University of Zurich[UZH]

# What Makes an Urban Public Space Popular? A Data-Based Analysis of Existing Urban Public Spaces in the City of Zurich

GEO 511 Master's Thesis


Author
Sophie Sturzenegger
16-705-634


Supervised by
Prof. Dr. Ross Purves

Faculty representative
Prof. Dr. Ross Purves

30.04.2022
Department of Geography, University of Zurich

*“Our spatial behaviour, which is defined by and defines the spaces around us, is an integral part of our social existence.”*

Madanipour (1999, p. 879)

# Acknowledgements

First and foremost, I would like to thank *Prof. Dr. Ross S. Purves* for providing just the right amount of support I needed to make this master's thesis a reality. From brainstorming together to putting the brakes on my drive to push the data analysis infinitely further, your contributions were always very valuable to me.

By providing me with access to their mobile phone data, Swisscom AG enabled my vision to carry out this work based on a completely new type of dataset. Thereby, a special thanks goes to *Yann Steimer*, who made this possible and also made it immensely easy for me to get started working with the data.

A big thank you also goes to *Stéphanie*, who was not only a true friend but also a perfect proofreader and thus helped me a lot when I myself could not see the forest for the trees. I am equally grateful to *Olga*, who not only helped me to learn about the concept of introduction and literature review in a completely new way, but also inspired me with her fascination for science to give my best until the end.

Even if we don't see each other for a while because our lives get in the way, I know I can always count on *my family*. That is incredibly valuable and gives me the energy to keep working. Many thanks especially to *my Dad*, who was also heavily involved in the proofreading, even though the subject does not really belong to his area of expertise.

And last but not least, I would like to thank the people who were not directly involved in the work, but whose presence in my life contributed no less to its completion: *my friends*. Thank you, *Thomas*, for reliably distracting my mind from all the work that needed to be done during the occasional table football tournament or pub quiz, and then for supporting me when I needed to finish my thesis on time. Thank you, *Aline*, *Alina*, *Sarah*, *Ronja* and *Victoria*, for being there for me for so many years. Even though we haven't known each other quite that long, you are also incredibly important to me, *Seraina* and *Sabrina*. Thank you, *Martina*, for your empathy and unbeatable hugs that have always motivated me to keep going. And finally, thank you to *Anna*, who despite our geographical distance is always there for me and always manages to show me life from new perspectives.

# Abstract

In the context of global urbanization and changes in our living habits, urban public spaces (UPSs) such as parks and squares are increasingly gaining in importance. They must therefore be designed to meet the needs of a wide variety of people. To better understand what actually makes UPSs attractive to the population and leads to their active use, correlations between design and popularity of existing UPSs can be explored. Various approaches have already been used for this purpose, but most of them were only able to describe these relationships in a temporally and spatially limited manner, since all information had to be recorded manually. The purely data-based approach used in this work, on the other hand, is easily scalable in both time and space. Using visitor densities derived from mobile phone data, the popularity of various UPSs in the city of Zurich is estimated. The resulting popularity is then compared to physical attributes of the UPSs that could make them more attractive to the population. The findings indicate that especially the number of shops and accessibility on foot within a neighborhood seem to have an influence on how popular a UPS is. Thus, data-based approaches, indeed have the potential to help urban planners plan in a more targeted, efficient, and population-oriented manner. However, it is also noted that there are some components and reasons for the attraction of UPSs that cannot be captured by data without additional semantic content, field observations, or surveys. In combination with various other research approaches, data-based analyses have the potential to provide valuable new insights into the relationship between the popularity of UPSs and their characteristics.

# Table of Contents

# Chapter 1 | Introduction

"Third Place Living" is what the Zukunftsinstitut calls the form of living we are increasingly adopting, referring to the shift of personal living space into semi-public or public spaces (Zukunftsinstitut, 2013, as cited in Frank, 2019). Cafés, parks or squares become temporary offices, just like public transport. The boundaries between the traditionally separate fields of work, leisure and mobility are increasingly dissolving and the spheres are converging both socially and spatially. Living in this flexible, temporary and de-centralized network of private, semi-public and public spaces and places, the residential environment assumes great relevance (Frank, 2019). Urban public space (UPS) suddenly has to fulfill even more functions for a wide variety of people and tasks. Accordingly, the interest in the design of public space has increased among the population, but also in the business community and in the various departments of the city administration. Moreover, in national and worldwide urban competition, the design quality of public urban space is rapidly becoming a key success criterion (WIDEL, 2006). In addition to changing lifestyles and needs in cities, the ongoing global urbanization is also bringing the relevance of public spaces and green areas increasingly into focus worldwide. Along with environmental sustainability, the issue of social sustainability is also becoming increasingly important in urban planning (Dempsey, et al., 2011). As a sub-goal of Sustainable Development Goal 11 to make cities and human settlements inclusive, safe, resilient, and sustainable, the United Nations (UN) defines ensuring universal access to green spaces and public spaces. A particular focus of this goal, to be achieved by 2030, is the usability of UPSs for women, children, the elderly, and people with disabilities (DESA, U.N., 2016). In order to meet the growing demands on public space and to plan in a target group-oriented manner, it is becoming increasingly imperative for urban planners to understand how UPSs should be designed in order to be attractive to different groups from the population.

One way to gain a conceptual understanding of the often complex relationships between the built environment and population behavior that has been receiving increasing attention in urban planning is through data analysis (e.g., Ivanov & Gnevanov, 2018). This is especially true as society has become highly digitized in recent decades and many cities have begun to pursue Smart City strategies. A key feature of this strategy is the strong promotion of technological infrastructure to drive the development of human and social capital in cities (Angelidou, 2016). In the process, different types of data are collected almost constantly by a wide variety of devices. While real-time data can provide valuable information for other fields, historical data is particularly pertinent to the field of urban planning (Rathore, et al., 2016). Analyzing past circumstances can increase understanding of the processes involved and provide insight into how they might develop in the future. In this way, planning can be done in a timely, sustainable and process-oriented manner.

## 1.1 Motivation and Goals

Data-based approaches promise a fundamentally new understanding of the complex processes and dynamics in cities (e.g., Thakuriah, 2017; Ivanov & Gnevanov, 2018). For an efficient and target group-oriented planning of the built environment in cities, it can therefore be very valuable to integrate such analyses into traditional urban planning processes. Especially places accessible to the public, such as squares and parks, must be very carefully thought out in their design and adapted to the requirements of the users. It is likely that UPSs that meet many of these requirements will ultimately be better received by the population and used more frequently.

The aim of this study is therefore to explore the opportunities and challenges of purely data-based analyses of existing UPSs. On the one hand, the aim is to find out whether there are certain characteristics that make existing UPSs more popular among the population. On the other hand, it will be investigated to what extent the approach of pure data analysis is really suitable to describe complex relationships between the design of places and the way this place is received by the population.

## 1.2 Structure

*Chapter 2* examines the background to the planning and design of urban public spaces. It discusses how the topic is debated in the literature and in policy, the particular importance of UPSs in urban planning, and design approaches. Thereby, the importance of UPSs in Switzerland is briefly examined in more detail. Approaches to evaluating existing UPSs by various researchers are then presented, followed by a discussion of the implications and opportunities of data-driven and GIS-based urban planning approaches. Methods for modelling the popularity of locations and the relationship with location factors are then examined in more detail, and research gaps are identified at the end. *Chapter 3* describes the datasets used for data analysis and how they were preprocessed. In addition, the part also includes an initial data exploration that was used to understand and validate the data. Finally, this chapter also demonstrates how correlations between attribute data and the number of visitors to UPSs were examined. All the results of the data exploration and the main analysis are then described in *Chapter 4*. The implications of these findings are subsequently reviewed in *Chapter 5*. This is also where the link back to the literature and the research questions is made. Furthermore, considerations are given to the methodology employed and the data used. The final *Chapter 6* concludes this thesis by reviewing the findings and making suggestions for future work.

# Chapter 2 | Background

## 2.1 Sustainable Urban Development

According to United Nations (UN) figures, since 2007, for the first time in human history, more people worldwide live in urban areas than in rural areas (Ritchie & Roser, 2018). By 2020, this was already over 56% of the world's population and by 2030 it is expected to exceed 60% (UN, 2018). Depending on the definition of an urban region, these figures are even higher. In its 2016 report, the EU claimed as much as 85% of the world's population were living in urban clusters (Pesaresi, et al., 2016). Regardless of the sources and definitions, however, one thing is evident: urbanization is in full swing worldwide. More and more people are moving to urban areas. This rapid social change poses a whole new set of challenges that cities must now address. This applies to ecological as well as social and economic aspects of daily life (Shahidehpour, et al., 2018). The necessity of sustainable development of cities has therefore become increasingly recognized and has been included in the agendas of many different institutions, organizations, and governments (Ahmadi & Toghyani, 2011). In 2015, the United Nations officially dedicated one of its 17 Sustainable Development Goals to this issue. With goal number 11, they are seeking to make cities and human settlements inclusive, safe, resilient and sustainable (DESA, U.N., 2016). However, there is hardly any general consensus on what sustainable development really is, and the term is often interpreted differently by different people depending on their backgrounds (Redclift, 1991). According to Dempsey et al. (2011), the discourse was quasi-exclusively concerned with environmental sustainability for an extended period of time. Only later did other dimensions of sustainability come more into focus. In addition to environmental sustainability, the aim was now to also achieve ecological as well as social sustainability (Davidson, 2010). While all of these factors must be taken into account to make a city as a whole truly sustainable, this work has a particular focus on the social dimension.

### 2.1.1 Social Sustainability

Social sustainability, like many terms related to sustainability, has no universally agreed upon definition. Dempsey et al. (2011) argue that the term should be viewed more as a dynamic concept that changes over time and in place depending on external factors such as general changes in social cohesion and interaction, the economy, the environment, and politics. Similarly, Eizenberg and Jabareen (2017) also see social sustainability as the integration of different social, ecological and economic aspects. Thereby, the concept of social sustainability is composed of the interrelated concepts of sustainable urban forms, equity, safety and eco-prosumption. A direct link is thus established here between urban forms, i.e. the built environment, and the achievement of social sustainability goals. It is this very connection that makes the general topic of social sustainability so highly relevant to the field of urban planning. However, social sustainability has only gradually become part of the general sustainability discourse in the context of urban planning, and thus there has not been done much research on it for a long time (Dempsey, et al., 2011). In policy debates, on the other hand, the issue has rapidly gained prominence in recent decades, with repeated attempts to define the term more clearly so as to gain a working understanding of it (Davidson, 2010). In the 2005 'Bristol Accord', EU ministers met informally to discuss and agree on the benefits of creating sustainable communities across Europe (ODPM, 2005). Although the concept of sustainable communities differs slightly from that of social sustainability, the two fields share a very large common sub-scope and mostly address very similar issues. For sustainable communities, the following definition was established in the framework of the 'Bristol Agreement':

*"Sustainable communities are places where people want to live and work, now and in the future. They meet the diverse needs of existing and future residents, are sensitive to their environment, and contribute to a high quality of life. They are safe and inclusive, well planned, built and run, and offer equality of opportunity and good services for all." – ODPM* (2005, p. 6)

Once again, the fundamental connection between sustainable communities and the physical urban context in which they exist can be recognized. The importance of good and sustainable urban planning and design is also highlighted by the EU itself, defining it as one of the eight main characteristics of a sustainable community (ODPM, 2005). But in what ways can an urban built context affect social coexistence in different neighborhoods? How must a neighborhood be designed to be 'socially sustainable'? And who decides to what extent this goal has been achieved in practice? Woodcraft (2012) emphasizes the importance of understanding how the concept of social sustainability is interpreted by different actors, as the understanding of social sustainability is still ultimately the result of what a "good life" and a "good city" means to urban populations from different social and ethnic backgrounds. Thus, a discourse between the various stakeholders and population groups regarding the different understanding is essential. Only when actual needs are identified can a planning framework be established that truly promotes social sustainability.
There are already individual urban planning initiatives by various cities, states or academics whose concepts are consciously or unconsciously strongly oriented toward ideas typically associated with social sustainability. One example is the liveable neighborhoods framework.

#### 2.1.1.1 Liveable Neighborhood Frameworks

Even though Wheeler's (2001) working paper on liveable communities in California was published four years before the 'Bristol Accord' and thus at a time when the concept of social sustainability was only slowly gaining importance, it already echoed many of the concepts later associated with social sustainability. Elements that contribute to a livable neighborhood are collected in the following list (Wheeler, 2001, p. 11):

- an attractive, pedestrian-oriented public realm
- low traffic speed, volume, and congestion
- decent, affordable, well-located housing
- convenient schools, shops, and services
- accessible parks and open space
- a clean natural environment
- diverse, legible, and educative built landscapes
- places that feel safe and accepting to all users
- places that emphasize local culture, history, and ecology
- environments that nurture human community and interaction

In the Western Australia region, the concept of the liveable neighborhood has not only been explored theoretically, but it represents an integral part of their planning strategy to this day. It is a dynamic framework that has been in use since 1998 and is constantly evolving (WAPC , 2009). Similar to Wheeler (2001), they define livable neighborhoods as ones that are safe, sustainable, attractive, and have a strong, site-specific identity that supports the local community. Although this framework describes in much greater detail how exactly to design a liveable neighborhood than does Wheeler's work, there is a great deal of overlap in the basic requirements and goals for the design and layout of new development.

Also Montgomery (1998) was an early adopter of the idea of a sustainable city, including a special focus on social sustainability. He argued that it was essential to achieve a sense of urbanity through activity in the newly created neighborhoods, otherwise they risked becoming increasingly lifeless, boring, and inert, i.e., increasingly suburban. To accommodate these activities, he proposes what he calls 'transaction bases', which provide space for social interactions such as cultural events, festivals, or everyday street life. To further enhance the vibrancy of a neighborhood, Maller (1999) advocates his urban design approach of structured accidentalness, based on a congested urban design that allows and even encourages chance encounters and events.

Indeed, it is the emphasis on promoting social interaction that is common to many frameworks and research related to social sustainability and urban planning, despite sometimes very different backgrounds and approaches. Public spaces, as settings that are open to all people, quite naturally represent a built measure that can provide a room for such interactions. Therefore, they play a major role in sustainable urban planning. Additionally, Emenike (2016) argues, that different types of public open spaces have a great impact on our general life, including health, well-being and physical development. This is in line with proposals for sustainable communities developed in the context of the 'Bristol Accord'. Thereafter, public spaces and green areas that promote health and prevent crime, thus giving people a sense of security, should be designed, and integrated into the urban environment (ODPM, 2005). Similarly, universal access to safe, inclusive, and accessible green and public spaces has also been identified by the UN as an official sub-goal of Sustainable Development Goal 11 (DESA, U.N., 2016).

## 2.2 Urban Public Spaces

Following a classification by Johnson and Glover (2013), four different categories of public space can be described based on ownership and accessibility (see Table 1). The first two categories are spaces that are privately owned. The private spaces can either be designed in such a way that it is easy to deny people access, as is the case in cafés or restaurants, for example. In other privately owned spaces, such as an easement used to piece together a trail system, it is difficult to deny people access. Categories three and four, on the other hand, are publicly owned spaces. Again, a distinction is made between spaces where it is easy to deny access, such as the municipal ice arena, and those where this is not the case (Johnson & Glover, 2013). The latter category includes public spaces that are freely accessible to everyone, such as communal open spaces like parks or squares.

**Table 1**: Categories of Urban Space after Johnson & Glover (2013)

| | Easy to deny access | Difficult to deny access |
|---|---|---|
| **Private ownership** | Private-public space (e.g., a coffee shop) | Common space (e.g., an easement) |
| **Public ownership** | Club space (e.g., a municipal ice arena) | Outwardly public space (e.g., public park) |

Besides squares and parks, there are of course additionally many other types of public spaces in cities that are freely accessible to all. Playgrounds, streets, or sidewalks are just some of the examples that can typically also be placed in this category (Emenike, 2016). However, especially in the context of developing socially and environmentally sustainable neighborhoods, public parks and squares have been repeatedly identified as critical in a variety of planning frameworks (e.g., Wheeler, 2001; ODPM, 2005; WAPC, 2009; DESA, U.N., 2016). For this reason, this work will focus exclusively on these two types of public spaces. The term 'urban public spaces' (UPSs) is used as a generic term for the totality of all squares and parks located in a city.

### 2.2.1 The Importance of UPSs for Cities

*"…without a transaction base cities and urban places become progressively more lifeless, dull and inert—that is to say more suburban. Without activity, there can be no urbanity."* – Montgomery (1998, p. 97)

Various scholars have highlighted the benefits of attractive UPSs for cities. Montgomery (1998) and Maller (1999) argue from a quality-of-life perspective. Activity on the street through pedestrian flows, events, general street life leads to more social interactions and thus to more urbanity. Urbanity is thereby strongly associated with the liveability of a place. Furthermore, Cattell et al. (2008) found that social interactions in public places provide relief from the daily routine, reinforce a sense of community, provide opportunities to maintain bonds or build bridges, and can have a direct impact on well-being by lifting people's spirits. The authors believe that public spaces are more than just places where people congregate; they have subjective meanings that develop over time and may help people meet their needs for safety, identity, and a sense of place (Cattell, et al., 2008). Furthermore, UPSs can also meet important socio-political needs of cities by challenging the ongoing process of socio-spatial fragmentation and promoting a concept of togetherness instead. (Madanipour, 1999). This is also why Madanipour (1999) argues that UPSs can even be seen as a potential cure for crime problems, as they provide a place of sociability and entertainment for disillusioned youth, who are often blamed for many urban ills.

Beyond the pure quality-of-life argument of the benefits of well-designed UPSs for a city, economic arguments also come into play. They also have an essential role to play in the global competition between cities for international tourists. They also play an integral role in the global competition between cities for international tourists, as they represent the places that shape a city's identity (Sharipov & Demirkol, 2018).

For UPSs to fulfill all of the functions described, they thus must be designed to be attractive to people from diverse populations, not only within the city but also nationally and internationally.

### 2.2.2 Designing Public Spaces

Many frameworks that emphasize the importance of urban public spaces do not reference clear design guidelines for them. In the 'Bristol Agreement', the EU explains for which demographic groups the park or square must function. Although the target group is clearly defined as the entire population, the focus here lies on children and the elderly (ODPM, 2005). To these two groups of the population, the UN adds in its SDGs those of women and people with disabilities, for whom such spaces should be optimized (DESA, U.N., 2016). The question, however, is how this can be implemented in terms of design and what characteristics a UPS must have in order to be suitable and attractive for the target groups.

The Project for Public Spaces (PPS), an interdisciplinary nonprofit organization that works to improve public spaces, has identified one of the biggest problems in designing public spaces is who is included in the decision what is needed (Metropolitan Planning Council, 2008). Top-down approaches are often used, where planners and officials did not consult the people who live in the neighborhood about issues they faced directly, but simply gathered feedback on the proposed designs. It is frequently attempted to equip new spaces with certain physical traits in order to attract specific target groups (Hjort, et al., 2018). However, this approach is not always effective, as it is very complex to predict which design features will attract which types of people in the long term. For this reason, the PPS also calls for traditional planning processes to be reconsidered and adapted to the actual needs of the population. They propose a bottom-up process they call Placemaking, defined as a process that "turns a public space into a living space" that should be more economical, efficient, and enjoyable for all stakeholders and help identify and meet real needs (Metropolitan Planning Council, 2008). The concepts of Placemaking echo many of the ideologies of existing sustainability and livability frameworks. Cilliers et al. (2015), who adopted the ideas of Placemaking for their approach to lively planning, argue that the concept of the 'power of 10' provides a good basis for planning the functions of a public place offered to the community. Therein, an argument is made that a great place must have at least ten things to do there, or ten reasons to be there. These might include, for example, a place to sit, art to touch, music to listen to, food to buy, historical information to learn, and books to read (Cilliers, et al., 2015). The PPS (2008) additionally maintain that 80% of the success of a public space has to do with its management. Only a public space that flexibly changes with the times and the population has a chance to remain popular in the long run (Metropolitan Planning Council, 2008).

### 2.2.3 Social Sustainability and Urban Public Spaces in Switzerland

Sustainability issues have also increasingly marked the urban planning and development discourse in Switzerland in recent years. Switzerland's official 2030 Sustainable Development Strategy, published in 2021 in conjunction with the United Nations Sustainable Development Goals, defines three priority themes. These are 'sustainable consumption and production', 'climate, energy and biodiversity', and 'equal opportunities and social cohesion' (Schweizerischer Bundesrat, 2021). While the Sustainable Development Strategy does describe some of the sub-goals of SDG 11 (sustainable cities and communities) and formulates them as goals for the entire country, it does not explicitly address the subgoal 11.7 of universal access to green spaces and public spaces.

In contrast, social sustainability and public urban spaces, which are important in this context, are mentioned in various private and public planning frameworks. Thus, a framework concept for social sustainability in settlement development commissioned by the Swiss Society for National Planning assigns special importance to the so-called network nodes in the path network of new settlements and residential yards (Drilling & Weiss, 2012). Furthermore, Swiss cities strongly emphasize the importance of public space for their populations. In 2006, the city of Zurich adopted the 'Strategy for Urban Spaces', in which the city declares the design principles according to which it intends to shape public urban space. In the preface, the then city councilor Martin Waser argues that after having become almost exclusively a place of movement, the importance of public space has recently increased greatly, and it is being reclaimed as a place to linger and live (WIDEL, 2006). UPSs are increasingly being used once again as meeting places, for recreation and a variety of activities. Due to the importance that the city of Zurich attaches to its UPSs, the strategy according to which it intends to develop them further is also very detailed. As a basis for the formulation of the strategy, the existing public spaces were subjected to a detailed SWOT (Strengths & Weaknesses, Opportunities & Threats) analysis. As one of the threats, the increasing pressure of use is particularly emphasized (WIDEL, 2006). Different user groups of a public space want to realize their own demands to the maximum, which could lead to conflicts given the limited space available. Thus, a good compromise must be found here to make UPSs work for diverse audiences. In general, the city of Zurich states that attractive urban spaces depend especially on a good quality of stay. They have presented a table showing which factors belong to this (see Table 2). In addition, the city of Zurich also defines very specific infrastructure elements, such as litter garbage cans or benches, that should generally be used in all public spaces (WIDEL, 2006).

**Table 2**: Checklist quality of stay for public space, as defined by the city of Zurich (WIDEL, 2006, p. 19; translated)

| Safety | Comfort | | Sensual Qualities |
|---|---|---|---|
| **Road Safety**<br>• Protection from accidents<br>• Protection from noise, pollution, emissions<br>• Clarity | **Walking**<br>• Sufficient space<br>• Attractive network<br>• Interesting facades<br>• Good surfaces<br>• Good accessibility for all<br>• No obstacles<br>• No path interruptions | **Stay**<br>• Zones for stay<br>• Opportunities to sit, rest, lean, look, be seen, enjoy<br>• Good local climate<br>• Inviting space edges and facades | **Climate**<br>• Protection against wind, rain, snow, heat and cold<br>• Allow sun<br>• Provide shadow<br>• Use heat and breeze to the extent pleasant |
| **Perception of Safety**<br>• Lively, used<br>• Social control present<br>• Uses that overlap in space and time | **Seeing, hearing, speaking**<br>• Comfortable walking distances<br>• Clear view, panoramic views<br>• Good lighting<br>• Low noise level<br>• Communicative arrangement of seats | **Activities**<br>• Inviting for sports, games, entertainment by day and night, summer and winter | **Aesthetic qualities**<br>• Good design<br>• Good materials<br>• Good lighting quality<br>• Views, sights<br>• Vegetation, Water<br>• Cleanliness<br>• Human scale |

### 2.2.4 Evaluating the Success of Urban Public Spaces

Many questions about the details that make up a good design of a public space are difficult to answer in purely theoretical terms due to the complexity of the subject matter. To better predict what design characteristics a UPS must have to be well received by the public, different researchers have attempted to evaluate existing public spaces.

In order to more accurately predict which factors can contribute to the success of public spaces, the PPS analyzed thousands of spaces around the world and used the findings to create what they call the Place Diagram (see Figure 1). This can help directly link the often qualitative, intangible characteristics used to describe spaces to quantitative, tangible measures, making it easier to evaluate and design public spaces (Metropolitan Planning Council, 2008). While a great many different measures are very clearly shown here, no quantities or weighting were suggested by the PPS that might be useful in classifying UPSs as well or poorly designed. To do this, Mehta (2014) weighted each of the tangible variables he used in his work and assigned a score between 0 (low) and 3 (high) for the presence of a given feature depending on the quantity in which it is found in a UPS. With such tools, a simple quantitative comparison can be made between different UPSs. This could then provide guidance on how UPSs need to be designed to ensure that they each work well with the population.

One drawback of the data-driven method Metha (2014) proposes is that it can be complex and laborious to apply, as it relies heavily on data that can only be collected through field observation, surveys, and interviews in public spaces. While this was indeed the case for a long time and much data still had to be collected manually, a new strategic dynamic has developed in major cities in recent years: the ideal of the Smart City is increasingly being targeted. One of the main goals of a Smart City is to increase efficiency and quality of life. To achieve this, a data-driven approach is often taken, meaning that cities collect and record enormous amounts of data about the urban environment (Ma, et al., 2018). In addition, all other areas of our lives are gradually being digitized (Ivanov & Gnevanov, 2018). This includes the use of various devices designed to make processes more efficient and improve our quality of life. Therefore, many private companies have also started to systematically collect data. The extensive databases created by this process promise exciting new opportunities for the field of urban planning.

