Data and Computational Journalism

Skills to Innovate in a Changing Media Landscape

Charles Berret

Brown Institute for Media Innovation

Columbia University

@cberret

Data, Computation, and Journalism:

Why Data Matters **More Than Ever** for Journalists and Educators

Charles Berret

Brown Institute for Media Innovation
Columbia University

@cberret

Today's Talk

1. What is data journalism and how is it being practiced in newsrooms today?

2. What are journalism schools doing to teach data skills?

3. What are the gaps in data journalism education and how can we address them?

4. What are some of the opportunities for young journalists?

"It's obviously a shocking outcome, in particular because the not just polling data, but all of the **self-proclaimed experts in data journalism**, this **new field of journalism** that has arisen that **claims to only view politics through an empirical lens** rather than through the dirty ideologies or partisan biases that everybody else is burdened with, **assured everybody** that it was overwhelmingly likely that Clinton would win."

Glenn Greenwald
Democracy Now
November 10th, 2016

FLASHBACK
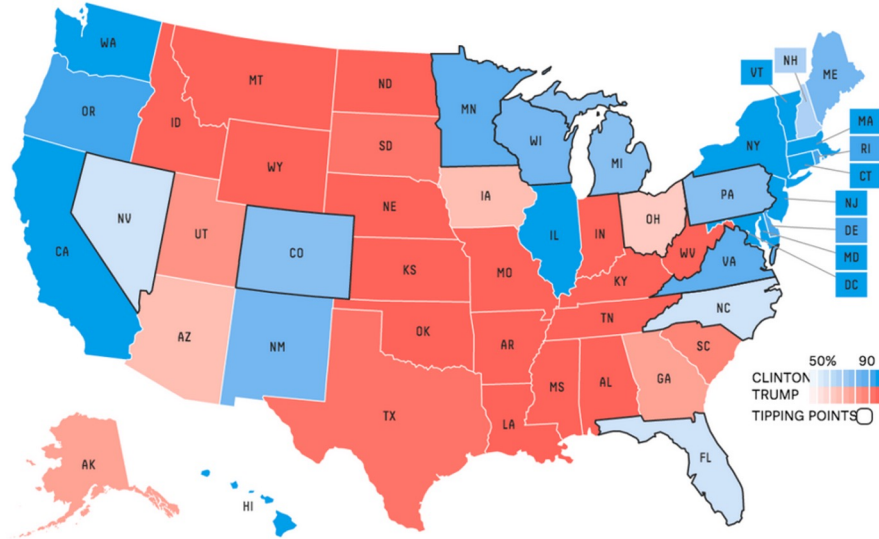
November 7, 2016

Chance of winning

Hillary Clinton
71.4%

Donald Trump
28.6%

Map legend:
50% — 90
CLINTON
TRUMP
TIPPING POINTS

Electoral votes

■ Hillary Clinton		302.2
■ Donald Trump		235.0
■ Evan McMullin		0.8
■ Gary Johnson		0.0

Popular vote

■ Hillary Clinton		48.5%
■ Donald Trump		44.9%
■ Gary Johnson		5.0%
■ Other		1.6%

:TheUpshot   Who Will Win? | Latest Polls | Senate Forecast | Get the Podcast | Primary Results

Ariz. Colo. Fla. Ga. Ill. Ind. Iowa Mich. N.C. N.H. Nev. Ohio Pa. Wis.

Who Will Be President?

By JOSH KATZ

Hillary Clinton has a **74% chance** of winning the presidency.

Last updated Friday, September 30 at 9:12 AM ET

CHANCE OF WINNING

74%
Hillary Clinton

26%
Donald J. Trump

:TheUpshot   Who Will Win? | Latest Polls | Senate Forecast | Get the Podcast | Primary Results

Ariz. Colo. Fla. Ga. Ill. Ind. Iowa Mich. N.C. N.H. Nev. Ohio Pa. Wis.

Who Will Be President?

By JOSH KATZ

CHANCE OF WINNING

Hillary Clinton has a 74% chance of winning the presidency.

Last updated Friday, September 30 at 9:12 AM ET

74%
Hillary Clinton

26%
Donald J. Trump

Forecast history ⌄   Recent changes ⌄   State by state ⌄   Other forecasts ⌄   Likely scenarios ⌄   Explore paths ⌄
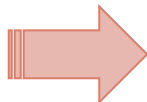
The Upshot's elections model suggests that Hillary Clinton is favored to win the presidency, based on the latest state and national polls. A victory by Mr. Trump remains quite possible: Mrs. Clinton's chance of losing is about the same as the probability that an N.F.L. kicker misses a 45-yard field goal.

Final Election Update: There's A Wide Range Of Outcomes, And Most Of Them Come Up Clinton

By Nate Silver

Filed under 2016 Election

Why the Election Is Close, and What Trump and Obama Have in Common

Nate Cohn @Nate_Cohn NOV. 6, 2016

Waiting to see Donald J. Trump at a campaign event in Hershey, Pa., on Friday.
Damon Winter/The New York Times

Four Key Areas of Data Journalism

Data Reporting

Data Visualizations and Interactives

Computational Journalism

Emerging Technologies

DRUGGING OUR KIDS

Children in California's foster care system are prescribed unproven, risky medications at alarming rates

Story by KAREN DE SÁ
Photographs and video by DAI SUGANO
Design by QIN CHEN

An investigation by this newspaper found that nearly 1 out of every 4 adolescents in California's foster care system is receiving these drugs — 3½ times the rate for all adolescents nationwide. Over the last decade, almost 15 percent of the state's foster children of all ages were prescribed the medications, known as psychotropics, part of a national treatment trend that is only beginning to receive broad scrutiny.

DRUGGING OUR KIDS

Children in California's foster care system are prescribed
unproven, risky medications at alarming rates

Story by KAREN DE SÁ
Photographs and video by DAI SUGANO
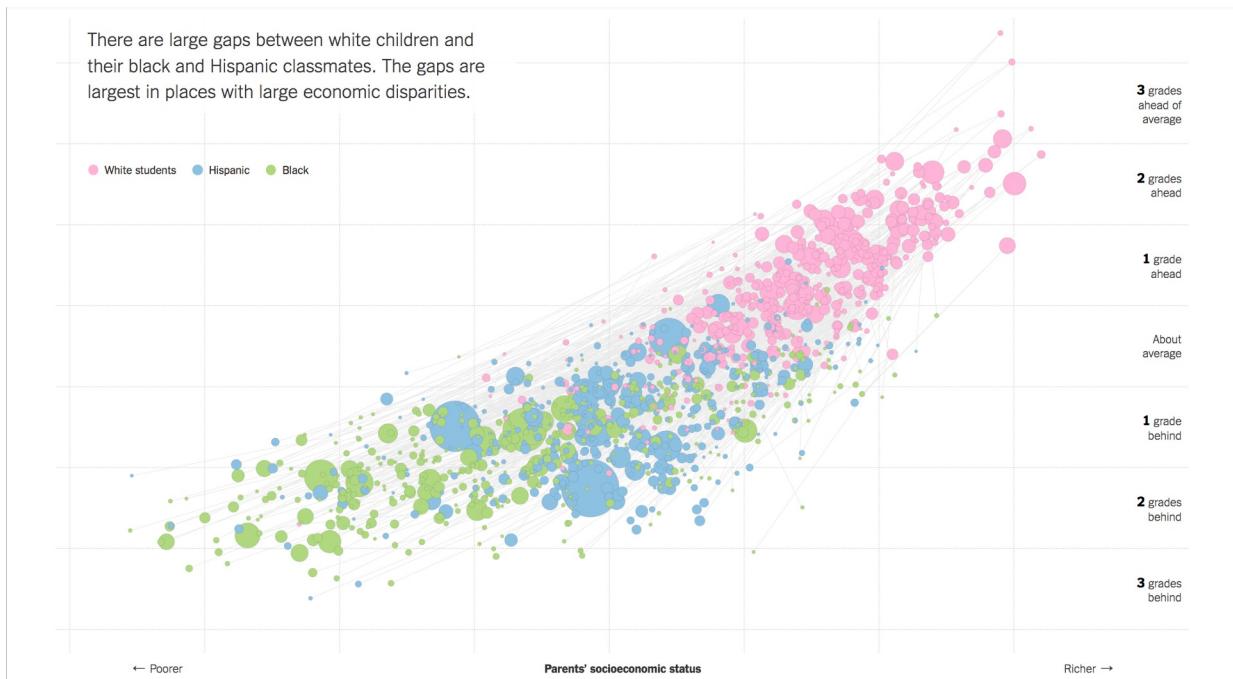Design by QIN CHEN

