

Kinds_in_charlotte

```
import requests
import MySQLdb
import re
import string
from bs4 import BeautifulSoup

# Open database connection
db = MySQLdb.connect("localhost","root","","kids_services" )

# prepare a cursor object using cursor() method
cursor = db.cursor()

given_url = "http://www.kidsincharlotte.com/"

r = requests.get(given_url)

soup = BeautifulSoup(r.content,'lxml')

csv_data = ''

for a in soup.find_all('a', attrs={"class": "ICONS"}):
    #This is to get the child page html
    child_url = requests.get(given_url+a['href'])
    #This is to get the soup ready
    child_url_soup = BeautifulSoup(child_url.content)
    print "Parent URL:", a['href']
    #Get all the child page url
    for child_a in child_url_soup.find_all('a', attrs={"class":
"ICONS"}):
        if '../other' not in child_a['href']:
            first_part_of_parent_link = a['href'].split('/')
            #This is to get the child page html
            second_child_url =
requests.get(given_url+first_part_of_parent_link[0]+'/' +child_a['href']
)
            #This is to get the soup ready
            second_child_url_soup =
BeautifulSoup(second_child_url.content)

            for data in second_child_url_soup.find_all('p',
attrs={"class": "DATATEXT"}):
                for second_data in
str(data).split('<br/>'):
                    if ('<a href="#Top">' not in
second_data) and (second_data) and ('<p class="DATATEXT"></p>' not in
second_data):
                        data_after_strip_tags =
re.sub('<[p|b|/][^>]*>', '', second_data)

                        if
len(data_after_strip_tags) > 1:
                            if 'View Map' in
data_after_strip_tags:
                                csv_data =
                                csv_data+data_after_strip_tags+'\n'
                            else:
```

```
csv_data+data_after_strip_tags+','
```

```
csv_data =
```

```
text_file = open("Output.txt", "w")  
text_file.write("%s" % csv_data)  
text_file.close()
```