

# Quo vadis?

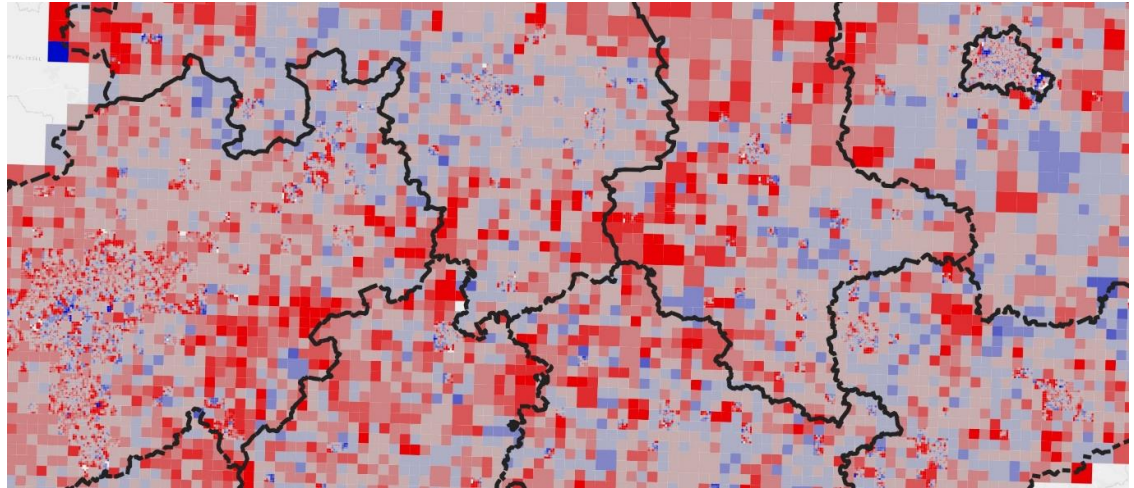
## How COVID changed the selecting of trip destinations

Project in cooperation with ARUP,  
funded by the Federal Institute for  
Research on Building, Urban Affairs  
and Spatial Development (BBSR)

Berd Academy

Data Challenge: Mobile Phone Data

12 November 2024



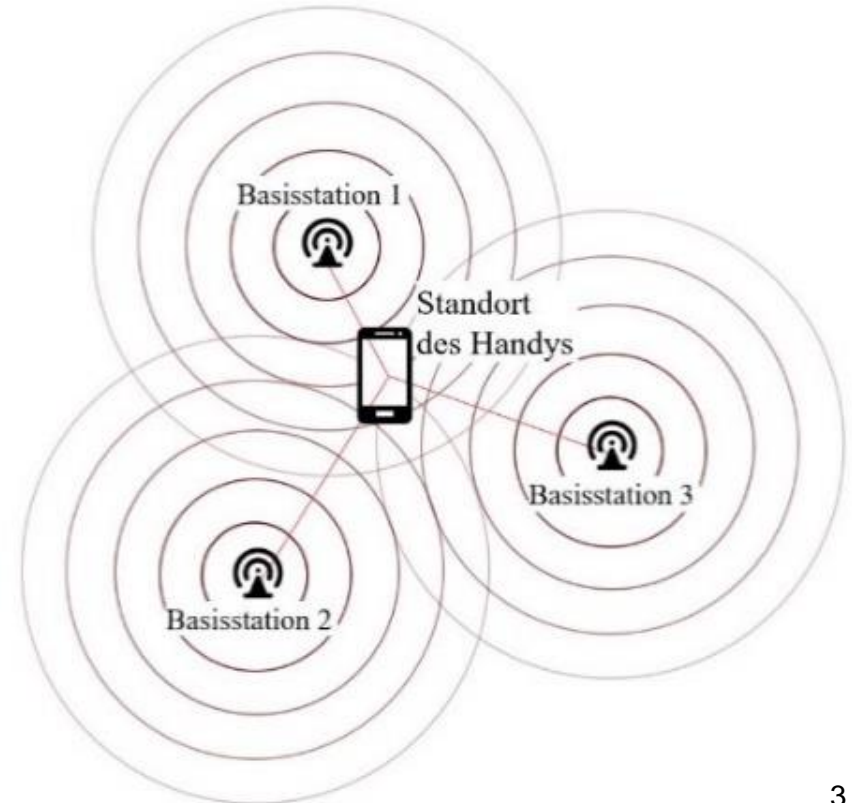
Rolf Moeckel | Professor of Travel Behavior | Department of Mobility Systems Engineering | Technical University of Munich  
Research partners: Nadine Blätgen (BBSR), Ali Saad (ARUP), Yangqian Cai (TUM)

# Problem statement

- It is well understood that the COVID pandemic had a **long-term impact on travel behavior**
- While the modal shift from transit to personal vehicles is well-established, it remains less clear how **preferences for destinations** might have changed
- Traditional household travel surveys (such as “Mobilität in Deutschland”, or Mobility in Germany) **are conducted too infrequently** to observe behavioral shifts in a timely manner

