Cardiff Metropolitan University
B.Sc. (Hons) in Business Information Systems
<u>Assignment Cover Sheet</u>

Student Details (Student should fill the content)			
Name	Mohamed Hirsi Saadh		
Student ID	ST20208732		
Scheduled unit details			
Unit code	CIS6008		
Unit title	Analytics and Business Intelligence		
Unit enrolment details	Year	3	
	Study period	2023	
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	3200(report)		
Due date / Time	06st August 2023		

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	Mohamed Hirsi Saadh	Date	06/09/2023

Submission	
Return to:	

Result				
Marks by 1st Assessor		Name & Signature of the 1st Assessor		**Agreed Mark**
Marks by2nd Assessor		Name & Signature of the 2nd Assessor		
Comments on the Agreed mark				

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

<table>
<tr><td colspan="2">STUDENT NAME: MOHAMED HIRSI SAADH</td><td>STUDENT NUMBER:
ST20208732</td></tr>
<tr><td colspan="2">Module Number & Title: Analytics and Business Intelligence</td><td>Semester: II</td></tr>
<tr><td colspan="3">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.</td></tr>
<tr><td colspan="3">For student use: Critical feedback on the individual progression towards achieving the assignment outcomes</td></tr>
<tr><td colspan="3"></td></tr>
<tr><td colspan="3"><u>For the Assessors’ feedback</u>
Indicate the Task number strength and Weaknesses and the marks for each task</td></tr>
<tr><td>Task No/Question No</td><td>Strengths (1st Assessor)</td><td>Strengths (2nd Assessor)</td></tr>
</table>

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 2ndAssessor :		**Date :**	
Name & Signature of the IV: (If any)		**Date :**	

M.H. SAADH - ST20208732 - BA.docx

ORIGINALITY REPORT

8%	4%	3%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Wales Institute, Cardiff Student Paper	2%
2	Submitted to The Robert Gordon University Student Paper	1%
3	www.coursehero.com Internet Source	1%
4	www.utm.my Internet Source	1%
5	Cecilia Soldatini, Yuri Vladimir Albores-Barajas, Tomas Lovato, Adriano Andreon et al. "Wildlife Strike Risk Assessment in Several Italian Airports: Lessons from BRI and a New Methodology Implementation", PLoS ONE, 2011 Publication	1%
6	apps.dtic.mil Internet Source	<1%
7	www.researchgate.net Internet Source	<1%

Info

Submission Details

Student ID	st20208732@outlook.cardiffmet....
Class Name	Analytics and Business Intelligen...
Class ID	6847392
Submission ID	212201499
Submission Date	03-Sep-2023 10:19PM (UTC+0100)
Submission Count	1
File Name	82592_HirsiMohomed_Saadh_M_...
File Extension	docx
File Size	29.06M
Character Count	19401
Word Count	3591
Page Count	37

ACKNOWLEDGEMENT

Foremost, I, Mohamad Hirsi Saadh, a student in batch 10 of the BSc in Computing and Software Engineering, would like to convey my heartfelt gratitude and appreciation to Mr. Asanka Dinesh, lecturer in Computational Intelligence for his unwavering support, patience, encouragement, and passion. Throughout the execution of my work, he provided me with a lot of information about this module. Aside from my lecturer, I'd want to express my gratitude to my other classmates and for their assistance and support in completing my work.

In addition, I'd want to thank the entire ICBT management and team for being so friendly and helpful, as well as for giving me this chance to showcase my skills via an assignment.

Finally, I'd want to convey my heartfelt appreciation to my family for their unwavering support and optimistic attitude throughout this report.

EXECUTIVE SUMMARY

This report presents a comprehensive analysis employing Geo Spatial and Statistical Analysis methods across various sectors. Divided into eight chapters, it delves into critical aspects of data analytics, each contributing valuable insights and solutions.

Chapter one investigates energy efficiency in building cooling in the USA, focusing on factors to reduce electricity consumption. Chapter two develops a statistical model for predicting cooling load using regression analysis based on Chapter 1's data.

Chapter three maps land suitable for wind power plants across Sri Lanka using data from the Renewable Energy Resource Development Plan and Chapter four highlights potential land for solar power development in Sri Lanka and assesses its role in addressing energy sector challenges.

The next chapter digitizes and preserves an informative area map in Hambantota, incorporating solar power plants, forests, trees, and suburbs using shapefiles. While chapter six performs Geo Spatial analysis to create a thematic map, visualizing data from the Sri Lanka Sustainable Energy Authority within the PostgreSQL database "SLPetrolium-2023."

Chapter seven identifies feasible locations for renewable energy generation in Sri Lanka, pinpointing them on a map using Google Earth Pro and QGIS. And chapter eight determines suitable land for the "Regional Research Center for Renewable Energy" in Kandy, Sri Lanka, based on specific criteria.

The report concludes with a synthesis of findings from each chapter. Appendices contain supporting evidence. This extensive analysis underscores the critical role of data analytics in diverse sectors, offering solutions and insights for energy efficiency and renewable energy development.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

QGIS	Quantum Geographic Information System
SL	Sri Lanka
USA	United States of America

INTRODUCTION

In today's world data analytics is crucial in almost every sector in the world. To maximize the use of analytics there are various methods. Some are Geo Spatial analysis and Statistical Analysis. Geo Spatial analysis mostly consists of maps and geographical data and Statistical Analysis consists of graphs, charts, and predictions along with other insights. This report consists of some both these analyses conducted in various sectors.

The first chapter consists of an analysis related to the energy efficiency in building cooling and its associations with various other factors to reduce residential and commercial electricity consumption. The data are obtained from a case study in the USA.

The second chapter is all about creating a statistical model to predict the cooling load by using the data obtained from chapter one. Regression analysis must be done to complete the chapter.

The third chapter consists of maps that are developed to showcase the lands that are available to develop wind power plants in each district. The data for the maps are obtained from the Renewable Energy Resource Development Plan of Sri Lanka.

The fourth chapter consists of a map that has been created to showcase the potential land for solar power development by using the data obtained from the Sri Lanka sustainable energy authority. Along with an analysis of how solar power energy would contribute for overcoming energy sector problems.

The fifth consists of the digitized informative area map of a place in Hambantota. The solar power plants, forests, trees, and other suburbs are digitized and saved as shape files in this chapter.

The sixth consists of geo spatial analysis done to visualize a thematic map classified by using the data provided by the Sri Lanka Sustainable Energy Authority. A database using PostgreSQL named "SLPetrolium-2023" is created and the data are visualized in the map.

The seventh chapter is about feasible locations for Renewable Energy Generation in Sri Lanka identified by the Sri Lanka Energy Authority. A map is drawn to pinpoint the exact feasible locations using Google Earth Pro and QGIS.

The final chapter consists of a map that is developed to find suitable land for the newly developing ‘*Regional Research Center for Renewable Energy*’ in Kandy. The suitable area in the map is developed using the criteria provided in the task and the total number of buildings within the suitable area at present, land area occupied by the buildings, and total suitable land area is found out.

Finally, the document is concluded with a conclusion of the report. All evidence related to the tasks are attached in the respective appendix.

CHAPTER ONE

R STUDIO ANALYTICS

This chapter consists of an analysis related to the energy efficiency in building cooling and its associations with various other factors to reduce residential and commercial electricity consumption. The goal is to analyze '*energy_efficiency_data.csv*' and provide critical discussions on how structural factors related to the shape of a building are affecting building cooling and have an impact in the energy generation.

Relative_Compactness	Surface_Area	Wall_Area	Roof_Area	Overall_Height	Orientation	Glazing_Area	Glazing_Area_Distribution	Heating_Load	Cooling_Load
0.98	514.5	294.0	110.25	7	2	0	0	15.55	21.33
0.98	514.5	294.0	110.25	7	3	0	0	15.55	21.33
0.98	514.5	294.0	110.25	7	4	0	0	15.55	21.33
0.98	514.5	294.0	110.25	7	5	0	0	15.55	21.33
0.90	563.5	318.5	122.50	7	2	0	0	20.84	28.28
0.90	563.5	318.5	122.50	7	3	0	0	21.46	25.38

Figure 1.1: Head of energy_efficient_data.csv

The above figure 1.1 depicts the head of the dataset that is to be analyzed. This data is extracted from a case study in the USA. The table consists of multiple variables such as *'Relative_Compactness'*, *'Surface_Area'*, *'Wall_Area'*, *'Roof_Area'*, *'Overall_Height'*, *'Orientation'*, *'Glazing_Area'*, *'Glazing_Area_Distribution'*, *'Heating_Load'*, and *'Cooling_Load.'* All these variables are crucial in cooling a building and can affect the efficiency of energy usage. By analyzing this dataset, the decision makers can get a good knowledge regarding how the structural factors related to the shape of a building are affecting energy generation.

Categorical	**Continuous**
Orientation	Relative Compactness
Glazing Area Distribution	Surface Area
	Wall Area
	Roof Area
	Overall Height
	Glazing Area
	Heating Load
	Cooling Load

Table 1.1: Categorical and continuous values of energy efficiency dataset

Building Cooling and Other Structural Factors related to the shape of the building.

1.1. Main Hypothesis Statements

H0: There is no association between building cooling and associated building structural factors related to the shape of the building.
H1: There is an association between building cooling and all associated building structural factors related to the shape of the building.
H2: There is an association between building cooling and associated building structural factors such as 'Relative Compactness', 'Surface Area', 'Wall Area', 'Roof Area', and 'Overall Height'.

Table 1.2: Main Hypothesis

Note: The heating load from the dataset won't be used to find relationships because it is not a structural factor of a building, and it does not affect the shape of the building.

- Significance level (x) = 0.05 (5%)
- Confidence level = 0.95 (95%)

1.2. Data Pre-Processing

```
#import the dataset
unclean.data <- read.csv("energy_efficiency_data.csv")

#data preprocessing
energy_data <- na.omit(unclean.data)
```

Figure 1.2: Importing data to RStudio.

Figure 1.2 depicts the code that is used to import and clean the dataset into RStudio.

1.3. Data Visualization

1.3.1. Dataset Size

```
> #data row count
> nrow(energy_data)
[1] 768
>
> #data column count
> ncol(energy_data)
[1] 10
```

Figure 1.3: Size of the dataset to be analyzed.

The above figure 1.3 shows the number are of columns and rows that are available in the dataset which are **10** and **768** respectively.

1.3.2. Data Visualization

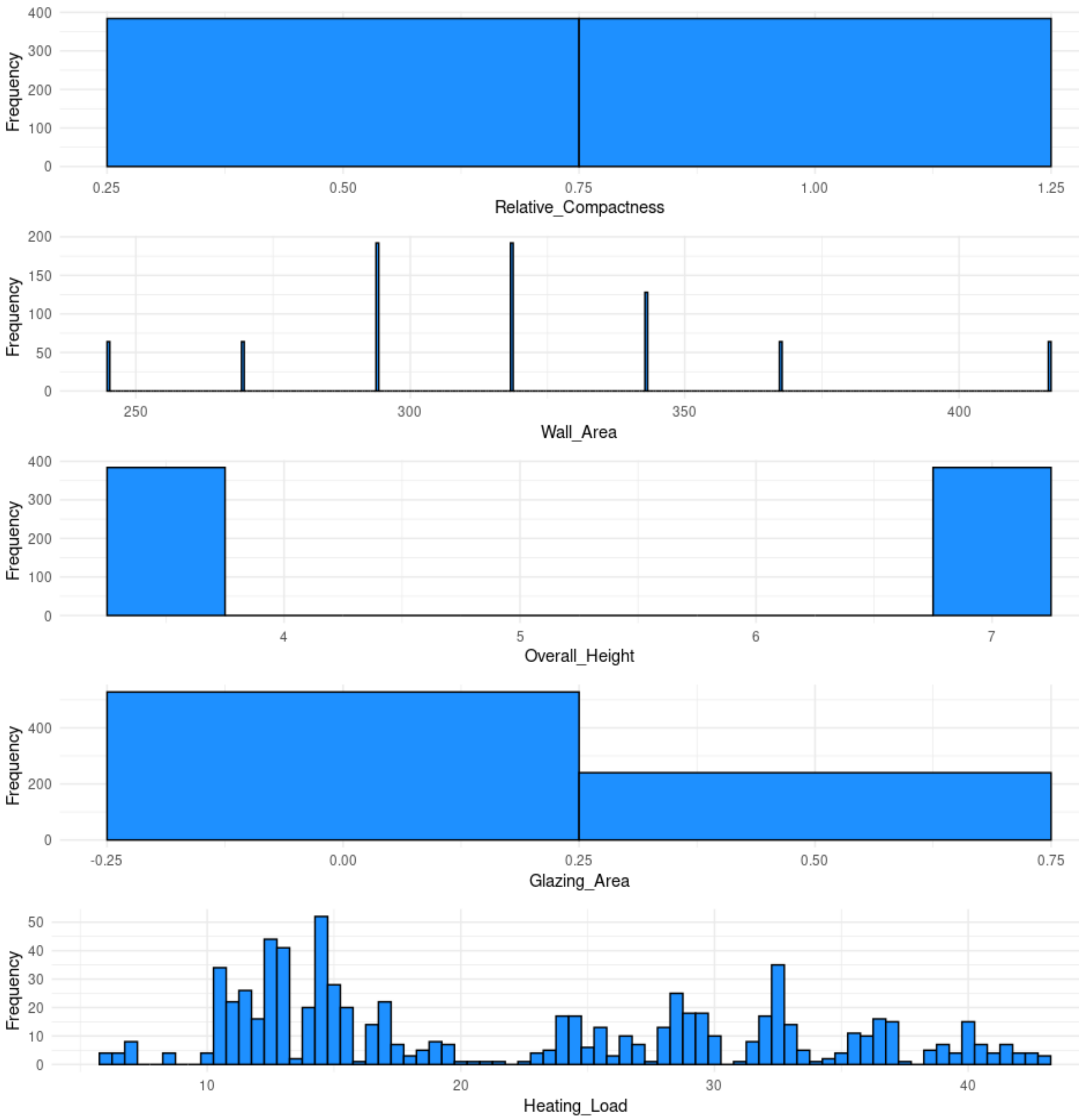


Figure 1.4: Energy Efficiency Data in Histogram - 1

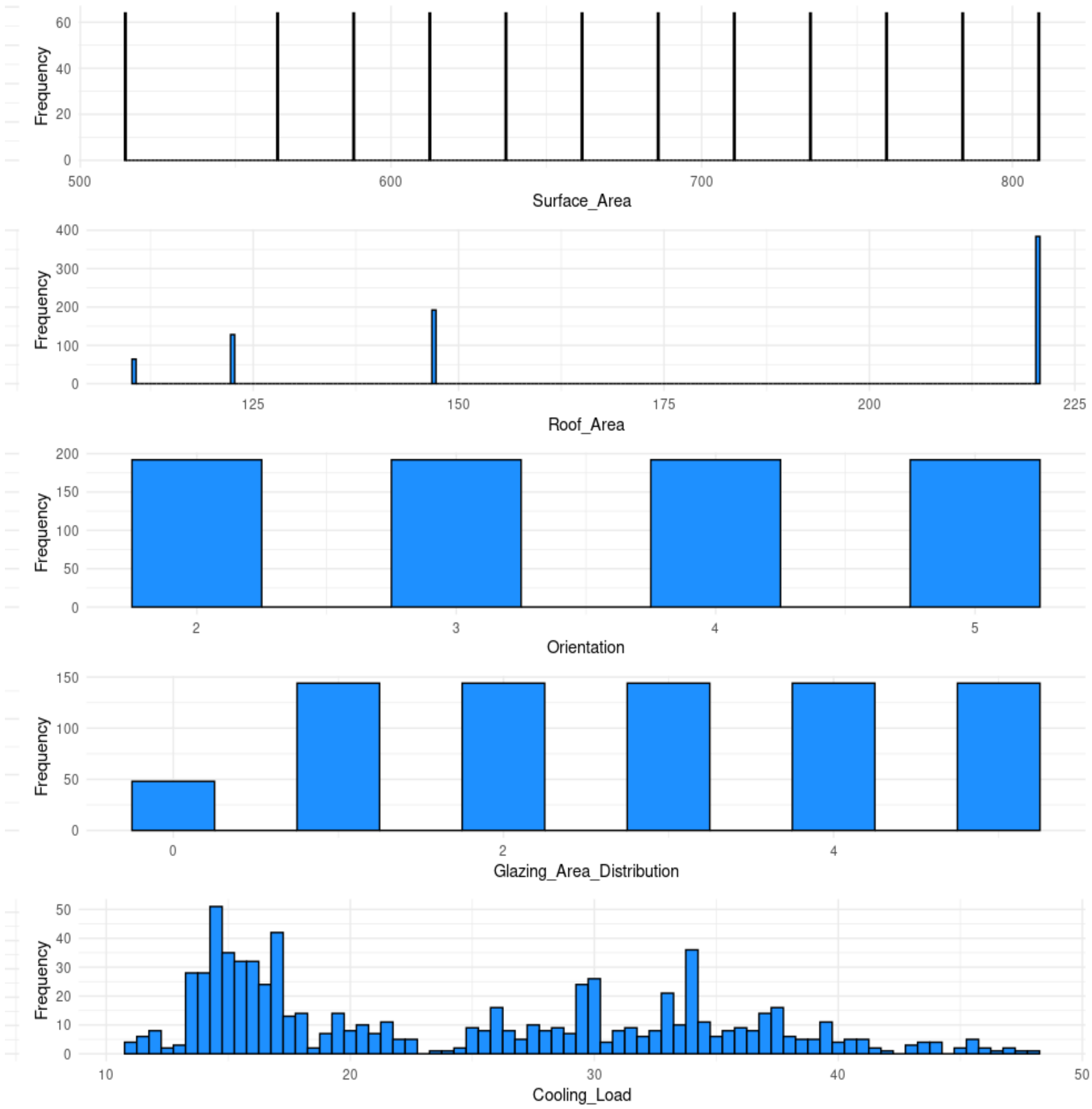


Figure 1.5: Energy Efficiency Data in Histogram - 2

As you can see in figures 1.4 and 1.5 the data are shown in histograms. When it comes to the Relative Compactness, Wall Area, Overall Height, Glazing Area, Surface Area, Roof Area, Orientation, and Glazing Area Distribution columns contain mostly repetitive values because there are few bars in their histograms. However, we can see that the cooling load data and heating load data have multiple values under them. The data can be further investigated by using the summary statistics function in R.

1.3.3. Summary statistics of the data

```
> summary(energy_data)
 Relative_Compactness  Surface_Area     Wall_Area       Roof_Area      Overall_Height
 Min.   :0.6200       Min.   :514.5   Min.   :245.0   Min.   :110.2   Min.   :3.50
 1st Qu.:0.6825       1st Qu.:606.4   1st Qu.:294.0   1st Qu.:140.9   1st Qu.:3.50
 Median :0.7500       Median :673.8   Median :318.5   Median :183.8   Median :5.25
 Mean   :0.7642       Mean   :671.7   Mean   :318.5   Mean   :176.6   Mean   :5.25
 3rd Qu.:0.8300       3rd Qu.:741.1   3rd Qu.:343.0   3rd Qu.:220.5   3rd Qu.:7.00
 Max.   :0.9800       Max.   :808.5   Max.   :416.5   Max.   :220.5   Max.   :7.00

  Orientation     Glazing_Area    Glazing_Area_Distribution  Heating_Load    Cooling_Load
 Min.   :2.00   Min.   :0.0000   Min.   :0.000             Min.   : 6.01   Min.   :10.90
 1st Qu.:2.75   1st Qu.:0.1000   1st Qu.:1.750             1st Qu.:12.99   1st Qu.:15.62
 Median :3.50   Median :0.2500   Median :3.000             Median :18.95   Median :22.08
 Mean   :3.50   Mean   :0.2344   Mean   :2.812             Mean   :22.31   Mean   :24.59
 3rd Qu.:4.25   3rd Qu.:0.4000   3rd Qu.:4.000             3rd Qu.:31.67   3rd Qu.:33.13
 Max.   :5.00   Max.   :0.4000   Max.   :5.000             Max.   :43.10   Max.   :48.03
```

Figure 1.6: Summary statistics of energy efficiency dataset

As depicted in figure 1.6, the summary function of R is used to get an idea of the energy efficiency dataset that will be analyzed for associations between the cooling load and structural factors.

1.4. Cooling Load and other factors

The cooling load factor will be compared with the other factors that could affect cooling by using scatter plots and their impact on the relationship can be seen.

1.4.1. Relative Compactness and the Cooling Load

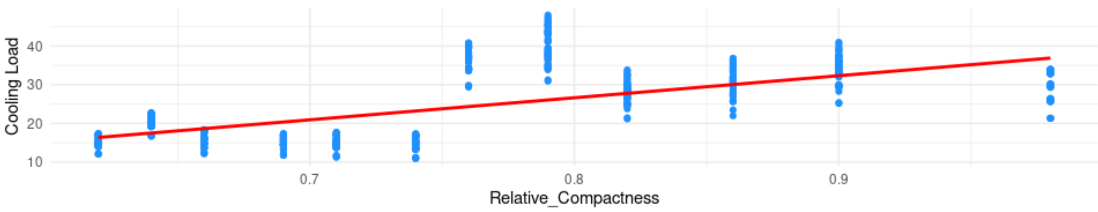


Figure 1.7: Scatter Plot for relative compactness with the cooling load

The above figure 7.1 depicts that there is a positively increasing relationship between relative compactness and the cooling load.

1.4.2. Surface Area and the Cooling Load

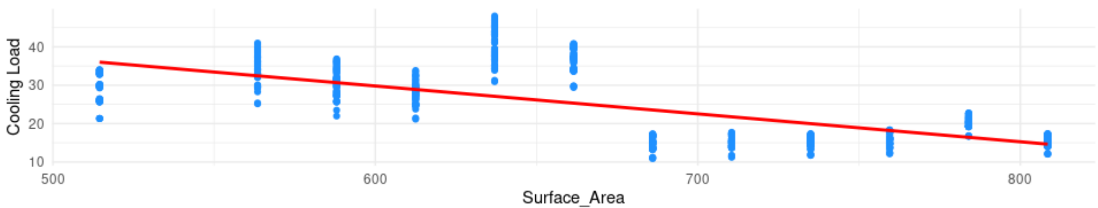


Figure 1.8: Scatter Plot for Surface Area with the cooling load

According to figure 1.8, it is known that when the surface area is increasing the cooling load decreases. There is a negatively decreasing relationship between surface area and cooling load.

1.4.3. Wall Area and the Cooling Load

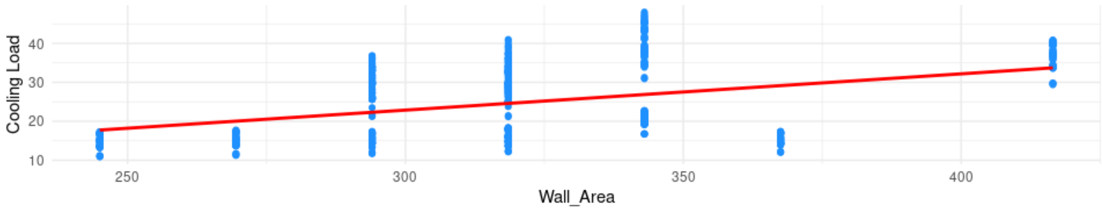


Figure 1.9: Scatter Plot for Wall Area with the cooling load

Figure 1.9 depicts that there is a positively increasing relationship between the cooling load and wall area.

1.4.4. Roof Area and the Cooling Load

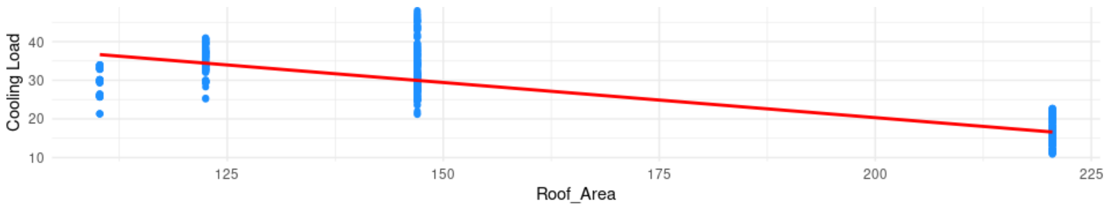


Figure 1.10: Scatter Plot for Roof Area with the cooling load

When the roof area increases the cooling load decreases, resulting in a negatively decreasing relationship according to figure 1.10.

1.4.5. Overall Height and the Cooling Load

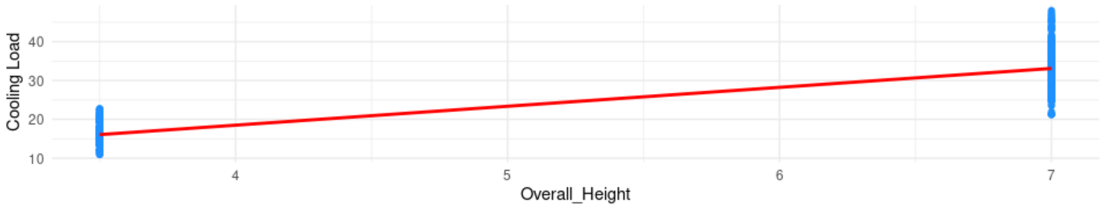


Figure 1.11: Scatter Plot for Overall Height with the cooling load

There is a positive relationship between the overall height and the cooling load of a building according to figure 1.11.

1.4.6. Glazing Area and the Cooling Load

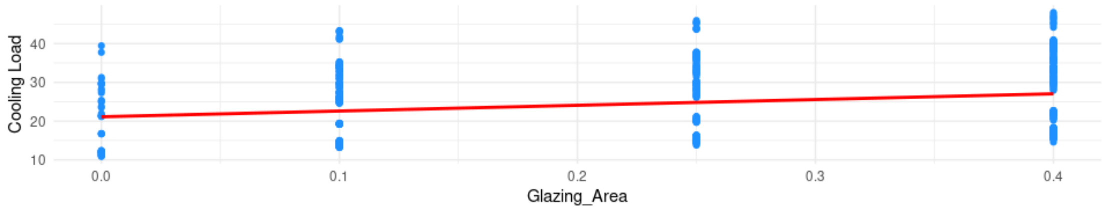


Figure 1.12: Scatter Plot for Glazing Area with the cooling load

According to figure 1.12 there is no significant positive relationship between the Glazing area and the cooling load, instead it shows a slightly positive increase. To know more about the relationships and associations correlation tests are conducted.

1.5. Correlation Tests

The correlation tests will be conducted to find the associations between the cooling load and structural factors affecting cooling load in a building.

1.5.1. Correlation Heatmap

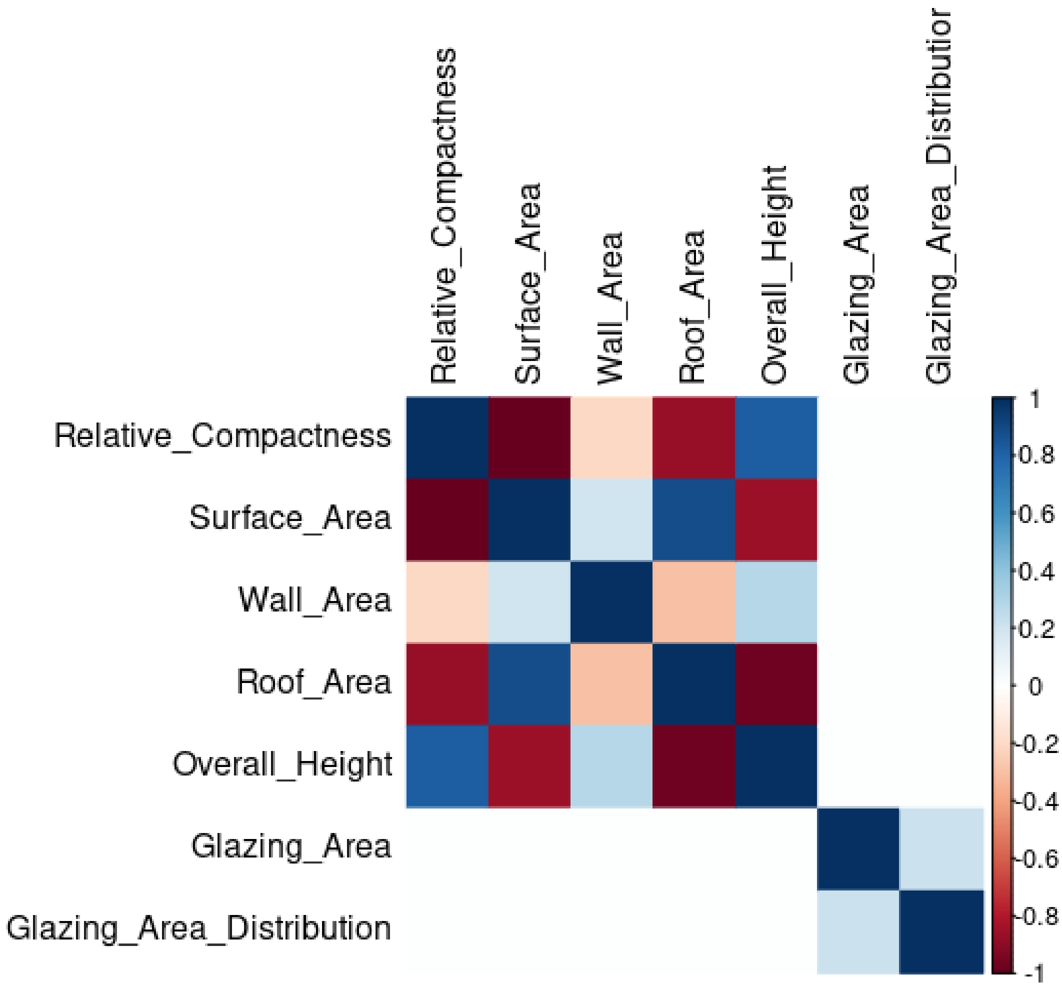


Figure 1.13: Correlational Heatmap

Figure 1.13 depicts the correlation matrix that is developed to uncover the associations between cooling load and other structural factors. According to the graph drawn Relative Compactness, Surface Area, Wall Area, Roof Area, and Overall Height have correlations with the cooling load. Refer to figure 1.14 below for detailed statistics on the correlational heatmap. The Glazing area and glazing area are not considered because there is no correlation to the cooling load.

	Relative_Compactness	Surface_Area	Wall_Area	Roof_Area	Overall_Height	Glazing_Area	Glazing_Area_Distribution
Relative_Compactness	1.000000e+00	-9.919015e-01	-0.2037817	-8.688234e-01	0.8277473	7.617400e-20	0.0000000
Surface_Area	-9.919015e-01	1.000000e+00	0.1955016	8.807195e-01	-0.8581477	4.664140e-20	0.0000000
Wall_Area	-2.037817e-01	1.955016e-01	1.0000000	-2.923165e-01	0.2809757	0.000000e+00	0.0000000
Roof_Area	-8.688234e-01	8.807195e-01	-0.2923165	1.000000e+00	-0.9725122	-1.197187e-19	0.0000000
Overall_Height	8.277473e-01	-8.581477e-01	0.2809757	-9.725122e-01	1.0000000	0.000000e+00	0.0000000
Glazing_Area	7.617400e-20	4.664140e-20	0.0000000	-1.197187e-19	0.0000000	1.000000e+00	0.2129642
Glazing_Area_Distribution	0.000000e+00	0.000000e+00	0.0000000	0.000000e+00	0.0000000	2.129642e-01	1.0000000

Figure 1.14: Correlation Matrix

1.5.2. Hypothesis Based Correlation Testing

This phase is to find out correlation between the structural factors and the cooling load.

1.5.2.1. Relative Compactness

H0: There is no significant relationship between cooling load and relative compactness
H1: There is a significant relationship between cooling load and relative compactness

Table 1.3: Hypothesis Statement One

```
[1] "Correlation Test between Cooling_Load and Relative_Compactness"

        Pearson's product-moment correlation

data:  energy_data$Cooling_Load and energy_data[[factor]]
t = 22.71, df = 766, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.5900749 0.6748011
sample estimates:
      cor
0.6343391
```

Figure 1.15: Pearson Cors test between Cooling Load and Relative Compactness

The Pearson correlation test between cooling load and relative compactness indicates that there is a statistically significant and moderately strong positive linear relationship based on the correlation coefficient '0.6343391' and a very small p-value rejecting the null hypothesis.

Decision: p-value = 2.2e-16/cor=0.6343391 > Reject H0

1.5.2.2. Surface Area

H0: There is no significant relationship between cooling load and surface area
H1: There is a significant relationship between cooling load and surface area

Table 1.4: Hypothesis Statement Two

```
[1] "Correlation Test between Cooling_Load and Surface_Area"

	Pearson's product-moment correlation

data:  energy_data$Cooling_Load and energy_data[[factor]]
t = -25.183, df = 766, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.7099423 -0.6323619
sample estimates:
       cor 
-0.6729989 
```

Figure 1.16: Pearson Cors test between Cooling Load and Surface Area

Figure 1.16 depicts that there is a strong and significant negative linear relationship between the cooling load and surface area due to the correlation coefficient of '-0.6729989' and a very small p-value of 2.2e-16, thus rejecting the null hypothesis.

<u>**Decision:** p-value = 2.2e-16/cor=-0.6729989> Reject H0</u>

1.5.2.3. Wall Area

H0: There is no significant relationship between cooling load and wall area
H1: There is a significant relationship between cooling load and wall area

Table 1.5: Hypothesis Statement Three

```
[1] "Correlation Test between Cooling_Load and Wall_Area"

	Pearson's product-moment correlation

data:  energy_data$Cooling_Load and energy_data[[factor]]
t = 13.074, df = 766, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.3674764 0.4832591
sample estimates:
     cor 
0.427117 
```

Figure 1.17: Pearson Cors test between Cooling Load and Wall Area

The output shows that there is a statistically significant and moderate positive linear relationship between the cooling load and wall area based on the coefficient of '0.427117' and the p-value of 2.2e-16. Rejecting the null hypothesis.

<u>**Decision:** p-value = 2.2e-16/cor =0.427117> Reject H0</u>

1.5.2.4. Roof Area

H0: There is no significant relationship between cooling load and roof area
H1: There is a significant relationship between cooling load and roof area

Table 1.6: Hypothesis Statement Four

```
[1] "Correlation Test between Cooling_Load and Roof_Area"

	Pearson's product-moment correlation

data:  energy_data$Cooling_Load and energy_data[[factor]]
t = -47.181, df = 766, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.8796165 -0.8432581
sample estimates:
       cor
-0.8625466
```

Figure 1.18: Pearson Cors test between Cooling Load and Roof Area

The above output depicts that there is a significant negative linear relationship between the cooling load and roof area based on the correlation coefficient of '-0.8625466' and a minimal p-value of 2.2e-16.

