

# Fairness on the Web: Alternatives to the Power Law

Based on :

Fairness on the Web: Alternatives to the Power Law. Jérôme Kunegis, Julia Preusse, In: Proc. Web Sci. Conf., pp. 175–184, 2012.

[ <https://dl.acm.org/citation.cfm?id=2380741> ]

Fairness



**1% OF THE CATS**

**OWN 99% OF THE  
CATNIP**

# The Pareto Principle

“20% of people own 80% of land.”

(In 1910s Italy)

(Pareto, 1919)



**Jérôme Kunegis**

@kunegis

#WebScience, #DataMining and #MachineLearning postdoc in Koblenz/DE. Network analysis and the math behind it. Koblenz, Germany · <http://uni-koblenz.de/~kunegis/>

Follow

353 TWEETS

130 FOLLOWING

135 FOLLOWERS



**Beyoncé Knowles**

@Beyonce

· <http://www.beyonce.com>

Follow

1 TWEET

14 FOLLOWING

4,520,227 FOLLOWERS

**UNFAIR**



**Steffen Staab**

@ststaab

Follow

398 TWEETS

1,550 FOLLOWERS



**Laszlo Barabasi**

@barabasi

Follow

569 TWEETS

587 FOLLOWERS

The Hidden Pattern Behind Everything We Do



Albert-László Barabási  
Author of 2002

Follow

37 TWEETS

1,247 FOLLOWERS

**8.9%** of bands make up **91.1%** of plays on Last.fm.

**14.6%** of user groups account for **75.4%** of group memberships on Flickr.

**17.4%** of movies receive **82.6%** of ratings on MovieLens.

**17.7%** of profiles receive **82.3%** of ratings on Czech dating site Libimseti.cz.

**19.7%** of all cats receive **80.3%** of all friendships on Catster.com.

**20.3%** of all users receive **79.7%** of “friend” and “foe” links on Slashdot.

**21.3%** of users receive **78.7%** of wall posts on Facebook.

**22.9%** of users make up **77.1%** of all “@” mentions on Twitter.

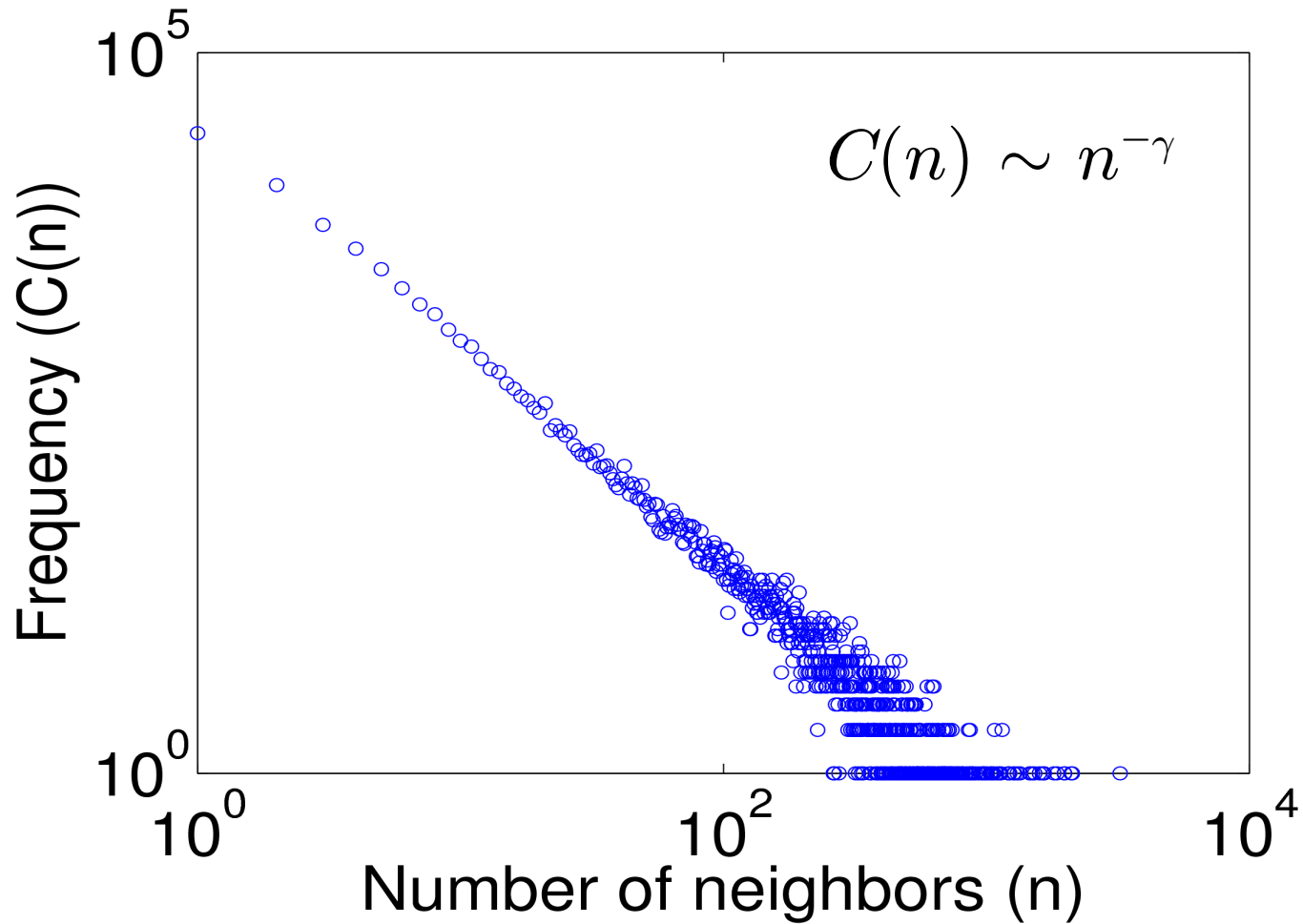
**23.1%** of projects make up **76.9%** of project memberships on Github.

