2021

# Litigation Analytics: A Framework For Understanding, Using & Teaching

Peter A. Hook

# A FRAMEWORK FOR UNDERSTANDING, USING & TEACHING LITIGATION ANALYTICS

Essential concepts for students, lawyers, and legal information professionals.

BY PETER A. HOOK

Litigation analytics are quantitative, replicable observations about litigation from which you may derive actionable insights. Attorneys use them to make data-driven predictions about litigation outcomes and to inform trial strategy. They provide evidence for client pitches, assist in-house corporate counsel when selecting outside counsel, and are used by mediators and arbitrators. Additionally, they provide helpful information to law students researching judges, attorneys, and law firms with whom they might be interviewing.

## A Brief History

Litigation analytics originated, in part, with the Stanford Intellectual Property Litigation Clearinghouse (IPLC) in 2008. The IPLC became a commercial product separate from Stanford University in 2009 and was renamed Lex Machina. It was subsequently purchased by LexisNexis in 2015. Bloomberg Law launched its litigation analytics platform in 2016. However, it had also launched a transactional analytics tool a year earlier. Litigation analytics on Westlaw were introduced as part of the Westlaw Edge search platform that was released in 2018. Analytics on Context, available for academic subscribers on the regular LexisNexis platform, also became available in 2018. While there are other litigation analytics products, some with different state coverage and topical focus, Westlaw, Bloomberg Law, Context, and sometimes Lex Machina, are the most widely available litigation analytics platforms to law school subscribers. Additionally, analytics about law firms and companies existed as early as 2009 on competitive intelligence tools such as Thomson Reuters's Monitor Suite and LexisNexis's atVantage.

## Types of Analytics

To get the most out of litigation analytics, you must know the breadth of analytics available and the types of questions they can answer. The general groupings of litigation analytics are discussed below.

### DOCKET ANALYTICS

Of the many different types of litigation analytics, almost all originate in the content found within case dockets. Docket analytics include case type (subject), the court in which the case was filed, the parties to the litigation, the law firms and attorneys representing the parties, the role of the attorney (representing the plaintiff or defendant), the assigned judge, and the dates of specific events in the case that allow for descriptive statistics such as the length of time to case completion. Much of the data about case type originates from Nature of Suit (NOS) codes required to be used in federal civil cases. However, as the topical granularity of NOS codes is inadequate for many purposes and only one NOS code is chosen per case, both LexisNexis and Westlaw supplement their case subject assignments with proprietary legal taxonomies employing artificial intelligence (AI) to make the subject assignments.

Basic docket analytics allow you to answer questions such as: How long do these types of cases usually take to complete? At what stage does a company, attorney, or law firm usually settle? How much experience does a particular firm or attorney have with a particular subject matter, opposing party, opposing law firm, court, or judge? Does a firm or attorney represent more plaintiffs or defendants?

### MOTION ANALYTICS

Motion analytics cover over 20 different motion types, their length of time to be decided, and outcome by judge (ruling for the plaintiff or defendant). They allow litigants to assess the likelihood of a particular judge granting a motion. For example, how often does a judge grant a motion to dismiss relative to the other judges on the same court or across different courts?

> "Litigation analytics allow novice attorneys to have the insights of their more experienced colleagues and allow these more experienced attorneys to have a global overview of a field that would not be possible otherwise."

### EXPERT WITNESS ANALYTICS

Expert witness analytics quantify instances a person has qualified as an expert and when they have been successfully challenged. They give the frequency with which an expert testifies in cases per year, the subjects of those cases, and their awarded damage amounts. Such data help attorneys identify experts for specific types of cases and provide examples and strategies as to how to successfully challenge them.

### OUTCOME ANALYTICS

Outcome analytics are of paramount concern to litigation strategy—what percentage of time do plaintiffs win in this particular jurisdiction as to this subject matter? As stated earlier, outcome analytics also apply in the context of motion practice. There are six specific case outcomes included in Westlaw's litigation analytics: uncontested dismissal, settlement, dispositive motions, verdict, the case is docketed elsewhere, and an "other" category.

