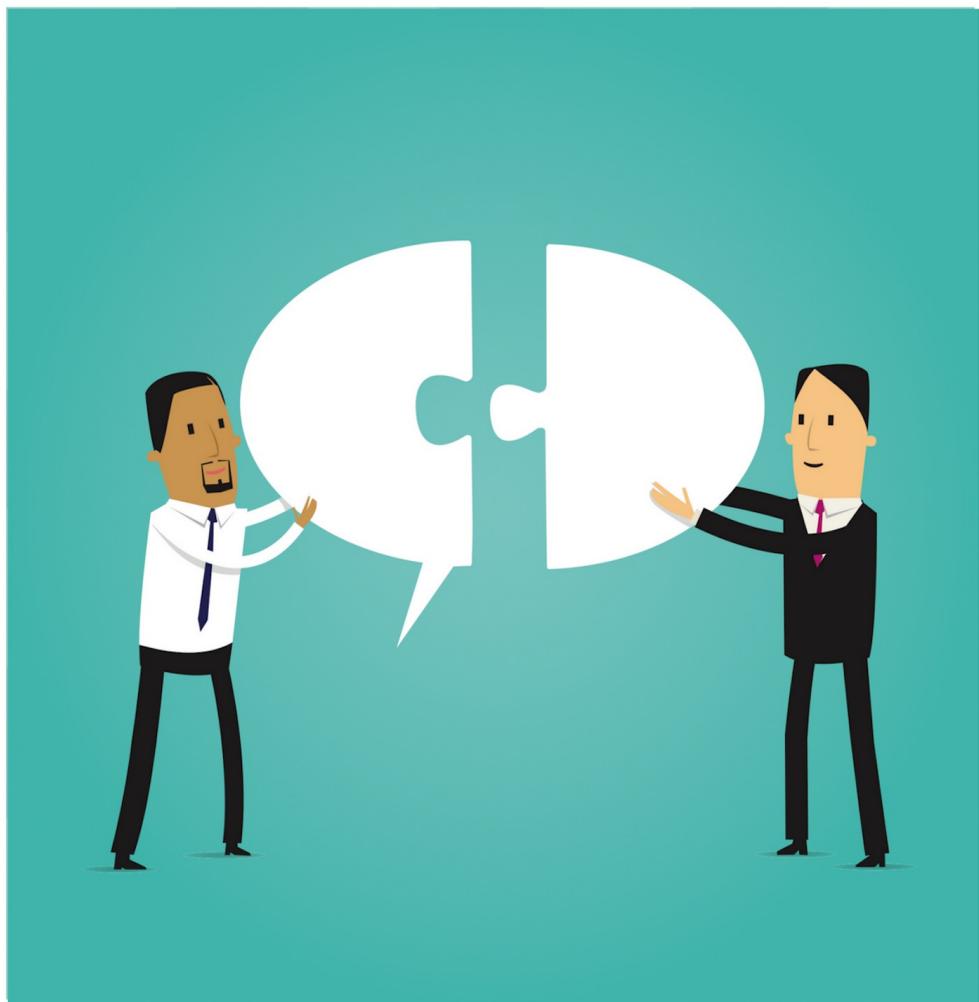


PEOPLE ANALYTICS

DATA AND TEXT ANALYTICS
FOR HUMAN RESOURCES



EDUARDO VALENCIA



People Analytics

Data and Text Analytics for Human Resources

Eduardo Valencia

This book is for sale at <http://leanpub.com/people-analytics>

This version was published on 2018-05-09



This is a [Leanpub](#) book. Leanpub empowers authors and publishers with the Lean Publishing process. [Lean Publishing](#) is the act of publishing an in-progress ebook using lightweight tools and many iterations to get reader feedback, pivot until you have the right book and build traction once you do.

© 2016 - 2018 Eduardo Valencia

Contents

Foreword	i
Jose Gonzalez: People Analytics the easy way	i
Burn After Reading!	iii
More about People Analytics	v
Is This Book for Me?	vi
Structured and Unstructured Data	vii
What Will You Find in This Book?	vii
Table of Instructions	x
Without further ado, I want you to know more about me and my flaws.	xi
Vampirizing <i>Data Coaching</i>	xiv
It's a fun book that you can learn from.	xviii
Illustrations by Marisa Echarri	xviii
I. An Introduction to People Analytics	1
01. What Is People Analytics About?	2
1. About Moneyball	2
2. The Most Valuable Asset is Also the Most Expensive.	7
3. Tomayto, Tomahto	8
4. Big Data for Human Resources	11
5. The Path to Transformation	12

CONTENTS

Recommended Reading	25
-------------------------------	----

Foreword

Jose Gonzalez: People Analytics the easy way

You may have seen a curious British film by Lasse Hallstrom: *Salmon Fishing in Yemen*. It's a rom-com about an Arab sheik's visionary (or perhaps ridiculous) dream to breed salmon in the heart of Yemen.

See the trailer of the film on Youtube¹



Salmon Fishing in Yemen. And analytics in human resources... Of course!

If you work in human resources (or “talent” or “people management”, if you prefer), for many years you’ve probably been hearing your colleagues in other areas (operations, sales, support, and even finance) speak about the growing importance of analytics and data science and their impact on business. And it must have crossed your mind more than once that the volume of information that your department manages, along with its elusive, diverse, and changing nature, would more than justify the application of data-based approaches. You may have also argued to those same colleagues that

¹<https://www.youtube.com/watch?v=JSYuTFK8Eas>

any such investment would have a multiplier effect on the business, by affecting aspects as essential as the fidelity (and happiness) of employees and attracting and retaining talent.

If that's your case, in this book by Eduardo Valencia you will find a full catalog of applications and an extensive repertoire of analytical tools, along with countless tips, precautions, and how-tos. And if that's not your case, this book will open your eyes. I guarantee it.