# New data potential

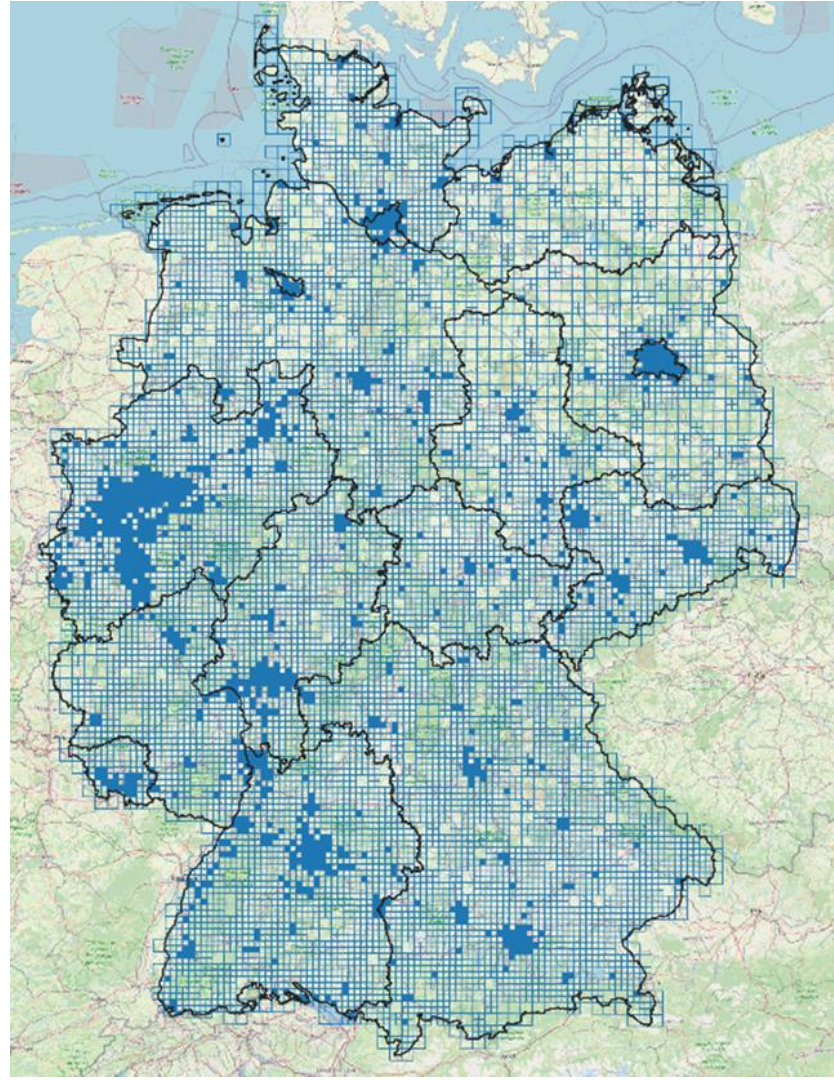
- Mobile phone data are collected continuously
- The sample size can be much larger than in traditional surveys
- Travel behavior changes can be observed almost instantaneously



# Mobile phone data

# Mobile phone data for Germany

- Teralytics provided mobile phone data for all of Germany
- Trip ends are provided by raster cells, their resolution ranges from 500x500 m in urban regions to 10x10 km in rural areas.
- Hourly temporal resolution from 2019 to 2022



# Direction of travel and distance

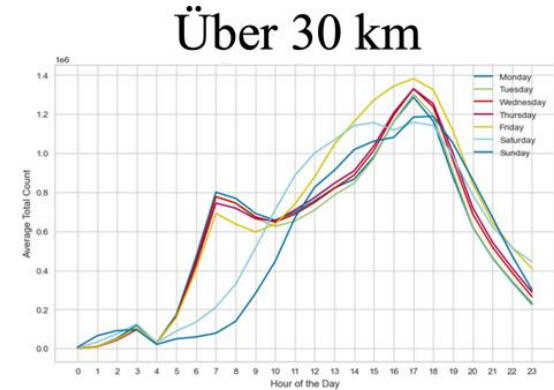
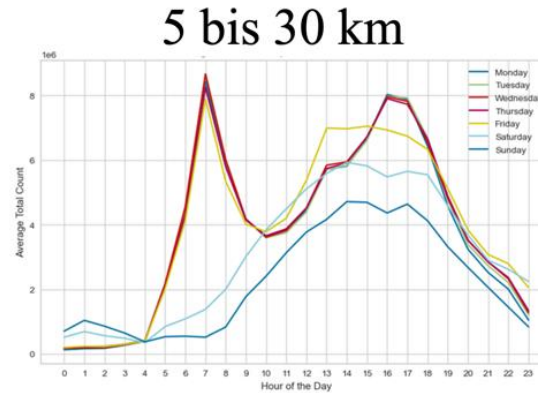
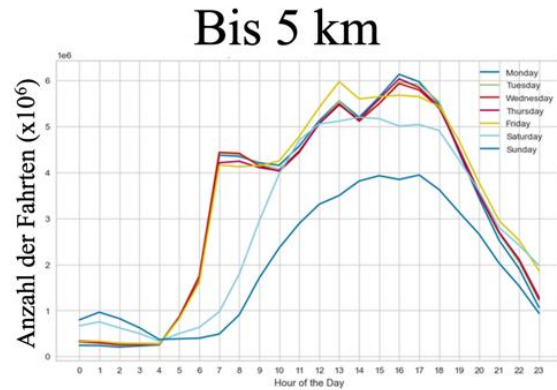
Data are provided for each raster cell by direction of travel and travel distance class:

|                   | <b>To a raster cell</b> | <b>From a raster cell</b> | <b>Roundtrip inside one raster cell</b> |
|-------------------|-------------------------|---------------------------|---|
| <b>0-5 km</b>     | 21.7%                   | 21.7%                     | 7.0%                                    |
| <b>5-30 km</b>    | 20.2%                   | 20.2%                     | 0.6%                                    |
| <b>&gt; 30 km</b> | 4.3%                    | 4.3%                      | 0.2%                                    |

*Share of trips by direction and distance class*

# Temporal resolution

- Data were provided in one-hour increments (0-1, 1-2, 2-3, ... 23-24)
- Data were provided for (almost) every day from Jan 1<sup>st</sup> 2019 to Dec 31<sup>th</sup> 2022

