Cartography M.Sc.

Master thesis

Visualization of spatial disparities in mobility service frequency with open public transit feed data in Germany

Md Imtiaz Uddin

Technical University of Munich

TECHNISCHE UNIVERSITÄT WIEN Vienna University of Technology

TECHNISCHE UNIVERSITÄT DRESDEN

UNIVERSITY OF TWENTE.

2022

TECHNISCHE UNIVERSITÄT DRESDEN

Visualization of spatial disparities in mobility service frequency with open public transit feed data in Germany

Submitted for the academic degree of Master of Science (M.Sc.)
Conducted at the Institute of Cartography, Department of Geosciences
Technical University of Dresden

Author:	Md Imtiaz Uddin
Study course:	Cartography M.Sc.
1st Supervisor:	Prof. Dipl.-Phys. Dr.-Ing. habil. Dirk Burghardt (TUD)
2nd Supervisor:	Dr. Ing. Sujit Kumar Sikder, Research Associate (IOER)
Reviewer:	Dr. Barend Köbben (UT)
Cooperation:	Leibniz Institute of Ecological Urban and Regional Development

Chair of the Thesis
Assessment Board: Prof. Dipl.-Phys. Dr.-Ing. habil. Dirk Burghardt (TUD)

Date of submission: 08.12.2022

Statement of Authorship

Herewith I declare that I am the sole author of the submitted Master’s thesis entitled:

“Visualization of spatial disparities in mobility service frequency with open public transit feed data in Germany”

I have fully referenced the ideas and work of others, whether published or unpublished. Literal or analogous citations are clearly marked as such.

Dresden, 08.12.2022 Md Imtiaz Uddin

Acknowledgment

First of all, thanks to almighty Allah the creator and deliverer of mankind, for enabling me to accomplish this work through his immense kindness. I feel exceedingly pleased to express my sincere gratitude and acknowledgment for the valuable advice, helpful criticism, and comments to my honorable supervisor Prof. Dipl.-Phys. Dr.-Ing. habil. Dirk Burghardt, Institute of Cartography, Technical University of Dresden. I want to express my gratitude and sincere thanks to Dr. Ing. Sujit Kumar Sikder, Research Associate, Leibniz Institute of Ecological Urban and Regional Development (IOER) for his immense support at the IOER as well as for the constant friendly cooperation and helpful criticism and comments during the research work. I am grateful to my reviewer Dr. Barend Köbben of the University of Twente for his valuable advice and feedback on the proposal defense and midterm presentation. Sincere thanks to Juliane Cron M.Sc. (TUM) for her continuous support during the program. I would like to express my gratitude towards Dr.-Ing Eva Hauthal (TUD) for her support throughout the semester at TU Dresden.

I appreciate the support of my sponsors Vereinigung der Bayerischen Wirtschaft e. V. (vbw), Santander Universitäten Deutschland, Gesellschaft von Freunden und Förderern der TU Dresden e.V. and Deutschlandstipendium selection committee of TUM and TUD for awarding me scholarship during my studies. I want to thank my classmates for their support during my studies. Moreover, I want to thank specially Sadhman Sadik, Estefania Ruiz Martinez, Sagnik Mukherjee, Mariam Gambashidze, Nargiz Kurumbayeva. Last but not least, I also would like to thank Farjana Akter, Md Arifur Rahman and, Marzan Tasnim Oyshi for their constant support during my sickness.

Finally, I would like to express my gratitude to my beloved parents, elder sister, brother-in-law, and younger brother and sister for their immeasurable sacrifice, blessing, and continuous inspiration throughout my life, which has led me to all my success in study and other spheres.

Abstract

Open data is a blessing to modern scientific exploration. This study was conducted using General Transit Feed Specification (GTFS) data of Germany. GTFS data has the potential to explore spatial disparities as it covers hourly data. Along with this OpenStreetMap and administrative data have been used in this study. This thesis seeks to visualize spatial disparities in mobility service frequency in a web-based map. A reproducible workflow for raw GTFS data processing has been developed in this study. Nevertheless, there is still scope for improvement in the designed workflow, especially Distance Indicator modeling needs quick processing. A systematic literature review process was adopted to find the state of the art. Data and visualization designs have been adopted from the literature. *PubTraDis Visualization* is the outcome of the study that uses ArcGIS online-based Dashboard technology. The user-centered design process has been followed in *PubTraDis Visualization*. Data and visualization designs have been adopted from the literature. The tool is suitable to perform comparisons between different Municipalities as well as between weekdays and weekends, and different times of the day. The commonly used disparity index Local Moran's Index has been calculated to show spatial disparities in grid-level Frequency Indicators. The expert-based evaluation shows that this tool would contribute to spatial disparity analysis in public transit research. However, there is scope for improvement in the current state of *PubTraDis Visualization*. Handling automatic updates of the GTFS data can be an important improvement of this tool.

Keywords: Open Data, GTFS, Public Transit Service, Interactive web-map, Spatial Disparities

Table of Contents

List of Tables

List of Figures

List of Equations

List of Models

Acronyms

API	Application Programming Interface
CO	Cluster/Outlier
GTFS	General Transit Feed Specification
IOER	Leibniz Institute of Ecological Urban and Regional Development
JSON	JavaScript Object Notation
OD	Origin-Destination
OTP	OpenTripPlanner
PTNs	Public Transport Networks
TVG	Time-Varying Graph
UCD	User-centered Design
VVO	Verkehrsverbund Oberelbe

Chapter One
INTRODUCTION

1.1 Motivation and Problem Statement

Open Data are publicly available data that are stored and distributed without any restrictions for usage and dissemination (Geiger & Von Lucke, 2012). With the progress of the internet, Open Data is evolving in the development process since the late 90's (Sikder et al., 2019). For the first time in 2005, Google and TriMet published transit data for the Portland Metro area by adopting Google Transit Feed Specification (GTFS) standard (Antrim & Barbeau, 2013). Later, GTFS was renamed as General Transit Feed Specification and has become the standard for the exchange of transit data[1]. GTFS data provide public transport information coupled with the location and direction of transit lines (Goliszek & Połom, 2016). These data are useful for spatial analysis as it provides hourly data, allowing automatic identification of variation in services, e.g., peak hours during the morning, afternoon, or night (Bast et al., 2014; Puchalsky et al., 2011). Third-party agencies are now using GTFS data for a variety of purposes, for example, real-time visualization of traffic movements, delay analysis in public transport, ridesharing, real-time information, and so on (Antrim & Barbeau, 2013). Nevertheless, GTFS comes with some restrictions involved in data format, e.g.: preparation and harmonization. Also, there is a lack of cross-border information that creates uncertainties and must be taken into consideration (Sikder et al., 2020). On the other hand, having comprehensive transit network data, particularly for large dense regions is difficult (Puchalsky et al., 2011).

Literature shows several studies were conducted at different scales and contexts using public transit data (Sikder et al., 2019; Bast et al., 2014; Puchalsky et al., 2011). Mobility services can contribute to disparity when an area is better served than the others. Rural areas face noteworthy contrast over urban areas in terms of essential opportunities. Gaps include accessibility to public services and resources that leads to social inequality (Camarero & Oliva, 2019). Some platforms use GTFS data for visualizing different phenomena related to transit service. However, the above discussion directs that there are opportunities to explore spatial disparities in public transit service frequency and interactively visualize them in a web-based map.

Taking the facts mentioned above into consideration, visualizing spatial disparities in mobility service frequency is therefore the main aim of this thesis. In this research "Spatial Disparity" refers to unequal transit frequency in different service areas. User-centered evaluation of a visualization tool is necessary for the field of Cartography to determine if the interactive design works (Roth et al., 2017). Therefore, the study concluded with the evaluation of the utility and usability of the visualization tool.

[1] https://beyondtransparency.org/part-2/pioneering-open-data-standards-the-gtfs-story/ (Accessed April 2022)

1.2 Research Objectives and Questions

The primary objective of this study is to visualize spatial disparities in mobility service frequency using open public transit feed data (GTFS) in Germany. This research work is of particular interest to the experts working in transportation, spatial planning, and cartographers intended to analyze disparities in transit service frequency. This will also be of special interest to cartographers concerning interactive web-map-based visualization. To complete the research, the study comprises with following objectives:

1. To identify and determine methods of exploring spatial disparities in public transit.

This objective will drive to aggregate different research work done have been done on spatial disparities in public transit. Before applying and developing visualizations these questions will help to understand available methods to explore spatial disparities in public transit research. Moreover, these research questions will help to understand the user-centric design requirements of an interactive web-map-based visualization considering existing tools for transit service visualization from GTFS data.

Research Questions:

A) What are the methods to explore spatial disparities in public transit?

B) What are the available visualization tools can be found online that used GTFS data?

C) What are the user requirements and design parameters for interactive web-map-based visualization?

2. To adopt open-source data for interactive web-map-based visualization of spatial disparities in mobility service frequency.

This objective will be dealing with the required data that will be needed to develop an interactive web-map-based visualization tool. This research will analyze the frequency of public transport for some specific time frame in Germany (please refer to section 3.3.3.1 for more details). For this research GTFS data, street network, administrative, and population data will be collected. The associate research question will depict how to process GTFS and other data and make use of available information for interactive web-map-based visualization.

Research Questions:

A) How can the GTFS data be integrated to calculate mobility service frequency in the best possible spatial resolution?

B) What are the major challenges to harmonize required multi-sourced input dataset in the calculation process of mobility service frequency in Germany?

3. To develop and evaluate a tool for interactive web-map-based visualization of spatial disparities in mobility service frequency.

This objective will complete the research by successful implementation and evaluation of a web-map-based interactive visualization application.

Research Questions:

A) How can the mobility service frequency be represented in the interactive web-map-based visualization application?

B) What are the evaluation outcomes of the designed application based on expert opinions?

1.3 Innovations Intended

There are several research that has been done on public transit data, in particular, real-time movement visualization, accessibility analysis, and transportation service. However, there exists (at least to the author's knowledge) no large-scale analysis and visualization of spatial disparities in mobility service frequency using open public transit feed data for the whole of Germany. This research intended to apply cartographic methods of spatial data visualization and create a tool that can interactively visualize spatial disparities in mobility service frequency. This will link cartographic research to the public transport, planning, and spatial analysis sector.

1.4 Thesis Structure

This thesis consists of five chapters. The contents of the chapters are as follows:

Chapter 1: This chapter consists of the general context of the thesis. It discusses the motivation and identifies the problem with the relevance to the topic. Research objectives and research questions are illustrated in this chapter followed by the intended innovation of the research. Finally, this chapter ends with the thesis structure.

Chapter 2: This chapter familiarizes the reader with the theoretical information of the research. It also addresses all the research questions of the first research objective. The first section explores the methods of exploring spatial disparities in public transit. Existing visualization tools are discussed in the second section. Finally, design principles for an interactive web map have been reviewed in the last section of this chapter.

Chapter 3: This chapter discusses briefly the method adopted in this research. This is further divided into 5 sections. The first section explains about theoretical approach, the second section reflects on the collected data and the third section discusses raw data processing, which also answers the first research question of the second objective. Interface design is introduced in the next section and this chapter ends with the method of interface evaluation.

Chapter 4: This chapter discusses the method and result of the visualization tool. The first three sections are about developing the visualization tool: its data design, visualization, and aggregation. After that user interface and interactive elements of the tool are discussed. This chapter ends with the outcomes of the evaluation.

Chapter 5: This chapter is concerned with discussion in the context of objectives. The first three-part summarizes the research questions. This chapter also discusses the limitations of the study. The last section of the chapter states the challenges of the study.

Chapter 6: The last chapter of the thesis discusses the prospects of the study and ends with the concluding statements.

Chapter Two

THEORETICAL FRAMEWORK

This chapter provides a brief description of theoretical information to the reader. This includes methods of exploring spatial disparities, a review of existing applications, and design principles.

2.1 Exploring Spatial Disparities in Public Transit

2.1.1 Taxonomy

2.1.1.1 Spatial Disparities

The concept of spatial disparities is the differences that are observed in different geographical locations within a country or region, for example, unequal distribution of infrastructure, economy, or other activities (Kabur and Venables, 2007 cited in Jain & Korzhenevych, 2017). Different studies have defined spatial disparities in various forms. For instance, differences in locational accessibility of medical facilities in various ethnic groups (Comber et al., 2008). In section (2.1.1.3), the term "Spatial Disparities" used in this research will be discussed.

2.1.1.2 Public Transit

In public transport research, the terms 'transport' and 'transit' are often used interchangeably, which includes buses, coaches, rail trams, ferries as well as shared cars (Balcombe et al., 2004; Parkan, 2002 cited in, Redman et al., 2013). However, 'public transport' is mostly used in Europe, Japan, and Australia to refer to the same. On the other hand, North America, and Southeast Asia use 'public transit' for the same case (Redman et al., 2013).

However, in this study 'Public Transit' refers to the collection of available transportation modes for a certain geographic region, specifically for Germany. The term "Mobility Service" is also used as "Public Transit" and vice versa throughout the study.

2.1.1.3 Spatial Disparities in Public Transit

As discussed in the first chapter spatial disparities are also observed in public transit. Combination of numerous factors are the reason for disparities in public transit. These include spatial mismatch, modal mismatch, etc.

Disparities may include differences in different public transit quality attributes. Redman et al., (2013) defined different attributes of public transit service quality.

Table 2.1: Definitions of Public Transit Service Quality Attributes

	Attribute	Definition
Physical	Reliability	How closely the actual service matches the route timetable
	Frequency	How often the service operates during a given period
	Speed	The time spent traveling between specified points
	Accessibility	The degree to which public transport is reasonably available to as many people as possible
	Price	The monetary cost of travel
	Information provision	How much information is provided about routes and interchanges
	Ease of transfers/ interchanges	How simple transport connections are, including time spent waiting
	Vehicle condition	The physical and mechanical condition of vehicles, including the frequency of breakdowns
Perceived	Comfort	How comfortable the journey is regarding access to seats, noise levels, driver handling, air conditioning
	Safety	How safe from traffic accidents passengers feel during the journey as well as personal safety
	Convenience	How simple the public transit service is to use and how well it adds to one's ease of mobility
	Aesthetics	The appeal of vehicles, stations, and waiting areas to users' senses

Adopted from (Redman et al., 2013)

Travel time is considered a key indicator to quantify public transit service quality (Xin et al., 2005) as shorter travel time can make public transport more attractive (Eriksson et al., 2008). While calculating travel time, it is important to consider the fluctuations in service provision of transit service (Farber et al., 2014). Similarly, price and frequency are the most important public transit quality attribute (Eboli & Mazzulla, 2008). Jain & Korzhenevych (2017) argued that there is a great difference in travel time in route choices and time of the day. Many studies have been done on different attributes listed in table (2.1). However, this study focuses only on spatial disparities in mobility service frequency. The following section discusses more details about it.

Public Transit Service Frequency: Public transit service frequency defines the number of times for an hour that a transit service user has access to transit mode (Xin et al., 2005). A single trip of public transit consists of all public transportation modes (bus, tram, train, subway, etc.) (Liao et al., 2020) which can vary from one transport service operator to another.

Xin et al. (2005) argued that to measure transit service frequency between two points, three issues need to be addressed. These are as follows:

1. How to recognize the different transit paths between two activity centers that are plausible from passengers' point of view,

2. How to identify the transit service frequency for a corridor that is connected by more than one path,

3. How to account for those routes that require transfers.

The following sections will provide detailed information on the studies related to Public Transport Research, the materials and method they have used, statistical evaluation involved in the study, and visualization techniques identified from the literature study.

2.1.2 Materials and Method

The literature primarily demonstrates three types of tools used to deal with public transit data. These are:

1) OpenTripPlanner
2) Esri ArcGIS Network Analyst
3) Graph-based network analysis

2.1.2.1 OpenTripPlanner (OTP)

OpenTripPlanner (OTP) is an open-source Java-based multi-modal routing engine (Credit et al., 2021). Several studies have used OTP for GTFS data analysis (Credit et al., 2021; Liao et al., 2020; Heppe and Liebig, 2017, Owen & Levinson, 2015; Liebig et al., 2014). OTP can provide routing, headway, and arrival information. This also can calculate the fastest door-to-door trip for every pair of origin-destination of a designated departure time. In this calculation consideration of all the available transport modes and walking speeds (Liao et al, 2020; Liebig et al., 2014) is possible. For example, Liao et al., (2020) set the maximum walking distance to 800 m and used a mixed sampling methodology of 15-minute resolution to prevent the Modifiable Temporal Unit Problem (MTUP).

An Application Programming Interface (API) and a web application (interfaces the API using RESTful services) are two components of OTP. Traffic network graph loading, and route calculation are done by the API. An interactive user interface fusing with the map is provided by the web application (Liebig et al., 2014).

2.1.2.2 Esri ArcGIS Network Analyst

Esri ArcGIS Network Analyst extension offers a wide range of powerful spatial modeling and analytical capabilities[1]. This has been used in different studies (Karner, 2018; Farber & Fu, 2017) to deal with GTFS data. It can calculate travel times (Farber & Fu, 2017), transit frequency, etc.

Moreover, with the OD Cost Matrix tool, researchers combined GTFS data and pedestrian network files to measure accessibility from an origin (Farber and Fu, 2017; Farber et al., 2014). A set of Origin-Destination (OD) and departure time is set to calculate the shortest travel time on the multimodal network (pedestrian and transit). It also allows the creation of a buffer zone to create walking distances of transit users around transit stops and stations. Ingress, egress, waiting, transfer, and in-vehicle journey times need to be considered in the calculation. A consistent speed of 4.8 km/h (Farber and Fu, 2017), and 5 km/h (Karner, 2018) was considered for ingress, egress, and transfer walking times along the pedestrian network. A Euclidian buffers of 0.4 km around bus stops and 0.8 km for the rails stations is recommended by The US Federal Transit Administration (Federal Transit Administration, 2012, cited in Karner, 2018).

On the other hand, calculating OD short path with ArcGIS Network Analyst seems to have efficiency issues subject to an increase in data components (all times of the day, larger network size) (Fayyaz et al., 2017). It will take 60 days on a quad-core processor in ArcGIS to calculate a network with 1400 stations with 100 transit routes (Farber et al., 2016). Several studies also used custom Python code along with ArcGIS Network Analyst for their studies (Farber and Fu, 2017; Farber et al., 2014).

2.1.2.3 Graph-based Network Analysis

Network science is a research field based on graph theory. This allows the researcher to study the connection and collaboration between elements in complex systems (Newman, 2010 cited in Luo et al., 2019). Many studies have been done on the public transit service using network science (Luo et al., 2019, Maduako et al., 2018). Transit network topological structure represents different transit features, for example, streets, stops, and dynamic features represent buses, trains, etc. on the transit network (Maduako et al., 2018). It can accomplish complex network analysis of public transit services such as service frequency and travel time (Lou et al., 2019).

Figure (2.1) shows two most used graph representations (Lou et al., 2019), named: L-space and P-space topology.

[1] https://pro.arcgis.com/en/pro-app/2.8/help/analysis/spatial-analyst/basics/what-is-the-spatial-analyst-extension.htm (Accessed July 2022)

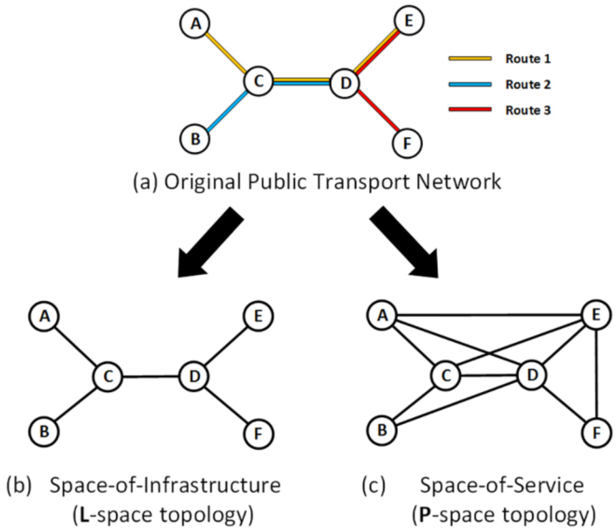


Figure 2.1: Two Common Topological Representations of the Public Transport Network adopted from (von Ferber et al., 2009 cited in Lou et al., 2019)

• L-space: This is defined as a simple representation of the Public Transport Networks (PTNs) from the perspective of structure. Each node of this graph symbolizes a stop. The link between two stops is made when they are nearby on at least one transportation section (i.e., road or rail).

• P-space: This representation is exclusively built on public transport service routes. In this graph, the nodes correspond to stops. They would be associated when they are operated by at least one ordinary route. Consequently, the neighboring node in this space can be reached without any shift.

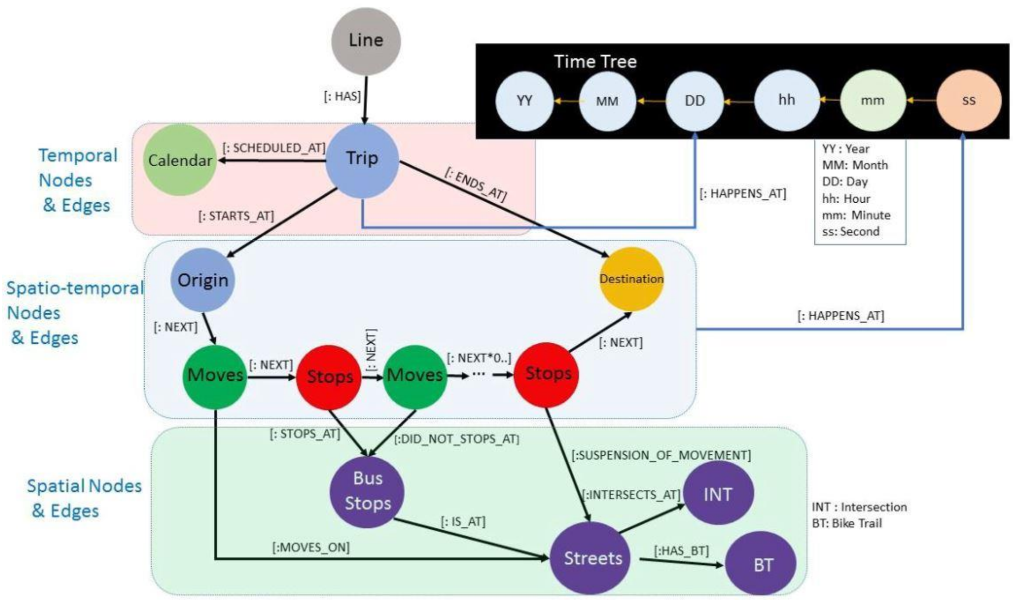


Figure 2.2: Logical Graph Model of the Time-Varying Graph (TVG) for Transit Network adopted from (Maduako et al., 2018)

Figure (2.2) represents the logical graph model of the transit network of the city of Moncton, New Brunswick, Canada TVG. This is a Neo4j graph database-based model. It shows the graph objects (nodes plus relationships) with semantic labels and grouping. This enables users to perform topological calculations such as the shortest path, network diameter, and mobility configuration identification for example traffic movement in a transit network.

2.1.3 Statistical Approach

Literature suggests that statistical evaluation should accompany the study of disparities. The following section portrays, different statistical measures applied to study disparities observed in the literature.

2.1.3.1 Standard Exploratory Data Analysis Techniques

Standard exploratory data analysis techniques have been used in most studies. These include

2.1.3.1.1 Measure of Central Tendency

Measures of central tendency are also identified as summary statistics. Mean, median, and mode are calculated to describe the data by identifying its central position[1].

Farber and Fu (2017) compared the number of transit stops, and routes between two-time frames to discuss disparities. Similarly, a number of studies (Huang, 2020; Luo et al., 2019; Maduako et al., 2018; Owen & Levinson, 2015) used different central tendency measurements for their studies.

2.1.3.1.2 Measures of Dispersion/Variability

These analyses are helpful to understand the data spread. Commonly used measures are coefficient of variation, standard deviation, and variation (Luo et al., 2019; Farber & Fu, 2017; Owen & Levinson, 2015; Farber et al., 2014).

2.1.3.2 Disparity Index

Along with standard exploratory data analysis techniques, some sophisticated statistical operations were also found in the literature. These are as follows:

2.1.3.2.1 Moran's Index I

Moran's Index I used to calculate the spatial autocorrelation of a dataset. Like the correlation coefficient, Moran's Index has a value ranging from -1 to 1. Statistical means of the values are as follows:

[1] https://statistics.laerd.com/statistical-guides/measures-central-tendency-mean-mode-median.php (Accessed July 2022)

- -1 is the perfect clustering of dissimilar values (perfect dispersion).
- 0 means no autocorrelation exists (absolute randomness.)
- +1 indicates perfect clustering of similar values (opposite of dispersion) (Stephanie, 2017).

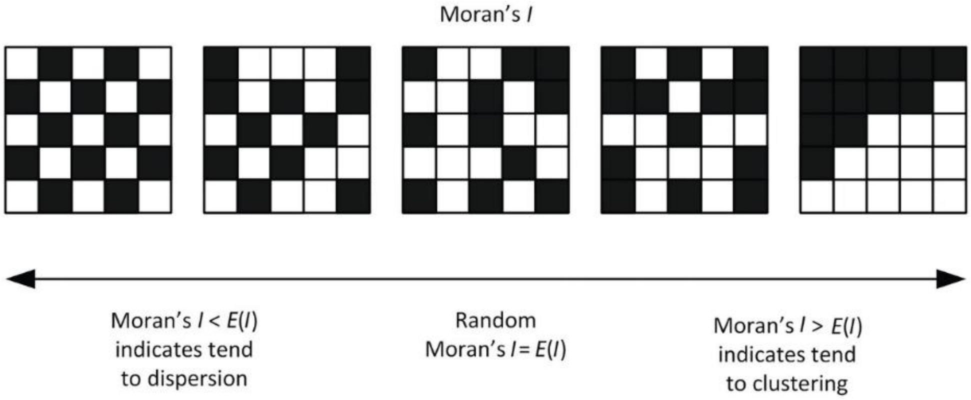


Figure 2.3: Level of Dispersion by Moran's Index adopted from (static.cambridge.org[1])

The equation to calculate Moran's I is shown in equation 2.1:

Equation 2.1: Moran's Index Calculation

$$I = \frac{n}{S_0} \sum_i \sum_j w_{ij}(x_i - \mu)(x_j - \mu) / \sum_i (x_i - \mu)^2$$

In this equation,

n = total number of observations,

$S_0 = \sum_i \sum_j w_{ij}$ is the sum of all the elements of the spatial weight matrix,

μ = mean of x over all the observations.

Note that if W is row-standardized, then $S_0 = n$, and the ratio n/S_0 is equal to one (Anselin & Piras, 2009).

Various transportation studies have used Moran's Index I for calculating spatial autocorrelation in their studies (Niedzielski, 2021; Allen & Farber, 2020; Huang, 2020; Wessel & Farber, 2019; Heppe & Leibig, 2017).

[1] https://static.cambridge.org/binary/version/id/urn:cambridge.org:id:binary:20200421144010724-0711:9781108614528:49898fig4_1.png?pub-status\u003dlive (Accessed July 2022)

2.1.3.2.2 Gini Coefficient

It is a statistical measure that quantifies the amount of inequality present in the distribution of an attribute (Stephanie, 2020). Gini-coefficient is used to capture the disparities in different studies (Panzera & Postiglione, 2020; Stępniak et al., 2019; Xing et al., 2018).

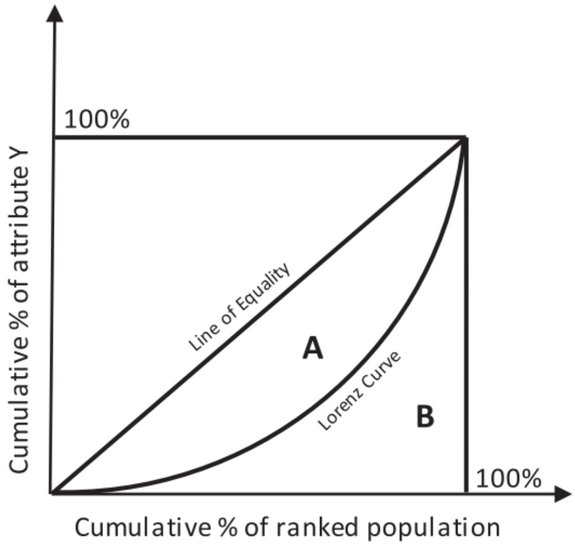


Figure 2.4: Graphical Representation of the Gini-coefficient adopted from (Ben-Elia & Benenson, 2019)

Note: The Gini index Gini (S) is calculated as the ratio between the areas - A/(A + B).

