Web Tools for Genomic Epidemiology - Paraguay 2023

Module Lead: Silvia Argimon & Julio Diaz Caballero

Table of contents

Background

Genomic epidemiology aims to understand the emergence and dissemination of high-risk clones within pathogen populations with the ultimate goal of implementing evidence-based interventions to protect public health. High-risk clones are subpopulations/strains/variants of a pathogen that carry risk elements, such as antibiotic resistance or virulence determinants, and thus pose a potential risk to public health.

The increased genetic resolution afforded by genomic data is useful from global to local geographic scales, and it has proven particularly useful for the investigation of pathogens that exhibit little genetic variation (e.g. Wong *et al.* 2016), and for outbreak investigations (e.g. Hendriksen *et al.* 2011, Eppinger *et al.* 2014). During outbreak investigations (and other genomic epidemiology studies), epidemiological data from patients is collected by healthcare professionals. Antimicrobial susceptibility data, species identification, and any further phenotypic or molecular characterization of the isolates is often generated by the laboratories linked to healthcare facilities and/or by the reference laboratory. Ideally, these different sources of data are stored in a centralised surveillance system and database, such as WHONET. However, these systems rarely incorporate genomic data produced by bioinformaticians. Genomic data may include sequence/assembly quality, genotyping/genoserotyping information, presence/absence of known risk elements, and clustering of the isolates based on genetic similarity –usually in the form of a phylogenetic tree or a minimum spanning-tree.

The job of a genomic epidemiologist often starts by combining data from diverse sources that might not be complete or standardised to facilitate the identification of relevant patterns for meaningful interpretation. Once integrated, the epi, lab, and genomic data can be interrogated over the structure of the tree to identify clusters of interest that will place the isolates in or out of the outbreak.

The **Centre for Genomic Pathogen Surveillance** develops free web applications for data collection, integration, visualisation, and analysis of genomic epidemiology data. **Epicollect5** is a mobile & web application for free and easy data collection. It provides both the web and mobile applications for the generation of forms (questionnaires) and freely hosted project websites for data collection. Projects are created by using the web application at five.epicollect.net, and then downloaded to the device to perform the data collection. Data are collected in the field using multiple devices and all data can be viewed on a central server (via map, tables, and charts). **Data-flo** is a system for customised integration and manipulation of diverse data via a simple drag and drop interface. Data-flo provides a visual method to design a reusable pipeline to integrate, clean, and manipulate data in a multitude of ways, eliminating the need for continuous manual intervention (e.g., coding, formatting, spreadsheet formulas, manual copy-pasting). **Microreact** allows you to upload, visualise and explore any combination of clustering (trees), geographic (map) and temporal (timeline) data. Other metadata variables are displayed in a table. You can specify colours and/or shapes to display on the map, tree and/or timeline. A permanent URL is produced for you to share your Microreact.

Introduction

Aim: To use web and mobile tools for the collection of epi data in the field, its integration with lab and bioinformatics data, and to finally visualise it together to detect a putative outbreak.

Note: This exercise is a work of fiction. Any resemblance to a real cholera outbreak is purely coincidental.

During January and February 2023, fifty-eight patients reported to eight hospitals in Asunción with acute watery diarrhoea and vomiting. Vibrio cholerae was isolated from the stool samples of fourty patients (all belong to serotype Inaba), and one colony from each confirmed case was sequenced on Illumina MiSeq. An epidemiological investigation revealed that more than half of the culture-confirmed cases regularly visited the Ñu Guazú Park, where water fountains and other water sources are accessible, and which is crossed by the Arroyo Itay. This suggested that the outbreak could be caused by contaminated water in the park. The epidemiological and demographic data is summarised on the file epi_data.csv.

A team of disease detectives will be deployed around Ñu Guazú Park today to investigate the potential source of the outbreak using Epicollect5, data-flo, and Microreact, free mobile and web applications developed by the Centre for Genomic Pathogen Surveillance for easy data collection, integration and visualisation.

Learning Outcomes

At the end of this session the participants will be able to:

1. Collect data in the field with Epicollect5.
2. Merge epidemiological, laboratory and genomic data with data-flo
3. Create a Microreact project to visualise epidemiological, laboratory and genomic data

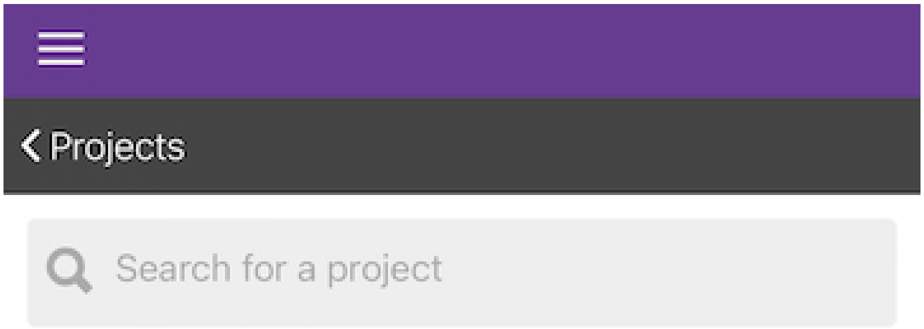
1. Collect field data with Epicollect

`https://five.epicollect.net/`

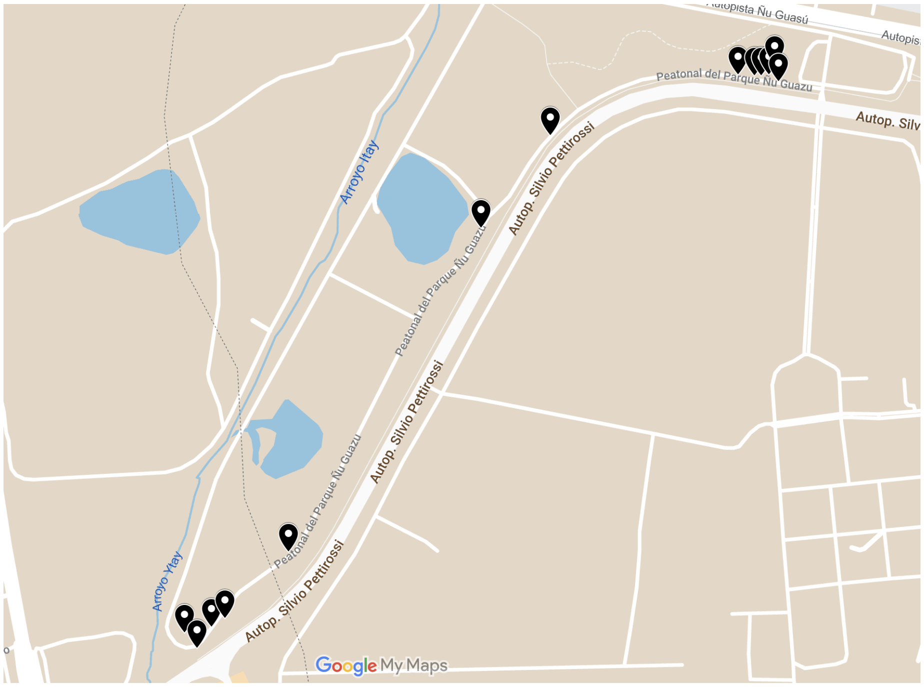
The disease detectives will work in pairs, and will be provided with a **collection tube** with a **barcode** sample identifier to collect a sample of water, and will use their **mobile phone** to collect associated metadata in the field.

Make sure your phone is connected to WiFi. Open the **Epicollect5** app on your mobile phone.

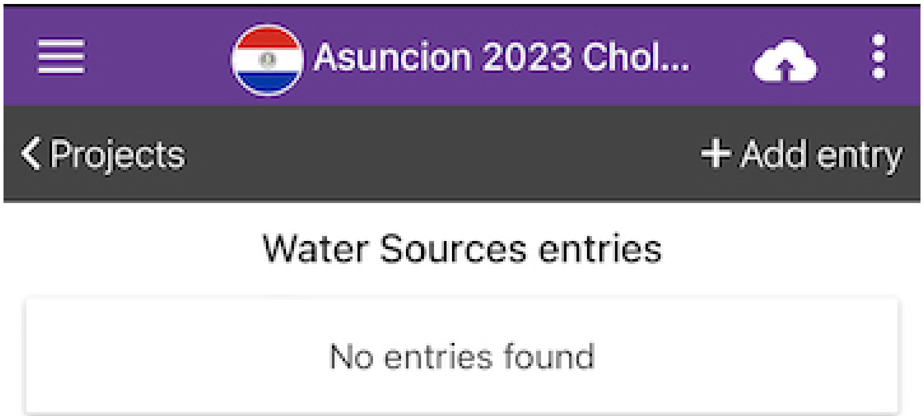
Add the project **Asuncion 2023 Cholera Outbreak** by typing into the search box –you will need to be connected to WiFi.



Using the map of the Ñu Guasú park provided, walk to the location of your collection point (water source). Locations are indicated by the red dots labelled a-j. You will not need internet access to collect data in the field.



Open the **Asuncion 2023 Cholera Outbreak** project and add an entry. Follow the form to collect a water sample and associated metadata.


Asuncion 2023 Chol...

Projects

Add entry

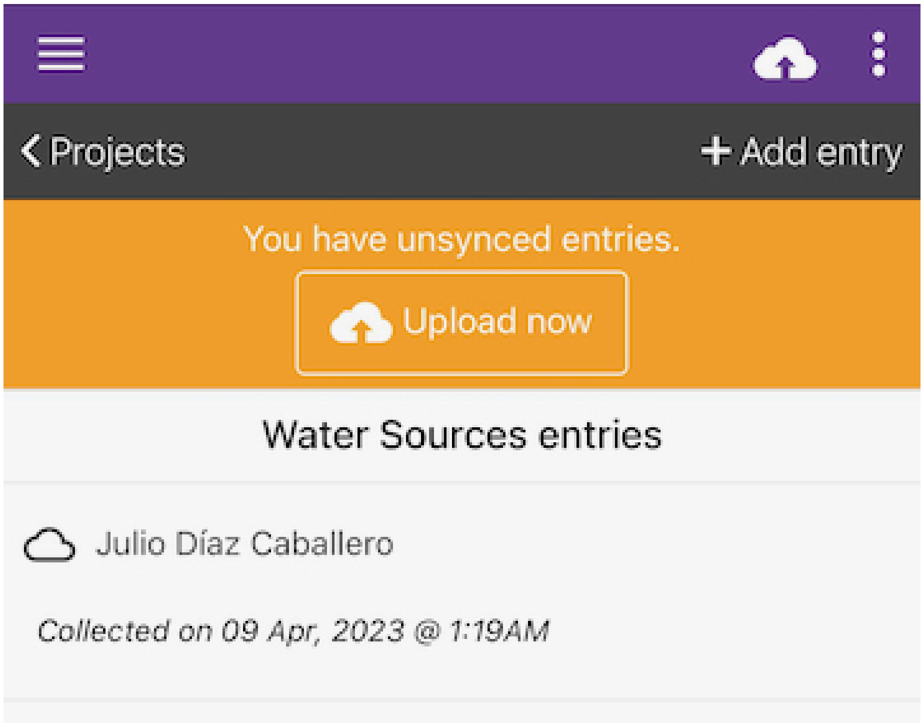
Water Sources entries

No entries found

Your team will collect only one water sample, but both of you can collect the associated data on Epicollect5. However, *make sure that only one of you uploads the data later on to avoid duplications.*

Return to the Epidemic Intelligence Center (i.e. the classroom).

One of the pair members only. **Upload your entry (and image)** to the Epicollect5 server –make sure you are connected to the WiFi.



Submit your water sample to the instructor, which will be sent to the reference lab for culture confirmation.

Once all entries are uploaded by the disease detectives we will take a look at the data together on **https://five.epicollect.net/project/asuncion-2023-cholera-outbreak**

Answer the following questions:

1. What is the most common type of water source?
2. Was it possible to collect a sample from all 13 sources? If not, what were the reasons?

2. Merge data with Data-flo

`https://data-flo.io/`

Note: you need to sign-up for data-flo and Microreact. See instructions in the Resources section. Creating your own account will allow you to manage and edit your projects.

The reference lab has sent you the culture results on an Excel file called `lab_results.xlsx`. The lab reported that **2 out of 13** water samples from Ñu Guasú park were positive for *Vibrio cholerae*. **This immediately prompted the closure of the 2 water sources.**

> Does this confirm that the source of the outbreak was contaminated water from Ñu Guasú park?

One colony from each source was sequenced on Illumina MiSeq by the reference lab. A maximum likelihood phylogenetic tree (`tree.nwk`) was inferred from the genomes of the 34 clinical samples and 2 environmental (water) samples. Six genomes from a previous outbreak (Jan-Feb 2020) were also included in the tree inference and their associated data added to the `epi_data.csv` file.

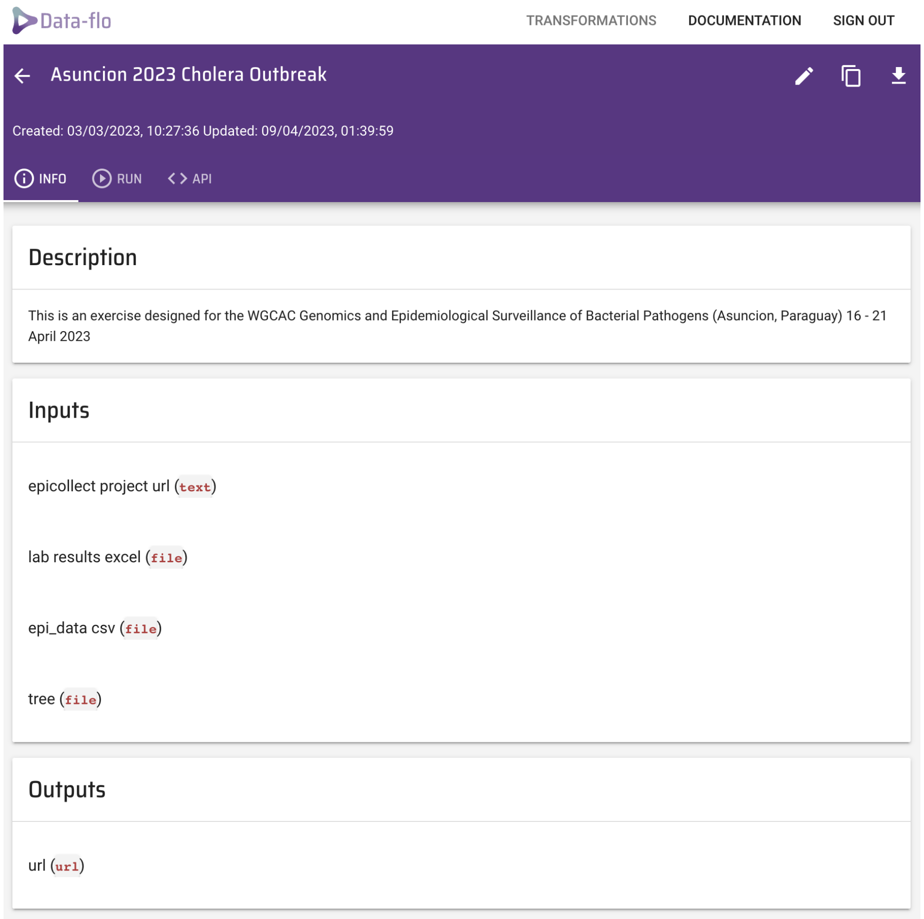
The disease detectives now have the information needed for the investigation in the following formats:

- `epi_data.csv` Epi data from 34 clinical cases and 6 cases from previous outbreak
- `Epicollect project` Metadata of 13 water sources from Ñu Guasú Park
- `lab_results.xlsx` Culture and serotyping results
- `tree.nwk` Phylogenetic tree of 34 clinical cases, 2 culture-positive water samples, and 6 cases from previous outbreak

The files are located in the VM folder **FOLDER NAME**.

