

How BIG DATA Equals Big Changes in Cases and the Bottom Line

Big data and smart data are terms that have woven themselves into our conversations, our courtrooms, our clients' industries, and our firms. The term "big data" refers both to datasets that are too large for analysis using traditional techniques and to the new and emerging cleaning and analytic tools we apply to such data. In fact, data is now so "big" that until 2003, the entire world had only generated 5 billion gigabytes of data. This same amount is now generated every 10 minutes, and 90 percent of the world's data has accumulated in the last five years.¹ Our collective data conversation is soon to change from one of gigabytes to zettabytes.

Numerous fields readily employ big data in their decision-making. President Barack Obama's 2012 election campaign famously used big data analytics and statistical experimentation to assemble a coalition of voters. Google used big data to track flu and Dengue fever epidemic trends in recent outbreaks, and a new initiative called Africa Open Data intends to use data to tackle health challenges like Ebola. A full spectrum of businesses—from big-box online retailers to bakeries—analyzes big customer data to target their marketing. It is more and more common for broad and deep information to be used as a guide to decision-making.

What does big data mean for the legal field? What are some best-practice tools for managing large amounts of information in a case, and when should attorneys rely on these tools? How can attorneys and firms best get to information that can guide their litigation strategies? How can understanding law firm data increase profitability, and what are the best tools to analyze law firm metrics?

The short answer is that law firms can benefit from data-driven decision-making, and attorneys can increase efficiency and profits by incorporating data into their legal practices.

This article addresses these questions and provides an overview of the intersection between data and the legal field.

MELISSA KOVACS has a Ph.D. in public policy and econometrics and is the principal and founder of FirstEval LLC, a data analytics and statistical consulting firm. Her clients include nonprofit organizations in the arts and social service sectors, governments, foundations, and attorneys. She can be reached at mkovacs@firsteval.com or via her website: <http://firsteval.com/>

eDiscovery and Multi-Matter Management

Many litigators use managed review services and electronic discovery (“eDiscovery”) technology tools to help them sift through big data in a case. And according to IDC, a leading provider of IT market intelligence and advice, there may be opportunities for the purchaser in the marketplace.

Sean Pike, IDC’s eDiscovery and Information Governance Program Director, says the market for eDiscovery vendors is rather saturated in the United States, and prices are decreasing as technology improves the document review process. He says, “For the five years prior to 2013, eDiscovery vendors saw a growth of 20 percent year over year. Now we are seeing growth in the single digits as the U.S. market has reached saturation. Managed review has become cheaper, as the supply of talent has grown and vendor competition has increased.” As more and more firms have entered the eDiscovery marketplace, competition among them for technological advancement has resulted in lower costs to litigators and shortened review time.

What’s more, some eDiscovery vendors are developing new tools to tackle multi-matter management (sometimes called cross-matter management), allowing in-house counsel to rely on previous work product for multiple corporate matters. For example, if a large corporation experiences serial litigation year after year, tracking these cases’ commonalities, key factors, and outcomes can eventually build a reliable dataset. Insights and useful nuggets can be gleaned from this dataset to inform future similar matters. In turn, relying on this data can reduce review time and help predict case outcomes and strategy for future similar matters.

In this sense, instead of conducting eDiscovery techniques on each case separately and comprehensively, multi-matter management tools can review a series of cases, resulting in reduced costs and review time. Getting to reliable insights will require a threshold of similar cases to comprise such a multi-matter dataset. Therefore, multi-matter management practices may not be useful for the small to medium-size practitioner, but it could be very informative for corporations and firms that see the same types of cases

over and over again.

Litigation strategy is hungry for data. Litigators should listen to the “lessons learned” stories their data can tell them to guide strategy in future similar matters.

External Data Sources

It’s not just the data from past matters and internal to firms that can be helpful in guiding strategy. Invaluable external data sources are becoming easier to use and more accessible to attorneys, especially in discrimination and damages cases. These include the incorporation of Census data in housing discrimination cases and external macroeconomic indicators to make demographic and economic arguments.

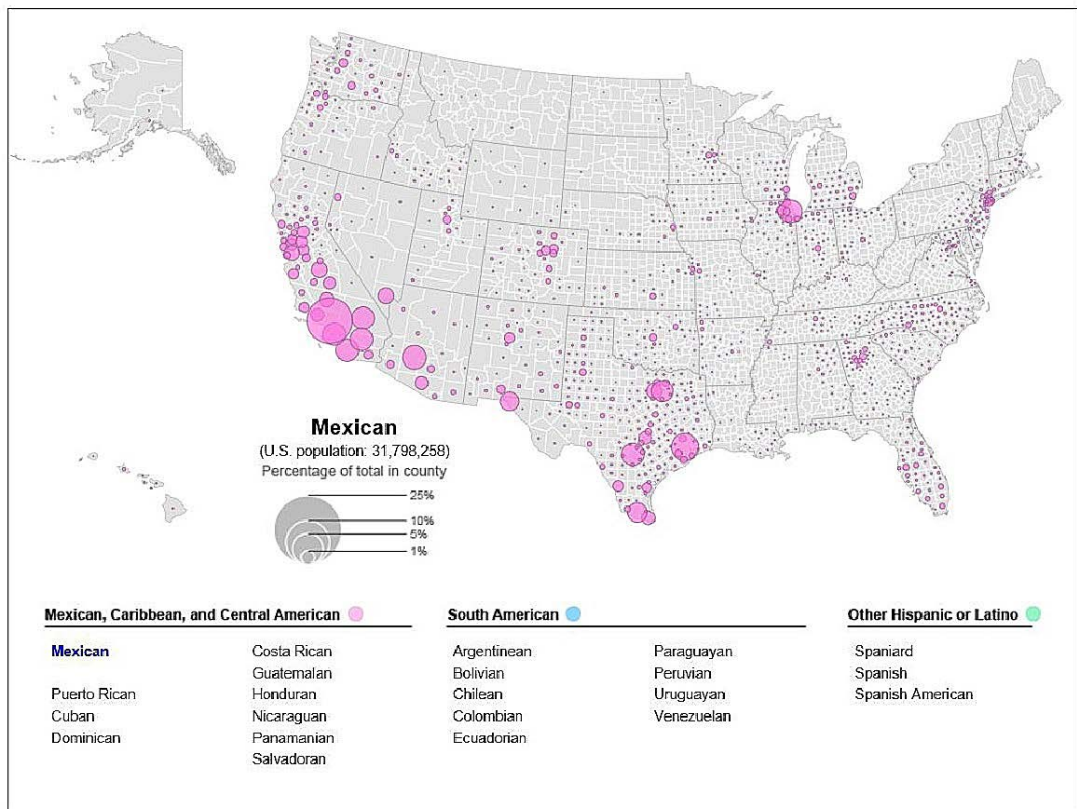
For example, the Census Bureau’s “Quick Facts” site is easier than ever to compare demographic factors like racial-ethnic population proportions, housing unit information, income, and business metrics among different geographic areas.² Kindly, the Census demographers have placed warnings on the site when you are attempting to compare two incomparable data pulls. The Census site also has a library of useful demographic data visualizations (examples on pp. 34 and 36), or you can create your own visualization that may be useful in the courtroom.

Numerous sources of macroeconomic indicator data also exist.³ Tracked over time, indicators such as GDP, the Consumer Price Index, and measures of employment can tell an economic story to supplement economic damages calculations. Easily accessible sources are available to even report the costs of environmental indicators like noise and air pollution, costs of crime and family changes, and costs of commuter time. As demographic and economic open data sources become easier to use, their incorporation into legal arguments will become easier to employ.

Distribution of Hispanic or Latino Population by Specific Origin: 2010

September 30, 2013

DATA VISUALIZATION EXAMPLE



The Business of Lawyering

Law firms are no different than other for-profit businesses in that their own financial data

BIG DATA RESOURCES FOR THE BUSINESS OF LAWYERING

and other metrics can guide profitability decision-making. Firms in all fields use business intelligence (“BI”) tools to varying degrees to make sense of real-time structured data to inform quick decision-making, better client understanding, find billing and invoice patterns, and analyze their human capital.

Specifically, law firms and corporate legal departments can use their own data to predict costs, negotiate prices, and even gauge attorney success. Data-driven outcomes can reduce firm costs, especially in the area of alternatives to traditional hourly fee structures.⁴ For example, one firm implemented BI tools to track partner and other billable time to determine how price and human capital affect client billing.⁵

But how do firms go about implementing a BI tool, and what does that mean?

BI tools provide data visualization and often interactive data dashboards, so that data can be viewed in real time and in picture form. While data scientists provide sophisticated services, dashboards allow non-data-savvy professionals to become data analysts, as anyone can pull up a dashboard

and shape a visualization toward their niche. BI tools sit within a broad spectrum, from open-source statistical computing platforms (such as R) to glossy interactive visualization dashboarding tools (such as Tableau and Spotfire) to more traditional statistical

analytic programs (such as SPSS and SAS) to well-known tools from well-known developers (such as Oracle, Microsoft, and LexisNexis). IT research and advisory firm Gartner provides a comprehensive guide to the BI tools market in its Magic Quadrant

R: An open-source programming language and environment for statistical computing and data mining, great at graphics and visualization. <https://www.r-project.org/>

Tableau: An interactive data visualization dashboarding tool, making it easy to see your data. <http://www.tableau.com/>

Similar to Tableau, **Spotfire** is data visualization and analytics software that makes it easy to see what’s in your data. <http://spotfire.tibco.com/>

SAS is a statistical computing package capable of advanced business intelligence analytics. It typically requires a programmer, but also has a user interface. www.sas.com

A bit more user-friendly, but not as sophisticated as SAS, IBM’s **SPSS** is also a statistical computing package that can easily handle business intelligence analytics. <http://www-01.ibm.com/software/analytics/spss/>

Oracle BI is a very popular suite of enterprise-level business intelligence analytics tools. <http://www.oracle.com/us/solutions/business-analytics/business-intelligence/overview/index.html>

Microsoft has **Power BI**, a dashboard creation tool to help with visualizing what is in your data. <https://powerbi.microsoft.com/>

LexisNexis has **InterAction Business Edge** that can compress multiple data sources to aid in law firm business development. <http://www.lexisnexis.com/intelligence/business-edge.page>

series, which can be useful for those wanting to use

BI to increase their bottom line.⁶

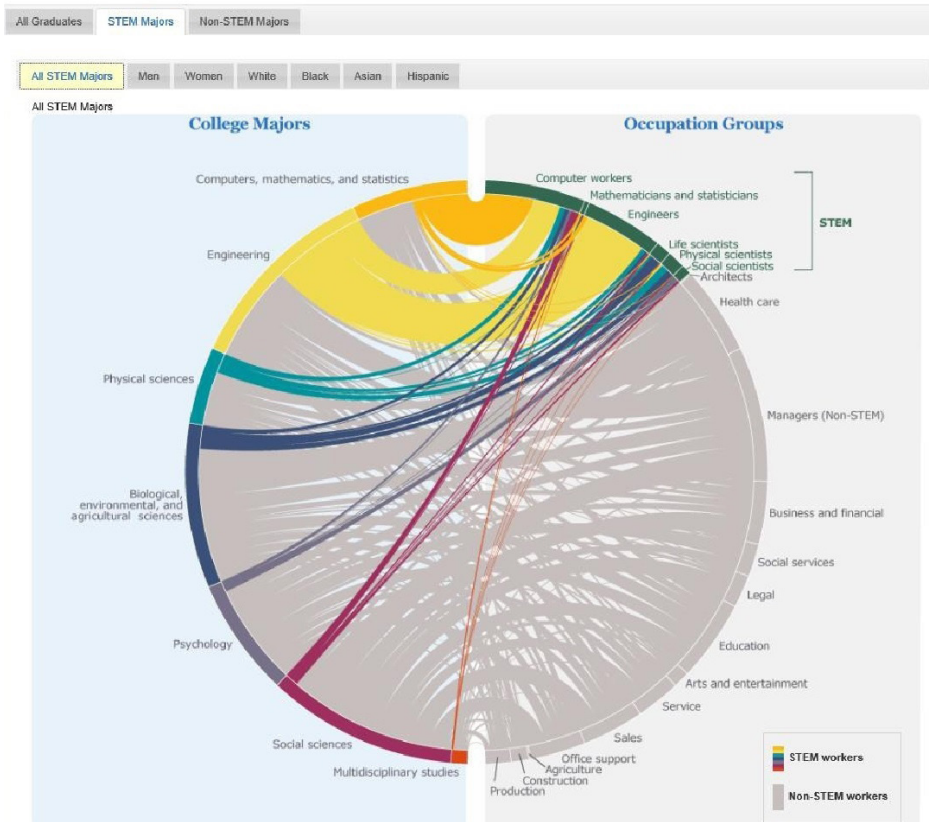
(For more detail on and links to affiliated businesses, see the sidebar on page 35.)

Technical tools, applications, and software that attorneys can use to mine data, organize evidence, and increase their profitability can be dizzying. Using the latest and greatest tool may not be what's best in every situation. Firms and corporate counsel departments should determine the highest return on investing in a BI tool dependent upon their size, needs, and access to data. Along with other factors, a BI tool should be chosen for its ease of use, value, and the service it will provide to a non-data-savvy attorney.

Big Data Price of Admission

To be blunt, there are few free tickets into the world of big data. While litigators are enjoying a decrease in technology-assisted review pricing, the cost of managing and analyzing discovery can still be a material expense in a case. When more in-depth analysis is needed, litigators may need to hire an expert, at expert hourly rates. However, numerous free external data sources exist for the ambitious attorney who wants to incorporate Census data or macroeconomic data into their arguments.

Beyond Excel (sitting at the least sophisticated end of the toolbox spectrum) and R (freeware programming language requiring expertise), BI tools lying in the middle of the sophistication spectrum will cost money. Such tools run a range of prices, and are often priced as software-as-a-service ("SAAS"), meaning that firm size will affect the number of software licenses required and the eventual price of implementing the




tool. There's also the cost of an analyst or consultant to help with BI data analytics, unless this knowledge exists in-house.

Overall, the prudent firm will need to commit some time to the homework of pricing analytic tools that match their needs.

Conclusion

The legal industry sits atop one commentator's list of industries that need big data—and need to understand it.⁷ Law firms and attorneys are sitting on a gold mine of BI data and past case data that can guide decision-making, boost case outcome successes, increase profitability and inform pricing

structures, and more efficiently distribute their human capital. Simultaneously, external open-data sources are becoming more user-friendly to the non-data-savvy, especially to benefit the legal areas of discrimination and economic damages.

The term "big data" is already on its way out, another reflection of the speed of technology.⁸ We will soon become accustomed to our data being "big" and turn our focus toward predictive and prescriptive analytics, citizen data science, and analytics marketplaces—all areas to which the legal field will want to monitor, participate, and contribute. 

endnotes

1. R. Bhargav, *What is Big Data, and Why Should You Care?* Aug. 25, 2015, available at www.datasciencecentral.com/m/blogpost?id=6448529%3ABlogPost%3A316182
2. See <http://quickfacts.census.gov>.
3. See, e.g., <http://genuineprogress.net>; www.census.gov/economic-indicators/; and the *Wall Street Journal's* Economic Indicators Archive.
4. See Ben Kerschberg, *Business Intelligence and Legal Matter Management*, FORBES, April 26, 2011, available at www.forbes.com/sites/benkerschberg/2011/04/26/business-intelligence-and-legal-matter-management/
5. Diann Daniel, *How a Global Law Firm Used Business Intelligence to Fix*

6. *Customer Billing Woes*, Jan. 8, 2008, available at www.cio.com/article/2437361/business-intelligence/how-a-global-law-firm-used-business-intelligence-to-fix-customer-billing-woes.html.
7. Rita L. Sallam *et al.*, *Magic Quadrant for Business Intelligence and Analytics Platforms*, Feb. 23, 2015, available at <https://www.gartner.com/doc/2989518/magic-quadrant-business-intelligence-analytics>.
8. Larry Alton, *5 Industries that Need Big Data*, July 24, 2014, available at <http://www.datasciencecentral.com/profiles/blogs/5-industries-that-need-big-data>.
9. William Vorhies, *Big Data Falls off the Hype Cycle*, Aug. 17, 2015, available at <http://www.datasciencecentral.com/profiles/blogs/big-data-falls-off-the-hype-cycle>.