BRIN
BADAN RISET
DAN INOVASI NASIONAL

ESCAP
Economic and Social Commission
for Asia and the Pacific

Training on Leveraging Large Language Models for Big Earth Data Processing: Mapping Flood Hotspots and Impacts

Level: Beginner

Pre-requisites: Basic knowledge of remote sensing and flood mapping

Introduction

The increased availability of freely available remote sensing data and other open geospatial datasets from space agencies, commercial entities, and other organizations has transformed space applications across disciplines, fostering multidisciplinary research and collaboration. Historical satellite imagery provides opportunities for time series analysis to detect changes over the past decades, while tools and cloud computing platforms like Google Earth Engine facilitate big data analytics. Cloud computing has revolutionized remote sensing applications by increasing the efficiency of image analysis, from preprocessing to image interpretation and analysis. The emergence of generative artificial intelligence (AI) has further increased the potential scope of the applications of Earth observation data by simplifying data processing through user prompts that signal the creation of geospatial data products for various disciplines.

Earth observation data plays a crucial role in disaster risk reduction, enabling the understanding of the disaster cycle from preparedness to rehabilitation and recovery. Flooding is one of the most common disasters due to severe weather events. It can damage infrastructure and ecosystems, disrupt livelihoods, and pose significant risks to human health and safety. Assessing flood impacts through satellite imagery analysis is vital to emergency response, designing rehabilitation schemes, and implementing recovery efforts.

Learning Objectives

In this short training, participants will take advantage of SatGPT's capability to map flooding and inundations hotspots using generative AI and Google Earth Engine. This dataset can be integrated with other geospatial data, such as land cover and socioeconomic datasets, to assess the potential impact of flooding in the selected area of interest. At the end of this hands-on session, participants are expected to replicate the methodology in assessing the extent of flooding, identifying affected areas, and estimating losses, using additional geospatial datasets related to agriculture, infrastructure, population, and relevant economic parameters.

Data Input

Data	Source	File Format
Inundation hotspot	SatGPT	GeoTIFF
Administrative boundaries	Humanitarian Data Exchange (HDX)	Esri Shapefile
Land cover	ESA WorldCover 2020	GeoTIFF
Inundation hotspot code (*flood_hotspot_gee_code.js*)	SatGPT	GEE JavaScript code

Software and Platforms

In this session, we will be using QGIS Desktop, an open-source software that allows users to create, edit, visualize, analyze, and publish geospatial information. The long-term release is recommended since this is the most stable release with the latest bugfixes.

SATGPT is an innovative solution that leverages the current capabilities of large language models (LLMs) and integrates them with cloud computing platforms and Earth Observation data. SATGPT represents a fully functional, next-generation spatial decision support system designed for rapid deployment, particularly in resource-limited contexts. Flooding and inundation hotspots will be mapped using SATGPT for the selected area of interest.

Google Earth Engine (GEE) is a cloud-based platform that allows users to access high-performance computing resources and process large Earth observation datasets and derived products. It is designed for users who may not have expertise in programming as well as developers to create data products from satellite image analysis and build interactive applications. Google Earth Engine also provides an extensive public data catalogue with analysis-ready geospatial data, complemented by a community-curated data catalog from researchers and other organizations (see https://gee-community-catalog.org/). It allows users to visualize, process, integrate, manipulate, edit, and create new spatial data products efficiently. We will be using GEE to download the inundation hotspot dataset for the selected area of interest and for advanced flood mapping analysis.

Flooding Scenario: Karawang (West Java), Indonesia

Floodwaters submerged houses in Karawang district, West Java during the February 2021 flooding. (Source: ANTARA Indonesia News Agency/Ibnu Chazar)

Karawang Regency, known as a major agricultural hub and the "national rice granary" of West Java, plays a significant role in Indonesia's rice production and export economy. However, its agricultural sustainability has been increasingly jeopardized by industrialization and urban expansion since the 1990s. However, industrialization and urban expansion since the 1990s have increasingly strained its agricultural land and infrastructure. Flooding has emerged as a persistent threat, impacting agricultural output and livelihoods. In 2017, severe flooding affected 12 subdistricts, with water levels reaching 100–150 cm in Jatisari and Telukjambe Barat.

While flood occurrences fluctuated in subsequent years, with some subdistricts witnessing temporary relief, extensive flooding returned in 2020, affecting almost all areas of Karawang. Effective flood mapping and mitigation strategies are essential to safeguard agricultural livelihoods and support the region's economic and food security contributions amidst increasing pressures from environmental and developmental changes. *(Source: Aruminingsih, A., Martono, D. N., Soesilo, T. E. B., & Tambunan, R. P. (2022). Flood Disaster Risk Model in Karawang Regency's Industrial Area, West Java Province, Indonesia. Indonesian Journal of Geography, 54(1), 70-82.)*

According to SatGPT:

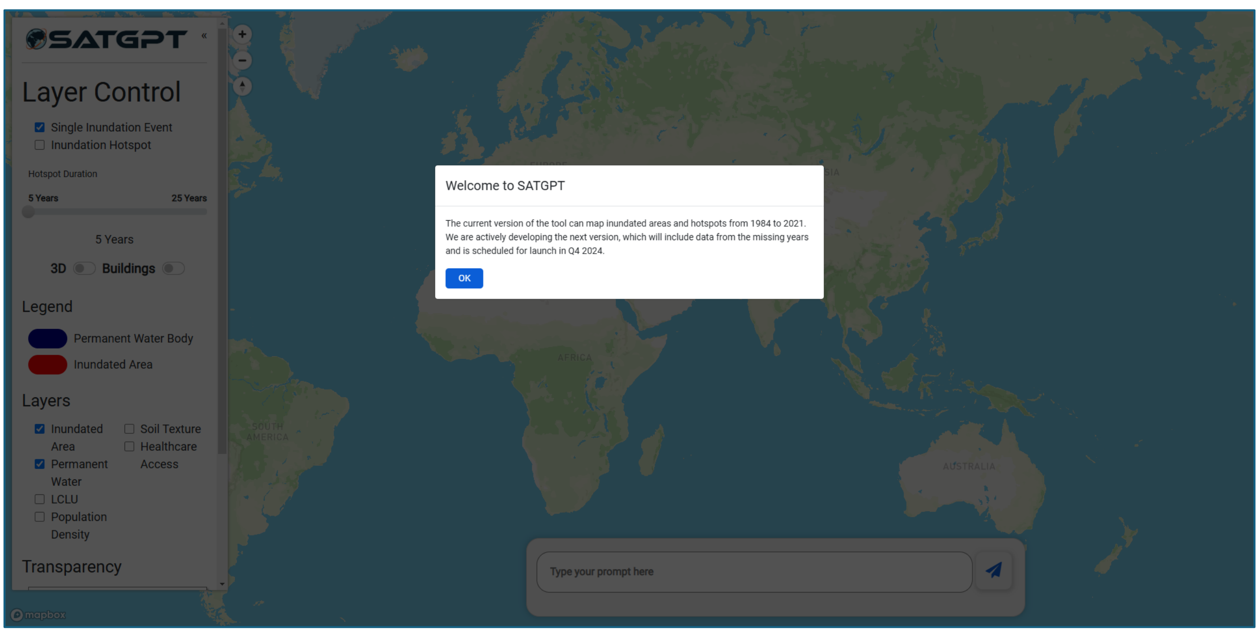
2007-01-15 to 2007-02-02: The floods in Karawang, Indonesia in 2007 caused significant damage to residential areas and agricultural lands. The start date of the flooding was January 15, 2007, and it lasted until February 2, 2007. Many communities in Karawang were affected, leading to disruptions in daily life and economic activities.

2015-12-20 to 2015-12-30: In December 2015, Karawang experienced severe flooding that lasted from December 20 to December 30. The floods inundated roads, homes, and businesses in the region, causing widespread damage. Local authorities worked to assist affected residents and minimize the impact on the community.

2020-01-02 to 2020-01-10: Karawang faced another flooding event in January 2020, with the start date on January 2 and end date on January 10. The floods affected various areas of the city, including residential neighborhoods and industrial zones. Emergency response teams were deployed to provide aid and support to those affected by the disaster.

Part I. Downloading Flood Data from SatGPT

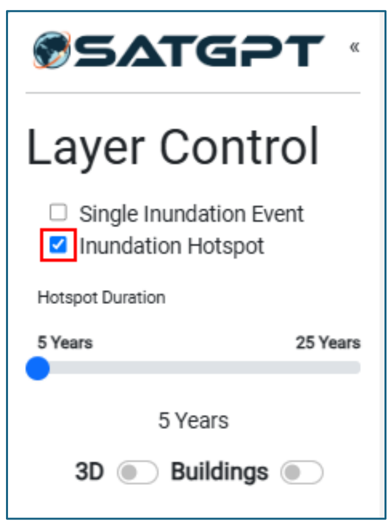
- **Open** your web browser.
- **Navigate** to the SatGPT website: https://satgpt.net/



Here is an overview of the SatGPT graphical user interface:

To get an overview on how to use SatGPT, you can check the user guide by **clicking** *Help* on the left panel.

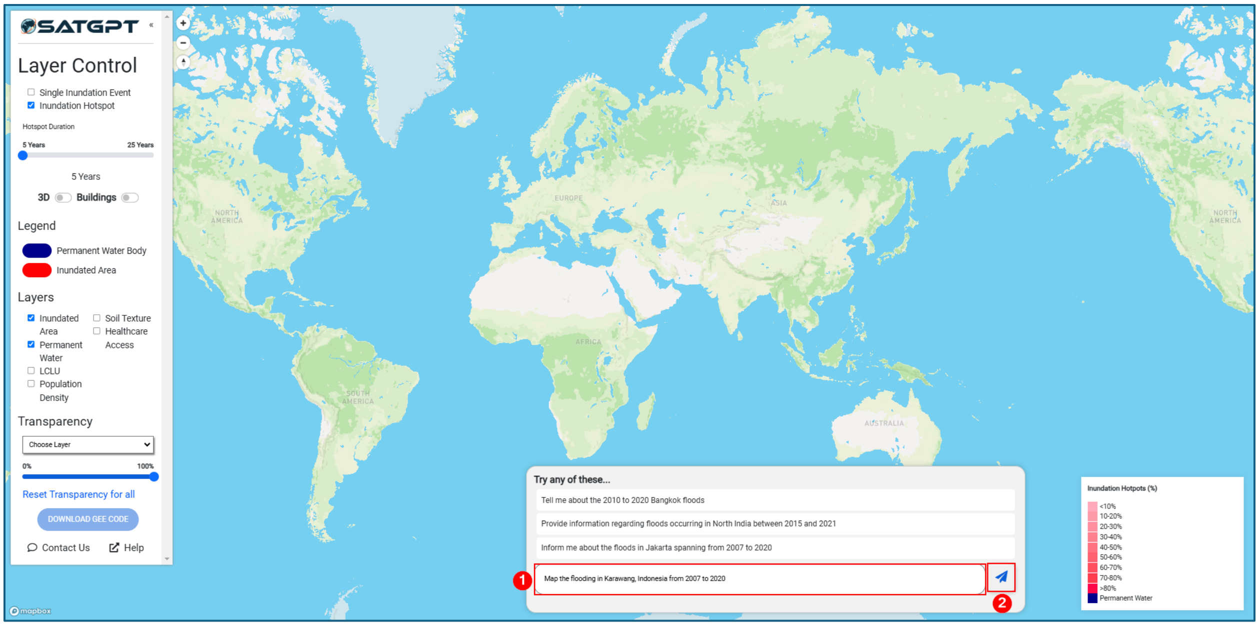
- On the left panel, **click** on the checkbox *Single Inundation Event* to remove this layer for the processing.
- Now, **click** on the checkbox *Inundation Hotspot* to select this layer for the processing.



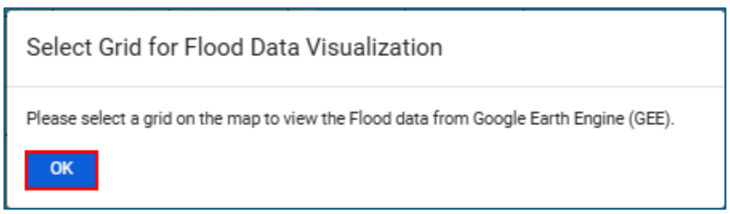
- In the user prompt, **input** the following:

Map the flooding in Karawang, Indonesia from 2007 to 2020

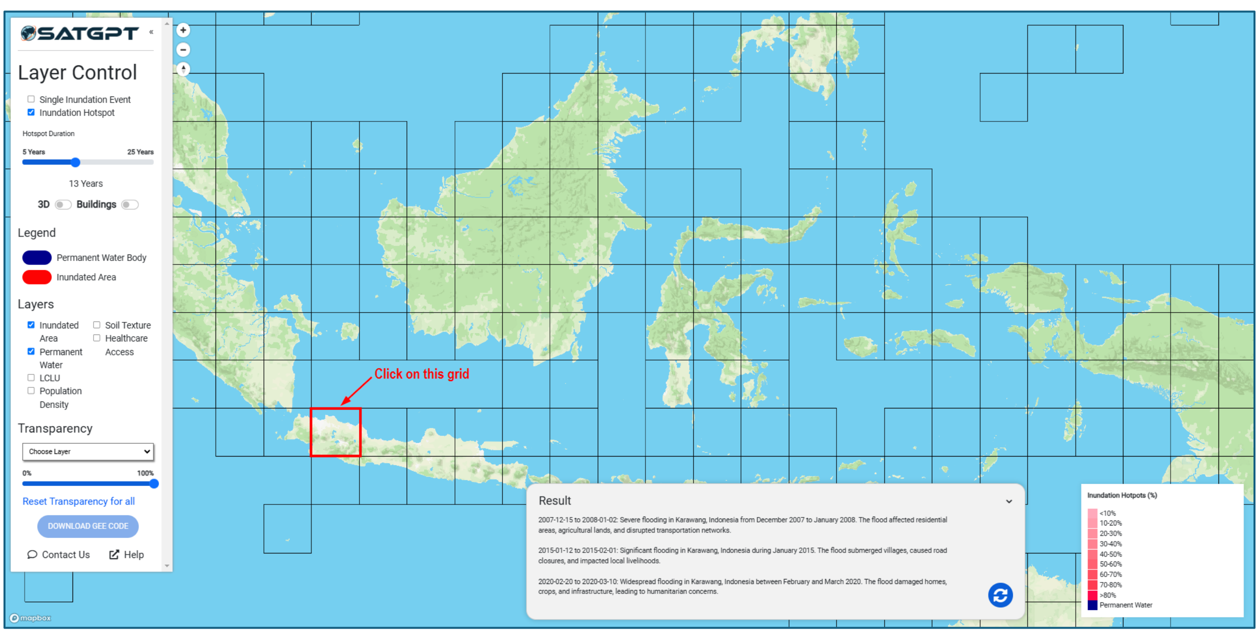
- **Click** the paper plane icon  beside the user prompt to process your request.