### DAMAGE ANALYTICS

Similar to what used to only be available in specialized jury verdict and settlement reporters, litigation analytics platforms provide damage awards and settlement amounts. Furthermore, now that the data are on interactive platforms, users can customize ("slice and dice") their results. On Lex Machina, you can obtain damage amounts limited to the following: employment cases; cases involving employment discrimination; federal district courts in New York; cases in which the damages are won at trial; and the specific type of damage claim of emotional distress. (Each of these is a selection point.) The customizable results allow you to add a column for damage amounts and then to sort by this column from highest to lowest amounts. Damage and settlement amounts inform core litigation strategy questions such as "How big of a damage award might my client recover if we are successful at trial?"

### JUDICIAL ANALYTICS

Judicial analytics include appeal analytics—the number of times that a judge has been affirmed, reversed, and affirmed in part/reversed in part. Other answerable questions include the following: How often has a particular judge granted motions for class certification? What types of cases do they hear the most? What types of cases do they adjudicate the fastest and slowest? What case does a particular judge cite most often as to a particular issue?

## Pivot Points

On interactive data platforms, a user can change perspectives from which all the data is organized. These are called pivot points. In other words, you can pivot on general categories in which the remaining data is subsequently rearranged. The major pivot points available on the litigation analytics platforms include represented entity, law firms, attorneys, court, judge, case types, experts, and patents. Not all platforms allow you to pivot by all of these types.

## Insight-Needs Categories

As the goal of litigation analytics is actionable insights, it is worth exploring the finite, exhaustive list

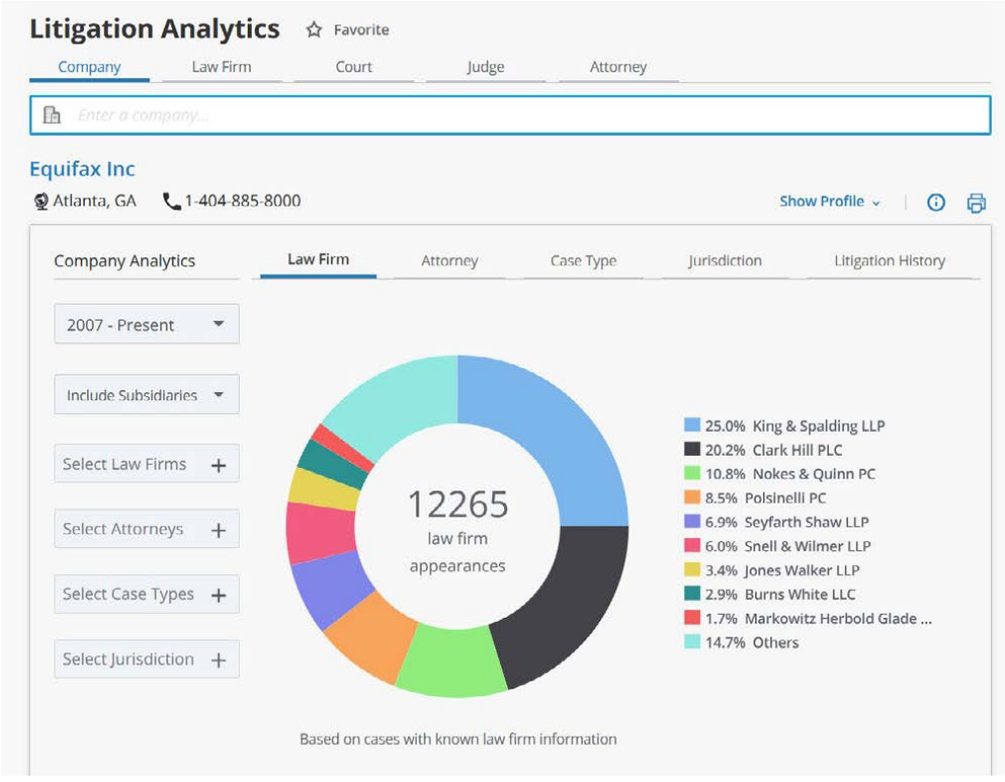


**Figure 1: Screenshot of Bloomberg Law's Represented Entity Analytics for Equifax Inc. (law firm appearances) taken on Oct. 1, 2021.**

