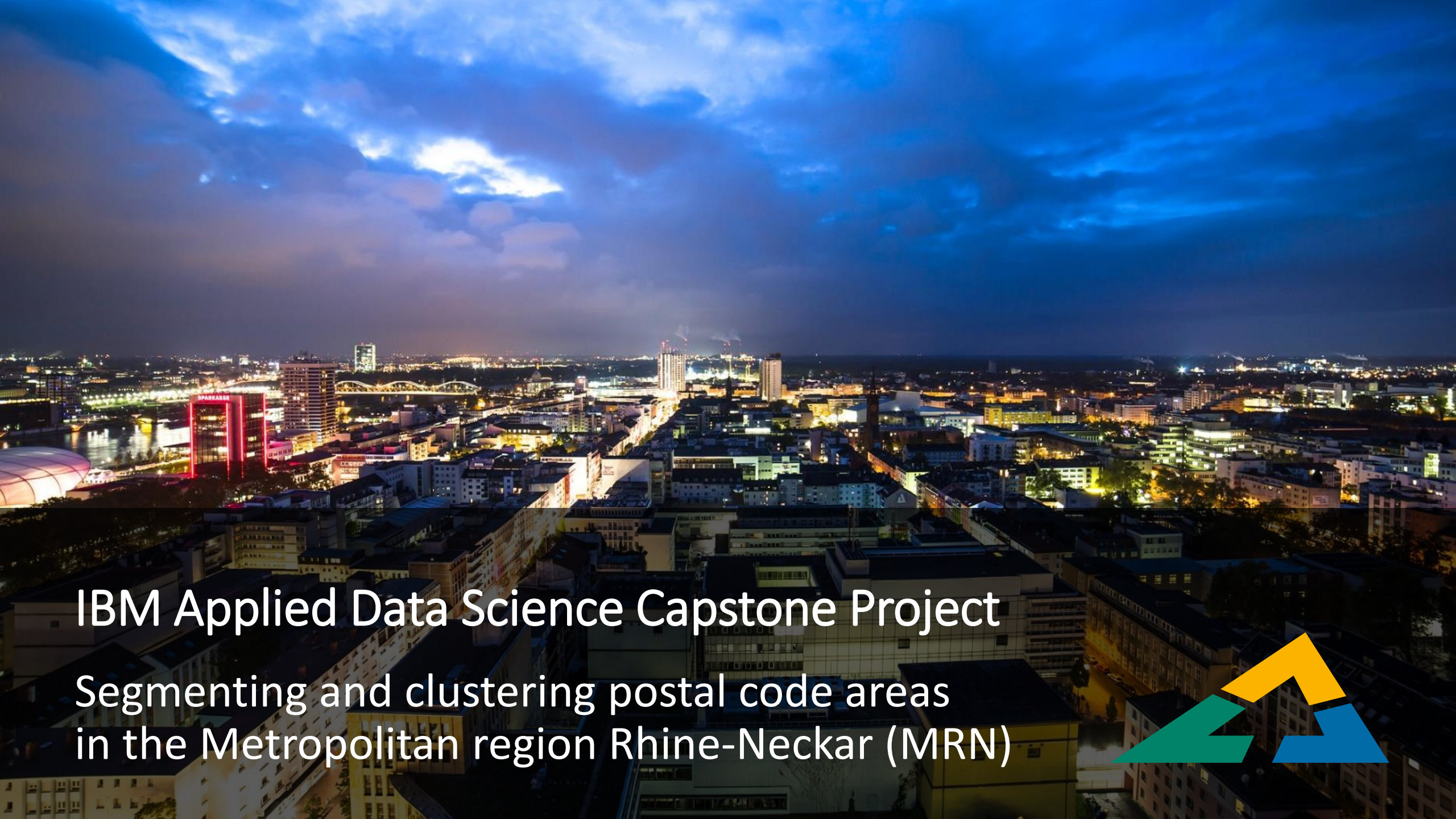
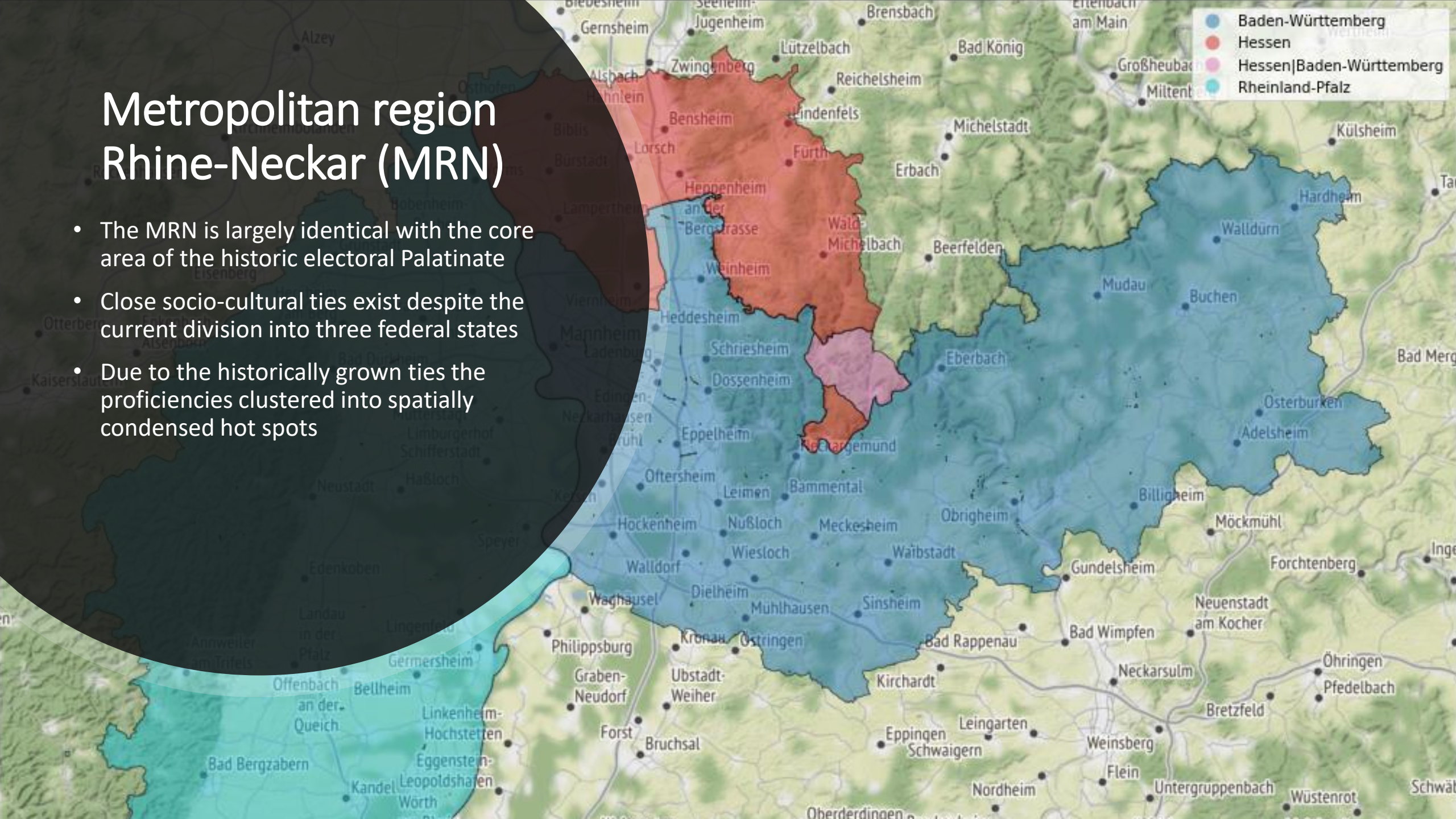


