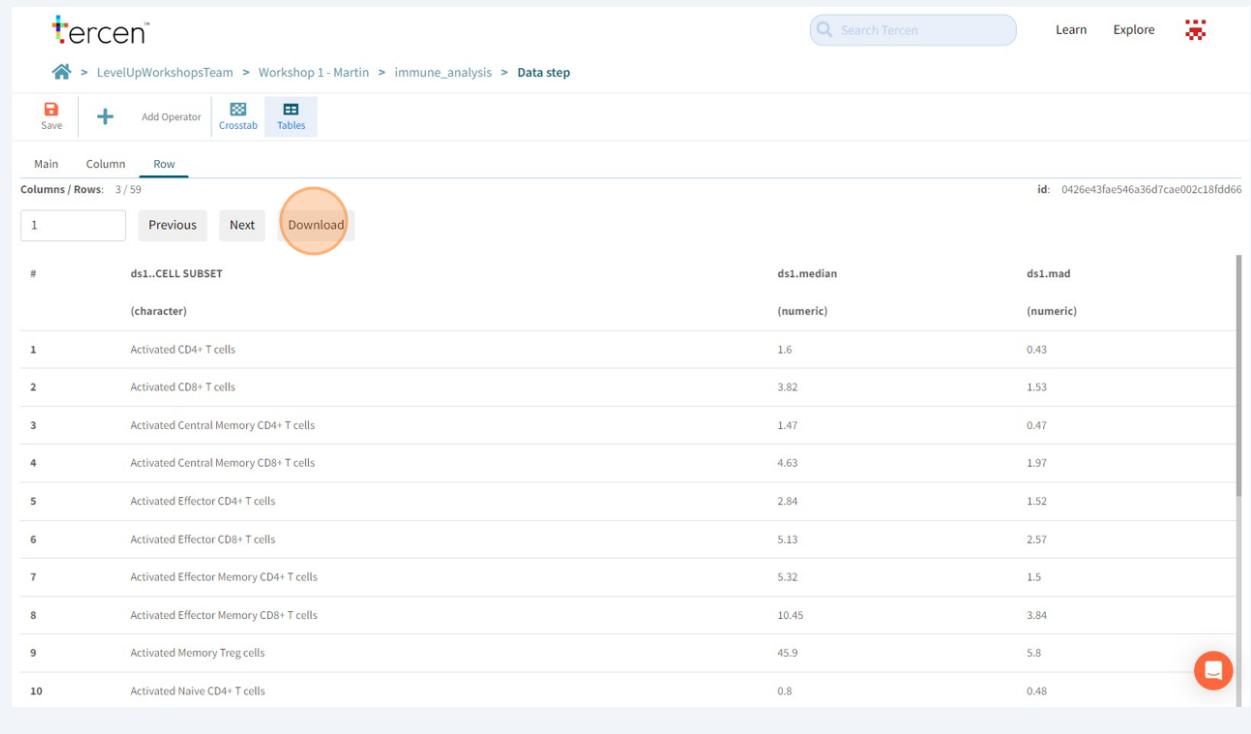
The screenshot shows the Tercen interface with the 'Row' tab highlighted by a yellow circle. The interface includes a toolbar with 'Save', 'Add Operator', 'Crosstab', 'Tables', 'Main', 'Column', and 'Row'. Below the toolbar, there is a 'Columns / Rows' section showing '3 / 59'. The main area displays a table with four columns: '#', 'ds1..CELL SUBSET', 'ds1.median', and 'ds1.mad'. The data rows are: Activated CD4+ T cells (1, 1.6, 0.43), Activated CD8+ T cells (2, 3.82, 1.53), Activated Central Memory CD4+ T cells (3, 1.47, 0.47), Activated Central Memory CD8+ T cells (4, 4.63, 1.97), Activated Effector CD4+ T cells (5, 2.84, 1.52), Activated Effector CD8+ T cells (6, 5.13, 2.57), Activated Effector Memory CD4+ T cells (7, 5.32, 1.5), Activated Effector Memory CD8+ T cells (8, 10.45, 3.84), Activated Memory Treg cells (9, 45.9, 5.8), and Activated Naive CD4+ T cells (10, 0.8, 0.48). A red circular icon with a white arrow is located in the bottom right corner of the table area.

