2.4_Analisis Geoespaciales

June 13, 2025

Creado por:

Isabel Maniega

[1]:
```
import warnings
warnings.filterwarnings('ignore')
```

1 Folium

[2]:
```
# pip install folium
```

[3]:
```
import pandas as pd
import folium
```

Extraemos la información de un dataset sobre las estaciones de las bicicletas en Dublín:

[4]:
```
location = "https://data.smartdublin.ie/dataset/
  ↪33ec9fe2-4957-4e9a-ab55-c5e917c7a9ab/resource/
  ↪2dec86ed-76ed-47a3-ae28-646db5c5b965/download/dublin.csv"
bike_station_locations = pd.read_csv(location)
```

[5]:
```
bike_station_locations
```

[5]:

	Number	Name	Address	Latitude	Longitude
0	42	SMITHFIELD NORTH	Smithfield North	53.349562	-6.278198
1	30	PARNELL SQUARE NORTH	Parnell Square North	53.353462	
2	54	CLONMEL STREET	Clonmel Street	53.336021	
3	108	AVONDALE ROAD	Avondale Road	53.359405	
4	56	MOUNT STREET LOWER	Mount Street Lower	53.337960	
..	…	…	…	…	
105	39	WILTON TERRACE	Wilton Terrace	53.332383	
106	83	EMMET ROAD	Emmet Road	53.340714	
107	92	HEUSTON BRIDGE (NORTH)	Heuston Bridge (North)	53.347802	
108	21	LEINSTER STREET SOUTH	Leinster Street South	53.342180	
109	88	BLACKHALL PLACE	Blackhall Place	53.348800	

```
1    -6.265305
2    -6.262980
3    -6.276142
4    -6.241530
..         …
105  -6.252717
106  -6.308191
107  -6.292432
108  -6.254485
109  -6.281637

[110 rows x 5 columns]
```

[6]:
```
# Creamos un avariable la localización de las estaciones:

bike_station_locations = bike_station_locations[["Latitude", "Longitude",␣
 ↪"Name"]]
```

[7]:
```
mapa = folium.Map(location=[bike_station_locations.Latitude.mean(),
                            bike_station_locations.Longitude.mean()],
                  zoom_start=14, control_scale=True)
```

[8]:
```
# Mostramos el área donde se encuentran las estaciones
mapa
```

[8]:
```
<folium.folium.Map at 0x7d71b1099b50>
```

[9]:
```
# Añadimos los puntos de cada una de las estaciones en el mapa:

for index, location_info in bike_station_locations.iterrows():
    folium.Marker([location_info["Latitude"],
                   location_info["Longitude"]],
                  popup=location_info["Name"]).add_to(mapa)
```

[10]:
```
# Visualizamos de nuevo el mapa completo:
mapa
```

[10]:
```
<folium.folium.Map at 0x7d71b1099b50>
```

1.1 Otro ejemplo: Choropleth maps

[11]:
```
url = (
    "https://raw.githubusercontent.com/python-visualization/folium/master/
 ↪examples/data"
)
state_geo = f"{url}/us-states.json"
state_unemployment = f"{url}/US_Unemployment_Oct2012.csv"
```

```
state_data = pd.read_csv(state_unemployment)

m = folium.Map(location=[48, -102], zoom_start=3)

folium.Choropleth(
    geo_data=state_geo,
    name="choropleth",
    data=state_data,
    columns=["State", "Unemployment"],
    key_on="feature.id",
    fill_color="YlGn",
    fill_opacity=0.7,
    line_opacity=0.2,
    legend_name="Unemployment Rate (%)",
).add_to(m)

folium.LayerControl().add_to(m)

m
```

[11]: <folium.folium.Map at 0x7d71b0ffc740>

---

2 Python packages for visualizing meteorological data

[12]:
```
# https://cds.climate.copernicus.eu/datasets
```

2.1 Metpy

[13]:
```
# pip install metpy
```

[14]:
```
# pip install cartopy
```

[15]:
```
from datetime import datetime

import pandas as pd

from metpy.cbook import get_test_data
import metpy.plots as mpplots
from metpy.units import units
from metpy.plots import MapPanel
```

[16]:
```
# Fetch the data and read it in
data = pd.read_csv(get_test_data('UPA_obs.csv', as_file_obj=False))
data
```

[16]:
```
     pressure  height  temperature  dewpoint  direction  speed station  \
0       500.0  5377.0        -28.7     -39.8      316.0   75.0     1M1
1       300.0  8859.0        -47.9       NaN      304.0   96.0     1M1
2       500.0  5110.0        -43.5     -54.5      310.0   23.0    CWPL
3       300.0  8420.0        -53.9     -62.9      300.0   27.0    CWPL
4       500.0  5470.0        -13.7     -14.1      215.0   96.0    CWQI
..        …      …             …         …          …     …      …
216     300.0  8852.0        -51.4       NaN       49.0   48.0    KYAK
217     500.0  5220.0        -23.7     -25.3      185.0   54.0    KYMW
218     300.0  8829.0        -42.9     -46.9      195.0   92.0    KYMW
219     500.0  5379.0        -30.1     -38.1      355.0   19.0    KYXY
220     300.0  8829.0        -52.9     -61.9       15.0   25.0    KYXY

           time     u_wind     v_wind   latitude  longitude
0    1993-03-14  52.099378 -53.950485        NaN        NaN
1    1993-03-14  79.587607 -53.682519        NaN        NaN
2    1993-03-14  17.619022 -14.784115  51.466667      -90.2
3    1993-03-14  23.382686 -13.500000  51.466667      -90.2
4    1993-03-14  55.063338  78.638596        NaN        NaN
..          …        …         …         …          …
216  1993-03-14 -36.226060 -31.490833        NaN        NaN
217  1993-03-14   4.706410  53.794514        NaN        NaN
218  1993-03-14  23.811352  88.865176        NaN        NaN
219  1993-03-14   1.655959 -18.927699        NaN        NaN
220  1993-03-14  -6.470476 -24.148146        NaN        NaN

[221 rows x 12 columns]
```

