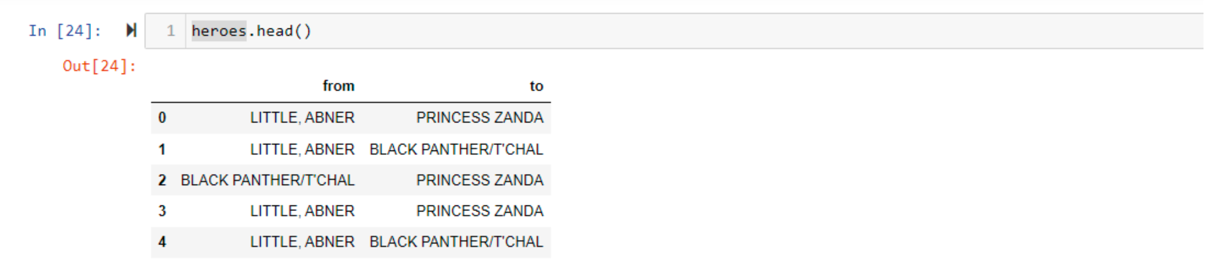
Assignment 6

I have use pandas libraries to import the dataframe used dataframe with heading ‘from’ and ‘to’ to make connections between marvel characters

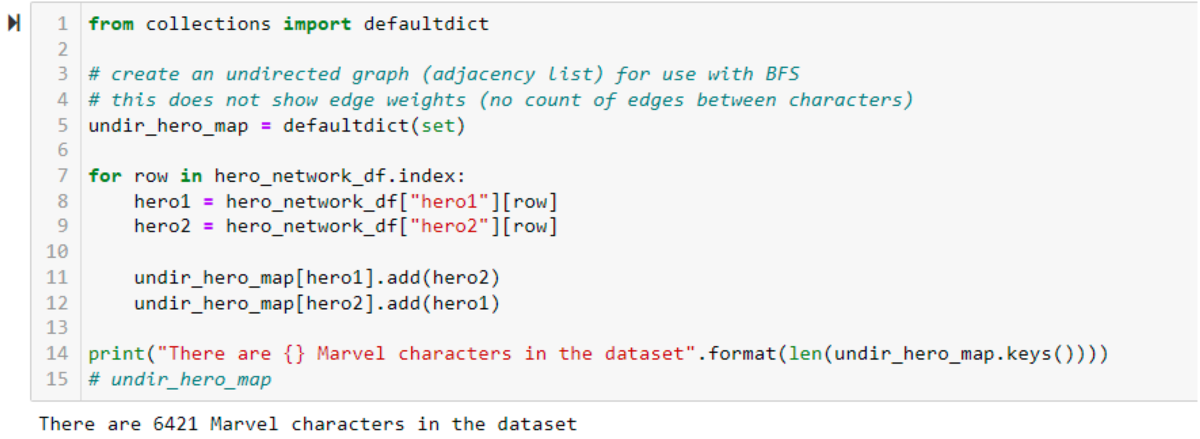

In [24]:

```
heroes.head()
```

Out[24]:

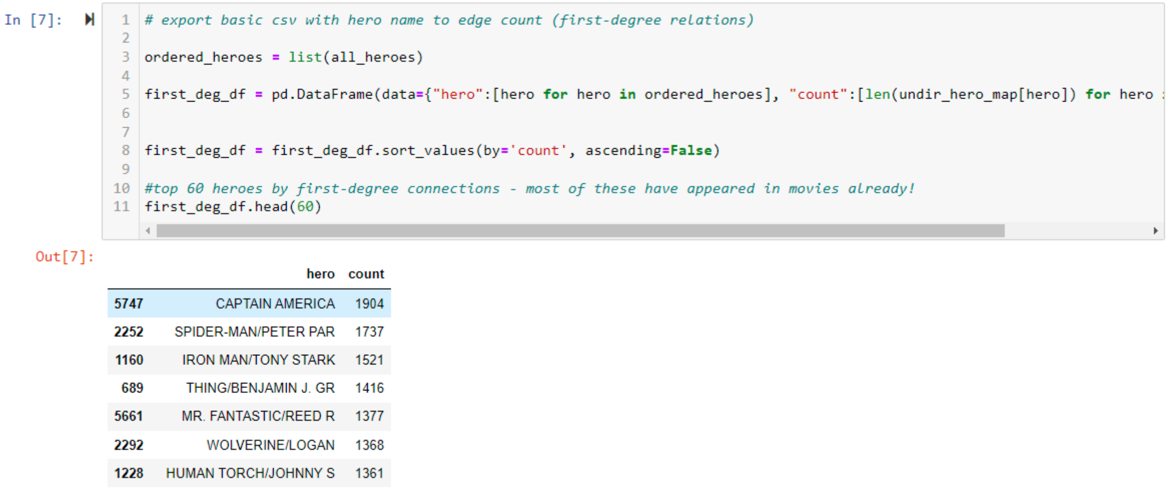
	from	to
0	LITTLE, ABNER	PRINCESS ZANDA
1	LITTLE, ABNER	BLACK PANTHER/T'CHAL
2	BLACK PANTHER/T'CHAL	PRINCESS ZANDA
3	LITTLE, ABNER	PRINCESS ZANDA
4	LITTLE, ABNER	BLACK PANTHER/T'CHAL

Here we try to make adjacency list a type of dictionary for marvel characters and there is total 6421 characters.


```
from collections import defaultdict

# create an undirected graph (adjacency list) for use with BFS
# this does not show edge weights (no count of edges between characters)
undir_hero_map = defaultdict(set)

for row in hero_network_df.index:
    hero1 = hero_network_df["hero1"][row]
    hero2 = hero_network_df["hero2"][row]

    undir_hero_map[hero1].add(hero2)
    undir_hero_map[hero2].add(hero1)

print("There are {} Marvel characters in the dataset".format(len(undir_hero_map.keys())))
# undir_hero_map
```

```
There are 6421 Marvel characters in the dataset
```

We order the dictionary in descending order and it shows that Captain America and Spider Man as expected are the characters with most interactions.

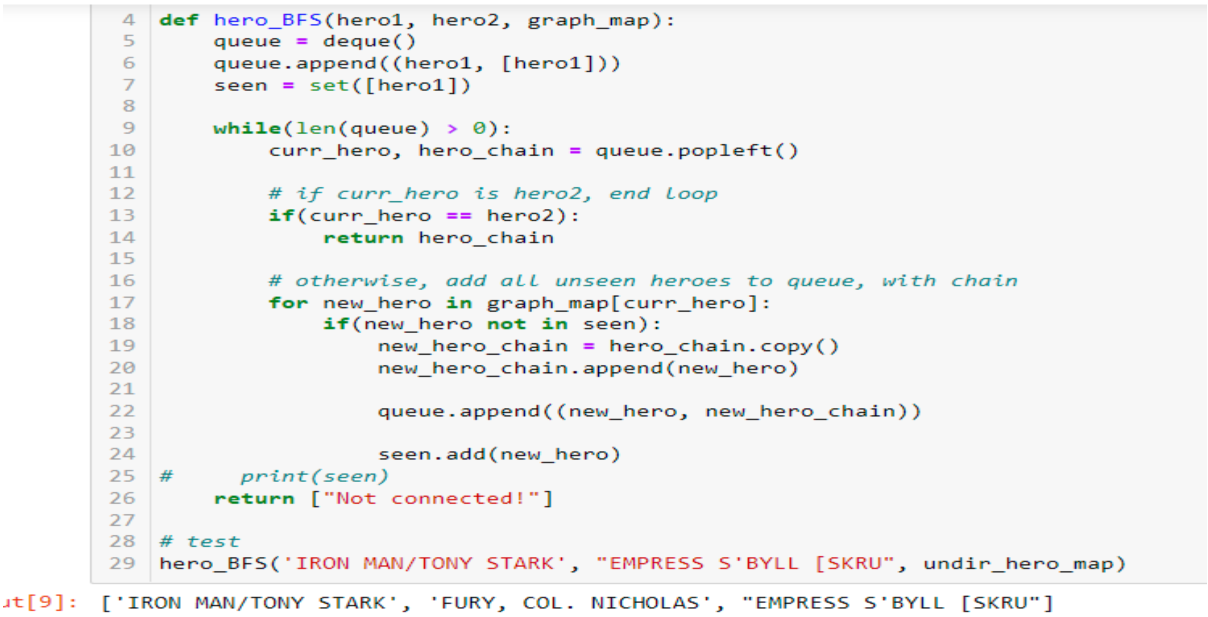

In [7]:

```
# export basic csv with hero name to edge count (first-degree relations)

ordered_heroes = list(all_heroes)

first_deg_df = pd.DataFrame(data={"hero":[hero for hero in ordered_heroes], "count":[len(undir_hero_map[hero]) for hero


first_deg_df = first_deg_df.sort_values(by='count', ascending=False)

#top 60 heroes by first-degree connections - most of these have appeared in movies already!
first_deg_df.head(60)
```

Out[7]:

	hero	count
5747	CAPTAIN AMERICA	1904
2252	SPIDER-MAN/PETER PAR	1737
1160	IRON MAN/TONY STARK	1521
689	THING/BENJAMIN J. GR	1416
5661	MR. FANTASTIC/REED R	1377
2292	WOLVERINE/LOGAN	1368
1228	HUMAN TORCH/JOHNNY S	1361

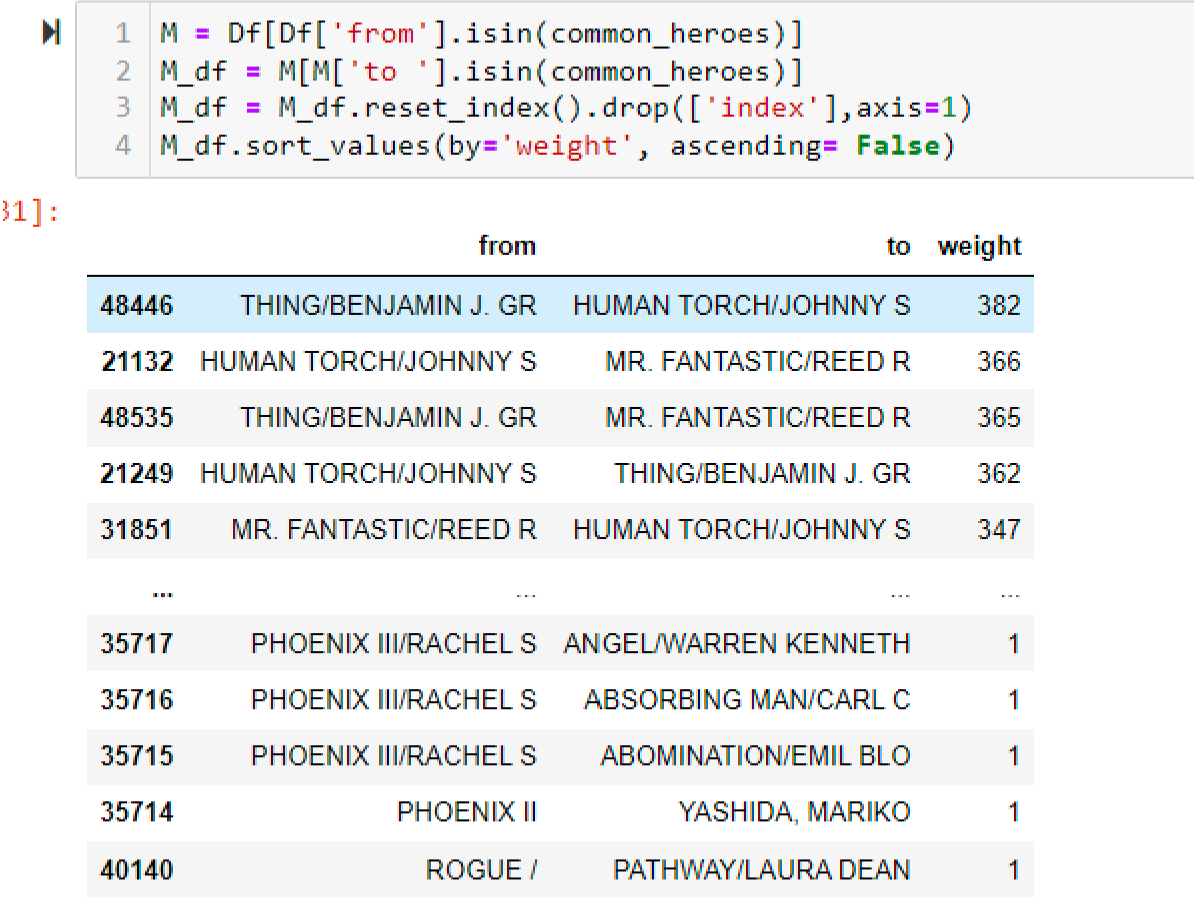
Using breadth first search we run through the dictionary o find connections between character for example between iron man and empress and find that a connection is there with fury in between.



```python
def hero_BFS(hero1, hero2, graph_map):
    queue = deque()
    queue.append((hero1, [hero1]))
    seen = set([hero1])

    while(len(queue) > 0):
        curr_hero, hero_chain = queue.popleft()

        # if curr_hero is hero2, end loop
        if(curr_hero == hero2):
            return hero_chain

        # otherwise, add all unseen heroes to queue, with chain
        for new_hero in graph_map[curr_hero]:
            if(new_hero not in seen):
                new_hero_chain = hero_chain.copy()
                new_hero_chain.append(new_hero)

                queue.append((new_hero, new_hero_chain))

                seen.add(new_hero)
#     print(seen)
    return ["Not connected!"]

# test
hero_BFS('IRON MAN/TONY STARK', "EMPRESS S'BYLL [SKRU", undir_hero_map)
```

```
ut[9]:  ['IRON MAN/TONY STARK', 'FURY, COL. NICHOLAS', "EMPRESS S'BYLL [SKRU"]
```

