



**HAL**  
open science

# Relationships among urban characteristics, real estate market, and spatial patterns of warehouses in different geographic contexts

Renata Oliveira, Matthieu Schorung, Laetitia Dablanc

► **To cite this version:**

Renata Oliveira, Matthieu Schorung, Laetitia Dablanc. Relationships among urban characteristics, real estate market, and spatial patterns of warehouses in different geographic contexts. [Research Report] Université gustave eiffel. 2021. halshs-03369462v1

**HAL Id: halshs-03369462**

**<https://shs.hal.science/halshs-03369462v1>**

Submitted on 7 Oct 2021 (v1), last revised 24 Nov 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

## **Relationships among urban characteristics, real estate market, and spatial patterns of warehouses in different geographic contexts**

### **Postdoctoral research development**

Authors : Renata Oliveira, Matthieu Schorung, Laetitia Dablanç

This report is a part of the research initiatives performed by the Logistics City Chair. The Logistics City Chair investigates urban logistics scientific concerns focusing on urban and peri-urban logistics real estate and trends and new consumer practices and their impact on urban logistics and real estate. The Chair was launched in 2019 and is a partnership between University Gustave Eiffel (France), Sogaris, Poste Immo, Region Ile-de-France.

## Abstract

This work is a contribution to topic 1.1 of the Logistics City Chair, and two hypotheses linking urban characteristics to the spatial structure of warehouses are explored: (i) the location of warehouses is closely related to the land/rent values of logistics facilities; and (ii) logistics sprawl is higher in cities with a high differential between land/rent values in city centers and peripheral areas. For that, we have collected data for 62 metropolitan areas and have considered 53 (complete data) to analyse the urban spatial structure in each metropolitan area and the relationship among urban variables, warehouse location, and real estate rent prices. We also provide a comparative analysis of the 53 metropolitan areas. The main results are (i) it is essential to classify metropolitan areas into a typology in order to perform comparative studies; (ii) warehouse location and rent prices are related to the concentration of urban activity; (iii) logistics sprawl is not significantly related to differential warehouse rental prices, but there are differences among metropolitan types.

## Table of content

---

1. Contextualisation and objectives .....	6
2. Research approach .....	9
2.1. Data collection and treatment .....	9
2.2. Spatial characterisation of urban activity and warehouse attributes in each metropolitan area .....	12
2.2.1. Spatial characterisation of urban activities.....	12
2.2.2. Spatial characterisation of warehouse location and rent prices .....	13
2.3. Exploring the hypotheses .....	17
2.3.1. Hypothesis 1: The location of warehouses is closely related to the rent values of logistics facilities.....	17
2.3.2. Hypothesis 2: Logistics sprawl is higher in cities with a high differential between activity hubs and peripheral activity zones land/rent values .....	18
3. Results and discussion.....	20
3.1. Data collection and treatment .....	20
3.2. Spatial characterisation of urban activity and warehouses attributes in each Metropolitan Region.....	20
3.3. Exploring the hypotheses .....	32
3.3.1. Hypothesis 1: The location of warehouses is closely related to the rent values of logistics facilities.....	32
3.3.2. Hypothesis 2: Logistics sprawl is higher in cities with a high differential between central and suburban land/rent values .....	36
4. Final discussion.....	42
5. References.....	45
APPENDIX A .....	47

## List of Tables

---

Table 1: Data sources for warehouse location, size, and rent prices .....	11
Table 2: Classification criteria for activity hubs and peripheral activity zones hexagon bins in each metropolitan region .....	13
Table 3: Relative classification criteria for warehouse location and rent prices in each metropolitan region.....	14
Table 4: Results and analytical elements for the exploratory spatial and nonspatial analysis .....	17
Table 5: Discrete data on each hexagon .....	21
Table 6: Continuous data on each metropolitan area .....	23
Table 7: Classification of cities regarding the differential rent prices and location of warehouses..	26
Table 8: Classification of cities regarding the logistics sprawl per year.....	29
Table 9: <i>Number of hexagons classified by the concentration of warehouses and urban classification</i> .....	32
Table 10: Number of hexagons classified by average prices of warehouses and urban classification .....	33
Table 11: Number of hexagons classified by the number of warehouses and warehouse average prices .....	34
Table 12: Classification of metropolitan areas.....	36
Table 13: Classification of metropolitan areas.....	38

## List of Figures

---

Figure 1: Methodological approach .....	19
Figure 2: Representation of the number of warehouses and the average rent price in each hexagon for Paris .....	21
<i>Figure 3: Representation of warehouse average rent prices for each metropolitan area.....</i>	<i>25</i>
<i>Figure 4: Representation of warehouse average rent prices for AH and PAZ for each metropolitan region.....</i>	<i>27</i>
<i>Figure 5: Representation of proportional rent prices differential for warehouses in each location for each metropolitan region.....</i>	<i>28</i>
<i>Figure 6: Representation of average distance to gravity center for t0 and t1, for each metropolitan region.....</i>	<i>30</i>
<i>Figure 7: Representation of yearly logistics sprawl.....</i>	<i>31</i>
Figure 8: Warehouse count x urban classification .....	33
Figure 9: Warehouse average prices x urban classification .....	34
Figure 10: Warehouse count and warehouse average prices.....	35
Figure 11: Differential classes and a scatterplot for DWP and YLS.....	37
<i>Figure 12: Density plot for YLS for each differential price classification .....</i>	<i>38</i>
<i>Figure 13: Density plot for DWP each differential price classification .....</i>	<i>39</i>
<i>Figure 14: Exploratory linear regression models for differential prices classes .....</i>	<i>41</i>

## 1. Context and objectives

The outsourcing of logistics activities has driven the emergence of a metropolitan logistics real estate market. Mostly, urban land and floor space scarcity and economies of scale have relocated logistics facilities towards less dense and more peripheral areas of cities (Cidell, 2010; Dablanc et al., 2014; Dablanc and Browne, 2019; Dablanc and Rakotonarivo, 2010; Heitz et al., 2017; Sakai et al., 2020). This process, known as 'logistics sprawl,' has compromised urban sustainability, livability, and economic growth (Dablanc et al., 2014; Dablanc and Browne, 2019; Dablanc and Rakotonarivo, 2010; Heitz et al., 2017). Many case studies of logistics sprawl (LS) and warehouses' locations in large urban areas have emerged in the past ten years. Investigations concerning urban form and function coordination, urban logistics, and real estate strategy have interested public and private actors.

The Logistics City Chair is focused on two scientific themes: (i) urban logistics real estate; and (ii) trends and new practices in consumption, production, and distribution impacting urban logistics and logistics real estate. This study is part of the first theme and included in the Chair's objective: "Logistics sprawl and urban logistics: analysis of territorial dynamics linked to the evolution of the location of logistics activities, at the 'macro' level" ("Logistics City Chair : Research," 2021).

Through this research, as a contribution to topic 1.1 of the Logistics City Chair, we intended to explore two of theme 1.1 hypotheses linking urban characteristics to the spatial structure of warehouses:

- i. The location of warehouses is closely related to the land/rent values of logistics facilities.
- ii. Logistics sprawl is higher in cities with a high differential between land/rent values in city centers and peripheral areas.

In other words, we increment the meta-analysis performed before by the Chair's team (Dablanc and Palacios-Argüello, 2020<sup>1</sup>) to compare the spatial patterns of warehouses and respective rent prices

---

<sup>1</sup> <https://www.lvmt.fr/wp-content/uploads/2021/09/Dablanc-Palacios-Arguello-2020.pdf>

practised by real estate agents in different cities around the world concerning the urban structure. This can also help understand the relationship between the evolution in the number and location of logistics facilities over time and the differential warehouse prices in activity hubs and peripheral activity zones.

To develop this research, we have considered the previous data collection performed by the Logistics City Chair concerning the metropolitan areas where logistics sprawl was investigated by research teams and published in scientific journals. In this previous dataset (Dablanc and Palacios-Argüello, 2020), there was information regarding Centographic logistics sprawl measures, timeframe and sources of the data collection, the population in the timeframes, metropolitan administrative information, information on the spatial structure of the metropolitan areas, its importance as a gateway in a regional scale, and aggregated information on logistics facilities rent prices. The meta-analysis (Dablanc and Palacios-Argüello, 2020) considered 74 case studies (metropolitan regions that have been studied in the literature regarding warehouses' location).

Additional data were collected to understand the relationships among logistics real estate rent prices and logistics facilities' spatial structure. We have managed to gather real estate data (rent prices) for 62 metropolitan areas worldwide. We could not consider all metropolitan areas for some of the analyses since some information was missing or there were no public rent prices or addresses for logistics facilities on real estate websites. This limitation is presented throughout this report whenever it occurs. For data collection in real estate websites, individual scripts were developed to scrape the warehouse information out of the *Html* file and assemble one dataset for each metropolitan area. For more than one metropolitan area, different websites were explored and, if considered for the data collection, are presented in this report.

This research's contribution is first methodological: developing a framework with homogeneous spatial units (hexagon bins) and disaggregated normalised data for comparing metropolitan regions considering the spatial pattern of logistics facilities, logistics real estate rental prices, and urban characteristics. The disaggregated information was obtained from the open data source OpenStreetMap (OpenStreetMap contributors, 2020). Understanding the dependence among logistics sprawl, real estate market, and urban attributes in the metropolitan regions under



investigation is another contribution of this work. To our knowledge, no published research has explored the relationship between rent prices, urban structure, and logistics sprawl.

Coordinating these dimensions is essential to support urban logistics stakeholders' needs, cities' livability, and the real estate market. This research brings an innovative framework for comparing metropolitan areas within different geographic contexts based on spatial analyses and disaggregated data. This work's results are reproducible and can induce local and regional public authorities to develop a more effective public policy addressing logistics land use and transportation planning.

## 2. Research approach

The methodological steps of this are based on two primary analyses: (i) the analysis of the urban spatial structure in each metropolitan area and the relationship among urban variables, warehouse location, and real estate rent prices; and (ii) a comparative analysis among the metropolitan areas under investigation, considering a meta-analysis of logistics sprawl in published studies. It is essential to highlight the previous meta-analysis regarding the Chair Logistics City research agenda (Dablanc, Palacios-Argüello, 2020a).

### 2.1. Data collection and treatment

Spatialized information was necessary to perform the investigation of the hypotheses designed for this work. Therefore, locational data concerning the urban areas' spatial structure and geographic data, along with logistics real estate information, were collected as the first step of this methodological proposal.

Initially, data concerning **geographic information, road infrastructure, and the location of points of interest** were collected. The data sources were the Openstreetmap database (OpenStreetMap contributors, 2020) for points of interest, road networks, and local public data agencies for demographic and geographic information. Points of interest are sites that accommodate different urban activities, such as touristic points, archaeological sites, ATMs, banks, bakeries, bars, restaurants, grocery and other food retailers, beauty shops, book shops, general retailing, medical care centers, and other health care facilities, dentists, parks, fire stations, florists, urban build environments elements, like benches and fountains, graveyards, hotels, libraries, museums, theaters, playgrounds, post offices, schools, toilets, wastebaskets, and other urban facilities. We use points of interest as a proxy for urban density. There are no facilities that concern logistics activities within the POI dataset for the metropolitan areas. The municipality boundaries were obtained from official statistics and geography institutes of each country and region.

OpenStreetMap (OSM) is an open database containing Voluntary Geographic Information (VGI). It is under the Open Database License and contains worldwide data, primarily collected and maintained by volunteers. OSM data is disaggregated and presents advantages over public authorities' data because they are continuously maintained by volunteers and are available for

different metropolitan areas. Despite the differences in data quality within regions of the world, the potential of OSM data to explore the concentration of urban activity is significant. The potential and quality analysis of the contributed street networks, Points of Interest, and the land use data reveal the vast potential of OSM data for various applications (Bakillah et al., 2014; Jokar Arsanjani et al., 2015; Klinkhardt et al., 2021; Mobasheri et al., 2017; OpenStreetMap contributors, 2020; Zhang and Pfoser, 2019).

Most of the nonspatial data of logistics facilities had been organised by the Logistics City Chair in previous work (Chair Logistics City, 2019; Dablanc and Palacios-Argüello, 2020; Palacios-Argüello and Dablanc, 2020). Additionally, data concerning rent prices of warehouses with locational information for the metropolitan areas have been collected from secondary sources directly or scraping reports and websites in this work.

Thus, we have consolidated and organised urban spatial information on all 74 metropolitan regions considered in theme 1.1 of Logistics City Chair previous works. Two metropolitan scales (therefore, two data sets) were considered for three metropolitan areas: Paris, Toronto, and Gothenburg; for the logistics real estate data collection, we have considered only the legal metropolitan area. Hence, we searched for real estate websites for each metropolitan area out of the 71 selected in the literature. For all the metropolitan areas, different real estate websites were explored. Some of the websites did provide prices for very few warehouse prices, and all the available information was collected. For Gothenburg, for instance, the websites did not provide information on the prices on the pages but only through consultation. The same happened for Colombian metropolitan areas. We could not find available warehouse price information for Tokyo and Chinese metropolitan areas. These case-studies will be explored in a subsequent study, which will also increase the number of European cities in the dataset.

Besides the unavailability of data sources for warehouse real estate information for eight metropolitan areas, another limitation was the spatial resolution of logistics real estate information for warehouses in the European metropolitan areas of Berlin, Brussels, and Bordeaux. For these cities, the collected data on warehouse rent prices was not enough for spatial differentiation. We have explored different resolutions (hexagons of different sizes) to represent this information.

However, since no spatial differentiation was possible for logistics real estate rent prices, we could not consider Berlin, Brussels, and Bordeaux for the differential analysis. Also, we did not consider Singapore, Montreal, Calgary, Halifax, Winnipeg, Zuid Holland, Flevoland, Noord Holland, and Utrecht due to incomplete data. Therefore, a dataset of complete data was consolidated for 53 metropolitan regions in the comparative methodological step. The dataset was composed of the information previously collected through the meta-analysis and the variables (i) warehouse rent price, (ii) location, and (iii) size, gathered through consultation on the sources presented in **Erreur ! Source du renvoi introuvable..**

**Table 1: Data sources for warehouse location, size, and rent prices**

<b>Metropolitan area location</b>	<b>Link</b>
USA	<a href="http://www.loopnet.com">www.loopnet.com</a>
Canada	<a href="http://www.loopnet.com">www.loopnet.com</a>
Brazil	<a href="http://www.zapimoveis.com.br">www.zapimoveis.com.br</a>
Netherlands	<a href="http://www.fundainbusiness.nl/en/">www.fundainbusiness.nl/en/</a>
Singapore	<a href="http://www.iproperty.com.sg/rent/warehouse/">www.iproperty.com.sg/rent/warehouse/</a>
France	<a href="http://bureaux-commerces.seloger.com/">bureaux-commerces.seloger.com/</a>

One of this work's contributions is the disaggregated spatial structure exploration from open data from OpenStreetMap (OpenStreetMap contributors, 2020) to classify the spatial patterns of warehouse location, rent prices, and urban structure. This disaggregation and standardisation of spatial units allow the comparison among metropolitan areas. Therefore, we have assembled all the data into a grid of hexagon cells with a short diagonal of five kilometers. Section road lengths were aggregated in each cell, and the number of points of interest was computed for these spatial units without classifying the activities.

A standard hexagon size was considered to compare metropolitan areas and eliminate the need to transform variables into area-related (density) ones. The ability to communicate and analyse the phenomena is stronger if the attributes are not calculated considering the area of the spatial units (such as density measures), especially in comparative studies.

## **2.2. Spatial characterisation of urban activity and warehouse attributes in each metropolitan area**

This methodological step aims at organising derivative indicators to categorise city centers (“urban activity hubs”) and peripheral areas (“peripheral activity zones”) and warehouse location and respective rent prices. Therefore, we characterise the urban spatial structure and warehouses available for rental at the time of data collection.

### **2.2.1. Spatial characterisation of city areas and urban activities**

For this characterisation to allow the subsequent comparison between metropolitan regions, it was necessary to standardise the spatial units and the information. Therefore, the chosen scale of analysis and spatial units (cells in grids) for this study were hexagons with a dimension of 5 kilometers in the minor diagonal (Ben-Joseph and Gordon, 2000; Birch et al., 2007; Crown et al., 2018; Ganis, 2015; Whitehand et al., 1996). The urban variables transformed into cells were the number of points of interest (POI) and the sum of road extensions. The points of interest were not stratified according to activities' categories since the objective of collecting this information and the location of the road infrastructure was to identify the centralities of each metropolitan region. Centralities are areas with higher intensity of activities and connectivity in relation to the spatial distribution of these facilities (Sarkar et al., 2020).

The territory of each metropolitan region, divided into hexagonal bins, was overlaid with the spatial structure of POIs and highways. For each cell, the length of highways (Quinn, 2013) and the number of POI contained in that unit were then calculated to consolidate the spatial differentiation among metropolitan areas. We do not use population density in the method for two reasons: (i) population information is aggregated in zones. These zones are usually more extensive than the hexagons. We would have a different aggregation level of this information for each metropolitan area, leading to difficulties in comparing metros; (ii) the other issue is that we would have to work with density variables and not absolute concentration. With homogeneous zoning, we could be more straightforward with the variables considered.

We then have normalised the number of POI and the road extension in each hexagon and **composed a relative indicator of urban activity intensity**, namely **Urban Activity Index (UAI)**, as a result of the sum of the **normalised variables for each hexagon  $\lambda$** . For the normalisation of the variables, we considered a min-max approach. We have then considered outliers (3.0 hinge) to differentiate the metropolitan spatial structure, considering each cell's UAI to categorise areas within the metropolitan regions (Table 2). Lower outliers and UAI lower than 95% of the overall information in each metropolitan area (percentile) were considered **peripheral activity zones (PAZ)**. UAI values higher or equal to the top 5% (percentile) and upper outliers were considered **activity hubs (AH)**.

**Table 2: Classification criteria for activity hubs and peripheral activity zones hexagon bins in each metropolitan region**

<b>Classification (percentiles)</b>	<b>Category</b>
Lower outlier	Peripheral activity zones
< 95%	Peripheral activity zones
>= 5 %	Activity hubs
Upper outlier	Activity hubs

### **2.2.2. Spatial characterisation of warehouse location and rent prices**

The same rationale adopted for the urban structure characterisation was performed for warehouse location and rent price. The variables were: **address and the number of warehouses for rental in the sources considered for this work, the average warehouse price (US\$/m<sup>2</sup>/year)**. The addresses for warehouses available for rent were geocoded and spatialised.

Rented warehouses may have a slightly different spatial structure in comparison to warehouses in general. For example, logistics facilities for rental are presumed to be newer, more modern, and more peripheral. This difference could not be verified for the data collected since disaggregated comprehensive data on all warehouses in each metropolitan area were not available. Due to the absence of comprehensive information on warehouse rent prices for different metropolitan areas and observing the limitations mentioned above, this study concerns the spatial structure of the

warehouses available for rental as the spatial structure of logistics facilities in each metropolitan region.

Also, no information on the real estate websites could differentiate warehouses and business facilities (*locaux d'activité*). Therefore, we performed an outlier analysis and excluded values classified as outliers to avoid considering these last facilities into the dataset. All the outliers and extreme values were identified for the warehouse concentration in the hexagons and the average prices. We have considered that values above  $Q3 + (1.5 \times IQR)$  or below  $Q1 - (1.5 \times IQR)$  were outliers. Values above  $Q3 + (3 \times IQR)$  or below  $Q1 - (3 \times IQR)$  were considered as extreme points (or extreme outliers). We then categorised the hexagons with non-zero observations according to Table 3.

**Table 3: Relative classification criteria for warehouse location and rent prices in each metropolitan region**

Classification (quantiles)	Category
< 25 %	Low
25 % - 50 %	Medium
50 % - 75%	Medium
> 75 %	High

In order to explore the hypotheses designed for this study, the average difference between warehouse rental prices in **activity hubs** and **peripheral activity zones** of the metropolitan areas was also calculated from the classification of the hexagonal bins, aiming at the comparability between the investigated areas. This indicator was entitled **Differential Warehouse Rent Prices (DWP)**, and concerns a continuous variable calculated from the ratio between the average warehouse rent price in **activity hubs (AH)** and **peripheral activity zones (PAZ)** (hexagon bins) (Equation 1).

$$DWP = \frac{\text{Average warehouse rent price}_{\text{AH}}}{\text{Average warehouse rent price}_{\text{PAZ}}} \quad (1)$$

Due to the lack of information for different timeframes concerning the urban activity, warehouse availability, and real estate information, we must state that these variables are static in time and do not allow computation of a logistics sprawl measure.

Since the timeframes considered in previous logistics sprawl studies are different (Centographic analysis of differentiation concerning warehouse location in metropolitan areas), the **Yearly Logistics Sprawl (YLS)** had been computed. The YLS was the difference between the more recent average distance to the barycenter (MDB) for logistics facilities and the previous measure (from the spatial structure of warehouses) divided by the period between these statistics (Equation 2).

$$YLS = \frac{MDB_{t1} - MDB_{t0}}{\text{time between } t0 \text{ and } t1} \quad (2)$$

The consolidation of the indicators presented before resulted in the composition of two datasets at different scales: (i) a dataset of hexagonal bins for each metropolitan region and (ii) a dataset of summary indicators for all metropolitan regions. The results of this exploratory spatial and nonspatial analysis of the consolidated data are presented in Table 4. The data collection, cleaning, and processing were performed in Free Open Source Software (FOSS): The R project 4.0.5 and GeoDa 1.14.0.

A graphical user interface (GUI) dashboard for visualising the results for each metropolitan area was deployed to communicate the results of this work. The R language and the package Shiny were the primary tools to build the web-based dashboard. The link to access the dashboard is [https://geodatascience.shinyapps.io/logistics\\_city\\_chair/](https://geodatascience.shinyapps.io/logistics_city_chair/).



**Methodological note:**

→ This analysis, in particular the price differential for renting warehouses, is based on a sample of prices from specialized websites that is only relatively representative. We do not have a complete database that lists all real estate and land values for all warehouses in a given geographic area.

→ This initial research will require further study, in particular by integrating more cases from Asian cities (in particular Japanese, Chinese and South Korean cities) and European cities, in order to obtain a more complete sample for a truly international analysis. The sample used in this research is very unbalanced in terms of North American cases.

→ The definition of metropolitan areas for mapping and data analysis (number of warehouses, prices) required delineation choices. These delimitations can in some cases modify the cartographic rendering and the final analysis, depending on the size of the metropolitan area for example.

→ The price differential between centers and peripheries may appear counter-intuitive in a number of cases (New Orleans, Portland). Where peripheral warehouses are more expensive than warehouses located in the central or pericentral area. We assume that in these cases, the logistics infrastructure in the city center is aging or obsolete and therefore less expensive, whereas in the peripheral areas, it is mostly newer warehouses that are more expensive.

