

# GIS4SG

## Explorace prostorových dat

podzim 2023

**Lukáš Herman**

[herman.lu@mail.muni.cz](mailto:herman.lu@mail.muni.cz)

**Laboratory on Geoinformatics and Cartography (LGC)  
Institute of Geography  
Masaryk University  
Czech Republic**

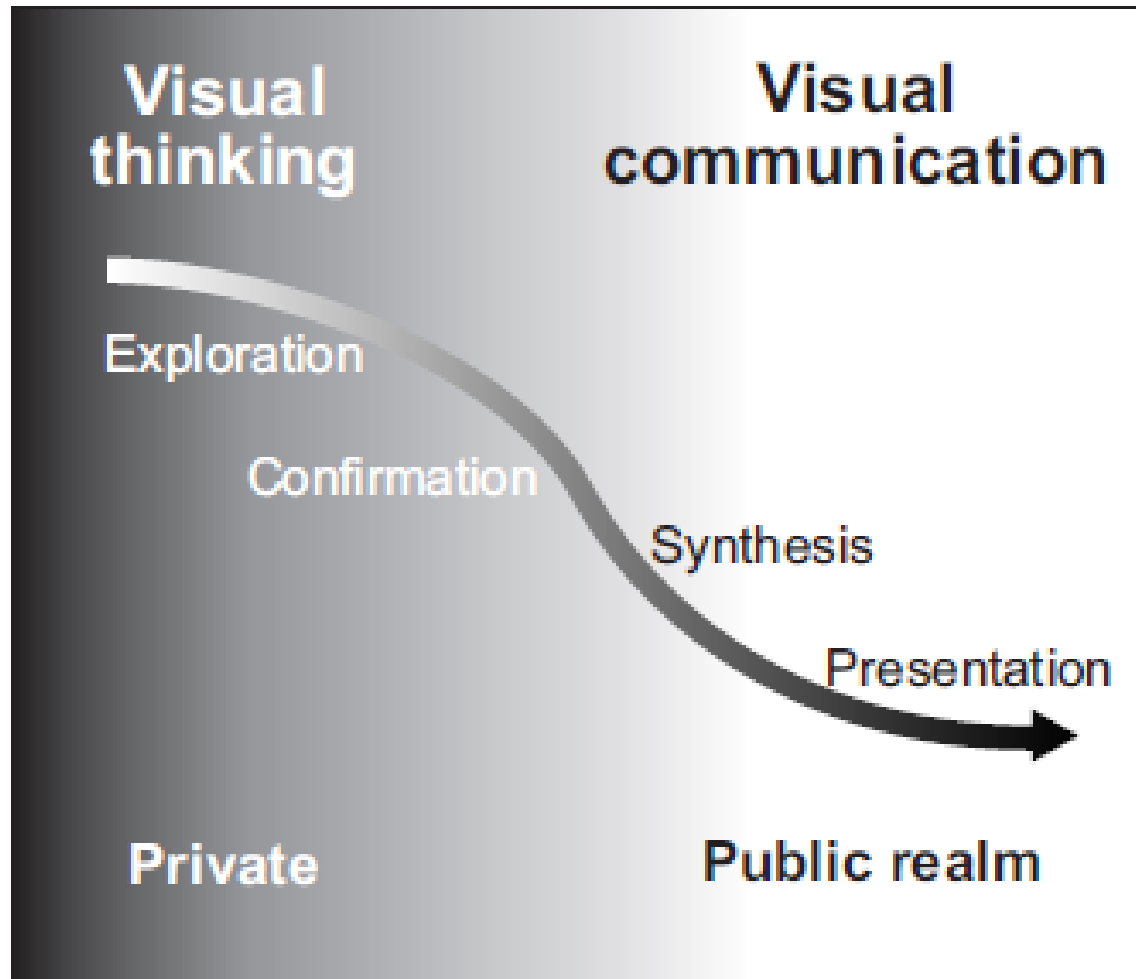


# Užití mapy (Map use)

- **Změna kartografického paradigmatu na začátku 90. let 20. století**
  - Jak mohou mapy lépe sloužit uživatelům?
- **Důraz na kartografickou vizualizaci**
- **Di Biase (1990) – vizualizace jako nástroj výzkumu.**
  - Viz další slide. Křivka vizualizace zobrazuje sekvenci výzkumných kroků a odlišnou roli vizualizace.
- **Na straně explorace hraje mapa a další typy vizualizace roli nástroje pro zdůvodnění (napomáhají vizuálnímu myšlení).**



# Mapa a vizualizace v geografickém výzkumu





# „Prostor“ užití mapy

- MacEachren(1994) – konceptuální pohled.
- Vizualizace (ve smyslu vizuálního myšlení) je doplněk komunikace (ve smyslu přenosu informace).
- Možnost souvislého měření podél tří os:
  - **Private** – public
  - **Revealing unknown** – presenting known
  - **High human-map interaction** – low interaction
- Kartografická vizualizace = užití mapy v části krychle pro soukromé využití, odhalující nové a s vysokou mírou interakce.



