

Unless otherwise noted, the content of this course material is licensed under a Creative Commons Attribution – Non-commercial – Share Alike 3.0 License.

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

Copyright © 2009, Mark McCabe.

You assume all responsibility for use and potential liability associated with any use of the material. Material contains copyrighted content, used in accordance with U.S. law. Copyright holders of content included in this material should contact open.michigan@umich.edu with any questions, corrections, or clarifications regarding the use of content. The Regents of the University of Michigan do not license the use of third party content posted to this site unless such a license is specifically granted in connection with particular content. Users of content are responsible for their compliance with applicable law. Mention of specific products in this material solely represents the opinion of the speaker and does not represent an endorsement by the University of Michigan. For more information about how to cite these materials visit <http://michigan.educommons.net/about/terms-of-use>.

Any medical information in this material is intended to inform and educate and is not a tool for self-diagnosis or a replacement for medical evaluation, advice, diagnosis or treatment by a healthcare professional. You should speak to your physician or make an appointment to be seen if you have questions or concerns about this information or your medical condition. Viewer discretion is advised: Material may contain medical images that may be disturbing to some viewers.

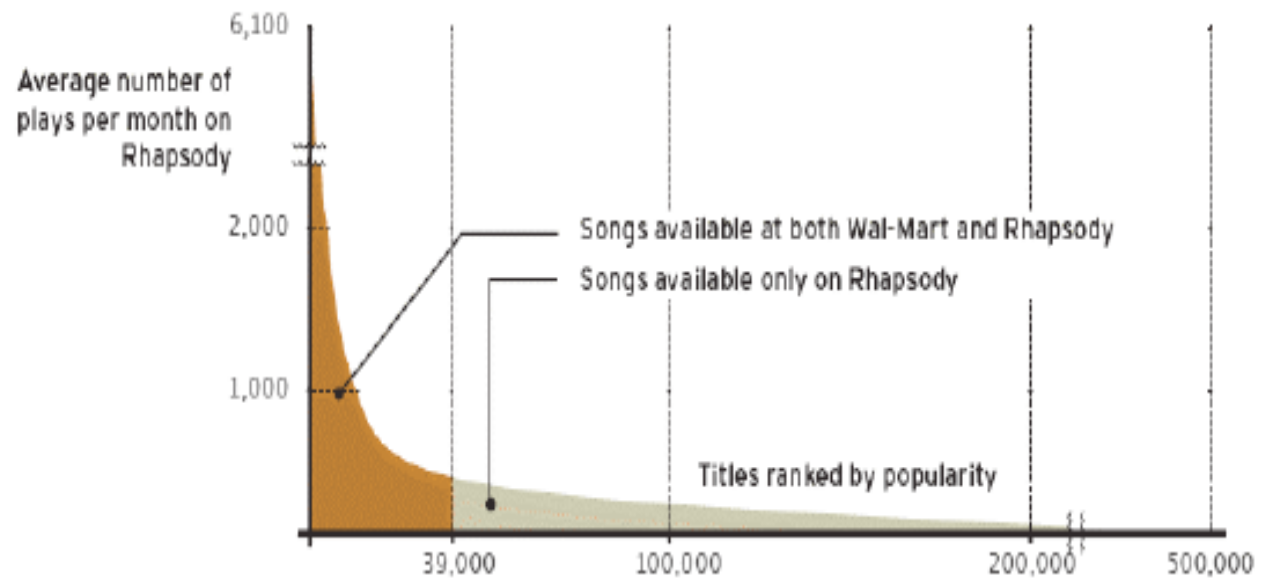
SI 646: Week 6

The Long Tail...

Today's Agenda:

1. Long Tail Stuff
2. Case: Long Wait for Long Tail?
3. Review Assignment 1, etc.
4. Next Time: Final Exam

Note: these slides are based partly on Jeff Mackie-Mason's 646 slides from Winter term 2008.