[17]:
```
data['time'].value_counts()
```

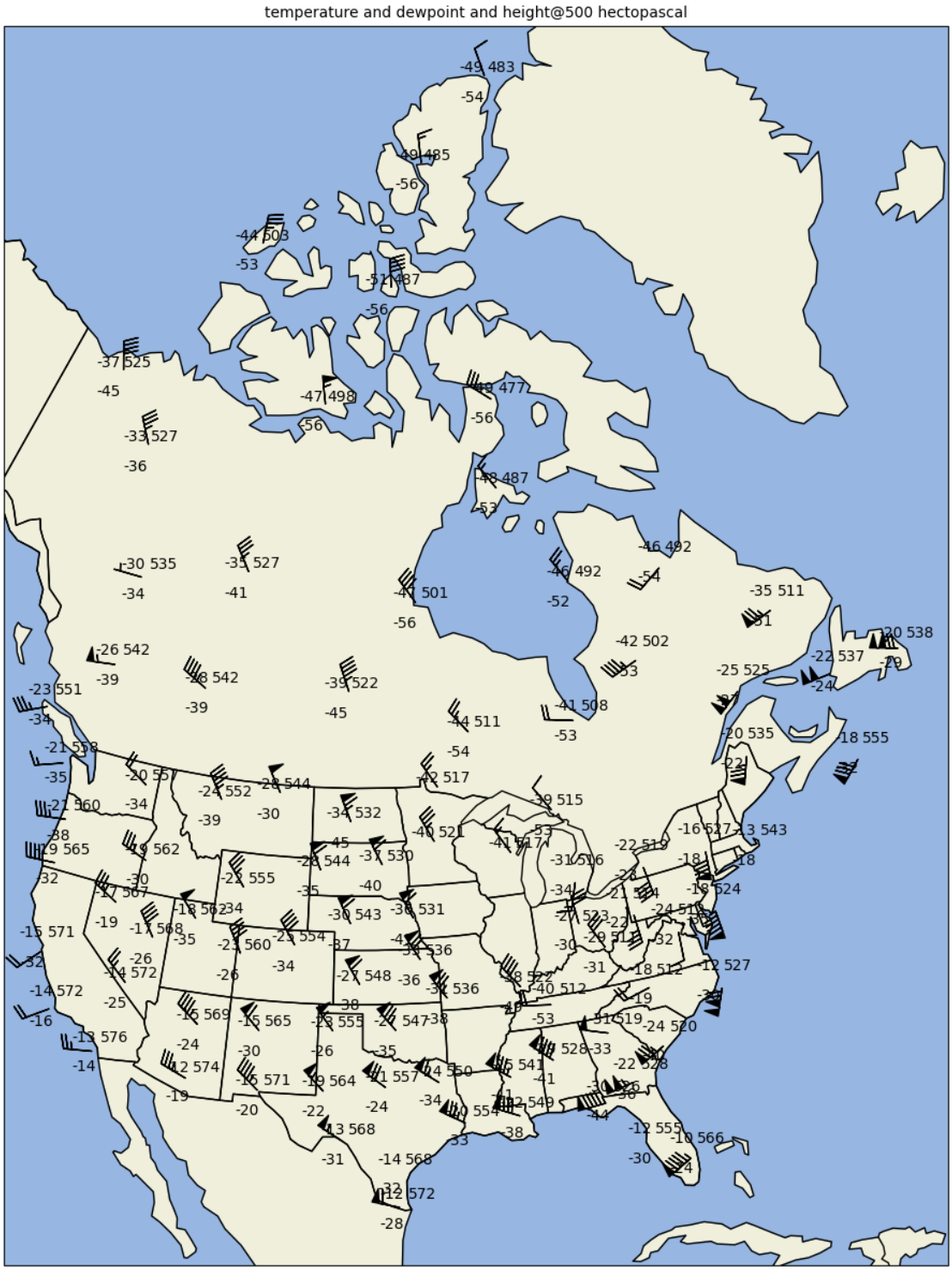
[17]:
```
time
1993-03-14    221
Name: count, dtype: int64
```

[18]:
```
# Plotting the Observations
obs = mpplots.PlotObs()
obs.data = data
obs.time = datetime(1993, 3, 14, 0)
obs.level = 500 * units.hPa
obs.fields = ['temperature', 'dewpoint', 'height']
obs.locations = ['NW', 'SW', 'NE']
obs.formats = [None, None, lambda v: format(v, '.0f')[:3]]
obs.vector_field = ('u_wind', 'v_wind')
obs.reduce_points = 0
```

[19]:
```
# Add map features for the particular panel
panel = MapPanel()
```

```
panel.layout = (1, 1, 1)
panel.projection = 'lcc'
panel.layers = ['coastline', 'borders', 'states', 'land', 'ocean']
panel.plots = [obs]
```

[20]:
```
# Collecting panels for complete figure
pc = mpplots.PanelContainer()
pc.size = (15, 15)
pc.panels = [panel]

# Showing the results
pc.show()
```


temperature and dewpoint and height@500 hectopascal

2.2 Tropycal

[21]:
```
# pip install tropycal
```

[22]:
```
import tropycal.tracks as tracks
import datetime as dt
```

[23]:
```
# read in HURTDAT2 dataset
basin = tracks.TrackDataset(basin='north_atlantic',include_btk=False)
basin
```

```
--> Starting to read in HURDAT2 data
--> Completed reading in HURDAT2 data (1.7 seconds)
```

[23]:
```
<tropycal.tracks.Dataset>
Dataset Summary:
    Basin:              north_atlantic
    Source:             hurdat
    Number of storms:   1991
    Maximum wind:       165 knots (Allen 1980)
    Minimum pressure:   882 hPa (Wilma 2005)
    Year range:         1851 - 2024
```

[24]:
```
basin.to_dataframe()
```

[24]:

season	all_storms	named_storms	hurricanes	major_hurricanes	ace	start_time	end_time
1851	6	6	3	1	36.1	1851-06-25 00:00:00	1851-10-19 18:00:00
1852	5	5	5	1	73.4	1852-08-19 00:00:00	1852-10-11 18:00:00
1853	8	8	4	2	76.4	1853-08-05 12:00:00	1853-10-22 06:00:00
1854	5	5	3	1	31.1	1854-06-25 00:00:00	1854-10-22 18:00:00
1855	5	5	4	1	18.2	1855-08-06 12:00:00	1855-09-17 06:00:00
…	…	…	…	…	…	…	…
2020	31	30	14	7	180.3	2020-05-16 18:00:00	2020-11-18 12:00:00
2021	21	21	7	4	145.3		
2022	16	14	8	2	94.7		
2023	21	20	7	3	148.0		
2024	18	18	11	5	161.6		

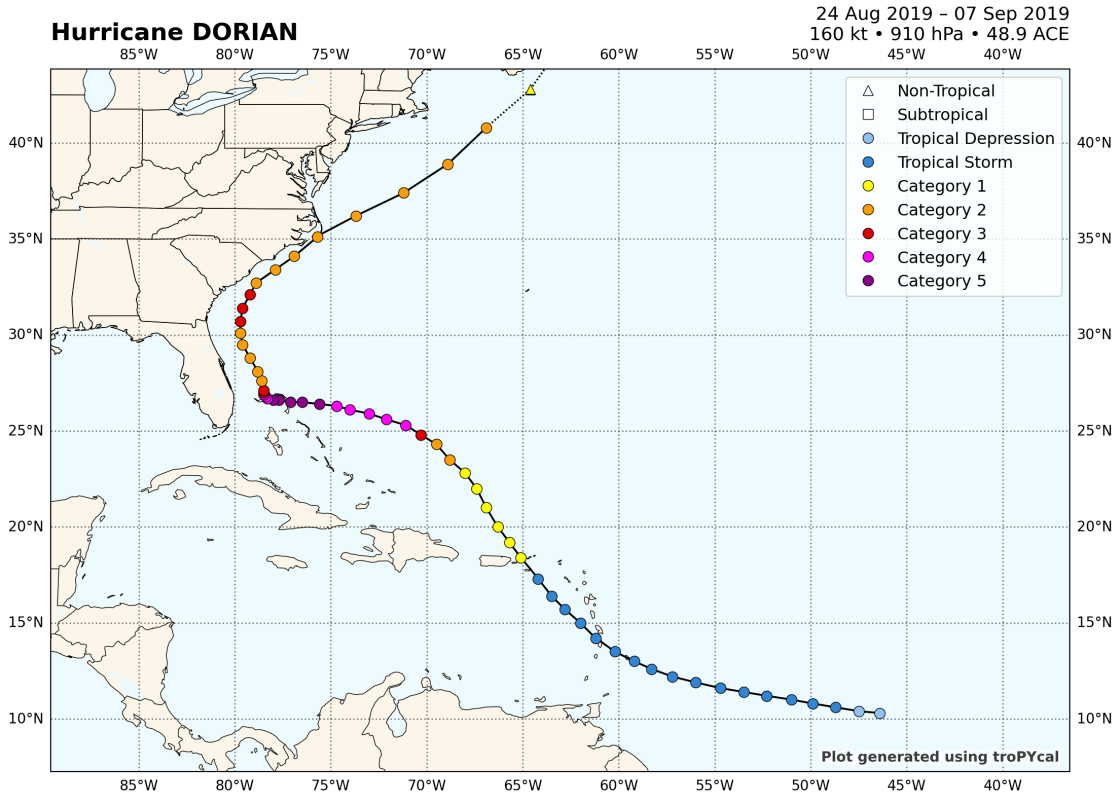
```
2021   2021-05-22 06:00:00 2021-11-07 06:00:00
2022   2022-06-05 00:00:00 2022-11-11 12:00:00
2023   2023-01-16 12:00:00 2023-10-28 18:00:00
2024   2024-06-19 12:00:00 2024-11-18 00:00:00

[174 rows x 7 columns]
```