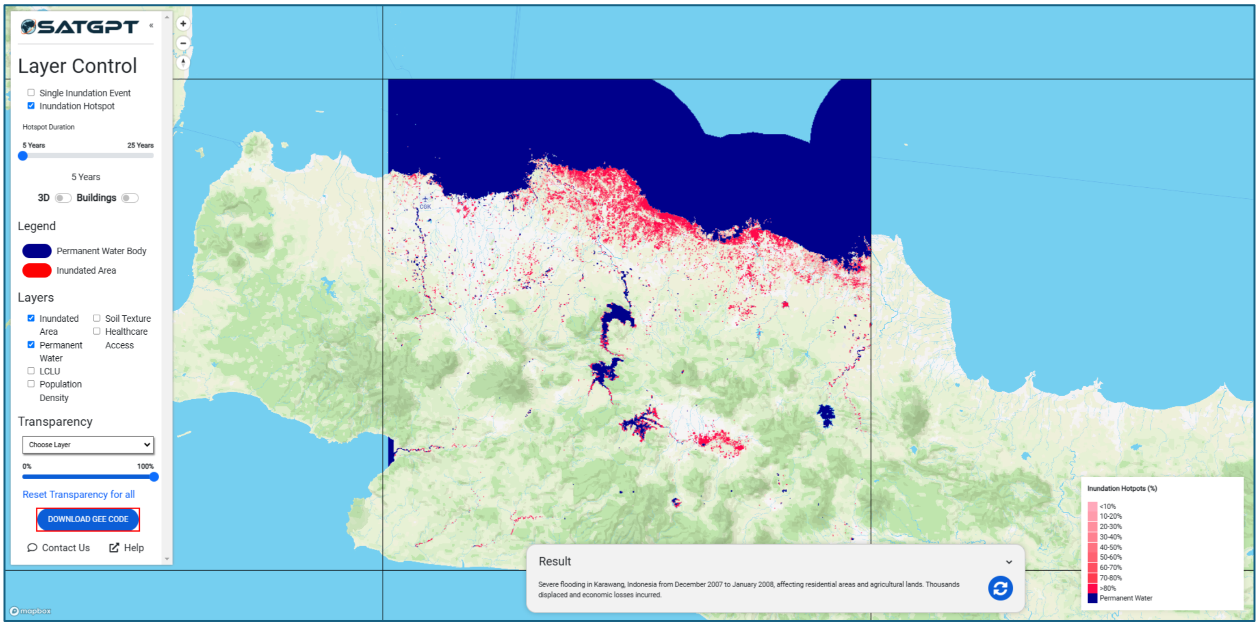
- The map will automatically zoom in to Indonesia. A dialog box will appear that will require you to select the grid that will define the area of interest for your request. **Click** *OK* to proceed.



- **Click** on the grid shown in the figure below to define your area of interest.



- Once the processing is finished, the inundation hotspot layer will be added to the map. An additional legend panel will also be shown on the lower left of the map for the inundation hotspots and permanent water body. Your result should be similar to the figure below. To proceed, **click** on *Download GEE* code.



SatGPT uses the Global Surface Water Mapping Layers developed by the Joint Research Centre (JRC) of the European Union to take into account historical information that classifies permanent and seasonal water bodies (to explore, go to https://global-surface-water.appspot.com/). Using this dataset, the flood extents are refined to only include areas that are most likely to be flooded for the time period defined by the user.

The GEE code will be automatically downloaded and based on your download settings, may either be saved directly to your *Downloads* folder or to the folder you select before you proceed to download. By default, the code will have the following filename: *flood_hotspot_gee_code.js*

Now that you have downloaded the JavaScript code, it's time to export the data in GEE!

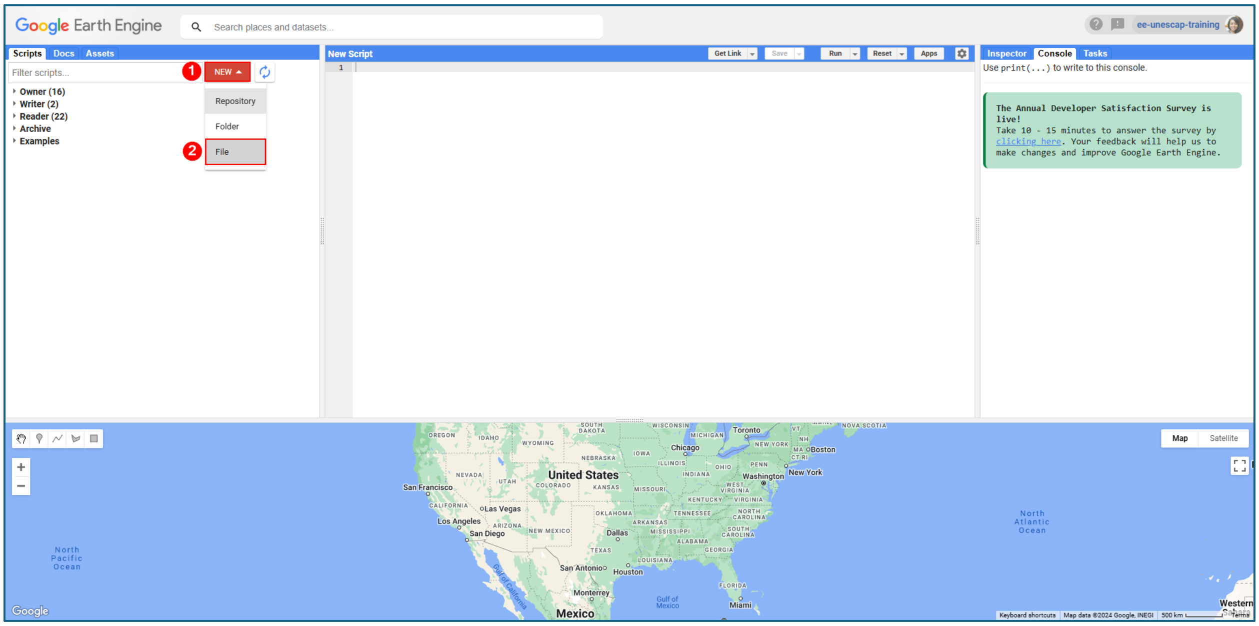
Part II. Exporting the Inundation Hotspot Data in GEE

- In your web browser, **go** to https://code.earthengine.google.com/

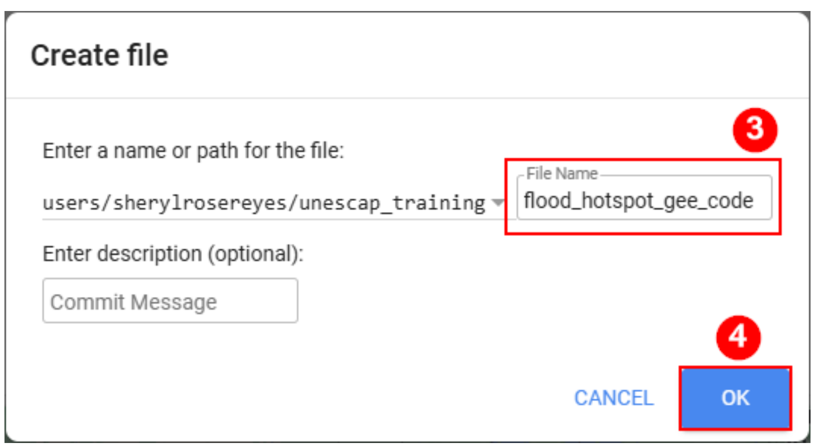
  In order to use Google Earth Engine, a Gmail account is required, and you need to sign up for this specific service. To sign up and start with GEE, you can go to this page: https://earthengine.google.com/

  Follow the steps in the succeeding pages, setup your Google Cloud Project, and you'll be ready to use GEE in no time!

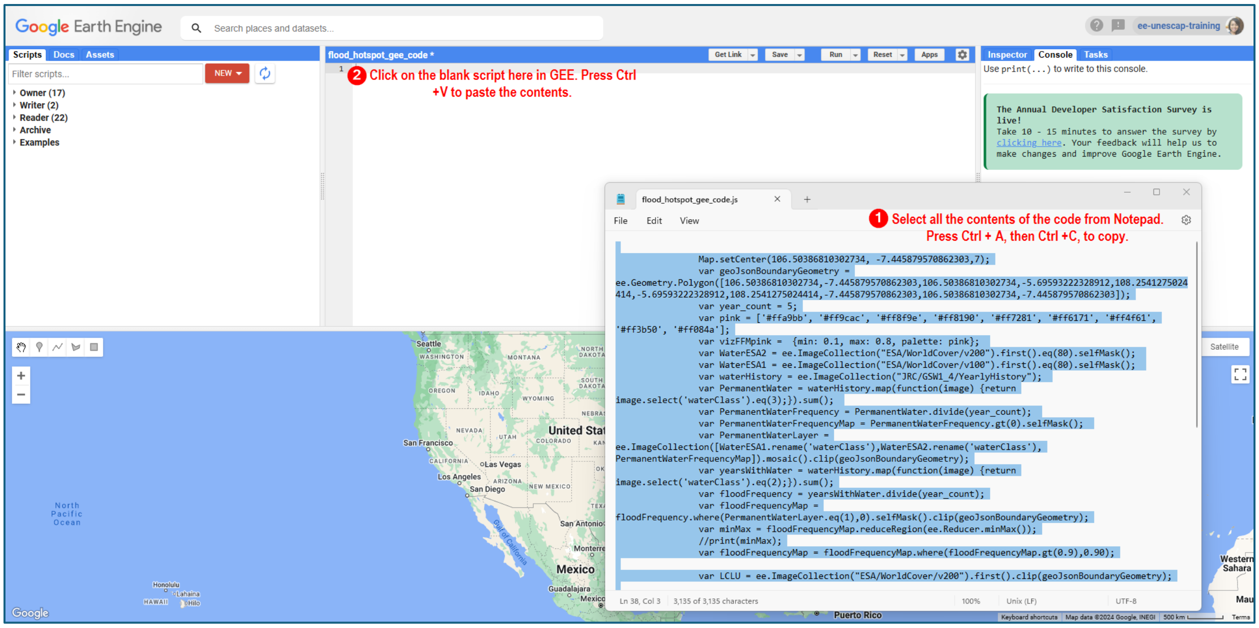
- By default, when you go to the Google Earth Engine Code Editor, a blank script is opened on the *Script Manager*. Alternatively, on the left panel, from the *Scripts* tab, **click** *New*, then **select** *File* to start a new script.



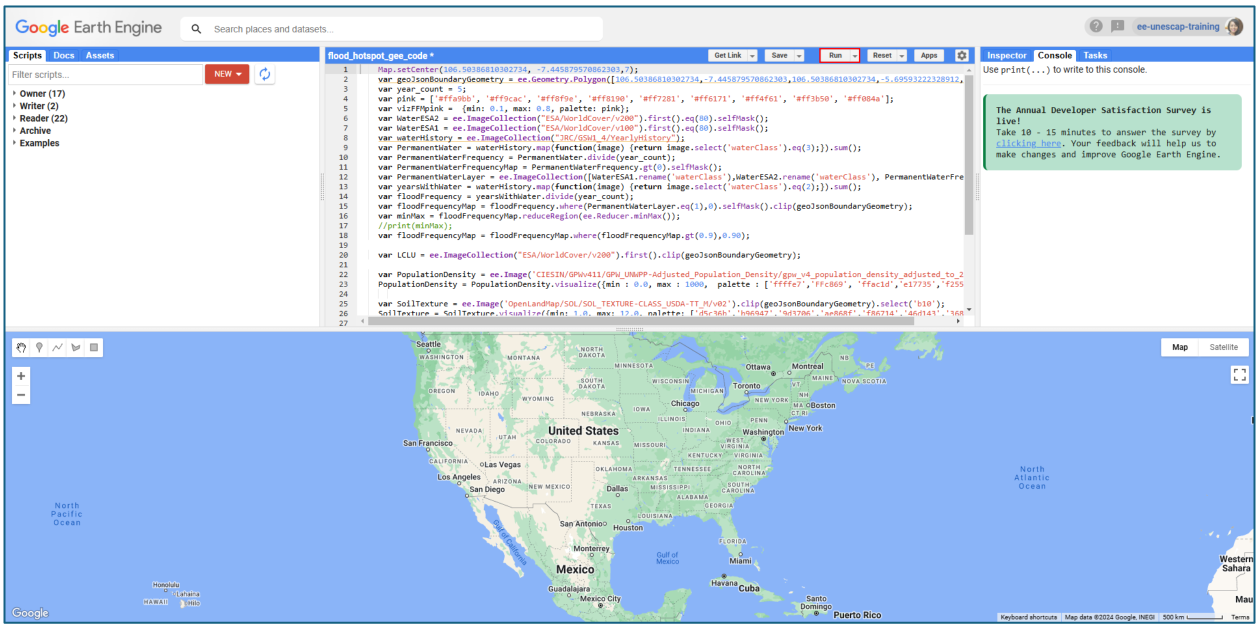
- You will be prompted to **save** this script and assign a filename. Let's use the default filename, *flood_hotspot_gee_code.*



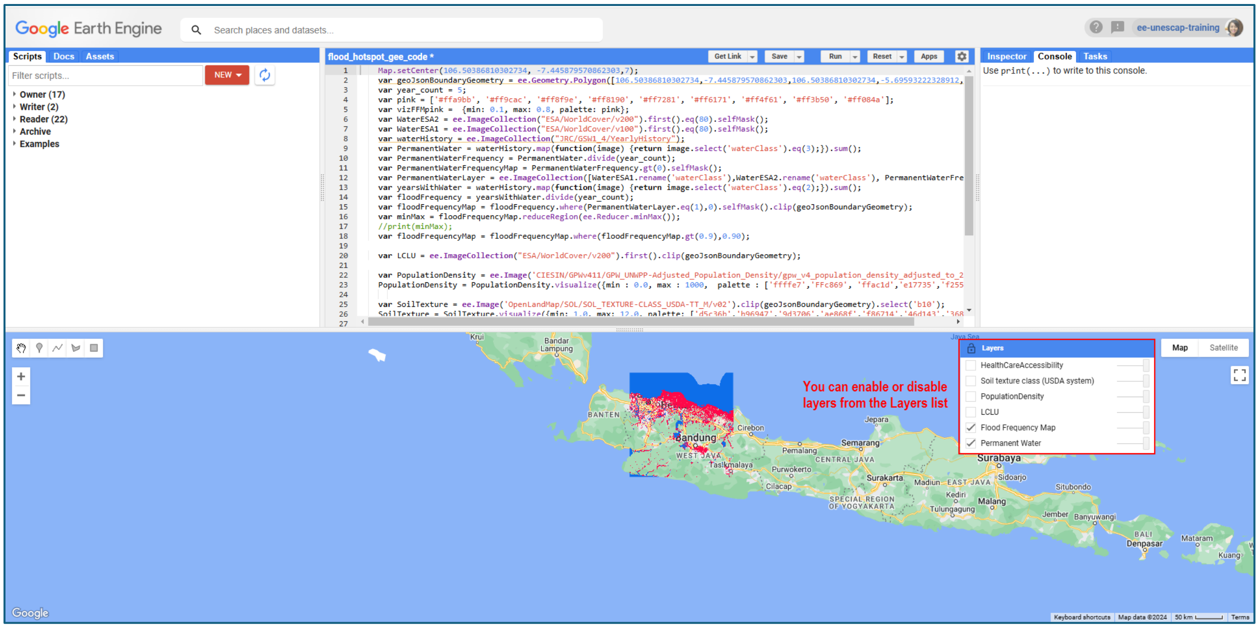
- **Open** the GEE code you downloaded from SATGPT, *flood_hotspot_gee_code.js*, in Notepad. To do this, you can **right-click** on the file, **select** *Open With > Notepad*.
- **Copy** the contents in your new GEE script.
- **Paste** the contents in the blank script you created previously in the *Script Manager*.



- Let's test if the code runs, **click** *Run* from the *Script Manager* window.