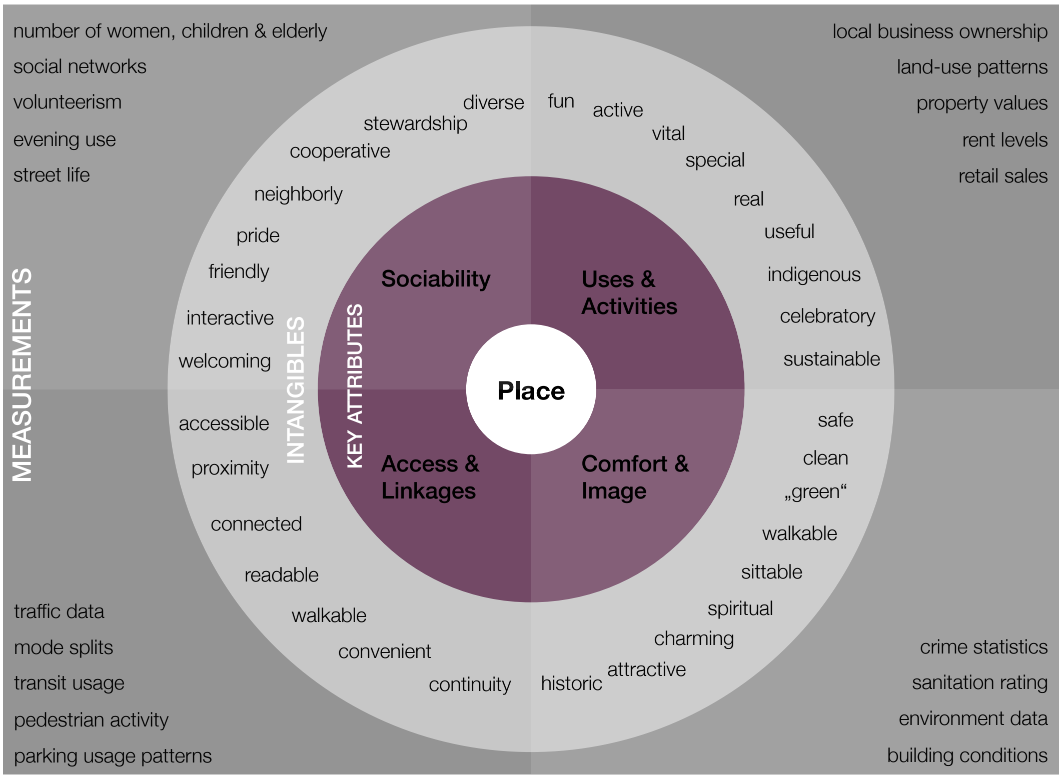



**Figure 1**: Place Diagram for successful public spaces, adapted from the PPS (Metropolitan Planning Council, 2008, p. 16)

## 2.3 Data-driven Urban Planning

Numerous scholars have advocated the increasing use of data-driven approaches in modern urban planning (e.g., Ivanov & Gnevanov, 2018; Rathore et al., 2016; Takuriah et al., 2015). Ivanov and Gnevanov (2018) even claimed that modern urban planners cannot exist without an understanding of Big Data and how to use it. This is because so many devices are in permanent use today, such as street cameras for surveillance systems, but also sensors for measuring air quality or GPS data from personal mobile devices that continuously produce data (Rathore, et al., 2016). The sum of all these electronic and Internet-connected devices can be summarized under the term Internet of Things (IoT) and the data collected through their use is usually very varied and exists in both static and dynamic forms. Ivanov and Gnevanov (2018) referred to them as “digital traces” that we constantly leave behind.

According to them, these traces, when appropriately analyzed, have the potential to provide a kind of “bird's eye view” of cities and thus contain an understanding of how life works. The research field of urban informatics deals precisely with these issues.

*“Urban informatics takes a transdisciplinary approach to understanding the city as an ecology that consists of technological, social, and architectural layers.”* - (Foth, et al., 2011, p. 4)

The processing of information, particularly via network technologies that incorporate a range of urban elements, is the focus of urban informatics (Foth, et al., 2011). As urban informatics seeks to gain a holistic view of the city and its dynamics, urban planning is also one of its topics (Thakuriah, et al., 2017). The possibilities of information and communication infrastructure and technology offer completely new opportunities to perfectly adapt urban systems to the life of the urban population and their needs (Ivanov & Gnevanov, 2018). While it is crucial for many typical Smart City applications that data is collected and analyzed in real time so that changes can be responded to as quickly as possible, historical data is of particular importance for the field of urban planning (Rathore, et al., 2016). On their basis, future resource planning can be carried out. Optimal locations for large infrastructure projects such as district heating can be determined by analyzing the current energy consumption of all buildings (Ali, et al., 2020). In addition to resource planning, the literature additionally provides quite a few other examples of how data analytics can support modern urban planning. Environmental sustainability is often a strong focus in the construction of entire new urban districts. To achieve this goal, for example, a combination of sustainable energy solutions and data-driven solutions can be used (Bibri, 2020). In addition, using an approach that employs anonymized call data records (CDR), Becker et al. (2011) have shown that typical challenges cities face in the wake of urbanization, such as highly congested roads, excessive development, and increasing pollution, can be solved by better understanding the day-to-day dynamics of a city.

### 2.3.1 GIS in Urban Planning

Becker et al. (2011) are not the only ones who have used location information and geographic analysis in the context of urban planning. Since the advent of geographic information systems, the two fields have become increasingly intertwined. In 1999, Maller expressed great optimism about this increasing fusion in his paper advocating the integration of structured accidentalness into urban design concepts. He argued that especially with the increasing power of GIS, the integration of a large number of layers and information would greatly facilitate the implementation of design concepts such as the structured accidentalness and congestion in future planning processes. Yeh (1999) also pointed out early on the increasing importance of GIS for planning support systems. He argued that the biggest constraints at that time were not technical problems anymore, but rather data availability, organizational changes, and staffing. However, with increasing digitalization, rising computing power and the introduction of IoT and Smart City concepts, geographical data in particular is being produced in large quantities and constantly. Thus, these constraints have greatly diminished over time, and today GIS is almost automatically part of many urban planning systems.

GIS approaches are often used to create, compare and optimize different planning scenarios. This can help better understand real planning problems and enhance quality and governance (Yaakup, et al., 2005). For example, Yeh und Li (2003) modeled different types of future urban growth for the Chinese city of Dongguan based on different planning goals.
To define which specifications and scenarios could actually work for a city, it is very important to determine the actual needs of the population. To achieve this, participatory approaches can be applied directly, or data can be collected and analyzed to provide insight into the behavior and opinions of the population. For instance, Kahila and Kyttä (2009) proposed the softGIS approach, which builds the bridge between participative GIS and planning support systems and thus enables place-based experiential knowledge of residents to be incorporated into planning decisions. But it is not only participatory approaches that are able to reveal information about the population and their needs in a particular location in the city. By combining demographic data and locations of everyday infrastructure, it is also possible, to make statements about where certain services need to be expanded (Carpio-Pinedo, et al., 2019). This makes it possible, for example, to predict where more schools, daycare centers, or hospitals need to be opened.
Demographic data based on information about residence can therefore provide valuable insights for building system-relevant infrastructures that are needed by everyone. However, when it comes to finding out more about where people choose to stay when they are on the move, other approaches and sources of information must be used.

## 2.4 Modelling the Popularity of Locations

In order to evaluate specific locations such as UPSs, as well as restaurants, stores, or other publicly accessible locations, the attributes of these locations are often compared to certain measures of their popularity among the general population. Private companies such as Starbucks or McDonald's, for example, expect higher profits by choosing locations with greater popularity (Hsieh, et al., 2019). Thereby, the exact or derived number of visitors to a place is often used as an approximation of its popularity (e.g., Hsieh et al., 2019, Tammet, et al., 2013). In turn, to be able to estimate visitor numbers, the locations and movements of people over time must be known. And while it is widely expected that the ability to predict human movement patterns will bring great benefits to our society (Soper, 2012), to date it has been difficult to get a comprehensive picture of the mobility of the population and the places where they spend time. Nevertheless, this has been attempted again and again by different scientists using different methods. Thereby, some have conducted their own observations manually and selectively, while others, have referred to different types of datasets generated using now widely available modern technologies. In the latter case, a general distinction can be made between user-generated content from social media platforms, which can be classified as a special form of voluntary geographic information, and data collected in the background, usually without our active knowledge, from mobile devices.
This chapter will elaborate on these various methods of determining human mobility focusing on the assessment of the popularity of urban public spaces and will highlight what has already been done in this area.

### 2.4.1 Manual Data Collection

In her study of different urban open spaces in Ljubljana and Edinburgh, Marušić (2011) wanted to find out how physical characteristics of the spaces affect the way they are used by the population. To do this, she decided to use a combination of behavior mapping and GIS techniques. In both cities, different UPSs were observed during four predefined time periods within a diurnal cycle on different days in the month of May and the activities performed were plotted as points on maps. Each 10-minute observation was made from a location that provided a good overview of the whole area, or in some cases the area was divided into sub-areas (Marušić, 2011). Rasidi et al. (2012), in their study on the design of urban green spaces in Malaysia and how it influences social interactions, have taken a very similar approach. They chose to conduct manual observations of park users in addition to a survey. In order to analyze certain usage patterns in more detail, they conducted systemic walks during four time periods of a day on different days during the week. Like Marušić (2011), they repeatedly took 10-minute breaks to visually scan a specific location and plotted all current users as a point on a map. Observations were conducted for three consecutive days of weekday, weekend, and public holiday (Rasidi, et al., 2012).

Ghavampour et al. (2017) also chose manual observations for their analysis of behavior in small public spaces in Wellington, New Zealand. Observations were made in all locations every 12 minutes on a sunny mid-week day. Unlike Marušić (2011) and Raisidi et al. (2012), however, they chose not to record visitors and usage directly on a map. Instead, they took regular photos of different sub-areas of a space and only later undertook coding for analysis. In doing so, they excluded individuals who were merely transiting from the analysis. The other individuals were each mapped as a single point per group of people, including information on group size, age, gender, and activity (Ghavampour, et al., 2017).

As can be seen from the example of these three studies, manual observation of people in UPSs brings some advantages. On the one hand, everyone has the possibility to get information about the number of visitors and activities in different UPSs. In addition, a special focus can be placed on obtaining exactly the information that is relevant for your own study. Additionally, direct attention can then be paid to ensuring that the data is as accurate as possible. On the other hand, there are very few major ethical or privacy concerns with the manual version, since the observed individuals do not involuntarily reveal any information about themselves other than their presence at a location at a given time.

However, it should also be noted that manual recording of visitors at a site can have significant drawbacks. Such observations are very time-consuming, and therefore it is very difficult to monitor sites at larger temporal and spatial scales. To address this problem, many researchers have based their analyses on existing datasets that provide information about where people spend time.

### 2.4.2 Automated Data Collection

Modern technology, which we are using more and more, continuously produces huge information databases. Whether we're sharing our experiences with other people on social media, shopping at the supermarket, or simply using our smartphone on the go, all of this information is stored by manufacturers and in some cases resold or shared for free, usually in anonymized form. In all three of the life situations described in which we probably all regularly find ourselves, data is generated that can provide information about our location. Precisely such datasets have also been used by some researchers to monitor the locations of populations over larger spaces and longer periods of time than would be possible with manual observations. While these are clear advantages of automatically collected data, there are also some disadvantages that should be considered, and perhaps could lead to manual observations being the better solution in some cases after all. For instance, such datasets are often difficult to obtain because of data protection issues, and their use still raises many ethical questions (Taylor, 2016). Nevertheless, in times of digitalization and the Smart City, they represent an increasingly important source of data.

The currently available automatically generated datasets can be subdivided, for example, according to the type of their origin. Thus, it can be further distinguished between active and deliberate sharing of geographic information, as is generally the case with Social Media, and passive sharing (Girardin, et al., 2008). In the latter, location data is shared without being explicitly initiated by the user, as happens, for instance, in the context of mobile communication.

Because these automatically generated datasets provide a consistent overview of larger spaces and sometimes even allow the tracking of individuals, they have been increasingly used in research, particularly in the context of mobility, activity, and movement.

#### 2.4.2.1 Active Sharing of Geographic Information

In the context of location information that is voluntarily shared by users, usually via social media platforms, the term volunteered geographical information (VGI) is often used (Goodchild, 2007). Location data from social media platforms are very popular and have been used in numerous studies. The advantages of this type of data are that they raise fewer ethical concerns than passively collected geographic data, since people share location information voluntarily, and that the data is often easily accessible (Dinkić, et al., 2016). They are also a relatively new way to gain insight into the activities and mobility patterns of visitors to specific locations (Heikinheimo, et al., 2017).

To find popular locations in different U.S. cities and to determine users' expected activities at these locations, Hasan et al. (2013a) analyzed hundreds of thousands of tweets (Twitter status messages) from tens of thousands of users. Dinkić et al. (2016) used a very similar approach to determine popularity in the Vračar region of Belgrade. In addition to the location information of the tweets, they also conducted a sentiment analysis of the content of the tweets.

The two studies demonstrate very well the benefits of working with datasets from social media platforms. On the one hand, one has immediate access to a huge amount of already existing data, including information voluntarily provided by the users. On the other hand, this data is

not only structured in nature, but includes additional unstructured user-generated content that can provide valuable additional information. This type of data, therefore, often has a lot of meaning, which Hasan et al. (2013a) cited as one of the biggest benefits of social media data for their study, as it allowed them to identify the activities of users in specific locations. In addition, such information can help to better understand how certain spaces are perceived, which is particularly relevant in the field of urban planning (Bahrehdar, et al., 2020).
However, geospatial data sourced from social media platforms and created by different users also have critical drawbacks. Content is often not created evenly across space and therefore different locations have limited comparability (Graham, 2011). Additionally, there are often biases in social media data because people do not post equally frequently on all activities and experiences (e.g., fewer work- and home-related tweets) (Hasan, et al., 2013a). Furthermore, there are also large differences in social media use between different age groups and ethnicities that must be taken into account so as not to greatly overestimate or underestimate any population groups. (Murthy, et al., 2016).
Thus, when working with social media datasets, there is always the question of how consistent the picture of the city and its residents that emerges from the analysis really is. However, it is very important, especially when designing UPSs that are intended to be attractive to as many population groups as possible, that the needs of all can be met. For the preliminary data analysis, it is therefore crucial that as little bias as possible toward individual population groups are included in the whereabouts data. One approach to circumventing this problem of heightened or diminished participation by individual groups would be to use data that is regularly collected in the background by widely used technologies without the user's active involvement.

#### 2.4.2.2 Passive Sharing of Geographic Information

Technologies that passively capture geographic information are manifold and are increasingly used for research purposes. Hasan et al. (2013b) analyzed smart card transactions (for public transportation) automatically registered when entering and exiting the subway to find popular locations in the city. They stated that they found the applied model to be a good and simple approximation of urban mobility and recommend its use also for simulations of disease spread in cities. It is in this area of disease spread that research related to human mobility has recently surged with the advent of COVID-19. However, in the context of this research area, another type of passive sharing of geographic information in particular has received much attention. Different research teams have emphasized the potential of analyzing mobile phone data. Although many uncertainties remain in working with such data (Grantz, et al., 2020), they should, for example, provide important insights for evaluating the effectiveness of various social distancing measures such as shutdowns or lockdowns (Oliver, et al., 2020). Several research teams thus have used this data and tested such analyses (e.g., Zhou et al., 2020 and Vinceti et al., 2020). Thereby, Vinceti et al. (2020) succeeded in establishing clear links between reduced human mobility due to exit restrictions in Italy and the sharp decrease in positive COVID-19 test results. Zhou et al. (2020) used mobile phone data to build a model for the Chinese city of Shenzhen that simulates different scenarios of mobility restriction and predicts how this will affect the control of COVID-19 spread. But mobile phone data are not only relevant in an epidemiological context. Such datasets have already been used several times in other studies, for example, as an indicator for socio-economic development (Pappalardo, et al., 2015) or criminal activity (Rummens, et al., 2021). There have also been a number of studies in the domain of urban planning that have made use of mobile communications data to better understand human mobility. Phithakkitnukoon et al. (2010) emphasize the importance of understanding human movement and characterizing human mobility patterns in particular to understand the inhabit dynamics of a city. Other research teams have also used similar approaches to better understand day-to-day dynamics in cities with the help of mobile phone data (Becker, et al., 2011). Steenbruggen et al. (2015) have studied the potential of mobile phone data and conclude that a better understanding of how, where, and when people move on a daily basis could lead to improvements in urban planning, especially in densely populated areas. However, they also point out that these data alone cannot fully represent the complexity of an urban context, and that they should be combined with other data for this purpose.

Mobile data is thus increasingly being used by various researchers, government agencies and private parties to gain new insights into population movements. However, there are still very many unresolved ethical issues with this type of data, as with all passively shared data that each user does not explicitly agree to share. As with many developments in IT, the availability of information about people's whereabouts has grown incredibly fast - faster than clear ethical frameworks or legal bases for the use of such information could be developed (Taylor, 2016). To protect the identity of individuals, many researchers are therefore calling for harmonized approaches to anonymizing mobile data sets. It is not enough to remove unique identifiers such as names, addresses, phone numbers or other specific identifiers. This only makes identification more difficult, but often cannot completely prevent it because of the possibilities of comparison with geotagged images and similar (de Montjoye, et al., 2014). Aggregation is one way to further anonymize the data, but it also has some potential for misinterpretation. Taylor (2016), brings up the issue of group privacy as a stretch of the current perception of privacy. In addition, she mentions the risk of miscomprehension of the data, by the people who typically process it (science community), as they do not necessarily have the same contextual perspective as others (social sciences).

Despite the aforementioned concerns about working with mobile phone data, this information, if used in an ethical manner, promises to provide entirely new insights. And although mobile phone data have been used for various applications and their usefulness for urban planning tasks has been repeatedly highlighted in the literature, there are not many studies that have actually used such datasets in the context of urban planning or specifically in determining the popularity of urban public spaces.

## 2.5 Modelling Suitability of Locations

In order to make a statement about the optimal design of a UPS, the visitor numbers of existing UPSs must be compared with their characteristics. If people were indeed to choose their stays in certain UPSs based on the existing infrastructure or the surrounding gastronomic or consumptive offer, the number of visitors to a UPS would correlate with certain attributes of its design. So far, there exist only a few approaches in literature on how to find correlations for exactly this use case, as not many researchers have investigated it in detail. Nguyen et al. (2019) focused thematically on a very similar case as this thesis. In their work, they utilized a check-in database from Instagram to examine the popularity of three different types of public spaces. For this, they used the geotags from Instagram posts to find out the visited venues of the users and then generated points of interest (POIs) and summed up the number of these POIs within predefined grid cells by type of public space. Finally, they used Pearson's correlation to compare the resulting grid value of POIs with the number of check-ins for those locations (Nguyen, et al., 2019). Here, Pearson's correlation is used to analyze correlations of three different types of public spaces. Searching for several individual attributes that make a UPS popular, however, requires statistical methods that can involve a much larger number of variables. While, to the best of my knowledge, such multi-criteria analyses have rarely been conducted in the field of urban planning to determine where people prefer to spend time, they are common practice, particularly in habitat suitability research. Standard regression analyses are often used to link the occurrence of a species to certain characteristics of the surrounding natural environment. An example of this approach is the study by Glenz et al. (2001). They fitted a stochastic model based on logistic regression to existing wolf population data in the northern Apennines. Using this model, they then examined habitat suitability for wolves in the canton of Valais to predict where wolves would be likely to reside and therefore where conflicts of interest might arise (Glenz, et al., 2001).

Especially with the rapid increase in computing power in recent decades, more complex algorithmic approaches, such as machine learning, have increasingly been used to calculate habitat suitability in some studies. In their case, Piri Sahragard et al. (2018) used a random forest algorithm to identify environmental factors affecting the distribution of plant species in southeastern Iran.

Although people clearly do not select the places where they spend their leisure time according to exactly the same criteria as animals or plants select their natural habitat, there is a strong overlap in the problem definition. When designing new UPSs, many academics and officials advocate equipping them with features that are attractive to the entire population or, in some cases, to very specific segments of the population. Thus, regression analyses of the type used for habitat suitability analyses could be useful in determining which of these features make one UPS more successful than another.

## 2.6 Research Gaps

The increasing urbanization of the world's population is causing cities to grow. In the process, ever more effort is being directed at ensuring that this growth is ecologically, economically and socially sustainable. The issue of social sustainability in particular is gaining in importance, even though many questions remain unanswered as to how exactly this goal can be achieved. Public spaces where people can spend time and socialize are now seen by many scholars and administrators as essential to achieving goals in the context of social sustainability. However, also in the design of urban public spaces, such as squares and parks, many questions remain unanswered to this day. UPSs are supposed to be built and designed in such a way that they work for a variety of people, but also for certain target groups from the population. We therefore need methods to find out which characteristics make a UPS attractive to which people.

One opportunity here, especially since the advent of Smart City ideologies and the increasing digitalization of our entire society, is an analysis of existing UPSs. These types of completely data-based analyses have rarely been used in research around UPSs until now. This is partly because for a very long time it was difficult to obtain reliable, non-manual data on visitor frequency at specific locations in the city. Today, however, it is possible to derive the dynamics of a city from automatically collected data, such as mobile phone data. Thereby, it is still unclear whether such data is at a level today where it can really provide relevant and reliable insights into the real situation for a problem in the context of urban design, such as the design of UPSs. Very generally, the fundamental question is whether purely data-based questions will be appropriate for answering questions that arise around human social interaction. There are currently few answers to these questions in the literature. With this work, I hope to contribute to their increasing clarification.

### 2.6.1 Research Questions

Based on the research gaps found, the research questions for this study are defined as follows:

**$RQ_1$: Are there specific attributes of an urban public space that increase its popularity?**

- What physical characteristics can be identified that show a correlation with an increased frequency of use?
- What can be done to make newly planned UPSs attractive for the whole community, but especially for women, children, and the elderly, and to invite them to spend time there?

**$RQ_2$: To what extent do the purely data-based methods used in this study allow us to capture and understand the relationships between popularity and characteristics in urban public space?**

- To what extent is the mobile phone data currently available in Switzerland suitable for analyzing number and type of visitors at specific locations in a city?

# Chapter 3 | Methods

## 3.1 Tools

In order to have access to a large number of readily available libraries while having the flexibility to perform various specific calculations, I decided to do almost all of the data work using Python. For this, I chose a virtual environment in a Juypter Notebook setup. For convenience, I also did some additional highly specialized calculations and integrations, as well as all map visualizations in the ArcGIS environment (Pro and Online).

## 3.2 Data

### 3.2.1 Geometries of Squares and Parks

Perhaps the most critical dataset for my analysis was one that included as many public squares and parks in the city of Zurich as possible. This is a dataset that does not yet exist or at least is not yet publicly available. Therefore, I decided to compile a dataset myself.
As a first step, I collected names of squares and parks in a list. In the case of the parks, I was able to draw on an existing list from Grünstadt Zürich (Stadt Zürich, 2021a). A compiled list of public places is not available anywhere. However, there is a list of all addresses in the city of Zurich (Stadt Zürich, 2021b). Since place names traditionally also include a designation for the type of location (e.g., Winterthurerstrasse - i.e., a location that is located on a street (in German: “Strasse”)), I took advantage of this and searched for all addresses that end with "Platz" (Eng. square) or "Hof" (Eng. courtyard). In addition, for validation and supplementation, I also conducted a minor internet search using the search term "Plätze Stadt Zürich". I then hand-coded the resulting squares and parks, saved the coordinates to the list, and created a point dataset from them. To get the boundaries of the UPSs, I intersected this point dataset with an existing polygon dataset of the city of Zurich, which visualizes the significance of different public spaces. This has the advantage that I did not have to define the boundaries for squares and parks myself, but could simply rely on the geometries that, according to the city of Zurich, represent the actual geographical extent of the individual UPSs.

### 3.2.2 Mobile Phone Data

A mobile phone dataset for Switzerland can be obtained from the webpage of the telecommunications company Swisscom AG (https://digital.swisscom.com/catalog). According to Swisscom, the data protection issues discussed in Section 2.4.2.2, which mobile data may exhibit in some cases, are not a concern here. All values are delivered aggregated in grid cells and are only given if at least 20 SIM cards are detected in the cell. In addition, measures are taken in advance to prevent individual SIM cards from being tracked over several days. To this end, the identification of the generated positions and journeys of a

SIM card changes on a daily basis, and the unit responsible for this and the additional anonymization is different and independent of the unit responsible for generating the mobile phone dataset provided to customers. The data is stored locally at Swisscom AG itself (Swisscom AG, 2022a). The protection of the identity of individuals should therefore be fulfilled.

For this work, Swisscom AG granted me access to their API, which can be used to retrieve data for the whole of Switzerland for the period between the respective previous day minus two years. They are thus delivered in near real time (Swisscom AG, 2022a). It is possible to obtain hourly values or cumulative values for the whole day. Since I was primarily interested in seeing if any general trends could be identified between the popularity of a square or park and its infrastructure, I decided to use only the daily values in this master's thesis.

In selecting the time period, I would use for my analysis, several factors had to be considered. On the one hand, the number of people outside in public UPSs is strongly related to seasonality. The population is much more likely to spend a large portion of their time outside during the warmer seasons. But weather also plays a large role within a season. When it rains, fewer people tend to choose to spend their leisure time in a park or a square. I therefore decided that I would use a period between late spring and early fall for my analysis.

