Name: Om Vaknalli

Roll No. 18376

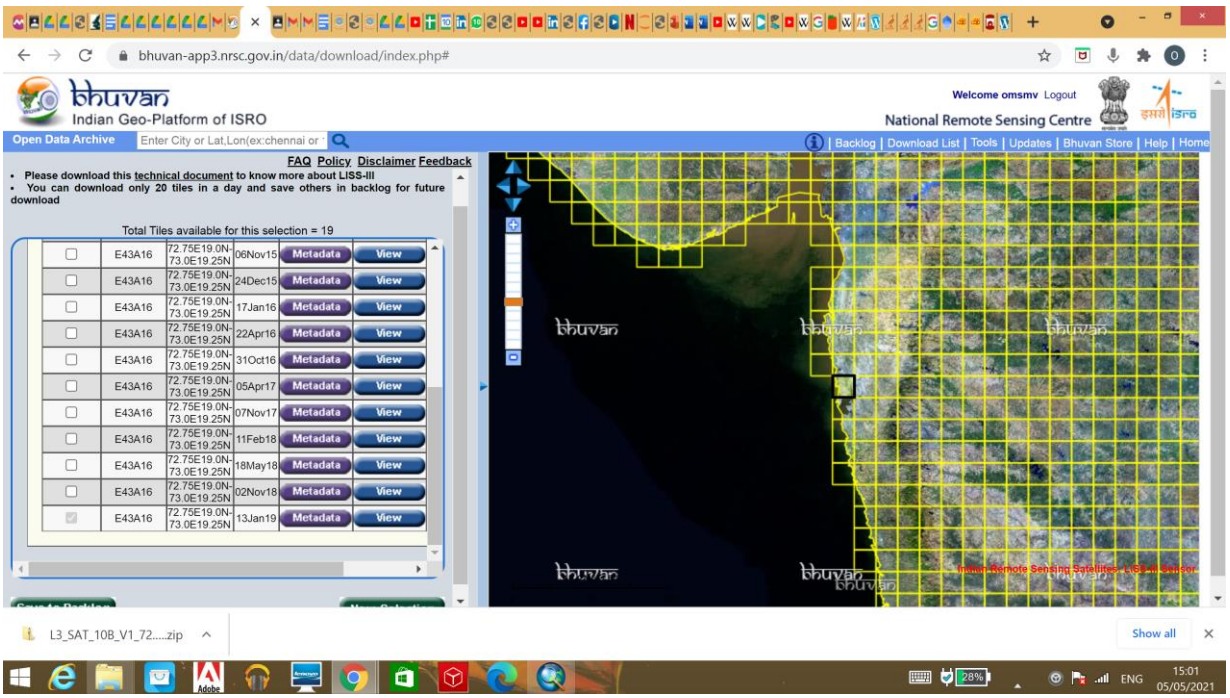
Date: 05/05/21

Subject: EES – 338, Remote Sensing and GIS Laboratory, Final Examination

Question 1 :-

Area of Interest: Suburban Mumbai

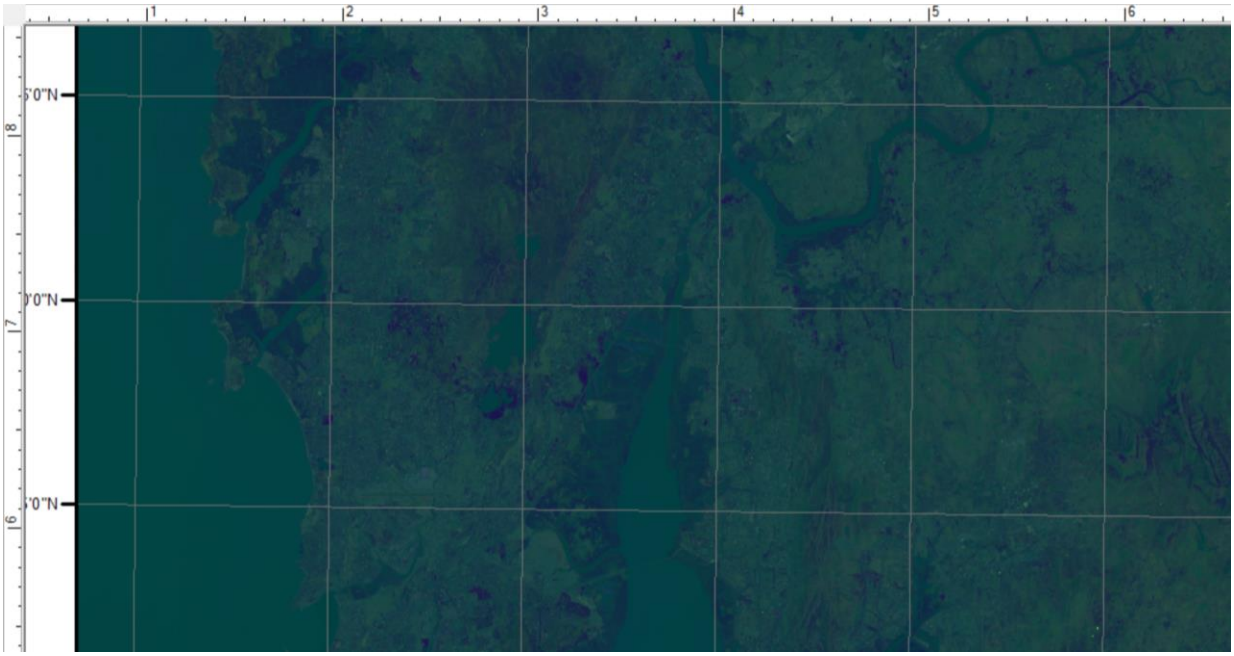
Bands chosen: L3_SAT_10B_V1_72.75E19N_E43A16_13Jan19 (ResourceSat – 1, LISS – 3, Bands 2 to 5)



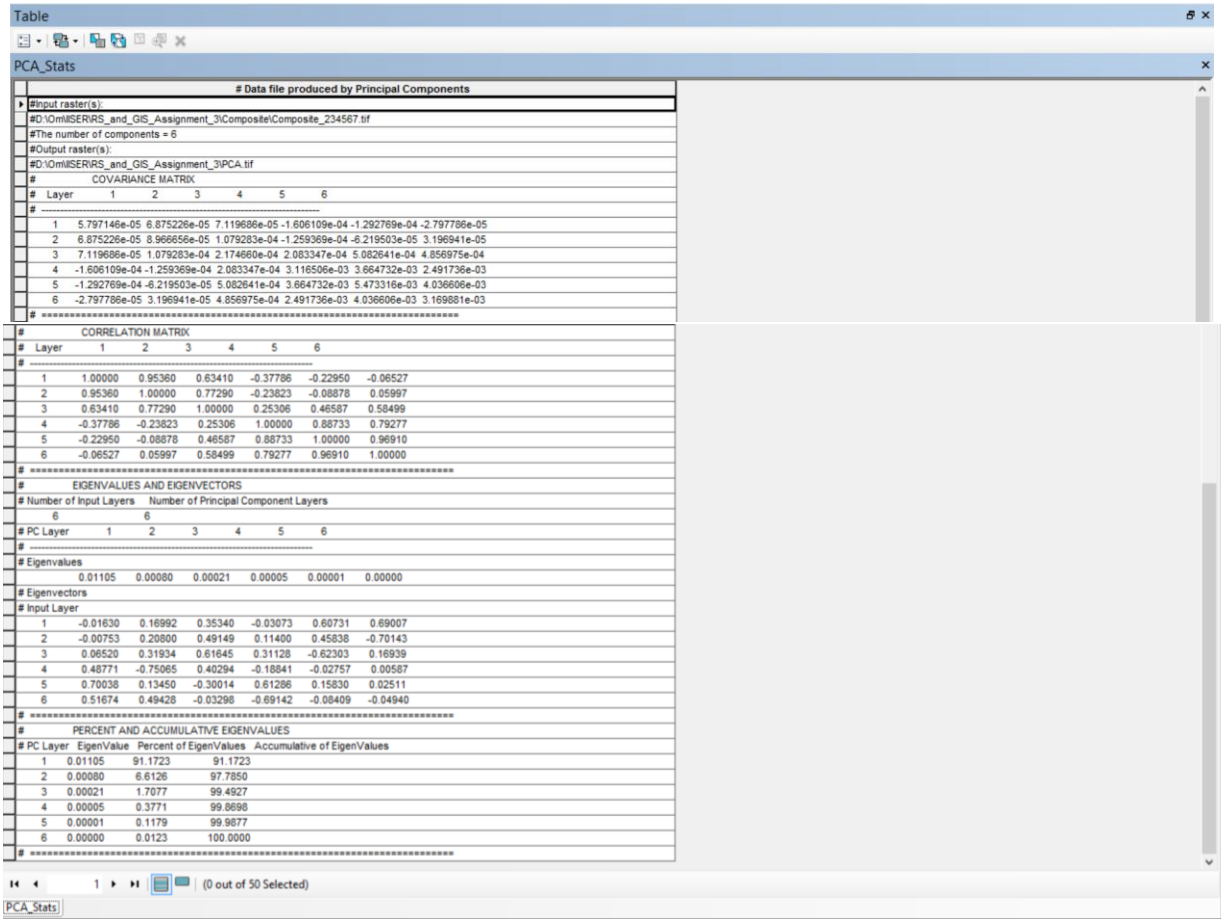
Data download from Bhuvan website

(a)