You should see the following results:



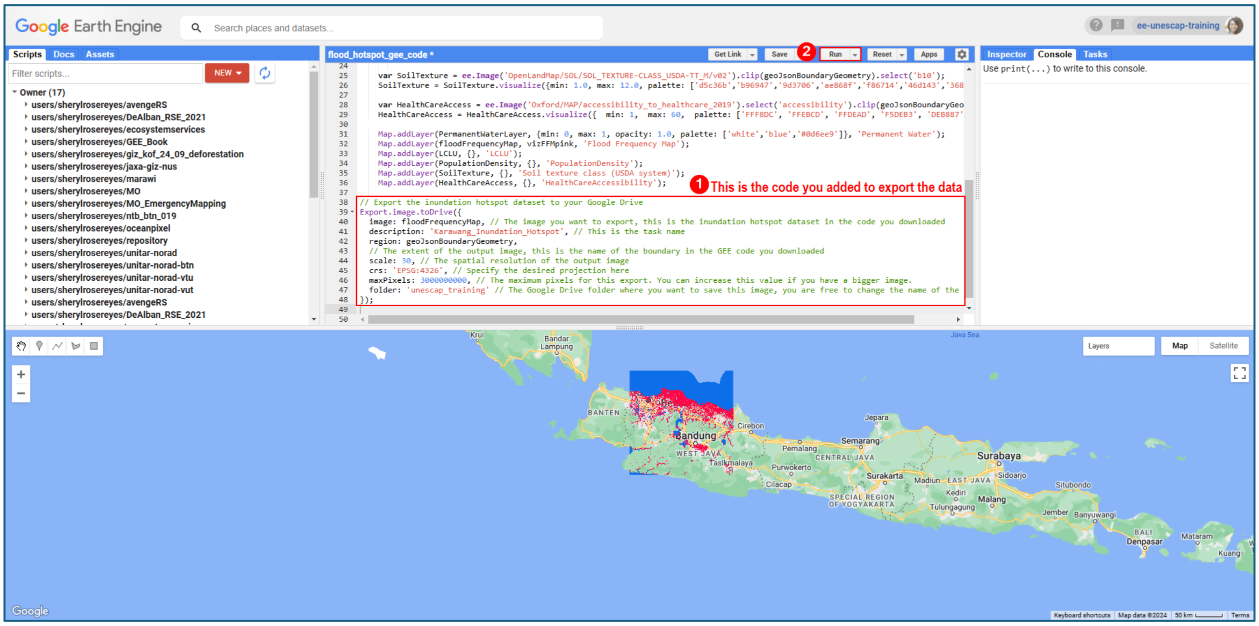
You can enable or disable the available layers from the *Layers* list in the map display as shown in the previous figure.

- Now that we have verified that there are no errors in the code, let's **add the following lines of code** at the end of the flood inundation code to initiate the task to export our inundation hotspot dataset in GEE. The green text are comments in the code to provide you additional information for each line of code. These comments are not part of the actual code but provide additional notes to users for traceability and additional information.

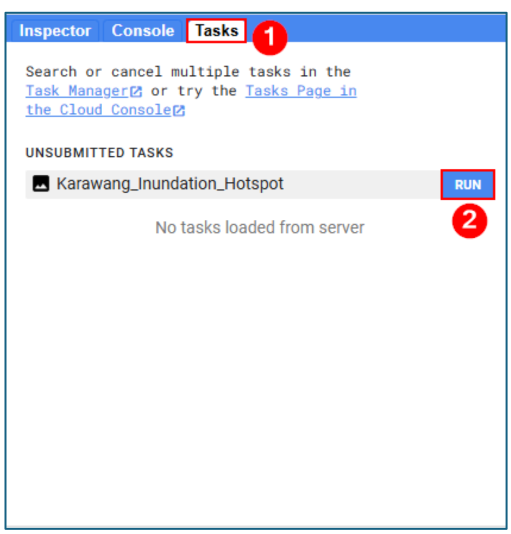
```
// Export the inundation hotspot dataset to your Google Drive
Export.image.toDrive({
  image: floodFrequencyMap, // The image you want to export, this is the inundation hotspot dataset in the code you downloaded
  description: 'Karawang_Inundation_Hotspot', // This is the task name
  region: geoJsonBoundaryGeometry,
  // The extent of the output image, this is the name of the boundary in the GEE code you downloaded
  scale: 30, // The spatial resolution of the output image
  crs: 'EPSG:4326', // Specify the desired projection here
  maxPixels: 3000000000, // The maximum pixels for this export. You can increase this value if you have a bigger image.
  folder: 'unescap_training' // The Google Drive folder where you want to save this image, you are free to change the name of the folder where you want to save the data
});
```

By exporting your processed image in GEE, you will be able to import it into another application and do additional processing and visualization. Take note that image downloads in GEE are limited to a maximum grid dimension of 10,000 pixels (for a total of 10,0000,000 pixels per image download).When exporting your image, there are several parameters that you can define for your output image including the subset of your image (region), the scale that defines the spatial resolution, the CRS (or coordinate reference system) to which your output image will be projected, the maximum pixels that you can adjust for larger exports, and the folder where you would like to save your output.

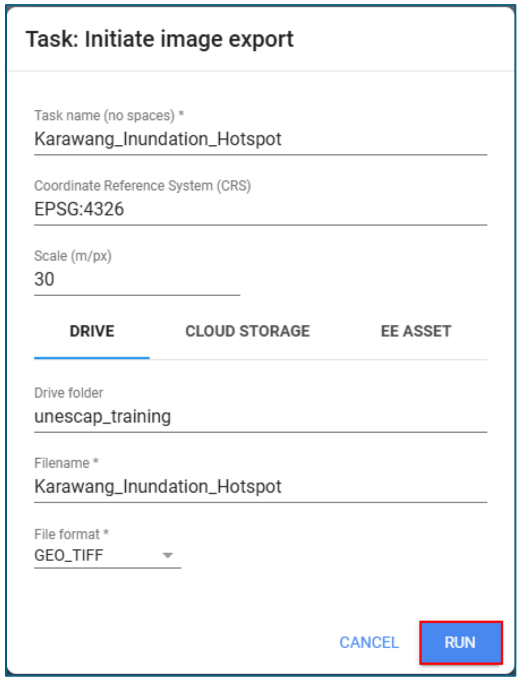
- **Click** *Run* to run your code and initiate the export task.



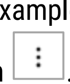
- Once the code is finished running, on the right panel, under the *Tasks* tab, you will find the task to export the inundation hotspot dataset, and the name of this task is the description you indicated in the code: *Karawang_Inundation_Hotspot*. **Click** *Run*.

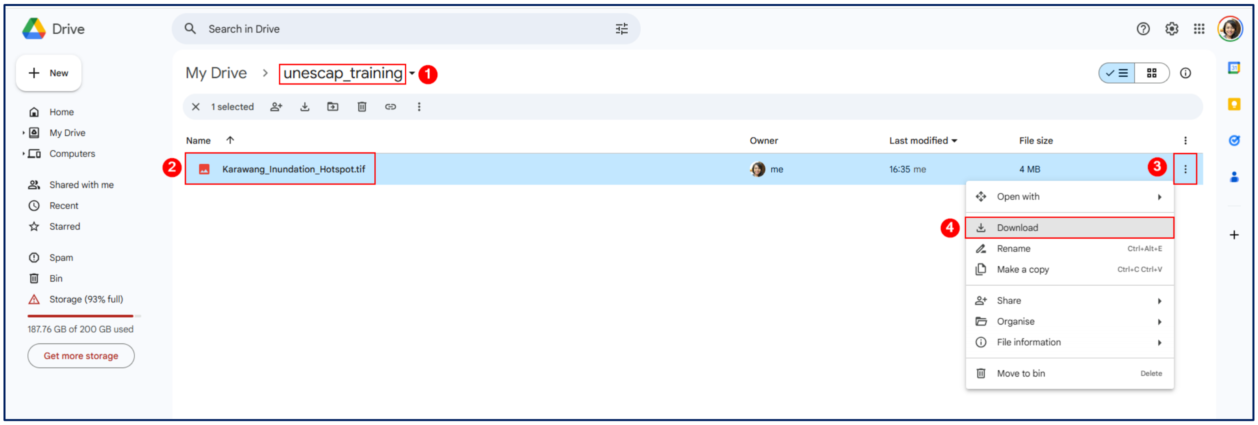


- The *Task: Initiate image export* dialog box will appear. You will notice that the parameters you set in the previous code are reflected here. If you did not define the CRS, it will default to the native coordinate system of the dataset. The suggested filename is based on the task name you assigned to this export. By default, images are exported as GEO_TIFF. You can still modify the task name, filename, scale, Google Drive folder to save your file, and the file format (GEO_TIFF or TF_RECORD_IMAGE). We will keep the parameters we set in the code for our export.
  Click *Run*.



  It will take some time for the export to finish. Once the processing is done, you can go to the Google Drive folder to download the data.

- **Go** to your Google Drive by following this link: https://drive.google.com/drive/my-drive
  Note that the Gmail account you are using for GEE is the same account for Google Drive used to save the export.
- **Navigate** to the folder where you saved the export. In this example, **go** to *unescap-training*
- **Select** on the inundation hotspot file, then **click** on this icon .
- **Click** *Download*.
  Again, based on your download your download settings, may either be saved directly to your Downloads folder or to the folder you select before you proceed to download.



Checkpoint

Downloading the land cover map from the ESA WorldCover 2021 website is only available for 60 x 60-degree tiles, which is ~9.2 GB of data for the area including our study site! Too big, right? Let's take advantage of the availability of this dataset in GEE and the platform's capabilities to subset and resample the exported data.

Now, based on the previous steps on how to export data from GEE, we can also download the corresponding land cover data for Karawang, Indonesia using the inundation hotspot code. Use the following questions to download the right layer from the code.

- *Can you find the variable corresponding to the land cover dataset? What is the variable name?*

- *Which parameters in the export code should you change? (HINT: these parameters refer to the source image and the output task)*

- *For the succeeding analysis, what should be the spatial resolution of your land cover data?*

You have now successfully downloaded the inundation hotspot and land cover datasets for Karawang, Indonesia. We are ready to move to the analysis in QGIS.

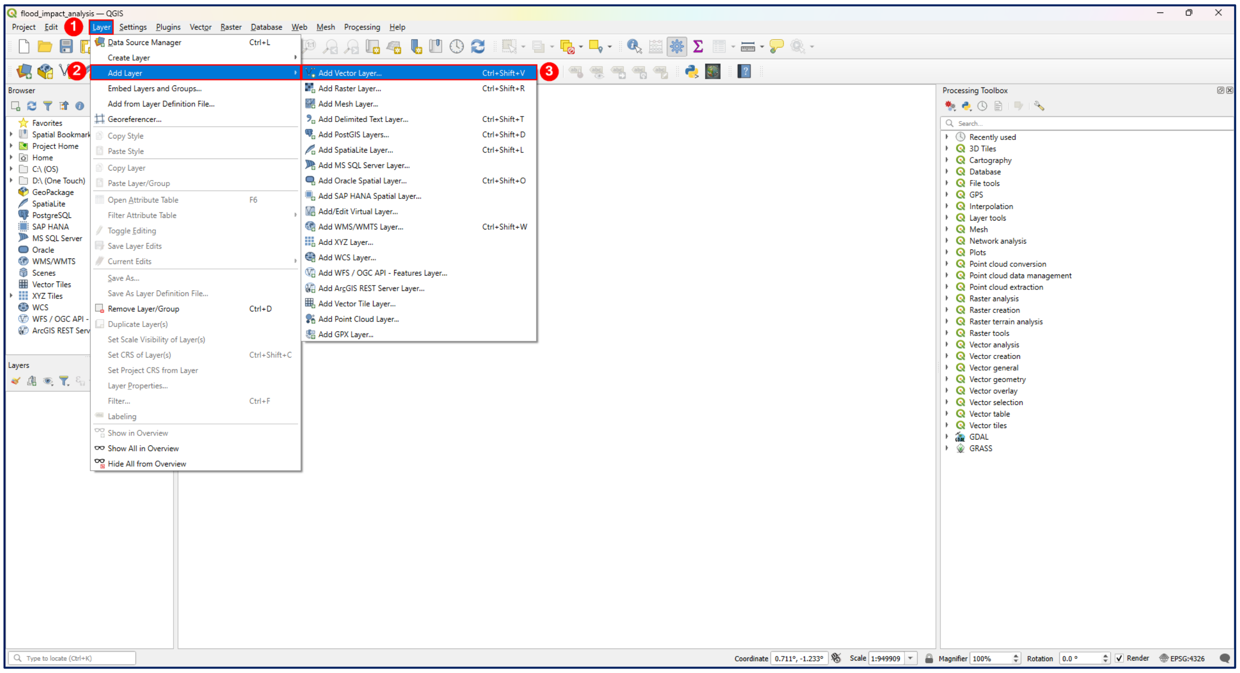
Part III. Mapping Flood Impacts in QGIS

Opening QGIS Desktop

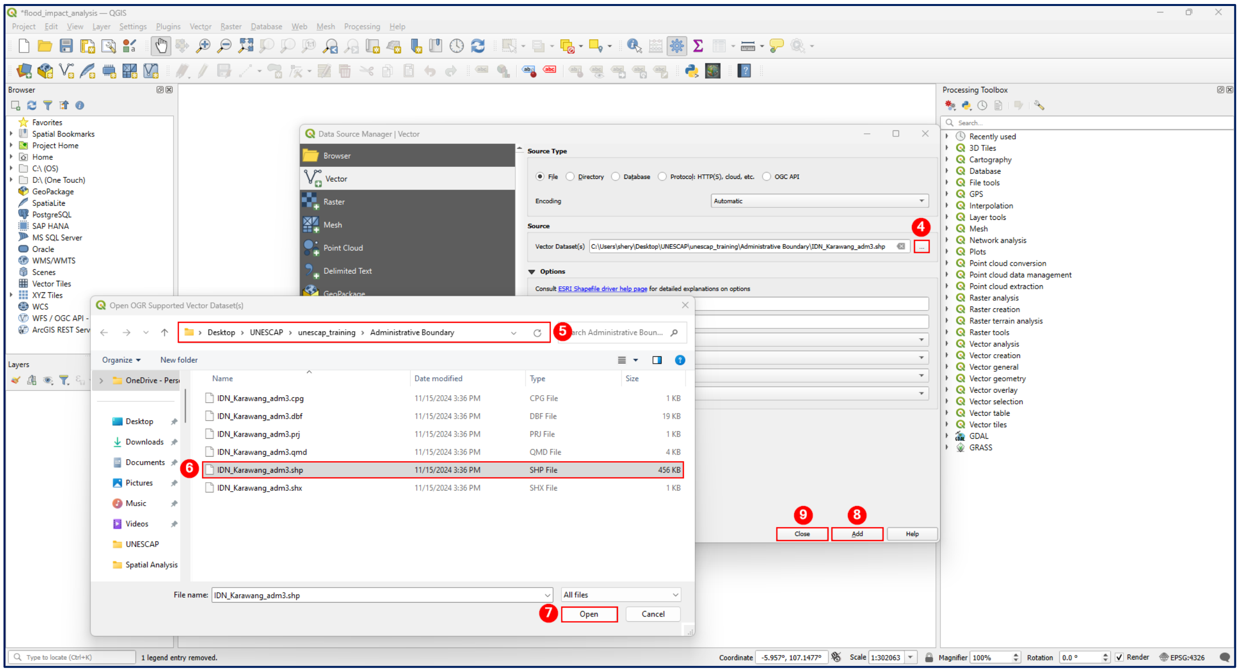
- If you need to install QGIS Desktop, you can **download** the latest version here: https://qgis.org/download/
- **Double-click** on the installation file and follow the prompts to **install** QGIS Desktop.
- **Open** your QGIS Desktop by **double-clicking** on the application icon QGIS Desktop 3.34.12. Alternatively, you can **select** QGIS Desktop from the *Start menu*.

