

```
import pandas as pd
import numpy as np
import os
from time import sleep
from bs4 import BeautifulSoup
import pickle

In [6]:

In [7]: Francisco Silveira, pases, Calle de Béjar, Madrid Moderno, Guindalera, Salamanca, Madrid, Comunidad de Madrid, 28028, España

In [8]:

48.4349564 -3.6745879

In [2]: #Empezamos trabajando la lista de direcciones url de restaurantessacadas del notebook 1_URL_Rest

with open('web_rest.pickle', "rb") as f:
    rest_url = pickle.load(f)

In [3]: len(rest_url)

Out[3]: 9869

In [4]: #copia de como queda
print(rest_url[58])

https://www.tripadvisor.es/Restaurant_Review-g187514-d1208969-Reviews-Ondinas_Do_Mendo-Madrid.html

In [5]: #para probar el codigo
shorter_list=rest_url[9600:9800]
len(shorter_list)

Out[5]: 200

In [6]: %time
Nombre=[]
Score=[]
Num_Reviews=[]
Posicion=[]
Rango_precio=[]
Tipo_Cocina=[]
Awards=[]
Score_Comida=[]
Score_Servicio=[]
Calidad_Precio=[]
Ambiente=[]
Direccion=[]
Horario=[]
Web_TripAdvisor=[]
Ultimo_review=[]
Opciones=[]
Excelente=[]
MuyBueno=[]
Normal=[]
Mal0=[]
Pesimo=[]
Contacto=[]

can
for url in shorter_list:

    print("Faltan (len(shorter_list)- c ) restaurantes")
    c+=1

    browser = helium.start_chrome(url, headless = True)
    soup = BeautifulSoup(browser.page_source, "html.parser")
    sleep(2)

    #web
    df_tripAdvisor.append(url)
    sleep(0.1)

    #Nombre#####
    try:
        nombre = soup.find("h1", class_="HjBfq").text
    except:
        nombre = np.nan
    Nombre.append(nombre)
    sleep(0.1)

    #Score#####
    try:
        score= soup.find("span", class_="ZDEqk").text.replace("\xa0", "")
    except:
        score=np.nan
    Score.append(score)
    sleep(0.1)

    #Numero_Reviews#####
    try:
        num_review= soup.find("span", class_ = "AfQz2").text.split()[0]
    except:
        num_review= np.nan
    Num_Reviews.append(num_review)
    sleep(0.1)

    #Posicion#####
    try:
        posicion= soup.find("div", class_="cNF1b").text.split()[1]
    except:
        posicion=np.nan
    Posicion.append(posicion)
    sleep(0.1)

    #Rango_precio#####
    try:
        rango_precio=soup.find_all("a", class_ = "d1MO3")[0].text
    except:
        rango_precio=np.nan
    Rango_precio.append(rango_precio)
    sleep(0.1)

    #Awards#####
    try:
        awards=soup.find("span", class_="1gtpa").text
    except:
        awards=""
    Awards.append(awards)
    sleep(0.1)

    #Tipo de cocina#####
    try:
        tipo_cocina=[]
        for x in soup.find_all("a", class_="d1MO3"):
            tipo_cocina.append(x.text)
        tipo_cocina=tipo_cocina[1:]
    except:
        tipo_cocina=np.nan
    Tipo_Cocina.append(tipo_cocina)
    sleep(0.1)

    #Score_comida#####
    try:
        punt_comida=soup.find_all('div', class_="DzMcU")[0]
        score_comida=float(punt_comida.find('span', class_="ui_bubble_rating").get('class')[-1].split('_')[-1])/10
    except:
        score_comida=np.nan
    Score_Comida.append(score_comida)
    sleep(0.1)

    #Score_servicio#####
    try:
        punt_servicio=soup.find_all('div', class_="DzMcU")[1]
        score_servicio=float(punt_servicio.find('span', class_="ui_bubble_rating").get('class')[-1].split('_')[-1])/10
    except:
        score_servicio=np.nan
    Score_Servicio.append(score_servicio)
    sleep(0.1)

    #Calidad_Precio#####
    try:
        punt_calidad=soup.find_all('div', class_="DzMcU")[2]
        calidad_precio=float(punt_calidad.find('span', class_="ui_bubble_rating").get('class')[-1].split('_')[-1])/10
    except:
        calidad_precio=np.nan
    Calidad_Precio.append(calidad_precio)
    sleep(0.1)

    #Ambiente#####
    try:
        punt_ambiente=soup.find_all('div', class_="DzMcU")[3]
        ambiente=float(punt_ambiente.find('span', class_="ui_bubble_rating").get('class')[-1].split('_')[-1])/10
    except:
        ambiente=np.nan
    Ambiente.append(ambiente)
    sleep(0.1)

    #Direccion#####
    try:
        direccion_detalle=soup.find_all("a", class_="AYWH")
        direccion=direccion_detalle[1].text.split(",")
    except:
        direccion=np.nan
    Direccion.append(direccion)
    sleep(0.1)

    #Puntuaciones Viajero
    #Excelente
    try:
        excelente= soup.find_all("div", class_ = "ui_checkbox item")[0].text.split("Excelente")[-1]
    except:
        excelente=np.nan
    Excelente.append(excelente)
    sleep(0.1)

    #Muy bueno
    try:
        muy_bueno=soup.find_all("div", class_ = "ui_checkbox item")[1].text.split("Muy Bueno")[-1]
    except:
        muy_bueno=np.nan
    MuyBueno.append(muy_bueno)
    sleep(0.1)

    #Bueno
    try:
        bueno=soup.find_all("div", class_ = "ui_checkbox item")[2].text.split("Normal")[-1]
    except:
        bueno=np.nan
    Normal.append(bueno)
    sleep(0.1)

    #Mal0
    try:
        malo=soup.find_all("div", class_ = "ui_checkbox item")[3].text.split("Mal0")[-1]
    except:
        malo=np.nan
    Mal0.append(malo)
    sleep(0.1)

    #Pesimo
    try:
        pesimo=soup.find_all("div", class_ = "ui_checkbox item")[4].text.split("Pesimo")[-1]
    except:
        pesimo=np.nan
    Pesimo.append(pesimo)
    sleep(0.1)

    #Ultimo_review#####
    try:
        comment=[]
        for i in soup.find_all("div", class_="entry"):
            comment.append(i.find_all("p")[0].text)
        #para quitar los "\n"
        for x in comment:
            cleaned_text = x.replace("\n", "")
        comentario=comment[0]
    except:
        comentario=np.nan
    Ultimo_review.append(comentario)
    sleep(0.1)

    #Opciones#####
    try:
        opc=[]
        for x in soup.find_all("div", class_ = "Sqgk8"):
            opc.append(x.text.replace("\xa0", ""))
        opciones=opc[1:]
    except:
        opciones=np.nan
    Ultimo_review.append(opciones)
    sleep(0.1)

Faltan 200 restaurantes
Faltan 199 restaurantes
Faltan 198 restaurantes
Faltan 197 restaurantes
Faltan 196 restaurantes
Faltan 195 restaurantes
Faltan 194 restaurantes
Faltan 193 restaurantes
Faltan 192 restaurantes
Faltan 191 restaurantes
Faltan 190 restaurantes
Faltan 189 restaurantes
Faltan 188 restaurantes
Faltan 187 restaurantes
Faltan 186 restaurantes
Faltan 185 restaurantes
Faltan 184 restaurantes
Faltan 183 restaurantes
Faltan 182 restaurantes
Faltan 181 restaurantes
Faltan 180 restaurantes
Faltan 179 restaurantes
Faltan 178 restaurantes
Faltan 177 restaurantes
Faltan 176 restaurantes
Faltan 175 restaurantes
