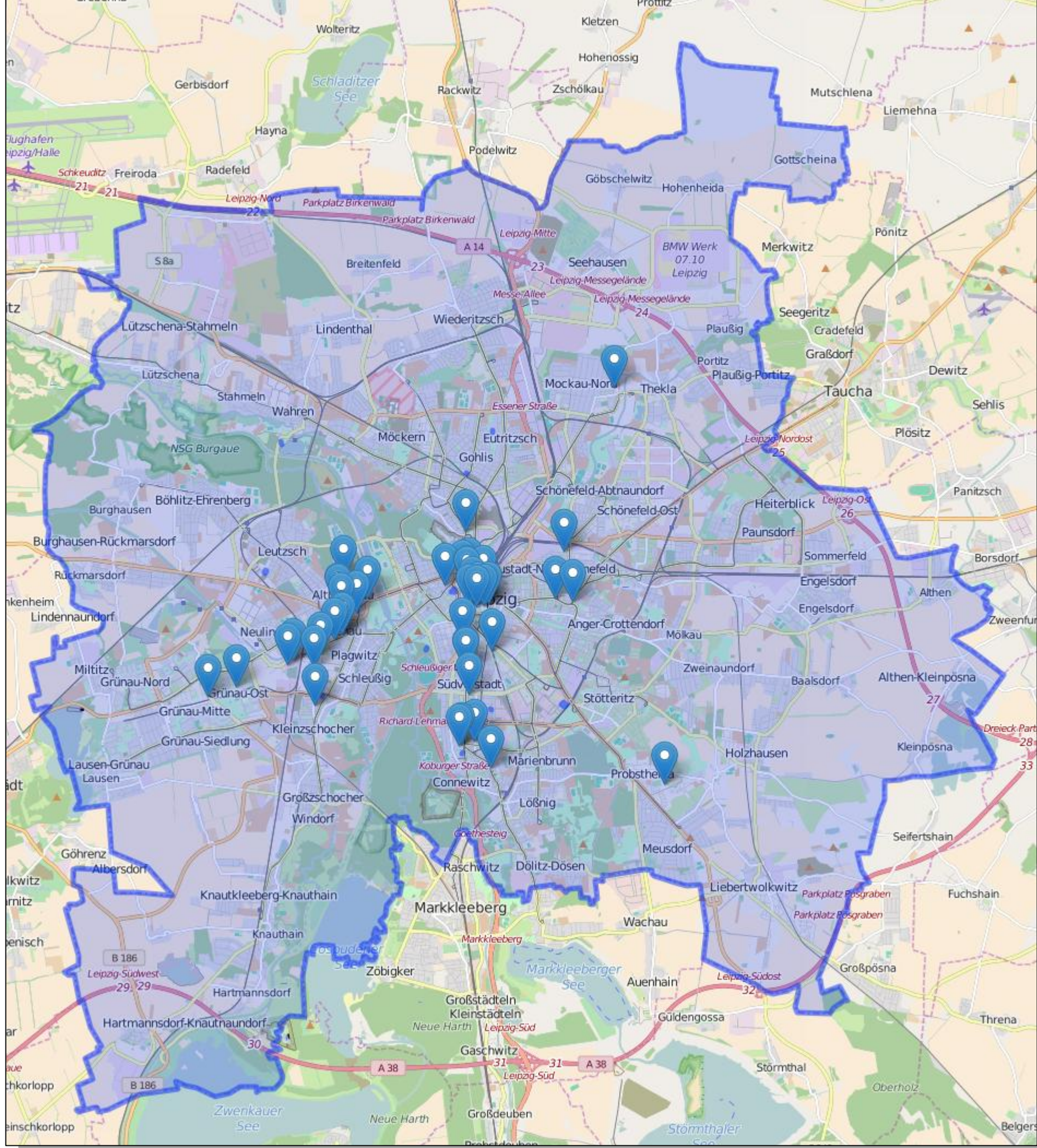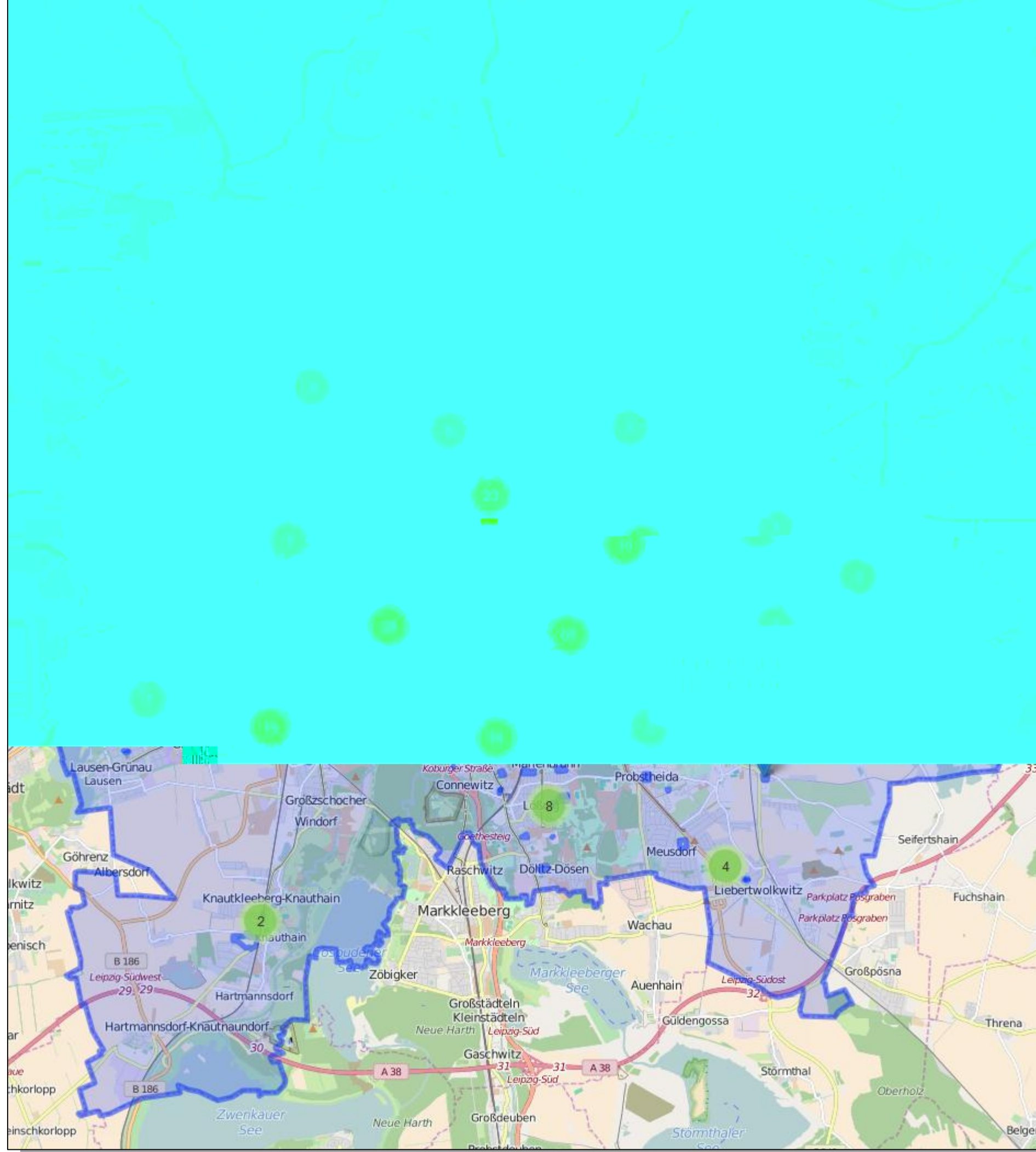# Extracting Geo-Semantics About Cities From OpenStreetMap

