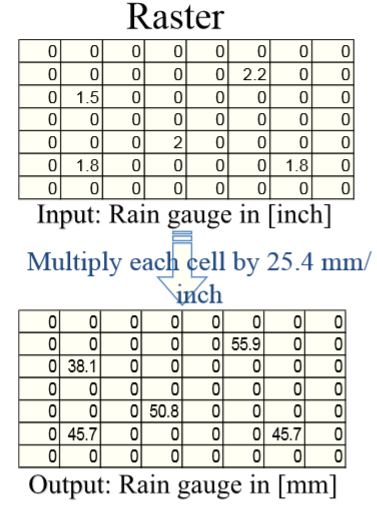
Martin Candee and Ran Wang

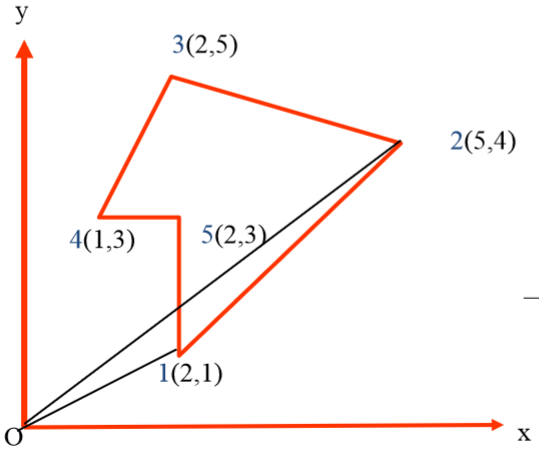
## Data Anlysis:

When talking about data analysis in GIS, one of the most important aspects of it is map overlay. Map overlay is the idea of creating new output features on a map by combining the features of two or more input layers. If there is more than two layers being combined together, then it is done step by step, meaning the first two input layers are combined to create an output layer, which is then combined with the third layer, and so on until you achieve the final layer. It is important when combining layers though, that they are done to the same scale and in the same projection. The two basic types of map overlay are arithmetic and logical, which can both be done using vector or raster data. For example, the image below shows an example of attribute combination of raster data using the arithmetic approach. The top table is shown in inches, and in order to achieve the final bottom table represented in millimeters, each cell must be multiplied by 25.4 mm. This shows how a simple arithmetic map overlay can convert and combine different layers of raster data.



However, the vector data map overlay is often much more complex than raster systems. This is because there is a multitude of possible overlay combinations that can be done, and new points and lines must be created when this is done. There are points, lines and polygons, and each one can be overlayed on one another. For example, a point can be combined with a polygon, will have the output of a point in polygon element. The process to get this output is quite complex, and first involves establishing the minimum boundary rectangle and the points that are outside of it, which can be derived from the half-line algorithm. However, a simple way to define the points in a polygon is that if the number of times a point intersects the polygon is odd, then it lies within the polygon.

There are many other aspects to data analysis that can be done in GIS. The area of a polygon can be determined by calculating many small triangles by placing the polygon on an XY axis and referring each small triangle back to the (0,0) point. The example below shows how different triangle areas are calculated. When these are determined, you can add these together and then subtract the area outside of the polygon to get it's area.



Other things such as buffers can be a useful tool that involves data analysis. Line intersection are also critical operations that are used in GIS. It is used not only in polygon overlay, but also when dissolving and merging both lines and polygons. When determining if lines intersect, first the equations of both lines must be calculated. The parametric equation of a line can be used, and then an intersection test must be performed. Then you can determine if the lines intersect, and where exactly they do. Line simplification is another important process of data analysis and can be using in reducing storage, plotting time, and creating better resolution of a line when you display the map on a smaller scale. The main two types of processing routines used in line simplification are local and global. Local processing is simple, easy and fast to do, but the disadvantage is that it doesn't account for extreme complexity in some lines, like global processing does. The Douglas-Peucker algorithm can be used as a global routine to determine which points on a line are not necessary.

Lastly, merge and dissolve are two data analysis operations that are extremely important for topography and things like road maps. For example, if you have two different data sets but want to connect them, then you want to make sure that the roads and building along the edge of each data set line up with the other. Edge matching can be used so that roads will appear as continuous when going from one data tile to another. Dissolve can be equally important when you want to remove boundaries between polygons. It is clear to see that data analysis is an important aspect of GIS programs.

