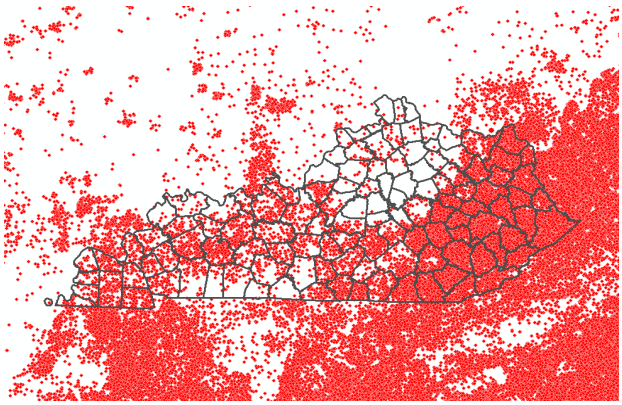
LAB 7 – ENVIRONMENTAL DISTRESS IN KENTUCKY: MAPPING MINES AND FIRES

What you'll Learn: How to use federal and state GIS data to analyze point locations of surface coal mining and wildfire occurrences in Kentucky. The new geoprocessing tools you'll use are **Clip** and **Spatial Join**. You'll revisit **Field Calculator** and **Summarize.**

What You'll Produce: Building off Lab 6, you'll make a series of maps that aggregate occurrences of wildfires and surface mining by census block groups. You'll make a three-map layout using these two environmental distress maps with the map you made in Lab 4, Kentucky Distressed Block Groups.

Data: You'll find necessary data on government agency websites at the links provided. In addition to select datasets from Lab 6, we will be downloading a **surface coal mining dataset** and a **wildfire dataset.** This is live data harvested in the wild.

Background: The surface coal mining data is a point dataset that shows surface mining permits (easternmost intersection of the access road with a public road). Latitude and longitude data are submitted by the permittee in the permit application. Every effort has been made to verify the location. In some cases, latitude and longitude have been field-verified by KY Dept. of Natural Resources personnel. The database is used for tracking all permitting and enforcement actions for coal mining activity in the Commonwealth of Kentucky. It is updated nightly.



The wildfire data contains a spatial database (GPKG) of wildfires that occurred in the United States from 1992 to 2015, generated for the national Fire Program Analysis (FPA) system with attributes: discovery date, cause, and final fire size. The wildfire records were acquired from the reporting systems of federal, state, and local fire organizations. The resulting product, referred to as the Fire Program Analysis fire-occurrence database (FPA FOD), includes nearly 1.6 million geo-referenced wildfire records, representing a total of 113 million acres burned during the 23-year period.

Lab naming conventions: Tools that you click will be bolded, e.g., **QGIS Menu > File > New** to create a new QGIS project file. Text that you'll type will have quotes around it, such as "MyNewProject.qgs" and names of existing datasets and directories will be *italicized*, e.g., *DataToUse.zip*. Key terms will be underlined. **Important tips and key instructions will be in bold red font.**

STEP 1: MAPPING SURFACE MINE PERMITS

Create a new workspace for Lab 7. Copy your data from Lab 6 into an appropriately named folder in your Lab 7 workspace. This data should include the following and all of their supporting files:

- Kentucky.shp
- KentuckyBlockGroups.shp
- KyDistressBG.shp
- KyCounties.shp

Next, create a new QGIS project titled Lab 7, and add these datasets to the project. Symbolize the distressed block groups layer the same way that you did in Labs 5 and 6.

It's important not to drag data files (e.g., shapefiles) from previous labs into a new folder. Likewise, it's important not to drag project files (e.g., .qgs files) from previous labs into a new folder. This will break the file paths that QGIS is reading in order to display the data, and cause more trouble for you! Instead, it is advisable to either 1) copy-paste data, or 2) simply create a new project file and load in data from previous workspaces without moving it.

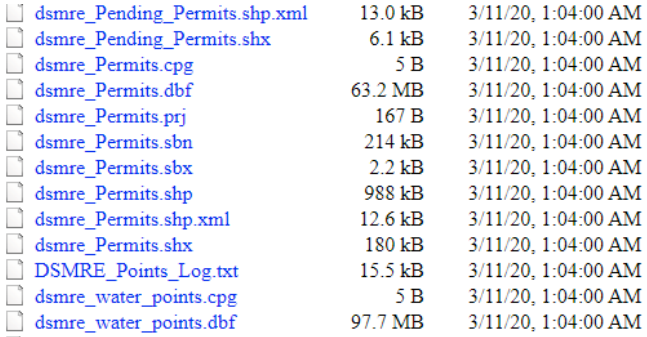
Now, on to downloading our mining permits!

The Kentucky Department of Natural Resources nightly publishes spatial data that contains locations of surface mine operations. The metadata explains that points are, "locations of surface mining permits (easternmost intersection of the access road with a public road)."

Download the data here:

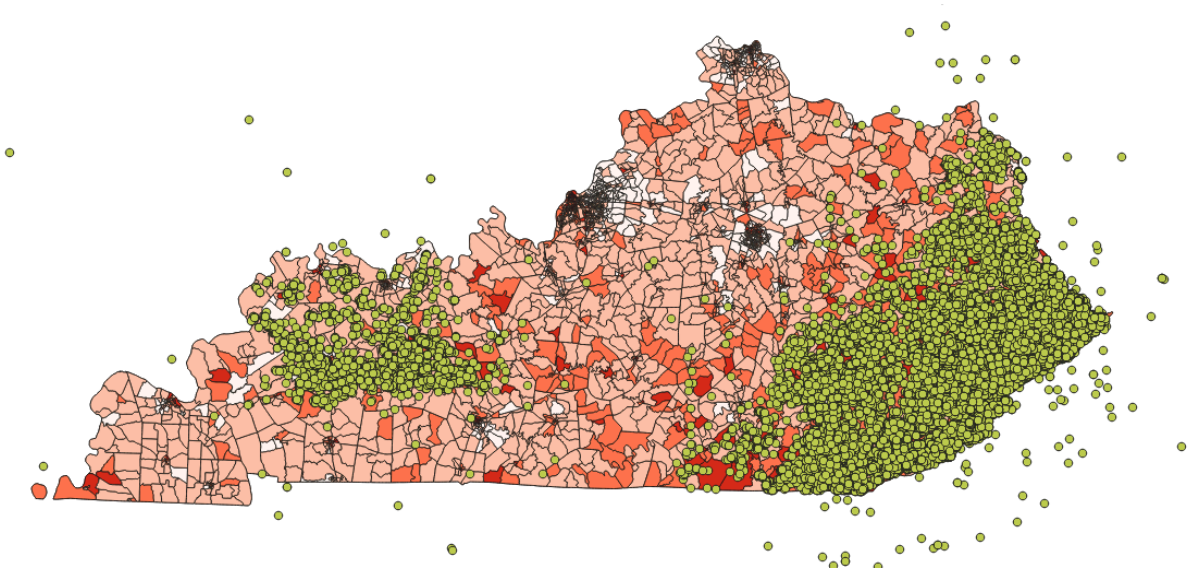
https://eec.ky.gov/Natural-Resources/Mining/Mine-Permits/Pages/maps-spatial-data.aspx

