

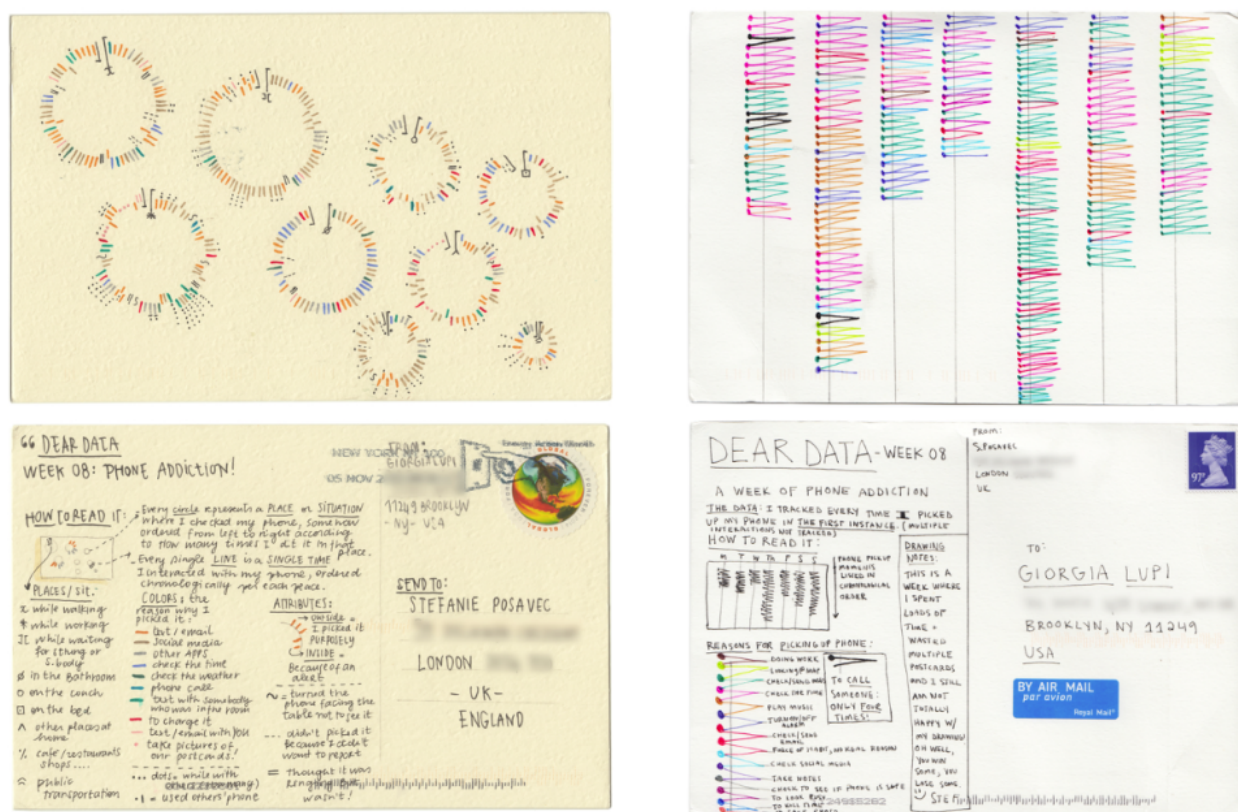
## 2015 年数据可视化十佳项目

2015 年，优秀的可视化作品生机勃勃，我可以确定，明年也会有很多好作品。横跨不同主题和应用形式的项目大量涌现，但如果让我选一个年度主题的话，那一定是“教学”，不管是通过解释说明，模拟说明还是深刻分析的方式。有时候会感到可视化创作者很大胆，试着让读者们不再用惯有的思维方式来理解数据和统计学。我很喜欢这一点。

