1 Capstone Project

1.1 Notebook 1: Data Collection

- Tausif Khan
- https://github.com/tausifKhan21
- tausifkit@gmail.com (mailto:tausifkit@gmail.com)

Table of Contents

2 Project Overview

The COVID-19 pandemic has severely affected the travel industry. International travel has been impacted, and in turn travel companies and travel websites have lost much of their engagement.

However, with the development of new vaccines for the virus, there is hope on the horizon for international travel and a time where life is somewhat back to normal. In order to increase engagement in the travel industry and increase excitement about travel opportunities, the Destination Dictionary was born!

The Destination Dictionary is a data product that allows future travelers to get a prediction for their perfect destination with the input of just a few words. Trained on over 28,000 unique text data points, the Destination Dictionary is able to predict a destination from 12 different popular cities with 81% accuracy based on text input of activities you want to do while on vacation.

2.0.1 Methodology & Data Used

This project utilized data from 12 top cities from TripAdvisor's list of Traveler's Choice destinations for Popular World Destinations 2020, which can be found via this link (https://www.tripadvisor.com/TravelersChoice-Destinations). The dataset was compiled by scraping the titles from Tripadvisor 'attractions' for each of the 12 cities. The final dataset included over 28,000 unique text values.

In [1]:
```python
# Import Statements

import requests
from bs4 import BeautifulSoup

import pandas as pd
```
executed in 662ms, finished 08:09:07 2021-01-26

2.0.2 Grab the attraction titles for one city

In [3]:
```python
# Use url for London from TripAdvisor
html_london = requests.get('https://www.tripadvisor.com/Attraction_Proc
soup = BeautifulSoup(html_london.content, 'html.parser')
```

In [4]:
```python
# Print the html code nicely
soup.prettify
```

Out[4]: <bound method Tag.prettify of <!DOCTYPE html>

<html lang="en" xmlns:fb="http://www.facebook.com/2008/fbml">
<head>
<meta content="text/html; charset=utf-8" http-equiv="content-type"/>
<link data-rup="long_lived_global_legacy" href="https://static.tacdn.com/
css2/build/concat/long_lived_global_legacy-v24030609022a.css" rel="styles
heet" type="text/css"/>
<link href="https://static.tacdn.com/favicon.ico?v2" id="favicon" rel="ic
on" type="image/x-icon"/>
<link as="font" crossorigin="" href="https://static.tacdn.com/css2/webfon
ts/TripAdvisor/TripAdvisor_Regular.woff2?v004.021" rel="preload" type="fo
nt/woff2"/>
<link crossorigin="" href="https://static.tacdn.com/css2/webfonts/TripSan
s/TripSans.css?v1.002" rel="stylesheet"/>
<link color="#000000" href="https://static.tacdn.com/img2/brand_refresh/a
pplication_icons/mask-icon.svg" rel="mask-icon" sizes="any">
<script data-rup="global_error" type="text/javascript">!function(){functi
on e(e){return"object"==typeof e&&"WARN"===e.level?"WARN":"ERROR"}functio

In [31]:
```python
1  # Grab the container with the attraction name inside
2  container = soup.find('div', id="ATTRACTION_PRODUCTS_LIST")
```

In [62]:
```python
1  # Find H2 elements within the container
2  attractions = container.find_all('h2')
3
4  attractions
```

Out[62]: [<h2><a href="/AttractionProductReview-g186338-d11468721-Jack_the_Ripper_
Tour_with_Ripper_Vision_in_London-London_England.html" onclick="(ta.prwid
gets.getjs(this,'handlers')).trackCancelLabel('23256P1','','true');(ta.pr
widgets.getjs(this,'handlers')).trackKidPricing('23256P1','','');(ta.prwi
dgets.getjs(this,'handlers')).trackLangSupport('23256P1','','true');(ta.p
rwidgets.getjs(this,'handlers')).trackAndOpenUrl(event, 'attraction_produ
ct_card_click', '/AttractionProductReview-g186338-d11468721-Jack_the_Ripp
er_Tour_with_Ripper_Vision_in_London-London_England.html', 'title', 'card
_0', '37352');" title="Jack the Ripper Tour with 'Ripper-Vision' in Londo
n">Jack the Ripper Tour with 'Ripper-Vision' in London</h2>,
 <h2><a href="/AttractionProductReview-g186338-d12078322-SEA_LIFE_London_
Aquarium_Admission_Ticket-London_England.html" onclick="(ta.prwidgets.get
js(this,'handlers')).trackCancelLabel('5097LONAQUA','','');(ta.prwidgets.
getjs(this,'handlers')).trackKidPricing('5097LONAQUA','','true');(ta.prwi
dgets.getjs(this,'handlers')).trackLangSupport('5097LONAQUA','','');(ta.p
rwidgets.getjs(this,'handlers')).trackAndOpenUrl(event, 'attraction_produ
ct_card_click', '/AttractionProductReview-g186338-d12078322-SEA_LIFE_Lond
on_Aquarium_Admission_Ticket-London_England.html', 'title', 'card_1', '69
25');" title="SEA LIFE London Aquarium Admission Ticket">SEA LIFE London

In [66]:
```python
1  # Find the a element for one h2 element
2  attractions[0].find('a')
```

Out[66]: <a href="/AttractionProductReview-g186338-d11468721-Jack_the_Ripper_Tour_
with_Ripper_Vision_in_London-London_England.html" onclick="(ta.prwidgets.
getjs(this,'handlers')).trackCancelLabel('23256P1','','true');(ta.prwidge
ts.getjs(this,'handlers')).trackKidPricing('23256P1','','');(ta.prwidget
s.getjs(this,'handlers')).trackLangSupport('23256P1','','true');(ta.prwid
gets.getjs(this,'handlers')).trackAndOpenUrl(event, 'attraction_product_c
ard_click', '/AttractionProductReview-g186338-d11468721-Jack_the_Ripper_T
our_with_Ripper_Vision_in_London-London_England.html', 'title', 'card_0',
'37352');" title="Jack the Ripper Tour with 'Ripper-Vision' in London">Ja
ck the Ripper Tour with 'Ripper-Vision' in London

In [70]:
```python
1  # Grab the attraction title from this a element
2  attractions[0].find('a').attrs['title']
```

Out[70]: "Jack the Ripper Tour with 'Ripper-Vision' in London"

In [82]:
```python
# Grab all titles with a for loop and store in a list

final_attractions = []
for attraction in attractions:
    att = attraction.find('a')
    if hasattr(att, 'title'):
        final_attractions.append(att.attrs['title'])

final_attractions
```

Out[82]: ["Jack the Ripper Tour with 'Ripper-Vision' in London",
 'SEA LIFE London Aquarium Admission Ticket',
 'Big Bus London Hop-On Hop-Off Tour and River Cruise with Optional Londo
n Eye',
 'Alcotraz: Prison Cocktail Experience (Shoreditch)',
 'Oxford Sightseeing River Cruise Along The University Regatta Course',
 'Phantom of the Opera',
 'Warner Bros. Studio: The Making of Harry Potter with Luxury Round-Trip
Transport from London',
 'Private DJ Class in London',
 'The Blood and Tears Walk: Serial Killers and London Horror',
 'Ghost Bus Tour of London',
 '2-Hour Cycle Tour of Oxford',
 'Emirates Airline Cable Car and One Way Uber Boat by Thames Clippers jou
rney',
 'Rock Cab Tours Presents: The Music Legends Tour of London',
 'Royal High Tea At Kensington Palace',
 'High-Speed Thames River RIB Cruise in London',
 'Charming London Private Tour: Little Venice Canals & Gardens Exploratio

▼ 2.0.3 Make a function to grab attraction titles for any city

In [86]:
```python
def get_attractions(html_path):
    """
    input html path from TripAdvisor as a string
    returns list of attraction names from that page
    """
    html = requests.get(html_path)
    soup = BeautifulSoup(html.content, 'html.parser')
    container = soup.find('div', id="ATTRACTION_PRODUCTS_LIST")
    attractions = container.findAll('h2')
    final_attractions = []
    for attraction in attractions:
        att = attraction.find('a')
        if hasattr(att, 'title'):
            final_attractions.append(att.attrs['title'])
    return final_attractions
```

In [88]:
```python
1  london_attractions = get_attractions('https://www.tripadvisor.com/Attra
2
3  london_attractions
```

