20. GIS Mapping

Dr. Shelton Padua

Marine Biodiversity and Environment Management Division
ICAR-Central Marine Fisheries Research Institute, Kochi

GIS is a tool for making maps. GIS helps us in using spatial information and it is mainly concerned with the location of the features as well as properties/attributes of those features. It helps us gather, analyse and visualize spatial data for different purposes. A GIS quantifies the locations of features by recording their coordinates which are the numbers that describe the position of these features on Earth. The uniqueness of GIS is its ability to do spatial analysis. GIS helps us analyse spatial relationships and interactions. Sometimes, GIS proves to be the only way to solve spatially-related problems and it is one of the most important tools that aid in the decision making process. GIS basically helps to answer three questions; How much of what is where? What is the shape and extent of it? Has it changed over time?

Globally, on an average, GIS tools save billions of dollars annually in the delivery of goods and services through proper route planning. GIS regularly helps in the day-to-day management of many natural and man-made resources, including sewer, water, power, and transportation networks. GIS helps us identify and address environmental problems by providing crucial information on where problems occur and who are affected by them. It also helps us identify the source, location and extent of adverse environmental impacts. GIS enables us to devise practical plans for monitoring, managing, and mitigating environmental damage. Human impacts on the environment, conflicts in resource use, concerns about pollution, and precautions to protect public health have spurred a strong societal push for the adoption of GIS.

GIS is composed of hardware, software, data, humans and a set of organizational protocols. The selection and purchase of hardware and software is often the easiest and quickest step in the development of a GIS. Data collection and organization, personnel development and the establishment of protocols for GIS use are often more difficult and time consuming endeavours. A fast computer, large data storage capacities and a high quality, large display form the hardware foundation of most GIS. GIS software provides the tools to manage, analyse, and effectively display and disseminate spatial information. GIS as a technology is based on geographic information science and is supported by disciplines like geography, surveying, engineering, space science, computer science, cartography, statistics etc.

In GIS, we handle the spatial and attribute data sets. Spatial data describes the absolute and relative location of geographic features while attribute data describes characteristics of the spatial features. These characteristics can be quantitative and/or qualitative in nature. Attribute data is also referred to as tabular data. Vector and raster are two different ways of representing spatial data. Raster data is made up of pixels (or cells), and each pixel has an associated value. A digital photograph is a simple example of a raster dataset where each pixel value corresponds to a particular colour. In GIS, the pixel values may represent elevation above/below sea level, or chemical concentrations, or rainfall etc. The key point is that all of this data is represented as a grid of (usually square) cells. Vector data consists of points, lines, and polygons. The individual points are stored as pairs of (x, y) co-ordinates. The points may be joined in a particular order to create lines, or joined into closed rings to create polygons, but all vector data fundamentally consists of lists of co-ordinates that define vertices, together with rules to determine whether and how those vertices are joined.

As with many other systems, GIS basically works on the principle of '*GIGO'* that is *garbage in garbage out*. Hence the quality of data that you feed into GIS is very important and it determines the quality of the end products. But, when used wisely, GIS can help us live healthier, wealthier, and safer lives.

Generating continuous surface data from point observations using Spatial Interpolation

There are over 40 spatial interpolation techniques available and the factors like sample size, sampling design and data properties affect the output of these methods. It has been reported that when the variation within the data increases, the accuracy of all the interpolation methods decreases. But in such cases, stratification and non-uniform sampling design is found to improve the accuracy of the results in case of environmental variables. (Li and Heap, 2008).

Here we will see two of the most common methods of interpolation namely IDWA and Spline.

IDWA

IDWA (Inverse Distance Weighted Average) is one of the simplest interpolation methods. In this method, the value for the un-sampled location is obtained by weighted averaging of the values of sampled locations in its neighborhood. The weights for the nearby sampled points that contribute to the prediction are determined by its distance from the un-sampled point. Weights determine how much influence a sampled point has on the un-sampled point. The basic assumption is that the sampled locations that are closer to the un-sampled location provide more representation of the value at the un-sampled location than the farther sampled locations. To put it in a simple way, there are two sampled points (A & B) and an un-sampled point X. A is 5 m away from X and B is 10 m away from X. In this case, the value of X will be more similar to the value of A. The weights depend on linear distances between the sampled and un-sampled locations and the spatial distribution of the observation points does not have any role in determining the weights. The weights assigned to each known points in the averaging process are proportional to the inverse of the distance between the sampled and the unknown (predicted) point (Hartkamp et al., 1999). That's why the method is called inverse distance weighted average.

The equation for IDWA is as follows:

$y(x) = \lambda i\ y(xi)$

Where:

y(x) = the predicted value of the variable at un-sampled location.

λi = the weights for the individual locations.

y(xi) = the values of variable evaluated in the sampled locations.

The weight λi is given by the equation:

λi= 1/dipi=1n1/dip

Where *d* is the distance between the un-sampled and sampled location, *p* is the power parameter, and *n* is the number of sampled points used for the estimation (neighborhood points). The main factor affecting the accuracy of IDWA is the value of the power parameter (Isaaks and Srivastava, 1989).

The power parameter *p* controls the influence of sampled locations on the un-sampled locations based on their distance. This parameter is a positive real number. A higher power will give more emphasis on the nearest sampled locations and hence, at higher powers, the interpolated value assumes the value of the closest sampled point and will give less smooth surface. We can say that IDWA behaves like nearest neighbor interpolation at higher powers. A lower value for power will give more influence to neighborhood points that are farther away, resulting in a smoother surface (Fig.2). In general, a power value of 2-5 is found to work satisfactorily. Increasing the number of sample points included in the interpolation by increasing the search radius will also result in a smooth surface. The neighborhood points (sampled points) for a particular un-sampled location is determined by the radial distance from the un-sampled location and is called search radius. Search radius size is provided as an input during the operation.

The most common problem with IDWA is that it may produce bull's eyes on sampled positions which can be seen in the Fig. 2. Otherwise, it's an easy method and helps to understand the influence of distance on the interpolated parameter. The IDWA works well when the sample points are regularly spaced (Isaaks and Srivastava, 1989). In IDWA, the interpolated values are never outside the range of sampled values.

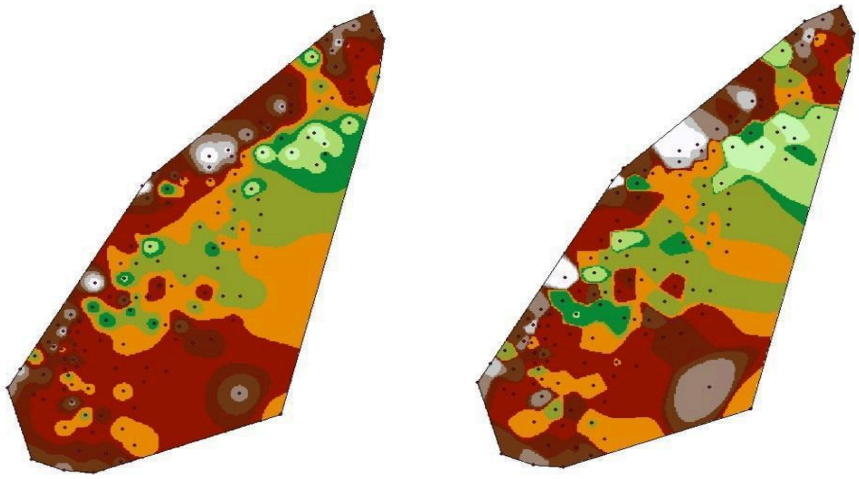


Fig. 2 Comparison of IDWA interpolation with different power parameters i.e. 2 and 10

Spline

It is a deterministic interpolation method. The spline interpolation tries to create a flexible surface passing through all sampled locations like a rubber sheet touching all the sampled locations and the value for un-sampled locations are extracted from this surface. Spline uses slope calculations (rate of change of value of the particular variable with the distance) to estimate the shape of the flexible rubber sheet. Because of this stretching, spline interpolator makes estimates that are outside the range of sampled values.

The splines consist of polynomial functions fitted through the sampled points and the polynomials are of degree p. Polynomials are fitted to a small number of data points exactly and describe the pieces of surfaces. The pieces of surfaces thus generated are smoothly joined together and form the entire surface. The places where the pieces join are called knots. If the degree of the polynomials is 1, it is called linear spline, if the degree is 2, then it is called quadratic spline and if the degree is 3, then it is called cubic spline (Burrough and McDonnell, 1998; Webster and Oliver, 2001).

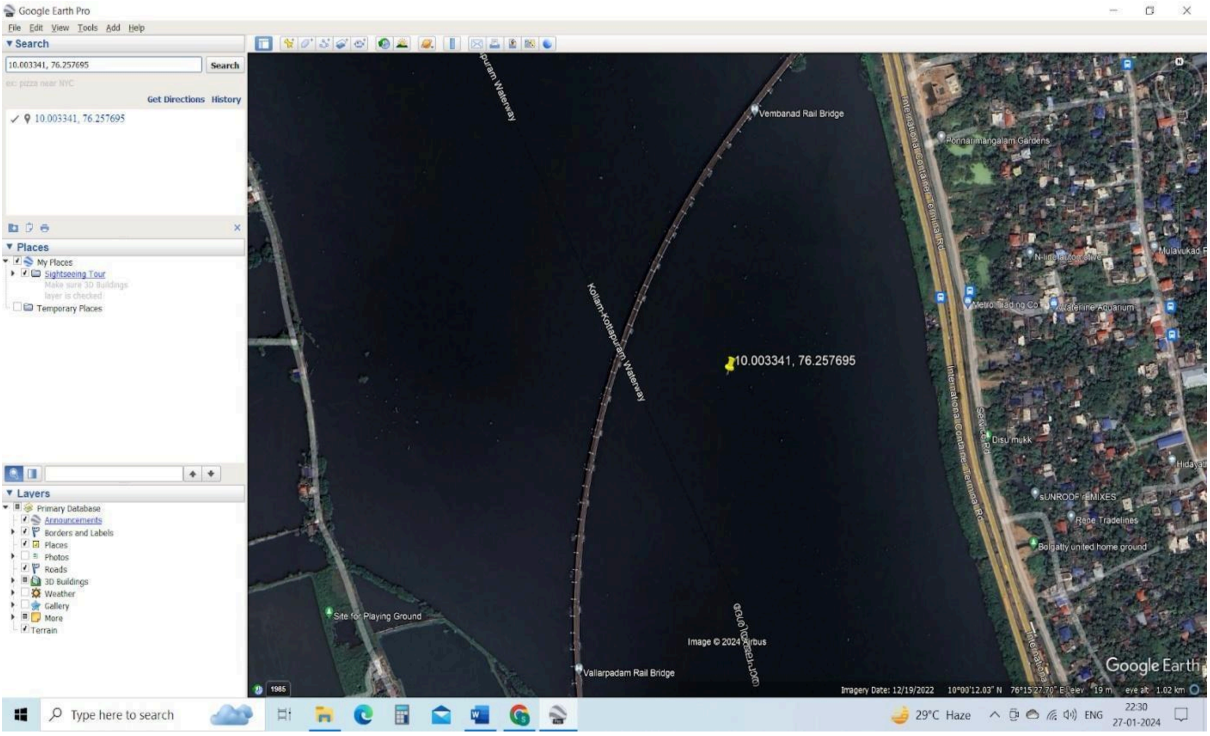
Spline is known to give unsatisfactory results in cases where the samples are closer and the range is very high. The spatial distribution of the samples also may affect its performance. Splines perform better when the samples are regularly spaced and dense. Irregularly spaced samples may reduce the performance of spline interpolation (Collins and Bolstad, 1996).