[25]:
```
# Retrieve an instance of Storm for Hurricane Michael
storm = basin.get_storm(('dorian',2019))
storm.to_dataframe()
```

[25]:
```
                 time  extra_obs special type   lat   lon  vmax  mslp  \
0  2019-08-24 06:00:00          0           TD  10.3 -46.4    25  1011
1  2019-08-24 12:00:00          0           TD  10.4 -47.5    30  1010
2  2019-08-24 18:00:00          0           TS  10.6 -48.7    35  1008
3  2019-08-25 00:00:00          0           TS  10.8 -49.9    35  1008
4  2019-08-25 06:00:00          0           TS  11.0 -51.0    35  1008
..                 …        …     …  …   …   …   …   …   …
65 2019-09-08 00:00:00          0           LO  45.2 -62.9    80   956
66 2019-09-08 06:00:00          0           EX  47.6 -61.9    75   960
67 2019-09-08 12:00:00          0           EX  49.4 -60.4    70   962
68 2019-09-08 18:00:00          0           EX  50.8 -57.9    60   966
69 2019-09-09 00:00:00          0           EX  51.6 -54.8    50   980

         wmo_basin
0   north_atlantic
1   north_atlantic
2   north_atlantic
3   north_atlantic
4   north_atlantic
..             …
65  north_atlantic
66  north_atlantic
67  north_atlantic
68  north_atlantic
69  north_atlantic

[70 rows x 9 columns]
```

[26]:
```
  # Customize our plot
storm.plot(domain='dynamic_tropical')
```

[26]:
```
<GeoAxes: title={'left': 'Hurricane DORIAN', 'right': '24 Aug 2019 - 07 Sep
2019\n160 kt • 910 hPa • 48.9 ACE'}>
```

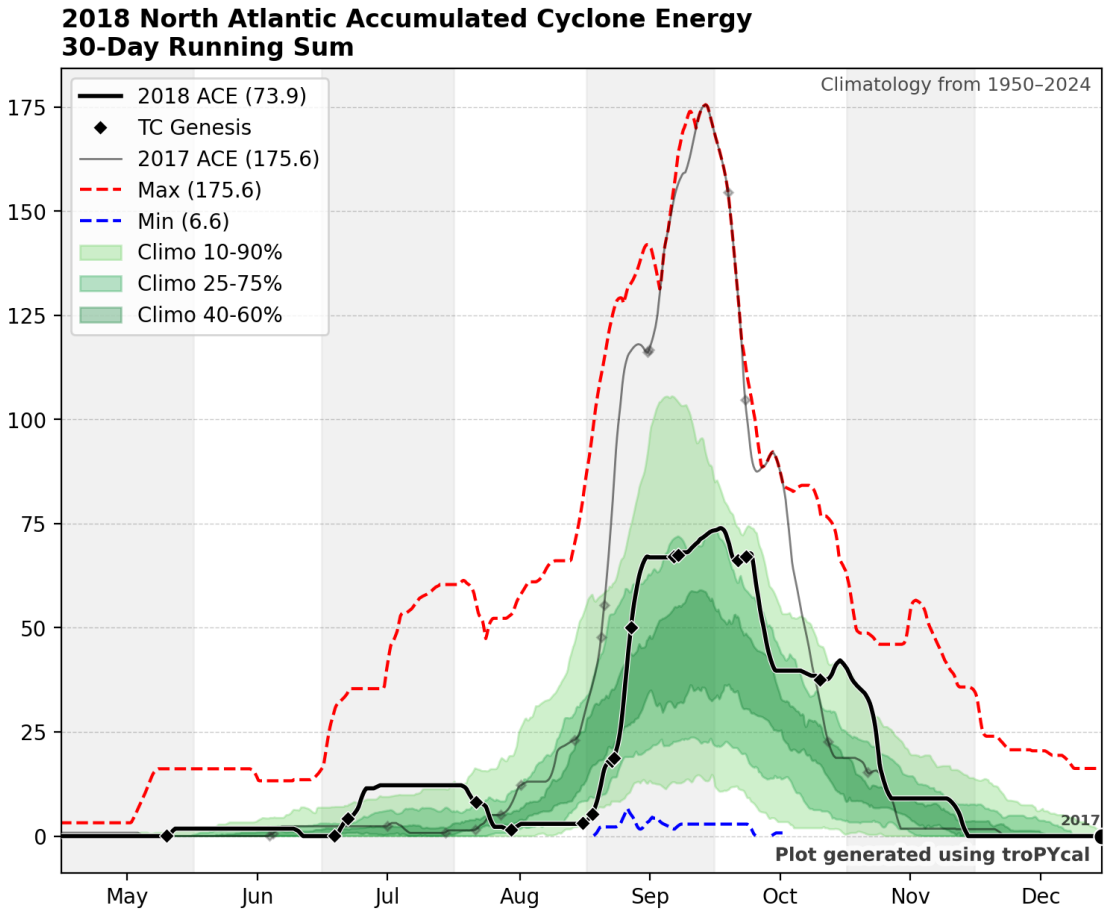


[27]:
```
basin
```

[27]:
```
<tropycal.tracks.Dataset>
Dataset Summary:
    Basin:              north_atlantic
    Source:             hurdat
    Number of storms:   1991
    Maximum wind:       165 knots (Allen 1980)
    Minimum pressure:   882 hPa (Wilma 2005)
    Year range:         1851 - 2024
```

[28]:
```
basin.ace_climo(rolling_sum=30,plot_year=2018,compare_years=2017)
```

[28]:
```
<Axes: title={'left': '2018 North Atlantic Accumulated Cyclone Energy \n30-Day
Running Sum'}>
```



3 Ejemplo en españa

[29]:
```
# pip install seaborn
```

[30]:
```
import pandas as pd, numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

[31]:
```
df = pd.read_csv('weather_features.csv')
df.dtypes
```

[31]:
```
dt_iso          object
city_name       object
temp           float64
temp_min       float64
temp_max       float64
pressure         int64
```

