

Welcome to TimesPeople
Get Started

TimesPeople Lets You Share and Discover the Best of NYTimes.com

9:28 AM

Recommend

HOME PAGE TODAY'S PAPER VIDEO MOST POPULAR TIMES TOPICS

Get Home Delivery Log In Register Now

The New York Times

Technology

Search All NYTimes.com

Go



WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS

Search Technology

Go

Inside Technology

Internet | Start-Ups | Business Computing | Companies

Bits Blog »

Personal Tech »

Cellphones, Cameras, Computers and more

VisitCurrituck.com

Advertise on NYTimes.com

For Today's Graduate, Just One Word: Statistics

By STEVE LOHR
Published: August 5, 2009

MOUNTAIN VIEW, Calif. — At Harvard, Carrie Grimes majored in anthropology and archaeology and ventured to places like Honduras, where she studied Mayan settlement patterns by mapping where artifacts were found. But she was drawn to what she calls “all the computer and math stuff” that was part of the job.

Enlarge This Image



Thor Swift for The New York Times

Carrie Grimes, senior staff engineer at Google, uses statistical analysis of data to help improve the company's search engine.

“People think of field archaeology as Indiana Jones, but much of what you really do is data analysis,” she said.

Now Ms. Grimes does a different kind of digging. She works at [Google](#), where she uses statistical analysis of mounds of data to come up with ways to improve its search engine.

Ms. Grimes is an Internet-age statistician, one of many who are changing the image of the profession as a place for dronish number nerds. They are finding themselves increasingly in demand — and even cool.

“I keep saying that the sexy job in the next 10 years will be statisticians,” said Hal Varian, chief economist at Google. “And I’m not kidding.”

Multimedia

Jon Kleinberg, 37
EDUCATION: Bachelor's degree in computer science, Cornell. Ph.D. in computer science, M.I.T.; a MacArthur Fellow.
CAREER: He is a professor of computer science at Cornell who mines huge data sets on the Web.

Graphic

Data Sleuths in an Internet Age

The rising stature of statisticians, who can earn \$125,000 at top companies in their first year after getting a doctorate, is a byproduct of the recent explosion of digital data. In field after field, computing and the Web are creating new realms of data to explore — sensor signals, surveillance tapes, social network chatter, public records and more. And the digital data surge only promises to accelerate, rising fivefold by 2012, according to a projection by IDC, a research firm.

Yet data is merely the raw material of knowledge. “We’re rapidly entering a world where everything can be monitored and measured,” said Erik Brynjolfsson, an economist and director of the [Massachusetts Institute of Technology's](#) Center for Digital Business. “But

Next Article in Technology (1 of 20) »

SIGN IN TO RECOMMEND

SIGN IN TO E-MAIL

PRINT

REPRINTS

SHARE

ARTICLE TOOLS SPONSORED BY

A story about two strangers.

Customize your headlines.

The New York Times **nytimes.com** | **LinkedIn**

The New York Times and LinkedIn have teamed up to give you customized industry news. Automatically.

[Click here to activate](#)

It takes away choice.

Today's Headlines Daily E-Mail

Sign up for a roundup of the day's top stories, sent every morning.

[Sign Up](#)

[See Sample](#) | [Privacy Policy](#)

Subscribe to Technology RSS Feeds

- Technology News
 - Internet
 - Start-Ups
 - Business
 - Companies
 - Computing
- Bits Blog
- Personal Tech
- Pogue's Posts

MOST POPULAR - TECHNOLOGY

E-MAILED BLOGGED



Daniel Rosenbaum for The New York Times
T-shirts for sale at the Joint Statistical Meetings in Washington this week.

the big problem is going to be the ability of humans to use, analyze and make sense of the data.”

The new breed of statisticians tackle that problem. They use powerful computers and sophisticated mathematical models to hunt for meaningful patterns and insights in vast troves of data. The applications are as diverse as improving Internet search and online advertising, culling gene sequencing information for cancer research and analyzing sensor and location data to optimize the handling of food shipments.

Even the recently ended [Netflix](#) contest, which offered \$1 million to anyone who could significantly improve the company's movie recommendation system, was a battle

waged with the weapons of modern statistics.

Though at the fore, statisticians are only a small part of an army of experts using modern statistical techniques for data analysis. Computing and numerical skills, experts say, matter far more than degrees. So the new data sleuths come from backgrounds like economics, computer science and mathematics.

They are certainly welcomed in the White House these days. “Robust, unbiased data are the first step toward addressing our long-term economic needs and key policy priorities,” [Peter R. Orszag](#), director of the [Office of Management and Budget](#), declared in a speech in May. Later that day, Mr. Orszag confessed in a [blog entry](#) that his talk on the importance of statistics was a subject “near to my (admittedly wonkish) heart.”

[I.B.M.](#), seeing an opportunity in data-hunting services, created a Business Analytics and Optimization Services group in April. The unit will tap the expertise of the more than 200 mathematicians, statisticians and other data analysts in its research labs — but that number is not enough. I.B.M. plans to retrain or hire 4,000 more analysts across the company.

In another sign of the growing interest in the field, an estimated 6,400 people are attending the statistics profession's annual conference in Washington this week, up from around 5,400 in recent years, according to the American Statistical Association. The attendees, men and women, young and graying, looked much like any other crowd of tourists in the nation's capital. But their rapt exchanges were filled with talk of randomization, parameters, regressions and data clusters. The data surge is elevating a profession that traditionally tackled less visible and less lucrative work, like figuring out life expectancy rates for insurance companies.