Faltan 174 restaurantes
Faltan 173 restaurantes
Faltan 172 restaurantes
Faltan 171 restaurantes
Faltan 170 restaurantes
Faltan 169 restaurantes
Faltan 168 restaurantes
Faltan 167 restaurantes
Faltan 166 restaurantes
Faltan 165 restaurantes
Faltan 164 restaurantes
Faltan 163 restaurantes
Faltan 162 restaurantes
Faltan 161 restaurantes
Faltan 160 restaurantes
Faltan 159 restaurantes
Faltan 158 restaurantes
Faltan 157 restaurantes
Faltan 156 restaurantes
Faltan 155 restaurantes
Faltan 154 restaurantes
Faltan 153 restaurantes
Faltan 152 restaurantes
Faltan 151 restaurantes
Faltan 150 restaurantes
Faltan 149 restaurantes
Faltan 148 restaurantes
Faltan 147 restaurantes
Faltan 146 restaurantes
Faltan 145 restaurantes
Faltan 144 restaurantes
Faltan 143 restaurantes
Faltan 142 restaurantes
Faltan 141 restaurantes
Faltan 140 restaurantes
Faltan 139 restaurantes
Faltan 138 restaurantes
Faltan 137 restaurantes
Faltan 136 restaurantes
Faltan 135 restaurantes
Faltan 134 restaurantes
Faltan 133 restaurantes
Faltan 132 restaurantes
Faltan 131 restaurantes
Faltan 130 restaurantes
Faltan 129 restaurantes
Faltan 128 restaurantes
Faltan 127 restaurantes
Faltan 126 restaurantes
Faltan 125 restaurantes
Faltan 124 restaurantes
Faltan 123 restaurantes
Faltan 122 restaurantes
Faltan 121 restaurantes
Faltan 120 restaurantes
Faltan 119 restaurantes
Faltan 118 restaurantes
Faltan 117 restaurantes
Faltan 116 restaurantes
Faltan 115 restaurantes
Faltan 114 restaurantes
Faltan 113 restaurantes
Faltan 112 restaurantes
Faltan 111 restaurantes
Faltan 110 restaurantes
Faltan 109 restaurantes
Faltan 108 restaurantes
Faltan 107 restaurantes
Faltan 106 restaurantes
Faltan 105 restaurantes
Faltan 104 restaurantes
Faltan 103 restaurantes
Faltan 102 restaurantes
Faltan 101 restaurantes
Faltan 100 restaurantes
Faltan 99 restaurantes
Faltan 98 restaurantes
Faltan 97 restaurantes
Faltan 96 restaurantes
Faltan 95 restaurantes
Faltan 94 restaurantes
Faltan 93 restaurantes
Faltan 92 restaurantes
Faltan 91 restaurantes
Faltan 90 restaurantes
Faltan 89 restaurantes
Faltan 88 restaurantes
Faltan 87 restaurantes
Faltan 86 restaurantes
Faltan 85 restaurantes
Faltan 84 restaurantes
Faltan 83 restaurantes
Faltan 82 restaurantes
Faltan 81 restaurantes
Faltan 80 restaurantes
Faltan 79 restaurantes
Faltan 78 restaurantes
Faltan 77 restaurantes
Faltan 76 restaurantes
Faltan 75 restaurantes
Faltan 74 restaurantes
Faltan 73 restaurantes
Faltan 72 restaurantes
Faltan 71 restaurantes
Faltan 70 restaurantes
Faltan 69 restaurantes
Faltan 68 restaurantes
Faltan 67 restaurantes
Faltan 66 restaurantes
Faltan 65 restaurantes
Faltan 64 restaurantes
Faltan 63 restaurantes
Faltan 62 restaurantes
Faltan 61 restaurantes
Faltan 60 restaurantes
Faltan 59 restaurantes
Faltan 58 restaurantes
Faltan 57 restaurantes
Faltan 56 restaurantes
Faltan 55 restaurantes
Faltan 54 restaurantes
Faltan 53 restaurantes
Faltan 52 restaurantes
Faltan 51 restaurantes
Faltan 50 restaurantes
Faltan 49 restaurantes
Faltan 48 restaurantes
Faltan 47 restaurantes
Faltan 46 restaurantes
Faltan 45 restaurantes
Faltan 44 restaurantes
Faltan 43 restaurantes
Faltan 42 restaurantes
Faltan 41 restaurantes
Faltan 40 restaurantes
Faltan 39 restaurantes