I met Eduardo a little over four years ago through a mutual friend. Since we first met, I thought he would be an extraordinary addition to our company. He has a rare combination of analytical reasoning, a creative mind, and work capacity. And to top it all off, he's funny. As for the book you are holding, you will see that it is also very instructional. Despite its length (which now doubles the estimate he gave me nine months ago), the book keeps you hooked continuously, even if it's only because of the tour he takes you on through a good collection of films that are part of the collective memory of a few generations, regardless of nationality, language, or creed, or because you enjoy the links to Youtube which make a nice pause in reading. Eduardo is also a movie buff (and, by the way, the one who suggested I see the salmon movie).

I have already confessed that Eduardo tricked me on the length of the book. But not only on that... We are talking about an orientation that would apply language technology (or natural language processing) to human resources. At the end of the day, this discipline, which brings together artificial intelligence, linguistics, and computation, makes up the foundation of our activities in MeaningCloud. But no: Eduardo had to make it an exercise of rationality and completeness to put this discipline in the context of the other analytical techniques: Fifteen chapters of context.

And here is the result at last. Dear Human Resources Management Professional, you are in luck.

Jose Gonzalez
CEO at MeaningCloud

Burn After Reading!

“People analytics” is the same as analytics and data for human resources. Or if you prefer, big data for human resources. Although, as you will see later if you continue reading, you shouldn’t say “big data” unless it really is big data.

As the proverb goes, *“Give a man a fish, and you feed him for a day; teach a man to fish, and you feed him for a lifetime.”* My unfinished business: Go from giving you fish (facts, suggestions, case studies, etc.) to teaching you how to fish by showing you how to put into practice what I’ve been preaching about.

I have worked in the area of analytics and data for quite a few years. Lately, I’ve been specializing in data for use in human resources departments.

I have spent a lot of time doing consulting work and training others. After working two years for its subsidiary in Spain, I just started working in New York for [MeaningCloud](#)², a company specializing in text analytics, where we have developed valuable resources to analyze unstructured data (that is, text) used in people analytics and in human resources in general.

I solemnly swear that MeaningCloud is a great place to work. It is full of highly qualified people where everyone is taken care of, even weirdos like me.

Currently, I also head up [People Analytics Pro](#)³, where we host a 140-hour Master’s program that teaches HR professionals to use data and analytics tools to make informed decisions that improve employee management in an organization. In other words, we help prepare HR analysts.

²<http://www.meaningcloud.com>

³<http://www.peopleanalyticspro.com>

The book you have just started reading was born from this experience. It's a practical tour of the most important points that are taught in the People Analytics Pro Master's program. In fact, this book is conceived of as a practical guide for how to put into place what I previously only talked about.

Indeed, for a long time, I was slightly frustrated with the results of my proselytizing. I would teach seminars and courses about people analytics. The content of these sessions was and still is, of the type: What is people analytics? What is it for?, Big data and human resources, application areas, methodology, case studies, blah blah blah... My most frequent takeaway from these sessions are comments from motivated attendees: "That's so interesting!", "You're completely right, Eddy, about making decisions based on data!" But they usually add, "and now what?", "How do I do it?", "How can I use data to improve employee selection or rotation in my company?"

My answers to these questions, in one way or another, once I took away the various bells and whistles that my years of salesmanship taught me to add, came down to the same thing:

"Call us, and we'll talk." Or to put it bluntly, **"Hire us, and you'll see how well it'll go!"**

While spouting off these lines, trying unsuccessfully to not look too pleased, I would hand them my business card. Asking to be hired is, I must confess, the main reason for these seminars, and I don't think there's anything necessarily wrong with that. So, I would think to myself, "Mission accomplished."

But, at the same time, I couldn't help noticing that what the attendees were actually asking was "How can we learn to do what you're telling us about?", "How can we learn to make predictive analytics for, as you say, to select people who are more productive, more faithful, and better aligned with the company's culture?" In short, "I want to learn" or "I want my company to learn it."

In the same style of "fish giving," now two years ago, I published in Spanish *Data Coaching. A little book about data, and profit.*⁴ in Spanish. The book talked about how much you can make by taking advantage of data, and such topics. Moral sermons sprinkled with funny anecdotes, often with brash language, and a few real cases. But once again, it stayed there in the realm of good ideas, without telling you how to make them a reality.

That is precisely what this book is for: to go from preaching about a concept to teaching you how to make it work.

To cover all the bases, this book contains both the dissemination-preaching part, like in my seminars and in **Data Coaching**, and also the step-by-step instructional part for those who want to do it themselves. Later in the "**Table of Instructions**," I tell you how to read this book according to your interests.

However, I recognize that this **Gotta Catch 'Em All** approach could possibly be another bad decision. Hakuna Matata! I'm used to it, so I can handle this frustration. I'll learn something for the next book. Or not.

More about People Analytics

People analytics is based on applying data science methodologies and techniques to the field of human resources with the aim of getting to know the people in our organization better and increasing their level of satisfaction and productivity.

⁴<http://www.eduardovalencia.com/p/data-coaching.html>



In people analytics we use the evidence that the data provides to respond to several questions:

Which candidate will be high-performing, effective, loyal, and aligned with the corporate culture?

How can we measure the economic impact of a training program?

How can I segment the workforce to make their actions more effective?

Which people are considering leaving the organization?

What net benefit will employees contribute throughout their time in a particular position?

How does employee commitment affect productivity and economic outcomes?

How can I design a study that is statistically and mathematically valid?

Is This Book for Me?

This book is aimed at professionals who have responsibilities over people and want to learn how data and analytics will help them make better decisions.

More specifically, it is aimed at:

- Managers of organizations who want to learn how to extract more value from the data collected by the HR department.
- HR leaders who wish to either incorporate analytics and data departments, or acquire new techniques of value extraction using data.

- HR pros who want to learn how to use people analytics in the daily tasks of their professional activity.
- Data analytics pros even if they are not specialized in human resources.

Structured and Unstructured Data

In 1998, Merrill Lynch claimed that about 80–90% of all potentially usable corporate information is unstructured data. Although no one has proven that this is the actual figure, we all accept and repeat it. Technically, "unstructured data" includes photographs or videos in addition to text, but those are beyond the scope of this book. For our purposes, "unstructured data" always refers to text.