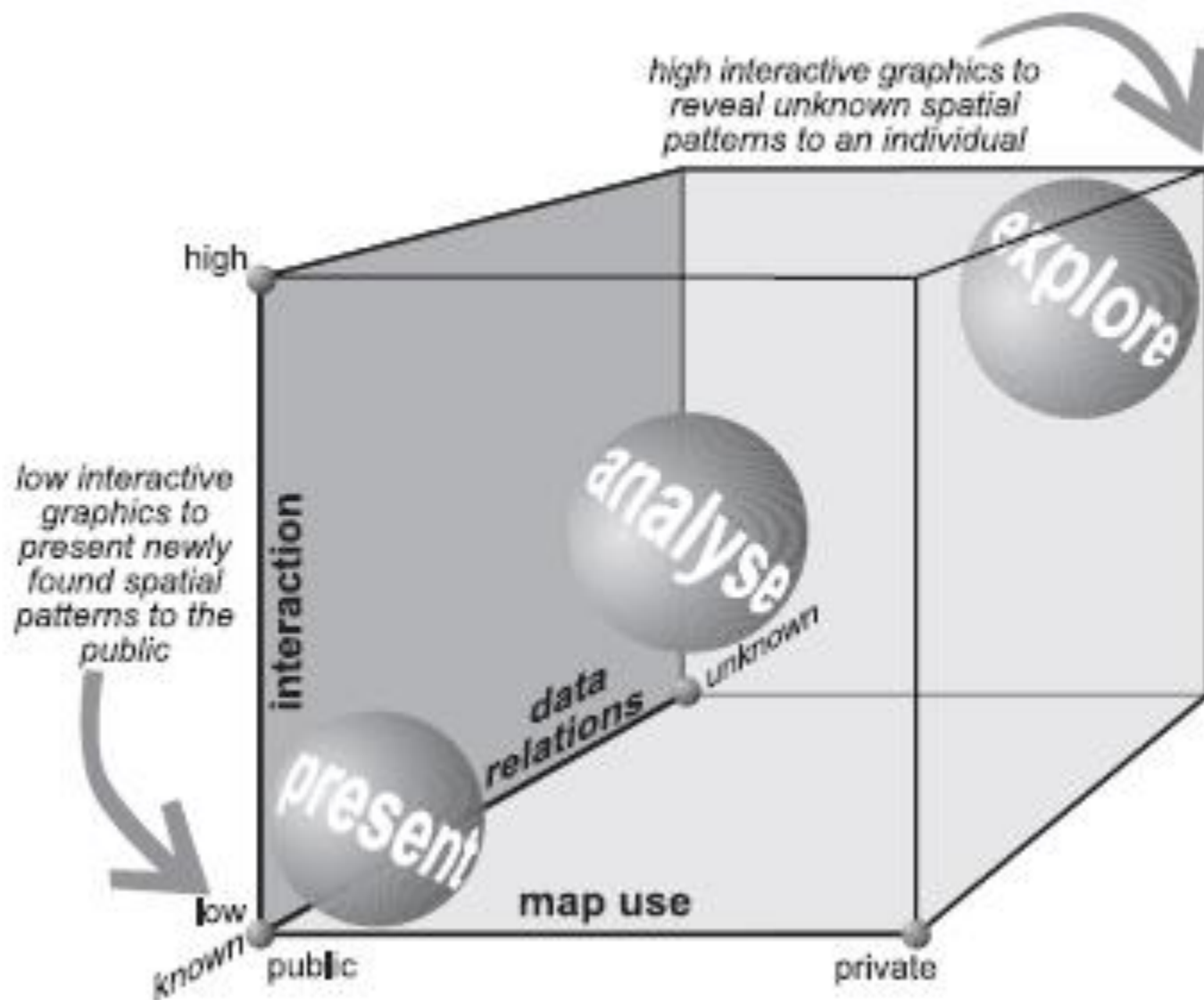
# Kraak a Ormeling (1996) širší pojetí v prostředí GIS

Širší interpretace termínu „vizualizace“:

- **explore** pro neznámá a často „surová“ nezpracovaná data;
- **analyse** (manipulate) – pro známá data;
- **present** (communicate) znalosti a informace.

Odpovídající změny umístění uvnitř MacEachrenovy krychle.

# Kraak a Ormeling (1996) širší pojetí v prostředí GIS



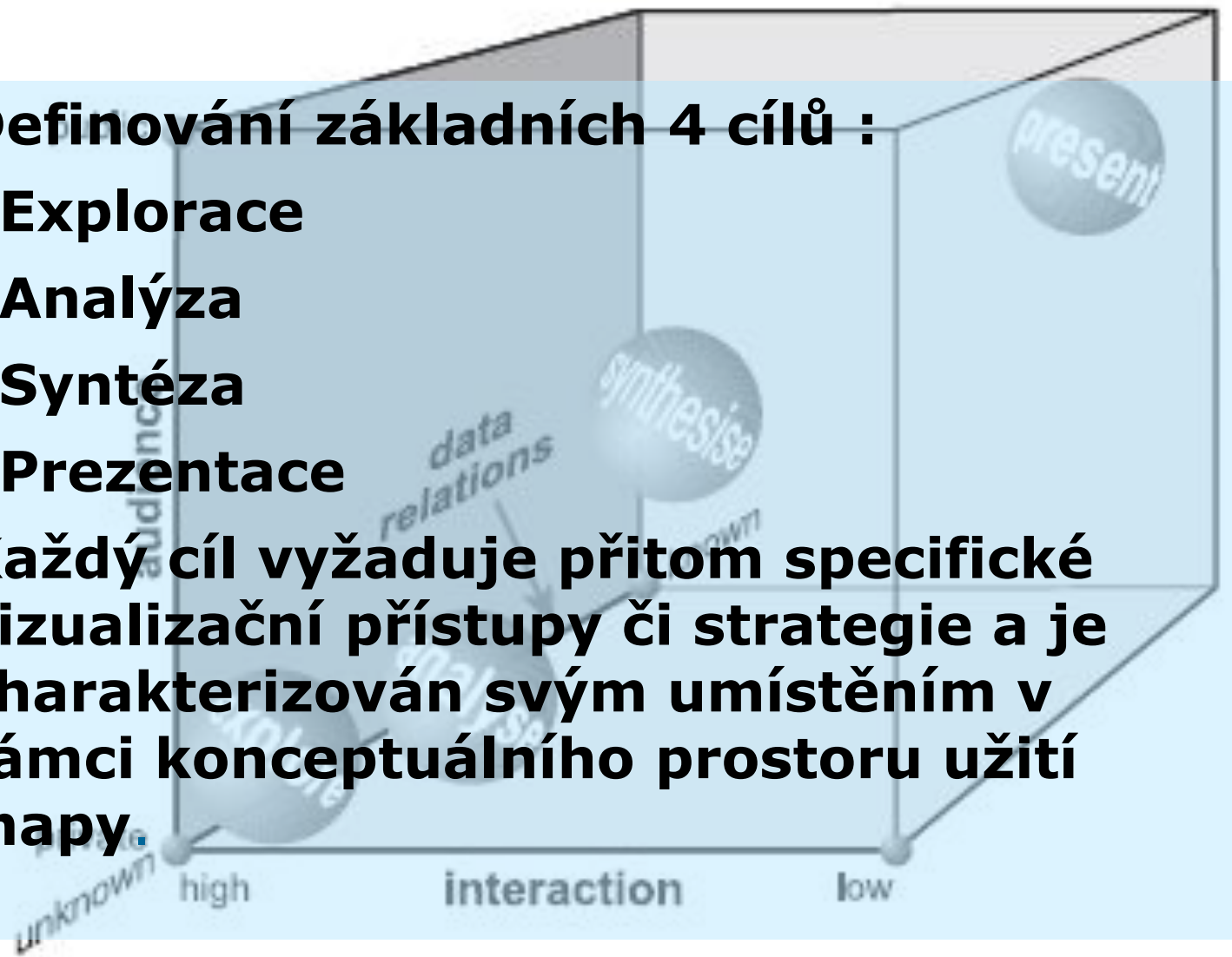


# Kartografická explorace – nová agenda (Kraak a MacEachren 1997)

**Definování základních 4 cílů :**

- **Explorace**
- **Analýza**
- **Syntéza**
- **Prezentace**

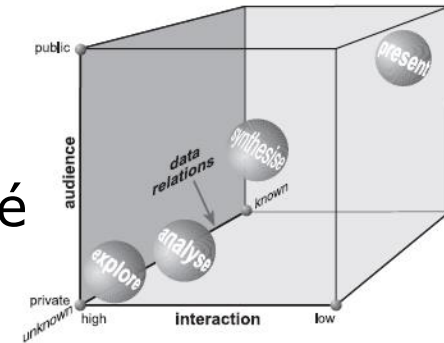
**Každý cíl vyžaduje přitom specifické vizualizační přístupy či strategie a je charakterizován svým umístěním v rámci konceptuálního prostoru užití mapy.**





