University of Zurich[UZH]

Characterising London City Parks Using Open Data.

GEO 511 Master's Thesis

Author
Dace Kirsteina
19-765-577

Supervised by
Prof. Dr. Ross Purves
Marie Mueller (marie.mueller.16@ucl.ac.uk)

Faculty representative
Prof. Dr. Ross Purves

04.10.2022
Department of Geography, University of Zurich

"Public parks and open spaces play a major part in helping to shape our cities, improve the quality of life and play a central role in making our cities great places. In our often very heavily developed and crowded cities, the parks and open spaces provide clean fresh air in which to play sport, get fit, meet with friends, attend memorable events and they also provide a place to appreciate local heritage, engage with nature and relax. As a result, it comes as no surprise that they are much loved and well used places."

Core Cities Parks and Greenspace Group, London

ABSTRACT

Nearly 70% of the world's population is expected to live in urban areas in 2050 (United Nations, World Urbanisation Prospects, 2018). Cities therefore must be safe and attractive places to live and work. Natural areas and urban parks are becoming increasingly important for the quality of life in cities. Urban green spaces provide residents, tourists, and municipalities with many environmental, social, and economic benefits. Crucially, they also have many physical and mental health benefits. Considering the promising role of green spaces, it is important to conduct studies on different types of parks and the characteristics of parks which are most valuable for people.

This study uses six factors to characterise urban parks: horizontal vegetation density, vertical vegetation density, the proportion of trails in parks, the proportion of water objects in parks, the proportion of buildings and other anthropogenic structures in parks, and the size of the park.

Three factors (horizontal vegetation density, vertical vegetation density, and proportion of anthropogenic structures) were acquired using LiDAR (Light Detection and Ranging) data.

K-Means clustering was applied with six factors to classify London's typology of urban parks. The results of this study can be used to classify urban parks based primarily on similarities between the spatial context and physical characteristics of the parks. The results showed that the six factors used to characterise urban parks grouped them precisely by their unique features.

ACKNOWLEDGEMENTS

I would like to thank all those who inspired, motivated and helped me throughout my studies and helped to achieve this master's thesis.

First, I would like to thank my supervisor Prof. Dr. Ross Purves, for his support throughout my research. I could not have imagined a better advisor for my studies. His support, guidance and insights during the study made my master's thesis research an exciting experience.

I would also like to thank Maria Muller for her encouragement, comments, and good ideas. Her advice was highly beneficial, especially at the beginning and end of the progress, where her inputs were very practical.

I would also like to thank all those who participated in the study's experiment.

Lastly, I thank my family for supporting me throughout my master's thesis. This thesis would not have been possible without their help, support and encouragement. Their continued support allowed me to relax and think about other things.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF CODES

LIST OF ABBREVIATIONS

EU:	European Union
GIS:	Geographic Information System
LiDAR:	Light Detection and Ranging
NPPF:	The National Planning Policy Framework
OSM :	Open Street Map
PCA:	Principal Component Analysis
PC:	Principal Component
PPG17:	Guidance 17: Planning for Open Space, Sport and Recreation
SDGs:	Sustainable Development Goals
UGS:	Urban Green Space
UK:	United Kingdom
HD:	Horizontal vegetation density
VD:	Vertical vegetation density
S:	Building and other anthropogenic structures
R:	Trails and paths
W:	Water objects

1. INTRODUCTION

Human life is closely linked to nature. Since the beginning of humans, nature has provided people with space for living, food, and resources. It has been a central part of human development. Nature-human relationships have changed and evolved, yet relationships still exist. In rural areas, nature has always been more accessible to residents. In urban areas, nature-human connections are less direct and usually related to the presence of green spaces in the city.

It is already clear that most of the world's population will live in cities in the future. According to World Urbanisation prospects, 55% of the people already reside in urban areas (Sulc et al., 2020). Nearly 70% of the world's population is expected to live in cities (United Nations, World Urbanisation Prospects, 2018) (Figure 1.1.).

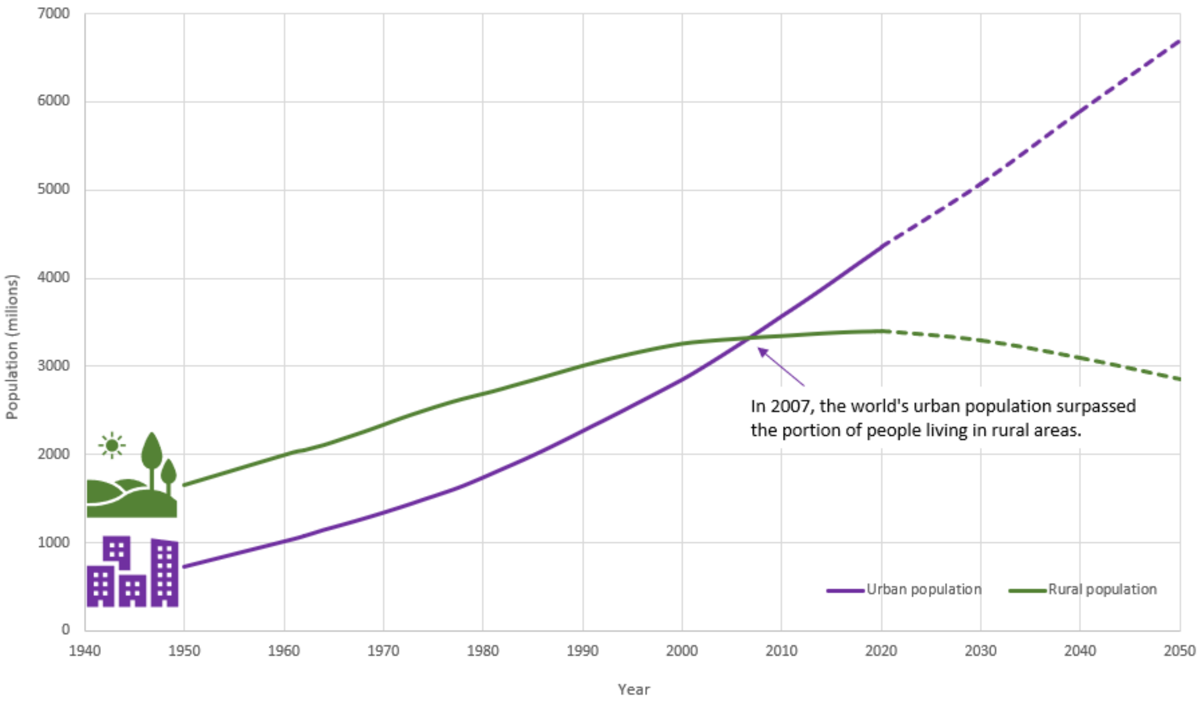


Figure 1.1. Growth of urban and rural world population since 1950, including future growth up to 2050 (Our World in Data, 2022; World Bank, 2018).

Cities need to increase their liveability and sustainability to ensure that urban spaces are safe and attractive places for living and working (European Commission, 2011). Natural areas and parks are becoming increasingly important for urban quality of life. Urban green spaces, such as urban parks, bring many environmental, social and economic benefits to citizens, tourists and municipalities. Physical and mental health benefits are also essential (Chiesura, 2004). Moreover, green spaces are traditionally considered "good places" for people to live in (Russo et al., 2018. Jansson, 2014). The United Nations have also recognised the importance of urban green spaces in their 2030 Sustainable Development Goals (SDGs).

Figure 1.2. Sustainable Development Goals, 2030.

Goal 11 explicitly states that "cities need to be made inclusive, safe, resilient and sustainable, with green and public spaces" (Sustainable Development Goals, 2030) (Figure 1.2.). In addition, urban green spaces also play a key role in achieving other goals. For example, Goal 3 promotes mental health and well-being. Goal 15 supports the sustainable use of terrestrial ecosystems, the sustainable use of forested land, and the halting of biodiversity loss.

It is therefore important to analyse why urban parks are perceived as so significant. While all parks are distinct in their own ways, optimal parks are often perceived to share common characteristics such as access to the surrounding community, various activities and experiences, vegetation of different types and densities, and picturesque views. In a perfect urban park, all of these elements blend to create a complex and enjoyable space for everyone. However, one could question whether there will always be a park that is "perfect" for everyone.

1.1. Motivation and Aim

Literature extensively discusses the importance and functions of green spaces in urban areas. (Baycan-Levent and Nijkamp, 2009; Lyytimäki and Sipilä, 2009; Harris et al., 2018). In an urban environment where a large proportion of the population needs spaces for recreation and relaxation, the importance of green spaces in promoting many social and environmental values is particularly apparent. Given the rapid urbanisation in the world and the popular approach to prevent urban sprawl by building dense and compact cities (Bibri et al., 2020), the importance of urban green spaces is crucial.

It is clear that urban green spaces (UGS) benefit not only people's everyday lives but also air quality in the city, local climate, recreation, and aesthetics (Konijnendijk van den Bosch et al., 2013). The characteristics of UGS, such as their location, size, availability of facilities, sports infrastructure, trails, and green features (such as the amount and density of vegetation), can determine the benefits or disadvantages visitors perceive. Understanding the qualities of UGS which create benefits can help to clarify the relationship between population choice towards local environments and the impact of urban green spaces.

Consequently, this study aims to investigate and analyse parks' characteristics and categorise them based on their composition. Many factors might characterise an urban park. In this research, six factors were used: horizontal vegetation density, vertical vegetation density, the proportion of trails, the proportion of water, the proportion of anthropogenic structures, and the size of the park. These are not the only factors that characterise a park. However, they are the ones that have emerged as distinct and relevant within this research project based on the literature review, defined research questions, and available open data.

Previous studies have been limited by only analysing vegetation in one dimension, i.e. horizontal vegetation density. This study fills this gap by integrating methods and approaches from forestry by analysing data from Light Detection and Ranging (LiDAR) technologies. LiDAR is an ideal technology for measuring the vertical structures of the tree canopy and its density. It is therefore worthwhile exploring the potential of using structural information from LiDAR data in the urban park context.

Furthermore, it is useful to understand what makes a park valuable for people. Study participants were asked to participate in a questionnaire, where they evaluated the importance of different park characteristics, aiding the understanding of the factors people value the most in parks. This also created a value of space measure based on ratings to assess how valuable different parks are for people.

Only open source data were used in the research project. Therefore, due to the widely available data, this research selected Greater London, with its 33 boroughs, as a case study. For each park in London larger than 0.5 Ha, the values of factors such as the park's size, horizontal vegetation density, vertical vegetation density, the proportion of waters, the proportion of trails and pathways, and anthropogenic structures, were calculated.

Three factors (horizontal vegetation density, vertical vegetation density, and proportion of anthropogenic structures) were acquired using LiDAR data. This optical remote-sensing technique uses laser light to sample the surface of the Earth densely. LiDAR produces highly accurate three-dimensional measurements of the shape of the Earth and its surface characteristics including terrain, tree canopy, sub-canopy, and vegetation (Dong et al., 2017). LiDAR yields information unavailable through two-dimensional landscape images provided by traditional multispectral remote sensing platforms. It provides detailed information about three-dimensional vegetation horizontal and vertical structure which is important for park and other greenspace assessments. Data for the other three factors (proportion of trails and pathways, proportion of water objects, and the park's size) were acquired using open source data (OpenStreetMap).

1.2. Research Questions

Building on the goals discussed above, the following research questions are examined by this study:

Research Question 1:

How does the structural information from LiDAR enhance the existing urban park data?

Research Question 2:

Can urban public parks be classified using multi-dimensional clustering that groups parks according to park attributes?

Research Question 3:

How do the clustering results correlate with respondents' perceptions of parks?

1.3. Outline

The following chapter introduces the theoretical background for this research. It reviews research on the classification of urban spaces, benefits and use of urban green space, and the use of LiDAR for the classification of vegetation. Chapter 3 provides an overview of the data used for this research. Chapter 4 outlines the methodology used in the study and gives an overview of the methods. Chapter 5 presents the results of the study. Chapter 6 discusses the results, relates them to the research questions and links them to the literature. Moreover, the limitations of the methodology and results are highlighted in this chapter. Finally, Chapter 7 concludes this research and provides recommendations for future research.

2. BACKGROUND

2.1. Classification of Urban Spaces

It is difficult to find a definition for "green space" that reflects all its characteristics, qualities and services. There are several definitions and classifications of urban green areas (Taylor and Hochuli, 2017; WHO Europe UGS, 2017). For the aim of this study, the definition and typology provided by Stanley et al. (2012) were found to be the most relevant. They propose a typology that reflects a set of forms and functions, including functionally specific and multifunctional categories. Stanley et al. (2012) (Figure 2.1.) divides open space into seven main types: transport facilities, streets, plazas, recreational space, incidental space, parks and gardens and food production areas. Open spaces within each type are further subdivided into three categories by a spatial scale: city scale, intermediate scale, and individual buildings scale. Stanley et al. attribute city scale to open spaces linked to the national or municipal level, or oriented to large population segments. The intermediate scale refers to a more local part or district of the city. Residence scale refers to individual buildings or residences with open space for local inhabitants, such as household gardens or enclosed courtyards.

This typology is based on Al-Hagla's (2008) distinction between "green space", which is defined as "open space consisting of any vegetated land or structure, water or geological feature in urban areas", "grey space", which is "open space which consists of urban squares, market places, and other paved areas with an urban function" and "grey/green space" as "open areas with mixed function of green and grey", like greenbelts, sports facilities and interior gardens.

The city's open space division into a green-grey sub-set covers above mentioned seven forms of open space. It is an essential variable in planning research, urban ecology, and human-environmental relations (Jenerette et al., 2011).

Stanley et al. (2012) typology was recognised as suitable for master's thesis research. As a basis of classification, major formal parks and garden spaces from the "City scale" and commons, community and institutional gardens, small parks and cemeteries from the "Intermediate scale" were chosen.

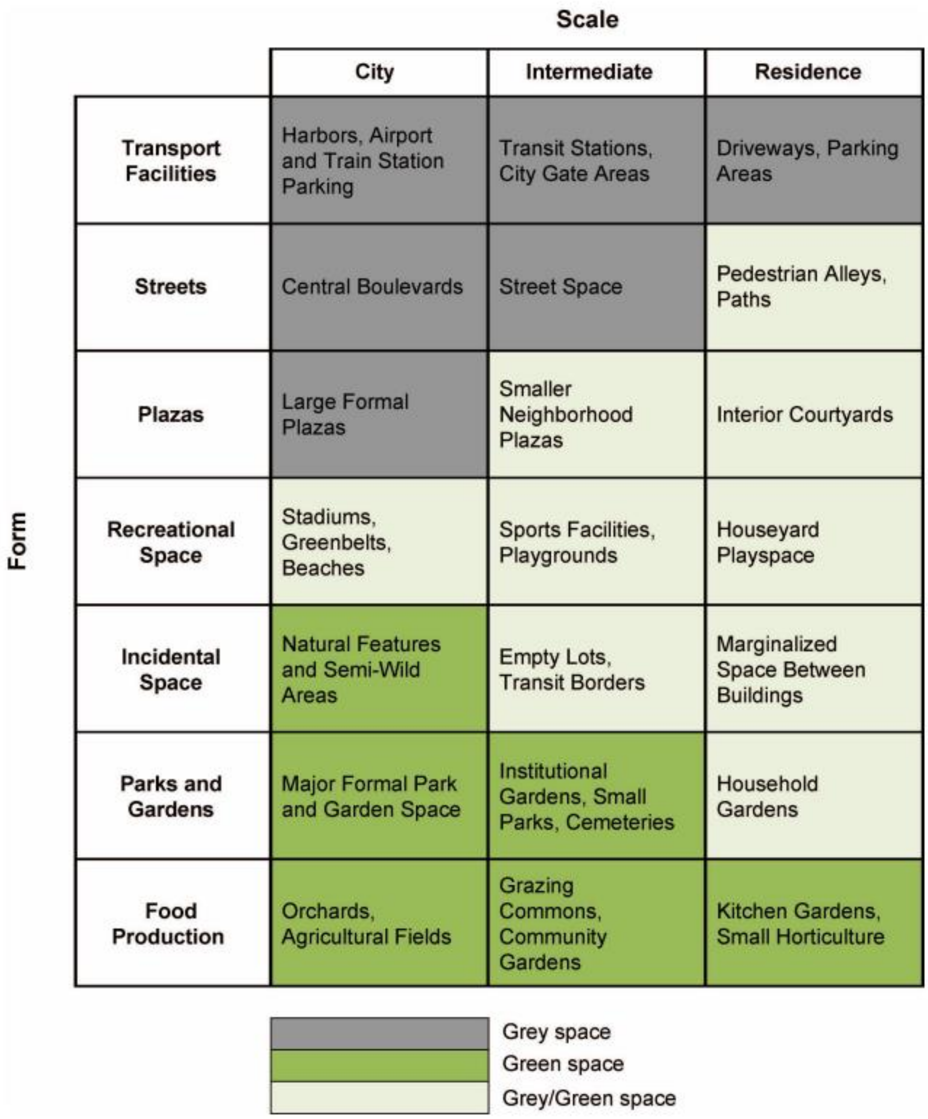


Figure 2.1. A transdisciplinary typology of urban open spaces (source: Stanley et al. (2012)).

In the context of this typology, parks and gardens are defined as green spaces, partly landscaped and intended for social, recreational activities or aesthetic purposes. Some parks and gardens can be highly specialised and institutionally designed for specific cultural functions. (Stanley et al., 2012).

It is important to note that this analytical separation is often limited and incomplete. Many places are a mix of open space types and are culturally assessed due to their multifunctional nature. For example, Richmond Park, the largest Royal Park in London, matches the parks and gardens category on the city-wide scale. However, Richmond Park is home to a diversity of spaces: landscaped gardens, recreational sports fields, streets used for transportation and recreation, parking areas, semi-wild forested areas, and several buildings.

2.2. "Green and Blue"

Green space, like parks, forests, vegetation corridors, playing fields, wetlands, urban forests, and cemeteries, is often described as "open, undeveloped land with natural vegetation" (Jansson, 2014; Mitchell et al., 2008). Unfortunately, this definition does not address whether to include water areas or blue spaces. Indeed, some authors consider blue spaces the same as green spaces (Al-Hagla, 2008). However, other scholars characterise water areas as separate (de Vries et al., 2003, White et al., 2010).

In the urban park context, it is more meaningful to define green space as all green areas of public value, including also areas of water such as lakes, rivers, canals and reservoirs, which offer significant opportunities for recreation or sport and are visually pleasing (Völker and Kistemann, 2011).

The literature clearly states that blue spaces are a unique element in the landscape that beautifies the environment. People have always been attracted to the environment with water (Nasar and Li, 2004). Blue space, complementing the landscape, creates a diversity of space that improves mood, reduces stress and performs emotional and spiritual healing (Völker and Kistemann, 2011). Water is often perceived as a symbol of purity in human mental and spiritual life and has a more significant impact on emotional well-being than other environments (Ulrich, 1981; Felsten, 2009). The combination of green and blue spaces has a much more positive impact on people's health than just green or blue (White et al., 2010; Williams, 2010).

Several studies about preference and perception refer to the value of blue space regarding plant and animal diversity and increased stimulation of visual, auditory, and olfactory sensations (Völker and Kistemann, 2011).

For this study, blue space such as rivers, canals, lakes, and ponds are considered significant and is one of the factors analysed.

2.3. Benefits and Use of Urban Green Space

Several studies have summarised the benefits of green space (Chiesura, 2004; Swanwick et al., 2003) and calculated its value in monetary terms (Miller, 1997; Bolund and Hunhammar, 1999; Boyd, 2007; Choumert and Salanié, 2008). Furthermore, the benefits of green space for people are often more difficult to calculate than their ecological value (Chiesura, 2004, Boyd, 2007). The aim of this study is not to calculate the benefits of UGS. Instead, it aims to understand which characteristics of UGS contribute to the benefits of green areas. It also seeks to understand people's preferences toward urban greenspaces better.

Urban green spaces are of significant importance for the quality of life of today's increasingly urbanised society, providing residents, tourists and municipalities with many environmental, social, and economic benefits as well as physical and mental health benefits (Jansson, 2014).

The presence of natural values in urban parks and forests, as well as various natural components, such as vegetation and water, contributes to the population's quality of life in many ways. In addition to providing important environmental services such as air and water purification, wind and noise filtering or stabilising the microclimate, natural areas offer social and psychological benefits. Urban dwellers need to have a pleasant experience in modern cities, and the presence of nature is what contributes to this. Green spaces reduce stress (Ulrich, 1981) and provide a sense of tranquillity (Kaplan, 1983). Contemporary research on urban parks and forests supports the theory of reducing population stress and mental health benefits (Hartig et al., 1991, Conway, 2000). Schroeder (1991) has proven that natural environments with vegetation and water evoke more pleasant emotions in observers than urban environments without vegetation. Besides the aesthetic, psychological, and health benefits mentioned above, cities' natural features also have social benefits. Nature encourages us to use the outdoors and enhances social integration and interactions among neighbours (Coley et al., 1997). Green areas also help people relax and feel restored, easing aggression (O'Connor, T (2008). In addition, urban nature can bring economic benefits to municipalities and citizens. Urban parks, for example, provide air purification by trees, which also reduces the costs of air pollution reduction and prevention measures. Besides, urban parks' aesthetic, historical, and recreational values increase the city's attractiveness and promote it as a tourist destination, generating additional revenue and employment. Finally, natural elements such as vegetation or water increase property values and tax revenues (Tagtow, 1990; Luttik, 2000).

To better understand why parks are essential for society and municipalities, this study adapts Rung et al. 's (2005) "conceptual benefit model" (Figure 2.2.) for understanding the benefits of parks.

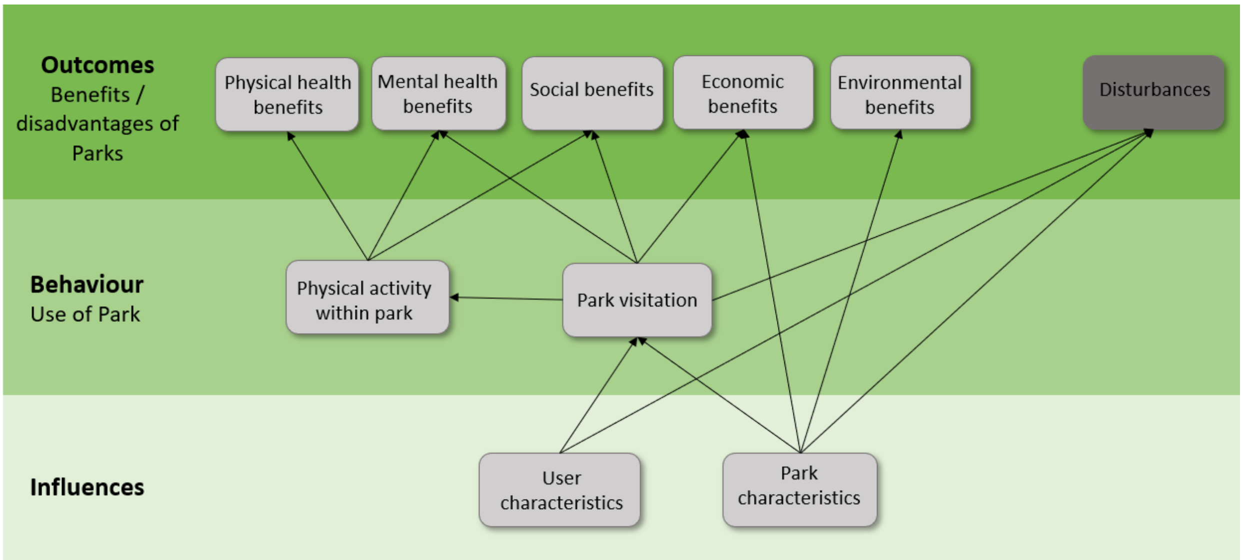


Figure 2.2. Conceptual benefit model (adapted from Rung et al., (2005)).