Follow the link to the FTP (file transfer protocol) site, under "Spatial Data." You'll see a screen like this:



dsmre_Pending_Permits.shp.xml	13.0 kB	3/11/20, 1:04:00 AM
dsmre_Pending_Permits.shx	6.1 kB	3/11/20, 1:04:00 AM
dsmre_Permits.cpg	5 B	3/11/20, 1:04:00 AM
dsmre_Permits.dbf	63.2 MB	3/11/20, 1:04:00 AM
dsmre_Permits.prj	167 B	3/11/20, 1:04:00 AM
dsmre_Permits.sbn	214 kB	3/11/20, 1:04:00 AM
dsmre_Permits.sbx	2.2 kB	3/11/20, 1:04:00 AM
dsmre_Permits.shp	988 kB	3/11/20, 1:04:00 AM
dsmre_Permits.shp.xml	12.6 kB	3/11/20, 1:04:00 AM
dsmre_Permits.shx	180 kB	3/11/20, 1:04:00 AM
DSMRE_Points_Log.txt	15.5 kB	3/11/20, 1:04:00 AM
dsmre_water_points.cpg	5 B	3/11/20, 1:04:00 AM
dsmre_water_points.dbf	97.7 MB	3/11/20, 1:04:00 AM

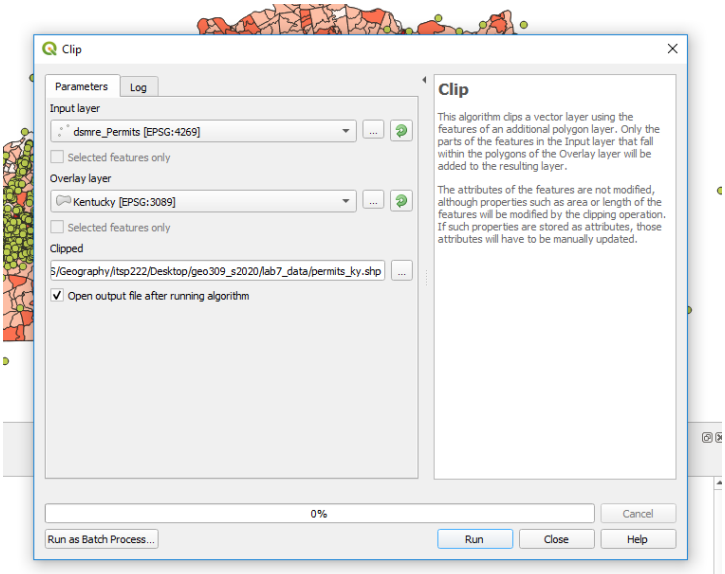
Download all of the files associated with “dsmre_Permits” shapefile (e.g., .cpg, .dbf, .prj, and so on). Save them in an appropriate folder in your workspace. You should then load the permits data into your Lab 7 QGIS project. The resulting layer should resemble the following:



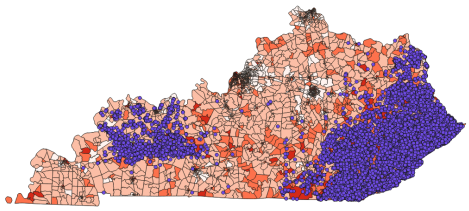
Question: how do we add up the number of permits per block group? The assumption is more permits, more mining. This is a big mess with too much point density to symbolize properly. Here’s where a **Spatial Join** is appropriate.

Begin by **Clipping** the surface mining points with the polygon outline of Kentucky. This will remove the erroneous points that fall outside of Kentucky. This will be especially helpful when you look at wildfire data.

To access the **Clip** tool, navigate to **Vector > Geoprocessing > Clip**, or enter “clip” into the processing toolbox’s search bar.



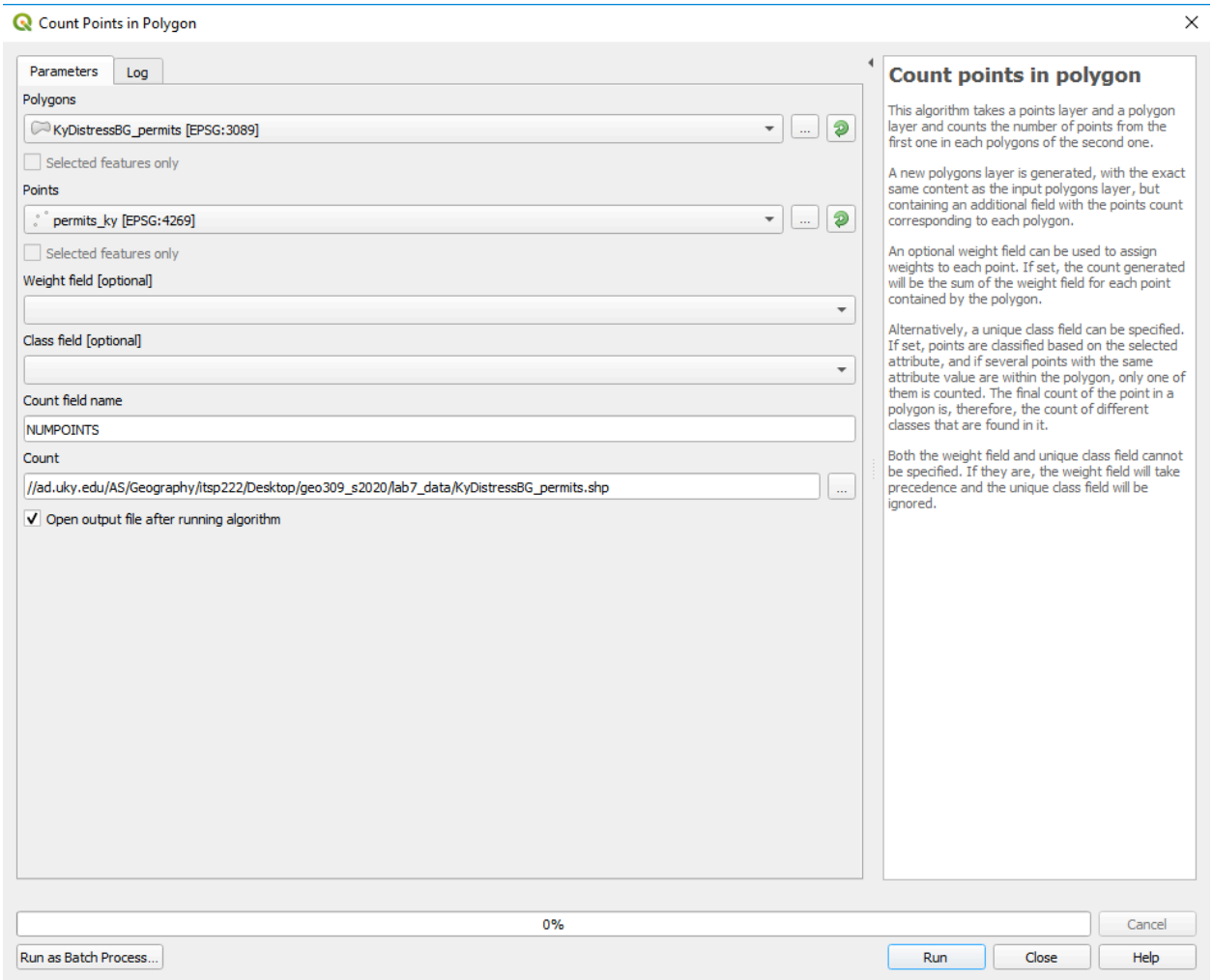
Save the output as “permits_ky” and add the layer to your project. The result should resemble the screenshot below:



Next, you’ll **Count Points in Polygon** for the permits to the *KyDistressBG* layer you created in Lab 5.

