LUISS

Corso di laurea in Data Science and Management

Cattedra Data Visualization

Decoding City Network
Exploring Urban Structure and Spatial Event Interplay through Multiple Centrality Assessment

Prof.ssa Blerina Sinaimeri

RELATORE

Prof. Giuseppe Italiano

CORRELATORE

Lorenzo Antolini

756861

CANDIDATO

Anno Accademico 2023/2024

Table of contents

Introduction **3**

Chapter one: Definitions of networks **4**

1.1 Spatial Networks **5**

1.2 Planar Graph **6**

1.3 Size and structure of a network **8**

1.3.1 Size of the Network 8

1.3.2 Network Density 9

1.3.3 Diameter of a Network 10

1.4 Primal and dual graph in urban network **12**

1.5 Understanding Network Complexity **13**

Chapter two: Network Analysis in the Urban context **14**

2.1 Planned and Self-organized cities **14**

2.2 Inter network index **17**

2.2.1 Gamma index 18

2.2.2 Cyclomatic Number 19

2.2.3 Maximum number of cycles 20

2.2.4 Redundancy Index 20

2.2.5 Characteristic Path Lenght 21

2.2.6 Global Efficiency 22

2.3 Intra network indices **23**

2.3.1 Degree Centrality 24

2.3.2 Betweenness Centrality 25

2.3.3 Closeness Centrality 26

2.3.4 Efficiency Centrality 27

2.3.5 Straightness Centrality 28

Chapter three: Decoding Urban Structures: Multiple Centrality Assessment **29**

3.1 Methodology **29**

3.2 Decoding Network structure: Planned and Self organized cities **30**

3.2.1 MCA Part 1 - Inter-Network Indices 34

3.2.2 MCA part 2 – Intra - Network Indices 38

3.3 The Real-World Implications of Centrality Measures on Urban Safety **51**

3.3.1 Using MCA to analyze car crash. 52

Conclusion **55**

Bibliography **57**

Introduction

This thesis explores the application of network science to urban planning, with a focus on understanding the structural dynamics of urban road networks through centrality measures and indices and analyzing their influence on car accidents as a critical geospatial dimension. This endeavor not only aims to shed light on the intricate patterns of urban traffic flows and safety but also seeks to contribute to the broader discourse on sustainable urban development and infrastructure resilience.

The significance of network science in urban contexts is underscored by a rich tapestry of scholarly works that collectively provide a multidimensional view of urban networks' complexity and their implications for city planning and management. Among the seminal contributions is the work of Sevtsuk (2017), who articulates the critical role of urban network analysis in enhancing the effectiveness of urban planning and management strategies. This comprehensive overview sets the stage for a deeper exploration of urban networks' intricate nature and their pivotal role in shaping urban experiences.

Porta, Crucitti, and Latora (2006) introduce groundbreaking methodologies for street network analysis, effectively bridging network theory with practical urban planning considerations. Their innovative approach has substantially influenced the field of urban studies, offering new perspectives on analyzing and conceptualizing urban spaces. Similarly, Kirkley, Barbosa, Barthelemy, and Ghoshal (2018) extend the concept of betweenness centrality within street networks, providing valuable insights into urban networks' structural characteristics and their implications for traffic flow and safety.

Zhang, Miller-Hooks, and Denny (2015) delve into the critical role of network topology in transportation resilience, a vital aspect of understanding how urban design impacts the functionality and robustness of urban transportation systems. This is complemented by Ingvardson and Nielsen's (2018) empirical examination of the interplay between urban density, network topology, and public transport ridership, highlighting the intricate relationship between urban form and transportation efficiency.

The literature also features innovative approaches to assessing urban infrastructure resilience, such as Scott, Novak, Aultman-Hall, and Guo's (2006) introduction of the Network Robustness Index. This methodological contribution emphasizes the importance of identifying critical links within transportation networks to ensure the resilience of urban infrastructure against disruptions.

In addition to these methodological and empirical contributions, the literature review encompasses a broad range of studies that offer insights into the dynamics of urban street networks (Crucitti, Latora, & Porta, 2006), the structural differences between self-organized and planned cities (Masucci & Molinero, 2016; Masucci, Smith, Crooks, et al., 2009), and the application of network science to various urban phenomena (Committee on Network Science for Future Army Applications, 2006).

This thesis situates itself within this rich scholarly landscape, aiming to bridge theoretical insights with practical analysis of car accident data within urban networks. By integrating the methodological frameworks and findings from these diverse studies, this research contributes to ongoing discussions in urban planning, network analysis, and public safety, offering new perspectives on leveraging network science to enhance urban livability and safety.

Chapter one: Definitions of networks

The initial part of this thesis explores the extensive domain of network research and elucidates its applications in several situations, encompassing social interactions, biology, and telecommunications. This text examines the essential elements of networks, namely nodes and edges, and their interplay in building intricate systems. The chapter highlights the significance of geographical networks, demonstrates the influence of the physical environment on network dynamics, and introduces planar graphs as an essential tool for representing networks without any edge intersections. Furthermore, it encompasses the dimensions of network size, structure, and complexity, furnishing details on network diameter, density, and challenges associated with measurement. Furthermore, it demonstrates the analytical consequences of comparing primal and dual graph representations in metropolitan networks. The first chapter presents a comprehensive framework for understanding the intricate interactions and behaviors that occur inside various networks.

Network science is a multidisciplinary field that studies the complexity of various interconnected systems. This concept is relevant to various domains, including computer systems, biological networks, communication, cognitive processes, semantic networks, and social interactions. Essentially, this discipline studies the distinct components of a network, known as nodes or vertices, as well as the connections between them, referred to as links or edges. This approach enables the examination of both the overall architecture of networks and their individual-level connections.

The field of network research encompasses a wide range of techniques and theoretical models that are tailored to the specific systems under investigation. These encompass statistical mechanics, a field rooted in physics that provides an understanding of the dynamics and characteristics of intricate systems.

Graph theory, derived from mathematics, provides the essential vocabulary and techniques for describing network topologies. In the field of computer science, data mining, and information visualization approaches enable the extraction and presentation of relevant patterns from vast collections. Research in sociology enhances our comprehension of human interactions and organizational networks. Statistical methods for inferential modeling facilitate the comprehension of connections and causality within network data.

The objective of network science, as defined by the US National Research Council, is to establish a comprehensive framework for analyzing and representing biological, physical, and social events through the use of network models. According to the 2006 Committee on Network Science for Future Army Applications, this endeavor is essential for the advancement of prognostic models capable of forecasting the behavior of intricate systems.

The distinctive features of networks allow for the quantification and analysis of their attributes. Understanding these characteristics is essential for comparing different network models and effectively implementing principles in network research. The initial section of this thesis conducts a comprehensive analysis of the fundamental concepts and principles that are important to the study of network research and graph theory, thus establishing the foundation for further discussions.

1.1 Spatial Networks

Spatial networks, which are distinguished by the incorporation of physical space in their structure and dynamics, hold a crucial position in the field of network science. In contrast to non-spatial networks, spatial networks consist of nodes that are positioned inside a geometric or physical space. The connections, or edges, between these nodes are greatly influenced by the spatial distances that separate them (Leskovec et al., 2007).

Spatial networks play a crucial role in the examination and enhancement of many essential infrastructures and systems. These networks exhibit distinct features such as the accurate positioning of nodes in a given area and the impact of physical distance on the likelihood and effectiveness of connections between these nodes. The spatial dimension is crucial in various networks, encompassing urban transportation systems and the intricate configuration of the brain (M. Barthélemy, 2011).

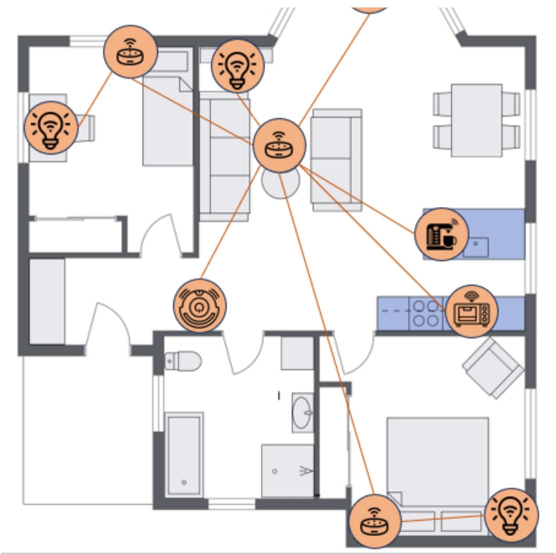
Examples of spatial networks:

Road Networks: The interconnected system of roads and crossings serves as an illustrative model for understanding how the arrangement of physical space influences travel paths, traffic movement, and the construction of infrastructure (*fig 1.1*).

Networks of Internet of Things (IoT): The efficiency of a smart home system is influenced by the spatial configuration of IoT devices, which are strategically positioned inside the physical space of the house.

Figure 1.1

Figure 1.2



The finding that the probability of a connection between two nodes generally diminishes as the distance between them grows (Leskovec et al., 2007) highlights the importance of distance in determining connectivity. This notion impacts the effectiveness of transportation routes and the ability of electricity grids to withstand and recover from disruptions, among other systems.

Considering M. Barthélemy's argument on the significance of the spatial component in network research, it is crucial to recognize instances where a network lacks this geographical characteristic. Non-spatial networks can be seen in various contexts. For instance, in an ecosystem, predator-prey relationships can be represented as nodes representing different species and edges symbolizing interactions such as predation (*fig 1.3*). Similarly, social networks can be represented by nodes representing persons and edges representing social ties between them (*fig 1.4*).

Figure 1.3

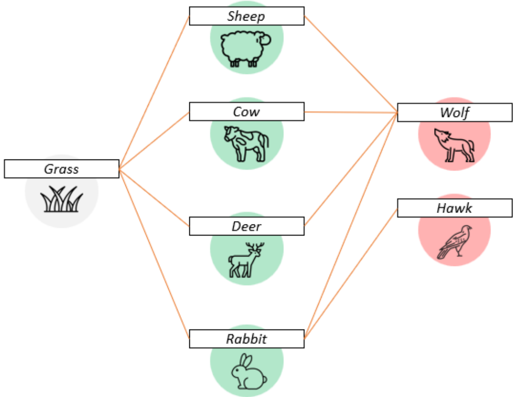
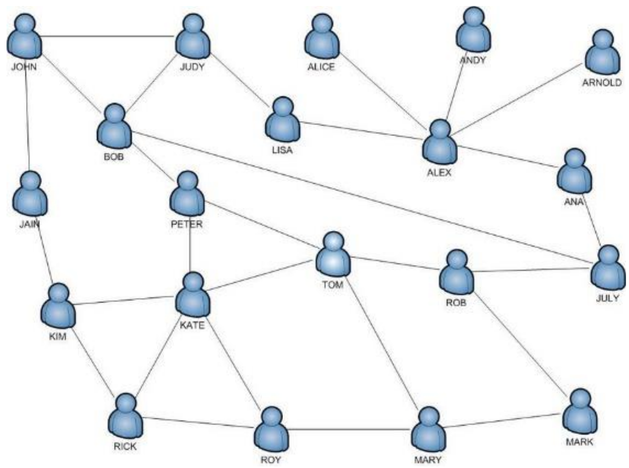


Figure 1.4



Diagrams are valuable tools that greatly improve the comprehension of spatial networks' intricacy, unveiling important network attributes such as density, connection, and possible bottlenecks.

The presence of geometric limitations greatly influences the arrangement of spatial networks, leading to unique connection patterns. These patterns encompass the phenomenon of clustering in social networks and the presence of hub-and-spoke structures in transportation networks.

The examination of geographical networks encounters distinct difficulties, specifically in handling spatial data and precisely representing physical limitations. The utilization of geographic information systems (GIS) and spatial statistics is essential for tackling these difficulties and revealing valuable information on spatial network features (M. Barthélemy, 2011; Bullmore et al., 2009).

Spatial networks provide useful insights into the functioning and structure of complex systems that are influenced by spatial linkages. By comprehending the distinctive attributes and intricacies of these networks, we may improve the structure and effectiveness of both organic and artificial systems.

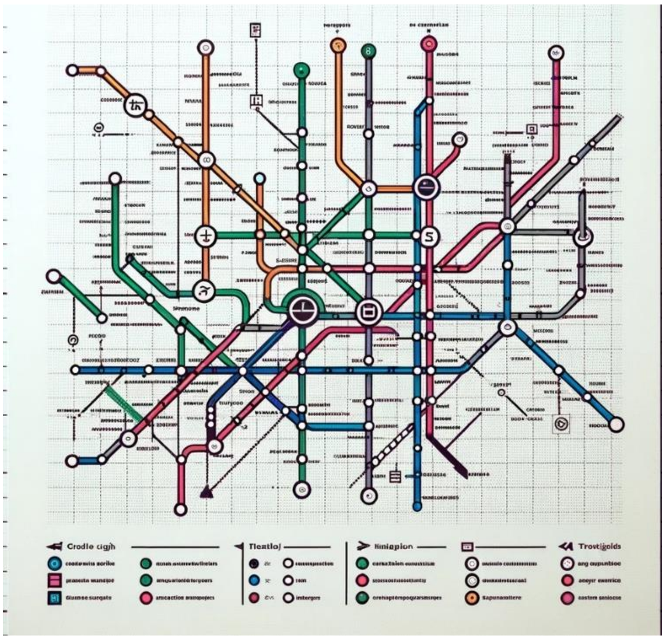
1.2 Planar Graph

Planar graphs are essential in both mathematics and physics for representing real-world networks that rely on spatial interactions. Graphs that can be shown on a plane without any edges crossing each other are known as planar graphs. These graphs are useful for simplifying the depiction of complex networks.

Planar graphs enable the visualization of networks in two dimensions, which is essential for various applications such as geographic mapping and theoretical physics. They are particularly useful in discretizing random surfaces in two-dimensional quantum gravity, as highlighted by Tutte (1963).

The usefulness of planar graphs in describing physical connections is exemplified by maps and transportation networks (*fig 1.5)*. The Kuratowski theorem offers a planarity test that is crucial for the practical application of these concepts (Barthélemy, 2011).

Figure 1.5



Euler's formula is a mathematical expression that is employed to analyze the property of planarity in a network. It demonstrates the inherent scarcity of edges in planar graphs. The simplicity of the formula restricts the level of complexity in network designs, highlighting its significance in planar graph theory (Barthélemy, 2011).

$$N - E + F = 2$$

The computational determination of planarity and its real-world applications present substantial obstacles. According to Barthélemy (2011), algorithms play a crucial role in the analysis of planar graphs, emphasizing the computational and practical aspects of network research.

Distinguishing spatial planar graphs from non-spatial networks elucidates the distinct factors involved in representing different types of networks. Planar graphs are specifically applicable to networks that are limited in a spatial manner, as opposed to networks involving social or ecological interactions (Barthélemy, 2011).

The study of planar graphs provides valuable insights across various fields, highlighting the significance of comprehending their features for applications in network research. Their theoretical and practical significance is emphasized by the referenced works, namely Barthélemy (2011), who offers a thorough examination of the applications and difficulties associated with planar graphs.

$$N - E + F = 2$$

1.3 Size and structure of a network

Understanding the structure and characteristics of a network involves examining its size, density, and diameter. These metrics provide insight into the network's complexity, connectivity, and the efficiency of information or resource flow within it.

1.3.1 Size of the Network

The size of a network is predominantly governed by the quantity of nodes (N) and edges (E). The interaction between nodes and edges in a network can differ depending on the type of network. This can range from simple, linked graphs to more intricate, directed graphs that may or may not have self-connections. For example, in a basic graph, the maximum number of edges (E_{max}) can be calculated using the formula:

$$\frac{N(N-1)}{2}$$

This represents a full network where each node is connected to every other node, excluding self-connections (Barthélemy, 2011).

Size of a network refers to either the count of nodes (**N**) or, less commonly, the count of edges (**E**). For connected graphs without multiple edges (*fig 1.6*), the number of edges ranges from N-1 (in the case of a tree) to the maximum value E_{max} (in a complete graph). In a simple graph (with at most one undirected edge between each pair of nodes and no self-connections), $E_{max}$is N(N-1)/2 (*fig 1.7*). For directed graphs without self-connections, E_{max} is N(N-1), and for directed graphs with self-connections (*fig 1.8*), $E_{max}$is N^2. In graphs with multiple edges between nodes (*fig 1.9*), E_{max} is infinite (M. Barthélemy, 2011).

Figure 1.6

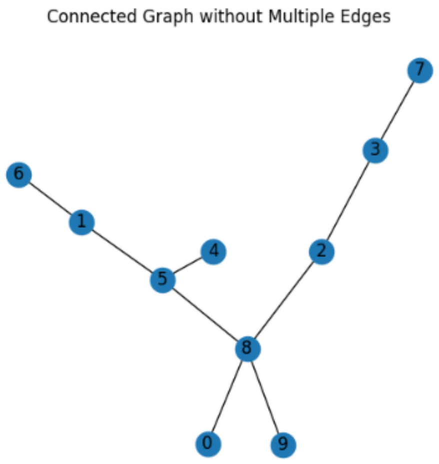


Figure 1.7

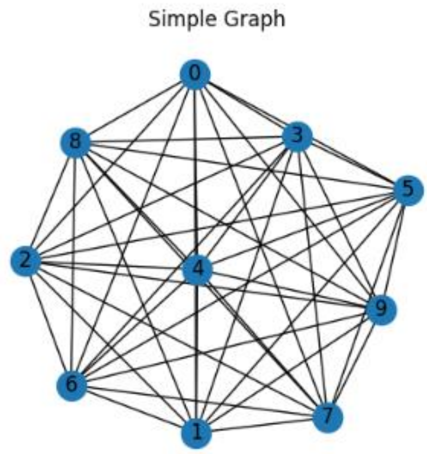


Figure 1.8

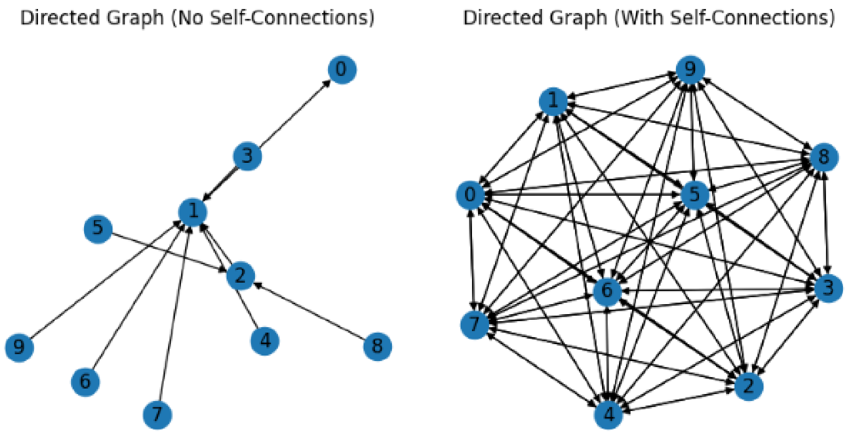
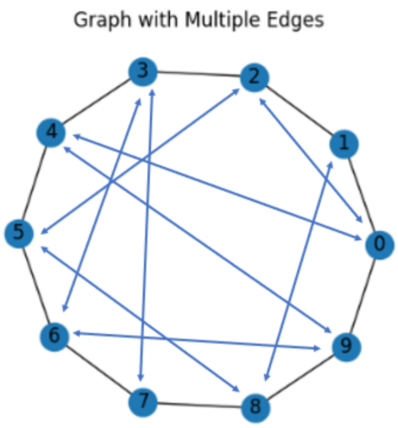


Figure 1.9



1.3.2 Network Density

Density (D) measures the degree of connectivity in a network. The connectivity ratio is a measure that represents the ratio of existing connections to potential connections. It ranges from 0, indicating no connections, to 1, indicating full connectivity.

$$D = \frac{E - E_{min}}{E_{max} - E_{min}}$$

E_{min} and E_{max} denote the lower and upper bounds on the number of edges in a linked network with N nodes. For simple graphs, the minimum number of edges, E_{min}, is equal to N-1. Density, on the other hand, is defined as:

$$D = \frac{2 * (E - N + 1)}{N * (N - 3) + 2}$$

This metric offers a deeper understanding of the network's connection and its capacity for communication and interaction between nodes. Density reflects the ratio between the total number of existing links and potential ties in a network and quantifies the connectedness of the network, with a value of unity indicating that every node is connected to every other node.

