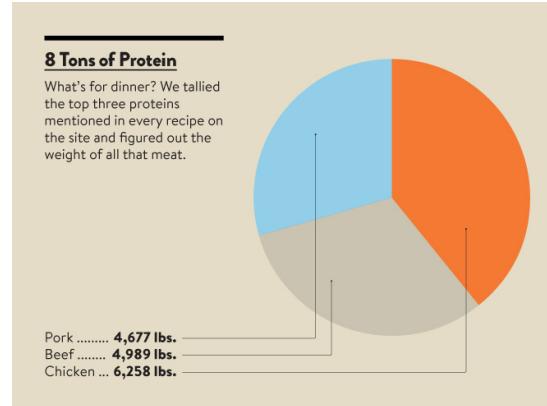




[CLIFF KUANG](#) DESIGN 10.29.13 6:30 AM

# 25 AMAZING FOOD INFOGRAPHICS, DRAWN FROM 49,733 RECIPES



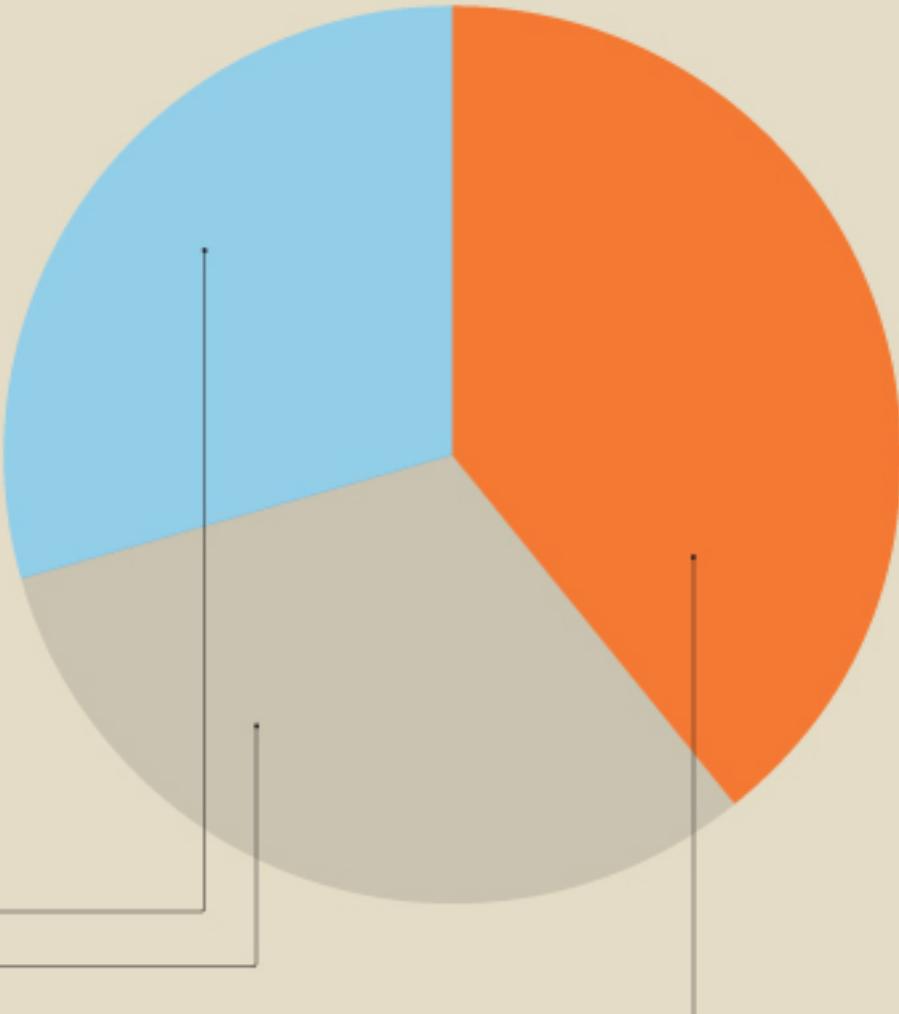
5 / 27 We figured that a good way to see what proteins people were eating was to add up the amounts of chicken and beef. Fish was far less popular.

---

EVERY DAY, GOOGLE'S servers crawl the web, quietly amassing a mirror image of the internet, so that the search behemoth can index everything and serve up lightning-fast responses to any question, no matter how inane. But there's a