Thin plate splines (TPS) also known as Laplacian smoothing splines are more robust as it uses a generalized cross validation function to determine the smoothing parameter. It has more predictive accuracy and is less influenced by the statistical assumptions about the data (Hutchinson, 1995).

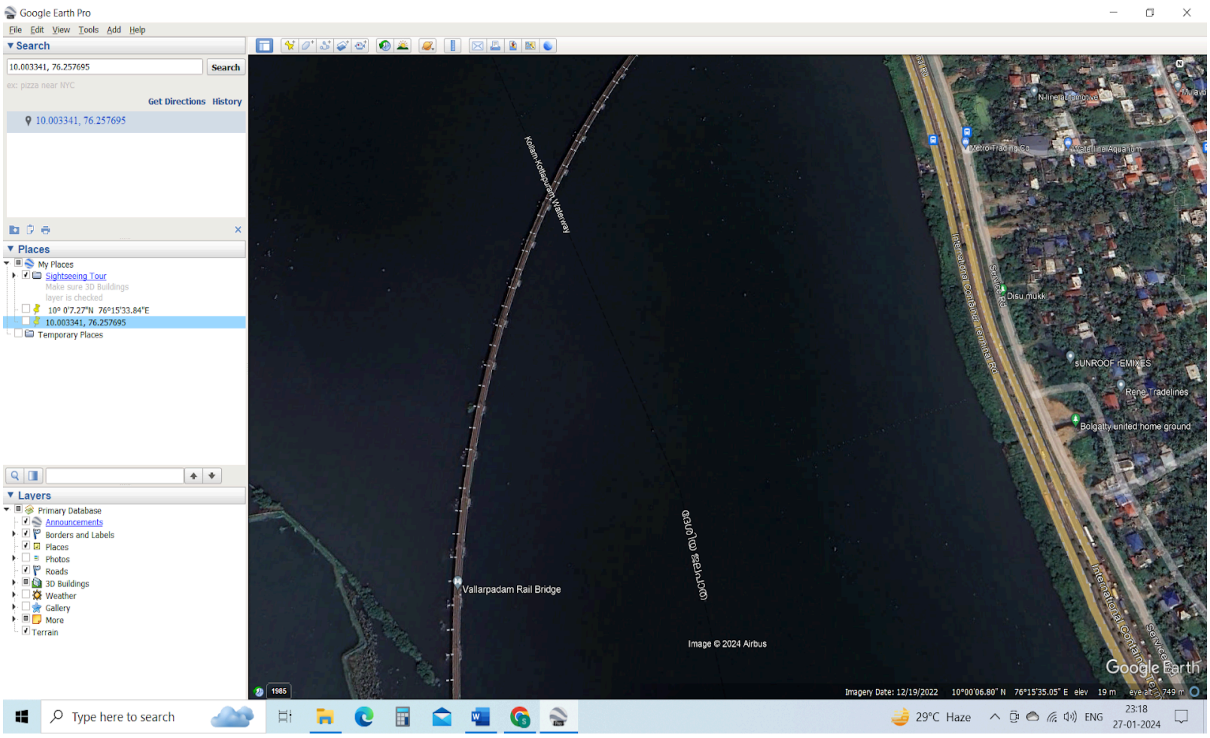
Now we will try to use two commonly available software (Google Earth and QGIS) for simple mapping.

Using Google Earth for Location Mapping

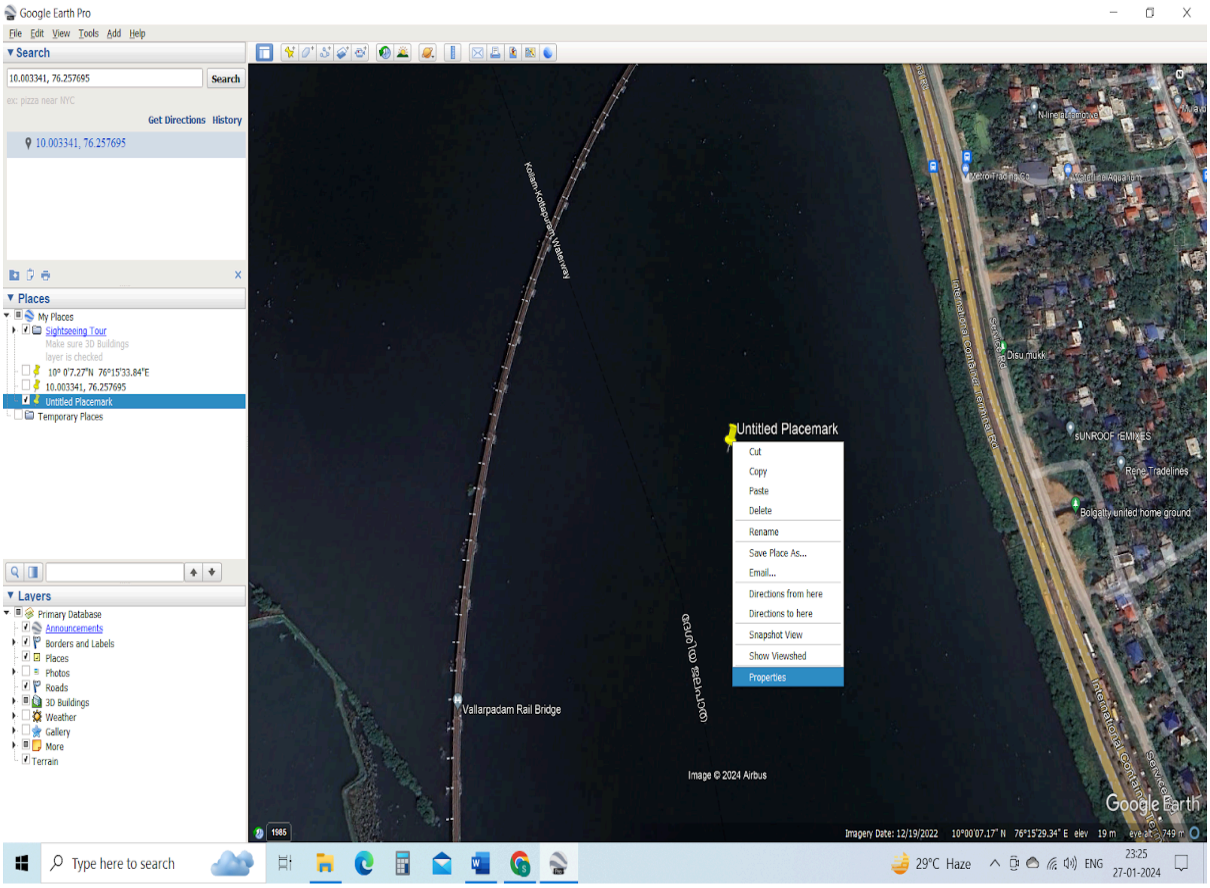
If you know the latitude and longitude of the location, type those coordinates in the search window and press the search button, a push button will be generated and it will be placed in the exact coordinates with the latitude and longitude visible.

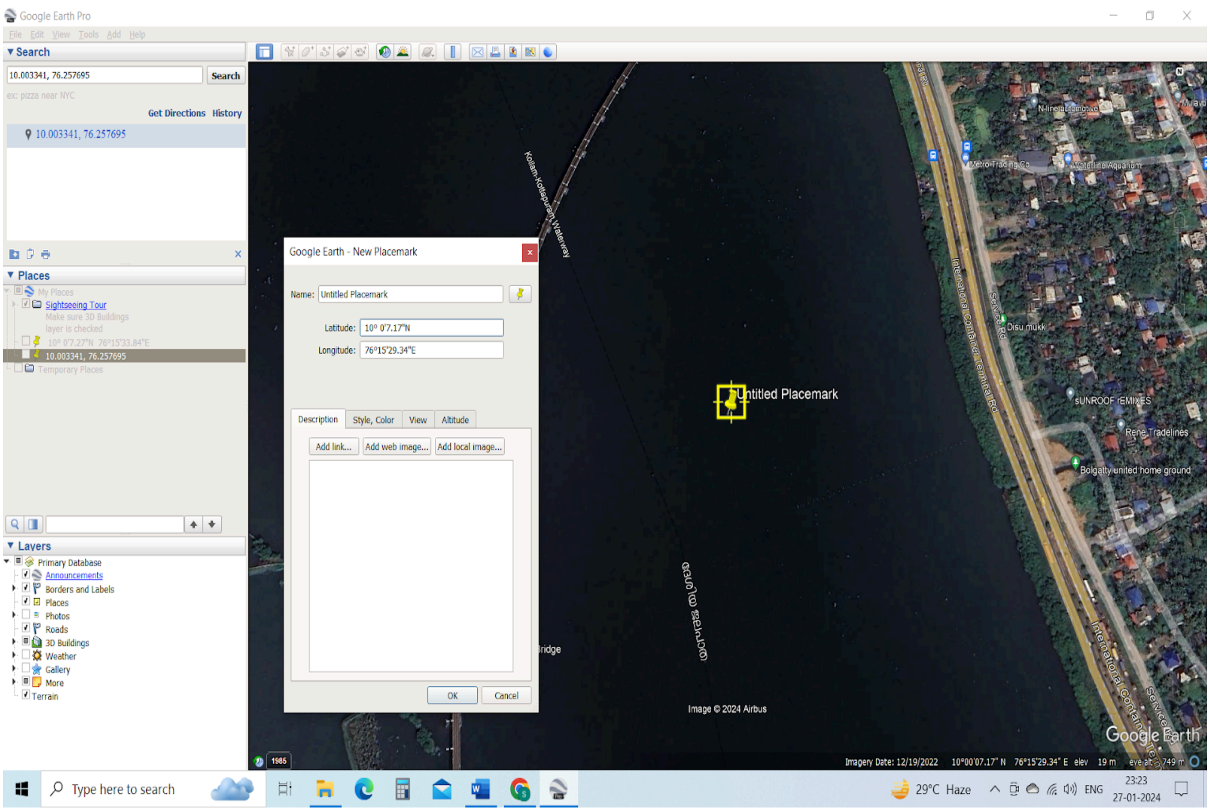


If you do not know the coordinates, but know the spot on the imagery, you can use the push button icon to add a place mark over the location. Using the properties menu of the place mark, one can do the required modifications like changing the place mark title, icon style etc.

