

**B.Sc. (Hons) in Business Information Systems Assignment
Cover Sheet****Student Details (Student should fill the content)**

Name	A.ASHRA FARVEEN	
Student ID	KG/BSCSD/19/07 (20328097)	

Scheduled unit details

Unit code	CIS6008	
Unit title	Analytics and Business Intelligence	
Unit enrolment details	Year	3
	Study period	2025
Lecturer	Assignment prepared by Induranga De Silva	
Mode of delivery	Full Time	

Assignment Details

Nature of the Assessment	Coursework (a report)		
Topic of the Case Study	Application of statistical and geospatial business analytics tools, techniques and methodologies to generate business intelligence essential for informed decision making in Power and Renewable Energy sector credible, efficient and effective informed decision making in Sri Lanka.		
Learning Outcomes covered	LO2, LO3, LO4		
Word count	5300(report)		
Due date / Time	19 st August 2025		
Extension granted?	Yes	No	Extension Date
Is this a resubmission?	Yes	No	Resubmission Date

Declaration

I certify that the attached material is my original work. No other person's work or ideas have been used without acknowledgement. Except where I have clearly stated that I have used some of this material elsewhere, I have not presented it for examination / assessment in any other course or unit at this or any other institution

Name/Signature	I.A Neranjan Prasad	Date	2023/09/06
----------------	---------------------	------	------------

Submission

Return to:	
------------	--

Result

Marks by 1 st Assessor		Name & Signature of the 1 st Assessor		Agreed Mark
Marks by 2 nd Assessor		Name & Signature of the 2 nd Assessor		
Comments on the Agreed mark				

CMU B.Sc. (HONS) BIS - ASSIGNMENT FEEDBACK SHEET –ICBT CAMPUS

STUDENT NAME:		STUDENT NUMBER:
Module Number & Title: Analytics and Business Intelligence		Semester: II
Assignment Type & Title: Coursework Application of statistical and geospatial business analytics tools, techniques and methodologies to generate business intelligence essential for informed decision making in Power and Renewable Energy sector credible, efficient and effective informed decision making in Sri Lanka.		
For student use: <i>Critical feedback on the individual progression towards achieving the assignment outcomes</i>		
For the Assessors' feedback Indicate the Task number strength and Weaknesses and the marks for each task		
Task No/Question No	Strengths (1st Assessor)	Strengths (2nd Assessor)

Task No / Question No	Weaknesses (1st Assessor)			Weaknesses (2nd Assessor)	
Areas for future improvement					
Comments by 1st Assessor			Comments by 2nd Assessor		
Marks					
Task /Question No	Marks by 1st Assessor	Marks by 2nd Assessor	Marks by IV (if any)	IV comments (If Any)	
Total Marks					
Name and the Signature of the 1st Assessor				Date:	
Name & Signature of the 2nd Assessor :				Date :	
Name & Signature of the IV: (If any)				Date :	

Task/assessment brief:**Purpose:**

This assignment is to assess student's ability to perform business analysis using Statistical and Geographic Information Systems(GIS) related tools, techniques and methodologies to find out applicable and useful intelligence for informed decision making in private and government sector institutions in the island and different parts of the countries in the world. The relevant higher level administrative officials of the those institutions can use GIS to generate maximum efficiency and benefits on informed business decision making while eliminating discrimination, ambiguity and uncertainty. Tasks Introduction Understand the given tasks based on Sri Lanka Higher Education Sector, using associated non geospatial and geospatial data models found in Data Science domain. Apply relevant tools, techniques and methodologies found in business analytics relevant to the module scope and conduct analysis on different subject matters with the support of source data provided in shape files (.shp), raster files(.asc),comma separated (.csv),text and excel file formats. The data analysis and visual demonstration required to be done using standard software tools recommended for the module (R, R-studio, R-commander, QGIS, PostgreSQL and Google Earth etc.).

Tasks: 1 – Report (100 Marks) The student required to do the following data analysis and visualizations based on datasets provided with this assignment using R, R-Studio, R-commander, QGIS, PostgreSQL, GOOGLE EARTH and other related supportive tools. All required datasets have been included within the “Data Sets” folder as separate subfolders per each question.

a) Ministry of Higher Education (MOHE) is working on various research studies to uncover new strategies to overcome inefficiencies in the higher education sector with the aid of local and global research data. It is very important to achieve improvements and developments of the country's higher education sector service quality improvements with the aid of strategic planning, informed decisions making and lessons learnt in the past. A dataset named University Quality Ranking (UQR.csv) and its data dictionary have been provided for you to uncover possible associations available between University Ranking Score and other related factors included in the dataset with

the aid of suitable, most appropriate statistical and business analytics tools such as R, R-studio and R-commander. As per the findings you are required to develop precise statistical models supported by graphical representations for supporting informed decision making. Your analysis should be followed by critical discussions of the findings relating to Sri Lanka's Higher Education Sector. (30 Marks)

- b) Develop a Sri Lanka map that visualizes information of Sri Lanka Schools in Table 2.2: Administrative Structure of Education by District in Annual School Sensus of Sri Lanka published by the Department of Census and Statistics.(provided as School_Census_Report_2022.pdf). The map should visualize information such as Administrative District Name, Number of National Schools and Provincial Schools retrieved via the newly created .csv file SL_Schools-2025. The new .csv file should contain only the aforesaid information. The map should be classified by the number of Total Schools. The map processing should be done using the provided vector data set with a suitable base map. The scenario exhibited by the map should be critically described with the support of all available resources. (20 Marks)
- c) Develop a digitized informative area map with suitable information about Ministry of Higher Education(MOHE) support digitization with QGIS open-layer plugins/Google Earth/Google Maps. It is mandatory to do the image geo-referencing before initiation of the digitization. Every vector layer attribute table should contain suitable data in columns id, name, type, and size. By analyzing the map, discuss critically how does MOHE geo-location contributes to improve efficient and effective service to the National Education System of the country. (For map development Use coordinates reference system as EPSG:5234). (10 Marks) d) Develop a Sri Lanka map contains 18 Universities belongs to the Ministry of Higher Education (MOHE) Sri Lanka. The map should visualize the information such as Location Label, District Name, Geo Location (Latitude and Longitude). The exact GPS locations information should be retrieved via Google Earth with the support of a KML/KMZ file. Develop a Postgres SQL Geospatial Database named SL_Uni_2025 and store all above associated data ,vector files and raster files in it.The required data for QGIS spatial analysis should be retrieved via the database created. A comprehensive critical discussion with recommendations should be included about aforesaid Universities contribution to the national higher education sector service improvements by considering their geo location, services and

contribution to the economic development in the Island and whether or not expected objectives as a country have been achieved so far. Supported authorized information can be found from <https://www.mohe.gov.lk> the official web presence of MOHE. (20 Marks)

e) Develop a map with the support of provided data set in order to find out the most suitable land for newly establishing Sri Lanka's first ever Research and Development Center for Space Sciences. The research center should be located in an area 1 Km away from the Ananda Primary School and 2 Km away from Sri Gunananda Vidyalaya. The ideal location should be selected considering nature of the soil in the land use area and no land should be selected that has already been used for traditional export crops. The map should be followed by a critical discussion of the feasibility of the decision of establishing the aforesaid Research Center in the identified area. The discussion should be followed by the following supportive information as well.

- Geo spatially identification of the entire area shown in the map.
- Total number of buildings situated within the suitability area at present.
- Total land area occupied by the buildings within the suitability area.
- Total suitable and available land area for the construction. (20 Marks)

ABI.pdf

ORIGINALITY REPORT

2%
SIMILARITY INDEX

1%
INTERNET SOURCES

1%
PUBLICATIONS

1%
STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Kaplan College Student Paper	<1%
2	phytotaxa.mapress.com Internet Source	<1%
3	rupress.org Internet Source	<1%
4	Alexander J. Kent, Doug Specht. "The Routledge Handbook of Geospatial Technologies and Society". Routledge, 2023 Publication	<1%
5	G. G. T. D. Wickramathilake, W. Hemachandra. "Chapter 5 Enhancing the Durability and Resilience of School Buildings in Sri Lanka: An Integrative Study of Common Defects and Remedial Measures", Springer Science and Business Media LLC, 2025 Publication	<1%
6	Submitted to International Business School Student Paper	<1%
7	Submitted to Asia Pacific Institute of Information Technology Student Paper	<1%
8	livrepository.liverpool.ac.uk Internet Source	<1%

Exclude quotes: Off
Exclude references: Off

Exclude matches: Off

Exclude bibliography: Off

Acknowledgment

I'm A. Ashra Farveen, a BSc in Software Engineering Batch 19 student, and I'd want to convey my heartfelt gratitude and admiration to our renowned Analytics & Business Intelligence lecturer, Mr. Induranga de Silva. His consistent counsel and support have been crucial as I embarked on my educational adventure. Thank you also to the dedicated instructors and personnel of my alma mater, whose thoughts and resources have been critical to my development as a student. In addition, I'd want to convey my heartfelt gratitude to the folks in my peer group who have supported me every step of the way.

My education has been greatly boosted by the friendships I've built with my classmates as we collaborated on numerous projects and responsibilities. Their constructive critiques and innovative suggestions have broadened my software engineering knowledge. Finally, I'd like to thank everyone who has been so supportive of me. Their unwavering faith in me and support have given me the courage to pursue my academic aspirations. I shall be eternally grateful to them for everything they've done to assist me reach my academic goals.

Executive Summary

This research conducts a thorough and data-driven investigation into the current state and strategic development potential of Sri Lanka's higher education system. The country faces both economic concerns and worldwide academic competitiveness, so educated, evidence-based policymaking is more important than ever. Recognizing this, the Ministry of Higher Education (MOHE) is investing in advanced business analytics (BA) and geospatial intelligence (GIS) to better uncover systemic inefficiencies and open up opportunities for change.

The initial phase of this research focuses on statistical analysis of university performance utilizing the University Quality Ranking (UQR) dataset. To identify the most influential factors driving university rankings, analytical techniques such as correlation analysis, hypothesis testing, and regression modeling were used with R, R-Studio, and R-Commander. Variables like research productivity, faculty composition, infrastructure, and student happiness were thoroughly studied. The findings provide unambiguous, data-driven insights into how various institutional traits lead to academic excellence—as well as potential deficiencies.

Beyond statistical research, the second phase uses Geographic Information Systems (GIS) tools like QGIS, Google Earth, and PostgreSQL/PostGIS to spatially map and evaluate Sri Lanka's educational infrastructure. The geospatial component addresses practical questions such as where national and provincial schools are concentrated and where resources are scarce, how accessible Sri Lanka's 18 public universities are across districts, and where the best locations for future developments—including a proposed Space Science Research Centre—might be based on terrain, land use, and proximity to critical services.

This comprehensive approach identifies large regional and institutional gaps, highlighting areas for policy intervention, infrastructure growth, and resource reallocation. The fundamental power of this study is its ability to translate raw data into actionable intelligence. By combining BA and GIS approaches, it provides policymakers and school planners with the resources they need to make fair and cost-effective decisions, foresee future requirements and issues, maximize the

return on public education investment, and encourage inclusive development that excludes no district.

Finally, the report shows how advanced analytics can go beyond descriptive data to guide significant higher education reforms. It highlights the value of digital transformation, data literacy, and systemic transparency. The strategic proposals presented not only complement the MOHE's current aims, but also serve as a solid foundation for Sri Lanka's education system to become smarter, fairer, and more globally competitive.

Table of Content

Acknowledgment.....	ii
Executive Summary	iii
Table of Content.....	iv
List of Figures	vi
List Of Tables.....	vi
Abbreviations	vii
Introduction	13
Chapter 1	15
1.1. Data Preparation	18
1.2. Dataset : University Quality Ranking. (UQR.csv)	18
1.3. Data Collection and Integration	18
1.4. Data Cleaning.....	18
1.5. Data Exploration.....	20
1.6. Data Transformation.....	21
1.7. Data Validation.....	21
1.8. Data Preparation for Analytic Tools	22
1.8.1. Summary of the Dataset	23
1.8.2. Install package	23
1.8.3. Scatter Plot	23
1.9. Regression Analysis	24
Chapter 2	25
2.1 Development of SriLanka Scholl network	25
2.2 Visualization Analysis School details	27
2.3 Digitization	28 11

2.4 Identification and Analysis of Educational Infrastructure in Sri Lanka.....	32
2.5 Mapping and Evaluation of Suitable Land for the Regional Research Center for Renewable Energy in Kandy.....	33
TASK – C Digitization and Geospatial Analysis	38
3.2 Introduction.	
3.3 Methodology – georeferencing and digitization	
3.4 Critical Discussion	
3.5 Conclusion	
4 TASK – D Geospatial and Economic	44
4.1 Critical discussion and recommendation	
5 TASK – E Geospatial and Economic	58
Conclusion	60
Future Recommendation.....	61
Reference.....	I
Appendices	III

Install Quick map service to illustrate maps

Select – KANDAWALA / SRILANAKA GRID

TASK – B

TASK – C

TASK – D

TASK – E

List of Figures

Figure 1 First 8 Data of that Data Set.....	3
Figure 2 First 8 Data Collection and Integration	3
Figure 3 Data Cleaning Code in R	3
Figure 4 Summary of the Dataset	3
Figure 5 install package	4
Figure 6 Scatter Plot.....	6
Figure 7 TASK B	6
Figure 8 TASK C	6
Figure 9 TASK D.....	6
Figure 10 TASK E	6

Abbreviations

Term	Definition
UQR	University Quality Ranking
R&D	Research and Development
GIS	Geographic Information System
QGIS	Quantum Geographic Information System
RMSE	Root Mean Squared Error
GPS	Global Positioning System
KML	Keyhole Markup Language
KMZ	Keyhole Markup Zipped

Introduction

Sri Lanka's higher education system is a pillar of the country's socioeconomic development, responsible for preparing qualified professionals, stimulating innovation, and contributing to research that advances national goals. In a period of rapid globalization and technological development, higher education institutions must adapt to fulfill rising demands for quality, relevance, and access. The Ministry of Higher Education (MOHE) recognizes the vital need to develop the sector's capability so that Sri Lankan universities can compete effectively on a regional and global scale.

Recent worldwide events have emphasized the importance of data-driven policymaking in improving the quality and impact of higher education. Countries that have adopted sophisticated business analytics (BA) and geospatial information systems (GIS) have been able to increase institutional performance, equity of access, and efficient allocation of scarce resources. These advances allow for detailed insights that go beyond traditional measurements by adding spatial, demographic, and infrastructure elements into strategic decision-making.

Sri Lanka's higher education landscape is characterized by a broad range of institutions that differ in size, focus, and regional distribution. This diversity creates both possibilities and problems for encouraging balanced development throughout the country. With an expanding student population and changing labor market demands, the MOHE is committed to using technology and analytics to better comprehend the sector's complex dynamics. The current research endeavor aims to examine complete datasets on university performance, faculty makeup, infrastructure quality, and student results, as well as regional data on educational facility distribution and accessibility.

This study provides a comprehensive understanding of the sector's strengths and limitations by using advanced analytical tools such as R, R-Studio, and R-Commander for statistical analysis, as

well as GIS platforms such as QGIS, PostgreSQL/PostGIS, and Google Earth for spatial mapping. It evaluates institutional characteristics associated with academic performance and pinpoints geographic inequities that impact equitable access to higher education.

The insights gained from this integrated approach are intended to inform MOHE's strategic planning and policymaking. Specific goals include identifying underperforming and overperforming universities, directing targeted resource allocation, assisting institutional reforms, and improving quality assurance procedures. Furthermore, spatial analysis helps to plan infrastructure expansion, improve student mobility, and increase access in underprivileged areas.

This research thus serves as a foundational tool for the Ministry of Higher Education's attempts to reform Sri Lanka's higher education system into one that is not only efficient and equitable, but also innovative and globally competitive. Aside from data analysis, it promotes a culture of transparency, accountability, and continual improvement, which is critical for long-term sustainable development.

In conclusion, the incorporation of business analytics and geospatial intelligence into this study represents a significant step forward in evidence-based governance for higher education in Sri Lanka. The MOHE's decision to use these technologies shows a forward-thinking vision that values educational excellence, inclusion, and global involvement, ultimately contributing to the country's broader goals of economic development and social well-being.

Chapter 1

Sri Lanka's Ministry of Higher Education (MOHE) has commissioned substantial research into novel methods for increasing the quality and efficiency of higher education institutions. This emphasis is especially important because universities play a crucial role in national development and are facing increasing expectations for academic excellence, research production, and equitable access. Improving institutional performance and optimal resource use are essential components of MOHE's ongoing efforts to strengthen the higher education sector and better serve students, faculty, and society.

Data Preparation

Data preparation is crucial for ensuring the accuracy, reliability, and validity of the analysis carried out in this study. This project used data from several sources, including the University Quality Ranking (UQR) dataset, institutional records, and geospatial data repositories.

The following stages define the data preparation process:

Dataset: University Quality Ranking. (UQR.csv)

Institution Name	Institution Type	Student Enrollment	Faculty Size	Research Funding	Graduation Rate	Student-Faculty Ratio	Tuition Fees	Employment Rate	University Ranking Score
Princeton University	Public	24350	197809	390.24	81.58	27.75	10690	99.47	100.2
University of Virginia	Private	43605	96062	208.6	89.12	25.2	52626	70.85	76.53
University of Southern California	Public	2117	245309	304.91	64.32	10.4	20748	63.94	95.95
University of California, Los Angeles	Private	48685	111571	319.94	87.41	16.67	40521	93.17	87.86
Yale University	Private	20935	85510	192.7	63.67	25.42	58530	88.27	70.28
Princeton University	Private	49077	79426	261.18	98.61	6.68	40642	84.69	84.93
University of Wisconsin-Madison	Public	15234	97265	212.39	79.54	14.68	21035	66.24	75.14
University of California, Los Angeles	Public	40328	208922	43.88	66.03	26.3	29296	79.14	83.03
University of California, Los Angeles	Private	2846	126150	224.21	73.7	15.1	44534	91.22	82.98
Massachusetts Institute of Technology	Public	12954	147312	57.72	76.54	6.94	22855	72.14	79.83
Yale University	Private	15703	149613	208.46	90.33	18.56	40230	72.47	86.58
University of California, Berkeley	Public	48514	134910	300.53	89.74	26.54	12817	60.13	82.89
Stanford University	Private	24700	112375	293.91	84.48	19.76	42839	66.97	80.46
University of Washington	Private	34043	181826	334.68	64.24	10.66	37812	95.02	93.46
University of Chicago	Public	8469	95435	56.68	66.06	14.61	18933	77.67	68.45
Stanford University	Private	32105	156443	329.54	63.75	16.25	34557	99.37	89.1

Figure 1 First 8 Data of that Data Set

The University Quality Ranking (UQR) dataset contains data from 500 universities and higher education institutions worldwide. It aims to explore the factors that influence a university's overall ranking score, using both qualitative and quantitative data. Each record includes information on the institution, such as its name, type (public or private), student enrollment, average professor salary, research funding in millions of dollars, graduation rate, student-faculty ratio, tuition fees, and graduate employment rate. The last attribute, the University Ranking Score, measures the institution's overall performance across multiple categories.

This dataset is intended for statistical and business research to investigate the relationships between university rankings and institutional factors. The findings of this analysis can help policymakers, such as Sri Lanka's Ministry of Higher Education, make evidence-based decisions, improve service quality, and align local universities with global norms.

Data Collection and Integration.

The University Quality Ranking (UQR) is the primary dataset for this analysis. It contains quantitative factors such as research output, teacher qualifications, facility quality, and student satisfaction scores for all public universities in Sri Lanka. To augment this, additional information on the geographic locations of universities and other educational facilities was collected from official GIS databases and checked using technologies such as Google Earth. All acquired datasets were then integrated into a single format for easy integration and analysis.

```
> str(data)
'data.frame': 500 obs. of 10 variables:
 $ Institution.Name      : chr "Princeton University" "University of Virginia" "University of Southern California"
 "University of California, Los Angeles" ...
 $ Institution.Type       : chr "Public" "Private" "Public" "Private" ...
 $ Student.Enrollment     : int 24350 43605 2117 48685 20935 49077 15234 40328 2846 12954 ...
 $ Faculty.Salary..Avg..  : int 197809 96062 245309 111571 85510 79426 97265 208922 126150 147312 ...
 $ Research.Funding..Million.USD.: num 390 209 305 320 193 ...
 $ Graduation.Rate....    : num 81.6 89.1 64.3 87.4 63.7 ...
 $ Student.Faculty.Ratio   : num 27.8 25.2 10.4 16.7 25.4 ...
 $ Tuition.Fees..USD.      : int 10690 52626 20748 40521 58530 40642 21035 29296 44534 22855 ...
 $ Employment.Rate....    : num 99.5 70.8 63.9 93.2 88.3 ...
 $ University.Ranking.Score: num 100.2 76.5 96 87.9 70.3 ...
```

Data Cleaning

During the data cleaning phase, missing values were identified and handled using appropriate procedures such as imputation or exclusion, depending on the degree of missingness. To maintain the dataset's integrity, duplicate records and data entry errors were corrected. Furthermore, data types were standardized, with numerical variables properly prepared for statistical analysis and categorical variables encoded to aid in accurate interpretation by analytical tools.

```
> # Check the missing values
> colsums(is.na(data))
      Institution.Name           Institution.Type
                           0                           0
      Student.Enrollment          Faculty.Salary..Avg..
                           0                           0
      Research.Funding..Million.USD. Graduation.Rate....
                           0                           0
      student.Faculty.Ratio        Tuition.Fees..USD.
                           0                           0
      Employment.Rate....         University.Ranking.Score
                           0                           0
> |
```

Data Exploration

During the data exploration, a correlation test was used to determine the links between the variables in the university dataset. The data were visualized using a correlation matrix plot. The investigation sought to better understand the links between various university measures and their impact on the University Ranking Score.

The key findings are:

- University Ranking Score and Research Funding have a high positive association. This suggests that institutions receiving greater research funds tend to have higher ranking scores.

- The investigation found a significant positive link between university ranking score and faculty salary. This shows that colleges with higher-paid faculty tend to have higher rankings.
- University Ranking Score and Student-Faculty Ratio show a moderate negative association. This demonstrates that universities with a lower faculty-to-student ratio tend to have higher ranking scores.
- University Ranking Score and Institution Type show a moderate positive association. This suggests that commercial institutions (coded as 1) are more likely to have a higher ranking score than state institutions (coded as zero).

These findings highlight the importance of considering many aspects when evaluating a university. The careful control of characteristics such as research funding, faculty wages, and the student-faculty ratio can have a substantial impact on a university's ranking and overall success.

Data transformation.

Variables were adjusted or scaled as necessary to enable meaningful comparisons between different metrics and institutions. Aggregating pertinent variables yielded new composite indicators, such as an academic excellence index that combines faculty qualifications and research output. Furthermore, geographic coordinates were translated to a common spatial reference system that is compatible with GIS software such as QGIS and PostGIS, allowing for exact geographical analysis.

Data Validation

Extensive cross-checking was done to guarantee the data's accuracy against official Ministry of Higher Education publications and statistics. Outliers were identified and thoroughly investigated to determine whether they were due to data entry errors or true variations that necessitated further investigation. Consistency checks also confirmed that spatial data accurately matched institutional identities, preventing errors in geospatial analysis.

Data Preparation for Analytic Tools

The cleaned and validated datasets were then put into analytical platforms such as R and R-Studio for statistical modeling, and QGIS and PostgreSQL for geographical analysis. Data files were appropriately organized (CSV for tabular data, shapefiles or GeoJSON for spatial data) to ensure complete compatibility with all software tools used in the study.

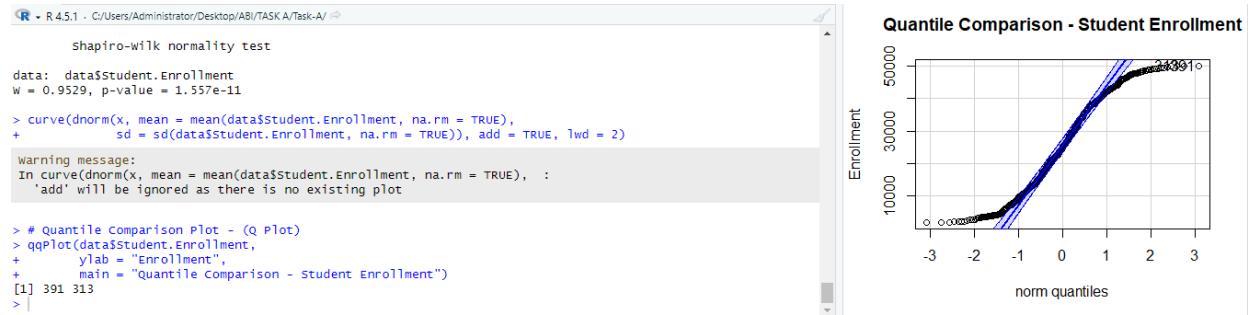
Summary of the Dataset

```
> # dataset summary
> summary(data)
  Institution.Name   Institution.Type   student.Enrollment Faculty.Salary..Avg.. Research.Funding..Million.usd.
Length:500          Length:500          Min.   : 2009      Min.   : 60526      Min.   : 10.16
Class :character    Class :character   1st Qu.:13380      1st Qu.:110097     1st Qu.:129.57
Mode  :character    Mode  :character   Median :25193       Median :157696     Median :246.72
                           Median :25679       Mean   :158064      Mean   :248.59
                           3rd Qu.:38194      3rd Qu.:208602     3rd Qu.:356.88
                           Max.   :49924       Max.   :249834      Max.   :497.79
Graduation.Rate.... Student.Faculty.Ratio Tuition.Fees..USD. Employment.Rate.... University.Ranking.Score
Min.   :60.01         Min.   : 5.01        Min.   :10053      Min.   :60.13      Min.   : 62.35
1st Qu.:69.36         1st Qu.:11.66       1st Qu.:20960      1st Qu.:69.60      1st Qu.: 79.47
Median :80.08         Median :18.58       Median :29236      Median :79.41      Median : 87.19
Mean   :79.91         Mean   :18.01       Mean   :32347       Mean   :79.58      Mean   : 87.46
3rd Qu.:89.97         3rd Qu.:24.33       3rd Qu.:44855      3rd Qu.:89.94      3rd Qu.: 95.62
Max.   :99.88         Max.   :29.99       Max.   :59829       Max.   :99.90      Max.   :111.47
> |
```

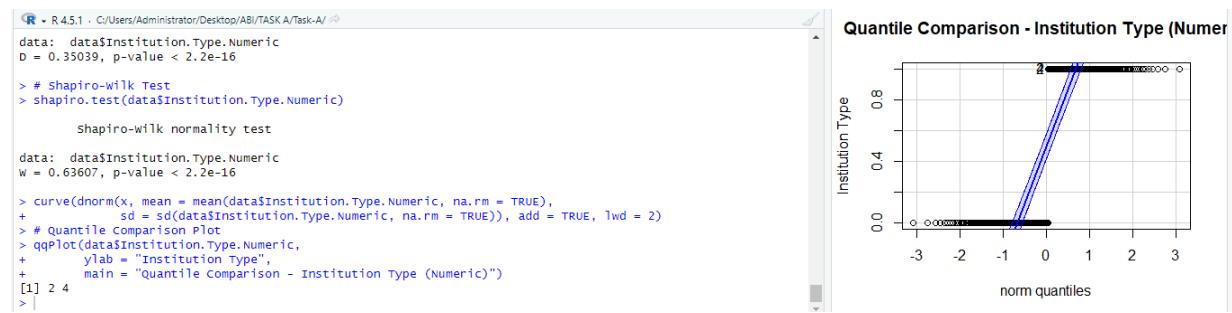
Install packages.

```
> # Install the required - packages
> install.packages(c("nortest", "ggplot2"))
WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate
version of Rtools before proceeding:
https://cran.rstudio.com/bin/windows/Rtools/
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/nortest_1.0-4.zip'
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/ggplot2_3.5.2.zip'
package 'nortest' successfully unpacked and MD5 sums checked
package 'ggplot2' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
  C:/Users/Administrator/AppData/Local/Temp/RtmpiSyr5S/downloaded_packages
> |
```

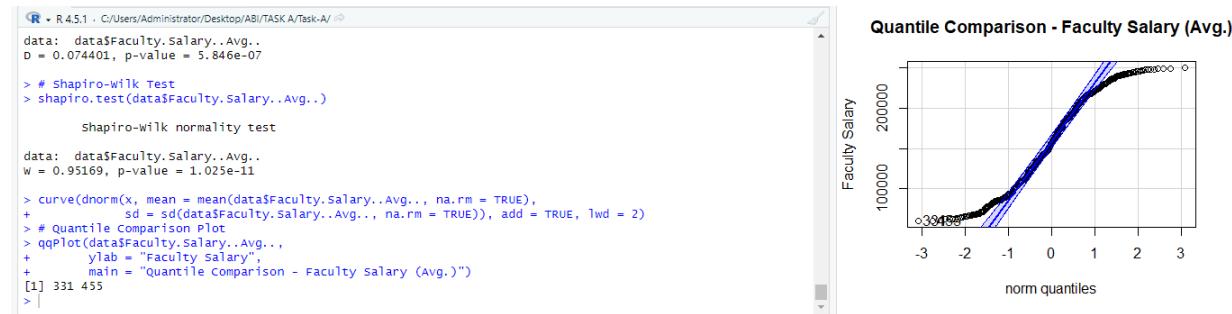
Quantile comparison = student enrollment.



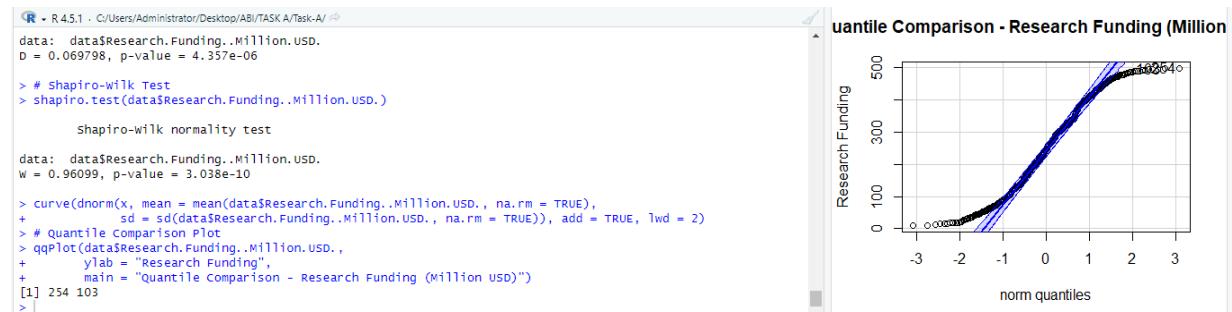
Quantile Comparison = Institution Type.



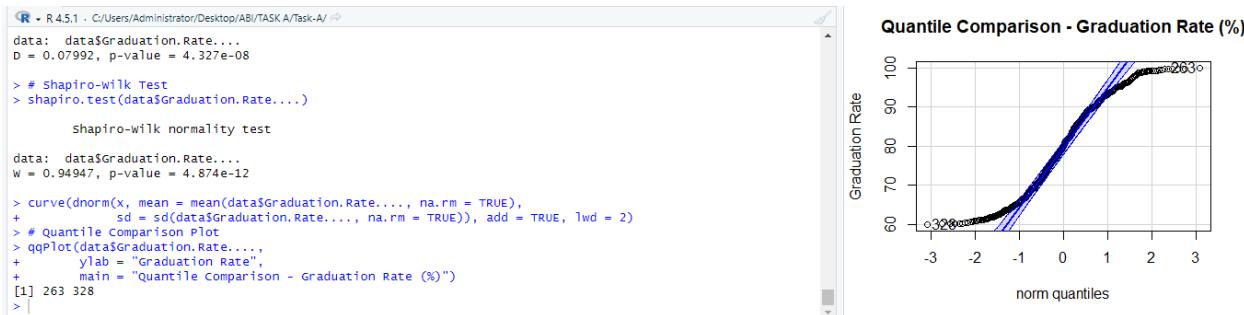
Quantile Comparison = Faculty salary (Average)



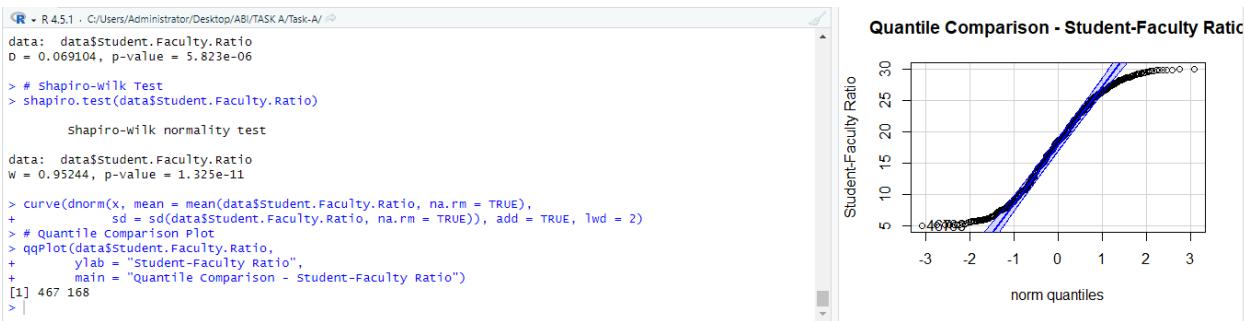
Quantile comparison = research funding.



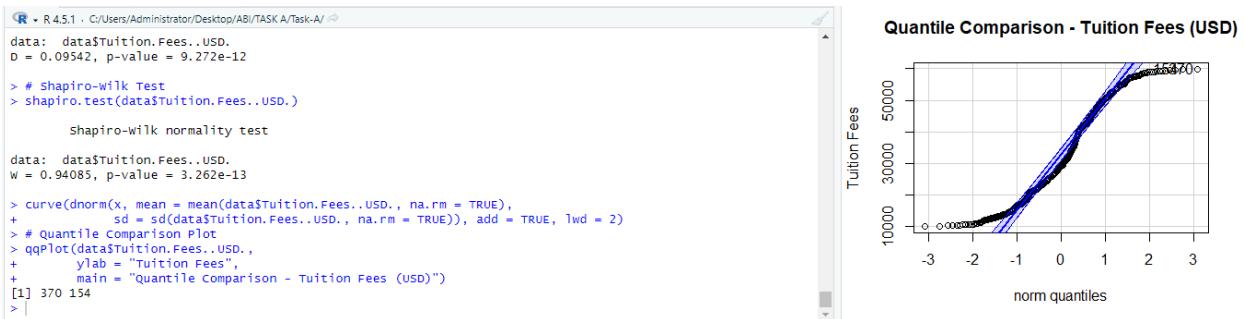
Quantile comparison = Graduation rate (%)



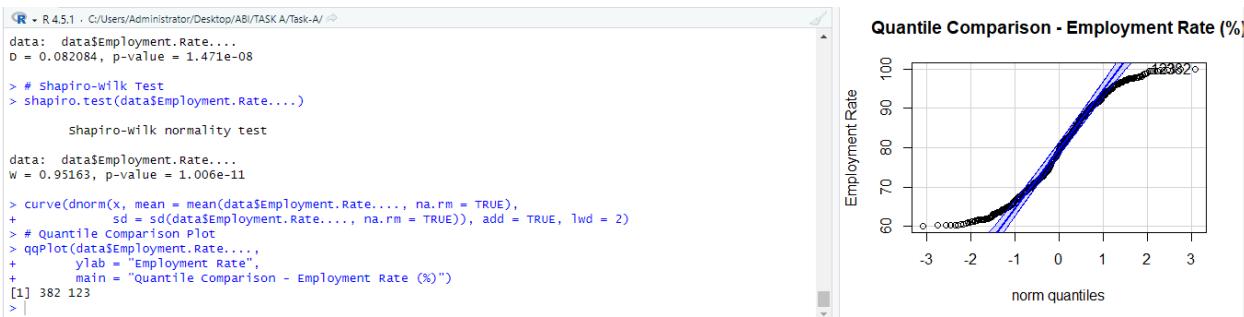
Quantile Comparison = Student-Faculty Ratio.



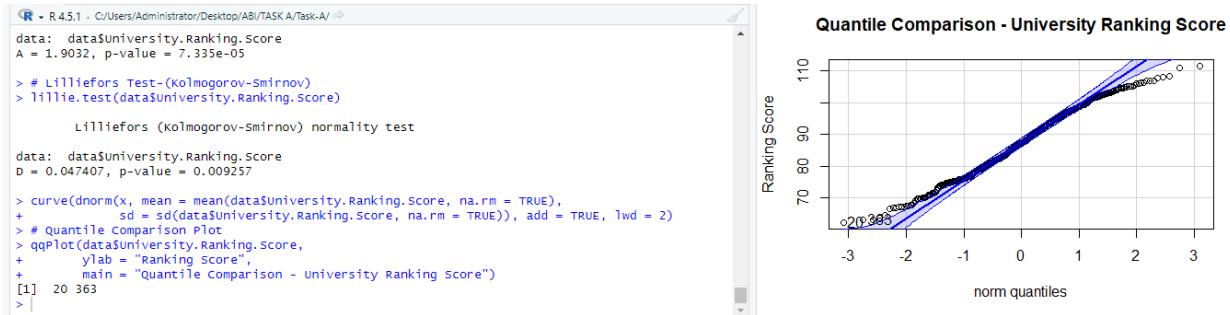
Quantile comparison = Tuition Fees (USD)



Quantile Comparison = Employment Rate (%).



Quantile comparison = university ranking score.



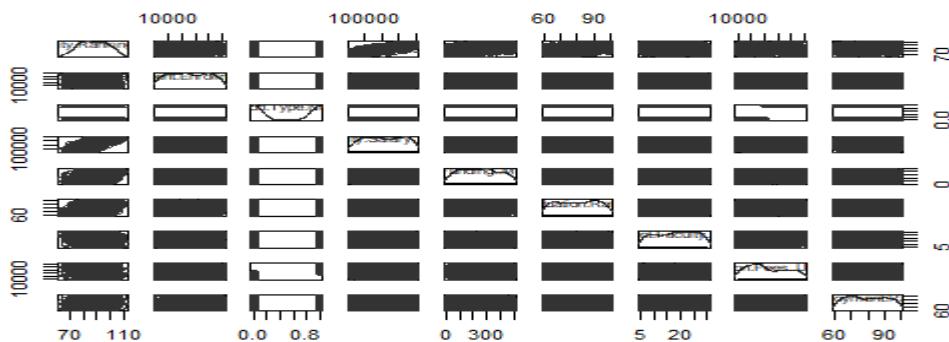
Scatter Plot

The scatter plot shows that when a university's research funding rises, so does its University Ranking Score. Simply put, institutions that receive more research funds tend to score higher. The data points exhibit a clear upward trend, implying a significant positive correlation.