We will combine data from these different sources with a data-flo workflow that takes the files above and the data from the Epicollect project as input, and creates as an output a Microreact project where the data can be visualised.

Open the data-flo workflow (**https://data-flo.io/run?prnvpvmKzq8rDKbPD9ZXAs**)



Asuncion 2023 Cholera Outbreak

Created: 03/03/2023, 10:27:36 Updated: 09/04/2023, 01:39:59

Description

This is an exercise designed for the WGCAC Genomics and Epidemiological Surveillance of Bacterial Pathogens (Asuncion, Paraguay) 16 - 21 April 2023

Inputs

epicollect project url (`text`)

lab results excel (`file`)

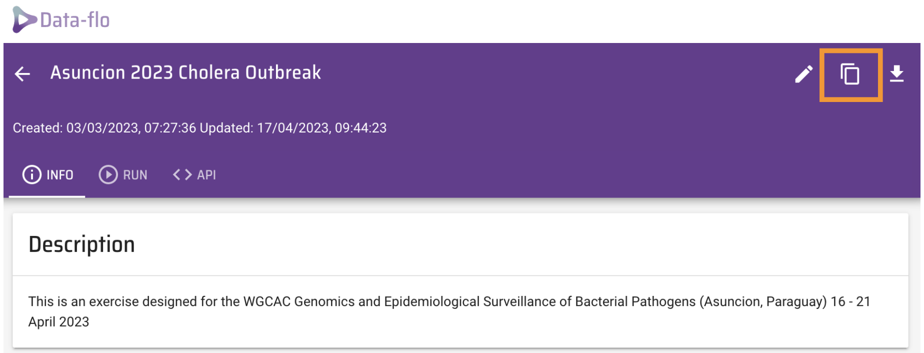
epi_data csv (`file`)

tree (`file`)

Outputs

url (`url`)

Copy the workflow to your own dataflo account.



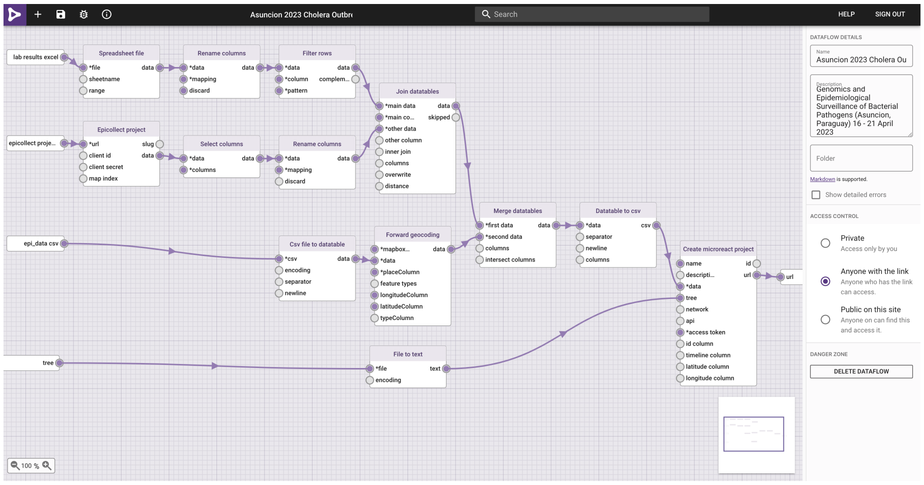
Asuncion 2023 Cholera Outbreak

Created: 03/03/2023, 07:27:36 Updated: 17/04/2023, 09:44:23

Description

This is an exercise designed for the WGCAC Genomics and Epidemiological Surveillance of Bacterial Pathogens (Asuncion, Paraguay) 16 - 21 April 2023

This will open a copy of this workflow in your dataflo account.



DATAFLOW DETAILS

Name: Asuncion 2023 Cholera Ou

Description: Genomics and Epidemiological Surveillance of Bacterial Pathogens (Asuncion, Paraguay) 16 - 21 April 2023

Folder

Markdown is supported.

□ Show detailed errors

ACCESS CONTROL

○ Private
Access only by you

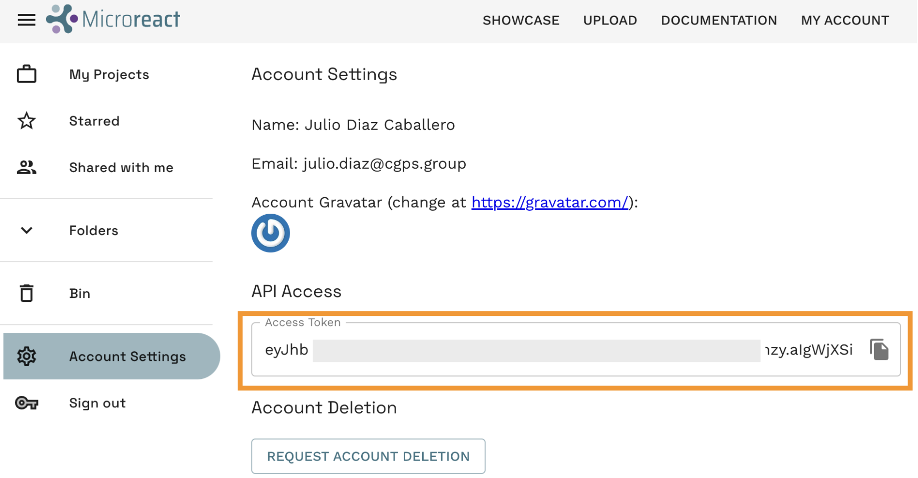
● Anyone with the link
Anyone who has the link can access.

○ Public on this site
Anyone on can find this and access it.

DANGER ZONE

DELETE DATAFLOW

On a different browser tab, get your microreact `API access token` at **https://microreact.org/my-account/settings** (you must already have created your microreact account).



Account Settings

Name: Julio Diaz Caballero

Email: julio.diaz@cgps.group

Account Gravatar (change at https://gravatar.com/):

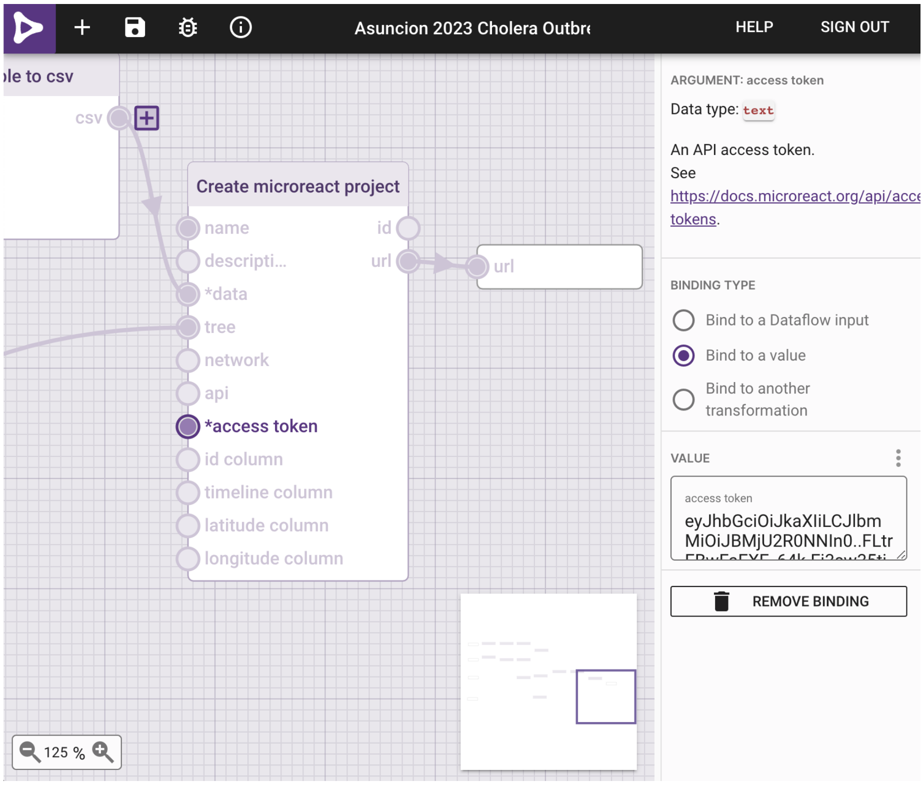
API Access

Access Token: eyJhb [redacted] ızy.aIgWjXSi

Account Deletion

Edit the workflow to include your own microreact `API access token`.

1. Click on `*access token` in the `Create microreact project` box
2. From the options on the right, select `Bind to value`
3. Paste your `API access token` in the `VALUE` box



ARGUMENT: access token

Data type: `text`

An API access token.
See
https://docs.microreact.org/api/acce
tokens.

BINDING TYPE

○ Bind to a Dataflow input

● Bind to a value

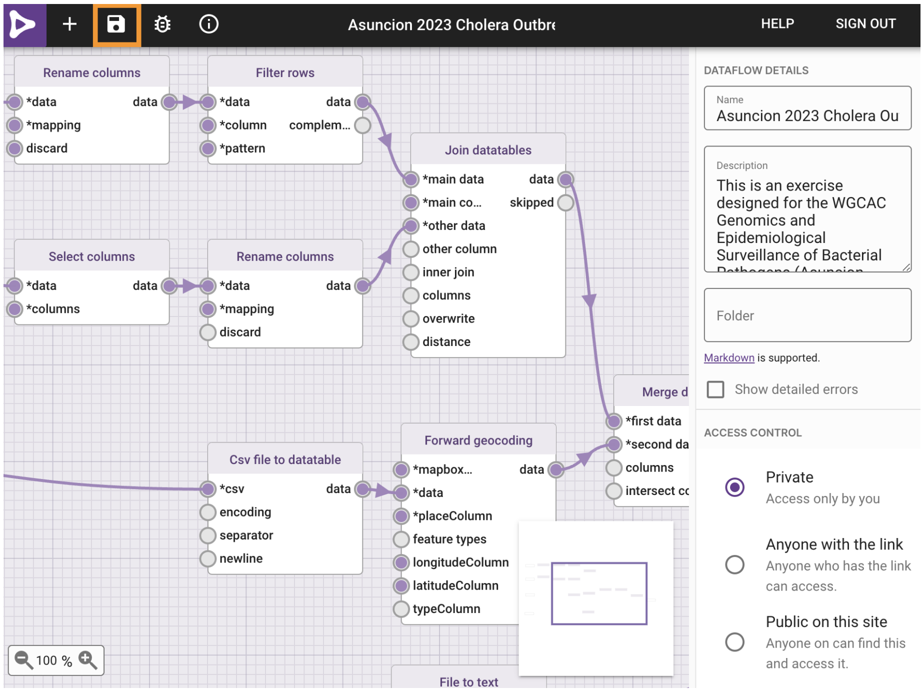
○ Bind to another transformation

VALUE

access token
eyJhbGciOiJkaXIiLCJlbm
MiOiJBMjU2R0NNIn0..FLtr

REMOVE BINDING

Save your newly edited dataflo project by clicking on the `save` icon.



DATAFLOW DETAILS

Name
Asuncion 2023 Cholera Ou

Description
This is an exercise designed for the WGCAC Genomics and Epidemiological Surveillance of Bacterial Pathogens (Asuncion

Folder

Markdown is supported.

☐ Show detailed errors

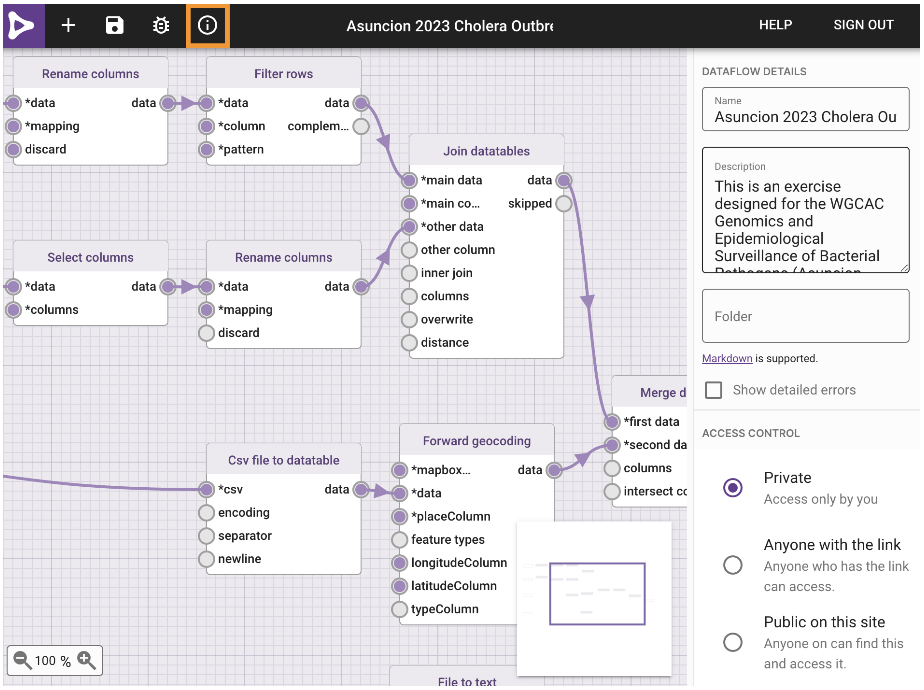
ACCESS CONTROL

● Private
Access only by you

○ Anyone with the link
Anyone who has the link can access.

○ Public on this site
Anyone on can find this and access it.

Now lets go to the implementation page by clicking on the **i** icon.



