



# Kartografická vizualizace ESDA + GeoDa

Simona Bočková

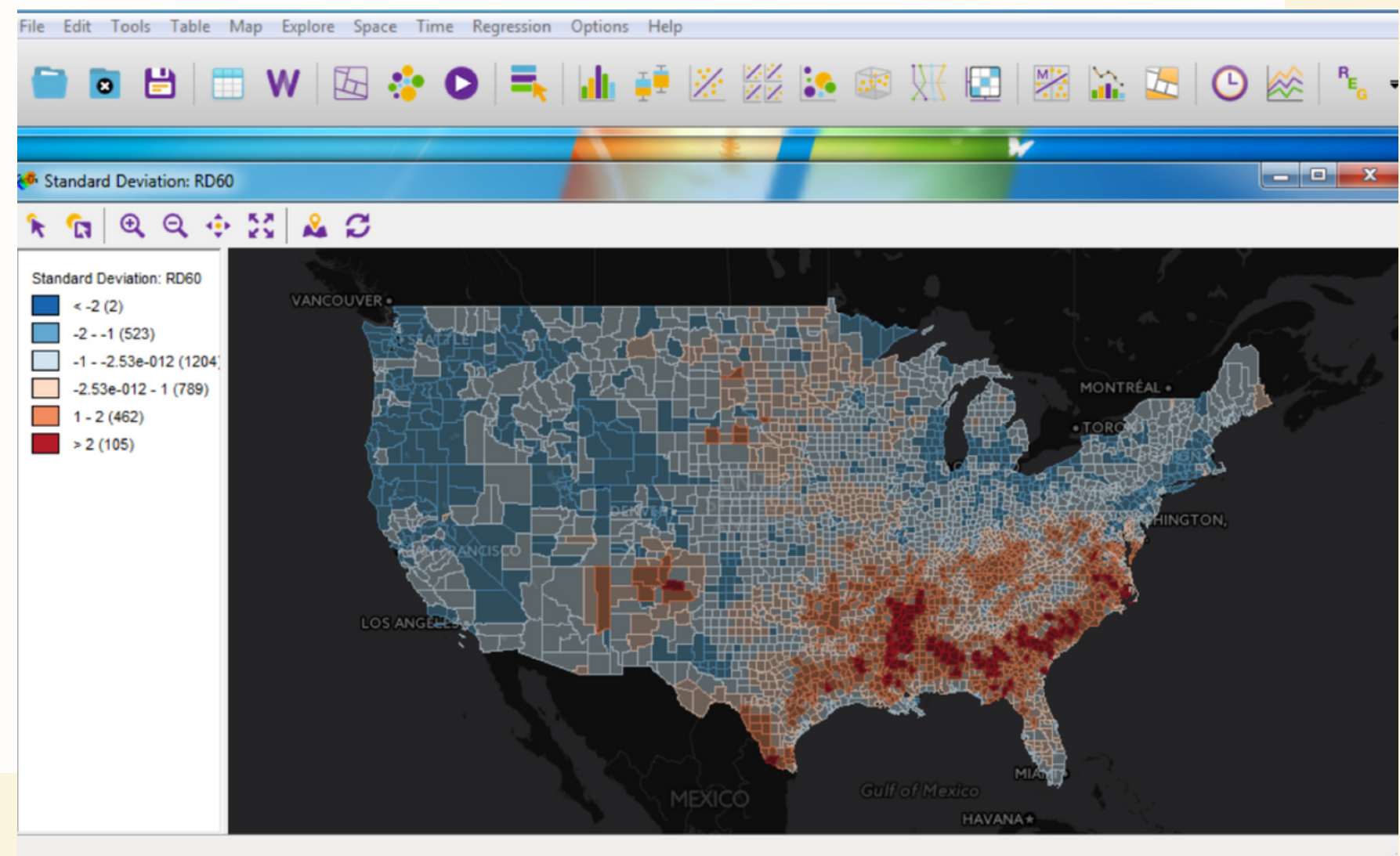
# ESDA (Exploratory Spatial Data Analysis)

= průzkumová analýza prostorových dat

- souhrn statistických metod používaných pro průzkum prostorových dat různého charakteru
- vztahuje se k prostoru - poloha a hodnota
- poskytuje informace o rozložení charakteristik

# GeoDa

- program pro metody ESDA
- podporuje různé datové formáty (shp, dbf, csv, kml, json, atd. )
- jednoduché na použití
- zdarma



# GeoDa - instalace

<https://geodacenter.github.io/download.html>

# zadání cvičení - data

- Vyberte si nějakou libovolnou datovou sadu (prostorový formát + informace obsažené v něm). Můžete si sadu i vytvořit spojením prostorových a popisných dat.
- Můžete využít data např. z Eurostatu, OSN nebo datové sady poskytované GeoDou (kromě datové sady použití v dokumentaci - NYC)
- tuto sadu nahrajte do programu **Geoda**

# zadání cvičení - zpracování

- histogram, na kterém budete zároveň demonstrovat brushing a linking
- Scatter plot
- Conditional plot - map
- parallel coordinate plot
- Dorling map