High-density networks (*fig 1.10*), such as those seen in telecommunication or the neural network of the brain, enable fast transmission of information, in contrast to the less dense connections (*fig 1.11*) found in citation or friendship networks (Zaballos et al, 2012).

Figure 1.10

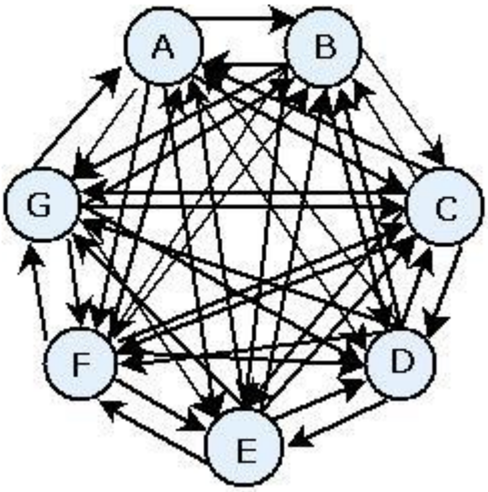
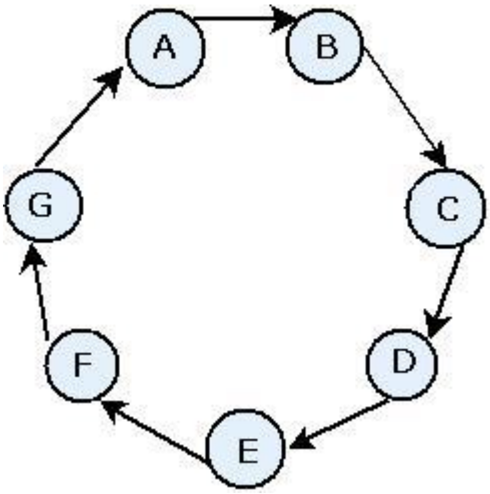


Figure 1.11

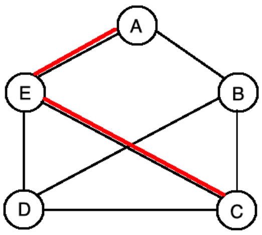


The formula above reflects the ratio between the total numbers of existing links and potential ties in a network. It quantifies the connectedness of the network, with a value of unity indicating that every node is connected to every other node (Zaballos et al, 2012).

1.3.3 Diameter of a Network

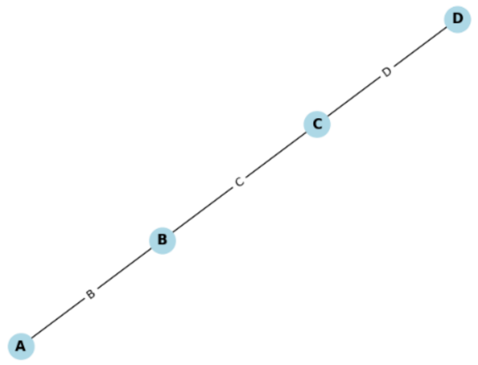
The network's diameter is the longest of all the shortest paths between any two nodes, offering a measure of the network's linear size. It reflects the maximum distance that separates any two nodes, providing insight into the network's extent and the efficiency of its connectivity.

Figure 1.12



The diameter reflects the linear dimensions of the network. For instance, if nodes A-B-C-D (*Fig 1.13*) are interconnected, the diameter is 3 (indicating 3 hops or links) when traversing from A to D.

Figure 1.13



The diameter of a graph is a measure of its size, representing the maximum eccentricity of any vertex in the graph. This means it is the greatest distance between any pair of vertices. To find the diameter of a graph, you first determine the shortest path between every pair of vertices. The longest length of any of these paths is considered the diameter of the graph.

In mathematical terms, for a network with N nodes, the diameter (**D**) is defined as the maximum among the shortest paths between any two nodes in the network. It can also be defined as the longest path (**p**) among the shortest paths between any two nodes, represented as:

$$D = \max\left(min_p\left[length\left(p_{ij}\right) - lenght(p)\right]\right)$$

The formula represents the length of the path between nodes i and j, and length(p) is a procedure that calculates the length of the path p. As an example, consider the diameter of a 4x4 Mesh network, which is 6.

In simpler terms, the diameter tells us how far apart the two most distant nodes in a network are, based on the shortest paths between them.

The concept of a network's diameter in an urban context can be reimagined by considering the longest possible route within a city's transportation network. Imagine navigating from the most remote suburb to the farthest industrial zone. This maximum travel distance showcases the urban network's reach and how effectively different city zones are connected. It underscores the urban planning challenge of ensuring accessibility across vast areas, aiming to reduce travel times and improve the quality of urban life through strategic infrastructure placement and efficient public transportation systems.

Conclusion

These three characteristics—size, density, and diameter—collectively inform the analysis of network structures, highlighting the diversity in network types and their implications for various applications, from infrastructure design to the study of social interactions.

By exploring these dimensions, we gain a comprehensive understanding of the intrinsic properties that define networks and their operational dynamics.

1.4 Primal and dual graph in urban network

The analysis of urban street networks can be approached from two main perspectives: primal and dual graphs. The primordial depiction emphasizes roadways as edges and junctions as nodes, adhering to a common methodology that accurately reflects the physical structure of urban layouts. In contrast, the dual graph technique takes the opposite perspective by considering street segments as nodes and junctions as edges. This method is commonly used in Space Syntax analysis to comprehend urban connections and the dynamics of movement (Porta et al., 2006; Hillier, 1996).

Primal graphs (*Fig 1.15*) utilize metric distances between nodes to depict the spatial linkages and connectivity of the urban environment. This method offers valuable information regarding the actual distances and efficiency of routes inside the network. Dual graphs (*Fig 1.16*) prioritize topological distances and highlight the relationships between elements rather than their physical separation. This viewpoint emphasizes the interconnectedness of the network's structural connections and how they either enable or hinder movement (Sevtsuk et al, 2012).

Figure 1.14

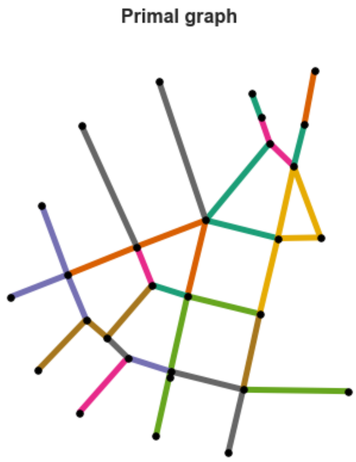


Figure 1.15

Figure 1.16

The selection between primal and dual representations has an impact on the understanding of network analysis, since each offers distinct perspectives on urban structure and operation. Primal graphs provide a straightforward assessment of the physical arrangement and ease of movement, whereas dual graphs expose the fundamental organization and effectiveness of urban connectedness.

Through a comparative analysis of these two methodologies, researchers and planners can get a thorough comprehension of urban networks, which can then be used to make informed decisions on development, navigation, and infrastructure enhancement.

1.5 Understanding Network Complexity

Network complexity refers to the complicated structures found in complex systems, where nonlinear interactions between components lead to the emergence of new phenomena and self-organization. The complexity of the network is not solely determined by the number of nodes but is significantly impacted by the diversity of its substructure (Wells, 2014).

The difficulty in measuring complexity is in differentiating between separate subgraphs and utilizing graph isomorphism to ascertain structural equivalence. Complexity metrics depend on distinct subgraph configurations obtained by adjustments such as removing edges.

Three distinct measures provide valuable insights into the complexity of a graph:

C_1e: Variability in subgraphs resulting from the removal of a single edge, assessed by the count of unique spanning trees.

C_2e e: Diversity in subgraphs post a single edge cut, evaluated through the spectra of the Laplacian and signless Laplacian matrices.

C_2e2: Subgraph variation following the removal of two edges, also analyzed via spectral differences.

These metrics illuminate the depth of interaction within networks, providing a oundation for comparing graphs of equivalent size but differing organizational complexities.

Chapter two: Network Analysis in the Urban context

2.1 Planned and Self-organized cities

The study of spatial networks can be traced back to Euler's investigation of the Königsberg seven-bridge problem in the 18th century. However, in the last decade, there has been a significant increase in the use of network analysis techniques in urban and regional studies (Turner, 2001; Porta et al., 2008; Hu et al., 2008; Crucitti et al., 2006; Scheurer et al., 2007; Jiang and Liu, 2011; Stahles et al., 2007; Ozbil and Peponis, 2011). Despite the establishment of the basic principles and measurement metrics for spatial network analysis several decades ago (Kansky, 1963; Hagget and Chorley, 1969; Tabor, 1976; Hillier and Hanson, 1984), the actual application of these methodologies remained computationally demanding, even for networks of relatively small size.

In contemporary times, the advancement of Information and Communication Technology (ICT) and Big Data has facilitated the utilization of network analysis principles in the urban design and planning process. Previously, the utilization of this type of network analysis was limited to certain applications that required extensive computational resources, such as addressing disaster planning issues, determining optimal locations for important facilities, and designing expensive utility and transportation infrastructure.

Regarding complex systems, the urban layout of a city can be classified as either a planned or a self-organized network, depending on the city. In this context, street networks can be seen as complex systems, with street intersections serving as nodes and street segments as links (Masucci, Molinero, 2016). Planned cities adhere to a predetermined layout, typically characterized by a grid-like form (Fig 2.1). Self-organized cities (Fig 2.2) have developed gradually in response to the requirements of its residents, leading to the emergence of more natural and unpredictable patterns.

Figure 2.1

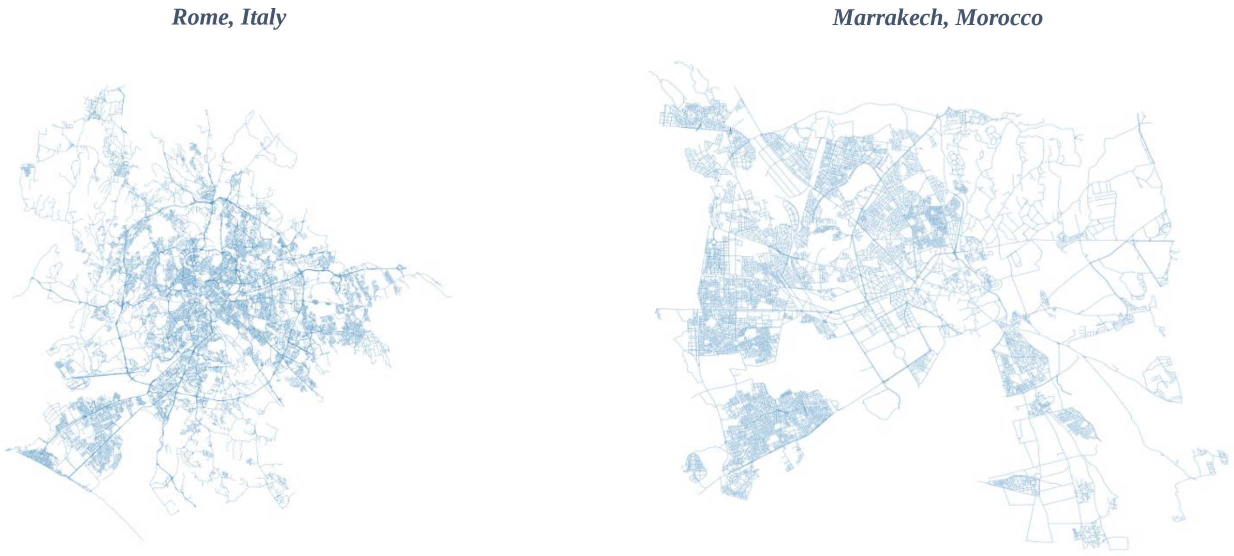


Figure 2.2

Self-organization, a characteristic feature of open and complex systems, is crucial in the spontaneous development of order, which is evident in several areas, such as urbanization. The inherent ability of a system to autonomously arrange its internal structure without being influenced by external factors is especially apparent in the development of cities. According to Porta et al. (2006), self-organized cities frequently exhibit scale-free characteristics similar to the degree distributions observed in non-spatial networks. This similarity is ascribed to their evolutionary trajectory, in which cities develop naturally over long periods of time, shaped by the requirements and choices of its residents rather than adhering to a preconceived, organized blueprint. The evolutionary processes of these cities make them similar to random graphs in terms of their structural construction.

Nevertheless, it is imperative to acknowledge that self-organization comprises a range of phenomena, such as nonlinearity, instability, fractal formations, and chaos. These phenomena are not only typical of intricate systems but also align with the core of urban life and urbanism, particularly in the late 20th century (Portugali, J.,1997). Self-organization is a term that goes beyond a particular theoretical framework. It is a broad concept that encompasses different theoretical methods. These approaches have common basic concepts but differ in how they specifically deal with these systems. These techniques prioritize different aspects of the processes, attributes, and subject issues being considered.

Various theoretical frameworks exist regarding various approaches of self-organization, such as dissipative cities, synergetic cities, chaotic cities, fractal cities, cellular automata cities, sandpile cities, and FACS and IRN cities (Portugali, J.,1997). This thesis does not try to explore all aspects of these various theoretical frameworks in detail, but rather seeks to offer a thorough comprehension of the main differentiation between urban network structure in self-organized and planned cities.

The configuration of streets and roads is of utmost significance as it enables residents to traverse the many functional elements of a city. Varied street configurations yield differing degrees of efficiency, accessibility, and utilization of transportation infrastructure. Therefore, elements such as design, arrangement, and connection of roads are attracting more attention in urban planning literature. The reference is from Kirkley et al. (2018). A roadway network can be viewed as an optimization problem aimed at minimizing transportation effort in both the Euclidean (or primal) space and the information space (Masucci, Smith, Crooks, 2016).

The majority of research in urban network analysis has depended on techniques that were originally created for social networks. Nevertheless, there are crucial modifications required when employing social network analysis techniques to urban environments. Within social networks, the interactions between network members are typically characterized topologically based on the degrees of separation between individuals. Consequently, the geographical and geometrical aspects of links hold minimal significance. Conversely, experts in the field of urban planning are typically focused on the accurate geographical connections within a spatial network. They consider factors such as distances, angles, and journey times as crucial elements in expressing the linkages and proximities between different locations (Porta et al, 2006). Furthermore, while the practice of assigning importance to individuals based on their personal traits has been uncommon in social networks, it is frequently crucial in geographical network studies.

Street networks exemplify planar graphs, wherein their edges symbolize a tangible link, distinct from the relational connections observed in several intricate networks. The geographical embedding greatly influences the network architecture, imposing limitations on the number of long-distance connections and the number of edges connected to a single node (its degree k). Although there has been much research conducted on these systems (Kirkley et al., 2018), degree-based network measures yield unremarkable results due to the pronounced peak in the degree distribution and the high values of other relevant metrics such as clustering and assortativity. On the other hand, measures that are not limited to a specific location, such as those based on network centralities, might offer further information. These measures show complex patterns in networks that are not spatially constrained and are closely related to the degree of connectivity.

The urban networks can be categorized into two types of network analysis indices based on these network metrics. The first category, known as inter-network indices, measures the characteristics of a spatial graph or subgraph. Their results acquire significance when juxtaposed with those of other networks. Intra-network indices refer to the relative relationship between each network element, such as a node or building, and the surrounding elements inside the same network. Sevtsuk (2017),

2.2 Inter network index

Inter-network indices are a useful tool for examining the general characteristics of a spatial network in a particular location. They provide valuable insights into the broader connectivity patterns between different areas. These indices are mostly used for conventional two-element networks that comprise of nodes and edges. In this context, nodes often represent street intersections, while edges indicate street segments. (Sevtsuk, 2017). Envision an urban cartography wherein the vertices represent the points of intersection and the edges symbolize the roadways.

By applying these indices to our city map, we can obtain a comprehensive understanding of the level of connectivity within the entire network. It resembles observing the map from a high altitude to discern the whole network of relationships. The numerical values obtained from these indices provide valuable insights into crucial aspects of the network, enabling us to gain a deeper understanding of the functioning of cities. Hence, the result obtained by applying these indices to our graph should be seen as a characteristic of our network at the "global" scale. Consequently, the graph's properties that are important for studying urban networks will be described by the value supplied by the evaluated Inter-network functions.

The gamma index is a significant indicator. The purpose of measuring the connectivity of intersections and roads is to assess their level of integration. On the other hand, the cyclomatic number quantifies the various possible routes between two locations (Scott et al., 2006). The presence of other pathways indicates that the city possesses enhanced capabilities in managing disturbances, such as the closure of roads. According to Sevtsuk (2017), a city is considered more resilient if it offers many routes to reach a goal.

The second index refers to the cyclomatic number, which is a significant measure in network research. It highlights the existence of several pathways connecting nodes rather than just one unique route (Ingvardson & Nielsen, 2018). Understanding the concept of network redundancy is crucial for comprehending the resilience of a network and its ability to retain functionality in the presence of disturbances. (Sevtsuk, 2017). Redundancy refers to the presence of additional alternatives or options. It is akin to possessing additional keys as a precautionary measure in the event of misplacing one. In the context of networks, redundancy refers to the presence of multiple routes, which becomes particularly crucial following disasters. Thus, in the event of road damage, alternative routes remain available for transportation. The reference is from Khademi et al. in 2015.

The cycle idea in network analysis refers to the maximum number of cycles, which indicates the independent cycles inside the network (Sevtsuk, 2017). Independent cycles serve as alternative routes in the event that certain paths get obstructed. They play a vital role in intricate networks. Consider them as alternative routes that enable us to continue progressing even when the primary path is obstructed. The cyclomatic number is a significant metric in network analysis that highlights the existence of several routes connecting nodes rather than singular connections.

The Characteristic Path Length (CPL) represents the mean distance that needs to be traveled between any two locations. A reduced CPL indicates a network that is both well-connected and efficient (Porta, Crucitti, & Latora, 2006). If one is able to traverse between two locations swiftly, it indicates that the urban area has been effectively planned to facilitate efficient mobility.

The Global Efficiency metric is analogous to the Common Path Length (CPL) metric, but it specifically emphasizes the ease of communication between all sites. The citation is from Porta, Crucitti, and Latora's work in 2006. The city is considered well connected as long as diverse parts can easily communicate with one other, even if not all connections are direct.

Ultimately, inter-network indices provide us with a comprehensive perspective on the interconnectedness of cities. They facilitate our comprehension of the seamless progression of events and the city's ability to adapt to alterations and disturbances.

2.2.1 Gamma index

The gamma index is a metric that quantifies the connectivity of a network by considering the ratio between the observed number of linkages (edges) and the maximum number of links that could exist. The index spans from 0 to 1 and offers a measurable value to demonstrate the extent of connectedness in the network (Scott et al. 2006). A gamma value of 1 indicates a network that is fully connected or complete, meaning that every node is directly linked to every other node. Nevertheless, in actual networks, the existence of a fully interconnected configuration is exceedingly uncommon, if not unattainable. In contrast, a score of 0 would indicate a network devoid of any connections.