DATAFLOW DETAILS

Name
Asuncion 2023 Cholera Ou

Description
This is an exercise designed for the WGCAC Genomics and Epidemiological Surveillance of Bacterial Pathogens (Asuncion

Folder

Markdown is supported.

☐ Show detailed errors

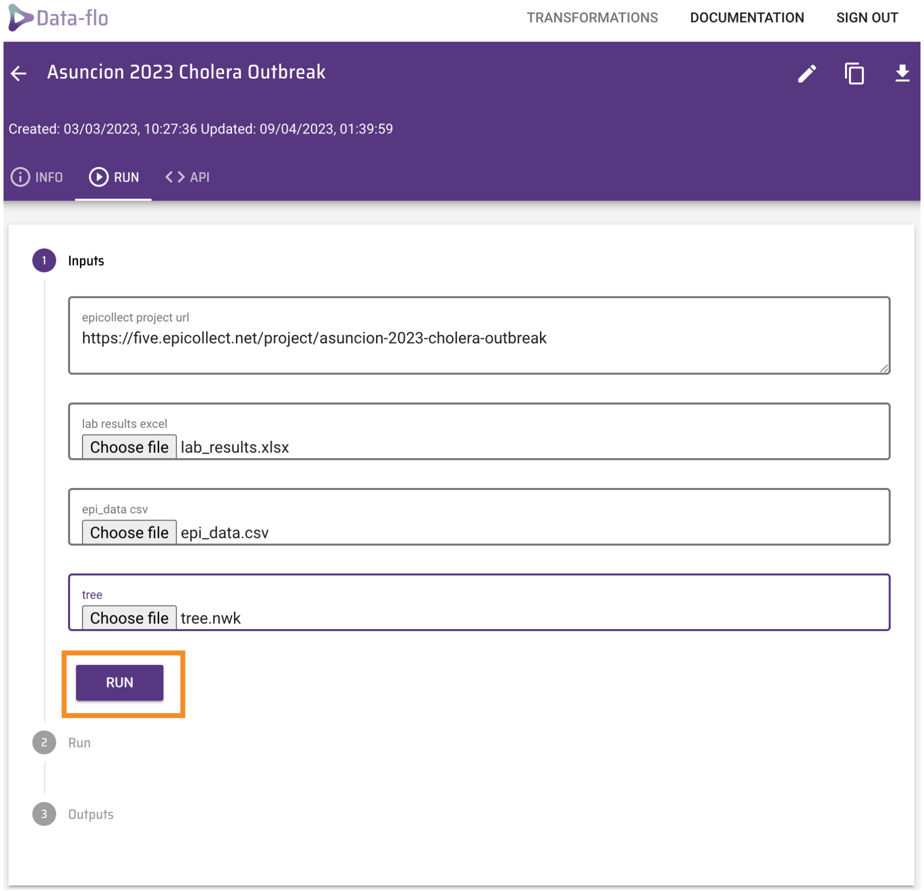
ACCESS CONTROL

● Private
Access only by you

○ Anyone with the link
Anyone who has the link can access.

○ Public on this site
Anyone on can find this and access it.

To run the workflow paste the url of the Epicollect5 project (**https://five.epicollect.net/project/asuncion-2023-cholera-outbreak**) and upload the files.



Asuncion 2023 Cholera Outbreak

Created: 03/03/2023, 10:27:36 Updated: 09/04/2023, 01:39:59

1 Inputs

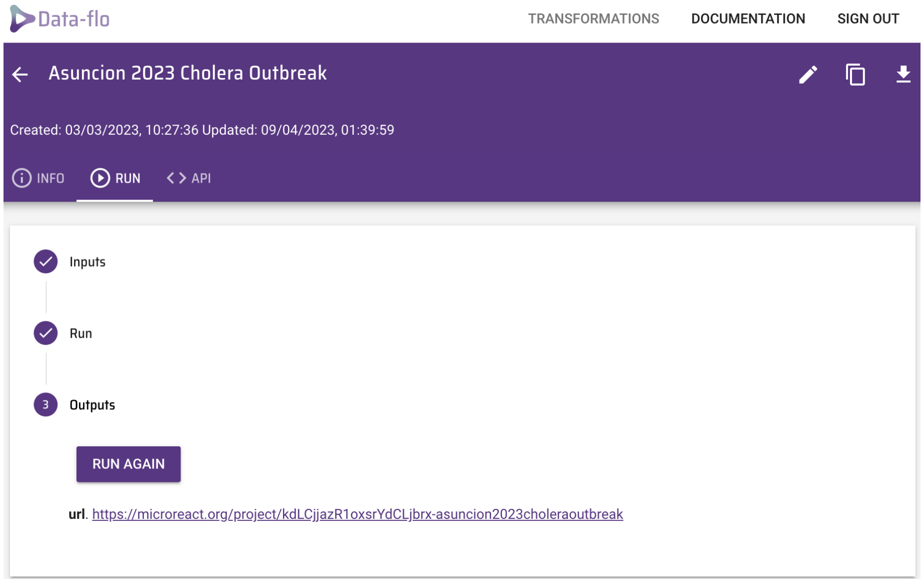
epicollect project url: https://five.epicollect.net/project/asuncion-2023-cholera-outbreak

lab results excel: Choose file lab_results.xlsx

epi_data csv: Choose file epi_data.csv

tree: Choose file tree.nwk

Click on **Run**. The Outputs box now shows the url of a Microreact project created by data-flo.



Asuncion 2023 Cholera Outbreak

Created: 03/03/2023, 10:27:36 Updated: 09/04/2023, 01:39:59

url. https://microreact.org/project/kdLCjjazR1oxsrYdCLjbrx-asuncion2023choleraoutbreak

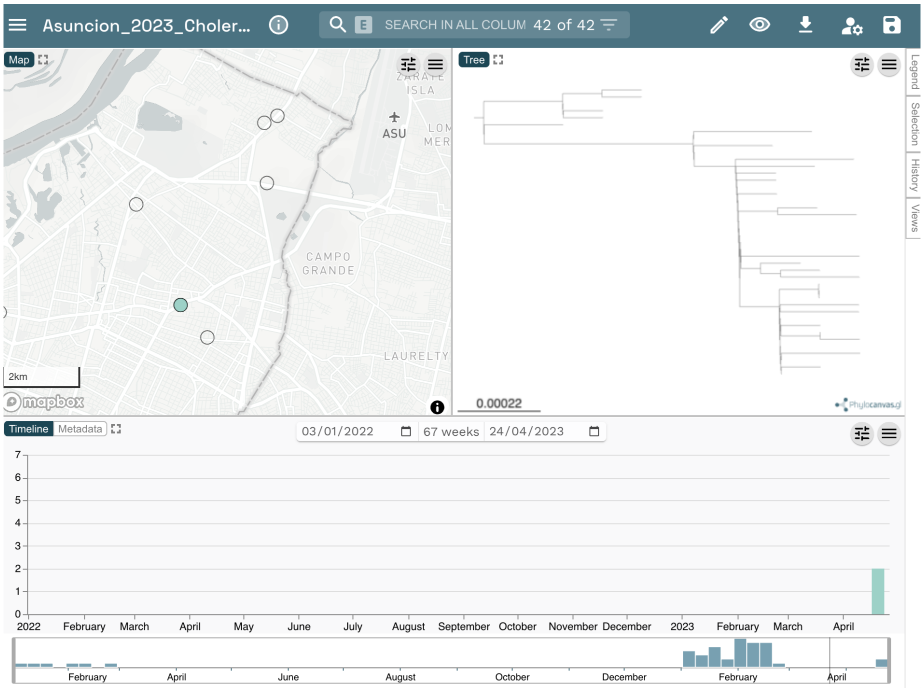
Before you open the Microreact project answer the following questions:

1. If later on we wanted to add additional entries (i.e. water samples) to the Epicollect project, could we use the same data-flo? If so, how would this be beneficial?
2. Do you use/know of any other methods to join data? If so, how do they compare with data-flo?

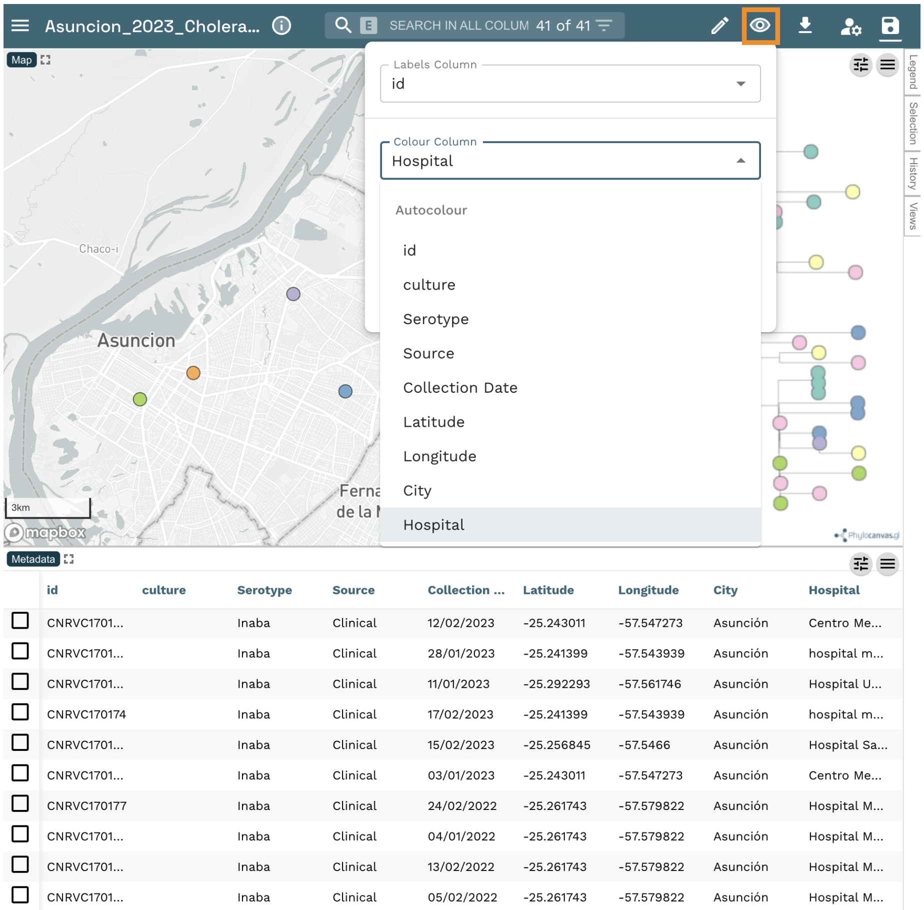
3. Microreact

`https://microreact.org/`

Open the Microreact link in the output from data-flo. It should look something like this:

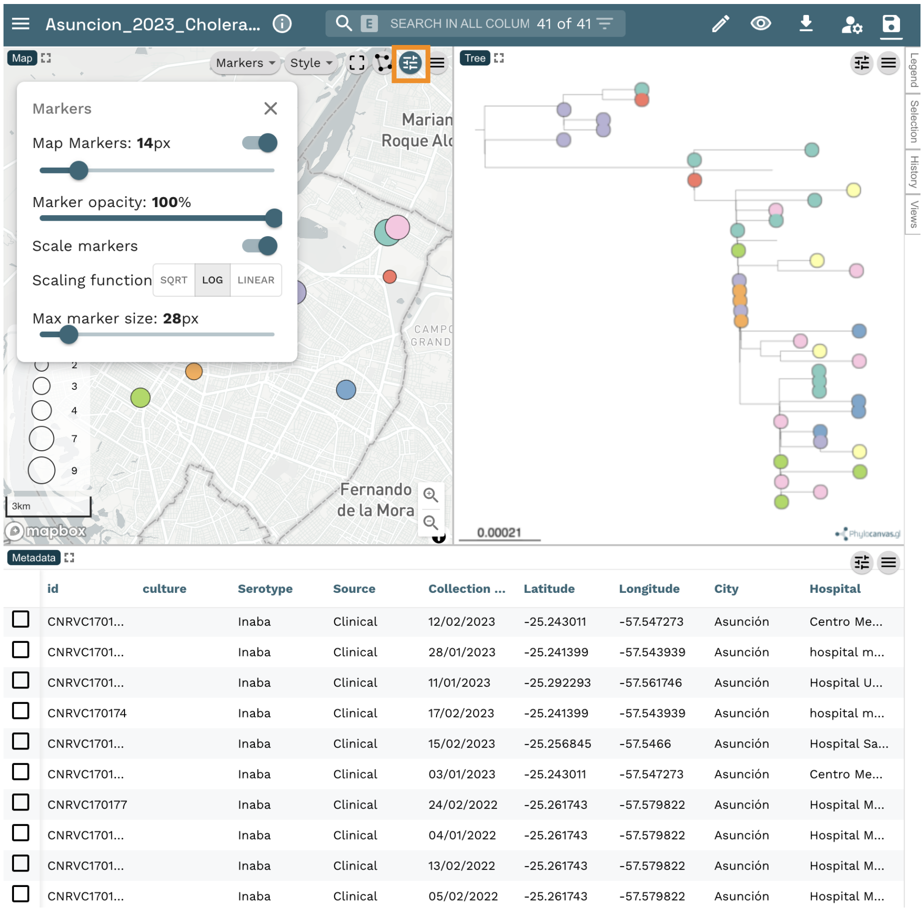


Now change the colour column to "Hospital". Click on the eye icon at the top left of the page. Select "Hospital from the `Colour Column` dropdown list.



	id	culture	Serotype	Source	Collection ...	Latitude	Longitude	City	Hospital
□	CNRVC1701...		Inaba	Clinical	12/02/2023	-25.243011	-57.547273	Asunción	Centro Me...
□	CNRVC1701...		Inaba	Clinical	28/01/2023	-25.241399	-57.543939	Asunción	hospital m...
□	CNRVC1701...		Inaba	Clinical	11/01/2023	-25.292293	-57.561746	Asunción	Hospital U...
□	CNRVC170174		Inaba	Clinical	17/02/2023	-25.241399	-57.543939	Asunción	hospital m...
□	CNRVC1701...		Inaba	Clinical	15/02/2023	-25.256845	-57.5466	Asunción	Hospital Sa...
□	CNRVC1701...		Inaba	Clinical	03/01/2023	-25.243011	-57.547273	Asunción	Centro Me...
□	CNRVC170177		Inaba	Clinical	24/02/2022	-25.261743	-57.579822	Asunción	Hospital M...
□	CNRVC1701...		Inaba	Clinical	04/01/2022	-25.261743	-57.579822	Asunción	Hospital M...
□	CNRVC1701...		Inaba	Clinical	13/02/2022	-25.261743	-57.579822	Asunción	Hospital M...
□	CNRVC1701...		Inaba	Clinical	05/02/2022	-25.261743	-57.579822	Asunción	Hospital M...