Decision: p-value = 2.2e-16/cor=-0.8625466> Reject H0

1.5.2.5. Overall Height

H0: There is no significant relationship between cooling load and overall height
H1: There is a significant relationship between cooling load and overall height

Table 1.7: Hypothesis Statement Five

```
[1] "Correlation Test between Cooling_Load and Overall_Height"

	Pearson's product-moment correlation

data:  energy_data$Cooling_Load and energy_data[[factor]]
t = 55.777, df = 766, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.8808626 0.9089291
sample estimates:
      cor
0.8957852
```

Figure 1.19: Pearson Cors test between Cooling Load and Overall Height

According to the output there is a significant and very strong positive linear relationship between cooling load and overall height based on the correlation coefficient '0.8957852' and a very small p-value of 2.2e16.

Decision: p-value = 2.2e-16/cor=0.8957852> Reject H0

1.6. Chapter Review

As per the analysis done previously it is visible that there is an association between the cooling load and Structural factors of a building such as the wall area, roof area, surface area, overall height, and the relative compactness. Therefore Main H0 is rejected and H1 can be rejected too because all structural factors do not have association between the cooling loads.

Decision: Reject Main Hypothesis' H0 and H1 and H2 is True

Therefore when a building is constructed 'Relative Compactness', 'Surface Area', 'Wall Area', 'Roof-Area', and 'Overall Height' must be considered. Because these factors mostly affect the cooling load of a building which in turn affects the energy consumption of a building. This way we can reduce the high energy consumption that takes place in Sri Lanka.

ALL CODES RELATED TO THIS CHAPTER ARE IN APPENDIX A

CHAPTER TWO

THE STATISTICAL MODEL

The following chapter is about creating a precise statistical model by using the findings from the previous chapter.

2.1. Regression Analysis

According to the hypothesis-based correlation tests done in the previous chapter it is believed that the factors associated with the cooling load are the relative compactness, surface area, wall area, roof area and overall height. Therefore by using these factors the regression analysis will be conducted.

2.1.1. Summary statistics for the linear regression model

```
Call:
lm(formula = Cooling_Load ~ Relative_Compactness + Surface_Area +
    Wall_Area + Roof_Area + Overall_Height, data = energy_data)

Residuals:
     Min       1Q   Median       3Q      Max
-12.1970  -2.2912  -0.2118   2.1972  13.6166

Coefficients: (1 not defined because of singularities)
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)          101.234810  24.391929   4.150 3.69e-05 ***
Relative_Compactness -70.787707  13.189195  -5.367 1.06e-07 ***
Surface_Area          -0.088245   0.021888  -4.032 6.09e-05 ***
Wall_Area              0.044682   0.008521   5.243 2.04e-07 ***
Roof_Area                    NA         NA      NA       NA
Overall_Height         4.283843   0.433242   9.888  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.761 on 763 degrees of freedom
Multiple R-squared:  0.8445,    Adjusted R-squared:  0.8437
F-statistic:  1036 on 4 and 763 DF,  p-value: < 2.2e-16
```

Figure 2.1: Summary statistics of linear regression function

According to figure 2.1 when the structural factors were fit into a linear regression model with the cooling load it seems to be significant with low p-values for the F-Statistics, and the predictor variables except for the roof area variable. Therefore the model can be created without the roof area variable. Also the Adjusted R-squared of '0.8437' suggests that the model has a substantial amount of variance in the cooling load.

Decision: Remove roof area from the model

```
Call:
lm(formula = Cooling_Load ~ Relative_Compactness + Surface_Area +
    Wall_Area + Overall_Height, data = energy_data)

Residuals:
     Min       1Q   Median       3Q      Max
-12.1970  -2.2912  -0.2118   2.1972  13.6166

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)
(Intercept)          101.234810  24.391929   4.150 3.69e-05 ***
Relative_Compactness -70.787707  13.189195  -5.367 1.06e-07 ***
Surface_Area          -0.088245   0.021888  -4.032 6.09e-05 ***
Wall_Area              0.044682   0.008521   5.243 2.04e-07 ***
Overall_Height         4.283843   0.433242   9.888  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.761 on 763 degrees of freedom
Multiple R-squared:  0.8445,	Adjusted R-squared:  0.8437
F-statistic:  1036 on 4 and 763 DF,  p-value: < 2.2e-16
```

Figure 2.2: Updated summary statistics of linear regression function

The above figure 2.2 depicts the summary statistics for the linear regression model after removing the roof area from the equation. All other values do not appear to be varied.

2.1.2. Linear regression model

```
#linear regression model
rep.statistical.model <- lm(Cooling_Load ~ Relative_Compactness + Surface_Area +
                              Wall_Area + Overall_Height , data = energy_data)
```

Figure 2.4: Created linear regression function.

```
Call:
lm(formula = Cooling_Load ~ Relative_Compactness + Surface_Area +
    Wall_Area + Overall_Height, data = energy_data)

Coefficients:
         (Intercept)  Relative_Compactness          Surface_Area             Wall_Area        Overall_Height
           101.23481             -70.78771              -0.08824               0.04468               4.28384
```

Figure 2.3: Printed linear regression model.

The coefficients in figure 2.3 provide insight into how each predictor variable contributes to the prediction of the cooling load in the linear regression model. They provide the direction (positive/negative) in how each variable affects the dependent variable along with the magnitude of the impact. The below table provides more insight to figure 2.3.

Variable	Coefficient	Interpretation
(Intercept)	101.23481	Estimated Cooling_Load when all predictor variables are zero.
Relative Compactness	-70.78771	Estimated change in Cooling_Load for a one-unit increase in Relative_Compactness, holding all other variables constant. A one-unit increase in Relative_Compactness is associated with a decrease of approximately 70.79 units in Cooling_Load.
Surface Area	-0.08824	Estimated change in Cooling_Load for a one-unit increase in Surface_Area, while keeping all other variables constant. A one-unit increase in Surface_Area is associated with a decrease of approximately 0.0882 units in Cooling_Load.
Wall Area	0.04468	Estimated change in Cooling_Load for a one-unit increase in Wall_Area, with all other variables held constant. A one-unit increase in Wall_Area is associated with an increase of approximately 0.0447 units in Cooling_Load.
Overall Height	4.28384	Estimated change in Cooling_Load for a one-unit increase in Overall_Height, while holding all other variables constant. A one-unit increase in Overall_Height is associated with an increase of approximately 4.28 units in Cooling_Load.

Table 2.1: Interpretation of Figure 2.3

2.1.3. Shapiro-Wilk Test

```
        Shapiro-Wilk normality test

data:  resid(rep.statistical.model)
W = 0.9775, p-value = 1.727e-09
```

Figure 2.5: Shapiro normality testing

According to the values obtained from the Shapiro test, it is known that the model is not normally distributed since the p-value is much smaller than the common significant level of 0.05. However the model can still be used in predictions and statistical models with some limitations.

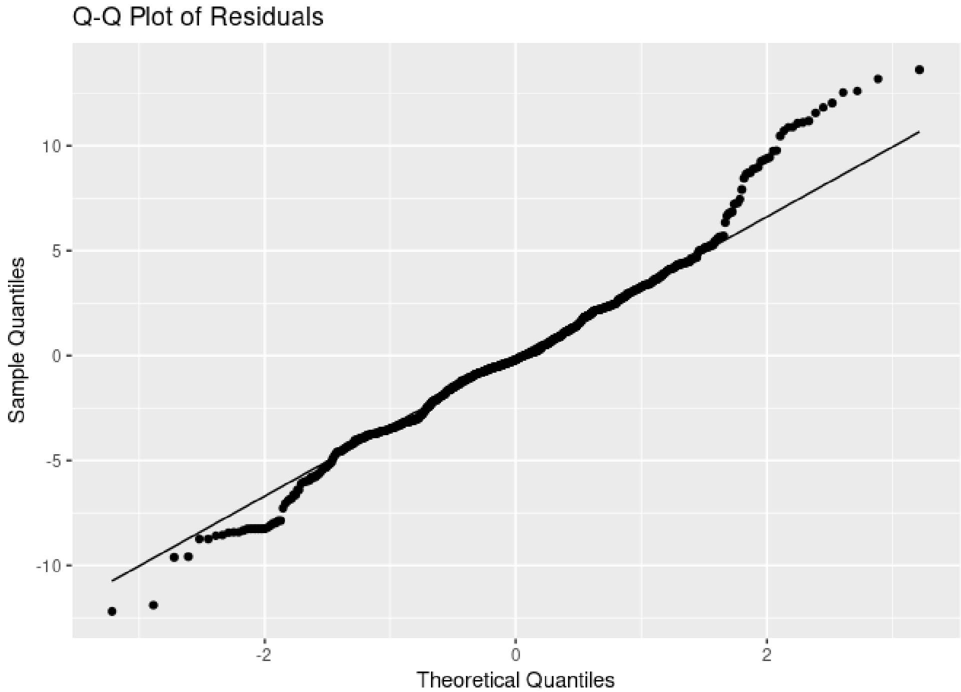


Figure 2.6: The model

2.2. Statistical Model

$$Y = X0 + X1 * \text{Relative Compactness} + X2 * \text{ Surface Area} + X3 * \text{ Wall Area} + X4 * \text{ Overall\_Height} + E$$

Equation 2.1: Statistical Model

Variable	Value	Description
Y	Cooling Load	Dependent Variable
X0	Intercept	The constant term
X1	Relative Compactness	Predictor Variable
X2	Surface Area	Predictor Variable
X3	Wall Area	Predictor Variable
X4	Overall Height	Predictor Variable
E	Error	Variability in the dependent variable that is not explained by the predictors

Table 2.2: Model Explanation

The goal of the above statistical model in equation 2.1 is to understand how changes to the predictor variables are associated with the changes in the cooling load. During the regression analysis the values needed for the variables are extracted from the lm () function.

$$Y = 101.23 - 70.79 * Relative\_Compactness - 0.0882 * Surface\_Area + 0.0447 * Wall\_Area + 4.28 * Overall\_Height$$

Equation 2.2: Actual Equation

The equation 2.2 depicts the equation with the values obtained from the linear regression analysis done previously. By using this equation it is possible to predict the Cooling Load of a building using the predictor variables.

This model can be used in Sri Lanka even before staring to construct the building and predict the cooling load. Where in turn we can use that to calculate the energy consumption that is required to cool down the building.

ALL CODES RELATED TO THIS CHAPTER ARE IN <u>APPENDIX B</u>

CHAPTER THREE

SRI LANKA WIND POWER DEVELOPMENT PLAN

The following chapter consists of maps that are developed to showcase the lands that are available to develop wind power plants in each district. The data for the maps are obtained from the Renewable Energy Resource Development Plan of Sri Lanka.

3.1. Land Available for wind power development data

District	Available Land extent (ha)									
	Scrub lands	Barren lands	Open forest	Sand	Homestead / Gardens	Sparsely used crop lands	Grass lands	Palmyrah	Coconut	Total
	SCRBA	BRRNA	FRSOA	SANDA	HOMSA	SPRSA	GRSLA	PLMRA	CCNTA	
Ampara	40,531	22,550	103,705	1,285	-	-	-	-	-	**168,071**
Anuradhapura	117,222	8,154	70,543	-	961	1,143	-	-	3	**198,026**
Badulla	126,647	3,742	127,656	-	-	-	-	-	-	**258,045**
Batticaloa	58,251	16,856	24,325	892	-	-	-	-	-	**100,324**
Colombo	1,066	162	861	191	-	-	-	-	-	**2,279**
Galle	6,454	181	5,941	252	-	-	-	-	-	**12,828**
Gampaha	5,121	746	1,570	375	-	-	-	-	-	**7,811**
Hambantota	53,217	2,222	37,456	1,555	-	-	-	-	-	**94,450**
Jaffna	9,928	7,750	1,096	5,920	34,778	3,313	2,449	701	874	**66,808**
Kalutara	16,468	151	14,080	239	-	-	-	-	-	**30,938**
Kandy	22,448	1,122	87,944	-	-	-	-	-	-	**111,515**
Kegalle	12,522	469	14,235	-	-	-	-	-	-	**27,226**
Kilinochchi	13,960	7,806	8,508	2,530	17,516	5,742	21	469	1,705	**58,256**
Kurunegala	38,396	1,027	20,206	1	-	-	-	-	-	**59,630**
Mannar	32,532	9,391	17,940	465	8,715	4,523	227	737	849	**75,379**
Matale	31,047	2,930	112,991	-	-	-	-	-	-	**146,968**
Matara	2,174	108	1,660	137	-	-	-	-	-	**4,078**
Moneragala	202,953	1,844	145,499	-	-	-	-	-	-	**350,296**
Mullaitivu	27,621	5,902	29,933	642	16,716	6,505	402	-	1,029	**88,749**
Nuwara Eliya	33,582	555	52,807	-	-	-	-	-	-	**86,944**
Polonnaruwa	60,743	10,996	57,998	-	-	-	-	-	-	**129,737**
Puttalam	42,660	5,650	18,174	2,332	-	-	-	-	-	**68,817**
Ratnapura	62,244	1,883	67,576	-	-	-	-	-	-	**131,703**
Trincomalee	46,942	18,139	62,628	563	-	4	-	-	-	**128,275**
Vavuniya	23,978	2,427	13,487	-	21,268	12,934	1,044	-	108	**75,247**
Total	**1,088,708**	**132,762**	**1,098,817**	**17,378**	**99,955**	**34,164**	**4,142**	**1,907**	**4,568**	**2,482,401**

Figure 3.1: Land available for wind power plant development (Sustainable Energy Authority, 2023)

The above figure 3.1 depicts the land that are available to develop wind power plants in Sri Lanka according to the Sustainable Energy Authority in Sri Lanka. The data is obtained and then saved as '*SLWindPowerLand_21_23.csv*' to be used in the mapping.

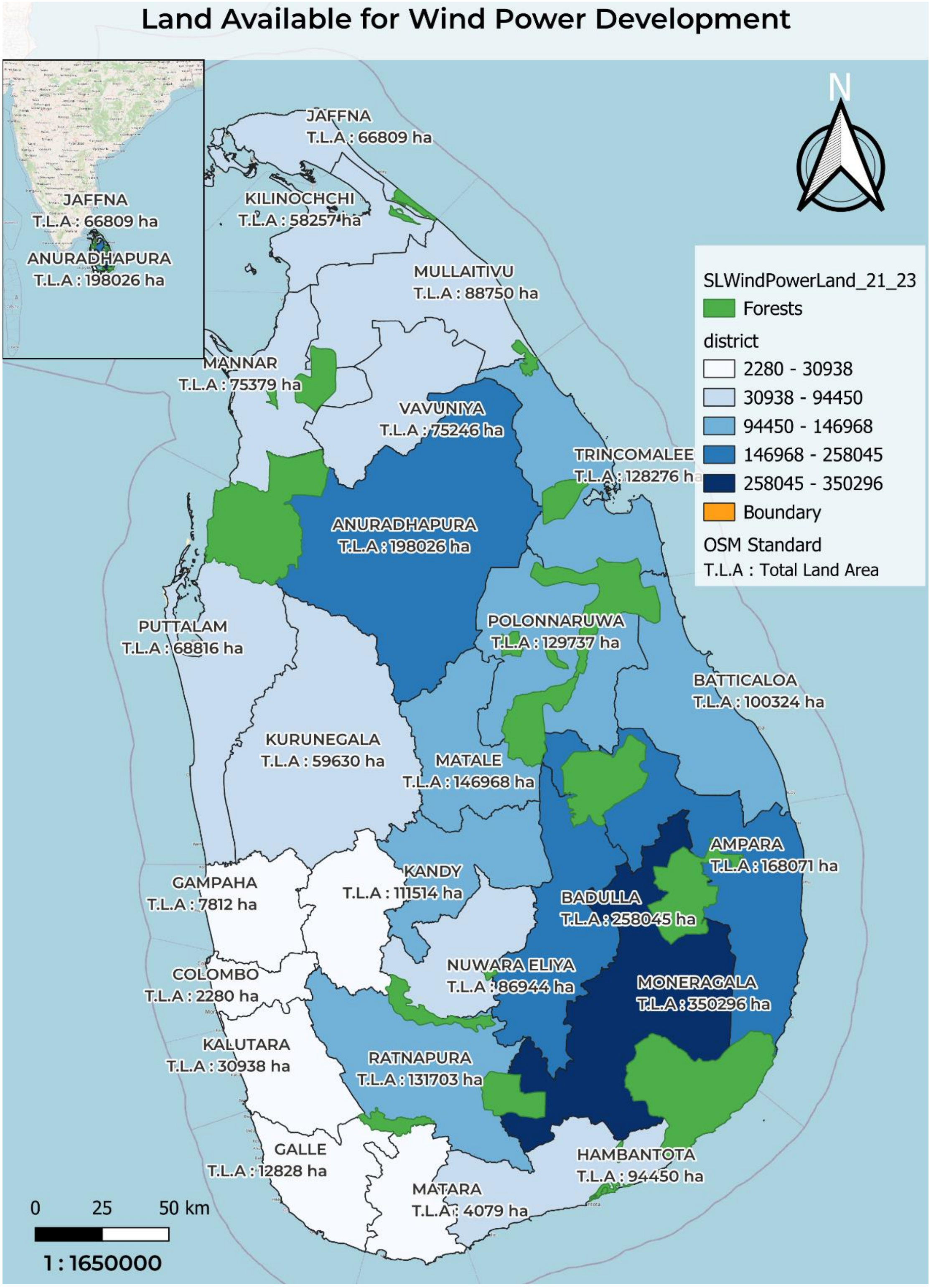


Figure 3.2: Total Land Available in each district for Wind Power development

The thematic map in figure 3.2 illustrates the availability of land for wind power development across the districts of Sri Lanka. The data is obtained from the Renewable Energy Development Plan 2021-2026 provided by the Sri Lanka Sustainable Energy Authority.

The district is color-coded to represent the availability of land by using darker colors for indicating larger available land areas. Districts with significant land availability for wind power development include Monaragala, Badulla, Anuradhapura, and Ampara. Least number of land available is represented by the Western Province and Galle district. The most promising district for development is Anuradhapura when considering the population and transmission of electricity.

This map serves as a valuable tool for investors and decision makers in the renewable energy sector providing a clear visual representation forecasting potential land for wind power development.

ALL STEPS RELATED TO THIS CHAPTER ARE IN <u>APPENDIX C</u>

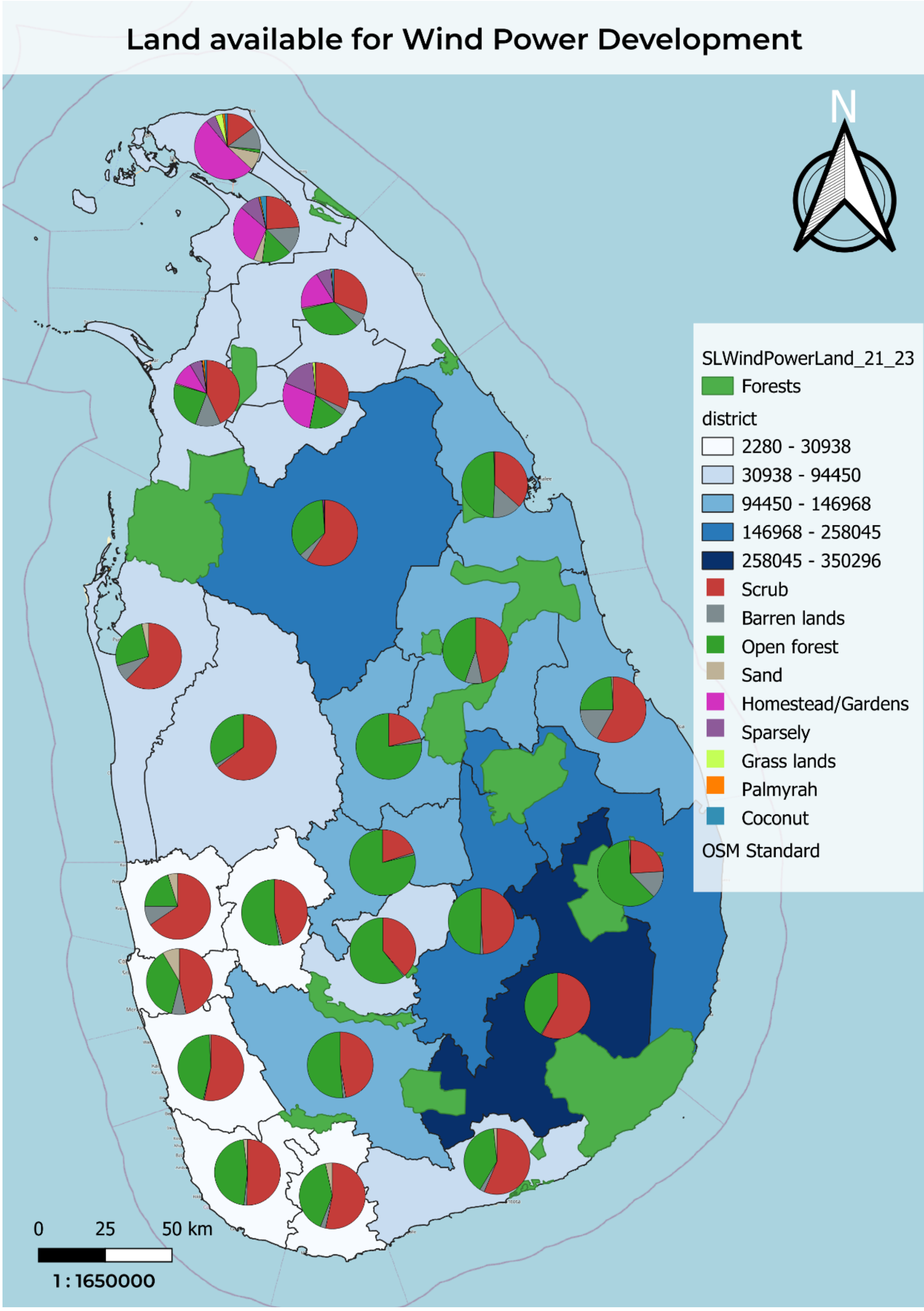


Figure 3.3: Types of land available in each district

The map in figure 3.3 shows pie charts above the districts to show the type of land that is available to develop wind power plants.

CHAPTER FOUR

POTENTIAL LAND FOR SOLAR POWER DEVELOPMENT

The following chapter consists of the thematic map that has been created to showcase the potential land for solar power development by using the data obtained from the Sri Lanka sustainable energy authority. Along with an analysis of how solar power energy would contribute for overcoming energy sector problems.

4.1. Land available for solar power development data

District	Total Area (ha)	Total Capacity (MW)	Site-wise capacity			Total estimated energy (GWh)
			10MW<x<25MW	25MW<x<100MW	>100MW	
Ampara	6,367	3,183	536	835	1,812	4,715
Anuradhapura	616	308	35	273	-	467
Badulla	10,103	5,052	831	2,288	1,933	7,522
Batticaloa	3,136	1,568	275	986	307	2,312
Colombo	26	13	13	-	-	19
Galle	302	151	35	116	-	226
Gampaha	415	207	117	90	-	295
Hambantota	1,976	988	121	388	479	1,436
Jaffna	962	481	39	144	297	738
Kalutara	1,160	580	121	176	283	852
Kandy	360	180	74	106	-	258
Kegalle	336	168	128	40	-	240
Kilinochchi	2,099	1,049	307	284	459	1,586
Kurunegala	3,452	1,726	476	826	424	2,525
Mannar	3,612	1,806	394	586	826	2,776
Matale	2,732	1,366	206	438	722	1,970
Matara	332	166	29	137	-	247
Moneragala	2,531	1,266	136	584	546	1,850
Mullaitivu	7,417	3,708	719	1,944	1,046	5,737
Nuwara Eliya	495	247	50	198	-	375
Polonnaruwa	982	491	124	367	-	746
Puttalam	2,520	1,260	274	653	333	1,879
Ratnapura	3,536	1,768	572	1,196	-	2,568
Trincomalee	1,957	979	164	264	551	1,464
Vavuniya	1,855	927	252	470	205	1,433
	59,278	**29,639**	**6,029**	**13,389**	**10,221**	**44,239**

Figure 4.1:Land available for solar power plant development (Sustainable Energy Authority, 2023)

The data in figure 4.1 shows the land available for solar power development along with the potential capacity and the estimated energy outputs from the plants in terms of each district (Sustainable Energy Authority, 2023).

4.2. Solar Power Potential and Land Usage in Sri Lanka

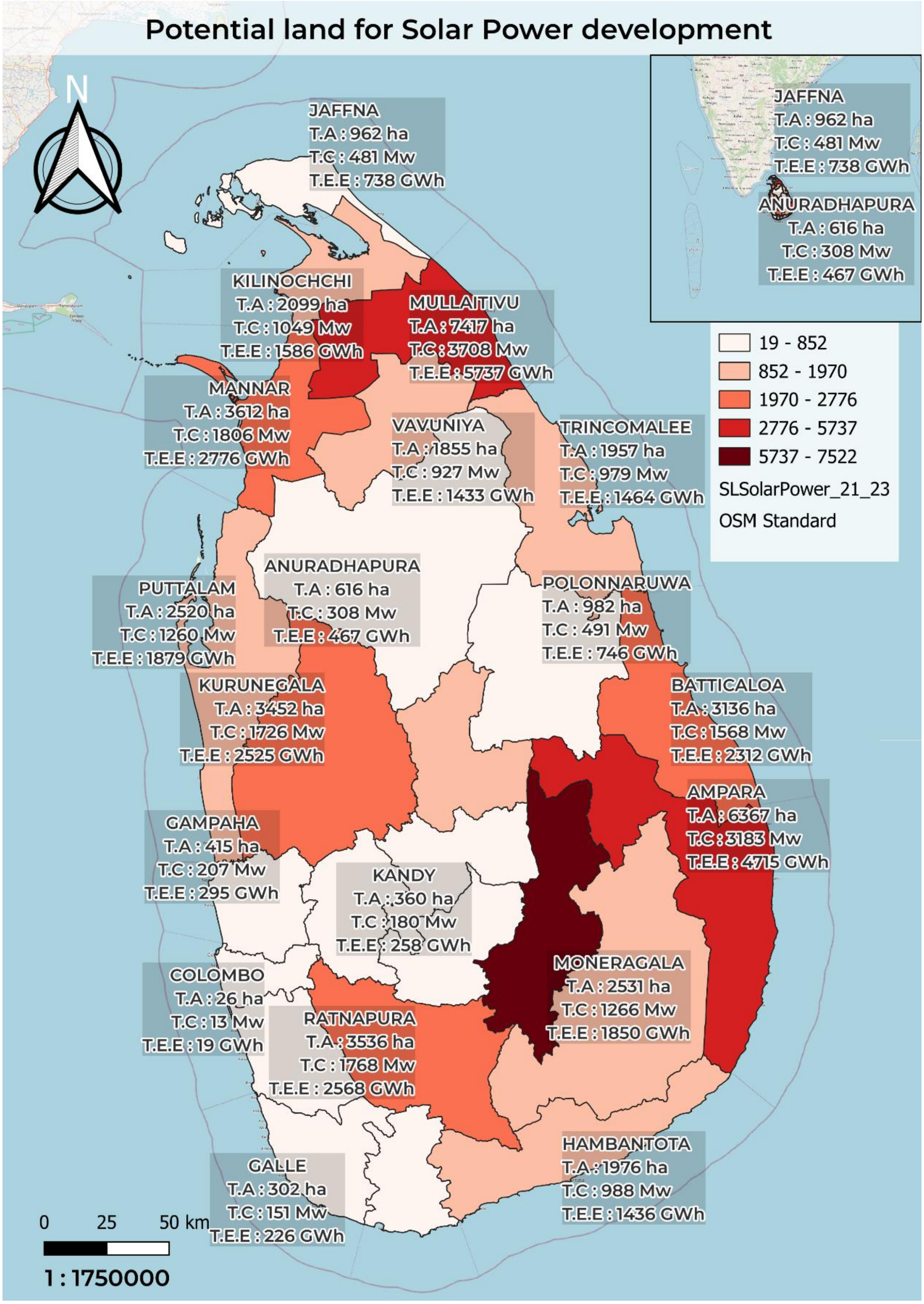


Figure 4.2: Potential land for Solar Power development

The figure 4.2 consists of a map of Sri Lanka with a visual representation of the solar power potential and land usage across the different districts in the country. The above map utilizes a thematic approach to highlight the district with the highest land available for solar power plants development.

The darker the color the higher the land available for solar power development. According to the map Badulla, Ampara and Mullaitivu stand out with their deep shades of color, indicating higher solar power potential. Urban and densely populated districts such as Colombo, Kandy and Gampaha have relatively low potential in solar power development.

This map utilizes the importance of sustainable energy development by highlighting areas where solar energy can be properly harnessed without extensive wastages and preserving nature.

ALL STEPS RELATED TO THIS CHAPTER ARE IN <u>APPENDIX D</u>

CHAPTER FIVE

DIGITIZED INFORMATIVE AREA MAP

The following chapter consists of the digitized informative area map of a place in Hambantota. The solar power plants, forests, trees, and other suburbs are digitized and saved as shape files in this chapter.

5.1. Pre digitalized map

Figure 5.1: Pre digitalized map

Figure 5.1 shows the map that should be digitized by creating vector layers/shape files.

5.2. Digitized area map

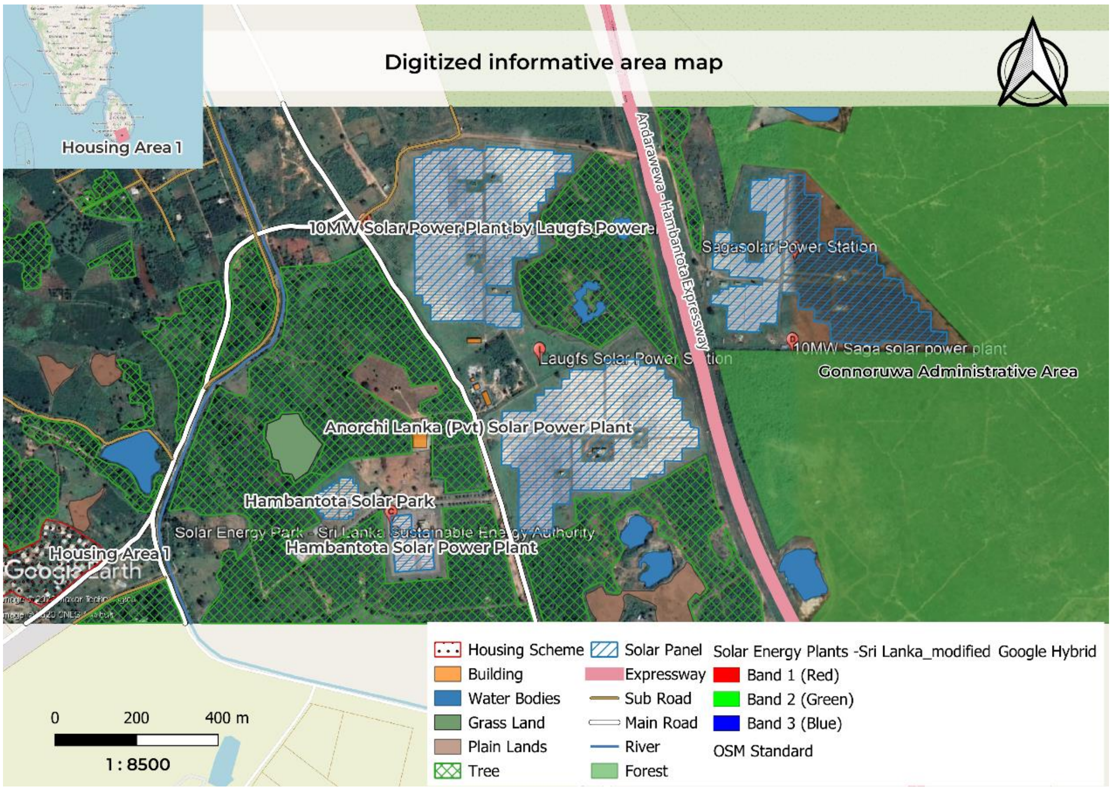


Figure 5.2: Digitized Area map

The digitalized map in figure 5.2 is captured from an area in the Hambantota district near the Gonnoruwa forest area. This area is mostly concentrated with solar power plants and forest areas. There are various water bodies and plain areas with trees in the above digitized area map.

5.3. Overcoming energy sector problems using solar power

Sri Lanka is a country with free sunshine the whole year. We can use it for the benefit of the country. For instance the small piece of land area depicted in figure 5.2 generates around 20Mw of electricity. Therefore by utilizing the land in Sri Lanka we can surpass our goal and generate surplus electricity. This could result in environmental benefits as well as reducing fuel imports in the country and can meet with the growing energy demand in the country.

ALL STEPS RELATED TO THIS CHAPTER ARE IN <u>APPENDIX E</u>

CHAPTER SIX

SRI LANKA PETROLEUM DISTRIBUTION

The following chapter consists of geo spatial analysis done to visualize a thematic map classified by using the data provided by the Sri Lanka Sustainable Energy Authority. A database using PostgreSQL named “SLPetrolium-2023” is created and the data are visualized in the map.

District	No of Sheds	Gasoline			Diesel		
		SD	Mean	CV_%	SD	Mean	CV_%
Ampara	54	117.5	1,806	6.5	1,079.7	3,015.7	35.8
Anuradhapura	47	182.7	3,360	5.4	2,199.0	6,160.8	35.7
Badulla	31	114.5	1,808	6.3	1,343.0	3,495.2	38.4
Batticaloa	39	114.3	1,323	8.6	660.0	1,756.2	37.6
Colombo	151	795.9	18,869	4.2	11,438.1	31,976.7	35.8
Galle	66	185.4	3,813	4.9	1,317.2	5,317.3	24.8
Gampaha	156	527.9	12,217	4.3	4,745.9	18,082.7	26.2
Hambantota	37	112.1	1,896	5.9	1689.8	4,082.9	41.4
Jaffna	59	96.5	1,634	5.9	825.7	2,671.1	30.9
Kalutara	56	216.7	4,471	4.8	1,421.8	6,150.7	23.1
Kandy	66	180.0	4,351	4.1	2,236.2	7,147.4	31.3
Kegalle	31	133.7	2,258	5.9	921.5	3,378.4	27.3
Kilinochchi	8	20.3	309	6.5	**573.4**	**1,031.8**	**55.6**
Kurunagela	111	302.6	6,892	4.4	2,683.5	9,992.4	26.9
Mannar	11	22.7	252	9.0	317.3	643.5	49.3
Matale	46	80.5	1,578	5.1	930.2	2,668.1	34.9
Matara	29	169.6	2,382	7.1	2,130.8	5,016.6	42.5
Monaragala	25	106.2	1,284	8.3	903.0	2,319.6	38.9
Mulativu	12	25.3	326	7.8	**688.1**	**1,206.4**	**57.0**
Nuwaraeliya	23	70.6	1,010	7.0	1,174.7	2,601.5	45.2
Polonnaruwa	**28**	**403.6**	**1,369**	**29.5**	1,259.5	2,931.2	43.0
Puttalam	72	122.1	3,023	4.0	1,617.0	5,124.4	31.6
Ratnapura	45	169.1	3,201	5.3	1,706.4	5,256.5	32.5
Trincomalee	**26**	**417.8**	**663**	**63.0**	1,169.2	2,432.4	48.1
Vavuniya	17	36.6	499	7.3	**1,156.2**	**2,060.1**	**56.1**

Figure 6.1: SL Petroleum Data

The table in figure 1.6 shows the data provided by the authority. The table comprises of columns named District, No. of Sheds, mean and standard deviation of gasoline and diesel distribution in the country. The map is developed to forecast these details.