Think of the valuable information that comes to us through the open-ended questions of a survey, email, or post on a social network.

By contrast, "structured data" includes things like salary, date of hire, commute distance, sex, marital status, or the ratings our colleagues give us. It refers to highly organized information that can be reflected in rows and columns as in a relational database or an Excel spreadsheet.

It's no secret that in this book we'll touch on both structured data and text analytics. The subtitle should make it fairly obvious.

What Will You Find in This Book?

People Analytics is divided into five parts and is a pedagogical tour quite similar to the one we do in the Master's program.

The book starts with general concepts but narrows to practical tasks of what needs to be done with each of our tools in order to solve the analytics problems that are presented.

In the first part of the book (I. An Introduction to People Analytics):

Chapter 1. What Is People Analytics about? I examine the case of Google. I walk you through the history of people analytics, from the beginning of the 20th century to the present day. I confess everything that we owe to marketing in people analytics.

Chapter 2. People Analytics: Why and How. I present the reason for using people analytics and its near-universal methodology.

Chapter 3. Big Data and People Analytics. I explain what big data really is and where it is applied in people analytics, beyond what we are told in the media. I make a small foray into the area of ethics.

In the second part of the book (II. Statistics, Errors, and Biases):

Chapter 4. Intuition vs. Data Analysis. I warn of the danger of cognitive biases in decision-making.

Chapter 5. Defense Against the Dark Arts: A Guide to Not Being Fooled by Data. I reveal how data can be deceptive. It includes the story of how the Bill and Melinda Gates Foundation squandered \$1.7 billion because it did not consider the Law of Small Numbers. It ends with a sensible guide to cleaning data.

In the third part (III. Strategy and Economics):

Chapter 6. Building a Business Case for Human Resources. A business case is used to justify an investment in a project. I'll teach you how to build one.

Chapter 7. Taking the Right Measurements. I go over the most frequently used metrics in people analytics, adding a guide to show you how to calculate the cost of attrition.

Chapter 8. Lifetime Value, the Gold Standard. I'll show you how to measure the net contribution that an employee makes to an organization while working there.

Chapter 9. Employee Experience, Engagement, and the Bottom Line. Results of engagement and climate surveys must be adapted according to changes in the bottom line.

Chapter 10. Performance and Compensations. I show you how to use data to evaluate performance. This will allow you to link incentives closer to the behaviors you want to reward.

In the fourth part (IV. Learning to Work with Real Cases of People Analytics):

I provide real, practical HR cases that work with structured data.

Chapter 11. Clear and Simple Algorithms for People Management. I explain the importance of understanding the real causes of a problem and of learning how to anticipate the direct and indirect consequences of a performance improvement action.

Chapter 12. Surveys and the Lingering Doubters. I demonstrate how science also needs urgent repairs in order to get reliable data. I teach you to use the "Cronbach's alpha" and the "Inter-Class Correlation."

Chapter 13. Segmentations: Divide and Conquer. Clustering helps to segment by actual behaviors and not only by socio-demographic variables.

Chapter 14. Predictive Selection. I'll teach you how to predict which candidate will be high-performing, effective, loyal, and aligned with the corporate culture.

Chapter 15. Turnover. I demonstrate how to use analytics to understand the causes that generate talent leakage in the organization and to build a predictive model to be able to anticipate people leaving.

In the fifth part (V. Text Analytics for Human Resources):

Chapter 16. Open-Text Analytics and the Voice of the Employee. The feedback that employees provide is extremely valuable for understanding strengths and weaknesses, and above all, areas where we can make a difference.

Table of Instructions

Now that you're familiar with the contents, I will suggest several different ways to read this book just like [Julio Cortázar](#)⁵ does in the "Table of Instructions" of [Hopscotch](#)⁶. My first suggestion is reading the book in its current form (starting with Chapter 1, followed by 2, and so on) until reaching Chapter 16. It wouldn't be bad to read it like that. But the truth is, I didn't even write it that way.

I bet some people will only read parts I, II, and III, mostly because the chapters in which I handle real cases and show the process of solving them seem somewhat more difficult. However, my advice is: if there are parts that are over your head, skip them for now, since the rest of the contents of those chapters in parts IV and V do not depend on those real examples and provide you with valuable information.

Finally, you can also choose between parts III, IV, and V, according to the professional area that interests you most. For example, you could go directly to the chapter on turnover (Chapter 15) if you want to learn about those tasks. Later on, you could read the chapter on predictive selection (chapter 14) and so on.

"So far so good", some of you might think, but...

⁵https://en.wikipedia.org/wiki/Julio_Cort%C3%A1zar

⁶[https://en.wikipedia.org/wiki/Hopscotch_\(Cort%C3%A1zar_novel\)](https://en.wikipedia.org/wiki/Hopscotch_(Cort%C3%A1zar_novel))

Without further ado, I want you to know more about me and my flaws.

My longest lasting relationship's name is Viki. We met on October 10, 2003. Our first evening together I did something that saved our relationship dozens of times. Even though it seemed like I was oversharing considering it was the very first day we met, I confessed, "Viki, I want you to know something about me: I'm terribly absent-minded. Definitely much more than anyone else you've ever known."

We were reminded of this simple confession every time I lost the keys, my wallet, or my suitcase; I forgot to close the fridge door or to lock the front door; I left home with something on the stove or without my ID when I really needed it; or I put unleaded gas in my diesel car. After the initial annoyance cooled off, she always said something like, "Good thing you warned me!"

Since I've received such positive results from making that confession early, I'll start by telling you a few things about me in hopes that it may save our relationship later. We'll see if this brutally honest start helps me when you find out the stark reality.

1. We all have a past.

Allow me to take the floor and confess to the support group, "My name is Eddy, and I was a linguist." That is to say, in college I didn't major in statistics, mathematics, or even programming. I didn't even study math in the last two years of high school. After finishing my undergrad, I worked for about ten years as an editor of the *General Basque Dictionary* of the Royal Academy of the Basque Language. Not very useful for data science! Until 1995, I lived among literature and languages. Also, more as a hobby than for work, I have been and still am very fond of history, philosophy, and film.