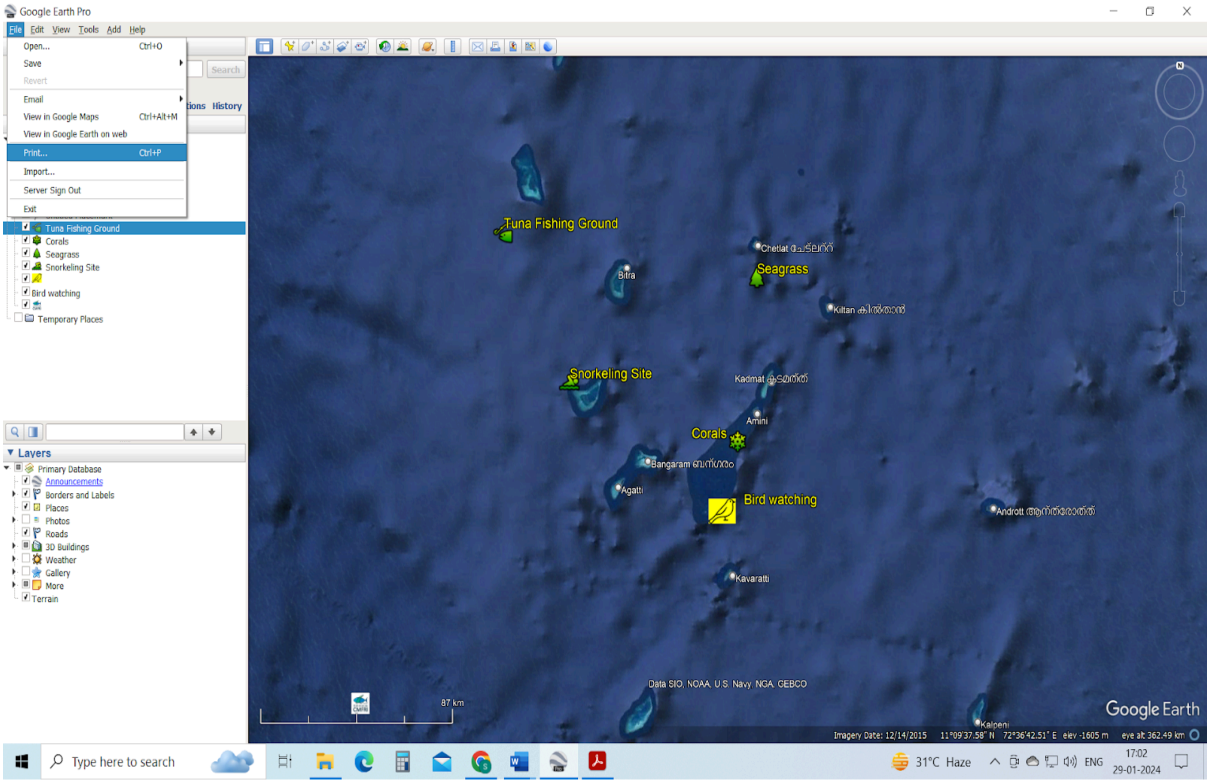


Right-clicking the placemark icon will open a drop-down menu showing the properties button using which you can change the placemark icon properties and do the editing.

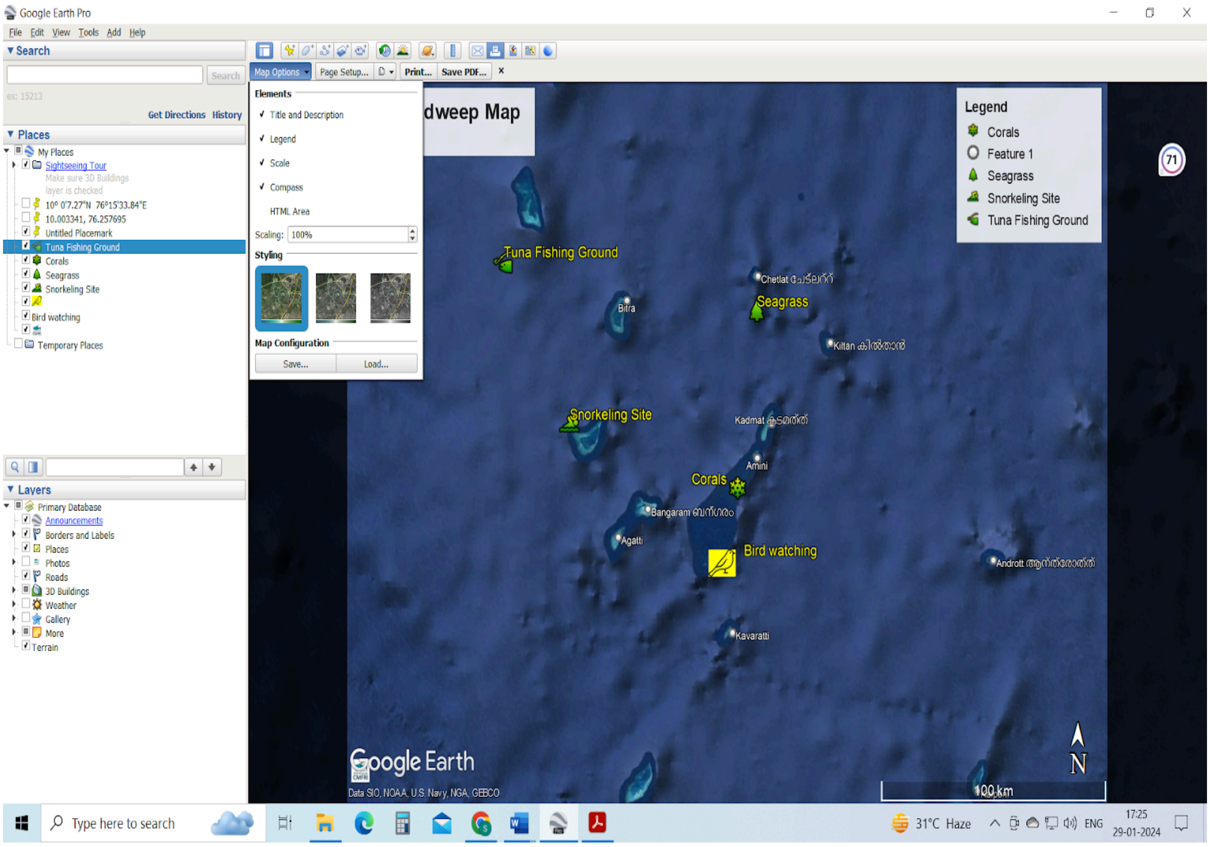




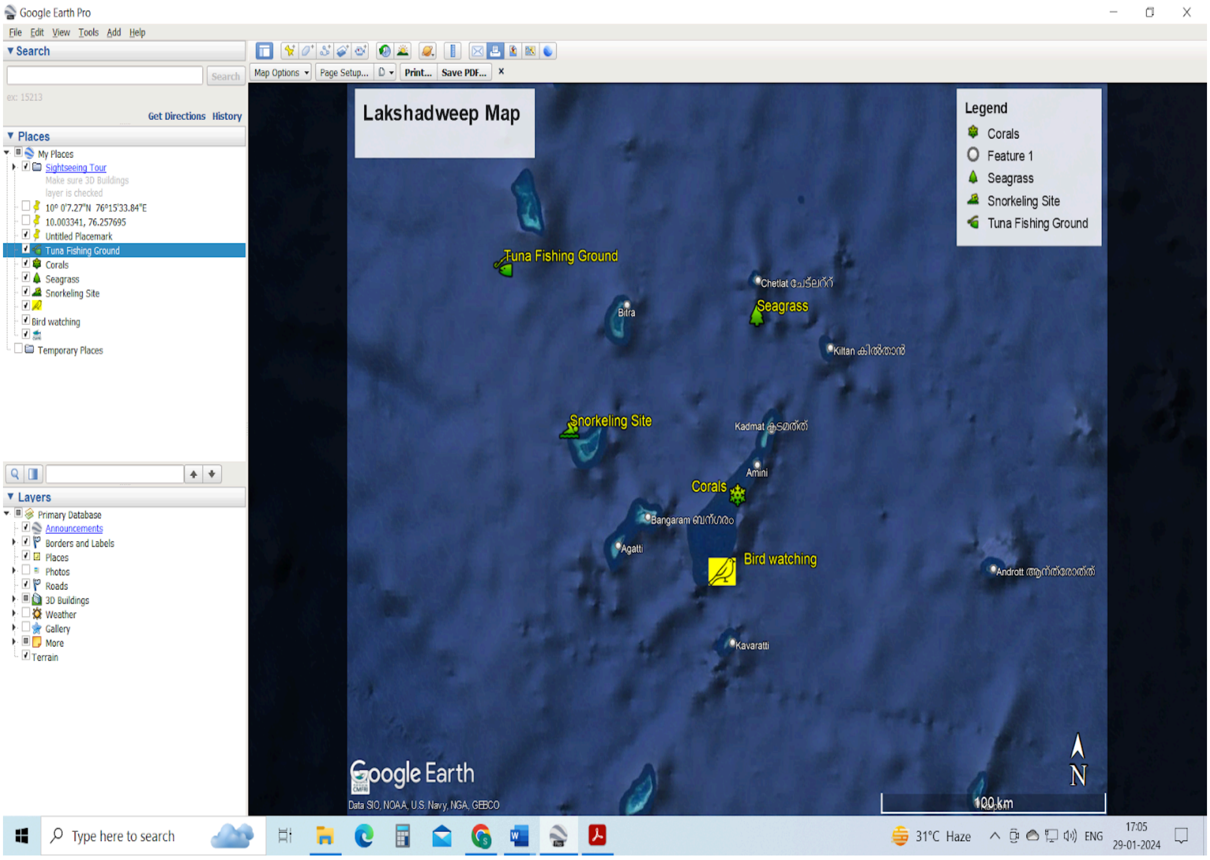
Once you are finished adding the placemarks, you can make the map composition from the print option of the File menu.

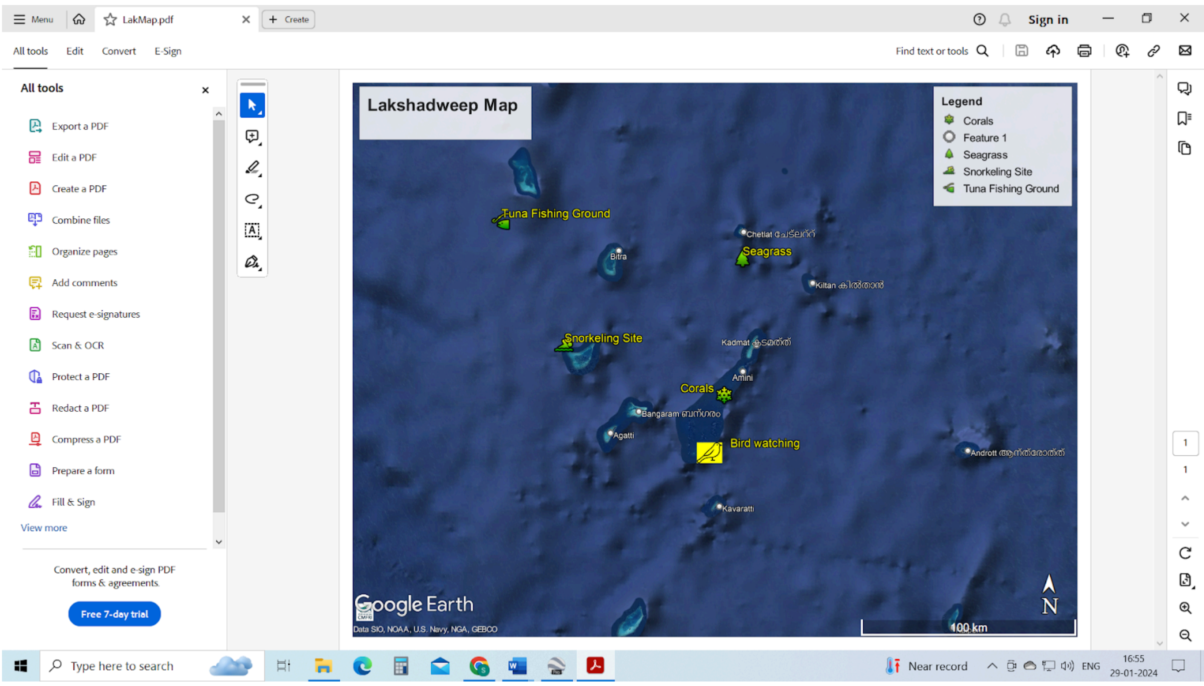


It will open up the print composition window using which you can add the map elements like Title, north arrow, legend and scale from the ‘Map options’ menu.