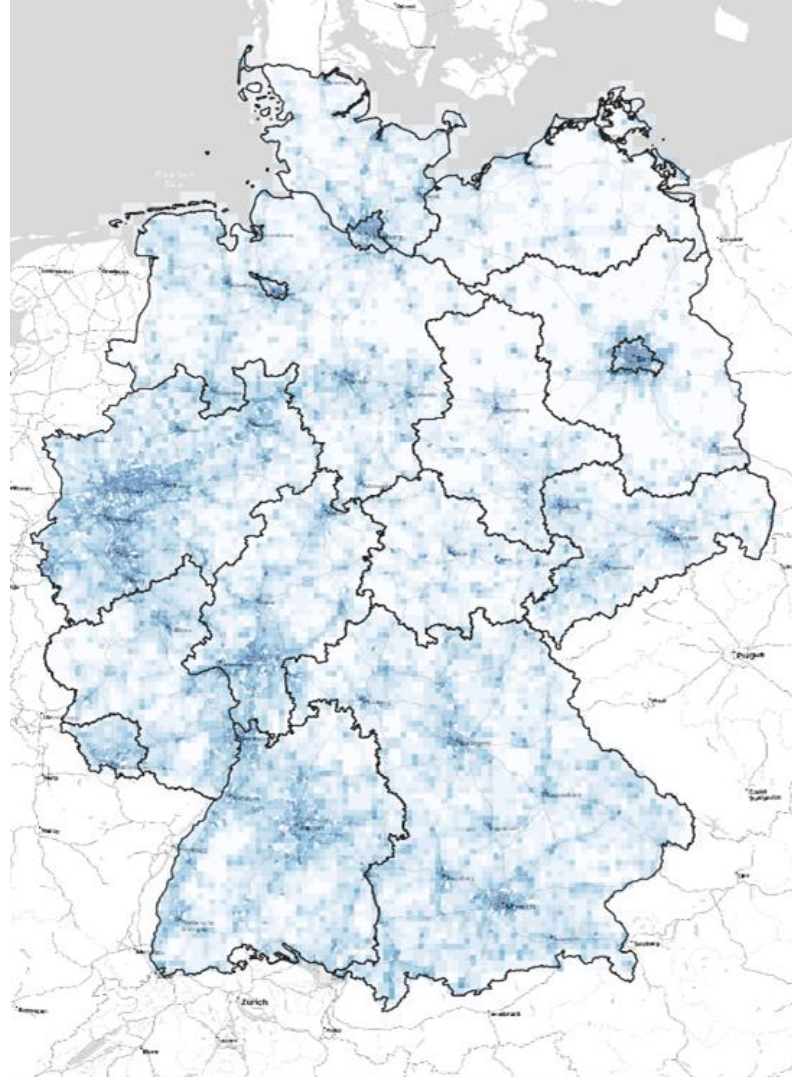


# Descriptive analysis



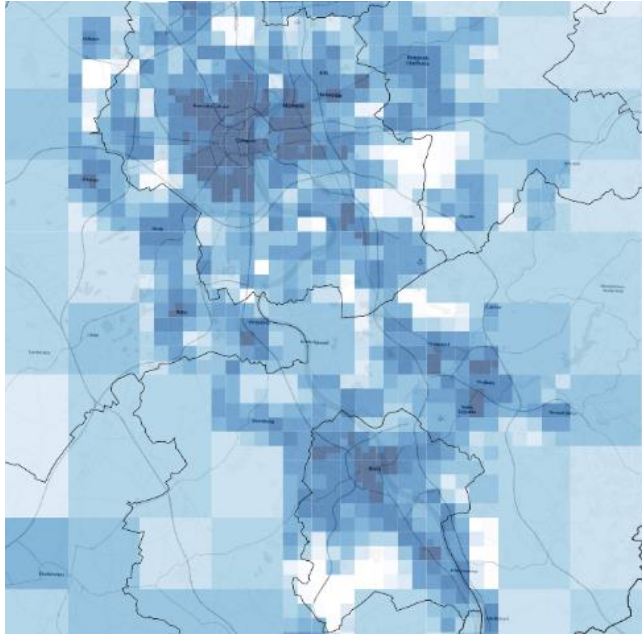
# Trip destinations

Jan 1<sup>st</sup>, 2019, all day

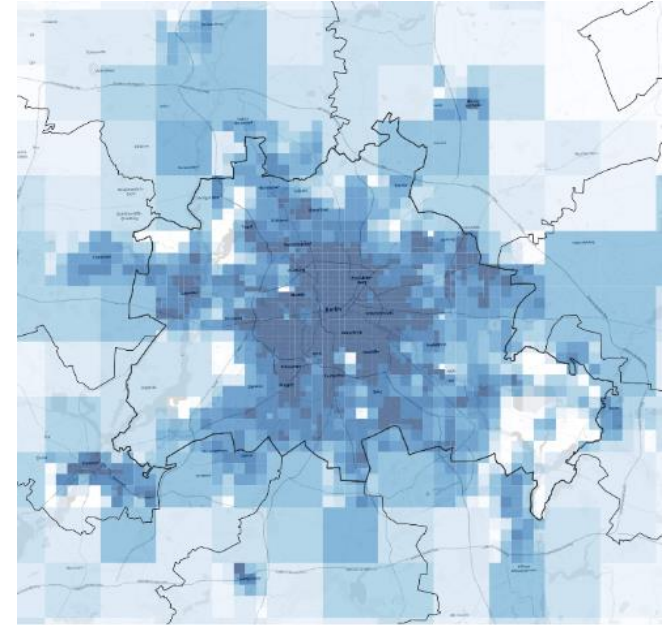


# Trip destinations

Cologne/Bonn



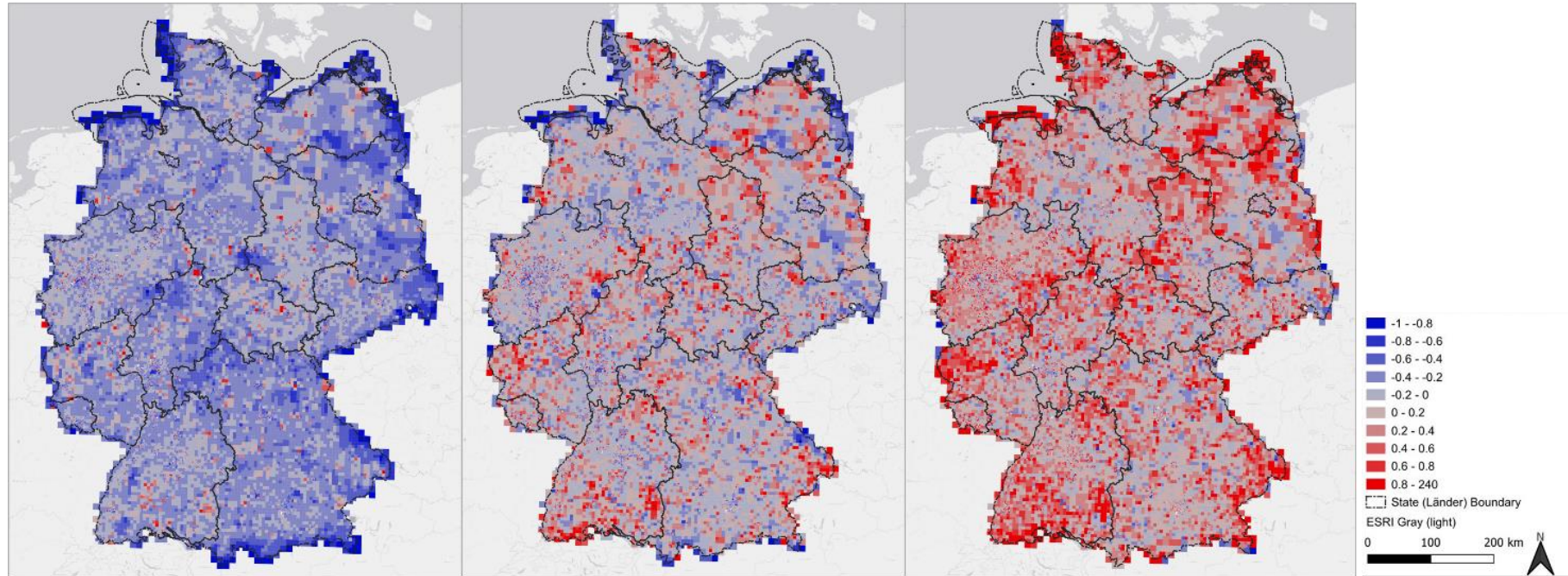
Berlin



# Seasonal change of number of trips



# Change in number of trips from year to year



April 2019 → April 2020

April 2020 → April 2021

April 2021 → April 2022

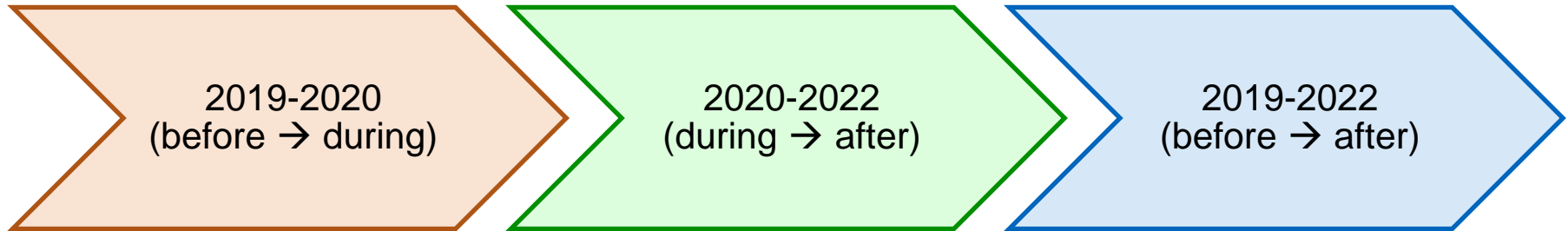
# Statistical analysis



# Definition of “before”, “during” and “after” COVID



## Time periods of interest



**Dependent variable:** Change of arriving trips per raster cell from 2019 to 2022



## Greenspace



## Population density



And a large number of  
variables from  
OpenStreetMap



# Aggregation of variables to reduce the dimensions

| Category            | Definition   | Ave. Density | 75 Percentile | Max. Density |
|---------------------|--|--------------|---------------|--------------|
| Restaurant          | 'bakery', 'cafe', 'fast_food', 'food_court', 'restaurant'  | 6.142        | 2             | 744          |
| Bars                | 'bar', 'nightclub', 'pub'  | 1.179        | 0             | 248          |
| Health              | 'chemist', 'clinic', 'dentist', 'doctors', 'hospital', 'nursing_home', 'optician', 'pharmacy', 'veterinary'  | 2.629        | 0.68          | 600          |
| Banks               | 'atm', 'bank'  | 0.888        | 0.04          | 128          |
| Beauty              | 'beauty_shop', 'hairstylist'   | 1.836        | 0.08          | 256          |
