

## Projeto 1- Teoria e Aplicação de grafos

Erick Hideki Taira – 222011525

Matheus Chagas Lopes – 222011599

Determine the connected components of this network and reduce the network to its largest connected component.

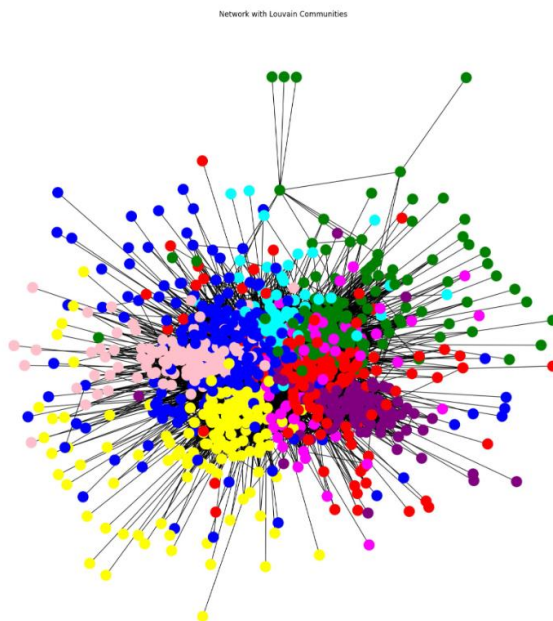
```
In [3]: #Componentes conectados
# Encontrar os componentes conectados
componentes_conectados = list(nx.connected_components(graph))

print("Componentes conectados:", len(componentes_conectados))
```

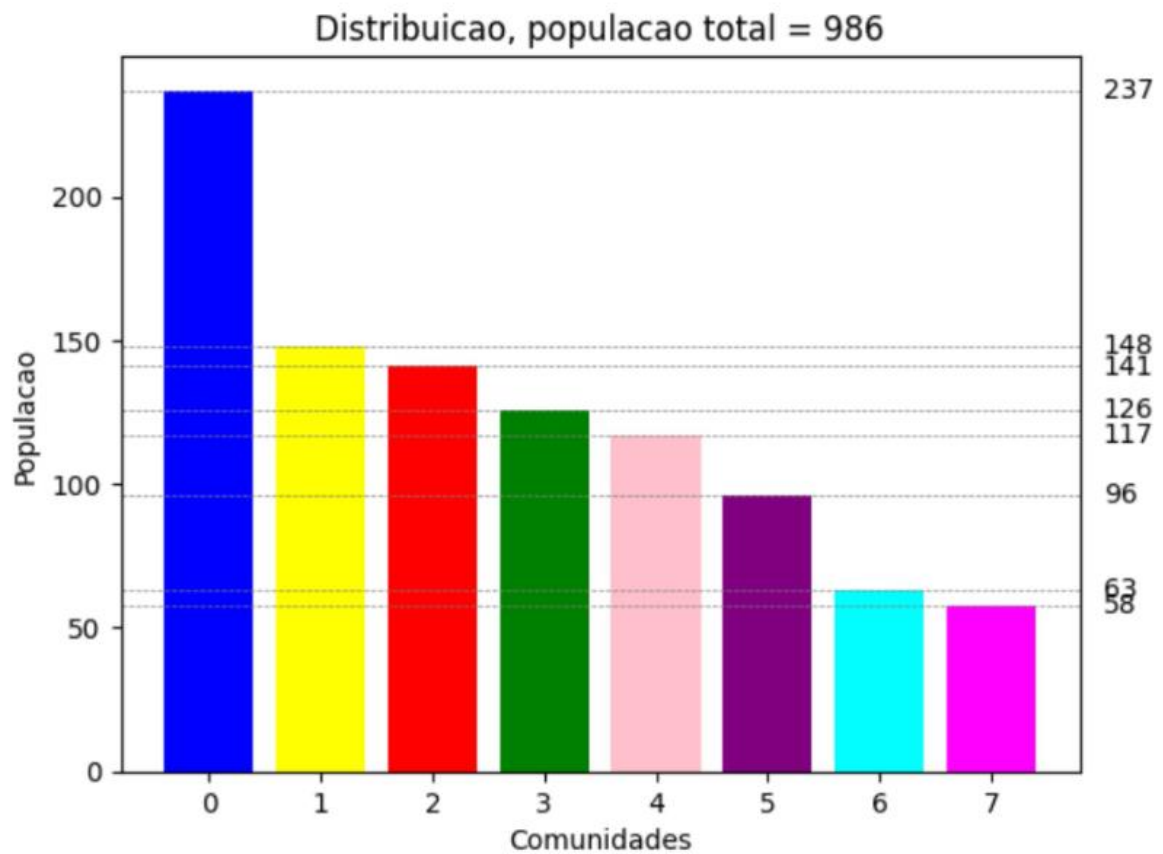
Componentes conectados: 1

“-Há apenas um componente conectado no grafo de edgelist.”

