

A HR ANALYTICS MANAGEMENT CYCLE

Report for participants
HR Analytics Project
Jan - Dec 2018

Dr. María Jesús Belizón
Assistant Professor Int. HRM



UCD Michael Smurfit
Graduate Business School



BRINGING DATA SCIENCE TO HR

It seems paradoxical that the growing scholarship on HR analytics in its current state does not draw from neither data science nor computer systems foundational building blocks. HR analytics as a line of enquiry in the HR field is indeed proliferating, however, it is for the most part theoretically underdeveloped (Angrave et al. 2016; Marler and Boudreau, 2017; Huselid, 2018). This research project attempts to bridge data science and HR by tapping into current HR analytics research but also by delving into the cross-disciplinary field of data analytics. The present report proposes a HR Analytics Management Cycle based on the preliminary findings of the research project with the idea to help HR professionals integrate HR analytics within HR decision-making.

GOALS Current scholarship on HR analytics is, overall, enormously detached from two other sciences that have de facto made HR analytics possible, namely, data science and computer systems, making existing assessments of HR analytics antecedents, purposes, impact and consequences largely skewed. Drawing on existing knowledge along with some insights from semi-structured interviews, a HR analytics management cycle is proposed on which to build HR analytics projects, delving into both the type of capabilities needed in each stage of the process and their relational nature (Minvaeva, 2018). The research questions that this paper addresses are:

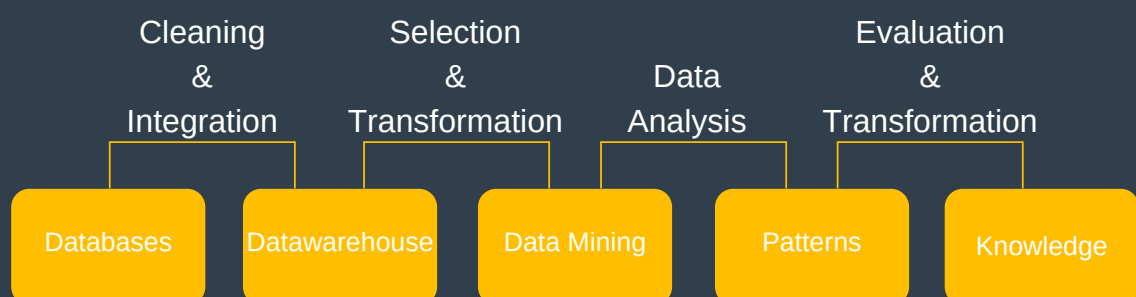
- (i) what are the stages of the HR Analytics cycle and capabilities needed therein?
- (ii) what are the challenges of managing such varied array of skillsets?.

For this purpose 6 organisations have been selected, five based in Ireland and one in the UK, all of them with relatively small HR Analytics teams (3 team members on average), using sophisticated analytics and operating across different industries.

TOWARDS A HR ANALYTICS MANAGEMENT CYCLE

As a first port of call, let us take Angrave and colleagues' definition of HR analytics, namely, "analytics [that] involves both traditional relational database and spreadsheet-based analysis, new forms of data software that allow very large quantities of data to be stored and organised more efficiently and new techniques for representing and understanding data through visualisation". In principle and in terms of the uses of a diverse array of databases, the definition seems to strike on sound ground, as is with the involvement of different pieces of software to store, manipulate and understand data. However, this definition falls short in different fronts. First, it is not able to capture all components of the data analytics cycle. If we turn to data science theoretical development, it was in the late 1990s that an accepted data mining cycle was proposed, the so-called 'Knowledge Discovery in Databases (KDD)' (Fayyad et al. 1995; Han et al. 2012), which is indeed largely echoed in Angrave et al. (2016)'s definition: it runs from the selection of data, pre-processing of data, transformation of data, identification of patterns and subsequent interpretation that eventually yields new knowledge (See Figure 1). And yet, what is a database?

F1: Knowledge Discovery in Databases (KDD)



Source: Adapted from Han et al. (2012)

BEYOND KNOWLEDGE DISCOVERY IN DATABASES

A relational database is “a collection of tables, each of which is assigned a unique name. Each table consists of a set of attributes (columns or fields) and usually stores a large set of tuples (records or rows)” (Han et al. 2012, p. 9).

A relational database can be query in terms of the attributes of the different unit of analysis, which are stored as a collection of tables. For instance, if our unit of analysis is our employees, we will have different attributes included in a wide array of tables: performance over time, employee demographic data, training history, etc. Query language (Structured Query Language - SQL) is the language used to interrogate relational databases and there exist software that through user-friendly interfaces allow us to represent these relationships in a more graphic fashion (e.g. Qlikview, Tableau are some examples).

A data warehouse, however, is a “repository of information collected from multiple sources, stored under a unified schema, and usually residing at a single site.” (p. 10).

Oftentimes, data warehouses store summarized inputs of databases. In this vain, data warehouses, for either HR or business data, offer specific information that has been selected carefully for business purposes and upon given compliance guidelines.

It is precisely data science literature, however, that tells us to look beyond databases and data storage. The KDD model has been superseded since with the advancement of more comprehensive cycles that incorporate three new elements: a business issue that needs to be addressed, the incorporation of computer systems that allow to manipulate data warehouses in the desired fashion and a subsequent action plan tailored to the specific business problem that is being tackled.