But seasonality was not the only influence that could cause large fluctuations in visitor numbers during the two years in 2020 and 2021 for which I was able to pull data. The COVID-19 pandemic was very much present throughout most of this period. As a result, complete or partial shutdowns were repeatedly imposed by the Swiss government. During these events, there were usually fewer people out and about, for example because of the home office obligation. On the other hand, stores, restaurants, bars, and clubs also had to close several times. Since I wanted to find out whether the presence of such establishments in the immediate vicinity of UPSs contributed to their being more popular than others, it would therefore not have been appropriate to use data collected during the period of these measures. In the summers of both 2020 and 2021, however, there was a time when the COVID-19 case numbers were rather low and therefore life could take place almost without restrictions. For this reason, I decided to work with data of a three-month period from July 1 to September 30.

Swisscom AG provides different datasets that contain distinct parts of the total available information of the mobile phone data. The differences result from the different pre-processing methods of the diverse datasets carried out by Swisscom before provision. A distinction is made between dwell times data and heatmap data. The former describes a frequency distribution of observed dwell times for a given geographical area and day. Values can be retrieved for grid cells with a size of 500 m x 500 m and represents the duration of people's stay in the area, for specific time buckets (Swisscom AG, 2022b). The heatmap data, in contrast, have a resolution of 100 m x 100 m and provide insights on population density for a given geographical area and during a given time period. In addition, they also contain certain demographic information (Swisscom AG, 2022a). Since some UPSs in the city of Zurich tend to be smaller than 100 m x 100 m, it was very important for my purposes to have the highest possible spatial resolution for the dataset providing an approximation for the number of visitors as well. In addition, demographic information in particular is very valuable and interesting for the specific design of UPSs. I therefore decided to use the heatmap dataset in this work and not pay attention to the actual dwell time at specific locations in the city. Within the heatmaps datasets, a further distinction is made between different information that can be accessed individually. Here, I chose to use data on dwell density and dwell demographics.

#### 3.2.2.1 Dwell Density

The dwell density parameter is an estimate of the average density of individuals. Swisscom AG uses a specific formula for pre-processing and density calculation from the raw mobile data. On the one hand, a probability is included that a SIM card has been within a certain tile (Swisscom AG, 2022a). This statement is made based on the number of observations of a SIM card, whereby a larger number of observations leads to smaller uncertainties. In addition, the dwell time of the SIM card within a tile is also included. In order to be able to estimate not only the number of people in a locality who use Swisscom as a telecommunications provider, but the overall population, the respective market share of the individual municipalities is factored in at the end. From all these factors, a single dwell density value per hour is calculated for each tile on the whole territory of Switzerland (Swisscom AG, 2022a). Such a calculated dwell density value therefore says nothing about how many individual people have even entered a UPS within an hour. This information, which records all SIM cards that are in a tile for at least 15 seconds, could be read from the Unique Counts dataset of the Swisscom API (Swisscom AG, 2022a). The value of the dwell density says much more about the permanent number of visitors. For instance, the calculated dwell density for an example tile where two Swisscom customers (market share in home municipality = 100%) are located with 100% certainty if they both spend half an hour there would be 1. To calculate the daily values, Swisscom then simply added up the hourly values. This means that with a constant visitor density over all 24 hours, the daily value of the example tile would be 24. Thus, the dwell density reveals information about the constancy with which places are visited.
From here on, the term "visitor density" will always refer to the average density of individuals as just described.

#### 3.2.2.2 Dwell Demographics

Based on the dwell density data, Swisscom also calculates the expected age and gender of the people who were in a particular tile at a particular time. This is done on the basis of the demographic distribution of the SIM card's municipality of origin, with age and gender sampled (simulated) (Swisscom AG, 2022a). Ages are split into four groups: 0-19, 20-39, 40-64, and >64. The demographic data used for this purpose are taken from the Swiss Federal Statistical Office (Swisscom AG, 2022a).

#### 3.2.2.3 Access and Processing Notes

When working with Swisscom's mobile phone data, it can happen that individual tiles do not contain any values. In the Zurich region, this is particularly true for demographic queries, i.e., age and gender. This is by design, since, as described earlier, for data protection reasons only results involving more than 20 SIM cards are delivered (Swisscom AG, 2022a). This must be kept in mind when working with the data in order to interpret the results correctly.

From the API I got back the daily density value for the three-month period I chose per tile in a table, with the tiles defined by the x and y values of the lower left and upper right corners of each tile. For my further analysis I created a vector dataset based on this information, where a polygon corresponded to a tile and had an attribute with a certain density or demographic ratio value. It must be noted here that this polygon dataset did not have a perfect topology. Some polygons overlapped by a few meters, which might be related to the fact that the coordinates of the tiles seem to be returned by the API only to seven decimal places and adjacent tiles are not aligned with each other, but only with themselves. Since it was not a very large overlap, I used density values, and since the methods used here can only provide approximations anyway, I did nothing further to correct for these inaccuracies. Nevertheless, in future projects where this is not the case, care should be taken to ensure that the topology is correct and that the tiles are placed next to each other in such a way that they share their edges and vertices exactly with the neighboring tiles on all sides.

### 3.2.3 Infrastructure and General Attractiveness

When planning new public spaces, an attempt is usually made to design them in such a way that they are attractive to the public. For this purpose, various objects from the categories of buildings, infrastructure and nature can be used. However, as stated earlier, there is still little detailed information on how a UPS must be designed in order to have a high probability of being well received by the population.

Since the goal of this thesis was to explore whether data-based analyses of existing places, their characteristics, and visitor numbers can be used to predict what characteristics a new UPS should have in order to be popular, I used several datasets that might be relevant to the attractiveness of a public space and are relatively reliable and easily accessible. Therefore, I only used datasets that are openly available for download via the Open Data Portal of the city of Zurich or via Open Street Map (OSM). Table 3 shows all these datasets, as well as their sources and geometry types.

In determining the relevance of various features of a UPS, I have relied in part on typical features that have been mentioned in the literature. The Project for Public Spaces (PPS) has attempted to summarize in its Place Diagram (Figure 1) some of what they consider to be the most important attributes for a successful UPS (Metropolitan Planning Council, 2008). These are divided into four different categories for which I retrieved and used data based on availability and feasibility. The categories are “Use & Activities”, “Comfort & Image”, “Access & Linkages”, and “Sociability” (Metropolitan Planning Council, 2008). The last of these categories deals directly with the number and type of visitors to specific UPSs, which is partially treated as a dependent variable in this thesis and cannot be derived from the built, infrastructural, and natural environment. Therefore, for the independent variables, I selected datasets that fit into one of the other three quality categories.

**Table 3**: Datasets used to evaluate the attractiveness of UPSs

| Category | Dataset | Geometry Type | Source |
|---|---|---|---|
| Use & Activites | Shops | Point | Open Data Stadt Zürich (2022a) & OSM (query: node["shop"]) |
| | Bars | Point | Open Data Stadt Zürich (2022b) & OSM (query: node["amenity"="bar"]) |
| | Restaurants | Point | Open Data Stadt Zürich (2022c) & OSM (query: node["amenity"="restaurant"]) |
| | Markets | Point | Open Data Stadt Zürich (2022a) |
| Access & Linkages | Two-wheel parking | Point | Open Data Stadt Zürich (2021c) |
| | ZüriVelo (public bikes) | Point | Open Data Stadt Zürich (2022d) |
| | Public transport quality | Polygon | Open Data Stadt Zürich (2021d) |
| | Public transport stops | Point | Open Data Stadt Zürich (2021e) |
| | Street parking | Point | Open Data Stadt Zürich (2019) |
| | Parking garages | Point | Open Data Stadt Zürich (2021f) |
| | Pedestrian and bike network | Line | Open Data Stadt Zürich (2021g) |
| | Traffic counts | Point | Open Data Stadt Zürich (2020) |
| Comfort & Image | Züri wie neu (damage reporting app) | Point | Open Data Stadt Zürich (2021h) |
| | Fountains | Point | Open Data Stadt Zürich (2021i) |
| | Trees | Point | Open Data Stadt Zürich (2021j) |
| | Significance | Polygon | Open Data Stadt Zürich (2011) |

#### 3.2.3.1 Uses and Activities

The opportunity to engage in various activities in a public space is repeatedly given high priority in the literature, also in the context of social sustainability. Montgomery (1998) even rates activity as key element for the urbanity of a place. To accommodate different activities of interest to various people, diverse ways of using the space on and in the immediate vicinity of the UPSs are required. Attractors here could be places for social gathering, such as restaurants and bars. In addition, retail such as various shops or markets may also be important to this attraction category.

Locations of attractions in this category are made available by the Zürich Tourismus organization via the Open Data Portal of the city of Zurich using an API. It should be noted, however, that Zürich Tourism has only included stores in its dataset that it probably considers to be of particular relevance to tourists. These are, for example, stores or restaurants that offer local specialties or souvenirs. Especially for the permanent population of the city of Zurich, these stores are likely to have little to no relevance, which is why I merged the datasets of bars, restaurants and shops by comparing the respective OSM ID with the corresponding datasets of OSM. This way I hoped to get a more comprehensive picture of the POIs in this category.

#### 3.2.3.2 Access and Linkages

Another category identified by the PPS as essential to public spaces is its "Access & Linkage". They assign intangible factors such as proximity, connectedness, walkability, and accessibility to this category (Metropolitan Planning Council, 2008). The Open Data Portal of the city of Zurich contains a relatively large number of datasets on these topics. The recorded figures from the traffic counting stations, which are distributed throughout the city, are provided at hourly intervals. I used these to gain insight into whether increased traffic has an impact on the popularity of a UPS and whether it is positive, or negative. There is also a line dataset that includes all pedestrian and bicycle routes in the city of Zurich. I wanted to integrate this network dataset into my analysis to have a measure of the centrality of the UPSs within their neighborhood. In general, in Switzerland, but especially in big cities like Zurich, public transport is also a very important means of transport. Therefore, I have additionally included public transport stops and an already elaborated spatial classification of the quality of public transport. The latter is provided by the canton of Zurich and is based on a combined calculation of the type of stop (rail node, rail line, bus, tram and associated course interval) and the distance to each of the stops for each location in space (< 300 m, 300-500 m, 501-750 m, 751-1000 m) (AFV, 2020).

In order to ensure that people arriving by private transportation can leave their vehicles and get to the UPS, parking spaces are required nearby. To assess how well each UPS performs in this regard, I used data on the location and number of on-street parking spaces and parking garages for cars. Afterwards I added up the two numbers to get a general overview of the availability of parking spaces for cars. In addition to the parking spaces for cars, I also included the parking spaces for bicycles near the UPSs in the analysis.

In Zurich, there has been a growing number of companies offering public bicycles for several years. One of the local companies that do this, and where the bikes have to be picked up and returned at fixed locations, is PubliBike "ZüriVelo". If a pick-up and drop-off point for these bikes is located directly at a UPS, this could increase its accessibility, which is why I decided to include this factor in my analysis as well.

#### 3.2.3.3 Comfort and Image

The last category is perhaps the most subjective of the PPS's Place Diagram. "Comfort & Image" can be perceived and interpreted quite differently within different populations and even between individuals. Nevertheless, there are common denominators proposed as metrics here as well. Intangible factors typically associated with this category include perceived safety, cleanliness, greening, general attractiveness, and cultural and historical significance, for example (Metropolitan Planning Council, 2008). On the one hand, on the Open Data Portal Zurich there is a complete dataset of all trees in the city, which can provide an indication of the greening. Fountains, on the other hand, could arguably contribute to the general attractiveness of a UPS while providing the population with access to free drinking water.

To obtain a measure of cleanliness, I used data from the damage reporting app "Züri wie neu". With the help of this app, the population can mark places where they find deficiencies in the infrastructure, where graffiti has been sprayed or where trash has been left lying around. Finally, to evaluate the image and overall size of the region for which a given place is significant, I used the polygon dataset of the significance of sites, which I had also used to extract the geometries of squares and parks.

## 3.3 Pre-Processing and Data Exploration

### 3.3.1 Pre-Processing of UPS Attributes

Most of the datasets I wanted to use as independent variables for my analysis were already in the appropriate form when I downloaded them. For some of them I had to do some pre-processing steps to make them suitable for my purposes. This includes the pedestrian and bike network datasets and the traffic counts.

#### 3.3.1.1 Traffic Counts

The traffic counts are huge CSV files that contain hourly records of all vehicles traveling in both directions for all stations in the city. Thereby, one file contains a year's worth of data. Since the data is far too detailed for my rather broad analysis, I decided to aggregate it heavily. To do this, I averaged the incoming and outgoing vehicles for each station over the entire year and finally added up the number of incoming and outgoing vehicles. This gave me a point dataset with one averaged hourly value per measurement location. Since there are only 91 measuring points in Zurich, I could not use them directly for my analysis, because there is not such a measuring point in the immediate vicinity of every UPS. If there was a proxy for average traffic for some UPSs and not for others, this would make comparability difficult. Therefore, to get at least an approximation of the traffic volume near each UPS, I calculated Thiessen Polygons from these points using ArcGIS Pro.

#### 3.3.1.2 Network Data

The pedestrian and bicycle network can be downloaded as a line dataset. The dataset is huge and could be useful for my purposes only if it could tell me something about the centrality of each UPSs in the network. In a first step I separated pedestrian and bicycle paths and then defined them as single networks. Within these two networks, I subsequently performed a closeness centrality analysis of the nodes. For the calculation of the closeness centrality, I used the momepy library (Fleischmann, 2019), which provides various options for centrality calculations. I decided to calculate a local instead of a global closeness, based on the assumption that urban public spaces should play an important role especially in neighborhoods in their immediate vicinity. For this purpose, I chose a threshold radius of 1200 meters, which corresponds to the rounded average distance of the centroid of all Zurich districts to their outer border. The result of this calculation was a point dataset in which the individual network nodes contained a closeness centrality value.

### 3.3.2 Calculating Visitor Density and Demographics per UPS

To be able to make a statement about the approximate number of visitors to individual UPSs, I needed a way to sample the density values from the Swisscom data, which were available in a 100 m x 100 m grid, down to the individual UPSs geometries. For this I intersected the two polygon datasets of the squares and the dwell densities of the mobile phone data. I then calculated what percentage of a given UPS lies in a given tile of the mobile phone dataset (piece of UPS/10'000 $m^2$) multiplied that by the measured visitor density of that tile. The weighted density values for the individual parts of a UPS that were located in different tiles were then reassembled. To do this, I calculated the sum of all the weighted density values of the individual parts that belonged to a particular UPS. Finally, to make the calculated numbers of the different UPSs comparable, I then divided this summed density value by the area of the UPS and multiplied it by 10'000, so that the values in the end were again in visitor density per hectare. I did this for all 92 days of my three-month time period and output the results into a table at the end. The pseudocode version of the calculations presented here can be seen in Figure 2.

```
for each date do
    intersection = ups.overlay(mobilePhoneData.date)
    weightedDensityIntersection = (intersection[Area]/10000)*intersection[Density]
    weightedDensity = weightedDensityIntersection.sum(on = ups[Name])
    resultingDensity =  (weightedDensity/ups[Area])*10000
end
```

**Figure 2**: Pseudocode of the applied methods for downsampling the visitor density values to the UPSs geometries

The demographic data do not include absolute values, but ratios for the different age groups and genders. Therefore, I have chosen a slightly different calculation approach for their processing. I also calculated a weighting by calculating what percentage of the total area of a UPS is intersected by a specific tile containing the demographic information. I then multiplied this weight by the respective ratio value of the tile and added up the shares of all individual tile parts for each UPS at the end. Due to Swisscom AG's privacy policy (more than 20 SIM cards need to have been found in a tile), especially for demographic data, some tiles did not contain data, so it was critical to develop a catch-all method for this case. I decided to handle these events by simply reusing the ratio of the last processed tile that was within the UPS and contained data when no data was available for a tile.

### 3.3.3 Edge Effects

The question arises as to how accurately the mobile phone data provided by Swisscom can provide information on the whereabouts of the population. This is especially critical because the data has a resolution of 100 m x 100 m, but many UPSs are smaller. Thus, the data must be downsampled, which could result in the distribution of visitor density numbers simply matching the general distribution throughout the surrounding region. Such edge effects would also imply that a correlation between the visitor frequency of UPSs and the existing

infrastructure would not allow to make a statement about the UPSs, but rather about their general surroundings. Therefore, I wanted to exclude from my analysis UPSs where such edge effects occur, and thus the number of visitors consequently does not differ significantly from those of the surrounding environment.
Since, so far, datasets similar to the mobile phone dataset I had received from Swisscom have very rarely been used in research for similar analyses, I have not found any existing methods for this distinction, despite an extensive search. Therefore, I have created my own method to predict at least with a certain degree of certainty that the visitor density assigned to a UPS is actually valid and distinctive. To do this, I constructed a buffer around the square, calculated the distribution of visitor density over the fixed period of three months within this buffer region, and then compared this to the visitor distribution of the UPS using a chi-square test. One difficulty here was the calculation of the buffer region. While it is very easy to calculate a buffer around a polygon based on a given fixed distance, in this case I wanted to calculate the buffer region as a multiple of the square area for comparability of different squares, which is considerably more complicated. This is the case because it is geometrically more complex to construct an arbitrary shape with a given area. To achieve this anyway, I programmed the buffer calculation to iteratively try different distances for calculating the buffer and compare the resulting buffer with a multiple of the UPS area until the area of the calculated buffer deviates from the target area by a maximum of 0.01 $m^2$.
Another difficulty was the determination of a meaningful multiple of the UPS area for the buffer calculation. I exploratively tried different buffer sizes, visualized them, and analyzed the resulting statistically significant visitor distributions for different UPSs, both for the year 2020 and 2021. With smaller buffer area, there were fewer UPS that showed a statistically significant visitor distribution over the three months. This is likely to be the case, as visitor density changes continuously across the entire urban area. Therefore, due to the downsampling method applied, which assumes that visitors are evenly distributed within tiles, it is clear that regions in the immediate vicinity of the UPS under consideration have a very similar visitor density distribution. Somewhat larger buffer regions were therefore desirable, although the size should not be so large that it covers too large an area and therefore no longer expresses anything about the specific visitor frequencies of a particular region of the city of Zurich. I have therefore exploratively tried out different buffer sizes. Above a certain buffer size, the number of UPSs for which a statistically significant visitor frequency could be determined barely changed. Therefore, I finally decided to use a buffer of 13 times the area of the UPSs around the UPSs as a reference for the chi-square test. In many cases, this was roughly equivalent to including all the nearest blocks around the UPSs.
In order to have the direct comparison, I analyzed the same buffer for both years 2020 and 2021 respectively and combined all statistically significant UPSs that were found. A total of 34 UPSs had a statistically significantly different distribution of visitor density than their surrounding region. 11 places were identified as statistically significant only for the year 2021. Other 23 were found for the year 2020 as well as for the year 2021 (see Figure 3). Since all UPSs found in 2020 were also found to be statistically significant in 2021, I decided to use all UPSs found in 2021 for further analysis. Additionally, because of this, I also decided to focus on the data from 2021 for further analysis.

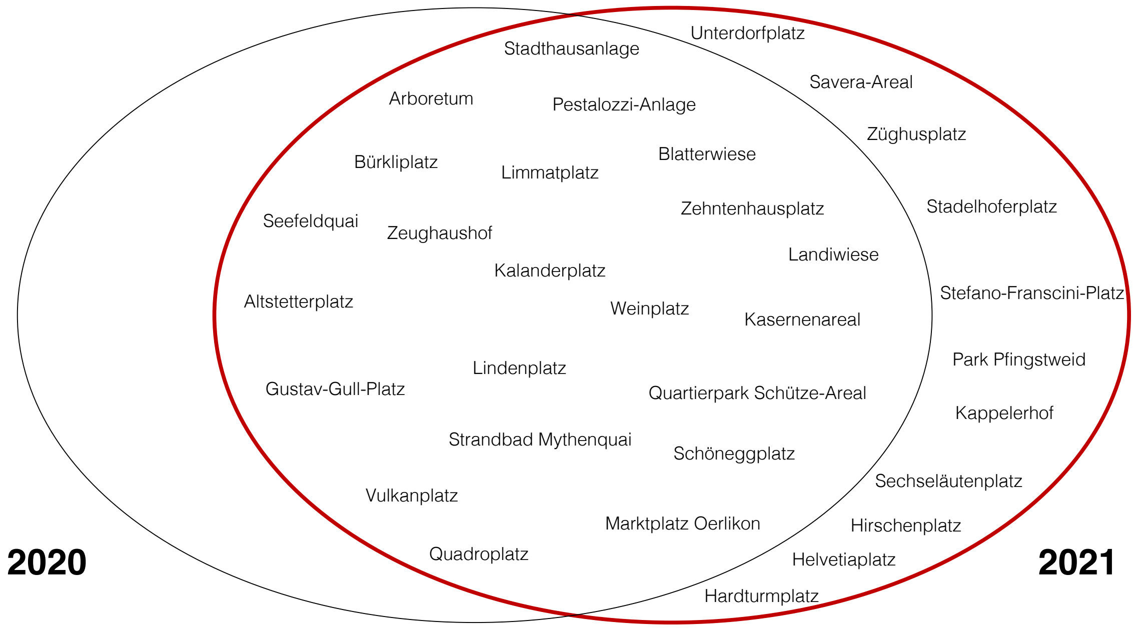



**Figure 3**: The 34 UPSs that were found to have a statistically significantly different distribution of visitor density than their surrounding area

### 3.3.4 Similarity of Visitor Density Patterns

Different squares and parks also serve different purposes for the population because of their geographic location. Some are heavily used during weekdays because they are surrounded by offices and other workplaces. Others, serve more for recreation on weekends. To better understand my data, I intended to find just such differences and divide the observed sites accordingly into groups with similar visitor density distributions. Thus, I first computed an average visitor density score for each square for each day of the week. From this value, I then subtracted the mean visitor density score of the entire time series. If the absolute value of this difference for one day of the week was larger than the standard deviation of the whole time series, I marked this day of the week as significantly different from the rest of the days of the week. Thereby, the visitor density value for one weekday could be classified as either distinctly higher than the other weekdays or distinctly lower. In a next step, I then classified the UPSs into five semantically similar (based on the pattern of their significant days) groups that I had predefined. These five groups and the rules used for classification can be found in Table 4.

**Table 4**: Predefined groups of UPS functionalities including rules used for classification (where 0 = not different, 1 = significantly higher, 2 = significantly lower)

| Group | Class | Function | Rule (pseudocode) |
|---|---|---|---|
| 1 | Same the whole week | Transport hubs, offices, and residence | If all_days == 0 |
| 2 | Saturday and Sunday lower | Many offices | If saturday == 2 & sunday == 2 |
| 3 | Saturday lowest | Offices, residence and recreation, no consumption | If saturday == 2 & all_others == 0 |
| 4 | Saturday highest, Sunday lowest | Recreation, no consumption | If sunday == 1 & sunday == 2 |
| 5 | Sunday lowest | Offices and consumption | If sunday == 2 & all_others == 0 |

## 3.4 Intersection of the Data

In order to assign all UPSs their attributes, and thus be able to perform a correlation analysis later, I intersected all the attribute datasets described above with the UPS geometries. Depending on the type of dataset to be intersected with the UPSs, I worked with a buffer around the original geometries of the UPSs. I chose this because, for example, trees or fountains are often located directly in the area of a UPS, but restaurants and stores may be located on adjacent properties around the square or park. Nevertheless, they can then serve as a point of attraction for visitors to the UPSs. I worked with different buffer sizes to intersect different datasets with the geometries of the UPSs. I set the buffer sizes based on a walk time calculation. To calculate the distance that could be walked in this time, I relied on the average walking speed of people in an unobstructed pedestrian flow of 1.34 m/s, as determined by Buchmüller and Weidmann (2006) in their report on the parameters of pedestrians, pedestrian traffic, and pedestrian facilities. I assumed that different POIs with a walk time of 0, 30, or 60 seconds would be relevant to a UPS.

After each intersection, I additionally normalized all point data that were not binary in nature and could be directly related to the size of the UPSs with the area of the UPSs (POIs/hectare). The datasets normalized in this way are the trees, fountains, damage reports (Züri wie neu), public transport stops, two-wheel parking, restaurants, bars and shops.

### 3.4.1 Walk Time 0 Seconds

I did not use a buffer for all the datasets that were either POIs that could be located directly on a square or park, or that contained polygons. The POIs included the trees, fountains, reports from the "Züri wie neu" damage reporting app, and markets. The polygon datasets consisted of the calculated Thiessen polygons of the traffic counts and the public transport quality classes. If a UPS intersected two or more polygons with different values, I adopted the higher value in each case. This is due to the assumption that the distances within a square or park in Zurich are generally not very large and therefore one is affected by the higher traffic value at almost all locations on the UPS anyway or can benefit from the higher public transport quality class.

### 3.4.2 Walk Time 30 Seconds

For most of the datasets, I used a foot interval of 30 seconds, which is about 40 meters. I chose this distance because I assumed that most of the POIs that we still perceive as part of a UPS must be situated in the first row of houses around the UPS. We should be able to reach these in about 30 seconds walking from the outer boundary of the UPS, even if we have to cross a street, for example. This applies to restaurants, bars and shops. In addition, I also assumed that a public transport stop could be perceived as belonging to the UPS at this distance, just like bicycle pumping stations. Furthermore, I presumed that bike parking and public bike stations could be used at this distance to reach a UPS in a few steps after parking the bike.
The lines of the pedestrian and bicycle network never cross the geometry of a UPS directly for reasons of topology. However, they often run right next to it. To include the calculated network nodes with closeness centrality values in my analysis, I thus also intersected this dataset with the geometries with 30-second buffers. Since I wanted to assign only one value per UPS for pedestrian and bicycle closeness, I proceeded again as with the polygon intersection of traffic counts and public transport quality classes. If several network points were located in the immediate vicinity of a UPS, I used the highest closeness centrality value as a reference.

