

Understanding History through Data: Cholera in Amsterdam, 1866

As you have seen in the map room at the University Library, Cholerakaart (see fig. 1) represented the number of deaths occurred in each district in 1866 when cholera stroked Amsterdam and caused more than 1000 death in merely five months. What does this map tell us, the readers in 21st century about this tragic event over a hundred years ago?



Figure 1 Cholerakaart

PART I: Understanding Cholerakaart

Cholerakaart shows the deaths of each *wijk* spanning across 20 weeks starting from the first of June, to 18th of October, 1866. The detailed number of deaths in each week is presented in a color-coded wheel. How to read these wheels/pie charts? Taking **wijk W** for example (fig. 2):

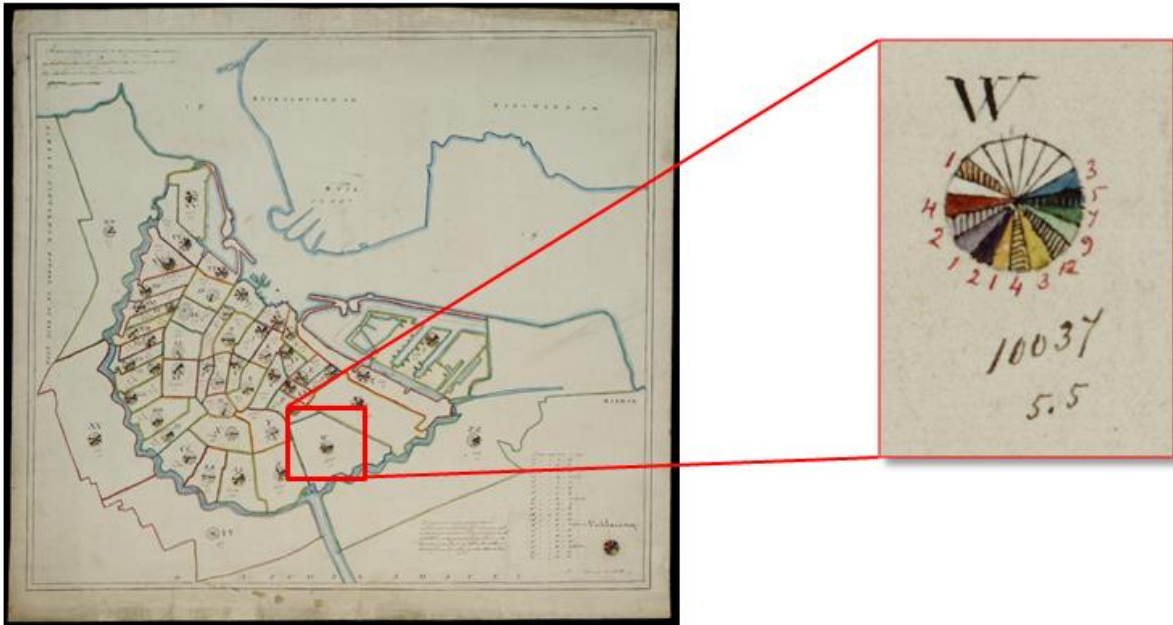


Figure 2 Wijk W

According to the legend in the lower right corner of the map (fig. 3), we see how the colors (and location) of each slide translated to a particular week. For instance, the light blue slide in the legend indicated week #5, 29-6-1866 to 5-7-1866. Looking at wijk W's wheel, the light blue slide (week 5) has a number "3" next to it, meaning there were 3 deaths in wijk W during week 5, between 29th June and 5th July, 1866. The uncolored slide means no death occurred during that week. The numbers below the wheel shows the population of wijk W and the total mortality rate (per thousand) during the entire five months.

1 ^{de}	Week van	1 tot	7 Juny.
2 ^{de}	"	8 "	14 "
3 ^{de}	"	15 "	21 "
4 ^{de}	"	22 "	28 "
5 ^{de}	"	29 "	5 July.
6 ^{de}	"	6 "	12 "
7 ^{de}	"	13 "	19 "
8 ^{de}	"	20 "	26 "
9 ^{de}	"	27 "	2 August.
10 ^{de}	"	3 "	9 "
11 ^{de}	"	10 "	16 "
12 ^{de}	"	17 "	23 "
13 ^{de}	"	24 "	30 "
14 ^{de}	"	31 "	6 September.
15 ^{de}	"	7 "	13 "
16 ^{de}	"	14 "	20 "
17 ^{de}	"	21 "	27 "
18 ^{de}	"	28 "	4 October.
19 ^{de}	"	5 "	11 "
20 ^{de}	"	12 "	18 "



Figure 3 Choleraart legends

Similarly, if we write down all the numbers in wijk W's wheel in an excel table, it looks like this:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
1	ID	Wijk1850	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Total	Population	Mortality Rate(/1k)
23	22	W	0	0	0	0	3	5	7	9	12	3	4	1	2	1	2	4	0	1	0	0		10037	5,5

In the total cell, you can calculate the total death in W over the twenty weeks:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	ID	Wijk1850	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Total	Population
23	22	W	0	0	0	0	3	5	7	9	12	3	4	1	2	1	2	4	0	1	0	0	=sum(C23:V23)	

PART II: A data view of Cholerakaart

Task 1: Fill in the data for wijk B-F

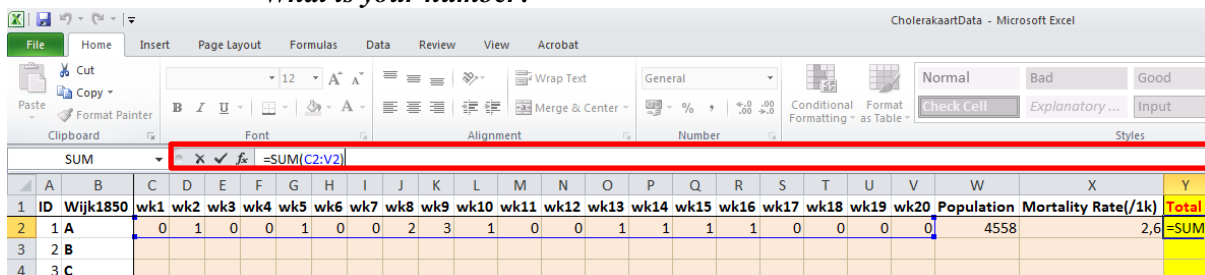
1. Open CholerakaartData.xlsx and go to sheet 'Cholerakaart'
2. Read Cholerakaart and fill the table for wijk B to F (A is filled as an example)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	ID	Wijk1850	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Population	Mortality Rate(/1k)
2	1	A	0	1	0	0	1	0	0	2	3	1	0	0	1	1	1	1	0	0	0	0	4558	2,6
3	2	B																						
4	3	C																						
5	4	D																						
6	5	E																						
7	6	F																						

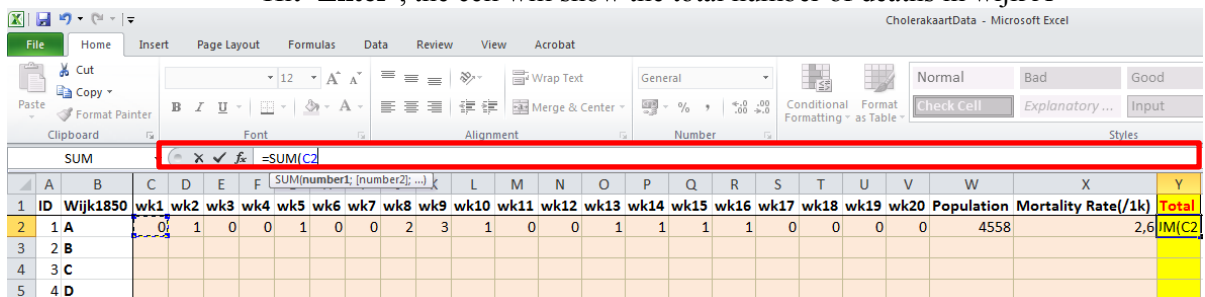
Task 2: Calculate the total number of death

The total number of deaths in wijk A is a sum of deaths in every week. The following steps show how to calculate total death for wijk A in Excel:

- o **Method 1:**
 - Click the cell under 'Total'
 - Click the formula line (red box)
 - Type in formula: '=sum(C2:V2)' (meaning this cell shows the value equals the sum from cell column C, row 2 to column V row 2)
 - Hit 'Enter', the cell will show the total number of deaths in wijk A
 - *What is your number?*



- o **Method 2:**
 - Click the cell under 'Total'
 - Click the formula line (red box)
 - Type in formula: '=sum('
 - Click cell C2 and drag your mouse to cell V2
 - Hit 'Enter', the cell will show the total number of deaths in wijk A



- Click cell Y2 and move your mouse to the lower right corner and the mouse arrow turns into a cross:
+
- Drag down to automatically apply the same formula to other cells