```
humidity                  int64
wind_speed                int64
wind_deg                  int64
rain_1h                 float64
rain_3h                 float64
snow_3h                 float64
clouds_all                int64
weather_id                int64
weather_main             object
weather_description      object
weather_icon             object
dtype: object
```

[32]:
```
df.isnull().sum()
```

[32]:
```
dt_iso                  0
city_name               0
temp                    0
temp_min                0
temp_max                0
pressure                0
humidity                0
wind_speed              0
wind_deg                0
rain_1h                 0
rain_3h                 0
snow_3h                 0
clouds_all              0
weather_id              0
weather_main            0
weather_description     0
weather_icon            0
dtype: int64
```

[33]:
```
df.city_name.value_counts()
```

[33]:
```
city_name
Madrid        36267
Bilbao        35951
Seville       35557
 Barcelona    35476
Valencia      35145
Name: count, dtype: int64
```

[34]:
```
df['city_name'] = df['city_name'].str.replace('Seville', 'Sevilla')
df
```

[34]:

	dt_iso	city_name	temp	temp_min	temp_max	\
0	2015-01-01 00:00:00+01:00	Valencia	270.475	270.475	270.475	
1	2015-01-01 01:00:00+01:00	Valencia	270.475	270.475	270.475	
2	2015-01-01 02:00:00+01:00	Valencia	269.686	269.686	269.686	
3	2015-01-01 03:00:00+01:00	Valencia	269.686	269.686	269.686	
4	2015-01-01 04:00:00+01:00	Valencia	269.686	269.686	269.686	
…	…	…	…	…	…	
178391	2018-12-31 19:00:00+01:00	Sevilla	287.760	287.150	288.150	
178392	2018-12-31 20:00:00+01:00	Sevilla	285.760	285.150	286.150	
178393	2018-12-31 21:00:00+01:00	Sevilla	285.150	285.150	285.150	
178394	2018-12-31 22:00:00+01:00	Sevilla	284.150	284.150	284.150	
178395	2018-12-31 23:00:00+01:00	Sevilla	283.970	282.150	285.150	

	pressure	humidity	wind_speed	wind_deg	rain_1h	rain_3h	snow_3h	\
0	1001	77	1	62	0.0	0.0	0.0	
1	1001	77	1	62	0.0	0.0	0.0	
2	1002	78	0	23	0.0	0.0	0.0	
3	1002	78	0	23	0.0	0.0	0.0	
4	1002	78	0	23	0.0	0.0	0.0	
…	…	…	…	…	…	…	…	
178391	1028	54	3	30	0.0	0.0	0.0	
178392	1029	62	3	30	0.0	0.0	0.0	
178393	1028	58	4	50	0.0	0.0	0.0	
178394	1029	57	4	60	0.0	0.0	0.0	
178395	1029	70	3	50	0.0	0.0	0.0	

	clouds_all	weather_id	weather_main	weather_description	weather_icon
0	0	800	clear	sky is clear	01n
1	0	800	clear	sky is clear	01n
2	0	800	clear	sky is clear	01n
3	0	800	clear	sky is clear	01n
4	0	800	clear	sky is clear	01n
…	…	…	…	…	…
178391	0	800	clear	sky is clear	01n
178392	0	800	clear	sky is clear	01n
178393	0	800	clear	sky is clear	01n
178394	0	800	clear	sky is clear	01n
178395	0	800	clear	sky is clear	01n

[178396 rows x 17 columns]

[35]:
```
df['city_name'] = df['city_name'].str.replace(' Barcelona', 'Barcelona')
df
```

[35]:

	dt_iso	city_name	temp	temp_min	temp_max	\
0	2015-01-01 00:00:00+01:00	Valencia	270.475	270.475	270.475	
1	2015-01-01 01:00:00+01:00	Valencia	270.475	270.475	270.475	

```
2       2015-01-01 02:00:00+01:00  Valencia  269.686   269.686   269.686
3       2015-01-01 03:00:00+01:00  Valencia  269.686   269.686   269.686
4       2015-01-01 04:00:00+01:00  Valencia  269.686   269.686   269.686
…                             …         …       …         …         …
178391  2018-12-31 19:00:00+01:00   Sevilla  287.760   287.150   288.150
178392  2018-12-31 20:00:00+01:00   Sevilla  285.760   285.150   286.150
178393  2018-12-31 21:00:00+01:00   Sevilla  285.150   285.150   285.150
178394  2018-12-31 22:00:00+01:00   Sevilla  284.150   284.150   284.150
178395  2018-12-31 23:00:00+01:00   Sevilla  283.970   282.150   285.150

        pressure  humidity  wind_speed  wind_deg  rain_1h  rain_3h  snow_3h  \
0           1001        77           1        62      0.0      0.0      0.0
1           1001        77           1        62      0.0      0.0      0.0
2           1002        78           0        23      0.0      0.0      0.0
3           1002        78           0        23      0.0      0.0      0.0
4           1002        78           0        23      0.0      0.0      0.0
…            …         …           …         …       …        …        …
178391      1028        54           3        30      0.0      0.0      0.0
178392      1029        62           3        30      0.0      0.0      0.0
178393      1028        58           4        50      0.0      0.0      0.0
178394      1029        57           4        60      0.0      0.0      0.0
178395      1029        70           3        50      0.0      0.0      0.0

        clouds_all  weather_id weather_main weather_description weather_icon
0                0         800        clear        sky is clear          01n
1                0         800        clear        sky is clear          01n
2                0         800        clear        sky is clear          01n
3                0         800        clear        sky is clear          01n
4                0         800        clear        sky is clear          01n
…              …         …           …                 …           …
178391           0         800        clear        sky is clear          01n
178392           0         800        clear        sky is clear          01n
178393           0         800        clear        sky is clear          01n
178394           0         800        clear        sky is clear          01n
178395           0         800        clear        sky is clear          01n

[178396 rows x 17 columns]
```

[36]:
```
df.city_name.value_counts()
```

[36]:
```
city_name
Madrid       36267
Bilbao       35951
Sevilla      35557
Barcelona    35476
Valencia     35145
Name: count, dtype: int64
```

[37]:
```
# pip install geopy
```

[38]:
```
from time import sleep
from geopy.geocoders import Nominatim
import numpy as np
import pandas as pd



def get_coords(neight):
    geolocator = Nominatim(user_agent="Spain")
    address = f'{neight}, España'
    location = geolocator.geocode(address)
    # print(location.address)
    if location is not None:
        latitude = location.latitude
        longitude = location.longitude
    else:
        latitude = longitude = np.NaN
    sleep(2)
    return pd.Series({"Latitud": latitude, "Longitude": longitude})




df_geo = pd.DataFrame({"city_name": ("Madrid", "Bilbao",
                                     "Sevilla", "Barcelona",
                                     "Valencia")})

lldata = pd.concat([df_geo, df_geo.city_name.apply(get_coords)], axis=1)
lldata
```

[38]:

	city_name	Latitud	Longitude
0	Madrid	40.416705	-3.703582
1	Bilbao	43.263002	-2.935004
2	Sevilla	37.388630	-5.995340
3	Barcelona	41.382580	2.177073
4	Valencia	39.469707	-0.376335

[39]:
```
def add_geoposition(row):
    for index, rows in lldata.iterrows():
        if row['city_name'] == rows['city_name']:
            return rows['Latitud'], rows['Longitude']
```