### 3.4.3 Walk Time 60 Seconds

Generally, I hypothesized that a walk time over 30 seconds to POIs would mean that they could no longer really be considered relevant attributes for a UPS. However, in the case of parking spaces for cars, it is often the case that people have to park a little further away due to space constraints. Therefore, I decided to apply a walk time buffer of 60 seconds for the measurement of the number of available street parking spaces and parking spaces in garages.

## 3.5 Correlation Analysis

In selecting the method that would be best to find out correlations between the number of visitors and the characteristics of a UPS, I relied on the methods described in chapter 2.5. Machine learning models have been increasingly used for analyses similar to mine in the last few years. However, since my final dataset was relatively small due to the limited number of UPSs in the city of Zurich and the additional exclusion of UPSs that did not show a statistically significant difference in the number of visitors to their immediate surroundings, I excluded this option. Instead, I decided to perform a simple regression analysis.
One of my goals in this thesis was to find out if there is a correlation between the attributes of a UPS and its popularity. Thereby I followed the hypothesis that a higher number of certain built, infrastructure, or natural elements could also lead to a higher attractiveness for the population. I assumed that correlations between the attributes and visitor frequencies would be linear. Because of this, I chose multiple linear regression (MLR) as the method for my correlation analysis. I used MLR for two different levels of aggregation of the data. On the one

hand, I calculated a score for each UPSs in the three different quality categories as defined by the PPS (Metropolitan Planning Council, 2008) and then checked if the scores of the categories show a correlation to the aggregated visitor densities of the UPSs. Then, to also find out if individual attributes were responsible for higher popularity, I performed the analysis for the individual independent variables.

### 3.5.1 Quality Categories

To calculate the scores of each UPS in the three quality categories, I scaled the data of each dataset belonging to a category ("Use & Activities", "Comfort & Image" and "Access & Linkage") between 0 and 1. Thus, the UPS that performed best had a scaled score of 1 and the UPS that performed worst had a scaled score of 0. Subsequently, I used the average of all scores calculated in this way to calculate the total score in a quality category. I thereafter used these resulting total scores for each category to conduct a MLR analysis with 3 independent variables. In addition, I summed up the scores of each quality category to give each UPS a specific total quality score.

### 3.5.2 Individual Attributes

In order to include as few variables as possible in the regression analysis that already show very strong correlations with another variable, in a first step I performed a qualitative correlation analysis of the independent variables. For this purpose, I created a correlation matrix using the Python library seaborn (Waskom, 2021).
Since I wanted to check a whole range of different UPS characteristics for their influence, I had to be very careful not to overfit the linear model. Therefore, I decided to use the backward selection method (JMP, 2022). In each step, the variable with the highest p-value is removed from the model, since it is assumed that it could be more precise without this variable (JMP, 2022). This is done until every single variable that is still in the model is found to be statistically significant ($\alpha < 0.05$).

# Chapter 4 | Results

## 4.1 Urban Public Spaces of Zurich

Based on the methods described in Section 3.2.1, I gathered the point coordinates of 180 squares and parks within the city boundaries of Zurich. However, after intersection with the polygon boundaries of the significance dataset of the city of Zurich, it became clear that some of these UPSs had been grouped and placed within a single polygon geometry by the city. In this case, I decided to use only one of the UPSs found for the respective polygon at a time, representing the entire group of UPSs. The grouping had the following structure:

- **Blatterwiese**: Blatterwiese, Seeuferweg, Zürichhorn
- **Bürkliplatz**: Bürkliplatz, Bürkliterrasse
- **Zwingliplatz**: Zwingliplatz, Grossmünsterplatz

Additionally, when looking closely at the geometries found, it became apparent that due to the applied way of searching squares, the place "Sunnige Hof", which resembles more a neighborhood street than a square, was included in the dataset. Since this work is specifically about UPSs that correspond to squares or parks, I excluded “Sunnige Hof” as well.
I thus entered the pre-processing and data exploration phase with polygon geometries of 175 UPSs. According to the boundary geometries used, the largest UPS in the city in terms of area is Irchelpark with an area of 84,527 m$^2$, and the smallest is Häringsplatz with an area of 245 m$^2$. The average area is 8758 m$^2$, with the median being lower at 3950 m$^2$.

## 4.2 Visitor Frequency Data

### 4.2.1 Data Availability

For data protection reasons (see also Section 3.2.2), Swisscom AG only includes density values in its dataset where more than 20 SIM cards were identified (Swisscom AG, 2022a). Therefore, it would be possible that in this case a visitor density (visitors/area of the UPS) of 0 was calculated for certain UPSs on certain days. If this would occur too often in fact, a meaningful analysis would not be possible at all. However, this was not the case. Of the total 16’100 visitor density values calculated for 175 UPSs over 3 months in summer 2021 (92 days), only 209 of them are 0. This corresponds to 1.3% of all values.
The range of all visitor densities sampled extends from 0 to 10368 visitor density per hectare, with the latter value for Helvetiaplatz measured on Saturday, 04.09.2021.

### 4.2.2 Statistically Significant Squares

A simple review of the data revealed that the calculated values for visitor density in the squares and parks may have been little to heavily influenced by normal visitor density in their immediate vicinity. To avoid such edge effects, I calculated the statistical significance of the visitor densities of the individual UPSs compared to their surroundings (see section 3.3.3). The 34 UPSs with statistically significant differences in visitor density per hectare from their surrounding areas are shown in Figure 4. In addition, Figure 4 also shows the respective average visitor density per hectare per day of each UPS.

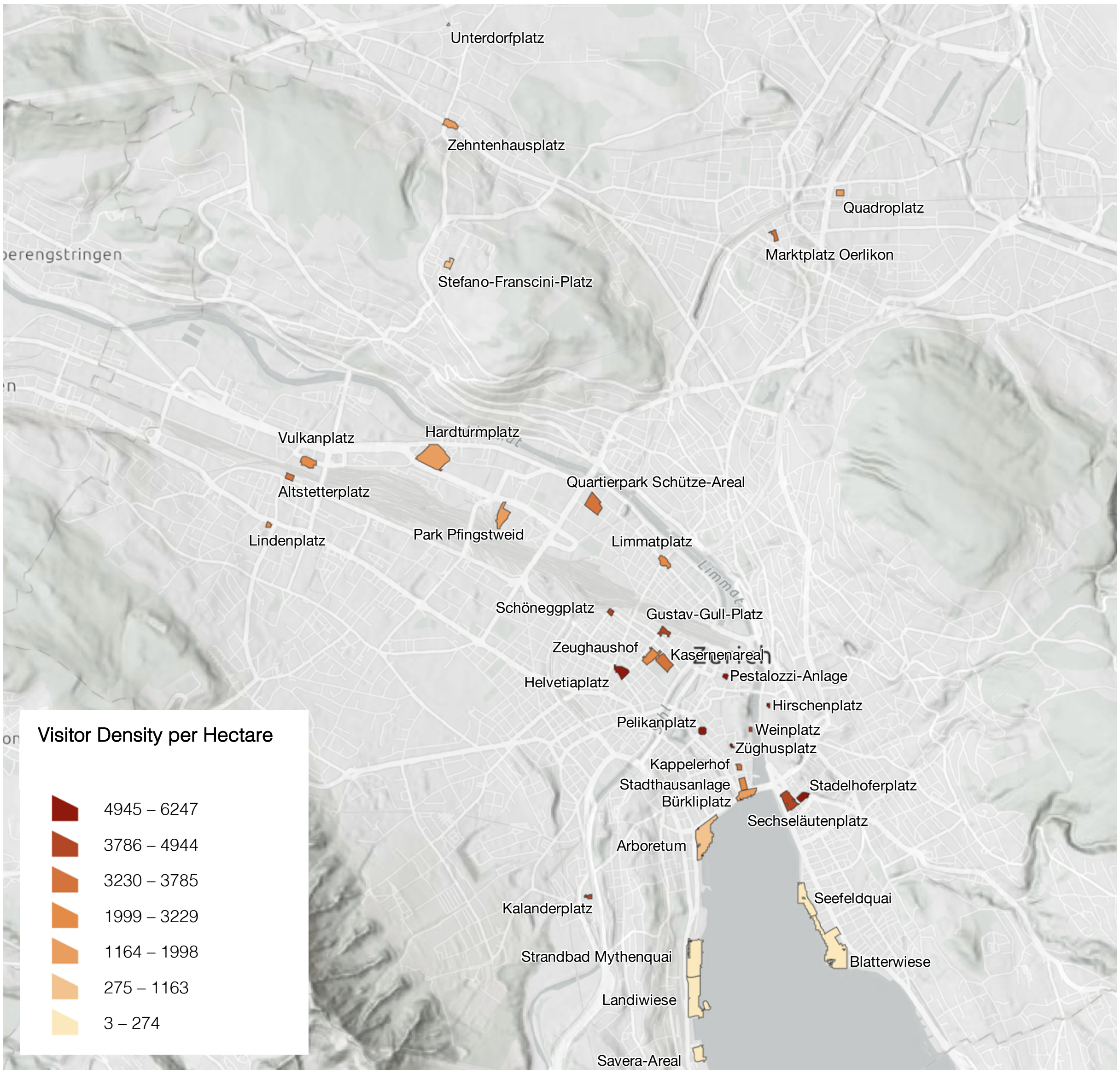



**Figure 4**: The 34 UPSs for which the daily visitor density per hectare is statistically significantly different from their surroundings (Data: Open Data Stadt Zürich (2021))

I had expected the visitor densities of larger UPSs in particular to be statistically significant, since the size of their area is more like the size of the tiles from the mobile phone dataset (100 m x 100 m) and the calculated values are therefore less approximated. Indeed, data from many rather larger UPSs were found to be statistically significant. Interestingly, however, there were also several UPSs with areas significantly smaller than the tiles in the mobile phone dataset for which visitor frequencies could be calculated that were statistically significant. The downsampling as applied thus also seems to work for smaller UPSs. The smallest UPS identified as significant is Unterdorfplatz with an area of only 367 $m^2$ (0.04 tiles), the largest is Blatterwiese with an area of 55,763 $m^2$ (5.6 tiles). The median area of the squares found is 6508 $m^2$ (Limmatplatz), which corresponds to 0.7 times the area of a tile in the mobile phone dataset. The average area of the 34 relevant UPSs found is 11494 $m^2$ (1.1 tiles).

### 4.2.3 Temporal and Spatial Differences in the Popularity of UPSs

#### 4.2.3.1 Comparison Weekdays and Weekend

Figure 5 shows the average visitor density on weekdays and weekends for the 34 UPSs, which are statistically significant. Large differences can be seen here. Stadelhoferplatz attracts by far the most people, with an average of 6773 people per hectare per day on a weekday. It is followed by Pestalozzi-Anlage and Pelikanplatz with 6387 and 5928 visitor density/ha, respectively. The fewest daily visitors during the week were detected at Strandbad Mythenquai (0), Landiwiese (1) and Blatterwiese (28).

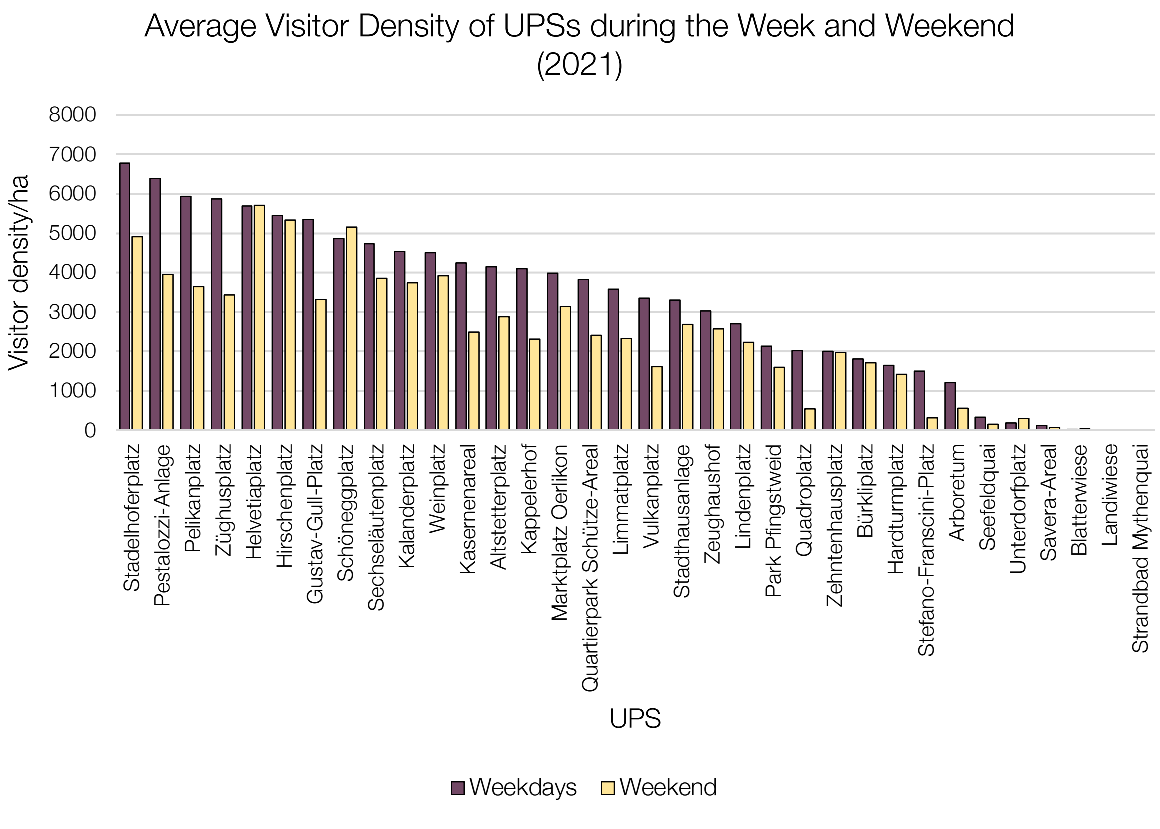



**Figure 5**: Comparison of the average visitor density of the statistically significant UPSs during weekdays and weekend

The distribution of visitor numbers on weekends is partially similar to weekdays, but the overall ranking of visitor densities changes slightly. This is due to the fact that some UPSs have much fewer visitors at the weekend than on weekdays, while others have even higher visitor numbers at the weekends. Helvetiaplatz has by far the highest visitor density (5706) of all UPSs in the city on weekends. On weekends, Landiwiese is at the bottom of the list with a visitor density of 8 per hectare. A total of six UPSs showed an increased density of visitors on weekends. These are Helvetiaplatz (+0.2%), Schöneggplatz (+5.9%), Blatterwiese (+42.9%), Unterdorfplatz (+53.7%), Landiwiese (+700%) and Strandbad Mythenquai. (+3600%). In the case of the Landiwiese and the Strandbad Mythenquai, the very strong rise is related to the fact that they very often showed no visitors at all on weekdays. The most significant decreases in visitor density over the weekend occurred at Stefano-Franscini-Platz (-79.2%), Quadroplatz (-73.1%), Seefeldquai (-54.5%), and the Arboretum (-54.1%). This makes sense considering their location in the city. Stefano-Franscini-Platz is located on the campus of the Swiss Federal Institute of Technology (ETH) and Quadroplatz is surrounded by large office buildings. Across all UPSs, there were 26.5% fewer visitors over the weekend.

#### 4.2.3.2 Differences in Daily Visitor Patterns

Looking at daily data for the UPSs in the three summer months of July, August, and September 2021, highly variable patterns in visitor frequency are evident. For illustrative purposes, I present here the temporal patterns of visitor frequency for four UPSs that exhibit particularly large differences among themselves or other unique features.

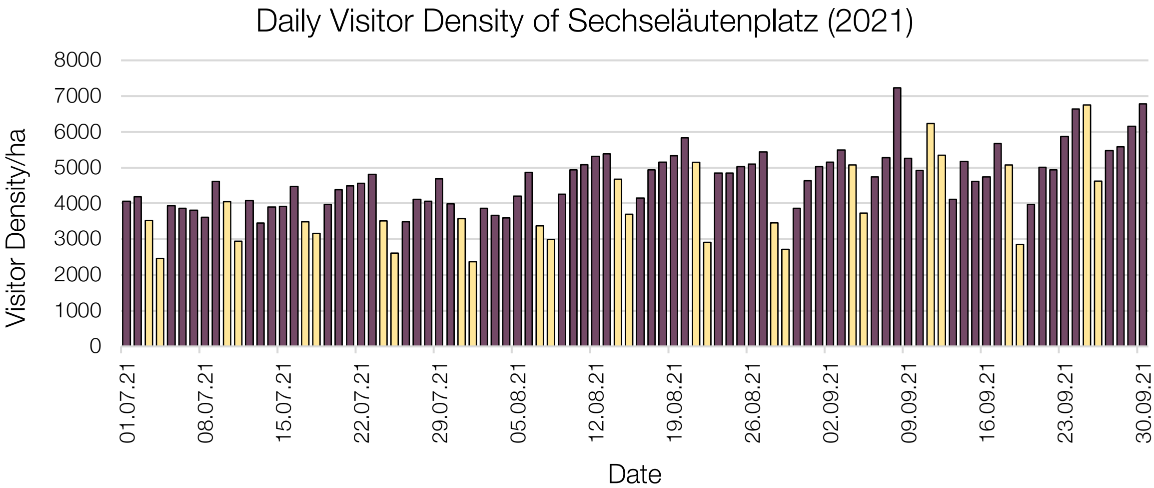



**Figure 6**: Daily visitor density of the Sechseläutenplatz for July, August, and September in 2021 (purple: weekdays, yellow: weekends)

In the case of Sechseläutenplatz (Figure 6), it can be seen that it is regularly least frequented on Sundays. Especially on Sunday, August 1, not many people seem to have been there, with a density value of 2369. This could be because August 1 is the Swiss national holiday and people prefer to spend this day at home or in places where it is possible to use fireworks, for example. Moreover, this day is in the middle of the summer school vacations. A clear peak is visible on Wednesday 8 September. On this day, the Weltklasse Zürich (athletics competition) took place at the Sechseläutenplatz (Zurich Diamondleague, 2021). On the Saturday with the highest density of visitors, 25.09.21, the Zurich Film Festival took place on and near Sechseläutenplatz (ZFF, 2021).

A completely different pattern in the visitor density distribution can be found for the Strandbad Mythenquai (Figure 7). No to relatively few visitors were registered there during the week. It can be clearly stated that the highest visitor density by far can be found on weekends. With 150, the highest visitor density could be determined on August 21.

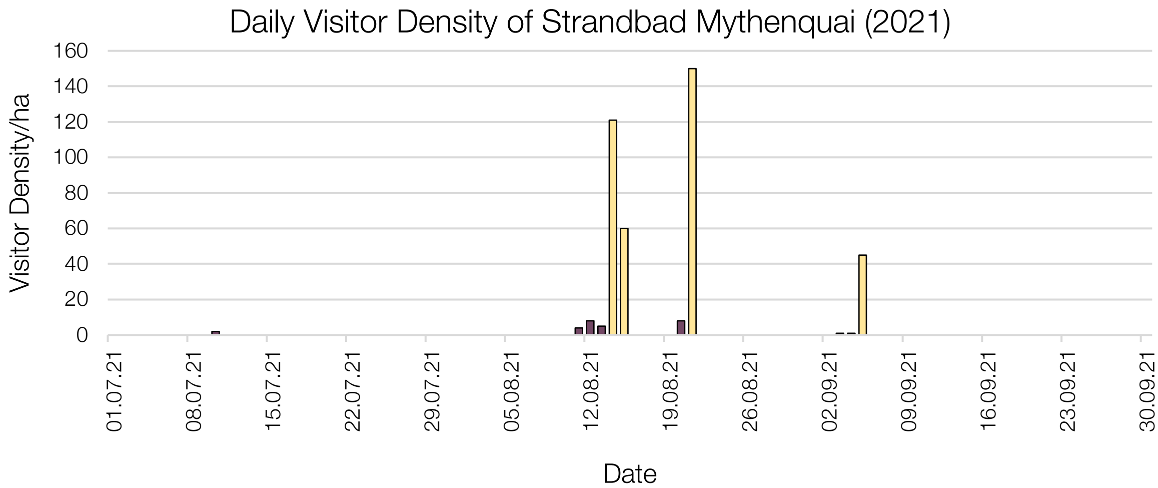



**Figure 7**: Daily visitor density of Strandbad Mythenquai for July, August, and September 2021 (purple: weekdays, yellow: weekends)

The summer of 2021 was a comparatively cold and wet summer. According to the Swiss national meteorological service, it was one of the wettest since the beginning of measurements (MeteoSchweiz, 2022). It therefore makes sense that the visitor numbers for the Strandbad Mythenquai, which is a lakeside bathing facility, are very low for the summer months in 2021. In addition, the lower limit of at least 21 SIM cards set by Swisscom AG under data protection law could play a role here for a tile to have any value at all. Nevertheless, the numbers for this UPS seem a bit too low and might need to be checked in future projects. A total of only 11 days were registered that contain visitor density information. It is also possible that this is a relic of the type of downsampling applied.

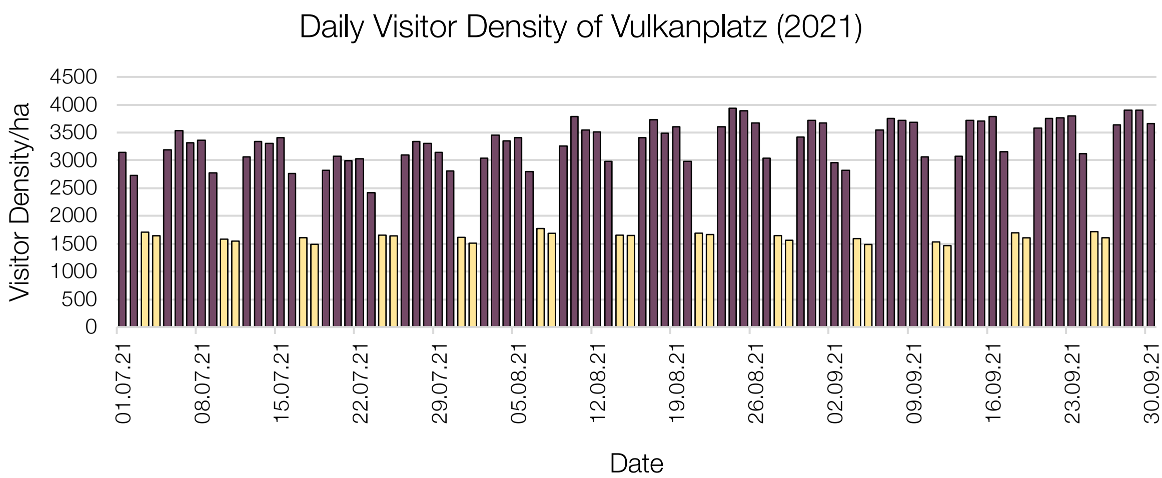



**Figure 8**: Daily visitor density of Vulkanplatz for July, August, and September 2021 (purple: weekdays, yellow: weekends)

Vulkanplatz shows a big drop in visitors (-52.9%) on weekends compared to weekdays (see Figure 8). This could be due to the fact that Vulkanplatz is located right next to a train station and is therefore probably frequented mainly when most people go to work. While weekend visitor numbers always remain relatively stable throughout the time period, weekday visitor numbers seem to be slightly lower roughly between July 17 and August 22 and become larger again end of August. There were school vacations (5 weeks) in the canton of Zurich during this time. So, it might be that the slightly lower numbers during the week are due to the fact that more people were on vacation or at least out of the office.
Visitor densities at other locations sometimes can also exhibit some clearly discernible anomalies in their temporal patterns. Helvetiaplatz saw a relatively high increase in visitor numbers on a single Saturday, September 4 (Figure 9). This Saturday the Zurich Pride took place, demonstrating for the rights of the LGBTQ community and for a yes to "marriage for all" on September 26 (SRF, 2021). Helvetiaplatz was the meeting ground for it.

