



# BIG DATA

## HR STILL GETTING IT WRONG BIG DATA & PREDICTIVE ANALYTICS THE RIGHT WAY

### OVERVIEW

**Research cited by Forbes estimates that more than half of companies sampled (over 60%) are investing in big data and predictive analytic tools for use in guiding Human Resources (HR) decisions.<sup>i</sup>**

With this surge in popularity coupled with pressures to keep ahead of the competition, authors and commentators describe the HR function as currently in a state of ‘transition’<sup>ii</sup> – moving from a concentration on meeting internal metrics (e.g., number hired, turnover number) to identifying the links between metrics (e.g., hiring the ‘right’ people to decrease turnover), and in this way optimizing HR processes and decisions.<sup>iii</sup> Problem is, there really isn’t time for transition; for companies to stay viable and navigate the competitive pressures they face, HR must do it now, and do it right.

By:  
Hannah Spell, Ph.D.,  
Director of Research & Analytics

Lauren Wegman, Ph.D.,  
Consultant



Given SMD's expertise in developing predictive analytics for HR, executives, and front-line leaders that drive actual business outcomes, the goals of this paper are threefold:

- Clarify what big data and predictive analytics should be in the field of HR.
- Identify appropriate methods of analysis and highlight common misconceptions.
- Demonstrate the business impact of predictive analytics with a case study.

## MAKING IT SIMPLE: BIG DATA AND PREDICTIVE ANALYTICS IN HR

### Big Data

Although often associated with complex analysis, big data is actually a simple concept – it is the collection and accumulation of numerous pieces of information that can be used to uncover connections between various concepts. For HR, these concepts may include employee behaviors, attitudes, skill or knowledge levels, performance metrics, turnover data, and much more. The ways the data accumulates can range from manual, such as the deployment of a selection procedure, to entirely automatic, such as machine scanning for resume selection. Yet, even manual systems are less manual today as they are inevitably aided through technological advancements (e.g., job knowledge assessment taken, scored, and stored electronically). Although the concept of big data is not new, the surge in the collection, administration, and accumulation of data have grown exponentially with advances in technology.

“ Just 10 percent of organizations in our study have taken the next step toward advanced analytics—helping business leaders solve their talent challenges through statistical analyses. A mere 4 percent are using predictive analytics to forecast future talent outcomes. ”

Bersin, O'Leonard, & Wang-Audia<sup>v</sup>

### Predictive Analytics

Despite the fact that it is possible to collect more and varied kinds of information, the collection itself is not that interesting or useful to organizations. For example, consider an organization that wants to decrease turnover in the coming year. Using attrition data gathered from HR, one can calculate how many individuals left the organization last year and the goal set for this year. But, without including other data inputs, not much else beyond examining baseline numbers can be done. Consequently, the real utility of big data comes when it is used in **predictive analytics** - the ability to show cause and effect links to real business outcomes. For example, if an organization is interested in identifying the key drivers behind employee turnover, predictive analytic methods can be used with data gathered from an employee survey coupled with HR data on employee turnover to determine which employee attitudes are most strongly linked to employee turnover (e.g., satisfaction, job fit, perceptions of managerial support). This allows organizational leaders to know which levers to pull to see the greatest impact on reducing employee turnover. In other words, predictive analytics, such as structural equation modeling, can identify what attitudes are “causes” of employee turnover.

## DON'T BE FOOLED: THE PREDICTIVE OR NOT TEST

It's logical that many organizations are embracing more data-driven approaches to HR due to the potential for impact, when done properly of course. However, few organizations are fully harnessing this potential of predictive analytics due to a number of misconceptions about predictive analysis methods. It is not uncommon for an organization to invest in a predictive analytic program that is actually not predictive. Below are three common analytic methods that are often misconstrued as being predictive in nature.