For an ordered sample S of the population attribute y of a size n, $y_1 \leq y_2 \leq \ldots \leq y_i < y_{i+1} \leq \ldots \leq y_{n-1} \leq y_n$, the Gini index is calculated by equation 2.2:

Equation 2.2: Gini Index Calculation

$$Gini(S) = 1 - \frac{2}{n-1}\left(n - \frac{\sum_{i=1}^{n} i y_i}{\sum_{i=1}^{n} y_i}\right)$$

The Gini index value lies between 0 and 1. Index value 0 means complete equality (each population member has the same value of an attribute y), on the contrary, value 1 suggests entire disparity (one member has all). Lower Gini index suggests equal distribution of characteristic y in a sample S (Ben-Elia & Benenson, 2019).

2.1.4 Result Visualization

In this section, the visualization of transport characteristics will be discussed. From the literature, it is evident that there are 3 major types of visualization used to show transport data. These are as follows:

1) Graph
2) Map

2.1.4.1 Graph

Almost everywhere graphs are used for data visualization. Similarly, transportation researchers used graphs for their result analysis. These figures (2.5) have been used to portray different information related to public transport studies.

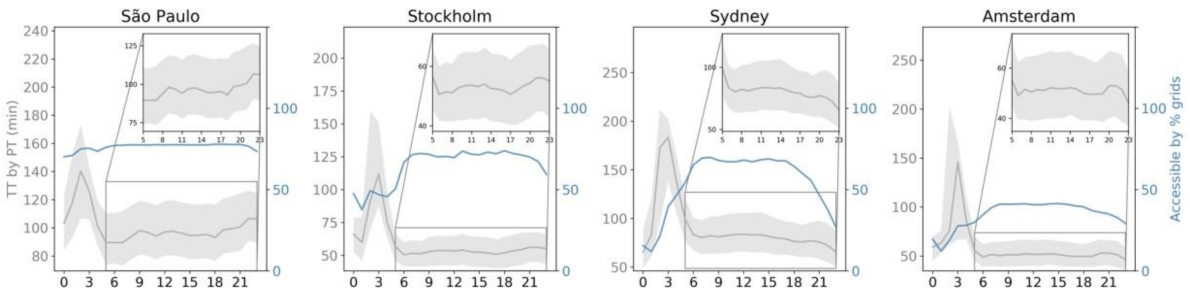


Figure 2.5: Temporal Variations of Travel Time by Public Transit adopted from (Liao et al., 2020)

Figure (2.5) is showing the temporal variation of travel time for a normal weekday by public transit in four different cities: Sao Paulo, Stockholm, Sydney, and Amsterdam. The travel time includes the average cross-departure location of each city weighted by its population density. In the figure, shaded area represents the 25th to 75th percentile range. On the right side, the percentage denotes the accessibility of each grid cell by public transit. Additionally, the inset figure shows the travel time from 5:00 am to 11:00 pm.

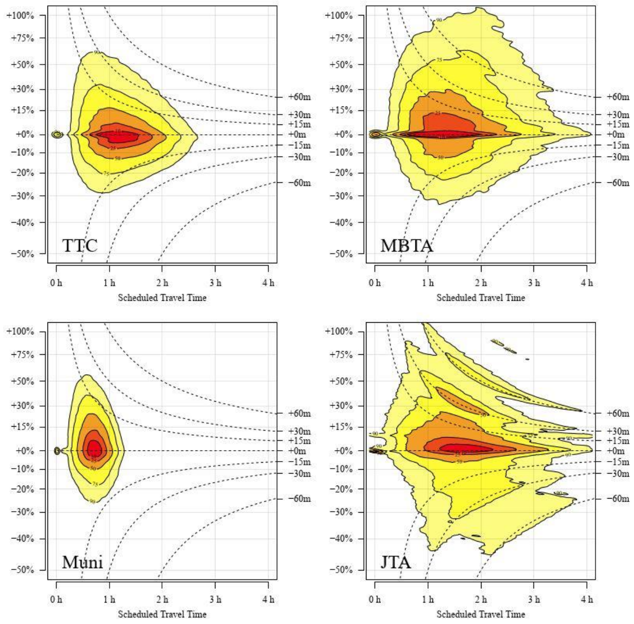


Figure 2.6: Scheduled Travel Times in Florida adopted from (Wessel & Farber, 2019)

This figure (2.6) is showing scheduled travel times of four different transit service agencies Toronto Transit Commission (TTC), the San Francisco Municipal Transportation Agency (SF Muni), the Massachusetts Bay Transportation Authority (MBTA), and the Jacksonville Transportation Authority (JTA) in Florida. It conveys a lot of information. For instance, the difference in average travel times between different transport agencies can be observed.

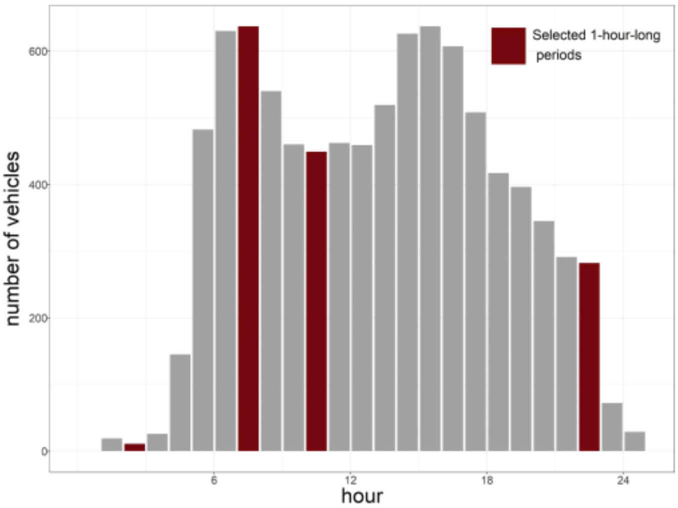


Figure 2.7: Frequency of Public Transport in Szczecin, Poland adopted from (Stępniak et al., 2019)

Figure (2.7) is used to show the total number of buses and trams per hour in the city of Szczecin, Poland. This type of histogram can provide the target user with the actual number of transit services available in an area.

From the above discussion, it can be concluded that graph visualization should accompany while working with transport-related data. It can provide supplementary information to the users through statistical analysis.

2.1.4.2 Map

A map is an important element of data visualization, especially when it comes to spatial data. Almost all the literature reviewed in this study uses maps for showing disparities in public transport frequency, distribution of transit stops, etc.

Figure 2.8: Results of Route Calculations adopted from (Liebig et al., 2014)

Figure (2.8) shows calculated routes depending on the traffic situation in Ireland. In this calculation, authors considered a weekday (Monday, 8th April 2013) and calculated routes from a permanent OD at different time stamps (from left to right: 7:00, 8:00, 8:30 in the morning). From this map, it can be said that basemap plays a significant role in data exploration. So, the visualization tool must have different basemap so that the users can choose them according to their preferences.

In figure (2.9), the author has shown the frequency of public transport for four different time frames (02:00-03:00, 07:00-08:00, 10:00-11:00, and 22:00-23:00) in the city of Szczecin, Poland. From the map, it is also evident that the associated legend of vehicles per hour helps the reader to differentiate between two-time frames.

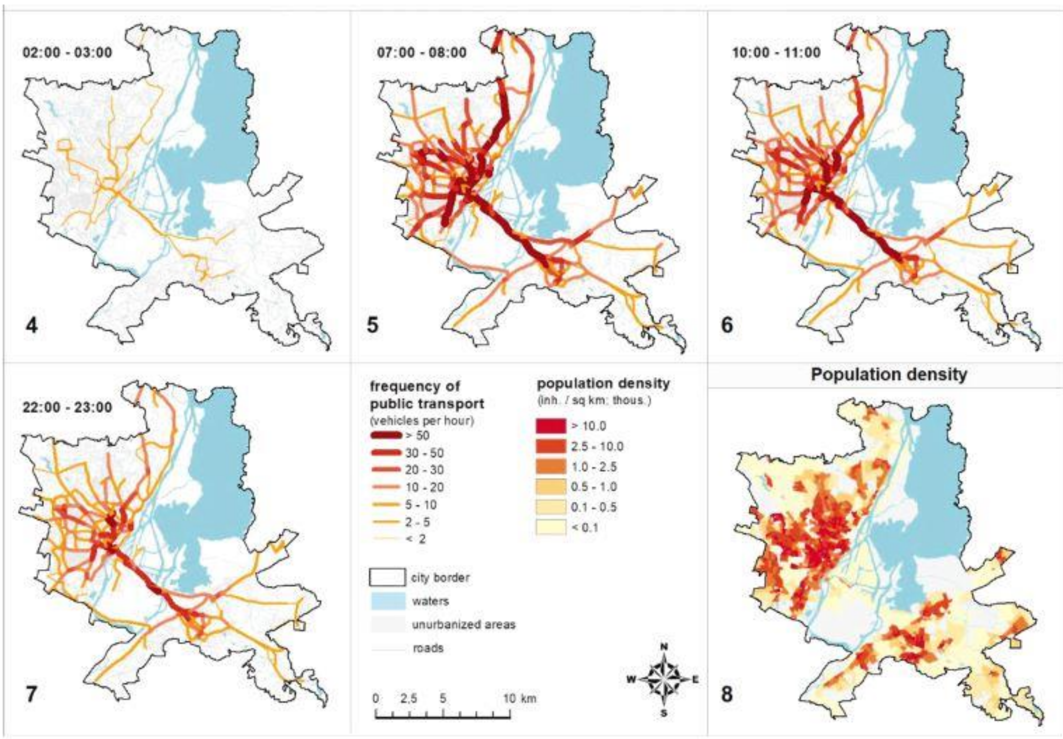


Figure 2.9: Frequency of Public Transport Network in Szczecin, Poland adopted from (Stępniak et al., 2019)

In this figure (2.10), the researchers have shown the average service frequency per hour in Cincinnati City in the US. They have used graduated color to symbolize different classes, while the map is showing the transit area boundary and facility locations (Supermarkets).

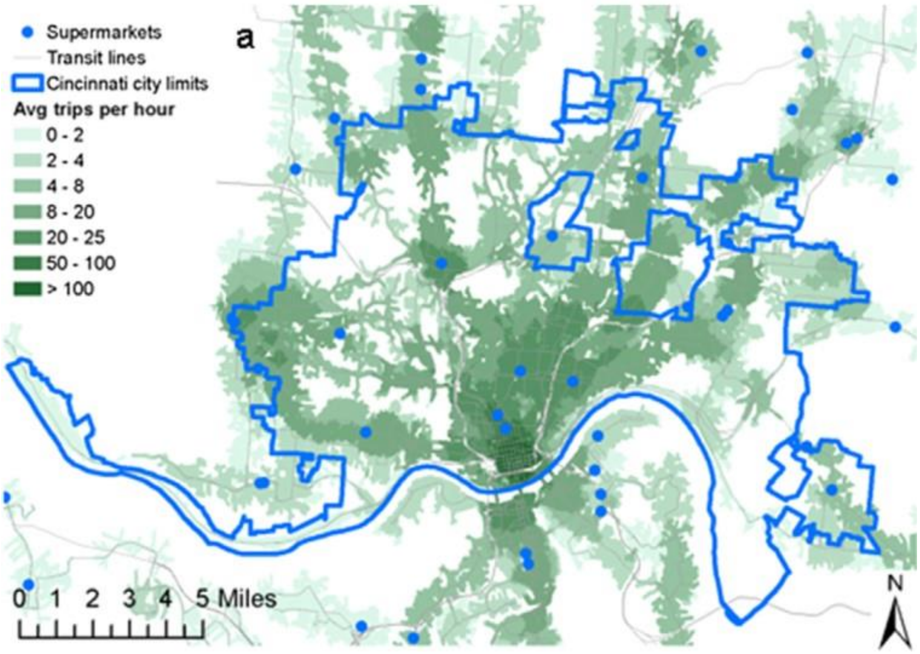


Figure 2.10: Transit Service Frequency Map of Cincinnati City adopted from (Farber et al., 2014)

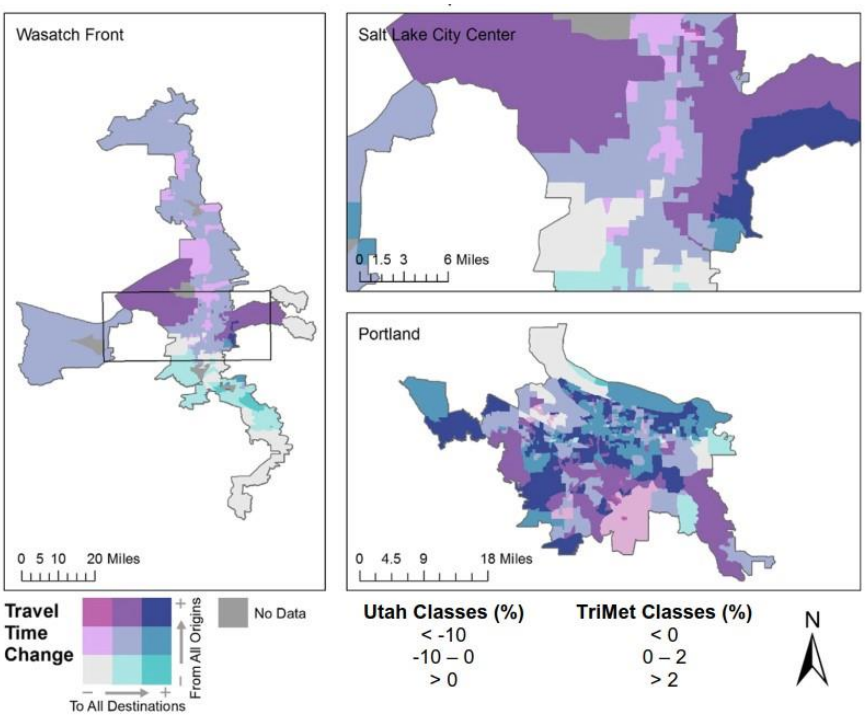


Figure 2.11: Average Travel Time Change Map adopted from (Farber & Fu, 2017)

In this figure (2.11), the authors have shown the changes in travel times of three different zones in the US. These changes are shown in the bivariate choropleth map in the percentage. This type of visualization would be useful to show grid-level frequency indicator in the visualization tool.

Figure (2.12) is showing accessibility ratio between transit and auto at 7:00 am in Miami-Dade County, US. This type of visualization would be useful to show grid level frequency indicator in the proposed tool.

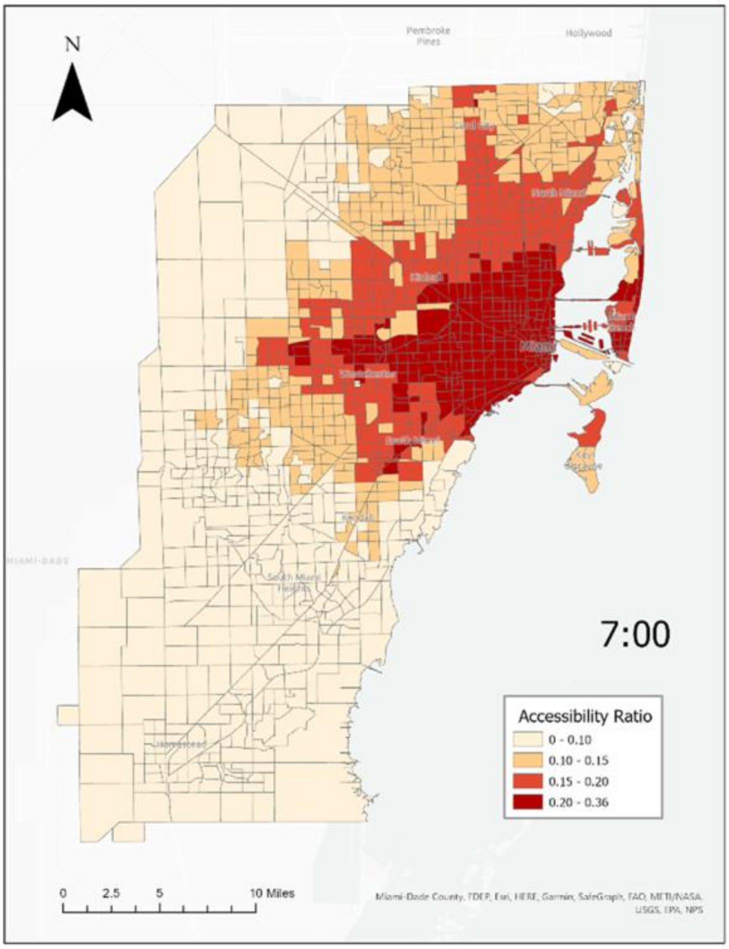


Figure 2.12: Transit and Auto Accessibility Ratio in Miami-Dade County, US adopted from (Yan et al., 2022)

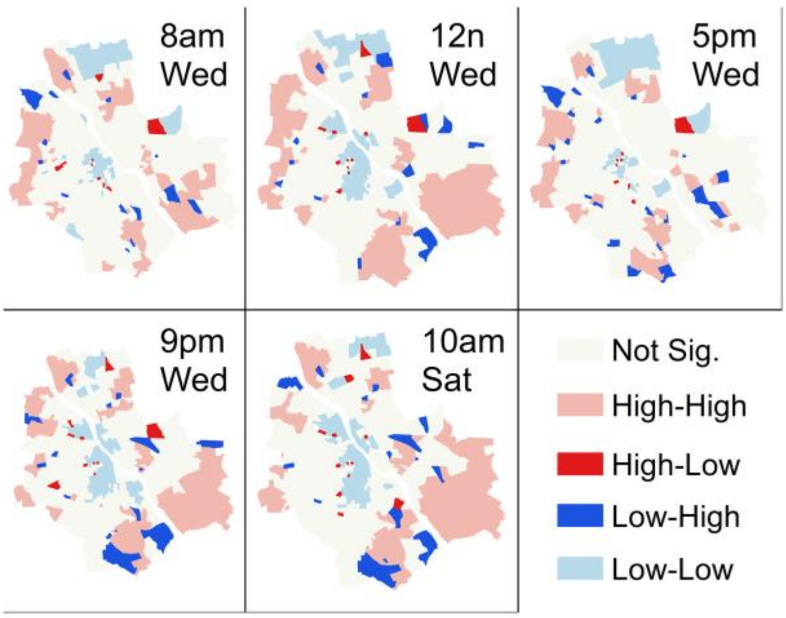


Figure 2.13: Local Moran's I Result by the Time of Day across All Metrics adopted from (Niedzielski, 2021)

Furthermore, Niedzielski (2021) used 5 different colors to show different clusters (figure 2.13). This technique would be particularly helpful for the user to understand the spatial disparity in public transit frequency.

Table 2.2: Selected Papers on Spatial Disparities in Public Transit

Bibtex Key/ Category	T1	T2	T3	T4	D1	D2	D3	D4	D5	2DG	2DM	3DV	TP1	TP2	TP3	SA	SA2	SA3	SA
Liao et al., 2020																			
Leibig et al., 2014																			
Farber and Fu, 2017																			
Farber et al., 2014																			
Heppe and Liebig, 2017																			
Credit et al., 2021																			
Fayyaz et al., 2017																			
Karner, 2018																			
Lou et al., 2019																			
Owen and Levinson,																			
Yan et al., 2022																			
Higgins et al., 2021																			
Chia and Lee, 2020																			
Niedzielski, 2021																			
Wessel and Farber, 2019																			
Allen and Farber, 2020																			
Stępniak et al., 2019																			
Lyons and Choi, 2021																			
Huang, 2020																			
Saghapour et al., 2016																			
Maduako et al., 2018																			
SUM	8	10	2	8	18	13	3	10	9	17	17	2	14	5	4	7	7	5	3

Note: T1-T3 (Tools): OpenTripPlanner, Esri ArcGIS Network Analyst, Graph based network; D1-D5 (Data): GTFS, Street network, GPS, Census, Others; 2DG: 2D Graph; 2DM: 2D Map; 3DV: 3D visualization; TP1-TP3 (Topic): Accessibility, Disparities/Equity, Others; SA1-SA4 (Spatial analysis): Central tendency, Dispersion/Variability, Local Moran's Index, Others.

2.2 Existing Applications

This section reviews available visualization tools found online that use GTFS data.

a) PubtraVis

Prommaharaj et al., (2020) created a visualization tool (PubtraVis) to display a spatiotemporal pattern of transit services from GTFS data. They used Calgary Transit as a case study. This tool can provide qualitative information to the user.

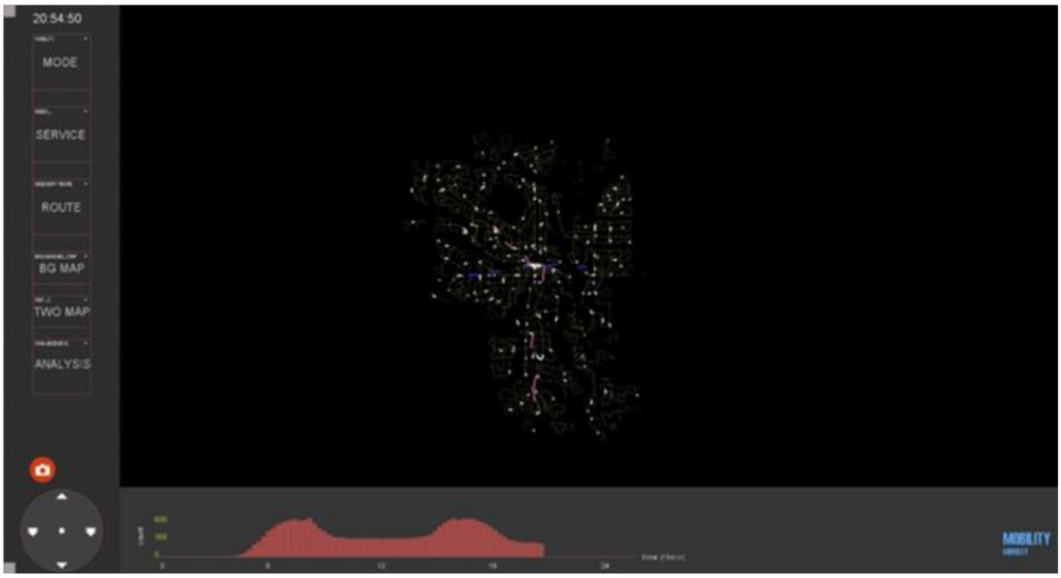


Figure 2.14: Interface of PubtraVis
adopted from (Prommaharaj et al., 2020)

The tool offers 6 visualization components that show the operational characteristics of the transit system. These are mobility, speed, flow, density, headway, and analysis. Side-by-side evaluation of two modules is possible in PubtraVis. The usability test was done with real-world users.

b) Swiss Railways Network

Swiss Railways Network is an animation-based map of the Swiss Federal Railways (SBB) network. The real-time data used in the visualization are obtained from (opentransportdata.swiss) in GTFS-RT format. Users can switch to 3 different basemaps including satellite view. Users can also choose specific transit service operators. The overview map also shows the number of trains operating during that period. Movement speed can be increased by 500 times. Zooming and searching for specific location is also possible in this tool. Different color has been used to distinguish between different train types.

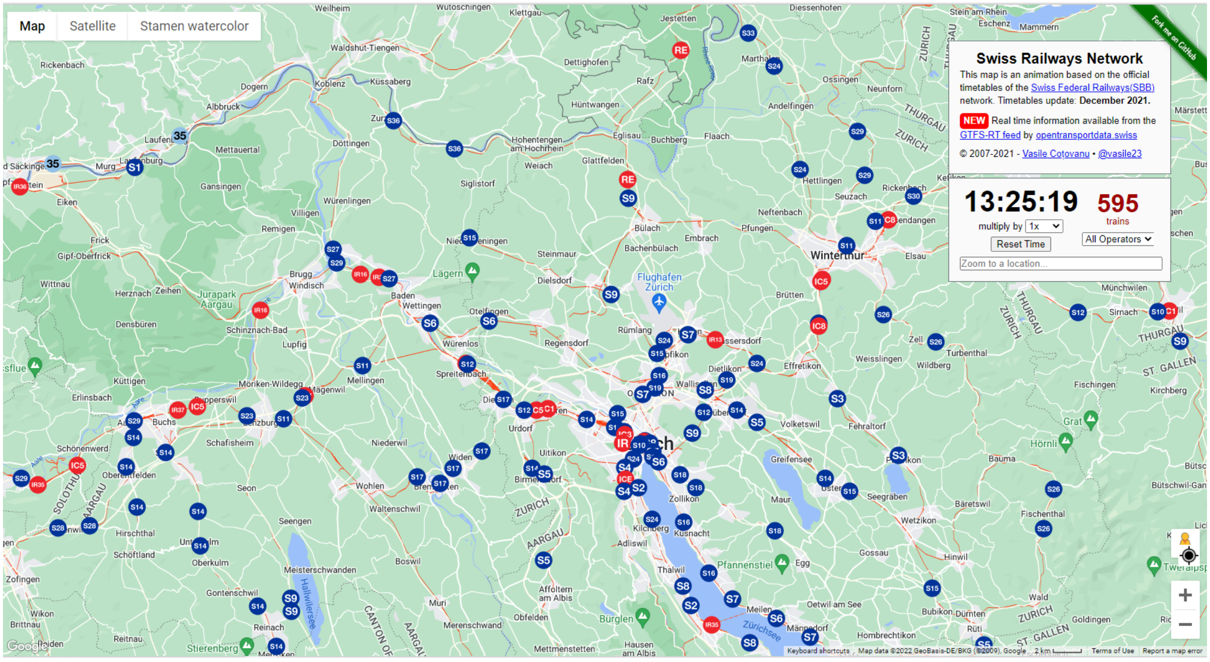


Figure 2.15: Interface of Swiss Railways Network
adopted from (https://maps.vasile.ch/transit-sbb/)

c) TRAVIC (Transit Visualization Client)

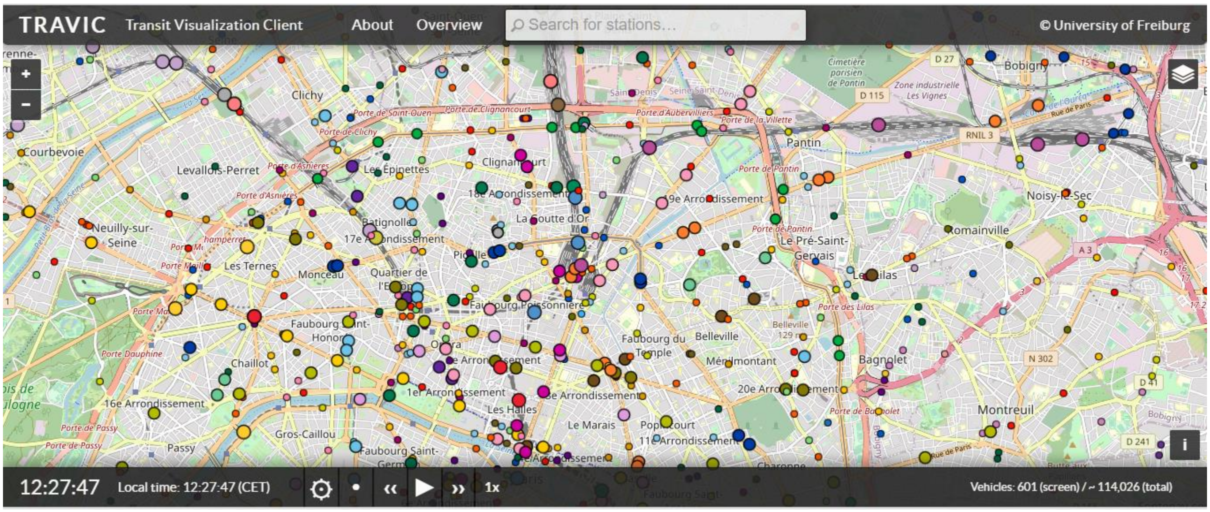


Figure 2.16: Interface of TRAVIC
adopted from (https://travic.app/)

Brosi (2014) has developed the TRAVIC (Transit Visualization Client) tool that visualizes the movement of public transit services from GTFS data. Users can select a base map from 6 available options. In this visualization tool, there is a possibility to overlay 3 different layers (Railway Map Infrastructure, Signals, and Maximum Speed). Identification of different transport modes (Trams, Subways, Rail, Bus, and Ferries) is possible in this tool. Different colors and sizes have been used to distinguish between different transit modes. Users can also select different zoom levels in the tool. However, the tool also provides the user to set a maximum 30x speed of the movement visualization.