The lower section of the model adapted from Rung et al. shows the factors affecting the use of parks. It includes the characteristics of the potential park user (e.g. age, gender, education, socio-economic and marital status) as well as the characteristics of the parks (size, features, attractiveness, biological diversity, type, distance to home or workplace, and available park facilities (e.g. trails, lighting, sports infrastructure, and benches)).

The middle part of the model describes the way the park is used. When users are in the urban park, they can be more or less physically and socially active during a visit.

The top section of the model represents the benefits or disadvantages of parks or the use of parks. Psychological and physical health and social benefits are associated with more active urban park use. Economic or ecological benefits will likely accumulate from a city's green space.

Next is a brief look at the park's benefits and disadvantages according to the conceptual model.

Physical health benefits

Raising the population's physical activity level is an essential health strategy, as almost two-thirds of the adult population does not currently reach the physical activity levels recommended by the World Health Organization (World Health Organization, 2006). However, the factors which motivate people to become physically active in parks remain unclear. Some studies show these factors could include how close the park is to home or work, the size of the park (Giles-Corti et al., 2005), and certain facilities, like trails and sports infrastructure.

Mental health benefits

Many people state that they feel better after walking in the park. Scientific research also suggests that spending time in parks improves mental health. The proximity of residents' housing or activities to the green space has been connected with fewer symptoms of depression and lower levels of stress (Sturm et al., 2014).

Social benefits

Green spaces provide places where people can meet and develop social connections (Jennings et al., 2019). It could involve a picnic, art event, or cultural program.

Economic benefits

For both governments and individuals, parks also create economic benefits. For example, several studies have found that parks positively relate to property value. These studies have concluded that the greater the distance of the residential property from the greenbelt, the lower its price (Crompton, 2001). The presence of parks in a community also helps to boost the local economy by attracting tourists, residents, and businesses.

Environmental benefits

Parks also preserve and purify the environment, including protecting plant and animal habitats, decreasing air pollution, and increasing water filtration (Nowak et al., 2006). Trees in urban parks also regulate temperatures, providing shade and cooling to an area, helping to reduce the risk of heat-related illnesses (Zander et al., 2020).

However, parks can also include aspects that can be perceived negatively. This can involve anthropogenic elements, such as waste, antisocial behaviour, or vandalism (Jorgensen et al., 2013). It can also involve natural factors, such as allergenic plants, mosquitoes, ticks, or intruding animals (Kasprzyk et al., 2019). Furthermore, different socio-demographic and cultural groups perceive additional factors negatively. Some of parks' negative aspects are closely related to the structure and density of vegetation. People have historically perceived dense vegetation as more dangerous than sparse vegetation and have tended to avoid shady parks with poor lightning (Jorgensen et al., 2007).

Based on the benefits and drawbacks that parks offer, it is clear that many factors influence society's perception of green space. The park's location, size, availability of sports or leisure facilities and green properties determine the benefits or disadvantages park visitors perceive. Understanding which components of parks contribute to which benefits can help meet city residents' various demands.

2.4. The Use of LiDAR for the Classification of Vegetation

LiDAR is a remote sensing method that uses light as a pulsed laser to measure variable distances to the Earth's surface (Shan and Toth, 2017). Light pulses generate precise, three-dimensional information about the shape of the Earth and its surface characteristics, such as terrain, tree canopy, subcanopy, and vegetation (Dong et al., 2017). Laser pulses are emitted from a lidar system and reflect objects on the ground surface. One emitted laser pulse can return to the sensor as one or many returns (see Figure 2.3.). The first returned laser pulse is the most significant return and is usually the tallest feature or object in the landscape, such as a treetop or the top of a building. The first return can also represent the ground, in which case only one return will be detected by the LiDAR system. In multiple returns, the intermediate returns usually indicate the structure of vegetation. The last return is used for bare-earth terrain models. However, the last return will not always be from a ground return. If the pulse does not reach the ground due to various obstacles, the last return is not from the ground but a barrier reflecting the entire laser pulse (Fernandez-Diaz, 2011; https://resources.arcgis.com/)

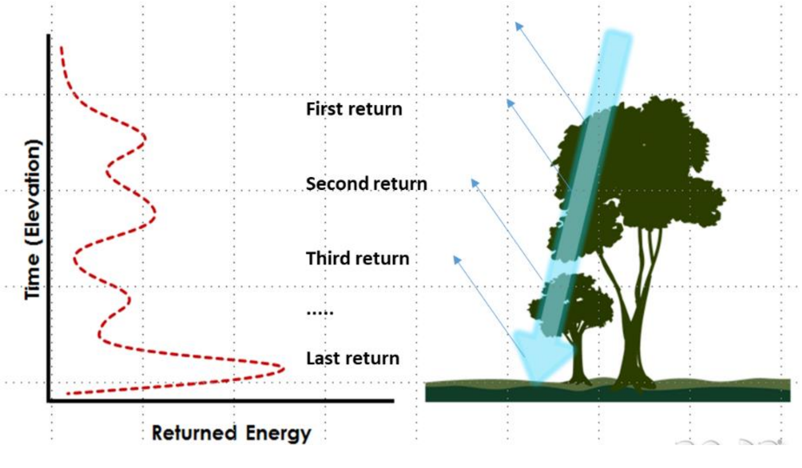


Figure 2.3. LiDAR returns. An example of LiDAR waveforms (source: https://pro.arcgis.com).

LiDAR data and technologies provide additional information for a more precise classification of vegetation compared with multispectral datasets (Garcia et al., 2011; Kim et al., 2020). The additional height information and other values extracted from LiDAR can improve the quality of vegetation classification. LiDAR data and data processing methodologies are widely used in classifying vegetation in forestry and other sectors.

One of the limiting factors of using LiDAR-based data and methods is the availability of LiDAR data. However, in recent years, LiDAR data availability has increased substantially. Pyszny et al. (2020) point out that the availability of high-resolution LiDAR data and the development of their analysis tools make it possible to create accurate maps of urban vegetation. Kim et al. (2020) also state that LiDAR-based methods give comparable or even better results than standard methodologies. They state that LiDAR allows for the classification of vegetation index from remote imagery data. LiDAR data was involved in the classification workflow to enhance information from aerial images in creating land use and land cover maps for wildlife and natural resource management in Florida, USA. Advances in LiDAR technologies have enabled remote sensing-based studies of vegetation canopies by providing a three-dimensional representation of vegetation structure throughout the canopy. Often three-dimensional information provided by LiDAR data is left unutilised, and LiDAR data and methodologies are focused on developing high-resolution Digital Evaluation Models, canopy heights, and occasionally understory density (Kumar et al., 2015).

Further, we provide three examples highlighting the benefits and promising functions of LiDAR.

Kumar et al. (2015) use an efficient Python-based data processing workflow to estimate the vertical canopy structure of a large forest area. LiDAR data was filtered to remove anomalous and noisy returns caused by atmospheric aerosols, dust, birds, or other unknowns. Low height (<1m) vegetation was removed. 1-meter vertical resolution was used to identify vegetation height from the ground surface to a maximum height. Normalised density profiles were created by computing the percent of total points in 1-meter vertical grids. This gridded dataset was used for the classification of vertical canopy structures. To verify the quality of this methodology, classified canopy structures were compared with vegetation maps of the same areas. LiDAR-derived canopy structures indicated significantly better results than vegetation maps that classify the region in traditional vegetation classes. This study states that forest managers can monitor the forest using repeated LiDAR surveys, which were previously impossible due to the complexity and volume of airborne LiDAR data sets.

Another study investigated the possibility of using LiDAR data to inspect urban trees (Chen et al., 2015). This research explores the correlation between green tree visibility and structural variables directly acquired from LiDAR data. The research area used was Cambridge, England, which has typical urban vegetation for cities in Western Europe. This methodology could be used for large-scale analysis of urban trees. Using remote sensing imagery is a widely used approach for greenery assessment. However, different tree species have a diversity of spatial structures. This means that a large crown area does not indicate a large height and additional vertical structure classification is necessary. This is particularly important for urban greenery because urban forests are important sources that provide people with visual greenness in urban areas. This research used LiDAR data and photographic images of sample trees to create a regressive model from LiDAR data variables that allow scalable estimating of green visibility and other essential tree characteristics.

Further research on LiDAR technologies was performed in Gothenburg, Sweden (Klingberg et al., 2017). In this study, LiDAR data were used to obtain the leaf area of urban trees to evaluate different aspects of the urban ecosystem, such as air quality, temperature regulation, and noise. Ground-based measurements of leaf areas are time-consuming and impractical. They are also expensive to perform over large areas, making remote sensing techniques attractive. Alongside the LiDAR-based methodology, two standard ground-based methods were used: commercial Plant Canopy Analyser and hemispherical photography. This approach allows the comparison and evaluation of results. The study concludes that estimates of leaf index in the urban environment based on LiDAR data have significant potential and reliable assessments.

2.5. Urban Green Spaces in London

As London parks will be analysed in this study, it is worthwhile investigating their history and the types of parks which exist within the city.

2.5.1. History of London's Parks

Many believe that London's parks are among the best in the world. They create greenery across the city, allowing Londoners to connect with nature, be active, and have fun (Greater London Authority, 2021). London has several categories of green areas, like royal parks, public parks, common land, pleasure gardens, churchyards, and purpose-built parks, all of which have historically different origins and were created for a variety of reasons.

Royal parks were the ones that developed first. These lands were initially used for the recreation of the royal family, which primarily involved hunting. They were not open to the public in the past. They are part of the hereditary possessions of The Crown, which are today managed by Royal Parks Limited (Thurston, 1974).

London's public parks began to form in the 19th century. The city became increasingly urbanised, and the British government attempted to create green spaces in the city. These parks were supposed to be enjoyed by all levels of society. They included vast open areas, pavilions, water objects and driveways. Many of these features still exist in these parks today (Smyth, 2019).

Common lands originated in the eleventh century and were properties typically owned by several persons or organisations. Others had certain traditional rights, such as allowing their livestock to graze on it and gather firewood or turf for fuel. There are more than a hundred commons across the city. Over time, some of this land has been sold. As London's population grew, public campaigns began to protect many of London's common grounds, which ensured that many commons remained open to the public. Today, unlike parks, which sometimes are gated, commons are open with unlimited access and are relatively wild and remote.

In the early 17th century, London began designing and creating gorgeous, purpose-built pleasure gardens designed solely for enjoyment. Some of these gardens were designed to provide both an indoor and outdoor space for entertainment. In contrast, others were created for medical purposes, to help people recover from various illnesses, which is why many purpose-built parks are located near water. (Smyth, 2019).

Churchyards are another category of parks in London city. At the end of the 19th century, having closed for burials because they were full, many became public Church Gardens managed by a local authority.

2.5.2. Open Space and Urban Park Classification in London

Although this study does not aim to explore London's planning policy landscape, it is helpful to investigate the official classification based on The National Planning Policy Framework (NPPF), which sets out planning policies for England (issued in July 2021) and Spatial Development Strategy for Greater London (published in May 2021). These outline a framework for how London will develop over the next 20-25 years.

NPPF has replaced PPG17. However, open space land uses are classified by previous PPG17 categories and the typology classifications still provide a valuable guide for categorising open space types. Within the NPPF and Spatial Development Strategy for Greater London, a hierarchy of urban green spaces is identified, and relevant policies are distinguished. Table 2.1. shows seven types of public open space and their descriptions. Each identified type of space contributes to recreation, biodiversity, landscape, quality of life, and climate adaptation objectives.

Table 2.1. Classification of public open spaces in London (Source: Greater London Authority, 2011)

Type	Description
Regional Parks	Large areas, corridors or networks of open space, the majority of which will be publicly accessible and provide a range of facilities and features offering recreational, ecological, landscape, cultural or green infrastructure benefits. Offer a combination of facilities and features that are unique within London, are readily accessible by public transport and are managed to meet best practice quality standards.
Metropolitan Parks	Large areas of open space that provide a similar range of benefits to Regional Parks and offer a combination of facilities at a sub regional level, are readily accessible by public transport and are managed to meet best practice quality standards.
District Parks	Large areas of open space that provide a landscape setting with a variety of natural features providing a wide range of activities, including outdoor sports facilities and playing fields, children's play for different age groups and informal recreation pursuits.
Local Parks and Open Spaces	Providing for court games, children's play, sitting out areas and nature conservation areas.
Small Open Spaces	Gardens, sitting out areas, children's play spaces or other areas of a specialist nature, including nature conservation areas.
Pocket Parks	Small areas of open space that provide natural surfaces and shaded areas for informal play and passive recreation that sometimes have seating and play equipment.
Linear Open Spaces	Open spaces and towpaths alongside the Thames, canals and other waterways; paths, disused railways; nature conservation areas; and other routes that provide opportunities for informal recreation. Often characterised by features or attractive areas which are not fully accessible to the public but contribute to the enjoyment of the space.

3. STUDY AREA AND DATA

3.1. Study Area

The study area for this work is Greater London. This city has 33 boroughs and is located in the South East corner of the United Kingdom. It is positioned on the Thames river at approximately 51°30'30.71"N, 0°7'32.66"E and encompasses an area of 1,569 km². London's parks cover almost 18 per cent of London's territory. London was chosen as a study area mainly because it has a wide range of open data.

3.2. Data

To assess the value of space in London parks, multiple datasets from multiple data sources were integrated into the GIS analysis. The complete list of the data sources can be found in Table 3.1.

Table 3.1. Data sources used for research.

Data layer/factor	Data source	Year
Borders of boroughs	**Ordnance Survey OpenData** Free OS OpenData Map Downloads | Free Vector & Raster Map Data | OS Data Hub	**2021**
Borders of parks	**OSM** Export | OpenStreetMap **Ordnance Survey Greenspace** OS Open Greenspace - data.gov.uk **Historic England Parks** Registered Parks and Gardens GIS Data - data.gov.uk	**2022** **2021** **2018**
Park size	**OSM** Export | OpenStreetMap **Ordnance Survey Greenspace** OS Open Greenspace - data.gov.uk **Historic England Parks** Registered Parks and Gardens GIS Data - data.gov.uk	**2022** **2021** **2018**
Vertical vegetation Density	**LiDAR** Defra Data Services Platform	**2018, 2020**

Horizontal vegetation density	**LiDAR** Defra Data Services Platform	**2018, 2020**
Trails and paths	**OSM**	**2022**
Water features	**OSM**	**2022**
Anthropogenic features	**LiDAR** Defra Data Services Platform	**2018, 2020**
Population	**Open Data London, Census data** Dataset Search - London Datastore	**2021**

3.2.1. Spatial Vector Data

This research used three different datasets for detecting the borders of London's parks:

- Open Street Map (OSM), a worldwide known, open dataset that is primarily user-generated and provides geographic information about various topics (OpenStreetMap, 2021).
- Ordnance Survey Open Greenspace database, which is an open source database which has been automatically generated and generalised from Ordnance Survey large-scale data. Greenspace data contains information on greenspace sites, such as parks and sports facilities, that are likely to be accessible to the public.
- Historic England Parks and Gardens database, which contains GIS spatial data for Registered Parks and Gardens. It is part of the National Heritage List for England.

Open Street Map was used to analyse paths and trail networks, and water objects.

3.2.2. LiDAR Data

LiDAR Point Cloud data for London city was acquired by airborne from November to April 2018 and 2020 under defoliated conditions (leaf-off). Data are supplied in *.laz format, and the discrete LiDAR returns were classified into the ground, low, medium and high vegetation classes using an automated classification process (National LiDAR Programme, 2022).

4. METHODOLOGY

4.1. Data preparation

This section explains how the data was prepared before the analysis process began. The data, namely park border data, trails and pathways, water features and LiDAR data are discussed in detail. The data preparation was mainly achieved using Python, R, QGIS, and ArcGIS. All Python and R scripts used for this research can be found in the GitHub repository (https://github.com/dkirsteina/MscThesis.git).

4.1.1. Vector Data Preparation

Geographical data were exported from OSM with Overpass Turbo (Appendix 1) by selecting the necessary information within an area of interest. The vectors were further pre-processed using the process described below.

4.1.1.1. The Borders of Parks

Determining the borders of parks was challenging. A significant number of qualitative data is available for London, but, as mentioned above, this study only used free and open data sources. Three datasets were useful for this aim: Open Street Map (OSM), the Ordnance Survey Open Greenspace database and the Historic England Park database. Unfortunately, none of these was complete or fulfilled data quality conditions. The Historic England Parks and Gardens database contained the parks and gardens from National Heritage List for England and was considered a data layer with high confidence. Unfortunately, only 166 London parks and gardens are included in this database. The Open Street Map data set and the Ordnance Survey Open Greenspace database had inconsistent quality. They missed out parks or gardens, provided duplicate polygons, and wrongly classified green areas. Precision and accuracy varied across London's territory. Therefore, the decision was made to use all three databases.

Open Street Map data were downloaded through Overpass Turbo, a web-based data mining tool for OpenStreetMap data. Appendix 1 shows the Overpass map query for extracting relevant information from Open Street Map database. The Ordnance Survey Open Greenspace database and Historic England Parks and Gardens database were downloaded through Open Data London portal. Since three data sources for park borders were used, they had to be integrated into one. The Geoprocessing workflow for developing park boundaries is shown in Figure 4.2.

Firstly, from Ordnance Survey Open Greenspace data and OpenStreetMap data, only those features that fulfilled conditions based on the selected layer by attribute function were selected. Following this, the geometric intersection of the combined database and the Historic England Parks and Gardens database was computed. This means that the attributes and geometry of the integrated database were updated by the Historic England Parks and Gardens database in the output feature class. Not all parks had names. Therefore, parks that did not have a name in the original database were given a unique name (for example, NoName1, NoName2). The Dissolve function was applied to eliminate the large number of neighbouring borders that resulted from merging databases. Only the parks with an area of more than 0.5 ha remained. The final database of London parks consisted of 1638 parks. The geometric intersection of London's borough database and park database was calculated to obtain information about which parks were located in which boroughs.

Figure 4. 1. London parks created in this study, extracted from Open Street Map, Ordnance Survey Open Greenspace database and Historic England Park database.

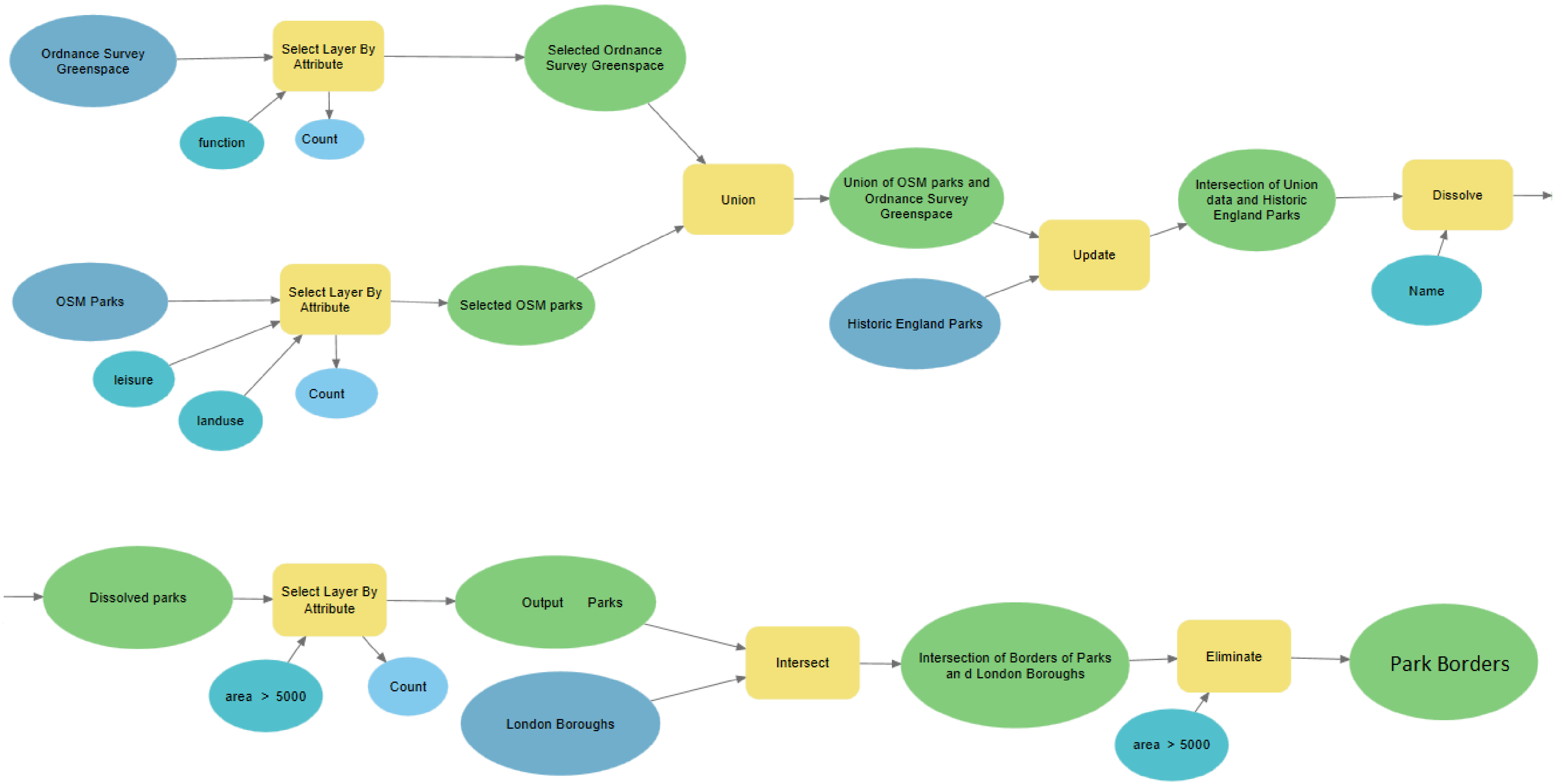


Figure 4.2. Geoprocessing workflow for developing Park boundaries. Created with ArcGIS Pro ModelBuilder.