1. *Descriptive Analysis, Group Comparisons, Tracking, and Data Visualization* – A descriptive analysis typically consists of averaging across items or displaying counts or frequencies for a given topic. Trends can be visualized across time by charting averages or frequencies across time points to obtain a trend line. Additionally, group comparisons can be made to determine if an organization has significantly increased or decreased on a given topic between Time one and Time two. Despite group comparisons’ usefulness in helping to gain an understanding of an organization’s progress on the given topic across time, trends based on descriptives alone cannot be projected with accuracy into the future. Unfortunately, HR has gravitated toward Data Visualization tools which, although effective at creating pretty trending pictures, are not predictive analysis and do little to actually move the needle on the business outcome being tracked (or “visualized”). In short, descriptives are helpful in tracking progress, and comparisons are helpful in determining whether or not a change is statistically significant, but neither are predictive analyses. HR must ask itself, does this show any ROI to the C-suite?
2. *Correlation and Simple Regression* – Both are analytic methods used to identify the strength and direction of relationship between two things. Although a correlation is a measure of relationship, it is not necessarily predictive. Just because a relationship exists between two variables does not mean that one causes the other. It’s the classic example of ice cream sales and shark attacks being correlated. No one would argue that shark attacks *cause* ice cream sales to go up nor that ice cream sales *cause* more shark attacks—they both just happen to increase during the summertime. Similarly, having high engagement and high business outcomes doesn’t mean that engagement causes the other. Because these methodologies do not tell if a true relationship actually exists, they are not considered strong enough for prediction. If investments are going to be made based on analysis, then correlation and simple regression are not up to the task of driving and securing a significant financial investment – these are not predictive analytics.
3. *Multiple Regression* – Multiple regression is closer to modeling real-world relationships because multiple factors can be tested as predictors of one outcome. This allows for the examination of each factors’ unique effects on the outcome. While regression does identify variables that predict (i.e., are antecedents to an outcome), they are not necessarily causal. For example, the winner of the Super Bowl is a strong predictor of stock market performance (an NFC champion predicts a bull market) – however this is obviously not causal in nature. As a result, there are limitations to the usefulness and interpretation of simple regression analyses. Although a step in the right direction, this method is still not the strongest method to determine cause and effect relationships because measurement error cannot be modeled; only one dependent variable can be included, and, most importantly, causation cannot be determined. This method gets closer, but it is still not true predictive analytics.

## **PREDICTIVE ANALYTICS DONE RIGHT**

The best method for making predictions in the HR space is an advanced modeling method called Structural Equations Modeling (SEM). SEM has four large advantages over the other analytic methods:

1. Multiple inputs or “causes” can be tested along with multiple outcomes concurrently.
2. An accurate assessment of ROI can be calculated.
3. It provides the ability to correct for measurement error.
4. Causation can be determined.

Taken together, SEM is the best approach to predictive analytics for HR. The caveat to utilizing SEM in HR is that it requires specialized statistical software and a highly trained statistician to be correctly implemented. However, this should not be a deterrent for any HR practitioner hoping to

leverage this type of analytics. Universities often have professors or graduate students with the skills to conduct this type of analysis, or the analysis can be outsourced to a consulting firm with expertise in predictive analytics. And, consider the mantra, anything worth doing is worth doing right.

### Comparison of Analytic Methods for HR Prediction

| Method  | Trend Visulation | Assesses R'Ships | Multiple Predictors | Multiple Outcomes | Calculate ROI | Determine Causation |
|---|------------------|------------------|---------------------|-------------------|---------------|---------------------|
| Descriptive Analysis, Group Comparisons, Tracking, and Data Visualization | ✓                | ✗                | ✗                   | ✗                 | ✗             | ✗                   |
| Correlation and Simple Regression   | ✓                | ✓                | ✗                   | ✗                 | ✗             | ✗                   |
| Multiple Regression   | ✓                | ✓                | ✓                   | ✗                 | ✗             | ✗                   |
| <b>Structural Equations Modeling (SEM)</b>                                | ✓                | ✓                | ✓                   | ✓                 | ✓             | ✓                   |

### HOW HR MAKES AN IMPACT WITH PREDICTIVE ANALYTICS

Predictive analytics can be applied to a variety of HR initiatives. Below are a few examples of ways that predictive analytics can be applied to improve the impact and value of these types of initiatives.

- *Employee Surveys* – Employee surveys are a great avenue for utilizing predictive analytics. Used in this way, predictive analytics can make an often under-utilized HR initiative more impactful to the organization. Employee survey data can be linked to business outcome data in any industry such as turnover, work unit performance metrics, or financial data and used in a predictive analysis. SEM can be used to identify which topics or categories from the employee survey are key drivers of the critical business outcomes of your organization. This step enables leaders to make targeted action plans based on what they know to be important in getting results rather than focusing ONLY on what areas they scored low. A topic that was low on the survey should only be targeted if it is impacting results; otherwise, valuable time and resources should be spent addressing more important factors that have a clear and direct link to business outcomes. Chasing engagement scores or benchmarks does not constitute predictive analysis, nor does it drive actual business outcomes.
- *360 Assessments* – 360 assessments often require a large allotment of organizational resources and time. As such, maximizing the utility and predictive impact of these assessments can bring great value and increased ROI to an organization. Predictive analytics can be applied to 360 results by linking competency/behavioral ratings with business outcome data and determining which competencies/behaviors are the strongest drivers of outcomes (typically performance data). This linkage allows the organization to make data-based decisions and to focus future training and development initiatives on those critical competencies/behaviors identified as impacting results.
- *Hiring Assessments* – Hiring assessments can be one of the best places to apply predictive analytics because of the ROI of increasing successful hiring decisions while also preventing poor hiring decisions. Making a bad hiring decision can be costly to organizations because poor hires are more likely to be below average performers and leave the organization in the short-term. The best way to prevent a poor hire is to apply predictive analytics to the selection process.

- First, utilize current employee data to determine which assessments, and which factors of those assessments, are the most strongly related to actual performance.
- Second, assess candidates on those factors found to be most critical in successful performance and base hiring decisions on candidate strength in those areas.
- Finally, apply predictive analytics to validate and adjust the selection program once it has been used to make hiring decisions (e.g., ensure that the factors used to select candidates are statistically significant predictors of new hire performance after a certain amount of time on the job). When taken together, predictive analytics can and should be present in the selection process at each step from beginning to end. Solid hiring tactics need to be used—not unstructured interviews—in order to assess the critical knowledge, skills, abilities, or competencies. Otherwise, interviewers are rolling the dice and gambling on the organization’s future.
- *Competency Models* – Competency models are a collection of knowledge, skills, and abilities that are required for a specific job. Predictive analytics can be used to identify the core competencies of a given job or those that stretch across the organization as a whole as well as to validate a model after it has been put into place. By applying SEM, HR leaders can determine which competencies are most important for critical business outcomes as well as determining if the full model is appropriate and predictive of employee and business outcomes. This will inform HR leaders if the current competencies are appropriate or if they should be adjusted, how to prioritize the competencies in the model, as well as to provide a demonstration of direct business impact. In applying predictive analytics in this way, it will solidify buy-in from senior leaders and frontline managers and will help drive training, hiring, and performance appraisal strategies because the competencies to focus, based on their importance to the bottom line, will be clear.

All of the examples outlined above can serve as a way for HR leaders to showcase their role as a strategic business partner to organizational leaders by linking HR initiatives to critical business outcomes and ultimately the bottom line.

## CASE STUDY

Proper application of predictive analytics can be extremely effective when done correctly by helping leaders make well-informed decisions backed by data – no more guessing. In one such example, a large, global software company needed to understand which of the knowledge, skills, abilities, and personality traits of their sales force were driving actual sales performance. The organizational leaders wanted to fully understand performance drivers so that they could make informed decisions on future hiring, training needs, and promotion or reorganization decisions.

The company collected data on several assessments of their sales force.

- First, a knowledge assessment was administered that provided information on the level of understanding on several key aspects of the product and sales tactics.
- Second, the sales force completed a personality inventory that assessed several personality factors thought to be important for sales performance.
- Third, the sales managers rated their sales force on a variety of behavioral competencies presumed to be important for sales performance.

The key goal was to identify which factors from all of these assessments were the key indicators of sales performance. The company provided several measures of sales performance data that SMD linked to the assessment data in order to conduct predictive analytics. This wealth of data enabled SMD to do the following:

1. Identify which factors from each assessment were key drivers of current sales rep performance by linking assessments to real sales outcomes such as new business revenue and quota attainment.