```
df[['Latitud', 'Longitud']] = df.apply(add_geoposition, axis=1,␣
 ↪result_type="expand")
df
```

[39]:

```
                        dt_iso city_name     temp  temp_min  temp_max  \
0       2015-01-01 00:00:00+01:00  Valencia  270.475   270.475   270.475
1       2015-01-01 01:00:00+01:00  Valencia  270.475   270.475   270.475
2       2015-01-01 02:00:00+01:00  Valencia  269.686   269.686   269.686
3       2015-01-01 03:00:00+01:00  Valencia  269.686   269.686   269.686
4       2015-01-01 04:00:00+01:00  Valencia  269.686   269.686   269.686
…                             …         …       …         …         …
178391  2018-12-31 19:00:00+01:00   Sevilla  287.760   287.150   288.150
178392  2018-12-31 20:00:00+01:00   Sevilla  285.760   285.150   286.150
178393  2018-12-31 21:00:00+01:00   Sevilla  285.150   285.150   285.150
178394  2018-12-31 22:00:00+01:00   Sevilla  284.150   284.150   284.150
178395  2018-12-31 23:00:00+01:00   Sevilla  283.970   282.150   285.150

        pressure  humidity  wind_speed  wind_deg  rain_1h  rain_3h  snow_3h  \
0           1001        77           1        62      0.0      0.0      0.0
1           1001        77           1        62      0.0      0.0      0.0
2           1002        78           0        23      0.0      0.0      0.0
3           1002        78           0        23      0.0      0.0      0.0
4           1002        78           0        23      0.0      0.0      0.0
…              …         …           …         …        …        …        …
178391      1028        54           3        30      0.0      0.0      0.0
178392      1029        62           3        30      0.0      0.0      0.0
178393      1028        58           4        50      0.0      0.0      0.0
178394      1029        57           4        60      0.0      0.0      0.0
178395      1029        70           3        50      0.0      0.0      0.0

        clouds_all  weather_id weather_main weather_description weather_icon  \
0                0         800        clear        sky is clear          01n
1                0         800        clear        sky is clear          01n
2                0         800        clear        sky is clear          01n
3                0         800        clear        sky is clear          01n
4                0         800        clear        sky is clear          01n
…              …         …           …                 …           …
178391           0         800        clear        sky is clear          01n
178392           0         800        clear        sky is clear          01n
178393           0         800        clear        sky is clear          01n
178394           0         800        clear        sky is clear          01n
178395           0         800        clear        sky is clear          01n

          Latitud  Longitud
0       39.469707 -0.376335
1       39.469707 -0.376335
2       39.469707 -0.376335
```

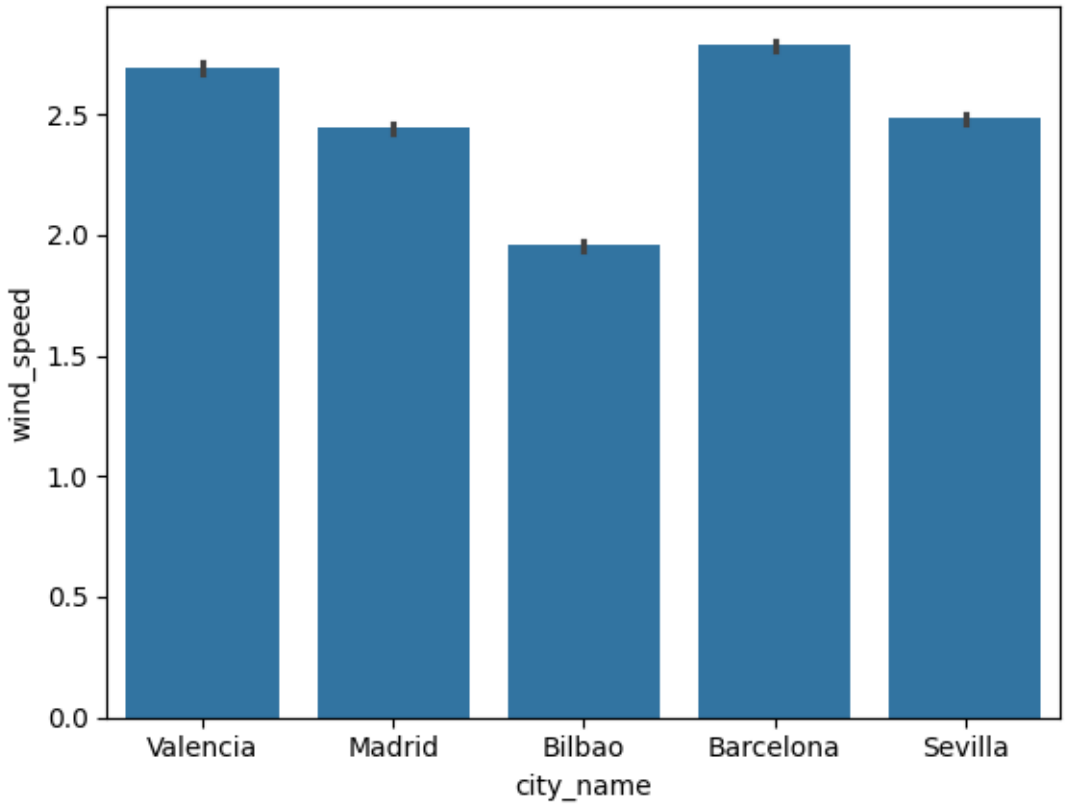
```
3       39.469707 -0.376335
4       39.469707 -0.376335
…            …        …
178391  37.388630 -5.995340
178392  37.388630 -5.995340
178393  37.388630 -5.995340
178394  37.388630 -5.995340
178395  37.388630 -5.995340

[178396 rows x 19 columns]
```

[40]:
```
df.dt_iso = pd.to_datetime(df.dt_iso)
df['date'] = [d.date() for d in df['dt_iso']]
df['time'] = [d.time() for d in df['dt_iso']]
df['year'] = pd.DatetimeIndex(df['date']).year
df['month'] = pd.DatetimeIndex(df['date']).month
df
```

[40]:
```
                        dt_iso city_name     temp  temp_min  temp_max  \
0       2015-01-01 00:00:00+01:00  Valencia  270.475   270.475   270.475
1       2015-01-01 01:00:00+01:00  Valencia  270.475   270.475   270.475
2       2015-01-01 02:00:00+01:00  Valencia  269.686   269.686   269.686
3       2015-01-01 03:00:00+01:00  Valencia  269.686   269.686   269.686
4       2015-01-01 04:00:00+01:00  Valencia  269.686   269.686   269.686
…                             …         …       …         …         …
178391  2018-12-31 19:00:00+01:00   Sevilla  287.760   287.150   288.150
178392  2018-12-31 20:00:00+01:00   Sevilla  285.760   285.150   286.150
178393  2018-12-31 21:00:00+01:00   Sevilla  285.150   285.150   285.150
178394  2018-12-31 22:00:00+01:00   Sevilla  284.150   284.150   284.150
178395  2018-12-31 23:00:00+01:00   Sevilla  283.970   282.150   285.150