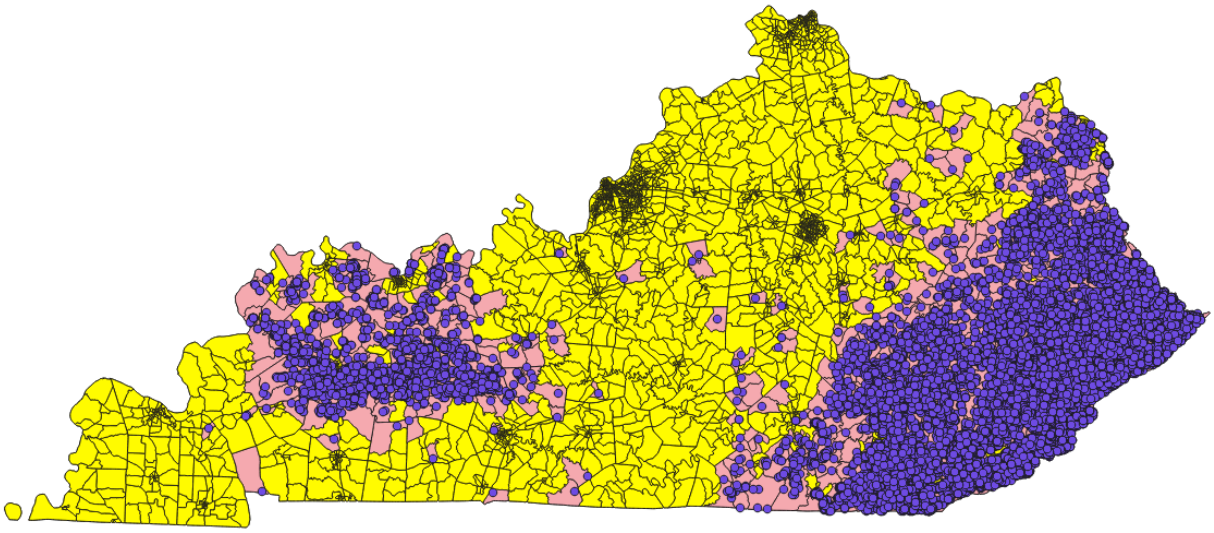
Count Points in Polygon is one way of completing a spatial join. Spatial joins, or attaching attributes from one layer to another based on their spatial relationship, are a fundamental problem of GIS work. In QGIS, **Count points in Polygon** produces useful summaries of the occurrence of one attribute (e.g., a point) within another (e.g., a polygon). Later on, we'll also perform a more robust spatial join using the **Join Attributes by Location** tool. It should be available in **Geoprocessing > Data Management > Join Attributes by Location**, or by searching "join attributes by location" in the processing toolbox's search bar.

In this case, we want to produce a map that shows mine permits per square mile for each block group. So, navigate to **Vector > Analysis tools > Count points in polygon**. Select *KyDistressBG* as your polygon and *permits_ky* as your points. Name the output "KyDistressBG_permits," and save it to your workspace, as shown below:



Open the attribute table and verify your have a "NUMPOINTS" field and you have values. If not, check your coordinate system! Be sure that both of your layers are properly projected in EPSG:3089.

Now select all values where NUMPOINTS = 0. These should be block groups where there are no instances of mining permits. The selection should resemble the following:



Now **Switch Selection** such that only values greater than 0 are selected. Open the **Statistics** panel on the "TotPop_" field to summarize the total population of Kentucky that lives in block groups that have had surface mining. Be sure that the "Selected features only" box is checked.

> **QUESTION 1:** What is the sum value of Kentucky's population living in block groups where mining is active?

Next, **Clear Selection** and open the attribute table, activating the Editor mode. Select **Add Field**, creating a new field called "PermitSqMi" with data type Float.

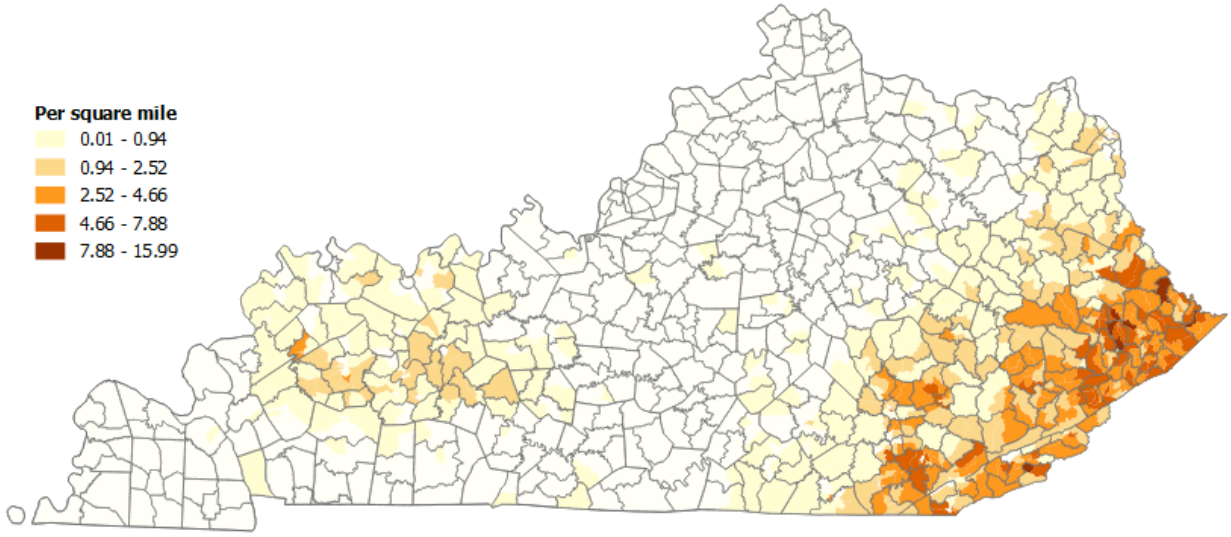
Once you've created your new field, **Field Calculate** it to make Permits per Sq. Mile. We've done this in previous labs and exercises: you're basically **normalizing data** here. To get permits per square mile, you'll need to divide the total instances of mines (e.g., "NUMPOINTS") by the square mileage (e.g., "Sq_Mi").'

Sort the "MinePermitsPerSqMi" so that you can find the block with the highest density of mine permits.

> **QUESTION 2:** What is the lowest density of permits per square mile? What is the highest density?

> **QUESTION 3:** What county is that block group in?

To finish this part make a map resembling the one below, which shows mine permits per square mile for each block group.