Make necessary editing in the map elements and save the map/print the map.

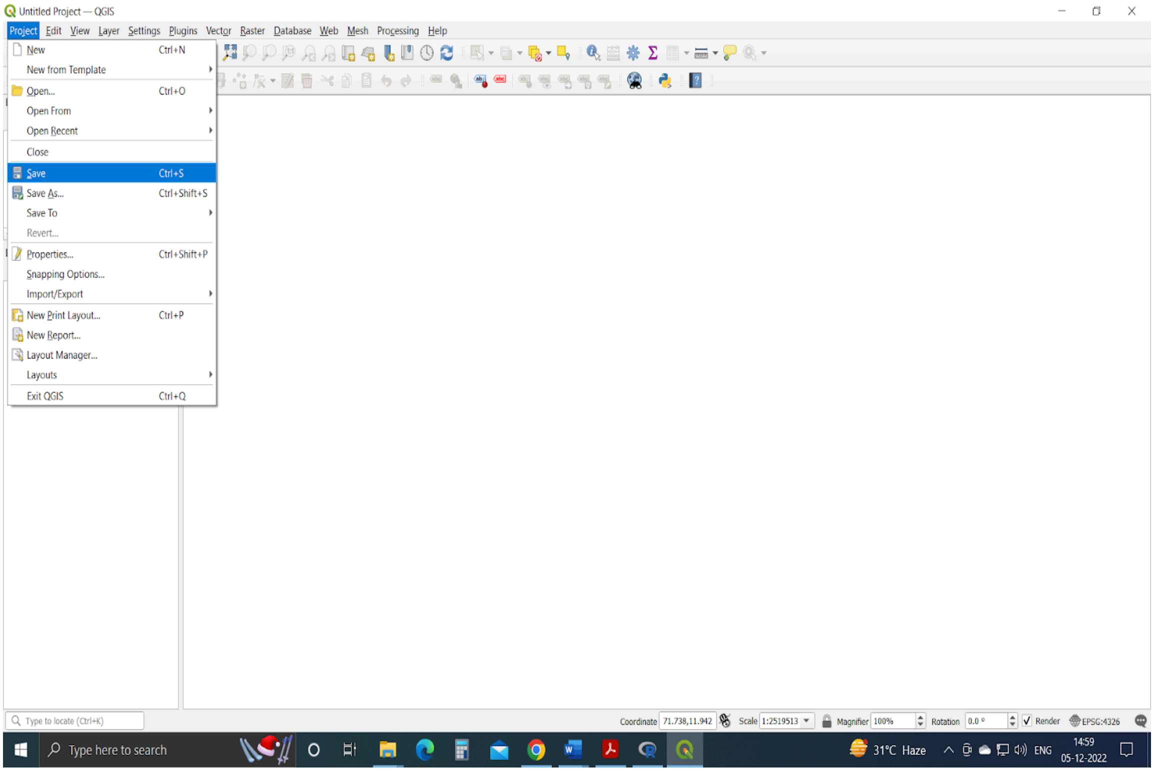




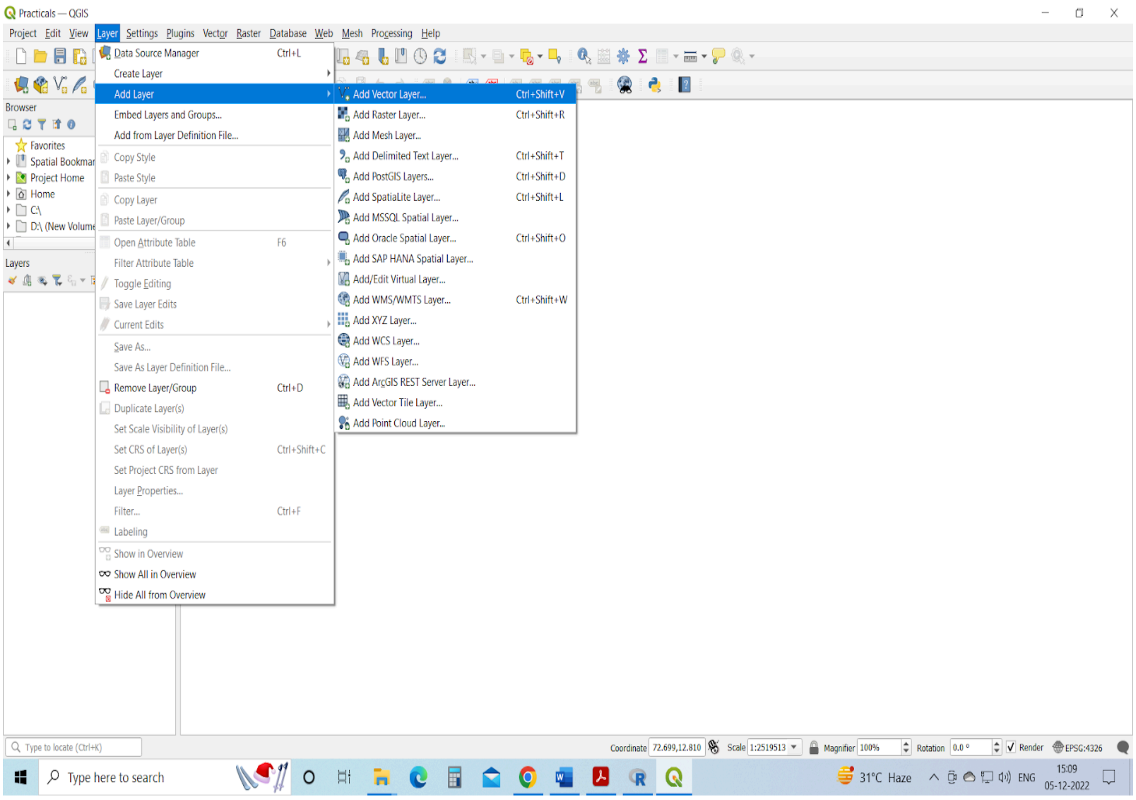
Making a Simple Location Map Using QGIS

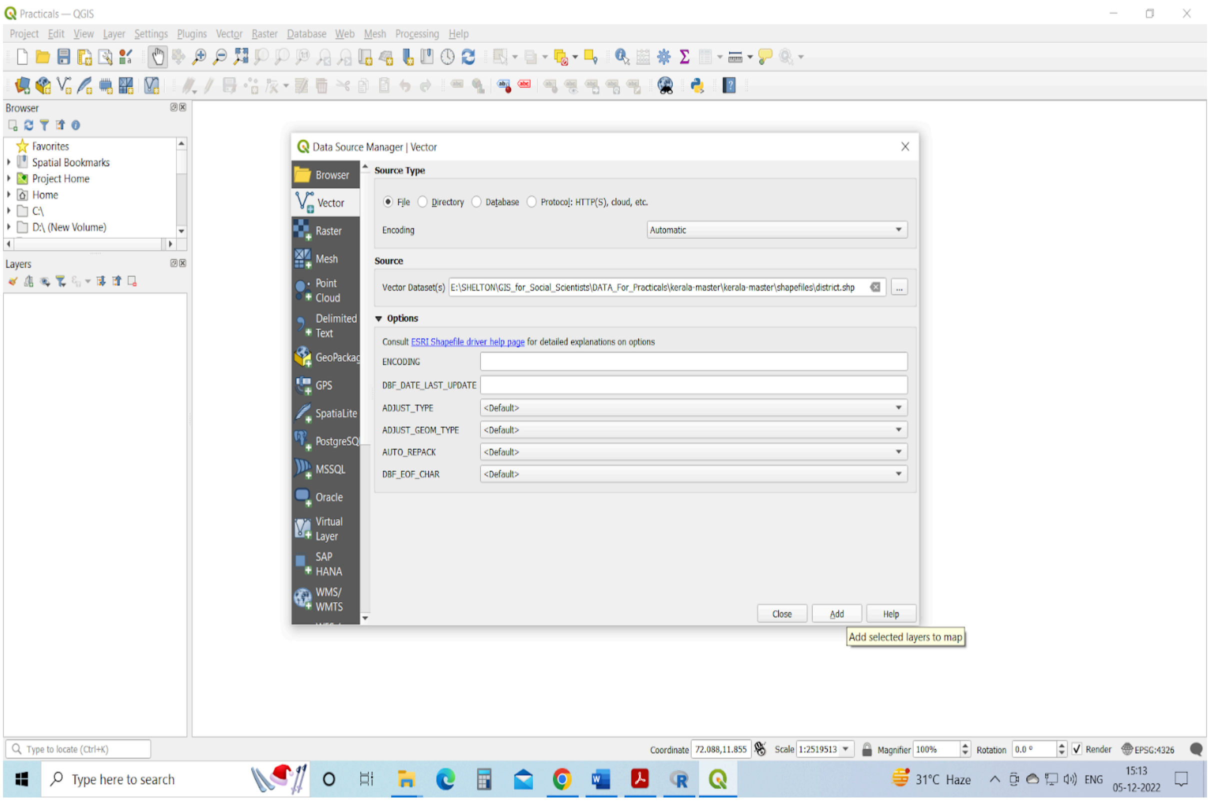
Here we learn to make a simple location map.

