

FIVE COSTLY MISTAKES COMPANIES MAKE WHEN WORKING WITH HUMAN CAPITAL ANALYTICS

Larissa Rabbiosi and Dana Minbaeva, CBS

Human capital analytics (HCA) enables organizations to anticipate business challenges and solve strategic problems. HCA can be used to identify the “why” questions that are strategic in nature — questions that focus on creating value – and to make well-informed decisions through the study of human capital data. However, the difficulties faced by managers wishing to answer the “why” questions have grown as the number of firms using HCA has exploded. The biggest challenge lies in the fact that organizations cannot fully exploit their data’s potential. Why do firms struggle? Our experience suggests that an understanding of five common – and potentially costly – mistakes can help firms get the most out of their HCA.

Mistake 1: Relying on data-driven definitions of the business problem

Too often, problem formulations are data driven. As firms recognize the importance of workforce data, they accumulate a mountain of it over time. Such data might, for example, cover employee demographics, employee-engagement surveys, customer satisfaction, and performance-related issues. Although exploiting available data can save time and money, the actual question in an analytical exercise should not *solely* be derived from the available data.

The exploration of large amounts of data in search of consistent patterns without any prior expectations of systematic relationships among the variables is called data mining. The goal of data mining is to produce a solution that can generate useful predictions without identifying the specific relationships among the variables. In other words, the exact nature of relationships among the variables is intentionally “black-boxed”.

This is why using the available data to find the relevant questions tends to only lead to incremental increases in value-added. As such, this practice diverts the organization’s attention from identifying solutions to the real business problems. In other words, data-driven questions tend to be less important for development of the firms’ long-term strategic goals.

Follow-up readings:

<https://www.linkedin.com/grp/post/5159783-6009461574106968067?trk=groups-post-b-title>

Mistake 2: Overlooking data organization

“I know that we have a lot of data, but I do not know what kind of data we have.” This is the most common response from managers when asked about their existing organizational data. Most firms do not know what types of data are already available or in what form. Moreover, they do not necessarily know where they can access that data. What data do we have? Where do we store the data? How has the data been collected? What rules have been applied? How can two different datasets be merged into one? What are the advantages and disadvantages of each dataset? Although these are basic questions, most firms do not have the answers. Such poor organization of firm data can be very costly. When formal, centralized coordination of data collection is lacking, we often see data duplication, wrong entries, etc. Moreover, such a situation makes it impossible to combine different datasets; creates unexplained breaks in time-series/longitudinal data; and leads to data inconsistencies due to the proliferation of various metrics, codings, or time frames. Accordingly, analyses based on such data are rarely comparable or combinable. Answers to complex business problems that rely on the analysis of different variables observed over several time periods and at different organizational levels (e.g., individuals, teams, departments, business units) are difficult to derive. Moreover, firms usually do not collect data documenting changes in the organization (e.g., business-unit reorganizations). However, as organizational change can modify the relationships under study, this failure to model such processes biases the analytics-based decision-making process.

Follow-up readings:

<https://www.bcgperspectives.com/content/articles/big-data-digital-economy-how-to-avoid-big-data-trap/>

Mistake 3: Always implying causality

We have heard this before: Correlation does not equal causation. For example, higher levels of employee engagement may drive improvements in firm performance. But the reverse might also be true. Although the majority of analytics projects allow only for identification of correlations, the results are often used to present a wide range of possible causal explanations. Indeed, we often hear users of HCA saying that “engagement drives performance” or “engagement is a predictor of performance” on the basis of a correlation pattern between engagement and performance. However, a failure to question causation claims can be very dangerous and bias the decision-making process – activities will be financed and implemented with the expectation of certain effects that, in all probability, will not be realized.

Other serious problems associated with assuming causation lie in the possible presence of an unobserved variable, which may determine both high levels of employee engagement and improvements in firm performance, and in the fact that highly engaged employees are more likely to choose to work for better (high-performing) firms. As employees are not randomly assigned to firms, teams, or departments, and because managerial practices influence both employees and performance, establishing causality is not easy. One possible solution is (again-again) to build theoretical models in advance; those that account for the alternative explanations. As the use of experiments and more advanced econometric techniques can also significantly alleviate endogeneity concerns, firms should also be encouraged to use these research settings.

Further explanation of endogeneity problem:

<https://www.youtube.com/watch?v=dLuTjoYmfXs>

Mistake 4: Forgetting to evaluate HCA implementation

Although firms are improving their abilities to act on the results of their HCA, too few collect data focused on the consequences of their analytics-based decisions and actions. What actions have been taken and where? How they have been operationalized? What changes are evident in the variables? The formal analysis of follow-up data reveals the effectiveness of the decisions and actions, helps identify how actions can be modified or changed to better achieve the expected output, and highlights those actions that are actually harmful and should therefore be stopped.

Moreover, reviews of the processes of data gathering, analysis, and action implementation serve as a source of learning. They make it easy to assess the entire analytics-based decision-making process, exploit opportunities, and avoid risks. It is important to consider the HCA process outside the specific business problem in order to leverage general lessons that can be applied to the analytics-based solution for the next business problem. All of these benefits will be missed if the follow-up stage is forgotten.

Mistake 5: Leaving data to IT systems alone

ERPs are useful for keeping data in order. However, the need to be aware of the data that is needed and ensure that the right data is collected requires a deeper understanding of how to model the “why” question. This requires extensive managerial participation, and collaboration between managers and analysts. Business-problem knowledge should be intertwined with analytical knowledge in order to develop a detailed model of the “why” question, which can then serve as the analytical framework for data collection. However, managers often lack analytical motivation or refrain from engaging in analytical discussions because they possess limited statistical and econometric skills. At the same time, analysts struggle to communicate with managers in terms of “telling the story” beyond p-values. Without such dialogue, the “why” question cannot be properly modeled. This implies that some key variables will not be identified, relevant data will not be collected, problems associated with the operationalization of the variables will not be anticipated, and potential pitfalls in empirical estimation will not be considered. In addition, analysts will be less efficient because of the costs associated with the collection of new data ex post, model adjustments, and delays in the analysis. Moreover, HCA could end up producing ineffective outputs that cannot support the decision-making process.

Best-practice example:

<https://www.linkedin.com/grp/post/5159783-5999598757845491714?trk=groups-post-b-title>

What do you think? Share your thoughts on LinkedIn – search for “*Human Capital Analytics Group*”