Here we find the interactions between character and how many times they have occurred (weights.

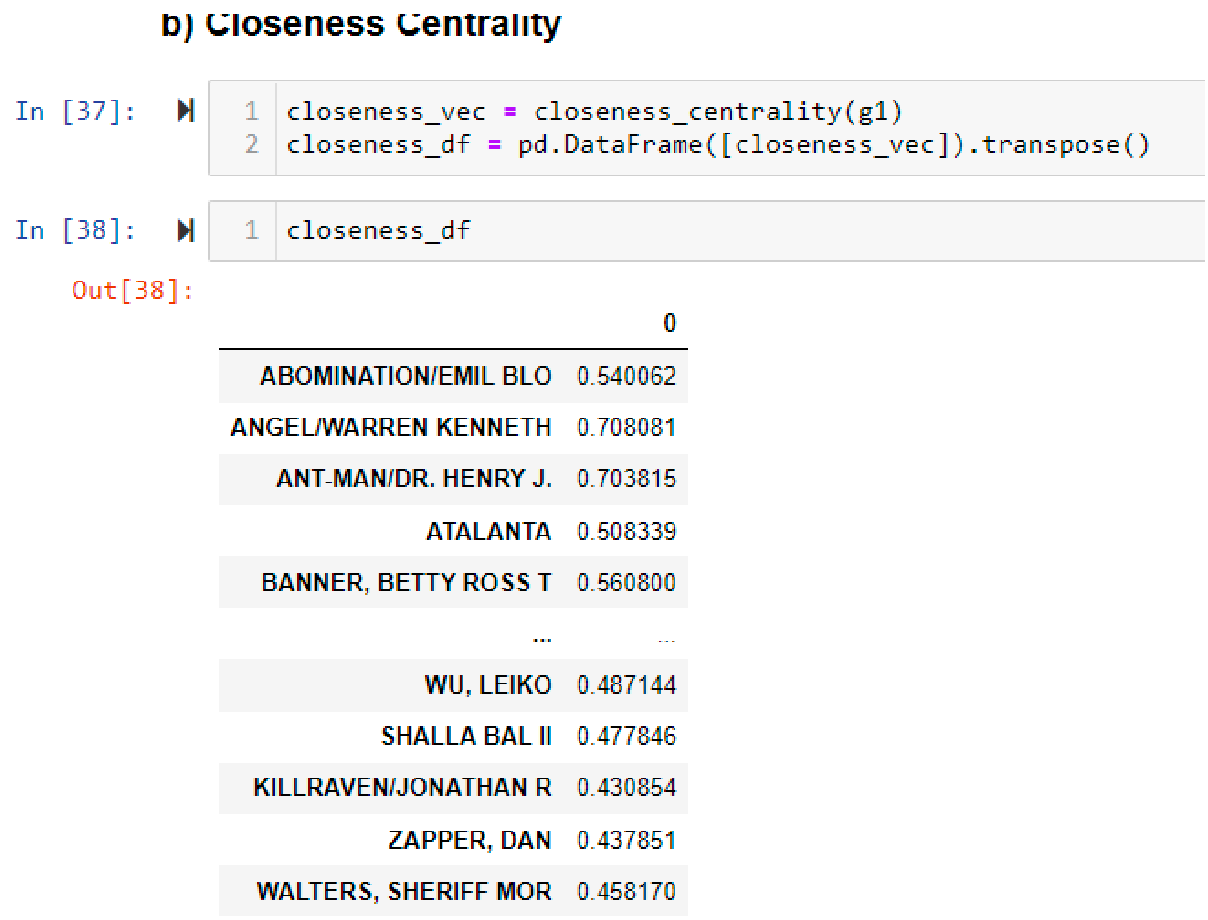


```python
M = Df[Df['from'].isin(common_heroes)]
M_df = M[M['to '].isin(common_heroes)]
M_df = M_df.reset_index().drop(['index'],axis=1)
M_df.sort_values(by='weight', ascending= False)
```