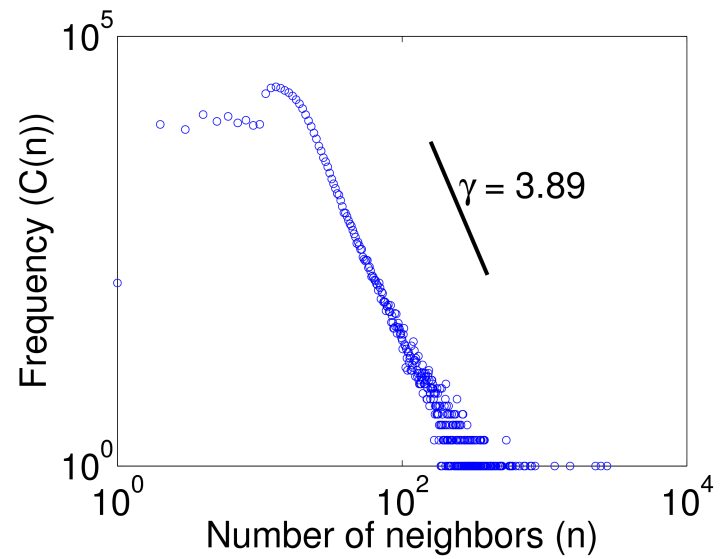
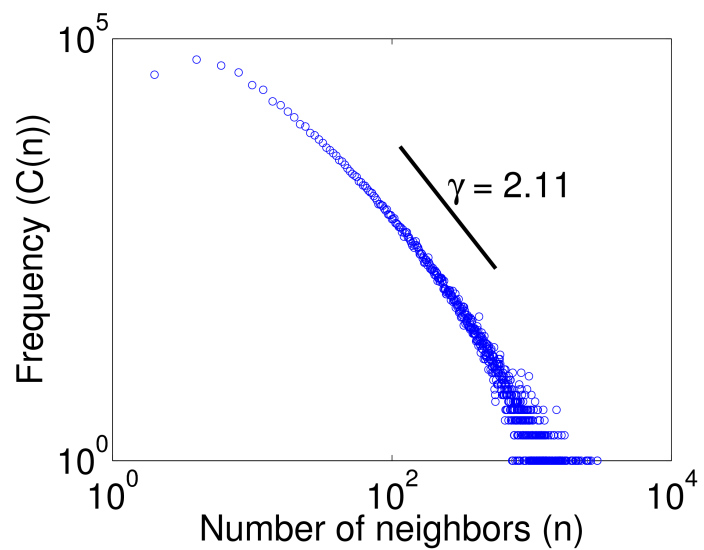
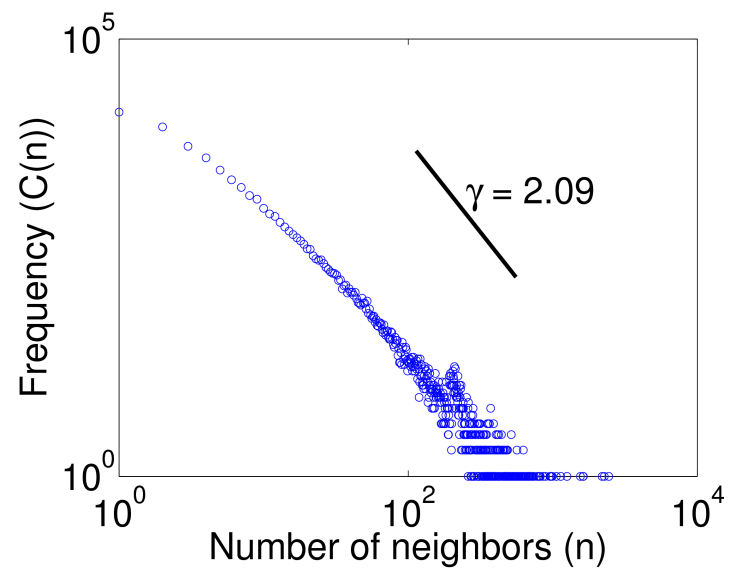
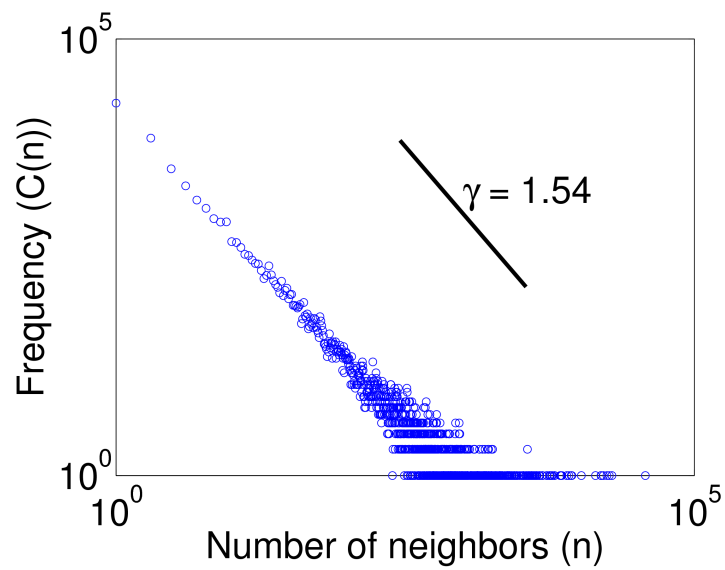
**27.3%** of users receive **72.7%** of replies on Digg.