2.3 Design Principles

2.3.1 Interactive Web-map Design

From the ancient wall-carved map to spatial diagrams, all maps are inherently interactive (Brown, 1949 cited in Roth, 2012). Researchers also argued that static maps and sketched maps are also interactive (Roth, 2013a; MacEachren & Ganter, 1990). For instance, a paper map user interacts by folding it to change mapped extent, bring it close or far from his or her eyes, and may be use color for important locations (Wallace, 2011 cited in Roth, 2012). However, the interactivity was limited and not impactful as it is today (Andrienko and Andrienko, 1999; Dykes, 2005; Harrower, 2008 cited in Roth, 2012). Creation of real-time interactive maps is now possible because of the advancement in personal computing and internet technologies (MacEachren & Monmonier, 1992 cited in Roth, 2012). In the following sections required map elements of an interactive web-map, cartographic interaction and operator primitives are discussed.

Map Elements

Table (2.3) convey information on important map elements and associated application in web mapping (Tolochko, 2016).

Table 2.3: Map Elements

Map Element	Possible web map adaptation(s)
Map title	Web maps can use a temporary splash screen that disappears when the user begins to interact with the map rather than using a static map title that remain in view.
Mapped area	Because of the ability to pan some web maps, the map uses can have control of which pact of the map they are looking at. Click-add-drag is the most common form of panning, or adjusting the mapped area in the current view. Many web maps with thematic data overlays often use base maps composed of map tiles, which may be provided by corporate or open-source contributors. Thematic design is discussed in more detail below.
Map scale	With interactive maps comes the possibility of multi-scale mapping, i.e., the design of a map to be seen at multiple zoom levels. Mapmakers should design their map differently for different zoom levels and only allow for zoom levels that are necessary for the purpose of the map.

	Zoom buttons are the most commonly visible map element to allow users to change the scale of the map. Some web maps display the given map scale at each zoom level, but many do not.
Supplemental information	Supplemental information can include embedded text, links, images, graphics, videos and be displayed interactively, e.g., when the user clicks on a button or map feature. This information can be presented as information windows or as a tooltip (on top of the map itself, moves with the pointer).
Labels	The quantity and size of place name labels can change at different zoom levels.
Inset/Locator map	The locator map can change to match the current view as the user interacts with the map. Locator map can help with zooming, allowing the user to zoom into the main map by drawing a rectangle of the desired area on the inset.
Map metadata	Metadata includes the cartographer(s)' name(s), data sources, map projection, etc. It does not need to be shown directly on the map but can be easily available in a splash screen or on a linked webpage.
North arrow	Rotates interactively as the user rotates the map. Many interactive maps do not have north arrows, as it is assumed that north is up.
Neat lines/ frame lines	For mobile, when the map takes up the full size of the screen, the neat line is the screen itself.
Supplemental graphs	Users maybe have a better understanding of thematic data through the inclusion of graphs or chats that are linked to the thematic map through simultaneous highlighting or brushing. These linked graphs or charts to the map should not be shown directly on top of the map, but rather in a side or bottom panel.
Legend	Allow for interaction with the legend to affect the map, such as being able to turn/off layers, adjust the timeline of temporal map data, etc.

*Menus	Provide the map user with additional options and interactivity in the form of menus.
*Help	Provide link to information for map users who need help learning how to use the map.

Adopted from (Muehlenhaus, 2013 cited in Tolochko, 2016). Map elements with (*) solely applicable to web maps.

Cartographic Interaction

Cartographic interaction can be specified as the communication of human and a map using a computing device (Roth, 2012). Figure (2.17) shows the components for a successful cartographic interaction.

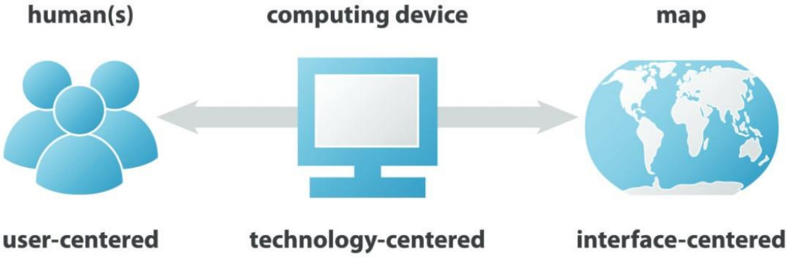


Figure 2.17: Components of Cartographic Interaction adopted from (Roth, 2012)

The (i) human (left) who is the same time develop and usage the interactive map (ii) the map (right) which is the end product and (iii) computing device which is the technology involved in the production.

The limitation for such kind of interactive maps lies on map user's objective, developer skills and hardware configuration (Gahegan, 1999).

Operator primitives

Following table (2.4) shows two major operator primitives adopted from Roth (2013b). Although these primitives play important role in the interactive web map design, are not necessarily offered in all maps hence developer should decide proper primitives to include or discard (Tolochko, 2016). Map maker can decide depending on the type of maps and purpose of map use.

Table 2.4: Cartographic Operator Primitives

Operator	Card Example
Enabling Operators	
Import	Get started by loading a stock map design of the world
Export	Export the maps as .pdf
Save	Save the map so that you can come back later to make a modification
Edit	Select a point to change the attribute data
Annotate	Mark up the map to show where to send resources
Work Operators	
Reexpress	Switch among multiple map representation strategies
Arrange	Arrange a large number of maps for simultaneous comparison
Sequence	Display one time slice after another on the map
Resymbolize	Change the relative sizing of circular proportional symbols
Overlay	Click on the layer panel to show layers of different types of crimes
Reproject	project the map using Albers equal area conic projection
Pan	Pan the map to different location
Zoom	Zoom in to see what is around the point source
Filter	Perform a query that specifies the range of contaminant concentration levels
Search	Enter search words into Google Maps to find target communities in Pittsburg
Retrieve	Brush over the first district of California to see how people voted
Calculate	Select two cities and calculate the distance between them

Adopted from (Roth, 2013b)

2.3.2 User-centered Map Design

User-centered Design (UCD) refers to the interface design while considering the user's needs. This includes special focus on ease-of-use and effectiveness of the designed interface (Norman, 1988; Nielsen, 1992; Nielsen 1993 cited in Roth et al., 2015). In this way the end product is induced by end-users (Abras et al., 2004).

The following table shows how to involve the target users in UCD process (Preece et al., 2002 cited in Abras et al., 2004)

Table 2.5: Users Integration in the UCD Process

Technique	Purpose	Stage of the design cycle
Background Interviews and questionnaires	Collecting data related to the needs and expectations of users; evaluation of design alternatives, prototypes and the final artifact	At the beginning of the design project
Sequence of work interviews and questionnaires	Collecting data related to the sequence of work to be performed with the artifact	Early in the design cycle
Focus groups	Include a wide range of stakeholders to discuss issues and requirements	
On-site observation	Collecting information concerning the environment in which the artifact will be used	
Role Playing, walkthroughs, and simulations	Evaluation of alternative designs and gaining additional information about user needs and expectations; prototype evaluation	Early and mid-point in the design cycle
Usability testing	Collecting quantities data related to measurable usability criteria	Final stage of the design cycle
Interviews and questionnaires	Collecting qualitative data related to user satisfaction with the artifact	Final stage of the design cycle

Adopted from (Preece et al., 2002 cited in Abras et al., 2004)

But, UCD is not always considered in designing, and it may be because of lack of access to user and time or money. But it can help the map developer to save resources by considering users opinion in the early stage of interface rather than making fundamental changes after development (Roth et al., 2015). Nielson (1988) emphasized the significance of iterative evaluation and revision during UCD (Roth et al., 2015). There are ten elements in his usability engineering lifecycle listed below:

(i) **Knowing the user:** Task or work domain analysis

(ii) **Competitive analysis:** Explore current application that support similar use case

(iii) **Setting the goal:** Use previous steps result to form a required functionality

(iv) **Participatory design:** Include target user in the design concept

(v) **Coordinated design:** Coordinate design to create a consistent product identity

(vi) **Guideline and heuristic analysis:** Include expert during development process to maintain design guidelines

(vii) **Prototyping:** Design static or interactive mockups of the user interface.

(viii) **Empirical testing:** Include target user to evaluate utility and usability

(ix) **Iterative design:** Improve the interface based on feedback

(x) **Collect feedback from field use:** Collect feedback about the user interface after the launch and use them in the next release.

On the other hand, Robinson et al., (2005) simplified these ten elements into six which is shown by figure (2.18).

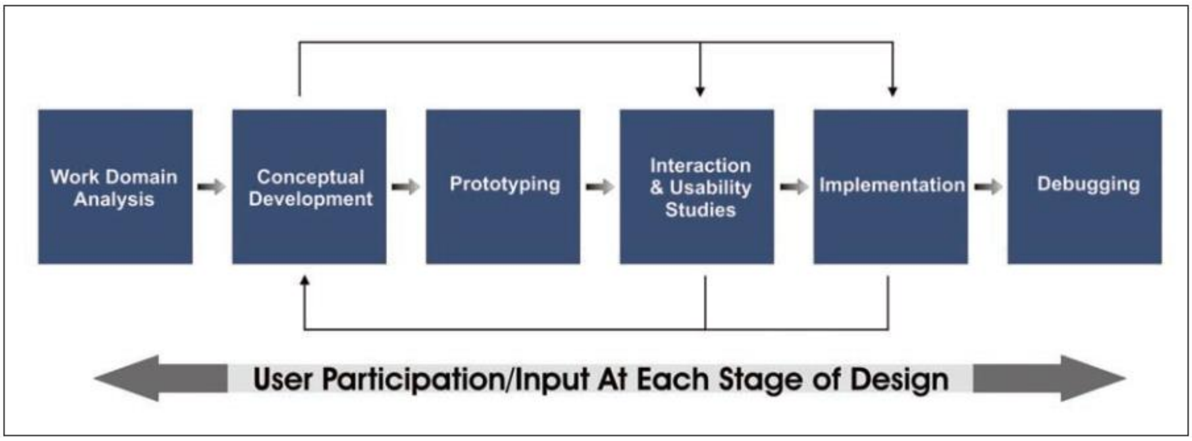


Figure 2.18: User-centred Design Process adopted from (Robinson et al., 2005)

In this process Robinson et al., (2005) adopted six stages to develop geo-visualization tools for Epidemiology. These stages are as follows:

a) **Work domain analysis:** It is the preliminary interaction of ideas and client requirements.
b) **Conceptual development:** Outline of the desired features such as layout, tools, and architecture.
c) **Prototyping:** Working model of the application.
d) **Interaction & usability studies:** Test the application with user incorporating formal and informal feedback.
e) **Implementation:** Adopting the result from the feedback and make required changes.
f) **Debugging:** The application is modified to increase stability and compatibility.

Chapter Three

METHODOLOGY

This chapter provides brief information about the methods used in this study. This discusses theoretical approach used in the study. Data processing also elaborately explained in this chapter. It also introduces the expected outcomes of the web-map and evaluation method.

The research questions are different in dimension, so it is important to choose an appropriate research design and method that fulfill the goal. Combination of different research design and adjustment for research purpose is possible (Bryman, 2012). To complete the research systematically, a mixed research method has been adopted to achieve the research objective and answer the research questions. Figure (3.1) illustrates the workflow of the methodology used in this research.

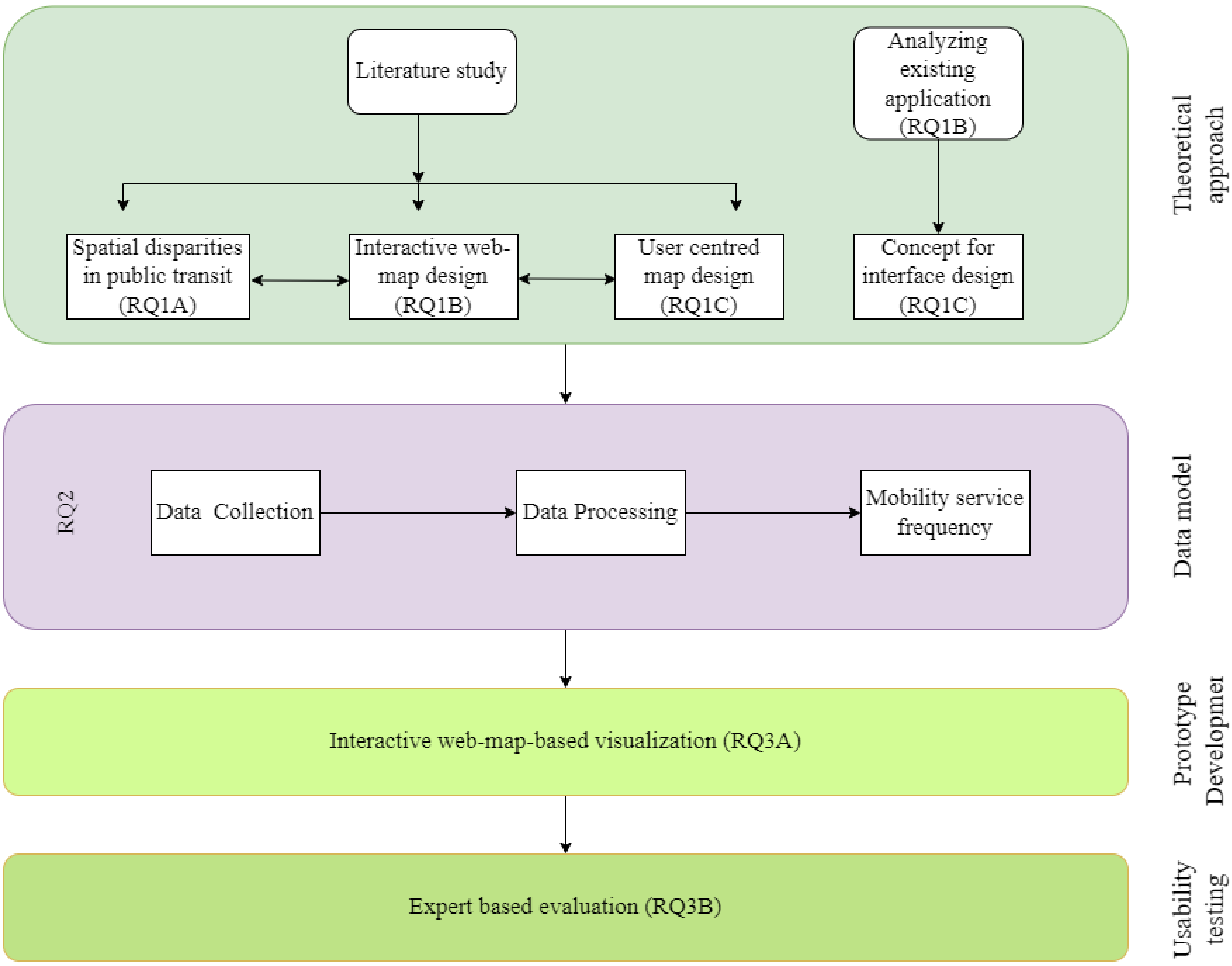


Figure 3.1: Workflow of the Methodology

3.1 Theoretical Approach

In this research systematic literature review (SLR) approach has been applied to identify the current state of art and answer the RQ1A. In this process Kitchenham's guidelines (Kitchenham et al., 2002) were followed. Initially 5 steps have been determined. These are (i) data sources and search strategy, (ii) inclusion and exclusion criteria, (iii) quality assessment, (iv) data extraction, and (v) selected studies.

3.1.1 Data Sources and Search Strategy

There are number of digital libraries and search engines that are available to collect papers (Chotisarn et al., 2020). As there are no dataset that present complete published materials, multiple databases (Web of Science and Scopus) have been chosen for this study (Xiao & Watson, 2019).

Keywords used for the search:

The keyword has been generated from the research question. The research question used for the search is "What are the methods to explore spatial disparities in public transit?". To obtain maximum number of articles, the search term also includes different synonym used in related research.

These terms were selected for final search:
"spatial disparity", "spatial disparities", "spatial inequality", "spatial equity", "public transit service", "public transport" and "public transit"
These keywords are the formulated as search terms according to the specific search engine guideline. Search has been done on 22nd June 2022.

Web of Science:

1st search (with title only)

((((((TI=(spatial disparity)) OR TI=(spatial disparities)) OR TI=(spatial inequality)) OR TI=(spatial equity)) AND TI=(public transit service)) OR TI=(public transport)) OR TI=(public transit)

2nd search (with abstract only)

((((((AB=(spatial disparity)) OR AB=(spatial disparities)) OR AB=(spatial inequality)) OR AB=(spatial equity)) AND AB=(public transit service)) OR AB=(public transport)) OR AB=(public transit)

Combine search (Title + Abstract): In third step the combine result of 1st and 2nd search was generated to extract most relevant papers. 4127 papers have been found from the search.

Scopus:

spatial disparity OR spatial disparities OR spatial inequality OR spatial equity AND public transit service OR public transport OR public transit

From Scopus 3940 papers have been found with the above search criteria.

3.1.2 Inclusion and Exclusion Criteria

According to (Kitchenham & Charters, 2007 cited in Xiao & Watson, 2019) inclusion and exclusion criteria should be stablished based on research question(s) and should be practical. This research is limited to open public transit data (GTFS), which has started from 2006. Consequently, papers which published before that period has been omitted. Moreover, in this method all papers other than article has been excluded. Final exclusion removed papers which are not in English.

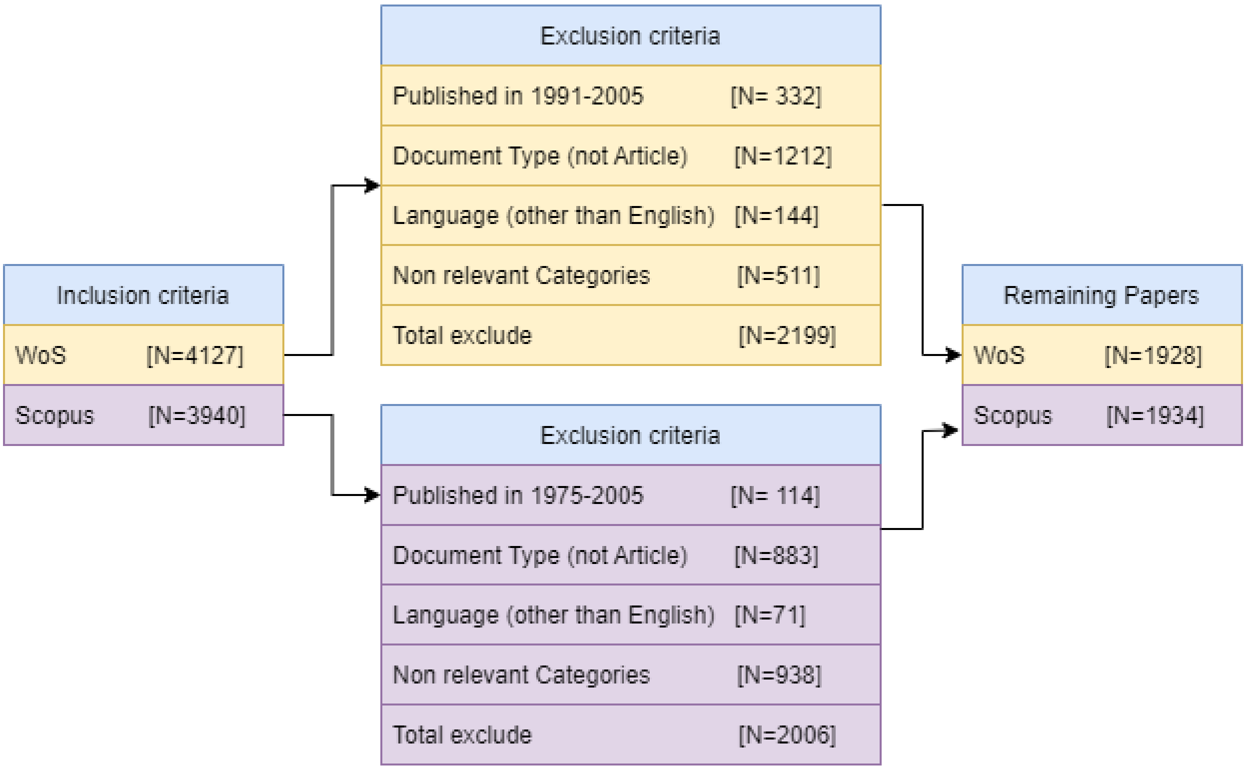


Figure 3.2: Stages of the Search Process and Number of Selected Studies at Each Stage

In this stage a details preview has been done on obtained research papers. The author has decided to work with papers retrieved from the Scopus database for further analysis. The decision has been made upon the following criteria: similarity between two data sources, time limitation of the research, and satisfactory collection of research papers in Scopus database search results.

3.1.3 Quality Assessment

Within the scope of this study and time limitation, the total number of articles to be reviewed needed to narrow down. Thus, quality assessment criteria have been set with the scientific guidelines (Chotisarn et al., 2020; Xiao and Watson, 2019) and adopted for the specific need. It is important to know the top journal in the relevant field. This will help to identify the proper literature for the study. Although initially 1934 articles were retrieved in the later stage top ten journals with more than or at least 30 articles have been considered for further step.

Table 3.1: Top 10 Journals & Their Impacts Based on the Number of Publications (≥ 30)

Rank	Sources	N	TC	C/A	h-in
1	Journal Of Transport Geography	242	6656	27.50	43
2	Sustainability (Switzerland)	190	1724	9.07	21
3	Transportation Research Part A: Policy And Practice	113	2900	25.66	29
4	Transport Policy	109	1886	17.30	28
5	Transportation Research Part D: Transport And Environment	82	838	10.22	17
6	Transportation Research Record	54	476	8.81	14
7	Journal Of Transport And Land Use	42	645	15.36	12
8	Sustainable Cities And Society	42	540	12.86	15
9	Case Studies On Transport Policy	36	173	4.81	8
10	Transportation	32	576	18	12

Note: N: Number of total publications; TC: Total citations; C/A: Average citation per article; h-in.: Hirsch (h)- index.

From table 3.1 it is observed that the Journal Of Transport Geography is leading with 242 articles, followed by Sustainability (Switzerland) (N = 190), Transportation Research Part A: Policy And Practice (N = 113), and Transport Policy (N= 108). In terms of h-index, the Journal Of Transport Geography has secured the highest h-index score of 43. Transportation Research Part A: Policy And Practice ranked second in this category with an h-index of 29 and Transport Policy ranked third with an h-index of 28. Table 3.1 also shows that the Journal Of Transport Geography and Transportation Research Part A: Policy And Practice secured the first two positions based on total number of citations. From the table, it is also evident that the Journal Of Transport And Land Use and Sustainable Cities And Society has 42 articles with more than 12 citations per article even though they have published a smaller number of articles. Moreover, Transportation has more than 18 C/A with only 32 articles, this suggests that they can be more influential and should consider for the literature study.

Although the keyword has been set up only for a related topic, it is observed that some papers returned irrelevant keywords. Thus, 491 papers have been omitted.

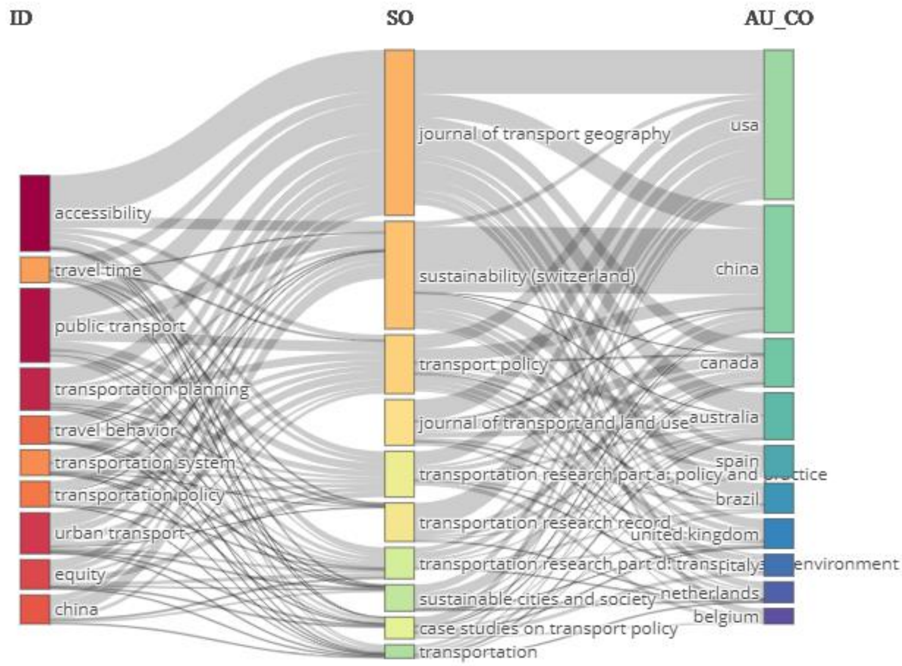


Figure 3.3: Three-field Diagram of the Collected Literature

Note: In this figure ID (left) represents the keywords, SO (middle) represents the journal name and AU_CO (right) represents the source country of the correspondent authors'.

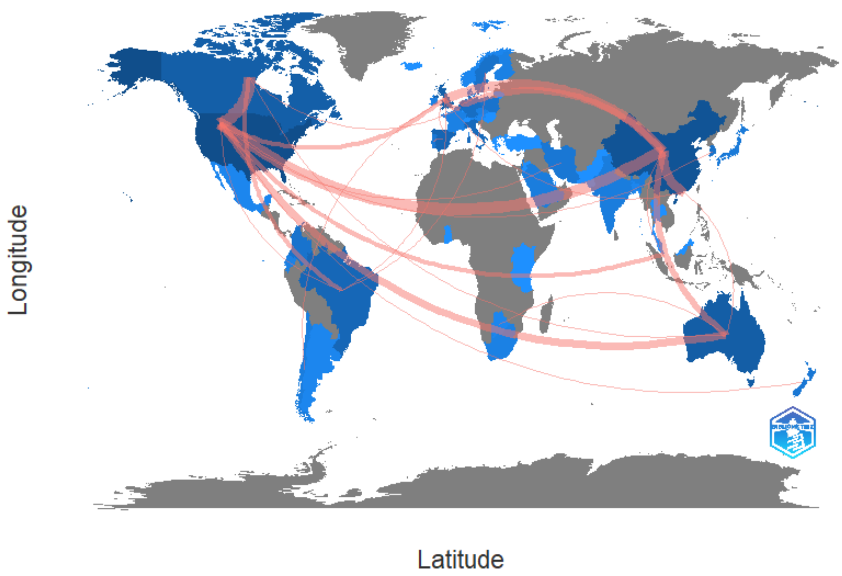


Figure 3.4: Collaboration Network

Figure (3.4) illustrates the correspondent authors' country. The countries that are in blue color represent contribution in papers that were selected in the previous stage.