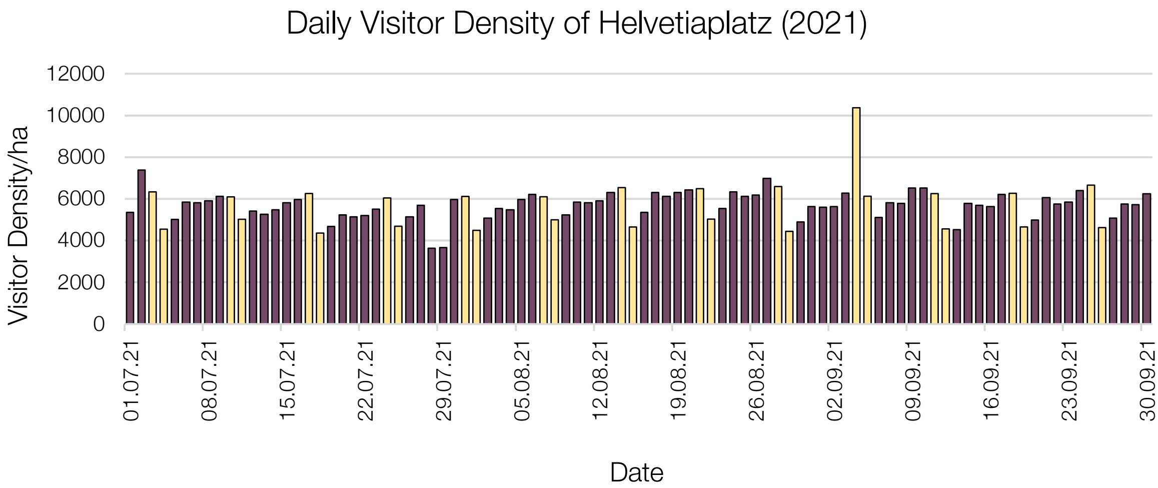
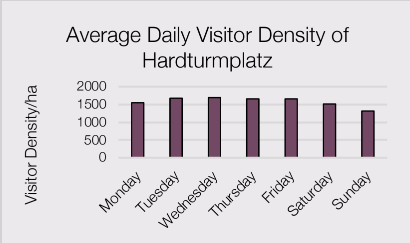
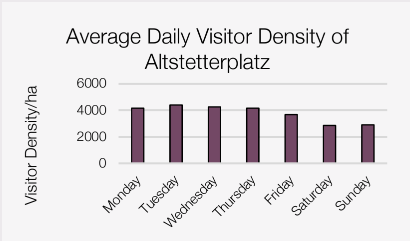
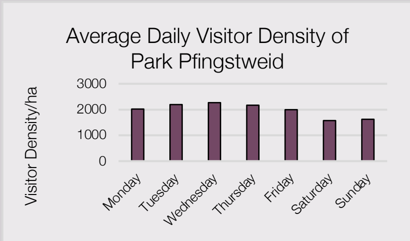
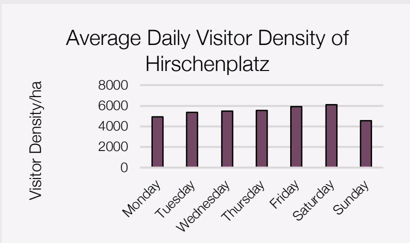
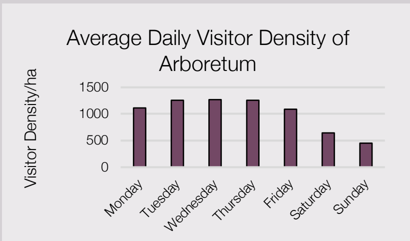



**Figure 9**: Daily visitor density of Helvetiaplatz for July, August, and September 2021 (purple: weekdays, yellow: weekends)

### 4.2.3.3 Temporal Patterns

The temporal distribution of visitor numbers of different UPSs is sometimes very similar. Therefore, they may be classified into groups of sites that may have similar physical characteristics or serve similar purposes for the population. In order to get a better understanding of these differences, and what can be read from the mobile phone data, I have classified the UPSs into five classes as described in section 3.3.4. The result of this classification is shown in Table 5.

**Table 5**: Different groups of UPSs, based on the distribution of their visitor density over the week

| Group | Distribution | Name |
|---|---|---|
| Other patterns |  | Savera-Areal<br>Blatterwiese<br>Schöneggplatz<br>Landiwiese<br>Unterdorfplatz<br>Hardturmplatz<br>Strandbad Mythenquai |
| Significantly lower densities on Saturdays and Sundays |  | Altstetterplatz<br>Vulkanplatz<br>Quadroplatz<br>Stefano-Franscini-Platz<br>Quartierpark Schütze-Areal |
| Significantly lower densities on Saturdays |  | Park Pfingstweid |
| Significantly higher densities on Saturdays and lower on Sundays |  | Hirschenplatz<br>Helvetiaplatz |
| Significantly lower densities on Sundays |  | Stadthausanlage<br>Arboretum<br>Pelikanplatz<br>Pestalozzi-Anlage<br>Seefeldquai<br>Bürkliplatz<br>Züghusplatz<br>Limmatplatz<br>Marktplatz Oerlikon<br>Stadelhoferplatz<br>Kalanderplatz<br>Kappelerhof<br>Weinplatz<br>Sechseläutenplatz<br>Gustav-Gull-Platz<br>Zehntenhausplatz<br>Lindenplatz<br>Zeughaushof<br>Kasernenareal |

The group to which the most UPSs were assigned is the one whose attendance is markedly the lowest on Sunday. A total of 19 of the 34 statistically significant UPSs belong to this category. In contrast, only 2 UPSs have the highest attendance exactly on Saturday, while Sunday has statistically significant low numbers. These are the Hirschenplatz and the Helvetiaplatz. The Park Pfingstweid is the only UPS that has lower attendance only on Saturday, while a total of 5 UPSs have significantly lower visitor numbers over the entire weekend than during the week. However, this difference between the two groups is hardly noticeable when looking at the visitor densities of the weekdays, since the Pfingstweid park also shows much fewer visitors on both weekend days. The visitor density on Saturday seems to have been just barely classified as significantly lower and that on Sunday as normal due to the chosen method of calculation. Finally, the last group consists of 7 UPSs hat show other patterns in weekly visitor density that are not clearly defined.

### 4.2.4 Demographic Data

In order to possibly find out more about the current target groups of the UPSs of the city of Zurich, I have also processed demographic figures contained in the mobile phone dataset by the Swisscom AG (Swisscom AG, 2022a). As with the dwell density data, the final demographic data for each UPS has been computed with a relatively high degree of approximation, on the one hand on the part of Swisscom AG, and on the other hand, on my part. Thus, in order to better estimate the data consistency, I have calculated some indicators before further processing. The results of this preliminary analysis can be seen in Table 6.
The demographic data available is a classification of the people present into 4 age groups and the binary gender classification of male or female. For these two pieces of information, I calculated a percentage indicating how many of the 92 total days examined actually contained data. In addition, I summed the average percentages of all age groups over the entire period and determined the average percentage of males for each UPS.
The results (Table 6) show that information about people's gender is much more available than information about their age group. While information on gender distribution was available on a daily basis for 15 of the 34 UPSs studied (days with results = 100%), this was the case for age groups for only 2 UPSs. However, an additional 11 UPSs more have this information for more than 90% of the days considered. 6 UPSs could not be assigned visitor age information on any day. These are UPSs that also had some of the fewest visitors compared to all UPSs. To know if the calculated ratios for the age groups of visitors to a UPSs could be roughly correct, I averaged the ratios of the age groups over the 3 months and then summed them all together. If this sum came close to 100%, and thus simply described all visitors to the UPS, I considered the numbers to be potentially meaningful. The resulting range of values is between 98.7% (Stadthausanlage) and 104.3% (Weinplatz) for all UPSs for which data on the distribution of visitors into age groups were available on at least one day. From this I concluded that, despite several approximations during preprocessing, relevant figures for the age distribution can still be extracted from the mobile phone data for the UPSs geometries.

**Table 6**: Results of the preliminary analysis of the consistency of the demographic numbers of each statistically significant UPS

| | Age Groups | | Gender | |
|---|---|---|---|---|
| Name | Days with Results [%] | Coverage of all age groups [%] | Days with Results [%] | Male proportion [%] |
| Altstetterplatz | 81.5 | 99.8 | 96.7 | 49.7 |
| Arboretum | 63 | 100 | 95.7 | 49.9 |
| Blatterwiese | 0 | | 34.8 | 50 |
| Bürkliplatz | 80.4 | 100.3 | 98.9 | 50 |
| Gustav-Gull-Platz | 100 | 100.9 | 100 | 49.9 |
| Hardturmplatz | 79.3 | 100 | 100 | 49.6 |
| Helvetiaplatz | 100 | 99.6 | 100 | 49.7 |
| Hirschenplatz | 88 | 100 | 89.1 | 49.6 |
| Kalanderplatz | 87 | 101.5 | 96.7 | 50.3 |
| Kappelerhof | 91.3 | 99.4 | 100 | 49.5 |
| Kasernenareal | 93.5 | 99.9 | 100 | 49.8 |
| Landiwiese | 0 | | 0 | |
| Limmatplatz | 92.4 | 100.2 | 100 | 49.8 |
| Lindenplatz | 60.9 | 100.3 | 98.9 | 49.7 |
| Marktplatz Oerlikon | 90.2 | 98.4 | 95.7 | 48.7 |
| Park Pfingstweid | 65.2 | 100.2 | 98.9 | 49.8 |
| Pelikanplatz | 98.9 | 99.8 | 100 | 49.7 |
| Pestalozzi-Anlage | 96.7 | 101.2 | 100 | 50 |
| Quadroplatz | 68.5 | 99.1 | 78.3 | 49.3 |
| Quartierpark Schütze-Areal | 93.5 | 100 | 100 | 49.7 |
| Savera-Areal | 0 | | 0 | |
| Schöneggplatz | 97.8 | 100.4 | 100 | 50 |
| Sechseläutenplatz | 100 | 100.2 | 100 | 49.9 |
| Seefeldquai | 0 | | 70.7 | 49.8 |
| Stadelhoferplatz | 97.8 | 100.6 | 100 | 50.2 |
| Stadthausanlage | 78.3 | 98.7 | 89.1 | 49.3 |
| Stefano-Franscini-Platz | 33.7 | 101.8 | 72.8 | 49.9 |
| Strandbad Mythenquai | 0 | | 0 | |
| Unterdorfplatz | 0 | | 0 | |
| Vulkanplatz | 69.6 | 99.1 | 100 | 49.4 |
| Weinplatz | 92.4 | 104.3 | 96.7 | 51.9 |
| Zehntenhausplatz | 30.4 | 100.3 | 100 | 49.4 |
| Zeughaushof | 97.8 | 100.2 | 100 | 49.8 |
| Züghusplatz | 78.3 | 100 | 92.4 | 49.6 |

Figure 10 shows the distribution of values of average ratios for all age groups for each UPS over the three months used for this analysis. It can be seen that there is relatively little variation in the distribution within age groups across all UPSs. The median value for the 0-19 age group is 0.1831, and for the 20-39 age group it is 0.2788. The 40-64 age group has the highest proportion, with a median value of 0.3740. The 65+ age group has the lowest proportion, with a median value of 0.1674. For comparison, Figure 11 shows the ratios for the respective age groups in Switzerland, the Canton of Zurich and the City of Zurich. The data, which are provided by the Federal Statistical Office, were only available for 2020, but are unlikely to have changed significantly within a year, which is why a direct comparison still makes sense.

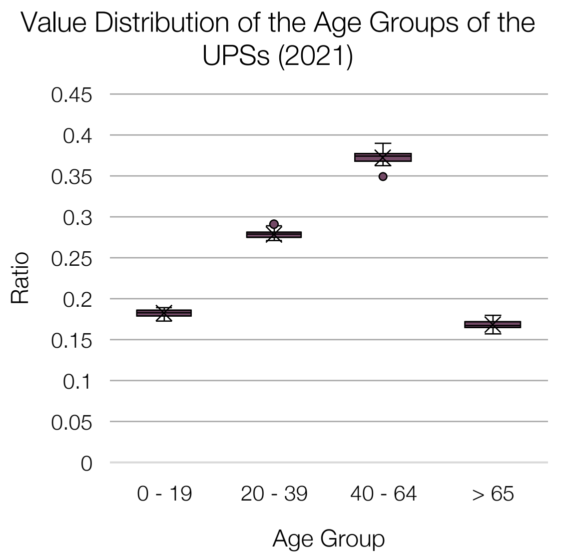



**Figure 10**: Age groups distribution of all UPSs, where age groups could be determined for at least one day

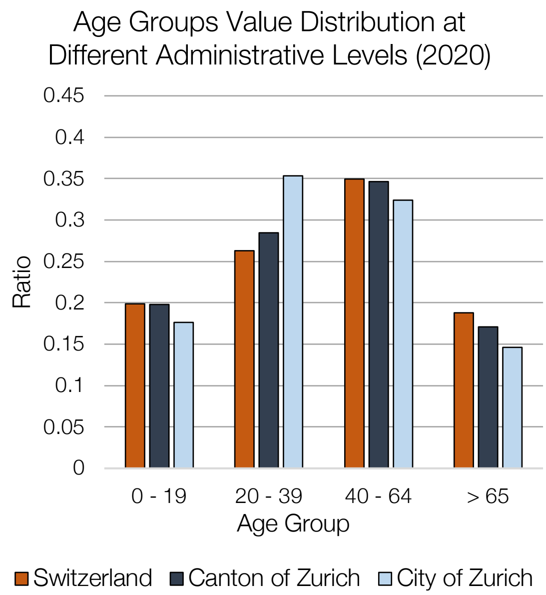



**Figure 11**: Age groups distribution at different administrative levels in the year 2020. Source: Federal Statistical Office (2021)

Here it can be seen that the ratios of the different age groups of all administrative levels are very similar to the ratios found for the UPSs. This makes sense, since Swisscom AG did not actually collect the demographic data, but merely simulated it by using the age and gender distributions of the respective home municipalities of the SIM card holders. It is striking that the proportion of young adults (20 - 39) that live in the city of Zurich is considerably higher than in the canton or in Switzerland as a whole (Figure 11). Within the city of Zurich, this share is the largest of all age groups. However, this overrepresentation is not found in the age groups identified for the UPSs in the city of Zurich. Instead, the distribution of age groups identified for the UPSs is more similar to that of the Canton of Zurich. This could indicate that many of the people who stayed in the UPSs were not registered in Zurich, but in another municipality.

With the binary gender classification into male and female visitors, more data is available for the different UPSs than is the case with the age classification. Of the permanent resident population of the city of Zurich in 2021, 49.9% were registered as female and 50.1% as male (Stadt Zürich Präsidialdepartement, 2022). Therefore, the average male proportion of each

UPS was expected to be around 50%. With values between 49% and 50%, the results of all UPSs come close to the expected male proportion. Thus, demographic figures on the gender distribution of people who stayed in a square or park also appear to be relatively reliable to calculate, despite various assumptions and simplifications. However, as with the age groups, there is not much variance among the different UPSs in terms of gender differences. The UPSs with the highest percentage of women (50.3%) are Quadroplatz and Stadthausanlage, while Weinplatz has the highest percentage of men (51.9%). In general, there seems to be a slight but probably negligible tendency towards female visitors, which again could be due to the many simplifying assumptions made in the calculation of these numbers. With the exception of 6 UPSs, a male proportion of slightly less than 50% was calculated for all UPSs. While both the age group and gender distributions are reasonable distributions, the question is whether or not the simulation nature of these parameters produces results that can be used for a true evaluation of the use of the various UPSs. Since this question could not be answered with 100% certainty before more information on these parameters is explored, I decided not to use the demographic data for this work to analyze the UPSs.

## 4.3 Characteristics of Urban Public Spaces

To determine if there are certain physical features that make UPSs more popular than others, I compared which POIs are located in their immediate vicinity and could therefore make them seem more attractive. In addition, I also compared other features that might enhance the quality of a UPS. A complete overview of all characteristics used can be found in the Section 3.2.3. The following Section illustrates the results related to these identified characteristics of UPSs.

### 4.3.1 Significance

Figure 12 shows the significance class that the city of Zurich assigned to the 34 UPSs. From the total of four categories used by the city for this classification, the visitor density of at least one UPS in each were found to be statistically significant. The importance of the Quartierpark Schütze-Areal is rated lowest from the point of view of the geographical range for which it is considered significant. Thus, according to the city, this UPS is mostly important for its own neighborhood. District-wide significance is attributed to the five UPSs Hardturmplatz, Quadroplatz, Unterdorfplatz, Schöneggplatz and Park Pfingstweid. With 15 associated UPSs, the largest group is represented by those places that have a citywide significance. Finally, 13 UPSs are considered to have international significance.

The UPSs that belong to this last group are all located around the lake basin or in the immediate vicinity of Zurich's main train station. One is also located in the immediate vicinity of the Oerlikon train station, which is also an important hub in Zurich's public transport network. For the other groups, such a spatial pattern is not immediately apparent.

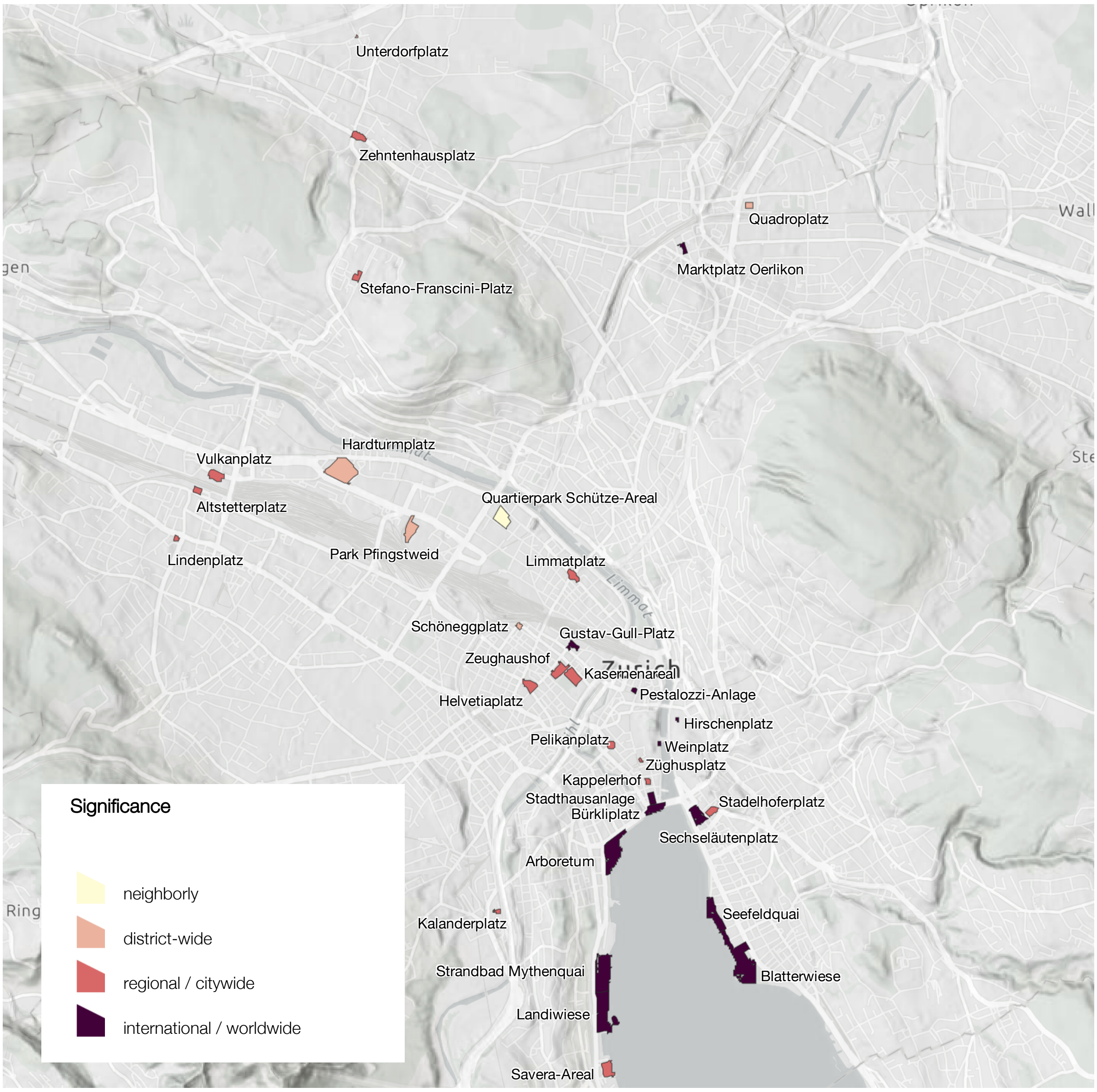



**Figure 12**: The significance assigned by the city of Zurich to the UPSs under consideration (Data: Open Data Stadt Zürich (2021))

### 4.3.2 Points of Interest

The points of interest I used to quantify some of the characteristics of different UPSs I included in the analysis via different buffers (30 and 60 seconds walk time). Examples of the data and the buffers used can be seen in Figure 13. A selection of interesting examples of UPSs is included. Table 7 moreover shows the calculated hectare values of a selection of POIs that are directly related to UPSs and their size. Values that were assumed to have a positive impact on the attractiveness of a UPS are highlighted in green, while values that could have a negative impact are highlighted in white. In all cases, except for the damage reports via the "Züri wie neu" app, the highest values are those that I classified as the best. It is noticeable that many damage reports were received, especially at Vulkanplatz and Zehntenhausplatz. Quadroplatz, Unterdorfplatz and Kalanderplatz did not receive any damage reports at all. However, for its size, Vulkanplatz also has the most fountains (54.5/ha). No other UPS has anywhere near as high a density of fountains. Vulkanplatz also ranks first for trees, followed by Zehntenhausplatz. Hirschenplatz clearly has the highest density of shops, with 275.1 per hectare.

**Table 7**: Density per hectare of POIs of different categories (green = best values, white = worst values)

| Name | Züri Wie Neu | Fountains | Trees | Shops | Bars | Restaurants | Two-Wheel Parking | ZüriVelo | Public Transport Stops |
|---|---|---|---|---|---|---|---|---|---|
| Altstetterplatz | 66.2 | 6 | 15.1 | 12 | 0 | 9 | 9 | 3 | 6 |
| Arboretum | 10.4 | 0.3 | 99.5 | 0.3 | 0.3 | 0.3 | 1.4 | 0.3 | 0.3 |
| Blatterwiese | 13.8 | 0.7 | 35.5 | 0 | 0 | 0.9 | 0 | 0.4 | 0.2 |
| Bürkliplatz | 44.9 | 0 | 23.8 | 0.9 | 0 | 0.9 | 0 | 0.9 | 1.8 |
| Gustav-Gull-Platz | 16 | 2 | 6 | 17.9 | 0 | 6 | 8 | 2 | 0 |
| Hardturmplatz | 3.6 | 0.3 | 0 | 0 | 0 | 0 | 0.3 | 0.3 | 0.5 |
| Helvetiaplatz | 43 | 1.2 | 20.9 | 7.4 | 4.9 | 7.4 | 7.4 | 1.2 | 1.2 |
| Hirschenplatz | 42.3 | 10.6 | 0 | 275.1 | 95.2 | 116.4 | 0 | 0 | 0 |
| Kalanderplatz | 0 | 0 | 0 | 120.1 | 5.2 | 26.1 | 5.2 | 0 | 5.2 |
| Kappelerhof | 4.6 | 0 | 27.6 | 101 | 0 | 4.6 | 13.8 | 0 | 4.6 |
| Kasernenareal | 4.7 | 0 | 33.7 | 0 | 0.8 | 3.1 | 0.8 | 0 | 0 |
| Landiwiese | 2 | 0 | 20.2 | 0 | 0 | 0 | 0.3 | 0 | 0.3 |
| Limmatplatz | 61.5 | 1.5 | 15.4 | 27.7 | 3.1 | 7.7 | 10.8 | 1.5 | 1.5 |
| Lindenplatz | 36.7 | 9.2 | 32.1 | 64.3 | 0 | 9.2 | 9.2 | 4.6 | 4.6 |
| Marktplatz Oerlikon | 23.4 | 5.8 | 43.9 | 90.7 | 0 | 8.8 | 2.9 | 2.9 | 0 |
| Park Pfingstweid | 18.1 | 2 | 94.7 | 0 | 1.3 | 0.7 | 4 | 0 | 0 |
| Pelikanplatz | 23.1 | 2.6 | 30.8 | 41 | 15.4 | 10.3 | 0 | 2.6 | 0 |
| Pestalozzi-Anlage | 57.7 | 6.4 | 32.1 | 96.2 | 6.4 | 12.8 | 6.4 | 0 | 0 |
| Quadroplatz | 0 | 0 | 0 | 3.1 | 0 | 9.3 | 6.2 | 0 | 0 |
| Quartierpark Schütze-Areal | 9.4 | 0.6 | 87.9 | 1.3 | 0 | 1.3 | 7.5 | 0.6 | 0 |
| Savera-Areal | 10.3 | 0 | 25.9 | 0 | 0 | 0 | 0.9 | 0.9 | 0.9 |
| Schöneggplatz | 58.8 | 4.9 | 29.4 | 4.9 | 19.6 | 14.7 | 14.7 | 0 | 0 |
| Sechseläutenplatz | 13.7 | 1.3 | 39.3 | 9.2 | 2.6 | 5.2 | 1.3 | 0 | 0.7 |
| Seefeldquai | 6.6 | 0.5 | 66 | 0 | 0.5 | 0.9 | 0 | 0 | 0 |
| Stadelhoferplatz | 63.6 | 3.9 | 27 | 48.2 | 5.8 | 15.4 | 1.9 | 0 | 3.9 |
| Stadthausanlage | 12.3 | 1.5 | 124.3 | 4.6 | 4.6 | 1.5 | 1.5 | 0 | 1.5 |
| Stefano-Franscini-Platz | 2.2 | 0 | 0 | 0 | 2.2 | 0 | 2.2 | 2.2 | 0 |
| Strandbad Mythenquai | 1.9 | 0 | 36.9 | 0 | 0 | 0.3 | 0.8 | 0 | 0.3 |
| Unterdorfplatz | 0 | 0 | 0 | 10.2 | 0 | 0 | 0 | 0 | 0 |
| Vulkanplatz | 218.2 | 54.5 | 1554.5 | 81.8 | 0 | 54.5 | 109.1 | 27.3 | 27.3 |
| Weinplatz | 2.1 | 1.1 | 2.1 | 41.5 | 0 | 2.1 | 0 | 0 | 1.1 |
| Zehntenhausplatz | 226.9 | 10.3 | 196 | 103.2 | 0 | 20.6 | 10.3 | 10.3 | 10.3 |
| Zeughaushof | 13.2 | 0 | 29.3 | 4.4 | 7.3 | 14.7 | 4.4 | 0 | 0 |
| Züghusplatz | 7.4 | 0.9 | 3.7 | 18.5 | 0.9 | 4.6 | 0 | 0 | 0 |

Hirschenplatz, Schöneggplatz and Pelikanplatz have the widest range of different bars. In terms of restaurants, Hirschenplatz also offers many options, as do Vulkanplatz and Kalanderplatz. In terms of the POIs that can be credited for the accessibility category, the two UPSs Vulkanplatz and Zehntenhausplatz have clear advantages over the other UPSs in all three categories "two-wheel parking", "ZüriVelo" and "public transport stops".
Noteworthy in this list are the figures for Quadroplatz. There is some data for it, but that is only the shops, restaurants and two-wheel parking. These are datasets that come at least in part from OSM (shops and restaurants) and were integrated via the buffer geometry and not directly via UPS geometry. Of the data points that are recorded and managed by the city of Zurich, such as the trees and fountains, none seem to be located at Quadroplatz at all. This is most likely not due to the fact that there are no trees in Quadroplatz, but that Quadroplatz, while open to the public, is managed privately and not by the city. For this reason, trees and fountains are not maintained by the city and therefore are not included in their dataset.