The gamma index provides a pragmatic method for monitoring the evolution of a network over time. Through the examination of variations in the gamma index, one can detect patterns in the development of a network's connection, indicating whether it is growing more interconnected or more fragmented.

Within the realm of spatial networks, such as the street network of a city, the gamma index serves as a metric to quantify the extent to which the network resembles a graph that is completely interconnected. The index (γ) is applicable to planar networks. The calculation is performed using the formula:

$$y = \frac{e}{e_{max}}$$

where $\boldsymbol{e}$ is the actual number of links in the network and $\boldsymbol{e_{max}}$ the maximum number of links in the network (i.e., all nodes are completely connected). In turn, $\boldsymbol{e_{max}}$s computed as

$$e_{max} = 3(v - 2)$$

Where v is the total number of nodes in the network. As the gamma index increases, the network's internal connectedness becomes stronger. Hence, a high gamma index indicates that a traveler can go more directly between various crossings in the street network, indicating a more efficient and accessible network layout (Sevtsuk, 2017). Hence, this metric might be crucial in urban planning, traffic engineering, and other related fields that necessitate a comprehensive comprehension of network connection.

2.2.2 Cyclomatic Number

The cyclomatic number is a significant metric in network analysis, highlighting the existence of many routes between nodes rather than singular connections. The concept of network redundancy, sometimes referred to as network resilience, is crucial for comprehending the network's ability to sustain functionality in the presence of disturbances.

The cyclomatic number, which is the count of fundamental circuits or loops in a network, is a useful indication of network robustness. This study is remarkable for its ability to clearly and effectively demonstrate the potential for movement and the overall strength of the network (Zhang, Miller-Hooks, & Denny, 2015). The term "fundamental circuit" pertains to the closed circuits present in a network that provide many potential routes for travel (Ingvardson & Nielsen, 2018).

The concept of circuits can be compared to the openings in a fishnet, with each one symbolizing a complete path or loop inside the network. The cyclomatic number can be calculated using the formula:

$$\mu = e - v + g$$

The equation defines the variables as follows: e represents the number of edges, v represents the number of nodes, and g indicates the number of connected components of the network. The cyclomatic number, within the framework of urban street networks, quantifies the number of possible route options and the resilience of the network by considering these fundamental elements (Sevtsuk, 2017).

The cyclomatic number is directly correlated with the robustness of a network. Networks with a larger cyclomatic number often have more resilience when the average degree and diameter are held constant. The explanation for this resides in the intrinsic redundancy of cycles: even if a single connection, or 'arc', is eliminated, the network's connectivity remains intact. The presence of alternate routes can be extremely helpful during disturbances, since it guarantees the uninterrupted operation of the network (Zhang, Miller-Hooks, & Denny, 2015). The availability of cycles in metropolitan networks is crucial for providing alternative routes when some street segments are inaccessible owing to disruptive events, maintenance work, or occurrences like a marathon.

2.2.3 Maximum number of cycles

The notion of independent cycles in a graph encapsulates the concept of the multitude of possible routes inside a network. In complicated networks, the presence of independent cycles, which are essentially alternate routes, becomes increasingly important. These cycles provide redundancy of paths, which is vital to ensure the stability and functionality of the system.

Within the framework of uncomplicated structures like tree graphs, the independent cycles hold no value as they are completely absent. A tree graph is characterized by having v-1 edges, where v represents the number of vertices or nodes. This arrangement establishes a distinct and shortest route between any two nodes in the tree, eliminating the possibility of alternative paths.

From an urban design standpoint, tree networks, which are frequently found in suburban areas, necessitate the minimal usage of road construction materials, such as asphalt, to connect a specific group of locations. Nevertheless, these networks frequently generate hierarchical organizational patterns and significantly restrict the possibilities for trip routes. Although this architecture is economical, it may impede efficient movement and render the network vulnerable to disturbances (Sevtsuk, 2017).

To get the maximum number of cycles in a network with a given number of vertices, you subtract the number of edges in a tree graph from the maximum number of edges in a complex graph.

$$Max.Cycles = \frac{(v^2 - v)}{2} - (v - 1)$$

This formula elucidates the basic disparity between uncomplicated tree networks and intricate urban networks. The complexity of a network increases as the number of cycles it contains increases, and conversely, the complexity decreases as the number of cycles decreases.

The quantity of cycles within a network is a significant indicator of the network's degree of advancement and intricacy. Within the framework of a transportation system, a greater quantity of cycles is generally associated with a sophisticated and well-developed system that offers a wide range of travel routes and is capable of withstanding potential disruptions (Sevtsuk, 2017).

2.2.4 Redundancy Index

Redundancy, in the context of engineering, focuses on duplicating essential components of a system to enhance the overall reliability of the system. Within the field of transportation network analysis, this principle is demonstrated by taking into account alternative routes in the event of disruptions to main roads, hence improving the network's ability to withstand and recover from such disruptions.

Khademi et al. (2015) employ a distinctive approach by examining the susceptibility and reaction of transportation networks during catastrophic events, specifically in the aftermath of intense earthquakes. This novel use highlights the importance of having alternative road connections, thus demonstrating the strength of the network in the face of potential disruptions.

The redundancy index is a quantitative measure of redundancy in a network. It is calculated by comparing the number of observed cycles (routes) to the maximum potential cycles in the network. The formula is as stated:

$$\frac{e - v + g}{\left[\left[\left(\frac{v^2 - v}{2}\right)\right] - v + 1\right]}$$

The formula utilizes the variables e, v, and g to respectively denote the number of edges or connections, the number of vertices or nodes, and the number of connected components in the network.

The redundancy index, which varies between zero and one, offers valuable information about the network's resilience. A redundancy index of zero signifies that the network is a tree network, characterized by the presence of only one distinct path connecting any two nodes. A redundancy index of one indicates a fully connected network, where each node is directly connected to every other node.

Hence, the redundancy index serves as a potent means to evaluate a network's susceptibility to disruptions, such as those induced by natural calamities like floods or mudslides. Urban planners can utilize this index to acquire crucial information on designing and managing transportation systems that can effectively endure and recover from such disruptive events Sevtsuk (2017).

2.2.5 Characteristic Path Lenght

The Characteristic Path Length (CPL) is a fundamental metric used to analyze urban transportation networks. It represents the average distance or duration of the shortest pathways between every pair of nodes in the network. The CPL, or Connectivity Performance Level, is a crucial metric that gauges the effectiveness, durability, and ease of use of a transportation system. A lower CPL often indicates a more efficient and interconnected network. The citation is from Porta, Crucitti, and Latora's work in 2006.

The CPL measures the mean journey distance or duration needed to go between any two points in the network. Therefore, it offers a thorough overview of the interconnectedness and convenience of travel throughout the entire urban transportation system, providing significant insights into its design, structure, and effectiveness.

In the field of network science, the characteristic path length (CPL), commonly denoted as L, is considered a key measure for assessing the connectivity characteristics of a network. When analyzing directed and tightly connected graphs, the concept of the CPL provides insights into the effectiveness of the transmission of information, resources, or traffic between every pair of nodes. This makes it a valuable tool for urban planners and transportation analysts.

Researchers utilize various strategies to ascertain the CPL. An extensively employed method utilizes Dijkstra's algorithm, a renowned algorithm that calculates the most efficient route between two nodes based on edge weights or journey times. This algorithm employs the notion of 'relaxation' to iteratively reduce the length of the computed shortest path between nodes until no shorter path can be discovered.

In this study, the characteristic path length (CPL) is determined by adding together the shortest distances (in meters) between every conceivable pair of nodes in the network, and then dividing this sum by the total number of pairings (Watts and Strogatz, 1998). The formula that represents this is:

$$L = \sum \frac{dij}{\left(N * (N-1)\right)}$$

The formula utilizes the variable dij to represent the shortest distance between nodes i and j, and N represents the total number of nodes in the network. Consequently, the resulting connectedness Path Length (CPL) serves as a crucial indicator of network connectedness, providing a quantitative assessment of the total accessibility of the urban transportation system (Porta, Crucitti, & Latora, 2006).

2.2.6 Global Efficiency

The Global Efficiency of a network is a reliable measure that assesses the efficiency of communication between nodes in the network. The Global Efficiency, like the Characteristic Path Length, offers valuable information into the connectivity of the network. However, it also provides a distinct benefit: it can be applied to networks that are not completely interconnected. The Global Efficiency feature is highly beneficial for assessing the overall interconnection and communication efficiency of various networks, such as social networks, computer networks, or transportation networks. (Porta, Crucitti, & Latora, 2006)

The Global Efficiency metric successfully addresses a constraint that is inherent in the calculation of Characteristic Path Length, which necessitates a network that is fully connected. The Global Efficiency hypothesis posits that the effectiveness of communication between two nodes is inversely related to the shortest distance (dij) between them. If the network (G) is not completely connected and there is no path linking nodes i and j, the shortest path length, dij, is assumed to be infinite. (Porta, Crucitti, & Latora, 2006)

The mathematical expression for the Global Efficiency of a network or graph is as follows:

$$e^{glob} = \frac{1}{N(N-1)} \sum_{i,j \in N; i \neq j} \frac{1}{d_y}$$

The formula uses N to denote the total number of nodes in the network, and d_ij to indicate the shortest path length between nodes i and j. The summing is performed across all unique pairings of nodes in the network (Porta, Crucitti, & Latora, 2006).

When working with metric systems or valued graphs, one can normalize the Global Efficiency by dividing it by the efficiency of an ideal system. In an optimal system, each pair of nodes, i and j, are directly linked by an edge with a length equivalent to the Euclidean distance between the nodes.

The calculation of the ideal efficiency is determined by the following equation: In this equation, N represents the total number of nodes in the network, while dij represents the shortest path length between nodes i and j. The summing is performed across all unique pairings of nodes in the network. (Porta, Crucitti, & Latora, 2006)

When working with metric systems or valued graphs, one can normalize the Global Efficiency by dividing it by the efficiency of an ideal system. In an optimal system, there is a direct connection between any two nodes, i and j, represented by an edge with a length equal to the Euclidean distance between the nodes.

The calculation of this optimal efficiency is determined by the following formula:

$$Ideal\ e^{glob} = \frac{1}{N(N-1)} \sum_{i \neq \in N} \frac{1}{dij^{Eucl}}$$

In this context, the symbol dij^{Eucl} denotes the Euclidean distance between nodes i and j, which can be seen as the direct length of a straight line connecting these nodes.

The Normalized Global Efficiency is calculated by dividing the Global Efficiency of the real network by the Global Efficiency of the ideal network, as indicated in the formula:

$$normalized\ e^{glob} = \frac{e^{glob}}{Ideal\ e^{glob}}$$

Such normalization offers a standardized measure that aids in the comparative analysis of different networks, improving our understanding of network efficiency and connectivity in a broader context. (Porta, Crucitti, & Latora, 2006)

2.3 Intra network indices

Intra-network indices offer a complete framework for evaluating the relative importance of individual nodes within a network. These indices provide a method for assessing the degree to which separate components within a specific network are connected to the larger network as a whole (Sevtsuk, 2017). In the subsequent part, we embark on an examination of six distinct indices that can be easily utilized for the evaluation of urban infrastructure networks.

As we progress, we lay the groundwork to define and analyze several measures of centrality. Freeman (1977; 1979) introduces three crucial indices under the concept of centrality. These indices can be generally classified into two distinct groups, as described by Latora and Marchiori (2004). Both

Degree Centrality (CD) and Closeness Centrality (CC) can be understood as measures of how close an individual is to others in a network, as discussed by Freeman (1977; 1979), Nieminen (1974), Sabidussi (1966), Scott (2003), and Shimbel (1953). Conversely, Betweenness Centrality (CB) quantifies centrality by assessing the extent to which one individual mediates connections among others (Anthonisse, 1971; Freeman, 1977; 1979; Freeman et al, 1991; Newman and Girvan, 2003).

The growing fascination with network research of intricate systems has stimulated the creation of innovative centrality indices. In this study, two indices, Efficiency Centrality (CE) and Straightness Centrality (CS), are important. Both of these indices are based on the concept of global efficiency. Efficiency Centrality (CE) is similar to closeness centrality, especially when used in the context of geographical graphs. Significantly, the normalization process involves comparing the lengths of the shortest pathways with those of hypothetical straight lines connecting the same nodes (Vragov et al, 2004). This phenomenon leads to the emergence of a new geographical concept called Straightness Centrality (CS). Straightness Centrality is defined as the degree of centrality that is determined by the capacity to directly reach all other nodes within the network.

The following section provides a strong foundation for comprehending the different measures of centrality and their significance in evaluating the interconnection and value of nodes within urban infrastructure networks. Afterwards, we thoroughly examine the six chosen indices, providing a detailed study of their practical uses and the impact they have on network analysis.

2.3.1 Degree Centrality

Degree Centrality (C^D) is a crucial statistic in network analysis, used to quantify the prominence or significance of a node inside a network. It measures a node's connectedness by considering the quantity of connections or relationships it has with other nodes. According to the logic of Degree Centrality, nodes with a greater number of connections, or higher degrees, are considered more central or important within the network.

The Degree Centrality of a node i can be mathematically defined as the total number of edges connected to that node, divided by the greatest number of connections it might have in the network. The formula for computing Degree Centrality can be stated as follows:

$$C_i^D = \frac{k_i}{N-1}$$

The variable $k_i$in this equation represents the degree of node i, which corresponds to the number of connections or edges it possesses. The symbol N represents the total number of nodes in the network.

Degree Centrality values span from 0 to 1. A Degree Centrality score of 1 signifies that a node has direct connections with all other nodes in the network. This indicates the highest level of importance and impact of that node in the network, as it acts as a central point of connection for all other nodes (Crucitti et al. 2006).

Within the framework of analyzing an urban transportation network, the Degree Centrality metric can provide very valuable insights. Through the computation of Degree Centrality for each node, we may obtain a comprehensive comprehension of its relative significance within the transportation system. Nodes with a greater Degree Centrality are likely to represent significant transportation hubs or important intersections, playing a crucial role in connecting and facilitating the flow within the network. Urban planners and transportation authorities can utilize this knowledge to inform their decision-making processes, specifically in areas such as allocating resources, optimizing networks, and planning for contingencies. It is important to note that degree centrality is not highly significant in primary urban networks, where the number of connections a node can have is restricted by geographic limitations (Crucitti et al., 2006).

2.3.2 Betweenness Centrality

Betweenness Centrality is a global metric in network theory that uses paths to analyze and anticipate congestion and load on networks. The main purpose of this measurement is to assess the ease of reaching a particular point within a network. It does so by evaluating the significance of a node based on the number of shortest paths that pass through it (Masucci et al, 2016). This measure of centrality quantifies. The betweenness centrality of a node in the dual graph, which depicts the connections between urban roads and crossings, indicates its importance in facilitating efficient traffic flow and connectivity throughout the city (Kirkley et al., 2018).

Betweenness Centrality (C^B) is a metric that quantifies the degree to which a node functions as a mediator or connector inside a network. Nodes with high Betweenness Centrality have a significant influence on the movement of information, goods, or people in a network. This is because they are located on the most efficient routes between other nodes.

The mathematical definition of the Betweenness Centrality of a specific node i is as follows:

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k \in N; j \neq k; j,k \neq i} \frac{n_{jk}(i)}{n_{jk}}$$

The equation definesn_{jk} as the total count of shortest paths connecting nodes j and k. Additionally, n_{jk} (i) represents the count of these pathways that contain the node i under consideration. Nodes with high Betweenness Centrality ratings play a crucial role in preserving the connectedness of the network, indicating their essentiality to the overall functionality of the network.

When implementing this metric in the analysis of an urban transportation network, we shall compute the Betweenness Centrality for every node. This enables us to pinpoint the nodes that play crucial roles in guaranteeing the efficient operation and strong connectivity of the urban transportation system.

Nonetheless, a thorough examination of Betweenness Centrality necessitates the inclusion of multiple additional criteria in order to precisely ascertain the shortest path. These criteria could include the number of connections in the purely topological scenario, the shortest straight-line distance between two sites if the edges are assigned weights based on such lengths, or edge weights determined by a cost function such as capacity or speed constraints in transportation networks. In more intricate situations, these considerations may be combined to determine the most optimal routes.

Furthermore, the utilization of Betweenness Centrality (C^B) can serve as a substitute for anticipated traffic movement in a transportation network by integrating this structural data into the edge weights. Consequently, nodes with high betweenness centrality are expected to manage larger amounts of traffic. Therefore, urban planners and policymakers can employ this metric to efficiently create and oversee transportation systems, thereby improving the movement of people and goods, decreasing traffic congestion, and guaranteeing the durability of the network.

2.3.3 Closeness Centrality

Closeness Centrality (C^C) is a crucial measure in network research that highlights the significance of a node by considering its proximity to all other nodes in the network. The importance of this measure lies in its ability to evaluate the velocity or effectiveness with which a node may connect with other nodes in the network. The accessibility or proximity of a node is a critical factor in determining its influence inside a network setting.

The Closeness Centrality of a certain node, represented as i, is mathematically defined as:

$$C_i^C = \frac{N-1}{\sum_{j \in G, j \neq i} dij}$$

The symbol dij represents the shortest path distance, also known as the geodesic, from node i to node j. It is crucial to remember that the Closeness Centrality measure is most significant for connected networks, where there is a path between every pair of nodes.

By utilizing Closeness Centrality, we can obtain a distinct viewpoint on the structure and functioning of the network. More precisely, it enables us to identify the nodes that are most efficient in distributing or transmitting information, items, or people because of their strategic position within the network. This holds great significance in diverse scenarios, including the spread of information in social networks, the effective distribution in supply chain networks, or the rapid evacuation in transportation networks during catastrophes (Crucitti et al. 2006).

Moreover, within the framework of urban transportation networks, nodes exhibiting high Closeness Centrality are likely to experience shorter average journey durations to all other nodes, rendering them optimal sites for establishments like transit hubs or distribution centers. On the other hand, regions of the network that have nodes with low Closeness Centrality can be focused on for enhancing connectivity in order to decrease average travel durations and enhance overall network effectiveness (Crucitti et al. 2006).

The investigation of the closeness centrality spectrum allows us to gain insights into the local behavior of the primal graph in terms of transportation efficiency on the original street network (Masucci et al, 2016). Hence, utilizing the Closeness Centrality metric can provide useful information for urban planners, network analysts, and policymakers who aim to enhance network performance, enhance connectedness, and promote efficient flow within a system.

2.3.4 Efficiency Centrality

The increasing interest in analyzing complex systems using network analysis has resulted in the development of novel centrality indices. Three relevant factors for this study are efficiency, straightness, and information, all of which are based on global efficiency. Efficiency centrality (C^E) is a measure of proximity that is specifically used for geographic graphs. It is normalized by comparing the length of the shortest pathways with the length of virtual straight lines between the same nodes (Vrago et al, 2004).

Efficiency Centrality (C^E) is a crucial network measure based on the idea that the effectiveness of communication between two separate nodes, labeled as i and j, is inversely proportional to the shortest distance dij between them. This statistic quantifies a node's capacity to efficiently and directly exchange information with other nodes. The Efficiency Centrality of a certain node i is calculated by taking the reciprocal of the total of the shortest path distances from node i to all other nodes in the network. The sum is then normalized by the direct Euclidean distance between these nodes, which represents the straight-line distance between them in a geometric space.