**Table 4: Results and analytical elements for the exploratory spatial and nonspatial analysis**

Item	Variables
A dataset containing all hex bins for all the metropolitan regions investigated	<ul style="list-style-type: none"> <li>- Urban activity index</li> <li>- Categorical variable to determine activity hubs and peripheral activity zones for each metropolitan area</li> <li>- Number of warehouses in each hex bin</li> <li>- Average warehouse price in each hex bin</li> <li>- Classification and outlier's identification for warehouse count and rent prices (categorical variable)</li> </ul>
A dataset containing the synthetic indicators for each metropolitan region	<ul style="list-style-type: none"> <li>- Population (t0 and t1) (Dablanc and Palacios-Argüello, 2020)</li> <li>- Metropolitan territorial area Dablanc and Palacios-Argüello, 2020)</li> <li>- Number of municipalities (Dablanc and Palacios-Argüello, 2020)</li> <li>- Number of warehouses (t0 and t1) (Dablanc and Palacios-Argüello, 2020)</li> <li>- Average distance to barycenter (t0 and t1) (Dablanc and Palacios-Argüello, 2020)</li> <li>- Binary variable for logistics sprawl (Dablanc and Palacios-Argüello, 2020)</li> <li>- Yearly logistics sprawl (YLS) from centrographic measures (continuous variables from the previous meta-analysis)</li> <li>- Differential rent prices (DWP) (continuous variable calculated considering the category of activity hubs/peripheral activity zones and average rent prices)</li> </ul>

### 2.3. Exploring the hypotheses

#### 2.3.1. Hypothesis 1: The location of warehouses is closely related to the rent values of logistics facilities

To explore this first hypothesis, we have related three categorical variables: the **number of warehouses**, the **average rent values**, and (iii) the **urban activity index**. All three variables were computed for the hexagon bins for all metropolitan areas. We then visualise the relationship and perform **Chi-square tests** to understand if these variables are dependent on each other. It is essential to highlight that 53 metropolitan areas were considered, and only complete data, where no missing values, were included in the analysis. The **Pearson's Chi-squared test** was done with simulated p-values (2000 replications) and a significance level of 5% to understand the **dependence** between the **urban classification and warehouse location and rent prices**.

### **2.3.2. Hypothesis 2: Logistics sprawl is higher in cities with a high differential of land/rental prices between urban and peripheral areas**

In this step, we develop different analyses to understand the relationship between urban structure and warehouse location and rent prices and address the second hypothesis proposed for this research: **logistics sprawl is higher in cities with a high differential between city centres (more exactly “activity hubs”, as defined above) and peripheral activity zones land/rent values.**

We have performed a **Pearson correlation analysis** (coefficient and test) to understand the relationship between **differential warehouse rent prices** and **yearly logistics sprawl**. Different representations of the variables were also performed. The metropolitan areas were classified according to the sprawling phenomenon for logistics facilities and the differential warehouse prices.

In Figure 1, we present the methodological approach developed for this work synthetically.

## 1. Data collection

### Urban structure

- poi - Points of Interest
- r - road network

### Warehouse

- wh – location of warehouse for rental
- wh\_price - warehouse rent price
- wh\_size - warehouse size

## 2. Spatial characterization of urban activities and warehouse attributes

### Spatial characterization of urban activities

- Transformation of variables into hexagonal homogeneous grid
- Calculate and represent measures of urban activity index as the sum of normalized (max-min) number of POIs and road extension for each hexagon  $\lambda$ .
  - Lower outliers and UAI lower than 95% of the overall information in each metropolitan area were considered as suburban areas.
  - UAI values higher or equal to than 5% and upper outliers were considered as central areas.
- Warehouse attributes in each hexagon were classified according to seven categories: extremely low, very low, low, medium, high, very high and extremely high.
- Calculate the rent prices differential values (continuous variable calculated considering the ratio of suburban/central for each hexagon bin)
- Consider yearly logistics sprawl (continuous variables from previous meta-analysis)

### Spatial characterization of warehouse location and rent prices

- Spatial representation of each variable, in each hexagon, within the metropolitan areas.
- Activity Hubs (AH) and Peripheral Activity Zones (PAZ) areas determined from the intensity of Points of interest and road connections (urban activity index)
- Dashboard with the representation of warehouse location, average rent prices and the urban classification regarding the UAI
- Dataset with variables for each metropolitan area with: (i) Urban Activity Index (numeric); (ii) urban classification regarding the UAI (categorical); (iii) warehouse attributes regarding the number of facilities and warehouse average rent price in each hexagon bin (categorical and numeric);
- Dataset with aggregated variables for all metropolitan areas with: (i) differential warehouse rent prices (numeric); (ii) logistics sprawl (numeric).

### Hypothesis 1: The location of warehouses is closely related to the land/rent values of logistics facilities.

Number of warehouses in each bin in relation to average rent values in the same bin

- In hexagons for all metropolitan areas with complete data.
- All bins from all metropolitan areas in the same analysis
- Categorical data
- Chi-square test.

### Hypothesis 2: Logistics sprawl is higher in cities with a high differential between central and suburban land/rent values.

Differential of warehouse rent prices (continuous variable calculated considering the ratio AH/PAZ) x yearly logistics sprawl (continuous)

- all bins in the same metropolitan area (metropolitan scale)
- Pearson correlation coefficient/Pearson correlation test.

Figure 1: Methodological approach

### 3. Results and discussion

This section presents the results and discussion concerning the methodological approach designed and described in chapter 2.

#### 3.1. Data collection and treatment

The results of this methodological step are presented together with the following steps since, after collection and treatment, data were gathered into the proposed spatial units.

#### 3.2. Spatial characterisation of urban activity and warehouses attributes in each metropolitan region

The main results of this methodological step are **the organisation of data into standard spatial areal units for all metropolitan regions**, the spatial characterisation of the urban activity was determined through the urban activity index – UAI, calculated from the sum of the normalised number of Points of Interest and road network length within each hexagon of each metropolitan area. The disaggregated spatial characterisation of warehouse location and rent prices was determined from the data collected, and, therefore, the generation of essential variables for answering the research questions proposed:

- (i) urban classification (Activity Hubs, Peripheral Activity Zones), as presented in Table 2 for the Urban Activity Index (UAI);
- (ii) warehouse location and rent prices classification according to Table 3;
- (iii) differential rent prices related to the categories through the UAI;
- (iv) yearly logistics sprawl from a previous meta-analysis (Palacios-Argüello and Dabanc, 2020a).

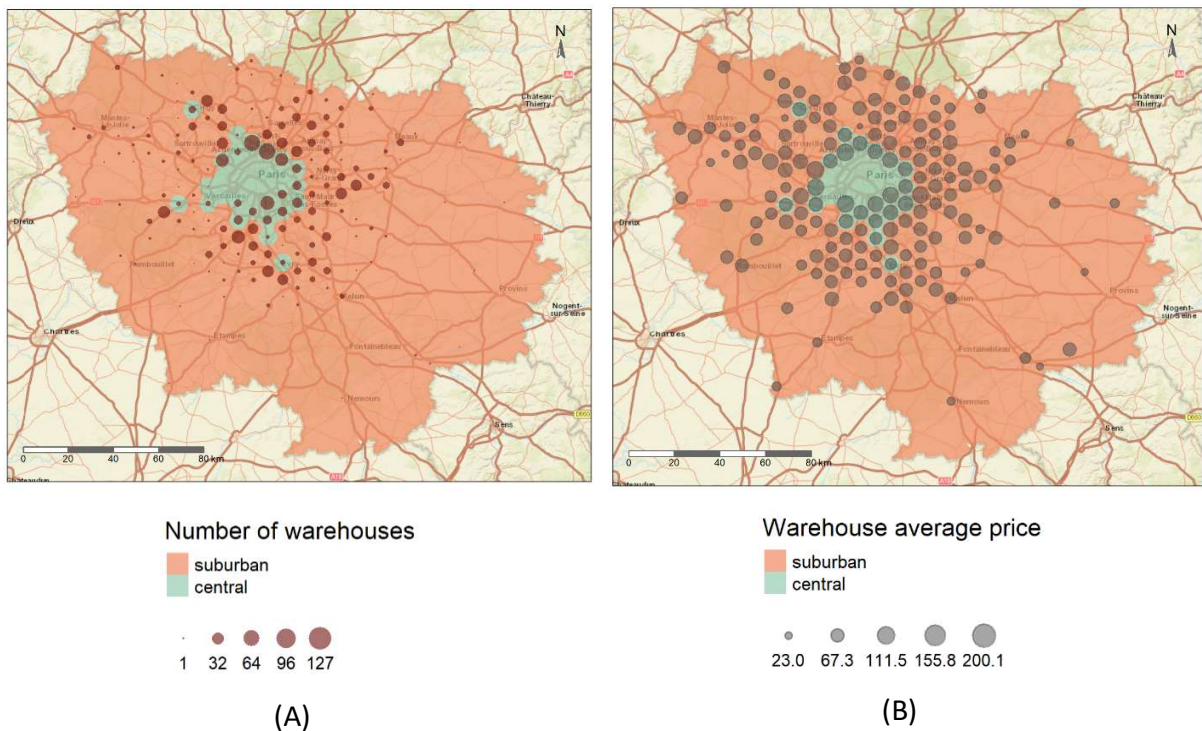
Concerning the 53 metropolitan areas, 42,443 hexagon bins are considered the sample for analysing the relationship between the location of warehouses and the rent values of logistics (first hypothesis). A header of this data (a sample to present the data structure) is presented in **Table 5**. In this example, we present the Paris region composed of 629 hexagon bins. It has 167 hexagon bins with warehouses in the collected data.

**Table 5: Discrete data on each hexagon (example)**

id	Metro	Number of warehouses *	Warehouse average rent price (\$/sq m/y)	Urban activity index	Urban activity index class	Number of warehouses class	Warehouse rent prices class
15	paris	5	112	0.23	PAZ	Medium	Medium
19	paris	1	99	0.19	PAZ	Low	Medium
29	paris	11	143	0.29	PAZ	Medium	High
43	paris	11	95	0.49	PAZ	Medium	Medium

\*In the real-estate website used

The dashboard available in [https://geodatascience.shinyapps.io/logistics\\_city\\_chair/](https://geodatascience.shinyapps.io/logistics_city_chair/) contains the spatial representation of the above-stated variables. To illustrate the data on the hexagon level, we present the Paris region in Figure 2 (A and B), respectively, for the number of warehouses and warehouse average rent price (US\$/m<sup>2</sup>/year). The maps for the other metropolitan areas are presented in Appendix A.



*Figure 2: Representation of the number of warehouses and the average rent price in each hexagon for Paris (region Ile-de-France)*

The data previously organised by the Logistics City Chair is presented in **Table 6**, and the logistics sprawl in each metropolitan area is represented in **Erreur ! Source du renvoi introuvable**. (Dabanc and Palacios-Arguello, 2020)). Additional variables resulted from the data collection

performed in this study concern: the average rent price in AH, the average rent prices in PAZ, the price differential, and the sprawl per year. In Table 6, we present the synthetic continuous data for the investigation of the hypotheses.

Brazilian cities (Belo Horizonte and São Paulo) have the lowest average warehouse rent prices, with 30.02 and 36.21 US\$/m<sup>2</sup>/year, respectively. Paris (106.06 US\$/m<sup>2</sup>/year) stands out, along with North American cities, which present the highest average value for this indicator. Metropolitan areas with the highest average warehouse rent prices are San Francisco, Los Angeles, San Diego, Miami, New York, Vancouver, Seattle, Austin, Washington, Boston, Denver, Toronto, Portland, Nashville. The average warehouse rent price in San Francisco is 178.49 US\$/m<sup>2</sup>/year.

**Table 6: Data on each metropolitan area**

Metropolitan area	Population 2000	Population 2015	Area	Year t0	Year t1	Time	# of WH t0	Average dist to gravity center t0	# of WH t1	Average dist to gravity center t1	Logistics sprawl	# of WH real estate	Average price city centre	Average price peripheral area	Diff price	Sprawl per year
albany	825920	879085	7607	2003	2013	10	54	27.87	48	28.32	0.45	46	59.24	69.75	0.85	0.04
atlanta	2621089	3603409	22890	1998	2008	10	132	28.65	401	32.99	4.34	163	83.75	77.43	1.08	0.43
austin	1041753	1534758	11151	2003	2013	10	38	16.17	50	18.22	2.05	95	122.01	121.09	1.01	0.20
belo horizonte	780000	5000000	9460	1995	2015	20	44	17.8	156	19	1.2	1040	32.08	28.88	1.11	0.06
birmingham	1051340	1144683	13913	2003	2013	10	47	20.6	51	24.54	3.94	38	66.65	68.97	0.97	0.39
boston	2634378	2872310	11906	2003	2013	10	290	45.11	294	47.72	2.61	141	156.34	106.01	1.47	0.26
buffalo	1170022	1133002	6263	2003	2013	10	57	16.29	57	16.29	0	36	63.62	60.13	1.06	0.00
charlotte	2091897	2712974	13407	2003	2013	10	124	22.9	145	25.97	3.07	117	95.07	71.13	1.34	0.31
chicago	2676215	3202509	25198	1998	2013	15	217	71	415	79.8	8.8	556	95.94	91.16	1.05	0.59
cincinnati	642221	683709	11043	2003	2013	10	112	23.5	122	23.06	-0.44	129	61.86	61.25	1.01	-0.04
cleveland	1978890	1898436	10458	2003	2013	10	148	37.5	150	37.55	0.05	157	56.03	60.27	0.93	0.00
columbus	1625491	1785971	12667	2003	2013	10	208	21.92	195	22.4	0.48	94	63.09	51.18	1.23	0.05
dallas	2846428	4321973	24982	2003	2013	10	338	30.8	402	29.98	-0.82	222	98.53	90.33	1.09	-0.08
dayton	269526	271791	3368	2003	2013	10	54	21.58	49	28.71	7.13	67	50.47	51.24	0.98	0.71
denver	2316068	2908463	21909	2003	2013	10	118	17.11	147	16.43	-0.68	198	113.61	111.48	1.02	-0.07
detroit	2201458	2350511	11190	2003	2013	10	196	31.53	210	36.55	5.02	309	66.85	74.54	0.90	0.50
grand rapids	250981	301251	10056	2003	2013	10	62	22.16	72	24.53	2.37	28	60.82	62.97	0.97	0.24
greensboro	720580	842338	5235	2003	2013	10	88	31.58	88	27.91	-3.67	39	52.23	42.06	1.24	-0.37
greenville	453531	566978	7220	2003	2013	10	101	32.03	97	28.82	-3.21	53	52.60	39.31	1.34	-0.32
houston	1193312	1762483	24640	2003	2013	10	221	23.69	298	26.15	2.46	365	94.39	100.72	0.94	0.25
indianapolis	824209	1033843	11328	2003	2013	10	121	28.9	171	28.9	0	46	68.65	72.39	0.95	0.00
kansas city	729993	796646	19199	2003	2013	10	159	19.91	153	24.37	4.46	172	77.93	75.76	1.03	0.45
las vegas	23541	22513	20889	2003	2013	10	51	10.73	80	20.54	9.81	89	96.65	78.79	1.23	0.98
los angeles	12365597	13234696	14761	1998	2009	11	220	41.69	515	51.44	9.75	666	174.91	160.50	1.09	0.89
louisville	1135588	1244745	9464	2003	2013	10	81	17.51	89	17.19	-0.32	42	65.10	55.13	1.18	-0.03
miami	5007956	5969135	16245	2003	2013	10	193	40.64	235	46.54	5.9	701	144.86	143.01	1.01	0.59
milwaukee	1401336	1459422	8798	2003	2013	10	101	23.74	98	30.98	7.24	53	52.60	56.04	0.94	0.72
nashville	512532	559425	16526	2003	2013	10	116	28.29	121	30.37	2.08	55	126.87	90.74	1.40	0.21



Metropolitan area	Population 2000	Population 2015	Area	Year t0	Year t1	Years	# of warehouses t0	Average distance to gravity center t0	# of warehouses t1	Average distance to gravity center t1	Logistics sprawl	# of warehouse real estate	Average price on AH	Average price on PAZ	Diff price	Sprawl per year
new orleans	1337740	1260281	20541	2003	2013	10	77	19.44	83	29.89	10.45	16	67.20	96.88	0.69	1.04
new york	14983782	16118232	28500	2003	2013	10	938	41.78	914	45.05	3.27	837	163.92	136.03	1.21	0.33
orlando	3442581	3985594	10512	2003	2013	10	75	19.7	91	19.33	-0.37	120	100.39	92.71	1.08	-0.04
paris	11360000	11900000	12012	2004	2012	8	713	17.4	955	21.5	4.1	3109	143.85	100.33	1.43	0.51
philadelphia	4264068	4564258	12717	2003	2013	10	288	33.89	340	35.9	2.01	205	89.78	85.85	1.05	0.20
phoenix	3251884	4578519	37834	1998	2015	17	41	17.86	183	20.6	2.74	276	98.88	95.38	1.04	0.16
pittsburg	1787955	1711755	13975	2003	2013	10	92	25.6	98	27.63	2.03	80	64.47	68.52	0.94	0.20
portland	1161090	1365871	18352	2003	2013	10	160	24.43	163	25.64	1.21	147	108.60	117.66	0.92	0.12
raleigh	781161	1238938	5563	2003	2013	10	76	31.16	77	31.3	0.14	42	114.67	92.13	1.24	0.01
richmond	990282	1186339	12204	2003	2013	10	58	18.83	87	20.28	1.45	23	80.66	81.28	0.99	0.15
rochester	259409	263632	13891	2003	2013	10	45	25.88	48	25.31	-0.57	44	62.09	51.09	1.22	-0.06
saint louis	2148575	2311690	21032	2003	2013	10	148	23.91	144	21	-2.91	97	71.12	75.17	0.95	-0.29
salt lake city	939169	1164912	21169	2003	2013	10	88	23.11	117	23.42	0.31	51	102.21	72.29	1.41	0.03
san antonio	1747863	2405335	19250	2003	2013	10	47	13.89	67	18.33	4.44	98	90.78	114.28	0.79	0.44
san diego	2813839	3280850	11730	2003	2013	10	84	21.68	86	20.74	-0.94	140	148.64	152.45	0.98	-0.09
san francisco	4123734	4647924	8897	2003	2013	10	305	56.07	349	57.29	1.22	244	194.17	172.98	1.12	0.12
sao paulo	15080000	21600000	7944	1992	2017	25	228	17.5	2066	17.6	0.1	1635	46.65	35.38	1.32	0.00
seattle	622023	789074	17230	1998	2009	11	85	20.6	212	19.31	-1.29	136	138.59	113.12	1.23	-0.12
tampa	2396038	2983928	8740	2003	2013	10	63	16.91	79	16.66	-0.25	119	97.51	88.11	1.11	-0.03
toronto	2432492	2745281	7930	2002	2012	10	165	16.7	228	17.9	1.2	43	128.07	109.19	1.17	0.12
tucson	843702	1009103	23847	2003	2013	10	33	56.66	55	45.8	-10.86	28	81.60	81.38	1.00	-1.09
tulsa	822560	939783	16741	2003	2013	10	39	19.6	37	23.63	4.03	43	62.02	66.78	0.93	0.40
vancouver	1536734	1776325	3034	2002	2012	10	135	12.8	134	17	4.2	26	145.05	129.45	1.12	0.42
virginia beach	1931738	2106945	10199	2003	2013	10	90	29.72	98	27.6	-2.12	20	88.41	96.88	0.91	-0.21
washington	4709434	5720217	17391	2003	2013	10	285	51.87	318	55.01	3.14	244	136.24	112.07	1.22	0.31

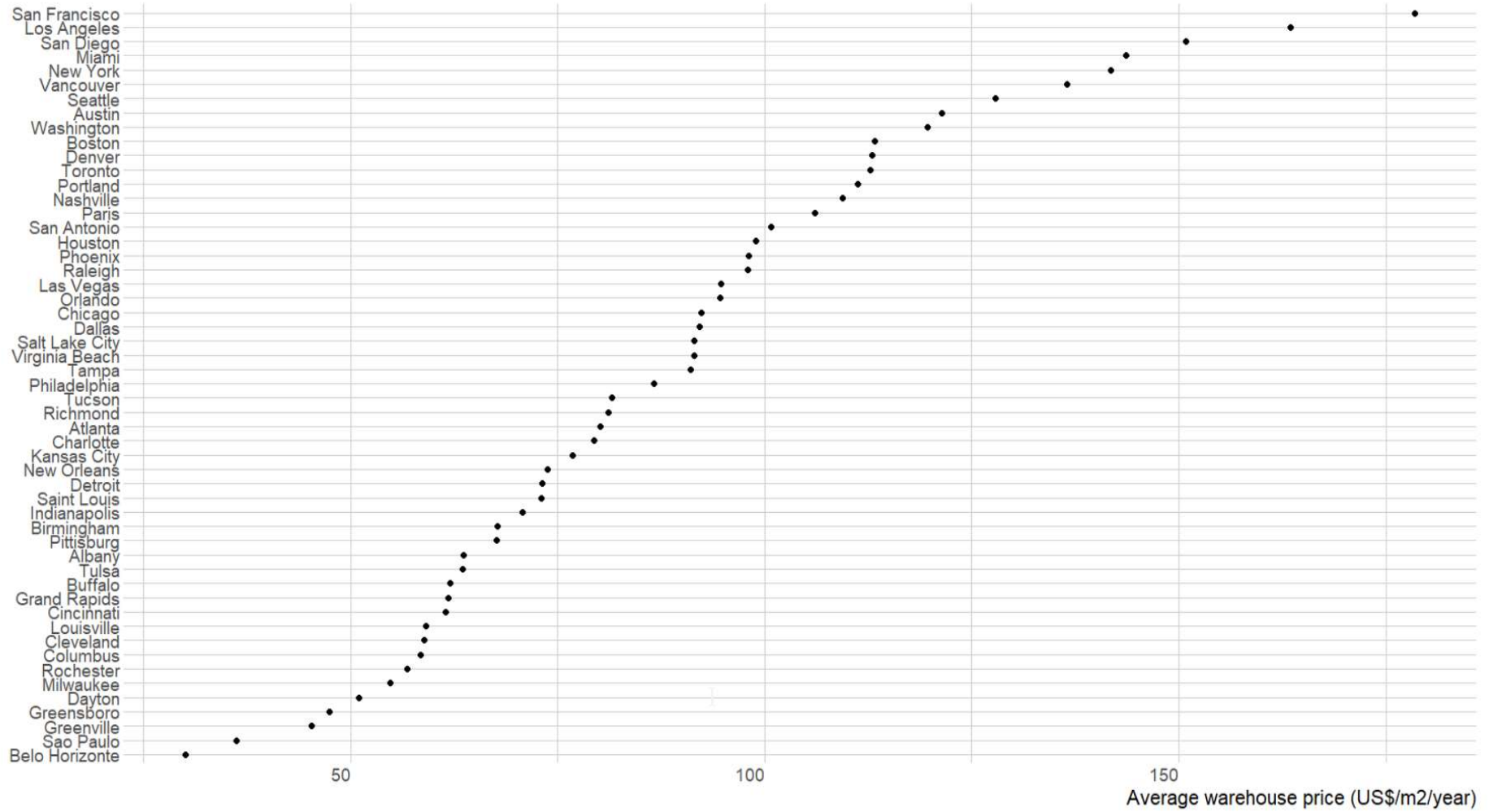


Figure 3: Representation of warehouse average rent prices for each metropolitan area

**Erreur ! Source du renvoi introuvable.** represents the average warehouse rent prices classified according to the location within the metropolitan area – differential rent price for AH and PAZ.