| Parks               | 'arts_centre', 'artwork', 'biopark', 'cinema', 'dog_park', 'museum', 'park', 'playground', 'theatre', ...    | 2.123        | 1             | 424          |
| Sports              | 'golf_course', 'ice_rink', 'pitch', 'stadium', 'swimming_pool', 'track'                                      | 0.560        | 0             | 112          |
| Retail              | 'beauty_shop', 'beverages', 'bicycle_shop', 'bookshop', 'butcher', 'car_dealership', 'clothes', ...          | 5.649        | 2             | 920          |
| Education           | 'college', 'kindergarten', 'library', 'school', 'university'   | 0.782        | 0.04          | 100          |
| Other               | 'laundry', 'travel_agent', 'wastewater_plant', 'water_works'   | 0.398        | 0             | 56           |
| Shared mobility     | 'bicycle_rental', 'car_rental', 'car_sharing'  | 0.478        | 0             | 80           |
| Religion            | 'wayside_shrine'   | 0.070        | 0             | 30           |
| Public services     | 'community_centre', 'courthouse', 'embassy', 'fire_station', 'police', 'post_office', 'public_building', ... | 0.522        | 0.08          | 136          |
| Hotels              | 'alpine_hut', 'camp_site', 'chalet', 'caravan_site', 'guesthouse', 'hotel', 'hostel', 'motel'                | 0.444        | 0.04          | 184          |
| Tourist attractions | 'archaeological', 'attraction', 'battlefield', 'castle', 'fort', 'monument', 'ruins', 'tower', 'windmill'    | 0.264        | 0.04          | 136          |