Scale the markers on the map by clicking on the sliders icoc, then in the opening options click on `Markers`. Finally, toggle the `Scale markers` option on.



Markers

Map Markers: **14**px

Marker opacity: **100**%

Scale markers

Scaling function SQRT LOG LINEAR

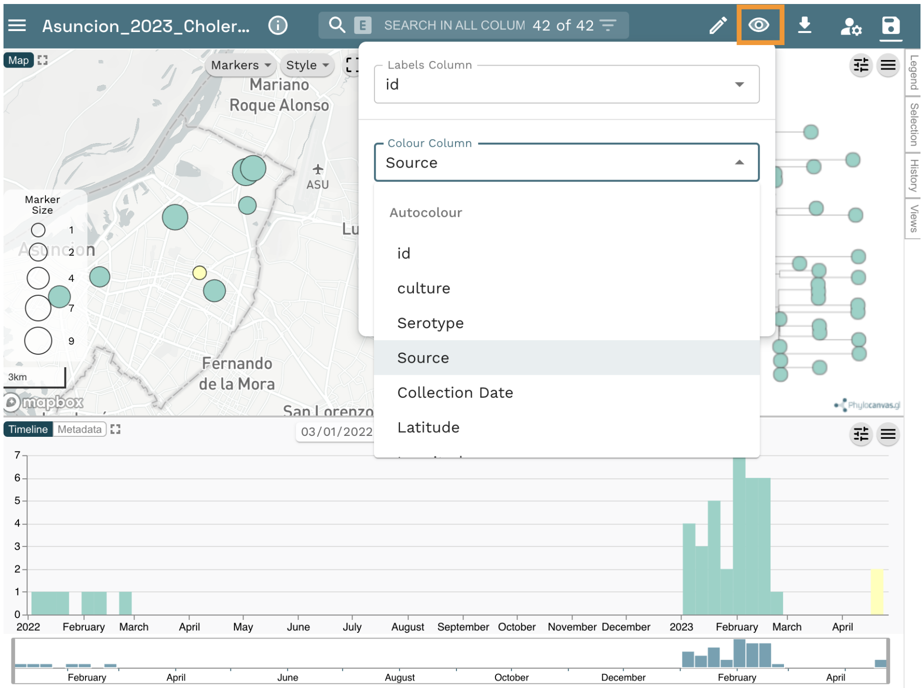
Max marker size: **28**px

	id	culture	Serotype	Source	Collection ...	Latitude	Longitude	City	Hospital
□	CNRVC1701...		Inaba	Clinical	12/02/2023	-25.243011	-57.547273	Asunción	Centro Me...
□	CNRVC1701...		Inaba	Clinical	28/01/2023	-25.241399	-57.543939	Asunción	hospital m...
□	CNRVC1701...		Inaba	Clinical	11/01/2023	-25.292293	-57.561746	Asunción	Hospital U...
□	CNRVC170174		Inaba	Clinical	17/02/2023	-25.241399	-57.543939	Asunción	hospital m...
□	CNRVC1701...		Inaba	Clinical	15/02/2023	-25.256845	-57.5466	Asunción	Hospital Sa...
□	CNRVC1701...		Inaba	Clinical	03/01/2023	-25.243011	-57.547273	Asunción	Centro Me...
□	CNRVC170177		Inaba	Clinical	24/02/2022	-25.261743	-57.579822	Asunción	Hospital M...
□	CNRVC1701...		Inaba	Clinical	04/01/2022	-25.261743	-57.579822	Asunción	Hospital M...
□	CNRVC1701...		Inaba	Clinical	13/02/2022	-25.261743	-57.579822	Asunción	Hospital M...
□	CNRVC1701...		Inaba	Clinical	05/02/2022	-25.261743	-57.579822	Asunción	Hospital M...

Answer the following questions:

1. What do the markers on the map represent?
2. Which hospital reported the index (first) case? Tip: Click on the earliest sample on the timeline.

Change the colour column to “Source”. Click on the eye icon at the top left of the page. Then select “Source” from the `Colour Column` dropdown list.

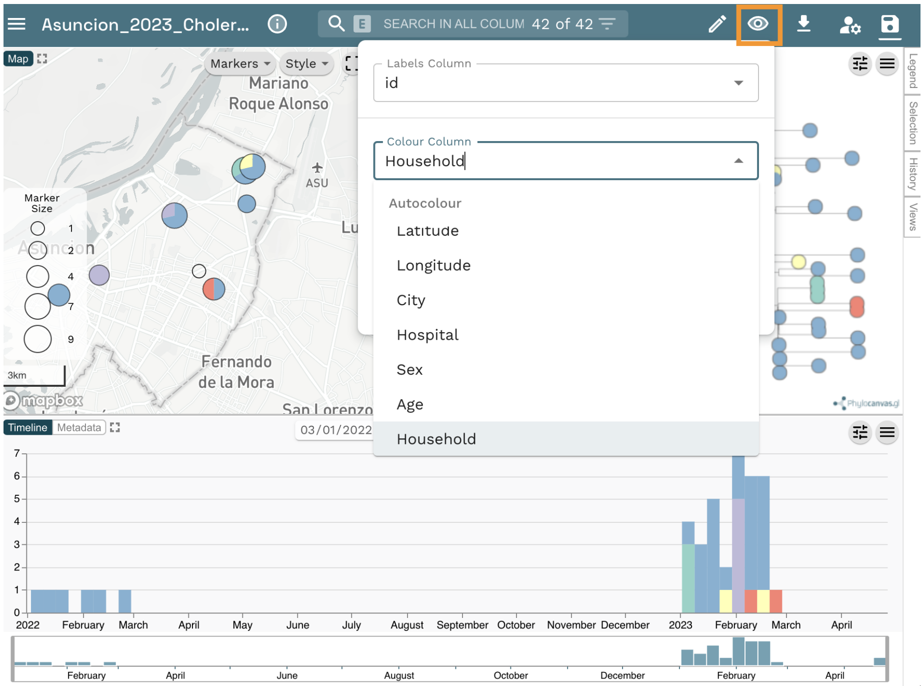


Labels Column: id

Colour Column: Source

3. Can you confirm that the source of the outbreak is the contaminated water from Ñu Guasú Park?
4. Why is it useful to include genomes from a previous outbreak in the analysis? Is the current outbreak related to the past outbreak from Jan-Feb 2020?

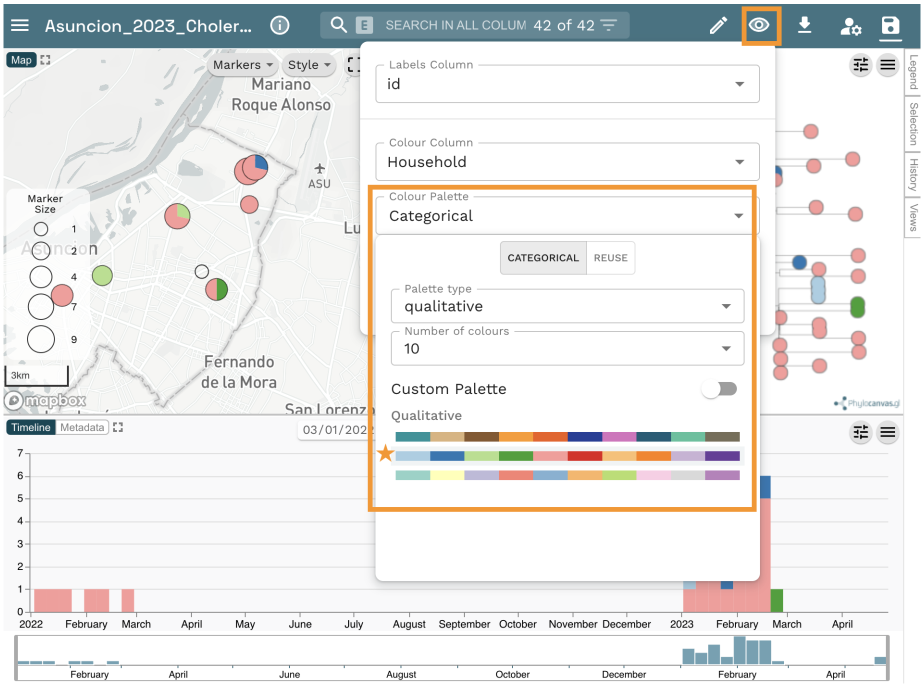
Microreact assigns colours automatically to every column in your metadata table, but you can customise them. Let's change the colour column to "Household". Click on the eye icon at the top left of the page. Then select "Household" from the `Colour Column` dropdown list.



Labels Column
id

Colour Column
Household

Now Select "Categorical" from the `Colour Palette` option. Next Select "qualitative" from the `Palette type` dropdown list and "10" from the `Number of colours` dropdown list. Finally, select the second option from the list of palettes.



Labels Column: id

Colour Column: Household

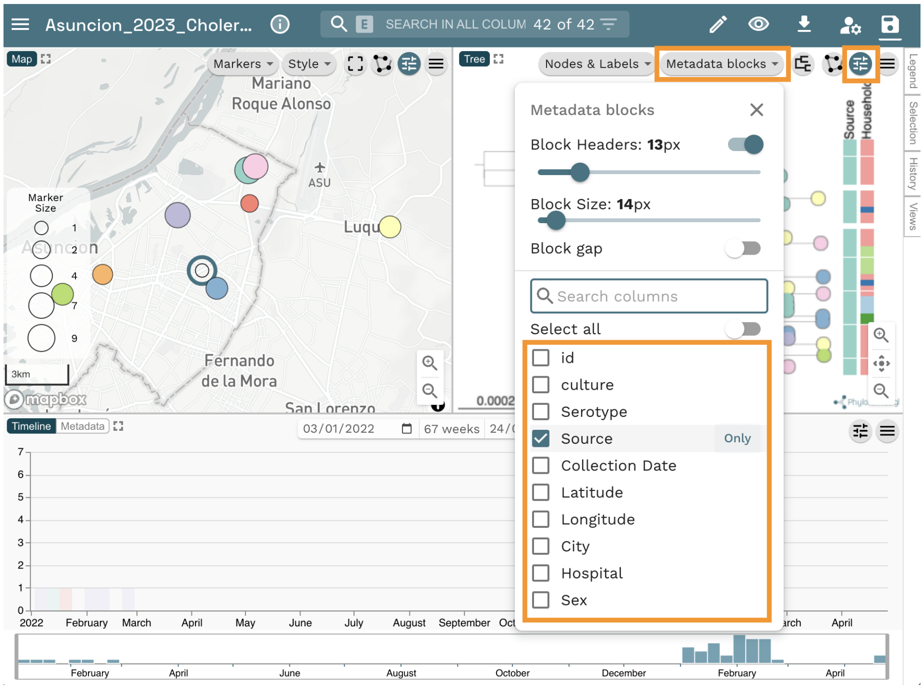
Colour Palette: Categorical

Palette type: qualitative

Number of colours: 10

Custom Palette

Change the colour column back to “Hospital” and add the columns “Source” and “Household” as metadata blocks. For this, select the slider icon from top right of the map panel, then, click on the `Metadata blocks` button. Finally, select “Source” and “Household” from the list of options.



Metadata blocks ✕

Block Headers: **13**px

Block Size: **14**px

Block gap

Search columns

Select all

□ id
□ culture
□ Serotype
☑ Source   Only
□ Collection Date
□ Latitude
□ Longitude
□ City
□ Hospital
□ Sex

5. Do any of the cases share a household? Taking into account the incubation period for cholera (a few hours to five days), does the data support intra-household transmission?

Resources

Install and test Epicollect5 on your mobile phone

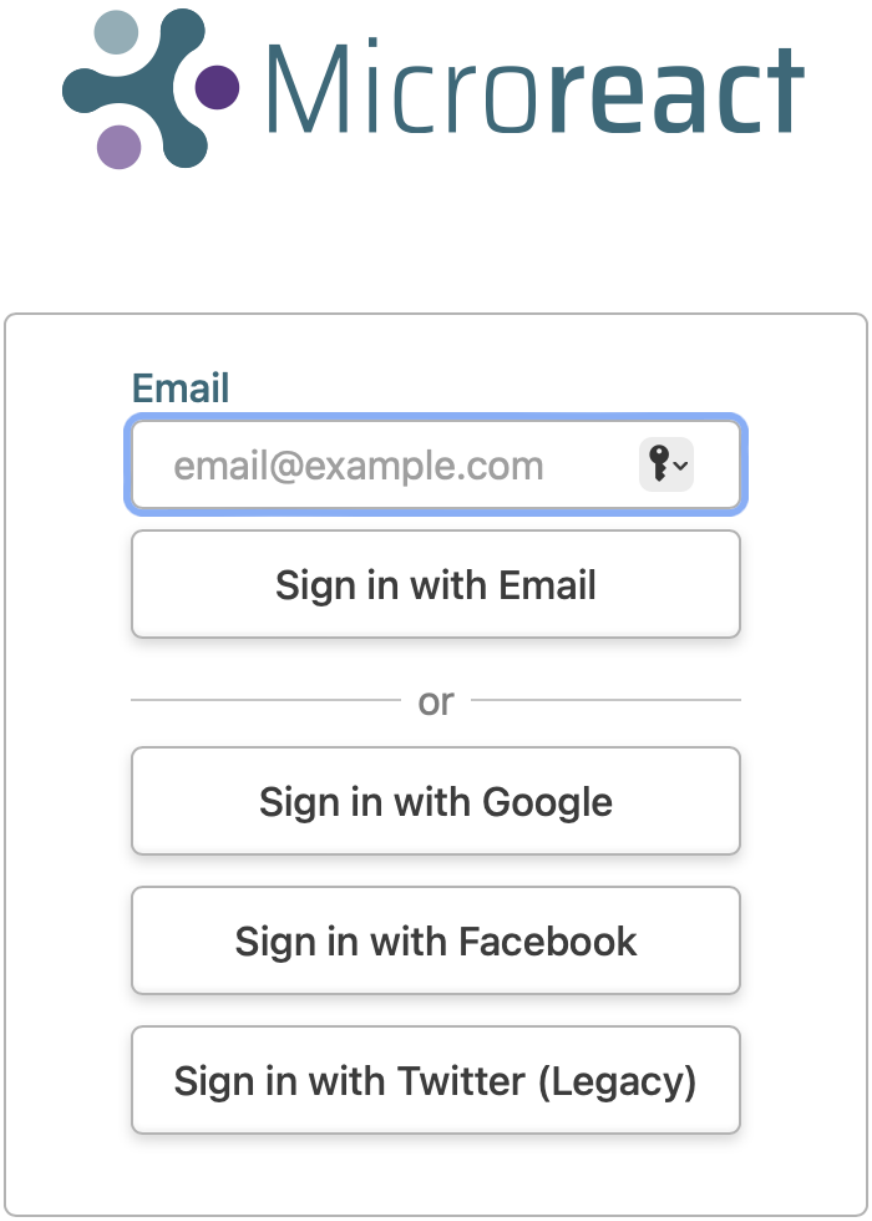
Epicollect5 is available on Android and iOS. To install it on your mobile device, get it from Google Play or the Apple App Store. To test Epicollect5 on your phone, open the application and click on the pre-loaded EC5 Demo Project. Click on `+ Add entry` and follow the questionnaire. Please answer all the questions, including taking a photo with your phone's camera. Once you've answered the questions, save and upload the entry. Please note that you must be connected to the internet (or to a mobile network) to be able to upload an entry.

