THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

DEPARTMENT OF LAND SURVEYING AND GEO-INFORMATICS
土地測量及地理資訊學系

2021-2022 Semester 2

LSGI3315: GIS Engineering

(Previous Course Name: LSGI3431A: System Customization and Development)

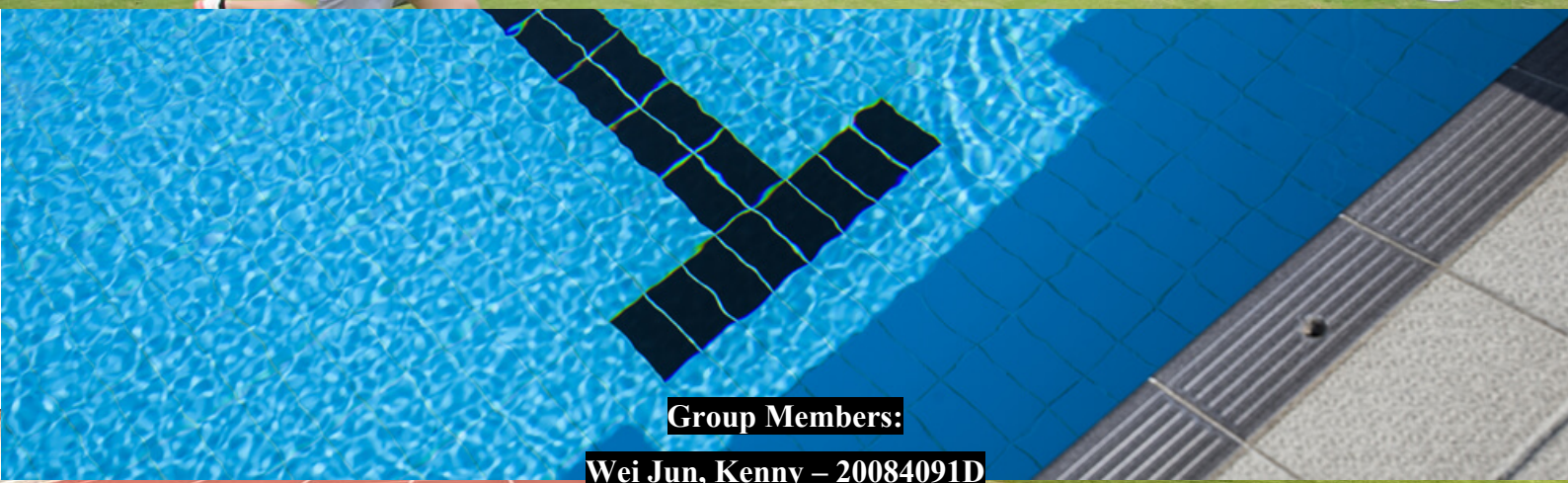
Subject Lecturer: Dr. Xu Yang

Tutor: Ms. Xinyue Chen and Ms. Mengqi Sun

Group 6 Report

Project Topic:

An Exploration of Spatial Analysis of Sports and Outdoor Facilities in Hong Kong via ArcPy



Group Members:

Wei Jun, Kenny – 20084091D

Tang Justin Hayse Chi Wing G. – 20016345D

Abstract

A variety of sports and outdoor facilities can enrich human's quantity of life and improve human physical and mental health. In this study, we explored the sports and outdoor facilities in Hong Kong through geospatial aspect and there are some key findings on it. At first, a majority of sports and outdoor facilities in Hong Kong are located in the tourist attractions, residential areas and new towns in Hong Kong. In addition, using badminton court as a specific facility type, a majority of the urban areas are within a reasonable walking distance (i.e., 500 m), such as the Western and Eastern part of Hong Kong Island, the Kowloon Peninsula and new towns in the New Territories. Also, the good coverage of badminton courts in Hong Kong are situated in Kwun Tong , Wong Tai Sin and Yau Tsim Mong District. Last but not least, a majority of the sports facilities are highly accessible in transportation within 100 meters, except country park. The group project is meaningful for students to perform various customized spatial analysis via Python programming.

Keywords: ArcPy, Python, ArcGIS, GIS, Spatial Analysis, Sports and Outdoor Facilities, Hong Kong, Geovisualization

1 – Background and Framework

To complete the Final Group Project of LSGI3315, we are required to develop a series of Python Functions and utilize the ArcPy syntaxes, so as to perform different spatial analysis for the sports and outdoor facilities in Hong Kong through geospatial perspective. The given sports and outdoor facilities data includes badminton courts, basketball courts, fitness centres, parks and gardens, sports grounds, swimming pools and country parks.

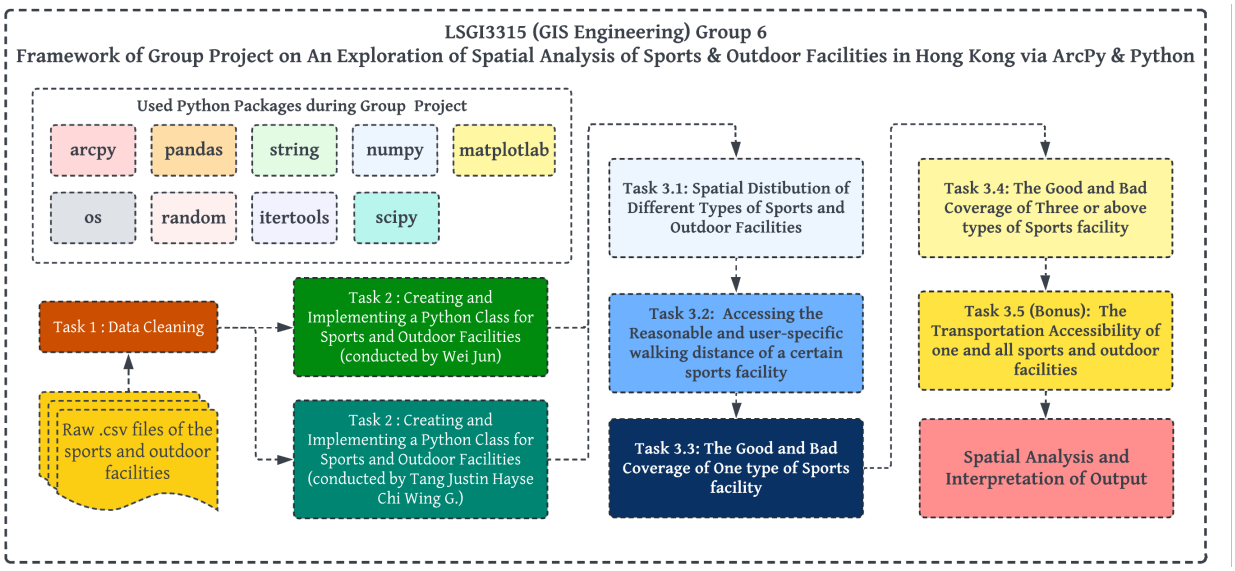


Figure 1. Framework of LSGI3315 Group Project.

Figure 1. shows the framework of our group work. The project contains three major tasks, namely Task 1 - data cleaning, Task 2 - creating and implementing a Python Class for the sport and outdoor facilities, and Task 3 - undertaking a series of spatial analysis to the facilitates (i.e., spatial distribution, buffer analysis and surface coverage analysis). In addition, our group have conducted the bonus part regarding accessing the performance of the transportation accessibility for all sports and outdoor facilities by using the line feature shapefile provided by the Transport Department, the HKSAR.

Figure 2 - 4. Different Types of Sports and Outdoor Facilities in Hong Kong.

2 - Preparation: Data Cleaning

Prior to conduct a series of spatial analysis through ArcPy and other Python package, data cleaning is of significance to select some important information and remove some unwanted data. The given .csv files contains GMID, facility names in Chinese and English, address, pairs of easting and northing in HK1980 Grid Coordinate System, pairs of latitude and longitude in WGS84 Coordinate System and the last updated information, website, email address, . A python code for the data cleaning is developed, so as to keep the important information for the spatial analysis. After data cleaning, Task 2 can be started. To be mentioned, Task 2 is conducted individually, therefore the our reports of Task 2 are separated and submitted individually.

```python
# LSGI3315 Group 6's Project - Task 1: Data Cleaning
# Group mate 1: Wei Jun, Kenny - 20084091D
# Group mate 2: Tang Justin Hayse Chi Wing G. - 20016345D
import pandas as pd
import os

print("========================================= Task 1: Data Cleaning =========================================")
data_folder = r"C:\Users\justi\Downloads\LSGI3315_Gp6_20016345D"  # working directory containing .csv files is set
os.chdir(data_folder)  # Changing the working directory
files = os.listdir(os.getcwd())  # receive all .csv files inside the folder
files = list(filter(lambda f: f.endswith('.csv'), files))  # find all ".csv" files in folder

out_path = r"C:\Users\justi\Downloads\LSGI3315_Gp6_Output_Files"  # Path to output directory

for file in files:  # Create a for-loop to loop through the data in .csv file
    df = pd.read_csv(file)  # Reading .csv files using pandas dataframe
    # subset the information which are kept
    df_subset = df[["Dataset", "Facility Name", "Address", "District", "Northing", "Easting", "Latitude", "Longitude"]]
    df_subset.to_csv(out_path + file)  # Save the data frames as .csv file as output
print("==================================== Task 1: Data Cleaning is completed! ====================================")
```

Figure 5. Python Code for Data Cleaning (Task 1).

	A	B	C	D	E	F	G	H	I	J
1	ENGLISH CATEGORY	中文類別	ENGLISH NAME	中文名稱	ENGLISH ADDRESS	中文地址	LONGITUDE	經度	LATITUDE	緯度
2	Badminton Courts	羽毛球場	North Kwai Chur	北葵涌鄧	292 Wo Yi Hop Roa	葵涌和宜合路292號	114-8-15	114-8-15	22-22-26	22-22-26
3	Badminton Courts	羽毛球場	Wong Nai Chung	黃泥涌體	4/F,Wong Nai Chun	跑馬地毓秀街2號黃泥	114-11-9	114-11-9	22-16-8	22-16-8
4	Badminton Courts	羽毛球場	Ap Lei Chau Spo	鴨脷洲體	8 Hung Shing Street,	鴨脷洲洪聖街8號	114-9-20	114-9-20	22-14-40	22-14-40
5	Badminton Courts	羽毛球場	Tsuen Wan Wes	荃灣西約	68 Hoi On Road, Tsu	荃灣海安路68號	114-6-2	114-6-2	22-22-14	22-22-14
6	Badminton Courts	羽毛球場	Fat Kwong Stree	佛光街體	18 Good Shepherd S	何文田牧愛街18號	114-10-52	114-10-52	22-18-53	22-18-53
7	Badminton Courts	羽毛球場	Tai Wo Hau Spo	大窩口體	39 Tai Ha Street, Ta	葵涌大窩口大廈街39	114-7-30	114-7-30	22-22-13	22-22-13
8	Badminton Courts	羽毛球場	Yau Oi Sports Ce	友愛體育	3 Hing On Lane, Tue	屯門興安里3號	113-58-18	113-58-18	22-23-8	22-23-8
9	Badminton Courts	羽毛球場	To Kwa Wan Spo	土瓜灣體	66 Ha Heung Road,	土瓜灣下鄉道66號	114-11-24	114-11-24	22-19-8	22-19-8
10	Badminton Courts	羽毛球場	Lei Yue Mun Spo	鯉魚門體	2/F to 5/F of Lei Yue	九龍油塘鯉魚門市政	114-14-19	114-14-19	22-17-31	22-17-31
11	Badminton Courts	羽毛球場	Ping Shan Tin Sh	屏山天水	1, Tsui Sing Road, Ti	元朗屏山天水圍聚星	114-0-17	114-0-17	22-26-51	22-26-51
12	Badminton Courts	羽毛球場	Canton Road Pla	廣東道遊	176 Canton Road, K	九龍廣東道176號	114-10-6	114-10-6	22-18-4	22-18-4
13	Badminton Courts	羽毛球場	Shek Kip Mei Par	石硤尾公	290 Nam Cheong St	石硤尾南昌街290號	114-10-12	114-10-12	22-20-14	22-20-14
14	Badminton Courts	羽毛球場	Queen Elizabeth	伊利沙伯	18 Oi Kwan Road, W	香港灣仔愛群道18號	114-10-44	114-10-44	22-16-31	22-16-31
15	Badminton Courts	羽毛球場	Tsing Yi Sports C	青衣體育	2/F, Tsing Yi Municip	青衣青綠街38號青衣	114-6-23	114-6-23	22-21-15	22-21-15
16	Badminton Courts	羽毛球場	Ma Mei Ha Playg	馬尾下遊	Ma Mei Ha, N.T.	新界馬尾下	114-10-37	114-10-37	22-31-6	22-31-6
17	Badminton Courts	羽毛球場	Tai Tau Leng Sitt	大頭嶺休	Tai Tau Leng Village	新界上水寶石湖路大	114-7-24	114-7-24	22-30-1	22-30-1
18	Badminton Courts	羽毛球場	Aberdeen Sports	香港仔體	6/F Aberdeen Munic	香港仔大道203號香	114-9-16	114-9-16	22-14-58	22-14-58
19	Badminton Courts	羽毛球場	Hang Hau Sports	坑口體育	1-3/F, Sai Kung Tseu	將軍澳培成路38號西	114-16-6	114-16-6	22-19-2	22-19-2
20	Badminton Courts	羽毛球場	Choi Hung Road	彩虹道羽	Sheung Hei Street, W	黃大仙雙喜街	114-11-49	114-11-49	22-20-15	22-20-15

Figure 6. Before Data Cleaning (Using badminton_court.csv as an example).