Colour composite of bands 2, 3, 4 and 5

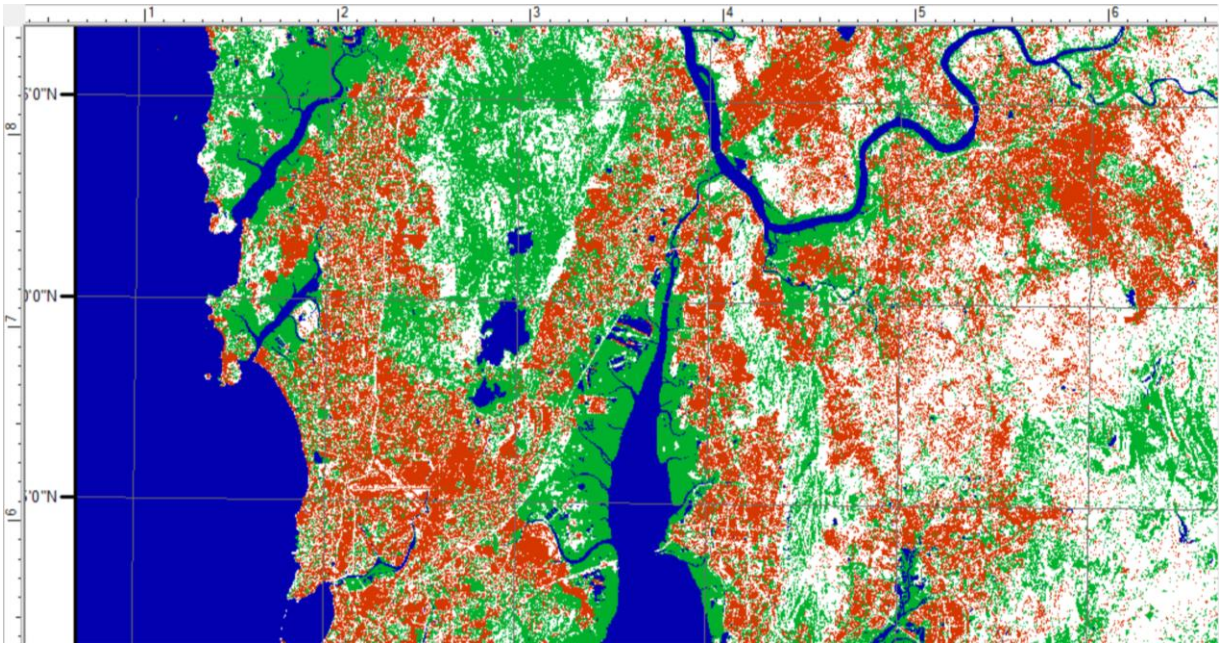


Result after application of Principal Component Analysis (PCA)

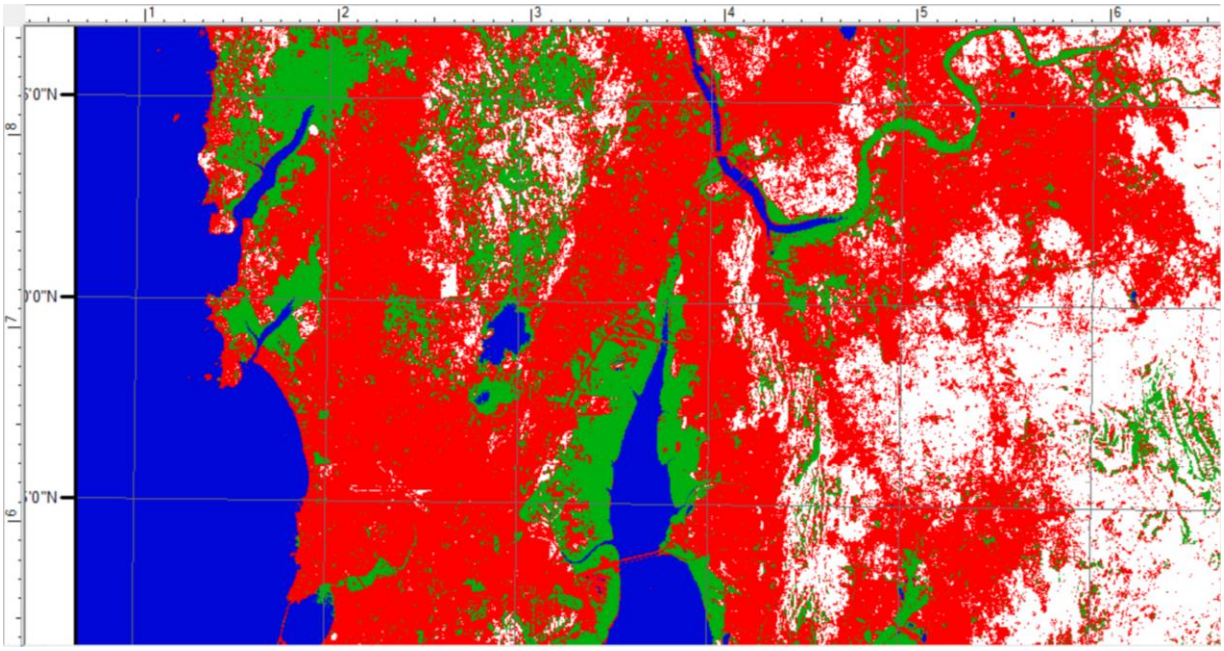


PCA Statistics

(b)



Result after application of Unsupervised Classification (using 25 classes)



Result after application of Supervised Classification

The accuracy of the Supervised Classification is greater than the accuracy of the Unsupervised Classifcation. This is fairly intuitive as we know that Supervised Classification works better when the data being analyzed is heterogeneous and poorly sorted/hetero-clustered (seen with urban

area accuracy comparison below), while Unsupervised Classification works better when the data is more homogeneous and well sorted (as can be seen in the comparison of waterbody accuracies). Considering that our data was more towards the heterogeneous side, it is only plausible to have better overall accuracy and Kappa coefficient of the Supervised Classification analysis.

(c)

Rowid_1	ROWID	CLASSIFIED	GRNDTRUTH-1	GRNDTRUTH1	GRNDTRUTH2	GRNDTRUTH4	GRNDTRUTH5
1	1	0	14	0	0	0	0
2	2	0	9	1	0	0	0
3	4	0	13	0	4	0	2
4	5	0	5	0	1	4	0

Sr. No.	Classification Type	Waterbodies	Forest Land	Open Land	Built-up	Ground Truth Values	Total Classified Pixel
1	Waterbodies	14	0	0	0	0	14
2	Forest Land	9	1	0	0	0	10
3	Open Land	13	0	4	0	2	17
4	Built-up	5	0	1	4	0	10
-	Total Ground Truth Pixels	41	1	5	4	-	51
a	User Accuracy	100	10	23.53	40	-	-
b	Producer Accuracy	34.15	100	80	100	-	-
c	Commission Error	0	90	76.47	60	-	-
d	Omission Error	65.85	0	20	0	-	-

Contingency table and accuracy analysis for Unsupervised Classification

Overall Accuracy = 45.09803922

Kappa Coefficient = 0.245243129

Rowid_1	ROWID	GRNDTRUTH1	GRNDTRUTH101	GRNDTRUTH201	GRNDTRUTH301
4	301	1	6	3	7
3	201	0	2	13	0
2	101	2	8	0	0
1	1	13	0	0	0

Sr. No.	Classification Type	Waterbodies	Forest Land	Open Land	Built-up	Total Classified Pixel
1	Waterbodies	13	0	0	0	13
2	Forest Land	2	8	0	0	10
3	Open Land	0	2	13	0	15
4	Built-up	1	6	3	7	17
-	Total Ground Truth Pixels	16	16	16	7	55
a	User Accuracy	100	80	86.67	41.18	-
b	Producer Accuracy	81.25	50	81.25	100	-
c	Commission Error	0	20	13.33	58.82	-
d	Omission Error	18.75	50	18.75	0	-

Contingency table and accuracy analysis for Supervised Classification

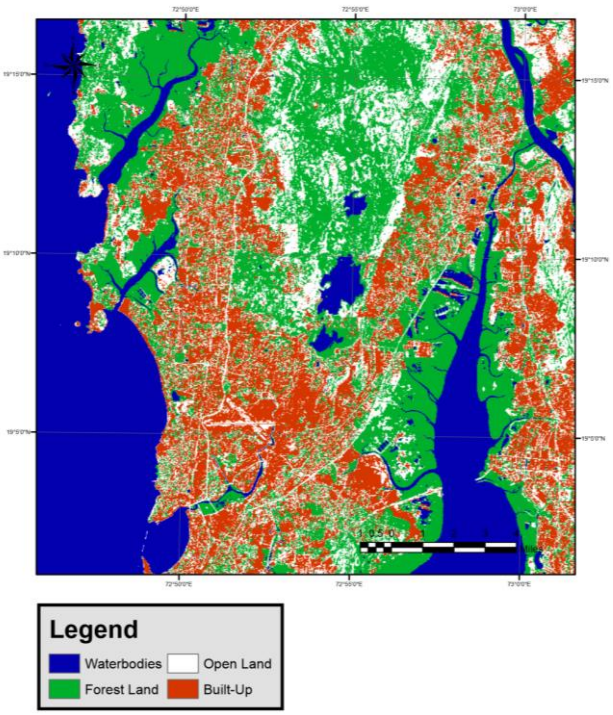
Overall Accuracy = 74.54545455

Kappa Coefficient = 0.6649260226

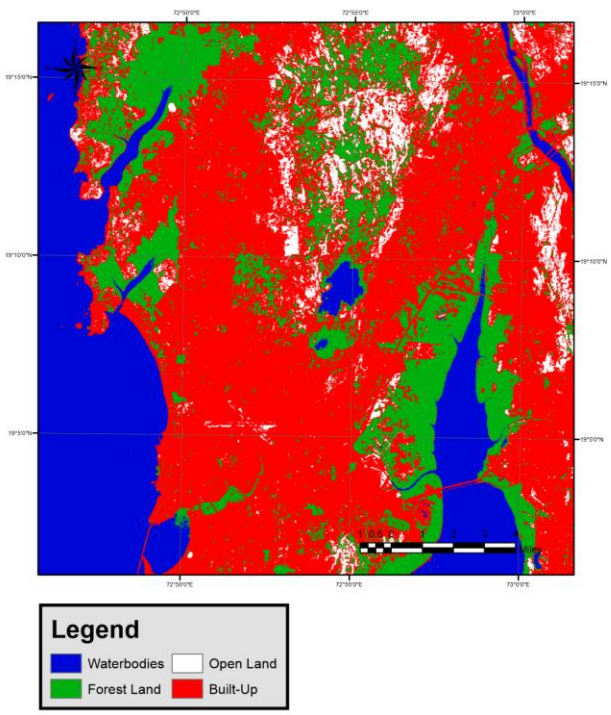
Results:

i. Classified maps (unsupervised/supervised) with a proper legend, coordinates, and scale:

Mumbai Topographical Map (Unsupervised Classification)



Mumbai Topographical Map (Supervised Classification)



The map does not contain agricultural lands as Mumbai's topography does not have enough of it to separately classify it. A larger no. of classes would have been required to incorporate the farming lands in unsupervised classification. To maintain uniformity, the supervised classification was also done by excluding this variable.

ii. A signature file containing the training samples that you have generated for supervised

Classification: Hyperlinked to the google drive containing the file.

iii. Error matrix showing with accuracy assessment reports :-

Error Matrices are included in section **1(c)**.

