

# COMPUTATION REPORT

Kushal Hebbar | C13031425 | khebbar@g.clemson.edu

- ❖ **Dataset used:** “as-22july06” from UFL matrix collection
- ❖ **Tasks:**
  - Plot the degree distribution histogram
  - Plot the cumulative degree distribution function
  - Compute power law parameters  $C$ , and  $\alpha$

## TASK 1: Plot the degree distribution histogram

Made use of R programming language along with igraph and ggplot2 libraries to handle the graph and plot the histogram.

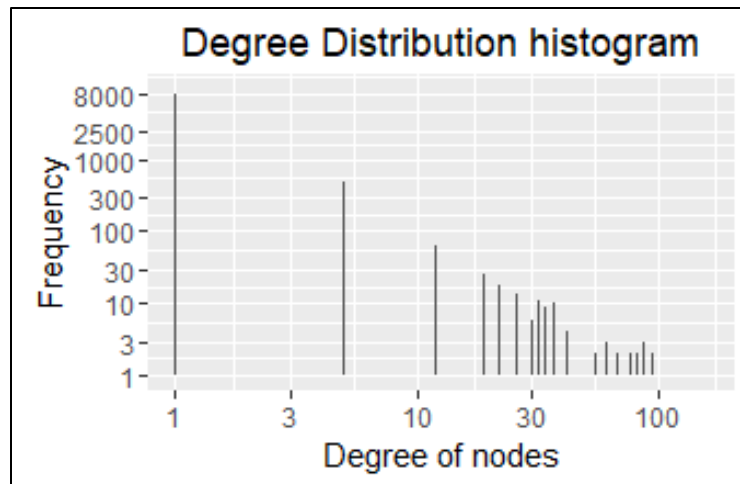
- Calculated the degrees of the graph G (as-22july06 dataset) and stored in degrees variable.
- Converted the structure degree table to a data frame with frequencies of degree occurrences and stored in degree.frequency
- Converted the first column of the data frame to numeric type to avoid conflict while plotting using log-log
- Plotted the histogram using ggplot with x-axis plotting degrees and the y-axis plotting frequencies

Used python and Networkx package to calculate and print out the degrees of the graph.

```
degrees = nx.degree(G)
print(degrees)
```