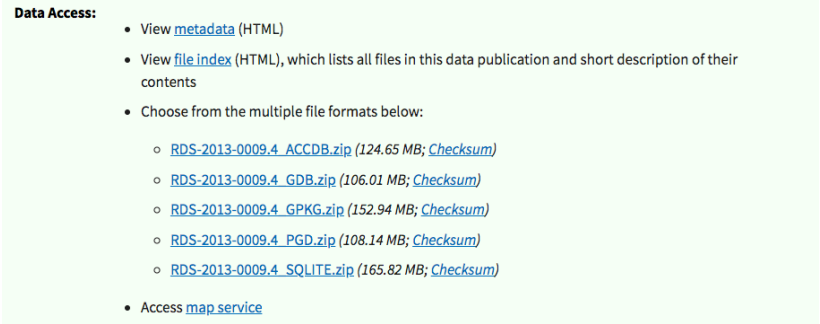
Your map should be more detailed than this, including:

- Author
- Date of publication
- Scale, north arrow (make sure each data frame is in same scale)
- Source and date range of data (e.g., Division of Mine Permits, Kentucky Department of Natural Resources, 1970s – February, 2015)

STEP 2: ADDING WILDFIRES TO THE ANALYSIS

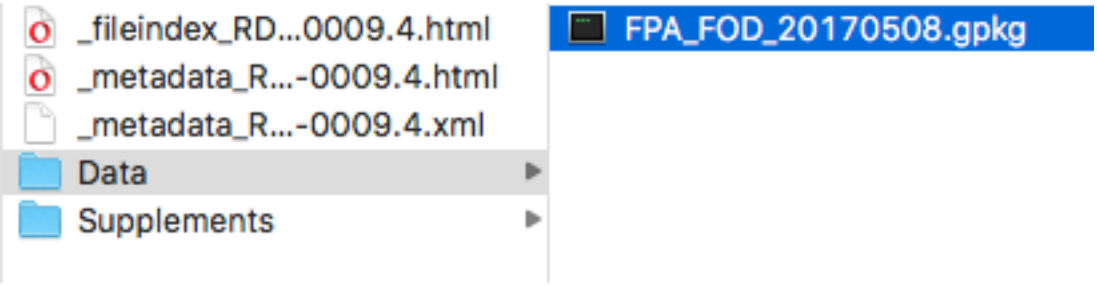
In this section, we'll be comparing the map of distress to instances of wildfires in Kentucky. The final step will bring all of these analyses together.

Navigate to the following link (https://www.fs.usda.gov/rds/archive/catalog/RDS-2013-0009.4) and download the file RDS-2013-0009.4_GPKG.zip.

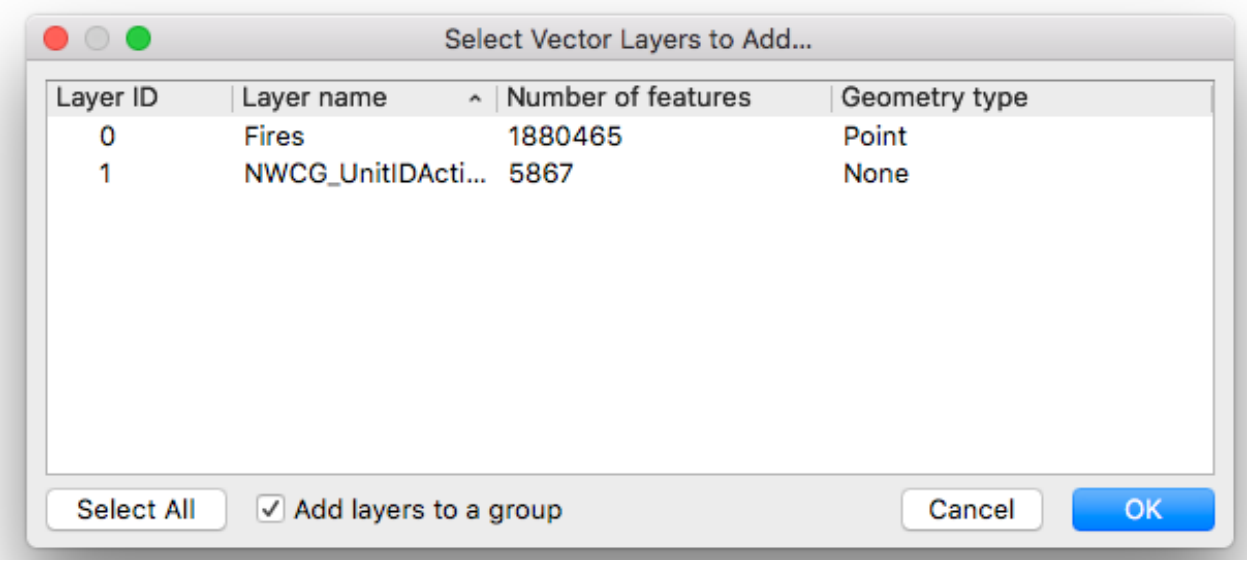


Make sure you download the GPKG file. This is a much bigger dataset and we need to be careful how we work it. Once you download it, unzip

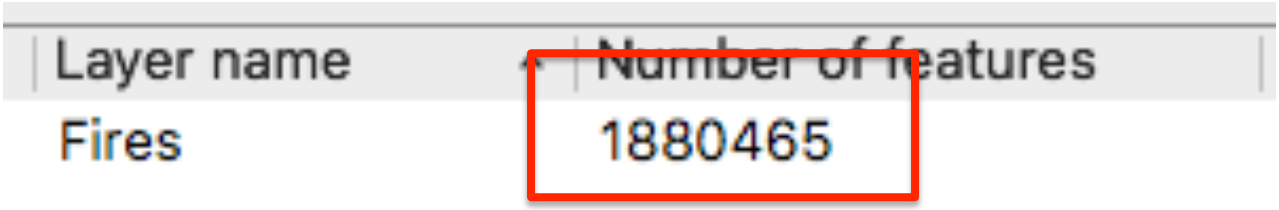
and move to your data folder in your workspace. You should see a file structure that resembles the following:



Add the file “FPA_FOD_20170508.gpkg” to your QGIS project. When you do so, you will be prompted with this popup:

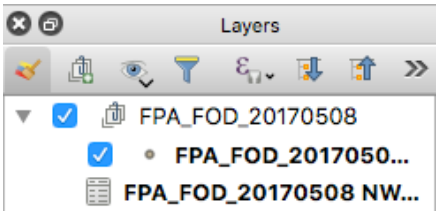


Wait a minute… am I reading this right???



Yep… that’s 1,880,465 point features. There’s no other way to say it: this is a Big Ass File. Go ahead and **Select All** and then click **OK**.

After adding, your table of contents should resemble the panel on the right. The GPKG file adds as a group that contains a tabular dataset and a point layer.



Right+click the point layer and zoom to it. You'll probably see something like this start to form:

After about ten minutes, it will probably look like this:

It draws very slowly. You can either watch the first three seasons of *Game of Thrones* while it loads, or simply turn the layer off. In the interest of finishing this lab on time, I recommend the latter (although I do think that *Game of Thrones* is an excellent series).

After clicking the layer off, let's **Clip** to Kentucky to remove unnecessary records. Use the wildfires layer as the input and the Kentucky layer as the overlay. It may take a minute to run, the gargantuan file that it is, but once you're done you can double-check to make sure it's correct by confirming that there are about 27,000 records. Make sure you've saved the output as "wildfires_ky" and added your new layer to the project.

Now we're going to perform a **Spatial Join**, but before we do so, we need to get rid of some of the fields that we don't need from our new KY wildfires dataset. Instead of deleting them outright, we can create a new file that the fields are cut from.