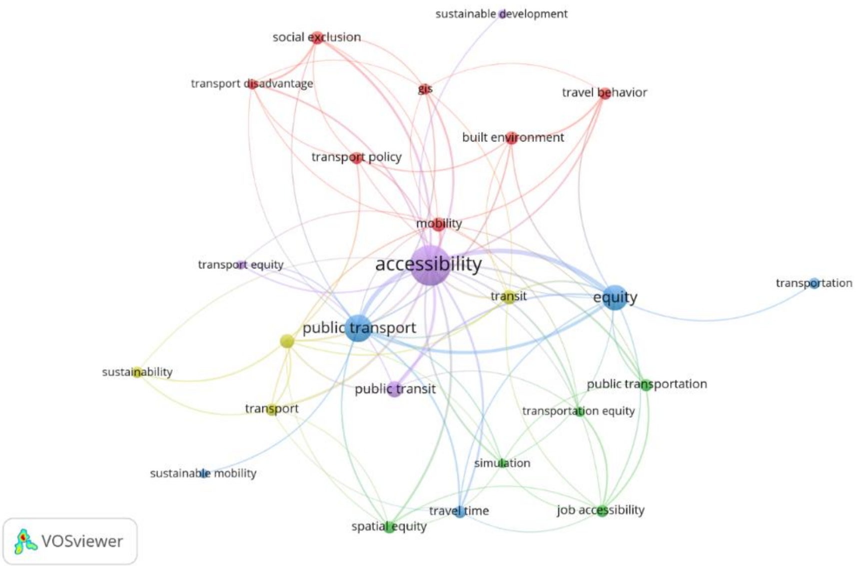


Figure 3.5: Co-occurrence Network among Authors' Keywords (occurrence ≥ 25)

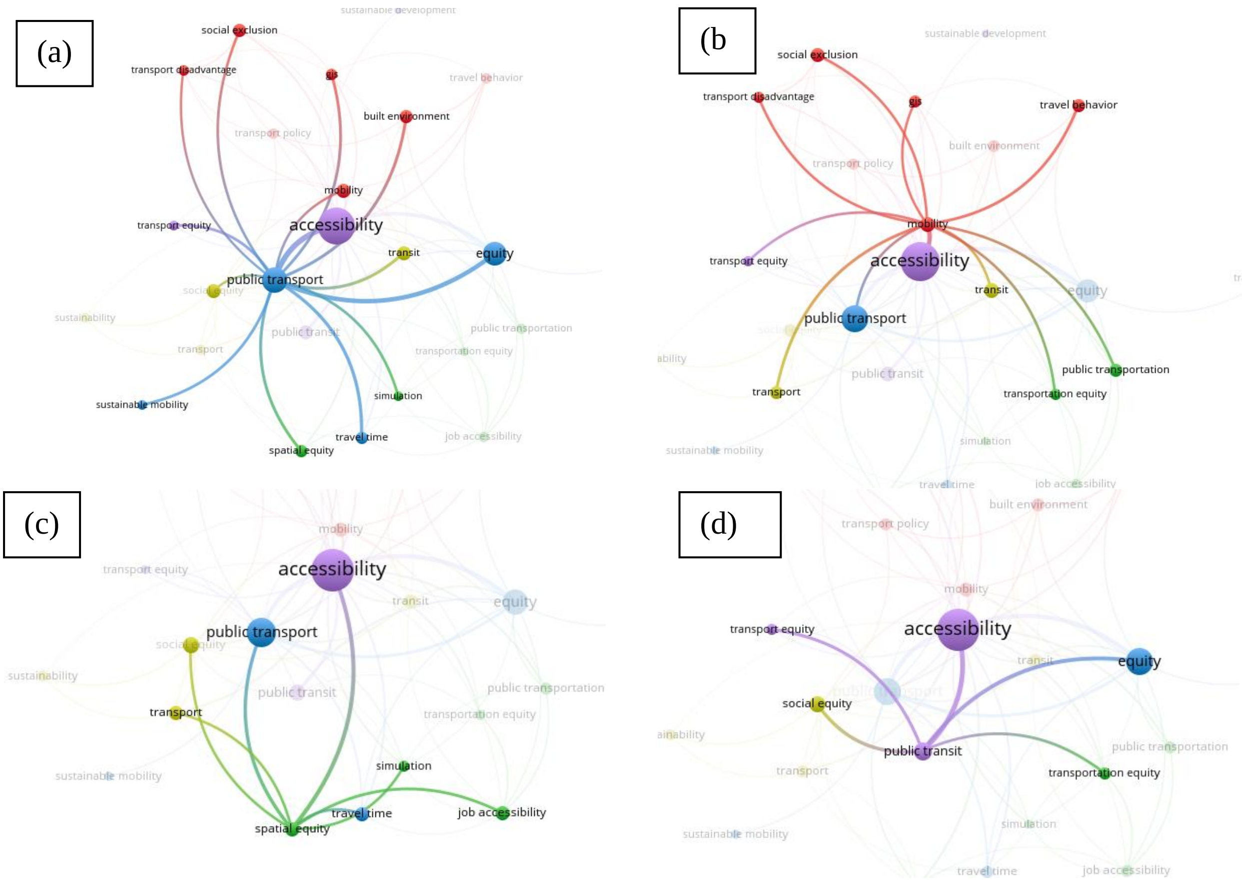


Figure 3.6: Relevant Keyword's Network

Note: In this figure, the 4 most relevant keywords for the study and their interconnection with other keywords has shown. (a) public transport, (b) mobility, (c) spatial equity and (d) public transit.

Lexical network analysis has been done to represent interconnection of keywords. In the analysis author's keywords were considered and minimum co-occurrence was set to 25. Figure (3.5) shows Co-occurrence network among authors' keywords. Also, most relevant keywords and their linking network for the study has been shown by figure 36 (a)-(d).

From the (3.6) it can be said that search result is scientifically eligible for the research aim. For each paper selected the author read the title and where it seems to be out of the scope for the current research has been discarded. Rest of the papers were collected for data extraction.

3.1.4 Data Extraction

From each paper selected author extract 4 items. Those are:

1) Theme
2) Materials and Method
3) Statistical Approach
4) Result Visualization

Extracted information from the papers has been discussed elaborately in the section 2.1.

3.1.5 Selected Studies

Finally, in this study 21 papers were selected from this process to explore spatial disparities in public transit. Along with these papers, this study also includes all the papers that were collected during extended research proposal writing.

3.2 Data

3.2.1 GTFS

Primary data for this study is the open public transit feed data (GTFS). The GTFS data was collected from (https://gtfs.de/de/feeds). They offer 3 different types of data for Germany. These are: 1) Long Distance Rail 2) Local Transit and 3) Regional Rail. For the current study local transit and regional rail data were downloaded. This data comes in *.zip* format. GTFS data of Germany contains 8 text files in *.txt* format. Table (3.2) shows details information of each text file. Though, it was supposed to be updated every week, author noticed during the study period same data was provided until 10.08.2022.

Table 3.2: Properties of the Collected GTFS Data

Feed Name	File size	Unique trips	Stops	Date
Local transit Germany	181	1.3 M	460 K	July, 10 2022
Regional trains Germany	6.7 MB	64 K	15 K	July, 10 2022

3.1.1.1 Structure of the Collected GTFS Data

It is seen from table (3.3) that each of the collected GTFS dataset has 7 *.txt* files, apart from them GTFS data also have *feed_info.txt* file. Details of each file are given below:

Table 3.3 Structure of the GTFS Files

File	rows	Field 1	Field 2	Field 3	Field 4	Field 5	Field 6	Field 7	Field 8	Field 9	Field 10
agency.txt	293* 22**	agency_id	agency_name	agency_url	agency_timezone	agency_lang					
calendar.txt	135* 128**	monday	tuesday	wednesday	thursday	friday	saturday	sunday	start_date	end_date	service_id
calendar_dates.txt	9* 9**	service_id	extension_type	date							
routes.txt	19944* 393**	route_long_name	route_short_name	agency_id	route_type	route_id					
stop_times.txt	1048576* 1001432**	trip_id	arrival_time	departure_time	stop_id	stop_sequence	pickup_type	drop_off_type			
stops.txt	450032* 15187**	stop_name	stop_id	stop_lat	stop_lon						
trips.txt	1048576* 65961**	route_id	service_id	direction_id	trip_id						

Note: Local transit Germany, Regional trains Germany***

agency.txt: It provides information about the public transport operator of this feed. They are uniquely identified with the *agency_id* field. This file also provides the timezone in the format of HH:MM:SS. It is important to the have same timezone when there are multiple agencies.

calendar.txt: This file provides service time for each day of the week. Starting and end dates of the feeds are also provided in this file. This file has a unique *service_id* that relates trips.txt file.

calendar_dates.txt: In this file, any exception of the service is provided. This allows to know considerable service variation. This file also relates the *trips.txt* file with a unique *service_id* field.

routes.txt: This file group trips into single services that are shown to users. The *route_type* field holds the mode of the transportation used for that route. Different route types available in the collected GTFS data are as follows:

Table 3.4: Route Type and Associated Transport Mode

route_type (GTFS field)	**Mode**
0	Tram, Streetcar, and Light rail. Any light rail or street level system within a metropolitan area.
1	Subway, Metro. Any underground rail system within a metropolitan area.
2	Rail. Used for intercity or long-distance travel.
3	Bus. Used for short- and long-distance bus routes.
4	Ferry. Used for short- and long-distance boat service.
7	Funicular. Any rail system designed for steep inclines.

Adopted from gtfs.org[1]

stop_times.txt: This provides the exact station sequence for each trip. *arrival_time* and *departure_time* fields provide arrival and departure time for each station. This relates the *stop.txt* file with the *stop_id* field.

stops.txt: This is a list of all available stops of the service areas served by different transport agencies. Stops are uniquely identified with the *stop_id* field. In this file, names of the stops *(stop_name)* and their geographical locations are also provided with *stop_lat* and *stop_lon* field.

trips.txt: This provides available trips of the areas served by transport agencies. This file is related to *routes.txt* file with *route_id* field and stop_times.txt file with *trip_id* field. The *service_id* field connects the *calendar.txt* and *calendar_dates.txt* files, this enables to identify in which days a specific trip operates.

[1] https://gtfs.org/schedule/reference/#routestxt (Accessed June 2022)

3.2.2 Street Network Data

Regarding street network data, for this study, it is important to model the transit service frequency in grid level. There are different sources for street network data. For this study, OSM data was collected from Geofabrik[1] for whole Germany as it is an open-source data made by volunteers and accessible without any complexity. This OSM data comes with the WGS 1984 projection.

3.2.3 Administrative Data

This study aims to visualize the spatial disparities in an interactive map. To do so, administrative data is necessary, so that users can get the information on different administrative boundary. In order to assign the transit service frequency layer to different administrative levels, "VG250" shapefile was downloaded from the Federal Agency for Cartography and Geodesy. This is available at 1:250000 scale. The dataset includes the administrative units of Germany from the country (state) to the "Gemeinden" (municipalities) level (Bundesamt für Kartographie und Geodäsie, 2021).

- Staat (country) STA
- Länder (states) LAN
- Administrative districts RBZ
- Districts KRS
- Administrative associations VWG
- Municipalities GEM

This data is updated once in a year. For this study latest available data was collected (01.01.2021). This dataset has the standard spatial reference for Germany, which is UTM projection in zone 32, Ellipsoid GRS80, Datum ETRS89 (Bundesamt für Kartographie und Geodäsie, 2021).

3.3 Data Processing

To calculate transit service frequency from GTFS data, ArcGIS Pro version 3.0.1 has been used for this study. Required license was provided from IOER. Details processing steps for the transit service frequency calculation are discussed in this section.

[1] https://download.geofabrik.de/europe/germany.html (Accessed June 2022)

System Properties

All the collected data were processed at IOER lab with the following system capabilities.

Table 3.5 System Properties

Device	PC I (ArcGIS Pro)	PC II (QGIS)
Processor	Intel(R) Core(TM) i5-8500 CPU @ 3.00GHz 3.00 GHz	Intel(R) Core(TM) i5-8400 CPU @ 2.80GHz 2.81 GHz
RAM	16 GB	16 GB
Operating System	Windows 10 Pro	Windows 10 Pro

3.3.1 Transit Service Stops Identification

It is important to include public transit stops location with their name in the visualization tool, so that users are able to know stops name when they select a specific stop. ArcGIS built-in Public Transit Tools offer transformation of stops.txt data to shapefile. GTFS Stops to Features is used in this regard to convert stops.txt file of local and regional GTFS data of Germany. The output feature class is georeferenced as WGS 1984, so it is important to change the projection. Both local and regional stops feature layers were then projected to IOER standard projection system (UTM projection in zone 32, Ellipsoid GRS80, Datum ETRS89). The following figure (3.7) is showing stops location for local and regional transport.

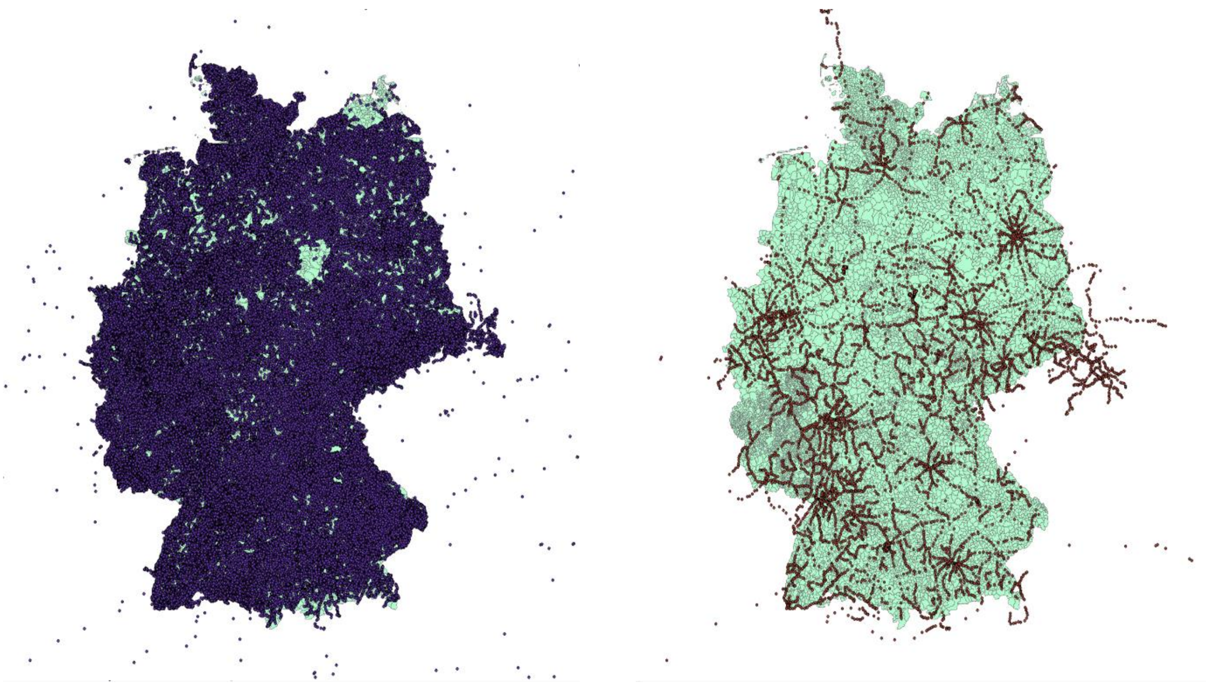


Figure 3.7: Stops Location in Germany, Note: local transport (left) & regional transport (right)

3.3.2 Public Transit Data Model

To calculate public transit service frequency in ArcGIS Pro, GTFS data must be fit in a way so that ArcGIS Pro can perform the analysis. As discussed earlier this study deals with two different GTFS data for Germany, it is necessary to create two different transit data model so as two geodatabases. So, to start the processing, two empty network geodatabases have been created for local and regional transport with defined spatial reference as the Germany and IOER standard projection system (UTM projection in zone 32, Ellipsoid GRS80, Datum ETRS89).

In this process whole GTFS folder is used as the input to create the transit network data model. The target feature dataset is the network geodatabase that was created earlier. The resulting Public Transit Network Data Model contains two feature classes (Stops, LineVariantElements) and 7 standalone tables namely:

1) Calendars
2) CalendarExceptions
3) Lines
4) LineVariants
5) Runs
6) ScheduleElements
7) Schedules

These tables are related to input GTFS data and modeled for further calculations.

The created network datasets were named identically (local & regional) so that no confusion arises in further processing.

3.3.3 Transit Service Frequency

In this study transit service frequency is modeled in two different levels. These are station level and grid levels. Following section provides detail processing steps for the modelling of transit service frequency at station and grid level.

3.3.3.1 Transit Service Frequency at Station Level

Public transit frequency can be calculated for any time frame and any day of the week as GTFS data contains 24 hours schedule of a week. Previous studies used specific time-frame for the calculation of different transit features (travel time) (Stępniak et al., 2019; Credit et al., 2021, Karner, 2018; Xin et al., 2005; Niedzielski, 2021, Owen & Levinson, 2015). To minimize the calculation time and maximize visualization tool performance, this study considered weekly and daily basis frequency. These are as follows:

Weekly Basis Frequency

Weekdays

Weekend

Daily Basis Frequency

Morning rush hour: 06:00 to 08:59

Afternoon rush hour: 14:00 to 16:59

It is possible to count arrival or departure in transit service frequency calculation. For this study author adopted departure count. In the output, prefix field AM used for morning rush hour and PM for evening rush hour. Consequently, the output of this processing step are 28 transit service frequency layers, of which 14 layers for local transport of each day of the week (Morning and Afternoon) and remaining 14 layers for regional transport.

The resulted feature layer contains number of trips visited by public transit services per hour in the selected time frame at every public transit stops.

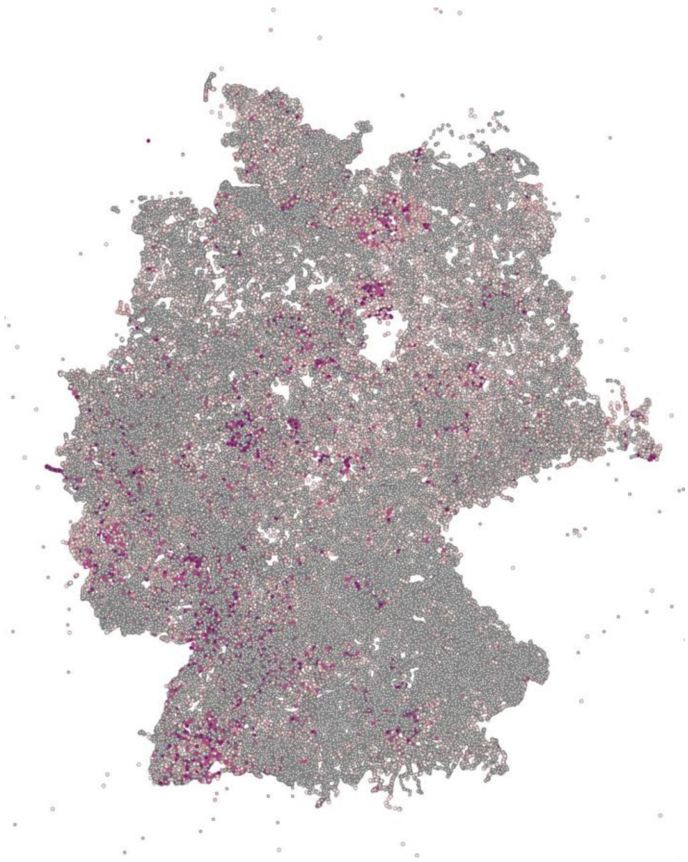
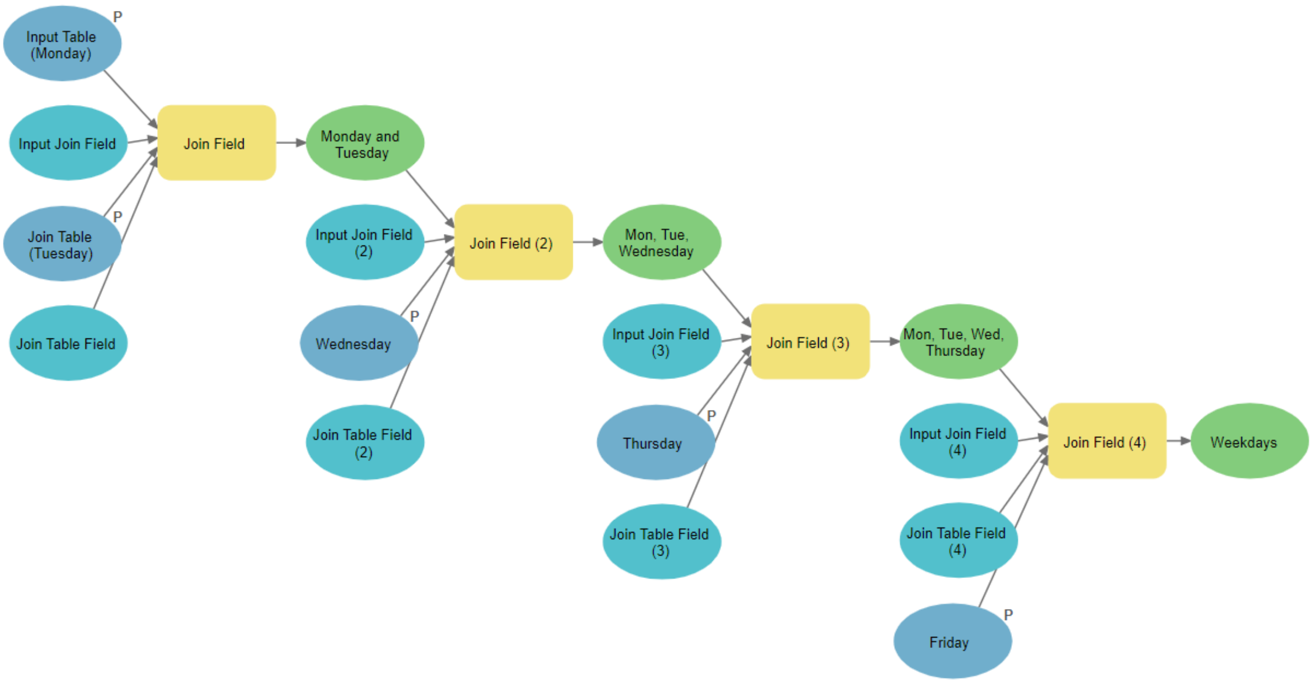


Figure 3.8: Station Level Transit Service Frequency

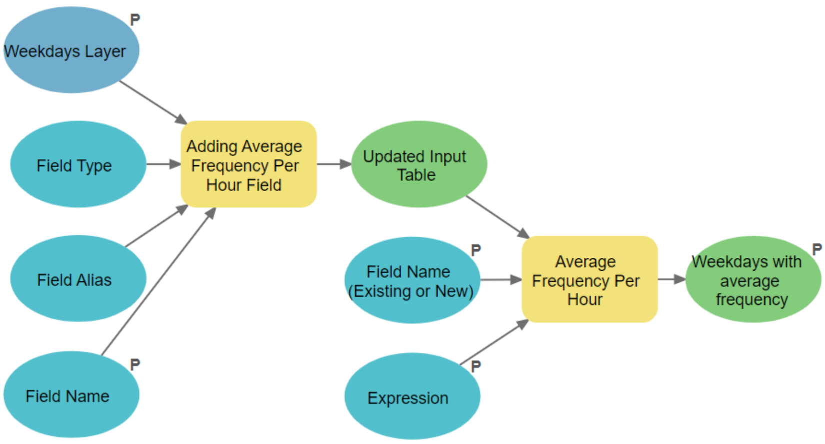
Note: From this figure (3.8) it is seen that default color palette for the visualization of transit service frequency per hour doesn't carry meaningful information. This problem has been solved and discussed in section 4.1.

3.3.3.1.1 Weekdays Layers Merge

In these processing steps the transit service frequency layers of weekdays are merged, consequently the resulted layers are reduced in number. This step saves computation time in further processing. The values of the transit service frequency layers are averaged. This step has been done with model builder in ArcGIS Pro. The following models (3.1) and (3.2) were used to merge weekday's frequency and calculate average frequency of weekdays.



Model 3.1: Merge Weekdays Frequency Layer



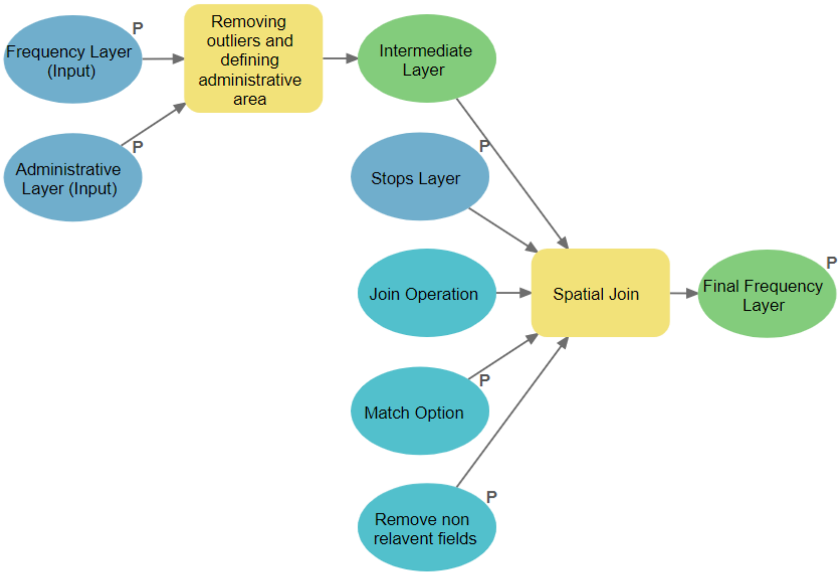
Model 3.2: Calculate Weekdays Average Frequency

3.3.3.1.2 Assignment of the Administrative Areas and Stops Name

It is seen from the figure (3.7) that there are some stops that lie outside of the Germany's administrative boundary. So, it is necessary to remove them before further processing.

In the visualization tool stops name will be shown in the pop-up. Which is why stops name has to be imported in the transit service frequency layers. Even though, section (3.3.2) generates stops feature still section (3.3.1) was carried out because stops feature from the Transit Data Model does not contain *stops_name* field.

Administrative areas also need to be included with the transit service frequency layer for user interactions. By the following model (3.3) transit stops outside of Germany's boundary has been removed as well as assignment of public transit stops name and administrative areas has been done.



Model 3.3: Assign Administrative Areas and Stops Name

Collected administrative data does not provide States and Municipalities name in same layer. Instead in **VG250_GEM** layer, *SN_L* field contains associated states number. So, in this step, States name were assigned from **VG250_LAN** with each frequency layer using ***Join Field*** in ArcGIS Pro. This tool was run in batch so that processing takes less effort and time.

3.3.3.2 Transit Service Frequency at Grid Level

To calculate transit service frequency at grid level, weekend frequency layers generated from section (3.3.3.1) and weekdays frequency layers from section (3.3.3.1.1) were used. In this section modelling of grid level transit service frequency are discussed.

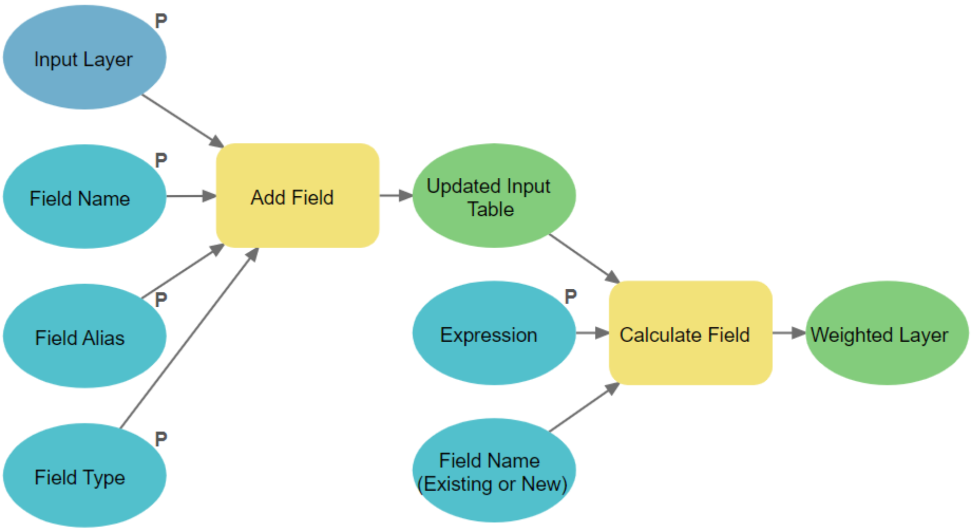
3.3.3.2.1 Assigning Weight in Transit Service Frequency

There are differences in capacities of transport modes used in local and regional transport. Ehrig (2020) used mode factor to solve this issue. Present study addresses this difference by using same mode factor for regional transport. In this step, transit service frequency of regional transport was multiplied by mode factor (equation 3.1).

Equation 3.1: Weighted Transit Service Frequency

$$F_W = F_{CH} * M.F$$

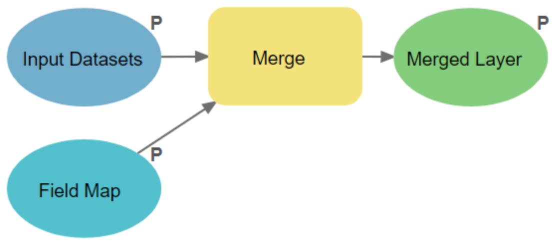
Note: In this equation, F_W refers to weighted frequency per hour for regional transport; F_{CH} refers to frequency per hour and M.F refers to mode factor which is 5 in this case.



Model 3.4: Calculate Weighted Transit Service Frequency

3.3.3.2.2 Merging Local and Regional Transit Service Frequency

For modelling transit service frequency at grid level local and weighted regional transit frequency were merged in this step. This has been done using following model (3.5).

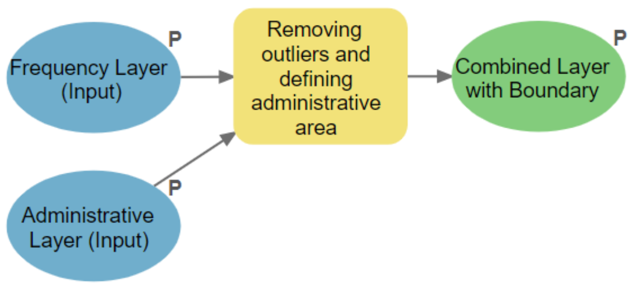


Model 3.5: Merge Local and Regional Transit Service Frequency

Note: During the model execution, field map was used to calculate sum of local and weighted values of regional transport service frequency.

3.3.3.2.3 Assignment of the Administrative Areas and Removing Outliers

In this step, administrative areas were assigned to combined frequency layers and frequencies that are outside of Germany's boundary were removed.



Model 3.6: Assign Administrative Areas and Remove Outliers

3.3.3.2.4 OSM Data Preparation

There are many available layers in OSM data, such as: highway, barrier, cycle way, track type, waterway, railway, airway, aerial way etc. (Al-Bakri, 2018). From the collected OSM dataset *gis_osm_roads_free_1* layer was used to build street network dataset. This data is then georeferenced as the IOER standard projection system used for Public Transit Data Model.

3.3.3.2.5 Modeling Distance Indicator

To model the pedestrian accessibility to the transit service stops location, origin-destination cost (distance indicator), stops locations of local and regional transport were merged.

Iso-Area as Interpolation (from Layer):

Following processing step has been done in QGIS version 3.6.0 using QNEAT3 IsoArea Algorithms (Raffler, 2018). The output is an interpolated distance that starts from multiple points (in this case stops location). The algorithm applies TIN-Interpolation method. The output is a distance-based raster layer. This method is useful for multi-criteria analysis related to accessibility1. This algorithm has been used in several studies (Demaria et al., 2022; Kucukali et al., 2022; Vendrametto Granzotti et al., 2022).

Parameters:

Network Layer: OSM filtered road data was used as the network layer.

Start Points Layer: Combined stops locations for Verkehrsverbund Oberelbe (VVO) service area was used as start point layer.

Size of Iso-Area: As discussed in previous section (2.1.2.2) a Euclidean buffer of 400m is recommended (Federal Transit Administration, 2012), previous studies used 450m (Ehrig, 2020), 700m (Fayyaz et al., 2017) and 800m (Liao et al., 2020) in their studies. This study considers 500m as the walking distance that a transit service user wants to consider as the maximum distance from his location to transit service stop. So, in this case 1000 m distance was used as the input of Iso-Area size.

Cellsize of Interpolation Raster: The cellsize depend on the analysis criteria. In this study 1000m (1KM) has been used, as modeling of the grid level transit service frequency will be done for 1KM grid.

Data Preparation:

This study proposed to calculate transit service frequency at grid level for Germany and until this step all processing was done for whole Germany. But, due to computational cost and time limitation only the VVO service area has been selected to model at station level frequency (please refer to section 5.4 for more details). The service area map was manually prepared by author, with reference to Topographic route network map[2] and *VG250_GEM, VG250_KRS* data. Combined stops location as well as OSM data were clipped using previously created VVO area layer for further processing.

[1] https://root676.github.io/IsoAreaAlgs.html (Accessed August 2022)
[2] https://www.vvo-online.de/img/VVO-Liniennetzplan-topografisch.jpg (Accessed August 2022)

OSM Data Filter:

As pedestrian do not have access to some roads, they are excluded from the calculation (Büttner, 2018). In this case “motorway”, “motorway_link”, “trunk” and “trunk_link”. The following block shows the filter used to exclude these road types.

```
"fclass" != 'motorway' AND "fclass" != 'motorway_link' AND
"fclass" != 'trunk' AND "fclass" != 'trunk_link'
```

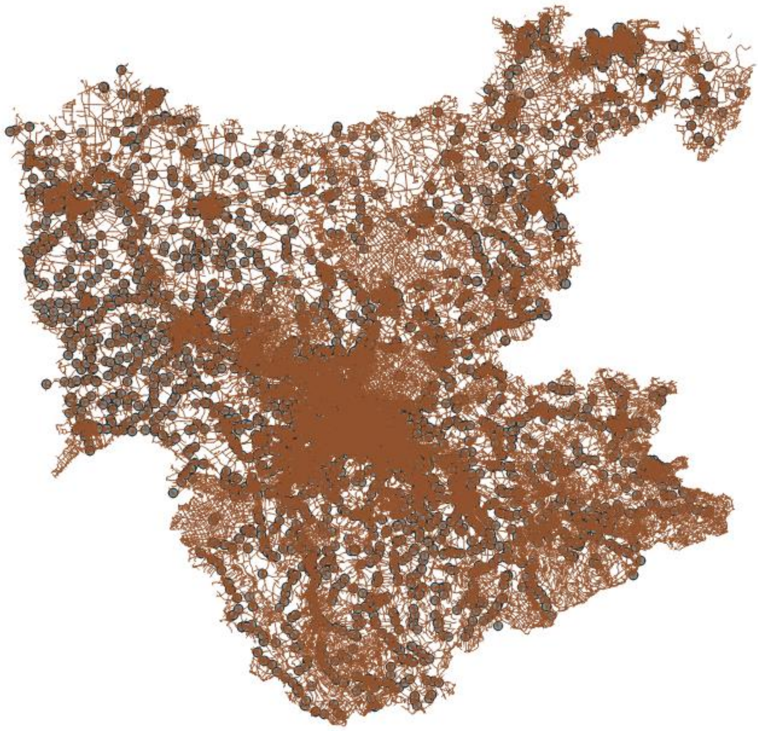


Figure 3.9: Filtered OSM Data and Transit Service Stops Location

Note: Above figure (3.9) shows the combined stops locations and filtered OSM road data at VVO service area.

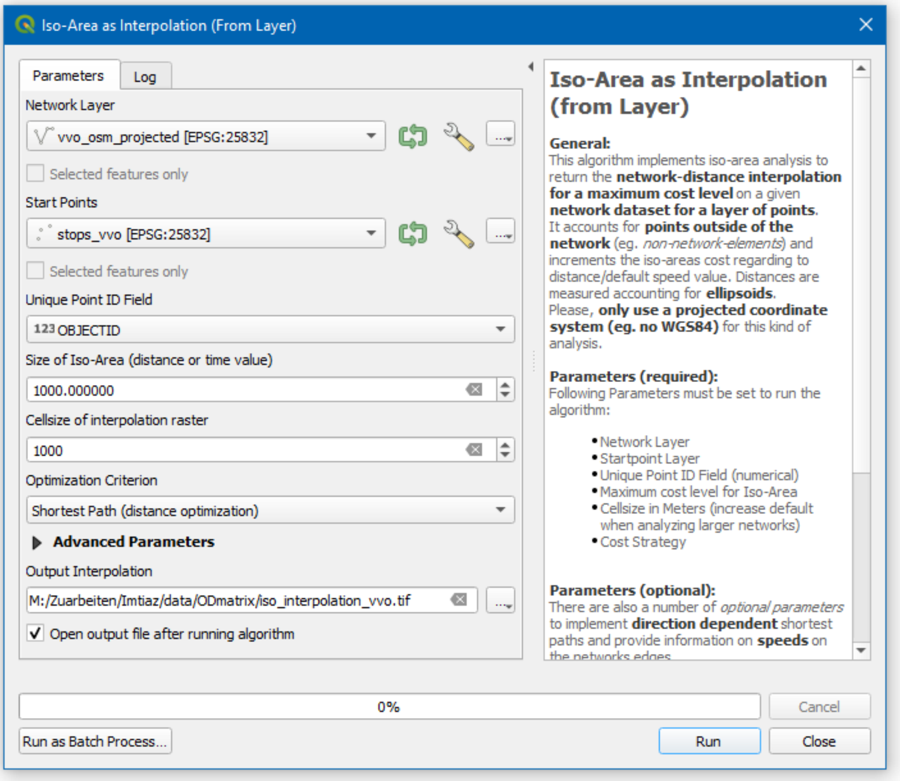


Figure 3.10: Modeling Distance Indicator

Note: Figure (3.10) shows the data and parameters of modeling Distance Indicator by ISO-Area as Interpolation

3.3.3.2.6 Data Transformation

The resulted output is a raster layer. To use the output raster in further processing, data transformation is mandatory, as ArcGIS dashboard doesn't support raster data. In this step raster output transformed into vector layer.

Float to integer: Before transforming the raster layer into vector layer, data conversion became necessary. Float values were converted into integer for further operation.

Raster to vector: In this step converted raster layer was transformed to vector layer using built QGIS function Polygonize (Raster to Vector).

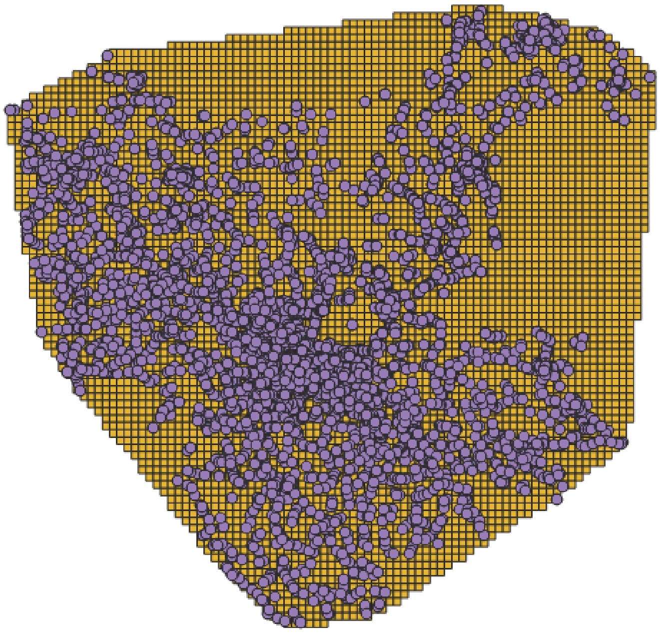
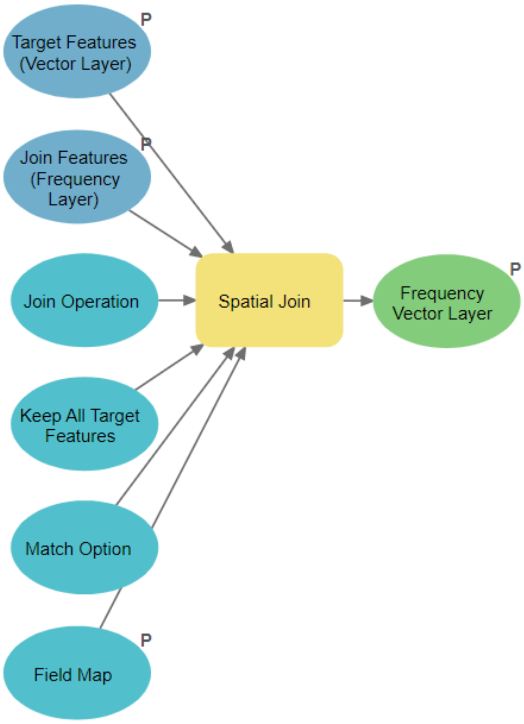


Figure 3.11: Distance Indicator Vector Layer and Stops Location

3.3.3.2.7 Joining Distance Indicator and Weighted Frequency Layer

The converted distance indicator vector layer was joined with each weighted frequency layers using following model (3.7).



Model 3.7: Join Distance Indicator and Weighted Frequency Layer

Note: During the model run, field map option was used to sum frequencies of all stops location in one grid cell.

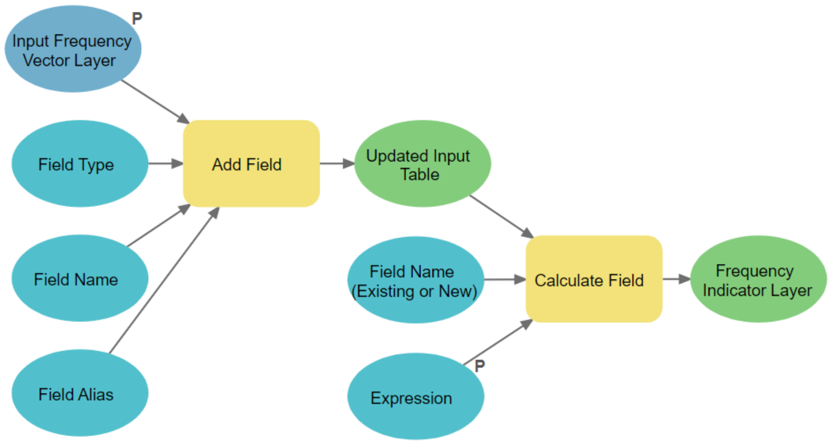
3.3.3.2.8 Transit Service Frequency Indicator

To calculate transit service Frequency Indicator the following model (3.8) has been used.

Equation 3.2: Transit Service Frequency Indicator

$$FI = \sum F_W / DI$$

Note: In this equation (3.2) FI refers to Frequency Indicator; F_W refers to weighted transit service frequency (calculated in section 3.2.2.2.1) and DI refers to Distance Indicator.



Model 3.8: Calculate Frequency Indicator

3.3.3.2.9 Disparity Index (Anselin Local Moran's I) Calculation

Studies shows that Local Moran's Index can show spatial disparities (please refer to section 2.1.3.2.1). In this study Anselin Local Moran's Index was used to analyze cluster and outliers. ArcGIS Pro built in tool ***Cluster and Outlier Analysis (Anselin Local Moran's I)*** used in this case.

Data Transformation: Moran's Index assume normally distributed values (Stieve, 2012 cited in Ehrig, 2020). To start the analysis normality is checked first. Figure (3.12) shows that the Frequency Indicator is not normally distributed. To obtain normal distribution the square root value was calculated from Frequency Indicator, still the value was not normally distributed. So square root performed for a second time. Figure (3.13) shows the normal distribution of Frequency Indicator after transformation.

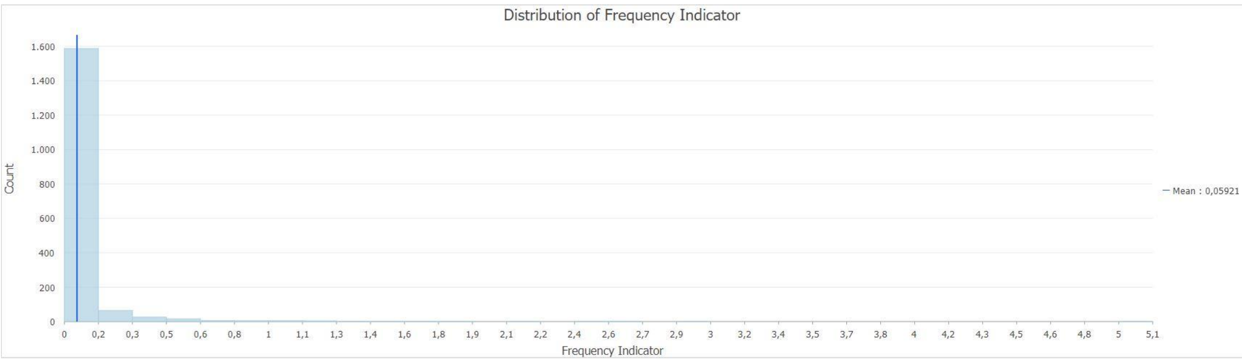


Figure 3.12: Frequency Distribution without Transformation

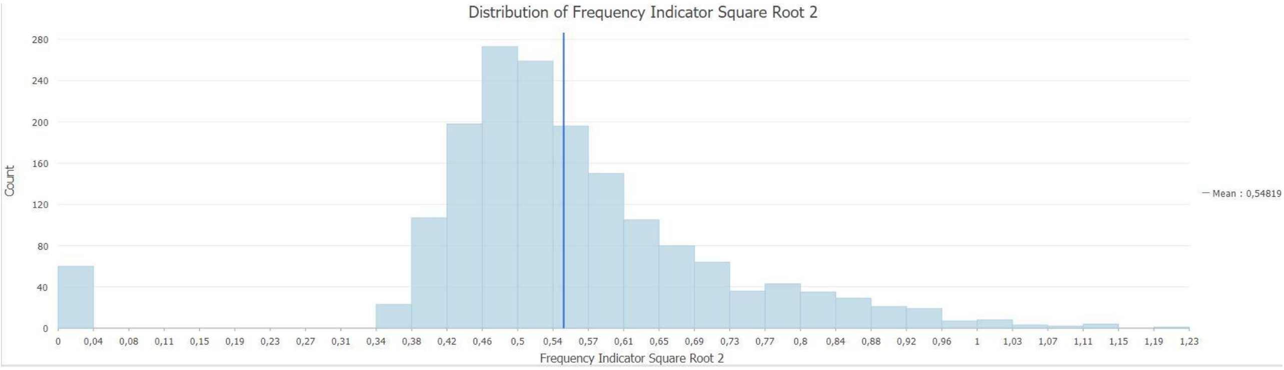


Figure 3.13: Frequency Distribution with Transformation

Parameters:

Distance Method: Euclidean

Distance Band or Threshold Distance: For polygon feature, centroids are used in distance computations. In this study, 2000m threshold was considered. Figure (3.14) represents a grid cell with its neighboring 8 cells within 2000m threshold.

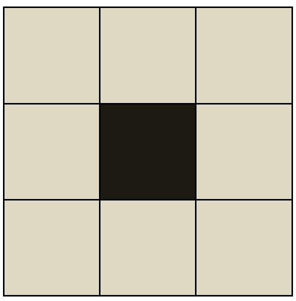
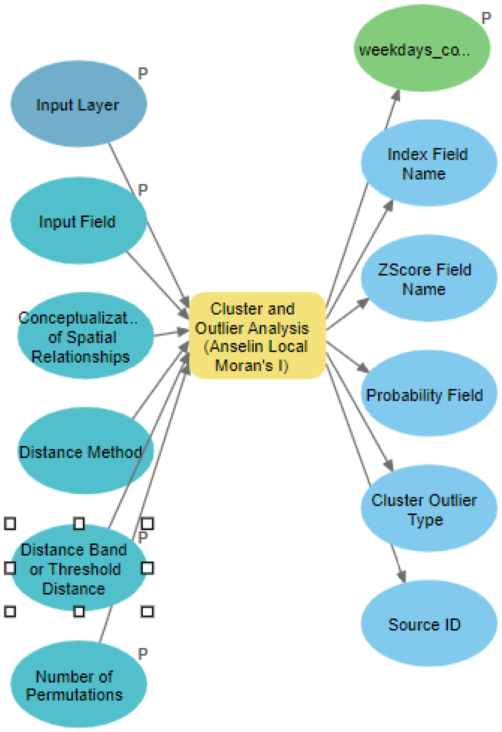


Figure 3.14: Distance Threshold for Local Moran's Index

Number of Permutations: This is used to determine how likely it would be to find the actual spatial distribution of the Frequency Indicator[1]. In this case permutations (9999) used for the calculation.



Model 3.9: Calculate Local Moran's Index

Even though the input weighted Frequency Indicator layer had administrative field, the output Local Moran's Index layer doesn't have administrative layer. So, in this step, State and Municipalities field were added to Local Moran's Index layers.

3.4 Interactive Web-map Design

It is crucial to know the design guideline to create an interactive design so that the users can manipulate according to their needs and preference of Cartographers. Web-map design was carried out from the outcomes of the theoretical background related to design principles (section 2.3) and considering existing applications (section 2.2).

[1] Cluster and Outlier Analysis (Anselin Local Moran's I) (Spatial Statistics)—ArcGIS Pro | Documentation (Accessed November 2022)

3.4.1 Expected Outcomes

After carefully reviewing literature and related visualizations author decided to include the following features in the web-map. With a view to provide necessary interactive functions the visualization application will offer following features.

a) Labels and Legends: Labels and legends will be included in the interface so that user can investigate.

b) Navigation Panel (Map Panel Name, Menu): Navigation panel will have Filter panel, Menu, Map panel name so that user can navigate through different interactivity.

c) Search and Filter Function (State, Municipalities): Search function allows the user to find any location on the map. Filter function enables the user to filter data on the basis of spatial aggregation.

d) Retrieve: The application will allow the users to see small amounts of information with pop-up.

e) Supplemental Graphs (Histogram): Supplemental graphs will show users the overall transit service frequency in municipalities level.

f) Supplemental Information (Index, Table): Index and table will allow the users to see generalized data on the basis of map extent.

g) Comparison Window: Comparison window will be included to allow the users to compare data on the basis of weekdays or different time of the day as well as within different location.

h) Customizable Interface: With this option users can manipulate the interface that will allow them more flexibility.

i) Overlay: Users will have opportunities to overlay some data to perform comparison and detailed investigation.

j) Download Summary Data: With this option users will be able to download summary data.

3.5 Evaluation of the Visualization Application

It is important to perform interaction/usability assessment activities to understand if the developed application work well or not (Robinson et al., 2005). It is a process that requires different actions such as collecting data, analyzing them, and suggesting solution (Ivory & Hearst, 2001). Traditionally usability tests are done in usability laboratory involving experts in user-interface design and testing (Abras et al., 2004). There are different approaches of interface evaluation (Roth et al., 2015; Ivory & Hearst, 2001). Roth et al., (2015) classified them into three categories. These are as follows:

a) **Expert-based methods:** Accept input and feedback from relevant expert.
b) **Theory-based methods:** Evaluate using theoretical framework.
c) **User-based methods:** Conduct user survey and accept input and feedback.

Expert-based evaluation: This method include expert in user interaction to analytically evaluate the developed interface and make decision about the usability design. The expert makes suggestion to improve the interface in case the developed design violate guidelines (Nielson, 1993 cited in Hix et al., 1999). This is an effective method in the cartographic design process, particularly for interactive map (Kostelnick et al., 2008). This method was used in different research (Roth et al., 2015; Kostelnick et al., 2008; Hix et al., 1999; Richard & Egenhofer, 1995). In this study the author has adopted expert-based methods to evaluate the visualization tool. This method is particularly helpful when small set of experts are available (Roth et al., 2015).

This section provides more details on the method of survey, participant information and evaluation method.

Participants: For the qualitative evaluation of the designed interface two participants were selected from IOER working on similar background (Visualization and Cartography). It is important select someone which is not part of the team (Roth et al., 2015). These participants were from same research area of IOER. But they were not communicated during the development phase, so they had no prior knowledge about the visualization tool.

Assessment method: The author has participated in a discussion which started with short presentation, following the similar approach in Kostelnick et al. (2008) and Richard & Egenhofer (1995). Presentation also included demo of designed interface. After the presentation the participants were involved in open discussions and gave feedback. They also had chance to explore the visualization tool. However, feedbacks were also collected through a questionnaire (please see appendix for more details). The participants were asked 5 questions about utility and 12 questions about usability of the application. There were 7 negative questions out of 17. These questions were constructed to get ratings on the scale of 1 to 5 where 1 represent disagree and 5 means agree.

Chapter Four

INTERACTIVE WEB-MAP DESIGN AND EVALUATION

In this chapter the methodology for creating interactive web-map-based visualization application is discussed in detail. The outlook of the designed application also explained in this section. Finally, the outcome of the expert-based evaluation is presented in this chapter.

4.1 Data Design

The processed data, described in section 3.3 were uploaded into ArcGIS online server so that they can be used to create ArcGIS Dashboard based visualization tool. This tool is named as *PubTraDis Visualization* which stands for Public Transit Disparity Visualization.

4.1.1 Web-map Design

ArcGIS Dashboard require web-map when it comes to map visualization. In this step four different web-maps were created so that they can be integrated into Dashboard.
These are:

a) Local (shows station level transit service frequency per hour for local transport),
b) Regional (shows station level transit service frequency per hour for regional transport),
c) Raster (1KM) (frequency indicator in 1KM raster grid),
d) Disparity Index (Local Moran's CO Type in 1KM raster grid).

4.1.1.1 Group Layer

To allow more user flexibility, layers are grouped into two. These layers can be turn on or off individually or by group. The following table shows the layer grouping for different web maps.

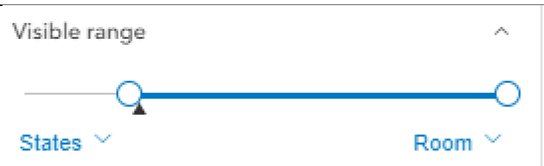
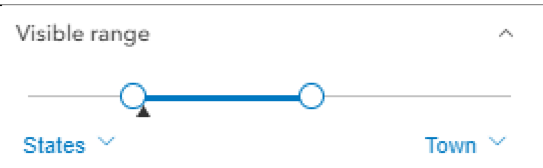
Table 4.1: Layer Groups in PubTraDis Visualization

Local & Regional map	Raster (1KM) map	Disparity Index map
▽ Dot Map	▽ Combined Frequency	▽ Combined Frequency
Weekdays Morning	Weekdays Morning	Weekdays Morning
Weekdays Afternoon	Weekdays Afternoon	Weekdays Afternoon
Saturday Morning	Saturday Morning	Saturday Morning
Saturday Afternoon	Saturday Afternoon	Saturday Afternoon
Sunday Morning	Sunday Morning	Sunday Morning
Sunday Afternoon	Sunday Afternoon	Sunday Afternoon
▽ Heat Map	▽ Frequency Indicator	▽ Local Moran's Index
Weekdays Morning	Weekdays Morning	Weekdays Morning
Weekdays Afternoon	Weekdays Afternoon	Weekdays Afternoon
Saturday Morning	Saturday Morning	Saturday Morning
Saturday Afternoon	Saturday Afternoon	Saturday Afternoon
Sunday Morning	Sunday Morning	Sunday Morning
Sunday Afternoon	Sunday Afternoon	Sunday Afternoon

4.1.1.2 Layer Visibility Range

In *PubTraDis Visualization* two different layers visibility range was set for Local and Regional Map. As not all data requires same zoom level and therefore doesn't need to be displayed in all levels[1], dot map and heat map layers visibility were set in different level.

Table 4.2 Visibility Range in PubTraDis Visualization

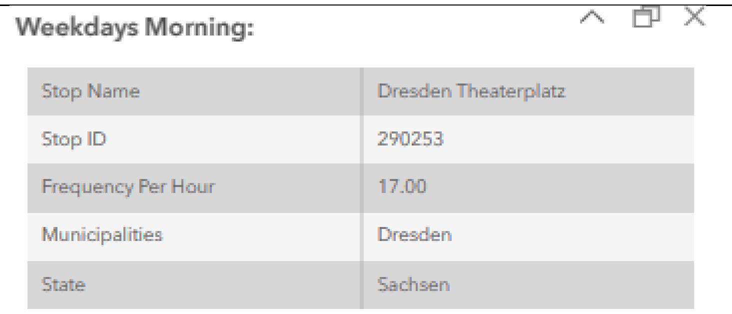
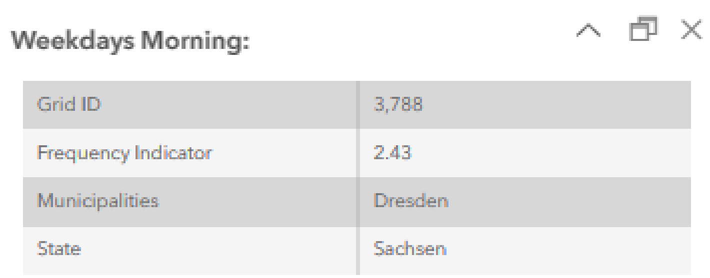
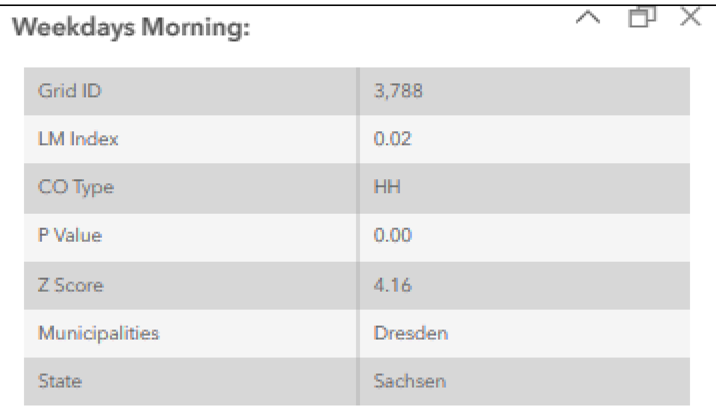
Dot Map	Heat Map
	

[1] https://doc.arcgis.com/en/arcgis-online/reference/set-visibility.htm (Accessed November 2022)

4.1.1.3 Pop-up Design

Not all the information is equally important to the users. So, to configure the pop-up only relevant information were chosen during the pop-up configuration. The following table shows the chosen field in each layer group for pop-up design.