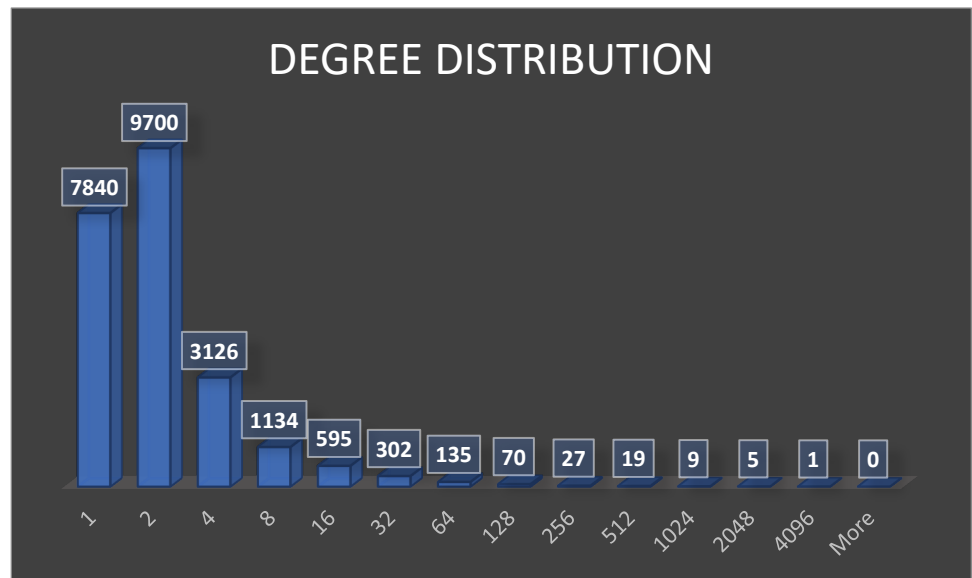
[('12263', 1), ('17174', 2), ('16424', 2), ('3086', 2), ('16932', 2), ('12474', 2), ('8560', 7), ('2288', 6), ('9432', 2), ('18005', 1), ('3774', 2), ('9816', 2), ('1758', 96), ('22412', 1), ('8797', 2), ('19586', 1), ('6174', 2), ('5904', 2), ('1871', 26), ('1195', 2), ('1469', 1), ('2609', 2), ('14727', 1), ('9797', 1), ('8790', 1), ('16339', 1), ('15079', 1), ('556', 2), ('7904', 3), ('9148', 2), ('14171', 3), ('16568', 1), ('12985', 2), ('324', 1), ('9203', 1), ('14814', 1), ('9831', 2), ('18075', 1), ('6241', 2), ('6820', 1), ('7420', 5), ('20503', 1), ('12061', 2), ('8423', 4), ('10843', 2), ('7514', 1), ('13788', 1), ('208', 2), ('18718', 1), ('19109', 1), ('22361', 1), ('2712', 9), ('2394', 3), ('4040', 1), ('22638', 1), ('493', 13), ('10778', 2), ('16783', 1), ('9427', 1), ('1823', 28), ('14304', 1), ('6062', 1), ('5507', 2), ('14679', 2), ('7424', 2), ('14558', 2), ('782', 1), ('6887', 1), ('17255', 2), ('22841', 1), ('18823', 1), ('4778', 1), ('362', 2), ('18710', 1), ('11090', 1), ('13578', 2), ('1503', 3), ('1187', 2), ('22494', 3), ('13520', 1), ('7869', 3), ('1791', 1), ('15266', 2), ('1746', 2), ('9053', 1), ('5148', 1), ('22617', 1), ('17025', 1), ('7170', 2), ('6626', 1), ('19747', 2), ('16114', 2), ('11284', 5), ('20244', 1), ('7756', 2), ('832', 2), ('14795', 2), ('9676', 3), ('21638', 1), ('19796', 1), ('8416', 1), ('5686', 1), ('16899', 1), ('4521', 1), ('419', 2), ('4995', 1), ('13683', 4), ('15927', 1), ('6153', 1), ('5761', 1), ('772', 3), ('1076', 3), ('9774', 1), ('16892', 1), ('2800', 3), ('5296', 5), ('8595', 1), ('20585', 1), ('2436', 11), ('1767', 2), ('18293', 2), ('8096', 1), ('10852', 1), ('12102', 2), ('20536', 1), ('12248', 1), ('7160', 2), ('20844', 2), ('1879', 58), ('8659', 2), ('13188', 1), ('2985', 2), ('22060', 1), ('7095', 2), ('21789', 1), ('18471', 2), ('4993', 1), ('11513', 2), ('21946', 1), ('5411', 2), ('14674', 1), ('21059', 1), ('17113', 2), ('22588', 3), ('2130', 1), ('14520', 1), ('243', 1), ('9305', 1), ('13629', 1), ('22647', 2), ('9426', 2), ('5324', 4), ('6997', 3), ('19645', 2), ('2722', 12), ('14525', 2), ('2957', 7), ('5467', 2), ('1039', 2), ('17328', 2), ('7699', 1), ('16020', 1), ('12048', 2), ('11244', 2), ('20904', 2), ('19012', 1), ('3582', 1), ('21994', 2), ('16351', 1), ('1173', 2), ('9810', 16), ('15299', 2), ('18827', 2), ('5292', 3), ('8168', 3), ('1

The following is the resulting histogram for the degree distribution.



Made use of excel data analysis package to verify the above result. The bins and the frequency formulated in excel are as follows.