4.1.1.2. Trails and Pathways

Trail and Pathway network lines for London parks were extracted from the OSM data through Overpass Turbo. Appendix 1 shows the Overpass map query where the line features with the "bridleway", "cycleway", "footway", "path", "pedestrian", "service", "steps", and "track" values were extracted.

Extracted trails and pathways were clipped with park borders and then aggregated and summarised, providing information about the overall length and the proportion of the trails in each park (Figure 4.3.).

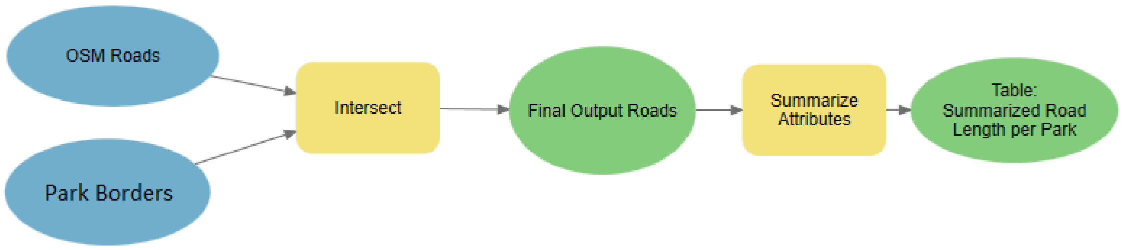


Figure 4.3. Geoprocessing workflow for Paths and Trails. Created with ArcGIS Pro ModelBuilder.

4.1.1.3. Water Features

Water bodies (lakes and ponds) and water lines (streams, rivers, creeks, canals, and ditches) were also collected from OSM data by using Overpass Turbo Query (Appendix 1). To ensure a uniform methodological approach, line-shaped objects (water lines) were buffered with a 2-meter buffer and thus created as polygons. Following this, the geometric intersection of the buffered water lines and water polygon features were computed. The Dissolve function was applied to eliminate the neighbouring borders that resulted from databases being merged. In the end, the geometric intersection of the park border database and the water features database was calculated to obtain information on how many water objects were in each park territory.

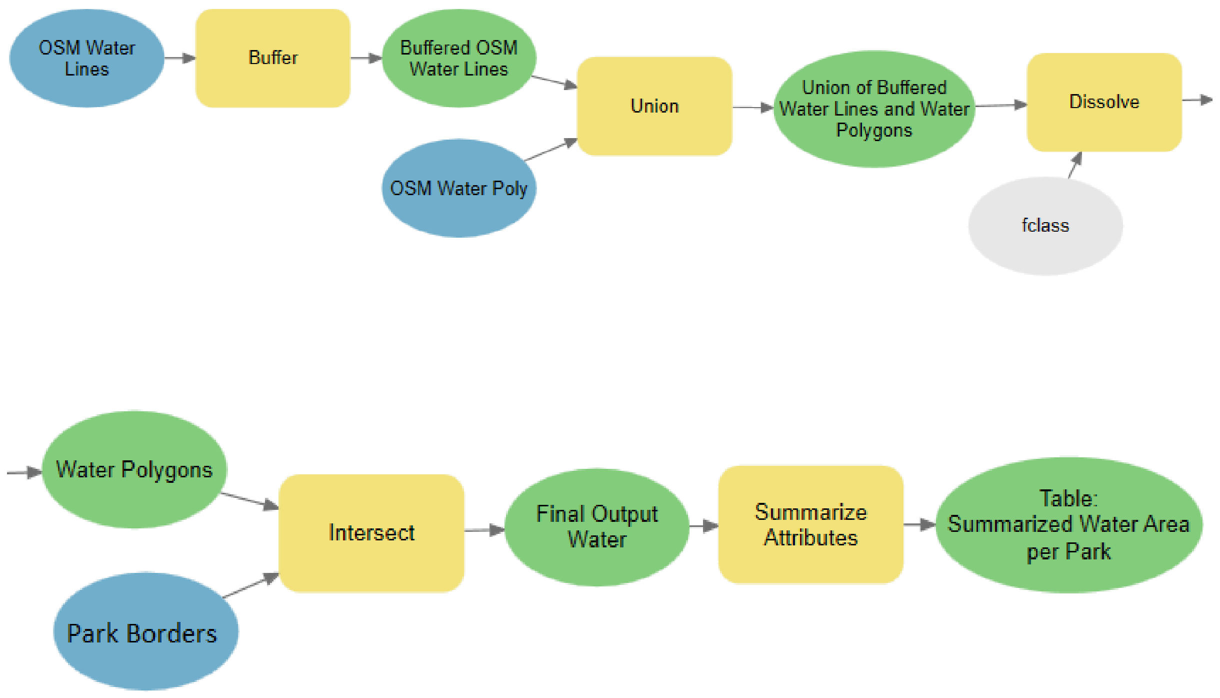


Figure 4 4. Geoprocessing workflow for Water Features. Created with ArcGIS Pro ModelBuilder.

4.1.2. LiDAR data processing

LiDAR data from UK National LiDAR Program were downloaded from the Open data portal (data.gov.uk) by 5 km data tiles. There were 67 tiles for the entire London area with a total data volume of 120 Gb.

Pre-processing LiDAR data involves several steps. Firstly, compressed LAZ files were converted to LAS files. Data was then cleaned by removing extraneous points, echoes, and deviations. Overlapping areas were filtered. The next step was merging tiles into a single point layer. This was necessary for further analysis.

4.1.2.1. Horizontal Vegetation Density

Horizontal vegetation density was calculated using first return points, which are the most significant returns and are associated with the highest feature in the landscape, such as a treetop. The first return points were filtered with medium and high vegetation class and then converted to raster with a three-meter grid cell (Figure 4.5. and Figure 4.6.)

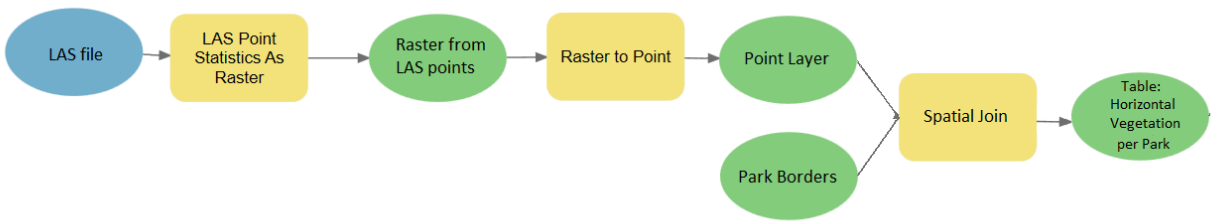


Figure 4. 5. Geoprocessing workflow for LiDAR data processing for horizontal vegetation density calculation. Created with ArcGIS Pro ModelBuilder.

Figure 4. 6. Process of rasterisation of vegetation horizontal density. The number in the grid cell indicates the first return LiDAR points. The value of light green grid cells is below the threshold and therefore is considered non-vegetated.

Grid cells contained between zero and several dozen points. By using an empirical approach, a threshold value was selected. Values above the threshold were considered to be medium or high vegetation. Following this, by using raster algebra, the horizontal density of vegetation per whole park territory was calculated.

4.1.2.2. Vertical Vegetation Density

To estimate the vertical density of vegetation, the ratio of the number of points representing vegetation (all returns from medium, and high vegetation) to the number of all points in a given grid cell area was calculated (Krzysztof et al., 2020). Vertical density was calculated only in areas that appeared as vegetation, based on horizontal density calculation (Figure 4.7.).

It is worth highlighting again that grass, as the base surface cover of a park, does not count as vegetation in this research.

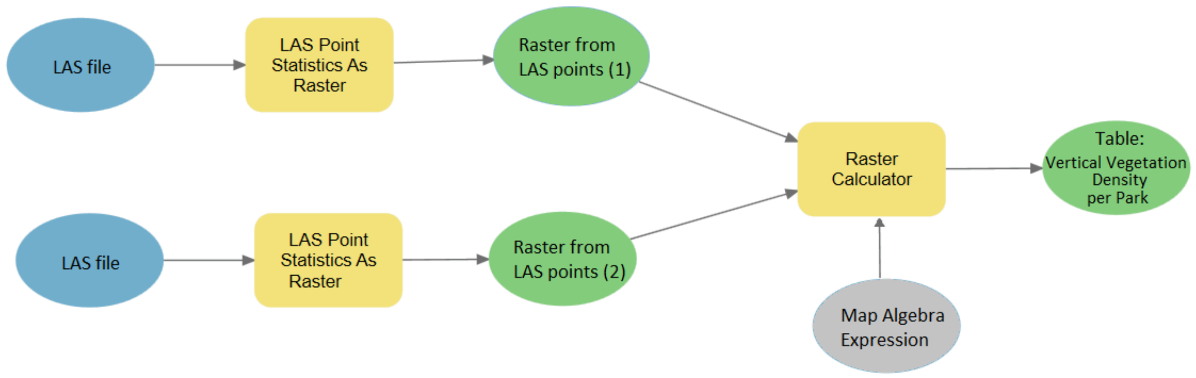


Figure 4. 7. Geoprocessing workflow for LiDAR data processing for vertical vegetation density calculation. Created with ArcGIS Pro ModelBuilder.

4.1.2.3. Buildings and other anthropogenic structures

The classification code representing building points from the first LiDAR laser return was selected to determine the proportion of buildings and other anthropogenic structures in parks. Then, spatial statistics were calculated within each park (Figure 4.8.).

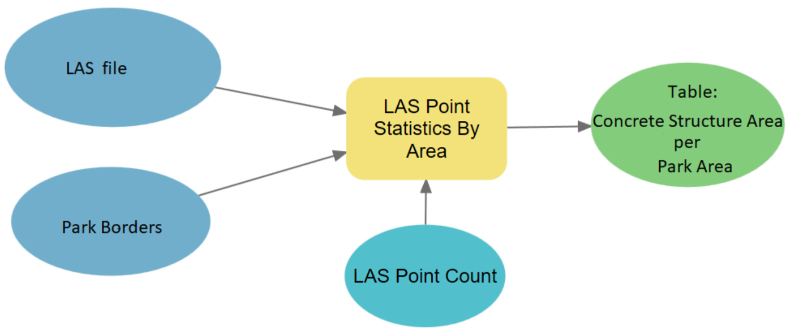


Figure 4. 8. Geoprocessing workflow for Buildings and other Anthropogenic Structures. Created with ArcGIS Pro ModelBuilder.

4.2. Cluster Analysis

One of the research questions for this study is to classify urban public parks based on important factors. To find parks with similar or related features, multi-dimensional clustering analysis, which is an unsupervised machine learning approach, was performed (Kongphunphin and Srivanit, 2021)

The methodology took place in three main stages, as shown in Figure 4.9.

The first phase identified and calculated key factors that may influence the value of space in urban parks. In the second phase, statistical analysis was applied to develop the urban park classification. In the third phase, cluster analysis was used as the statistical method for generating and mapping the groups of urban parks.

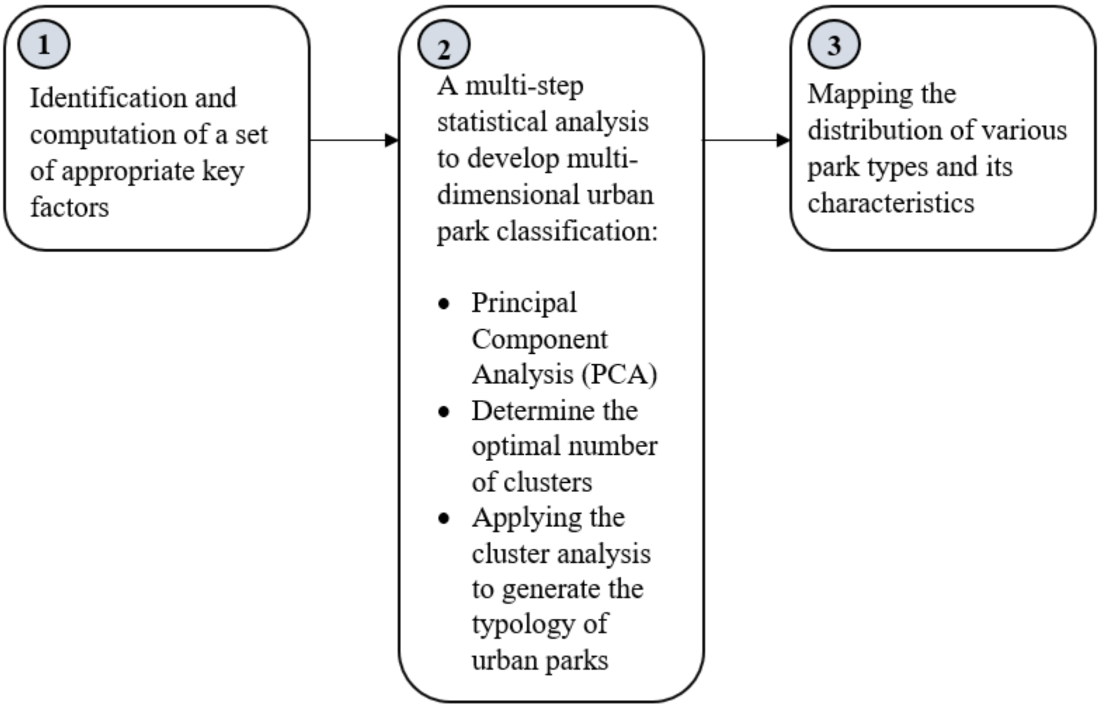


Figure 4.9. Three stages of the methodology of cluster analysis.

Based on the literature review, defined research questions, and available open data, six applicable variables were selected (Table 4.1.), and their values were calculated. These are not the only factors that characterise a park; however, they are the ones that emerged as distinct and relevant in this research project.

Table 4.1. Description of the factors influencing the value of space of urban parks.

Factor	Abbreviation	Description	Units
Park size	Area	The size of the urban park	Hectares (Ha)
Trail and pathway density	R	The ratio of the length of the trail and pathway network inside the park to the park area.	Meters on square meter (m/m^2)
Coverage of water features	W	Percentage of water features (ponds, lakes, streams, rivers, creeks, canals and ditches) area inside the park.	Percentage (%)
Buildings and other anthropogenic structures	S	Computed factor based of LiDAR data; shows how many buildings and other anthropogenic structures area were inside the park.	Lidar points per park area square meter ($points/m^2$)
Horizontal vegetation density	HD	Computed factor based of LiDAR data; shows percentage of medium and high vegetation (bushes, plants, trees) area inside the park. Grass, as the base surface cover of a park, does not count as vegetation in this research.	Percentage (%)
Vertical vegetation density	VD	Computed factor based of LiDAR data; shows how dense is the vegetation inside vegetated area.	Lidar points per vegetated area square meter ($points/m^2$)

4.3. Participant Study

An experiment in the form of a questionnaire was carried out to compare the value of space or urban parks described in the previous section to questionnaire participants' assessments of the characteristics of the parks they value most.

4.3.1. Participants

The study was carried out with thirty participants. They were recruited via Twitter by one of the thesis supervisors, Marie Mueller. The Geocomputation unit retweeted the tweet at the Department of Geography University of Zurich. Furthermore, the study was advertised to the students of the GIS for Environmental Monitoring module (Geo888, 2022). Further, some participants were recruited through acquaintances.

4.3.2. Design and Procedure

The questionnaire aimed to gather survey participants' opinions on park values. The full questionnaire can be found in Appendix 2.

The questionnaire consisted of:

- An introduction part, where the purpose of the research work and questionnaire was described.
- Introduction questions about participants' gender, professional experience, and interest in park/green areas policy-making or management.
- Behavioural questions about frequency and barriers respondents experience when visiting parks and/or other urban greenspaces.
- In the main part of the questionnaire, respondents were asked to evaluate and rank the six factors by their importance for their "ideal" urban parks and/or other urban greenspaces.

Some questions were multiple choice, some were open-ended, and the main question was a ranking question. All data assessed by the questionnaire were analysed using the computer software R.

5. RESULTS

5.1. Descriptive Results of Factors.

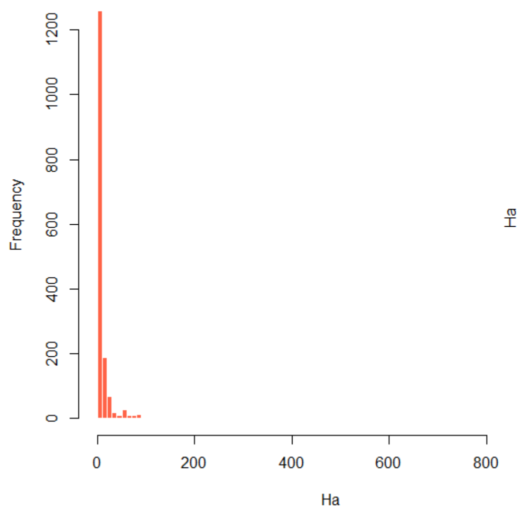
The following section provides descriptive statistics about the parks.

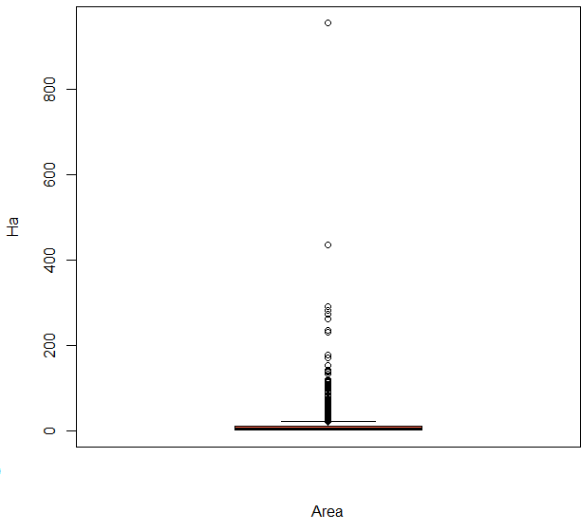
5.1.1. Park Size

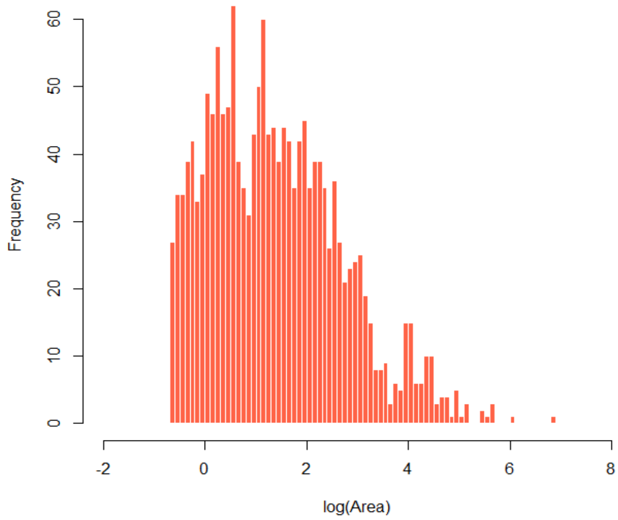
Figure 5.1. shows a histogram of distribution, box-plot, log transformation histogram, and park size statistical values. These show that data had a strongly skewed distribution typical for natural origin data. Data had wide distribution between 0.50 Ha and 954.67 Ha; however, the histogram and the box-plot show that most data were closer to a minimal value. There were a few outliers with extremely high area values (for example, Richmond Park, around 955 Ha, and Bushy Park, around 435 Ha).

Since distribution histograms and box-plots do not provide much information about data, the logarithmic transformation of data was used to spread out clumps of data, bring together spread-out data, and reduce the inconsistency of original park area data. Logarithmic transformation showed that data had two frequency peaks, around 2 Ha and 3 Ha and that most parks were smaller than 60 Ha, despite data being spread between 0.5 and 954 Ha.

Furthermore, box-plots showed that most parks are small, meaning that few data outliers affect the data set, lowering the mean value of the data, median, and interquartile range.





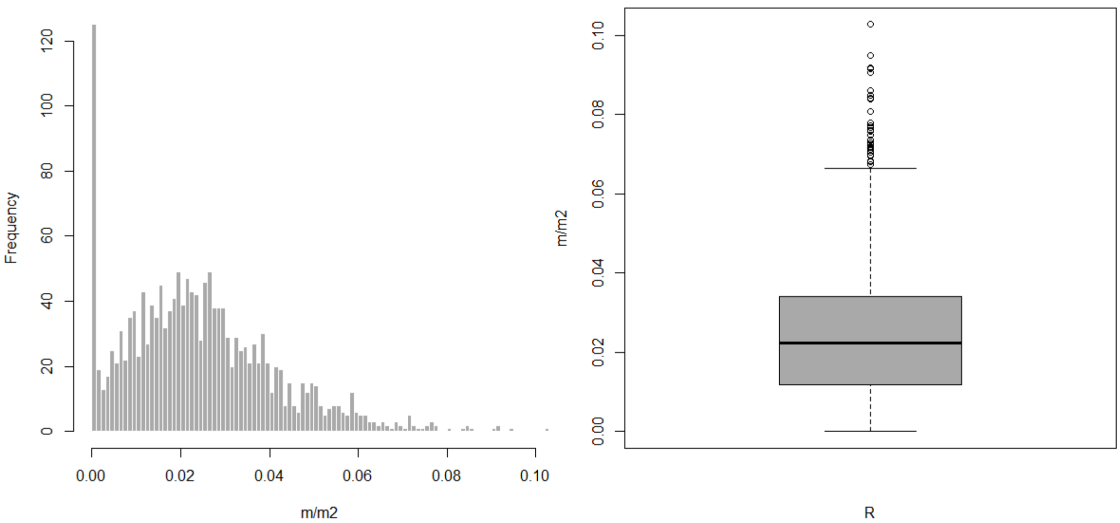


	Area
Min	0.503
1st Q.	1.381
Median	3.369
Mean	11.660
3rd Q.	9.348
St.Dev.	35.534
Max	954.673

Figure 5. 1. Histogram of distribution, box-plot, histogram of log transformation, and statistical values of variable "Area" (size of parks).

5.1.2. Trails and Pathways

As seen in Figure 5.2, the trails and pathways density factor was near normally distributed. Data were between 0 to 0.10 meters of trails and pathways per m^2 of the park. Data higher than 0.07 m/m2 were classified as outliers. 75% of data were up to a value of 0.03 m/m^2. Therefore, 25% of parks had dense trail and pathways network (e.g. more than 0.07 m/m^2), but 75% could be considered medium to low density.



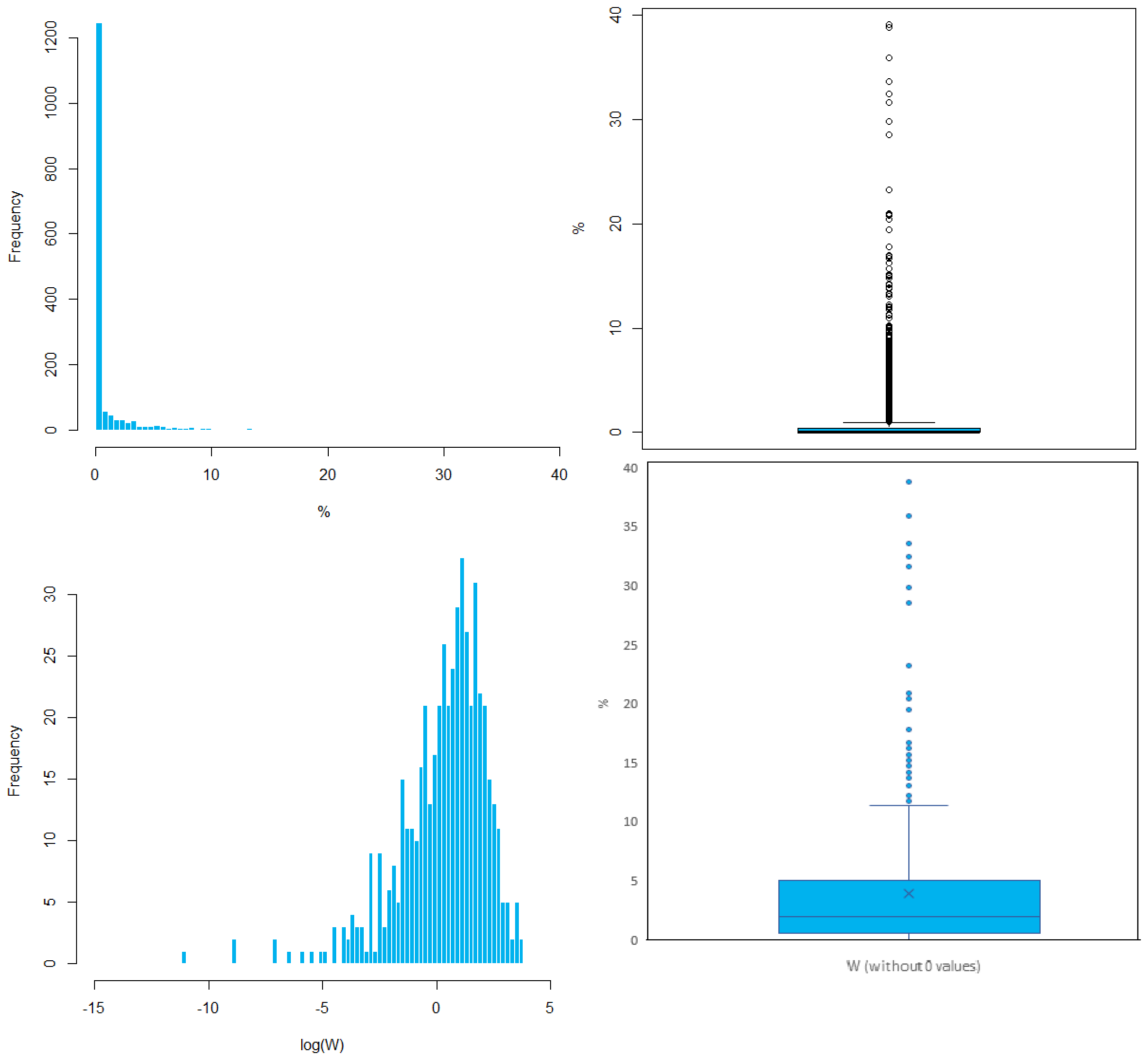
	R
Min	0.000
1st Q.	0.012
Median	0.022
Mean	0.024
3rd Q.	0.034
St.Dev.	0.017
Max	0.103

Figure 5. 2. Histogram of distribution, box-plot and statistical values of variable "R" (trails and pathways).

5.1.3. Water Features

Figure 5.3. displays a histogram of distribution, box-plot, log transformation, and statistical values of water features. The histogram of distribution and box-plot shows that data were strongly skewed. Data was distributed between 0 and 39.05% of water features per park territory. Both histogram and box-plot show that the dominated quantity of water features was concentrated around 0 value. 68% of data had 0 value, meaning that 2/3 of all parks had no water features. Data outliers affected the water feature data set, downgrading the mean value of the data, median, and interquartile range.

Logarithmic transformation showed that the data set had a detectable frequency peak of around 3-6% of water features within parks. Excluding parks which contained no water, it can be observed that in 75% of parks, water features cover about five per cent of the total park area. All values above 12% were considered to be outliers.



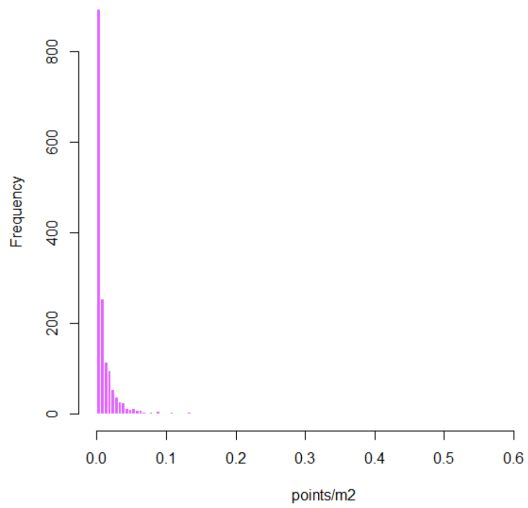
	W
Min	0.000
1st Q.	0.000
Median	0.000
Mean	1.235
3rd Q.	0.382
St.Dev.	3.691
Max	39.045

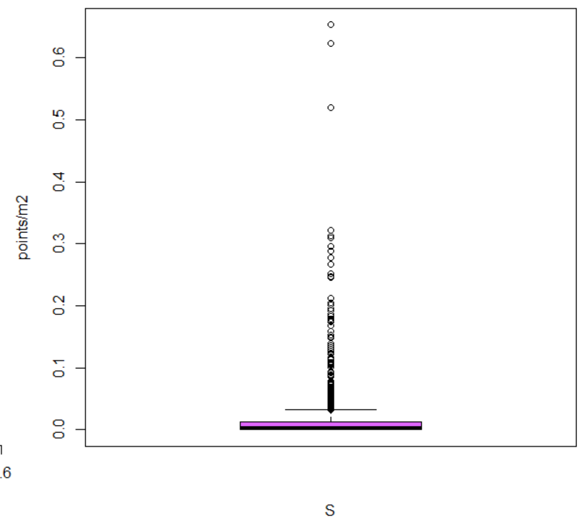
Figure 5. 3. Histogram of distribution, box-plot, histogram of log transformation, box-plot without 0 values, and statistical values of variable "W" (water features).

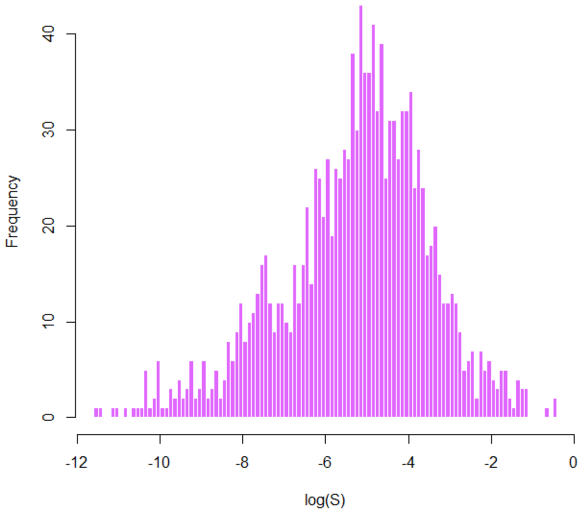
5.1.4. Buildings and Other Anthropogenic Structures

This factor was calculated based on LiDAR data. Therefore, its unit of measurement was the Lidar points of buildings and other anthropogenic structures per square meter of park area (points/m2).

Figure 5.4. shows similar data patterns to the water feature dataset. Building and anthropogenic structure data were also heavily skewed: almost 20% of all parks had no buildings or other anthropogenic structures. Many parks had very few or no buildings, with a resultant density of zero. This resulted in very low means, medians, and interquartile ranges.







	S
Min	0.000
1st Q.	0.000
Median	0.004
Mean	0.015
3rd Q.	0.013
St.Dev.	0.042
Max	0.653

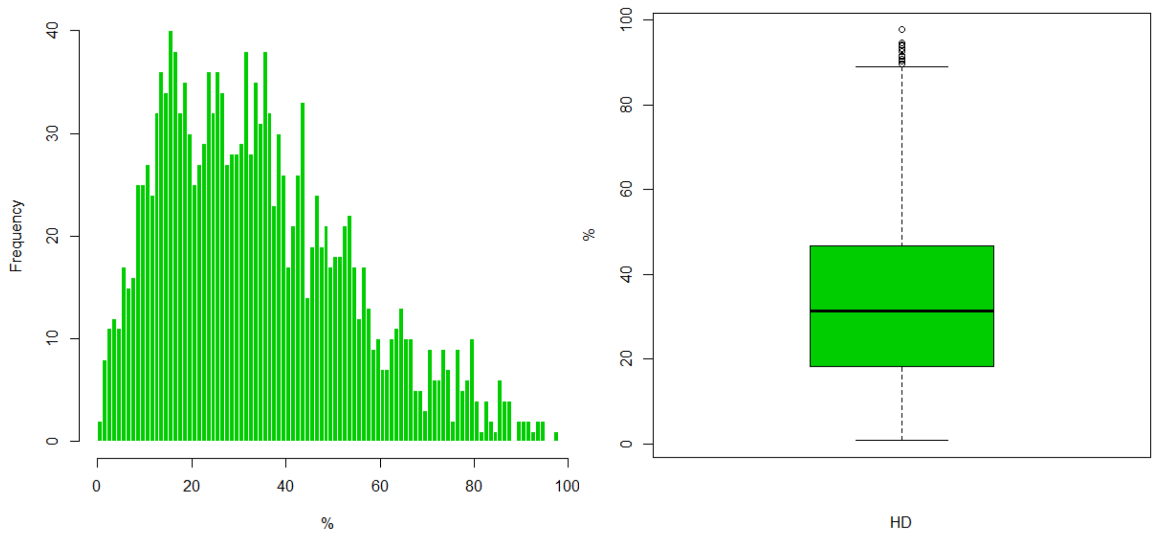
Figure 5. 4. Histogram of distribution, box-plot, histogram of log transformation and statistical values of variable "S" (buildings and other anthropogenic structures).

As can be observed from the box-plot, there were three parks with very high buildings and other anthropogenic structures values. These are the Barbican, Golden Lane Estate Designed Landscape, and The Water Gardens Designed Landscape areas, a unique mixture of residential and multi-art complexes with greenspaces and water features.

Log transformation of data showed that the data set had a frequency peak of around 0.04 points/m2, and most of the data were in the range of 0.001 points/m2 and 0.13 points/m2. Therefore, most parks did not have many buildings or other anthropogenic structures.

5.1.5. Horizontal Vegetation Density

As can be observed from Figure 5.5., the horizontal vegetation density factor was near normally distributed. Data were between 0.78 to 97.66 % of vegetation per park. Parks covered by more than 90% vegetation were therefore considered outliers. 50% of all data range from 18.16% to 46.66% of vegetation per park area. From the data, it is clear that parks were 30-35% covered by vegetation (grass, as the base surface cover of a park, does not count as vegetation in this research).

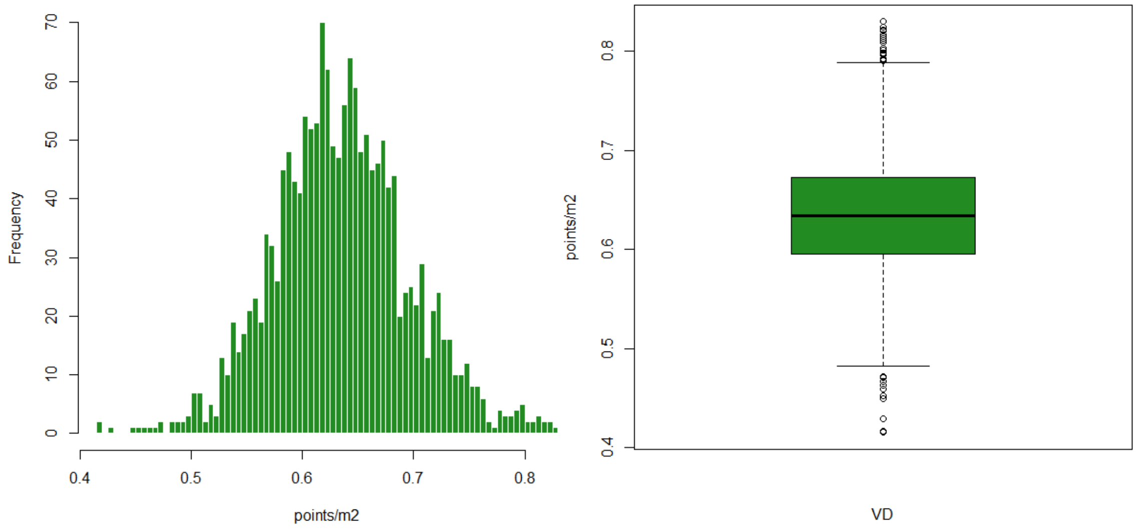


	HD
Min	0.774
1st Q.	18.161
Median	31.413
Mean	34.152
3rd Q.	46.659
St.Dev.	19.928
Max	97.664

Figure 5. 5. Histogram of distribution, box-plot and statistical values of variable "HD" (horizontal vegetation density).

5.1.6. Vertical Vegetation Density

Figure 5.6. shows that the vertical vegetation density factor was normally distributed. Data were between 0.42 to 0.83 LiDAR points per vegetated area square meter (points/m^2). Data lower than 0.49 points/m2 and higher than 0.78 points/m2 were classified as outliers. Half of the data ranged from 0.60 points/m2 to 0.67 points/m2. Therefore, it can be concluded that some individual parks had a mix of dense and sparse vertical density vegetation. However, most parks were of medium vertical density.



	VD
Min	0.415
1st Q.	0.595
Median	0.634
Mean	0.635
3rd Q.	0.673
St.Dev.	0.061
Max	0.830

Figure 5. 6. Histogram of distribution, box-plot, and statistical values of variable "VD" (vertical vegetation density).

5.2. Principal Component Analysis: Study of Influencing Factors.

Based on the six factors discussed above, a statistical analysis was used to develop the multi-dimensional classification of urban parks in London.

Before performing the variable classification analysis, values for all computed variables were normalised (min-max scaled) to transform original values from 0 to 1 to be on the same scale. Values were then standardised to create values centred around the mean with a unit standard deviation. This was based on the approach of Freedman et al. (2007).

To explore six selected factors and identify correlated variables, principal component analysis (PCA) was performed. This is a mathematical transformation of possibly correlated variables into a number of uncorrelated variables called principal components. The resulting components were defined so that the first principal component had the highest variance and accounted for most of the variability in the data (Felipe et al., 2021).

Theoretically, limiting the number of components to the number that accounted for a certain fraction of the total variance was possible. Figure 5.7 shows that the first five principal components explained 92% of the variation. This means that the first five principal components retained more than 90% of the information variances contained in the data.

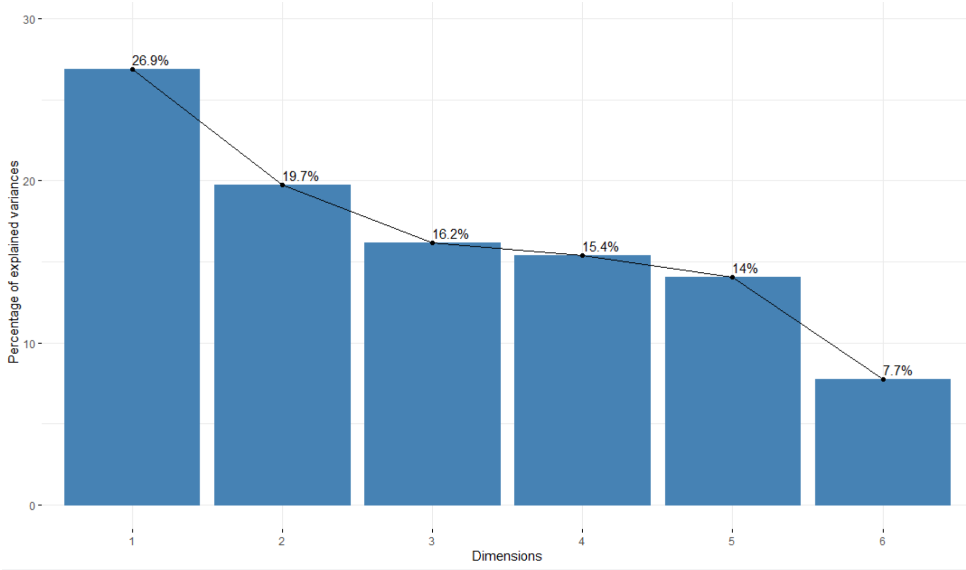


Figure 5. 7. Scree plot of eigenvalues ordered from largest to smallest. The first five principal components explain 92% of the variation.

However, there is no accepted objective way to decide how many principal components are enough (Peres-Neto et al., 2005). Furthermore, I would add that this study is interested in exploring all six factors. Therefore, all six components are used.

Table 5.1. shows the correlation between a variable and a dimension, which is used as the coordinates of the variable on the PC. The representation of variables differed from the plot of the observations. The observations were represented by their projections, but the variables were defined by their correlations (Abdi and Williams 2010).

Table 5.1. Correlations between variables and principal component.

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6
R	-0.387	0.547	-0.240	-0.488	0.468	0.189
W	-0.308	-0.458	-0.449	-0.543	-0.444	-0.035
S	-0.265	0.427	-0.659	0.514	-0.196	-0.106
HD	-0.824	-0.045	0.281	0.000	0.150	-0.467
VD	-0.787	-0.162	0.235	0.271	-0.168	0.444
Area	-0.029	-0.682	-0.377	0.230	0.580	0.034

Variable project values on each PC show how much weight they had on that PC. Table 5.1. and Figure 5.8. show that horizontal vegetation density (HD) and vertical vegetation density (VD) strongly influenced PC1 (Dimension 1). By contrast, the area of the parks (Area) and trails and pathways (R), as well as water features (W) and buildings and other anthropogenic structures (S), were more influential in PC2 (Dimension 2). Buildings and other anthropogenic structures (S) and water features (W) were also dominant variables for PC3 (Dimension 3) and PC4 (Dimension 4). In PC5 (Dimension 5) area of the parks (Area) had the strongest weight. Contrastingly, both vertical and horizontal vegetation densities strongly influenced PC6 (Dimension 6).

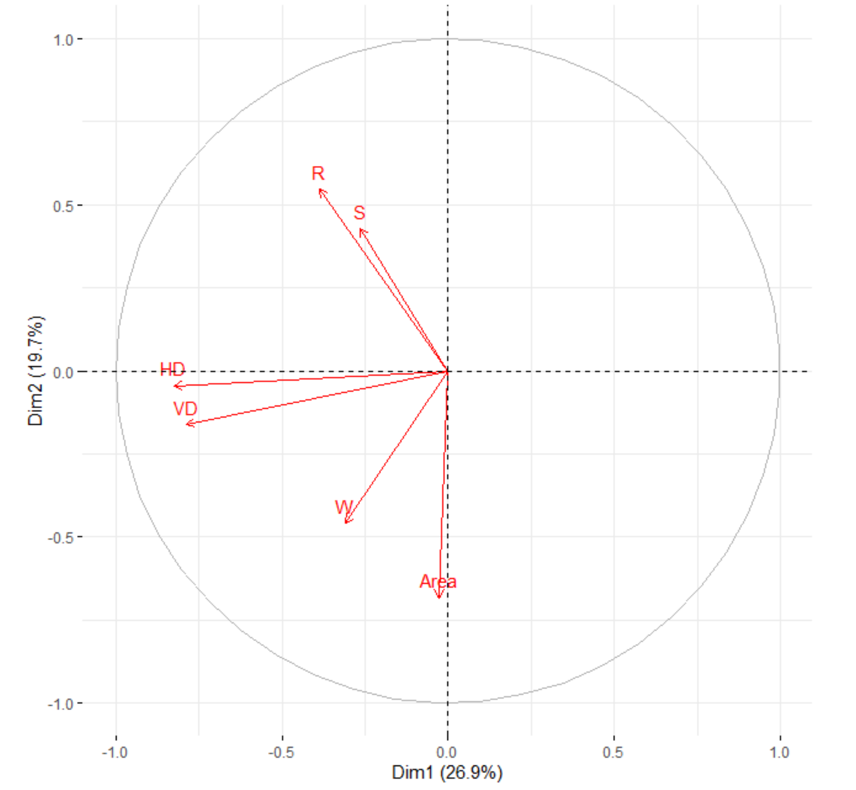


Figure 5. 8. Variable correlation plot for first two PC (explains 46.6% of the variation).

Figure 5.8. depicts a plot of the first two principal components, which explained 46.6% of the variation. It is therefore possible to estimate how variables correlated with one another based on the angles between the vectors. Although the first two principal components explained under 50% of the variation, they shared a similar correlation trend between variables with Pearson's correlation matrix (Figure 5.9.).

Horizontal vegetation density (HD) was positively correlated with vertical vegetation density (VD). Furthermore, trails and pathways (R) were positively correlated with buildings and other anthropogenic structures (S). In addition, horizontal and vertical vegetation densities (HD and VD) showed the highest correlation values, which makes sense as horizontal and vertical vegetation densities are similar in nature.

Park area (Area) was negatively correlated with trails and pathways (R). Additionally, buildings and other anthropogenic structures (S) and water features (W) were negatively correlated with buildings and other anthropogenic structures (S). This is evident because these sets of variables were positioned on opposite sides of the plot origin.

The area of the park (Area) and horizontal and vertical vegetation densities (HD and VD) are unlikely to be correlated. Furthermore, water features (W) and buildings with other anthropogenic structures (S) are unlikely to be correlated. This is evident as these sets of factors meet each other at a 90° angle on the plot or close to zero at the correlation matrix.

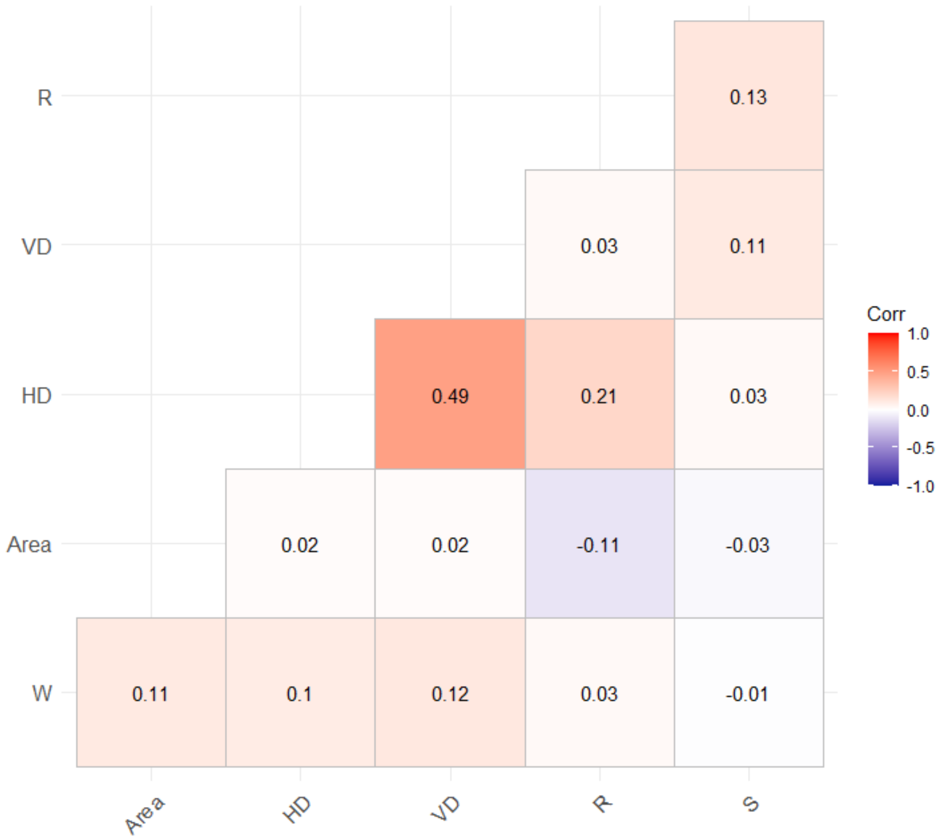
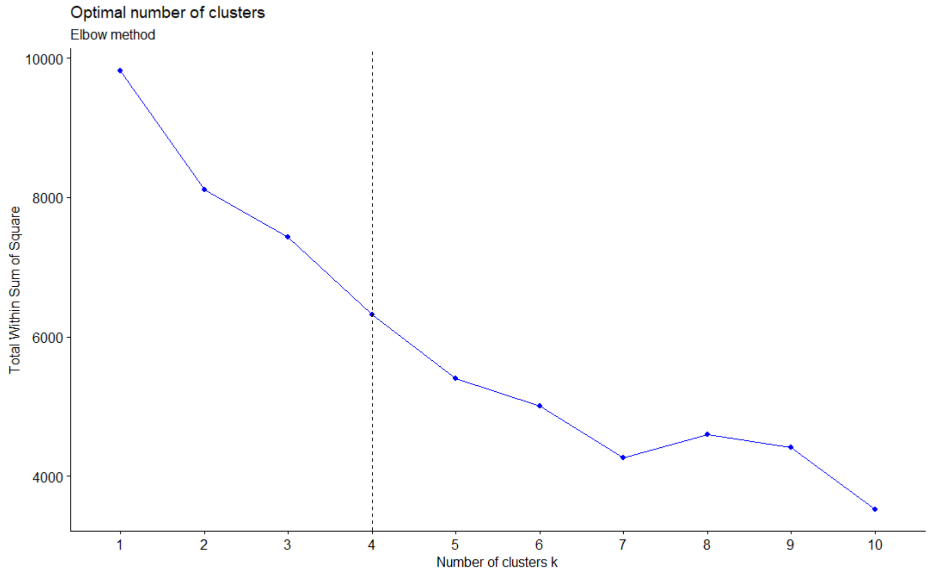


Figure 5. 9. Pearson's correlation matrix of all six variables.

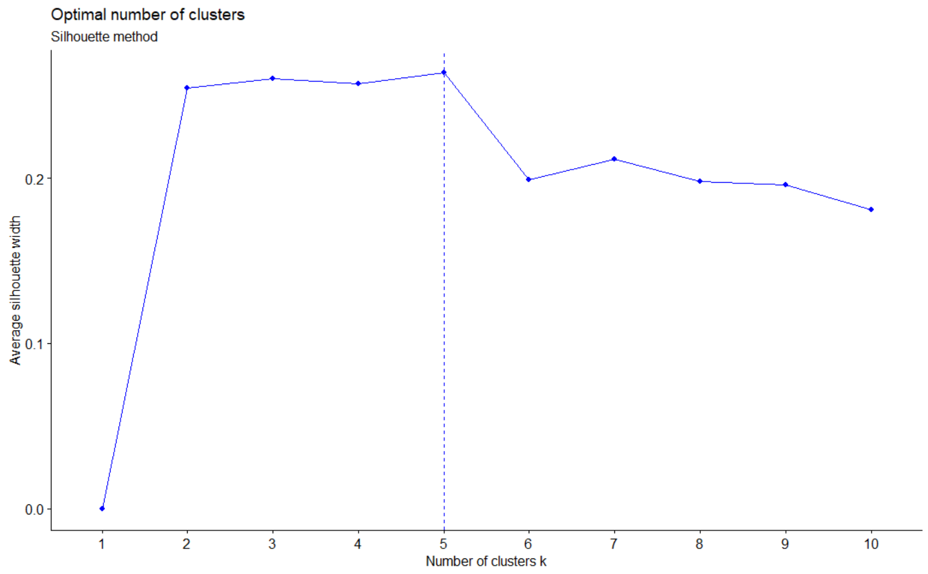
5.3. Cluster Analysis: Classification of the Typology of Urban Parks and Their Characteristics

K-Means clustering analysis outcome was used for the typology classification of 1639 urban parks. The parks were sorted into clusters or groups that shared similarities and were dissimilar to the objects belonging to another cluster.

Two different methods were used to determine the number of clusters in a data set. As can be seen from Figure 5.10., by applying the elbow method, four cluster groups were identified. By contrast, the silhouette method suggested using five clusters (Kumar, 2021). Both methods were applied, and the results showed that four clusters more accurately sorted parks by similarities.



(a)



(b)

Figure 5. 10. (a) Elbow method and (b) Silhouette method to determine the optimal number of clusters.

Four cluster groups were found by applying the elbow method, as indicated by the cluster plot in Figure 5.11.

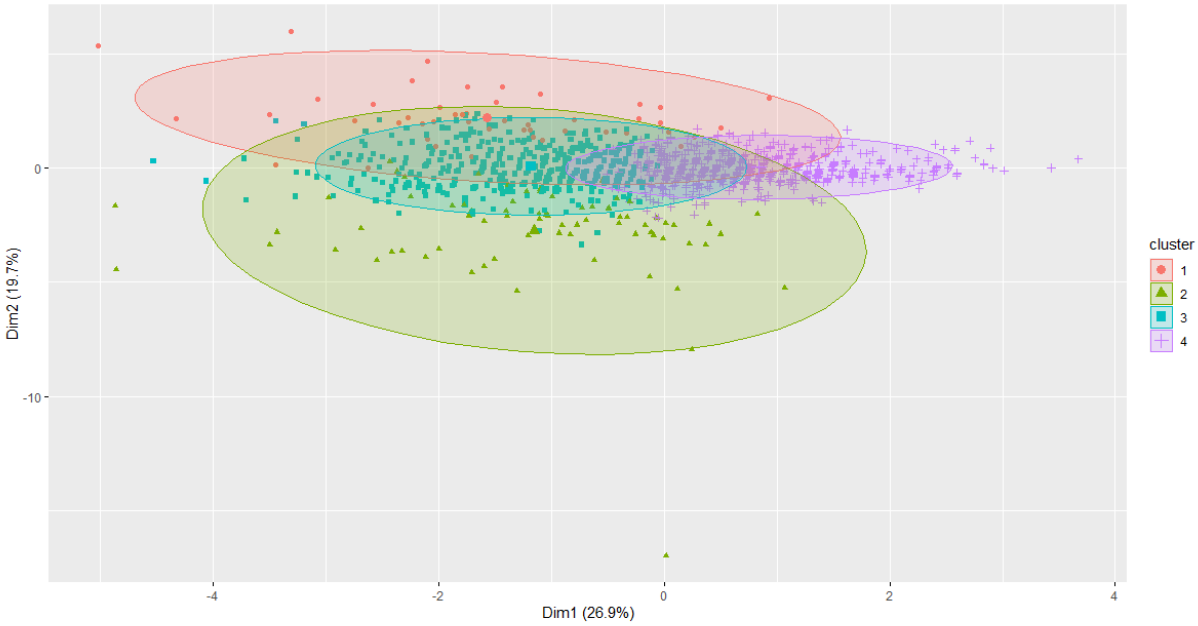


Figure 5. 11. The cluster plot of variables.

The classification result was presented on a map (Figure 5.12.) to display the spatial distribution of each urban park in London.

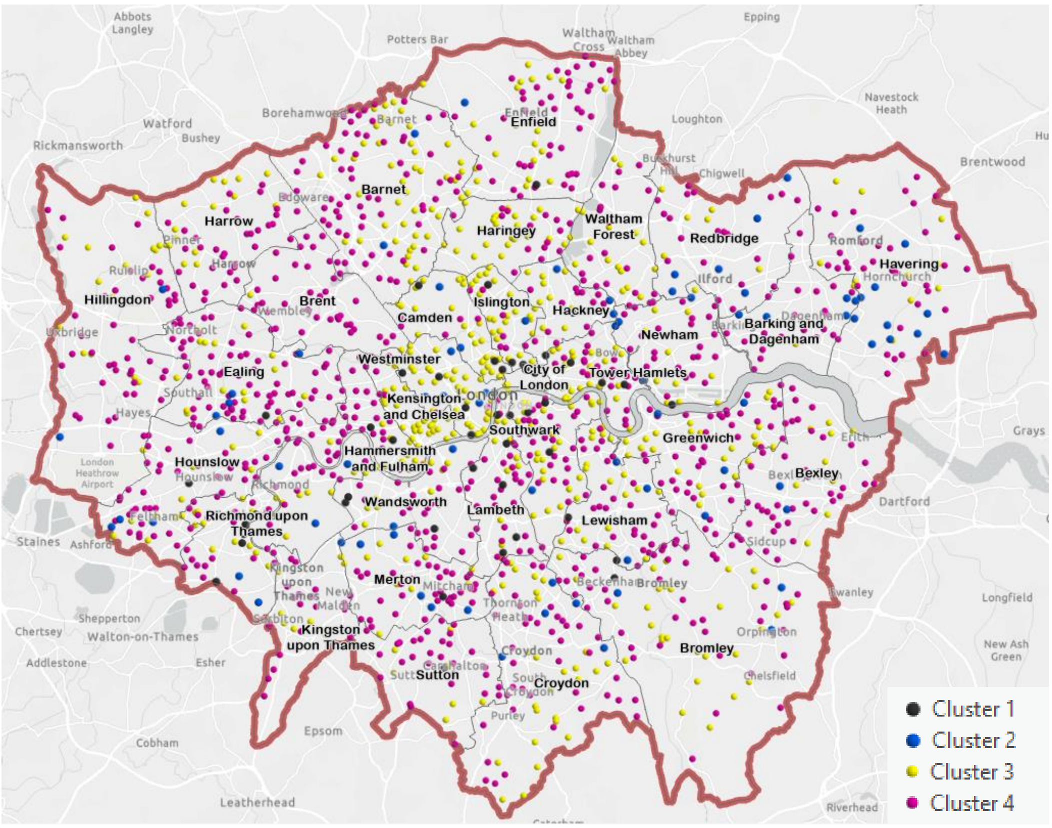


Figure 5. 12. Locations of the 1639 urban public parks in London are classified into four clusters.

Figure 5.17 shows a summary table for each cluster group, their essential characteristics, box plots, and the most typical park for each cluster.

The size of each group suggests that the highest number of parks was in cluster 4, which had 965 parks, 58.9% of the total studied parks. Cluster 3 had 551 parks, which was 33.6% of the total. Cluster 2 consisted of 74 parks, which was 4.5 % of the total. Finally, cluster 1 had the smallest number of parks, 48 parks, which was almost 3% of all parks included in this research.

5.3.1. Cluster 1

According to Figure 5.13 and Figure 5.17, cluster 1 contained parks with the highest average values of buildings and other anthropogenic structures among all groups. These parks were relatively small (75% of all parks in cluster 1 were less than 10 ha) but had dense trail and pathway networks. Most parks in cluster 1 had no or small water features, but some were outliers with as much as 9% of the park area. The parks were relatively high in horizontal vegetation density, with an average of 30% to 50% of the park's areas covered by medium or high vegetation with fairly high vertical density.

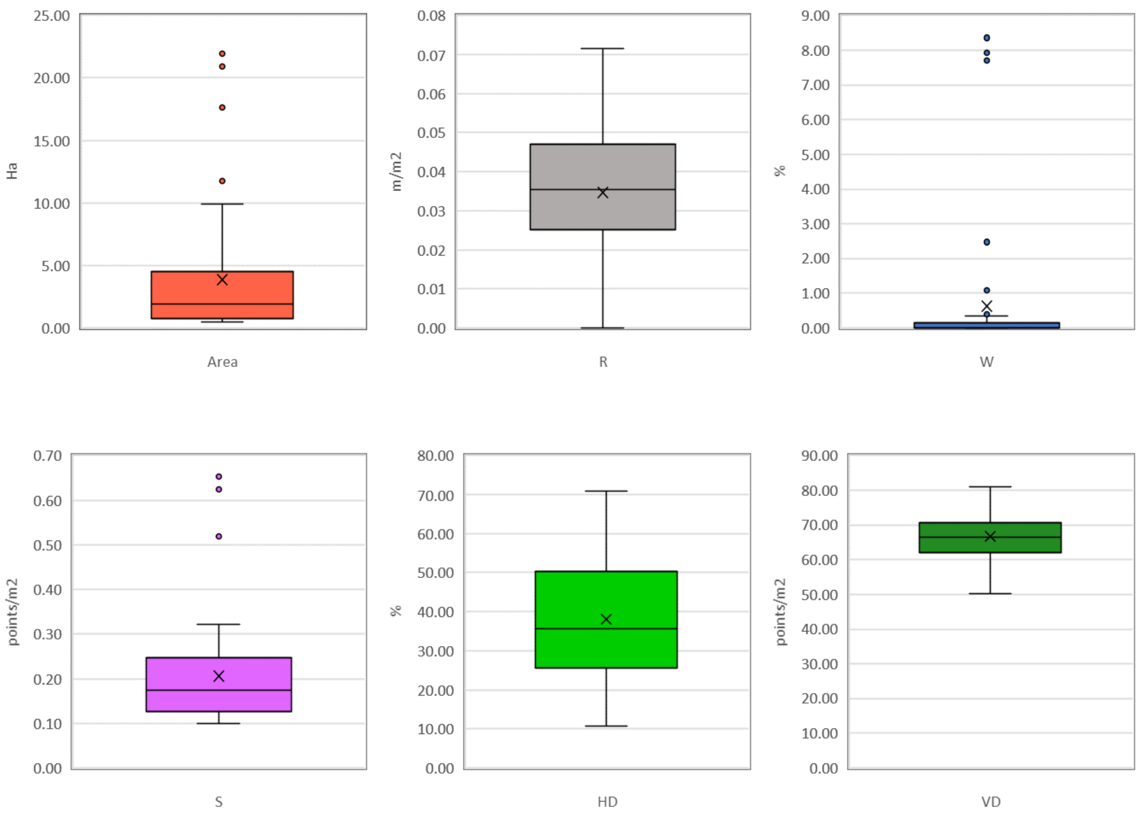


Figure 5. 13. Box-plots of the cluster 1 variables (original values of variables).

By examining each cluster, the most typical representatives of each cluster were found by using the median values of variables.

Further investigating parks in the cluster 1 suggested two types of parks in the group. The first are the palace and manor gardens, old Church gardens, and Churchyards. Examples include Lambeth Palace Garden, Pope's Garden, and York House Gardens. The second type is urbanised parks and estate gardens, such as Barbican Gardens, Brunel Estate, Golden Lane Estate Designed Landscape, and the Grays INN. These are outdoor spaces characterised as common parks, landscapes, and gardens of the estate.

5.3.2. Cluster 2

Cluster 2 had the second lowest number of parks: 74 parks. As seen in Figure 5.14. and Figure 5.17. this group included London's largest parks with the highest proportion of water features compared to other clusters.

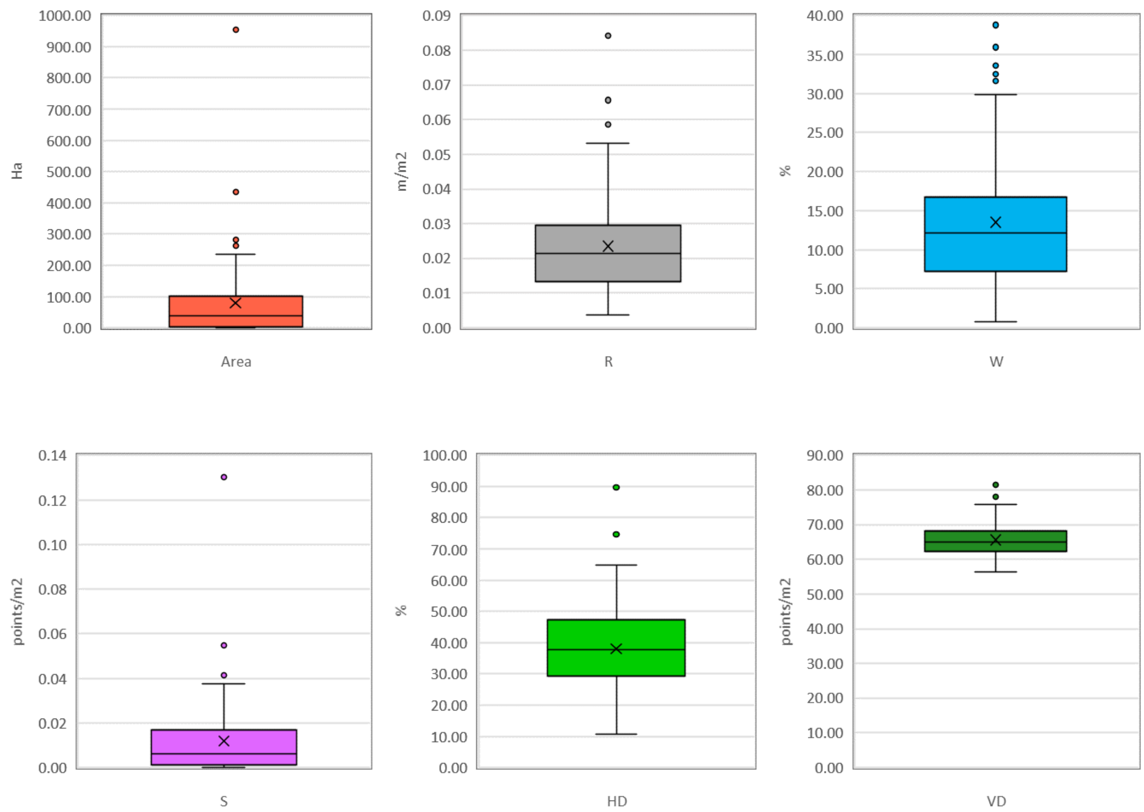


Figure 5. 14. Box-plots of the cluster 2 variables (original values of variables).

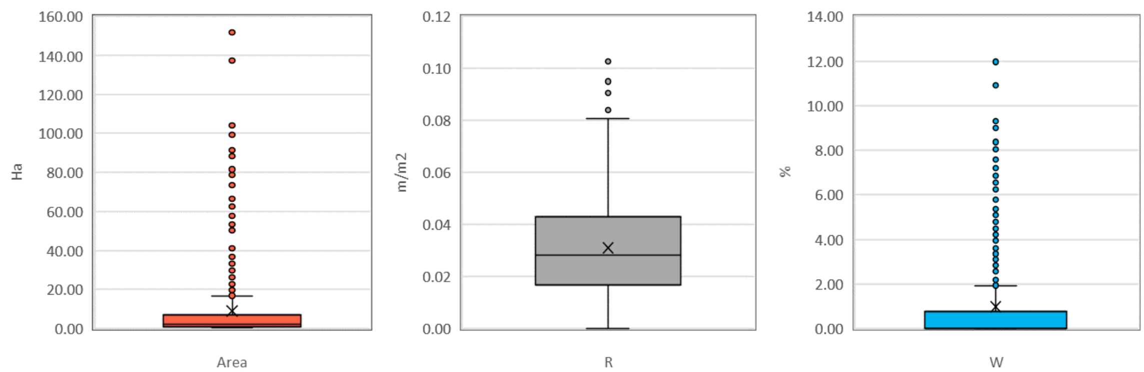
This cluster has many historical and cultural parks when compared to other groups. The defining feature of this group's parks is a wide range of recreational facilities that perhaps attract more visitors due to the larger area of parks than in other clusters. Parks had many open spaces but had fewer buildings and anthropogenic structures than the first cluster. Inside the parks, the landscape was full of natural elements, particularly water features like ponds or lakes. These were mixed with big trees, creating a green-blue infrastructure. Moreover, most parks had various facilities for various uses, including active and passive leisure activities, such as sports, recreation, education, and art. Cluster 2 is exemplified by Hyde Park, Kensington's Gardens, Regent's Park, Richmond Park, Battersea Park, and others.

5.3.3. Cluster 3

One-third of all parks in this research fitted within cluster 3. The parks in this group were diverse, as shown in Figure 5.15. and Figure 5.17. Most parks were relatively small. 75% of cluster 3 parks were under 18 Ha. However, there were some outliers which were up to 160 ha.

Another characteristic of this group is that they have few or almost no water features and few buildings and other anthropogenic structures. On the other hand, cluster 3 had the highest values of horizontal vegetation density, with an average of 40% to 65% of the park areas covered by medium or high vegetation and the highest vertical vegetation density values. Cluster 3 parks also had the second highest density of trails and paths.

Cluster 3 parks were rich in trees of varying sizes and vegetation, with plenty of shaded spaces for walks, outdoor exercise, and physical activity. These parks can therefore support relaxing after-work activities for residents.