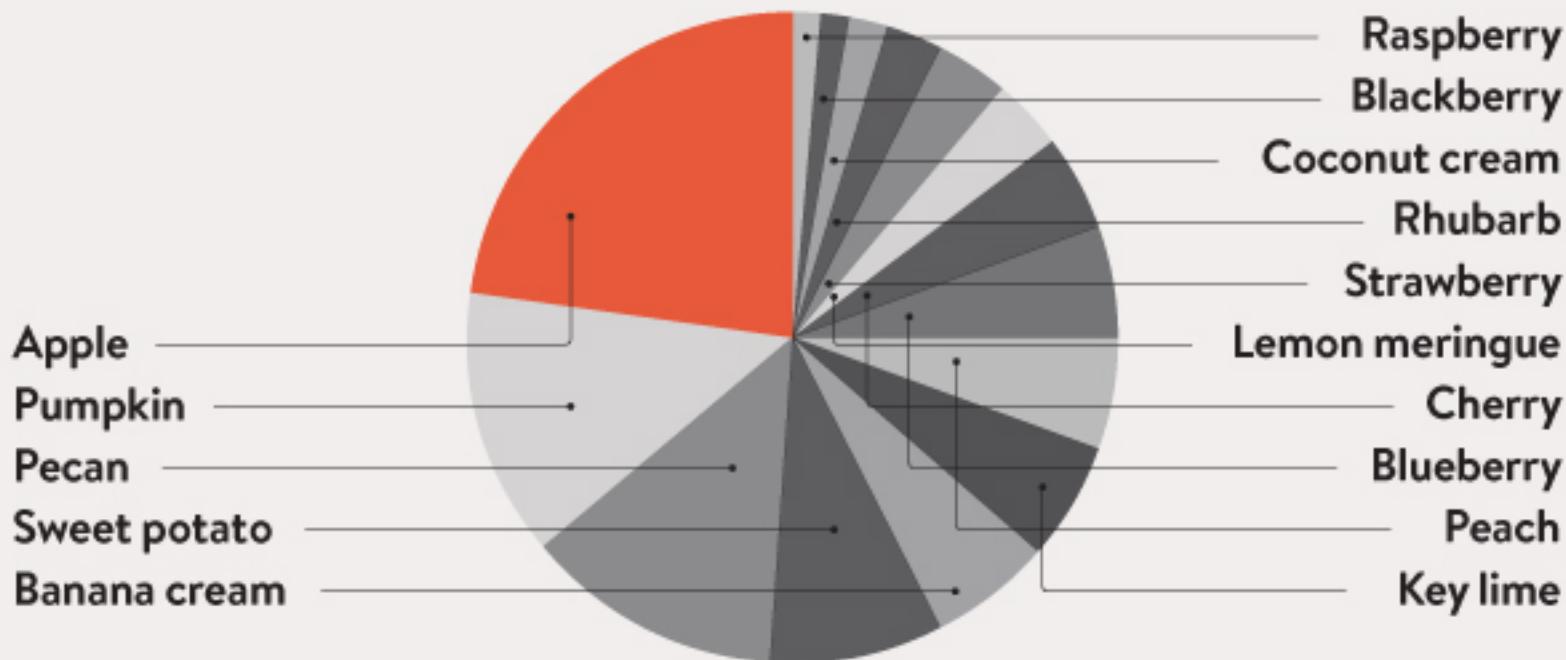
## **8 Tons of Protein**

What's for dinner? We tallied the top three proteins mentioned in every recipe on the site and figured out the weight of all that meat.



## A Pie Chart of Pies

All of Food Network's 468 pie recipes, broken down by delicious filling.



dark side to this process: The countless marketing companies and hackers who write web-crawling scripts to gather massive data sets that serve their own ends.