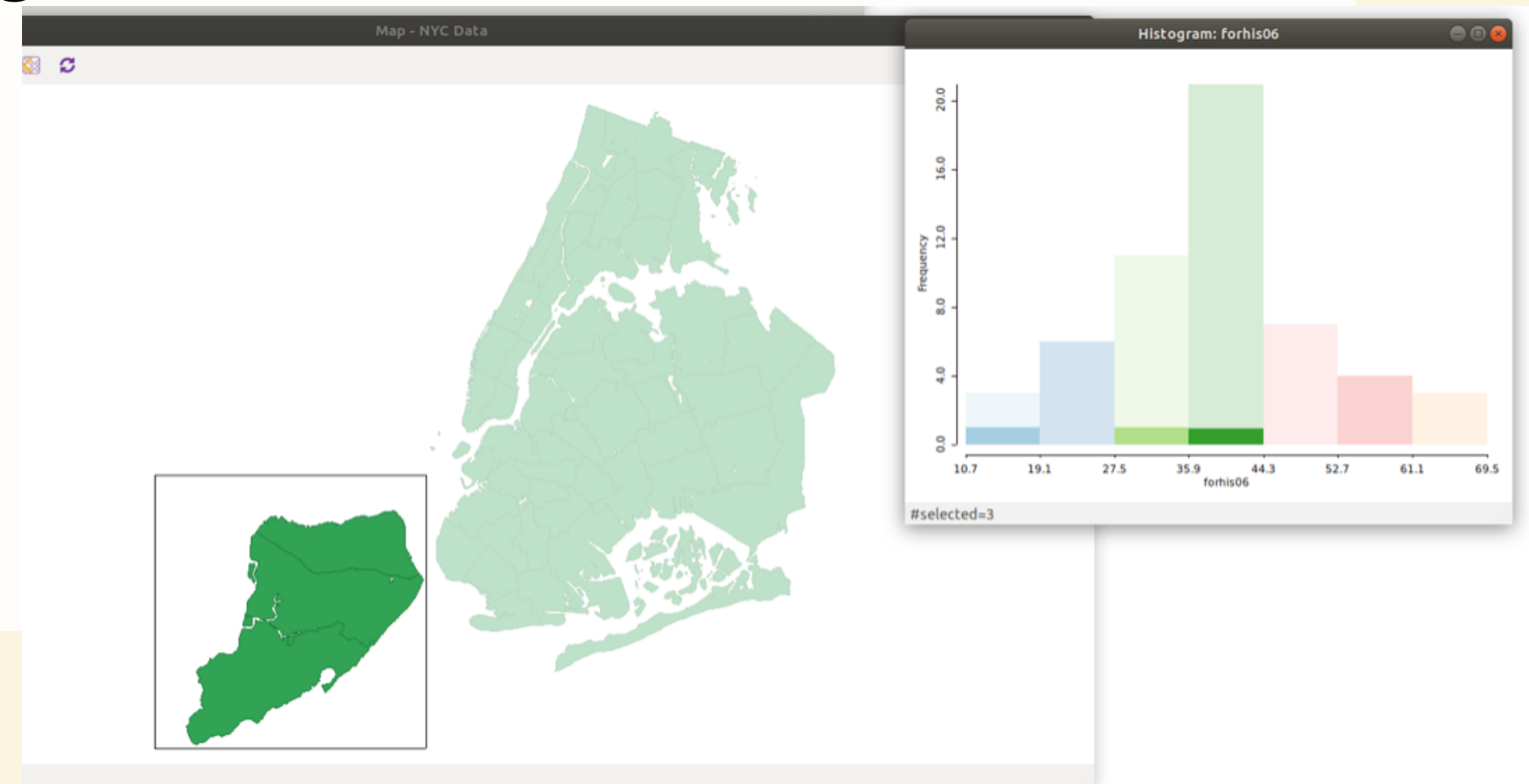


# zadání cvičení - výstupy

- Pro každou z těchto metod okomentujte její princip a okomentujte i charakteristiky a trendy, které bude zobrazovat v datech
- Následně vysvětlete a proveďte shlukování pomocí metody K-means. Vyberte počet shluků, který bude správně reprezentovat daný trend v rozložení hodnot
- **Nakonec proveďte analýzu hlavních komponent (PCA), okomentujte výsledky a popište k čemu se dá PCA využít**