ID	Wijk1850	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Population	Mortality Rate(/1k)	Total
1 A		0	1	0	0	1	0	0	2	3	1	0	0	1	1	1	1	0	0	0	0	4558	2,6	12
2 B																								0
3 C																								0
4 D																								0
5 E																								0
6 F																								0

Task 3: Calculate the total number of death using mortality rate

An alternative way to calculate the total number of deaths in a wijk is to use the population and mortality rate. Mortality rate in the Cholerakaart indicates the number of death per thousand people in a wijk. Therefore we can calculate the total deaths by population and mortality rate:

$$\text{Total death} = \text{Total population} \times \text{per thousand Mortality Rate} / 1000$$

ID	Wijk1850	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Total	Population	Mortality Rate(/1k)	Total deaths
1 A		0	1	0	0	1	0	0	2	3	1	0	0	1	1	1	1	0	0	0	0	12	4558	2,6	=X2*Y2/1000
2 B																									
3 C																									
4 D																									
5 E																									
6 F																									

- Click the cell under 'Total deaths'
- Type in formula: `'=X2*Y2/1000'`
- Or type '=' and then click cell X2, type '*', click cell Y2 and type '/1000'
- Hit **Enter**, the cell will show the calculated number of deaths in wijk A

LET OP: the calculated number of deaths may not be an integer, but the number of death has to be an integer since a person can either be dead or alive. We cannot have a 0.3 person died. Therefore the calculated deaths (usually with decimals) need to be rounded into integer. In this case of wijk A:

$$\text{Total deaths in wijk A} = \frac{4558 \times 2.6}{1000} = 11.85 \approx 12$$

We thus say the calculated deaths in wijk A is 12, instead of 11.85.

Question:

Now we have the sum of all deaths in each week or the calculated number using population and mortality rate. Which number do you think is more accurate? Please justify your answer.

Hint: think about where the mortality rate comes from.

Task 4: Calculate the missing morality rate



Figure 4 details for wijk RR and NN

Due to some reasons, wijk RR did not have a morality rate, although the population and ‘death wheel’ available.

Can you calculate the morality rate in Excel for wijk RR?

- Going back to **CholeraData.xlsx**
- Fill in the morality rate cell for row RR (marked in orange)

Task 5: Dealing with missing data

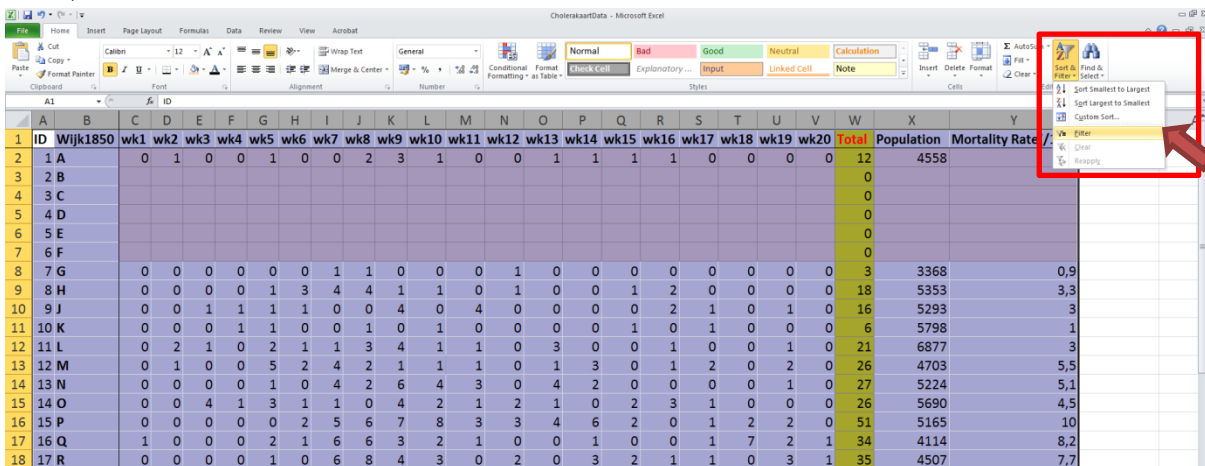
For a map that is over 150 years old, it is not surprising that the map is not in a perfect state. As you can see from figure 4, some numbers around the color wheel for wijk NN is unreadable. Luckily the population and morality rate number are intact.

Based on what we have learned from the calculated total deaths from morality rate and population, how would you deal with this unreadable data?

Task 6: Rank by attributes

We would like to know which wijk had the highest total death number and we want to know the top 5 wijks that have the highest death rate.

Select the table you want to, go to **Home** menu, to the right end, click on **Sort & Filter**, in the stroll down menu, click **Filter**:



Click the arrow on **Total** column, then in the stroll down menu, click **Sort Largest to Smallest**:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	
1	IC	Wijk185C	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Total	Population	Mortality Rate(/1k)	Total death	
2	1A		0	1	0	0	1	0	0	2	3	1	0	0	1	1	1	1						4558		2,6	11,85
3	2B																										
4	3C																										
5	4D																										
6	5E																										
7	6F																										
8	7G		0	0	0	0	0	0	1	1	0	0	0	1	0	0	0	0						3368		0,9	
9	8H		0	0	0	0	1	3	4	4	1	1	0	1	0	0	1	2						5353		3,3	
10	9J		0	0	1	1	1	0	0	4	0	4	0	0	0	0	0	2							5293		3
11	10K		0	0	0	1	1	0	0	1	0	1	0	0	0	0	1	0							5798		1
12	11L		0	2	1	0	2	1	1	3	4	1	1	0	3	0	0	1							6877		3
13	12M		0	1	0	0	5	2	4	2	1	1	1	0	1	3	0	1							4703		5,5
14	13N		0	0	0	0	1	0	4	2	6	4	3	0	4	2	0	0							5224		5,1
15	14O		0	0	4	1	3	1	1	0	4	2	1	2	1	0	2	3							5690		4,5
16	15P		0	0	0	0	0	2	5	6	7	8	3	3	4	6	2	0							5165		10
17	16Q		1	0	0	0	2	1	6	6	3	2	1	0	0	1	0	0							4114		8,2
18	17R		0	0	0	0	1	0	6	8	4	3	0	2	0	3	2	1							4507		7,7

The wijks will be ranked in descending order based on the total number of deaths. You can rank the table in an ascending or descending order for any attributes you wish.¹

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	IC	Wijk185C	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Total	Population	Mortality Rate(/1k)	Total death
2	18S		0	2	2	1	5	1	6	10	13	9	8	8	14	2	7	5	5	2	0	0	100	10401		9,6
3	19T		0	0	1	1	4	0	2	3	2	7	7	2	16	11	11	7	1	2	0	0	77	11631		6,6
4	18S		0	0	0	2	15	8	7	15	5	2	4	1	0	3	1	0	0	1	0	4	68	3625		18,7
5	22W		0	0	0	0	3	5	7	9	12	3	4	1	2	1	2	4	0	1	0	0	54	10037		5,5
6	15P		0	0	0	0	0	2	5	6	7	8	3	3	4	6	2	0	1	2	2	0	51	5165		10
7	32GG		0	1	0	1	3	1	8	6	7	5	2	2	3	3	2	2	2	1	0	1	50	7037		7,1
8	21V		0	1	1	1	3	5	10	9	3	2	2	0	0	1	1	0	0	0	2	1	41	5290		7,7
9	17R		0	0	0	0	1	0	6	8	4	3	0	2	0	3	2	1	1	0	3	1	35	4507		7,7

If we rank the table by mortality rate, the top wijks change:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1	IC	Wijk185C	wk1	wk2	wk3	wk4	wk5	wk6	wk7	wk8	wk9	wk10	wk11	wk12	wk13	wk14	wk15	wk16	wk17	wk18	wk19	wk20	Total	Population	Mortality Rate(/1k)	Total death
2	18S		0	0	0	2	15	8	7	15	5	2	4	1	0	3	1	0	0	1	0	4	68	3625		18,7
3	50ZZ		0	0	0	1	2	3	2	2	2	2	1	0	0	0	0	0	0	0	0	0	15	1139		13
4	15P		0	0	0	0	0	2	5	6	7	8	3	3	4	6	2	0	1	2	2	0	51	5165		10
5	41QQ		0	2	2	1	5	1	6	10	13	9	8	8	14	2	7	5	5	2	0	0	100	10401		9,6
6	46VV		0	0	0	0	0	0	1	1	0	0	0	0	1	8	1	1	3	0	0	1	18	2252		8,4
7	16Q		1	0	0	0	2	1	6	6	3	2	1	0	0	1	0	0	1	7	2	1	34	4114		8,2

Comparing the two rankings, we see wijk S has the highest death rate but also ranks high in the absolute number of deaths while wijk ZZ has a high death rate but the absolute number of death is low.

Any other observations you come across?

