University of Zurich[UZH]

Impact of Urban Structure on Mobility during COVID-19: A Polycentricity Perspective

GEO 511 Master's Thesis

Author
Adrian Nicolas Grossenbacher
16-054-363

Supervised by
Dr. Cheng Fu

Faculty representative
Prof. Dr. Robert Weibel

25.08.2023
Department of Geography, University of Zurich

Impact of Urban Structure on Mobility during COVID-19: A Polycentricity Perspective

Abstract

The COVID-19 pandemic has caused immense disruptions, particularly affecting urban mobility as a crucial aspect of infection containment efforts. While numerous studies have investigated various factors driving mobility changes, a substantial gap exists in understanding the influence of spatial structure in this context. This study addresses this gap by investigating the connection between spatial structure, particularly polycentricity, and mobility patterns during the pandemic. The polycentric structure of 384 U.S. Metropolitan Statistical Areas (MSAs) is assessed by employing a novel application of whole graph embedding on dynamic human mobility flow data. Utilizing dimensionality reduction and clustering techniques, the MSAs are categorized into monocentric, intermediate, and polycentric groups. The findings reveal a larger reduction within areas characterized by a higher degree of polycentricity. Despite these significant results, the applied regression model highlights the dominance of factors such as education, employment density, and public transportation. The results emphasize the complex nature of mobility and its drivers. When considering the broader concept of spatial structure, the applied model demonstrates a notable 12 to 25 % enhancement in R^2 performance, underscoring the importance of spatial structure on mobility reduction. This study not only offers valuable insights into how spatial structure, especially polycentricity, affected mobility during the pandemic, but also demonstrates the effectiveness of whole graph embedding in modeling the complexity of urban dynamics. The findings have the potential to shape spatial planning strategies, public health policies, and economic activities of urban space.

Keywords: Urban Spatial Structure, Polycentricity, Human Mobility Flows, COVID-19, Whole Graph Embedding, Machine Learning

Acknowledgements

I would like to take this opportunity to express my gratitude to everyone who has helped me in one way or another to complete this master's thesis. First and foremost I want to express my gratitude to my supervisor Dr. Cheng Fu for his guidance and support throughout the whole research project. His expertise and feedback have helped me understand the subject and improved this work immensely. I also want to thank the faculty representative Professor Dr. Robert Weibel for helping me set up the project. Considerable gratitude is extended to my friends and family for their emotional support and understanding, which kept me motivated. Lastly, I want to thank my partner for her continuous encouragement and for proofreading this thesis.

Table of Contents

Chapter 1: Introduction **1**

1.1 Motivation and Background ... 1

1.2 Research Aim ... 2

Chapter 2: State of Current Research **3**

2.1 Urban Structure ... 3

2.1.1 Theoretical Models ... 3

2.1.2 Urban Structure in the U.S. ... 4

2.2 Measuring Polycentricity ... 4

2.2.1 Traditional Indices ... 5

2.2.2 Whole Graph Embedding ... 6

2.3 Mobility during COVID-19 ... 7

2.3.1 Changes in Mobility ... 7

2.3.2 Explanatory Research ... 7

2.4 Research Gaps and Questions ... 8

Chapter 3: Methodology **11**

3.1 Data ... 12

3.1.1 Data for Whole Graph Embedding ... 12

3.1.2 Data for Regression Analysis ... 14

3.2 Preprocessing ... 14

3.3 Whole Graph Embedding ... 15

3.3.1 Whole Graph Algorithm ... 15

3.3.2 Scattering Feature Reduction ... 16

3.4 Cluster Analysis ... 16

3.4.1 Qualitative Analysis ... 16

3.4.2 Quantitative Analysis ... 17

3.4.3 Temporal Analysis ... 18

3.4.4 Further Explorative Analysis ... 19

3.5 Regression Analysis ... 19

3.5.1 Impact of Spatial Structure ... 19

3.5.2 Impact of Polycentricity ... 19

3.5.3 Model Building and Selection ... 20

Chapter 4: Results **22**

4.1 Whole Graph Embedding and Feature Clustering ... 22

4.2 Cluster Analysis ... 23

4.2.1 Qualitative Analysis ... 23

4.2.2 Quantitative Analysis 28
4.2.3 Temporal Change 31
4.2.4 Further Exploratory Analysis 33
4.3 Impact of Spatial Structure 34
4.3.1 Model Selection 34
4.3.2 Model Comparison 35
4.3.3 Model Interpretation 37
4.4 Impact of Polycentricity 39
4.4.1 Model Selection 39
4.4.2 Model Interpretation 39
4.5 Summary of Results 41

Chapter 5: Discussion **43**
5.1 Whole Graph Embedding Clusters 43
5.2 Impact of Spatial Structure 45
5.3 Impact of Polycentricity 45
5.4 Limitations 47

Chapter 6: Conclusion **48**
6.1 Summary and Implications 48
6.2 Future Work 49

References **50**

Appendix A: Figures **56**

Appendix B: Tables **60**

List of Figures

Figure 1: Morphological and functional polycentricity. 5
Figure 2: Example of various graph embedding areas. 6
Figure 3: Simplified workflow. 11
Figure 4: Absolute (above) and relative (below) mobility reduction for all MSA. 13
Figure 5: PCA transformed embedding vectors. 22
Figure 6: Overview of the selected MSA. 24
Figure 7: Lynchburg MSA, Virginia. 25
Figure 8: Baton Rouge MSA, Louisiana. 26
Figure 9: New Orleans-Metairie-Kenner MSA, Louisiana. 27
Figure 10.1: Social network measures; a) Burger et al. (2011); b) Green (2007); c) Limtanakool et al. (2007). 29
Figure 10.2: Rank-size distribution measure by Burger & Meijers (2012). 30
Figure 11: PCA embedded vectors before (left) and during (right) mobility restrictions. 32
Figure 12: Mobility reduction per cluster with standard deviation (above) and difference between clusters (below). 32
Figure 13: Estimated linear regression between PC_1 and a) total number of nodes, b) mean commuting time, c) employment density and d) cumulative COVID-19 cases per 1'000 people. 34
Figure 14.1: Graphic node features plotted against pairwise distance; a) Albany, GA; b) Austin, TX. 36
Figure 14.2: Graphic node features plotted against pairwise distance; a) Chicago, IL; b) New York, NY. 37
Figure 15: Feature importance of the WGE model (a) and the Base model (b). 38
Figure 16: Feature importance of the RF model. 41
Figure A.1: Enid MSA, Oklahoma. 56
Figure A.2: Houston-Sugar Land-Baytown MSA, Texas. 57
Figure A.3: Jacksonville MSA, Florida. 58
Figure A.4: North Port-Bradenton-Sarasota MSA, Florida. 59

List of Tables

Table 1: Overview of the mobility flow data. 12
Table 2: Polycentricity measures. 17
Table 3: Overview of variables and their sources. 20
Table 4: Pearson correlation between external polycentricity indices. 30
Table 5: Pearson correlation between external indices and principal components. 31
Table 6: Model selection. 35
Table 7: Model comparison. 35
Table 8: Pearson correlation coefficients of relative mobility reduction and node/edge features. 38
Table 9: Individual WGE feature importance of the top 61 features. 39
Table 10: Model selection. 39
Table 11: Pearson correlation coefficients between relative mobility reduction and various variables. 40
Table B.1: Overview of outlier MSAs. 60
Table B.2: Overview of polycentric MSAs. 60
Table B.3: Overview of intermediate MSAs. 61
Table B.4: Overview of monocentric MSAs. 64

Chapter 1: Introduction

1.1 Motivation and Background

At the end of 2019, the SARS-CoV-2 virus, which is believed to have originated in the Hubei province of China (Wu et al., 2020), changed our daily life. Despite early containment measures, the virus quickly managed to spread to other Asian countries and subsequently across the globe. After cases were confirmed in 113 countries, on March 11th, 2020 the World Health Organization (WHO) declared COVID-19 a pandemic because of its alarming levels of spread and severity (WHO, 2020). The United States was among the countries that have been hit the hardest by the virus. Due to weak early responses, the virus was able to spread rapidly across the country, and by March 12th, 47 states had already been infected. On the next day, President Trump declared the outbreak a national emergency (Whitehouse, 2020). Towards the end of March, the U.S. surpassed China in total confirmed cases, making it the country with the highest number of cases in the world (McNeil, 2020). Since the disease is an airborne infection and no vaccine was available yet, reducing mobility and thus social contact remained the main non-pharmaceutical intervention to halt the spread of the virus (Zhang et al., 2022). Densely populated urban areas, in which the majority of the population resides, are especially vulnerable to the spread of such contagious diseases. Numerous policies were introduced in order to slow down the outbreak and reduce its impact on public health. Restrictions were implemented on large gatherings and traveling; schools were closed, and stay-at-home requirements were issued (Hallas et al., 2021).

A lot of effort has been made for a better understanding of mobility and its influencing factors during the pandemic. Most research is hereby focused on various demographic and socio-economic factors (Fraiberger et al., 2020; Roy & Kar, 2020). On the contrary, the impact of urban spatial structure, such as the degree of polycentricity, has been largely neglected. Nonetheless, the degree of polycentricity significantly influences mobility patterns by shaping the spatial distribution of economic centers, particularly with respect to commuting (Schwanen et al., 2004). Moreover, polycentricity can be conceptualized as an outcome of dynamic interactions, such as mobility flows among local areas (more in Section 2.2). As such, the degree of polycentricity can also be affected by changes in mobility, due to sudden mobility restrictions, as seen during the pandemic. The structure of a city can also have a great influence on the spread of a disease. In fact, experiences from past outbreaks such as cholera have already significantly shaped the form of modern urban space (Bereitschaft and Scheller, 2020). Filling this research gap in regard to spatial structure is the main goal of this thesis. Understanding the impact of spatial structure on mobility could provide valuable information not only for policy assessment, but also for various aspects of urban spatial planning. Future urban space needs to be more resilient to pandemic outbreaks, as the frequency of emerging infectious diseases has increased in the last decades (Jones et al., 2008).

In order to assess the impacts of mobility-restricting policies such as stay-at-home orders or school closures, a lot of human mobility data has been collected and evaluated. The large amount of information gained from the Global Navigation Satellite Systems (GNSS) of mobile devices can be used to analyze the mobility behavior of individuals while guaranteeing anonymity. One such dataset by Kang et al. (2020) has been generated by

tracking millions of mobile phone users' trajectories in the U.S., collected by SafeGraph[1]. The set has some advantages over other available ones, such as a finer resolution on the census tract scale and integrated human mobility flow matrices to catch intra-city dynamics. With this data, the goal is to examine the impact of the urban spatial structure on mobility reduction during the COVID-19 pandemic in U.S. metropolitan areas.

1.2 Research Aim

Human mobility plays an important role in the direct transmission of viruses, and its management has proved an effective means of controlling outbreaks especially in the earlier stages. Extensive research has been conducted to understand the factors influencing mobility patterns. However, spatial aspects, such as the degree of polycentricity, have received limited attention. This master's thesis aims to analyze the influence of spatial structure, specifically polycentricity, on mobility during the COVID-19 pandemic outbreak in U.S. metropolitan areas. In order to achieve this, large-scale human mobility data will be analyzed, with a focus on comparing mobility patterns between monocentric and polycentric cities. This requires the quantitative assessment of polycentricity for various cities. Existing polycentricity indices have some shortcomings in regard to the complexity they are able to model (Section 2.2). For this reason, a novel whole graph embedding approach will be applied, which promises improvements in this aspect.

The rest of this study is structured as follows:

Chapter 2 provides an overview of the current state of research, addressing topics such as urban structure, measures of polycentricity, and mobility changes during the pandemic. The chapter closes by identifying research gaps and formulating research questions. Chapter 3 offers insights into the data, methods, and procedures applied in order to answer the research questions. Chapter 4 presents the obtained results, which are discussed in more detail in Chapter 5, where potential study limitations are also addressed. Lastly, Chapter 6 concludes the study by summarizing key findings and implications, while also outlining potential directions for future research.

[1] https://www.safegraph.com/

Chapter 2: State of Current Research

2.1 Urban Structure

The shape of urban space has changed dramatically since the industrial revolution. The rapid growth led to various challenges in regard to city planning (Knox and McCarthy, 2012). Urban economics and geography have since examined city structures extensively and introduced several theoretical frameworks that could help in efficiently planning a city (Anas et al., 1998). In the coming sections, some theoretical models are presented and the historic development of U.S. cities in regard to their structure is broadly explained. Lastly, it is discussed how polycentricity can be measured quantitatively.

2.1.1 Theoretical Models

The monocentric city model, first introduced by William Alonso (1964), has been the dominating model for urban analysis for decades. It remains a prominent model with its simplistic approach that can explain the urban structure and the land patterns within. At the core of the model lies the dominance of the Central Business District (CBD). In the monocentric model, the CBD lies at the center of a city where all economic activities take place and everyone commutes for work. Land close to the CBD is limited and in high demand, because people generally prefer lower commuting times. Because the land is expensive, owners prefer to build higher and more compact buildings to maximize their rent income. Towards the edge of the city, on the other hand, the land becomes more abundant, and the demand decreases due to longer commuting times. People can substitute housing costs with transportation costs based on their preferences. Although the model is still used today, its assumption of a central business district as the single employment center has been challenged by continued city sprawl and the emergence of multiple subcenters (Arribas-Bel and Sanz-Gracia, 2014).

This new urban form can be better described by the polycentric model, which is usually defined by the presence of one or more employment subcenters besides the CBD (McMillen, 2001). Although the historic CBD often remains as the largest employment center, it can happen that a subcenter outgrows it and becomes larger. The concept of land and housing prices remains as in the monocentric model, although now for several subcenters within the urban area.

There is a wide range of other urban models that try to capture certain aspects of urban structure, some of which have already been proposed before the monocentric model. This includes noteworthy models such as the concentric zone model (Burgess, 2015/1925), the sector model (Hoyt, 1939) or the multiple nuclei model (Harris and Ullman, 1945). There are further newer models, such as the maximum disorder model, the mosaic of live-work communities model and the constrained dispersal model (Angel and Blei, 2015). They try to better describe areas that cannot be assigned as easily to either the monocentric or polycentric model. Although further theoretical models exist, most of them can be seen as a hybrid form between the monocentric and polycentric models. Out of all the existing models, these two remain among the most prominent ones.

Although there has been a lot of debate about which city form should be aspired, most of them are based on opinions rather than on quantitative understanding. Nevertheless, the

transition from monocentric to polycentric has been observed in many cities. It is argued that this transition occurs due to increasing difficulties with congestion and pollution that would otherwise negatively affect the monocentric city (Louf and Barthelemy, 2014). This suggests that polycentric urban structure has a positive influence on the efficiency of the mobility network. However, the empirical findings on this are not conclusive. Sun et al. (2016) report a negative relation between average commuting times and the degree of polycentricity in Chinese cities. Another study conducted in the Netherlands reports significantly longer commuting times and distances for auto drivers in polycentric cities (Schwanen et al., 2004). Due to the definition of polycentric cities, most studies only examine commuting behavior but neglect general mobility in their analysis. Besides mobility, polycentricity has also been proven to influence socio-economic features, such as the gross domestic product (Meijers and Burger, 2010; Liu and Wang, 2016).

2.1.2 Urban Structure in the U.S.

In the case of the U.S., the transition of monocentric cities towards a polycentric structure has been well observed. The shift of the economic structure around the 1920s from manufacturing towards a service-oriented sector was one of the main factors. The industrial sector was traditionally located in the central places of a city. The new office buildings for the service sector were built as high-rise and high-density buildings, which led to the concentrated CBD we know today. The old industrial areas were virtually deserted and deprecated. Living in the main city became unattractive. The white-collar workers that followed the new service industry after 1945 wanted to live outside the city. The affordability of automobiles and, most importantly, of low-cost houses led to a high growth of the suburbs, surpassing the growth of the city itself. The people that lived outside the city were soon followed by supermarkets and other economic activities. Some of these suburbs continued growing until they became small cities themselves. These cities are called edge cities and could grow larger than the original main city itself (Knox and McCarthy, 2012).

The polycentric nature of a city in modern times is mainly dependent on the importance of its CBD compared to other employment subcenters in its nearer territory. A metropolitan area remains monocentric when most of the workers from the suburbs commute towards the CBD. Contrary, the upcoming new employment subcenters outside the city center would increase the polycentric structure depending on the number and size of the subcenters. Arribas-Bel and Sanz-Gracia (2014) looked at the structure of employment centers in 359 metropolitan areas at three different points in time (1990, 2000, and 2010). They found that the majority (around 60 %) of the areas still have a monocentric structure, especially smaller or medium-sized cities. They further show that polycentric areas are generally larger and have higher densities than monocentric ones.

2.2 Measuring Polycentricity

To appropriately analyze the spatial structure of urban space, the classification of the cities' mobility network in regard to polycentricity is essential. There are various different indices that measure polycentricity either from a *morphological* or *functional* perspective (Burger and Meijers, 2012). An illustration of these two dimensions is given in Figure 1. The morphological perspective measures the spatial distribution and size of urban attributes, e.g., the distribution of residents or workplaces. Whereas the functional polycentricity measures the interactions between different regions, e.g., the commuting flows.

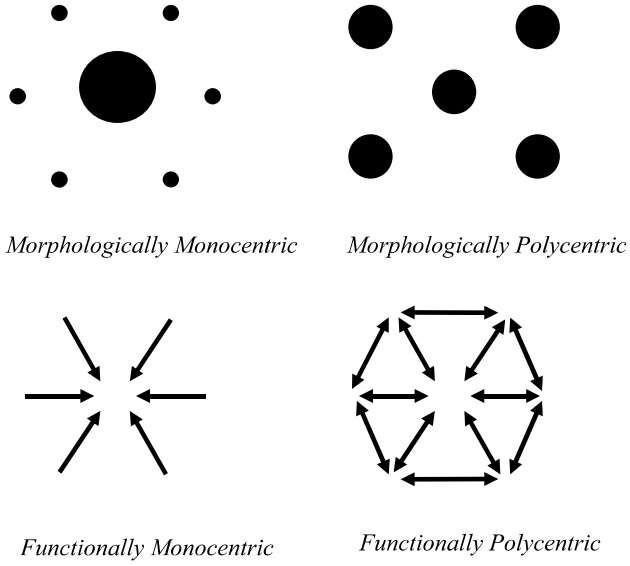


Figure 1: Morphological and functional polycentricity (Burger and Meijers, 2012, p. 1134).

2.2.1 Traditional Indices

The commonly used polycentricity indices can be broadly categorized into three main groups (Liu et al., 2016): 1) indices originating from social network theory; 2) measurements that use the slope of a regression line (rank-size distribution); 3) measures that compare observed polycentricity with an ideal theoretical model. The first group of indices usually uses social network statistics such as nodality or in-/external centrality to examine the polycentricity. The majority of all polycentricity indices can be categorized within this group. The second group believes that polycentricity is best described by the relation between the absolute and relative importance of centers. The rank-size distribution is therefore assumed to be a good indicator of polycentricity. The slope of a linear regression is used as the main measurement, whereas a flat slope indicates a higher degree of polycentricity. This group also utilizes the nodality and centrality of networks. An important aspect of the rank-size distribution is the number of nodes that are used to calculate the slope. Although this choice remains largely subjective, the use of a small but fixed number of nodes has been recommended (Burger and Meijers, 2012). The third group of indices derives the polycentricity by comparing an observed metric in the real world to an ideal model, such as the gravity model applied by Hanssens et al. (2014). This type of index generally relies on large amounts of data and requires fine-grained information. It is often limited to one-case studies and thus usually excluded in comparative studies on polycentricity (Bartosiewicz and Martcińczak, 2020).

As the study by Bartosiewicz and Martcińczak (2020) shows, these different types of indices are not consistent and come to different conclusions for the same city. In order to achieve satisfying results, the proper index has to be selected for the city in question, which still requires prior expert knowledge. Moreover, these indices only capture a single perspective of urban dynamics, missing out on the more complex patterns.

2.2.2 Whole Graph Embedding

Embedding techniques have become a powerful tool to convert discrete data into a low dimensional space as continuous vector representations, which are better suited for many analytical tools such as machine learning. A key feature of embeddings is that similar data entities are located closer within this vector space.

Embedding covers a broad spectrum of different fields of research. One of the earlier applications was in the field of Natural Language Processing (NLP). Words and phrases cannot be directly processed by machine learning algorithms. Word embedding methods such as Word2Vec (Mikolov et al., 2013) are able to extract vector representations of words. In the vector space, words with similar meanings are located closer to each other. These embeddings can then be used as direct input for machine learning for tasks such as sentiment analysis, text classification, or paraphrase detection. The same concept can be applied to sentences or even whole documents. Doc2Vec (Le and Mikolov, 2014) is a direct extension of Word2Vec that can vectorize documents. Besides the already extensively researched NLP, the concept has reached new areas of application such as image processing (Karpathy and Fei-Fei, 2015), audio (Luo et al., 2017), bioinformatics (Elnaggar et al., 2022), times series data (Kazemi et al., 2019) and graph analysis.

In the context of graph embedding, a few subcategories can be distinguished for certain parts of a graph. There is the embedding of nodes like the Node2Vec method (Grover and Leskovec, 2016) that only embeds the nodes of a graph as vectors. There are also specific methods that only embed the edges of graphs and hybrid methods that embed a combination of graph features, e.g., nodes and edges. Lastly, Whole Graph Embedding (WGE) does what its name suggests and embeds the whole graph. An exemplary overview of these different graph embedding categories can be seen in Figure 2.

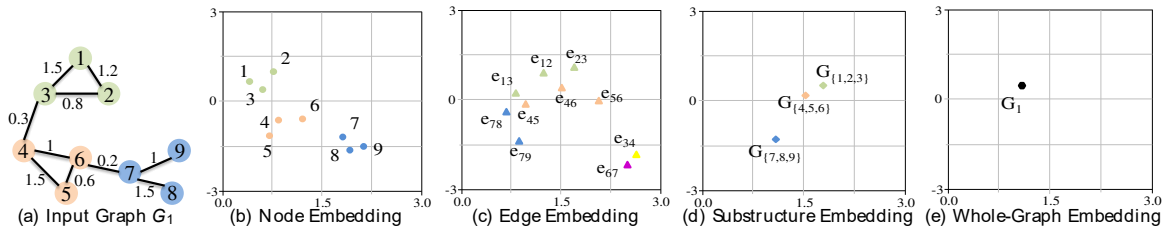


Figure 2: Example of various graph embedding areas (Cai et al., 2018, p. 2).

Whole graph embedding techniques can be differentiated into three main groups (Fu et al., 2021). The first group such as graph2vec (Narayanan et al., 2017) is graph-kernel based and combines a neural network with inputs from random walks through the graphs. The second group utilizes deep learning models such as neural networks (Niepert et al., 2016), and the last group uses spectral representation (Tsitsulin et al., 2018). There are various types of graphs that can be directed/undirected, weighted/non-weighted, or with/without node attributes. For this reason, different whole graph embedding methods are specialized to model a certain type of graph.

A novel application of whole graph embedding by Fu et al. (2021) promises to model the broader complexity of urban dynamics. This method can represent complex, non-Euclidean graphs as an embedding vector in an Euclidean space. Non-Euclidean graphs do not represent

a physical structure in two or three-dimensional space, but rather the entities and the relationships between them. Important is the relation between the entities and not the distance between them. Compared to these non-Euclidean graphs, their vector space representations are compact, capture similarity and perform better as input to machine learning algorithms (Chamberlain et al., 2017). Since whole graph embedding is able to map both node and edge features in the same vector, it can be used to model both morphological and functional polycentricity in an integrated manner, potentially improving measurement results.

2.3 Mobility during COVID-19

2.3.1 Changes in Mobility

Mobility research helped to predict and control the spread of the virus and was essential in evaluating implemented policies aimed at reducing human mobility. There is a large amount of literature describing this reduction in detail, most of which uses mobile GPS data from various sources. A global study across 52 countries by Nouvellet et al. (2021) shows a minimum mean reduction of 63 % from the baseline on March 11th, 2020. The variation between countries is quite large and shows relative numbers from 37 % up to 83 %. Although the difference between countries is high, the minimum reduction is still a substantial change. Besides the overall mobility reduction, changes in the mode of transportation and commuting habits can also be identified. A study in New Zealand and Australia found that especially public transportation was affected negatively (Thomas et al., 2021). A similar study in the United Kingdom reported the same trend, with a significant increase in walking or travel by bicycle (Harrington and Hadjiconstantinou, 2022). Another central aspect of many other studies lies in the evaluation of policies that were implemented to reduce mobility and thus the outbreak. Public transport closure, public event cancellation, workplace closure, and stay-at-home orders have been identified to be most effective in reducing mobility (Li et al., 2021).

Human mobility data during the pandemic has been the focus of numerous studies across the world. Regarding the U.S., reduction estimates are in the range of 35 – 63 % compared to normal conditions (Badr et al., 2020). This is a large difference considering it takes place within a single country. The political responses within the U.S. varied greatly among different regions. While some regions were rather restrictive, others opposed protective measures or wanted to lift measures too early. This can be explained by the differences in risk perception between the political groups (Bruine de Bruin et al., 2020). Despite these differences, the overall mobility reduction was substantial. Especially working conditions have changed ever since the technical possibility of working from home, completely eliminating the need for commuting. Brynjolfsson et al. (2020) report that between February and May 2020, over one third of workers in the U.S. switched to remote work from home. This is naturally not possible for all types of jobs, as not all work can be done from home (Huang et al., 2022).

2.3.2 Explanatory Research

There have been various studies that try to find explanatory variables that can shed light on the complex and dynamic nature of human mobility. Understanding which aspects influence the reduction in mobility is also instrumental to policymakers in the special context of the recent pandemic. A lot of focus has been placed on socio-demographic, economic, and cultural variables, especially in the U.S. Social and ethnical disparities are a hot topic in the U.S. political discourse. There are various studies that show significant evidence for

disproportionate mobility reduction between different socio-demographic groups. Several works find that the ability to reduce one's mobility is strongly dependent on the financial situation. Bigger changes in social distancing and mobility have been observed by increased family income or per capita income (Weill et al., 2020; Elarde et al., 2021; Carella et al., 2022). Besides racial differences in the pandemic on income or health, several studies also point out that certain ethnical groups exhibit differences in social distancing and general mobility reduction (Benitez et al., 2020; Coven and Gupta, 2020; Hu et al., 2021; Hu et al., 2022). Other socio-demographic aspects like partisanship or the level of education have also been identified to impact mobility during the pandemic (Grossman et al., 2020; Bollyky et al., 2023).