Out[88]: ['SEA LIFE London Aquarium Admission Ticket',
 'Big Bus London Hop-On Hop-Off Tour and River Cruise with Optional Londo
n Eye',
 'Alcotraz: Prison Cocktail Experience (Shoreditch)',
 'Oxford Sightseeing River Cruise Along The University Regatta Course',
 'Phantom of the Opera',
 'Warner Bros. Studio: The Making of Harry Potter with Luxury Round-Trip
Transport from London',
 'Private DJ Class in London',
 'The Blood and Tears Walk: Serial Killers and London Horror',
 'Ghost Bus Tour of London',
 '2-Hour Cycle Tour of Oxford',
 'Emirates Airline Cable Car and One Way Uber Boat by Thames Clippers jou
rney',
 'Rock Cab Tours Presents: The Music Legends Tour of London',
 'Royal High Tea At Kensington Palace',
 'High-Speed Thames River RIB Cruise in London',
 'Charming London Private Tour: Little Venice Canals & Gardens Exploratio
n Game',

2.0.4 Get all attractions from all 94 pages

In [108]:
```python
1  # Figure out the pattern with page numbers
2  page_1 = 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.' +
3  page_2 = 'https://www.tripadvisor.com/Attraction_Products-g186338-a_sor
4  page_3 = 'https://www.tripadvisor.com/Attraction_Products-g186338-a_sor
```

In [110]:
```python
1  # Test pattern (30 attractions per page)
2  base = 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.'
3  page = '-oa' + str(30)
4  end = '-London_England.html#ATTRACTION_LIST'
5
6  page_2 = base + page + end
7  page_2
```

Out[110]: 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa30-London_E
ngland.html#ATTRACTION_LIST'

In [111]:

```python
1   # Create list of all htmls for London
2   london_htmls = ['http://tripadvisor.com/Attraction_Products-g186338-a_s
3   ]
4   for i in range(2, 95):
5       base = 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.'
6       page = '-oa' + str(30*i - 30)
7       end = '-London_England.html#ATTRACTION_LIST'
8       html = base + page + end
9       london_htmls.append(str(html))
10
11  print(london_htmls)
```

['http://tripadvisor.com/Attraction_Products-g186338-a_sort.-London_Engla
nd.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g18
6338-a_sort.-oa30-London_England.html#ATTRACTION_LIST', 'http://tripadvis
or.com/Attraction_Products-g186338-a_sort.-oa60-London_England.html#ATTRA
CTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-
oa90-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attrac
tion_Products-g186338-a_sort.-oa120-London_England.html#ATTRACTION_LIST',
'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa150-London_
England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Product
s-g186338-a_sort.-oa180-London_England.html#ATTRACTION_LIST', 'http://tri
padvisor.com/Attraction_Products-g186338-a_sort.-oa210-London_England.htm
l#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a
_sort.-oa240-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.co
m/Attraction_Products-g186338-a_sort.-oa270-London_England.html#ATTRACTIO
N_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa30
0-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attractio
n_Products-g186338-a_sort.-oa330-London_England.html#ATTRACTION_LIST', 'h
ttp://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa360-London_En
gland.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-
g186338-a_sort.-oa390-London_England.html#ATTRACTION_LIST', 'http://tripa
dvisor.com/Attraction_Products-g186338-a_sort.-oa420-London_England.html#
ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_s
ort.-oa450-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/
Attraction_Products-g186338-a_sort.-oa480-London_England.html#ATTRACTION_
LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa510-
London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_
Products-g186338-a_sort.-oa540-London_England.html#ATTRACTION_LIST', 'htt
p://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa570-London_Engl
and.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g1
86338-a_sort.-oa600-London_England.html#ATTRACTION_LIST', 'http://tripadv
isor.com/Attraction_Products-g186338-a_sort.-oa630-London_England.html#AT
TRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sor
t.-oa660-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/At
traction_Products-g186338-a_sort.-oa690-London_England.html#ATTRACTION_LI
ST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa720-Lo
ndon_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Pr
oducts-g186338-a_sort.-oa750-London_England.html#ATTRACTION_LIST', 'htt
p://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa780-London_Engl
and.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g1
86338-a_sort.-oa810-London_England.html#ATTRACTION_LIST', 'http://tripadv
isor.com/Attraction_Products-g186338-a_sort.-oa840-London_England.html#AT
TRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sor
t.-oa870-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/At
traction_Products-g186338-a_sort.-oa900-London_England.html#ATTRACTION_LI
ST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa930-Lo

ndon_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Pr
oducts-g186338-a_sort.-oa960-London_England.html#ATTRACTION_LIST', 'htt
p://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa990-London_Engl
and.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g1
86338-a_sort.-oa1020-London_England.html#ATTRACTION_LIST', 'http://tripad
visor.com/Attraction_Products-g186338-a_sort.-oa1050-London_England.html#
ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_s
ort.-oa1080-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.co
m/Attraction_Products-g186338-a_sort.-oa1110-London_England.html#ATTRACTI
ON_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1
140-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attract
ion_Products-g186338-a_sort.-oa1170-London_England.html#ATTRACTION_LIST',
'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1200-London
_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Produc
ts-g186338-a_sort.-oa1230-London_England.html#ATTRACTION_LIST', 'http://t
ripadvisor.com/Attraction_Products-g186338-a_sort.-oa1260-London_England.
html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g18633
8-a_sort.-oa1290-London_England.html#ATTRACTION_LIST', 'http://tripadviso
r.com/Attraction_Products-g186338-a_sort.-oa1320-London_England.html#ATTR
ACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.
-oa1350-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Att
raction_Products-g186338-a_sort.-oa1380-London_England.html#ATTRACTION_LI
ST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1410-L
ondon_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_P
roducts-g186338-a_sort.-oa1440-London_England.html#ATTRACTION_LIST', 'htt
p://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1470-London_Eng
land.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g
186338-a_sort.-oa1500-London_England.html#ATTRACTION_LIST', 'http://tripa
dvisor.com/Attraction_Products-g186338-a_sort.-oa1530-London_England.html
#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_
sort.-oa1560-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.co
m/Attraction_Products-g186338-a_sort.-oa1590-London_England.html#ATTRACTI
ON_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1
620-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attract
ion_Products-g186338-a_sort.-oa1650-London_England.html#ATTRACTION_LIST',
'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1680-London
_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Produc
ts-g186338-a_sort.-oa1710-London_England.html#ATTRACTION_LIST', 'http://t
ripadvisor.com/Attraction_Products-g186338-a_sort.-oa1740-London_England.
html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g18633
8-a_sort.-oa1770-London_England.html#ATTRACTION_LIST', 'http://tripadviso
r.com/Attraction_Products-g186338-a_sort.-oa1800-London_England.html#ATTR
ACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.
-oa1830-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Att
raction_Products-g186338-a_sort.-oa1860-London_England.html#ATTRACTION_LI
ST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1890-L
ondon_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_P
roducts-g186338-a_sort.-oa1920-London_England.html#ATTRACTION_LIST', 'htt
p://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa1950-London_Eng
land.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g
186338-a_sort.-oa1980-London_England.html#ATTRACTION_LIST', 'http://tripa
dvisor.com/Attraction_Products-g186338-a_sort.-oa2010-London_England.html
#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_
sort.-oa2040-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.co
m/Attraction_Products-g186338-a_sort.-oa2070-London_England.html#ATTRACTI
ON_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2
100-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attract

ion_Products-g186338-a_sort.-oa2130-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2160-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2190-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2220-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2250-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2280-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2310-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2340-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2370-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2400-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2430-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2460-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2490-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2520-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2550-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2580-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2610-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2640-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2670-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2700-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2730-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2760-London_England.html#ATTRACTION_LIST', 'http://tripadvisor.com/Attraction_Products-g186338-a_sort.-oa2790-London_England.html#ATTRACTION_LIST']

In [112]:
```python
# Build another function that loops through the list of urls and grabs
def get_all_attractions(html_list):
    """
    input list of strings of html paths from TripAdvisor
    returns list of attraction names from all pages
    """
    final_attractions = []
    for html in html_list:
        html_ = requests.get(html)
        soup = BeautifulSoup(html_.content, 'html.parser')
        container = soup.find('div', id="ATTRACTION_PRODUCTS_LIST")
        attractions = container.findAll('h2')
        new_attractions = []
        for attraction in attractions:
            att = attraction.find('a')
            if hasattr(att, 'title'):
                new_attractions.append(att.attrs['title'])
        final_attractions.extend(new_attractions)
    return final_attractions
```

```
In [105]:  1  # Warning! Takes a few mins to run
           2  all_london_attractions = get_all_attractions(london_htmls)
```

```
In [106]:  1  len(all_london_attractions)
```
Out[106]: 2803

```
In [107]:  1  all_london_attractions[:5]
```
Out[107]: ['SEA LIFE London Aquarium Admission Ticket',
 'The Jack The Ripper Walking Tour in London ',
 'Ghost Bus Tour of London',
 'Big Bus London Hop-On Hop-Off Tour and River Cruise with Optional Londo
n Eye',
 'The Blood and Tears Walk: Serial Killers and London Horror']

```
In [114]:  1  # Make into a DataFrame
           2  london_df = pd.DataFrame(all_london_attractions, columns=['Attraction']
```

```
In [118]:  1  london_df['City'] = 'London, United Kingdom'
```

```
In [119]:  1  london_df
```
Out[119]:

