Appendix 2b – Data preparation: Building preparation

Why building preparation?

* To create the Detour Index (DI) and the Local Significance (LS) index, we are only interested in the residential buildings.
* By downloading all buildings from OSM that are located inside a certain area, we have also downloaded buildings with non-residential function, though.
* In addition to filtering out non-residential buildings, we need population data attached to buildings to calculate the LS values.
* These tasks will be executed during the building preparation process.

Building preparation

* During the process of downloading the building data from the Overpass\_API, we retrieved a plethora of polygons representing all buildings in a city.
* Attached to the building data there is often some information on the building type or use.
* Since the OSM contributors that create the building data have different levels of experience and different motivations, the annotation of these information is rather inconsistent, so we cannot solely rely on it.
* To circumvent the inconsistent annotation, we will compare the OSM buildings with the urban atlas (UA) data that we have downloaded manually.
* UA land-use data provides information on residential areas of cities on a European level.
* We can utilize the information on residential areas to filter out all OSM buildings that are located outside of these areas.
* This way we try to retrieve all residential buildings of a city.
* A side-effect of this procedure is, that we can in the same step join the population data that is contained in the UA dataset to the OSM buildings.
* Information on the population is mandatory for calculating the LS index.
* For further steps in the analysis it is also required, that we proceed with the centroids of the buildings, rather than the polygons.
* We have bundeled these tasks in the *buildingPrep* function of our data preparation workflow.

Functionality

Population data:

* Copernicus: Urban atlas data from 2018
* Population mostly aggregated on block level
  + Multiple buildings aggregated
* How does Copernicus get / disaggregate population data?

Buildings:

* OpenStreetMap
* Downloaded for UA FUAs and clipped to city core (Manu’s Layer)
* Buildings selected inside UA residential classes
  + Classes: 11100, 11210, 11220, 11230, 11240, 11300
* UA population disaggregated by building area

1. First approach:
   1. Building area of all buildings summed in each UA polygon
   2. UA population distributed to buildings based on their share of overall building area
2. For buildings that are located inside UA residential polygons with population 0:
   1. calculated average population per building area for each UA residential class in each downloaded OSM tile
   2. multiplied resulting value with building area