31]:

	from	to	weight
48446	THING/BENJAMIN J. GR	HUMAN TORCH/JOHNNY S	382
21132	HUMAN TORCH/JOHNNY S	MR. FANTASTIC/REED R	366
48535	THING/BENJAMIN J. GR	MR. FANTASTIC/REED R	365
21249	HUMAN TORCH/JOHNNY S	THING/BENJAMIN J. GR	362
31851	MR. FANTASTIC/REED R	HUMAN TORCH/JOHNNY S	347
...	...	...	...
35717	PHOENIX III/RACHEL S	ANGEL/WARREN KENNETH	1
35716	PHOENIX III/RACHEL S	ABSORBING MAN/CARL C	1
35715	PHOENIX III/RACHEL S	ABOMINATION/EMIL BLO	1
35714	PHOENIX II	YASHIDA, MARIKO	1
40140	ROGUE /	PATHWAY/LAURA DEAN	1

Here we find the closeness centrality using the centrality library.



b) Closeness Centrality

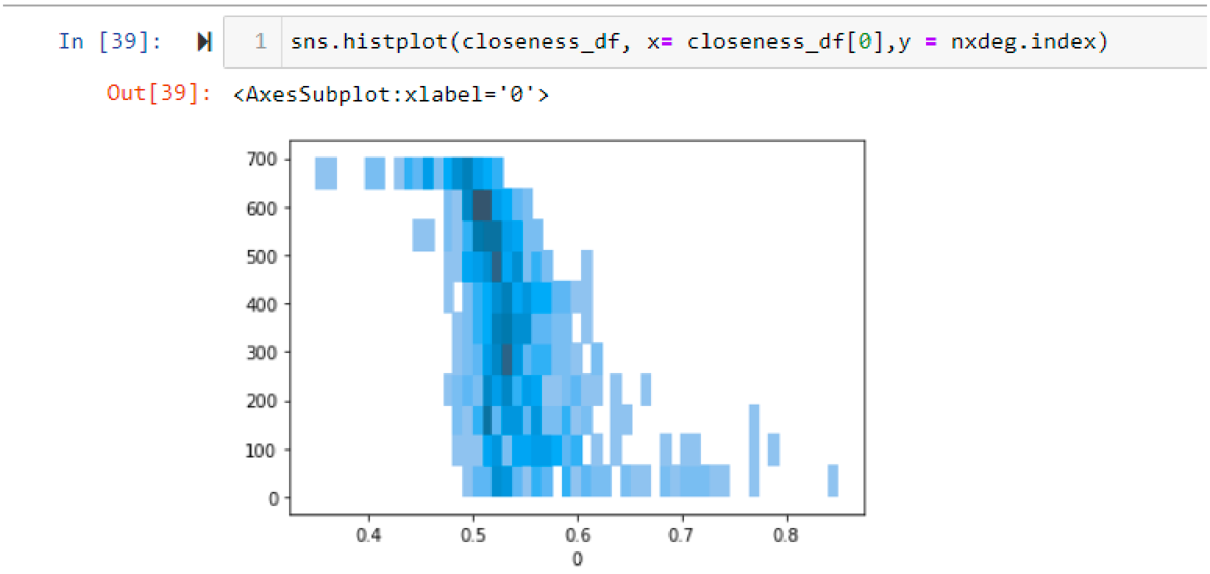
```
In [37]:  closeness_vec = closeness_centrality(g1)
          closeness_df = pd.DataFrame([closeness_vec]).transpose()
```

```
In [38]:  closeness_df
```

