


```
<class 'geopandas.geodataframe.GeoDataFrame'>
MultiIndex: 4306 entries, ('node', 3762437414) to ('way', 507071972)
Data columns (total 83 columns):
 #   Column           Non-Null Count Dtype  
--- 
 0   barrier          8 non-null     object  
 1   geometry         4306 non-null  geometry 
 2   leisure           3528 non-null  object  
 3   addr:city        69 non-null    object  
 4   addr:district    62 non-null    object  
 5   addr:housenumber 10 non-null    object  
 6   addr:neighbourhood 81 non-null  object  
 7   addr:street      45 non-null    object  
 8   name              619 non-null   object  
 9   access            92 non-null    object  
 10  highway           1 non-null    object  
 11  nodes             4216 non-null  object  
 12  landuse           790 non-null   object  
 13  source            479 non-null   object  
 14  height            3 non-null    object  
 15  museum            1 non-null    object  
 16  name:de           1 non-null    object  
 17  name:en           6 non-null    object  
 18  tourism           2 non-null    object  
 19  wikipedia:en     1 non-null    object  
 20  wikipedia:tr     1 non-null    object  
 21  wheelchair        47 non-null   object  
 22  sport              1124 non-null  object  
 23  addr:postcode     23 non-null   object  
 24  surface            113 non-null  object  
 25  operator           35 non-null   object  
 26  operator:type     12 non-null   object  
 27  ways               297 non-null  object  
 28  type               297 non-null  object  
 29  note               6 non-null    object  
 30  phone              2 non-null    object  
 31  website            5 non-null    object  
 32  leaf_cycle        8 non-null    object  
 33  leaf_type          110 non-null  object  
 34  natural            24 non-null   object  
 35  layer               1 non-null   object  
 36  wikidata           33 non-null  object  
 37  building           4 non-null   object  
 38  wikimedia_commons 10 non-null   object  
 39  wikipedia          14 non-null  object  
 40  rest               1 non-null   object  
 41  opening_hours       14 non-null  object  
 42  name:tr            1 non-null   object  
 43  hoops              28 non-null  object  
 44  lit                65 non-null  object  
 45  roundabout         1 non-null   object  
 46  int_name           1 non-null   object  
 47  addr:housename     1 non-null   object  
 48  garden:type        10 non-null  object  
 49  max_age            10 non-null  object  
 50  description         16 non-null  object  
 51  dog                4 non-null   object  
 52  smoking            3 non-null   object  
 53  start_date         2 non-null   object  
 54  fee                7 non-null   object  
 55  covered             3 non-null   object  
 56  indoor              3 non-null   object  
 57  min_age            11 non-null  object  
 58  playground:theme    6 non-null   object  
 59  fence_type          1 non-null   object  
 60  area               1 non-null   object  
 61  traffic_calming   2 non-null   object  
 62  amenity            3 non-null   object  
 63  opening_hours:covid19 1 non-null  object
```

```

64 boundary          1 non-null    object
65 opening_date      1 non-null    object
66 alt_name          4 non-null    object
67 trees             2 non-null    object
68 image              1 non-null    object
69 charge             1 non-null    object
70 email              1 non-null    object
71 construction       1 non-null    object
72 payment:cash       1 non-null    object
73 parking             1 non-null    object
74 supervised          1 non-null    object
75 landcover           1 non-null    object
76 name_1              1 non-null    object
77 cuisine             1 non-null    object
78 gaelic_games:football 1 non-null    object
79 fitness_station     1 non-null    object
80 league              1 non-null    object
81 location             1 non-null    object
82 place               2 non-null    object
dtypes: geometry(1), object(82)
memory usage: 2.9+ MB

```

```
In [5]: green_izmir.isnull().sum()
```

```

Out[5]: barrier        4298
geometry         0
leisure          778
addr:city        4237
addr:district    4244
...
gaelic_games:football 4305
fitness_station   4305
league            4305
location          4305
place             4304
Length: 83, dtype: int64