# Explorace

- **Pozice – private - high interaction – revealing of unknown.**
- **Otázky :**
  - Jaká je podstata datové sady?
  - Které z vybraných datových sad mají podobné prostorové vzory?
  - Co když...?
- **Nástroje umožňující uživateli :**
  - Zkoumat prostorová data vizuálně – animace,
  - Identifikovat vztahy mezi proměnnými – propojená okna, „brushing“
  - Pohlížet na data s více pohledů – jak **prostorových** , tak **konceptuálních**.
- **Dynamické mapovací metody. Vzhled ovlivnitelný uživatelem.**







# Kartografická explorace

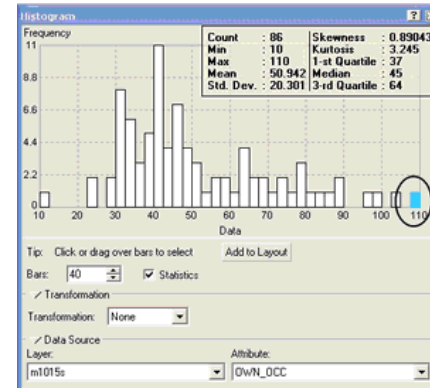
- Je účinná a funkční **v prvotní fázi** výzkumného procesu.
- Datové sady jsou prohlíženy a zkoumány při **měnících se** kartografických **podmínkách**.
- Cílem je rozpoznat **prostorové vzory** (patterns) a **trendy**, které jsou přítomny a případně ohodnotit jejich platnost.
- Snaží se podpořit vznik **nových myšlenek** a nápadů, hlavním cílem není prezentování závěrů.
- Poskytuje **dynamické zobrazení – linking and brushing**.
- Podporuje **experimentování** s různými kombinacemi dat a grafických symbolů.



# **Příklady exploračních kartografických funkcí a metod**

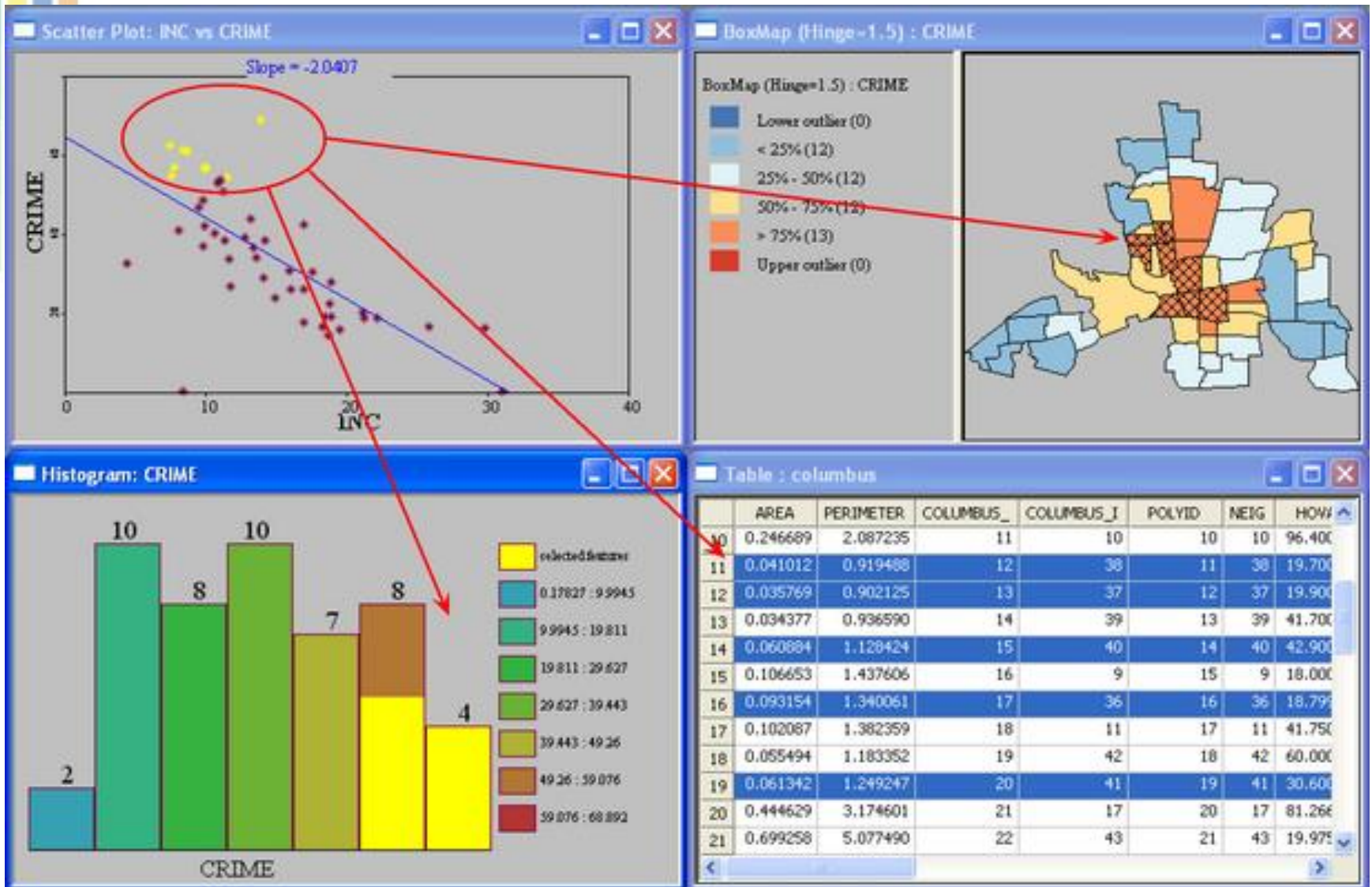
# Základy explorační analýzy dat

- Základní forma explorační datové analýzy zahrnuje výpočet základních statistických ukazatelů jednotlivých datových atributů.
- Grafické znázornění/vizualizace výsledků využívá především formu:
  - Histogram
  - Pie charts,
  - Box plots
  - Parallel coordinate plot
- Neposkytují explicitně prostorový pohled na dat, avšak lze je propojit s mapovým výstupem/vizualizací a vytvořit nový nástroj pro vstupní exploraci.
- Brushing and linking





# Linking a brushing

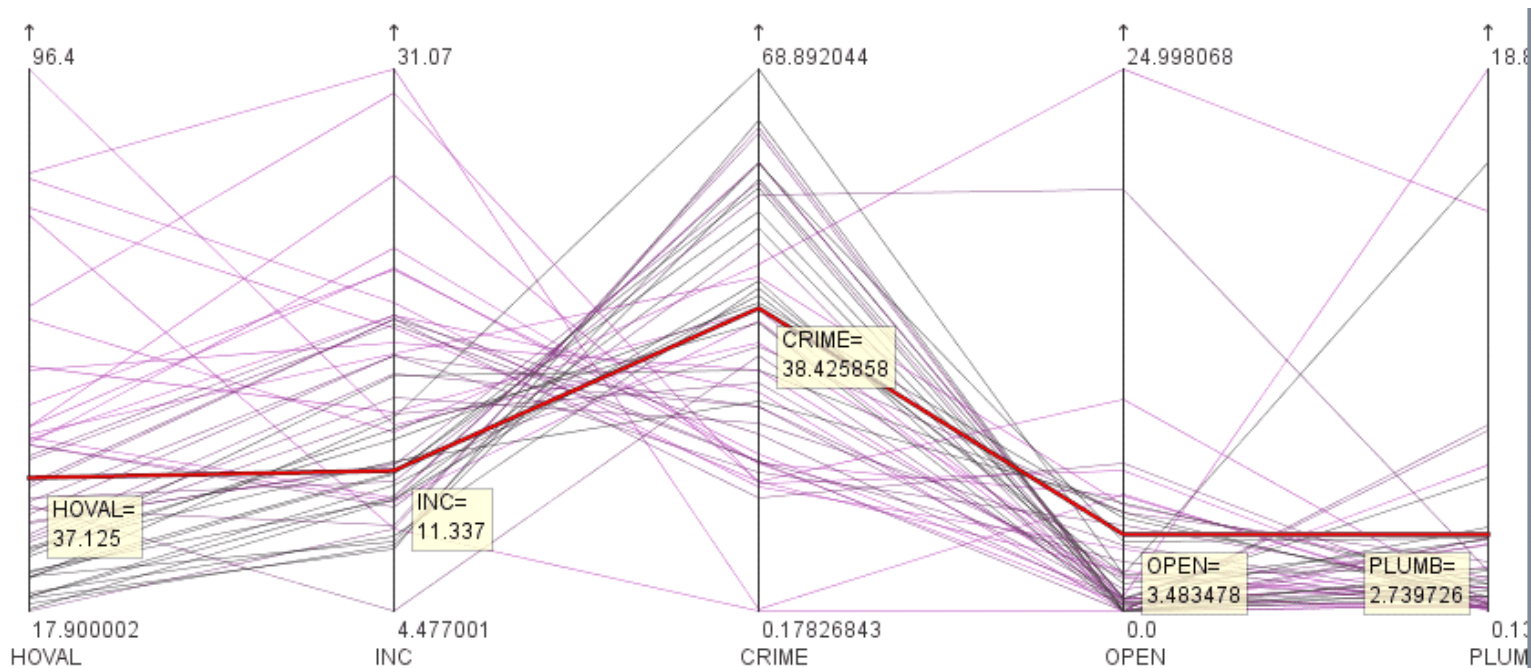




# Parallel coordinate plot (PCP)

## Rovnoběžné souřadnice

- Zaměřen na exploraci vícerozměrných datových sad.
- Každá proměnná má vlastní osu s [min, max] vertikálním rozsahem a spojitou linií vyjadřující průběh změn proměnných v jednotlivých prostorových jednotkách.
- Lze vybrat jednotlivé „podpisy“ jednotek podle volby uživatele a zvýraznit jejich hodnoty a případně umístění v mapě.





# Star plot

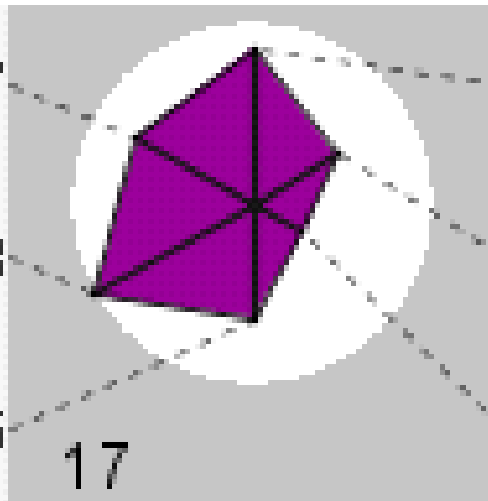
## Paprskový graf/hvězdicový či pavučinový graf

- Každá proměnná je vynesena na jednu osu,
- Počet os je roven počtu proměnných, délka osy je proporčně upravena podle rozsahu proměnných, případně přizpůsobena dle pravidel normalizace.
- Každá mapovaná oblast má individuální tvar grafu a lze je použít ve formě kartodiagramu.

PLUMB = 13.849287

OPEN = 24.998068

CRIME = 43.962486



NEIG = 42.0

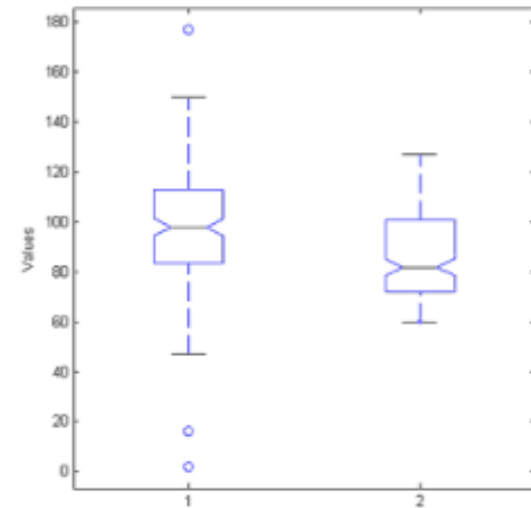
HOVAL = 60.0

INC = 13.185

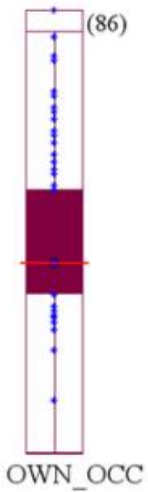
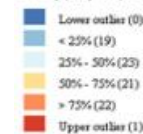
17

# Box plot Krabicový graf/diagram

- Představuje jeden ze základních grafů – jak je konstruován?
- Horní a spodní hranice „krabice/box“ – 25 a 75 percentil vzorku. Vzdálenost mezi nimi - *inter-quartile range (IQR)*
- Linie uvnitř boxu - medián vzorku, jeho umístění indikuje zešikmení (skewness) souboru.
- Vodorovné čárky (*whiskers*) zobrazují rozsah zbytku souboru za předpokladu, že nejsou přítomny odlehlé hodnoty (*outliers*).
- Odlehlá hodnota?
- Outlier více než 1.5 krát IQR od hranic krabice, možno definovat i jinak.
- Symboly mimo rozsah souboru (obvykle jinak barevné tečky či kruhy) indikují odlehlé hodnoty.
- Zdroje odlehlých hodnot?



BoxMap (Hinge=1.5) : OWN\_OCC





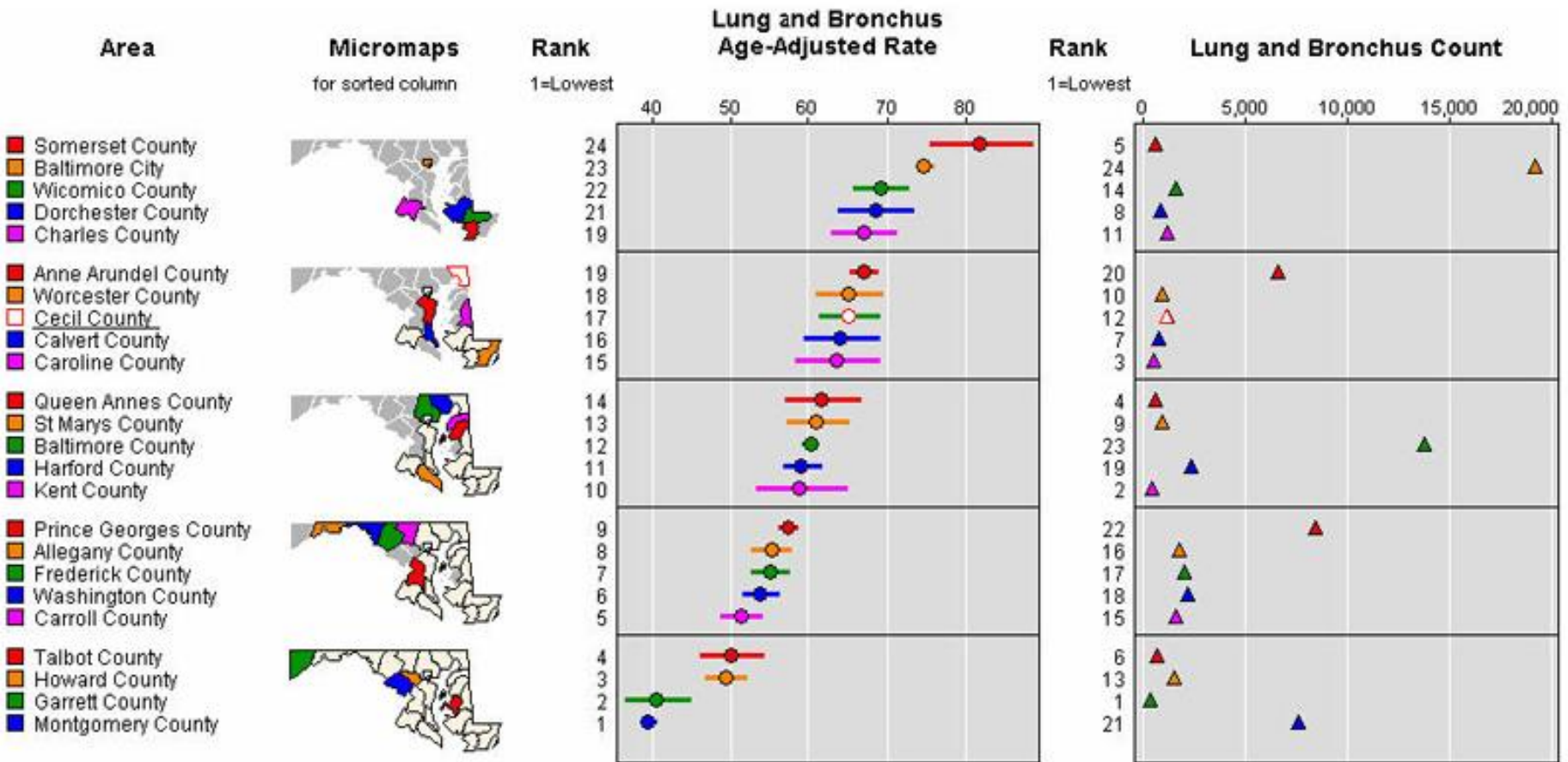
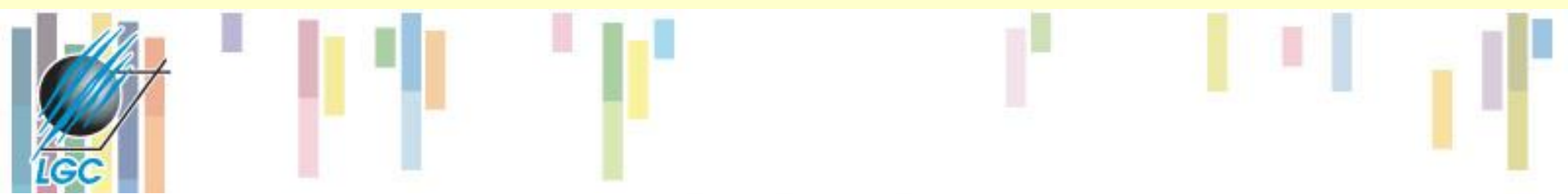
# **Příklady exploračních kartografických nástrojů**





# Linked Micromaps

- Umožní pohled na více proměnných v jednom okamžiku a srovnání atributů napříč vymezenými jednotkami v prostoru a čase.
- Podporuje 6 typů vizualizace:
  - **Sloupcové grafy;**
  - **Box plots;**
  - **Tabulky dat;**
  - **Bodové symboly;**
  - **Bodové symboly se šipkou (trendem);**
  - **Bodové symboly s intervalem spolehlivosti.**
- <http://gis.cancer.gov/tools/micromaps/>





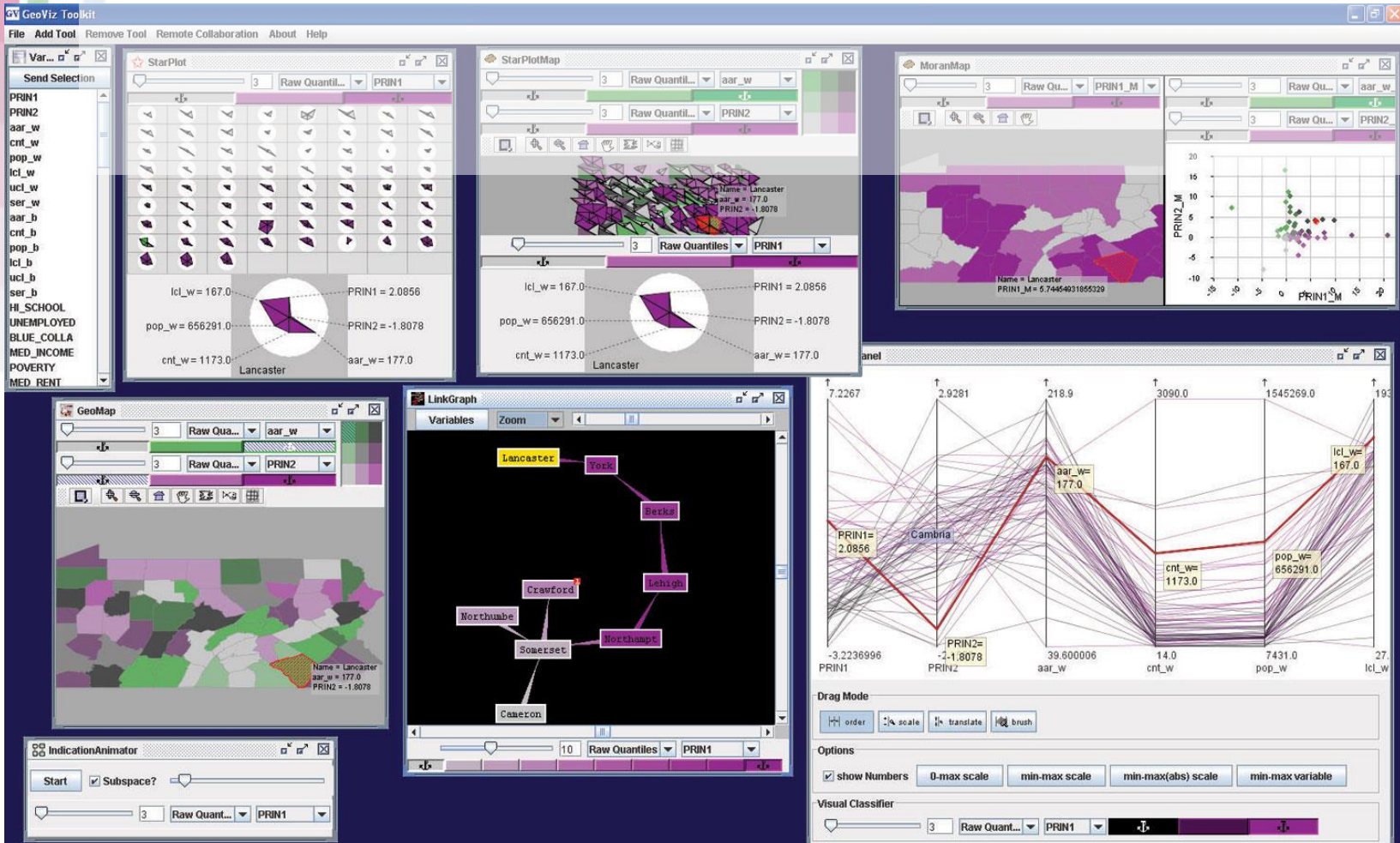
# Geoviz Toolkit

- **Frank Hardisty et al., Pen State university.**
- **Systematická analýza prostorových, časových a atributových pohledů na data.**
- **Dynamicky propojený a multivariační pohled na datové sady.**
- **Vývojové prostředí pro doplňování nástrojů.**
- **Možnost vstupu vlastních dat \*.shp**
- **Omezené vysvětlení funkcionality.**



# GeoViz Toolkit

<http://www.geovista.psu.edu/geoviztoolkit/>



- GeoDa - Dr. Luc Anselin.
- Určen pro ESDA na vektorových datech.
- Grafické rozhraní pro popisnou geografickou analýzu, ale také prostorovou autokorelaci a prostorovou regresní analýzu.
- Stále ve vývoji – původně postaveno na MapObjects.
- <http://geodacenter.asu.edu/projects/opengeoda>



# GEoDa funcionalita

- In broad terms, the functionality can be classified into six categories:
- spatial data manipulation and utilities: data input, output, and conversion,
- data transformation: variable transformations and creation of new variables,
- mapping: choropleth maps, cartogram and map animation,
- EDA: statistical graphics,
- spatial autocorrelation: global and local spatial autocorrelation statistics, with inference and visualization,
- spatial regression: diagnostics and maximum likelihood estimation of linear spatial regression models.