        pressure  humidity  wind_speed  wind_deg  rain_1h  …  weather_id  \
0           1001        77           1        62      0.0  …         800
1           1001        77           1        62      0.0  …         800
2           1002        78           0        23      0.0  …         800
3           1002        78           0        23      0.0  …         800
4           1002        78           0        23      0.0  …         800
…              …         …           …         …        …  …           …
178391      1028        54           3        30      0.0  …         800
178392      1029        62           3        30      0.0  …         800
178393      1028        58           4        50      0.0  …         800
178394      1029        57           4        60      0.0  …         800
178395      1029        70           3        50      0.0  …         800

        weather_main  weather_description  weather_icon    Latitud  Longitud  \
0              clear         sky is clear           01n  39.469707 -0.376335
1              clear         sky is clear           01n  39.469707 -0.376335
```

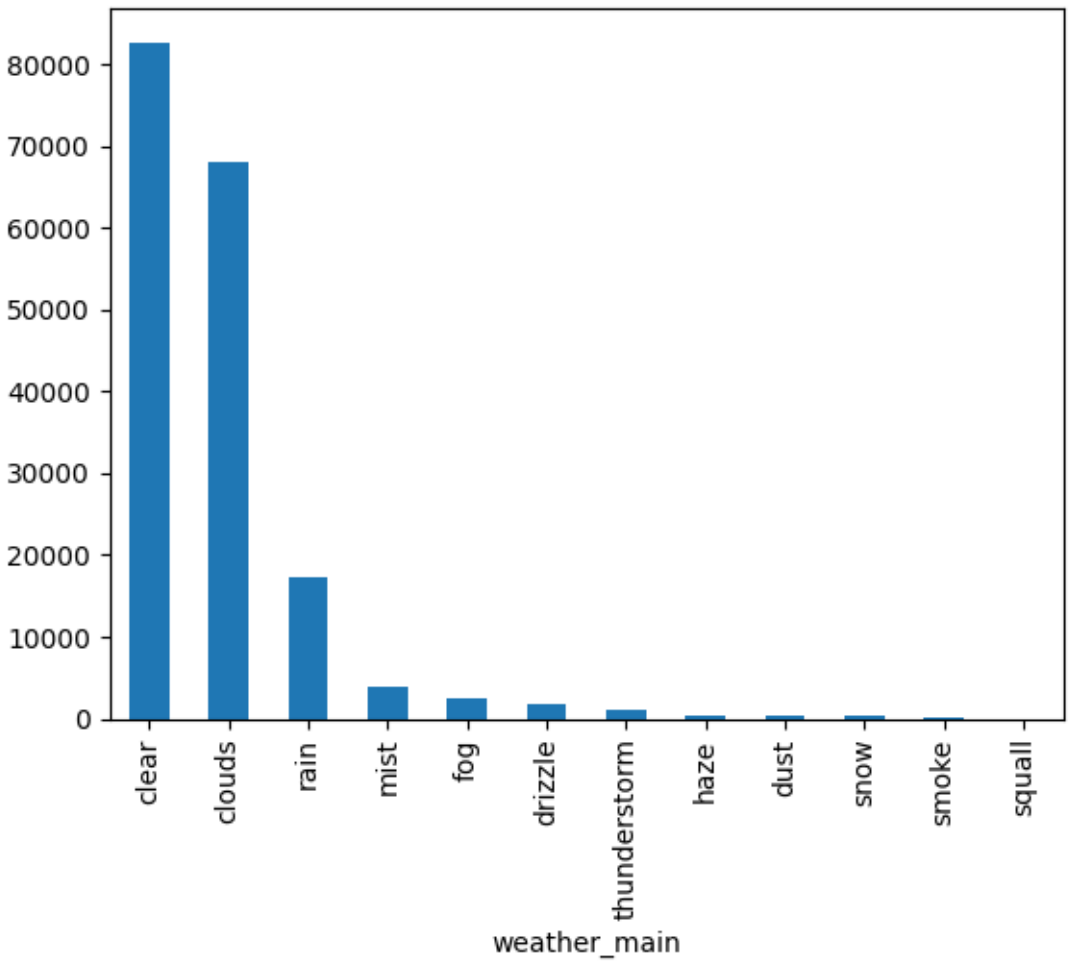
```
2              clear         sky is clear           01n  39.469707 -0.376335
3              clear         sky is clear           01n  39.469707 -0.376335
4              clear         sky is clear           01n  39.469707 -0.376335
…                …                  …          …      …         …
178391         clear         sky is clear           01n  37.388630 -5.995340
178392         clear         sky is clear           01n  37.388630 -5.995340
178393         clear         sky is clear           01n  37.388630 -5.995340
178394         clear         sky is clear           01n  37.388630 -5.995340
178395         clear         sky is clear           01n  37.388630 -5.995340

                date      time  year month
0         2015-01-01  00:00:00  2015     1
1         2015-01-01  01:00:00  2015     1
2         2015-01-01  02:00:00  2015     1
3         2015-01-01  03:00:00  2015     1
4         2015-01-01  04:00:00  2015     1
…                …         …     …    …
178391    2018-12-31  19:00:00  2018    12
178392    2018-12-31  20:00:00  2018    12
178393    2018-12-31  21:00:00  2018    12
178394    2018-12-31  22:00:00  2018    12
178395    2018-12-31  23:00:00  2018    12