From a mathematical standpoint, this can be represented as:

$$C_i^E = \frac{\sum_{j \in N; j \neq i} \frac{1}{d_{ij}}}{\sum_{j \in N; j \neq i} \frac{1}{d_{Eucl_{ij}}}}$$

In this equation, d_{ij} represents the minimum distance between nodes i and j, while $d_{Eucl_{ij}}$ denotes the Euclidean distance between these nodes.

Efficiency Centrality provides a novel viewpoint on the dynamics of network connectivity. A node with a high Efficiency Centrality score indicates its ability to efficiently communicate with neighboring nodes across relatively short pathways. This indicates a significant degree of efficacy in the transmission of information, allocation of resources, or movement of transportation. Nodes with high Efficiency Centrality are thus expected to serve as crucial hubs in the network, effectively linking various sections of the network.

In the context of urban transportation networks, nodes with high Efficiency Centrality are important junctions or intersections that facilitate the smooth movement of traffic throughout the city. These nodes might be regarded as pivotal sites for overseeing the movement of traffic, establishing transportation centers, or allocating resources for infrastructure development.

Hence, the utilization of Efficiency Centrality proves highly advantageous for urban planners, network analyzers, and policymakers in discerning these pivotal nodes, optimizing network architecture, and augmenting the network's total performance.

2.3.5 Straightness Centrality

Straightness Centrality (C^S)) is a metric derived from Efficiency Centrality that employs a different way of normalization to assess the importance of nodes in a network. This measure specifically indicates the extent to which the paths connecting a node to all other nodes in the network stray from being direct or linear.

Straightness Centrality for a given node, represented as i, is determined by dividing the Euclidean distance to the shortest path distance from node i to all other nodes. Subsequently, the aforementioned ratio is multiplied by the total count of nodes in the network, except one. This can be mathematically represented as:

$$C_i^S = \sum_{j \in N; j \neq i} \left(\frac{d_{Eucl_{ij}}}{d_{ij}} \right) * (N-1)$$

In this expression, d_{ij} signifies the minimum distance between nodes i and j, while $d_{Eucl_{ij}}$ indicates the geometric distance between these nodes. The Euclidean distance between two nodes is the direct distance between them, assuming they are present in a geometric space.

Straightness Centrality is unique in its ability to measure the extent to which the paths connecting a certain node i to all other nodes diverge from being direct or straight. A high Straightness Centrality score signifies that the connections between a node and all other nodes in the network closely resemble direct, linear routes. To clarify, a node is considered to have a high Straightness Centrality when it is positioned on paths that are as straight as possible, considering the overall configuration of the network.

This measure can be very advantageous in comprehending the function of nodes in path selection. For example, in transportation networks, nodes with high Straightness Centrality are crucial intersections that enable efficient and direct routes around the city. This level of understanding can assist urban planners in formulating efficient route plans and improving the overall efficiency of the network.

Chapter three: Decoding Urban Structures: Multiple Centrality Assessment

The study of urban structures and their influence on human behavior and functionality greatly relies on understanding street networks. Urban street analysis encompasses various scientific disciplines, including transportation planning, land-use planning, and economic geography. This section provides an overview of the motivations for investigating street networks, followed by an analysis of relevant literature in this specific field. Following that, a multiple centrality assessment (MCA) will be provided for various city street networks. The analysis commences with the research conducted by Porta, S., Crucitti, P., and Latora, V. in 2006. The analysis will investigate the relationship between car-related road accidents and different network characteristics. The features mentioned are characterized by the intra and inter network measurements outlined in chapter 2. These measures are computed for the cities under investigation.

Examining street networks in urban and territorial settings has been a long-standing tradition and is a primary area of interest in the field of urban studies. It is an essential instrument for comprehending the complexities of urban systems, particularly in relation to accessibility, spatial arrangement, and the interrelationships among different urban components. Researchers have analyzed various aspects of urban streets, including their roles in transportation networks, the proximity of key locations, network integration, connectivity, transportation costs, and the level of effort required for transit within metropolitan areas.

As we proceed to the next section, we shift our focus towards applying the theoretical framework developed in the previous chapters to conduct a practical analysis of particular towns. The aforementioned inter-network and intra-network indices will have a vital role in this endeavor. The aim of this analysis is to provide a thorough assessment of urban networks, including both quantitative data and a qualitative understanding of how these networks affect the daily lives of city dwellers.

This approach represents a shift from theoretical principles to practical investigation, linking abstract concepts with actual urban situations, with a particular focus on car accidents as spatial incidents to be evaluated in the research.

The purpose of this chapter is to clarify the relationship between network indices and real-life events that happen on a daily basis, and to present them as important data in existing models and data-driven decision-making processes. The upcoming chapters are anticipated to improve our understanding of urban systems, offering insights that are both methodologically robust and contextually complex.

3.1 Methodology

The methodology employed in this thesis for analyzing urban traffic safety and car crash data integrates cutting-edge tools and datasets to construct, analyze, and visualize the complexities of city road networks and their relation to traffic accidents. Initially, the urban network of the city was constructed using OSMnx, a Python package that facilitates the download, modeling, analysis, and visualization of street networks from OpenStreetMap. OSMnx allows for the efficient modeling of walking, driving, or biking networks with a minimal coding effort, providing a robust foundation for subsequent analysis. This tool not only enables the analysis of street networks but also offers functionalities for working with urban amenities, building footprints, transit stops, and various other geospatial features, making it an invaluable resource for comprehensive urban analysis.

Following the construction of the city network, the study employed NetworkX, another Python package, to compute various centrality measures. NetworkX is designed for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks, facilitating a detailed examination of the centrality measures that are crucial for understanding the roles of specific roads and intersections within the urban traffic network. This analysis helps in identifying key nodes and pathways that significantly influence traffic flow and safety within the city.

The final component of the methodology involved integrating car accident data from a comprehensive countrywide traffic accident dataset. This dataset, which covers 49 states of the United States since February 2016, compiles information from multiple data providers, including APIs that provide real-time traffic event data. The dataset, currently encompassing approximately 1.5 million accident records, is sourced from a variety of entities such as the US and state departments of transportation, law enforcement agencies, traffic cameras, and sensors within the road networks. The integration of this accident data into the city network model allowed for a nuanced analysis of accident locations in relation to the network's structural characteristics. By mapping accidents to the nearest nodes in the network, this study explores the correlation between centrality measures and accident frequencies, providing insights into the spatial distribution of traffic accidents and highlighting potential areas for safety improvements.

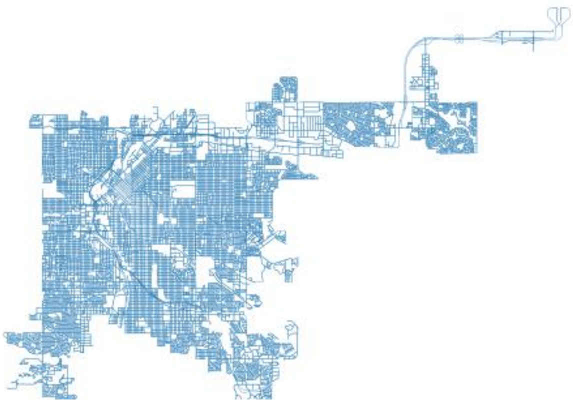
This methodological approach, combining OSMnx and NetworkX for network analysis with comprehensive accident data integration, offers a detailed and nuanced understanding of urban traffic safety, enabling the identification of critical factors contributing to traffic accidents within urban environments.

3.2 Decoding Network structure: Planned and Self organized cities

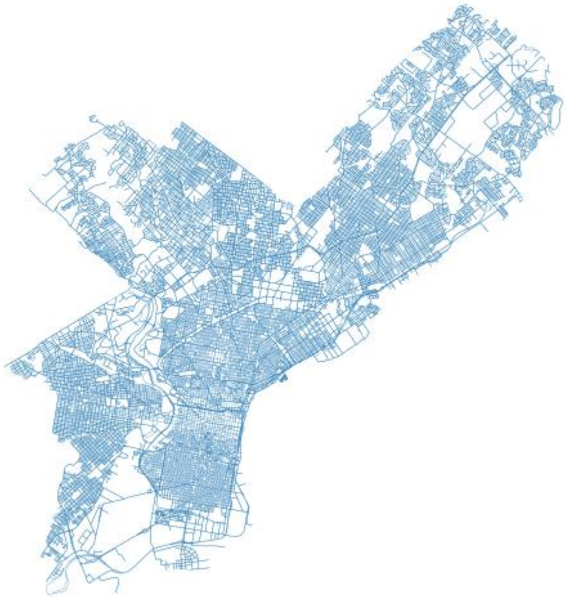
The start of our journey involves a basic differentiation between two typical types of urban development: planned cities and self-organized cities. This division not only provides insight into the historical and cultural influences that shape urban environments, but also establishes a foundation for a detailed examination of their spatial attributes.

Planned cities, characterized by their organized and predictable structures, provide a distinct chance to examine the influence of intentional design on urban existence. In contrast, the natural development of self-organized cities offers a different viewpoint, emphasizing the influence of spontaneous occurrences in forming metropolitan areas. This contrast establishes the foundation for a more thorough examination of how these distinct urban configurations impact many facets of urban existence, ranging from transportation to the ability to withstand and recover from disturbances.

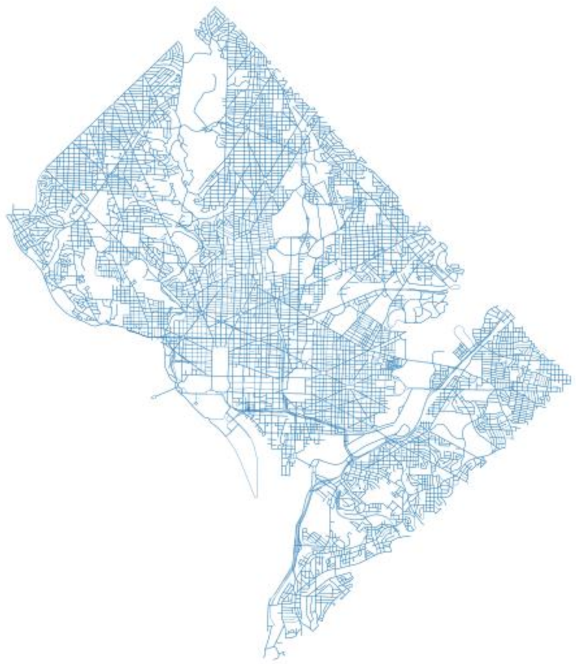
Urban settlements that have been carefully designed and organized in advance: The cities chosen for this analysis exemplify the concepts of intentional urban development, distinguished by their purposeful design and well-structured layouts. These planned towns usually follow a particular architectural vision or urban planning philosophy. They are characterized by organized and frequently grid-like street patterns, which represent a logical approach to urban design. The structured organization of urban areas typically promotes easy navigation and even distribution of urban services, demonstrating the intention and foresight of urban planners. However, it may occasionally lack the flexibility to naturally adjust to changing urban needs.



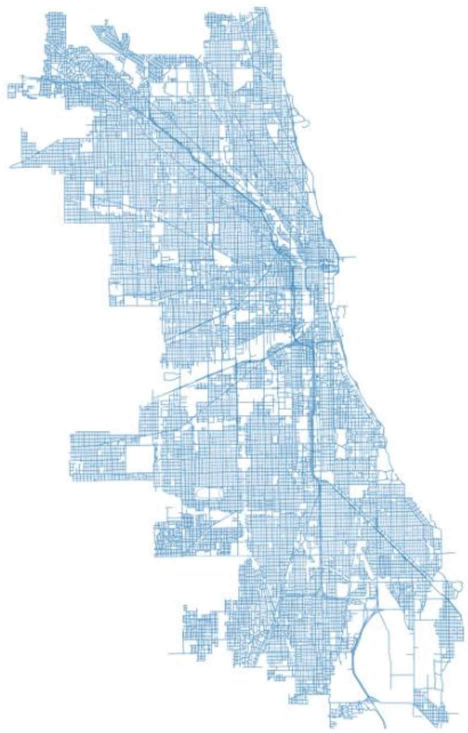
Denver, Colorado: *Denver was planned with a grid street pattern, and its streets are oriented to the four cardinal directions. The city's layout was designed for simplicity and ease of navigation.*



Philadelphia, Pennsylvania: *One of the oldest cities in the USA, Philadelphia was designed with a grid street system by William Penn. The city's layout was innovative for its time, featuring organized blocks and five public squares.*



Washington D.C.: *As the capital, Washington D.C. was meticulously planned by Pierre Charles L'Enfant. It is known for its broad, tree-lined avenues, uniform city blocks, and prominent landmarks and monuments.*

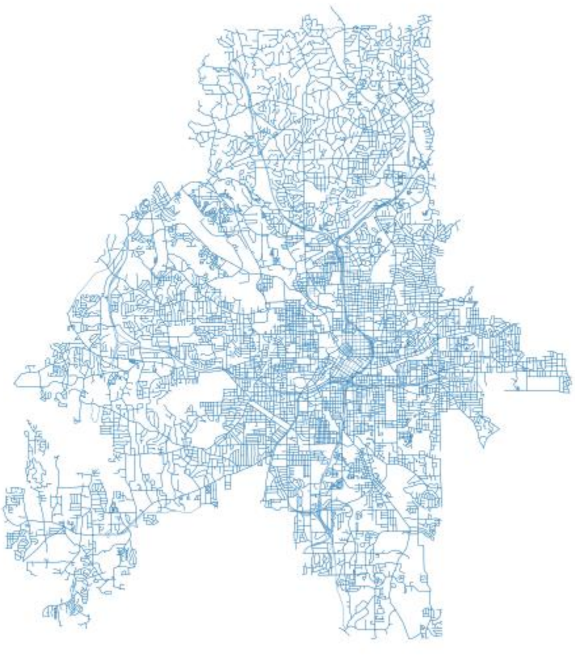


Chicago, Illinois: *After the Great Chicago Fire of 1871, much of the city was rebuilt using a planned grid system. The layout was carefully planned to accommodate future growth, with wide streets and avenues.*

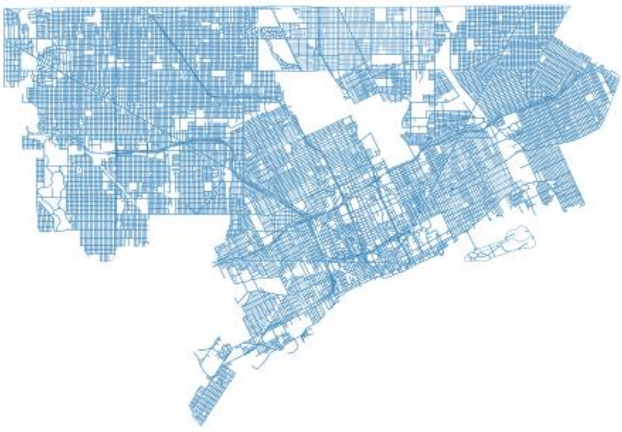
Self-Organized Cities: In contrast, self-organized cities develop in a more natural manner, influenced by the collective activities and decisions of their residents over a period of time. These cities frequently have a more intricate and less foreseeable arrangement of streets, which mirrors the unplanned and adaptable character of their development. These patterns emerge as a result of the specific needs, limitations, and circumstances related to the geography, history, and socio-economic conditions of each city. Although this might result in a more complex and perhaps less effective urban structure, it frequently gives these cities a distinct personality and flexibility to adjust to local circumstances.

The analyzed cities below serve as prime examples of self-organized urban development, wherein the growth and arrangement of the city are shaped by a blend of natural landscapes, economic factors, historical occurrences, and community-led expansion, resulting in distinctive and frequently intricate urban structures.

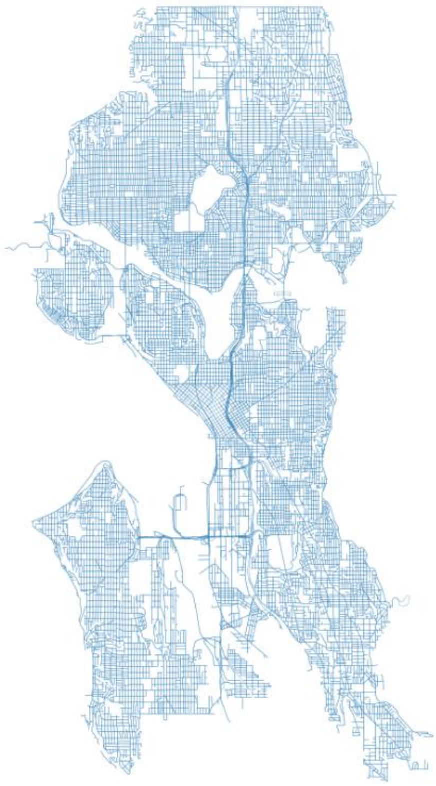
San Francisco, California: *Influenced by the Gold Rush and rapid population growth, San Francisco's urban layout developed organically. The city's streets adapt to its hilly terrain, resulting in a unique and varied urban fabric*



Atlanta, Georgia: *Atlanta's growth was significantly influenced by its status as a major railroad hub and later by the proliferation of automobile traffic. Its urban layout, particularly outside the planned Midtown and Downtown areas, evolved organically, leading to a sprawl-like pattern with a mix of dense urban centers and extensive suburban areas.*



Detroit, Michigan: *Originally a hub for the automotive industry, Detroit's urban layout evolved with the growth of the automobile sector. The city's development was influenced by industrial needs, economic booms and declines, and the resulting population shifts, leading to a street network that reflects a combination of planned segments and more organically developed areas.*



Seattle, Washington: *Seattle's urban development was largely influenced by its geographical setting between Puget Sound and Lake Washington. The city's layout reflects a combination of grid-like patterns in some areas and more organically developed zones, especially in neighborhoods shaped by the city's hilly terrain and waterways.*

3.2.1 MCA Part 1 - Inter-Network Indices

In this chapter, we embark on a comprehensive exploration of Multiple Centrality Assessment (MCA) with a focus on Inter-Network Indices. The objective is to dissect and compare the properties of various urban street networks across different U.S. cities. This analysis is crucial as it provides a deeper understanding of urban morphology and its implications for urban planning, transportation, and social interaction within these cities.

The study utilizes a diverse array of data points to paint a detailed picture of the urban networks. These include:

Number of Nodes: This metric represents the total count of intersections or junction points within the network, serving as a fundamental indicator of network complexity and connectivity.

Number of Edges: The count of street segments between intersections, which helps in understanding the structure and density of the street network.

Average Degree (k_avg): A key metric indicating the average number of edges connected to a node, reflecting the network's density and connectivity.

Total Edge Length: The cumulative length of all edges in the network, providing insights into the network's scale and extent.

Average Edge Length: This measures the average length of individual street segments, offering a sense of scale and granularity of the network.

Average Streets per Node: The average count of streets emanating from a node, an indicator of local connectivity and network complexity.

Intersection Count: A straightforward count of the intersections, crucial for understanding network navigability and complexity.

Total Street Length: The overall length of streets, which speaks to the extent of the network.

Street Segment Count: The total number of street segments, indicating the subdivision and intricacy of the street layout.

Average Street Length: This metric provides an average measure of street lengths, contributing to an understanding of block sizes and network granularity.

Circuity Average: A measure of how much paths deviate from a straight line, with values greater than 1 indicating more circuitous routes.

Self Loop Proportion: The proportion of edges that start and end at the same node, revealing aspects of network design and connectivity.

Gamma Index (γ): A crucial measure of network connectivity, indicating the ratio of actual to maximum possible edges.

Cyclomatic Number: This represents the number of independent cycles within the network, highlighting the potential for multiple routes and resilience.