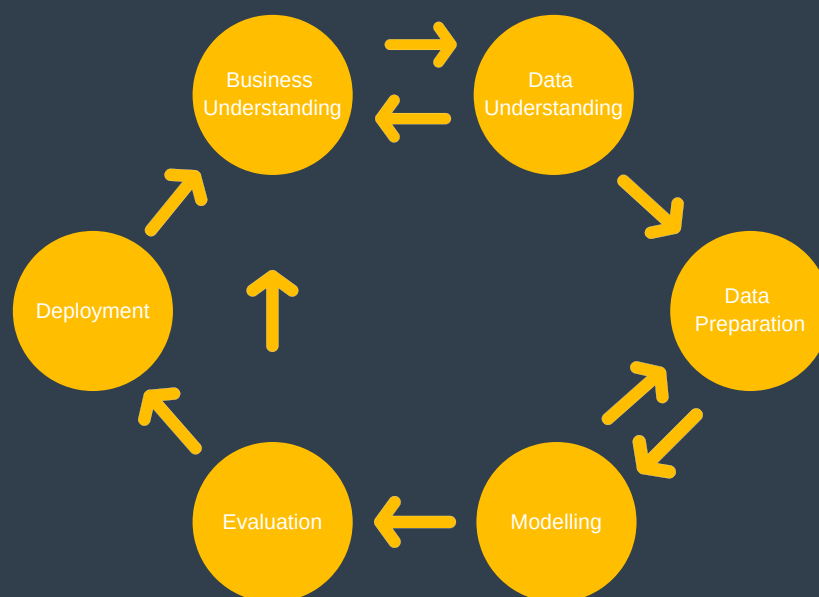


THE CRISP-DM MODEL

Data-mining, which is defined as the “the process of discovering interesting patterns and knowledge from large amounts of data” (Han et al. 2012, p. 8), constitutes just a component of a broader process of data analysis (See Figure 2). An example of data analysis model that superseded the KDD is the CRISP-DM (Cross Industry Standard Process for Data Mining) model proposed by industry experts and funded by the EU (Shearer, 2000). It commences with the understanding of a business problem and ends with

the deployment of an appropriate implementation and evaluation (See Figure 2). The CRISP-DM model was an attempt to conceive the data analysis cycle in a more comprehensive manner and it was developed by four European companies, namely, Integral Solutions Ltd. (a provider of commercial data mining solutions, later acquired by SPSS), NCR (a database provider), Daimler AG (the automobile manufacturer), and OHRA (an insurance company).

F2: CRISP-DM model cycle



Source: Adapted from Shearer (2000)

HR ANALYTICS: A BASELINE MODEL

Based on these developments, one can easily observe that the first model aims to make use of data through a multi-stage process in order to identify and extract patterns for the delivery of new knowledge. The latter data mining model, such as the CRISP-DM, however, encompasses a broader cycle that begins with a business problem and finds closure with a business solution or action plan (Shearer, 2000; Kurgan and Musilek, 2006).

The former model places the accent on the actual data analysis, the latter one focuses on facilitating the solution for a business problem through data-led decision-making.

In this vein, the definition of HR analytics provided by Marler and Boudreau (2017, p. 15) seems to be more accurate: “A HR practice enabled by information technology that uses descriptive data, visual, and statistical analysis of data related to HR processes, human capital, organizational performance, and external economic benchmarks to establish business impact and enable data-driven decision making”.

And yet, this definition becomes lengthy and somewhat complicated. For instance, why are external economic benchmarks included as a condition sine qua non for HR analytics?

The HR Analytics Management Cycle Baseline Model

Current scholarship emphasises the need for a HR analytics process starting with a HR problem (Levenson, 2018; Huselid, 2018) such as, high employee turnover or low employee retention levels, high sick leave or absenteeism rates, or alternatively with a HR question such as why do our employees choose to stay with us? (Davenport et al. 2010) or what are the predictors of store performance in our organisation? (Simon and Ferreiro, 2018). Thereafter, once the business problem has been fully understood in its proper context, a research design needs to be undertaken with a view to understand the data to be collected, data types and data sources that should be involved in studying the HR problem under investigation.

This second phase requires of a combination of capabilities, including HR acumen and statistical knowledge. The following stage of data preparation and curation, however, involves a separate skillset altogether, namely, data management, which draws from computer systems and it aims at the integration of data from different databases. Currently, there exist open source integration systems such as Hadoop, which through SQL or HIVE language data management professionals are able to interrogate data from Excel, Human Resource Information System (HRIS), and other Management Information System (MIS) (or enterprise IT systems), manipulate them and merged them into unified system from which data analysts can draw the ready-to-analyse data that needs to be modelled and analysed.