Right+click the layer *wildfires_ky* and select **Export > Save Features As…** Name the output "fires_ky_short", save it as a shapefile, and in the dialog box where you can

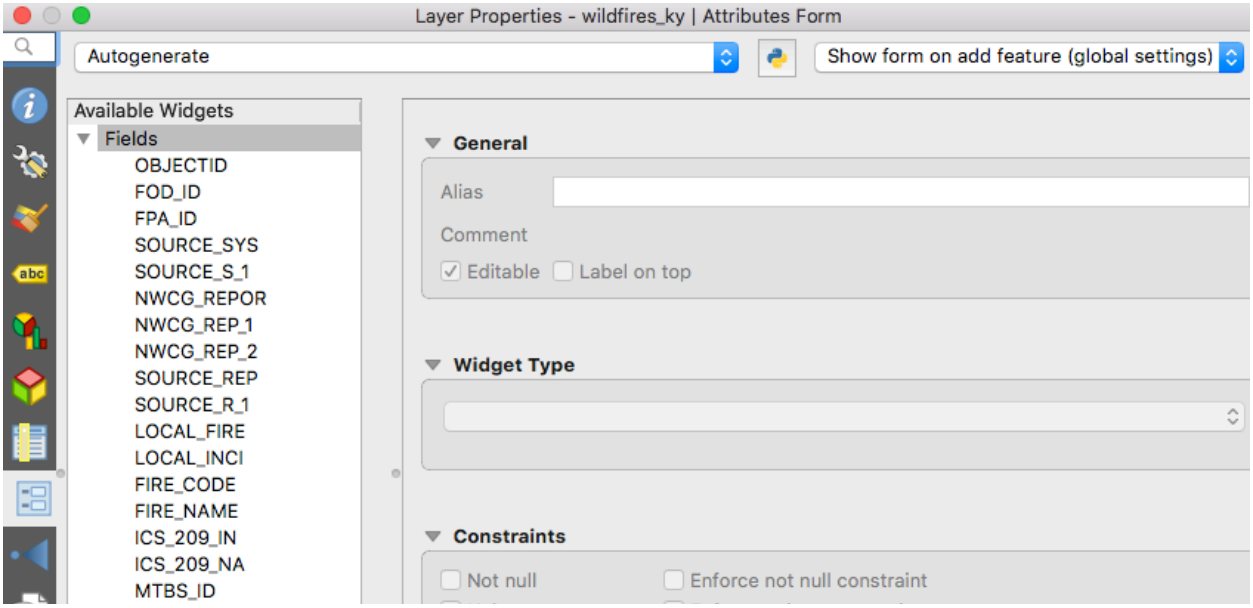
select fields to export, check the following and uncheck the rest (note that you can use the “deselect all” option to make this task much, much easier:

- FIRE_NAME
- FIRE_YEAR
- STAT_CAU_1
- FIRE_SIZE
- FIRE_SIZE_

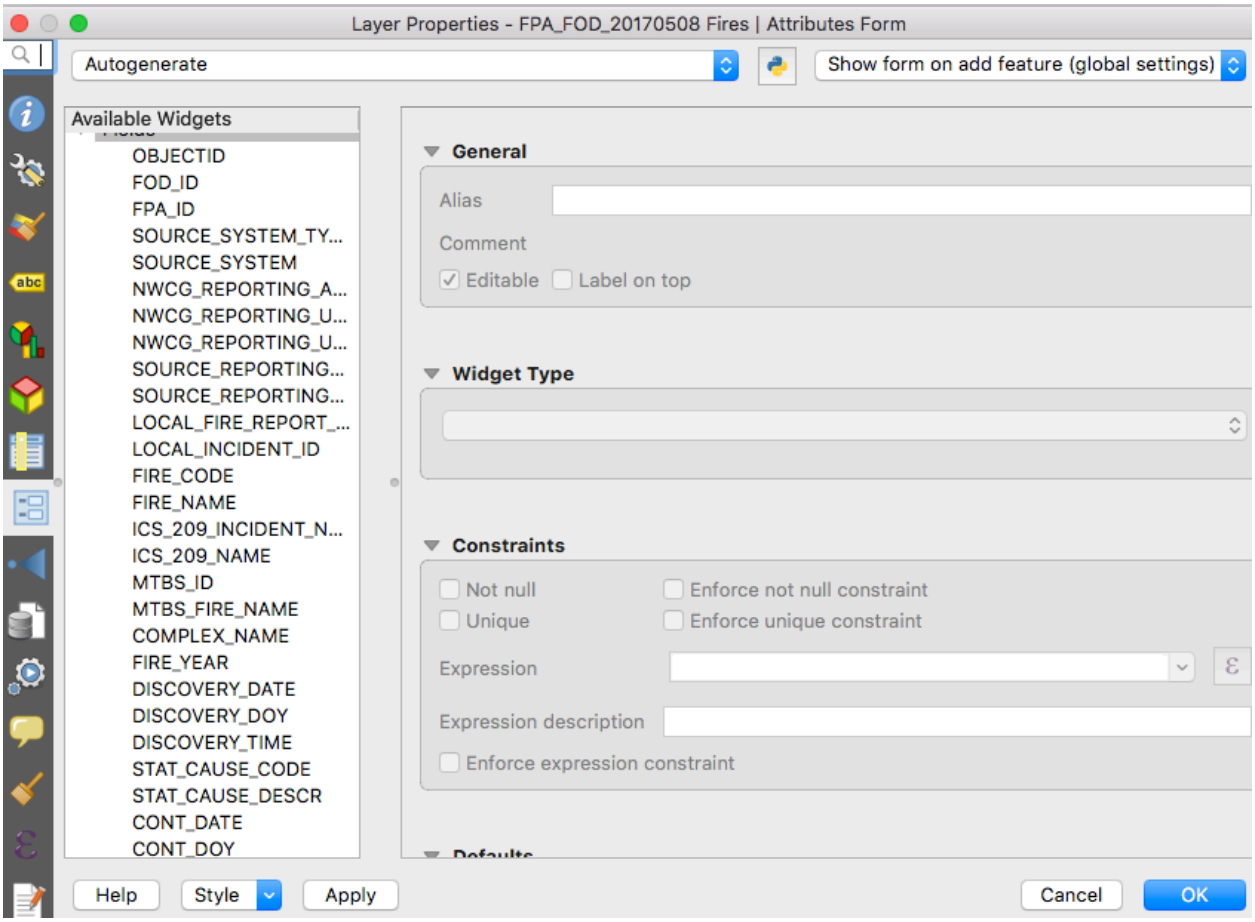
Your output should resemble the screenshot below, and be sure that you’re exporting the CRS to EPSG:3089. Add the new layer to your map.

It is worth pausing to talk about field names at this point. The field names we selected are strange at best, and incomprehensible at worst. **This is an example of the limitations of using shapefiles in QGIS**. In QGIS, shapefile field names can only handle ten characters. If you are running any kind of geoprocessing, the field names will be cut off at 10, and duplicate names will be automatically appended with an underscore (“_”) or an underscore + 1 (“_1”). This can be seen in our case, with STAT_CAU_1 and FIRE_SIZE_.