	Attraction	City
0	SEA LIFE London Aquarium Admission Ticket	London, United Kingdom
1	The Jack The Ripper Walking Tour in London	London, United Kingdom
2	Ghost Bus Tour of London	London, United Kingdom
3	Big Bus London Hop-On Hop-Off Tour and River C...	London, United Kingdom
4	The Blood and Tears Walk: Serial Killers and L...	London, United Kingdom
...	...	...
2798	Stonehenge & Winchester	London, United Kingdom
2799	Private Transfer: Heathrow to Southampton Crui...	London, United Kingdom
2800	2 Day Stay Package Tour in London with English...	London, United Kingdom
2801	Bicester Village Shopping Outlet Private Tour ...	London, United Kingdom
2802	Big Bus London Christmas Light Open-Top Night ...	London, United Kingdom

2803 rows × 2 columns

```
In [123]:  1  # Save DF to repo
           2  london_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project
```

2.1 Repeat for all top 12 cities

2.1.1 Paris

In [187]:

```
1   paris_html = 'https://www.tripadvisor.com/Attraction_Products-g187147-P
```

In [188]:

```python
# Create list of all htmls for Paris
paris_htmls = [paris_html]
for i in range(2, 91):
    base = 'https://www.tripadvisor.com/Attraction_Products-g187147-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Paris_Ile_de_France.html#ATTRACTION_LIST'
    html = base + page + end
    paris_htmls.append(str(html))

print(paris_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g187147-Paris_Ile_de_Fr
ance.html', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_so
rt.-oa30-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvi
sor.com/Attraction_Products-g187147-a_sort.-oa60-Paris_Ile_de_France.html
#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1871
47-a_sort.-oa90-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.t
ripadvisor.com/Attraction_Products-g187147-a_sort.-oa120-Paris_Ile_de_Fra
nce.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produc
ts-g187147-a_sort.-oa150-Paris_Ile_de_France.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa180-Paris_I
le_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g187147-a_sort.-oa210-Paris_Ile_de_France.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa24
0-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g187147-a_sort.-oa270-Paris_Ile_de_France.html#ATTR
ACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_
sort.-oa300-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripa
dvisor.com/Attraction_Products-g187147-a_sort.-oa330-Paris_Ile_de_France.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
187147-a_sort.-oa360-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa390-Paris_Ile_d
e_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_P
roducts-g187147-a_sort.-oa420-Paris_Ile_de_France.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa450-Pa
ris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Att
raction_Products-g187147-a_sort.-oa480-Paris_Ile_de_France.html#ATTRACTIO
N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.
-oa510-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g187147-a_sort.-oa540-Paris_Ile_de_France.html#
ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g18714
7-a_sort.-oa570-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.t
ripadvisor.com/Attraction_Products-g187147-a_sort.-oa600-Paris_Ile_de_Fra
nce.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produc
ts-g187147-a_sort.-oa630-Paris_Ile_de_France.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa660-Paris_I
le_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g187147-a_sort.-oa690-Paris_Ile_de_France.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa72
0-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g187147-a_sort.-oa750-Paris_Ile_de_France.html#ATTR
ACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_
sort.-oa780-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripa
dvisor.com/Attraction_Products-g187147-a_sort.-oa810-Paris_Ile_de_France.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
187147-a_sort.-oa840-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa870-Paris_Ile_d

e_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_P roducts-g187147-a_sort.-oa900-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa930-Pa ris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Att raction_Products-g187147-a_sort.-oa960-Paris_Ile_de_France.html#ATTRACTIO N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort. -oa990-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadviso r.com/Attraction_Products-g187147-a_sort.-oa1020-Paris_Ile_de_France.html #ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1871 47-a_sort.-oa1050-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://ww w.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1080-Paris_Ile_de _France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr oducts-g187147-a_sort.-oa1110-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1140-P aris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/At traction_Products-g187147-a_sort.-oa1170-Paris_Ile_de_France.html#ATTRACT ION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sor t.-oa1200-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadv isor.com/Attraction_Products-g187147-a_sort.-oa1230-Paris_Ile_de_France.h tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1 87147-a_sort.-oa1260-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https:// www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1290-Paris_Ile_ de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_ Products-g187147-a_sort.-oa1320-Paris_Ile_de_France.html#ATTRACTION_LIS T', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa13 50-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.co m/Attraction_Products-g187147-a_sort.-oa1380-Paris_Ile_de_France.html#ATT RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a _sort.-oa1410-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tri padvisor.com/Attraction_Products-g187147-a_sort.-oa1440-Paris_Ile_de_Fran ce.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product s-g187147-a_sort.-oa1470-Paris_Ile_de_France.html#ATTRACTION_LIST', 'http s://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1500-Paris_ Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attract ion_Products-g187147-a_sort.-oa1530-Paris_Ile_de_France.html#ATTRACTION_L IST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa 1560-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor. com/Attraction_Products-g187147-a_sort.-oa1590-Paris_Ile_de_France.html#A TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147 -a_sort.-oa1620-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.t ripadvisor.com/Attraction_Products-g187147-a_sort.-oa1650-Paris_Ile_de_Fr ance.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ cts-g187147-a_sort.-oa1680-Paris_Ile_de_France.html#ATTRACTION_LIST', 'ht tps://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1710-Pari s_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra ction_Products-g187147-a_sort.-oa1740-Paris_Ile_de_France.html#ATTRACTION _LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.- oa1770-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadviso r.com/Attraction_Products-g187147-a_sort.-oa1800-Paris_Ile_de_France.html #ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1871 47-a_sort.-oa1830-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://ww w.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1860-Paris_Ile_de _France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr oducts-g187147-a_sort.-oa1890-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa1920-P aris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/At traction_Products-g187147-a_sort.-oa1950-Paris_Ile_de_France.html#ATTRACT

```
ION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sor
t.-oa1980-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadv
isor.com/Attraction_Products-g187147-a_sort.-oa2010-Paris_Ile_de_France.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1
87147-a_sort.-oa2040-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa2070-Paris_Ile_
de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_
Products-g187147-a_sort.-oa2100-Paris_Ile_de_France.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa21
30-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g187147-a_sort.-oa2160-Paris_Ile_de_France.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a
_sort.-oa2190-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g187147-a_sort.-oa2220-Paris_Ile_de_Fran
ce.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g187147-a_sort.-oa2250-Paris_Ile_de_France.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa2280-Paris_
Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attract
ion_Products-g187147-a_sort.-oa2310-Paris_Ile_de_France.html#ATTRACTION_L
IST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa
2340-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.
com/Attraction_Products-g187147-a_sort.-oa2370-Paris_Ile_de_France.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147
-a_sort.-oa2400-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.t
ripadvisor.com/Attraction_Products-g187147-a_sort.-oa2430-Paris_Ile_de_Fr
ance.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ
cts-g187147-a_sort.-oa2460-Paris_Ile_de_France.html#ATTRACTION_LIST', 'ht
tps://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa2490-Pari
s_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g187147-a_sort.-oa2520-Paris_Ile_de_France.html#ATTRACTION
_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187147-a_sort.-
oa2550-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g187147-a_sort.-oa2580-Paris_Ile_de_France.html
#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1871
47-a_sort.-oa2610-Paris_Ile_de_France.html#ATTRACTION_LIST', 'https://ww
w.tripadvisor.com/Attraction_Products-g187147-a_sort.-oa2640-Paris_Ile_de
_France.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g187147-a_sort.-oa2670-Paris_Ile_de_France.html#ATTRACTION_LIST']
```

In [189]:
```python
1  all_attractions_paris = get_all_attractions(paris_htmls)
2  print(len(all_attractions_paris))
3  print(all_attractions_paris[:5])
```

```
2689
['Bateaux Parisiens Seine River Gourmet Dinner & Sightseeing Cruise', 'Lo
uvre Museum Skip the Line Access Guided Tour with Venus de Milo & Mona Li
sa', "Paris' Louvre Museum Virtual Guided Tour", 'See 30 Plus Top Paris S
ights with a Fun Guide', 'Paris Sightseeing Guided Bike Tour Like a Paris
ian with a Local Guide']
```

In [190]:
```python
1  paris_df = pd.DataFrame(all_attractions_paris, columns=['Attraction'])
2  paris_df['City'] = 'Paris, France'
3  paris_df.head()
```

Out[190]:

	Attraction	City
0	Bateaux Parisiens Seine River Gourmet Dinner &...	Paris, France
1	Louvre Museum Skip the Line Access Guided Tour...	Paris, France
2	Paris' Louvre Museum Virtual Guided Tour	Paris, France
3	See 30 Plus Top Paris Sights with a Fun Guide	Paris, France
4	Paris Sightseeing Guided Bike Tour Like a Pari...	Paris, France

In [191]:
```python
1  paris_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project/
```

▼ 2.1.2 Crete, Greece

In [183]:

```python
# Create list of all htmls for Crete
crete_htmls = ['https://www.tripadvisor.com/Attraction_Products-g189413

for i in range(2, 36):
    base = 'https://www.tripadvisor.com/Attraction_Products-g189413-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Crete.html#ATTRACTION_LIST'
    html = base + page + end
    crete_htmls.append(str(html))

print(crete_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g189413-Crete.html', 'h
ttps://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa30-Cret
e.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g189413-a_sort.-oa60-Crete.html#ATTRACTION_LIST', 'https://www.tripadvis
or.com/Attraction_Products-g189413-a_sort.-oa90-Crete.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa12
0-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g189413-a_sort.-oa150-Crete.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g189413-a_sort.-oa180-Crete.html#ATTRAC
TION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_so
rt.-oa210-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g189413-a_sort.-oa240-Crete.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa270-Crete.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1
89413-a_sort.-oa300-Crete.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g189413-a_sort.-oa330-Crete.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa36
0-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g189413-a_sort.-oa390-Crete.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g189413-a_sort.-oa420-Crete.html#ATTRAC
TION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_so
rt.-oa450-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g189413-a_sort.-oa480-Crete.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa510-Crete.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1
89413-a_sort.-oa540-Crete.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g189413-a_sort.-oa570-Crete.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa60
0-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g189413-a_sort.-oa630-Crete.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g189413-a_sort.-oa660-Crete.html#ATTRAC
TION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_so
rt.-oa690-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g189413-a_sort.-oa720-Crete.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa750-Crete.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1
89413-a_sort.-oa780-Crete.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g189413-a_sort.-oa810-Crete.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa84
0-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g189413-a_sort.-oa870-Crete.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g189413-a_sort.-oa900-Crete.html#ATTRAC
TION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g189413-a_so
rt.-oa930-Crete.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g189413-a_sort.-oa960-Crete.html#ATTRACTION_LIST', 'http

```
s://www.tripadvisor.com/Attraction_Products-g189413-a_sort.-oa990-Crete.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1
89413-a_sort.-oa1020-Crete.html#ATTRACTION_LIST']
```

In [184]:
```python
1  all_attractions_crete = get_all_attractions(crete_htmls)
2  print(len(all_attractions_crete))
3  print(all_attractions_crete[:5])
```

```
1023
['Minoans World 3D Museum & 9D Cinema', 'Crete Wine and Olive Oil Tour',
'Skip the Line: Knossos Palace and Archaeological Site Entrance Ticket',
'Georgioupolis Safari Off Road Tours Outdoor Activities Land Rover Defend
er', 'Spinalonga island']
```

In [185]:
```python
1  crete_df = pd.DataFrame(all_attractions_crete, columns=['Attraction'])
2  crete_df['City'] = 'Crete, Greece'
3  crete_df.head()
```

Out[185]:

	Attraction	City
0	Minoans World 3D Museum & 9D Cinema	Crete, Greece
1	Crete Wine and Olive Oil Tour	Crete, Greece
2	Skip the Line: Knossos Palace and Archaeologic...	Crete, Greece
3	Georgioupolis Safari Off Road Tours Outdoor Ac...	Crete, Greece
4	Spinalonga island	Crete, Greece

In [186]:
```python
1  crete_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project/
```

2.1.3 Bali, Indonesia

In [179]:
```python
# Create list of all htmls for Bali
bali_htmls = ['https://www.tripadvisor.com/Attraction_Products-g294226-

for i in range(2, 168):
    base = 'https://www.tripadvisor.com/Attraction_Products-g294226-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Bali.html#ATTRACTION_LIST'
    html = base + page + end
    bali_htmls.append(str(html))

print(bali_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g294226-Bali.html', 'ht
tps://www.tripadvisor.com/Attraction_Products-g294226-a_sort.-oa30-Bali.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g2
94226-a_sort.-oa60-Bali.html#ATTRACTION_LIST', 'https://www.tripadvisor.c
om/Attraction_Products-g294226-a_sort.-oa90-Bali.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g294226-a_sort.-oa120-Ba
li.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g294226-a_sort.-oa150-Bali.html#ATTRACTION_LIST', 'https://www.tripadvi
sor.com/Attraction_Products-g294226-a_sort.-oa180-Bali.html#ATTRACTION_LI
ST', 'https://www.tripadvisor.com/Attraction_Products-g294226-a_sort.-oa2
10-Bali.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g294226-a_sort.-oa240-Bali.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g294226-a_sort.-oa270-Bali.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g294226-a_sor
t.-oa300-Bali.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attract
ion_Products-g294226-a_sort.-oa330-Bali.html#ATTRACTION_LIST', 'https://w
ww.tripadvisor.com/Attraction_Products-g294226-a_sort.-oa360-Bali.html#AT
TRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g294226-
a_sort.-oa390-Bali.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/At

In [180]:
```python
all_attractions_bali = get_all_attractions(bali_htmls)
print(len(all_attractions_bali))
print(all_attractions_bali[:5])
```

5000
['Hotel to Hotel:Bali Private Transfer (Daytime)', 'Bali White Water Raft
ing', 'Bali Snorkeling Tour at Blue Lagoon Beach: All-Inclusive', 'Privat
e Bali Hotel Transfer: Hotel to Hotel', 'Inter hotel transfer in Bali (Ku
ta, Seminyak and Nisa dua) to ubud (PVT)']

In [181]:
```python
bali_df = pd.DataFrame(all_attractions_bali, columns=['Attraction'])
bali_df['City'] = 'Bali, Indonesia'
bali_df.head()
```

Out[181]:

	Attraction	City
0	Hotel to Hotel:Bali Private Transfer (Daytime)	Bali, Indonesia
1	Bali White Water Rafting	Bali, Indonesia
2	Bali Snorkeling Tour at Blue Lagoon Beach: All...	Bali, Indonesia
3	Private Bali Hotel Transfer: Hotel to Hotel	Bali, Indonesia
4	Inter hotel transfer in Bali (Kuta, Seminyak a...	Bali, Indonesia

In [182]:
```python
bali_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project/b
```

2.1.4 Rome, Italy

In [175]:
```python
# Create list of all htmls for Rome
rome_htmls = ['https://www.tripadvisor.com/Attraction_Products-g187791-

for i in range(2, 168):
    base = 'https://www.tripadvisor.com/Attraction_Products-g187791-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Rome_Lazio.html#ATTRACTION_LIST'
    html = base + page + end
    rome_htmls.append(str(html))

print(rome_htmls)
```

```
['https://www.tripadvisor.com/Attraction_Products-g187791-Rome_Lazio.htm
l', 'https://www.tripadvisor.com/Attraction_Products-g187791-a_sort.-oa30
-Rome_Lazio.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attractio
n_Products-g187791-a_sort.-oa60-Rome_Lazio.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187791-a_sort.-oa90-Rome_Laz
io.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g187791-a_sort.-oa120-Rome_Lazio.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g187791-a_sort.-oa150-Rome_Lazio.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187791
-a_sort.-oa180-Rome_Lazio.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g187791-a_sort.-oa210-Rome_Lazio.html#ATTRACTIO
N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187791-a_sort.
-oa240-Rome_Lazio.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Att
raction_Products-g187791-a_sort.-oa270-Rome_Lazio.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g187791-a_sort.-oa300-Ro
me_Lazio.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_P
roducts-g187791-a_sort.-oa330-Rome_Lazio.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g187791-a_sort.-oa360-Rome_Lazio.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
```

In [176]:
```python
all_attractions_rome = get_all_attractions(rome_htmls)
print(len(all_attractions_rome))
print(all_attractions_rome[:5])
```