IBM Applied Data Science Capstone Project

Segmenting and clustering postal code areas in the Metropolitan region Rhine-Neckar (MRN)

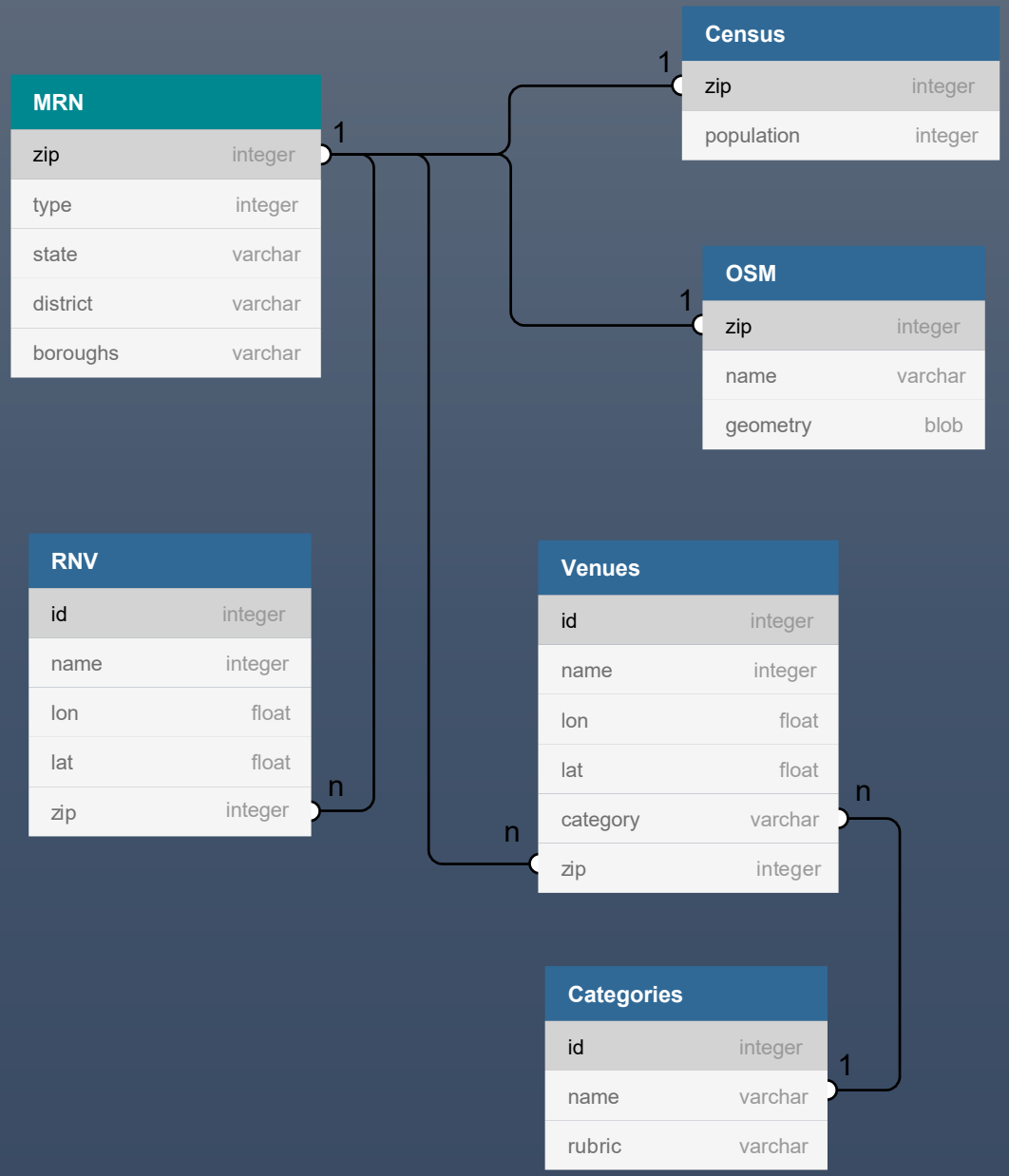
Metropolitan region Rhine-Neckar (MRN)

