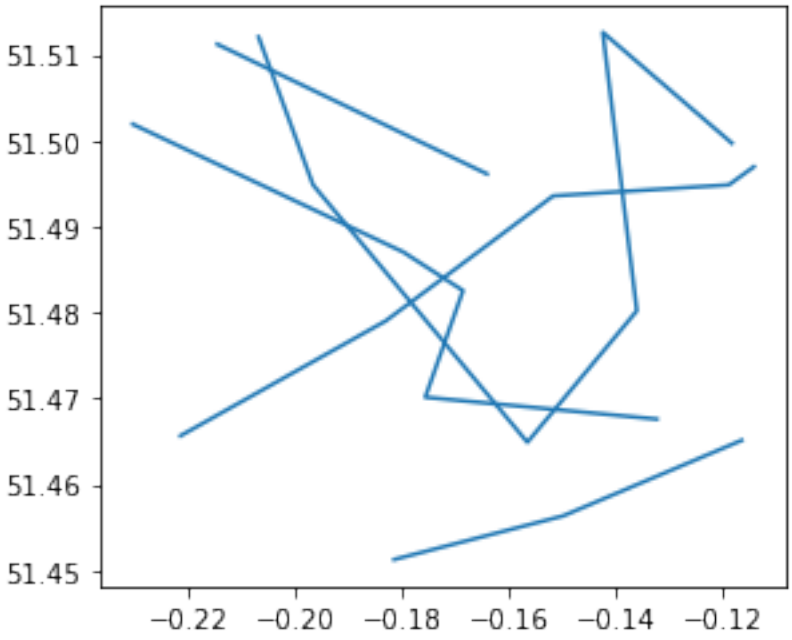
notebook

May 25, 2022

[ ]:
```
import geopandas as gpd

gdf = gpd.read_file('selfintersects.geojson')
gdf.plot()
```

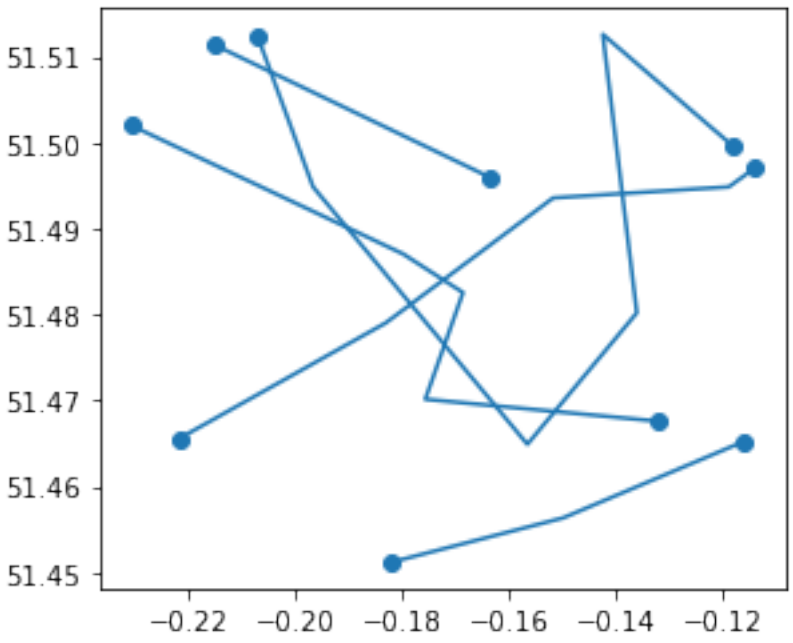
[ ]: <AxesSubplot:>



[ ]:
```
def get_endpoints(gdf):
    from shapely.geometry import Point
    startpoint = gdf.geometry.apply(lambda x: x.coords[0])
    endpoint = gdf.geometry.apply(lambda x: x.coords[-1])

    startpoints = [Point(i) for i in startpoint]
    endpoints = [Point(i) for i in endpoint]

    return startpoints, endpoints
```