Out[38]:

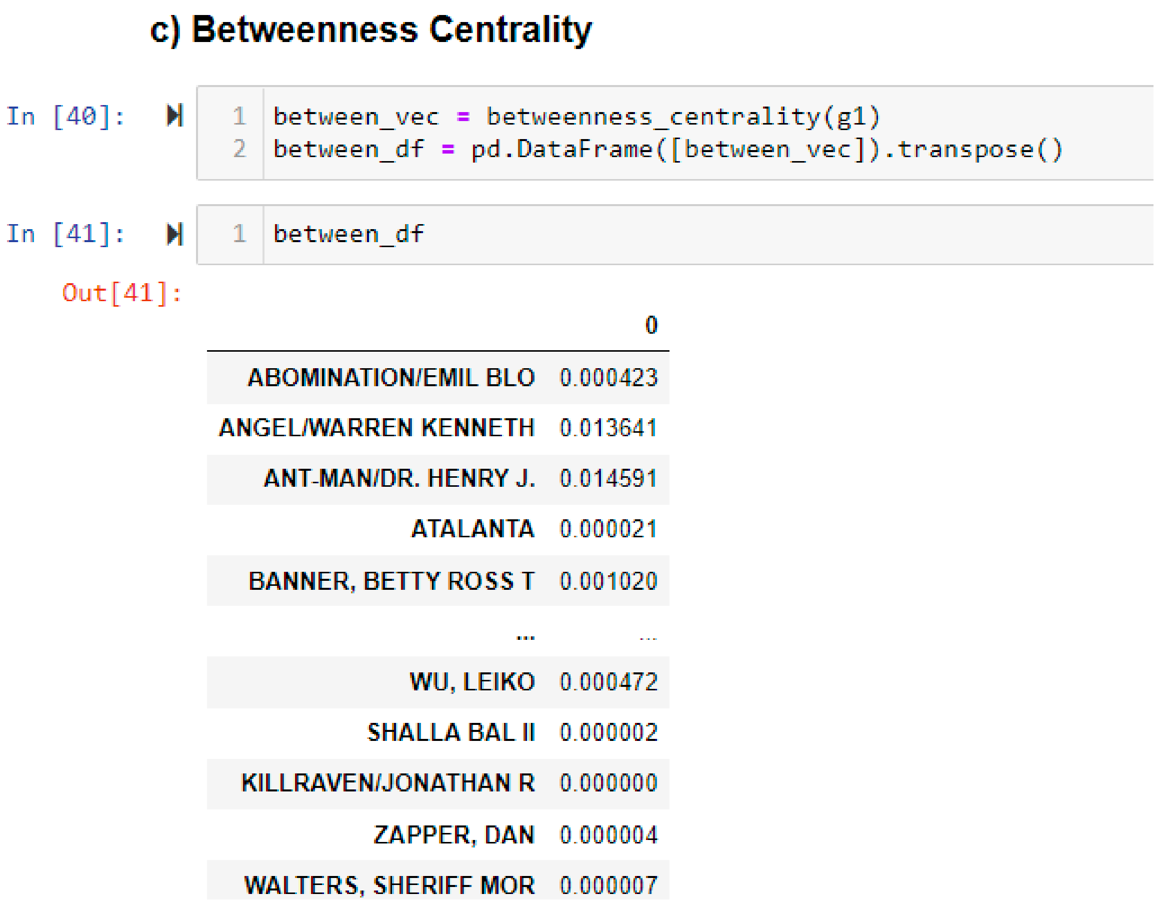
	0
ABOMINATION/EMIL BLO	0.540062
ANGEL/WARREN KENNETH	0.708081
ANT-MAN/DR. HENRY J.	0.703815
ATALANTA	0.508339
BANNER, BETTY ROSS T	0.560800
...	...
WU, LEIKO	0.487144
SHALLA BAL II	0.477846
KILLRAVEN/JONATHAN R	0.430854
ZAPPER, DAN	0.437851
WALTERS, SHERIFF MOR	0.458170