[178396 rows x 23 columns]
```

[41]:
```
from numpy import mean

g = sns.barplot(x='city_name', y='wind_speed', data = df, estimator=mean)
```



[42]:
```
df['weather_main'].value_counts().plot(kind='bar')
```

[42]:
```
<Axes: xlabel='weather_main'>
```



[43]:
```
df[df.humidity == df.humidity.max()]
```

[43]:

	dt_iso	city_name	temp	temp_min	temp_max	\
1887	2015-03-20 14:00:00+01:00	Valencia	283.884	283.884	283.884	
1888	2015-03-20 15:00:00+01:00	Valencia	283.884	283.884	283.884	
1889	2015-03-20 16:00:00+01:00	Valencia	283.746	283.746	283.746	
1890	2015-03-20 17:00:00+01:00	Valencia	283.746	283.746	283.746	
1893	2015-03-20 20:00:00+01:00	Valencia	284.186	284.186	284.186	
…	…	…	…	…	…	
178283	2018-12-27 07:00:00+01:00	Sevilla	277.330	276.150	279.150	
178285	2018-12-27 09:00:00+01:00	Sevilla	278.050	276.150	279.150	
178305	2018-12-28 05:00:00+01:00	Sevilla	277.330	276.150	279.150	
178306	2018-12-28 06:00:00+01:00	Sevilla	276.940	276.150	278.150	
178307	2018-12-28 07:00:00+01:00	Sevilla	276.940	276.150	278.150	

	pressure	humidity	wind_speed	wind_deg	rain_1h	…	weather_id	\

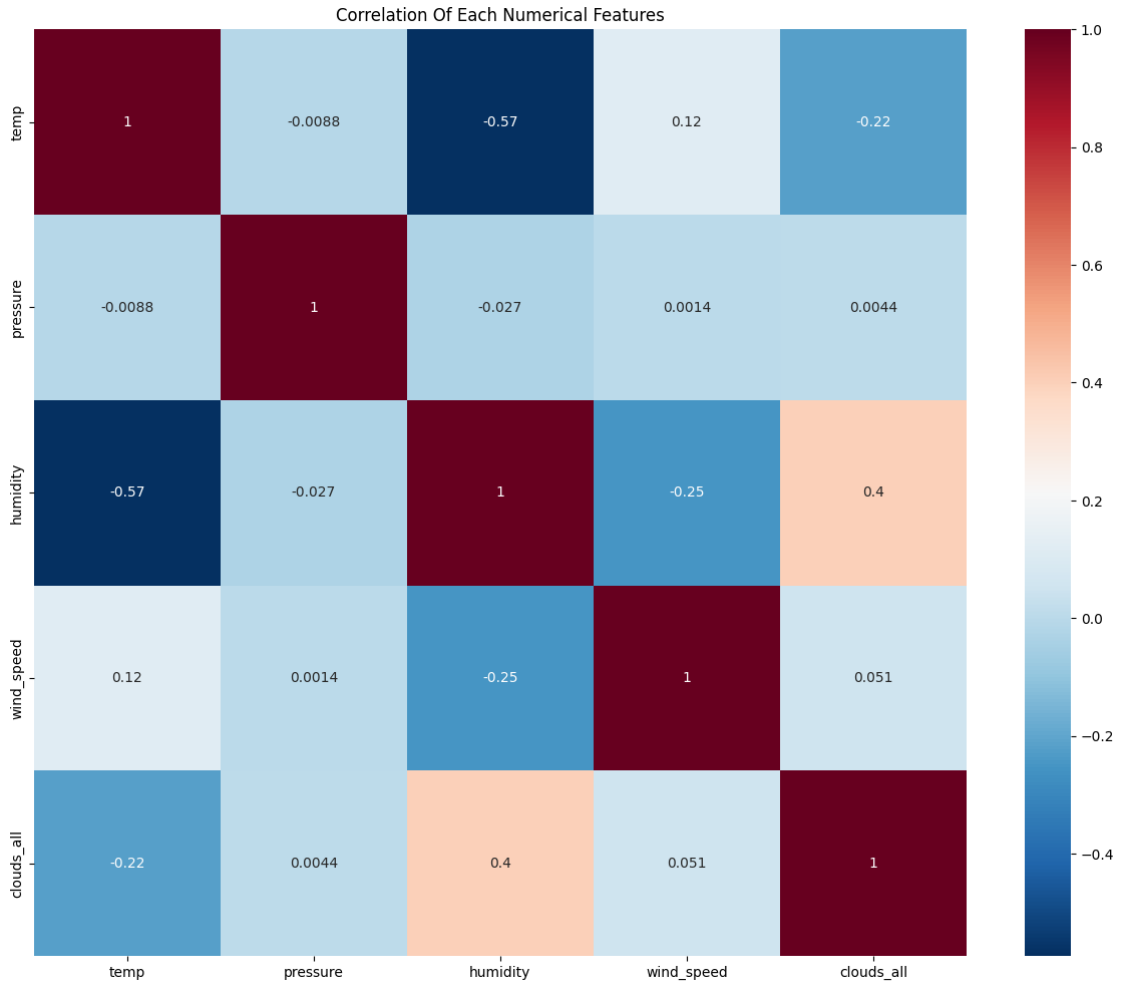
```
1887         979       100           8        69      0.9  …         501
1888         979       100           8        69      0.9  …         501
1889         978       100           8        54      0.9  …         501
1890         978       100           8        54      0.9  …         501
1893         978       100           7        54      0.0  …         804
…            …         …             …        …       …    …         …
178283      1028       100           1        30      0.0  …         802
178285      1028       100           2        30      0.0  …         701
178305      1026       100           1        32      0.0  …         701
178306      1026       100           1       320      0.0  …         701
178307      1026       100           2        35      0.0  …         701

        weather_main  weather_description  weather_icon    Latitud  Longitud  \
1887            rain        moderate rain           10d  39.469707 -0.376335
1888            rain        moderate rain           10d  39.469707 -0.376335
1889            rain        moderate rain           10d  39.469707 -0.376335
1890            rain        moderate rain           10d  39.469707 -0.376335
1893          clouds      overcast clouds           04n  39.469707 -0.376335
…              …                   …           …        …       …
178283        clouds     scattered clouds           03n  37.388630 -5.995340
178285          mist                mist           50n  37.388630 -5.995340
178305          mist                mist           50n  37.388630 -5.995340
178306          mist                mist           50n  37.388630 -5.995340
178307          mist                mist           50n  37.388630 -5.995340

              date      time  year month
1887    2015-03-20  14:00:00  2015     3
1888    2015-03-20  15:00:00  2015     3
1889    2015-03-20  16:00:00  2015     3
1890    2015-03-20  17:00:00  2015     3
1893    2015-03-20  20:00:00  2015     3
…              …       …    …    …
178283  2018-12-27  07:00:00  2018    12
178285  2018-12-27  09:00:00  2018    12
178305  2018-12-28  05:00:00  2018    12
178306  2018-12-28  06:00:00  2018    12
178307  2018-12-28  07:00:00  2018    12

[10128 rows x 23 columns]
```

[44]:
```
num_cols = ['temp', 'pressure', 'humidity', 'wind_speed', 'clouds_all']
corrs = df[num_cols].corr()
corrs
```

[44]:
```
                temp  pressure  humidity  wind_speed  clouds_all
temp        1.000000 -0.008833 -0.573542    0.115307   -0.221331
pressure   -0.008833  1.000000 -0.027458    0.001379    0.004443
```

```
humidity    -0.573542 -0.027458  1.000000   -0.250336    0.400483
wind_speed   0.115307  0.001379 -0.250336    1.000000    0.051049
clouds_all  -0.221331  0.004443  0.400483    0.051049    1.000000
```

[45]:
```
plt.figure(figsize=(15,12))
sns.heatmap(df[num_cols].corr(),annot=True,cmap='RdBu_r')
plt.title('Correlation Of Each Numerical Features')
plt.show()
```



[46]:
```
df
```

[46]:
```
                       dt_iso city_name     temp  temp_min  temp_max  \
0  2015-01-01 00:00:00+01:00  Valencia  270.475   270.475   270.475
1  2015-01-01 01:00:00+01:00  Valencia  270.475   270.475   270.475
2  2015-01-01 02:00:00+01:00  Valencia  269.686   269.686   269.686
3  2015-01-01 03:00:00+01:00  Valencia  269.686   269.686   269.686
```