Mónica Posada-Sánchez · Stefan Bischof · Axel Polleres
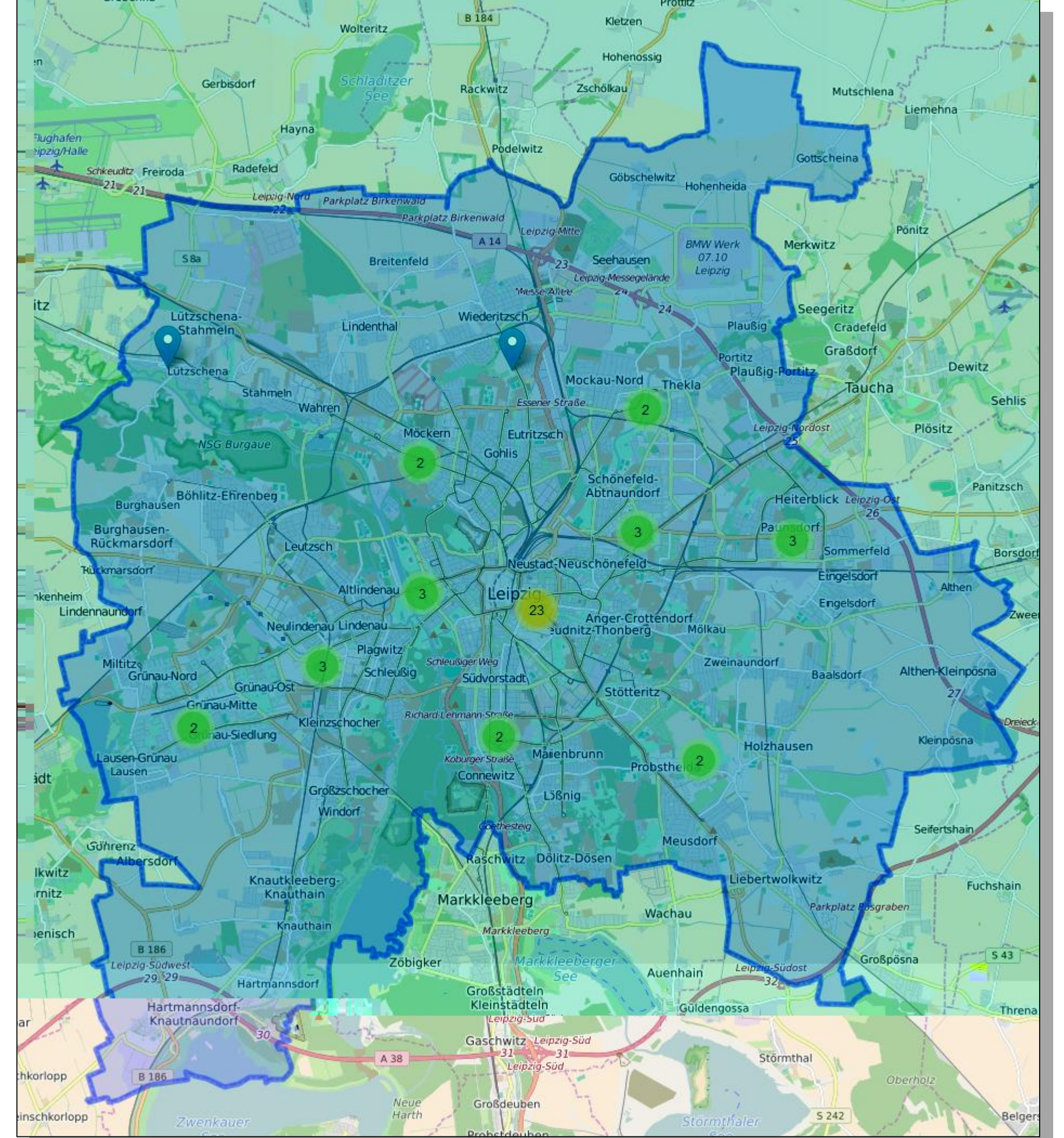
Leipzig culture OSM feature layers