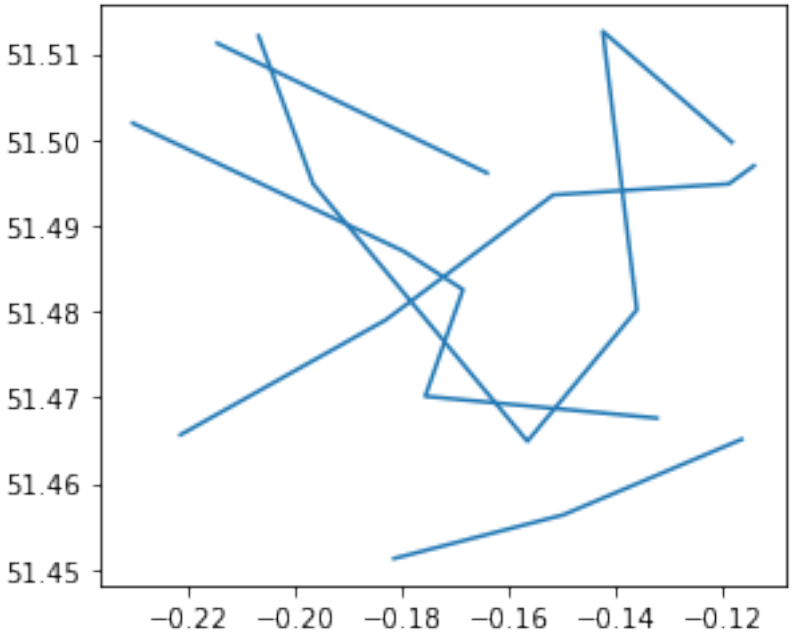
```
def create_endpoints(startpoints, endpoints):
    geom = []
    for a,b in zip(startpoints, endpoints):
        from shapely.geometry import Point
        geom.append(a)
        geom.append(b)

    endpoints = gpd.GeoDataFrame({'id': range(0, len(geom))}, crs=gdf.crs,␣
 ↪geometry=geom)
    return endpoints

startpoints, endpoints = get_endpoints(gdf)
endpoints = create_endpoints(startpoints, endpoints)
import matplotlib.pyplot as plt

fig, ax = plt.subplots()
gdf.plot(ax=ax)
endpoints.plot(ax=ax)
```

[ ]: <AxesSubplot:>



[ ]:
```
union_geom = gdf.unary_union
union = gpd.GeoDataFrame({'id':[0]}, crs=gdf.crs, geometry=[gdf.unary_union])
union.plot()
```