Only a few studies have also examined spatial variables in their mobility analysis. Zhu et al. (2021), for example, include the built environment such as gross employment density or street intersection density in their study. They found all built environment factors to be relevant for the tested counties, with employment density being one of the overall most important variables. Another study by Aguilar et al. (2020) has specifically examined the influence of polycentricity on spreading patterns and on the efficiency of mitigation measures. Although the more centralized cities favor spreading both in speed and extent, mitigation policies are shown to be more effective than in polycentral cities. They classify the cities by computing a metric called the flow hierarchy, which is derived entirely from trip-flows. However, they only examine the impact of polycentricity on epidemic severity but not on mobility directly.

2.4 Research Gaps and Questions

In regard to the related work, the following research gaps have been identified:

- The spatial structure of urban areas may play a crucial role in shaping mobility behavior during the pandemic. However, existing research has largely focused on socio-demographic variables, neglecting the impact of spatial features. Understanding the impact of spatial structure is essential for developing effective mobility reduction strategies and containment measures for the specific urban context.
- Polycentricity as a key aspect of urban structure has been largely absent in mobility studies during the pandemic. The impact of polycentricity on mobility is generally not well understood and existing studies report contradicting findings. By analyzing the importance of polycentricity in various urban areas, this research aims to find valuable insights into its influence on mobility behaviors.
- Comparative studies of polycentricity measures show contradicting findings and the selection of the appropriate indices requires prior expert knowledge of the study area. Traditional indices only capture either the functional or the morphological dimension of polycentricity, while modeling the polycentricity as a whole may give new insights. To address these shortcomings, this study aims to apply a novel whole graph embedding method that can model both dimensions.
- Most studies that aim to capture the degree of polycentricity have focused on a few well-known cities, leading to a limited understanding of urban spatial structure. To overcome this limitation, this study aims to analyze a broader, more complete set of metropolitan areas in the U.S. By considering diverse spatial structures of urban locations, new insights and trends may be uncovered.
- Changes in the degree of polycentricity have only been researched in a historical context over a larger timespan. However, the sudden and large reduction in mobility

due to the pandemic threat presents a unique opportunity to examine the change in the functional dimension of polycentricity. New insights can be gained by exploring how urban areas dynamically adapted their functional structure during the pandemic.

With these research gaps in mind, the following research questions (RQ) have been formulated:

RQ1: To what extent does the degree of polycentricity influence mobility reduction during the COVID-19 pandemic in U.S. metropolitan areas, and did the degree of polycentricity change after the implementation of mobility regulations?

More specifically, the relative change in general mobility flows will be used as the main indicator, with a focus on comparing polycentric and monocentric cities to identify any differences in mobility behavior. The change in the degree of polycentricity will be quantified using Principal Component Analysis (PCA) for two different time frames, which will enable a direct comparison.

Hypothesis 1: Assuming that an increased risk of virus spread leads to an increased regulatory mobility reduction, monocentric areas are expected to experience higher reduction rates than their polycentric counterparts. The high concentration of transportation and people in the center within monocentric areas increases the risk of infection (Aguilar et al., 2020). Conversely, decentralized polycentric areas with less crowded spaces are anticipated to show lower reduction rates due to a reduced infection risk.

Hypothesis 2: After the implementation of mobility restrictions to contain the virus, urban areas are expected to become more functionally polycentric. In order to halt the spread of the virus, the high mobility concentrations in central areas need to be restricted. This may lead to a shift in mobility away from the center, resulting in a more functionally polycentric structure than before.

RQ2: How important is the general spatial structure of U.S. metropolitan areas, as represented by whole graph embeddings, for mobility reduction during the COVID-19 pandemic?

One of the main benefits of whole graph embedding is its ability to capture the spatial structure of the graphs. The importance of spatial structure can be estimated by comparing the WGE with a baseline model that does not include spatial structure.

Hypothesis 3: Spatial structure greatly dictates the course of transportation routes for a given city, which therefore is essential for any mobility flows. As a consequence, spatial structure significantly influenced mobility reduction during the COVID-19 pandemic.

RQ3: How important is the degree of polycentricity for mobility reduction during the COVID-19 pandemic in U.S. metropolitan areas, compared to other socio-demographic variables such as ethnicity, education, or employment?

Spatial structure alone will not be enough to understand the complex nature of human mobility. It is also important to compare the contribution of spatial structure on mobility reduction with other influential factors. A regression model will be built to assess the influence of polycentricity and other important socio-demographic variables on mobility reduction.

Hypothesis 4: Polycentric areas exhibit a lower risk of viral transmission in comparison to monocentric areas (Aguilar et al., 2020). Therefore, the degree of polycentricity is expected to be a good indicator of the overall mobility reduction and will be among the most important variables in the regression model.

Chapter 3: Methodology

This chapter outlines the systematic approach and research design used to address the research objectives. For a better overview of the processes and results, a simplified workflow is shown in Figure 3. In the first step the raw data were *preprocessed* (Section 3.2) such that they can be used as input for the *Whole Graph Embedding* (WGE) (Section 3.3). The scattering features resulting from the whole graph algorithm were used in a *regression analysis* (Section 3.5.1) for model comparison and as further input for the *feature reduction* (Section 3.3.2). Within the feature reduction, clustering was applied to group data points with similar spatial structures together, enabling the *analysis* (Section 3.4) of differences and patterns between the groups. Besides the clusters, the feature reduction also generated several principal components. These components were then used in another, different *regression analysis* (Section 3.5.2), where their contributions were compared to various other socio-demographic and mobility-related variables. The following sections will describe each step in more detail. Both regression analyses will be discussed individually within the same section, as they have similarities in model building and model selection.

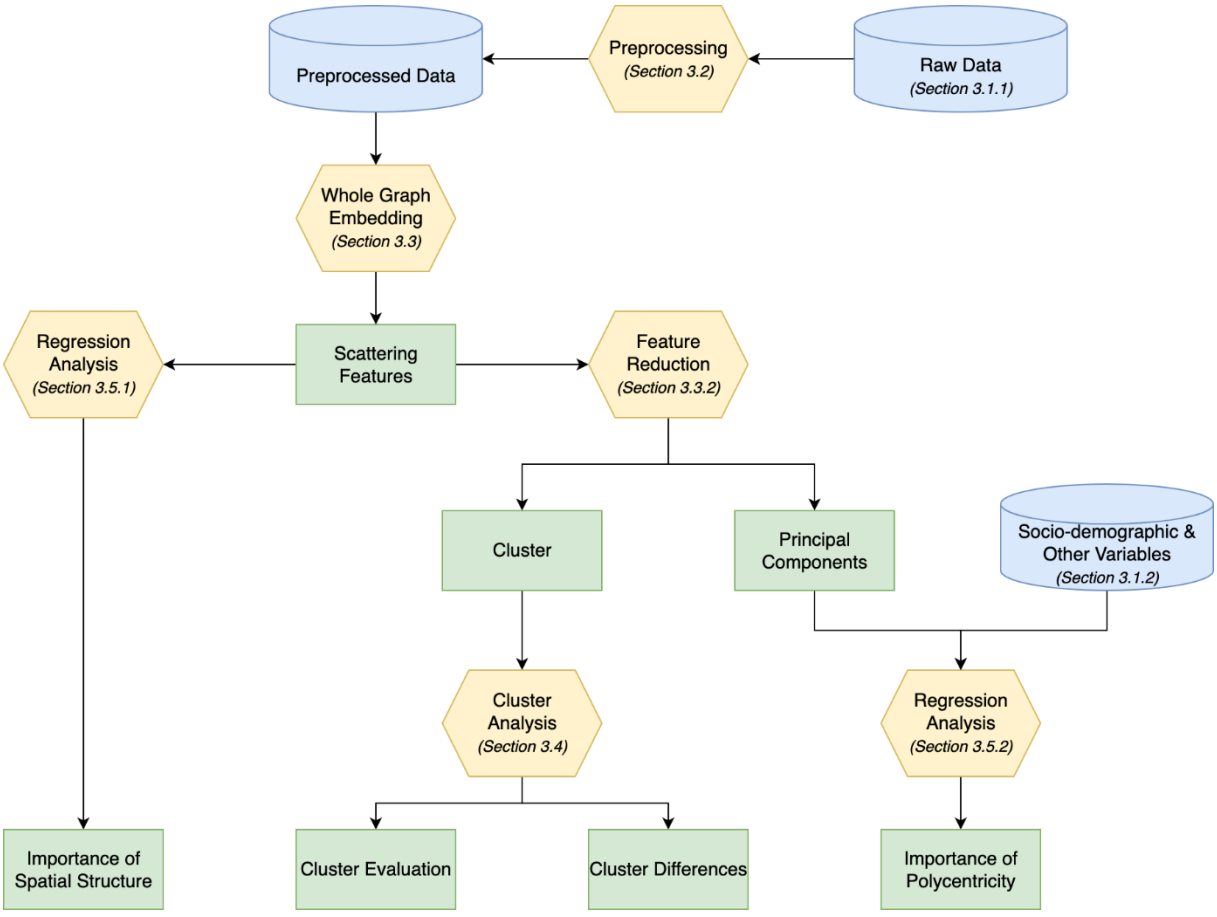


Figure 3: Simplified workflow; blue indicates data, yellow a process and green results.

3.1 Data

The upcoming two subsections describe the data used in this thesis. Section 3.1.1 describes the mobility data used for the whole graph embedding, and Section 3.1.2 introduces the variables used in the regression analysis.

3.1.1 Data for Whole Graph Embedding

The mobility dataset used in this thesis contains daily and weekly origin-destination (OD) flows starting from January 1st, 2019 to December 27th, 2021 in the U.S. (Kang et al., 2020). The whole dataset is publicly available and can be downloaded directly from GitHub[2]. The visitor flows are directly calculated from the mobile phone GPS data collected by SafeGraph (2020). The data is computed from millions of mobile phone users' visits to various places and is aggregated at three geographic scales: census tract, county, and state.

Table 1: Overview of the mobility flow data.

Variable	Description	Example
geoid_o	Unique ID of the origin	25027728400
geoid_d	Unique ID of the destination	25027708100
lng_o	Latitude of the geometric centroid of the origin (GCS NAD 83)	-71.85086158823391
lat_o	Longitude of the geometric centroid of the origin (GCS NAD 83)	42.38162773835911
lng_d	Latitude of the geometric centroid of the destination (GCS NAD 83)	-71.90265475768281
lat_d	Longitude of the geometric centroid of the destination (GCS NAD 83)	42.55122458473228
date_range	Date range of the records	01/20/20 – 01/26/20
visitor_flows	Estimated number of visitors detected by SafeGraph	28
pop_flows	Estimated population flows, inferred from visitor_flows	554

Table 1 provides an overview of the attributes of the mobility data. The identifier of the census tracts is the same as the Federal Information Processing Standard (FIPS), which is used to uniquely identify any census tract within the U.S. The country is divided into several geographical areas for the purpose of data collection and analysis. The whole nation consists of several states that are divided into multiple counties and each county can be further divided into smaller units called census tracts. The identifier mentioned above consists of an 11-digit code, whereby the first two digits represent the state, the next three the county, and the last six the individual census tract. The data also includes the dynamic population flows, which are inferred from the visitor flows. To get the population flows, the visitor flows are multiplied by

[2] https://github.com/GeoDS/COVID19USFlows.com

the ratio of the total population and the total number of mobile phone devices detected. However, this work focuses on visitor flows only.

There are several mobility datasets available for public usage provided by different sources such as Google, Facebook, or Twitter. However, these come with three main limitations. They do not come with human mobility flow matrices, have a lower spatial resolution, and provide only a sample of the entire population. The chosen dataset addresses two of the three mentioned limitations. However, data bias remains a common issue as not everyone has a mobile phone or uses a smartphone application with activated GPS tracking. It is estimated that the data represents around 10 % of the whole population.

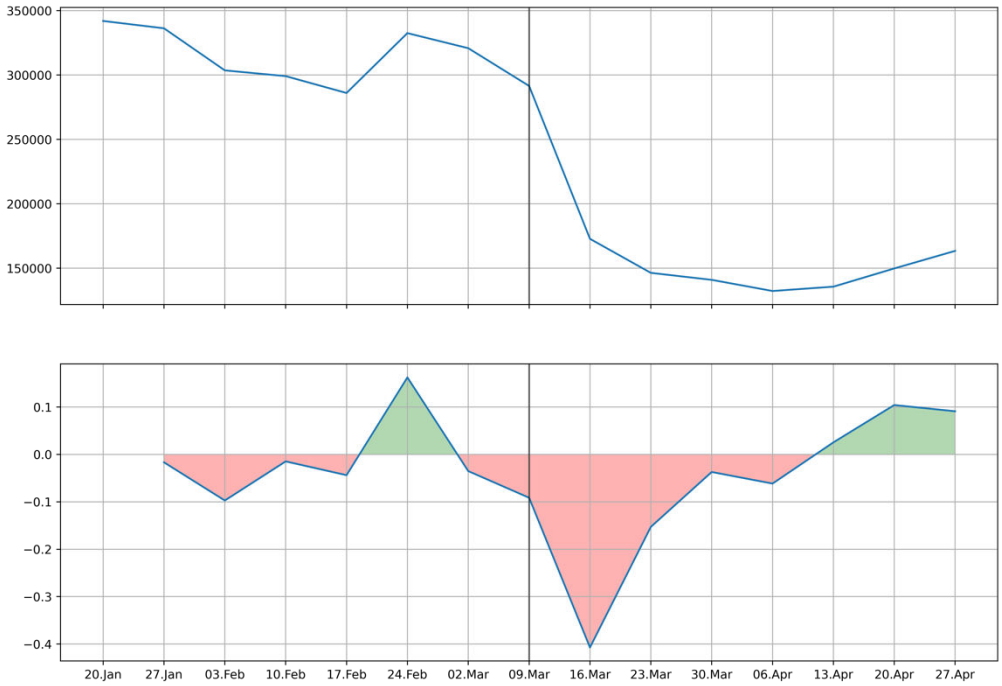


Figure 4: Absolute (above) and relative (below) mobility reduction for all MSA.

Although the U.S. administrative areas reacted very differently towards the pandemic thread, a uniform decline in mobility flows can be seen after the announcement of a national emergency by the president in the second week of March 2020. This week acts as the transition point between the two time frames before and after regulations are implemented. As shown in Figure 4, a total of seven weeks before and after the transition week have been included in this study. The time periods are sufficiently long to capture the main drop in mobility as well as the start of a slow recovery phase.

Besides mobility data, employment information is also utilized in the WGE. This is described in more detail in the following section, since the regression analysis uses the same employment data.

3.1.2 Data for Regression Analysis

A range of different variables is needed to model the complex dynamic of urban mobility flows, and to make a statement about the importance of spatial variables in comparison. Most of the included variables are from the U.S. Census Bureau[3], which provides a number of different datasets about the nation and its people. These datasets can be directly accessed through their Application Programming Interface (API). The American Community Survey (ACS) is one of the more extensive sets that grant insight into many socio-demographic areas. The public transportation rate is the ratio of people that use public transportation to commute to work. The data regarding the political party comes from the Massachusetts Institute of Technology[4] (MIT). The dataset contains results from presidential elections from 2000 to 2020 on a county level. The most recent available data for each county will be used for the variable. When including employment data, it is important to know whether the number of jobs is counted in the workplace area or the residential area, e.g., if the data reflects where the employees live or where they actually work. For that reason, employment data from a subset of the Longitudinal Employer-Household Dynamics (LEHD) is used, called the LEHD Origin-Destination Employment Statistics (LODES) dataset. The LODES set has a special focus on commuting and employment patterns and contains both information regarding workplace characteristics as well as residential areas. Any densities are computed by dividing the variables by the area of the respective census tract. These area sizes are extracted from the Census Bureau's Planning Database (PDB).

3.2 Preprocessing

The raw mobility data needs to be cleaned and organized before it can be used as input for the whole graph algorithm, which is described in the next section. After downloading the files for the chosen time frames, the data for the Metropolitan Statistical Areas (MSAs) was selected. An MSA is defined as an urbanized area of at least 50'000 population, plus adjacent territory that has a high degree of social and economic interaction as measured by commuting ties (Office of Management and Budget, 2021). An MSA is made up of one or more counties. MSAs account for around 85 % of the total U.S. population. The integration of commuting in the definition makes MSA the ideal geographic unit for this study. The selection was done with the help of a delineation file from the U.S. Census Bureau[5], which lists all counties to their relevant metropolitan area. By linking the first five digits of the census tract identifier with the county identifier from this delineation file, only census tracts within an MSA were kept. For the purpose of this work, areas in Puerto Rico have been excluded, because of the inherent differences to other U.S. states, including the spatial structure of their urban areas. Flows between different MSAs have also been excluded, such that the data only included sets of origins and destinations within the same MSA. Finally, the mean for each area was calculated for the two separate time periods of seven weeks each. The preprocessed data reflects information from a total of 384 metropolitan areas within the U.S.

The conceptual model for whole graph embedding first represents the graph as a collection of nodes and edges, where nodes correspond to entities, and edges capture relationships between these entities. In this thesis, the nodes are represented by the census tracts and the edges by the mobility flows. With this data, the structural information of the transportation network can be modeled. In addition to the graph structure, the WGE model can take advantage of node

[3] https://www.census.gov/
[4] https://dataverse.harvard.edu/
[5] https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html

features or attributes. These features describe certain characteristics of the nodes, which can provide valuable information that helps enhance the quality of the graph embedding. Several other node features were used as input for the whole graph algorithm in this study, including the *total volume of inflows*, *self-loop flows*, *clustering coefficient*, *employment*, and *degree centrality*. The first three features have been successfully applied in the study by Fu et al. (2021). The degree centrality is a graphic node feature used in the original whole graph algorithm (Gao et al., 2019). Employment, as the total number of workplaces per census tract, is the last node feature used. Employment data is a core feature commonly used by other indices to assess the degree of polycentricity.

3.3 Whole Graph Embedding

The central purpose of the WGE applied in this thesis is to capture the spatial structure of the urban areas in an Euclidean space as described in Section 2.2.2. This allows for further analytical application of clustering and machine learning. This section includes the main whole graph algorithm used in this thesis (Section 3.3.1), as well as the feature reduction and clustering (Section 3.3.2) of its output.

3.3.1 Whole Graph Algorithm

The geometric-scattering-based algorithm by Gao et al. (2019) was applied in this thesis. The algorithm basically extracts scattering features from a graph by applying cascading multi-layer wavelet transforms. Although there are also other options as mentioned in a previous section, this algorithm is able to embed weighted directed graphs, such as OD graphs. Meaning that both morphological (node features) as well as functional (edge features) polycentricity can be modeled.

The algorithm is available on GitHub[6] together with a couple of example scripts to showcase its uses. The whole graph algorithm uses two inputs: edge features and node features. The edge features are given as a list of adjacency matrices and the node features are also given as a list. The created matrices show the directed flows for each pair of census tracts within an MSA. The calculation of the clustering and degree node features first required the construction of graphs. These were created with the network analysis library NetworkX[7]. Once the weighted and directed graphs were created, employment data was also added as an attribute for each node within the graphs. With the adjacency matrices and the corresponding graphs, all node features could now be calculated. The inflow and self-loop flows could simply be derived from the adjacency matrices. The employment data could also be read directly from the node attributes of the graphs. The clustering coefficient and the node degree as the two graphic features were computationally more expensive to generate.

With all node and edge features prepared, the actual whole graph algorithm could be applied. The algorithm produces 64 scattering features for each individual node feature given as input. Using five node features resulted in 320 scattering features. These scattering features always have the same values for each node feature, regardless of the combination of features, meaning that there is no randomness involved and one could use the algorithm for each node feature separately and then add them together afterwards.

[6] https://github.com/matthew-hirn/geo-scattering-graph-data
[7] https://networkx.org/

3.3.2 Scattering Feature Reduction

The scattering features resulting from the whole graph algorithm still have a high dimensionality, which needs to be reduced for clustering purposes. Firstly, the multi-correlation between the features was inspected. For this purpose, the Variance Inflation Factor (VIF) was examined as a measurement of collinearity. The VIF was computed for each feature whereas higher values indicate a higher collinearity. The feature with the highest value was then iteratively removed until the highest value was no greater than a chosen threshold. As a rule of thumb, this threshold is often set at 5, which was also applied in this context.

To further reduce these dimensions for clustering purposes, Principal Component Analysis (PCA) was applied. PCA increases the interpretability of the data while preserving most of its information. For this work, principal components that can explain at least 90 % of the variance were kept.

In order to cluster the principal components obtained from the PCA, a k-means algorithm was applied, and the best k was estimated using the elbow method. The first two principal components were used for clustering in a two-dimensional space, and the clusters were color coded according to the results of the k-means clustering.

3.4 Cluster Analysis

The aim of the cluster analysis is firstly to label the clusters according to the degree of polycentricity, and secondly, to examine and explain any behavioral differences between the clusters in regard to mobility reduction. Assessing feature clustering results without any ground truth is always a challenging task, especially, when no expert knowledge about the areas is present. Nevertheless, an effort was made to assess the clustering results. In the upcoming sections, the clusters were analyzed qualitatively (Section 3.4.1) as well as quantitatively (Section 3.4.2). Furthermore, the cluster-specific temporal changes were examined (Section 3.4.3), and further explorative analysis was conducted (Section 3.4.4).

3.4.1 Qualitative Analysis

To get a good first impression of the distribution of the degree of polycentricity within the vector space, some individual MSAs were examined qualitatively. Often, an aerial view of the urban area can give a good first impression of the overall spatial structure. To visually examine such aerial views, QGIS was used to display Google satellite images of some areas. The borders of the MSA as well as of the individual census tracts were also displayed for a better understanding of the network structure. The satellite images of the selected MSA were compared in regard to their position within the clusters, and some simple demographics were also examined. The position of the individual points in the clusters can further be assessed with other indices, such as the urban sprawl index by Smart Growth America (2014). For each main cluster, the MSA closest to the centroid was examined. This MSA is the most representative of its cluster. Additionally, the extreme areas that are most different from each other were examined as well. Any outliers were also checked.

3.4.2 Quantitative Analysis

To complement the qualitative analysis and make a broader statement about the polycentricity of all MSAs and their clusters, a more comprehensive quantitative analysis was conducted as well. For each MSA, several commonly used polycentricity indices were calculated. These indices were then used to color-code the respective data points within the vector space. This visual representation further improved the understanding of the overall patterns of polycentricity within the clustering results. The similarities between the external indices and the principal components were further examined by comparing their correlation coefficients.

Table 2: Polycentricity measures.

Index	**Morphological Polycentrictiy**	**Functional Polycentricity**
Green (2007)	$P_M = 1 - \frac{\delta_F}{\delta_{Max}}$	$P_F = 1 - \frac{\delta_F}{\delta_{Max}} * \Delta$
	δ_F: The standard deviation of total inflow in the network.	$\Delta = \frac{L}{L_{Max}}$
	δ_{Max}: The standard deviation of inflow in a two-node network where one node has zero and the other has the maximum observed inflow.	Delta represents the network density where L is the total inflow and L_{Max} equals the maximum possible value of inflow for a single node.
Burger et al. (2011)	PM = 1 – ratio of employment in the largest center and the total employment in the region.	PF = 1 – ratio of inflow into the largest center and the total inflow in the region.
Limtanakool et al. (2007)	—	$El = -\sum_{i=1}^{L} \frac{(Zl)Ln(Zl)}{Ln(L)}$
		For z = 0 holds that (z)Ln(z) = 0
		L: Total number of links in the network ($l = 1,2,3, \dots, L$).
		Zl: Journeys on link l in relation to the total number of journeys in the network.
Burger and Meijers (2012)	The rank-size distribution of the nodality scores, where nodality is counted as the total employment.	The rank-size distribution of internal centrality scores, where internal centrality is counted as the total inflow.
	The 54 largest nodes are chosen to calculate the slope of the regression.	The 54 largest nodes are chosen to calculate the slope of the regression.

A total of four different indices from two groups of measures were applied for the data and used for comparison. Table 2 provides an overview of the measurements used, their formulae as well as the dimension of polycentricity measured. The first three (Green 2007; Limtanakool et al., 2007; Burger et al., 2011) are all methods that originate from social network theory. The method by Burger and Meijers (2012) measures polycentricity with a rank-size distribution. The third group of polycentricity indices compares observed polycentricity with an ideal model. However, their application is limited to cases with fine-grained data of the population, and as Bartosiewicz and Marcińczak (2020) point out: *"..., they are virtually never employed in comparative studies on polycentric urban development and/or on the effects of polycentricity on social and economic phenomena"* (Bartosiewicz and Marcińczak, 2020: 3). For this reason, this group of measures was not included in this work.

The formulae have been mostly applied, as in the study by Bartosiewicz and Marcińczak (2020), with the exception of the measure by Burger and Meijers (2012). Originally, they estimate the slope of the second, third, and fourth largest nodes in the network and then calculate the average of these three scores. The scale hereby is substantially different, as they use whole cities as nodes in a national network. They point out that the sample size of nodes used to calculate the slope of the regression should best be based on a fixed number of centers. In an earlier work by Meijers and Sandberg (2008), they set the sample size as the average number of nodes across all networks. If a network should have less than the fixed sample size, then all nodes are included. The same strategy was applied to the calculations in the adjusted rank-size formula. The median value of 54 has been selected as the sample size, as the distribution across all networks is skewed. Another small adjustment was made to the morphological polycentricity index by Burger et al. (2011). Bartosiewicz and Marcińczak (2020) use commuting flows for both morphological and functional index, perhaps due to data availability reasons. However, the original study uses employment data for the morphological dimension. Since such data was available for this study, it was applied as in the original formula. For an easier direct comparison, all polycentricity indices were adjusted such that their values range from 0 to 1, where a higher value indicates a higher degree of polycentricity.