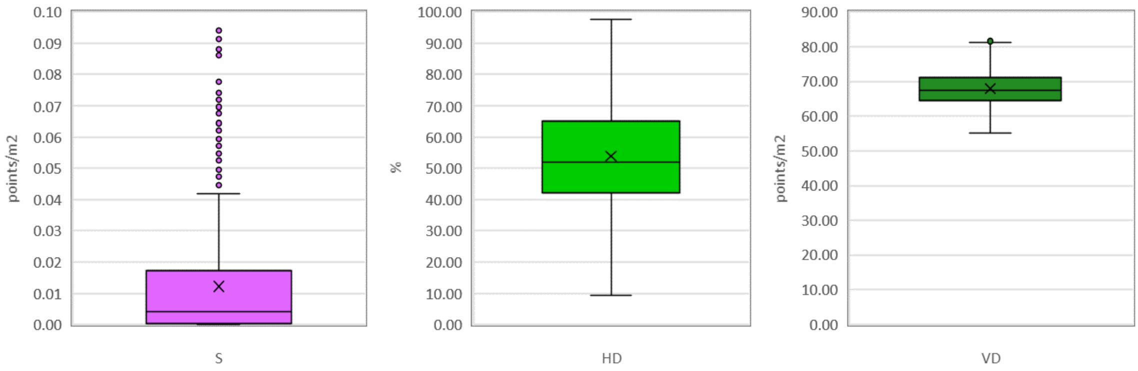
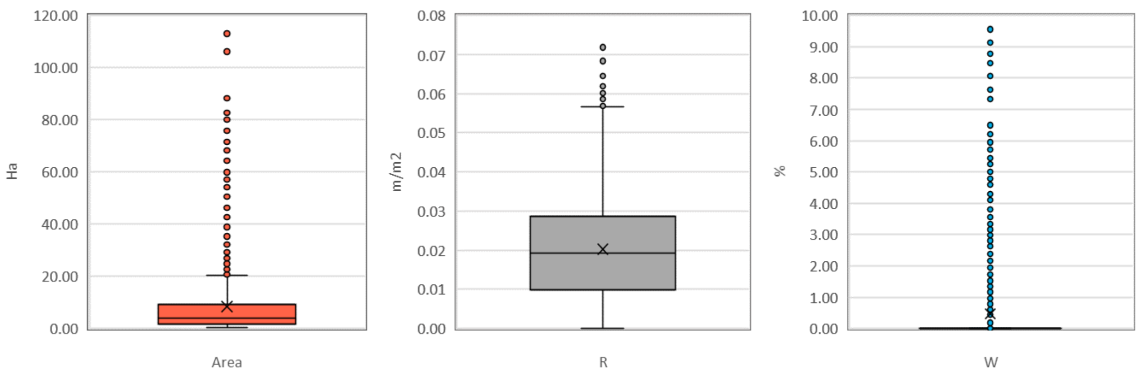


Figure 5. 15. Box-plots of the cluster 3 variables (original values of variables).

The typical parks of cluster 3 were Holland Park, Mayow Park, and forest parks, such as Cherry Tree Wood.

5.3.4. Cluster 4

Cluster 4 had the highest number of parks, containing almost 59% of all studied parks. Some features of cluster 4 parks were similar to those of cluster 3, such as park size and their limited number of water features. A typical characteristic of cluster 4 is the lower vertical and horizontal density of vegetation, shown in Figure 5.16. In 50% of cluster 4 parks, between 13% and 31% of the park area was covered by relatively sparse vegetation.



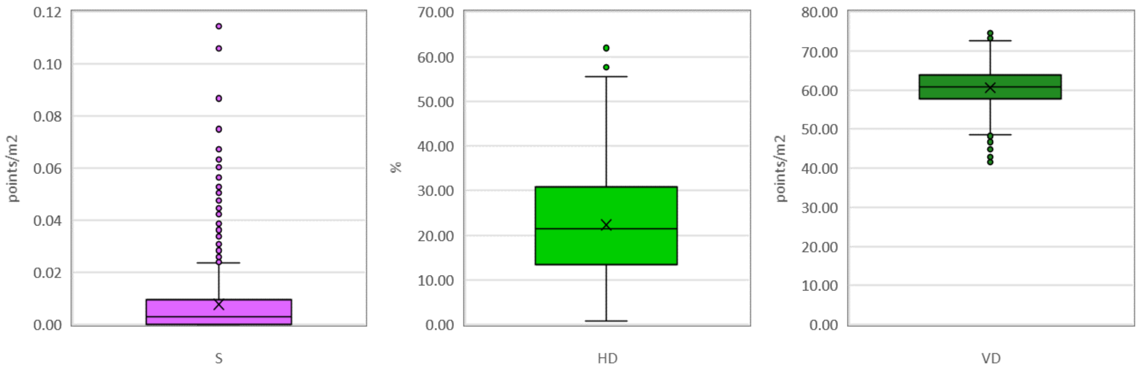
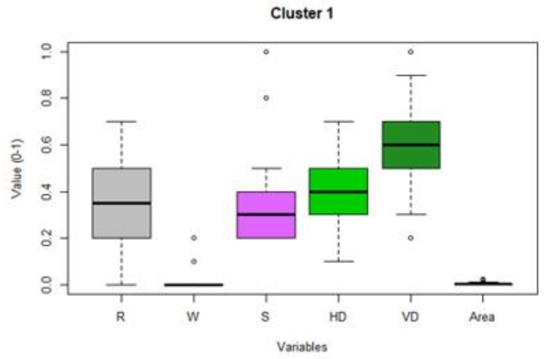
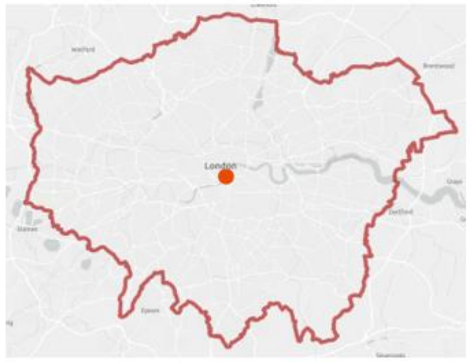
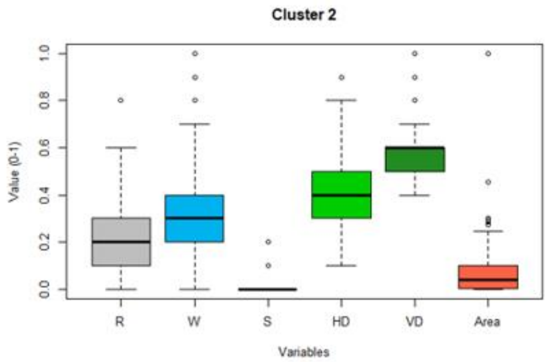
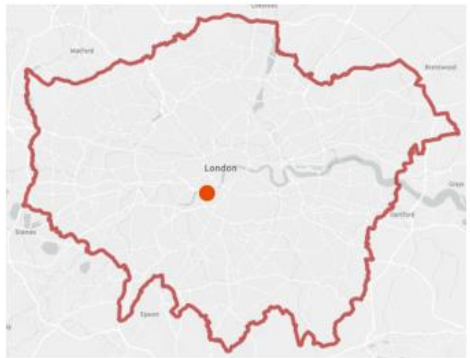


Figure 5. 16. Box-plots of the cluster 4 variables (original values of variables).

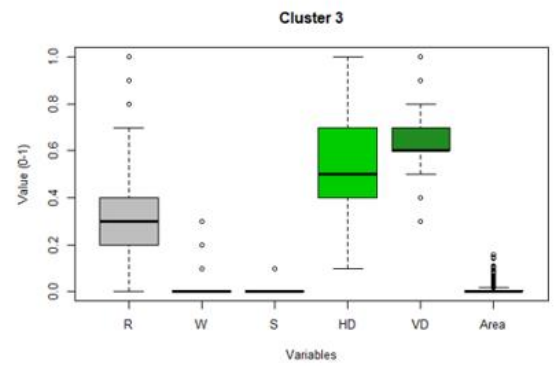
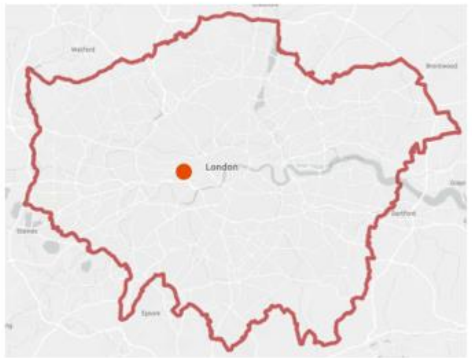
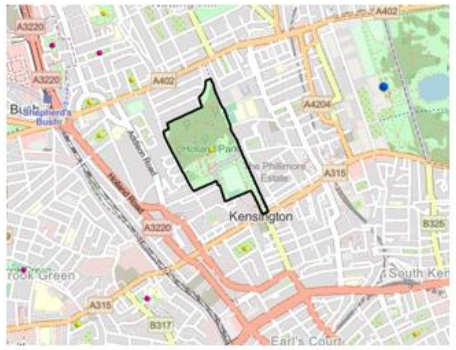
Cluster 4 parks play a significant role in the lives of local residents because they are evenly scattered across the city's boroughs. Furthermore, they have been designed to have beautiful surroundings to promote local aesthetics. Most of these parks can be used as play areas, as places for small events, or for taking lunch breaks. These parks could be a place of quiet refuge and escape from the busy city life. These parks can significantly increase the value of space. Creating additional water features, buildings, or other anthropogenic structures would further increase the value of these parks. These parks are exemplified by Acton Park, Maylands Fields, and others.

<table>
<tr><th rowspan="2">Cluster</th><th colspan="2">Description of Cluster</th><th colspan="3">Typical Examples of Parks</th></tr>
<tr><th>Box-plots</th><th>Dominant characteristics</th><th>Name and Location of the Park</th><th>Visualization in the map</th><th>Content</th></tr>
<tr><td>1

n = 48</td><td></td><td>1. Small in size
2. A high proportion of buildings and other anthropogenic structures
3. Dense trail and pathway networks
4. There are no or few water bodies
5. Relatively high in horizontal vegetation density
6. High vertical density</td><td>Lambeth Palace Garden
</td><td></td><td>
rowanwilliams.archbishopofcanterbury.org/gallery.php/71/lambeth-palace-garden</td></tr>
<tr><td>2

n = 74</td><td></td><td>1. Large area
2. Few buildings and other anthropogenic structures
3. The average density of trail and pathway network
4. Rich in water objects
5. Well-balanced vertical and horizontal vegetation density</td><td>Battersea Park
</td><td></td><td>
https://www.popsugar.co.uk/smart-living/photo-gallery/41360060/image/41360061/Battersea-Park</td></tr>
</table>

<table>
<tr>
<td>3

n = 551</td>
<td>
</td>
<td>1. Relatively small in size
2. Few buildings and other anthropogenic structures
3. There are no or few water bodies
4. A high density of trails and path networks
5. High horizontal vegetation density
6. High vertical vegetation density</td>
<td>Holland Park

</td>
<td>
</td>
<td>
https://www.expedia.com.au/Holland-Park-London.d6168475.Attraction?pwaDialog=gallery-open</td>
</tr>
<tr>
<td>4

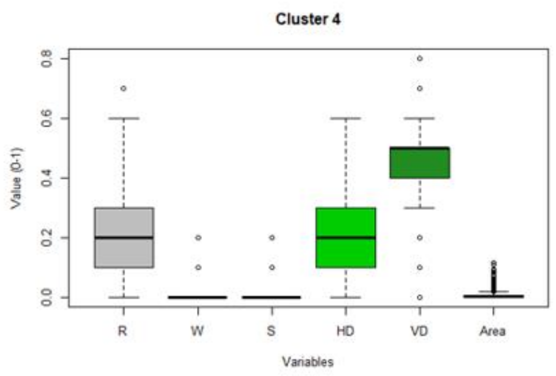
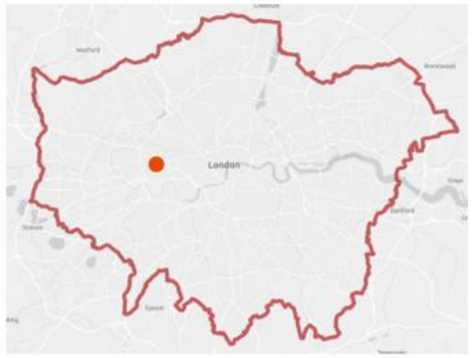
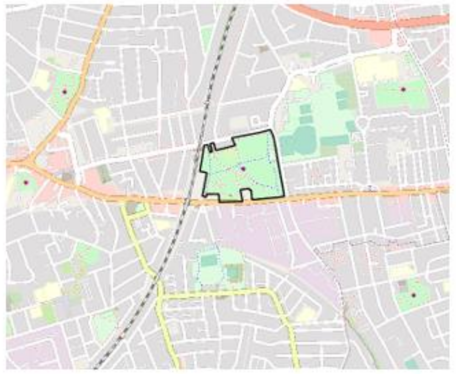
n = 965</td>
<td>
</td>
<td>1. Relatively small in size
2. Few buildings and other anthropogenic structures
3. There are no or few water bodies
4. The average density of trails and path networks
5. Many open spaces
6. Average vertical vegetation density</td>
<td>Acton Park

</td>
<td></td>
<td>
https://www.playfinder.com/london/venue/acton-park/tennis-outdoor-32514</td>
</tr>
</table>

Figure 5. 17. Summary table for each cluster group, their essential characteristics, box plots, and examples of the typical urban park for each cluster.

5.4. Urban Parks at the Borough Level

This research has analysed 1667 urban parks, showing that London is undoubtedly a green city. However, some of its boroughs offer more green space per person than others.

Figure 5.18 shows the area of urban parks per borough resident in square meters. It is important to note that only parks larger than 0.5 ha were analysed.

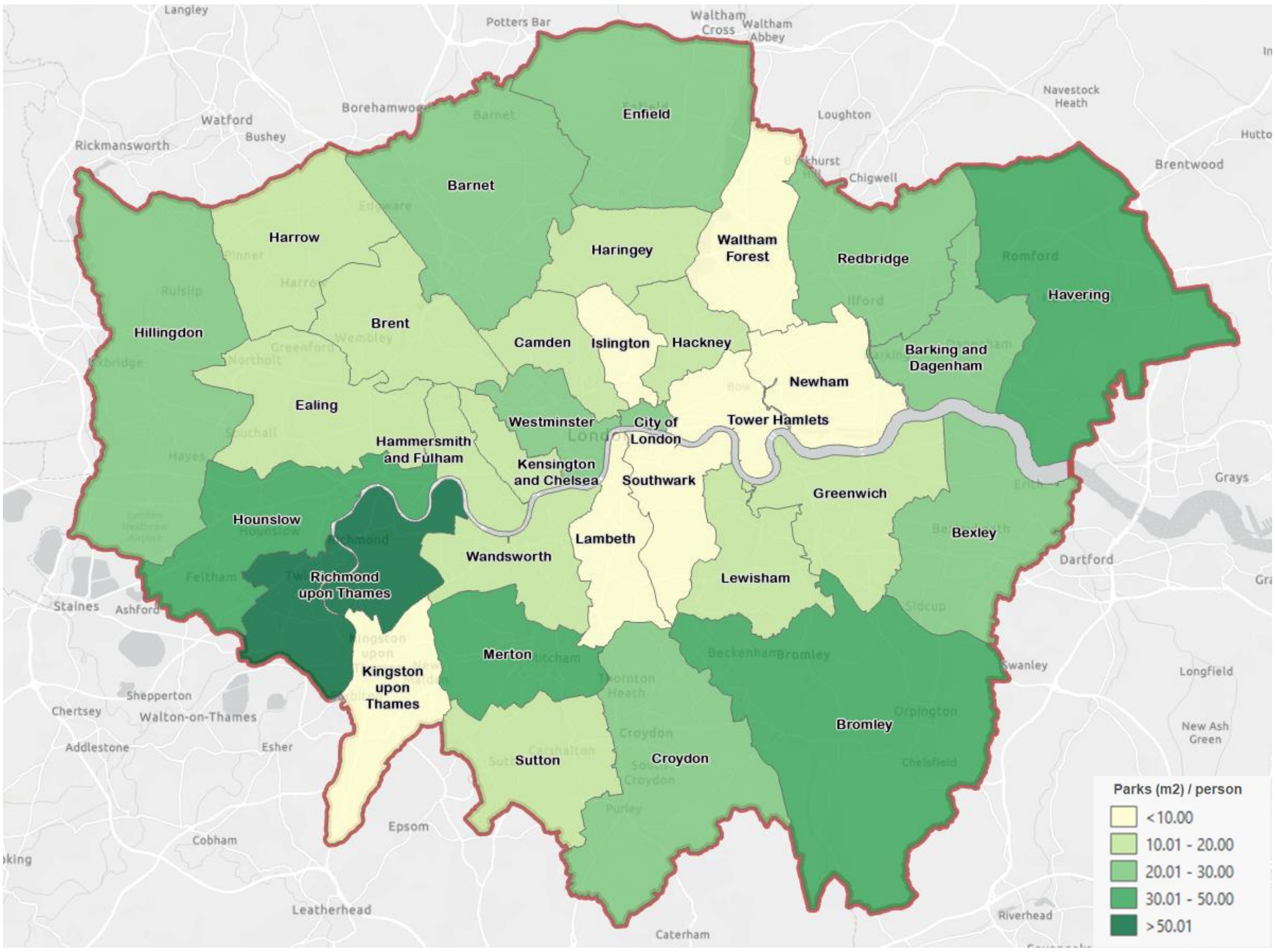
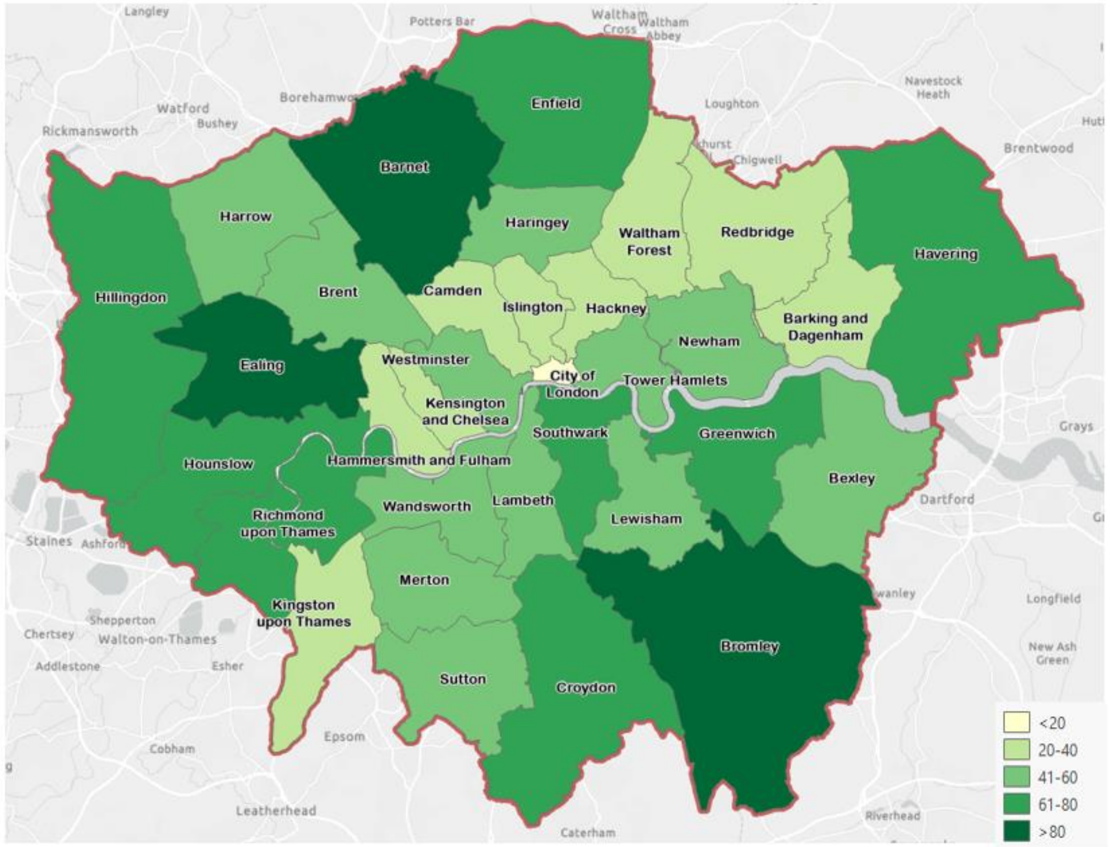
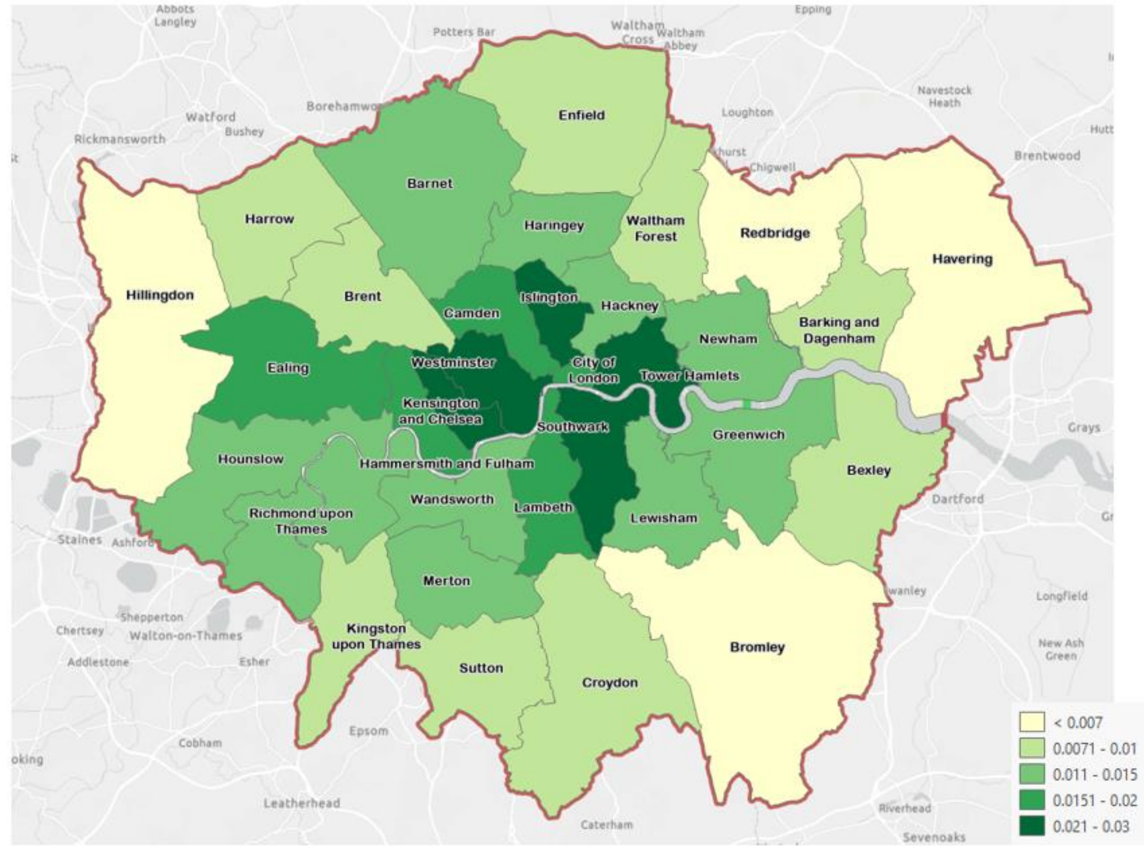


Figure 5. 18. The area of urban parks in square meters per borough resident.

The proportion of parks per person is higher in outer London boroughs, meaning districts farther away from the city centre. The City of London and Westminster are the only exceptions within the inner London boroughs: they contain between 20 and 30 m^2 of green space per borough resident. Richmond upon Thames is the richest London borough with sixty-six m^2 of parks per person.



(a) The number of urban parks per borough (pc.).



(b) The number of urban parks per borough, normalized by the area of the borough. (pc/Ha).

	Barking and Dagenham	Barnet	Bexley	Brent	Bromley	Camden	City of London	Croydon	Ealing	Enfield	Greenwich	Hackney	Hammersmith and Fulham	Haringey	Harrow	Havering	Hillingdon	Hounslow	Islington	Kensington and Chelsea	Kingston upon Thames	Lambeth	Lewisham	Merton	Newham	Redbridge	Richmond upon Thames	Southwark	Sutton	Tower Hamlets	Waltham Forest	Wandsworth	Westminster
Cluster 1					2	3	3		1	1		2	2					1	2			6	1	1	1		5	5	1	3		3	5
Cluster 2	5	2	2	1	3	2		5	3	1	4	3				11	2	4	1	1	1		3	5	3	6	5	1		1		2	5
Cluster 3	2	46	10	2	29	25	2	27	14	21	25	9	10	24	14	10	20	13	23	31	8	13	17	9	11	6	20	24	9	27	15	11	32
Cluster 4	20	65	38	37	48	5		42	69	46	43	14	21	15	34	37	47	56	6	4	21	30	30	34	27	22	36	30	31	23	19	25	3

(c) Distribution of the number of urban parks per cluster in boroughs.

Figure 5. 19. The number of urban parks per borough (a), the number of urban parks per borough, normalised by park area (b) and distribution of the number of urban parks per cluster in boroughs (c)

Looking at the number of parks in the boroughs (Figure 5.19. (a)), it is clear that Barnet, with its 113 parks, has a great collection of urban parks. Ealing and Broomley boroughs also have more than 80 parks. There are only five parks larger than 0.5 ha in the City of London, which is to be expected, because the City of London has the smallest area.

Therefore, analysing the number of urban parks is worthwhile by normalising them to the area of boroughs (Figure 5.19. (b)). This shows that most of inner London have the highest density of parks, whereas Havering, Redbridge, Bromley, and Hillington have the lowest density of urban parks.

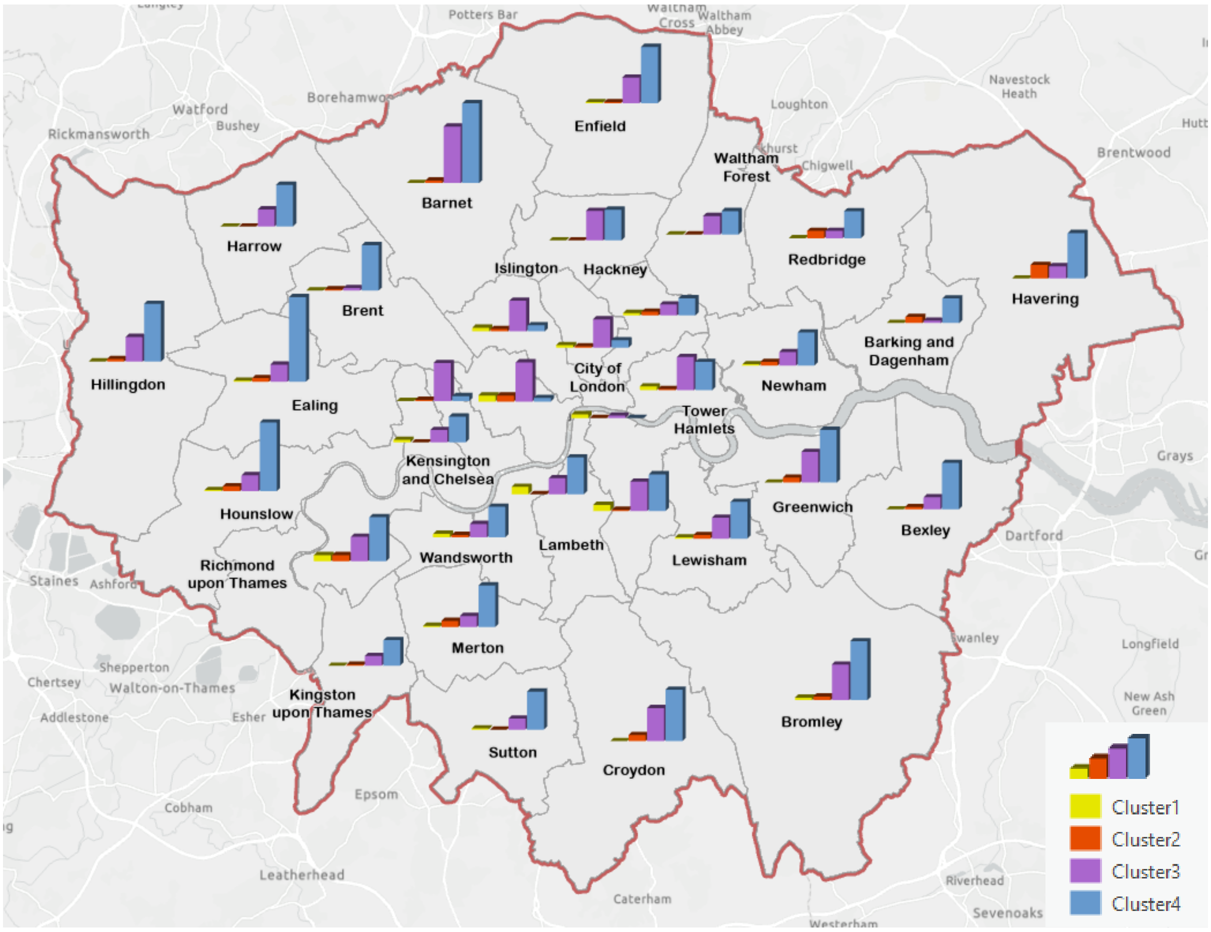


Figure 5. 20. The number of parks in boroughs by clusters.

Figure 5.20. further shows an analysis of the number of parks in boroughs by clusters.

By analysing the results of Figure 5.19. (c) and Figure 5.20, it is clear that just over half of boroughs have parks within cluster 1. These are relatively small parks with dense trail and pathway networks and many buildings and other anthropogenic structures.

It is not surprising that almost all inner London boroughs have cluster 1 parks, except Greenwich and Kensington and Chelsea.

Only one-fifth of all boroughs do not have cluster 2 parks, which are London's biggest parks with the highest proportion of water features. All eight Royal Parks of London are found within this cluster. Havering borough has 11 cluster 2 parks, which is the highest number of all boroughs.

As mentioned in sections 5.3.3 and 5.3.4, parks in cluster 3 have the highest values of horizontal vegetation density and a dense network of trails and paths. While some features of cluster 4 parks are similar to those of cluster 3, their dominant feature is sparser vegetation, which is less dense. Most parks with dense horizontal and vertical vegetation are in the boroughs of Westminster, Kensington, and Chelsea. In contrast, parks in the Ealing and Hounslow boroughs have sparser vegetation, which is typical for cluster 4 parks. Cluster 3 and cluster 4 parks dominate the Barnet borough.

5.5. Participant Study

Thirty people participated in the survey experiment to determine their assessments of the value of space within urban parks described in the previous sections and to understand which urban park cluster the respondents prefer.

As we see from the survey results in Appendix 3, of the total number of participants, twenty-three women (76.7% of participants) and seven men (23.3% of participants) took part in the survey.

Nine participants (30% of the total number) had professional experience, interest, or were involved in research on parks or green areas policy-making, management, evaluation, or monitoring. Of these nine professionals, only one was male.

Twenty-eight participants (73.3% of the total) were active visitors to urban parks or other urban greenspaces and visited them once or several times a week. Six participants (20% of the total) visited parks once or twice a month. Only two participants (6.7% of the total) said that they visit parks less than once a month. Professionals visit parks even more often. Eight participants (88.9% of the total) visited the park one or more times a week, and only one visited once or twice a month.

Respondents were asked if they would like to visit urban parks or green spaces even more often. Only four participants said they did not want to visit parks more often. The role that urban parks play in people's lives is therefore clear. People who visited parks several times a week were willing to visit them even more often. Seven respondents said they experience some barriers to visiting urban parks or green spaces. Most concerns were about the availability and accessibility of urban parks. Participants mentioned that they do not have nice, well-maintained green spaces or parks close enough to their homes. Close enough referred to parks being within walking distance. Some cited a lack of time to visit parks or an indoor lifestyle, and a lack of greenspace culture among friends. These responses once again demonstrate the importance of urban parks or green spaces being available within walking distance of where people live or where they work.

The participants were asked to imagine their ideal park and rank the six factors in order of importance, with #1 being the most important and #6 being the least important. By dividing respondents into two parts - professionals and non-professionals - two distinct opinion groups can be observed within the professionals (Figure 5.21. and Figure 5.22.).

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
HD	1	1	2	1	6	6	6	6	6
VD	2	2	4	4	2	5	5	5	5
W	4	5	5	2	3	4	3	2	2
R	3	3	1	5	5	3	2	3	3
S	5	6	6	3	1	1	1	1	1
Area	6	4	3	6	4	2	4	4	4

Figure 5. 21. Main question answer matrix. Group of nine respondents - professionals. A lower value indicates a more important factor for the respondent.

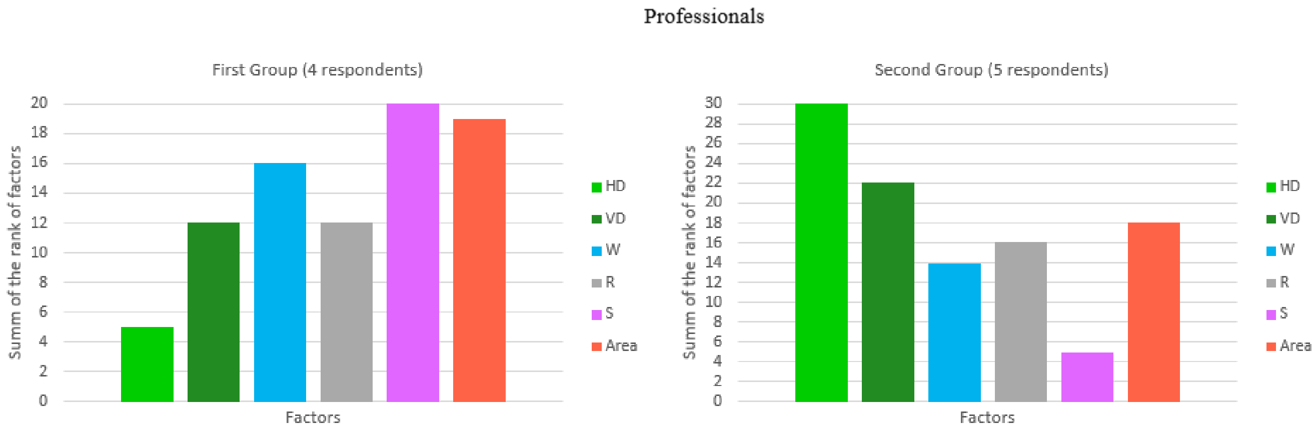


Figure 5. 22. Distribution of responses of two groups of professionals. A lower sum of factors indicates a more important factor for the respondents.

The professionals of the first group ranked and assessed the vertical and horizontal density of vegetation as the most important factor in parks. The least important to this group was the size of the parks and the presence of buildings and other anthropogenic structures in the parks. They would rather have city parks with a lot of vegetation, and with plenty of trails to walk and relax. Cluster 2 and cluster 3 parks correspond to this description.

By contrast, the most important factor to professionals in the second group was their preference for large parks with human impacts, such as buildings and other anthropogenic structures. Therefore, the density of horizontal and vertical vegetation was less important to them. This group valued urbanised parks and estate gardens, such as those within cluster 1, the most.

The responses of non-professionals were less disparate (Figure 5.23. and Figure 5.24.).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	10	16	17	18	19	20	21
HD	1	1	2	5	3	1	2	1	5	1	1	6	1	1	1	2	1	2	2	6	5
VD	2	6	3	4	1	6	1	2	6	2	2	4	2	2	4	1	2	1	3	3	6
W	4	3	1	6	2	2	5	4	4	3	5	2	4	4	3	3	3	3	4	4	4
R	5	2	4	2	5	5	4	5	2	5	3	3	5	5	5	4	5	5	6	2	3
S	6	4	6	1	4	3	6	6	1	6	6	1	6	6	6	5	6	6	1	1	1
Area	3	5	5	3	6	4	3	3	3	4	4	5	3	3	2	6	4	4	5	5	2

Figure 5. 23. Main question answer matrix. Group of twenty-one respondents – non-professionals. A lower value indicates a more important factor for the respondent.

As can be seen in Figure 5.24, natural elements, such as vertical and horizontal vegetation density and water features, were those most valued by non-professionals in urban parks. After this, these individuals valued the size of parks, the proportion of trails and paths, buildings, and other anthropogenic structures.

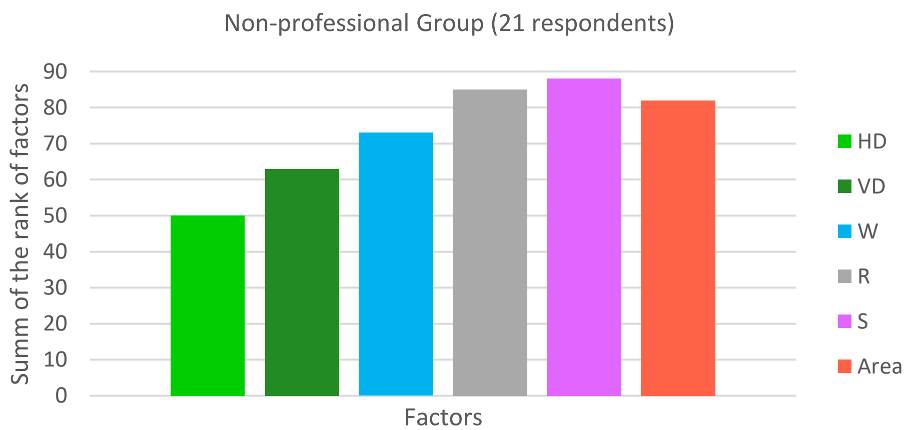


Figure 5. 24. Distribution of responses of non-professionals. A lower sum of factors indicates a more important factor for the respondents.

It can be concluded that this group would prefer the urban parks of cluster 2, which are rich in various natural elements, water resources, vegetation, and open spaces. These parks have few buildings and other anthropogenic structures.

6. DISCUSSION

In this chapter, the results presented in the previous chapter are discussed in relation to the research questions stated in chapter 1.2. Each research question is discussed and the specific uncertainties and limitations are pointed out at the end of each section. Finally, the study's general limitations are discussed in section 6.5.

6.1. RQ1: How Does the Structural Information from LiDAR Enhance the Existing Urban Park Data?

Interpretation of Results

The results support the theory that LiDAR is a powerful remote sensing dataset that is unique in its ability to characterise vegetation structure (Garcia et al., 2011; Kumar et al., 2015; Kim et al., 2020). The results of the study indicate that, using LiDAR data, it is possible to accurately analyse and resolve the horizontal and vertical vegetation structure, alongside more traditional urban land cover classes, such as buildings and other anthropogenic structures.

Building and horizontal vegetation density for this study were also determined using LiDAR data. However, this could have been done using other data resources, such as classic spectral satellite images or vector data. LiDAR data is unique in its ability to determine the vertical density of vegetation. Often three-dimensional information provided by LiDAR data is left unutilised (Kumar et al., 2015). Research proved that the additional height information and the other values extracted from LiDAR improved the quality of vegetation classification. The study identified how vertical vegetation density varies across London's urban parks. Still, it would be useful to carry out a structural analysis of vegetation not only for the park as a whole, but by analysing individual parts of the park.

Uncertainties and Limitations.

Although LiDAR data are a unique dataset, they have some limitations. One of the limiting factors for LiDAR use is the availability of data. There are still cities for which LiDAR data is unavailable or for which availability is limited. However, the availability of LiDAR data has increased substantially in recent years. Issues also occurred because of the massive amount of data. For this study, a Python-based data processing workflow was used. LiDAR data were filtered to remove anomalous and noisy returns. In the next step, overlay and duplicate points were removed. LiDAR tiles needed to be merged to successfully perform the data processing process, which was time and resource-consuming.

6.2. RQ2: Can Urban Public Parks be Classified Using Multi-Dimensional Clustering that Groups Parks According to Park Attributes?

Interpretation of Results

In this study, a multi-dimensional clustering approach was used to classify urban parks based on their characteristics, such as horizontal vegetation density, vertical vegetation density, the proportion of trails, the proportion of water, the proportion of anthropogenic structures, and park size. Two different methods were used to determine the number of clusters in a data set. The Elbow method suggested four cluster groups; however, the silhouette method recommended using five clusters. As explained in chapter 5.3, both methods were applied, and the results showed that four clusters more accurately sorted parks by similarities. Despite the size of the parks dataset (1,638 parks), the clustering of parks was easy to implement, and the results were credible and verifiable.

Uncertainties and Limitations

There were some limitations in the multi-dimensional clustering analysis, which may affect the accuracy of data analysis in classifying urban parks. One limitation is data outliers. All four clusters contained both very typical cluster members as well as less typical ones. This is due to the K-mean clustering being sensitive to data outliers. In some cases, removing outliers would be worthwhile before clustering is performed. In this study, each urban park was important and had a unique character and characteristics, so this was not carried out. A further limitation is the data sources used for this research. The integration of additional data (such as information about park infrastructure, including lighting and benches, possible activities in the park, solar exposure, and socio-economic data) could allow for a more comprehensive investigation. However, studies with existing data still offer valuable insight into the composition and distribution of urban parks.

6.3. RQ3: How do the clustering results correlate with respondents' perceptions about parks?

Interpretation of Results

As described in Chapter 5.5, the survey results showed that respondents were split into three groups regarding their valuation of city park qualities. Respondents who do not consider themselves as professionals related to urban greenspaces, most highly valued elements of nature, such as vegetation and water features, showing that they are looking for nature in urban parks. Of the four clusters of parks, the best match for these respondents would be cluster 2, which has parks rich in various natural elements, water resources, vegetation, and open spaces. Cluster 2 has few buildings and other anthropogenic structures. Battersea park, Bushy park, Kensington gardens, and Richmond park are some of the urban parks that non-professional respondents would therefore prefer.

The responses of professionals were divided into two opposing groups. The first group tended to agree with non-professionals partly and most highly rated more natural parks with trails. This group would enjoy park clusters 2 and 3. The second group of professionals valued human influences, buildings, and other anthropogenic structures. Cluster 1 parks would be the most enjoyable for respondents in this group.

Some of the participants gave a broader explanation about their choices and the qualities that they value in parks. One respondent stated that the "design" of the park was most important (such as the way that elements are created, maintained, and combined) and "affordances" were also important (for example, the possibilities for play, exercise, and socialising). Another respondent said that it was difficult to rank the characteristics because some could be rated equally. Often, participants' choice of park solely depended on how close the park was to their home. Safety was also an issue. Additionally, the importance of trails depended on the activity which the participants wanted to use the park for. Another participant stated that sometimes people go to parks to use the facilities in the anthropogenic structures, such as the sports centre. Although respondents did not rate anthropogenic structures highly, they were often the main reason for visiting the park in the first place. Therefore, if questions had been worded differently, or participants had been allowed to rank park characteristics equally, the results could have been different. On the other hand, then the study would risk that they rate all or most characteristics the same.

Uncertainties and Limitations

Future research could involve more survey respondents to obtain more reliable and representative survey results. Furthermore, there are also qualitative limitations regarding the participants' previous knowledge. Less than one-third of participants considered themselves experts with professional experience, interest, or involvement in research on parks or green areas policy-making, management, evaluation, or monitoring. It would be interesting to involve more experts because they were only 1/3 of all those questioned in the current survey.

6.4. General Limitations

This study has some limitations and weaknesses. The first limitation of the study relates to the determination of park boundaries. As mentioned in previous chapters, this study only intended to use open-source data. Open Street Map, the Ordnance Survey Open Greenspace database, and Historic England Park databases were combined to obtain a complete urban park database. Unfortunately, both the Open Street Map data set and the Ordnance Survey Open Greenspace database had inconsistent quality. They missed out parks or gardens, duplicated polygons, and wrongly classified green areas. Precision and accuracy varied across London's territory. It would therefore be advisable for park boundaries to use official data to ensure the data is accurate.

Another limitation directly related to park boundaries and open source data is the land use classification used in Open Street Map data. Although park-related objects (parks, gardens, and cemeteries) were selected from the OSM database, closer examination of the selection on Google Maps often showed that the classification was ambiguous. For example, golf courses were often classified as parks. Forests were also often classified as parks. Weaknesses in such classification affected the determination of park boundaries and categorising clusters. Furthermore, future research could include additional variables that would allow for a more in-depth data analysis.

7. CONCLUSION

Because of global urbanisation, natural areas and parks in urban contexts are becoming increasingly important for the quality of life in cities. This master's thesis has researched different types of parks in Greater London and the characteristics of these parks that people value most.

Six factors were used to characterise urban parks: horizontal vegetation density, vertical vegetation density, the proportion of trails, the proportion of water, the proportion of anthropogenic structures, and park size. Two primary open source data were used for the research: LiDAR data for extracting Horizontal and vertical vegetation information, anthropogenic structures, and an Open Street map for trail and path network and water object data layer. Open Street Map, the Ordnance Survey Open Greenspace database, and the Historic England Parks and Gardens database were used to detect the borders of London's parks. K-Means clustering was applied to classify the typology of urban parks along with six factors that influence the usage of the city public parks in London. A total of 1638 public parks were included in this study. The results of this research can be used to classify public parks, primarily based on the similarity of the parks' spatial context and physical characteristics, which were sorted out into four clusters. The results showed that the six factors used for characterising the urban parks for clustering accurately grouped parks by their unique characteristics.