Sign-in to data-flo

https://docs.data-flo.io/introduction/getting-started-sign-in

Sign-in to Microreact

Head over to **https://microreact.org/my-account**, and sign up using your email address or any other authenticating option



Web colours

Web colours are used on web pages, such as the microreact.org projects, and are usually specified in hexadecimal format preceded by a number sign (or hashtag sign if you were born before the Jurassic period). For example, `#FFFFFF` is the hexadecimal code (or hex code) for the colour “white”, `#000000` is the hex code for the colour “black”, and `#FF00FF` is the hex code for the colour “magenta”. For more information on the format of the hexadecimal code see Wikipedia page **https://en.wikipedia.org/wiki/Web_colors**

Useful websites to work with web colours:

ColorBrewer

https://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3 Colour palettes for maps

ColorHexa

https://www.colorhexa.com/ Get complementary colours, colour gradients, etc.

Coolors

https://coolors.co/palettes/trending Colour palette generator

Wes Anderson colour palettes

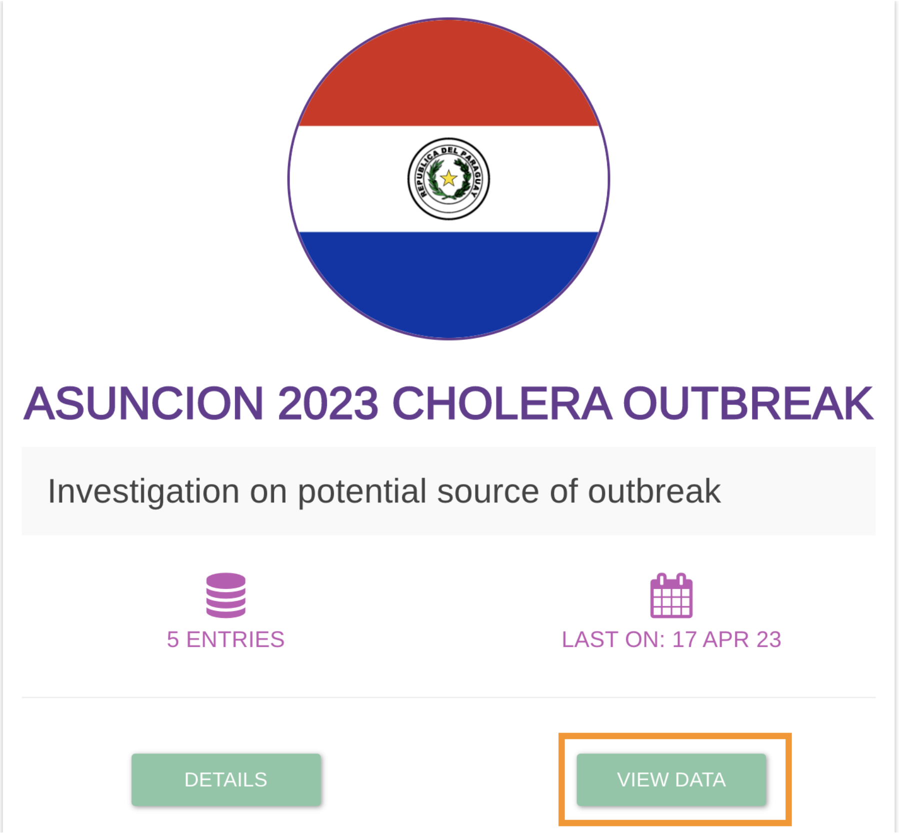
https://github.com/karthik/wesanderson Hipster colour palettes inspired by Wes Anderson movies

Bonus activities

1. Download data from the Epicollect5 project on the web.

a. Go to the **epicollect link provided above**.

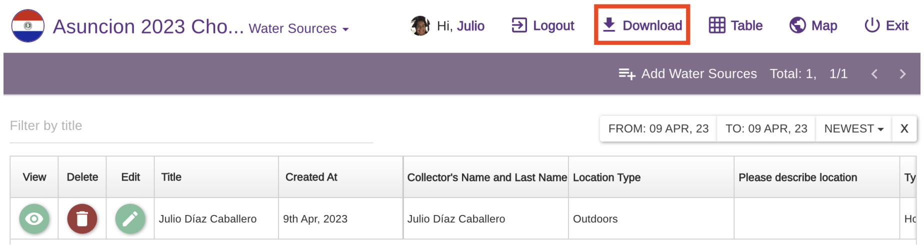
b. Select `VIEW DATA`.



ASUNCION 2023 CHOLERA OUTBREAK

Investigation on potential source of outbreak

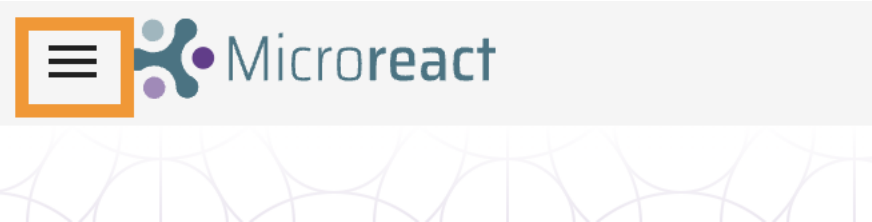
c. Click on `Download` and select your format of choice.



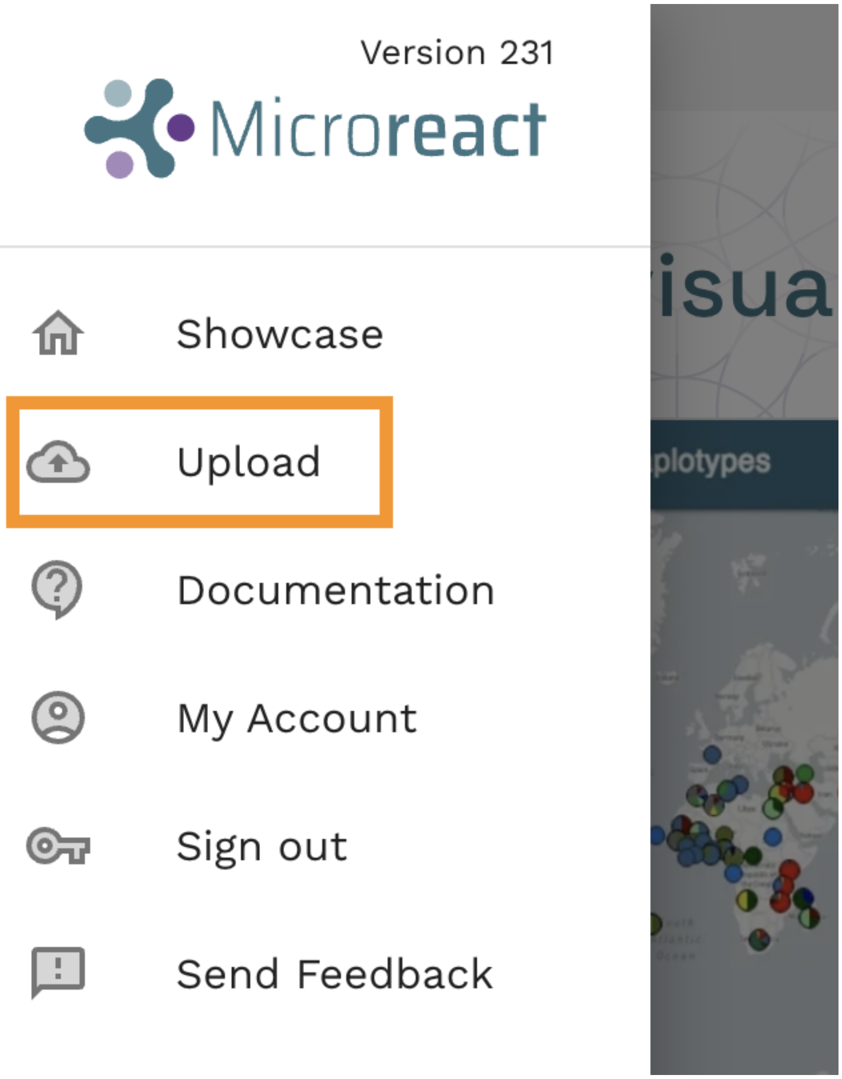
View	Delete	Edit	Title	Created At	Collector's Name and Last Name	Location Type	Please describe location	Ty
👁	🗑	✎	Julio Díaz Caballero	9th Apr, 2023	Julio Díaz Caballero	Outdoors		Ho

2. Create a microreact project using a metadata in csv format and a phylogenetic tree.

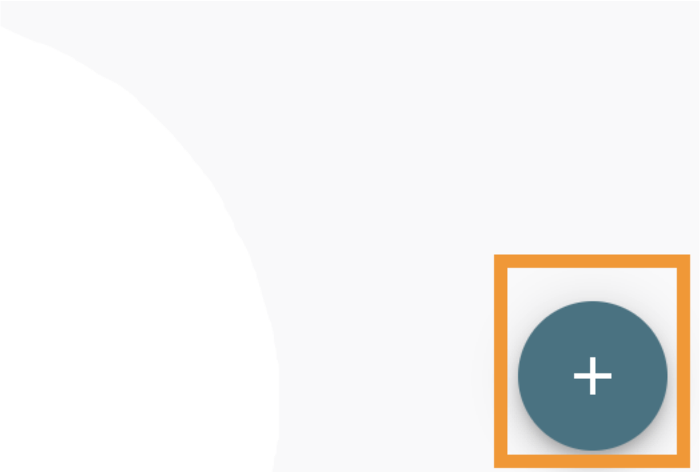
a. Click on the menu icon at the top left of the screen



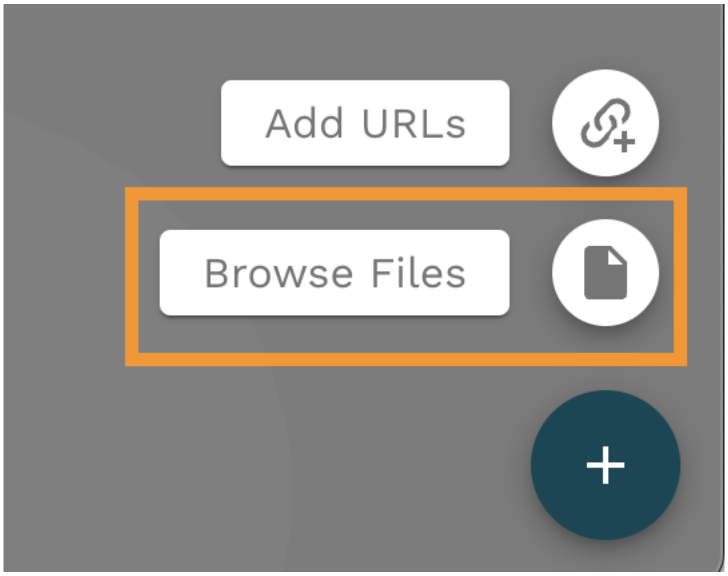
b. Click on `Upload`



c. Click on the + symbol at the bottom right of the screen.



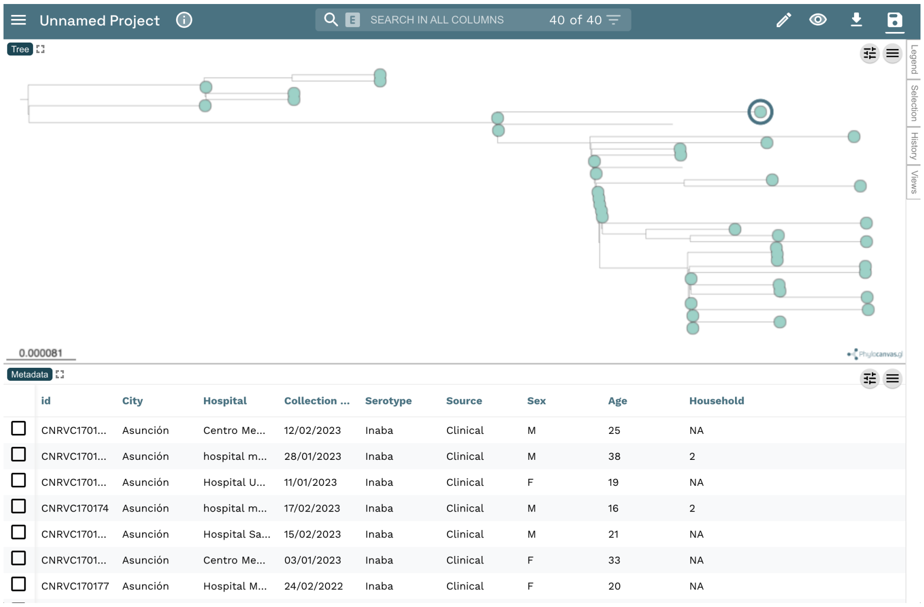
d. Select `Browse Files`



e. Choose the files `tree.nwk` and `epi_data.csv`, and click on the `Open` button (These files were downloaded in the dataflo section).