3.4.3 Temporal Analysis

In order to examine any potentially different cluster-specific behavior over time, their individual change was analyzed before and after mobility restrictions were implemented. These changes were then compared to determine whether there are any differences in how the clusters reacted to the mobility reduction. This was done in regard to mobility and the degree of polycentricity. For mobility, the relative change in mobility flows was investigated with basic statistical methods. The relative change in polycentricity required a bit more preparation. The datasets from the two different time periods were joined together into a single dataset, and the same process as in Section 3.3.2 was applied. First, the features with high multicollinearity were removed, and then the principal component analysis was applied to further reduce the dimensionality of the data. The resulting principal components could then be visualized and compared within the same space. The data entries were colored according to the previously obtained clusters, allowing a comparison of cluster-specific movement. The centroids of the clusters were used to give a good quantitative measure of how much the cluster as a whole has moved.

3.4.4 Further Explorative Analysis

To gain a better understanding of any possible differences between the clusters, the relation between the first principal component and other, mainly spatial variables was examined. Some of these variables were also used in the second regression analysis later on. The relation was examined by estimating a linear regression between the variables. The slope of this regression gave insights into the relation of the variables that were being compared and helped form an explanation for the different behaviors of the clustering groups.

3.5 Regression Analysis

Two different regression analyses were conducted in the scope of this thesis with the aim of better understanding the relationship between spatial structure (w.r.t. polycentricity) and mobility reduction during the pandemic. In the first regression analysis, a model based on results from the WGE was compared to an aggregated baseline model. The difference in model performance was used to illustrate the importance of spatial structure for mobility reduction as modeled by the whole graph algorithm. The second regression analysis was used to compare the importance of independent variables within a single model. More specifically, the focus was put on the comparison of polycentricity with other socio-demographic, spatial, and travel-related variables.

3.5.1 Impact of Spatial Structure

To assess the importance of spatial structure for the mobility reduction, two models were compared with each other. The general idea was to compare a whole graph model with a base model. For both models the dependent variable was the relative mobility reduction. The whole graph model used the scattering features obtained from the whole graph algorithm as explanatory variables, whereas the base model used mean values of the raw node and edge features. Since the whole graph algorithm can model the spatial structure of a graph, the difference in performance between these two models will reflect the impact of the spatial structure. Furthermore, the impact of graphic node features was examined by removing graphic node features in both models. This resulted in a total of four models being compared with each other. Besides the performance metrics, the feature importance, as well as the Pearson correlation coefficient of the independent variables and the mobility reduction, was also compared. All variables were normalized prior to the regression analysis.

3.5.2 Impact of Polycentricity

To get a better understanding of polycentricity and its impact on mobility reduction, it was compared to a range of mostly socio-demographic variables. An overview of the variables and their respective source is shown in Table 3.

*Table 3: Overview of variables and their sources. Note: * square root transformed.*

Category	Variable	Source
	Explanatory variables	
Socio-demographic	Median age	U.S. Census Bureau
	Sex ratio	
	Black or African American	
	Asian *	
	Hispanic or Latino	
	Unemployment rate	
	Poverty rate	
	Gini index	
	Higher education	
	Dominant political party	Massachusetts Institute of Technology
Travel-related	Mean travel time	U.S. Census Bureau
	Public transportation rate *	
Spatial	Employment density *	U.S. Census Bureau
	Polycentricity (principal components)	Own work
	Dependent variable	
Mobility data	Relative mobility reduction	SafeGraph

Variables with high collinearity and low overall feature importance have been removed. Any variables with a skewed distribution were transformed by a square root function. The first, second, and forth principal component (PCs) have also been transformed using a square root function. Additionally, a power transformation was applied on the ninth PC. The skewness was assessed using the skew and kurtosis function from the SciPy[8] library. All variables were normalized before applying the regression analysis. As in the other regression analysis, the feature importance as well as the Pearson correlation coefficient were used for further comparison.

3.5.3 Model Building and Selection

The general workflow of the regression analysis was the same for both prediction tasks. The hyperparameter tuning of different algorithms was performed with the GridSearchCV method from the scikit-learn[9] library using a 10-fold cross-validation. The best set of hyperparameters was then used in a nested cross-validation to measure the model performance. 10-fold cross-validation was applied in both the outer and inner loop of the nested cross-validation. To choose the best model, a range of performance metrics was compared. This includes the R^2 score, Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). For all metrics, the mean value and the standard deviation from the nested cross-validation were reported. In earlier testing, ensemble methods showed the overall best performance. These methods also have an integrated feature importance measure, which is useful for assessing the impact of certain variables. For this reason, a Random Forest (RF) model, an AdaBoost (AB) model, and a Gradient Boosting (GB) model were compared in more detail. The best performing model out of these three was selected based on the metrics

[8] https://scipy.org/
[9] https://scikit-learn.org/stable/

and then examined in more detail. The main parameters that were fine-tuned for RF and GB were the number of trees used in the model and the number of features to consider when looking for the best split. The AB model underwent fine-tuning for the number of trees utilized and the weight assigned to regulate the impact of individual weak learners.

Chapter 4: Results

In the upcoming sections, the results to answer the given research questions (Section 2.4) are presented. Section 4.2 covers RQ1 by examining the clustering results, which are presented in Section 4.1. RQ2 will be answered in Section 4.3, in which the importance of spatial structure is examined. In Section 4.4 the impact of the degree of polycentricity will be analyzed in order to answer the final RQ3. Lastly, a summary of the findings is presented in Section 4.5.

4.1 Whole Graph Embedding and Feature Clustering

The removal of features with a high multicollinearity resulted in a total decrease from 320 to 50 features. The further application of the dimensionality reduction with PCA resulted in 9 principal components that can retain 90 % of the variance, whereas the first component alone can explain above 50 % of the total variance. This adds up to 65 % of the variance together with the second principal component. The results of the k-means clustering with the best k of 4 are shown in Figure 5. A complete list of individual Metropolitan Statistical Areas (MSAs) with their respective cluster and other important information is shown in Appendix B.

Figure 5: PCA transformed embedding vectors.

Three main groups can be identified with an additional small outlier cluster. The distribution of areas per cluster is skewed. The majority of the MSAs are placed within Group 2 (blue) on the very left. This cluster includes 306 out of the total 384 areas, coming close to 80 % of all examined MSA. The cluster labeled Group 1 (orange) consists of 60 MSAs, Group 4 (green) of 16, and the outlier Group 3 (red) of only 2 areas. It is noticeable that the clusters differ in terms of their compactness or spread. Group 2 is very compact with a mean spread of 0.68 from the cluster centroid. Group 1 is more extended with a mean spread of 2.17, and Group 4 is very widespread with a mean distance from the centroid of 5.85. Most of the visible difference of the three main groups is shown along the PC 1 axis.

4.2 Cluster Analysis

The aim of this section is to appropriately label the four clusters identified in the previous section based on their degree of polycentricity. This is accomplished through a combination of qualitative evaluation using satellite images and quantitative analysis applying various polycentricity measures. After labeling the clusters according to their degree of polycentricity, a comparison is conducted between the different groups in terms of their mobility behavior and structural changes. Lastly, an effort is made to explain any observed differences between the clusters by examining factors that may have influenced their mobility patterns.

4.2.1 Qualitative Analysis

A total of seven selected MSA are examined in this section. Figure 6 provides an overview of the MSA and their position within the clusters. Due to space reasons, only the images of the three areas Lynchburg, VA, Baton Rouge, LA, and New Orleans, LA, are displayed within this section. These three are closest to the centroid of their respective cluster. The images of the other areas are attached in Appendix A. For each MSA, two images are displayed, one with only MSA borders and one with additional census tract borders.

The first area examined is the MSA of Lynchburg in Virginia (Figure 7). It is closest to the centroid of Group 2 (blue) on the very left. It is a quite small metropolitan area with only 60 network nodes and a total population of 263’937. The area looks quite rural with a lot of green spaces and only a few smaller urban patches. One main center can be seen in the middle with a few smaller settlements scattered around. The census tracts are relatively equal in size, compared to other MSAs. Most of the main urban core seems to be made up of just a single census tract. The area is listed among the most sprawling small metropolitan areas in the Measuring Sprawl 2014 report by Smart Growth America. However, a city can still be monocentric if it experiences the sprawl adjacent to its already existing built-up area. The sprawl would only indicate increased polycentricity if it expands outside of the existing built-up area, which does not appear to be the case based on the aerial view.

The second image in Figure 8 depicts the MSA of Baton Rouge, Louisiana. It is at the center of the orange cluster in the middle. It is significantly larger with 157 nodes and a population of 856’779. The census tracts are visibly smaller in the urban center and quite large towards the border area. In regard to the sprawl report, the area is listed as the most sprawling medium sized metropolitan area.

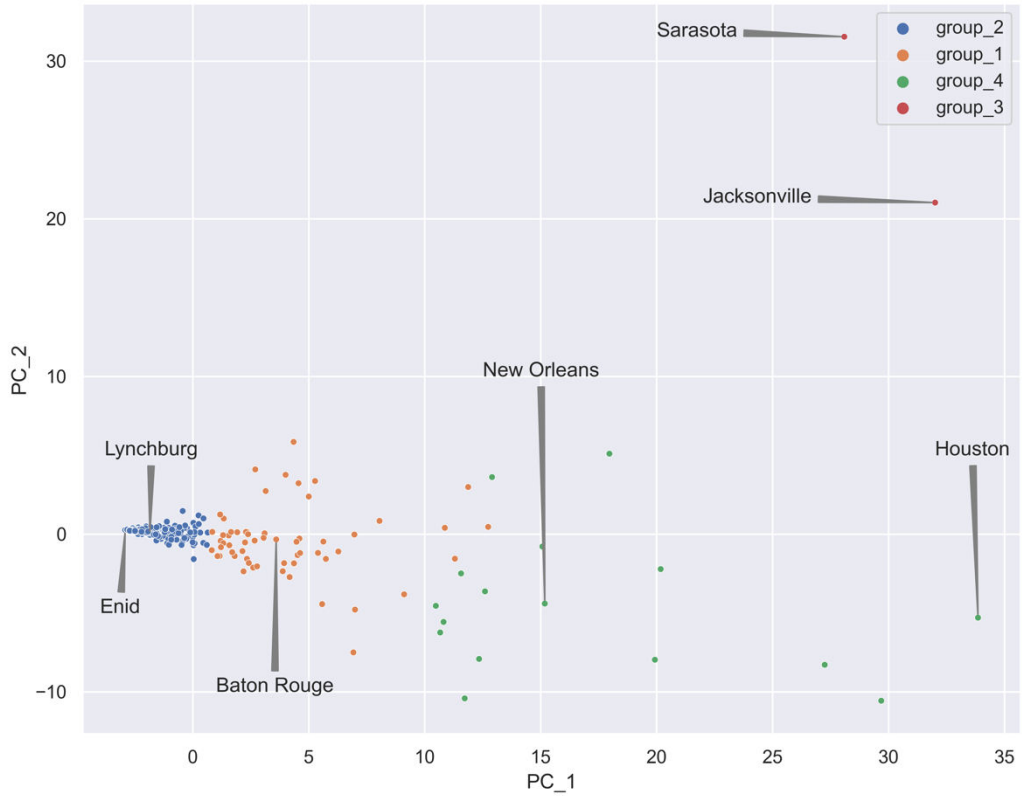


Figure 6: Overview of the selected MSA.

The last area presented in Figure 9 is the MSA of New Orleans-Metairie-Kenner, Louisiana. It is the central MSA of the green cluster to the right. This area is adjacent to the south-east of Baton Rouge. It is the largest of the three with 403 network nodes and a population of 1'271'651. Besides the main urban core, at least two other medium sized settlements can be identified. Most census tracts seem to be very small in the urban center with only a few larger ones towards the border. New Orleans is ranked among the less sprawling metropolitan areas in the report of 2014. The area has a few water bodies in or adjacent to its territory, which could be a natural limit to its sprawl. Its close proximity to the Gulf of Mexico brings challenges in regard to natural hazards such as hurricanes. New Orleans was in fact severely hit by Hurricane Katrina in 2005, where most of the city was completely flooded. This will most likely have limited sprawling activity for several years, which could also explain the low sprawl index from 2014.

The two MSAs Enid in Oklahoma and Houston in Texas have been included in the appendix in Figure A.1 and A.2, respectively. Both represent extreme cases with Enid being the MSA placed on the far left and Houston being on the far right of the PC 1 axis. Enid's structure is clearly monocentric, while the structure of Houston is very polycentric. Texas is known for its economic-oriented spatial planning policy, and Houston itself is famous for having no zoning ordinance (Qian, 2010). It is therefore not surprising that Houston is also the number one sprawling large MSA in the urban sprawl report. There is no sprawl index for Enid, as the report excluded smaller areas.

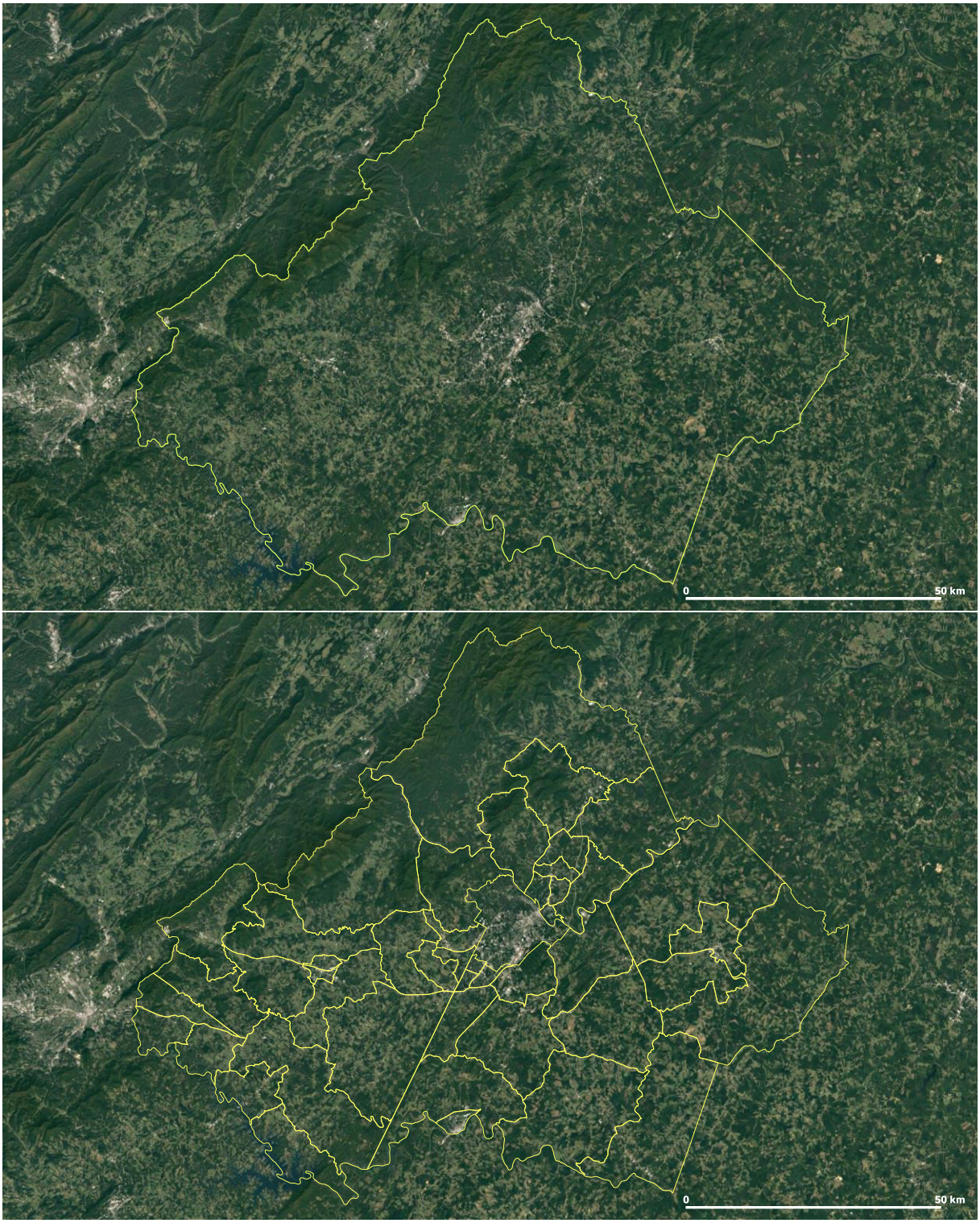


Figure 7: Lynchburg MSA, Virginia; MSA border (above) and census tract borders (below).

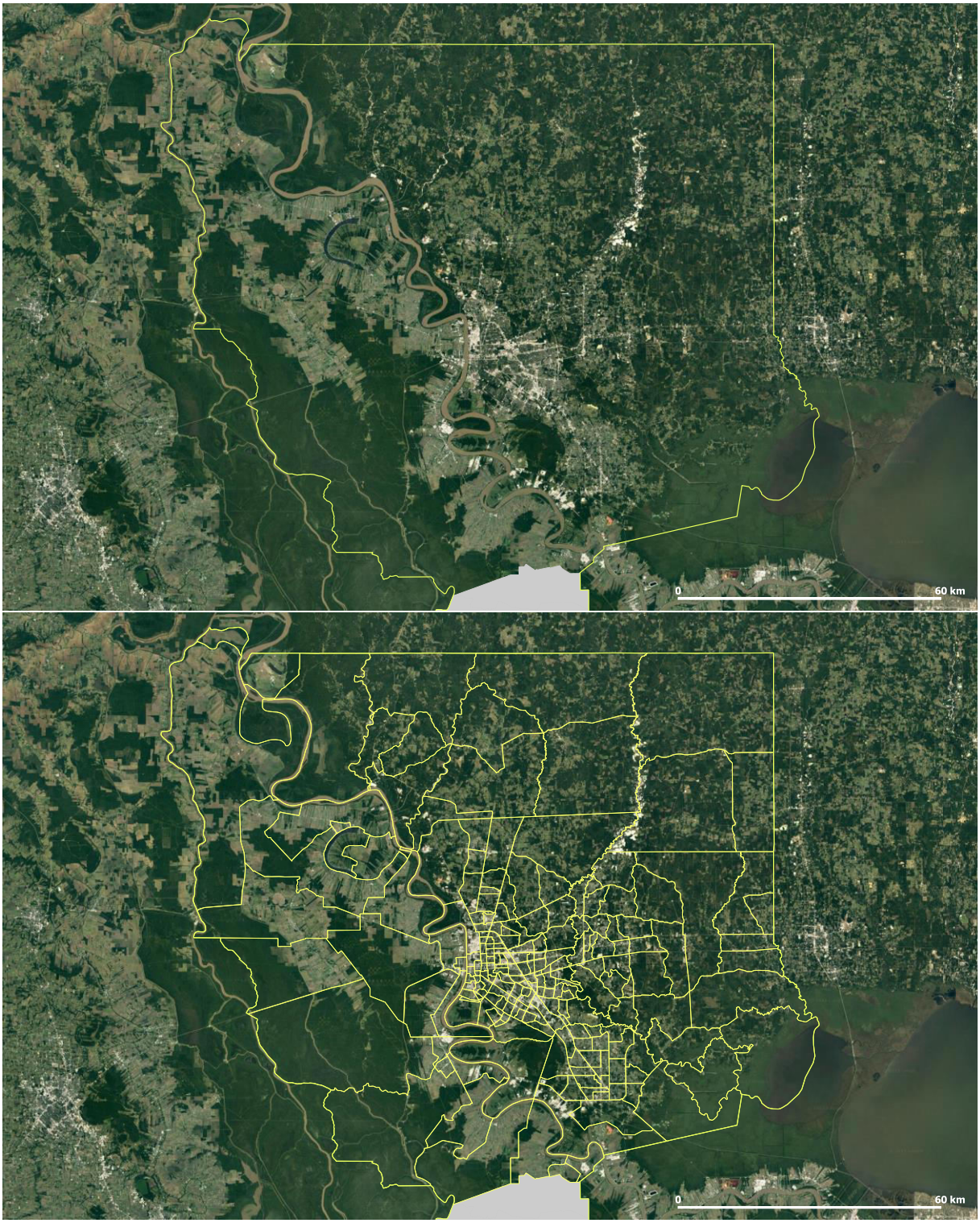


Figure 8: Baton Rouge MSA, Louisiana; MSA border (above) and census tract borders (below); the grey area is part of a larger combined statistical area, but not the MSA itself.

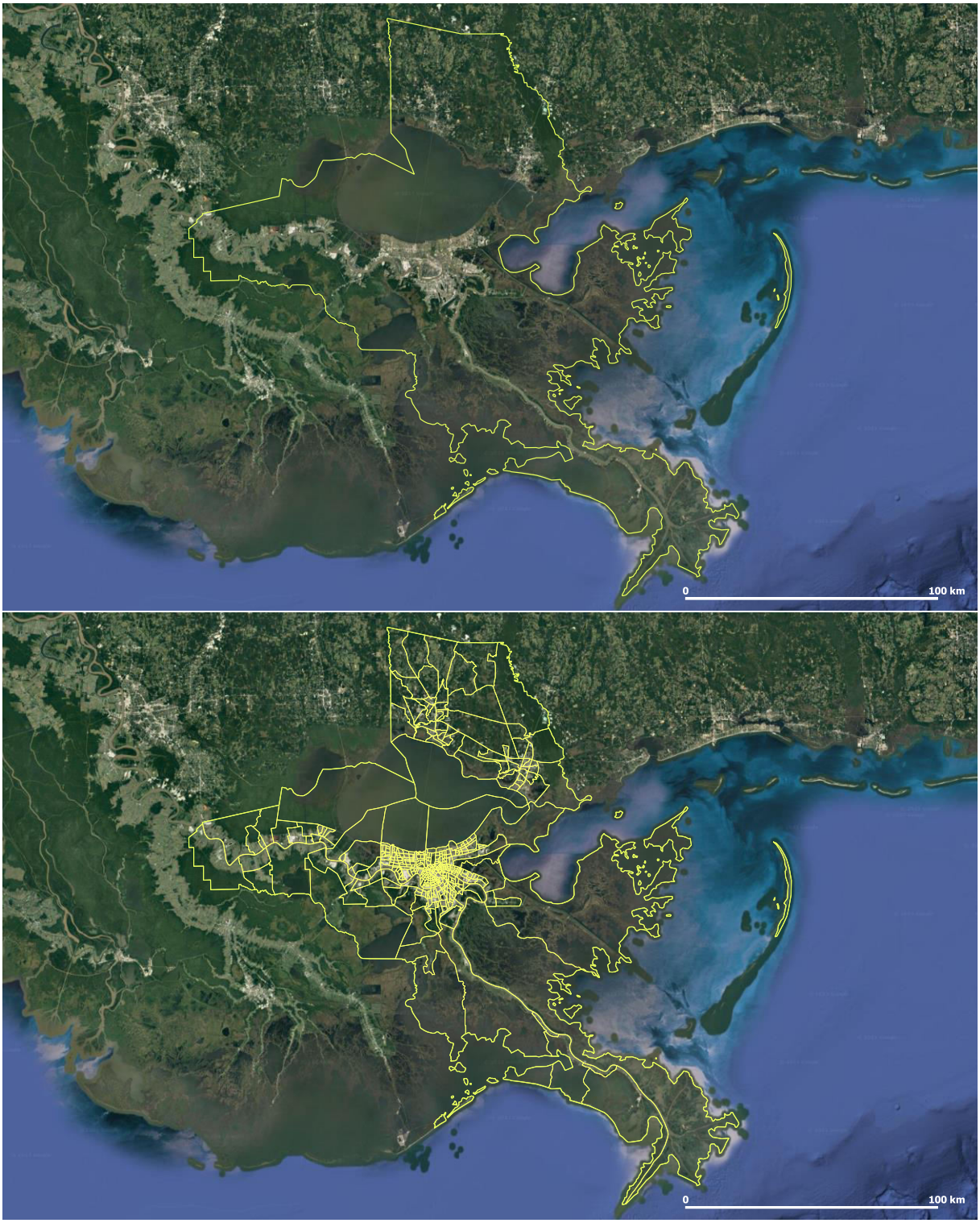


Figure 9: New Orleans-Metairie-Kenner MSA, Louisiana; MSA border (above) and census tract borders (below).

A clear trend can be seen when looking at the five MSAs discussed above. The polycentric structure seems to increase mainly along the first principal component. This suggests that the three main clusters found are monocentric, intermediate, and polycentric from left to right. Another trend that becomes apparent is the increase in the network size, i.e., the total number of nodes within a network along the first principal component. These two patterns will be further investigated in Section 4.2.4.

The two outlier MSAs are shown in the appendix in Figure A.3 and A.4. Both Jacksonville and Sarasota are located in Florida. The two MSA have further similarities in their location as both are located on the coast. Both areas seem to be polycentric, Sarasota more so than Jacksonville as it is more spread along the coast. This is also supported by their urban sprawl index with Jacksonville having a significantly higher value. However, judging from the satellite images alone, no clear feature can be identified that could explain their outlier position within the clustering.

The findings of this subsection can be summarized as follows:

- The degree of polycentricity as well as the overall size of the MSAs seems to increase mainly along the first principal component.
- These results suggest that the clusters are monocentric, intermediate, and polycentric from left to right, with a small outlier cluster that also seems to be polycentric.

4.2.2 Quantitative Analysis

The aforementioned external polycentricity indices in Section 3.4.2 were applied, and the results can be seen in Figure 10.1 and 10.2, respectively. Although some differences between the indices can be seen, they all generally show a similar pattern. The degree of polycentricity increases from left to right mostly along the axis of the first principal component. The index by Green (2007) seems to have a more different pattern with values that are generally more monocentric. Differences between morphological and functional polycentricity are overall very small. Only the index by Burger et al. (2011) shows a lower polycentricity in the functional dimension compared to the morphological one. No significant difference between the social network indices and the rank-size distribution measure can be identified. Both outliers from the clustering are classified as polycentric for the most part. The compact cluster on the very left has the vast majority of monocentric areas, whereas the cluster on the right has mostly polycentric ones. There is a mixed transition space between them that lies primarily within the orange cluster in the middle.