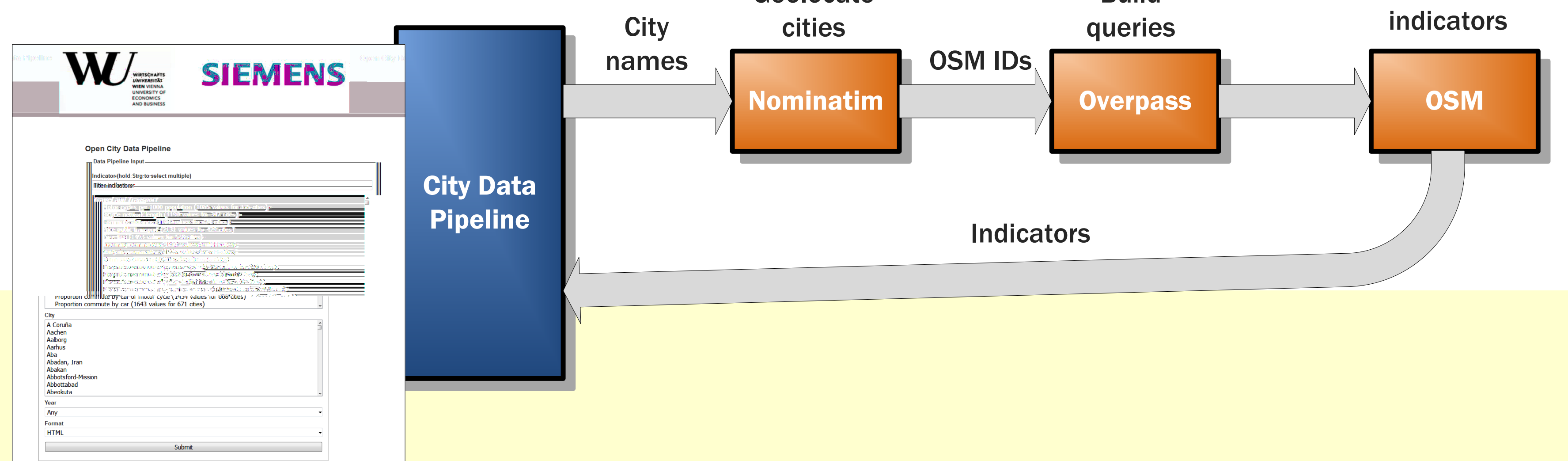
Leipzig education ways layer
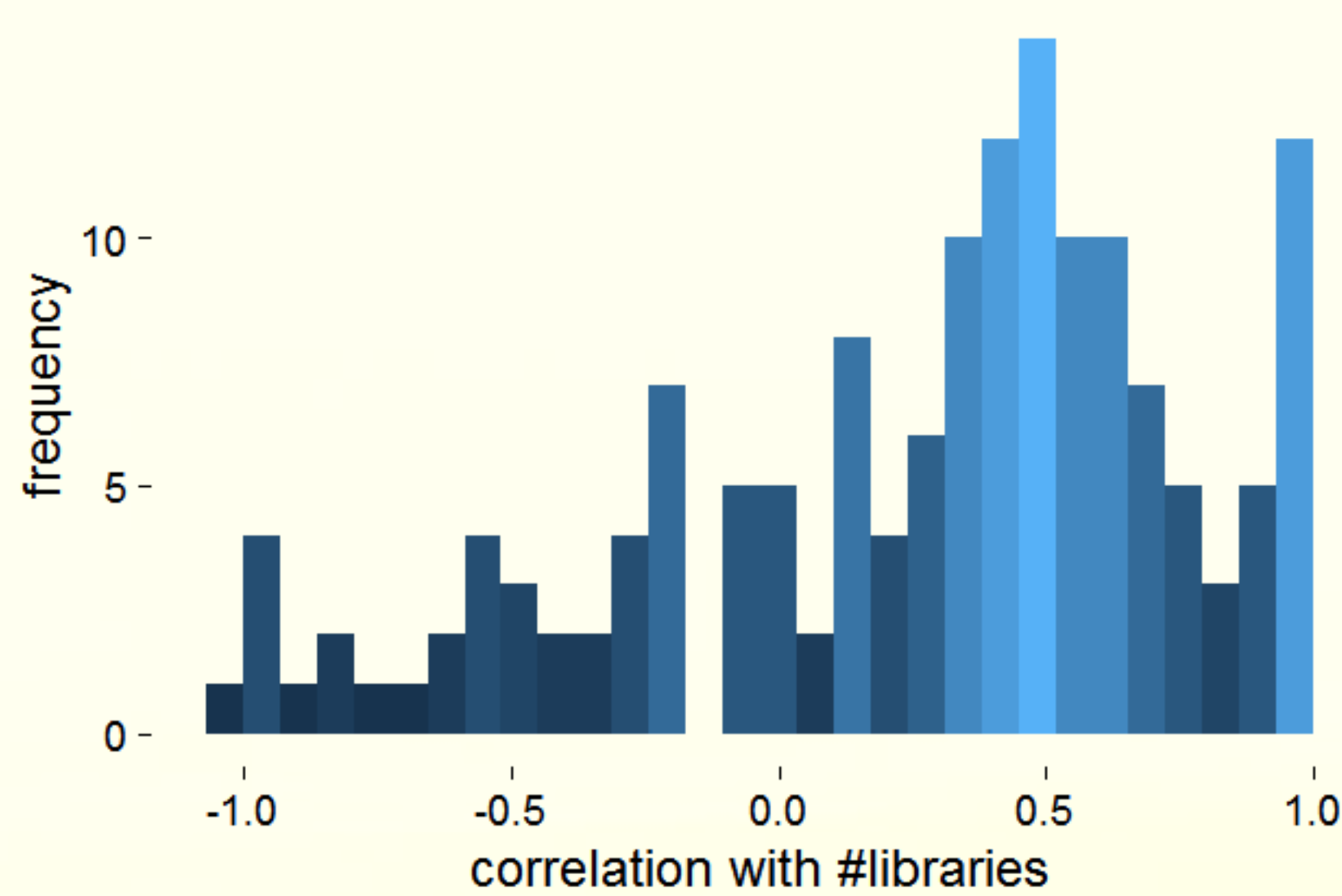
Leipzig libraries nodes layer