```
5000
["Faster Than Skip-the-Line: Vatican, Sistine Chapel and St. Peter's Basi
lica Tour", 'Skip the Line: Colosseum Small Group Tour with Roman Forum &
Palatine Hill ', 'Colosseum Underground Guided Tour with Palatine Hill &
Roman Forum Access', "Skip-the-Line Group Tour of the Vatican, Sistine Ch
apel & St. Peter's Basilica", 'Fast- track- entry ticket to the Vatican M
useums & the Sistine Chapel']
```

In [177]:

```python
1  rome_df = pd.DataFrame(all_attractions_rome, columns=['Attraction'])
2  rome_df['City'] = 'Rome, Italy'
3  rome_df.head()
```

Out[177]:

	Attraction	City
0	Faster Than Skip-the-Line: Vatican, Sistine Ch...	Rome, Italy
1	Skip the Line: Colosseum Small Group Tour with...	Rome, Italy
2	Colosseum Underground Guided Tour with Palatin...	Rome, Italy
3	Skip-the-Line Group Tour of the Vatican, Sisti...	Rome, Italy
4	Fast- track- entry ticket to the Vatican Museu...	Rome, Italy

In [178]:

```python
1  rome_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project/r
```

▼ 2.1.5 Phuket, Thailand

In [171]:

```python
 1  # Create list of all htmls for Phuket
 2  phuket_htmls = ['https://www.tripadvisor.com/Attraction_Products-g29392
 3
 4  for i in range(2, 42):
 5      base = 'https://www.tripadvisor.com/Attraction_Products-g293920-a_s
 6      page = '-oa' + str(30*i - 30)
 7      end = '-Phuket.html#ATTRACTION_LIST'
 8      html = base + page + end
 9      phuket_htmls.append(str(html))
10
11  print(phuket_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g293920-Phuket.html',
'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa30-Phu
ket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produc
ts-g293920-a_sort.-oa60-Phuket.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g293920-a_sort.-oa90-Phuket.html#ATTRACTION
_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-
oa120-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g293920-a_sort.-oa150-Phuket.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa180-Phuket.html
#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g29393
20-a_sort.-oa210-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.c
om/Attraction_Products-g293920-a_sort.-oa240-Phuket.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa27
0-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_P
roducts-g293920-a_sort.-oa300-Phuket.html#ATTRACTION_LIST', 'https://www.
tripadvisor.com/Attraction_Products-g293920-a_sort.-oa330-Phuket.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a
_sort.-oa360-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/A
ttraction_Products-g293920-a_sort.-oa390-Phuket.html#ATTRACTION_LIST', 'h
ttps://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa420-Phuk
et.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g293920-a_sort.-oa450-Phuket.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g293920-a_sort.-oa480-Phuket.html#ATTRACTIO
N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.
-oa510-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attract
ion_Products-g293920-a_sort.-oa540-Phuket.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa570-Phuket.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
293920-a_sort.-oa600-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvis
or.com/Attraction_Products-g293920-a_sort.-oa630-Phuket.html#ATTRACTION_L
IST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa
660-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction
_Products-g293920-a_sort.-oa690-Phuket.html#ATTRACTION_LIST', 'https://ww
w.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa720-Phuket.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920
-a_sort.-oa750-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g293920-a_sort.-oa780-Phuket.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa810-Ph
uket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ
cts-g293920-a_sort.-oa840-Phuket.html#ATTRACTION_LIST', 'https://www.trip
advisor.com/Attraction_Products-g293920-a_sort.-oa870-Phuket.html#ATTRACT
ION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sor
t.-oa900-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g293920-a_sort.-oa930-Phuket.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa960-Phuket.

```
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
293920-a_sort.-oa990-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvis
or.com/Attraction_Products-g293920-a_sort.-oa1020-Phuket.html#ATTRACTION_
LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-o
a1050-Phuket.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g293920-a_sort.-oa1080-Phuket.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293920-a_sort.-oa1110-Phuke
t.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g293920-a_sort.-oa1140-Phuket.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g293920-a_sort.-oa1170-Phuket.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293920-a_sor
t.-oa1200-Phuket.html#ATTRACTION_LIST']
```

In [172]:
```
1  all_attractions_phuket = get_all_attractions(phuket_htmls)
2  print(len(all_attractions_phuket))
3  print(all_attractions_phuket[:5])
```

```
1201
['Phi Phi, MaiTon & Khai Islands by Speedboat', 'Phi Phi Islands Adventur
e Day Trip with Seaview Lunch', 'Phi Phi Island by Premium Speedboat incl
uding Buffet Lunch', 'Phang Nga Bay Sea Cave Canoeing & James Bond Island
w/ Buffet Lunch by Big Boat', 'A Morning with the Elephants at Phuket Ele
phant Sanctuary']
```

In [173]:
```
1  phuket_df = pd.DataFrame(all_attractions_phuket, columns=['Attraction']
2  phuket_df['City'] = 'Phuket, Thailand'
3  phuket_df.head()
```

Out[173]:

	Attraction	City
0	Phi Phi, MaiTon & Khai Islands by Speedboat	Phuket, Thailand
1	Phi Phi Islands Adventure Day Trip with Seavie…	Phuket, Thailand
2	Phi Phi Island by Premium Speedboat including …	Phuket, Thailand
3	Phang Nga Bay Sea Cave Canoeing & James Bond I...	Phuket, Thailand
4	A Morning with the Elephants at Phuket Elephan…	Phuket, Thailand

In [174]:
```
1  phuket_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project
```

2.1.6 Sicily, Italy

In [167]:

```python
# Create list of all htmls for Sicily
sicily_htmls = ['https://www.tripadvisor.com/Attraction_Products-g18788

for i in range(2, 68):
    base = 'https://www.tripadvisor.com/Attraction_Products-g187886-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Sicily.html#ATTRACTION_LIST'
    html = base + page + end
    sicily_htmls.append(str(html))

print(sicily_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g187886-Sicily.html', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa30-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa60-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa90-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa120-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa150-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa180-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa210-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa240-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa270-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa300-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa330-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa360-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa390-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa420-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa450-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa480-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa510-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa540-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa570-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa600-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa630-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa660-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa690-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa720-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa750-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa780-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa810-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa840-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa870-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa900-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa930-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa960-Sicily.

```
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
187886-a_sort.-oa990-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvis
or.com/Attraction_Products-g187886-a_sort.-oa1020-Sicily.html#ATTRACTION_
LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-o
a1050-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g187886-a_sort.-oa1080-Sicily.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa1110-Sicil
y.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187886-a_sort.-oa1140-Sicily.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g187886-a_sort.-oa1170-Sicily.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sor
t.-oa1200-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attr
action_Products-g187886-a_sort.-oa1230-Sicily.html#ATTRACTION_LIST', 'htt
ps://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa1260-Sicil
y.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187886-a_sort.-oa1290-Sicily.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g187886-a_sort.-oa1320-Sicily.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sor
t.-oa1350-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attr
action_Products-g187886-a_sort.-oa1380-Sicily.html#ATTRACTION_LIST', 'htt
ps://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa1410-Sicil
y.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187886-a_sort.-oa1440-Sicily.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g187886-a_sort.-oa1470-Sicily.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sor
t.-oa1500-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attr
action_Products-g187886-a_sort.-oa1530-Sicily.html#ATTRACTION_LIST', 'htt
ps://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa1560-Sicil
y.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187886-a_sort.-oa1590-Sicily.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g187886-a_sort.-oa1620-Sicily.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sor
t.-oa1650-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attr
action_Products-g187886-a_sort.-oa1680-Sicily.html#ATTRACTION_LIST', 'htt
ps://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa1710-Sicil
y.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187886-a_sort.-oa1740-Sicily.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g187886-a_sort.-oa1770-Sicily.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sor
t.-oa1800-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attr
action_Products-g187886-a_sort.-oa1830-Sicily.html#ATTRACTION_LIST', 'htt
ps://www.tripadvisor.com/Attraction_Products-g187886-a_sort.-oa1860-Sicil
y.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187886-a_sort.-oa1890-Sicily.html#ATTRACTION_LIST', 'https://www.tripad
visor.com/Attraction_Products-g187886-a_sort.-oa1920-Sicily.html#ATTRACTI
ON_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187886-a_sor
t.-oa1950-Sicily.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attr
action_Products-g187886-a_sort.-oa1980-Sicily.html#ATTRACTION_LIST']
```

In [168]:
```python
all_attractions_sicily = get_all_attractions(sicily_htmls)
print(len(all_attractions_sicily))
print(all_attractions_sicily[:5])
```