2. By identifying which factors were key drivers of performance SMD developed a profile for a successful sales force to make data-based staffing and talent management decisions.
3. By using current sales rep data, the company was able to use current sales force scores to identify sales representatives that needed focused training related to the key drivers.

**Step 1:**

Access the current Sales Team on knowledge, competencies, and personality



**Step 2:**

Determine which factors are the strongest predictors (Key Drivers) of sales performance



**Step 3:**

Use the analytics to drive targeted development, identify high performers for promotion and placement decisions

All of these goals were met by utilizing predictive analytics. By applying SEM to the assessment and sales performance data, SMD was able to not only identify which factors were statistically related to sales performance, but also determine which factors had the strongest relationships with sales performance. This linkage allowed SMD to prioritize and weight specific competencies, personality traits, and knowledge factors to create a predictive algorithm that could be used to make developmental, promotion, and placement decisions, as well as hiring decisions.

With this customized approach that used the company's own data and determined the specific factors that were driving their sales performance (rather than taking an off-the-shelf model of broad sales competencies), SMD was able to increase the company's ability to make a successful decision more than twice as often. Success was achieved because specific, tailored approaches are by nature more predictive in that they take into account the uniqueness of the organization.

From the competencies identified as key drivers in sales performance, the company developed a tailored training curriculum to target these specific skills. After implementation of the new training curriculum, the company evaluated the performance of those that completed the training and those that had not.

They were able to see an increase in sales revenue of \$60 million and an increase in quota attainment of \$100 million for the sales force that completed the training courses.

In sum, the use of this organization's rich data enabled SMD to utilize predictive analytic methods to provide a tailored approach to the assessment of the sales force, their hiring, promotion, and placement decisions, as well as what they focus on and how they structure their training initiatives. While the time to gather the needed data, conduct this analysis, and put recommendations into place will be greater than if using an off-the-shelf product that provides competencies for "sales" in general, the ROI of having something that is tailor-made for an organization will make it well worth the time and effort. At the end of the day, demonstrating ROI is critical. If ROI can't be shown and the bottom line isn't impacted, any time and effort spent will be wasted.

## CONCLUSION

To bring all organizations into 2015, HR must begin understanding and adopting true predictive analytics. Big data coupled with predictive analytic capabilities can provide a myriad of benefits to organizations—if done correctly. Technology has eased the speed and accumulation of organizational data, providing many valuable and unique internal and external sources for evaluation and prediction. There is great

power in tailored approaches specifically designed for a particular organization's goals and culture – allowing for the selection, assessment, and management of individuals who will succeed in that unique work environment. Whatever the organizational goal (e.g., reduce turnover, increase hiring success, increase ROI of your HR initiatives), predictive analytics will not only allow for the identification of key drivers, but will also allow for their prioritization, making planning and resource allocation straightforward and giving a company a massive competitive advantage – the quality of people.

#### References:

<sup>i</sup>Bersin, J. (Oct. 2013). Big data in Human Resources: A world of haves and have-nots. Forbes. Retrieved from <http://www.forbes.com/sites/joshbersin/2013/10/07/big-data-in-human-resources-a-world-of-haves-and-have-nots/>

<sup>ii</sup>Ferrar, J. (Dec. 2014). Predictive analytics: What big data means for the future of HR. Recruiting Daily. Retrieved from <http://recruitingdaily.com/big-data-analytics-hr/>

<sup>iii</sup>Feffer, M. (Oct. 2014). HR moves toward wider use of predictive analytics. Society for Human Resource Management (SHRM). Retrieved from <http://www.shrm.org/hrdisciplines/technology/articles/pages/more-hr-pros-using-predictive-analytics.aspx>

<sup>iv</sup>Bersin, J., O'Leonard, K., & Wang-Audia, W. (Sept. 2013). High-Impact Talent Analytics: Building a World-Class Measurement and Analytics Function. Bersin by Deloitte. Retrieved from <http://marketing.bersin.com/rs/bersin/images/hita100113sg.pdf>

---

**SMD is the leader in predictive analytics for employee assessments. SMD Link is its patented technology that links employee data to business outcomes and provides these key insights and actions in all manager reports. Learn more at [www.smdhr.com](http://www.smdhr.com) and contact them at [info@smdhr.com](mailto:info@smdhr.com).**

