# Map visualisation

July 12, 2023

### Map visualisation

Prepare data of sentiments as a GeoJSON FeatureCollection for using on a OpenStreetMap with the vanilla-js-web-component-leaflet-geojson to look like this.

### Setup

[ ]:
```
!pip install pandas
```

### Preparando los datos

[3]:
```
import pandas as pd

versos_al_paso_geo_url = 'https://github.com/migupl/sentimientos-al-paso/raw/
 ↪main/notebooks/output/versosalpaso_sentiment_text-davinci-003_geo.csv'
versos_al_paso_geo = pd.read_csv(versos_al_paso_geo_url, sep=';',␣
 ↪encoding='utf-8')
versos_al_paso_geo.columns
```

[3]:
```
Index(['Unnamed: 0', 'id', 'latitud', 'longitud', 'autor', 'barrio', 'verso',
       'direccion', 'sentiment', 'quarter', 'district', 'city'],
      dtype='object')
```

[6]:
```
sentiments_by_district = versos_al_paso_geo.groupby('district')pd.
 ↪pivot_table(versos_al_paso_geo[['district', 'sentiment']],␣
index='district', columns='sentiment', aggfunc=len, fill_value=0)
sentiments_by_district
```

[6]:
```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x7f80774f0490>
```

[13]:
```
no_of_districts = len(versos_al_paso_geo.district.unique())
print(f'There are {no_of_districts} districts')
```

```
There are 21 districts
```

[14]:
```
pd.pivot_table(versos_al_paso_geo[['district', 'sentiment']], index='district',␣
 ↪columns='sentiment', aggfunc=len, fill_value=0)
```

[14]:

| sentiment | negative | neutral | positive |
|---|---|---|---|
| district | | | |
| Arganzuela | 1 | 4 | 48 |
| Barajas | 0 | 0 | 11 |
| Carabanchel | 2 | 7 | 72 |
| Centro | 1 | 2 | 67 |
| Chamartín | 0 | 4 | 71 |
| Chamberí | 1 | 0 | 63 |
| Ciudad Lineal | 0 | 1 | 77 |
| Fuencarral-El Pardo | 2 | 4 | 59 |
| Hortaleza | 0 | 5 | 51 |
| Latina | 1 | 7 | 67 |
| Moncloa-Aravaca | 3 | 2 | 58 |
| Moratalaz | 0 | 1 | 30 |
| Puente de Vallecas | 0 | 5 | 53 |
| Retiro | 1 | 1 | 45 |
| Salamanca | 1 | 3 | 59 |
| San Blas - Canillejas | 1 | 1 | 51 |
| Tetuán | 2 | 2 | 39 |
| Usera | 0 | 4 | 42 |
| Vicálvaro | 0 | 1 | 22 |
| Villa de Vallecas | 1 | 0 | 11 |
| Villaverde | 1 | 2 | 30 |

### Define GeoJSON Feature

[7]:
```
import json

def feature_json(latitude: float, longitude: float, author: str, verse: str,␣
 ↪sentiment: str) -> json:
    colors = {'negative': 'red', 'neutral': 'yellow', 'positive': 'green'}
    return {
        'type': "Feature",
        'geometry': {
            'type': "Point",
            'coordinates': [longitude, latitude]
        },
        'properties': {
            'popupContent': f'<strong>{verse}</strong><br>- <cite>{author}</
 ↪cite>',
            'icon': {
                "iconUrl": f'https://raw.githubusercontent.com/migupl/
 ↪svg-vectors-and-icons/main/heart-like/heart-{colors[sentiment]}.png',
                "iconSize": [41, 41],
                "iconAnchor": [20, 41],
                "popupAnchor": [1, -34]
            },
```

```
        }
    }
```

Prepare the JSON of FeatureCollection by district

[9]:
```
geojson = {}
for name, group in sentiments_by_district:
    features = []
    for index, row in group.iterrows():
        feature = feature_json(row.latitud, row.longitud, row.autor, row.verso,␣
 ↪row.sentiment)
        features.append(feature)

    geojson[name] = {
        'type': "FeatureCollection",
        'features': features
    }

keys = geojson.keys()
keys
```

[9]:
```
dict_keys(['Arganzuela', 'Barajas', 'Carabanchel', 'Centro', 'Chamartín',
'Chamberí', 'Ciudad Lineal', 'Fuencarral-El Pardo', 'Hortaleza', 'Latina',
'Moncloa-Aravaca', 'Moratalaz', 'Puente de Vallecas', 'Retiro', 'Salamanca',
'San Blas - Canillejas', 'Tetuán', 'Usera', 'Vicálvaro', 'Villa de Vallecas',
'Villaverde'])
```

Test that there are 21 districts

[10]:
```
import unittest

tc = unittest.TestCase()
no_of_keys = len(keys)
tc.assertEqual(21, no_of_keys)
```

Save JSON for later use

[12]:
```
json_file_path = './output/sentiments_by_district_geo.json'
with open(json_file_path, "w") as outfile:
    json.dump(geojson, outfile)
```