In Figure 6, we represent the proportional warehouse differential (ratio between AH/PAZ) areas. The areas with no significant differential were the ones with less than 10% of the proportional difference of warehouse rent prices between AH and PAZ. Therefore, metropolitan areas with a ratio of less than 0.9 were classified as having higher AH and significant differential. The ones with a ratio greater than 1.1 were classified as metropolitan areas with higher warehouse prices in PAZ. The remaining regions were classified as having no significant differential. Three groups of cities can be derived from this differential, presented in Table 7.

**Table 7: Classification of cities regarding the differential rent prices and location of warehouses**

Warehouse prices are higher in the Activity Hubs areas	Areas with no significant differential	Warehouse prices are significantly higher in the Peripheral Activity Zones areas
Boston	Atlanta	Albany
Charlotte	Austin	Detroit
Columbus	Belo Horizonte	New Orleans
Greensboro	Birmingham	San Antonio
Greenville	Buffalo	
Las Vegas	Chicago	
Louisville	Cincinnati	
Nashville	Cleveland	
New York	Dallas	
Paris	Dayton	
Raleigh	Denver	
Rochester	Grand Rapids	
Salt Lake City	Houston	
San Francisco	Indianapolis	
São Paulo	Kansas City	
Seattle	Los Angeles	
Toronto	Miami	
Vancouver	Milwaukee	
Washington	Orlando	
	Philadelphia	
	Phoenix	
	Pittsburg	
	Portland	
	Richmond	
	Saint Louis	
	San Diego	
	San Antonio	
	Tampa	
	Tucson	
	Tulsa	
	Virginia Beach	

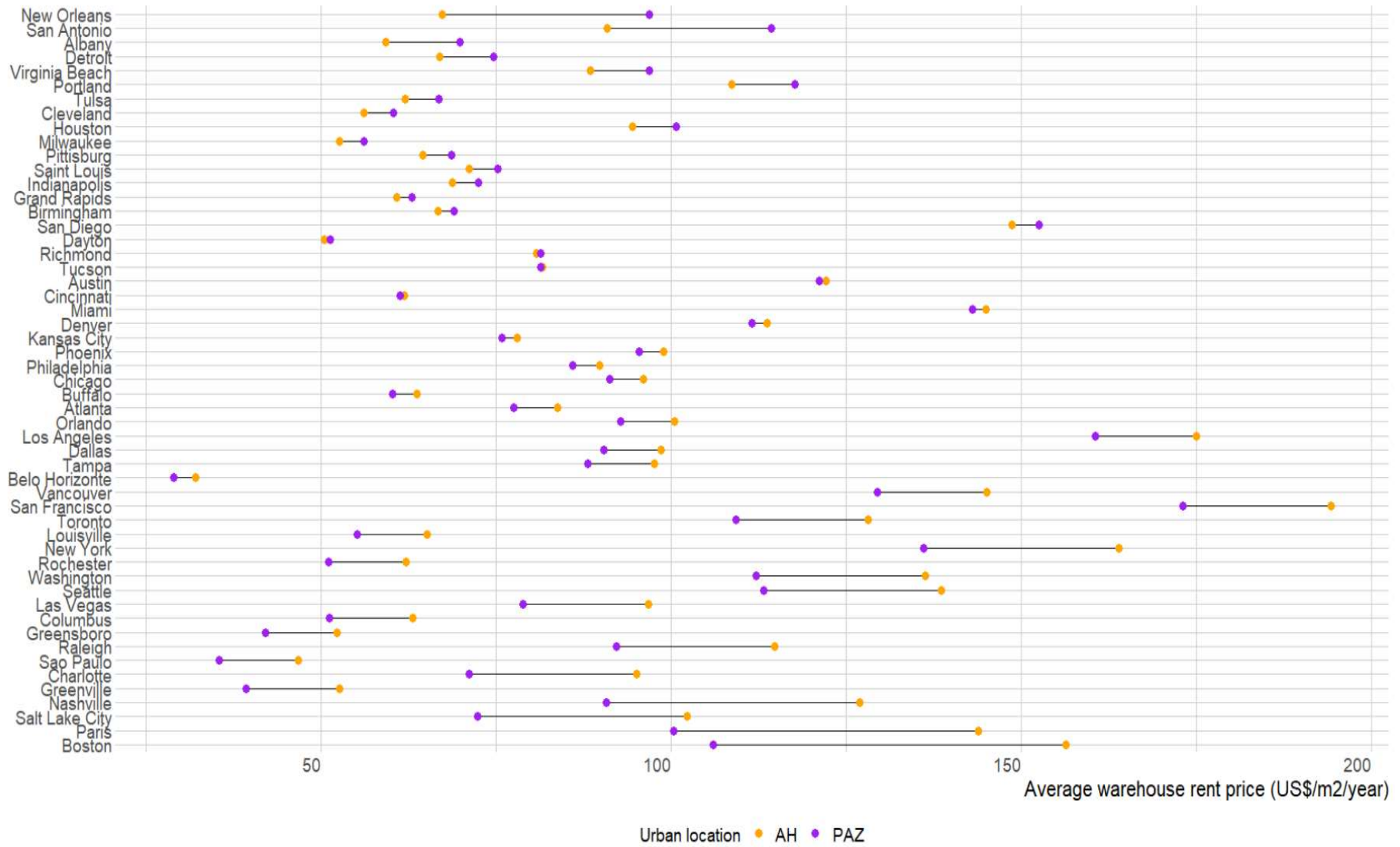


Figure 4: Representation of warehouse average rent prices for AH and PAZ for each metropolitan region

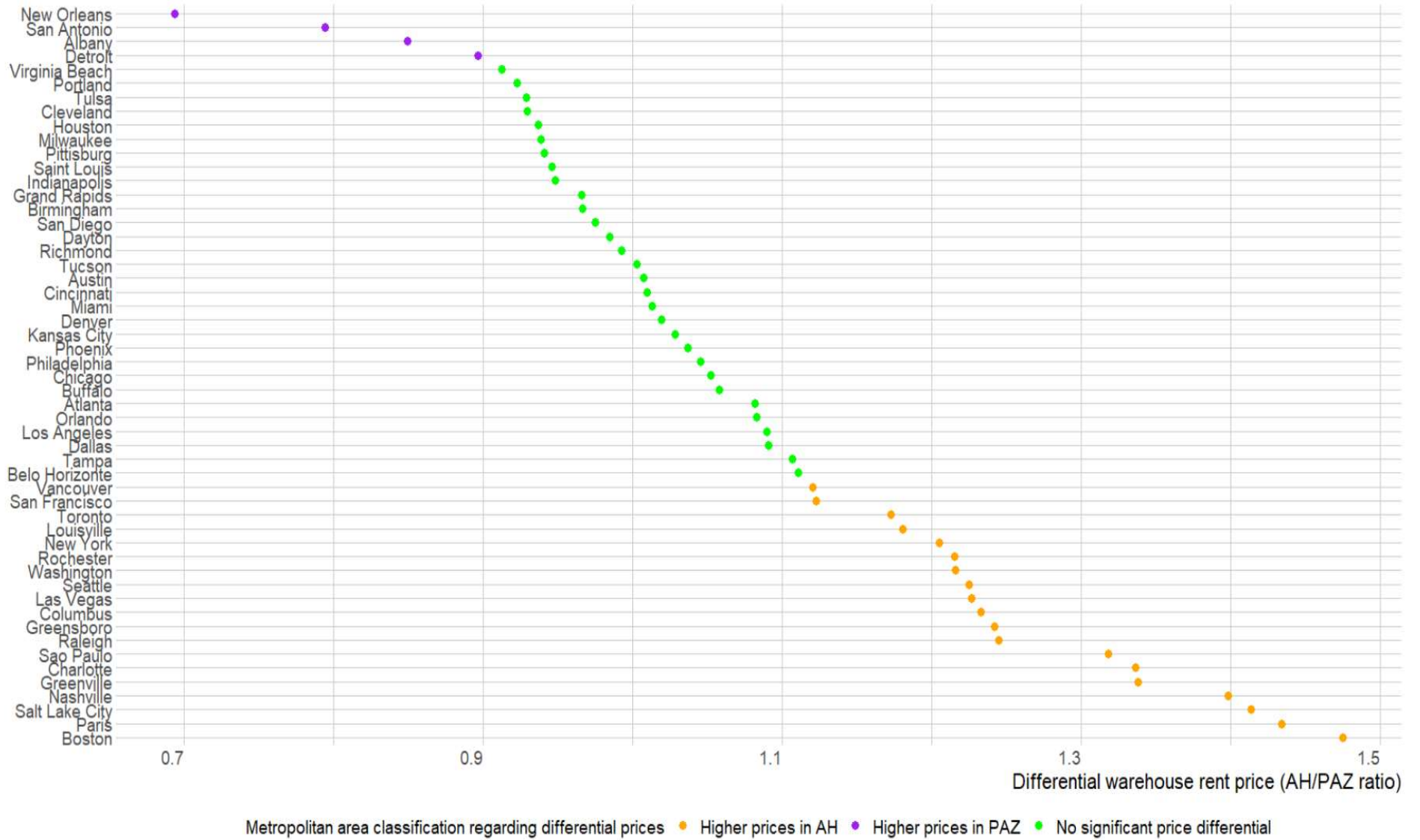


Figure 5: Representation of proportional rent prices differential for warehouses in each location for each metropolitan region

Finally, Figure 6 presents the average distance to the gravity center (logistics sprawl measure) for the two timeframes considered in the studies for the metropolitan areas reviewed through the meta-analysis. Furthermore, Figure 7 presents the yearly logistics sprawl. In Table 8, the classification of metropolitan areas regarding the yearly sprawl is presented. The three categories are (i) areas with negative indicators; (ii) areas with no significant yearly sprawl ( $-1 < YLS < 1$ ), and; (iii) areas with positive YLS.

**Table 8: Classification of cities regarding the logistics sprawl per year**

Negative yearly sprawl	Areas with no significant yearly sprawl	Positive yearly sprawl
Greensboro	Albany	Atlanta
Greenville	Belo Horizonte	Austin
Saint Louis	Buffalo	Birmingham
Seattle	Cincinnati	Boston
Tucson	Cleveland	Charlotte
Virginia Beach	Columbus	Chicago
	Dallas	Dayton
	Denver	Detroit
	Indianapolis	Grand Rapids
	Louisville	Huston
	Orlando	Kansas City
	Raleigh	Las Vegas
	Rochester	Los Angeles
	Salt Lake City	Miami
	San Diego	Milwaukee
	São Paulo	Nashville
	Tampa	New Orleans
		New York
		Paris
		Philadelphia
		Pittsburgh
		Phoenix
		Portland
		Richmond
		San Antonio
		San Francisco
		Toronto
		Tulsa
		Vancouver
		Washington

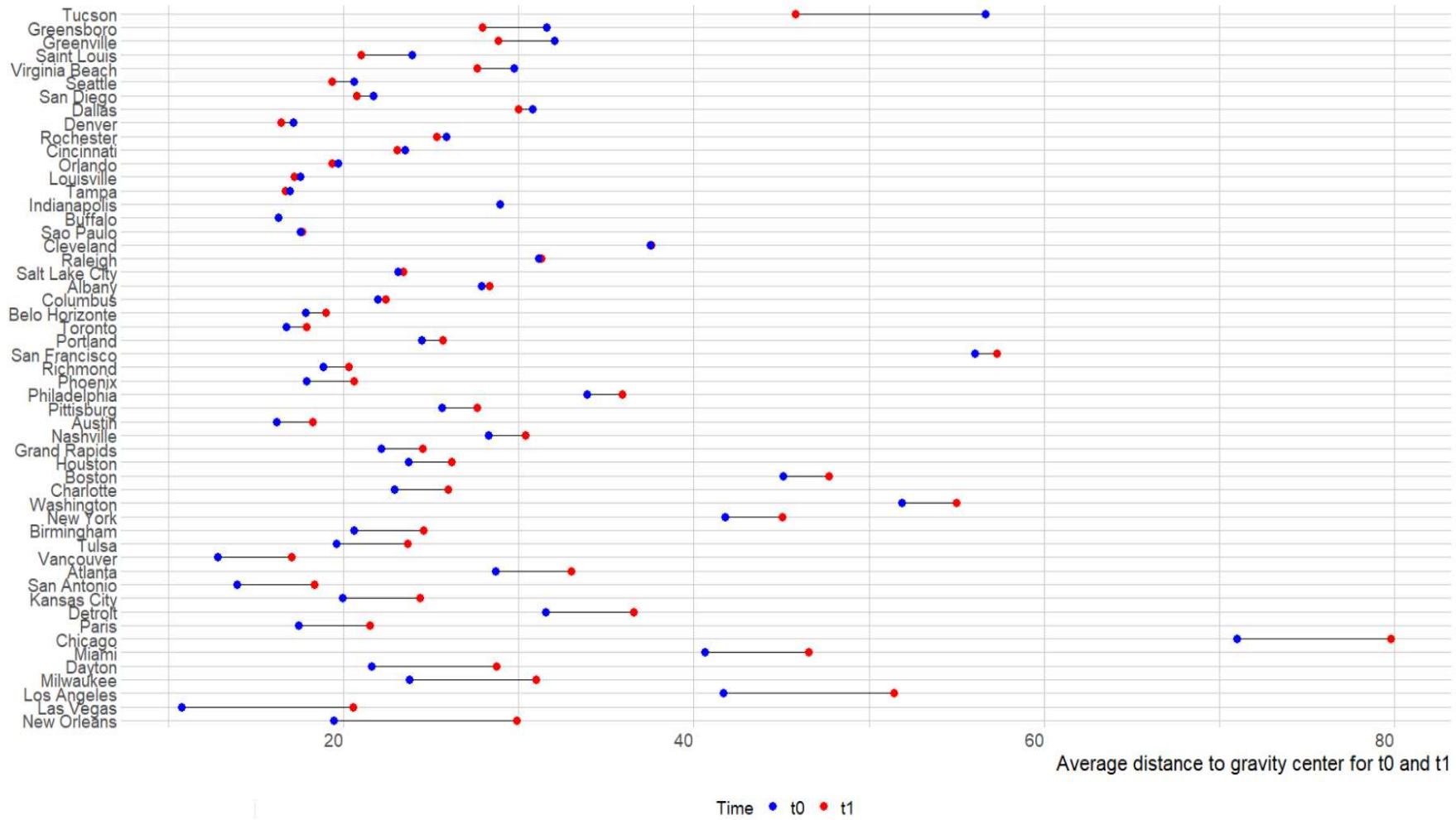


Figure 6: Representation of average distance to gravity center for t0 and t1, for each metropolitan region

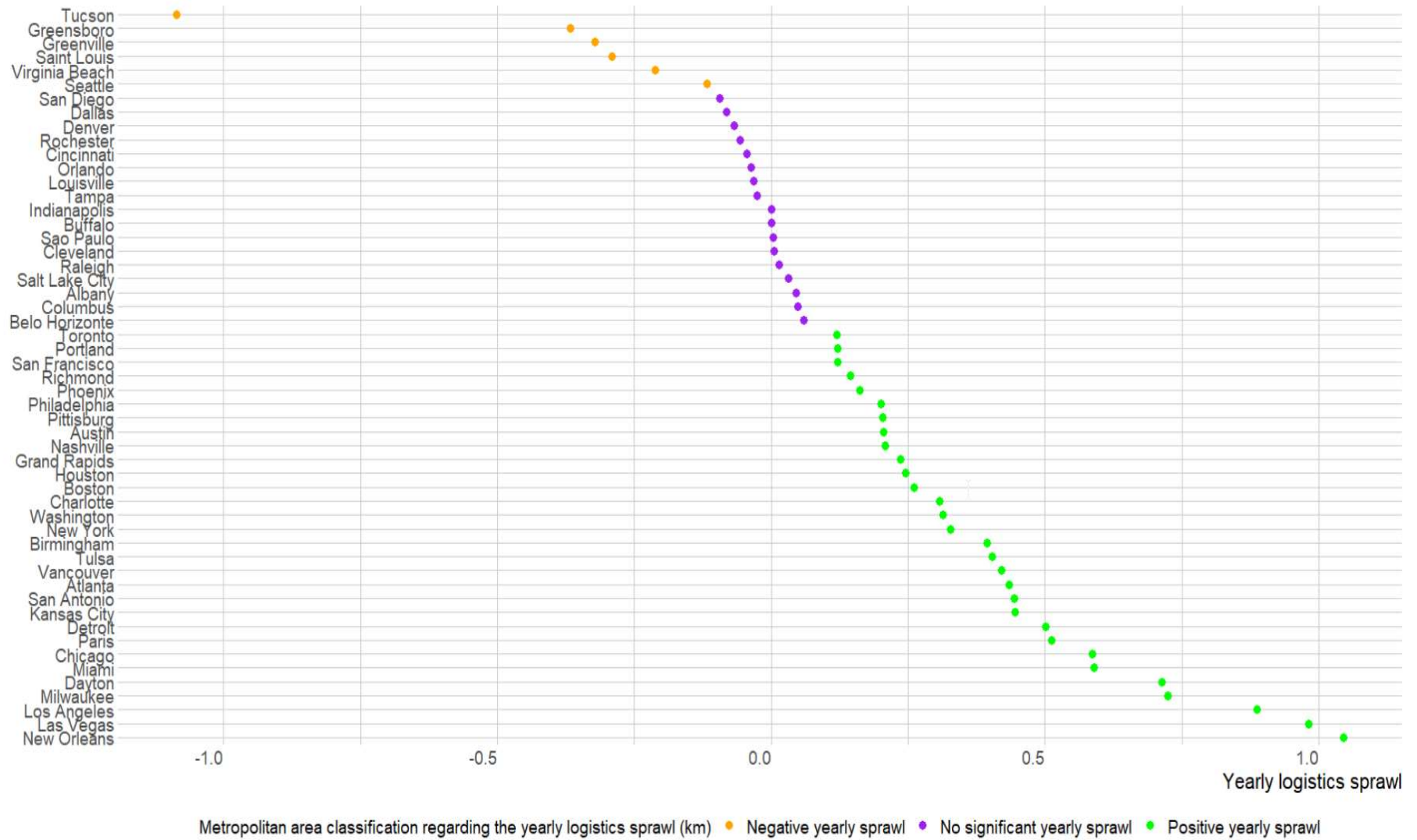


Figure 7: Representation of yearly logistics sprawl



### 3.3. Exploring the hypotheses

#### 3.3.1. Hypothesis 1: The location of warehouses is closely related to the rent values of logistics facilities

For this analysis, we present the Chi-square test regarding the warehouse location (Urban Activity Index), warehouse concentration (number of warehouses), and average rent values for all hexagon bins, disregarding the metropolitan region to where they belong. This approach was designed to statistically test hypothesis 1: Are warehouses' location and rent prices related to urban activity? For this analysis, we have considered only the hexagon bins with warehouses in them. Thus, 2,783 hexagon bins were taken into account (all complete cases). In **Erreur ! Source du renvoi introuvable.** to 10, we present the visual relations among variables, and the Chi-Square test is presented in Tables 9, 10, and 11.