Abstract

Access to high quality and updated data is crucial to assess and contextualize city state of affairs. The City Data Pipeline uses diverse Open Data sources to integrate statistical information about cities. The resulting incomplete dataset is not directly usable for data analysis. We exploit data from a geographic information system, namely OpenStreetMap, to obtain new indicators for cities with better coverage. We show that OpenStreetMap is a promising data source for statistical data about cities.

Workflow

City names → Geolocate cities — **Nominatim** → OSM IDs → Build queries — **Overpass** → Get indicators — **OSM**

City Data Pipeline

Indicators

Relation to City Data Pipeline Indicators

frequency / correlation with #libraries

OSM derived indicator definitions

Theatre	[amenity=theatre],
libraries	["amenity"~"library|public_bookcase"],
Education	["amenity"~"library|public_bookcase|college|kindegarten|school|music_school|language_school|university"],
schools	["amenity"~"school|kindegarden|kindegarten|guarderia"],
sport_ctr	["leisure"="sports_centre"],
culture	["amenity"~"theatre|cinema|arts_centre"],
market	[amenity=market_place],
police	[amenity=police],
greenery	["landuse"~"forest|grass|greenfield|meadow|orchard|plant_nursery|village_green"],
hotels	["tourism"~"hotel|hostel|motel|apartment|guest_house|motel"],
tourism	["tourism"!~"hotel|hostel|motel|apartment|guest_house|motel"]

Results

Name	OSM id	area	population	#culture	#education	#greenery	#hotels	#libraries	#markets	#police	#schools	#sport ctr	#theatre	#tourism
A Coruna	346810	38.82	244810	10	116	474	33	11	0	8	69	39	5	122850
Aarhus	1784663	477.52	256018	12	121	1093	30	22	0	2	92	51	5	510
Brasov	222220923		253200	6	54	98	54	5	0	7	22	20	3	259869
Galway	1390623	50.61	50800	8	61	778	3	5	0	3	42	19	6	228509
Ghent	897671	158.21	237000	23	179	2660	46	11	0	10	132	58	17	2795557
Graz	109163	133.025	274000	48	284	1159	70	55	0	16	144	52	22	643546
Karlsruhe	62518	175.88	283959	32	278	3979	70	48	0	13	195	127	15	940703

Reference

Stefan Bischof, Christoph Martin, Axel Polleres, and Patrik Schneider. Collecting, Integrating, enriching and republishing open city data as linked data. In Proceedings of the 14th International Semantic Web Conference (ISWC'15), Part II, pages 57–75. Springer, 2015.