Open the QGIS software and give a name to the project. (Open QGIS>Project Menu>Save>give file name (Practicals.qgz)

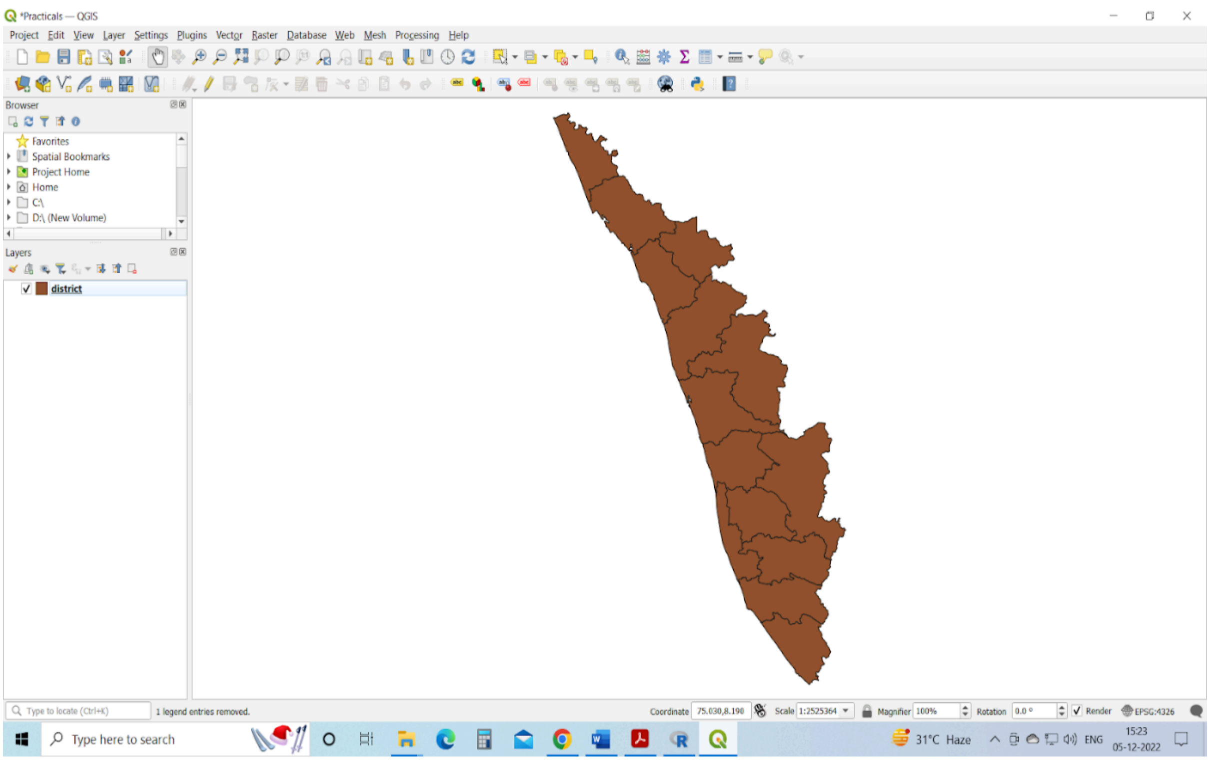


Now, load the shape file (District.shp). Layer Menu > Add Layer> Add Vector Layer> select the file name (District.shp)>Add.



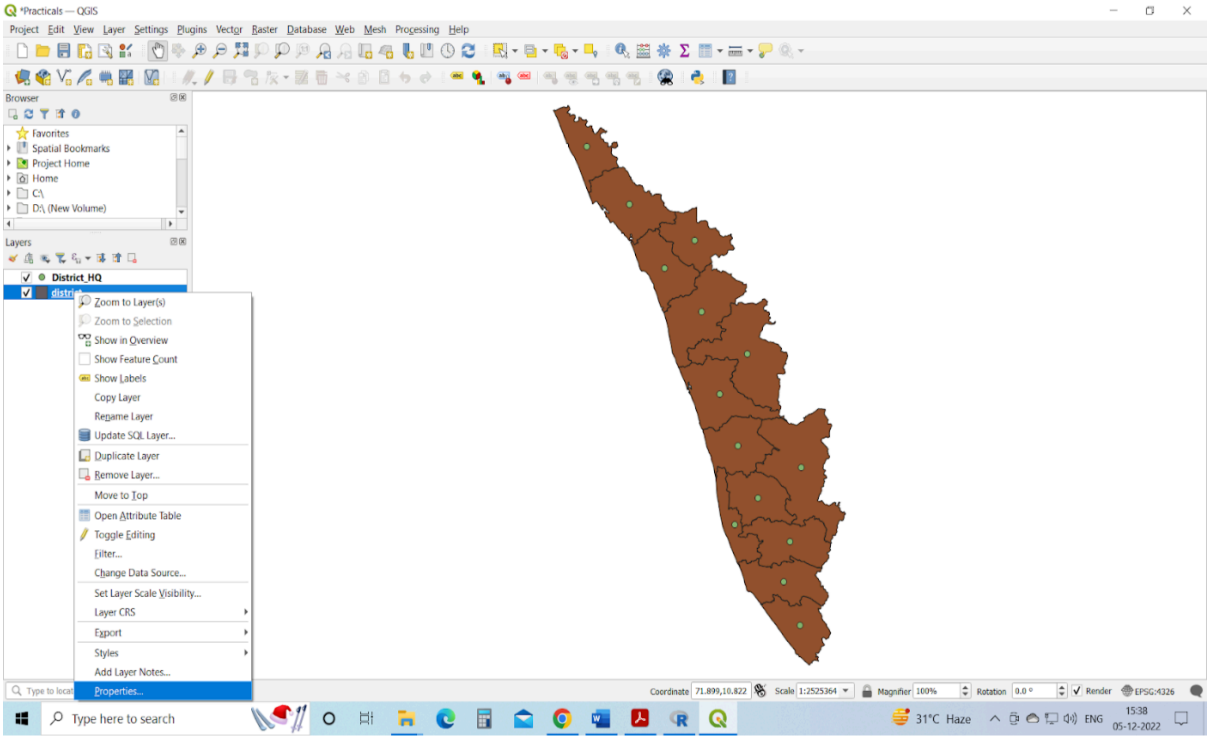


Now the Districts will appear in the map window and we can modify the appearance of the districts using the Layer panel.

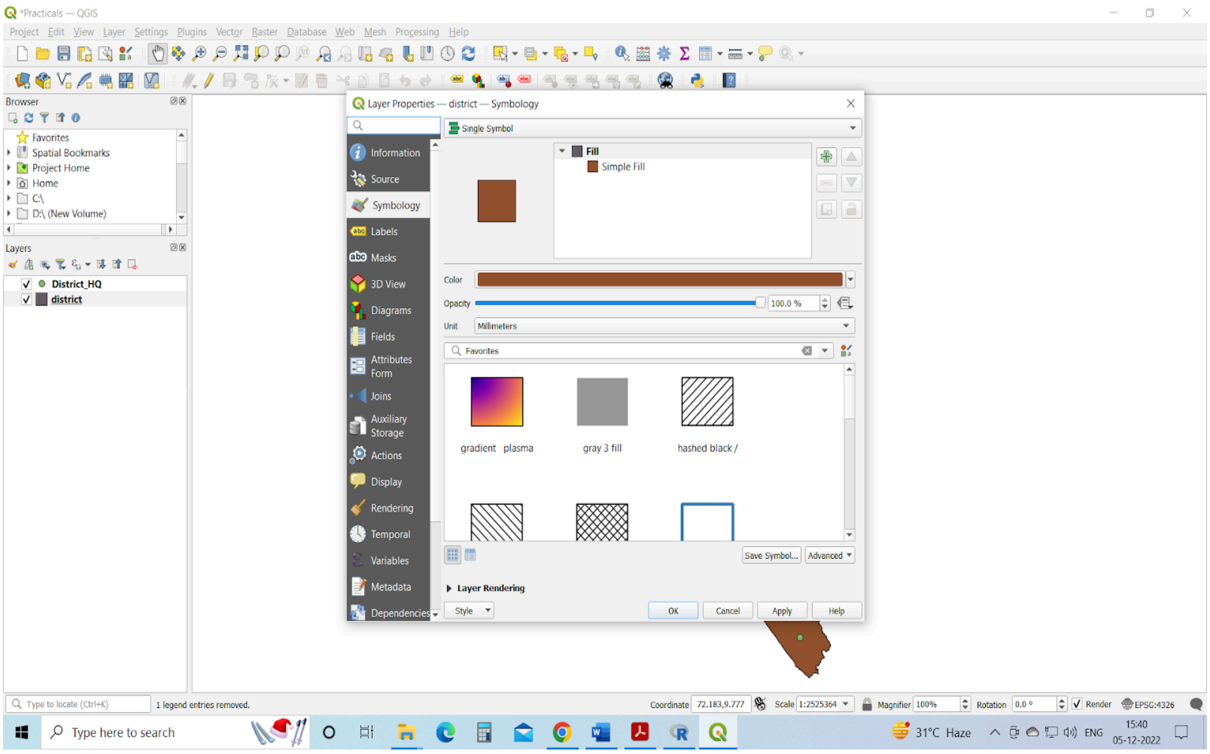


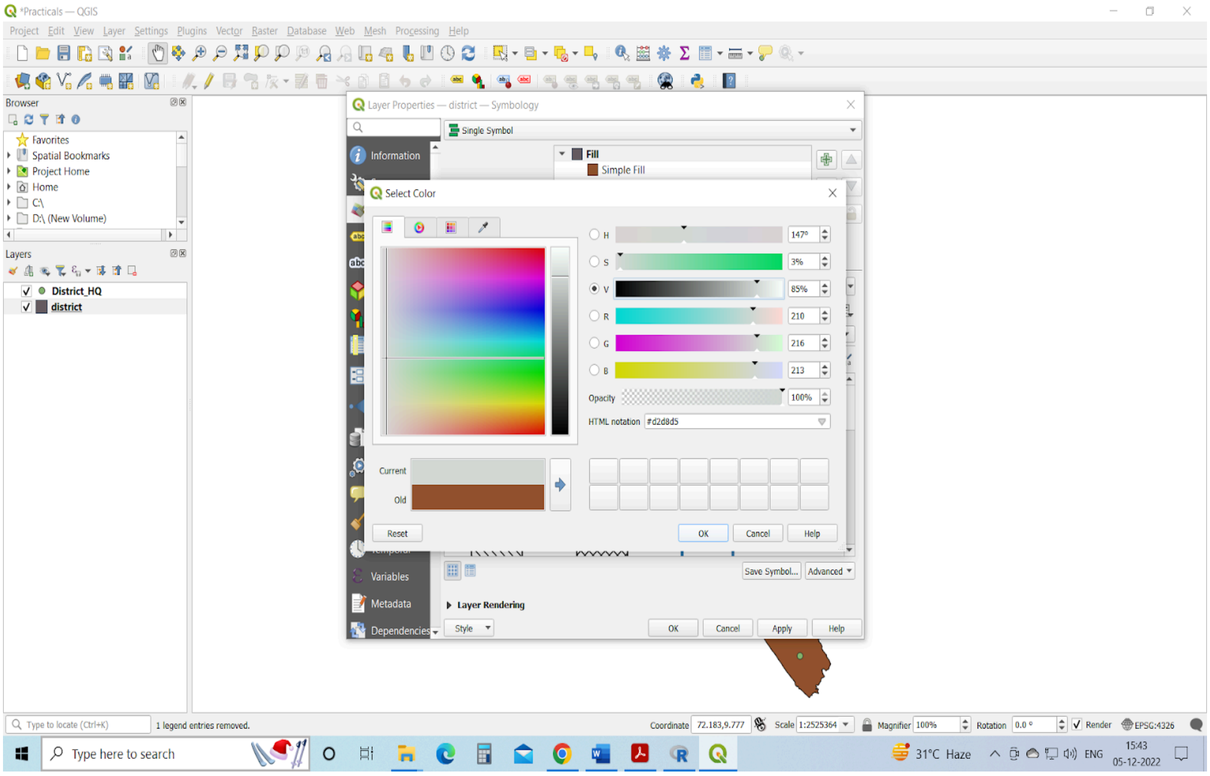
Likewise, add the file District_HQ.shp (Layer Menu > Add Layer> Add Vector Layer> select the file name (District_HQ.shp)>Add).