Use the Louvain algorithm to determine a vertex partition/community structure with optimal modularity in this network.



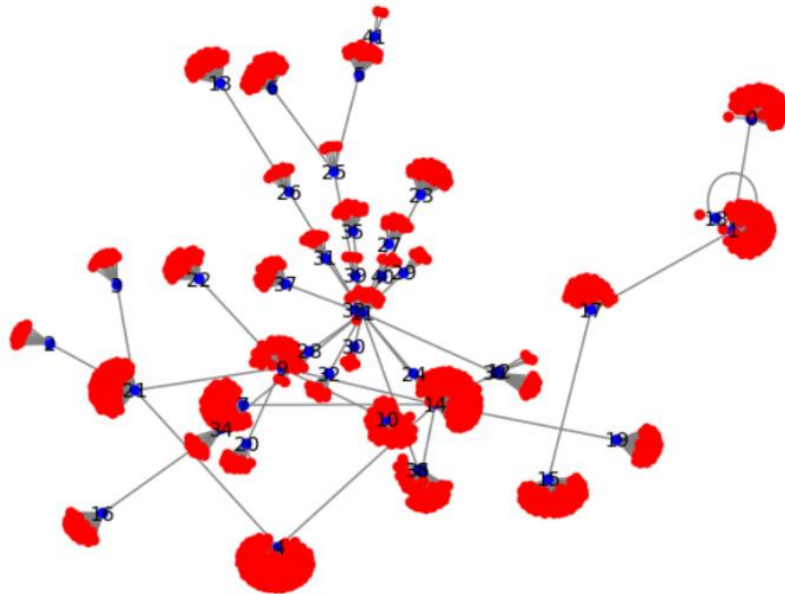
“-ao aplicar o Louvain Algorithm o grafo foi separado em comunidades de diferentes cores.”



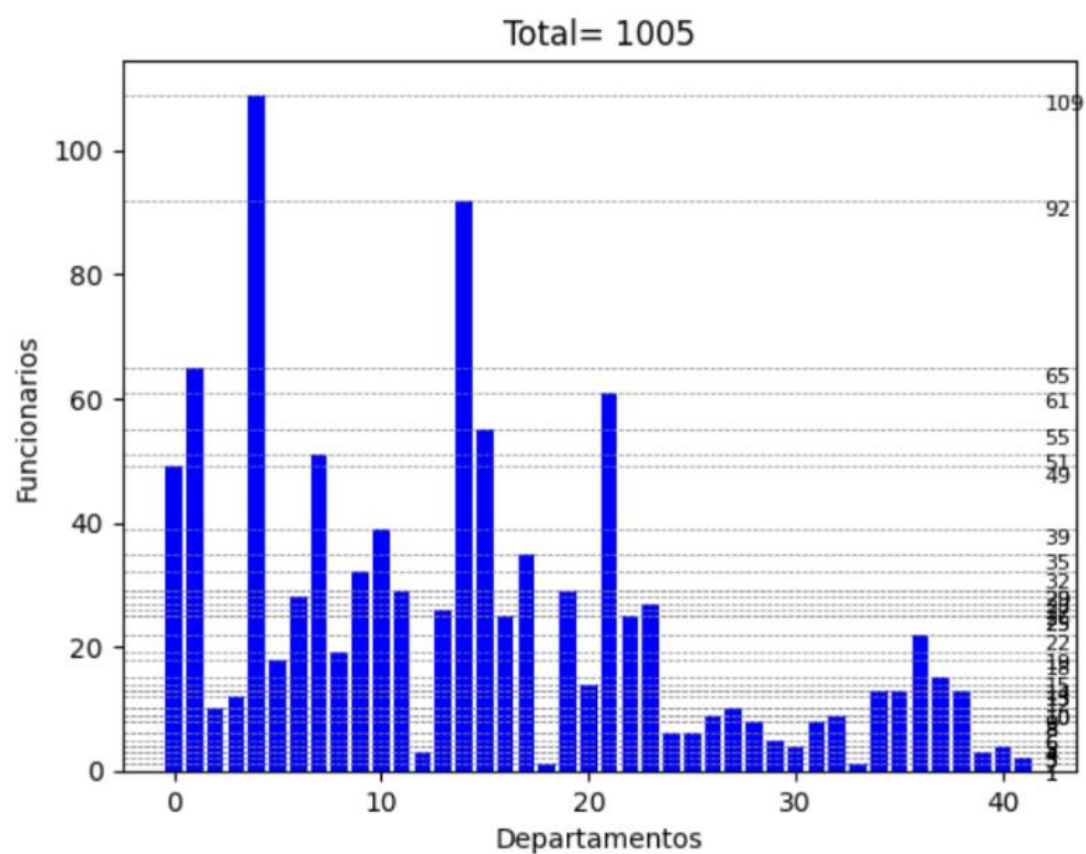
“-Distribuição dos funcionários por comunidade, as cores da tabela representam os nodes do grafo.”

Visualize the graph color-coded by the Louvain community, and then visualize the graph separately color-coded by the ground truth department. Compare the

visualizations. Can you describe any of the Louvain communities in terms of departments?

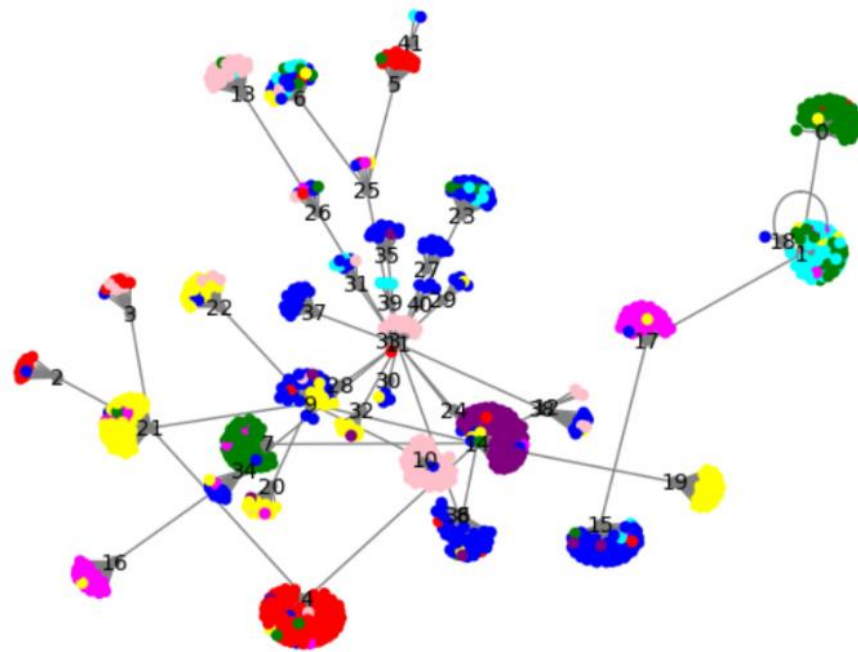


“-Grafo do ground\_truth, os nós em azul representam o departamento e os nós vermelhos os funcionários associados a esses departamentos. Não é possível visualmente identificar alguma relação com esse grafo com o grafo anteriormente mostrado.”



“-Distribuição dos funcionários.”

Create a dataframe containing the community and department for each vertex. Manipulate this dataframe to show the percentage of individuals from each department in each community. Try to visualize this using a heatmap or other style of visualization and try to use this to describe the communities in terms of departments.



“-Ao mesclar ambos os gráficos se tornam mais evidente a relação das comunidades e os departamentos em que os indivíduos estão inseridos. Pode-se observar que em geral indivíduos do mesmo departamento também estão na mesma comunidade.”