<u>**Adding and Visualizing Vector Data**</u>

- From the main menu, **select** *Layer*.
- **Select** *Add Layer*, then **click** on *Add Vector Layer…*

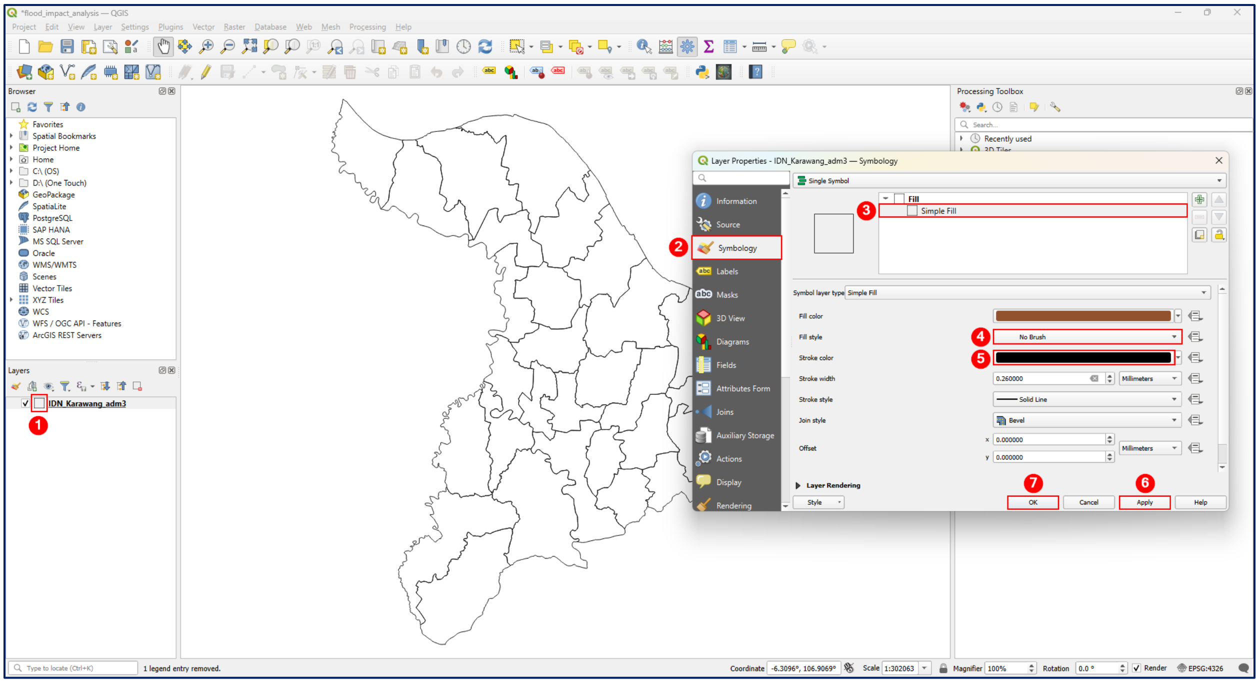


- The *Add Vector Layer* dialog box will appear. **Click** on the ellipsis  to **navigate** to the folder where you saved the data inputs for this flood impact analysis.
- **Select** the folder where you saved the administrative boundary for Karawang, Indonesia: *IDN_Karawang_adm3*
- **Select** the file with the *.shp* extension.
- **Click** *Open*.
- **Click** *Add*, then *Close*.



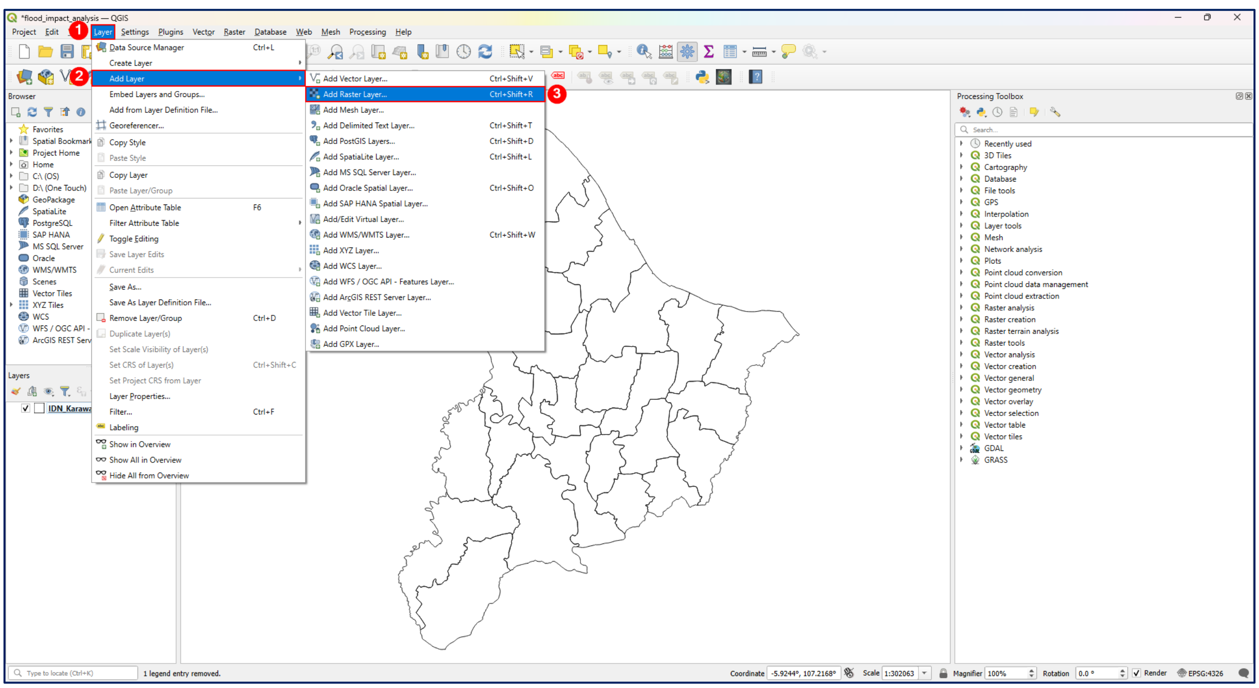
The administrative boundary for Karawang, Indonesia will now appear in your map display. Let's quickly do a visualization of this layer.

- **Double-click** on the polygon symbol beside the layer name, *IDN_Karawang_adm3*
- This will direct you to the *Layer Properties* dialog box, which will allow you to change the symbology for this layer.
- Under *Single Symbol* > *Fill*, **select** *Simple Fill*.
- To properly overlay the administrative boundary with the other data inputs, let's **set the symbology** to *No Brush* for the *Fill Style*.
- You can **choose** any *Stroke color* that you prefer, as long as it will allow you to visualize the boundaries with the other layers that we are going to add in this QGIS project. In this example, we are choosing a black stroke color.
- **Click** *Apply*, then *OK*.



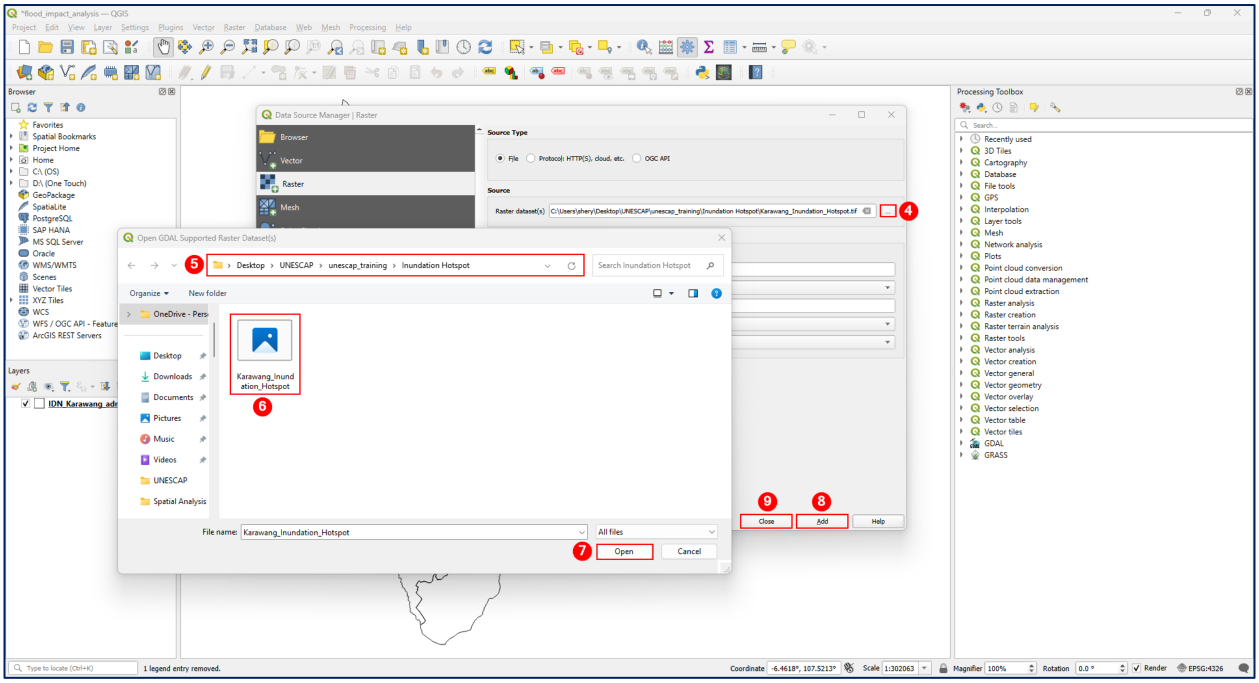
Adding and Visualizing Raster Data

- From the main menu, **select** *Layer*.
- **Select** *Add Layer*, then **click** on *Add Raster Layer…*



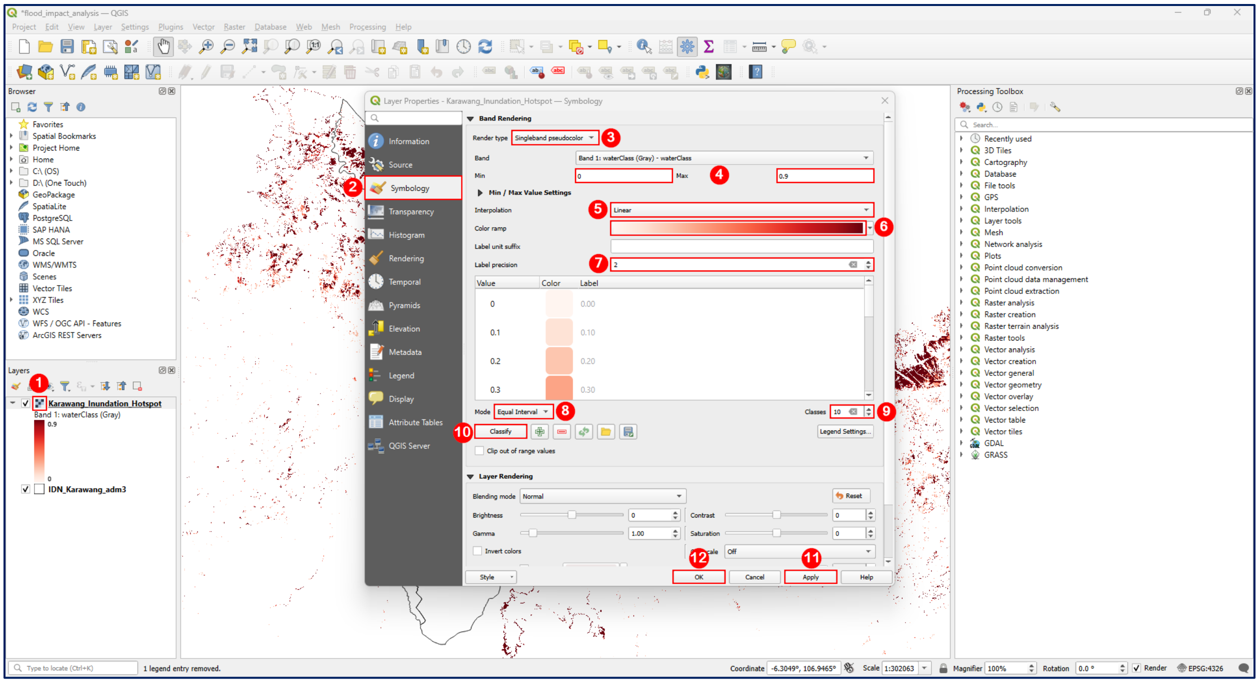
- The *Add Raster Layer* dialog box will appear. **Click** on the ellipsis … to **navigate** to the folder where you saved the data inputs for this flood impact analysis.
- **Select** the folder where you saved the inundation hotspot data for Karawang, Indonesia.
- **Select** the file: *Karawang_Inundation_Hotspot.tif*

- **Click** *Open*.
- **Click** *Add*, then *Close*.



By default, raster data such as this inundation hotspot map will be visualized in greyscale in the map display. Let's quickly change the minimum and maximum values for this raster and mimic the color ramp used in SATGPT.

- **Double-click** on the raster symbol beside the layer name, *Karawang_Inundation_Hotspot*
- This will direct you to the *Layer Properties* dialog box, which will allow you to change the symbology for this layer.
- Under *Render typel*, **select** *Singleband Pseudocolor*.
- **Change** the *minimum* and *maximum* values to 0 and 0.9, respectively.
- **Select** *Linear* as the *Interpolation* method.
- For the *Color ramp*, **select** *Reds*.
- **Set** the *Label precision* to 2.
- **Choose** *Equal interval* for the *Mode*.
- Then, **create** 10 classes.
- **Click** *Classify* to classify the raster.
- **Click** *Apply*, then *OK*.



- Following the methods in adding a raster layer in QGIS, **add** the land cover data to your project. Unlike the inundation hotspot layer, the land cover data will automatically be visualized with the default colors assigned by the data provider. See the following table for the land cover classes and their corresponding raster values.

Value	Color	Description
10	#006400	Tree cover
20	#ffbb22	Shrubland
30	#ffff4c	Grassland
40	#f096ff	Cropland
50	#fa0000	Built-up
60	#b4b4b4	Bare / sparse vegetation
70	#f0f0f0	Snow and ice
80	#0064c8	Permanent water bodies
90	#0096a0	Herbaceous wetland
95	#00cf75	Mangroves
100	#fae6a0	Moss and lichen

Checkpoint

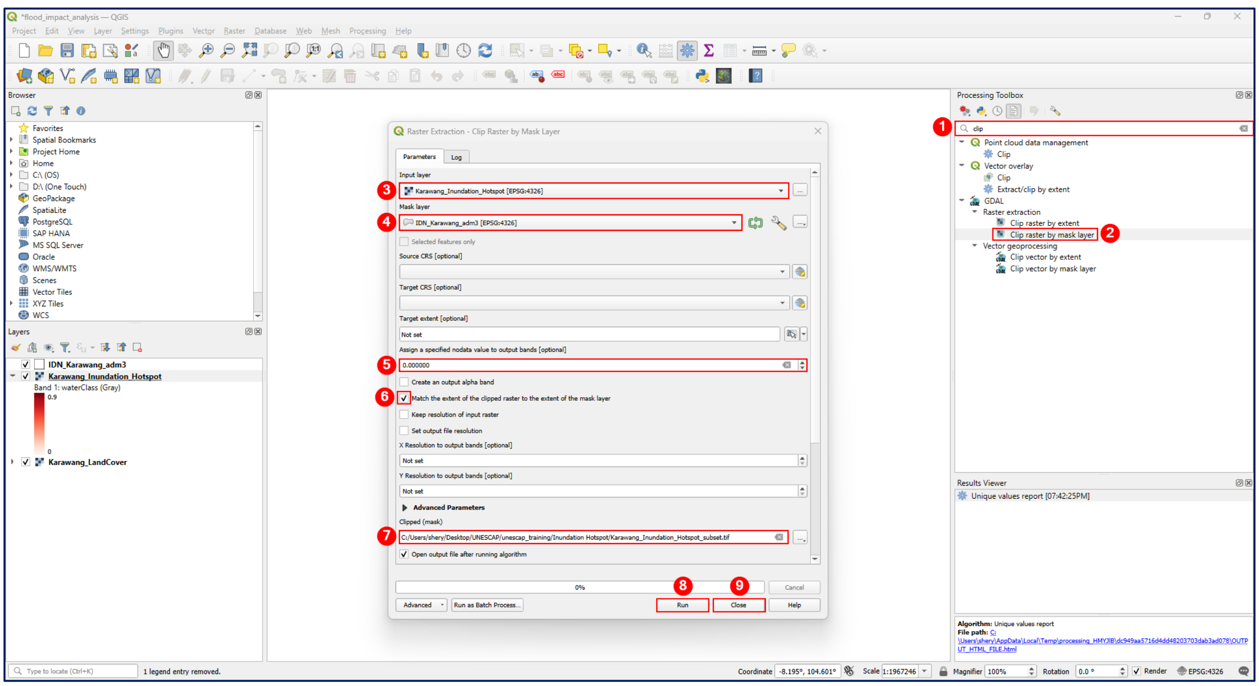
Open the *Layer Properties* for the land cover data and delete the values that do not correspond to any land cover classes. *(HINT: You need to click on a symbol the subtracts)*

Subset Raster Layers

Our raster layers extend to other nearby areas in Karawang. We also need to reproject our layers to a projected coordinate system prior to our analysis. For Karawang, Indonesia we will use the assigned Universal Transverse Mercator projection for this area, which is UTM Zone 48S.