If you navigate to **Properties > Attribute Form** in the KY wildfires layer, you’ll note that many of the fields have been cut off in just this way. That happened when we exported as a shapefile. See below:



Note how instead, if you navigate to **Properties > Attribute Form** in the original national wildfires dataset, all of the field names are much longer. That’s because this file is stored in a **Geopackage format**. The QGIS geopackage (GPKG) format can indeed handle field names longer than ten characters. Considering this, we can look through the field names we selected and get a bit more clarity. See below:

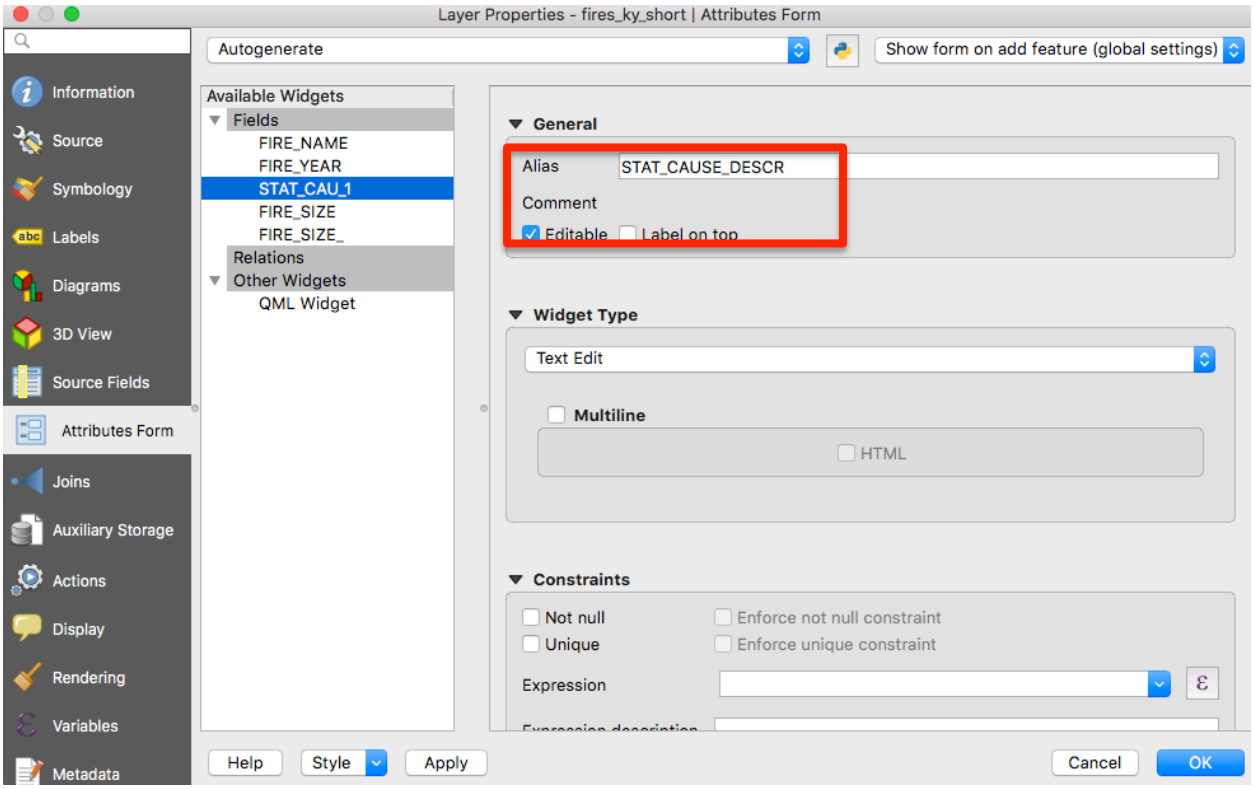


By inspecting and comparing the two attribute table forms, it becomes clear that the field names in our *fires_ky_short* layer are actually derived as follows:

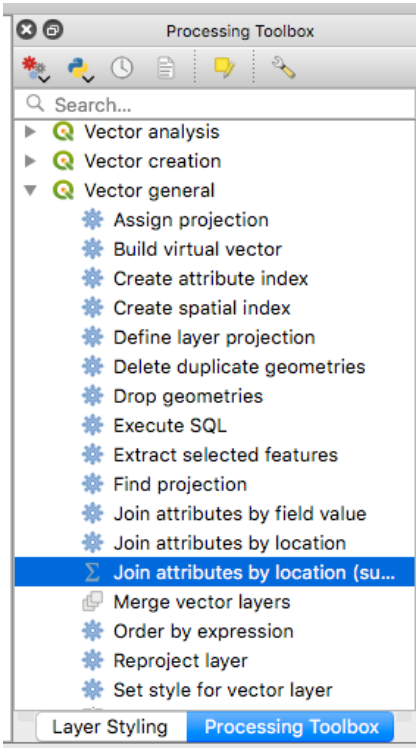
- FIRE_NAME = FIRE NAME
- FIRE_YEAR = FIRE_YEAR
- STAT_CAU_1 = STAT_CAUSE_DESCR
- FIRE_SIZE = FIRE_SIZE
- FIRE_SIZE_ = FIRE_SIZE_CLASS

To fix this confusion, navigate once again to **Properties > Attribute Form** in the *fires_ky_short* layer, and rename STAT_CAU_1 and FIRE_SIZE_ to more appropriate titles by typing those titles into the “Alias” field and clicking **Apply**.

Once you’re done, you should see those changes reflected in the actual attribute table.



At this point, you can remove the national fires layer and the original KY fires layer.

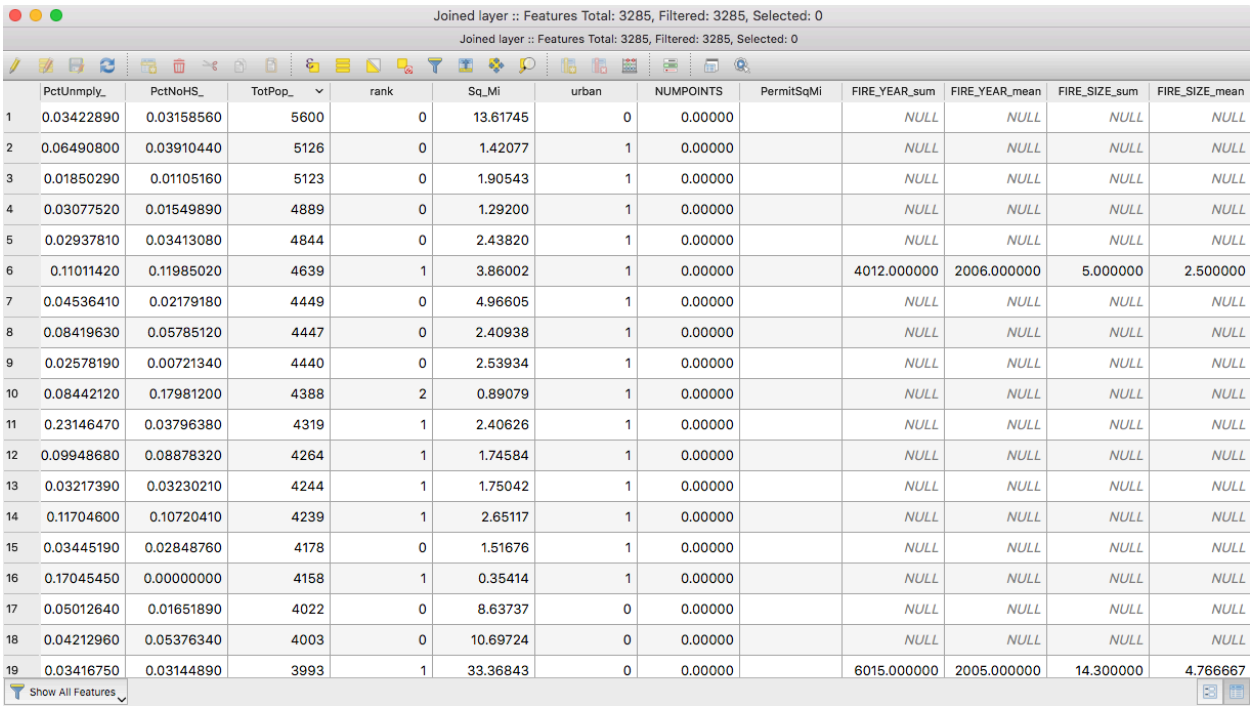


Now it's time to perform a spatial join on *KyDistressBG_permits*. In the processing toolbox, double-click the **Join attributes by location (summary)** tool, as seen on the right.