## Lab

This small project involves many knowledge regarding data analysis in terms of GIS. As we are supposed to inspect all the roads and building roofs on the campus, we can reuse the database Building and Road from previous assignment. The task asks us to classify the roads to three conditions with regards to how much it is worn: received too much salt, damaged by sown plow and in good condition. We add one more Field called “level” in “Character” Type to the Table Road, which includes all the three conditions (See figure 1). We do the same to the Table Building by adding a new Field “roofs” that covers two types of roof installation- sun terrace and greenhouse (See figure 2)

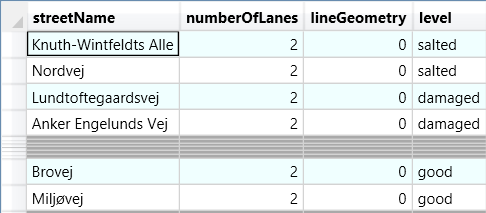


Figure Modified Road Table

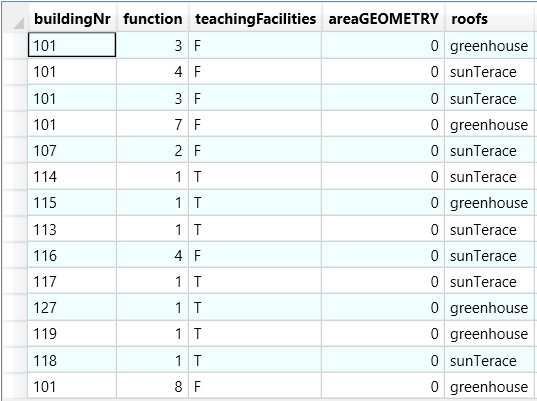


Figure Modified Building Table

Once the tables are done, we can make thematic maps to illustrate the new features of Road and Building. For instance, we color the road condition with three colors, and we use two intuitive patterns to mimic greenhouse and sun terrace (See figure 3).

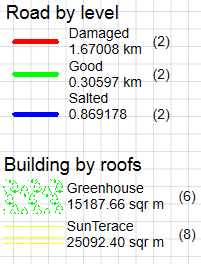


Figure Part of legend with different themes

The road length and roof area must be the interesting information for the administration to estimate the cost for repairment and installation. We use the SQL to calculate the total length of road and total area of roof for each classifications. The algorithms are made in SQL, so that we can easily make operations on the data we have (See figure 4, 5).

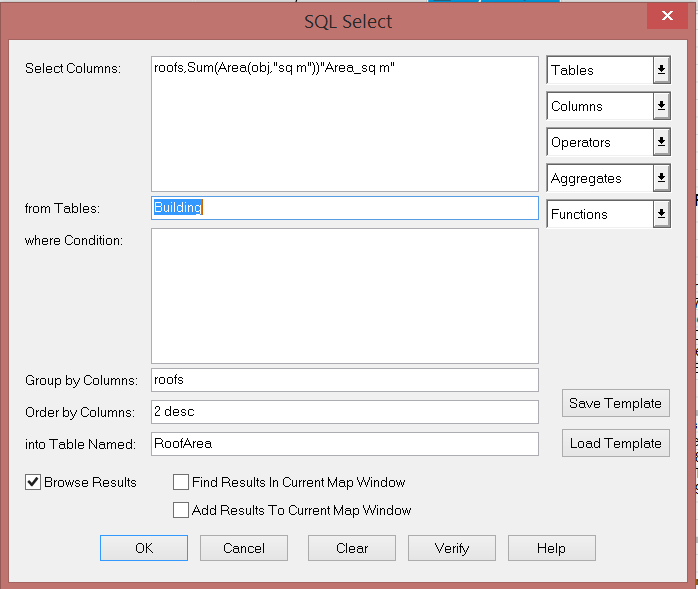


Figure SQL for roof area

Moreover, we make the orthophoto less visible in favor of the visibility of the roads and buildings. We create a new layout to include the map, legend, title, scale bar and north arrow and save it as a png file (See figure 6).