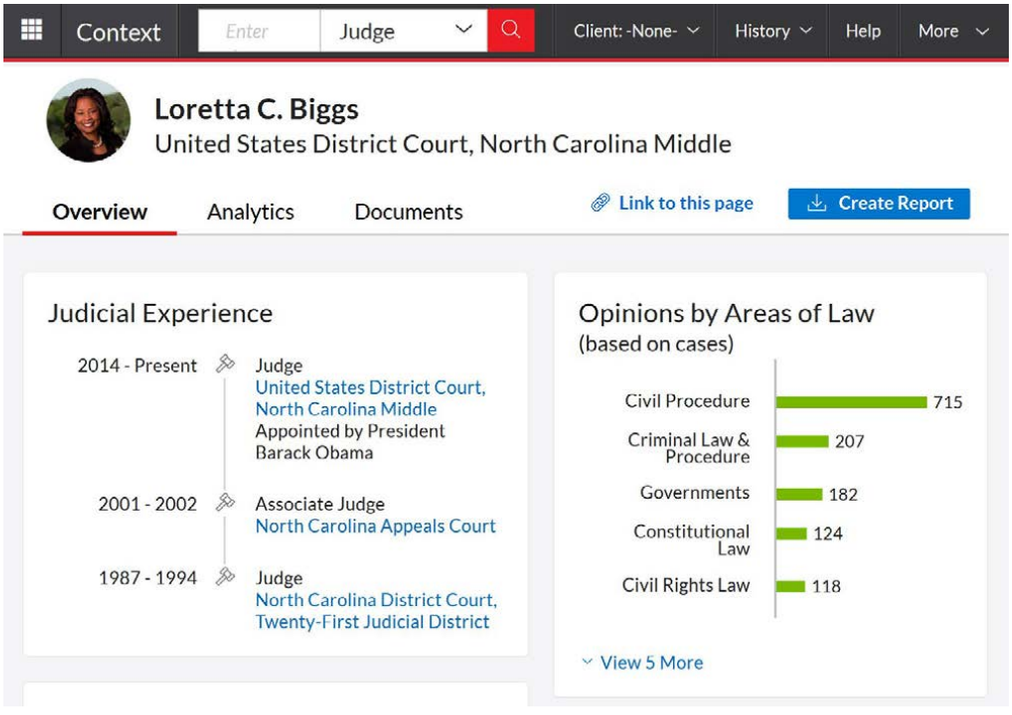


**Figure 2: Screenshot of LexisNexis's Context's Judicial Analytics Overview page, for the Honorable Loretta C. Biggs, United States District Court, Middle District of North Carolina, taken on Oct. 1, 2021.**

of insight-needs categories that have been identified by scholars: (1) categorizing and clustering; (2) ordering, ranking, and sorting; (3) distribution; (4) comparison; (5) trends; (6) geospatial location; (7) composition; and (8) relationships. (See Katy Börner's, *Atlas of Knowledge: Anyone Can Map* at bit.ly/ND21Atlas.) The section below contextualizes these insight-needs with specific examples from litigation analytics. Such lists remind you of all possible insights to be gained and help ensure that no important questions go unasked.

### 1. CATEGORIZING AND CLUSTERING

Categorization is at the heart of litigation analytics. Similar items are placed in similar categories ("buckets") so you gain insight from their frequency. One example of categorization is using NOS codes and other subject headings to put like cases together by subject. Additional categorization occurs through groupings by firm, attorney, company, judge, court, representation type (plaintiff/defendant), motion type, etc. Categorizing and clustering is a means of achieving insight through simplification.

### 2. ORDERING, RANKING, AND SORTING

Almost all litigation analytics, when displayed, are ordered, ranked, and sorted. This allows users to quickly ascertain the superlatives—most, fewest, largest, and smallest. This provides answers to a variety of questions, such as the following: What law firm represents Home Depot the most? What attorney in Indianapolis represents the most plaintiffs in employment discrimination disputes? What judge is overturned the most on appeal?

### 3. DISTRIBUTION

This category includes statistical distributions including the range of values in the dataset (highs and lows), anomalies, and gaps (missing data). It includes descriptive statistical measures such as central

tendency—mean, median, and mode. Analytics are most helpful when they contextualize a specific entity's numbers with what is normal, or average, for the context. For instance, if it takes one federal judge two years on average to complete employment discrimination cases, and the average for the rest of the judges in a particular district is one-and-a-half years, that difference might be meaningful and actionable.

#### 4. COMPARISON

As alluded to above, comparisons are one of the main ways to gain actionable insights from analytics. Seeing how a particular firm, attorney, or judge's numbers stand out against averages for the whole allows you to know how likely a particular firm or attorney is to settle, how likely a particular judge is to grant a motion for summary judgment, or that a particular case type usually takes much longer than average. Lex Machina has good comparison apps that are designed to make it easier for users to compare data for selected entities.

#### 5. TRENDS

Trends are usually time based. However, they can be spatially based as well, such as some phenomena increasing steadily from east to west. Trend data also include bursts—items with a spike of increased activity against their normal background frequencies. On litigation analytics platforms, a common trend feature is the number of litigated cases over time. In the case of the corporation Equifax Inc., a graph of the number of cases per year over time reveals a spike of increased litigation after its well-publicized data breach in 2017.

#### 6. GEOSPATIAL LOCATION

This insight-needs category relates to where things happen in a defined space. This is most often utilized on litigation analytics platforms to identify where in the United States firms or attorneys represent litigants in courts. Similarly, geospatial location data quickly reveal the places particular parties are most often sued.

#### 7. COMPOSITION

Composition as applied to litigation analytics includes proportions of the whole. "Parts of the whole" graphics, such as pie charts, doughnut charts, stacked bar charts, tree maps, etc., may be the most common type of information graphic appearing on litigation analytics platforms.

#### 8. RELATIONSHIPS

Relationships are frequently illustrated as node-link diagrams. While common when representing citation networks such as Ravel View on LexisNexis, they are not common yet on litigation analytics platforms. Other relationship data possibilities include correlation and causation. These are also seldom explicitly expressed on litigation analytics platforms.

### When Predictive Becomes Restrictive

It is possible that the predictive nature of analytics might stifle the evolution and development of the law. You can envision a future in which a client enters a law office, presents a legal scenario complete with damages, and the lawyer quickly employs litigation analytics to recommend to the client how to proceed. If the analytics indicate that a litigant has only a 45 percent chance of winning a verdict, risk-averse parties might not continue with the lawsuit and instead elect to settle for a fraction of the damages. When the analytics are recalculated involving years of this feedback cycle, they might indicate that a litigant now has only a 10 percent chance of winning a verdict if meritorious claims were disproportionately excluded. Alternatively, litigants might have an 80 percent chance of winning if only strong cases were brought to trial in the interim. If the law in the specific area were to develop without the influence of analytics, future plaintiffs might have won a much different percentage of cases. The challenge for future analytics designers is to detect and prevent these distortions.

### Looking Ahead

Litigation analytics allow novice attorneys to have the insights of their more experienced colleagues and allow these more experienced attorneys to have a global overview of a field that would not be possible otherwise. Because of their utility, analytics should be taught in the 1L legal research and writing curriculum. Students should know about them to prepare for interviews. Students should know what case precedents a judge relies upon so they can incorporate or distinguish those cases in their briefs. Furthermore, students should know about the wealth of content within case dockets that can be used as exemplars for their own legal writing. To that end, I have used the content of this article as the basis of a CLE (continuing legal education) course and as part of an experiential, advanced legal research course entitled, Analytics for Lawyers: Leveraging Social Science Research for Effective Advocacy. The syllabus for this analytics course, which includes an analytics-related final project, may be found in the most recent Research, Instruction & Patron Services Special Interest Section *Teach in Kit* at bit.ly/ND21kit. ■

IM Information Management
R+A Research + Analytics
T+T Teaching + Training

**PETER A. HOOK**
**ASSOCIATE DIRECTOR**
Kresge Law Library
Notre Dame Law School
Notre Dame, Indiana
phook2@nd.edu