以下是我选出的 2015 最佳项目。按照惯例，排名不分先后。同时，也有很多不在这个名单上的作品，它们同样很优秀。  
我们一起来看看它们。

## 亲爱的数据

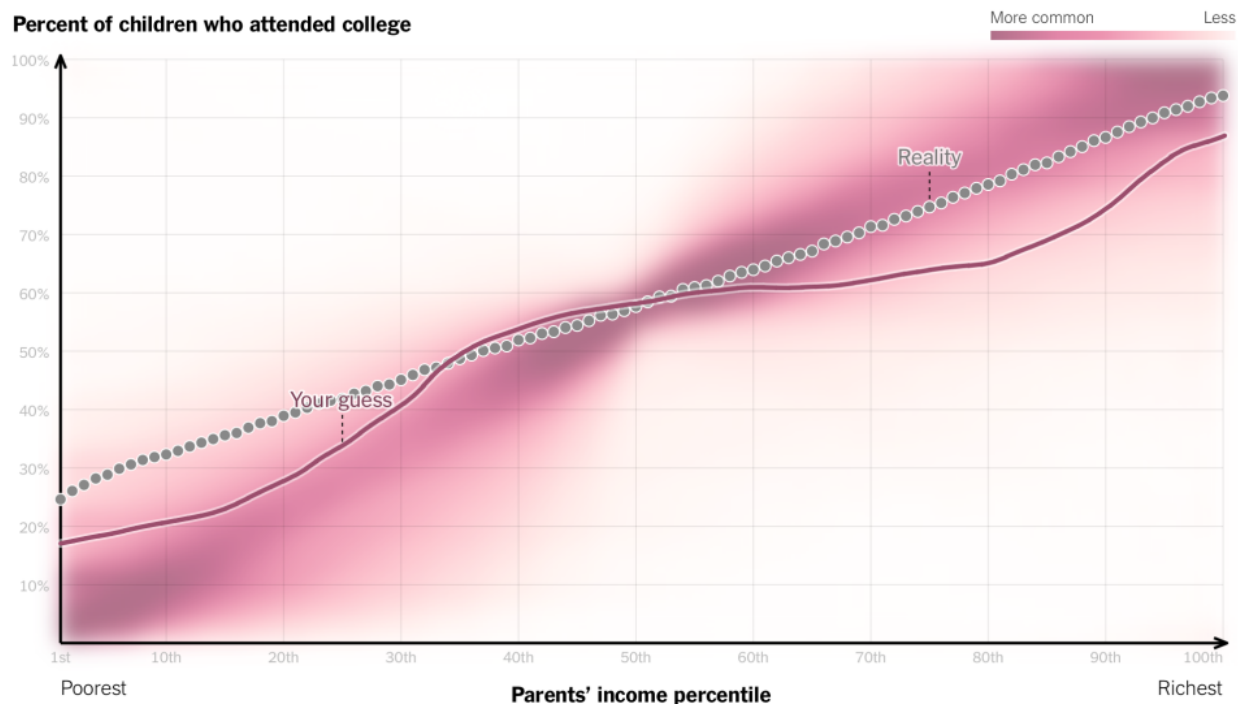
这是一个值得跟踪的有趣项目，它涉及的两个话题——可视化和自我监测——引起了我的注意。



亲爱的数据是一个 Stefanie Posavec 和 Giorgia Lupi 做的一年项目。每个人会追踪一周中每天发生的事情，比如每个人接多少次电话，然后把这些数据可视化在一张明信片上。然后他们把这些明信片寄给对方——Lupi 现居纽约而 Posavec 住在伦敦

### 你来画：家庭收入如何预测孩子的大学入学几率

感觉今年似乎是挑战读者如何在一个更高的统计学视角上理解数据的一年。



纽约时报 Upshot 的 Gregor Aisch, Amanda Cox 和 Kevin Quealy 请读者画一条反映家庭收入和上大学孩子数百分比的线。所以你能看到你自己的线，真实数据画出的线和其他人是怎么看待这个关系的。

关于 Upshot/New York Times 和数据假设，参见 [this quick puzzle to test your pattern-finding ability](#)（快速测试你的模式识别能力）和 [3-D chart of the economy's future](#)（经济未来发展 3D 图）。

## “黑”出你的科学荣耀

纽约时报的 538 也通过他们的可视化互动进行了一些数据科学教学活动。其中最好的一节课是教人们如何通过“窃取 P 值”（p-hacking）来从同一数据集中得到你想要的结果。

# Hack Your Way To Scientific Glory



You're a social scientist with a hunch: **The U.S. economy is affected by whether Republicans or Democrats are in office.** Try to show that a connection exists, using real data going back to 1948. For your results to be publishable in an academic journal, you'll need to prove that they are "statistically significant" by achieving a low enough p-value.