**27.6%** of all hamsters receive **72.4%** of all friendships on Hamsterster.com.

**35.7%** of all Twitter users receive **64.3%** of all follows.

# Degree distribution

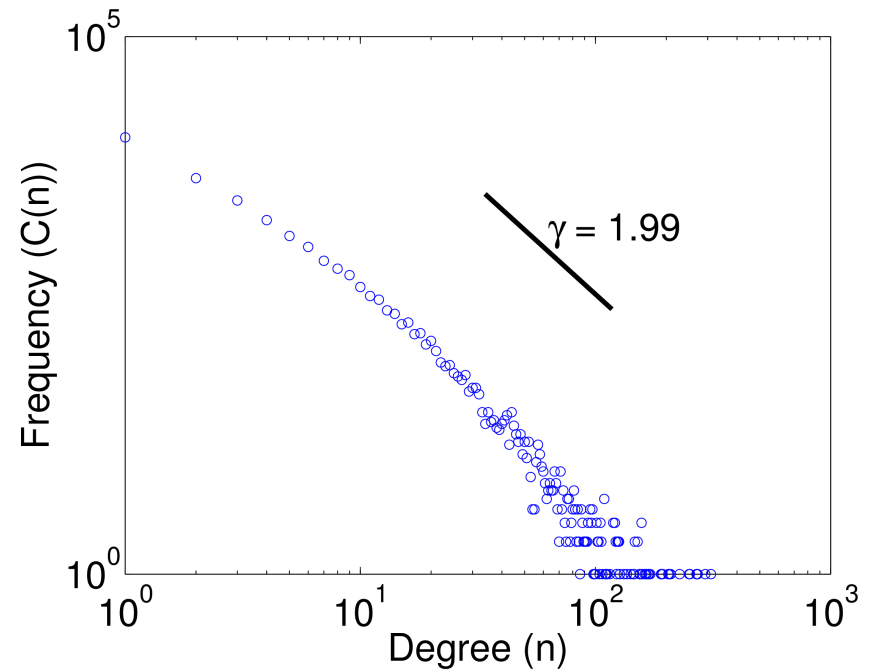
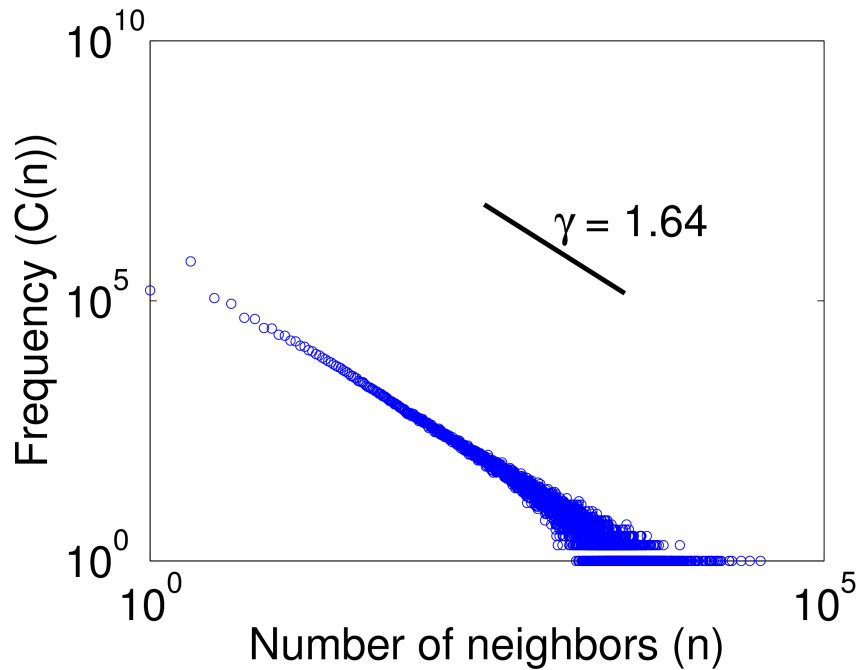




