The Core of Correlation

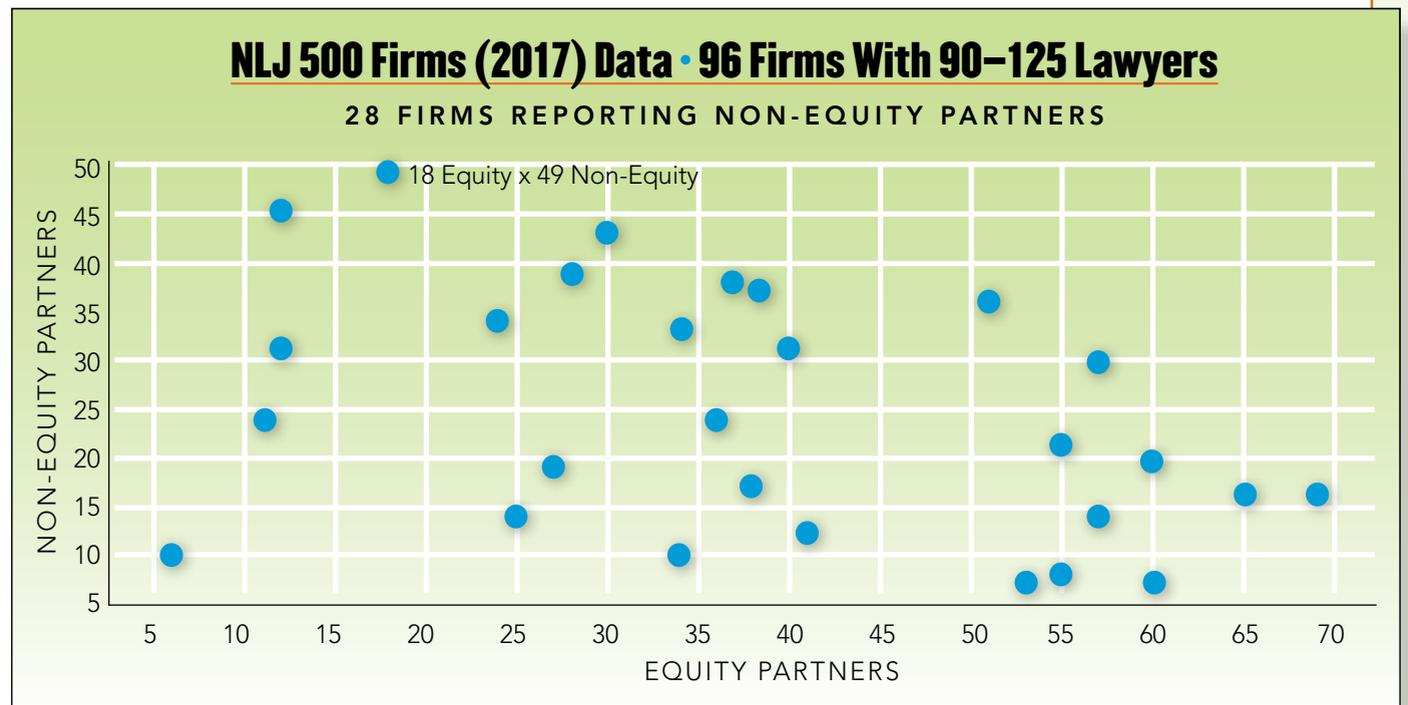## By Rees Morrison / Altman Weil, Inc.

Managers of law departments (and of law firms) often believe that there is an identifiable connection between one set of numbers and another. Perhaps they sense that the size of the plaintiff's law firm has some bearing on the cost of defending a lawsuit; they feel the number of patents applied for rises and falls with their company's R&D investment; or they've noticed that client satisfaction scores relate to keeping close to budget. Fortunately, those types of subjective impressions of managers can be tested and quantified.

Correlation is the statistical tool to quantify such possible connections – the degree to which two sets of numbers (called *variables*) are associated with each other – and it can also clarify the relationship between them.

Insightful correlations are plentiful for legal managers if they collect the data. Consider law firms working with a law department: A general counsel might suspect a connection between their overall effective billing rate and their number of lawyers. Or perhaps a lawyer's gut feeling is that the more years practicing law an in-house counsel has, in general, the less closely she or he manages their outside counsel spend. A billing partner might suspect that the longer an invoice goes unpaid, the lower the realization rate. After collecting some data, someone can run a correlation and test those suppositions.

To explain correlation with real data, let's hypothesize that the more equity partners a law firm has, the more non-equity partners it is likely to have. We can find out by using data from the NLJ 500 list (ALM). To have a more homogenous set, we focus on the 96 law firms that reported between 90 and 125 lawyers, where it turns out that only a third of them (28 firms) reported non-equity partners.

