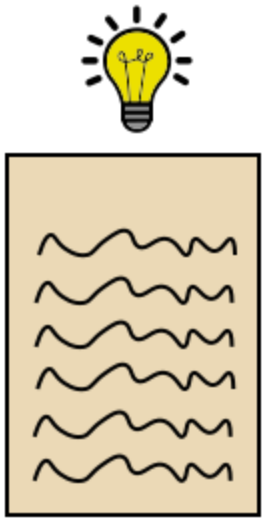
Scratch Data Tools Extension



User Documentation

Getting Started

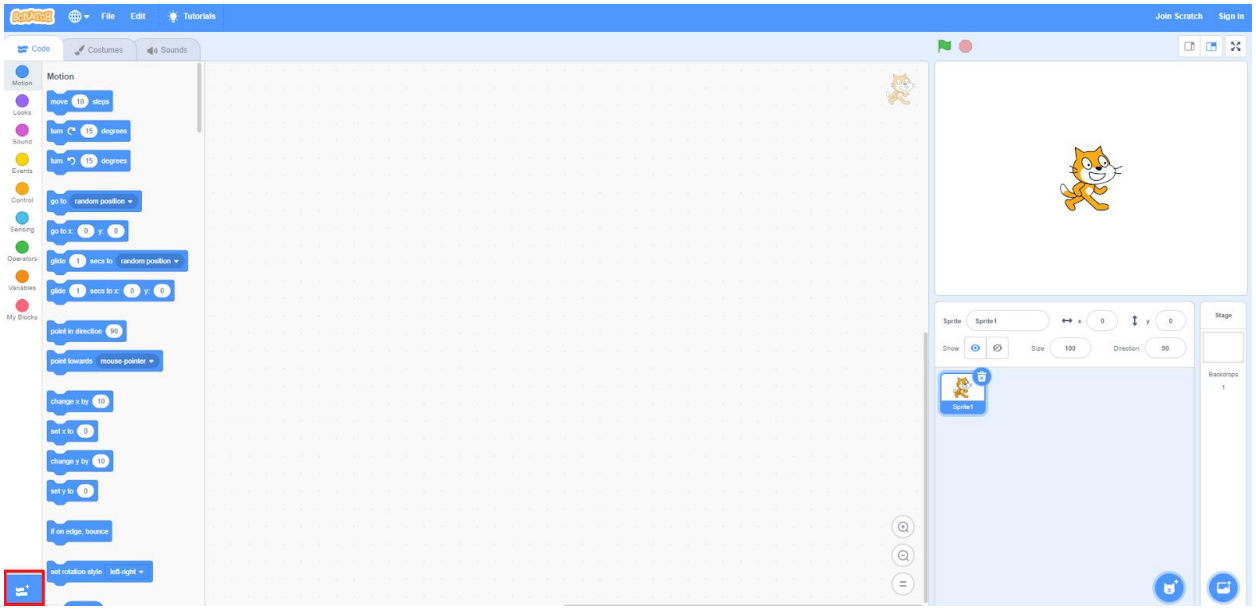
Currently, the only way to use the extension is to host Scratch locally using the modified version of Scratch 3.0 hosted on github by the KSU Scratch Data Tools team. For those technologically inclined, you can find the repositories here:
https://github.com/papaphil/DataTools-scratch-vm
https://github.com/papaphil/DataTools-scratch-gui
https://github.com/papaphil/DataTools-scratch-blocks
Simply install scratch locally as you would scratch, but using these repositories. Instructions on how to install Scratch locally are maintained by the LLK developers and can be found here:
https://github.com/LLK/scratch-gui/wiki/Getting-Started

In the future, the KSU Scratch Data Tools team hopes to host the modified version of Scratch publicly on the Kansas State University network. For the time being if you wish to use scratch you must go to https://scratch.mit.edu/.

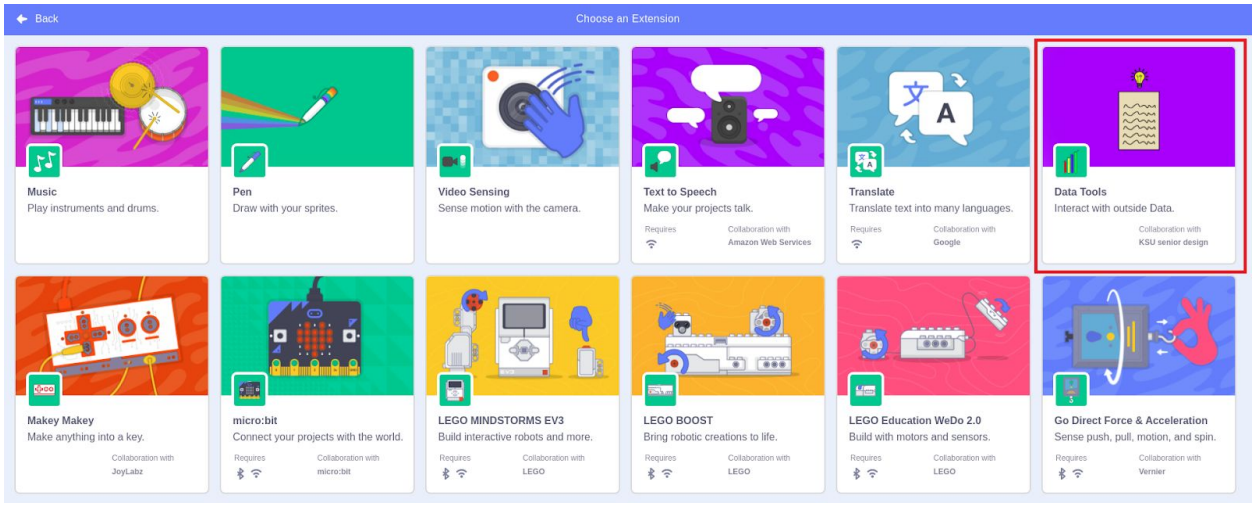
This is a living document and serves to illustrate the current features added to the in development Scratch Data Tools extension, and how a user might use them.

Adding the Extension

To add the extension, first click on the “Add Extension” button in the bottom left.



Click on the “Data Tools” extension to add it to your scratch project.



When you first add the extension, it will prompt you to upload a dataset file. See the “Upload a Dataset File” section for more information.