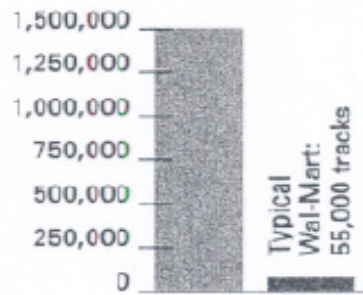


Rhapsody: More than 2 million tracks
40% of revenue from tracks not at Wal-Mart

The new growth market: Products you can't find anywhere but online

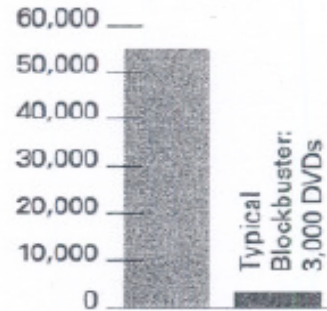
Rhapsody

Total inventory:
1.5 million tracks



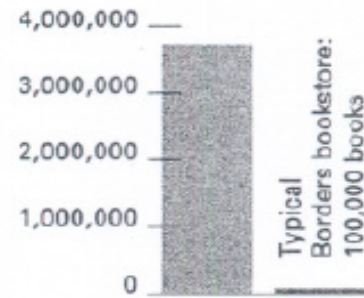
Netflix

Total inventory:
55,000 DVD titles

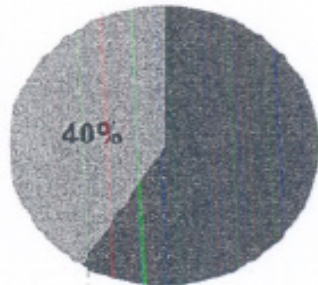


Amazon

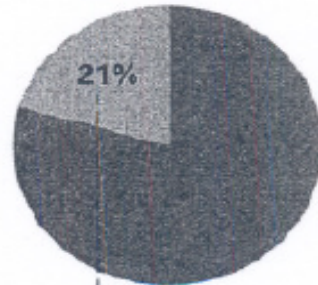
Total inventory:
3.7 million book titles