Comunidade Bluey			
	Dept	QuantidadeDePessoas	Porcentagem
0	2	1	0.421941
1	3	1	0.421941
2	4	3	1.265823
3	6	12	5.063291
4	7	1	0.421941
5	8	18	7.594937
6	9	28	11.814346
7	10	1	0.421941
8	14	2	0.843882
9	15	48	20.253165
10	17	1	0.421941
11	18	1	0.421941
12	21	1	0.421941
13	22	2	0.843882
14	23	19	8.016878
15	24	3	1.265823
16	25	1	0.421941
17	26	3	1.265823
18	27	10	4.219409
19	29	4	1.687764
20	30	3	1.265823
21	31	4	1.687764
22	34	11	4.641350
23	35	12	5.063291
24	36	17	7.172996
25	37	15	6.329114
26	38	10	4.219409
27	40	4	1.687764
28	41	1	0.421941

## Comunidade Yellow

	Dept	QuantidadeDePessoas	Porcentagem
0	0	1	0.675676
1	1	3	2.027027
2	4	1	0.675676
3	6	3	2.027027
4	16	1	0.675676
5	17	1	0.675676
6	19	29	19.594595
7	20	11	7.432432
8	21	53	35.810811
9	22	19	12.837838
10	24	2	1.351351
11	25	3	2.027027
12	28	8	5.405405
13	29	1	0.675676
14	30	1	0.675676
15	32	8	5.405405
16	34	1	0.675676
17	36	1	0.675676
18	38	1	0.675676

### Comunidade Red

	Dept	QuantidadeDePessoas	Porcentagem
0	0	2	1.418440
1	2	8	5.673759
2	3	5	3.546099
3	4	99	70.212766
4	5	17	12.056738
5	6	1	0.709220
6	8	1	0.709220
7	9	1	0.709220
8	14	1	0.709220
9	15	1	0.709220
10	21	1	0.709220
11	25	1	0.709220
12	26	1	0.709220
13	33	1	0.709220
14	36	1	0.709220

### Comunidade Green

	Dept	QuantidadeDePessoas	Porcentagem
0	0	46	36.507937
1	1	16	12.698413
2	4	2	1.587302
3	5	1	0.793651
4	6	4	3.174603
5	7	48	38.095238
6	13	1	0.793651
7	15	1	0.793651
8	21	1	0.793651
9	23	3	2.380952
10	24	1	0.793651
11	26	1	0.793651
12	36	1	0.793651



### Comunidade Pink

	Dept	QuantidadeDePessoas	Porcentagem
0	2	1	0.854701
1	3	6	5.128205
2	4	1	0.854701
3	6	3	2.564103
4	9	1	0.854701
5	10	38	32.478632
6	11	29	24.786325
7	12	3	2.564103
8	13	24	20.512821
9	20	1	0.854701
10	22	4	3.418803
11	26	2	1.709402
12	31	1	0.854701
13	36	1	0.854701
14	38	2	1.709402

### Comunidade Purple

	Dept	QuantidadeDePessoas	Porcentagem
0	4	1	1.041667
1	9	1	1.041667
2	14	88	91.666667
3	15	2	2.083333
4	20	1	1.041667
5	32	1	1.041667
6	35	1	1.041667
7	36	1	1.041667

Comunidade Cyan				
	Dept	QuantidadeDePessoas	Porcentagem	
0	1	43	68.253968	
1	6	5	7.936508	
2	13	1	1.587302	
3	15	2	3.174603	
4	23	5	7.936508	
5	31	3	4.761905	
6	39	3	4.761905	
7	41	1	1.587302	

Find the largest clique size in the graph. How many such largest cliques are there?  
What do you think a clique represents in this context?

