In [ ]:
```
# TODO I will no longer be updating this project since I'm not using it anymore
import requests
from bs4 import BeautifulSoup
import csv
import json
from operator import itemgetter
import pandas as pd
import re
```

In [ ]:
```
URL = "https://housing.uga.edu/explore-options/"
page = requests.get(URL)
# exploreHall hosts name
#explorePhoto hosts photo
soup = BeautifulSoup(page.content, "html.parser")
```

In [ ]:
```
hallList = []
for item in soup.find_all('div', id='exploreHall'):
    item = item.get_text()
    item = str(item)
    hallList.append(item)


# multiple hall names, internal text.
URL_INTERNAL = []

for i in range(len(hallList)):
    URL_INTERNAL.append("https://housing.uga.edu/explore-options/" +
                        hallList[i].strip().replace(' ', '-') + "/")
```

In [ ]:
```

```

In [ ]:
```
#don't touch under, all of these work already.

substring = "/sa_images/featured/"
photoList = []
# ,id="explorePhoto"
for picture in soup.find_all("img"):
    #picture = str(picture)
    if substring in str(picture):
        picture = picture['src']
        photoList.append(picture)
```

In [ ]:
```
res = {}
for key in hallList:
    for value in photoList:
        res[key] = value
        photoList.remove(value)
        break
```

In [ ]:
```
# with open('some.csv', 'w') as csv_file:
#     writer = csv.writer(csv_file)
#     for key, value in res.items():
#        writer.writerow([key, value])
```

In [ ]:
```
# csvfile = open('some.csv', 'r')
# jsonfile = open('file.json', 'w')
```

In [ ]:
```
# fieldnames = ("Name of Hall ", "Photo of Hall ")
# reader = csv.DictReader(csvfile, fieldnames)
# for row in reader:
#     json.dump(row, jsonfile)
#     jsonfile.write('\n')

# jsonfile.close()
```

In [ ]:
```
# for each key in hallist, go to the link of the hall, and inner information (t
inner_information = ["Visitation", "Open for Breaks", "Bus Stops", "Parking Lot
```

In [ ]:
```
# this is the ACTUAL information for each hall
spans = []

for i in range(len(URL_INTERNAL)):
    page_inside = requests.get(URL_INTERNAL[i])
    soup_inside = BeautifulSoup(page_inside.content, "html.parser")
    for i in range(len(URL_INTERNAL)):
        spans.append(soup_inside.find_all('span', {'class': 'nearby-right'}))
```

In [ ]:
```
#each value in spans add it to df
new = []
for i in range(len(spans)):
    for j in range(len(spans[i])):
        new.insert(j,soup_inside.find_all(
            'span', {'class': 'nearby-right'})[j].get_text())
```

In [ ]:
```
df = pd.DataFrame(columns=hallList, index=inner_information)
```

In [ ]:
```
# fill in the dataframe with values from new

for i in range(len(new)):
    df.loc[i] = new[i]

df
```