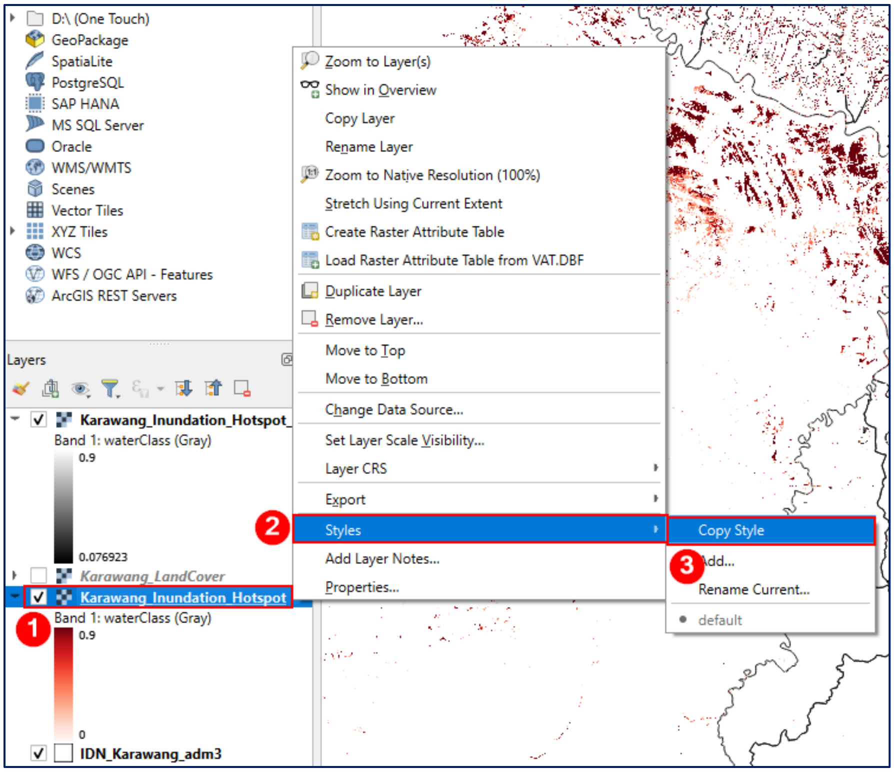
Let's quickly subset and reproject our data to Karawang, Indonesia only.

- From the *Processing Toolbox*, **search** for 'clip'.
- Under *GDAL*, **select** *Clip raster by mask layer*. The *Clip Raster by Mask Layer* dialog box will appear.
- For the *Input layer*, **select** the *Karawang_Inundation_Hotspot* raster layer.
- **Select** *IDN_Karawang_adm3* as the *Mask layer*.
- **Assign** the *no data value* to *0*.
- **Check** the *Match the extent of the clipped raster to the extent of the mask layer* to subset the raster layer to the administrative boundary of Karawang.
- **Save** your file, **append** 'subset' to your new filename to distinguish this from the original layer.
- **Click** *Run*, then *Close*.



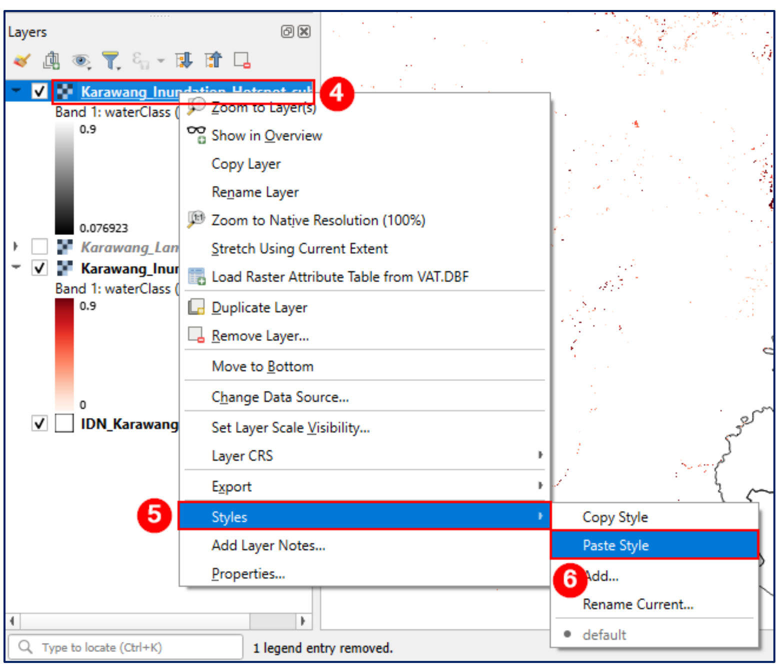
By default, raster data will again be visualized in greyscale in the map display. Let's replicate the previous symbology by doing the following:

- **Right-click** on the original layer, *Karawang_Inundation_Hotspot* to open the context menu.
- **Hover** to *Styles*, then **click** on *Copy Style*.



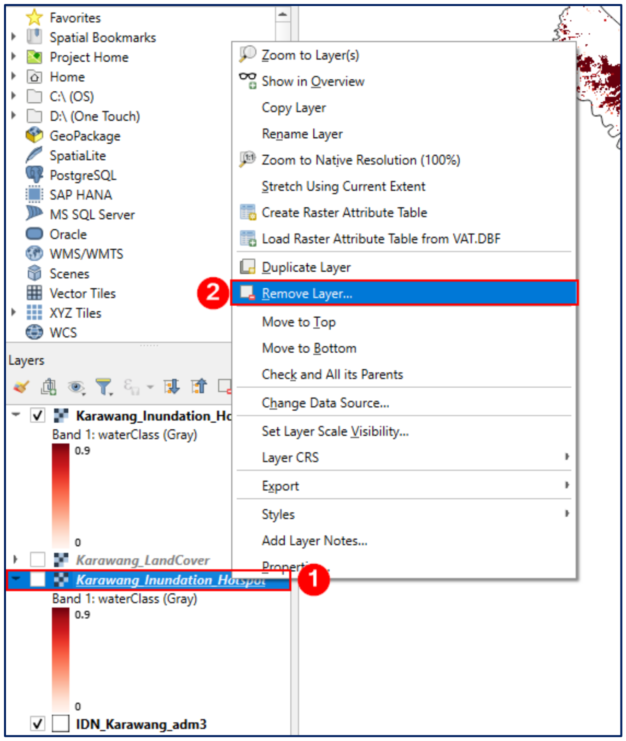
The style of this layer is now copied to your clipboard.

- **Right-click** on the new layer, *Karawang_Inundation_Hotspot_subset* to open the context menu.
- **Hover** to *Styles*, then **click** on *Paste Style*.



The subset of your inundation hotspot data will now have the same symbology as the original data. Let's cleanup a bit:

- **Right-click** on the new layer, *Karawang_Inundation_Hotspot_subset* to open the context menu.
- **Click** *Remove Layer…* to avoid confusion



Checkpoint

Follow the same steps to subset the land cover layer. Do not forget to clean up to avoid confusion.

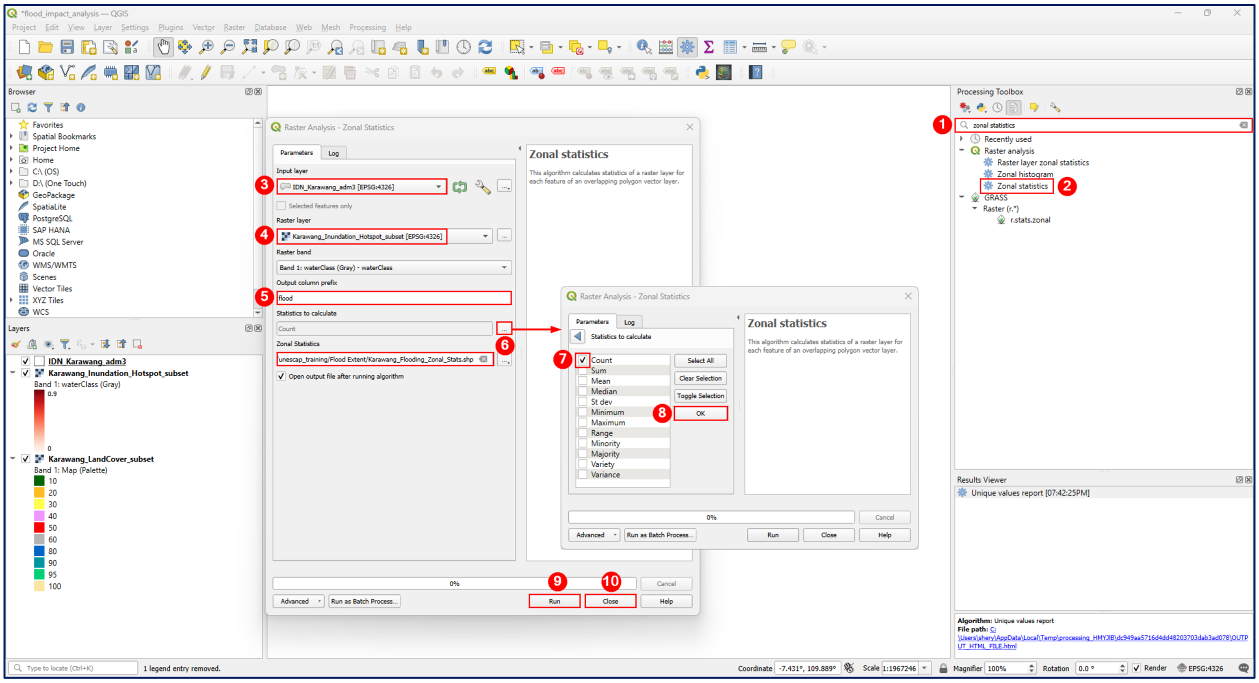
Calculating the Flooded Area by Administrative Units

- From the *Processing Toolbox*, **search** for 'zonal statistics'.
- **Click** on *Zonal statistics* under *Raster analysis*.

With the Zonal Statistics tool in QGIS, you can analyze the results of a thematic classification. It allows you to calculate several values of the pixels of a raster layer with the help of a polygonal vector layer. Choosing a color band, the plugin generates output columns in the vector layer with a user-defined prefix and calculates for each polygon, statistics on pixels that are within. Some of the available statistics are *Count* (to count the number of pixels), *Sum* (to sum the pixel values), and *Mean* (to get the mean of pixel values).
(Source: https://docs.qgis.org/2.18/en/docs/user_manual/plugins/plugins_zonal_statistics.html)

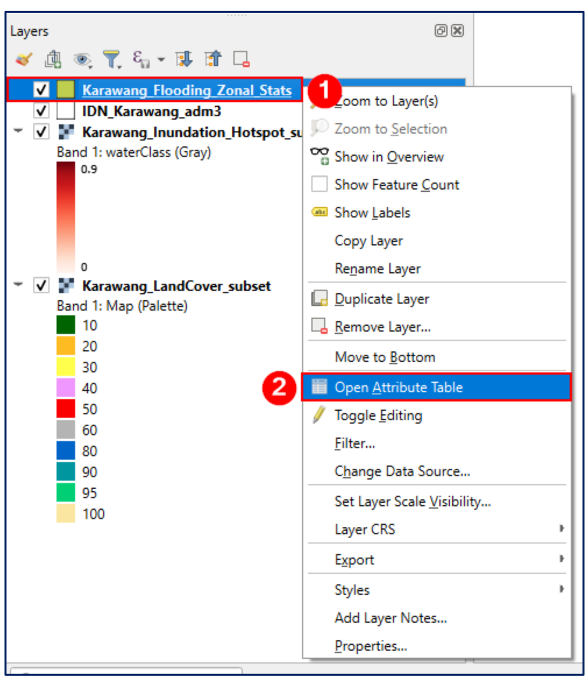
- For the *Input layer*, **use** the administrative boundary for Karawang, *IDN_Karawang_adm3*.
- **Set** *Raster layer input* to the *Karawang_Inundation_Hotspot_subset*.
- For the *Output column prefix*, **enter** 'flood'.
- **Click** on the ellipsis [...] under *Statistics to calculate…*
- **Uncheck** all other parameters and **make sure to check** only the *Count* parameter.
- **Click** *OK* to close the parameter dialog box.

- **Save** your zonal statistics file. For this example, we are using *Karawang_Flooding_Zonal_Stats.shp*
- **Click** *OK* then *Close*.



We are only interested in the number of pixels that fall under the subdistricts for Karawang, Indonesia to estimate the area affected by flooding.

- The outputs of the zonal statistics calculation will be added to your Layers panel. **Right-click** to open the context menu, then **select** *Open Attribute Table.*.



The attribute table should look similar to the following figure. We are only interested in the *flood_count* column.

Karawang_Flooding_Zonal_Stats — Features Total: 30, Filtered: 30, Selected: 0

	ADM3_EN	ADM3_PCODE	ADM3_REF	ADM3ALT1EN	ADM3ALT2EN	ADM2_EN	ADM2_PCODE	ADM1_EN	ADM1_PCODE	ADM0_EN	ADM0_PCODE	flood_coun
1	Tegalwaru	ID3215011	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	45.0000000000...
2	Kotabaru	ID3215072	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	106.000000000...
3	Pangkalan	ID3215010	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	109.000000000...
4	Karawang Timur	ID3215112	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	241.000000000...
5	Cikampek	ID3215050	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	266.000000000...
6	Purwasari	ID3215051	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	284.000000000...
7	Tirtamulya	ID3215060	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	405.000000000...
8	Telukjambe Tim...	ID3215031	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	560.000000000...
9	Telukjambe Barat	ID3215032	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	623.000000000...
10	Ciampel	ID3215020	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	826.000000000...
11	Klari	ID3215040	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	855.000000000...
12	Majalaya	ID3215111	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	893.000000000...
13	Jatisari	ID3215070	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	913.000000000...
14	Karawang Barat	ID3215113	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	2081.00000000...
15	Talagasari	ID3215100	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	3081.00000000...
16	Lemahabang	ID3215090	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	4810.00000000...
17	Banyusari	ID3215071	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	4888.00000000...
18	Rengasdengklok	ID3215150	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	5068.00000000...