The waterbody land-use class had the best accuracy through both the machine learning methods simply because the Mumbai AOI spanned about 40% of waterbodies. This made it easier to collect larger samples of training data in these regions for Unsupervised Classification. The rest of the classes had similar order of errors. The complexity of the images was high, i.e., the sorting of the various land-use elements was very poor, which led to these results.

In Supervised Classification, the user accuracies of waterbodies, vegetation and open areas were greater than the producer accuracies of the same. This signifies that the algorithm identified lesser points in these land use classes than that which were actually present. On the other hand, the urban area showed a higher producer accuracy compared to the user accuracy, which meant that a higher percentage of points were excessively identified in this class. A similar logic can also apply to the Unsupervised Classification wherein the waterbodies and urban areas show similar trends as above while the vegetation and open areas accuracy trends have reversed when compared with Supervised Classification. The lesser the difference between the user and producer accuracies, the better is the classification of that specific class. Hence, when the table of accuracy values are compared for the methods of classification, the winner is quite obvious; it is the unsupervised classification.

iv. Detailed information of the image data used for this exercise. For example, satellite mission, resolutions, date of acquisition, and a brief introduction of the study area :-

Resourcesat – I is a remote sensing and earth observation satellite made and launched into space by the Indian Space Research Organisation (ISRO). It is part of the Indian Remote Sensing Programme and is the 10^{th} satellite installment of this series. Its purpose is to act as the successor for its predecessors and maintaining their data services, while also providing a significant increase in the satellite image quality. It was launched on 17^{th} October 2003 from the Sriharikota launch pad. Its intended life-span was 5 years, however it has substantially outlived it by presently operating for 17+ years. It has a 817 km high polar sun-synchronous orbit and majorly tracks the data from the Indian Subcontinent. The satellite majorly controls 3 multispectral pushbroom scanner units:

1. the Linear Imaging Self Scanner (LISS – 4): It has a high resolution (5.8 m spatial resolution and 5-day temporal resolution), 23.9 km swath and operates in the visible and near infrared (NIR) ranges using 3 spectral bands.
2. the Linear Imaging Self Scanner (LISS – 3): It has a moderate resolution (23.5 m spatial resolution and 24-day temporal resolution), 141 km swath and operates in the visible, NIR and shortwave infrared (SWIR) ranges using 4 spectral bands. Mentioned below are the details of these spectral bands.

Spectral Band	Wavelength	Resolution
Band 1	0.52 - 0.59 µm	23.5 m
Band 2	0.62 - 0.68 µm	23.5 m
Band 3	0.77 - 0.86 µm	23.5 m
Band 4	1.55 - 1.70 µm	23.5 m

3. the Advanced Wide Field Sensor (AWiFS): It has a lower spatial resolution of 56 m, 740 km swath and operates in the visible, NIR and SWIR ranges using 4 spectral bands.. Mentioned below are the details of these spectral bands.

Spectral Band	Wavelength	Resolution
Band 1	0.52 - 0.59 µm	56 m
Band 2	0.62 - 0.68 µm	56 m
Band 3	0.77 - 0.86 µm	56 m
Band 4	1.55 - 1.70 µm	56 m

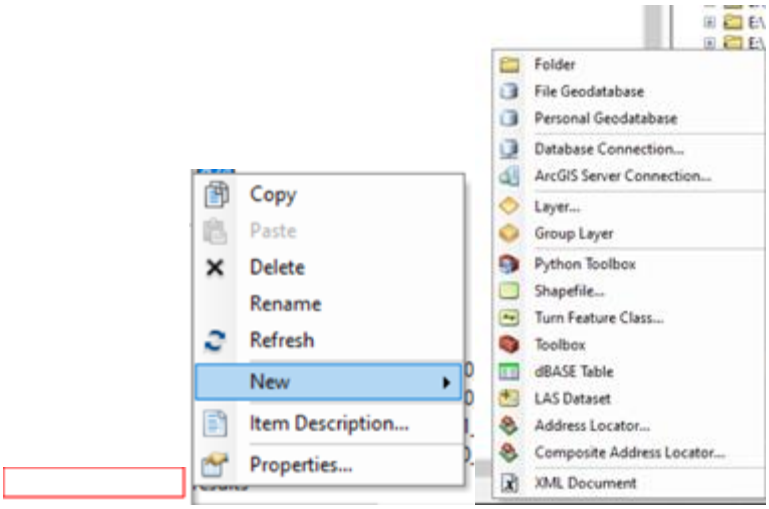
The raw data images used from the Resourcesat – I are captured by the LISS – 3 multispectral scanner on 13th January 2019 at the E43rd path and the A16th row. It uses the WGS-84 dating system.

Suburban Mumbai is a densely populated metropolitan city located on the west coast of the Indian Subcontinent. The city is surrounded partially surrounded by the Arabian Sea on its western and southern fronts while being land-locked on its eastern and northern sides. Topographically speaking, the terrain is majorly urbanised by buildings, roadways, slums and other civil constructions. With that being said, the city houses some dense vegetation in its northern areas (owing to the Sanjay Gandhi National Park and the Aarey Milk Colony) as well as its coastal regions (by Mangrove Forests). Few open lands are scattered throughout and around the city in the form of barren hilly regions (part of the northern end of the deccan traps) and large play and dumping grounds. Apart from the Arabian Sea, Mumbai envelopes several large lakes (Powai Lake, Vihar Lake, Tulsi Lake, etc) as well as a few rivers (Mithi River, Dahisar River, Oshiwara River, Poisar River, Tanasa River, etc) and plentiful creeks (Mahim Creek, Worli Creek, Bhayandar Creek, Vasai Creek, Malad Creek, etc). The city has negligible to no agricultural lands due to the scarcity of space.

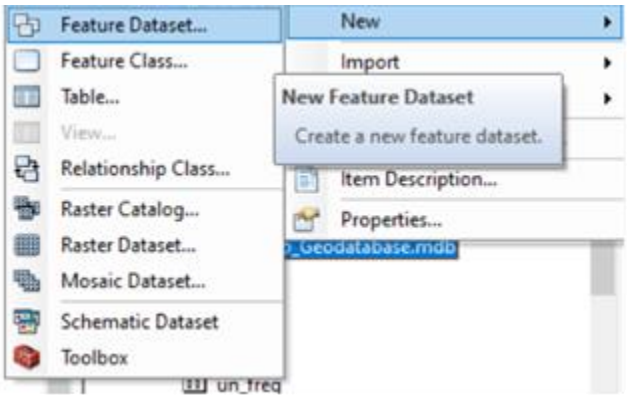
Question 2 :-

(a)

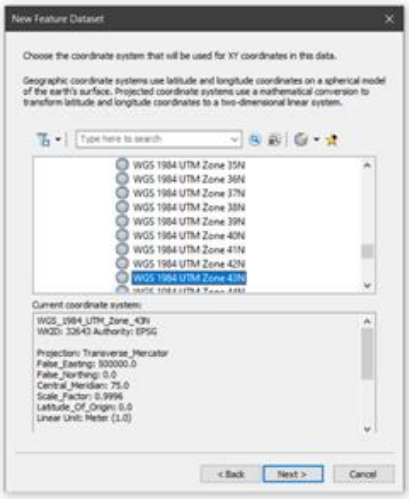
Data base creation:



In arc catalogue, go to working directory in ArcCatalog tree. Right click on folder and Select Click File Geodatabase.



Renamed the database as “Topo_Geodatabase”. Now Right click on Geodatabase and then Feature Dataset.

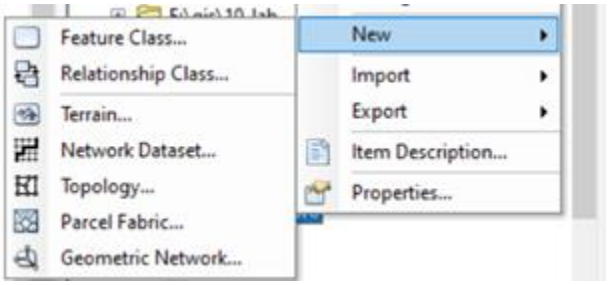


Put name the feature dataset “RS_pro” in dialog box. Select the coordinate system: WGS 1984 UTM Zone 43N (meters). Leave rest to defaults.

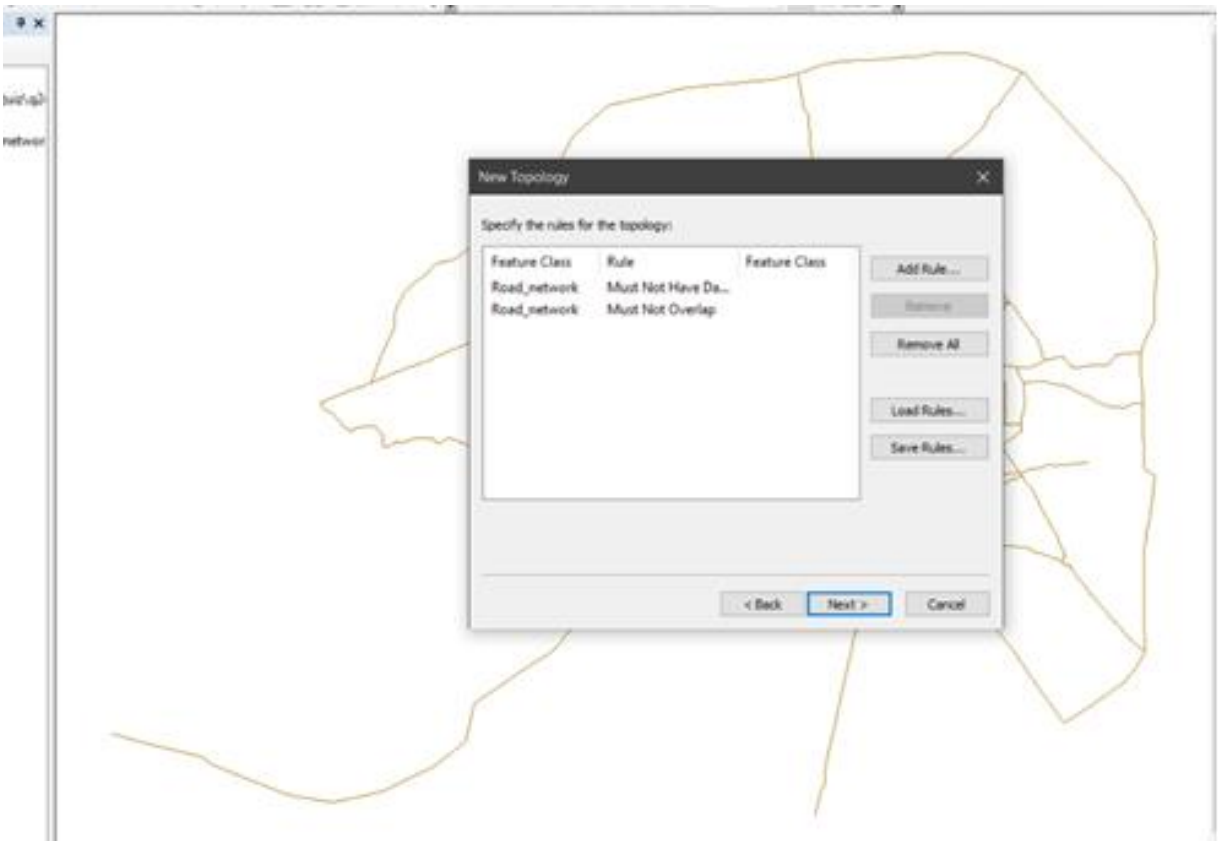
Importing network shapefile:

On Geodatabase, Right click and select Feature Class (multiple) from Import.