I usually define myself as an "optimistic simpleton." I can't help but see the positive side of almost everything, a little bit like Forrest Gump, but plumper and with no known tendency to run. However, as an optimistic simpleton, I've turned this disadvantage around.

This first flaw of my past (being a linguist) can be interpreted as a story of self-improvement. If a slightly overweight humanities major has been able to learn enough statistics, mathematics, and programming to call himself a data scientist, teach others what he has learned, and even help many organizations earn money by analyzing data, there is certainly hope.

In the world of human resources, I find that many people committed the same original sin, having studied humanities instead of sciences. Others who have studied for more scientific careers such as psychology, sociology, or economics leave statistics at the door when they enter the HR department. Am I wrong? So, if you're concerned that your guide along this transformative path to improving your company with data analysis can't put himself in your shoes, rest assured that I've been there and done that.

And finally, as you'll see later, my linguistics background comes in handy in the area of text analytics, which is becoming more and more valuable.

2. Extravagant connections

The long-term effects of having studied humanities can still be seen in my unstoppable tendency to connect historical facts, languages, philosophy, film, personal anecdotes, or popular culture with the realities that I have to analyze or write about in the field of people analytics.

What's more, I actually believe that I see the relationship between these very disparate worlds, and I go out of my way to explain the connections as best I can. That's why, in addition to data,

algorithms, and human resources, I talk about Descartes, Socrates, baseball, politics, films, shows, and things that are further off topic.

I'm afraid that, as I confessed, this process is unrestrainable. Although I am aware of my excesses, I don't want to stop making the same mistakes. I'm so naive as an optimistic simpleton that I imagine that there is a hidden and silent minority who enjoys, like I do, both the practicality of the analytics culture and its scientific approach as well as a peculiar sense of humor, historical references, philosophy, and film. Viki says I always have my head in the clouds, and she's certainly right, because she knows me as few others do.

For those of you who like to get to the point and can't stand the musings of a sleep-deprived humanities nerd, know that I always (or almost always) let you know when I'm about to make a peculiar connection or eccentric digression. The section titles will also help to identify upcoming zaniness and can save you from a questionably useful paragraph. For example, titles like "**Descartes Explains the Reason for the X Files**," "**The True Story of the Master Who Taught Me the Great Lesson of What for**," or "**Digression On Questions, Philosophy, Obstetrics, and Criticism**," don't fare well for those who hate these kinds of detours.

Underneath all this digression, there is surely a somewhat immature desire to have fun with what I do, to laugh, and to make others smile. I'm not exactly sure; maybe I'll have it analyzed.

3. Business, Business, Business!

This probably doesn't fit the profile that you've made of me so far, but I do analyze and explain analytics concepts to organizations with a tremendously practical mindset, oriented entirely to their benefit.

Since 1995, when I left the Dictionary, I have been an entrepreneur and CEO for a better part of these 20 plus years. I know what I'm talking about. I understand the importance of spending resources

wisely, because for years I also had the responsibility of writing paychecks at the end of the month.

In Chapter 2, I share my experience with a world divided into two camps:

1. **The humanists:** Human resources is there to make employees happy. They're the majority. They believe that their activity is aimed at improving the well-being and happiness of the people within an organization.
2. **The pragmatists or cynics:** Human Resources is there, like the rest of the departments of the company, to increase value for shareholders, or, in a nutshell, to make the company more money. I admit that I have the membership card of this group of soulless cynics.

But don't oversimplify. The story doesn't end in this seemingly irreconcilable dialectical confrontation. In **Chapter 2. Why People Analytics? And How?**, I lay out a proposal to synthesize these two opposing positions.

Vampirizing *Data Coaching*

As I said, two years before releasing this book you have just started to read, I published in Spanish *Data Coaching. A little book about data and profit*.

"Don't kid yourself," I said in *Data Coaching*, "having data does not increase profits."

Because believing that collecting data is enough to make money is like believing that buying a scale, weighing ourselves daily, and keeping a record of our weight will be enough to lose weight



I've brought back several nuggets from that little book to this thicker one.

Judging by the number of downloads of **Data Coaching**, it is rather unlikely that you will be affected by these **Data Coaching** rip-offs in **People Analytics**. That is to say, that even if there are identical paragraphs from the previous book, the chances of you being one of the few who 1) downloaded the book, 2) read it, 3) went back for more despite having endured it, and 4) remember some of what you read are slim to none. To add insult to injury, **Data Coaching** remains in Spanish, untranslated.

Indeed, in entire paragraphs, I haven't bothered with quoting myself.

I've already published a great deal of what you will find here in posts from [my LinkedIn profile](#)⁷.

Both books, along with the blog, share common concepts, passion, and the same playful style that I've become used to. In any case, there are new and original things in this book.

⁷<https://www.linkedin.com/in/eduardovalencia/>

DATA COACHING

UN LIBRITO SOBRE DATOS Y BENEFICIOS

Eduardo Valencia



It's a fun book that you can learn from.

The few who have read *People Analytics* also assure me that although it is close to 400 pages, they read it almost without stopping, because they found it so enjoyable. Some readers confessed to me they skipped the *Do It Yourself* guides. It's okay, I get it.

Illustrations by Marisa Echarri

In all the chapters, you will find my best friend Marisa's illustrations, one of the best things in the book. I know it's unusual to find these kinds of colorful, suggestive, and fun drawings in a technical and practical book like this. But nothing is what it should be in this book. Or so it seems...

THE ABANDONMENT

“Here’s the breakdown of figures
on personnel turnover.”



The Abandonment. Illustration by Marisa Echarri

I. An Introduction to People Analytics

“Love is the most important thing in the world, but baseball is pretty good, too.”

Yogi Berra (18-time All-Star, a Yankees legend)

01. What Is People Analytics About?

Let's start off on the *wrong* foot. This chapter starts with a section on baseball and uses some informal language.