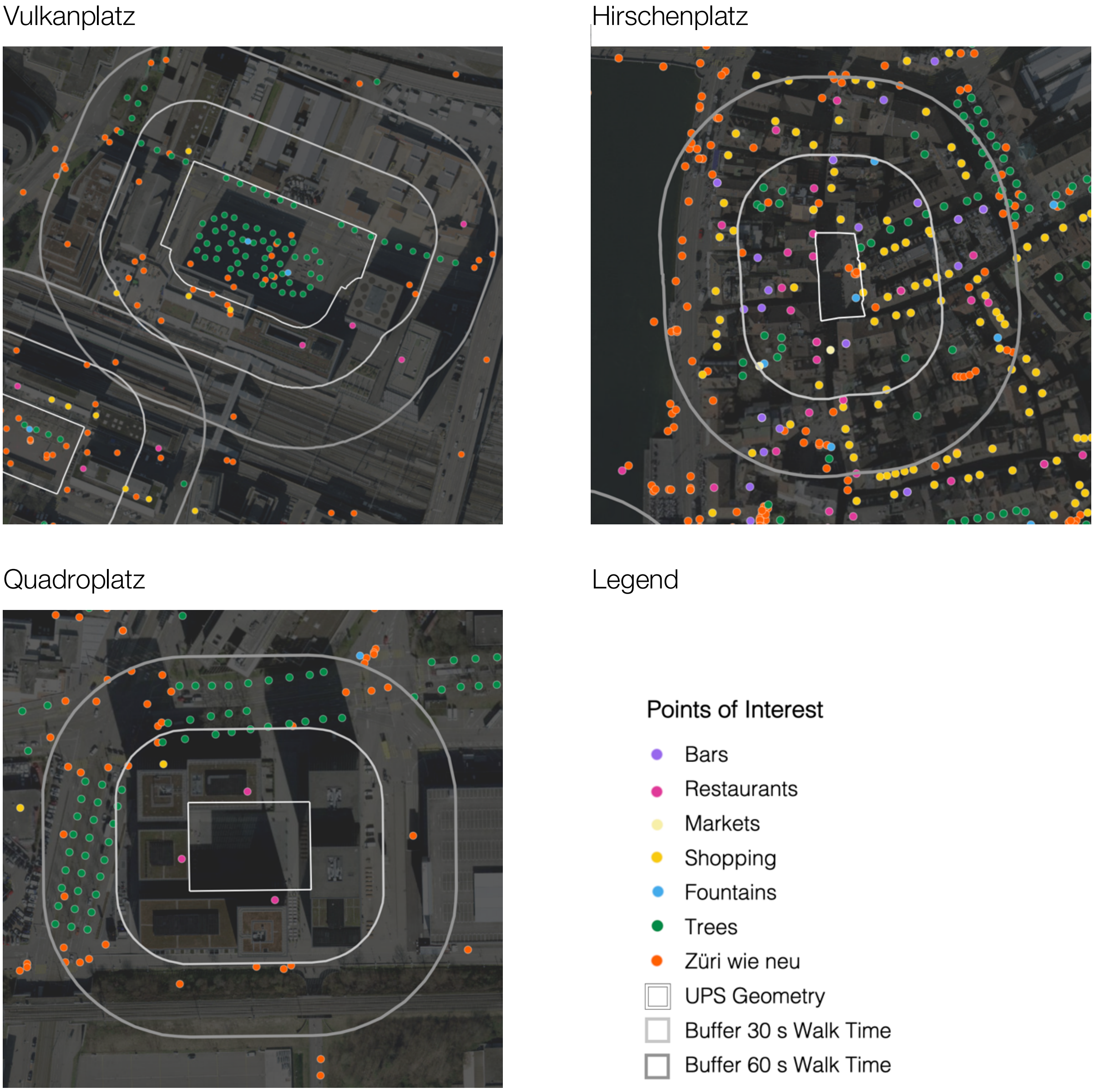



**Figure 13**: Samples of different UPSs, the POIs and the buffers used for the intersection around the UPSs geometries (Data: OpenStreetMap contributors & Open Data Stadt Zürich (2021))

## 4.3.3 Accessibility

### 4.3.3.1 Centrality

To learn about the accessibility of the different UPSs for pedestrians and cyclists, I calculated the local closeness centrality based on a radius of 1200 meters (see chapter 3.3.1.2). Figure 14 shows the calculated centrality index values for the pedestrian network nodes.

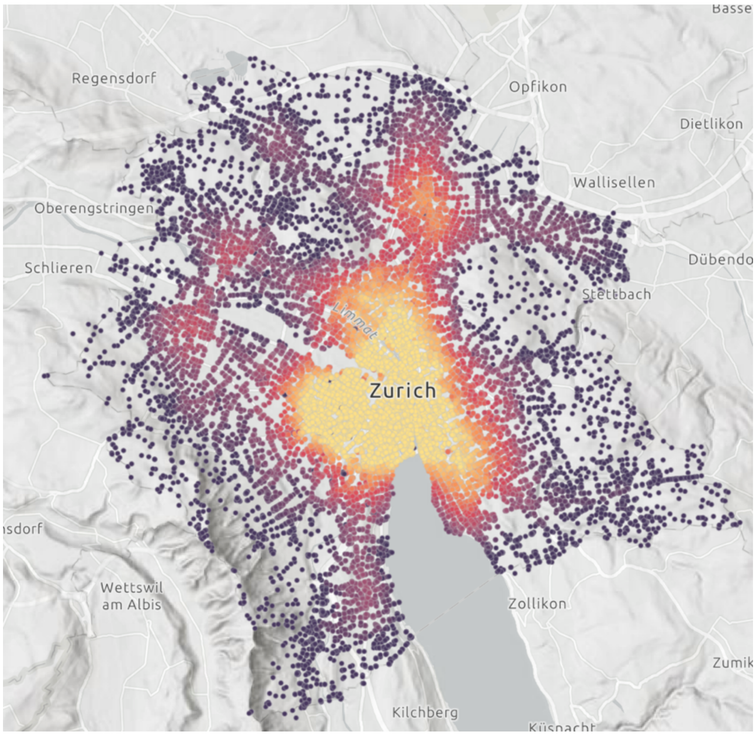



**Figure 14**: Calculated local closeness centrality for pedestrians for a radius of 1200 meters (yellow: high local closeness centrality index, dark purple: low local closeness centrality index)

As expected, the highest values are found in the area of the city center. But also in the northeast of the city, around the Oerlikon train station, there is an increased accessibility of network nodes for pedestrians. Accessibility also appears to be somewhat higher in the northwest of the city, in the Altstetten and Hönggerberg region. This pattern can also be seen for closeness centrality in the bicycle network.

As a reference for the reachability of each UPS, I took the highest calculated value of the network nodes located in the immediate vicinity of the UPS. I did this for both pedestrian and bicycle accessibility. Figure 15 displays a ranking of UPSs according to their different accessibility. Here, it can be seen that the UPSs' closeness centrality ratings for pedestrians and cyclists are similar, but still show some discrepancies, especially for the UPSs in the middle range. The bicycle accessibility to the various UPSs generally appears to be lower than pedestrian accessibility, but there is less variation between sites.

Pelikanplatz is the most accessible UPS by foot, while Unterdorfplatz is the least accessible. After Pelikanplatz, the most accessible UPSs by foot are Pestalozzi-Anlage and Weinplatz. The most accessible by bike is Helvetiaplatz, while the most difficult to reach is Landiwiese. Pelikanplatz and Sechseläutenplatz the next most accessible.

No bike centrality index could be assigned to Weinplatz and Hirschenplatz, which means that the next network node of the bicycle network is too far away (= more than 30 seconds walk time) from them and therefore was not included in the analysis. Therefore, and because the pedestrian centrality index shows many similarities with the bike centrality index anyway, I decided not to use the bike centrality in the final analysis and to approximate centrality only by the pedestrian centrality.

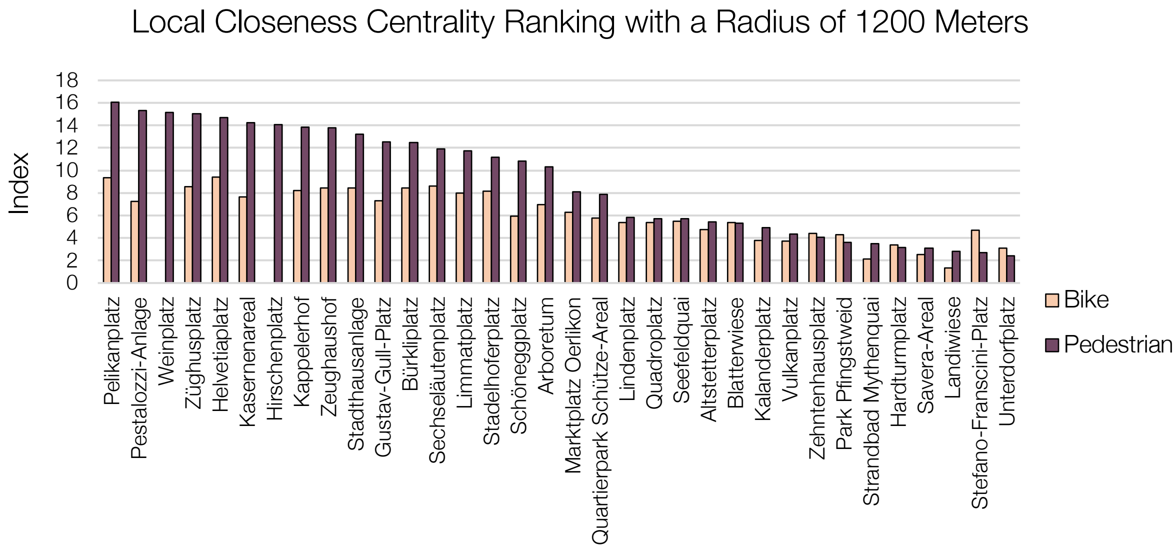



**Figure 15**: Accessibility by bicycle or on foot, approximated through the calculation of the local closeness centrality (radius 1200 meters) of network nodes

#### 4.3.3.2 Traffic and Car Parking

Since traffic counting stations have been installed in the city only at certain locations, it was not possible for me to estimate the traffic directly at each UPS. Instead, I constructed Thiessen Polygons from the existing traffic data to estimate the traffic in the neighborhoods of the UPSs. The results of this calculation can be seen in Figure 16.

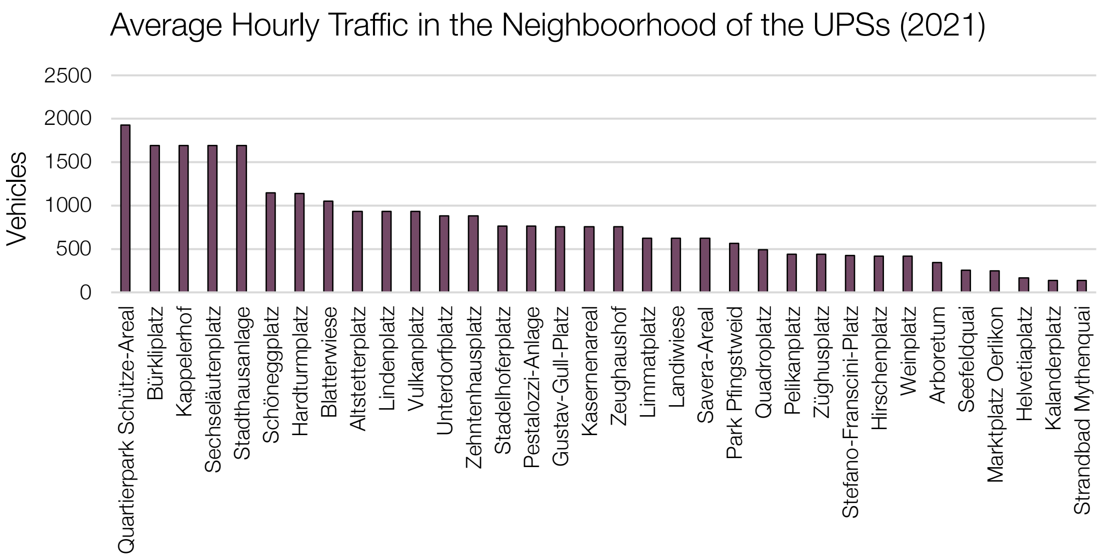



**Figure 16**: Averaged hourly traffic counts in the year 2021 in the neighborhoods (Thiessen Polygons) around the studied UPSs

With an hourly average of 1931 vehicles in 2021, Quartierpark Schütze-Areal is clearly in the busiest area. The Strandbad Mythenquai, on the other hand, is located in a low-traffic area. With an hourly average of 1692 vehicles, Kappelerhof and Sechseläutenplatz, along with Bürkliplatz, rank second among UPSs at busy locations in the city. The three squares have exactly the same traffic counts because they are spatially relatively close to each other and thus fall within the same calculated Thiessen Polygon. The same is true for Altstetterplatz, Lindenplatz and Vulkanplatz (932 vehicles/hour), which are located around Altstetten train station. Similarly, the Kasernenareal, Gustav-Gull-Platz and the Zeughaushof (753 vehicles/hour) as well as the Landiwiese and the Savera-Areal (622 vehicles/hour) also have the same traffic figures. Finally, this also applies to the two UPSs on the edge of the city, Kalanderplatz and Strandbad Mythenquai (136 vehicles/hour).
If there are people who would like to use a private car to get to a UPS, it would also be important that it can be parked in the immediate vicinity. The total number of available parking spaces, both on-street and in parking garages, in close proximity (60 seconds walking time or 80 m distance) to the UPSs is shown in Figure 17.

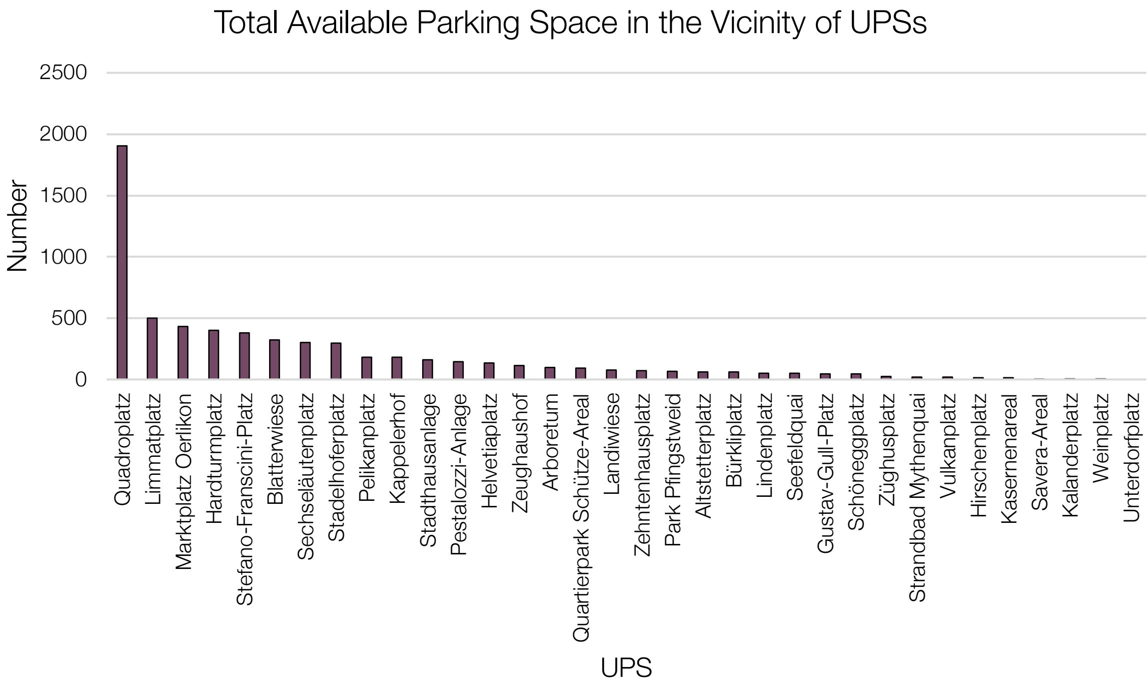



**Figure 17**: Total available parking, both on-street and in parking garages within a 60-second walk from the UPSs

Here, Quadroplatz is clearly in first place with 1907 available parking spaces. This is the case because the parking garage for the large Messe Zürich congress center is located a few meters away. This is followed by Limmatplatz with 498 and Marktplatz Oerlikon with 434 parking spaces. The fewest parking spaces are available around the Savera-Areal, Kalanderplatz, Weinplatz and Unterdorfplatz.

#### 4.3.3.3 Public transport

Since public transport plays an important role in Swiss daily life and policy, I used datasets on accessibility by public transport in addition to measurements of individual traffic (traffic counts, parking spaces, pedestrian and bicycle accessibility) to evaluate the UPSs. These were a polygon dataset showing the public transport quality class of locations and a point dataset of public transport stops. The datasets are interrelated to some extent, since the public transport quality classes depend, among other things, on the locations of the stops (for more information see Section 3.2.3.2). Accordingly, the results of the evaluation of squares and parks here are also similar for both datasets. There is not much variance between the accessibility of the different UPSs by public transport. 30 of the 34 UPSs were classified in public transport quality class "A", i.e., the best possible in the canton of Zurich. 3 were classified in the next best class "B". These are the UPSs Blatterwiese, Seefeldquai, and Strandbad Mythenquai. Only one UPS was classified in category "C", namely Unterdorfplatz, which is located at the city boundary.

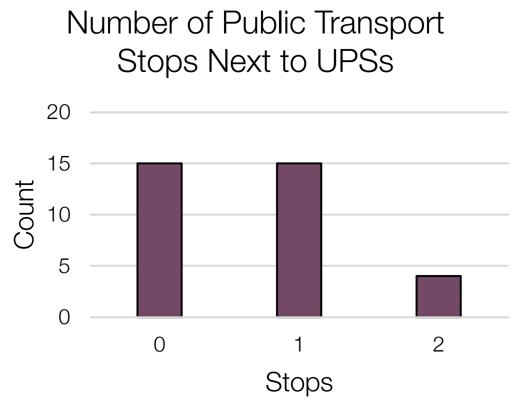



**Figure 18**: Absolute number of public transport stops in the immediate vicinity of the 34 UPSs

Still, for 15 UPSs, the nearest public transport stop is more than 30s walk time (= 40 m) away from their outer boundary (see Figure 18). For another 15, 1 stop is within this distance. In the case of 4 UPSs, there are even two stops directly adjacent to them. These are Altstetterplatz, Bürkliplatz, Hardturmplatz and Stadelhoferplatz.
When comparing the number of public transport stops per hectare, Vulkanplatz is far at the top (27.3 stops/hectare). It is followed by Zehntenhausplatz (10.3 stops/hectare)

## 4.4 Modelling Popularity Based on UPS Characteristics

### 4.4.1 Quality Categories

#### 4.4.1.1 Quality Scores of UPSs

Figure 19 shows the calculated quality index for the three quality categories of UPSs used for this work. The order from left to right corresponds to the ranking (descending) of the UPSs according to their achieved total quality score. Here it can be seen that Hirschenplatz performs best of all 34 UPSs when all three categories are considered together. With a value of 0.75, it is also the UPS that scores best in the "Uses & Activities" category, i.e., the presence of shops, bars, restaurants and markets. In the "Uses & Activities" quality category Hirschenplatz is followed by Marktplatz Oerlikon and Lindenplatz. In the "Comfort & Image" category, Vulkanplatz comes out far ahead with a total score of 0.68. It is followed by the Stadthausanlage (0.51) and Marktplatz Oerlikon (0.5). In the quality category "Access & Linkages", Vulkanplatz scores best with 0.62, followed by Kappelerhof and Stadthausanlage.

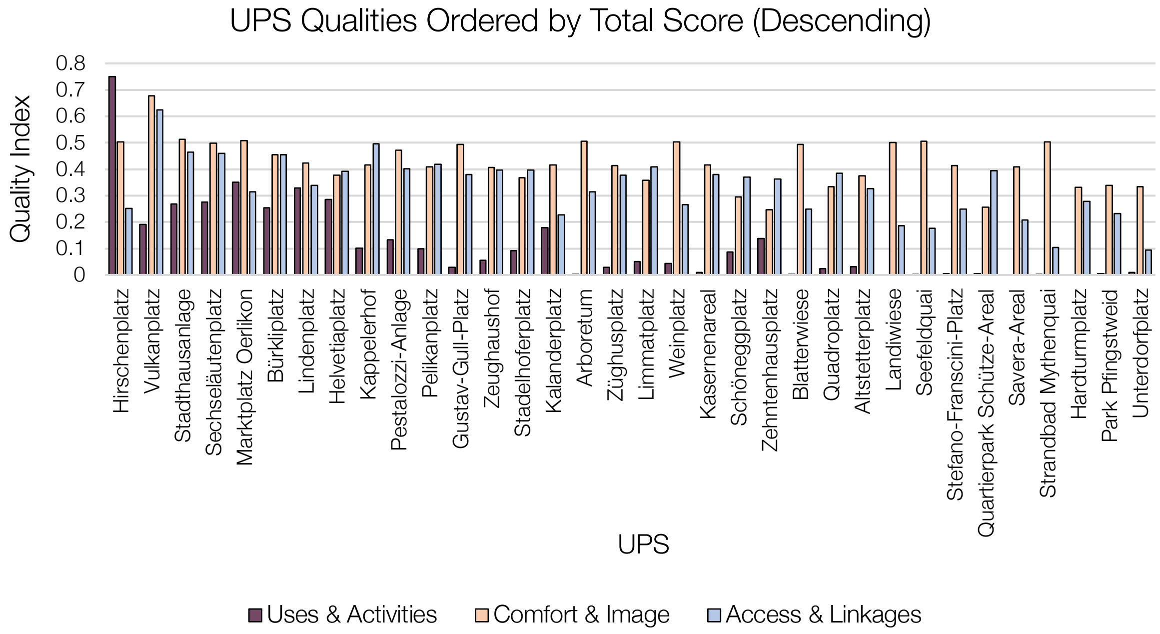



**Figure 19**: Index of quality for the three quality categories of an UPS, ordered by total score (descending)

UPSs that received a low total quality score have almost no POIs that fall into the "Uses & Activities" category. These include the UPSs of Unterdorfplatz, Seefeldquai, Park Pfingstweid Hardturmplatz and Strandbad Mythenquai.

Of the UPSs that are in the top 5 in the calculated quality score, none are also in the top 5 in terms of average weekday visitor density per hectare. In contrast, 3 of the 5 UPS that received the worst quality score also have the lowest density of visitors on working days. Thus, in purely qualitative terms, certain correspondences between UPS quality and visitor numbers can be determined here. However, this cannot be proven quantitatively, as shown in Figure 20, which compares the total quality score and the average daily visitor density per hectare over the entire week, including weekends.

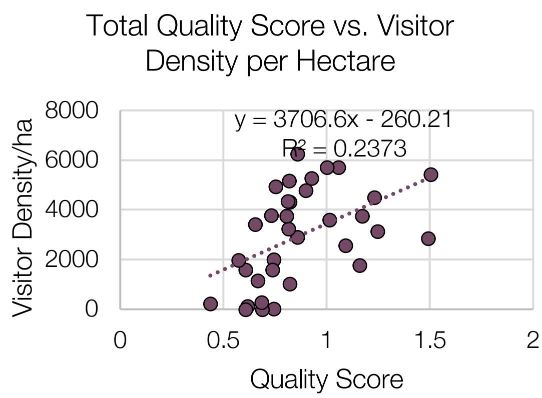



**Figure 20**: Correlation between the calculated total quality score and average daily visitor density per hectare of each UPS

#### 4.4.1.2 Popularity on Weekdays

At the selected significance level of $\alpha < 0.05$, statistical significance was found for the weekdays for the quality categories "Access & Linkages" and "Uses & Activities" (see Table 8). Especially the category "Access & Linkages" has a relatively strong influence on the number of visitors with a coefficient of 1034. However, the calculated R-squared and the adjusted R-squared of the model are not very high. The question therefore arises as to whether it makes sense at all to search for correlations between the quality categories and the visitor numbers with a linear model and how relevant the correlations found actually are.

**Table 8**: Results of the multiple linear regression analysis for the three quality indices on weekdays

| R-Squared | | Adjusted R-Squared | | F-Statistic | | |
|---|---|---|---|---|---|---|
| 0.419 | | 0.361 | | 7.220 | | |
| Variable | Coefficient | Standard Error | t | P > \|t\| | [0.025 | 0.975] |
| Constant | 3216.0000 | 281.131 | 11.439 | 0.000 | 2641.853 | 3790.147 |
| Uses & Activities | 647.5057 | 299.159 | 2.164 | 0.039 | 36.542 | 1258.469 |
| Comfort & Image | -396.0300 | 293.258 | -1.350 | 0.187 | -994.943 | 202.883 |
| Access & Linkages | 1033.9467 | 289.476 | 3.572 | 0.001 | 442.758 | 1625.135 |

#### 4.4.1.3 Popularity on Weekends

As on weekdays, the quality categories "Access & Linkages" and "Uses & Activities" proved to be statistically significant for visitor density of UPSs also on weekends. Here, however, the "Uses & Activities" category seems to be much more important compared to the "Access & Linkages" category than was the case on weekdays. In addition, the "Comfort & Image" quality category is also statistically significant for weekends. Here, however, a higher score in this category leads to a lower visitor density per hectare. With an adjusted R-square of 0.409, this model also raises the question of its relevance.

**Table 9**: Results of the multiple linear regression analysis for the three quality indices on the weekend

| R-Squared | | Adjusted R-Squared | | F-Statistic | | |
|---|---|---|---|---|---|---|
| 0.463 | | 0.409 | | 8.627 | | |
| Variable | Coefficient | Standard Error | t | P > \|t\| | [0.025 | 0.975] |
| Constant | 2362.7941 | 220.606 | 10.710 | 0.000 | 1912.257 | 2813.331 |
| Uses & Activities | 897.3676 | 234.752 | 3.823 | 0.001 | 417.940 | 1376.795 |
| Comfort & Image | -474.9601 | 230.122 | -2.064 | 0.048 | -944.931 | -4.989 |
| Access & Linkages | 549.4144 | 227.154 | 2.419 | 0.022 | 85.505 | 1013.324 |

## 4.4.2 Individual Attributes

### 4.4.2.1 Correlations between Attributes

Aside from the two public transportation datasets and the pedestrian and bicycle centrality measures, attributes were chosen under the assumption that they were independent. Nevertheless, it is of course possible that some of them correlate with each other, especially in environments heavily modified by humans. In order to detect such cases and not include them later in the multiple linear regression analysis, I performed a purely qualitative correlation analysis. For this I created a correlation matrix, which is displayed in Figure 21.

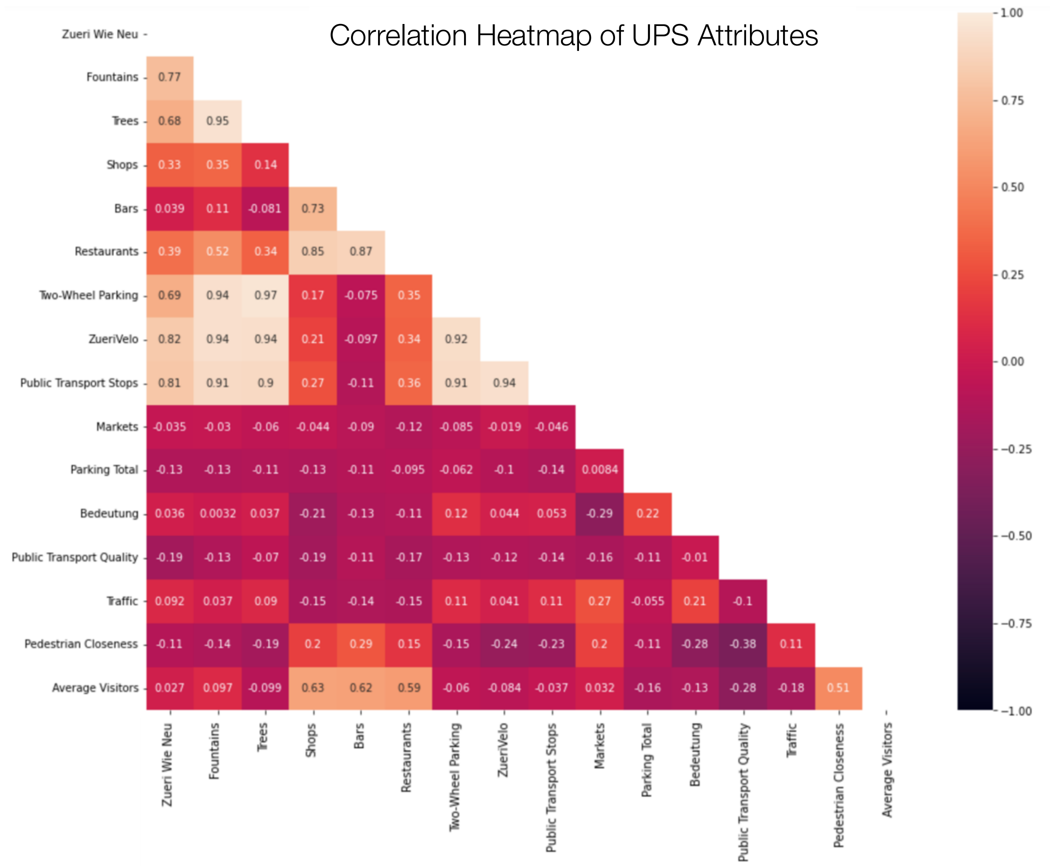



**Figure 21**: Correlation matrix of UPS attributes

Here it can be seen that indeed many of the attributes are correlated among themselves to some degree. I decided that I did not want to include variables in the MLR that had a correlation value of > 0.75, and therefore were highly correlated.

In particular, the three attributes "Two-Wheel Parking", "ZüriVelo" and "Public Transport Stops" correlate very strongly with several other attributes. Therefore, I decided not to include them in the regression analysis. In addition, the fountains and trees also correlate relatively strongly (0.95). Thus, I used only the trees. Finally, restaurants, shops and bars also correlate relatively strongly with each other. Here I decided to use only shops and bars for the MLR.

#### 4.4.2.2 Popularity on Weekdays

The result of the MLR for the weekdays shows that only three attributes are statistically significant. This can be seen in Table 10.

**Table 10**: Results of the multiple linear regression analysis for all (preselected) variables during weekdays

| R-Squared | | Adjusted R-Squared | | | F-Statistic | |
|---|---|---|---|---|---|---|
| 0.695 | | 0.665 | | | 22.79 | |
| Variable | Coefficient | Standard Error | t | P > \|t\| | [0.025 | 0.975] |
| Constant | 3216.0000 | 203.731 | 15.786 | 0.000 | 2799.927 | 3632.073 |
| Shops | 626.9690 | 210.603 | 2.977 | 0.006 | 196.859 | 1057.079 |
| Significance | -589.2996 | 215.058 | -2.740 | 0.010 | -150.093 | -1028.506 |
| Pedestrian Centrality | 1541.1569 | 214.755 | 7.176 | 0.000 | 1102.568 | 1979.746 |

These are the shops, the significance of the UPSs and the pedestrian centrality. Here, accessibility on foot has by far the greatest influence on visitor density. It is 2.5 times more important than the presence of shops and the significance of UPSs. In contrast to the aggregate quality scores, this model appears to be more reliable, with an adjusted R-squared of 0.665.

#### 4.4.2.3 Popularity on Weekends

Table 11 shows the MLR results for the weekend.

**Table 11**: Results of the multiple linear regression analysis for all (preselected) variables on the weekend

| R-Squared | | Adjusted R-Squared | | | F-Statistic | |
|---|---|---|---|---|---|---|
| 0.646 | | 0.610 | | | 18.22 | |
| Variable | Coefficient | Standard Error | t | P > \|t\| | [0.025 | 0.975] |
| Constant | 2362.7941 | 179.228 | 13.183 | 0.000 | 1996.762 | 2728.826 |
| Shops | 598.3251 | 185.274 | 3.229 | 0.003 | 219.945 | 976.705 |
| Significance | -420.5985 | 189.193 | -2.223 | 0.034 | -34.216 | -806.981 |
| Pedestrian Centrality | 1149.9280 | 188.926 | 6.087 | 0.000 | 764.089 | 1535.767 |

The same three attributes were also found to be statistically significant on weekends. Here, too, accessibility in the form of pedestrian centrality plays a role. On weekends, pedestrian accessibility also plays the largest role of the three attributes but has a slightly lower importance compared to that of shops than on weekdays. This is the case because the importance of the presence of shops on weekends seems to increase. While the average number of visitors to all UPSs decreases by 26% on weekends compared to weekdays, the coefficient of shops on weekends compared to weekdays only decreases by 4%. The decline in contributions to visitor density on UPSs of significance as well as pedestrian centrality, however, is roughly in line with the general decline in weekend visitors.

# Chapter 5 | Discussion

After the individual results of this work were presented in the previous chapter, they will be critically examined in this chapter. In addition, a connection is drawn to the background of the research topic, which was presented in chapter 2, and the answers found to the research questions posed at the beginning are presented. $RQ_1$ cannot be answered completely based on the methodology used in this thesis, whereas $RQ_2$ can. In addition, the limitations and uncertainties of applying the methods used are also discussed and what the findings imply for the future of urban planning and research in the field.

## 5.1 Reflections on the Results

### 5.1.1 Attributes that Increase the Popularity of UPSs

In the more practical part of my research question, I wanted to find out if it is possible, based on data analysis, to identify physical characteristics of a UPS that increase its popularity among the population. In particular, I also wanted to find out if there are very specific attributes that can be targeted to make urban places and parks more attractive to the audiences of women, children, and the elderly. The research question posed was therefore as follows:

**$RQ_1$: Are there specific attributes of an urban public space that increase its popularity?**

- What physical characteristics can be identified that show a correlation with an increased frequency of use?
- What can be done to make newly planned UPSs attractive for the whole community, but especially for women, children, and the elderly, and to invite them to spend time there?

#### 5.1.1.1 UPS Characteristics

The results show that to some extent it is indeed possible to determine attributes that make UPSs more attractive than others. For both weekdays and weekends, features belonging to the calculated quality categories “Uses & Activities” and "Access & Linkages" were attributed a statistically significant influence on the visitor density of different UPSs. Additionally, on weekends, the “Comfort & Image” quality category was also found to be statistically significant. However, it must be noted here that for the linear models created for the quality categories, R-Squares and adjusted R-Squared are not particularly high (< 0.5). This raises the question of the extent to which the correlations found can really represent the relationship between the quality categories and the visitor density.
Zooming from the summarized quality categories to individual attributes, we find that the models found here are more stable than those built based on the quality categories (adjusted

R-squares > 0.6). One attribute from the category "Uses & Activities" is statistically significant for weekdays and weekends. This is the density of shops in the immediate vicinity of UPSs. Here, too, we can observe what was already observed for the aggregated quality categories: on weekends, the shops have a greater influence on the observed visitor density in direct comparison with the other attributes. The fact that shops attract more people on weekends who spend a longer time on site makes sense. Many people don't have to go to work on the weekend and therefore have more time to linger. However, it is interesting to note that on weekends, most shops in the city of Zurich are only open on Saturday. On Sunday, only a few grocery stores are open, especially near train stations. This means that if the shops really do play such a large role in visitor density at the weekend, then this would have to be the case on Saturday in particular. Restaurants that are typically open on Sunday, on the other hand, were not included in the MLR because their density correlated too strongly with the density of shops and bars. In a future continuation of the research on UPSs, it would therefore be of interest to also examine the relationship between restaurant and visitor density in more detail. One attribute from the "Access & Linkages" category was also found to be statistically significant for weekdays and weekends. This is the pedestrian centrality of the UPSs within a radius of 1200 meters. The 1200 meters here corresponds to the approximate average distance from the center of all districts of the city of Zurich to their boundaries. Thus, the centrality of a UPS in its neighborhood also plays a large role in its popularity. This seems plausible, since cities are defined by their short distances to various facilities and activities compared to more rural areas. People do not want to travel far to reach their leisure activities. Furthermore, following Frank's (2019) argumentation UPSs becoming more and more like our extended living room, accessibility on foot is strongly relevant. In addition, this finding also fits with the principles set out in several frameworks on social sustainability and liveability, which emphasize that it is particularly important that the public realm be designed in a pedestrian-oriented manner (e.g., Wheeler, 2001; WAPC, 2009).

The attribute of significance of a UPS, which belongs to the quality category "Comfort & Image", was another attribute which proved to be statistically significant in determining visitor density for both weekdays and weekends. It is interesting to note that a higher significance from a spatial perspective (e.g., international) resulted in fewer visitors. This is probably due to the fact that a large proportion of the UPSs classified as internationally important are to be found around the lake basin. They are also the ones with the lowest daily visitor densities per hectare. Thus, it seems that the significance assigned by the city of Zurich to the UPSs under consideration is not made dependent on the number of visitors and that the significance classes are really to be understood in purely qualitative terms and an international significance does not mean at the same time that a high density of both of local and international guests can be expected. But the negative correlation with significance may also be strongly related to assumptions made for simplification. In order to be able to use more data and because the two types of public spaces have also been analyzed together in previous studies (Grimm-Pretner, 2004), I included both squares and parks in the city of Zurich in the category of UPSs. In the process, these two variants of UPSs were processed, evaluated, and analyzed in exactly the same way. While this can make sense from a pragmatic point of view and can provide general insights into what makes public space more attractive, the results might be more informative if these two categories were treated separately. After all, they are also used, at least in part, for different activities and purposes. Parks with a lot of green areas and nature are often much larger than squares placed in between streets and buildings. Additionally, attributes such as scenery, views, and the like could be very important, as parks are used more for recreational purposes, sporting activities, or relaxation (Bühler, et al., 2010), and less for transit, or short breaks.

#### 5.1.1.2 Targeting Specific Population Groups

In a second step of this work, I wanted to further find out which attributes of a UPS could be particularly attractive to the groups of women, children and elderly people. Several reasons made it impossible to answer this question using the methods employed. On the one hand, the demographic information that can be obtained from mobile data is merely simulated based on the demographics of the SIM card owners' municipality of origin. The age and gender distributions found for the UPSs studied are therefore also very similar to the general age and gender distribution in the Swiss population and the variance of the results for the different UPSs is rather low. Besides the general nature of the data, this additionally could be a product of the downsampling approximations applied to translate the data from the tiles to the UPSs geometries. With such small differences in the proportions of different age groups or genders of visitors to different UPSs, it would have made it very difficult to find correlations with certain characteristics of UPSs that resonate better with certain populations. In addition, although small children, for example, are represented by the simulation of age groups, their actual number is completely uncertain, since many of them do not yet own cell phones anyway. Therefore, there would be a lot of assumptions involved here. In addition, also the data used to characterize the UPSs do not show much variation in their nature. Since they were partly chosen because of their ease of availability, many of them are probably amenities that would likely be used by almost all population groups anyway. To be able to really say something about which attributes of a UPS are particularly important for certain groups, more diverse attributes would have to be included, such as playgrounds or accessibility for wheelchair users. Including a stronger temporal dimension, on the one hand for attributes such as shops and restaurants (e.g., opening hours), and on the other hand for the number and type of visitors (i.e., hourly visitor density numbers), could also lead to more relevant results.

### 5.1.2 Data-Based Analyses for a Better Understanding of UPSs

The second and more theoretical research question I addressed was the effectiveness of the methods used to answer $RQ_1$. The question was whether a pure data analysis can indeed help to understand real processes and dynamics in urban public space. Additionally, I wanted to know how well the used mobile data can be used to evaluate the number of visitors as a proxy for the popularity of a UPS and if meaningful results can be found. The defined research question relating to this was as follows:

**$RQ_2$: To what extent do the purely data-based methods used in this study allow us to capture and understand the relationships between popularity and characteristics in urban public space?**

- To what extent is the mobile phone data currently available in Switzerland suitable for analyzing number and type of visitors at specific locations in a city?

#### 5.1.2.1 Using Mobile Phone Data to Analyze the Visitor Numbers of UPSs

The mobile phone data used for the analysis is available at any hour over the whole of Switzerland for two years. Therefore, it was easily possible to analyze the visitor densities of all UPSs in the city of Zurich. In addition, this could readily be scaled up, for example to the whole canton, or even the whole of Switzerland. Furthermore, visitor density can be analyzed on a daily basis, as I have done in this work, or hourly information can be used to perform analyses based not only on static but also on dynamic near real-time numbers. Thus, it can be clearly stated that mobile phone data bring entirely new opportunities to work at larger scales than previously possible, both temporally and spatially. This is consistent with previous findings when working with mobile phone data (Wesolowski, et al., 2016).
The insights gained by analyzing the mobile phone data for the various UPSs are very interesting. For many UPSs, the calculated visitor densities per hectare per day seem to be feasible from personal experience. In a future project, it might make sense to randomly measure this directly on site to confirm. However, it can be said with certainty that despite the downsampling of the data from the 100 m x 100 m grid to the UPSs geometries applied in this work, anomalies in the visitor densities clearly stand out. Thus, in the case of Vulkanplatz, which seems to be mainly frequented by people on their way to work during weekdays, it was possible to observe the school vacations in summer. Moreover, two events on single days at Sechseläutenplatz (Weltklasse Zürich) and Helvetiaplatz (Zurich Pride) were evident. The very low visitor density figures for the bathing location Strandbad Mythenquai can probably be attributed to the comparatively poor weather for the months of July, August and September. However, with this UPS in particular, as well as other UPSs considered that are located directly on the lake, the question also arises as to how reliably the mobile phone data combined with the methods employed worked in their case. Many of them have a very low density of visitors per hectare. In fact, the visitor density at Landiwiese and Strandbad Mythenquai is often (on 90% of all days considered) 0 per hectare, which does not seem feasible. It may be much more likely that these and other park-like UPSs located at the edge of the city regularly do not exceed the privacy-related threshold of a daily visitor density of

more than 20 per tile. This could be a limitation when trying to achieve a uniform information picture across all UPSs of a city. It is therefore critical to keep this fact in mind when working with the data. For the analysis carried out here, however, this did not cause any major issues, as it was still possible to get a comparison with the other UPSs and to find out the approximate visitor density patterns, including particularly busy days.

Although privacy of individuals does not seem to be an issue with the data available, Taylor (2016) fundamentally argues that extreme caution should always be used when working with mobile data. This is because it is still possible for certain populations to be grossly over- or under-represented. This is because, for example, young children generally do not yet own smartphones, while certain adults own several, both for personal and business use. Therefore, depending on the study, it is extremely important to obtain more detailed knowledge about the society under investigation. Since this study only used approximate values for visitor density, such details were not considered essential. In particular, if more emphasis is to be placed on demographic information in future studies, it is highly recommended to focus more on the actual distribution of cell phone usage of different population groups.

In general, it can be said that the mobile phone data, which can currently be acquired in a 100 m x 100 m grid from Swisscom AG, is very valuable for estimating the approximate number of visitors to a location in the city, such as UPSs. In contrast, the information on the type of visitors that can be obtained from mobile phone data (demographics) is somewhat more limited in its usefulness for the purpose of this work, as described earlier.

#### 5.1.2.2 Data-Based Methods for the Evaluation of UPSs

Using the data analysis applied in this thesis to find out more about the popularity and connection to characteristics of UPSs, three attributes of urban public spaces were actually found to influence visitor density. That is, with the applied approach it was possible to obtain certain insights into the relationship between the built environment and the average number of people present. Nevertheless, the three attributes found that could influence the attraction that a UPS has on the population are not enough to be of real use in a new planning or redesign of a UPS. It would therefore be important to study the relationship further. More and different attributes could be used for this purpose. For example, not only could trees be used as an indicator of the greenness of a UPS, but in addition lawns and flower beds could be included as information. In this context, the proximity to water features, the view, or the tranquility would also be something that could be considered relevant for the popularity of a UPS. Additionally, more detailed information about some of the POIs could be integrated. For example, to understand the relationship between attractiveness and popularity of UPSs, it would also be of interest to know not only the number of existing shops, but also the type of shops. Grocery stores could have a completely different impact here than clothing stores. The number of public transport stops also provides only a partial insight into how well connected a UPS actually is. At one stop there could be several means of public transport (streetcar, bus, train) running at regular intervals of, for example, 10 minutes, or there could be only one means of transport running once an hour. To some extent, this information could be taken from the public transport quality dataset, which, however, shows relatively little variation within the city of Zurich. Also, as mentioned earlier, more attention could be paid to daily and hourly dynamics in the measured visitor densities. In this work, many numbers were aggregated for simplicity, which could blur certain findings.

All of the aforementioned adjustments could easily be incorporated into the analysis in the future, provided the necessary data are available. Indeed, data analytics enables precisely this kind of incremental expansion of models and continuous increase in overall complexity. These are also the advantages of the pure data analysis over other research methods that have already been employed to evaluate UPSs. In this direct comparison with other methods, however, some drawbacks of pure data analysis also stand out. The proportion of real circumstances, which the model can reproduce, depends very strongly on the condition of the input data. If these are wrong, or incomplete, or their connection and influence is wrongly defined in the model, the result will also not be reliable. In this work, for example, it was noticed that many UPSs around the lake basin could be assigned almost no POIs that fall into the category "Uses & Activities" (shops, bars, restaurants and markets). They also received a very low overall quality score because of this. However, once one has seen the locations, it becomes clear that there are actually many gastronomic offerings there, but often in the form of mobile stalls. Temporary facilities of this kind are in fact very rarely integrated into static datasets. Such discrepancies between the data used and the real conditions can strongly distort the results of a data analysis when evaluating public spaces. In addition, there are uncertainties in the preprocessing and weighting of certain data. Information about traffic in the city of Zurich can only be read from individual measuring points irregularly distributed over the city. In order to get a continuous picture of the whole city, the data must be interpolated in some way. In my analysis I have chosen the Thiessen Polygons. Of course, these do not reveal anything about the traffic directly at the UPSs, but only about the entire district in which they are located. Thus, the question remains whether this information is sufficient to describe the way traffic affects the individual UPS. Furthermore, there is also the question of whether more traffic means that the UPS is more accessible and could therefore attract more people, or whether it is perceived as a nuisance and should therefore be weighted negatively when calculating overall summary scores, as in the "Access & Linkages" quality category.

Finally, there is also the fundamental question of the extent to which it makes sense to use the mere number of certain objects as an indicator of the attractiveness of a place, and to what extent their presence and number is also inversely determined by the number of people who are typically in this place. Madanipour (1999, p. 879) describes this interaction as follows: “Our spatial behaviour, which is defined by and defines the spaces around us, is an integral part of our social existence.” Especially with time, it may be that the environment is constantly adapted to the clientele. Bars can only exist if they have customers. Bicycle parking is only necessary if many bicycles have to be parked regularly. When analyzing data and trying to find correlations between attributes of existing UPSs and their popularity, it is therefore important to keep this in mind and to critically question results and implications for future UPSs.

In conclusion, it can be said that data analytics can be a simple, easily scalable method that can be useful in determining characteristics that future UPSs should exhibit in order to resonate well with the population. Nevertheless, the limitations described above should be considered. It might be useful to combine data analysis with other methods, such as Social Media data that contain some semantic information, interviews, or our own observation for very detailed analyses of UPSs and their context. This finding is consistent with findings from previous approaches to make certain components of urban planning more efficient through data analysis (Steenbruggen, et al., 2015).

## 5.2 Limitations

Since approaches similar to those used in this work have not been as common in the field of urban planning, many of the data preparation methods used tend to be exploratory and no definitive conclusions can be made about their validity and implications. For example, to circumvent the influence of edge effects, it is questionable whether it is necessary and useful to compare the visitor density per hectare of each UPS with the immediate region to determine if UPSs' numbers are statistically significantly different. Also, the method used, which uses an optimized buffer of 13 times the UPS area as a reference, may need to be reconsidered.

While data-based work brings many advantages in terms of scalability in the dimensions of time and space, there are also some disadvantages. On the one hand, of course, the reliability and consistency of data from different providers must always be questioned. Since I mainly worked with two professional data sources in this thesis, the city of Zurich and Swisscom AG, it can be assumed that the data used is of a rather high standard. Nevertheless, errors can sometimes occur in such datasets. In addition, datasets may not be complete, may not fit the analysis, or may not exist at all. For example, the data set that the city of Zurich refers to as "shopping" via Zürich Tourismus apparently consists only of shops that might be of particular importance to tourists. For the analysis applied in this work, the inclusion of only these shops would of course have made little sense and would have skewed the results. Therefore, this dataset was combined with a dataset from OSM. Since OSM is a collaborative project where volunteers contribute geographic information, it cannot be said with complete certainty that this dataset includes all shops in the city of Zurich (Graham, 2011). Furthermore, the city of Zurich only records POIs of trees, fountains and others for their publicly available data, which they manage themselves. UPSs that are therefore managed by other entities, such as private individuals (Quadroplatz) or the canton (Botanical Garden), could not be evaluated in the same way as the other UPSs, even though they are also located within the city of Zurich. Here the questions arose whether such UPSs should be directly excluded in a next project. However, given the current data situation on public places and parks, this effort could prove difficult. Since there is currently no dataset publicly available that shows the locations and geometries of all the squares and parks, I had to create this myself. For this, I was able to access a pre-made list with the names of all the public parks in the city. To the best of my knowledge, however, there is no such list for the squares of the city of Zurich. Therefore, I took advantage of a linguistic and cultural custom and searched the street register for all street names that ended with "Platz" or "Hof". While according to my personal experience, the most important UPSs were all recorded with this method, it is not entirely clear here whether really all of them could be found. In addition, very occasionally places were found that do not represent UPSs, such as the neighborhood street "Sunnige Hof". I was able to locate and rule out this particular example by visually analyzing the geometries of the UPSs. However, especially for larger future analyses of UPSs, consideration should be given to using automatic verification of geometries. If it is assumed that squares and parks often follow certain shapes, and for example are rarely very elongated, the polygon compactness could be calculated, for example.