Faltan 38 restaurantes
Faltan 37 restaurantes
Faltan 36 restaurantes
Faltan 35 restaurantes
Faltan 34 restaurantes
Faltan 33 restaurantes
Faltan 32 restaurantes
Faltan 31 restaurantes
Faltan 30 restaurantes
Faltan 29 restaurantes
Faltan 28 restaurantes
Faltan 27 restaurantes
Faltan 26 restaurantes
Faltan 25 restaurantes
Faltan 24 restaurantes
Faltan 23 restaurantes
Faltan 22 restaurantes
Faltan 21 restaurantes
Faltan 19 restaurantes
Faltan 18 restaurantes
Faltan 17 restaurantes
Faltan 16 restaurantes
Faltan 15 restaurantes
Faltan 14 restaurantes
Faltan 13 restaurantes
Faltan 12 restaurantes
Faltan 11 restaurantes
Faltan 9 restaurantes
Faltan 8 restaurantes
Faltan 7 restaurantes
Faltan 6 restaurantes
Faltan 5 restaurantes
Faltan 4 restaurantes
Faltan 3 restaurantes
Faltan 2 restaurantes
Faltan 1 restaurantes
CPU times: total: 14.6 s
Wall time: 29min 2s
```

```
In [7]: df_tripAdvisor_Rest_Madrid=pd.DataFrame()
```

```
In [8]: df_tripAdvisor_Rest_Madrid["Nombre"]=Nombre
df_tripAdvisor_Rest_Madrid["Score"]=Score
df_tripAdvisor_Rest_Madrid["Excelente"]=Excelente
df_tripAdvisor_Rest_Madrid["MuyBueno"]=MuyBueno
df_tripAdvisor_Rest_Madrid["Normal"]=Normal
df_tripAdvisor_Rest_Madrid["Mal0"]=Mal0
df_tripAdvisor_Rest_Madrid["Pesimo"]=Pesimo
df_tripAdvisor_Rest_Madrid["Num_Reviews"]=Num_Reviews
df_tripAdvisor_Rest_Madrid["Posicion"]=Posicion
df_tripAdvisor_Rest_Madrid["Rango_precio"]=Rango_precio
df_tripAdvisor_Rest_Madrid["Tipo_Cocina"]=Tipo_Cocina
df_tripAdvisor_Rest_Madrid["Awards"]=Awards
df_tripAdvisor_Rest_Madrid["Score_Comida"]=Score_Comida
df_tripAdvisor_Rest_Madrid["Score_Servicio"]=Score_Servicio
df_tripAdvisor_Rest_Madrid["Calidad_Precio"]=Calidad_Precio
df_tripAdvisor_Rest_Madrid["Ambiente"]=Ambiente
df_tripAdvisor_Rest_Madrid["Direccion"]=Direccion
df_tripAdvisor_Rest_Madrid["Web_TripAdvisor"]=Web_TripAdvisor
```

```
In [9]: df_tripAdvisor_Rest_Madrid.head()
```

	Nombre	Score	Excelente	MuyBueno	Normal	Mal0	Pesimo	Num_Reviews	Posicion	Rango precio	Tipo_Cocina	Awards	Score_Comida	Score_Servicio	Calidad_Precio	Ambiente	Direccion	Web_TripAdvisor
0	La Manduga de Andorra	4.5	331	186	41	12	9	651	301	€€€€	[Mediterránea, Europa, España]	-	4.5	4.5	3.5	4.0	[Calle Sagasta, 14 BAJO, 28004 Madrid España]	https://www.tripadvisor.es/Restaurant_Review-g-28004-Ma-9-
1	Bule Bule	4.5	2	0	1	0	1	8	353	€€-€€€	[Internacional, Mediterránea, Rusia]	-	3.5	3.5	3.5	NaN	[Calle Marqués de Valtierra, 6, 28004 Ma-	https://www.tripadvisor.es/Restaurant_Review-g-28004-Ma-9-
2	Restaurante La Mison	4.0	106	146	45	15	19	342	1175	€€-€€€	[Mediterránea, Europa, España]	-	4.0	4.0	3.5	4.0	[C. José Silva, 22, 28043 Madrid España]	https://www.tripadvisor.es/Restaurant_Review-g-28043-Ma-9-
3	Barhop Madrid	4.5	7	1	0	0	1	9	244	€	[Americana]	-	NaN	NaN	NaN	NaN	[Calle Apodaca 11, 28004 Madrid España]	https://www.tripadvisor.es/Restaurant_Review-g-28004-Ma-9-
4	Plan Based - Madrid	5.0	3	0	0	0	0	3	403	Saludable	[]	-	NaN	NaN	NaN	NaN	[Calle de Luisa 28008 Madrid E.	https://www.tripadvisor.es/Restaurant_Review-g-28008-Ma-9-

```
In [10]: archivo= open("Trip_Rest_9600_9800.pk1", "wb")
pickle.dump(df_tripAdvisor_Rest_Madrid, archivo)
archivo.close()
```

```
In [ ]: #####
```

```
In [ ]: 
```