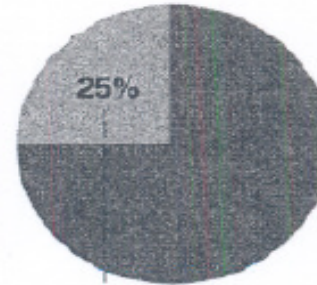
Total sales



Total sales



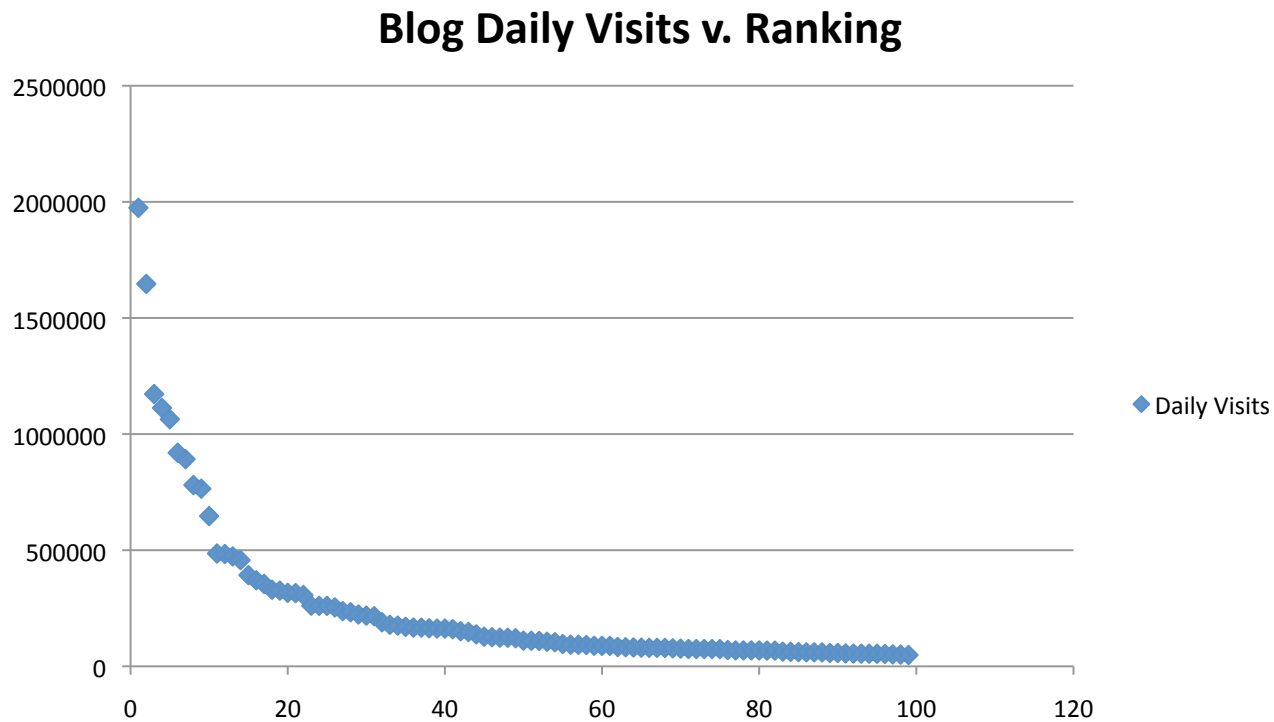
Total sales



Products not available in offline retail stores

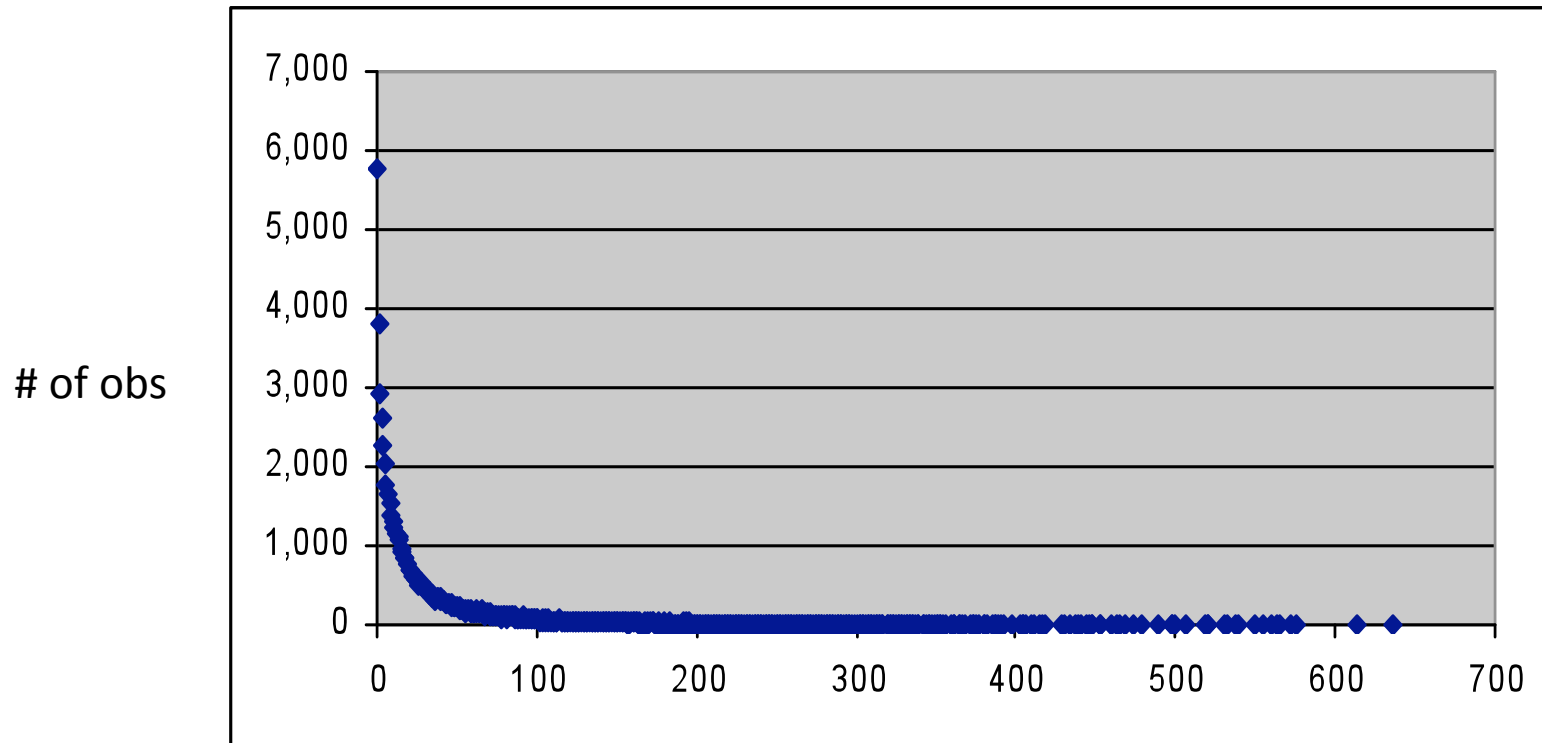
Visits to Top 100 Blogs

(based on data obtained from ["the truth laid bear"](#) , 11 Feb 2009)



Note: top 20 blogs responsible for 62% of traffic...