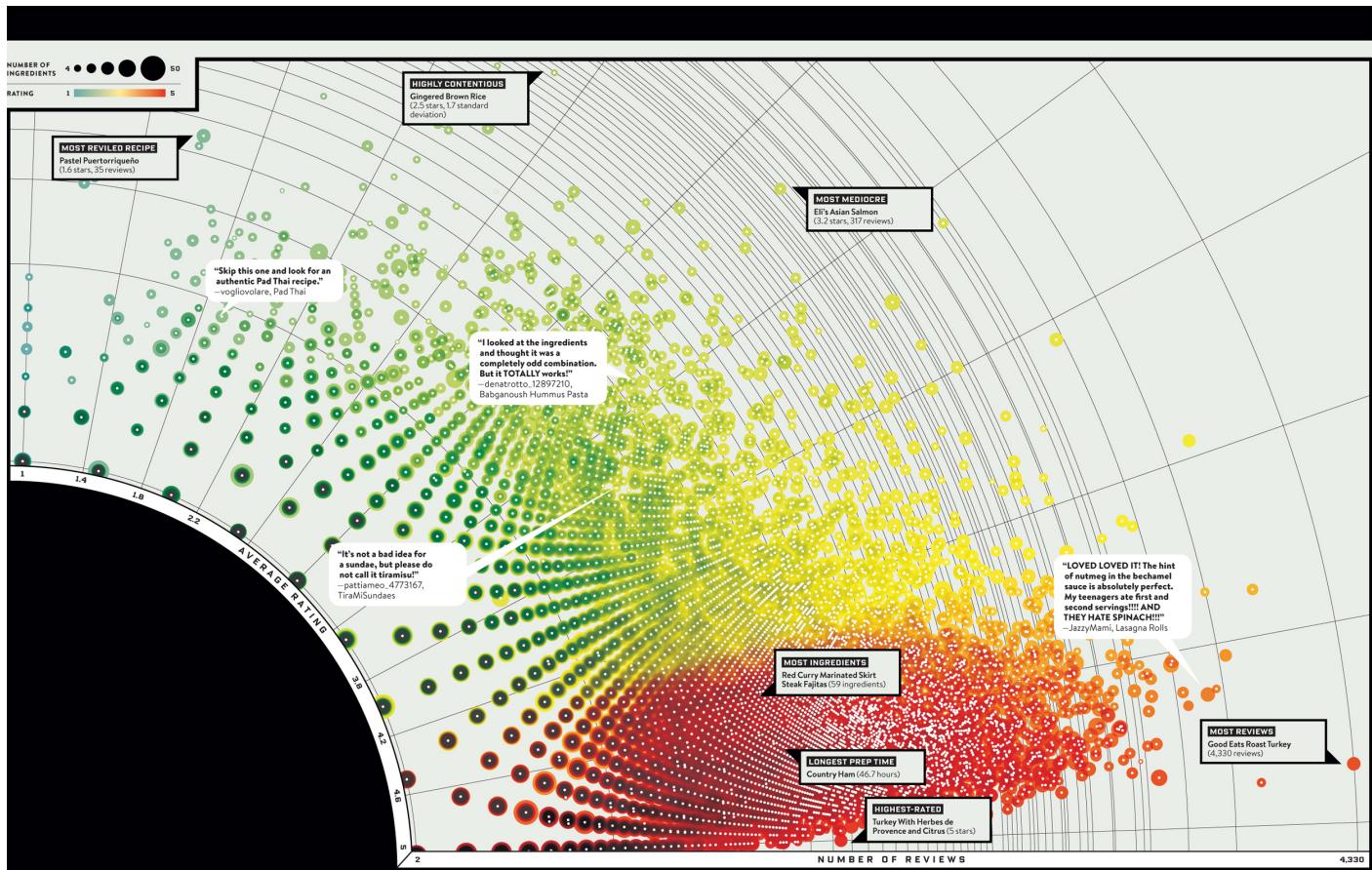
So we wondered: How could we take that same web-crawling process and subvert it? Could we scrape a decently massive dataset and produce something wonderful?

We hit upon a ripe target: [Food Network](#) has amassed one of the richest repositories of cookery available today: Its website racks up over 200 million pageviews a month. But go try and find the perfect Bolognese recipe in 10 minutes. You can't. There's simply too much information, and it's virtually impossible to extract any trends or heuristics from the dumb progression of web pages. This is the state of the web in a nutshell.

Things quickly got complicated. You can't simply go out and scrape a massive site like the Food Network's without getting sued—those voluminous terms of service agreements that you find at the bottom of most websites are designed to prevent anyone from taking data and republishing it. So we asked Food Network very, very nicely: Would you be willing to let us scrape your data, with the aim of creating as many infographics as we can dream up? Pretty please? Amazingly, Food Network agreed. (Thanks Danielle!)

Then we got to work. First, we hired a world-class data-miner, [Dylan Fried](#). He employed tools that are fairly common on the web, if you know where to look. In particular, he used a bunch of [Python web-scraping scripts](#), to crawl all 49,733 Recipes and 906,539 Ratings on Foodnetwork.com, then he dumped those into Mongo, a non-relational database that let us do all kinds of crazy queries.

Maybe the most visually stunning thing we were able to create was a chart showing the structure behind every single one of those 49,733 recipes. On the x-axis are the number of ratings; on the curved axis are the average ratings for each recipe:



Joseph Reyes

As you can see, there's structure there that you would ordinarily never be able to see. You can spy outliers, and spot the clusters where the data gets super dense. You can see how there's a massive clumping of recipes that are all in the range of being pretty good—that is, which have an average rating somewhere above four stars.

Obviously, we didn't stop there. With just a few lines of code, we were able to ask the database some wacky questions, such as: How do all the celebrity chefs on Food Network stack up? Which foods are popular in various regions across the U.S.? And of course, is everything really better with bacon? The 26 infographics you see above, created by Josef Reyes and Catalogtree, represent some of our coolest findings, gathered over a three month period that involved hundreds of different queries, and many false leads. Our aim was to shed light on how Americans eat, using the database we had amassed: We figured out a way to create a report card for all the chefs on the network; visualized all the top recipes you might make for Thanksgiving; and dissected the food trends that have waxed and waned in the last six years. You can see

the fruit of all that labor in the slides above, and captions detailing how each one was made. There's plenty of tasty nuggets. Enjoy!

DATA MINING: DYLAN FRIED; INFOGRAPHICS: JOSEF REYES; DATA VISUALIZATION: CATALOGTREE

---

#GALLERY

---

---

[VIEW COMMENTS](#)

---

## SPONSORED STORIES



SMART ASSET

How Much House Can You Afford? The Answer May Surprise You.



PC MAG

Steam Machines Finally Available for Pre-Order



BBC NEWS

Millions of Millennials making Millions



KIAWAH ISLAND REAL ESTATE

Kiawah Island Ranks as a Best Place for Second Homes Five Years Straight



10 Awesome Philly Attractions With Free Admission