Now, we will change the appearance of the above vector files. Right-click the District.shp file in the layers menu to open its properties.



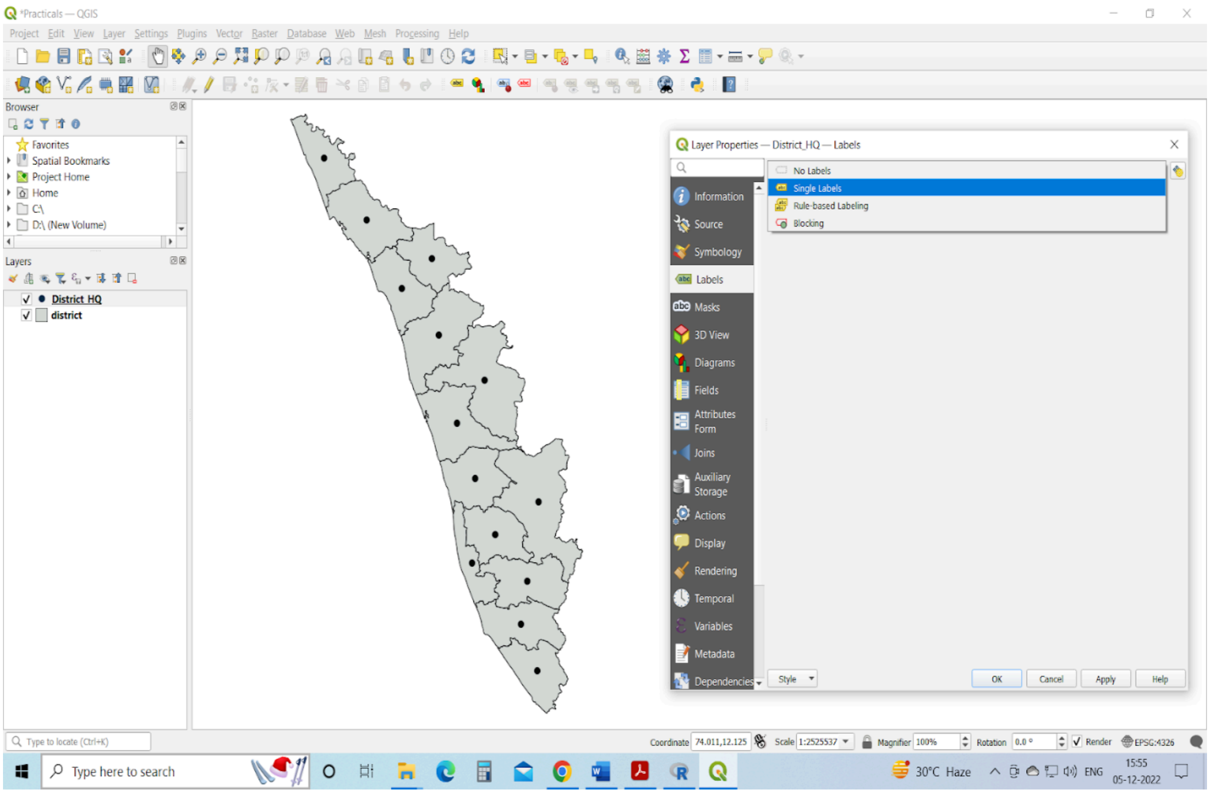
Then the Layer Properties window will appear. Then select the Symbology menu and change the colour of the polygon by clicking the colour bar in the window.



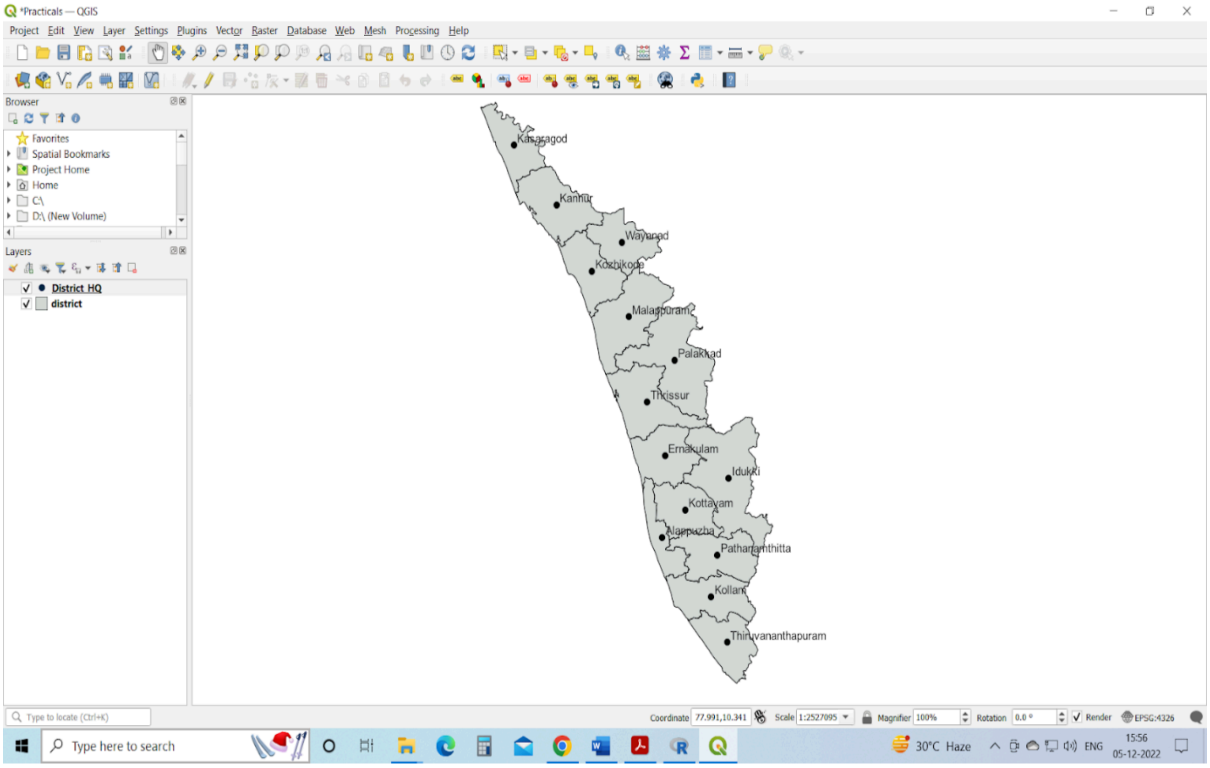


Likewise, change the colour of the District_HQ.shp file.

Now, we will add names to the points (District_HQ.shp). Right-click the District.shp file in the layers menu to open its properties. *Layer Properties* window will appear. Then select the Labels menu and change the selection from *No Labels* to *Single Labels*.