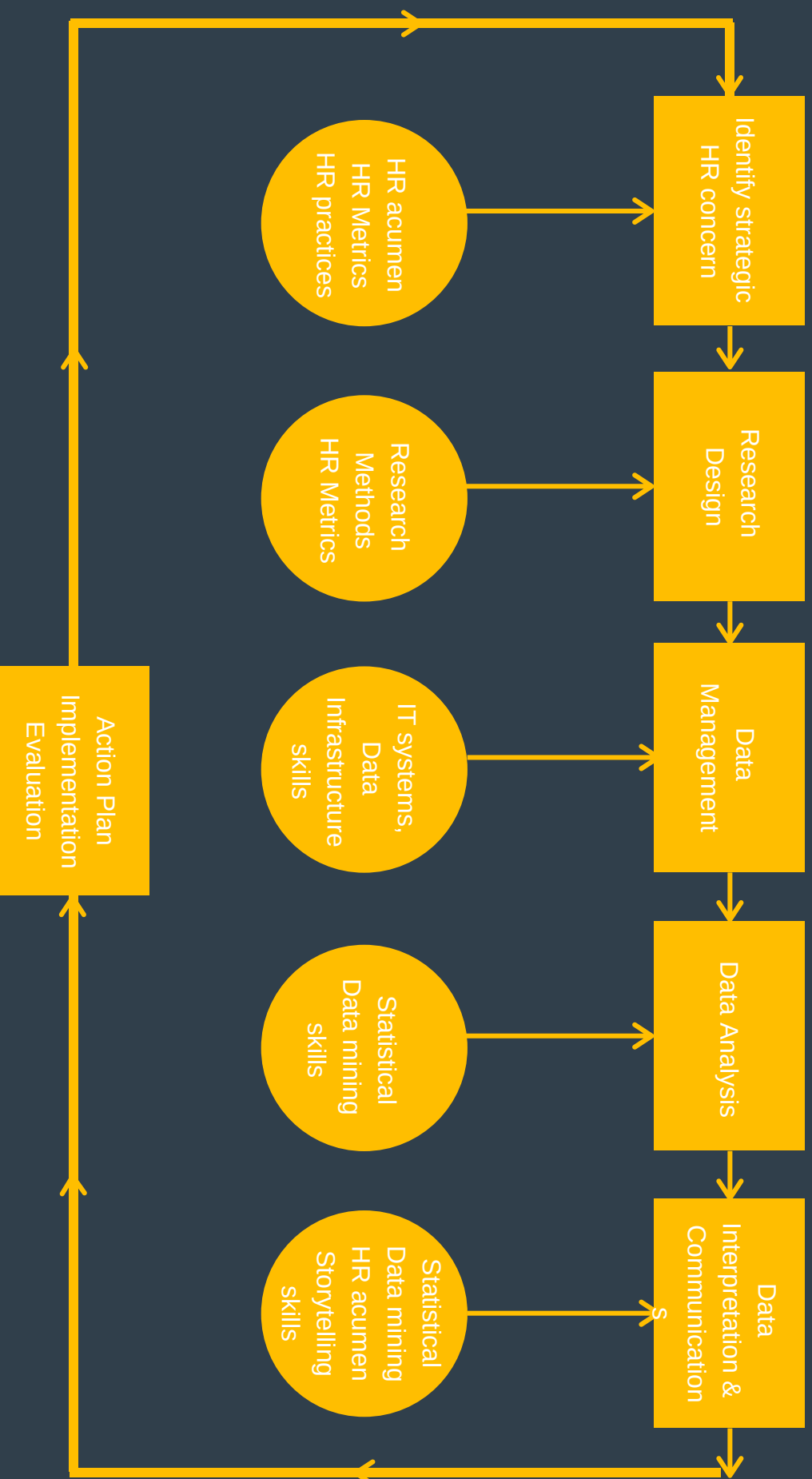
Finally, upon a careful statistical analysis, there it comes an output that needs to be interpret and explained, again coupling together statistical and HR acumen.

The definition of HR analytics that is proposed here tries to embrace both the data analysis management cycle and the diversity of capabilities needed:

“HR analytics is the process of addressing a strategic HR concern making use of HR data (and business and external data if necessary) and encompassing the following components: research design for the specific HR issue, data management, data analysis, data interpretation and communication, a subsequent action plan and evaluation” (see Figure 3).

The key to a decent deployment of HR analytics resides on knowing that a HR analytics team draws from a diverse range of capabilities that need to be managed (Minvaeva, 2018).

F3: HR Analytics Management Cycle: stages and capabilities



RESEARCH METHODS

Given the embryonic state and the variety of professional backgrounds in the HR Analytics field, the research approach taken in this project is the interpretivist approach, whereby the researcher not only engages in interpreting phenomena but also tries to understand the subjective perception of actors (Denzin, 2001; Lee and Ling, 2008), in this case, Heads of HR Analytics and HR analytics consultants. Invitations were sent through LinkedIn upon a profile search with the following keywords in their current job title and/or job description 'HR analytics', 'People analytics', 'Workforce analytics' and 'Talent analytics'.

For the purpose of building a HR Analytics management cycle baseline model, this report builds on 6 organisations, five based in Ireland and one in the UK, all of them with relatively small HR Analytics teams (3 team members on average), using sophisticated analytics and operating across different industries such as IT, aviation, finance, food and retail. It also includes the expert opinions of two HR analytics consultants and a HR Analytics expert and content creator based in the Netherlands with worldwide presence. All participants were interviewed in between April and June 2018, sixty minutes on average.

T1: Selected organisations

Features	Alpha	Beta	Gamma	Delta	Epsilon	Zeta
Nationality	US	Ireland	Ireland	Ireland	US	UK
Industry	IT	Food	Retail	Finance	IT	Aviation
Size of the organisation	+25,000	+20,000	+65,000	+10,000	+15,000	+40,000
Multinational or Domestic	MNC	MNC	MNC	Domestic	MNC	MNC
Respondent	HR Analytics and Insights Partner	Global HR Analytics Manager	HR Reporting and Analytics Manager	HR Reporting and Data Analytics Manager	Workforce Analytics Manager	Head of HR Analytics
Professional background of the respondent	General management and HRM	Mathematician, Data Analytics	Mathematician, Data Analytics	Physicist, Data Analytics	Talent Management	Computer Systems

IDENTIFYING A STRATEGIC HR CONCERN

The first stage of the HR analytics process is to pinpoint a HR issue that needs to be addressed in order to support the business strategy. Allocation of resources becomes a priority, even more so for comparatively low-budgeted business functions such as HR (Huselid, 2018). This concern will become the purpose of the HR analytics process or project. As Levenson (2018, p. 687) points out: “Competitive advantage analytics comes first, and involves prioritizing the business issues to ensure that any group-level and individual-level analyses are closely aligned with specific business challenges, and to ensure that the value of addressing a specific business challenge is understood relative to addressing alternative business challenges”.

In relation to the need for understanding the business and HR context with a view to support strategy, one of the participants commented:

“I suppose some of it is trying to get the guys ahead of things. I would sit down with the Head of HR to talk about what’s happening in the business, what’s coming over or what are the emerging challenges so that we can get ahead of us where possible” (Head of HR Analytics, Alpha).

A HR strategic priority has to be identified, one that can be studied at an aggregated level (organizational level or team level) and ideally that can be actionable at individual level (Levenson, 2018). The Head of Analytics of an Irish-owned multinational gives an excellent example of this:

“What I’d love to be able to get is a point where we can have a trending for particular employees that leave us. Let’s say, for example, we’re looking at odd days off here and there, or people who’ve been here a certain length of time, and they haven’t been promoted in this length of time, and they haven’t had a pay raise in this length of time. So that we can say there’s a risk of 10% of our specific population that potentially might be leaving, and these are the 10 people you need to be focused on” (Head of HR Analytics, Beta).