RECOMMENDATIONS FOR FURTHER RESEARCH

Several suggestions for future research on the classification of urban parks can be made:

- It would be advisable to use official data for park boundaries to ensure data accuracy and to avoid other land uses being included within the park category.
- It would be useful to include more additional factors which influence urban park classification, such as recreation facilities, public transport services, number of roles that the urban park plays (such as recreation, heritage, health and wellness, natural environment, and education), and borough components (for example, socio-economic parameters).
- Future research could perform grid-based park analysis because parks properties and characteristics are not homogeneous. Therefore, average values are not necessarily representative of parks. It would be useful to calculate the average value of properties characterising the parks and the percentage of the park area with low values and the ratio of the park area with high values. This would provide a better understanding of the characteristics of the parks.

REFERENCES

General References

- Al-Hagla, K. (2008). Towards a sustainable neighborhood: The role of open spaces. International Journal of Architectural Research, 2, pp.162–177
- Baycan-Levent, T., and Nijkamp, P. (2009). Planning and management of urban green spaces in Europe: Comparative analysis. Journal of Urban Planning and Development, 135
- Bibri, S., Krogstie J. and Kärrholm, M. (2020). Compact city planning and development: Emerging practices and strategies for achieving the goals of sustainability, Developments in the Built Environment, 4 https://doi.org/10.1016/j.dibe.2020.100021
- Bolund, P. and Hunhammar, S. (1999). Ecosystem services in urban areas. Ecological Economics, 29(2), pp. 291−301
- Boyd, J. (2007). Nonmarket benefits of nature: What should be counted in green GDP? Ecological Economics, 61(4), pp. 716−723
- Burns, R. and Burns, R. (2008). Business Research Methods and Statistics Using SPSS. Los Angeles: Sage Publishing, pp. 145–149
- Chiesura, A. (2004). The role of urban parks for the sustainable city. Landscape and Urban Planning, 68(1), pp.129-138 https://doi.org/10.1016/j.landurbplan.2003.08.003
- Choumert, J. and Salanié, J. (2008). Provision of urban green spaces: some insights from economics. Landscape Research, 33(3), pp. 331−345
- Dong, P. and Chen, Q. (2017). Principles of LiDAR Remote Sensing. 10.4324/9781351233354-2
- Felipe, L., Gustavo, G., Ferreira, H., De Arruda, F., Silva, N., Comin, D., and Da Costa., L. (2021). Principal Component Analysis: A Natural Approach to Data Exploration. ACM Computing Survey, 54(4), https://doi.org/10.1145/3447755
- Fernandez-Diaz, J. (2011). Lifting the Canopy Veil - Airborne LiDAR for Archeology of Forested Areas. Imaging Notes, 26(2)
- Freedman, D., Pisani, R. and Purves, R. (2007). Statistics: Fourth International Student Edition. W.W. Norton & Company. ISBN 9780393930436
- Garcia, M., Riaño, D., Salas, C. and Danson, F. (2011). Multispectral and LiDAR Data Fusion for Fuel Type Mapping Using Support Vector Machine and Decision Rules. Remote Sensing of Environment 115(6), pp.1369–1379. doi:10.1016/j.rse.2011.01.017
- Harris, V., Kendal, D., Hahs, A. and Threlfall, C. (2018) Green space context and vegetation complexity shape people's preferences for urban public parks and residential gardens. Landscape Research, 43(1), pp.150-162, DOI: 10.1080/01426397.2017.1302571

- Jansson, M. (2014). Green space in compact cities: The benefits and values of urban ecosystem services in planning. Nordic Journal of Architectural Research, 26, pp.139–160
- Jenerette, G., Harlan, S., Stefanov, W. and Martin, C. (2011) Ecosystem ser-vices and urban heat riskscape moderation: Water, green spaces, and social inequality in Phoenix, USA. Ecological Applications, 21, pp.2637–2651
- Jiyeon K., Popescu S., Lopez, R., Wu, X. and Silvy, N. (2020) Vegetation mapping of No Name Key, Florida using lidar and multispectral remote sensing. International Journal of Remote Sensing, 41(24), pp.9469-9506, DOI: 10.1080/01431161.2020.1800125
- Klingberg J., Konarska, J., Lindberg, F., Johansson, L. and Thorsson, S. (2017). Mapping leaf area of urban greenery using aerial LiDAR and ground-based measurements in Gothenburg, Sweden. Urban Forestry & Urban Greening, 26, pp.31-40 https://doi.org/10.1016/j.ufug.2017.05.011
- Konijnendijk van den Bosch, C., Annerstedt, M., Nielsen, A. and Maruthaveeran, S. (2013). Benefits of urban parks. A systematic review. A Report for IFPRA. Copenhagen & Alnarp
- Kongphunphin, C. and Srivanit, M. (2021). A Multi-Dimensional Clustering Applied to Classify the Typology of Urban Public Parks in Bangkok Metropolitan Area, Thailand. Sustainability, 13 https://doi.org/10.3390/ su132011426
- Kumar, J., Weiner, J., Hargrove, W., Norman, S., Hoffman, F. and Newcomb, D. (2015). Characterization and Classification of Vegetation Canopy Structure and Distribution within the Great Smoky Mountains National Park Using LiDAR. 10.1109/ICDMW.2015.178
- Lyytimäki, J. and Sipilä, M. (2009). Hopping on one leg – The challenge of ecosystem disservices for urban green management. Urban Forestry & Urban Greening 8, pp.309-315
- Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. International Journal of Psychology Research, 3, pp.97–110
- Pataki, D., Carreiro, M., Cherrier, J., Grulke, N., Jennings, V., Pincetl, S., Pouyat, R., Whitlow, T. and Zipperer, W. Coupling biogeochemical cycles in urban environments: ecosystem services, green solutions, and misconceptions. Frontiers in Ecology and the Environment 9(1), pp.27-36
- Peres-Neto, P., Jackson, D. and Somers, K. (2005) How Many Principal Components? Stopping Rules for Determining the Number of Non-Trivial Axes Revisited. Computational Statistics and Data Analysis, 49, pp.974–997. http://dx.doi.org/10.1016/j.csda.2004.06.015
- Pyszny, K., Sojka, M. and Wróżyński, R. (2020). LiDAR based urban vegetation mapping as a basis of green infrastructure planning. E3S Web of Conferences. 171. 10.1051/e3sconf/202017102008

- Russo, A. and Giuseppe, C. (2018). Modern Compact Cities: How Much Greenery Do We Need? International Journal of Environmental Research and Public Health, 15 doi:10.3390/ijerph15102180
- Shan, J. and Toth, C. (2017). Topographic Laser Ranging and Scanning: Principles and Processing. (2nd ed). CRC Press. https://doi.org/10.1201/9781315154381
- Stanley, B., Stark, B., Johnston, K. and Smith, M. (2012). Urban Open Spaces in Historical Perspective: A Transdisciplinary Typology and Analysis. Urban Geography. 33, pp.1089-1117. 10.2747/0272-3638.33.8.1089
- Sulc, I., Morgado, S., Dordevic, Z., Gasparovic, S., Radovic, V., Keranova, D. (2020). Societal Issues and Environmental Citizenship. In: , et al. Conceptualizing Environmental Citizenship for 21st Century Education. Environmental Discourses in Science Education, vol 4. Springer, Cham. https://doi.org/10.1007/978-3-030-20249-1_4
- Swanwick, C., Dunnett, N., and Woolley, H. (2003). Nature, role and value of green spaces in towns and cities: an overview. Built Environment, 29(2), pp. 94-106
- Taylor, L. and Hochuli D. (2017). Defining greenspace: Multiple uses across multiple disciplines. Landscape and Urban Planning, 158, pp.25-38 https://doi.org/10.1016/j.landurbplan.2016.09.024
- Thurston, Hazel. Royal Parks for the People: London's Ten. UK and USA: David and Charles. Vancouver: Douglas, David and Charles. 1974. ISBN 0-7153-6454-5
- White, M., Smith, A., Humphryes, K., Pahl, S., Snelling, D. and Depledge, M. (2010). Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes. Journal of Environmental Psychology, 30(4), pp. 482−493
- Ziyue C., Bing X., Bingbo, G. (2015). Assessing visual green effects of individual urban trees using airborne Lidar data. Science of The Total Environment, 536, pp.232-244 https://doi.org/10.1016/j.scitotenv.2015.06.142

Online References

- Greater London Authority (2021). Available at: Green and Resilient Spaces Fund | GLA (london.gov.uk)
- Kumar, A. (2021). Elbow Method vs Silhouette Score – Which is Better? Data Analytics, Available at: https://vitalflux.com/elbow-method-silhouette-score-which-better
- Number of people living in urban and rural areas. Our World in Data. Available at: https://ourworldindata.org/grapher/urban-and-rural-population. [Accessed: 4-April-2022].
- https://resources.arcgis.com/

- Ritchie, H. and Roser, M. (2018). Urbanization. Published online at OurWorldInData.org. Available at: https://ourworldindata.org/urbanization [Online Resource]
- O'Connor, T (2008) The benefit of nature on nurture. Available at: https://www.smh.com.au/environment/the-benefit-of-nature-on-nurture-20080520-2gfs.html [Online Resource]
- Smith J. (2018) A Brief History of London's Parks. Available at: A Brief History of London's Parks (parkgrandlancastergate.co.uk)

Planning Documents

- European Commission. (2013). Sustainable development in the European Union: 2011 monitoring report of the EU sustainable development strategy: 2011 edition, Publications Office. Available at: https://data.europa.eu/doi/10.2785/1538
- Miller, R. (1997). Urban Forestry – planning and managing urban greenspaces. (2nd ed). New Jersey: Prentice Hall
- United Nations, Department of Economic and Social Affairs, Population Division. (2019). World Population Prospects 2019: Highlights. ST/ESA/SER.A/423.
- World Development Indicators – World Bank. (2021). World Bank based on World Population Prospects - UN Population Division
- United Nations, Department of Economic and Social Affairs, Population Division (2018)
- United Nations. (2018). The 2030 Agenda and the Sustainable Development Goals: An opportunity for Latin America and the Caribbean (LC/G. 2681-P/Rev. 3), Santiago
- WHO Europe Urban Green Space Expert Panel. (2017). Urban green spaces: a brief for action. Available at: https://www.who.int/europe/publications/i?healthtopics=248648e5-e2d3-49b2-8fd1-c5a542adcca4

Data Sources

- Ordnance Survey OpenData: Available at: https://osdatahub.os.uk/downloads/open
- OSM data: Available at: https://www.openstreetmap.org/
- Ordnance Survey Greenspace data: Available at: https://www.data.gov.uk/dataset/5d009d8a-702b-4a88-bf71-d4d6df87df53/os-open-greenspace
- Historic England Parks data: Available at: https://www.data.gov.uk/dataset/88cfe0de-85cd-431f-9836-2bee841d8165/registered-parks-and-gardens-gis-data
- LiDAR data: Available at: https://environment.data.gov.uk/
- Open Data London, Census data: Available at: https://data.london.gov.uk/dataset

Code

- https://www.github.com/ dkirsteina/MscThesis.git

Appendix 1

```
[out:json][timeout:200];
area["name"="Greater London"][admin_level=5]->.search;
(
way[leisure="park"](area.search);
way[leisure="garden"](area.search);
way[landuse="cemetery"](area.search);
    );
out tags geom;
relation[leisure="park"](area.search)->.relations_parks;
(
way(r.relation);
     );
out tags geom;
```

Code 1. Overpass query for extracting park features.

```
[out:json][timeout:200];
area["name"="Greater London"][admin_level=5]->.search;
(
way[highway="bridleway"](area.search);
way[highway="cycleway"](area.search);
way[highway="footway"](area.search);
way[highway="path"](area.search);
way[highway="pedestrian"](area.search);
way[highway="service"](area.search);
way[highway="steps"](area.search);
way[highway="track"](area.search);
       );
out geom;
```

Code 2. Overpass query for extracting trail and pathway network.

```
[out:json][timeout:200];
area["name"="Greater London"][admin_level=5]->.search;
(
way[natural="water"](area.search);
way[waterway="stream"](area.search);
way[waterway="river"](area.search);
way[waterway="ditch"](area.search);
way[waterway="canal"](area.search);
way[waterway="drain"](area.search);
                    );
out geom;
```

Code 3. Overpass query for extracting water objects.

Appendix 2

Section 1 of 3

Questionnaire to evaluate the importance of different urban park characteristics

Form description

Introduction

It is expected that by 2050, almost 70% of the world's population will be living in urban areas. Urban spaces need to be safe and attractive places for living and working. Natural areas and parks in urban contexts are becoming increasingly important for the quality of life in cities. Urban green spaces provide residents, tourists, and municipalities with many environmental, social, and economic benefits, and, crucially, have physical and mental health benefits. Considering the promising role of green spaces, it is important to carry out studies on different types of parks and what characteristics of parks are most valuable for people.

In this work, I use six factors to characterize urban parks: horizontal vegetation density, vertical vegetation density, proportion of trails, proportion of water, proportion of human-made structures, and park size. Of course, these are not the only factors that characterize a park, but they are the ones that have emerged as distinct and relevant in my MSc research project.

Three of the factors (horizontal vegetation density, vertical vegetation density and proportion of human-made structures) were acquired by using LiDAR (Light Detection and Ranging), an optical remote-sensing technique that uses laser light to densely sample the surface of the Earth producing highly accurate three-dimensional measurements about the shape of the Earth and its surface characteristics like terrain, tree canopy, sub-canopy and vegetation. LiDAR yields information that is not available in two-dimensional images of the landscape provided by traditional multi-spectral remote sensing platforms, therefore, providing detailed information about three-dimensional vegetation horizontal and vertical structure that is important to park and another greenspace assessment.

The rest of factors (proportion of trails, proportion of water objects and the size of the park) were acquired by using open source data (OpenStreetMap).

In this short questionnaire, you will be asked to rate six park characteristics from most to least important. At the start, you will also be asked four questions about your person. The completion of this questionnaire will take no longer than 10 minutes. Thank you for your participation!

Description (optional)

Gender *

- Female
- Male
- Queer/non-binary
- Prefer not to say

Do you have professional experience, interest or are you involved in research on park/green areas policy-making, management, evaluation or monitoring? *

- Yes
- No

How often do you visit urban parks and/or other urban greenspaces? *

- Several times a week
- About once a week
- Once or twice a month
- Less than once a month
- Almost never

Would you like to visit urban parks and/or other urban greenspaces more often? *

- Yes
- No

Do you experience barriers to visiting urban parks and/or other urban greenspaces? *

- Yes
- No

After section 1 Continue to next section

Section 2 of 3

Section title (optional)

Description (optional)

If you experience barriers to visiting urban parks and/or other urban greenspaces, what are they?

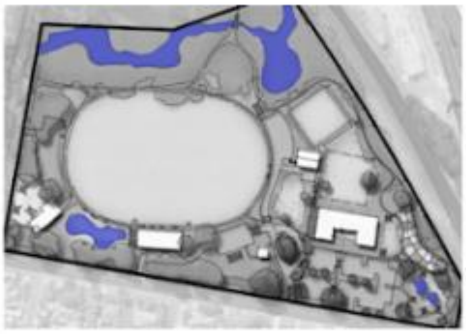
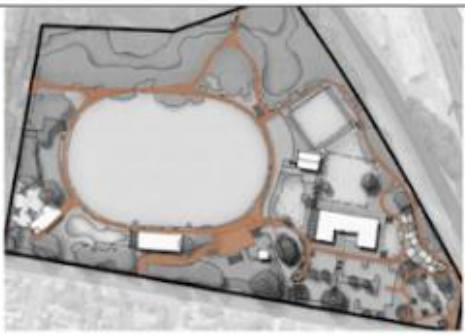
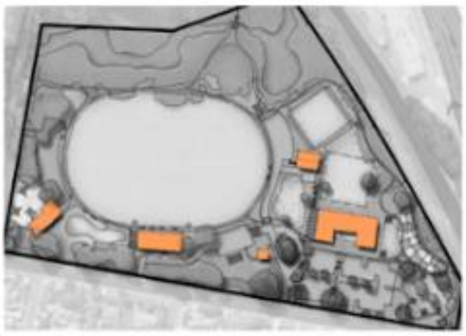
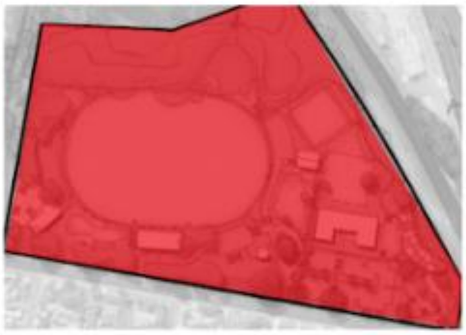
Long answer text

Description of factors

Description (optional)

Image t...

Factor	Description	Visualization
Horizontal Vegetation Density	This factor refers to how much of the park area is covered by vegetation. Only medium vegetation (e.g., shrubs) and high vegetation (e.g., trees) count. Grass, as the base surface cover of a park, does not count as vegetation in this project.	
Vertical Vegetation Density	This factor refers to the density (or thickness) of the available vegetation.	 Dense vegetation  Sparse vegetation

Proportion (amount) of water objects in the park	This factor refers to how much of the park area is covered by water objects (i.e., rivers, creeks, ponds, and lakes).	
Proportion (amount) of trails in the park	This factor refers to the proportion of trails and paths in the park territory.	
Proportion (amount) of human-made structures in the park	This factor refers to the proportion of human-made structures (e.g., houses, sculptures and other concrete objects) in the park territory. Concrete surface objects are detected by using Lidar.	
Size of the park		

Main image sources:

https://www.wangaratta.vic.gov.au/

When you think of your "ideal" urban park and/or other urban greenspace, which of the above mentioned six factors do you value the most?

Description (optional)

Please rank the six factors in order of importance, with #1 being the most important to you and #6 being the least important to you.

In this question, "important" means that you "value" a factor: i.e., an increase in a particular factor would increase the value of a park or other urban greenspace to you (or, simply put, more of that factor would improve a park for you).

Question *

	1	2	3	4	5	6
Horizontal V...	○	○	○	○	○	○
Vertical Veg...	○	○	○	○	○	○
Proportion (...	○	○	○	○	○	○
Proportion (...	○	○	○	○	○	○
Proportion (...	○	○	○	○	○	○
Size of the p...	○	○	○	○	○	○

Your comments or suggestions:

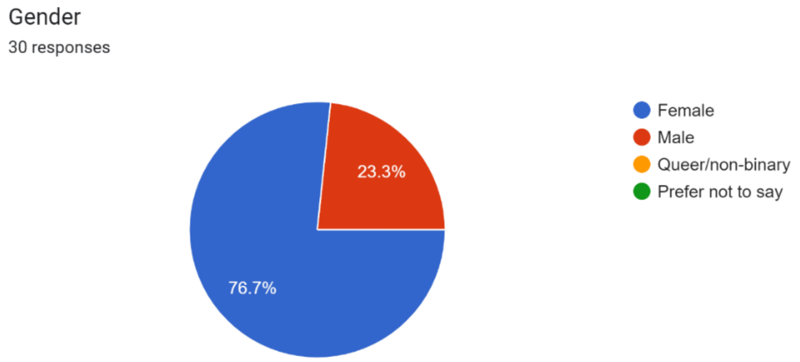
Long answer text

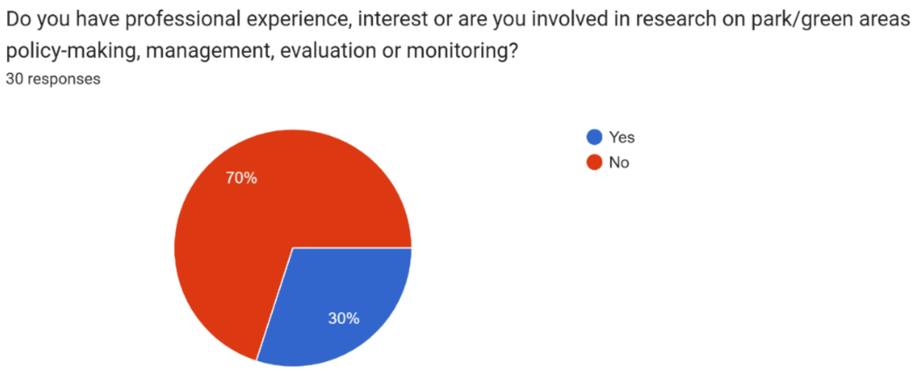
Thank you very much for your participation!

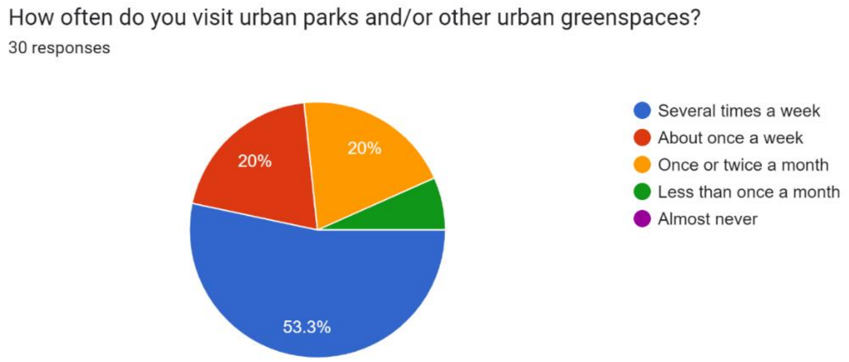
Description (optional)

Appendix 3

Results of the survey.

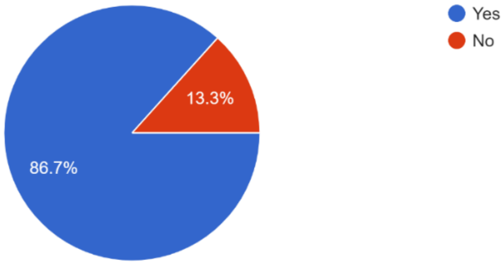






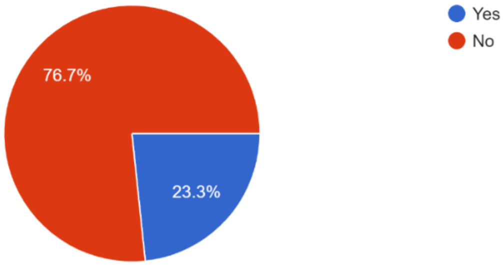
Would you like to visit urban parks and/or other urban greenspaces more often?

30 responses



Do you experience barriers to visiting urban parks and/or other urban greenspaces?

30 responses



Personal Declaration

Personal declaration: I hereby declare that the material contained in this thesis is my own original work. Any quotation or paraphrase in this thesis from the published or unpublished work of another individual or institution has been duly acknowledged. I have not submitted this thesis, or any part of it, previously to any institution for assessment purposes.

Signature: Dace Kirsteina