In general, the work was based on many simplifications and approximations that could be reconsidered in future studies. One of the basic assumptions of the analysis was that the popularity of a UPS can be assessed by the mere presence of people at that location. While this is true to some extent, it still raises the question of whether this measure can be used as the sole indicator of popularity. After all, some of the UPSs that had particularly high daily visitor densities are located in very busy places where many people might be working, shopping, or simply passing by on their daily commute. Although calculating the statistical significance of the number of visitors to these UPSs might indicate that people were actually staying at the UPS, it could be that people simply chose that UPS for practical reasons. It could well be that they simply chose the spatially closest UPS for their lunch break or to meet friends after work for pragmatic reasons. However, real popularity is not merely related to practical reasons, but to emotions. We speak of popularity when something is sincerely liked and enjoyed by many people (Cambridge Dictionary, 2022). Thus, the question arises whether purely quantitative information, as mobile phone data can provide, is sufficient to express these relationships to a location. To better understand the background of these relationships, it might be useful to collect more semantic information. Surveys can be conducted for this purpose, for example, but then of course the advantage of scalability that data analyses bring with them is lost to some extent. In order to preserve this, especially in the initial analysis phases in which very rough general findings are to be obtained, other types of already existing datasets could be used instead. For instance, a study by Koblet and Purves (2020) has shown that text analysis of Social Media posts or other participatory media can provide insights into how people perceive landscapes. Sentiment analyses such as those conducted by Dinkić et al. (2016) could also be relevant in this context. Incorporating such information into the analysis of UPSs, in addition to visitor density information, could potentially lead to a more accurate representation of popularity. In general, it has proven difficult to represent components that influence the more emotional and social part of human behavior in this work. The PPS Place Diagram (Figure 1) actually contains four categories, which as independent dimensions should have an influence on how popular a UPS is among the population. For this work, only the three categories, “Access & Linkages”, “Uses & Activities” and “Comfort & Image”, were used, for which it was relatively easy to highlight and evaluate the different physical qualities of the UPSs. The category "Sociability", which arguably also has a strong influence on which places we feel comfortable in, was not represented as a decision characteristic. For more insight into this component, qualitative approaches should be used to complement the quantitative methods used.

## 5.3 Implications

It has been shown that mobile phone data indeed have great potential for mapping and analyzing the general spatial behavior of the population. This is consistent with findings from numerous other studies (e.g., Grantz et al., 2020; Phithakkitnukoon et al., 2010; Steenbruggen et al., 2015). Mobile phone data therefore also offer the possibility to get to know and analyze the dynamics of a city in a completely new way. In combination with other data, this opens up many possibilities that can be easily scaled in time and space to provide insights into urban life as it happens on UPSs, for example. It was found that a larger number of shops and pedestrian centrality lead to higher visitor densities on UPSs. Therefore, the use of UPSs can be increased if they have a wide range of stores in close proximity and are well integrated into the pedestrian network. The analysis also showed that the significance of squares seemed to have a negative impact on visitor numbers. This is a good example of why the data analysis conducted here, as well as data analysis in general, should always be treated with caution. Input data and weighting play large roles in outcomes. Mobile objects are often not entered into static datasets, and many social phenomena are also poorly represented via analysis of purely quantitative data. It would therefore make sense to include data with semantic information or to conduct additional qualitative interviews or observations. Combined with other methods, data analyses of the kind carried out in this work represent a real opportunity for efficient, sustainable and population-oriented planning of the UPSs of the future.

# Chapter 6 | Conclusion

Urban public space is constantly gaining importance. This is the result of ongoing global urbanization on the one hand and slowly changing trends in lifestyles and work patterns on the other. Furthermore, as issues of sustainability become increasingly important, there is also a growing emphasis on ensuring that cities and urban space are ecological and provide space for social interaction. Public spaces in the city, such as squares and parks, must therefore meet many requirements and accommodate the needs of a wide range of people. For efficient and sustainable urban planning, it is therefore crucial to understand how such UPSs need to be designed so that they are used by the entire population and specifically by certain target groups. In past studies, qualitative small-scale approaches such as surveys were often conducted to get direct feedback on what physical attributes of a UPS resonate well with people. The enormous amounts of data collected today through the adoption of Smart City concepts and the general digitization of our society offer completely new opportunities to explore the issue over much larger scales.

In this work, a fully data-driven approach was taken to find out which attributes make a UPS popular among the population. As an indicator for the popularity of a UPS, the visitor density derived from a mobile phone dataset was used. It has been found that mobile phone data is very promising for tracking human movement patterns and dynamics. For example, events that had taken place on UPSs could be easily identified in the temporal progression of daily visitor densities. Furthermore, a statistically significant correlation with daily visitor densities was found for three characteristics of UPSs. These are the density of shops in the immediate vicinity (30 seconds walk time from the outer boundary) of the UPS, the geographical significance (neighborly/district-wide/citywide/international) assigned to them by the city of Zurich, and the centrality for pedestrians within the neighborhood (radius 1200 meters).

While the mobile data could provide unique insights, the question arises to what extent they are suitable for representing popularity. Thus, in addition to visitor numbers, future work could also integrate semantic information, for example from Social Media posts, to get a direct impression of how the UPSs are really perceived by their users. Moreover, the model used could be improved by including further attributes such as views, tranquility, proximity to water, green spaces and playgrounds. In particular, a focus should be placed on how the social component can be better depicted as an attractiveness criterion. For this it might make sense to carry out qualitative methods such as surveys directly on site. When combined with other methods, data analyses such as those carried out in this work promise to make a real contribution to informing the urban planning of the future.

# Literature


AFV, Amt für Verkehr des Kantons Zürich, 2020. *ÖV-Güteklassen.* Zürich

Ahmadi, F. & Toghyani, S., 2011. The role of urban planning in achieving sustainable urban development. *OIDA International Journal of Sustainable Development,* 2(11), pp. 23-26.

Ali, U. et al., 2020. A data-driven approach to optimize urban scale energy retrofit decisions for residential buildings. *Applied Energy,* Band 267, p. 114861.

Angelidou, M., 2016. Four European smart city strategies. *International Journal of Social Science Studies,* 4(4), p. 18.

Bühler, E., Kaspar, H. & Ostermann, F., 2010. *Sozial nachhaltige Parkanlagen: Forschungsbericht des Nationalen Forschungsprogramms NFP 54" Nachhaltige Siedlungs-und Infrastrukturentwicklung"*

Bahrehdar, A. R., Adams, B. & Purves, R. S., 2020. Streets of London: Using Flickr and OpenStreetMap to build an interactive image of the city. *Computers, Environment and Urban Systems,* Band 84, p. 101524.

Becker, R. A. et al., 2011. A tale of One city: using cellular network Data for urban planning. *IEEE Pervasive Computing,* 10(4), pp. 18-26.

BFS, B. f. S., 2021. *Ständige Wohnbevölkerung nach Alter, Kanton, Bezirk und Gemeinde, 2010-2020.* [Online]. Available at: https://www.bfs.admin.ch/bfs/de/home/statistiken/kataloge-datenbanken/tabellen.assetdetail.18344310.html, [Accessed 17 April 2022].

Bibri, S. E., 2020. The eco-city and its core environmental dimension of sustainability: green energy technologies and their integration with data-driven smart solutions. *Energy Informatics,* 3(1), pp. 1-26.

Buchmüller, S. & Weidmann, U., 2006. Parameters of pedestrians, pedestrian traffic and walking facilities. *IVT-Report,* Volume 132.

Cambridge Dictionary, 2022. *Meaning of Popularity in English.* [Online]. Available at: https://dictionary.cambridge.org/dictionary/english/popularity, [Accessed 25 April 2022].

Carpio-Pinedo, J., De Gregorio Hurtado, S. & Sánchez De Madariaga, I., 2019. Gender Mainstreaming in Urban Planning: The Potential of Geographic Information Systems and Open Data Sources. *Planning Theory & Practice,* 20(2), pp. 221-240.

Cattell, V., Dines, N., Gesler, W. & Curtis, S., 2008. Mingling, observing, and lingering: Everyday public spaces and their implications for well-being and social relations. *Health & place,* 14(3), pp. 544-561.

Cilliers, E. J., Timmermans, W., van den Goorbergh, F. & Slijkhuis, J. S. A., 2015. Designing public spaces through the lively planning integrative perspective. *Environment, development and sustainability,* 17(6), pp. 1367-1380.

Davidson, M., 2010. Social sustainability and the city. *Geography Compass,* 4(7), pp. 872-880.

de Montjoye, Y.-A., Kendall, J. & Kerry, C. F., 2014. Enabling Humanitarian Use of Mobile Phone Data. *Issues in Technology Innovation,* pp. 1-11.

Dempsey, N., Bramley, G., Power, S. & Brown, C., 2011. The social dimension of sustainable development: Defining urban social sustainability. *Sustainable development,* 19(5), pp. 289-300.

DESA, U.N., 2016. *Transforming our world: The 2030 agenda for sustainable development.*

Dinkić, N. et al., 2016. Using sentiment analysis of Twitter data for determining popularity of city locations. In: *International Conference on ICT Innovations.* Cham: Springer, pp. 156-164.

Drilling, M. & Weiss, S., 2012. *Soziale Nachhaltigkeit in der Siedlungsentwicklung.* Bern: VLP-ASPAN.

Eizenberg, E. & Jabareen, Y., 2017. Social sustainability: A new conceptual framework. *Sustainability,* 9(1), p. 68.

Emenike, A. I., 2016. Developing sustainable and liveable neighbourhoods: the role of public open spaces. *WIT Transactions on Ecology and the Environment,* Band 204, pp. 257-264.

Fleischmann, M., 2019. momepy: Urban Morphology Measuring Toolkit. *Journal of Open Source Software,* 4(43), p. 1807.

Foth, M., Choi, J. H.-j. & Satchell, C., 2011. Urban informatics. In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work.* Hangzhou, China, pp. 1-8.

Frank, S., 2019. Die Digital Natives und das Wohnen in der Stadt der Zukunft: Widersprüche und Ambivalenzen. In: R. G. Heinze, S. Kurtenbach & J. Üblacker, Hrsg. *Digitalisierung und Nachbarschaft..* Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, pp. 165-184.

Ghavampour, E., Del Aguila, M. & Vale, B., 2017. GIS mapping and analysis of behaviour in small urban public spaces. *Area,* 49(3), pp. 349-358.

Gim, T.-H. T., 2017. Full random coefficients multilevel modeling of the relationship between land use and trip time on weekdays and weekends. *Sustainability,* 9(10), p. 1824.

Girardin, F. et al., 2008. Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive computing,* 7(4), pp. 36-43.

Glenz, C., Massolo, A., Kuonen, D. & Schlaepfer, R., 2001. A wolf habitat suitability prediction study in Valais (Switzerland). *Landscape and Urban Planning,* 55(1), pp. 55-65 .

Goodchild, M. F., 2007. Citizens as sensors: the world of volunteered geography. *GeoJournal,* 69(4), pp. 211-221.

Graham, M., 2011. Wiki space: Palimpsests and the politics of exclusion. *Critical point of view: A Wikipedia reader,* pp. 269-282.

Grantz, K. H. et al., 2020. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. *Nature communications,* 11(1), pp. 1-8.

Grimm-Pretner, D., 2004. Public Parks and Squares in Viennese Urban Renewal Areas–Sites for Everyday Life. In: *Proceeding of Open Space: People Space Conference.* Edinburgh, UK.

Hasan, S., Schneider, C. M., Ukkusuri, S. V. & González, M. C., 2013b. Spatiotemporal patterns of urban human mobility. *Journal of Statistical Physics,* 151(1), pp. 304-318.

Hasan, S., Zhan, X. & Ukkusuri, S. V., 2013a. Understanding urban human activity and mobility patterns using large-scale location-based data from online social media. In: *Proceedings of the 2nd ACM SIGKDD international workshop on urban computing.* New York: Association for Computing Machinery, pp. 1-8.

Heikinheimo, V. et al., 2017. User-Generated Geographic Information for Visitor Monitoring in a National Park: A Comparison of Social Media Data and Visitor Survey. *ISPRS International Journal of Geo-Information,* 6(3), p. 85.

Hjort, M., Martin, W. M., Stewart, T. & Troelsen, J., 2018. Design of Urban Public Spaces: Intent vs. Reality. *International Journal of Environmental Research and Public Health,* 15(4), p. 816.

Hsieh, H. P.et al., 2019. Temporal popularity prediction of locations for geographical placement of retail stores. *Knowledge and Information Systems,* 60(1), pp. 247-273.

Ivanov, N. & Gnevanov, M., 2018. Big data: perspectives of using in urban planning and management. In: *MATEC Web of Conferences.* EDP Sciences, p. 01107.

JMP, 2022. *Variable Selection in Multiple Regression.* [Online]. Available at: https://www.jmp.com/en_ch/statistics-knowledge-portal/what-is-multiple-regression/variable-selection.html, [Accessed 29 March 2022].

Johnson, A. J. & Glover, T. D., 2013. Understanding urban public space in a leisure context. *Leisure Sciences,* 35(2), pp. 190-197.

Kahila, M. & Kyttä, M., 2009. SoftGIS as a bridge-builder in collaborative urban planning. In: *Planning support systems best practice and new methods.* Dordrecht: Springer, pp. 389-411.

Koblet, O. & Purves, R. S., 2020. Doctoral dissertation. *Extracting perceived landscape properties from text sources.* Universität Zürich.

Madanipour, A., 1999. Why are the design and development of public spaces significant for cities?. *Environment and planning B: Planning and Design,* 26(6), pp. 879-891.

Maller, A., 1999. Structured accidentalness: An argument in favour of liveable urban congestion. *Journal of Urban Design,* 4(2), pp. 131-154.

Ma, R., Lam, P. T. & Leung, C., 2018. Big data in urban planning practices: Shaping our cities with data. In: *Proceedings of the 21st international symposium on advancement of construction management and real estate.* Singapore: Springer, pp. 365-373.

Marušić, B. G., 2011. Analysis of patterns of spatial occupancy in urban open space using behaviour maps and GIS. *Urban design international,* 16(1), pp. 36-50.

Mehta, V., 2014. Evaluating Public Space. *Journal of Urban Design,* 19(1), pp. 53-88.

MeteoSchweiz, 2022. *Klimabulletin Jahr 2021,* Zürich: Schweizerische Eidgenossenschaft.

Metropolitan Planning Council, 2008. *A Guide to Neighborhood Placemaking in Chicago.* Chicago: Project for Public Spaces Inc, Metropolitan Planning Council.

Montgomery, J., 1998. Making a city: Urbanity, vitality and urban design. *Journal of urban design,* 3(1), pp. 93-116.

Murthy, D., Gross, A. & Pensavalle, A., 2016. Urban social media demographics: An exploration of Twitter use in major American cities. *Journal of Computer-Mediated Communication,* 21(1), pp. 33-49.

Nguyen, T. V. T., Han, H. & Sahito, N., 2019. Role of urban public space and the surrounding environment in promoting sustainable development from the lens of social media. *Sustainability,* 11(21), p. 5967.

ODPM, 2005. Bristol Accord. In: *Conclusions of Ministerial Informal Meeting on Sustainable Communities in Europe.* Bristol: ODPM, pp. 1-20.

Oliver, N. et al., 2020. Mobile phone data and COVID-19: Missing an opportunity?. *arXiv preprint arXiv:2003.12347.*

Pappalardo, L., Pedreschi, D., Smoreda, Z. & Giannotti, F., 2015. Using big data to study the link between human mobility and socio-economic development. In: *2015 IEEE International Conference on Big Data (Big Data).* IEEE, pp. 871-878.

Pesaresi, M., Melchiorri, M., Siragusa, A. & Kemper, T., 2016. *Atlas of the Human Planet - Mapping Human Presence on Earth with the Global Human Settlement Layer,* Luxembourg: Publications Office of the European Union.

Phithakkitnukoon, S. et al., 2010. Activity-aware map: Identifying human daily activity pattern using mobile phone data.. In: *International workshop on human behavior understanding.* Berlin, Heidelberg: Springer, pp. 14-25.

Piri Sahragard, H., Ajorlo, M. & Karami, P., 2018. Modeling habitat suitability of range plant species using random forest method in arid mountainous rangelands. *Journal of Mountain Science,* 15(10), pp. 2159-2171.

Rasidi, M. H., Jamirsah, N. & Said, I., 2012. Urban green space design affects urban residents' social interaction. *Procedia-Social and Behavioral Sciences,* Band 68, pp. 464-480.

Rathore, M. M., Ahmad, A., Paul, A. & Rho, S., 2016. Urban planning and building smart cities based on the internet of things using big data analytics. *Computer networks,* Band 101, pp. 63-80.

Redclift, M., 1991. The multiple dimensions of sustainable development. *Geography,* pp. 36-42.

Ritchie, H. & Roser, M., 2018. *Urbanization. Our world in data.* [Online]. Available at: https://ourworldindata.org/urbanization, [Accessed 21 January 2022].

Rummens, A. et al., 2021. Do Mobile Phone Data Provide a Better Denominator in Crime Rates and Improve Spatiotemporal Predictions of Crime?. *ISPRS International Journal of Geo-Information,* 10(6), p. 369.

Schweizerischer Bundesrat, 2021. *Strategie Nachhaltige Entwicklung 2030.* Bern

Setz, M., Frank, J. & Suter, S., 2019. *regiosuisse Monitoring - Die regionalwirtschaftliche Entwicklung in der Schweiz,* Bern: regiosuisse - Netzwerkstelle Regionalentwicklung.

Shahidehpour, M., Li, Z. & Ganji, M., 2018. Smart cities for a sustainable urbanization: Illuminating the need for establishing smart urban infrastructures. *IEEE Electrification Magazine,* 6(2), pp. 16-33.

Sharipov, S. & Demirkol, H. G., 2018. Analysing the flow pattern in the tourism city of Eskisehir. *International Journal of Tourism Cities,* 4(3), pp. 342-354.

Soper, D., 2012. Is human mobility tracking a good idea?. *Communications of the ACM,* 55(4), pp. 35-37.

SRF, 2021. *20'000 fordern in Zürich ein Ja zu «Ehe für alle».* [Online]. Available at: https://www.srf.ch/news/abstimmungen/ehe-fuer-alle/abstimmung-vom-26-september-20-000-fordern-in-zuerich-ein-ja-zu-ehe-fuer-alle, [Accessed 18 April 2022].

Stadt Zürich Präsidialdepartement, 2022. *Bevölkerung nach Alter, Geschlecht und Herkunft, seit 1993.* [Online]. Available at: https://www.stadt-zuerich.ch/content/prd/de/index/statistik/themen/bevoelkerung/alter-geschlecht/geschlecht.html, [Accessed 5 April 2022].

Stadt Zürich, 2011. *Bedeutungsplan.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_bedeutungsplan, [Accessed 8 October 2021].

Stadt Zürich, 2019. *Öffentlich zugängliche Strassenparkplätze OGD.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_oeffentlich_zugaengliche_strassenparkplaetze_ogd, [Accessed 8 October 2021].

Stadt Zürich, 2020. *Daten der Verkehrszählung zum motorisierten Individualverkehr (Stundenwerte), seit 2012.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/sid_dav_verkehrszaehlung_miv_od2031, [Accessed 8 October 2021].

Stadt Zürich, 2021a. *Parkanlagen von A–Z.* [Online]. Available at: https://www.stadt-zuerich.ch/ted/de/index/gsz/natur-erleben/park-und-gruenanlagen/parkanlagen-von-az.html, [Accessed 13 May 2021].

Stadt Zürich, 2021b. *Strassennamen.* [Online]. Available at: https://www.stadt-zuerich.ch/ted/de/index/geoz/adressierung/strasssennamen.html, [Accessed 13 May 2021].

Stadt Zürich, 2021c. *Zweiradparkierung.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_zweiradparkierung, [Accessed 8 October 2021].

Stadt Zürich, 2021d. *ÖV-Güteklassen (OGD) (kantonaler Datensatz).* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/ktzh_oev_gueteklassen__ogd_, [Accessed 10 October 2021].

Stadt Zürich, 2021e. *Haltestellen des öffentlichen Verkehrs (OGD) (kantonaler Datensatz).* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/ktzh_haltestellen_des_oeffentlichen_verkehrs___ogd_, [Accessed 8 October 2021].

Stadt Zürich, 2021f. *Öffentlich zugängliche Parkhäuser.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_oeffentlich_zugaengliche_parkhaeuser, [Accessed 8 October 2021].

Stadt Zürich, 2021g. *Fuss- und Velowegnetz.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_fuss__und_velowegnetz, [Accessed 8 October 2021].

Stadt Zürich, 2021h. *Züri wie neu.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_zueri_wie_neu, [Accessed 8 October 2021].

Stadt Zürich, 2021i. *Brunnen.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_brunnen, [Accessed 8 October 2021].

Stadt Zürich, 2021j. *Bauminventar.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_bauminventar, [Accessed 8 October 2021].

Stadt Zürich, 2021. *Open Data Portal.* [Online]. Available at: https://data.stadt-zuerich.ch, [Accessed 8 October 2021].

Stadt Zürich, 2022a. *Shopping in Zürich (Zürich Tourismus).* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/zt_shopping, [Accessed 14 February 2022].

Stadt Zürich, 2022b. *Bars und Lounges in der Region Zürich (Zürich Tourismus).* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/zt_bars, [Accessed 14 February 2022].

Stadt Zürich, 2022c. *Gastronomie in Zürich (Zürich Tourismus).* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/zt_gastronomie, [Accessed 14 May 2022].

Stadt Zürich, 2022d. *Standorte ZüriVelo.* [Online]. Available at: https://data.stadt-zuerich.ch/dataset/geo_standorte_zuerivelo, [Accessed 17 February 2022].

Stadt Zürich, 2022k. *Botanischer Garten.* [Online]. Available at: https://www.stadt-zuerich.ch/ted/de/index/gsz/natur-erleben/park-und-gruenanlagen/parkanlagen-von-az/botanischer-garten.html, [Accessed 10 April 2022].

Steenbruggen, J., Tranos, E. & Nijkamp, P., 2015. Data from mobile phone operators: A tool for smarter cities?. *Telecommunications Policy,* 39(3-4), pp. 335-346.

Swisscom AG, 2022a. *Heatmaps - Understand where people are.* [Online]. Available at: https://digital.swisscom.com/products/heatmaps/faq, [Accessed 23 March 2022].

Swisscom AG, 2022b. *Dwell Times - See how long people stay.* [Online]. Available at: https://digital.swisscom.com/products/dwelltimes/info, [Accessed 23 March 2022].

Tages-Anzeiger, 2020. *Protest ohne Schutzmasken, Ausschreitungen in der Nacht.* [Online]. Available at: https://www.tagesanzeiger.ch/protest-ohne-schutzmasken-394525637659, [Accessed 1 April 2022].

Tammet, T., Luberg, A. & Järv, P., 2013. Sightsmap: crowd-sourced popularity of the world places. In: *Information and communication technologies in tourism 2013.* Berlin, Heidelberg: Springer, pp. 314-325.

Taylor, L., 2016. No place to hide? The ethics and analytics of tracking mobility using mobile phone data. *Environment and Planning D: Society and Space,* 34(2), pp. 319-336.

Thakuriah, P. V., Tilahun, N. Y. & Zellner, M., 2017. Big data and urban informatics: innovations and challenges to urban planning and knowledge discovery. In: *Seeing cities through big data.* Cham: Springer, pp. 11-45.

UN, 2018. *World Urbanization Prospects: The 2018 Revision.* United Nations, Department of Economic and Social Affairs, Population Division: custom data acquired via website.

Vinceti, M. et al., 2020. Lockdown timing and efficacy in controlling COVID-19 using mobile phone tracking. *EClinicalMedicine,* Band 25, p. 100457.

WAPC , 2009. *Liveable Neighbourhoods. A Western Australian Government sustainable cities initiative.* Perth: Western Australian Planning Commission.

Waskom, M. L., 2021. seaborn: statistical data visualization. *Journal of Open Source Software,* 6(60), p. 3021.

Wesolowski, A., Buckee, C. O., Engø-Monsen, K. & Metcalf, C. J. E., 2016. Connecting Mobility to Infectious Diseases: The Promise and Limits of Mobile Phone Data. *The Journal of infectious diseases,* 214(Suppl 4), pp. S414-S420.

Wheeler, S., 2001. Livable communities: Creating safe and livable neighborhoods, towns, and regions in California.

WIDEL, 2006. Stadträume 2010 - Strategie für die Gestaltung von Zürichs öffentlichem Raum. Stadt Zürich.

Woodcraft, S., 2012. Social sustainability and new communities: Moving from concept to practice in the UK. *Procedia-Social and Behavioral Sciences,* Band 68, pp. 29-42.

Yaakup, A., Ludin, A. N. M., Sulaiman, S. & Bajuri, H., 2005. GIS in urban planning and management: Malaysian experience. In: *International Symposium & Exhibition on Geoinformation, Geospatial Solutions for Managing the Borderless World.* Pulau Pinang: pp. 27-29.

Yeh, A. G., 1999. Urban planning and GIS. *Geographical information systems,* 2(877-888), p. 1.

Yeh, A. G.-O. & Li, X., 2003. Simulation of development alternatives using neural networks, cellular automata, and GIS for urban planning. *Photogrammetric Engineering & Remote Sensing,* 69(9), pp. 1043-1052.

ZFF, Z. F., 2021. *Programm - 17. Zürich Film Festival.* [Online]. Available at: https://s3.zff.com/medialibrary/2021/09/ZFF21_Programmheft_web.pdf?_ga=2.112011335.2380543.1650273762-1788978176.1650273762, [Accessed April 2022].

Zhou, Y. et al., 2020. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. *The Lancet Digital Health,* 2(8), pp. e417-e424.

Zurich Diamondleague, 2021. *Weltklasse Zürich auf dem Sechseläutenplatz.* [Online]. Available at: https://zurich.diamondleague.com/en/info/besucherguide/besucherguide-auf-dem-sechselaeutenplatz/, [Accessed 18 April 2022].

## Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the thesis.

Sophie Sturzenegger
Zurich, 30.04.2022