Citations to journal-year observations in **economics**



Citations in Year Y to journal-publication-year X

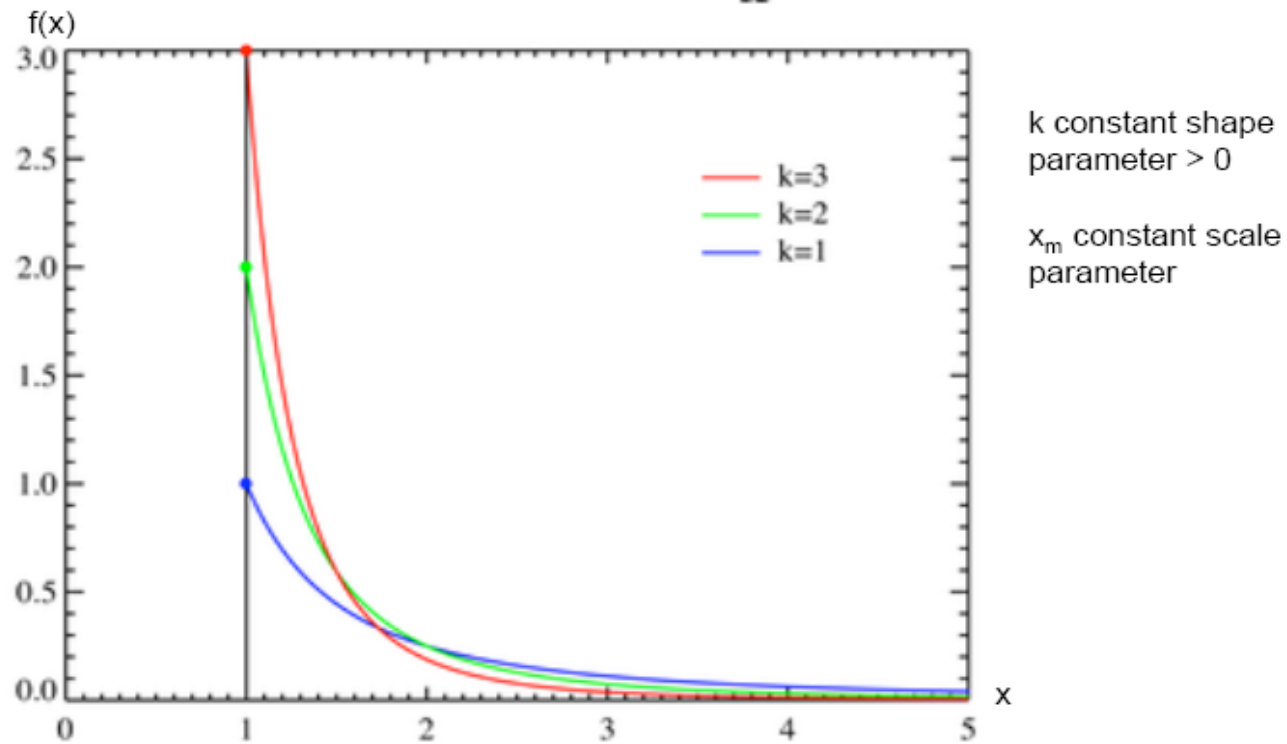
Source: McCabe and Snyder, 2008

Pareto distribution:

What is the probability that a person has wealth x ?

- Vilfredo Pareto, Italian economist and social theorist, 1848-1923

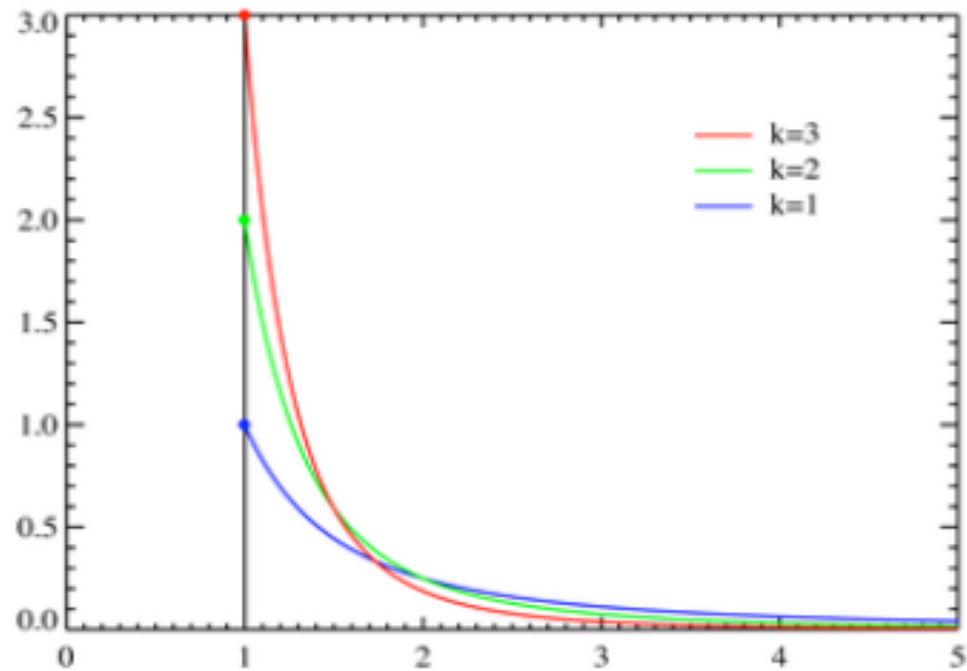
$$f(x; k, x_m) = k \frac{x_m^k}{x^{k+1}} \text{ for } x \geq x_m.$$