#	ds1..CELL SUBSET	ds1.median	ds1.mad
1	Activated CD4+ T cells	1.6	0.43
2	Activated CD8+ T cells	3.82	1.53
3	Activated Central Memory CD4+ T cells	1.47	0.47
4	Activated Central Memory CD8+ T cells	4.63	1.97
5	Activated Effector CD4+ T cells	2.84	1.52
6	Activated Effector CD8+ T cells	5.13	2.57
7	Activated Effector Memory CD4+ T cells	5.32	1.5
8	Activated Effector Memory CD8+ T cells	10.45	3.84
9	Activated Memory Treg cells	45.9	5.8
10	Activated Naive CD4+ T cells	0.8	0.48

## 41 Click "Download"

A CSV file with your selected data will be generated and downloaded to your local drive.

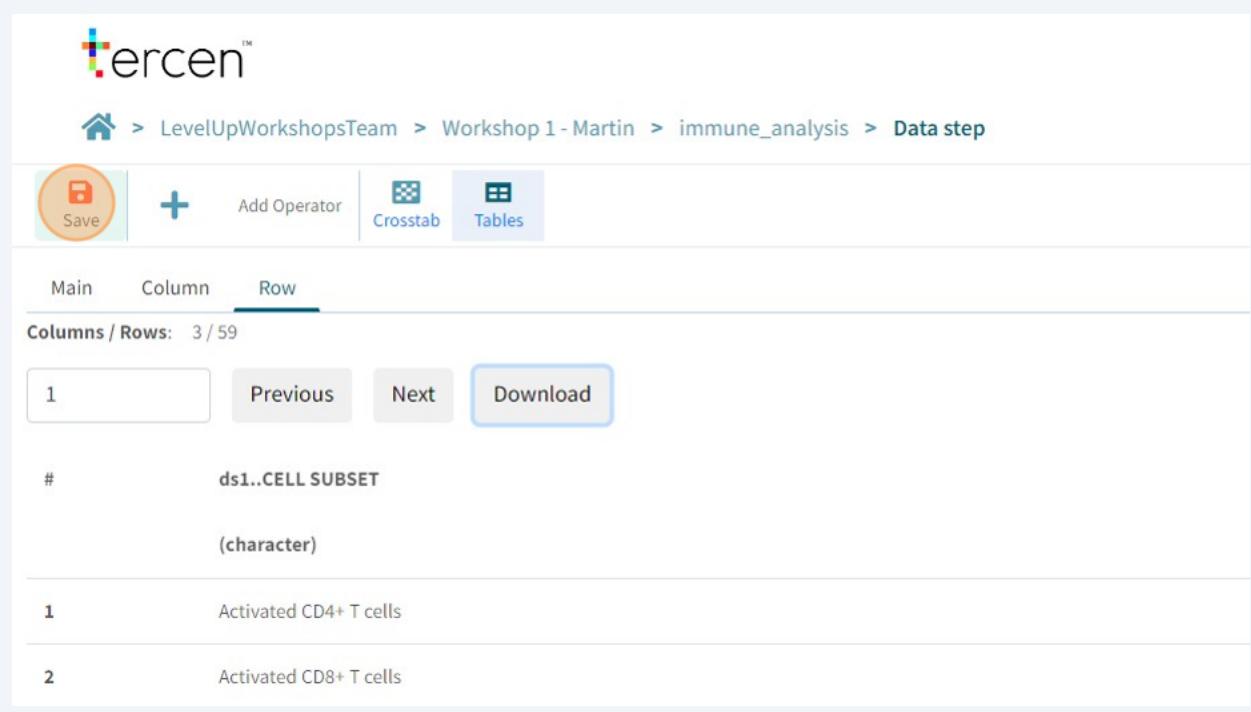
Open it in Excel to review the output.



The screenshot shows a Tercen interface with a search bar and navigation links. Below is a table titled 'Data step' with columns: '#', 'ds1..CELL SUBSET', 'ds1.median', and 'ds1.mad'. The table lists 10 rows of activated T cell subsets. A 'Download' button at the top right of the table is highlighted with an orange circle. The URL in the address bar is 'https://app.tercen.com/LevelUpWorkshopsTeam/Workshop 1 - Martin/immune\_analysis/Data step'.

#	ds1..CELL SUBSET	ds1.median	ds1.mad
1	Activated CD4+ T cells	1.6	0.43
2	Activated CD8+ T cells	3.82	1.53
3	Activated Central Memory CD4+ T cells	1.47	0.47
4	Activated Central Memory CD8+ T cells	4.63	1.97
5	Activated Effector CD4+ T cells	2.84	1.52
6	Activated Effector CD8+ T cells	5.13	2.57
7	Activated Effector Memory CD4+ T cells	5.32	1.5
8	Activated Effector Memory CD8+ T cells	10.45	3.84
9	Activated Memory Treg cells	45.9	5.8
10	Activated Naive CD4+ T cells	0.8	0.48

## 42 Click "Save"



The screenshot shows the same Tercen interface as above, but the 'Save' button in the top navigation bar is highlighted with an orange circle. The table below is identical to the one in the previous screenshot, showing activated T cell subsets. The URL in the address bar is 'https://app.tercen.com/LevelUpWorkshopsTeam/Workshop 1 - Martin/immune\_analysis/Data step'.