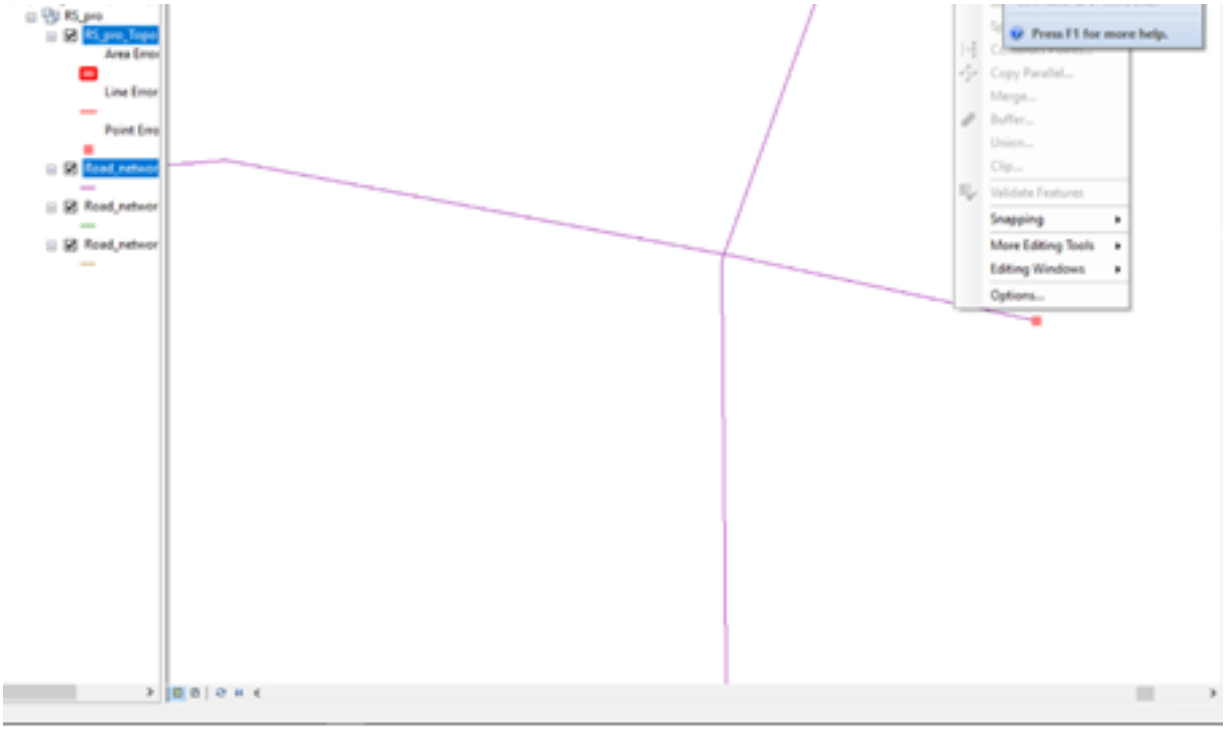
Select, drag and drop the file over the window.



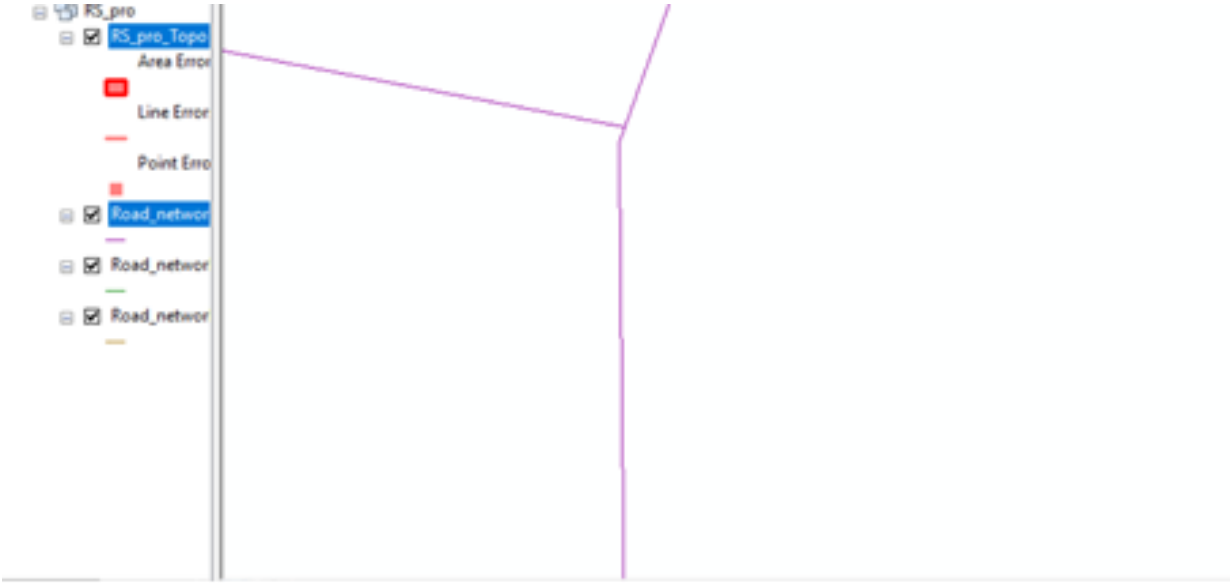
Start editing and create new topology.



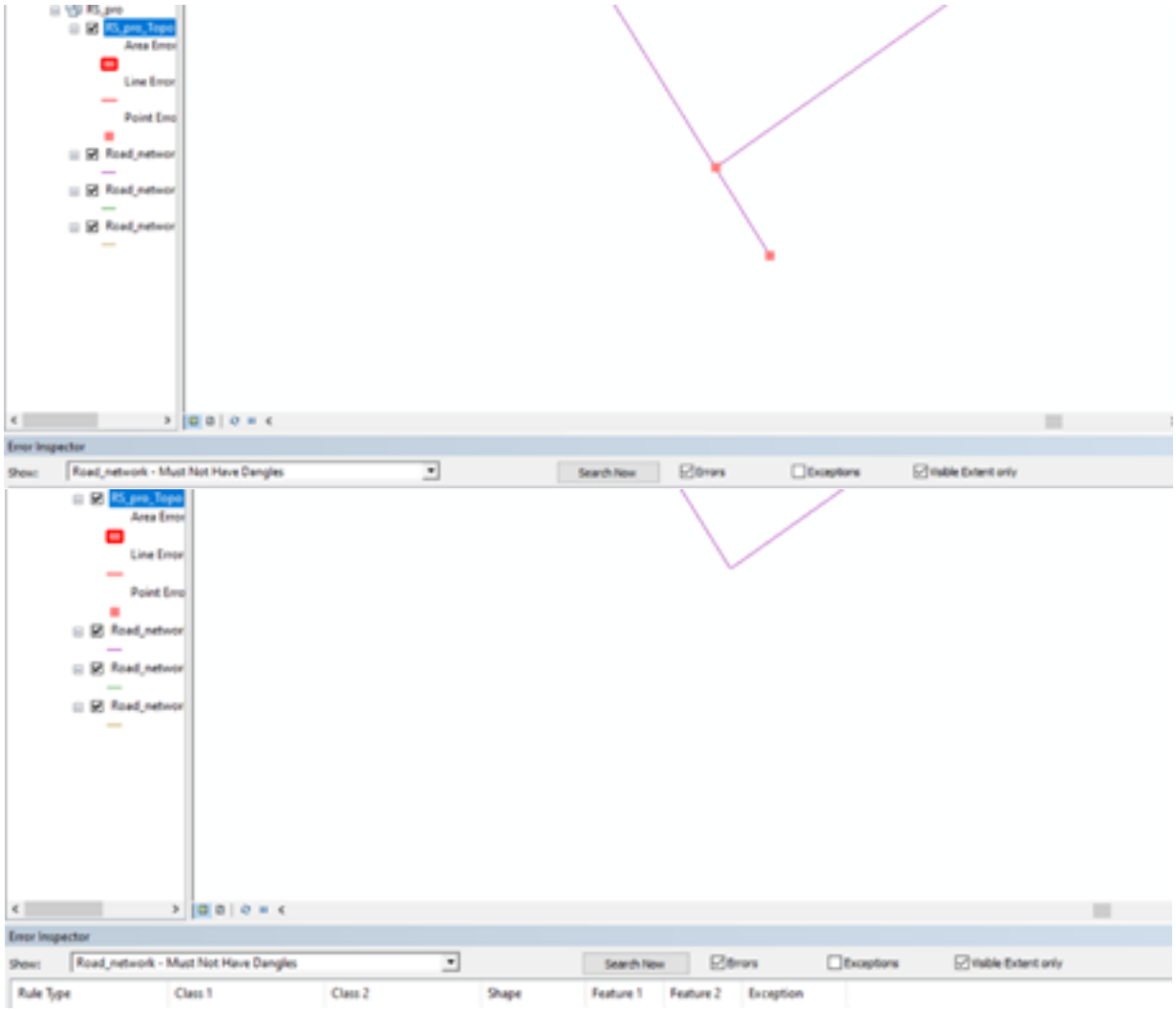
Set error rules to get preferred error heighlighted.