Enter the following parameters in the dialog:

- Use *KyDistressBG_permits* as the input layer.
- Use *fires_ky_short* as the join layer.
- Set the geometric predicate to "intersects."
- Leave "Fields to summarize" empty, because we want to use all fields.
- Under "Summaries to calculate," click the ellipsis and select "Sum" and "Mean."
- **Leave the final parameter empty such that you create a temporary layer (the output will be titled "Joined layer")**
- Click **Run**!

Open the attribute table of the *Joined layer* output. You should see some new fields that resemble the screenshot below:



	PctUnmply_	PctNoHS_	TotPop_	rank	Sq_Mi	urban	NUMPOINTS	PermitSqMi	FIRE_YEAR_sum	FIRE_YEAR_mean	FIRE_SIZE_sum	FIRE_SIZE_mean
1	0.03422890	0.03158560	5600	0	13.61745	0	0.00000		NULL	NULL	NULL	NULL
2	0.06490800	0.03910440	5126	0	1.42077	1	0.00000		NULL	NULL	NULL	NULL
3	0.01850290	0.01105160	5123	0	1.90543	1	0.00000		NULL	NULL	NULL	NULL
4	0.03077520	0.01549890	4889	0	1.29200	1	0.00000		NULL	NULL	NULL	NULL
5	0.02937810	0.03413080	4844	0	2.43820	1	0.00000		NULL	NULL	NULL	NULL
6	0.11011420	0.11985020	4639	1	3.86002	1	0.00000		4012.000000	2006.000000	5.000000	2.500000
7	0.04536410	0.02179180	4449	0	4.96605	1	0.00000		NULL	NULL	NULL	NULL
8	0.08419630	0.05785120	4447	0	2.40938	1	0.00000		NULL	NULL	NULL	NULL
9	0.02578190	0.00721340	4440	0	2.53934	1	0.00000		NULL	NULL	NULL	NULL
10	0.08442120	0.17981200	4388	2	0.89079	1	0.00000		NULL	NULL	NULL	NULL
11	0.23146470	0.03796380	4319	1	2.40626	1	0.00000		NULL	NULL	NULL	NULL
12	0.09948680	0.08878320	4264	1	1.74584	1	0.00000		NULL	NULL	NULL	NULL
13	0.03217390	0.03230210	4244	1	1.75042	1	0.00000		NULL	NULL	NULL	NULL
14	0.11704600	0.10720410	4239	1	2.65117	1	0.00000		NULL	NULL	NULL	NULL
15	0.03445190	0.02848760	4178	0	1.51676	1	0.00000		NULL	NULL	NULL	NULL
16	0.17045450	0.00000000	4158	1	0.35414	1	0.00000		NULL	NULL	NULL	NULL
17	0.05012640	0.01651890	4022	0	8.63737	0	0.00000		NULL	NULL	NULL	NULL
18	0.04212960	0.05376340	4003	0	10.69724	0	0.00000		NULL	NULL	NULL	NULL
19	0.03416750	0.03144890	3993	1	33.36843	0	0.00000		6015.000000	2005.000000	14.300000	4.766667

There are a lot of NULL values in our four new fields. That's okay for now. Enter **editing mode**. The field FIRE_YEAR_sum is nothing, so let's delete it. The other three contain valuable summary information that was generated through attaching attribute data between different layers, based on spatial parameters:

- FIRE_YEAR_mean: the average year a fire occurred in a block group, between 1992 and 2015
- FIRE_SIZE_sum: total acres burned per block group between 1992 and 2015

If we save this layer as a shapefile, a lot of these field names will be cut off. Let's save ourselves the hassle, and instead, save it as a geopackage. **Right+click** the *Joined layer* and select **Export > Save data as…** Choose a geopackage as the file type, and name the output *MinePermits_Fires_BG*. Add the new layer to your project and notice that the attribute table field names haven't been cut off.

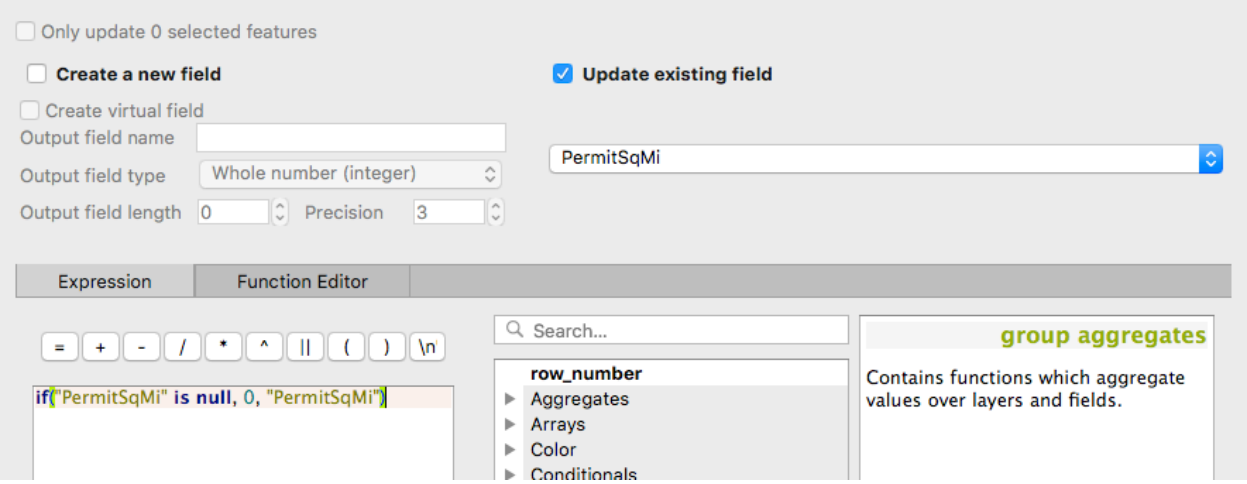
To take a step back again, you should begin to see some contours of a larger research project emerge. A few labs ago, we were looking into socioeconomic distress in KY. This lab we have used data about mining permits and wildfires to create datasets of environmental distress. **Now, we have produced a spatial dataset that has socioeconomic and environmental distress in one attribute table.**

Let's take a moment to get rid of these empty values in the fields. Open the attribute table for *MinePermits_Fires_BG*. **Enter editing mode**. You want to be in editing mode for this, because if you enter the function incorrectly and erase a bunch of important data, you'll want to be able to quit editing mode without and discard those changes.

Once you're in editing mode, open the **Field calculator,** check "update existing field," and select "PermitSqMi" as the field. Then enter the following in the Expression dialog:

if("PermitSqMi" is null, 0, "PermitSqMi")

Your screen should resemble the following almost exactly:



This expression should replace all NULL values in your "PermitSqMi" field with 0 values.

Once you've clicked **OK**, confirm that it worked properly in the attribute table. If it worked (e.g., it didn't erase all your other values), save your changes by closing editing mode and saving changes. Then, run the same field calculation for the other three fields: "FIRE_YEAR_mean", "FIRE_SIZE_sum", and "FIRE_SIZE_mean".