- The MRN is largely identical with the core area of the historic electoral Palatinate
- Close socio-cultural ties exist despite the current division into three federal states
- Due to the historically grown ties the proficiencies clustered into spatially condensed hot spots



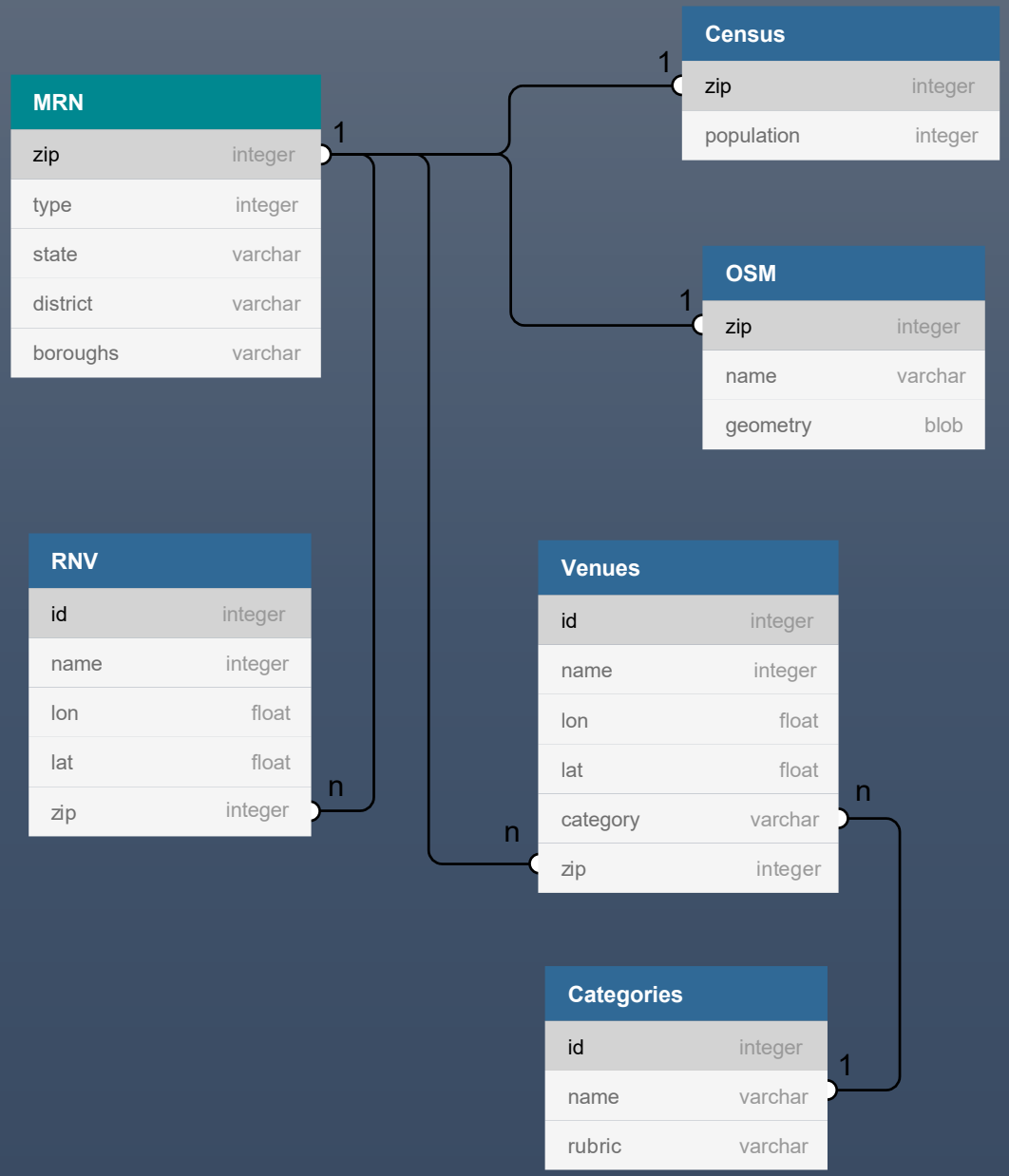
Data Integration

- Rhein Neckar Wiki
- Rhein-Neckar-Verkehr GmbH
- Federal Statistical Office of Germany
- OpenStreetMap
- Foursquare Labs, Inc.