Show All Features

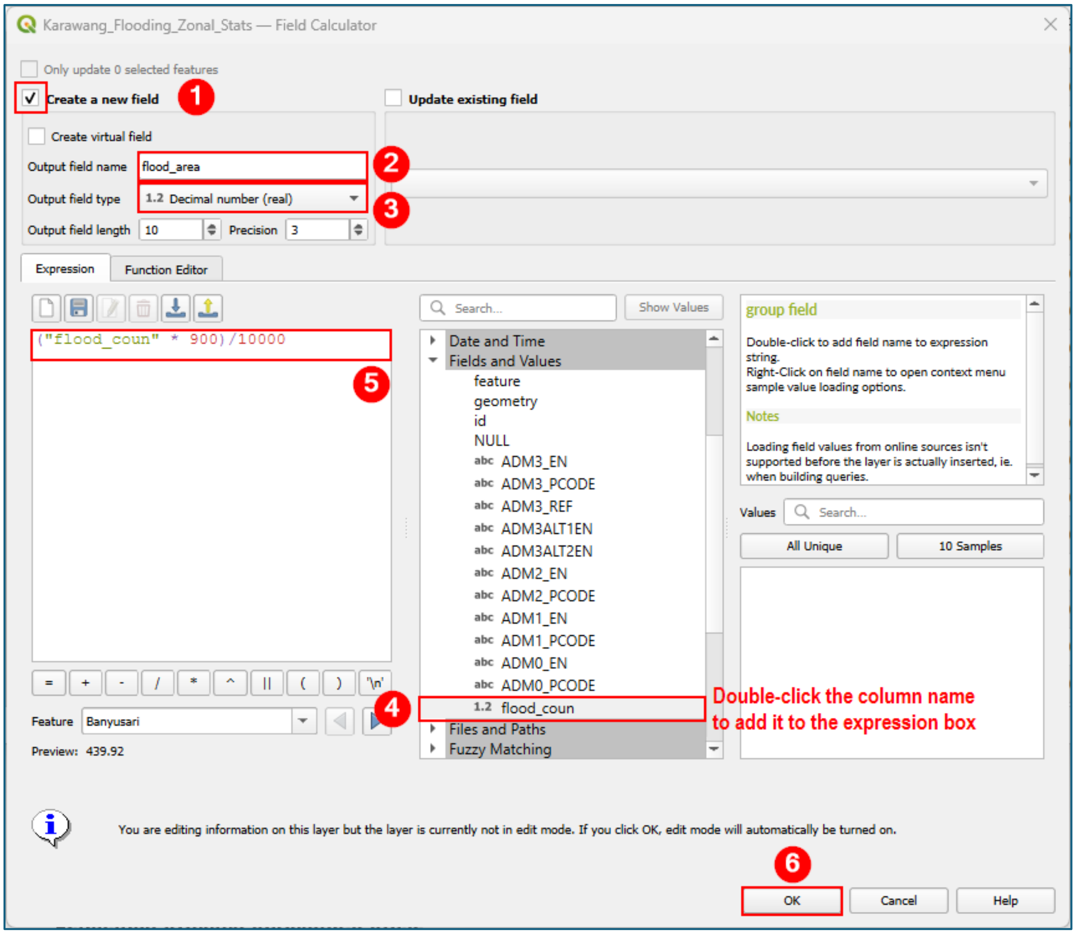
- From the *Attribute Table* window, **click** on the *Field Calculator* icon. Once you open the *Field Calculator* dialog box, this layer will automatically be in editing mode.
- **Check** *Create new field.*
- **Enter** *flood_area* for the *Output field name*.
- **Select** *Decimal number (real)* for the *Output field type.*
- **Select** the *flood count* column under *Fields and Arrays*.
- **Enter** the following expression in the expression box:

$$("flood\_count" * 900)/10000$$

The "flood_count" is the column name for the pixel count that was computed using the *Zonal Statistics* tool. In this example, the column name is *flood_coun*. You may need to replace this with the correct column name in case you used a different name from the example.

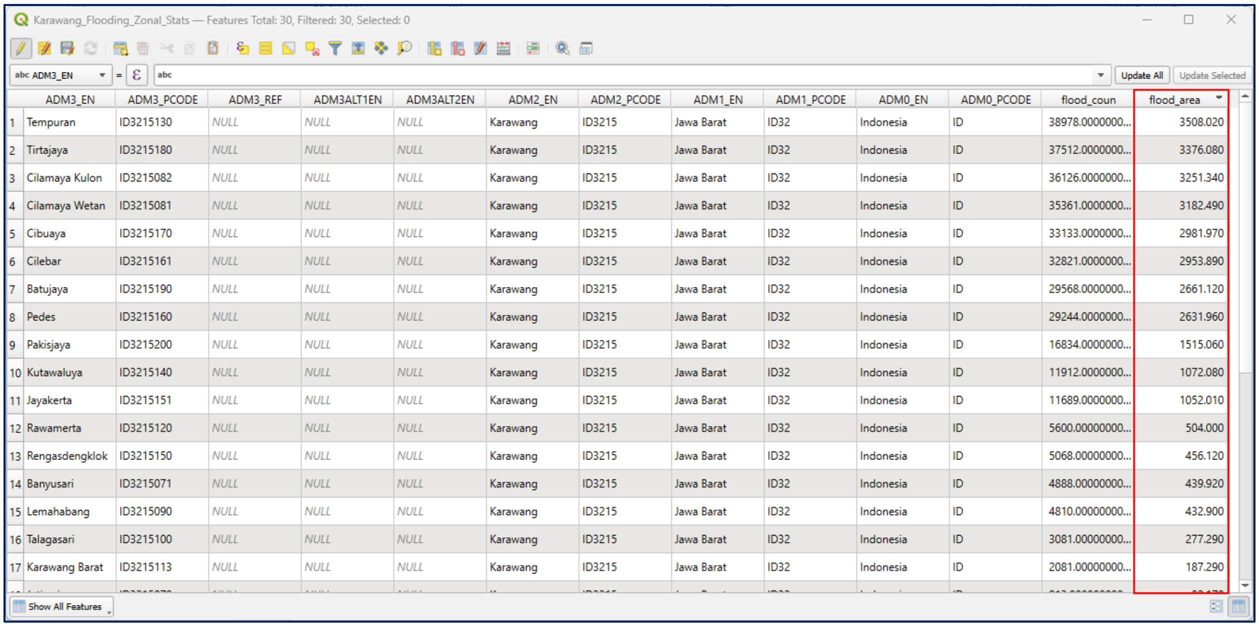
This expression calculates the area based on the pixel count. The spatial resolution of the inundation hotspot map is 30 meters, which means that one pixel is equivalent to an area of 30 x 30 meters (900 square meters). We then add the hectare conversion factor of 10,000 square meters per one hectare to get the values in hectares.

- **Click** *OK*.



A new column named *flood_area* will be added to your attribute table.

The Field Calculator in the attribute table allows you to perform calculations based on existing attribute values or defined functions, for instance, to calculate length or area of geometry features. The results can be used to update an existing field or written to a new field (that can be a virtual one).
The Field Calculator is available on any layer that supports editing. When you click on the field calculator icon the dialog opens. If the layer is not in edit mode, a warning is displayed and using the field calculator will cause the layer to be put in edit mode before the calculation is made. Based on the Expression Builder dialog, the Field Calculator dialog offers a complete interface to define an expression and apply it to an existing or a newly created field.
(Source: https://docs.qgis.org/3.34/en/docs/user_manual/working_with_vector/attribute_table.html#using-the-field-calculator)



	ADM3_EN	ADM3_PCODE	ADM3_REF	ADM3ALT1EN	ADM3ALT2EN	ADM2_EN	ADM2_PCODE	ADM1_EN	ADM1_PCODE	ADM0_EN	ADM0_PCODE	flood_coun	flood_area
1	Tempuran	ID3215130	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	38978.0000000...	3508.020
2	Tirtajaya	ID3215180	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	37512.0000000...	3376.080
3	Cilamaya Kulon	ID3215082	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	36126.0000000...	3251.340
4	Cilamaya Wetan	ID3215081	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	35361.0000000...	3182.490
5	Cibuaya	ID3215170	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	33133.0000000...	2981.970
6	Cilebar	ID3215161	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	32821.0000000...	2953.890
7	Batujaya	ID3215190	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	29568.0000000...	2661.120
8	Pedes	ID3215160	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	29244.0000000...	2631.960
9	Pakisjaya	ID3215200	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	16834.0000000...	1515.060
10	Kutawaluya	ID3215140	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	11912.0000000...	1072.080
11	Jayakerta	ID3215151	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	11689.0000000...	1052.010
12	Rawamerta	ID3215120	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	5600.00000000...	504.000
13	Rengasdengklok	ID3215150	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	5068.00000000...	456.120
14	Banyusari	ID3215071	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	4888.00000000...	439.920
15	Lemahabang	ID3215090	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	4810.00000000...	432.900
16	Talagasari	ID3215100	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	3081.00000000...	277.290
17	Karawang Barat	ID3215113	NULL	NULL	NULL	Karawang	ID3215	Jawa Barat	ID32	Indonesia	ID	2081.00000000...	187.290

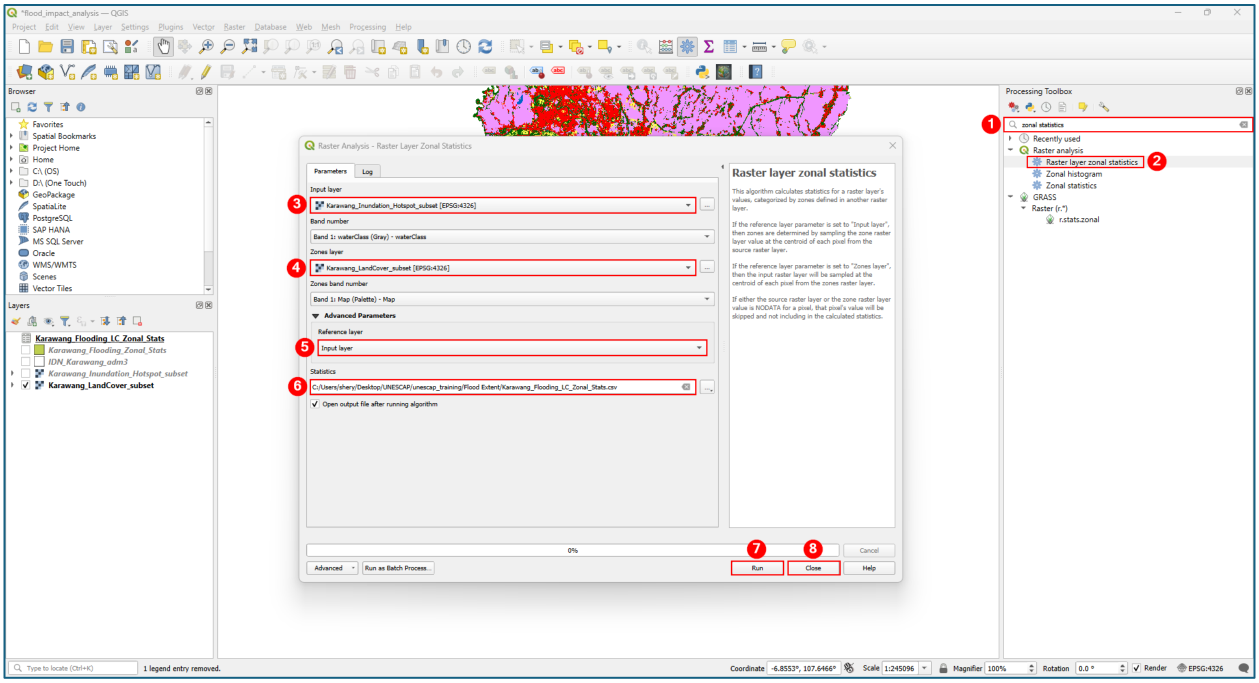
- Make sure to save your area calculation by **clicking** on the *Save Edits* icon in this attribute table.

You have now calculated the area affected by flooding in the subdistricts of Karawang, Indonesia.

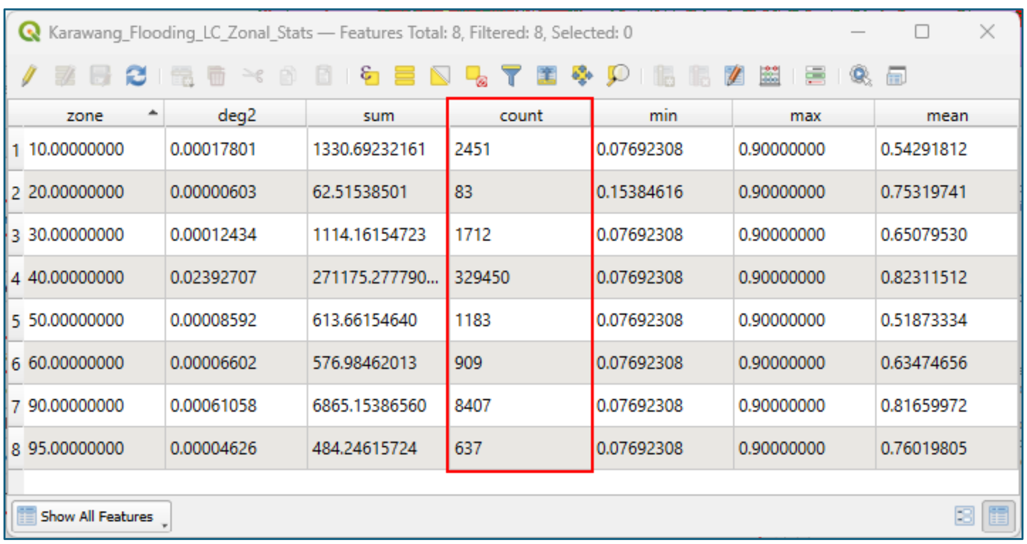
Estimating and Visualizing Agricultural Areas Affected by Flooding

Now, let's try to estimate the area affected by flooding by land cover type.

- From the *Processing Toolbox*, **search** for 'zonal statistics'.
- **Click** on *Raster layer zonal statistics* under *Raster analysis*.
- For the *Input layer*, **use** the administrative boundary for Karawang, *Karawang_Inundation_Hotspot_subset*.
- **Set** *Zones layer* to the *Karawang_LandCover_subset*.
- **Keep** the *Reference layer* to the *Input layer*.
- **Save** your zonal statistics file as CSV. For this example, we are using *Karawang_Flooding_LC_Zonal_Stats.csv*
- **Click** *OK* then *Close*.



- The CSV file will be added to the *Layers* panel. You can **open** it like an attribute table. Your attribute table should be similar to the following figure. Take note of the *count* column.