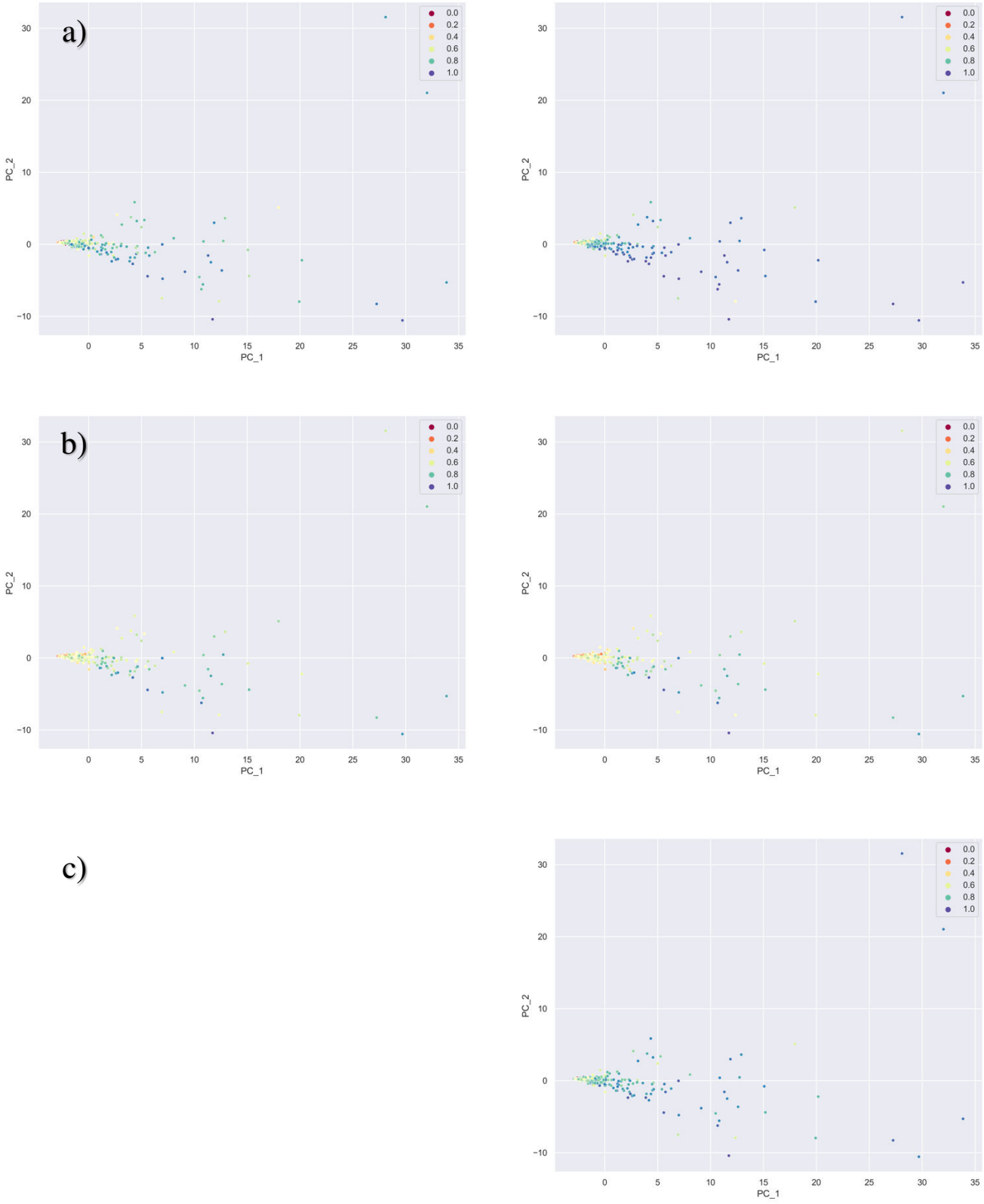


Figure 10.1: Social network measures; morphological polycentricity (left), functional polycentricity (right); a) Burger et al. (2011); b) Green (2007); c) Limtanakool et al. (2007) (only applicable for functional polycentricity).

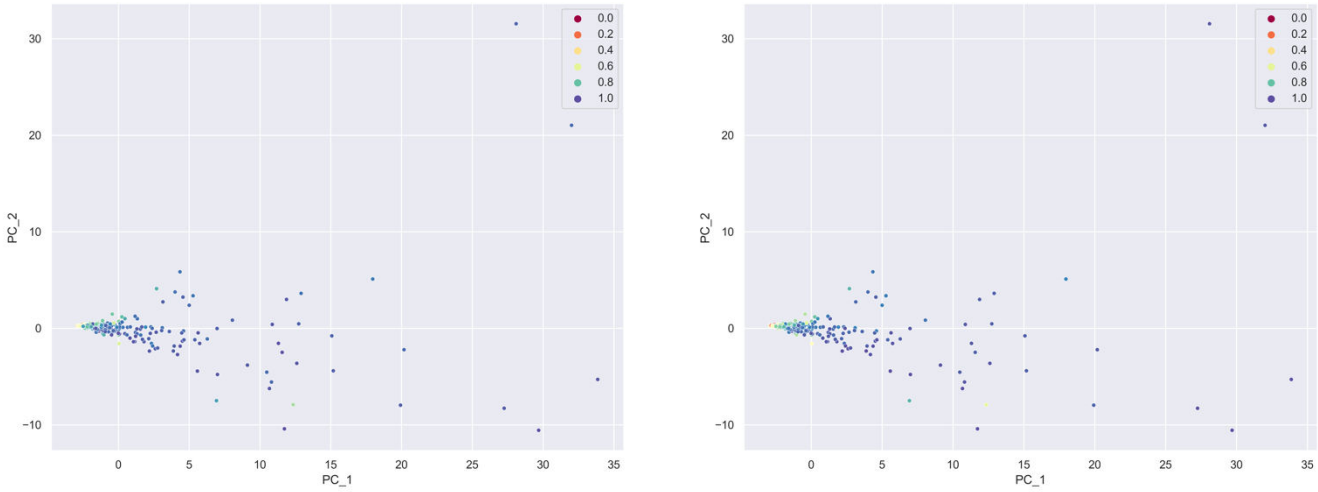


Figure 10.2: Rank-size distribution measure by Burger & Meijers (2012); morphological polycentricity (left), functional polycentricity (right).

The correlation coefficients shown in Table 4 present a similar picture as the visualizations above. The external indices show a medium to high, positive correlation between each other. The morphological and functional measures by Green (2007) and Burger & Meijers (2012) show a very high correlation close to 1. Especially the rank-size distribution index seems to correlate highly with all other indices. Contrary to expectations, not all functional indices correlate better with other functional ones and the same is true in regard to the morphological measures.

Table 4: Pearson correlation between external polycentricity indices. Note: $p < 0.001$ for all values; Suffix F for functional and M for morphological.

	Burger_F	**Burger_M**	**Green_F**	**Green_M**	**Lim_F**	**BM_F**	**BM_M**
Burger_F	1.0	0.681	0.515	0.444	0.861	0.789	0.769
Burger_M	0.681	1.0	0.625	0.61	0.599	0.684	0.709
Green_F	0.515	0.625	1.0	0.986	0.452	0.763	0.732
Green_M	0.444	0.61	0.986	1.0	0.402	0.736	0.709
Lim_F	0.861	0.599	0.452	0.402	1.0	0.694	0.66
BM_F	0.789	0.684	0.763	0.736	0.694	1.0	0.964
BM_M	0.769	0.709	0.732	0.709	0.66	0.964	1.0

The correlation between the external indices and the principal components is shown in Table 5. The principal components have a mostly low correlation to the external indices. The first principal component is the only one that has a higher correlation. It is also the only one that shows a positive correlation. The p-value of the fourth and ninth principal components are too high for the results to be considered significant. The first, third, and sixth principal components are the only ones that show a higher and more significant correlation overall. Some components like PC 5 or PC 2 seem to capture certain indices better than others.

Especially the second component shows low correlation coefficients, despite retaining around 12 % of the variance.

*Table 5: Pearson correlation between external indices and principal components. Note: * p < 0.05, ** p < 0.01, *** p < 0.001; Suffix F for functional and M for morphological.*

	Burger_F	**Burger_M**	**Green_F**	**Green_M**	**Lim_F**	**BM_F**	**BM_M**
PC_1	0.463 ***	0.388 ***	0.507 ***	0.479 ***	0.4 ***	0.422 ***	0.396 ***
PC_2	-0.116 *	-0.108 *	-0.175 ***	-0.166 **	-0.075	-0.075	-0.08
PC_3	-0.322 ***	-0.263 ***	-0.344 ***	-0.316 ***	-0.32 ***	-0.304 ***	-0.281 ***
PC_4	-0.099	-0.078	-0.098	-0.082	-0.075	-0.078	-0.066
PC_5	-0.167 **	-0.073	-0.078	-0.066	-0.083	-0.212 ***	-0.201 ***
PC_6	-0.292 ***	-0.224 ***	-0.289 ***	-0.26 ***	-0.244 ***	-0.289 ***	-0.274***
PC_7	-0.132 **	-0.143 **	-0.151 **	-0.144 **	-0.108 *	-0.163 **	-0.166 **
PC_8	-0.155 **	-0.145 **	-0.117 *	-0.104 *	-0.108 *	-0.151 **	-0.145 **
PC_9	-0.013	-0.027	-0.009	-0.007	-0.003	-0.015	-0.014

The results of this subsection can be summarized as follows:

- All external polycentricity indices depict an increasing degree of polycentricity mainly along the first principal component, confirming the findings of Section 4.2.1.
- Whereas the external indices have a medium to high correlation with each other, only the first principal component shows a higher positive level of correlation with the external measures.

4.2.3 Temporal Change

The results of the temporal analysis are shown in Figure 11. Due to the results from the previous two sections, new aliases have been assigned to the groups: *Intermediate* for Group 1 (orange), *Monocentric* for Group 2 (blue), *Outlier* for Group 3 (red), and *Polycentric* for Group 4 (green). For the temporal analysis, the outlier cluster has been excluded, due to its small size. The results show a general movement towards the left, which is the same for all clusters, although in different relations. The higher the value of the first principal component, the larger is the change in distance between the two time periods. The centroid of the polycentric cluster (green) shows the largest change with an Euclidean distance of 13.266. The intermediate cluster (orange) follows with a distance of 3.470, and the monocentric cluster (blue) with a distance of 0.299 comes in last. This suggests that MSAs with a higher degree of polycentricity experienced a larger change.

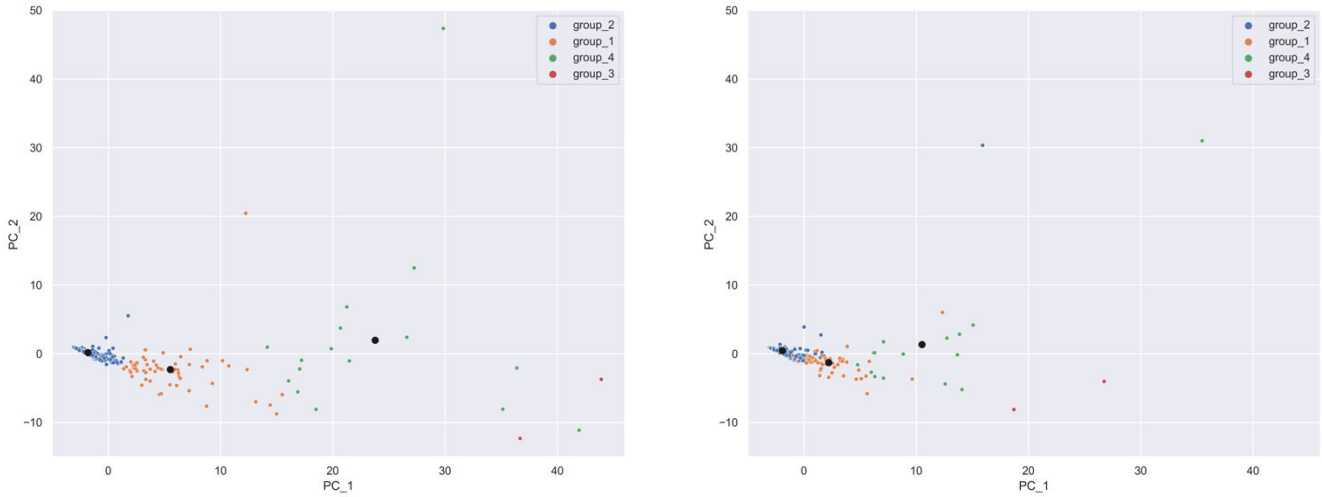


Figure 11: PCA embedded vectors before (left) and during (right) mobility restrictions.

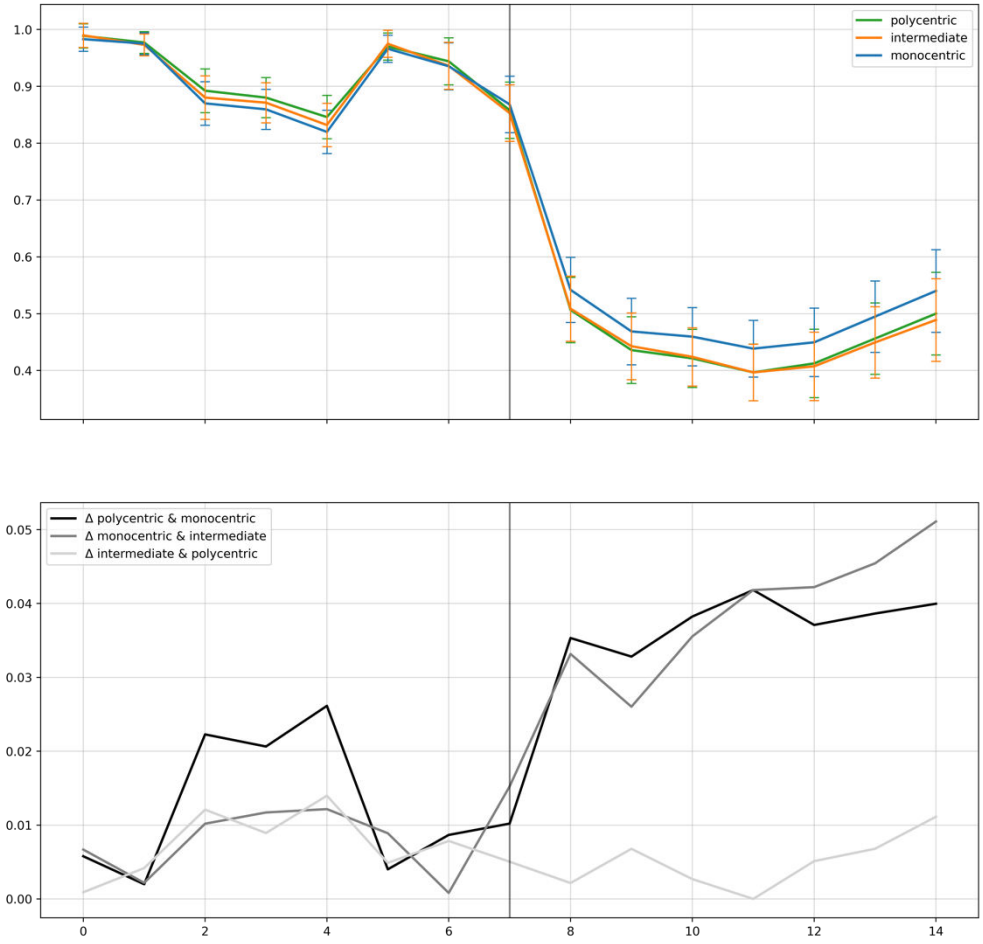


Figure 12: Mobility reduction per cluster with standard deviation (above) and difference between clusters (below).

When inspecting the mean mobility reduction within each cluster in Figure 12, the overall pattern looks very similar. A small reduction before the regulations with a following recovery is apparent. The main reduction after regulations have been implemented is severe for all cluster types. The beginning of a recovery from the regulations is also visible for all groups. However, some differences can be seen towards the end of the second period. The difference between the monocentric and polycentric clusters reaches 4 %, whereas most change occurs rapidly around the transition week. The difference between the monocentric and intermediate clusters reaches just above 5 %, with a very similar pattern as the polycentric cluster. The gap between polycentric and intermediate clusters is low throughout the whole time period and stays under 1 % for the most part. The spread of all three clusters is very similar and shows a general increase when the mean value drops. The mean values always stay within the spread of the other clusters, indicating that the three clusters may not be significantly different from each other. To assess whether the mean relative reduction is statistically significant, an independent t-test was performed. The difference between polycentric and monocentric ($p < 0.05$), as well as the difference between intermediate and monocentric clusters ($p < 0.001$) are significant. Only the difference between the polycentric and intermediate clusters is not significant ($p > 0.95$).

The findings of this subsection can be shortly summarised as follows:

- The overall degree of polycentricity was reduced among all clusters after mobility restrictions were implemented, whereas the polycentric cluster experienced the largest change.

- In regard to the relative mobility reduction, the polycentric and intermediate cluster show a similarly higher reduction of 4 to 5 % compared to the monocentric cluster.

4.2.4 Further Exploratory Analysis

In order to explain the findings of the earlier sections within this chapter, linear regression is estimated between the first principal component and some other variables. Some of these variables will also be used in the regression analysis in Section 4.4. The results of these plots can be seen in Figure 13, whereas all relations are significant ($p < 0.01$). Both the total number of nodes and the mean commuting time clearly increase along with the first principal component. These two spatial variables are an estimation of the total size of the network, meaning that the networks generally grow larger in size along the first principal component. This confirms the findings of Section 4.2.1. The same statement can be made about employment density. Lastly, a weaker yet positive relation can be seen between the cumulative COVID-19 cases[10] and the first principal component.

[10] https://github.com/nytimes/covid-19-data

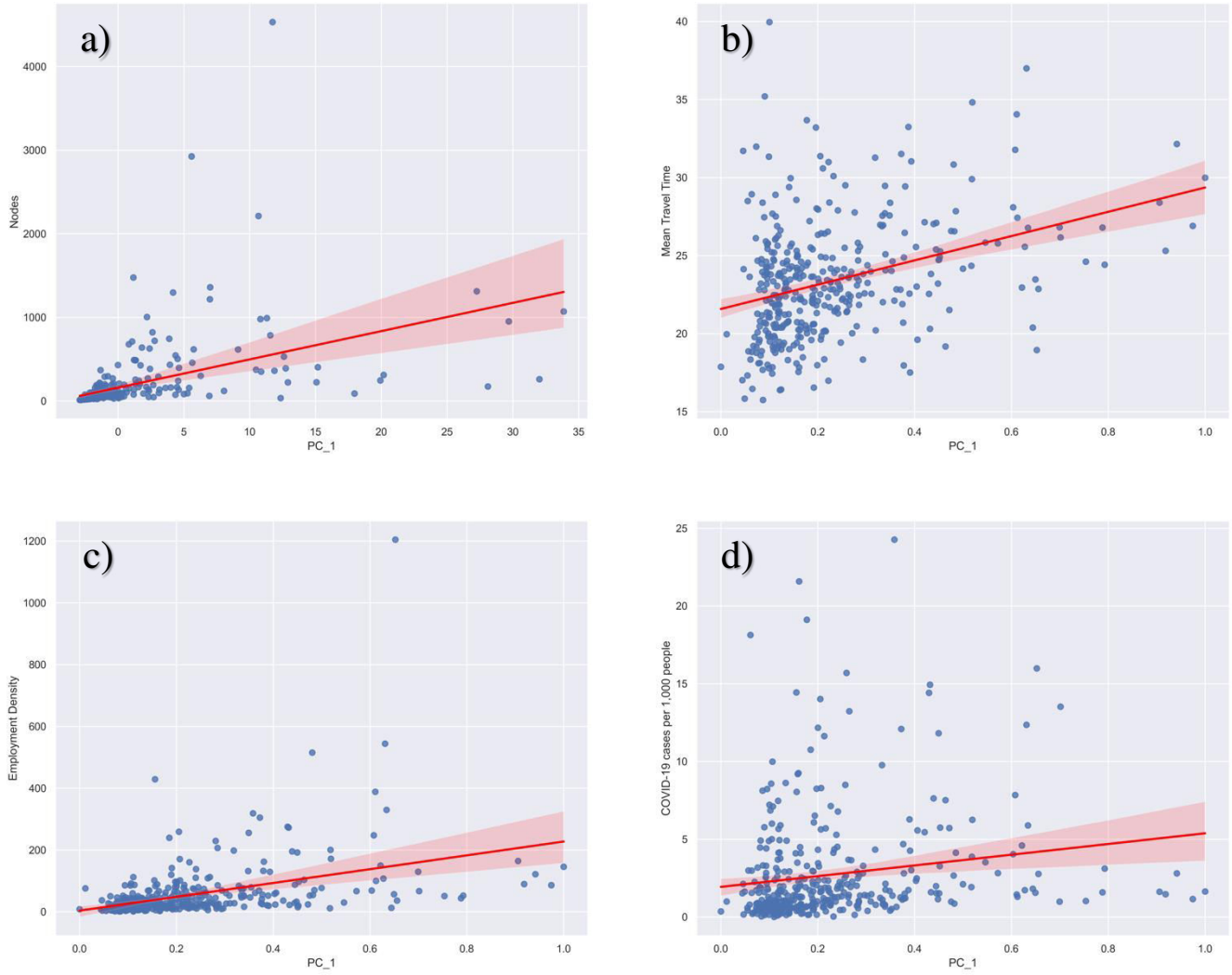


Figure 13: Estimated linear regression between PC_1 and a) total number of nodes, b) mean commuting time, c) employment density and d) cumulative COVID-19 cases per 1'000 people.

4.3 Impact of Spatial Structure

For a better understanding of the importance of urban spatial structure on mobility reduction, a regression analysis has been conducted. After the best performing algorithm has been selected, the Whole Graph Embedding (WGE) model is compared with a baseline model to assess the impact of spatial structure on mobility reduction. The models are further interpreted by comparing the correlation and feature importance.

4.3.1 Model Selection

Among the ensemble methods tested for both WGE and base model, Gradient Boosting (GB) performed best overall. The performance metrics of each method are shown in Table 6. AdaBoost (AB) generally performed worse than the other two methods. Although Random Forest (RF) performed better than GB for the WGE model, it performed worse for the base model. Based on these results, the GB will be analyzed in more detail in the following sections.

Table 6: Model selection. Note: The content in brackets shows the best hyperparameters used.

	WGE Model			Base Model		
	RF (128, 1.0)	GB (64, 1.0)	AB (256, 0.3)	RF (100, sqrt)	GB (30, 1.0)	AB (30, 0.2)
R^2	0.4336 ± 0.0215	0.4150 ± 0.0241	0.3853 ± 0.0176	0.3465 ± 0.0176	0.3694 ± 0.018	0.329 ± 0.0222
MSE	0.0163 ± 0.0004	0.0166 ± 0.0004	0.0176 ± 0.0004	0.019 ± 0.0002	0.0183 ± 0.0004	0.0194 ± 0.0003
MAE	0.1 ± 0.001	0.1013 ± 0.0013	0.1061 ± 0.0015	0.1028 ± 0.0008	0.1018 ± 0.0013	0.1063 ± 0.0015
RMSE	0.1262 ± 0.0015	0.1279 ± 0.0014	0.1316 ± 0.0015	0.1359 ± 0.0012	0.1336 ± 0.0013	0.1377 ± 0.0017

4.3.2 Model Comparison

The WGE model outperformed the base model, both with and without graphic node features, as demonstrated by all selected performance metrics in Table 7. The improvement in performance is particularly evident in the R^2 score, with an increase of 12 %, respectively 25 % when comparing the models without any graphic node features. The other three metrics (MSE, MAE, and RMSE) are consistently correlated with the R^2 score, although to a lesser extent. Especially the MAE displays marginal improvements compared to the enhancements observed in the R^2 score. The inclusion of graphic features significantly enhanced the performance of the base model. In contrast, the WGE model experiences a reduction in performance when graphic node features are included.

Table 7: Model comparison. Note: The content in brackets shows the best hyperparameters used.

	WGE model (64, 1.0)	**Base model** (30, 1.0)	**Non-graphic WGE model** (64, sqrt)	**Non-graphic Base model** (30, sqrt)
R^2	0.4150 ± 0.0241	0.3694 ± 0.0180	0.4467 ± 0.0252	0.3570 ± 0.0186
MSE	0.0166 ± 0.0004	0.0183 ± 0.0004	0.0158 ± 0.0005	0.0187 ± 0.0003
MAE	0.1013 ± 0.0013	0.1018 ± 0.0013	0.1000 ± 0.0016	0.1024 ± 0.0007
RMSE	0.1279 ± 0.0014	0.1336 ± 0.0013	0.1246 ± 0.0021	0.1349 ± 0.0014

A possible explanation for this could be that whole graph embedding models the graphic features implicitly. It could also be that the spatial structure modeled by the WGE has a high correlation with the graphic node features, making them irrelevant to the model. This second hypothesis can be tested by inspecting the relation between the pairwise distance as a spatial indicator and the graphic node features. The sum of pairwise distance is calculated for a sample of areas and their correlation with the graphic features is examined. Four MSAs have been chosen at random with an increasing number of nodes. The results can be seen in Figure 14.1 and 14.2. The plots suggest no clear relation between the graphic node features and the pairwise distance across all tested MSAs.