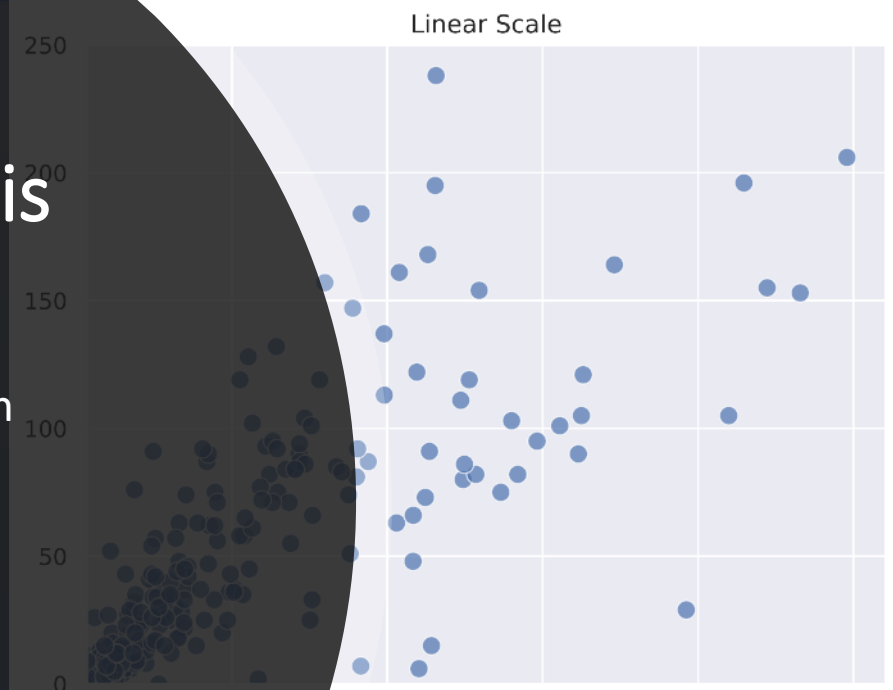
Data Transformation

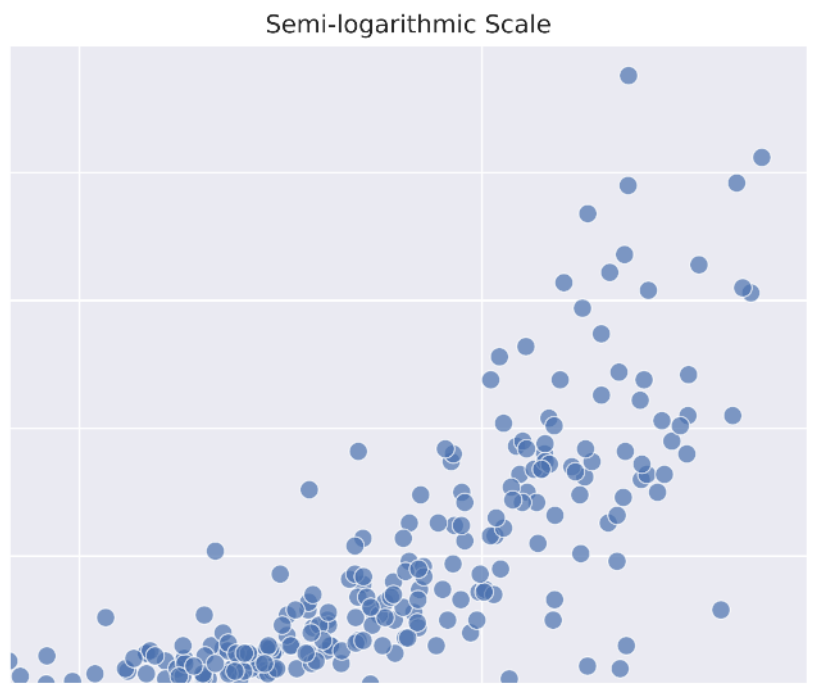
- Area estimation
- Population density estimation
- RNV stop count, area density and per capita density
- Foursquare group frequency, area density, per capita density and rubric ratio

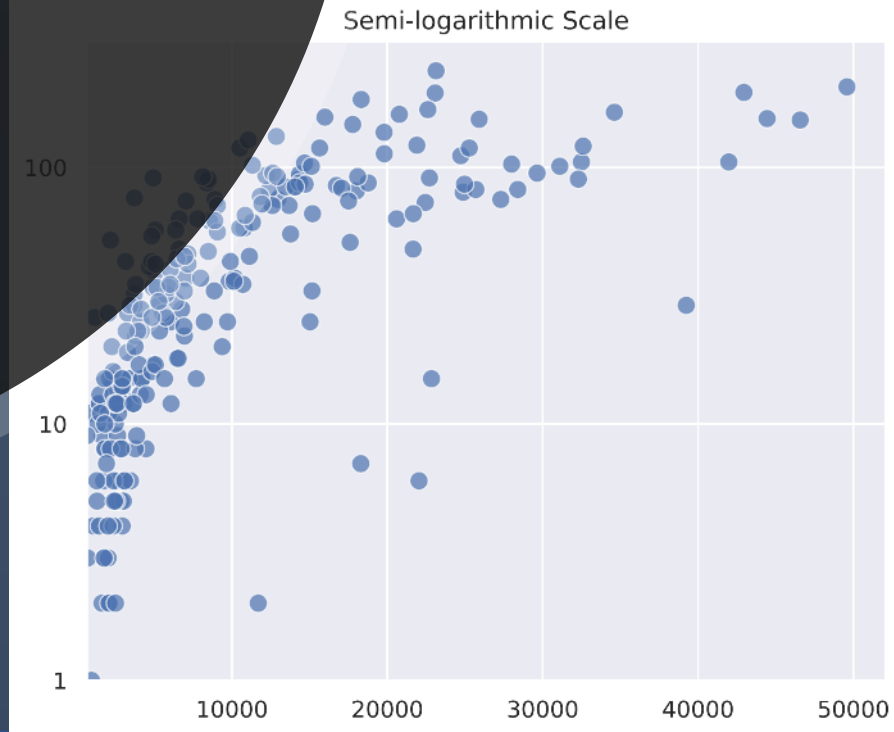


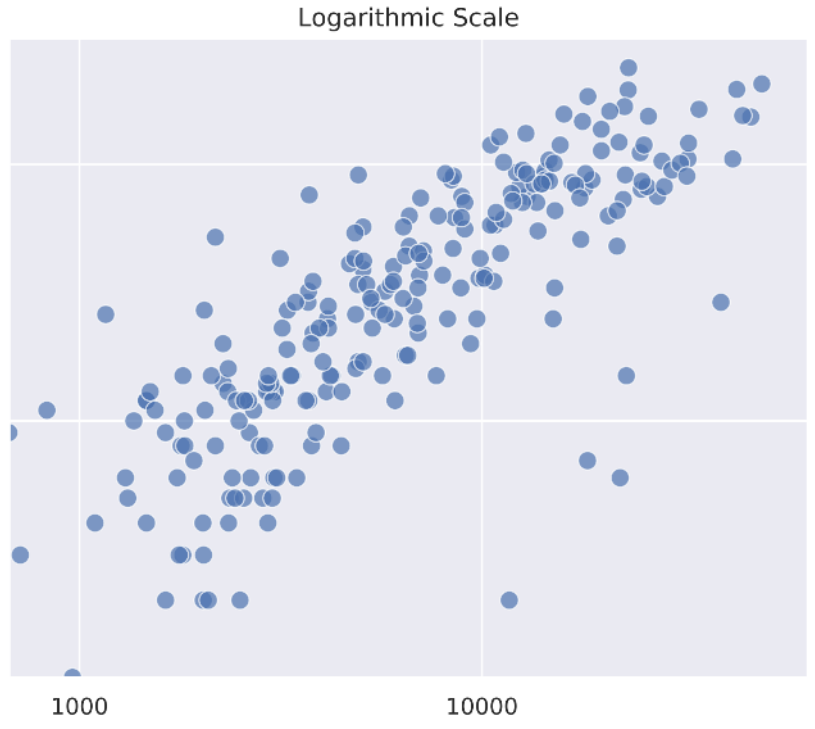
Explorative Analysis in venue supply

- visual search for patterns within Population and Venues
- Linear Scales
- Semi-logarithmic Scales
- Logarithmic Scales



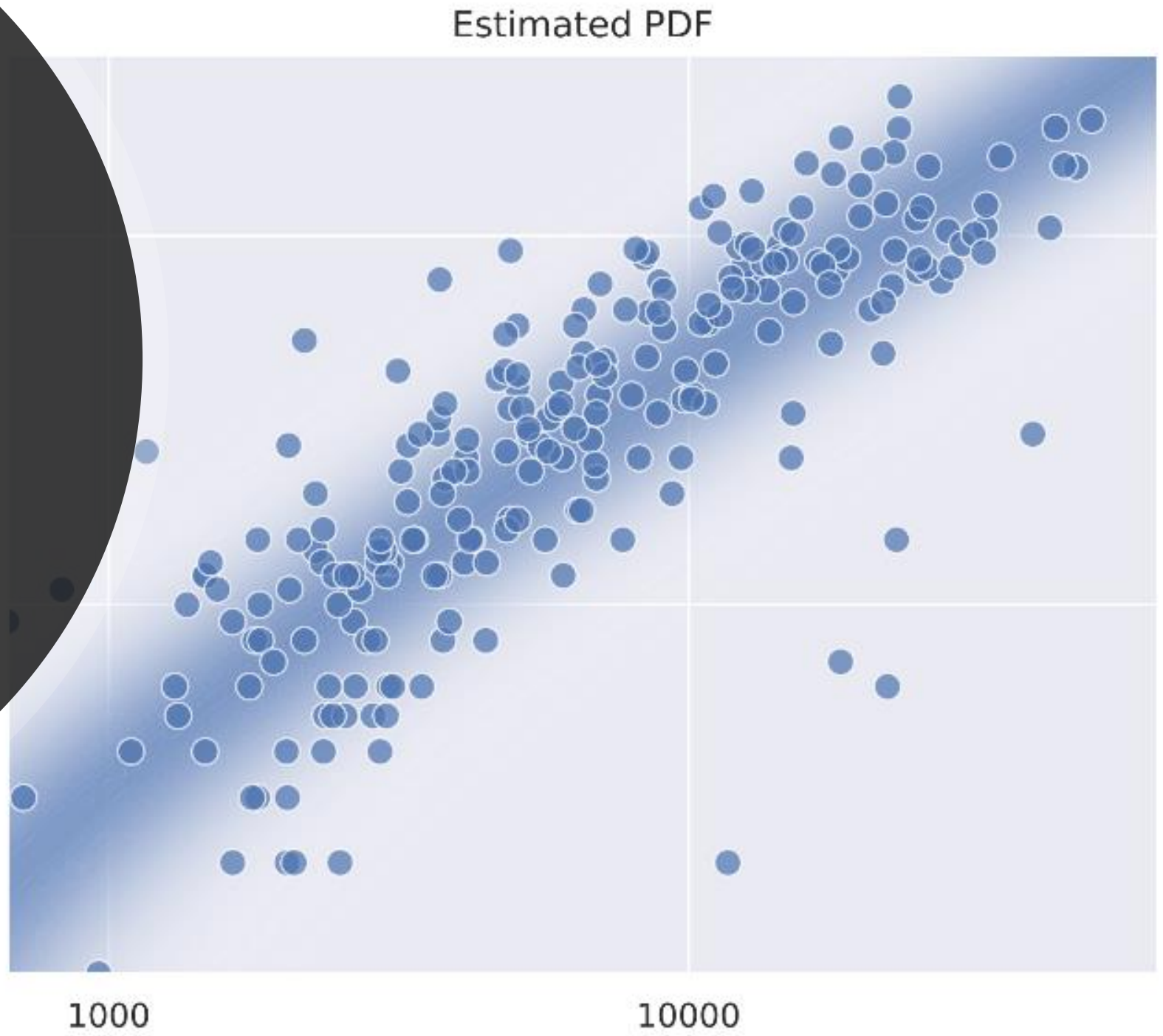






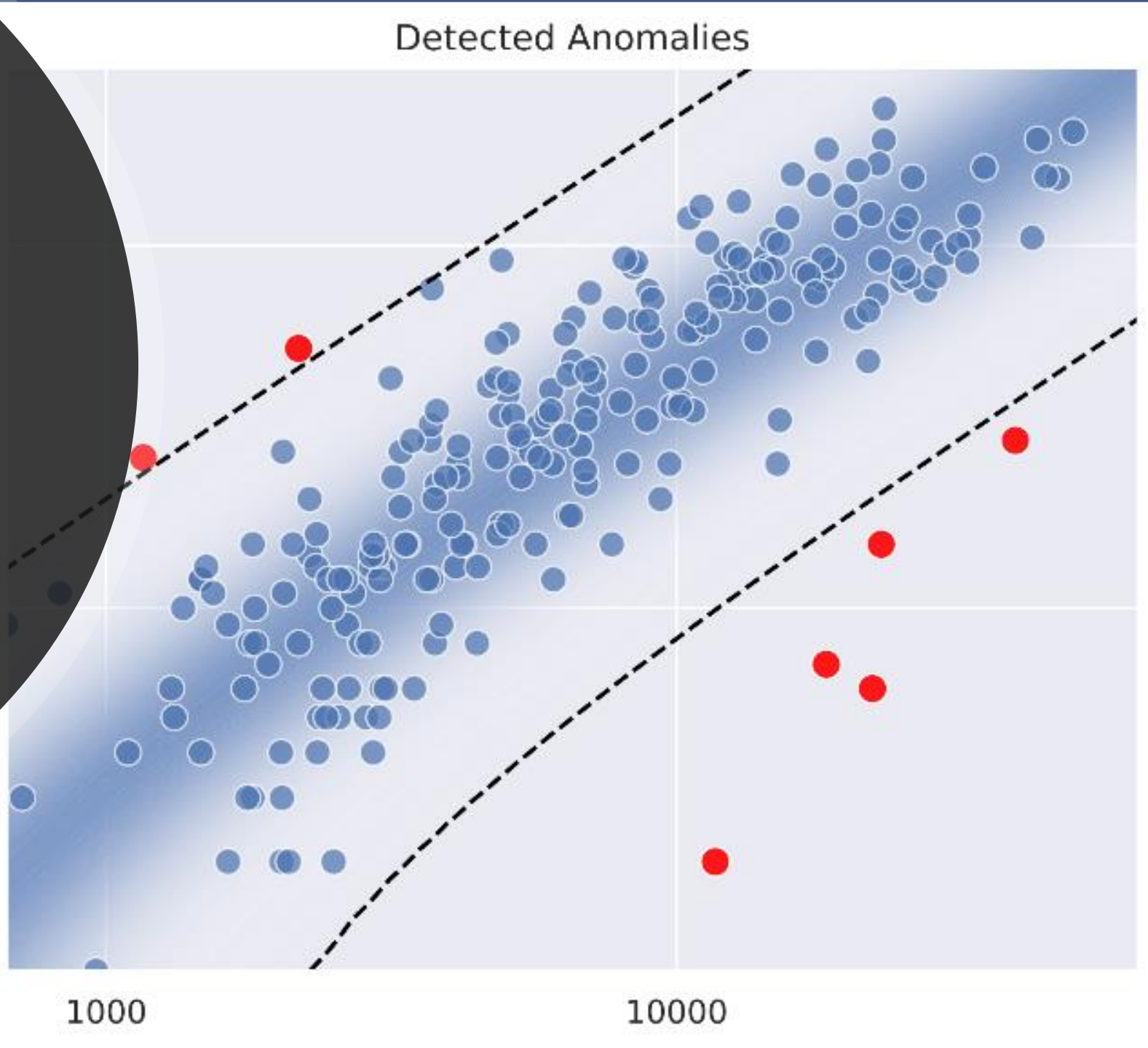
Density Estimation in venue supply

- Derivation of Principal Components in log-log scale
- Projection of (Venue, Population) to second principal component
- Modeling of normal distribution along second principal component



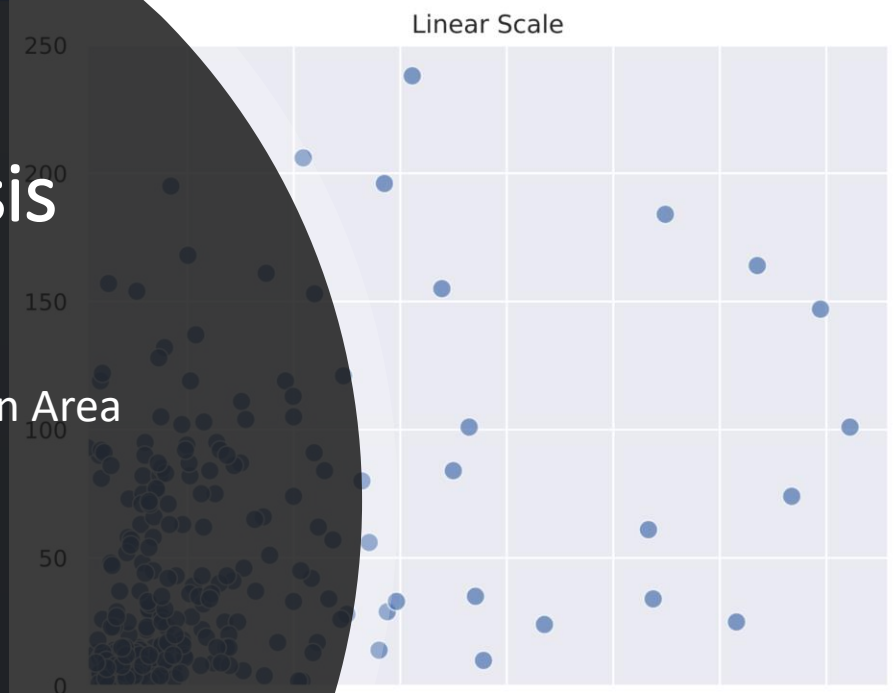
Anomaly Detection in venue supply

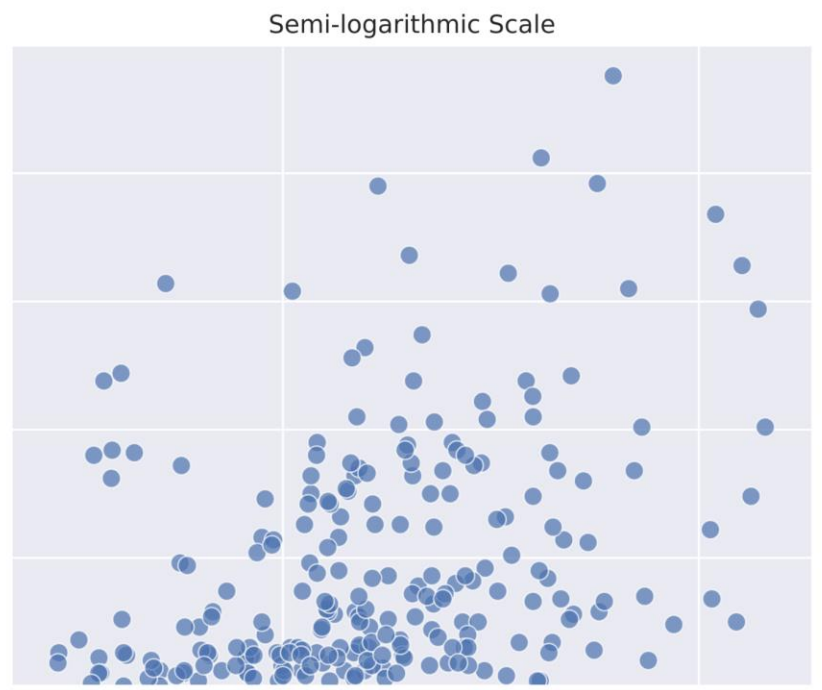
- p-value < 0.05
- Upper Group: Tourism & Pub Culture in the Palatinate wine-growing region
- Lower Group: Suburban Areas