And Rochelle Trochtenberg, now 31 and living in Eureka, still struggles to bring a glass to her lips because her hands are so shaky from the years she spent on a shifting mix of lithium, Depakote, Zyprexa, Haldol and Prozac, among others. When people ask, she tries to cover it up with remarks about a possible hereditary condition.

The truth is too painful to explain, she said. "I don't want to tell people I have a tremor because I was drugged for my whole adolescence."

Money, Race and Success: How Your School District Compares

By **MOTOKO RICH, AMANDA COX** and **MATTHEW BLOCH**   APRIL 29, 2016

Sixth graders in the richest school districts are four grade levels ahead
of children in the poorest districts.

There are large gaps between white children and
their black and Hispanic classmates. The gaps are
largest in places with large economic disparities.

● White students   ● Hispanic   ● Black

3 grades
ahead of
average

2 grades
ahead

1 grade
ahead

About
average

1 grade
behind

2 grades
behind

3 grades
behind

← Poorer   **Parents' socioeconomic status**   Richer →

Dollars for Docs

How Industry Dollars Reach Your Doctors

By Charles Ornstein, Lena Groeger, Mike Tigas, and Ryann Grochowski Jones, ProPublica. Updated March 17, 2016

Pharmaceutical and medical device companies are now required by law to release details of their payments to a variety of doctors and U.S. teaching hospitals for promotional talks, research and consulting, among other categories. Use this tool to search for general payments (excluding research and ownership interests) made from August 2013 to December 2014. | Related Story: Now There's Proof: Docs Who Get Company Cash Tend to Prescribe More Brand-Name Meds »

Has Your Doctor Received Drug or Device Company Money?

| 🔍 | All States | Search |

For example: Andrew Jones, Boston, 10013

$3.49B in disclosed payments

681,020 doctors

1,135 teaching hospitals

1,565 companies

Top 50 Companies

Click on a company to see how its payments break down by drug, device or doctor. Or, see all companies »

COMPANY	PAYMENTS
Genentech, Inc.	$388M
DePuy Synthes Products LLC	$94.7M
Topera, Inc.	$93.1M
Stryker Corporation	$90.8M
AstraZeneca Pharmaceuticals LP	$90.7M

Highest-Earning Doctors

NAME	PAYMENTS
SUJATA NARAYAN Family Medicine	$43.9M
KAREN UNDERWOOD Pediatric Critical Care Medicine	$28.5M
STEPHEN BURKHART Orthopaedic Surgery	$24M
SANJAY YADAV Cardiovascular Disease	$23.1M
KEVIN FOLEY Neurological Surgery	$22M
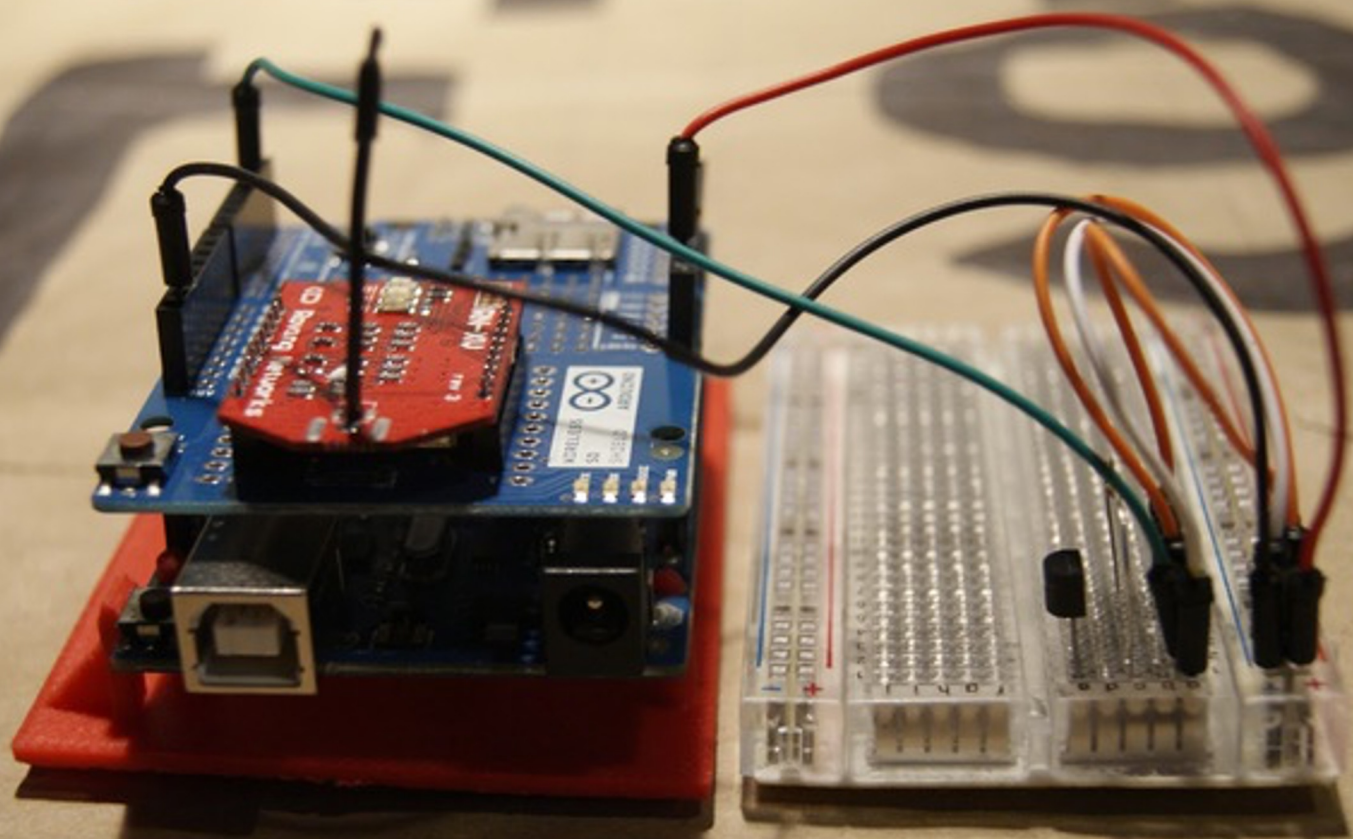
"It's obviously a shocking outcome, in particular because the not just polling data, but all of the **self-proclaimed experts in data journalism**, this new field of journalism that has arisen that **claims to only view politics through an empirical lens** rather than through the dirty ideologies or partisan biases that everybody else is burdened with, **assured everybody** that it was overwhelmingly likely that Clinton would win."