6.1. Fuel statistics of Sri Lanka by district

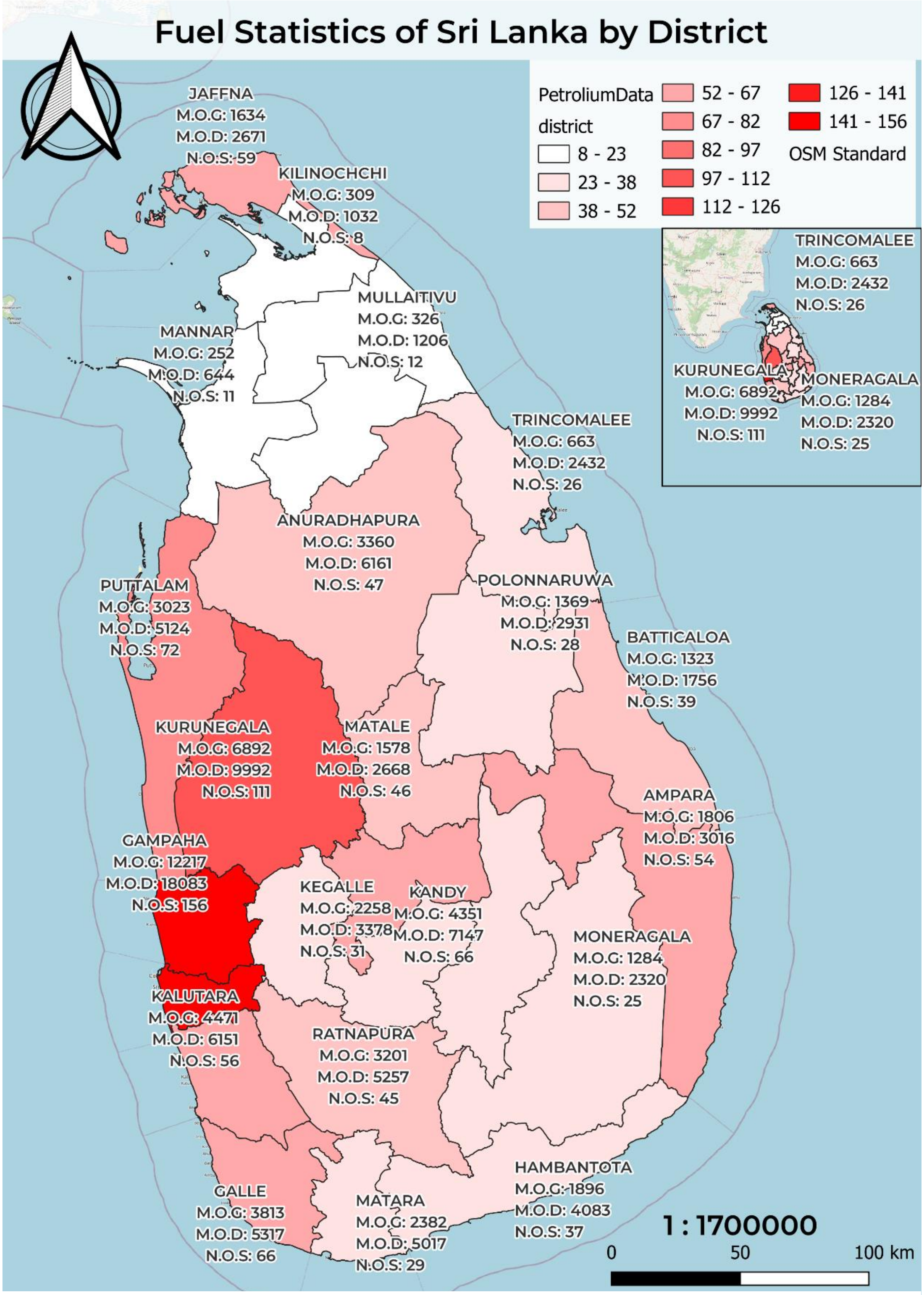


Figure 6.2: Fuel statistics of Sri Lanka by District

The map provided in figure 6.2 shows the thematic map of Sri Lanka by the no. of sheds in each district. According to the graph, the highest number of sheds are available in the Gampaha and Colombo districts. The lowest number of sheds are available in the Kilinochchi, Vavuniya, Mullaitivu and Mannar Districts. Also the means of gasoline distribution and diesel distribution are also provided in the labels of the map. The darker color shows the highest number of sheds, and the lighter colors show the least number of sheds.

Therefore it is safe to assume that the places with dense population have more no. of sheds and more distribution of diesel and gasoline.

6.2. Geo Spatial Database

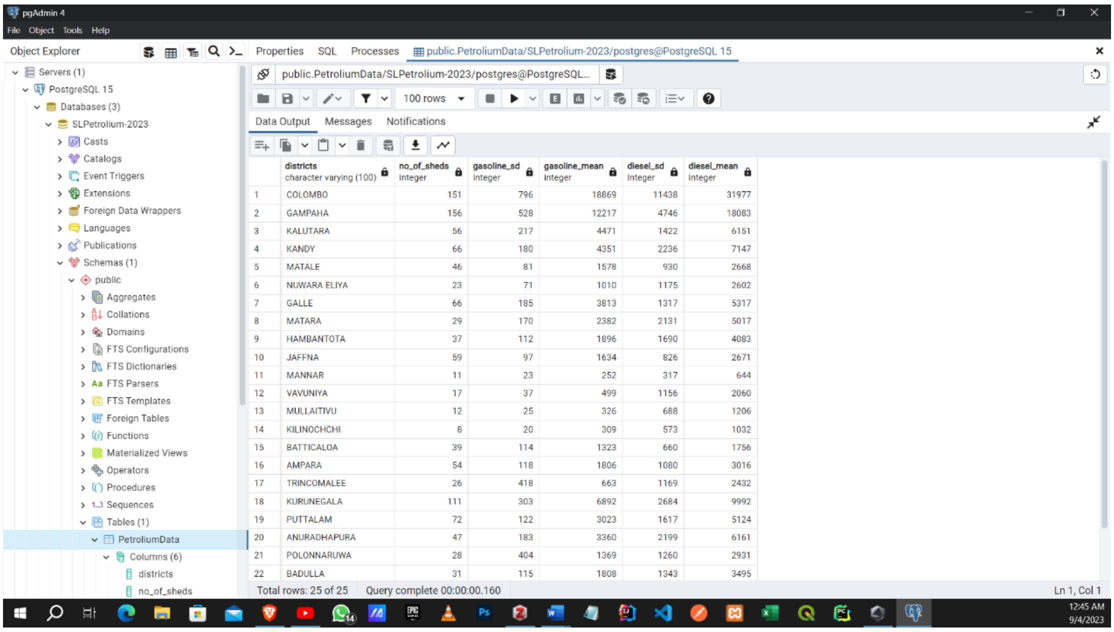


Figure 6.3: Geo Spatial database created for SL Petroleum distribution.

The figure 6.3 depicts the geo spatial database created for the SL Petroleum distribution data provided by the Sri Lanka Sustainable Energy Authority.

ALL STEPS RELATED TO THIS CHAPTER ARE IN <u>APPENDIX F</u>

CHAPTER SEVEN

LOCATIONS FOR RENEWABL ENERGY GENERATION

The following chapter is about feasible locations for Renewable Energy Generation in Sri Lanka identified by the Sri Lanka Energy Authority. A map is drawn to pinpoint the exact feasible locations using Google Earth Pro and QGIS.

Most common renewable energy generation methods in Sri Lanka are Wind Power, Solar Power, and Biomass power. Therefore the most suitable and feasible locations for the development of these sites must be identified. Sri Lanka Sustainable Energy Authority have provided some such locations for the development of renewable energy.

7.1. Factors considered when selecting a location.

Sri Lanka Sustainable Energy Authority have considered multiple factors when selecting a location for the development. For instance for selecting solar power development locations they tend to check the place with sunlight coverage throughout the year, easy to maintain places and places with least harm to the environment. When considering factors for wind power they have considered places with the most windfall. In addition to that the Monaragala, Polonnaruwa, and Batticaloa district is selected for the development of Biomass power plants.

7.2. Feasible Locations

7.2.1. Solar Power

Kilinochchi Solar Power Area	Latitude: 9°26'33.19"N Longitude: 80°23'1.94"E	Name: Kilinochchi Solar Power Area District: Kilinochchi
Ampara Solar Power Area	Latitude: 6°56'30.93"N Longitude: 81°51'12.18"E	Name: Ampara Solar Power Area District: Ampara
Hambantota Solar Power Area	Latitude: 6°11'45.53"N Longitude: 81°13'52.59"E	Name: Hambantota Solar Power Area District: Hambantota
Puttalam Solar Power Area	Latitude: 8°10'17.89"N Longitude: 79°56'14.75"E	Name: Puttalam Solar Power Area District: Puttalam

Table 7.1: Feasible locations for Solar Power

7.2.2. Wind Power

Mannar Wind Power 1	Latitude: 9° 3'13.67"N Longitude: 79°49'22.89"E	Name: Mannar Wind Power 1 District: Mannar
Mannar Wind Power 2	Latitude: 8°48'29.77"N Longitude: 79°55'4.26"E	Name: Mannar Wind Power 2 District: Mannar
Kilinochchi Wind Power	Latitude: 9°17'18.94"N Longitude: 79°59'42.22"E	Name: Kilinochchi Wind Power District: Kilinochchi
Jaffna Wind Power 1	Latitude: 9°28'47.15"N Longitude: 79°43'9.82"E	Name: Jaffna Wind Power District: Jaffna
Puttalam Wind Power	Latitude: 8° 8'30.25"N Longitude: 79°50'38.67"E	Name: Puttalam Wind Power District: Puttalam
Trincomalee Wind Powe	Latitude: 8°30'6.89"N Longitude: 81°18'26.78"E	Name: Trincomalee Wind Power District: Trincomalee
Anuradhapura Wind Power	Latitude: 8° 7'7.55"N Longitude: 80°43'54.55"E	Name: Anuradhapura Wind Power District: Anuradhapura
Hambantota Wind Power	Latitude: 6°18'22.96"N Longitude: 81°27'51.93"E	Name: Hambantota Wind Power District: Hambantota

Table 7.2: Feasible locations for wind power development

7.2.3. Biomass

Batticaloa Biomass Power	Latitude: 7° 1'44.54"N Longitude: 81°48'2.66"E	Name: Batticaloa Biomass Power District: Batticaloa
Polonnaruwa Biomass Power	Latitude: 7°55'17.09"N Longitude: 81°12'35.38"E	Name: Polonnaruwa Biomass Power District: Polonnaruwa
Monaragala Biomass Power	Latitude: 6°58'25.07"N Longitude: 81°25'47.25"E	Name: Monaragala Biomass Power District: Monaragala
Hambantota Biomass Power	Latitude: 6°23'14.32"N Longitude: 81°26'22.86"E	Name: Hambantota Biomass Power District: Hambantota

Table 7.3: Feasible locations for biomass power development

7.3. Mapping the locations

By considering all these locations as potential locations for renewable energy development they are mapped in Google Earth Pro and a graph representing the locations are drawn and depicted in figure 7.1 below.

According to the figure in 7.1, there are multiple locations suitable for renewable energy generation and they are marked with certain symbols. The triangle symbol represents potential land for solar power development. Pink circle represents feasible land for wind power development and black circle represents biomass power development.

Almost all of the wind power locations are situated in coastal areas and solar power locations are situated in uninhabited places.

ALL STEPS RELATED TO THIS CHAPTER ARE IN <u>APPENDIX G</u>

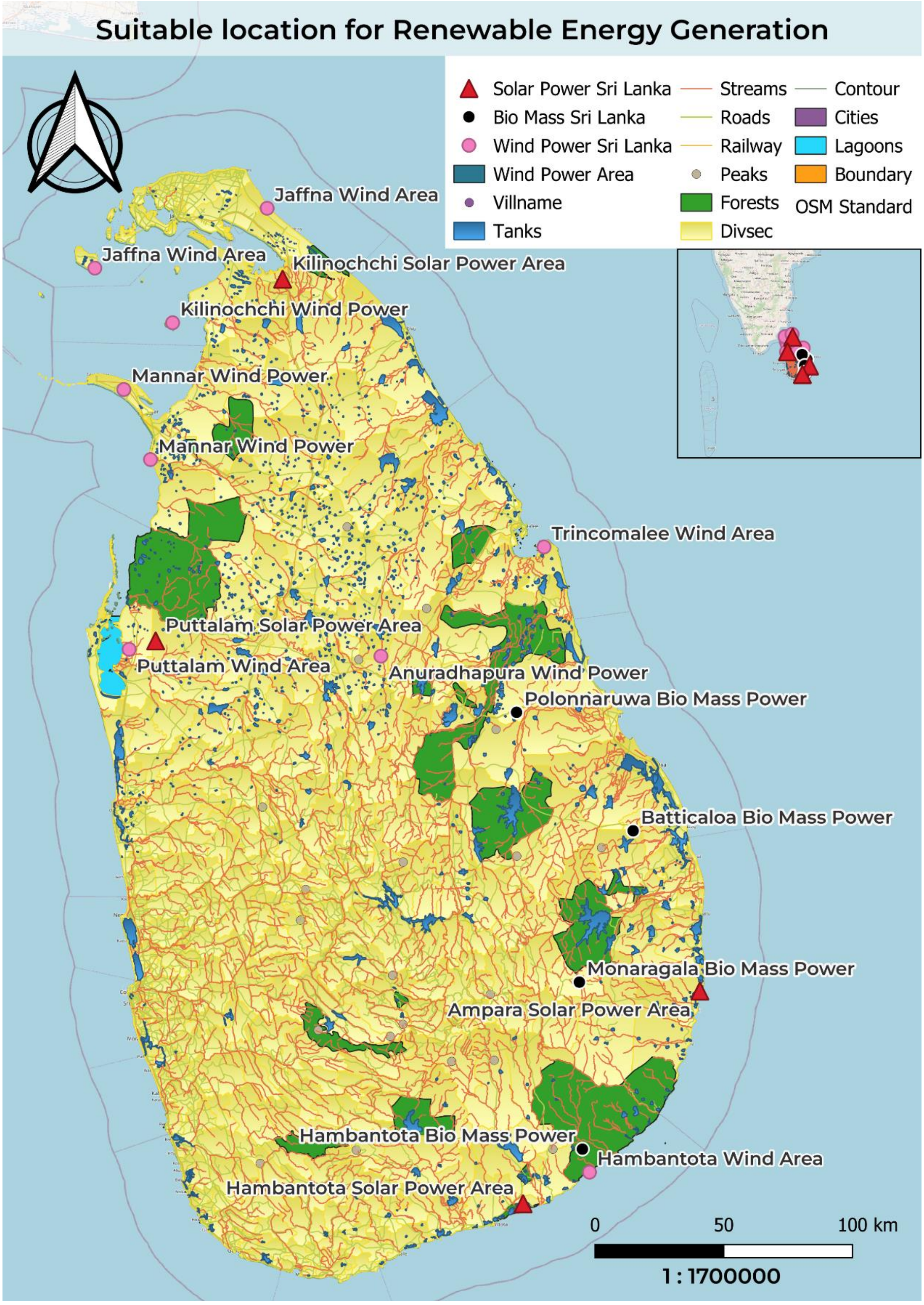


Figure 7.1: Feasible locations for renewable energy development

CHAPTER EIGHT

RENEWABLE ENERGY RESEARCH CENTER FOR KANDY

The following chapter consists of the map that is developed to find the suitable land for the newly developing ‘*Regional Research Center for Renewable Energy*’ in Kandy. The suitable area in the map is developed using the criteria provided in the task and the total number of buildings within the suitable area at present, land area occupied by the buildings, and total suitable land area is found out.

8.1. Criteria for the suitable Area

500m away from the Uruwala Primary School
600m away from Industrial Development Center
700m away from the Jayathilaka Hall

Table 8.1: criteria for the suitable area

The above table 8.1 mentions the requirements of the land area where the new research center is to be developed.

8.2. Suitable area for the newly developing research center

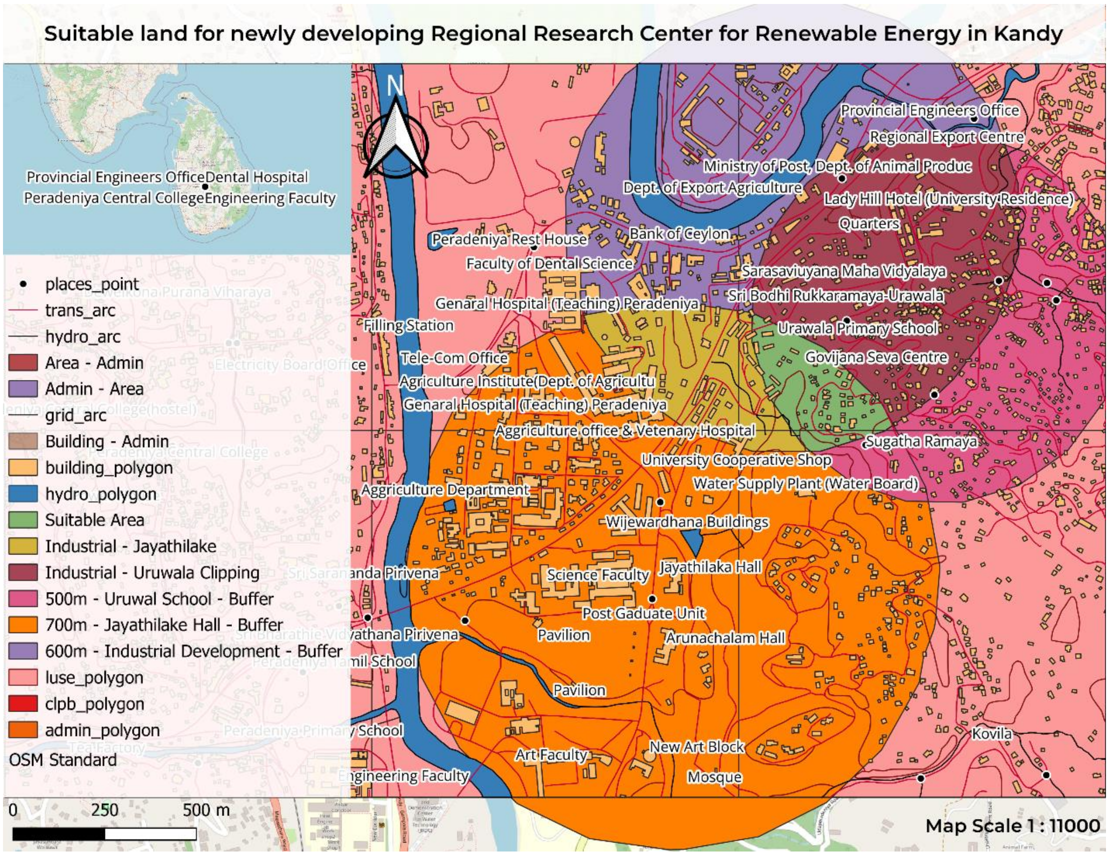


Figure 8.1: Suitable area for the research center

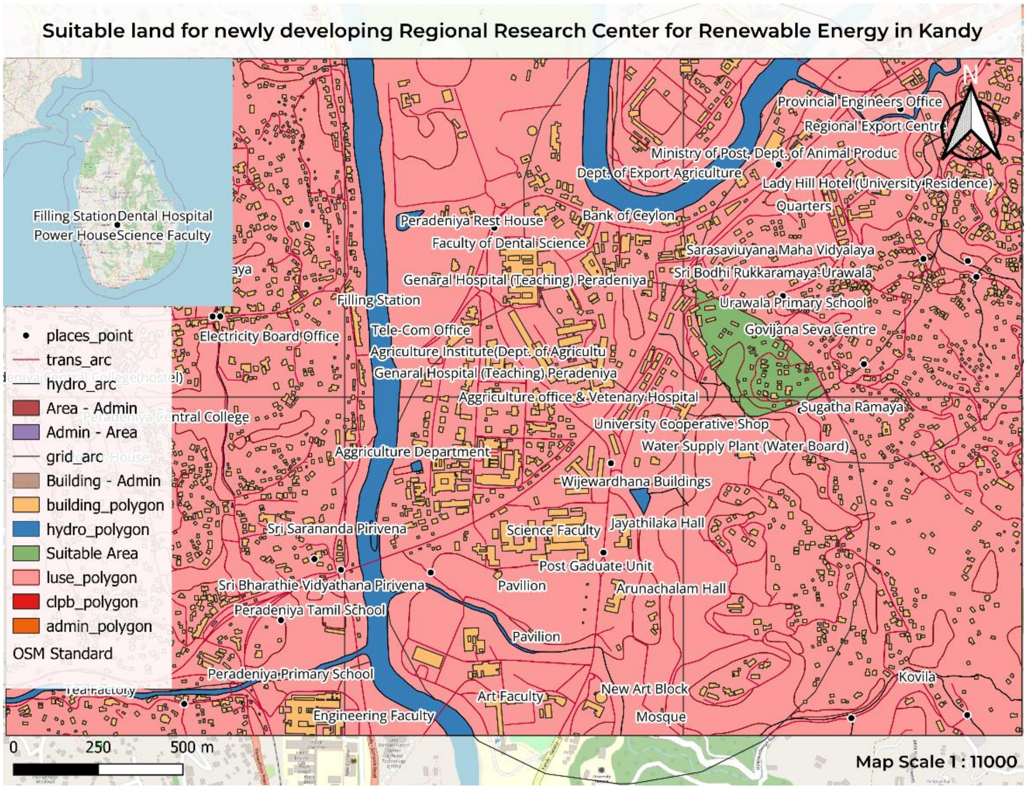


Figure 8.2: Suitable area

Figure 8.1 shows the suitable area that is selected for the development of regional research center in green color. This piece of land is selected by meeting all conditions put forth by the department. The orange circle visible in the map is the buffer drawn to show the 700m radius of Jayathilaka hall, 600m radius circle drawn around the Industrial development center is purple and Uruwala primary school's circle is pink with a radius of 500m. The places are then intersected to find the most suitable piece of land for the newly developed research center.

Statistics of the suitable area	
Buildings situated within the selected area	55 Buildings
Total Land area occupied by the buildings	7928.76m^2
Total Suitable Land Area	73421m^2

Table 8.2: Statistics of the suitable location

ALL STEPS RELATED TO THIS CHAPTER ARE IN APPENDIX H

CONCLUSION

The above report comprises the use of geo-spatial analysis and statistical analysis. This report has approached combining geospatial and statistical analysis to explore critical aspects of energy efficiency and renewable energy development.

The first chapter consists of associations between building cooling with various factors affecting residential and commercial electricity consumption. This chapter laid the foundation for a deeper understanding of energy-saving strategies.

By using the data collected from Chapter One a statistical model to predict cooling load through regression analysis is provided. This model provides a powerful tool for optimizing energy consumption within buildings, contributing to the broader goal of energy conservation.

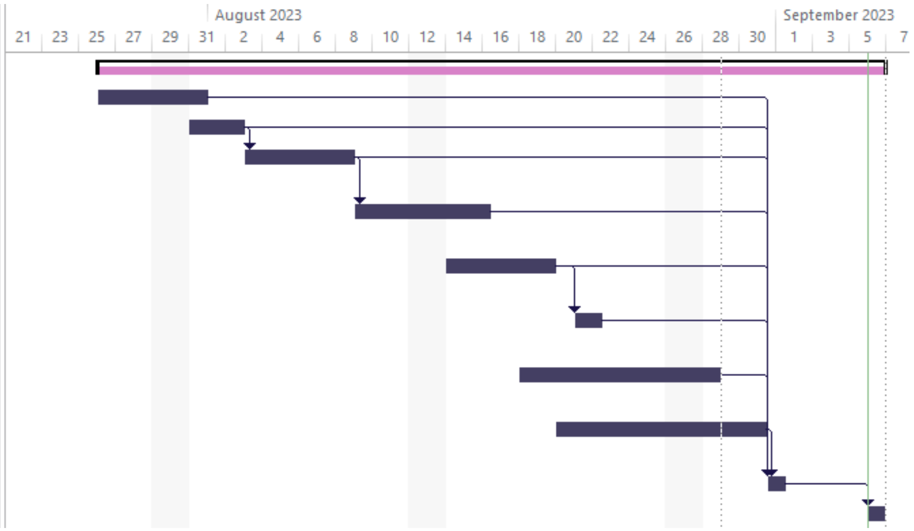
In Chapter Three, we uncovered lands ripe for wind power plant development across Sri Lanka's districts, leveraging data from the Renewable Energy Resource Development Plan. In Chapter Four, we examined the potential for solar power development and analyzed how solar energy can be a potent force in mitigating energy sector challenges.

Chapter five provided a small area of Hambantota, which is digitized by highlighting features such as solar parks, forests, water bodies and others by creating shape files. In Chapter Six, geo spatial analysis is done by creating a database in Postgres SQL and presenting the fuel distribution data in a thematic map.

Chapter Seven consists of the feasible locations for renewable energy development. Locations are pinned using Google Earth pro and the mapping is done using QGIS. The final chapter analysis is done to select the most suitable location for the newly developing regional research center in Kandy. And finally, the report ended with the conclusion.

GANTT CHART

Task Name	Duration	Start	Finish
Analytics and Business Intelligence	**31 days**	**Wed 7/26/23**	**Wed 9/6/23**
Chapter One - R-Studio Analytics	4 days	Wed 7/26/23	Mon 7/31/23
Chapter Two - The Statistical Model	3 days	Mon 7/31/23	Wed 8/2/23
Chapter Three - Sri Lanka Wind Power Development	4 days	Thu 8/3/23	Tue 8/8/23
Chapter Four - Potential land for Solar Power Development	5.26 days	Wed 8/9/23	Wed 8/16/23
Chapter Five - Digitized Informative Area Map	6 days	Mon 8/14/23	Sat 8/19/23
Chapter Six- Sri Lanka Petroleum Distribution	1.5 days	Mon 8/21/23	Tue 8/22/23
Chapter Seven - Locations for Renewable Energy Generation	7 days	Fri 8/18/23	Mon 8/28/23
Chapter Eight - Renewable Energy Research Center for Kandy	9.5 days	Sun 8/20/23	Thu 8/31/23
Documentation	1 day	Thu 8/31/23	Fri 9/1/23
Submission	1 day	Wed 9/6/23	Wed 9/6/23



REFERENCES

Sustainable Energy Authority, 2023. *New Renewable Energy*. [online] Available at: <https://www.energy.gov.lk/index.php/en/renewable-energy/new-renewable-energy> [Accessed 28 August 2023].

APPENDIX A

Chapter One R Studio Codes

```
#import the dataset
unclean.data <- read.csv("energy_efficiency_data.csv")

#data preprocessing
energy_data <- na.omit(unclean.data)

#data row count
nrow(energy_data)

#data column count
ncol(energy_data)

#head of the dataset
head_energy <- head(energy_data)

print(head_energy)


library(ggplot2)

# Select numerical columns for visualization
numerical_cols <- c("Relative_Compactness", "Surface_Area", "Wall_Area", "Roof_Area", "Overall_Height",
                    "Orientation", "Glazing_Area", "Glazing_Area_Distribution", "Heating_Load", "Cooling_Load")

# Create histograms for all numerical columns
histograms <- lapply(numerical_cols, function(col) {
  ggplot(energy_data, aes(x = .data[[col]])) +
    geom_histogram(binwidth = 0.5, fill = "dodgerblue", color = "black") +
    labs(x = col, y = "Frequency") +
    theme_minimal()
})
# Arrange histograms in a grid
library(gridExtra)
grid.arrange(grobs = histograms, ncol = 2)


# Select numerical columns for visualization
numerical_cols <- c("Relative_Compactness", "Surface_Area", "Wall_Area", "Roof_Area", "Overall_Height",
                    "Orientation", "Glazing_Area", "Glazing_Area_Distribution", "Heating_Load", "Cooling_Load")

# Subset of numerical columns to create scatterplots with Cooling_Load
scatterplot_cols <- numerical_cols[!numerical_cols %in% c("Cooling_Load")]

# Create scatterplots for each column against Cooling_Load
scatterplots <- lapply(scatterplot_cols, function(col) {
  ggplot(energy_data, aes(x = .data[[col]], y = Cooling_Load)) +
    geom_point(color = "dodgerblue") +
    geom_smooth(method = "lm", color = "red", se = FALSE) +  # Add linear regression line
    labs(x = col, y = "Cooling Load") +
    theme_minimal()
})

# Arrange scatterplots in a grid
grid.arrange(grobs = scatterplots, ncol = 2)


# Import the necessary packages
install.packages("psych")
library(psych)

# Calculate correlation coefficients and their significance levels
correlation_results <- cor.test(energy_data$Cooling_Load, energy_data$Relative_Compactness)
print(correlation_results)

# Perform correlation tests for other factors
factors <- c("Relative_Compactness", "Surface_Area", "Wall_Area", "Roof_Area", "Overall_Height", "Glazing_Area", "Glazing_Area_Distr

for (factor in factors) {
  correlation_results <- cor.test(energy_data$Cooling_Load, energy_data[[factor]])
  cat("Correlation Test between Cooling_Load and", factor, "\n")
  print(correlation_results)
  cat("\n")
}


# Calculate the correlation matrix
correlation_matrix <- cor(energy_data[, factors])

# Print the correlation matrix
print(correlation_matrix)

library(corrplot)
#correlation heatmap
corrplot(correlation_matrix, method = "color", tl.col = "black")
```

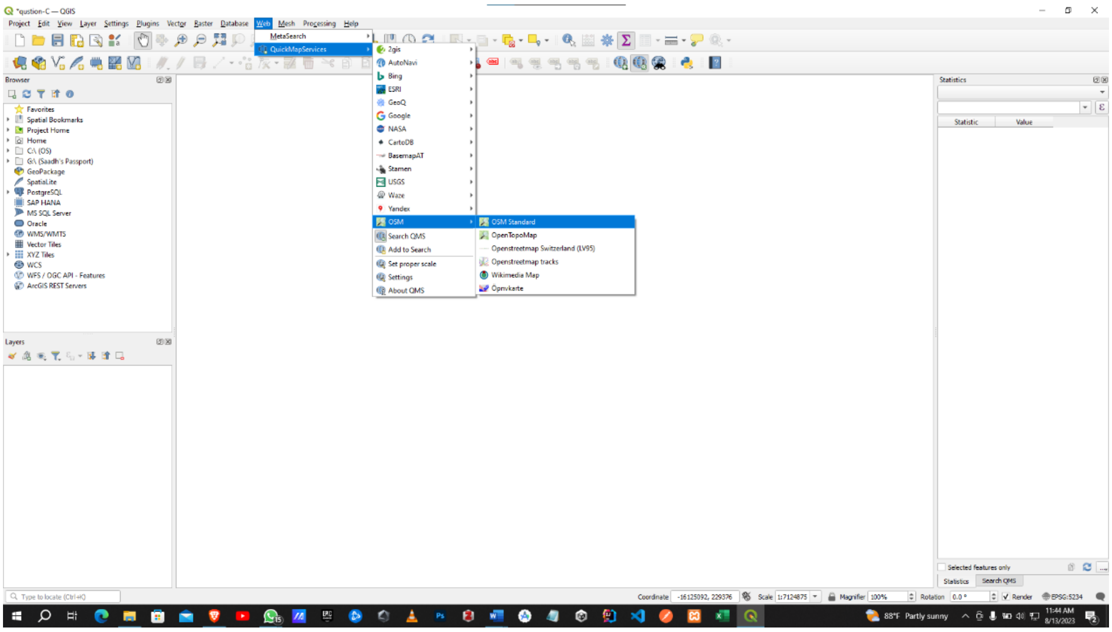
APPENDIX B

Chapter Two R Studio Codes

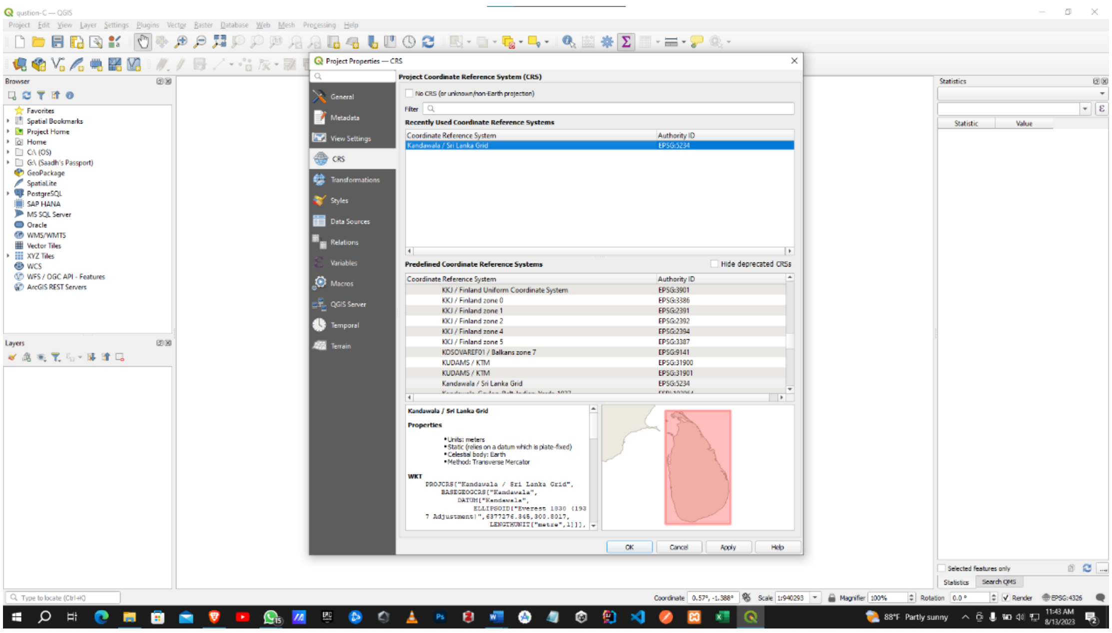
```
#relationship among
statistical.model <- lm(Cooling_Load ~ Relative_Compactness + Surface_Area + Wall_Area + Roof_Ar
print(summary(statistical.model))

#linear regression model
rep.statistical.model <- lm(Cooling_Load ~ Relative_Compactness + Surface_Area +
                              Wall_Area + Overall_Height , data = energy_data)

print(summary(rep.statistical.model))

print(rep.statistical.model)

# Shapiro-Wilk test for normality
shapiro.test(resid(rep.statistical.model))

# Create a Q-Q plot
ggplot(data = NULL, aes(sample = resid(rep.statistical.model))) +
  geom_qq() +
  geom_qq_line() +
  labs(title = "Q-Q Plot of Residuals", x = "Theoretical Quantiles", y = "Sample Quantiles")
```