Before figuring out the correlation that will answer our hypothesis, sound statistical practice says to plot the data. Our eyes easily spot oddities

NLJ 500 Firms (2017) Data · 96 Firms With 90–125 Lawyers
28 FIRMS REPORTING NON-EQUITY PARTNERS
(Scatter plot: x-axis EQUITY PARTNERS from 5 to 70; y-axis NON-EQUITY PARTNERS from 5 to 50. Annotation: "18 Equity x 49 Non-Equity")

in the data and even patterns. We particularly want to avoid wasting time on correlations where the variables are not linearly related to each (such as if they include squares or cubes). Correlation is an appropriate numerical measure only for linear relationships, and is sensitive to *outliers* (values at the extremes).

In the plot, the horizontal, x axis tells how many equity partners each of the 28 firms reported, while the y axis tells the firm's number of non-equity partners. For example, the highest number of non-equity partners was 49 at a firm with 18 equity partners. (We could have added a best-fit *regression* line, but that's for another column.)

At a glance you can see that this scatter plot suggests no outliers and displays a *linear relationship* between the two variables. So we can proceed and calculate the *correlation coefficient*, which measures the strength and direction of a linear relationship between two quantitative variables. For the 28 law firms, the correlation between the number of equity and non-equity partners is -0.41, a moderately strong negative correlation. Our hypothesis was wrong: Actually, as law firms of this size have more equity

partners, they tend to have fewer non-equity partners.

Correlation coefficients always range between -1 and 1. Values near -1 indicate a strong negative linear relationship (as is our negative 0.41 value), values near 0 indicate a weak linear relationship, and values near 1 indicate a strong, positive linear relationship. A negative correlation means that one variable is less likely to be below the average of that variable when the other variable is below its average.

Let's extend our understanding of correlation with some additional observations.

1. Correlation disregards the units of measurement of the variables. Inches can correlate to dollars, or speed with age. The correlation coefficient can be computed and interpreted for any two numeric variables.

2. Because all correlations range between -1 and 1, you can compare correlations between different pairs of variables. So, we might find and say that there is a stronger correlation between the number of equity partners and the number of associates.

3. We can calculate correlations between ratios, such as between R&D spending as a percentage of revenue and patent lawyers as a percentage of all the lawyers in a law department.

4. If certain assumptions are met, we can interpret the *confidence interval* of the correlation coefficient (what range the coefficient would likely fall in if we redid the calculation on lots of samples from the data) and test the *null hypothesis*, that there really is no correlation between the two variables (and any correlation you observed is a consequence of random sampling).

5. If both variables are normally distributed (think bell curve), the proper method is the *Pearson coefficient* [the covariance (see below) of the two variables divided by the product of their standard deviations]. Another method, *Spearman's rank correlation*, does not assume that the underlying data is distributed in a statistically normal way.

6. Usually there is an intuitive explanation for a strong positive or negative correlation. With our scenario, firms that believe in full ownership and equity participation of their partners are probably less likely to create a tier of non-equity partners.

7. Correlation differs from regression. Correlation is not predictive. That is, correlation can only tell us how closely one set of numbers is associated with another set. By contrast, regression enables us to predict one variable based on another.

Simply to know the direction and strength of the relationship between two metrics gives legal managers useful insights. But even if no significant correlation exists between two metrics, that finding too can help managers. For example, what if it turns out that there is no statistically meaningful correlation between the average number of lawyers in a law department per office location and total legal spending? That realization might suggest that it doesn't much matter about the number of in-house lawyers at various locations. We should point out, however, that a lack of correlation does not necessarily imply a lack of relationship; the variables

Rees Morrison *is a Partner at Altman Weil. He focuses on law department consulting and data analytics for lawyers. He recently published a book on LeanPub, "Data Graphs for Legal Management." He can be reached at rwmorrison@altmanweil.com.*

If you capture and analyze data carefully, it will point you in the right direction.
— Matt Kivlin

items and then compares them against the billing guidelines. It determines which are most likely to require an adjustment, and then our billing analysts take over. These are legal professionals with expertise in billing analysis. The machine learning helps our experts prioritize where they should be reviewing potential issues or non-compliance in invoicing. From there, our experts will identify possible adjustments.

This creates an opportunity for greater communication with outside counsel. We offer the outside counsel the ability to address any non-compliance issues. It's fair. Then, if appropriate, they will readjust that invoice. It's just a really principled process with good communication that continuously improves the firms' billing performance. We're improving compliance and in many cases saving millions of dollars for our clients in legal spend – sometimes even millions each month.

As we head into the new year, what advice would you offer to general counsel or legal operations managers who want to improve their legal spend programs?