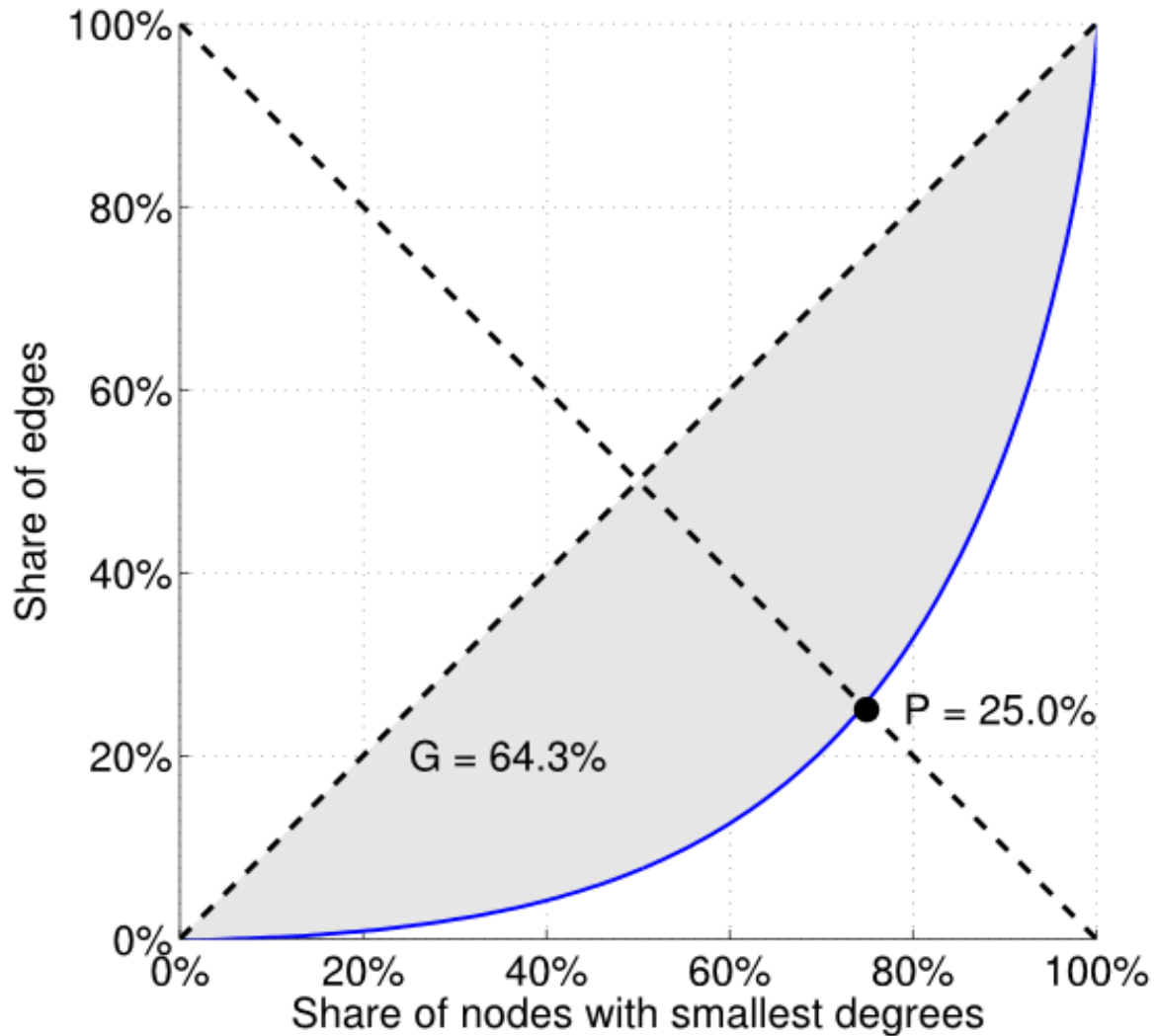


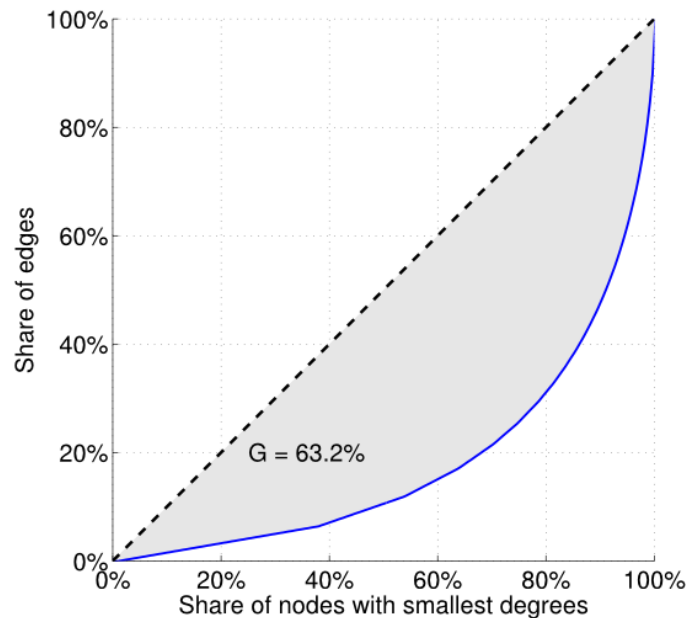
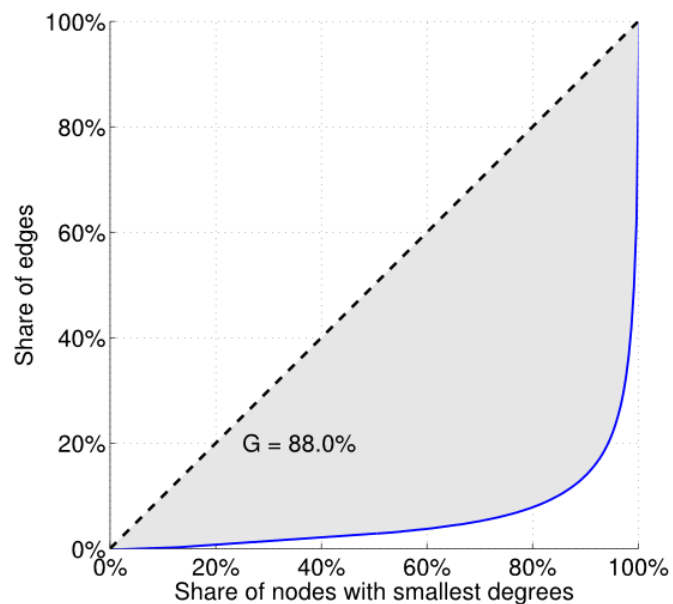
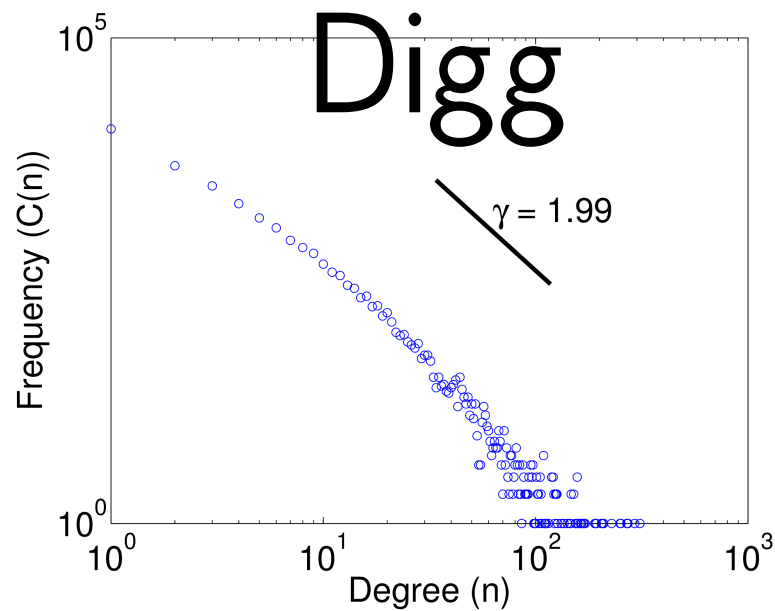
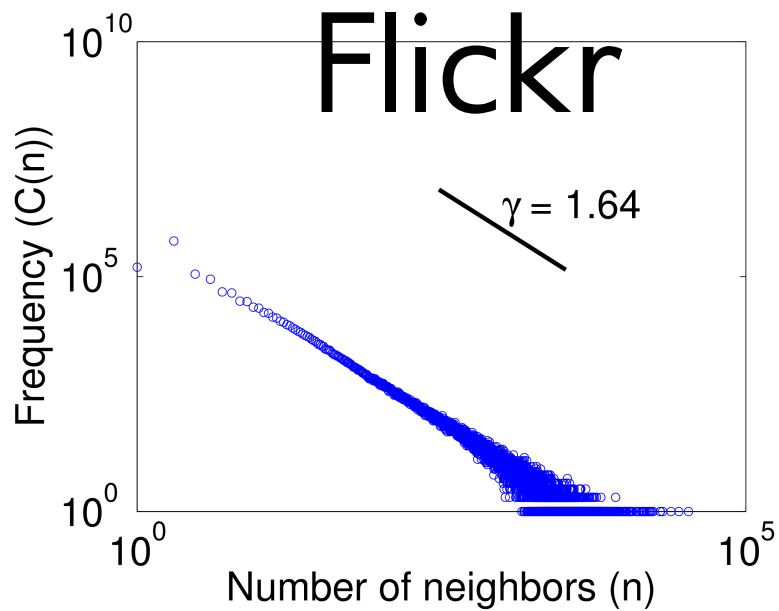
# Which distribution is fairer?