	A	B	C	D	E	F	G	H
1	Dataset	Facility Name	Address	District	Northing	Easting	Latitude	Longitude
2	Badminton Courts	Aberdeen Sports Centre	6/F Aberdeen Municipal Servi	SOUTHERN	812311	833947	22.24956	114.15436
3	Badminton Courts	Ap Lei Chau Sports Centre	8 Hung Shing Street, Apleicha	SOUTHERN	811748	834056	22.24448	114.15542
4	Badminton Courts	Boundary Street Sports Centre Na	200 Sai Yee Street, Mong Kok	YAU TSIM MONG	820730	835654	22.32559	114.17091
5	Badminton Courts	Boundary Street Sports Centre Na	200 Sai Yee Street, Mong Kok	YAU TSIM MONG	820784	835614	22.32608	114.17053
6	Badminton Courts	Bowen Road Temporary Playgrou	Bowen Road, Wan Chai, Hon	WAN CHAI	814730.2	836058.5	22.27141	114.17484
7	Badminton Courts	Canton Road Playground	176 Canton Road, Kowloon	YAU TSIM MONG	818015.1	835391.5	22.30108	114.16837
8	Badminton Courts	Chai Wan Sports Centre	6 Yee Shun Street, Chai Wan	EASTERN	814005	842765	22.26485	114.23991
9	Badminton Courts	Che Kung Temple Sports Centre	No. 10, Sha Tin Tau Road, Sha	SHA TIN	825894	837164	22.37223	114.18557
10	Badminton Courts	Cheung Chau Sports Centre	Nam She Tong, Cheung Chau	ISLANDS	807661	821245	22.20751	114.03118
11	Badminton Courts	Cheung Fat Sports Centre	4/F, Cheung Fat Shopping Cer	KWAI TSING	824815.8	828627.9	22.36247	114.1027
12	Badminton Courts	Cheung Sha Wan Sports Centre	J/O Hing Wah Street & Cheun	SHAM SHUI PO	822100	833857	22.33797	114.15347
13	Badminton Courts	Choi Hung Road Badminton Cent	Sheung Hei Street, Wong Tai	WONG TAI SIN	822052	838341	22.33753	114.197
14	Badminton Courts	Choi Hung Road Sports Centre	Sheung Hei Street, Wong Tai	WONG TAI SIN	822007	838285	22.33713	114.19645
15	Badminton Courts	Chuk Yuen Sports Centre	Chuk Yuen North Estate, Chu	WONG TAI SIN	822937	837982	22.34552	114.19351
16	Badminton Courts	Chun Wah Road Sports Centre	Top Floor, Lok Nga Court Car	KWUN TONG	820385	840764	22.32247	114.22051
17	Badminton Courts	Fa Yuen Street Sports Centre	13/F Fa Yuen Street Municipa	YAU TSIM MONG	820197	835632	22.32078	114.1707
18	Badminton Courts	Fat Kwong Street Sports Centre	18 Good Shepherd Street, Ho	KOWLOON CITY	819538	836704	22.31483	114.18111
19	Badminton Courts	Fu Heng Sports Centre	1/F, Fu Heng Shopping Centre	TAI PO	835424	835700	22.45829	114.17136
20	Badminton Courts	Fu Shin Sports Centre	Multi Car-Park Building, Fu Sh	TAI PO	834936	835984	22.45388	114.17412

Figure 7. After Data Cleaning (Using badminton_court.csv as an example).

3 - Methodology

3.1 –Question 1 (Core): Spatial Distribution of Different Types of Sports and Outdoor Facilities

3.1.1 Workflow for Python Programming

For the first question, we are asked to explore the spatial distribution of various sports and outdoor facilities via ArcPy. In order to familiar with the spatial distribution of all types of sports and outdoor facilities, we first define the spatial reference by arcpy.SpatialReference() and create a fishnet arcpy.management.CreateFishnet() in which covers the whole Hong Kong. Afterwards, using the point data stored in Python List (done in Task 2) to conduct the spatial join function arcpy.analysis.SpatialJoin(), so as to import all points of sports and outdoor facilitates into the specific width of cells in the fishnet. The fishnet contains a number of join-count from all point data and it will be saved in the attribute table. Eventually, the join-count data can be geo-visualized and displayed as square cells in ArcGIS Pro.

To be remarked, as the square cells will generate excess from the continent to the sea, and this phenomenon is in the reality, thus an additional ArcPy function (arcpy.analysis.clip) is adopted, so as to remove the cells located in the sea. Considering the highest number of join-count is seven, thus the join-count data is ranging from one to seven.

Figure 8. illustrates the workflow of Function SpatialDistribution_Facilities () and it will be discussed as follows.

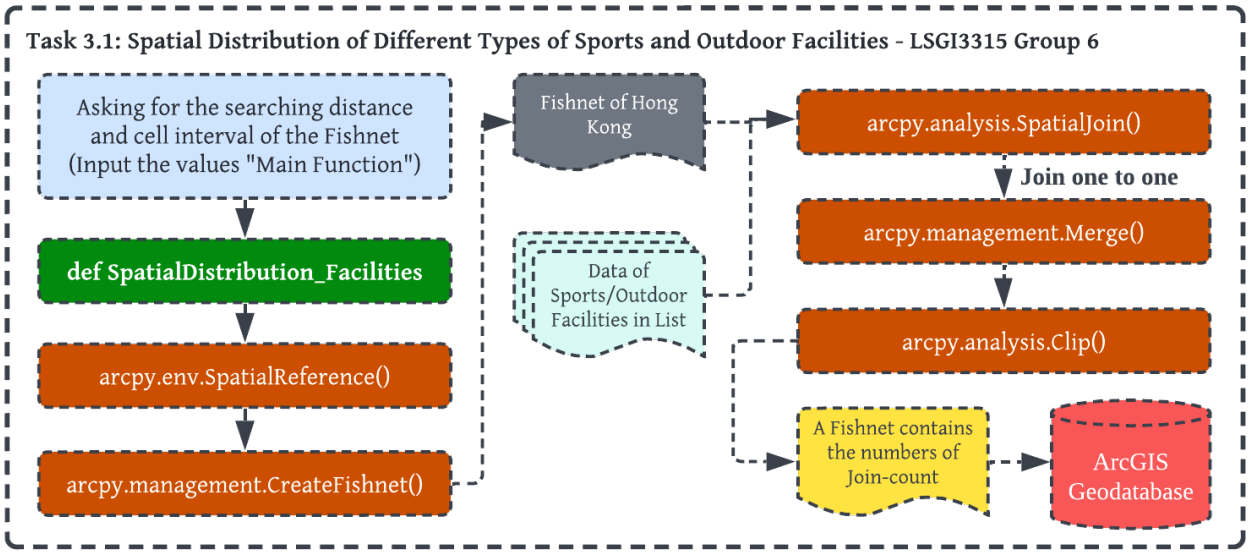


Figure 8. Workflow of Task 3.1: Spatial Distribution of Sports and Outdoor Facilities in Hong Kong .

3.1.2 Progress of Python Code

The below screen captures shows the programming mission of Question 1 conducted by various ArcPy syntaxes.

```
206  # Task 3.1: Spatial distribution of Difference types of sports and outdoor facilities
207  def SpatialDistribution_Facilities(searching_distance, cell_interval):
208      inputted_distance = searching_distance  # Save the user input of searching distance
209      inputted_interval = cell_interval  # Save the user input of cell interval for fishnet
210      temp_list = []  # Create an Empty List
211
212      #  Using HK1980 Grid Coordinate System to create a Fishnet is easier because it use meter as the unit
213      spatial_reference = arcpy.SpatialReference(2326)  # EPSG 2326 = HK1980 Grid Coordinate System
214      arcpy.env.outputCoordinateSystem = spatial_reference  # set this coordinate system temporarily
215
216      # Create a Fishnet of rectangular cells by inputting the essential parameters
217      arcpy.management.CreateFishnet("fishnet", '801010.096208300 801656.030632100', '801010.096208300 801666.030632100',
218                                     inputted_interval, inputted_interval, '', '', '863615.312086720 846943.245240580',
219                                     'LABELS', '', '')
220
221      arcpy.Delete_management("fishnet")
222
223      # Input the data from the List
224      fc_list = ['Badminton_court', 'Basketball_court', 'country_parks', 'fitness_center',
225                 'parks_gardens', 'sports_grounds', 'swimming_pools', 'other_recreation_sports_facilities']
226
```

Figure 9. Python Code for Task 3.1: Spatial Distribution of Sports and Outdoor Facilities in Hong Kong.

```python
    # Create a For-loop for looping all types of sports and outdoor facilities
    for index, value in enumerate(fc_list):

        # Define the output coordinate system again. To be remarked, the coordinate system will be changed at the end.
        arcpy.env.outputCoordinateSystem = spatial_reference

        # Using Spatial Join Function to let the point data import into the specific width of cells
        arcpy.analysis.SpatialJoin("fishnet_label", str(value), "fishnet_label_temporary" + str(index), 'JOIN_ONE_TO_MANY',
                                   'KEEP_COMMON', '', 'WITHIN_A_DISTANCE', str(inputted_distance) + ' meters')

        arcpy.DeleteIdentical_management("fishnet_label_temporary" + str(index), "TARGET_FID")

        # Append the temporary output
        temp_list.append("fishnet_label_temporary" + str(index))

    # Merge the temporary output
    arcpy.management.Merge(temp_list, 'temp_merged')
    # Using Spatial Join Function to let the point data import into the specific width of cells
    arcpy.analysis.SpatialJoin("fishnet_label", 'temp_merged', "fishnet_final", 'JOIN_ONE_TO_ONE', 'KEEP_COMMON', '',
                               'INTERSECT', '', '')

    # Remove the cells located in the sea or irrelevant areas in Hong Kong
    arcpy.analysis.Clip("fishnet_final", HK_shapefile, "clipped_fishnet_final")
```

Figure 10. Python Code for Task 3.1 (Cont.).

3.2 - Questions 2 - 4 (Core): Given a specific type of sports and outdoor facility, which areas are within reasonable walking distance and within user-specified distance? Which areas have a good coverage of a type of Facilities?

3.2.1 Workflow for Python Programming: Reasonable Walking Distance

Indeed, the compulsory questions regarding the reasonable walking distance, user-specific distance, and good coverage of a type of sports and outdoor facilities can be programmed simultaneously by using ArcPy syntaxes and nested for-loop. Therefore, several missions are combined and the methodological framework is shown below. In the framework, it is conspicuous that there are three defined functions in green colour. The function Buffer_analysis_One () is used to reveal the reasonable and user-specificed walking distance of one type of sports and outdoor facilities. Instead, the functions of Intersect_analysis_and_calc_parm () and copyResultToXzq () are utilized to find the good coverage.