Glenn Greenwald
"Democracy Now"
November 10th, 2016

THE DRONE PAPERS

The Intercept has obtained a cache of secret documents detailing the inner workings of the U.S. military's assassination program in Afghanistan, Yemen, and Somalia. The documents, provided by a whistleblower, offer an unprecedented glimpse into Obama's drone wars.

NOT SO SECURUS

Massive Hack of 70 Million Prisoner Phone Calls Indicates Violations of Attorney-Client Privilege

What are journalism schools doing with data and computation?

The School of Journalism in Columbia University[1]

A Review of Criticisms and Objections—Reflections Upon the Power, the Progress and the Prejudices of the Press—Why Specialized Concentration and Education at College Would Improve the Character and Work of Journalists and So Promote the Welfare of the Republic.

" The man who writes, the man who month in and month out, week in and week out, day in and day out, furnishes the material which is to shape the thoughts of our people, is essentially the man who more than any other determines the character of the people and the kind of government this people shall possess."—PRESIDENT ROOSEVELT, April 7, 1904.

THE editor of *The North American Review* has asked me to reply to an article recently printed in its pages criticising the College of Journalism which I have endowed as part of Columbia University. In complying with his request I have enlarged the scope of the reply to include all other criticisms and misgivings, many honest, some shallow, some based on misunderstanding, but the most representing only prejudice and ignorance. If my comment upon these criticisms shall seem to be diffuse and perhaps repetitious, my apology is that—alas !—I am compelled to write by voice, not pen, and to revise the proofs by ear, not eye— a somewhat difficult task.

"Our Republic and its press will rise or fall together."

"An able, disinterested, public-spirited press, with trained intelligence to know the right and courage to do it, can preserve that public virtue without which popular government is a sham and a mockery."

"A cynical, mercenary, demagogic press will in time produce a people as base as itself."

arbitrate." Before an arbitration tribunal questions are discussed on their merits. Appeals to prejudice, to class or national animosities, to cynical self-interest, are dropped. Every such hearing is a lesson in order and civilization.

There is always a tendency on the part of the weaker side to ask for arbitration and on that of the stronger to refuse it. Here is the opportunity of the press to bring its moral force into the dispute and overcome the obstinacy of brute strength by the pressure of public opinion.

The literature of arbitration is already immense. The workings of experiments in compulsory arbitration, of boards of conciliation, of permanent State arbitration tribunals, of standing arbitration agreements between labor-unions and employers, and of the long line of international settlements leading up to the establishment of the world's court of arbitration at The Hague, would furnish material in themselves for a full and most valuable course of study for a journalist.

Everybody says that statistics should be taught. But how? **Statistics**

Statistics are not simply figures. It is said that nothing lies like figures—except facts. You want statistics to tell you the truth. You can find truth there if you know how to get at it, and romance, human interest, humor and fascinating revelations as well. The journalist must know how to find all these things— truth, of course, first. His figures must bear examination. It is much better to understate than to overstate his case, so that his critics and not himself may be put to confusion when they challenge him to verify his comparisons.

He must not read his statistics blindly; he must be able to test them by knowledge and by common sense. He must always be on the alert to discover how far they can actually be trusted—and what they really mean. The analysis of statistics to get at the

39

"Everybody says that statistics should be taught, but how?"

"You can find truth there if you know how to get at it, and romance, human interest, humor, and fascinating revelations as well. The journalist must know how to find all these things — truth, of course, first."

Joseph Pulitzer, 1904, *The North American Review*

arbitrate." Before an arbitration tribunal questions are discussed on their merits. Appeals to prejudice, to class or national animosities, to cynical self-interest, are dropped. Every such hearing is a lesson in order and civilization.

There is always a tendency on the part of the weaker side to ask for arbitration and on that of the stronger to refuse it. Here is the opportunity of the press to bring its moral force into the dispute and overcome the obstinacy of brute strength by the pressure of public opinion.

The literature of arbitration is already immense. The workings of experiments in compulsory arbitration, of boards of conciliation, of permanent State arbitration tribunals, of standing arbitration agreements between labor-unions and employers, and of the long line of international settlements leading up to the establishment of the world's court of arbitration at The Hague, would furnish material in themselves for a full and most valuable course of study for a journalist.

Everybody says that statistics should be taught. But how? **Statistics**

Statistics are not simply figures. It is said that nothing lies like figures—except facts. You want statistics to tell you the truth. You can find truth there if you know how to get at it, and romance, human interest, humor and fascinating revelations as well. The journalist must know how to find all these things— truth, of course, first. His figures must bear examination. It is much better to understate than to overstate his case, so that his critics and not himself may be put to confusion when they challenge him to verify his comparisons.

He must not read his statistics blindly; he must be able to test them by knowledge and by common sense. He must always be on the alert to discover how far they can actually be trusted—and what they really mean. The analysis of statistics to get at the
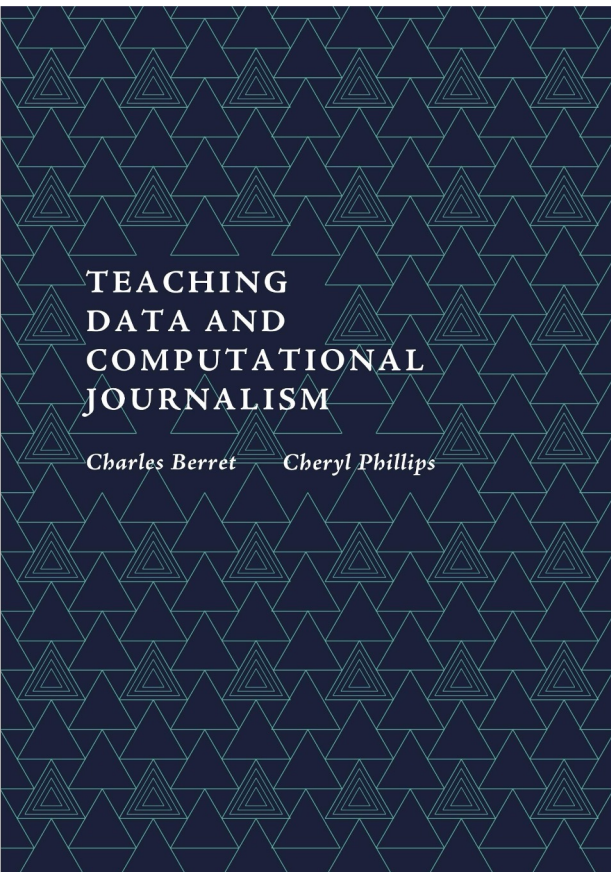
39

"He must not read his statistics blindly; he must be able to test them by knowledge and by common sense. He must always be on the alert to discover how far they can actually be trusted — and what they really mean."

