

DESCARTES™



Unlocking the Analytics Advantage: **Part 1**

# The Democratization of Analytics for Transportation and Logistics



# Table of Contents

- The Analytics Advantage** .....1
  - Data-driven Organizations Realize an Analytics Advantage .....2
  - The Democratization of Analytics .....2
  
- What Keeps Companies from Realizing Their Analytics Advantage?** .....3
  - Data Science Talent Shortage.....3
  - The IT Bottleneck.....4
  - Data Wrangling Challenges .....4
- Leveraging Solution Providers’ Expertise and Pre-integrated Datasets.....5
  - Dedicated Data Science Team.....5
  - Robust Analytic Technology .....5
  - Customer-specific Data Integration Services.....6
  - Data Wrangling Capabilities .....6
  - Comprehensive Suite of Transportation and Logistics Solution Functionality .....7
  - Finding the Right Solution Partner Makes a Difference.....7
  
- Unlocking the Analytics Advantage**.....8



This is the first in a three-part series on using analytics in transportation and logistics to achieve a competitive advantage:

***Part One: The Democratization of Analytics for Transportation and Logistics***

Here in part one, we explore the kinds of data being produced by transportation and logistics systems, how that can be used to create a data-driven enterprise, the substantial obstacles to achieving an analytic advantage, and how those obstacles can be overcome.

***Part Two: Analytics for Private Fleet and Driver Performance***

In part two, we look at how analytics uses real-time location data, combined with orders, plans, proof-of-delivery, vehicle data, and more to drive significant improvements to fleet and driver performance.

***Part Three: Analytics for Improving Carrier Performance and Leveraging Trade Data***

In the third and final part, we discuss how analytics can create improvements to carrier performance while substantially reducing costs, as well as how trade data can help importers and exporters gain competitive insights, manage supply chain risk, optimize total landed costs, and more.

# The Analytics Advantage

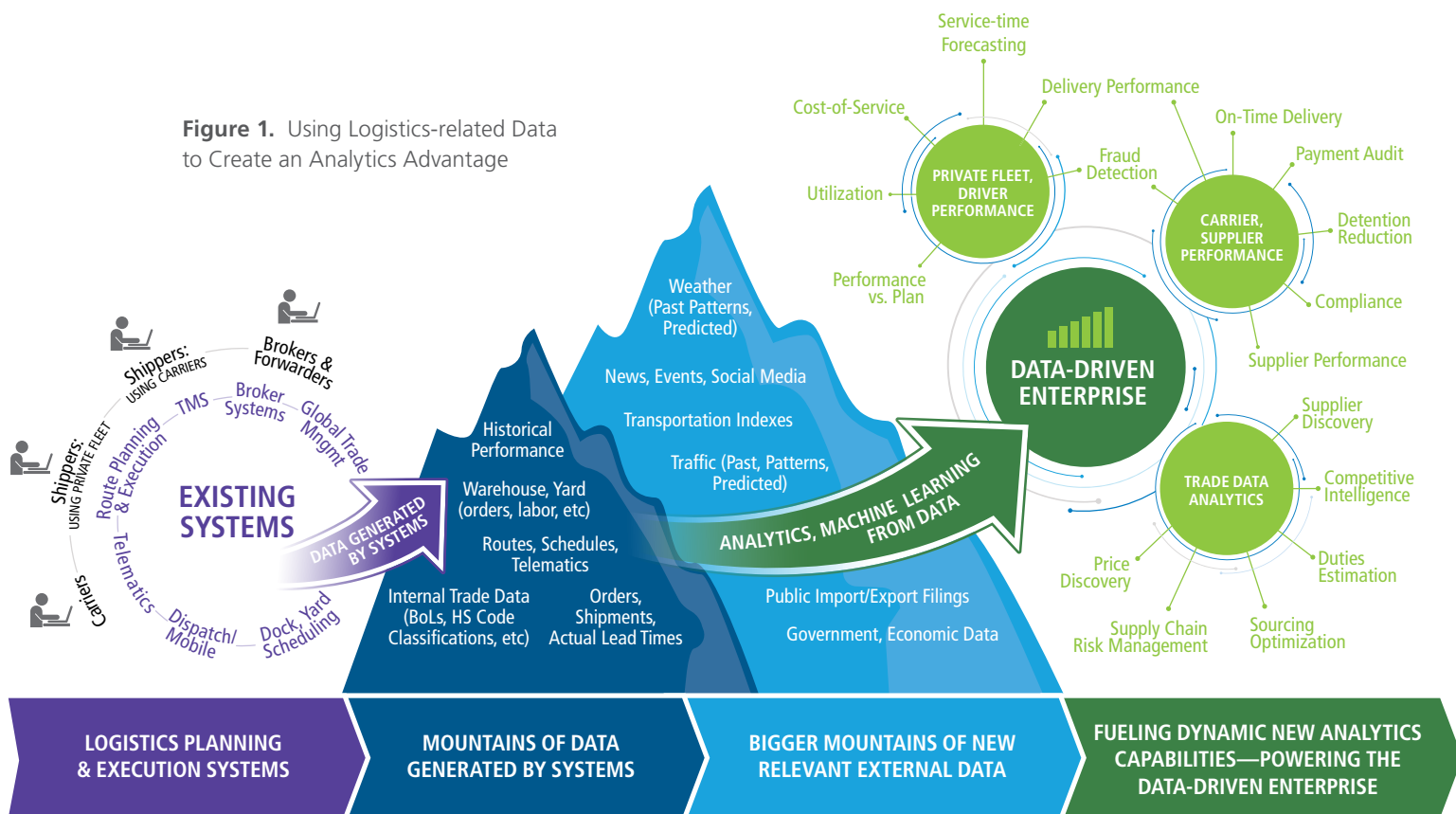
As companies continue to digitize their supply chain and logistics operations, they are beginning to recognize that they are sitting on a wealth of information that has an enormous, yet largely untapped potential to help increase their performance and improve their customer service.

Companies are beginning to recognize that they are sitting on a wealth of information that has an enormous, yet largely untapped potential.

Increasingly, there are examples of how companies have gained significant value through the data that is available to them. Not only do their internal systems generate volumes of relevant data, but there are external sources that can be leveraged to give a comprehensive view of their business, supply chain and even their entire industry. This level of insight is profoundly changing how companies use their supply chain and logistics operations and the value they deliver.

Supply chain and logistics systems generate a wealth of valuable data. Core planning and execution systems such as a transportation management system (TMS) or route planning, real-time visibility, mobile, and telematics are prime data generators. For companies that operate internationally, importing or exporting of goods often uses global trade management systems to classify goods, submit filings, and ensure compliance, creating additional valuable supply chain and logistics data. All of the data from these systems needs to be brought together into a cohesive interlinked set, in order to provide deep insights into operational performance and customer service. Companies that are able to intelligently use this data, become data-driven enterprises that realize a powerful 'analytics advantage.'

**Figure 1.** Using Logistics-related Data to Create an Analytics Advantage



## Data-driven Organizations Realize an Analytics Advantage

Data-driven, analytics-capable organizations are able to ask the right questions and extract tremendous value from the mountains of data they already have. They are able to find very specific ways to improve their purchased transportation or fleet performance, evaluate their ability to better serve their customers and the cost to do it, understand and optimize the total landed cost of their international sources, improve how well they are complying with regulatory mandates, and much more. Becoming data-driven fundamentally changes an organization’s culture, decision-making processes, and the role of metrics.

	Traditional Organization		Data-Driven Organization
<b>Culture:</b>	Metrics-driven, KPI-oriented culture	⇒	Insights-driven, Outcomes-oriented culture
<b>Diffusion:</b>	‘Ivory-tower Analytics’, done by a few	⇒	‘Democratized Analytics’, done by the masses
<b>Decision-Making:</b>	Decisions based primarily or solely on experience and intuition	⇒	Data and analytics play strong roles in decision-making processes
<b>Data Literacy:</b>	Analytics seen as a ‘black art’	⇒	Empowered, data-literate employees
<b>Data Organizing:</b>	Siloed data, one-off data-gathering projects, inadequate resources	⇒	Integrated data, ongoing data-wrangling resources/activities (internal and/or external)

**Table 1.** Characteristics of Traditional vs. Data-Driven Organizations

### The Democratization of Analytics

What really sets a data-driven organization apart is how widely and deeply diffused the use of data and analytics is throughout the enterprise. Analytics should not be the exclusive domain of a few data scientists isolated in their ivory tower. Data-driven insights and the ability to ask questions should be available throughout the organization, at the point wherever decisions are made. Analytics are democratized by infusing them in the systems and processes used by professionals throughout the organization.

Analytics should be accessible to all—the people at the front lines making the day-to-day decisions and getting things done, as well as to the managers and executives responsible for functional and business unit effectiveness. Broad access to good analytics ‘lifts the blinders off’ for an organization. Employees gain much more specificity into where and how performance can be improved, as well as the tools for early identification and correction of execution issues. The organization becomes much more agile and competitive—quickly sensing issues and quickly adjusting course, solving those issues before they become bigger problems. Employees are more empowered to solve problems directly and effectively. As a result, morale, individual performance, and job satisfaction improve. Democratization is a powerful force.



Analytics should be accessible to all—the people at the front lines making the day-to-day decisions and getting things done, as well as to the managers and executives responsible for functional and business unit effectiveness.

# What Keeps Companies from Realizing Their Analytics Advantage?

As more and more competitors become data-driven, those who don't do so will be at a major disadvantage. Many companies are constrained in their use of analytics due to a set of interrelated challenges:

## 1 Data Science Talent Shortage

While there are shortages of IT talent across the board, data scientist and big data analytics specialists top the list as the scarcest.<sup>1</sup>

## 2 The IT Bottleneck

For organizations lucky enough to have some inhouse talent, the backlog<sup>2</sup> of projects can stretch for months or years.

## 3 Data Wrangling Challenges

Data wrangling typically consumes 50%–80% of data scientist/analytics project resources, with precious little time and resources left for building models and doing analysis.

# 1

## Data Science Talent Shortage

The shortage of data scientists has made it hard for even deep-pocketed companies to hire and retain an adequately staffed inhouse analytics team. [LinkedIn's 2018 Workforce Report](#) said "Demand for data scientists is off the charts." The report singled out data scientists over all the other jobs.

The number of data scientist jobs grew by over 6.5X from 2012–2017<sup>3</sup> and continues to grow at over 50%/year. Most estimates indicate there are more than twice as many data science job openings as there are qualified candidates. The consensus is that demand for these skills will continue to increase faster than the supply of qualified candidates for many years to come.



## 2

**The IT Bottleneck**

Hiring good data scientists is only part of the battle. Most IT departments are already stretched thin and being asked to do more with less. Over 70% of their time (on average) is consumed maintaining and supporting existing systems, rather than helping implement new functionality or delivery of new types of analytics and reports. IT department project backlogs often extend to over a year (in some cases several years). Analytics projects just have to get in line. That is the opposite of being agile for a business trying to keep up.

## 3

**Data Wrangling Challenges**

Data wrangling involves finding, accessing, organizing, cleaning, and enriching data from many different sources into a cohesive data set that is useable for analytics.<sup>4</sup> The effort required to do this is usually vastly underestimated. For more details on the activities involved and what consumes all that time, see the sidebar *What Makes Data Wrangling So Time-Consuming?*

Given all these challenges, individuals often manually gather whatever data they can and 'grind out answers' in Excel. They can use spreadsheets to try to find underperforming drivers, carriers, and so forth, to get a better handle on what is going on. Yes, they can get answers this way, but it is a very manual approach. A more scalable and sustainable approach is needed. Partnering with the right transportation and logistics solutions provider can be the answer for many companies.

**What Makes Data Wrangling So Time-Consuming?**

Activities<sup>5</sup> that make data wrangling challenging and time-consuming include:

- **Define and clarify the use cases:** Working with the end users, creating clear definitions of what problems you are trying to solve, what kinds of entities are involved (people, transactions, products, batches, loads, etc.), what data is likely to be relevant, potential sources for that data, etc.
- **Obtain access to the data/ensure proper security:** Obtaining access to confidential data and putting mechanisms in place to ensure it remains secure; may mean exposing only derived inferences, while confidential data remains hidden and inaccessible. Figuring all this out, setting it up properly, and doing the necessary security testing/audits takes time.
- **Resolve key entities/data elements:** This includes deduplication which can be a non-trivial exercise when entities have multiple identities or identifiers (like different spellings or part numbers).
- **Identify the relationships between the data:** When pulling data from many different sources, the relationship between the data needs to be understood and incorporated into the analytic model. This can involve identifying natural keys and how they relate. The data owners will need to confirm that the relationships are correct.
- **Discover any changes to syntax or semantics:** If the syntax or semantics of any of the data has changed over time, that needs to be taken into account, or any results incorporating data from those time periods will be incorrect.
- **Find and incorporate dispersed data about actions, outcomes:** Data about actions taken and the outcomes that resulted may be dispersed across individual spreadsheets, emails, and other documents. Finding and incorporating this data is time-consuming.
- **Identify and adjust for biased data:** Algorithms are only as good as the data they are given. Selection bias occurs when the sample data available is not an accurate representation of the real world.<sup>6</sup> There are many other types of bias<sup>7</sup> in the data as well. These can be adjusted for, but it is labor-intensive.
- **Feature Engineering:** Feature engineering is the process of selecting and transforming data into attributes for a machine learning model to use. It is one of the most important and time-consuming activities for successful machine learning.
- **Correlation/redundancy analysis:** Data that are redundant should be removed to reduce the complexity of the model. Examples might be when you have different fields for the same attribute in different languages, or a numerical value that fully represents another text attribute.

## Leveraging Solution Providers' Expertise and Pre-integrated Datasets

Data wrangling is best done by those who know the data the best—those who work with the data and systems every day. One option is to see whether the software company that provides your transportation and logistics solution has the required data science expertise, has already done the data-wrangling for a wide array of common use cases, and has the tools and services to rapidly incorporate your unique data. This can save a tremendous amount of time and money, allowing an organization to get started with advanced analytics capabilities almost right away.



Data wrangling is best done by those who know the data the best.

Some things to look for when assessing whether or not a logistics software solution provider has the right capabilities to provide the analytic capabilities you need include:

- Dedicated team of data scientists
- Robust analytic technology
- Customer-specific data integration services
- Data wrangling capabilities
- Comprehensive suite of transportation & logistics solution functionality

Below are some questions to ask and things to look for when assessing solution provider's analytics services and capabilities.

### Dedicated Data Science Team

How many data scientists does the solution provider have who are dedicated to providing these services and what is their experience? As this is the scarcest resource, it can be an important differentiator to achieving results.

### Robust Analytic Technology

The analytics technology from a solution provider may be a mix of software they developed in-house and ISV-developed tools.<sup>8</sup> Some things to look for include:

- Visualization capabilities that are versatile, intuitive, and easy to use
- Customizable live data dashboards
- Ad hoc analysis and reporting
- Database connectivity preconnected to a wide variety of databases, especially the ones your organization uses and cares about
- Support for NoSQL, unstructured, and streaming data (e.g. IoT data feeds)
- Scalability for ingesting and managing large volumes of data (depending on your needs)
- Ease-of-use for business users (business analysts, power users, and casual users)
- Granular security and personalization
- Data quality tools for finding and correcting deficiencies in data with minimum manual effort
- Responsive UI and use cases to include mobile analytics and mobile data collection
- Embedding of analytics into applications and workflows—prebuilt into the solution providers applications, but also the ability to embed analytics into your own and third-party applications
- Active user and developer communities, can make a big difference in finding answers and support





### Customer-specific Data Integration Services

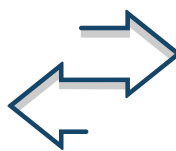
We recommend looking for a solution provider that offers tools and services to integrate your own company-specific data, such as your enterprise transactional data, including orders, shipments, BOLs, customs filings, etc., as well as master data (customers, suppliers, carriers), planning data, and any other data you possess that might be useful for analytics. This may include data in spreadsheets and other individual documents.

The provider's integration services should also incorporate unique external data you may use or need such as weather, traffic, and data that may be specific to your industry or company's needs (like specific commodity prices or event schedules). This should be more than a set of APIs the solution provider hands to you and says, 'have at it.' Unless your IT shop is set up for and wants the additional workload of building and maintaining those integrations, you may prefer a provider that can do the integration work and maintain the integrations over time. Ask about their integration methodology, services, tools, and ongoing support to better understand what they really provide.