Upload a Dataset File

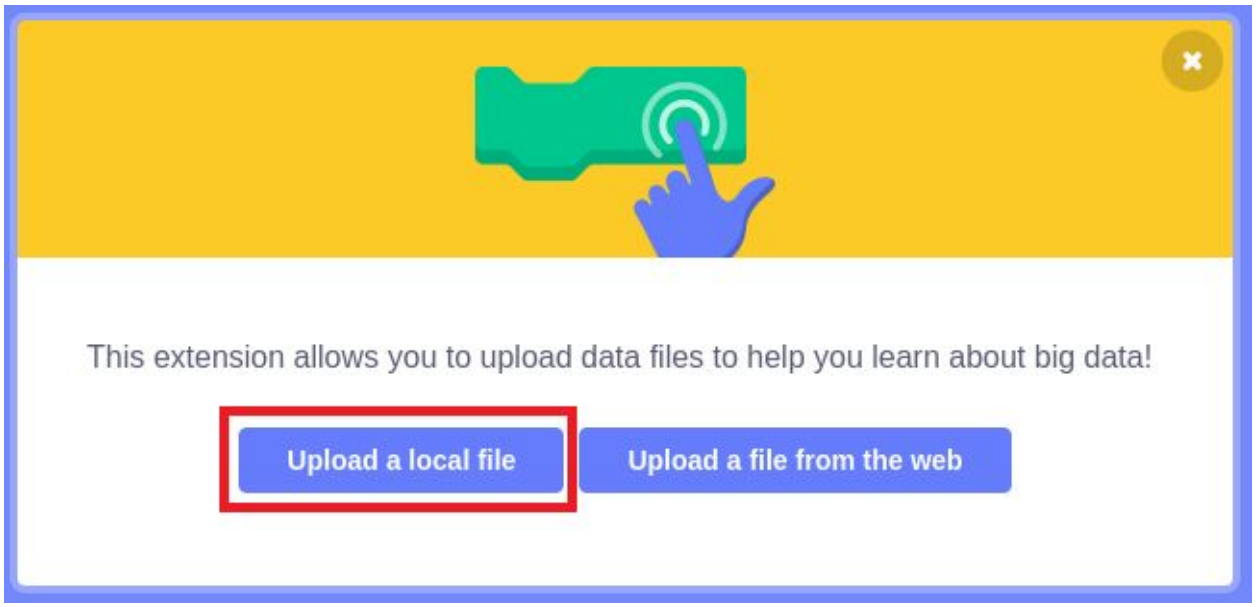
Before you can use the extension, you will need a dataset to work with. To do this, you must upload the dataset to your scratch project. There are two ways to do this, from a local file on your computer, or with a link to a file on the web.

Upload a Local File

There are two methods you can use to upload a local dataset file. One way is to use the popup when you first add the extension, the other is to use the upload button.

Using the Popup

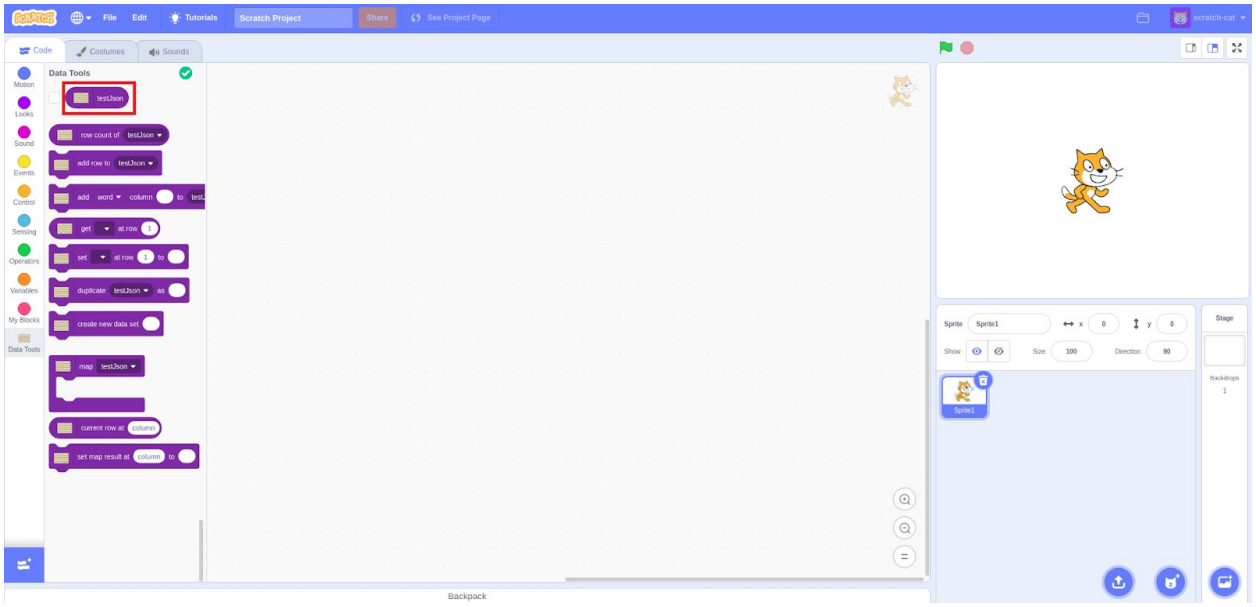
When you first add the extension it will create a popup that will prompt you to upload a dataset file. To upload a local dataset file, click on the “Upload a local file” button.



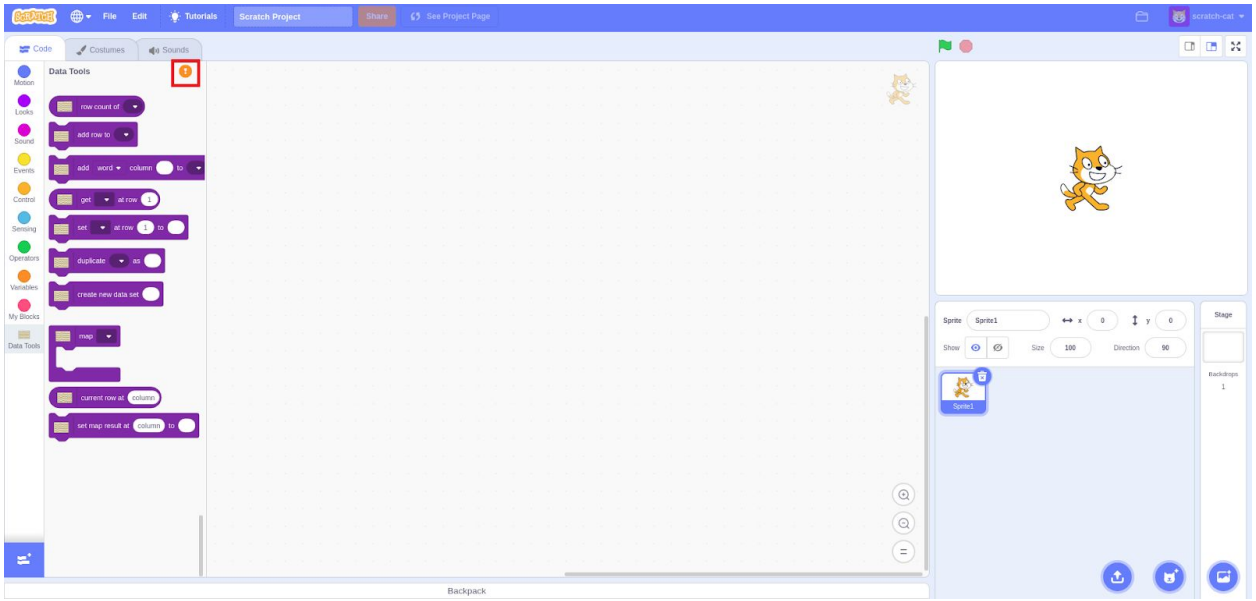
It will open your file explorer. Navigate through your computer to where the dataset file you want to upload is, then select it for upload.