Joseph Pulitzer, 1904, *The North American Review*

NICAR

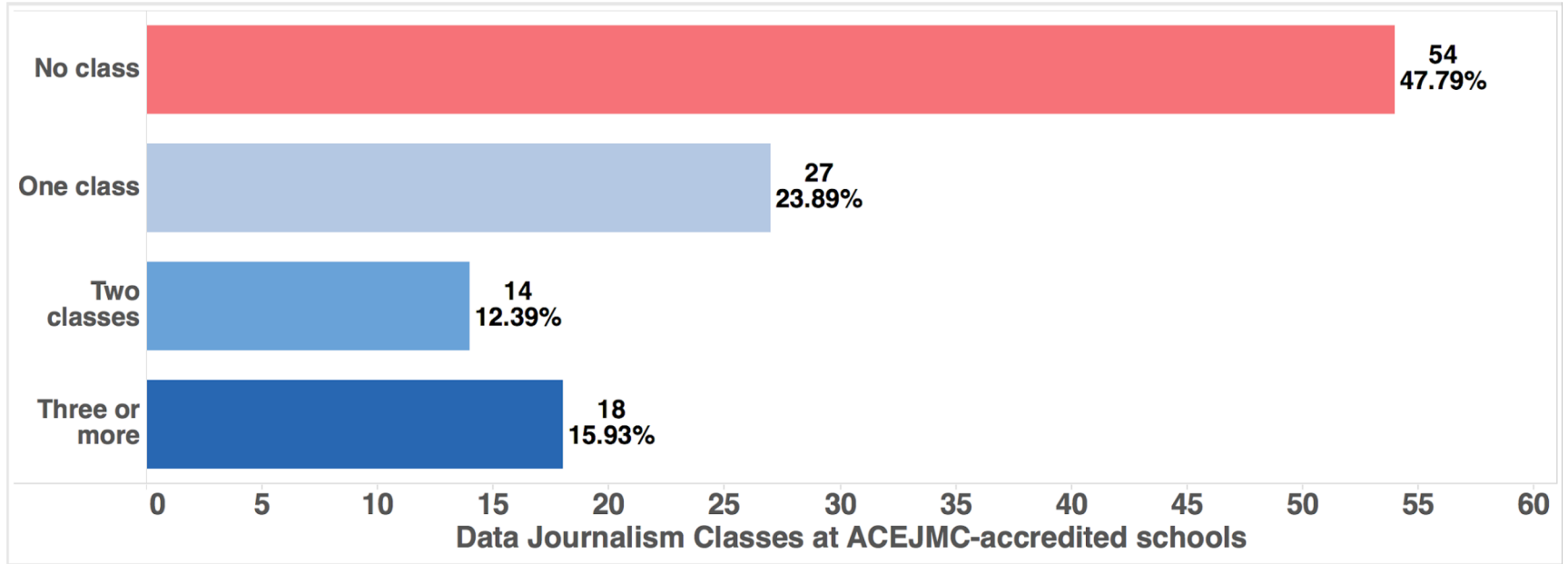training digital sorcerers since 1989

Key Findings

- Nearly half of journalism programs (48 percent) offer no classes that teach data-driven reporting.

- Data classes are largely introductory, and advanced data course offerings are rare.

- Digital multimedia journalism classes are offered at nearly every program (87 percent).

- Most journalism schools (88 percent) offer no instruction in programming (beyond HTML/CSS)

TEACHING
DATA AND
COMPUTATIONAL
JOURNALISM

Charles Berret   *Cheryl Phillips*

Knight Foundation

Methodology

- **Sample**: the 113 graduate and undergraduate journalism programs accredited by ACEJMC.

- **Gathering**: We collected course offerings from each school through their online catalogs and websites, then verified with an administrator at each school.

- **Confirmation**: Even when no data journalism class was listed, we followed up directly with the school to double-check.

- **Coding**: we labeled each class for whether it taught: basic numeracy; data-driven reporting; multimedia production; web coding; data visualization; computational methods; or emerging media technologies.
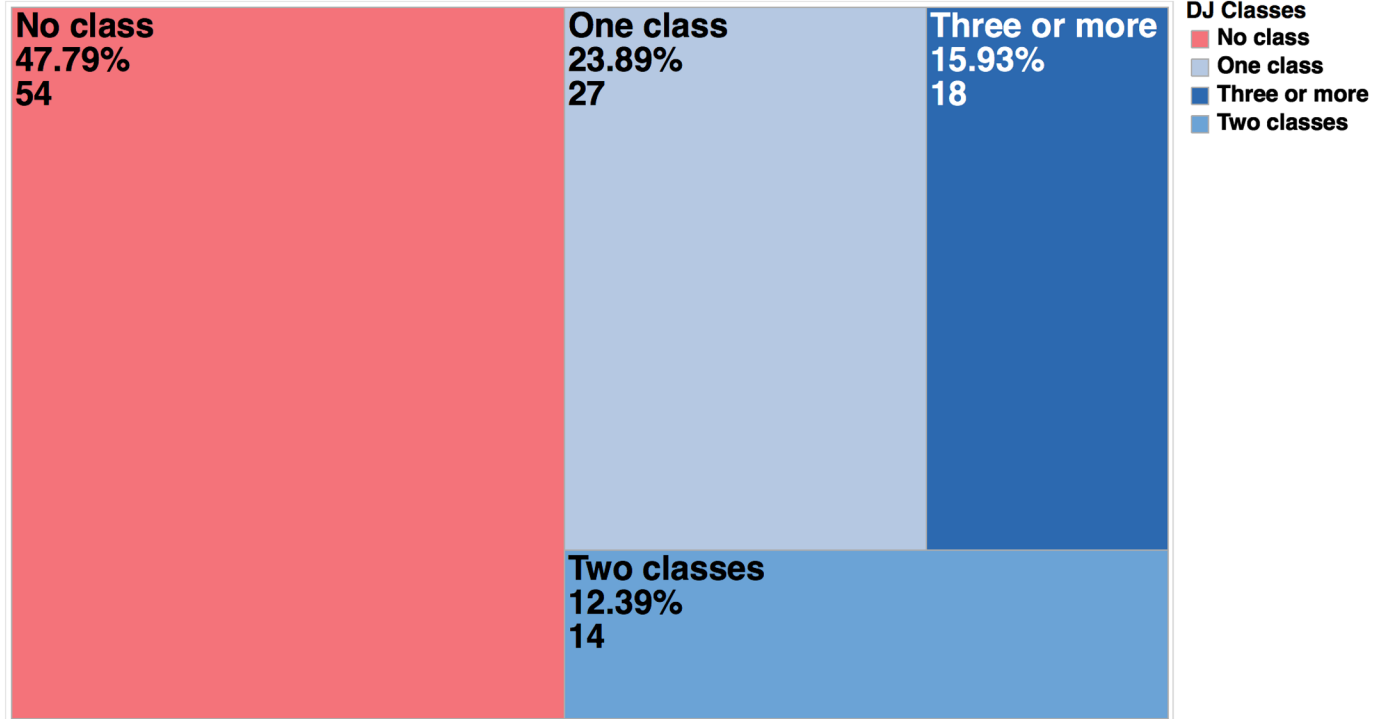
Data Journalism
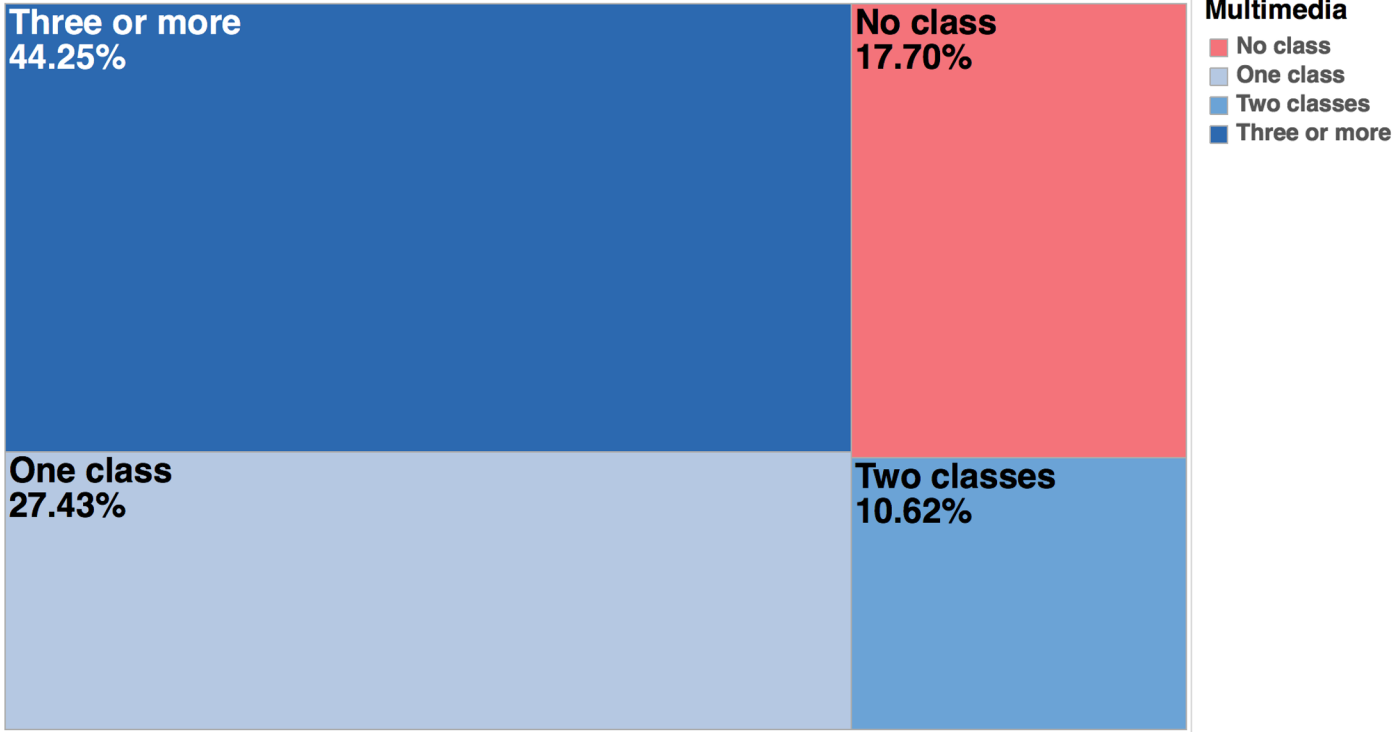## Classes Offered in Accredited US Journalism Programs

Data Journalism Classes at ACEJMC-accredited schools

Category	Count	Percentage
No class	54	47.79%
One class	27	23.89%
Two classes	14	12.39%
Three or more	18	15.93%

Programming Classes (beyond HTML/CSS)
Offered in Accredited US Journalism Programs

Programming
- No class
- One class
- Two classes
- Three or more

No class — 87.61% 99
One class — 5.31% 6
Two classes — 4.42% 5
Three or more — 2.65% 3

Programming beyond HTML/CSS
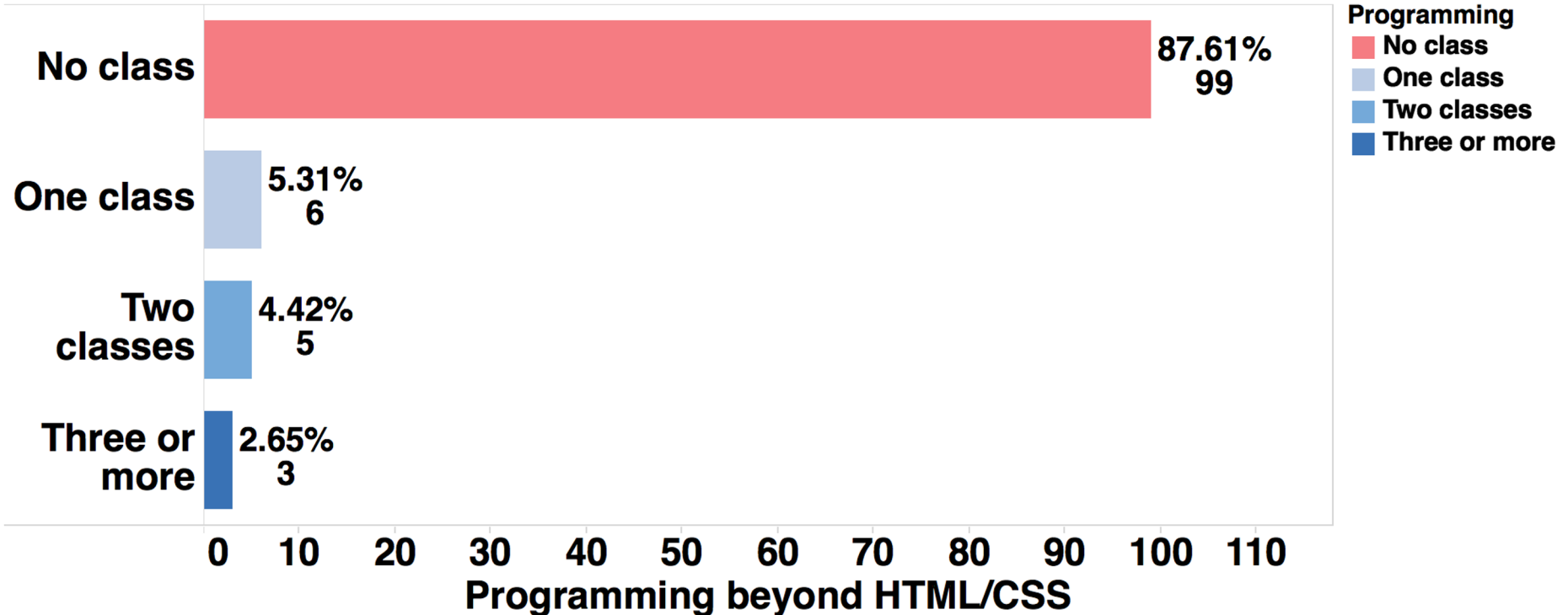
Syllabus Analysis

Topics represented in a collection of 63 data journalism syllabi

Chart showing "Number of Classes Teaching Each Concept" with the following values:

Concept	Number of Classes
Programming Concepts	7
Statistics	10
Design Concepts	11
Relational Databases	13
Spreadsheets	22
Journalistic/Critical Thinking	26
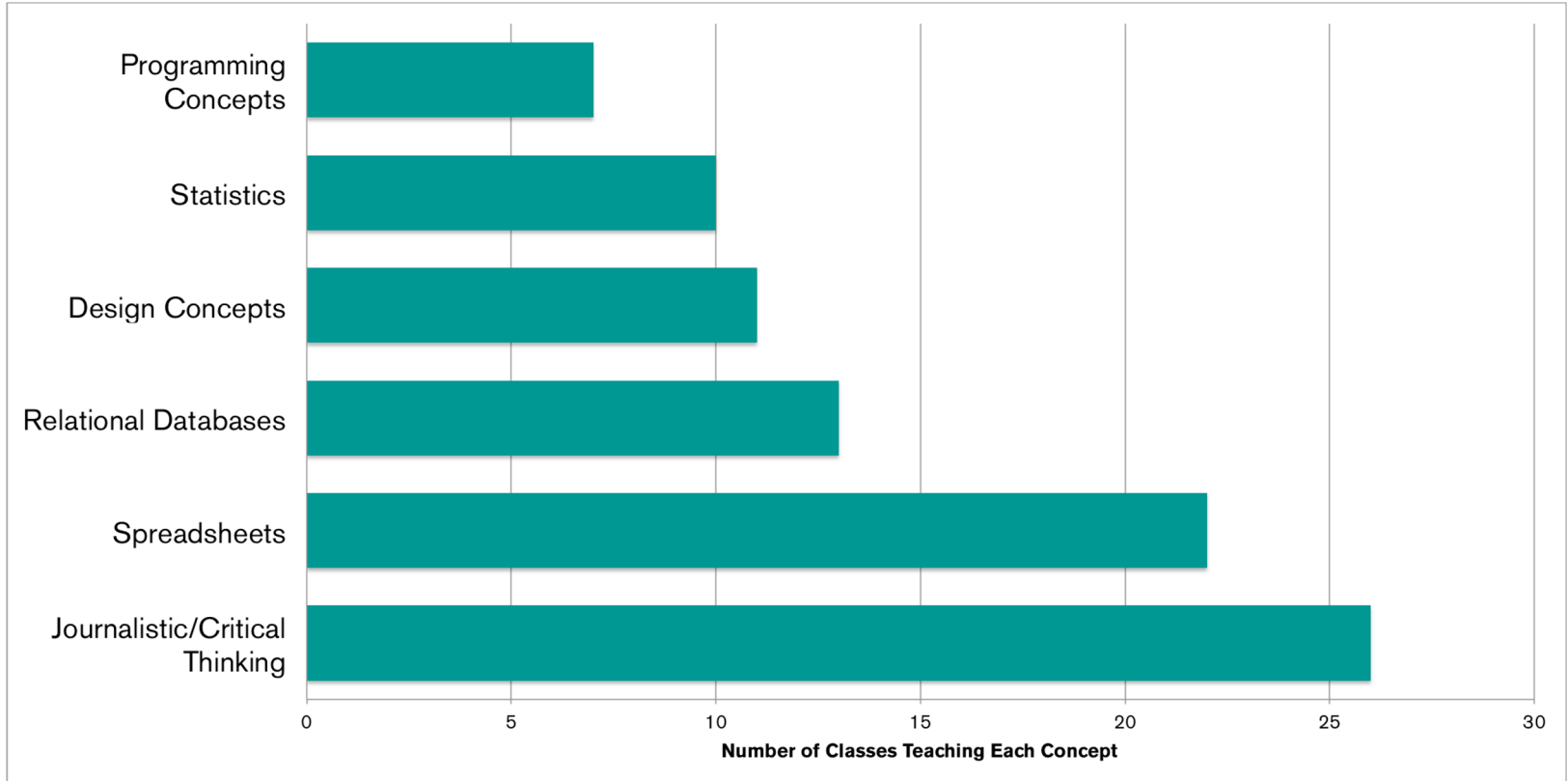
Textbooks

The readings most often assigned in the 63 syllabi we gathered.
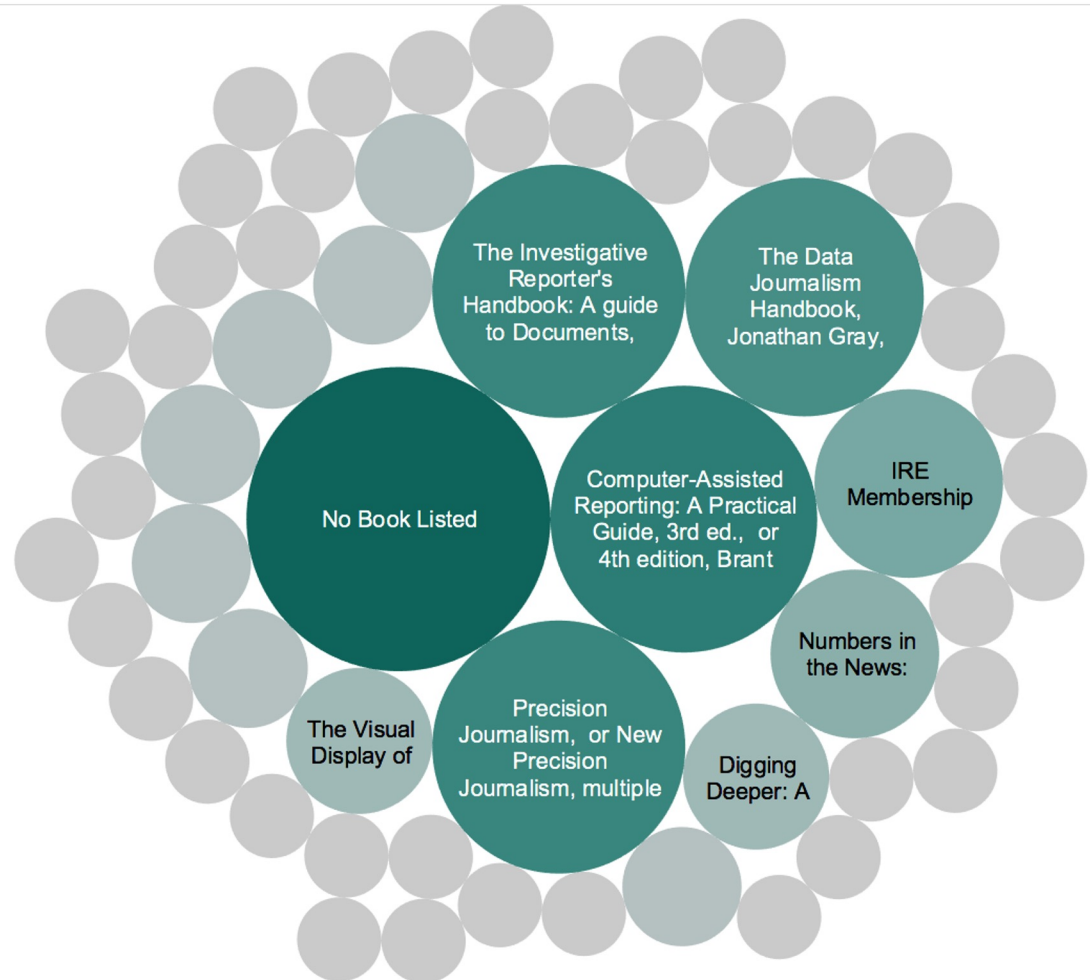
Model Curricula

1. ## One Required Course in Data Journalism

-   **Skills:** data acquisition, data cleaning, exploratory data analysis, basic visualization.
-   **Purpose:** to know when to seek out data as a reporter, and to use it wisely.

1. ## Integrating Data into Existing Courses

-   Media Law and Ethics
-   History of Journalism
-   Investigative Reporting
-   Feature Writing and Long-Form

-   Social Media
-   Business Reporting
-   Global Reporting
-   Science and Environmental Reporting

Model Curricula

3. Concentration in Data & Computational Journalism

So that students may develop a focused proficiency in data and computation.

4. Expertise-Driven Reporting on Data and Computation

Mid-career instruction focused on writing with and about data.

5. Emerging Journalistic Techniques and Technologies

Focus on new platforms and media innovation.

Institutional Challenges

Resources

Some students don't have necessary equipment. Using lab computers often not ideal.

Faculty Expertise

Finding qualified faculty is often difficult. Working journalists not incentivized to adjunct.

**Incoming Skill Levels and Literacies**

Many students actively avoid math, stats, and other technical matters.

Student Engagement

How to get students interested in learning data skills.

Among Our Recommendations

- Every journalism program should have a core, required course on analyzing data to find stories.

- Journalism schools should collaborate across the University to develop new, interdisciplinary reporting capacities.