Karawang_Flooding_LC_Zonal_Stats — Features Total: 8, Filtered: 8, Selected: 0

	zone	deg2	sum	count	min	max	mean
1	10.00000000	0.00017801	1330.69232161	2451	0.07692308	0.90000000	0.54291812
2	20.00000000	0.00000603	62.51538501	83	0.15384616	0.90000000	0.75319741
3	30.00000000	0.00012434	1114.16154723	1712	0.07692308	0.90000000	0.65079530
4	40.00000000	0.02392707	271175.277790...	329450	0.07692308	0.90000000	0.82311512
5	50.00000000	0.00008592	613.66154640	1183	0.07692308	0.90000000	0.51873334
6	60.00000000	0.00006602	576.98462013	909	0.07692308	0.90000000	0.63474656
7	90.00000000	0.00061058	6865.15386560	8407	0.07692308	0.90000000	0.81659972
8	95.00000000	0.00004626	484.24615724	637	0.07692308	0.90000000	0.76019805

Show All Features

Checkpoint

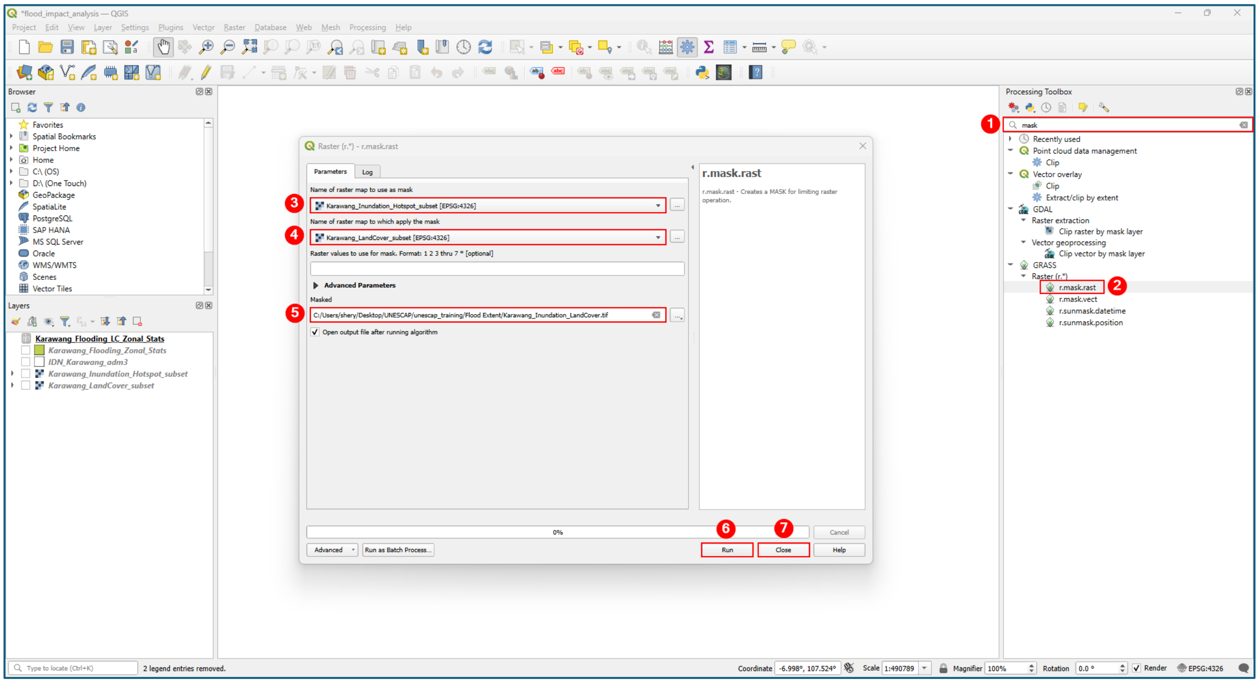
Follow the same steps to calculate the agricultural area affected by flooding using the *Field Calculator*. Here are some guide questions:

- *What is the corresponding map class in the land cover layer for agricultural area?*

- *Can you name another land cover class that is severely affected by flooding?*

Let's try to visualize the agricultural areas that are affected by flooding.

- From the *Processing toolbox*, **search** for 'mask'.
- Under *GRASS > Raster (r.*)*, **select** the function *r.mask.rast*. A dialog box will appear.
- For the *Name of raster map to use as mask*, **use** the layer *Kawarang_Inundation_Hotspot_subset*.
- For the *Name of raster map to which apply the mask*, **use** the layer *Kawarang_LandCover_subset*.
- **Save** the file. In this example, save as *Karawang_Inundation_LandCover.tif*
- **Click** *Run* then *OK*.



Checkpoint

Visualize the new land cover map with the administrative boundary. Create a map layout indicating the subdistricts and land cover classes with the required map elements.

Analysis Guide and Questions

1. *Visualize the inundation hotspot map with the administrative boundary of Karawang, Indonesia. Add labels to your administrative boundary by using the Layer Properties dialog box and set the label to ADM3_EN. Compare this map to the area you calculated using the Zonal Statistics tool. Are there any significant differences?*

2. *What are the five most affected subdistricts?*

3. *What is the total agricultural area affected by flooding in Karawang, Indonesia?*

____

4. *Can you list the five most affected subdistricts by agricultural area?*

____

Part IV. Practice Session

Each group will follow these steps to conduct your assigned flood impact analysis:

- **Select an Area of Interest (AOI)**: Choose a region for analysis that is known for flooding or is relevant to your work.
- **Generate Flood Hotspot Maps**: Use SatGPT to create flood hotspot maps within your AOI, identifying areas with high flood vulnerability, download the data to be used in QGIS or GEE.
- **Analyze Impacts in Google Earth Engine (GEE) or QGIS**: Once you have your flood hotspot map, use QGIS or GEE to complete the specific impact analysis assigned to your group.

Here is a list of data and their corresponding sources for your reference. Note that the use of some datasets is not limited to their category. Some datasets may also be used for multiple assessments such as infrastructure and economic loss, or infrastructure and population.

Dataset and Description	Spatial Resolution	Source
Land cover		
Copernicus Global Land Cover Provides at global level spatial information on different types (classes) of physical coverage of the Earth's surface, e.g. forests, grasslands, croplands, lakes, wetlands for the 2019 base year. The data are updated annually and are available for the 2015-2019 years.	100 meters	https://land.copernicus.eu/en/products/global-dynamic-land-cover/copernicus-global-land-service-land-cover-100m-collection-3-epoch-2019-globe
MODIS Land Cover Type (MCD12Q1) The Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer (MODIS) Land Cover Type (MCD12Q1) Version 6.1 data product provides global land cover types at yearly intervals (2001-2022). The MCD12Q1 Version 6.1 data product is derived using supervised classifications of MODIS Terra and Aqua reflectance data. Land cover types are derived from the International Geosphere-Biosphere Programme (IGBP), University of Maryland (UMD), Leaf Area Index (LAI), BIOME-Biogeochemical Cycles (BGC), and Plant Functional Types (PFT) classification schemes.	500 meters	https://lpdaac.usgs.gov/products/mcd12q1v061/#tools
Population		
WorldPop WorldPop produces estimates of populations with age/sex breakdowns for each 100m x 100m grid square on the planet. These function as default, open access datasets for UN agencies planning humanitarian and development interventions, and help governments fill census gaps.	100 meters	https://www.worldpop.org/datacatalog/

Gridded Population of the World The Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11 consists of estimates of human population density (number of persons per square kilometer) based on counts consistent with national censuses and population registers, for the years 2000, 2005, 2010, 2015, and 2020.A proportional allocation gridding algorithm, utilizing approximately 13.5 million national and sub-national administrative units, was used to assign population counts to 30 arc-second grid cells.	1 kilometer	https://sedac.ciesin.columbia.edu/gpw-v411-app/?shuid=gpw-v4-population-density-rev11
Global Human Settlement Layer (GHSL) The Global Human Settlement Layer (GHSL) project is supported by European Commission, Joint Research Centre and Directorate-General for Regional and Urban Policy. The GHSL produces new global spatial information, evidence-based analytics, and knowledge describing the human presence in the planet. The collection consists of three main types of products: built-up (GHS-BUILT), population (GHS-POP) grids, and city model (GHS-SMOD). The datasets are published as raster files together with pyramids (i.e., TIF and OVR files).	10 meters	https://human-settlement.emergency.copernicus.eu/download.php
Infrastructure		
OpenStreetMap OpenStreetMap relies mostly on data collected by project members using their GPS devices and entered in the central database with specialized editors. For some areas, third party data has been imported. The focus is mainly on transport infrastructure (streets, paths, railways, rivers), but OpenStreetMap also collects a multitude of points of Interest, buildings, natural features and land use information, as well as coastlines and administrative boundaries.	-	https://www.geofabrik.de/data/download.html
Global Roads Inventory Project (GRIP) The GRIP dataset is mainly aimed at providing a roads dataset that is easily usable for scientific global environmental and biodiversity modelling projects. GRIP4 is based on many different sources (including OpenStreetMap) and to the best of our ability we have verified their public availability, as a criterion in our research. The UNSDI-Transportation data model was applied for harmonization of the individual source datasets.	8 kilometers	https://www.globio.info/download-grip-dataset
Economic Data		
VIIRS Nighttime Lights The Suomi National Polar-orbiting Partnership (SNPP) Visible Infrared Imaging Radiometer Suite (VIIRS) supports a Day-Night Band (DNB) sensor that provides global daily measurements of nocturnal visible and near-infrared (NIR) light that are suitable for Earth system science and applications. The VIIRS DNB's ultra-sensitivity in lowlight conditions enable us to generate a new set of science-quality nighttime products to better monitor both the magnitude and signature of nighttime phenomena, and anthropogenic sources of light emissions.	500 meters	https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/VNP46A2/

Below are your group-specific instructions and the recommended datasets for each task:

Group 1 - **Agriculture impact** - estimate affected crop areas by overlaying flood maps with land cover or crop-specific layers (**Land Cover**: Use **Copernicus Global Land Cover** (100m or 300m) or **MODIS Land Cover Type** (MCD12Q1) for general classifications and for **Crop-Specific Data** use.

Group 2 - **Population exposure** - calculate the population affected by overlaying flood extent maps with population density grids (**WorldPop** – high-resolution population data or **Gridded Population of the World (GPWv4)** – provides population counts and densities or **Global Human Settlement Layer (GHSL)** – offers additional data on urban and population extents.)

Group 3 - **Infrastructure exposure** - estimate the impact on roads, buildings, or other infrastructure by using vector layers for these features and assessing flood intersections (**OpenStreetMap (OSM)** – detailed data on roads, buildings, and other infrastructure, **Global Human Settlement Layer (GHSL)** – for urban areas, **Global Roads Inventory Project (GRIP)** – provides global road networks.)

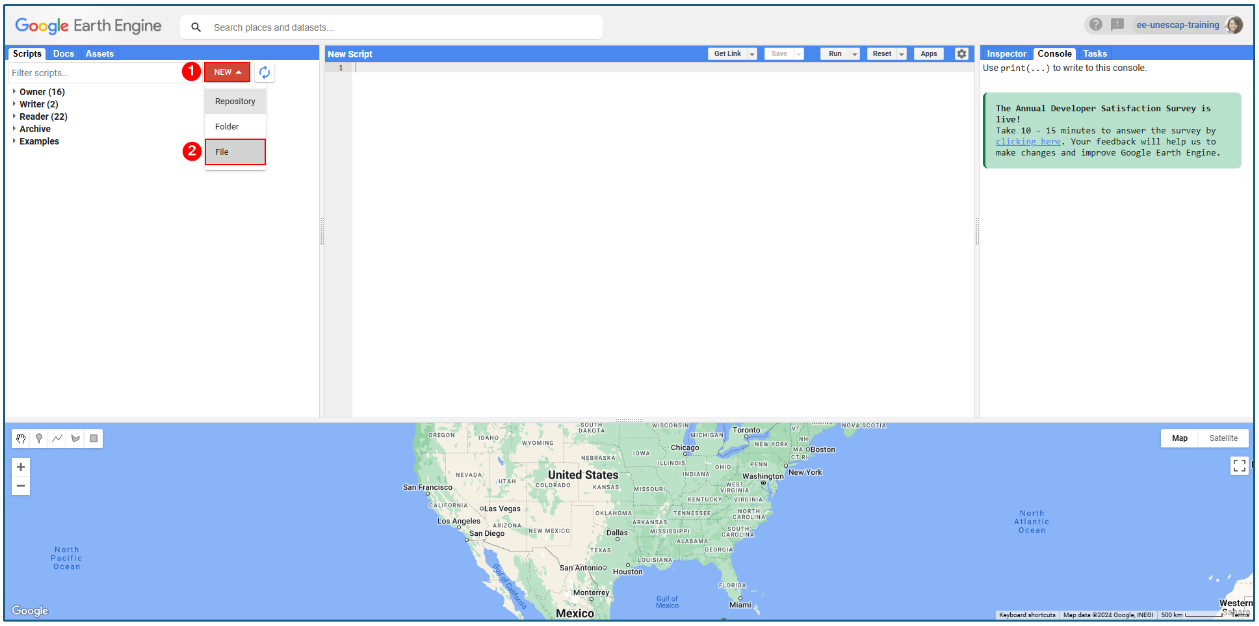
Group 4 - **Economic cost estimation** – use average economic values per land cover or infrastructure type, participants can estimate potential economic losses (**Land Cover**: Use **Copernicus Global Land Cover** (100m or 300m) or **MODIS Land Cover Type** (MCD12Q1) for general classifications, **Nighttime Lights Data (VIIRS)** – as an economic activity indicator, **Global Human Settlement Layer (GHSL)** – useful for evaluating urban areas, **Gridded Population of the World (GPWv4)** – can support population-based economic estimates.)

Challenge: Flood Impact Assessment in Google Earth Engine

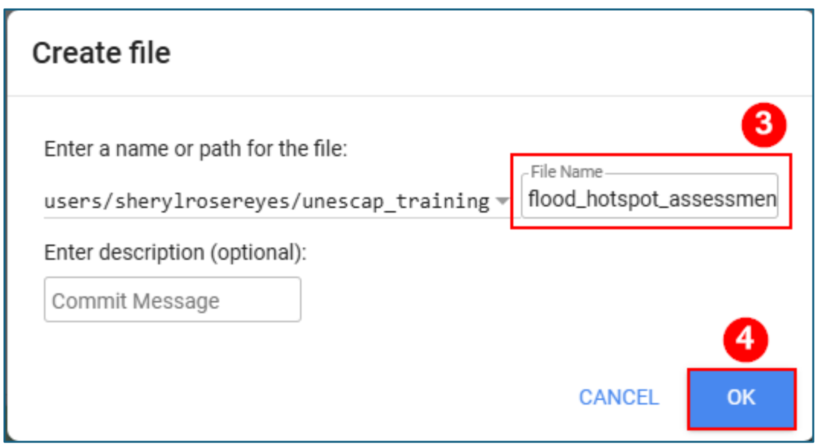
Let’s try a more advanced version of the flood impact assessment in Google Earth Engine.

Start a New Script

- **Go** to the *Scripts* tab, then **click** *New*.



- **Select** *File* under *New*. The *Create file* dialog box will appear.
- **Enter** the filename *flood_hotspot_assessment* for this new script.
- **Click** *OK* to create this script in your repository.



Modified SatGPT Code

Let's refine the inundation hotspot code from SatGPT. Here is the modified code:

```
/*******************************************
  A. INUNDATION HOTSPOT CODE FROM SATGPT
*******************************************/

var aoi = ee.FeatureCollection('projects/ee-unescap-training/assets/IDN_Karawang_adm3');
Map.centerObject(aoi, 10);

var year_count = 5;
```

```
var pink = ['#ffa9bb', '#ff9cac', '#ff8f9e', '#ff8190', '#ff7281', '#ff6171', '#ff4f61',
'#ff3b50', '#ff084a'];
var vizFFMpink =  {min: 0.1, max: 0.8, palette: pink};

var WaterESA2 = ee.ImageCollection("ESA/WorldCover/v200").first().eq(80).selfMask();
var WaterESA1 = ee.ImageCollection("ESA/WorldCover/v100").first().eq(80).selfMask();
var waterHistory = ee.ImageCollection("JRC/GSW1_4/YearlyHistory");

var PermanentWater = waterHistory.map(function(image) {return
image.select('waterClass').eq(3);}).sum();
var PermanentWaterFrequency = PermanentWater.divide(year_count);
var PermanentWaterFrequencyMap = PermanentWaterFrequency.gt(0).selfMask();
var PermanentWaterLayer =
ee.ImageCollection([WaterESA1.rename('waterClass'),WaterESA2.rename('waterClass'),
PermanentWaterFrequencyMap]).mosaic().clip(aoi);
var yearsWithWater = waterHistory.map(function(image) {return
image.select('waterClass').eq(2);}).sum();

var floodFrequency = yearsWithWater.divide(year_count);
var floodFrequencyMap =
floodFrequency.where(PermanentWaterLayer.eq(1),0).selfMask().clip(aoi);
var minMax = floodFrequencyMap.reduceRegion(ee.Reducer.minMax());
//print(minMax);

var floodFrequencyMap = floodFrequencyMap.where(floodFrequencyMap.gt(0.9),0.90);

Map.addLayer(PermanentWaterLayer, {min: 0, max: 1, opacity: 1.0, palette:
['white','blue','#0d6ee9']}, 'Permanent Water');
Map.addLayer(floodFrequencyMap, vizFFMpink, 'Flood Frequency Map');
```

In this modified code, the area of interest has been modified to the administrative boundary of Kawarang, Indonesia. Land cover and population layers were moved to the next sections of the code, while the soil texture and health facilities were removed since these layers will not be used in this example.