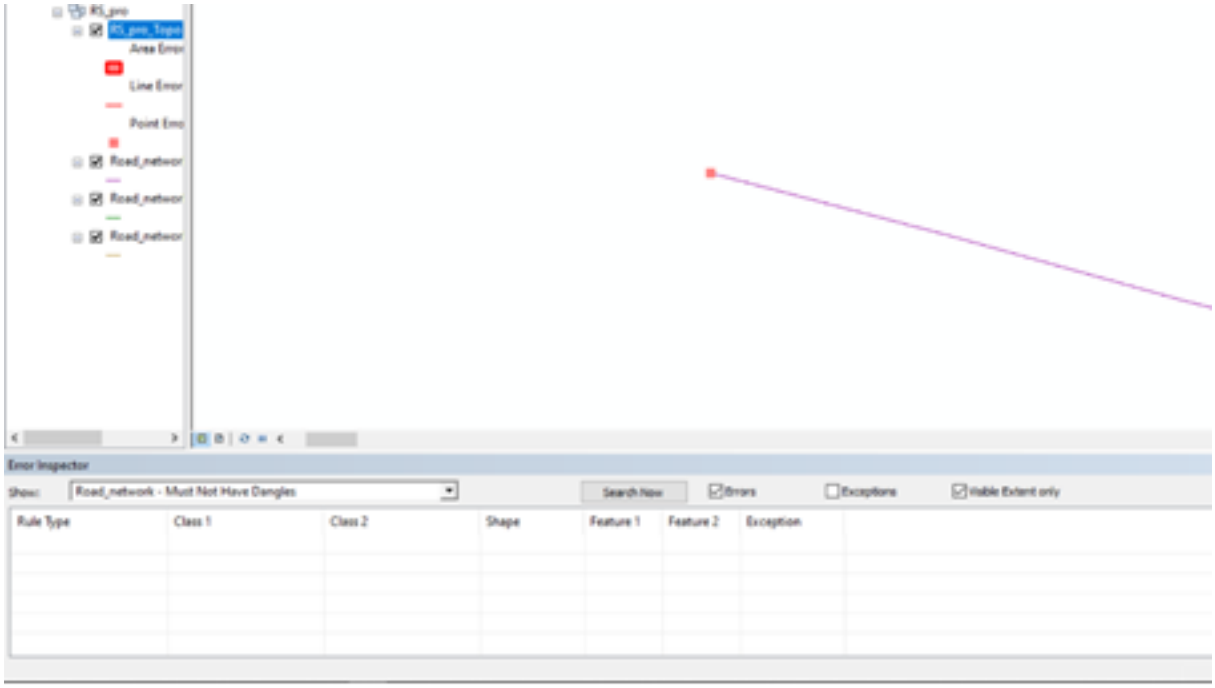
Select Error Inspector and correct the highlighted errors.



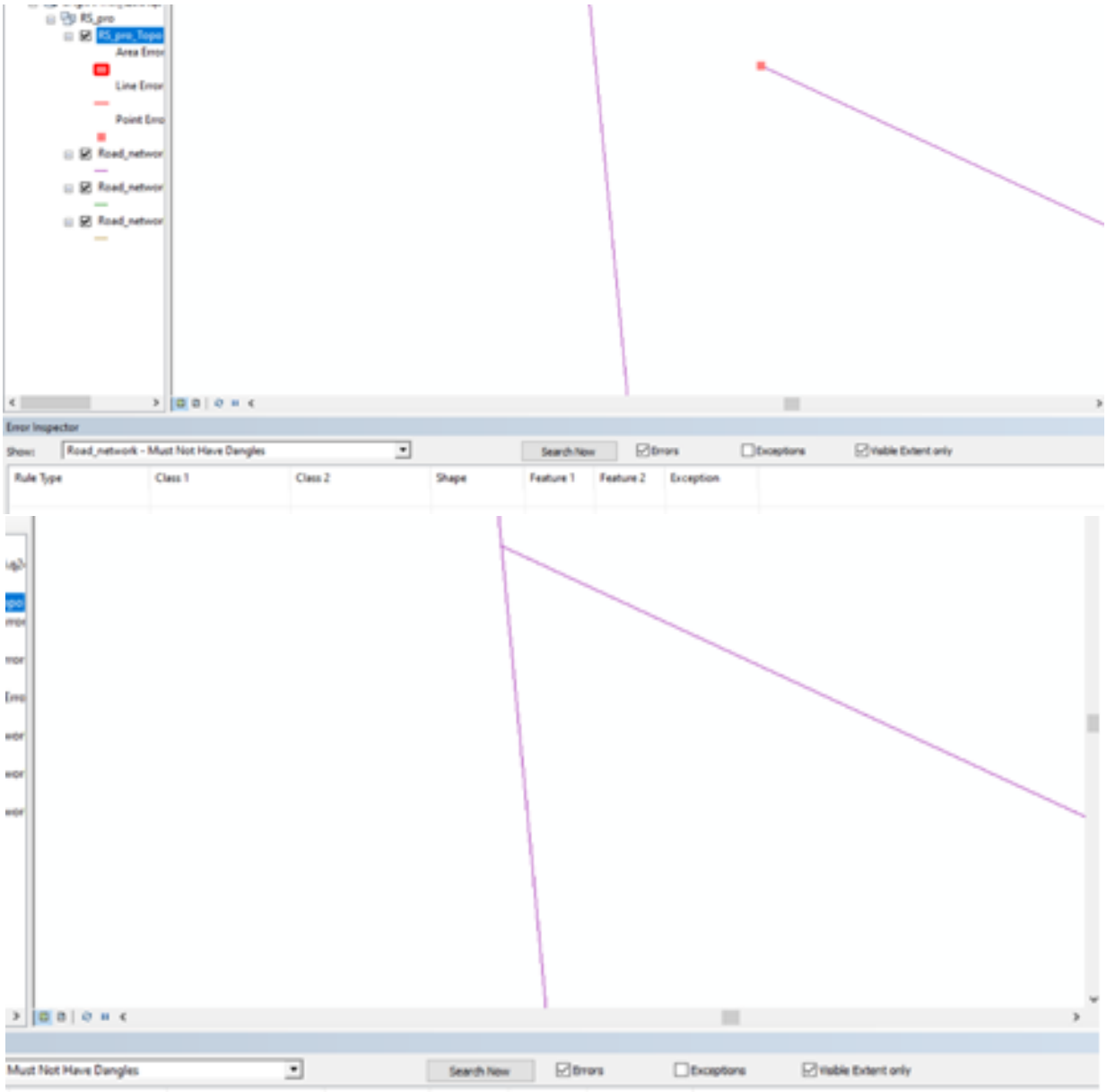
The overshoot error



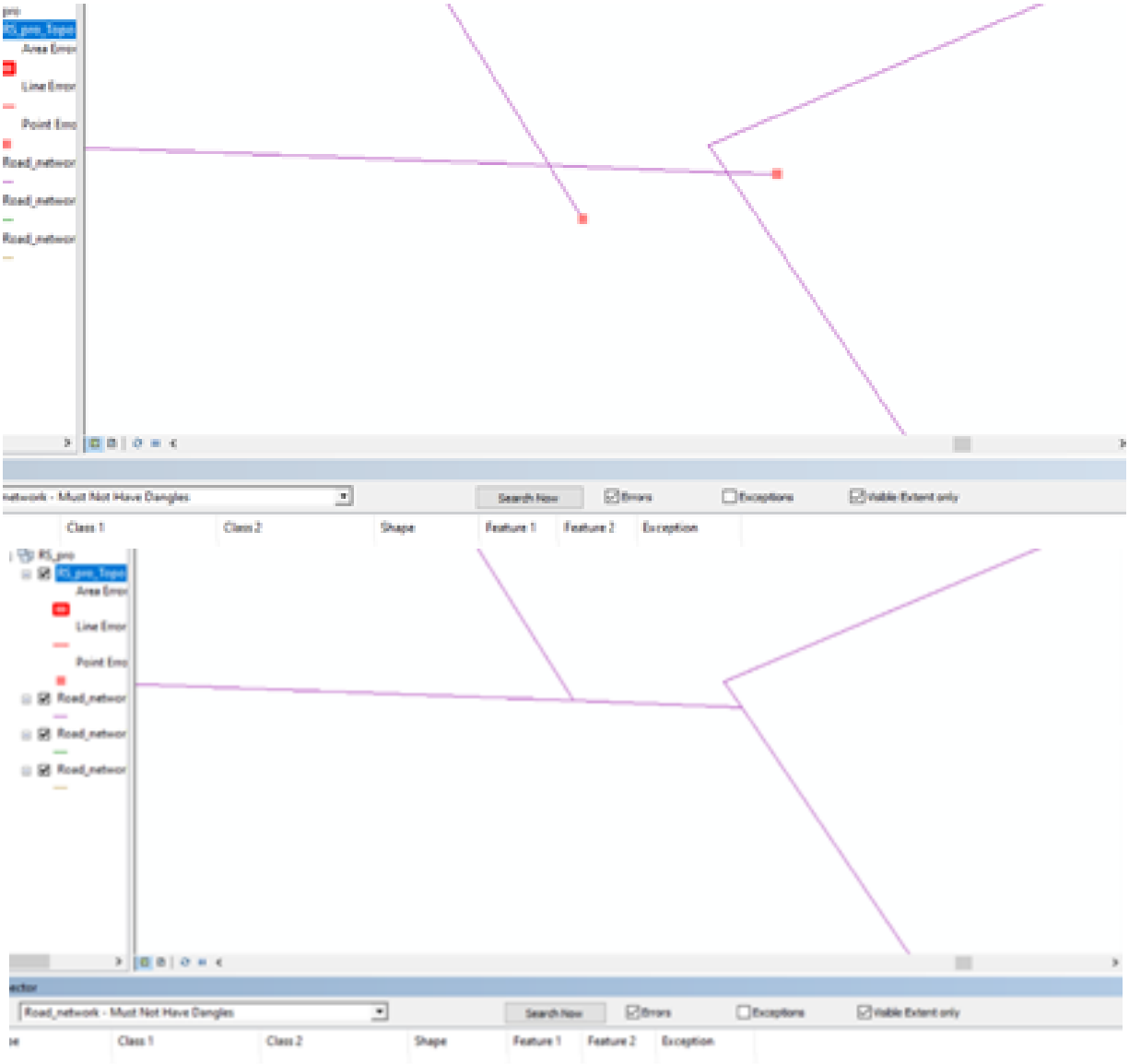
The overshoot error



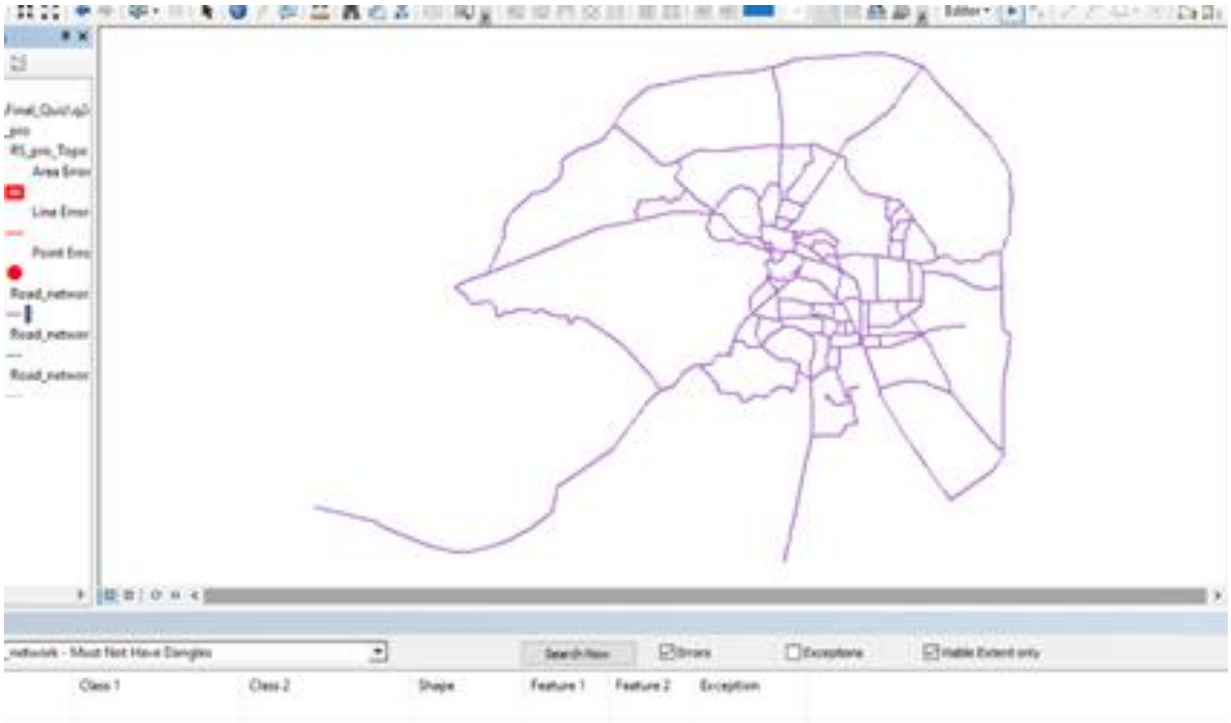
Not an error, have to ignore



Undershoot error

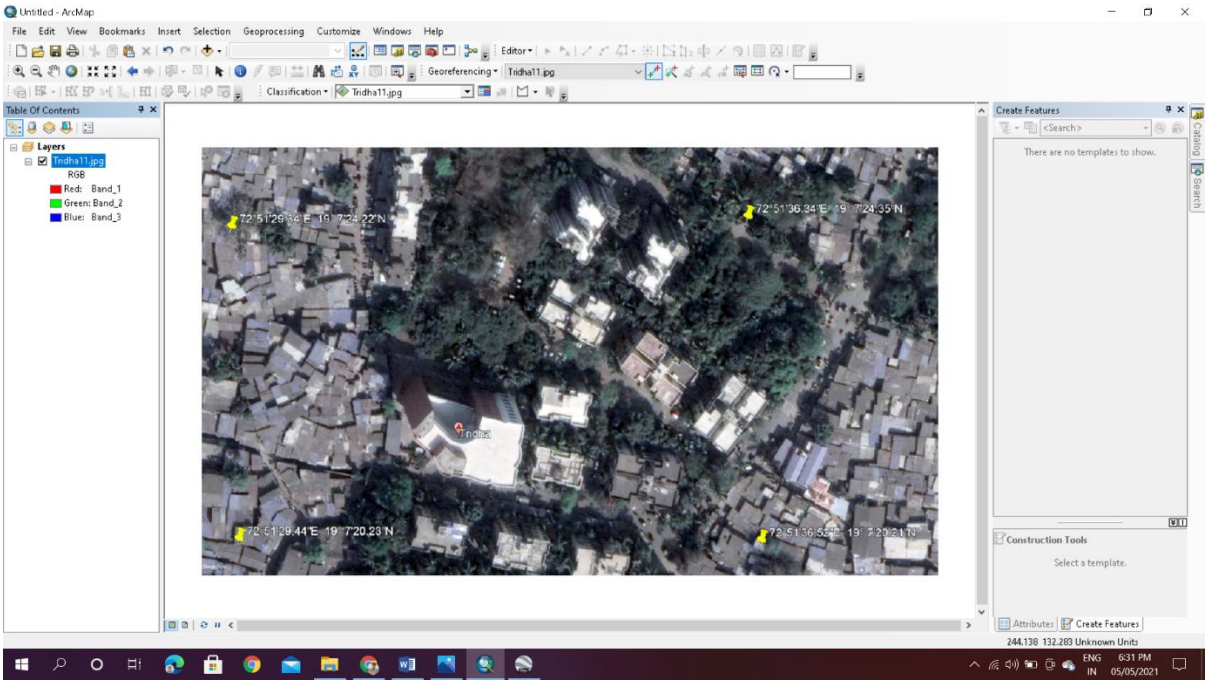


save the edits and stop editing.

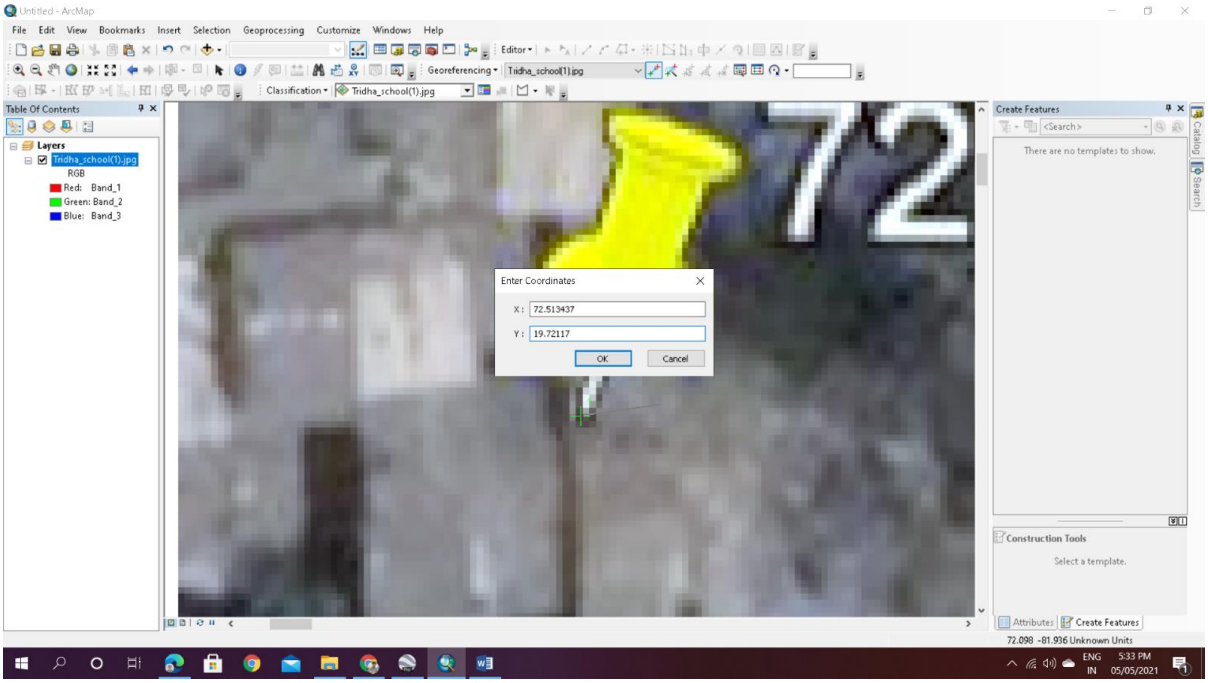


Shapefile after fixes.

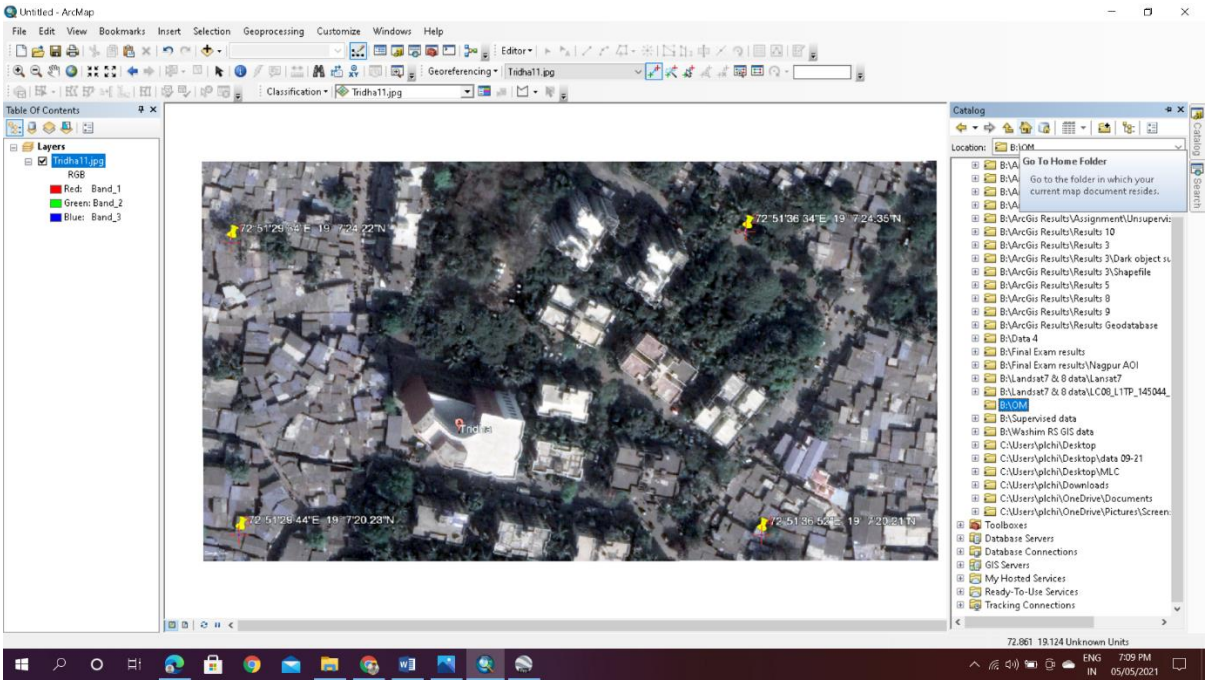
(b)