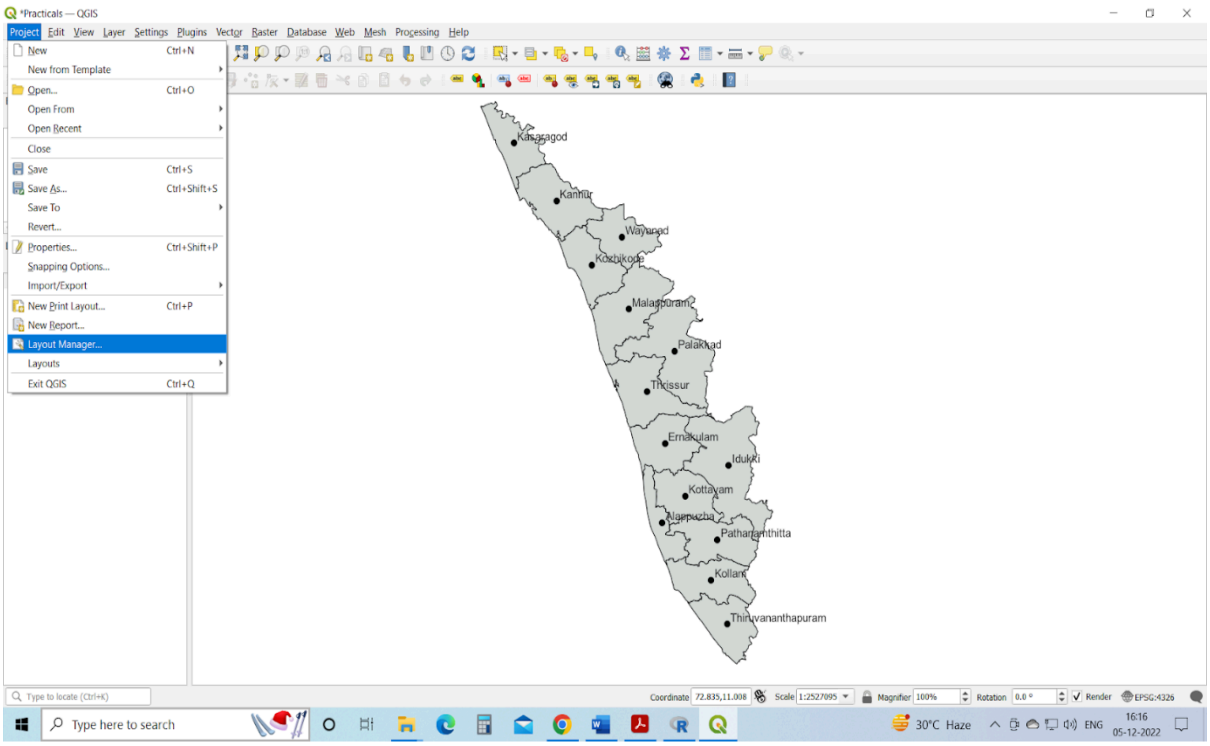


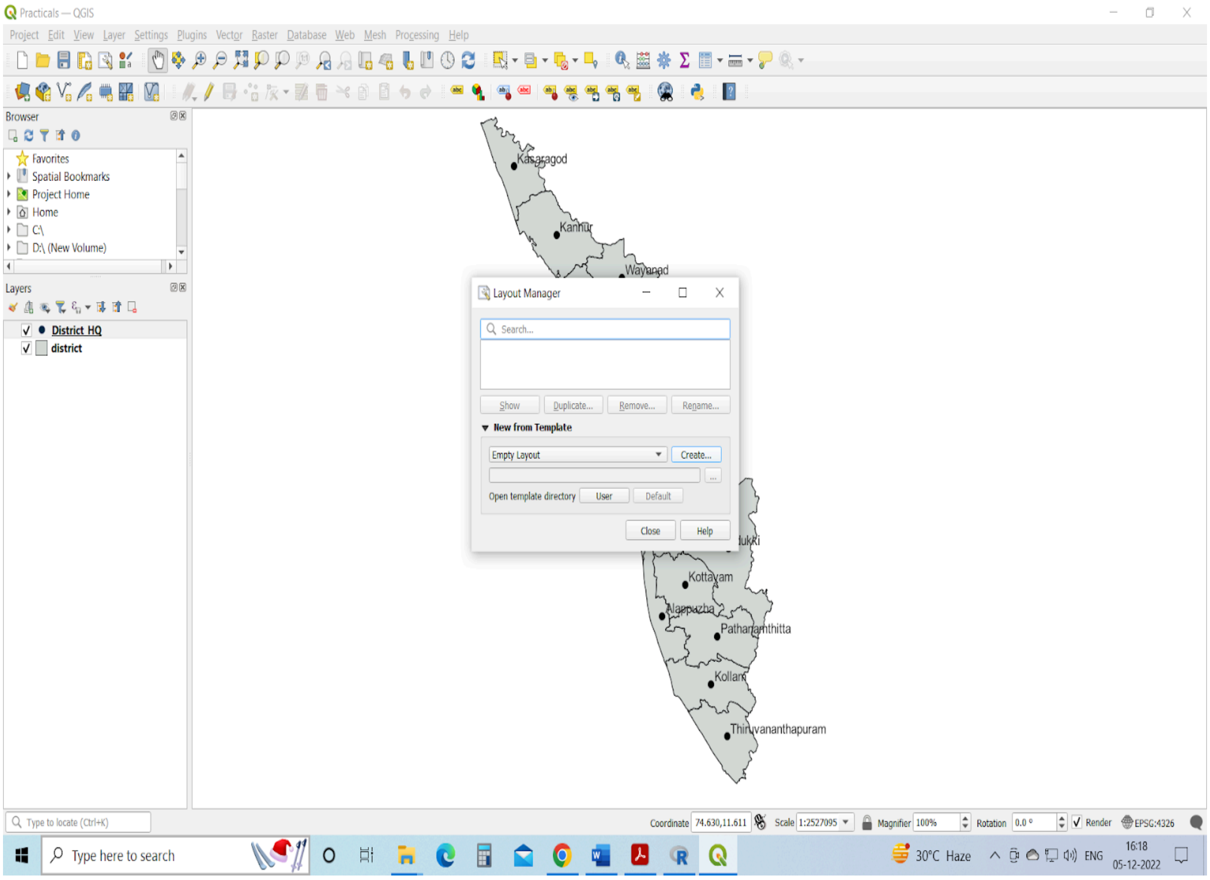
And click *Apply Button*. Now the District_HQ names will appear on the Map Window.

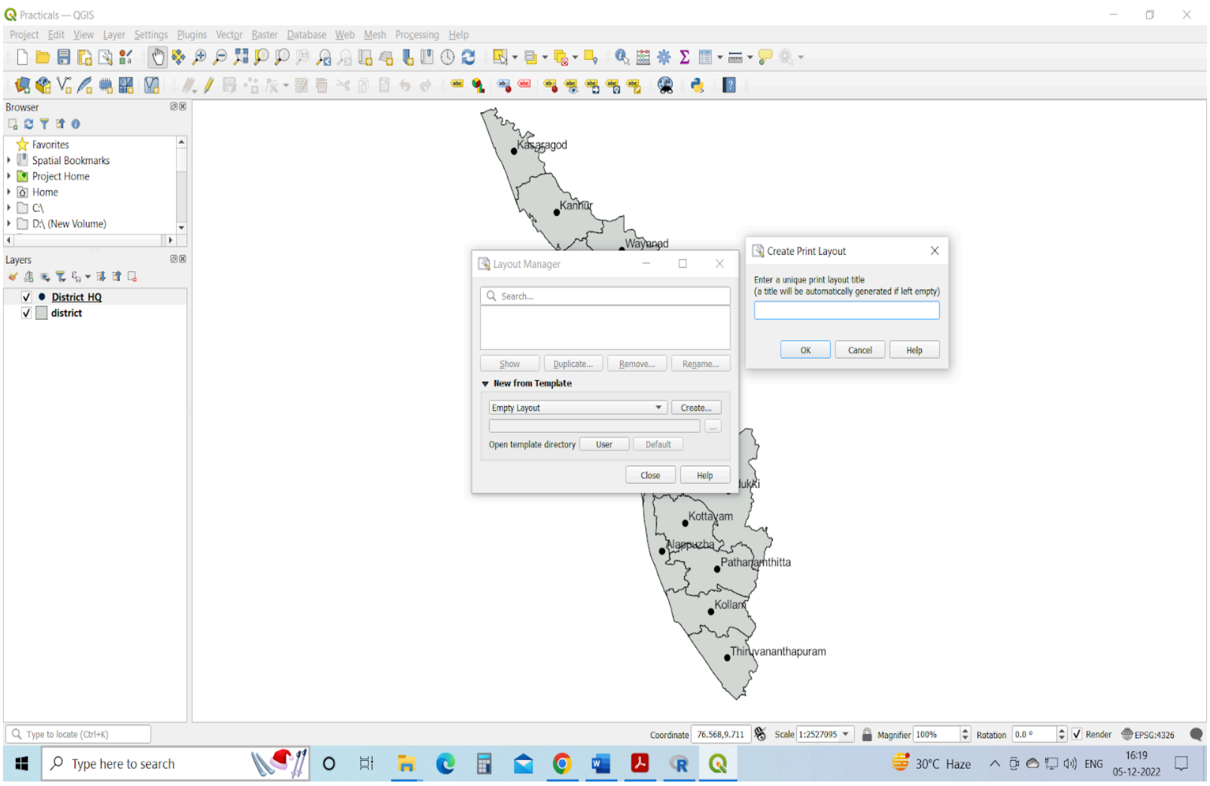


Now we will make a map layout from this data.

Go to the Layout Manager in the Project Menu (Project Menu>Layout Manager) and create an Empty Layout by clicking the create button in the Layer Manager, give a name for the Layout (*MyMap*) and press OK.