### Data Wrangling Capabilities

When evaluating a transportation and logistics solution provider's support for analytics, it is good to find out what level of data wrangling they have already done and are willing to do. Some specific questions to consider:

- What is their overall approach, methodology, tools, and resources for data wrangling?
- What use cases have they already implemented? What data and systems—both from the solution provider and from the customer—are involved in those use cases?
- What kinds of tools and methodologies do they have for cleaning up and organizing the data?
- What is their security architecture and approach for data in the analytics data set?
- What varieties of data do they handle?
- What steps do they take to ensure performance and scalability, such as denormalization?
- What steps, if any, do they take to adjust for bias in the data?



Look for a solution provider that offers tools and services to integrate your own company-specific data, as well as unique external data specific to your industry or company's needs.



### Comprehensive Suite of Transportation and Logistics Solution Functionality

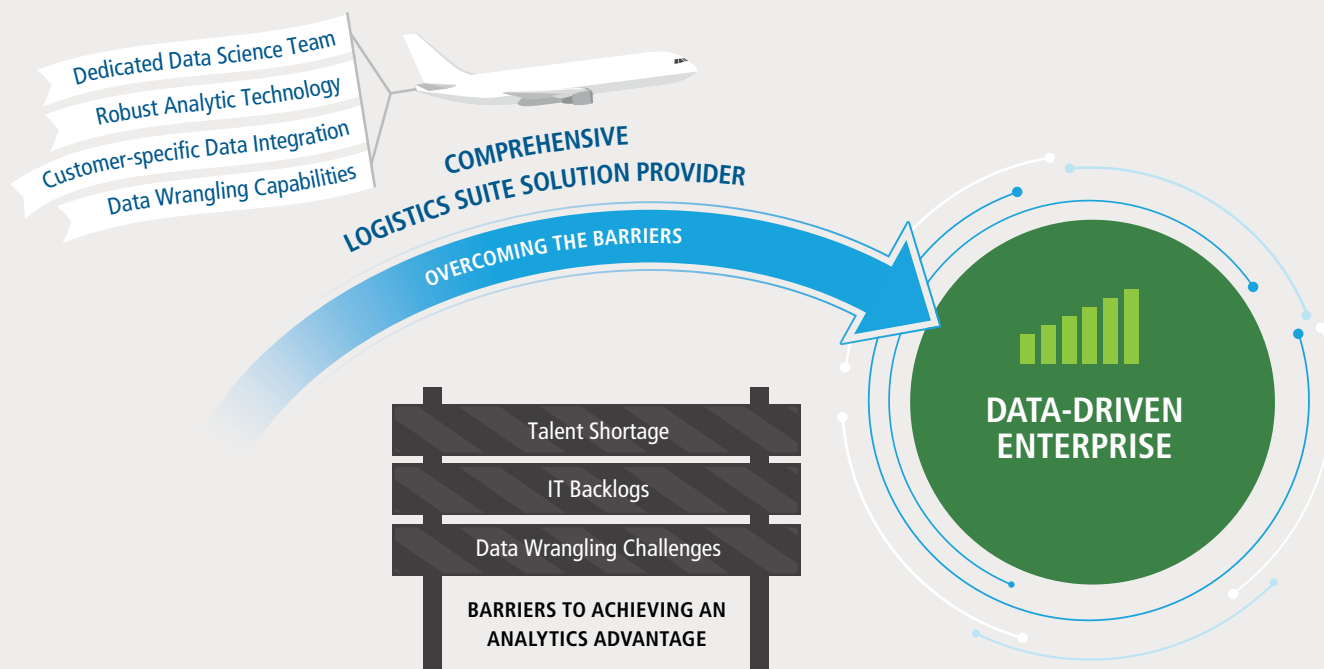
It is better to try to find a solution provider that can fulfill all of your transportation and logistics needs and that has already done the data wrangling and provides the kinds of analytic capabilities and integration services described above.