Maximum Number of Cycles: The theoretical maximum number of independent cycles, indicating the potential complexity of the network.

Redundancy Index: A ratio indicating the network's redundancy, essential for understanding the robustness and resilience of the urban network.

By analyzing these indices, we aim to uncover the underlying structural patterns of urban networks and their potential impact on urban functionality. This chapter sets the stage for a detailed comparison and discussion of these diverse metrics, enabling us to draw meaningful conclusions about the nature and efficiency of urban networks in different U.S. cities.

In the following images, the results are printed in a table layout to be easily readable. Comments on result are also present.

Planned Cities

Structured Network Layout: Planned cities display a lower circuity average (mean: 1.026), aligning with the more structured and grid-like street patterns.

Connectivity and Network Design: The average degree (k_avg) is slightly lower (mean: 5.345), reflecting the deliberate planning in the network's connectivity. The Gamma Index is also lower (mean: 0.892), suggesting less overall network connectivity compared to self-organized cities.

Intersection and Street Layout: There is a high intersection count (mean: 19088.75) and substantial street lengths (mean: 4.23 million meters), which are indicative of the planned and systematic layout of these cities.

Network Redundancy and Complexity: The planned cities have a lower self-loop proportion (mean: 0.0025) and a higher Redundancy Index (mean: 0.0002), which might reflect the planned nature of these networks with fewer ad-hoc connections but a higher level of deliberate redundancy.

Philadelphia, PA

Statistic	Value
Number of Nodes	24831
Number of Edges	61313
k_avg	4.9384
edge_length_total	6668783.44
edge_length_avg	108.77
streets_per_node_avg	3.3266
intersection_count	23786
street_length_total	4490721.44
street_segment_count	41180
street_length_avg	109.05
circuity_avg	1.0292
self_loop_proportion	0.0038
gamma_index	0.8231
cyclomatic_number_directed	36483
max_cycles	308252035
redundancy_index	0.000118

Chicago, IL

Statistic	Value
Number of Nodes	28536
Number of Edges	75892
k_avg	5.3190
edge_length_total	10363526.57
edge_length_avg	136.56
streets_per_node_avg	3.3804
intersection_count	26557
street_length_total	6825212.52
street_segment_count	47951
street_length_avg	142.34
circuity_avg	1.0131
self_loop_proportion	0.0012
gamma_index	0.8866
cyclomatic_number_directed	47357
max_cycles	407108845
redundancy_index	0.000116

Washington, DC

Statistic	Value
Number of Nodes	9906
Number of Edges	26740
k_avg	5.3987
edge_length_total	3207131.61
edge_length_avg	119.94
streets_per_node_avg	3.2936
intersection_count	9373
street_length_total	1920680.68
street_segment_count	16197
street_length_avg	118.58
circuity_avg	1.0343
self_loop_proportion	0.0038
gamma_index	0.9000
cyclomatic_number_directed	16835
max_cycles	49049560
redundancy_index	0.000343

Denver, CO

Statistic	Value
Number of Nodes	16815
Number of Edges	48369
k_avg	5.7531
edge_length_total	6418392.09
edge_length_avg	132.70
streets_per_node_avg	3.3069
intersection_count	15765
street_length_total	3676422.93
street_segment_count	27543
street_length_avg	133.48
circuity_avg	1.0280
self_loop_proportion	0.0012
gamma_index	0.9590
cyclomatic_number_directed	31555
max_cycles	141346891
redundancy_index	0.000223

Self-Organized Cities

Network Complexity and Connectivity: These cities tend to have a higher circuity average (mean: 1.035), indicating more circuitous routes. This aligns with the self-organized nature of these cities where the street layout evolved more organically.

Network Density: The average degree (k_avg) is relatively high (mean: 5.457), suggesting dense connectivity within the urban street networks.

Intersection Density and Road Length: They have a substantial number of intersections (mean: 14149.25) and total street lengths (mean: 3.12 million meters), reflecting the intricate and dense urban fabric typical of self-organized cities.

Network Design and Redundancy: These cities show a higher self-loop proportion (mean: 0.0043) and a moderate Gamma Index (mean: 0.910), indicating a balance between connectivity and redundancy in the network design.

Atlanta, GA

Statistic	Value
Number of Nodes	12647
Number of Edges	33193
k_avg	5.2491
edge_length_total	5105432.07
edge_length_avg	153.81
streets_per_node_avg	2.8872
intersection_count	10868
street_length_total	2787970.50
street_segment_count	18093
street_length_avg	154.09
circuity_avg	1.0690
self_loop_proportion	0.0142
gamma_index	0.8750
cyclomatic_number_directed	20547
max_cycles	79954335
redundancy_index	0.000257

Seattle, WA

Statistic	Value
Number of Nodes	19060
Number of Edges	50283
k_avg	5.2763
edge_length_total	5516054.33
edge_length_avg	109.70
streets_per_node_avg	3.0455
intersection_count	17080
street_length_total	2971820.51
street_segment_count	28992
street_length_avg	102.50
circuity_avg	1.0274
self_loop_proportion	0.0008
gamma_index	0.8795
cyclomatic_number_directed	31224
max_cycles	181613211
redundancy_index	0.000172

Detroit, MI

Statistic	Value
Number of Nodes	20738
Number of Edges	59627
k_avg	5.7505
edge_length_total	8379917.22
edge_length_avg	140.54
streets_per_node_avg	3.3671
intersection_count	19665
street_length_total	4874506.40
street_segment_count	34655
street_length_avg	140.66
circuity_avg	1.0124
self_loop_proportion	0.0009
gamma_index	0.9585
cyclomatic_number_directed	38890
max_cycles	215001216
redundancy_index	0.000181

Pl

San Francisco, CA

Statistic	Value
Number of Nodes	9791
Number of Edges	27187
k_avg	5.5535
edge_length_total	3113690.70
edge_length_avg	114.53
streets_per_node_avg	3.3142
intersection_count	8984
street_length_total	1851088.22
street_segment_count	16218
street_length_avg	114.14
circuity_avg	1.0330
self_loop_proportion	0.0014
gamma_index	0.9258
cyclomatic_number_directed	17397
max_cycles	47917155
redundancy_index	0.000363

Network Layout and Complexity: Self-organized cities tend to have more complex and less predictable street networks, reflected in their higher circuity averages and self-loop proportions. In contrast, planned cities exhibit more structured, grid-like layouts with lower circuity and self-loop proportions.

Connectivity and Density: Self-organized cities generally show denser connectivity, as indicated by their higher average degree. Planned cities, while still well-connected, follow a more uniform connectivity pattern.

Redundancy and Robustness: Planned cities tend to have a slightly higher redundancy index, which might be a result of deliberate urban design to accommodate future growth and ensure network resilience.

In conclusion, the analysis of Inter-Network Indices reveals distinct patterns in the urban street networks of self-organized and planned cities, reflecting their unique development histories and urban planning philosophies. Self-organized cities exhibit characteristics of organic growth with complex and dense networks, while planned cities showcase structured, uniform, and resilient network designs.

3.2.2 MCA part 2 – Intra - Network Indices

This section builds upon the theoretical framework established in the previous chapters, focusing on Intra-network indices which serve as critical tools for analyzing and understanding the dynamics of urban street networks. These indices enable us to evaluate the significance and interconnectedness of individual nodes within the network, thereby providing a granular perspective on the structure and functionality of urban spaces. The upcoming analysis will be anchored around three specific indices: Degree Centrality, Betweenness Centrality, Closeness Centrality.

This index provides insights into the potential vibrancy and social interaction facilitated by a node's location in the urban fabric.

The application of these intra-network indices to urban street networks will involve a detailed analysis of selected cities, encompassing both planned and self-organized urban forms. This analysis aims to uncover the underlying patterns of connectivity, accessibility, and efficiency within these urban landscapes. By comparing the centrality measures across different types of cities, this study seeks to draw meaningful conclusions about the impact of urban planning and development strategies on the functionality and resilience of urban networks.

In addition to theoretical exploration, this section will also present case studies to illustrate the practical implications of these indices. These case studies will include analyses of car accidents per node to explain how MCA can drive decision maker to choose smart solution for urban development. The findings from this analysis will offer valuable insights for urban planners, policymakers, and researchers in understanding and shaping the future of urban environments.

In summary, this chapter delves into the heart of urban network analysis, providing both a theoretical foundation and practical tools for understanding the complex dynamics of city street networks. Through the lens of intra-network indices, we gain a deeper appreciation of the interplay between urban form and functionality, setting the stage for more informed and effective urban planning and policy-making.

Degree Centrality

Objective: Assess the connectivity of each node within the network.

Method: Compare the degree centrality values across cities to identify which have more interconnected nodes. Higher values indicate nodes with more connections, suggesting major hubs or intersections in the transportation network.

Application: Identify potential areas for traffic congestion or key points for infrastructure investment. High degree centrality nodes in urban networks are critical for network efficiency and can be focal points for urban development and planning.

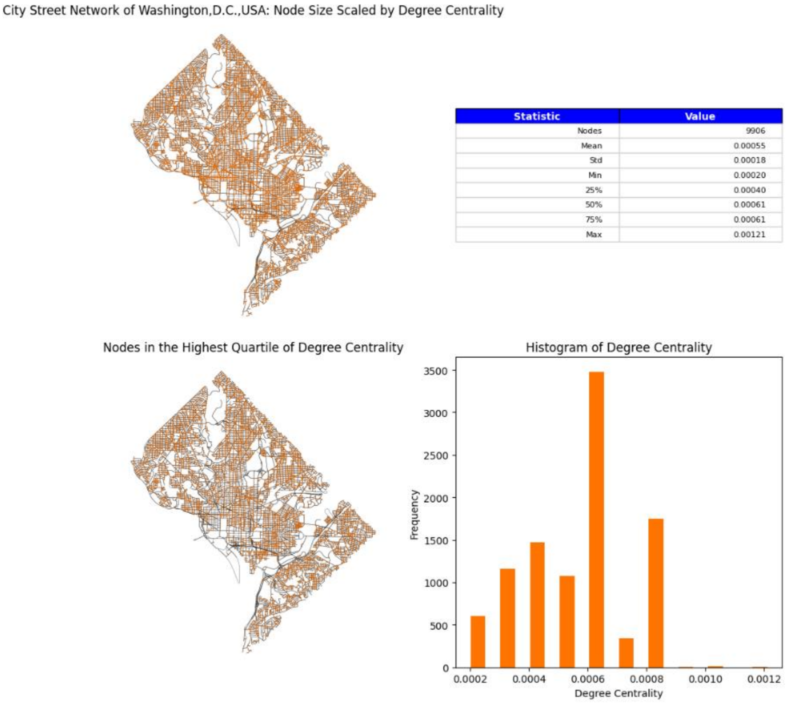
Planned Cities:

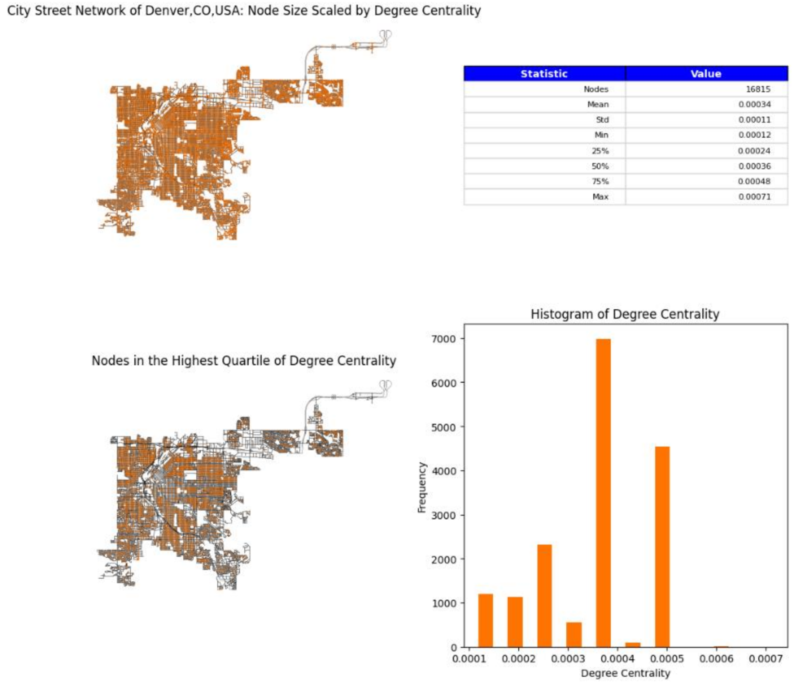
Denver, CO, USA: The Degree Centrality Mean (0.000342) indicates well-distributed connectivity, characteristic of planned cities.

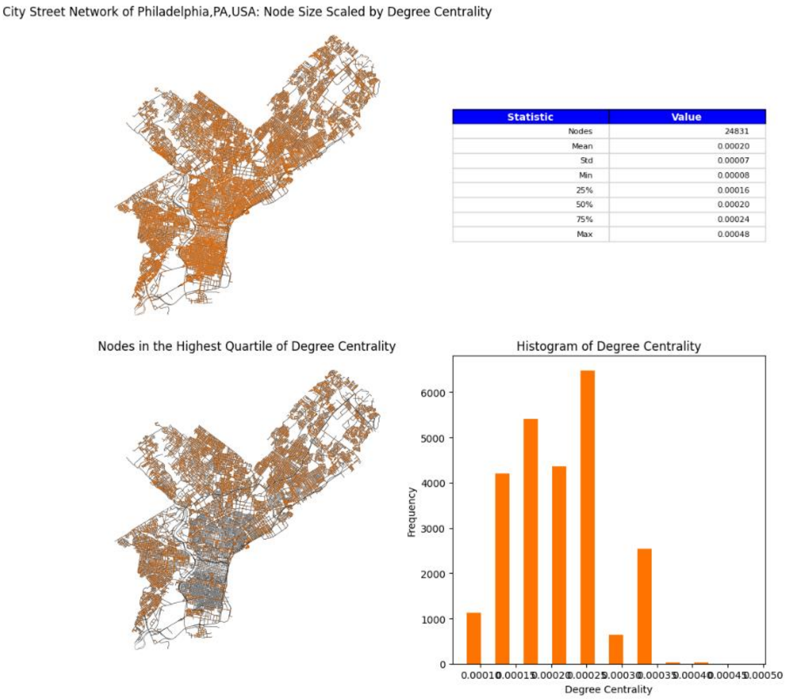
Washington D.C., USA: High Degree Centrality Mean (0.000545), reflecting the city’s planned structure with numerous intersections and hubs.

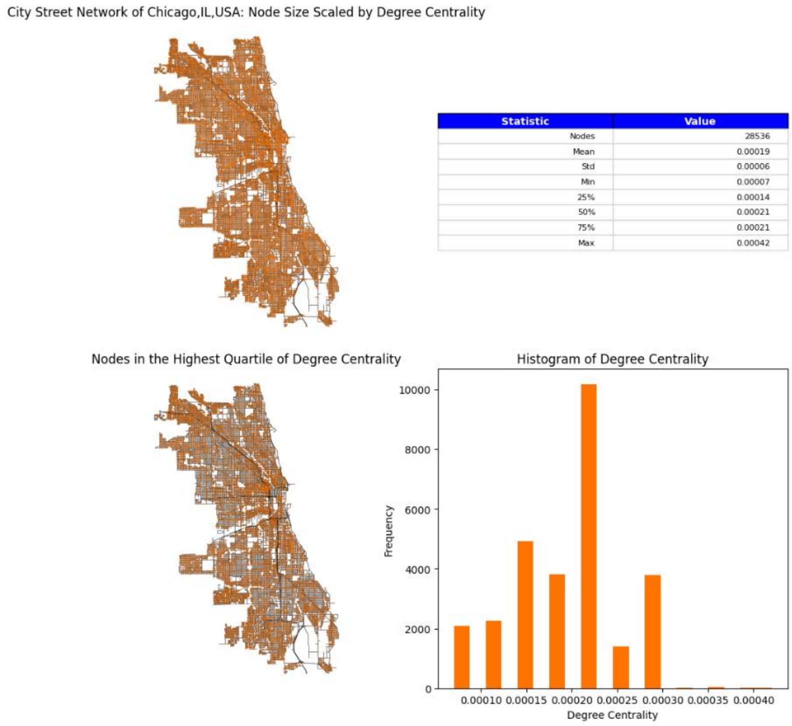
Chicago, IL, USA: Shows the lowest Degree Centrality Mean (0.000186), suggesting a less dense network, possibly due to its large geographical size.

Philadelphia, PA, USA: Similar to Chicago, it has a lower Degree Centrality Mean (0.000199), indicating a less dense urban network.









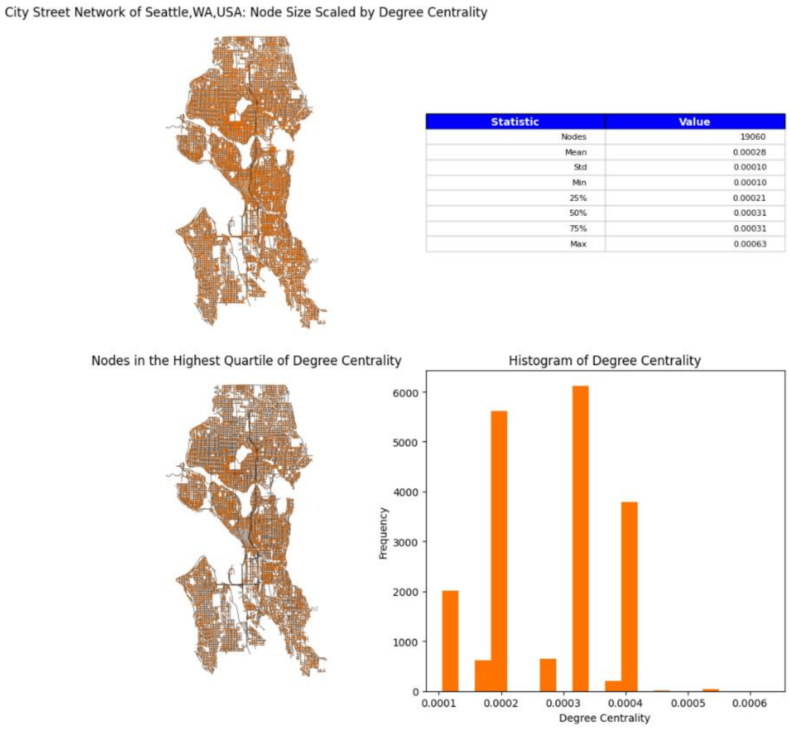
Self-Organized Cities:

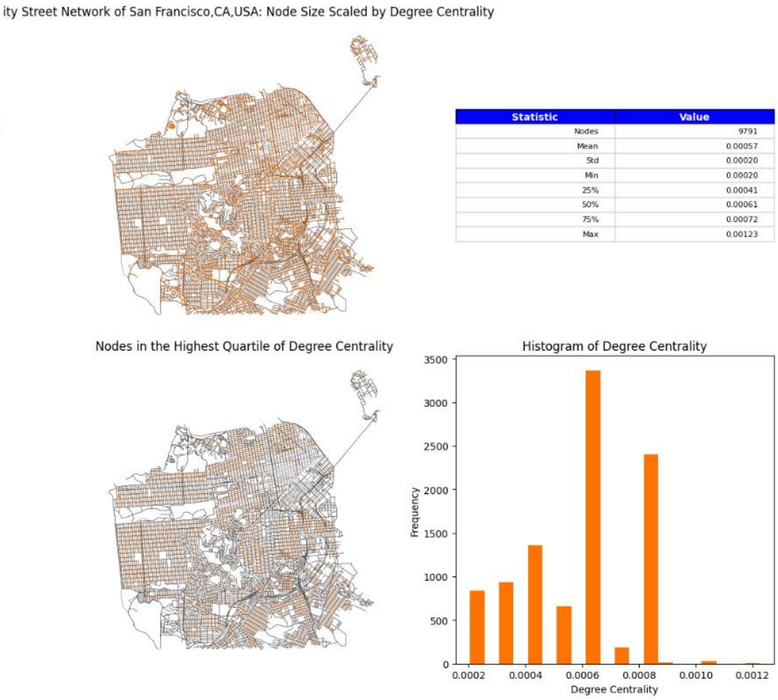
San Francisco, CA, USA: Exhibits a high Degree Centrality Mean (0.000567), suggesting dense connectivity and numerous intersections. The Maximum Degree Centrality (0.001226) is also quite high, indicating the presence of major hubs within the city.

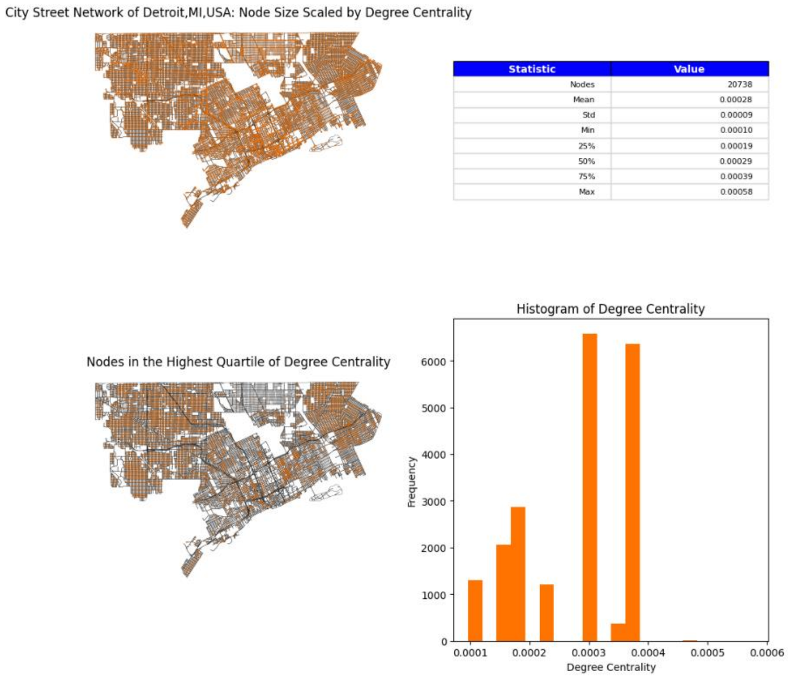
Detroit, MI, USA: Shows a lower Degree Centrality Mean (0.000277) compared to San Francisco, indicating a less dense network but still significant connectivity.

Atlanta, GA, USA: Has a moderate Degree Centrality Mean (0.000415), reflecting a balance between dense urban areas and less connected regions.

Seattle, WA, USA: Degree Centrality Mean (0.000277) is similar to Detroit, suggesting a moderate level of connectivity across the city.







Network Density: Self-organized cities are characterized by denser and more complex networks, whereas planned cities exhibit more uniform and structured networks.

Urban Vibrancy vs Congestion: The denser networks in self-organized cities can lead to vibrant urban areas but may also be prone to congestion. In contrast, the uniform network distribution in planned cities can mitigate congestion but might not have the same level of vibrancy.

Strategic Urban Planning: Planned cities demonstrate the effectiveness of strategic urban planning in creating efficient, well-connected networks, while self-organized cities reveal the adaptability of urban areas to various influences over time.

Betweenness Centrality

Objective: Evaluate the importance of nodes in terms of their role in facilitating traffic flow.

Method: Analyze the distribution of betweenness centrality scores across the network. Nodes with high scores are those that frequently appear on the shortest paths between other nodes, indicating their strategic importance in the network.

Application: This measure is crucial for understanding traffic dynamics and planning for congestion management. Nodes with high betweenness centrality are often critical for maintaining network connectivity and are potential points for traffic optimization and control.

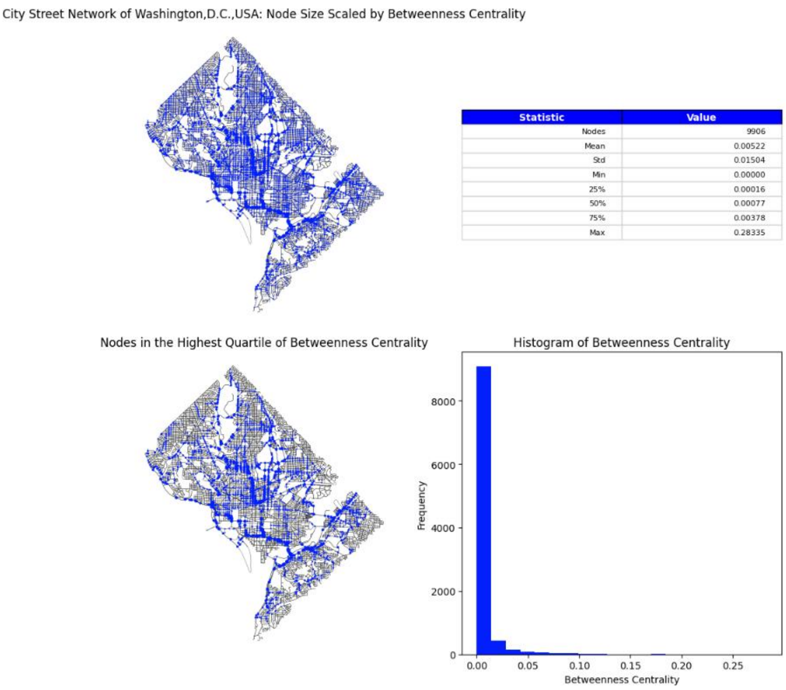
Planned Cities:

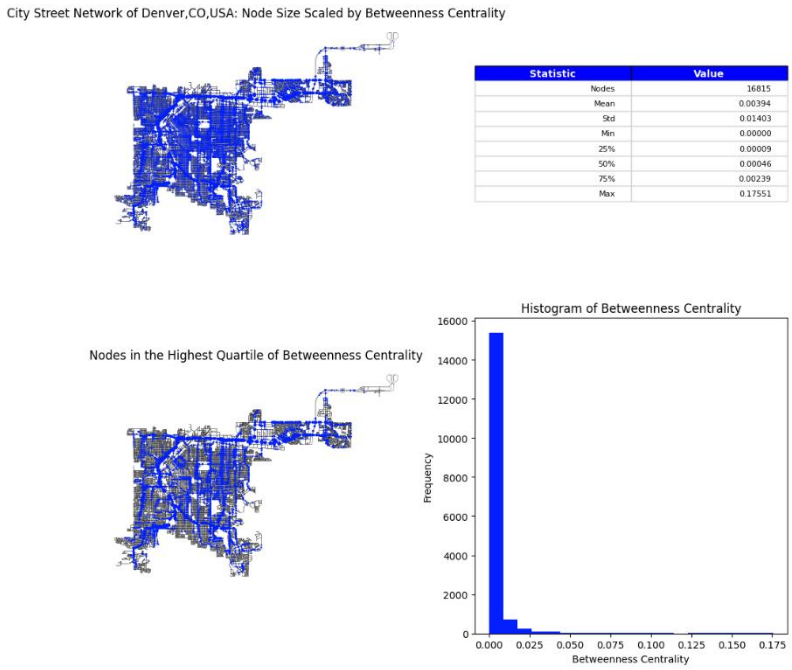
Denver, CO, USA: The city has a moderate Betweenness Centrality Mean (0.003943), reflecting its planned nature with a well-distributed network of important traffic nodes.

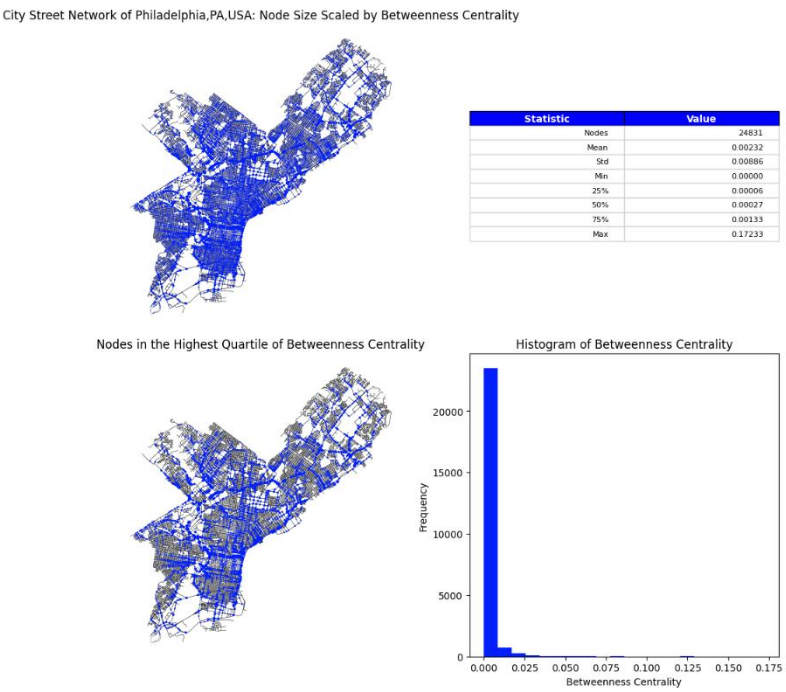
Washington D.C., USA: Exhibits the highest Betweenness Centrality Mean (0.005251) among the cities analyzed, indicative of a highly strategic network where certain streets play a crucial role in maintaining connectivity across the city.

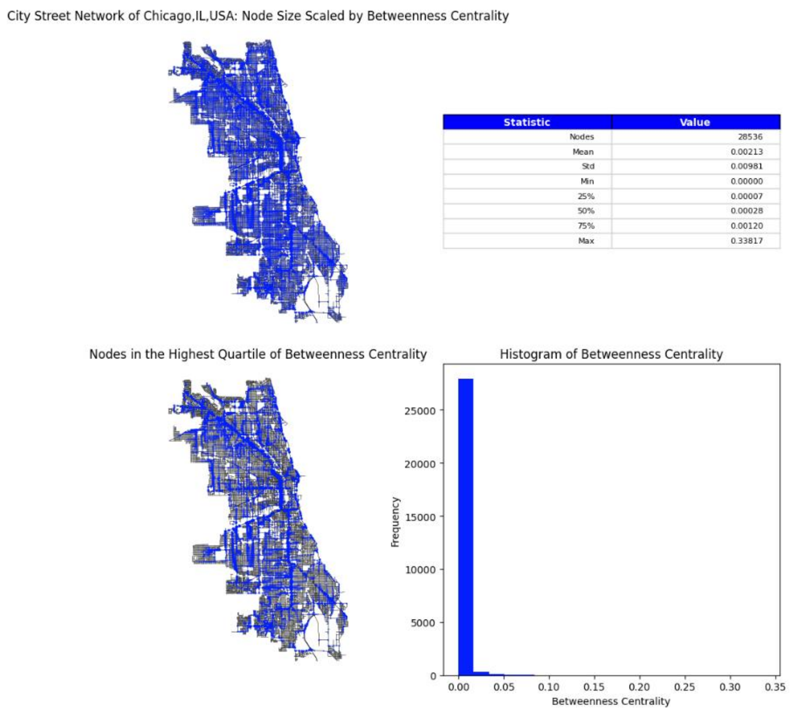
Chicago, IL, USA: Shows a relatively lower Betweenness Centrality Mean (0.002126), suggesting a less complex network with fewer critical traffic nodes, possibly due to its expansive geographic layout.

Philadelphia, PA, USA: Has a Betweenness Centrality Mean (0.002316) similar to Chicago, reflecting a network with a balanced distribution of traffic flow across various nodes.









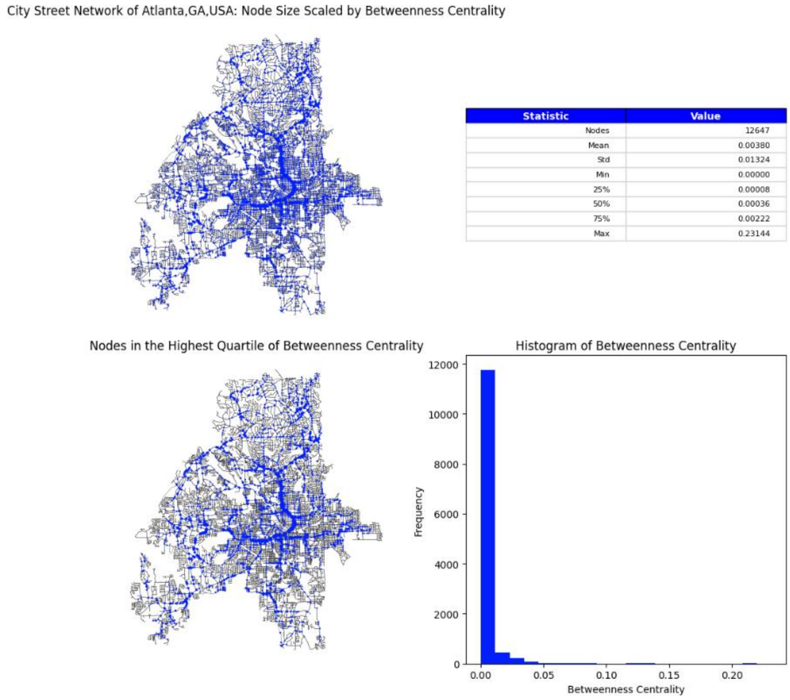
Self-Organized Cities:

San Francisco, CA, USA: The city demonstrates a high Betweenness Centrality Mean (0.004394), indicating numerous nodes are critical in facilitating traffic flow across the network. This suggests that many streets in San Francisco play a vital role in connecting different parts of the city.

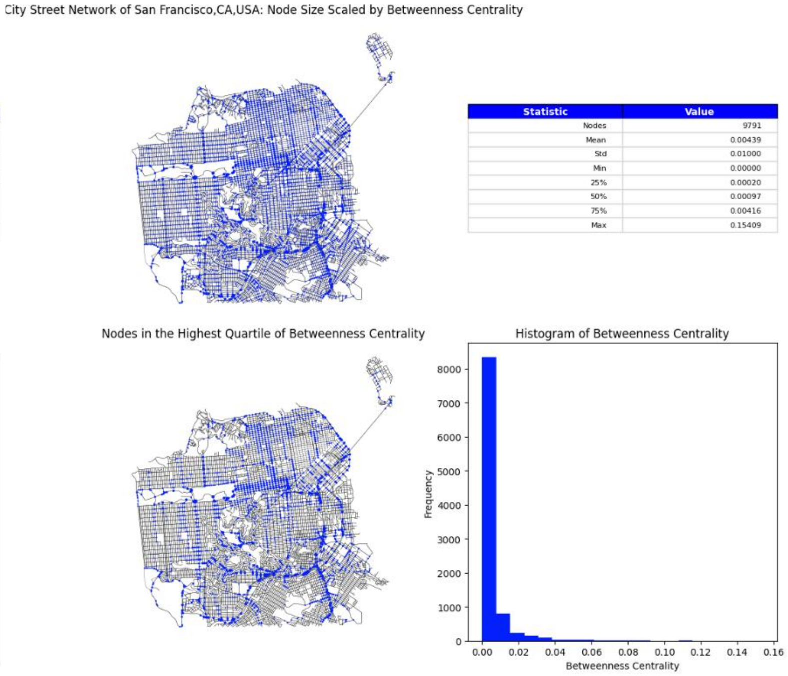
Detroit, MI, USA: Has a lower Betweenness Centrality Mean (0.002488) compared to San Francisco. This reflects a network with fewer key nodes dominating traffic flow, which is typical for cities with a less complex street layout.

Atlanta, GA, USA: Shows a moderately high Betweenness Centrality Mean (0.003804), indicating the presence of several important traffic nodes. This suggests a balance between major thoroughfares and smaller streets in the city's traffic network.

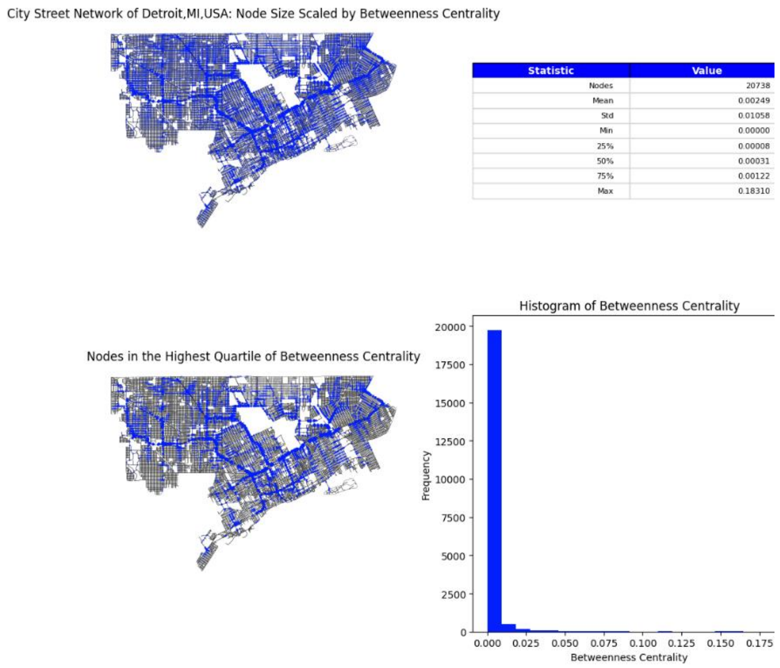
Seattle, WA, USA: Exhibits a Betweenness Centrality Mean (0.002904) that is higher than Detroit but lower than San Francisco and Atlanta. This implies that while Seattle has significant nodes for traffic flow, they are not as dominant as in San Francisco.San Francisco, CA, USA: The city demonstrates a high Betweenness Centrality Mean (0.004394), indicating numerous nodes are critical in facilitating traffic flow across the network. This suggests that many streets in San Francisco play a vital role in connecting different parts of the city.



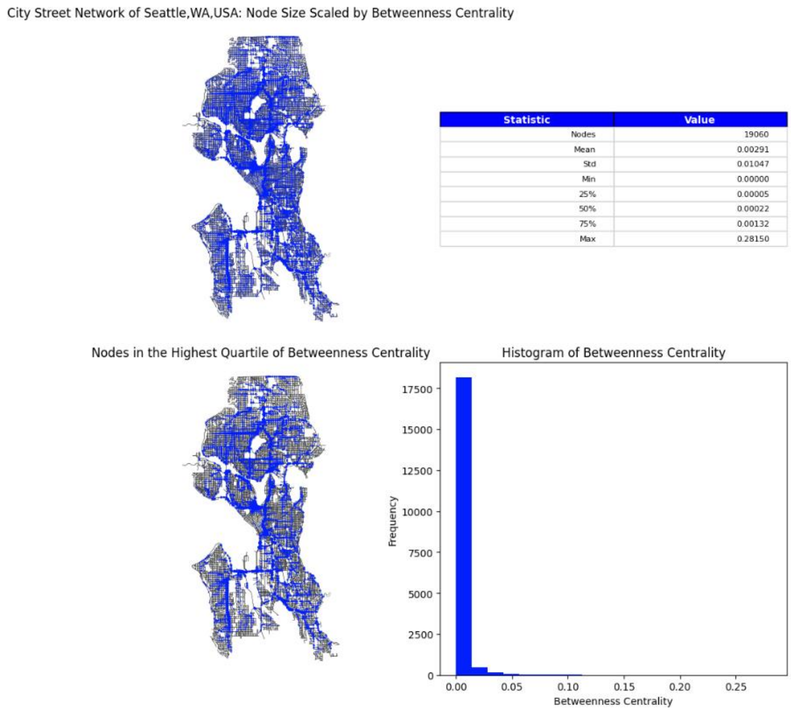
Statistic	Value
Nodes	12647
Mean	0.00380
Std	0.01324
Min	0.00000
25%	0.00008
50%	0.00036
75%	0.00222
Max	0.23144



Statistic	Value
Nodes	9791
Mean	0.00439
Std	0.01000
Min	0.00000
25%	0.00020
50%	0.00097
75%	0.00416
Max	0.15409



Statistic	Value
Nodes	20738
Mean	0.00249
Std	0.01058
Min	0.00000
25%	0.00008
50%	0.00031
75%	0.00122
Max	0.18310



Statistic	Value
Nodes	19060
Mean	0.00291
Std	0.01047
Min	0.00000
25%	0.00005
50%	0.00022
75%	0.00132
Max	0.28150

Comparative Analysis: Self-Organized vs Planned Cities

Self-Organized Cities: Generally exhibit higher Betweenness Centrality Means, indicative of complex networks with numerous critical nodes for traffic flow. This complexity is a result of organic city growth and development.

Planned Cities: Tend to show a more balanced distribution of Betweenness Centrality, reflecting their structured and strategic urban planning. However, cities like Washington D.C. demonstrate the presence of highly significant nodes in the network.

Closeness Centrality

Objective: Determine the efficiency of the network in terms of the proximity of nodes to all others.

Method: Calculate and compare the closeness centrality for each node across different cities. Higher closeness centrality indicates that a node is closer to all other nodes in the network, signifying efficient information or traffic flow.

Application: This measure is useful for identifying areas in the network that are strategically placed for rapid access to various parts of the city. It can inform decisions on emergency services placement, urban accessibility improvements, and public transport network design.

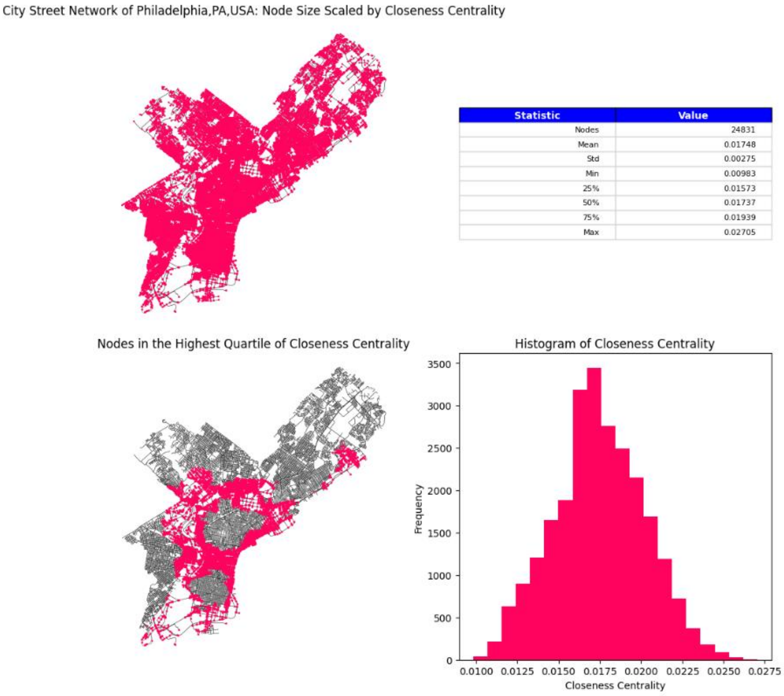
Planned Cities:

Denver, CO, USA: With a Closeness Centrality Mean (0.015384), Denver shows efficient connectivity, which is a characteristic of its planned nature. However, it is slightly lower compared to some self-organized cities, indicating a potential trade-off between uniformity and maximum accessibility.

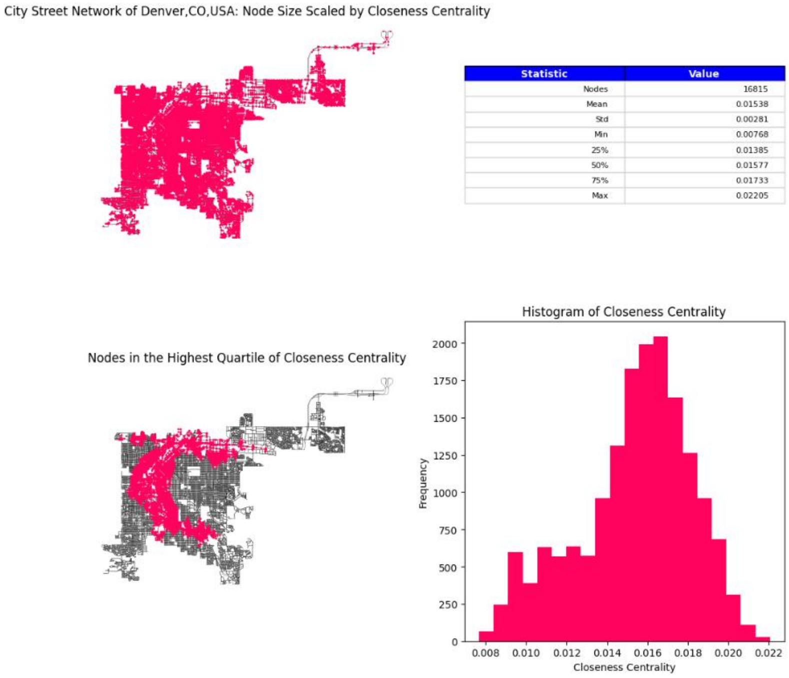
Washington D.C., USA: Exhibits a relatively high Closeness Centrality Mean (0.019715), reflecting the city's effective planning in ensuring accessibility and connectivity across the network.

Chicago, IL, USA: The city has a moderate Closeness Centrality Mean (0.016619), suggesting good overall connectivity, though it may not be as efficient as some of the other cities in terms of network accessibility.

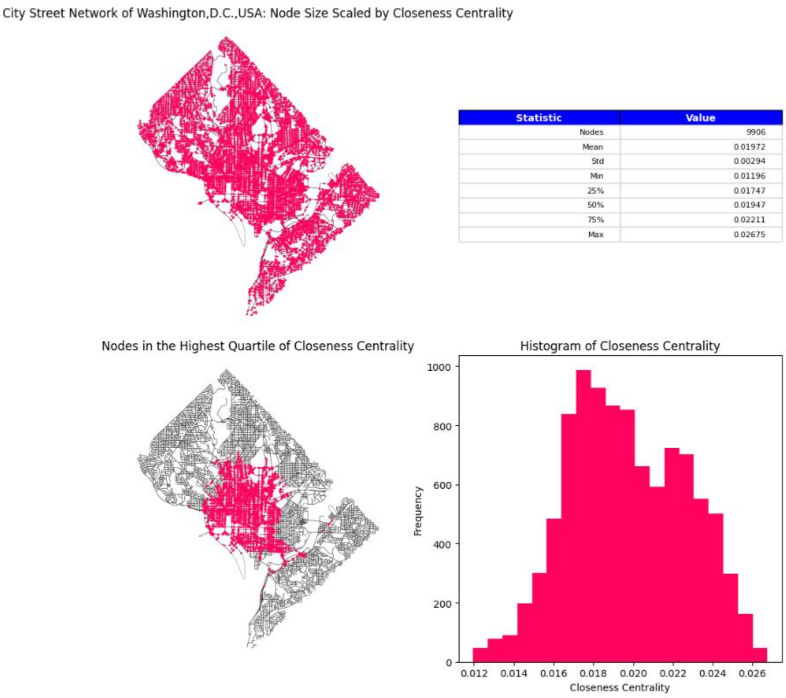
Philadelphia, PA, USA: Shows a Closeness Centrality Mean (0.017481) that is comparable to Chicago, indicating a well-planned network that facilitates movement across the city.



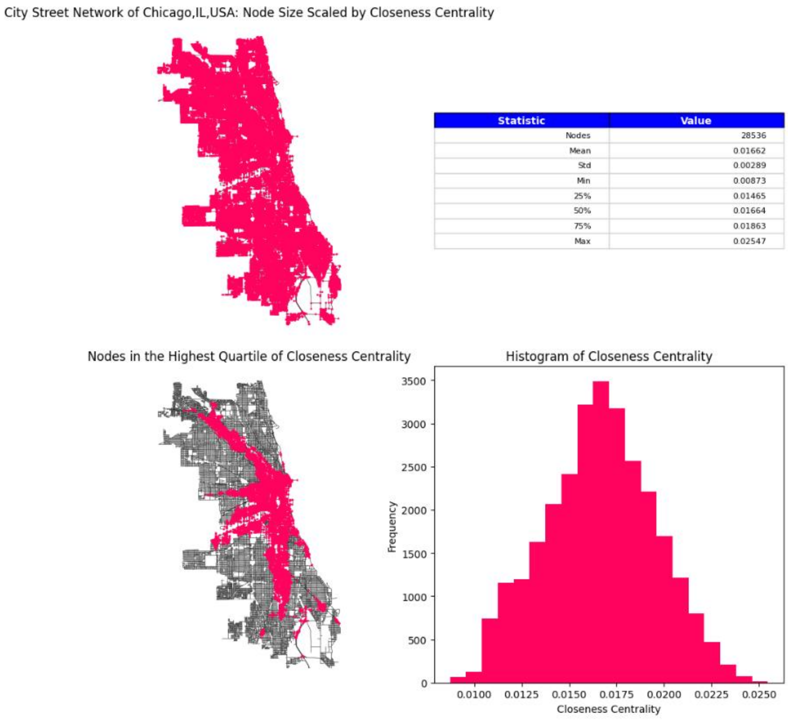
Statistic	Value
Nodes	24831
Mean	0.01748
Std	0.00275
Min	0.00983
25%	0.01573
50%	0.01737
75%	0.01939
Max	0.02705



Statistic	Value
Nodes	16815
Mean	0.01538
Std	0.00281
Min	0.00768
25%	0.01385
50%	0.01577
75%	0.01733
Max	0.02205



Statistic	Value
Nodes	9906
Mean	0.01972
Std	0.00294
Min	0.01196
25%	0.01747
50%	0.01947
75%	0.02211
Max	0.02675



Statistic	Value
Nodes	28536
Mean	0.01662
Std	0.00289
Min	0.00873
25%	0.01465
50%	0.01664
75%	0.01863
Max	0.02547

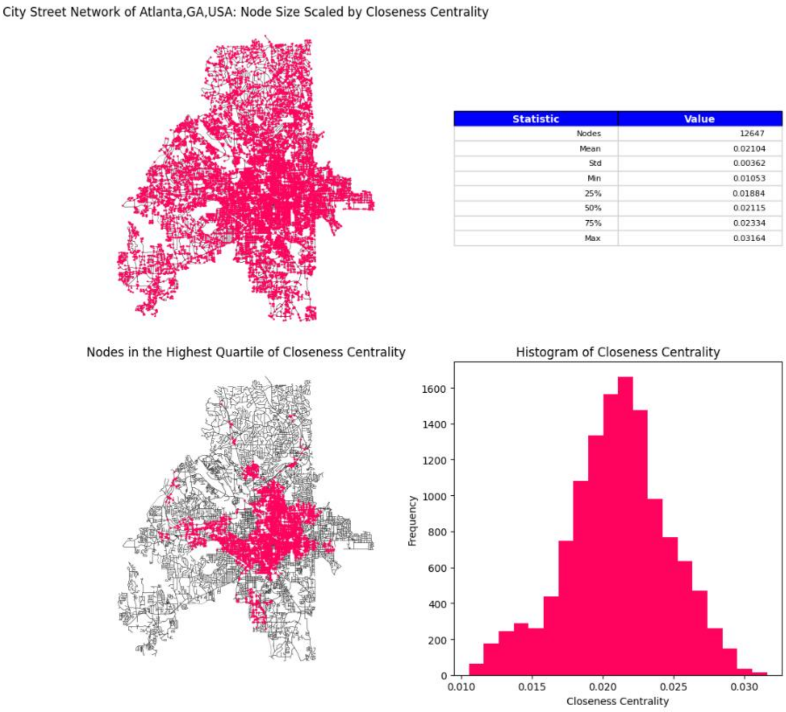
Self-Organized Cities:

San Francisco, CA, USA: With a high Closeness Centrality Mean (0.023288), San Francisco demonstrates efficient network connectivity, indicating that most parts of the city are easily accessible from any given node. The Maximum Closeness Centrality (0.031927) further suggests that certain areas are exceptionally well-connected.

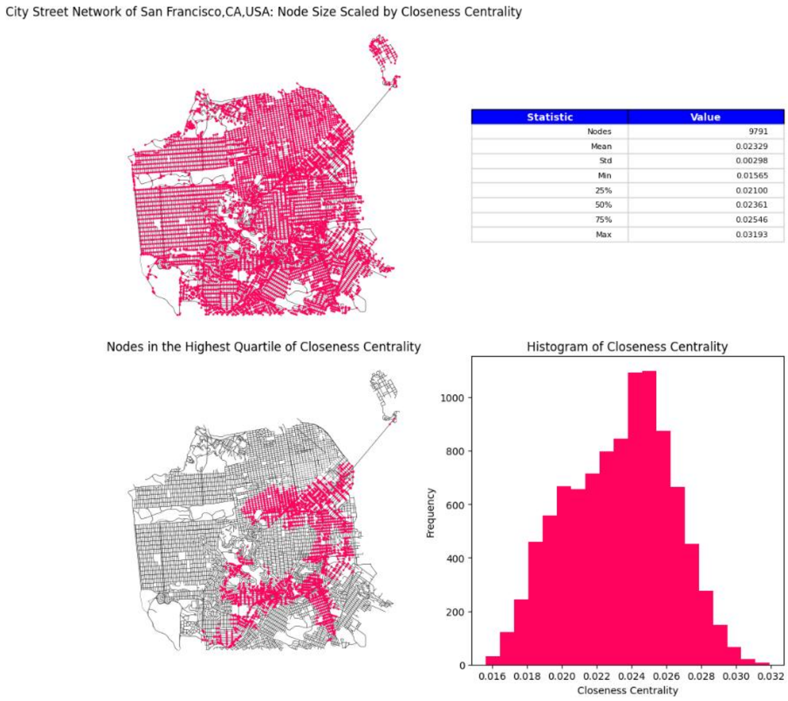
Detroit, MI, USA: Has a moderate Closeness Centrality Mean (0.019221), reflecting reasonable accessibility throughout the city. However, it is not as high as San Francisco, indicating some variation in connectivity across different parts of the city.

Atlanta, GA, USA: Exhibits a fairly high Closeness Centrality Mean (0.021039), suggesting efficient connectivity and ease of movement across the city. This indicates a well-interconnected urban fabric, despite the city's organic growth.

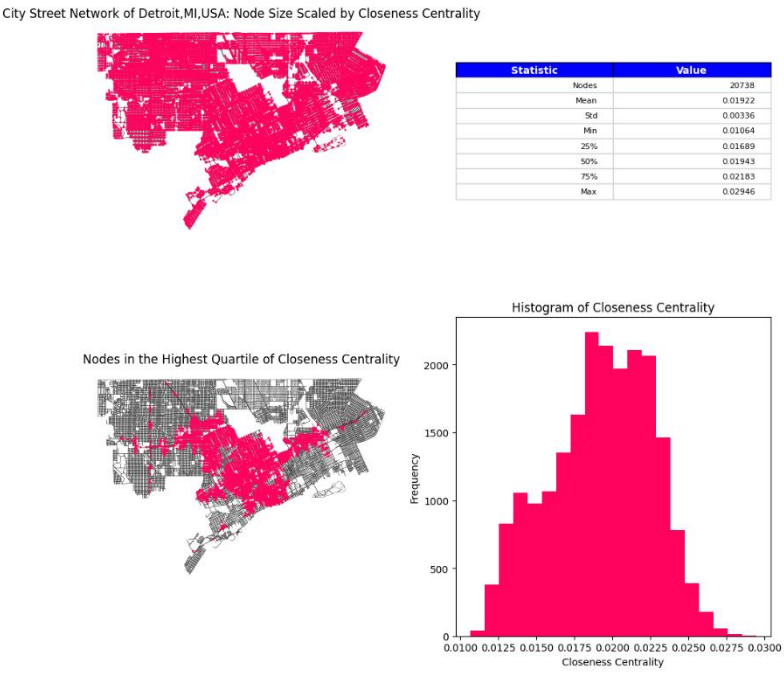
Seattle, WA, USA: Shows a Closeness Centrality Mean (0.018649) that is slightly lower than Atlanta, indicating good but slightly less uniform accessibility across the city compared to Atlanta and San Francisco.



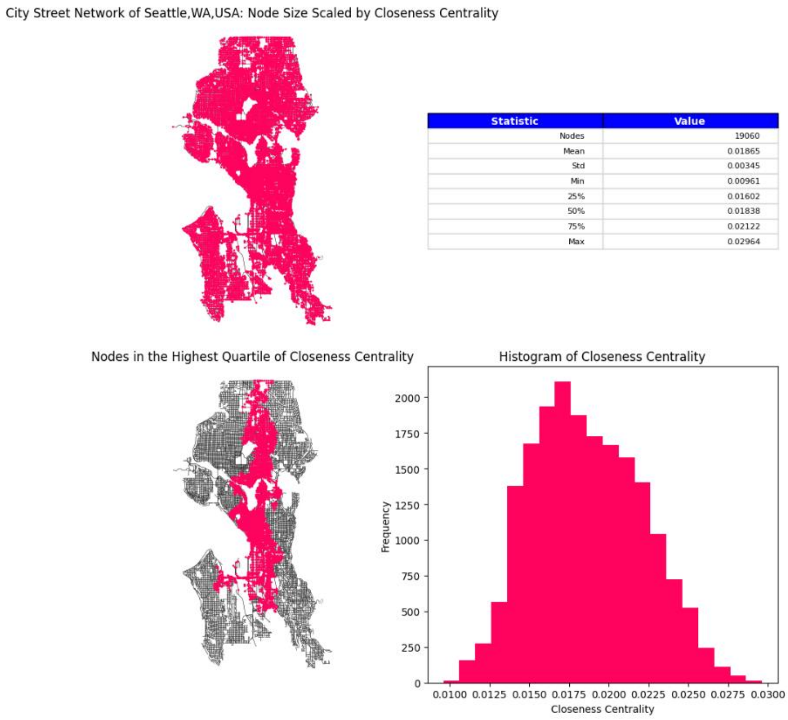
Statistic	Value
Nodes	12647
Mean	0.02104
Std	0.00362
Min	0.01053
25%	0.01884
50%	0.02115
75%	0.02334
Max	0.03164



Statistic	Value
Nodes	9791
Mean	0.02329
Std	0.00298
Min	0.01565
25%	0.02100
50%	0.02361
75%	0.02546
Max	0.03193



Statistic	Value
Nodes	20738
Mean	0.01922
Std	0.00336
Min	0.01064
25%	0.01689
50%	0.01943
75%	0.02183
Max	0.02946



Statistic	Value
Nodes	19060
Mean	0.01865
Std	0.00345
Min	0.00961
25%	0.01602
50%	0.01838
75%	0.02122
Max	0.02964

Comparative Analysis: Self-Organized vs Planned Cities

Self-Organized Cities: Tend to exhibit higher Closeness Centrality Means, indicating efficient and well-connected networks. This could be due to the organic adaptation of the street network to the city's evolving needs, leading to a complex but highly accessible urban fabric.