In this Chapter:

1. Lessons learned from *Moneyball*.
2. Where does the name "people analytics" come from?
3. Areas where you can apply people analytics.
4. History of people analytics.
5. What people analytics owes to marketing.

1. About Moneyball

In 2003, *Moneyball: The Art of Winning an Unfair Game* was published. The book, written by Michael Lewis, is based on a true story about the Oakland Athletics baseball team and its general manager, Billy Beane. Through an analytical, evidence-based approach, Billy Beane brings together a competitive baseball team, despite Oakland's disadvantaged income situation. A film based on the book was released in 2011, starring Brad Pitt and Jonah Hill.

The gap between rich and poor in baseball is much greater than in any other professional sport in the United States. In "La Liga," the national Spanish soccer league (Messi, Cristiano Ronaldo, etc.), this gap is ten times greater than in US baseball, by the way. At the start

of the 2002 season, the richest team, the New York Yankees, had a payroll of \$126 million, while the two poorest teams, the Oakland Athletics and the Tampa Bay Devil Rays, had payrolls of less than a third, about \$40 million. The disparities in budgets meant that only rich teams could afford the best players.

However, by reanalyzing the strategies that result in victories in the field (which we would call an "innovative performance analysis"), The Oakland Athletics in 2002, with a budget of \$44 million in salaries, was able to compete with teams with much more resources like the New York Yankees, who had invested \$125 million in payroll that same season.

The Oakland Athletics talent search system still works today. This innovative analytical strategy led the Oakland Athletics to success in 2002 and 2003. The Boston Red Sox, which had won the championship for the last time in 1918, followed this "Moneyball" strategy and won the World Series in 2004.

Therefore, *Moneyball* is based on a true story that tells us about innovation in the evaluation of performance and recruitment, the conflict between intuition, data, and predictive analytics. It's certainly the film that best illustrates the spirit of people analytics.

"Moneyball" has entered the lexicon of baseball and the legend of people analytics. In fact, teams that seem to value analytical concepts are often said to be "playing Moneyball."

Drawing from a scene in the film, I'm going tell you something that I hope doesn't bother you too much:



No Offence!

Some HR departments still have a **medieval way of thinking**, and they are stuck in the Dark Ages of hunches and intuition.

To paraphrase the scene I share on the next page, there is an epidemic failure within these departments to understand what is really going on. And this leads those who run HR departments to waste valuable resources and mismanage people.

Before you angrily toss the book into the recycling bin, or close **Acrobat Reader** or **Kindle** or whatever you are using to read this book, hold on for a few more lines because I have an excuse for the remarks I've made. Later, you can do as you wish.

I took the "medieval way of thinking" from *Moneyball* (the movie). Consider it an homage. On the same day that Peter, the statistician (played by Jonah Hill) and Billy Beane, the manager of the Oakland Athletics (Brad Pitt) meet, they're having a conversation in a garage. Billy Beane has an epiphany in that scene that leads to the miraculous (scientific, rather) transformation of the lowly Oakland Athletics team into one of the greats of the First Division of American baseball.

YouTube. Link to the Moneyball scene. Bill meets Peter⁸.



QR. YouTube. Bill meets Peter

⁸<http://tiny.cc/pa-moneyball>

PETER

Baseball thinking is medieval. It's stuck in the Dark Ages. I have a more scientific view of the game.

BILLY

Keep going, Peter.

PETER

There is an epidemic failure within the game to understand what's really happening. And it leads people who run major league teams to misjudge their players and mismanage their players.



Moneyball. Illustration by Marisa Echarri

2. The Most Valuable Asset is Also the Most Expensive.

In human resources, they often repeat that people are the most valuable asset of the organization.

According to Accenture, the total cost of human capital accounts for up to seventy percent of all business expenses.

Since organizations want to manage this rather expensive asset in the best way possible, they are asking to be able to lean on data in order to make these important decisions for optimization. This request translates into the need to eliminate the conjectures of the talent management cycle and apply analytics to every key process of the HR department: recruitment, training, performance analysis, labor climate, compensation, and loyalty.

2.1. Companies Are Looking for HR Professionals Who Know about Data.

A great opportunity for everyone in human resources: The best companies are desperately looking for people who know how to interpret data and turn it into improvements.

In the area of management positions, a study of the *Intelligence Unit* of *The Economist*, in collaboration with *PwC*, stated that:

"seventy-two percent of managers believe that familiarity with data-based decision-making is an essential requirement for access to management positions."

The number of job offers makes this phenomenon evident. We saw it at the most important conference on HR analytics in the world, the "Wharton People Analytics Conference." Hours before the start of the conference, through the event app, the attendees had already

shared dozens of job offers to work in HR Analytics positions in companies like Facebook, Google, or ESPN.

Thus, learning from people analytics is a great opportunity for HR professionals. Knowledge of the field of human resources is precisely [the main competence⁹](#) sought by organizations for People Analytics positions¹⁰.

3. Tomayto, Tomahto

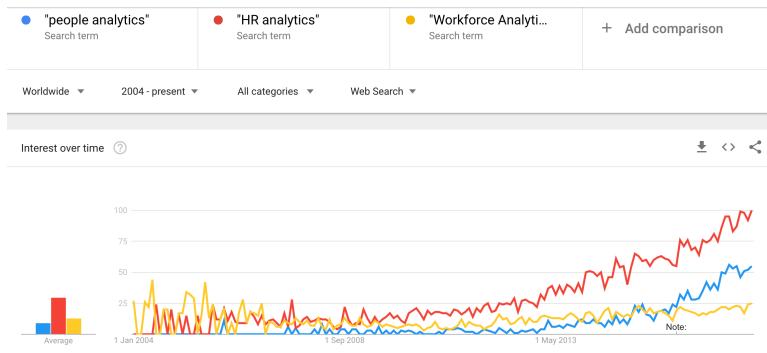
If today we call it "people analytics" instead of "HR Analytics," "Talent Analytics," or "Workforce Analytics," it's Google's fault.

They (Google) rechristened their HR department as "People Operations," and the organization's people-centered analysis is called "people analytics." The name stuck. "People Operations" has also become a popular name for HR departments: Dropbox, Facebook, LinkedIn, and more than twenty companies have adopted the name.