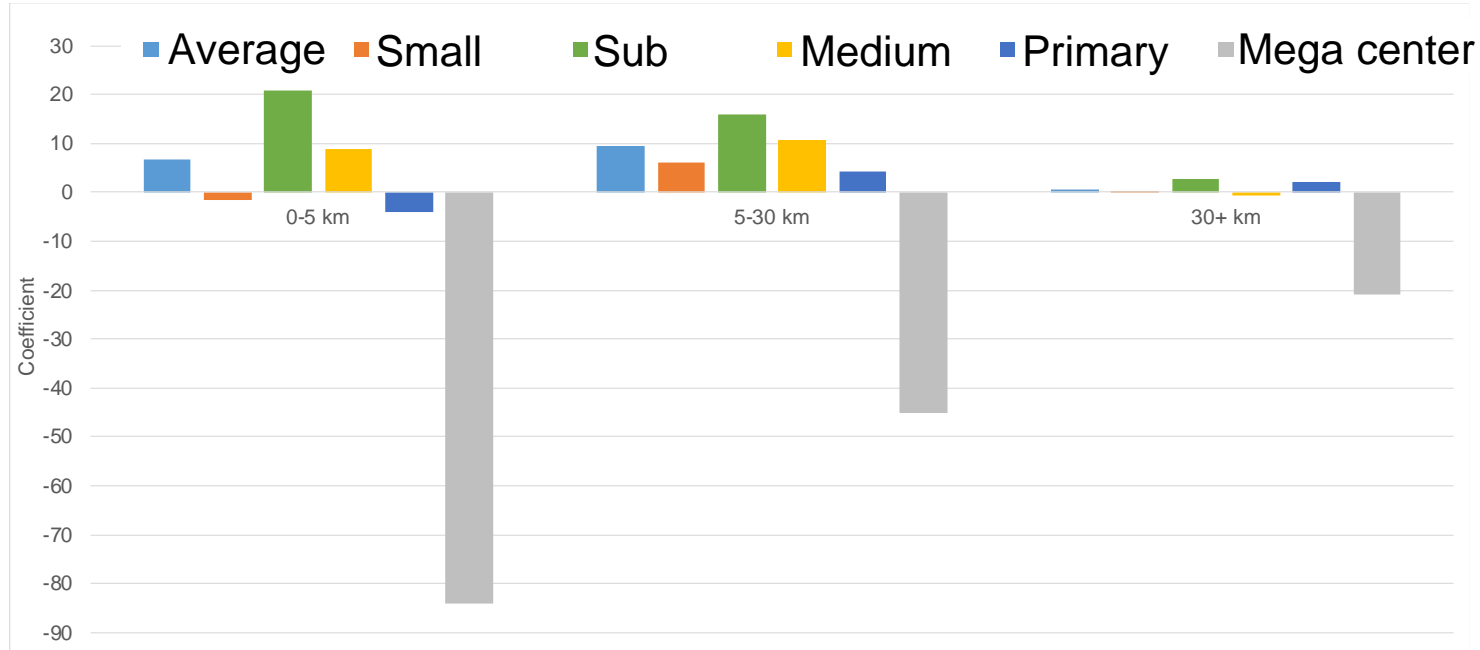
# Statistical methods applied

1. Regression analysis
2. Generalized Additive Model
3. XGBoost (Machine Learning)
4. Geographically weighted regression

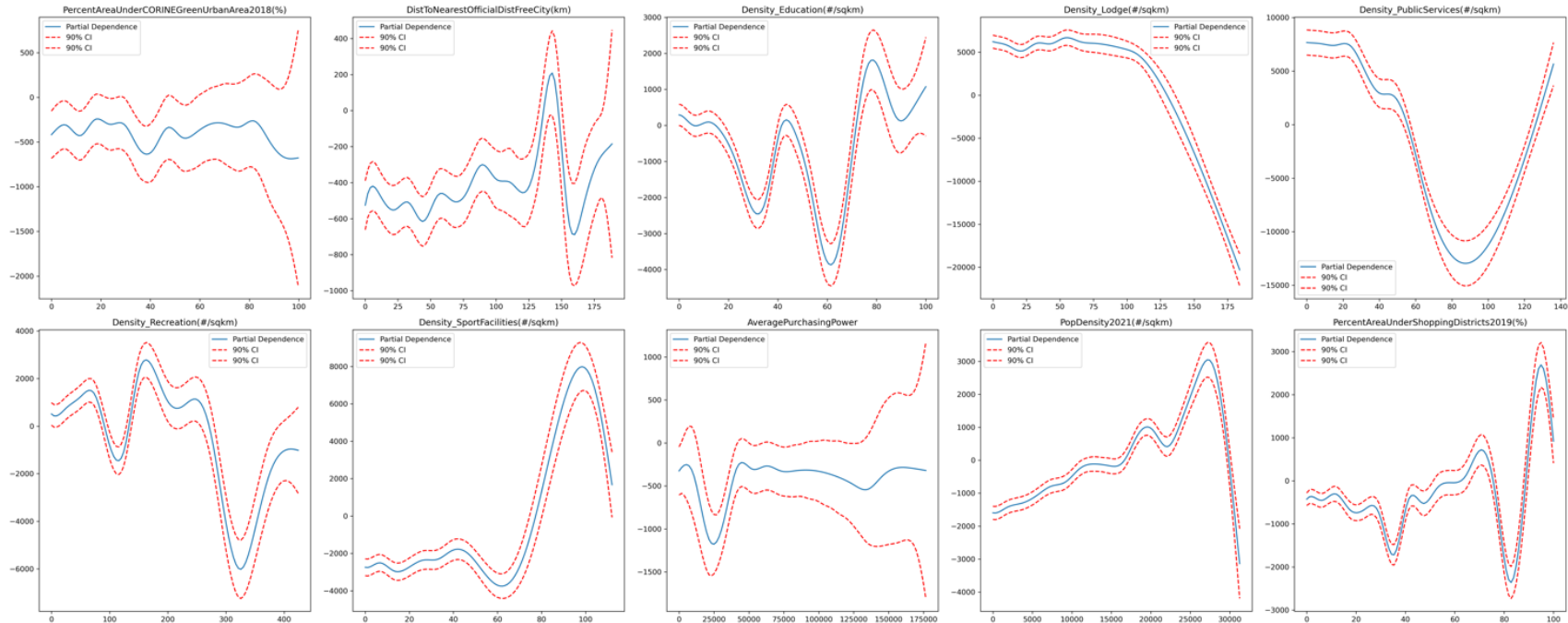
## Regression analysis of trip attraction (trips < 5 km)