Source: [wikipedia](https://en.wikipedia.org/wiki/Power-law_distribution)

Pareto distribution:

Most people have low wealth, few have very high



Other Examples:

- Human settlement sizes
- File sizes transferred over Internet
- Sizes of oil fields
- Rates of return on corporate equities (stocks)
- Areas burnt in forest fires

Zipf's law (1949):

“In a corpus of natural language utterances, the frequency of any word is roughly inversely proportional to its rank in the frequency table”

More generally, the size of the r'th largest occurrence of the event is inversely proportional to it's rank:

$$y = a r^{-b}$$

with $b > 1$, but close to unity

In the “Brown” corpus,

“the” accounts for 7% = .07/1

“of” for 3.5% = .07/2

“and” for 2.8% = .07/2.5

The first 135 words account for 50%

Based on a 1967 analysis of the

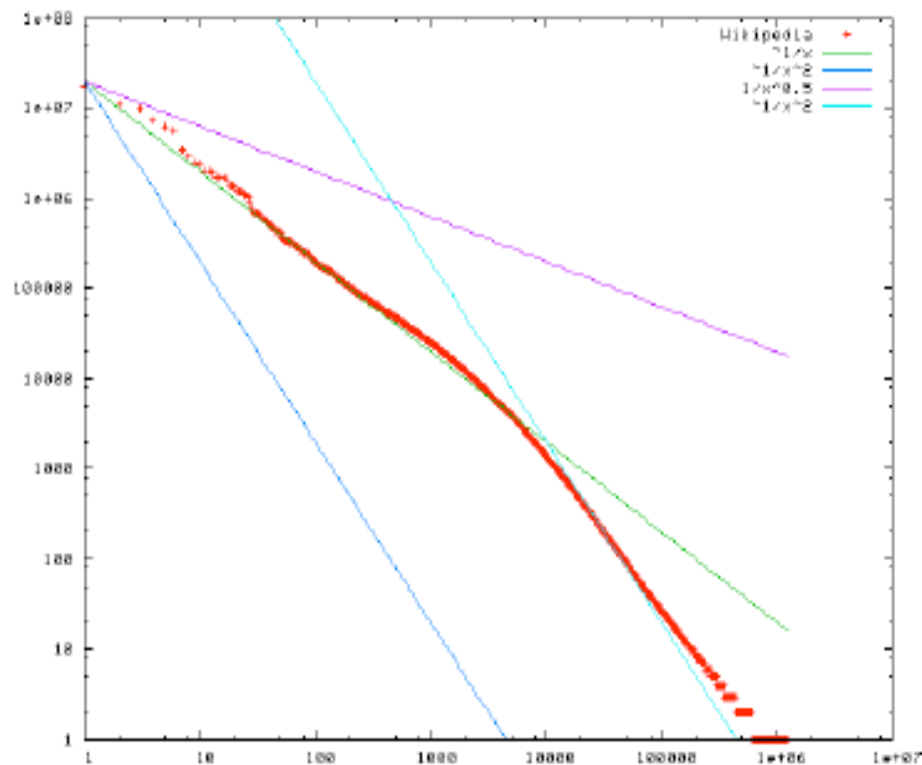
Brown University Standard Corpus of Present-Day American English

For more info go to the following Wikipedia [link](#)

Pareto & Zipf are both examples of a
power law:

$$y = a x^k$$

Take logs of both sides: $\log y = \log a + k \log x$
which is log-log linear: $z = a + b v$