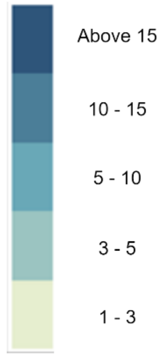
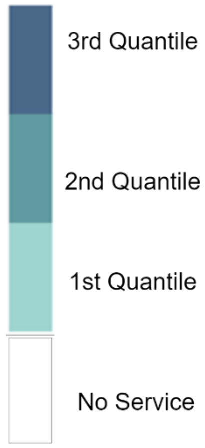
Table 4.3: Pop-up Design in PubTraDis Visualization

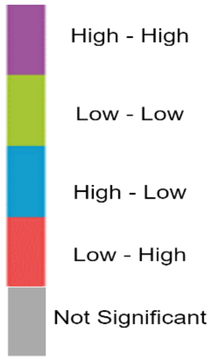
<table>
<tr><td>Frequency Layer: Local, Regional
Combined Frequency Layer: Raster (1KM), Disparity Index</td><td>
</td></tr>
<tr><td>Frequency Indicator Layer: Raster (1KM)</td><td>
</td></tr>
<tr><td>Local Moran’s Index Layer: Disparity Index</td><td>
</td></tr>
</table>

4.2 Data Visualization

As discussed in section (4.1.1) *PubTraDis Visualization* has 4 different maps groups, which represent different data with different visualization styles. The following table shows the visualization style applied in different layers and associated visual variable with their color scheme. To make the data understandable to the users, different data classification methods were chosen with respective to data properties. For example, data were manually classified in to 5 class (less than 1 transit service per hour excluded) for Dot Map. On the other hand, quantile-based classification was chosen for Frequency Indicator layers.

Table 4.4: Visualization Style Used in PubTraDis Visualization

Visualization Style	Layer Name	Map Name	Visual variable	Color scheme
Color (Sequential)	Dot map	Local, Regional	Color	
Color (Sequential)	Heat map	Local, Regional	Color	
Color (Sequential)	Frequency Indicator	Raster (1KM)	Color	

Color (Diverging)	Local Moran's Index	Disparity Index	Color	

4.2.1 Map Visualization

4.2.1.1 Dot Map

PubTraDis Visualization offers 12 dot maps (with visibility range from States-Room) visualizing station level transit service frequency for local and regional transport. To reduce visual density at smaller scales, this tool adopted scale-based symbol sizing while retaining an appropriate relative size at larger scales.

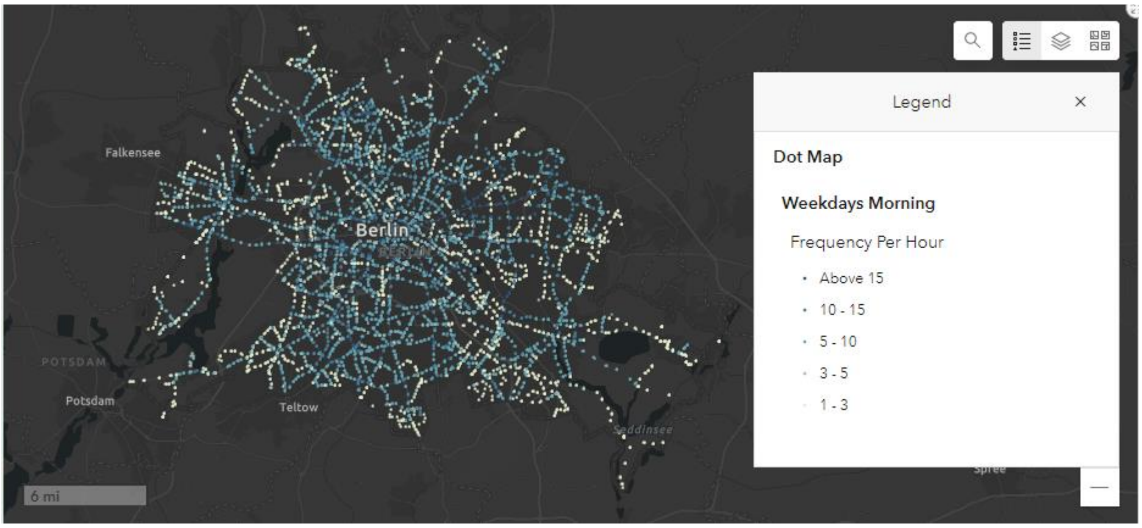


Figure 4.1: Dot Map Showing Berlin during Weekdays Morning

Note: This dot map is showing transit service frequency per hour for local transport in Berlin, Germany for weekdays morning.

4.2.1.2 Heat Map

Transit service frequency are calculated along the station level for the whole Germany, therefore Heat map in *PubTraDis Visualization* will allow the users to have an overview of spatial disparities among different areas. There are in total 12 heat maps along with visibility range from

States-Town. Layer blending was kept normal, and transparency was set to 30% for all heat map layers.

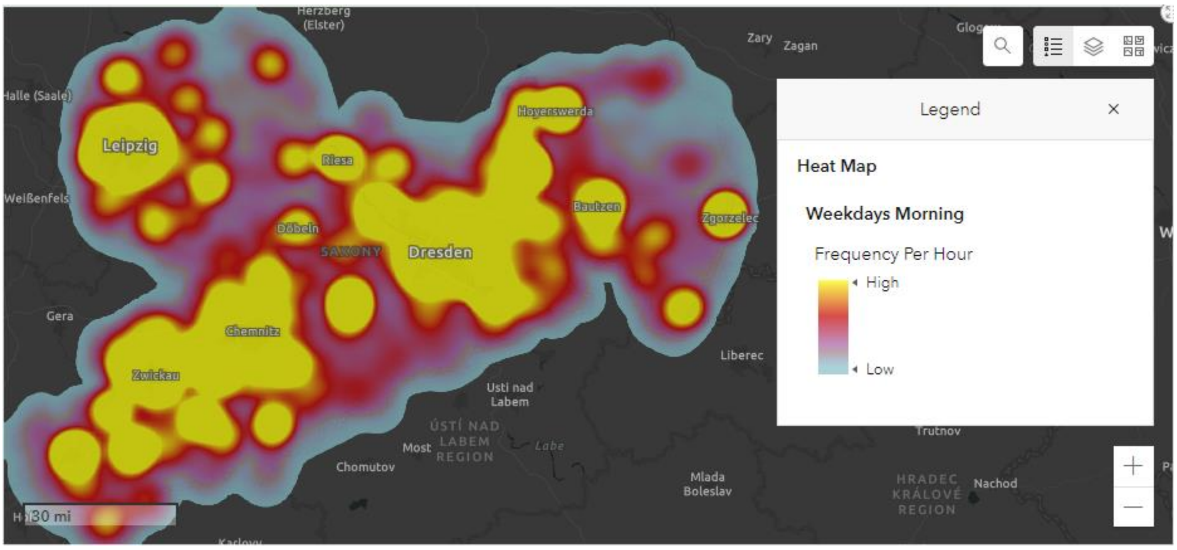


Figure 4.2: Heat Map Showing Saxony during Weekdays Morning

Note: This heat map showing transit service frequency per hour for local transport in Saxony state, Germany for weekdays morning.

4.2.1.3 Choropleth Map

PubTraDis Visualization offers 12 choropleth maps for the users. 6 of them visualize Frequency Indicator in Raster (1KM) map and others show Local Moran's Cluster/Outlier (CO) Type in Disparity Index map group.

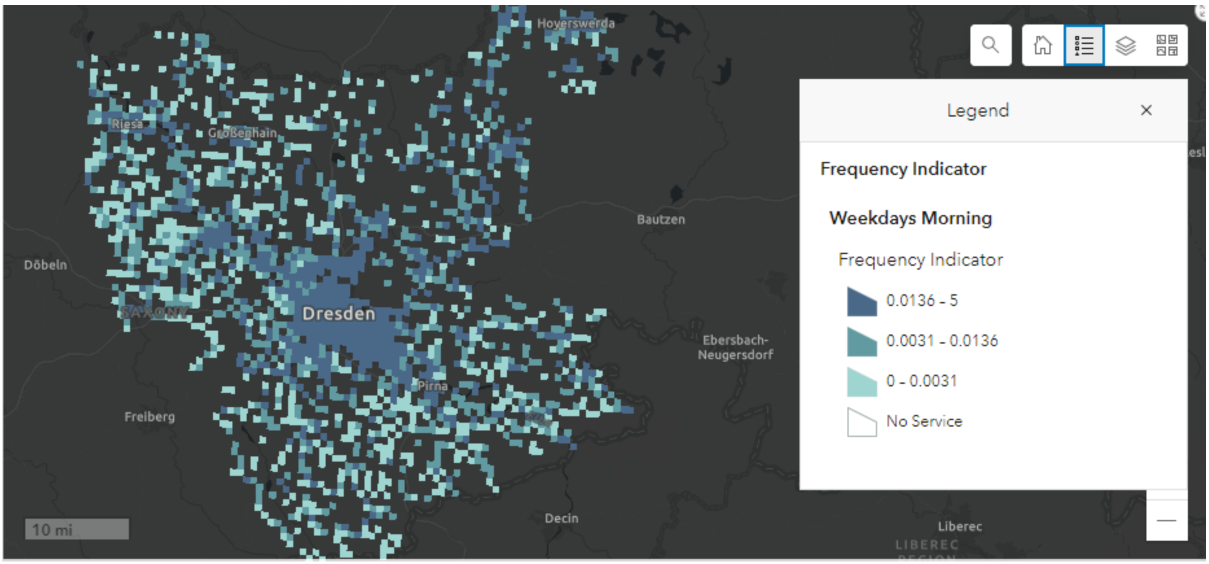


Figure 4.3: Choropleth Map Showing Frequency Indicator

Note: This choropleth map is showing frequency indicator in VVO area for weekdays morning.

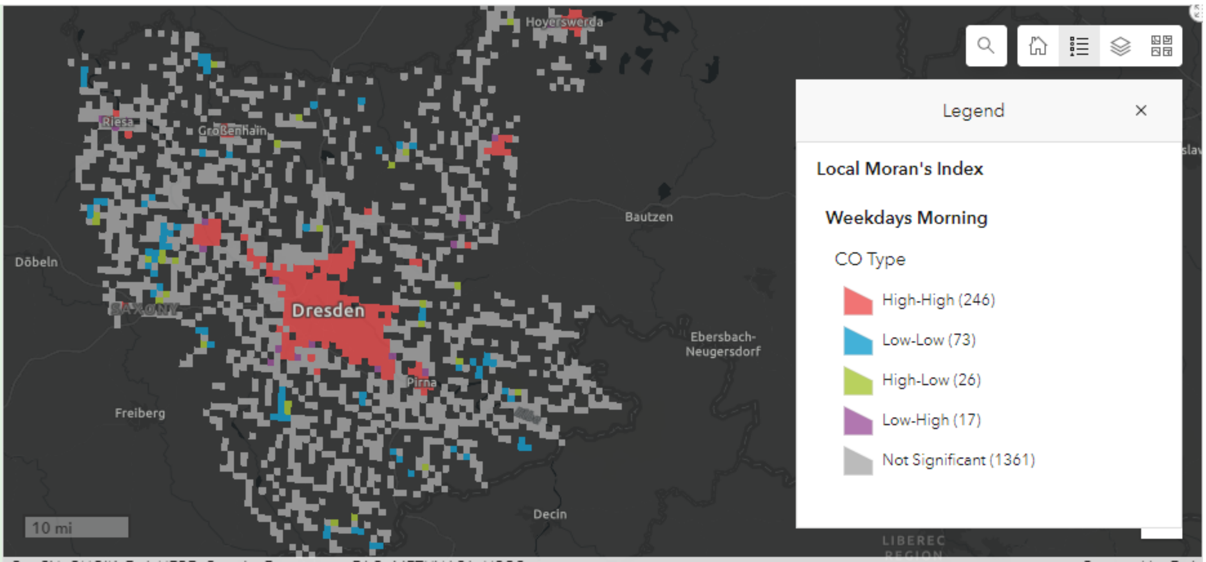


Figure 4.4: Choropleth Map Showing CO Type

Note: This choropleth map is showing CO Type in VVO area for weekdays morning.

4.3 Data Aggregation

4.3.1 Spatial Aggregation (Map)

The interface offers 24 indicator tabs, which are by default, aggregated along the map interface. The aggregation dynamically changes when users perform zooming, panning, or scrolling.

4.3.1.1 Administrative Boundary (State, Municipalities) Based Aggregation

PubTraDis Visualization will allow the user to see different degrees of details. Processed transit service dataset is aggregated based on administrative level. As originally the data is for whole Germany, it is primarily aggregated on country level which provide overview of visualized dataset. Further aggregation has done based on State and Municipalities level. For the aggregation field related to states and municipalities were used. This boundary-based aggregation allows the users to use two different filter options to select states and municipalities in *PubTraDis Visualization.*

Administrative boundary based also produce 12 charts and 12 tables in interface. These has been done using municipalities field alias in the associated layers.

4.3.1.2 Grid Level Aggregation

Transit service frequency per hour for local and regional transport have been aggregated in Grid level as well (please refer to section 3.3.3.2 for details methodologies). The Frequency Indicator and Disparity Index map group *PubTraDis Visualization* shows the Grid level aggregated data.

4.4 User Interface

This section discusses about the interface of the *PubTraDis Visualization*, accessible from *(https://tinyurl.com/2pmz6hw7*). Screenshots are added from the interface to provide examples.

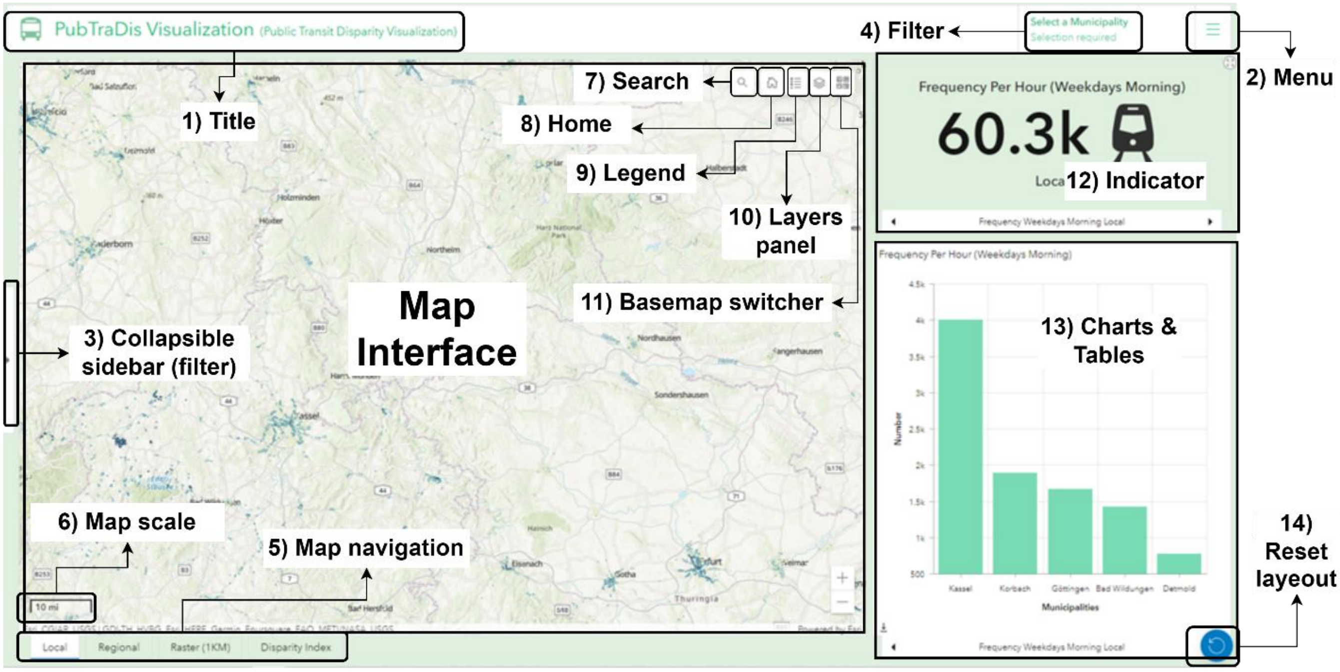


Figure 4.5: User Interface of PubTraDis Visualization

1) Title: Title of the tool is placed on the top-left corner of the interface.

2) Menu: Menu is placed on top-right corner of the interface. In the menu, user can choose Comparison window, details information of the tool from About PubTraDis & User Guide. Comparison window will take the user to a new interface where they can perform a comparison between two map elements. Details are discussed in section 4.5 (i).

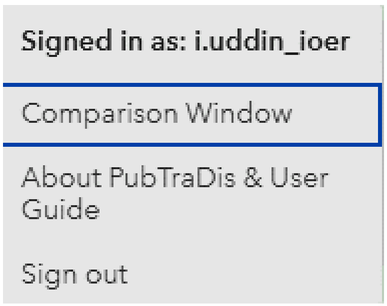


Figure 4.6: Menu in PubTraDis Visualization

3) Collapsible Sidebar (Filter): This is placed in mid-left section of the interface. This will allow the user to filter States.

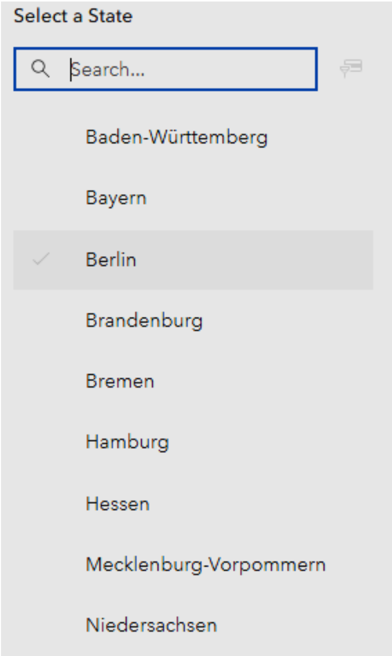


Figure 4.7: Collapsible Sidebar in PubTraDis Visualization

4) Filter: On the left side of the menu, there is a filter option placed. User can select desired municipalities from this filter option.

5) Map Navigation: Map navigation is placed at lower left corner of the interface. From this panel user can switch between Local, Regional, Raster (1KM) and Disparity Index map.

Map Interface: Map Interface comprises with elements 6 to 11.

6) Map Scale: Map scale is placed just above the map navigation.

7) Search: Search is placed at the upper side of the map interface (4th from the right to left side).

8) Home: Home is placed at the right side of the search option.

9) Legend: Legend is just right side of the home option.

10) Layers Panel: Layers panel is placed at the right side of the legend option.

11) Basemap Switcher: Basemap switcher is situated at the right side of the layers panel.

12) Indicator: Indicator is placed below filter and menu options. There are 24 tabs for indicator. These are named after the information shown by them. For example, Frequency Indicator Weekdays Morning Local means the number of total frequency per hour of the map area (subject to change when a filter is active). Users need to select the respective day and time manually to investigate the relation in the map layer.

Indicator Parameters:

Frequency: Total number of frequency per hour
Frequency Indicator: Frequency Indicator (average)
Local Moran's index: Local Moran's Index (average) (Greater than or equal to 0)

13) Charts & Tables: Charts and Tables are placed beneath Indicator in the interface. There are 12 charts and 12 tables in the interface. Each chart shows the sum of frequency per hour by municipalities in the map interface (subject to change when a filter is active). Only the top 5 are shown here. The charts names are identical to the transit service frequency layers. For example, Frequency Weekdays Morning Local represent transit service frequency of local transport during weekday's morning.

The frequency Indicator table provides statistics of the average Frequency Indicator at the grid level and is grouped by Municipalities in the map interface (subject to change when a filter is active). The table names are also identical to the related layers. For example, Frequency Indicator Weekday's Morning represent Frequency Indicator Weekday's Morning layer information.

Table Parameters:
Local Moran's Index table provides statistics of LM Index (Average) (Greater than or equal to 0), Z Score (Average), and P Value (Average) grouped by Municipalities in map interface (subject to change when a filter is active).

14) Reset Layout: This is placed at bottom right corner of the interface.

4.5 Interactive Elements

a) **Zoom-in & Zoom-out** Users can zoom in and zoom out in *PubTraDis Visualization*. They can either scroll or use icon + — from the map interface.

b) Default Map View: This option will allow the user to return default map view of *PubTraDis Visualization*.

c) Retrieve (Pop-up): *PubTraDis Visualization* allows user for deeper examination by pop-up. When they click a station or grid in the map interface, a pop-up window shows details about that feature (please refer to table no 4.3 for more details).

d) Search and Filter (State & Municipalities): Users can search any place in *PubTraDis Visualization* with search option.

State Selector (Collapsible Bar): Users can select one or more States at a time. To deselect any State, they can double-click on that State name or use the reset option.

Municipality Selector: It will be activated after at least one state is selected. Users can select only one municipality at a time. For example, with the Bavaria state selected users can select any municipality from Bavaria.

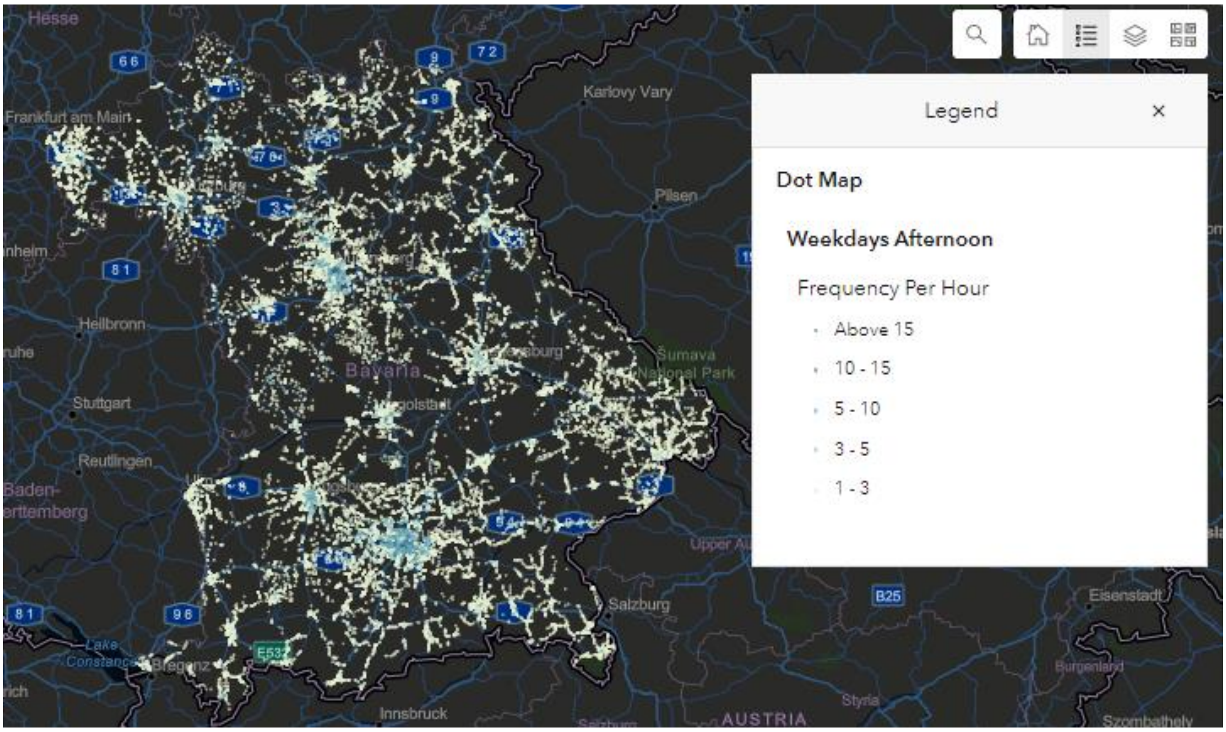


Figure 4.8: Transit Service Frequency in Bavaria.

Note: Above figure (4.8) is showing transit service frequency per hour in Bavaria during weekdays morning.

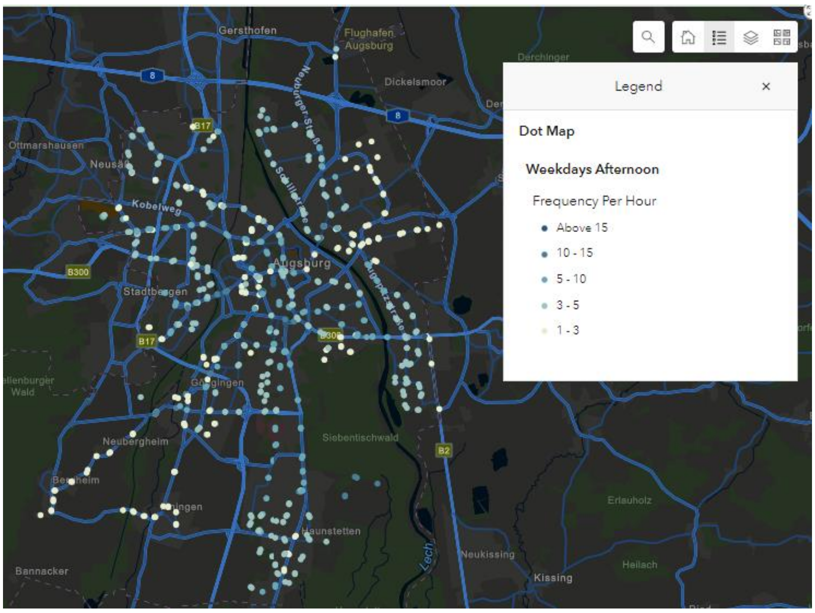


Figure 4.9: Transit Service Frequency in Augsburg, Bavaria.

Note: Above figure (4.9) is showing transit service frequency per hour in Augsburg, Bavaria during weekdays morning.

e) Layer Visibility: *PubTraDis Visualization* offers the user to turn on and off any layer group as well as an individual layer within the group.

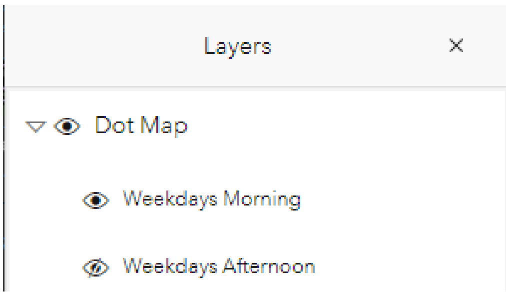


Figure 4.10: Layer Visibility Option in PubTraDis Visualization

f) Basemap Switcher: *PubTraDis Visualization* offers users to select a number of basemap. This enables them to investigate the data with more flexibility. This is particularly helpful for performing overlay.

g) Overlay: Some of the data are possible to overlay on top of one another. For example, in the "Raster (1KM)" map users can overlay the combined frequency layer on top of the Frequency Indicator layer. This helps them to investigate the changes of Frequency Indicator in different grid. Same goes for "Disparity Index" map.

Figure 4.11: Overlay of Combined Frequency and Local Moran's Index layer

Note: Above figure (4.11) is showing the overlay of Combine Frequency and Local Moran's Index layer in Disparity Index map group. With this option users will be able to see the combined transit service frequency per hour of each station in certain grid cell. And relate the disparity index.

h) Customizable Interface: The interface is customizable. Users can change the size of the map, index, and chart tab.

i) Comparison Interface: Comparison interface is one of the important interactivities available in *PubTraDis Visualization*. The following figure (4.12) is showing the interface of comparison window.

Comparison: Users can select any layer from the left map interface and either the same layer or a different layer from the right map interface depending on their comparison criteria.

Filter: State and Municipalities filter also independent in both maps. For example, they can select Berlin on the left map and Munich on the right map. Like the main interface of *PubTraDis Visualization*, Municipality filter of comparison window activate after State selection and depends on the selected State.

To investigate all information related to a map layer users need to select the same name in the indicator and charts/table panel. These names contain either left or right which is identical to the map. For example, a user wants to investigate Weekdays Morning Frequency Indicator for VVO on the left map and Saturday Morning Frequency Indicator on the right map. Then user need to turn on the layer Weekdays Morning in Frequency Indicator group layer from the left side "Raster (1KM)" map tab and Saturday Morning in Frequency Indicator group layer from the right side "Raster (1KM)" map. From the charts/table tab they also need to select Frequency Indicator Weekdays Morning (Left) for the left map statistics and Frequency

Indicator Saturday Morning (Right) for the right map statistics. The same goes for the indicator tab.