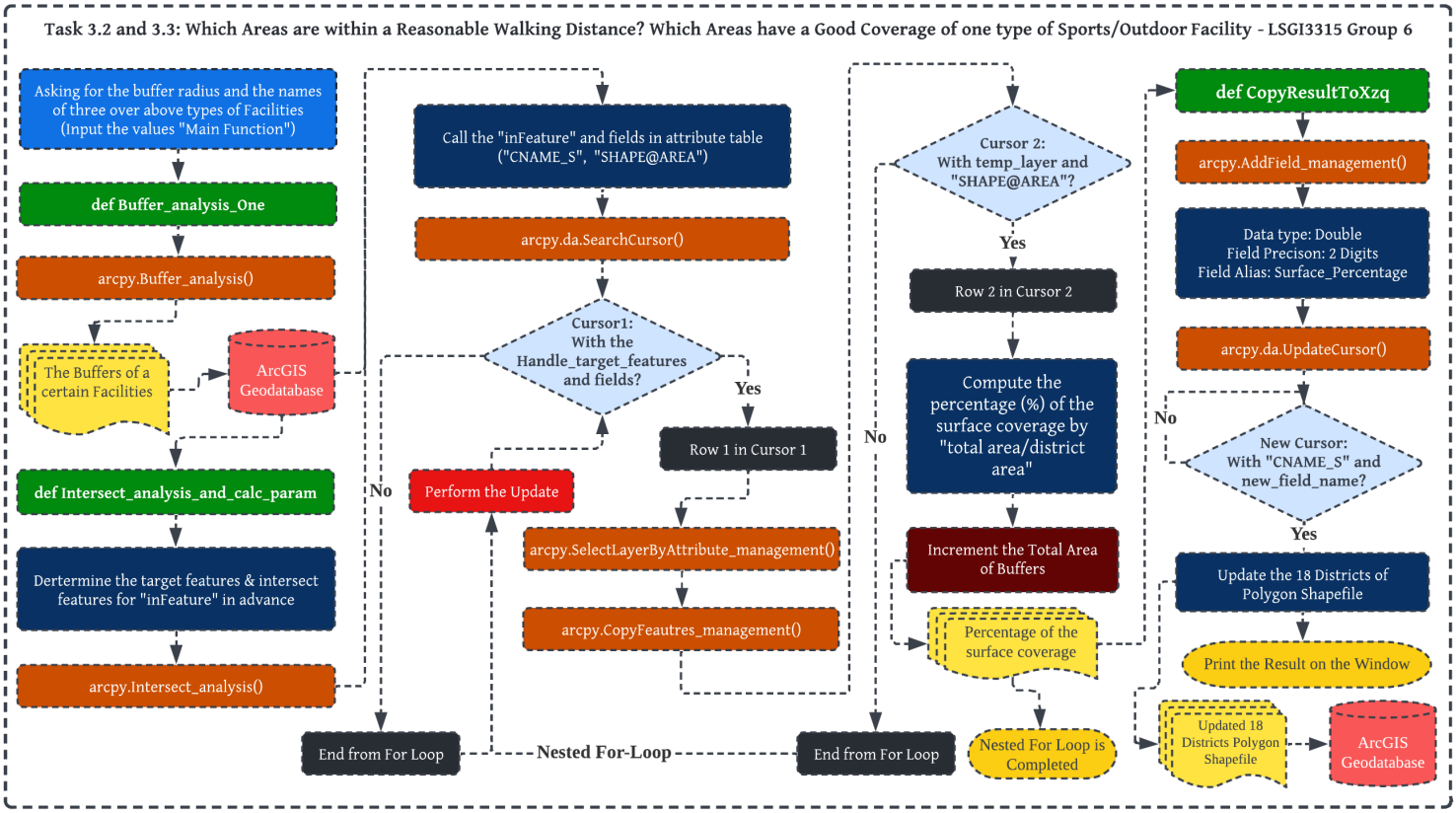


Figure 11. Workflow of Task 3.2-3.3: Reasonable Walking Distance and Good Coverage of a type of Sports Facilities.

To commence with the buffer analysis, the program is started by calling the function named Buffer_analysis_One () and using arcpy.Buffer_analysis() to create buffer areas for the chosen facility. In terms of the buffer distance, it can be determined as one of the user-defined parameters of the function. The radius of one kilometre is adopted because this distance takes around **15 to 20 minutes on foot**, which is acceptable and reasonable to define the user-specific distance. The dissolve option (a parameter) is also user-defined. However, it should be written as true most of time, to remove the overlapping areas of different buffer zones.

The buffer zones are generated successfully, but the buffer areas are not reliable enough for computing surface coverage because the buffers may include the areas where located in other districts or sea and thus exaggerate the surface coverage. As a consequence, intersect analysis is of significance to remove the impractical areas. The function named Intersect_analysis_and_calc_parm () is responsible for clipping the data and compute the surface coverage by using Hong Kong eighteen districts polygon shapefile as well as arcpy.Intersect_analysis() syntax. To be remarked, the good coverage will be computed by total the total buffer areas (surface area) divided by district areas. The original buffer zones in blue color and buffer zones after intersecting analysis in yellow colour are shown in Figure 12. To visualize the coverage of the buffer zones with different radius, in Figure 13., four buffer zones with radii ranging from 500m, 1000m, 1500m to 2000m are generated by simply change the value from the main function.

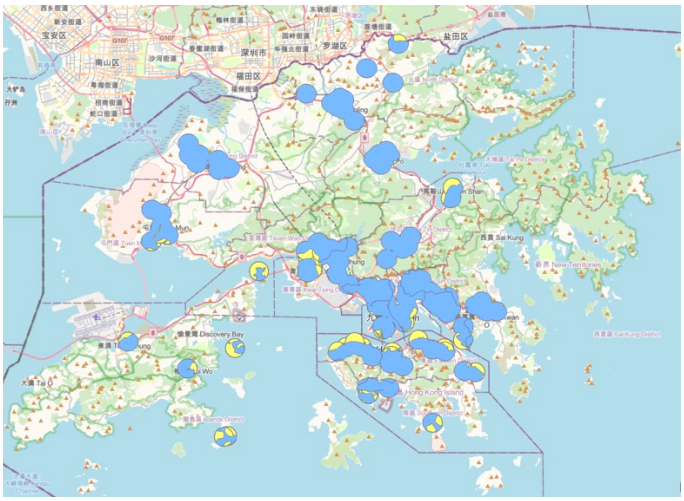


Figure 12 (Left). Impractical Buffer areas are shown in Yellow Colour.
Figure 13 (Right). Buffer Analysis of Badminton Courts using 500, 1000, 1500 and 2000 meters.

3.2.2 Workflow for Python Programming: Coverage Analysis of One Type of Facility

After generating the buffers and clipping the unreasonable parts of buffers, the surface coverage of 18 districts can be computed by using syntaxes of arcpy.da.SearchCursor() and arcpy.SelectLayerByAttribute_management(). In particular, a **Nested For-loop** (containing two for-loops at the same time) is developed, to compute the total area of buffers and increment the total area in each district. For the area of buffers located in the same district, its value of area will be incremented and compute the its surface coverage is computed by Formula 1:

$$surface\ coverage = \frac{area\ of\ buffers}{area\ of\ district} \quad (1)$$

The output of the proportion will be returned to the dictionary. Subsequently, by calling function copyResultToXzq (), the result will be saved into the field named "*Surface_percentage(%)*" in the given polygon shapefile. In order to easily distingulish the shapefile, this function initially copies the feature using arcpy.CopyFeatures_management(). Then, arcpy.Addfield_management() and arcpy.da.UpdateCursor() are used to create and update a field for storing the proportion of the surface coverage.To be remarked, the source of shapefile is the given "*HKDistrict18.shp*"; a new shapefile named "*Hong Kong.shp*" will be generated and saved into the temporary ArcGIS Geodatabase. The data can be viewed in the attribute table in ArcGIS Pro and geo-visualized through three-dimensional perspective. The results will be discussed in Section 4.3.

3.2.3 Progress of Python Code

The Python code shows the three functions to explore the reasonable walking distance and good coverage of one facility.

```python
# Conduct the Simple Buffer Analysis for a type of Sports and Outdoor Facility
def buffer_analysis_One(in_features, out_feature_class, buffer_distance, dissolve_option=None):
    try:
        handle_in_features = project_data(in_features)
        # Conducting the buffer analysis and the same set of values in their Dissolve Field(s) will be dissolved together.
        arcpy.Buffer_analysis(handle_in_features, out_feature_class, buffer_distance, dissolve_option=dissolve_option)
        arcpy.Delete_management(handle_in_features)
    except Exception as result:
        raise Exception(result)
```

Figure 14. Function of Buffer_analysis_One ( )

```python
# Task3.2 (2): Conduct the Intersect Analysis for the Buffers which are currently generated
def Intersect_analysis_and_calc_param(target_features, intersect_features):
    try:
        # Processing the data by using intersect tool,
        # thus the surface areas which are outside the corresponding district will be removed.
        handle_target_features = project_data(target_features)
        handle_intersect_features = project_data(intersect_features)
        # Save the target and intersect features in the "inFeature"
        inFeatures = [handle_target_features, handle_intersect_features]
        # Create the output name from the intersect analysis
        intersect_output = "task_two_intersect_result"
        # Conduct the intersect analysis
        arcpy.Intersect_analysis(inFeatures, intersect_output)
        print("Intersect analysis completed.")
        rDict_double = {}

        # Compute the percentage by surface area(buffer)/District area
        fields = ["CNAME_S", "SHAPE@AREA"]  # Create a field contains name and the area of buffers
        with arcpy.da.SearchCursor(handle_target_features, fields) as cursor:
            #  Create a Nested For-loop to compute the total area of buffers
            for row in cursor:  # Create a First for-loop to scan the attribute table
                xzqmc = row[0]
                total_area = 0
                # Select the information from the attribute
                temp_layer = arcpy.SelectLayerByAttribute_management(intersect_output, "NEW_SELECTION",
                                                                     '"CNAME_S" = ' + "'" + xzqmc + "'")
                arcpy.CopyFeatures_management(temp_layer, xzqmc + "_tempLayer")
                # Find the "SHAPE@AREA" from the attribute
                with arcpy.da.SearchCursor(temp_layer, ["SHAPE@AREA"]) as cursor2:
                    for row2 in cursor2:  # Create a second for-loop to increment the total area
                        total_area += row2[0]  # add the total area of buffers

                # The formula to compute the surface coverage
                rDict_double[xzqmc] = (total_area / row[1]) * 100
                arcpy.Delete_management(xzqmc + "_tempLayer")
        arcpy.Delete_management(handle_target_features)
        arcpy.Delete_management(handle_intersect_features)
        print("Buffer analysis completed.")
        return rDict_double  # return the output of the surface coverage

    except Exception as result:
        raise Exception(result)
```

Figure 15. Function of Intersect_analysis_and_calc_parm ( )

```python
# Task3.2 (3): Copy the results to The Shapefile of Hong Kong 18 Districts.
def copyResultToXzq(new_field_name, source_layer, result_layer, dict):
    try:
        arcpy.CopyFeatures_management(source_layer, result_layer)
        # Add the output surface coverage to the new field "Surface_percentage"
        arcpy.AddField_management(result_layer, new_field_name, "DOUBLE", 2, field_alias="Surface_percentage(%)")
        with arcpy.da.UpdateCursor(result_layer, ["CNAME_S", new_field_name]) as cursor:
            # Create a for loop to loop the surface percentage to all districts
            for row in cursor:
                layer_xzqmc = str(row[0])
                row[1] = float(dict[layer_xzqmc])
                cursor.updateRow(row)  # update the row
```

Figure 16. Function of copyResulttoXzq( )

3.3 - Questions 5: Which areas have a good coverage of three or above Sports and Outdoor Facilities?

3.3.1 Workflow for Python Programming

The fourth question is similar to Task 3.3 while the requirement is to find the good coverage of three or above sports and outdoor facilities. Therefore, we adopted buffer analysis again and further develop some sophisticated functions on the Python code. At the beginning, the user can input the value of buffer radius and input three or above names of sports and outdoor facilities in the Main Function. Then, based on the customized input, the python code will be executed.

To explicitly distinguish the functions, the functions of buffer analysis and multiple facilities are separated. After the user typed the customized buffer radius and the three (or above) names of sports and outdoor facilities, the function will initially conduct the buffer analysis arcpy.Buffer_analysis and arcpy.MultipartToSinglepart_Management for the three (or above) facilities by calling the buffer analysis function. Subsequently, this function will only keep the buffers which intersect with three buffers from three facilities using arcpy.SelectLayerByAttribute_management and arcpy.CopyFeatures_management syntax. The output layers will be recorded as "XXXXXXXXX(Facility Name)_Filtered". To explore further, each intersect of the remained buffers (intersected by the three facilities) have a number of heterogeneous intersections stored in attributed table, and it can represent as whether the buffers which intersect the most and intersect the less. Therefore, based on the number of intersections, we can determine the where are the good coverages containing three or above facilities. It uses defined interval method for the coverage classification.