# Histogram

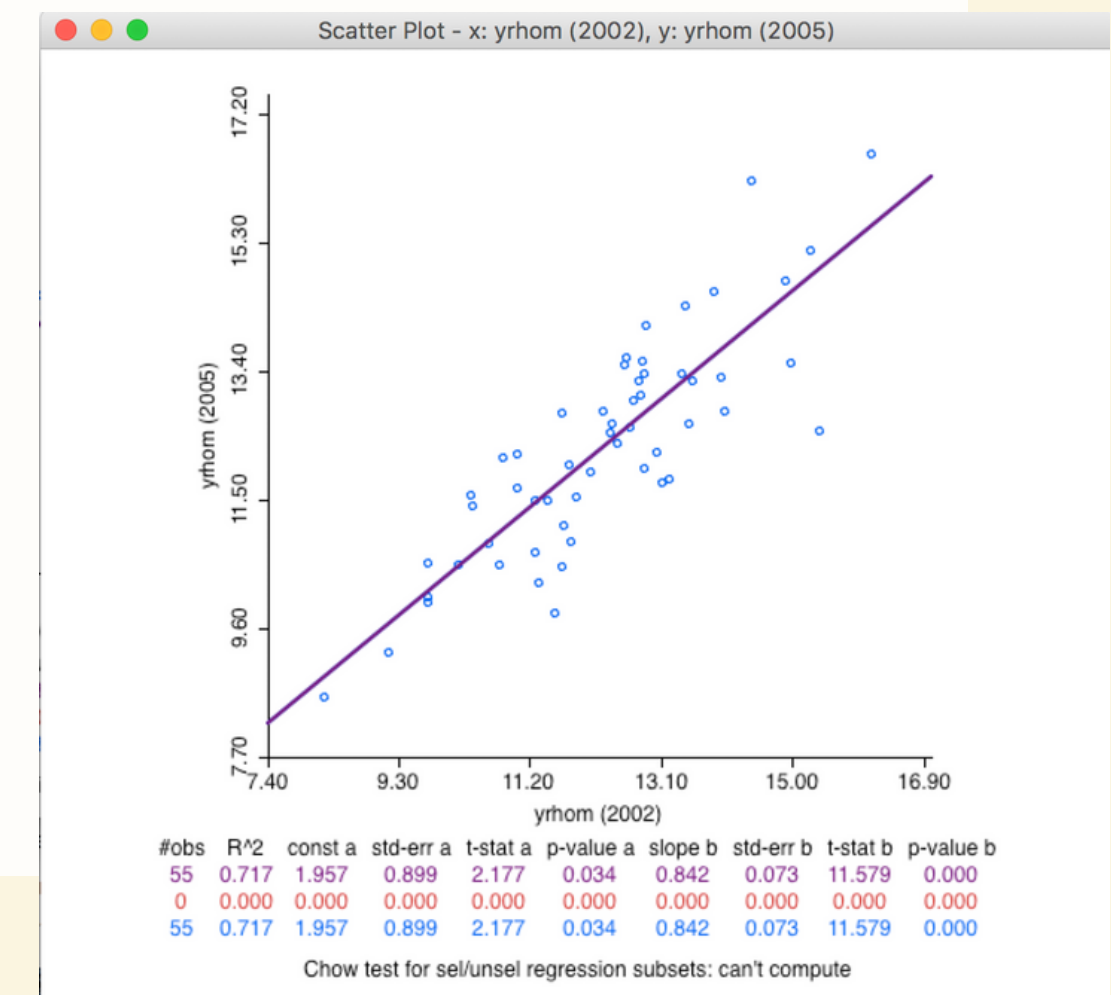
- Rozložení určitého jevu v datovém souboru
- Brushing - výběr konkrétních objektů v mapovém okně i okně histogramu
- Linking - zobrazení těchto objektů i v připojeném okně.
- Geoda --> Explore --> Histogram