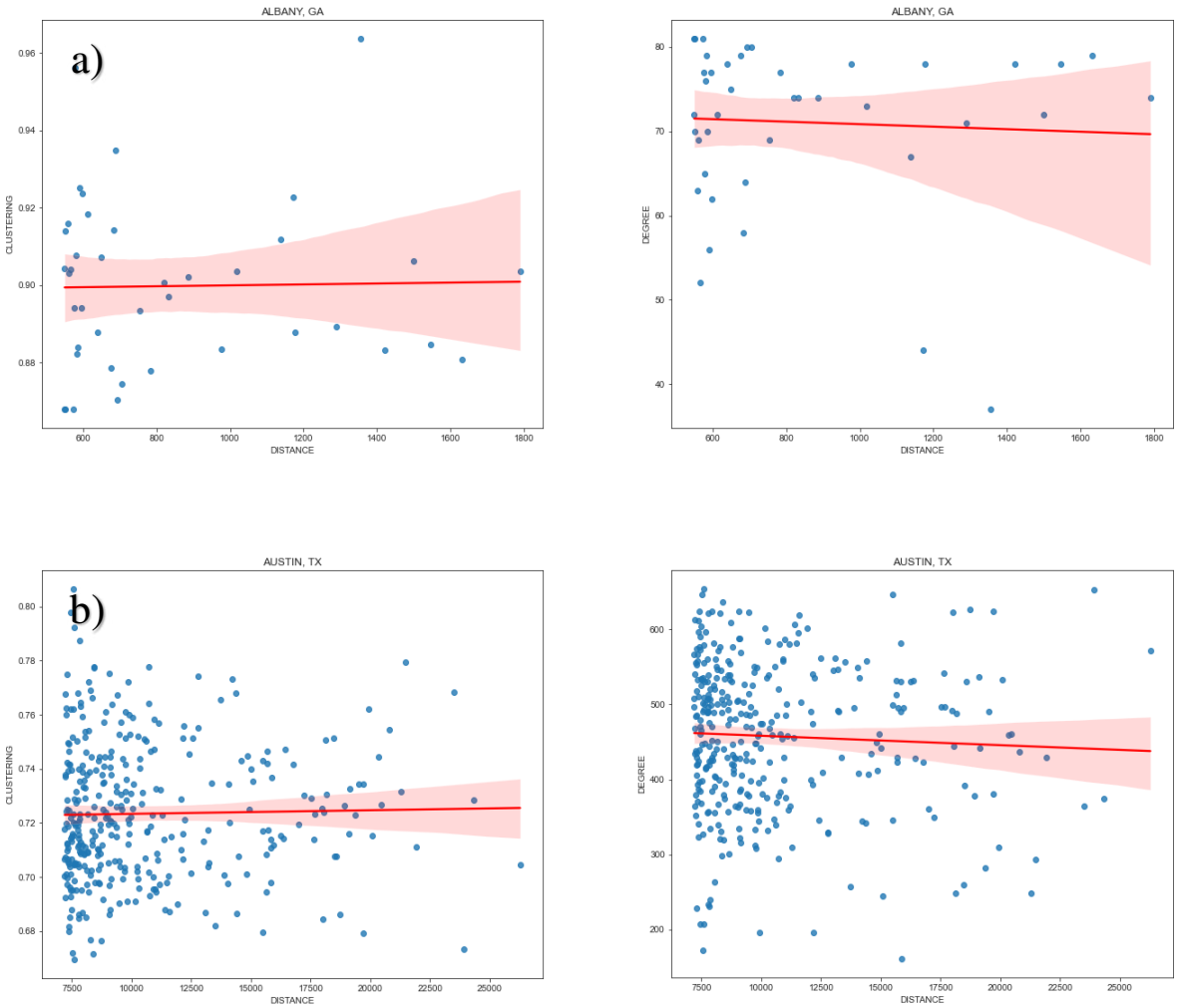


Figure 14.1: Graphic node features plotted against pairwise distance; a) Albany, GA; b) Austin, TX.

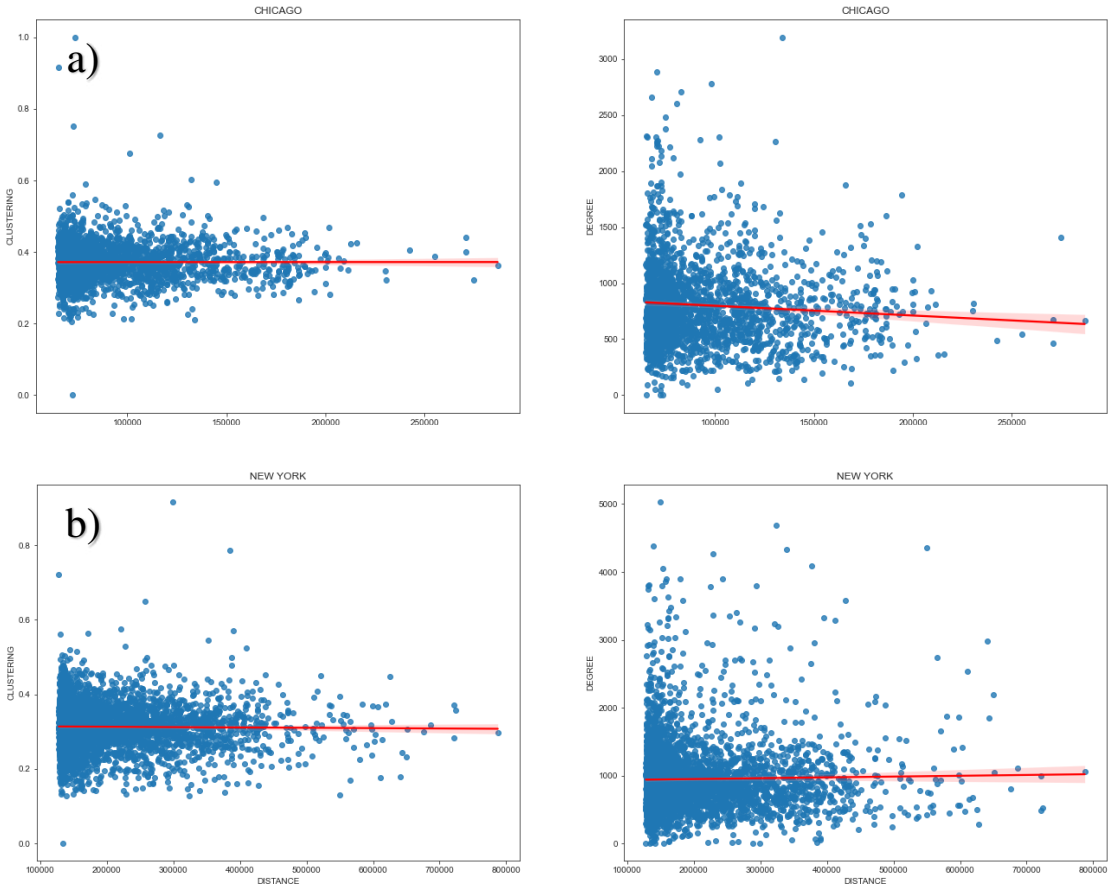


Figure 14.2: Graphic node features plotted against pairwise distance; a) Chicago, IL; b) New York, NY.

This subsection can be summarised in the following two points:

- Spatial structure as represented by the WGE leads to a model performance increase of 12 to 25 % in terms of the R^2 score.
- The WGE model performs better without graphic node features, which may be due to its ability to also model graphic node features implicitly.

4.3.3 Model Interpretation

The Pearson correlation coefficients of relative mobility reduction and node/edge features for both models can be seen in Table 8. The values for the WGE model are aggregated and thus have no single p-value. The values are generally higher for the base model with the mobility flows showing the highest correlation.

*Table 8: Pearson correlation coefficients of relative mobility reduction and node/edge features. Note: *** $p < 0.001$, values of the WGE model are aggregated.*

Node features	WGE model	Base model
Inflow	0.164	-0.174 ***
Self-loop	0.042	-0.279 ***
Clustering	0.364	-0.464 ***
Employment	0.096	0.302 ***
Degree	0.178	0.403 ***
Flows	-	-0.447 ***

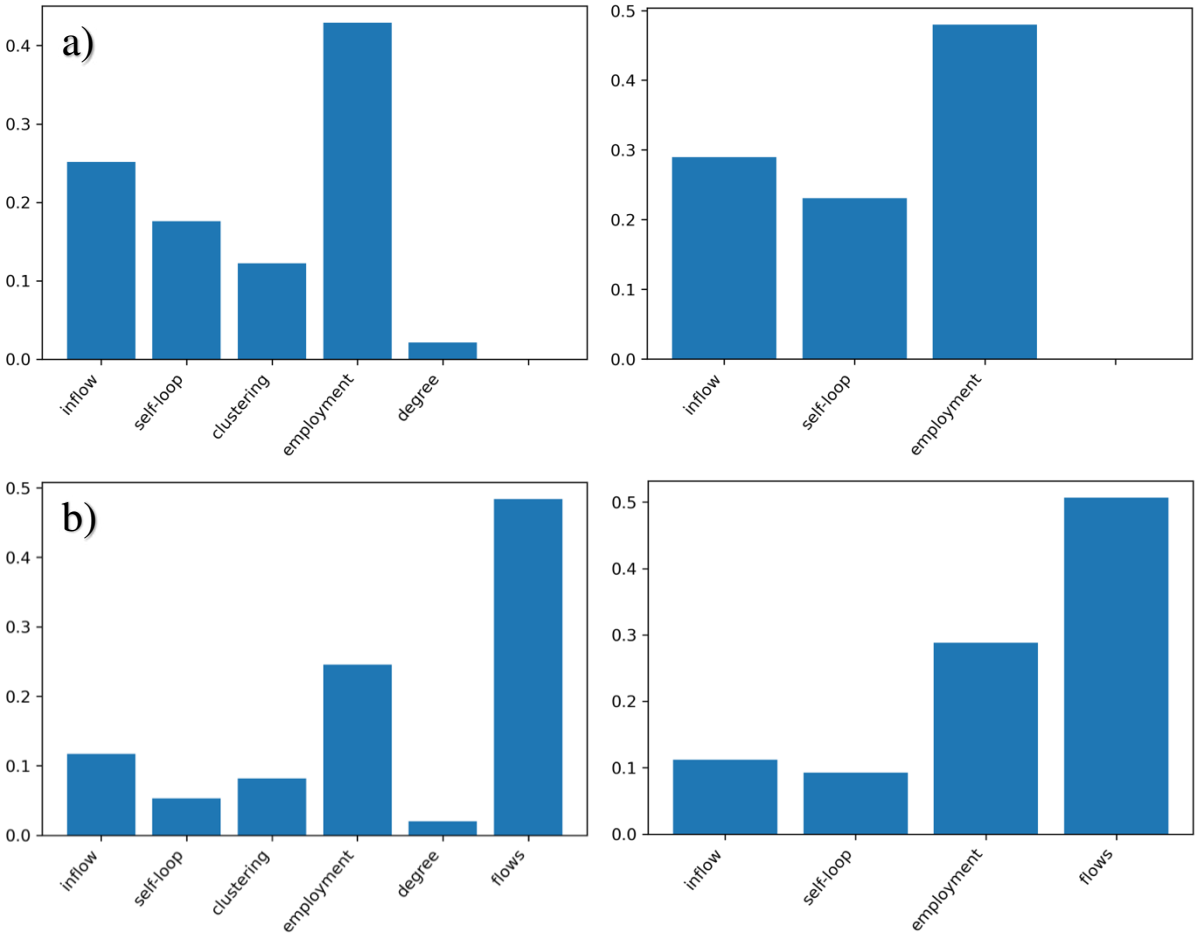


Figure 15: Feature importance of the WGE model (a) and the Base model (b).

The feature importance of both models can be seen in Figure 15. The feature importance of the WGE model is aggregated and displays the sum of 64 values from the scattering features. For the WGE model, employment is the most important feature, followed by inflow, self-loop, clustering, and degree. This pattern remains unchanged when excluding graphic node features. For the base model, flows are most important, followed by employment, inflow, clustering, self-loop, and degree. Again this remains unchanged for the base model when excluding graphic node features. The pattern for employment, inflow, and self-loop remains the same for the most part across all models. The clustering is more important than self-loop in the base model and degree has a low importance across all models.

Table 9: Individual WGE feature importance of the top 61 features.

	Inflow	**Self-loop**	**Clustering**	**Employment**	**Degree**
Total	16	16	10	17	2
Mean	0.0135	0.0089	0.0096	0.0236	0.0048
Median	0.0117	0.0049	0.0073	0.0122	0.0048

The individual feature importances of the WGE are also examined for any specific pattern and to compare with the aggregated values shown before (see Table 9). The feature importances are sorted descending and the cumulative sum shows an inflection point at 61, meaning that the top 61 values account for the majority of the total feature importance. The results confirm the findings of the aggregated feature importance for the most part. Degree has only two values and a low mean and median. Employment is both in total and average numbers the most important, followed by inflow in second place. Although clustering has fewer values than self-loop, its mean and median are higher. Self-loop has many values overall, but a lot of them are comparatively low.

4.4 Impact of Polycentricity

4.4.1 Model Selection

For the second regression analysis, the Random Forest model performed best, followed by the Gradient Boosting and the AdaBoost model. The results are shown in Table 10 below. Thus, the RF model has been used for further analysis in the following section.

Table 10: Model selection. Note: The content in brackets shows the best hyperparameters used.

	RF (90, 1.0)	**GB** (70, sqrt)	**AB** (50, 0.3)
R^2	0.4963 ± 0.0232	0.4670 ± 0.0265	0.4267 ± 0.0138
MSE	0.0143 ± 0.0002	0.0152 ± 0.0004	0.0165 ± 0.0002
MAE	0.0894 ± 0.0008	0.0916 ± 0.0017	0.0986 ± 0.0011
RMSE	0.1178 ± 0.0008	0.1214 ± 0.0018	0.1266 ± 0.0007

4.4.2 Model Interpretation

For a first impression of the variables used in the regression analysis, their correlations with mobility reduction are inspected. The Pearson correlation coefficients can be seen in Table 11. When looking at the PCs, only 5 out of 9 values can be considered significant. The overall correlation of these is low, with PC 1 having the highest value of 0.257. This is also quite low when compared to other variables like education, public transportation, employment density, or the Asian population.

The results from the correlation analysis are complemented by the feature importance shown in Figure 16. Education is by far the most important, with above 30 % model importance. The

three variables—public transportation rate, employment density, and the Asian population—also hold comparatively high importance. Most variables besides the ones mentioned already have a low impact on the model. This also includes the nine principal components. Interestingly, the first principal component has a lower importance compared to the other PCs. After summing up, all PCs take almost 15 % of the total model importance. The results seem in line with the correlation values, except for the dominant political party, which has low importance, despite having a somewhat high correlation.

*Table 11: Pearson correlation coefficients between relative mobility reduction and various variables. Note: * $p < 0.05$,** $p < 0.01$,*** $p < 0.001$.*

Variables	**Pearson Coefficient**	**PC's**	**Pearson Coefficient**
Median age	-0.122 *	PC_1	0.257 ***
Sex ratio	-0.069	PC_2	-0.18 ***
Black or African American	-0.153 **	PC_3	0.213 ***
Asian	0.469 ***	PC_4	-0.102 *
Hispanic	0.08	PC_5	0.047
Unemployment rate	-0.188 ***	PC_6	-0.175 ***
Poverty rate	-0.205 ***	PC_7	-0.016
Gini index	0.2 ***	PC_8	-0.029
Higher education	0.615 ***	PC_9	-0.078
Party	0.47 ***		
Mean travel time	0.234 ***		
Public transportation rate	0.562 ***		
Employment density	0.484 ***		

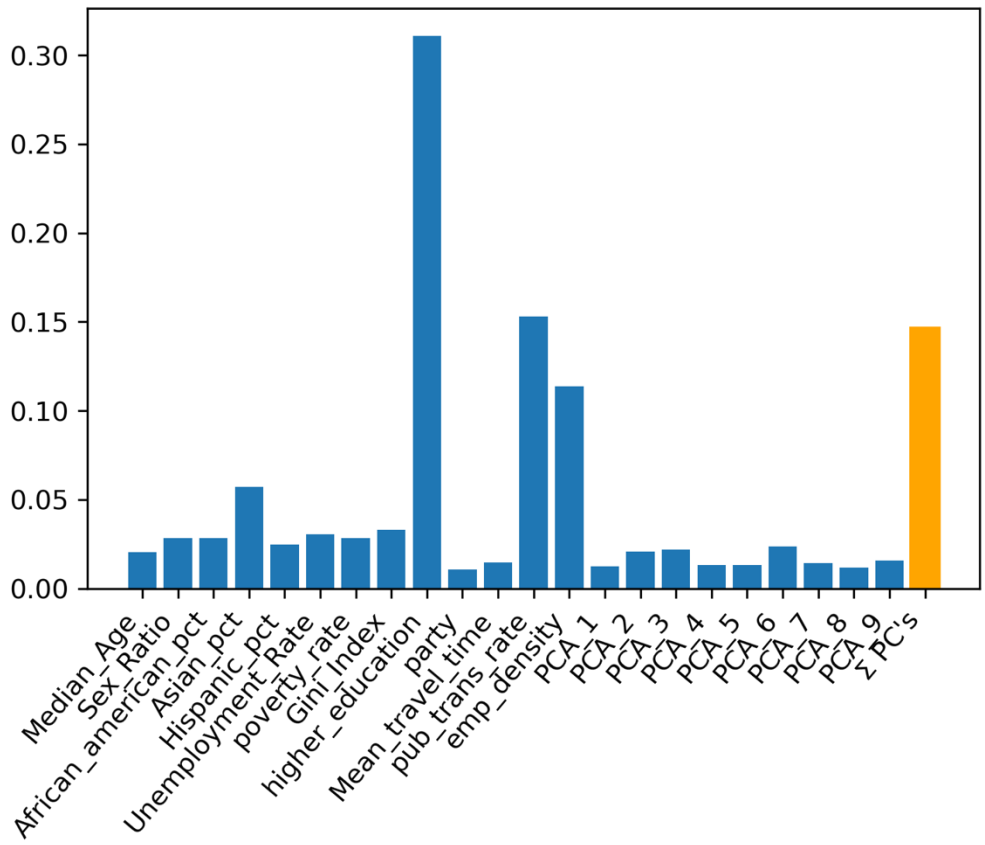


Figure 16: Feature importance of the RF model. Orange: sum of importance for all principal components.

The subsection can be shortly summarized as follows:

- The impact of individual principal components is limited on the model. Other variables such as education, public transportation rate, employment density, or the Asian population are dominant.
- However, the general spatial structure, as presented by the sum of the principal components' importance, is among those dominant variables.

4.5 Summary of Results

The main results can be summarized in the following points:

- Three main clusters (Group 2, Group 1, and Group 4) and one outlier cluster (Group 3) have been identified, with the majority of MSAs being within only one cluster.
- The qualitative and quantitative analysis suggests that the three main clusters are monocentric, intermediate, and polycentric. The outlier cluster also seems to be polycentric.
- The temporal analysis shows a larger change for cities with an increased degree of polycentricity. MSAs with higher polycentricity experienced a larger relative mobility reduction throughout the time period analyzed.

- By comparing a WGE model with a base model, the impact of spatial structure on mobility reduction could be quantified. The spatial structure brings forth a model improvement of around 12 – 25 % in terms of R^2 score.
- The overall impact of the individual principal components is low in the chosen Random Forest regression. Education is by far the most impactful variable, followed by the public transportation rate and employment density. However, the principal components added together amount to about 15 % of the overall importance, in alignment with other important variables.

Chapter 5: Discussion

This chapter covers the interpretation of the research findings, offering a comprehensive analysis of the results in the context of the research questions. By examining the outcomes and comparisons to existing literature, this chapter aims to provide insights and contributions for a broader understanding of the research area. The clusters of the embeddings and differences in mobility behavior are discussed in Section 5.1. The impact of spatial structure is covered in Section 5.2, and the impact of polycentricity in Section 5.3. Lastly, in Section 5.4 possible limitations of this study are addressed. As a reminder, the research questions, as well as the hypotheses, are restated at the beginning of each subsection.

5.1 Whole Graph Embedding Clusters

RQ1: To what extent does the degree of polycentricity influence mobility reduction during the COVID-19 pandemic in U.S. metropolitan areas, and how did it change after the implementation of mobility regulations?

Hypothesis 1: Assuming that an increased risk of virus spread leads to an increased regulatory mobility reduction, monocentric areas are expected to experience higher reduction rates than their polycentric counterparts. The high concentration of transportation and people in the center within monocentric areas increases the risk of infection (Aguilar et al., 2020). Conversely, decentralized polycentric areas with less crowded spaces are anticipated to show lower reduction rates due to reduced infection risk.

Hypothesis 2: After the implementation of mobility restrictions to contain the virus, urban areas are expected to become more functionally polycentric. In order to halt the spread of the virus, the high mobility concentrations in central areas need to be restricted. This may lead to a shift in mobility away from the center, resulting in a more functionally polycentric structure than before.

The four clusters identified from the whole graph embedding could be appointed to a monocentric, intermediate, polycentric, and outlier cluster. The majority (≈ 80 %) of Metropolitan Statistical Areas (MSAs) are included within the monocentric cluster, while polycentric areas only account for around 5 % and the intermediate for the remaining 15 %. Arribas-Bel & Sanz-Gracia (2014) also report that the majority of the 359 MSAs they examined had a monocentric structure. However, there are other studies that report the opposite, with polycentricity being the most found urban form within U.S. MSAs (Hajrasouliha and Hamidi, 2017). Most other studies only focus on a selected few well-known cities to measure the polycentricity, which makes a direct comparison difficult.

Nonetheless, the qualitative analysis shows that the embeddings can differentiate MSAs of varying degrees of polycentricity. This was confirmed quantitatively by comparing the principal components (PCs) of the scattering features with other conventional polycentricity indices, which measure either only the functional or the morphological dimension of polycentricity. When looking at the correlation coefficient between these conventional polycentricity indices and the principal components, it becomes apparent that the first principal component has a higher correlation. It is also the only one with a positive correlation, while all others have a low to very low negative correlation. That being said, the first PC does not seem to be a duplicate of any of the existing indices as its unique correlation

pattern indicates. The first PC arguably includes both functional and morphological dimensions, as the level of correlation is not higher for either the functional nor the morphological indices. Interestingly, the second principal component, which accounts for a relatively large amount of the total variance, has correlations close to zero with all other indices. This component apparently captures something important that is not explained by the conventional indices.

The analysis of cluster-specific behavior between the two different time periods yielded unexpected results. Against the first hypothesis stated, the monocentric MSAs experienced the smallest relative change. Although minor, the relative mobility reduction was 4 to 5 % larger in the polycentric and intermediate clusters. The difference between polycentric and intermediate clusters was statistically not significant ($p > 0.95$), indicating no real difference in their mobility behavior. These findings are in conflict with the results from Aguilar et al. (2020), who report that the spread of an epidemic, as well as the response to mobility restrictions, is much higher in monocentric cities, where flows are highly concentrated. A couple of reasons can be named to try to explain these different findings. The results from Aguilar et al. (2020) are on a smaller city level, compared to the wider scale of MSAs used in this study. They further use a fix sized grid to measure mobility, whereas this study uses census tracts of varying sizes that inflate densely populated areas. For these reasons, a direct comparison of the findings is not feasible. Polycentric MSAs have also been found to be generally denser in employment and larger than their monocentric counterparts, which matches the findings by Arribas-Bel & Sanz-Gracia (2014). Polycentric areas further showed a higher number of cumulative COVID-19 cases per 1'000 persons. This may explain the higher reduction compared to monocentric MSA. More people interacting in a small space is an ideal situation for the spread of a virus. In order to keep infection rates in check, polycentric areas had to implement overall more restrictive mobility regulations, whereas monocentric areas could concentrate their efforts in the main CBD. However, the relationship between COVID-19 cases and the first principal component is not isolated, and many more variables could potentially be more important.

The temporal change in terms of the degree of polycentricity also acted against the second hypothesis. The movement of the clusters indicates a general trend towards a more monocentric structure rather than becoming more polycentric. This movement is larger for MSAs with a higher degree of polycentricity, meaning that polycentric areas experienced a larger change in their structure. Perhaps general mobility was reduced more than commuting, which can be altered less freely depending on the working conditions. That being the case, the overall pattern would become more monocentric, since commuting flows are more focused towards the employment center.

As for the first research question: A higher degree of polycentricity could be associated with a higher relative mobility reduction in U.S. metropolitan areas during the COVID-19 pandemic. Both the polycentric and intermediate clusters showed higher reduction values than the monocentric cluster. The overall spatial structure of MSAs, as measured by whole graph embeddings, has become more monocentric after mobility restrictions were implemented. These structural changes were also higher with an increased degree of polycentricity.

5.2 Impact of Spatial Structure

RQ2: How important is the general spatial structure of U.S. metropolitan areas, as represented by whole graph embeddings, for mobility reduction during the COVID-19 pandemic?

Hypothesis 3: Spatial structure greatly dictates the course of transportation routes for a given city, which therefore is essential for any mobility flows. As a consequence, spatial structure significantly influenced mobility reduction during the COVID-19 pandemic.

The Whole Graph Embedding (WGE) model shows an overall performance improvement of 12 – 25 % in terms of R^2 score, compared to the base model. This is a significant improvement that reflects the importance of the general spatial structure for mobility reduction. Although the WGE model performed substantially better than the base model, it has to be mentioned that the values for the base model have been largely aggregated. Whereas the WGE uses data on census tract level, the base model uses average values for the whole MSA. It is possible that the performance improvement is influenced by the finer granularity of the input data.

The WGE model performs better without graphic node features, which can be explained by its ability to model them implicitly. Another possibility is that the spatial structure has a high correlation with these graphic features and makes them obsolete for the regression model. The second possibility was tested by estimating a linear regression between the graphic features and the pairwise distance for a few selected locations. Since there was no clear relation between them, it has to be assumed that the WGE does model the graphic node features implicitly.

The grouped feature importance of all 9 principal components in Figure 16 presents the overall importance of the spatial structure as obtained from the whole graph embedding. This includes spatial dimensions other than polycentricity that we may not know yet. These added values account for up to 15 % of total model importance, ranking among the most impactful variables included in the model. Although not the most important variable, the general spatial structure had a significant importance for the model used, thus confirming the third hypothesis.

To answer the second research question: The general spatial structure of U.S. metropolitan areas, as presented by whole graph embeddings, had a significant impact on mobility reduction during the COVID-19 pandemic. By comparing a WGE model that includes spatial structure to a base model that does not, performance improvements of 12 to 25 % in terms of R^2 score have been achieved. Further investigation of the feature importance in the second regression analysis provided an overall feature importance of around 15 %. These are significant values that indicate the value of spatial structure in modeling mobility reduction.