Bin	Frequency
1	7840
2	9700
4	3126
8	1134
16	595
32	302
64	135
128	70
256	27
512	19
1024	9
2048	5
4096	1
More	0

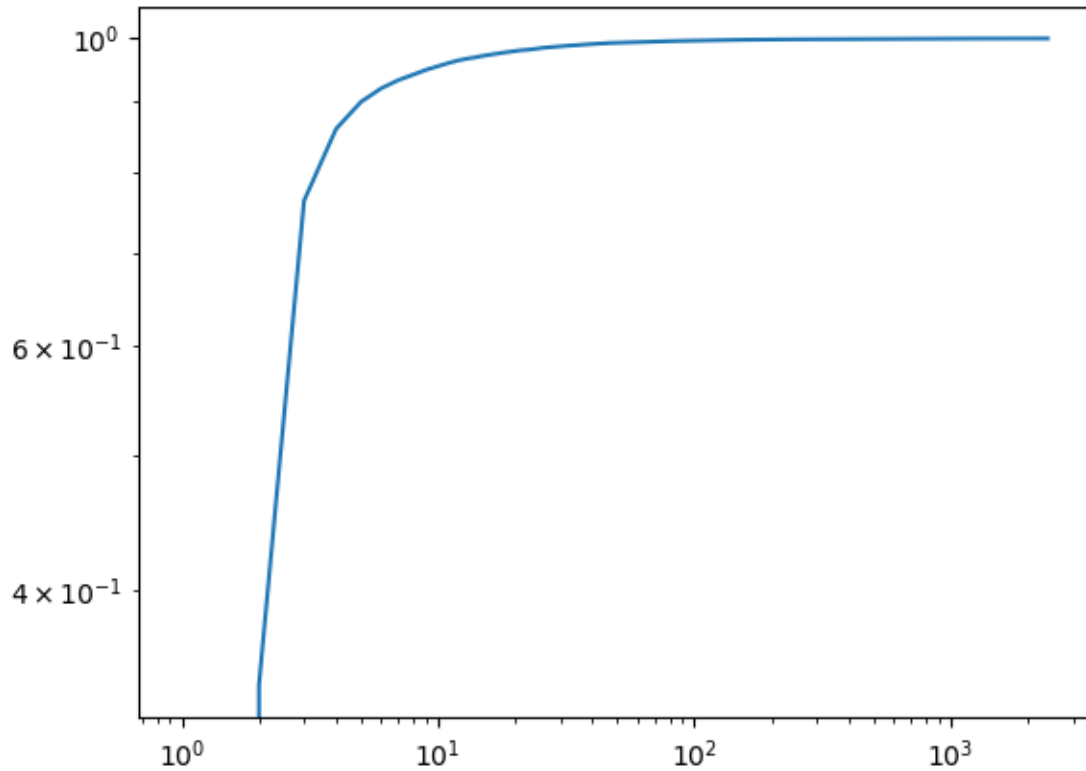


The code is in a file named degree\_histogram.

## TASK 2: Plot the cumulative degree distribution function

- Made use of the powerlaw package in python available at: <https://github.com/jeffalstott/powerlaw>
- Read the network as a weighted edgelist and sorted the degrees of the nodes
- Used function cdf and plot\_cdf that are included in the powerlaw package to plot the cumulative degree distribution

The following is the cumulative degree distribution plot.



The code for the above graph is present as `cumulative_degree_distribution.py`.

### TASK 3: Compute power law parameters $C$ , and $\alpha$

The parameters  $C$  and  $\alpha$  are constants in the power-law distribution and are used to determine the probability of the distribution. I calculated the cumulative degree distribution in excel using a logarithmic binning, calculated the frequency and the cumulative percentage. The last column shown in the figure below depicts the inverse of cumulative percentage.

<i>Bins</i>	<i>Frequency</i>	<i>Cumulative %</i>	
1	7840	34.14%	1
3	11919	86.05%	65.86%
9	2173	95.51%	13.95%
27	705	98.58%	4.49%
81	226	99.56%	1.42%
243	66	99.85%	0.44%
729	26	99.97%	0.15%
2187	7	100.00%	0.03%
6561	1	100.00%	0.00%
More	0	100.00%	

The graph below depicts the plot for the cumulative degree distribution for the above binning and also provides us the equation for  $y$  and R-squared value from which we can conclude that the value of  $C = 1.7111$  and  $\alpha = 1.126$ .