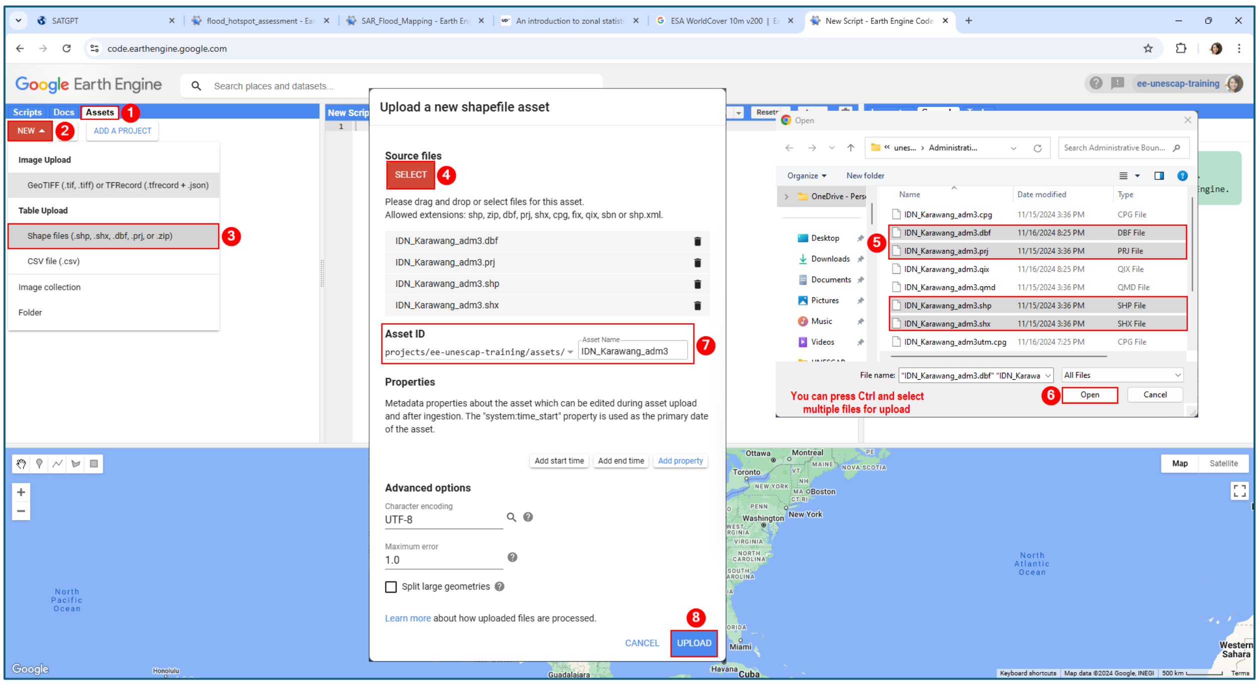
Upload an Asset

Let's upload the actual boundary of Kawarang, Indonesia. This will be an asset in GEE. You will replace the location of the `aoi` after you have uploaded the asset in your Google Cloud Project.

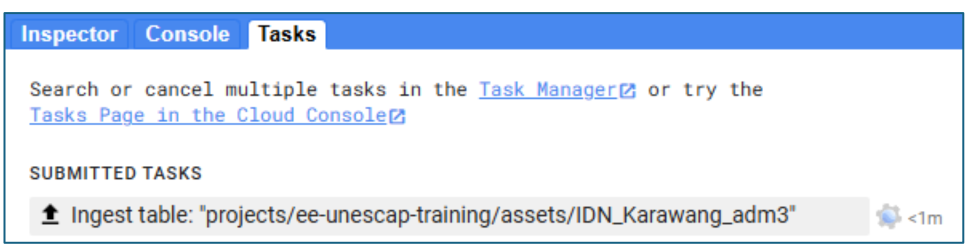
Earth Engine assets are images and tables as well as image collections and folders that you upload to GEE and can be imported to your codes. These assets are associated with a Google Cloud Project. GEE users before the cloud integration will find Legacy Assets in their Assets tab.
(Source: https://developers.google.com/earth-engine/cloud/assets).

- **Go** to the *Assets* tab, **click** *New*.
- Under *New*, **select** *Table Upload*. **Choose** *Shape files (.shp, shx, .dbf, .prj, or.zip)*.
- The *Upload a new shapefile asset* dialog box will appear. Click on Select under Source files. Navigate to the folder where you saved the administrative boundary for Kawarang, Indonesia.

- **Select** *IDN_Karawang_adm3.dbf*, *IDN_Karawang_adm3.prj*, *IDN_Karawang_adm3.shp*, and *IDN_Karawang_adm3.shx*. All these files must be uploaded to properly access and import your shapefile in GEE.
- Under *Asset ID*, **make sure** that your file is saved under the cloud project you previously created. By default, GEE assigns an asset name to your file, typically copying its original filename. You can modify the asset name, but make sure that there are no spaces in your selected filename.
- **Click** *Upload*.

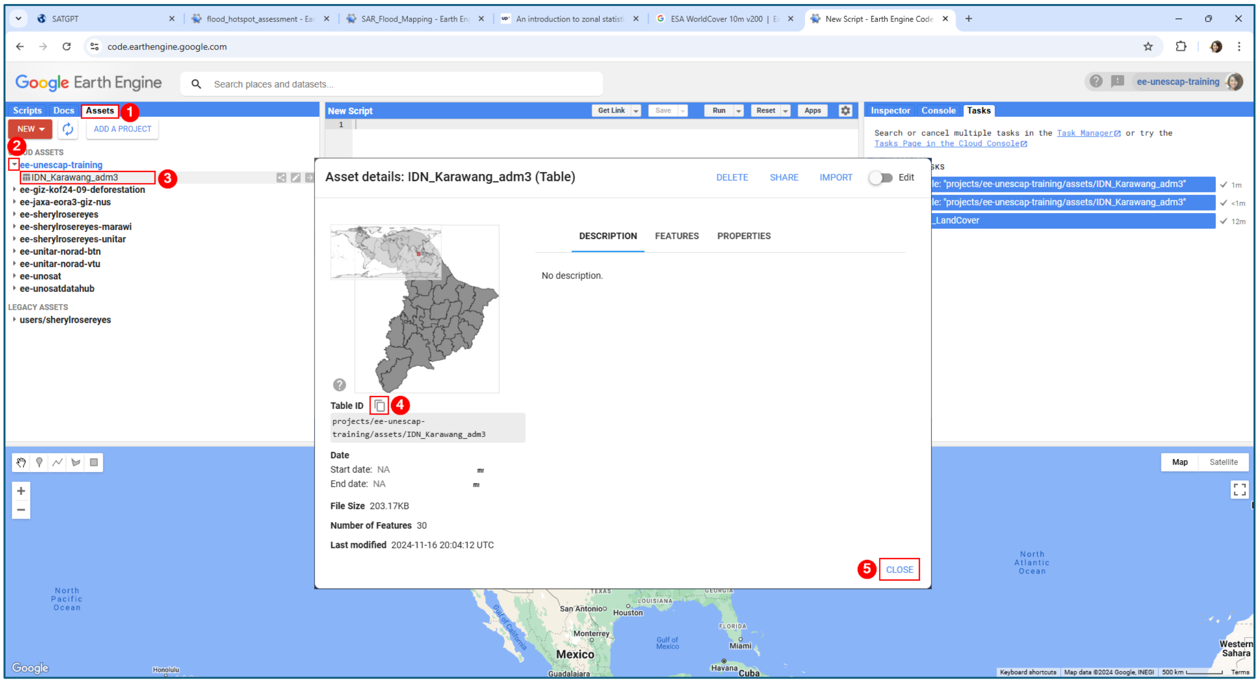


- **Navigate** to the *Task manager* tab. You will now see that your shapefile is being uploaded. Once the upload is finished, a check mark will replace the moving Earth Engine symbol on the right.



Groups of related features can be combined into a FeatureCollection, to enable additional operations on the entire set such as filtering, sorting and rendering. Besides just simple features (geometry + properties), feature collections can also contain other collections.
(Source: https://developers.google.com/earth-engine/guides/feature_collections)

- **Click** on the *Assets* tab and **select** the newly uploaded asset.
- **Copy** the file path of the asset for your area of interest by **clicking** on the copy icon.
- **Click** *Close*.



- **Paste** the location for the `aoi` in the Code Editor. **Replace** the file path highlighted in yellow to your asset's file path.

```
/*****************************************
  A. INUNDATION HOTSPOT CODE FROM SATGPT
******************************************/

var aoi = ee.FeatureCollection('projects/ee-unescap-training/assets/IDN_Karawang_adm3');
Map.centerObject(aoi, 10);
```

- **Click** *Save* and **click** *Run*.

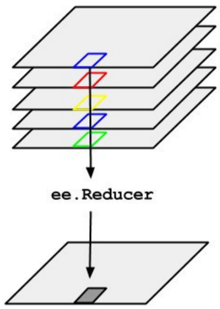
Flood Impact Assessment: Exposed Population

In this challenge, we will take advantage of the data products available in GEE. Instead of downloading, subsetting, and reprojecting each dataset on a step-by-step basis, these datasets will be processed using the clip and reproject functions in GEE. The raster layer that is being assessed, in this case, population count, is used to subset the inundation hotspot map to limit the data where there are inhabitants.

```
/**************************************
  B. FLOOD IMPACT ASSESSMENT
**************************************/
//---Exposed Population Density ---//

// Load GPWv411: Population Count (Gridded Population of the World Version 4.11)
```

```
// Resolution is at 927.67 meters (approximated to 1000 meters)
var population_count = ee.ImageCollection("CIESIN/GPWv411/GPW_Population_Count")
  .first()
  .clip(aoi);

// Calculate the amount of exposed population
// get GPW projection
var GPWprojection = population_count.projection();

// Reproject flood layer to GPW CRS
var floodFrequencyMap_reproj = floodFrequencyMap
    .reproject({
    crs: GPWprojection
  });

// Create a raster of the exposed population based on the resampled flood layer
var population_exposed = population_count
  .updateMask(floodFrequencyMap_reproj)
  .updateMask(population_count);

// Sum the pixel values of exposed population raster
var pop_stats = population_exposed.reduceRegion({
  reducer: ee.Reducer.sum(),
  geometry: aoi,
  scale: 1000,
  maxPixels:1e9
});

// Get the number of exposed people as integer
var number_pp_exposed = pop_stats.getNumber('population_count').round();
```

Similar to the QGIS flood impact analysis, layers that are used in the analysis should be in the same map projection and spatial resolution. The CRS of the data products are used as the reference for the inundation hotspot map. In GEE, data is aggregated using reducers.



Reducers are the way to aggregate data over time, space, bands, arrays and other data structures in GEE. The `ee.Reducer` class specifies how data is aggregated. The reducers in this class can specify a simple statistic to use for the aggregation (e.g. minimum, maximum, mean, median, standard deviation, etc.), or a more complex summary of the input data (e.g. histogram, linear regression, list). Reductions may occur over time (`imageCollection.reduce()`), space (`image.reduceRegion()`, `image.reduceNeighborhood()`), bands (`image.reduce()`), or the attribute space of a FeatureCollection `featureCollection.reduceColumns()` or FeatureCollection methods that start with `aggregate_`).

Reducers take an input dataset and produce a single output. When a single input reducer is applied to a multi-band image, Earth Engine automatically replicates the reducer and applies it separately to each band. As a result, the output image has the same number of bands as the input image; each band in the output is the reduction of pixels from the corresponding band in the input data. Some reducers take tuples of input datasets. These reducers will not be automatically replicated for each band.
(*Read more on https://developers.google.com/earth-engine/guides/reducers_intro*)

Flood Impact Assessment: Affected Agricultural Area

Again, similar to the QGIS flood impact analysis, we need to extract the map class from the ESA WorldCover 2021 data that is equivalent to the agricultural area. The land cover class is 'Cropland' and this is also used to subset the inundation hotspot map to limit the data to agricultural areas.

```
//----Affected Agricultural Areas---//

// Using ESA WorldCover 10m v200
// Spatial resolution at 10 meters
var landcover = ee.ImageCollection("ESA/WorldCover/v200")
  .first()
  .clip(aoi);

// Extract only cropland pixels, map class equal to 40
var cropmask = landcover
  .eq(40);
var cropland = landcover
  .updateMask(cropmask)

// get ESA WorldCover projection
var LCprojection = landcover.projection();

// Reproject flood layer to ESA WorldCover CRS
var floodFrequencyMap_reproj1 = floodFrequencyMap
    .reproject({
    crs: LCprojection
  });
```

```
// Calculate affected cropland using the resampled flood layer
var cropland_affected = floodFrequencyMap_reproj1
  .updateMask(cropland)

// get pixel area of affected cropland layer
var crop_pixelArea = cropland_affected
  .multiply(ee.Image.pixelArea()); //Calculate the area of each pixel

// sum pixels of affected cropland layer
var crop_stats = crop_pixelArea.reduceRegion({
  reducer: ee.Reducer.sum(), //sum of the area of all pixels
  geometry: aoi,
  scale: 30,
  maxPixels: 1e9
  });
```

Visualizing Outputs in GEE

Visualization parameters, including the minimum and maximum values to display and color palette, should be defined in GEE prior to adding layers in the map display. Layers that will not be loaded when the code is implemented will be set to false.

```
/***************************************
  C. DISPLAY OUTPUTS
***************************************/
// Population Count
var populationCountVis = {
  min: 0,
  max: 200.0,
  palette: ['060606','337663','337663','ffffff'],
};
Map.addLayer(population_count, populationCountVis, 'Population Count', false);

// Exposed Population
var populationExposedVis = {
  min: 0,
  max: 200.0,
  palette: ['yellow', 'orange', 'red'],
};
Map.addLayer(population_exposed, populationExposedVis, 'Exposed Population');

// ESA WorldCover
Map.addLayer(landcover, {}, 'Land Cover 2021', false);

// Cropland
var croplandVis = {
  min: 0,
  max: 14.0,
  palette: ['30b21c'],
```

```
};
Map.addLayer(cropland, croplandVis, 'Cropland', false)

// Affected Cropland
Map.addLayer(cropland_affected, croplandVis, 'Affected Cropland');
```

You have now successfully completed a quick flood impact analysis in GEE. You can explore other datasets in GEE and follow the same methodology to assess the potential impact of flooding for different socio-economic parameters.

Checkpoint: Printing Values and Exporting Your Data

Use the print function in GEE to show the outputs from the reducers. Review Part II of this training module. Adopt the export code that you previously developed to export the datasets of the exposed population and affected agricultural areas.

Revise your flood assessment code by adding the codes corresponding to the abovementioned tasks.

Questions

What other datasets in GEE that can you use for estimating the potential impact of flooding in Kawarang, Indonesia?

Are there any other visualization features in GEE that you think is suitable in effectively communicating the impacts of flooding? Why are these visualization features able to communicate the extent of the damages due to flooding?