Downloaded image of my school “Tridha” (before georeferencing)



Selecting ground control points and georeferencing



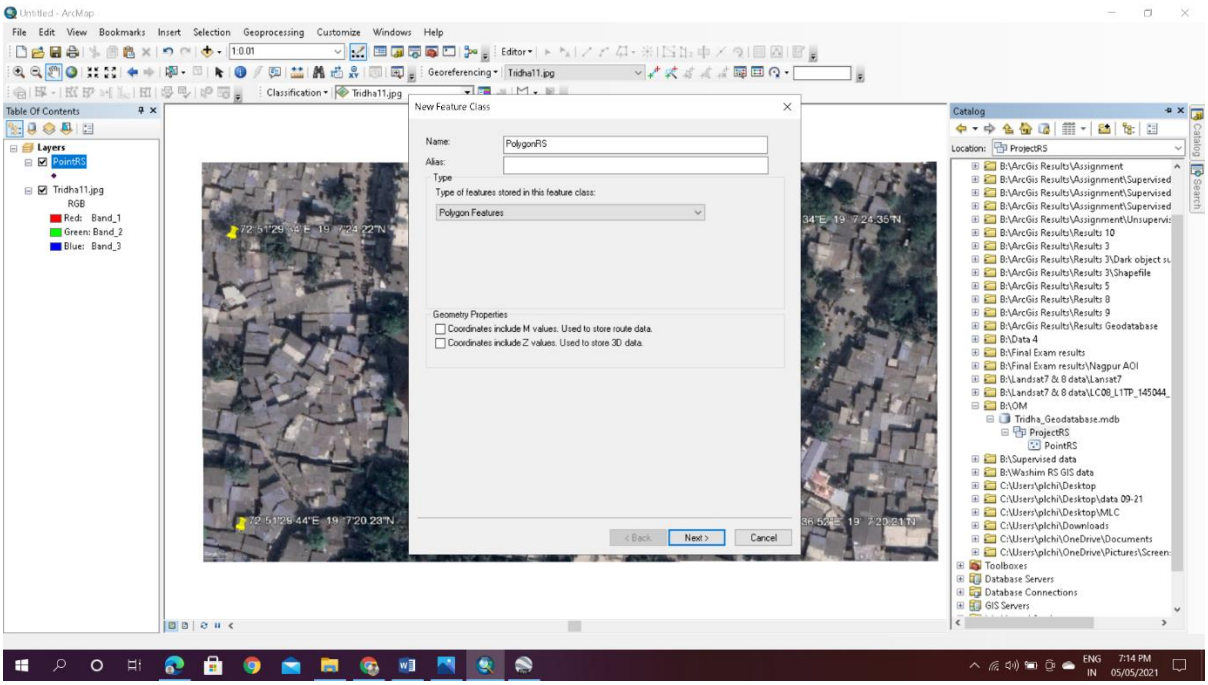
Result after georeferencing

Link

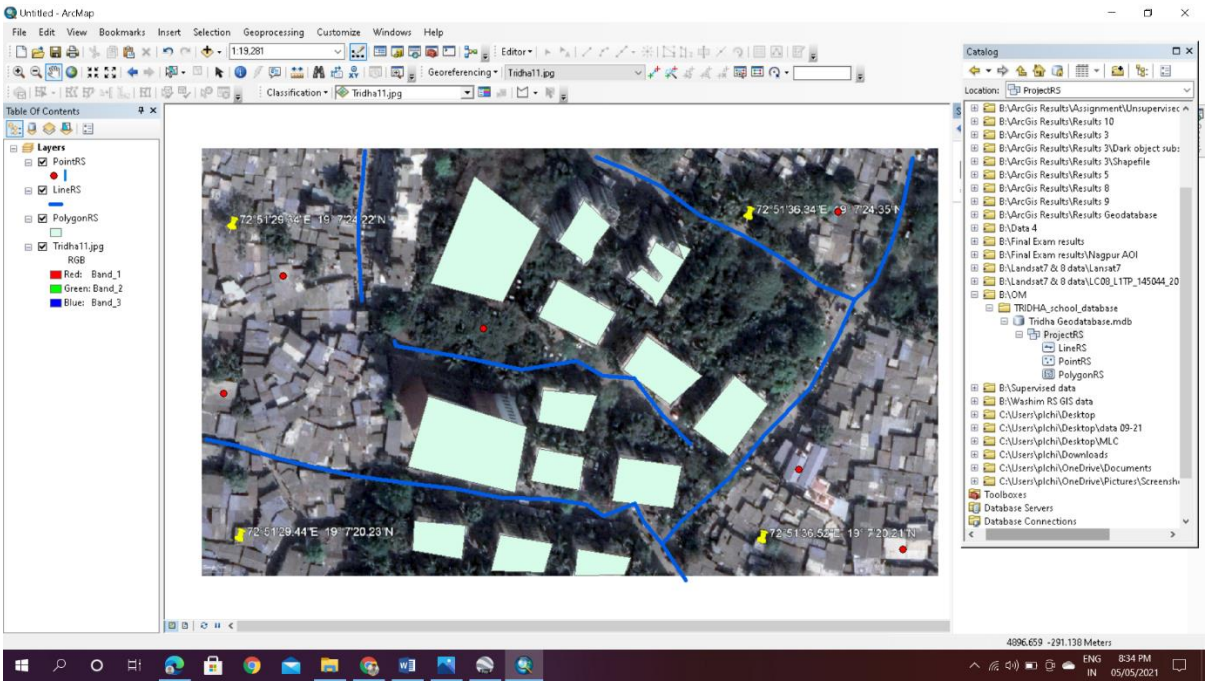
Total RMS Error: Forward:8.03735e-08

	Link	X Source	Y Source	X Map	Y Map	Residual_x	Residual_y	Residual
☑	1	188.02173913	-561.51449275	72.85815000	19.12339000	-0.00000000	-0.00000008	0.00000008
☑	2	3550.13210415	-490.00478747	72.86009000	19.12343000	0.00000000	0.00000008	0.00000008
☑	3	232.37275289	-2597.93858882	72.85818000	19.12229000	0.00000000	0.00000008	0.00000008
☑	4	3642.74829982	-2599.90598132	72.86014800	19.12229000	-0.00000000	-0.00000008	0.00000008

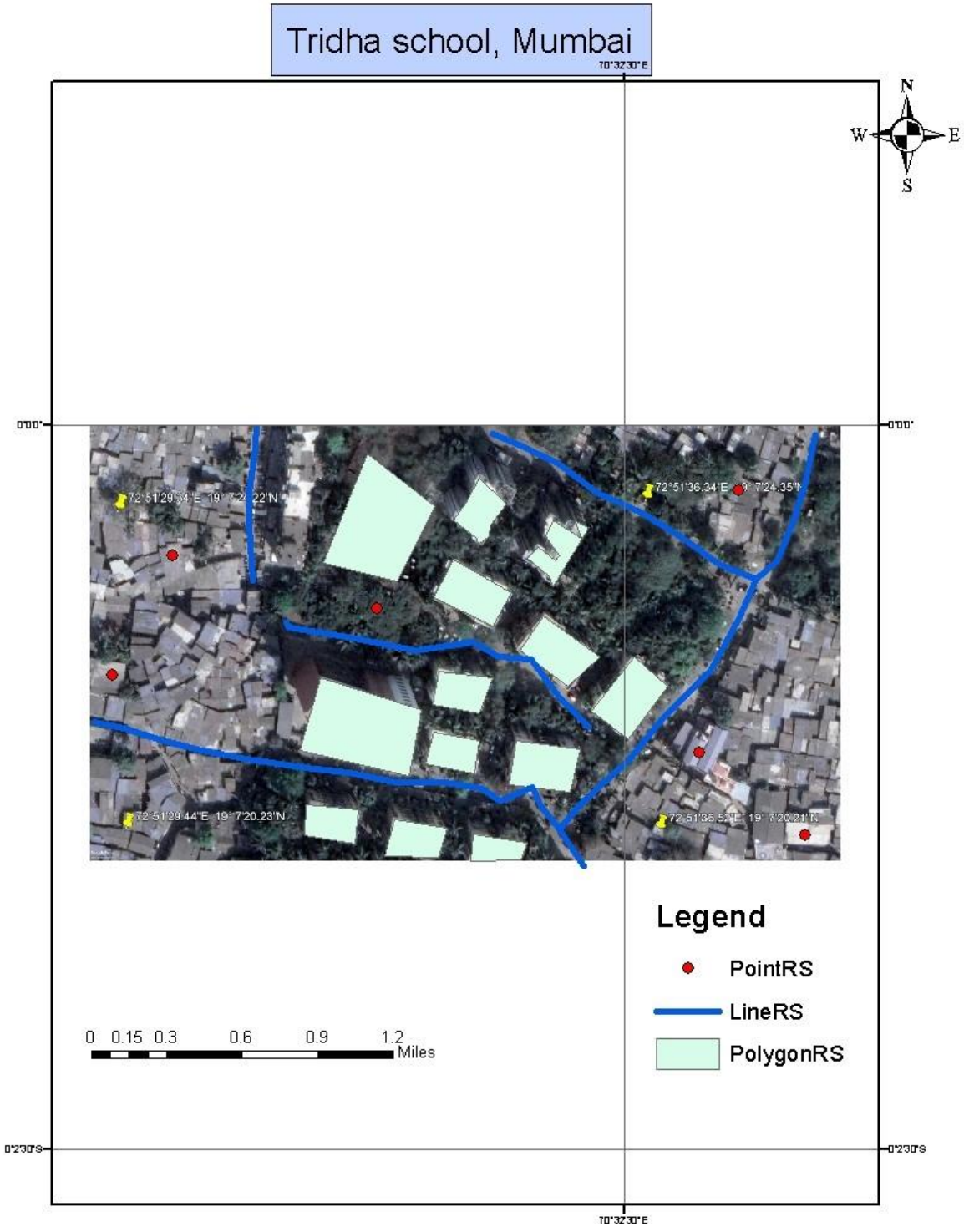
Ground control points



Creating feature classes



Result after digitisation



Result of Digitised School Map

Question 3

- Creating Table:

Table 1: StateCovidDetails

```
CREATE TABLE "StateCovidDetails" (
    "ID"    TEXT,
    "District_Name" TEXT,
    "Confirmed_Case"    INTEGER,
    "Active_Case"   INTEGER,
    "Recovered_Case"    INTEGER,
    "Deceased"  INTEGER
);
```

Table 2: HealthInfrastructure

```
CREATE TABLE "HealthInfrastructure" (
    "ID"    INTEGER,
    "District_Name" TEXT,
    "Name_of_Hospital"  TEXT,
    "Hospital_type" TEXT,
    "Hospital_with_vaccine_facility"    TEXT,
    "Total_population_districtwise" INTEGER,
    "Vaccine_dose_done" INTEGER
);
```

- Based on your understanding, say which candidate key in both the tables, you will use as a primary key and why? Also select an appropriate data type to store the attributes

    - For table 1, District_key should be used as it is unique for all districts.
    - For table 2, Hospital_ID should be used as it is unique for all the hospitals even in the same district.