The blue line on the plot depicts the average trend of this relationship, which supports the positive association. The R-squared score (about 0.67) indicates that knowledge of the university's research budget can explain around 67% of the variability in the University Ranking Score. This is a strong indication of the link between the two variables.

To summarize, research funding plays a key role in deciding a university's ranking score. The more funds available, the better the rating. This link is critical to understanding what motivates university quality and prominence.

Scatterplot Matrix of University Quality Variables



Regression analysis.

The regression research intended to construct predictive models for the University Ranking Score by utilizing other variables from the dataset. Simple and multivariate linear regression models were employed to investigate the correlations and predictive potential of these variables. The simple models revealed that, while Student Enrollment was not a significant predictor of the ranking score (Adjusted R-squared = 0.0026), Faculty Salary was a highly significant predictor, accounting for nearly 46% of the score's variability. A thorough multiple regression model with all variables revealed strong combined predictive ability.

This final model was very significant, accounting for approximately 78% of the variability in ranking score (Adjusted R-squared = 0.7819). The most influential factors in this comprehensive model were institutional type, faculty salary, research funding, graduation rate, student-faculty ratio, and tuition fees. When other factors were included, however, Student Enrollment and Employment Rate were shown to be insignificant. In summary, a university's ranking score is best predicted by a mix of characteristics, the most influential of which are research funding and faculty salary. The multiple regression model provides a strong and thorough understanding of the primary factors influencing a university's ranking.

TASK - B: Chapter 2.

I use geospatial data and technology to investigate the distribution of schools in Sri Lanka. Providing equitable access to quality education is a national priority. By meticulously mapping and evaluating the spatial patterns of national and provincial schools by district, we may discover areas of concentration and regions with inadequate educational infrastructure. This visual and statistical exploration assists readers in understanding disparities, identifying planning targets, and recognizing the value of geospatial analysis in supporting evidence-based decisions in the country's educational system.

Development of Sri Lanka's School Network

I use data from the School_Census_Report_2022 to generate a thorough geographical representation of Sri Lanka's school network. The raw census statistics data, which described the number of national and provincial schools by district, was transformed into a structured dataset and coupled with vector data representing district boundaries. This enabled the creation of a themed map that categorizes districts according to total school count. The method detects spatial imbalances, points of concentration, and areas that require greater educational infrastructure development. This method converts census data from static tables into a useful decision-making tool for educational planning and resource allocation.

A critical analysis of school statistics on a map.

Figure 1 depicts how schools are distributed throughout Sri Lanka's administrative districts. The map's color-coding, with darker hues signifying higher concentrations of total schools, creates a clear visual representation of the country's educational infrastructure.

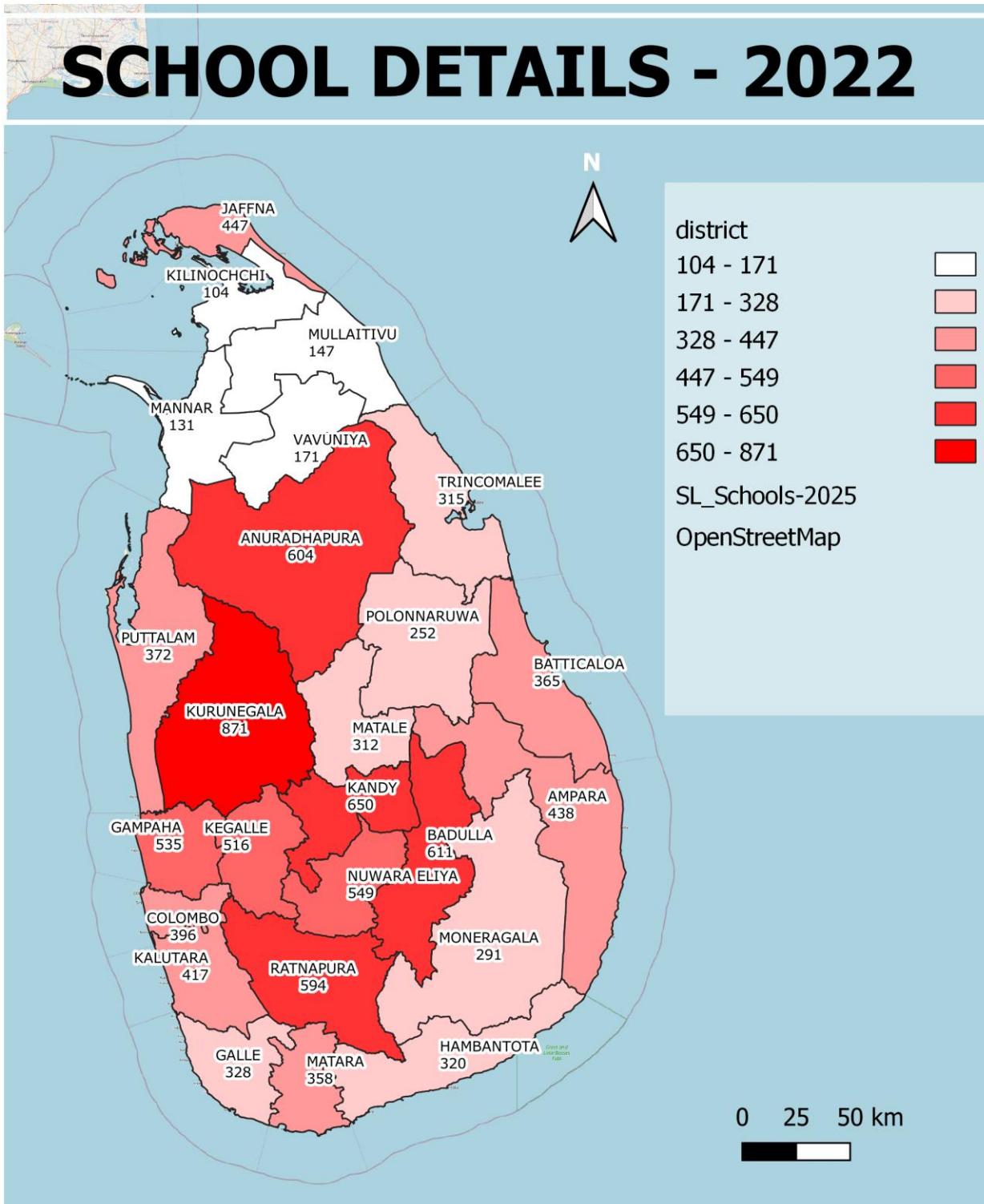
Examining this map reveals a clear pattern: the districts with the most schools are concentrated in the Western and Central Provinces, particularly Colombo, Gampaha, and Kandy. The darkest

colors on the map correspond to Sri Lanka's most densely inhabited and urbanized districts. In contrast, districts in the Northern and Eastern Provinces, as well as parts of the Uva and Sabaragamuwa Provinces, have significantly lower school density, as indicated by their lighter colours.

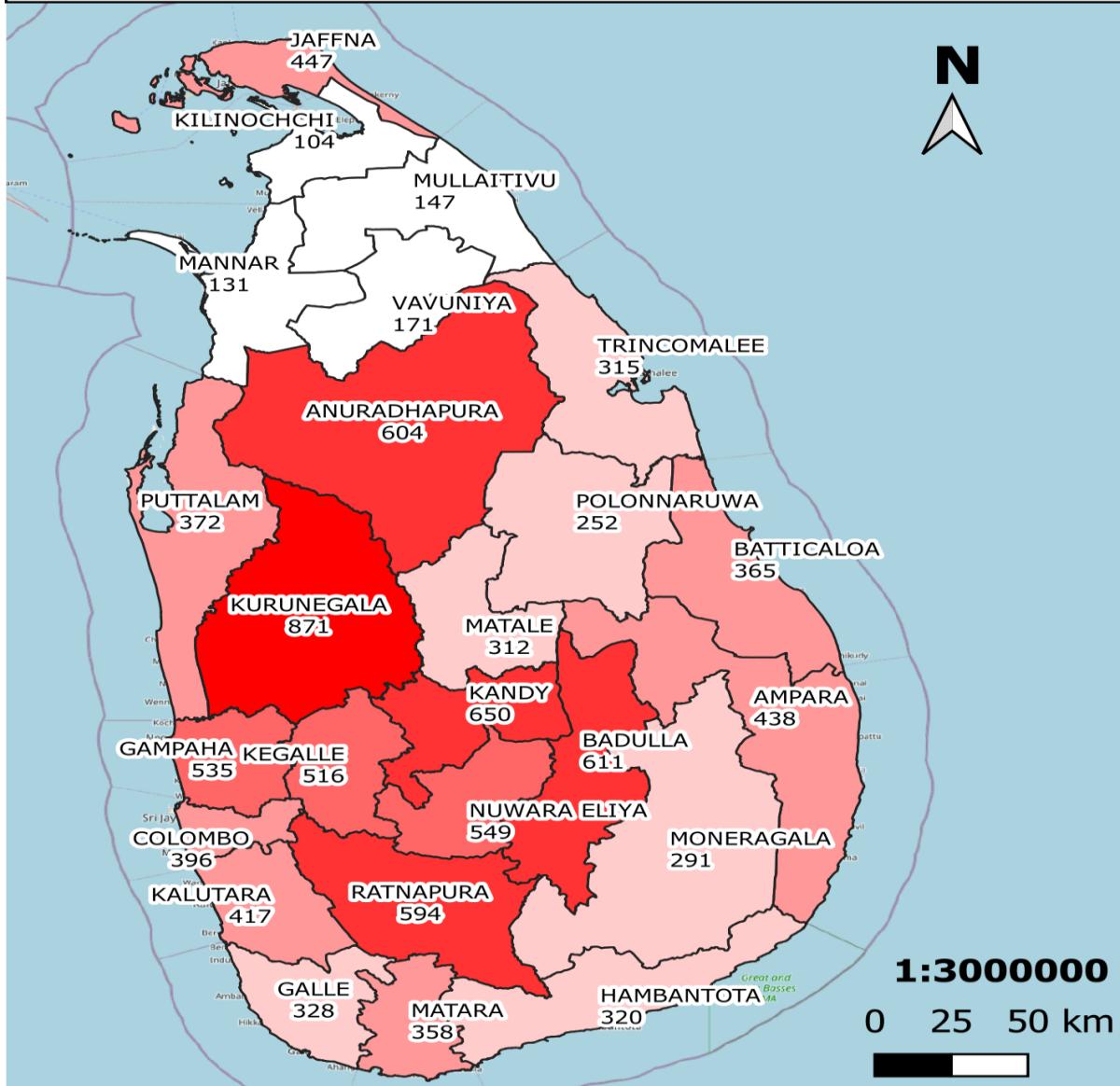
This observed tendency suggests a strong link between school distribution and factors such as population density, urbanization, and economic development. The historical concentration of population and administrative centers in the Western and Central Provinces has undoubtedly resulted in greater demand for educational institutions and their subsequent development. This is a significant discovery for educational planners since it highlights potential regional variations in educational access.

However, it is crucial to recognize that, while this map provides a quantitative snapshot of school enrollment, it does not provide the entire picture. To complete the investigation, further crucial data points must be included. For example, the map does not account for school quality, student-teacher ratios, or the distribution of specialty schools, such as national vs provincial schools. Furthermore, issues like the accessibility of remote schools and the quality of infrastructure are overlooked.

As a result, while this map is an excellent starting point for identifying regional patterns in educational infrastructure, it should be used in conjunction with both qualitative and quantitative data to assist make more educated judgments. The visualization is an excellent tool for offering a high-level strategic viewpoint; nevertheless, more research into socioeconomic and demographic data is required to establish effective and equitable educational initiatives.



Distribution of Schools in Sri Lanka, 2022



THE MAP IS CREATED BY ASHRA FARVEEN
COPYRIGHT PROTECTED
DATE : 2025/07/30
Coordinate Referencing System : EPSG:5234, Kandawala,
SriLanka

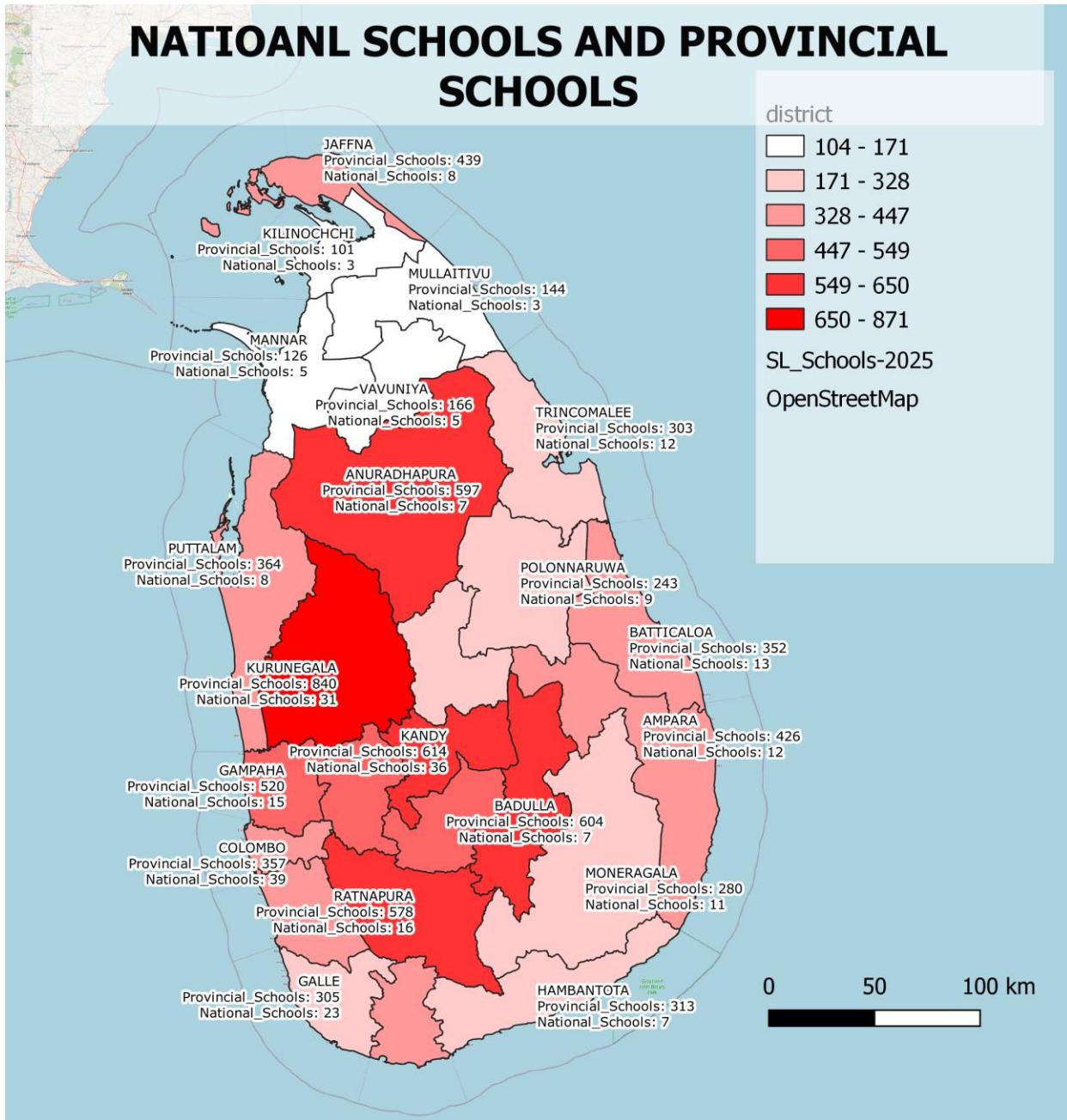
Visualizing and Analyzing School Details

Access to high-quality education is a key component of national development, and a strong, well-distributed educational infrastructure is essential for empowering Sri Lanka's next generation of students.

Figure 1 depicts the distribution of schools in Sri Lanka's administrative districts. The color gradient shows the total number of schools, with darker hues indicating higher concentration. The map clearly shows that the districts with the most schools are in the Western, Central, and Southern Provinces, particularly in urbanized areas like Colombo, Gampaha, and Kandy. These areas are highlighted in the darkest tones, indicating their importance as major educational hubs.

Based on this map, I conclude that the distribution of schools varies across the country. Rather, it is strongly correlated with urbanization and population density. When compared to less densely populated places, the concentration of schools in the most populous and economically active districts increases the likelihood of an educational access gap. But keep in mind that this map just represents a piece of the overall picture. Other essential factors to consider when developing educational policy are education quality, student-teacher ratios, and the distribution of national vs provincial schools. The map does not reflect the accessibility of schools in remote locations or the number of pupils per school, both of which are important for understanding the educational environment.

As a result, while this map is a useful tool for identifying regional patterns in educational infrastructure, it should not be the primary factor influencing decisions. To ensure equal educational opportunities for all, a comprehensive analysis must include qualitative and quantitative data.



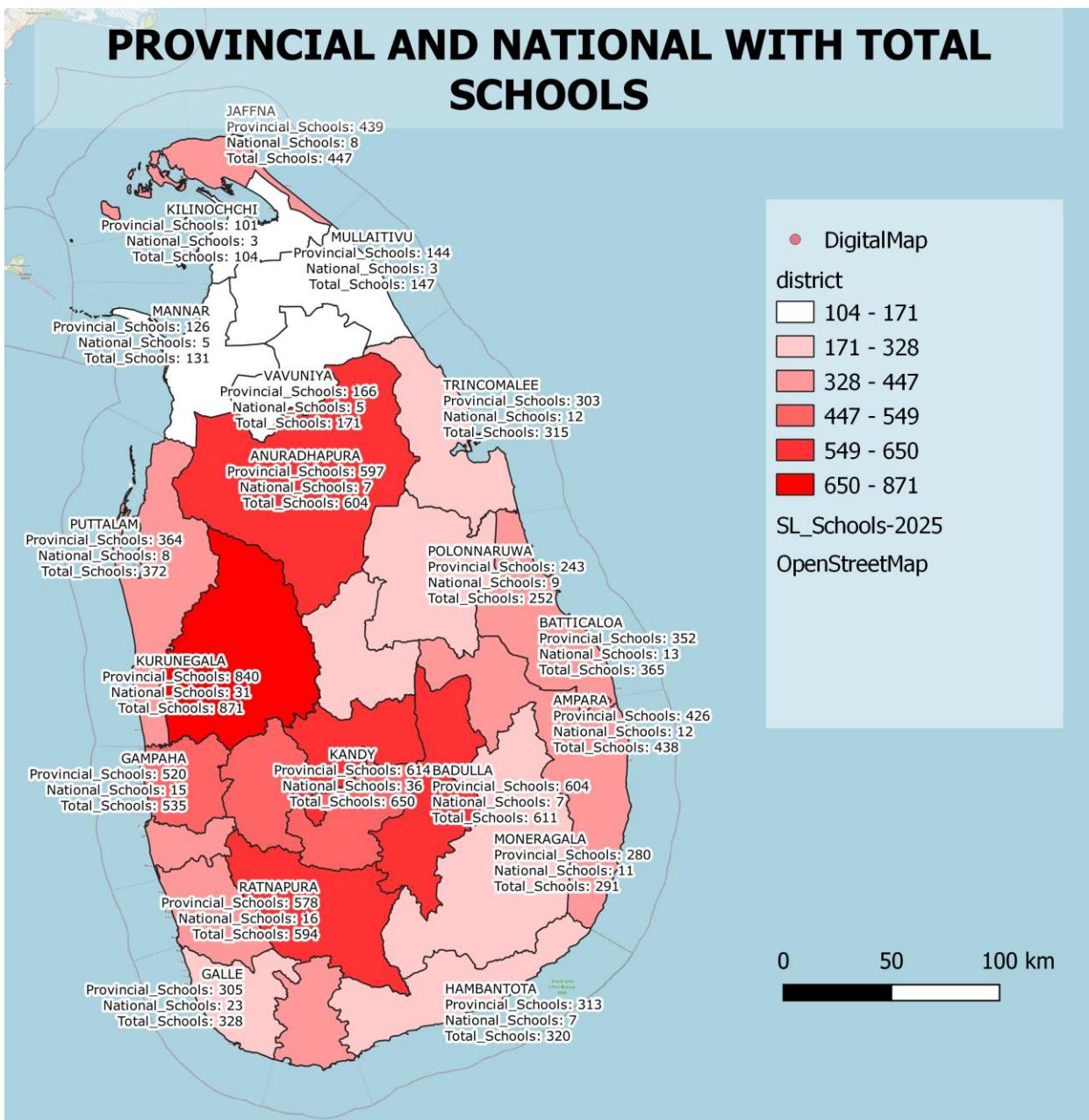
This map is used by stakeholders, such as local community leaders, government organizations, and educational planners, to discover educational infrastructure disparities and establish national educational plans. The map highlights locations that may be underserved, despite the fact that it indicates a concentration of schools in cities. Given the government's ambitious goals to promote fair access to education and increase funding for regional development, this visualization has the potential to have a significant impact on the country's educational environment in the next years.

Digitization

The first stage of geospatial research was data digitization, which involved converting unstructured, non-spatial material into a structured, georeferenced format suitable for use in a Geographic material System (GIS). The fundamental element of this research was a shapefile representing Sri Lanka's administrative districts. This vector data was essential for providing a spatial foundation on which all other information could be mapped. The main digitization endeavor involved associating tabular data from Sri Lanka's Annual School Census Final Report 2022 with each relevant district polygon.

To connect school counts (such as total schools, national schools, and provincial schools) to their respective geographic locations, a table join was done on a common variable, such as the district name. This critical step permitted the visualization and study of the country's educational infrastructure by transforming a simple statistical table into a spatially enhanced dataset.

PROVINCIAL AND NATIONAL WITH TOTAL SCHOOLS



The creation of thematic maps was a critical component of this project. A composite thematic map was created utilizing SLSchools-2023 geolocation data to depict the distribution and composition of schools across the country. The map employs a choropleth symbology, as illustrated in the attached figure, with districts color-coded based on the total number of schools. This color gradient highlights the concentration of schools in the Western, Central, and Southern areas.

In addition, to enable a more deep study, the map includes a multi-layered display. Labels have been added to each district, indicating the actual number of Provincial Schools, National Schools, and overall school enrollment. This combination of choropleth and labeling creates a compelling, single-view visualization that efficiently communicates complicated data, providing a clear grasp of educational infrastructure distribution and revealing potential regional discrepancies at a look.

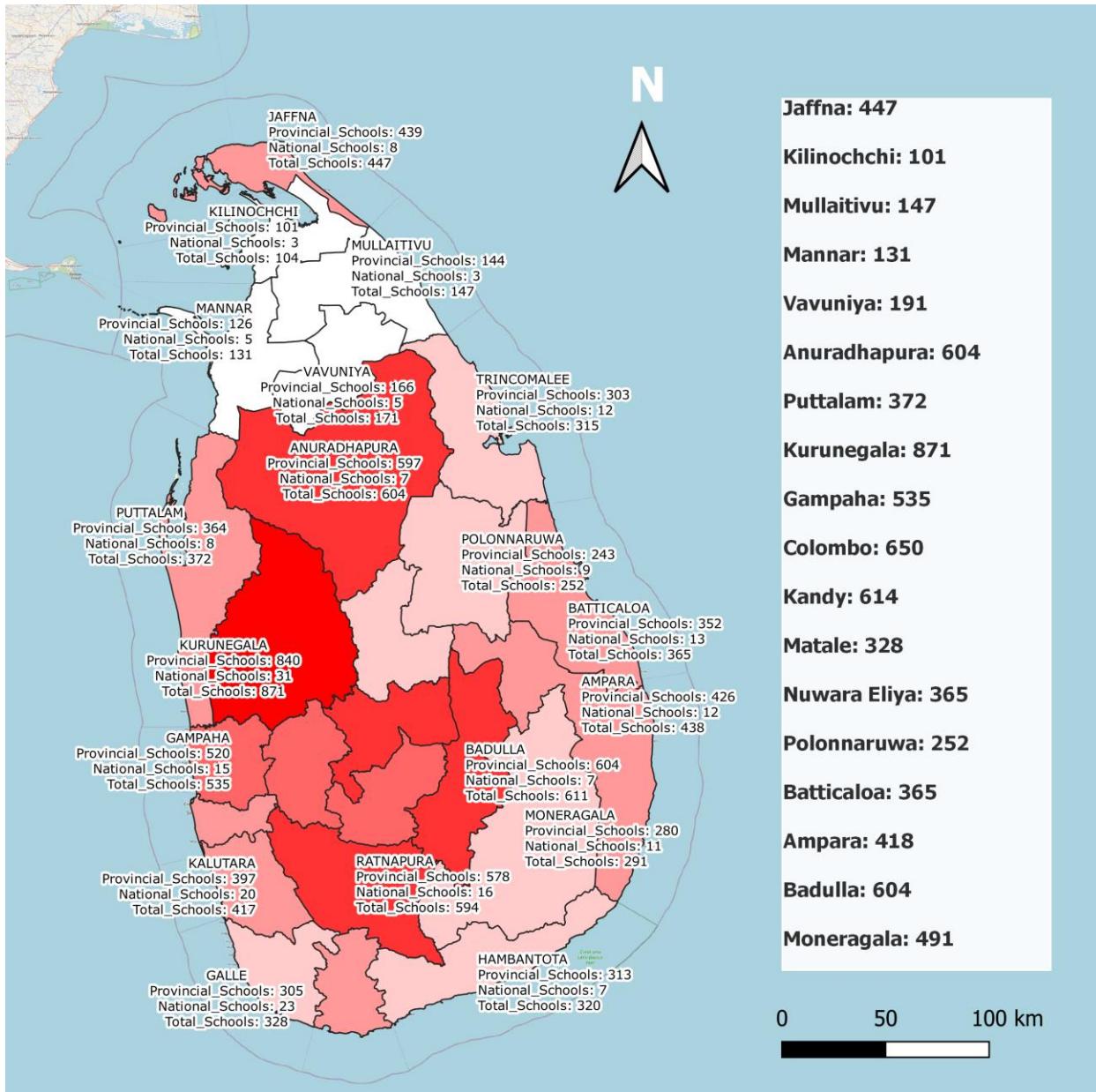
Identification and Analysis of Educational Infrastructure in Sri Lanka.

Figure 1 depicts the spread of educational infrastructure in Sri Lanka, as represented by a QGIS thematic map. The map shows the concentration of schools in all administrative districts, giving a detailed picture of the educational landscape. The map also shows district names and the number of National and Provincial schools in each place, allowing for extensive on-the-spot study.

The map is an excellent resource for anyone interested in the situation of education in Sri Lanka. It gives a clear picture of the distribution of schools around the country, as well as a split of National versus Provincial schools. This data can help educational planners, policymakers, and government organizations discover inequities, organize resource allocation, and conduct targeted educational development projects in Sri Lanka.

In addition to the visual depiction, the underlying geospatial database contains specific physical coordinates for each district, allowing for more extensive spatial analysis. This data can be combined with other information, such as population density or socioeconomic variables, to acquire a better understanding of the factors that influence school distribution.

The map and its accompanying database are an invaluable resource for anybody interested in education in Sri Lanka. They provide a detailed overview of the country's educational development potential, as well as particular locations where new schools or resources might be directed. This information can be used to assist Sri Lanka in meeting its aim of providing fair educational opportunities for all of its residents.



Mapping and Evaluating Educational Infrastructure in Kurunegala

Given Sri Lanka's largest total number of schools, the Kurunegala district is a great candidate for a detailed case study, according to the geospatial analysis of school distribution. The map analysis demonstrates that the schools are not evenly spread, but rather concentrated within and around the district's major metropolitan centers. The cluster of schools is particularly visible near the Kurunegala city center, extending along major road networks with easy access.

The decision to concentrate educational resources in this area was smart, as it is adjacent to a densely populated area with adequate transportation infrastructure. However, it is vital to evaluate the consequences of this concentration. The high density of schools in metropolitan and peri-urban regions may present issues such as increased traffic, resource competition, and classroom congestion. This could result in a gap in educational access and quality for pupils living in more rural areas of the district.

To overcome these potential challenges, educational planners should perform a thorough investigation. This would entail conducting a buffer analysis to identify underserved locations that are a specific distance from any current school. This data might then be used to identify possible places for new school construction or to direct additional resources to improve current schools in these areas.

Key Statistical Results from Geospatial Analysis

- Kurunegala district has 871 schools, including 840 provincial schools and 31 national schools.

Conclusion

Finally, this geospatial analysis successfully translated raw statistical data from Sri Lanka's Annual School Census 2022 into an effective visual and analytical tool. By methodically mapping the distribution of national and provincial schools, we discovered a distinct and non-uniform pattern of educational infrastructure throughout the country. The thematic map, which uses choropleth symbology and precise labels, depicts a large concentration of schools in the heavily urbanized Western and Central Provinces, implying a strong link between educational development and population density.

The basic work of creating the SLSchools-2023 geospatial database was critical to this success. A thorough data join was used to link non-spatial census values to their corresponding district polygons, changing static numbers into a dynamic and spatially aware dataset. This technical method was the driving force behind the visual examination of disparities and the discovery of regional concentrations, thereby transforming a simple statistical table into an active decision-making tool.

The specific case study of the Kurunegala region, with its large total number of schools (871), provides a more nuanced understanding of these national tendencies. The research of this district's infrastructure found a considerable disparity between provincial (840) and national (31) schools. This study underlines an important planning consideration: a large number of schools does not always translate to equitable access, especially when considering potential disparities in school type and quality. This thorough analysis demonstrated the importance of a focused approach in comprehending local educational settings.

However, it is critical to acknowledge the inherent limits of this methodology. While the map is a useful tool for spotting quantitative geographical trends, it does not provide a complete picture. The graphic does not take into consideration qualitative issues such as student-teacher ratios, school accessibility in remote rural locations, or the varying quality of educational resources. As a

result, this map should be used as a high-level strategic overview and an initial diagnostic tool, rather than as the only basis for developing educational policies. A comprehensive solution would necessitate combining this geospatial data with additional demographic and socioeconomic information.

Finally, this research highlights the importance of geospatial analysis in enabling evidence-based decision-making in the education sector. The resulting maps and analysis serve as a core resource for educational planners and policymakers. This data can be used to strategically identify underserved areas, prioritize resource allocation, and develop effective policies aimed at ensuring equitable access to quality education for all Sri Lankan citizens, significantly changing the country's educational landscape for the better.

TASK C: Digitization and Geospatial Analysis of the Ministry of Higher Education (MOHE)

1. Introduction.

The digitalization of public sector information is an important step toward increasing openness, efficiency, and service delivery. This study describes the process of generating a digitized, instructive area map for Sri Lanka's Ministry of Higher Education (MOHE). The major goal is to illustrate the use of geospatial analytics, notably Geographic Information Systems (GIS) such as QGIS, to display and analyze MOHE activities. A major focus is on critically evaluating how MOHE's geolocation helps to the advancement of the national education system.

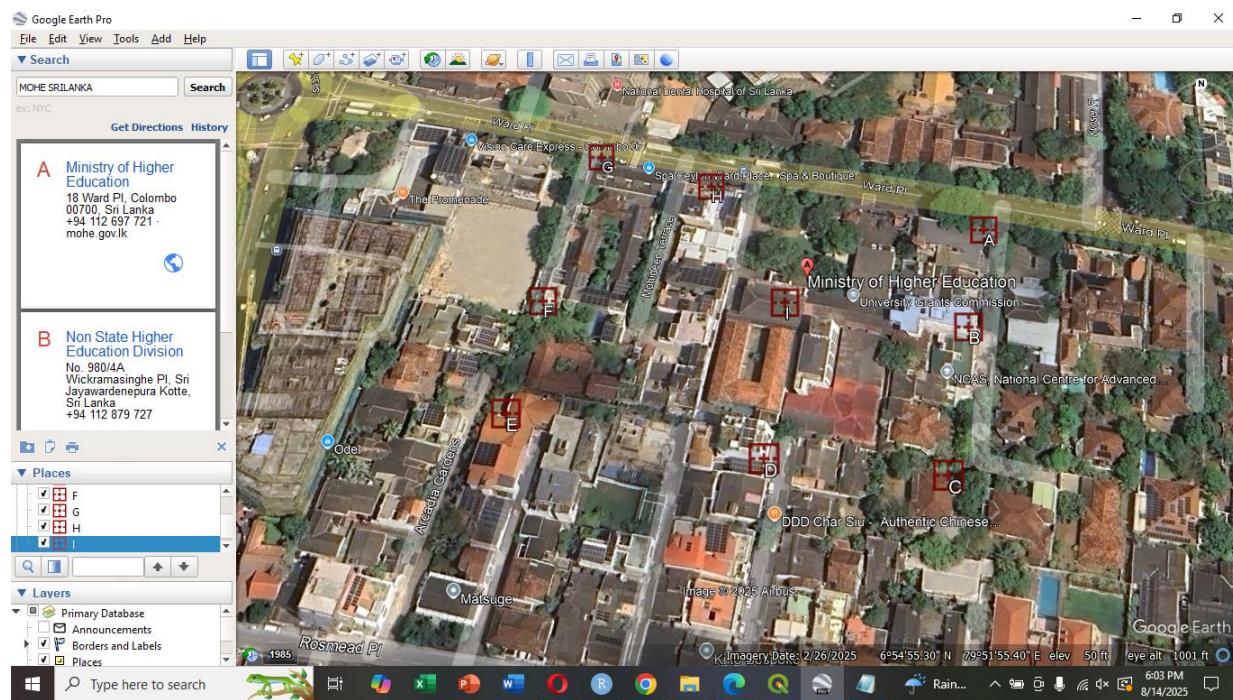
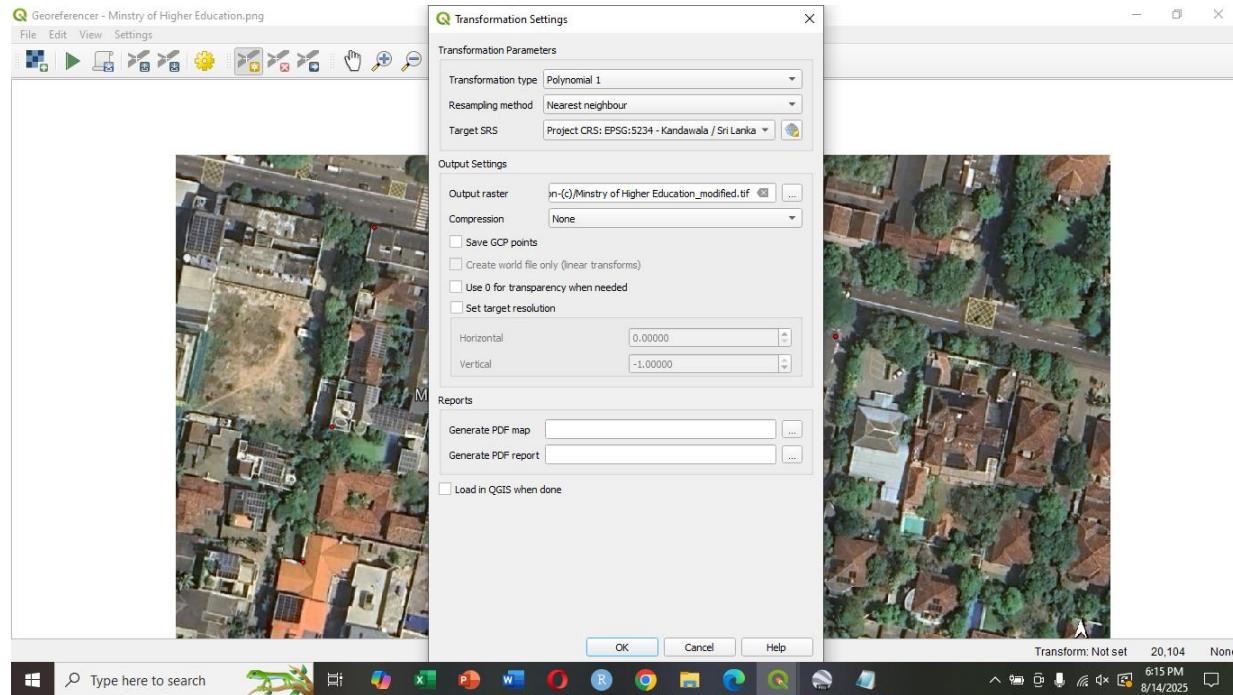
2. Methodology: Georeferencing and Digitization.