Now click on “Open” to upload the file.

The dataset should now be loaded into your project, and you should see a new block for the dataset.



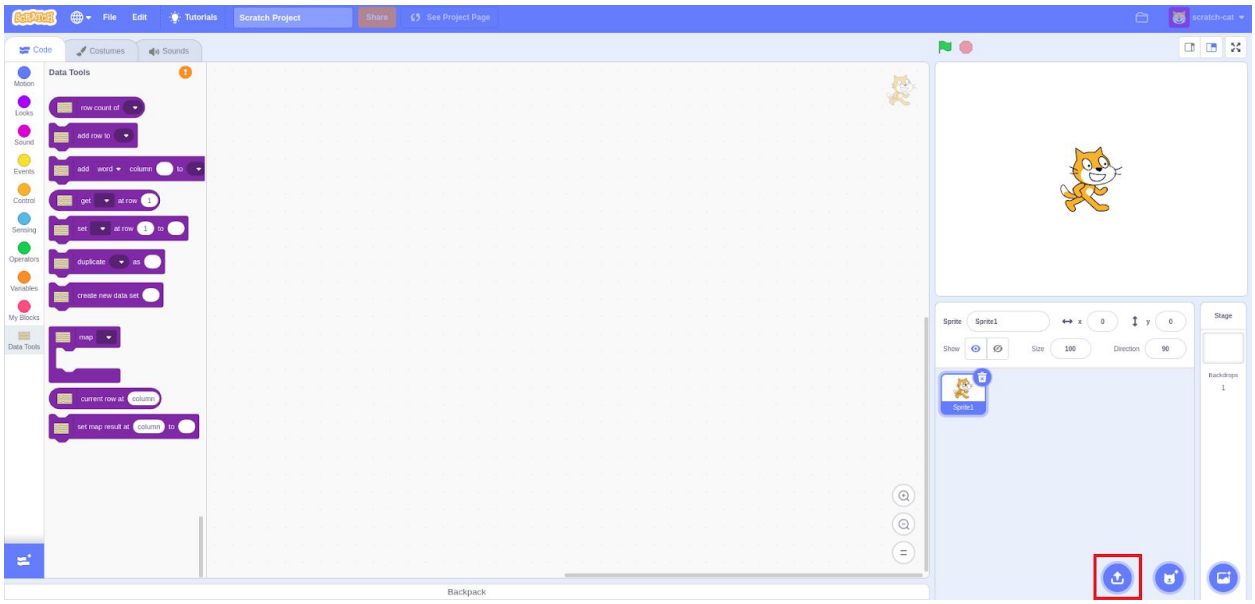
If you accidentally close the popup or want to open it again, you can click on this button.



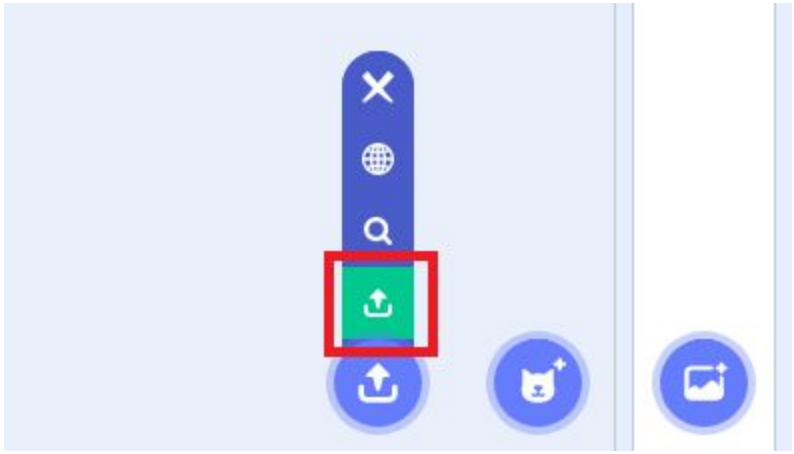
The button will change from an orange exclamation mark into a teal check mark once you have uploaded at least one dataset file, but will still reopen the popup.

Using the Button

Once you have added the extension to your scratch project, there should be a new button in the bottom right of scratch.



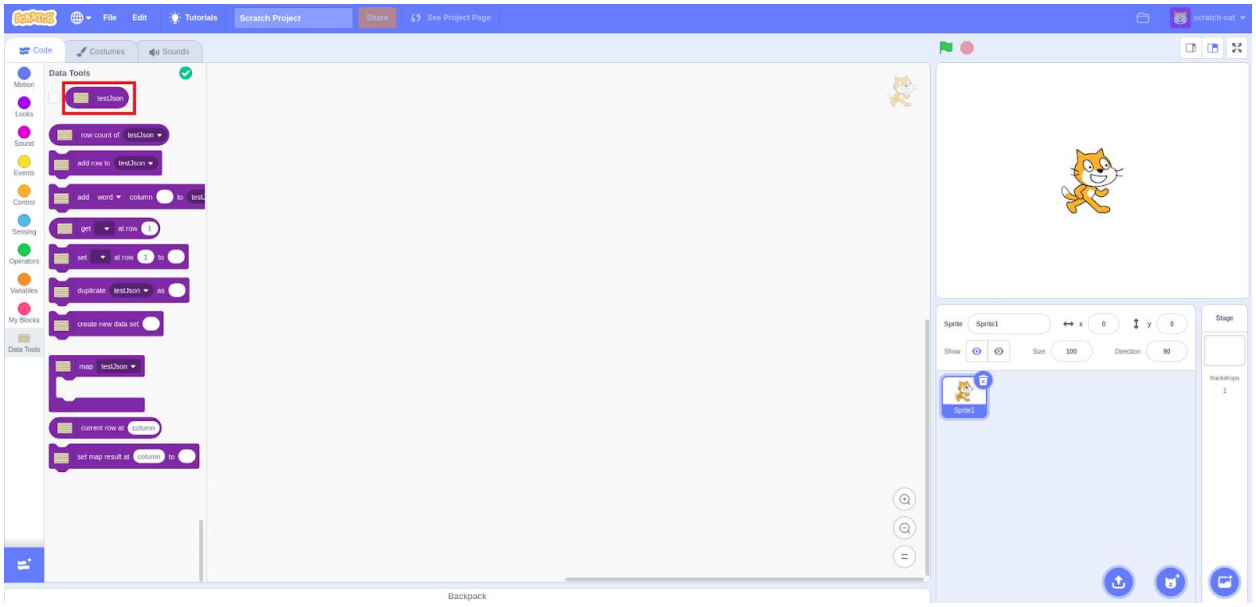
To upload a local file, move your cursor over the button and then click on the “Upload Local File” option.



It will open your file explorer. Navigate through your computer to where the dataset file you want to upload is, then select it for upload.

Now click on “Open” to upload the file.

The dataset should now be loaded into your project, and you should see a new block for the dataset.



Upload a File from the Web

This feature is currently not fully functioning and won't be for the foreseeable future. For now, please upload your dataset file from your local computer.

There are two methods you can use to upload a dataset file from the web. One way is to use the popup when you first add the extension, the other is to use the upload button.

Using the Popup

When you first add the extension it will create a popup that will prompt you to upload a dataset file.

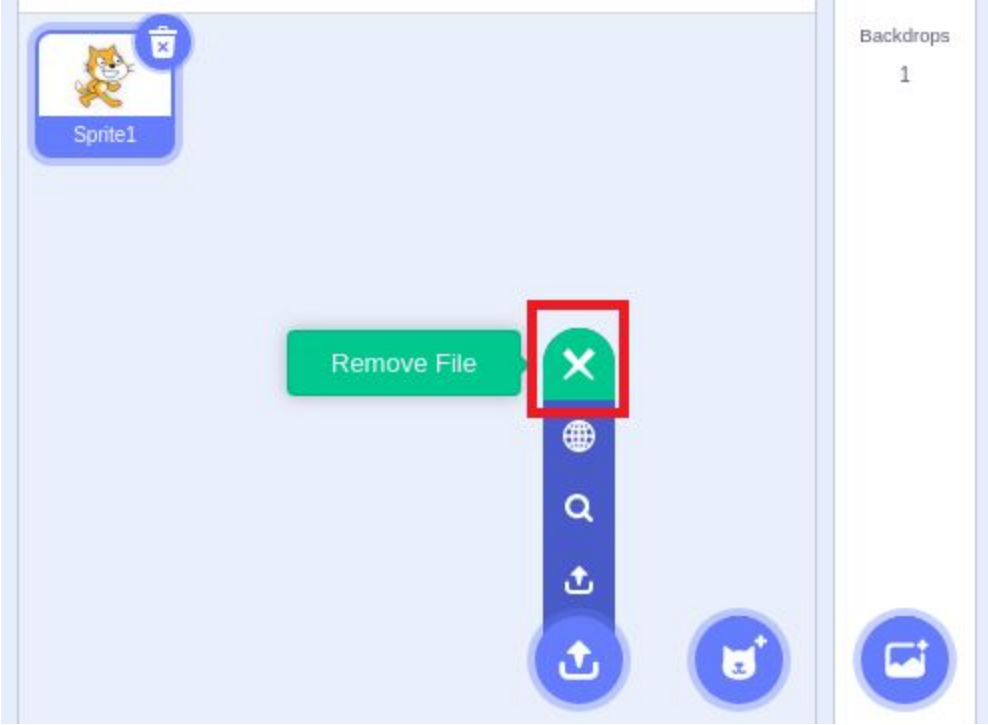
Using the Button

Once you have added the extension to your scratch project, there should be a new button in the bottom right of scratch.

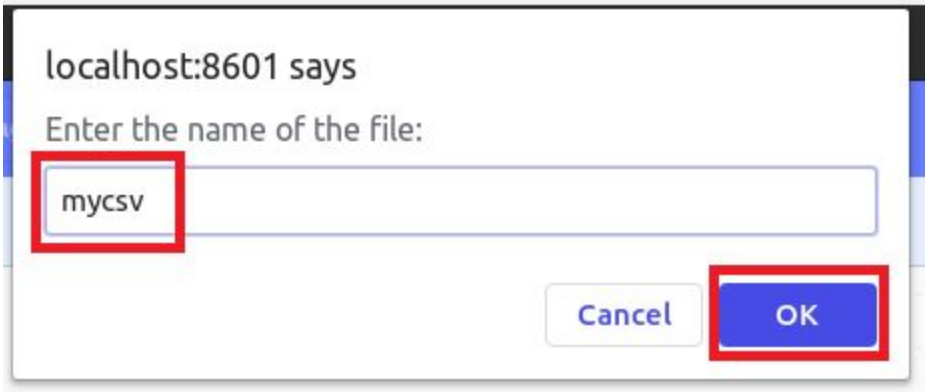
To upload a file from the web, move your cursor over the button and then click on the "Upload Web File" option.

Removing a Dataset

Once you have uploaded a dataset file, you might want to remove it. In order to remove it, you must move your cursor over the new button added by the extension in the bottom right of scratch, and click on the “Remove File” option.

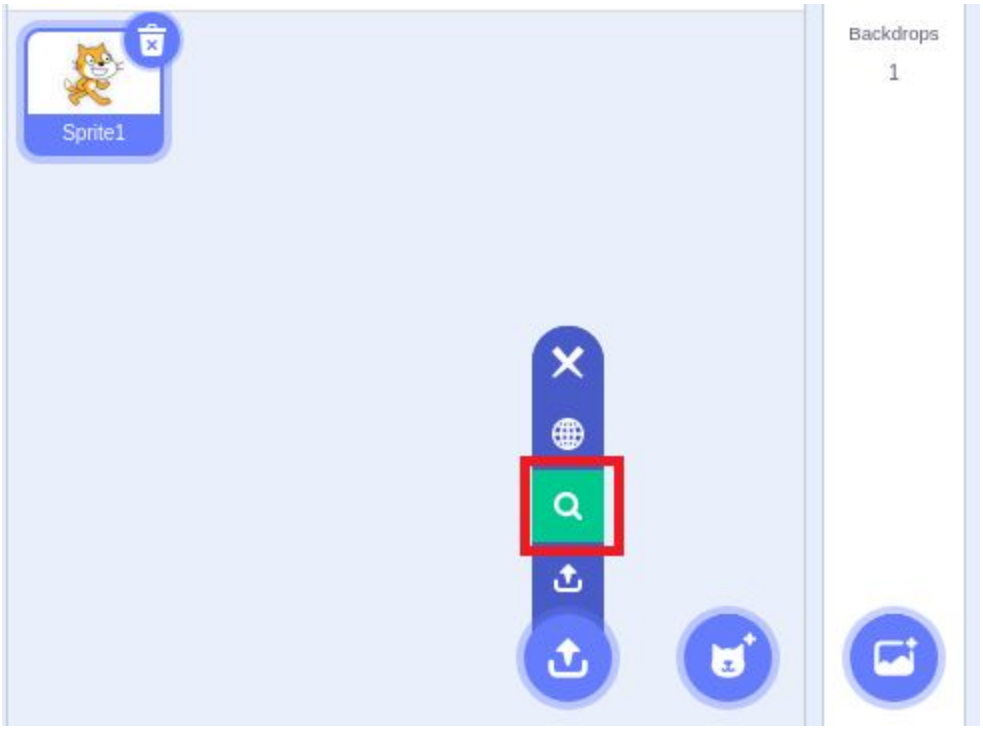


This will open a popup that prompts you to enter the name of the dataset file to remove. You must type in the exact name of the dataset you want to remove, as it appears on the block in scratch, into the prompt to remove it. Once you have done this, click the “OK” button to remove the dataset file.