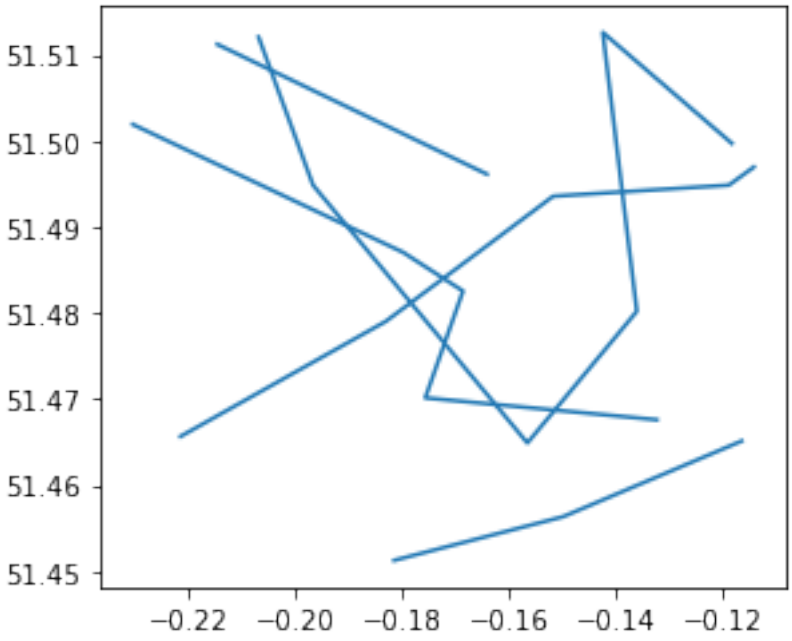
[ ]: <AxesSubplot:>



[ ]:
```
from shapely.ops import linemerge

lm = gpd.GeoDataFrame({'id':[0]}, crs=gdf.crs,␣
 ↪geometry=[linemerge(union_geom)]).explode().reset_index(drop=True)
lm.plot()
```

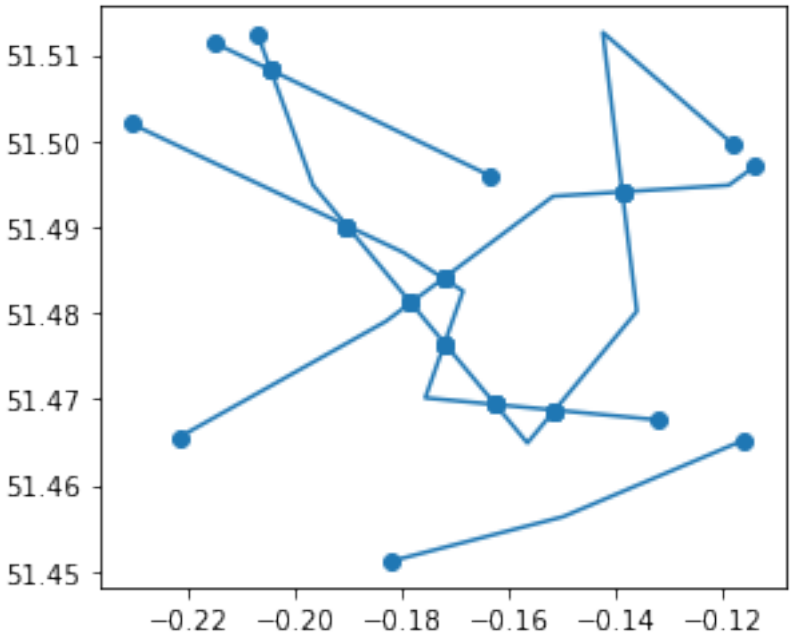
```
C:\Users\SutanMufti\AppData\Local\Temp\ipykernel_31876\1472115873.py:3:
FutureWarning: Currently, index_parts defaults to True, but in the future, it
will default to False to be consistent with Pandas. Use `index_parts=True` to
keep the current behavior and True/False to silence the warning.
  lm = gpd.GeoDataFrame({'id':[0]}, crs=gdf.crs,
geometry=[linemerge(union_geom)]).explode().reset_index(drop=True)
```

[ ]: <AxesSubplot:>



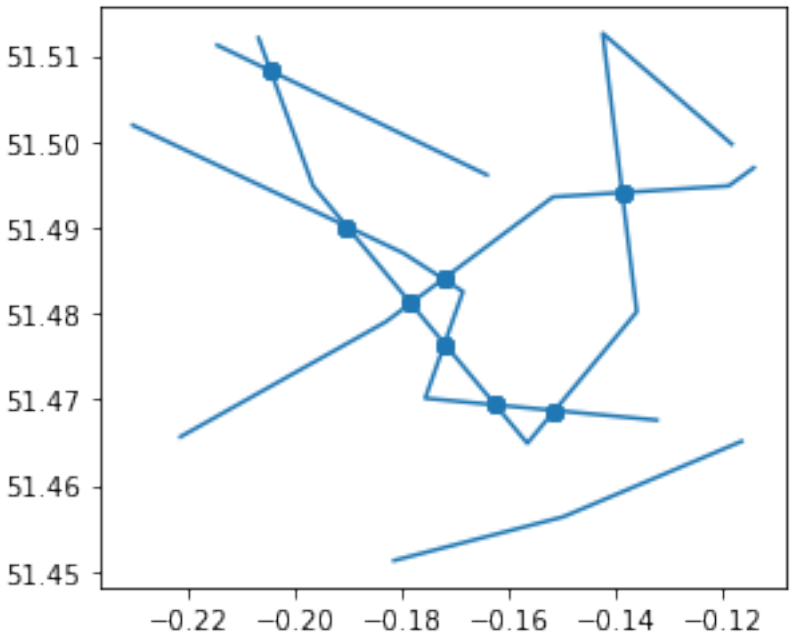
[ ]:
```
startpoints, endpoints = get_endpoints(lm)
endpoints = create_endpoints(startpoints, endpoints)

# cleansing with snap
from shapely.ops import snap
endpoints['geometry'] = endpoints.geometry.apply(lambda x: snap(x, union_geom,␣
↪0.00001))

fig, ax = plt.subplots()
gdf.plot(ax=ax)
endpoints.plot(ax=ax)
```