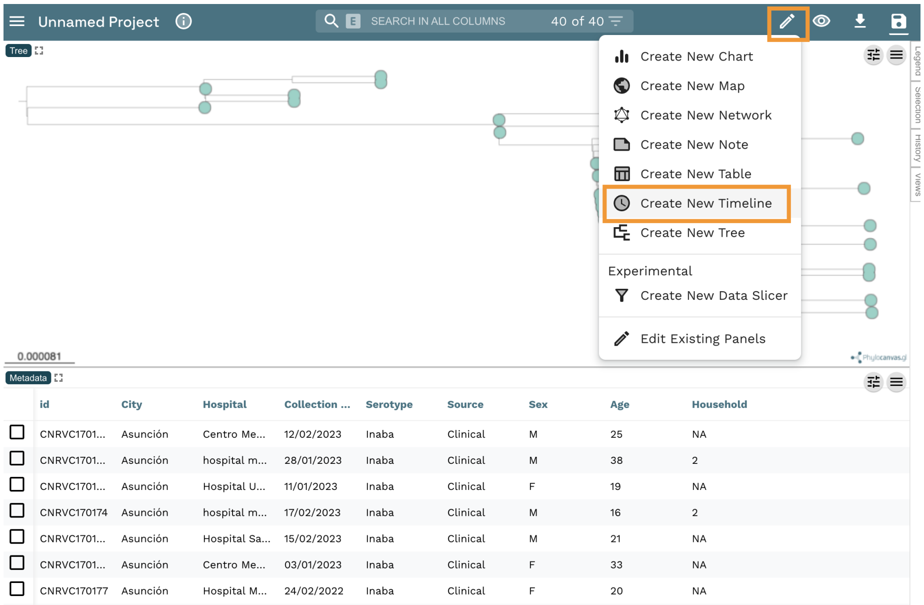
f. Click on `Continue` in the next prompts.

g. Your view will look like this:



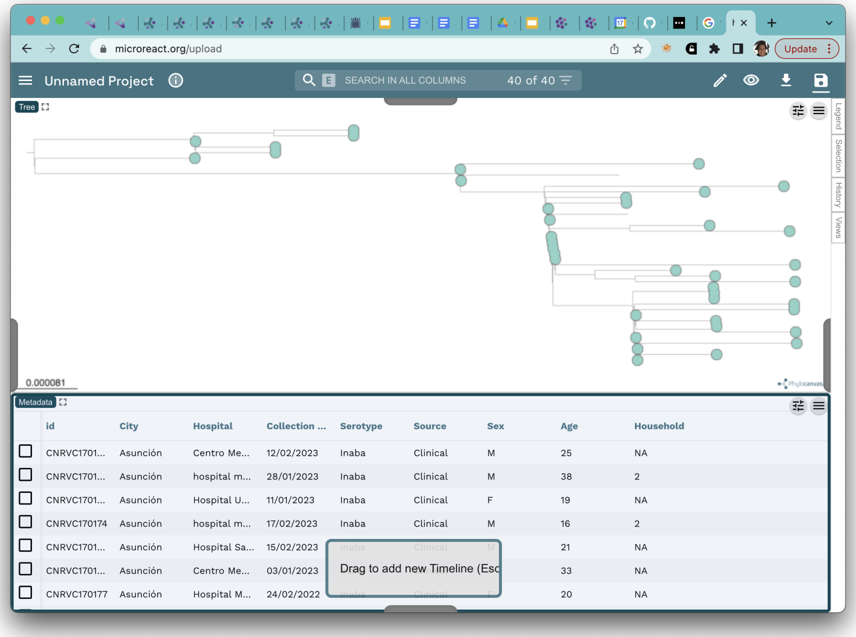
	id	City	Hospital	Collection ...	Serotype	Source	Sex	Age	Household
□	CNRVC1701...	Asunción	Centro Me...	12/02/2023	Inaba	Clinical	M	25	NA
□	CNRVC1701...	Asunción	hospital m...	28/01/2023	Inaba	Clinical	M	38	2
□	CNRVC1701...	Asunción	Hospital U...	11/01/2023	Inaba	Clinical	F	19	NA
□	CNRVC170174	Asunción	hospital m...	17/02/2023	Inaba	Clinical	M	16	2
□	CNRVC1701...	Asunción	Hospital Sa...	15/02/2023	Inaba	Clinical	M	21	NA
□	CNRVC1701...	Asunción	Centro Me...	03/01/2023	Inaba	Clinical	F	33	NA
□	CNRVC170177	Asunción	Hospital M...	24/02/2022	Inaba	Clinical	F	20	NA

h. You can add a Timeline by clicking on the `pencil` icon at the top left of the creen, and clicking on `Create New Timeline`



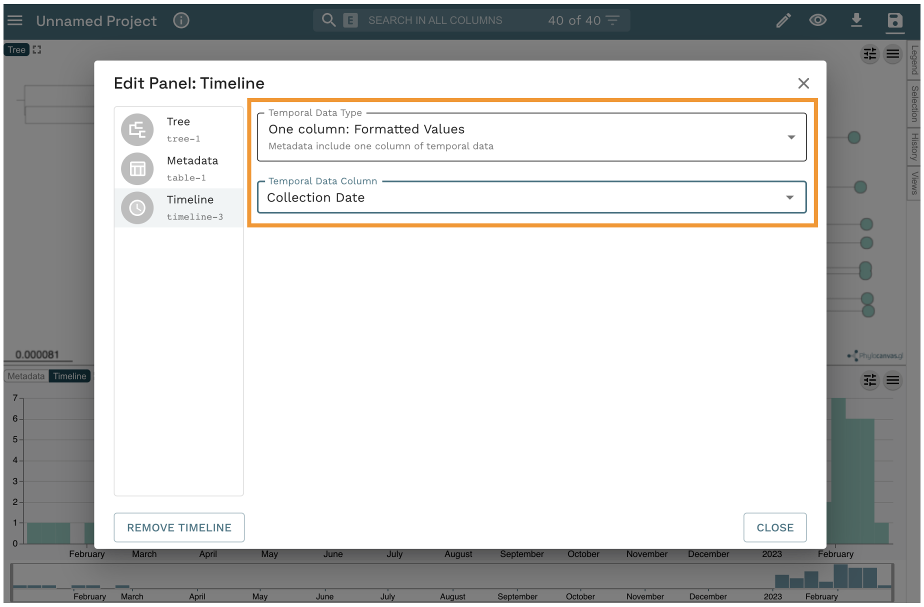
	id	City	Hospital	Collection ...	Serotype	Source	Sex	Age	Household
□	CNRVC1701...	Asunción	Centro Me...	12/02/2023	Inaba	Clinical	M	25	NA
□	CNRVC1701...	Asunción	hospital m...	28/01/2023	Inaba	Clinical	M	38	2
□	CNRVC1701...	Asunción	Hospital U...	11/01/2023	Inaba	Clinical	F	19	NA
□	CNRVC170174	Asunción	hospital m...	17/02/2023	Inaba	Clinical	M	16	2
□	CNRVC1701...	Asunción	Hospital Sa...	15/02/2023	Inaba	Clinical	M	21	NA
□	CNRVC1701...	Asunción	Centro Me...	03/01/2023	Inaba	Clinical	F	33	NA
□	CNRVC170177	Asunción	Hospital M...	24/02/2022	Inaba	Clinical	F	20	NA

i. Then position the new panel by dragging the pointer to the bottom panel.



	id	City	Hospital	Collection ...	Serotype	Source	Sex	Age	Household
□	CNRVC1701...	Asunción	Centro Me...	12/02/2023	Inaba	Clinical	M	25	NA
□	CNRVC1701...	Asunción	hospital m...	28/01/2023	Inaba	Clinical	M	38	2
□	CNRVC1701...	Asunción	Hospital U...	11/01/2023	Inaba	Clinical	F	19	NA
□	CNRVC170174	Asunción	hospital m...	17/02/2023	Inaba	Clinical	M	16	2
□	CNRVC1701...	Asunción	Hospital Sa...	15/02/2023	Inaba	Clinical	M	21	NA
□	CNRVC1701...	Asunción	Centro Me...	03/01/2023				33	NA
□	CNRVC170177	Asunción	Hospital M...	24/02/2022	Inaba	Clinical	F	20	NA

j. Finally, select “One column: Formatted Values” from the `Temporal Data Type` dropdown list, and “Collection Date” from the `Temporal Data Column` dropdown list. Click on `CLOSE`.

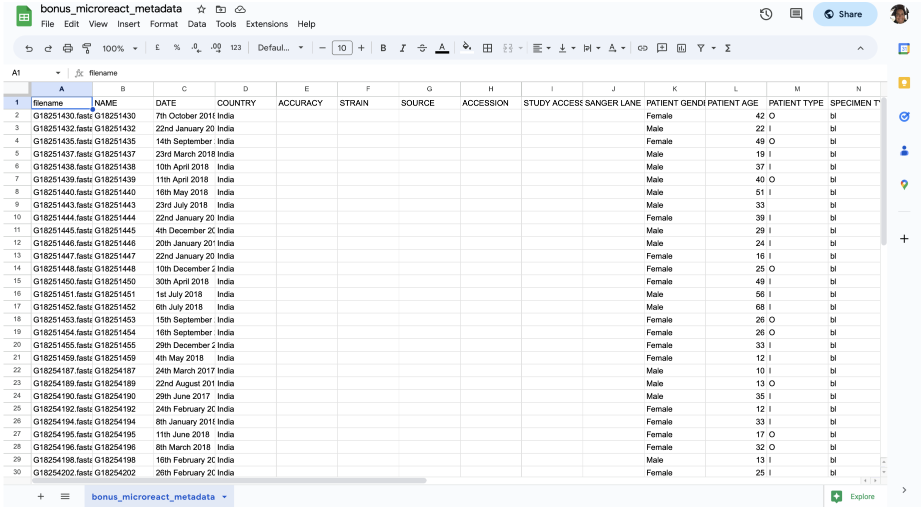


Temporal Data Type
One column: Formatted Values
Metadata include one column of temporal data

Temporal Data Column
Collection Date

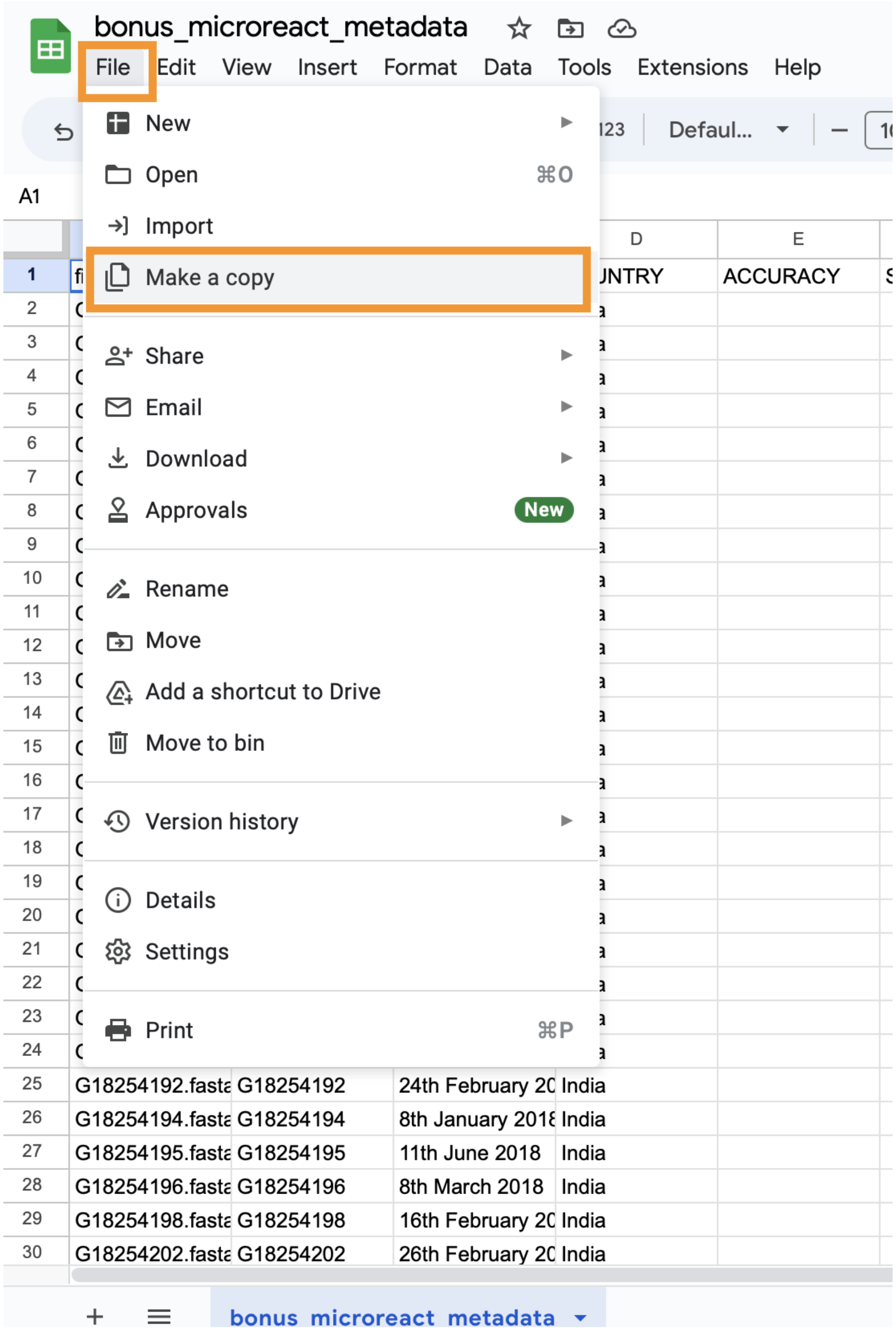
3. Create a microreact project from a Google Spreadsheet

a. Open this **Google spreadsheet**

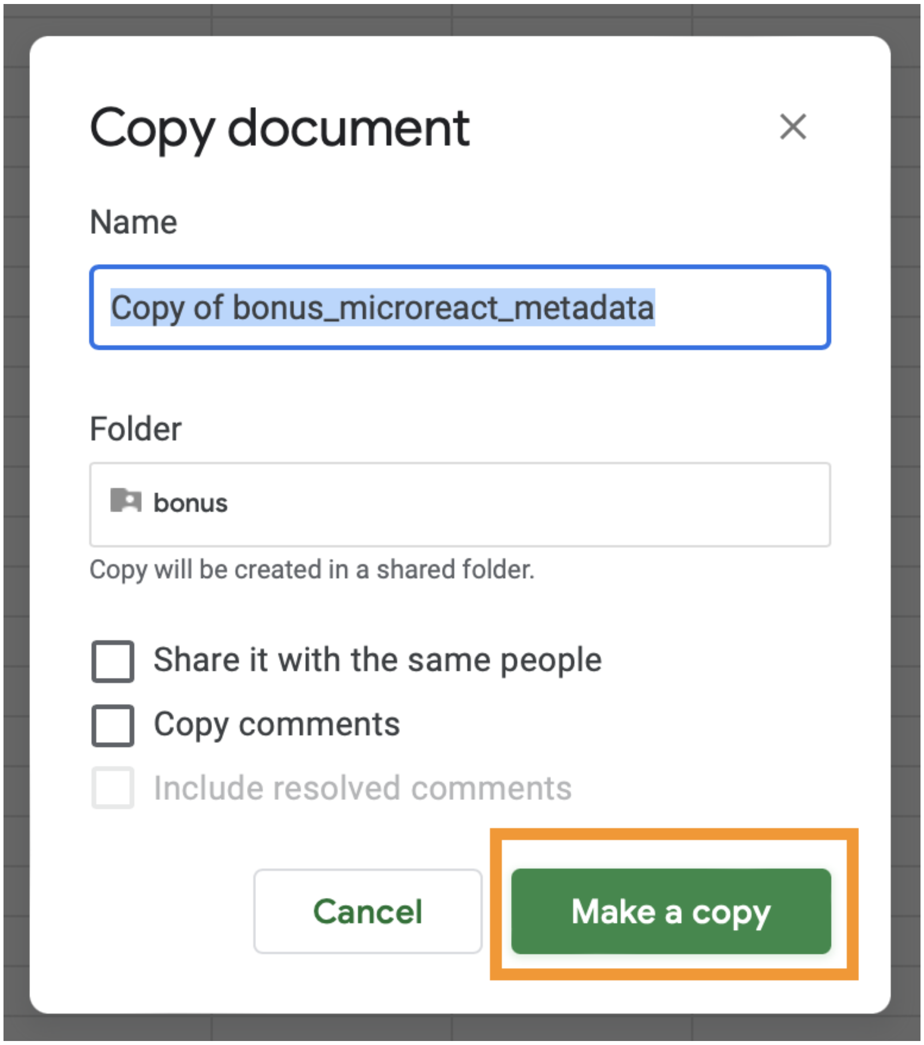


	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	filename	NAME	DATE	COUNTRY	ACCURACY	STRAIN	SOURCE	ACCESSION	STUDY ACCESS	SANGER LANE	PATIENT GENDI	PATIENT AGE	PATIENT TYPE	SPECIMEN T
2	G18251430.fasta	G18251430	7th October 2018	India							Female	42	O	bl
3	G18251432.fasta	G18251432	22nd January 20	India							Male	22	I	bl
4	G18251435.fasta	G18251435	14th September	India							Female	49	O	bl
5	G18251437.fasta	G18251437	23rd March 2018	India							Male	19	I	bl
6	G18251438.fasta	G18251438	10th April 2018	India							Male	37	I	bl
7	G18251439.fasta	G18251439	11th April 2018	India							Male	40	O	bl
8	G18251440.fasta	G18251440	16th May 2018	India							Male	51	I	bl
9	G18251443.fasta	G18251443	23rd July 2018	India							Male	33		bl
10	G18251444.fasta	G18251444	22nd January 20	India							Female	39	I	bl
11	G18251445.fasta	G18251445	4th December 20	India							Male	29	I	bl
12	G18251446.fasta	G18251446	20th January 201	India							Male	24	I	bl
13	G18251447.fasta	G18251447	22nd January 20	India							Female	16	I	bl
14	G18251448.fasta	G18251448	10th December 2	India							Female	25	O	bl
15	G18251450.fasta	G18251450	30th April 2018	India							Female	49	I	bl
16	G18251451.fasta	G18251451	1st July 2018	India							Male	56	I	bl
17	G18251452.fasta	G18251452	6th July 2018	India							Male	68	I	bl
18	G18251453.fasta	G18251453	15th September	India							Female	26	O	bl
19	G18251454.fasta	G18251454	16th September	India							Female	26	O	bl
20	G18251455.fasta	G18251455	29th December 2	India							Female	33	I	bl
21	G18251459.fasta	G18251459	4th May 2018	India							Female	12	I	bl
22	G18254187.fasta	G18254187	24th March 2017	India							Male	10	I	bl
23	G18254189.fasta	G18254189	22nd August 201	India							Male	13	O	bl
24	G18254190.fasta	G18254190	29th June 2017	India							Male	35	I	bl
25	G18254192.fasta	G18254192	24th February 20	India							Female	12	I	bl
26	G18254194.fasta	G18254194	8th January 2018	India							Female	33	I	bl
27	G18254195.fasta	G18254195	11th June 2018	India							Female	17	O	bl
28	G18254196.fasta	G18254196	8th March 2018	India							Female	32	O	bl
29	G18254198.fasta	G18254198	16th February 20	India							Male	13	I	bl
30	G18254202.fasta	G18254202	26th February 20	India							Female	25	I	bl

b. Make a copy of this in your own Google account by selecting `Make a copy` from the `File` menu item. When prompted, click on `Make a copy`. This will open a separate tab where the copy will be available.



				D	E
1	fi			JNTRY	ACCURACY
25	G18254192.fasta	G18254192	24th February 20	India	
26	G18254194.fasta	G18254194	8th January 2018	India	
27	G18254195.fasta	G18254195	11th June 2018	India	
28	G18254196.fasta	G18254196	8th March 2018	India	
29	G18254198.fasta	G18254198	16th February 20	India	
30	G18254202.fasta	G18254202	26th February 20	India	



Copy document

Name

Copy of bonus_microreact_metadata

Folder

bonus

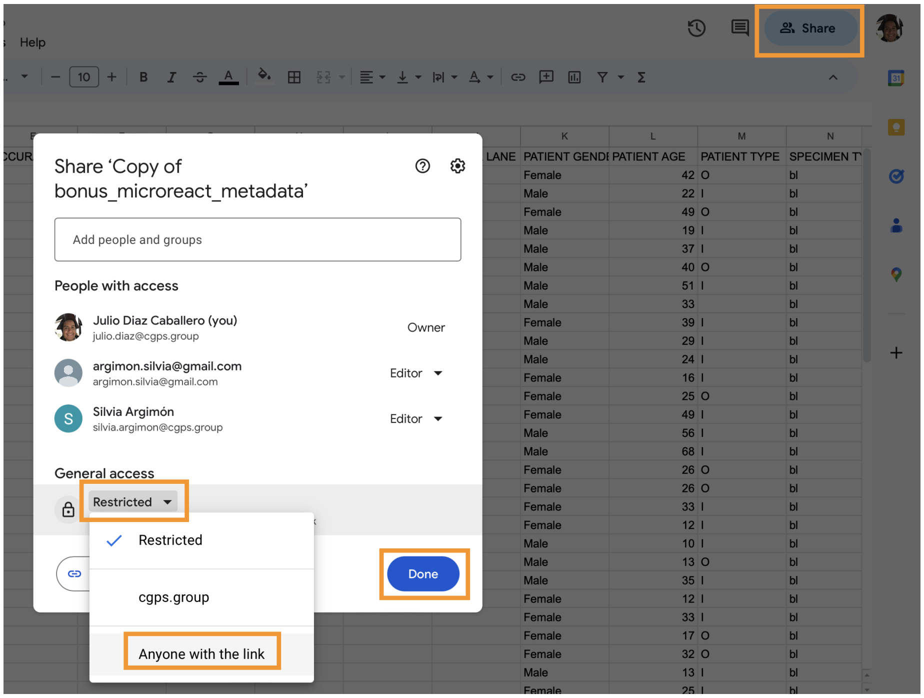
Copy will be created in a shared folder.

□ Share it with the same people

□ Copy comments

□ Include resolved comments

c. Set access to shareable by clicking on the `Share` button at the top right of the screen. A dialogue screen will popoup, here, click on the `Restricted` button, select `Anyone with the link`, and click on `Done`



Share 'Copy of bonus_microreact_metadata'

Add people and groups

People with access

Julio Diaz Caballero (you)
julio.diaz@cgps.group — Owner

argimon.silvia@gmail.com
argimon.silvia@gmail.com — Editor

Silvia Argimón
silvia.argimon@cgps.group — Editor

General access

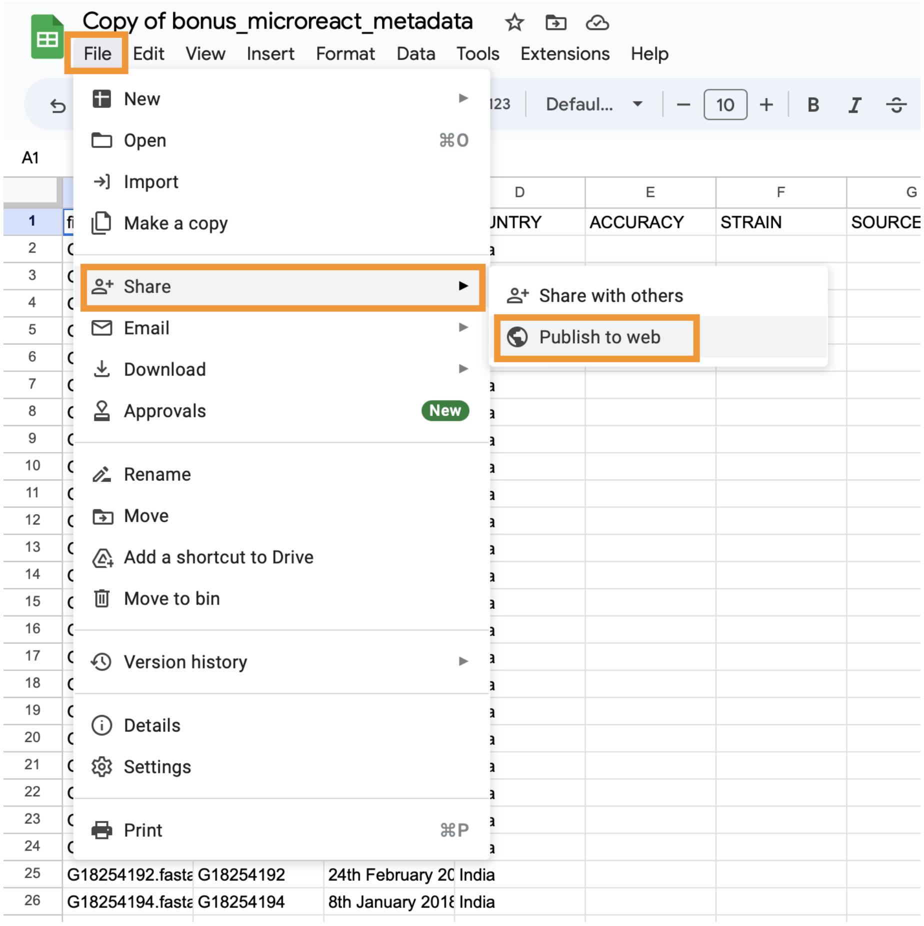
Restricted

- ✓ Restricted
- cgps.group
- Anyone with the link

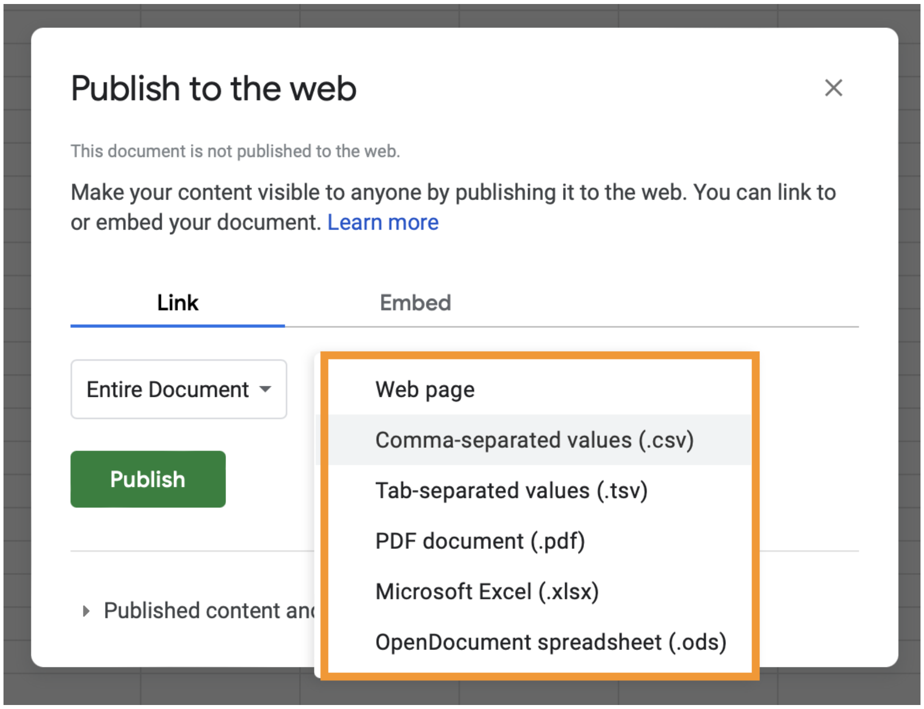
Done

LANE	PATIENT GENDI	PATIENT AGE	PATIENT TYPE	SPECIMEN T\
	Female	42	O	bl
	Male	22	I	bl
	Female	49	O	bl
	Male	19	I	bl
	Male	37	I	bl
	Male	40	O	bl
	Male	51	I	bl
	Male	33		bl
	Female	39	I	bl
	Male	29	I	bl
	Male	24	I	bl
	Female	16	I	bl
	Female	25	O	bl
	Female	49	I	bl
	Male	56	I	bl
	Male	68	I	bl
	Female	26	O	bl
	Female	26	O	bl
	Female	33	I	bl
	Female	12	I	bl
	Male	10	I	bl
	Male	13	O	bl
	Male	35	I	bl
	Female	12	I	bl
	Female	33	I	bl
	Female	17	O	bl
	Female	32	O	bl
	Male	13	I	bl
	Female	25	I	bl

d. Now to publish the google spreadsheet click on the `File` menu item, select `Share`, and click on `Publish on web`



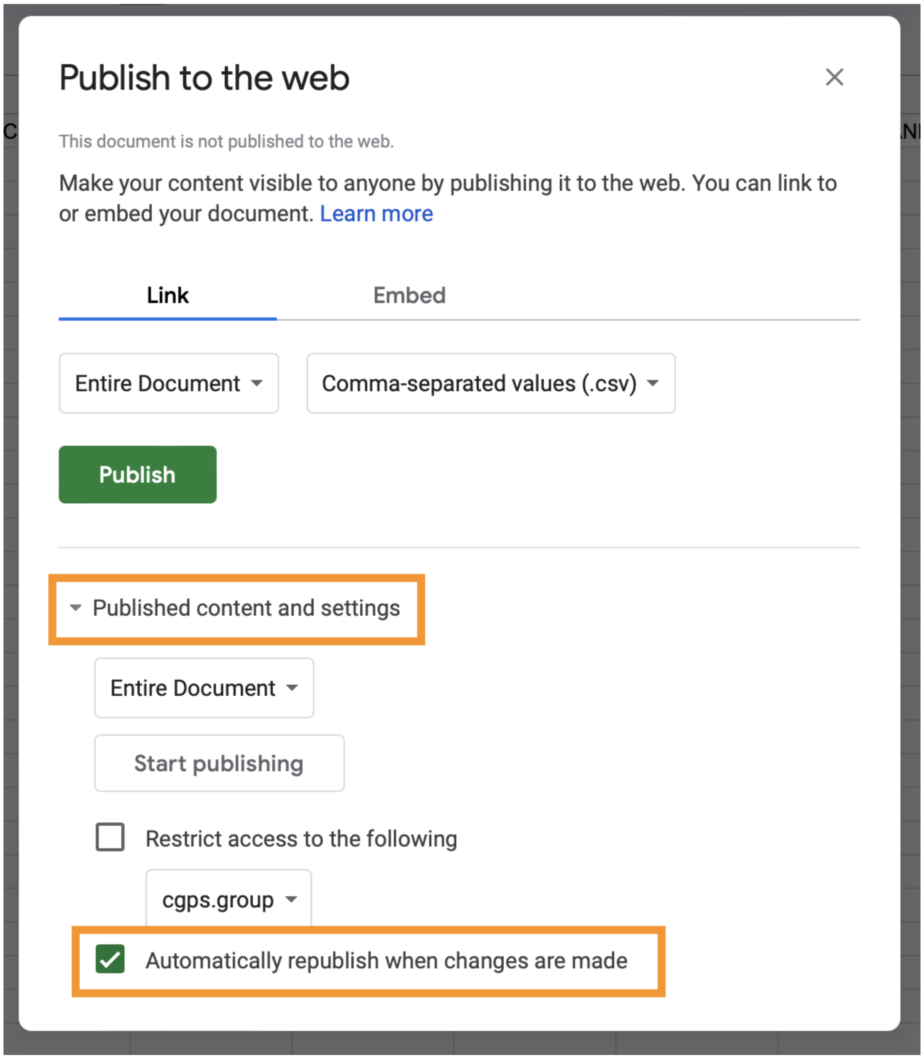
e. In the popup message click on `Web page`, and select `Comma-separated values (.csv)`. Also, make sure the `Automatically republish when changes are made` option is turned on under the `Published content and settings` section



Publish to the web

This document is not published to the web.

Make your content visible to anyone by publishing it to the web. You can link to or embed your document. Learn more



Publish to the web

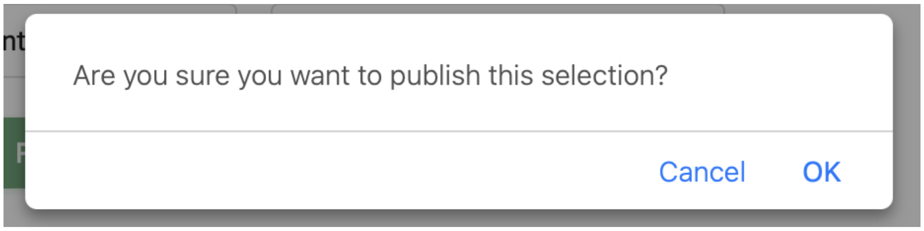
This document is not published to the web.

Make your content visible to anyone by publishing it to the web. You can link to or embed your document. Learn more

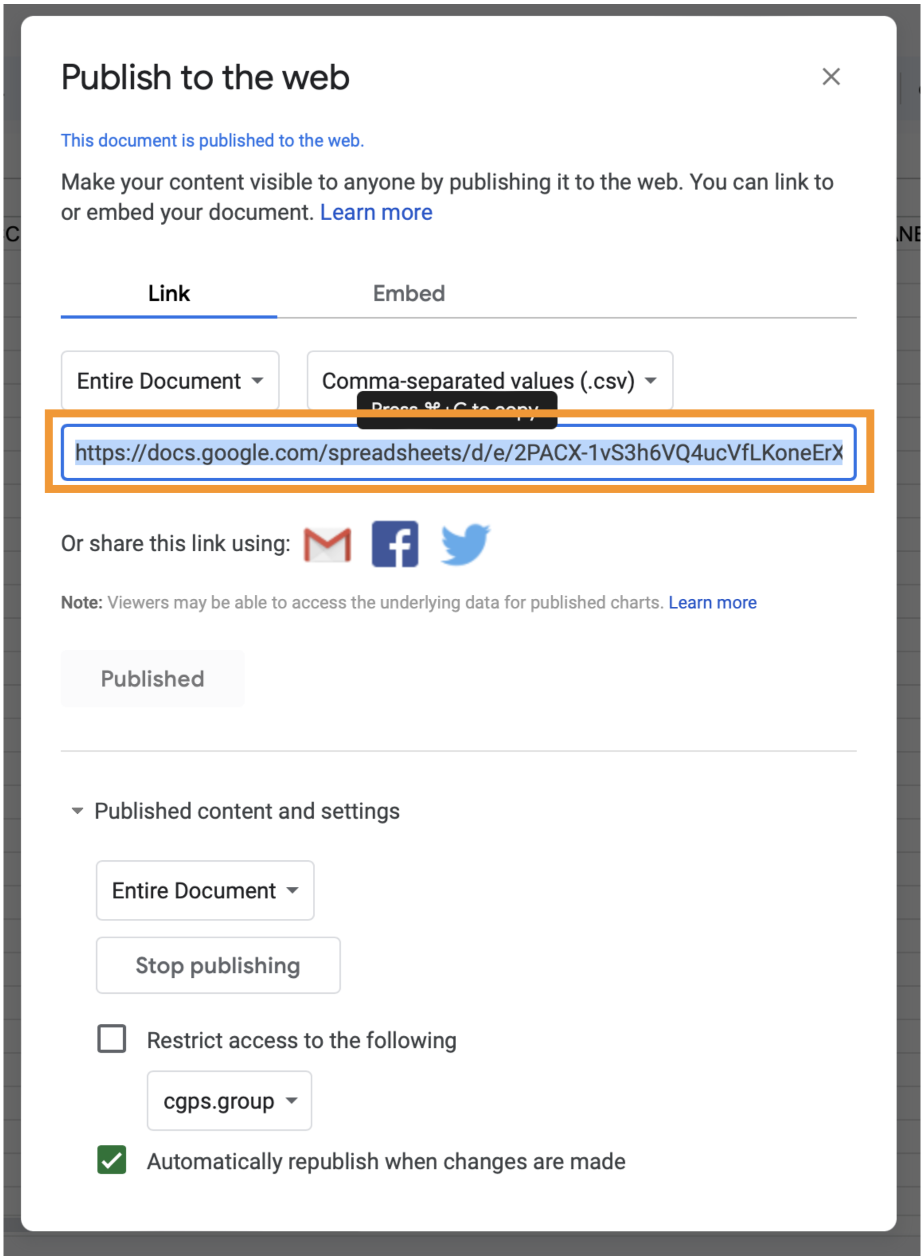
□ Restrict access to the following

☑ Automatically republish when changes are made

f. Confirm your choices in the popup message clicking on `OK`. This will provide more details about your Google spreadsheet, copy the url.



Are you sure you want to publish this selection?



Publish to the web

This document is published to the web.

Make your content visible to anyone by publishing it to the web. You can link to or embed your document. Learn more

Entire Document ▾ Comma-separated values (.csv) ▾

Press ⌘+C to copy

https://docs.google.com/spreadsheets/d/e/2PACX-1vS3h6VQ4ucVfLKoneErX

Or share this link using:

Note: Viewers may be able to access the underlying data for published charts. Learn more

▾ Published content and settings

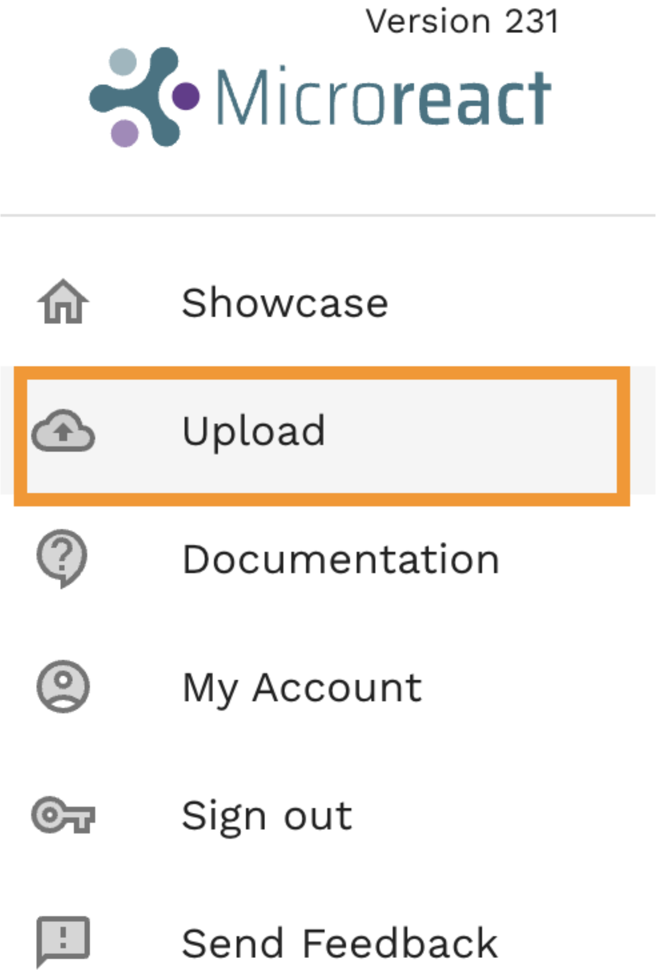
Entire Document ▾

□ Restrict access to the following

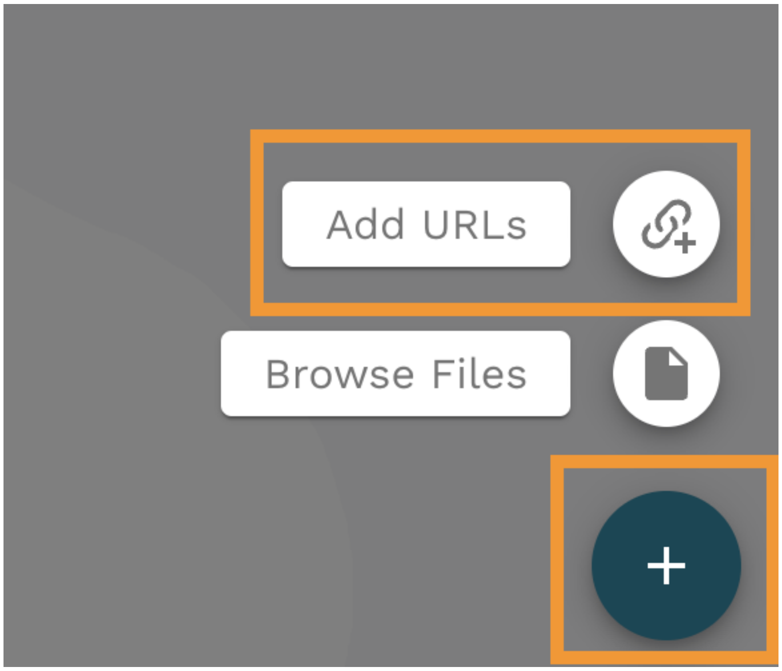
cgps.group ▾

☑ Automatically republish when changes are made

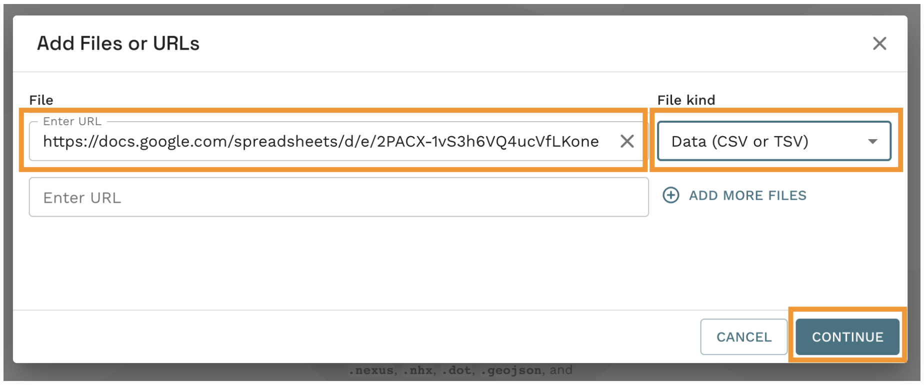
g. Go to **microreact.org**, and select `Upload` from the main menu.



h. Select the plus icon at the bottom right, and click on `Add URLs`



i. Paste the url from step *f* and select `Data (CSV or TSV)` under `File kind`. Finally click on `CONTINUE` to see your microreact.



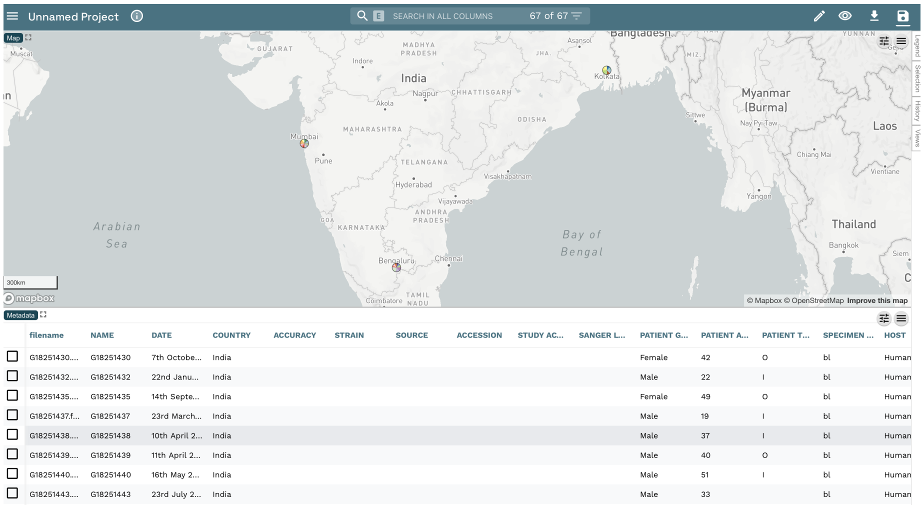
Add Files or URLs

File

Enter URL: https://docs.google.com/spreadsheets/d/e/2PACX-1vS3h6VQ4ucVfLKone

File kind: Data (CSV or TSV)

j. You should get a screen like this!



	filename	NAME	DATE	COUNTRY	ACCURACY	STRAIN	SOURCE	ACCESSION	STUDY AC...	SANGER L...	PATIENT G...	PATIENT A...	PATIENT T...	SPECIMEN ...	HOST
□	G18251430....	G18251430	7th Octobe...	India							Female	42	O	bl	Human
□	G18251432....	G18251432	22nd Janu...	India							Male	22	I	bl	Human
□	G18251435....	G18251435	14th Septe...	India							Female	49	O	bl	Human
□	G18251437.f...	G18251437	23rd March...	India							Male	19	I	bl	Human
□	G18251438....	G18251438	10th April 2...	India							Male	37	I	bl	Human
□	G18251439....	G18251439	11th April 2...	India							Male	40	O	bl	Human
□	G18251440....	G18251440	16th May 2...	India							Male	51	I	bl	Human
□	G18251443....	G18251443	23rd July 2...	India							Male	33		bl	Human