Thereafter, a time for hypotheses development comes. One of the participants notes:

“Generally speaking I take the approach that in trying to form a hypothesis and the quantitative analysis side, I would try and at least have some listening sessions or some initial inputs from the business to make sure that it is not just my impression or it wasn’t just what I was able to research myself from other companies because every company has a different context and a different narrative [...] In trying to understand, to define the HR problem and to define hypotheses against that, you would do some research online, you’re trying to find out and understand. ‘Okay, how are other companies trying to understand this? What elements do we need to look at?’” (Head of HR Analytics, Epsilon).

Wondering if HR analytics practitioners make use of academic research for hypotheses development, one of the participants commented:

“Yes, sometimes I would use academic work [...] to get a sort of hypothesis of something I am studying, to get a theory or an idea of how something should be or how something is and if the research is already done and I can then use that to apply it internally” (Head of HR Analytics, Gamma).

Hypotheses development

< Generally, companies do not have access to academic journals where a good deal of HR concepts have been developed >

However, what practitioners call academic work becomes blurred and there it emerges the fact that most HR analytics professionals do not have access to purely academic work, namely, access to top tier academic journals. One of the participants explains it quite succinctly:

“I use CEB (Global CEB – Gartners) quite a bit and I get... I suppose Twitter has quite a bit on Harvard Business Review, some of those. I wouldn’t be the best in terms of researching but I do look at them” (Head of HR Analytics, Beta)

Another participant tells us: *“Well, whatever is available on the net [...] We have access to some specific stuff like CEB (Global CEB – Gartners) but I don’t know how much it does proper academic research” (Head of HR Analytics, Gamma).*

At this point, it becomes clearer that the use of high-quality academic research in HR analytics seems to be rare, not because of a lack of professionalism of the part of HR analytics teams but simply because HR analytics professionals do not seem to have access to academic journals, with some exceptions such as the Harvard Business Review. It is important to note that academic research is generally more rigorous than any other industry-oriented research and more importantly, this quality will pay off in terms of saving time in the process of defining a specific theoretical model on which to build the hypotheses. Only one participants in the study had access to academic journal articles not through their organisation but via their student status while taking an analytics course at a third level education institution.

One of the consultants interviewed affirms:

“We need the papers. I try to make as much use as possible of academic journals. Of course, we need to do something that has been done many times and there is a quite good internal knowledge and we can go back for that [...] but we are trying not just do analytics doing business, but evidence-based business. I have access to academic journals, personally I have but not through my organisation” (CONS2).



1

Hypotheses development in industry can benefit from academic research

2

HR Analytics professionals can benefit from academic literature review training

RESEARCH DESIGN

Once the HR concern has been narrowed down to a specific question, the HR analytics team will undertake a careful research design that incorporates the identification of the dependent variable, independent variables and the types of relationship existing among them. It is here when the HR analytics team will develop hypotheses, explore types of variables and measurements and define the type of data that they need and where to source it. Oftentimes, they might have the data already internally stored (through an integrated data warehouse, an integrated HRIS, a series of data platforms across HR practices or just Excel files) whereas other times, they might have to gather their data specifically for that particular HR analytics project in order to analyse it on its own or in combination with stored HR and/or business data.

When asked about the research design process, one of the participants points out the following:

"I suppose in terms of the design I would concentrate on the business problem. Getting people together and following a loose methodology around customer-centric design, where we get lots of input from people about framing the issue, the topic. What's the problem? What's the bypasses around the problem? Where is the problem? How much of a problem is it? [...] Once I have that frame it's only then we'll start looking at data that has come up around parts and pockets and trying to build a data structure and release and then, to find if we can answer those questions. We end up, I suppose, the output from that analysis will be a set of questions that need to be answered about the dependent variable [...] We would build a data model on paper" (Head of HR Analytics, Delta).



In-house or outsourced HR metrics?

In terms of the capabilities needed at this stage, HR acumen becomes crucial, coupled with an extensive knowledge on generic HR metrics. Oftentimes, the HR metrics needed to be modelled and tested are easily to find or generate in-house. However, other times, these metrics might be provided by an HR analytics consultant.

< Budgetary limitations can be a reason to hire a consultant. However, most times it will depend on the type of HR data we need and whether it has to be benchmarked against a broader population >

"In most cases, HR comes to us and we deploy the methodologies and the technology to mine the data. [...] what most HRIS provide I describe it as extrinsic data, extrinsic to the person. What we provide is intrinsic data. We work from an inside out approach first. [...] we have analytics that defines 12 dimensions of leadership paradox. One dimension of leadership paradox is communication. For organisations, it would be helpful then to get some metrics around where their leadership is at against those dimensions. [...] then, we can also get a measurement of the team itself and see what the weak points are. Then, we know where we like them to be, now we have a very clear picture of what we need to do to put in place development interventions to close the gap, both at individual and team level. The great things about the analytics that we use is that we can not only get a measure of where the team is at now, and where the individual leaders are now, but we can measure that again in six months' time and measure it again in a year's time" (CEO, CONS1)

DATA MANAGEMENT

This phase of the HR analytics project begins once the collection of data have taken place. Sources of data are varied and range from internal data to data that can be collected through ordinary HR information, surveys or artificial intelligence tools (e.g. text analysis, personality and personal qualities assessments, voice recognition, etc.).

Once the sources of data have been selected, there is a time of data curation, preparation and storage into the appropriate repository, a data warehouse.

These will possibly entail the interrogation of a diverse array of databases and an integration of the data into either a sole database or data warehouse platform.

"If you look at the life cycle of HR analytics, I suppose, from the data creator to the data consumer, I would say the data gathering and data integration pieces was probably 80% of where the effort of the team was lying as the starting point. The data analysis piece and the visualisation delivery piece would be less than 10% of the effort because there was no time [...] We've switched that around now. We've carried a huge piece of work around the data. The data structures in particular. Making sure we integrated our data, automated all the different systems. We load systems into one place. Create that golden source of data" (Head of HR Analytics, Delta)

"We use SQL to interrogate the database and we use Tableau and Excel and other products to display the information"

Head of HR Analytics, Zeta

OPEN SOURCE TOOLS?