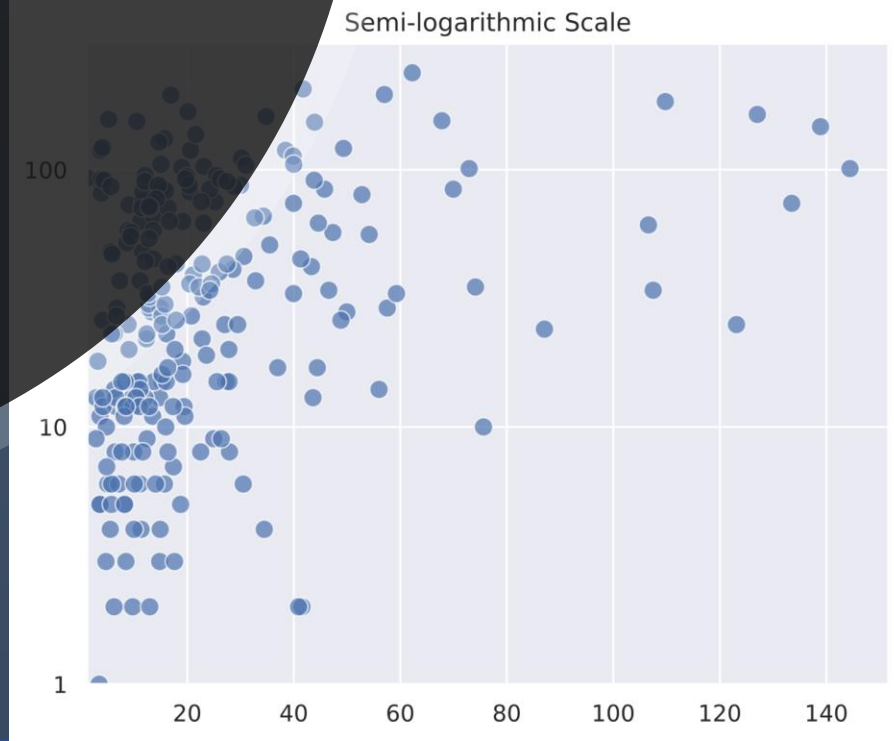


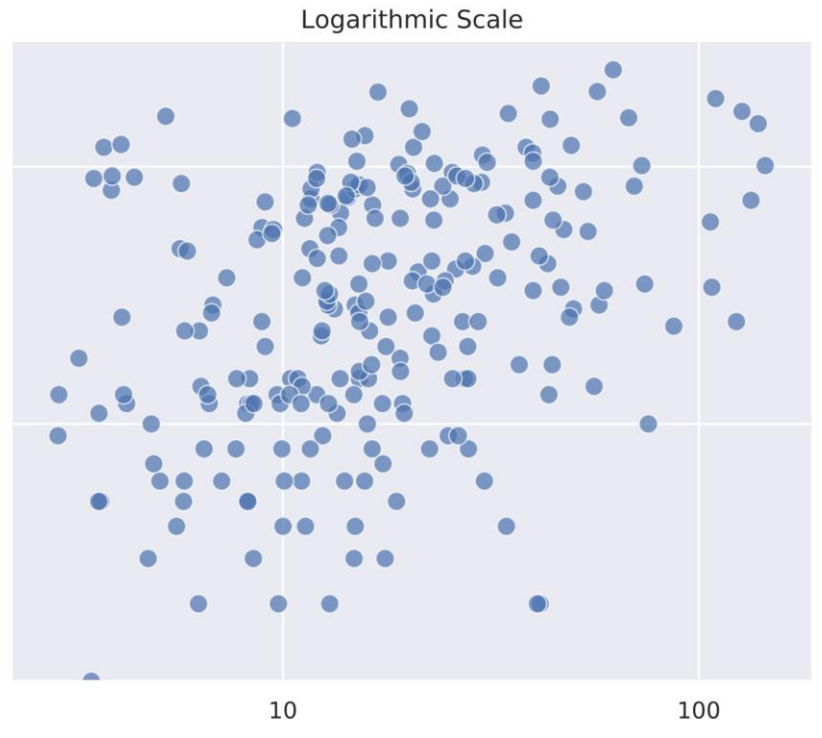
Explorative Analysis in venue density

- visual search for patterns within Area and Venues
- Linear Scales
- Semi-logarithmic Scales
- Logarithmic Scales



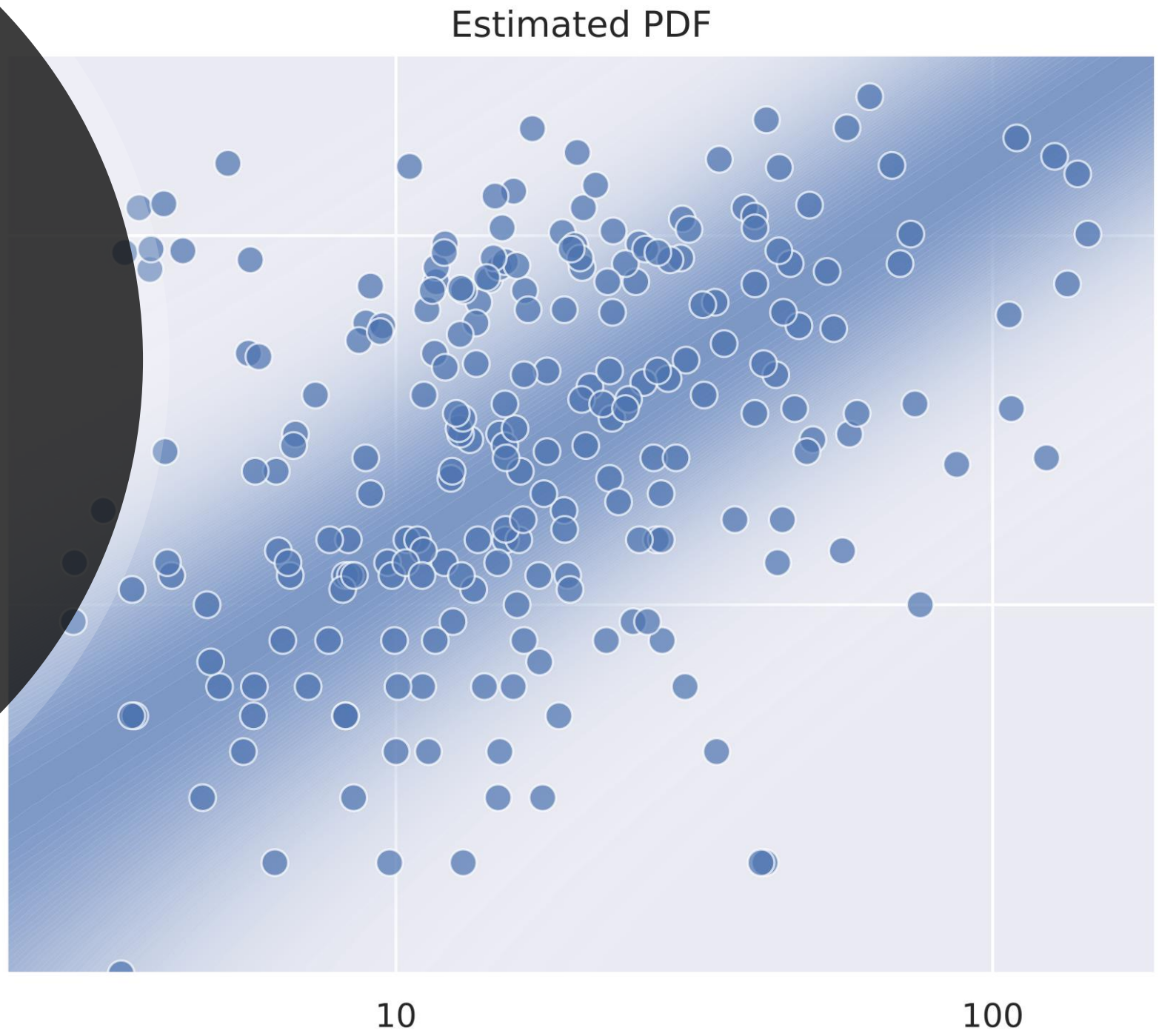






Density Estimation in venue density

- Derivation of Principal Components in log-log scale
- Projection of (Venue, Area) to second principal component
- Modeling of normal distribution along second principal component



Anomaly Detection in venue density

- p-value < 0.05
- Upper Group: Centres of largest cities in the MRN
- Lower Group: Large areas with low population density