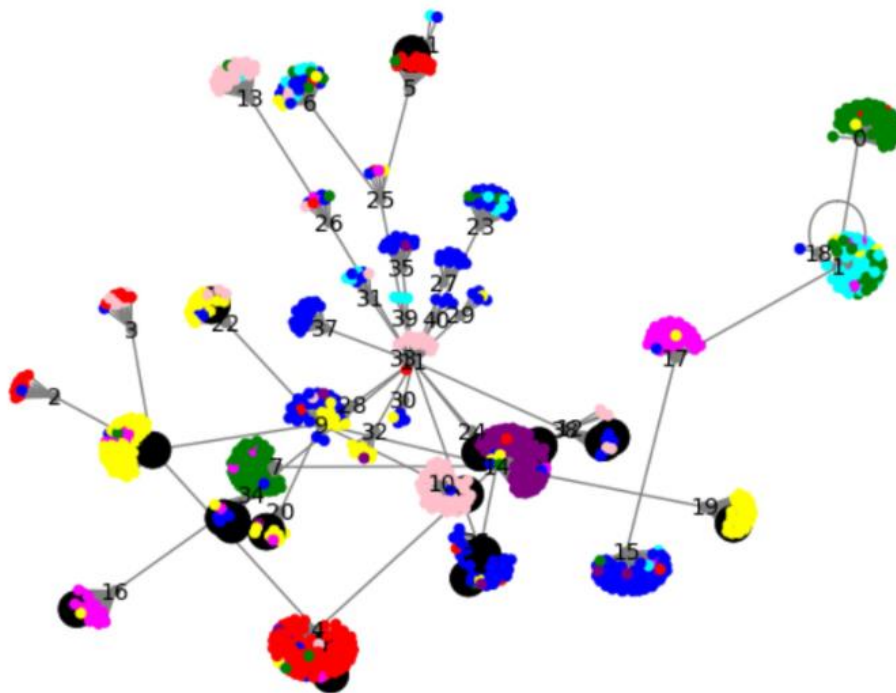
```
graph = nx.from_pandas_edgelist(edgelist,
source = 'from', target = 'to')
cliques = nx.find_cliques(graph)
maximal_cliques = sorted(cliques, key = len)
```

```
#Maior clique
cliques = maximal_cliques[len(maximal_cliques) - 1]
print("O grafo tem " + str(len(maximal_cliques)) + " cliques maximais")
print("O maior clique tem tamanho " + str(len(cliques)))
```

```
O grafo tem 42709 cliques maximais
O maior clique tem tamanho 18
```

“-Nesse contexto, acreditamos que os cliques maximais representam grupo extremamente unidos, ou seja, funcionários que se comunicam constantemente.”

Try to visualize the members of these cliques in the context of the entire graph. What can you conclude?



"- os nodes do maior clique estão representados pelas bolas pretas, elas foram aumentadas para maior identificação."

Use the Leiden community detection algorithm to find a vertex partition with optimal modularity. How many communities does the Leiden algorithm detect?

```
from cdlib import algorithms
graph = nx.from_pandas_edgelist(edgelist,
source = 'from', target = 'to')

# Aplicar o algoritmo Leiden
leiden = algorithms.leiden(graph)

for ind, comm in enumerate(leiden.communities):
    print(str(ind) + ": " + str(len(comm)))
```

```
0: 314
1: 144
2: 134
3: 130
4: 111
5: 97
6: 56
```

Compare the Leiden partition modularity to the Louvain partition modularity.

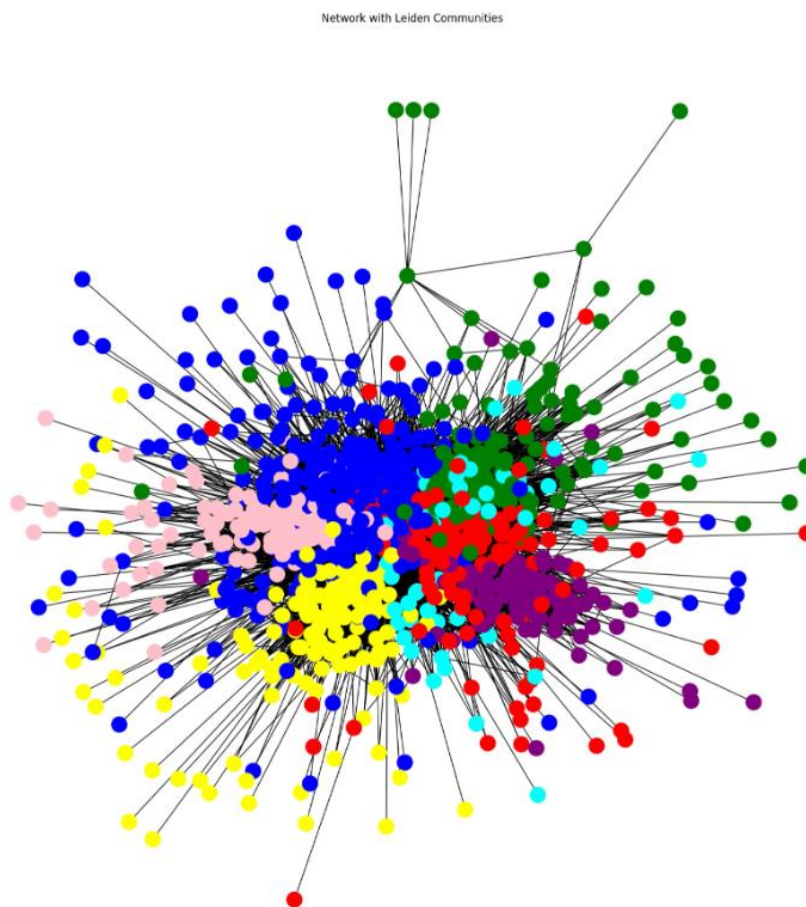
```
louvain_comms.newman_girvan_modularity()
```