**Table 9: Number of hexagons classified by the concentration of warehouses and urban classification**

	Extremely high	Very High	High	Medium	Low
AH	29	13	386	476	3
PAZ	13	5	453	1330	75

Pearson's Chi-squared test with simulated p-value (based on 2000 replicates):  $\chi^2 = 168.30$ , p-value = 0.0004998



Figure 8: Warehouse count x urban classification

**Table 10: Number of hexagons classified by average prices of warehouses and urban classification**

	High	Medium	Low
AH	295	420	192
PAZ	437	951	488

Pearson's Chi-squared test with simulated p-value (based on 2000 replicates):  $\chi^2 = 28.07$ , p-value = 0.0004998

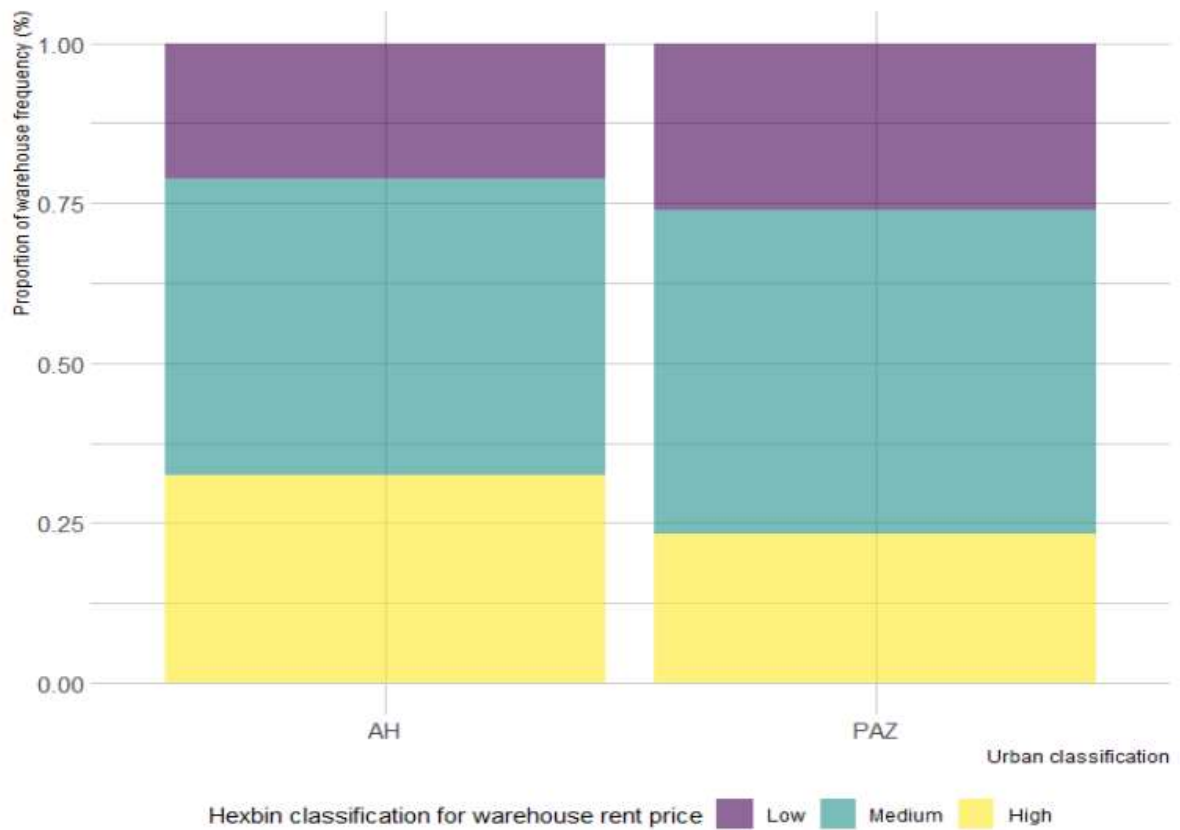


Figure 9: Warehouse average prices x urban classification

Table 11: Number of hexagons classified by the number of warehouses and warehouse average prices

Warehouse count	High	Medium	Low
Extremely high	14	23	5
Very high	7	7	4
High	246	463	130
Medium	445	862	499
Low	20	16	42

Pearson's Chi-squared test with simulated p-value (based on 2000 replicates):  $\chi^2 = 92.6019$ , p-value = 0.0004997

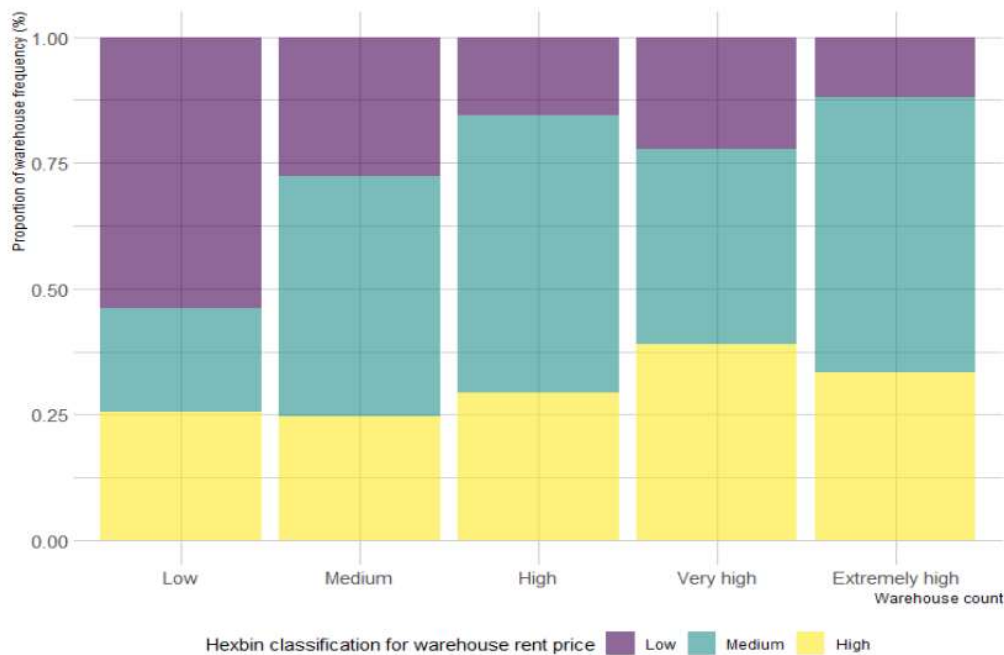


Figure 10: Warehouse count and warehouse average prices

Considering all metropolitan areas (Figure 8), 33% of the warehouses are in AH and 67% in PAZ. From the  $\chi^2$  we can reject the independence hypotheses and, therefore, assume that, in this case, that the **concentration of warehouses depends on the urban activity index**. Also, AH have a higher proportion of hexagons with a high number of warehouses than PAZ.

Figure 9 shows the relationship between average warehouse rent prices and the location within the city. Also, considering all metropolitan areas, the average warehouse prices decrease as we move closer to PAZ, and we can reject the hypothesis that these variables are independent. In other words, considering a significance level of 5%, **the warehouse rent prices depend on the location in the metropolitan areas**.

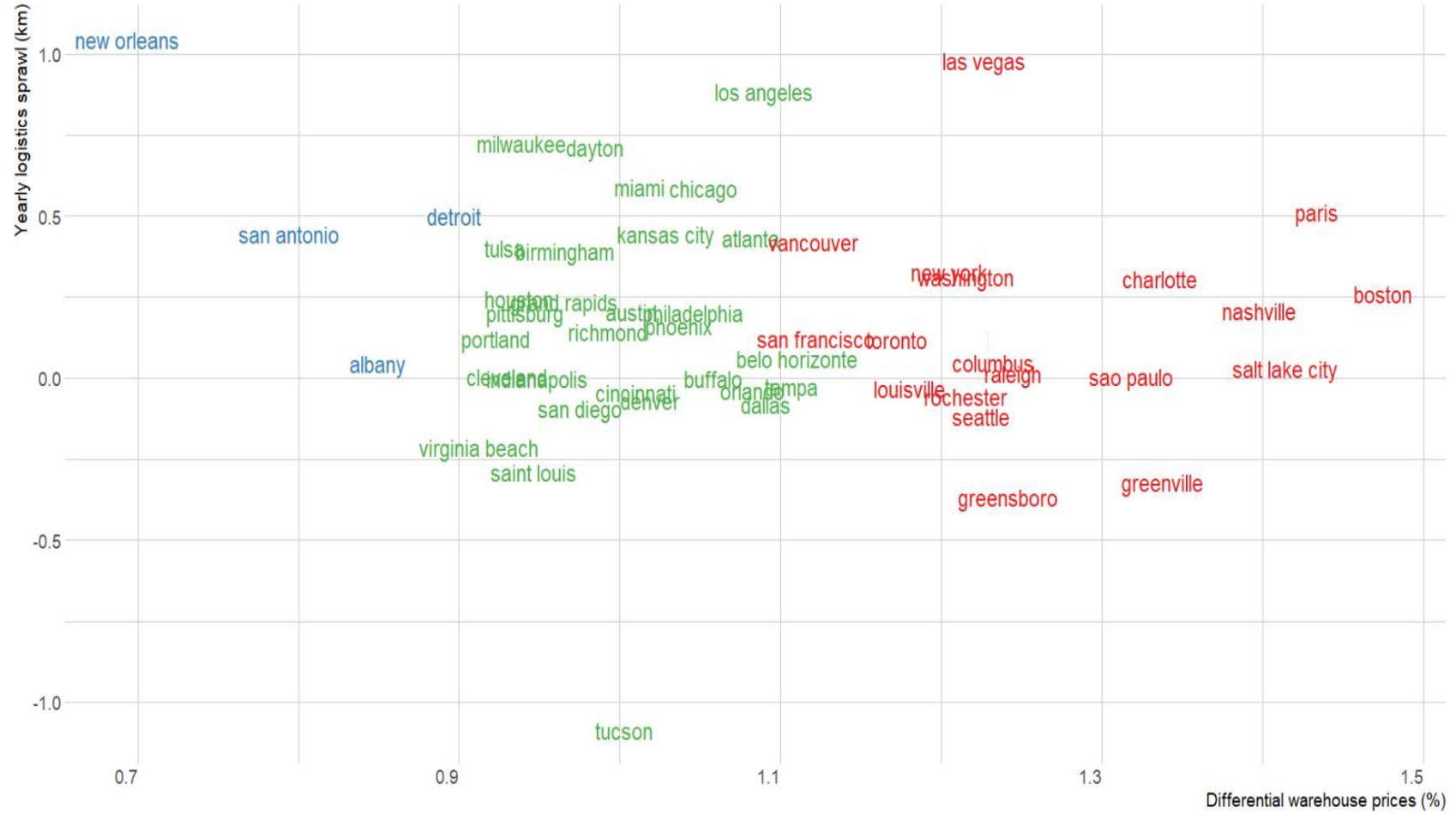
Synthetically, considering all metropolitan areas investigated, it is possible to state that **the rent prices of the warehouses depend on the spatial concentration of these logistics facilities**. This interpretation is related to a possible effect of agglomeration economy or locational decisions performed for similar reasons. Therefore, this work's **first hypothesis cannot be rejected from the tests performed considering the investigated metropolitan areas and methods**.

### 3.3.2. Hypothesis 2: Logistics sprawl is higher in cities with a high differential between central and suburban land/rent values

This section compares the urban attributes, logistics spatial structure, and real estate practices in metropolitan areas. For that, two variables were used: (i) differential warehouse prices (DWP) and (ii) yearly logistics sprawl. The first was the ratio between the average warehouse price in suburban areas and central areas. The second, the difference of average distance to the mean center between years in analysis divided by the number of years considered. Figure 11 shows the classification of metropolitan areas according to the differential price category (three groups) and the relationship between the YLS and the DWP. These areas are listed according to the classification in Table 12 (relating YLS and DWP categories).

**Table 12: Classification of metropolitan areas**

Differential warehouse price	Yearly logistics Sprawl		
	Positive YLS	No significant YLS	Negative YLS
<b>Lower prices in AH</b>	Detroit New Orleans San Antonio	Albany	-
<b>No significant differential prices</b>	Atlanta Austin Birmingham Chicago Dayton Grand Rapids Houston Kansas City Los Angeles Miami Milwaukee Philadelphia Phoenix Pittsburg Portland Richmond Tulsa	Belo Horizonte Buffalo Cincinnati Cleveland Dallas Denver Indianapolis Orlando San Diego Tampa	Saint Louis Tucson Virginia Beach
<b>Higher prices in AH</b>	Boston Charlotte Las Vegas Nashville New York Paris San Francisco Toronto Vancouver Washington	Columbus Louisville Raleigh Rochester Salt Lake City São Paulo	Greensboro Greenville Seattle



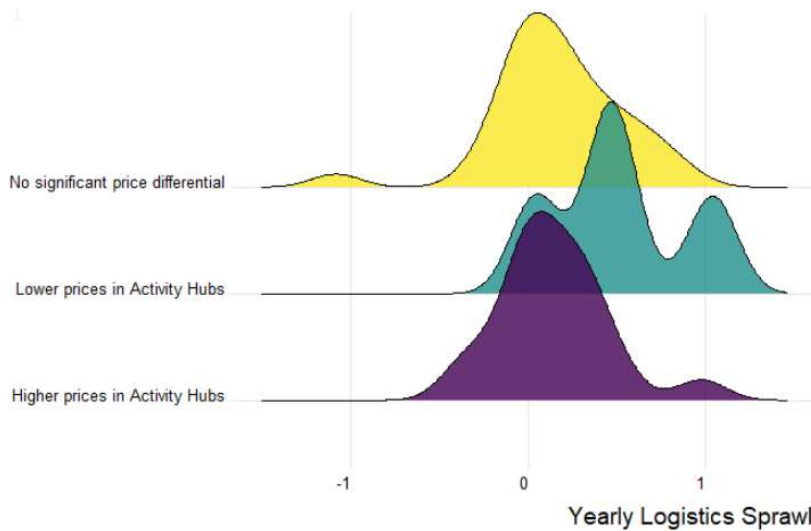
a Higher prices in Activity Hubs    a Lower prices in Activity Hubs    a No significant price differential

Figure 11: Differential classes and a scatterplot for DWP and YLS

When calculating the correlation coefficient, it is assumed that at least one of the variables is Normally distributed. For that, we have performed the Shapiro-Wilk normality test for DWP and YLS in each category of differential prices. The results are presented in Table 13. We could only reject the null hypothesis for DWP, at a significance level of 5% ( $p > 0.05$ ), for the metropolitan areas classified as having *no significant differential price* and regarding the YLS. In other words, we can say that all metropolitan area variables considering the differential price classification are normally distributed besides the combination of *no significant differential price* and YLS. Therefore, it is possible to assess the variable DWP and YLS relationship for the metropolitan area groups. The representation of these variables is in Figures 12 and 13.

**Table 13: Classification of metropolitan areas**

Normality test	Classification of differential warehouse price		
	Lower prices in AH	No significant differential price	Higher prices in AH
Differential warehouse prices (DWP)	W = 0.96843 p-value = 0.8317	W = 0.94201 p-value = 0.103	W = 0.92697 p-value = 0.1523
Yearly Logistics Sprawl (YLS)	W = 0.95821 p-value = 0.7677	W = 0.92233 p-value = 0.03084	W = 0.94459 p-value = 0.3184



*Figure 12: Density plot for YLS for each differential price classification*

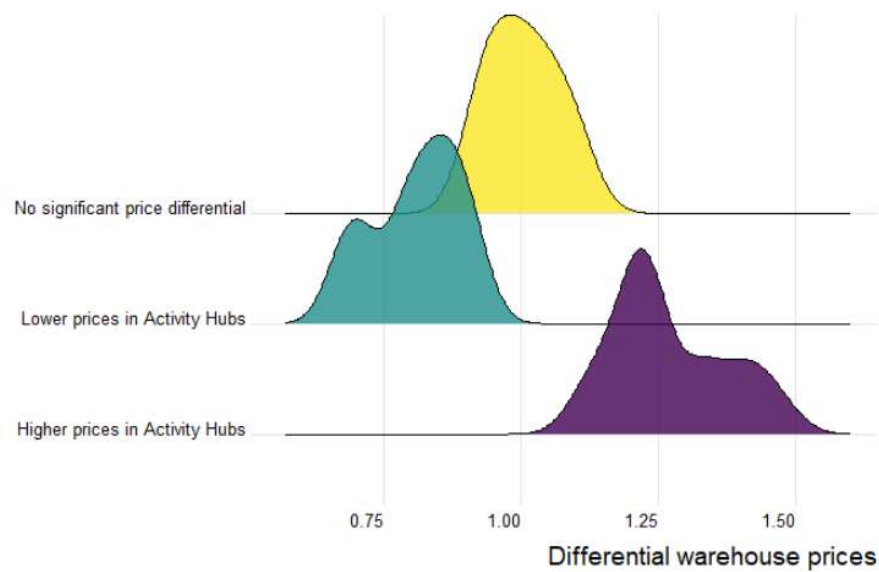


Figure 13: Density plot for DWP each differential price classification

Still analysing the first hypothesis (The location of warehouses is closely related to the land/rent values of logistics facilities), considering this classification of differential prices, we have then performed the **Pearson Correlation** test for all the **hexagon datasets** (53 metropolitan areas) and each combination of categories, relating the variables (i) **average warehouse price; and (ii) urban activity index**.

For all the metropolitan areas, the Pearson coefficient was 0.484 (p-value < 2.2e-16). Since the p-value is lower than 0.05 (5% of significance level), the null hypothesis that the relationship between the average warehouse rent prices and the urban activity index is not significant can be rejected. Therefore, the greater the urban activity index, the higher the rent prices.

If we consider the dataset for each category, relationships are still significant but less intensive, and the correlation index was:

- Higher prices in activity hubs:  $r = 0.274$
- Lower prices in activity hubs:  $r = -0.1809918$
- No significant differential prices:  $r = 0.1584641$

Regarding the second hypothesis (Logistics sprawl is higher in cities with a high differential between land/rent values in Activity Hubs and Peripheral Activity Zones), we have computed the



Pearson correlation index to investigate the relationship between **YLS** and **DWP**. Considering the **DWP** and **YLS** for **each metropolitan area** (dataset with 53 metropolitan areas), the  $r = -0.1447$  (p-value = 0.3009). For each category, we have:

- No significant price differential:  $r = 0.0668$  (p-value = 0.7257)
- Lower prices in Activity Hubs:  $r = -0.7449$  (p-value = 0.2551)
- Higher prices in Activity Hubs:  $r = 0.0265$  (p-value = 0.9141)

In Figure 14, we present exploratory linear regression models for the categories of differential prices. Considering the values for the index  $r$  and the p-values, the **null hypothesis that the relationship between the differential warehouse prices (DWP) and the yearly logistics sprawl (YLS) is not significant cannot be rejected** for all the data combinations.

**There is no evidence that differential warehouse rent prices are related to the yearly sprawl. Therefore, the second hypothesis is rejected, considering the method to relate yearly logistics sprawl with differential warehouse rent prices.**

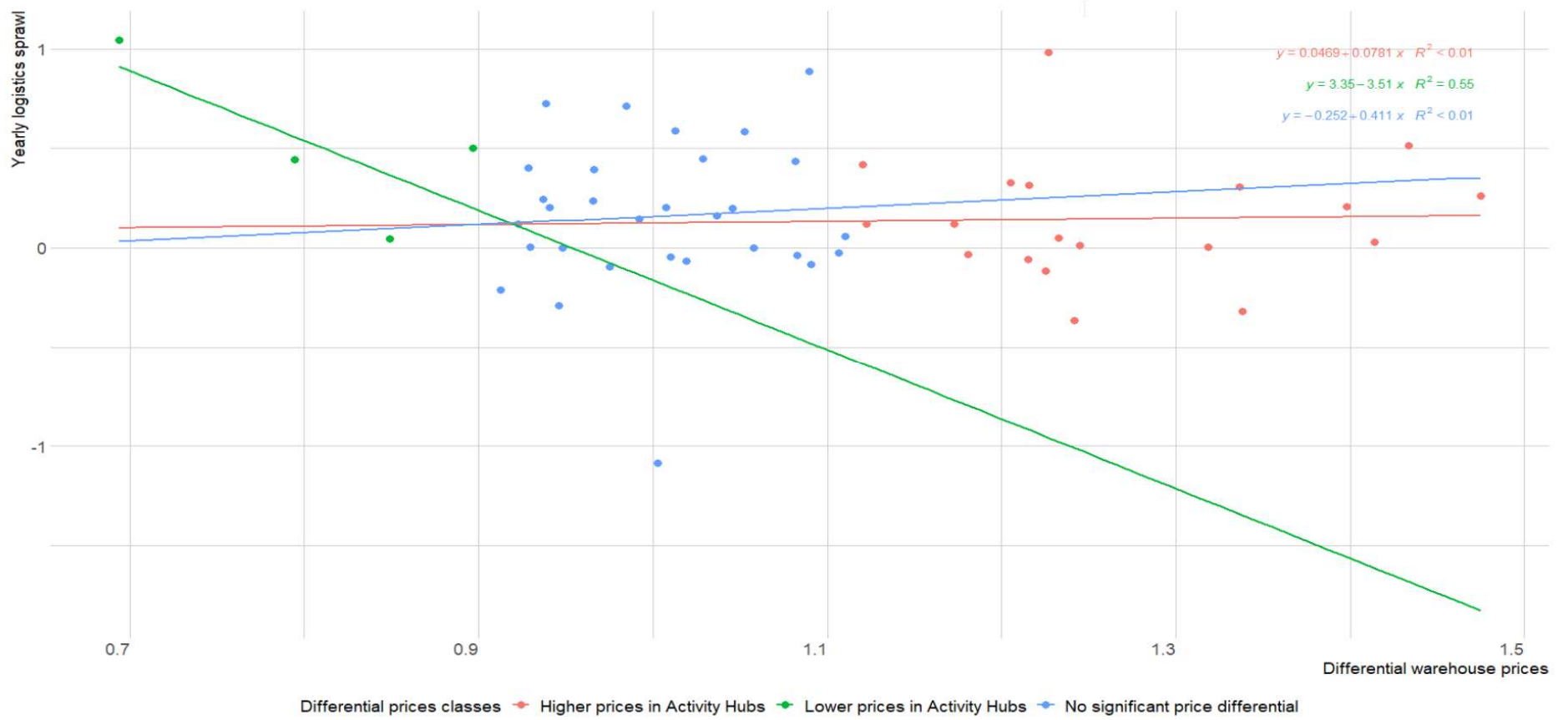


Figure 14: Exploratory linear regression models for differential prices classes

## 4. Final discussion

In this work, we present a disaggregated methodological approach to explore the hypothesis: (i) the location of warehouses is closely related to the land/rent values of logistics facilities, and (ii) logistics sprawl is higher in cities with a high differential between central and suburban land/rent values.

Regarding the first hypothesis, which aimed at investigating whether warehouse rent prices were higher in more central places within the metropolitan areas, we can state that there is statistically significant evidence that the location of warehouses and the average rent prices are not independent.

For the second hypothesis, we explore the correlation between differential rent values and yearly logistics sprawl. Considering all metropolitan regions, we identify that the null hypothesis (the relationship between the differential warehouse prices and the yearly logistics sprawl is not significant) cannot be rejected. Nevertheless, since the metropolitan areas are significantly different in many dimensions, this statement cannot be generalised when we gather them into more similar groups.

The typology adopted to classify metropolitan areas is to further investigate by the correlation and dependency results where: (i) areas with higher warehouse rent prices in Activity Hubs; (ii) area with lower rent prices in Activity Hubs; and (iii) areas with no significant warehouse rent differential prices.

Finally, we conclude that: (i) it is essential to classify metropolitan areas into a typology in order to perform comparative studies; (ii) warehouse location and prices are related to the urban activity; (iii) logistics sprawl is not significantly related to differential warehouse rental prices, but there are differences among metropolitan categories.

This work brings a methodological contribution since we present an innovative framework for comparing metropolitan regions considering the spatial pattern of logistics facilities and urban characteristics regarding disaggregated open data. This method is reproducible for other

approaches and other city scales. It can subsidise local and regional public authorities in decision-making oriented to developing more effective public policy addressing logistics land use and transportation planning. Coordinating these dimensions is essential to support urban logistics stakeholders' needs, cities' livability, and the real estate market.

Some limitations of this work concern:

- The real estate data was collected considering only the available warehouses for rental, not consolidating a comprehensive dataset;
- We used only a limited number of commercial real estate websites to identify warehouse rental prices. These websites may not be representative enough of the actual market of warehouses. There is a need for further investigation.
- The real estate prices for logistics facilities were collected for one static timeframe, which compromises the analysis of the dynamic relationship between warehouse structure and rent prices. We recommend a periodic collection of this information to consolidate a dataset that can reflect this dynamic phenomenon;
- The typology might change if we include other metropolitan areas and cities and if we include or exclude variables. There is a need for further investigation to improve the classification of these urban areas;
- Economies of agglomeration concerning infrastructure and economies of scale can be essential for locational decisions and should be further investigated.
- The number of warehouses for each metropolitan area is significantly different. This issue can result in bias while analysing the areas in comparative ways, but we managed to address these issues through homogeneous disaggregated spatial units and normalisation of variables.
- Cities in Asia are missing from the analysis (due to difficulties in collecting rental prices). There is a need for further investigation. Similarly, there are not enough European cities in the dataset.