- Journalism programs must embrace research as well as teaching in their missions for both the school and the profession to evolve.
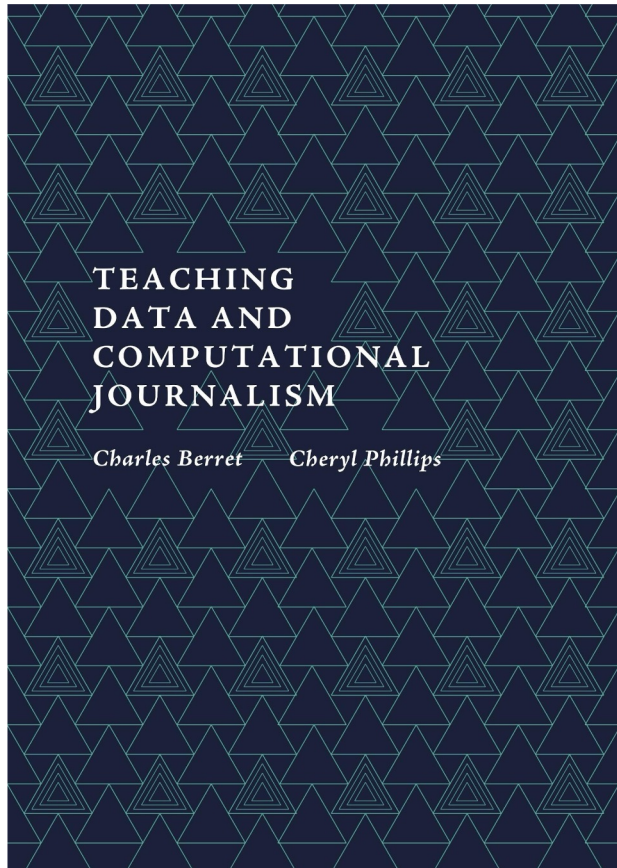
"Teaching Data and Computational Journalism"

By Charles Berret and Cheryl Phillips

Free Download. GitBook and PDF:
http://brwn.co/teachingdata

Data Journalism: A Hot Job Market

Nov. 16, 2016

Narratives with Numbers

"increased role that numerical data plays in producing and distributing information in the digital age"

https://en.wikipedia.org/wiki/Data_journalism
Feb 9, 2016 - **Data journalism** is a **journalism** specialty reflecting the increased role that numerical **data** is used in the production and distribution of information in the digital era. It reflects the increased interaction between content producers (**journalist**) and several other fields such as design, computer science and statistics.

As data becomes increasingly ubiquitous, companies are desperately searching for talent with these data skills. LinkedIn recently reported data analysis is one of the hottest skill categories over the past two years for recruiters, and it was the only category that consistently ranked in the top 4 across *all* of the countries they analyzed. Interestingly, much of the current hiring emphasis has centered on the data preparation and analysis skills—not the "last mile" skills that help convert insights into actions. Many of the heavily-recruited individuals with advanced degrees in economics, mathematics, or statistics struggle with communicating their insights to others effectively—essentially, telling the story of their numbers.

2016 Survey of Data Journalism Jobs

Most popular skills listed on DataJournalismJobs.com
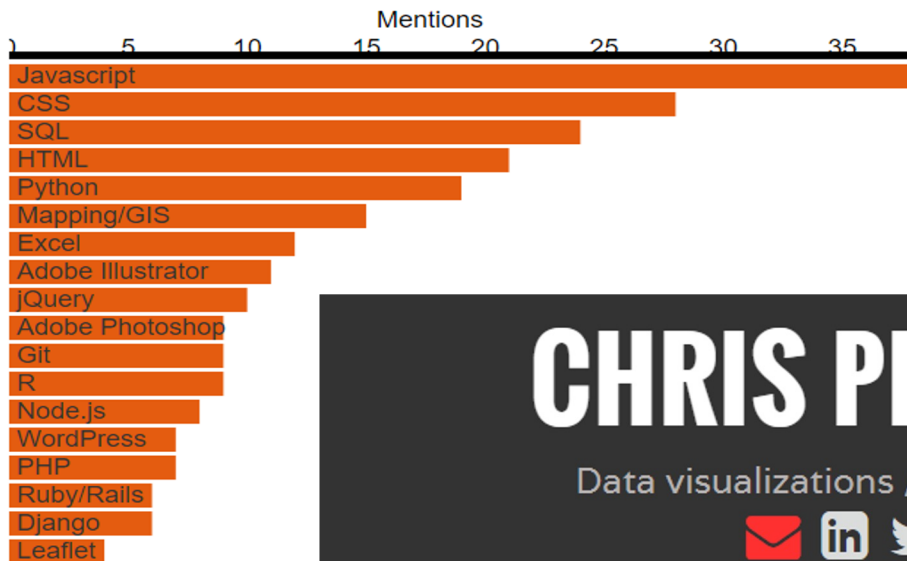
CHRIS PERSAUD

Data visualizations / Data journalism

Data Ninjas Needed

Or this opening from Reuters seeking a data reporter, posted in September. It asks for applicants skilled in Excel, SQL and R.

We want a data ninja who can hunt and slay data, resulting in stories that uncover business secrets, drive markets and challenge titans of industry. You must enjoy collaboration and have a sixth-sense for finding stories that matter. Our data reporters do not merely serve up data; they work with other reporters and editors to conceive stories and shape our reporting.

Like this one from the Los Angeles Times posted in October for a summer 2016 internship, which asks for applicants skilled in SQL and R, among other programming languages.

The Data Desk is a team of reporters and computer programmers that works with journalists in The Times newsroom to collect, organize, analyze and present large amounts of information. In short, we use data to find and tell stories.

Data Journalism Jobs Submit new job

Job	Company	Location	Salary	Skills	Details	Apply	Date submitted
Designer/Developer (Sandbox)	McClatchy	Washington, DC		Please see full job description via posted link.	A core part of the new video division is the "Sandbox" team -- a small group charged with advancing video storytelling and technology far into the future. They will work to enhance existing products and to build new ones. They'll explore the space where text and video merge, the development of new editorials tools and imagine new contexts in which video will play. This group will include a mix of creative thinkers, hackers, visual journalists and product designers -- and you will be the lead developer. We're looking for a full-stack developer with the ability to work across a broad array of technologies.	Apply here	11/4/2016
Data reporter	Las Vegas Review-Journal	Las Vegas, NV		Applicants should have experience with programs needed for data manipulation and presentation -- including Excel, Access, Illustrator, Tableau, Google Fusion – as well as experience with data mapping tools. Intermediate or better coding skills are a definite plus. Applicants should have at least a bachelor's degree in journalism, communications or a related field, reporting experience and a record of delivering results in a deadline environment.	The Las Vegas Review-Journal is seeking a data-driven journalist to work with reporters and editors throughout the newsroom to produce information-rich stories across all platforms. This person must know how to find, mine and analyze data, and ideally, will have the reporting and writing skills to help readers understand the data behind the news. The ideal candidate will have familiarity with existing public data sets and government data sources.	pjohnson@reviewjournal.com	10/26/2016
Editorial Designer	The Marshall Project	New York		The ideal candidate is a well-rounded journalist and designer with 3-5 years' experience in media and/or publishing. The Editorial Designer should have a wide range of visual interests and experience and the ability to juggle a variety of tasks in a small but ambitious newsroom.	The Marshall Project seeks a full-time Editorial Designer to join us in our office in New York and oversee the award-winning visuals of the publication, including digital design, photography and illustration.	Apply here	10/25/2016
Data Visuals Developer	Texas Tribune	Austin, TX		You'll be asked to gather your own data, decide how to best visualize it, and build out user-friendly, informative stories on a regular basis. We write code in Node.js/JavaScript and Python, and use tools and platforms like D3, Webpack and Django in our newsroom. Familiarity with these languages and tools (and/or the desire to learn them) is crucial.	Our editorial Data Visuals team works in tandem with the Tribune's newsroom and engineering group to tell politics and policy stories, build and maintain data projects, and help our readers better access and understand their state government. We create large-scale data explorers on everything from public schools to public salaries, craft custom story templates for hard-hitting investigative pieces, develop clever and informative data visualizations, and keep an ever-watchful eye on the conflicts and interests of elected and appointed officials.	jobs@texastribune.org	9/29/2016
Data Journalism	Golding Properties	ohio	$22/Hour	. Exceptionally strong writing and storytelling skills in English 2. Experience working with survey data 3. Interest in current affairs	As a Data Journalism Intern at Dalia, you will 1. Propose your own ideas for the projects and execute/manage them independently 2. Research, write and edit content related to public policy, social affairs, economics and business	samuelcolton10777@gmail.com	9/26/2016
Data scientist	BuildZoom	San		Python SQL D3 Machine learning Digital mapmaking	If you are a data scientist with machine learning skills, e.g. you know your random forests	issi@buildzoom.com	8/25/2016

- Strong quantitative skills and proficiency with Microsoft Excel

Seeking Legislative Analyst

Bloomberg, DC, Washington

Apply Now   ⊘ Save   ⊏ Share   🖶 Print

Date Posted	November 03, 2016
Industry	Government
Specialty	Politics
Job Status	Full-time
Salary	Not Specified
Website	https://careers.bloomberg.com/job/detail/53264

Bloomberg

Description:

When government is at its best, diverse ideas are at work. The same is true of Bloomberg Government. Our ability to innovate and serve our clients means diversity and inclusion are essential. If you're eager to join a multi-dimensional team that celebrates and leverages difference, if you have the vision to see how information can transform "business as usual" and if you are hungry to create-to build a service transforming how things get done-then stop reading this and join our team.

Our Team:
Delivering news, analytics and data-driven decision tools, Bloomberg Government's digital workspace gives an intelligent edge to government affairs and contracting professionals influencing government action.

Individuals who embrace hard work, act with urgency and collaborate without reserve thrive at Bloomberg Government. You are on the front lines of revolutionizing the information services industry in Washington-as high as our expectations will be of you, yours must be higher. We are a start-up, backed and nurtured by the powerful incubation of the worldwide leader in information services, Bloomberg L.P. If you are dogged, innovative and interested in work that contributes to causes greater than yourself, you belong at Bloomberg Government.
We are looking for a candidate who can write legislative analysis and who will work with the director of legislative analysis and the team to develop content about public policy for Bloomberg Government. Analysts also work with the Congress Tracker team and other staff to cover legislation moving through Congress. The successful candidate will produce rich, in-depth analysis on the implications of government action, with a focus on congressional activity.

Global Opportunities

Data Journalism Jobs by Location

[from silk.co](#)
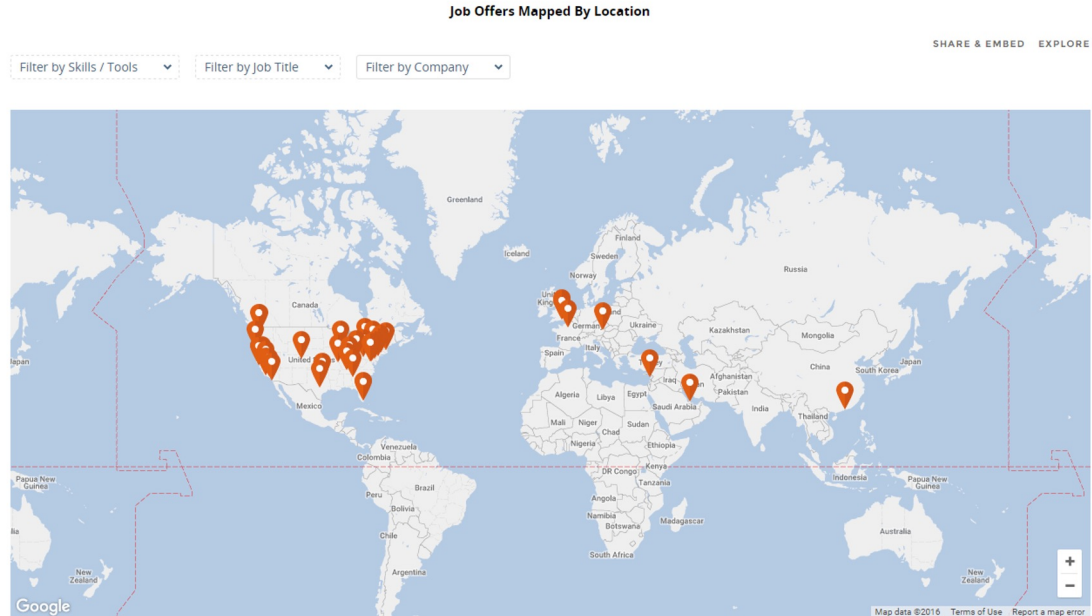
B

Broadcast Journalist (Data Unit) - BBC News

Business Data Reporter - Minneapolis/St. Paul Business Journal

G

General Assignment Data Reporter - Sun Sentinel

Graphic Artist - NJ Advance Media

Graphic Artist, Data Visualisation - The Guardian

Graphic Design Fellow - Huffington Post

Graphic Designer / Visual Journalist - Reddit

Graphic Journalist - San Diego Tribune

Graphic Journalist - Vox

Graphics & Interactive Designer - Visulizing Impact

Graphics journalist - The New York Times

Graphics/Multimedia Editor - Quanta Magazine

C

California Health Data Fellowship - USC Annenberg

Cartographer, Senior Production - National Geographic

Computational Journalist - The Dallas Morning News

Computer-Assisted-Reporting Producer - E.W. Scripps Washington Bureau

Content Editor (Graphics) - The Globe and Mail

I

Infographics Editor - Independent Digital

Information Architect - Visualizing Impact

Interaction Designer, Data Visualization/News - New York Public Radio

Interactive Content Editor - BuzzFeed

Interactive data journalist - Raycom Media

Interactive Designer - Baltimore Sun

D

Data and Interactives editor - Huffington Post

Data and Visual Director - Wisconsin Center for Investigative Journalism

Data Desk Internship - Los Angeles Times

Data Driven / Visual Editor - Quartz

Data Driven / Visual Reporter - Quartz

Data Driven / Visual Reporter - Quartz (2)

Data Editor - MapLight

Data Editor - The Chicago Reporter

Data Editor - The Guardian

Data Editor Reporter - Reuters

Data Editor/Computer-Assisted Reporting Specialist - The Connecticut Mirror

Data Fellow - ProPublica

Data Intern - ProPublica

Data Journalism Internship - PBS NewsHour Weekend

E

Editorial Experience Designer - ProPublica

Editorial Graphics Producer - ESPN

Editorial Trainee - Reuters

F

Front End Developer, Data Visualizations

H

Head of Data Journalism - The Telegraph

K

Knight Chair in Data and Explanatory Journalism - University of Syracuse

Knight-Mozilla Fellow - Knight-Mozilla OpenNews

L

Lead developer - California Civic Data Coalition

Lead Interaction Designer - The New York Times

Lead Interactive Designer/Developer - Huffington