APPENDIX C

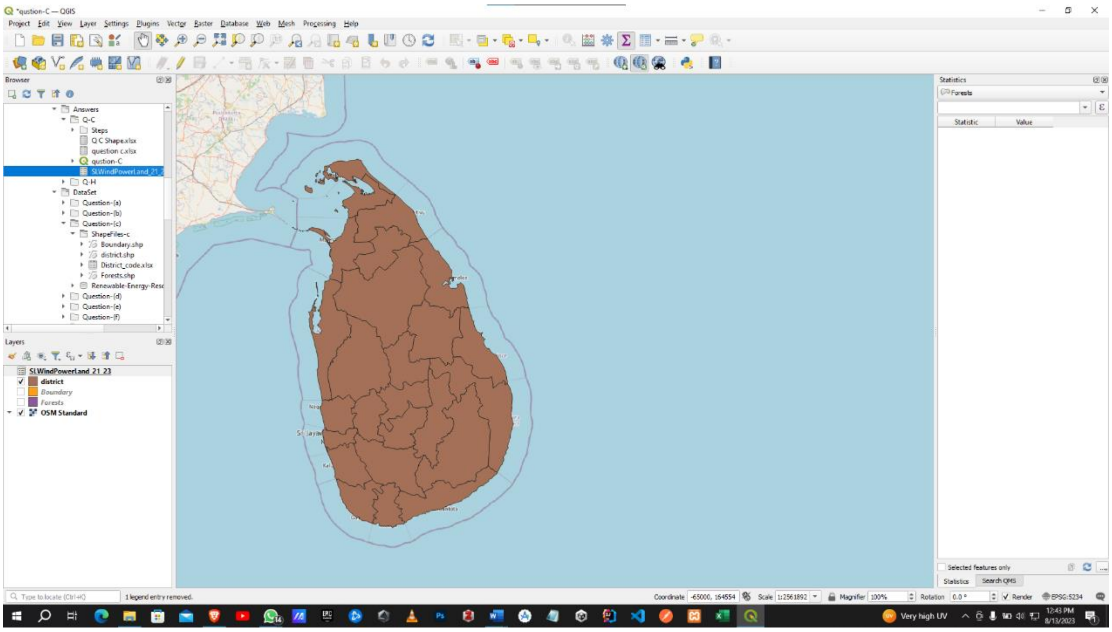
Adding OSM Standard Map



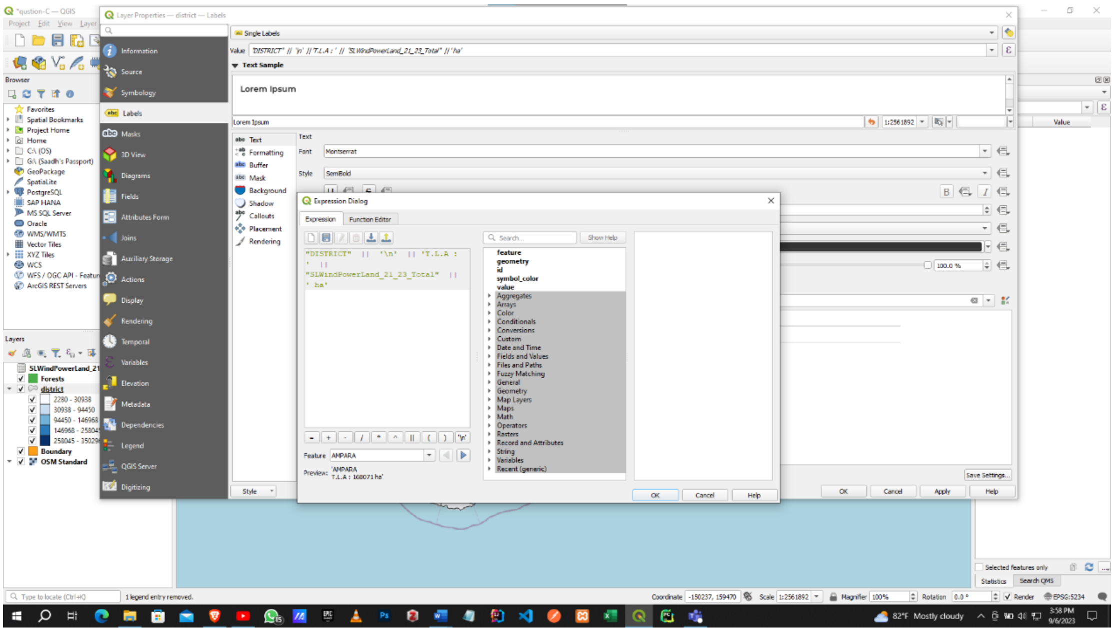
Selecting Grid system



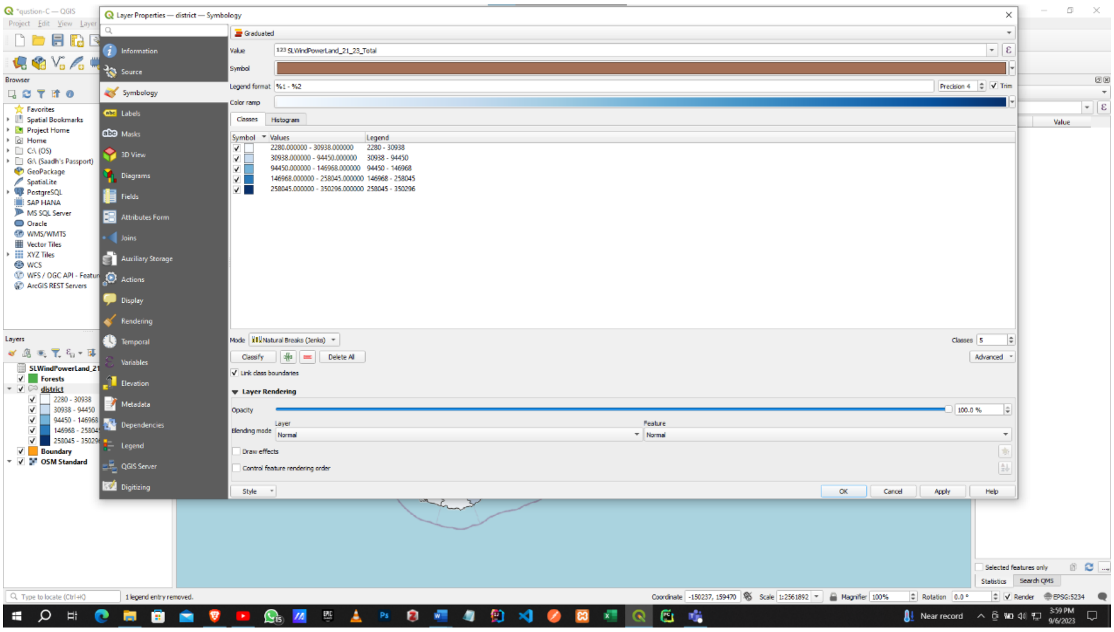
Adding Shape files



Adding labels.

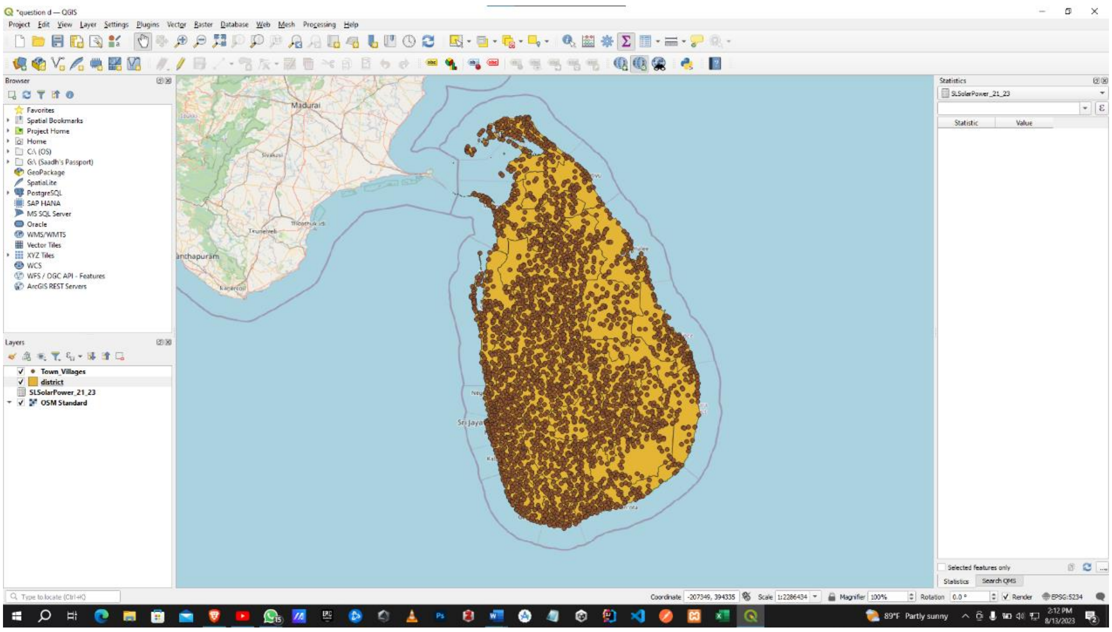


Adding Symbology

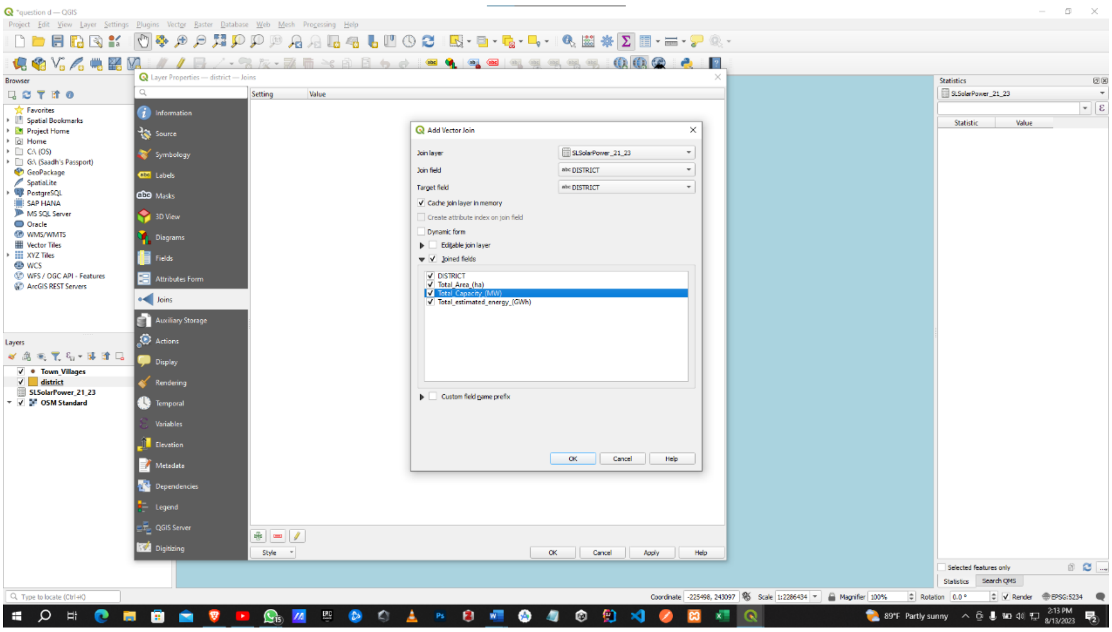


APPENDIX D

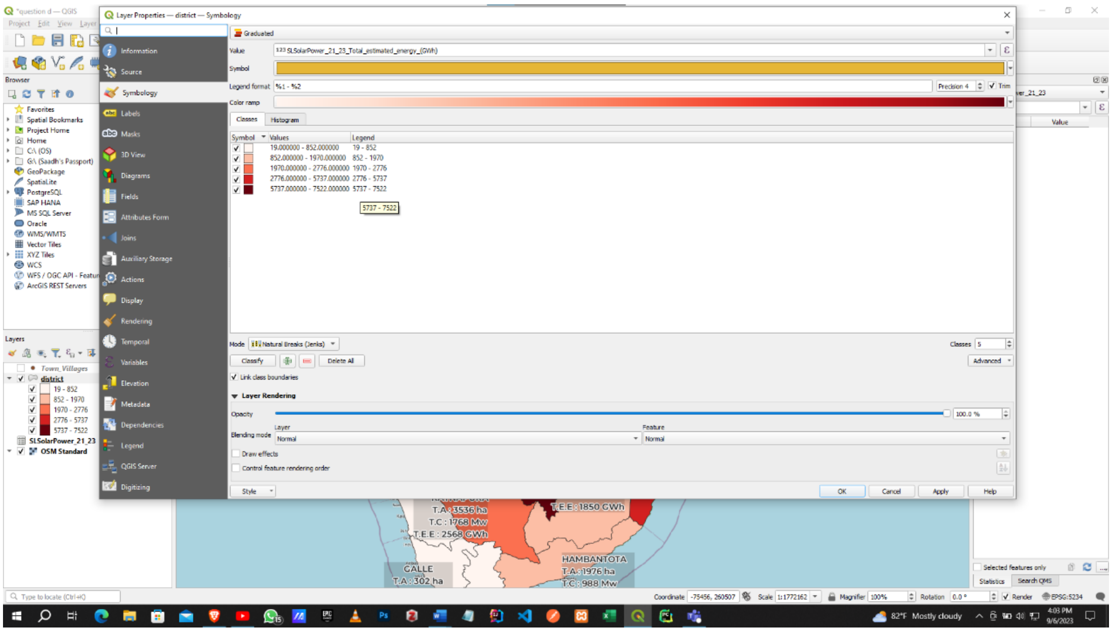
Added the shape files and SLSolarPower21_23.csv dataset.



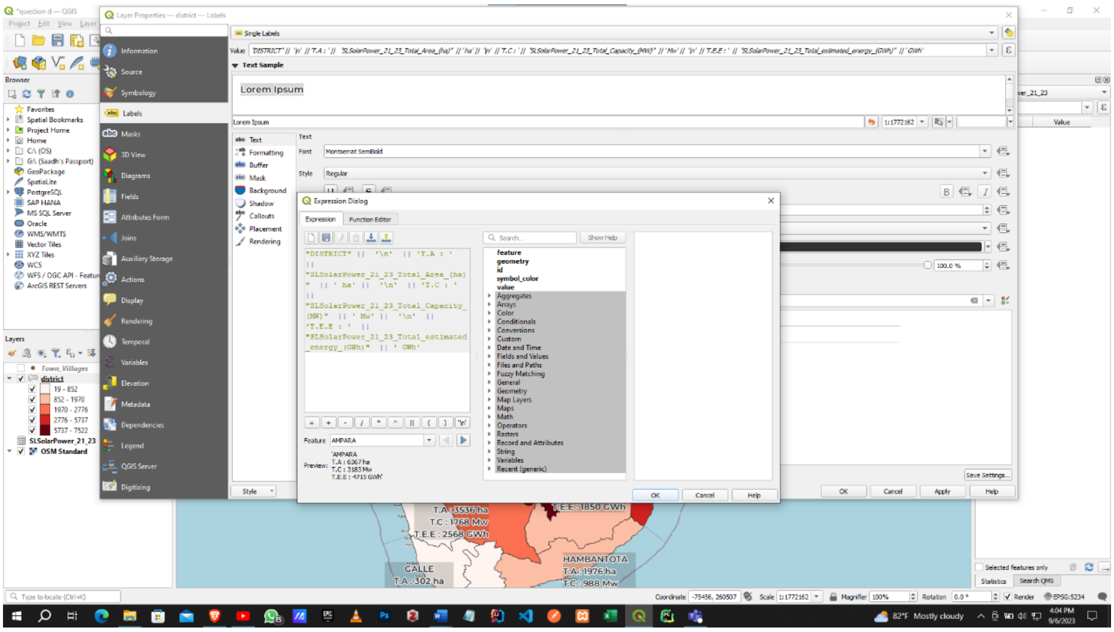
Joining the shape file and dataset



Adding Symbology

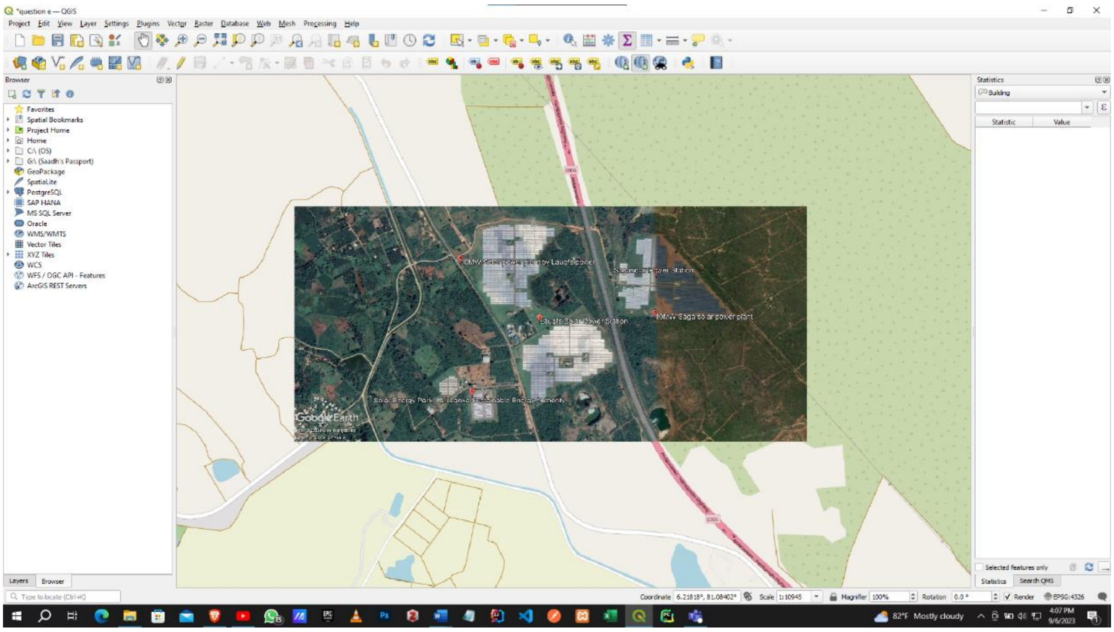


Adding labels

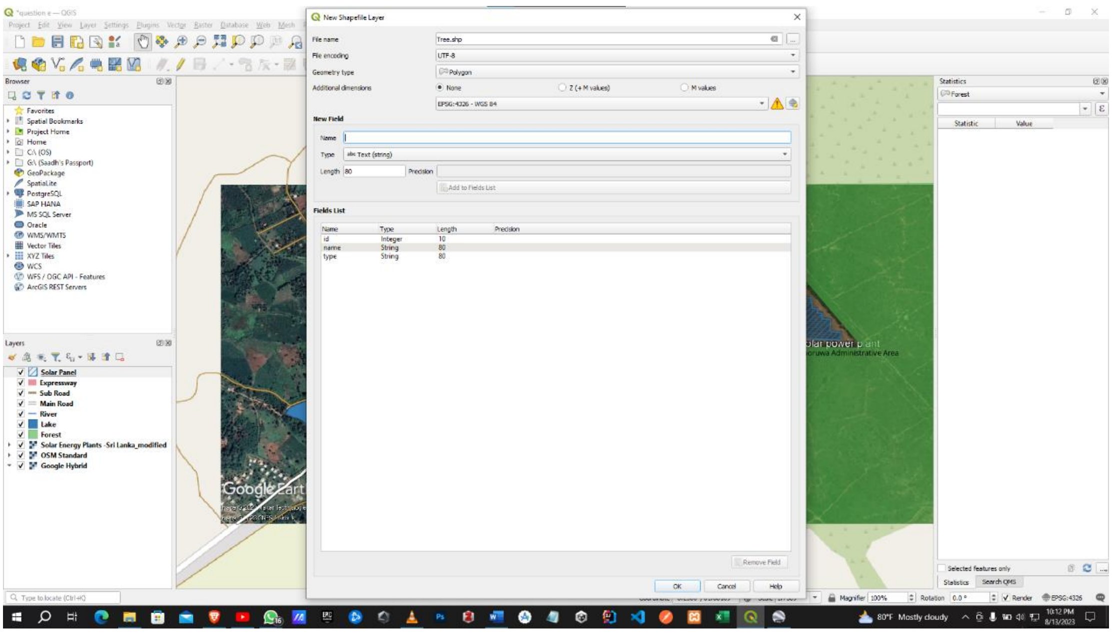


APPENDIX E

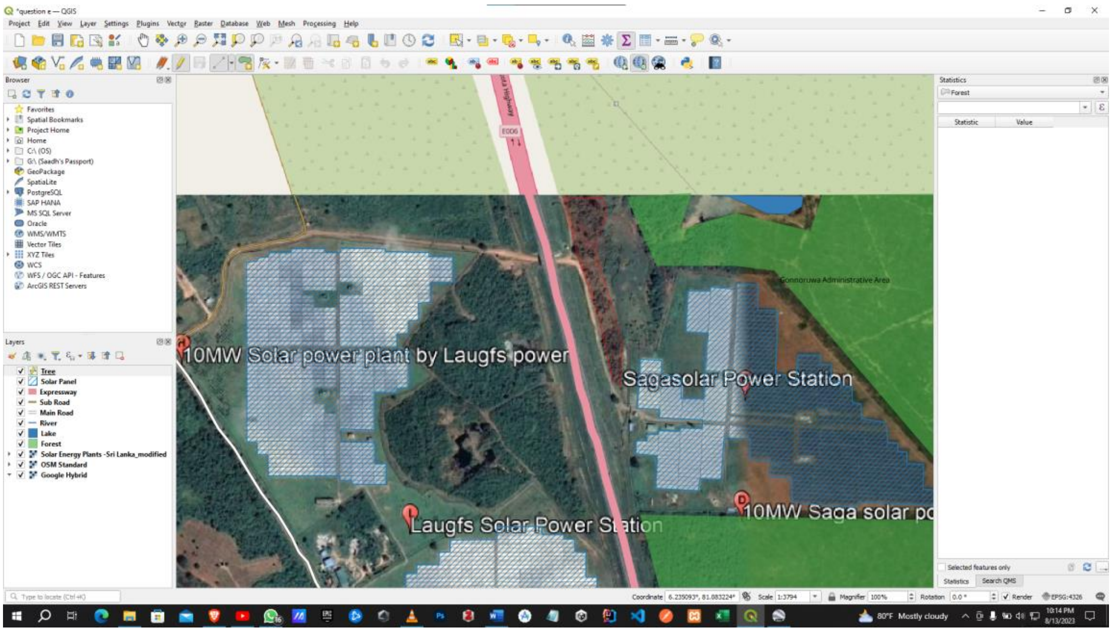
Added the image on the OSM standard map.



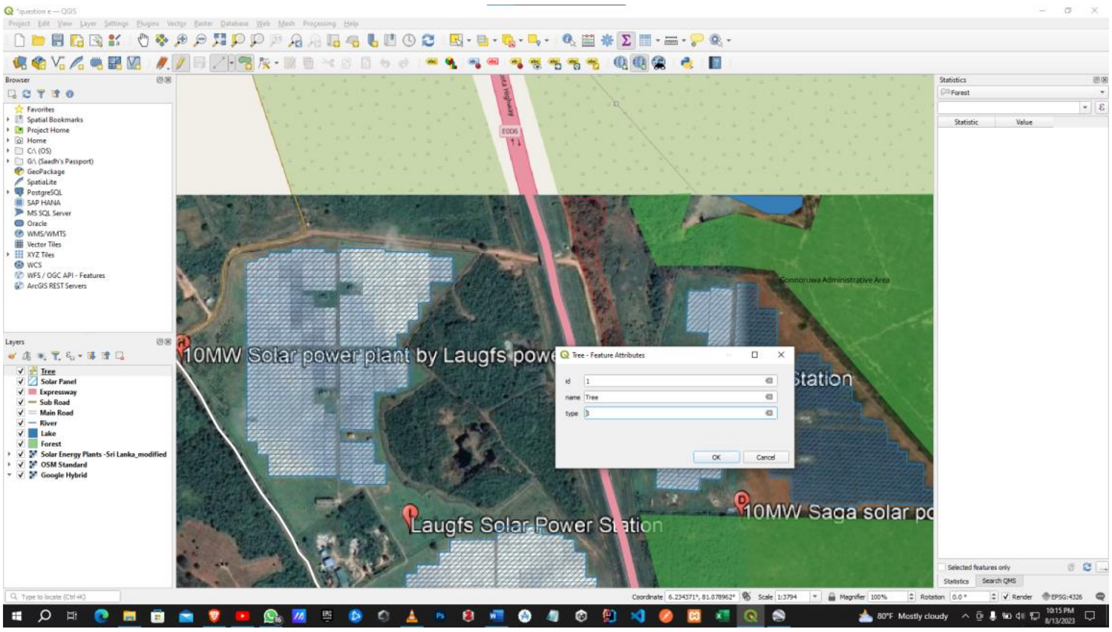
Creating shape files



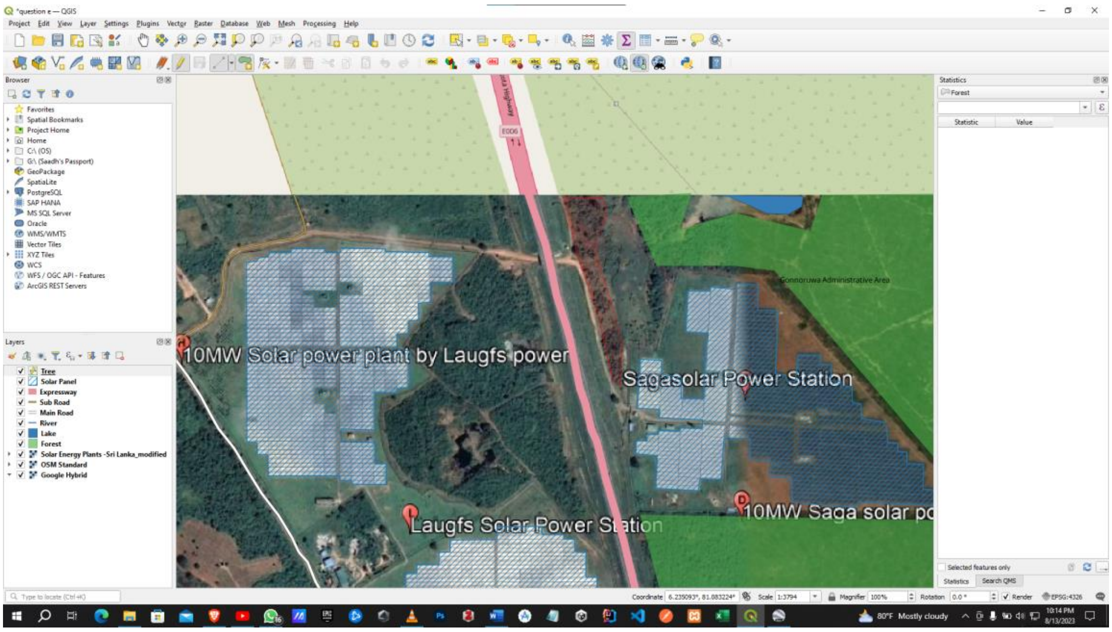
Selecting Areas for the shape file



Saving Attributes

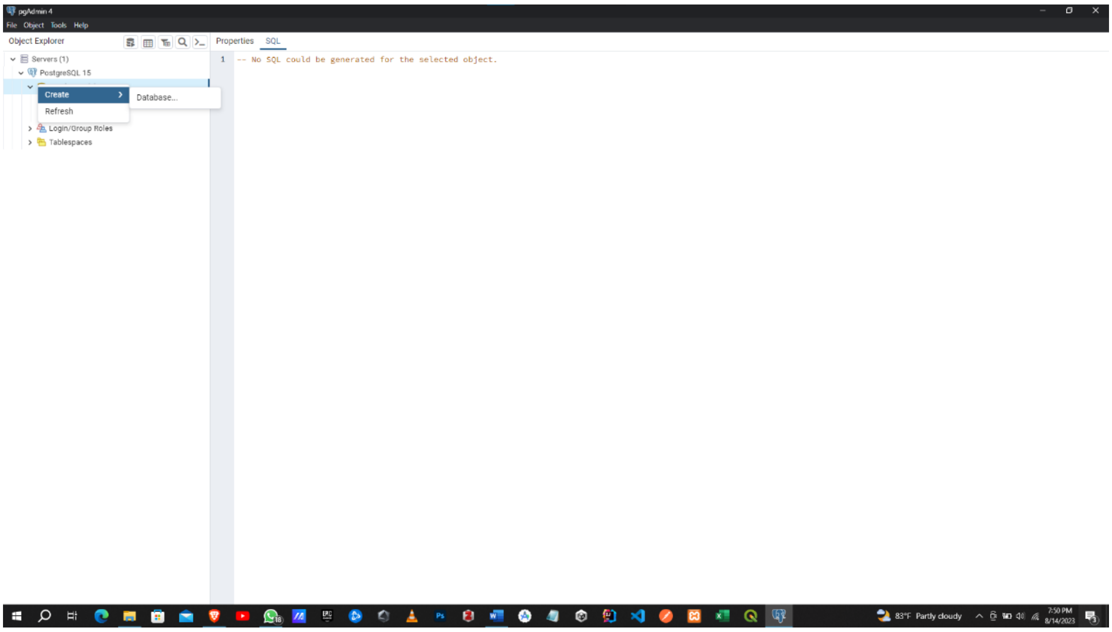


Created shape files.

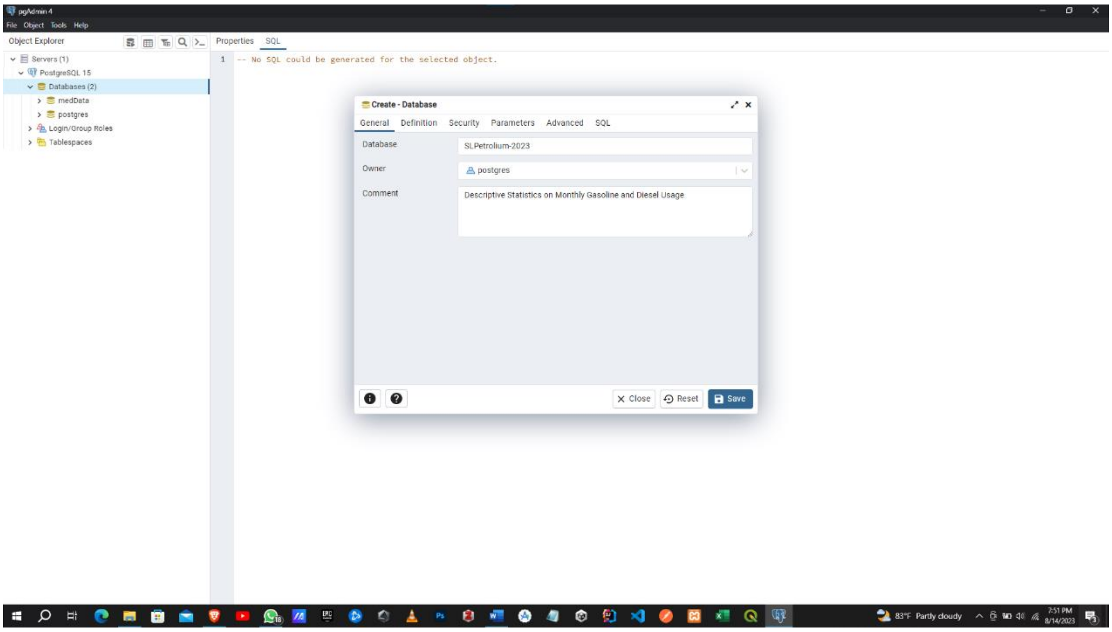


APPENDIX F

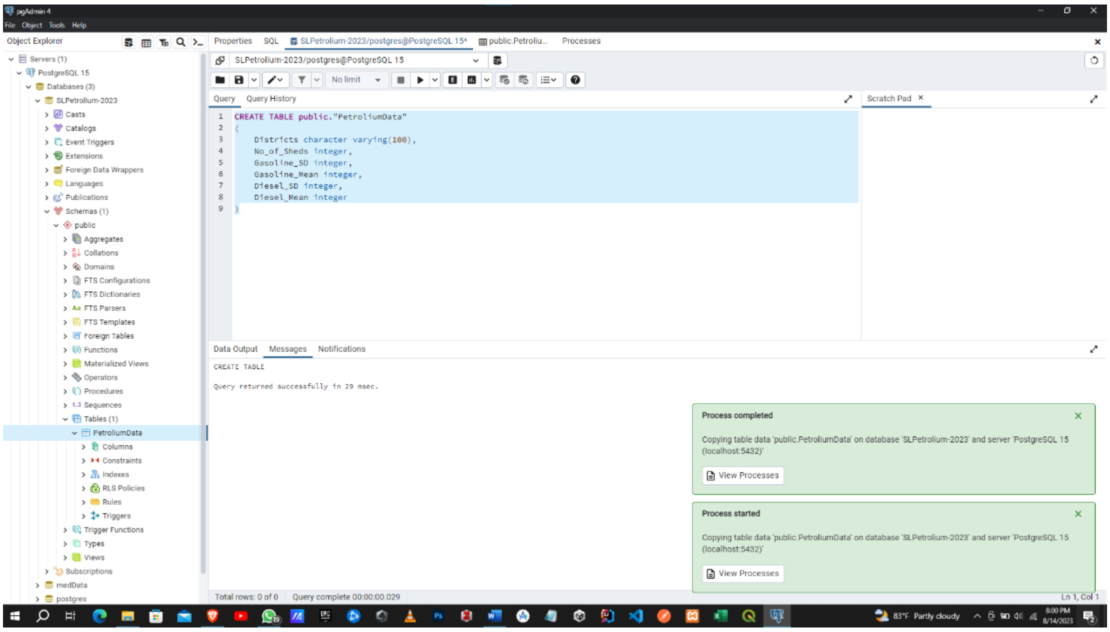
Create Postgres SQL database.



