plotly

Product

Company

API Libraries

Learn

Workspace

The Plotly Blog

A "Startlingly Neat & Simple" Rule & Five Graphs About Patterns That Might Surprise You

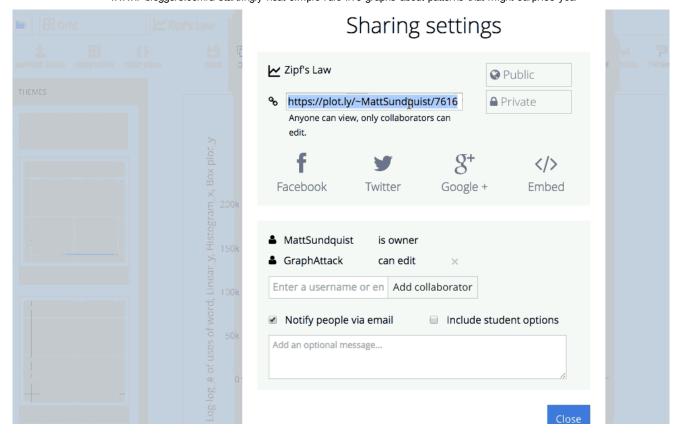
George Zipf popularized an idea–Zipf's Law–that approximates populations of cities, distribution of money in counties, and how frequently words are used. Nobel Prizewinning columnist Paul Krugmans wrote of Zipf's Law that

"the usual complaint about economic theory is that our models are oversimplified — that they offer excessively neat views of complex, messy reality. [In the case of Zipf's law] the reverse is true: we have complex, messy models, yet reality is startlingly neat and simple."

Read on to learn more. Let us know if you want to run Plotly Enterpriseon-premise.

A Zipfian Distribution: How Often Words Appear

A Zipfian distribution is a type of power law. A power law occurs when one event varies as a power of another. One application of Zipf's law states that in texts of natural language (e.g., books), each word is used twice as often as the next most commonly occuring word. The graph below applies the rule to word usage in 29 UK books below. "The" occurred 225,300 uses, and was the most commonly used word. Note that the graph is interactive; you can press the "play with this data" link to edit, embed, and share your own version.



Evaluating Power Laws

We can test for a power law by plotting frequency (y-axis) against rank (x-axis) on a double log axis. Then check for a straight line. The graph below shows three attempts to fit a power law function to datasets. The plot on the left is a good fit. The plot in the middle is a decent fit. The plot on the right is not a good fit.

Evaluating Zipfian Distributions For City Populations

population of cities (y-axis) and the rank of each city. In this dataset, New York has the highest population and is ranked first.

GDP Of Nations

We are approaching a Zipfians distribution for country GDP vs rank.

Evaluating Power Laws For Many Datasets

Researchers use power laws to determine how much inftrasture a city needs, examine the number of gas stations required in a city, and much more.

If you liked what you read, please consider sharing. Find us at feedback@plot.ly and @plotlygraphs.

R