Ms. Grimes, 32, got her doctorate in statistics from Stanford in 2003 and joined Google later that year. She is now one of many statisticians in a group of 250 data analysts. She uses statistical modeling to help improve the company's search technology.

For example, Ms. Grimes worked on an algorithm to fine-tune Google's crawler software, which roams the Web to constantly update its search index. The model increased the chances that the crawler would scan frequently updated Web pages and make fewer trips to more static ones.

1. [For Today's Graduate, Just One Word: Statistics](#)
2. [Sony to Cut E-Book Prices and Offer New Readers](#)
3. [Online Scammers Prey on the Jobless](#)
4. [State of the Art: New Entry in E-Books a Paper Tiger](#)
5. [Bits: Has Google Found a Stairway to Heaven in Brooklyn?](#)
6. [Gadgetwise: App of the Week: A Concertgoer's Band Aide](#)
7. [Gadgetwise: Three Apps to Take to the Wine Seller](#)
8. [Bits: Google's Android Jumps to the Living Room](#)
9. [Bits: LEDs Are As Energy Efficient as Compact Fluorescents](#)
10. [Bits: Internet Ad Spending Shrinks Again](#)

[Go to Complete List »](#)



Top 10 summer sunglasses

ALSO IN T MAGAZINE »

[Reinventing a sofa](#)
[How to get haute couture hair](#)

[nytimes.com](#)

MAGAZINE



Ads by Google

[what's this?](#)

[Spend Data Analysis Tool](#)

Easy, Web Based Spend Data Analysis App. Low Cost. Fast Deployment.

www.Ariba.com/go/Visibility

[Study Background Painting](#)

Learn Background Painting at an Art Institutes School. Apply Now!

www.artinstitutes.edu

[MBA Degree Online](#)

The Most Affordable Online Master's Business Administration
www.WGU.edu

INSIDE NYTIMES.COM

The goal, Ms. Grimes explained, is to make tiny gains in the efficiency of computer and network use. "Even an improvement of a percent or two can be huge, when you do things over the millions and billions of times we do things at Google," she said.

It is the size of the data sets on the Web that opens new worlds of discovery.

Traditionally, social sciences tracked people's behavior by interviewing or surveying them. "But the Web provides this amazing resource for observing how millions of people interact," said Jon Kleinberg, a computer scientist and social networking researcher at [Cornell](#).

For example, in [research just published](#), Mr. Kleinberg and two colleagues followed the flow of ideas across cyberspace. They tracked 1.6 million news sites and blogs during the 2008 presidential campaign, using algorithms that scanned for phrases associated with news topics like "lipstick on a pig."

The Cornell researchers found that, generally, the traditional media leads and the blogs follow, typically by 2.5 hours. But a handful of blogs were quickest to quotes that later gained wide attention.

The rich lode of Web data, experts warn, has its perils. Its sheer volume can easily overwhelm statistical models. Statisticians also caution that strong correlations of data do not necessarily prove a cause-and-effect link.

For example, in the late 1940s, before there was a polio vaccine, public health experts in America noted that polio cases increased in step with the consumption of ice cream and soft drinks, according to David Alan Grier, a historian and statistician at [George Washington University](#). Eliminating such treats was even recommended as part of an anti-polio diet. It turned out that polio outbreaks were most common in the hot months of summer, when people naturally ate more ice cream, showing only an association, Mr. Grier said.

If the data explosion magnifies longstanding issues in statistics, it also opens up new frontiers.

"The key is to let computers do what they are good at, which is trawling these massive data sets for something that is mathematically odd," said Daniel Gruhl, an I.B.M. researcher whose recent work includes mining medical data to improve treatment. "And that makes it easier for humans to do what they are good at — explain those anomalies."

Andrea Fuller contributed reporting.

[Sign in to Recommend](#)

A version of this article appeared in print on August 6, 2009, on page A1 of the New York edition.

[Next Article in Technology \(1 of 20\) »](#)



[Click here to enjoy the convenience of home delivery of The Times for 50% off.](#)

SIGN IN TO E-MAIL

PRINT

REPRINTS

Ads by Google

[what's this?](#)

[Faster Data Analysis](#)

Find out why WinterCorp is Calling Vertica The Future.
www.vertica.com

Statistician

Professional to assist you with project, research, or dissertation.
www.ElaineBellucci.com

SAS® Statistical Analysis

SAS Provides an Integrated Approach to Data Analysis- Info Kit Explains
www.SAS.com

Related Articles

FROM THE NEW YORK TIMES

Digital Eyes Will Chart Baseball's Unseen Skills
(July 10, 2009)

Answering Baseball's What-Ifs
(April 7, 2009)

F.A.Q.
(February 1, 2009)

Data Analysts Are Mesmerized by the Power of Program R
(January 7, 2009)

FROM AROUND THE WEB

What's This?

MEMEX 1.1

Lies, damn lies andâ€¦ er, damn good jobs

MOTHER JONES

Revenge of the Nerds

PACIFIC EPOCH

MoC Discloses Top Game Export Earners

4 Related Articles »

Powered by Blogger

WORLD »



A Little Bit Country, a Little Bit Wine 'n' Cheese

HOME & GARDEN »



Low-Maintenance Living

ART & DESIGN »



A Loincloth to Set Parisians Aflutter

OPINION »

Op-Ed: Democracy Behind Bars
How prison populations skew the census.

FASHION & STYLE »



Could 'Intervention' Have Helped a Star?

OPINION »



Errol Morris: Seven Lies About Lying (Part 1)