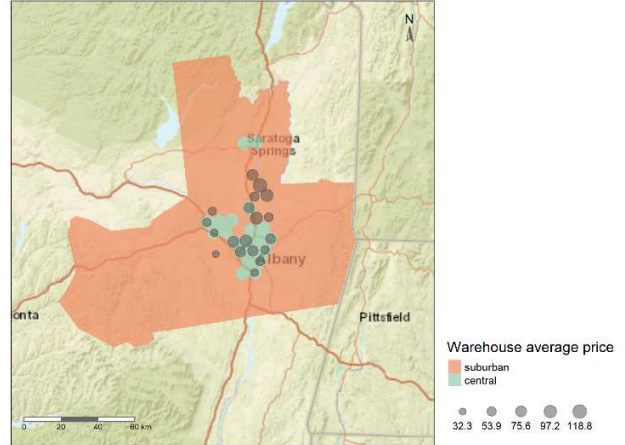
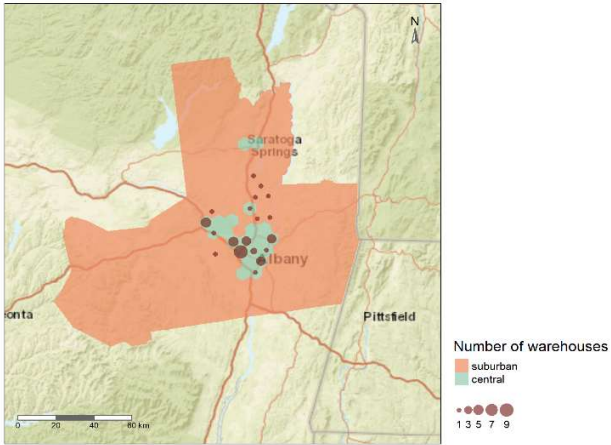
## 5. References

- Bakillah, M., Liang, S., Mobasheri, A., Jokar Arsanjani, J., Zipf, A., 2014. Fine-resolution population mapping using OpenStreetMap points-of-interest. *Int. J. Geogr. Inf. Sci.* 28, 1940–1963. <https://doi.org/10.1080/13658816.2014.909045>
- Ben-Joseph, E., Gordon, D., 2000. Hexagonal planning in theory and practice. *J. Urban Des.* 5, 237–265. <https://doi.org/10.1080/713683965>
- Birch, C.P.D., Oom, S.P., Beecham, J.A., 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecol. Model.* 206, 347–359. <https://doi.org/10.1016/j.ecolmodel.2007.03.041>
- Chair logistics City, 2019. Welcome to Logistics City - White paper of the new urban logistics.
- Cidell, J., 2010. Concentration and decentralisation: The new geography of freight distribution in US metropolitan areas. *J. Transp. Geogr.* 18, 363–371. <https://doi.org/10.1016/j.jtrangeo.2009.06.017>
- Crown, W.D., Caldwell, J.H., Rosenstein, J.G., 2018. Standard 7 — Geometry and Spatial Sense. *N. J. Math. Curric. Framew.* 209–249.
- Dablanc, L., Browne, M., 2019. Introduction to special section on logistics sprawl. *J. Transp. Geogr.* 0–1. <https://doi.org/10.1016/j.jtrangeo.2019.01.010>
- Dablanc, L., Ogilvie, S., Goodchild, A., 2014. Logistics sprawl. *Transp. Res. Rec.* 2410, 105–112. <https://doi.org/10.3141/2410-12>
- Dablanc, L., Palacios-Argüello, L., 2019. Chaire Logistics City.
- Dablanc, L., Rakotonarivo, D., 2010. The impacts of logistics sprawl: How does the location of parcel transport terminals affect the energy efficiency of goods' movements in Paris and what can we do about it? *Procedia - Soc. Behav. Sci.* 2, 6087–6096. <https://doi.org/10.1016/j.sbspro.2010.04.021>
- Ganis, M., 2015. Planning Urban Places, Planning Urban Places. <https://doi.org/10.4324/9781315760742>
- Heitz, A., Dablanc, L., Tavasszy, L.A., 2017. Logistics sprawl in monocentric and polycentric metropolitan areas: The cases of Paris, France, and the Randstad, the Netherlands. *Region* 4, 93–107. <https://doi.org/10.18335/region.v4i1.158>
- Jokar Arsanjani, J., Mooney, P., Zipf, A., Schauss, A., 2015. Quality Assessment of the Contributed Land Use Information from OpenStreetMap Versus Authoritative Datasets, in: Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (Eds.), *OpenStreetMap in GIScience*, Lecture Notes in Geoinformation and Cartography. Springer International Publishing, Cham, pp. 37–58. [https://doi.org/10.1007/978-3-319-14280-7\\_3](https://doi.org/10.1007/978-3-319-14280-7_3)
- Klinkhardt, C., Woerle, T., Briem, L., Heilig, M., Kagerbauer, M., Vortisch, P., 2021. Using OpenStreetMap as a Data Source for Attractiveness in Travel Demand Models. *Transp. Res. Rec. J. Transp. Res. Board* 036119812199741. <https://doi.org/10.1177/0361198121997415>
- Logistics City Chair : Research [WWW Document], 2021. URL <https://www.lvmt.fr/en/chaieres/logistics-city/> (accessed 6.2.21).
- Mobasheri, A., Sun, Y., Loos, L., Ali, A., 2017. Are Crowdsourced Datasets Suitable for Specialized Routing Services? Case Study of OpenStreetMap for Routing of People with Limited Mobility. *Sustainability* 9, 997. <https://doi.org/10.3390/su9060997>
- OpenStreetMap contributors, 2020. Planet dump retrieved from <https://planet.osm.org>.
- Palacios-Argüello, L., Dablanc, L., 2020a. Logistics city chair: Logistics sprawl characterisation. Chair Logistics City, Paris.
- Palacios-Argüello, L., Dablanc, L., 2020b. Logistics City Chair: Logistics Sprawl characterization. Paris.
- Quinn, P., 2013. Road density as a proxy for population density in regional-scale risk modeling. *Nat. Hazards* 65, 1227–1248. <https://doi.org/10.1007/s11069-012-0379-y>

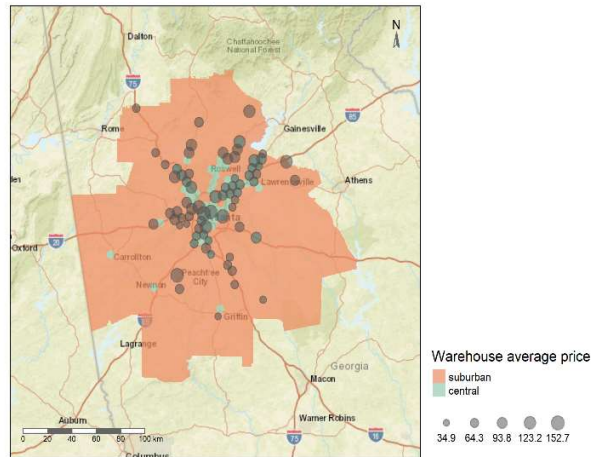
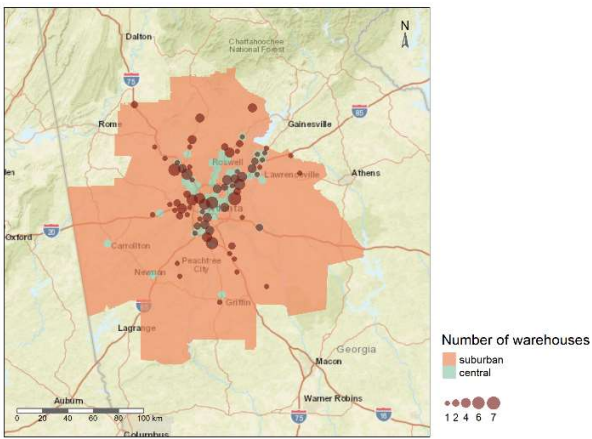
- Sakai, T., Beziat, A., Heitz, A., 2020. Location factors for logistics facilities: Location choice modeling considering activity categories. *J. Transp. Geogr.* 85, 102710. <https://doi.org/10.1016/j.jtrangeo.2020.102710>
- Sarkar, S., Wu, H., Levinson, D.M., 2020. Measuring polycentricity via network flows, spatial interaction and percolation. *Urban Stud.* 57, 2402–2422. <https://doi.org/10.1177/0042098019832517>
- Whitehand, J.W.R., Batty, M., Longley, P., 1996. Fractal Cities: A Geometry of Form and Function. *Geogr. J.* 162, 113. <https://doi.org/10.2307/3060277>
- Zhang, L., Pfoser, D., 2019. Using OpenStreetMap point-of-interest data to model urban change—A feasibility study. *PLOS ONE* 14, e0212606. <https://doi.org/10.1371/journal.pone.0212606>

# APPENDIX A

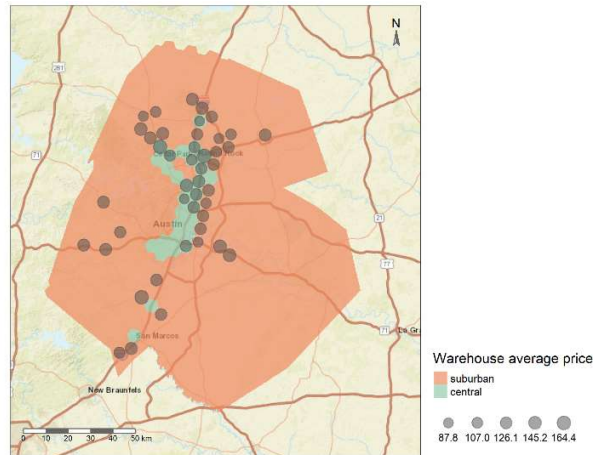
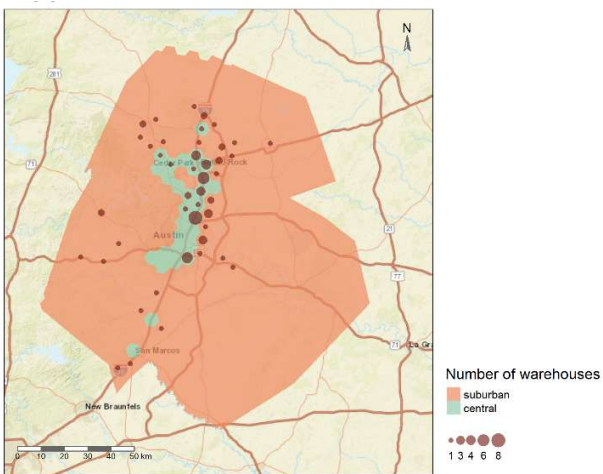
## Albany



## Atlanta

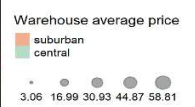
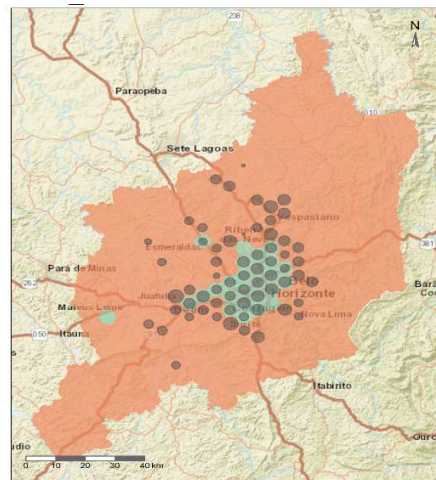
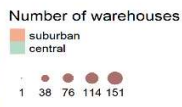
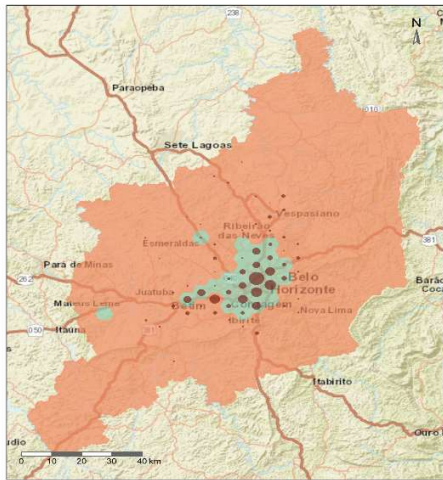


## Austin

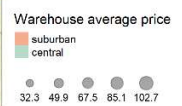
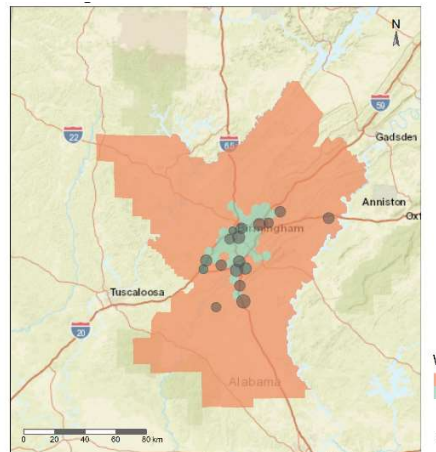
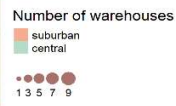
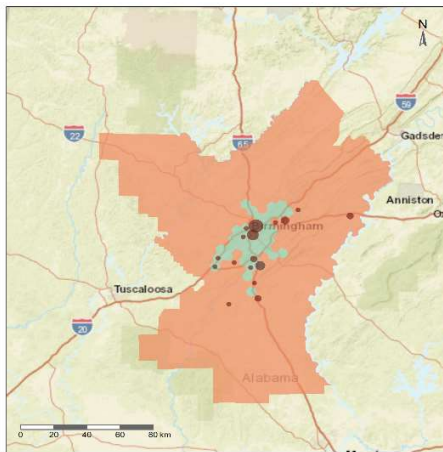




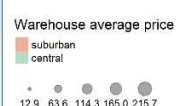
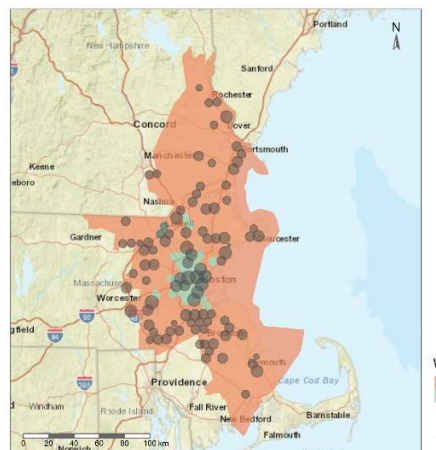
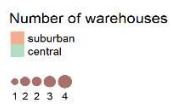
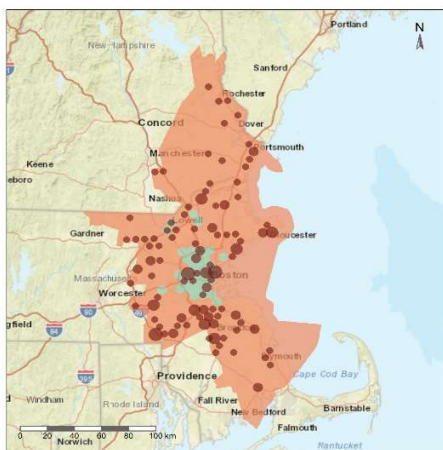
## Belo Horizonte



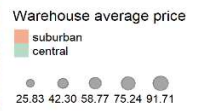
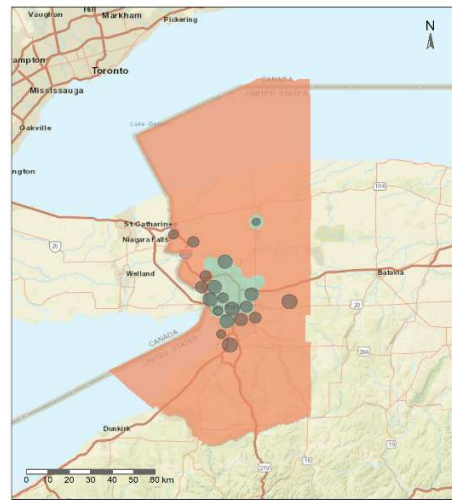
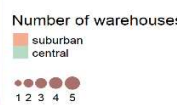
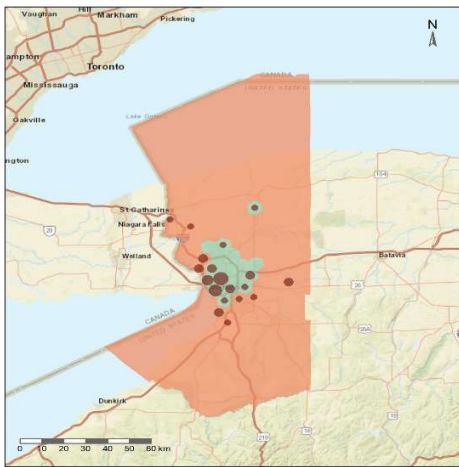
## Birmingham



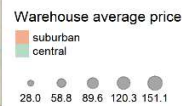
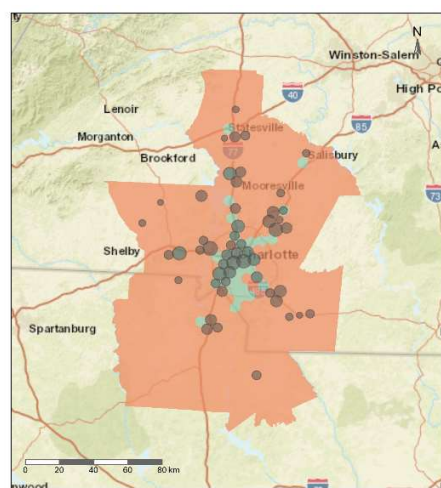
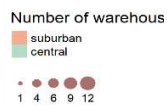
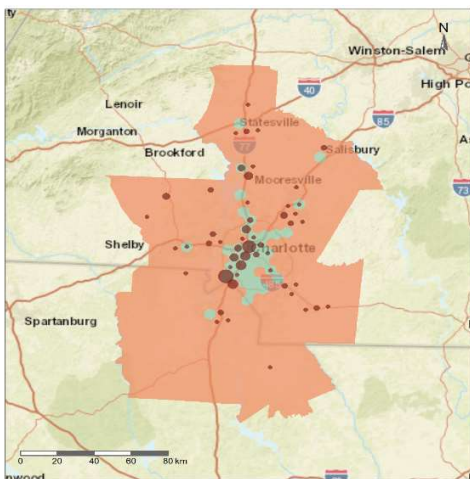
## Boston



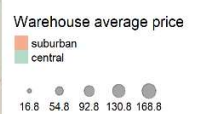
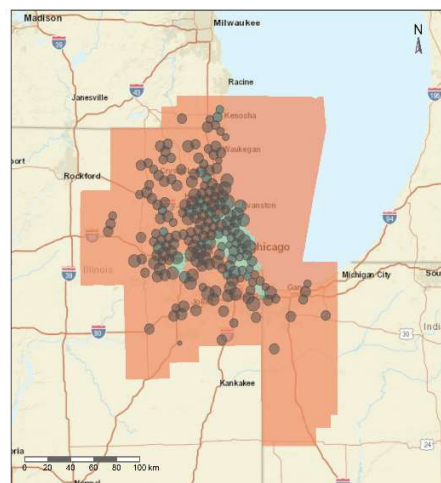
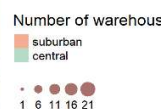
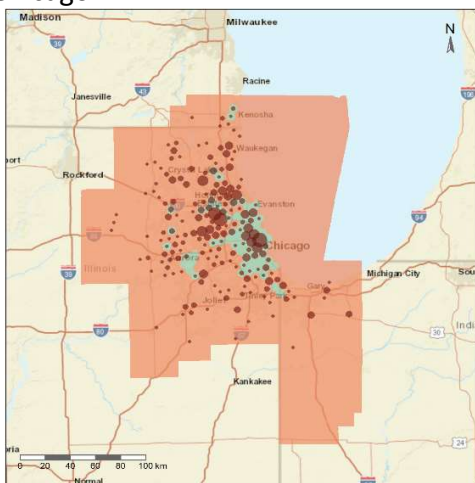
## Buffalo



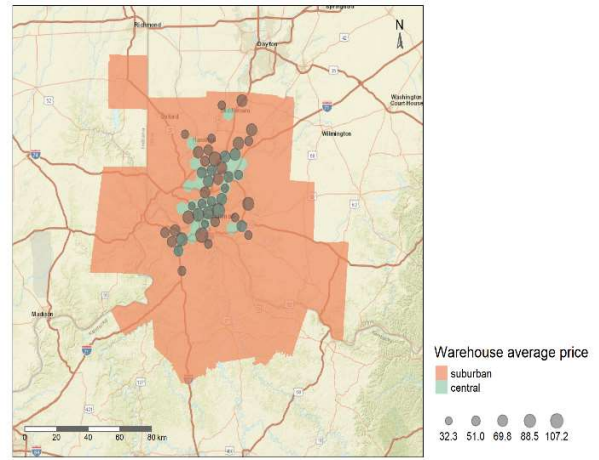
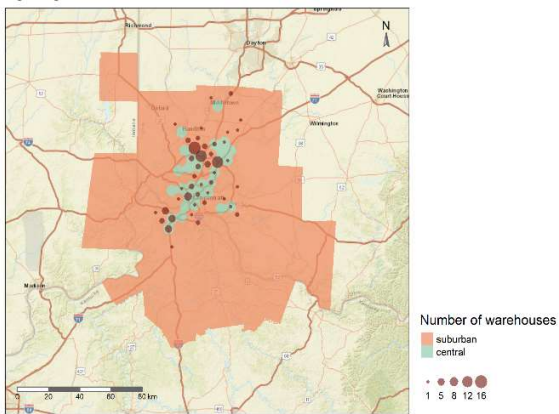
## Charlotte