5.3 Impact of Polycentricity

RQ3: How important is the degree of polycentricity for mobility reduction during the COVID-19 pandemic in U.S. metropolitan areas, compared to other socio-demographic variables such as ethnicity, education, or employment?

Hypothesis 4: Polycentric areas exhibit a lower risk of viral transmission in comparison to monocentric areas (Aguilar et al., 2020). Therefore, the degree of polycentricity is expected to

be a good indicator of the overall mobility reduction and will be among the most important variables in the regression model.

The Random Forest regression had overall the best R^2 score of the three tested algorithms, with a value close to 0.5. Although the R^2 score is not very high, it can explain a significant part of overall mobility reduction. After all, the goal was not to predict the reduction as best as possible, but to be able to explain the impact of polycentricity compared to other variables.

The feature importance of the model is clearly dominated by education, followed by the public transportation rate, employment density, and the percentage of the Asian population. The dominance of education in both correlation coefficient and feature importance is surprising. Although proven to influence mobility in other studies such as in Zhu et al. (2021), its importance was mediocre. Perhaps higher-educated people were more aware of the pandemic threat and acted accordingly. Education at higher levels can also be taught online with video conference platforms, which gained high popularity during the pandemic. Lastly, higher educated people usually have occupations that do not require much manual labor and can easier be done via home office.

Public transportation has a very crucial role in the spread of an airborne disease such as COVID-19. They provide optimal conditions for infections due to low airflow and high accumulation of people in small spaces. Whenever possible, people would try to avoid these crowded spaces of public transportation and switch to alternative means such as working from home. Moreover, many cities reduced or even closed their public transportation services as a reaction to increasing infection rates.

The gross employment density was also among the most important features in the study by Zhu et al. (2021). Commuting is one of the most common drivers of mobility, whereby workers travel to work and back home on a daily basis. It follows that the higher the employment density in a given place, the higher will be commuting flows in that area. Due to the pandemic, a lot of people switched to work from home, completely circumventing the need to leave the house. Additionally, the rise in unemployment may have reinforced the reduction in commuting. It is therefore not surprising that employment density has a high feature importance.

In regard to ethnicity, both the Hispanic and African-American population do not seem to have influenced the mobility reduction. The Asian population on the other side correlates clearly with an increased reduction and has an overall high feature importance for the regression model.

The 9 principal components in the model have individually a low feature importance. The first PC seems to include both functional and morphological dimensions of polycentricity, although it shows the lowest importance value among the PCs. The second, third, and sixth PCs have the overall highest impact among the PCs. The correlation coefficient between the PCs and the mobility reduction is also limited. It is again only the first principal component that shows higher and positive values, which are overall still low compared to other variables such as education or employment density. Against the last hypothesis, the degree of polycentricity has only a modest impact on mobility reduction. Compared to the much more significant impact of general spatial structure, it seems that there are other spatial aspects besides the degree of polycentricity that have been much more influential.

In regard to the last research question: Polycentricity had an overall low impact on the mobility reduction of U.S. metropolitan areas during the COVID-19 pandemic. Both feature

importance and correlation coefficient do not imply a stronger influence. Other features such as education, employment, public transportation rate, and the Asian population are much more important for modeling mobility reduction.

5.4 Limitations

There are some limitations that need to be addressed in the scope of this thesis. One of the main aspects is the data itself. As already mentioned, the data comes with inherent bias and only represents an estimated 10 % of the population. The GPS data requires a mobile phone and activated GPS tracking. Only this specific part of the population has produced the entire dataset. Especially older adults and children are unlikely to be included in the data.

Another common issue regarding mobile phone data is privacy concerns, and subsequently, the bias introduced by necessary data aggregation. Data entries are excluded when there is only one recorded visitor per geographical unit. Moreover, to enhance differential privacy, visitor counts between two and four are recorded but only shown as four. Moreover, the aggregation of flows from lower to higher levels may lead to visitor duplication.

A crucial aspect in spatial studies is always the scale of the examined area. In the context of this study, mobility networks have been constructed using census tracts as nodes. However, the size of census tracts varies widely based on the density of the settlement. This variability results in the presence of very small census tracts within highly populated areas, and conversely, very large census tracts in regions with extremely low population densities. Consequently, this variation in census tract sizes inflates the network's representation compared to the actual spatial structure, leading to a distortion of the perceived urban layout.

The time period of 14 weeks with an additional transition week that was covered in this thesis was at an early stage of the pandemic. The pandemic was drawn-out over a much larger timeframe with several waves of reoccurring infections. A more in-depth analysis would require a longer timeframe or examination of several timeframes along the peaks of infections.

Chapter 6: Conclusion

In this final chapter, a concise and clear summary of the key findings from this study is presented in Section 6.1, highlighting the contributions to the existing body of knowledge and potential applications in real-world scenarios. Additionally, potential directions for further exploration and research are discussed in Section 6.2.

6.1 Summary and Implications

This study has offered profound insights into the complex relationship between spatial structure (w.r.t. polycentricity) and mobility patterns within U.S. metropolitan statistical areas (MSAs) amid the COVID-19 pandemic. By using a novel Whole Graph Embedding (WGE) application, the research aimed to analyze mobility behavior changes and structural shifts in response to varying degrees of polycentricity. Furthermore, the study aimed to examine the importance of spatial structure on mobility, while comparing the influence of polycentricity with other socio-demographic, travel-related, and spatial variables.

Through the application of the WGE technique, the study effectively differentiated MSAs based on their degrees of polycentricity. The resulting clustering detected three primary groups – monocentric, intermediate, and polycentric. It was found that both the polycentric and intermediate clusters demonstrated distinct mobility patterns compared to the monocentric group. Contrary to initial hypotheses, the monocentric group experienced less reduction in mobility flows after the introduction of restrictions compared to more polycentric groups. It is likely the higher population density of polycentric cities that raised the risk of infection, resulting in a more restrictive mobility reduction in order to halt contagion.

By applying machine learning, the spatial structure, as represented by the whole graph embeddings, was found to play a significant role in reducing mobility flows. Compared to a baseline model, the integration of spatial structure improved performance in the range of 12 to 25 % in terms of R^2 score. This gain was further supported by the added feature importance of the principal components, accounting for around 15 % of total model importance. The findings promote the significance of spatial layout, which influences mobility patterns, transportation systems, economic activities, travel efficiency, and travel mode selection.

The study reveals that polycentricity, as one aspect of spatial structure, exerted a comparatively modest influence on mobility reduction during the early stage of the pandemic. Other socioeconomic factors, such as education, public transportation rate, employment density, and the Asian population, held greater significance. Although the clusters with different degrees of polycentricity showed significant dissimilarities, other variables appear to have been more relevant for the mobility reduction.

This study is able to gain valuable insights into the relevance of spatial structure and polycentricity on the urban mobility network of U.S. MSAs during the pandemic. In this field of research, examining such a large number of urban areas is a rarity. This was possible due to the novel application of whole graph embedding, which challenges conventional polycentricity indices by modeling more complexity of urban dynamics. Although more research is still required in this regard, its successful application has been demonstrated. The findings have an interdisciplinary relevance for economics, sociology, public health, and urban planning and design. The findings can be incorporated into creating more adaptable and resilient cities.

6.2 Future Work

There are still several open questions that can be addressed in future work. With respect to the WGE algorithm, it remains unclear what node features are best suited to describe the polycentric structure of a network. This should be tested further in more detail by comparing a broad range of different node features. One limitation mentioned is the varying size of the census tracts. This could be addressed by adding the land area as a node attribute.

In this work only two time frames were considered at an earlier stage of the pandemic. Further research about an extended time frame over the whole course of the pandemic with its reoccurring infection waves could bring more insight into the behavior of the different urban structures. It would be especially interesting to examine the recovery period of mobility to determine whether polycentric and monocentric areas behaved differently.

This extension can also be applied to other study areas. Although the U.S. offers a large number of differently sized urban spaces, they all share a similar historical and cultural background that has shaped their structure. Regions with very old historical structures like Europe or regions that experience enormous urban growth like Asia will certainly yield different results.

The studied time period was very special in regard to the COVID-19 pandemic and changed the daily routine of many people. It offered a possibility to examine the change of polycentricity. Nonetheless, its results stem from an uncommon time period. With the pandemic in the past, a more general examination of more usual circumstances may yield different results.

The change in the degree of polycentricity was mainly analyzed in regard to the functional polycentricity with changing mobility flows. The morphological dimension of the node features such as employment remained unchanged. The impact of the pandemic on the employment market was therefore not considered, although huge unemployment statistics were recorded. By also including the change in this regard, a more throughout analysis of structural changes could be conducted.

References

Aguilar, J., Bassolas, A., Ghoshal, G., Hazarie, S., Kirkley, A., Mazzoli, M., Meloni, S., Mimar, S., Nicosia, V., Ramasco, J. J., & Sadilek, A. (2020). Impact of urban structure on COVID-19 spread. arXiv preprint arXiv:2007.15367, 58.

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*. Harvard University Press.

Anas, A., Arnott, R., & Small, K. A. (1998). Urban spatial structure. *Journal of Economic Literature*, *36*(3), 1426-1464.

Angel, S., & Blei, A. M. (2016). The spatial structure of American cities: The great majority of workplaces are no longer in CBDs, employment sub-centers, or live-work communities. *Cities*, *51*, 21-35. https://doi.org/10.1016/j.cities.2015.11.031

Arribas-Bel, D., & Sanz-Gracia, F. (2014). The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. *Urban Geography*, *35*(7), 980-997. https://doi.org/10.1080/02723638.2014.940693

Badr, H. S., Du, H., Marshall, M., Dong, E., Squire, M. M., & Gardner, L. M. (2020). Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. The Lancet Infectious Diseases, 20(11), 1247–1254. https://doi.org/10.1016/S1473-3099(20)30553-3

Bartosiewicz, B., & Martcińczak, S. (2020). Investingating polycentric urban regions: Different measures - Different results. *Cities*. 105, 102855. https://doi.org/10.1016/j.cities.2020.102855

Benitez, J., Courtemanche, C., & Yelowitz, A. (2020). Racial and Ethnic Disparities in COVID-19: Evidence from Six Large Cities. Journal of Economics, Race, and Policy, 3(4), 243–261. https://doi.org/10.1007/s41996-020-00068-9

Bereitschaft, B., & Scheller, D. (2020). How Might the COVID-19 Pandemic Affect 21st Century Urban Design, Planning, and Development? Urban Science, 4(4), Article 4. https://doi.org/10.3390/urbansci4040056

Bollyky, T. J., Castro, E., Aravkin, A. Y., Bhangdia, K., Dalos, J., Hulland, E. N., Kiernan, S., Lastuka, A., McHugh, T. A., Ostroff, S. M., Zheng, P., Chaudhry, H. T., Ruggiero, E., Turilli, I., Adolph, C., Amlag, J. O., Bang-Jensen, B., Barber, R. M., Carter, A., … Dieleman, J. L. (2023). Assessing COVID-19 pandemic policies and behaviours and their economic and educational trade-offs across US states from Jan 1, 2020, to July 31, 2022: An observational analysis. The Lancet, 401(10385), 1341–1360. https://doi.org/10.1016/S0140-6736(23)00461-0

Bruine de Bruin, W., Saw, H.-W., & Goldman, D. P. (2020). Political polarization in US residents' COVID-19 risk perceptions, policy preferences, and protective behaviors. Journal of Risk and Uncertainty, 61(2), 177–194. https://doi.org/10.1007/s11166-020-09336-3

Brynjolfsson, E., Horton, J. J., Ozimek, A., Rock, D., Sharma, G., & TuYe, H.-Y. (2020). COVID-19 and Remote Work: An Early Look at US Data (Working Paper 27344). National Bureau of Economic Research. https://doi.org/10.3386/w27344

Burger, M. J., de Goei, B., van der Laan, L., & Huisman, F. J. M. (2011). Heterogeneous development of metropolitan spatial structure: Evidence from commuting patterns in English and Welsh city-regions, 1981–2001. Cities, 28(2), 160–170. https://doi.org/10.1016/j.cities.2010.11.006

Burger, M., & Meijers, E. (2012). Form Follows Function? Linking Morphological and Functional Polycentricity. *Urban Studies*, 49(5), 1127–1149. https://doi.org/10.1177/0042098011407095

Burgess, E. W. (2015). The Growth of the City: An Introduction to a Research Project. *The city reader* (pp. 212-220). Routledge. (Original work published 1925)

Cai, H., Zheng, V. W., & Chang, K. C.-C. (2018). A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications. *IEEE Transactions on Knowledge and Data Engineering*, *30*(9), 1616-1637. https://doi.org/10.1109/TKDE.2018.2807452

Carella, G., Pérez Trufero, J., Álvarez, M., & Mateu, J. (2022). A Bayesian Spatial Analysis of the Heterogeneity in Human Mobility Changes During the First Wave of the COVID-19 Epidemic in the United States. The American Statistician, 76(1), 64–72. https://doi.org/10.1080/00031305.2021.1965657

Chamberlain, B. P., Clough, J., & Deisenroth, M. P. (2017). Neural Embeddings of Graphs in Hyperbolic Space (arXiv:1705.10359). arXiv. https://doi.org/10.48550/arXiv.1705.10359

Coven, J., & Gupta, A. (2020). Disparities in mobility responses to COVID-19. New York Univ, 1.

Elarde, J., Kim, J.-S., Kavak, H., Züfle, A., & Anderson, T. (2021). Change of human mobility during COVID-19: A United States case study. PLOS ONE, 16(11), e0259031. https://doi.org/10.1371/journal.pone.0259031

Elnaggar, A., Heinzinger, M., Dallago, C., Rehawi, G., Wang, Y., Jones, L., Gibbs, T., Feher, T., Angerer, C., Steinegger, M., Bhowmik, D., & Rost, B. (2022). ProtTrans: Toward Understanding the Language of Life Through Self-Supervised Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(10), 7112–7127. https://doi.org/10.1109/TPAMI.2021.3095381

Fraiberger, S. P., Astudillo, P., Candeago, L., Chunet, A., Jones, N. K. W., Khan, M. F., Lepri, B., Gracia, N. L., Lucchini, L., Massaro, E., & Montfort, A. (2020). Uncovering socioeconomic gaps in mobility reduction during the COVID-19 pandemic using location data (arXiv:2006.15195). arXiv. https://doi.org/10.48550/arXiv.2006.15195

Fu, C., Nanni, M., Yeghikyan, G., & Weibel, R. (2021). *Measuring Polycentricity: A Whole Graph Embedding Perspective*. https://doi.org/10.25436/E2TG6C

Gao, F., Wolf, G., & Hirn, M. (2019). Geometric Scattering for Graph Data Analysis. Proceedings of the 36th International Conference on Machine Learning, 2122–2131. https://proceedings.mlr.press/v97/gao19e.html

Green, N. (2007). Functional Polycentricity: A Formal Definition in Terms of Social Network Analysis. Urban Studies, 44(11), 2077–2103. https://doi.org/10.1080/00420980701518941

Grossman, G., Kim, S., Rexer, J. M., & Thirumurthy, H. (2020). Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. Proceedings of the National Academy of Sciences, 117(39), 24144–24153. https://doi.org/10.1073/pnas.2007835117

Grover, A., & Leskovec, J. (2016). node2vec: Scalable Feature Learning for Networks (arXiv:1607.00653). arXiv. https://doi.org/10.48550/arXiv.1607.00653

Hajrasouliha, A. H., & Hamidi, S. (2017). The typology of the American metropolis: Monocentricity, polycentricity, or generalized dispersion? Urban Geography, 38(3), 420–444. https://doi.org/10.1080/02723638.2016.1165386

Hallas, L., Hatibie, A., Majumdar, S., Pyarali, M., & Hale, T. (2021). Variation in US states' responses to COVID-19. University of Oxford.

Hanssens, H., Derudder, B., Van Aelst, S., & Witlox, F. (2014). Assessing the Functional Polycentricity of the Mega-City-Region of Central Belgium Based on Advanced Producer Service Transaction Links. *Regional Studies*, *48*(12), 1939-1953. https://doi.org/10.1080/00343404.2012.759650

Harrington, D. M., & Hadjiconstantinou, M. (2022). Changes in commuting behaviours in response to the COVID-19 pandemic in the UK. Journal of Transport & Health, 24, 101313. https://doi.org/10.1016/j.jth.2021.101313

Harris, C. D., & Ullman, E. L. (1945). The Nature of Cities. The ANNALS of the American Academy of Political and Social Science, 242(1), 7–17. https://doi.org/10.1177/000271624524200103

Hoyt, H. (1939). *The Structure and Growth of Residential Neighborhoods in American Cities*. US Government Printing Office.

Hu, S., Luo, W., Darzi, A., Pan, Y., Zhao, G., Liu, Y., & Xiong, C. (2021). Do racial and ethnic disparities in following stay-at-home orders influence COVID-19 health outcomes? A mediation analysis approach. PLOS ONE, 16(11), e0259803. https://doi.org/10.1371/journal.pone.0259803

Hu, S., Xiong, C., Younes, H., Yang, M., Darzi, A., & Jin, Z. C. (2022). Examining spatiotemporal evolution of racial/ethnic disparities in human mobility and COVID-19 health outcomes: Evidence from the contiguous United States. Sustainable Cities and Society, 76, 103506. https://doi.org/10.1016/j.scs.2021.103506

Huang, X., Lu, J., Gao, S., Wang, S., Liu, Z., & Wei, H. (2022). Staying at Home Is a Privilege: Evidence from Fine-Grained Mobile Phone Location Data in the United States during the COVID-19 Pandemic. Annals of the American Association of Geographers, 112(1), 286–305. https://doi.org/10.1080/24694452.2021.1904819

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L., & Daszak, P. (2008). Global trends in emerging infectious diseases. Nature, 451(7181), Article 7181. https://doi.org/10.1038/nature06536

Kang, Y., Gao, S., Liang, Y., Li, M., Rao, J., & Kruse, J. (2020). Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. Scientific Data, 7(1), Article 1. https://doi.org/10.1038/s41597-020-00734-5

Karpathy, A., & Fei-Fei, L. (2015). Deep Visual-Semantic Alignments for Generating Image Descriptions (arXiv:1412.2306). arXiv. http://arxiv.org/abs/1412.2306

Kazemi, S. M., Goel, R., Eghbali, S., Ramanan, J., Sahota, J., Thakur, S., Wu, S., Smyth, C., Poupart, P., & Brubaker, M. (2019). *Time2Vec: Learning a Vector Representation of Time* (arXiv:1907.05321). arXiv. https://doi.org/10.48550/arXiv.1907.05321

Knox, P. L., & McCarthy, L. (2012). Urbanization : An introduction to urban geography (3rd ed.). Pearson.

Le, Q., & Mikolov, T. (2014). Distributed Representations of Sentences and Documents. *Proceedings of the 31st International Conference on Machine Learning*, 1188–1196. https://proceedings.mlr.press/v32/le14.html

Li, Y., Li, M., Rice, M., Zhang, H., Sha, D., Li, M., Su, Y., & Yang, C. (2021). The Impact of Policy Measures on Human Mobility, COVID-19 Cases, and Mortality in the US: A Spatiotemporal Perspective. International Journal of Environmental Research and Public Health, 18(3), Article 3. https://doi.org/10.3390/ijerph18030996

Limtanakool, N., Dijst, M., & Schwanen, T. (2007). A Theoretical Framework and Methodology for Characterising National Urban Systems on the Basis of Flows of People: Empirical Evidence for France and Germany. Urban Studies, 44(11), 2123–2145. https://doi.org/10.1080/00420980701518990

Liu, X., & Wang, M. (2016). How polycentric is urban China and why? A case study of 318 cities. *Landscape and urban planning*, *151*, 10-20. https://doi.org/10.1016/j.landurbplan.2016.03.007

Liu, X., Derudder, B., & Wu, K. (2016). Measuring Polycentric Urban Development in China: An Intercity Transportation network Perspective. *Regional Studies*, *50*(8), 1302-1315. https://doi.org/10.1080/00343404.2015.1004535

Louf, R., & Barthelemy, M. (2014). How congestion shapes cities: From mobility patterns to scaling. Scientific Reports, 4(1), Article 1. https://doi.org/10.1038/srep05561

Luo, Y., Chen, Z., Hershey, J. R., Le Roux, J., & Mesgarani, N. (2017). Deep clustering and conventional networks for music separation: Stronger together. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 61-65. https://doi.org/10.1109/ICASSP.2017.7952118

McMillen, D. P. (2001). Nonparametric Employment Subcenter Identification. *Journal of Urban Economics*, *50*(3), 448-473. https://doi.org/10.1006/juec.2001.2228

McNeil, D. G. Jr. (2020, March 26). The U.S. Now Leads the World in Confirmed Coronavirus Cases. The New York Times. https://www.nytimes.com/2020/03/26/health/usa-coronavirus-cases.html

Meijers, E., & Sandberg, K. (2008). Reducing Regional Disparities by Means of Polycentric Development: Panacea or Placebo? Scienze Regionali : Italian Journal of Regional Science : 7, Supplemento 2, 2008, 71–96. https://doi.org/10.3280/SCRE2008-SU2005

Meijers, E. J., & Burger, M. J. (2010). Spatial Structure and Productivity in US metropolitan Areas. *Environment and planning A: Economy and Space*, *42*(6), 1383-1402. https://doi.org/10.1068/a42151

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space (arXiv:1301.3781). arXiv. https://doi.org/10.48550/arXiv.1301.3781

Narayanan, A., Chandramohan, M., Venkatesan, R., Chen, L., Liu, Y., & Jaiswal, S. (2017). graph2vec: Learning Distributed Representations of Graphs. *arXiv preprint arXiv:1707.05005*. https://doi.org/10.48550/arXiv.1707.05005

Niepert, M., Ahmed, M., & Kutzkov, K. (2016). Learning Convolutional Neural Networks for Graphs. *Proceedings of The 33rd International Conference on Machine Learning*, 2014–2023. https://proceedings.mlr.press/v48/niepert16.html

Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K. E. C., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N. F., Cattarino, L., Cooper, L. V., Coupland, H., Cucunuba, Z. M., Cuomo-Dannenburg, G., Dighe, A., Djaafara, B. A., Dorigatti, I., Eales, O. D., van Elsland, S. L., Nascimento, F. F., … Donnelly, C. A. (2021). Reduction in mobility and COVID-19 transmission. Nature Communications, 12(1), Article 1. https://doi.org/10.1038/s41467-021-21358-2

Office of Management and Budget. (2021, July 16). 2020 Standards for Delineating Core Based Statistical Areas. Federal Register. https://www.federalregister.gov/documents/2021/07/16/2021-15159/2020-standards-for-delineating-core-based-statistical-areas

Qian, Z. (2010). Without zoning: Urban development and land use controls in Houston. Cities, 27(1), 31–41. https://doi.org/10.1016/j.cities.2009.11.006

Roy, A., & Kar, B. (2020). Characterizing the spread of COVID-19 from human mobility patterns and SocioDemographic indicators. Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities, 39–48. https://doi.org/10.1145/3423455.3430303

SafeGraph (2020). The impact of coronavirus (COVID-19) on foot traffic. [No longer accessible]

Schwanen, T., Dieleman, F. M., & Dijst, M. (2004). The Impact of Metropolitan Structure on Commute Behavior in the Netherlands: A Multilevel Approach. Growth and Change, 35(3), 304–333. https://doi.org/10.1111/j.1468-2257.2004.00251.x

Smart Growth America. (2014). Measuring Sprawl 2014. Retrieved August 22, 2023, from https://smartgrowthamerica.org/resources/measuring-sprawl-2014/

Sun, B., He, Z., Zhang, T., & Wang, R. (2016). Urban spatial structure and commute duration: An empirical study of China. *International Journal of Sustainable Transportation*, 10(7), 638-644. https://doi.org/10.1080/15568318.2015.1042175

Thomas, F. M. F., Charlton, S. G., Lewis, I., & Nandavar, S. (2021). Commuting before and after COVID-19. Transportation Research Interdisciplinary Perspectives, 11, 100423. https://doi.org/10.1016/j.trip.2021.100423

Tsitsulin, A., Mottin, D., Karras, P., Bronstein, A., & Müller, E. (2018). NetLSD: Hearing the Shape of a Graph. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2347–2356. https://doi.org/10.1145/3219819.3219991

Weill, J. A., Stigler, M., Deschenes, O., & Springborn, M. R. (2020). Social distancing responses to COVID-19 emergency declarations strongly differentiated by income. Proceedings of the National Academy of Sciences, 117(33), 19658–19660. https://doi.org/10.1073/pnas.2009412117

Whitehouse. Proclamation on Declaring a National Emergency Concerning the Novel Coronavirus Disease (COVID-19) Outbreak (13 March 2020). Retrieved August 22, 2023, from https://trumpwhitehouse.archives.gov/presidential-actions/proclamation-declaring-national-emergency-concerning-novel-coronavirus-disease-covid-19-outbreak/

World Health Organization. (2020). WHO Director-General's opening remarks at the media briefing on COVID-19. (No Title). Retrieved August 22, 2023, from https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., Yuan, M.-L., Zhang, Y.-L., Dai, F.-H., Liu, Y., Wang, Q.-M., Zheng, J.-J., Xu, L., Holmes, E. C., & Zhang, Y.-Z. (2020). A new coronavirus associated with human respiratory disease in China. Nature, 579(7798), Article 7798. https://doi.org/10.1038/s41586-020-2008-3

Zhang, M., Wang, S., Hu, T., Fu, X., Wang, X., Hu, Y., Halloran, B., Li, Z., Cui, Y., Liu, H., Liu, Z., & Bao, S. (2022). Human mobility and COVID-19 transmission: A systematic review and future directions. Annals of GIS, 28(4), 501–514. https://doi.org/10.1080/19475683.2022.2041725

Zhu, G., Stewart, K., Niemeier, D. & Fan, J. (2021). Understanding the Drivers of Mobility during the COVID-19 Pandemic in Florida, USA Using a Machine Learning Approach. ISPRS International Journal of Geo-Information, 10(7), 440. https://doi.org/10.3390/ijgi10070440

Appendix A: Figures

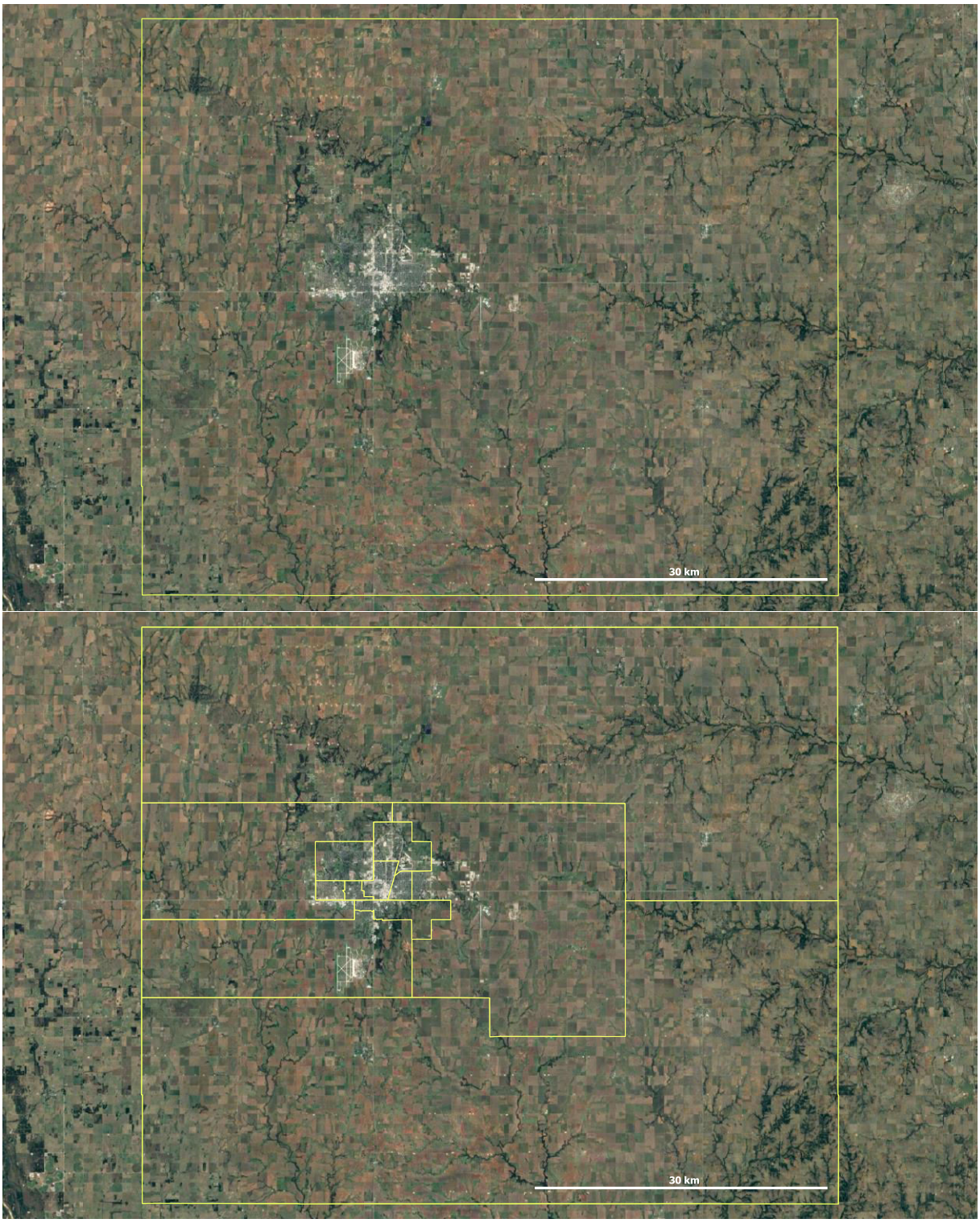


Figure A.1: Enid MSA, Oklahoma.

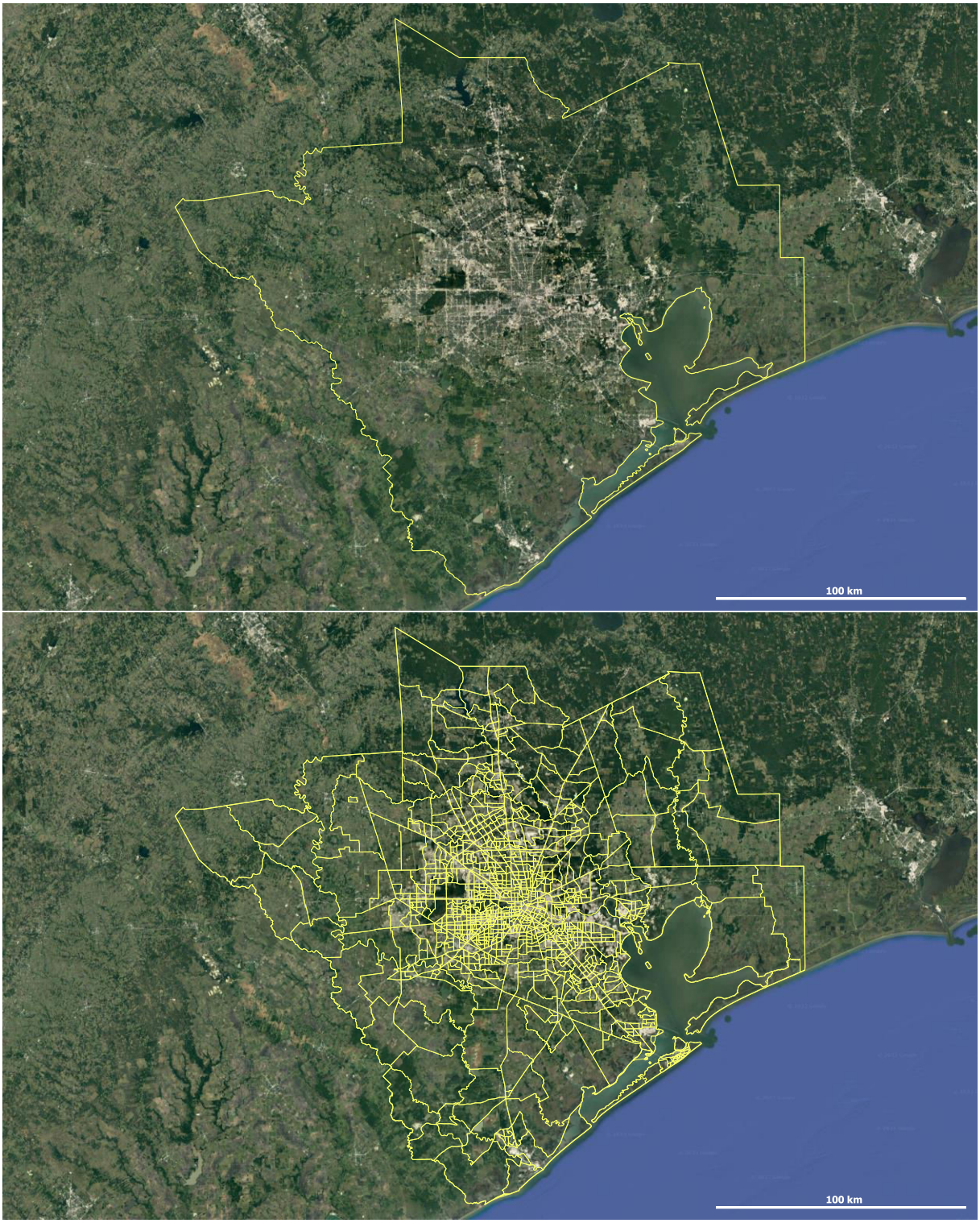


Figure A.2: Houston-Sugar Land-Baytown MSA, Texas.

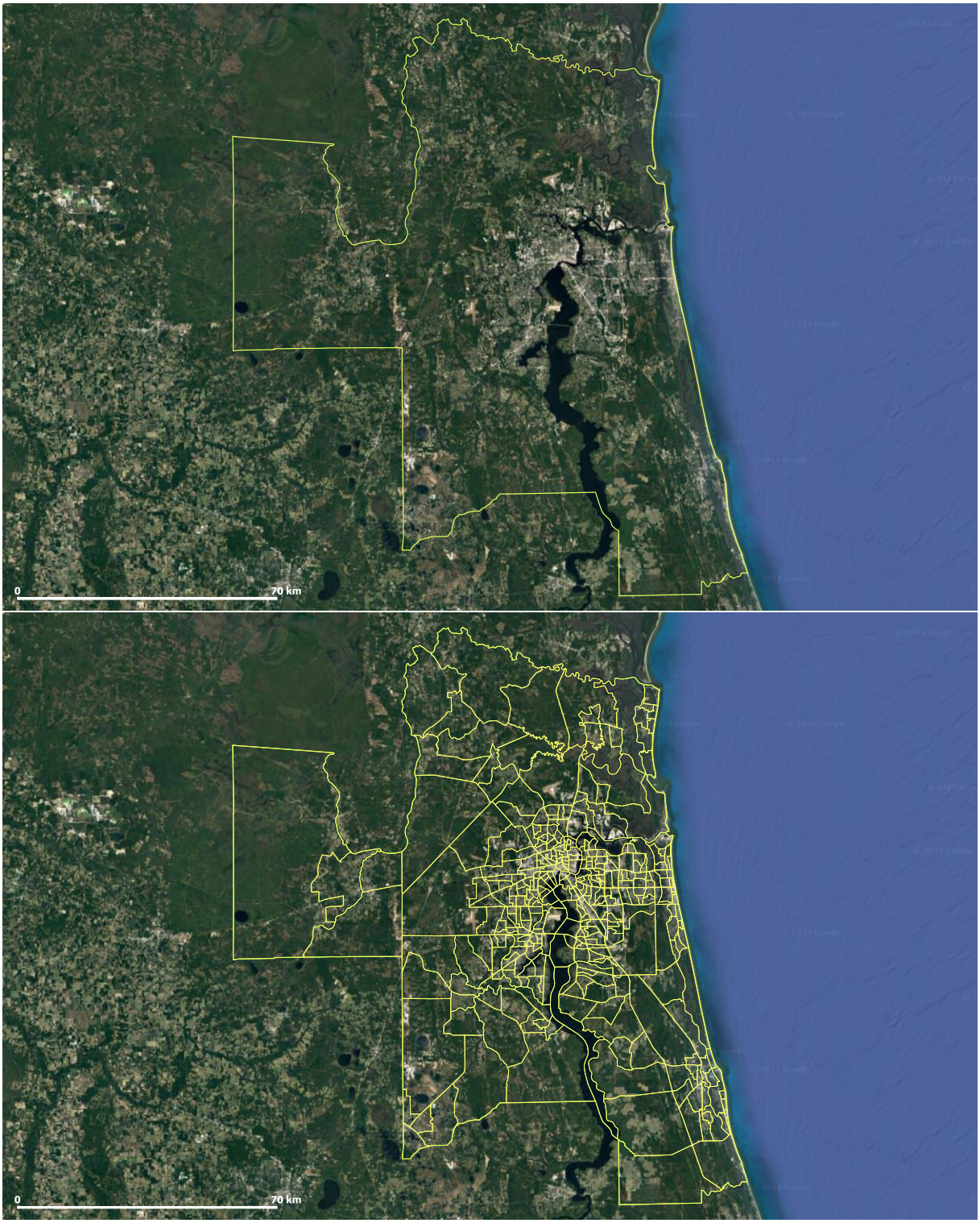


Figure A.3: Jacksonville MSA, Florida.

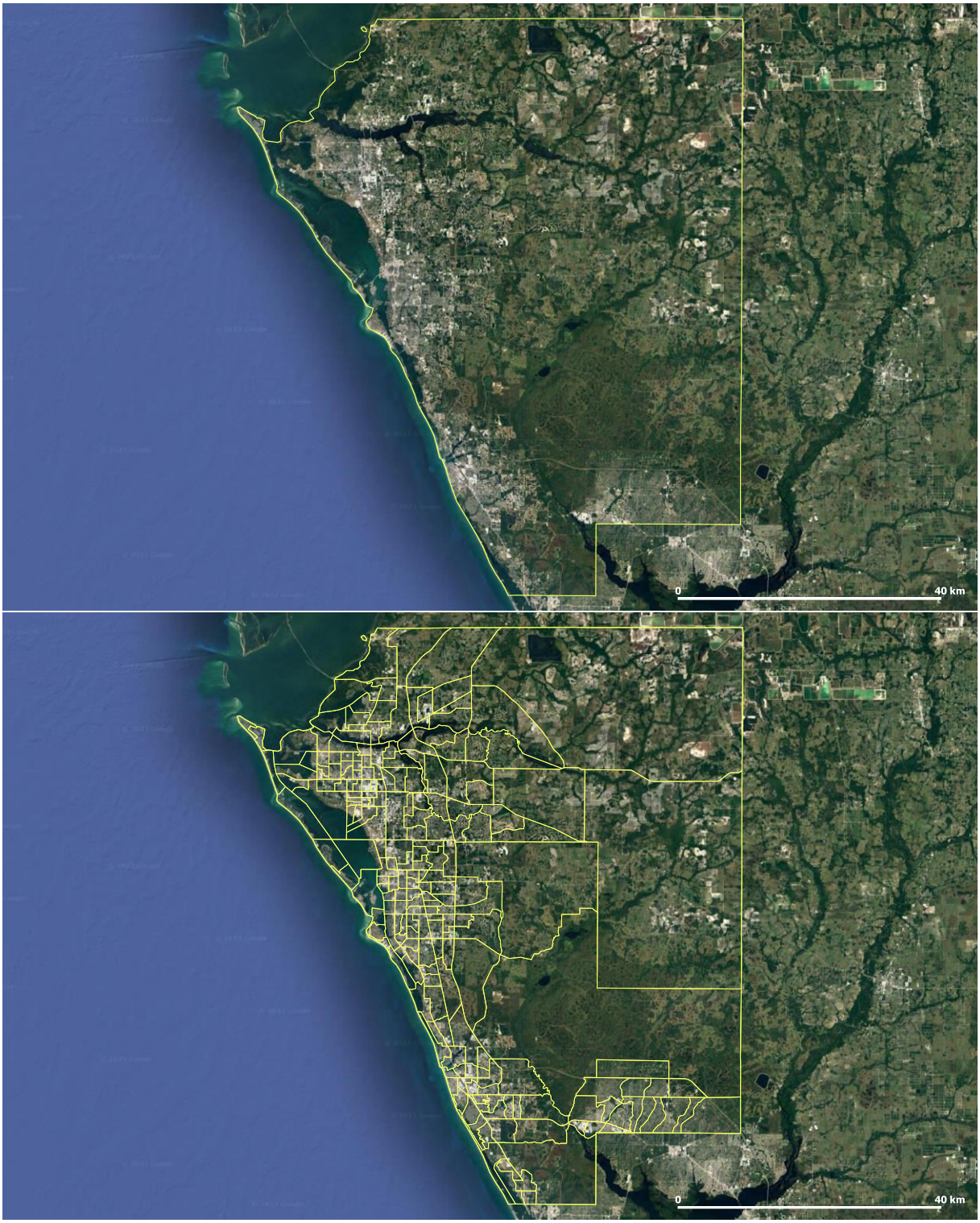


Figure A.4: North Port-Bradenton-Sarasota MSA, Florida.

Appendix B: Tables

Table B.1: Overview of outlier MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
1.	Jacksonville	Florida	5	260	1'533'796	8'291	32.011	21.02
2.	North Port-Sarasota-Bradenton	Florida	2	173	821'613	3'364	28.095	31.538

Table B.2: Overview of polycentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
1.	Houston-The Woodlands-Sugar Land	Texas	9	1'070	6'979'613	21'389	33.86	-5.301
2.	Atlanta-Sandy Springs-Alpharetta	Georgia	29	951	5'947'008	22'497	29.687	-10.568
3.	Dallas-Fort Worth-Arlington	Texas	11	1'312	7'451'858	22'457	27.25	-8.288
4.	Memphis	Arkansas, Mississippi, Tennessee	8	311	1'343'150	11'857	20.177	-2.221
5.	Birmingham-Hoover	Alabama	6	246	1'088'170	11'625	19.927	-7.968
6.	Sebring-Avon Park	Florida	3	89	390'211	3'471	17.966	5.097
7.	New Orleans-Metairie	Louisiana	8	403	1'271'651	8'293	15.18	-4.407
8.	Raleigh-Cary	North Carolina	3	224	1'362'997	5'487	15.076	-0.789
9.	San Francisco-Oakland-Berkeley	California	2	223	1'215'955	19'901	12.905	3.614
10.	Kansas City	Kansas, Missouri	14	530	2'144'129	18'792	12.602	-3.634
11.	Valdosta	Georgia	4	34	146'462	4'117	12.347	-7.918
12.	New York-Newark-Jersey City	New York, New Jersey, Pennsylvania	23	4'532	19'261'570	17'319	11.728	-10.418
13.	Minneapolis-St. Paul-Bloomington	Minnesota, Wisconsin	15	785	3'605'450	18'255	11.569	-2.496
14.	Santa Fe	New Mexico	5	979	4'709'220	6'399	10.814	-5.571
15.	Chicago-Naperville-Elgin	Illinois, Indiana, Wisconsin	14	2'212	9'478'801	18'640	10.669	-6.242
16.	Nashville-Davidson–Murfreesboro–Franklin	Tennessee	13	374	1'904'186	14'736	10.482	-4.551

Table B.3.1: Overview of intermediate MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
1.	Orlando-Kissimmee-Sanford	Florida	4	390	170'924	1'126	12.739	0.457
2.	Oklahoma City	Oklahoma	7	363	284'698	1'870	11.877	2.984
3.	Phoenix-Mesa-Chandler	Arizona	2	991	1'397'040	14'275	11.304	-1.563
4.	Austin-Round Rock-Georgetown	Texas	5	350	2'173'804	10'930	10.87	0.397
5.	Salt Lake City	Utah	15	615	2'806'349	20'366	9.115	-3.821
6.	Mobile	Alabama	2	120	430'313	5'982	8.053	0.837
7.	Washington-Arlington-Alexandria	Maryland, District of Columbia, Virginia, West Virginia,	25	1'360	6'250'309	17'002	6.998	-4.79
8.	Miami-Fort Lauderdale-Pompano Beach	Florida	3	1'217	6'129'858	13'151	6.97	-0.03
9.	Tuscaloosa	Alabama	4	61	251'878	9'050	6.927	-7.503
10.	Louisville/Jefferson County	Indiana, Kentucky	10	299	1'262'287	8'382	6.278	-1.106
11.	Denver-Aurora-Lakewood	Colorado	10	618	2'928'437	21'617	5.741	-1.574
12.	San Luis Obispo-Paso Robles	California	8	457	2'510'211	18'940	5.632	-0.477
13.	Los Angeles-Long Beach-Anaheim	California	2	2'926	13'211'027	12'558	5.58	-4.44
14.	Charleston-North Charleston	South Carolina	3	156	790'955	6'704	5.4	-1.194
15.	South Bend-Mishawaka	Indiana, Michigan	3	93	397'590	6'719	5.274	3.364
16.	Columbus	Alabama, Georgia	7	83	83'280	1'054	5.003	2.382
17.	Indianapolis-Carmel-Anderson	Indiana	11	397	2'050'933	11'154	4.629	-1.201
18.	Jackson	Mississippi	7	134	596'287	14'001	4.599	-0.28
19.	Tucson	Arizona	1	241	1'038'476	23'795	4.561	3.226
20.	Cincinnati	Indiana, Kentucky, Ohio	16	503	304'584	5'588	4.52	-1.332
21.	Little Rock-North Little Rock-Conway	Arkansas	6	164	740'602	10'581	4.475	-0.487
22.	Charlotte-Concord-Gastonia	North Carolina, South Carolina	11	545	2'595'027	14'500	4.365	-1.855
23.	Naples-Marco Island	Florida	1	74	379'345	5'176	4.349	5.844

Table B.3.2: Overview of intermediate MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
24.	Detroit-Warren-Dearborn	Michigan	6	1'298	4'317'384	10'071	4.178	-2.724
25.	Elmira	New York	2	162	841'602	14'462	4.003	3.756
26.	Columbus	Ohio	10	433	429'120	3'968	3.94	-1.843
27.	Tampa-St. Petersburg-Clearwater	Florida	4	744	186'580	4'946	3.881	-2.358
28.	Baton Rouge	Louisiana	10	157	856'779	11'307	3.598	-0.337
29.	Colorado Springs	Colorado	2	136	207'248	4'440	3.145	2.73
30.	Montgomery	Alabama	4	96	373'552	7'027	3.099	0.049
31.	Richmond	Virginia	17	289	1'282'067	11'302	3.056	-0.238
32.	Sherman-Denison	Texas	3	720	3'928'498	15'210	2.773	-2.047
33.	Fargo	Minnesota, North Dakota	2	46	243'966	7'279	2.697	4.1
34.	Columbia	South Carolina	6	191	319'643	7'217	2.669	-0.413
35.	Riverside-San Bernardino-Ontario	California	2	822	4'600'396	70'612	2.605	-2.126
36.	Santa Maria-Santa Barbara	California	2	383	1'985'926	6'939	2.418	-1.834
37.	Santa Cruz-Watsonville	California	1	628	3'323'970	10'896	2.386	-0.01
38.	Durham-Chapel Hill	North Carolina	5	121	636'256	5'932	2.337	-1.571
39.	Champaign-Urbana	Illinois	2	47	226'418	3'718	2.327	0.162
40.	Des Moines-West Des Moines	Iowa	6	140	690'585	9'361	2.296	0.139
41.	Tulsa	Oklahoma	7	272	996'141	16'238	2.254	-0.531
42.	Boston-Cambridge-Newton	Massachusetts, New Hampshire	7	1'004	4'854'808	9'033	2.191	-2.366
43.	Greensboro-High Point	North Carolina	3	168	767'467	5'164	2.145	-1.08
44.	Huntsville	Alabama	2	89	464'607	3'527	1.919	0.119
45.	Cleveland-Elyria	Ohio	5	638	161'676	3'223	1.808	-1.392
46.	Virginia Beach-Norfolk-Newport News	North Carolina, Virginia	19	427	1'768'956	9'166	1.7	-1.147
47.	Knoxville	Tennessee	8	199	861'872	8'340	1.651	0.134
48.	Urban Honolulu	Hawaii	1	243	979'682	1'556	1.578	-0.707

Table B.3.3: Overview of intermediate MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
49.	Cape Coral-Fort Myers	Florida	1	167	756'570	2'032	1.553	-0.088
50.	Las Vegas-Henderson-Paradise	Nevada	1	487	2'228'866	20'439	1.337	-0.605
51.	Albuquerque	New Mexico	4	202	915'986	24'042	1.337	0.984
52.	Omaha-Council Bluffs	Iowa, Nebraska	8	255	2'560'260	9'010	1.316	-0.574
53.	Sacramento-Roseville-Folsom	California	4	485	2'338'866	13'194	1.307	-0.068
54.	Portland-South Portland	Maine	7	491	2'472'774	17'311	1.216	-0.831
55.	Madison	Wisconsin	4	133	660'212	8'572	1.203	-0.432
56.	Portland-Vancouver-Hillsboro	Oregon, Washington	2	78	481'334	2'889	1.18	1.241
57.	Philadelphia-Camden-Wilmington	New Jersey, Pennsylvania, Delaware, Maryland	11	1'476	4'860'338	37'726	1.158	-1.389
58.	Pittsburgh	Pennsylvania	7	711	2'324'447	13'679	1.069	-1.396
59.	Wichita Falls	Texas	4	149	639'668	10'747	0.843	0.139
60.	Baltimore-Columbia-Towson	Maryland	7	679	2'800'427	6'738	0.819	-1.021

Table B.4.1: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
1.	Beaumont-Port Arthur	Texas	3	104	394'268	5'441	0.644	0.097
2.	Winston-Salem	North Carolina	5	150	671'156	5'203	0.607	-0.687
3.	St. Cloud	Minnesota	2	56	698'537	4'776	0.465	0.993
4.	Greenville-Anderson	South Carolina	4	195	908'680	7'021	0.456	-0.559
5.	Boise City	Idaho	5	95	730'483	30'473	0.323	0.095
6.	Rockford	Illinois	2	84	336'928	2'057	0.262	0.637
7.	Abilene	Texas	3	47	171'354	7'106	0.251	1.182
8.	Gulfport-Biloxi	Mississippi	4	84	414'662	5'740	0.122	0.473
9.	Grand Rapids-Kentwood	Michigan	4	207	1'069'696	6'960	0.08	-0.585
10.	St. Joseph	Kansas, Missouri	5	91	134'409	1'030	0.076	0.127
11.	Gainesville	Florida	3	70	327'329	6'069	0.064	-0.043
12.	El Paso	Texas	3	34	152'275	3'083	0.042	-1.586
13.	Crestview-Fort Walton Beach-Destin	Florida	2	55	278'479	5'097	0.037	-0.707
14.	Alexandria	Louisiana	2	38	152'715	5'079	0.028	0.713
15.	Milwaukee-Waukesha	Wisconsin	4	430	1'576'525	3'768	-0.005	-0.522
16.	Chattanooga	Georgia, Tennessee	6	119	561'055	5'411	-0.011	-0.23
17.	Pensacola-Ferry Pass-Brent	Florida	2	97	403'747	8'642	-0.018	0.099
18.	Tallahassee	Florida	4	84	3'152'928	6'510	-0.106	0.104
19.	Davenport-Moline-Rock Island	Illinois, Iowa	4	104	380'274	5'880	-0.208	-0.151
20.	Ann Arbor	Michigan	1	100	368'385	1'829	-0.248	0.367
21.	Lansing-East Lansing	Michigan	4	147	547'786	5'772	-0.279	-0.307
22.	Pine Bluff	Arkansas	3	30	89'464	5'258	-0.327	0.543
23.	Dayton-Kettering	Ohio	3	209	805'929	3'321	-0.331	-0.347
24.	Bakersfield	California	1	151	892'458	21'062	-0.351	0.424
25.	Toledo	Ohio	4	177	558'635	10'986	-0.358	-0.25
26.	Fort Collins	Colorado	1	73	350'523	6'724	-0.371	0.376
27.	Rochester	New York	6	270	1'071'784	8'460	-0.406	-0.38
28.	Albany	Georgia	4	41	147'431	4'120	-0.434	1.464
29.	Springfield	Massachusetts	2	86	322'494	2'456	-0.462	-0.59
30.	Worcester	Connecticut, Massachusetts	2	197	943'312	5'242	-0.482	-0.69
31.	Lafayette	Louisiana	5	93	490'220	8'829	-0.511	0.196
32.	Augusta-Richmond County	Georgia, South Carolina	7	119	605'303	9'015	-0.532	0.037
33.	Wilmington	North Carolina	2	60	293'339	2'749	-0.575	0.439
34.	Provo-Orem	Utah	2	130	633'129	13'975	-0.612	0.059