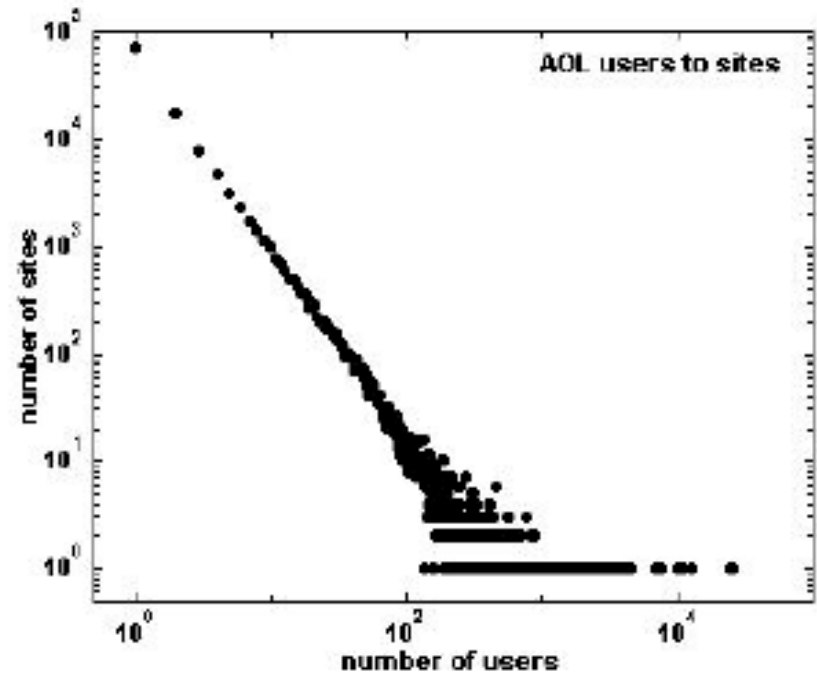
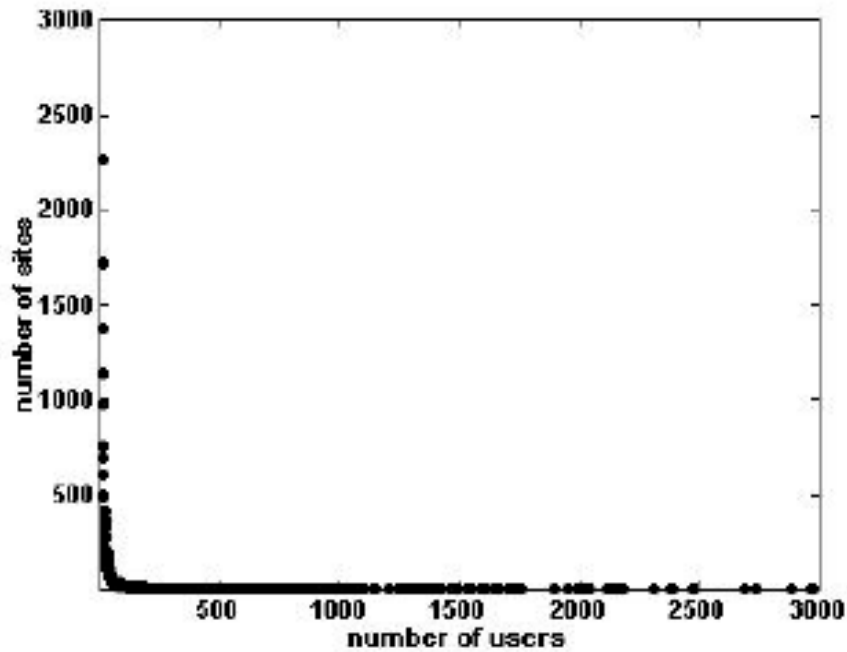
A plot of word frequency in Wikipedia (November 27, 2006).

The plot is in log-log coordinates.

x is rank of a word in the frequency table; y is the total number of the word's occurrences. Most popular words are "the", "of" and "and", as expected.

Zipf's law corresponds to the upper linear portion of the curve, roughly following the green $\log y = 1.3 \times 10^7 - \log x$ line.

http://en.wikipedia.org/wiki/Zipf's_law



Sites visited by AOL users, December day 1997.

L. Adamic, "Zipf, Power-laws, and Pareto - a ranking tutorial",

<http://www.hpl.hp.com/research/idl/papers/ranking/ranking.html>

Note: this resembles the journal citation plot – not a standard long tail plot

80 - 20 Rule?

For many phenomena, 80% of the consequences stem from 20% of the causes (Pareto principle)

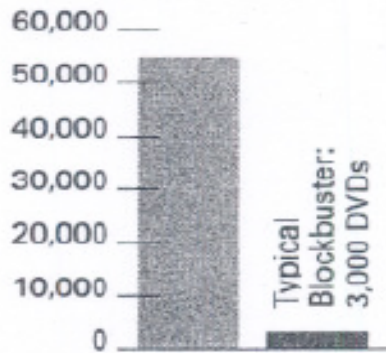
Doesn't need to add to 100! These are percents of different things.