```
4       2015-01-01 04:00:00+01:00  Valencia  269.686   269.686   269.686
…                             …       …      …       …       …
178391  2018-12-31 19:00:00+01:00   Sevilla  287.760   287.150   288.150
178392  2018-12-31 20:00:00+01:00   Sevilla  285.760   285.150   286.150
178393  2018-12-31 21:00:00+01:00   Sevilla  285.150   285.150   285.150
178394  2018-12-31 22:00:00+01:00   Sevilla  284.150   284.150   284.150
178395  2018-12-31 23:00:00+01:00   Sevilla  283.970   282.150   285.150

        pressure  humidity  wind_speed  wind_deg  rain_1h  …  weather_id  \
0           1001        77           1        62      0.0  …         800
1           1001        77           1        62      0.0  …         800
2           1002        78           0        23      0.0  …         800
3           1002        78           0        23      0.0  …         800
4           1002        78           0        23      0.0  …         800
…              …       …          …       …      …  …          …
178391      1028        54           3        30      0.0  …         800
178392      1029        62           3        30      0.0  …         800
178393      1028        58           4        50      0.0  …         800
178394      1029        57           4        60      0.0  …         800
178395      1029        70           3        50      0.0  …         800

        weather_main  weather_description  weather_icon    Latitud  Longitud  \
0              clear         sky is clear           01n  39.469707 -0.376335
1              clear         sky is clear           01n  39.469707 -0.376335
2              clear         sky is clear           01n  39.469707 -0.376335
3              clear         sky is clear           01n  39.469707 -0.376335
4              clear         sky is clear           01n  39.469707 -0.376335
…              …                  …           …        …       …
178391         clear         sky is clear           01n  37.388630 -5.995340
178392         clear         sky is clear           01n  37.388630 -5.995340
178393         clear         sky is clear           01n  37.388630 -5.995340
178394         clear         sky is clear           01n  37.388630 -5.995340
178395         clear         sky is clear           01n  37.388630 -5.995340

              date      time  year month
0       2015-01-01  00:00:00  2015     1
1       2015-01-01  01:00:00  2015     1
2       2015-01-01  02:00:00  2015     1
3       2015-01-01  03:00:00  2015     1
4       2015-01-01  04:00:00  2015     1
…              …       …   …   …
178391  2018-12-31  19:00:00  2018    12
178392  2018-12-31  20:00:00  2018    12
178393  2018-12-31  21:00:00  2018    12
178394  2018-12-31  22:00:00  2018    12
178395  2018-12-31  23:00:00  2018    12
```

[178396 rows x 23 columns]

3.1 Plotly

[47]:
```
# pip install plotly
```

[48]:
```
from plotly.offline import init_notebook_mode, iplot
import plotly.express as px

init_notebook_mode(connected=True)


fig = px.scatter_map(df, lat="Latitud", lon="Longitud", hover_name="city_name",
                     hover_data=["humidity"], animation_frame="date",␣
 ↪size='humidity',
                     color_discrete_sequence=["fuchsia"], zoom=3, height=300)
fig.update_layout(map_style="open-street-map")
fig.update_layout(margin={"r":0,"t":0,"l":0,"b":0})

iplot(fig)
```

[49]:
```
df['temp'] = round(df['temp'] / 10, 2)
df['temp_min'] = round(df['temp_min'] / 10, 2)
df['temp_max'] = round(df['temp_max'] / 10, 2)
df
```

[49]:

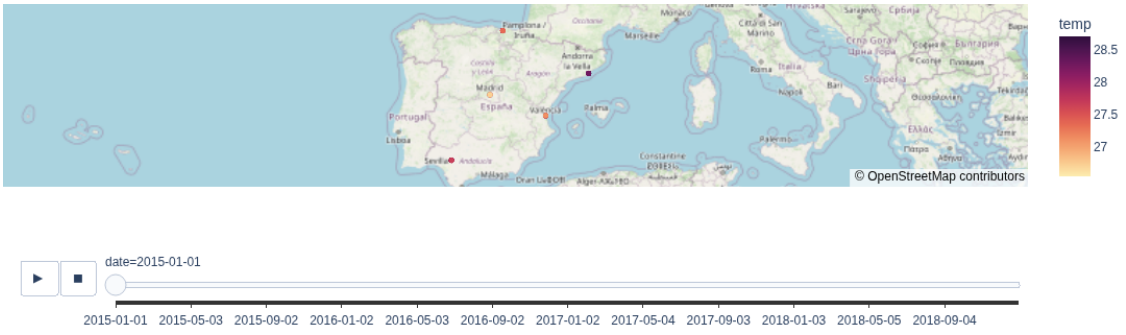
	dt_iso	city_name	temp	temp_min	temp_max	\
0	2015-01-01 00:00:00+01:00	Valencia	27.05	27.05	27.05	
1	2015-01-01 01:00:00+01:00	Valencia	27.05	27.05	27.05	
2	2015-01-01 02:00:00+01:00	Valencia	26.97	26.97	26.97	
3	2015-01-01 03:00:00+01:00	Valencia	26.97	26.97	26.97	
4	2015-01-01 04:00:00+01:00	Valencia	26.97	26.97	26.97	
…	…	…	…	…	…	
178391	2018-12-31 19:00:00+01:00	Sevilla	28.78	28.71	28.82	
178392	2018-12-31 20:00:00+01:00	Sevilla	28.58	28.51	28.62	
178393	2018-12-31 21:00:00+01:00	Sevilla	28.51	28.51	28.51	
178394	2018-12-31 22:00:00+01:00	Sevilla	28.42	28.42	28.42	
178395	2018-12-31 23:00:00+01:00	Sevilla	28.40	28.21	28.51	

	pressure	humidity	wind_speed	wind_deg	rain_1h	…	weather_id	\
0	1001	77	1	62	0.0	…	800	
1	1001	77	1	62	0.0	…	800	
2	1002	78	0	23	0.0	…	800	
3	1002	78	0	23	0.0	…	800	
4	1002	78	0	23	0.0	…	800	
…	…	…	…	…	…	…	…	
178391	1028	54	3	30	0.0	…	800	

```
178392      1029        62           3        30      0.0  …         800
178393      1028        58           4        50      0.0  …         800
178394      1029        57           4        60      0.0  …         800
178395      1029        70           3        50      0.0  …         800

        weather_main  weather_description  weather_icon    Latitud  Longitud  \
0              clear         sky is clear           01n  39.469707 -0.376335
1              clear         sky is clear           01n  39.469707 -0.376335
2              clear         sky is clear           01n  39.469707 -0.376335
3              clear         sky is clear           01n  39.469707 -0.376335
4              clear         sky is clear           01n  39.469707 -0.376335
…                  …                    …           …        …         …
178391         clear         sky is clear           01n  37.388630 -5.995340
178392         clear         sky is clear           01n  37.388630 -5.995340
178393         clear         sky is clear           01n  37.388630 -5.995340
178394         clear         sky is clear           01n  37.388630 -5.995340
178395         clear         sky is clear           01n  37.388630 -5.995340

              date      time  year month
0       2015-01-01  00:00:00  2015     1
1       2015-01-01  01:00:00  2015     1
2       2015-01-01  02:00:00  2015     1
3       2015-01-01  03:00:00  2015     1
4       2015-01-01  04:00:00  2015     1
…              …         …    …    …
178391  2018-12-31  19:00:00  2018    12
178392  2018-12-31  20:00:00  2018    12
178393  2018-12-31  21:00:00  2018    12
178394  2018-12-31  22:00:00  2018    12
178395  2018-12-31  23:00:00  2018    12

[178396 rows x 23 columns]
```

[50]:
```
from plotly.offline import init_notebook_mode, iplot
import plotly.express as px

init_notebook_mode(connected=True)


fig = px.scatter_map(df, lat="Latitud", lon="Longitud", hover_name="city_name",
                     hover_data=["temp"], animation_frame="date", color='temp',
                     color_continuous_scale='matter', # https://plotly.com/
 ↪python/builtin-colorscales/
                     zoom=3.5, height=300)
fig.update_layout(map_style="open-street-map")
fig.update_layout(margin={"r":0,"t":0,"l":0,"b":0})
```

```
iplot(fig)
```



Creado por:

Isabel Maniega