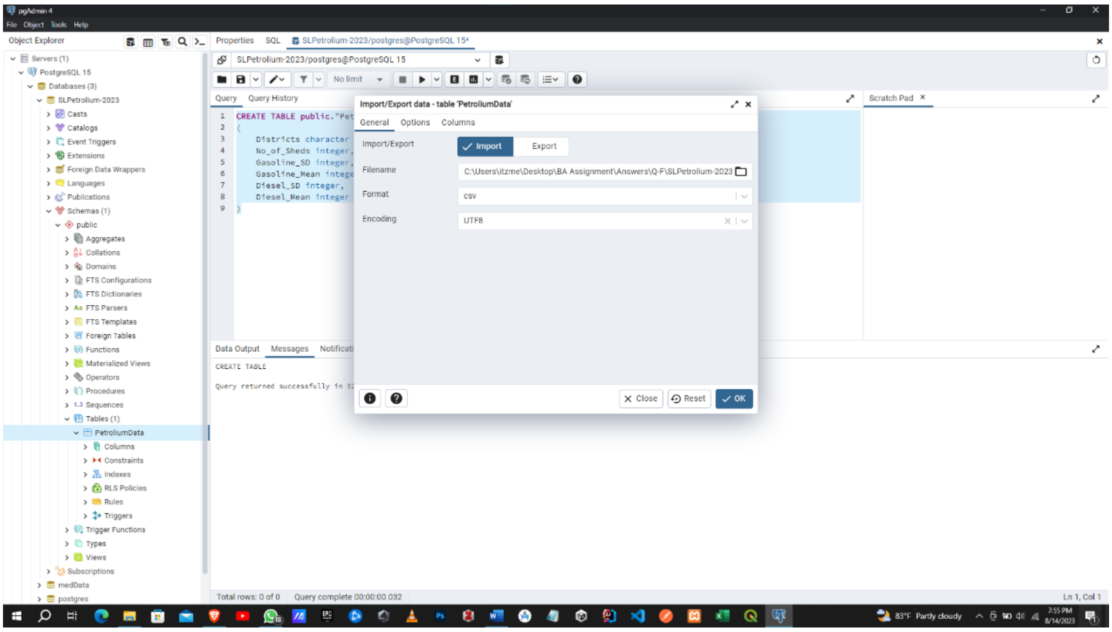
Give the name to the database.



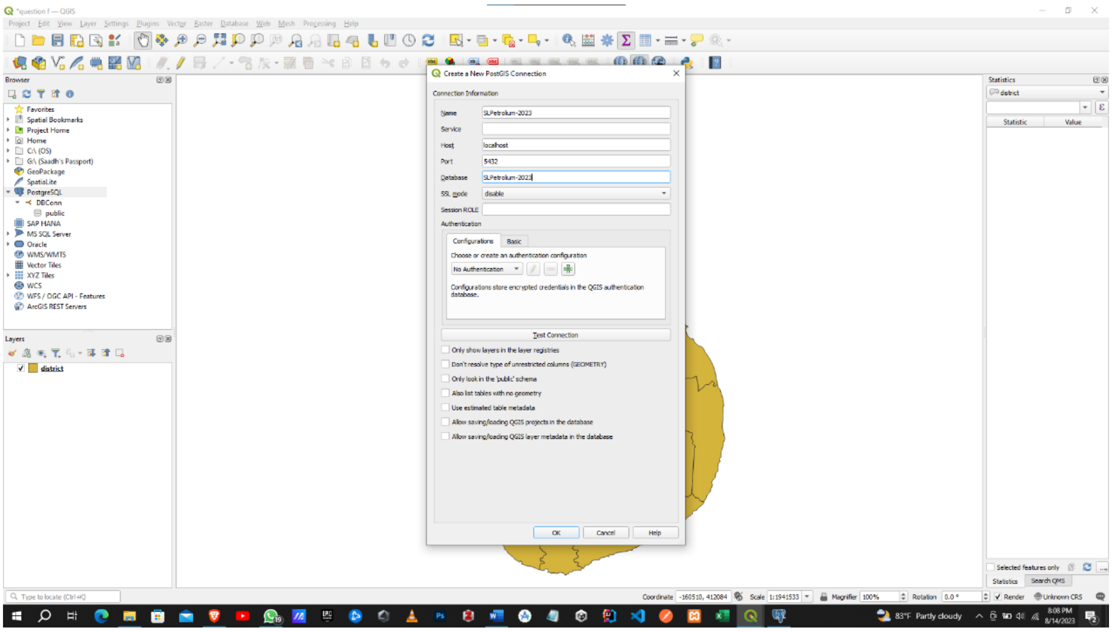
Create table.



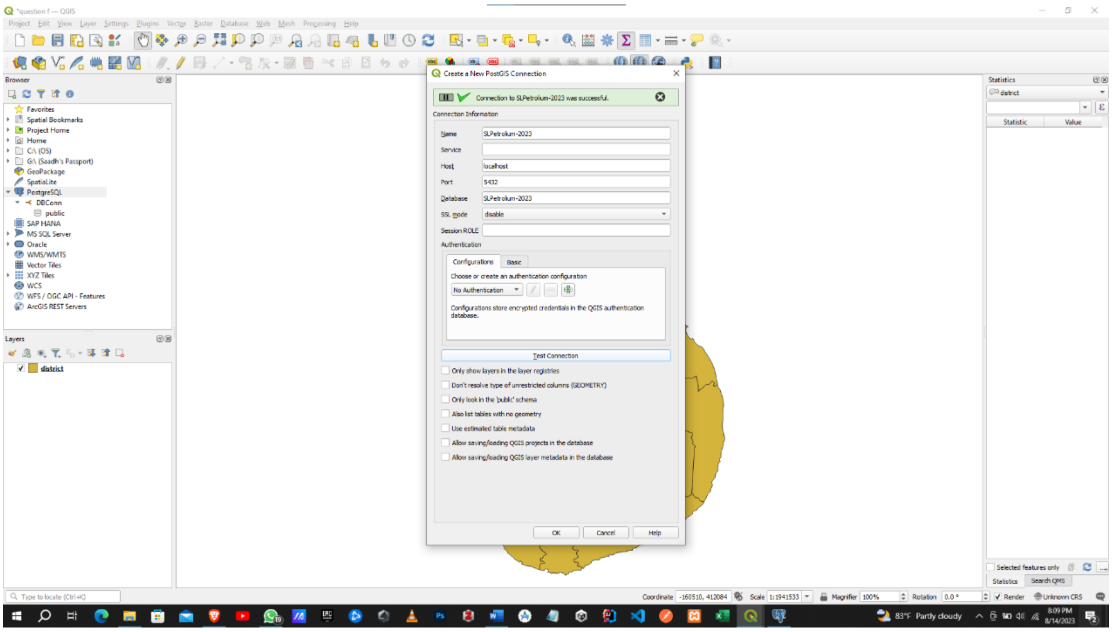
Import the dataset.



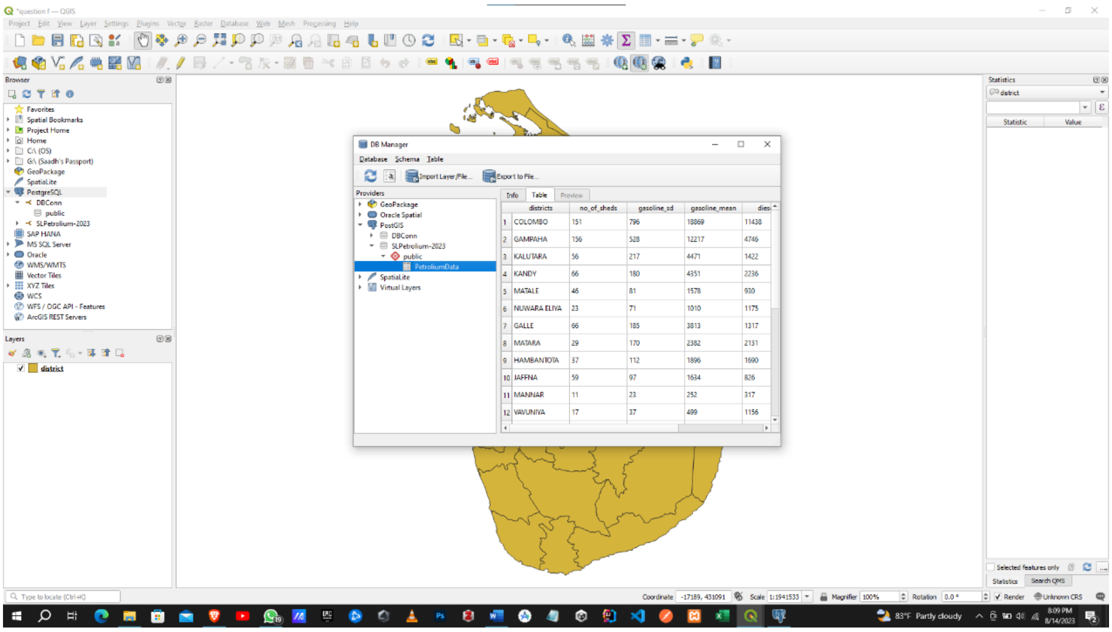
Link the database with QGIS.



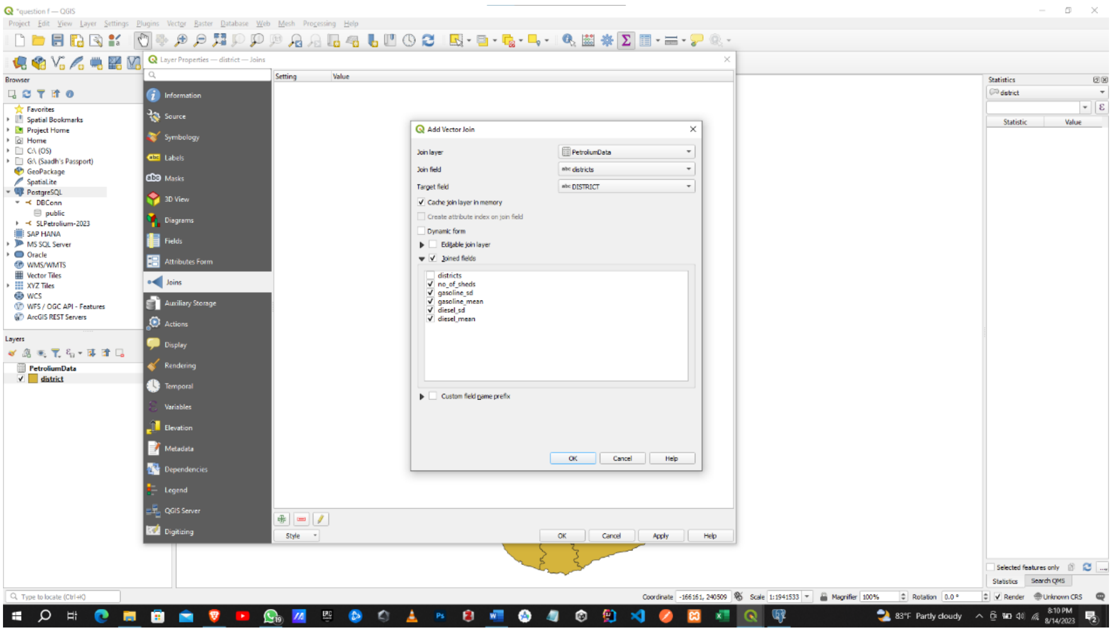
Test the DB Connection



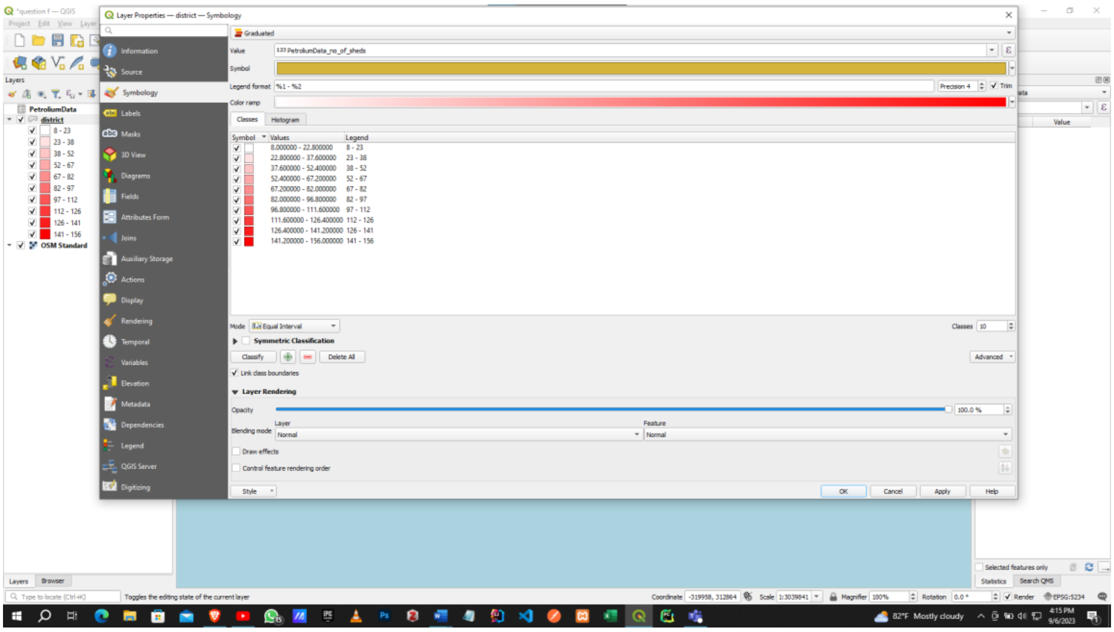
Check whether the data is retrieved.



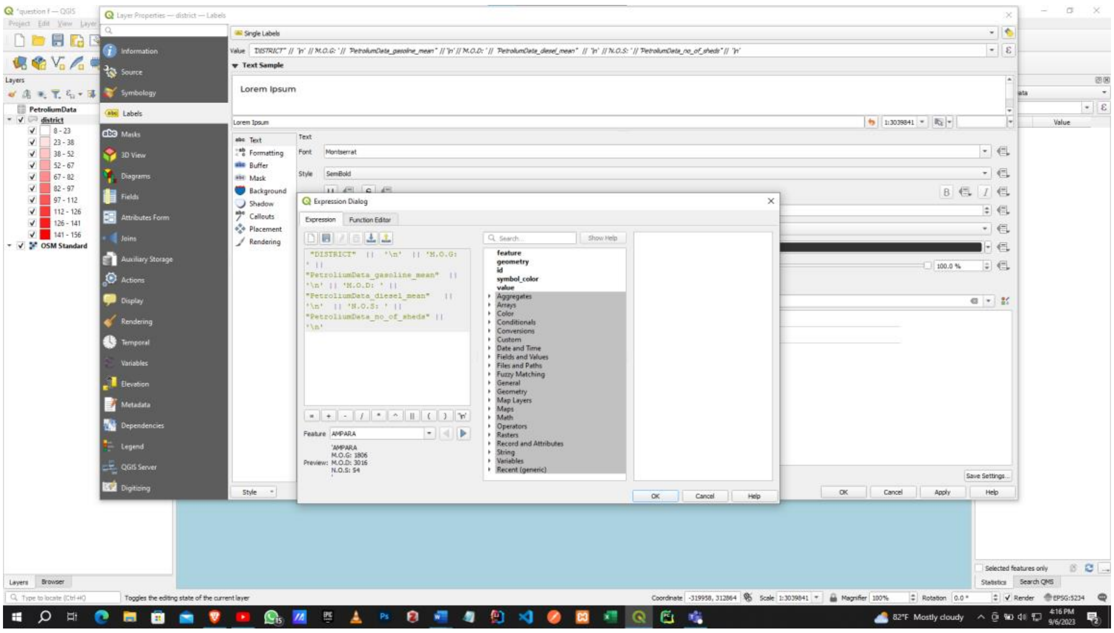
Join it with the shape layer.



Adding Symbology

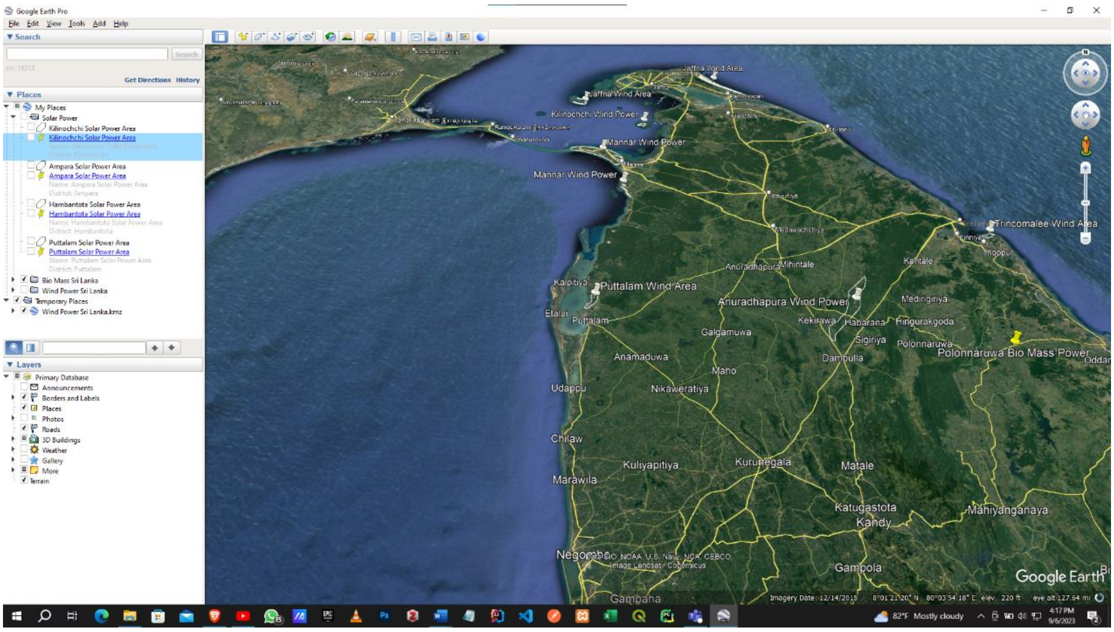


Adding Labels

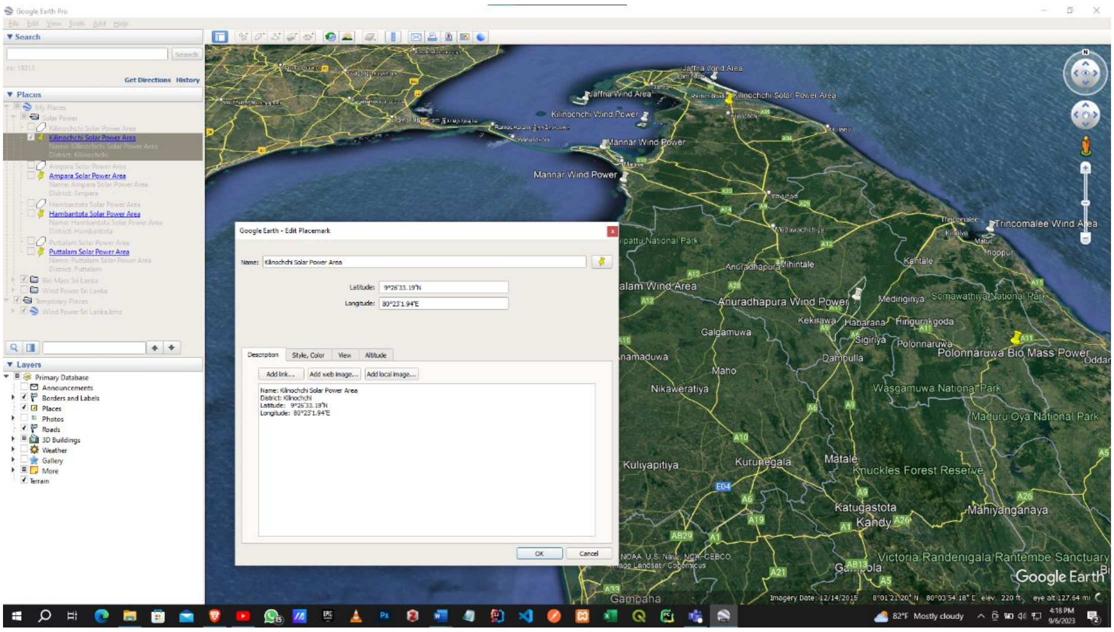


APPENDIX G

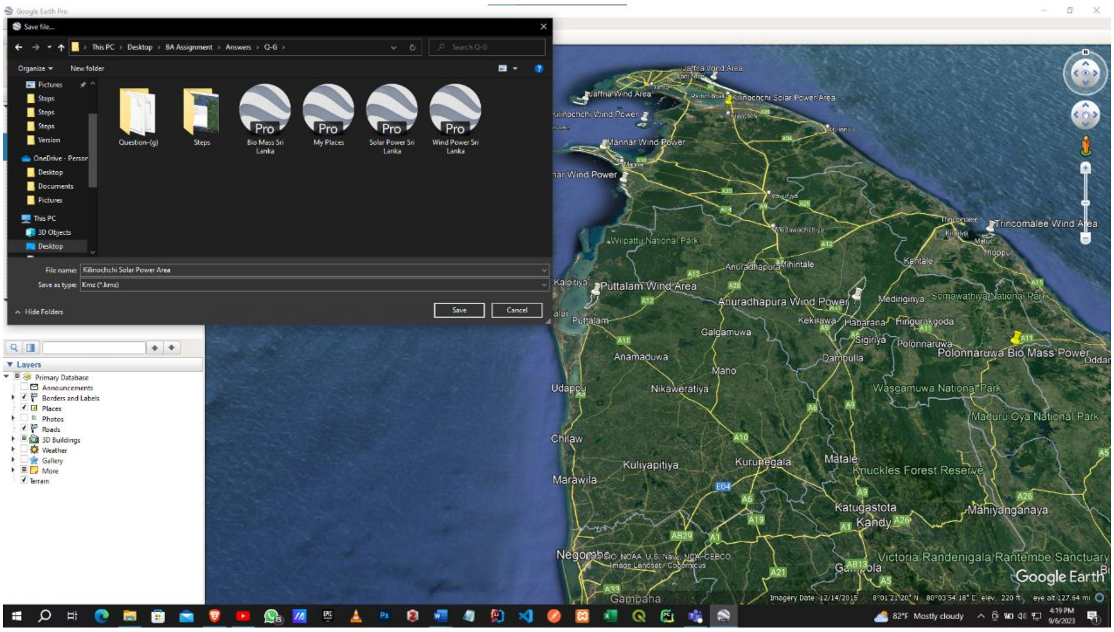
Add points and polygons using Google Earth Pro



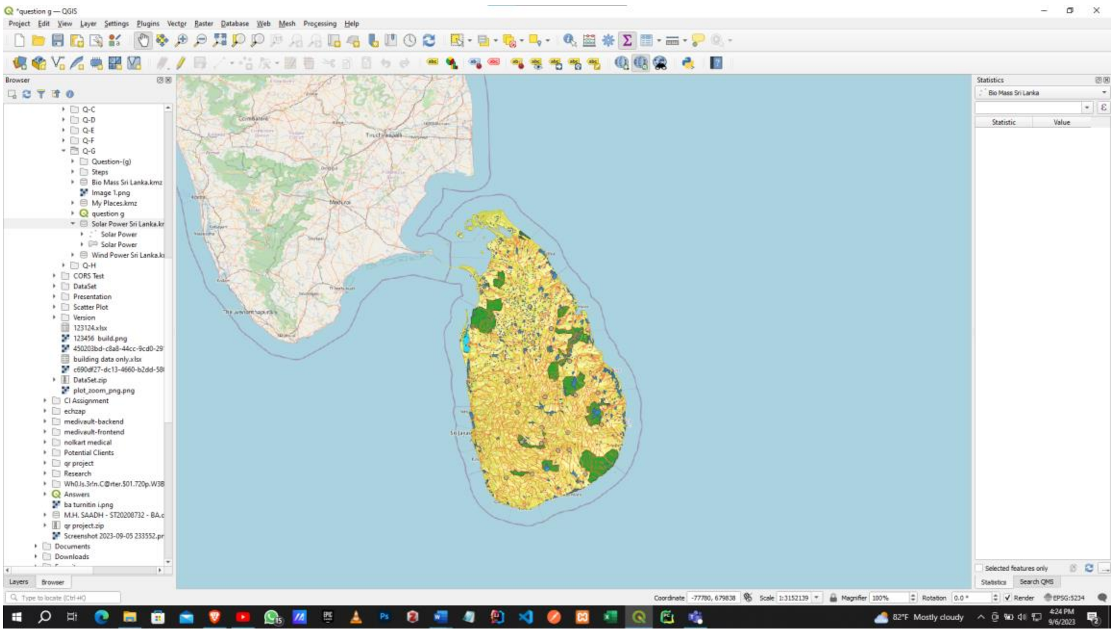
Add place details.



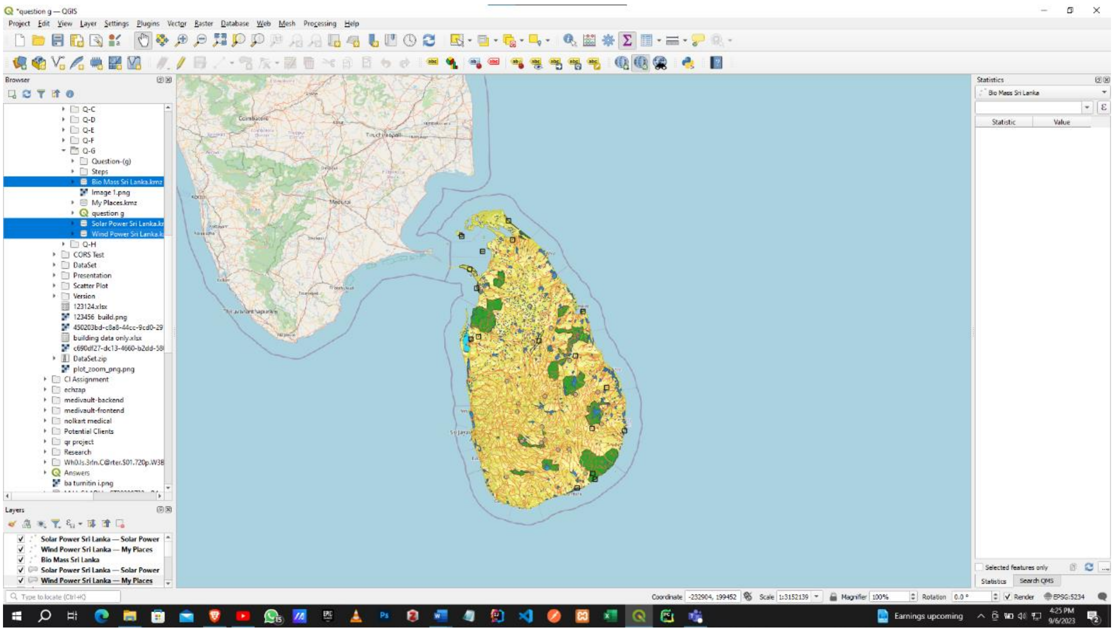
Save them as KML files



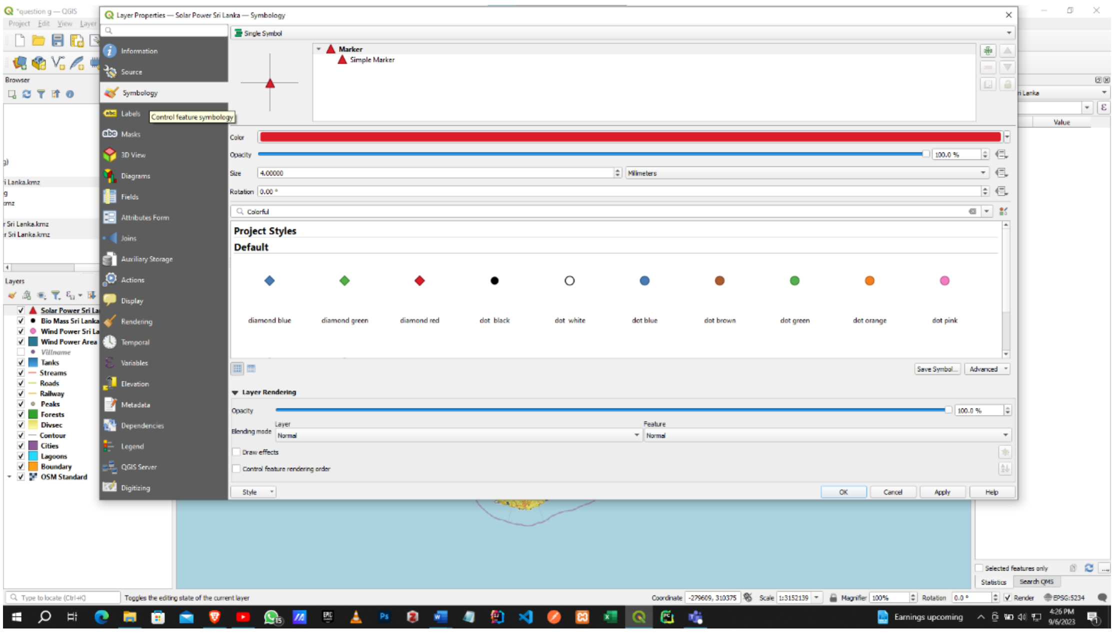
Open QGIS and added the shape files



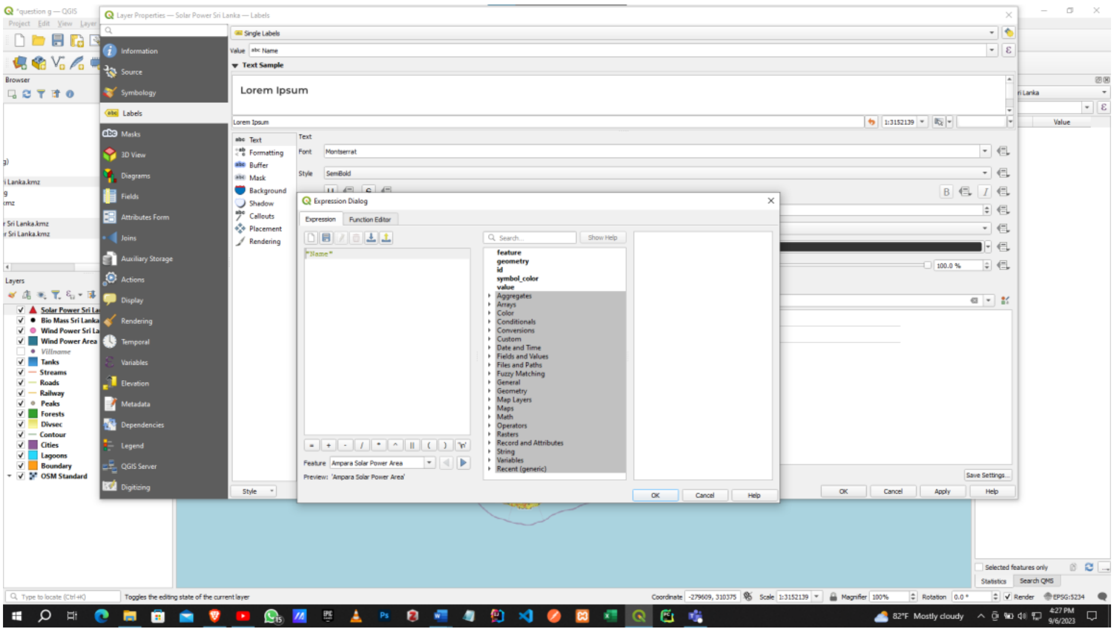
Add the KML Files



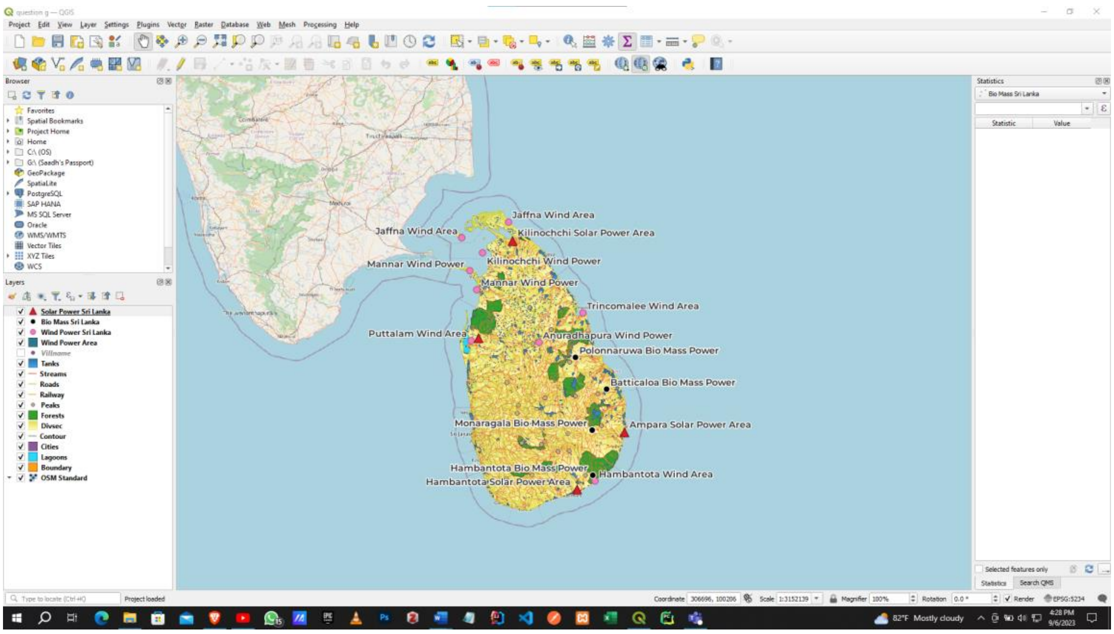
Add Symbology.



Add labels.

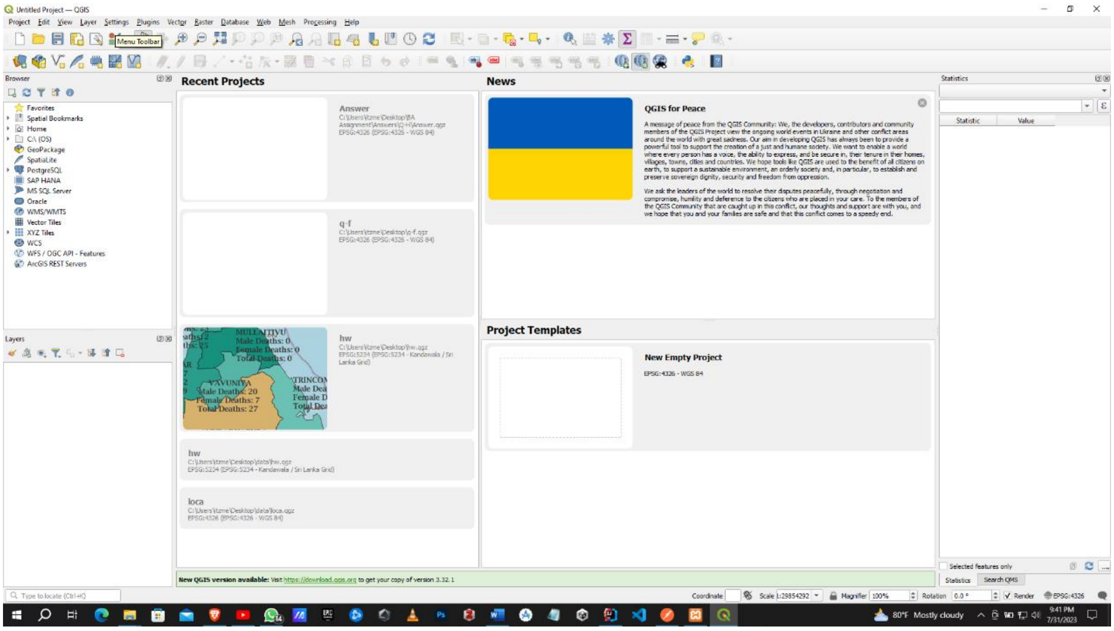


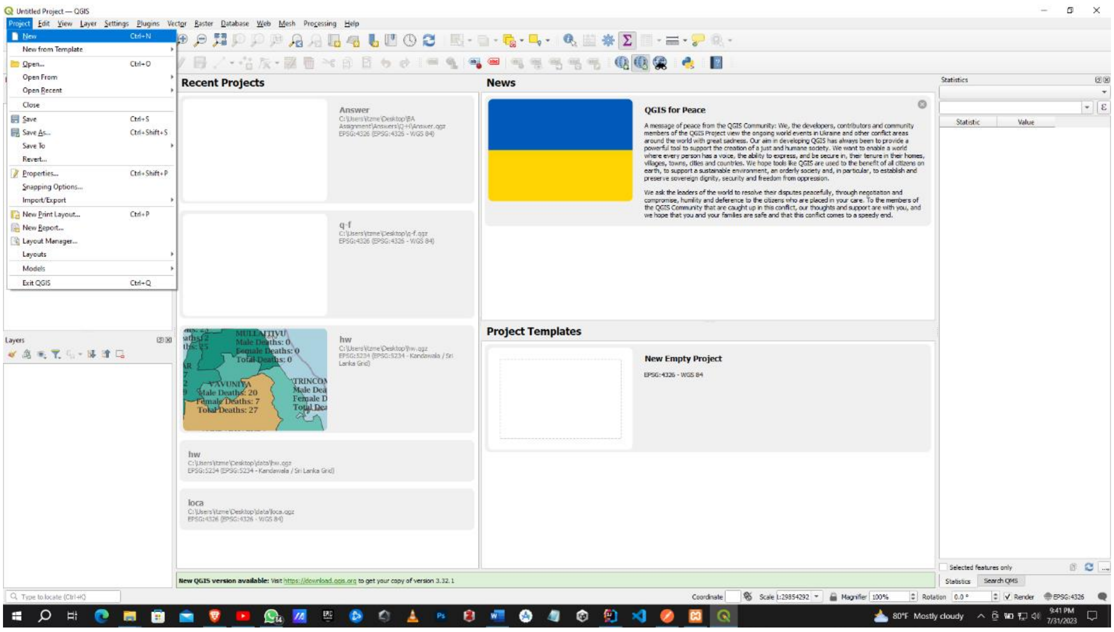
Final

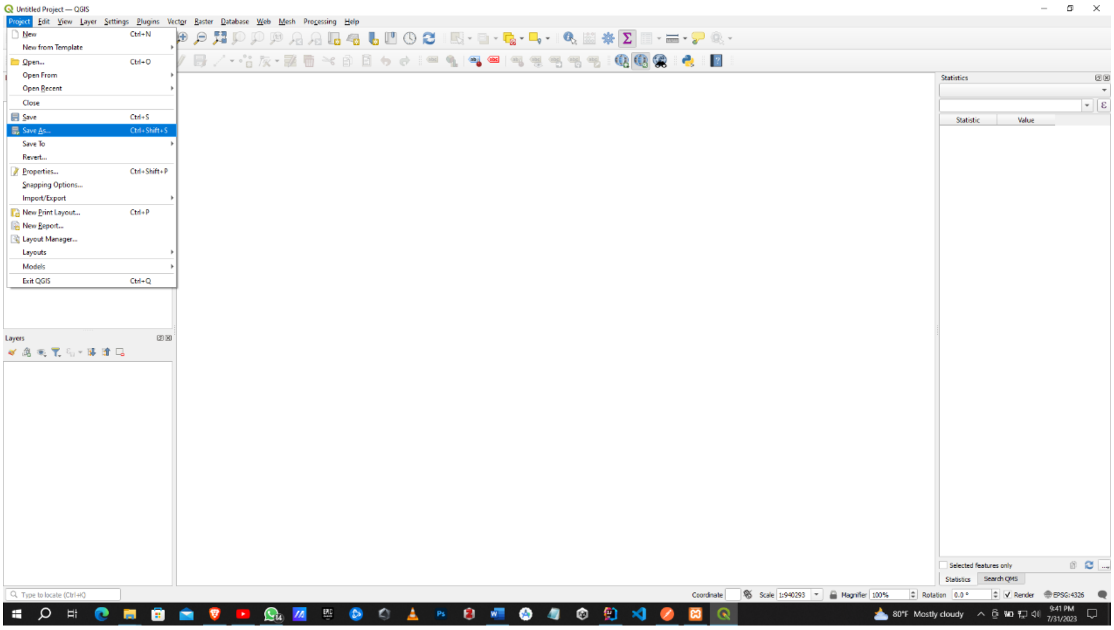


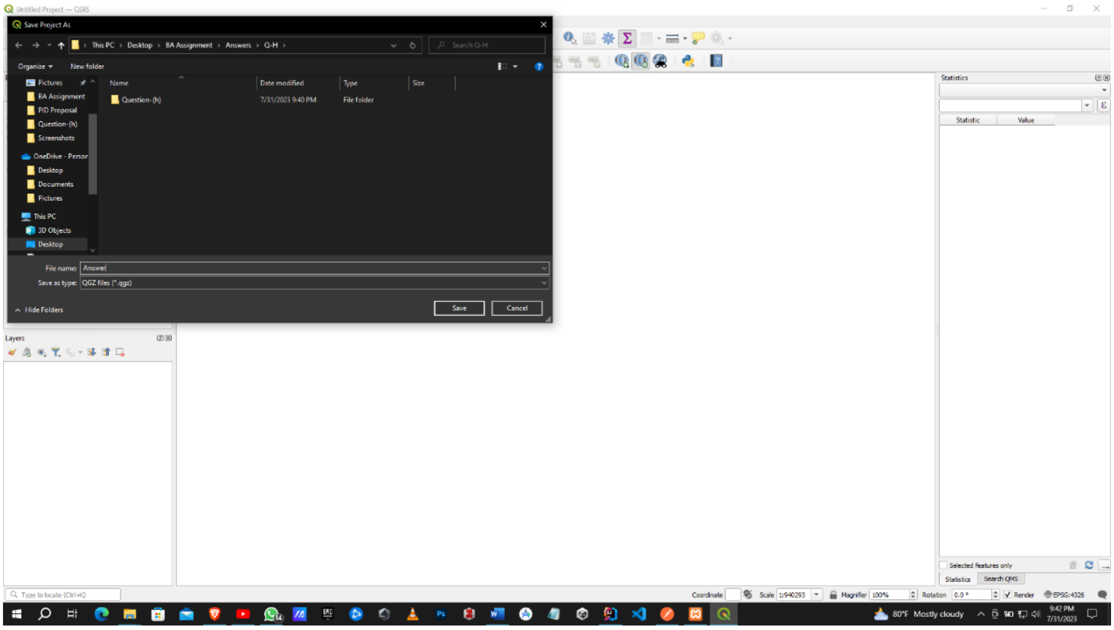
APPENDIX H

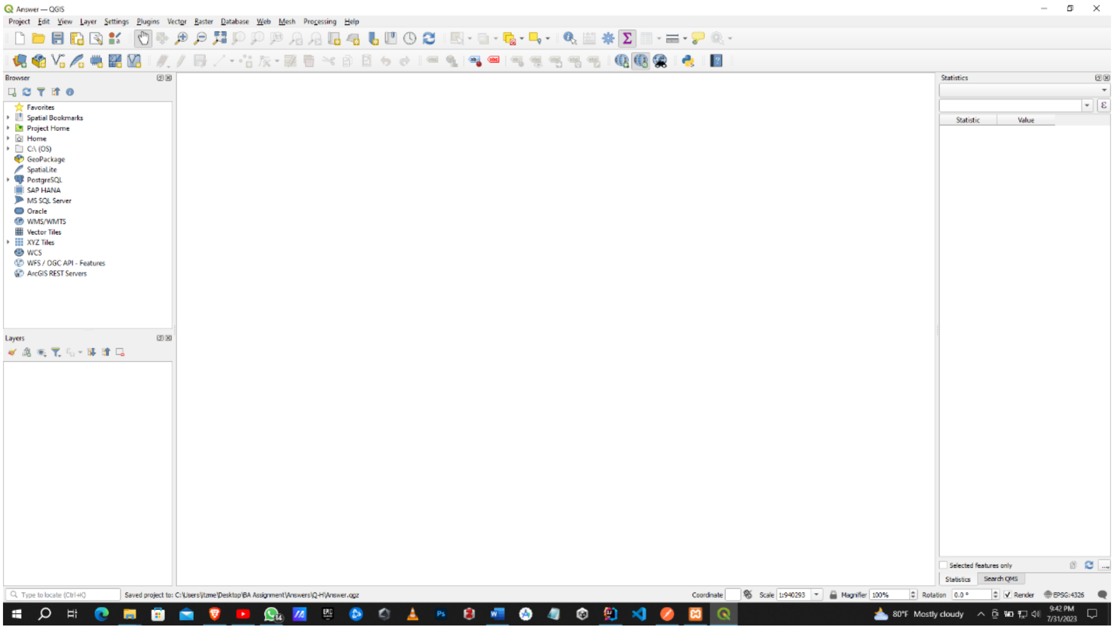
All steps related to question H are in order

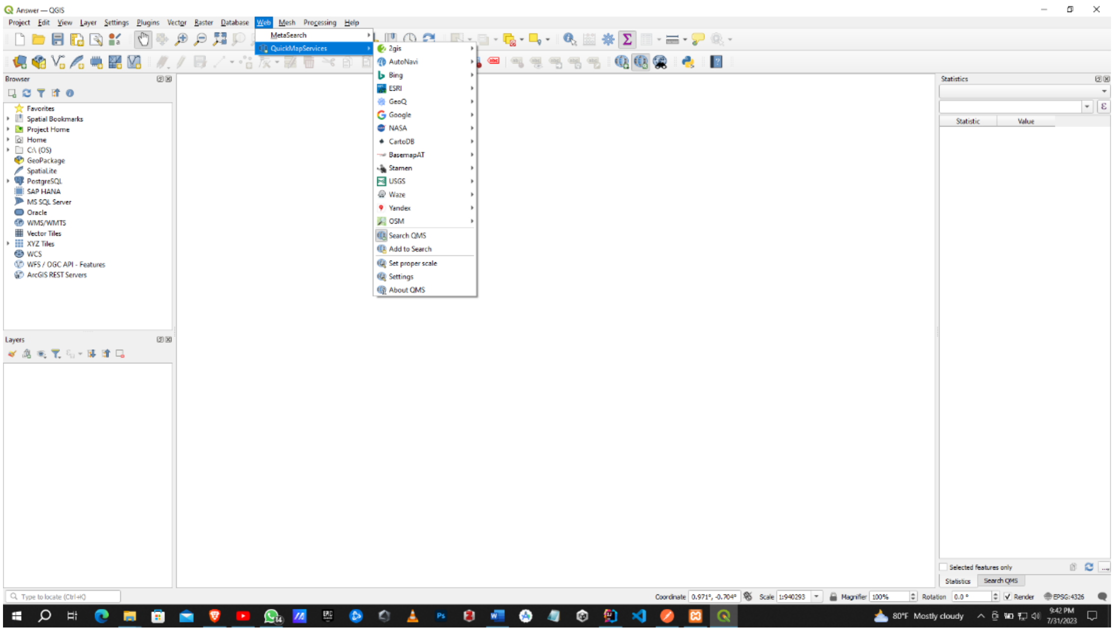


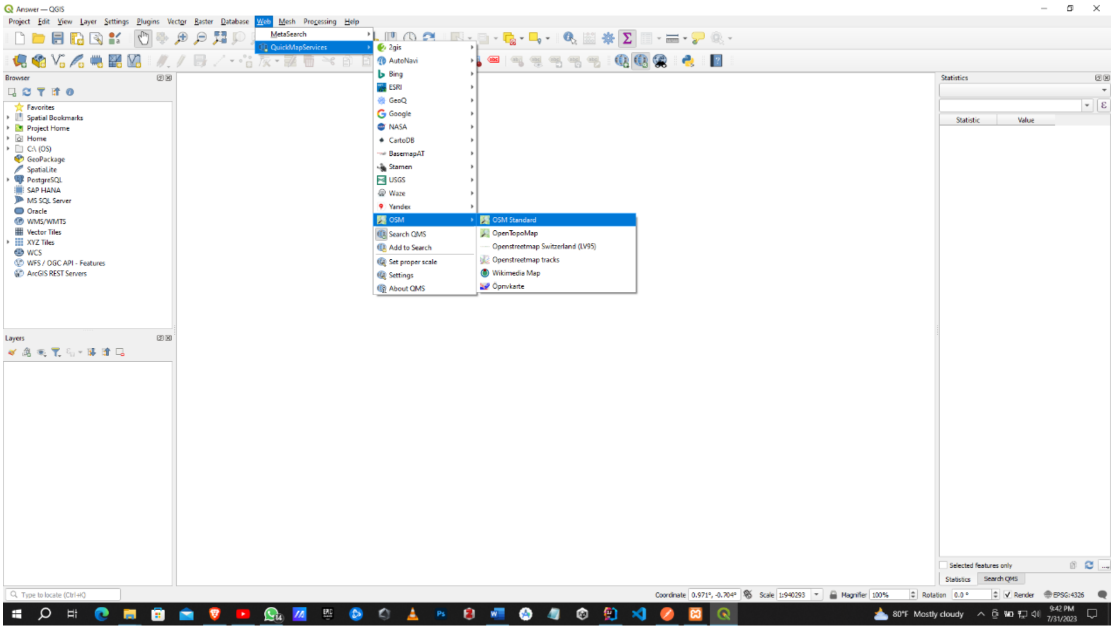


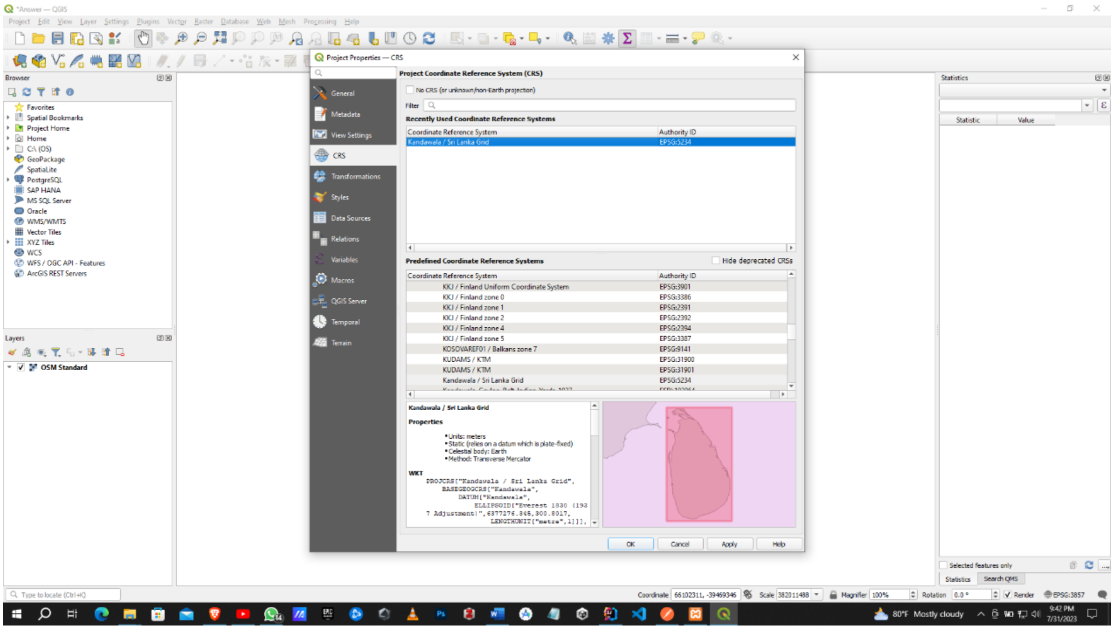

Untitled Project — QGIS
Project
Edit
View
Layer
Settings
Plugins
Vector
Raster
Database
Web
Mesh
Processing
Help
New
Ctrl+N
New from Template
Open...
Ctrl+O
Open From
Open Recent
Close
Save
Ctrl+S
Save As...
Ctrl+Shift+S
Save To
Revert...
Properties...
Ctrl+Shift+P
Snapping Options...
Import/Export
New Print Layout...
Ctrl+P
New Report...
Layout Manager...
Layouts
Models
Exit QGIS
Ctrl+Q
Layers
Statistics
Statistic
Value
Selected features only
Search QMS
Coordinate
Scale
Magnifier
Rotation
Render
EPSG:4326
80°F Mostly cloudy

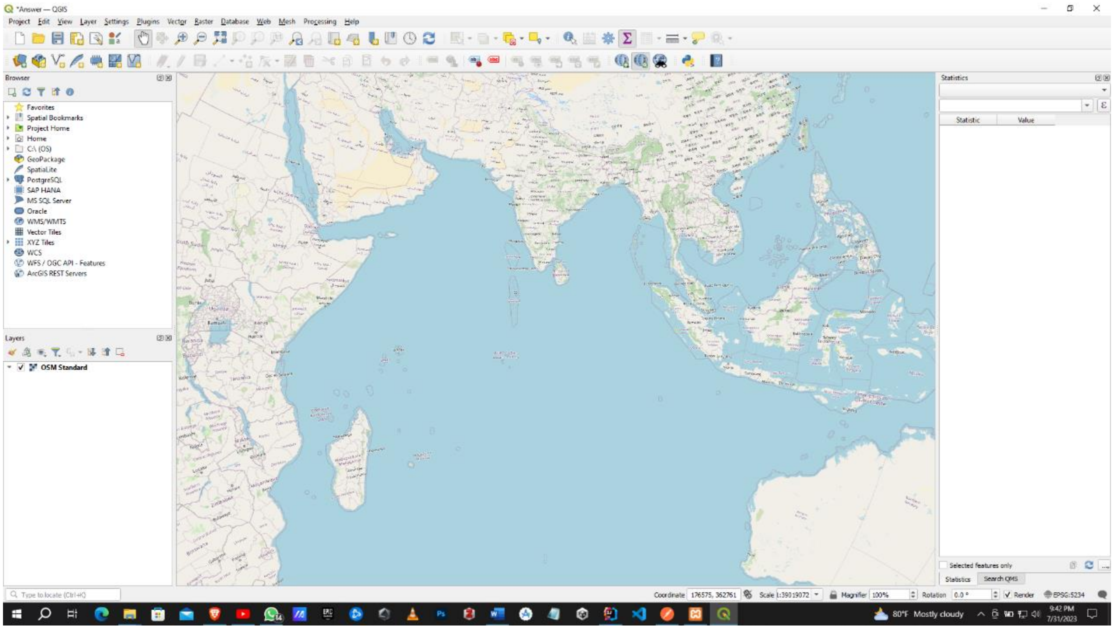

Untitled Project — QGIS
Save Project As
This PC
Desktop
BA Assignment
Answers
Q-H
Organize
New folder
Pictures
BA Assignment
PID Proposal
Question-(h)
Screenshots
OneDrive - Person
Desktop
Documents
Pictures
This PC
3D Objects
Desktop
Name
Date modified
Type
Size
Question-(h)
7/31/2023 9:40 PM
File folder
File name:
Answer
Save as type:
QGZ files (*.qgz)
Hide Folders
Save
Cancel
Layers
Statistics
Statistic
Value
Selected features only
Search QMS
Coordinate
Scale
Magnifier
Rotation
Render
EPSG:4326
80°F Mostly cloudy

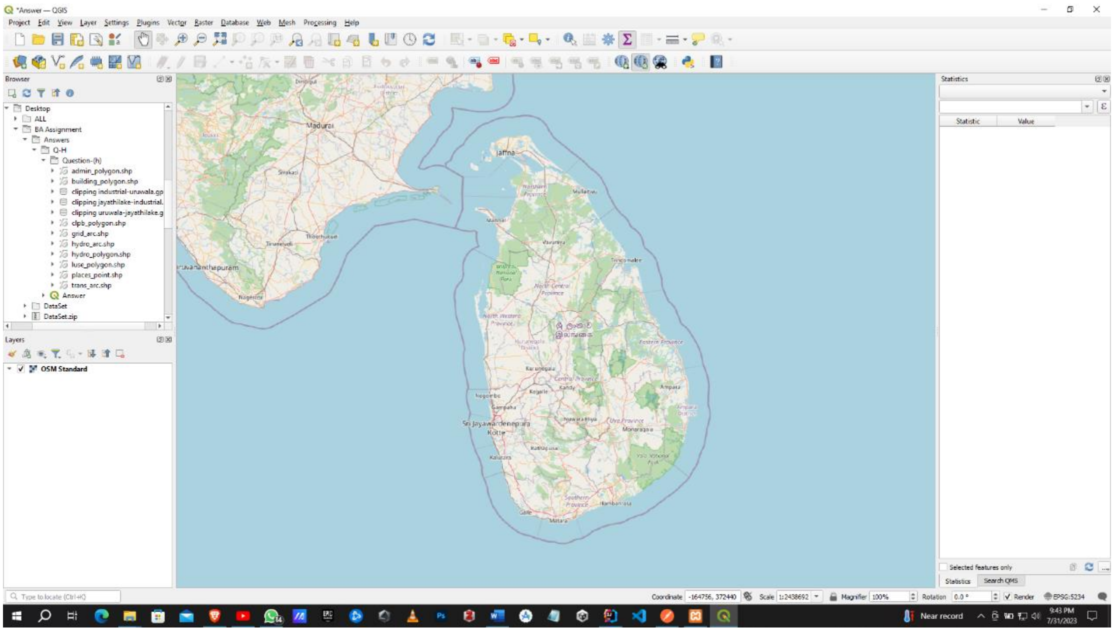


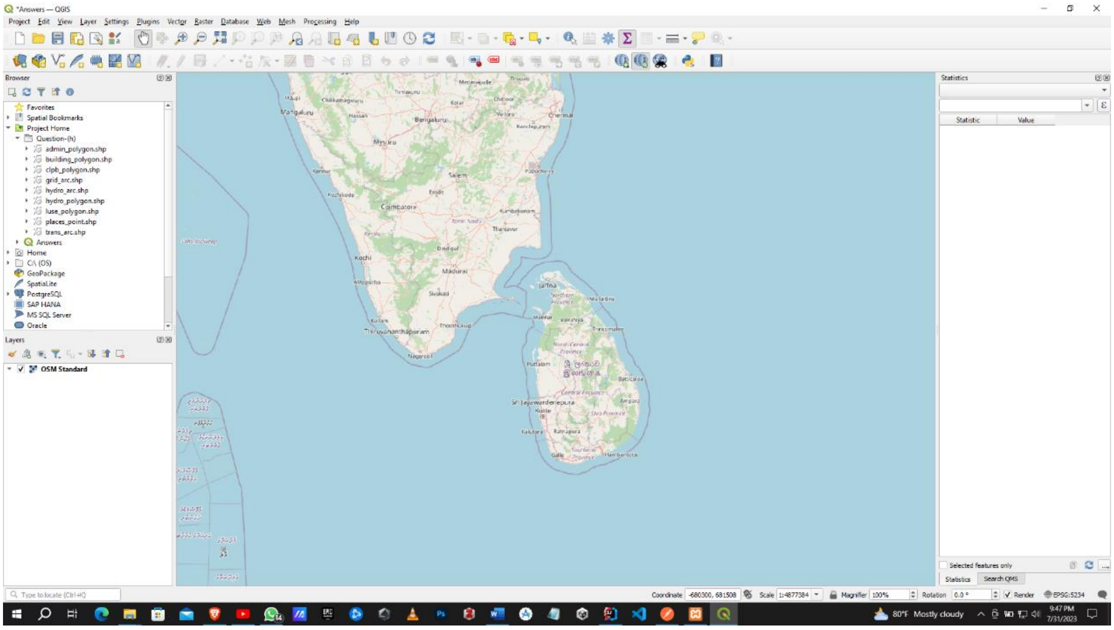


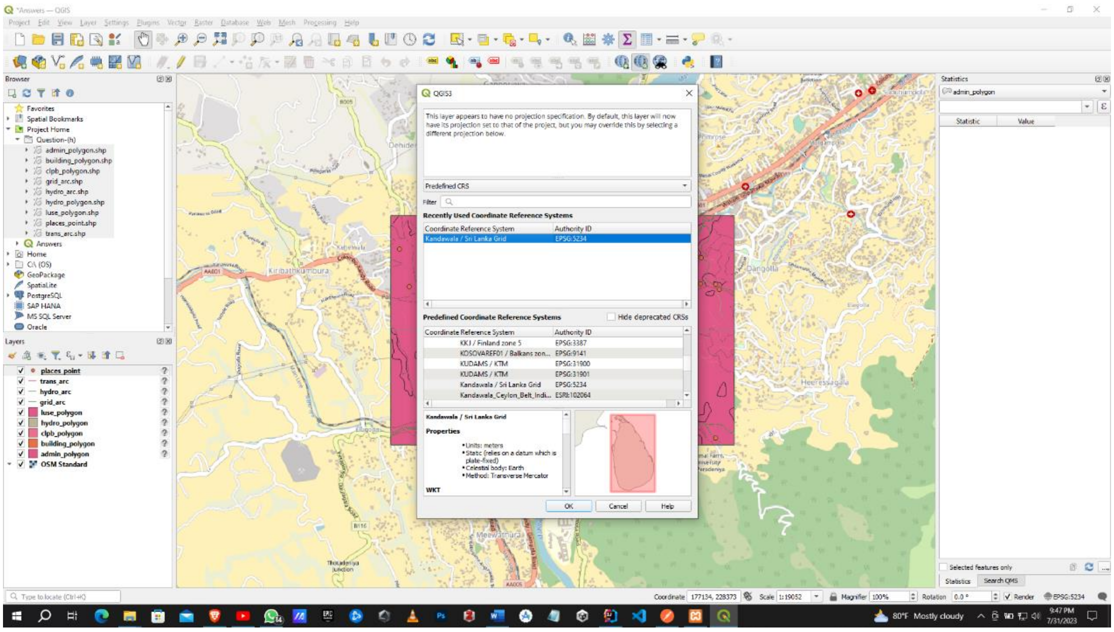


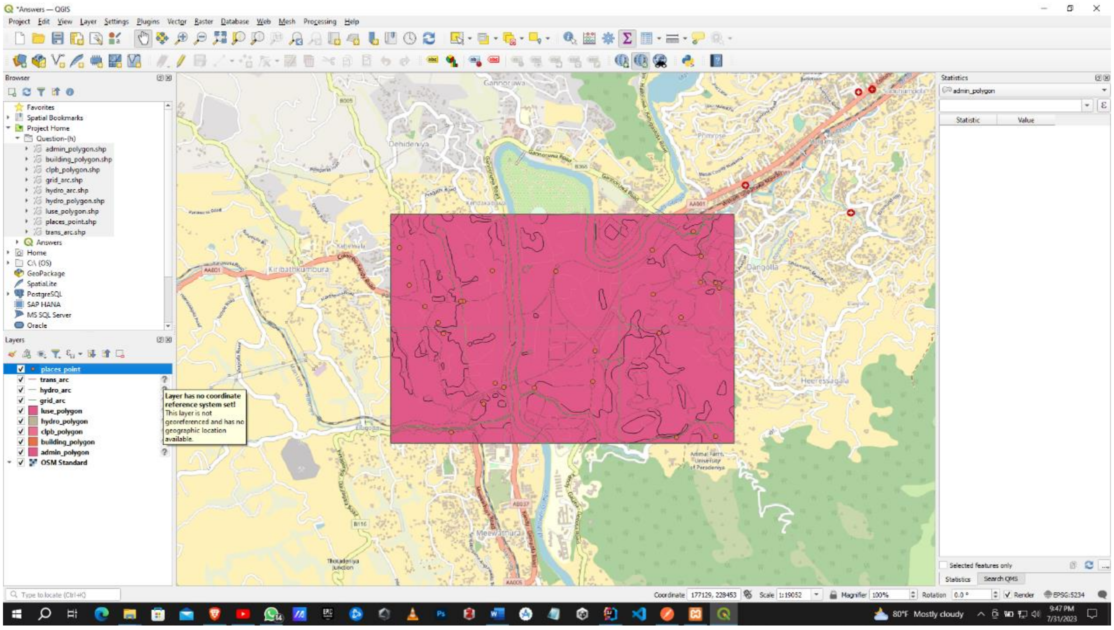


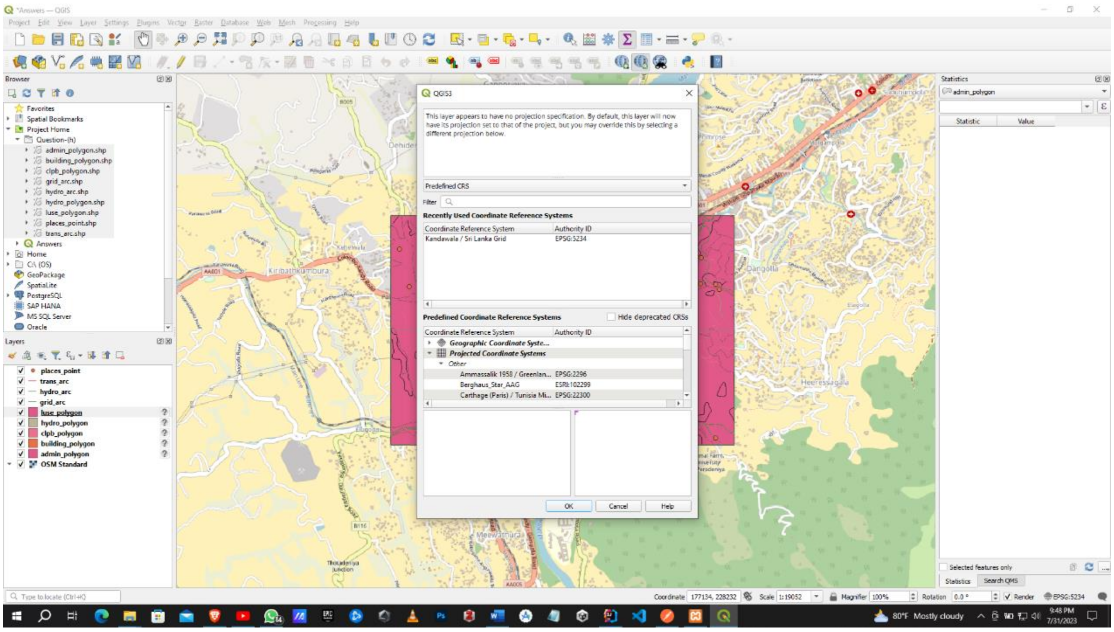


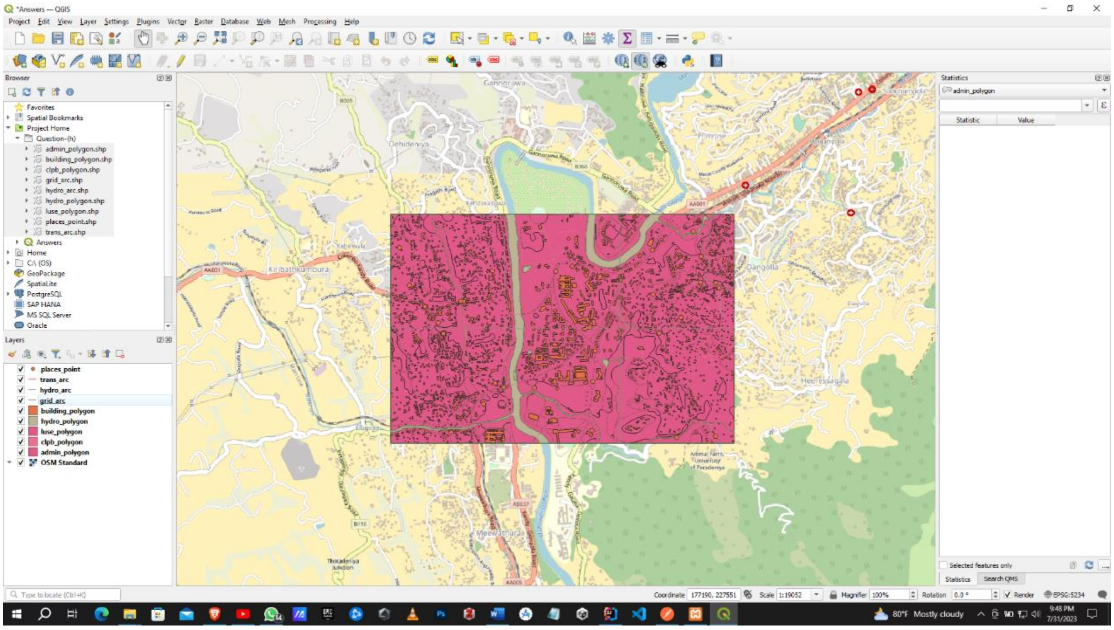


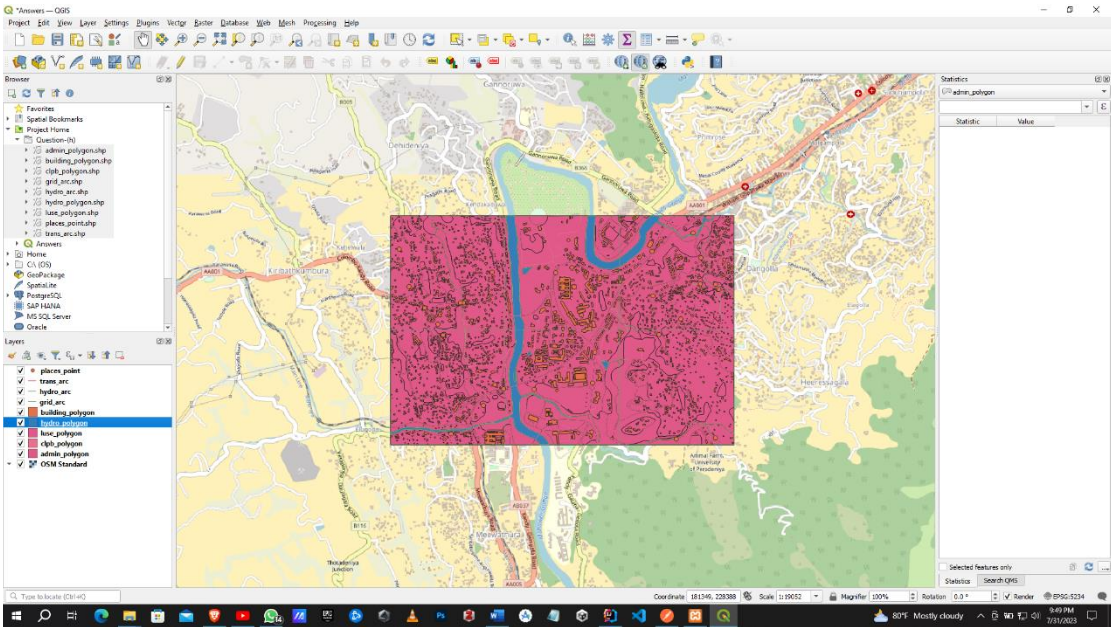


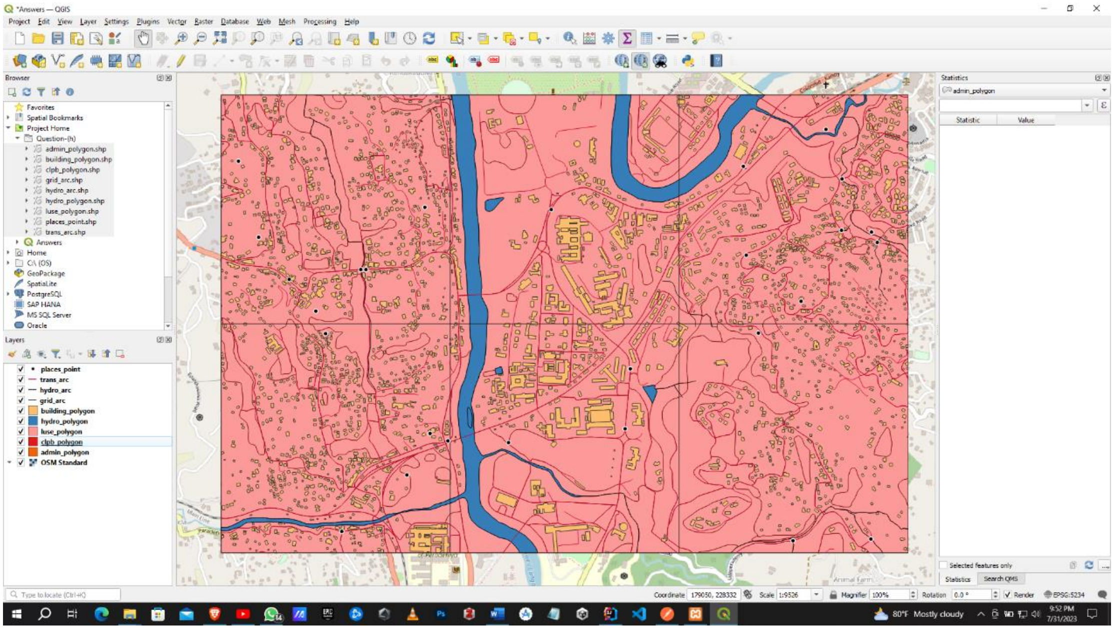


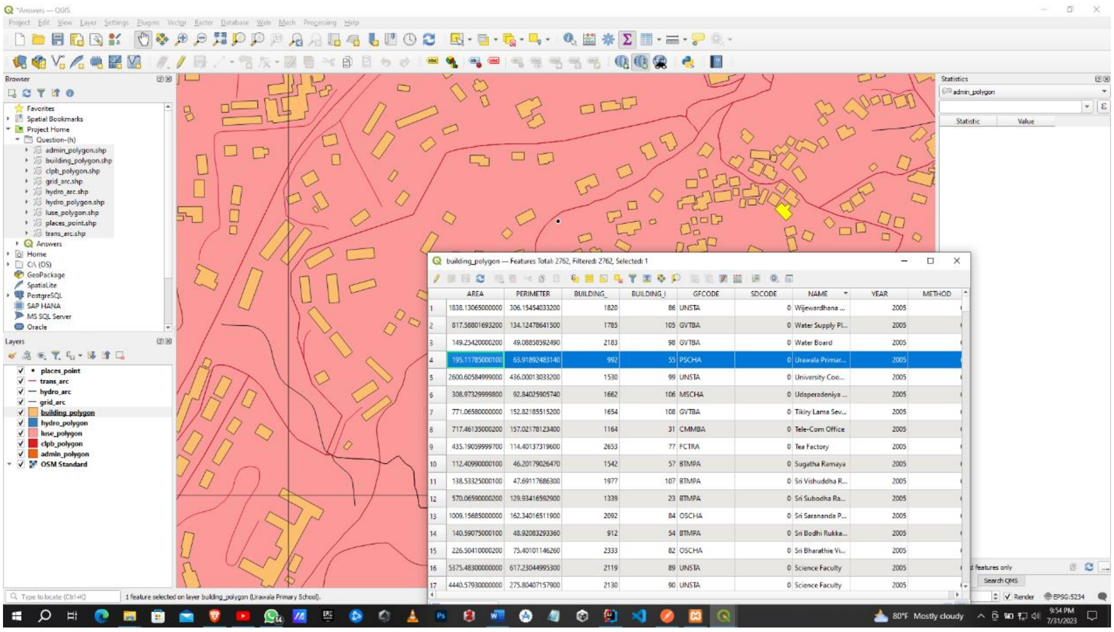

*Answers — QGIS
Layer has no coordinate reference system set! This layer is not georeferenced and has no geographic location available.
places_point
trans_arc
hydro_arc
grid_arc
luse_polygon
hydro_polygon
clpb_polygon
building_polygon
admin_polygon
OSM Standard

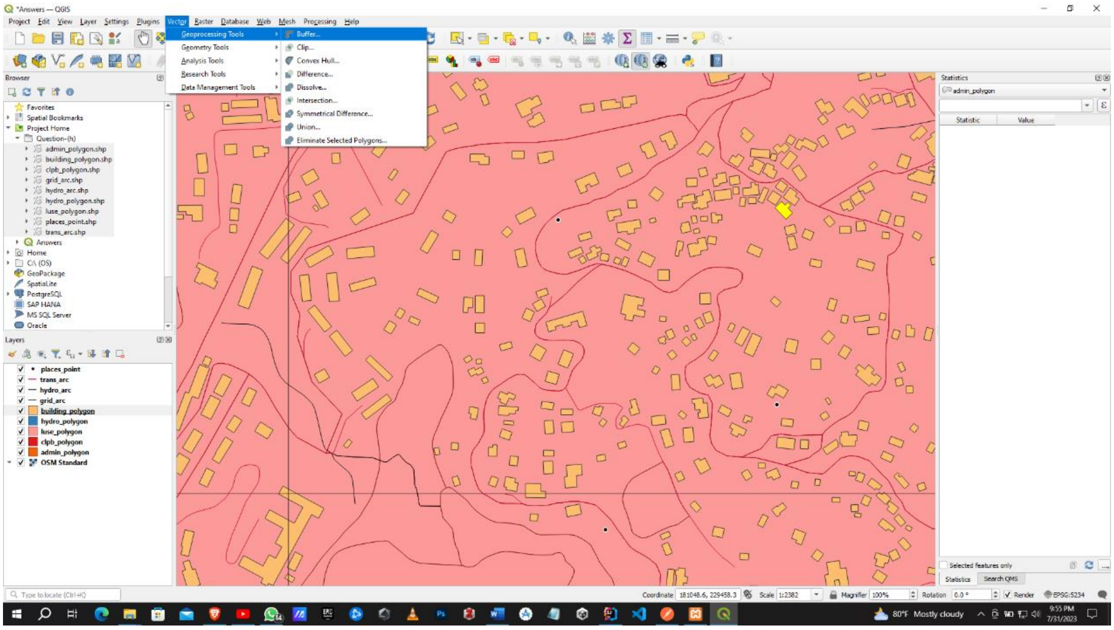

*Answers — QGIS
QGIS3
This layer appears to have no projection specification. By default, this layer will now have its projection set to that of the project, but you may override this by selecting a different projection below.
Predefined CRS
Recently Used Coordinate Reference Systems
Kandawala / Sri Lanka Grid
EPSG:5234
Predefined Coordinate Reference Systems
OK
Cancel
Help

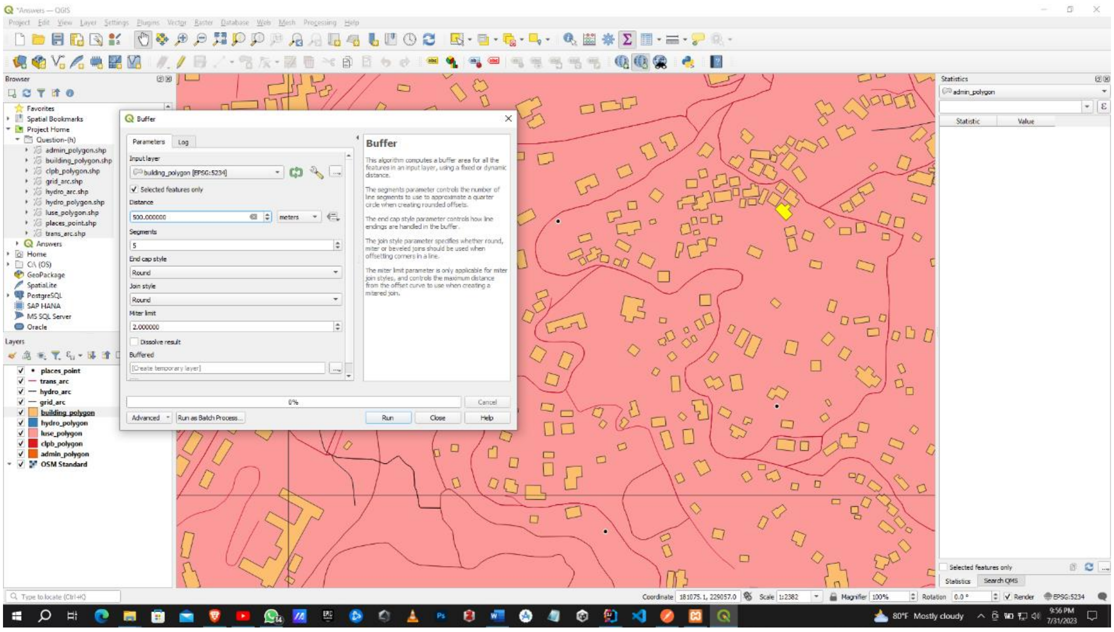


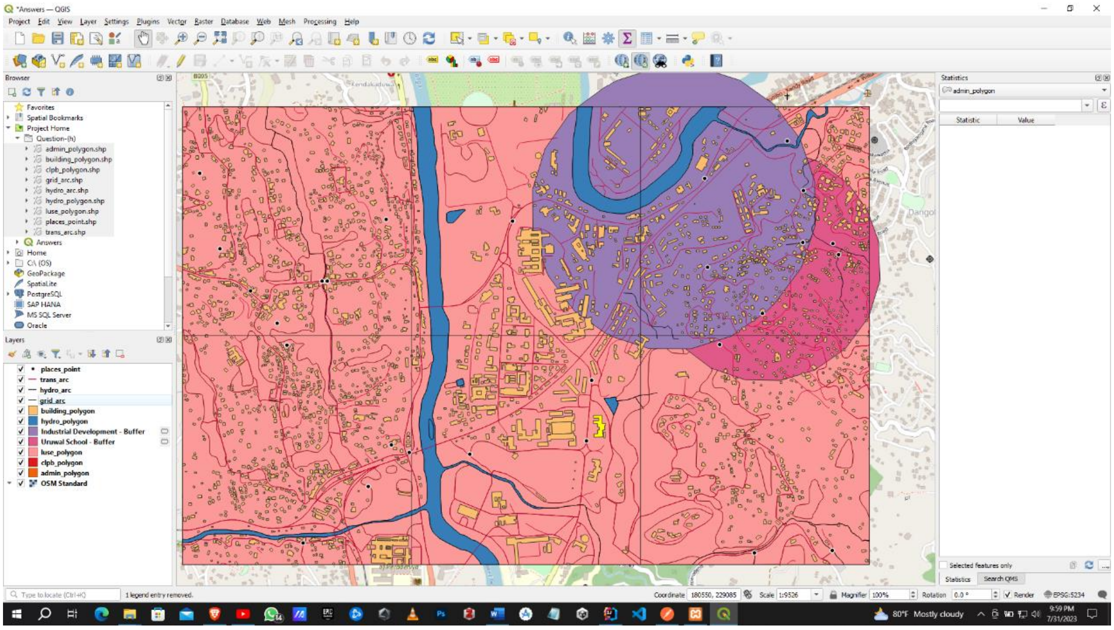


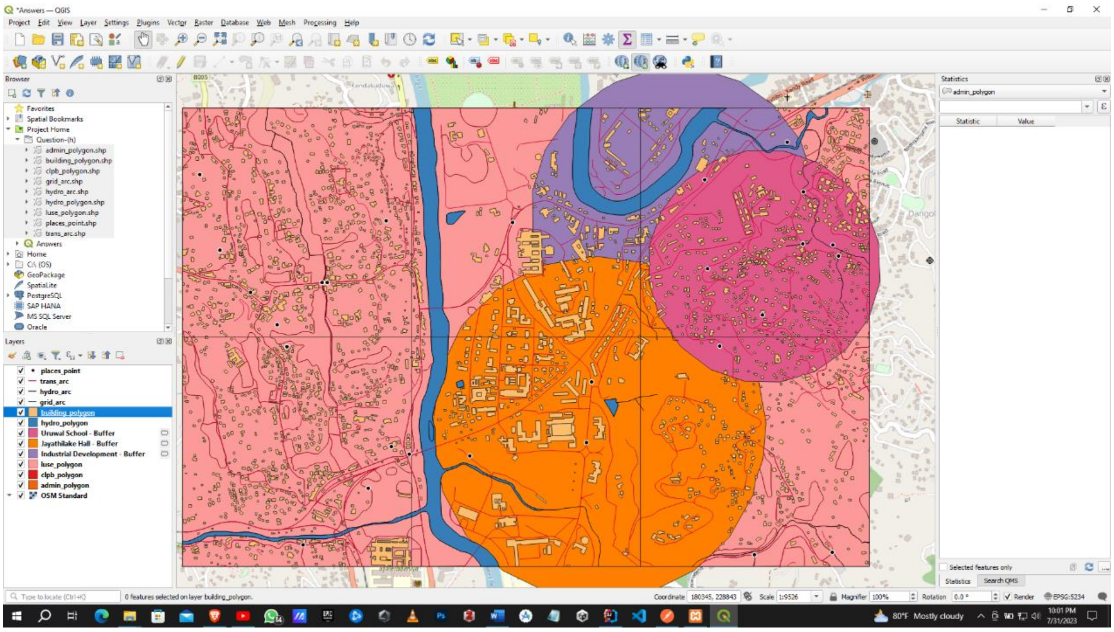

*Answers — QGIS
Browser
Favorites
Spatial Bookmarks
Project Home
Question-(h)
admin_polygon.shp
building_polygon.shp
clpb_polygon.shp
grid_arc.shp
hydro_arc.shp
hydro_polygon.shp
luse_polygon.shp
places_point.shp
trans_arc.shp
Answers
Home
GeoPackage
SpatiaLite
PostgreSQL
SAP HANA
MS SQL Server
Oracle
Layers
places_point
trans_arc
hydro_arc
grid_arc
building_polygon
hydro_polygon
luse_polygon
clpb_polygon
admin_polygon
OSM Standard
Statistics
admin_polygon
Statistic
Value
Selected features only
80°F Mostly cloudy
9:52 PM
7/31/2023

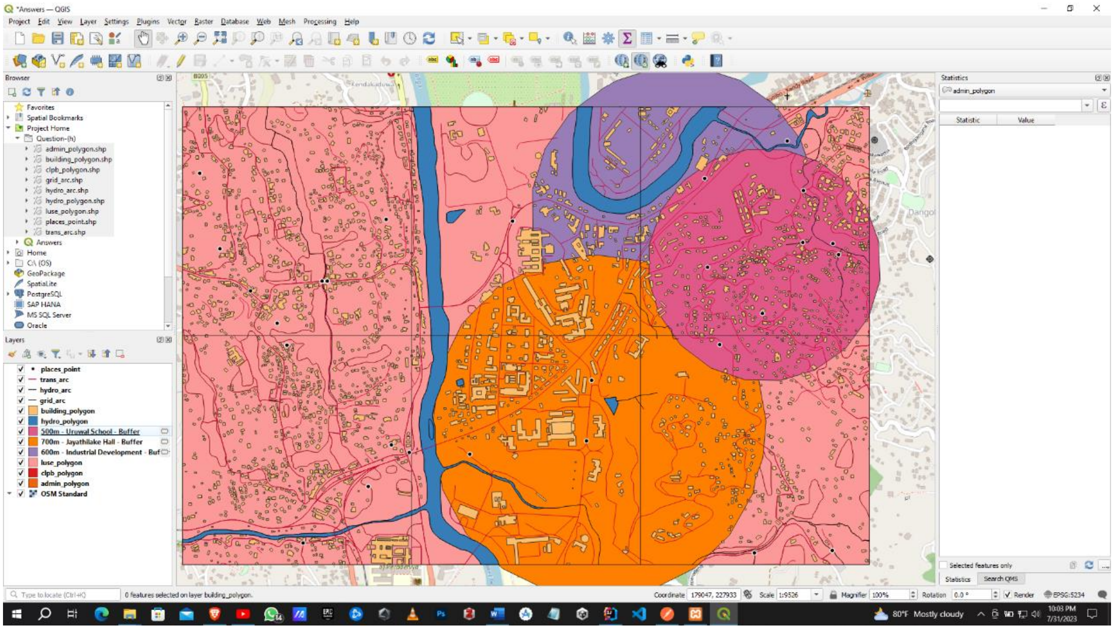

*Answers — QGIS
Layers
places_point
trans_arc
hydro_arc
grid_arc
building_polygon
hydro_polygon
luse_polygon
clpb_polygon
admin_polygon
OSM Standard
building_polygon — Features Total: 2762, Filtered: 2762, Selected: 1
AREA
PERIMETER
GFCODE
SDCODE
NAME
YEAR
METHOD
1 feature selected on layer building_polygon (Urawala Primary School).
80°F Mostly cloudy
9:54 PM
7/31/2023

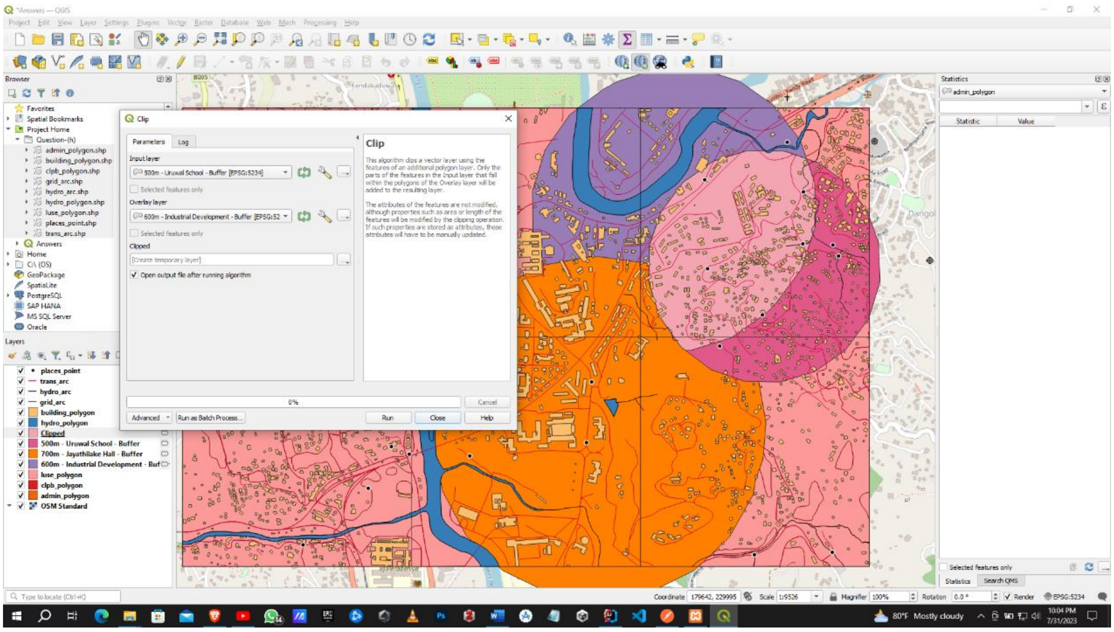


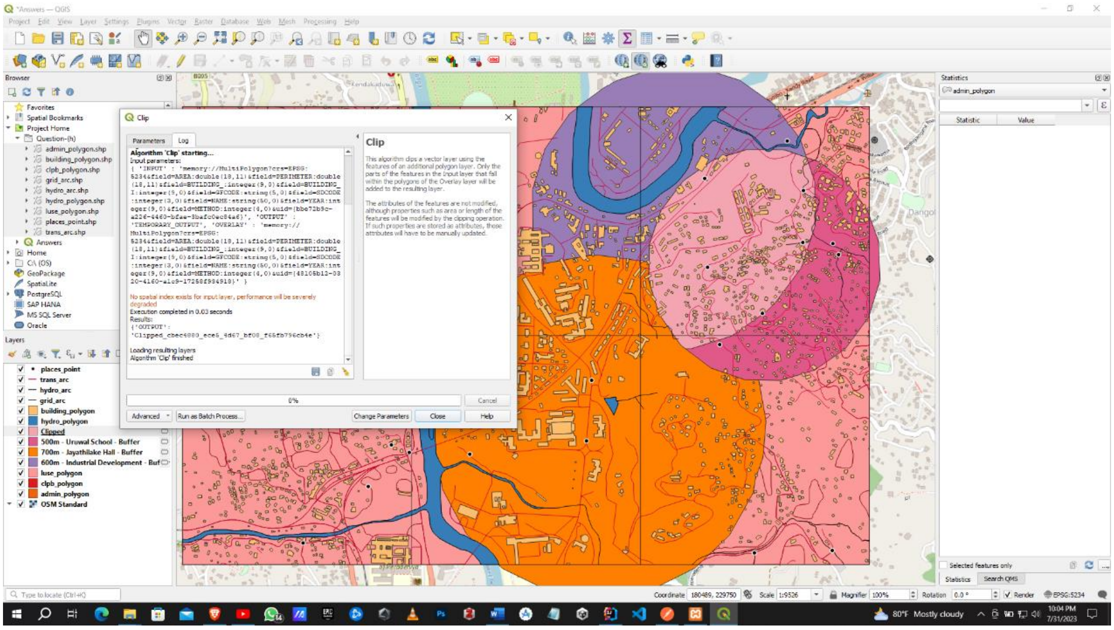


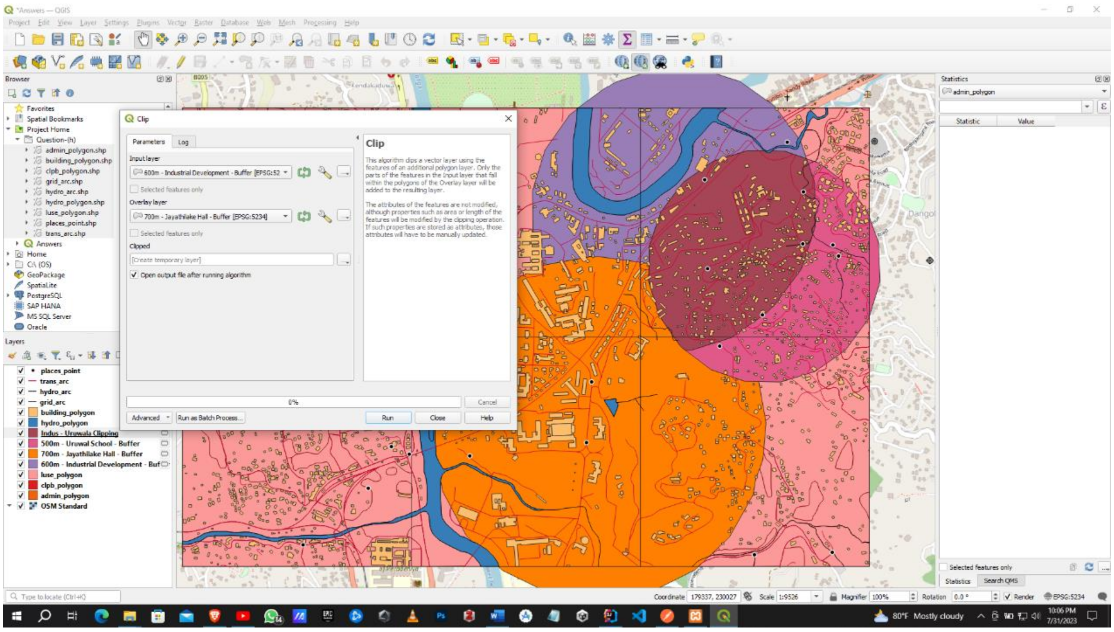


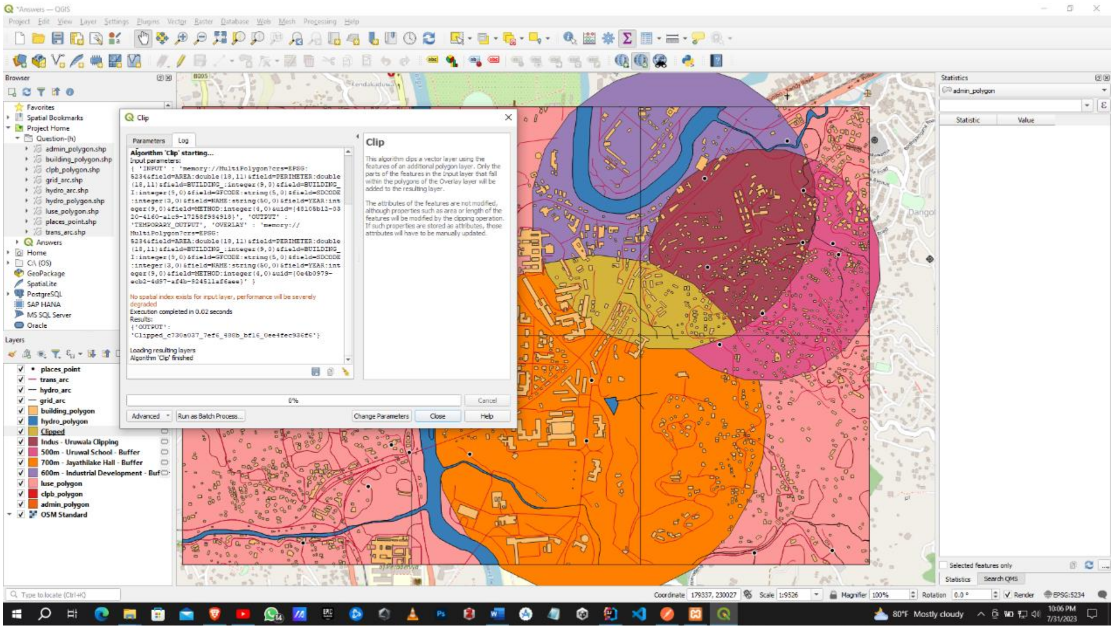


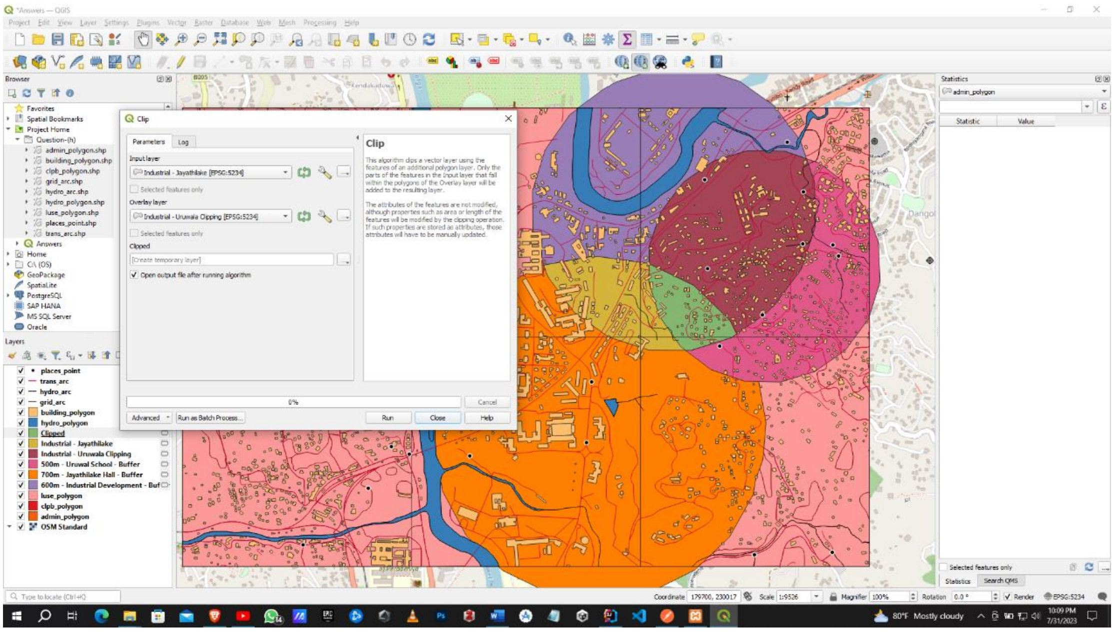


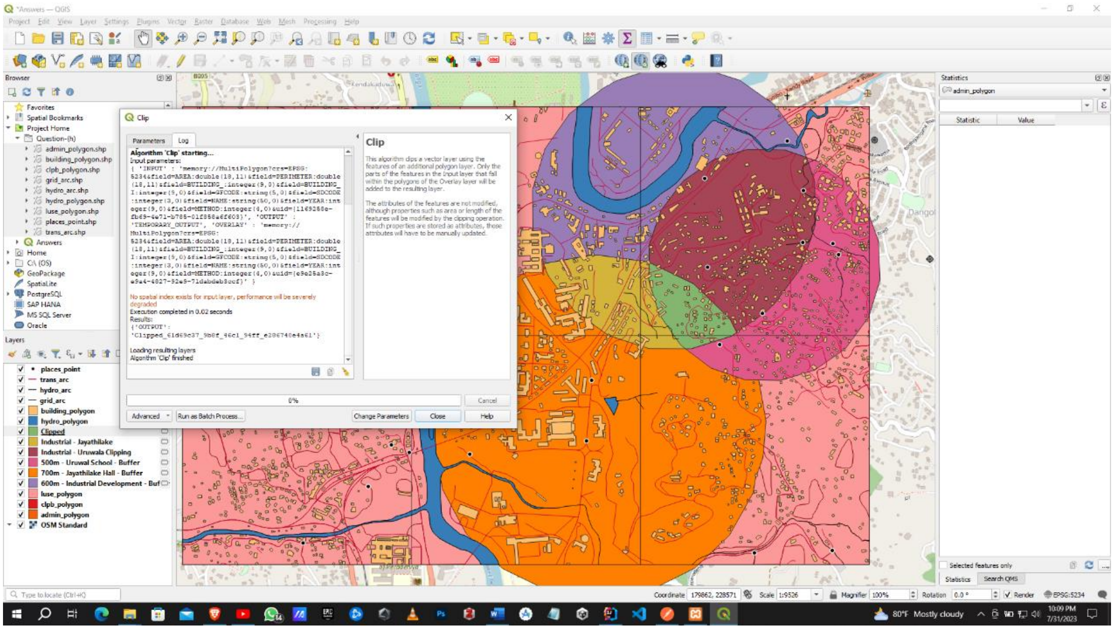


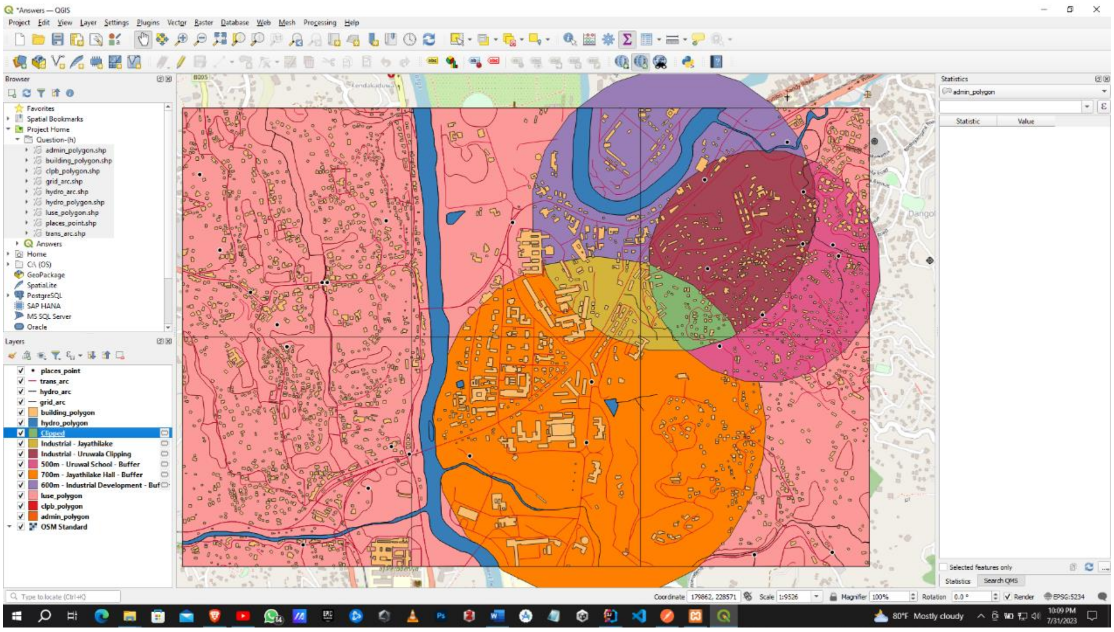

Clip
Parameters
Log
Advanced
Run as Batch Process...
Change Parameters
Close
Help
Cancel