EXAMPLE: Inventory typically 10% of products account for 80% of inventory
(or 20% of products account for 95% of inventory)

**“Small fraction of X responsible
for large fraction of Y”**

Netflix

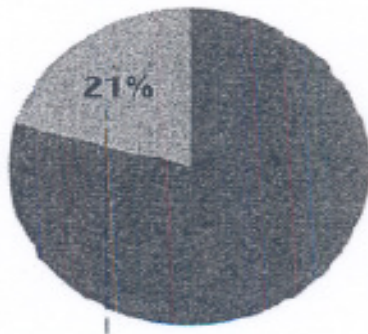
Total inventory:
55,000 DVD titles



3000 out of 55,000 titles: 5.5%

79% of sales

Total sales



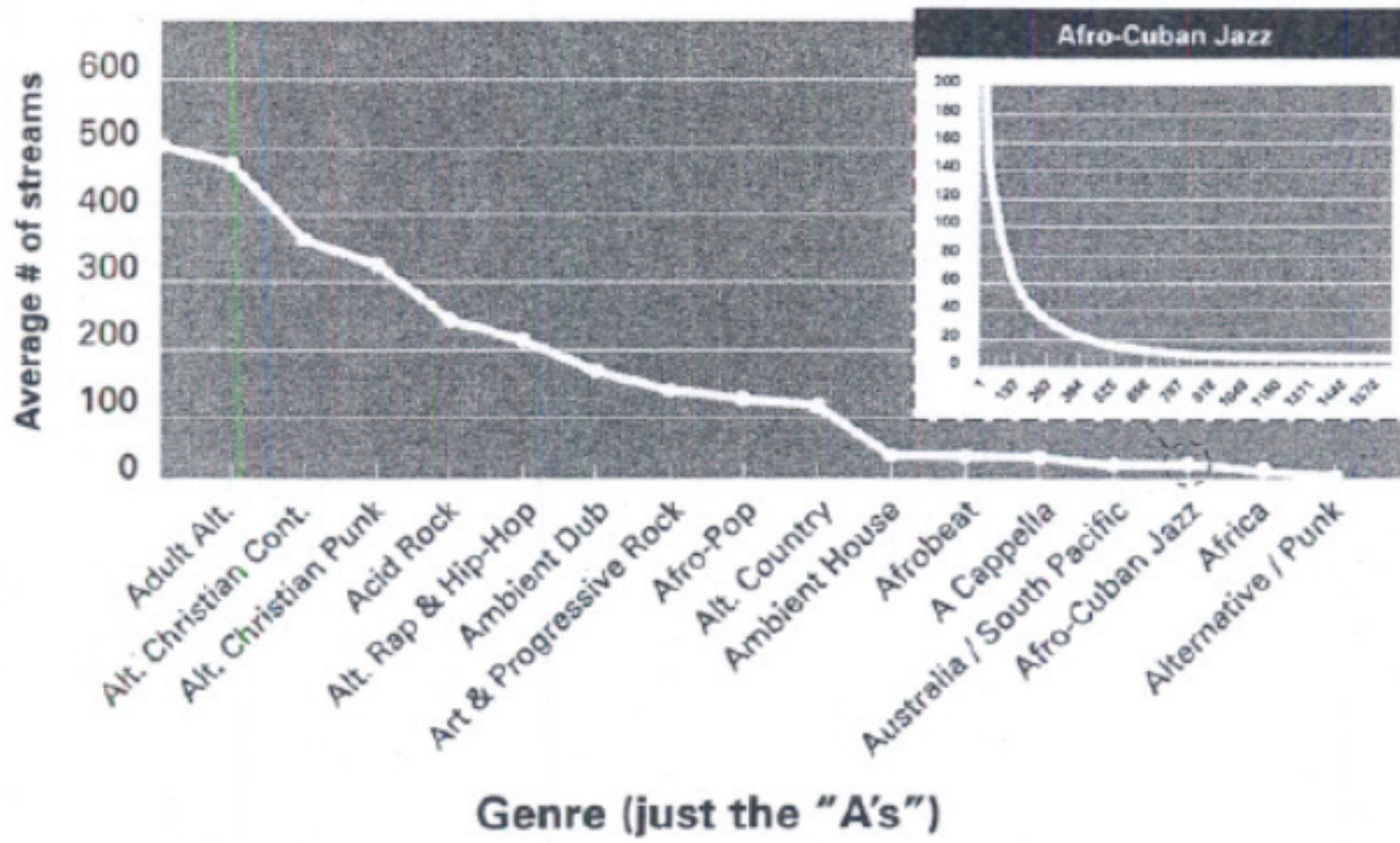
For Netflix, “80 – 6 rule”

What causes power laws?

- Variety
- Quality variation (drives differences in popularity)
- Network effects (arising from reputation effects)
- Scarcity (without scarce time/attention/resources would consume everything!)

Generally, goods power laws apply to subcategories (e.g., genres) as well...