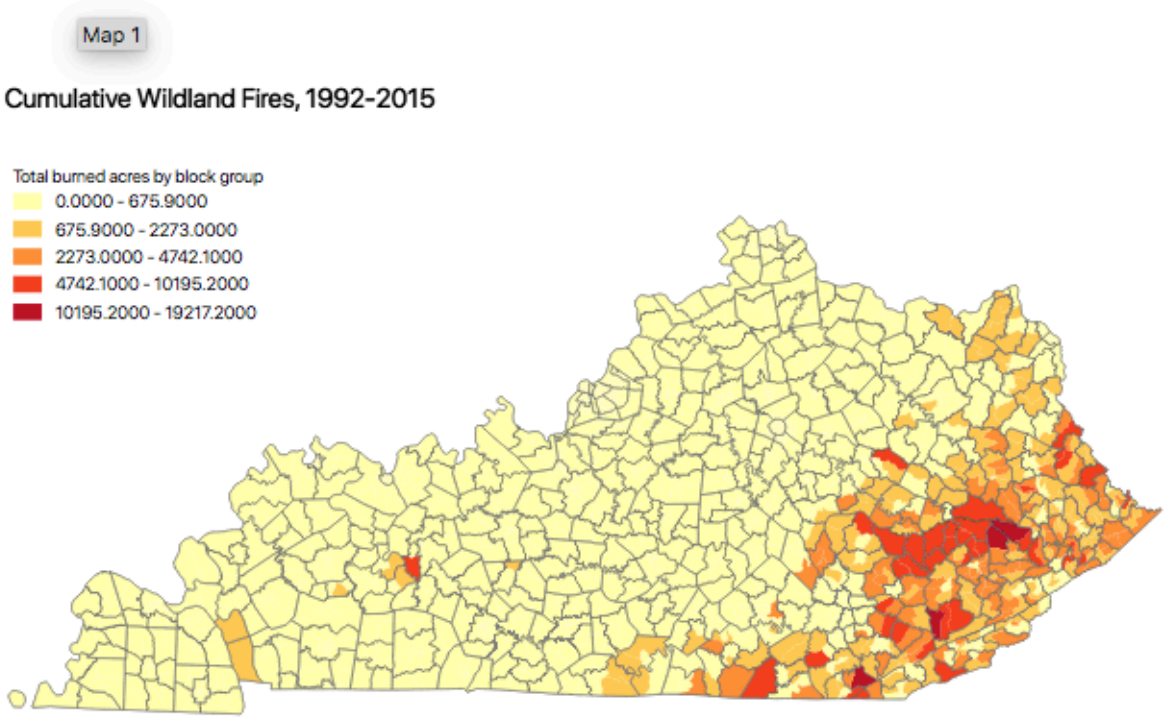
TIP: I can't emphasize how important it is to use the editing function very intentionally during field calculations like this. If you screw something up outside of the editing mode, it will take a lot of time to re-do things. Use the editor smartly!

QUESTION 4: What is the sum value of Kentucky's population living in block groups where wildfires are active? TIP: refer to Q1 for how to answer this question using the Statistics tool.

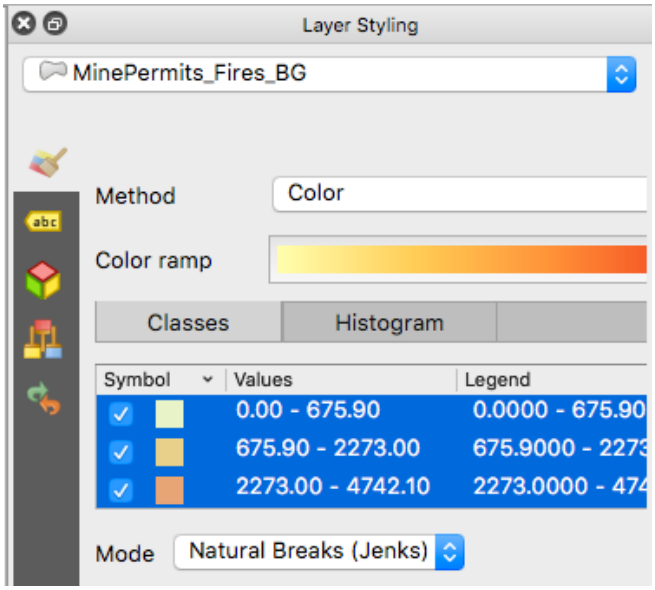
QUESTION 5: Sort your attribute table by the column "FIRE_SIZE_mean." In what county is the block group with the greatest average fire size?

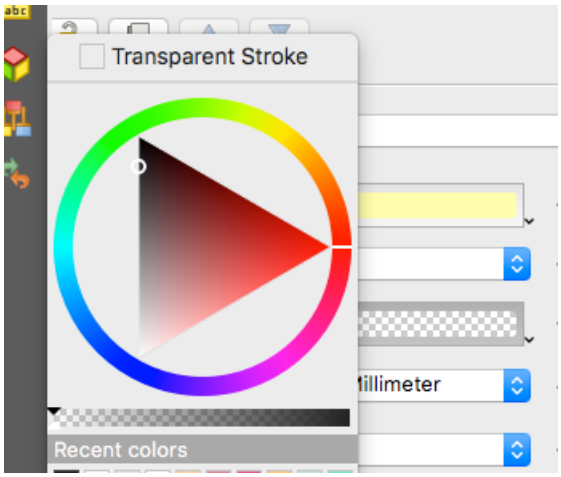
Now that you've got your dataset all sorted, make map showing cumulative acres of wildfire burn areas by block group. The field we want to map is "FIRE_SIZE_sum" which is the cumulative acres burned by block group over the time period of our dataset, 1992-2015. You might think we can normalize fire size by block group to get a percent burned. Sorta, it won't be a literal percent burned, since a small area could have many repeated fires. But nonetheless, we'll make

a map! Open the **Layer styling panel** and create a map that resembles the following:



TIP: Remove the outlines of block groups by selecting all classes in your layer styling panel, **double+click** one of them while holding the shift key, and change the stroke color of the simple fill to “Transparent Stroke.”





As with the map of surface mining permits, your map should be more detailed than this, including:

- Author
- Date of publication
- Scale, north arrow (make sure each data frame is in same scale)
- Source and date range of data (e.g., Division of Mine Permits, Kentucky Department of Natural Resources, 1970s – February, 2015)

STEP 3: REVIEW & SUBMISSION

You should submit completed materials for Lab 7 via Canvas by Monday, 3/23 at 11:59pm. The completed materials include:

- A Word document containing answers to the questions posed throughout the lab
- Two maps, formatted in print composer w/ all requisite elements, in .pdf or .png
    - One map of mining permit sites
    - One map of wildfires

Be in touch if you have any questions!

For extra credit worth 2 points, create three-map layout that contains the following maps showing distress by block group:

1. Socioeconomic Distress by block group from Labs 5 and 6
2. Surface coal mine permits per sq. mile
3. Cumulative acres burned by wildfires

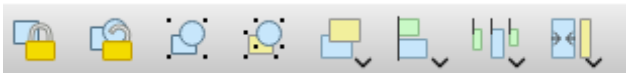
Make sure you have the following elements for each map

1. Descriptive Title (e.g., Surface Coal Mine Permits Per Square Mile)
2. Legend
3. Source and date range of data (e.g., Division of Mine Permits, Kentucky Department of Natural Resources, 1970s – February, 2015)

And for the layout as a whole, you'll need to add

1. Main Publication Title
2. Author
3. Date of publication
4. Scale, north arrow

Order you maps evenly and make them aligned. In Print composer mode, use the options below:



An example map can be found below: