

Making Everything Easier!<sup>TM</sup>

Alteryx Special Edition

# Big Data Analytics

FOR  
**DUMMIES**<sup>®</sup>

A Wiley Brand

## Learn to:

- Understand what Big Data is and why it is important
- Maximize value from Big Data using analytic applications
- Improve Big Data decision making using analytic platforms and predictive analysis

Brought to you by

alteryx

Michael Wessler, OCP & CISSP



# About Alteryx

Alteryx provides an indispensable and easy-to-use analytics platform for enterprise companies making critical decisions that drive their business strategy and growth. Alteryx Strategic Analytics runs analytic applications that empower executives to identify and seize market opportunities, outsmart their competitors, increase customer loyalty, and drive more revenue. It humanizes Big Data by enabling business analysts and Data Artisans to combine Big Data with market knowledge, location insight, and business intelligence; easily perform predictive and spatial analytics; and produce analytic apps that can be shared via the private cloud or the Alteryx Analytics Gallery public cloud. Customers like Experian Marketing Services and McDonald's rely on Alteryx daily. Headquartered in Irvine, California, and with offices in Boulder and Silicon Valley, Alteryx empowers 250+ customers and 200,000+ users worldwide.

Alteryx, the Leading Platform for Strategic Analytics™

For more information or a free download visit:  
[www.alteryx.com](http://www.alteryx.com).

*Big Data*  
*Analytics*  
FOR  
**DUMMIES**

A Wiley Brand

ALTERYX SPECIAL EDITION

**by Michael Wessler, OCP & CISSP**

**WILEY**

These materials are the copyright of John Wiley & Sons, Inc. and any dissemination, distribution, or unauthorized use is strictly prohibited.

## **Big Data Analytics For Dummies®, Alteryx Special Edition**

Published by  
**John Wiley & Sons, Inc.**  
111 River St.  
Hoboken, NJ 07030-5774  
[www.wiley.com](http://www.wiley.com)

Copyright © 2013 by John Wiley & Sons, Inc., Hoboken, New Jersey

Published by John Wiley & Sons, Inc., Hoboken, New Jersey

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permissions>.

**Trademarks:** Wiley, the Wiley logo, For Dummies, the Dummies Man logo, A Reference for the Rest of Us!, The Dummies Way, Dummies.com, Making Everything Easier, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. Alteryx is a registered trademark of Alteryx, Inc. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

**LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER AND AUTHOR ARE NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.**

For general information on our other products and services, please contact our Business Development Department in the U.S. at 317-572-3205. For details on how to create a custom *For Dummies* book for your business or organization, contact [info@dummies.biz](mailto:info@dummies.biz). For information about licensing the *For Dummies* brand for products or services, contact [BrandedRights&Licenses@Wiley.com](mailto:BrandedRights&Licenses@Wiley.com).

ISBN 978-1-118-60704-6 (pbk); ISBN 978-1-118-60890-6 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

# Table of Contents

---

<b>Introduction .....</b>	<b>1</b>
About This Book .....	1
Foolish Assumptions .....	2
How This Book Is Organized .....	2
Icons Used in This Book.....	3
Where to Go from Here .....	4
<b>Chapter 1: Understanding the Big Data Landscape .....</b>	<b>5</b>
What Big Data Is .....	5
How Data Is Changing.....	6
Shift in Processing Due to Big Data .....	8
Big Data Is Everywhere .....	9
<b>Chapter 2: Getting Started with Big Data Analytics... .</b>	<b>11</b>
Changing Focus with Big Data.....	11
The Role of the Data Analyst .....	12
Implementing Big Data Analytics within	
an Organization Using Alteryx.....	13
Blending Data from Multiple Sources.....	14
Looking at Alteryx Designer Desktop .....	15
<b>Chapter 3: Analyzing Big Data in Context. . . . .</b>	<b>17</b>
Focus on Context, Not Just Integration.....	17
Combining Big Data with Spatial Data.....	18
Leveraging External Data Provider Resources.....	19
<b>Chapter 4: Getting Value from Predictive</b>	
<b>Analytics and Big Data .....</b>	<b>21</b>
Why Do Predictive Analytics on Big Data? .....	21
Moving Predictive Analytics to the Front Lines.....	23
Gaining Real Business Value from Predictive Analysis.....	24
<b>Chapter 5: Humanizing Big Data Analytics..... .</b>	<b>25</b>
Putting Big Data in the Hands of Those Who Need It.....	26
Humanizing Data Design Principles.....	27
Humanizing Big Data Analytics Workflow.....	28
Considering Consumerization of Big Data Analytics.....	31
Getting an Alteryx Analytics Gallery Overview .....	32
Publishing Data and Analytics to Cloud Service.....	33
Focusing on Consuming Applications .....	34
The Best Platform for Strategic Analytics.....	35
<b>Chapter 6: Ten (Okay, Nine) Things to Consider</b>	
<b>with Big Data Analytics..... .</b>	<b>37</b>

## Publisher's Acknowledgments

We're proud of this book and of the people who worked on it. For details on how to create a custom For Dummies book for your business or organization, contact [info@dummies.biz](mailto:info@dummies.biz). For details on licensing the For Dummies brand for products or services, contact [BrandedRights&Licenses@Wiley.com](mailto:BrandedRights&Licenses@Wiley.com).

Some of the people who helped bring this book to market include the following:

### ***Acquisitions, Editorial, and Vertical Websites***

**Senior Project Editor:** Zoë Wykes

**Editorial Manager:** Rev Mengle

**Acquisitions Editor:** Amy Fandrei

**Business Development Representative:**  
Kimberley Schumacker

**Custom Publishing Project Specialist:**  
Michael Sullivan

### ***Composition Services***

**Senior Project Coordinator:** Kristie Rees

**Layout and Graphics:** Jennifer Henry

**Proofreader:** Dwight Ramsey

**Alteryx Contributors:** Dipesh Patel,  
Paul Ross, Rick Schultz

---

### **Publishing and Editorial for Technology Dummies**

**Richard Swadley**, Vice President and Executive Group Publisher

**Andy Cummings**, Vice President and Publisher

**Mary Bednarek**, Executive Director, Acquisitions

**Mary C. Corder**, Editorial Director

### **Publishing and Editorial for Consumer Dummies**

**Kathleen Nebenhaus**, Vice President and Executive Publisher

### **Composition Services**

**Debbie Stailey**, Director of Composition Services

### **Business Development**

**Lisa Coleman**, Director, New Market and Brand Development

# Introduction



**B**ig Data is changing how businesses do business. Data is growing at an explosive rate, coming into the company from different areas, and in myriad formats. Social media, sensor data, spatial coordinates, and external data resource providers are just some of the new data vectors companies must now address. The result is that existing analytic and Business Intelligence (BI) practices must be rethought in the context of Big Data.

Yet, despite these challenges, Big Data offers great opportunities. Powerful analytic platforms such as Alteryx allow data analysts to rapidly build and deploy analytic applications to business decision makers. Alteryx Designer Desktop and Alteryx Analytics Gallery are among the fastest ways to gain insight into Big Data. Together they provide context with internal data resulting in perspectives and vision that would otherwise not be available. The end result is better data being used to make better business decisions to take advantage of business opportunities.

## *About This Book*

Big Data is changing how we manage data and how we use it in our businesses. Big Data comes in many forms, and from new sources such as mobile devices (smart phones for example), scientific sensors, and the cloud, and it's coming at fire hose speed. Smart companies realize that the rules of data are changing, and they need to improve how they manage Big Data to remain relevant and competitive in the marketplace.

The focus of this book is how data analysts can use powerful analytic tools to take advantage of Big Data and create powerful analytic applications for rapid deployment to business decision makers.

## *Foolish Assumptions*

It's been said that most assumptions have outlived their usefulness, but I'll assume a few things nonetheless! Mainly, I assume that you know a little something about Business Intelligence and analytics and want to improve your business decision making by using data in a smarter way. As such, this book is written primarily for those who understand basic IT principles and have heard of Big Data but want to find out whether using analytic processing tools with Big Data can help them make better, more informed business decisions.

## *How This Book Is Organized*

This book consists of six conveniently distilled chapters filled with just the information you need. Here's a brief look at what unfolds!

### *Chapter 1: Understanding the Big Data Landscape*

The book begins with an overview of what Big Data is and explains why it's such a hot topic for businesses trying to make the most of their data. You see how Big Data is transforming analytic processing and what makes it a natural fit for cloud architecture.

### *Chapter 2: Getting Started with Big Data Analytics*

This chapter details the role of the data analyst and explains why this person is the most important person working with Big Data. It also explores the Alteryx Designer Desktop that you can use to quickly build and deploy powerful analytic applications.

## *Chapter 3: Analyzing Big Data in Context*

This chapter delves into Big Data in the context of the business and internal data sources to make the best possible decisions. You see how spatial data and external data resources help to identify exciting business opportunities.

## *Chapter 4: Getting Value from Predictive Analytics and Big Data*

Chapter 4 explores what predictive analytics is and how it lends itself to getting real value out of Big Data for businesses. You get a look at the predictive analytics tools within the Alteryx Designer Desktop.

## *Chapter 5: Humanizing Big Data*

Here, I talk about humanizing Big Data and why it is important. You find out how to put Big Data in the hands of those who need it with tools such as Alteryx Analytics Gallery.

## *Chapter 6: Ten (Okay, Nine) Things to Consider with Big Data Analytics*

The classic endpoint in every *For Dummies* book is the famous Part of Tens chapter. This chapter covers nine items that you'll want to know as you explode into the exciting world of Big Data Analytics!

## *Icons Used in This Book*

Throughout this book, you occasionally see special icons that call attention to important information. You won't find smiley faces or any other cute little emoticons, but you'll definitely want to take note! Here's what you can expect:



This icon points out things you'll be glad I mentioned later on. This is the stuff you want to remember when you start using the material on your own.



I try to keep the techie stuff to a minimum, but I am a techie person at heart and old habits die hard! These are technical tidbits that aren't essential, but they are nice to know.



This icon points out pieces of sage wisdom that I wish someone had told me when I was learning this subject.



Learning "the hard way" makes for good experience and sometimes funny stories, but it also sometimes hurts. Take heed in these warnings, and you may just avoid making the mistakes this book talks about in the first place!

## Where to Go from Here

Someone once said, "Having lost sight of our objective, we will redouble our efforts." How often have you seen that mentality at work, usually by a frustrated manager after an embarrassing mistake? People promise to work harder and smarter, but they still aren't really quite sure what they are doing or why.

Not understanding where you are going, what you want to do, or how to get there is fun for a carefree vacation, but it's no way to approach anything that you want to be successful.

That's certainly true if you are trying to learn a new paradigm such as Big Data, but fortunately with *Big Data Analytics For Dummies*, Alteryx Special Edition, you have help to guide you on this exciting journey!

If you don't know where you're going, any chapter will suffice — but Chapter 1 might be a good place to start! However, if you see a particular topic that interests you, feel free to jump ahead to that chapter. Each chapter is written to stand on its own, so feel free to start reading anywhere or to skip around! Read this book in any order that suits you (although I don't recommend upside down or backwards). I promise you'll put the book down thinking, "Wow, I didn't know this stuff could be so easy!"

## Chapter 1

# Understanding the Big Data Landscape

### *In This Chapter*

- ▶ Understanding what Big Data is and why it is important
- ▶ Looking at how data is changing at such an incredible rate
- ▶ Identifying the paradigm shift in analytic processing
- ▶ Gaining insights into cloud computing and the impact on Big Data

**B**ig Data is important if you want to be successful in analytic processing. But, why is that important? The answer is that success in a highly competitive, fast-moving marketplace is determined by who can capitalize on business opportunities before everyone else seizes the same opportunity. The way to be agile (and therefore successful) is to spot trends, opportunities, and risks via analytic processing of data, and in modern times, with Big Data. Thus, if you want to be successful, you must understand Big Data and how to quickly extract from it the business critical information that your business requires.

This chapter looks at what Big Data is and how the overall data landscape is changing.

## *What Big Data Is*

Many people believe Big Data is simply a large amount of data, but it is defined by more than just size. Leading IT industry research group Gartner defines Big Data as:

*Big Data are high-volume, high-velocity, and/or high-variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization.*

Data is described within the Gartner definition (and within the IT industry) based on the three Vs:

- ✓ **Volume:** Size of data (how big it is)
- ✓ **Velocity:** How fast data is being generated
- ✓ **Variety:** Variation of data types to include source, format, and structure

In terms of the three Vs, the Gartner definition effectively says that:

*There is a lot of data, it is coming into the system rapidly, and it comes from many different sources in many different formats.*

The definition may seem vague given that it is describing a technical item, but to accurately capture the scope of Big Data the definition itself must be “big.”



IT companies are investing billions of dollars into research and development for Big Data, Business Intelligence (BI), data mining, and analytic processing technologies. This fact underscores the importance of accessing and making sense of Big Data in a fast, agile manner.

Big Data is important; those who can harness Big Data will have the edge in critical decision making. Companies utilizing advanced analytics platforms to gain real value from Big Data will grow faster than their competitors and seize new opportunities.

## How Data Is Changing

It is not a secret that data is changing in both quantity (volume) and format (variety). Explosive growth (velocity) is the most obvious example of data change as evidenced by these two statistics:

- ✓ IBM estimates 2.5 quintillion bytes of data are generated each day.
- ✓ Ninety percent of the data in the world is less than two years old.

Traditional, corporate data internal to the organization is being overwhelmed by a new generation of data external to the datacenter. Reasons for the data explosion are largely due to new technologies generating and collecting vast amounts of data. These sources include

- ✓ Scientific sensors such as global mapping, meteorological tracking, medical imaging, and DNA research
- ✓ Point of Sale (POS) tracking and inventory control systems
- ✓ Social media such as Facebook posts and Twitter Tweets
- ✓ Internet and intranet websites across the world

Explosive data growth by itself, however, does not accurately describe how data is changing; the format and structure of data are changing. Rather than being neatly formatted, cleaned, and normalized data in a corporate database, the data is coming in as raw, unstructured text via Twitter Tweets on smart phones, spatial data from tracking devices, Radio Frequency Identification (RFID) devices, and audio and image files updated via smart devices.

Much of the data generated by new technology is unstructured or in a semi-structured data format that makes it more difficult to manage and process. Furthermore, while a social media post may be relatively small text, data related to images and audio input is very large. The increased size of unstructured data and increased complexity managing to it is a difficult task requiring enhancements of how data is managed and analyzed.



NASA reportedly has accumulated so much data from space probes, generating such a data backlog, that scientists are having difficulty processing and analyzing data before the storage media it resides on physically degrades.

## *Shift in Processing Due to Big Data*

Traditionally, large datasets would reside on a corporate mainframe or in a data warehouse in a well-defined format, often managed by an advanced Relational Database Management System (RDBMS). This is a tried-and-true configuration, but it does not reflect the changing nature of Big Data. As the data has changed, so must how it is processed.

Traditional BI tools that rely exclusively on well-defined data warehouses are no longer sufficient. A well-established RDBMS does not effectively manage large datasets containing unstructured and semi-structured formats. To support Big Data, modern analytic processing tools must

- ✓ Shift away from traditional, rearward-looking BI tools and platforms to more forward-thinking analytic platforms.
- ✓ Support a data environment that is less focused on integrating with only traditional, corporate data warehouses and more focused on easy integration with external sources.
- ✓ Support a mix of structured, semi-structured, and unstructured data without complex, time-consuming IT engineering efforts.
- ✓ Process data quickly and efficiently to return answers before the business opportunity is lost.
- ✓ Present the business user with an interface that doesn't require extensive IT knowledge to operate.

Fortunately, IT vendors and the IT open source community are stepping up to the challenge of Big Data and have created tools that meet these requirements. Popular software tools include

- ✓ **Hadoop:** Open-source software from Apache Software Foundation to store and process large nonrelational data sets via a large, scalable distributed model. Commercialized Hadoop distributions are available from companies such as Hortonworks and Cloudera.

- ✓ **NoSQL:** A class of database systems that are optimized to process large unstructured and semi-structured data sets. Commercialized NoSQL distributions are available from companies such as 10gen and Couchbase.

One platform in particular that is discussed in this book is the Alteryx Strategic Analytics platform. Alteryx specializes in Big Data Analytics with an emphasis on bringing Big Data within reach of the people who can best use it.

## *Big Data Is Everywhere*

Big Data comes from multiple sources — often from technologies that until recently did not exist. Increasingly, Big Data is coming from handheld smart devices, complex scientific sensors, and retail, inventory, and sales tracking devices.

Big Data also resides in multiple locations. Gone are the days where data only exists in a chilly datacenter behind the locked doors of the IT department. Today the most valuable data is outside the company where it is hosted by external entities and shared (or purchased) by those wise enough to seek it.

The paradigm has shifted from storing all data and managing all of IT exclusively in-house to a more open (yet secure) model. This model is called *cloud computing*, which references the idea that the end user simply accesses their data or applications from the “cloud” without concern to where the IT resources physically reside.

With all the data, applications, and resources in the cloud, it is shared by all and accessible anywhere and anytime (with proper security). Furthermore, the cloud is a limitless computing environment where size and capacity issues do not exist. High Availability (HA) is provided by redundancy of cloud components; if one component fails, another takes its place.

Cloud computing can be less expensive for a company as well. Instead of purchasing and supporting a complex IT infrastructure in-house, companies pay for only the resources they use within the shared cloud. The metered service feature allows

companies to pay for what they need and actually use, not overhead they don't need or want.

Clouds are classified based on their deployment model, which relates to who has access to the cloud and its resources:

- ✓ Private clouds are exclusive to a specific organization; the public is not invited. This is the most secure form of cloud computing.
- ✓ Community clouds are restricted to departments within a company or agency, multiple government agencies, or a group of companies within a specific industry.
- ✓ Public clouds are exposed to everyone. However, security features are in place. These clouds offer maximum flexibility for the services offered and accessed by cloud consumers.

The impact of cloud computing on Big Data is huge. Data sources can be from public, private, or community clouds. For example, customer demographic data can come from a public cloud, but complex scientific collection information or industry-sensitive data would be from community clouds. Any Big Data Analytic platform must be able to access any cloud platform and be able to publish results to any cloud environment in a fast and secure manner.

## Chapter 2

---

# Getting Started with Big Data Analytics

---

### *In This Chapter*

- ▶ Redirecting your focus with Big Data
  - ▶ Unlocking value with your data analysts
  - ▶ Implementing Big Data in your company
  - ▶ Using Alteryx Designer Desktop to rapidly build and publish powerful analytic apps
- 

**S**tarting any initiative is the most difficult step; the key is in knowing how to get started. With Big Data, knowing what to focus on is a fundamental first step. Next, you need to utilize your most powerful asset — the data analysts at your company. Finally, you must ensure that Big Data is easy to use by those who need it, and that must be done using a powerful analytic platform.

This chapter shows you how to begin the journey with Big Data Analytics and how to start using it within your company.

## *Changing Focus with Big Data*

As discussed in Chapter 1, the three Vs of data are defined as Volume (size), Velocity (how fast it is generated), and Variety (variation). However, when implementing Big Data in organizations, the three Vs fall short; a fourth V for Value must become the driving focus.

Unlocking the value in data is the key to providing value to the business. Too often IT infrastructure folks focus on data capacity or throughput speed. Business Intelligence vendors extol the benefits of executive-only dashboards and visually stunning graphical reports. While both perspectives have some merit, they only play a limited role in the overall mission of bringing real value to those in the company who need it.

Value is added by using an approach and platform to bring Big Data into the hands of those who need it in a fast, agile manner to answer the right business questions at the right time. Knowing what data is needed to answer questions and where to find it is critical; having the analytic tools to capitalize on that knowledge is even more critical. It is through those platforms that real value is realized from Big Data.

## *The Role of the Data Analyst*

The most powerful data expert in a company likely isn't in the IT department, doesn't hold advanced computer engineering degrees, and likely doesn't hold the title "data scientist." The most powerful data expert is in the business department and understands fundamental IT concepts, but the real knowledge is in the business processes and data that the company relies upon on a daily basis. This person probably has a better idea of what operational data is needed to support the company than the CEO does. This person — the *data analyst* — can be the hero of your Big Data Analytics experience.



In-depth data analysis is as much of an art form as it is a science.

The data analyst knows the business inside and out, but also knows where all the key data is located. Here are some examples of key data:

- ✓ Internal, corporate databases and data warehouses
- ✓ Spreadsheets and documents stored on local workstations and shared network drives
- ✓ External data feeds that the company receives on a daily basis
- ✓ Data that would be valuable to have but is currently not available

The data analyst is the person who Big Data Analytics tools need to empower because it is the data analyst who actually uses data on a daily basis. This person is either making business decisions or providing data to the decision makers. Simply put, if the data analyst is not able to access Big Data, that missing data will not be part of any decision-making process.

Humanizing Big Data has the largest positive impact on data analysts and business decision makers. The concept of *humanizing Big Data* is the ability to combine Big Data with market knowledge, location insight, and business intelligence while performing predictive and spatial analysis to produce analytic applications that are shared with decision makers. Humanizing Big Data is critical to a successful implementation within a company and using the right analytic platform (such as Alteryx) makes that process possible.

## *Implementing Big Data Analytics within an Organization Using Alteryx*

Technology alone doesn't generate real value from Big Data. Data analysts, empowered with the right analytic technology platform, humanize Big Data, which is how companies realize value.

Analytic platforms such as Alteryx make extracting value from Big Data possible. Important benefits to businesses that the Alteryx Strategic Analytics platform provides include

- ✓ Improving the self-sufficiency of decision makers to run and share analytic applications with other data users. Data analysts who understand the business should develop good analytic applications that are shared for everyone's benefit.
- ✓ Injecting Big Data into strategic decisions without waiting months for an IT infrastructure and data project. Alteryx gets the data into the hands of decision makers so that businesses can identify and capitalize on opportunities.

- ✓ Delivering the power of predictive analytics to everyone, not just a few executive decision makers far removed from operations. Ensuring that the right data is readily available to all authorized parties leads to making the best possible decisions.

Alteryx is a powerful platform to humanize data into the hands of data analysts to expose and capitalize on the real value of Big Data.

## *Blending Data from Multiple Sources*

The nature of Big Data is large data, usually from multiple sources. Some data will come from internal sources, but increasing data is coming from outside sources. These outside sources include

- ✓ Social media data feeds such as Twitter and Facebook
- ✓ Point of Sale and customer loyalty tracking programs
- ✓ Government agency sources such as census data
- ✓ Spatial data from mobile devices and satellite mapping feeds
- ✓ Consumer demographic data brokers, such as Experian
- ✓ Any number of public, private, or community clouds

*Data blending* is the process of combining multiple heterogeneous data sources and blending them into a single, usable analytic dataset. The purpose of data blending is to create analytic datasets to answer business questions using data that is not bound by the control and lengthy timelines of traditional IT processes. An example of data blending is when the data analyst integrates packaged, external data from the cloud with internal data sources to create a very business-specific analytic dataset.

Data blending is important, and any Big Data Analytics platform must support this function. With data blending, the complete

scope of Big Data becomes available to the data analyst. One such platform that excels at data blending is Alteryx Designer Desktop.

## *Looking at Alteryx Designer Desktop*

Alteryx Designer Desktop is a powerful yet easy-to-use analytic workflow design environment that allows data analysts to access Big Data and then quickly put together analytic applications. It is the fastest way to build and publish analytic applications. Benefits of Alteryx Designer Desktop include

- ✓ Easy-to-use, single point of entry graphical interface with visual representation of analytic workflows.
- ✓ Menus with prebuilt calculations, joins, utilities, reports, spatial tools, and predictive tools.
- ✓ Access to predefined external data sources such as salesforce.com, as well as integration with internal data sources. Data blending is a key feature.
- ✓ Capability to securely publish analytic applications to Alteryx Analytics Gallery or your own internal private cloud.
- ✓ Ability to download and open preexisting applications to use as templates for your new applications.
- ✓ Robust data analyst support infrastructure for questions or problems.

The interface is easy, and the tools are powerful. Alteryx Designer Desktop is designed to empower the data analyst without the need for advanced IT skill sets or long development timeframes.

Figure 2-1 shows the Alteryx Designer Desktop with a sample dataflow opened.

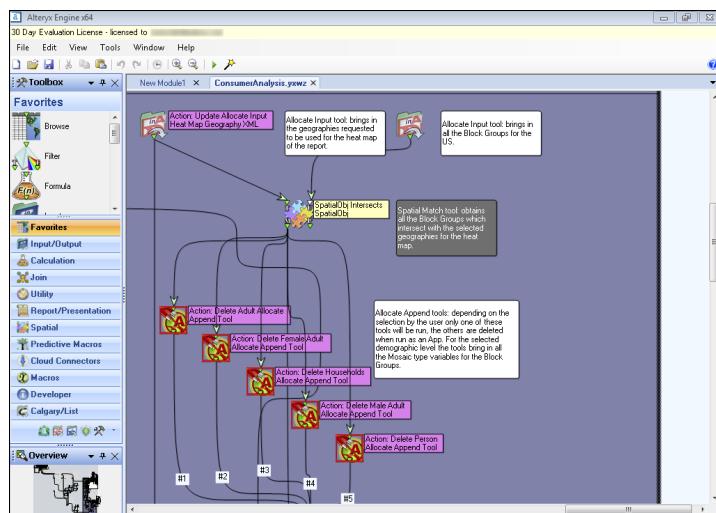


Figure 2-1: Alteryx Designer Desktop opened with a sample application.

How does a data analyst get started with Alteryx Designer Desktop? Here are the steps to take in an example use case. The data analyst:

1. Takes one of the sample analytic applications and opens it to find a visual workflow with data points.
2. Updates the data feeds to use different data sources and potentially incorporates additional tools, filters, or formulas already included by Alteryx.
3. Runs test iterations until the desired data output is generated.

The finished analytic application is then published to the user base that needs it within the company via Alteryx Analytic Gallery or private cloud.

Advanced analytic capabilities are available, but the key benefit is that the difficult work has already been done by Alteryx and is included in the Designer Desktop. Spatial utilities, predictive and calculation tools, utilities, and integration with almost any data source imaginable are available. These capabilities, packaged in a single, easy-to-use interface are what make Alteryx Designer Desktop the fastest way to deliver analytic applications to business users.

# Chapter 3

---

# Analyzing Big Data in Context

---

## *In This Chapter*

---

- ▶ Putting Big Data in context
  - ▶ Capitalizing on spatial attributes present in Big Data
  - ▶ Taking advantage of external data resources
- 

**A**n often overlooked aspect of using Big Data is context. Knowing how your data sources relate to each other is very important. Integrating the spatial attributes of data into the overall context along with valuable, external data resources elevates analytic decision making to a new level.

This chapter discusses how to put data into context to include spatial and external resources and includes a real-world case study example.

## *Focus on Context, Not Just Integration*

Data integration is a big piece of the picture that often receives a great deal of attention, but it is not the most important aspect of working with Big Data. Putting Big Data into context with all your other sources of data is what is important.

Big Data is obviously important (obvious since you are reading this book) but don't forget your internal data sources. Your company's internal databases, corporate data warehouses, and myriad of spreadsheets and documents hold a wealth

of information. The limitless sources of Big Data need to be pared down to what is relevant and be placed in context with your internal data sources.

Using analytic tools such as Alteryx allows you to place your Big Data into context with internal data sources promoting maximum efficiency and agility. Alteryx does this in a simple manner and is driven by the data analysts (refer to Chapter 2).



Decision makers must make decisions based on all relevant data available: Amazon, Experian, TomTom, demographics, spatial attributes, and internal data sources such as spreadsheets and data warehouses.

So where does data integration come into the picture and what are the best ways to integrate? Here are some guidelines:

- ✓ Integrate Big Data with all required data including local data, corporate data stores such as Teradata, and data from cloud applications such as Salesforce.com.
- ✓ Integrate packaged market and customer data from leading data vendors such as Experian and Dun & Bradstreet (D&B).
- ✓ Work to process the data without explicit, lengthy IT integration involvement whenever possible.
- ✓ Allow data analysts to build analytic applications that incorporate the complete context for decision making.

Focusing on context first with integration second allows for a better, more confident decision-making experience without the risk of becoming bogged down in a long integration project.



Use all the relevant data in an agile manner. You need to understand what data you really need and then leverage it quickly.

## Combining Big Data with Spatial Data

*Spatial data* is defined as data that identifies the geographic locations of objects or features on Earth. The widespread deployment of Global Positioning Satellite (GPS) devices has

coincided with the practice of *geotagging* which has been growing at an explosive rate. Geotagging is the process of adding spatial metadata to generated content such as photographs, Short Message Service (SMS) texts, Rich Site Summary (RSS) feeds, and social media posts. This process provides a continual stream of data identifying where people are and what they are doing and is a major part of the growing wave of Big Data. Analytic platforms that understand and process this type of data without additional, specialized technologies give the business user a powerful edge.

What are the benefits of spatial data for companies using Big Data Analytics? Knowing where your customers are, when they are there, and what is near them is all critical intelligence for sales and marketing. However, the complex nature of spatial data and GIS software requires a powerful analytic platform to make sense of that data. The Alteryx Strategic Analytics platform with native spatial capability leverages spatial attributes in Big Data to:

- ✓ Deliver deep spatial and location understanding as part of the unified workflow to answer contemporary business questions more effectively.
- ✓ Provide a more complete understanding of transactions and market and customer interactions.
- ✓ Improve the outcome of critical customer, market, and investment decisions without the requirement for complex and specialized GIS software.

Taking spatial data attributes into context with your other data sources is the best method to leverage your Big Data using the Alteryx analytic platform. Blending the spatial data provided by geotagging with your internal, customer, and demographic data resources opens the doorway to realizing the full value of Big Data Analytics in decision making.

## *Leveraging External Data Provider Resources*

Data aggregation service providers are companies that assist clients by capturing and leveraging data about their target market, customer base, and industry. These companies are

dedicated to providing a wealth of specific data to clients who use that data to gain a competitive advantage, for a price.

External, packaged market data provides a company with customer and market context that individual companies cannot obtain. These data sources provide companies with information that cannot easily be recreated in a format tailored to their requirements. Companies such as Experian that provides detailed customer data and Target Smart that provides customer and political campaigning are examples of external data providers.

Powerful analytic platforms need to integrate easily with external data providers. Alteryx goes a step further and includes packaged demographic, spatial, firmographic, and household data with the core platform. Fast and simple access to valuable third-party data sources allows data analysts to blend internal corporate data with external data resources for the best possible analytic results and decision making.



Alteryx provides packaged demographic, firmographic, household, and spatial datasets for use with its analytic platform.

## Chapter 4

# Getting Value from Predictive Analytics and Big Data

### *In This Chapter*

- ▶ Understanding why predictive analytics is important
- ▶ Placing the power of predictive analytics in the hands of the people who need it most
- ▶ Getting the most real business value from predictive analysis

**K**

nowing what is likely to happen in the future based on data trends puts the decision maker in the driver's seat when business opportunities develop. Applying predictive analytics to Big Data gives the decision maker forward-looking insights that make a real difference in a highly competitive marketplace.

This chapter delves into why predictive analysis is important and how it can bring value to the business.

## *Why Do Predictive Analytics on Big Data?*

*Predictive analytics* is the process of using a set of sophisticated analytic tools to develop models and estimations of what the environment will do *in the future*. This capability opens up these exciting benefits:

- ✓ You discover new ideas and opportunities much faster, allowing you to be first to market with solutions.
- ✓ You can anticipate customer needs and wants and therefore be better prepared for the business at hand.
- ✓ You can provide rapid turnaround times as environmental conditions alter current requirements.

Figure 4-1 shows how predictive analytics is supported by Alteryx, which offers multiple options for using the R statistical programming language for predictive analytics. The platform provides native integration with R, which allows data analysts to write their own program code or cut and paste code from other projects. The platform also includes sets of prebuilt drag-and-drop tools that don't require any prior R programming language experience or training. The Alteryx platform allows for an easy entry point into using predictive analytics.



For more information about how Alteryx enables predictive analytics, visit [www.alteryx.com/predictive-analytics](http://www.alteryx.com/predictive-analytics).

### Alteryx Predictive Analytic Capabilities

Data Preparation / Investigation Tools

Multivariate Predictive Models / Assessment Tools

Time Series Models / Assessment Tools

Clustering Models / Assessment Tools

Native Integration with  
Full Set of R Capabilities

**Figure 4-1:** Alteryx capabilities and tools for predictive analytics.

Predictive analytics is an extremely powerful capability that has the potential of yielding a substantial advantage for those who can leverage its findings.

## *Moving Predictive Analytics to the Front Lines*

Predictive analytics is powerful, but the concept is not new. For years, early technology generations using predictive analytics have faced these challenges:

- ✓ Data structures favored historic data, and software tools were more suited to reporting what happened in the past than predicting what will occur in the future.
- ✓ Predictive analytics is a complex science that historically required specialized programming techniques, a deep understanding of statistics, and highly trained IT staff.
- ✓ Long lead times were required to set up necessary infrastructure and to answer questions. This increased costs and decreased the value of the results.

Alteryx has identified and remedied these shortcomings with the Alteryx predictive analytics tools included in the Designer Desktop analytic workflow design environment. Alteryx has made predictive analytics attainable for data analysts by:

- ✓ Providing a framework focusing on forward-looking Big Data and trends instead of being limited to a historical perspective.
- ✓ Prebuilding a rich library of powerful predictive tools that leverage the power of the R statistical programming language, available for immediate use.
- ✓ Eliminating the need to wait for highly trained and over-worked statisticians and IT specialists to build out necessary data infrastructure.
- ✓ Providing an easy-to-use, drag-and-drop interface for building sophisticated analytic workflows.

The effect of these improvements is that the power and benefit of predictive analytics has moved forward to the business users and decision makers positioned at the leading edge of the business. No longer is the technology limited to only a few highly trained specialists with deep IT resources; data analysts can leverage predictive analytics as part of their normal analytic development process. Now both operational and executive-level decision makers can easily enjoy the benefits of predictive analytics.

## *Gaining Real Business Value from Predictive Analysis*

Smart businesses apply predictive analytics to their decision-making process to obtain large benefits and advantages in their marketplace. Examples include

- ✓ Taking a proactive rather than a reactive position based on real-time data trends and predictions to ensure the leadership position in the market.
- ✓ Identifying and responding to trends on social media outlets to take advantage of opportunities. For example, if a company detects rapidly increasing interest in a product, the company proactively can increase inventory in the impacted geographic area as identified by the spatial data.
- ✓ Quantifying and leveraging the actual value of social media comments on the business. Blend the context of Big Data with user sentiment analysis and quantitative data to gain an accurate picture of the environment and market direction.

Predictive analysis is a powerful tool for decision makers. With Alteryx Designer Desktop and predictive tools, this capability is available to deliver valuable insights for decision makers at all levels.

## Chapter 5

# Humanizing Big Data Analytics

### *In This Chapter*

- ▶ Making Big Data accessible to the common user
- ▶ Humanizing Big Data for everyone
- ▶ Appreciating the workflow of Big Data Analytics
- ▶ Looking at consumerization of Big Data Analytics
- ▶ Unlocking value with the Alteryx Analytics Gallery
- ▶ Working with data analytics and the cloud
- ▶ Putting your focus on application consumption
- ▶ Finding the best platform for Strategic Analytics

**p**lacing Big Data into the hands of those who need it is an often overlooked, but critical, step. Data must be humanized to have real value to a company (refer to Chapter 2). This humanizing is done by deploying analytic applications to the cloud. These applications must be as easy and intuitive to use as the ones business users use at home. Consumerization of Big Data Analytics with the Alteryx platform plays a key role in your company's future. In selecting the right platform for strategic analytics, you'll be on the path for success.

This chapter takes a look at how to humanize Big Data with the Alteryx Strategic Analytics platform and gain the greatest value from Big Data.

## *Putting Big Data in the Hands of Those Who Need It*

For too long, Big Data has been exclusively in the hands of highly trained data scientists and large IT departments. Long, expensive, and complex IT projects to capture and integrate Big Data often meet with failure. Expensive Business Intelligence (BI) dashboards can provide data to a very small number of high-level executives, but that still isn't sufficient because midlevel and operational level staff require Big Data for their daily work too.

To be useful, Big Data must be accessible to nontechnical people within the company. Fortunately, Alteryx frees Big Data for use by all users. The analytic platform brings Big Data Analytics out from isolation in a data lab and into the business community of end users. Alteryx does so by:

- ✓ Placing Big Data in the hands of normal business users at all levels.
- ✓ Providing a single, easy-to-use interface into Big Data. Extensive IT experience is not required.
- ✓ Enabling instant access to Big Data without long-running IT projects. All the complex engineering work has already been done by Alteryx and is at the users' fingertips.
- ✓ Exchanging ideas and analytic applications within a community of data analysts to ensure that the corporate value from analytics continues to improve.
- ✓ Making analytic applications for business end users and decision makers an immediate reality.

These capabilities place Big Data directly in the hands of those who need it on a daily basis: business end users and decision makers.

# Humanizing Data Design Principles

Rendering vast amounts of Big Data into relevant information for those who need it is the core of humanization. Combining Big Data with market knowledge, location insight, and Business Intelligence while quickly creating analytic applications utilizing predictive and spatial analytics is how Alteryx humanizes Big Data.

Humanizing Big Data is built upon two crucial elements:

- ✓ Making Big Data easy to access by all users who make critical decisions at both the operational and strategic levels.
- ✓ Helping Big Data tell its story by enriching it with business context from all the data sources available and applying powerful analytic capabilities when needed without the need for advanced IT knowledge.

Essentially, humanization puts Big Data into the hands of those who need it and brings out the real business value of Big Data.

Critical design principles for humanizing Big Data are as follows:

- ✓ **Ingest and integrate data from anywhere.** This includes social media, sensor data, external data resources, and internal corporate data sources.
- ✓ **Seek patterns.** Predicting future outcomes is dependent on detecting patterns. Be flexible and creative to fuse the unstructured data with structured quantitative data within a business context.
- ✓ **Make insights available at the point of decision.** Those who make decisions on a daily basis are those most able to benefit from powerful analytic applications.

✓ **Reuse analytic intellectual property.** Data analysts create and share their Big Data analytic work with the full range of decision makers and data users in their company. Each time that their creation is used, it matures in value and context to become a more valuable piece of analytic property.



Leveraging the design principles humanizes Big Data in a faster, more effective manner to maximize the value returned to the company.

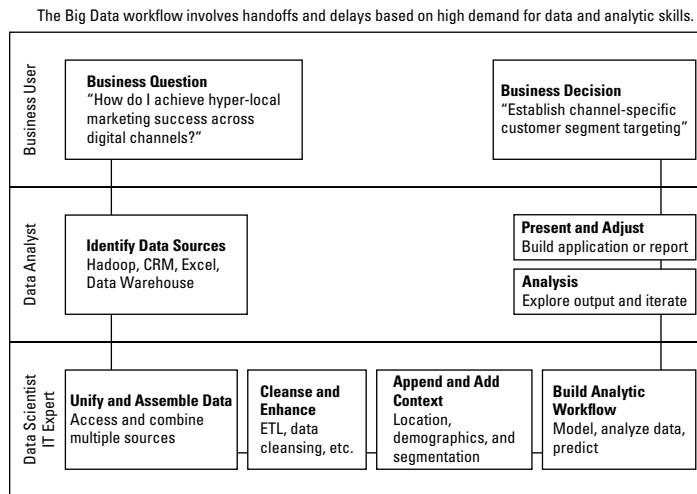
## Humanizing Big Data Analytics Workflow

For most companies, the current workflow for integrating the use of Big Data is highly manual and cumbersome. Value that could be derived is absent from the decision-making process. Several reasons exist for this dysfunctional situation:

- ✓ A vast and turbulent ocean of data exists: structured data across internal databases, large corporate data warehouses, myriads of spreadsheets and documents on workstations, unstructured data feeds, external data resource providers, ever-growing sensor and spatial data, and the wild world of social media content.
- ✓ Outdated BI tools, dashboards, and canned reports that focus well on the *past* but give no real insight into the *future*. Tools and perspective need to be updated for the current and future requirements of business.
- ✓ Too much focus on internal, historical data. There is little market insight, competitive intelligence, location data, and social media analysis. As a result, there is little data context to the present market environment so the data is telling only part of the total story.

- ✓ Traditional IT practices to extract value from Big Data are time-consuming, complex, expensive, and not guaranteed to provide the desired result. Extract, Transform, and Load (ETL) practices for databases are time-consuming. Highly skilled experts are expensive and in short supply.
- ✓ Centralized ownership by the IT department. These efforts need to be led by the data analyst working in conjunction with the business users — not behind the closed doors of the datacenter.
- ✓ Too many steps and handoffs in the overall process, adding to complexity, duration, and increasing the likelihood of mistakes.

Figure 5-1 shows you a typical Big Data workflow and the problems inherent to its design.



**Figure 5-1: Suboptimal Big Data workflow.**

What you see in Figure 5-1 is a structured but complex multi-step and multiparty process that involves so much data requiring so many steps and people that the timely value of Big Data is lost.

If the current process is broken, what is the solution? Alteryx is designed from the ground up to meet and address the problems inherent to typical Big Data workflows. Alteryx does this by:

- ✓ Deploying prebuilt and high performance analytic capabilities tailored to Big Data. This eliminates the need for a team of expensive data scientists to “reinvent the wheel.” The impact is reduced costs, faster timelines, and more consistent results.
- ✓ Placing the data analyst and business user in the driver’s seat rather than taking a backseat to the IT department. Putting the data in the hands of the people who know what they need ensures that the final product will meet the needs of the actual users.
- ✓ Streamlining the workflow process to involve fewer steps with greater involvement by those most impacted by end product. This allows for fewer hand-offs between parties resulting in much faster implementation with fewer opportunities for mistakes.
- ✓ Providing a powerful yet easy-to-use visual toolset that allows the data analysts to seamlessly craft workflows that capture internal and external data, and then use pre-built filters and analytic tools as needed. This allows for rapid creation and deployment of analytic applications to extract real context and value from the data and put it on decision makers’ desktops.

As shown in Figure 5-2, Alteryx applies the Big Data workflow while adhering to humanizing design principles.

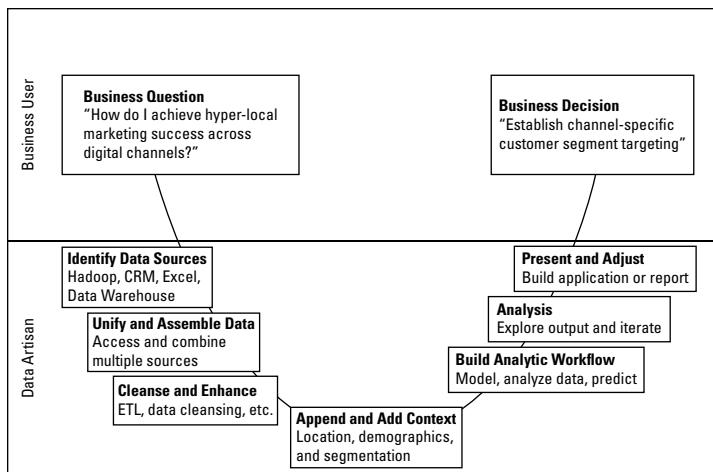


Figure 5-2: Optimal Big Data workflow.

You can see how an Alteryx-designed Big Data workflow is far more streamlined and business centric than typical IT workflows. The result is a much faster, less complex and expensive, and more useful usage of Big Data for those who need it. Alteryx fully implements the vision of humanizing Big Data by rapidly providing it to those who need it while allowing Big Data to add real value by telling its story in context of the business requirements.

## *Considering Consumerization of Big Data Analytics*

*Consumerization* is widely defined as the practice of taking an IT product that is popular in the consumer market space and introducing it into the workplace. Alteryx is a leader of the

consumerization of Big Data Analytics as it brings the power of Big Data into the hands of business users in the workplace.

The Alteryx platform makes using powerful Big Data Analytics an engaging and social experience for the business users via easy-to-use and simple interfaces. Specifically, Alteryx aids in the consumerization of Big Data by:

- ✓ Leveraging cloud technology as a social experience for decision makers to run and share their applications anytime and from anywhere. This contributes to the self-sufficiency and independence of the user.
- ✓ Bringing Big Data context into the decision-making process immediately without waiting for a large IT project.
- ✓ Placing the power of predictive analytics into the hands of every decision maker without adding complexity.

Alteryx consumerization of Big Data allows for better decisions to be made faster and with less complexity by those who need the data the most.

## *Getting an Alteryx Analytics Gallery Overview*

The primary method for consumerization of Big Data by data analysts is through Alteryx Analytics Gallery. The Gallery is a web-based cloud community where data analysts securely access and share their analytic applications deployed to the cloud. To access the Alteryx Analytics Gallery, go to <http://gallery.alteryx.com>.

In Figure 5-3, you see the Alteryx Analytics Gallery front page.



Figure 5-3: Alteryx Analytics Gallery front page.

As you can see in Figure 5-3, the Alteryx Analytics Gallery has been built for consumerized analytics with a highly intuitive and visually appealing interface. The interface makes it as easy to access and run analytic applications at work as the apps decision makers use at home.

If you do not have an account, you can create an account for free and try out the publicly available applications! This is an easy and fun way to see how Alteryx will make your analytic decision making better and more efficient.

## *Publishing Data and Analytics to Cloud Service*

After the data analyst has created an analytic application using Alteryx Designer Desktop, it must be deployed as quickly and easily as possible for use by the business users. The optimal deployment environment is the cloud, from which applications:

- ✓ Are available to designated users regardless of location.
- ✓ May be shared with other data analysts or end users in the community.

- ✓ Are secure so only authorized users can access the application and resulting data.
- ✓ Are fast, highly available, and independent of traditional IT infrastructure.

More than just an application is being deployed; the analytic intellectual property of the company is encapsulated within the application. Fortunately, the cloud is secure and the management and sharing of the analytic components are controlled. Using Alteryx, the data analyst has two major cloud deployment options available:

- ✓ **Alteryx Analytics Gallery:** A public cloud environment with secure login and security restrictions to access applications.
- ✓ **Internal private cloud:** Using Alteryx Private Cloud Server software, a company may host its internal private cloud architecture and deploy applications with the highest degree of security.

Alteryx cloud services prove to be a fast, reliable, and secure deployment framework for applications and sharing of analytic assets within the designated communities.

## *Focusing on Consuming Applications*

Big Data Analytics consumerization is important because it makes analytics accessible and easy for the end user. Several key considerations exist for successful consumerization of applications:

- ✓ Easy access to Big Data for end users is essential. Access must be easily obtained without complex IT involvement.
- ✓ Consumerization should focus on the decision-maker experience and consuming applications. Consider the perspective of the decision maker when developing applications.

- ✓ Specific devices such as mobile devices and smart phones should not be the primary focus. Technology is continually changing and the underlying technological stack should be abstracted from the users' view.
- ✓ All relevant data sources should be consumed. This is where data analysts show their value by excluding extraneous data while not missing what is critical.
- ✓ Easy-to-use parameterized wizards and interfaces that solicit user input to enhance and socialize the user experience. This is how the user "drives" the application so the interface must make sense from a user's perspective.
- ✓ Complex and sophisticated algorithms should be contained within predictive analytic tools without forcing the user to have a deep understanding of the behind-the-scenes technical details. Hiding complexity is achieved by prebuilt analytic tools, and formulas at the data analyst's fingertips.

Shifting the focus from IT and complex data design to intelligently consuming applications so Big Data is easily accessible to business users and decision makers will ensure a successful Big Data Analytics implementation.

## *The Best Platform for Strategic Analytics*

*Strategic analytics* is the sophisticated analysis used to make critical decisions that drive business strategy and growth. This analysis is a direct contrast to the static, prebuilt, and predefined reports and executive dashboards focused on *past* data and performance. Modern strategic analytics products look at the current environment and use Big Data context with advanced analytic tools to gain insights into the *future*. This is important because the forward-looking approach offers the greatest avenue to identify and capitalize on key business opportunities.

Alteryx supports strategic analytics with Big Data via these components:

- ✓ **Alteryx Designer Desktop** to rapidly design, build, and publish powerful analytic applications.
- ✓ **Alteryx Analytics Gallery** to publish, access, share, and run deployed applications in a secure cloud environment.
- ✓ **Alteryx Private Cloud Server** to host and run deployed applications within a secure, private cloud hosted by the customer's organization.

Alteryx is a leading force in the consumerization and humanization of Big Data Analytics. This focus ensures that powerful and complex analytic applications are easily being deployed to those who need them.

Alteryx is also one of the fastest avenues for the design and publishing of analytic applications. In the hands of the data analyst, Alteryx Designer Desktop is used to quickly create and deploy powerful applications that blend data from multiple internal and external sources while allowing Big Data to add context. The results are analytic applications that truly answer the key business questions necessary for the support and growth of the business. The applications are then deployed to the cloud within Alteryx Analytics Gallery for secure and easy access by the decision makers of the company.

## Chapter 6

# Ten (Okay, Nine) Things to Consider with Big Data Analytics

### *In This Chapter*

- Getting the most out of Big Data Analytics

**L**everaging any technology takes some effort, but the payoffs are huge. Knowing a few tips, tricks, and guidelines to make your efforts easier and maximize the benefit is always advisable.

This chapter looks at ways to ensure that your Big Data experience is successful and that you gain the largest competitive advantage possible out of your Big Data Analytics investment.

## *Ensure That the Big Data Analytics Platform Combines Cloud Experience with Sophisticated Analytics*

Cloud computing provides the consumer with multiple benefits: rapid deployment, easy access from anywhere, security, and reduced IT overhead. However, it isn't sufficient to only access lightweight, low-capability BI and dashboards located in the

cloud. Business decision makers require high-performance analytics, but with the convenience and cost benefits of the cloud.

Make sure that your sophisticated, analytic platforms are cloud friendly. Don't assume that because your analytics are powerful, they are not suitable for the cloud. On the contrary, your most powerful processing tools are the best candidates for cloud computing.

Remember to leverage the limitless processing power of the cloud to run your sophisticated analytic tools so they always will be easily accessible to your Data Artisans and business decision makers.

## *Access All the Relevant Data to Make the Best Possible Strategic Business Decisions*

Data comes in many different forms. At one end of the spectrum, data can be structured neatly inside a well-defined corporate database. At the opposite end, the data can be social media with no structured format at all. The data also can be somewhere in between as semi-structured data. Additionally, the data can come from traditional IT corporate data warehouses, desktop workstation documents and spreadsheets, automated smart devices and sensor equipment, or the cloud. Regardless of source or format, all relevant data has value to the decision maker and must be accessible to be useful.

A decision based on bad, incomplete, or unrepresentative data is likely to be a bad decision. Use analytic tools that access all the relevant data available regardless of structure, source, or format. The analytic platform must blend the data seamlessly from the various sources and structures. Furthermore, ensure that the software tools implemented will access the myriad of data sources quickly and easily so that critical business decisions are not delayed.

## *Use a Single Platform for the Complete Analytic Process*

The journey from accessing a multitude of different large data sources to quickly answering key business questions is indeed technically complex. Historically, highly specialized hardware and software to support data management, Extract Transform Load (ETL), integration, access, analysis, reporting, and presentation were common. Each component was highly specialized and required a specific skill set and integration process to transform raw data to valued business decisions. Besides cost and complexity, this process simply takes too long in today's business environments where opportunities exist for only short periods of time before they are seized by competitors.

Be sure to select a platform that wraps up the complete end-to-end analytic process, not just a mix of complex, disjointed components. The platform selected must recognize the workflow associated with analytics and manage that workflow and associated processes from a unified, easy-to-understand user interface.

## *Leverage an Analytic Platform to Access and Make Business Sense of Big Data*

Gleaning real value out of Big Data is critical, otherwise data collection and processing is pointless. However, getting real value quickly is not a trivial task. The complexities associated with analytics have historically limited access to the value of Big Data, but fortunately that situation has changed with platforms such as Alteryx.

To get real value from Big Data, an analytic platform needs to do the following:

- ✓ Ensure availability of all types and formats of Big Data but without long wait times for access, integration, and processing. Big Data sources must be easily and

seamlessly integrated with other, more-traditional data sources without requiring highly skilled technical staff.

- ✓ Leverage your existing Data Artisans, without a massive retraining effort. The target platform needs to be simple and intuitive enough for Data Artisans to use quickly and effectively.

## *Move from Social Media Feedback to Real Business Knowledge*

Many companies monitor social media sources — such as Facebook, Twitter, Foursquare, or Yelp — for consumer comments. Seeing a number of Facebook “Likes” for a product or service is a powerful indicator, but what does it actually mean in terms of real business value? Can you put a dollar amount on a “Like”?

The solution is integration of social media data with more traditional and structured data sources to gather a complete picture of the consumer environment. For example, by themselves, a series of Facebook “Likes” or Twitter comments are only qualitative feedback. However, integrating the geospatial characteristics of the social data with more concrete data from point-of-sale and customer loyalty programs will quantify the true value of those social media inputs. Be sure that your analytic toolset can support this kind of integration so that you capture a complete, quantitative view of the data and financial value.

## *Gain Value from Big Data Sooner, Rather than Later*

Better data yielding better decisions is an easy concept for people to understand. Translating those better decisions into competitive advantage and increased sales is also an easy sell for most business leaders. The problem occurs when the development and implementation path proves to be long,

tedious, and expensive. Frequently, a complex analytic solution deployed in-house by an already overworked IT department can take many months or even years to yield quantifiable benefits for everyday business users.

Platforms such as Alteryx avoid this pitfall by being immediately usable by the Data Artisans and decision makers. Leveraging access to external Big Data sources, integration with internal data stores, and an easy-to-use single toolset bring the value of Big Data to the company immediately. Rather than embarking on a long IT project, decision makers and Data Artisans are actively using Big Data Analytics with quantifiable results in a very short timeframe.

## *Know Where Your Business and Customers Are Located*

Operating without a profile of your customers is obviously a path to failure. Demographic data has proven itself critical, and many demographic data sources have been available for years. However, one of the most important data attributes historically has been overlooked due to technical challenges; fortunately, the use of geospatial data is now possible by everyone.

Geospatial intelligence tells you where your customers are, where they visit, where they buy your products, and where competitors are located. Social media tags data with geospatial tags and analytic processing tools such as Alteryx Strategic Analytics take advantage of these tags to identify and analyze this data. Ensure that spatial intelligence is a part of any analytic processing toolset you utilize so that this data will enhance critical decision making.

## *Value Your Data Artisans and What They Bring to the Table*

The most important component in any system is always the people, and the same is true especially in BI and analytic processing. Of your staff, the most critical folks are your Data

Artisans. These are your business experts, typically attached to a business unit, who truly understand your business and what data is critical. They understand IT, but technical knowledge is not how they bring value. The ability to know what data the decision maker needs, where to find that data, and how it needs to be analyzed and processed is the critical, game-changing skill set of the Data Artisans.

As with any advantage, a smart business leader wants to maximize that advantage for maximum impact. The way to make the most of your Data Artisans is to empower them and to provide the tools that best support their work. Agile, powerful, and easy-to-use platforms such as Alteryx Strategic Analytics are the best way to leverage the inherent knowledge of your Data Artisans.

## Notes

## Notes

# HUMANIZING BIG DATA

Big Data continues to spark excitement, but most businesses are still searching for a way to get value from it

More employees need access to Big Data to make informed decisions

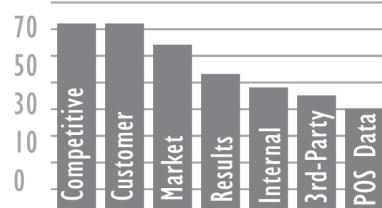
77%

Top types of business decisions that would benefit from the inclusion of Big Data



Market Opportunity    Customer Retention and Segmentation

**Big Data Analytics drive new market opportunities and customer retention**



Not only is there a need for more relevant Big Data, but there is also the need for data to be analyzed in combination with the appropriate traditional data. This in addition to broader and better access will allow Big Data to fulfill its purpose: assisting business leaders in making better decisions with the information they have.

## Make better and faster business decisions with Big Data using powerful analytics

Big Data holds incredible value to businesses that know how to access and use it. Social media, cloud data, and enormous customer data providers all provide the backdrop for Big Data to tell its story. For that story to be heard, however, powerful analytic platforms must quickly and easily delve into the relevant data to find the right information to answer business questions. Use predictive analytics and a cloud architecture to unlock the real value of Big Data for your business!

- ***Understand Big Data — find out how Big Data volume, velocity, and variety are changing analytic processing***
- ***Deploy analytic applications — build powerful Big Data analytic applications with easy-to-use desktop tools and deploy them to the cloud***
- ***Leverage high-performance capabilities — use the fastest methods possible to unlock value from Big Data for your key decision makers***



**Open the book and find:**

- **What exactly Big Data is and how it works**
- **How to maximize the use of Big Data with powerful analytic applications**
- **Why humanizing Big Data improves your analytic capabilities**
- **How to combine Big Data with predictive and spatial analytics**
- **Why Big Data with analytic processing enables better business decision making**

**Go to [Dummies.com](http://Dummies.com)**  
for videos, step-by-step examples,  
how-to articles, or to shop!

**WILEY**

ISBN 978-1-118-60704-6  
Book not for resale