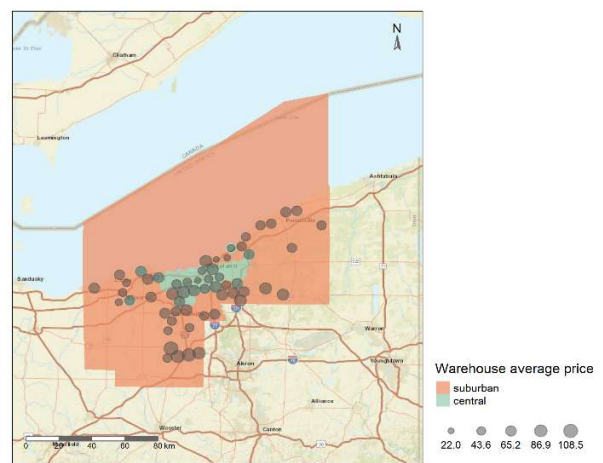
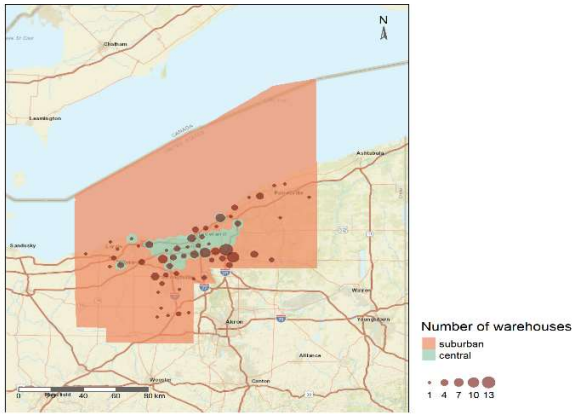
## Chicago



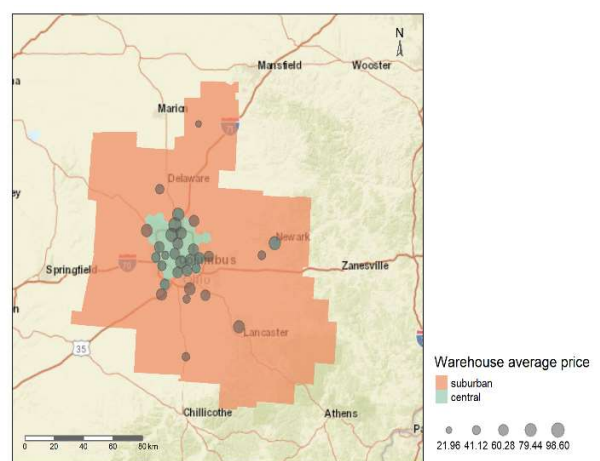
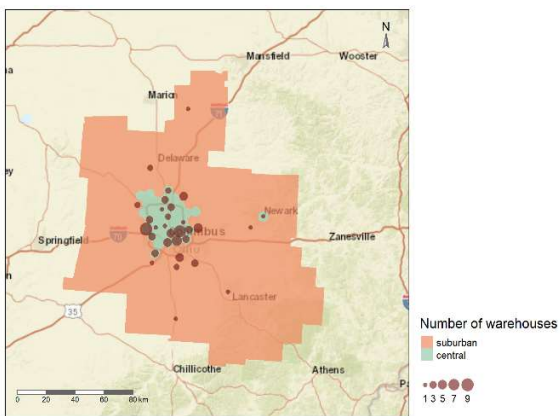
### Cincinnati



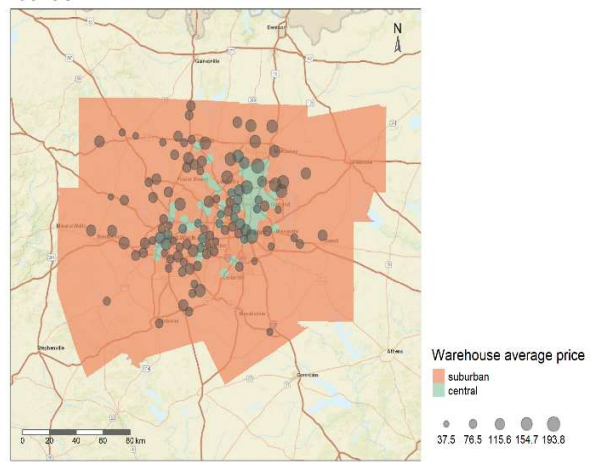
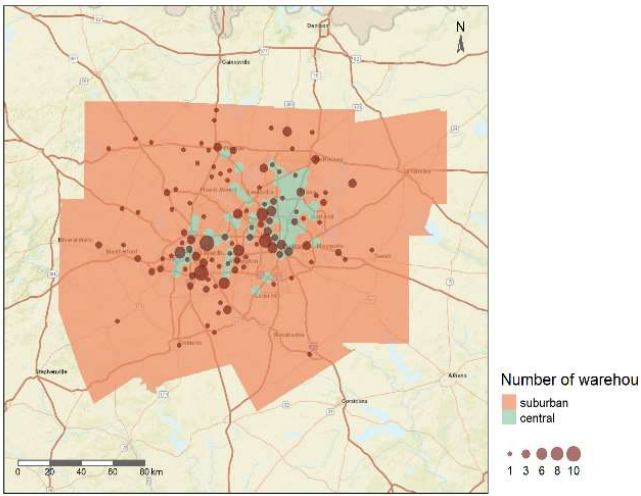
### Cleveland



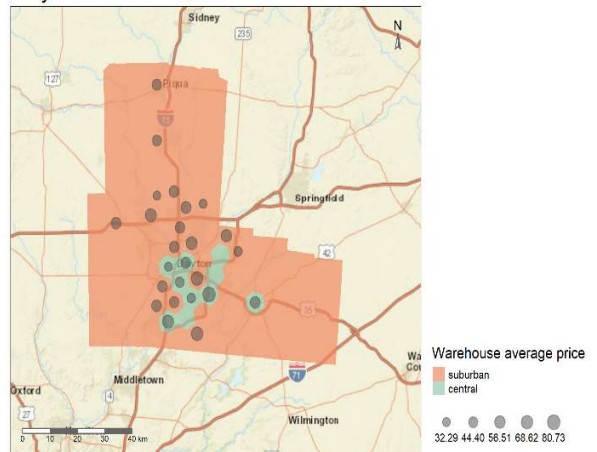
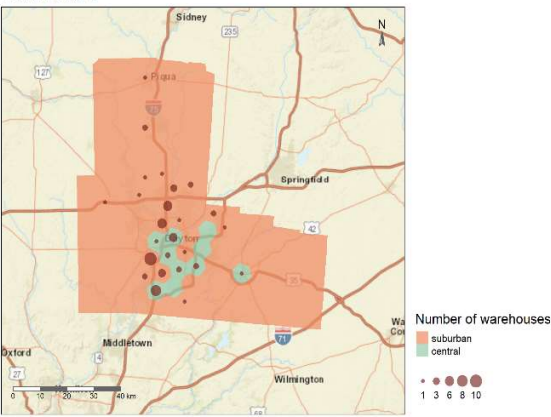
### Columbus



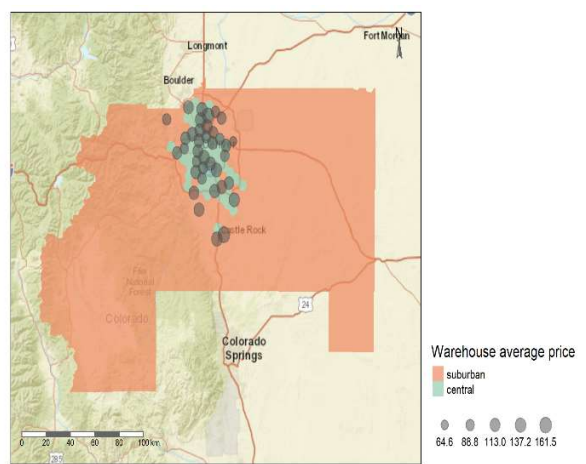
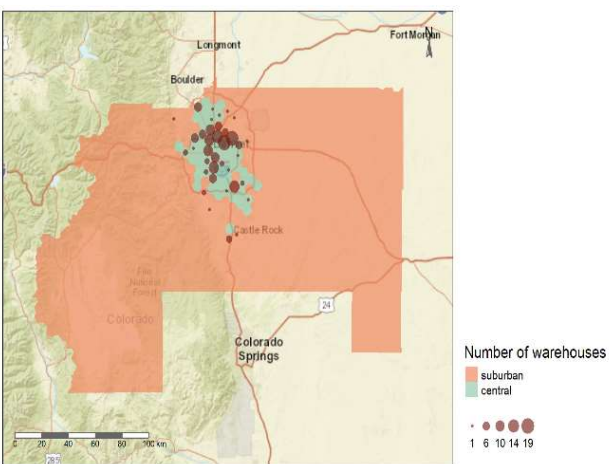
## Dallas



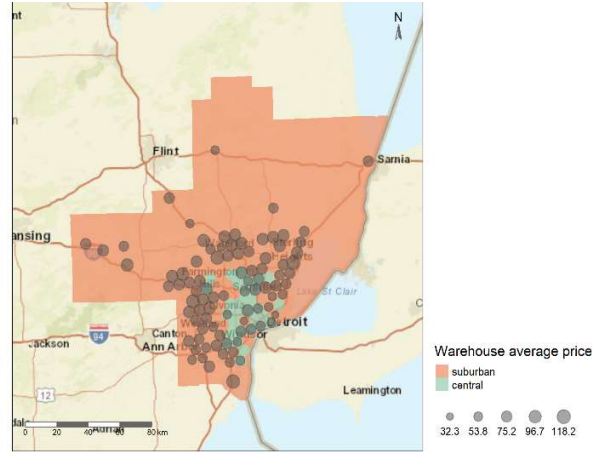
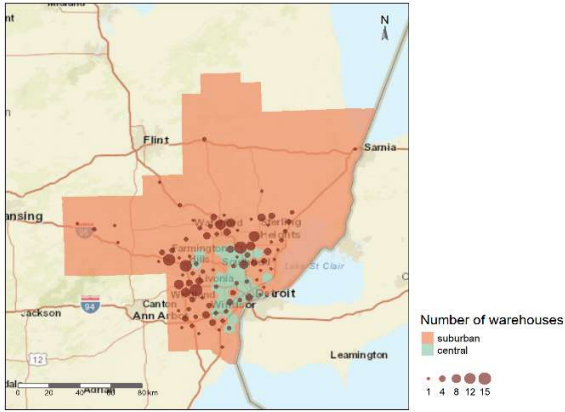
## Dayton



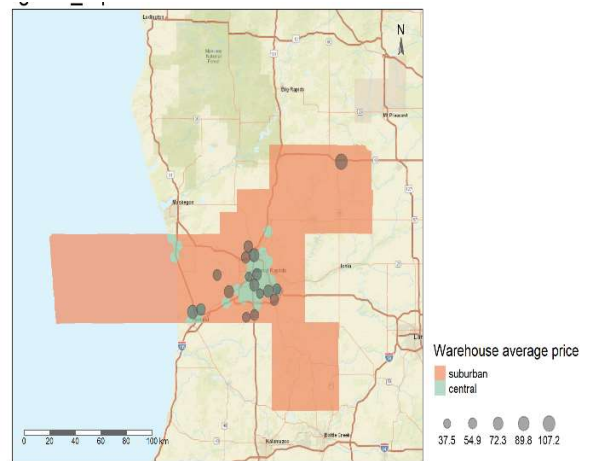
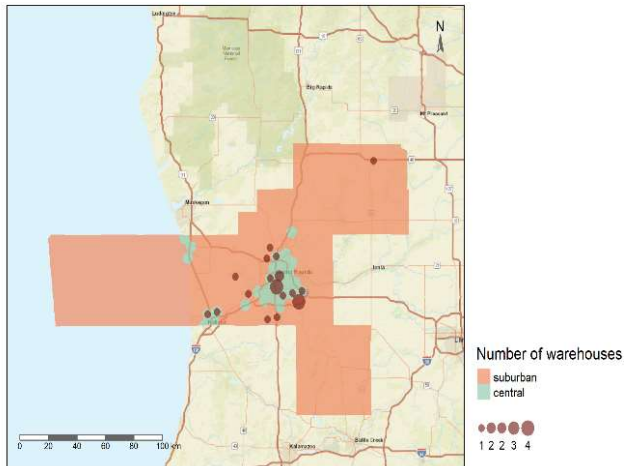
## Denver



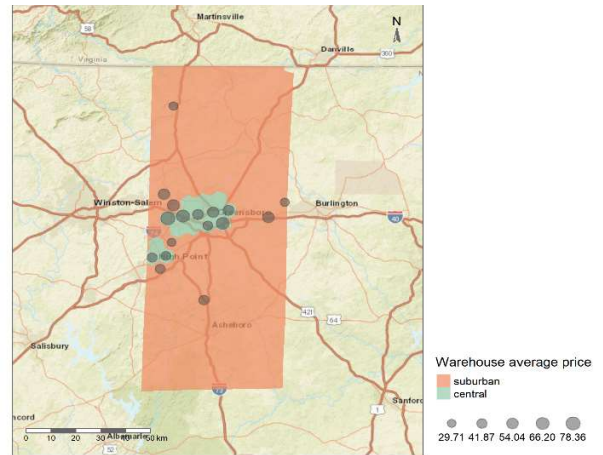
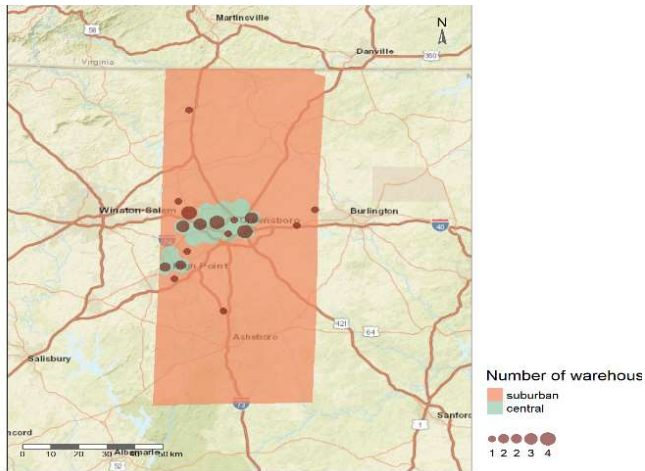
### Detroit



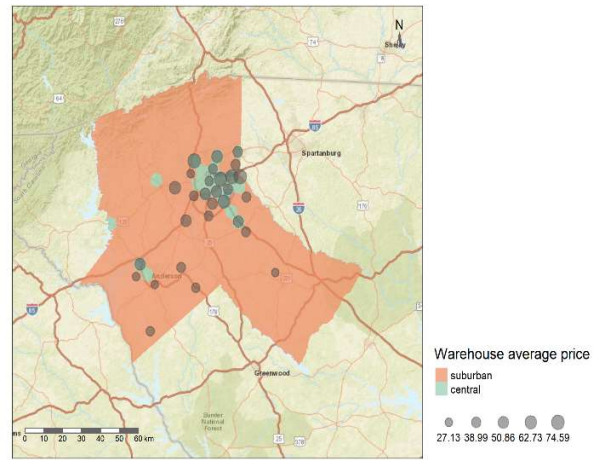
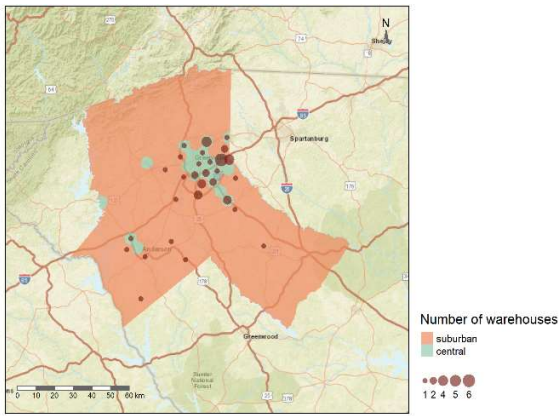
### Grand Rapids



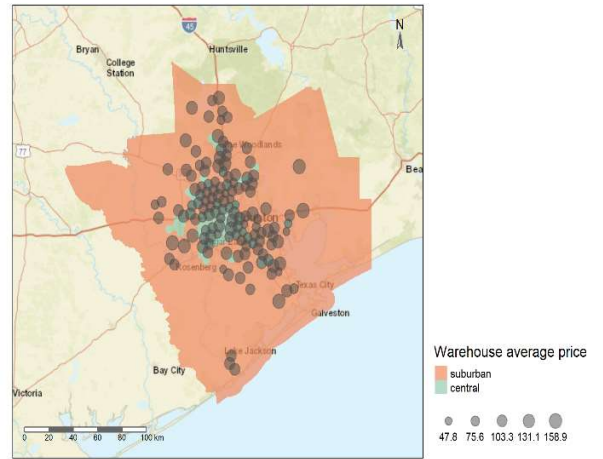
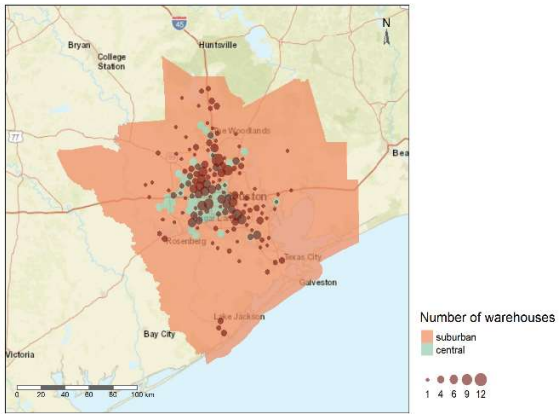
### Greensboro



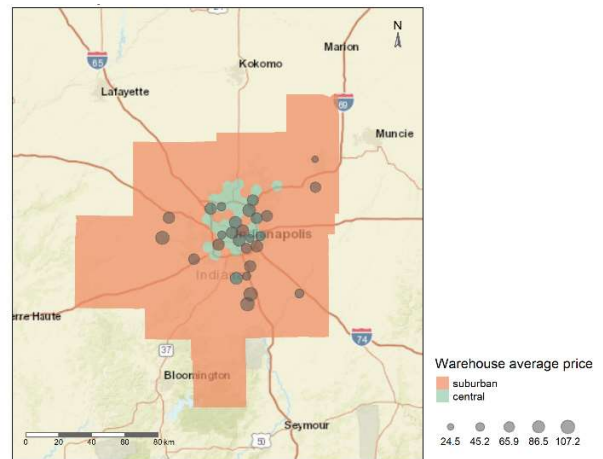
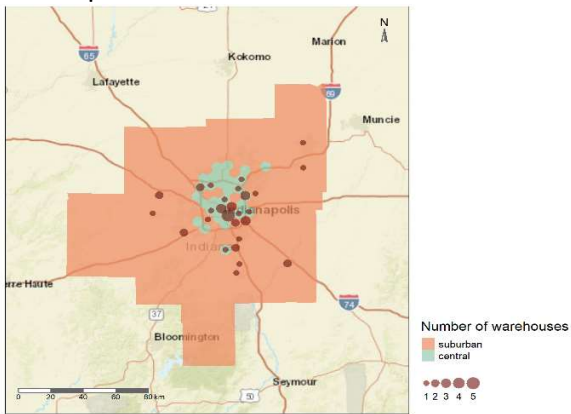
### Greenville



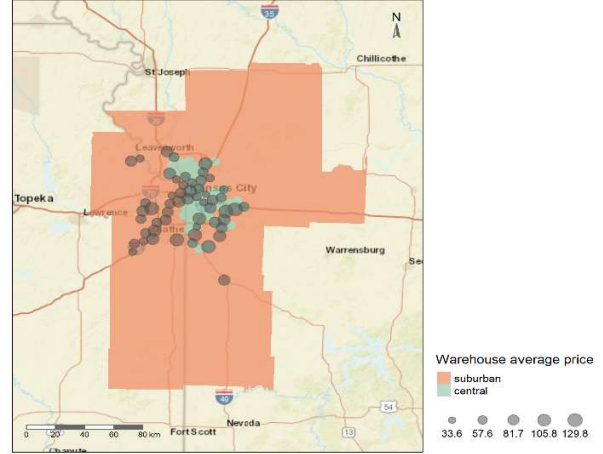
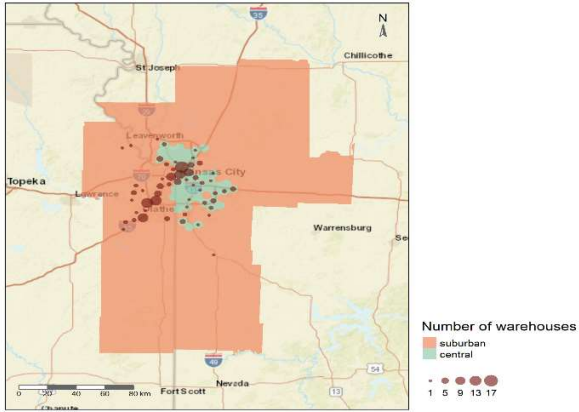
### Houston



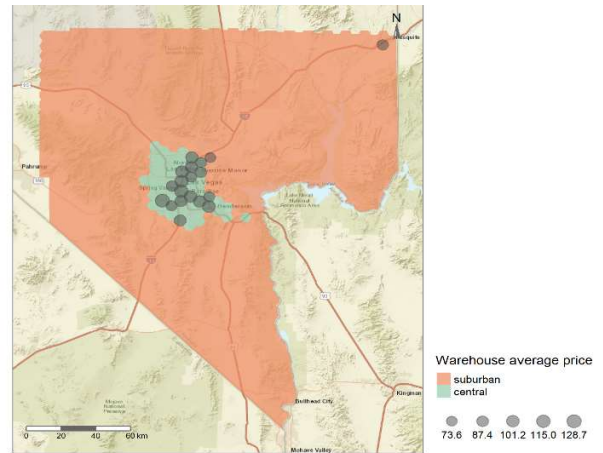
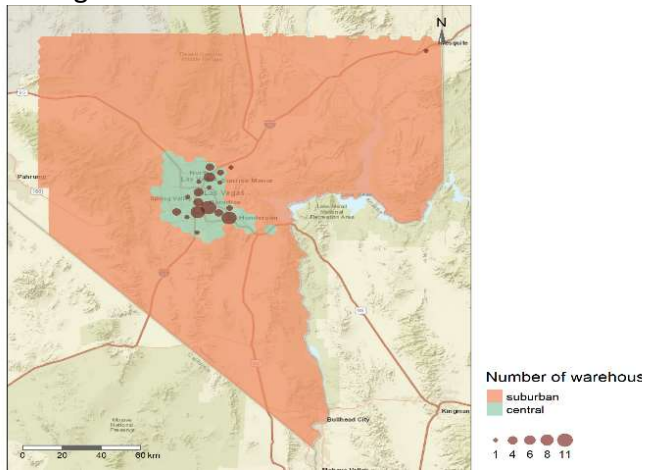
### Indianapolis