Flickr friendships OR Digg replies



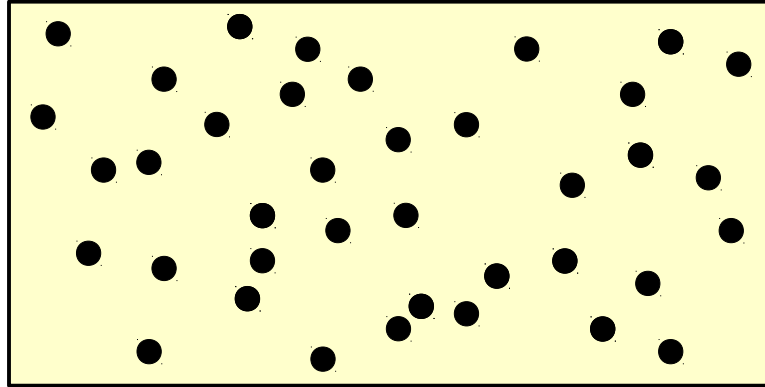
# Lorenz Curve – Gini Coefficient



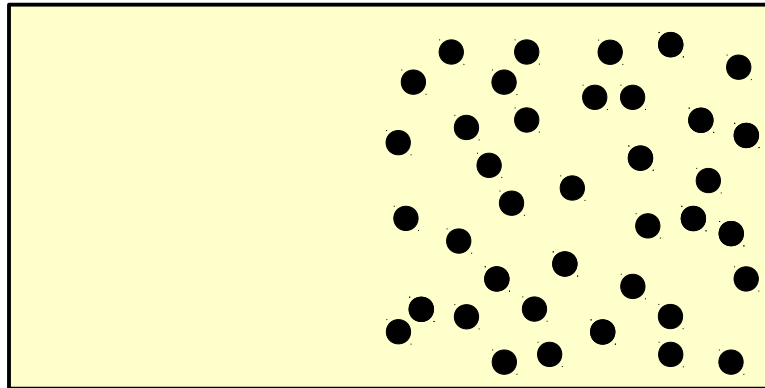




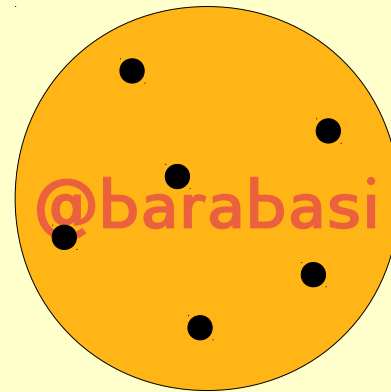
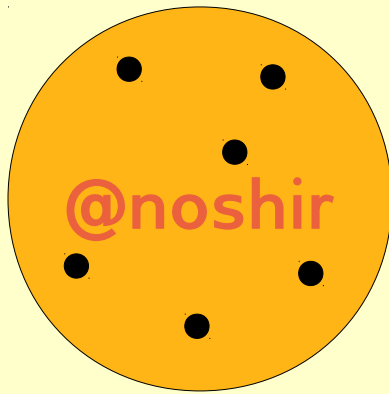
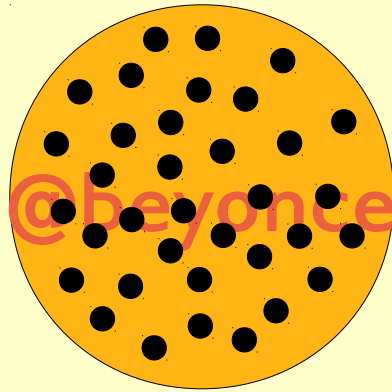
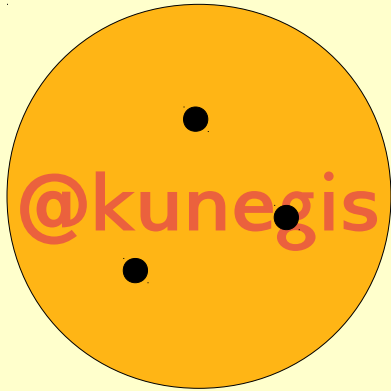
# Entropy