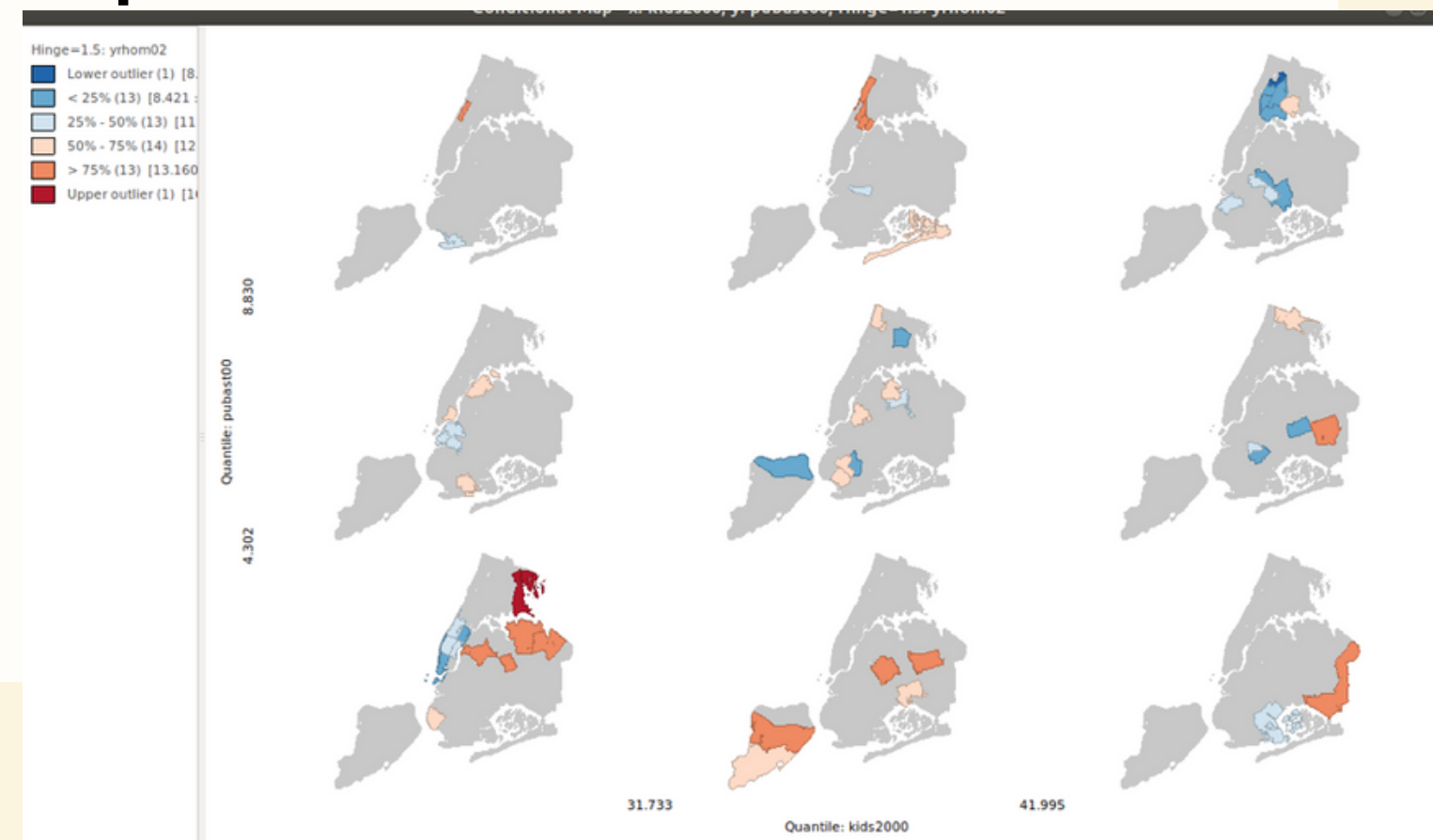
# Scatter plot

- vizualizace vztahu dvou proměnných
- každá proměnná je zobrazena na jedné ose
- lze zjistit jestli spolu některé proměnné korelují, odhalit trend nebo najít odchylky v datech
- Geoda --> Explore --> Scatter plot



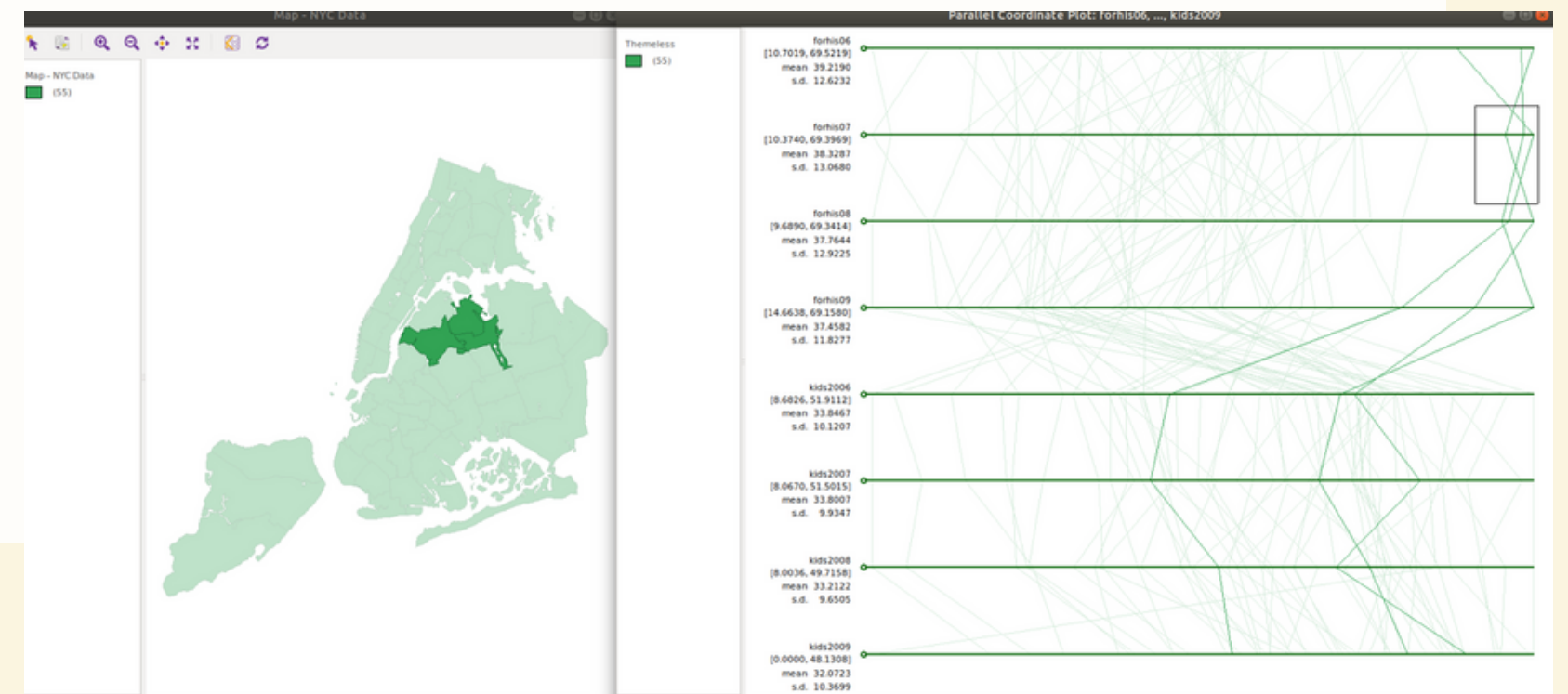
# Conditional plot - mapa

- standardně mapa 3x3
- vybrání proměnné na ose x a y
- defaultně obsahuje 2 zlomy - možnost změnit způsob a výpočet zlomů (Horizontal/vertical bins break)
- Geoda --> Conditional plot --> Map



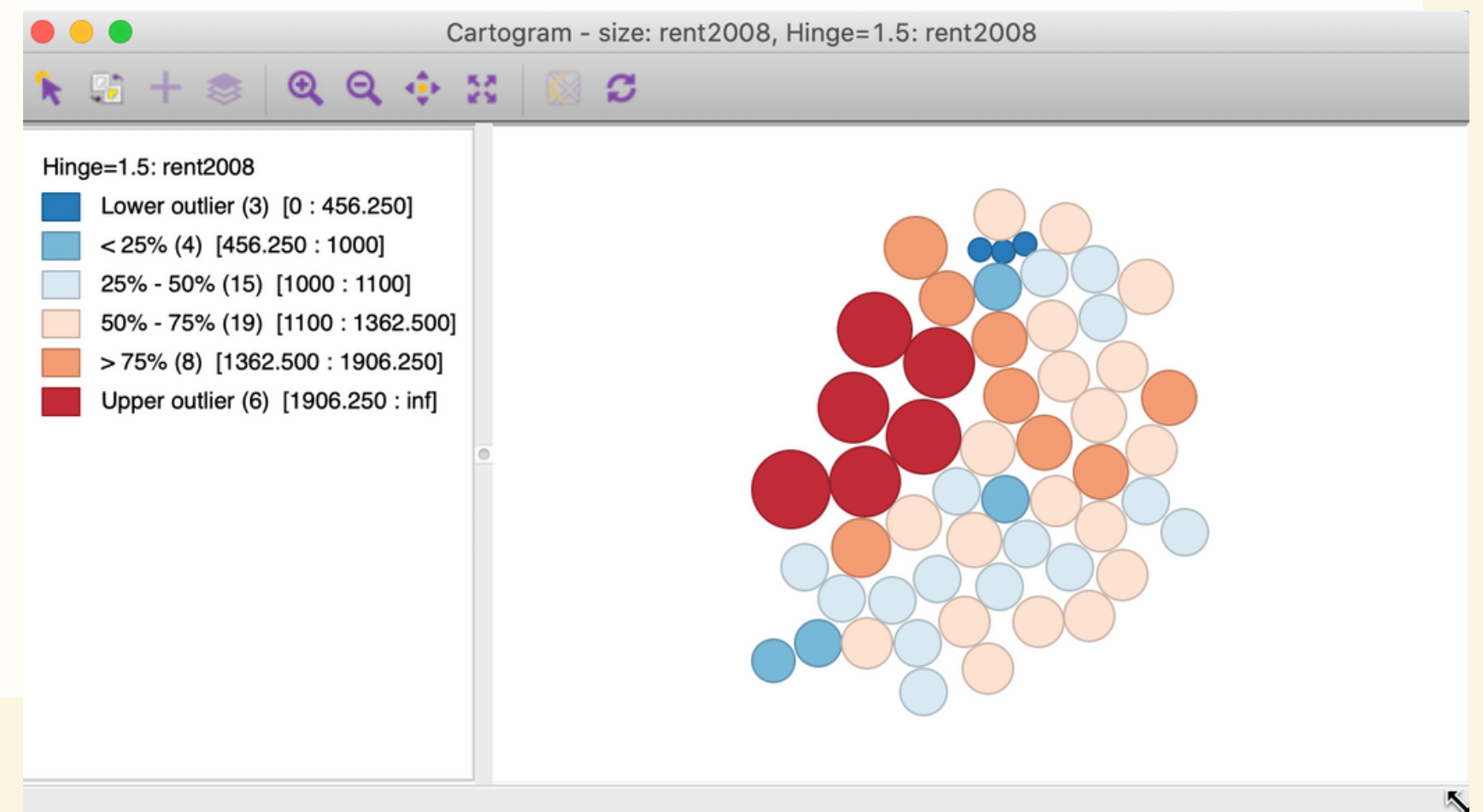
# Parallel coordinate plot

- vynesení více proměnných do grafu
- možnost sledovat změny jednotlivých charakteristik napříč objekty
- možnost odhalení extrémně vysokých a nízkých hodnot
- sledování vývoje hodnot v čase
- Explore --> Parallel Coordinate Plot



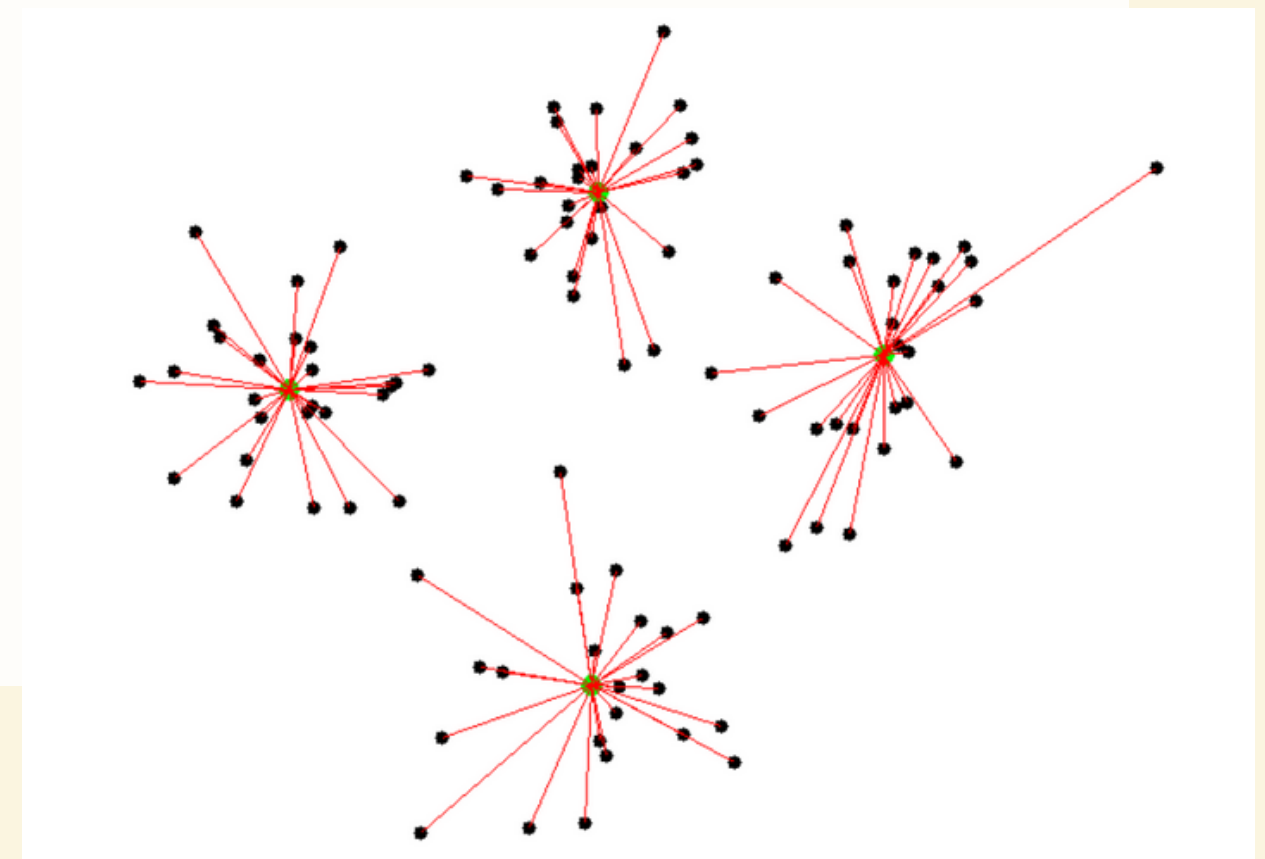
# Dorling map

- druh anamorfózy
- Zachovávají se pouze středy objektů (nahrazeny jiným objektem)
- lze odhalit prostorové souvislosti sledovaných proměnných
- Map -> Cartogram



# Shlukování

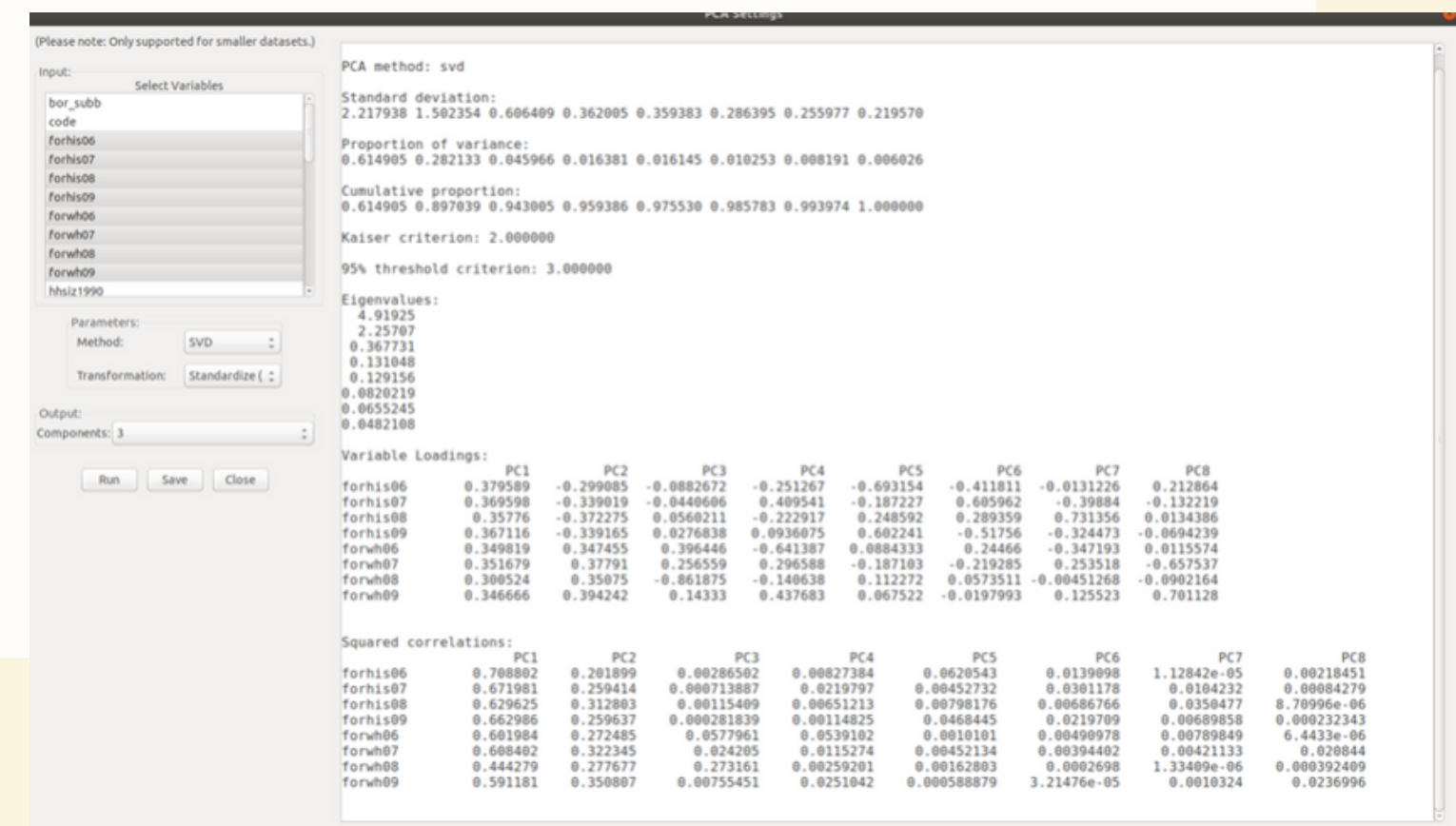
- jedna z nejjednodušších metod je K-means
  - definuje náhodné pozice centroidů (cílový počet shluků)
  - další iterace --> nové přepočítané pozice
  - ukončeno jakmile se pozice nemění nebo je ukončen počet iterací
- Clusters --> K Means.





# PCA analýza hlavních komponent

- redukce dimenzionality vícerozměrných dat při zachování co největší variability v datech
- princip: rotace osu ve směru nejvyšší variability dat
  - nejvíce původní informace + největší úspora dimenzionality (počet proměnných)
- Clusters -> PCA





# inspirace I.

- popis jednotlivých metod Geody a jejich ukázka na mapových výstupech - DP (2016)

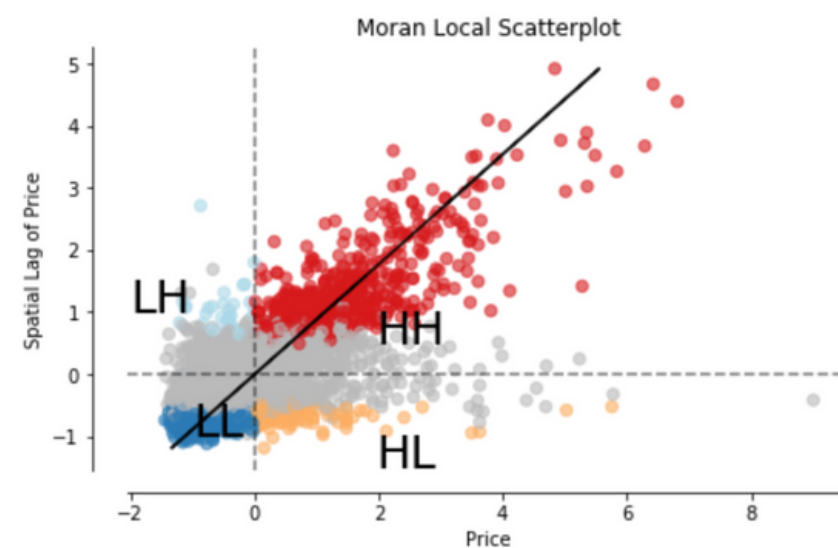
[https://is.muni.cz/th/cpxqx/Diplomova\\_prace\\_Havlickova\\_finish](https://is.muni.cz/th/cpxqx/Diplomova_prace_Havlickova_finish)

- dokumentace k metodám Geody

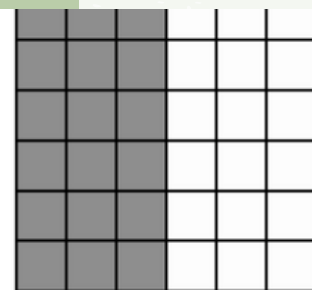
<https://geodacenter.github.io/documentation.html>

# inspirace II.

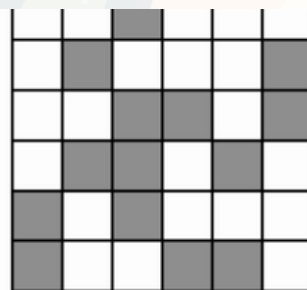
- použití ESDA pro analýzu Airbnb v Londýně:  
<https://towardsdatascience.com/what-is-exploratory-spatial-data-analysis-esda-335da79026ee>



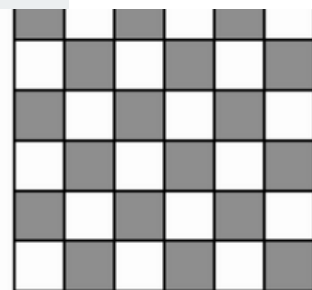
Moran Local Scatter Plot — LISA



Positive spatial autocorrelation

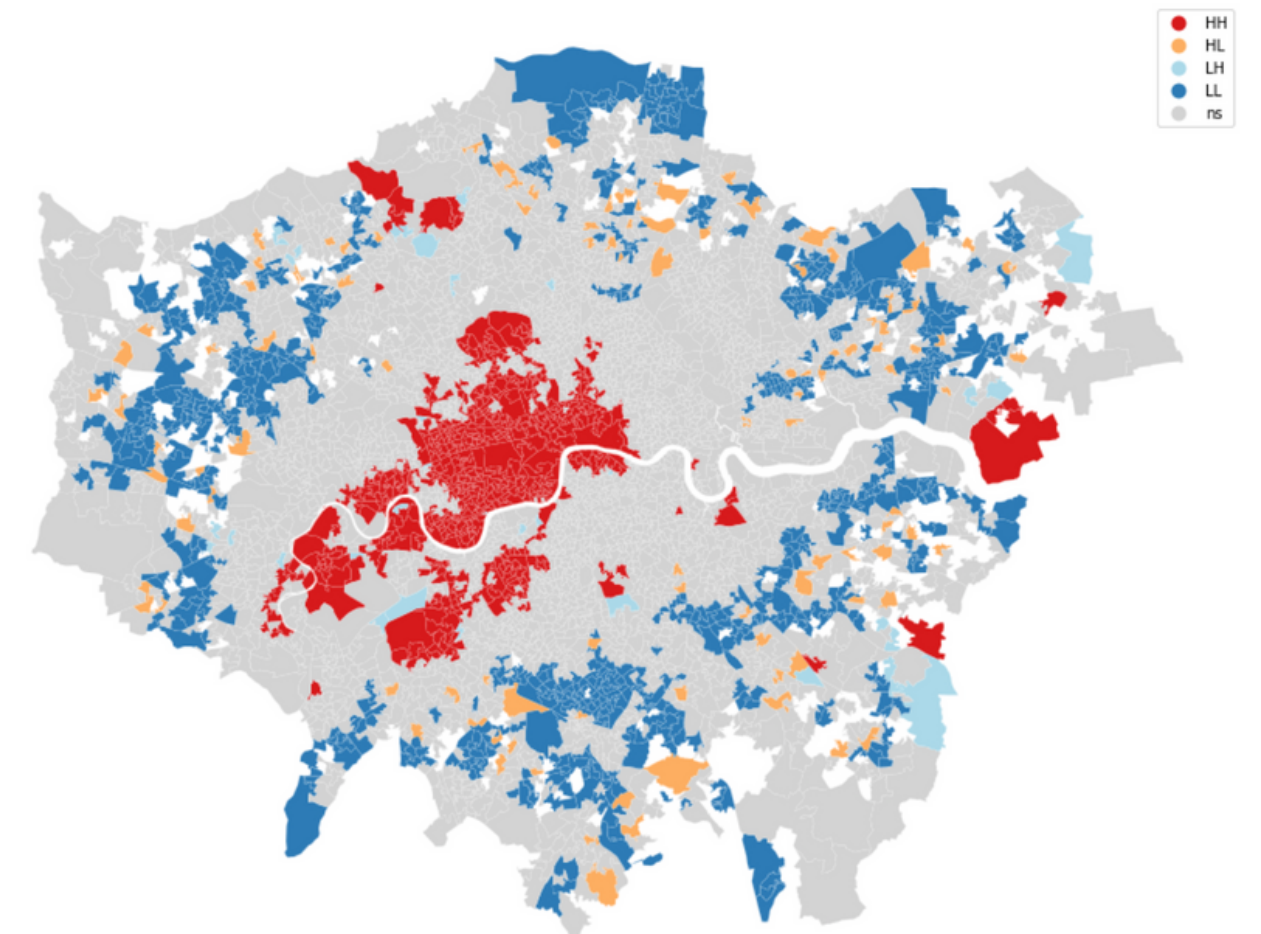


No spatial autocorrelation



Negative spatial autocorrelation

Spatial Autocorrelation: [Source](#)



LISA Cluster Map -Airbnb Average price per neighborhood.

konec