```
2004
['Etna and Taormina Full-Day Tour from Catania', 'Palermo Street Food Wal
king Tour', 'Sailing to Cyclops Islands Group Tour', 'Sunset Trekking on
Stromboli - Ashàra Volcanological Guides', 'Excursion around Ortigia Isla
nd and sea caves']
```

In [169]:
```python
sicily_df = pd.DataFrame(all_attractions_sicily, columns=['Attraction']
sicily_df['City'] = 'Sicily, Italy'
sicily_df.head()
```

Out[169]:

	Attraction	City
0	Etna and Taormina Full-Day Tour from Catania	Sicily, Italy
1	Palermo Street Food Walking Tour	Sicily, Italy
2	Sailing to Cyclops Islands Group Tour	Sicily, Italy
3	Sunset Trekking on Stromboli - Ashàra Volcanol…	Sicily, Italy
4	Excursion around Ortigia Island and sea caves	Sicily, Italy

In [170]:
```python
sicily_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project
```

2.1.7 Majorca, Balearic Islands

In [161]:
```python
# Create list of all htmls for Majorca
majorca_htmls = ['https://www.tripadvisor.com/Attraction_Products-g1874

for i in range(2, 17):
    base = 'https://www.tripadvisor.com/Attraction_Products-g187462-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Majorca_Balearic_Islands.html#ATTRACTION_LIST'
    html = base + page + end
    majorca_htmls.append(str(html))

print(majorca_htmls)
```

```
['https://www.tripadvisor.com/Attraction_Products-g187462-Majorca_Baleari
c_Islands.html', 'https://www.tripadvisor.com/Attraction_Products-g187462
-a_sort.-oa30-Majorca_Balearic_Islands.html#ATTRACTION_LIST', 'https://ww
w.tripadvisor.com/Attraction_Products-g187462-a_sort.-oa60-Majorca_Balear
ic_Islands.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction
_Products-g187462-a_sort.-oa90-Majorca_Balearic_Islands.html#ATTRACTION_L
IST', 'https://www.tripadvisor.com/Attraction_Products-g187462-a_sort.-oa
120-Majorca_Balearic_Islands.html#ATTRACTION_LIST', 'https://www.tripadvi
sor.com/Attraction_Products-g187462-a_sort.-oa150-Majorca_Balearic_Island
s.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g187462-a_sort.-oa180-Majorca_Balearic_Islands.html#ATTRACTION_LIST', 'h
ttps://www.tripadvisor.com/Attraction_Products-g187462-a_sort.-oa210-Majo
rca_Balearic_Islands.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/
Attraction_Products-g187462-a_sort.-oa240-Majorca_Balearic_Islands.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187462
-a_sort.-oa270-Majorca_Balearic_Islands.html#ATTRACTION_LIST', 'https://w
ww.tripadvisor.com/Attraction_Products-g187462-a_sort.-oa300-Majorca_Bale
aric_Islands.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g187462-a_sort.-oa330-Majorca_Balearic_Islands.html#ATTRACTIO
N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187462-a_sort.
-oa360-Majorca_Balearic_Islands.html#ATTRACTION_LIST', 'https://www.tripa
dvisor.com/Attraction_Products-g187462-a_sort.-oa390-Majorca_Balearic_Isl
ands.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ
cts-g187462-a_sort.-oa420-Majorca_Balearic_Islands.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g187462-a_sort.-oa450-Ma
jorca_Balearic_Islands.html#ATTRACTION_LIST']
```

In [162]:
```python
all_attractions_majorca = get_all_attractions(majorca_htmls)
print(len(all_attractions_majorca))
print(all_attractions_majorca[:5])
```

```
454
['Caves of Genova Admission in Palma de Mallorca', 'Shore Excursion:Palma
Hop-On Hop-Off Bus Tour with Optional Boat & Belver Castle', 'Exclusive s
ailing trip half day sailing Palma Bay, Playa de Palma, Can Pastilla)',
'Mallorca in One Day Sightseeing Tour with Boat Ride and Vintage Train',
'Electric Scooter Tour in Palma de Mallorca']
```

In [163]:
```python
majorca_df = pd.DataFrame(all_attractions_majorca, columns=['Attraction
majorca_df['City'] = 'Majorca, Balearic Islands'
majorca_df.head()
```

Out[163]:

	Attraction	City
0	Caves of Genova Admission in Palma de Mallorca	Majorca, Balearic Islands
1	Shore Excursion:Palma Hop-On Hop-Off Bus Tour ...	Majorca, Balearic Islands
2	Exclusive sailing trip half day sailing Palma ...	Majorca, Balearic Islands
3	Mallorca in One Day Sightseeing Tour with Boat...	Majorca, Balearic Islands
4	Electric Scooter Tour in Palma de Mallorca	Majorca, Balearic Islands

In [164]:
```python
majorca_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-projec
```

2.1.8 Barcelona, Spain

In [159]:
```python
# Create list of all htmls for Barcelona
barcelona_htmls = ['https://www.tripadvisor.com/Attraction_Products-g18

for i in range(2, 63):
    base = 'https://www.tripadvisor.com/Attraction_Products-g187497-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Barcelona_Catalonia.html#ATTRACTION_LIST'
    html = base + page + end
    barcelona_htmls.append(str(html))

print(barcelona_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g187497-Barcelona_Catal
onia.html', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_so
rt.-oa30-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvi
sor.com/Attraction_Products-g187497-a_sort.-oa60-Barcelona_Catalonia.html
#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1874
97-a_sort.-oa90-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.t
ripadvisor.com/Attraction_Products-g187497-a_sort.-oa120-Barcelona_Catalo
nia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produc
ts-g187497-a_sort.-oa150-Barcelona_Catalonia.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa180-Barcelo
na_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g187497-a_sort.-oa210-Barcelona_Catalonia.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa24
0-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g187497-a_sort.-oa270-Barcelona_Catalonia.html#ATTR
ACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_
sort.-oa300-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripa
dvisor.com/Attraction_Products-g187497-a_sort.-oa330-Barcelona_Catalonia.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
187497-a_sort.-oa360-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa390-Barcelona_C
atalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_P
roducts-g187497-a_sort.-oa420-Barcelona_Catalonia.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa450-Ba
rcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Att
raction_Products-g187497-a_sort.-oa480-Barcelona_Catalonia.html#ATTRACTIO
N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.
-oa510-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g187497-a_sort.-oa540-Barcelona_Catalonia.html#
ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g18749
7-a_sort.-oa570-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.t
ripadvisor.com/Attraction_Products-g187497-a_sort.-oa600-Barcelona_Catalo
nia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produc
ts-g187497-a_sort.-oa630-Barcelona_Catalonia.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa660-Barcelo
na_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g187497-a_sort.-oa690-Barcelona_Catalonia.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa72
0-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g187497-a_sort.-oa750-Barcelona_Catalonia.html#ATTR
ACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_
sort.-oa780-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripa
dvisor.com/Attraction_Products-g187497-a_sort.-oa810-Barcelona_Catalonia.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
187497-a_sort.-oa840-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://

www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa870-Barcelona_C
atalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_P
roducts-g187497-a_sort.-oa900-Barcelona_Catalonia.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa930-Ba
rcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Att
raction_Products-g187497-a_sort.-oa960-Barcelona_Catalonia.html#ATTRACTIO
N_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.
-oa990-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g187497-a_sort.-oa1020-Barcelona_Catalonia.html
#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1874
97-a_sort.-oa1050-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://ww
w.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa1080-Barcelona_Ca
talonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pr
oducts-g187497-a_sort.-oa1110-Barcelona_Catalonia.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa1140-B
arcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/At
traction_Products-g187497-a_sort.-oa1170-Barcelona_Catalonia.html#ATTRACT
ION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sor
t.-oa1200-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadv
isor.com/Attraction_Products-g187497-a_sort.-oa1230-Barcelona_Catalonia.h
tml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1
87497-a_sort.-oa1260-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa1290-Barcelona_
Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_
Products-g187497-a_sort.-oa1320-Barcelona_Catalonia.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa13
50-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g187497-a_sort.-oa1380-Barcelona_Catalonia.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a
_sort.-oa1410-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g187497-a_sort.-oa1440-Barcelona_Catalon
ia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g187497-a_sort.-oa1470-Barcelona_Catalonia.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa1500-Barcel
ona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attract
ion_Products-g187497-a_sort.-oa1530-Barcelona_Catalonia.html#ATTRACTION_L
IST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa
1560-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.
com/Attraction_Products-g187497-a_sort.-oa1590-Barcelona_Catalonia.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497
-a_sort.-oa1620-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.t
ripadvisor.com/Attraction_Products-g187497-a_sort.-oa1650-Barcelona_Catal
onia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ
cts-g187497-a_sort.-oa1680-Barcelona_Catalonia.html#ATTRACTION_LIST', 'ht
tps://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-oa1710-Barc
elona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attra
ction_Products-g187497-a_sort.-oa1740-Barcelona_Catalonia.html#ATTRACTION
_LIST', 'https://www.tripadvisor.com/Attraction_Products-g187497-a_sort.-
oa1770-Barcelona_Catalonia.html#ATTRACTION_LIST', 'https://www.tripadviso
r.com/Attraction_Products-g187497-a_sort.-oa1800-Barcelona_Catalonia.html
#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g1874
97-a_sort.-oa1830-Barcelona_Catalonia.html#ATTRACTION_LIST']