⁹<http://bit.ly/2uPH4uY>

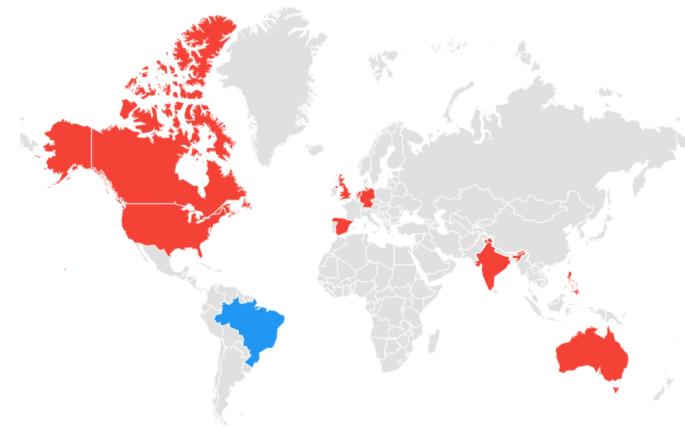
¹⁰<http://bit.ly/2uPH4uY>

3.1. Who Uses What?



Google Trends (September 26, 2017). People analytics, HR analytics, workforce analytics

If you look at the distribution of "people analytics" vs. "HR analytics" in Google Trends (09/26/2017), you will see that "HR analytics" is searched twice as often as "people analytics" and four times more often than "workforce analytics." "HR analytics" is predominant in all major countries, except in Brazil.



Google Trends (September 26, 2017). People analytics (blue) vs. HR analytics (red) by countries

3.2. What's in a Name?

Given the number of names that compete to define this activity, in *The Power of People: Learn How Successful Organizations Use Workforce Analytics to Improve Business Performance*, Nigel Guenole defends the term "workforce analytics" to refer to what we call "people analytics." He starts his chapter by quoting Confucius: "*The beginning of wisdom is to call things by their authentic name.*"

He adds that, without a common language, professionals are at risk of confusion about exactly what the function does.

In my opinion, the name is not as important as the content. As if I were in a Pokémon battle with Confucius, "Shakespeare, I choose you." In *Romeo and Juliet*, the young woman in love says:

What's in a name? that which we call a rose

By any other name would smell as sweet

In "A rose with any other name could smell so sweet," Juliet claims that it does not matter that Romeo is a Montague. In other words, the names of things do not affect what they really are.

4. Big Data for Human Resources

We have more information at our disposal about tastes, opinions, and certain behaviors that could be very useful in making decisions about performance, salaries, turnover rate, leaves of absence, etc.

Each day, employees show multiple signs that measure their performance, degree of satisfaction with their work, and commitment to the organization. Technology has tremendously simplified the measurement and collection of objective parameters on work activities (attendance records or "clocking in and out," systematic employee performance reviews, number of emails sent, time spent in meetings, conference room reservations, shared calendars, and more). It can also measure subjective parameters (employee feedback about the company, its products or services, or work relationships, with different metrics specific to the organization or more standardized ones like the *Employee Net Promoter Score*). Employees could also be sharing very relevant information on a daily basis about how they feel towards the organization through their interactions with co-workers through different channels of internal or external social media networks, which could be tapped into and analyzed.

Accordingly, people analytics processes that information through analysis tools that allow you to look at talent management decisions in a new way. People analytics is based on data analysis as opposed to traditional methods of relationships or personal experiences.

That's why we sometimes say that "**big data for human resources**" is synonymous with people analytics. In Chapter 3, I explain that the use of the word "big data" is reserved for a more limited context.

5. The Path to Transformation

People analytics is based on applying data science techniques to human resources with the aim of getting to know the people of our organization better and increasing their degree of satisfaction and productivity.

Large organizations are so complex, and their operations so remote from each other that it is virtually impossible for management to see some of the vital connections and influences. Current people analytics techniques include developing algorithms to analyze, and statistically balance the connections between internal and external factors, while optimizing processes, delivering services, and measuring performance.

A typical example of how people analytics can intervene is in analyzing and improving turnover rate. A predictive model can identify which employees are at risk of leaving the organization, and even when they will leave, and the reasons why.

Unlike other assets in the company, the value of human capital grows over time. As each employee gains experience, their talent and experience increase. The more experience they have in the company, the greater the loss when they leave.

On average, every employee who leaves an organization costs the company one or two times their annual salary. Turnover analyses reveal some unforeseen key differences between employees who stay and those who leave:

1. Whether they have had a track record of not staying long with any job or position
2. Their demographic, their job position, or even their commute distance
3. Whether they are within or near a period that typically has high turnover (e.g. their first four months or their second or third year in the company)

There are also some signs that can help anticipate employee discouragement and risk of leaving the company, such as absences, coming in late, or declining performance ratings.

From this and other data, you can apply an algorithm that classifies all people according to how likely they are to leave the company. Those who rank higher in the model and are more valuable to the organization will be the first to receive special attention.

5.1. Managing the Entire Cycle with Data

Each company faces unique challenges to meeting its objectives and priorities, but no matter what their circumstance, they can find value in any of the phases of talent management: recruitment, retention, training, commitment, or compensation.

There's a lot at stake. Let's go back to turnover. **Harvard Business Review** claims that 80% of employee turnover is due to poor hiring decisions.

According to Josh Bersin, *"the cost of losing an employee is 1.5 to 2.0 times the person's annual salary. And if we talk about a top executive or highly qualified staff, this cost can rise rapidly to beyond double."*

People analytics offers a unique opportunity for human capital professionals to position themselves as strategic senior management partners, using analytically proven techniques to hire and retain employees who drive higher value in companies.

The main innovation to a more conventional approach is the intensive use of data for assessment and decision-making in the people department: who to hire, who to promote, what the ideal structure of the company is, etc.

Any people analytics solution should support decision-making in human resources. It should help those responsible by integrating

modeling tools and human knowledge, especially in areas of uncertainty or incomplete information to avoid making risky decisions based on human judgment and personal preferences.

5.2 From Metrics to Predictive Analytics

The purpose of this book is to teach you how to overcome this "pre-metric" state of hunches and bias, and to take it a step further. If you are already part of that growing group in which metrics have begun to direct the daily management of human resources, don't let the opportunity pass you by. You can use predictive analytics to understand the real causes and anticipate what is going to happen, or at least what it is most likely to happen.

5.3. A Little History, and a Few More Books

HR departments that assume that their work has value without providing the data to prove it, are somewhat out of tune with the rest of the company.

Other departments (such as marketing or logistics) are able to demonstrate decision-making factors with data and provide greater value in their services. When we cannot provide data, management is obligated to allocate funds that correspond to the findings of the other departments.

Although these human resource departments don't seem to know it, there is a more scientific view of people management that has been going on since the time of Taylor at the beginning of the 20th century.

- In 1911, Frederick Taylor tried to measure worker productivity and improve industrial efficiency. He pioneered scientific management using studies on time and movement. He developed a method to capture and measure the effectiveness of

employees' work, and summarized techniques for efficiency in his book, *The Principles of Scientific Management*¹¹.

- In the 1920s, Elton Mayo was studying the impact of lighting conditions on workers, exposing some workers to higher levels of light than others. Mayo was a social theorist, sociologist, and industrial psychologist specialized in the theory of organizations, human relations, and the human relations movement. When Mayo measured post-intervention productivity, he found that workers were 25% more productive compared to when the experiments began. But surprisingly, this improvement occurred in both of the two groups: those who worked with more light, and **also in those that had less light**. All participants became more productive. What happened? Simply by asking the workers to participate in their experiments, the Mayo team empowered them. The workers found themselves to be an important group whose help and advice was valuable to the company. Mayo introduced a behavioral element into the productivity equation and gave rise to the human relations movement, which proposed that workers would be more productive when their social conditions were met. *The human problems of an industrial civilization*¹².
- In 1978, Jac Fitz-Enz published the first HR metrics. And in 1984, he printed his findings in the book **How to Measure Human Resources Management**¹³, a reference in the field.
- In 2001, Brian E. Becker, in his book **The HR Scorecard**¹⁴, stressed that HR performance indicators show that aligning HR activities with corporate strategy and activity improves the results of the organization.
- It was only in about 2010 that predictive analytics began to appear in the HR departments of some leading companies. In

¹¹<http://tiny.cc/eb6mmy>

¹²<http://tiny.cc/yb6mmy>

¹³<http://tiny.cc/7b6mmy>

¹⁴<http://tiny.cc/tc6mmy>

2014 Jac Fitz-enz published [Predictive Analytics for Human Resources¹⁵](#).

5.4. The Unbearable Fragility of Poorly Developed HR Departments



By Gleilson Miranda/Secretaria de Comunicação do Estado do Acre - CC BY 2.5 br, <http://tiny.cc/yu6mmy>

According to [New Scientist¹⁶](#), in 2013 there were more than 100 isolated tribes worldwide. These isolated tribes lack immunity to common illnesses that can kill a large percentage of their people after contact with other humans.

There are many more HR departments who are just as isolated, still unaware of the basic metrics from the 1970s. Needless to say, they haven't even heard of "predictive analytics."

¹⁵<http://tiny.cc/vc6mmy>

¹⁶<http://bit.ly/2wzSphI>

According to Josh Bersin, only four percent of Fortune 1000 companies were using predictive analysis in 2014. But in doing so, this group's actions surpassed those of their peers of *Standard and Poor's 500* by thirty percent.

And, by the way, that's good news! The number of companies that perform predictive HR analytics has doubled from 2014 to 2015, and again in 2016.

In the midst of highly competitive markets, HR departments that have not integrated analytics and data processes into their business practices are a liability for the overall performance of a company.

In fact, we often witness how upper management pressures HR departments to produce data-based ideas and strategies for critical talent issues such as predicting who will be the best leaders for the future, and how to best prepare them for those leadership positions. It is also essential to know which people are at risk of leaving and what the company is doing to keep them.

5.5. Learning from Marketing



Human resources is following in marketing's footsteps.

In the 1990s, marketing was practically the same as what some HR departments are now: a creative, artistic discipline where decisions were made based on personal experience and intuition.

However, marketing has already integrated analysis and data into its day-to-day management.

Now HR departments of all organizations recognize the need to become strategic units and integrate analytics into daily activity. Just like in marketing, they are already beginning to attract the best employees, group them into teams to be more effective, develop their talents, make them loyal, and reward them using knowledge acquired from people analytics.

While talking about the techniques employed in people analytics (HR analytics), we owe a lot to marketing for having paved the way.

The area of customer analytics can probably boast of the most well-developed analytics procedures to date.

Today, in both online and offline marketing, we measure and

analyze everything under the sun. Every sale, every mouse click, and every conversion on a website are logged. You can predict the value that a customer can bring to your company, the risk of losing each client, and the likelihood that a segment of your customers will buy a product if you launch a certain campaign. This chapter includes return on investment, analysis of client loss, reviews, NPS, lifetime value, RFM, clustering, and much more; later chapters explain how to put them into practice.

In human resources, employees are often referred to as "internal clients." The common denominator between marketing and human capital management is analyzing human behavior.

People are the overlap of marketing and HR analytics. It's no wonder we've been copying all the significant advances in the field.

5.5.1. Marketing Has Been There and Done That.

Marketing has undergone a fundamental change by integrating analytics into business practices.

The old "spray and pray" method of acquiring and retaining customers quickly became obsolete. The best companies gradually began to put aside those old instinct-driven practices and began using analytics.

Today in most successful companies, marketing is inseparable from analytics and constantly leverages the power of data intelligence.

5.5.2. Return on Investment

HR leaders, not just the CEO and CFO, need to think in terms of return on investment (ROI) and profit. An essential marketing ratio is the ROI — the profit you make from an investment. In purely economic terms, it is a way of considering gains in relation to the capital invested. The human capital of a company also needs investment.