The question at this juncture is: is a data warehouse the dream of organisations but only financially attainable by large firms? Not necessarily if the organisation has the right capabilities in place, namely, data management skills.

The skills needed to integrate the data differ from those needed to analyse the data. Thus, it becomes a priority to identify what skillsets are indispensable for the type of data we are dealing with, the specific data systems we have (if any) and above all, to figure out whether the composition and skill profile of our HR analytics team is the appropriate in each case.

For those organisations that do not have the capabilities needed to use open source data warehouses and statistical software, there are indeed solutions that offer certain functionality. However, HR analytics teams should be mindful that these analytical tools have limited functionality (e.g. they only operate under a particular programming language, they only interrogate data from certain HRIS, etc.).

"We use mostly Alteryx and Tableau [for the analysis of data]. These are the two we use, which is a very nice combination and we get a lot of our data from Terada, it's our data warehouse" (Head of HR Analytics, Gamma).



Open Source

Efficient HR analytics and insights do not have to be done under a substantial budget necessarily: work with open source software and programming languages

Terminology:

Ecosystem or Data Lake: combination of programs that together supports a large scale app. Here raw and unstructured data is stored (e.g. Hadoop).

Data Warehouse: platform where data is stored, processed, aggregated and partitioned (e.g. Teradata, Cloudera, Apache Hive).

Query Language: language used to interrogate a database, modify, join or merge data (e.g. SQL, HIVE).

Statistical packages: used to analyse data (e.g. R, SPSS, Python-based packages).

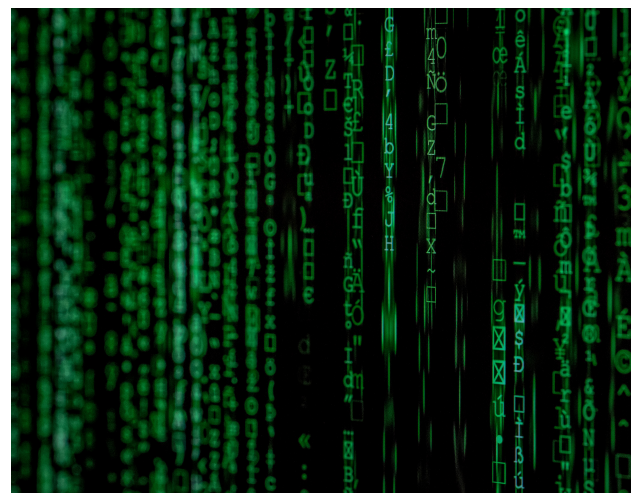
Visualisation tools: used to transmit the findings through a user-friendly visualisation tool (e.g. Tableau

As a way of example, Teradata is also a data science solution provider that offers software to manage cloud systems, data warehouse ecosystems, and other analytical solutions applied to certain business functions such as operations, value chains or customer services. They provide the databases barebones or structure along with a data engine that collects data from the business operations. These data bases skeletons are unique to every industry, and thus the Teradata Database differs across sectors such as financial, manufacturing, healthcare, retail, etc.

These ready-to-make data bases engines are indeed a suitable approach for companies that are not interested in hiring a data manager or a data management team. To a certain extent and depending on the sophistication of the organisations' data bases, acquiring a data management solution might be a way to upgrade their data systems. On the other hand, these solutions are always closed systems to a certain degree and this might limit the freedom of creativity when designing the organisation's own data warehouse, which as we know can be created through open source software tools.

The key to a successful data management is to figure out whether it is necessary and worth designing and generating the organisation's unique data warehouse versus purchasing a solution that creates it based on a built-in database structure suitable and designed specifically for your industry. In this vein, another Head of HR Analytics notes:

"Because of my understanding of systems and my understanding of the data, I then was approached by the HR team to take on the role of manager of HR analytics and I think the key thing at that point when I joined the team in 2013 was actually not data analysis in itself whilst we spent a lot of time looking at the systems, the way in which the data was held in the warehouse, we hadn't spent much time on-- I did have a warehouse of data, which was probably, fair to say, pretty ill-designed, not very efficient, riddled with complexity. We spent much of the last five years trying to sort it out so that analytics becomes an easier job" (Head of HR Analytics, Zeta).



Challenges with HR Data Analysis

"First of all, the specialist [HR] data analytics people are going to be nearly impossible to find. They just haven't developed. It's such a new field, people haven't gone through the "Okay, I've got three years' experience". [...] There is more analytics people coming at us because of the courses offered in Irish third level education institutions, and everywhere but they're snapped up pretty quickly. So if the Head of HR tells you "you're doing a great job, go find three more people", it could take me six months, could take me a year because, again, you really want the right person and the right person isn't going to be hanging around with their feet up, but saying that I would be very comfortable hiring people straight out of an analytics degree that did not necessarily have any HR analytics specialisation. I would be very comfortable just hiring the smart, young, dynamic people that have new ideas, not necessarily have to have fifteen years of experience" (Head of HR Analytics, Gamma).

HR Analytics skills shortage appears top of the list. Some organisations make use of data scientist that are working for other business functions such as finance. When asked about the presence of HR data scientists or analysts, one of the participants explains the following:

"We have two data scientists in charged of HR data analysis. They're not HR data scientist, they're finance guys. Again, I can give them so much and anything I give them has to be anonymised or they do not understand because it's HR and they are specialised in finance" (Head of HR Analytics, Beta).

This shortage of skills within the HR department can pose concerns as to how accurate data scientists working for other management functions understand the specific HR problem we are trying to solve.