| Variable                                       | Parameter | Std. Error | t Value | P(> t ) |
|--|-----------|------------|---------|---------|
| Constant                                       | -70.646   | 23.77      | -2.972  | 0.003   |
| Share green area (2018)                        | 1.3812    | 1.109      | 1.246   | 0.213   |
| Distance to next county-free city (km)         | 0.7353    | 0.24       | 3.07    | 0.002   |
| Education facilities / km <sup>2</sup>         | -29.85    | 2.398      | -12.445 | 0       |
| Hotels / km <sup>2</sup>                       | -36.8172  | 2.888      | -12.748 | 0       |
| Public services / km <sup>2</sup>              | -38.9345  | 3.825      | -10.179 | 0       |
| Parks and leisure facilities / km <sup>2</sup> | 0.9558    | 0.916      | 1.044   | 0.297   |
| Sports facilities / km <sup>2</sup>            | 14.3915   | 2.441      | 5.896   | 0       |
| Average purchasing power                       | 0.0007    | 0          | 1.614   | 0.107   |
| Population / km <sup>2</sup> (2021)            | 0.1114    | 0.003      | 40.172  | 0       |
| Share retail center area (2017)                | 6.7441    | 1.377      | 4.898   | 0       |
| Transit stops / km <sup>2</sup>                | 2.488     | 0.387      | 6.421   | 0       |
| Adjusted R <sup>2</sup> : 0.061                |           |            |         |         |

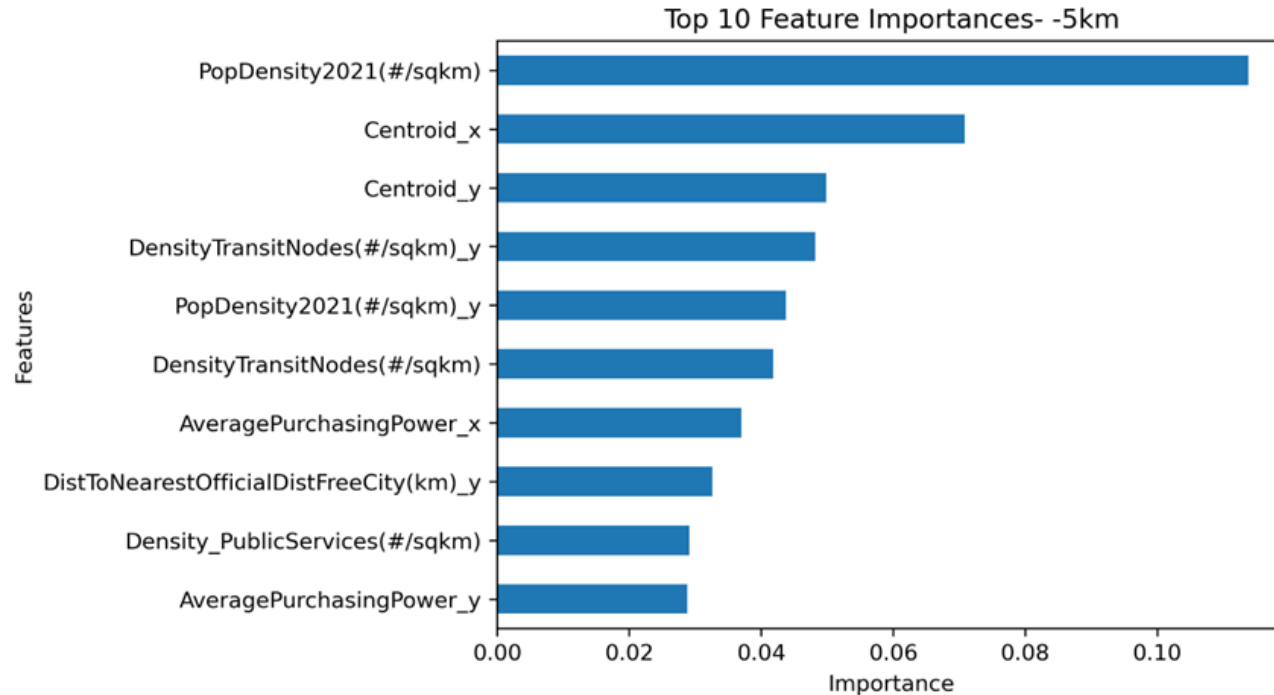
# Regression analysis: Effect of retail centers by size