Table 1. Classification of The Degree of Coverage for Three or above Facilities

Interval of the Intersections	Degree of Coverage
0% - 20% of Data	Least Dense
20% - 50% of Data	Less Dense
50% - 80% of Data	Dense
80% - 90% of Data	Very Dense
90% - 100% of Data	The Densest

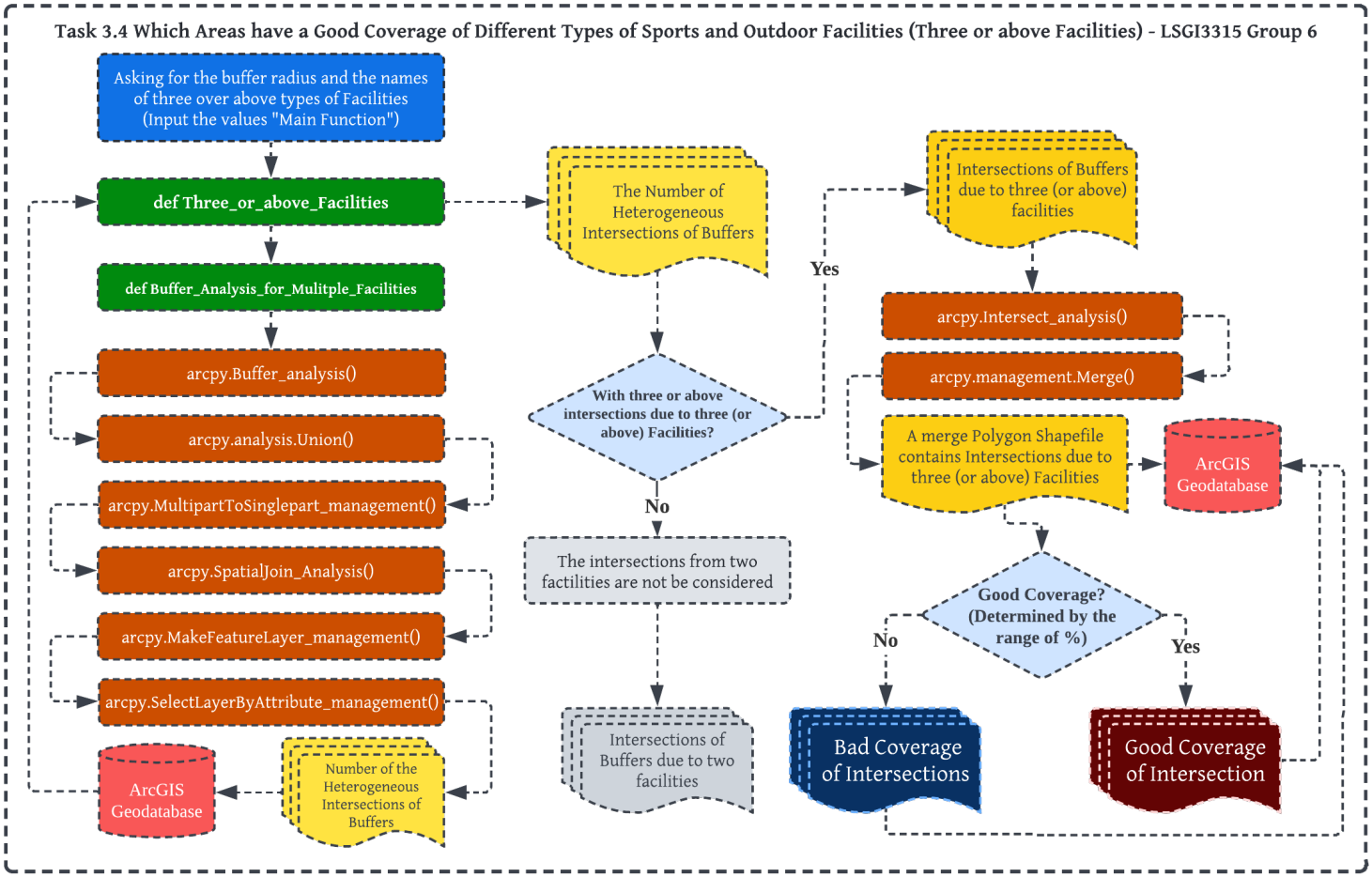


Figure 17. Workflow of Task 3.4:The Coverage Analysis of Three or Above Sports and Outdoor Facilities.

3.3.2 Progress of Python Code

The below screen captures shows the programming mission of Question 5 undertaken by different ArcPy syntaxes.

```
# Task 3.4: Which Areas have a Good Coverage of three types of above sports and facilities in Hong Kong?
def Three_or_above_Facilities(input_search_radius, input_facility_list):
    input_buffer_distance = input_search_radius
    input_buffer_distance = int(input_buffer_distance)  # set as integer data type
    types_of_facilities = input_facility_list

    print("Start to conduct the buffer analysis of types of sports and outdoor facilities.")
    # Create a for loop to loop through the point data for creating buffers
    for feature_class_in_list in types_of_facilities:
        Buffer_Analysis_for_Multiple_Facilites(input_buffer_distance, feature_class_in_list)  # Call Function
    print("The types of Buffer analysis (Called Function) are completed!")
    names_list = []  # Create an empty name list

    # Create a for loop to loop through the filtered buffers
    for temp_fac in types_of_facilities:
        name = temp_fac + "_join_filtered"  # add the "_filtered"
        names_list.append(name)  # Append the results

    three_or_above_facilities_list = []  # Create an empty list for three and above facilities
    for comb in combinations(names_list, 3):  # Default three types of facilities
        un_list = list(comb)  # record the for loop
        three_or_above_facilities_list.append(un_list)  # Append the result
```

```
    print("Start to conduct the buffer analysis of three types of sports and outdoor facilities.")
    individual_names_list = []  # Create a list for three or above types of feature coverage
    i = 1
    # Create a for loop for
    for combination in three_or_above_facilities_list:
        arcpy.Intersect_analysis(combination, "multiFeatureCoverageArea_" + str(i), "ONLY_FID")
        individual_names_list.append("multiFeatureCoverageArea_" + str(i))
        i = i + 1
    # Merge the three coverages of buffer into one merged feature coverage
    arcpy.management.Merge(individual_names_list, "multiFeatureCoverage")
    # Copy the merged feature to the ArcGIS Geodatabase
    arcpy.CopyFeatures_management(base_path + "\\Raw_Data\\Facility_data.gdb" + "\\multiFeatureCoverage",
                                  temp_gdb + "\\multiFeatureCoverage")
    i = 1
    # Create a for loop to remove the excess files
    for combination in three_or_above_facilities_list:
        arcpy.Delete_management("multiFeatureCoverageArea_" + str(i))
        arcpy.Delete_management("multiFeatureCoverageArea_clip" + str(i))
        i = i + 1
    print("The Coverage analysis of the three or above types of Sports and Outdoor Facilities is completed.")
    print("Please open the ArcGIS Geodatabase in ArcGIS Pro to check the Degree of Coverage.")
```

Figure 18. Function of Three_or_above_Facilities ( )

3.4 – Question 6 (Bonus Part): The Transportation Accessibility for all types of Sports and Outdoor Facilities

3.4.1 Workflow for Python Programming

Apart from the coverage of the facilities which is based on the reasonable walking distance, we are also interested in the transportation accessibility for a certain type of sports and outdoor facilities. With the aim of evaluating transportation accessibility, the geo-referenced road network shapefile of whole Hong Kong provided by the Transport Department, HKSAR is adopted, in order to evaluate the transportation accessibility of different types of sports and outdoor facilities. A map shown rightwards is the visualization of the transport network.

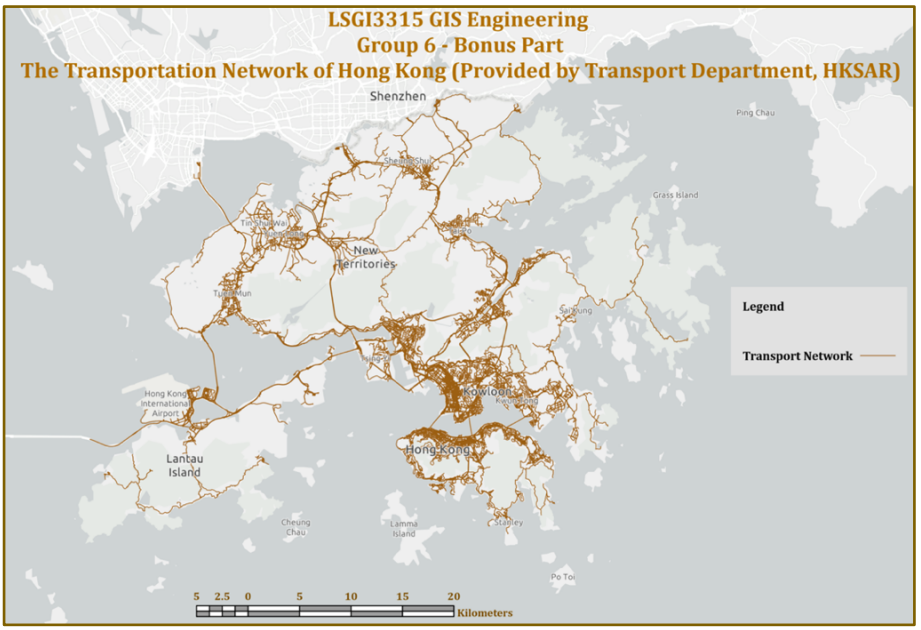


Figure 19. Transport Network in Hong Kong.

The additional function is developed and named **Traffic_access_analysis ()**. The user can input the searching distance and one type of facilities, to access its transportation accessibility in Hong Kong. Prior to perform the feature searching, the first step is to check the coordinate system of the input features by **arcpy.Describe(road_network)** and **If...Else Statement**, to determine the type of distance (either geodesic distance or projected distance) should be taken. The features will be selected based on the corresponding distance and use **arcpy.SelectLayerByLocation_management** to select the features. The features will be saved into another layer in the temporary GDB. Last but not least, using the syntax of **arcpy.GetCount_management()** to count the number of facilities within the searching distance and save the summation (sum) to the attribute table and display the output on the window. Further, the performance of all types of sports and outdoor facilities is undertaken, to compare the accessibly in a high-rise high-dense Hong Kong city.

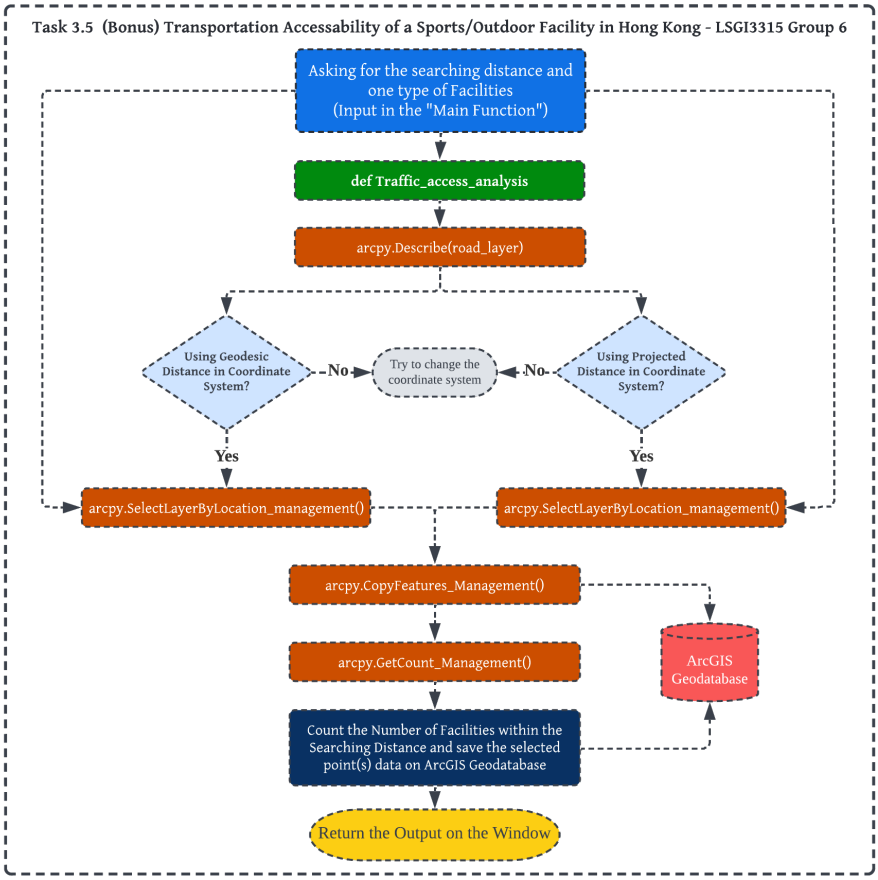


Figure 20. Workflow of the Bonus Part: Transportation Accessibility for a type of Sports and Outdoor Facilities.