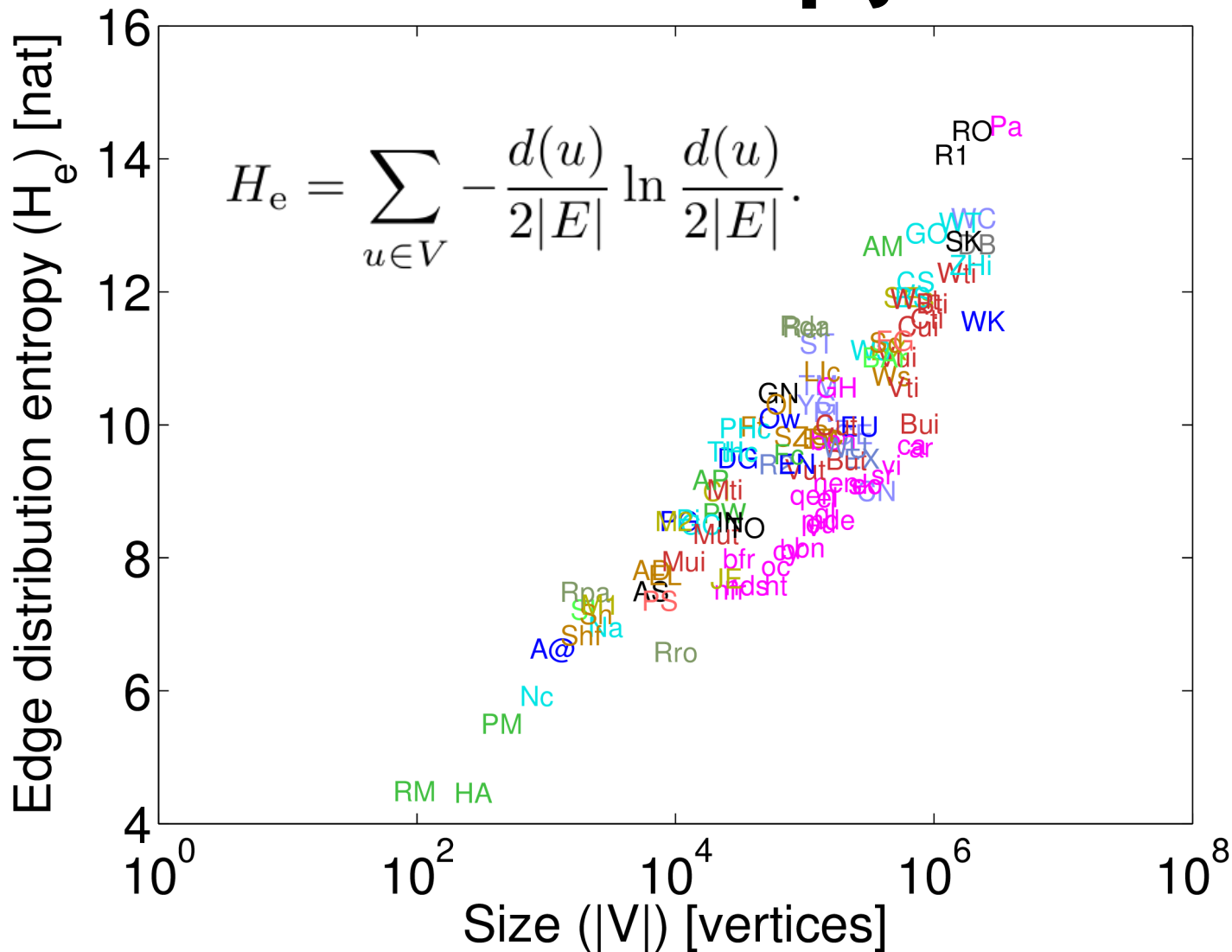
High entropy

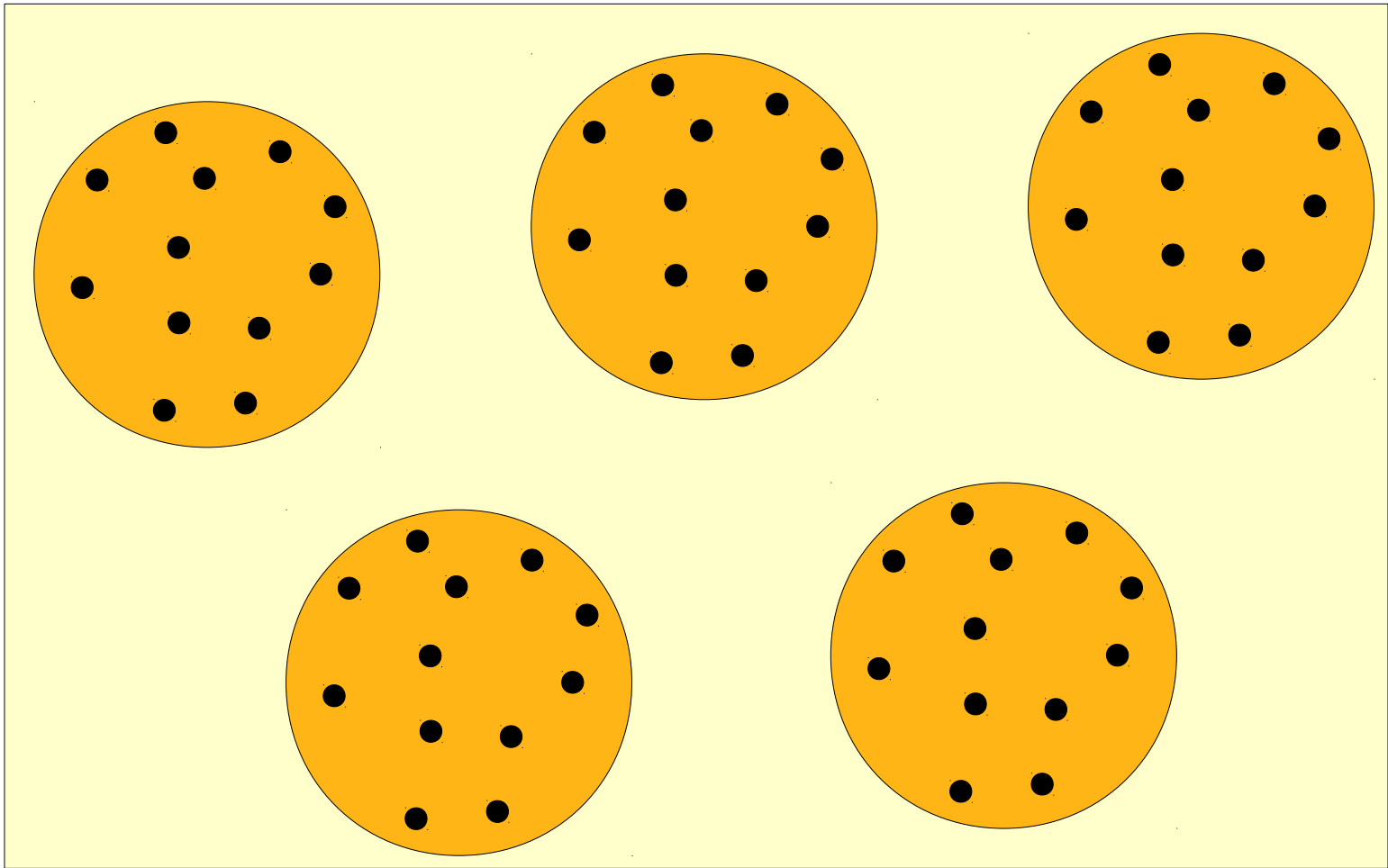


Low entropy



# Entropy

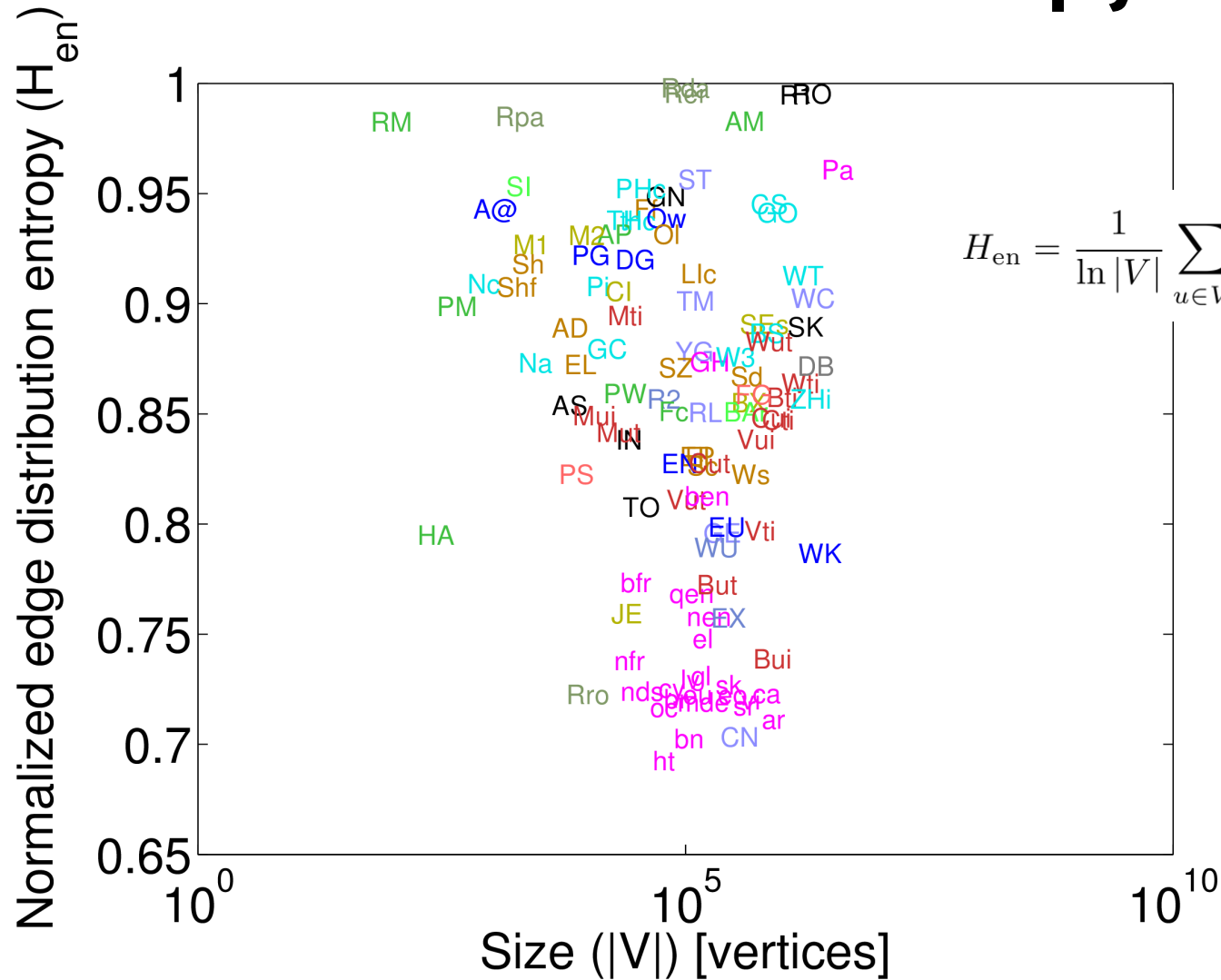




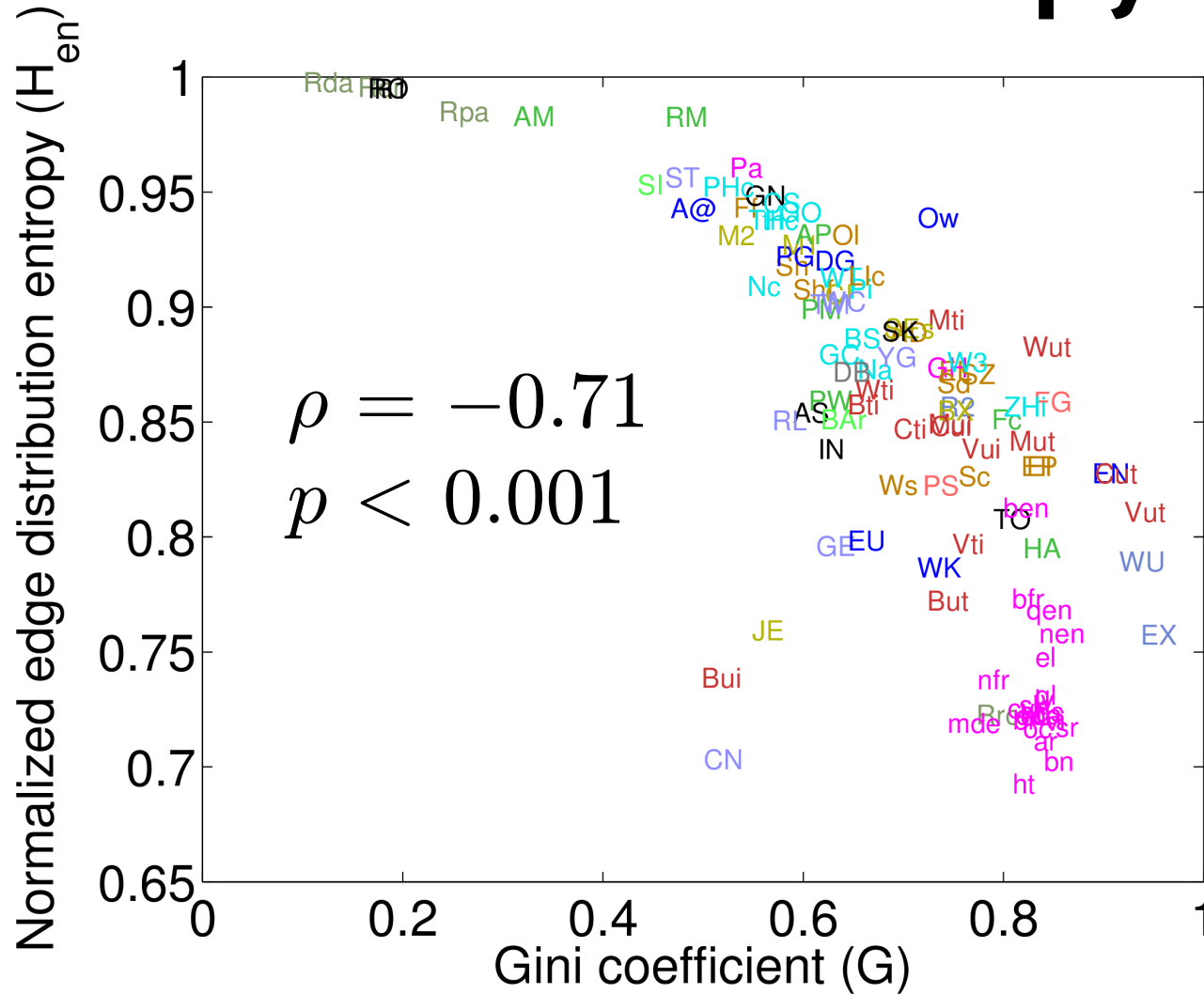
$$\max H_e = \ln |V|$$



# Normalized Entropy



# Gini vs Entropy



<b>Network</b>	<b>Vertices</b>	<b>Edges</b>	<b>Density</b>	$\gamma$	$G$	$P$	$H_e$	$H_{en}$	$H_d$
Advogato	6,535	51,397	15.725	3.510	0.705	0.221	7.809	0.889	3.316
Amazon	402,439	3,387,388	16.794	3.391	0.331	0.386	12.686	0.983	3.516
arXiv astro-ph	18,772	396,160	42.208	2.601	0.610	0.265	9.166	0.931	3.902
arXiv hep-ph	28,093	6,296,894	448.289	1.471	0.611	0.266	9.573	0.935	6.283
arXiv hep-th	22,908	4,889,596	426.890	1.471	0.670	0.242	9.185	0.915	6.102