Here is the graph showing a lot characters are located in between 0.5 and 0.6 range.



```
In [39]:  sns.histplot(closeness_df, x= closeness_df[0],y = nxdeg.index)
Out[39]: <AxesSubplot:xlabel='0'>
```

Here we find the betweenness centrality using the centrality library.



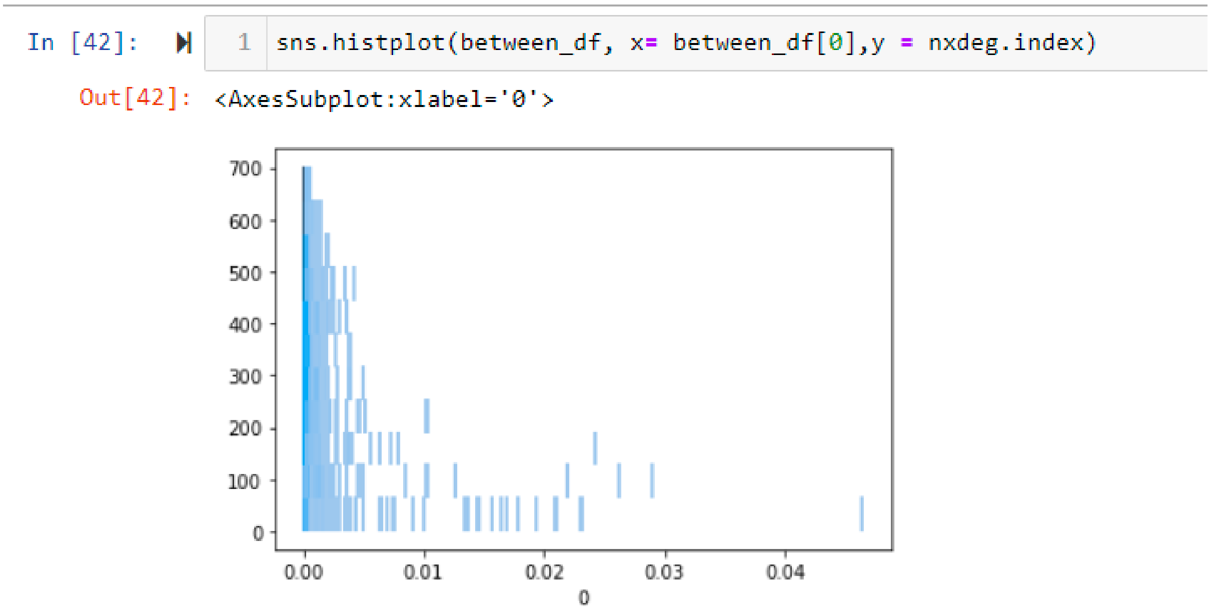
c) Betweenness Centrality

```
In [40]:
between_vec = betweenness_centrality(g1)
between_df = pd.DataFrame([between_vec]).transpose()
```

```
In [41]:
between_df
```

```
Out[41]:
                                  0
ABOMINATION/EMIL BLO       0.000423
ANGEL/WARREN KENNETH       0.013641
ANT-MAN/DR. HENRY J.       0.014591
            ATALANTA       0.000021
BANNER, BETTY ROSS T       0.001020
                 ...            ...
           WU, LEIKO       0.000472
       SHALLA BAL II       0.000002
KILLRAVEN/JONATHAN R       0.000000
         ZAPPER, DAN       0.000004
WALTERS, SHERIFF MOR       0.000007
```