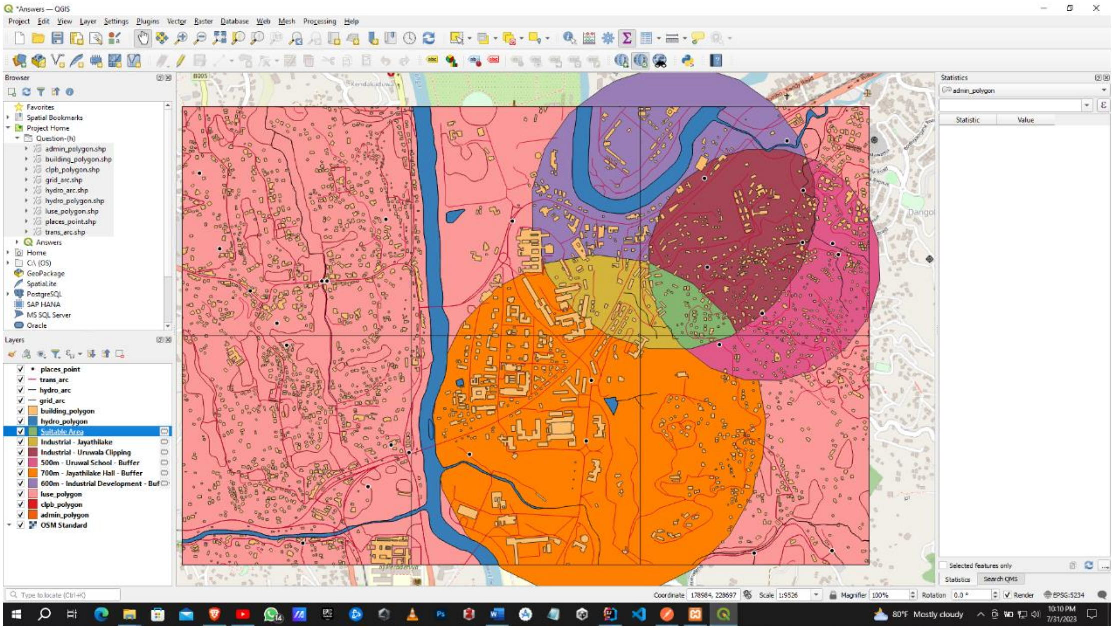

Clip
Parameters
Log
Input layer
Overlay layer
Clipped
Open output file after running algorithm
Advanced
Run as Batch Process...
Run
Close
Help
Cancel

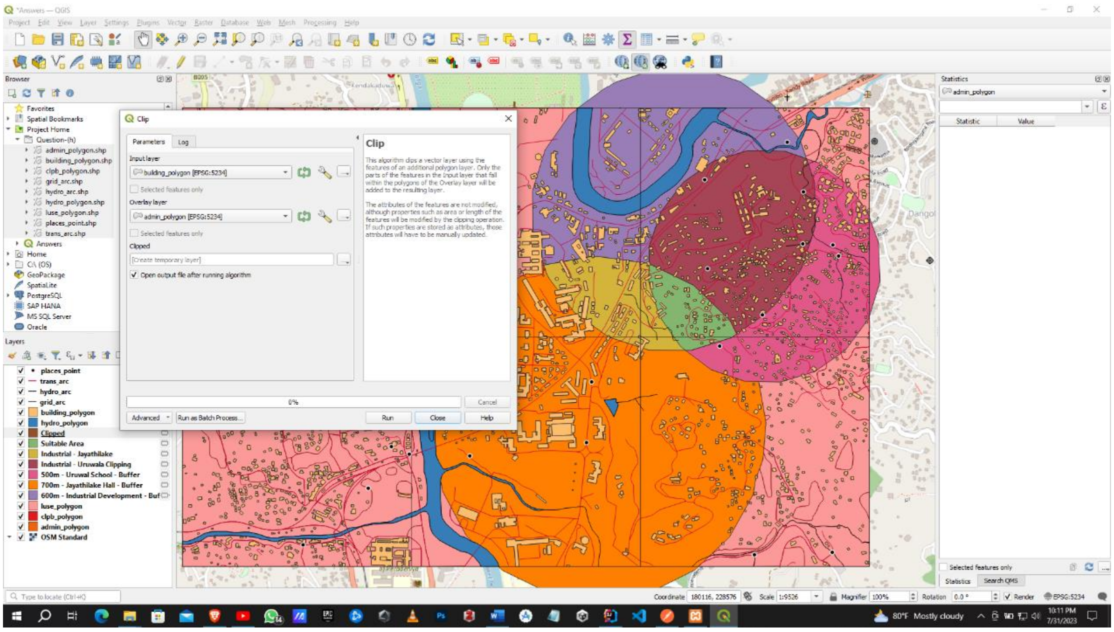


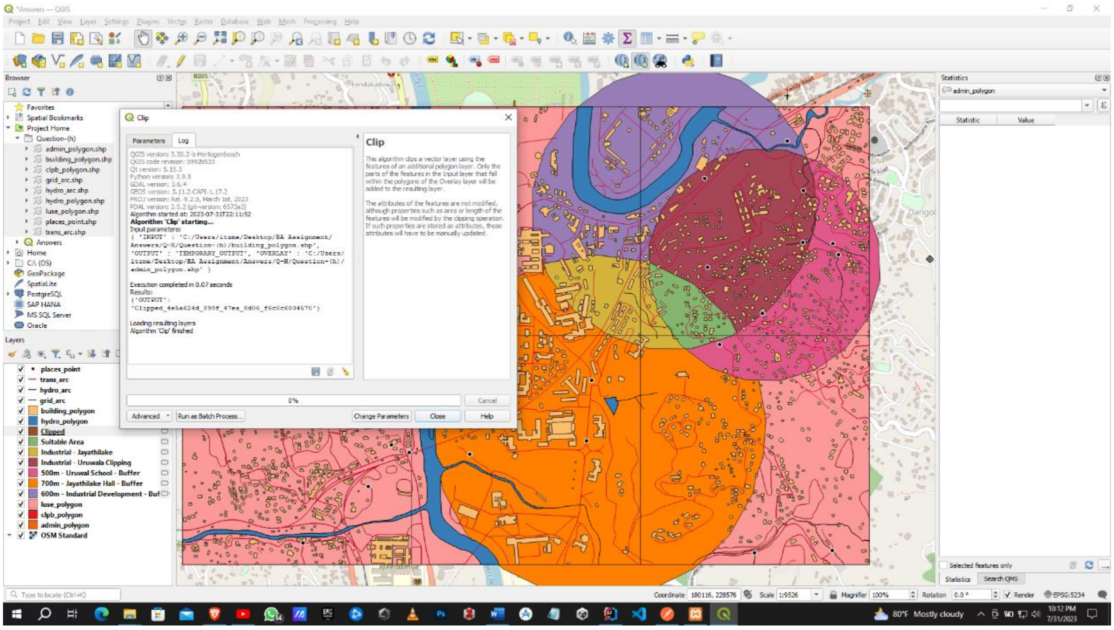


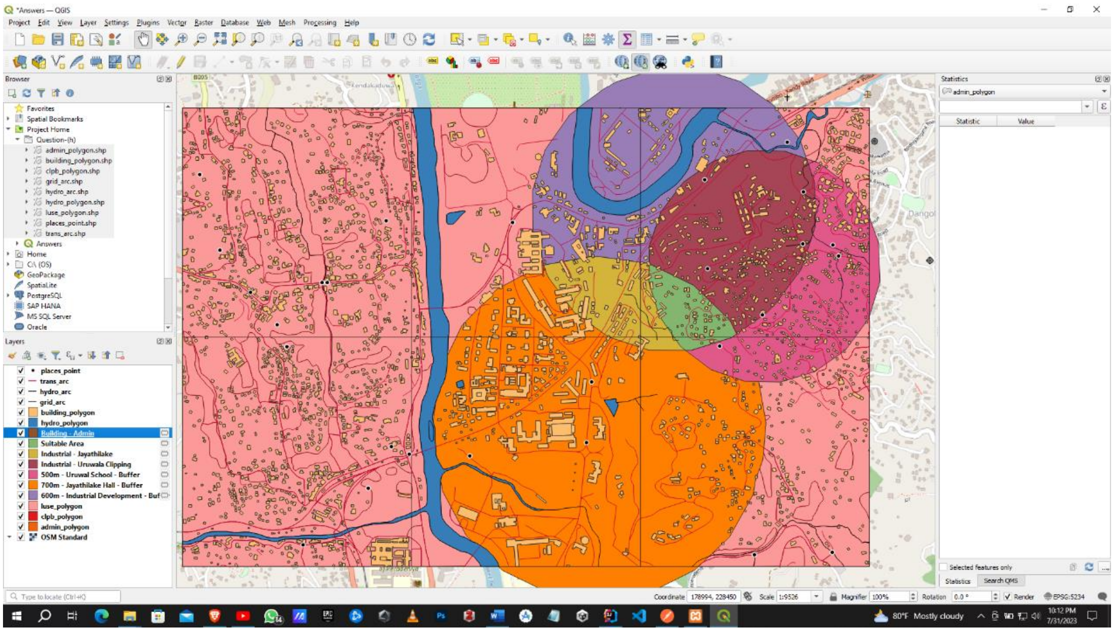


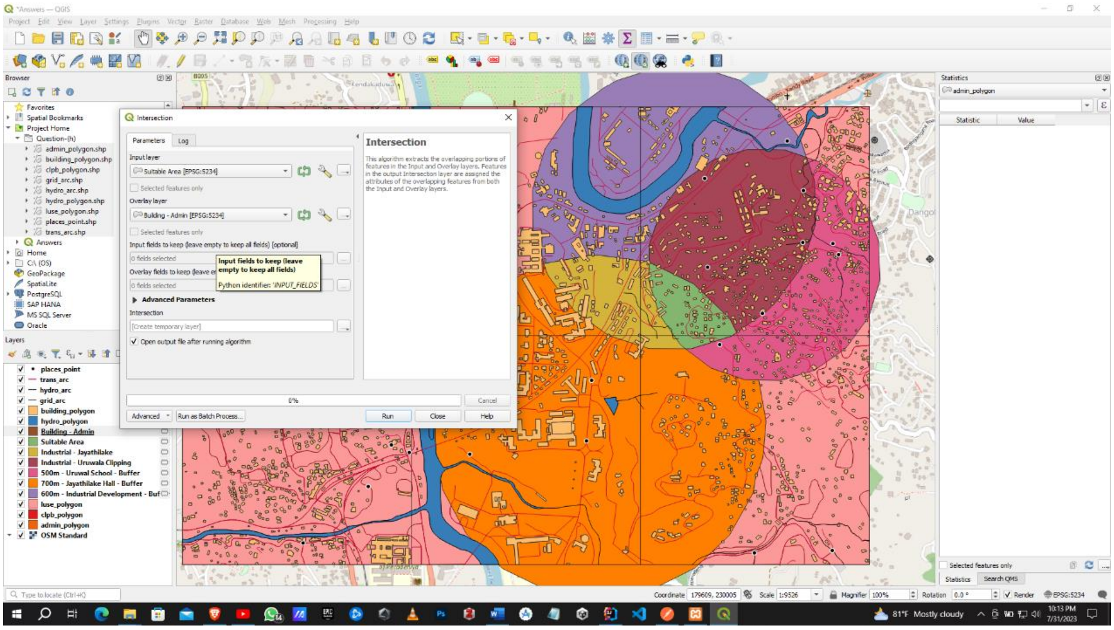


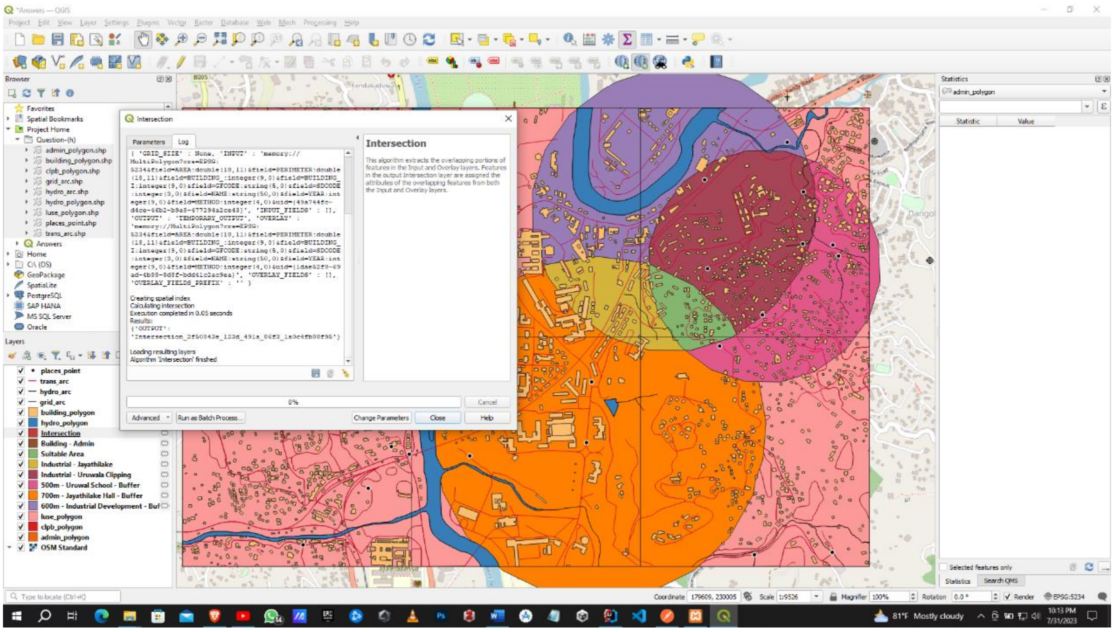


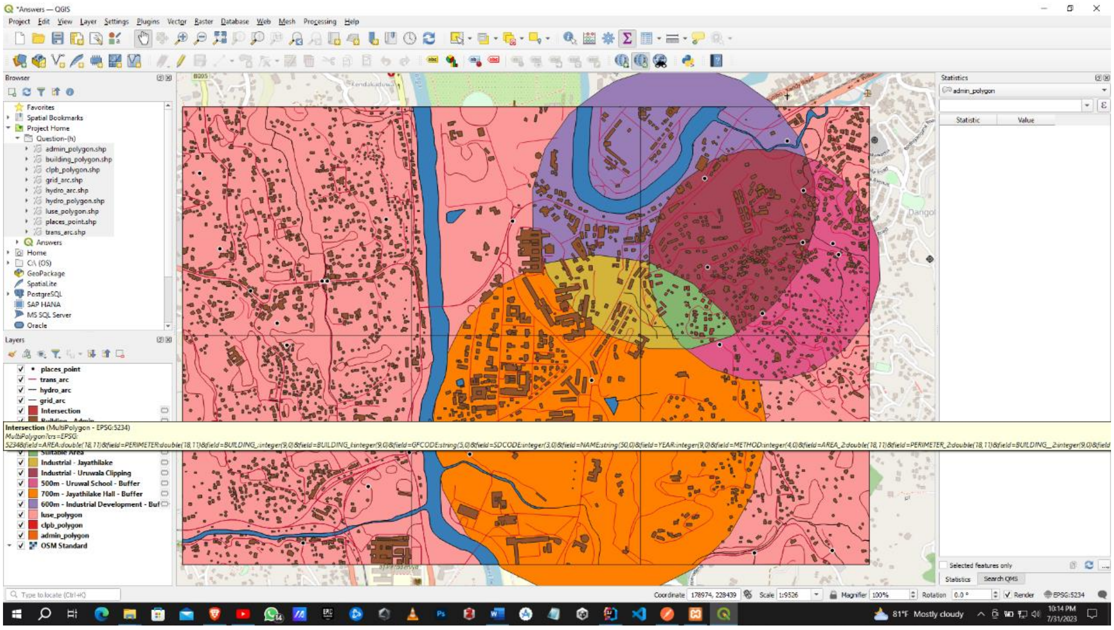


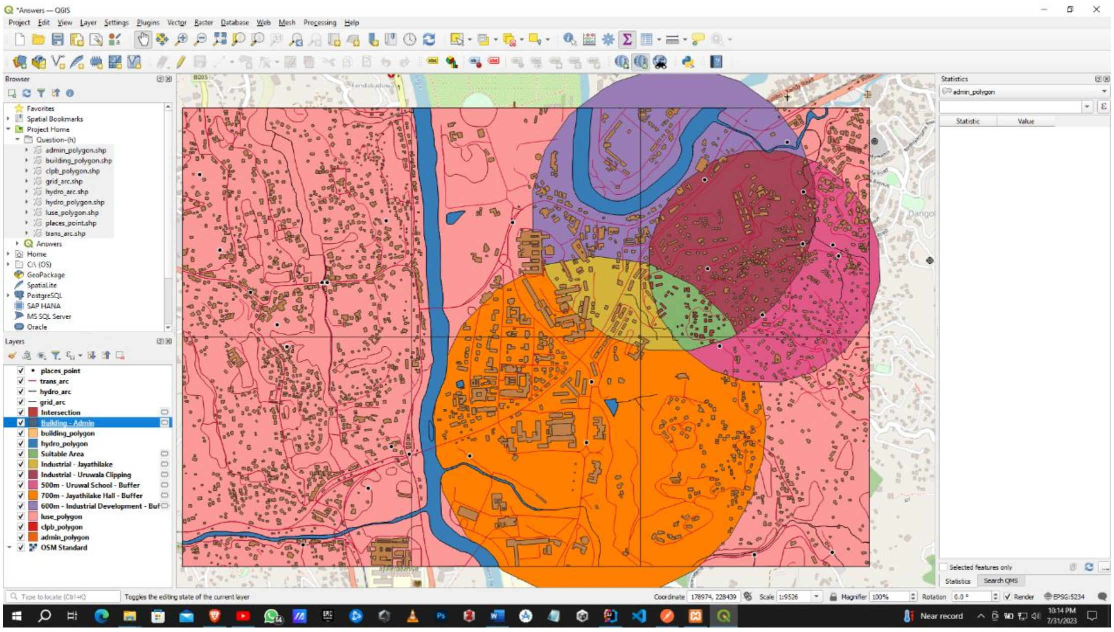


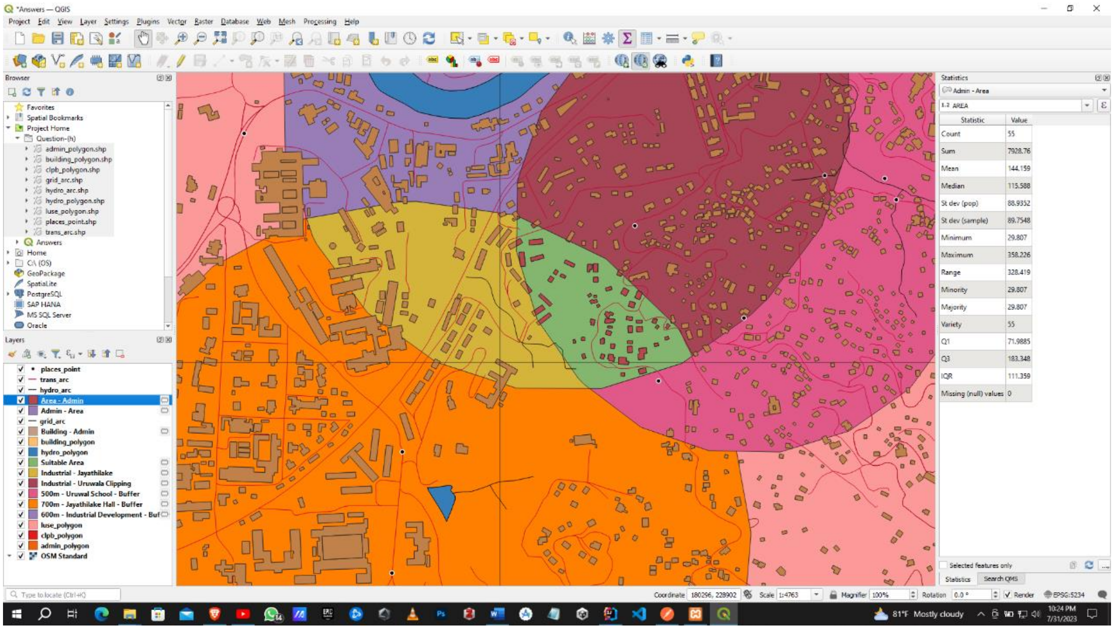


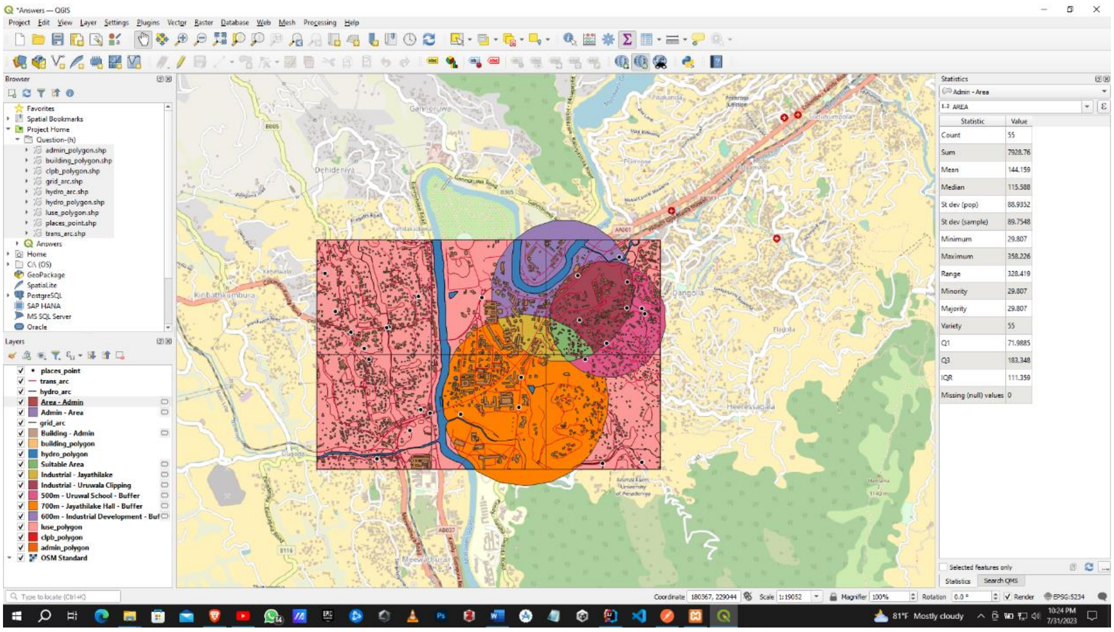


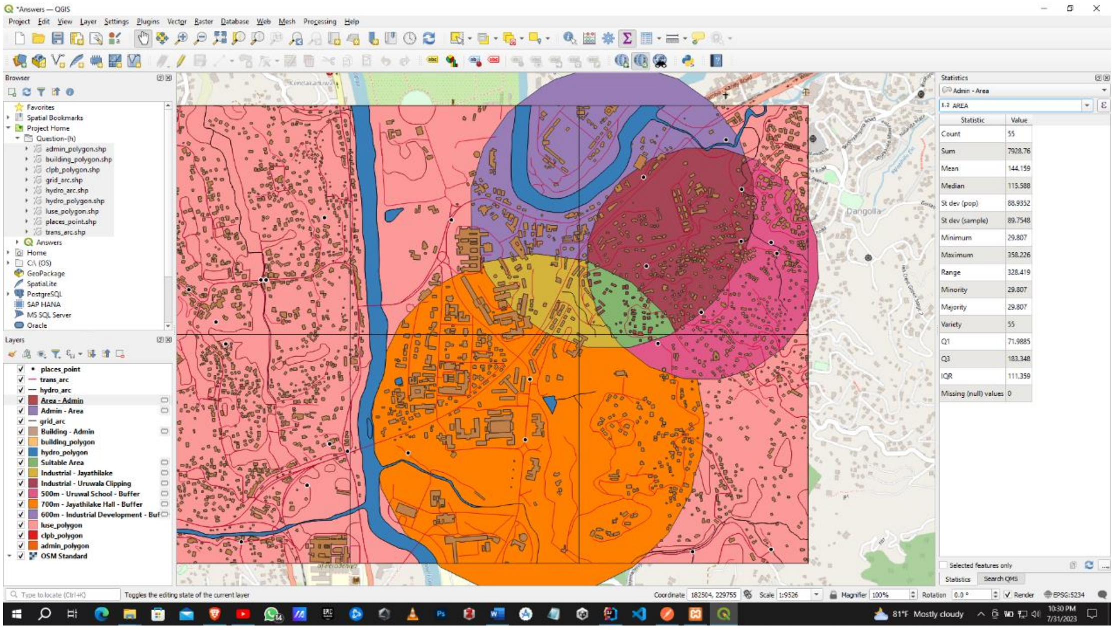


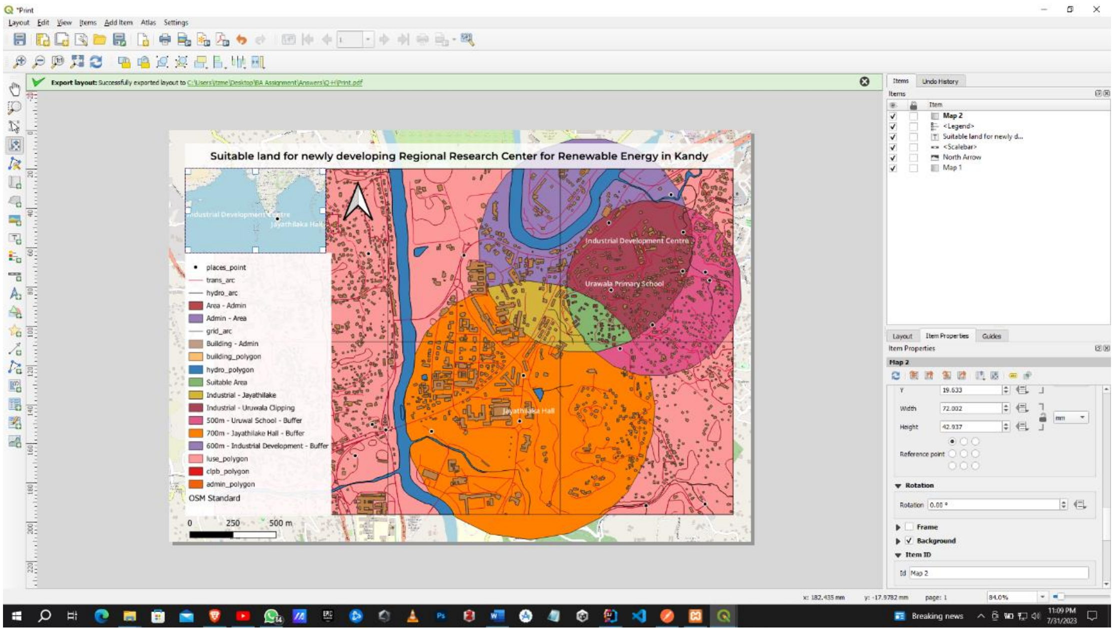


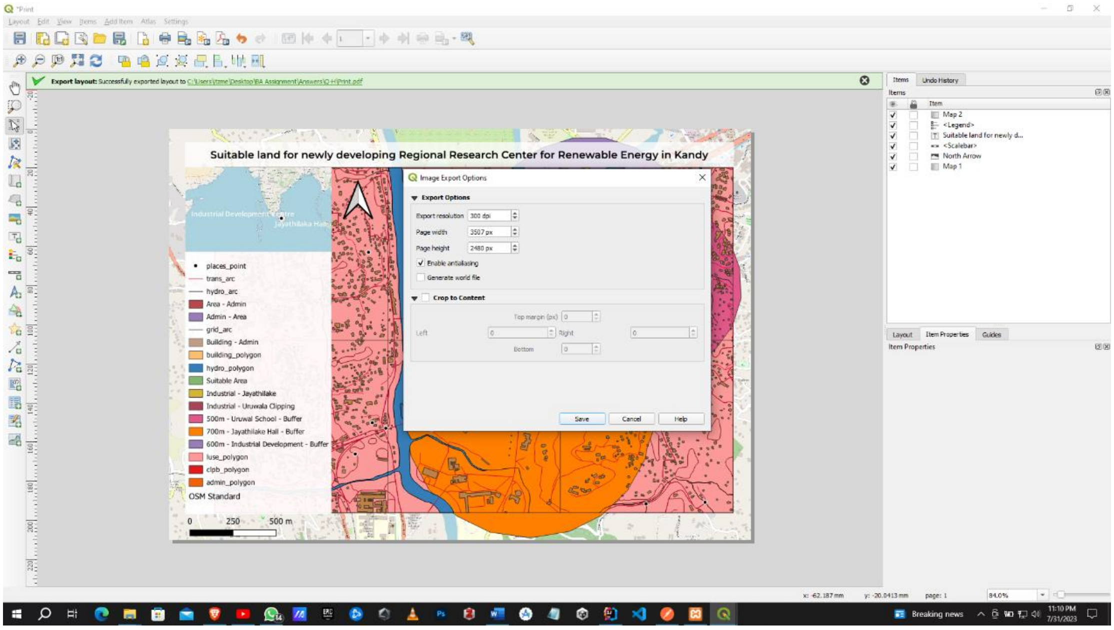


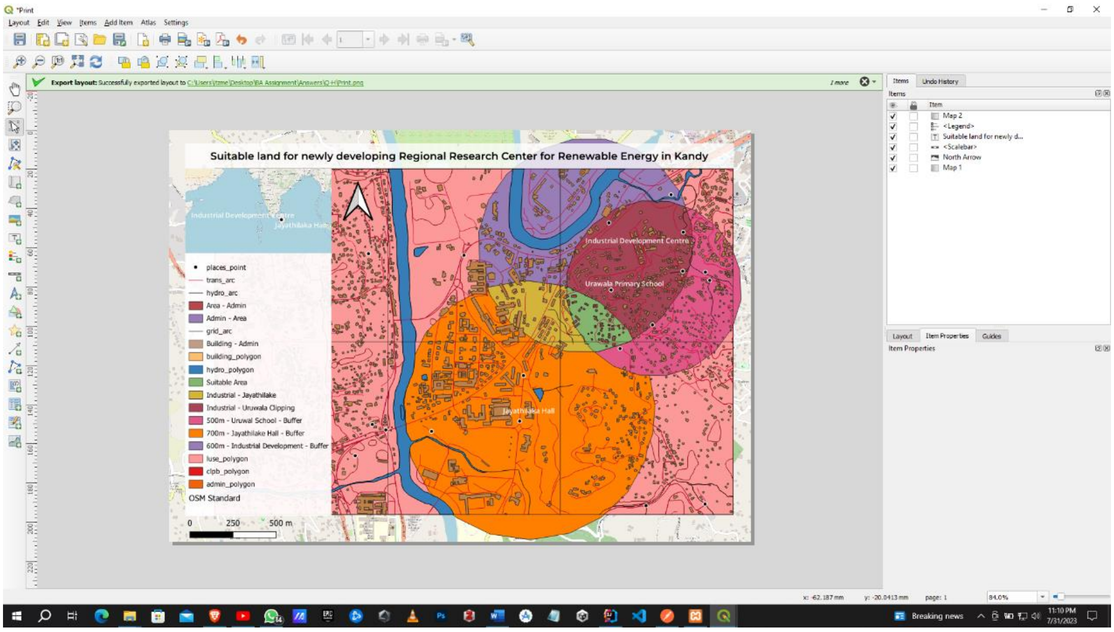


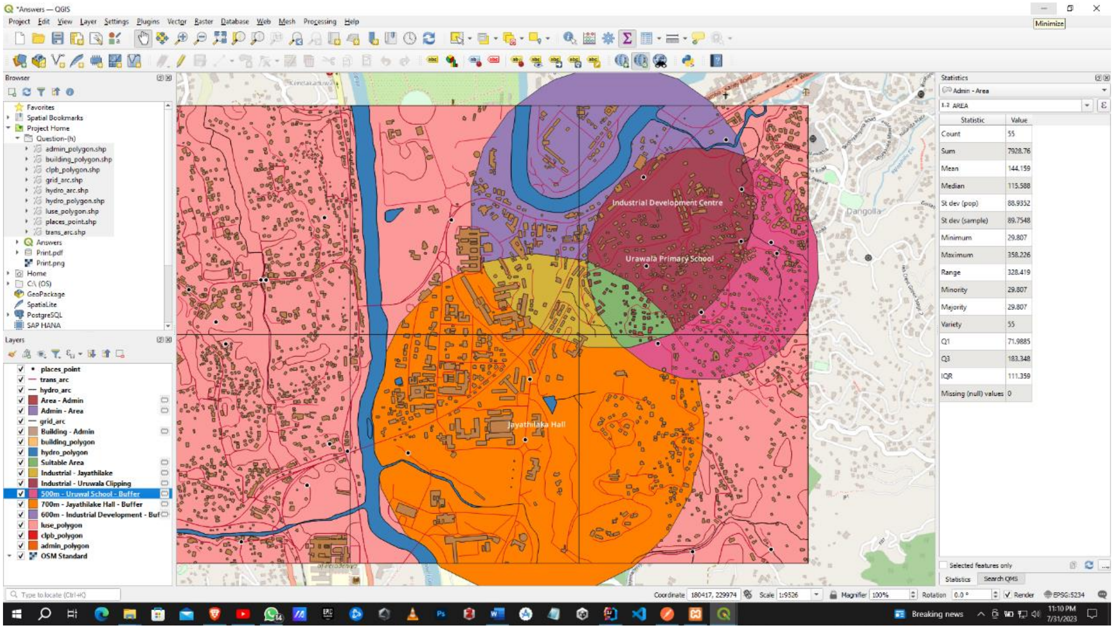





Suitable land for newly developing Regional Research Center for Renewable Energy in Kandy


Suitable land for newly developing Regional Research Center for Renewable Energy in Kandy
Image Export Options
Export Options
Crop to Content
Save
Cancel
Help


Suitable land for newly developing Regional Research Center for Renewable Energy in Kandy
places_point
trans_arc
hydro_arc
Area - Admin
Admin - Area
grid_arc
Building - Admin
building_polygon
hydro_polygon
Suitable Area
Industrial - Jayathilake
Industrial - Uruwala Clipping
500m - Uruwal School - Buffer
700m - Jayathilake Hall - Buffer
600m - Industrial Development - Buffer
luse_polygon
clpb_polygon
admin_polygon
OSM Standard
Industrial Development Centre
Urawala Primary School
Jayathilaka Hall


*Answers — QGIS
Industrial Development Centre
Urawala Primary School
Jayathilaka Hall
Statistics
Admin - Area
AREA
Statistic Value
Count 55
Sum 7928.76
Mean 144.159
Median 115.588
St dev (pop) 88.9352
St dev (sample) 89.7548
Minimum 29.807
Maximum 358.226
Range 328.419
Minority 29.807
Majority 29.807
Variety 55
Q1 71.9885
Q3 183.348
IQR 111.359
Missing (null) values 0