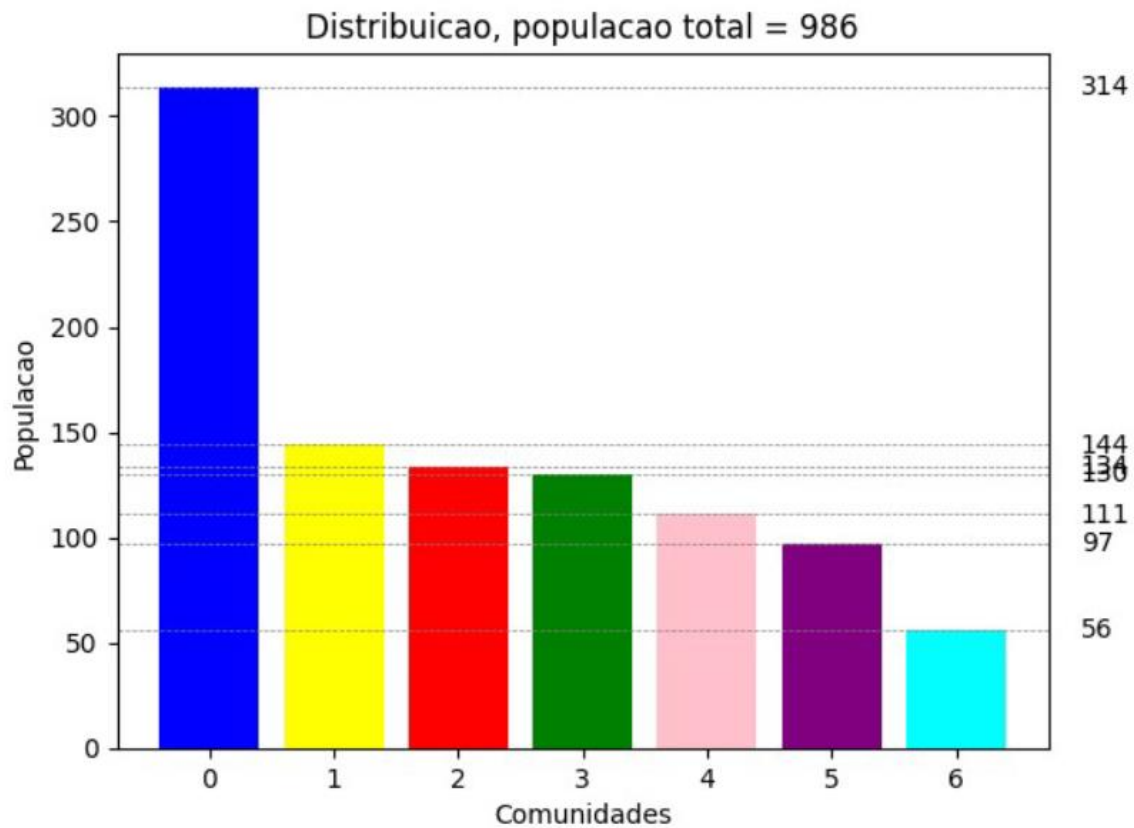
```
FitnessResult(min=None, max=None, score=0.41621670842774317, std=None)
```

```
leiden.newman_girvan_modularity()
```

```
FitnessResult(min=None, max=None, score=0.41737940969873, std=None)
```

Try to use visualization or data exploration methods to determine the main differences between the Leiden and Louvain partitions.





“-Ao usar Louvain Algoritmo o grafo foi dividido em 7 comunidades, já ao usar Leiden o grafo foi dividido em 6 comunidades. Além disso a comunidade azul no algoritmo Leiden cresceu abruptamente.”

“Observação, é provável que o mesmo código ao ser inicializado em diferentes momentos apresente diferentes resultados, em virtude das mudanças de resultados do Louvain Algorithm e Leiden Algorithm.”