Kivlin: We'd suggest a few things. First, don't let your data go to waste. If you capture and analyze your data carefully, it's going to point you in the right direction and it will help you improve your overall results. Also, involve your law firms. They're a key part to your mutual success. Share your goals with them, most importantly, so they know what "good" looks like in your eyes. This will help them better meet your expectations and ensure that you mutually get to

that right balance between costs and the best outcome. Then, when it comes to using data and analytics to improve performance, focus. For example, if you want to improve budgeting, make that a focus and put a program around it. Or if you want to improve how you are managing your panels, use spend data to build a program for that. Then, whatever your goal is, communicate it well and measure your results often.

Linda Hovanec *is a Senior Director in product management, global business intelligence and analytics at Wolters Kluwer's ELM Solutions, with a focus on leading the company's LegalVIEW® product line. Hovanec has worked in the business of law for over 25 years. She has been at ELM Solutions since 2001, helping corporate legal and claims organizations drive better outcomes in partnership with outside counsel. She also has experience as a litigation paralegal and a law firm administrator. Her extensive background gives her a deep understanding of the realities and challenges of corporate law departments and claims organizations. She can be reached at linda.hovanec@wolterskluwer.com.*

Matt Kivlin *leads the Product Management team responsible for identifying, prioritizing and validating offerings in Wolters Kluwer's ELM Solutions' core markets. In this role, he oversees the product portfolios for corporate legal departments and insurance claims defense organizations. With a strong history of incubating and launching successful new products, he has a proven track record for product strategy, new solutions definition, cross-functional alignment, and exceeding customer expectations. He can be reached at Matt.Kivlin@wolterskluwer.com.*

Correlation

could have a non-linear (e.g., curved, step or bimodal) relationship. Using our data set, the correlation turns out not to be statistically significant, meaning that it could be a function of chance.

Always remember that correlation does not prove causation. Just because the degree of regulation of an industry correlates with higher legal spend per billion of revenue does not mean that regulation is the cause of those legal expenses (nor do bulging legal bills cause regulations!). Other factors may be at work such as political influence, case-law development, maturity of the industry, or specific legislation. Sometimes a third factor accounts for the close relationship between one metric and another, such as concentration in an industry. Statisticians refer to those behind-the-scenes factors as *lurking variables.*

One way to remove the effect of a possible lurking variable from the correlation of two other variables is to calculate a partial correlation. A *partial correlation* is the correlation that remains between two variables after removing the correlation that is due to their mutual association with one or more other variables. By looking at a graphic of the partial correlation coefficient matrix among the variables (like a heat map), it may be visually clear that the partial correlation between some of them is quite high.

Let's add one more math concept related to correlation. *Covariance* tells how much two sets of variables vary together. It's similar to *variance*, but where variance tells about a single variable's distribution, covariance tells you how two variables vary together. The larger the values of the variables, the larger the covariance. For our data set of 28 law firms, the covariance of equity partners and non-equity partners is -91.8. If we divide that covariance by the product of both variables' standard deviations we arrive at the correlation coefficient given above.

We emerge from this brief dive into mathematical methods. The key takeaway from this column is that correlation tools stand ready to help managers of lawyers to more crisply understand data variables that might vary together. Managers need not rely only on their own experiential sense of two sets of numbers changing in relation to each other. Correlation can put precision to intuitive insights and can clarify the direction of the variation and its credibility.

Law Department Operations

has already been previewed and make sure the rate didn't change arbitrarily in the middle of the month – that a $150 rate didn't suddenly become $475 without a conversation about it. That's the kind of thing that can be easy to miss when you're looking at a bill more holistically. When you look at a whole law department, those little misses add up to quite a bit of money.

Most of the spend management systems out there that are just focused on electronic billing will tell you that they're finding anywhere from 3 to 7 percent savings, so that's a pretty simple business case for law departments to put them into place. Say we spend X amount of dollars on a legal management system in the cloud, and it reduces our outside counsel expenditures from last year by 5 percent, that's probably somewhere between 7 and 14 times return on investment. These systems also allow the general counsel to analyze their spend and offer practical business cases for additional savings. For example, they could show the value of hiring additional practicing lawyers in-house, and thus reduce their outside counsel spend.

What do you think will be the biggest trend in law department operations for 2018?

Ruderman: The biggest trend I see for 2018 is an increased reliance on information and data analysis – and it doesn't matter whether it is people doing the analysis or machines. The real change will be a focus on the questions that matter and a move toward acknowledging the inherent biases implicit in doing things the way they've always been done. This will lead law departments to hire more professionals beyond practicing attorneys that can bring analytical and other business skills. We will also see increased adoption of cloud-based software, outsourced legal processing, alternative fees and dedicated legal sector procurement managers. As the law department acts more and more like the other business units in terms of efficiency and client satisfaction, it will also partner better with IT, sourcing and external vendors to gain a better picture of operations and a firmer grip on spend.