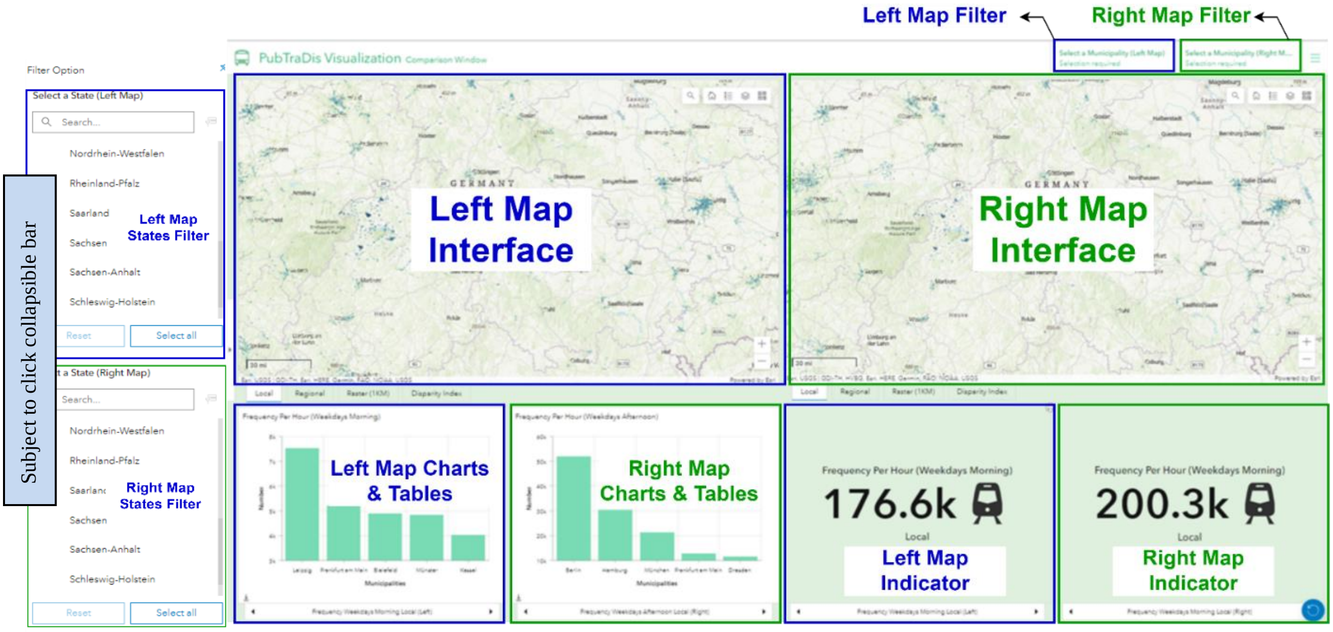


Figure 4.12: Comparison Window in PubTraDis Visualization

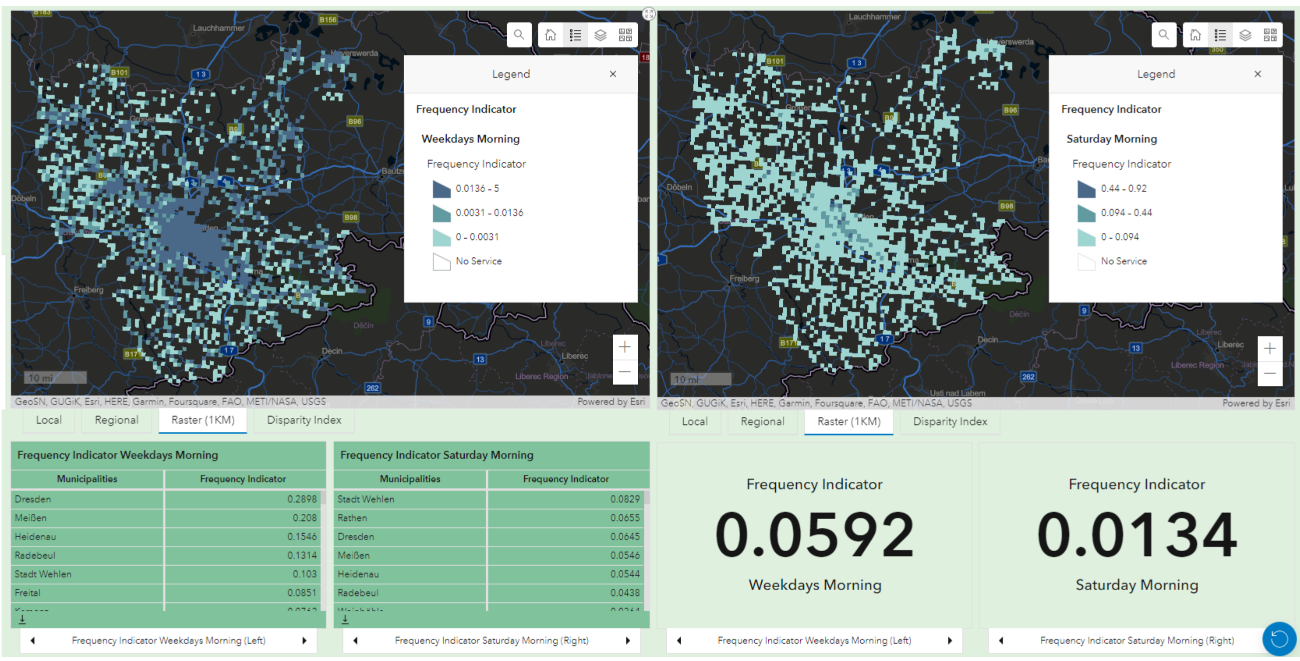


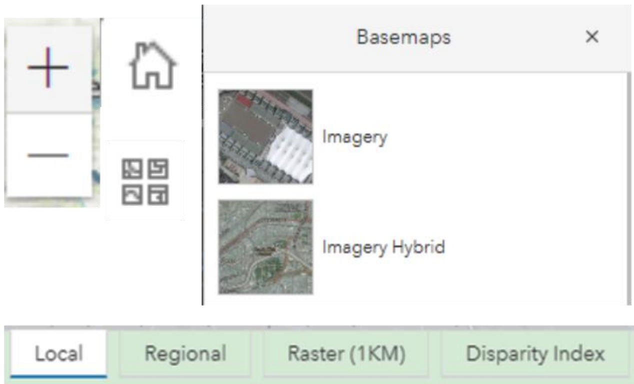
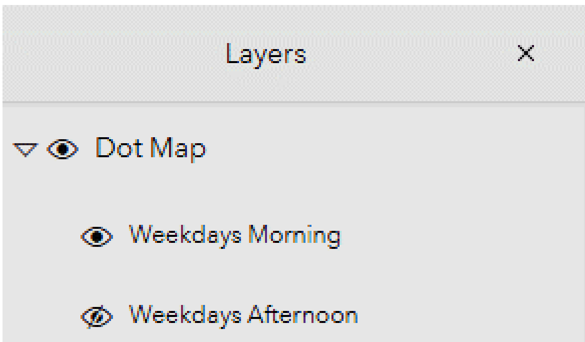
Figure 4.13: Comparison between Weekdays and Weekend

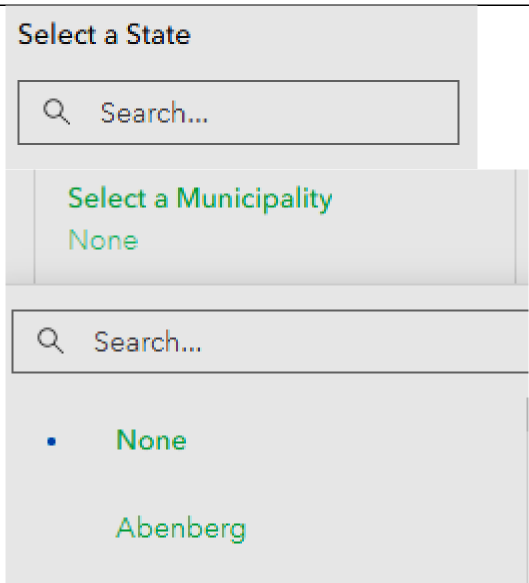
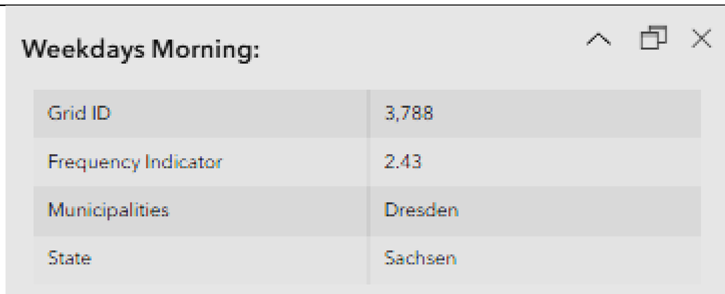
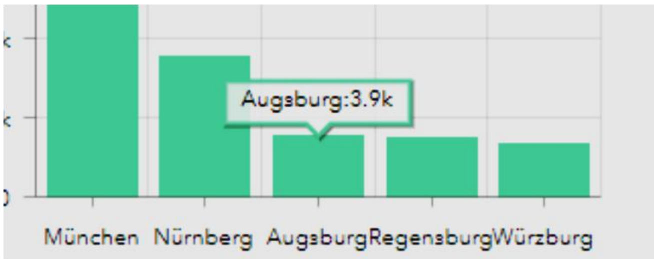
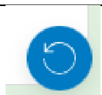
Note: Above figure (4.13) is showing the comparison window in *PubTraDis Visualization*. In this comparison, Weekdays Morning Frequency Indicator is selected on left map and Saturday

Morning Frequency Indicator is selected on the right map as well as associated tables and indicators tab.

j) Reset Layout: Reset layout  option allows the users to reset the interface after any changes made by them. This will also reset any filter or selection made.

Table 4.5: Interactive Functionalities in PubTraDis Visualization

Tasks	**Interaction**	**Snapshot from PubTraDis Visualization**
Zoom	Clicking, Scrolling,	
Default map view	Clicking	
Navigation	Clicking	
Basemap switch		
Layer visibility	Clicking	
Search	Typing	

Filter	Clicking, typing	
Retrieve	Clicking, hovering	
Overlay	Clicking, relating	
Customizable interface	Clicking	
Comparison	Clicking, zooming, relating	
Reset layout	Clicking	

4.6 Evaluation Outcomes

4.6.1 General Information of the Participants

The participants have more than 10 years of experience in the field. They were asked about the frequency of interactive web-based map use. Results shows that both of them uses interactive web-based map at weekly basis.

4.6.2 Interface Interaction

The participants had chance to interact with the interface. They were asked five questions on the basis of their interaction.

4.6.2.1 General Outlook

They were asked about the preference of general settings of the interface. It is seen from the figure (4.14) that both of them prefer map at left and supplementary information at right which is also present outlook of the interface.

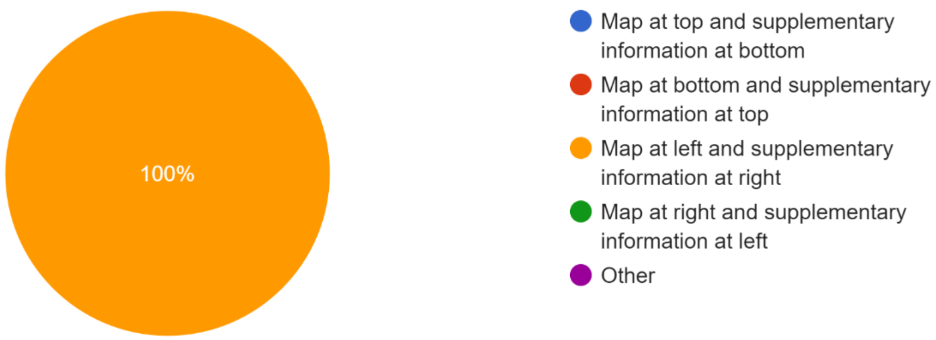


Figure 4.14: Interface Outlook Preference

4.6.2.2 Visualization Style

Mixed response found from the participants about the visualization style. One of them prefer only Dot Map for the station level frequency. On the other hand, another opined that he would prefer both Dot map and Heat map for the station level frequency.

For the visualization of Disparity Index one participant agree prefer the current CO Type map visualization and another opined that he would prefer Local Moran Index (LM Index) map instead of CO Type map.

4.6.3 Utility and Usability Rating

As the written questionnaire formulated to get feedback from the participants, the following section shows the outcome of the evaluation.

4.6.3.1 Utility

In this evaluation, 4 positive and 1 negative question were asked to get insights about the utility of the *PubTraDis Visualization*. The following table 4.6 shows the result of utility rating.

Table 4.6: Participants' Response to Utility Ratings

S.N.	Utility Rating	1	2	3	4	5	Average
		Disagree			Agree		
1	I would use *PubTraDis Visualization* frequently for my work		1	1			2.5
2	This tool would contribute to spatial disparity analysis in public transit research					2	5
3	*PubTraDis Visualization* is a novel approach to investigate spatial disparities in mobility service frequency			1		1	4
4	It has sufficient information to explore spatial disparity in mobility service frequency			1	1		3.5
5	*PubTraDis Visualization* does not provide useful interactivity to analyze spatial disparity in mobility service frequency	1		1			2
Average Rating for Positive Questions (4)							3.8
Rating for Negative Questions (1)							2
Overall Average with Negative Questions Inversed							3.4

Overall ratings on the utility of *PubTraDis Visualization* were mixed (table 4.6). Positively phrased questions received an average of 3.8 out of "5" and negatively phrased question received 2 out of "5". Thus, overall average rating was 3.4 just slightly above the "3" or meaning which is midpoint on the scale. It is seen from the table that the participant strongly agreed on the second question related to contribution to spatial disparity analysis in public transit research. Also question no. 3 and 4 got positive support. These questions were related to the novelty and competence of *PubTraDis Visualization*.

4.6.3.2 Usability

In this questionnaire equal number of positive and negative questions were asked to get insights about the usability of the *PubTraDis Visualization*. The following table (4.7) shows the result of usability rating.

Table 4.7: Participants' Response to Usability Ratings

S.N.	Usability Rating	1	2	3	4	5	Average
		Disagree			Agree		
1	*PubTraDis Visualization* was easy to use			2			3
2	Additional information required to use the tool	1				1	3
3	I was confused with the interactive tools and felt uncomfortable		1		1		3
4	Customizable interface was unnecessary	1		1			2
5	Filter options are convenient to use and add value to the tool				2		4
6	Pop-up windows add extra value to the tool				1	1	4.5
7	Supplementary information (Chart/Table) does not add any value in the interface	1	1				1.5
8	Data download option is not important for the tool	1				1	3
9	The comparison window is easy to use and add value to the tool			1		1	4
10	Comparison window does not provide adequate options for detail analysis		1	1			2.5
11	The visual design of *PubTraDis Visualization* is well standard		1			1	3.5
12	*PubTraDis Visualization* comply with basic cartographic principles				2		4
Average Rating for Positive Questions (6)							3.8
Rating for Negative Questions (6)							2.5
Overall Average with Negative Questions Inversed							3.2

From the table (4.7) it is seen that there is no similar pattern of the ratings, and they are mixed. Positively phrased questions received an average of 3.8 out of "5" while negatively phrased questions got an average of 2.5 out of "5". It can be said that the participants were neither agree nor disagree on the usability of the *PubTraDis Visualization* as the overall average ratings was 3.2 out of "5". Participants were agreed (4 out of "5") on the compliance of basic cartographic principles of *PubTraDis Visualization*. Question no. 6 got highest ratings which was about pop-up. On the other hand, question no. 7 got lowest rating (1.5 out of “5”) which was negatively phrased. They agreed that comparison window was easy to use and add extra value to the visualization tool (4 out of “5”). Similarly, participants also agreed that filter options were convenient to use and add value to the tool. Interestingly, participants ratings on question no. 2 and 8 were completely opposite to one another. These questions were about the necessity of additional information to use the tool and data download options of the tool.

4.6.4 Suggestions

Participants also gave some suggestions during the discussions. These are as follows:

a) Set same scale for both map interface in comparison window and link them together so that same scale follows in both interfaces.

b) Set a bookmark location so that map can load fast.

c) Make sub layer for the tab name in the interface.

d) Integrate functionality so that after user selection into one map layer reflect related chart and indicator tab and automatically visualize them on top.

e) Performance could be better for end user.

Chapter Five

DISCUSSION

This chapter discusses the results in context to research objectives and research questions. Limitations and challenges of the study also included in the last part of the chapter.

5.1. Objective 1

To identify and determine methods of exploring spatial disparities in public transit.

5.1.1 Exploring Spatial Disparities

RQ1A: What are the methods to explore spatial disparities in public transit?

Within the scope of present study several materials and method identified for exploring spatial disparities in public transit. As for the tools researcher relied mostly on Esri ArcGIS Network Analyst and OTP. From the study it is found that 10 studies used the first one and 8 used OTP for their analysis (please refer to table 2.2 for more details).

Tools: In this study ArcGIS Pro was used for processing the raw data and GIS automation techniques were used to develop a fast and reproducible data processing framework. All the tools developed for this study are discussed in section (3.3) and submitted to IOER as per requirement.

In this study the author summarized different statistical approaches that used to explore spatial disparities in public transit. Central tendency measurement was used in 7 studies whereas equal number of studies also used measures of dispersion or variability. Local Moran’s Index used in 5 studies for the quantification of spatial disparities in public transit studies. Nevertheless, Gini-coefficient also used in 3 studies for the spatial disparities analysis.

From the literature it is found that multiple types of graphs and maps used to display spatial disparities in public transit. In this study, histogram was used to show aggregated data on municipalities level. Moreover, section (2.1.4.2) reveals that Choropleth map used to show disparities in different studies (Yan et al., 2022; Niedzielski, 2021; Farber & Fu, 2017; Farber et al., 2014).

5.1.2 Existing Applications Summary

RQ1B: What are the available visualization tools can be found online that used GTFS data?

A number of visualization tools are available on online that uses GTFS data. These tools are mainly focused on visualizing spatiotemporal pattern. They are developed in order to visualize operational characteristics of specific transit system (i.e., PubtraVis). That shows speed, flow, density etc. of Calgary Transit system. On the other hand, Swiss Railways Network and TRAVIC visualize movement of transit service.

There is a definite need of a visualization tool that can show spatial disparities in public transit. *PubTraDis Visualization* is the first approach that can contribute to public transit research that visualize spatial disparities in mobility service frequency.

5.1.3 User-centered Web-map Design

RQ1C: What are the user requirements and design parameters for interactive web-map-based visualization?

Theoretical approach in combination with existing application played important role in work domain analysis. Required functionalities for *PubTraDis Visualization* were identified from the findings of work domain analysis. Also, data design and visualization ideas considered from the literature and existing applications (please refer to section 2.2 for existing visualizations examples)

5.2 Objective 2

To adopt open-source data for interactive web-map-based visualization of spatial disparities in mobility service frequency.

5.2.1 GTFS Data Integration

RQ2A: How can the GTFS data be integrated to calculate mobility service frequency in best possible spatial resolution?

GTFS data has been used in different studies (please refer to table 2.2 for more details). Findings from section (2.1) suggests that no effort has been made so far in visualizing spatial disparities in mobility service frequency. Within the scope of the current study, the author has developed reproducible workflow so that the GTFS data can be integrated to calculate mobility service frequency for whole Germany. The following figure (5.1) summarizes the workflow to calculate station level frequency and figure (5.2) represent the Frequency Indicator and Disparity Index calculation.

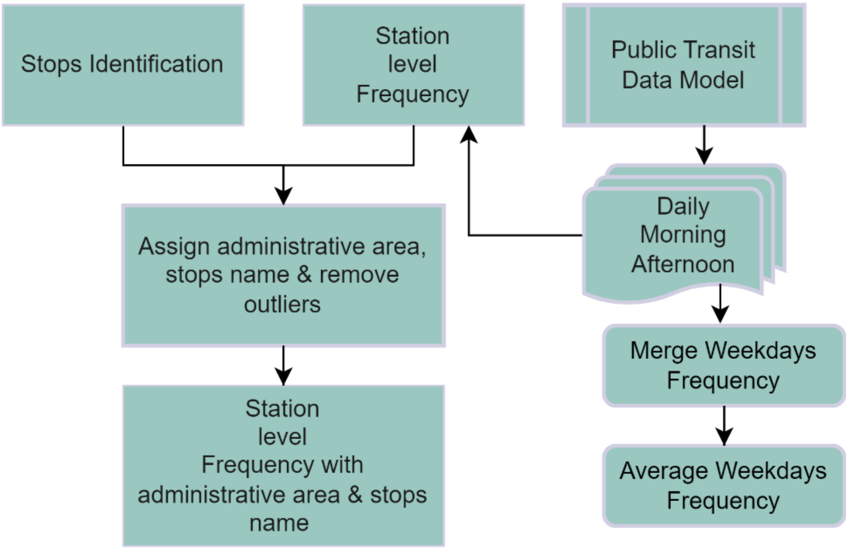


Figure 5.1: Workflow to Calculate Station Level Frequency

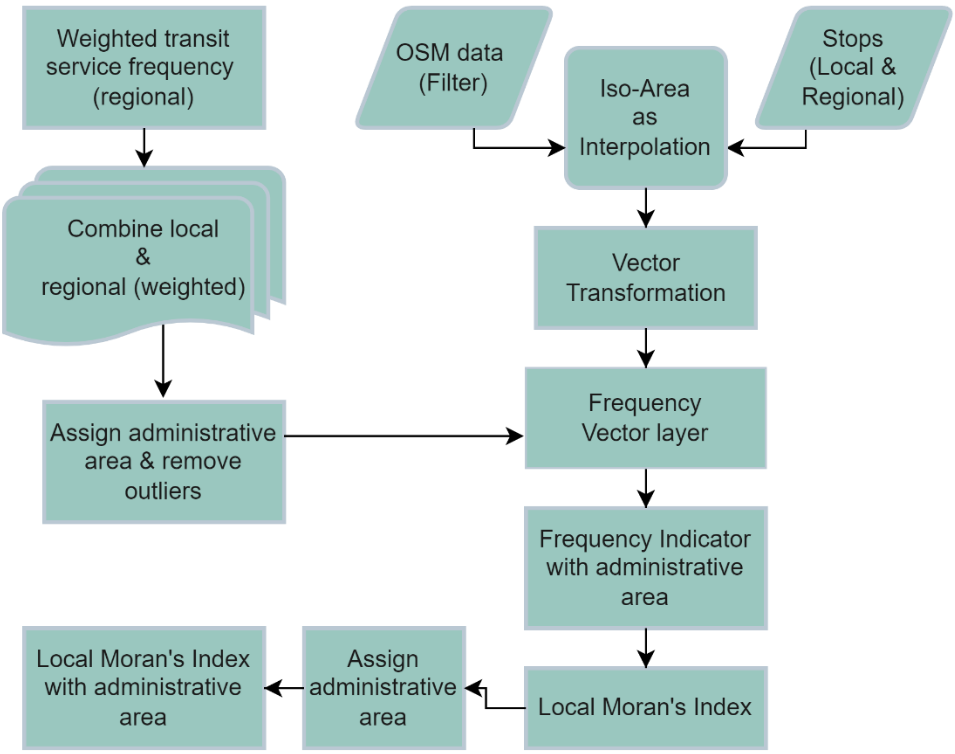


Figure 5.2: Workflow to Calculate Frequency Indicator and Disparity Index

Though in this study Frequency Indicator and Disparity index calculated only for VVO area due to time limitation, the developed workflow will be applicable for any administrative level (Country, States, Municipalities etc.)

5.2.2 Challenges in Multi-source Data Harmonization

RQ2B: What are the major challenges to harmonize required multi-sourced input dataset in the calculation process of mobility service frequency in Germany?

5.2.2.1 Missing Data

Collected GTFS data of Germany comes with some missing data. Figure (5.3) depicts the locations of missing data. In this study, no attempt has been made to calculate the percentage of the missing data as there was no reliable source for exact number of public transits stops.

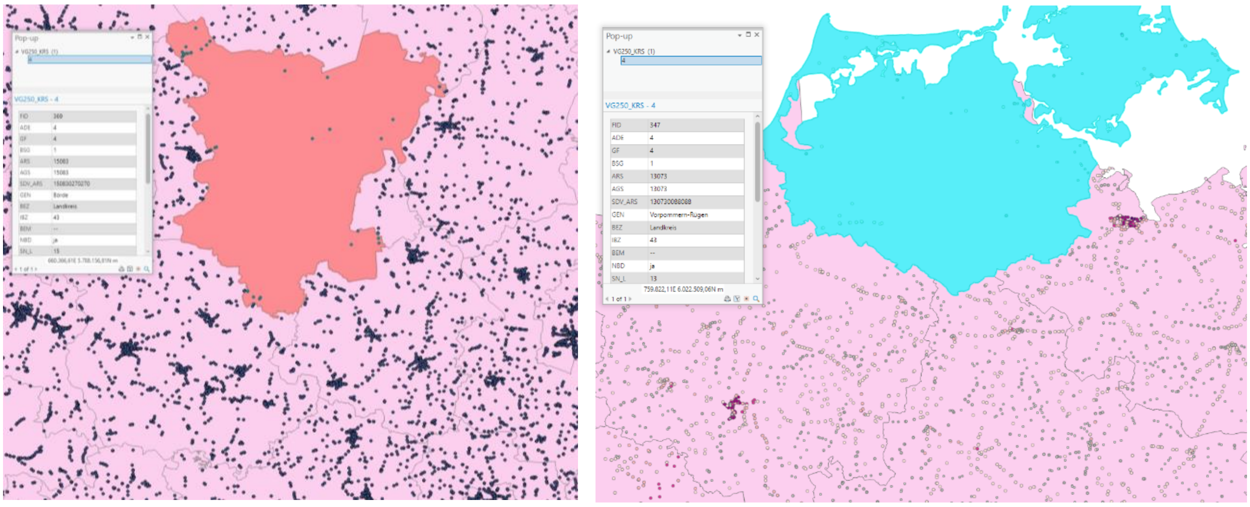


Figure 5.3: Missing Data in GTFS

5.2.2.2 Geographical Reference

One of the challenges of multi-source data is the geographic reference of spatial data. Data should maintain certain quality standard. Collected dataset for the study have different geographical reference. For instance, OSM data comes with WGS 1984 projection, whereas administrative data is projected in UTM zone 32, Ellipsoid GRS80, Datum ETRS89. In this study, IOER standard geographical reference system has been followed to ensure the current practice and further continuation of the research.

5.3 Objective 3

To develop and evaluate a tool for interactive web-map-based visualization of spatial disparities in mobility service frequency.

5.3.1 Web-map-based Visualization

RQ3A: How can the mobility service frequency be represented in the interactive web-map-based visualization application?

In this study the author has adopted the UCD process from Robinson et al., (2005), where they implemented a six stages of design process that involves user at each stage of the design.

5.3.1.1 Work Domain Analysis

To know the preset state of the research and user requirement and design principles this study initiated with theoretical background which also fulfill the first objective.

5.3.1.2 Conceptual Development

This refers to the required feature outline derived from first stage (Robinson et al., 2005). The author identified the required feature as well as the layout in this stage. The layout design and functionalities iterate through the feedback of the supervisors.

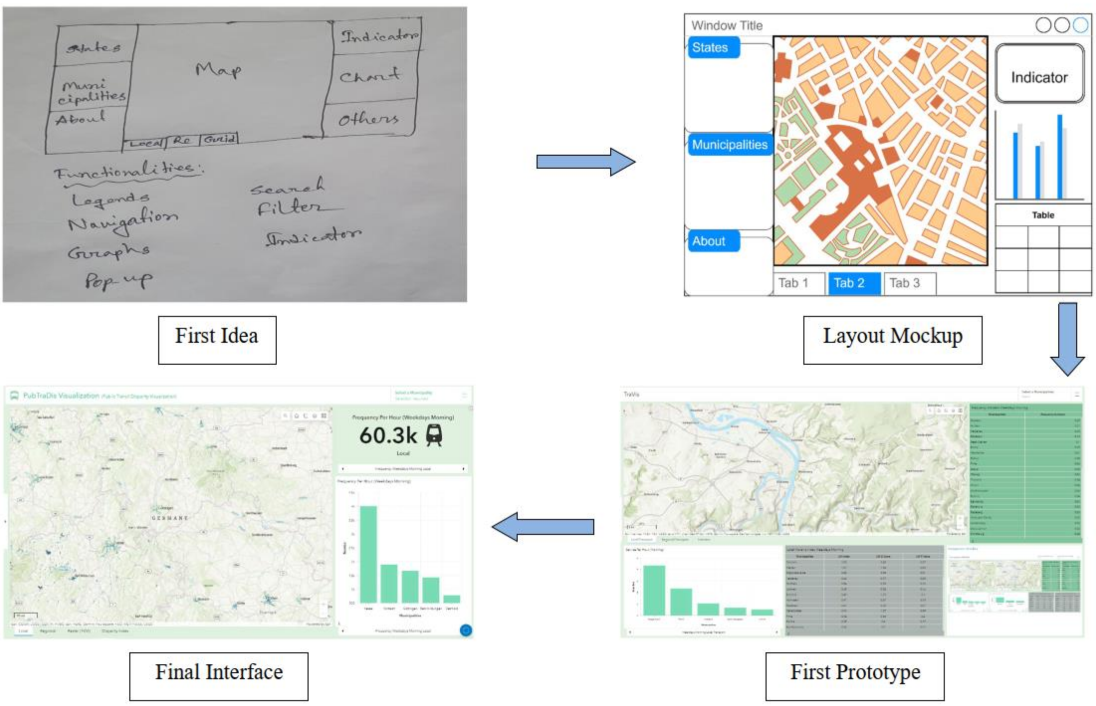


Figure 5.4: Conceptual Development to Final Interface of PubTraDis Visualization

5.3.1.3 PubTraDis Visualization

At this stage *PubTraDis Visualization* developed. In this study the final interface the author adopted some feedback from the supervisors (First Prototype to Final Interface in figure 5.4) on feature and functionality. Due to time limitation all suggestion were not implemented. JavaScript Object Notation (JSON) file of the *PubTraDis Visualization* submitted to IOER as per requirement.