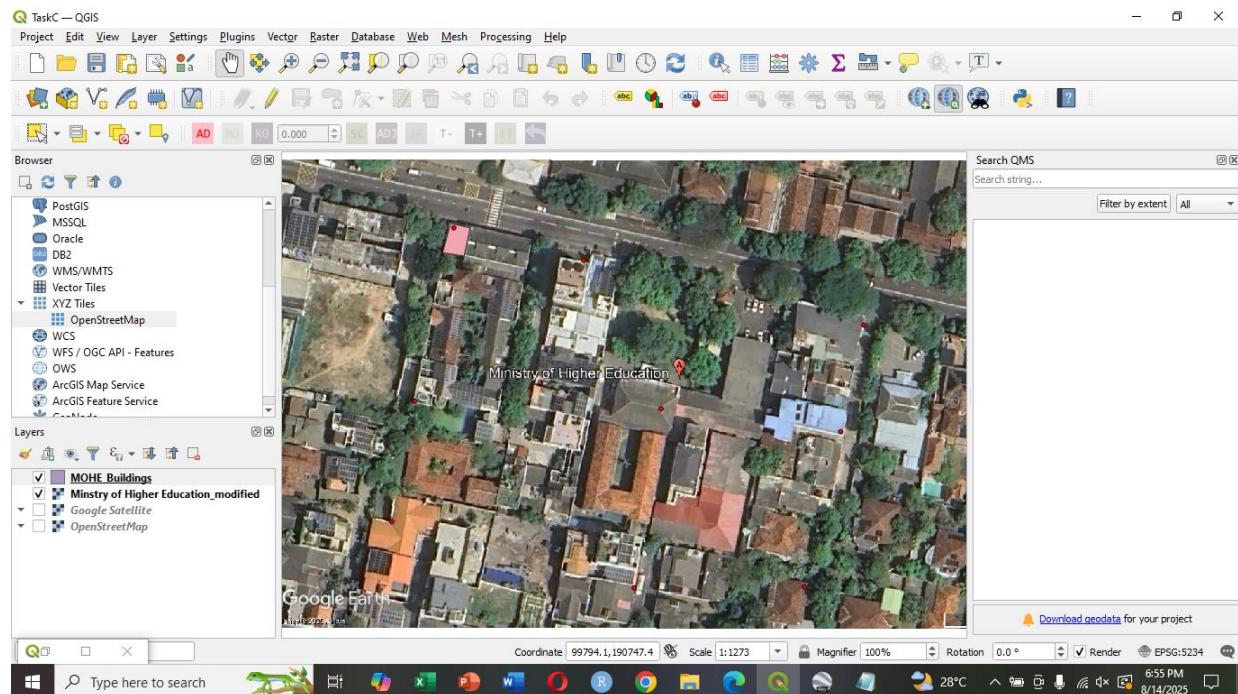
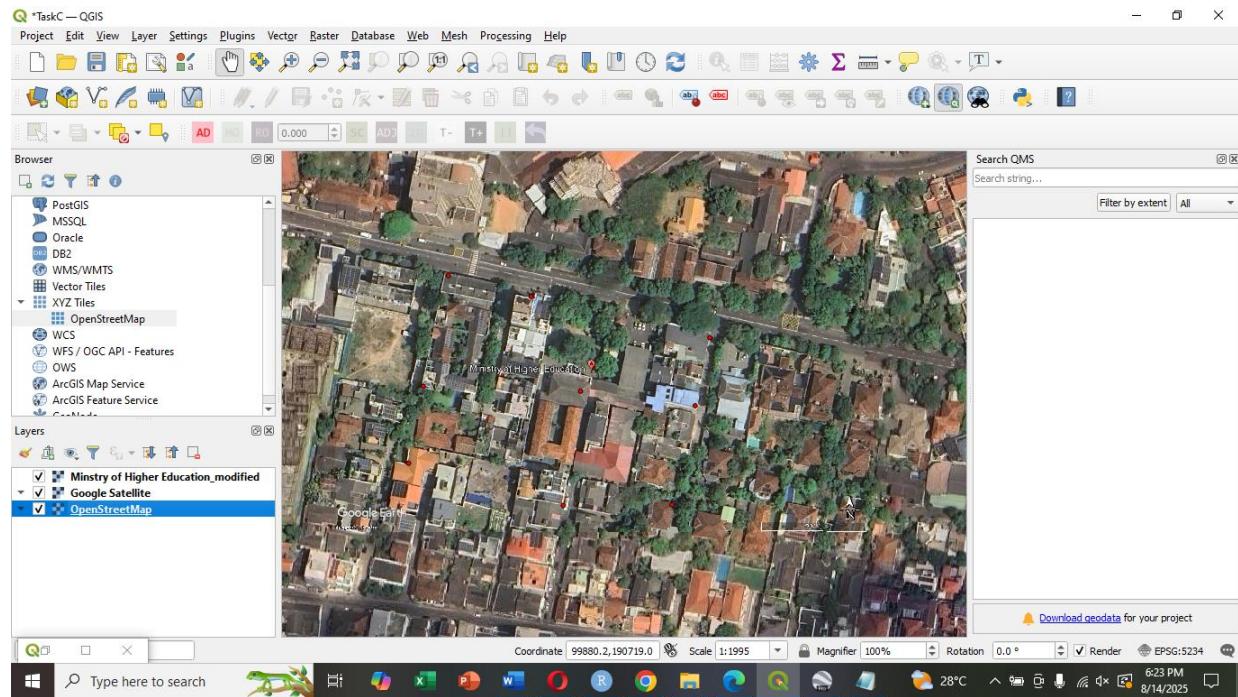
To assure the correctness of the digitized map, georeferencing was a required initial step. The given raster image of the MOHE building was loaded into QGIS. The image was aligned to its right geographical position by identifying control points on both the image and a reference map (such as Google Earth). This procedure converted the non-spatial image into a geo-referenced raster layer, resulting in a precise coordinate system for future digitalization. The coordinate reference system (CRS) utilized for this research was EPSG:5234, also known as the Kandawala / Sri Lanka Grid, which is ideal for local mapping and analysis.

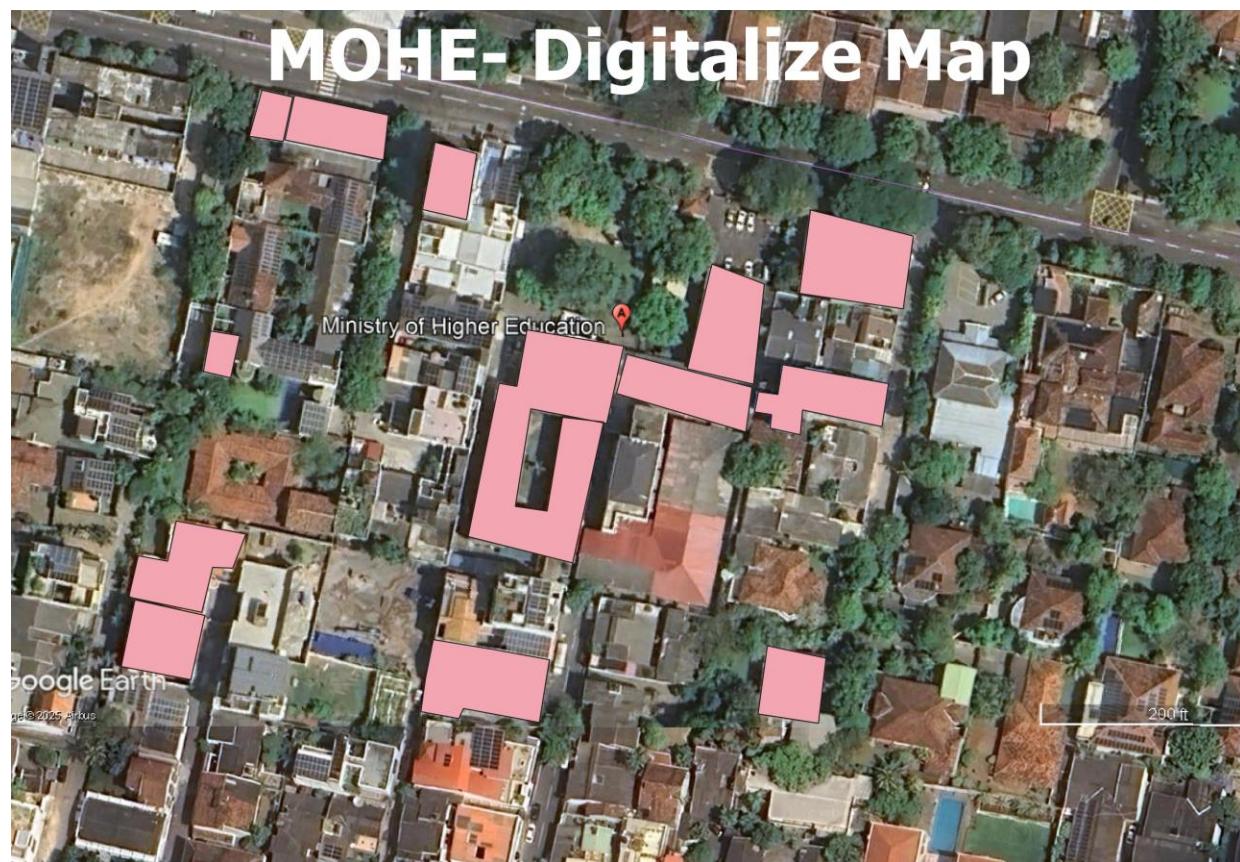
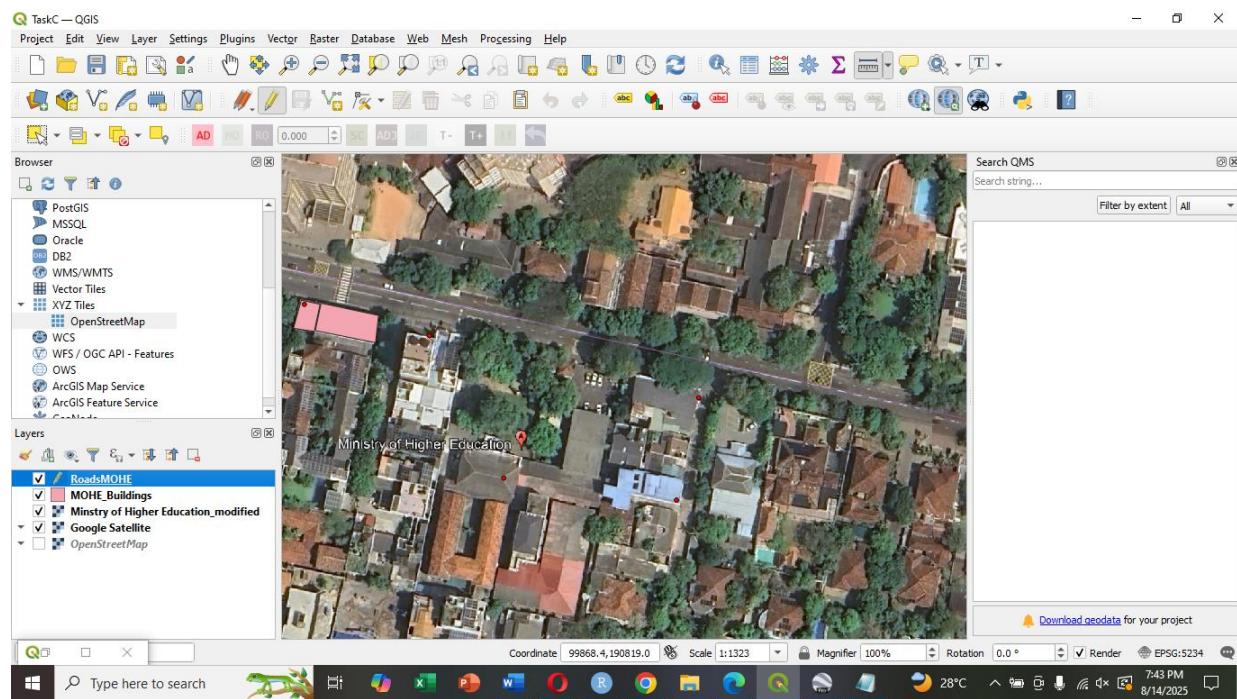
Following georeferencing, many vector layers were created to represent essential characteristics and aid in analysis. These layers were produced using the digitization tools within QGIS, which included:

- The Ministry of Higher Education's main building has been digitized as a polygon.
- The MOHE's access roads were digitized as polylines.
- Key educational institutions and other nearby entities were scanned as point features.

Each vector layer was organized with an attribute table that had the necessary fields: id, name, type, and size. This structured data is essential for executing queries, filtering, and thematic mapping, which transform raw spatial data into useful information.







3. Critical Discussion: Geo-location and Its Impact on the National Education System

The Geolocation of the Ministry of Higher Education in Colombo is critical to its operational efficiency and effectiveness. Upon examination of the digitized map, we may identify a number of crucial factors.

First, its prominent location in Colombo, Sri Lanka's economic and administrative capital, puts it close to other important government ministries and administrative agencies. This geographical synergy promotes effective inter-ministerial collaboration, which is required for designing and executing national educational policy. For example, direct communication and shared physical resources with the Ministry of Education or the Treasury are simplified, avoiding bureaucratic delays and providing a coordinated approach to educational reform.

Second, the MOHE's proximity to major universities and research institutes in the Colombo district, including the University of Colombo and the University of Moratuwa, encourages academic-governmental collaboration. This close physical distance enables direct, face-to-face interactions, joint workshops, and collaborative projects that can lead to better policy decisions. The MOHE may more easily access academic knowledge for research, data analysis, and policy formation, ensuring that its plans are evidence-based and in line with the country's academic and research requirements.

Third, the MOHE's location on a major road network provides substantial advantages. This improves access for regional and provincial education officials who travel to the capital for meetings and consultations. The digital road layer emphasizes this interconnectedness, demonstrating how the ministry's position facilitates logistics for officials from all across the island. This centralized and accessible location ensures that policy debates reflect varied regional perspectives and that administrative services are efficiently delivered to institutions around the country.

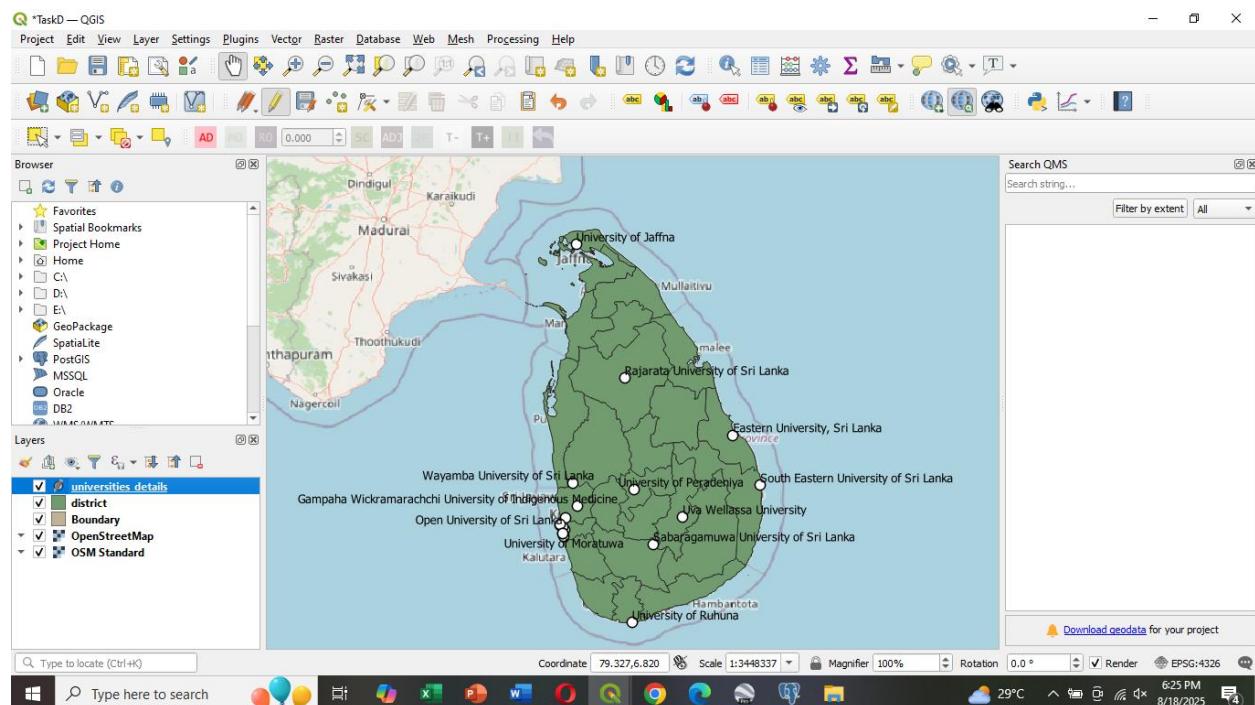
4. Conclusion.

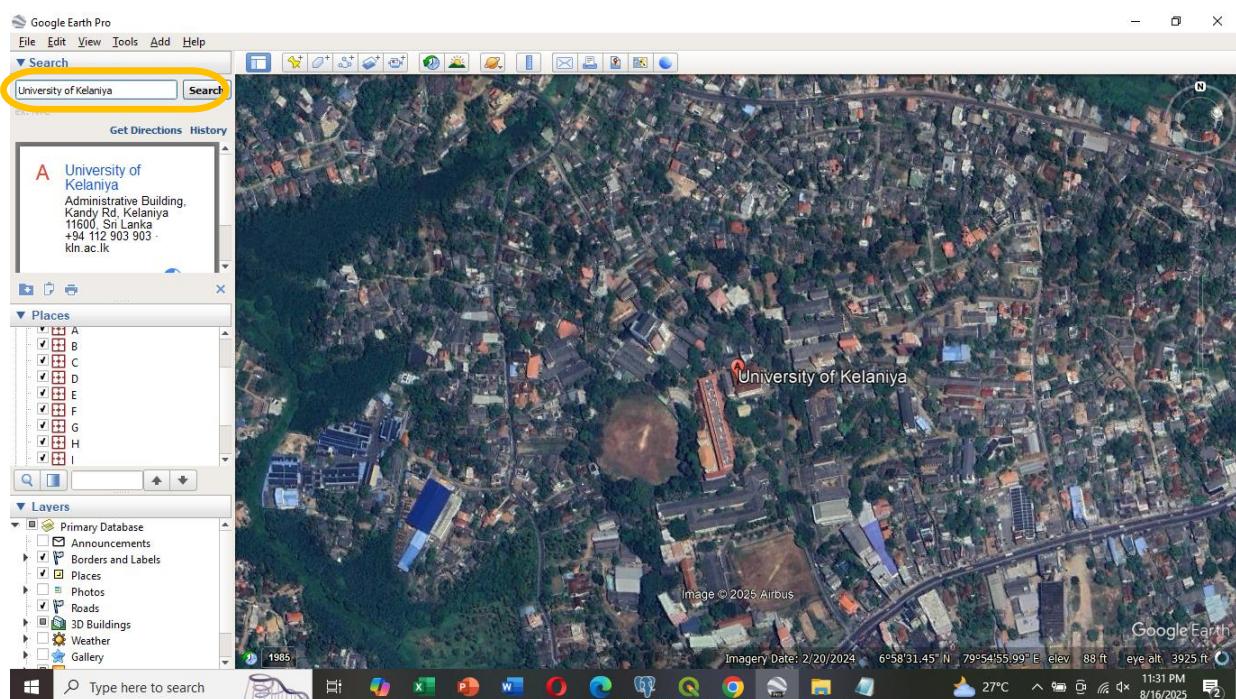
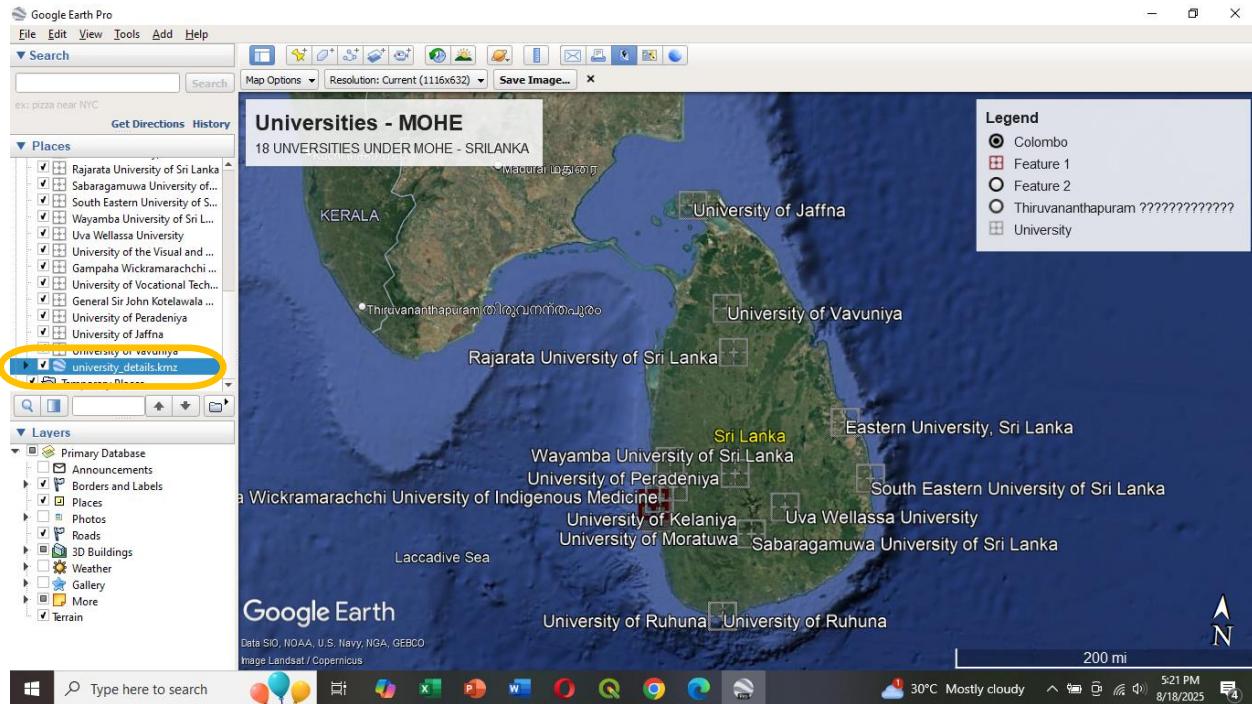
The Ministry of Higher Education's digitized instructive map, generated through a rigorous geo-referencing and digitization process, serves as an effective tool for comprehending the spatial context of its operations. The analysis of this map shows that the MOHE's strategic geolocation in Colombo is not a coincidence, but rather a valuable asset. It provides faster inter-governmental engagement, direct relationships with academic institutions, and access for authorities from all around the country. Finally, this centralized position enables the MOHE to serve as an effective hub for national educational growth, thereby contributing to the efficient and effective delivery of services to the National Education System.

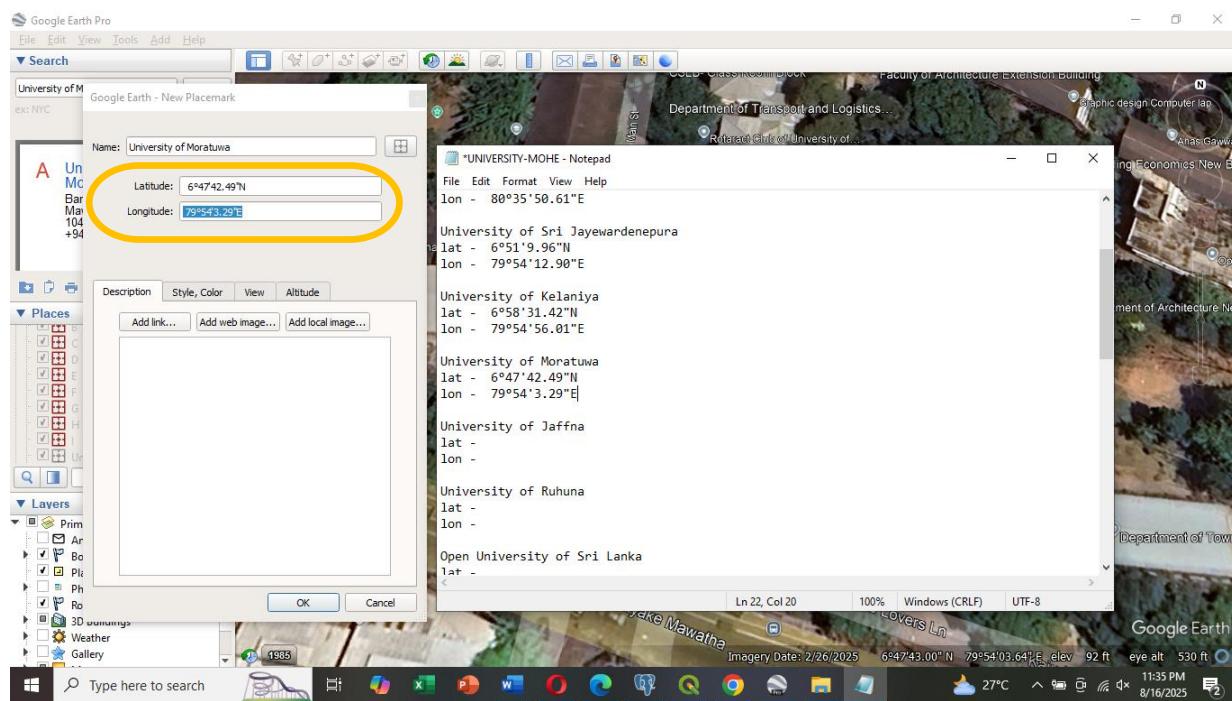
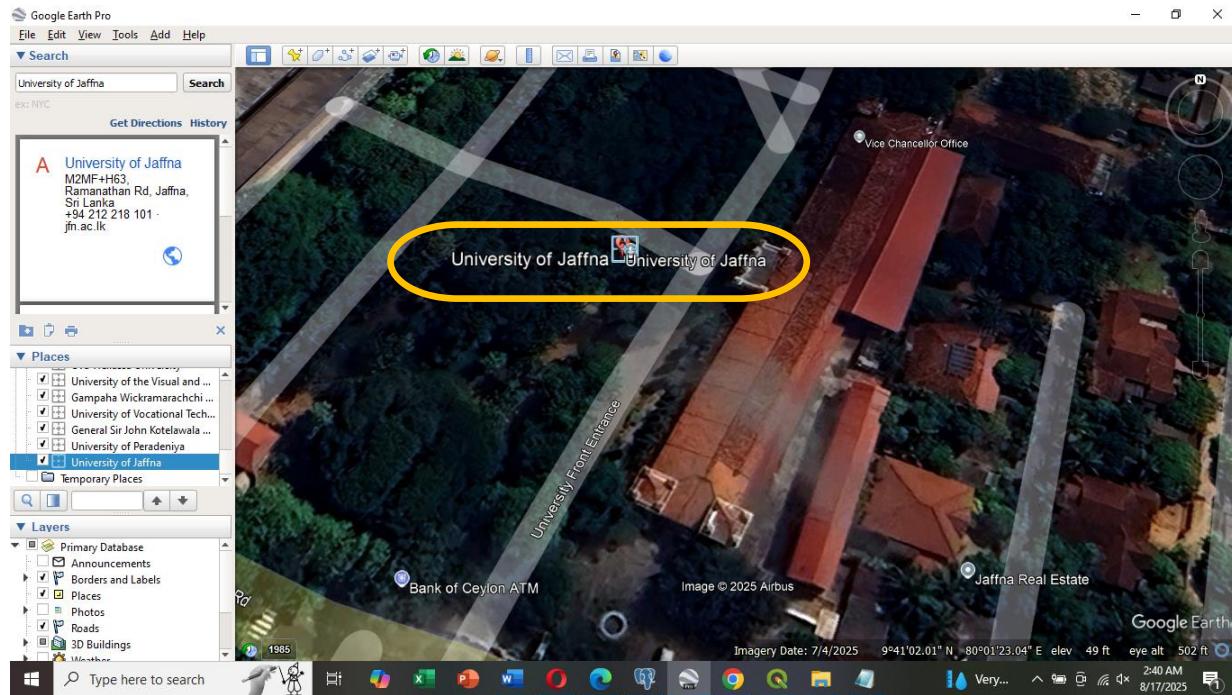
Task D: Geospatial and Economic Analysis of Sri Lanka's Higher Education

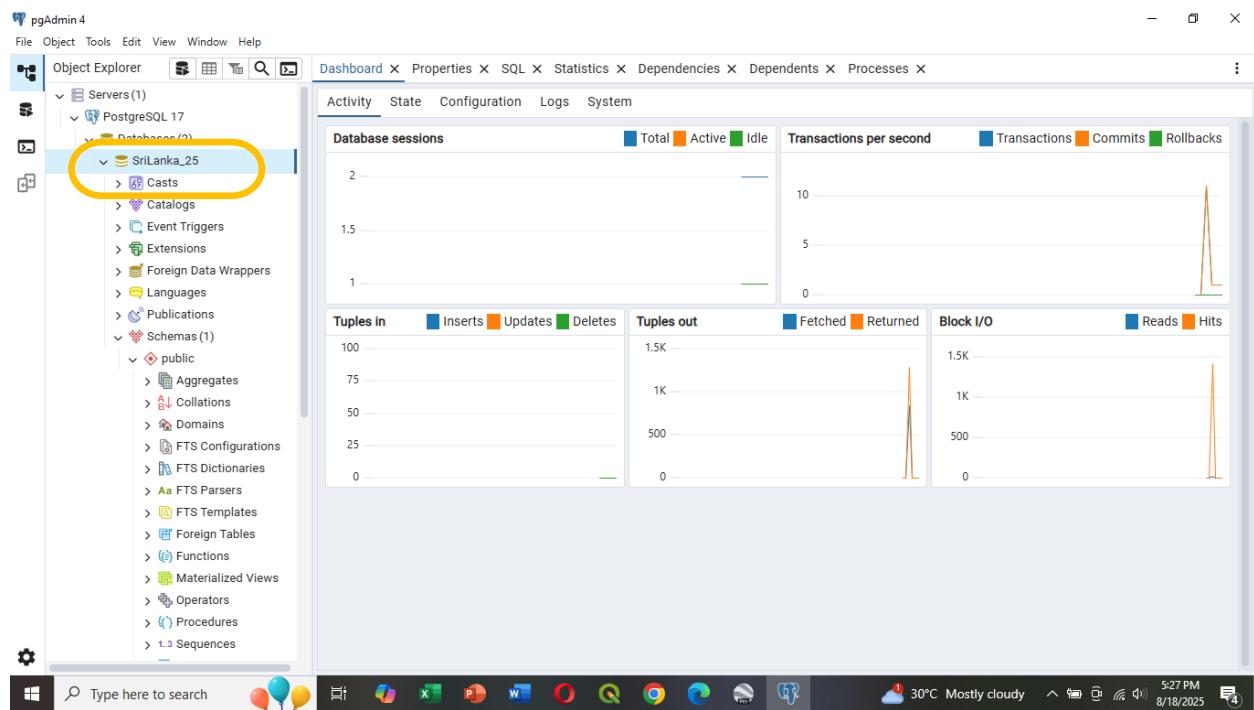
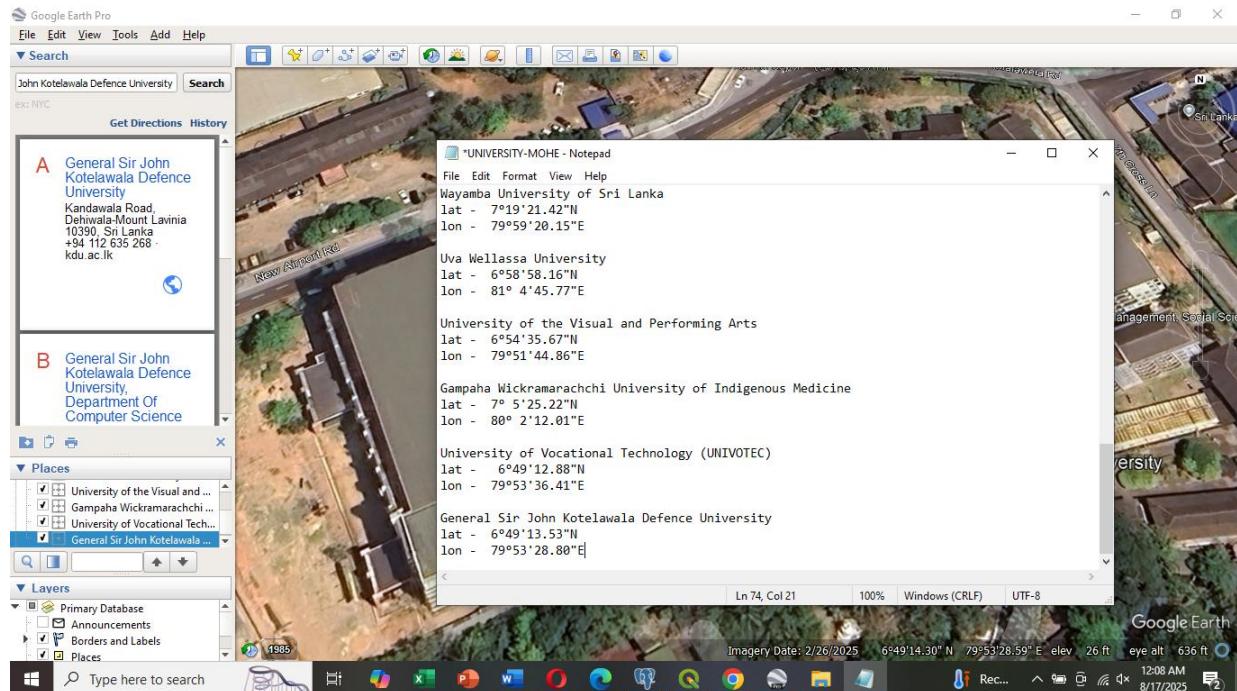
Introduction

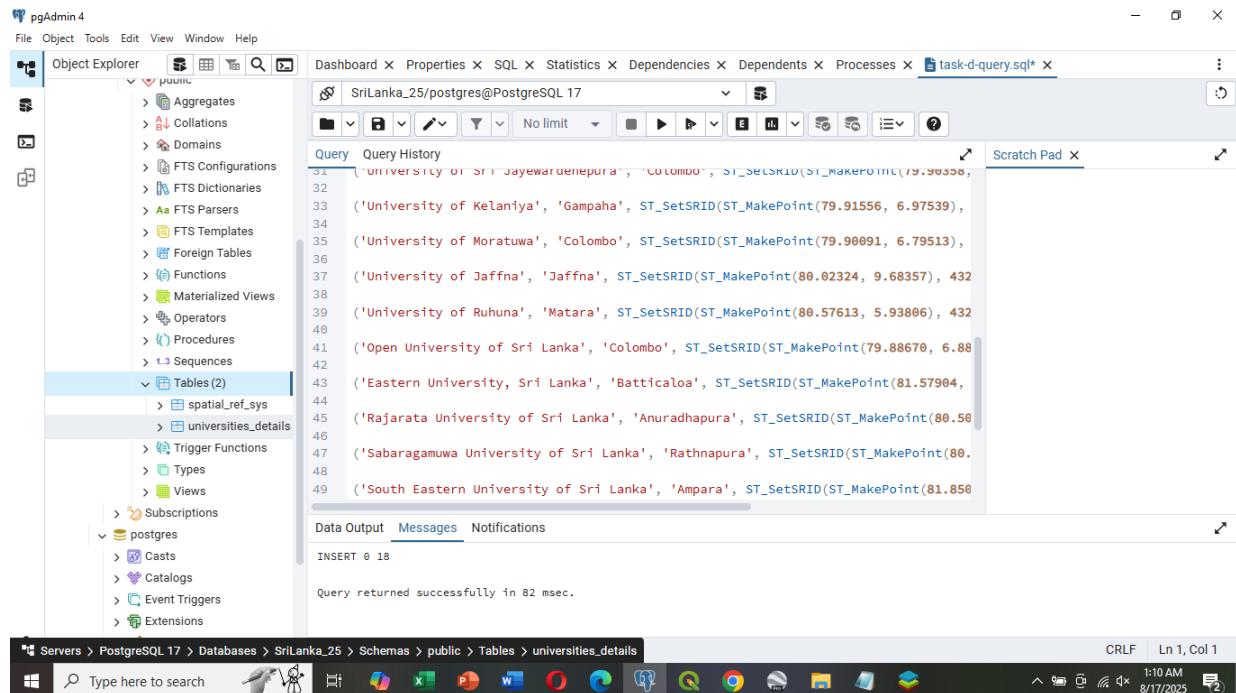
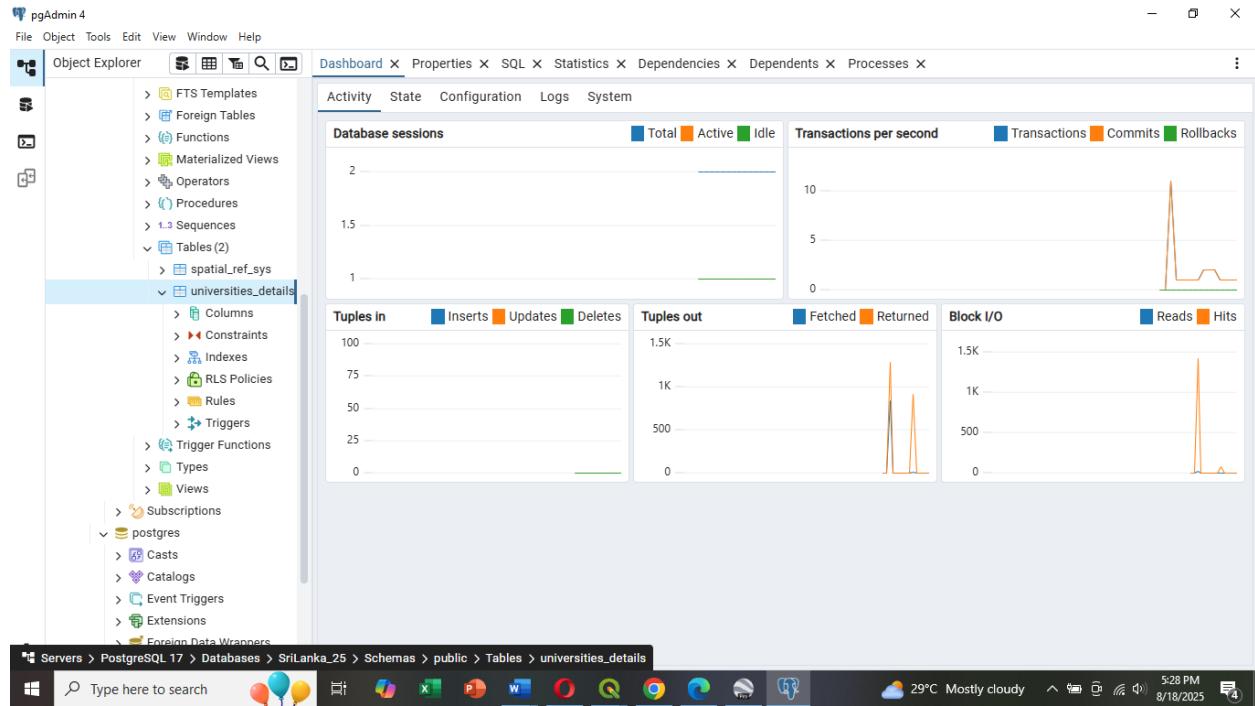
Digitization of public sector information is a vital step toward more openness, efficiency, and service delivery. This study describes the process of digitizing an instructive area map for Sri Lanka's Ministry of Higher Education (MOHE). The major goal is to illustrate the use of geospatial analytics, specifically Geographic Information Systems (GIS) such as QGIS, to visualize and analyse MOHE's operations. A significant focus is on critically assessing how MOHE's geolocation helps to the advancement of the national education system.