Berkeley/Stanford	685,230	7,600,595	22.184	2.451	0.659	0.250	11.911	0.886	3.677
Caenorhabditis elegans	382	4,596	9.007	2.211	0.618	0.269	5.497	0.899	2.589
CAIDA	26,475	106,762	8.065	2.121	0.628	0.269	8.535	0.838	1.617
California	1,965,206	5,533,214	5.631	7.021	0.186	0.438	14.420	0.995	1.257
CiteSeer	384,413	1,751,492	9.113	2.751	0.579	0.285	12.155	0.945	2.308
CiteULike	885,046	2,411,819	5.285	2.291	0.707	0.227	11.599	0.847	1.948
DBLP	12,591	49,759	7.904	3.351	0.658	0.235	8.570	0.908	2.631
DBLP	3,578,447	5,344,649	2.987	2.371	0.542	0.300	14.497	0.961	1.737
DBLP	916,319	11,303,522	14.305	2.161	0.680	0.233	13.042	0.950	2.751
DBpedia	2,152,642	7,494,124	6.506	2.151	0.649	0.263	12.709	0.872	2.386
Digg	30,398	87,627	5.685	3.261	0.632	0.255	9.496	0.920	2.383
Enron	87,273	1,148,072	7.377	1.661	0.908	0.101	9.415	0.828	1.722
Epinions	876,252	13,668,320	31.197	1.951	0.769	0.197	11.337	0.829	3.748
EU institution	265,214	420,045	3.168	2.631	0.663	0.240	9.972	0.798	0.896
Facebook New Orleans	60,102	1,545,686	48.507	2.381	0.643	0.250	10.304	0.931	4.580
Filmtipset	75,360	1,266,753	31.975	1.581	0.803	0.171	9.557	0.851	3.480
Flickr	499,610	8,545,307	34.208	1.451	0.849	0.144	11.267	0.859	3.467
Gnutella	62,586	147,892	4.726	4.831	0.563	0.264	10.475	0.948	2.081

...

# Comparison

	Power Law	Gini	Norm. Entropy
Generality	Power-law	All networks	All networks
Interpretation	Many (!)	Economy	Physical
Runtime	Slow	Fast	Fast
Coverage	$d(u) \geq d_{\min}$	All	All