## 1 CHOOSE A POLITICAL PARTY

Republicans

Democrats

## 2 DEFINE TERMS

Which politicians do you want to include?

☒ Presidents

☒ Governors

☐ Senators

☐ Representatives

How do you want to measure economic performance?

☐ Employment

☒ Inflation

☒ GDP

☒ Stock prices

Other options

☒ Factor in power

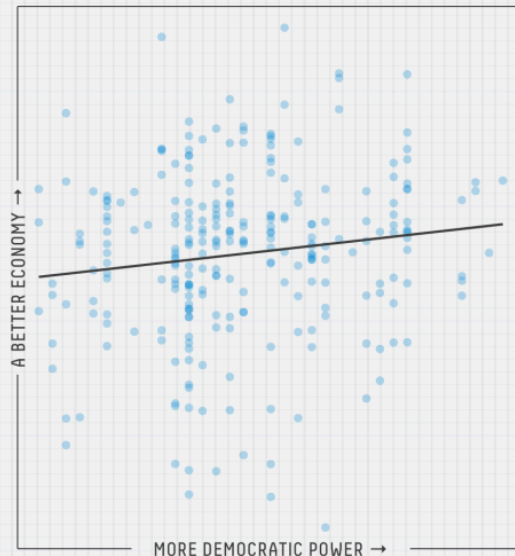
Weight more powerful positions more heavily

☐ Exclude recessions

Don't include economic recessions

## 3 IS THERE A RELATIONSHIP?

Given how you've defined your terms, does the economy do better, worse or about the same when more Democrats are in power? Each dot below represents one month of data.



## 4 IS YOUR RESULT SIGNIFICANT?

If there were no connection between the economy and politics, what is the probability that you'd get results at least as strong as yours? That probability is your p-value, and by convention, you need a **p-value of 0.05 or less** to get published.



**Result: Publishable**

You achieved a p-value of **0.03** and showed that **Democrats** have a **positive** effect on the economy. Get ready to be published!

If you're interested in reading real (and more rigorous) studies on the connection between politics and the economy, see the work of Larry Bartels and Alan Blinder and Mark Watson.

Data from The @unitedstates Project, National Governors Association, Bureau of Labor Statistics, Federal Reserve Bank of St. Louis and Yahoo Finance.

这个项目发布的时候正是一个刚毕业的学生被曝出伪造数据的时候（[hyperlink](#)）。Christie Aschwanden 和 Ritchie King 的重点不是怀疑一个荒唐的结果是如何通过了严格的同行评审的，相反，他们想说的是，科学地做研究并解释数据才是真正的难点（阅读更多）。

## 制作歌曲“你现在在哪里”



纽约时报的这个团队做了一个有关 Justin Bieber 的访谈，不仅不错，还很有趣。



尽管 Bieber 更像一个配角，因为 Diplo 和 Skrillex 才是详尽研究如何制作一首大卖歌曲的人，但是这个配在视频旁边的音乐可视化作品可以帮助你更好地理解音乐家们到底在说什么（阅读更多）。

## 麻疹如何在接种了疫苗的儿童中传播及何时不传播

卫报的 Rich Harris, Nadja Popovich 和 Kenton Powell 展示了当一个国家的孩子们不接种麻疹疫苗时会发生什么。

 vaccinated
  susceptible
  vaccinated but susceptible
  infected
  contact with an infected person



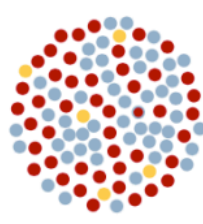
NOT PROTECTED

10.0% vax rate



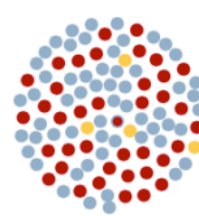
NOT PROTECTED

30.0% vax rate



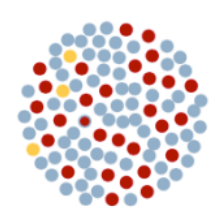
NOT PROTECTED

50.0% vax rate



NOT PROTECTED

58.5% vax rate, similar to Okanagan County, WA



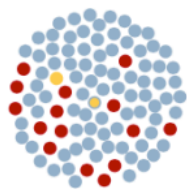
NOT PROTECTED

68.9% vax rate, similar to Thurston County, WA



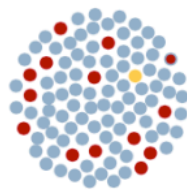
NOT PROTECTED

74.4% vax rate, similar to Island County, WA



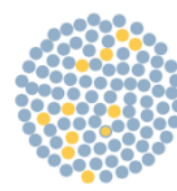
NOT PROTECTED

83.8% vax rate, similar to Santa Cruz County, CA



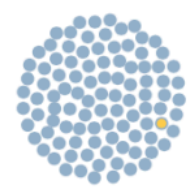
NOT PROTECTED

86.0% vax rate, similar to Los Angeles County, CA



PROTECTED

90.0% vax rate, similar to Orange County, CA



PROTECTED

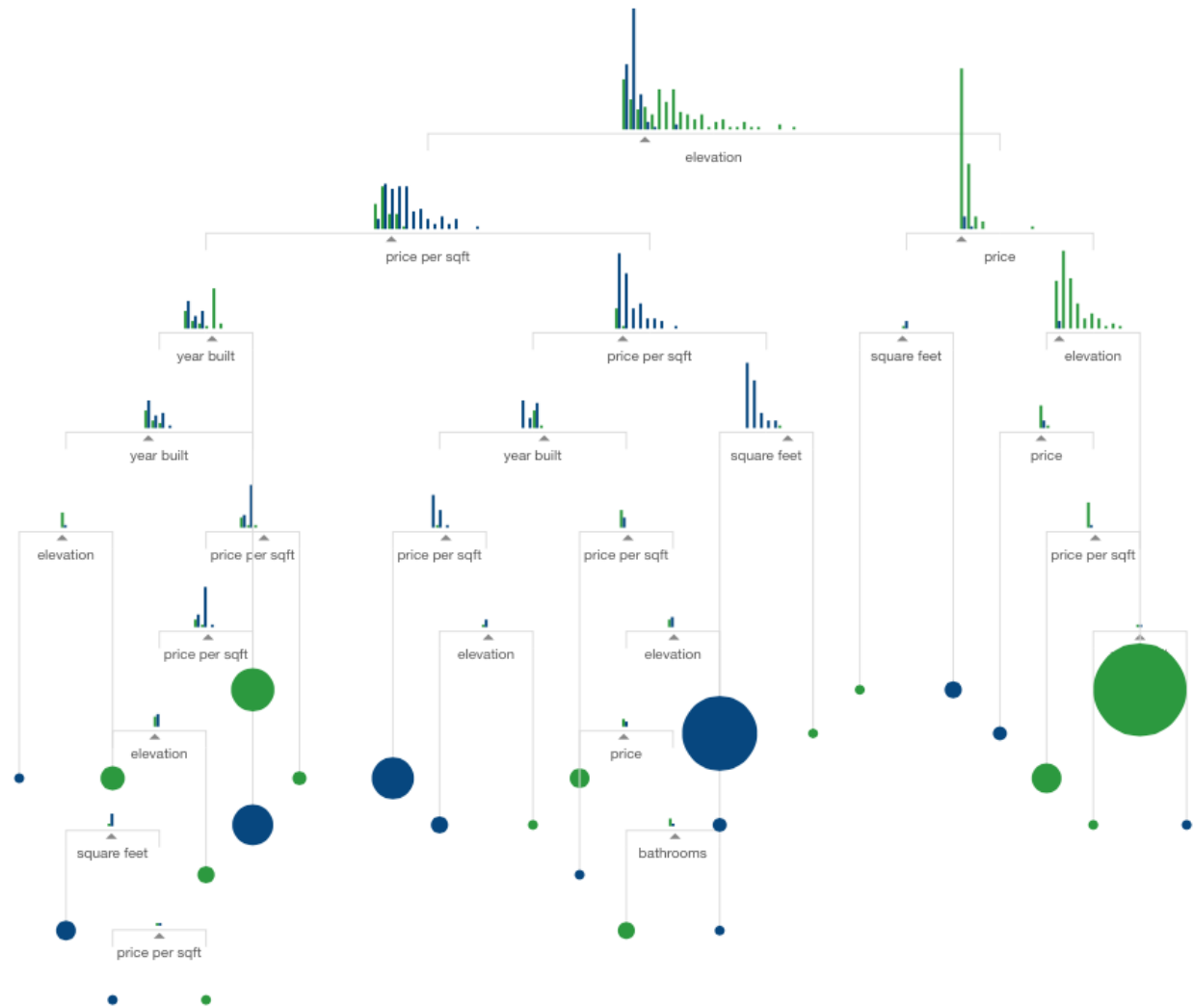
99.7% vax rate, similar to Gadsden County, FL

作为父母，我想把这整个列表做成交互式的。

## 机器学习的可视化介绍

机器学习似乎像一个有魔力的概念，仿佛意味着一个机器人可以不受你的教导去做奇怪的事。

Stephanie Yee 和 Tony Chu 对此用一个可视化例子进行了解谜。

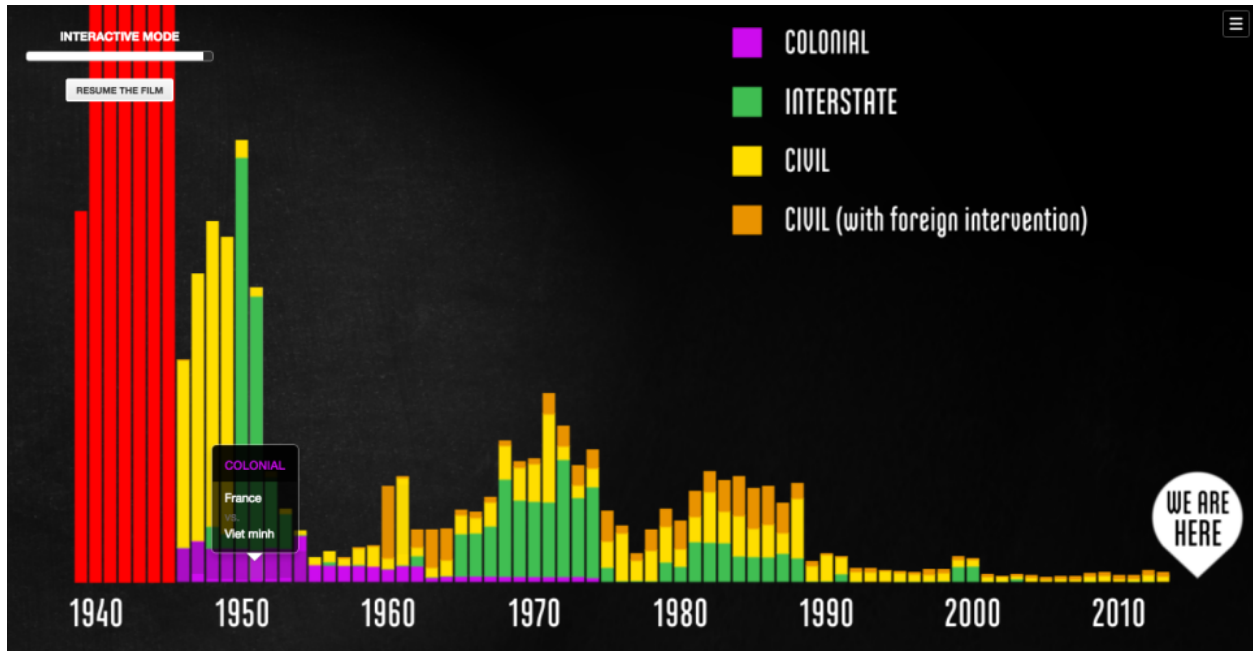


这个可视化例子如卷轴般带着你一步一步了解机器是如何“学习”的。过渡图表让整个图衔接得非常顺畅。现有成果似乎是一个系列项目的第一部分，但是我们可能得等一段时间才能看到后面的。