For example, a solution suite that can manage parcel, TL, and LTL; can manage both private fleet and purchased transportation ; across all modes ; has global trade management capabilities ; supports pooled distribution ; offers telematics hardware and software ; dock scheduling and yard management ; and ecommerce integration. Going with a single analytics-capable provider that can handle all of your transportation and logistics needs will simplify your analytics strategy and implementation, compared to trying to stitch together offerings from multiple providers.



### Finding the Right Solution Partner Makes a Difference

Partnering with the right solution provider that has the right combination of transportation and logistics software capabilities, analytics tools and expertise, and data wrangling services can help jump start analytics efforts, dramatically reducing risk, upfront investment, and time-to-value to achieve an analytics advantage.





## Unlocking the Analytics Advantage

The process of digitizing a company's supply chain and logistics processes and systems is a journey.

Even early steps on that journey start to generate valuable data that can be mined. In fact, many companies are sitting on a goldmine of largely untapped transportation and logistics data. Part of the reason the value remains locked up is the scarcity of data science talent and the time and resources to wrangle data. These challenges can be overcome by choosing a full-suite transportation solution provider that has invested the time and resources to acquire the necessary data science talent and has done the heavy lifting involved in wrangling data across their suite, as well as their customer's systems.

Applying analytics can be like a near-sighted person putting on a pair of glasses for the first time. All of a sudden everything that was blurry comes into focus. As the use of analytics matures in a company, it can become an 'insights engine' for them, highlighting not just where problems are, but the 'why,' what is causing those problems, what are the potential solutions, and the tradeoffs between those potential solutions. As a company becomes more adept at leveraging these insights, a true 'analytics advantage' is realized.



Many companies are sitting on a goldmine of largely untapped transportation and logistics data.

## Notes

- 1 In the [Harvey Nash/KPMG CIO survey](#), CIOs were asked “*Which functions do you feel suffer from a skills shortage?*” Presented with a list of 22 different IT-related functions, *Big Data/Analytics* topped the list in both last year’s survey ([What’s Driving Data Science Hiring in 2019?](#)) and this year’s survey (see chart in the [results here](#)—scroll down to question 16).
- 2 IT backlogs are not just due to the workload on data scientists from analytics projects, but all of the IT staff from a variety of demands such as maintenance and support of existing systems (which consume on average over 70% of IT budget and resources), massive implementations or major upgrades to ERP or other systems, integration and data cleanup efforts, and so forth.
- 3 See [LinkedIn’s 2017 U.S. Emerging Jobs Report](#)
- 4 Data wrangling may also be used to prepare data for AI and machine learning uses
- 5 For more on these activities, see [What is Data Wrangling and Why Does it Take So Long](#).
- 6 For example, suppose you have rich data about a subset of customers, but that subset is not a representative sample of your broader customer base. Perhaps you have certain data only from opt-in customers, and maybe they are younger or of a certain mix of persona types that is different from your overall base of customers. For more, see [Picking Favorites: A Brief Introduction to Selection Bias](#)
- 7 Other types of bias in the data include seasonal bias, linearity bias, confirmation bias, recall bias, survivor bias, observer bias, and reinforcement bias.
- 8 ISV = [Independent Software Vendor](#). One source for evaluating these kinds of tools is the [Gartner Magic Quadrant for Analytics and Business Intelligence Platforms](#). The top four ISVs in Gartner’s 2019 evaluation were, in this order: 1) Microsoft Power BI, 2) Tableau, 3) Qlik, 4) ThoughtSpot. It is notable that Microsoft is rated quite a bit ahead of all the others and has been ranked #1 in Gartner’s evaluations for the past 12 years in a row.



#### **About ChainLink Research**

ChainLink Research, Inc. is a Supply Chain research organization dedicated to helping executives improve business performance and competitiveness through an understanding of real-world implications, obstacles and results for supply-chain policies, practices, processes, and technologies. The ChainLink 3Pe Model is the basis for our research; a unique, multidimensional framework for managing and improving the links between supply chain partners.

#### **For more information, contact ChainLink Research at:**

321 Walnut Street, Suite 442, Newton, MA 02460-1927

[www.clresearch.com](http://www.clresearch.com)

[info@clresearch.com](mailto:info@clresearch.com)

(617) 762-4040

**DESCARTES™**

[www.descartes.com](http://www.descartes.com)

[info@descartes.com](mailto:info@descartes.com)