¹ The current rank is without the wijk B-F, after filling in the table, the rank may subject to change

PART III: Creating your own Cholerakaart in QGIS

With all the calculation done in Excel, now we move on to QGIS and see if we can visualize the excel table in space.

Task 1: Install QGIS and introduction

1. Installation:

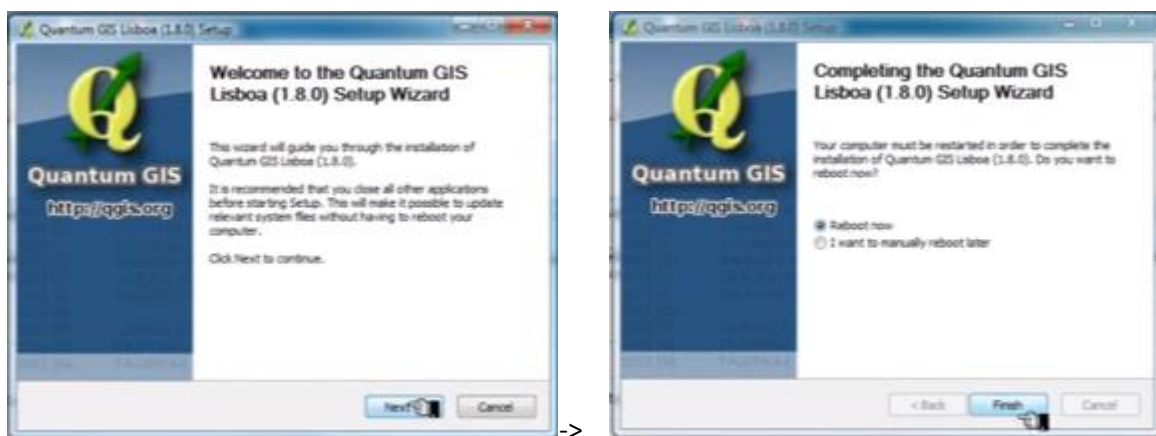
For Windows users:

Download and run the standalone installer, available here:

<http://qgis.org/en/site/forusers/download.html> (QGIS Standalone Installer Version 2.12)

When it gives you the option, you do not need to download the sample datasets, although feel free to and use them for other tutorials or examples.

Double click your installation package and follow the steps.



Video guide: <https://www.youtube.com/watch?v=loNPpcOb-sE>

For Mac OS users:

OS 10.7 and Higher: Follow the installation instructions provided here:

<http://www.kyngchaos.com/software/qgis>

OS 10.6.8: Try using the latest experimental build for Mac OS 10.6.8 provided here:

<http://qgis.dakotacarto.com/> (scroll down to and click on “Nightly for Snow Leopard - Mac OS X 10.6”).

Alternatively, try compiling with HomeBrew (<http://brew.sh/>), following instructions here:

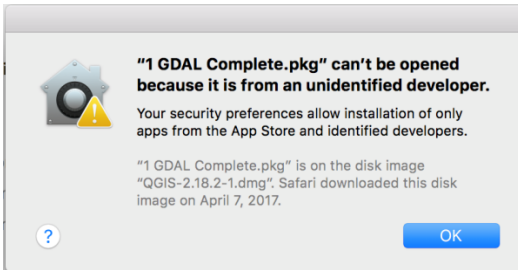
<https://github.com/OSGeo/homebrew-osgeo4mac> .

Notes for Mac OS users:

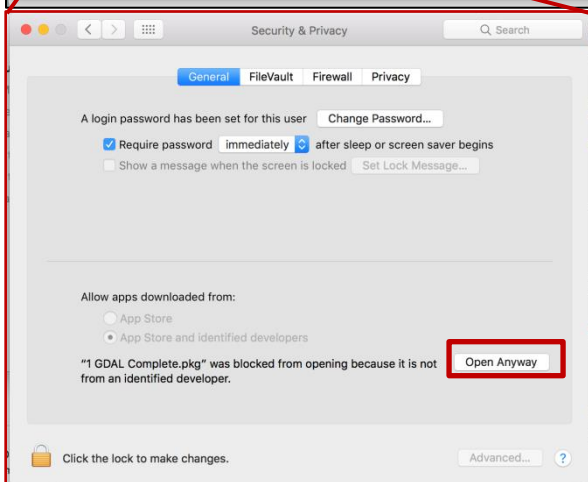
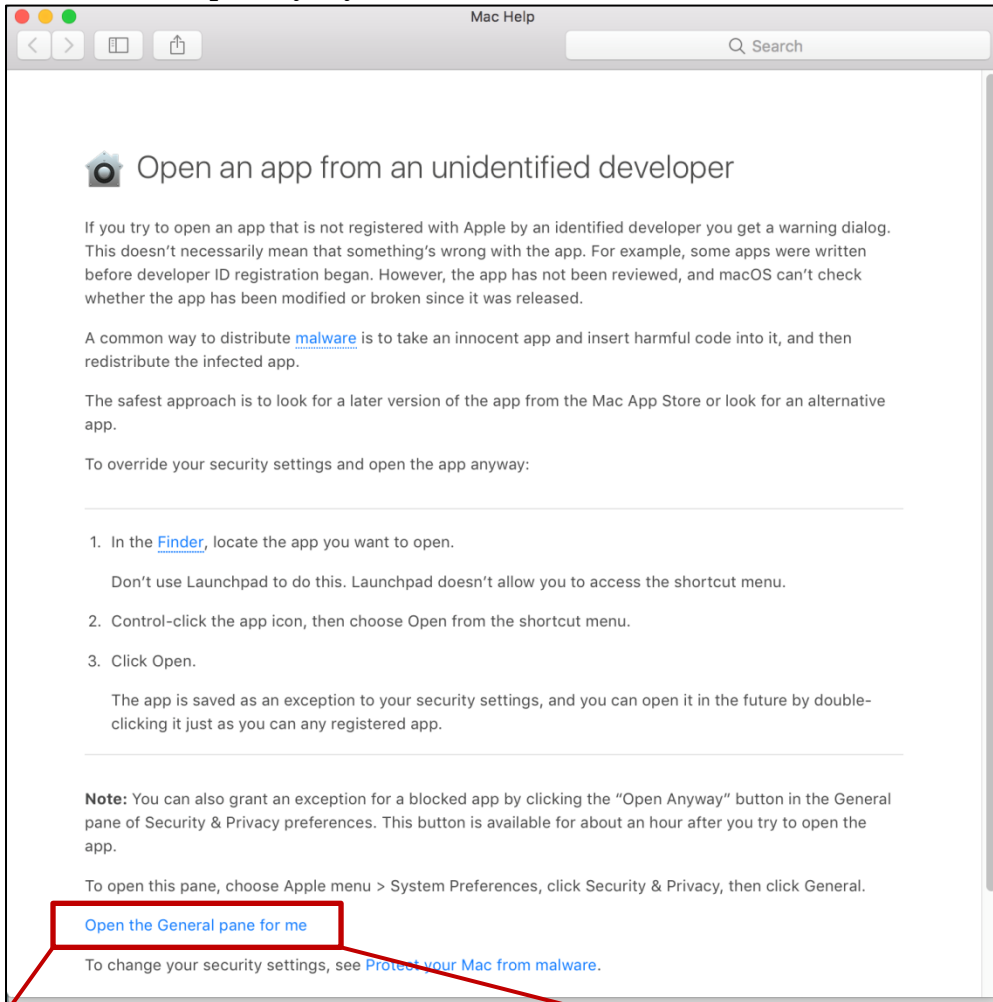
Unlike Windows, QGIS has 4 packages for Mac OS to install. They are numbered in the package your just downloaded, so please install following the sequence.