Viewing a Dataset

Once you have uploaded a dataset, you will find viewing it in the editor significantly easier than displaying it's values using scratch blocks. To view the file in the editor, once again hover over the new button added by the extension in the bottom right, and click on the "View File" option.



Now you should be able to view the dataset in spreadsheet form. To switch between uploaded files, click on the name of the dataset you want to view at the top of scratch.

	year	number	incomeCurrentPerCapita	income2018PerCapita
1	2018	324356	36080	36080
2	2017	323144	35048	35902
3	2017	323156	34489	35330
4	2016	320372	33205	34746
5	2015	318868	31653	33549
6	2014	316168	30176	32039
7	2013	313443	30027	32420
8	2013	313395	28829	31127
9	2012	311116	28281	30988
10	2011	308827	27554	30830
11	2010	306553	26558	30653
12	2009	304280	26530	31126
13	2008	301483	26964	31525
14	2007	299106	26804	32541
15	2006	296824	26352	32900
16	2005	293834	25036	32270
17	2004	291166	23857	31792
18	2003	288280	23276	31856
19	2002	285933	22794	31905
20	2001	282082	22851	32489
21	2000	279517	22346	32675

If your dataset has too many columns to legibly display at once, then the dataset will be displayed such that it is legible and to view more columns you must change pages. You can view the next page of columns by clicking on the single “>” button, or the single “<” button to go back one page. If you wish to skip to the first page click on the double “<<” button, or to skip to the last page click the double “>>” button.

To stop viewing the datasets and return to the blocks screen, click the “Back” button in the top left.

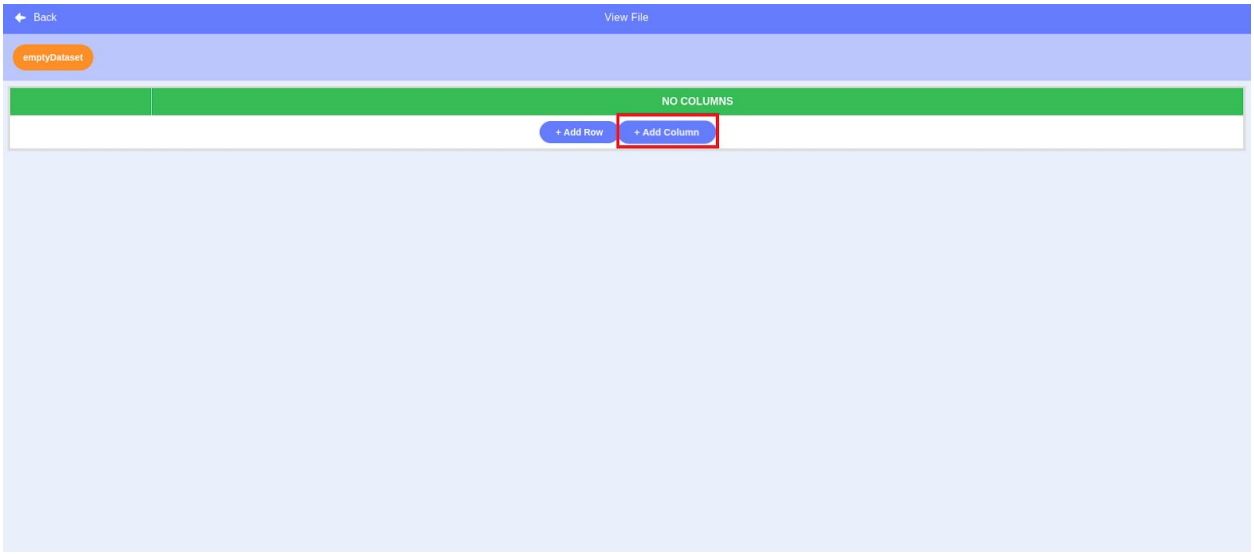
	year	number	incomeCurrentPerCapita	income2018PerCapita
1	2018	324356	36080	36080
2	2017	323144	35048	35902
3	2017	323156	34489	35330
4	2016	320372	33205	34746
5	2015	318868	31653	33549
6	2014	316168	30176	32039
7	2013	313443	30027	32420
8	2013	313395	28829	31127
9	2012	311116	28281	30988
10	2011	308827	27554	30830
11	2010	306553	26558	30653
12	2009	304280	26530	31126
13	2008	301483	26964	31525
14	2007	299106	26804	32541
15	2006	296824	26352	32900
16	2005	293834	25036	32270
17	2004	291166	23857	31792
18	2003	288280	23276	31856
19	2002	285933	22794	31905
20	2001	282082	22851	32489
21	2000	279517	22346	32675

Editing a Dataset

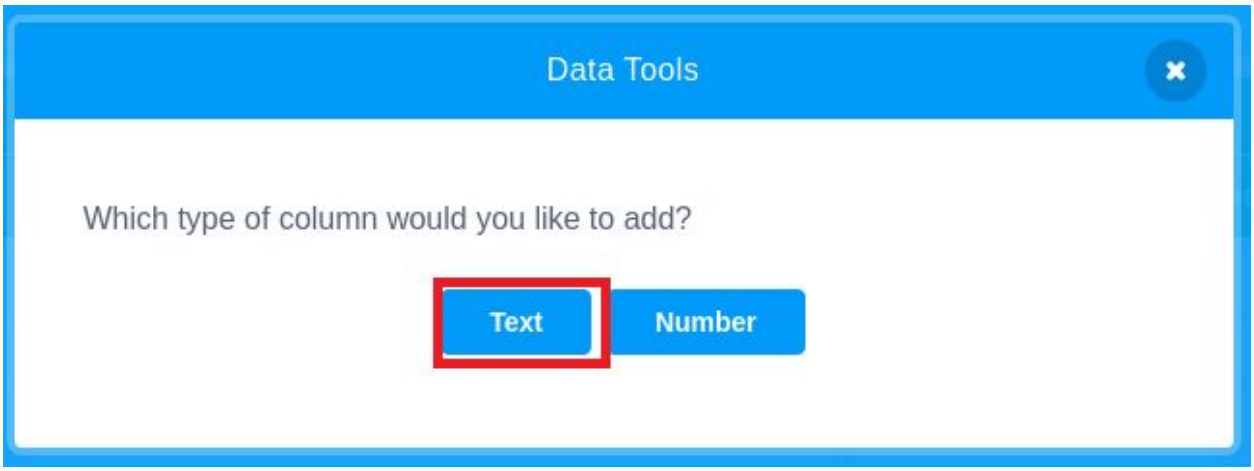
To make it easier to analyze and edit the datasets uploaded, it is possible to edit them with or without the use of blocks. The following section will explain how to edit a dataset without using blocks. To see information on how to edit a dataset with blocks, see the "Using the Blocks" section.