```

Data Filtering

Izmir OSM green space dataset consists of a total of 4.306 records and 83 columns. Since the analysis focuses on the number of green areas per neighbourhood and borough, there is no need to keep all the columns. We will keep only necessary columns and remove the rest.

```
In [6]: green_izmir_filtered = green_izmir[["geometry", "leisure", "landuse", "natural", "name"]]
```

```
In [7]: green_izmir_filtered.head()
```

		geometry	leisure	landuse	natural	name
element_type	osmid					
node	3762437414	POINT (27.37389 37.94258)	playground	NaN	NaN	NaN
	3999523005	POINT (27.27746 37.93217)	playground	NaN	NaN	NaN
	4002504031	POINT (27.37345 37.96300)	playground	NaN	NaN	NaN
	4002504032	POINT (27.37315 37.96162)	playground	NaN	NaN	NaN
	6790231538	POINT (27.35375 38.15291)	playground	NaN	NaN	NaN

Geometries Type

The entries in the geometry column have different types of geometries; Polygon, MultiPolygon, Point. Since only the number of green spaces are needed, all geometries will be converted to Point type.


```
In [16]: green_points_izmir_proj.geom_type.value_counts()
```

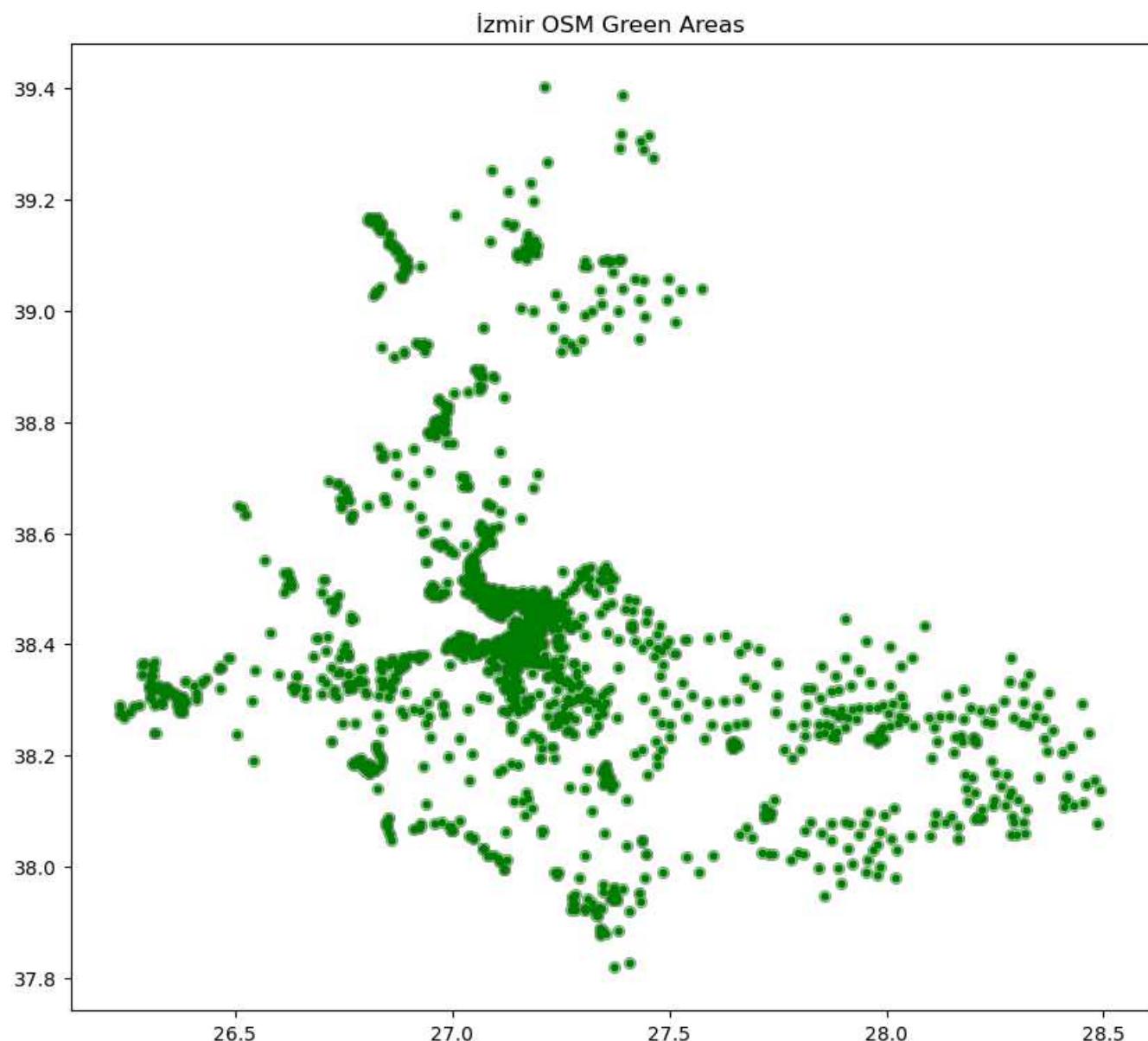
```
Out[16]: Point    4306  
Name: count, dtype: int64
```

```
In [17]: green_points_izmir_deg = green_points_izmir_proj.to_crs(green_points_izmir.crs)
```

```
In [18]: green_points_izmir_deg.crs
```

```
Out[18]: <Geographic 2D CRS: EPSG:4326>  
Name: WGS 84  
Axis Info [ellipsoidal]:  
- Lat[north]: Geodetic latitude (degree)  
- Lon[east]: Geodetic longitude (degree)  
Area of Use:  
- name: World.  
- bounds: (-180.0, -90.0, 180.0, 90.0)  
Datum: World Geodetic System 1984 ensemble  
- Ellipsoid: WGS 84  
- Prime Meridian: Greenwich
```

```
In [19]: fig, ax = plt.subplots(figsize=(10,10))  
green_points_izmir_deg.plot(ax=ax, color = "lightgreen", edgecolor = "gray", linewidth = 0.5)  
green_points_izmir_deg.plot(ax = ax, color="green", markersize = 10)  
plt.title("İzmir OSM Green Areas")  
plt.show()
```



```
In [20]: green_points_izmir_deg.to_file("İzmir_green_data.gpkg", layer='points', driver="GPKG")
```

```
In [21]: gdf = gpd.read_file("Izmir_green_data.gpkg")
```

```
In [22]: gdf.head()
```

```
Out[22]:
```

	element_type	osmid	leisure	landuse	natural	name	geometry
0	node	3762437414	playground	None	None	None	POINT (27.37389 37.94258)
1	node	3999523005	playground	None	None	None	POINT (27.27746 37.93217)
2	node	4002504031	playground	None	None	None	POINT (27.37345 37.96300)
3	node	4002504032	playground	None	None	None	POINT (27.37315 37.96162)
4	node	6790231538	playground	None	None	None	POINT (27.35375 38.15291)

```
In [23]: gdf.info()
```

```
<class 'geopandas.geodataframe.GeoDataFrame'>
RangeIndex: 4306 entries, 0 to 4305
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   element_type 4306 non-null   object 
 1   osmid        4306 non-null   int64  
 2   leisure       3528 non-null   object 
 3   landuse      790 non-null   object 
 4   natural      24 non-null    object 
 5   name         619 non-null   object 
 6   geometry     4306 non-null   geometry
dtypes: geometry(1), int64(1), object(5)
memory usage: 235.6+ KB
```

Outcome

The dataset is now ready with consistent point geometries, proper CRS, and only relevant attributes. This ensures accurate spatial analysis in the next step.

Transition to Step-2

The next step will focus on preparing the neighbourhood shapefile for the spatial join, which includes filtering, cleaning, and generating borough-level boundaries from the neighbourhood data.

```
In [ ]:
```