Video guide: <https://www.youtube.com/watch?v=MnXi5b2pqDo>

LET OP: Since QGIS is a third-party software (i.e., not from Apple store), you *may* need an extra step to authorize the installation. When you double click on the first package, you will see this window:




Click [OK], and this window will pop up, then click [Open the General panel for me], and in the following window, click [Open anyways]:

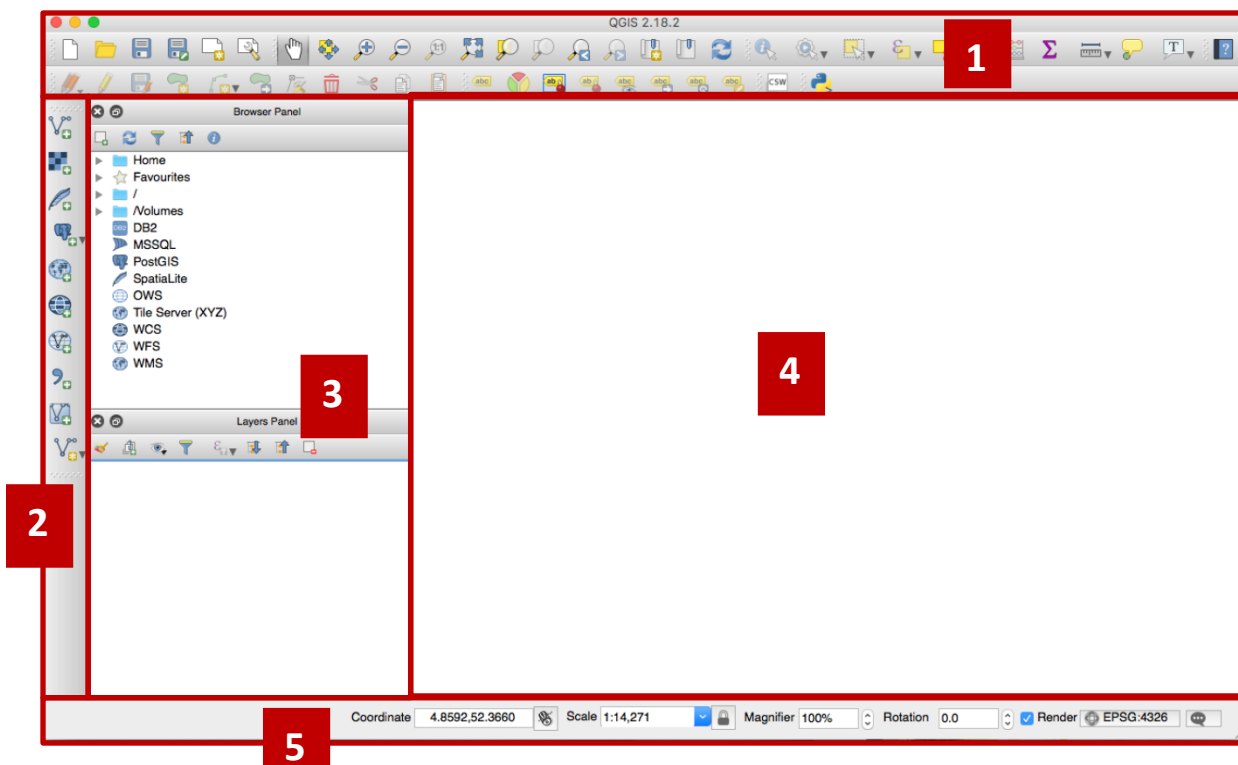


After you click [**Open anyways**], the new window will have a [Open] button. Click it and do the same for all four packages. Then follow the instructions to install.



After successfully install QGIS, you can open it by clicking: 

2. The basics of the QGIS user interface:




The elements identified in the figure above are:

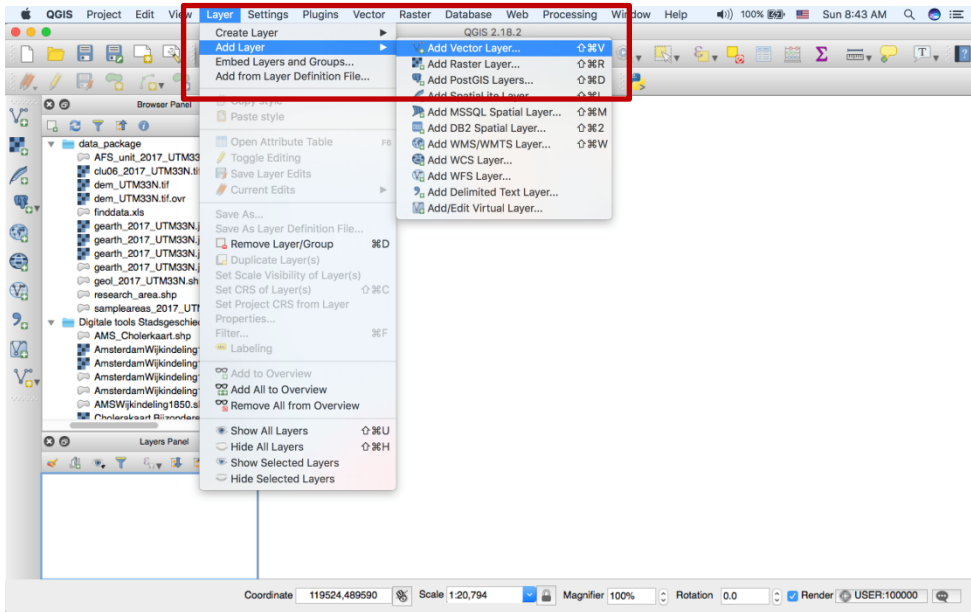
1. Toolbars
2. Side Toolbar
3. Layers List / Browser Panel
4. Map canvas
5. Status bar

You will use the sections for the next tasks.

In the following sections, Text that [**Looks like this**] directs you through menus and buttons.

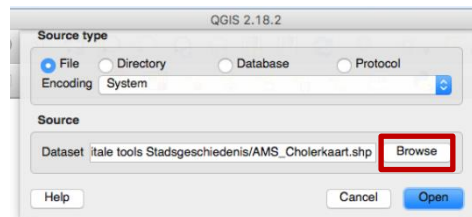
Task 2: Create a new map and load the shapefile

- Open QGIS and create a blank file (default)
- Then go to **[Layer] -> [Add Layer] ->  [Add Vector Layer]**




In the pop up window, click **[Browse]** and selecting a shapefile `AMS_Cholerkaart.shp` from the list and clicking **[Open]** loads it into QGIS.

(You only need to load **.shp** file, while the other files in the shapefile folder provide supporting information for **.shp** file.)

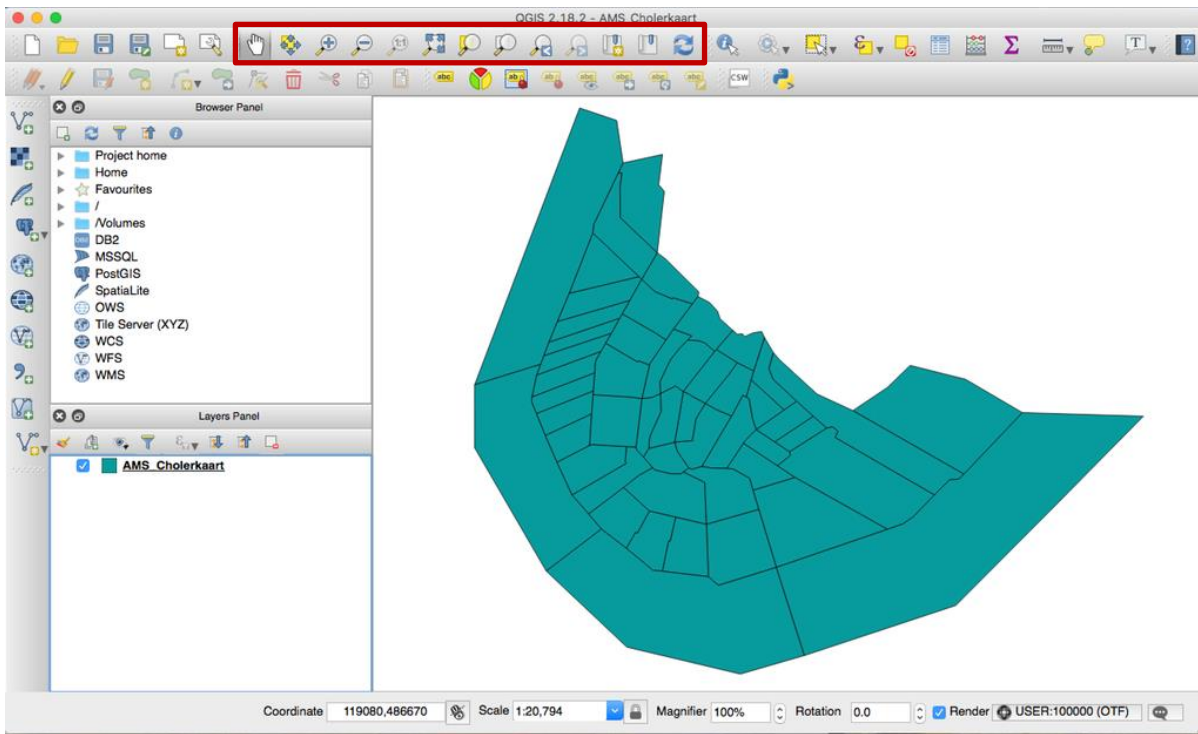



Congratulations! You now have a basic map. Now would be a good time to save your work.

- Click on the **[Save as]** button: 
- Save the map under the folder of your choice. Your map will be saved in `.qgs` format.

Once a shapefile is loaded, you can see the 50 districts of Amsterdam in simple, (random) single-color polygons. Now you can zoom around it using the map navigation tools (in the red bar area).

Tips: Apart from using the  pan and  zoom-in /  zoom-out icons on the toolbar with the mouse, navigating can also be done with the mouse wheel, spacebar and the arrow keys.



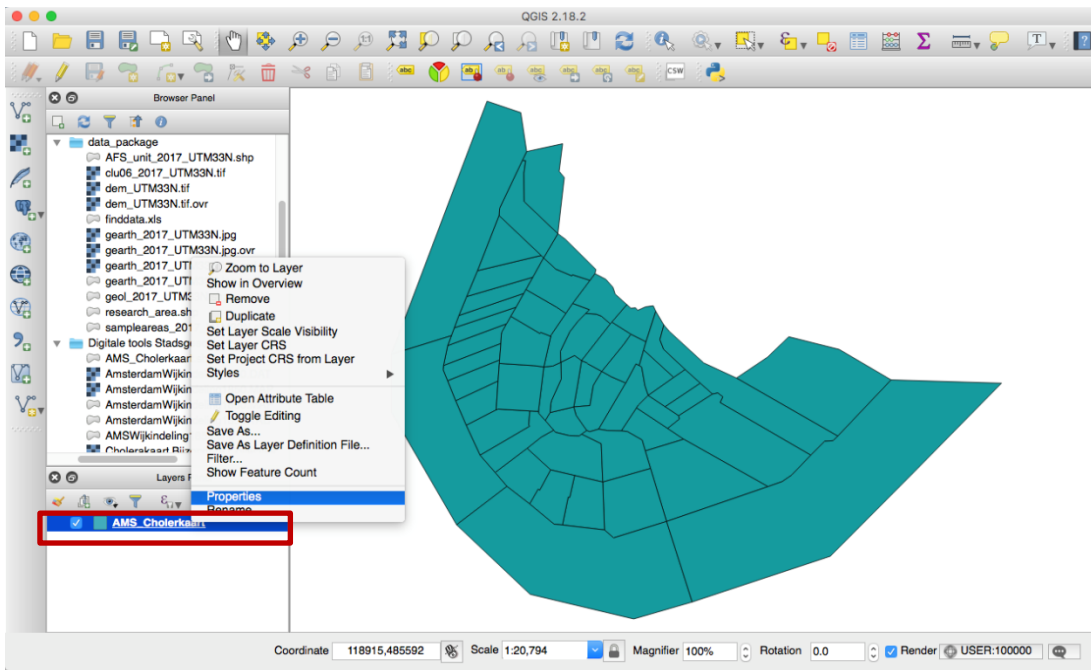
[Tip] in case you 'lost' the map (by zooming out too much or move it out of your sight), you can always click  [zoom full] to bring the shape file to your screen.

Task 3: Labeling the wijks

Labels can be added to a map to show any information about an object. Any vector layer can have labels associated with it. These labels rely on the attribute data of a layer for their content.

Now, go to the lower left side of the window, you will see the layer **AMS_Cholerkaart** is

- Double click the layer's name: **AMS_Cholerkaart**
- Or right click on **AMS_Cholerkaart**, then click on **[Properties]**

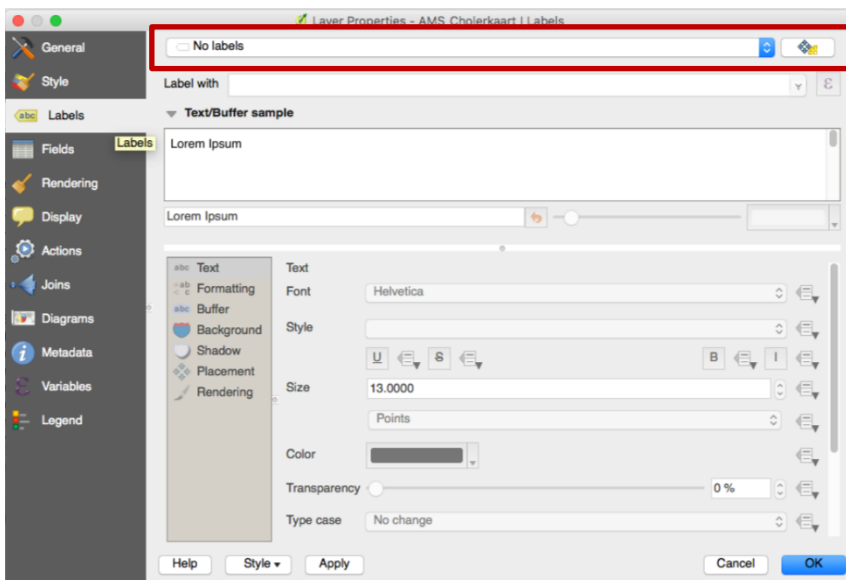


This gives you *Layer Properties* dialogue.

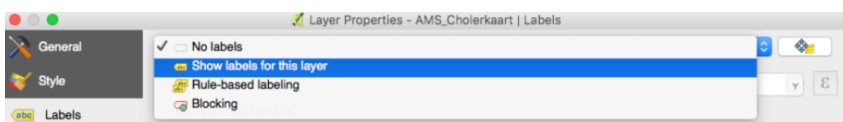
As we see from the map above, the Wijks are shown as simple polygons. We want to see the name of the zone appears on the map.

In the *Layer Properties* dialogue:

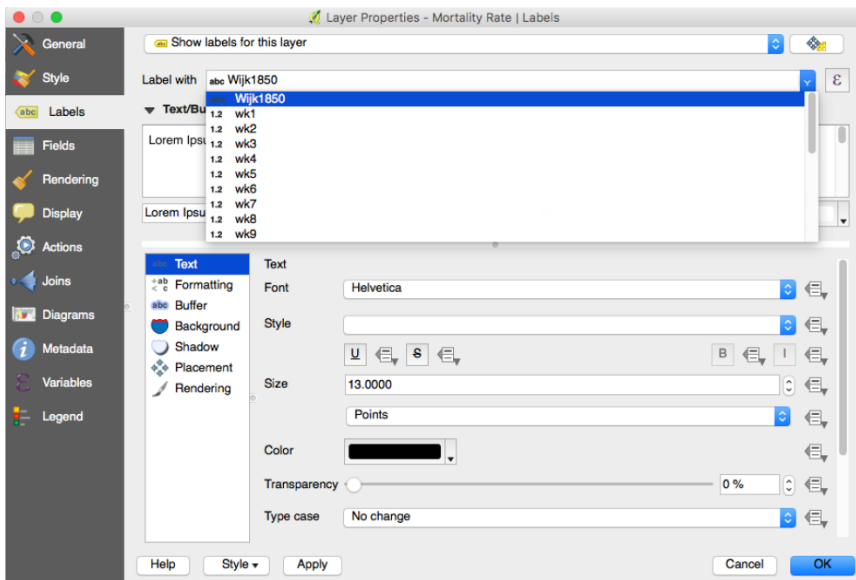
- Click the [Labels] on the left side bar



- In the stroll down menu above, choose [Show labels for this layer]



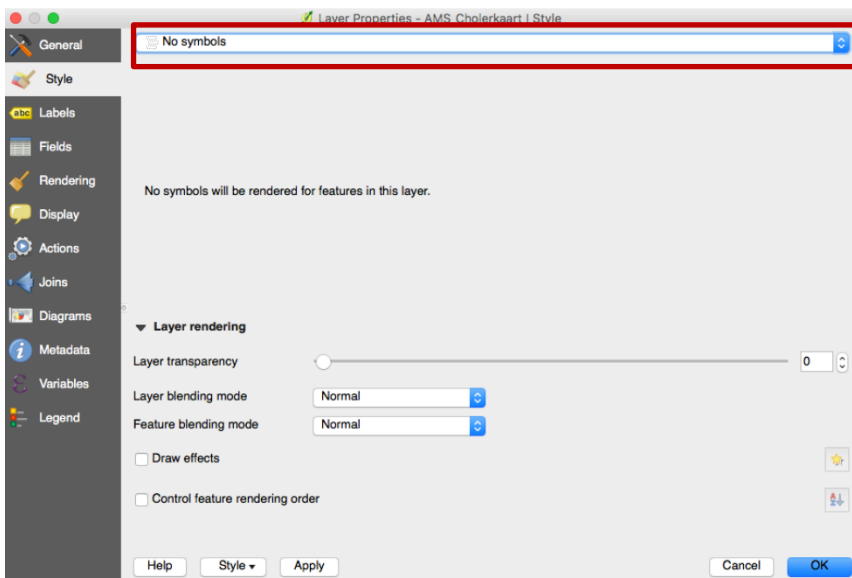
- In the [Label with] menu, stroll down and select [Wijk1850], then click [OK]



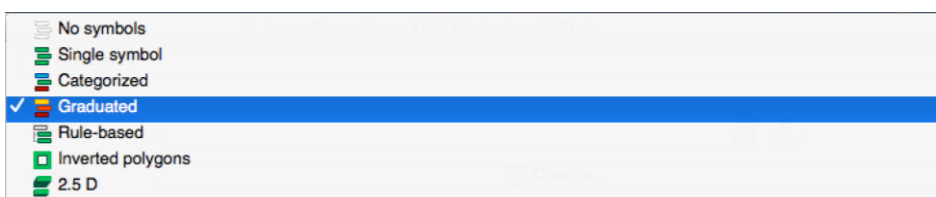
Then each zone in your map will be labeled with its Wijk name.

Task 4: Color-coded the wijks based on the total number of death

- Double click the layer's name: **AMS_Cholerkaart** (or right click on **AMS_Cholerkaart**, then click on **[Properties]**)
- Click on **[Style]**

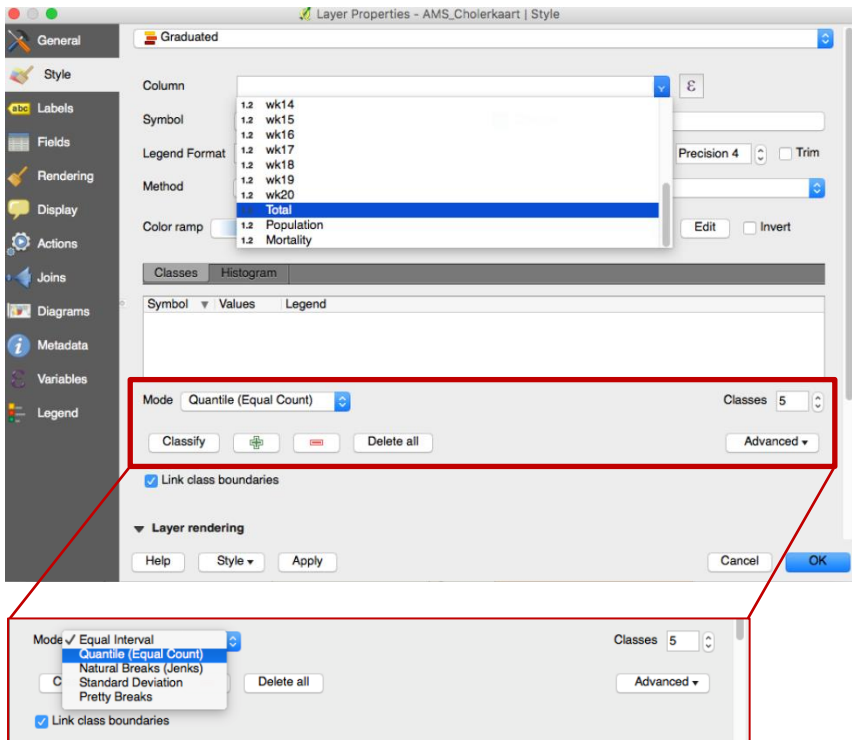


- Select **[Graduated]** from **[No symbols]** drop-down menu

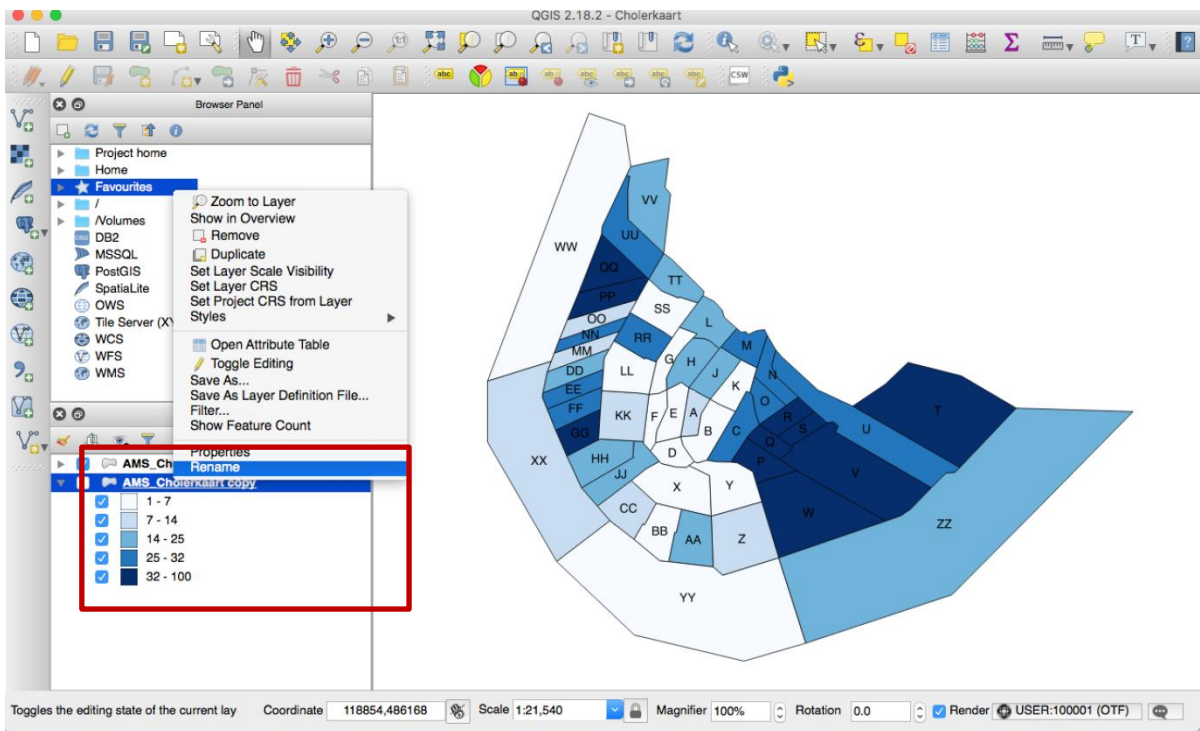


- Select **[Total]** (which is the total number of death we calculated in part 1) from **[Column]** drop-down menu – the **[Column]** option defines which attribute will be presented on the map.

- Then in the [Mode] menu, select [Quantile (Equal Count)] (You can also define your own categories in the [Advanced] options.)



- Click [OK], you will see the wijks are colored in different colors
- Right click on the layer name, then select [Rename] and change the layer name to 'Total Death'



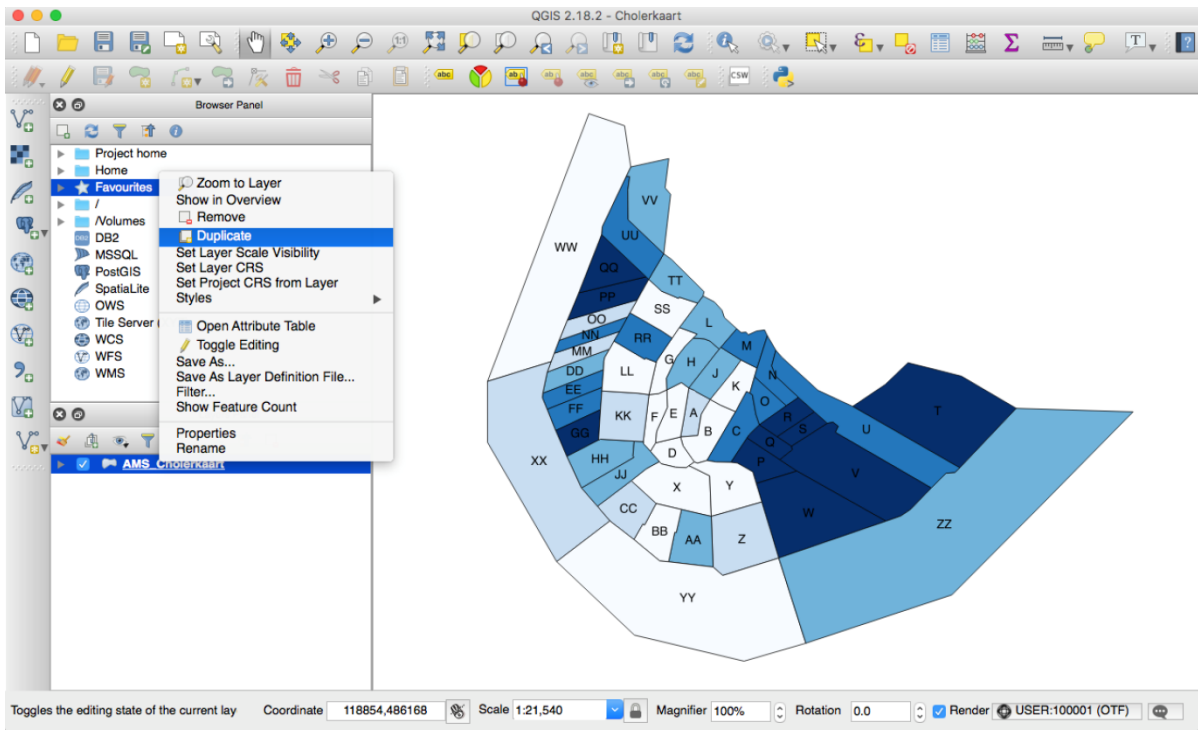
Question: What can you learn from this map?

Task 5: Add a new layer/ Copy a layer

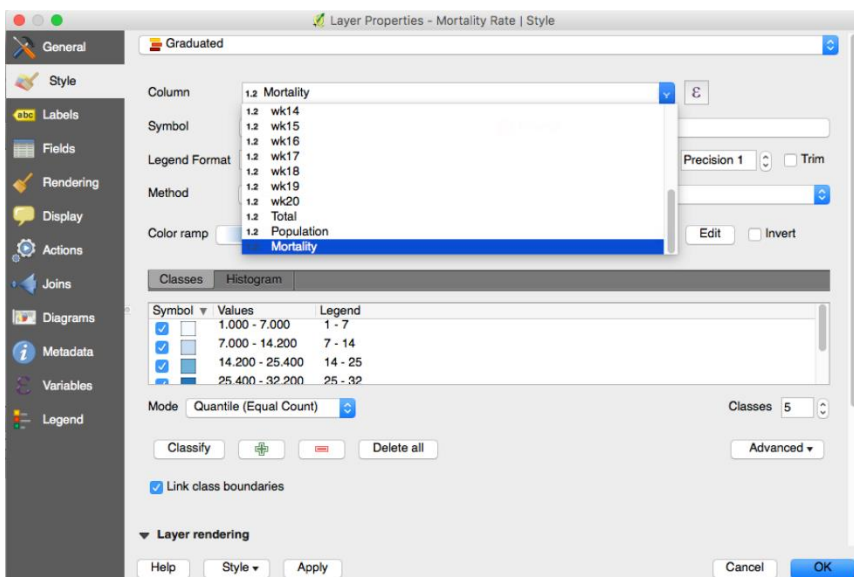
Now we have a map showing the total number of death. We also want to learn how the mortality rate varies from wijk to wijk.

Instead of creating another map file from scratch, we can simply duplicate the one we just made and change the display to a different attribute, in this case, the mortality rate.

- Right click on **AMS_Cholerkaart** (or **Total Death**, if you renamed it already), then click on **[Duplicate]**



- Right click on the layer name, then select **[Rename]** and change the layer name to ‘**Mortality Rate**’
- Double click the new layer’s name, then click on **[Properties]**

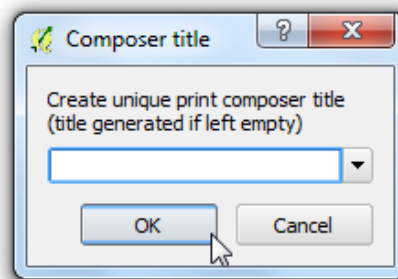


Again, in the *Layer Properties* dialogue, change the [Column] attribute to [Morality], meaning this new layer will present morality rate in different categories marked by different color. In order to distinguish this new layer from the old one, you may want to change the color from [Color ramp] menu.

Question: What does your new map look like? Is it very different from the total death map? What can it tell you?

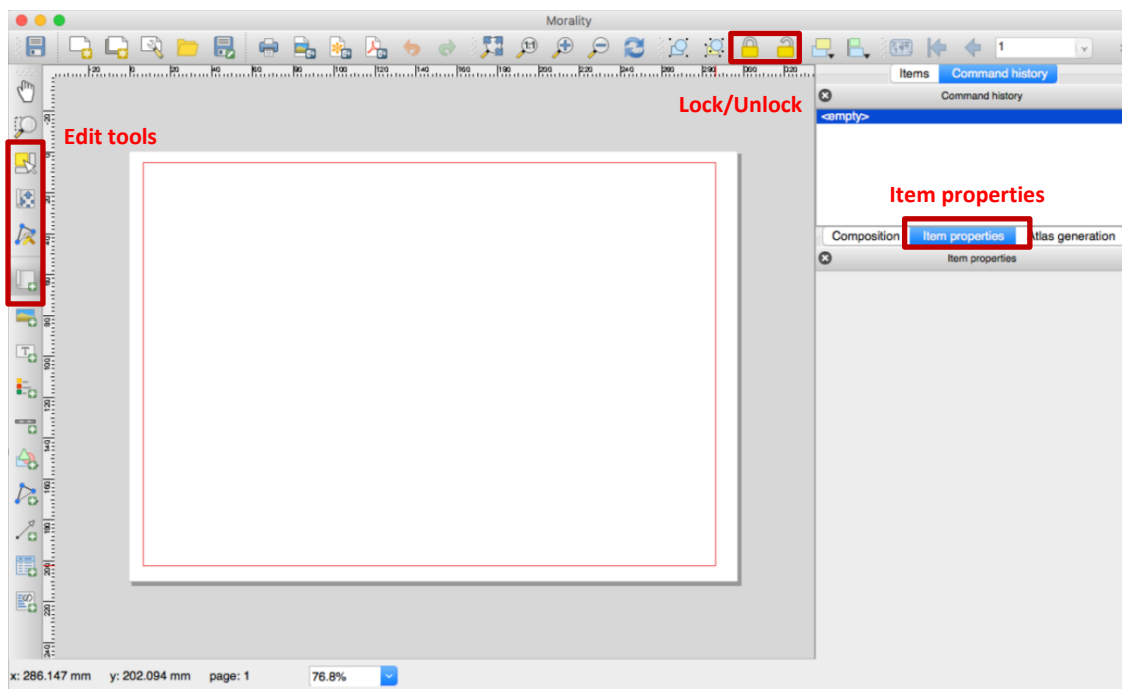
Task 6: Export a map

Before you export a map or to start to work with the *Print Composer*, you need to load the QGIS map canvas and adapt their properties to suit your own convenience. After everything is rendered and symbolized to your liking, click the  icon in the toolbar or choose [File]>[New Print Composer]. You will be prompted to choose a title for the new Composer.





Add a map area



Click on the  toolbar button in the Print Composer toolbar to add the QGIS map canvas. Now, drag a rectangle onto the Composer canvas with the left mouse button to add the map.




Edit a map area

You can resize the map item by clicking on the  Select/Move item button, selecting the element, and dragging one of the blue handles in the corner of the map. This button also helps to move the map to another place. Select the item and while holding the left mouse button, move to the new place and release the mouse button.

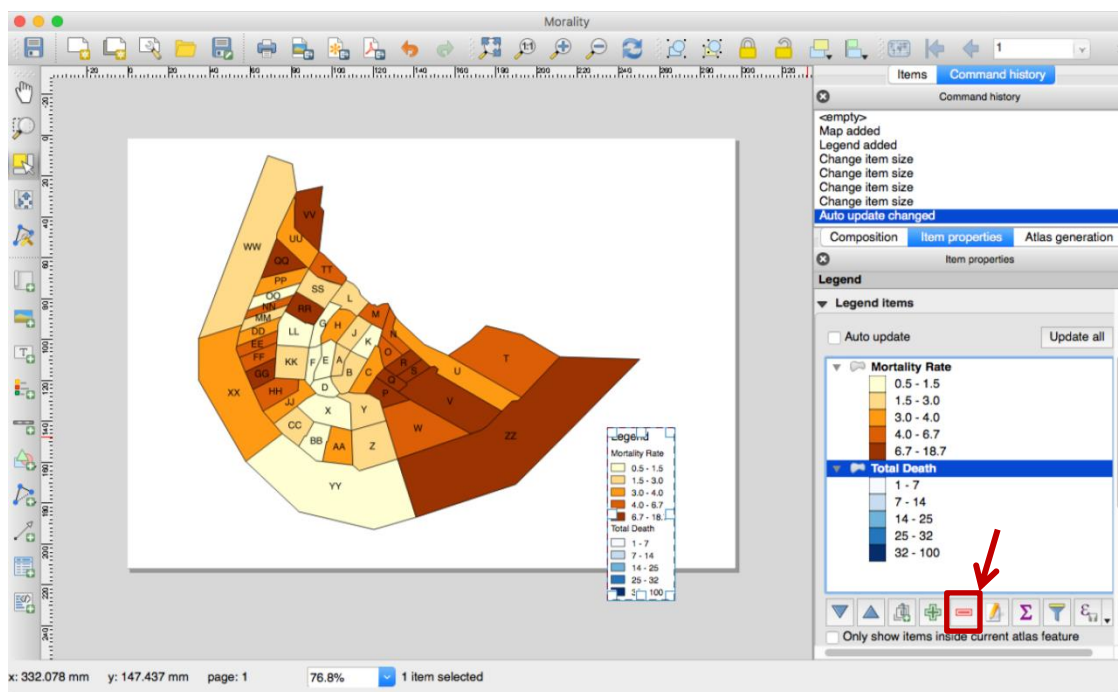
To move layers within the map element, select the map element, click the  Move item content icon and move the layers within the map item frame with the left mouse button.

[Tip] After you have found the right place for an item, you can lock the item position within the Print Composer canvas. Select the map item and use the toolbar  Lock Selected Items (on the top of this Map Composer window). A locked item can only be selected using the **[Items]** tab. Once selected you can use the **[Items]** tab to unlock individual items. The  Unlock All Items icon will unlock all locked composer items. With the map selected, you can now adapt more properties in the map **[Item properties]** tab.

Add a legend


To add a map legend, click the  Add new legend icon, place the element with the left mouse button on the Print Composer canvas and drag to position. To customize the appearance in the legend **[Item properties]** tab on the right side of the window.

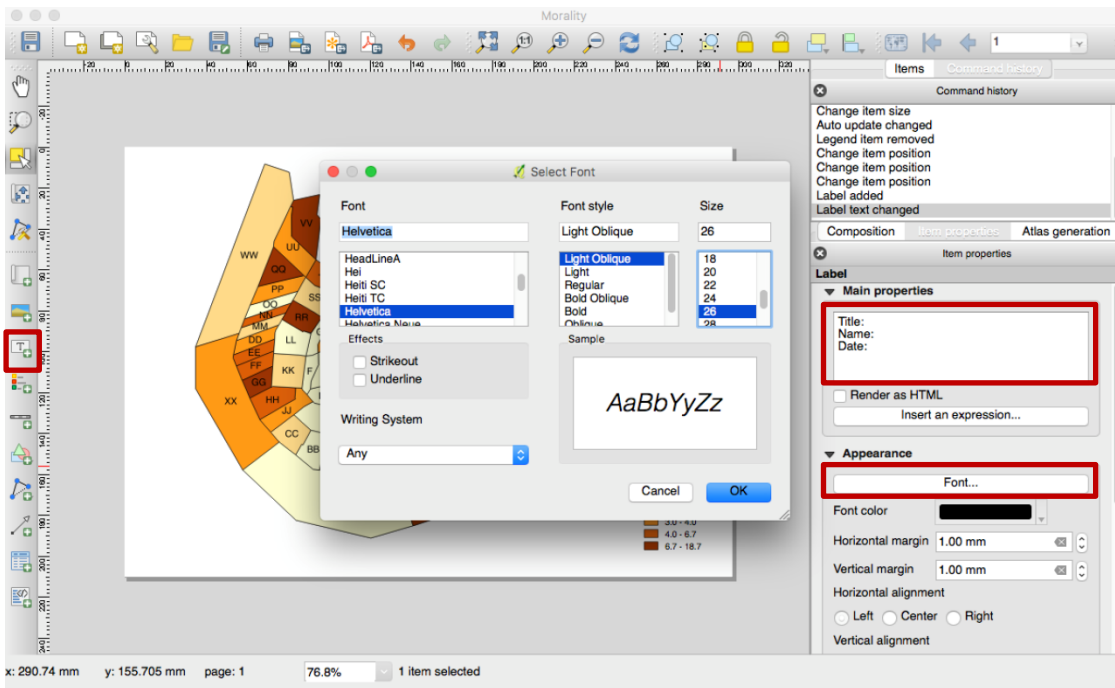
Because we are only showing the morality rate in this map, we use the **[minus]** button to remove the *Total Death* layer.



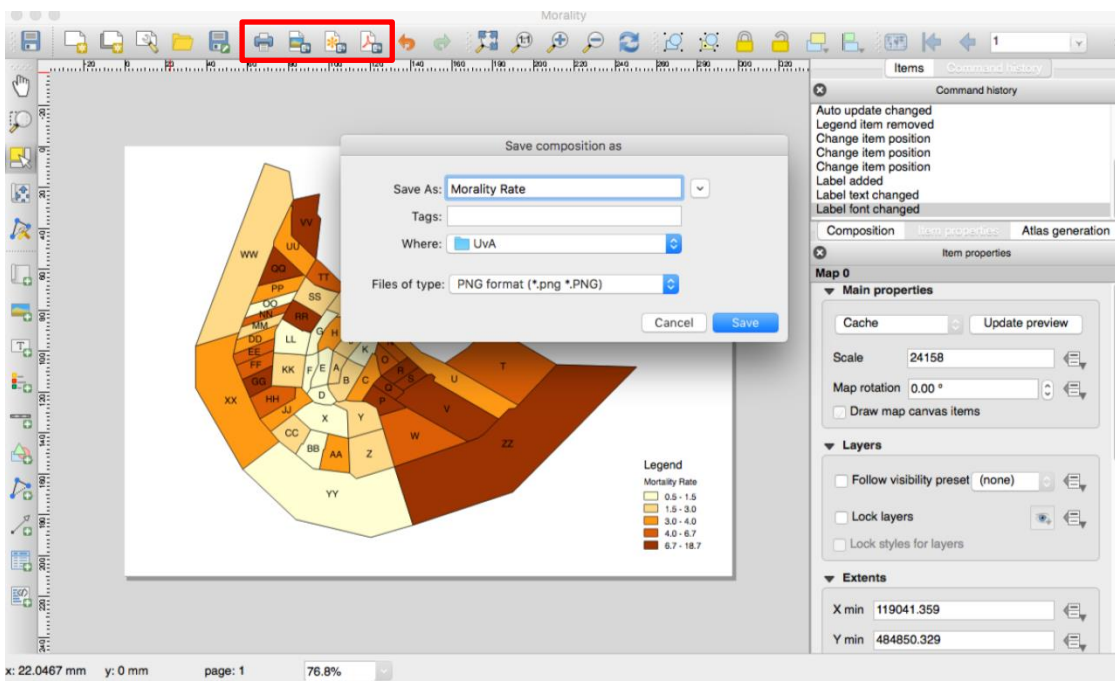
Add a title

It is like adding a text box in a PowerPoint.

- Click on  add text in the sidebar
- Type in: Title of your map, your name and the date into the box on the right
- You can also change the font size by clicking on **[Font...]** button on the right




Export your map:



- Select on type for export:

The  Print icon allows you to print the layout to a connected printer or a PostScript file, depending on installed printer drivers.

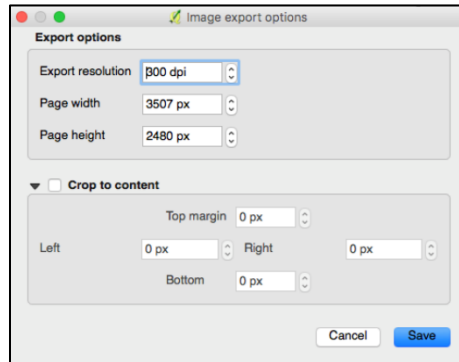
The  Export as image icon exports the Composer canvas in several image formats, such as PNG, BPM, TIF, JPG,...

The  Export as SVG icon saves the Print Composer canvas as an SVG (Scalable Vector Graphic).

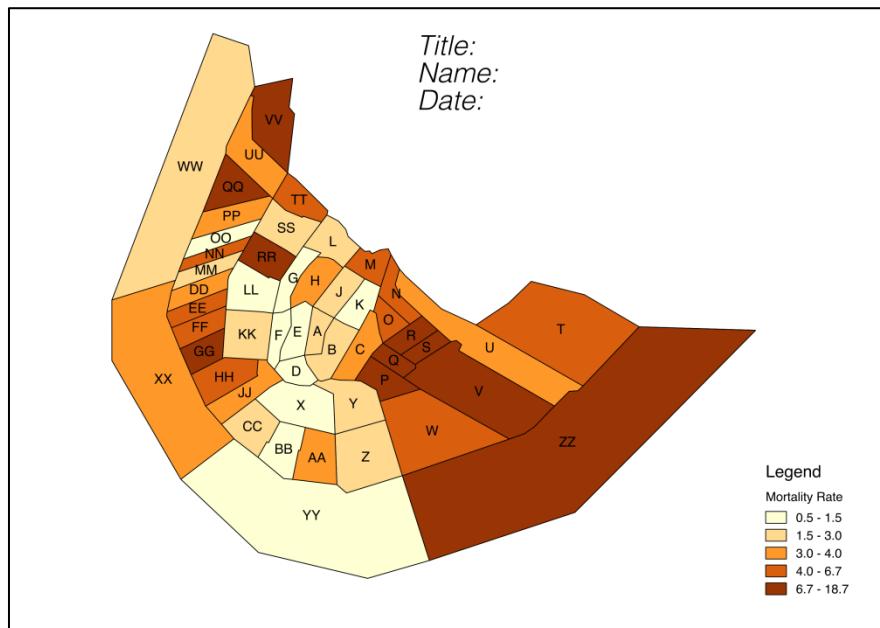
The  Export as PDF icon saves the defined Print Composer canvas directly as a PDF.

(NOTE: the  Save Project icon can only save the features of your design in QGIS, but not to export as a map.)

- Enter the name of you map. You can also select the File of type menu for your desired file type (.jpg, .png, etc.)
- Click **[Save]**, then in the pop up window you can change other options, the default is usually preferred. Then click on **[Save]** to finish export.



Your final map may look like the following:



Task 7: Your own question and solution

With all the tools you have learned (labeling the polygons, color-coded the wijks based on the information you wanted to present and to select certain attributes based on the criteria you are interested.) What question(s) that you are interested? Can you use the tools to answer your own question?