In [160]:
```python
all_attractions_barcelona = get_all_attractions(barcelona_htmls)
print(len(all_attractions_barcelona))
print(all_attractions_barcelona[:5])
```

```
1855
['Interactive Spanish Cooking Experience in Barcelona', 'Flamenco Show Ti
cket at Theatre Barcelona City Hall', 'Best of Barcelona Live Virtual Tou
r', '2 hour Sunset Sailing Cruise Barcelona', 'Camp Nou - F.C.Barcelona O
pen Date Ticket']
```

In [165]:
```python
barcelona_df = pd.DataFrame(all_attractions_barcelona, columns=['Attrac
barcelona_df['City'] = 'Barcelona, Spain'
barcelona_df.head()
```

Out[165]:

	Attraction	City
0	Interactive Spanish Cooking Experience in Barc...	Barcelona, Spain
1	Flamenco Show Ticket at Theatre Barcelona City...	Barcelona, Spain
2	Best of Barcelona Live Virtual Tour	Barcelona, Spain
3	2 hour Sunset Sailing Cruise Barcelona	Barcelona, Spain
4	Camp Nou - F.C.Barcelona Open Date Ticket	Barcelona, Spain

In [166]:
```python
barcelona_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-proj
```

2.1.9 Istanbul, Turkey

In [192]:
```python
# Create list of all htmls for Istanbul
istanbul_htmls = ['https://www.tripadvisor.com/Attraction_Products-g293

for i in range(2, 78):
    base = 'https://www.tripadvisor.com/Attraction_Products-g293974-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Istanbul.html#ATTRACTION_LIST'
    html = base + page + end
    istanbul_htmls.append(str(html))

print(istanbul_htmls)
```

['https://www.tripadvisor.com/Attraction_Products-g293974-Istanbul.html',
'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa30-Ist
anbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Prod
ucts-g293974-a_sort.-oa60-Istanbul.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g293974-a_sort.-oa90-Istanbul.html#ATTR
ACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_
sort.-oa120-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/
Attraction_Products-g293974-a_sort.-oa150-Istanbul.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa180-Is
tanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pro
ducts-g293974-a_sort.-oa210-Istanbul.html#ATTRACTION_LIST', 'https://www.
tripadvisor.com/Attraction_Products-g293974-a_sort.-oa240-Istanbul.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974
-a_sort.-oa270-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.c
om/Attraction_Products-g293974-a_sort.-oa300-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa33
0-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction
_Products-g293974-a_sort.-oa360-Istanbul.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa390-Istanbul.ht
ml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g29
3974-a_sort.-oa420-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvis
or.com/Attraction_Products-g293974-a_sort.-oa450-Istanbul.html#ATTRACTION
_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-
oa480-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attrac
tion_Products-g293974-a_sort.-oa510-Istanbul.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa540-Istanbu
l.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g293974-a_sort.-oa570-Istanbul.html#ATTRACTION_LIST', 'https://www.tripa
dvisor.com/Attraction_Products-g293974-a_sort.-oa600-Istanbul.html#ATTRAC
TION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_so
rt.-oa630-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/At
traction_Products-g293974-a_sort.-oa660-Istanbul.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa690-Is
tanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Pro
ducts-g293974-a_sort.-oa720-Istanbul.html#ATTRACTION_LIST', 'https://www.
tripadvisor.com/Attraction_Products-g293974-a_sort.-oa750-Istanbul.html#A
TTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974
-a_sort.-oa780-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.c
om/Attraction_Products-g293974-a_sort.-oa810-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa84
0-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction
_Products-g293974-a_sort.-oa870-Istanbul.html#ATTRACTION_LIST', 'https://
www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa900-Istanbul.ht
ml#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g29
3974-a_sort.-oa930-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvis

or.com/Attraction_Products-g293974-a_sort.-oa960-Istanbul.html#ATTRACTION
_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-
oa990-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attrac
tion_Products-g293974-a_sort.-oa1020-Istanbul.html#ATTRACTION_LIST', 'htt
ps://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa1050-Istan
bul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produc
ts-g293974-a_sort.-oa1080-Istanbul.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g293974-a_sort.-oa1110-Istanbul.html#AT
TRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-
a_sort.-oa1140-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.c
om/Attraction_Products-g293974-a_sort.-oa1170-Istanbul.html#ATTRACTION_LI
ST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa1
200-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attracti
on_Products-g293974-a_sort.-oa1230-Istanbul.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa1260-Istanb
ul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g293974-a_sort.-oa1290-Istanbul.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g293974-a_sort.-oa1320-Istanbul.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a
_sort.-oa1350-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g293974-a_sort.-oa1380-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa14
10-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attractio
n_Products-g293974-a_sort.-oa1440-Istanbul.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa1470-Istanb
ul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g293974-a_sort.-oa1500-Istanbul.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g293974-a_sort.-oa1530-Istanbul.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a
_sort.-oa1560-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g293974-a_sort.-oa1590-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa16
20-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attractio
n_Products-g293974-a_sort.-oa1650-Istanbul.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa1680-Istanb
ul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g293974-a_sort.-oa1710-Istanbul.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g293974-a_sort.-oa1740-Istanbul.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a
_sort.-oa1770-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g293974-a_sort.-oa1800-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa18
30-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attractio
n_Products-g293974-a_sort.-oa1860-Istanbul.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa1890-Istanb
ul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g293974-a_sort.-oa1920-Istanbul.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g293974-a_sort.-oa1950-Istanbul.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a
_sort.-oa1980-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g293974-a_sort.-oa2010-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa20
40-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attractio
n_Products-g293974-a_sort.-oa2070-Istanbul.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa2100-Istanb
ul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Product
s-g293974-a_sort.-oa2130-Istanbul.html#ATTRACTION_LIST', 'https://www.tri

```
padvisor.com/Attraction_Products-g293974-a_sort.-oa2160-Istanbul.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g293974-a
_sort.-oa2190-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g293974-a_sort.-oa2220-Istanbul.html#ATTRACTION_LIS
T', 'https://www.tripadvisor.com/Attraction_Products-g293974-a_sort.-oa22
50-Istanbul.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attractio
n_Products-g293974-a_sort.-oa2280-Istanbul.html#ATTRACTION_LIST']
```

In [193]:
```
1  all_attractions_istanbul = get_all_attractions(istanbul_htmls)
2  print(len(all_attractions_istanbul))
3  print(all_attractions_istanbul[:5])
```

```
2307
['Bosphorus Sunset Cruise on Luxury Yacht', '90-Minute Istanbul Bosphorus
Cruise, Free Refreshment and Audio Guide App ', 'Best of Istanbul: 1, 2 o
r 3-Day Private Guided Tour', 'Istanbul Full Day Old City Tour', 'Taste o
f Two Continents Food Tour 5 People Max']
```

In [194]:
```
1  istanbul_df = pd.DataFrame(all_attractions_istanbul, columns=['Attracti
2  istanbul_df['City'] = 'Istanbul, Turkey'
3  istanbul_df.head()
```

Out[194]:

	Attraction	City
0	Bosphorus Sunset Cruise on Luxury Yacht	Istanbul, Turkey
1	90-Minute Istanbul Bosphorus Cruise, Free Refr...	Istanbul, Turkey
2	Best of Istanbul: 1, 2 or 3-Day Private Guided...	Istanbul, Turkey
3	Istanbul Full Day Old City Tour	Istanbul, Turkey
4	Taste of Two Continents Food Tour 5 People Max	Istanbul, Turkey