[ ]:
```
<AxesSubplot:>
```



[ ]:
```
sjoin = endpoints.sjoin(gdf, how='left')

fig, ax = plt.subplots()
gdf.plot(ax=ax)
sjoin[sjoin['index_right'].isna()].plot(ax=ax)
```

[ ]: <AxesSubplot:>



[ ]:
```
sjoin
```

[ ]:

	id	geometry	index_right
0	0	POINT (-0.23054 51.50201)	2.0
1	1	POINT (-0.19028 51.49015)	NaN
2	2	POINT (-0.22161 51.46556)	1.0
3	3	POINT (-0.17832 51.48120)	NaN
4	4	POINT (-0.21475 51.51131)	4.0
5	5	POINT (-0.20446 51.50822)	NaN
6	6	POINT (-0.20685 51.51227)	0.0
7	7	POINT (-0.20446 51.50822)	NaN
8	8	POINT (-0.20446 51.50822)	NaN
9	9	POINT (-0.19028 51.49015)	NaN
10	10	POINT (-0.20446 51.50822)	NaN
11	11	POINT (-0.16376 51.49603)	4.0
12	12	POINT (-0.19028 51.49015)	NaN
13	13	POINT (-0.17832 51.48120)	NaN
14	14	POINT (-0.19028 51.49015)	NaN
15	15	POINT (-0.17223 51.48403)	NaN
16	16	POINT (-0.18179 51.45123)	3.0
17	17	POINT (-0.11621 51.46524)	3.0
18	18	POINT (-0.17832 51.48120)	NaN
19	19	POINT (-0.17223 51.48403)	NaN
20	20	POINT (-0.17832 51.48120)	NaN
21	21	POINT (-0.17202 51.47649)	NaN
22	22	POINT (-0.17223 51.48403)	NaN
23	23	POINT (-0.13885 51.49407)	NaN
24	24	POINT (-0.17223 51.48403)	NaN
25	25	POINT (-0.17202 51.47649)	NaN
26	26	POINT (-0.17202 51.47649)	NaN
27	27	POINT (-0.16253 51.46939)	NaN
28	28	POINT (-0.17202 51.47649)	NaN
29	29	POINT (-0.16253 51.46939)	NaN
30	30	POINT (-0.16253 51.46939)	NaN
31	31	POINT (-0.15145 51.46873)	NaN
32	32	POINT (-0.16253 51.46939)	NaN
33	33	POINT (-0.15145 51.46873)	NaN
34	34	POINT (-0.15145 51.46873)	NaN
35	35	POINT (-0.13885 51.49407)	NaN
36	36	POINT (-0.15145 51.46873)	NaN
37	37	POINT (-0.13201 51.46759)	2.0
38	38	POINT (-0.13885 51.49407)	NaN
39	39	POINT (-0.11398 51.49710)	1.0
40	40	POINT (-0.13885 51.49407)	NaN

```
    41  41  POINT (-0.11810 51.49955)          0.0
```

[ ]: