Visualising Geo-Spatial Data with Python

Beneth Musiime

Overview

- Introduction to spatial data
- Python based Tools for spatial data
- Importing data from source
- Visualisation Demo
  - Setting the Mapping area
  - Plotting points
  - Adding tooltips and
  - Plotting polygons
  - Combining points and polygons

Key takeaways

- Have a better understanding of tools used for geospatial data
- Filtering and transforming Spatial Data
- Merging and Joining multiple GeoDataFrames
- Visualise Points, Lines and polygons

Geospatial Data

Points

Lines

Polygones

Prerequisites

- python 3.4.5
- pandas 0.24.2
- geopandas 0.4.1
- spaciy 1.6.4
- follium 0.9.1
- matplotlib 3.1.0

Data Source

Airbnb Listing in Berlin:
https://www.kaggle.com/brittabettendorf/berlin-airbnb-data#listings.csv

```
import pandas as pd

df = pd.read_csv("filename.csv", index_col=0)
df.head()
```

Importing Spatial Data

Map of neighborhoods in Berlin:
https://github.com/funkeinteraktiv/Berlin-Geodaten/blob/master/berlin_bezirke.geojson

```
import geopandas as gpd

df_geo = gpd.read_file('geodata_filename.geojson')

df_geo.head()
```

Data Transformation

Filtering

Merging DataFrames

Joining spatial data

Get the center of the polygon

Calculate area for each polygon

Merging DataFrames

pandas.merge

how : {‘left’, ‘right’, ‘outer’, ‘inner’}, default ‘inner’

Specifies the type of join that will occur

on : label or list(*must be found in both DataFrames*)

Specifies the column or index level names to join on.

```
merged_df = pd.merge(gdf, df, how='inner', on = 'name')
```

Joining spatial data

geopandas.sjoin

op: (*intersects,within,contains*) specifies how geopandas will join the attributes of one object to another

how : specifies the type of join that will occur and which geometry is retained in result geodataframe.

```
joined_gpd = gpd.sjoin(polygon_gpd, points_gpds,
                       how='inner', op = 'within')
```

Create a map with folium

Basic Parameters

location – Latitude(Northing), Longitude(Easting).

tiles (str, default 'OpenStreetMap')

zoom_start (int, default 10) – Initial zoom level for the map.

Heat maps with choropleth

Arguments:

geo_data - the source data for the polygons

name - geometry column to be plotted

data - the source DataFrame or Series for the normalized data

columns - a list of columns: polygons, and plot values

Questions?

And discussions

ThoughtWorks®