In [195]:
```
1  istanbul_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-proje
```

2.1.10 Goa, India

In [196]:
```python
# Create list of all htmls for Goa
goa_htmls = ['https://www.tripadvisor.com/Attraction_Products-g297604-

for i in range(2, 13):
    base = 'https://www.tripadvisor.com/Attraction_Products-g297604-a_s
    page = '-oa' + str(30*i - 30)
    end = '-Goa.html#ATTRACTION_LIST'
    html = base + page + end
    goa_htmls.append(str(html))

print(goa_htmls)
```

```
['https://www.tripadvisor.com/Attraction_Products-g297604-Goa.html', 'htt
ps://www.tripadvisor.com/Attraction_Products-g297604-a_sort.-oa30-Goa.htm
l#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g297
604-a_sort.-oa60-Goa.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/
Attraction_Products-g297604-a_sort.-oa90-Goa.html#ATTRACTION_LIST', 'http
s://www.tripadvisor.com/Attraction_Products-g297604-a_sort.-oa120-Goa.htm
l#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g297
604-a_sort.-oa150-Goa.html#ATTRACTION_LIST', 'https://www.tripadvisor.co
m/Attraction_Products-g297604-a_sort.-oa180-Goa.html#ATTRACTION_LIST', 'h
ttps://www.tripadvisor.com/Attraction_Products-g297604-a_sort.-oa210-Goa.
html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g
297604-a_sort.-oa240-Goa.html#ATTRACTION_LIST', 'https://www.tripadvisor.
com/Attraction_Products-g297604-a_sort.-oa270-Goa.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g297604-a_sort.-oa300-Go
a.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products
-g297604-a_sort.-oa330-Goa.html#ATTRACTION_LIST']
```

In [197]:
```python
all_attractions_goa = get_all_attractions(goa_htmls)
print(len(all_attractions_goa))
print(all_attractions_goa[:5])
```

```
349
['Fontainhas Heritage Walk', 'Sunset Cruise on Paradise', 'North Goa Sigh
tseeing Full Day Tour ( 09 am - 05 pm )', 'Dudhsagar Waterfall Trip in Go
a', 'Water Sports Goa']
```

In [198]:
```python
goa_df = pd.DataFrame(all_attractions_goa, columns=['Attraction'])
goa_df['City'] = 'Goa, India'
goa_df.head()
```

Out[198]:

	Attraction	City
0	Fontainhas Heritage Walk	Goa, India
1	Sunset Cruise on Paradise	Goa, India
2	North Goa Sightseeing Full Day Tour (09 am - ...	Goa, India
3	Dudhsagar Waterfall Trip in Goa	Goa, India
4	Water Sports Goa	Goa, India

In [199]:
```python
goa_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project/go
```

2.1.11 Dubai, United Arab Emirates

```
In [201]:    1  # Create list of all htmls for Dubai
             2  dubai_htmls = ['https://www.tripadvisor.com/Attraction_Products-g295424
             3
             4  for i in range(2, 125):
             5      base = 'https://www.tripadvisor.com/Attraction_Products-g295424-a_s
             6      page = '-oa' + str(30*i - 30)
             7      end = '-Dubai_Emirate_of_Dubai.html#ATTRACTION_LIST'
             8      html = base + page + end
             9      dubai_htmls.append(str(html))
            10
            11  print(dubai_htmls)
```

```
['https://www.tripadvisor.com/Attraction_Products-g295424-Dubai_Emirate_o
f_Dubai.html', 'https://www.tripadvisor.com/Attraction_Products-g295424-a
_sort.-oa30-Dubai_Emirate_of_Dubai.html#ATTRACTION_LIST', 'https://www.tr
ipadvisor.com/Attraction_Products-g295424-a_sort.-oa60-Dubai_Emirate_of_D
ubai.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ
cts-g295424-a_sort.-oa90-Dubai_Emirate_of_Dubai.html#ATTRACTION_LIST', 'h
ttps://www.tripadvisor.com/Attraction_Products-g295424-a_sort.-oa120-Duba
i_Emirate_of_Dubai.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/At
traction_Products-g295424-a_sort.-oa150-Dubai_Emirate_of_Dubai.html#ATTRA
CTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g295424-a_s
ort.-oa180-Dubai_Emirate_of_Dubai.html#ATTRACTION_LIST', 'https://www.tri
padvisor.com/Attraction_Products-g295424-a_sort.-oa210-Dubai_Emirate_of_D
ubai.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/Attraction_Produ
cts-g295424-a_sort.-oa240-Dubai_Emirate_of_Dubai.html#ATTRACTION_LIST',
'https://www.tripadvisor.com/Attraction_Products-g295424-a_sort.-oa270-Du
bai_Emirate_of_Dubai.html#ATTRACTION_LIST', 'https://www.tripadvisor.com/
Attraction_Products-g295424-a_sort.-oa300-Dubai_Emirate_of_Dubai.html#ATT
RACTION_LIST', 'https://www.tripadvisor.com/Attraction_Products-g295424-a
_sort.-oa330-Dubai_Emirate_of_Dubai.html#ATTRACTION_LIST', 'https://www.t
```

```
In [202]:    1  all_attractions_dubai = get_all_attractions(dubai_htmls)
             2  print(len(all_attractions_dubai))
             3  print(all_attractions_dubai[:5])
```

```
3694
['Premium Red Dunes & Camel Safari with BBQ at Al Khayma Camp', 'Red Dune
s By Quad Bike, Sandsurf, Camels & BBQ at Al Khayma Camp ', 'Dubai Marina
Luxury Yacht Tour with Breakfast or BBQ', 'Dubai Red Dune Desert Safari:
Camel Ride, Sandboarding & BBQ Options', 'Atlantis Aquaventure & Lost Cha
mbers Ticket']
```

In [203]:
```python
1  dubai_df = pd.DataFrame(all_attractions_dubai, columns=['Attraction'])
2  dubai_df['City'] = 'Dubai, United Arab Emirates'
3  dubai_df.head()
```

Out[203]:

	Attraction	City
0	Premium Red Dunes & Camel Safari with BBQ at A...	Dubai, United Arab Emirates
1	Red Dunes By Quad Bike, Sandsurf, Camels & BBQ...	Dubai, United Arab Emirates
2	Dubai Marina Luxury Yacht Tour with Breakfast ...	Dubai, United Arab Emirates
3	Dubai Red Dune Desert Safari: Camel Ride, Sand...	Dubai, United Arab Emirates
4	Atlantis Aquaventure & Lost Chambers Ticket	Dubai, United Arab Emirates

In [204]:
```python
1  dubai_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project/
```

▼ 2.2 Concatenate all the DFs

In [207]:
```python
1  cities_list = [london_df, paris_df, crete_df, bali_df, rome_df, phuket_
2                 sicily_df, majorca_df, barcelona_df, istanbul_df, goa_df
3  cities_df = pd.concat(cities_list)
4  cities_df.head()
```

Out[207]:

	Attraction	City
0	SEA LIFE London Aquarium Admission Ticket	London, United Kingdom
1	The Jack The Ripper Walking Tour in London	London, United Kingdom
2	Ghost Bus Tour of London	London, United Kingdom
3	Big Bus London Hop-On Hop-Off Tour and River C...	London, United Kingdom
4	The Blood and Tears Walk: Serial Killers and L...	London, United Kingdom

In [208]:
```python
1  # See all 12 cities in the DF
2  cities_df['City'].unique()
```

Out[208]:
```
array(['London, United Kingdom', 'Paris, France', 'Crete, Greece',
       'Bali, Indonesia', 'Rome, Italy', 'Phuket, Thailand',
       'Sicily, Italy', 'Majorca, Balearic Islands', 'Barcelona, Spain',
       'Istanbul, Turkey', 'Goa, India', 'Dubai, United Arab Emirates'],
      dtype=object)
```

In [209]:
```python
1  # Save to CSV
2  cities_df.to_csv('/Users/tiaplagata/Documents/Flatiron/capstone-project
```

▼ 2.3 Analysis continued in [EDA and Modeling Notebook (https://github.com/tiaplagata/capstone-project/blob/main/Notebooks/EDA_and_Modeling.ipynb)](https://github.com/tiaplagata/capstone-project/blob/main/Notebooks/EDA_and_Modeling.ipynb)