3.4.2 Progress of Python Code

The below screen captures demonstrate the Function Traffic_access_analysis (), so as to perform the transportation accessibility in Hong Kong. An end-user can simply input the customized searching distance as well as one type of sports and outdoor facilities in the Main Function. The recommended searching distances are ranging from 20m to 500m for the facilities located in urban area because Hong Kong is a high dense city where contains a variety of sports facilities, not only indoor, but also outdoor.

```
# Task 3.5 (Bonus) Evaluating the Transportation Accessibility of one type of Sports/Outdoor Facilities
def traffic_access_analysis(road_layer, search_distance, facility_layer, analysis_result_layer):
    print("Searching the features of Sports and Outdoor Facilities...")
    try:
        # Check the coordinate system of the shapefile(geodesic coordinate system/projected coordinate system)
        desc = arcpy.Describe(road_layer)
        spatial_reference = desc.spatialReference  # Determine the Coordinate System

        # select the features within the user-defined distance
        if spatial_reference.type == u'Geographic':
            # use geodesic distance for the geodesic coordinate system
            temp_layer = arcpy.SelectLayerByLocation_management(facility_layer,
                                                                overlap_type='WITHIN_A_DISTANCE_GEODESIC',
                                                                select_features=road_layer,
                                                                search_distance=search_distance)
        else:
            # use projected distance for the projected coordinate system
            temp_layer = arcpy.SelectLayerByLocation_management(facility_layer, overlap_type='WITHIN_A_DISTANCE',
                                                                select_features=road_layer,
                                                                search_distance=search_distance)

        # Output the selected features to the temporary Geodatabase
        arcpy.CopyFeatures_management(temp_layer, analysis_result_layer)
        print("Searching features completed.")
        # Compute the coverage rate of the road network
        total_count = arcpy.GetCount_management(facility_layer)  # total number of a type of facility
        within_count = arcpy.GetCount_management(analysis_result_layer)  # number of the facility within search radius
        result = str(round(round(int(within_count[0]) / int(total_count[0]), 2) * 100)) + "%"  # formula of coverage rate
        return result
    except Exception as result:
        raise Exception(result)
```

Figure 21. Function of Traffic_access_analysis ( )

4 - Result

4.1 Spatial Distribution of Difference Types of Sports and Outdoor Facilities

The clipped join-count data is then saved in the ArcGIS geodatabase. The data can be inputted in the ArcGIS Pro and conduct some cartographic techniques, to geo-visualize the join-count data in an attractive and pleasant manner. The thematic map of spatial distribution of sports and outdoor facilities is produced using dark-mode and shown as follows.

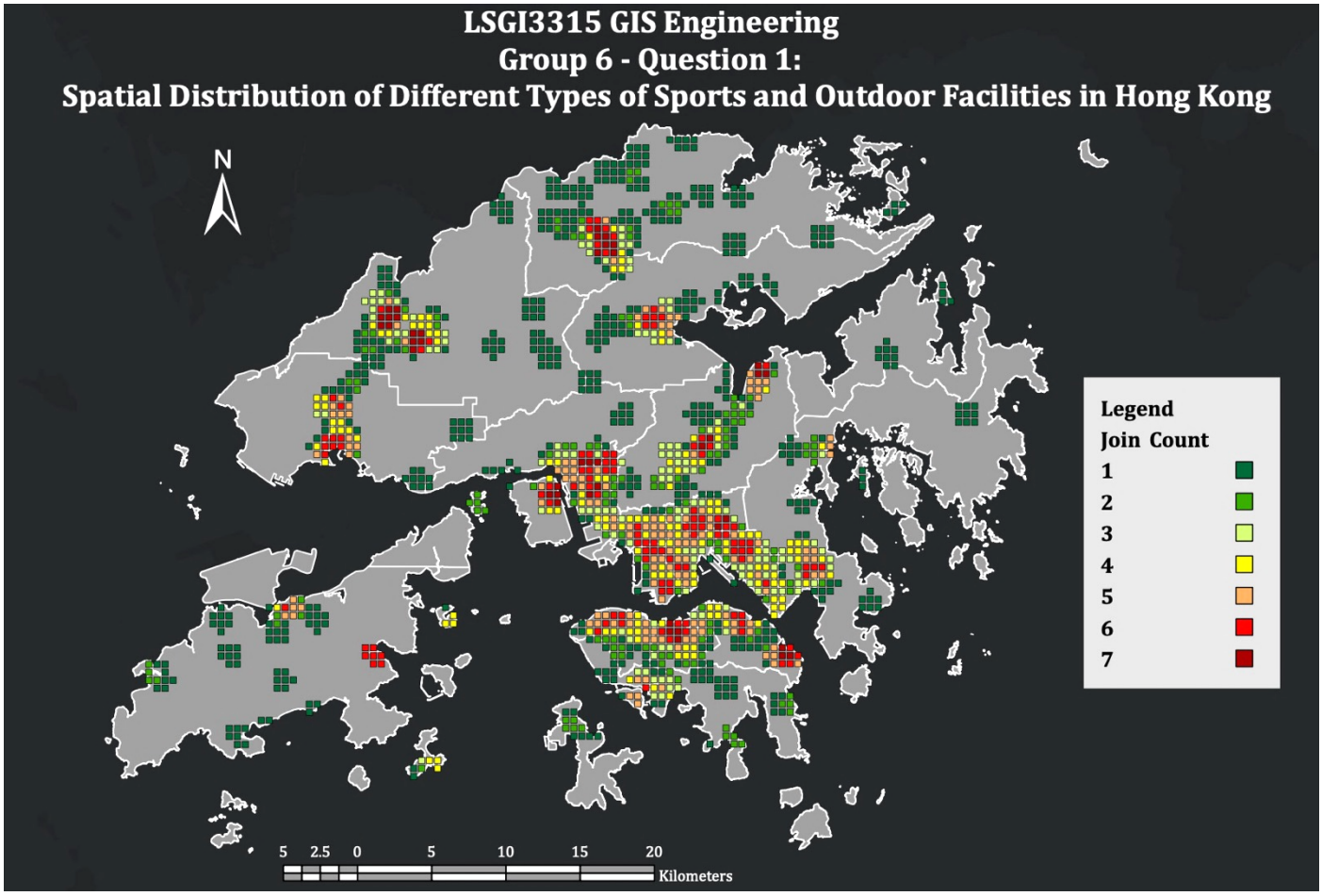


Figure 22. Spatial Distribution of Different Types of Sports and Outdoor Facilities in Hong Kong.

Spatial Analysis and Interpretation:

Figure 22. shows the spatial distribution of different types of sports facilities in Hong Kong. The map utilizes the join-count method on the fishnet, to explore which areas have the highest and lowest number of facilities. If an area does not contain any sports facilities, the square cells will be removed. The colour ranges from forest green (1), green (2), lime (3), yellow. (4), orange (5), red (6) to dark red (7), to represent the number of join-count in a pleasant manner.

By observation, a variety of sports and outdoor facilities spread across the built environment in Hong Kong. To discuss the areas with six to seven types of sports and outdoor facilities, the square cells shown in red and dark red are mostly located in the tourist hotspots (e.g., Causeway Bay, Cantral, Mong Kok, Tsim Sha Tsui, Penny's Bay), the new towns (e.g., Yuen Long, Tin Shui Wai, Tsuen Wan, Sha Tin, Ma On Shan, Kwai Chung, Tung Chung) and several residential areas (e.g., Fanling, Tai Po,Chai Wan, Quarry Bay). There are two possible reasons to explain why sports and outdoor facilites are mostly located in these areas. (1) Firstly, developing various sports and outdoor facilities in tourist hotspots can attract more locals and tourists, so as to gain the popularity easily. (2) Another reason is to satisfy the quality of life for Hong Kong residents. Different types of sports and outdoor facilities situated in residential areas nearby can promote a healthy lifestyle in the community. On the contrary, the least join-count areas are mostly in the rural area (country parks) or the places far from the city centre. It can be explained that developing various types of facilities in rural areas or the least densely populated places could be regarded as a waste of resources and resulting in low usage of facilities.

In overall, considering the **tourism management** and **improving the quality of life**, a majority of sports and outdoor facilities in Hong Kong are located in the **tourist attractions**, **residential areas** and **new towns in Hong Kong**.

4.2 Reasonable Walking Distance

In this section, it presents the map of accessing reasonable walking distance through buffer analysis of **badminton courts** in Hong Kong. A series of buffer zones are defined and generated 250 m, 500 m, 1000 m, 1500 m and 2000 m of radius, and the walking distances are represented in red, orange, yellow, sky blue and blue, respectively. Specifically, the area beyond the continent are removed.

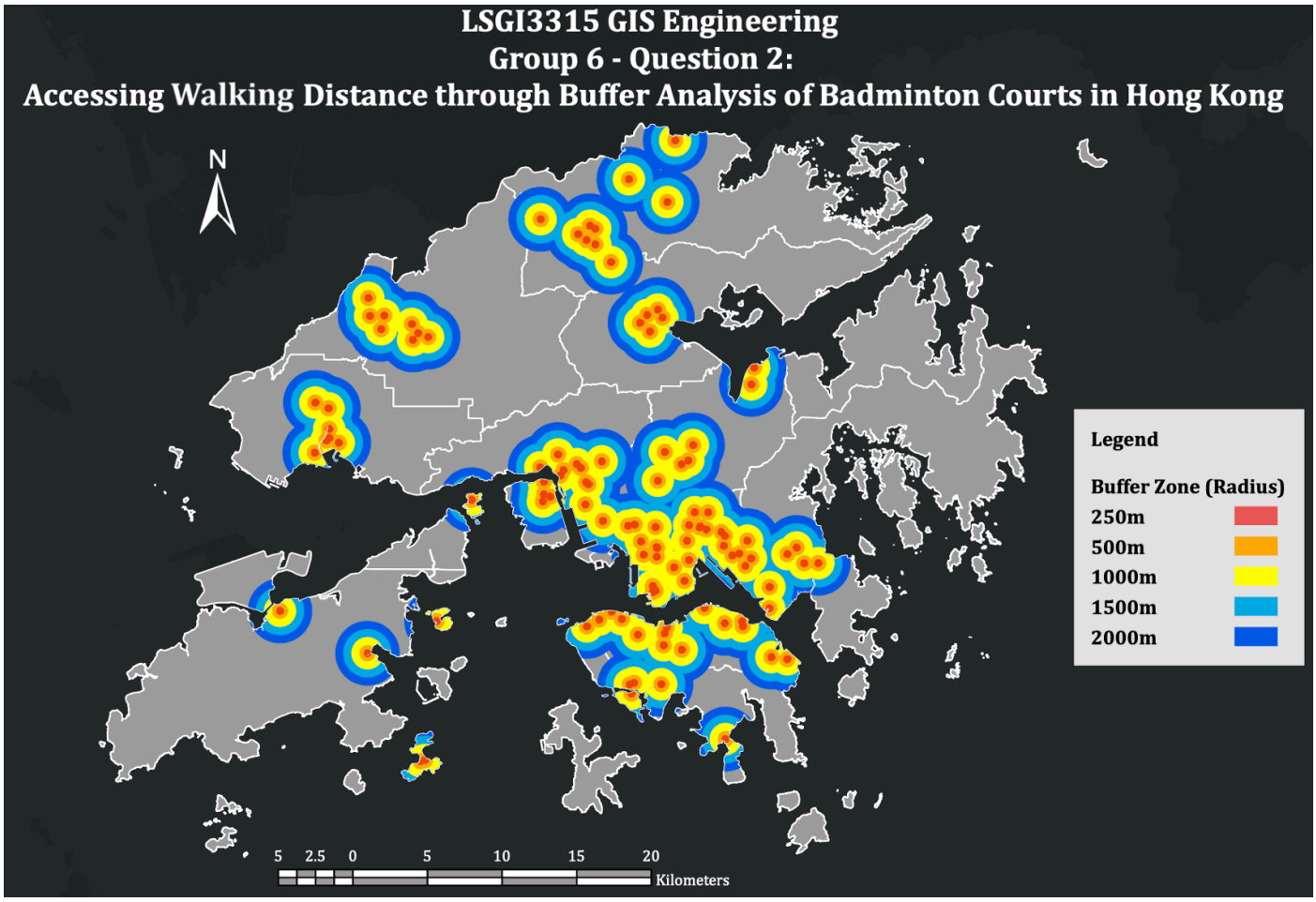


Figure 23. Accessing different walking distances of Badminton Courts through Buffer Analysis via ArcPy.

Spatial Analysis and Interpretation:

The second significant finding is related to the reasonable walking distance to a certain sports and outdoor facilitates, and the study selected the badminton court as the specific sports facility type.

To commence with the radius of 250-500 m, the shorter walking distance. It can be observed a large portion of the red and orange buffers are nearly stick with another badminton courts, particularly in Central, Sheung Wan, Wan Chai, Causeway Bay, Aberdeen and the Eastern District on Hong Kong Island; Mong Kok, Kwai Chung, Tsing Yi, Kwun Tong and Kowloon Bay in Kowloon; Yuen Long, Tuen Mun and Tai Po in the New Territories. The results suggest that the popular cities and the residential areas have a shorter walking distance and highly accessible to the badminton court in Hong Kong. On the other hand, for the longer walking distance (i.e., 1000 m) of badminton courts, they are mainly located and merged in both sides of Hong Kong Victoria Harbour and the Kowloon Peninsula. Several merged buffers are located in Yuen Long district, Tai Po District and Sheung Shui and Fanling in the New Territories. Nonetheless, the for the areas with the longest walking distances (i.e., 1500 m and 2000 m), these are considered as unreasonable and non-user-specific walking distance for people and it is highly recommended to take the modes of transport, rather than going to the badminton court on foot.

Therefore, the various buffers shown on the map of can be interpreted that a majority of the urban areas are within a reasonable walking distance (i.e., 500 m), such as the **Western and Eastern part of Hong Kong Island**, **nearly the whole coverage of Kowloon Peninsula** and several new towns in the New Territories (i.e., **Yuen Long District** and **Tuen Mun District**) can access to badminton court within around 500 to 1000 meters.

4.3 Coverage of One Type of Sports and Outdoor Facilities

For the third section, it aims at exploring which areas have a good coverage and bad coverage of a type of sports and outdoor facilities. Badminton court is taken as the specific sports and outdoor facility. As abovementioned, the good coverage is accessed by buffer analysis **in district level**. The below Figure 24. tabulated surface percentages computed by Formula 1 (See Section 4.2.2). In addition, the data of surface percentage is then geo-visualized through the state-of-the-art 3D GIS in ArcGIS. To be mentioned, the Z value (height) of 3D model is changed to "*Serface_percentage(%)*".

OBJECTID_1 *	Shape *	ENAME	SHAPE_Leng	Shape_Length	Shape_Area	Surface_percentage(%)
1	Polygon	WONG TAI SIN	0.156228	0.156228	0.000816	35.519745
2	Polygon	KOWLOON CITY	0.273668	0.273668	0.000884	29.49739
3	Polygon	KWUN TONG	0.220475	0.220475	0.000987	42.941167
4	Polygon	SAI KUNG	3.1547	3.1547	0.011963	2.037607
5	Polygon	NORTH	1.671697	1.671697	0.01207	4.331728
6	Polygon	CENTRAL & WESTERN	0.231894	0.231894	0.001091	22.64455
7	Polygon	WAN CHAI	0.168909	0.168909	0.000867	21.255729
8	Polygon	EASTERN	0.313439	0.313439	0.001631	17.375907
9	Polygon	TUEN MUN	0.85943	0.85943	0.007402	4.452075
10	Polygon	YUEN LONG	0.790774	0.790774	0.012297	3.493701
11	Polygon	SOUTHERN	0.976238	0.976238	0.003469	5.980359
12	Polygon	ISLANDS	3.056692	3.056692	0.015631	1.189698
13	Polygon	SHAM SHUI PO	0.246298	0.246298	0.000822	31.480385
14	Polygon	YAU TSIM MONG	0.201043	0.201043	0.000614	34.031715
15	Polygon	KWAI TSING	0.487708	0.487708	0.002046	18.538836
16	Polygon	TSUEN WAN	0.830662	0.830662	0.005474	5.156053
17	Polygon	TAI PO	2.013693	2.013693	0.012998	1.984943
18	Polygon	SHA TIN	0.491069	0.491069	0.006066	6.096753

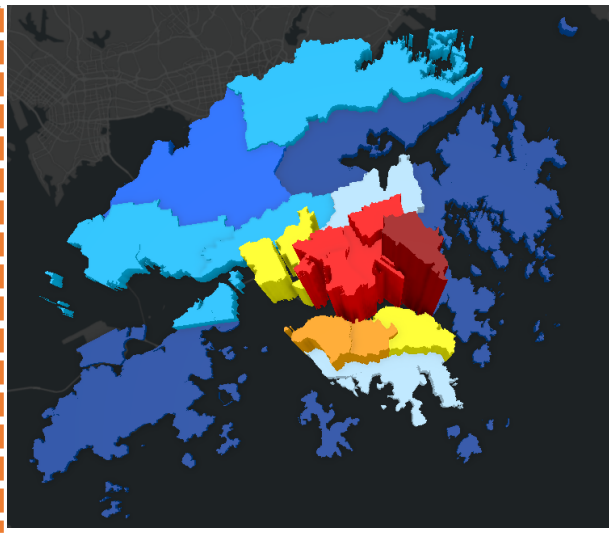
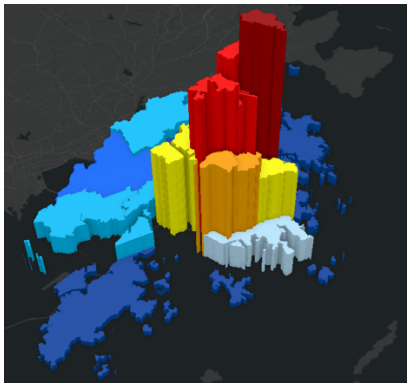
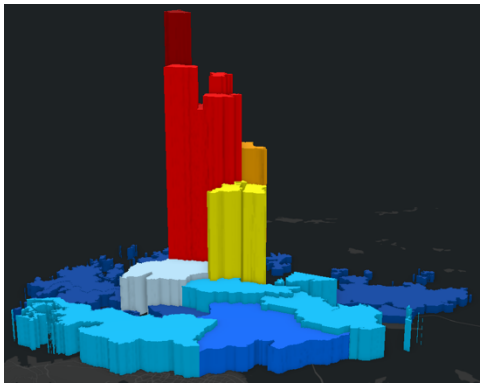


Figure 24 (Left). An Attribute Table contains the computed Surface Percentage (The Degree of Coverage).
Figure 25 (Right). The Surface Percentage is geo-visualized in 3D manners.





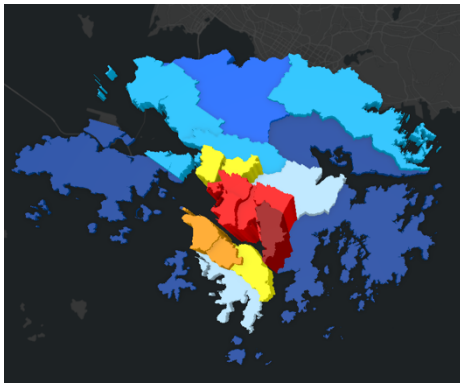


Figure 26 - 29. More Screenshots on the 3D GIS visualising Surface Percentage (The Degree of Coverage).

Spatial Analysis and Interpretation:

The highest surface percentage represents a good coverage and the lowest surface percentage represents a bad coverage. The degree of coverage is determined by the a spectrum of colours, red represents the highest surface coverage and navy blue represents the lowest surface coverage. By observing the 3D GIS model, it can observed that main higher surface percentages areas in district level are situated in the sides of Victoria Harbour, while the lower surface percentage are mostly in the New Territories and the Southern District of Hong Kong Island.

Based on the attribute table of badminton court in ArcGIS Pro shown in Figure 24. and 3D visualization in Figure 25., the good coverage is in Kwun Tong District, with 42.94% of the coverage. The second and third highest surface percentage are in Wong Tai Sin (35.51%) and Yau Tsim Mong (34.03%), respectively. These three districts with good coverage are also located in Kowloon. On the other hand, the lowest surface percentage is the Island District, with barely 1.19% of coverage. The Island district includes Lantau Island, Lamma Island, Cheung Chau, Discovery Bay, Tai O, Ma Wan and Peng Chau, to name but a few. Furthermore, the second and third lowest surface percentage are in the Tai Po District (1.98%) and the Sai Kung District (2.04%), respectively.

In conclusion, the good coverages of badminton courts in Hong Kong are situated in Kowloon, including the **Kwun Tong District, Wong Tai Sin District** and **Yau Tsim Mong District**. However, the comparatively bad coverages of badminton courts in Hong Kong are located in the **Island District, Tai Po District** and **Sai Kung District**.

4.4 Coverage of Three or above Types of Sports and Outdoor Facilities

As abovementioned, the number of intersections of buffers which attributed to the three or above sports and outdoor facilities and it can displayed in ArcGIS Pro. To make the thematic map more eye-catching, this map uses a spectrum of rainbow colours to represent the degree of coverage. In this study, a good coverage includes three or above facilities is defined and categorised in yellow (dense), orange (very dense) and red (the densest). On the contrary, a bad coverage is defined as sky blue (very dense) and blue (least dense). The corresponding spatial analysis is discussed in next section.

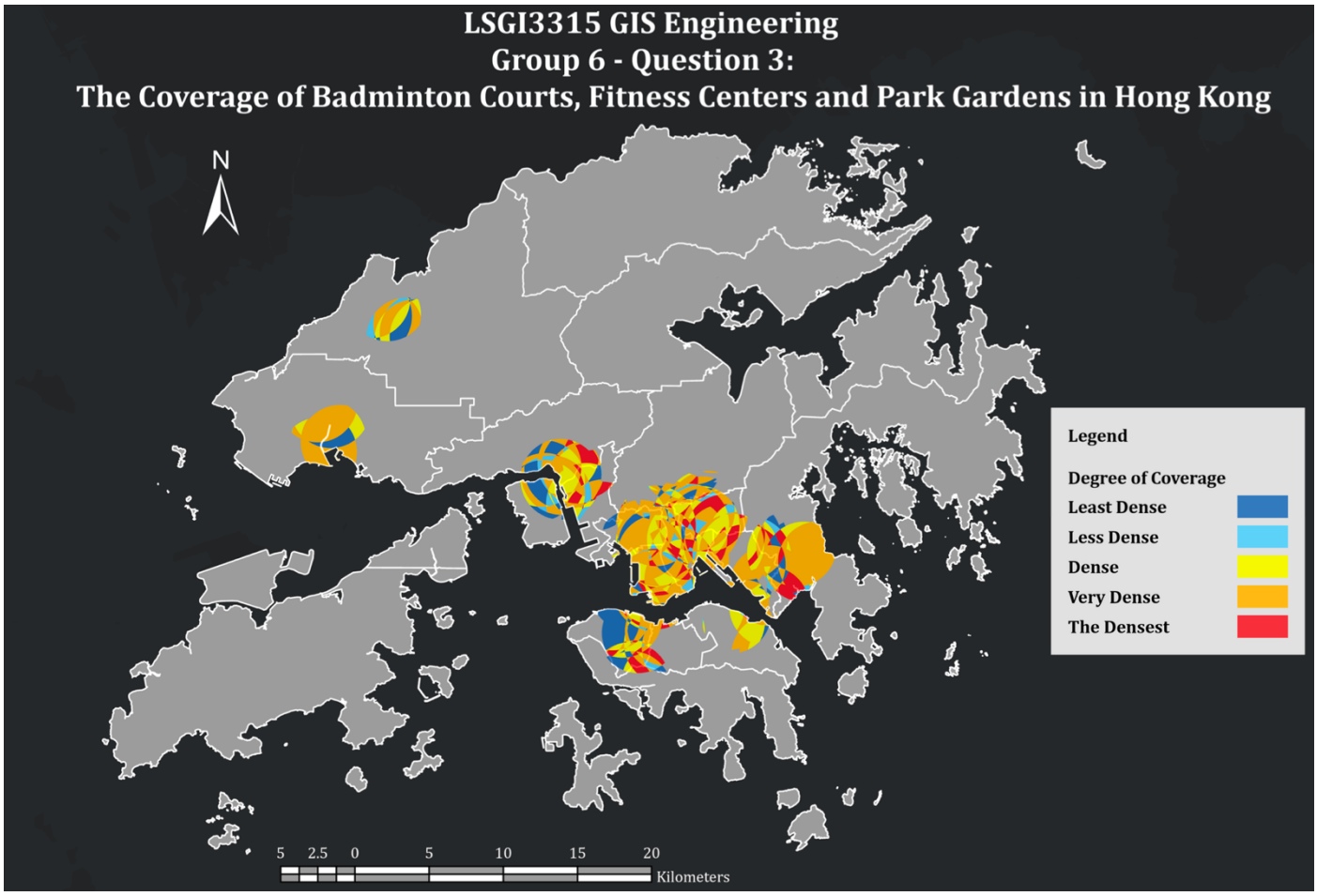


Figure 30. Coverage Analysis of Badminton Courts, Fitness Centre and Park Gardens in Hong Kong.

Spatial Analysis and Interpretation:

Figure 30. shows the degree of coverage of three types of sports and outdoor facilities in Hong Kong, namely badminton courts, fitness centres and park gardens, based on the heterogenous intersections of buffers from three specific facilities. In general, the coverages contain three facilities are located in Central business district, North Point on Hong Kong Island, nearly the whole Kowloon Peninsula, Kwai Chung District, Tseung Kwan O, Yuen Long District and Tuen Mun District. The below will further discuss the degree of coverage, whether the areas within three sports facilities are the densest or the least dense.

To begin with the densest area with three specific sports facilities (represent in red colour). Indeed, the densest areas are rare in Hong Kong, they are scattered in central, a small portion area of Kwai Chung district, Mong Kok, Kowloon Bay and Tseung Kwan O. Nevertheless, the areas with very dense of three facilities are spread through in Kowloon Peninsula and Tuen Mun District. On the other hand, the least dense areas of three types of facilities are located in Sheung Wan, Tsang Yi and the suburbs of Yuen Long and Tuen Mun.

Based on the above findings, it can be concluded that **the Kowloon Peninsula, Tseung Kwai O, Tuen Mun District and Yuen Long District** have comparatively high dense coverage of badminton courts, fitness centres and park gardens, and these places can be considered as good coverage through coverage analysis using ArcPy.

4.5 Bonus Part: Evaluating the Transportation Accessibility of All Sports and Outdoor Facilities

This is the **BONUS part** in this project. The additional section aims at studying the transportation accessibility of the sports and outdoor facilities. The user can input the searching distance for the accessibility and one type of facilities, to reveal its transportation accessibility. For the reference scenario, we study the badminton court within 100 meters and the related thematic map is shown below. Afterwards, we also explore the accessibility for all types of sports and outdoor facilities and the outputs are illustrated in line graph using **Matplotlib Package in Python library**.

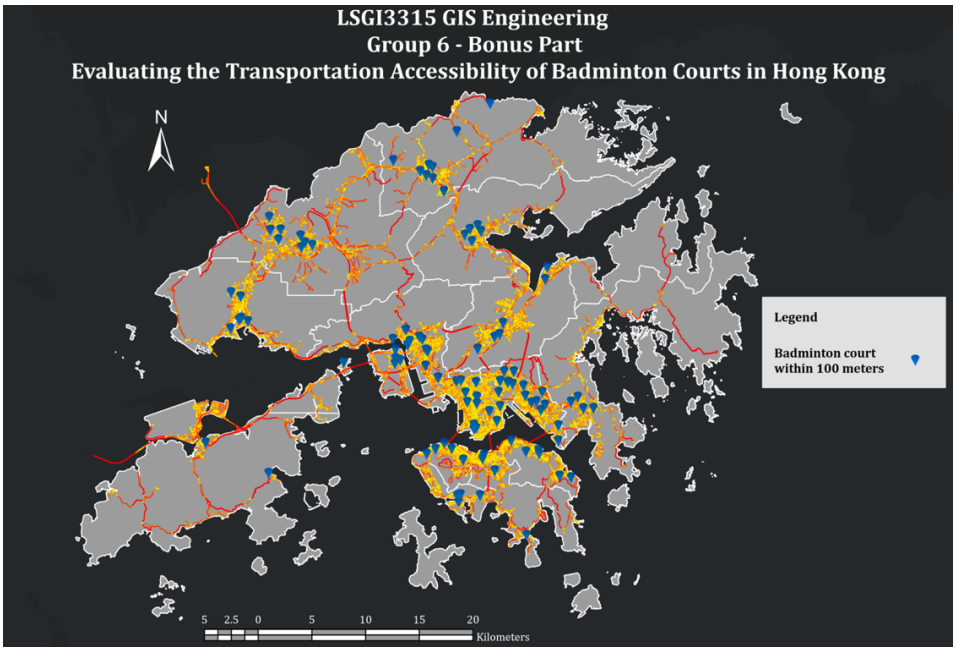


Figure 31. Evaluating the Transportation Accessibility of Badminton Courts in Hong Kong.

Spatial Analysis and Interpretation:

Figure 31 and 32. displayed **around 88% of badminton courts are within 100 meters** and they are mainly located in the urban area. However, the empirical findings seem fewer in this study. As a result, we conducted all transport accessibility for all types of sports and outdoor facilities in Hong Kong by simply change the value in the Main Function. To be mentioned, the miscellaneous of facilities is not included, because they have not been explicitly classified the types, and thus it could be meaningless to undertake the assessment.

Figure 32. displayed the transportation assessment for all sports and outdoor facilities, and the customized searching distance ranges from 20 meter to 200 meter. Remarkably, **a majority of the sports and outdoor facilities are highly accessible in transport, except country park**. For example, basketball courts (84%), badminton counts (89%) and fitness centres (90%) has a pleasant transport accessibility within 80 meters. It suggests one can access those facilities by transport modes and get off with walking very short distance to access the sports facilities.

However, the results of country park shows a very low transportation accessibility, with **only 8% of country parks are nearly transport network**. Indeed, this output is reasonable because the country parks are normally located in countryside with limited mobility, resulting in the an extremely low transportation accessibility.

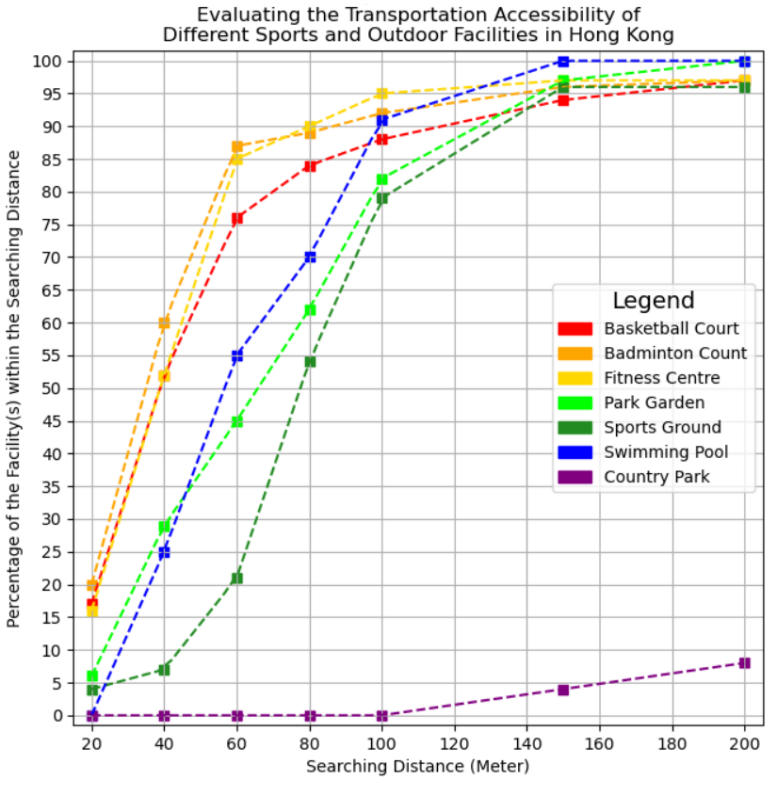


*Figure 32. Transportation Accessibility of All Types of the Sports and Outdoor Facilities (Drawn by the **Matplotlib Package** in Python).*

5 – Conclusion

For the Final Group Project of LSGI3315 (GIS Engineering), we explored the sports and outdoor facilities in Hong Kong and there are some significant findings can be found through geospatial perspective.

Question 1: What are the spatial distribution of different types of sports and outdoor facilities?
For the first compulsory question regarding the spatial distribution of different types of sports and outdoor facilities. With the consideration of the tourism management as well as improving the quality of life for Hong Kong residents, a majority of sports and outdoor facilities in Hong Kong are located in the tourist attractions, residential areas and new towns in Hong Kong.

Question 2 - 3: Given a specific facility type, which areas are within a reasonable walking distance to at least one of such facility, and which areas are within a user-specified distance?
Afterwards, in terms of reasonable walkable distance, using badminton court as a specific facility type, a majority of the urban areas are within a reasonable walking distance (i.e., 500 m), such as the Western and Eastern part of Hong Kong Island, nearly the whole coverage of Kowloon Peninsula and several new towns in the New Territories (i.e., Yuen Long District and Tuen Mun District) can access to badminton court within around 500 to 1000 meters.

Question 4: Given a specific facility type, which areas have a good coverage of it and which areas do not?
The third compulsory question is about the good and bad coverage of a type of sports facility. By using badminton court as a specific facility type, the good coverages of badminton courts in Hong Kong are situated in Kowloon, including the Kwun Tong District, Wong Tai Sin District and Yau Tsim Mong District. However, the comparatively bad coverages of badminton courts in Hong Kong are located in the Island District, Tai Po District and Sai Kung District.

Question 5: Which areas have a good coverage of different types of sports and outdoor facilities (e.g., equal or more than three types)?
Furthermore, for the question of the good and bad coverage regarding three or above types of sports facilities. By using Badminton Courts, Fitness Centre and Park Gardens as the three input, Kowloon Peninsula, Tseung Kwai O, Tuen Mun District and Yuen Long District have comparatively high dense (good) coverage. However, the bad coverage of these three types of facilities are located in Sheung Wan, Tsang Yi and the suburbs of Yuen Long and Tuen Mun.

Question 6 (Bonus): The Performance of Transportation Accessibilities for All Types of Sports and Outdoor Facilities?
Last but not least, the bonus question is about the transportation accessibility of the sports and outdoor facilities. Indeed, a majority of the sports and outdoor facilities are highly accessible in transportation, except country parks (with only 8% number of country parks are nearly transport network).

--- THIS IS THE END OF LSGI3315 (GIS ENGINEERING)'s GROUP 6 PROJECT. THANK YOU. ---

Reference List

Pimpler, E. (2013). *Programming ArcGIS 10.1 with Python cookbook over 75 recipes to help you automate geoprocessing tasks, create solutions, and solve problems for ArcGIS with Python (1st ed.).* Birmingham: Packt Pub.

Toms, S., Saldanha, A., & Shetty, A. (2015). *ArcPy and ArcGIS, geospatial analysis with python : Use the ArcPy module to automate the analysis and mapping of geospatial data in ArcGIS (1st ed., Community experience distilled)*. Birmingham, England ; Mumbai, [India]: Packt Publishing.

Division of Labour (Contribution)

Python Work and Report Writing	Wei Jun, Kenny	Tang Justin Hayse Chi Wing G.
Task 1: **Data Cleaning**	✓	✓
Task 2: **Implement a Python Class**	✓	✓
Task 3.1: **Spatial Distribution of Sports and Outdoor Facilities**	✓	✓
Task 3.2: **A certain facility within a Reasonable Walking Distance**	✓	✓
Task 3.3: **Exploring the Good and Bad Coverage of a certain Facility**	✓	✓
Task 3.4: **Exploring the Coverage of Multiple Sports/Outdoor Facilities**	✓	✓
Task 3.5 (Bonus Part): **Evaluating the Transportation Accessibility of a certain Facility**	✓	✓

Appendix

Considering the programming script can be easily customized by simply change the input in the Main Function, several different map productions of Task 3.2 and 3.4 are generated and the outputs are shown below.

Task 3.2: A certain facility within a Reasonable Walking Distance

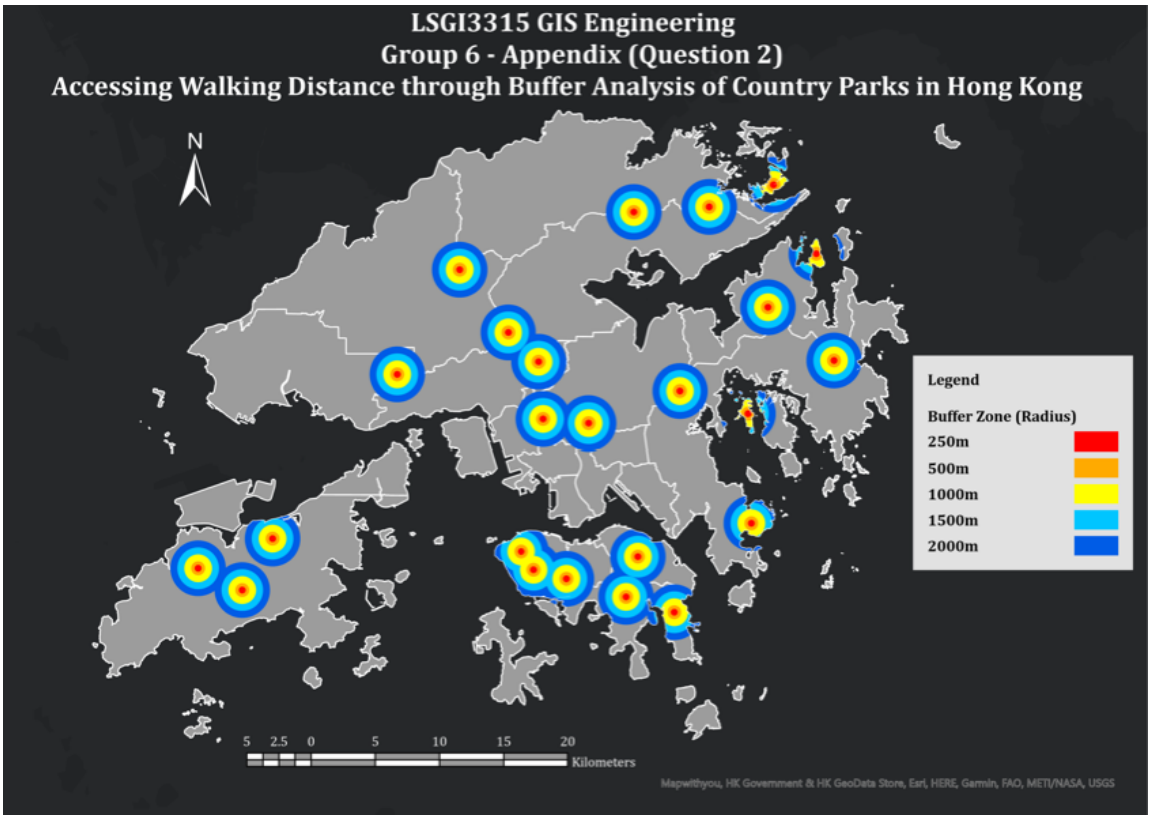


Figure A1. Accessing different walking distances of Country Parks through Buffer Analysis via ArcPy.

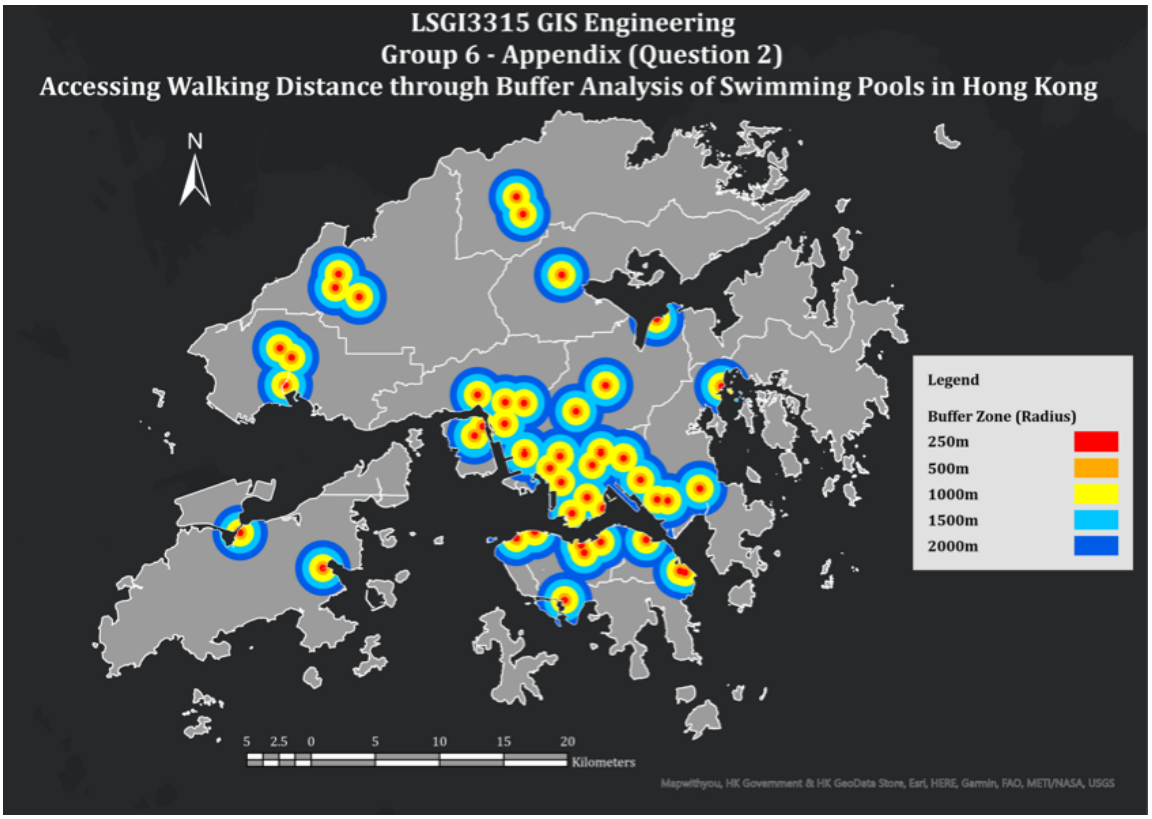


Figure A2. Accessing different walking distances of Swimming Pools through Buffer Analysis via ArcPy.

Task 3.4: Exploring the Coverage of Multiple Sports/Outdoor Facilities

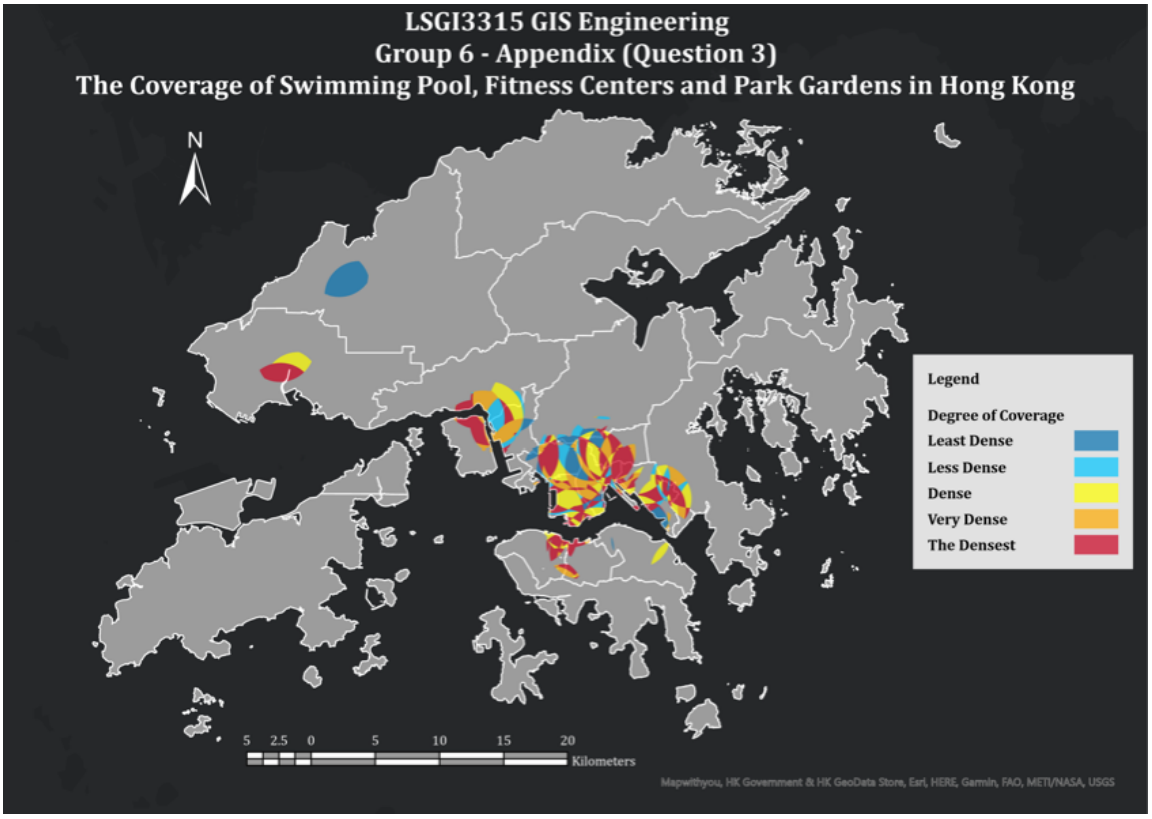


Figure A3. Coverage Analysis of Swimming Pools, Fitness Centre and Park Gardens in Hong Kong.

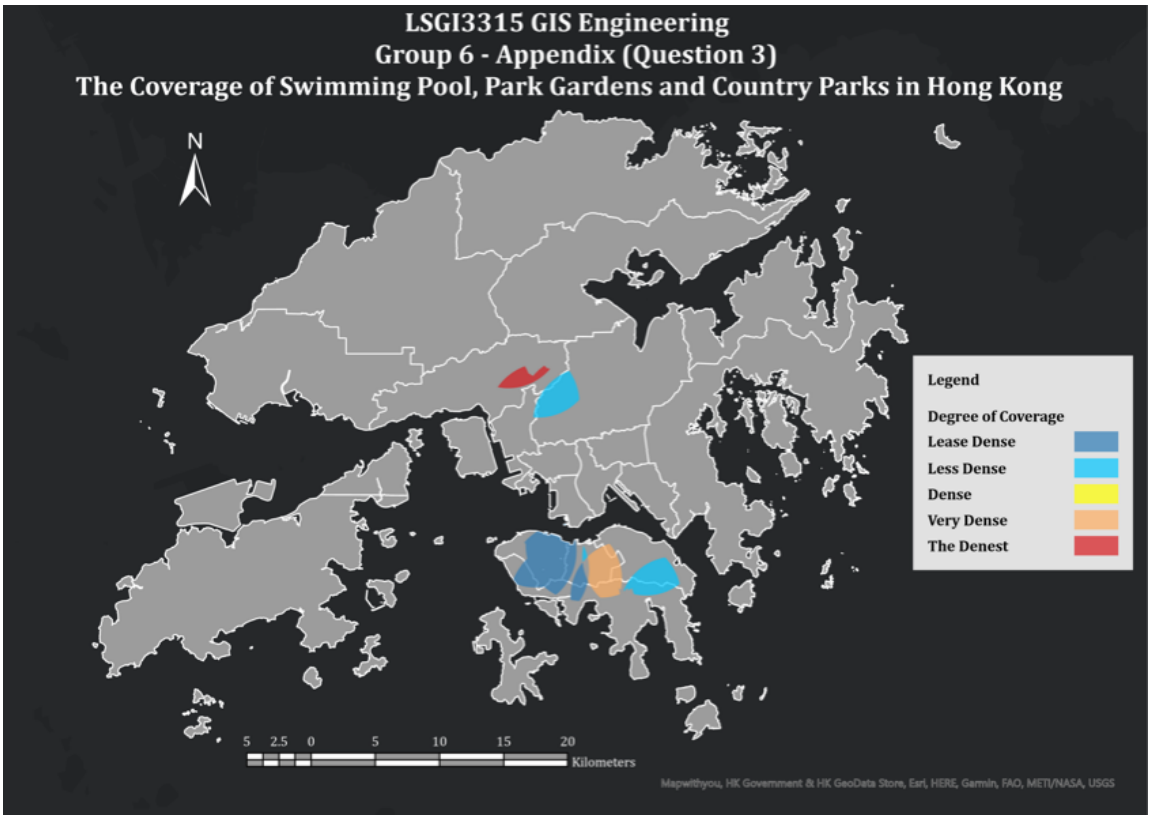


Figure A4. Coverage Analysis of Swimming Pools, Park Gardens and Country Parks in Hong Kong.

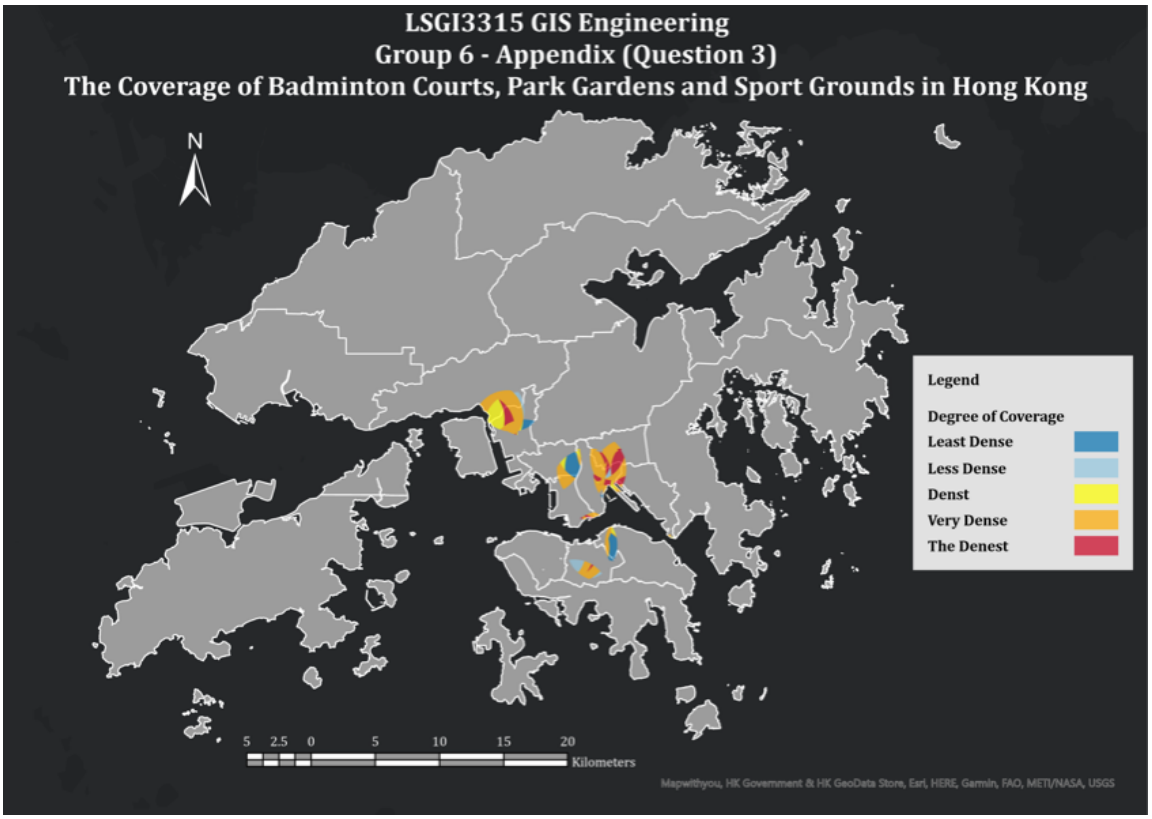


Figure A5. Coverage Analysis of Badminton Courts, Park Gardens and Sport Ground in Hong Kong.