Adding a Column

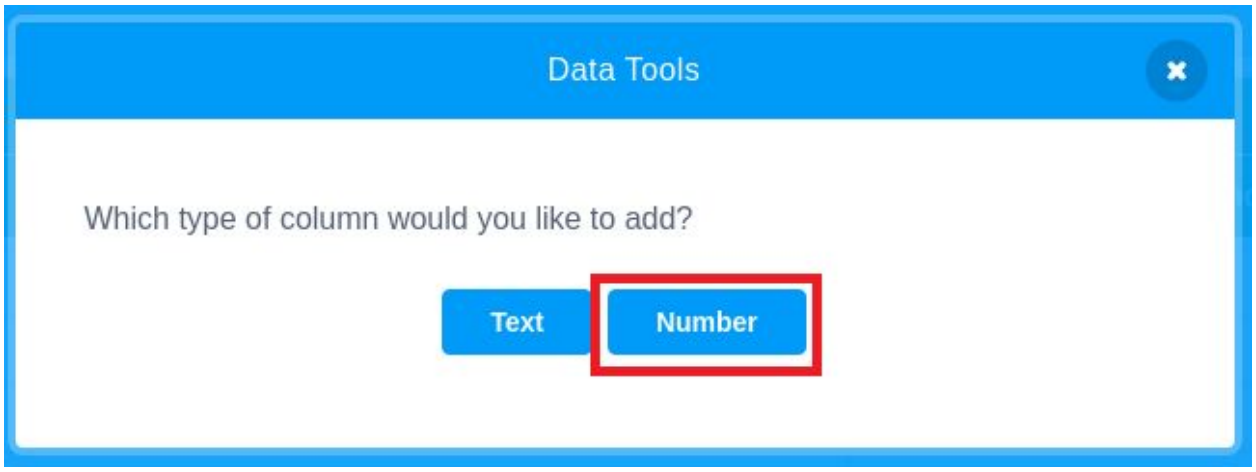
To add a column to your dataset, click on the "Add Column" button that can be found at the bottom of your dataset.



A popup will appear. Next, select the kind of values you would like that column to represent; text or numbers.
If you would like the values stored in that column to be text, click on the "Text" button.

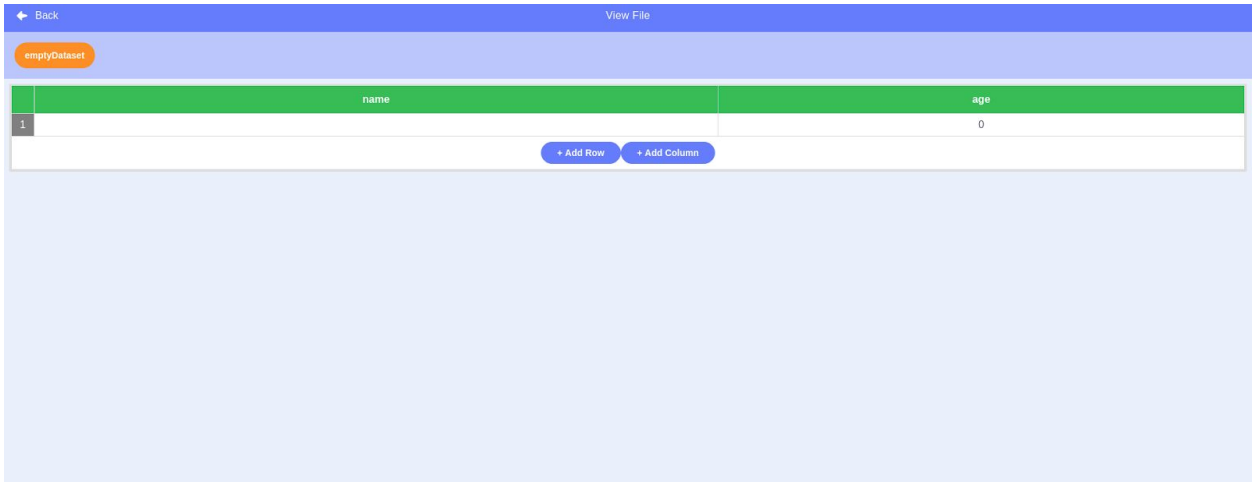


Or, if you would like the values stored in that column to be numbers, click on the “Number” button.



After selecting the kind of values to be stored in that column, type in the name you want for the column and click the “Enter” button.

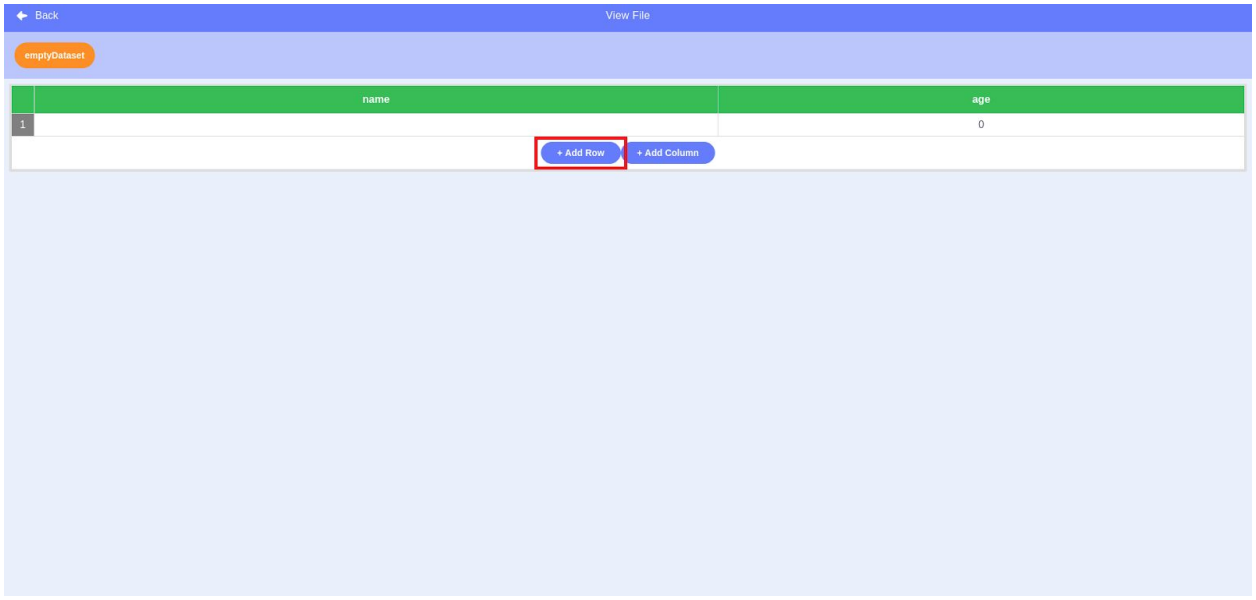
You should now see your new column. If the column has “Number” type values, then the default value that was used as a placeholder for your data was 0. Otherwise, there is no value.