Data design: In this study author combined standard exploratory data analysis techniques and sophisticated statistical approach for the *PubTraDis Visualization* user. Frequency Indicator data are aggregated by mean (Yan et al., 2022; Lou et al., 2019; Farber & Fu, 2017; Farber et al., 2014). On the other hand, Local Moran's Index was calculated and visualized to show spatial disparity in public transit in grid level (Niedzielski, 2021; Allen & Farber, 2020; Huang, 2020; Wessel & Farber, 2019; Heppe & Leibig, 2017).

Map Visualization: Following the findings from theoretical approach, the author adopted Choropleth map visualization to show spatial disparities in grid level. Moreover, *PubTraDis Visualization* implemented Yan et al., (2022) techniques for the visualization of Frequency Indicator. On the other hand, author has chosen CO Type data for displaying Local Moran's Index output following Niedzielski (2021).

5.3.1.4 Expert Evaluation

After the development of *PubTraDis Visualization* expert base evaluation carried out (please refer to section 4.6 for more details).

5.3.1.5 Implementation

Once the feedback is collected, implementation stage begins. Due to time limitation the author could not address all the suggestion in this stage (please refer to table 5.1 for more details).

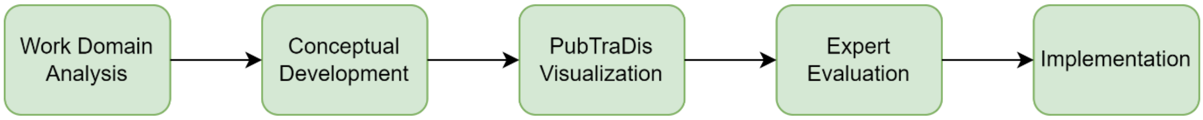


Figure: 5.5 UCD Process Followed in PubTraDis Visualization

5.3.2 Evaluation Summary

RQ3B: What are the evaluation outcomes of the designed application based on expert opinions?

PubTraDis Visualization is the first visualization tool of its kind, that show spatial disparity in mobility service frequency using public transit feed (GTFS) data. Expert base evaluation outcomes suggests that general interface design is preferred by both of them. Though visualization style followed previous studies, mix response found on that.

Utility outcomes indicate that *PubTraDis Visualization* is a novel approach and would contribute to spatial disparity analysis in public transit research. Similarly, experts also agreed on the compliance of basic cartographic principles in *PubTraDis Visualization.*

Experts' suggestions on *PubTraDis Visualization* are listed in table (5.1) along with the implementation state.

Table 5.1: Experts' Suggestions on PubTraDis Visualization and Implementation

Suggestion	**Fixed?**	
	Yes	**No**
a) Same scale/link both windows		No
b) Bookmark location	Yes	
c) Sub layer for tab name		No
d) Functionality for linked selection		No
e) Performance	Yes (to some extent)	

5.4 Limitations of the Study

a) Though, the research aim was to model spatial disparities for Germany, it could not carry out grid level frequency calculation for the whole Germany. According to the initial plan, OSM data was collected and processed for Germany. Due to computational cost Distance Indicator could not calculated for whole Germany. It would take 14 days to complete with the high configured computer facilities at IOER, which was not possible for author to monitor. On second attempt, author reduced the study area to Saxony state which was also not possible to finish within desired timeframe. Finally, relatively smaller Verkehrsverbund Oberelbe (VVO) service area in Saxony were chosen. Following figure (5.6) showing the Distance Indicator modelling for Saxony.

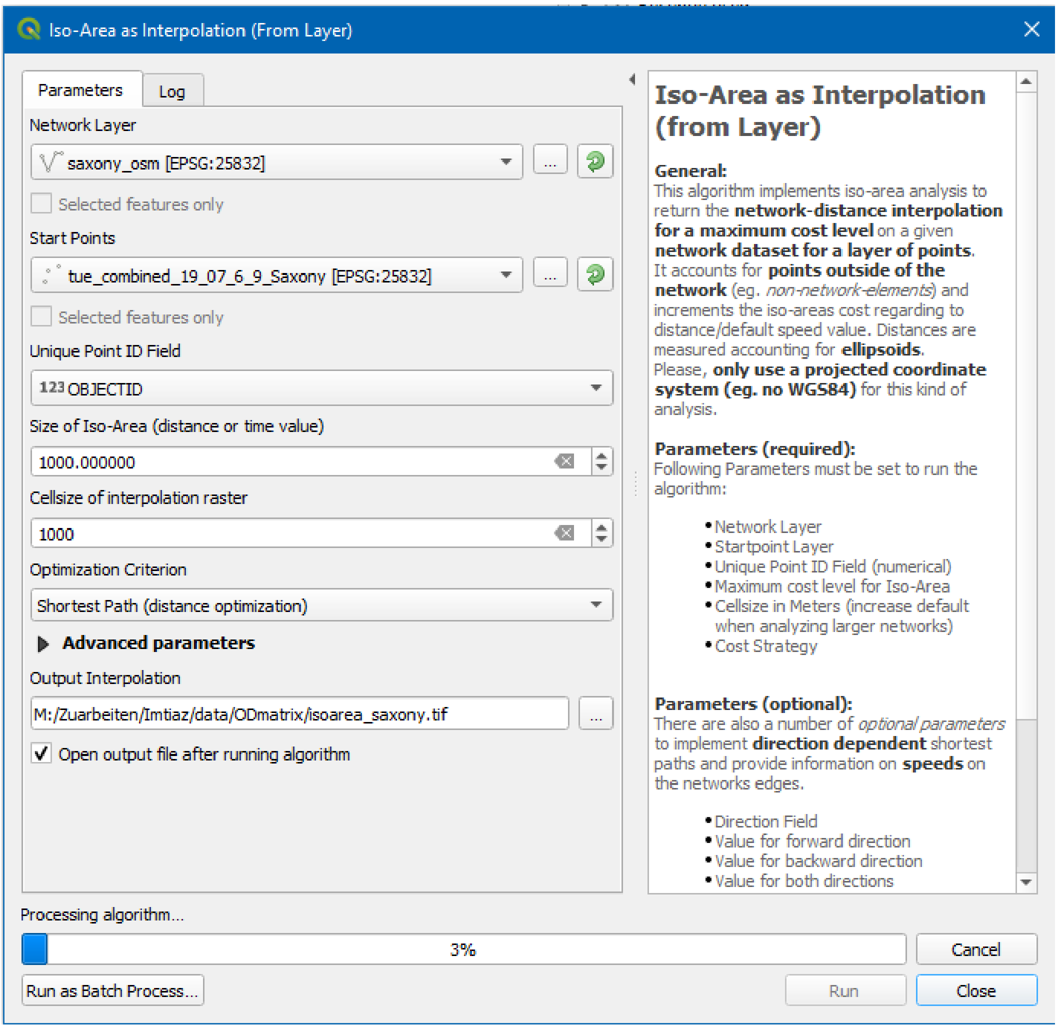


Figure 5.6 Distance Indicator by Iso-Area as Interpolation for Saxony

b) This research could not implement most of the suggestions from experts due to time shortage.

c) Since there was no new available GTFS data during the study period, automate update of *PubTraDis Visualization* could not done, but in theory it is possible[1,2]

[1] https://learn.arcgis.com/en/projects/automate-updates-to-web-feature-layers-without-scripting/ (Accessed December 2022)
[2] https://www.esri.com/arcgis-blog/products/arcgis-online/administration/updating-arcgis-com-hosted-feature-services-with-python/(Accessed December 2022)

5.5 Challenges of the Study

a) This research started with an understanding that it will adopt previously developed tool for the raw data processing. But the tool was intended for specific transit service authority. After several attempt, the author could not fit the tool to use that for GTFS data of whole Germany. Consequently, author had to develop new workflow and tool for raw GTFS data analysis.

b) In the first phase of the study, author collected only local transit data for Germany (03-10th July 2022). In the later stage, decision made to include regional transport data in the study. By that time GTFS data was updated to (17-24th July 2022) at the source website. Unfortunately, there was no data archive repository available for the previous GTFS data. So, author could not collect regional transport data for the same period. As it was not known before author had to collect new data for both local and regional transport and start data processing from the beginning. A data repository should be developed so that one can make comparison between two-time frames. This will also allow the researcher to make analysis on seasonal basis (if there is any difference) as well as to find the ratio of missing data in each GTFS data.

Chapter Six

CONCLUSION & OUTLOOK

6.1 Future Work

This study will contribute to further research related spatial disparities in mobility service frequency. Current state of the *PubTraDis Visualization* can be used to perform analysis between different States & Municipalities (station level frequency) as well as grid level analysis in different Municipalities of VVO area. Nevertheless, following the same workflow, one can perform country scale analysis. Further research would be possible by extending the present study in different way. Some of the possible future research scopes are discussed here:

A) With the current Frequency Indicator one can analyze population impact in mobility services frequency, for instance, to find out if population plays any role for high or low frequency.

B) Moreover, the Frequency Indicator can be combined with building level data. For instance, research can be done to find the disparities in residential area and industrial are.

C) Furthermore, one can perform quantitative analysis based on current data to find transit service users satisfaction on current service as well as measure the walk time index and classify them in different zone.

D) One can create a reproducible workflow to automate new GTFS data integration (upon new data availability) in *PubTraDis Visualization* from the current JSON file with little effort.

6.2 Conclusion

This study started with a systematic literature review to answer the method of exploring spatial disparities in public transit. Interface ideas and required functionalities were then identified from the analysis of the existing applications and reviewing design principles. This study uses ArcGIS Pro and QGIS for the data analysis. GIS automation techniques were applied to enhance processing and to develop reproducible workflow. At present, there is no visualization application related to spatial disparities in mobility service frequency. Consequently, data design and other visualization method fixed from findings of the theoretical framework. Evaluation outcomes suggests that *PubTraDis Visualization* will contribute to relevant field. This novel approach can be extended for further research that can play a role to explore spatial disparities in mobility service in any region.

References

Abras, C., Maloney-Krichmar, D. & Preece, J. (2004). User-centered design. *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications*, 37, 445-456.

Al-Bakri, D. M. (2018). *Matching assessment of road network objects of volunteered geographic information. 25*(1), 15.

Allen, J., & Farber, S. (2020). Planning transport for social inclusion: An accessibility-activity participation approach. *Transportation Research Part D: Transport and Environment*, *78,* 102212. https://doi.org/10.1016/j.trd.2019.102212

Anselin, L., & Piras, G. (2009). *Approaches Towards the Identification of Patterns in Violent Events, Baghdad, Iraq:* Defense Technical Information Center. https://doi.org/10.21236/ADA500195

Antrim, A., & Barbeau, S. J. (2013). The Many Uses of GTFS Data – Opening the Door to Transit and Multimodal Applications.

http://www.locationaware.usf.edu/wp-content/uploads/2010/02/The-Many-Uses-of-GTFS-Data-%E2%80%93-ITS-America-submission-abbreviated.pdf

Bast, H., Brosi, P., & Storandt, S. (2014). Real-time movement visualization of public transit data. In Y. Huang, M. Schneider, M. Gertz, J. Krumm, & J. Sankaranarayanan (Eds.), *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 331–340). New York, NY, USA: ACM. https://doi.org/10.1145/2666310.2666404

Ben-Elia, E., & Benenson, I. (2019). A spatially-explicit method for analyzing the equity of transit commuters' accessibility. *Transportation Research Part A: Policy and Practice, 120,* 31–42. https://doi.org/10.1016/j.tra.2018.11.017

Brosi, P. (2014). *Real-Time Movement Visualization of Public Transit Data* (dissertation). Retrieved July 2022, from https://ad-publications.cs.uni-freiburg.de/theses/Master_Patrick_Brosi_2014.pdf.

Bryman, A. (2012). *Social Research Methods - 4th Ed*. OXFORD University Press. https://www.academia.edu/38228560/Alan_Bryman_Social_Research_Methods_4th_Edition_Oxford_Universi ty_Press_2012_pdf

Bundesamt für Kartographie und Geodäsie. (2021). Documentation. Administrative Areas 1:25000, VG250 and VG250-EW

Büttner, B., Kinigadner, J., Ji, C., Wright, B., & Wulfhorst, G. (2018). The TUM Accessibility Atlas: Visualizing Spatial and Socioeconomic Disparities in Accessibility to Support Regional Land-Use and Transport Planning. *Networks and Spatial Economics, 18*(2), 385–414. https://doi.org/10.1007/s11067-017-9378-6

Camarero, L., & Oliva, J. (2019). Thinking in rural gap: Mobility and Social Inequalities. *Palgrave Communications*, *5*(1). https://doi.org/10.1057/s41599-019-0306-x

Chotisarn, N., Merino, L., Zheng, X., Lonapalawong, S., Zhang, T., Xu, M., & Chen, W. (2020). A systematic literature review of modern software visualization. *Journal of Visualization*, *23*(4), 539–558. https://doi.org/10.1007/s12650-020-00647-w

Comber, A., Brunsdon, C., & Green, E. (2008). Using a GIS-based network analysis to determine urban greenspace accessibility for different ethnic and religious groups. *Landscape and Urban Planning*, *86*(1), 103–114. https://doi.org/10.1016/j.landurbplan.2008.01.002

Credit, K., Dias, G., & Li, B. (2021). Exploring neighbourhood-level mobility inequity in Chicago using dynamic transportation mode choice profiles. *Transportation Research Interdisciplinary Perspectives*, *12*, 100489. https://doi.org/10.1016/j.trip.2021.100489

Demaria, K., Sturmberg, B. C., Riley, B., & Markham, F. (2022). Exploring the feasibility of electric vehicle travel for remote communities in Australia. *Australian Geographer*, *53*(2), 201–222. https://doi.org/10.1080/00049182.2022.2086720

Eboli, L., & Mazzulla, G. (2008). A stated preference experiment for measuring service quality in public transport. *Transportation Planning and Technology*, *31*(5), 509–523. https://doi.org/10.1080/03081060802364471

Ehrig, N. (2020, 10, 21). Erhebung und Bewertung der ÖPNV-Versorgung mittels offener Daten. [Unpublished Master's thesis] Technical University of Dresden.

Eriksson, L., Friman, M., & Gärling, T. (2008). Stated reasons for reducing work-commute by car. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11*(6), 427–433. https://doi.org/10.1016/j.trf.2008.04.001

Farber, S., & Fu, L. (2017). Dynamic public transit accessibility using travel time cubes: Comparing the effects of infrastructure (dis)investments over time. *Computers, Environment and Urban Systems*, *62*, 30–40. https://doi.org/10.1016/j.compenvurbsys.2016.10.005

Farber, S., Ritter, B., & Fu, L. (2016). Space–time mismatch between transit service and observed travel patterns in the Wasatch Front, Utah: A social equity perspective. *Travel Behaviour and Society*, *4*, 40–48. https://doi.org/10.1016/j.tbs.2016.01.001

Farber, S., Morang, M. Z., & Widener, M. J. (2014). Temporal variability in transit-based accessibility to supermarkets. *Applied Geography*, *53*, 149–159. https://doi.org/10.1016/j.apgeog.2014.06.012

Fayyaz S., S. K., Liu, X. C., & Zhang, G. (2017). An efficient General Transit Feed Specification (GTFS) enabled algorithm for dynamic transit accessibility analysis. *PLOS ONE*, *12*(10), e0185333. https://doi.org/10.1371/journal.pone.0185333

Gahegan, M. (1999). Four barriers to the development of effective exploratory visualisation tools for the geosciences. *International Journal of Geographical Information Science*, *13*(4), 289–309. https://doi.org/10.1080/136588199241210

Geiger, C. P. & Von Lucke, J. (2012). Open Government and (Linked) (Open) (Government) (Data). *JeDEM - eJournal of eDemocracy and Open Government, 4*(2), 265-278. https://jedem.org/index.php/jedem/article/view/143/115

Goliszek, S., & Połom, M. (2016). The use of general transit feed specification (GTFS) application to identify deviations in the operation of public transport at morning peak hours on the example of Szczecin. *Europa XXI*, *31*, 51–60. https://doi.org/10.7163/Eu21.2016.31.4

Heppe, L., & Liebig, T. (2017). Real-Time Public Transport Delay Prediction for Situation-Aware Routing. In G. Kern-Isberner, J. Fürnkranz, & M. Thimm (Eds.), *KI 2017: Advances in Artificial Intelligence* (Vol. 10505, pp. 128–141). Springer International Publishing. https://doi.org/10.1007/978-3-319-67190-1_10

Hix, D., Swan, J. E., Gabbard, J. L., McGee, M., Durbin, J., & King, T. (1999). User-centered design and evaluation of a real-time battlefield visualization virtual environment. *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*, 96–103. https://doi.org/10.1109/VR.1999.756939

Huang, R. (2020). Transit-based job accessibility and urban spatial structure. *Journal of Transport Geography*, *86*, 102748. https://doi.org/10.1016/j.jtrangeo.2020.102748

Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, *33*(4), 470–516. https://doi.org/10.1145/503112.503114

Jain, M., & Korzhenevych, A. (2017). Spatial Disparities, Transport Infrastructure, and Decentralization Policy in the Delhi Region. *Journal of Urban Planning and Development, 143*(3), 05017003. https://doi.org/10.1061/(ASCE)UP.1943-5444.0000379

Karner, A. (2018). Assessing public transit service equity using route-level accessibility measures and public data. *Journal of Transport Geography*, *67*, 24–32. https://doi.org/10.1016/j.jtrangeo.2018.01.005

Kitchenham, B. A., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El Emam, K., & Rosenberg, J. (2002). Preliminary guidelines for empirical research in software engineering. *IEEE Transactions on Software Engineering*, *28*(8), 721–734. https://doi.org/10.1109/tse.2002.1027796

Kostelnick, J. C., Dobson, J. E., Egbert, S. L., & Dunbar, M. D. (2008). Cartographic Symbols for Humanitarian Demining. *The Cartographic Journal*, *45*(1), 18–31. https://doi.org/10.1179/000870408X276585

Kucukali, A., Pjeternikaj, R., Zeka, E., & Hysa, A. (2022). Evaluating the pedestrian accessibility to public services using open-source geospatial data and QGIS software. *Nova Geodesia*, *2*(2), 42. https://doi.org/10.55779/ng2242

Liao, Y., Gil, J., Pereira, R. H. M., Yeh, S., & Verendel, V. (2020). Disparities in travel times between car and transit: Spatiotemporal patterns in cities. *Scientific Reports*, *10*(1), 4056. https://doi.org/10.1038/s41598-020-61077-0

Liebig, T., Piatkowski, N., Bockermann, C., & Morik, K. (2014). Route Planning with Real-Time Traffic Predictions. *In Proceedings of the LWA 2014 Workshops: KDML, IR, FGWM*, 83-94.

Luo, D., Cats, O., van Lint, H., & Currie, G. (2019). Integrating network science and public transport accessibility analysis for comparative assessment. *Journal of Transport Geography*, *80*, 102505. https://doi.org/10.1016/j.jtrangeo.2019.102505

MacEachren, A. M., & Ganter, J. H. (1990). A PATTERN IDENTIFICATION APPROACH TO CARTOGRAPHIC VISUALIZATION. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *27*(2), 64–81. https://doi.org/10.3138/M226-1337-2387-3007

Maduako, I., Cavalheri, E., & Wachowicz, M. (2018). Exploring the use of time-varying graphs for modelling transit networks. *ArXiv, abs/1803.07610*.

Niedzielski, M. A. (2021). Grocery store accessibility: Different metrics – Different modal disparity results and spatial patterns. *Journal of Transport Geography*, *96*, 103160. https://doi.org/10.1016/j.jtrangeo.2021.103160

Owen, A., & Levinson, D. M. (2015). Modeling the commute mode share of transit using continuous accessibility to jobs. *Transportation Research Part A: Policy and Practice*, *74*, 110–122. https://doi.org/10.1016/j.tra.2015.02.002

Panzera, D., & Postiglione, P. (2020). Measuring the Spatial Dimension of Regional Inequality: An Approach Based on the Gini Correlation Measure. *Social Indicators Research*, *148*(2), 379–394. https://doi.org/10.1007/s11205-019-02208-7

Prommaharaj, P., Phithakkitnukoon, S., Demissie, M. G., Kattan, L., & Ratti, C. (2020). Visualizing public transit system operation with GTFS data: A case study of Calgary, Canada. *Heliyon*, *6*(4), e03729. https://doi.org/10.1016/j.heliyon.2020.e03729

Puchalsky, C., Joshi, D., & Scherr, W. (2011). Development of a Regional Forecasting Model Based on Google Transit Feed. 13th TRB Planning Application Conference, May 2011, Reno, NV. https://www.researchgate.net/publication/260256429_Development_of_a_Regional_Forecasting_Model_Based_on_Google_Transit_Feed/stats

Raffler, C. 2018. "QGIS Network Analysis Toolbox 3." May, 2018. https://root676.github.io/index.html.

Redman, L., Friman, M., Gärling, T., & Hartig, T. (2013). Quality attributes of public transport that attract car users: A research review. *Transport Policy*, *25*, 119–127. https://doi.org/10.1016/j.tranpol.2012.11.005

Richard, J. R., & Egenhofer, M. J. (1995). A Comparison of two Direct-Manipulation GIS User Interfaces for Overlay. *Geographical Systems*, *2*(4), 267-290.

Robinson, A. C., Chen, J., Lengerich, E. J., Meyer, H. G., & MacEachren, A. M. (2005). Combining usability techniques to design geovisualization tools for epidemiology. *Cartography and Geographic Information Science*, *32*(4), 243–255. https://doi.org/10.1559/152304005775194700

Roth, R. E. (2013a). Interactive maps: What we know and what we need to know. *Journal of Spatial Information Science*, *6*, 59–115. https://doi.org/10.5311/JOSIS.2013.6.105

Roth, R. E. (2013b). An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2356–2365. https://doi.org/10.1109/tvcg.2013.130

Roth, R. E. (2012). Cartographic Interaction Primitives: Framework and Synthesis. *The Cartographic Journal*, *49*(4), 376–395. https://doi.org/10.1179/1743277412Y.0000000019

Roth, R. E., Çöltekin, A., Delazari, L., Filho, H. F., Griffin, A., Hall, A., Korpi, J., Lokka, I., Mendonça, A., Ooms, K., & van Elzakker, C. P. J. M. (2017). User studies in cartography: Opportunities for empirical research on interactive maps and Visualizations. *International Journal of Cartography*, *3*(sup1), 61–89. https://doi.org/10.1080/23729333.2017.1288534

Roth, R., Ross, K., & MacEachren, A. (2015). User-Centered Design for Interactive Maps: A Case Study in Crime Analysis. *ISPRS International Journal of Geo-Information*, *4*(1), 262–301. https://doi.org/10.3390/ijgi4010262

Sikder, S. K., Ehrig, N., Herold, H., & Meinel, G. (2020). Analyse der ÖPNV-Versorgung mittels offener Fahrplandaten – Potenziale, Herausforderungen und Lösungsansätze. In Meinel, G., Schumacher, U., Behnisch, M., Krüger, T. (Hrsg.): Flächennutzungsmonitoring XII mit Beiträgen zum Monitoring von Ökosystemleistungen und SDGs. Berlin: Rhombos, IÖR Schriften 78, S. 263- 270. 10.26084/12dfns-p026

Sikder, S. K., Herold, H., Meinel, G., Lorenzen-Zabel, A., & Bill, R. (2019). Blessings of Open Data and Technology: E-Learning Examples on Land Use Monitoring and E-Mobility. *Proceedings of the STS Conference Graz 2019*. Verlag der Technischen Universität Graz. https://doi.org/10.3217/978-3-85125-668-0

Stephanie. (2020, June 8). *Gini coefficient: Simple definition*. Statistics How To. Retrieved July 17, 2022, from https://www.statisticshowto.com/gini-coefficient/

Stephanie. (2017, November 14). *Moran's I: Definition, examples*. Statistics How To. Retrieved July 17, 2022, from https://www.statisticshowto.com/moransi/#:~:text=What%20is%20Moran's%20I%3F,the%20observations%20are%20not%20independent.

Stępniak, M., Pritchard, J. P., Geurs, K. T., & Goliszek, S. (2019). The impact of temporal resolution on Public Transport Accessibility Measurement: Review and case study in Poland. *Journal of Transport Geography*, *75*, 8–24. https://doi.org/10.1016/j.jtrangeo.2019.01.007

Tolochko, R. C. (2016). *Contemporary professional practices in Interactive Web Map Design* (dissertation). Retrieved June 2022, from https://minds.wisconsin.edu/bitstream/handle/1793/74969/Tolochko%20Robin%20C%202016.pdf?sequence=1&isAllowed=y.

Vendrametto Granzotti, R., Agostinho, A. A., & Bini, L. M. (2022). Drivers and spatial patterns of population synchrony of fish species in a floodplain. *Freshwater Biology*, *67*(5), 857–872. https://doi.org/10.1111/fwb.13886

Wessel, N., & Farber, S. (2019). On the accuracy of schedule-based GTFS for measuring accessibility. *Journal of Transport and Land Use, 12*(1). https://doi.org/10.5198/jtlu.2019.1502

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456x17723971

Xin, Y., Fu, L., & Saccomanno, F. F. (2005). Assessing transit level of service along travel corridors. *Transportation Research Record: Journal of the Transportation Research Board*, *1927*(1), 258–267. https://doi.org/10.1177/0361198105192700129

Xing, L., Liu, Y., & Liu, X. (2018). Measuring spatial disparity in accessibility with a multi-mode method based on park green spaces classification in Wuhan, China. *Applied Geography*, *94*, 251–261. https://doi.org/10.1016/j.apgeog.2018.03.014

Yan, X., Bejleri, I., & Zhai, L. (2022). A spatiotemporal analysis of transit accessibility to low-wage jobs in Miami-Dade County. *Journal of Transport Geography*, *98*, 103218. https://doi.org/10.1016/j.jtrangeo.2021.103218

Appendix

Questionnaire for Expert based evaluation of the PubTraDis Visualization

I. General Information

1) Year of Experience

a) Less than 10 b) 10-20 c) More than 20

2) How frequent you work with interactive web-based map?

a) Daily b) Weekly c) Other

II. Interface Interaction

1) What do you prefer?

a) Map at top and supplementary information at bottom

b) Map at bottom and supplementary information at top

c) Map at left and supplementary information at right

d) Map at right and supplementary information at left

e) Other

2) Which type of visualization do you prefer for station level frequency?

a) Dot Map b) Heat Map c) Both

3) Do you prefer CO Type map for Disparity Index Visualization?

a) Yes b) No

If no, which information you prefer

a) LM Index b) Z Score c) P Value

III. Utility & Usability Rating

PubTraDis Visualization = Name of the designed interface

Please answer the following questions:

a) Utility Rating

S.N.	Utility Rating	1	2	3	4	5
		Disagree			Agree	
1	I would use *PubTraDis Visualization* frequently for my work					
2	This tool would contribute to spatial disparity analysis in public transit research					
3	*PubTraDis Visualization* is a novel approach to investigate spatial disparities in mobility service frequency					
4	It has sufficient information to explore spatial disparity in mobility service frequency					
5	*PubTraDis Visualization* does not provide useful interactivity to analyze spatial disparity in mobility service frequency					

b) Usability Rating

S.N.	Usability Rating	1	2	3	4	5
		Disagree			Agree	
1	*PubTraDis Visualization* was easy to use					
2	Additional information required to use the tool					
3	I was confused with the interactive tools and felt uncomfortable					
4	Customizable interface was unnecessary					
5	Filter options are convenient to use and add value to the tool					
6	Pop-up windows add extra value to the tool					
7	Supplementary information (Chart/Table) does not add any value in the interface					
8	Data download option is not important for the tool					
9	The comparison window is easy to use and add value to the tool					
10	Comparison window does not provide adequate options for detail analysis					
11	The visual design of *PubTraDis Visualization* is well standard					
12	*PubTraDis Visualization* comply with basic cartographic principles					

IV. Suggestions

Do you have any suggestions ?……………………………………………………………………...

……………………………………………………………………………………………………
……………………………………………………………………………………………………