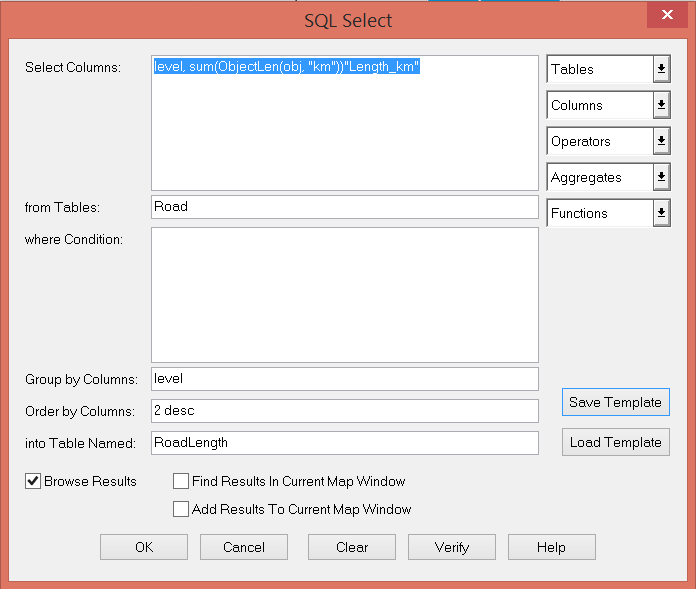


Figure SQL for road length

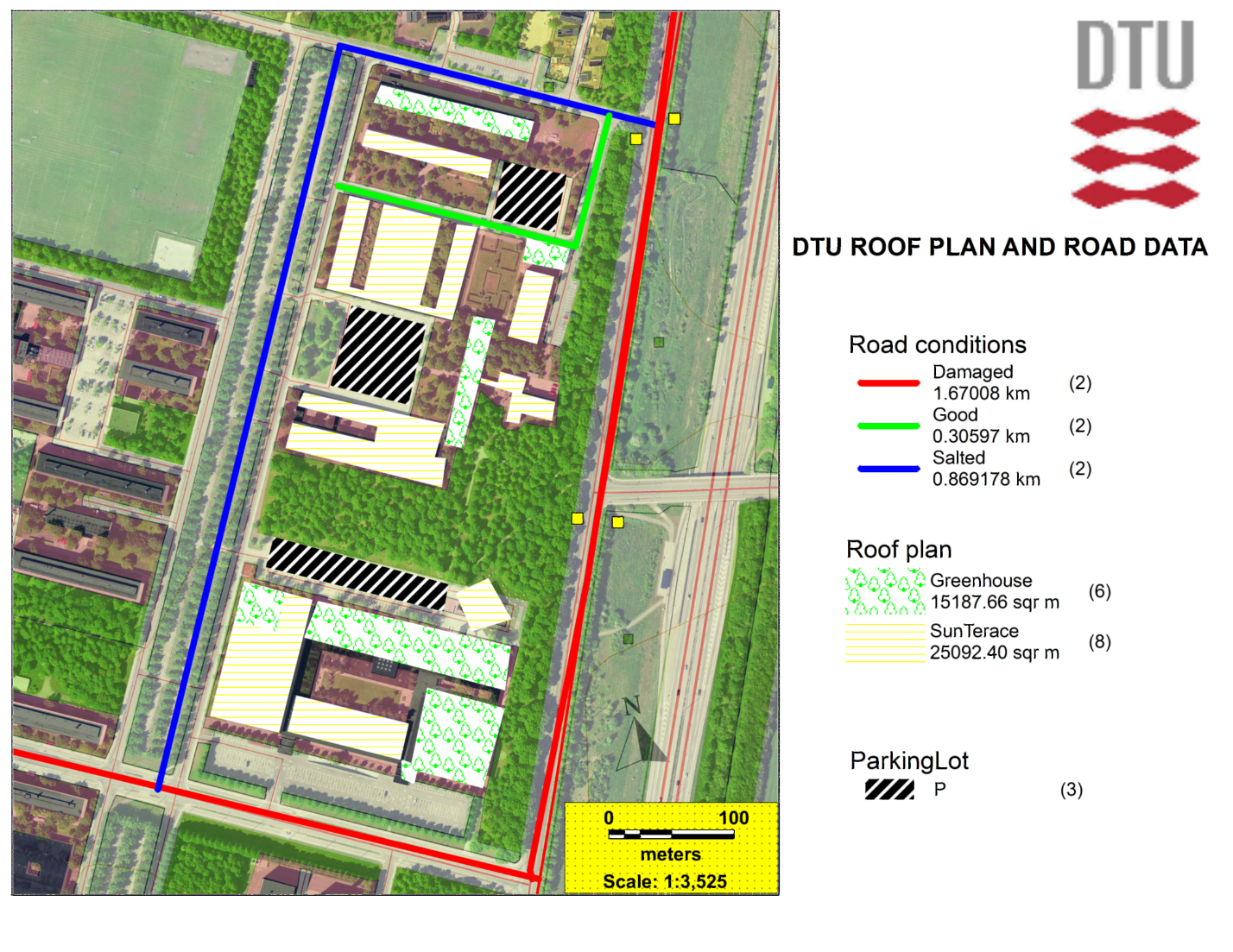


Figure Final map

## Conclusion

This assignment gives a good example of how the data can be used in practice administrative affairs. During this assignment we have practiced a lot on data analysis including data reconstruction, algorithms in SQL, themes reflecting features, which equipped us with a better understanding about data analysis in both theory and approaches (tools) in Mapinfo. The SQL is a strong tool to make many operations on the data with regards to the purpose. So far, we only used it to calculate line length and polygon area, we will be glad to discover more functions of it in the coming assignment.