A challenge for data analysis robustness is sample size. There are some scholars that have highlighted that perhaps small and medium enterprises (SMEs) do not have HR big data (Cappelli, 2017). One of our participants, the creator of a HR analytics community and a HR analytics content creator says:

"I'm always very suspicious when people start selling analytics to small companies. Yes, you can do some analytics but if you have ten employees... When you have ten people you have very different worries. You need a certain economy of scale in order to make analytics work and that's not happening with ten people. I've worked with a hundred people, with a thousand people. Then, you have the economy of scale you need to make an impact. I think that the companies that are selling analytics to small companies... I think that, to put it mildly, content wise is close to being unethical. I can't say it in a black and white manner, but it feels a bit like they're selling snake oil" (HR Analytics Expert & Content Creator).

1

Shortage of HR Analysts

2

*Small samples, particularly in
SMEs*

However, some participants have pointed out an alternative to this shortage of data in SMEs. A consultant notes:

“It doesn’t mean that analytics are not useful for smaller companies. It’s just that it is really a problem of size. The bigger you are, the better, the more you can collect. Actually, I think it would be great if smaller companies that work in the same business within the same area could organise analytics as a group. You could have, for example, hotels in a specific region that share a lot of common characteristics.

They share similar workforce, but each of them has 30, 40, 50 employees. Altogether, they would make a good sample size. That’s a way in which smaller companies could make use, could use analytics and work well” (HR Data Scientist, CONS2).



The fact of the matter is that it all depends on the type of data and the type of analysis needed to solve your HR concern. If you are employing analytics to identify training needs across 10 sales associates, analysing a few hundred of their phone conversations, then sample size is not a problem since your unit of analysis is a compilation of voice files that will allow you to make interventions in a personalised fashion, tailored to the training needs of each sales associate. The same might be the case of text analysis for an employee engagement survey given that a suitable number of open questions have been answered and can undergo qualitative or quantitative analysis.

DATA INTERPRETATION & STORYTELLING

Interpreting the data is an important part of this process and it is the summit of the HR analytics business case where direct and indirect impact of the HR analytics project is spelt out. In order to do so, there seems to be a need to put together the business acumen (including HR acumen) and the data analysis capabilities. One the participants tells us:

"We sit down on a monthly basis and look at the system's analytics and then we say, "What does that mean?" We talk through together what it might mean. [...] I'm sitting down with one of the senior HR VPs. I'm doing a very specific piece of work on our employees and working very much closely with him on understanding and working out what the data is telling us. I'll press him on what he sees and he'll press me on what I see and we'd sort it together." (Head of Analytics, Gamma).

A quantifiable effect or impact to present to the organisation leadership seems to become essential when heading towards the front end of the HR analytics management cycle.



One of the participants explains this with a graphic example:

"We had done some regressions and we were able to predict for every unit increase in employee happiness that you could potentially have more product made. When we presented that to the site leadership, they were very skeptical about the engagement tool at first. They thought this is very fluffy. This is very airy-fairy. We were able to say that if you can improve one unit of happiness on the scale of 1 to 10 from where it is now, which was I think 6.8 to 7.8, you look at 11.5% extra productivity from the shift on a quarterly basis. It was quantifiable. We did say that, look, we had data available to us" (Head of HR Analytics, Epsilon).

Most participants agreed on the paramount priority of communicating the findings to HR and the organisation leadership in a palatable manner. The statistical output needs to be communicated and presented in such a way that makes sense for your audience, namely, HR professionals and CEOs in most cases.

They describe the necessity to develop a story for the business, based on the significant data and parameters they uncovered. One of our participants explains it this way:

“My view is because HR is the poor relative to everything else to actually generate income. We don't have much money to spend. I guess, we can't find it, we've got business case to make some changes on the data because there's no, often no business case because it's all about making it easier to spin up the data. I think for me, the storytelling piece is really important because of both HR and the business, not just HR. The CEO needs to understand why we need to spend money doing this so that we can look at that. When you stop playing in the fact that your employee cost is the wrong controllable cost then you have above anything else... Doing the storytelling, helping people understand the why's, and the what for's, and the so what's, is really important.

I think Marketing can get away with it from the revenue generating area. You can be very bright and smart and you look at the ways in which you can change the pricing of your product and that's real key. I think the selling part is less required. But for me, possibly we're the poor relatives and so we have to justify it to ourselves a little bit harder” (Head of HR Analytics, Gamma).

THE OLD HR BUSINESS CASE?



ACTION PLAN & EVALUATION

Upon the delivery of findings and having clarified the direct and indirect effects of the statistical output, practitioners mention the development of an action plan, usually in the form of interventions, that is, changes in HR policies and practices. These changes are generally different according to teams, employee types, etc. The action plan is, therefore, elaborated at the individual level of analysis (Levenson, 2018). One of our participants provides an illustrative example referring to the impact of flexible working arrangements on employee retention in his organisation:

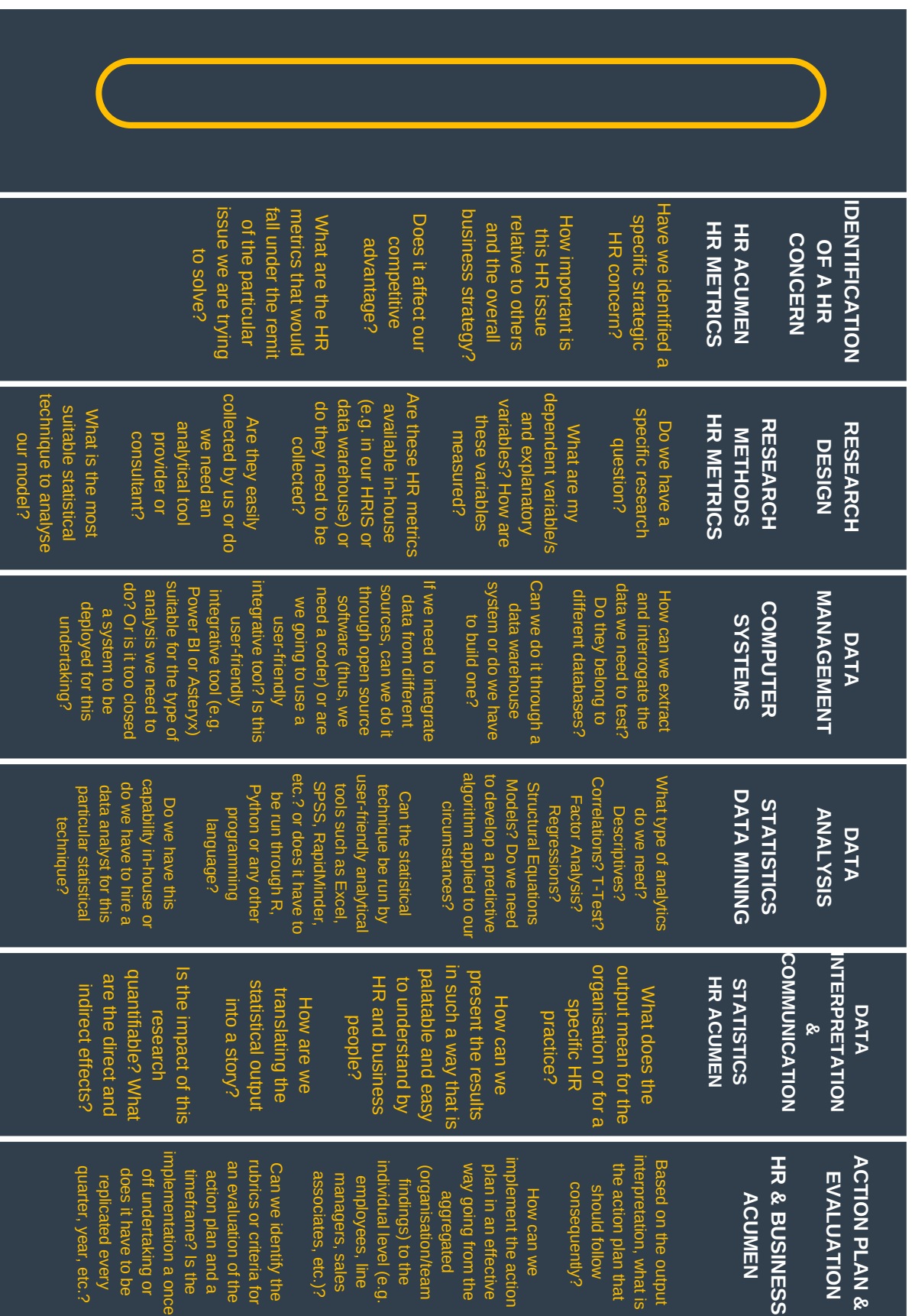
“That input is now taken into the HR policy, specifically trying to come up with a uniform flexible working policy, which seems to be the most significant predictor of retention in our organisation. Now, there will be different components to that policy for different workers, but in the context of those unskilled workers who are working on shift, the main problem for them was that they were not in an office job that if they needed to log on from home for the afternoon because they had to pick up one of their kids, they could do that. They were working in the factory, they had to do [physical] work. The pain for them was having to arrange a swap with someone

else. That wasn't always available. The policy for that group that may be implemented may be different from the other group, but at least it will be consistent” (Head of HR Analytics, Epsilon).

Another participant was asked whether the HR practices were changing based on the data outputs her team had provided in the past:

“They absolutely are. I think the thing that has worked in the last year was making sure my relationship with the HR VP team is completely tight and solid. I meet with the head of HR at least once a fortnight, definitely to go through the analytics. They are putting interventions in place to address specific issues that they are finding. The thing for me is about making sure that the analytics is robust and repeatable and just runs, because they might put one intervention in place that has absolutely no impact whatsoever so then they have to go back to think what is it that we can do next to come and put down whatever we want to do. I would say to them, “I've done the easy bit, all I'm doing is support the business. You need to find out what's going to shift this” (Head of HR Analytics, Gamma).

T2: A HR Analytics Management Cycle



CHALLENGES MANAGING THE HR ANALYTICS CYCLE

There is certainly a good deal of merit on the work of Davenport and colleagues on the uses of HR analytics and the work of John Bersin on his so-called 'talent analytics maturity model'. There is indeed great variety of sophistication levels across organisations (Davenport et al. 2010). As one participant explain, upgrading on this maturity scale can be the source of new blood into the HR analytics function:

"There is a danger, there is an inherent danger in analytics just as in Agile, that analytics is such a loaded term and means different things to different people. Depending on what an organisation is used to, an analytics solution that is better than what they had could be a chart on an Excel form, because if all they've had before is Excel reports of tables, then something that is slightly visual is an improvement and that's their first step on the journey" (Head of HR Analytics, Epsilon).

A cognitive effort is needed to comprehend the different stages of the HR analytics process and the diverse array of skills key to each stage.



There seems to be also some truth on the relevant role of data infrastructure and data management sophistication. When asked about the proportion of effort dedicated to building the data management of the participants notes:

"Yes. It's definitely 80% data, 20% analytics" (Head of HR Analytics, Gamma).

Finally, given the relational nature of the HR Analytics management cycle, managing a cross-disciplinary team is crucial to the success of HR Analytics projects. Can our HR professionals and research designers talk to our data managers? Can our data managers understand what our data analyst or scientist need to run the statistical techniques? Can the findings be presented to the Head of the HR in a comprehensible manner so that s/he can make sense of it? Thus, the HR Analytics management cycle does not just depend on the sophistication or maturity of the analytics but also on a holistic view of the process, of the required skillset across the different stages and on a successful management of a cross-disciplinary team.



Acknowledgements

My most sincere thanks to all participants, those based in Ireland and abroad: for their time, experiences, comments and feedback. Thanks to Prof. John Geary, Prof. Bill Roche, Prof. Anne Keegan and colleagues from the HRM&ER Subject Area in the UCD College of Business for their unconditional support.