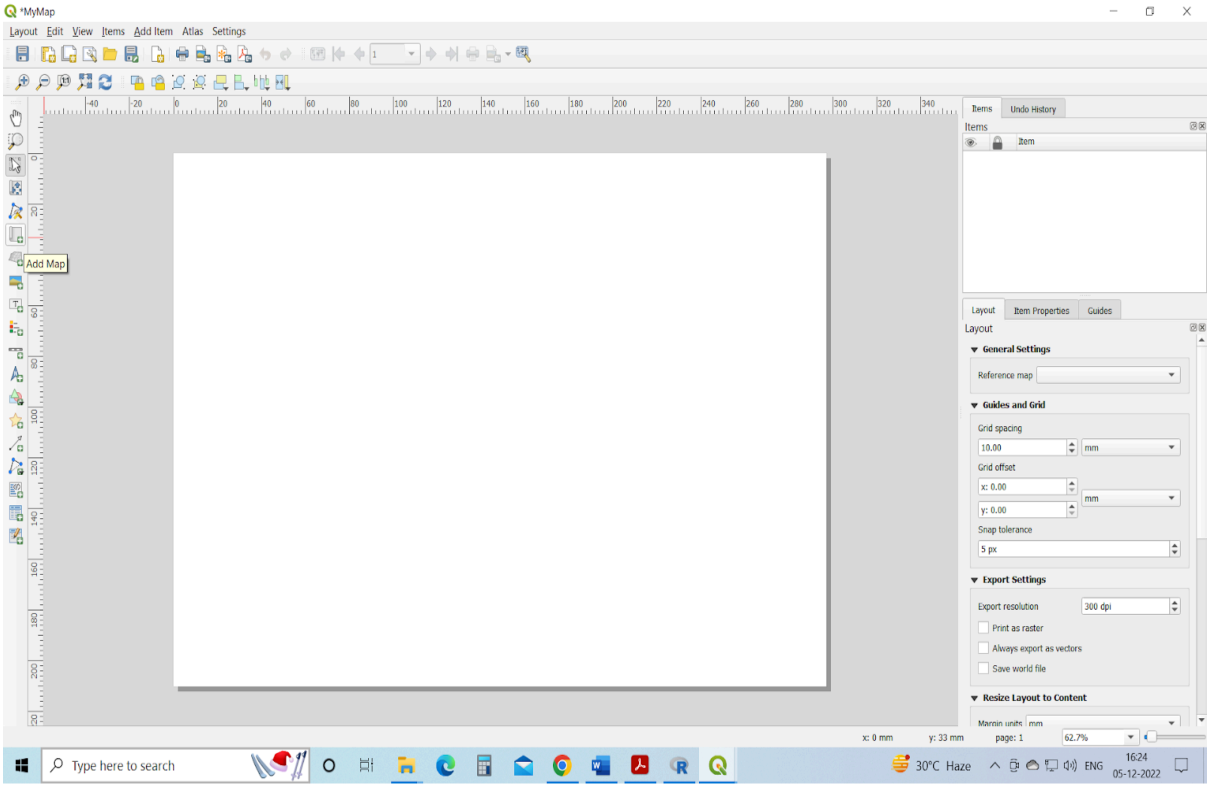


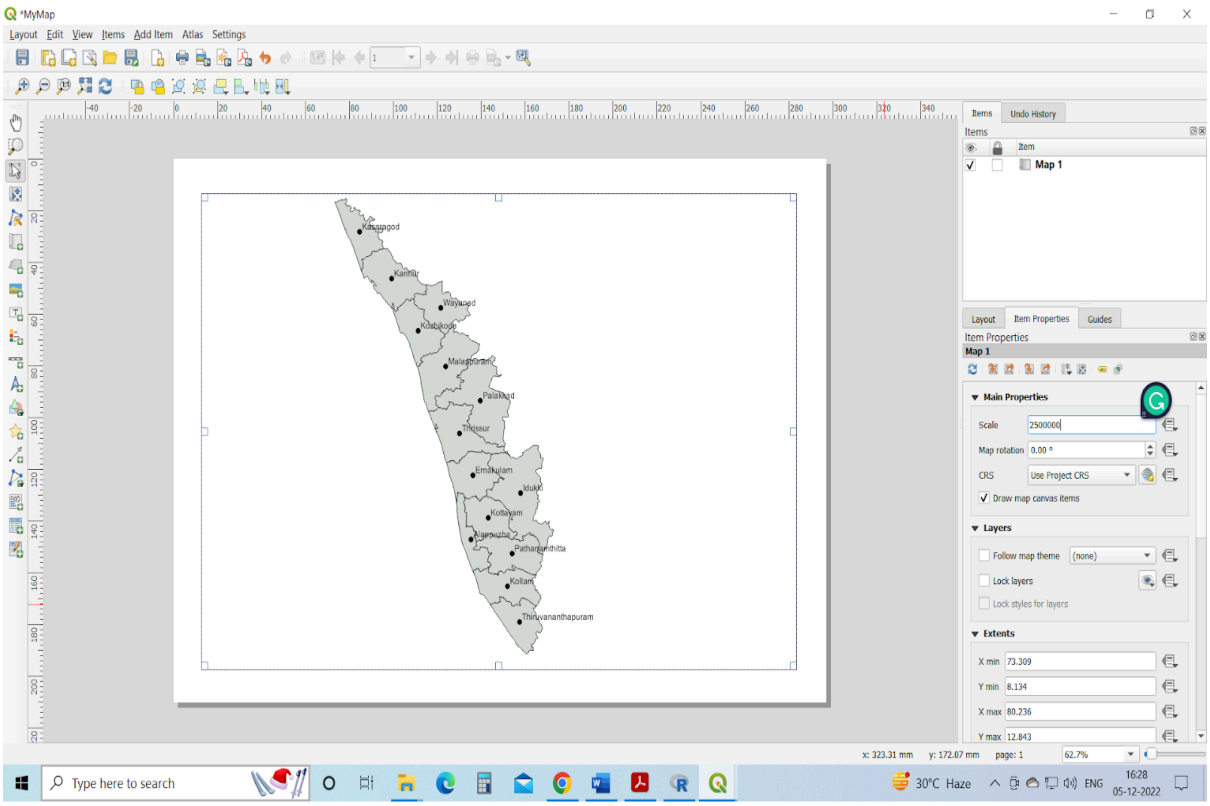




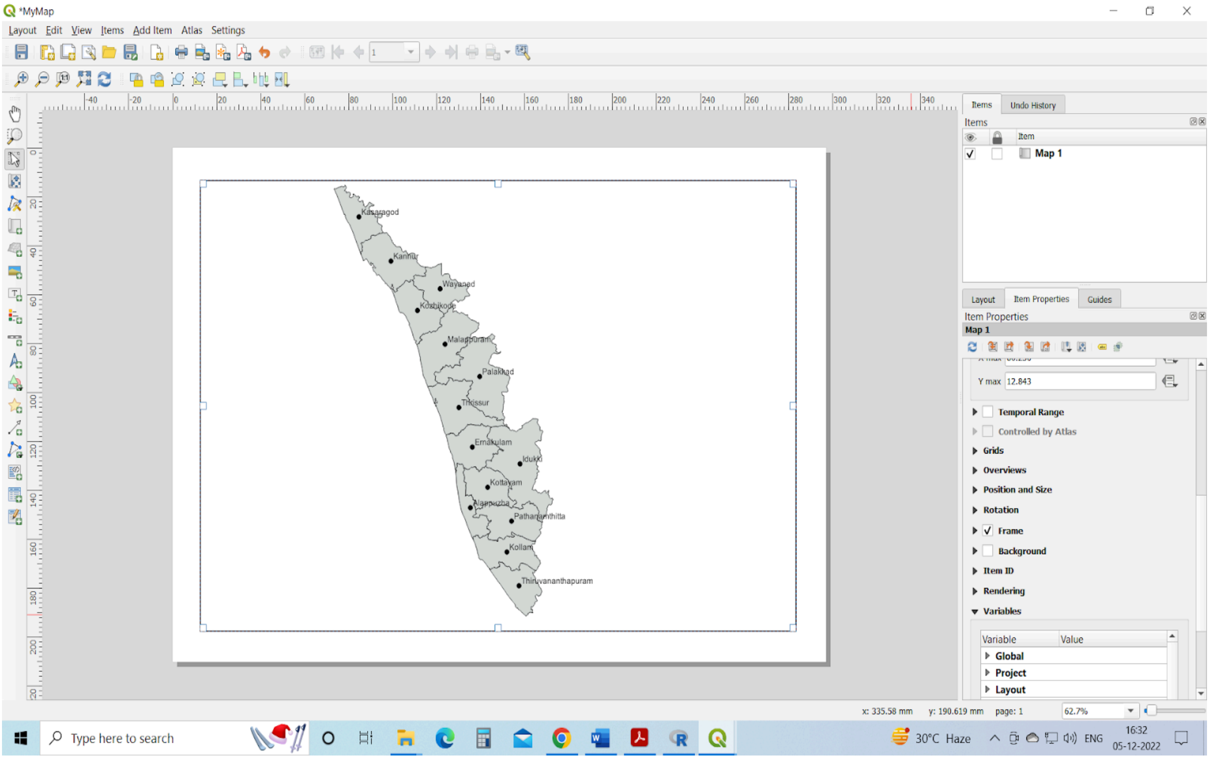
Now the MyMap layout will open up and one can make the necessary manipulations in it to make the map.

At first, we have to add the map to the *MyMap* layout by clicking the *Add Map* icon in the left panel of the map layout window and by clicking and dragging in the canvas.

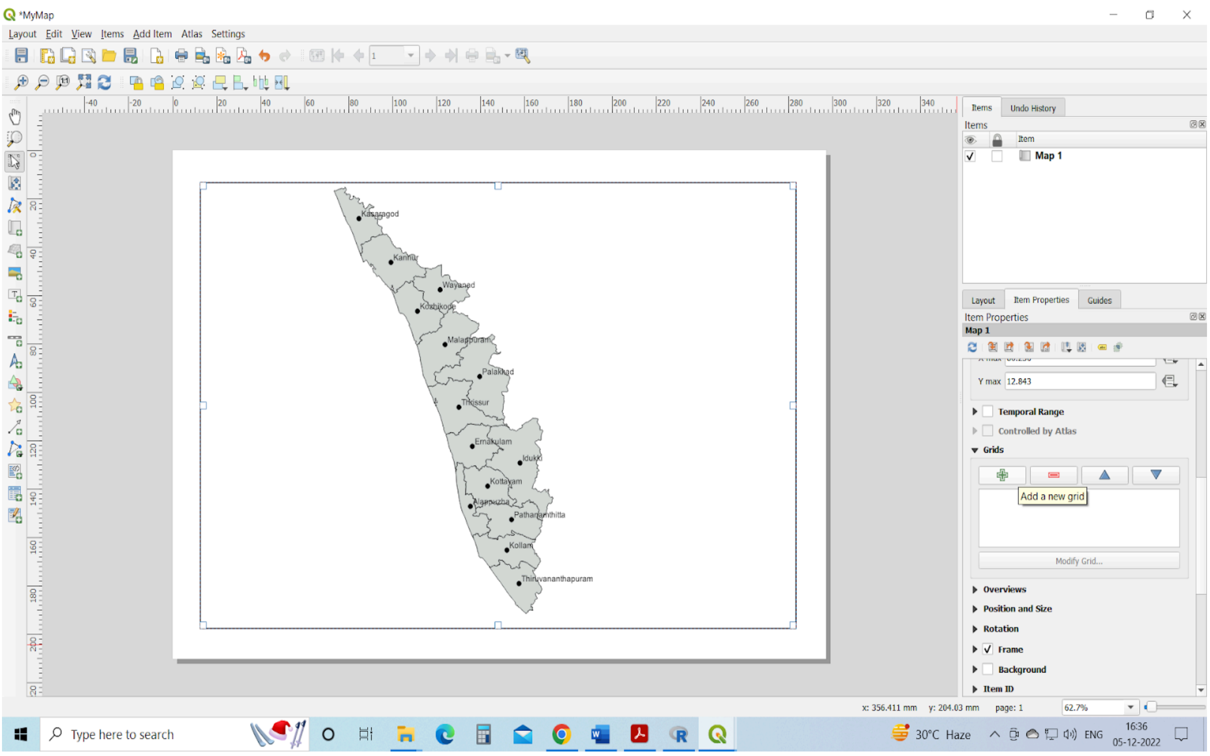




You can give a frame to the map by ticking the *Frame* option in the *Item Properties Menu* in the panel on the right side.



Now, we can add the lat-long grids to the map by adding a new grid to the map layout and changing its properties to suit the requirements. We can also add other map elements like North Arrow, Title, Legend, Scale Bar etc., using the icons in the left side of the Map Layout Window. Once the map is ready, it can be exported as an image using the *Export As Image* icon at the top of the Map Layout Window.



SUGGESTED READINGS

1. Principles of Geographical Information Systems.: By Peter A. Burrough, Rachael A. McDonnell, and Christopher D. Lloyd, Oxford University Press, 2015.

2. An Introduction to GIS: http://www.paulbolstad.net/5thedition/samplechaps/Chapter1_5th_small.pdf

3. Ferreira, J., João, P. and Martins, J. "GIS for Crime Analysis - Geography for Predictive Models" The Electronic Journal Information Systems Evaluation Volume 15 Issue 1 2012, (pp36 -49) www.ejise.com/issue/download.html?idArticle=817

4. Overman, Henry G. (2010) GIS a job: what use geographical information systems in spatial economics. Journal of regional science, 50 (1). pp. 165-180. ISSN 0022-4146; http://eprints.lse.ac.uk/30784/1/Gis_a_job_%28LSERO_version%29.pdf

5. QGIS: A Free and Open Source Geographic Information System; http://www.qgis.org/en/site/

6. Padua, S., Kripa, V., Prema, D. et al. Assessment of ecosystem health of a micro-level Ramsar coastal zone in the Vembanad Lake, Kerala, India. Environ Monit Assess 195, 95 (2023). https://doi.org/10.1007/s10661-022-10692-7

7. Robinson, T. P. and Metternicht, G., 2006. Testing the performance of spatial interpolation techniques for mapping soil properties *Computers and Electronics in Agriculture,* vol. 50, pp. 97–108.

8. Dinkins, C. P. and Jones, C., 2008. Soil Sampling Strategies. A self-learning resource from MSU Extension. MT200803AG New 4/08.

9. Li, J. and Heap, A.D., 2008. A Review of Spatial Interpolation Methods for Environmental Scientists. Geoscience Australia, Record 2008/23, pp. 1-137.

10. Hartkamp, A. D., De Beurs, K., Stein, A. and White, J.W., 1999. Interpolation Techniques for Climate Variables. NRG-GIS Series 99-01. Mexico, D.F.: CIMMYT.

11. Isaaks, E.H. and Srivastava, R.M., 1989. Applied Geostatistics. Oxford University Press, New York, 561 pp.

12. Hutchinson, M.F., 1995. Interpolating mean rainfall using thin plate smoothing splines. International Journal of Geographical Information Systems, 9(4): 385- 403.

13. Webster, R. and Oliver, M., 2001. Geostatistics for Environmental Scientists. John Wiley & Sons, Ltd, Chichester, 271 pp.

14. Burrough, P.A. and McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford, 333 pp.

15. Collins, F.C. and Bolstad, P.V., 1996. A comparison of spatial interpolation techniques in temperature estimation, Proceedings, Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, NM. Santa Barbara, CA: National Center for Geographic Information and Analysis, Santa Barbara.

16. Negrieros et. al., 2010. Geographical information systems principles of ordinary kriging interpolator. Journal of applied sciences, 10 (11), 852-867.

17. Jernigan, R. W., 1986. A primer on kriging. U.S. Environmental Protection Agency, Washington D. C. pp. 89.

18. Rossiter, D. G., 2012. A minimal introduction to geostatistics with R/gstat. University of Twente, Faculty of Geo-Information Science & Earth Observation (ITC), pp. 69.

19. Bivand, R. S., Pebesma, E. J., Gómez-Rubio, V. 2008. Applied Spatial Data Analysis with R. Springer Science+Business Media, LLC. pp.378