## 第二次世界大战的结束

Neil Halloran 的这个项目的一部分是记录，另一部分是交互式可视化，二者无缝衔接在了一起。

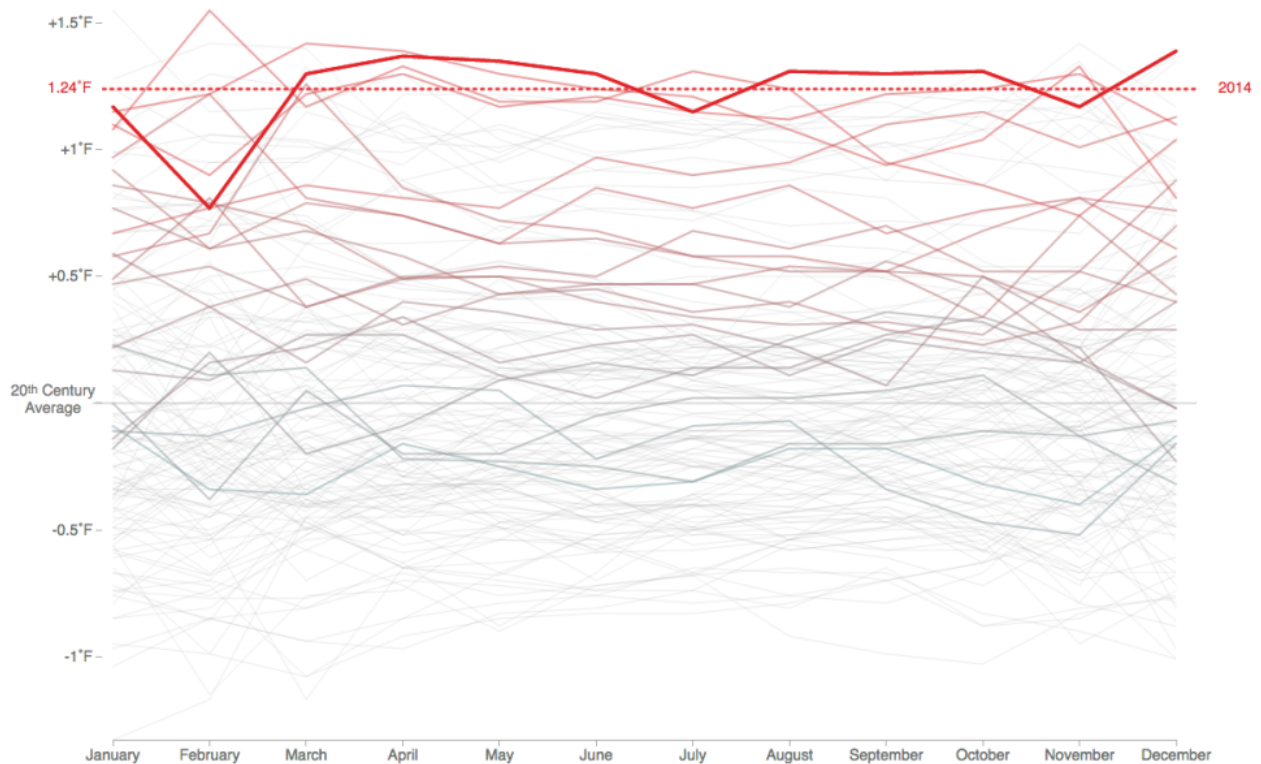




我很惊讶地发现并没有很多人做这个类型的项目。当我意识到他们在做这样一个项目时，我非常希望这个能继续下去。

## 2014 年是有记录以来最热的一年

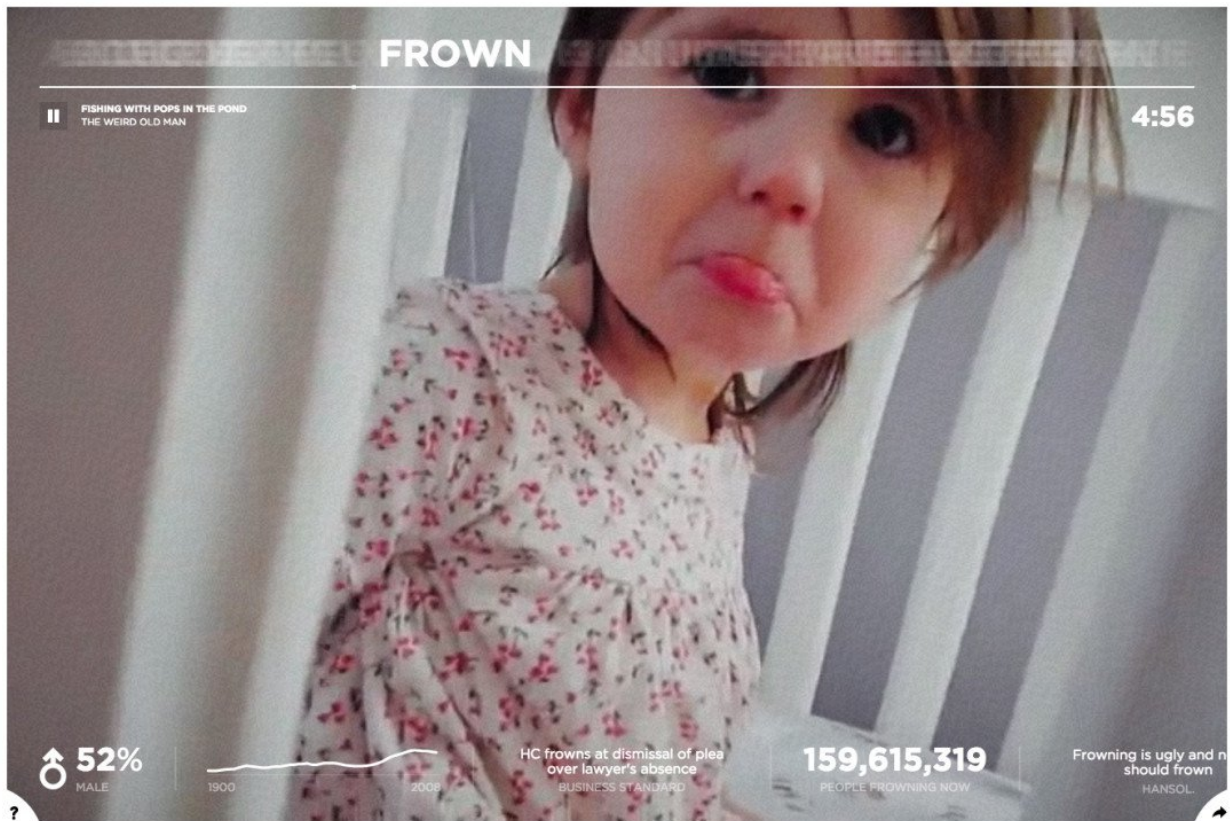
最直观的可视化这类数据的方式是单线图。但把线进行分解后，我们可以得到更多的信息。



彭博社的 Tom Randall 和 Blacki Migliozi 做的这张动态图展示了月度平均气温。每条线代表一个完整的年份，随着时间由远到近，这条线在几英寸几英寸地升高。

## 网络效应

距离上一次看到 Jonathan Harris 的这类项目已经过去一段时间了



他和 Greg Hochmuth 合作了这个项目，“网络效应”是对整个互联网的点评，可以让人通过一种奇妙而引人入胜的方式了解互联网的方方面面，一次花几分钟）。

## 常用比喻

“比喻修辞是写作者常用的，在读者脑海中有特定形象的，一种工具和写作习惯”。Bocoup 数据可视化团队做的“常用比喻”项目研究了比喻修辞中常用的词。

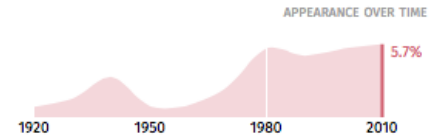




## Mama Bear

Bears usually won't attack humans — but get between a mother bear and her cub, and she'll tear straight through you. Threaten her children, and you are in for a world of hurt. Righteous awesomeness will ensue, and the heroine reclaims her child with a tearful embrace. Oftentimes, when a previously perceived meek mother (or My Beloved...

[Read more on tvtropes.org >>](#)



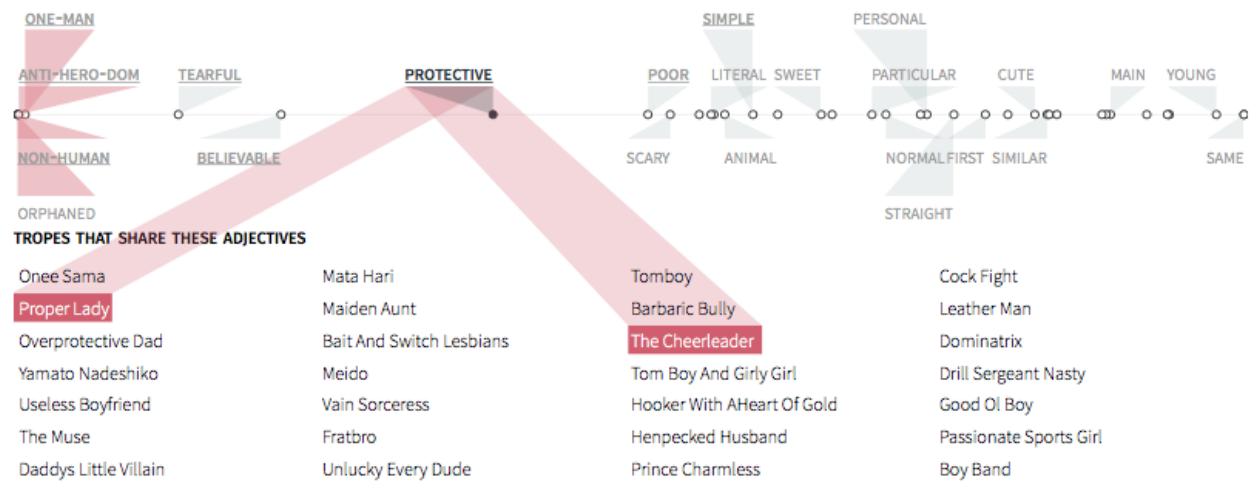
Mama Bear appears in 37 films of the 2010s (5.7% of films in this decade).

### ADJECTIVES DESCRIBING THIS TROPE

The description above is a short snippet of what was written about this trope on [tvtropes.org](#). We wanted to examine more closely the adjectives used in the full description. Adjectives closer to the left are more associated only with this trope, and ones further to the right are more generic. Those adjectives that have blue or pink triangles beneath them appear only in this trope, and nowhere else. Read about how we found these associations [here](#).

← More strongly associated with this trope

Less strongly associated with this trope →



如果你了解电影里的性别角色和人物性格，这个项目正是你要找的那个。

原文链接: <http://flowingdata.com/2015/12/22/10-best-data-visualization-projects-of-2015/>

原文作者: NATHAN YAU

译者: Fibears