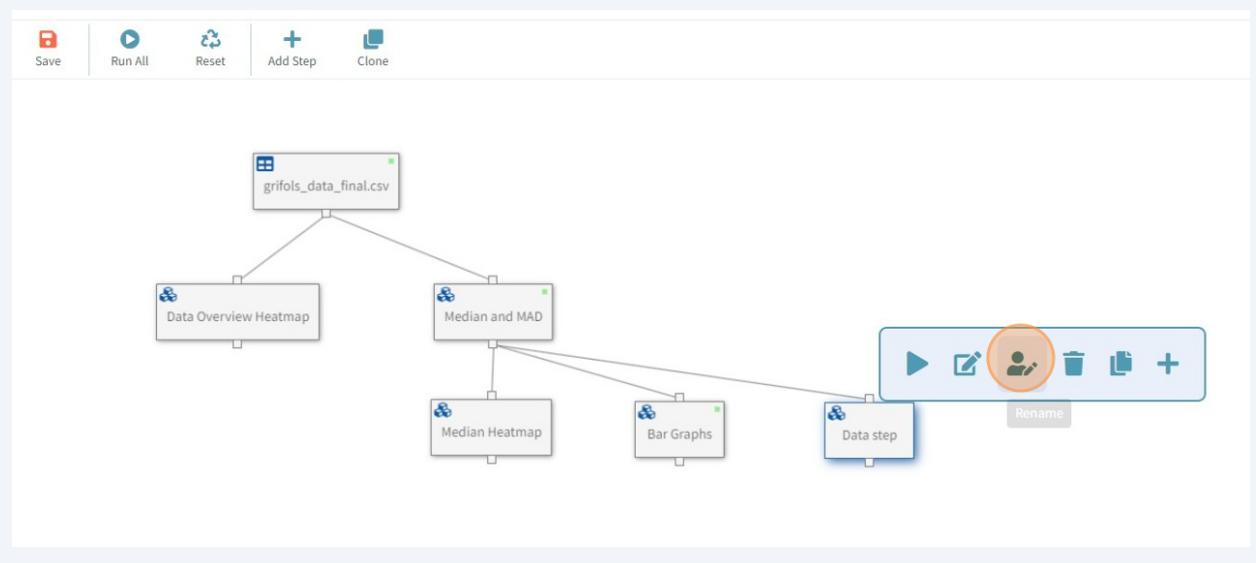
#	ds1..CELL SUBSET	ds1.median	ds1.mad
1	Activated CD4+ T cells	1.6	0.43
2	Activated CD8+ T cells	3.82	1.53
3	Activated Central Memory CD4+ T cells	1.47	0.47
4	Activated Central Memory CD8+ T cells	4.63	1.97
5	Activated Effector CD4+ T cells	2.84	1.52
6	Activated Effector CD8+ T cells	5.13	2.57
7	Activated Effector Memory CD4+ T cells	5.32	1.5
8	Activated Effector Memory CD8+ T cells	10.45	3.84
9	Activated Memory Treg cells	45.9	5.8
10	Activated Naive CD4+ T cells	0.8	0.48

**43** Return to the "immune\_analysis" workflow canvas.

The screenshot shows the Sercen workflow canvas interface. At the top, there is a breadcrumb navigation: > LevelUpWorkshopsTeam > Workshop 1 - Martin > immune\_analysis > Data step. A large orange circle highlights the 'Data step' link in the breadcrumb. Below the navigation, there are three buttons: 'Add Operator' (with a plus sign), 'Crosstab' (with a grid icon), and 'Tables' (with a table icon). The 'Tables' button is selected. There are two tabs at the top: 'Column' and 'Row', with 'Row' being the active tab. Below the tabs, it says 'Rows: 3 / 59'. There are three buttons: 'Previous', 'Next', and 'Download'. The main area displays a table with three rows:

ds1..CELL SUBSET	(character)
Activated CD4+ T cells	1.
Activated CD8+ T cells	3.

**44** Rename your data step to "Export Table"



**45** Don't forget to Save your project.