A high ROI in human capital means that the money invested in the people of the organization has generated high profits for the company.

5.5.3. Business Case

A "business case" is a document that summarizes the key aspects of a transformative project, and is often used to justify investment in a project. Other departments (such as marketing) are able to demonstrate with data that they provide greater value with their services. If the HR department cannot provide the data that demonstrates the positive economic impact of their proposed investments, management is obligated to allocate funds to other departments that are able to do so. Therefore, it is critical to present "business cases" to justify investments.

In Chapter 6 we delve into business cases.

5.5.4. Recommendation Systems... of Careers

Recommender systems have become more and more popular in different areas of marketing. There are recommender platforms for films, music, news, and all kinds of products and services. One of the better-known ones are the shopping cart recommendation systems for e-commerce sites. In people analytics, we use recommendation algorithms to reveal the most relevant features for each position in an organization and generate automatic recommendations for career development.

5.5.5. Simulations "What if...?"

A "What if?" simulation is a powerful business analytics tool that assesses how strategic, tactical, or operational changes can impact the business. Through different scenarios, you can perform a real analysis of the processes before putting them into practice. People analytics can answer questions like:

- *How would the manufacturing time be reduced if we doubled the number of people?*
- *What would the cost/benefit ratio be if we reduced the process time of a specific activity?*
- *How would altering shifts affect operational costs and service levels?*

You can take a look at some simulations in chapter 10.

5.5.6. Employee Net Promoter Score

For many companies, the Net Promoter Score (NPS) has become the fundamental metric to quantify customer satisfaction. Along that same line, the mission of having engaged and happy employees also has its metric: the Employee Net Promoter Score (eNPS).

Companies that have achieved substantial improvements in customer experience through the NPS have already begun to apply the same methodology to measure and improve the employee experience. Although fewer studies have been made, they are finding that the eNPS has the same degree of correlation between the metric and company earnings as what they discovered when analyzing the data of the NPS (for customers). I cover it in more detail in Chapter 9.

5.5.7. Employee Lifetime Value

Employee lifetime value (ELTV) is a prediction of the net profit attributed to an employee through the permanence of that person in a given position. Once again, this metric is taken from marketing. In an ELTV, there are usually three metrics: cost, performance, and job abandonment. All three are combined to help you make decisions. It's covered in more detail in Chapter 8.

5.5.8. Survival and Churn Analysis

Survival analysis, which originated in healthcare but is now widely used in marketing, is generally defined as a set of methods to analyze the time until an event happens. The event may be the death of a patient or the loss of an employee. All have one principle in common: There is a common pattern of deterioration that affects any population. In any single year of your life, there is an associated probability that an event will happen. Likewise, in an organization, staying longer than a month is something associated with over 90% of employees. But staying longer than ten years in the same company, only 10%. Survival analysis helps to more accurately segment groups of employees that are above or below an organization's average survival ratios.

I go over survival in Chapter 15.

5.5.9. RFM Analysis

RFM analysis is a marketing technique used to segment customers using three fundamental metrics: recency, frequency, and monetary value. These three metrics are the most important factors in customer segmentation.

The RFM technique is widely used in the field of marketing. It has not traditionally been applied in any other field. In people analytics, we have borrowed this technique for the field of commercial segmentation.

A good part of Chapter 13 is dedicated to RFM analysis.

5.5.10. Segmentation with Clustering

Just like with customers, we can group employees into similar categories to help define loyalty actions that best suit their specific needs. Clustering allows us to segment by real behaviors, based

on interests, motivations, knowledge, and attitudes, and not by sociodemographic criteria.

By analyzing the available employee data (performance evaluations, previous experience, surveys, assignments, etc.) clustering algorithms find associations that allow us to group the employees into segments of people who share similar characteristics.

Chapter 13 deals with clustering, too.

5.5.11. Predictive Models for Selection

Several predictive algorithms that have had a long history in marketing are used to help identify the particular attributes of the candidates in order to improve the selection of people in an organization. For example, the variables associated with competencies, learning ability, or motivation contribute to the prediction of which employees will have high levels of performance and loyalty.

In **Chapter 14: Predictive Selection**, I teach you how to predict which candidates will be high-performing, productive, and loyal employees aligned with the corporate culture.

5.5.12. Predictive Models for Job Abandonment

When a valuable person leaves your company, there is always a loss involved, both economic (the cost of hiring another person and providing all the necessary training that is specific to your organization) and emotional, because of all the ties that that person had created with other co-workers.

Customer loss analysis has been used reliably for marketing. In people analytics, we have successfully implemented a job abandonment analysis that predicts the potential loss of an employee and the causes that led to it in time to take preventative action.



Takeaways:

1. There is a great opportunity for everyone in human resources: The best companies are desperately looking for people who know how to interpret data and turn it into improvements.
2. If today we call it "people analytics" instead of "HR analytics," "Talent Analytics," or "Workforce Analytics," it's Google's fault.
3. HR departments that assume that their work has value without providing the data to prove it, are somewhat out of tune with the rest of the company. Although these HR departments don't seem to know it, there is a more scientific view of people management that has been going on since the time of Taylor at the beginning of the 20th century.
4. In order to attract, segment, grow, retain, and reward customers, marketing took full advantage of analytics to fully understand its customer base and optimize the whole customer life cycle and customer relationship management. That's how marketing analytics was born. Human resources is now going the same route. Many of the techniques that were successfully used over the nearly 30 years of this transformation in marketing are now being implemented in human resources with the same transformative results.

Recommended Reading

1. Human Capital Management (HCM) Part 2 of 4: [Controlling Workforce Costs Using Total Cost of Workforce \(TCOW\)](http://tiny.cc/836mmy).¹⁷
2. HBR: It's Better to Avoid a Toxic Employee than Hire a Superstar.¹⁸
3. [Toxic Workers](http://hbs.me/2vQX7G7).¹⁹

¹⁷<http://tiny.cc/836mmy>

¹⁸<http://bit.ly/2vRiV4y>

¹⁹<http://hbs.me/2vQX7G7>