# Generalized Additive Model (trips < 5 km)

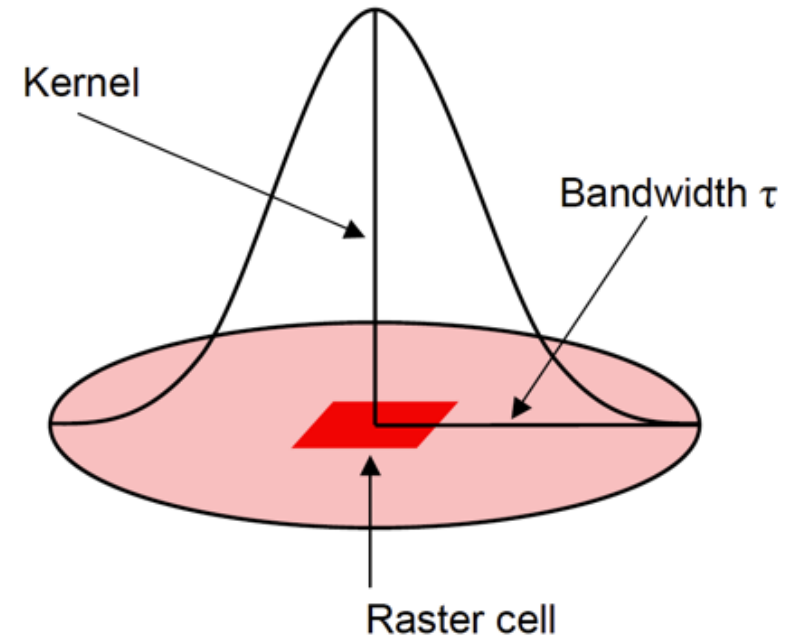


## XGBoost (trips < 5 km)



# Geographically weighted regression

Calculate coefficients based on current raster cell and neighboring raster cells.



# Geographically weighted regression estimation results

| Variable                                       | Average | Std. Dev. | Min     | Median | Max    |
|--|---------|-----------|---------|--------|--------|
| Constant                                       | -0.041  | 1.219     | -20.128 | -0.044 | 9.245  |
| Share green area (2018)                        | 0.003   | 0.244     | -5.789  | 0.005  | 2.465  |
| Distance to next county-free city (km)         | -0.01   | 1.409     | -22.809 | 0.001  | 12.012 |
| Education facilities / km <sup>2</sup>         | -0.04   | 0.649     | -6.479  | -0.035 | 6.515  |
| Hotels / km <sup>2</sup>                       | 0.008   | 0.622     | -3.608  | 0.008  | 2.924  |
| Public services / km <sup>2</sup>              | -0.005  | 0.454     | -2.275  | -0.011 | 5.6    |
| Parks and leisure facilities / km <sup>2</sup> | 0.092   | 0.64      | -3.703  | 0.034  | 5.236  |
| Sports facilities / km <sup>2</sup>            | -0.035  | 0.83      | -6.244  | -0.009 | 12.227 |
| Average purchasing power                       | -0.025  | 0.211     | -3.883  | -0.003 | 2.544  |
| Population / km <sup>2</sup> (2021)            | 0.224   | 0.549     | -2.557  | 0.177  | 3.844  |
| Share retail center area (2017)                | -0.037  | 0.577     | -3.526  | -0.041 | 3.147  |
| Transit stops / km <sup>2</sup>                | -0.055  | 0.54      | -2.926  | -0.027 | 4.228  |