Tails within Tails



- **Network effects and recommender services are most effective at genre or subcategory level**

Let's do the basic economics of variety:
demand for, supply of

Why the long tail now?

**Didn't consumers demand variety
before?**

Need to know about and find variety.

YAHOO! SHOPPING

The screenshot shows the Yahoo! Shopping homepage. At the top, there are navigation tabs for HOME, CLOTHING, ELECTRONICS, COMPUTERS, HOME & GARDEN, and MORE. Below these is a search bar with the text "SHOP FOR:" followed by an empty input field, the word "in", a dropdown menu set to "all departments", and a yellow "Search" button. The main content area is divided into several sections. On the left, there are two columns of category links, each with a sub-link: Clothing & Accessories (Women's, Men's, Shoes, more...), Computers (Laptops, Desktops, Software, more...), Electronics (iPods, Cameras, TVs, Phones, more...), DVDs, Music & Books (Top Movies, Billboard 100, New Books), Flowers & Gifts (Gifts, Roses, Gourmet Foods, more...), and Fragrances & Beauty (Perfumes, Makeup, Skin Care, more...). The right column includes Home & Garden (Bed & Bath, Kitchen, Furniture, more...), Jewelry & Watches (Diamonds, Gold, Watches, more...), Sales & Bargains Center (Free Shipping, Coupons, Sales, more...), Sports & Outdoors (Athletic Shoes, Fitness, Sports, more...), Toys & Baby (Dolls, Car Seats, Board Games, more...), and More Categories (Autos, Health, Tools, Video Games). At the bottom of this section are three buttons: "Shop by Brand", "Shop by Store", and "Shopsphere". On the right side, there are two promotional banners. The top one features the Target logo and text: "Let us do the legwork. Free shipping on select furniture." with a "Shop Target.com" button and an image of a lamp. The bottom one is titled "YAHOO! SHOPPING BUYING GUIDES" and includes the text: "Need help Shopping? Visit our Buying Guides for expert advice and help choosing the product that is right for you." with a list of categories: Digital Cameras, MP3 Players, Laptop Computers, Television & Video, Cell Phones, and Desktop Computers. A small image of a woman's face is visible in the bottom right corner of the buying guides banner.

Internet can help with this in big way, increasing demand for variety.

Other big changes are in cost of providing variety...

e.g. costs of distrib, mfg, inventory, mktg, etc

Suppose there are distribution fixed costs

- $\pi = (p - mc)Q - F$
- Produce if expect $\pi > 0 \rightarrow Q > F / (p - mc)$
- If F gets smaller, more goods offered
- Retail display, warehousing, some marketing are mostly fixed.
- e.g., need to rent 1 inch of CD storage per title regardless of how many are sold.

So, fixed distribution costs favor mass market hits. Lower fixed costs favor niche goods.

See MacKie-Mason, Shenker and Varian, "Service Architecture and Content Provision: The Network Provider as Editor," in *Telecommunications Policy*, vol. 20, no. 3, April 1996: 203-17.

Does marketing to the Long Tail imply a shorter “head”?

- Availability of variety: people shift attention and consumption away from head
- Lower search costs of finding what you want
- Sampling to overcome risk of purchasing less familiar experience goods
- **EVIDENCE: Brynjolfsson et al (2005), women’s clothing company, lower search**

costs for fixed inventory:

- Bottom 80% of products yielded 15.7% of catalog sales
- but 28.8% of online sales

Does marketing to the Long tail increase demand or just shift it?

- Time and attention are limited (did Amazon increase avg. hours of reading?)
- But, by providing more desirable products, we may shift time and attention from other activities.
- Also, better pre-selection may help us consume more efficiently: “I may not read any more words than I once did, but they’re more likely to be meaningful to me.

Should prices be higher or lower for products down the tail?

- Depends on demand elasticity, which can go either way.
- For example, demand for niche books is often less elastic (supports higher prices) because of specific preferences. Discounting higher on best sellers (why?)
- Music demand, on the other hand, might be more elastic further down: willing to experiment, but only if price lower.

Anderson: Long Tail spawns two imperatives.

- 1. Make everything available
 - Driven by reduced distribution costs
- 2. Help me find it
 - Driven by reduced costs of finding and recommending

Anderson's nine rules

- A good synthesis review of SI 646!
- But Didn't need to discover Long Tail to learn them.
No new ideas....

1. Move inventory way in...or way out
2. Let customers do the work
3. One distribution method doesn't fit all
4. One product doesn't fit all (unbundle/rebundle)
5. One price doesn't fit all (price discrimination!)
6. Share information (lose control)
7. Think "and", not "or" (variety!)
8. Trust the market to do your job (related to 2)
9. Understand the power of free (experience goods)