Table B.4.2: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
35.	Charlottesville	Virginia	5	44	217'455	4'259	-0.626	0.364
36.	Lincoln	Nebraska	2	78	333'193	3'650	-0.66	0.316
37.	McAllen-Edinburg-Mission	Texas	1	113	861'137	4'069	-0.684	0.035
38.	Harrisburg-Carlisle	Pennsylvania	3	124	574'691	4'201	-0.699	-0.356
39.	Evansville	Indiana, Kentucky	4	78	315'186	3'794	-0.761	0.328
40.	Greeley	Colorado	1	77	315'389	10'327	-0.766	0.532
41.	Macon-Bibb County	Georgia	5	60	229'565	4'462	-0.767	0.329
42.	Buffalo-Cheektowaga	New York	2	297	1'129'018	4'054	-0.784	-0.144
43.	Palm Bay-Melbourne-Titusville	Florida	1	113	180'076	1'965	-0.795	0.183
44.	Monroe	Louisiana	3	54	202'138	5'912	-0.82	0.25
45.	Lexington-Fayette	Kentucky	6	129	514'273	3'804	-0.851	-0.049
46.	Jackson	Tennessee	4	49	178'601	4'432	-0.89	0.266
47.	Stockton	California	1	139	140'540	3'295	-0.894	0.29
48.	Peoria	Illinois	6	106	6'092'403	11'920	-0.922	-0.17
49.	Syracuse	New York	3	185	384'783	6'185	-0.926	-0.357
50.	Fayetteville	North Carolina	3	104	523'480	4'243	-0.956	0.332
51.	Fresno	California	1	199	990'204	15'432	-1.013	-0.06
52.	Hartford-East Hartford-Middletown	Connecticut	3	289	1'205'842	3'923	-1.022	-0.234
53.	Ogden-Clearfield	Utah	4	117	672'948	18'725	-1.024	0.035
54.	Scranton–Wilkes-Barre	Pennsylvania	1	50	150'319	4'946	-1.028	-0.68
55.	Laredo	Texas	1	61	274'847	8'707	-1.05	0.404
56.	Lakeland-Winter Haven	Florida	1	153	705'735	4'657	-1.068	-0.255
57.	Hagerstown-Martinsburg	Maryland, West Virginia	3	50	285'990	2'611	-1.097	-0.574
58.	Lubbock	Texas	3	74	320'031	6'962	-1.099	0.276
59.	Reno	Nevada	2	111	468'268	17'005	-1.109	0.175
60.	Texarkana	Arkansas, Texas	3	34	129'938	1'417	-1.116	0.791
61.	Springfield	Missouri	1	69	313'791	2'093	-1.143	0.338
62.	Waco	Texas	2	57	271'326	4'669	-1.187	0.31
63.	Atlantic City-Hammonton	New Jersey	1	70	264'650	1'440	-1.19	0.486
64.	Springfield	Ohio	2	117	208'224	3'064	-1.208	0.118
65.	Deltona-Daytona Beach-Ormond Beach	Florida	2	134	658'961	4'109	-1.228	-0.164
66.	Lake Charles	Louisiana	2	46	209'821	6'083	-1.229	0.402
67.	Springfield	Illinois	4	57	264'437	6'672	-1.234	0.169

Table B.4.3: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
68.	Myrtle Beach-Conway-North Myrtle Beach	North Carolina, South Carolina	2	105	481'489	5'131	-1.238	0.095
69.	Brownsville-Harlingen	Texas	1	86	422'135	2'308	-1.26	0.425
70.	Modesto	California	1	94	546'235	3'872	-1.279	0.155
71.	San Antonio-New Braunfels	Texas	1	94	432'977	8'497	-1.28	-0.109
72.	Corpus Christi	Texas	2	98	92'168	1'751	-1.328	0.015
73.	Providence-Warwick	Massachusetts, Rhode Island	6	369	1'621'099	4'111	-1.336	-0.129
74.	Amarillo	Texas	5	67	264'345	13'338	-1.346	0.234
75.	Port St. Lucie	Florida	3	115	536'314	5'387	-1.347	-0.062
76.	Athens-Clarke County	Georgia	4	46	210'810	2'655	-1.36	-0.005
77.	Bridgeport-Stamford-Norwalk	Connecticut	1	211	944'306	1'619	-1.364	-0.3
78.	Albany-Schenectady-Troy	New York	5	218	880'766	7'282	-1.406	-0.049
79.	Boulder	Colorado	1	68	324'682	1'882	-1.425	-0.079
80.	Manchester-Nashua	New Hampshire	1	86	415'305	2'270	-1.431	-0.178
81.	St. George	Utah	3	157	466'897	7'788	-1.432	-0.175
82.	Flint	Michigan	1	131	406'770	1'650	-1.434	0.013
83.	Lawton	Oklahoma	2	34	127'157	4'409	-1.455	0.503
84.	Fort Smith	Arkansas, Oklahoma	4	49	250'260	6'235	-1.473	0.156
85.	Allentown-Bethlehem-Easton	New Jersey, Pennsylvania	4	179	841'265	3'764	-1.474	-0.171
86.	Fort Wayne	Indiana	2	103	409'419	2'572	-1.484	0.094
87.	Vallejo	California	1	96	444'538	2'129	-1.497	-0.118
88.	Sheboygan	Wisconsin	3	170	554'787	4'525	-1.53	0.03
89.	Topeka	Kansas	5	57	232'248	8'373	-1.549	0.027
90.	Wichita	Kansas	3	43	151'352	6'785	-1.557	0.433
91.	Green Bay	Wisconsin	3	69	320'827	4'844	-1.557	0.239
92.	Seattle-Tacoma-Bellevue	Washington	1	90	444'895	7'084	-1.567	-0.407
93.	Killeen-Temple	Texas	3	89	452'428	7'294	-1.573	0.119
94.	Akron	Ohio	2	170	703'286	2'332	-1.576	-0.04
95.	Muskegon	Michigan	1	42	173'679	1'294	-1.582	0.437
96.	Fayetteville-Springdale-Rogers	Arkansas	3	85	526'101	6'796	-1.582	-0.128
97.	Clarksville	Kentucky, Tennessee	4	66	123'731	1'978	-1.593	0.333
98.	Oxnard-Thousand Oaks-Ventura	California	1	174	594'001	2'631	-1.594	0.024

Table B.4.4: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
99.	Youngstown-Warren-Boardman	Ohio, Pennsylvania	3	155	538'115	4'410	-1.597	-0.02
100.	Cape Girardeau	Illinois, Missouri	3	23	97'026	3'709	-1.605	-0.085
101.	Santa Rosa-Petaluma	California	1	54	282'517	8'544	-1.606	0.002
102.	Tyler	Texas	1	41	230'184	2'387	-1.627	0.389
103.	Lafayette-West Lafayette	Indiana	4	49	230'353	4'256	-1.627	-0.041
104.	New Haven-Milford	Connecticut	1	190	855'733	1'566	-1.647	-0.068
105.	York-Hanover	Pennsylvania	1	90	447'628	2'342	-1.679	-0.04
106.	Eugene-Springfield	Oregon	1	86	377'749	11'793	-1.694	0.053
107.	Columbia	Missouri	3	37	832'925	9'591	-1.724	0.075
108.	Cedar Rapids	Iowa	3	57	271'734	5'203	-1.752	0.041
109.	Poughkeepsie-Newburgh-Middletown	New York	2	158	675'601	4'163	-1.754	0.072
110.	Huntington-Ashland	Kentucky, Ohio, West Virginia	7	95	359'481	6'475	-1.763	0.008
111.	Appleton	Wisconsin	2	51	236'834	2'476	-1.767	0.105
112.	Rochester	Minnesota	4	50	219'848	6'416	-1.787	0.104
113.	Visalia	California	1	78	463'955	12'495	-1.817	0.44
114.	Lynchburg	Virginia	5	60	262'937	5'474	-1.848	-0.05
115.	Coeur d'Alene	Idaho	1	25	261'701	5'440	-1.857	0.473
116.	Charleston	West Virginia	5	75	260'920	6'855	-1.884	-0.044
117.	Utica-Rome	New York	2	93	290'812	6'796	-1.895	0.159
118.	Rapid City	South Dakota	2	28	140'653	16'182	-1.898	0.366
119.	Anchorage	Alaska	2	79	399'450	68'150	-1.901	0.068
120.	Roanoke	Virginia	6	65	313'289	4'840	-1.903	0.07
121.	Panama City	Florida	1	44	90'151	1'551	-1.908	0.378
122.	Iowa City	Iowa	2	29	172'919	3'064	-1.913	0.296
123.	Canton-Massillon	Ohio	2	93	398'711	2'512	-1.923	-0.017
124.	Wheeling	Ohio, West Virginia	3	47	140'199	2'444	-1.925	0.162
125.	Bowling Green	Kentucky	4	39	176'852	4'183	-1.928	0.092
126.	Binghamton	New York	2	65	240'473	3'172	-1.934	0.144
127.	Asheville	North Carolina	4	103	459'344	5'266	-1.935	0.114
128.	Longview	Texas	4	59	285'554	6'942	-1.938	0.144
129.	Spokane-Spokane Valley	Washington	4	35	144'152	5'372	-1.953	0.144
130.	Kalamazoo-Portage	Michigan	1	57	264'322	1'455	-1.969	0.189
131.	Waterloo-Cedar Falls	Iowa	3	50	169'107	3'894	-1.969	0.105
132.	Erie	Pennsylvania	1	72	272'046	2'070	-1.972	0.145
133.	San Diego-Chula Vista-Carlsbad	California	4	97	411'137	5'436	-1.985	0.158
134.	Gadsden	Alabama	1	30	102'721	1'386	-1.994	0.396
135.	Hattiesburg	Mississippi	4	32	168'646	5'244	-1.995	0.225

Table B.4.5: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
136.	Duluth	Minnesota, Wisconsin	4	88	289'276	27'253	-2.002	0.164
137.	San Angelo	Texas	2	70	428'472	4'981	-2.005	0.082
138.	Midland	Texas	2	29	176'914	4'702	-2.007	0.487
139.	College Station-Bryan	Texas	3	52	735'480	6'952	-2.012	0.059
140.	Reading	Pennsylvania	1	90	419'062	2'219	-2.012	0.091
141.	Bloomington	Illinois	1	41	172'164	3'065	-2.014	0.27
142.	Trenton-Princeton	New Jersey	1	77	368'085	582	-2.018	0.086
143.	Hilton Head Island-Bluffton	South Carolina	2	45	219'290	3'190	-2.03	0.057
144.	Lancaster	Pennsylvania	1	98	543'050	2'445	-2.033	0.074
145.	Warner Robins	Georgia	2	29	182'819	1'362	-2.039	0.379
146.	Pueblo	Colorado	1	54	167'412	6'180	-2.075	0.195
147.	Lima	Ohio	1	33	102'808	1'043	-2.08	0.314
148.	Anniston-Oxford	Alabama	1	30	114'324	1'570	-2.121	0.266
149.	Ocala	Florida	1	61	360'210	4'104	-2.125	0.152
150.	Bremerton-Silverdale-Port Orchard	Washington	1	55	268'945	1'023	-2.149	0.19
151.	Sebastian-Vero Beach	Florida	1	99	496'801	4'082	-2.155	0.111
152.	Brunswick	Georgia	3	21	118'149	3'332	-2.161	0.12
153.	Norwich-New London	Connecticut	1	65	266'868	1'723	-2.177	0.136
154.	Medford	Oregon	1	41	218'781	7'210	-2.178	0.215
155.	Yuba City	California	2	35	173'839	3'197	-2.182	0.3
156.	Bloomington	Indiana	2	36	168'172	2'020	-2.19	-0.007
157.	Burlington-South Burlington	Vermont	3	47	219'764	3'243	-2.192	0.15
158.	Jacksonville	North Carolina	1	31	198'377	1'976	-2.196	0.267
159.	Yuma	Arizona	1	54	211'931	14'282	-2.207	0.15
160.	Decatur	Illinois	1	34	104'688	1'504	-2.209	0.32
161.	Sioux Falls	South Dakota	1	26	133'527	2'416	-2.222	0.19
162.	Odessa	Texas	1	28	162'067	2'326	-2.226	0.345
163.	Kankakee	Illinois	1	29	109'924	1'753	-2.235	0.293
164.	Prescott Valley-Prescott	Arizona	1	42	232'396	21'040	-2.242	0.175
165.	Olympia-Lacey-Tumwater	Washington	1	49	940'163	11'266	-2.243	0.124
166.	Dothan	Alabama	3	34	148'825	4'445	-2.252	0.152
167.	Jackson	Michigan	1	38	158'174	1'818	-2.254	0.272
168.	Bangor	Maine	1	46	151'696	8'800	-2.257	0.159
169.	Shreveport-Bossier City	Louisiana	1	30	156'964	1'303	-2.265	0.292
170.	Victoria	Texas	2	25	99'622	4'492	-2.269	0.304
171.	Florence	South Carolina	2	49	205'095	3'526	-2.275	0.105
172.	Hickory-Lenoir-Morganton	North Carolina	4	73	367'982	4'241	-2.284	0.114
173.	Florence-Muscle Shoals	Alabama	2	36	147'827	3'265	-2.286	0.202
174.	Terre Haute	Indiana	5	48	149'482	5'291	-2.288	0.131

Table B.4.6: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
175.	Muncie	Indiana	1	30	114'461	1'016	-2.299	0.271
176.	Decatur	Alabama	2	36	152'321	3'290	-2.307	0.155
177.	La Crosse-Onalaska	Minnesota, Wisconsin	2	30	136'838	2'600	-2.311	0.196
178.	Lawrence	Kansas	1	22	121'304	1'181	-2.325	0.301
179.	Ames	Iowa	2	27	123'736	2'964	-2.34	0.13
180.	Cheyenne	Wyoming	1	21	99'272	6'957	-2.341	0.426
181.	Johnstown	Pennsylvania	1	42	131'611	1'783	-2.348	0.077
182.	Houma-Thibodaux	Louisiana	2	45	209'277	5'958	-2.354	0.127
183.	Pittsfield	Massachusetts	1	39	125'927	2'401	-2.359	0.001
184.	Dubuque	Iowa	1	26	97'193	1'576	-2.365	0.259
185.	Ocean City	New Jersey	1	33	92'701	652	-2.372	0.152
186.	Redding	California	1	48	179'267	9'779	-2.374	0.206
187.	Kennewick-Richland	Washington	2	50	294'396	7'622	-2.374	0.218
188.	Cumberland	Maryland, West Virginia	2	30	98'049	1'948	-2.374	0.227
189.	Billings	Montana	3	40	180'641	16'777	-2.375	0.246
190.	State College	Pennsylvania	1	31	122'714	2'596	-2.375	0.186
191.	Hot Springs	Arkansas	1	20	99'043	1'756	-2.382	0.301
192.	Johnson City	Tennessee	3	44	203'147	2'212	-2.382	0.189
193.	Joplin	Missouri	2	34	178'816	3'272	-2.384	0.197
194.	Dover	Delaware	1	32	179'124	1'519	-2.39	0.259
195.	Saginaw	Michigan	1	56	191'166	2'073	-2.391	0.184
196.	Bellingham	Washington	1	34	224'538	5'457	-2.393	0.182
197.	Salem	Oregon	2	38	200'264	4'537	-2.397	0.169
198.	The Villages	Florida	1	19	643'724	4'192	-2.398	0.237
199.	Michigan City-La Porte	Indiana	1	28	110'026	1'550	-2.401	0.188
200.	Carbondale-Marion	Illinois	3	33	136'837	3'492	-2.403	0.199
201.	Salisbury	Delaware, Maryland	4	34	124'946	4'288	-2.404	0.193
202.	Parkersburg-Vienna	West Virginia	2	28	496'278	4'321	-2.409	0.245
203.	Auburn-Opelika	Alabama	1	27	163'461	1'574	-2.414	0.101
204.	Owensboro	Kentucky	3	29	845'599	4'774	-2.416	0.182
205.	Oshkosh-Neenah	Wisconsin	1	41	118'951	2'328	-2.426	0.183
206.	Mount Vernon-Anacortes	Washington	1	30	127'442	4'484	-2.432	0.228
207.	Las Cruces	New Mexico	1	41	217'696	9'862	-2.434	0.248
208.	Savannah	Georgia	1	52	273'170	1'153	-2.44	0.164
209.	St. Louis	Illinois, Missouri	1	44	162'264	2'875	-2.441	0.208
210.	Merced	California	1	49	273'661	5'012	-2.445	0.202
211.	Kahului-Wailuku-Lahaina	Hawaii	1	36	166'657	3'009	-2.445	0.236
212.	San Jose-Sunnyvale-Santa Clara	California	3	27	121'472	9'058	-2.449	0.294

Table B.4.7: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
213.	Daphne-Fairhope-Foley	Alabama	1	32	218'289	4'118	-2.452	0.148
214.	Danville	Illinois	1	24	76'704	2'327	-2.454	0.238
215.	Elizabethtown-Fort Knox	Kentucky	1	36	205'184	1'200	-2.457	0.215
216.	Lake Havasu City-Kingman	Arizona	1	43	210'998	34'476	-2.457	0.108
217.	Idaho Falls	Idaho	3	26	148'811	13'446	-2.458	0.265
218.	Chico	California	1	51	2'214'265	11'793	-2.466	0.186
219.	Wausau-Weston	Wisconsin	2	37	163'172	6'279	-2.466	0.192
220.	Burlington	North Carolina	1	36	166'144	1'098	-2.47	0.208
221.	Dalton	Georgia	2	26	143'911	1'645	-2.473	0.242
222.	Blacksburg-Christiansburg	Virginia	4	33	167'201	2'778	-2.475	0.165
223.	Lewiston-Auburn	Maine	1	28	107'958	1'212	-2.477	0.236
224.	Grand Junction	Colorado	1	29	152'962	8'622	-2.478	0.25
225.	Morgantown	West Virginia	2	32	139'806	2'613	-2.482	0.162
226.	Yakima	Washington	1	45	250'649	11'126	-2.487	0.2
227.	New Bern	North Carolina	3	28	124'416	3'927	-2.489	0.206
228.	Gainesville	Georgia	1	36	201'434	1'018	-2.493	0.212
229.	Logan	Idaho, Utah	2	28	140'072	4'736	-2.493	0.168
230.	Manhattan	Kansas	3	26	131'571	4'754	-2.498	0.175
231.	Goldsboro	North Carolina	1	26	123'785	1'433	-2.504	0.29
232.	Barnstable Town	Massachusetts	1	57	213'505	1'020	-2.505	0.187
233.	Vineland-Bridgeton	New Jersey	1	34	150'085	1'253	-2.515	0.271
234.	Kingsport-Bristol	Tennessee, Virginia	5	75	307'267	5'206	-2.522	0.147
235.	Harrisonburg	Virginia	2	30	134'696	2'245	-2.526	0.208
236.	Williamsport	Pennsylvania	1	29	114'014	3'183	-2.537	0.187
237.	Eau Claire	Wisconsin	2	31	168'307	4'265	-2.538	0.204
238.	Mansfield	Ohio	1	30	121'043	1'283	-2.541	0.232
239.	Sumter	South Carolina	2	35	650'211	6'177	-2.541	0.239
240.	East Stroudsburg	Pennsylvania	1	33	168'824	1'576	-2.542	0.23
241.	Pocatello	Idaho	2	24	94'377	6'518	-2.549	0.259
242.	Hammond	Louisiana	1	20	133'753	2'050	-2.55	0.282
243.	Jefferson City	Missouri	4	31	151'094	5'822	-2.557	0.152
244.	Jonesboro	Arkansas	2	24	132'828	3'796	-2.561	0.176
245.	Ithaca	New York	1	23	102'237	1'230	-2.561	0.229
246.	Punta Gorda	Florida	1	38	185'926	1'762	-2.564	0.218
247.	Napa	California	1	40	138'572	1'939	-2.569	0.221
248.	Missoula	Montana	1	20	119'062	6'717	-2.57	0.263
249.	Rocky Mount	North Carolina	2	32	146'356	2'709	-2.573	0.2
250.	Longview	Washington	1	24	108'399	2'953	-2.574	0.243
251.	Albany-Lebanon	Oregon	1	21	127'216	5'932	-2.576	0.181

Table B.4.8: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
252.	Farmington	New Mexico	1	33	125'608	14'279	-2.576	0.125
253.	Greenville	North Carolina	1	32	179'961	1'689	-2.578	0.215
254.	Janesville-Beloit	Wisconsin	1	38	162'532	1'860	-2.582	0.19
255.	Hinesville	Georgia	2	16	81'275	2'306	-2.587	0.368
256.	Racine	Wisconsin	1	45	195'859	862	-2.588	0.212
257.	Glens Falls	New York	2	36	125'491	4'399	-2.59	0.189
258.	El Centro	California	1	31	180'580	10'818	-2.591	0.235
259.	Sierra Vista-Douglas	Arizona	1	26	104'574	2'634	-2.591	0.27
260.	Monroe	Michigan	1	39	150'000	1'423	-2.594	0.218
261.	Niles	Michigan	1	48	153'797	1'471	-2.595	0.202
262.	Staunton	Virginia	3	24	751'615	3'604	-2.609	0.224
263.	Altoona	Pennsylvania	1	34	122'495	1'362	-2.613	0.221
264.	Rome	Georgia	1	20	97'805	1'321	-2.614	0.257
265.	Lebanon	Pennsylvania	1	31	140'410	938	-2.616	0.218
266.	Homosassa Springs	Florida	1	27	147'938	1'507	-2.617	0.232
267.	Grand Forks	Minnesota, North Dakota	2	28	101'627	8'826	-2.632	0.219
268.	Salinas	California	1	21	172'127	6'285	-2.636	0.224
269.	Kingston	New York	1	47	178'371	2'912	-2.639	0.213
270.	Weirton-Steubenville	Ohio, West Virginia	3	37	117'223	1'503	-2.642	0.229
271.	Spartanburg	South Carolina	1	32	126'442	15'970	-2.647	0.191
272.	Bismarck	North Dakota	3	25	128'589	11'089	-2.649	0.225
273.	Morristown	Tennessee	3	26	141'909	1'855	-2.651	0.191
274.	Cleveland	Tennessee	2	24	2'053'137	5'174	-2.653	0.224
275.	Hanford-Corcoran	California	1	27	151'090	3'599	-2.661	0.268
276.	Flagstaff	Arizona	1	28	142'254	48'223	-2.664	0.154
277.	Bloomsburg-Berwick	Pennsylvania	2	19	83'568	1'589	-2.67	0.209
278.	Mankato	Minnesota	2	23	101'544	3'099	-2.674	0.235
279.	Battle Creek	Michigan	1	39	133'943	1'830	-2.675	0.211
280.	Bend	Oregon	1	24	191'749	7'818	-2.676	0.22
281.	Bay City	Michigan	1	26	103'506	1'146	-2.681	0.238
282.	Elkhart-Goshen	Indiana	1	22	84'115	1'056	-2.684	0.263
283.	Sioux City	Iowa, Nebraska, South Dakota	1	27	115'152	1'325	-2.689	0.229
284.	Fond du Lac	Wisconsin	1	20	102'654	1'864	-2.699	0.233
285.	Chambersburg-Waynesboro	Pennsylvania	1	27	154'954	2'001	-2.7	0.227
286.	Corvallis	Oregon	1	18	223'344	4'239	-2.703	0.259
287.	Fairbanks	Alaska	1	19	98'455	19'006	-2.713	0.279
288.	Watertown-Fort Drum	New York	1	25	111'454	3'286	-2.714	0.246
289.	Winchester	Virginia, West Virginia	3	24	139'270	2'754	-2.715	0.215

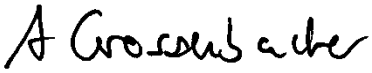
Table B.4.9: Overview of monocentric MSAs, ranked by PC_1 value.

Rank	Metropolitan Title	State(s)	Counties	Census Tracts	Population	Area (km2)	PC_1	PC_2
290.	Beckley	West Virginia	2	29	117'539	3'282	-2.72	0.211
291.	Wenatchee	Washington	2	22	119'173	12'276	-2.733	0.245
292.	Grants Pass	Oregon	1	16	87'097	4'247	-2.746	0.198
293.	Great Falls	Montana	1	22	81'576	6'989	-2.755	0.239
294.	Kokomo	Indiana	1	20	82'486	760	-2.762	0.242
295.	Gettysburg	Pennsylvania	1	23	102'627	1'344	-2.762	0.253
296.	Grand Island	Nebraska	3	19	75'592	4'146	-2.773	0.216
297.	Midland	Michigan	1	19	83'445	1'338	-2.781	0.243
298.	Twin Falls	Idaho	2	19	110'272	6'523	-2.793	0.239
299.	Madera	California	1	23	155'925	5'535	-2.798	0.243
300.	Casper	Wyoming	1	18	80'067	13'832	-2.8	0.255
301.	Walla Walla	Washington	1	12	60'785	3'290	-2.822	0.287
302.	Columbus	Indiana	1	15	2'101'543	12'423	-2.83	0.259
303.	California-Lexington Park	Maryland	1	18	113'182	926	-2.833	0.262
304.	Lewiston	Idaho, Washington	2	16	63'104	3'845	-2.839	0.273
305.	Carson City	Nevada	1	13	55'244	375	-2.905	0.267
306.	Enid	Oklahoma	1	12	61'555	2'742	-2.91	0.26

Personal Declaration

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

Binningen, 25.08.2023



Adrian Nicolas Grossenbacher