Planned Cities: While also showing good connectivity, the Closeness Centrality in planned cities is sometimes slightly lower than in self-organized cities. This could be a result of the structured and uniform design, which, while efficient, may not always achieve the highest levels of accessibility seen in more organically developed areas.

Conclusion

Planned Cities

Degree Centrality: Planned cities show varying degrees of centrality. Some cities like Washington D.C. exhibit high centrality, suggesting a well-connected network. This reflects the deliberate planning in constructing efficient and accessible urban networks.

Betweenness Centrality: The betweenness centrality is generally high, indicating the presence of key nodes critical for maintaining network connectivity. This is a result of intentional planning to create major transportation routes and hubs.

Closeness Centrality: The closeness centrality values are similar to those in self-organized cities, implying that despite their planned nature, these cities maintain efficient connectivity within their networks.

Self-Organized Cities

Network Connectivity (Degree Centrality): Both self-organized and planned cities demonstrate a significant degree of network connectivity. However, the denser and more intricate patterns are often more pronounced in self-organized cities.

Critical Nodes for Traffic Flow (Betweenness Centrality): Planned cities often have higher betweenness centrality, reflecting the presence of strategically important nodes in their transportation networks. Self-organized cities, while having important nodes, may not exhibit as clear a pattern due to their organic development.

Accessibility and Efficiency (Closeness Centrality): Both types of cities offer efficient connectivity, with nodes generally being accessible to each other. This indicates that irrespective of the development style, urban networks tend to evolve towards optimizing connectivity.Accessibility and Efficiency (Closeness Centrality): Both types of cities offer efficient connectivity, with nodes generally being accessible to each other. This indicates that irrespective of the development style, urban networks tend to evolve towards optimizing connectivity.

Self-Organized vs Planned Cities

Network Connectivity (Degree Centrality): Both self-organized and planned cities demonstrate a significant degree of network connectivity. However, the denser and more intricate patterns are often more pronounced in self-organized cities.

Critical Nodes for Traffic Flow (Betweenness Centrality): Planned cities often have higher betweenness centrality, reflecting the presence of strategically important nodes in their transportation networks. Self-organized cities, while having important nodes, may not exhibit as clear a pattern due to their organic development.

Accessibility and Efficiency (Closeness Centrality): Both types of cities offer efficient connectivity, with nodes generally being accessible to each other. This indicates that irrespective of the development style, urban networks tend to evolve towards optimizing connectivity.

In summary, the intra-network analysis underscores the complexities and strengths inherent in both self-organized and planned urban environments. While self-organized cities display dense and complex network structures, planned cities show clear patterns of connectivity with strategically placed key nodes. Both types of cities, however, manage to maintain efficient overall network connectivity, highlighting their adaptability and resilience in urban planning and development.

3.3 The Real-World Implications of Centrality Measures on Urban Safety

The importance of network analysis in urban environments has been amplified by progress in data science and computer methods, particularly in the field of traffic safety and mobility. The utilization of extreme value theory (EVT) models and Bayesian hierarchical approaches for conducting real-time safety evaluations in metropolitan road networks represents a notable progress in improving road safety (Zheng et al., 2014) This methodological approach leverages data from road user trajectories, namely those gathered from drones, to educate algorithms that create safer routes. It represents a pioneering application of network analysis to enhance urban safety (Ghoul et al., 2023).

The fluctuating nature of accident risk, which varies in both space and time, emphasizes the intrinsic complexity of urban networks. Traditional methods of evaluating road safety typically rely on retrospective crash data, which are often reactive and insufficiently timely to effectively address the immediate and evolving factors that contribute to the risks of accidents. Implementing EVT models to analyze traffic conflict data obtained from road user trajectories introduces a proactive and predictive methodology. This approach allows for the calculation of safety measures in real-time, which is crucial for the ever-changing conditions of urban roads (Kamel et al., 2022; Zheng & Sayed, 2020).

Researchers in Athens, Greece conducted a study where they used a Bayesian hierarchical EVT model to analyze data collected by drones (Ghoul, 2023). This allowed them to create a routing algorithm that can determine the safest route through a network of intersections, both with and without traffic signals, in real-time. This methodology is distinguished for its ability to measure the likelihood of accidents on different routes and to include the length of time spent in hazardous conditions in the safety evaluation. Urban areas necessitate the crucial consideration of multiple factors, since road users encounter a multitude of swiftly altering variables.

The finding that the routes with the highest level of safety were, on average, 22% safer, but taking 11% longer to travel, highlights the fundamental compromise between safety and efficiency in urban transportation (Zheng & Sayed, 2020). The equilibrium of this matter is a pivotal concern for urban planners and transportation engineers. The creation of a multi-objective routing system that takes into account user preferences for safety in relation to journey duration signifies a notable advancement. This indicates a growing recognition of the need for customized and flexible transportation options in urban environments.

Furthermore, the utilization of real-time data in determining routing choices indicates a crucial transition towards more advanced intelligent transportation systems (ITS). These systems have the objective of enhancing safety, efficiency, and sustainability in urban transportation networks. The ability to adaptively modify routing based on real-time evaluations of crash risk has the potential to transform urban mobility, enhancing safety, responsiveness, and user-centricity.

Overall, the utilization of network analysis, including EVT models and Bayesian methodologies, for urban road safety signifies a significant progress in our ability to understand and reduce crash risks in real-time. This research makes significant contributions to the academic disciplines of data science and urban planning. Additionally, it presents practical strategies that can be used in cities worldwide to improve road safety. Through the utilization of advanced analytics and up-to-the-minute data, urban planners and transportation authorities may create transportation networks that are safer, more efficient, and more responsive, thereby effectively addressing the requirements of their communities.

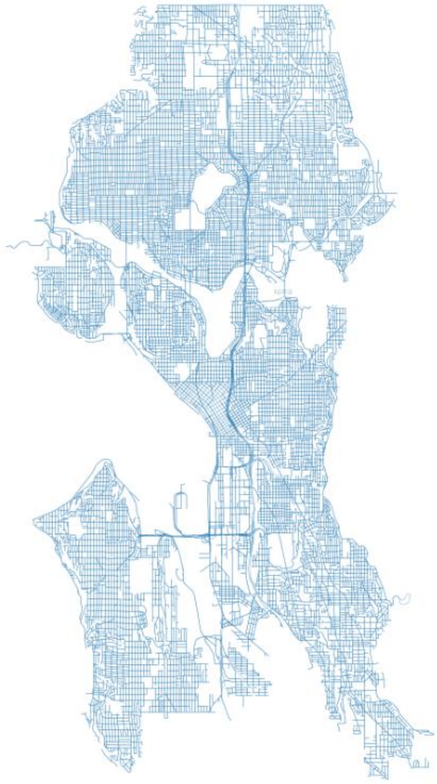
It is also important to examine the significance of centrality measures, specifically betweenness centrality, in identifying heavily trafficked roads and their relationship with the distribution of accidents in urban regions. Betweenness centrality, a fundamental term in network analysis, is a potent technique for identifying roadways that have significant influence on urban traffic flow. This metric calculates the frequency at which a certain node (such as an intersection or road segment) serves as a bridge on the shortest route connecting two other nodes. Roads with high betweenness centrality ratings are typically used as important connectors or main routes in the urban traffic network. This often corresponds to higher traffic volumes and, potentially, increased risk of accidents.

Analyze the distribution of major roads based on betweenness centrality enables urban planners and traffic safety researchers to identify specific locations where safety initiatives would have the greatest impact. For instance, roads that have a high betweenness centrality play a crucial role in facilitating the movement of traffic within the metropolitan road network. Therefore, these roads may be given priority for safety enhancements in order to reduce the risk of accidents. This technique is in line with the proactive and predictive methodology emphasized by Kamel et al. (2022) and Zheng & Sayed (2020), providing a detailed comprehension of urban dynamics that can guide the creation of safer routing algorithms and intelligent transportation systems (ITS).

3.3.1 Using MCA to analyze car crash.

Utilizing Multiple Correspondence Analysis (MCA) to examine car crash data within urban environments offers a nuanced understanding of how road network characteristics influence accident distribution. This methodological approach was applied to Seattle, a city emblematic of the self-organized urban forms commonly found in Europe, where the historical layering of infrastructure and development patterns often precludes systematic urban planning.

The methodology commenced with the construction of the city's road network graph through OpenStreetMap data. This foundational step enabled a detailed analysis of the road network's structural properties, following the analytical framework provided by Porta, S., Crucitti, P., & Latora, V. (2006) in their seminal paper on the network analysis of urban streets using a primal approach. This method prioritizes the direct analysis of the street network, treating intersections as nodes and streets as edges, to understand the structural and functional complexities of urban mobility networks.



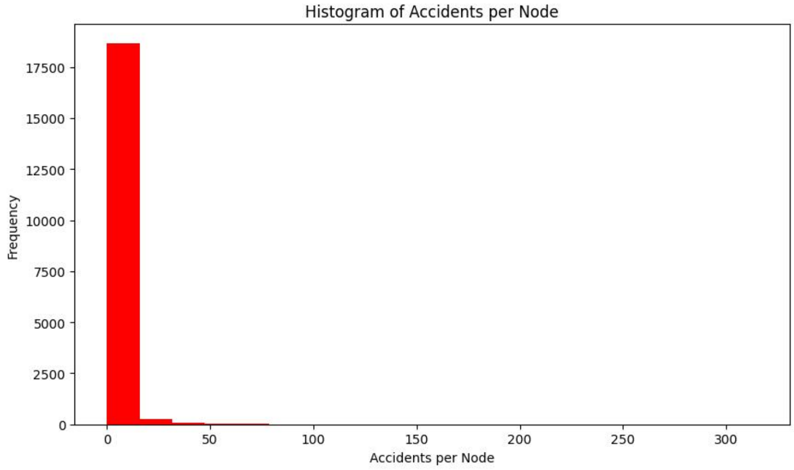
Subsequent to the construction of the road network graph, accident data was integrated from the US-Accidents dataset, a comprehensive collection of traffic accidents spanning 49 states in the United States, compiled since February 2016. This dataset, acknowledged for its breadth and depth, provided a robust foundation for analyzing accident occurrences within the network. The accident locations were mapped onto the city graph by assigning each accident to the nearest node based on their x and y coordinates, ensuring a precise overlay of accident data onto the urban street network.

City Street Network: Node Size Scaled by Normalized Accident Counts

Statistic	Value
Count	19058.00000
Mean	1.60667
Std	7.61975
Min	0.00000
25%	0.00000
50%	0.00000
75%	1.00000
Max	315.00000

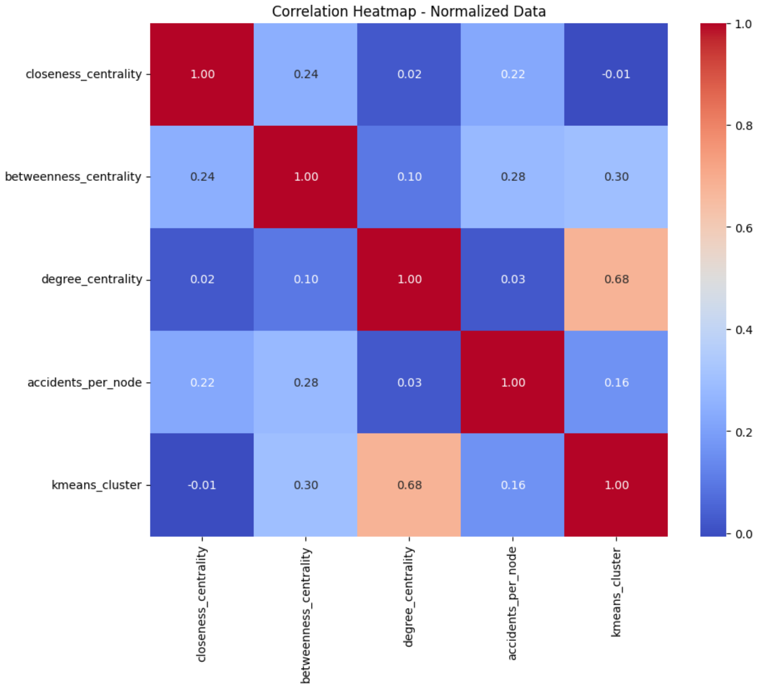
Nodes in the Highest Quartile of Accidents

Histogram of Accidents per Node



The final phase of the analysis involved creating a correlation matrix and distribution plot to examine the relationship between centrality measures, specifically betweenness centrality, and car accident frequencies. Betweenness centrality measures the extent to which a node lies on paths between other nodes, serving as an indicator of a road segment's importance to the flow of traffic throughout the network. The analysis revealed a weak correlation between betweenness centrality and car accidents, with an average correlation coefficient of 0.30.

This weak correlation suggests that while roads with high betweenness centrality are integral to the urban traffic flow, contributing to higher exposure and potential risk, they are not the sole determinants of accident locations. Other factors, such as road design, traffic volume, and driver behavior, likely play significant roles in influencing accident occurrences. This insight challenges simplistic assumptions about the direct relationship between traffic flow centrality and accident risk, highlighting the complexity of urban traffic systems and the multifaceted nature of road safety. Such findings underscore the importance of incorporating diverse data sources and analytical techniques in urban safety studies to capture the intricate dynamics of city road networks and their impact on traffic accidents.



Conclusion

The exploration of urban structures through the lens of network science, centrality measures, and indices has unearthed profound insights into the dynamics of urban life and its implications for traffic safety and overall city livability. This thesis has demonstrated the invaluable role of network analysis in dissecting the complexities of urban environments, offering a pathway to not only understand but also optimize the lifestyle of those dwelling within these intricate systems.

The utilization of Multiple Correspondence Analysis (MCA) to scrutinize car crash data within the context of Seattle's self-organized urban form has highlighted the nuanced relationship between network characteristics and accident occurrences. This approach, grounded in the pioneering work of Porta, Crucitti, and Latora (2006), has provided a methodical framework for evaluating the interplay between urban network structures and car-related accidents. The findings reveal a weak correlation between betweenness centrality and car accidents, suggesting that while high-traffic roads are crucial for urban flow, they do not singularly predict accident sites. This revelation underscores the multifaceted nature of urban traffic safety, urging a holistic consideration of various factors beyond mere network centrality.

The distinction between planned and self-organized cities, as explored through the comparative analysis of urban networks, further enriches our understanding of urban form and functionality. The inherent differences in network design between these city types illuminate the diverse challenges and opportunities they present for urban planning and safety interventions. Whether through the organized grid of planned cities or the organic sprawl of self-organized ones, the study confirms the critical impact of urban layout on accessibility, connectivity, and, by extension, safety.

This research not only contributes to the academic discourse on urban planning and network analysis but also lays the groundwork for practical applications that can significantly enhance urban living conditions. By integrating advanced tools like OSMnx and NetworkX with comprehensive accident data, this thesis provides a robust methodology for city planners, policymakers, and researchers to assess and improve urban networks. Such analyses are pivotal for designing safer, more efficient, and sustainable urban environments that respond dynamically to the evolving needs of their inhabitants.

Looking ahead, the imperative to continue exploring network science applications in urban planning cannot be overstated. The potential to refine our understanding of urban systems, predict future challenges, and devise innovative solutions is boundless. As cities grow and transform, the insights derived from network science will be instrumental in shaping urban policies and practices that prioritize safety, efficiency, and the overall quality of life. This ongoing inquiry into the nexus of network science and urban planning is not merely academic; it is a crucial endeavor towards fostering cities that not only accommodate but also enrich the lives of those who call them home.

Bibliography

1. Sevtsuk, A. (2017). Analysis and Planning of Urban Networks. In Encyclopedia of Social Network Analysis and Mining.
2. Porta, S., Crucitti, P., & Latora, V. (2006). The network analysis of urban streets: a primal approach. Environment and Planning B: Planning and Design, 33, 705-725.
3. Kirkley, A., Barbosa, H., Barthelemy, M., & Ghoshal, G. (2018). From the betweenness centrality in street networks to structural invariants in random planar graphs. Nature Communications.
4. Zhang, X., Miller-Hooks, E., & Denny, K. (2015). Assessing the role of network topology in transportation network resilience. Journal of Transport Geography, 46, 35-45.
5. Ingvardson, J. B., & Nielsen, O. A. (2018). How urban density, network topology and socio-economy influence public transport ridership: Empirical evidence from 48 European metropolitan areas. Journal of Transport Geography, 72, 50-63.
6. Scott, D. M., Novak, D. C., Aultman-Hall, L., & Guo, F. (2006). Network Robustness Index: A new method for identifying critical links and evaluating the performance of transportation networks. *Journal of Transport Geography*, 14, 215-227.
7. Crucitti, P., Latora, V., & Porta, S. (2006). Centrality in networks of urban streets. *Chaos*, 16, 015113
8. Masucci, A. P., & Molinero, C. (2016). "Robustness and closeness centrality for self-organized and planned cities." *The European Physical Journal B*, 89(2), 53.
9. Masucci, A., Smith, D., Crooks, A. et al. Random planar graphs and the London street network. Eur. Phys. J. B 71, 259–271 (2009).
10. Committee on Network Science for Future Army Applications (2006). Network Science. National Research Council. doi:10.17226/11516. ISBN 978-0309653886. S2CID 196021177.
11. Loder, A., Ambühl, L., Menendez, M. et al. Understanding traffic capacity of urban networks. Sci Rep 9, 16283 (2019)
12. Capello, R., & Camagni, R. (2000). Beyond Optimal City Size: An Evaluation of Alternative Urban Growth Patterns. Urban Studies, 37(9), 1479–1496. https://doi.org/10.1080/00420980020080221
13. World Resources Institute. (1994). World Resources 1994–95: A Guide to the Global Environment. Oxford University Press.
14. Richardson, H. W. (1972). Optimal city size: Theory and an empirical example. Journal of Urban Economics, 4(1), 69-85.
15. van den Berg, L., Drewett, R., Klaassen, L. H., & Rossi, A. (1983). A Comparative Study of Metropolitan Economic Development in Western Europe. Regional Studies, 17(6), 437–456.
16. Sevtsuk, A., & Mekonnen, M. (2012). Urban network analysis: A new toolbox for ArcGIS.
17. Barthelemy, M. (2010). Spatial Network. *Physics Reports-review Section of Physics Letters*, 499
18. Bullmore, E., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10, 186-198.

19. Zeballos, L. J., Gomes, M. I., Barbosa-Povoa, A. P., & Novais, A. Q. (2012). Optimum Design and Planning of Resilient and Uncertain Closed-Loop Supply Chains. In I.D.L. Bogle & M. Fairweather (Eds.), *Computer Aided Chemical Engineering* (Vol. 30, pp. 407-411). Elsevier. ISBN 9780444594310. ISSN 1570-7946

20. Kim, J., & Wilhelm, T. (2008). What is a complex graph? *Physica A: Statistical Mechanics and its Applications*, 387(11), 2637-2652
21. Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. 2007. Graph evolution: Densification and shrinking diameters. ACM Trans. Knowl. Discov. Data 1, 1 (March 2007), 2–es.

22. Portugali, J. (1997). Self-organizing cities. *Futures*, *29*(4–5), 353-380.
23. Ghoul, T., Sayed, T., & Fu, C. (2023). Real-time safest route identification: Examining the trade-off between safest and fastest routes. *Department of Civil Engineering, The University of British Columbia, Canada; School of Transportation Science and Engineering, Harbin Institute of Technology, China*.