**Yelp Reviews: Data Mining using Yelp FUSION API**

In [52]:

**import** **pandas** **as** **pd**  
**import** **requests**

**Creating Request to Search for Physicians on Yelp.com**[**¶**](#gjdgxs)

You will require an API Key to extract data, used two functions: request() and getreviews()

To create an API Key, do the following:

* sign up on Yelp.com to create an account, ignore if already have an account
* Login to Yelp.com
* Go to Create App
* In the create new app form, enter dummy information about your app(App Name, Industry, Email, Description), then agree to Yelp API Terms of Use and Display Requirements. Then click the Submit button.
* You will now have an API Key.
* Replace the Fake API Key in both functions with your API Key

In [41]:

**def** request():  
 *"""Given your API\_KEY, send a GET request to the API.*  
 *Args:*  
 *url (str): The path of the API with the domain.*  
 *api\_key (str): Your API Key.*  
 *url\_params (dict): An optional set of query parameters in the request.*  
 *Returns:*  
 *dict: The JSON response from the request.*  
 *Raises:*  
 *HTTPError: An error occurs from the HTTP request.*  
 *"""*  
 api\_key = "ajndlajsnashjksdbjknsdghjklljfghjkldfghjklgfdghjkhgfdghjkhgfdhhg" *# Fake API Key*  
 url = "https://api.yelp.com/v3/businesses/search"  
 headers = {  
 'Authorization': 'Bearer **%s**' % api\_key,  
 }  
 url\_params = {  
 'term': "physician".replace(' ', '+'),  
 'location': "New York, NY".replace(' ', '+'),  
 'limit': '50'  
 }  
  
 print(u'Querying **{0}** ...'.format(url))  
  
 response = requests.request('GET', url, headers=headers, params=url\_params)  
  
 **return** response.json()

In [43]:

yelp\_data = request()

Querying https://api.yelp.com/v3/businesses/search ...

**Get Reviews for Each Physician using Yelp Fusion API**[**¶**](#30j0zll)

In [68]:

**def** get\_reviews(business\_id):  
 *"""Given your API\_KEY, send a GET request to the API.*  
 *Args:*  
 *url (str): The path of the API with the domain.*  
 *api\_key (str): Your API Key.*  
 *Returns:*  
 *dict: The JSON response from the request.*  
 *Raises:*  
 *HTTPError: An error occurs from the HTTP request.*  
 *"""*  
 url = "https://api.yelp.com/v3/businesses/**{buss\_id}**/reviews".format(buss\_id=business\_id)  
 api\_key = "ajndlajsnashjksdbjknsdghjklljfghjkldfghjklgfdghjkhgfdghjkhgfdhhg" *# Fake API Key*  
  
 headers = {  
 'Authorization': 'Bearer **%s**' % api\_key,  
 }  
 print(u'Querying **{0}** ...'.format(url))  
  
 response = requests.request('GET', url, headers=headers)  
 **return** response.json()

**Filtering Required Data into Pandas DataFrame**[**¶**](#1fob9te)

In [143]:

yelp\_dataframe = pd.DataFrame()  
**for** data **in** yelp\_data["businesses"]:  
 avg\_rating = data["rating"]  
 title = data["categories"][0]["title"]  
 alias = data["categories"][0]["alias"]  
 city = data["location"]["city"]  
 country = data["location"]["country"]  
 address = ", ".join(data["location"]["display\_address"])  
 state = data["location"]["state"]  
 zip\_code = data["location"]["zip\_code"]  
 name = data["name"]  
 buss\_id = data["id"]  
 response\_dict = get\_reviews(buss\_id)  
 **if** response\_dict:  
 reviews\_list = response\_dict["reviews"]  
 **for** review **in** reviews\_list:  
 review["user"] = review["user"]["name"]  
 review["physician"] = name  
 review["avg\_rating"] = avg\_rating  
 review["title"] = title  
 review["alias"] = alias  
 review["city"] = city  
 review["country"] = country  
 review["address"] = address  
 review["state"] = state  
 review["zip\_code"] = zip\_code  
 yelp\_dataframe = yelp\_dataframe.append(reviews\_list)

Querying https://api.yelp.com/v3/businesses/laura-guderian-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/mohawk-village-medical-new-york-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/michael-raffinan-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/inna-grishina-md-brooklyn-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/xiaoyan-song-md-flushing/reviews ...  
Querying https://api.yelp.com/v3/businesses/bruce-chung-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/weill-cornell-medicine-primary-care-west-side-new-york-3/reviews ...  
Querying https://api.yelp.com/v3/businesses/socrates-t-kangadis-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/karis-cho-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/jason-h-dolinsky-md-new-york-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/michelle-li-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/aldo-alleva-md-brooklyn-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/alan-pinyavat-md-manhattan/reviews ...  
Querying https://api.yelp.com/v3/businesses/sally-h-kwa-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/amy-bleyer-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/lam-terry-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/zina-kroner-do-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/viplov-mehta-md-brooklyn/reviews ...  
Querying https://api.yelp.com/v3/businesses/tammy-leopold-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/shum-chi-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/leaf-medical-brooklyn-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/louis-c-barricelli-md-brooklyn/reviews ...  
Querying https://api.yelp.com/v3/businesses/michael-zuckman-md-sunnyside/reviews ...  
Querying https://api.yelp.com/v3/businesses/paul-knoepflmacher-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/smart-health-medical-elmhurst/reviews ...  
Querying https://api.yelp.com/v3/businesses/bailyn-douglas-h-md-new-york-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/elena-klimenko-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/pei-gao-md-queens/reviews ...  
Querying https://api.yelp.com/v3/businesses/scafuri-frank-md-iii-staten-island/reviews ...  
Querying https://api.yelp.com/v3/businesses/allison-fox-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/astrahealth-urgent-and-primary-care-hoboken-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/lola-iskhakova-md-ozone-park/reviews ...  
Querying https://api.yelp.com/v3/businesses/ann-jin-qui-md-manhattan/reviews ...  
Querying https://api.yelp.com/v3/businesses/neighborhood-natural-medicine-brooklyn-2/reviews ...  
Querying https://api.yelp.com/v3/businesses/glasser-jeffrey-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/forest-hills-medical-services-forest-hills/reviews ...  
Querying https://api.yelp.com/v3/businesses/gerald-h-leventhal-md-pc-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/st-claire-medical-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/moitri-savard-md-queens/reviews ...  
Querying https://api.yelp.com/v3/businesses/gentile-ronald-dr-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/dr-melanie-joannow-ditmars-steinway/reviews ...  
Querying https://api.yelp.com/v3/businesses/macaulay-grant-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/westside-family-medicine-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/zhanna-tsukerman-md-brooklyn/reviews ...  
Querying https://api.yelp.com/v3/businesses/arturo-constantiner-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/david-r-kaplowitz-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/rebecca-press-md-mph-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/alexander-shifrin-md-brooklyn/reviews ...  
Querying https://api.yelp.com/v3/businesses/jeffrey-m-engel-md-new-york/reviews ...  
Querying https://api.yelp.com/v3/businesses/alexander-krishtul-md-new-york/reviews ...

In [144]:

yelp\_dataframe.reset\_index(inplace=**True**)

In [145]:

yelp\_dataframe.drop('index', axis=1, inplace=**True**)

In [148]:

yelp\_dataframe.head(2)

Out[148]:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **address** | **alias** | **avg\_rating** | **city** | **country** | **physician** | **rating** | **state** | **text** | **time\_created** | **title** | **url** | **user** | **zip\_code** |
| **0** | 35 E 21st St, Fl 7, New York, NY 10010 | internalmed | 4.5 | New York | US | Laura Guderian, MD | 5 | NY | I've been delighted with Dr. Guderian as my pr... | 2017-09-06 13:21:59 | Internal Medicine | https://www.yelp.com/biz/laura-guderian-md-new... | Bonnie E. | 10010 |
| **1** | 35 E 21st St, Fl 7, New York, NY 10010 | internalmed | 4.5 | New York | US | Laura Guderian, MD | 1 | NY | I really wanted to like One Medical Group. I a... | 2016-09-25 16:00:50 | Internal Medicine | https://www.yelp.com/biz/laura-guderian-md-new... | Yelpish N. | 10010 |

**Exporting Pandas DataFrame to Local CSV**[**¶**](#3znysh7)

In [149]:

yelp\_dataframe.to\_csv("yelp\_reviews.csv", index=**False**, encoding="UTF-8")