Adding a Row

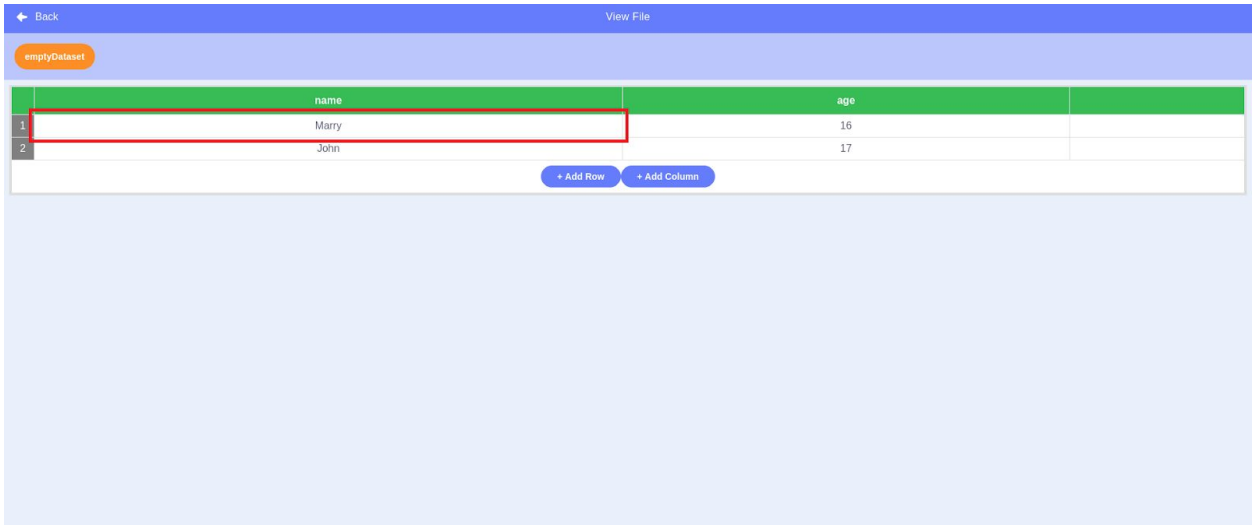
Before you can add a new row to a dataset, your dataset must first have a column. If your dataset does not already have a column, add one using the instructions found in the "Adding a Column" section.

To add a row to your dataset, click on the "Add Row" button that can be found at the bottom of your dataset.

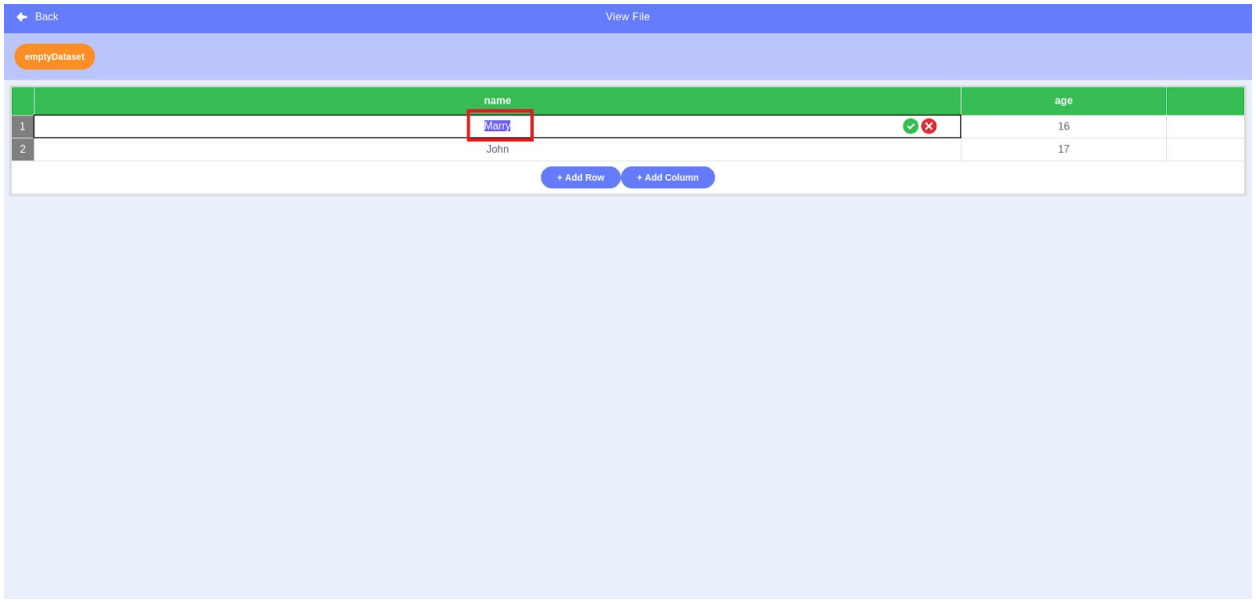


Editing a Value

To edit the value of a cell at a particular row and column, hover over the cell you would like to edit. Then simply click on the cell.



The previous value will be selected automatically. Begin typing and the old value will be overwritten by the new value typed. Press the “enter” key on your keyboard to save the new value and stop editing the cell.



Additionally, you may click the green checkmark-button on the right-hand side of the cell to save the new value and stop editing the cell.

If you want to cancel your current changes and stop editing the cell, reverting the value back to what it was before you began editing the cell, you may click on the red x-button on the right-hand side of the cell.

Blocks

Stack Blocks

<sb> add row to [dataset v] </sb>
<sb> add [text v] column [ ] to [dataset v] </sb>
<sb> set [[dataset v] column v] at row (1) to [ ] </sb>
<sb> duplicate [dataset v] as [ ] </sb>
<sb> create new data set [ ] </sb>
<sb> map: set result at [column] to [ ] </sb>
<sb> filter: return [true v] </sb>
<sb> reduce: update accumulator by (0) </sb>

C Blocks

<sb> [map v] ([dataset v]) </sb>

Reporter Blocks

<sb> (dataset) </sb>
<sb> (row count of [dataset v]) </sb>
<sb> (get [[dataset v] column v] at row (1)) </sb>
<sb> (current row at [column]) </sb>
<sb> (save [dataset v] as [filename]) </sb>

add row to [dataset v] (Block)

The **add row to [dataset v]** block is a Stack Block in the Scratch Data Tools Extension that adds a new row to the specified dataset. The new row will be filled with default values based upon the type of value for each column.