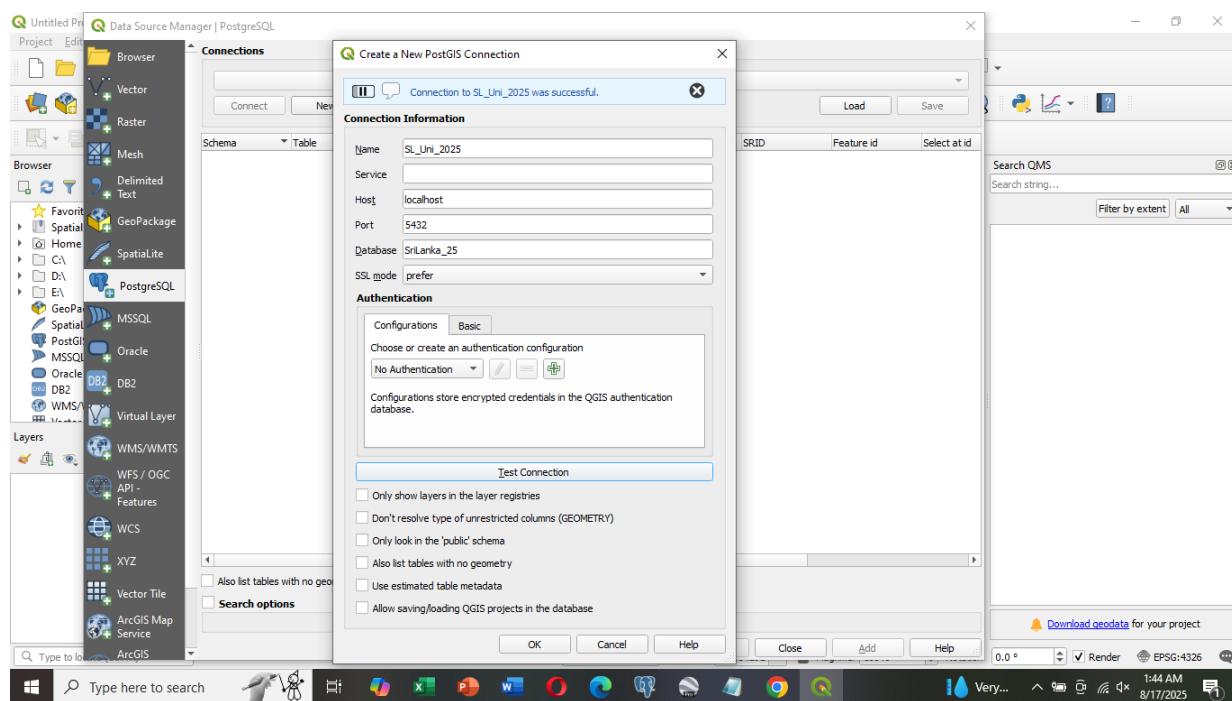
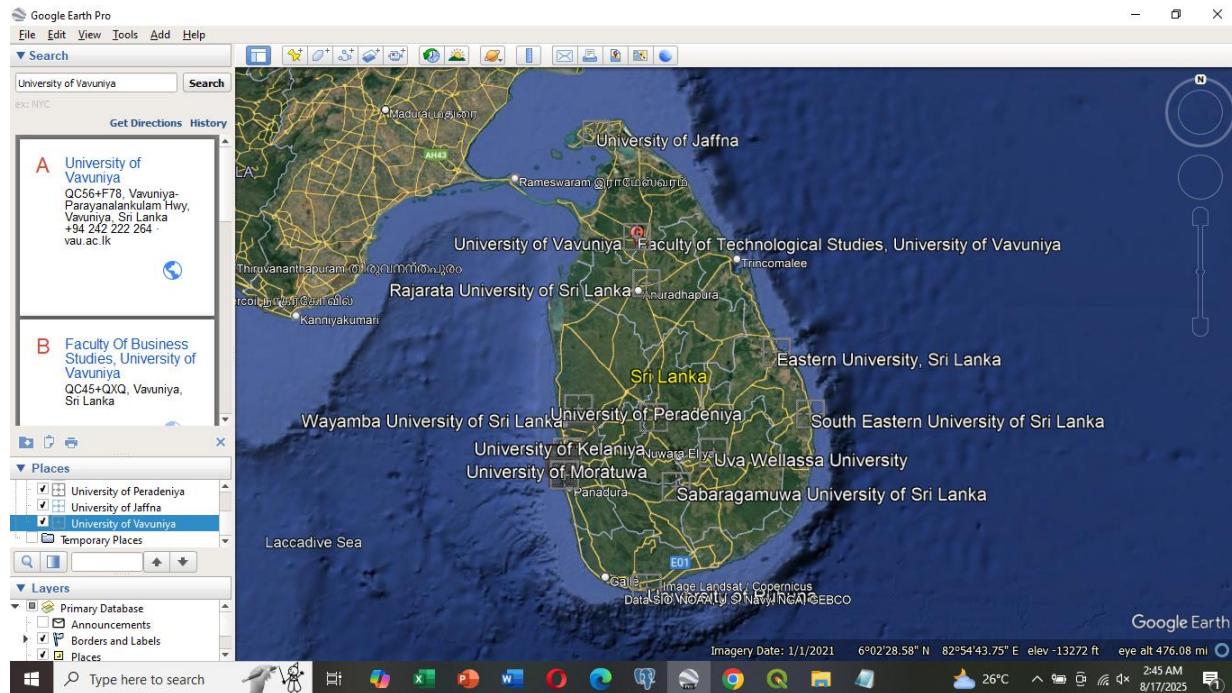












Critical discussion and recommendations

Higher education has an important strategic role in Sri Lanka.

Higher education institutions in Sri Lanka are more than just academic centers; they are key drivers of national development, strategically positioned to solve the country's economic, social, and regional imbalances. The Ministry of Higher Education (MOHE) and subsequent administrations have historically leveraged the development and expansion of universities to promote a more equitable distribution of knowledge and opportunity. The geographic distribution of these 18 state institutions is thus not a coincidental event, but rather a purposeful policy decision aimed at promoting regional growth and assuring accessibility. This debate seeks to critically examine these institutions' contributions to the national higher education system, analyzing their influence on economic growth, regional equity, and service improvements, with an emphasis on whether the country's goals have been met.

Analysis of geographical distribution and accessibility.

Based on a geographical study of university sites, a distinct pattern appears. The Western and Central provinces have a high concentration of universities, notably in and around important cities like as Colombo and Kandy. The Colombo district is home to several educational institutions, including the University of Colombo, the University of Sri Jayewardenepura, and the University of Moratuwa. This pattern is not just a result of historical evolution, but it also represents the current concentration of people and economic activity. However, a deeper examination of the map indicates that institutions have been purposefully constructed in other places to solve historical and geographical inequalities. Institutions such as the University of Jaffna in the north, Eastern University in the east, and the University of Ruhuna in the south are part of a deliberate effort to spread higher education throughout the island, with the goal of fostering national integration and peacebuilding in post-conflict regions.

While this regional grouping was initially effective, it has had serious consequences. The concentration of top-tier colleges in the country's most developed areas may increase regional

disparities. Students from rural and underdeveloped areas, such as Uva and the North Central Province, face significant challenges to accessing these universities, including greater travel and housing expenditures. This financial load might be prohibitive for many, making higher education a privilege rather than a right for all. Furthermore, the concentration of educational resources in cities adds to "internal brain drain," in which brilliant students and graduates leave their home regions in pursuit of better prospects, leaving rural areas with a shortage of competent workers. While initiatives to decentralize education have been implemented, the persistent dominance of urban-based colleges indicates that the objective of a fully egalitarian and geographically balanced education system has only been partially achieved.

Contribution to National Higher Education and Economic Development.

Sri Lanka's universities have contributed significantly to the country's economic growth by providing a highly qualified workforce, promoting research, and serving as local economic catalysts. They play an important role as human capital incubators, with several specialized to address specific industry demands. For example, the University of Moratuwa is well-known for its Engineering and Architecture colleges, which provide skilled labor directly to Sri Lanka's critical IT, construction, and manufacturing industries. Similarly, the University of Ruhuna, located in an agriculturally rich region, places a major emphasis on agriculture and medicine, with research having a direct influence on regional economic activity and public health services. This expertise is critical to the country's economic diversity and resiliency.

These institutions serve as research and innovation hubs in addition to human resource development. Universities such as the University of Peradeniya and the University of Colombo account for a large amount of the country's academic research output. This study adds to national knowledge pools by tackling local concerns in areas ranging from climate change adaptation to public health. Furthermore, colleges serve as local economic hubs. The presence of a prominent university draws businesses, creates jobs in auxiliary services (such as housing, food, and transportation), and stimulates local growth. This beneficial spillover effect is especially noticeable

in smaller towns where the university is the largest and most powerful institution, contributing considerably to the local GDP.

TASK – E

Critical discussion on the feasibility of the proposed R&D center.

The GIS study indicates a very acceptable location for the planned Research and Development Center for Space Sciences by using geospatial techniques to satisfy particular project criteria. This process not only identified a desirable location, but it also offered measurable facts to back up the conclusion, assuring an efficient and educated choice.

Methodology and Criteria Fulfillment

A multi-step GIS geoprocessing workflow resulted in the identification of the proposed location, which was titled Proposed_R&D_Site.shp. Initially, a buffer study was carried out to determine the appropriate proximity zones: a 1 km buffer surrounding Ananda Primary School and a 2 km buffer around Sri Gunananda Vidyalaya. These buffer zones resulted in a single, merged polygon reflecting all prospective land locations that met the distance criteria.

Following that, a Difference geoprocessing tool was used to narrow the choices. Using the LU_Landuse_Pg.shp layer and filtering for regions labeled 'AGRCA' (Agricultural farms), the analysis effectively excluded all land utilized for conventional export crops. This critical phase guarantees that the project does not displace vital agricultural operations, therefore alleviating a key project restriction. The resultant polygon depicts a region that meets both the proximity and land-use criteria.

Geospatial and quantitative data are used to determine feasibility.

The identified location is very viable for the following reasons, as substantiated by the data gathered from the analysis.

The location strikes a balance between safety and proximity, allowing for future partnerships and community participation. This balance minimizes interruption to the educational environment while offering a convenient location for a national research institution.

Sustainable land use: The exclusion of agricultural land reflects a commitment to long-term development while still respecting existing economic activity. The remaining available area is not in competition with typical export crops, which simplifies land acquisition and development.

The research identifies suitable land and infrastructure for new building, notwithstanding the presence of existing structures on the site. The Buildings_in_Proposed_R&D_Site_statistic.html file shows 391 structures having a land area of 36,956.66 m². This exact information is crucial for project planning and prospective relocation initiatives. To get the total available land for development, subtract the building footprint area from the entire size of the Proposed_R&D_Site. This estimate gives project managers an accurate quantity of useable land, demonstrating the site's suitability for the planned R&D facility.

\tImplicit Soil Suitability: By excluding agricultural land, the study identifies regions with soil characteristics that are more favorable for development than agriculture. While a more extensive geotechnical study is required to establish load-bearing capability and other engineering criteria, the GIS model has already completed the critical first pass screening.

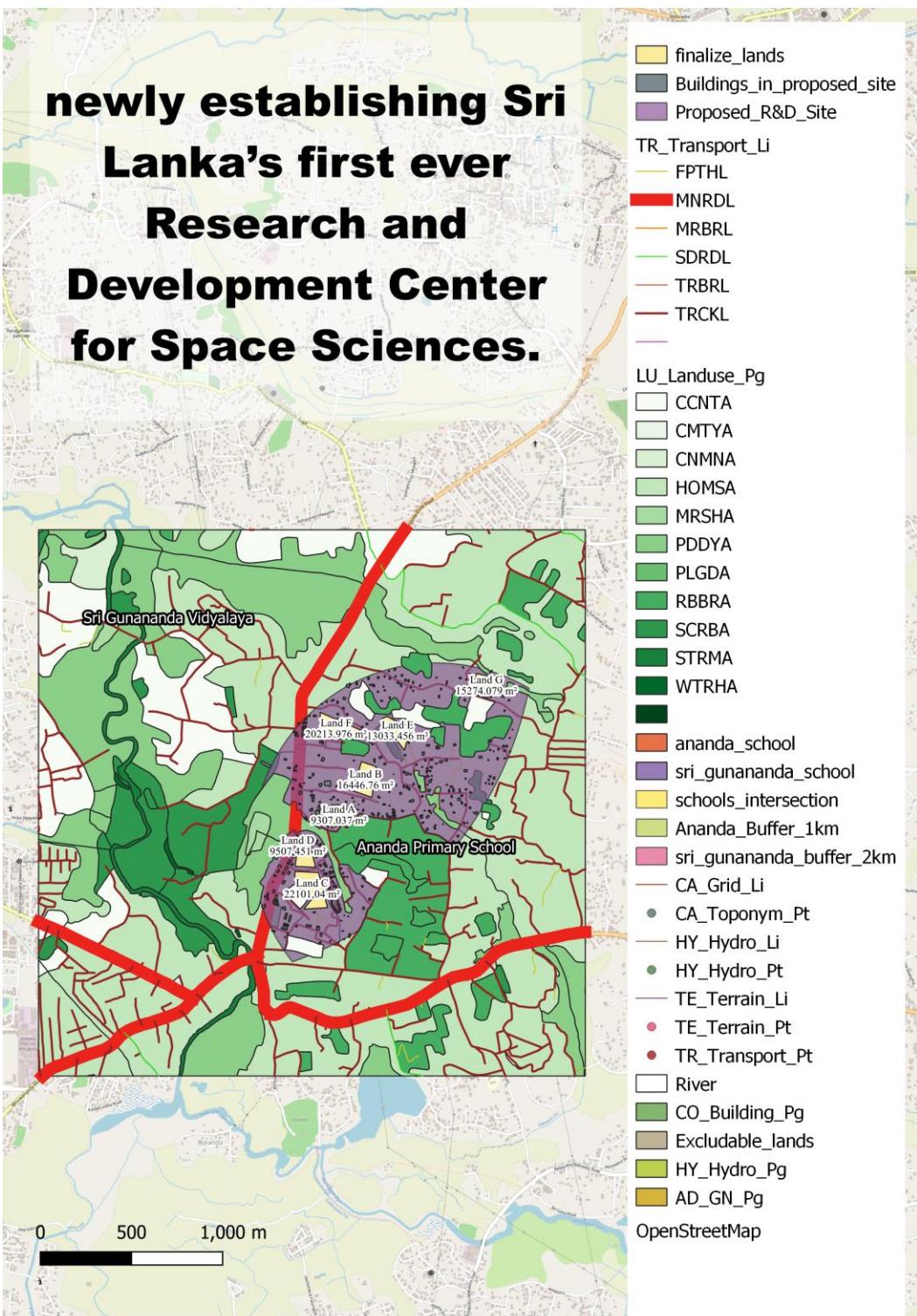
Human Capital Development and Community Engagement: Being close to educational institutions is not just a safety precaution, but also a strategic advantage. The facility is conveniently located near Ananda Primary School and Sri Gunananda Vidyalaya, making it ideal for educational

outreach initiatives, student visits, and prospective partnerships. This closeness has the potential to excite the next generation of scientists and engineers, so creating a talent pipeline for the center and the nation's space sector.

Economic and environmental viability: The use of the Difference geoprocessing tool to eliminate area utilized for conventional export crops (AGRCA) demonstrates the project's commitment to sustainable development. This judgment assures that the R&D center's installation will neither harm local agriculture businesses or cause severe socioeconomic disturbance. It also means a reduction in the requirement for complex land acquisition agreements and associated compensation concerns, which may save time and costs.

Data-Driven Decision Making: GIS analysis provides quantifiable data, such as the number of existing buildings (391) and total occupied area ($36,956.66 \text{ m}^2$), transforming the project from a theoretical notion to a feasible design. This data enables accurate cost-benefit analysis, thorough project budgeting for destruction or repurposing existing structures, and exact computation of the remaining accessible acreage for development. This level of information, made possible by the GIS analysis, minimizes uncertainty while strengthening the entire project proposal. The final computation of total appropriate and available land area yields a clear, actionable indicator for project implementation.

The geospatial analysis makes a persuasive case for the selected site's practicality. The chosen area is strategically located, ecologically conscientious, and quantitatively feasible for a large research institution. The investigation included GIS methods such as Buffering and Difference to discover a site that met all requirements, including proximity to schools and the absence of agricultural area. The data, including 391 existing buildings with an occupied area of $36,956.66 \text{ m}^2$, gives a strong, evidence-backed reasoning. This strategy assures that the project is consistent with national development goals, environmental stewardship, and solid project management principles, translating a complicated problem into a well-defined, practical solution.



Google Satellite

Conclusion

In conclusion, this geospatial analysis successfully transformed raw statistical data from the *Annual School Census of Sri Lanka 2022* into a powerful visual and analytical tool. By systematically mapping the distribution of national and provincial schools, we have identified a clear and non-uniform pattern of educational infrastructure across the country. The thematic map, employing choropleth symbology and detailed labels, vividly illustrates a significant concentration of schools in the highly urbanized Western and Central Provinces, suggesting a strong correlation between educational development and population density.

The foundational work of developing the `SriLanka_2025` geospatial database was instrumental to this achievement. By performing a meticulous data join, the non-spatial census figures were linked to their corresponding district polygons, transforming static numbers into a dynamic and spatially aware dataset. This technical process was the linchpin that enabled the visual exploration of disparities and the identification of regional concentrations, effectively turning a simple statistical table into an actionable decision-making tool.

The focused case study of the Kurunegala district, with its high total number of schools (871), provided a deeper, more nuanced understanding of these national trends. The analysis of this district's infrastructure revealed a significant imbalance between Provincial (840) and National (31) schools. This finding highlights a crucial planning consideration: a high volume of schools does not automatically equate to equitable access, particularly when considering the potential for disparities in school type and quality. This detailed investigation proved the value of a targeted approach in understanding local educational landscapes.

However, it is crucial to recognize the inherent limitations of this analysis. While the map is an excellent tool for identifying quantitative geographical trends, it does not tell the entire story. The visualization does not account for qualitative factors such as the student-to-teacher ratio, the accessibility of schools in remote rural areas, or the varying quality of educational resources. Therefore, this map must serve as a high-level strategic overview and an initial diagnostic tool, not as the sole basis for educational policy formulation. A holistic approach would require integrating this geospatial data with additional demographic and socio-economic information.

Ultimately, this project demonstrates the critical role of geospatial analysis in supporting evidence-based decision-making for the education sector. The generated maps and analyses provide a foundational resource for educational planners and policymakers. This information can be leveraged to strategically identify underserved areas, prioritize resource allocation, and formulate effective policies aimed at ensuring equitable access to quality education for all citizens of Sri Lanka, thereby significantly shaping the nation's educational landscape for the better.

Future Recommendation:

1. Integrate demographic and socioeconomic data.

Your current study has effectively identified the distribution of schools. A critical next step is to combine demographic data, such as population density by age group, with socioeconomic data, such as poverty rates or family income. This allows you to do an analysis of unmet demand. For example, you may look for districts with a high school-age population but a small number of schools, which would be a good sign of neglected regions.

2. Perform a Buffer Analysis for Accessibility

To offer a more realistic and meaningful advice, conduct a buffer analysis to determine school accessibility. This would include establishing a buffer (e.g., a 2-5km radius) surrounding each school. The generated map would graphically show places that are outside any school's buffer zone, suggesting areas where kids may face severe travel difficulties. This is an effective method for determining the exact sites for future new school constructions.

3. Compare National and Provincial School Distribution

Your study already touches on this, however a further look is highly recommended. You might make distinct theme maps for national and provincial schools. By displaying the two levels side by side, you can better highlight the differences in access to these various school kinds. This might be used to support policy suggestions on how to enhance the quality of provincial schools or strategically elevate them to national status in order to increase equity.

4. Develop a Web-based Interactive Map

Create a simple web-based map to make your study more accessible to a larger audience. Using technologies such as QGIS's qgis2web plugin, you may convert your maps to an interactive version that can be viewed in a browser. This will enable stakeholders, such as local community leaders and government officials, to study the data without requiring sophisticated GIS software. This type of assignment would show a very practical application of your expertise.

References

- Department of Census and Statistics (2022) *Statistical Information of Sri Lanka 2022*. Colombo: Department of Census and Statistics.
- ESRI (2019) *GIS for a Sustainable World: Geographic Information Systems and Sustainable Development*. Redlands, CA: ESRI Press.
- Goodchild, M. F. (2010) 'Geographic information science: current status and future trends'. In GIS, Human Geography, and the Social Sciences. Springer, pp. 3-10.
- ICBT Campus (2024) *ICBT CIS6008 S2 Sri Writing 1 Nov-2024 Main 2024-25*. [Assessment Brief]. Available at: [Insert your file URL or mention as "provided document"].
- Ministry of Education (2022) *Annual School Census of Sri Lanka Final Report - 2022*. [PDF document]. Colombo: Statistics Branch of Ministry of Education.
- Ministry of Education, Sri Lanka (2023) *National Education Policy Framework*. Colombo: Ministry of Education.
- QGIS (2025) *QGIS Version 3.28*. [Computer software]. Available at: <https://qgis.org/en/site/> (Accessed: 15 August 2025).
- Thematic Map Visualization (2025) *Distribution and Composition of Educational Infrastructure in Sri Lanka*. [Geospatial map visualization]. Created in QGIS for academic project.
- World Bank (2020) *Sri Lanka Education Sector Assessment: From Youth to Adult*. Washington, DC: World Bank. Available at: [Insert URL if you used an online version].
- Google Earth Pro, (version 7.3). 'Sri Lanka Satellite Imagery'. Available at: <https://earth.google.com/> (Accessed: 2023/08/22). | Google Maps, (Year). 'Maps of Sri Lanka'. Available at: <https://www.google.com/maps> (Accessed: 2023/08/22).

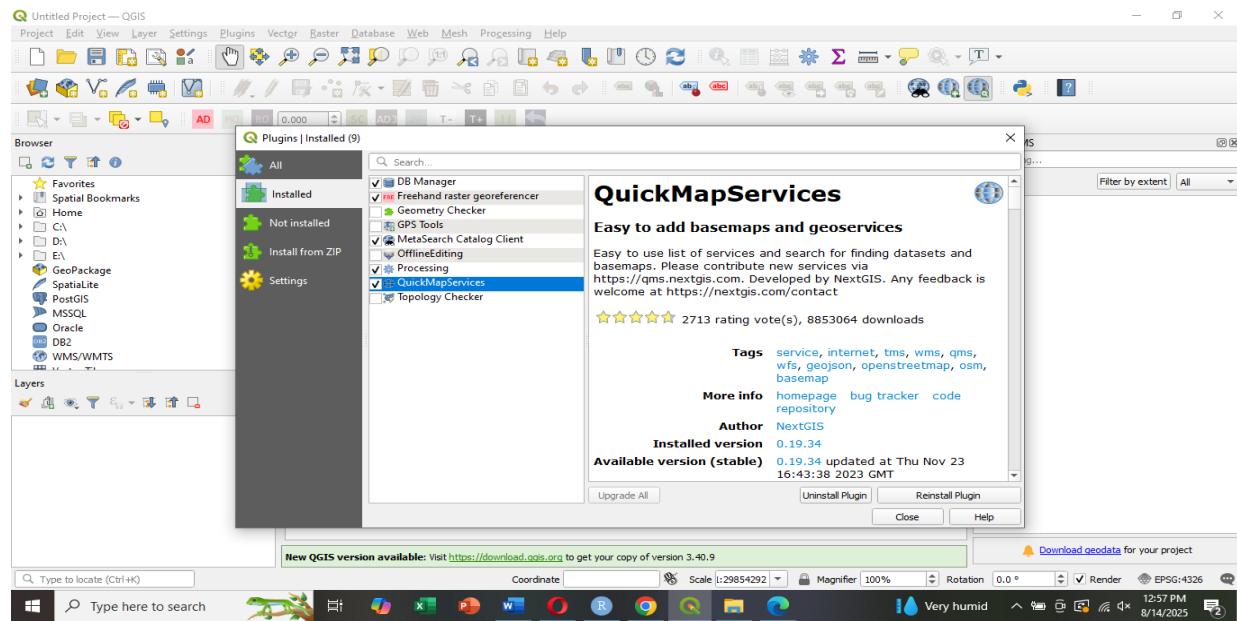
- Central Environmental Authority, Sri Lanka, (Year). 'Environmental Guidelines for Planning in Sri Lanka'. Available at: <https://www.cea.lk/web/en>(Accessed: 2023/08/22).
- QGIS Development Team, (Year). 'QGIS Geographic Information System'.Open Source Geospatial Foundation Project. Available at:<https://www.qgis.org/> (Accessed: 2023/08/22).
- RStudio Team, (Year). 'RStudio: Integrated Development for R'. RStudio, Inc., Boston, MA. Available at: <http://www.rstudio.com/> (Accessed: 2023/08/22).

Appendices

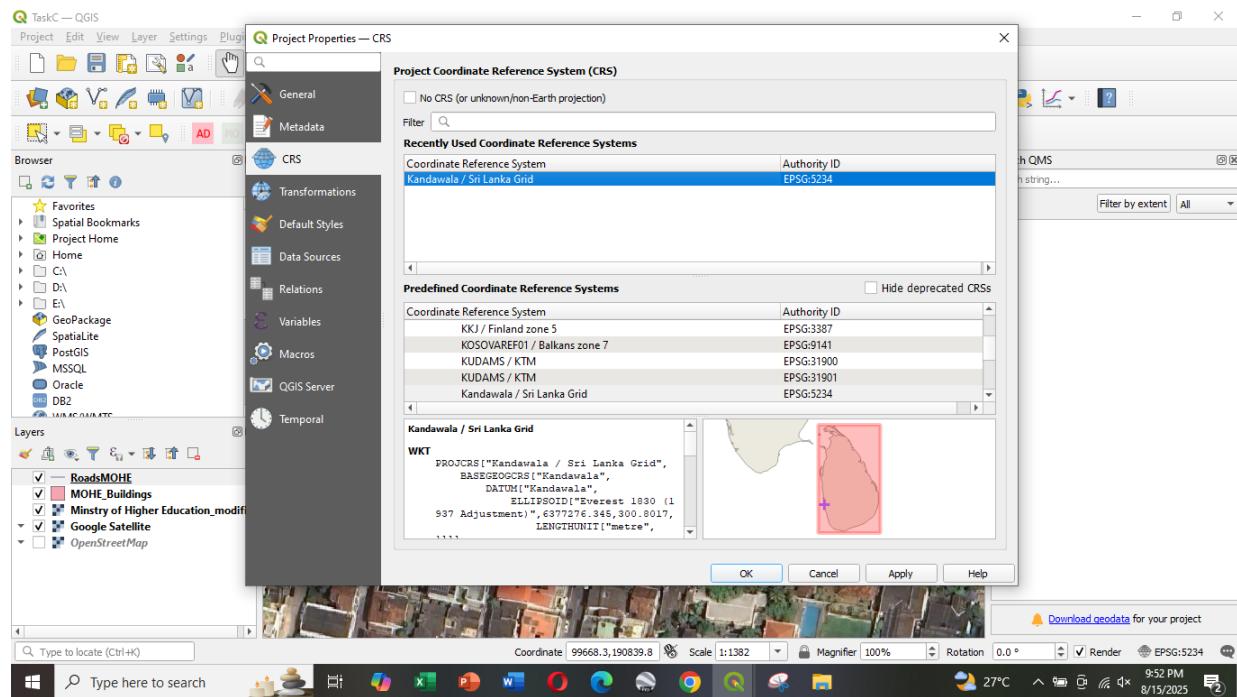
Github link: <https://github.com/ashraff96/ABI-2025>.

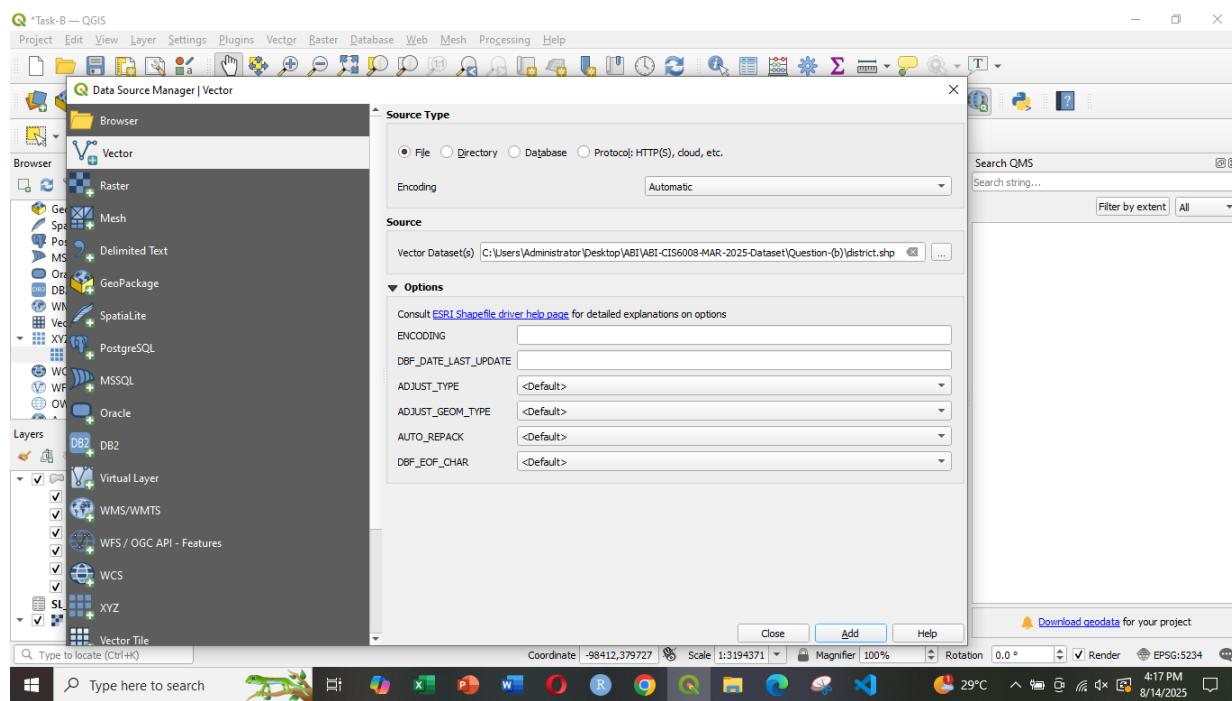
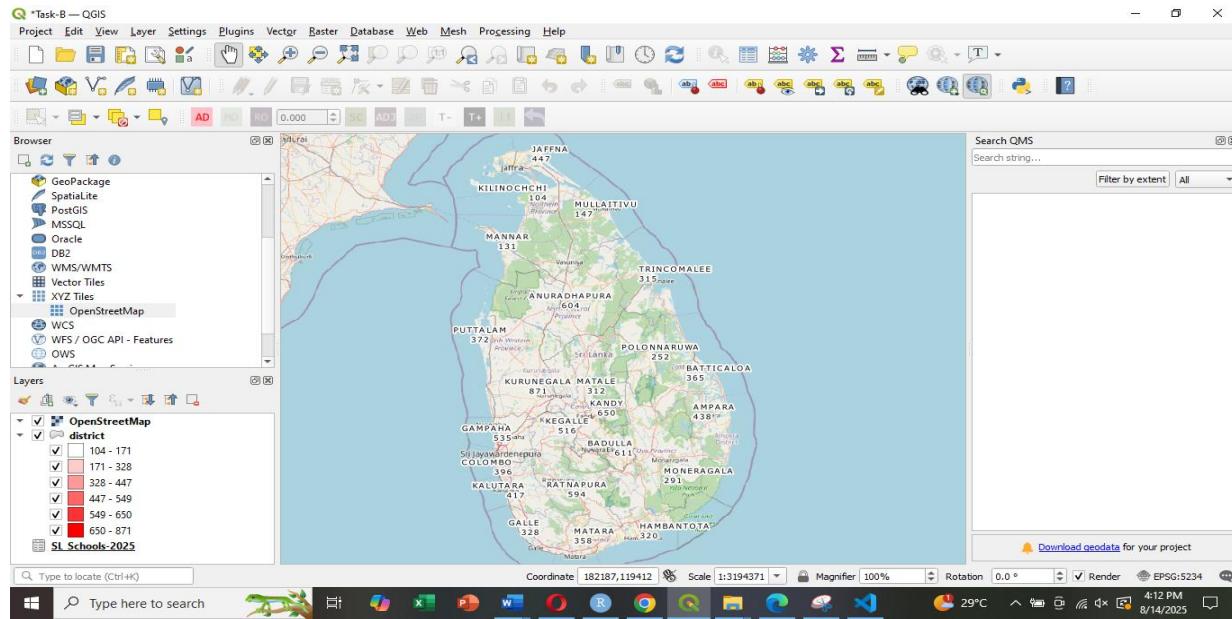
Map Development: Question B

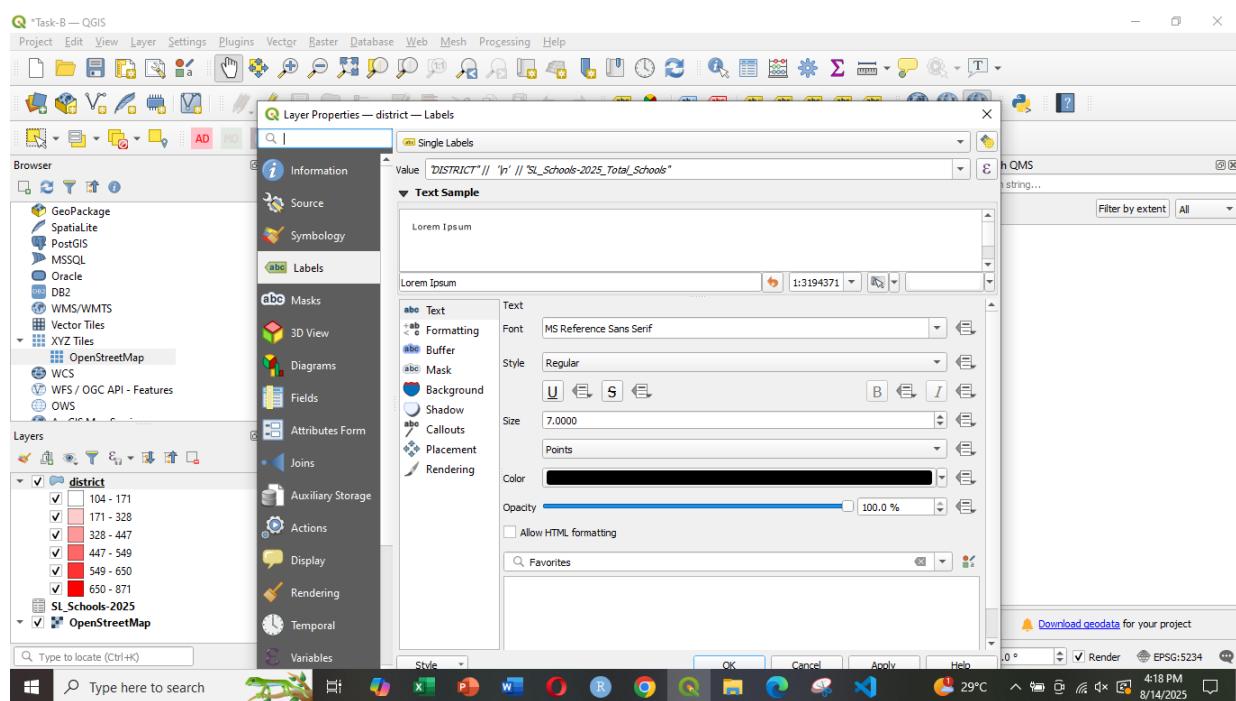
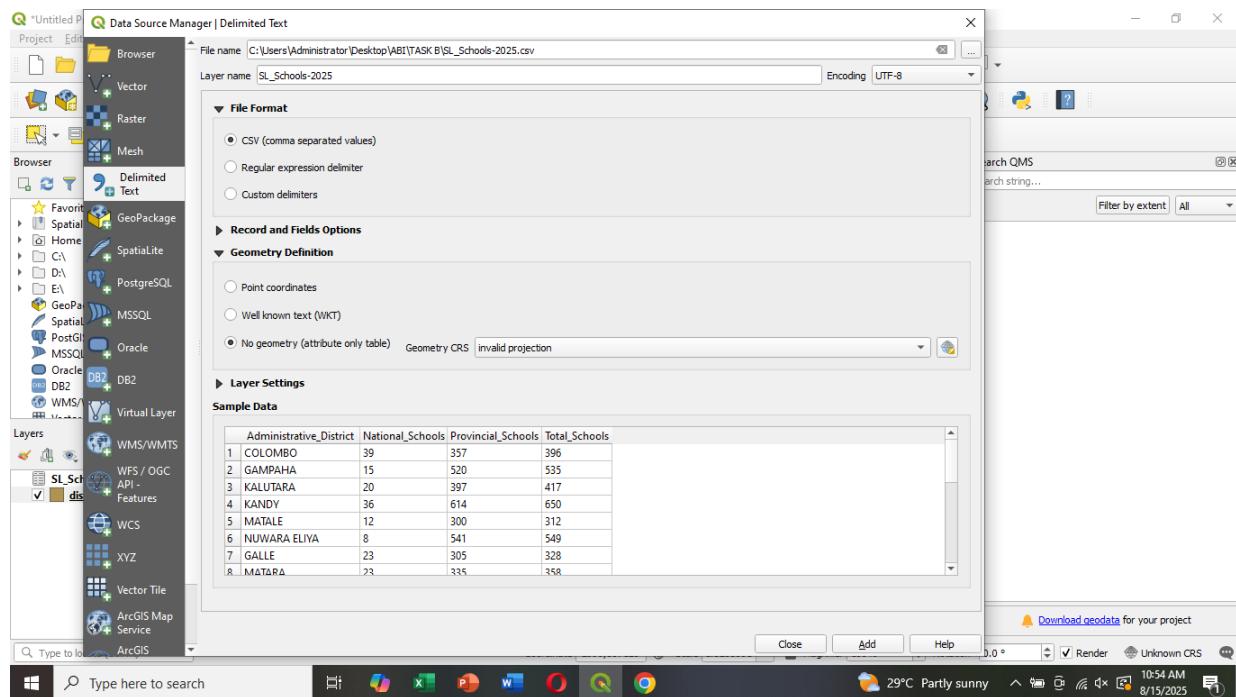
Step 1: Install Quick Map Services to illustrate maps.

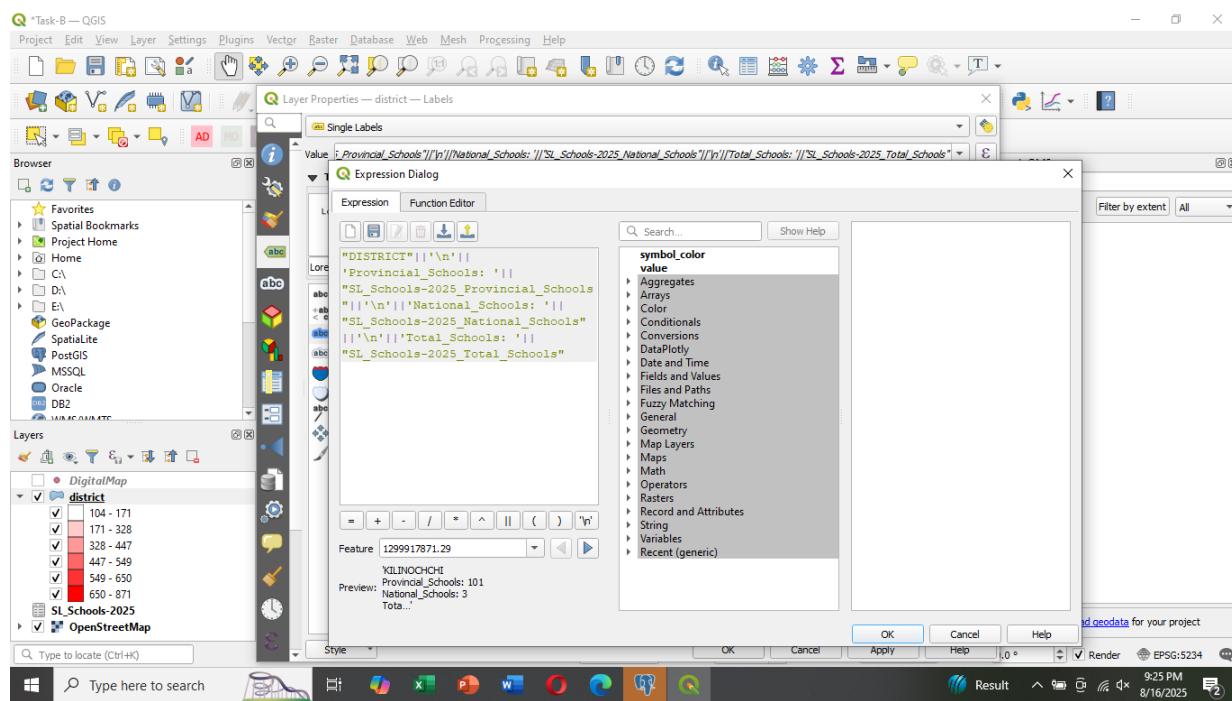
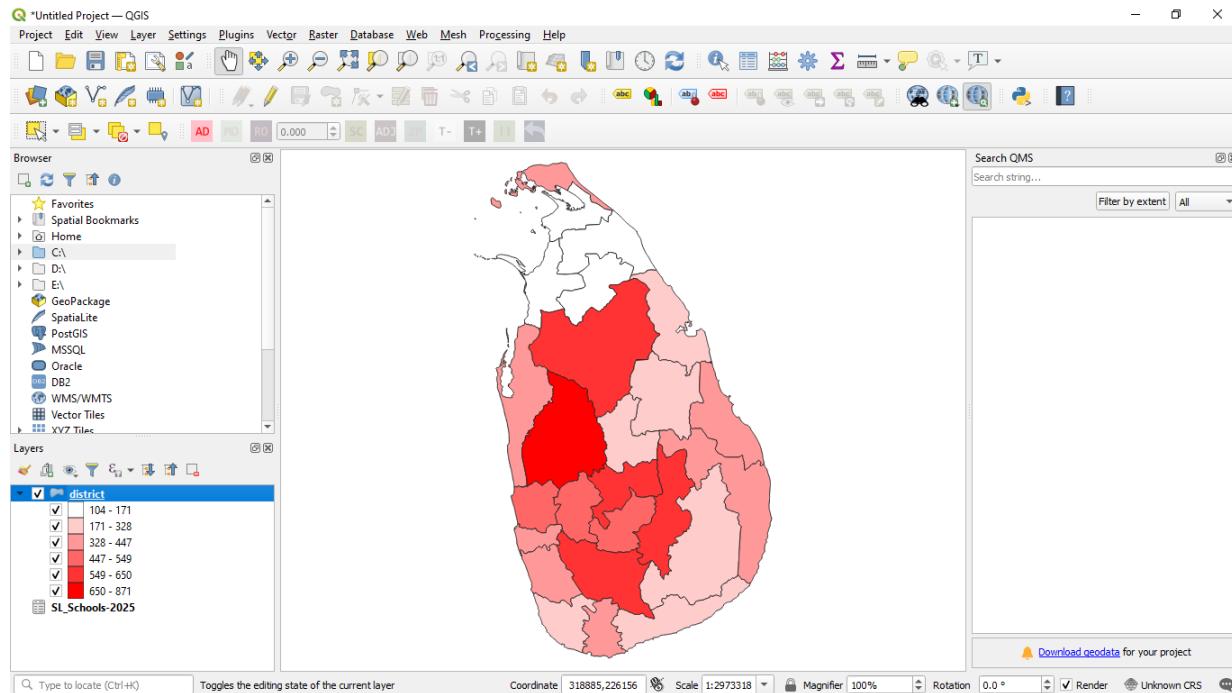


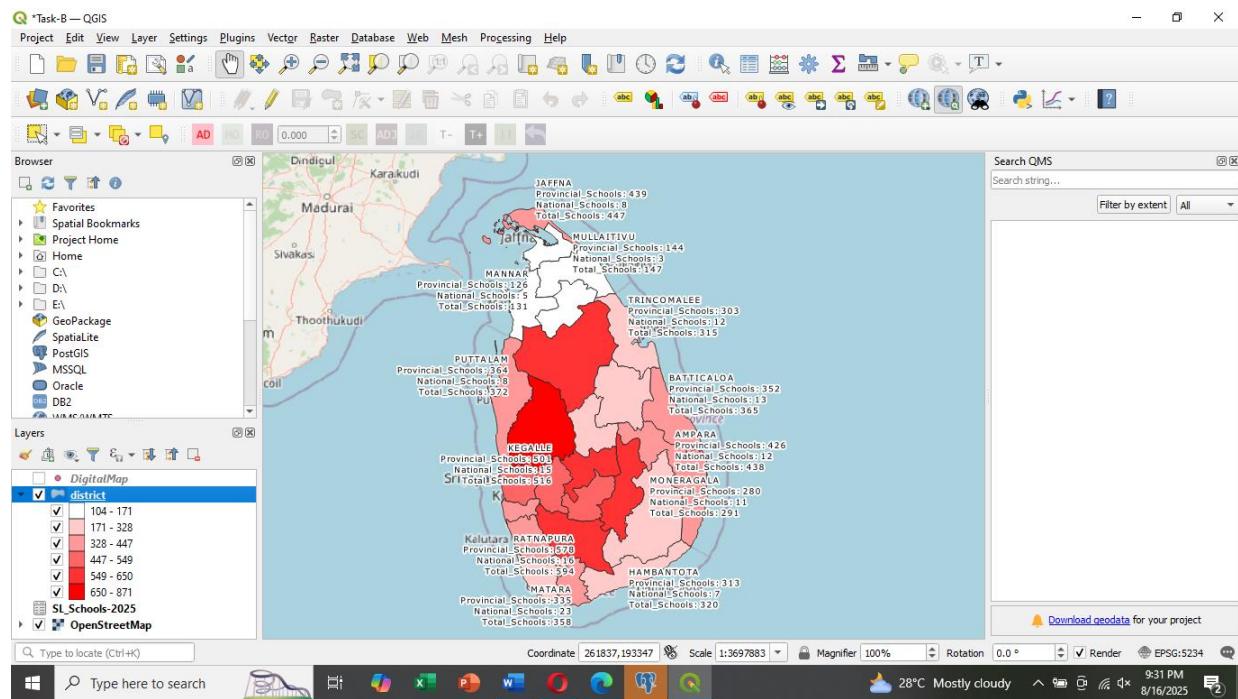
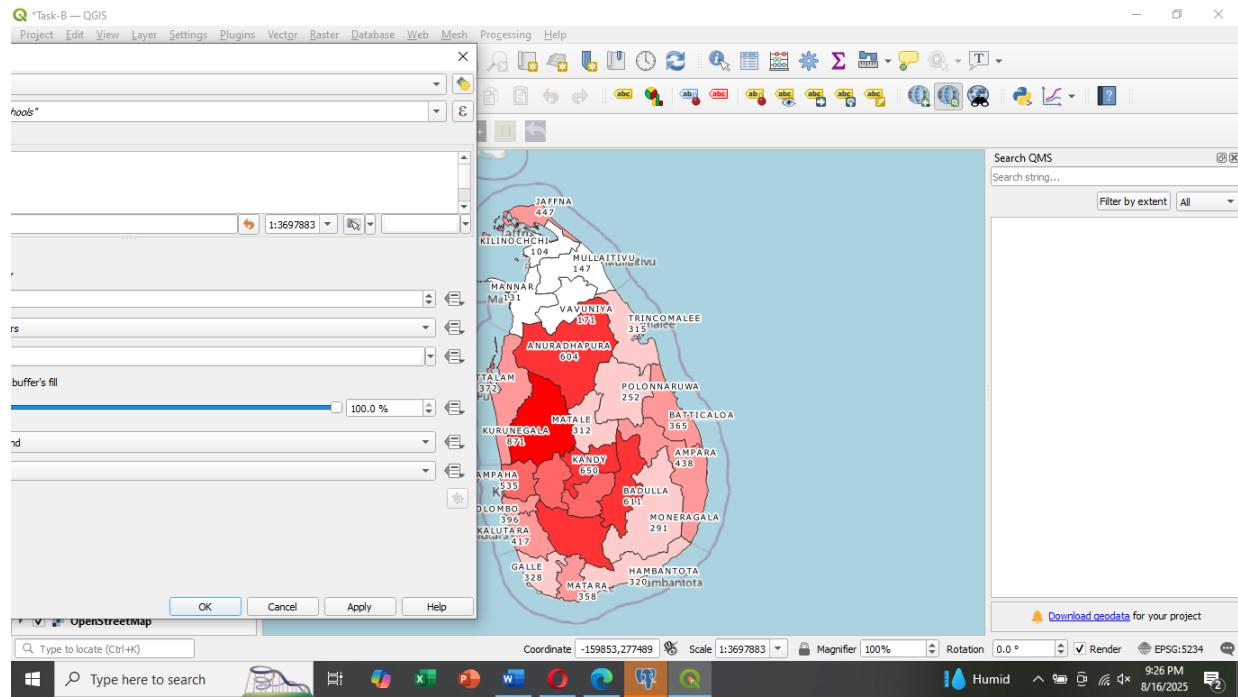
Select – KANDAWALA / SRILANKA GRID

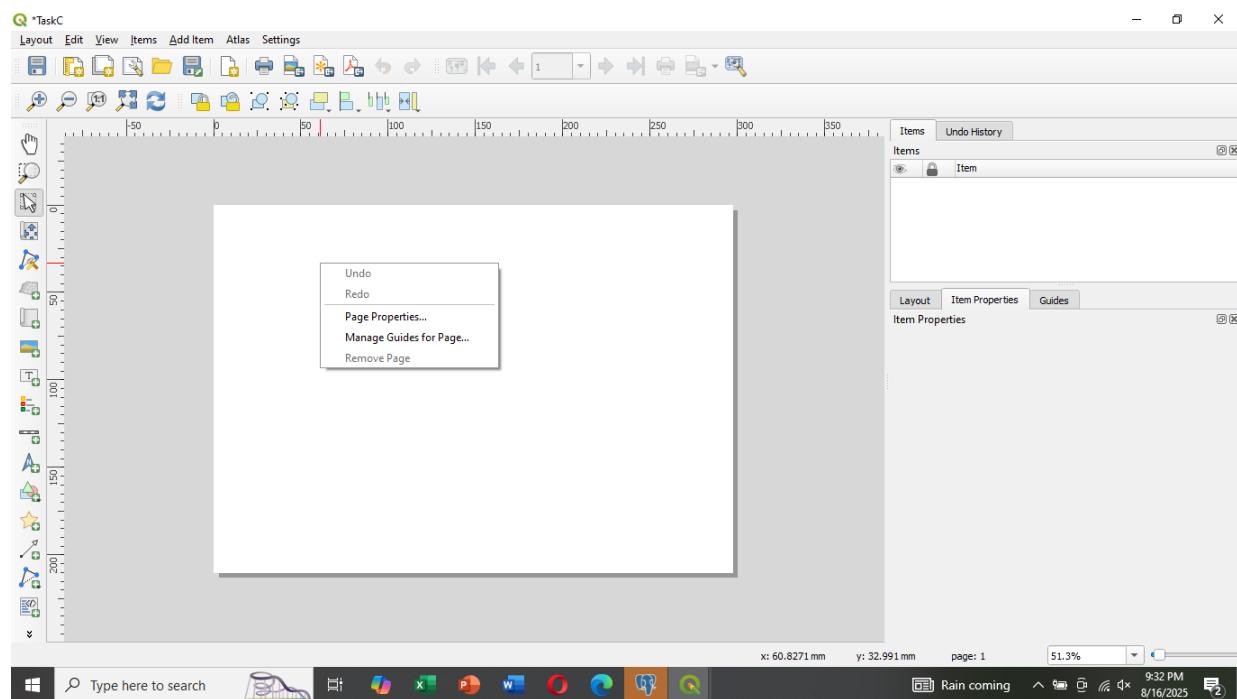
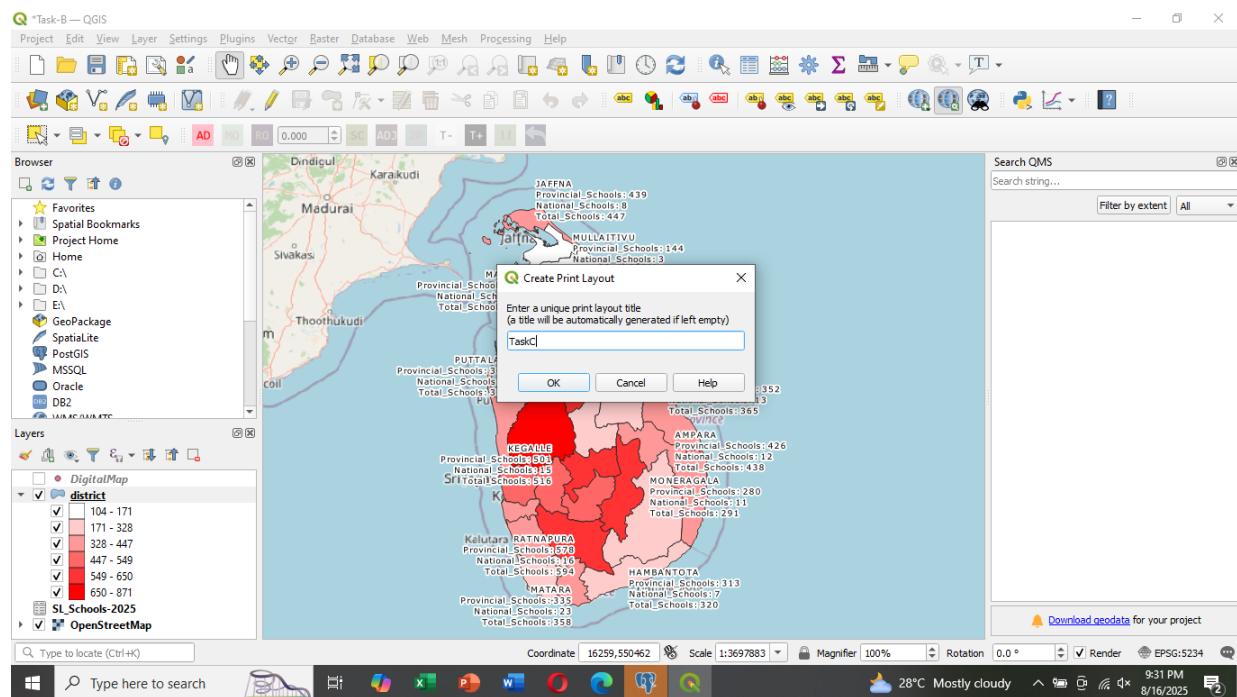


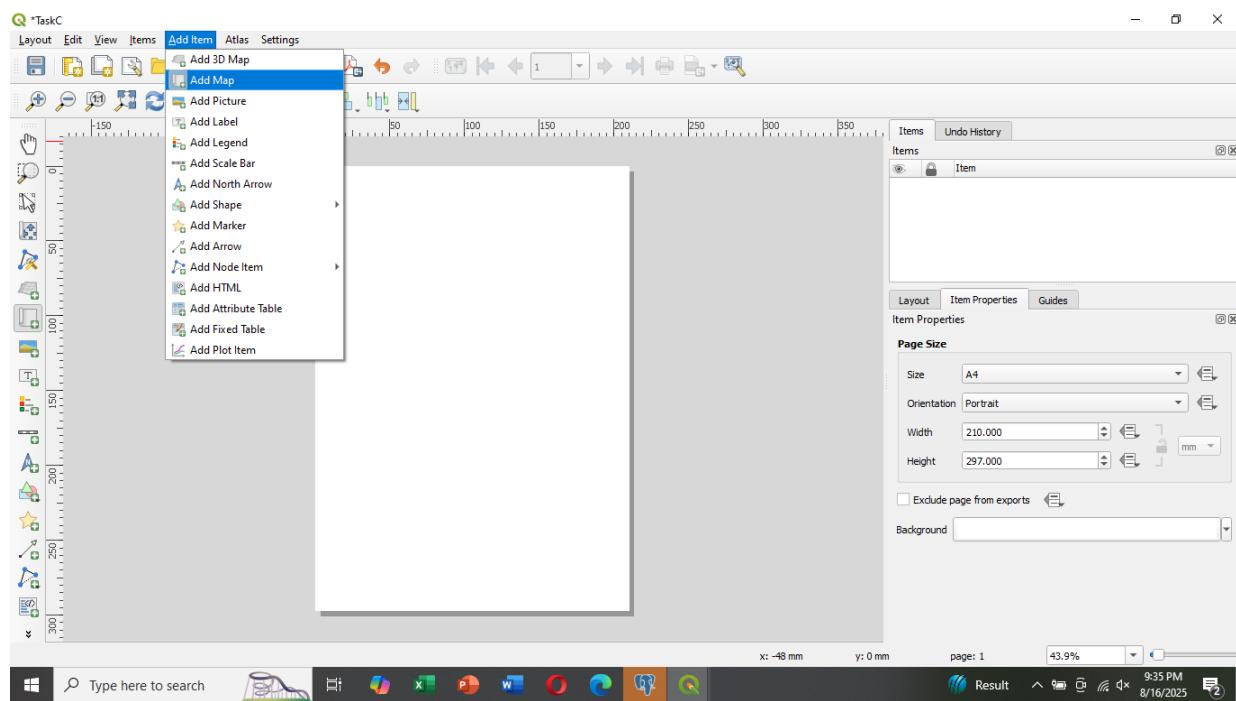
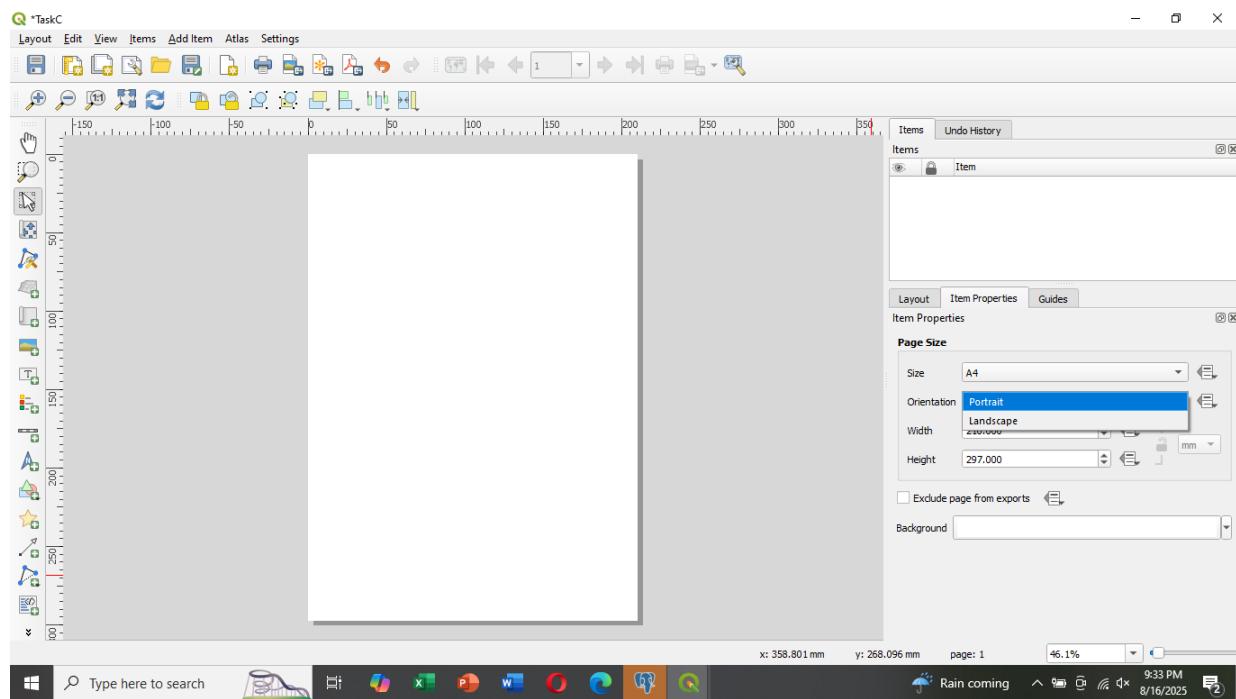


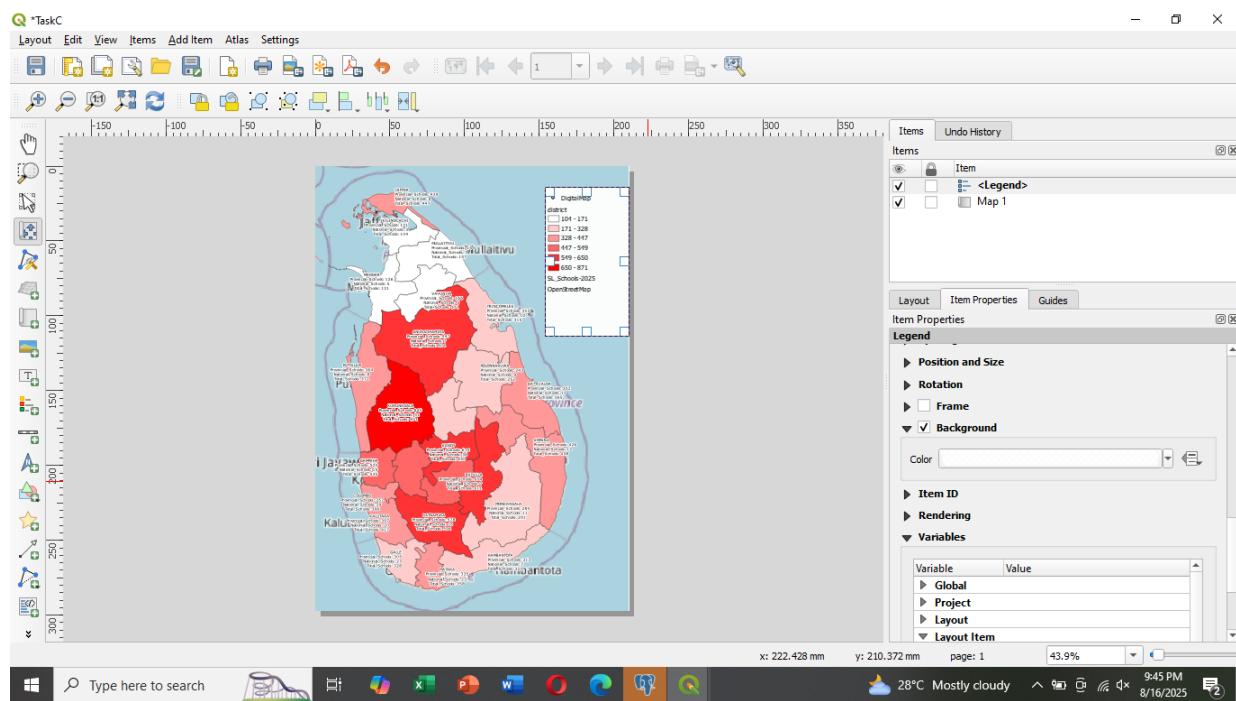
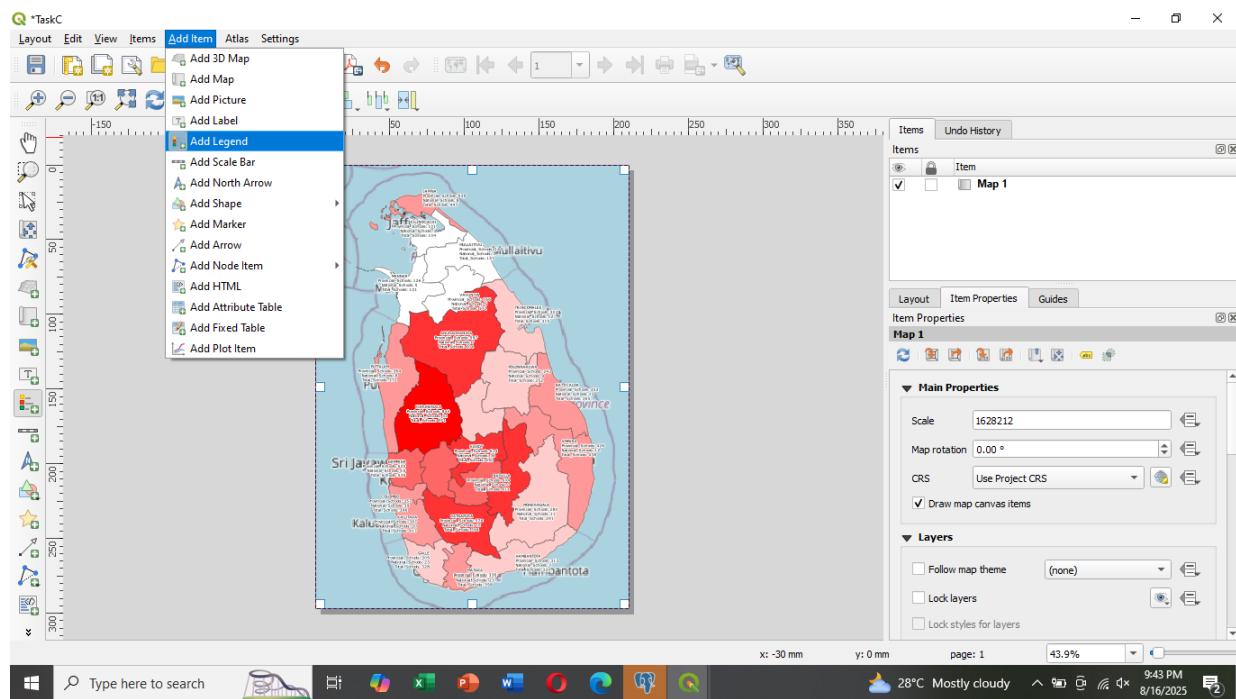


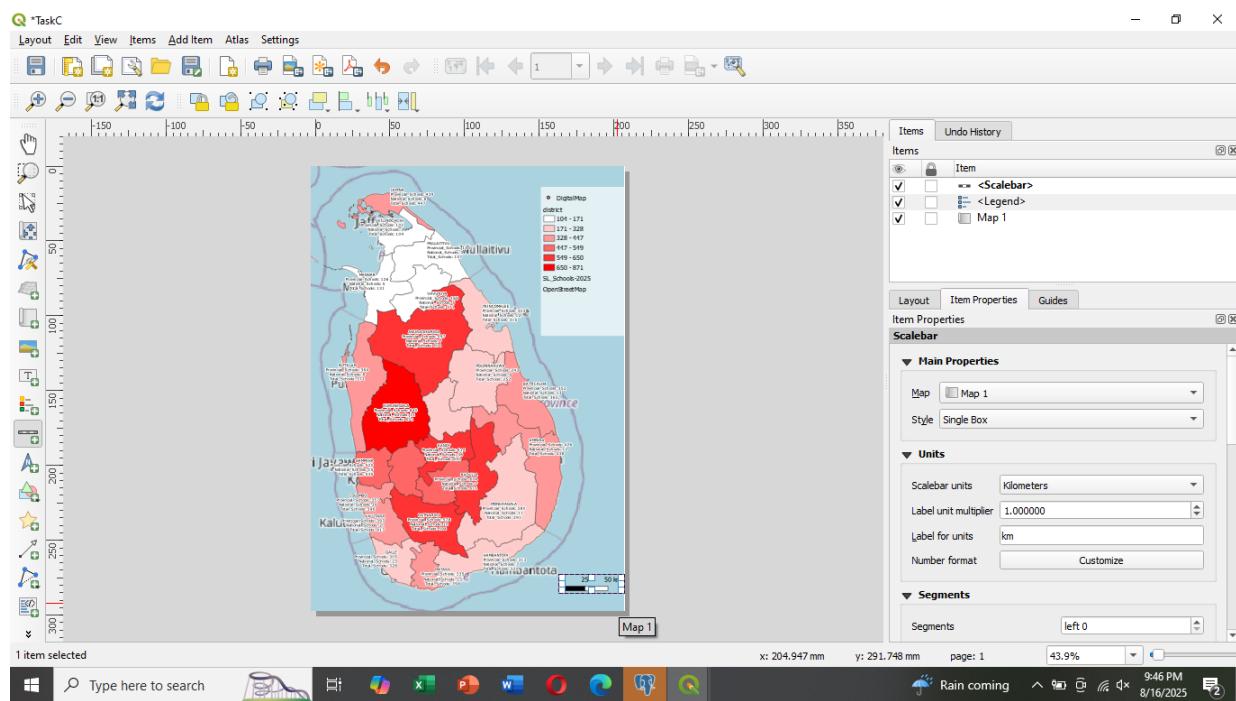
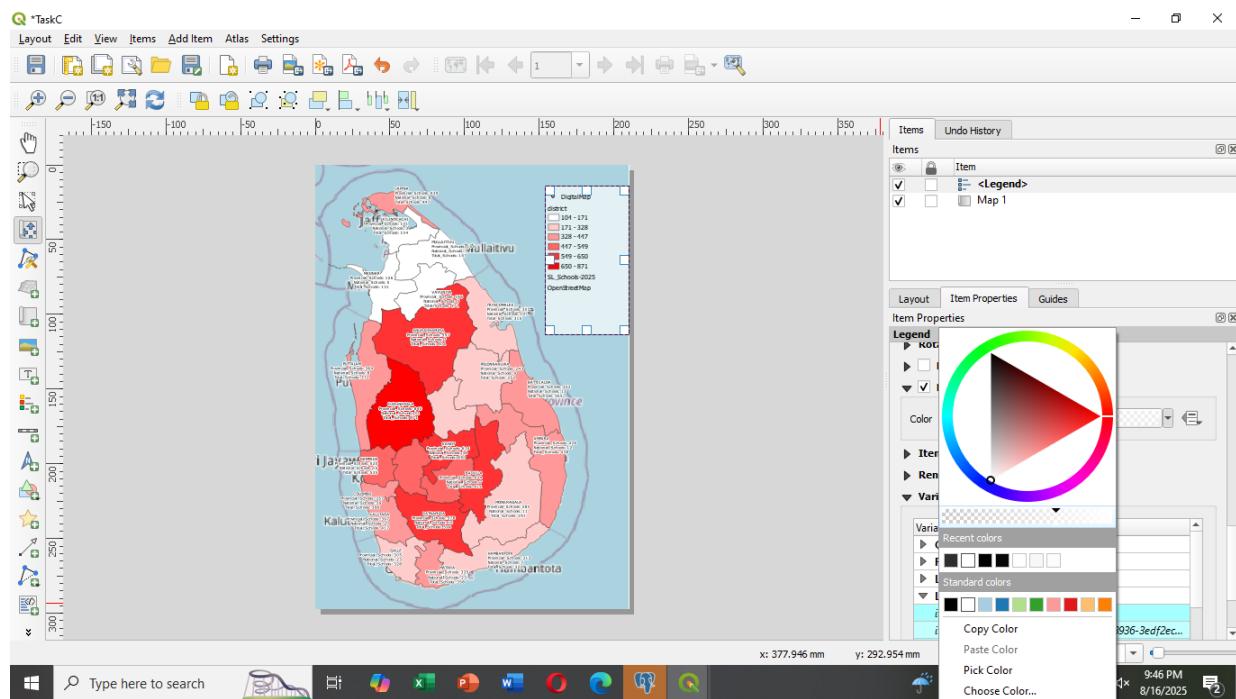


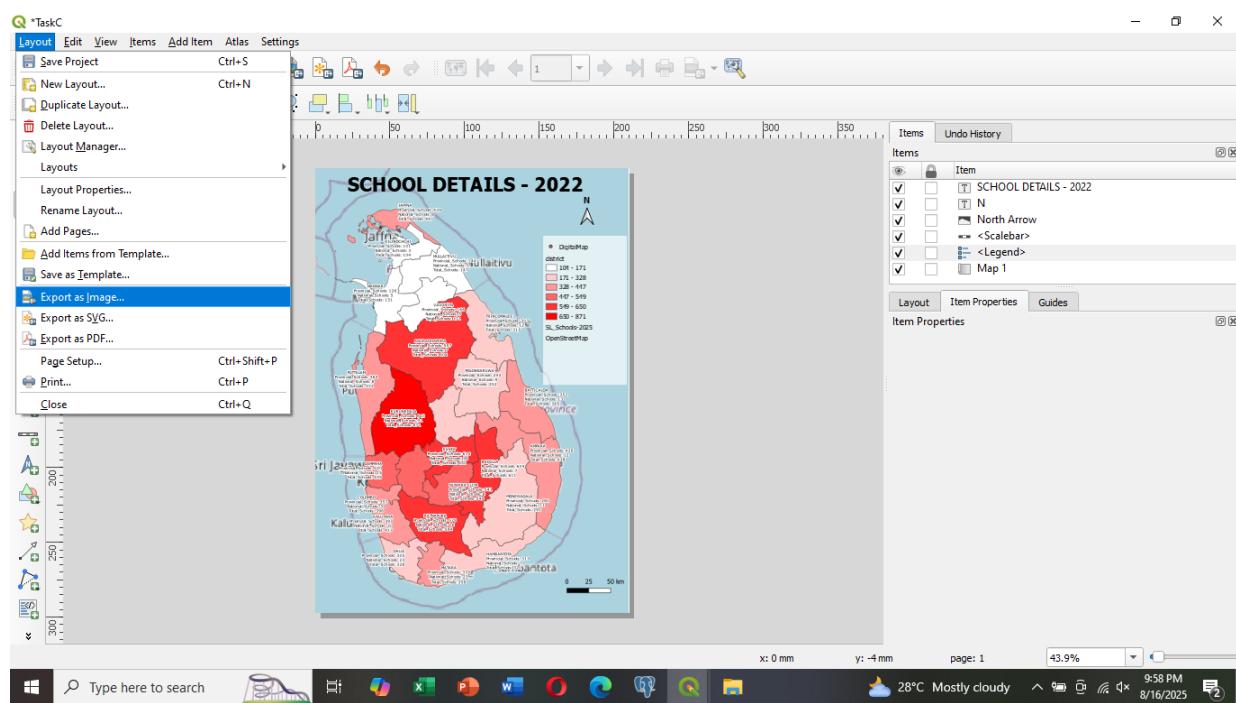
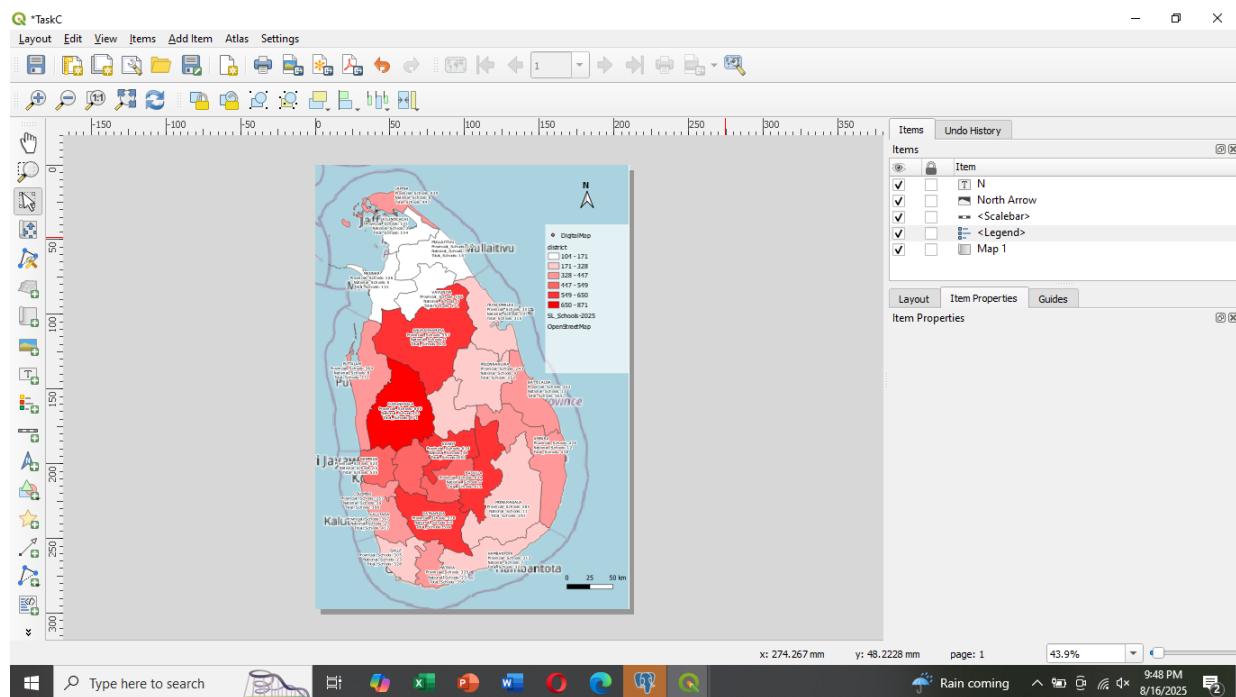


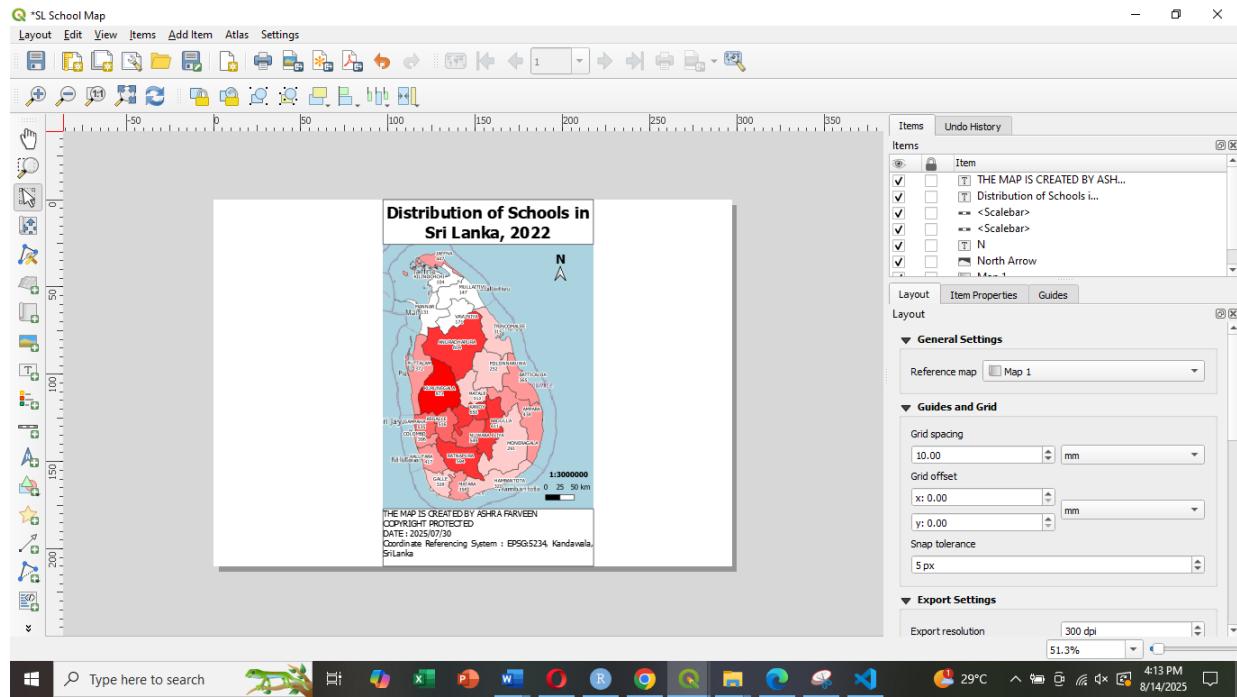




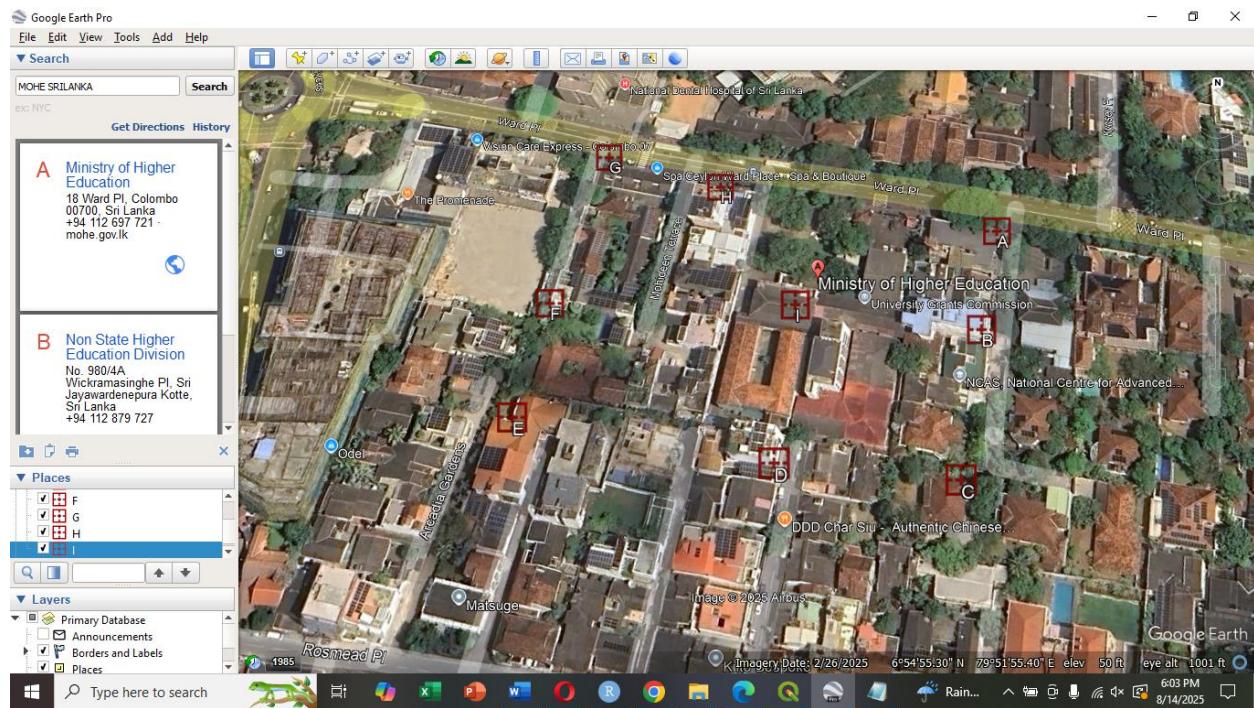


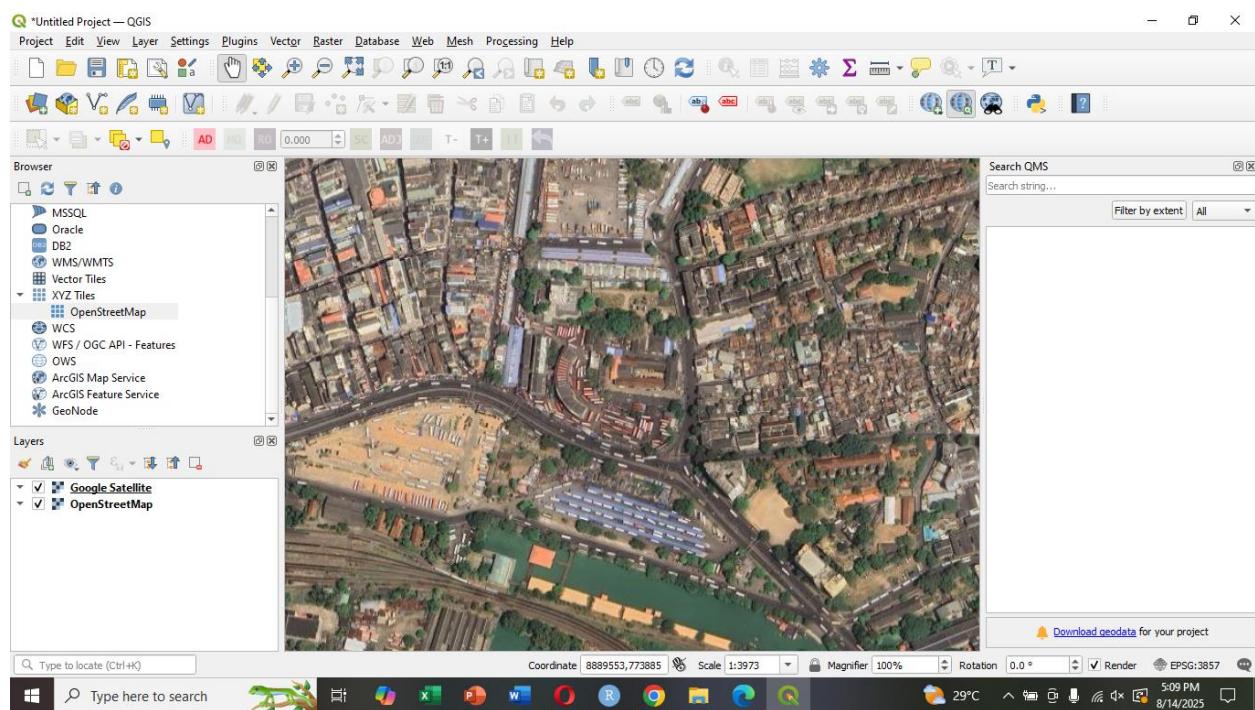
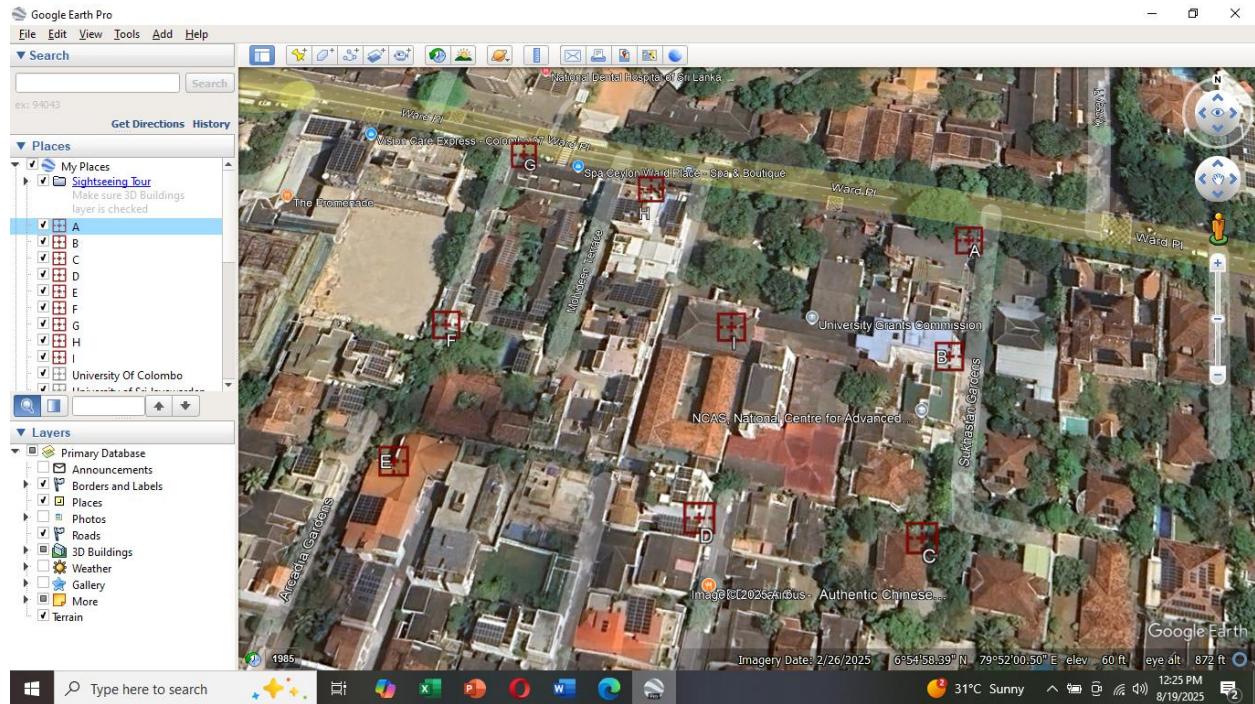


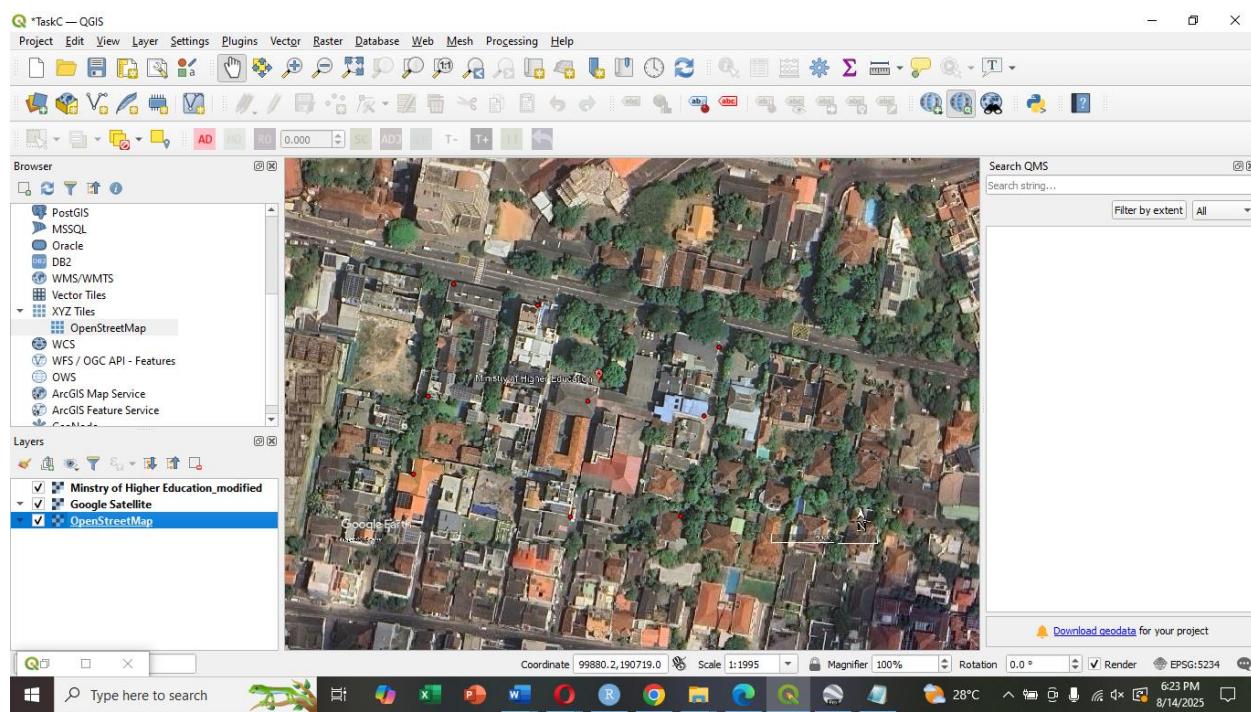
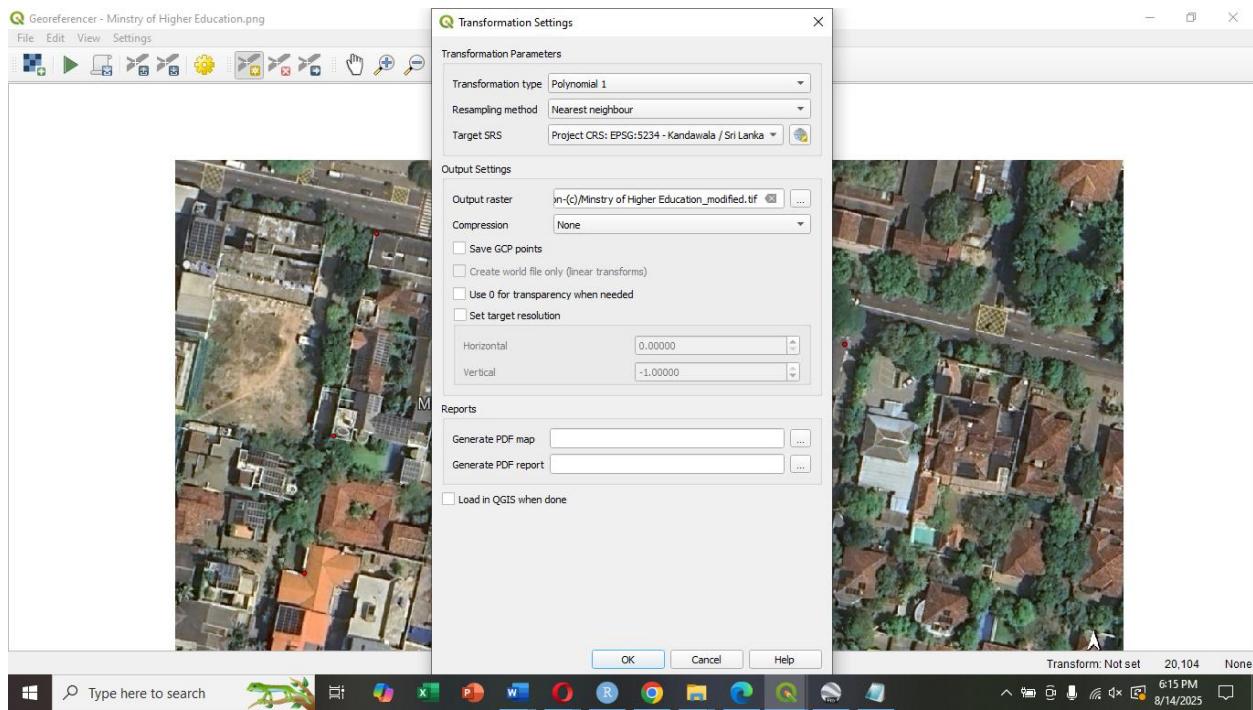


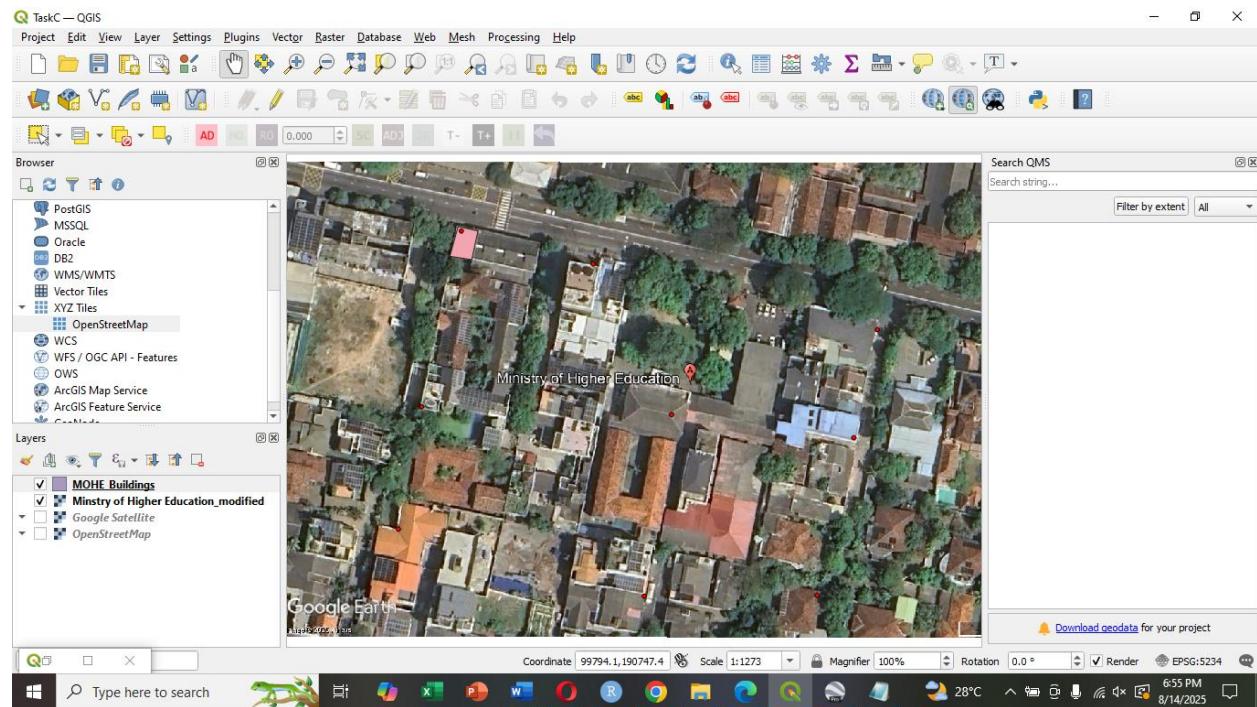
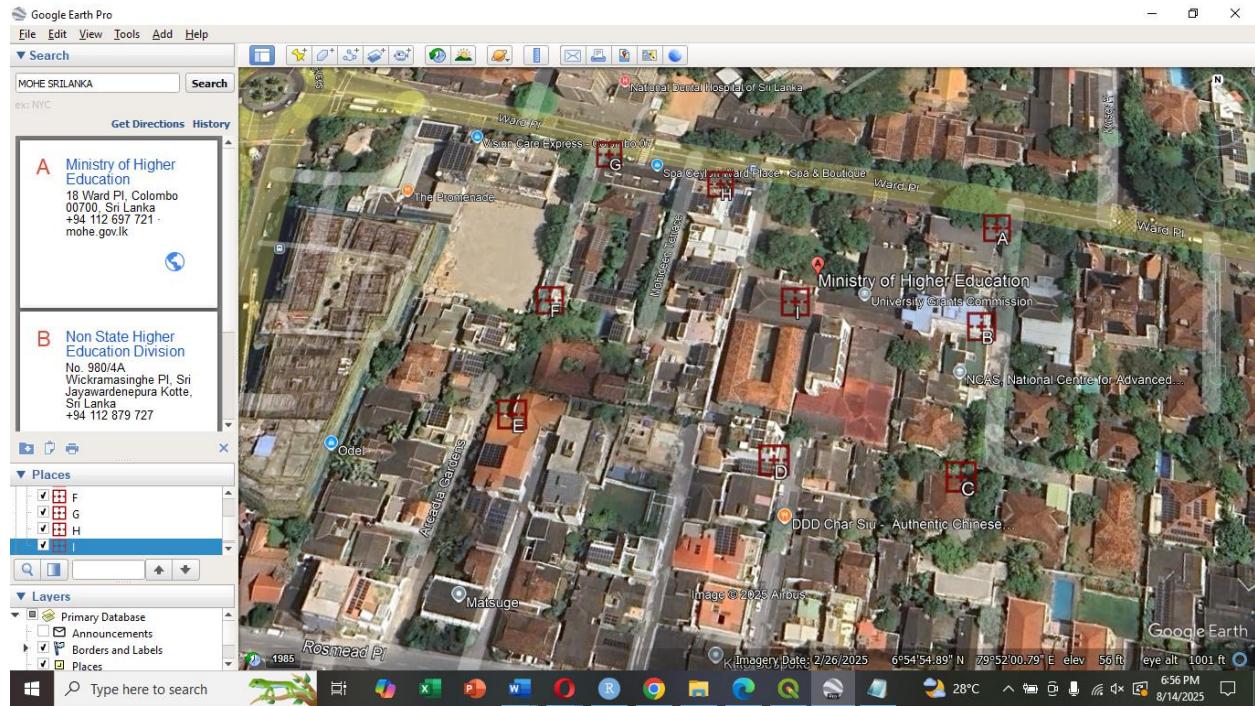


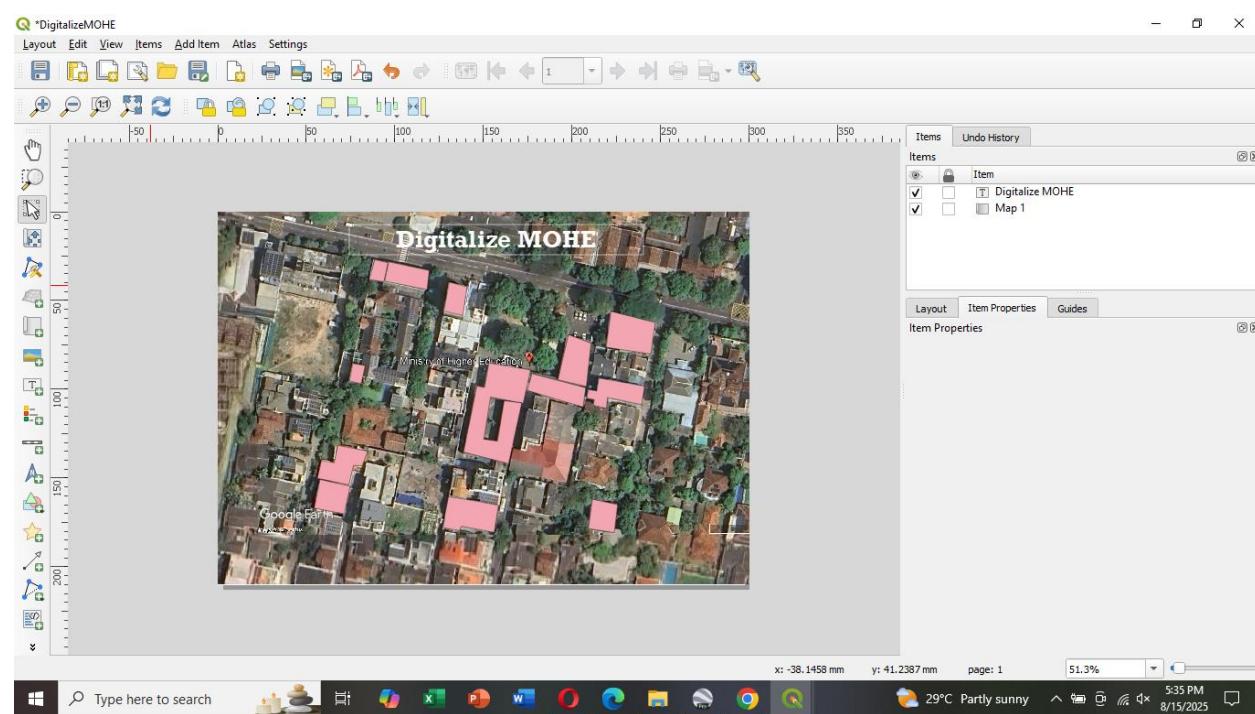
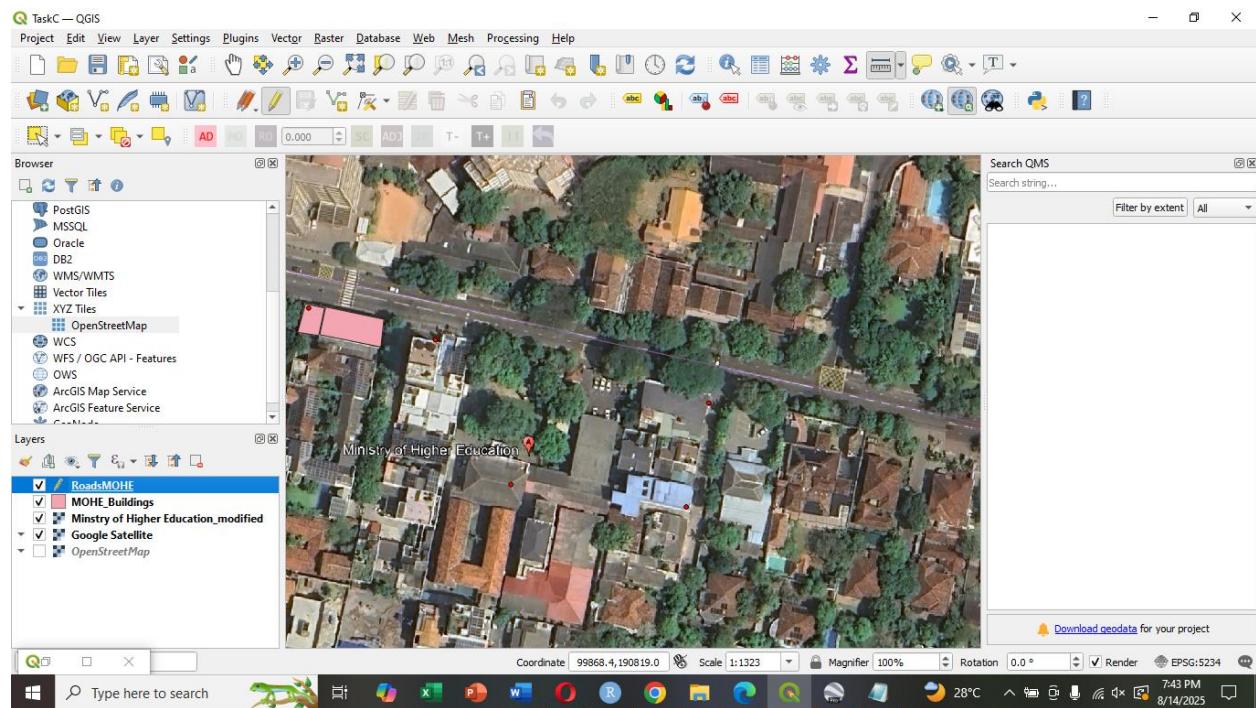
TASK – C

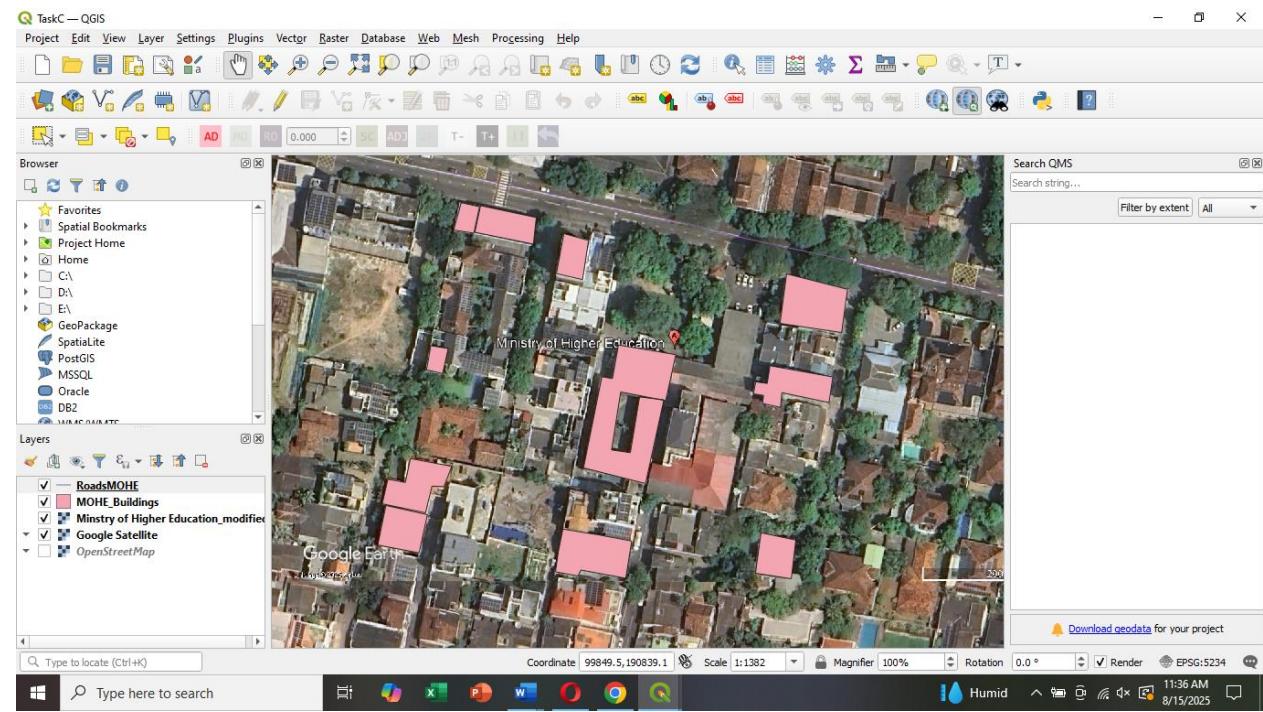
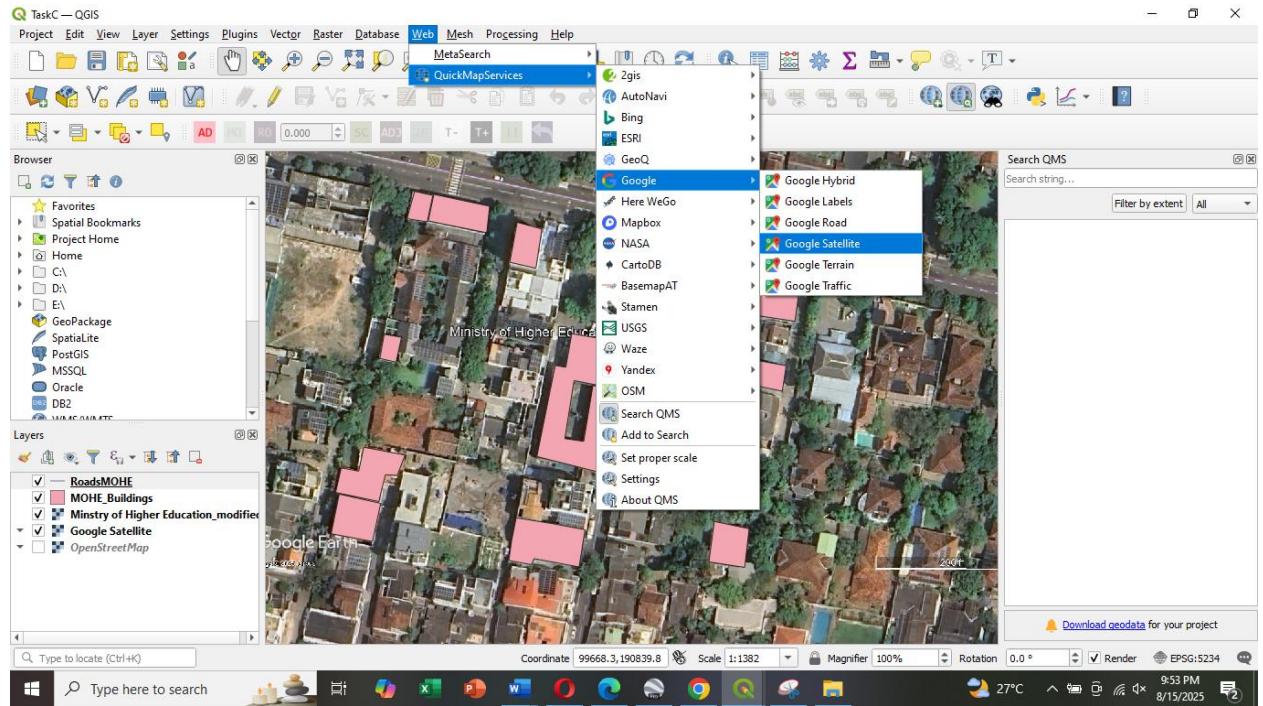


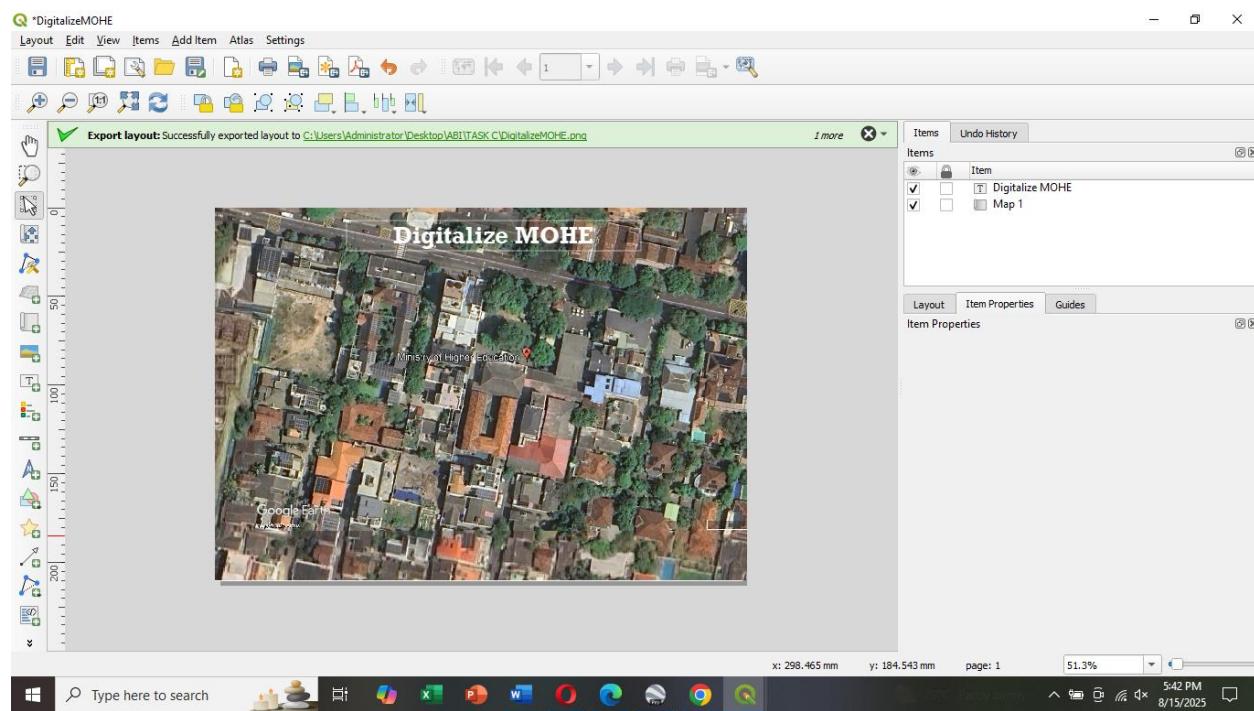
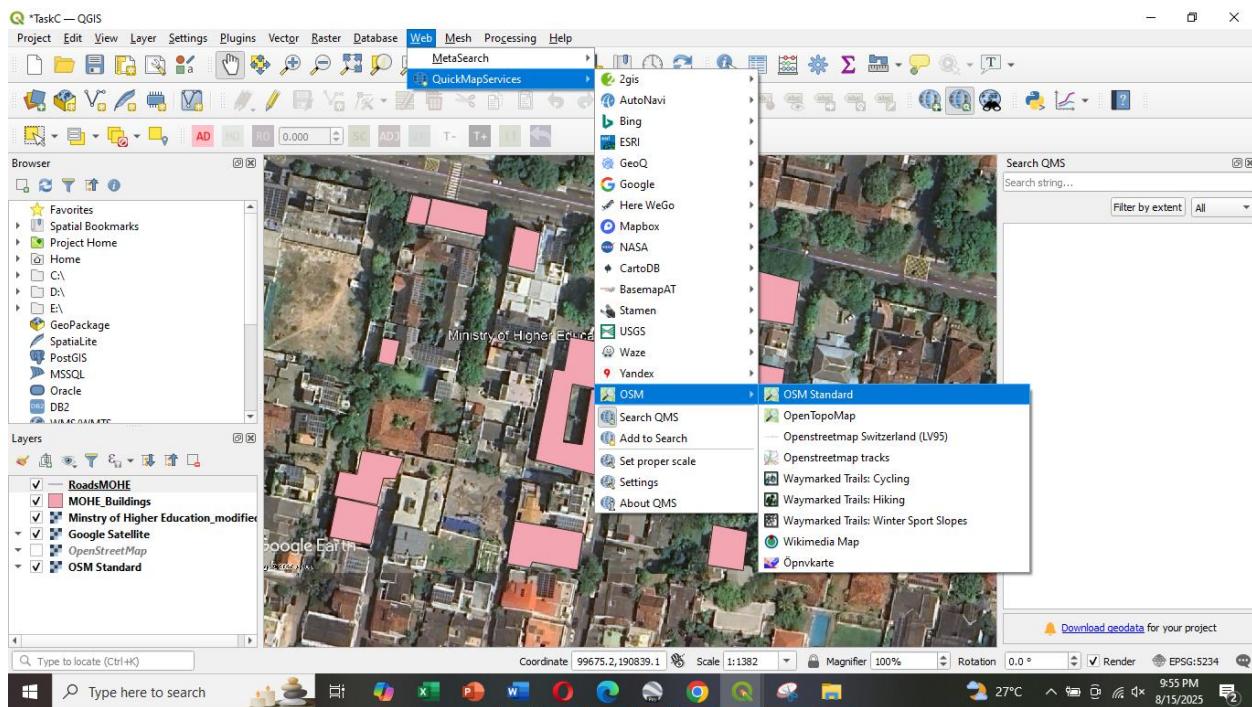




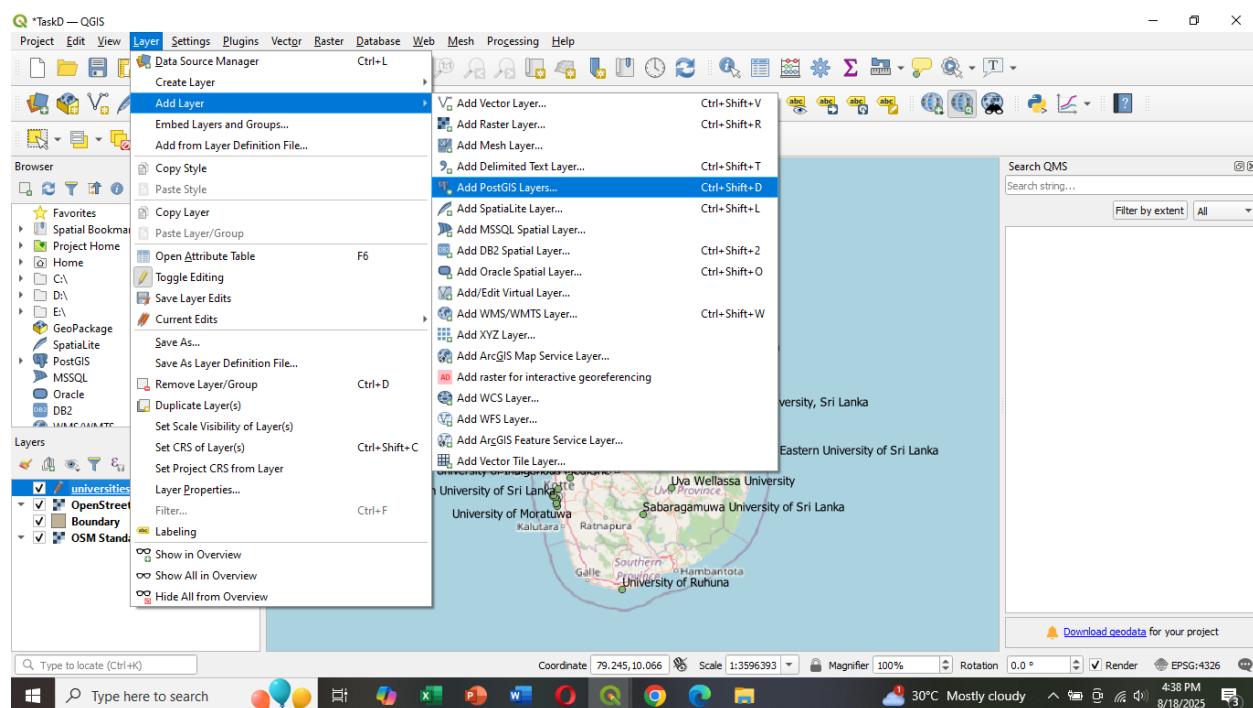
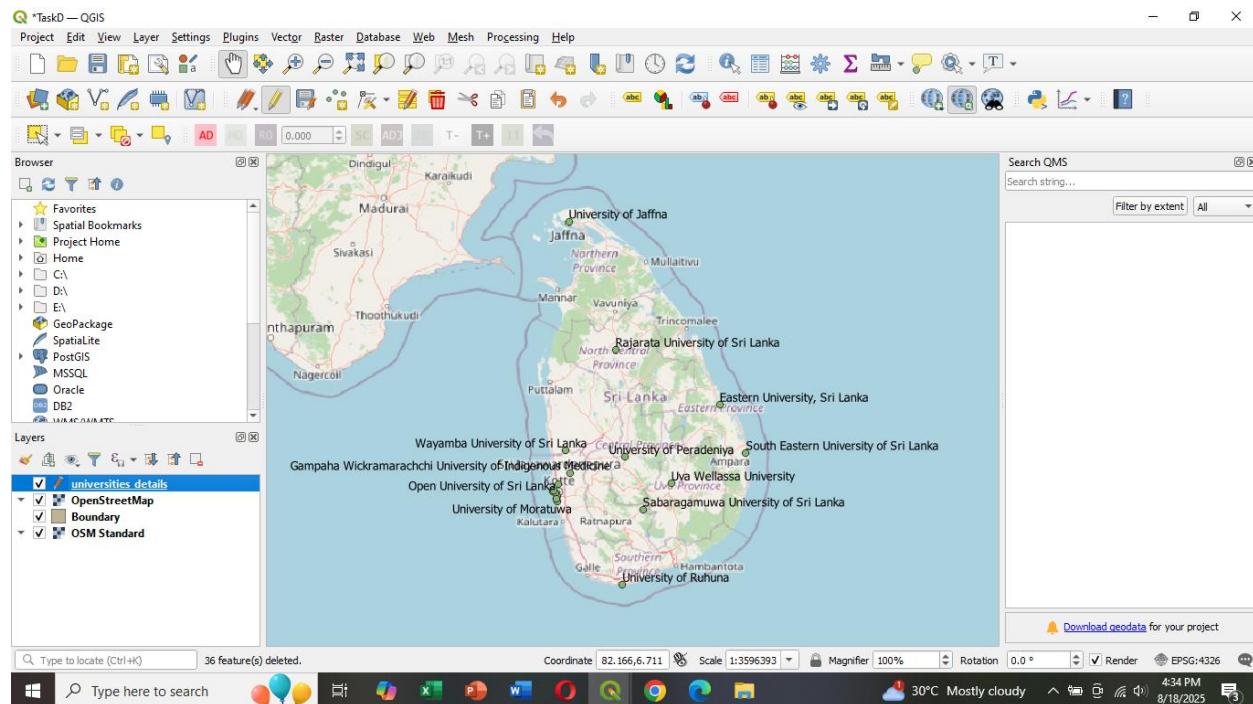


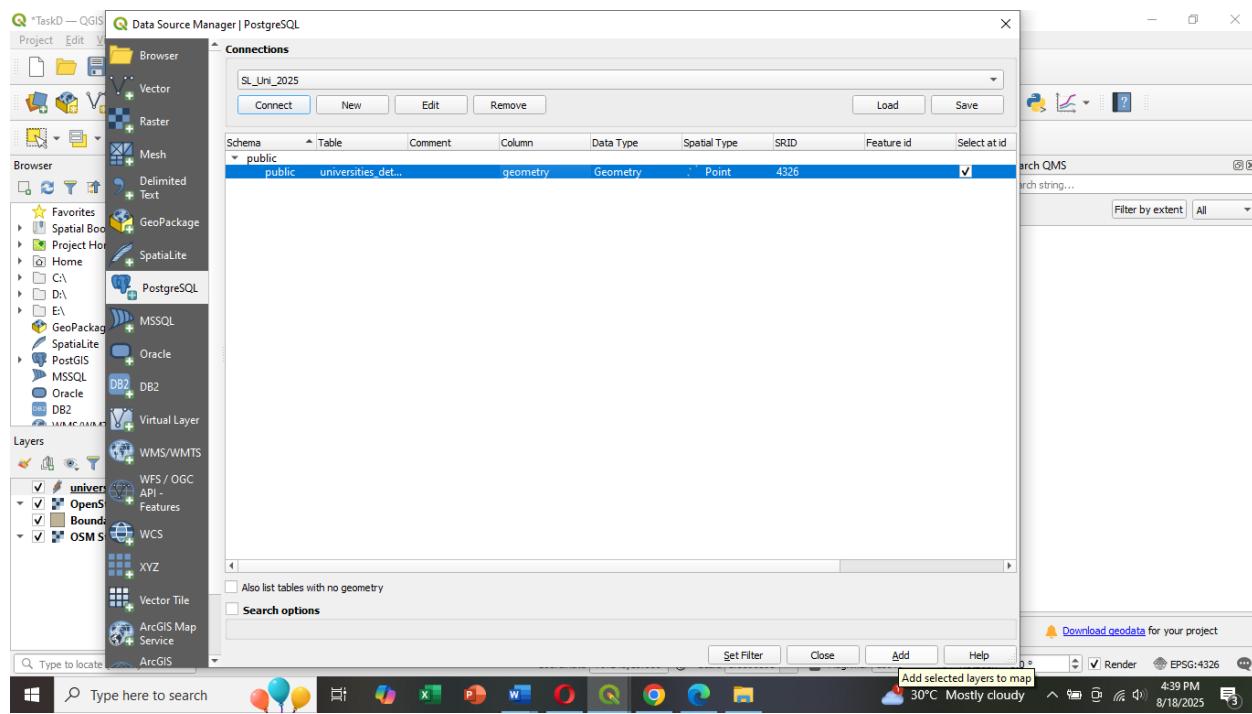
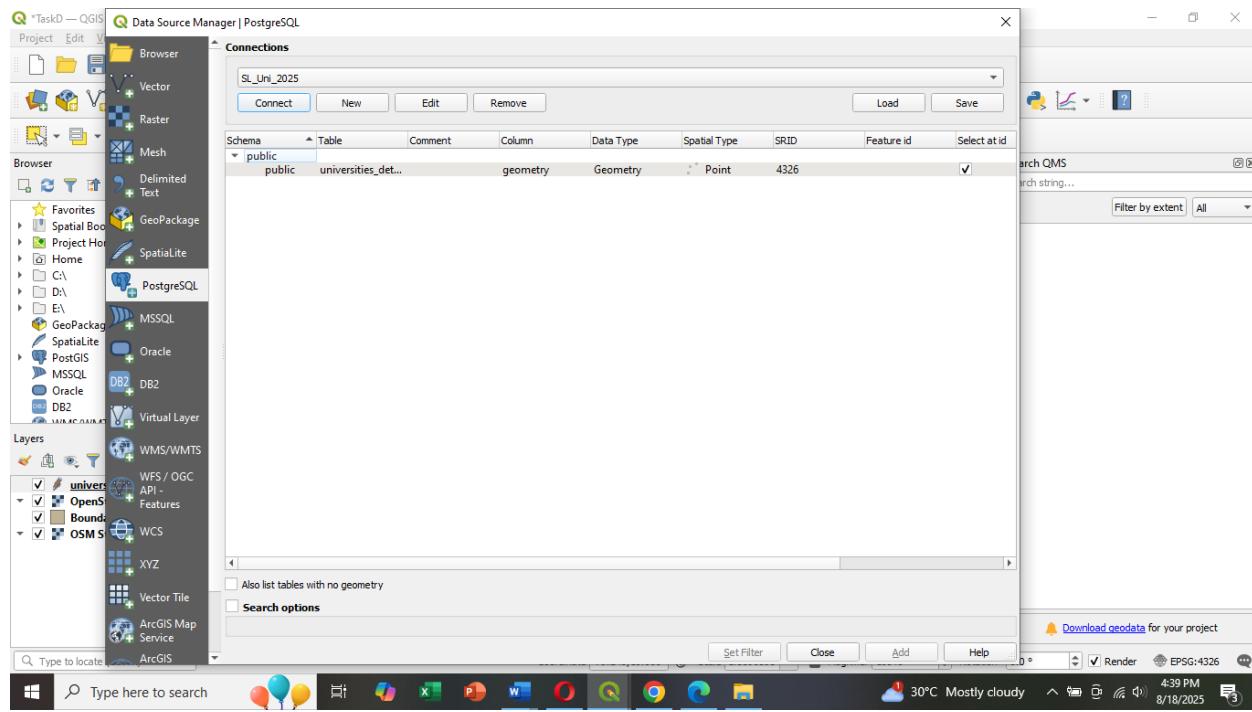


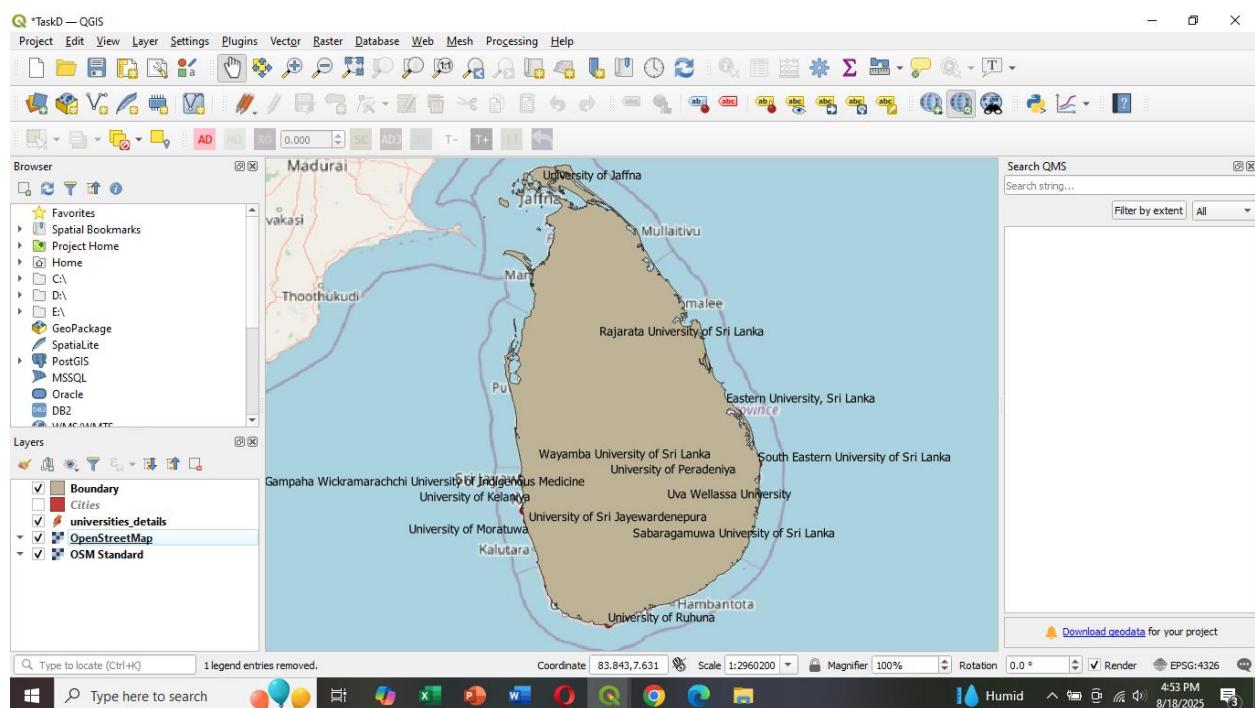
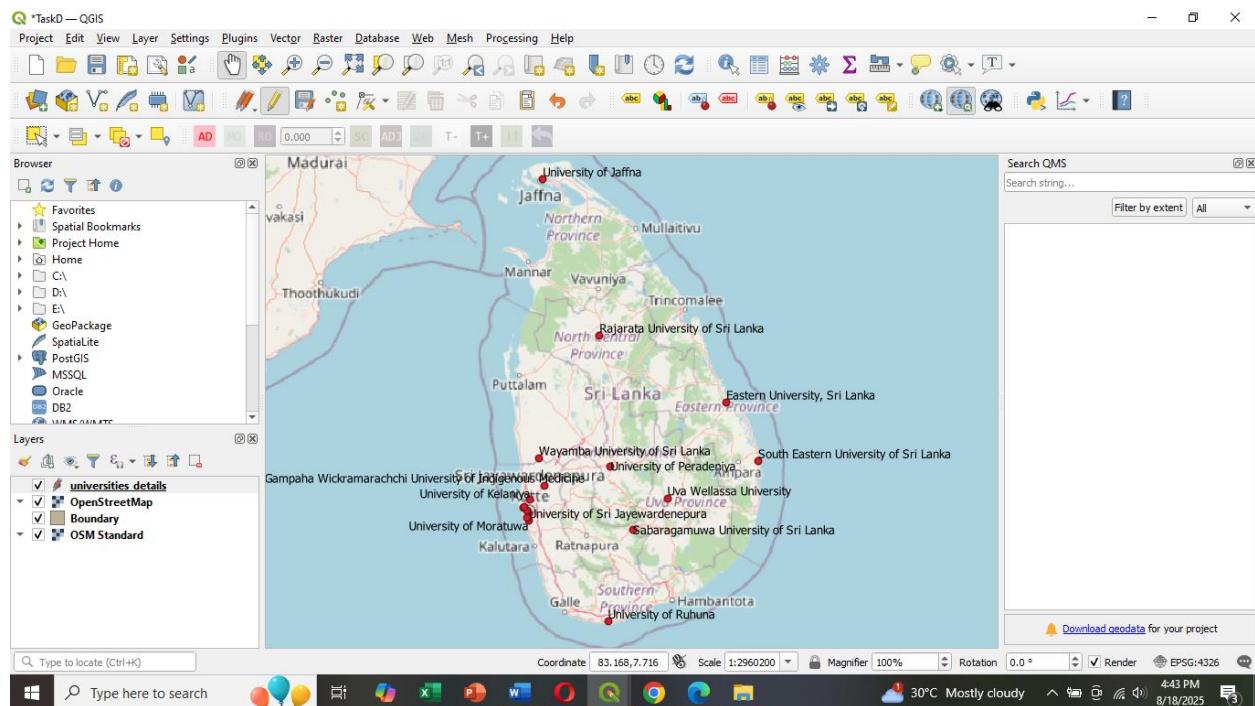


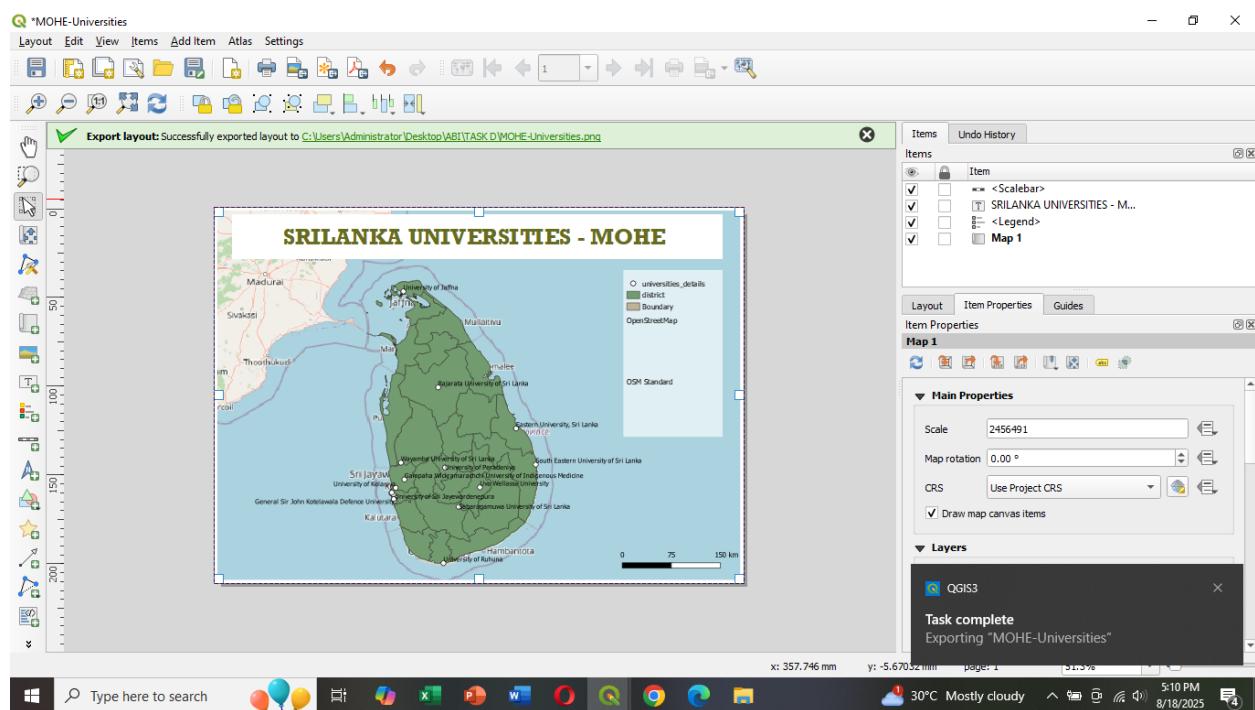
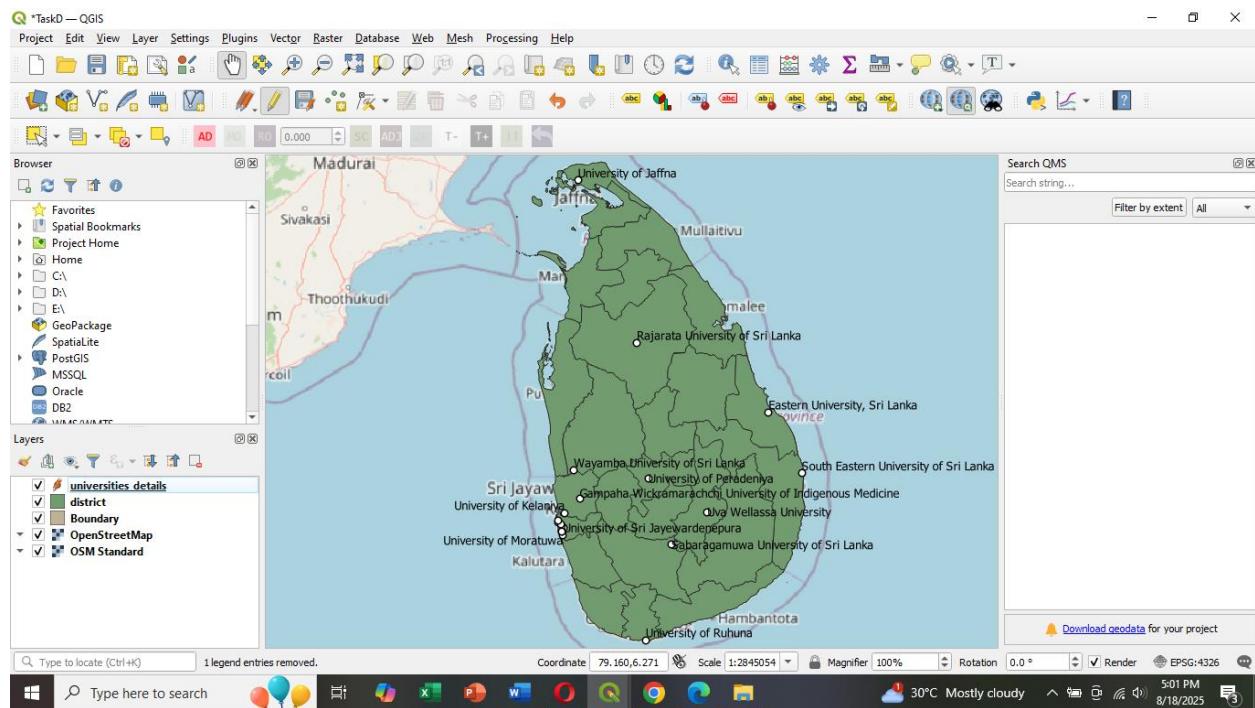


TASK – D

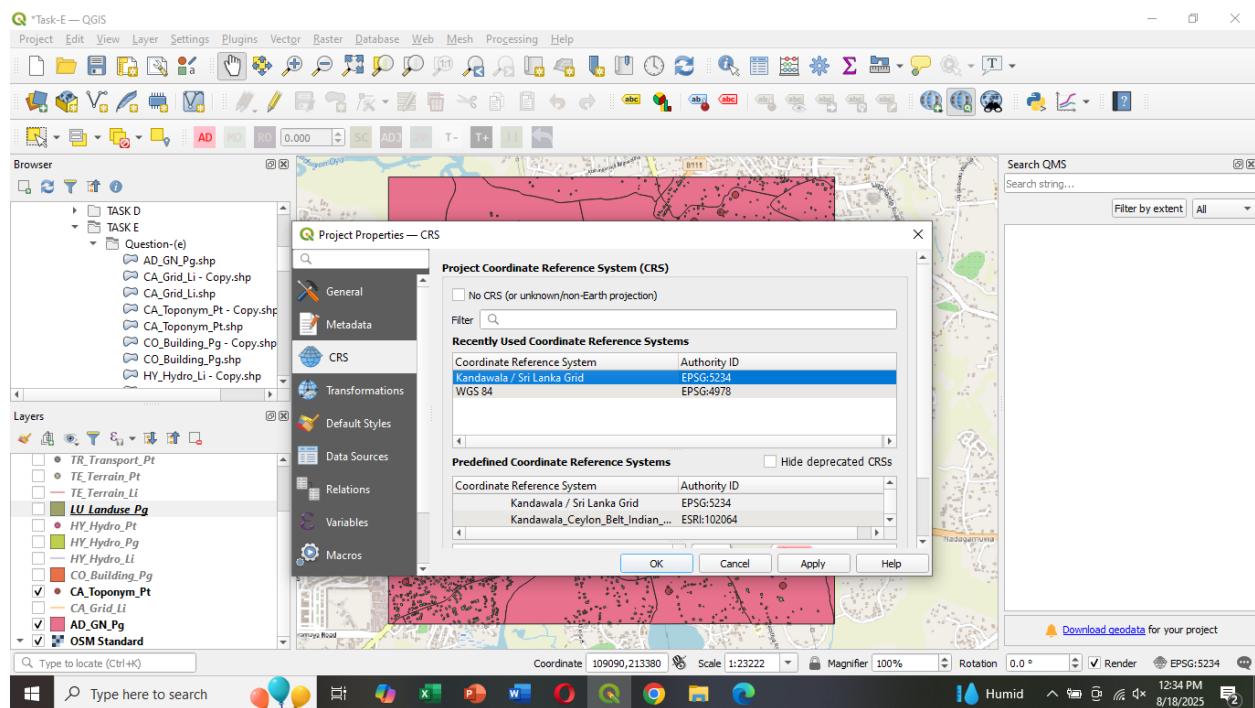
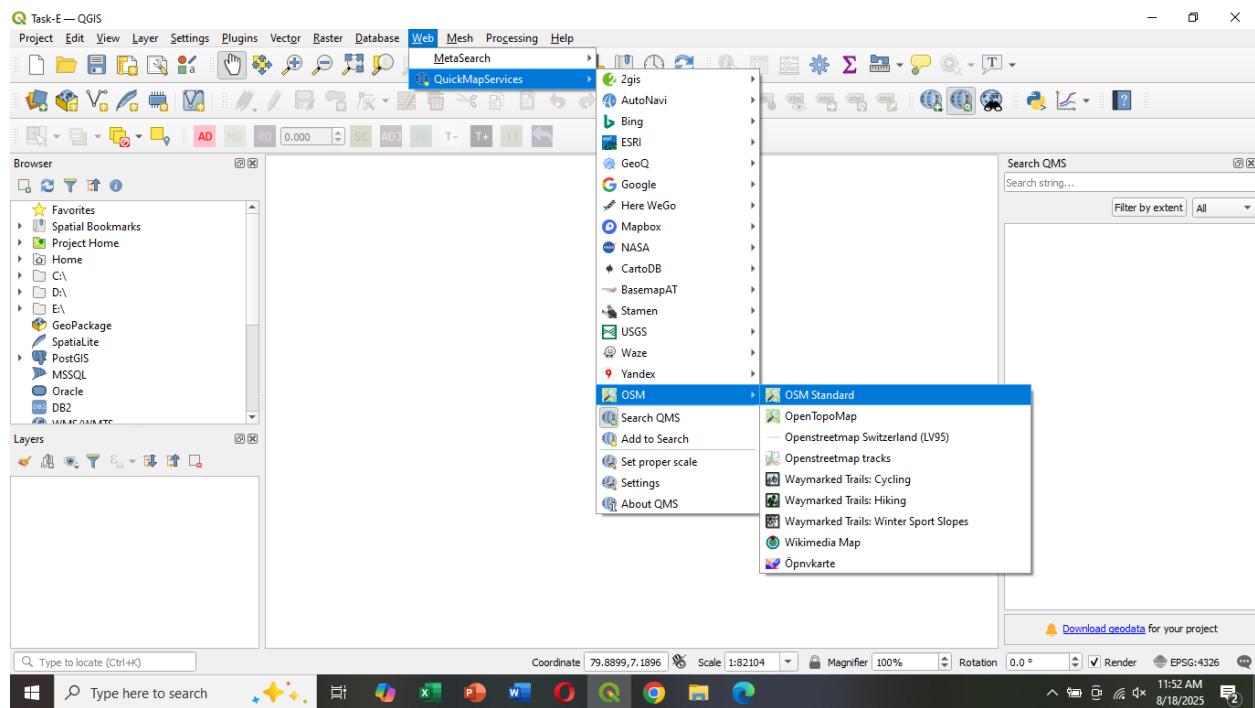


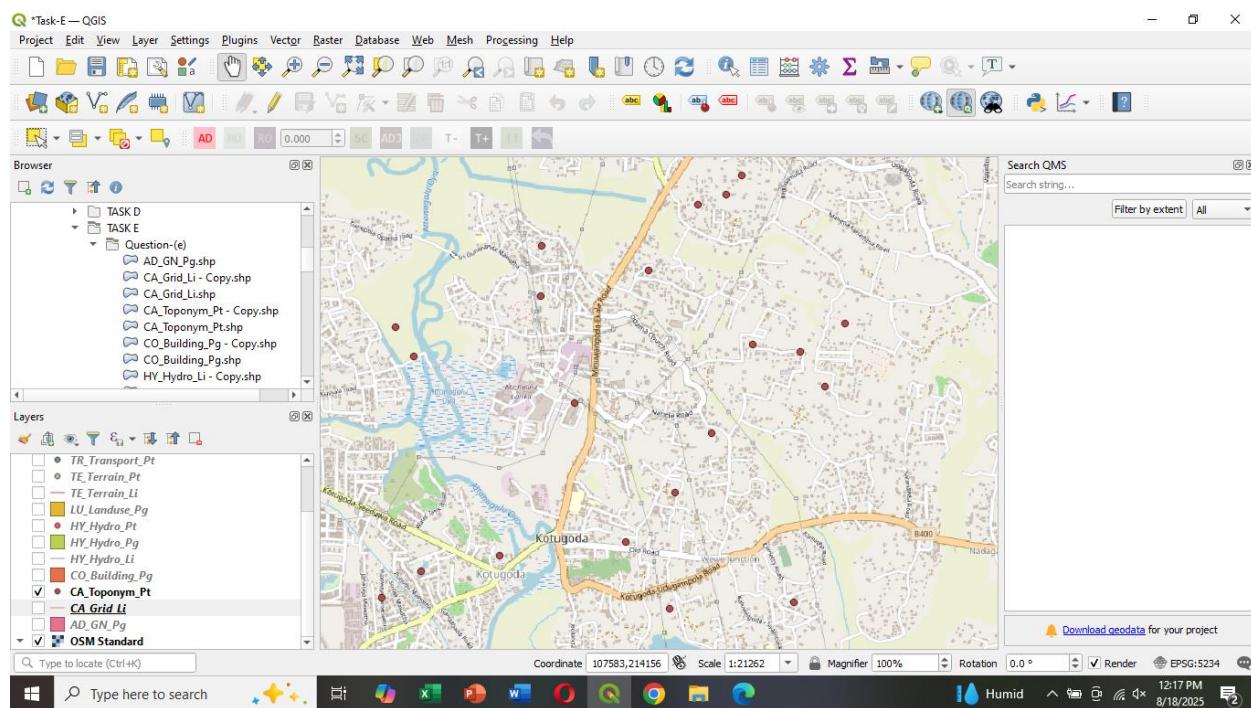
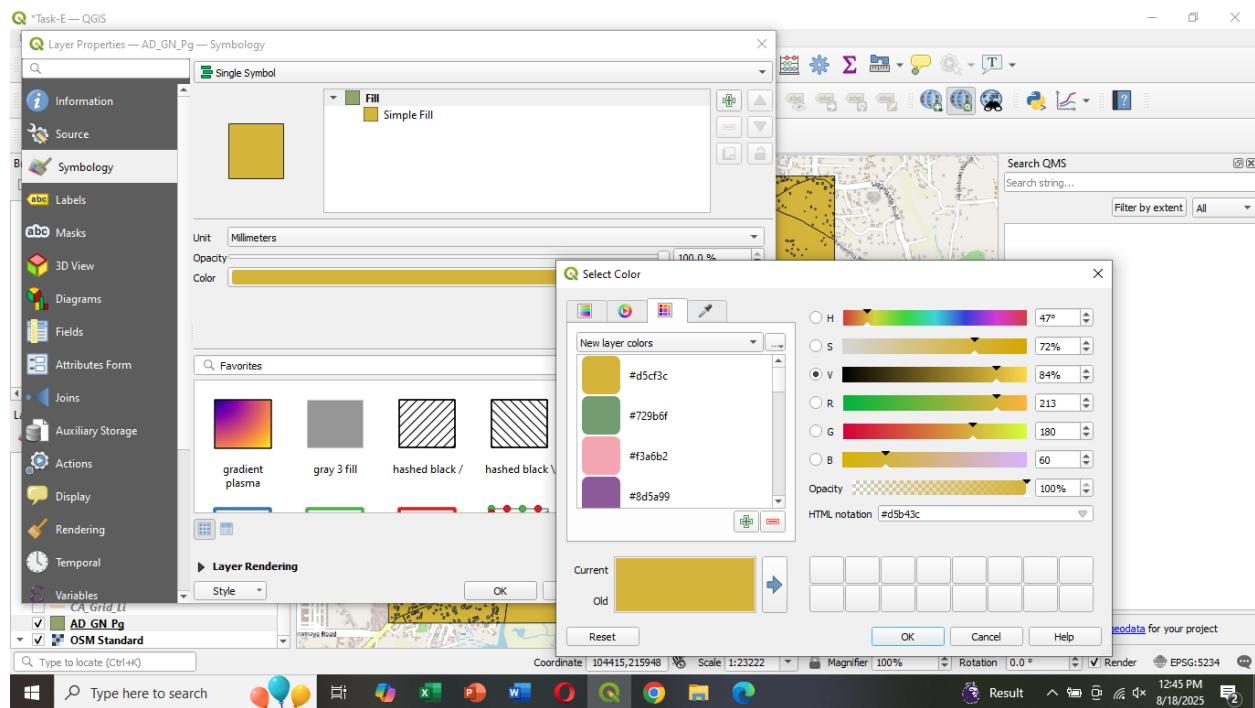


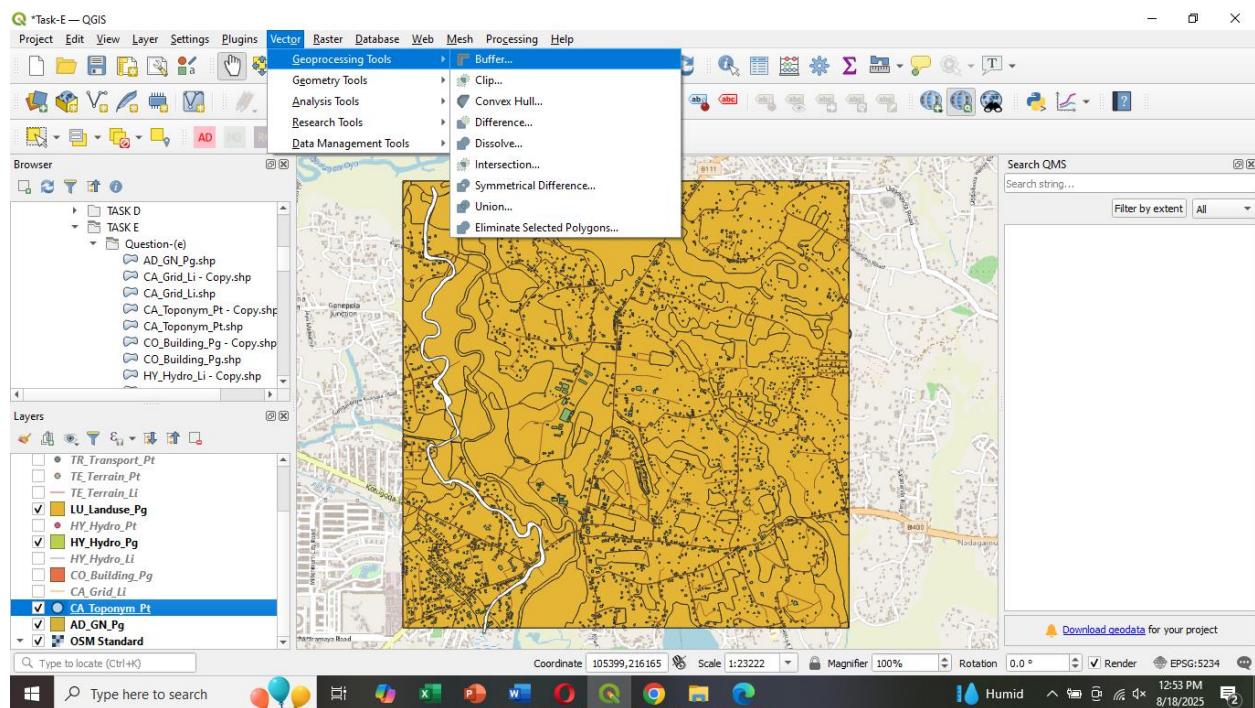
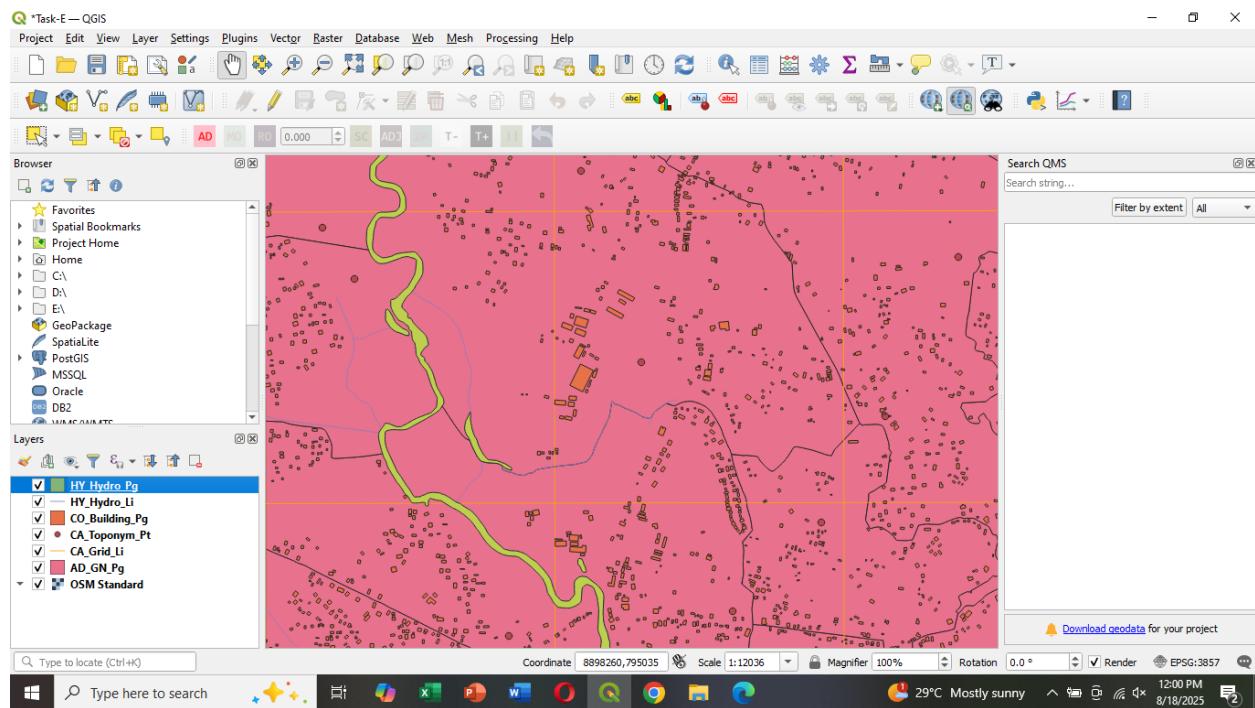


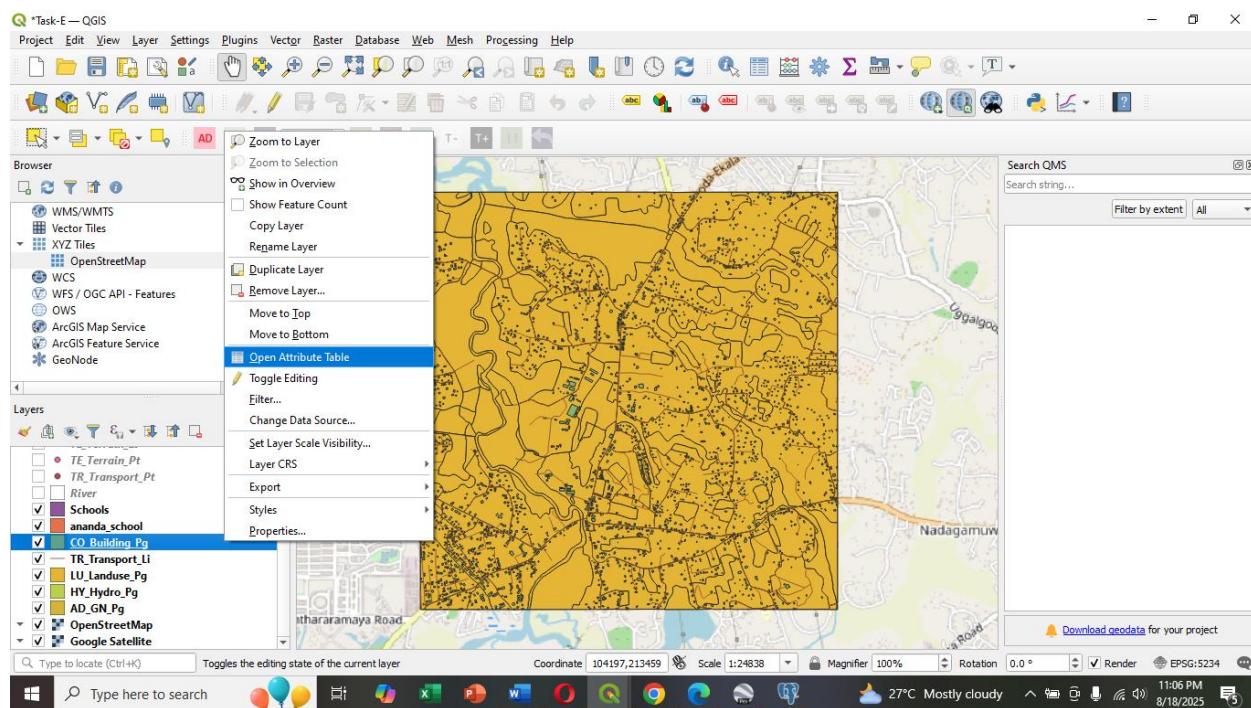
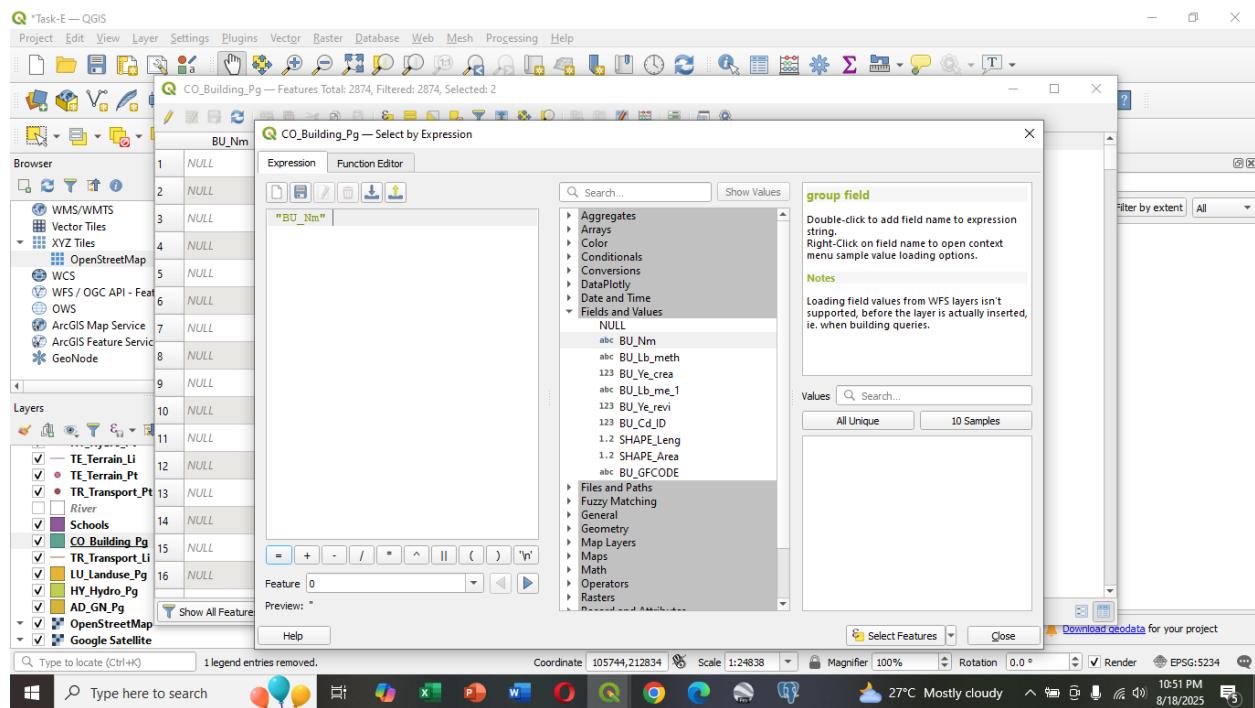


TASK – E









Q *Task-E — QGIS

Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help

CO_Building_Pg — Features Total: 2874, Filtered: 2874, Selected: 2

BU_Nm BU_Lb_meth BU_Ye_crea Select/filter features using form (Ctrl+F) U_Cd_ID SHAPE_Leng SHAPE_Area BU_GFCODE

1	NULL	8	2004	NULL	0	0	43.28749137710	109.67833931900	BLDG
2	NULL	8	2004	NULL	0	0	43.12944305460	91.68077699440	BLDG
3	NULL	8	2004	NULL	0	0	44.26791846440	113.06284972600	BLDG
4	NULL	8	2004	NULL	0	0	40.81239871820	104.08513918000	BLDG
5	NULL	8	2004	NULL	0	0	13.63317496000	11.53646593920	BLDG
6	NULL	8	2004	NULL	0	0	35.49403693310	78.17204004790	BLDG
7	NULL	8	2004	NULL	0	0	27.83361359760	48.35574194970	BLDG
8	NULL	8	2004	NULL	0	0	31.89003590750	62.85703762270	BLDG
9	NULL	8	2004	NULL	0	0	41.61864665390	101.12302024000	BLDG
10	NULL	8	2004	NULL	0	0	39.31719150240	96.12208605340	BLDG
11	NULL	8	2004	NULL	0	0	22.99241064770	27.78089125050	BLDG
12	NULL	8	2004	NULL	0	0	27.84749813180	48.46647105500	BLDG
13	NULL	8	2004	NULL	0	0	21.63538668430	27.80061052630	BLDG
14	NULL	8	2004	NULL	0	0	29.47640864440	54.28503279040	BLDG
15	NULL	8	2004	NULL	0	0	35.70089373370	79.01155294950	BLDG
16	NULL	8	2004	NULL	0	0	123.37702491500	831.60563800900	BLDG

Filter by extent All

Type to locate (Ctrl+K) 1 legend entries removed. Coordinate 104082,214392 Scale 1:24838 Magnifier 100% Rotation 0.0° Render EPSG:5234

Download geodata for your project

10:49 PM 8/18/2025

Q *Task-E — QGIS

Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help

Intersection

Parameters Log

Input layer Ananda_Buffer_1km [EPSG:5234]

Overlay layer eri_gunananda_buffer_2km [EPSG:5234]

Selected features only

Input fields to keep (leave empty to keep all fields) [optional] 0 options selected

Overlay fields to keep (leave empty to keep all fields) [optional] 0 options selected

Advanced Parameters

Intersection Users/Administrator/Desktop/ABI/TASK E/schools_intersection.gpkg

Open output file after running algorithm

Run as Batch Process... Run Close Help

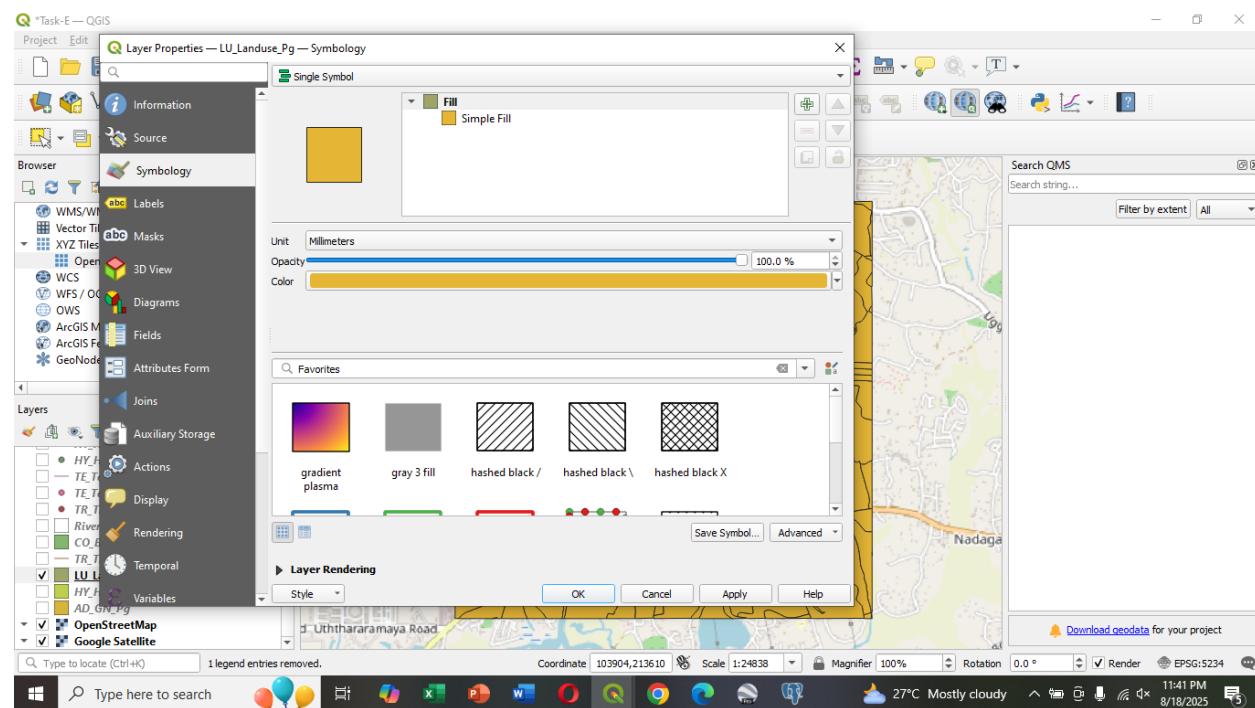
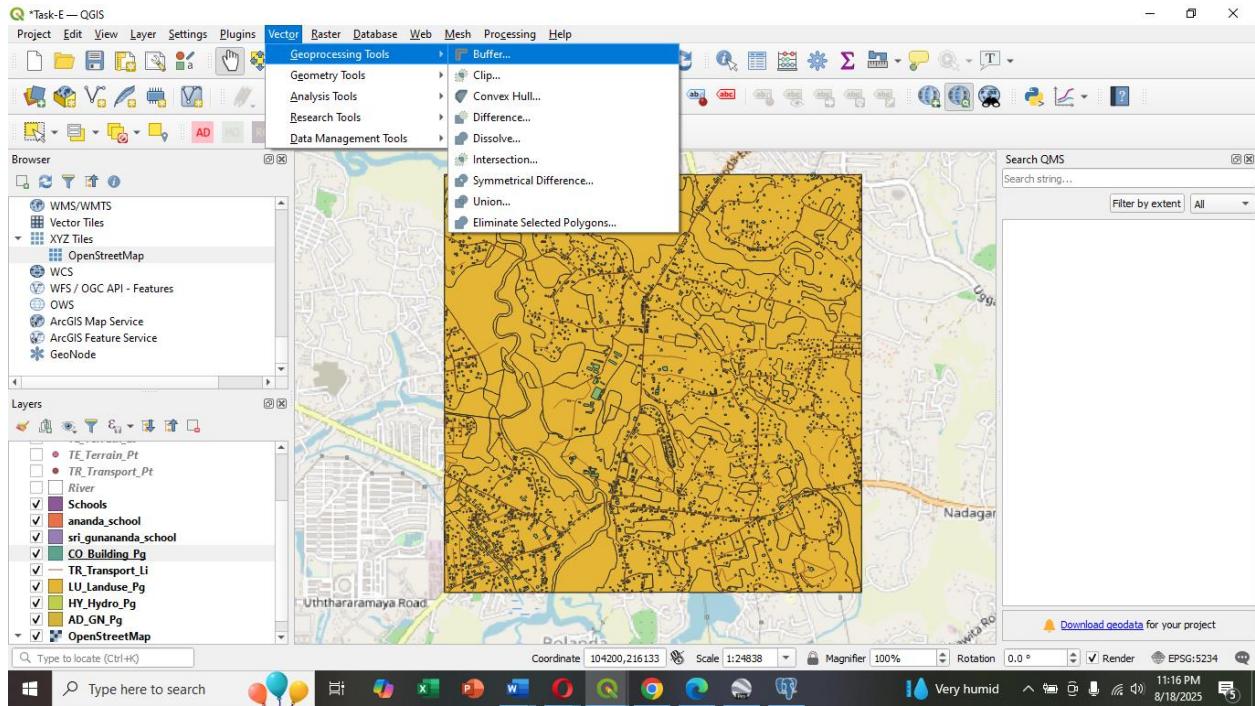
0%

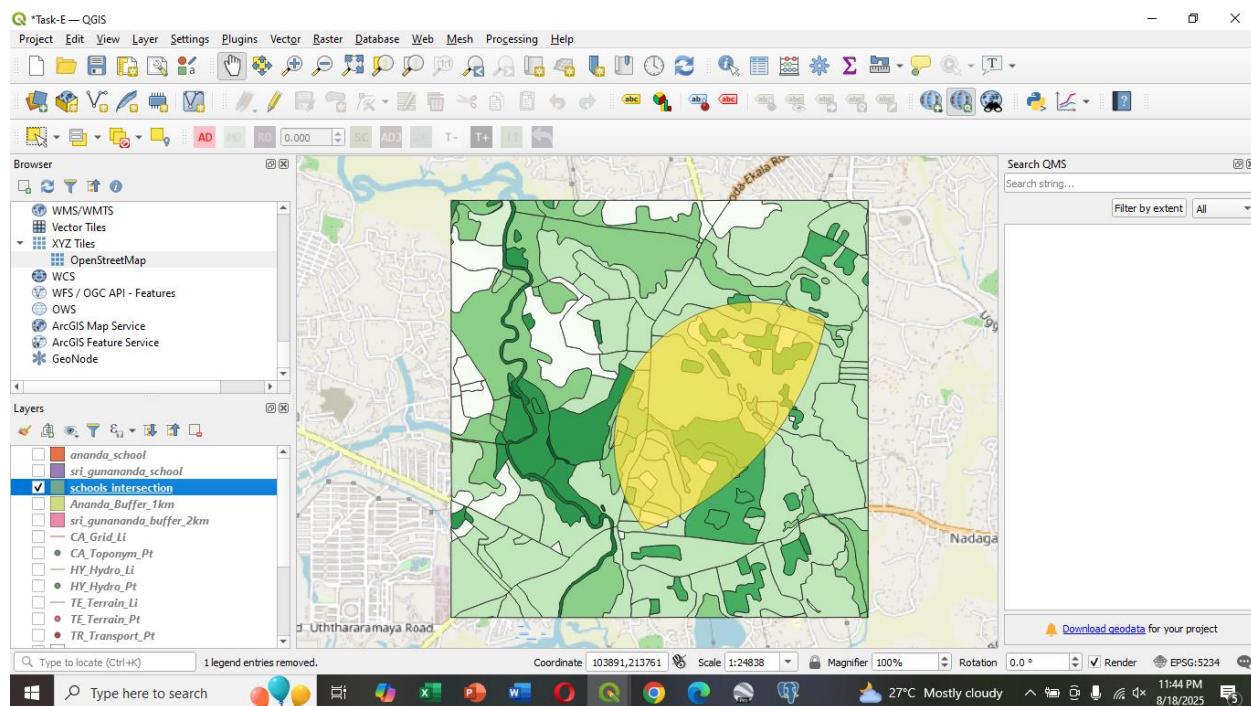
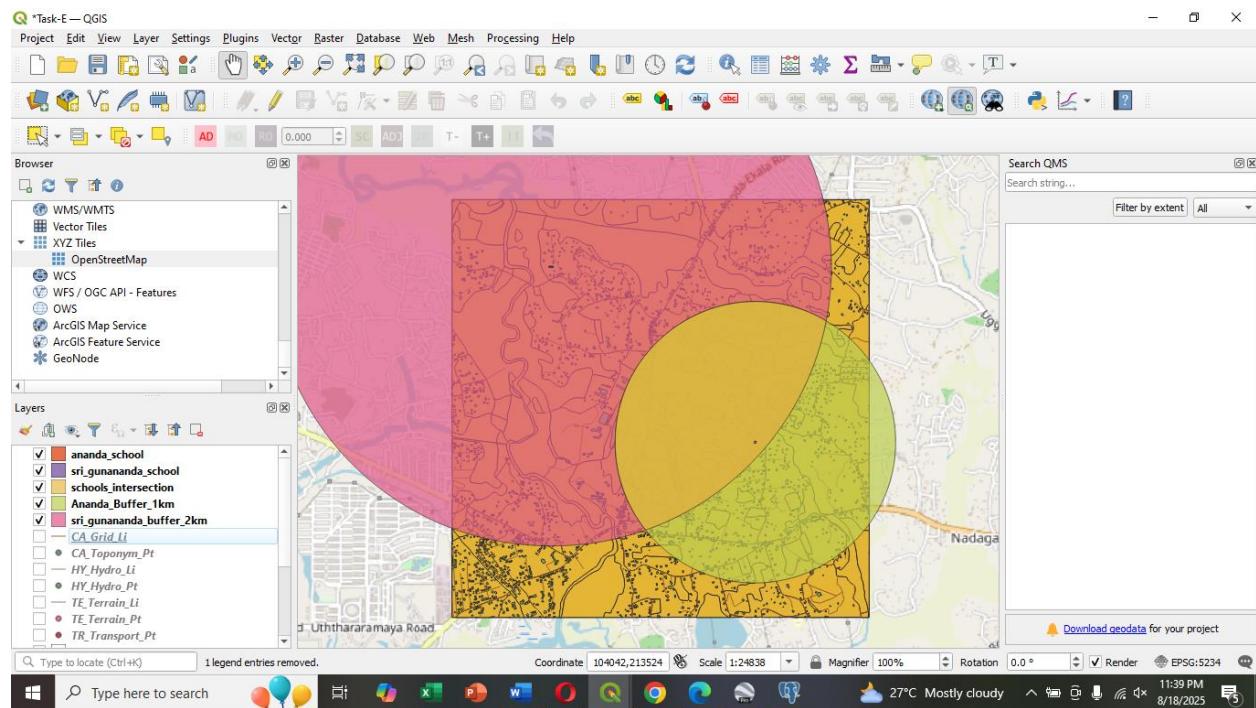
Search QMS Search string... Filter by extent All

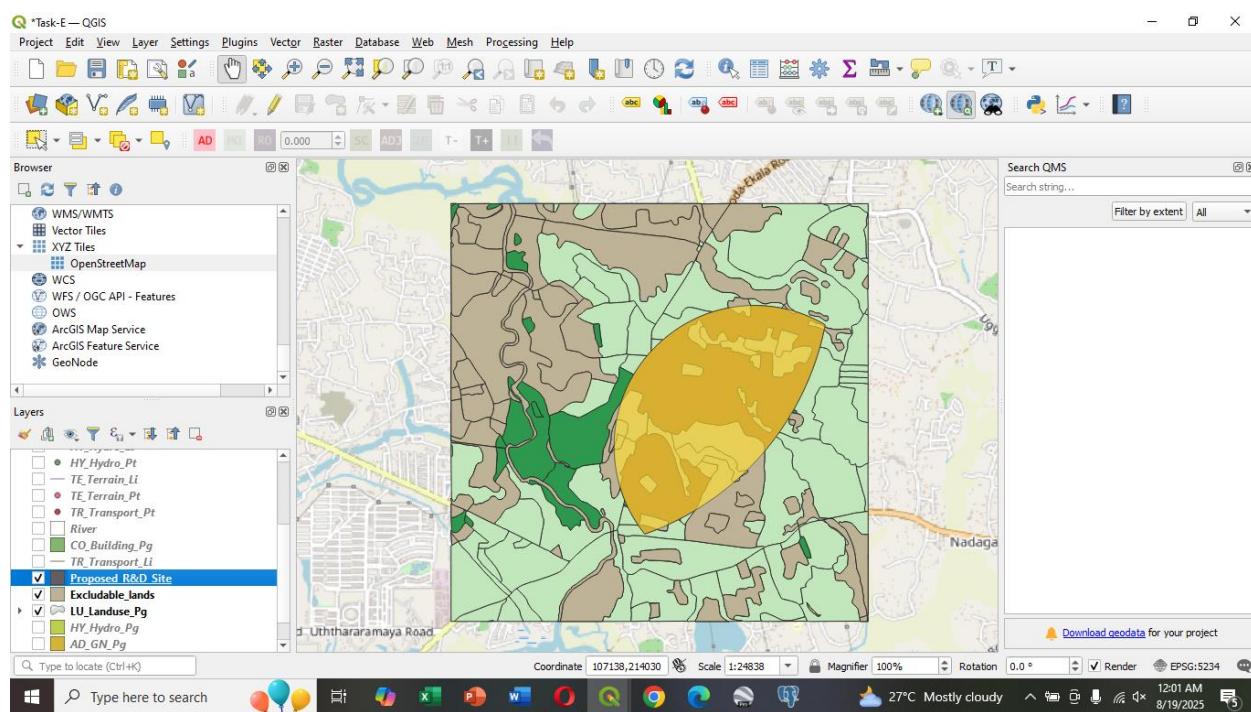
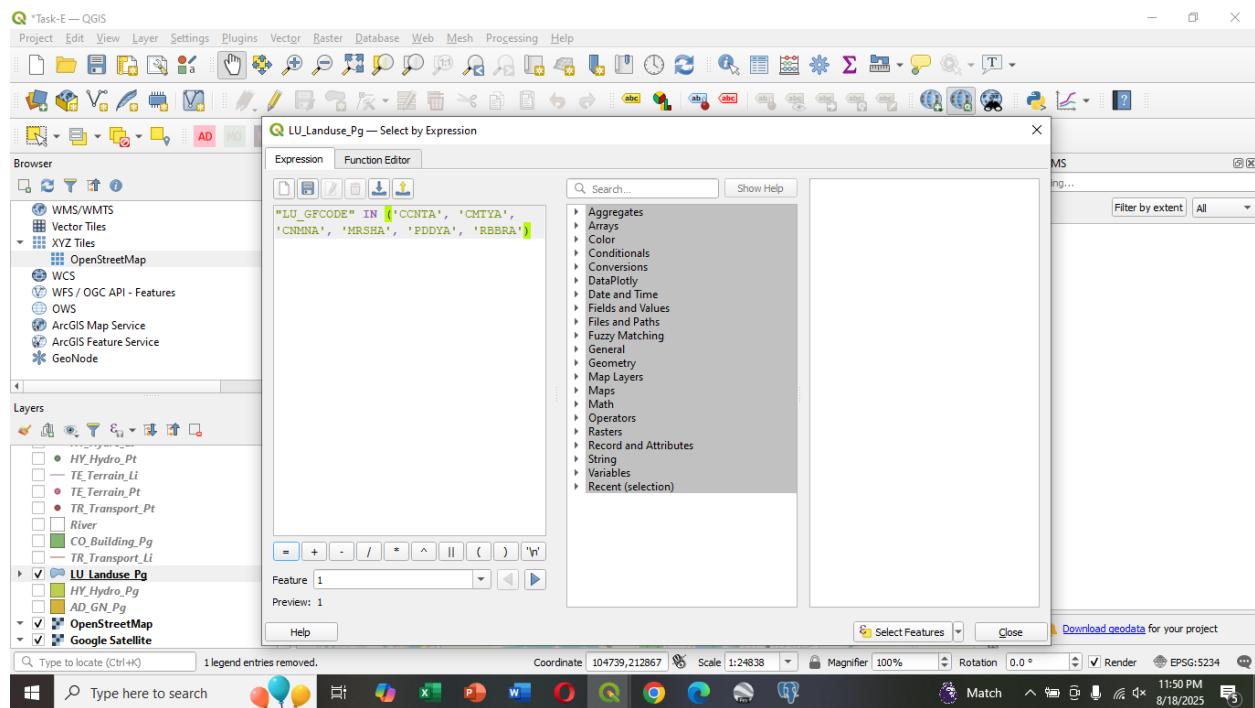
Download geodata for your project

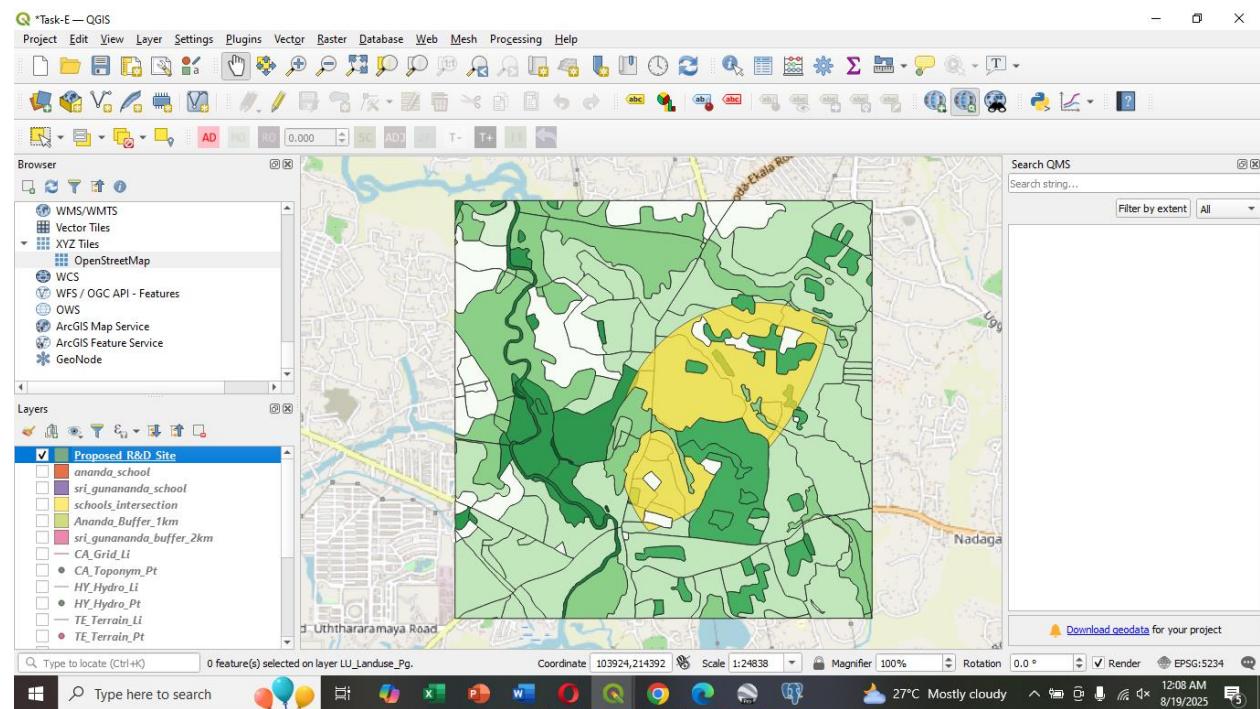
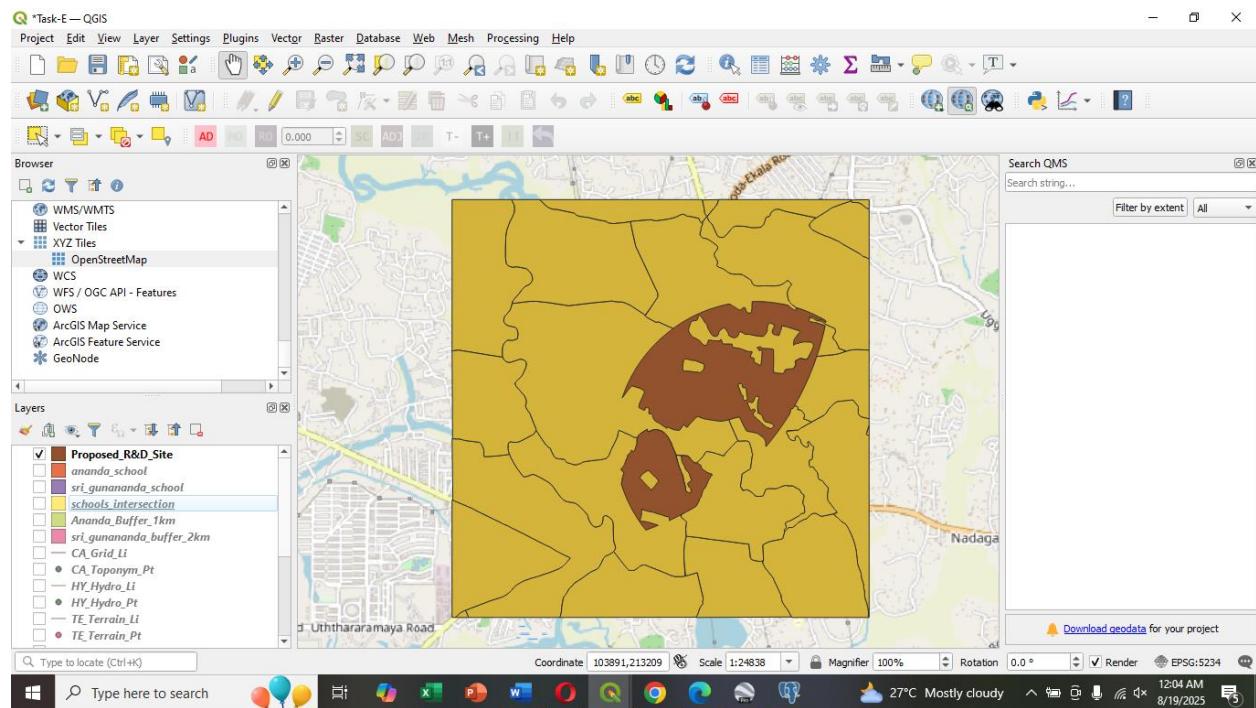
Type to locate (Ctrl+K) Coordinate 104259,216160 Scale 1:24838 Magnifier 100% Rotation 0.0° Render EPSG:5234

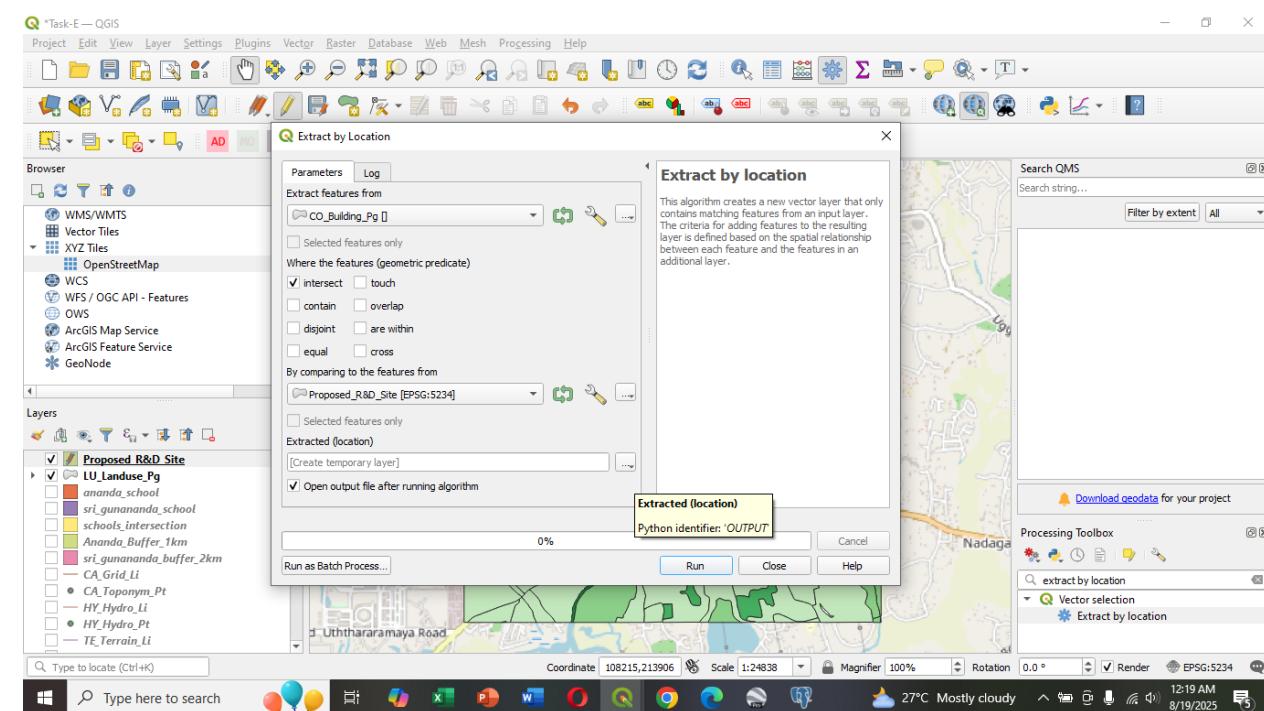
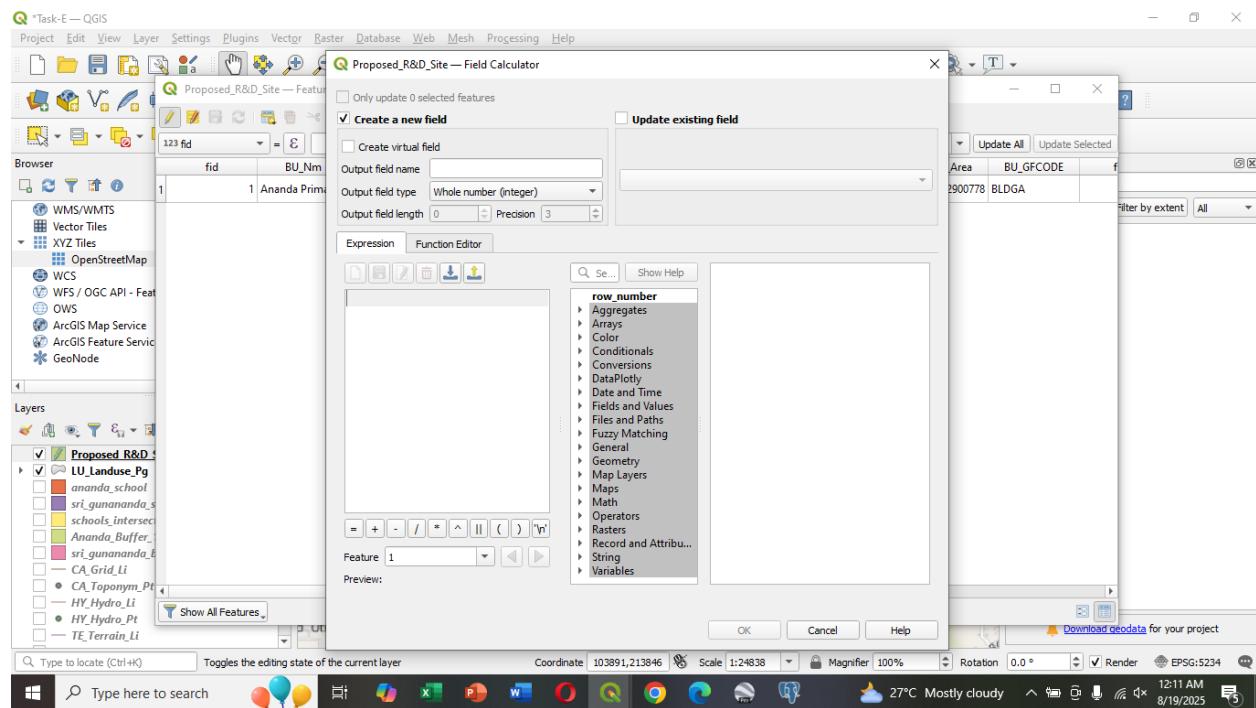
27°C Mostly cloudy 11:31 PM 8/18/2025











Q *Task-E — QGIS

Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help

Buildings_in_proposed_site — Features Total: 391, Filtered: 391, Selected: 0

123 fid	BU_Nm	BU_Lb_meth	BU_Ye_crea	BU_Lb_me_1	BU_Ye_revi	BU_Cd_ID	SHAPE_Leng	SHAPE_Area	BU_GFCODE	Area_sqm	
1	1	NULL	8	2006	NULL	0	0	36.1238420017	64.8142815214	BLDG	64.81023876214...
2	2	NULL	8	2006	NULL	0	0	41.223472176	104.903971957	BLDG	104.8961989916...
3	3	NULL	8	2006	NULL	0	0	18.6686691804	21.7643509497	BLDG	21.76287419480...
4	4	NULL	8	2006	NULL	0	0	31.266393721	60.8440882334	BLDG	60.83986937382...
5	5	NULL	8	2004	NULL	0	0	30.0136335471	42.6781590246	BLDG	42.67590362804...
6	6	NULL	8	2004	NULL	0	0	43.9198818548	116.678724303	BLDG	116.6691515816...
7	7	NULL	8	2004	NULL	0	0	41.4860708353	87.2982534475	BLDG	87.2920272103148
8	8	NULL	8	2004	NULL	0	0	25.0490315295	34.5982271037	BLDG	34.59589573985...
9	9	NULL	8	2004	NULL	0	0	61.3095856933	192.725178131	BLDG	192.7103665324...
10	10	NULL	8	2004	NULL	0	0	67.3723228929	233.869485807	BLDG	233.8520527371...
11	11	NULL	8	2004	NULL	0	0	23.643997339	33.5449504064	BLDG	33.5427100306...
12	12	NULL	8	2004	NULL	0	0	33.2795625941	69.0481194665	BLDG	69.04324534470...
13	13	NULL	8	2004	NULL	0	0	42.8855584339	107.058488996	BLDG	107.0508023573...
14	14	NULL	8	2004	NULL	0	0	26.8696408319	43.0339942473	BLDG	43.03105306842...
15	15	NULL	8	2004	NULL	0	0	23.9909182611	35.966620088	BLDG	35.96417953775...

Extract by location

Type to locate (Ctrl+K) Toggles the editing state of the current layer Coordinate 103891,216179 Scale 1:24838 Magnifier 100% Rotation 0.0° Render 12:28 AM 8/19/2025

Q *Task-E — QGIS

Project Edit View Layer Settings Plugins Vector Raster Database Web Mesh Processing Help

Buildings_in_proposed_site — Features Total: 391, Filtered: 391, Selected: 0

Search QMS

Processing Toolbox

Results Viewer

Almond tree Statistics

27°C Mostly cloudy 1:09 AM 8/19/2025