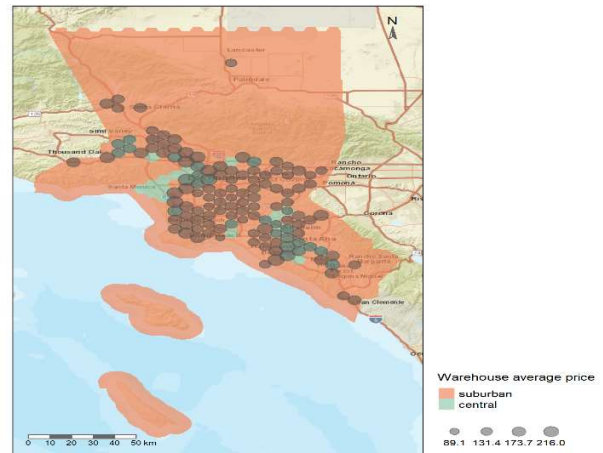
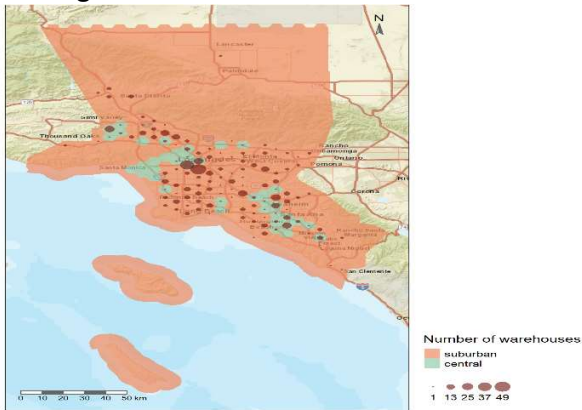
### Kansas City



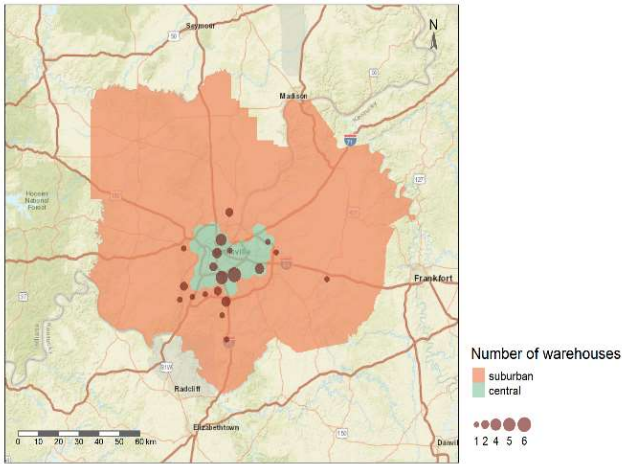
### Las Vegas



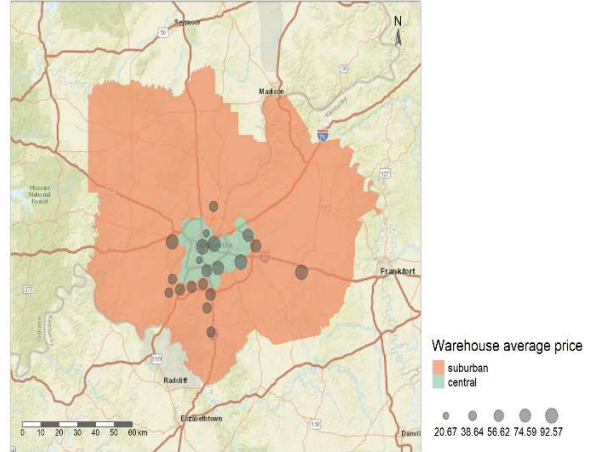
### Los Angeles



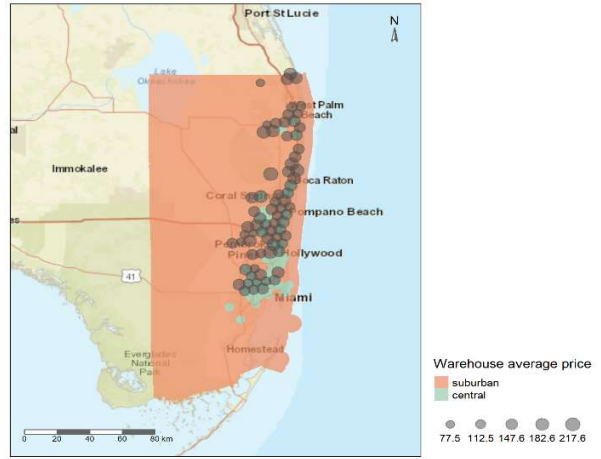
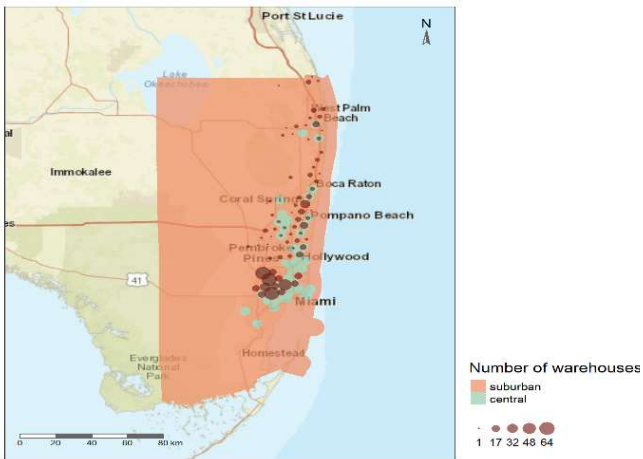
### Louisville



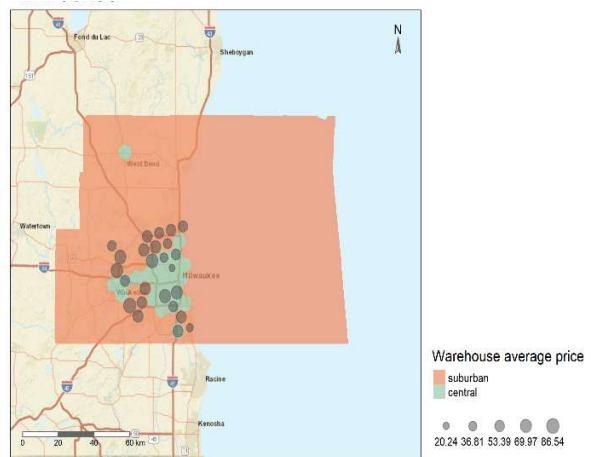
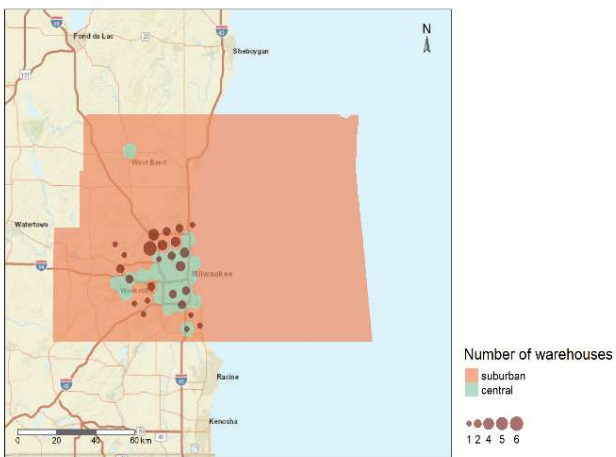
### Louisville



### Miami

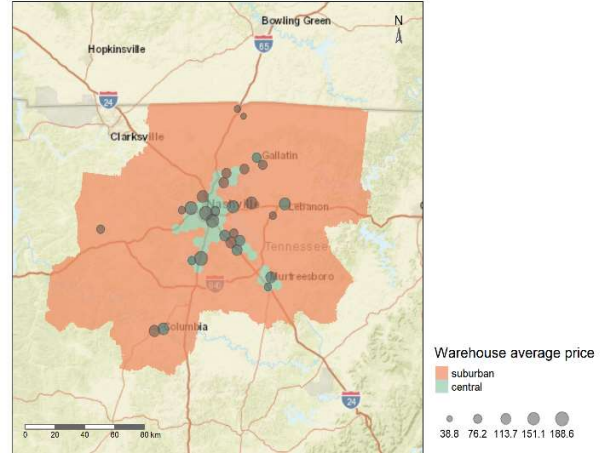
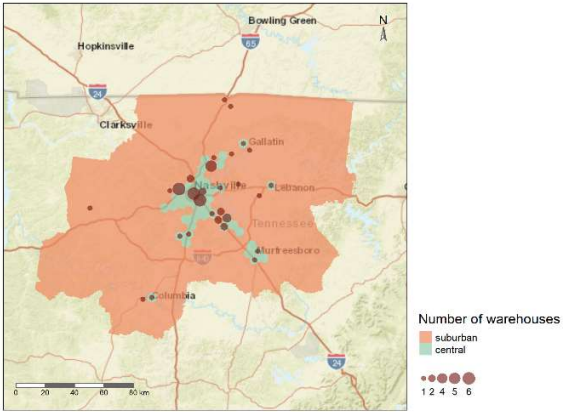


### Milwaukee

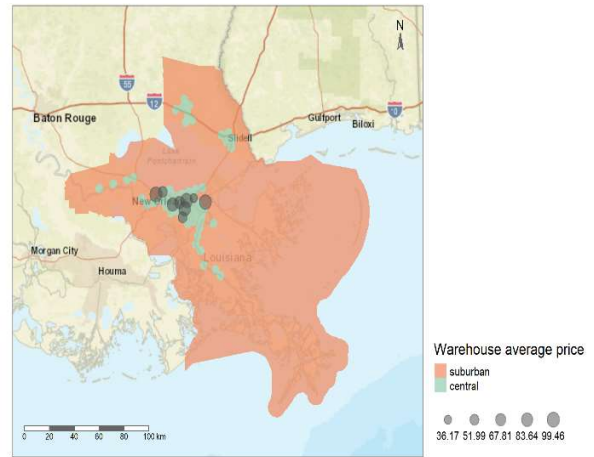
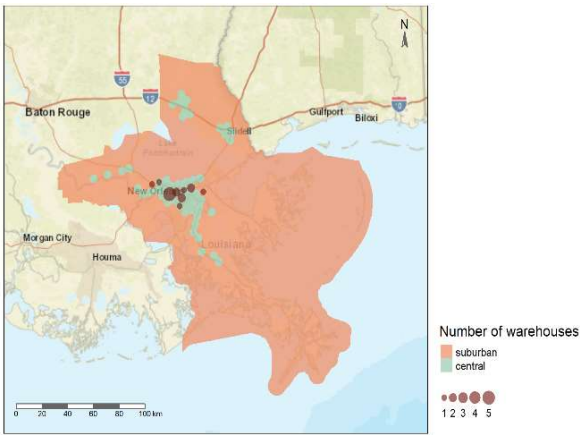




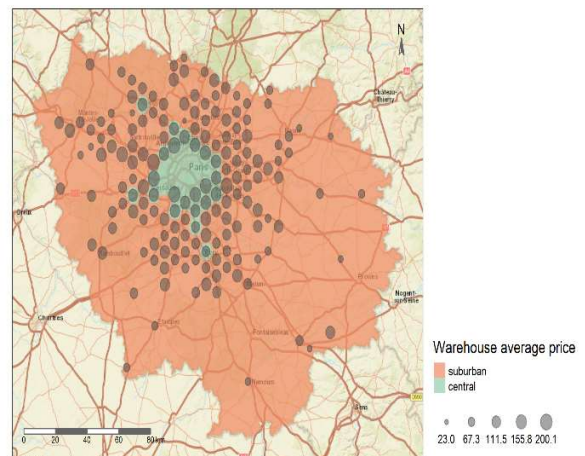
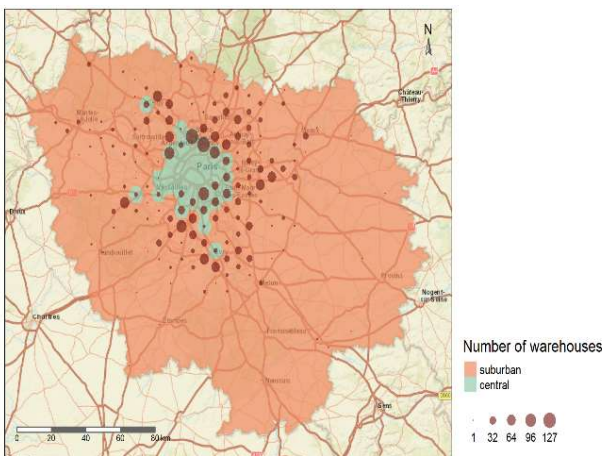
## Nashville



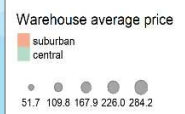
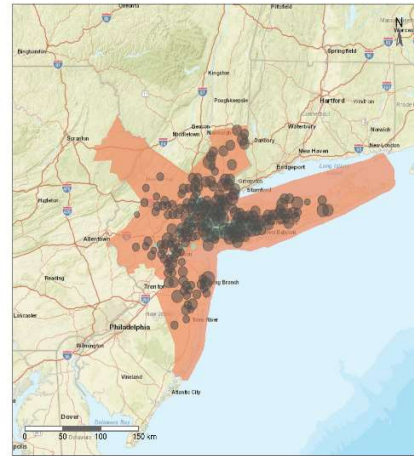
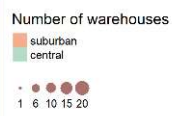
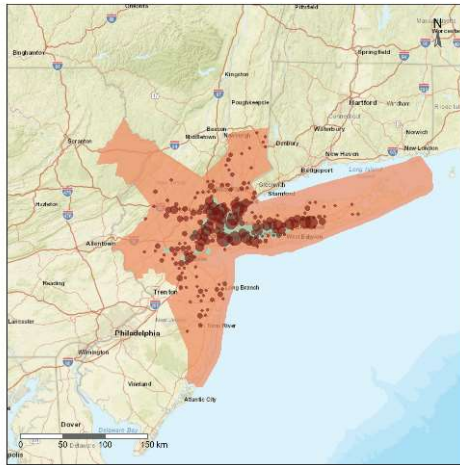
## New Orleans



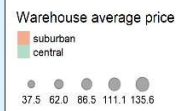
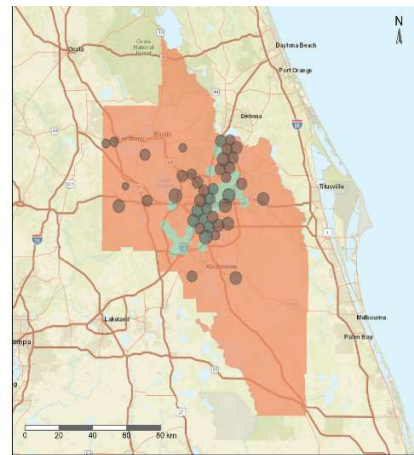
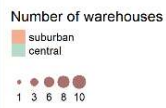
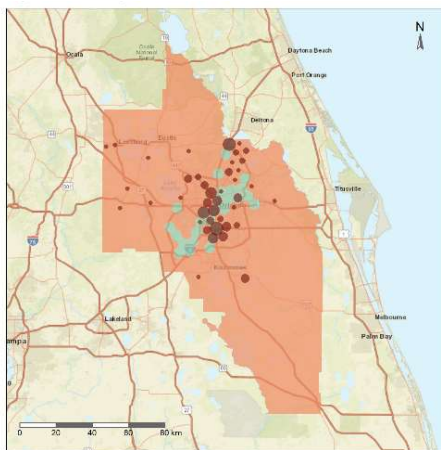
## Paris



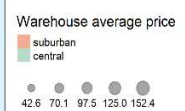
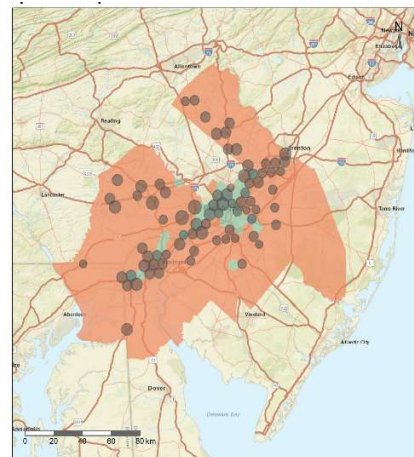
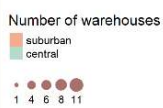
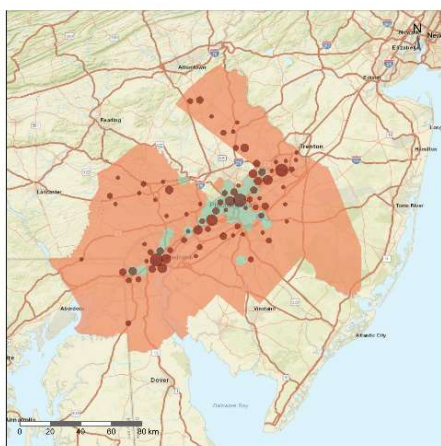
## New York



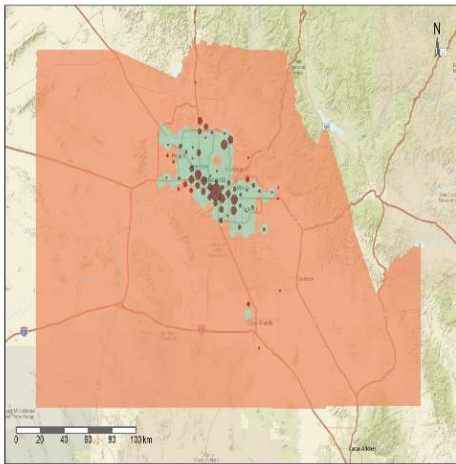
## Orlando



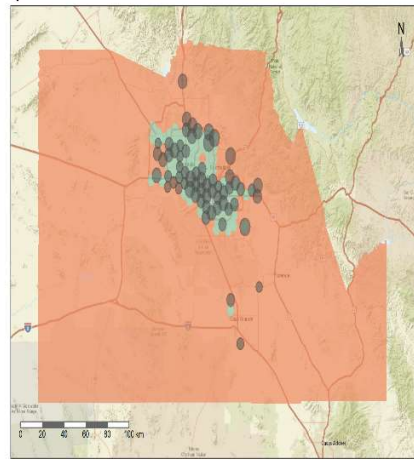
## Philadelphia



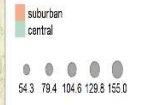
## Phoenix



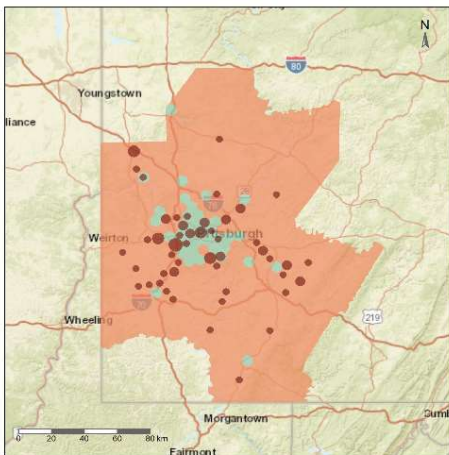
Number of warehouses



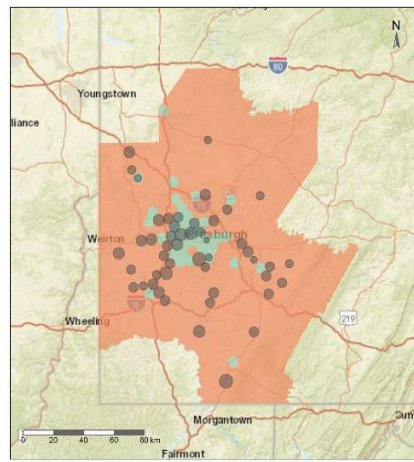
Warehouse average price



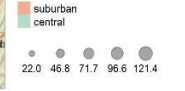
## Pittsburg



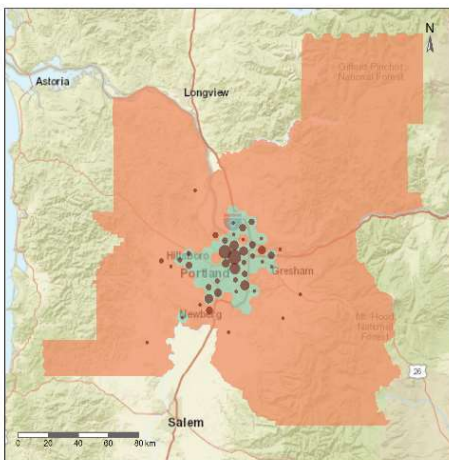
Number of warehouses



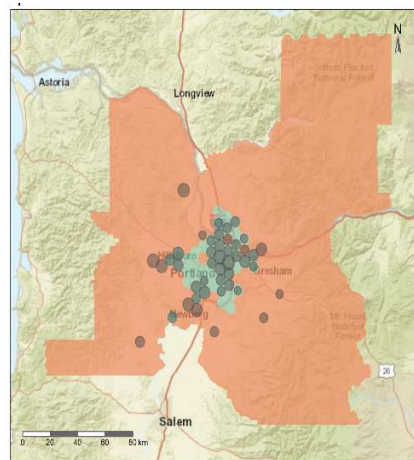
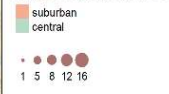
Warehouse average price



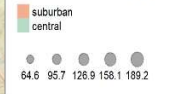
## Portland



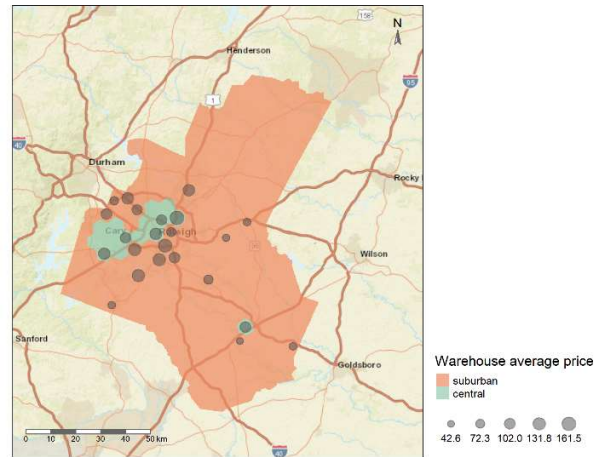
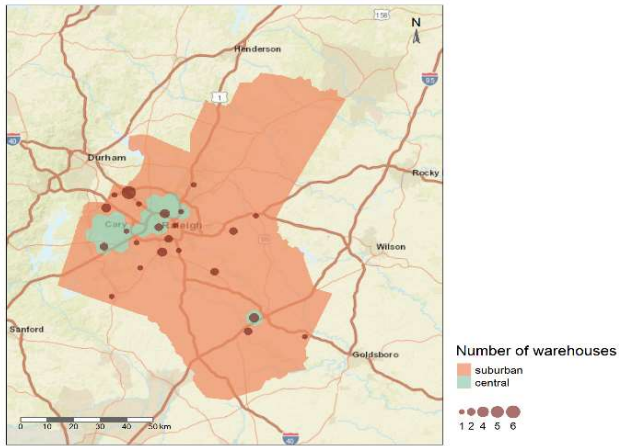
Number of warehouses