Using the above tables, answer the following questions;

I) The query to show the district with highest number of cases in State:

```
SELECT *
```

```
FROM StateCovidDetails
ORDER BY Active_Case DESC
LIMIT 1;
```

According to query Pune has the highest number of active cases on 1/04/2021

ID	District_Name	Confirmed_Case	Active_Case	Recovered_Case	Deceased
MH_Pune	Pune	544287	64648	471296	8343

II) The query to show the number of hospitals with vaccine facility with highest number of active cases:

```
SELECT COUNT(Hospital_with_vaccine_facility)
FROM HealthInfrastructure
WHERE District_Name = “Pune”
```

III) SQL query to arrange the districts with active covid cases in ascending order:

```
SELECT *
FROM StateCovidDetails
ORDER BY Active_Case
```

IV) SQL query to show the top five worst affected districts in terms of the confirmed covid case in your state:

```
SELECT *
```

```
FROM StateCovidDetails
ORDER BY Confirmed_Case DESC
LIMIT 5;
```

V) SQL query to show the district which has the least number of hospitals with vaccine facilities:

```
SELECT MIN(SELECT COUNT(Hospital_with_vaccine_facility)
FROM HealthInfrastructure
WHERE HealthInfrastructure.District_Name = StateCovidDetails. District_Name)
As Least_Facility
FROM StateCovidDetails
```

Table Desciption:

ID	District_Name	Confirmed_Case	Active_Case	Recovered_Case	Deceased
Filter	Filter	Filter	Filter	Filter	Filter
MH_Ahmednagar	Ahmednagar	95323	9858	84244	1221
MH_Akola	Akola	28822	4058	24297	467
MH_Amravati	Amravati	49379	2985	45744	650
BR_Aurangabad	Aurangabad	85371	13496	70498	1377
MH_Beed	Beed	26379	4305	21453	621
MH_Bhandara	Bhandara	18625	3114	15190	321
MH_Buldhana	Buldhana	28133	4138	23710	285
MH_Chandrapur	Chandrapur	29542	2414	26683	445
MH_Dhule	Dhule	26474	5456	20653	365
MH_Gadchiroli	Gadchiroli	10137	439	9588	110
MH_Gondia	Gondia	16189	932	15075	182
MH_Hingoli	Hingoli	7110	1867	5140	103
MH_Jalgaon	Jalgaon	86508	6994	77915	1599
MH_Jalna	Jalna	23322	774	22137	411
MH_Kolhapur	Kolhapur	51397	761	48939	1697
MH_Latur	Latur	33904	5935	27217	752
MH_Mumbai	Mumbai	423419	55767	355944	11708
MH_Nagpur	Nagpur	233883	48851	181113	3919

Table-1: Information of Covid cases state-wise

ID	District_Name	Name_of_Hospital	Hospital_type	Hospital_with_vaccine_facility	Total_population_districtwise	Vaccine_dose_done
Filter	Filter	Filter	Filter	Filter	Filter	Filter
NULL	Mumbai	Breach Candy ...	Private	NULL	NULL	602225
NULL	Mumbai	Cumbala Hill Hospit...	Private	NULL	NULL	602225
NULL	Mumbai	Asian Heart Institut...	Private	NULL	NULL	602225
NULL	Mumbai	Lilavati Hospital an...	Private	NULL	NULL	602225
NULL	Mumbai	Tata Memorial ...	Private	NULL	NULL	602225
NULL	Mumbai	Bhatia General ...	Private	NULL	NULL	602225
NULL	Mumbai	Seven Hills Hospital	Private	NULL	NULL	602225
NULL	Mumbai	Bombay Hospital A...	Private	NULL	NULL	602225
NULL	Mumbai	P. D.Hinduja Nation...	Private	NULL	NULL	602225
NULL	Mumbai	Saifee Hospital	Private	NULL	NULL	602225
NULL	Mumbai	Fortis Hospitals, ...	Private	NULL	NULL	602225
NULL	Mumbai	Jaslok Hospital And ...	Private	NULL	NULL	602225
NULL	Mumbai	Guru Nanak Hospital	Private	NULL	NULL	602225
NULL	Mumbai	Criticare ...	Private	NULL	NULL	602225
NULL	Mumbai	Surana Sethia ...	Private	NULL	NULL	602225
NULL	Mumbai	Balaji Hospital	Private	NULL	NULL	602225
NULL	Mumbai	Kokilaben Dhirubhai...	Private	NULL	NULL	602225
NULL	Mumbai	Aditya Jyot Eye ...	Private	NULL	NULL	602225
NULL	Mumbai	Shree Bidada ...	Private	NULL	NULL	602225
NULL	Mumbai	Bombay City Eye ...	Private	NULL	NULL	602225
NULL	Mumbai	The B.D.Petit Parse...	Private	NULL	NULL	602225
NULL	Mumbai	Centre For Sight, ...	Private	NULL	NULL	602225
NULL	Mumbai	Asian Heart Institute	Private	NULL	NULL	602225
NULL	Pune	Aditya Birla Memori...	Private	NULL	NULL	466959
NULL	Pune	Columbia Asia ...	Private	NULL	NULL	466959

Table-2: Information about health infrastructure in Maharashtra.

Have null values due to unavailability of Data.

SQL logs:

```
1  -- EXECUTING ALL IN 'SQL 1'
4   Insert into StateCovidDetail(ID,District_Names,Confirmed_Cases,Recovered,Deceased)
5   Select District_Key,District,Confirmed,Recovered,Deceased
6   from Covid_Raw_Data
7   WHERE Dates = "2021-04-01" AND State = "Maharashtra"
8  -- Result: no such table: StateCovidDetail
12  Insert into StateCovidDetails(ID,District_Names,Confirmed_Cases,Recovered,Deceased)
13  Select District_Key,District,Confirmed,Recovered,Deceased
14  from Covid_Raw_Data
15  WHERE Dates = "2021-04-01" AND State = "Maharashtra"
16 -- Result: table StateCovidDetails has no column named District_Names
20  Insert into StateCovidDetails(ID,District_Name,Confirmed_Case,Recovered_Case,Deceased)
21  Select District_Key,District,Confirmed,Recovered,Deceased
22  from Covid_Raw_Data
23  WHERE Dates = "2021-04-01" AND State = "Maharashtra"
24 -- Result: no such column: Dates
28  Insert into StateCovidDetails(District_Name,Confirmed_Case,Recovered_Case,Deceased)
29  Select District,Confirmed,Recovered,Deceased
30  from districts
31  WHERE Dates = "2021-04-01" AND State = "Maharashtra"
32 -- Result: no such column: Dates
36  Insert into StateCovidDetails(District_Name,Confirmed_Case,Recovered_Case,Deceased)
37  Select District,Confirmed,Recovered,Deceased
38  from districts
39  WHERE Date = "2021-04-01" AND State = "Maharashtra"
40 -- Result: query executed successfully. Took 40ms, 36 rows affected
44  Insert into StateCovidDetails(Active_Case)
45  select Confirmed_Case-Recovered_Case-Deceased
46  from StateCovidDetails
47 -- Result: query executed successfully. Took 0ms, 36 rows affected
51  Insert into StateCovidDetails(Active_Case)
```

```
8  -- Result: no such table: StateCovidDetail
12   Insert into StateCovidDetails(ID,District_Names,Confirmed_Cases,Recovered,Deceased)
13   Select District_Key,District,Confirmed,Recovered,Deceased
14   from Covid_Raw_Data
15   WHERE Dates = "2021-04-01" AND State = "Maharashtra"
16  -- Result: table StateCovidDetails has no column named District_Names
20   Insert into StateCovidDetails(ID,District_Name,Confirmed_Case,Recovered_Case,Deceased)
21   Select District_Key,District,Confirmed,Recovered,Deceased
22   from Covid_Raw_Data
23   WHERE Dates = "2021-04-01" AND State = "Maharashtra"
24  -- Result: no such column: Dates
28   Insert into StateCovidDetails(District_Name,Confirmed_Case,Recovered_Case,Deceased)
29   Select District,Confirmed,Recovered,Deceased
30   from districts
31   WHERE Dates = "2021-04-01" AND State = "Maharashtra"
32  -- Result: no such column: Dates
36   Insert into StateCovidDetails(District_Name,Confirmed_Case,Recovered_Case,Deceased)
37   Select District,Confirmed,Recovered,Deceased
38   from districts
39   WHERE Date = "2021-04-01" AND State = "Maharashtra"
40  -- Result: query executed successfully. Took 40ms, 36 rows affected
44   Insert into StateCovidDetails(Active_Case)
45   select Confirmed_Case-Recovered_Case-Deceased
46   from StateCovidDetails
47  -- Result: query executed successfully. Took 0ms, 36 rows affected
51   Insert into StateCovidDetails(Active_Case)
52   select Confirmed_Case-Recovered_Case-Deceased
53   from StateCovidDetails
54  -- Result: query executed successfully. Took 0ms, 72 rows affected
58  UPDATE StateCovidDetails(
```