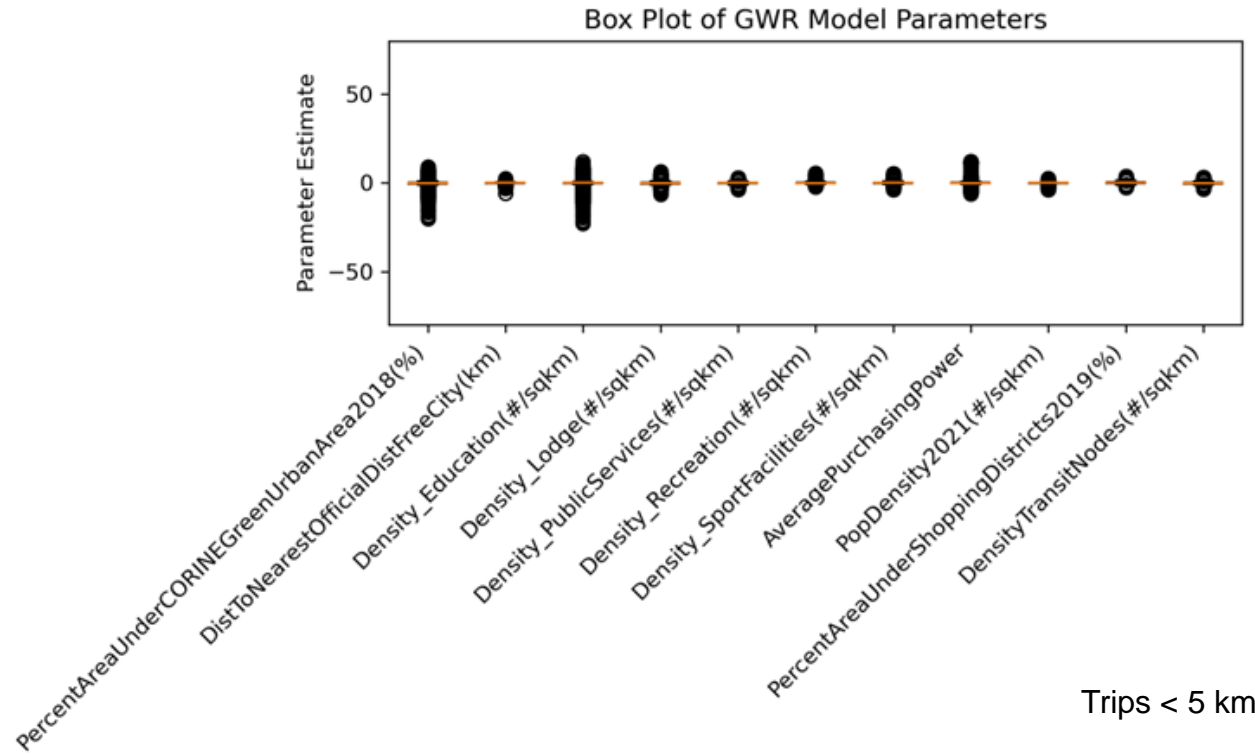
Bandwidth t: 179

Trips < 5 km

Adj. R<sup>2</sup>: 0,459

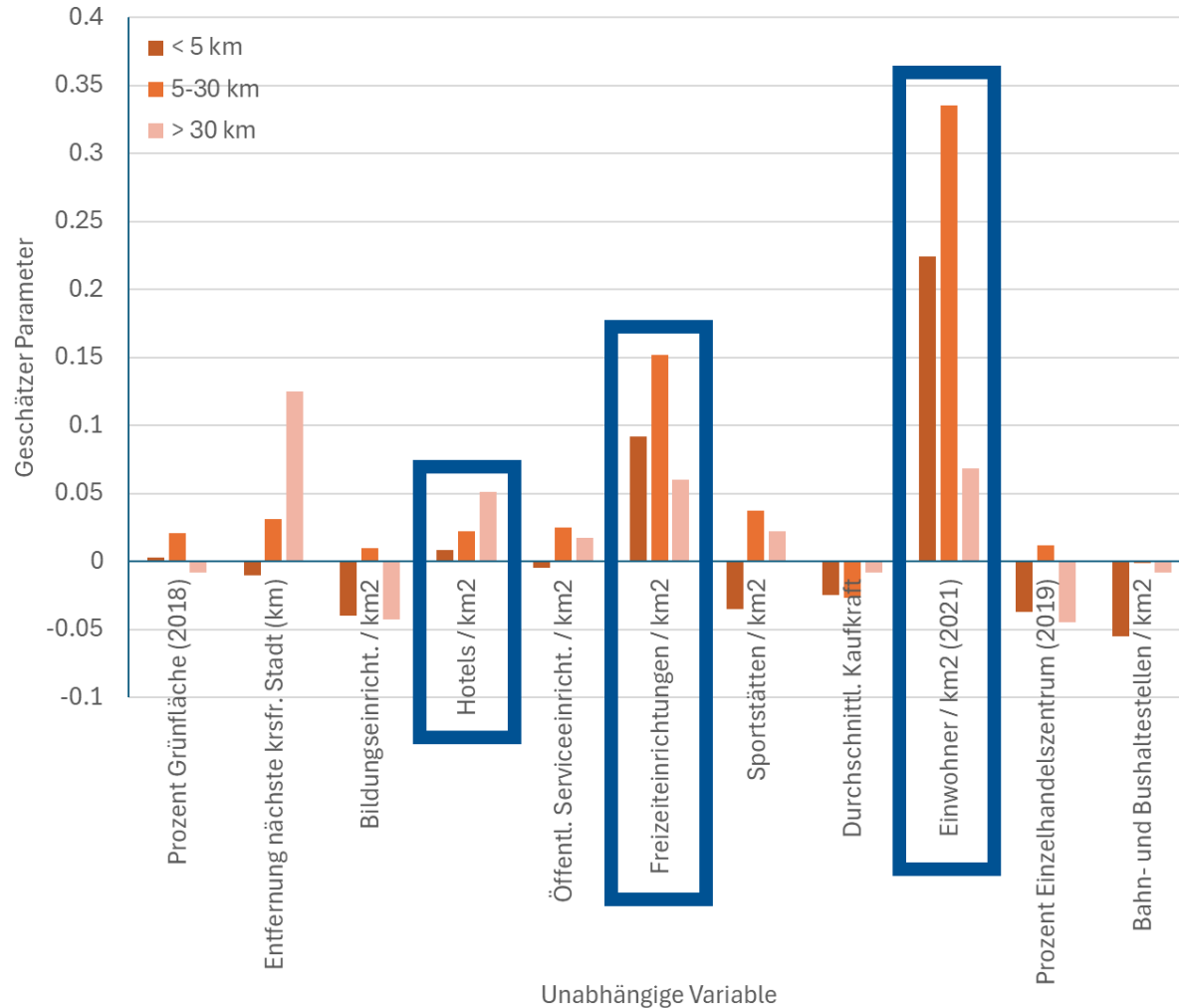


# Geogr. weighted regression: Range of estimated parameters



# Summary of statistical analyses

|  | < 5 km |       |       |       | 5 to 30 km |       |       |       | > 30 km |       |       |       |
|--|--------|-------|-------|-------|------------|-------|-------|-------|---------|-------|-------|-------|
| Model  | LR     | GAM   | XGB   | GWR   | LR         | GAM   | XGB   | GWR   | LR      | GAM   | XGB   | GWR   |
| Adj. R <sup>2</sup> or Pseudo R <sup>2</sup> | 0.067  | 0.120 | 0.280 | 0.459 | 0.134      | 0.175 | 0.266 | 0.445 | 0.012   | 0.075 | 0.253 | 0.377 |
| Moranindex                                   |        | 0.216 | 0.156 | 0.026 |            | 0.097 | 0.086 | 0.012 |         | 0.131 | 0.114 | 0.036 |



Normalized mean parameters of geographically weighted regression

# Conclusions

- Densely populated areas and leisure facilities gained in popularity for trip attraction from pre- to post-COVID
- Mixed impact of retail: mega shopping centers lost in attractiveness, but other shopping centers gained
- **Major strength** of mobile phone data: Massive data available almost instantaneously
- **Major weakness** of mobile phone data: Unknown processing of data provider and limited spatial resolution