REFERENCES

- Angrave, D, Charlwood, A, Kirkpatrick, I, Lawrence, M, Stuart, M (2016) HR and analytics: why HR is set to fail the big data challenge. *Human Res. Man. J.*, 26 (1), 1-11.
- Bersin, J, O'Leonard, K, Wang-Audia, W (2013) *High-Impact Talent Analytics: Building a World-Class HR Measurement and Analytics Function*, Bersin, Deloitte Consulting LLP.
- Cappelli, P. (2017) There is no such thing as big data in HR, *Harvard Bus. Rev.*, June 2017.
- Davenport, TH, Harris, JG, Morison, R (2010) *Analytics at work*(Harvard Bus. Sch. Press, Cambridge, MA)
- Denzin, NK (2001) *Interpretive Interactionism* (Sage, London).
- Fayyad, UM, Piatetsky-Shapiro, G, Smyth, P (1995) From Data Mining to Knowledge Discovery: An Overview. In U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (eds.). *Adv. in Know. Disc. and Data Min.* MIT Press, Cambridge, MA.
- Huselid, M (2018) The science and practice of workforce analytics: Introduction to the HRM special issue, *Human Res. Man.*, 57 (3): 679-684.
- Kukla, A (2000), *Social Construction and the Philosophy of Science*(Routledge, London).
- Kurgan, LA, Musilek, P (2006) A Survey of Knowledge Discovery and Data Mining Process Models. *The Know. Eng. Rev.*, 21 (1): 1-24.
- Latour, B. and Woolgar, S. (1979). *Laboratory Life: The Construction of Scientific Facts*, (Princeton University Press, Princeton).
- Levenson A. (2018) Using workforce analytics to improve strategy execution. *Hum. Res. Man.*, 57 (1): 685–700.

Marler, JH, Boudreau, JW (2017) An evidence-based review of HR Analytics. *The Int. J. of Hum. Res. Man.*, 28 (1): 3-26.

Minbaeva, DB (2018) Building credible human capital analytics for organizational competitive advantage. *Hum. Res. Man.*, 57(3): 701-713.

Rubin. HJ, Rubin, IS (1995) *Qualitative Interviewing: The Art of Hearing Data* (Sage: London)

Shearer, C (2000) The CRISP-DM Model: The New Blueprint for Data Mining, *J. of Data Wareh.*, 5 (4): 13-22.

Simón, C, Ferreiro, E (2018). Workforce analytics: A case study of scholar-practitioner collaboration. *Hum. Res. Man.*, 57 (3): 781-793.



HR ANALYTICS

- A game-changer if one knows how to use it -

WHAT IS THE BUSINESS PROBLEM?



Can we pin down the specific business issue we need to tackle?

What are the business and HR metrics that are involved particular issue we are trying to solve?
Is it so clear that we can build a testing model?

RESEARCH DESIGN



Do we have a research question?

What are my dependent and explanatory variables?
How are these variables measured? Are they available or do we need to collect the data?
Do we have the capability in-house or should we hire a consultant? What statistical technique would be useful?

DATA MANAGEMENT



Ecosystem, Database, or both?

How can we extract and interrogate the data we need to test? Do they belong to different databases?
Can we do it through a data warehouse system or do we have to build one?
What are the skills needed for the extraction, cleaning, manipulation and merging of data?

DATA SYSTEMS



Putting out a fire or long term infrastructure?

Whether our data infrastructure is big or small, is it designed for the long run?
Do we need an ample budget or can we use open source technology?
Can we work towards a business intelligence ecosystem or is it too early?
Do we have support from the leadership team?

DATA KNOW-HOW & ANALYSIS

Data terminology is not interchangeable

Ecosystem or Data Lake: combination of programs that together supports a large scale app. Here we store raw and unstructured data (Example: Hadoop).

Data warehouse: platform where we store processed, structured, aggregated and partitioned data (Example: Teradata, Alteryx, Cloudera, or Apache Hive).



Query language: language used to interrogate a database, modify, join, add or merge data.

SQL (sequel): structured query language used to interrogate databases.

Hive: query language based on SQL and designed by Facebook.

Statistical analysis of big data: traditional statistical techniques run through packages able to manage big data: Python, R, SPSS.

Python-based packages: Jupyter, Spider, NumPy, StatsModel, etc.

DATA VISUALISATION

An image goes a long way



Are we able to transmit the findings through user-friendly visualisation tools?

Can our colleagues understand the graphics with ease?

Are our visuals able to convey actionable meaning?

Tableau & QlikView can be manipulated with Python, which is open source.

INTERPRETATION & STORYTELLING

Communicate impact



What does the output mean for the organisation or for a specific HR practice?

SKILLS OF THE HR ANALYTICS PARTNER

Unique profile



Business and HR acumen
HR metrics
Exposure to HRIS (e.g. Workday, SuccessFactors)
Research design, data collection, data management and statistical analysis (descriptives, correlations, factor analysis, t-tests, regressions)
Knowledgeable of the HR Analytics management cycle from end to end
Project Management skills
Knowledgeable of data science tools available in terms of ecosystems, databases and languages that are used for data analytics
Knowledgeable and some exposure to open source/licenced data science tools (SQL, Excel, SPSS, Tableau...)
Ability to identify business and HR needs in the data
Ability to uncover stories in HR data to redefine HR strategy and practices
Leverage the work of leaders and colleagues through the insights the team delivers
Training and mentoring other leaders
Ability to manage cross-disciplinary projects (data infrastructure engineers, data scientists and HR colleagues / Team Leads)
Compelling storytelling able to drive action

BE MINDFUL OF...

Small print



GDPR and ethical considerations in the use of HR data

The most efficient HR analytics and insights do not operate under a substantial budget necessarily: work with open source software and programming languages

There is no need for a sophisticated ecosystem to produce interesting data-driven decision making. The data infrastructure is important though, count with the help of data infrastructure / computer systems engineers

The HR analytics partner's skillset is unique and rare, PhDs across disciplines using quantitative data analysis can be a good starting point, provided they have some business/HR acumen or exposure

The HR analytics team should enjoy considerable autonomy to design, undertake and manage projects

© Maria Belizon 2018