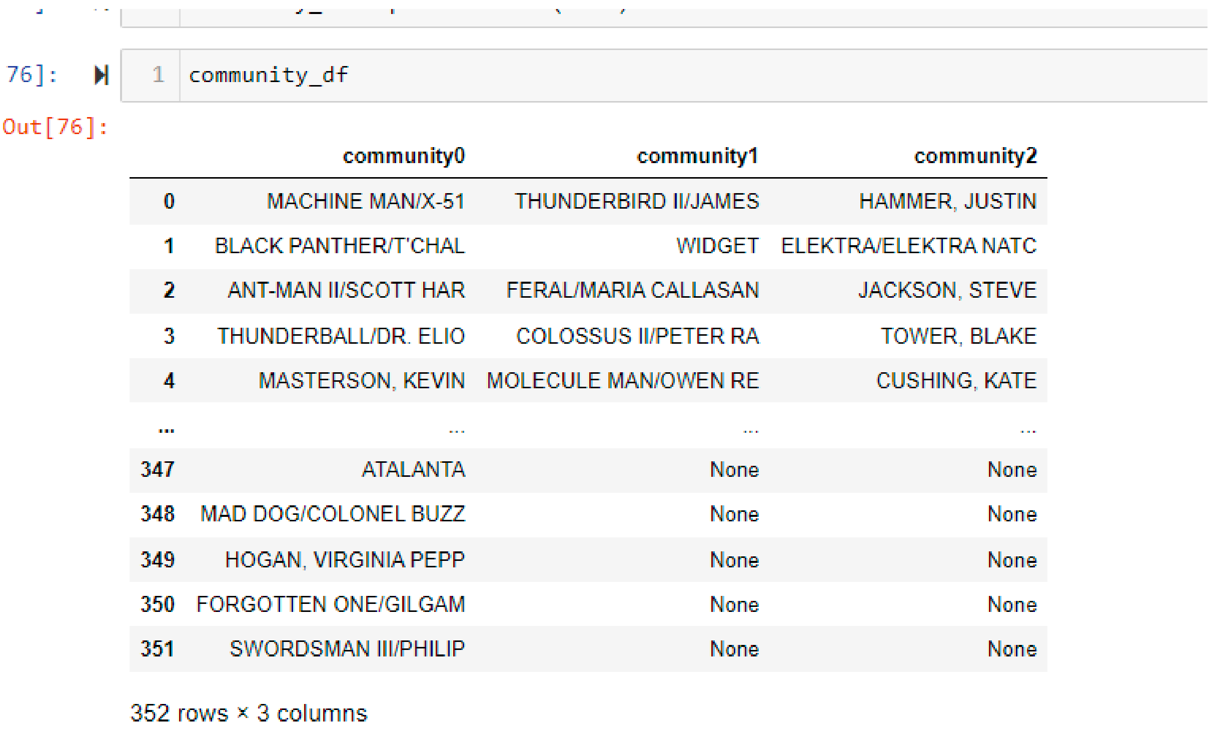
Here is the graph showing a lot characters are located in between 0.0 and 0.01 range.



```
In [42]:
sns.histplot(between_df, x= between_df[0],y = nxdeg.index)
```

```
Out[42]: <AxesSubplot:xlabel='0'>
```

Here we find try find the communities the marvel character has built and they have made 3 communities which means in each communities the characters have more likely chance to meet its own community member and if a member meets another community member it would be crossover.



```
76]:   community_df
```

Out[76]:

	community0	community1	community2
0	MACHINE MAN/X-51	THUNDERBIRD II/JAMES	HAMMER, JUSTIN
1	BLACK PANTHER/T'CHAL	WIDGET	ELEKTRA/ELEKTRA NATC
2	ANT-MAN II/SCOTT HAR	FERAL/MARIA CALLASAN	JACKSON, STEVE
3	THUNDERBALL/DR. ELIO	COLOSSUS II/PETER RA	TOWER, BLAKE
4	MASTERSON, KEVIN	MOLECULE MAN/OWEN RE	CUSHING, KATE
...	...	...	...
347	ATALANTA	None	None
348	MAD DOG/COLONEL BUZZ	None	None
349	HOGAN, VIRGINIA PEPP	None	None
350	FORGOTTEN ONE/GILGAM	None	None
351	SWORDSMAN III/PHILIP	None	None

352 rows × 3 columns