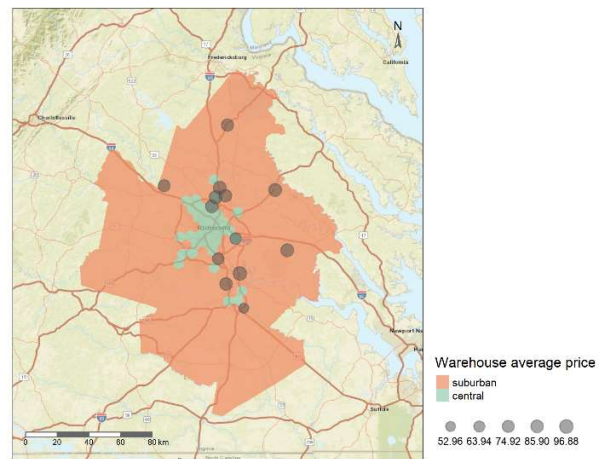
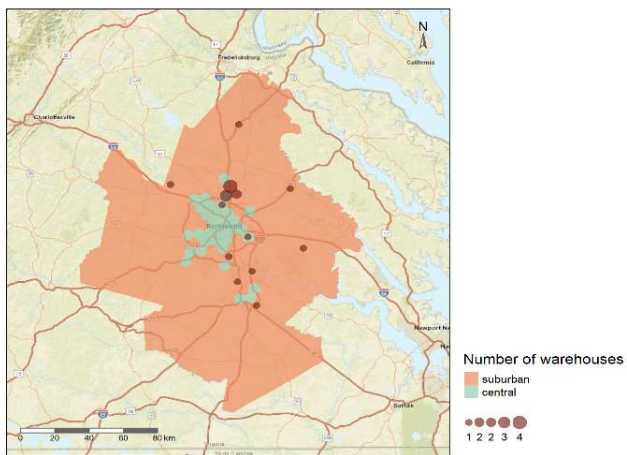
Warehouse average price



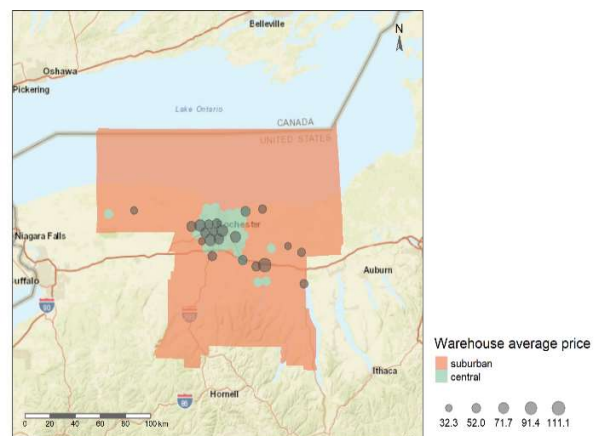
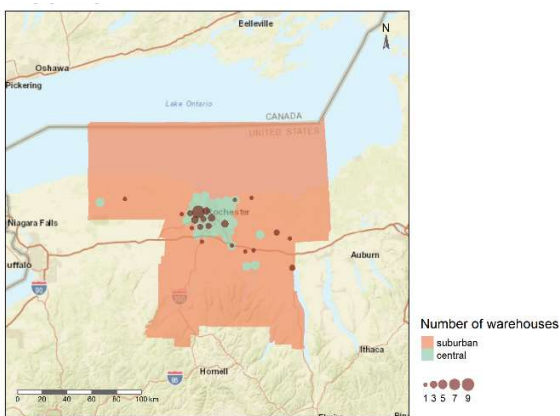
## Raleigh



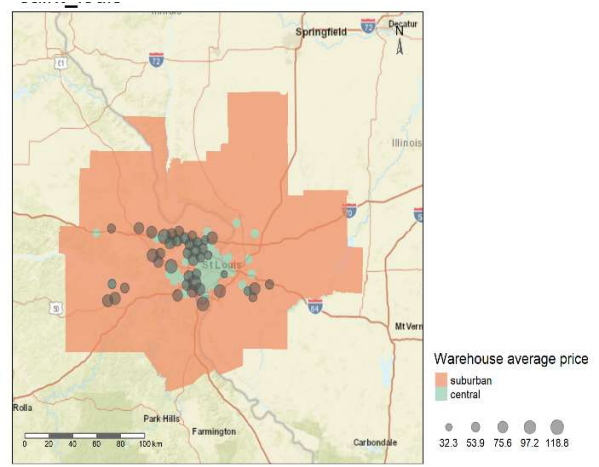
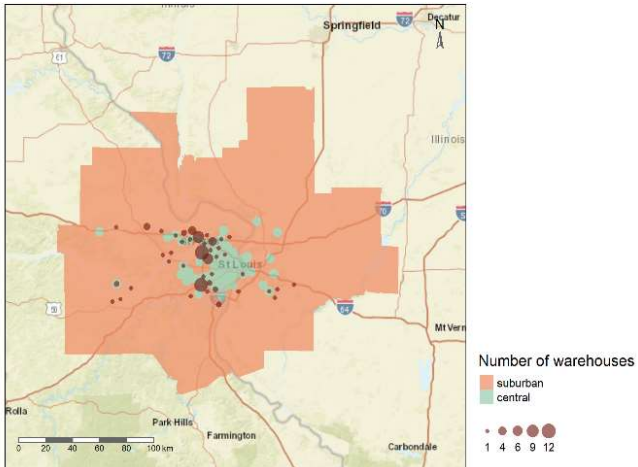
## Richmond



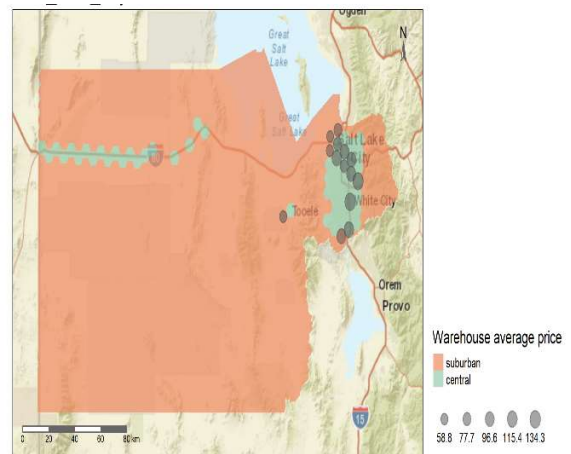
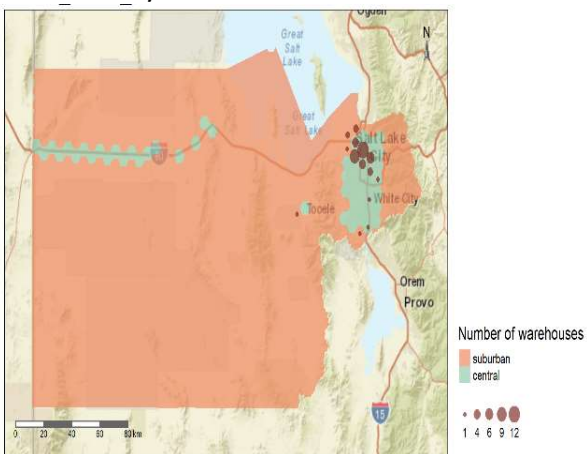
## Rochester



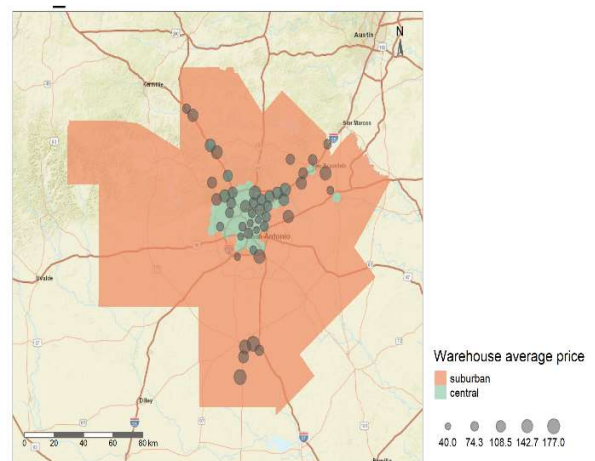
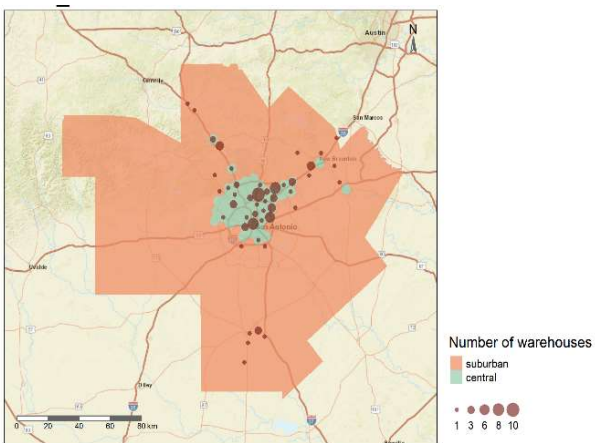
### Saint Louis



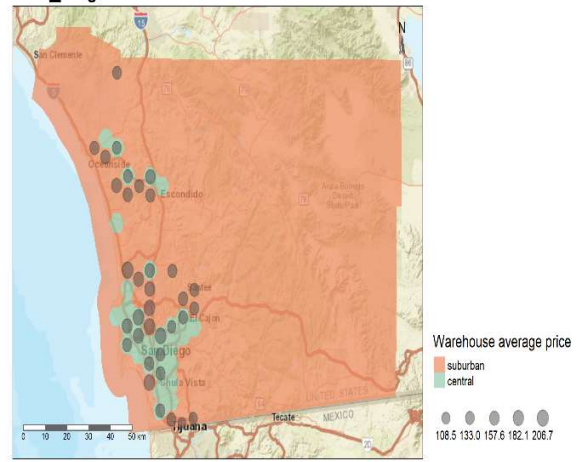
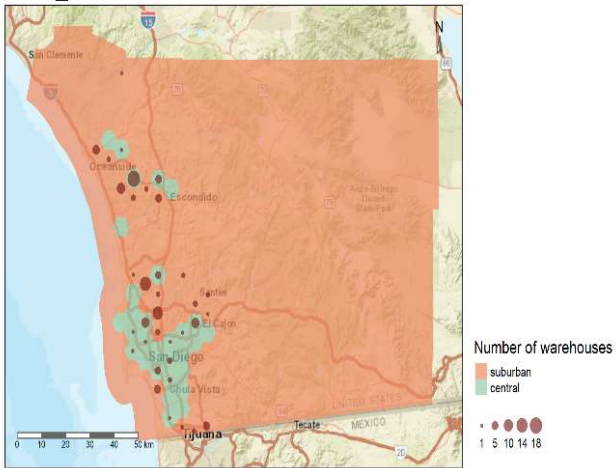
### Salt Lake City



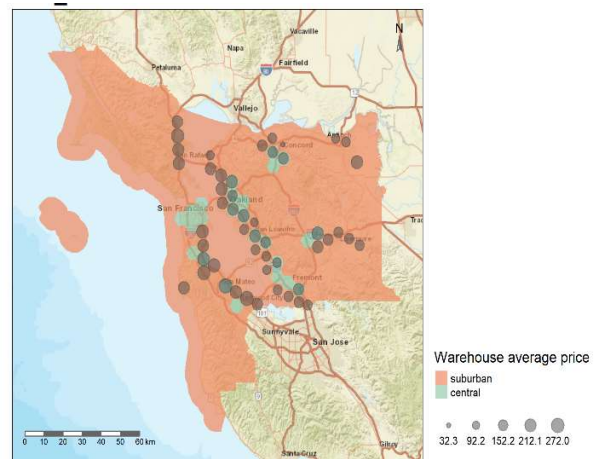
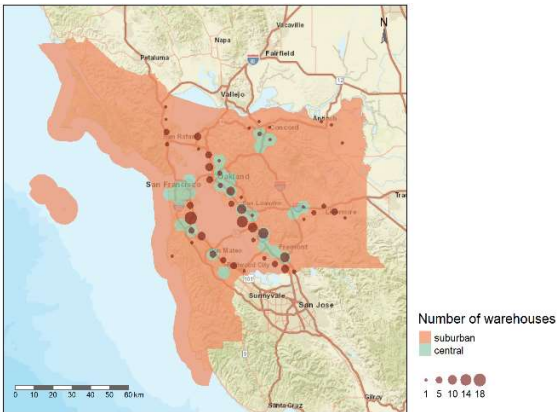
### San Antonio



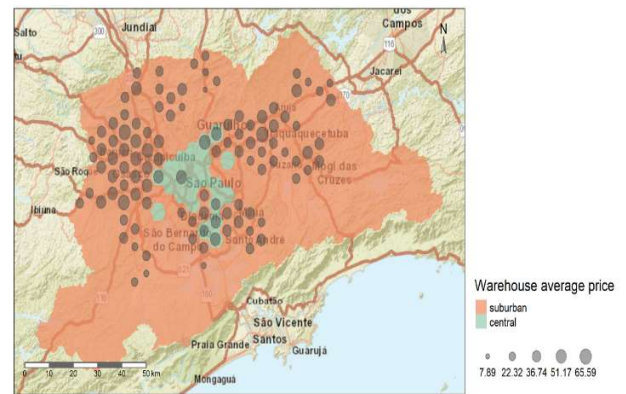
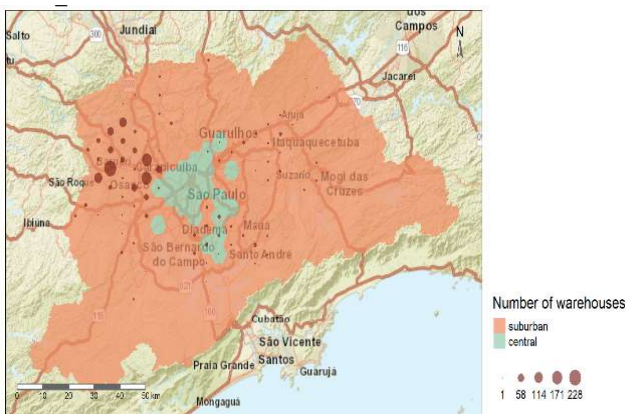
## San Diego



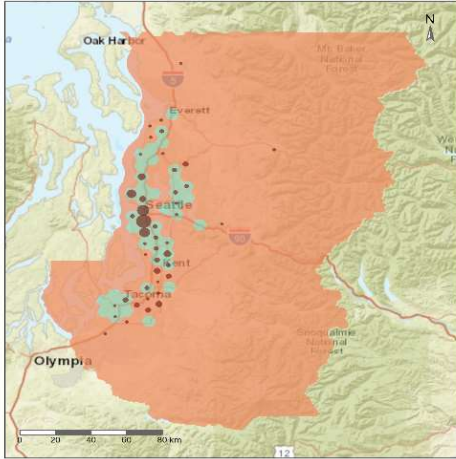
## San Francisco



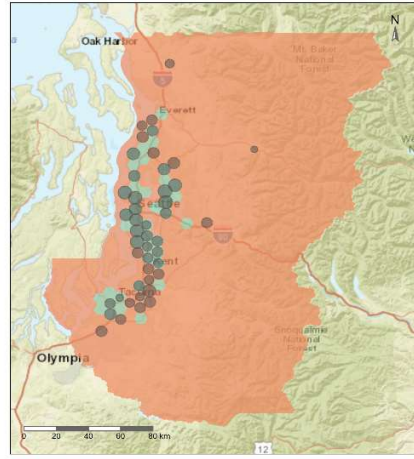
## Sao Paulo



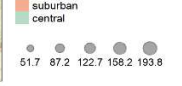
## Seattle



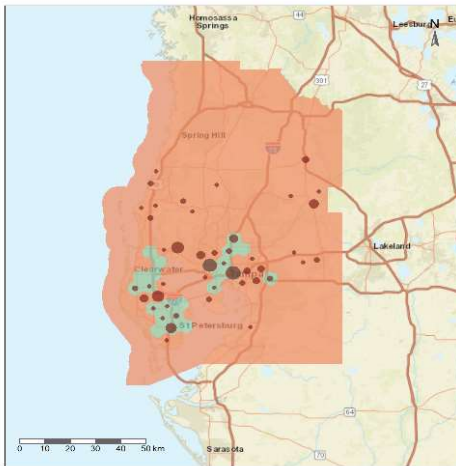
Number of warehouses



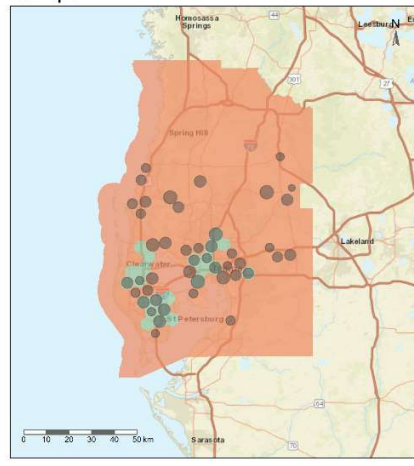
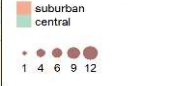
Warehouse average price



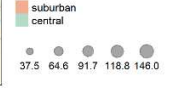
## Tampa



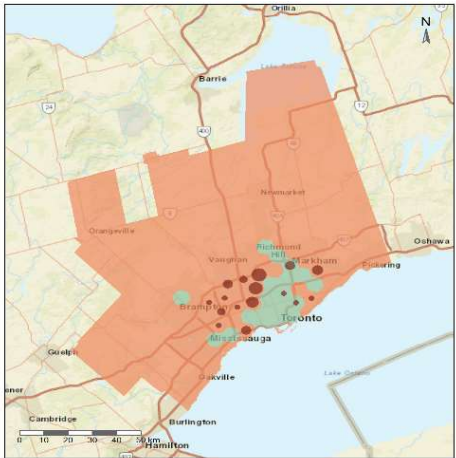
Number of warehouses



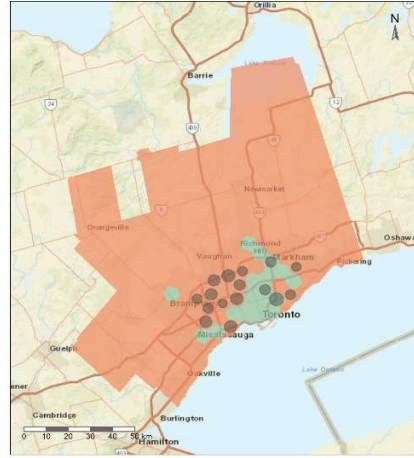
Warehouse average price



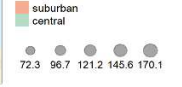
## Toronto



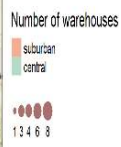
Number of warehouses



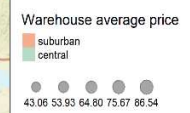
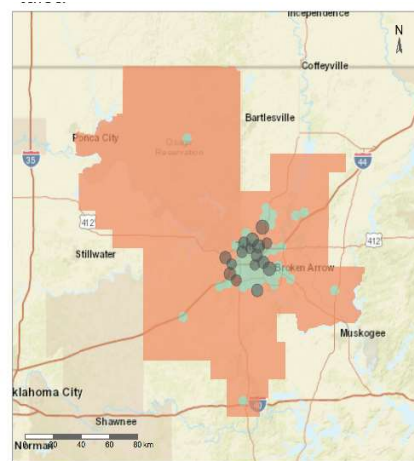
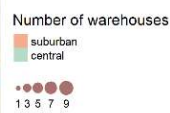
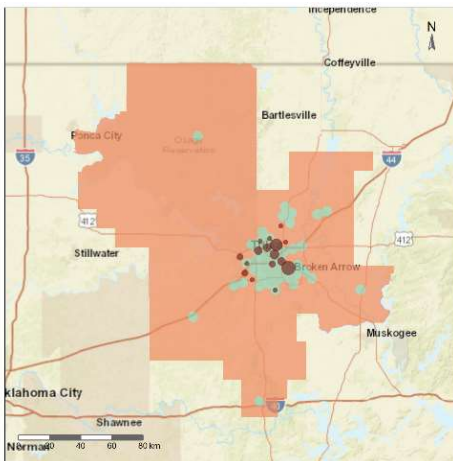
Warehouse average price



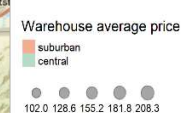
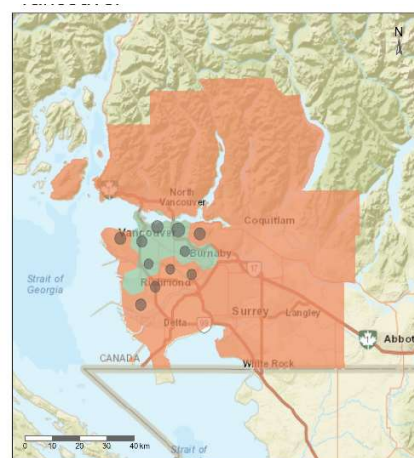
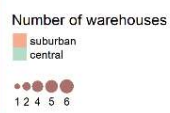
### Tucson



### Tulsa

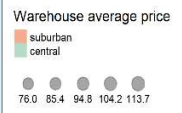
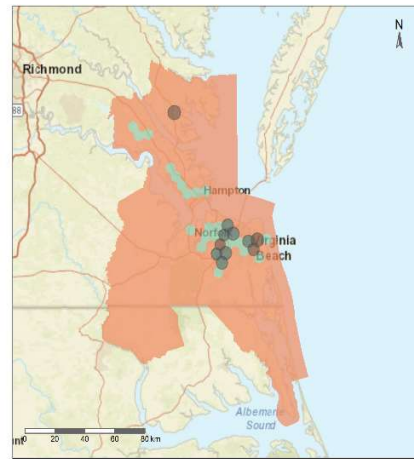
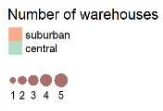
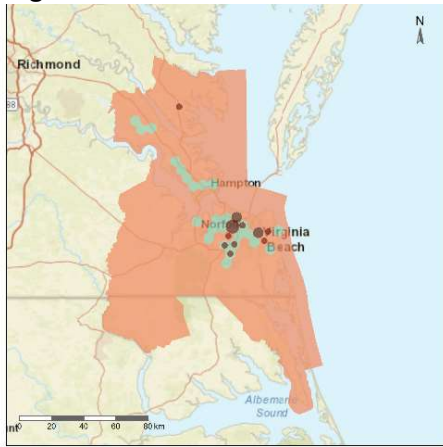


### Vancouver





## Virginia Beach



## Washington