Example Uses

Adding a new row to the dataset “students.”
<scratchblocks>

add row to [students v]

</scratchblocks>

add [text v] column [ ] to [dataset v] (Block)

The **add [text v] column [ ] to [dataset v]** block is a Stack Block in the Scratch Data Tools Extension that adds a new column of the selected type (text or number), with the given name, to the selected dataset. The values in that column in each row will be filled with the default value based upon the column type.

Example Uses

Adding a new column named “name” that is “text” type to the dataset “students.”
<scratchblocks>

add [text v] column [name] to [students v]

</scratchblocks>

set [[dataset v] column v] at row (1) to [ ] (Block)

The **set [[dataset v] column v] at row (1) to [ ]** block is a Stack Block in the Scratch Data Tools Extension that modifies the value at a particular row and column to a provided value. It takes the dataset, column in said dataset, and the row number and replaces the stored value there with the given value.

Example Uses

Sets the grade, the given column, for the student from the “students” dataset at row one to an “A.”
<scratchblocks>

set [[students v] grade v] at row (1) to [A]

</scratchblocks>

duplicate [dataset v] as [ ] (Block)

The **duplicate [dataset v] as [ ]** block is a Stack Block in the Scratch Data Tools Extension that duplicates the given dataset, creating a new copy of it with the given name.

Example Uses

Duplicates the “students” dataset, creating a copy of it as a new dataset with the name “studentsNew.”
<scratchblocks>

duplicate [students v] as [studentsNew]

</scratchblocks>

create new data set [ ] (Block)

The **create new data set [ ]** block is a Stack Block in the Scratch Data Tools Extension that creates a new, blank, dataset with the given name.

Example Uses

Creates a new, blank, dataset that will be named “students.”
<scratchblocks>

Create new data set [students]

</scratchblocks>

map: set result at [column] to [ ] (Block)

The **map: set result at [column] to [ ]** block is a Stack Block in the Scratch Data Tools Extension that sets value at a given column, where the current row in the map operation this block is within is used to determine the row, to the given input.

A **map: set result at [column] to [ ]** block can only be used within a **[map v] ([dataset v])** block where the chosen option in the first drop-down is “map.”

Example Uses

Sets the value of the “GPA” column in each row, within the “students” dataset to the given value “4.”

<scratchblocks>

[map v] ([students v])
    map: set result at [GPA] to [4]
end

</scratchblocks>

filter: return [true v] (Block)

The **filter: return [true v]** block is a Stack Block in the Scratch Data Tools Extension that

A **filter: return [true v]** block can only be used within a **[map v] ([dataset v])** block where the chosen option in the first drop-down is “filter.”

Example Uses

<scratchblocks>

</scratchblocks>

reduce: update accumulator by (0) (Block)

The **reduce: update accumulator by (0)** block is a Stack Block in the Scratch Data Tools Extension that

A **reduce: update accumulator by (0)** block can only be used within a **[map v] ([dataset v])** block where the chosen option in the first drop-down is “reduce.”

Example Uses

<scratchblocks>

</scratchblocks>

[map v] ([dataset v]) (Block)

The **[map v] ([dataset v])** block is a C Block in the Scratch Data Tools Extension that allows the user to apply operations on every row in a given dataset, since the block iterates over every row and performs the operations contained within the block. Essentially, it is a Repeat () block that repeats for every row in the dataset, but also allows accessing the current row based on iterations.

For each option within the **[map v] ([dataset v])** block within the first dropdown (map, filter, or reduce), in order to function they require their associated Stack Blocks:
map: set result at [column] to [ ]
filter: return [true v]
reduce: update accumulator by (0)

Example Uses

Sets the value of the “GPA” column in each row, within the “students” dataset to the given value “4.”
<scratchblocks>

[map v] ([students v])
    map: set result at [GPA] to [4]
end

</scratchblocks>

dataset (Block)

The **dataset** block is a Dataset Block and Reporter Block in the Scratch Data Tools Extension that represent a dataset. Each version of the block represents a different dataset. The block itself is used as input into other blocks for determining which dataset to work with. However, most such blocks that can take it as input also have a drop-down menu for selecting a dataset from, so the dataset block is not needed for other blocks to function.

Example Uses

Adding a new row to the given dataset “students,” which was passed in from the “students” variable block.
<scratchblocks>

add row to [(students) v]

</scratchblocks>

row count of [dataset v] (Block)

The **row count of [dataset v]** block is a Reporter Block in the Scratch Data Tools Extension that retrieves the number of rows in the given dataset.

Example Uses

Have the sprite say the number of rows in the given dataset “dataset.”
<scratchblocks>

say [(row count of [dataset v])]

</scratchblocks>

get [[dataset v] column v] at row (1) (Block)

The **get [[dataset v] column v] at row (1)** block is a Reporter Block in the Scratch Data Tools Extension that retrieves the value for a given column and row combination in the given dataset.

Example Uses

Have the sprite say the “name,” the given column, of row one in the given dataset “students.”
<scratchblocks>

say [(get [[students v] name v] at row (1))]

</scratchblocks>

save [dataset v] as [filename] (Block)

The **save [dataset v] as [filename]** block is a Reporter Block in the Scratch Data Tools Extension that saves a dataset from a **[map v] ([dataset v])** block either for use within another **[map v] ([dataset v])** block or to save the final result of said block to a new dataset; this block allows for nested **[map v] ([dataset v])** blocks. The block is used as the dataset input for the **[map v] ([dataset v])** block, and the **[map v] ([dataset v])** block to be nested is used as the dataset input for it. It can serve the same use as a **duplicate [dataset v] as [ ]** block when given no **[map v] ([dataset v])** block as input for a dataset but instead just a dataset to duplicate.

Example Uses

Have the sprite say the “name,” the given column, for each row in the dataset “students.”
<scratchblocks>

```
[map v] ([
        save [
                [map v] ([students v])
                        map: set result at [GPA] to [3]
                end
        v] as [students2]
 v])
        map: set result at [GPA] to [4]
end
```

</scratchblocks>

current row at [column] (Block)

The **current row at [column]** block is a Reporter Block in the Scratch Data Tools Extension that retrieves the value for a given column and row, based on the current row in the containing **[map v] ([dataset v])** block, combination in the dataset currently being operated on by the **[map v] ([dataset v])** block this block is contained within.

A **current row at [column]** block can only be used within a **[map v] ([dataset v])** block.

Example Uses

Have the sprite say the “name,” the given column, for each row in the dataset “students.”
<scratchblocks>

[map v] ([students v])
say [(current row at [name])]
end

</scratchblocks>