

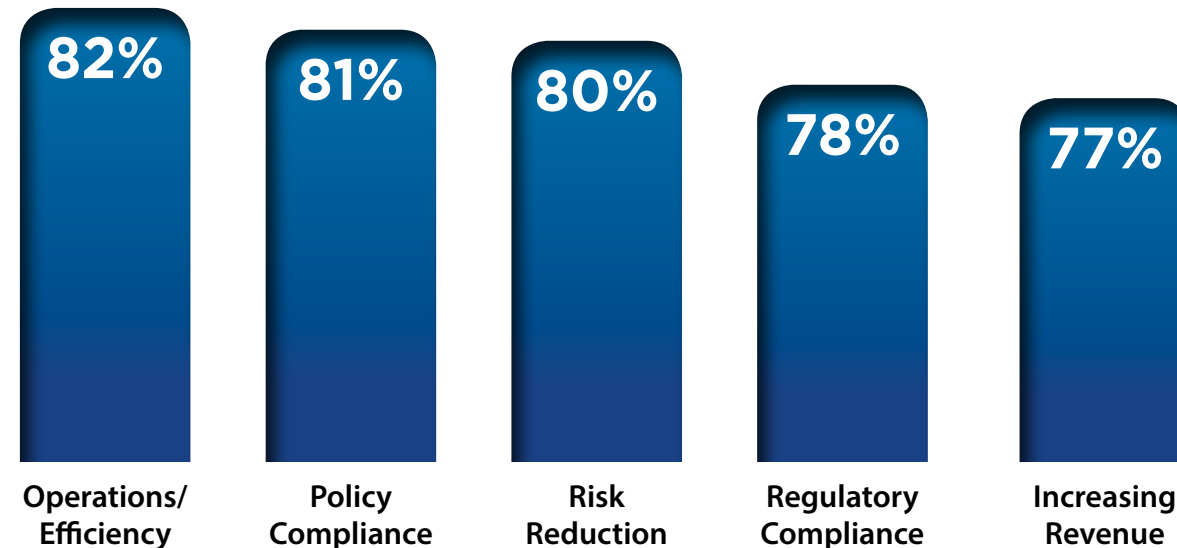


The State of Data Discovery and Cataloging

An IDC InfoBrief, Commissioned by Alteryx | January 2018

Data discovery: The ability to locate, understand, access and trust data is a key enabler of business in the era of digital transformation

Importance of Data Discovery to Business Drivers



No **one** driver stands out. Data discovery is important for business. Period.

Even though data discovery and integrity is important for business, 30-50% of organizations are not where they want to be

Top 5 current realities:

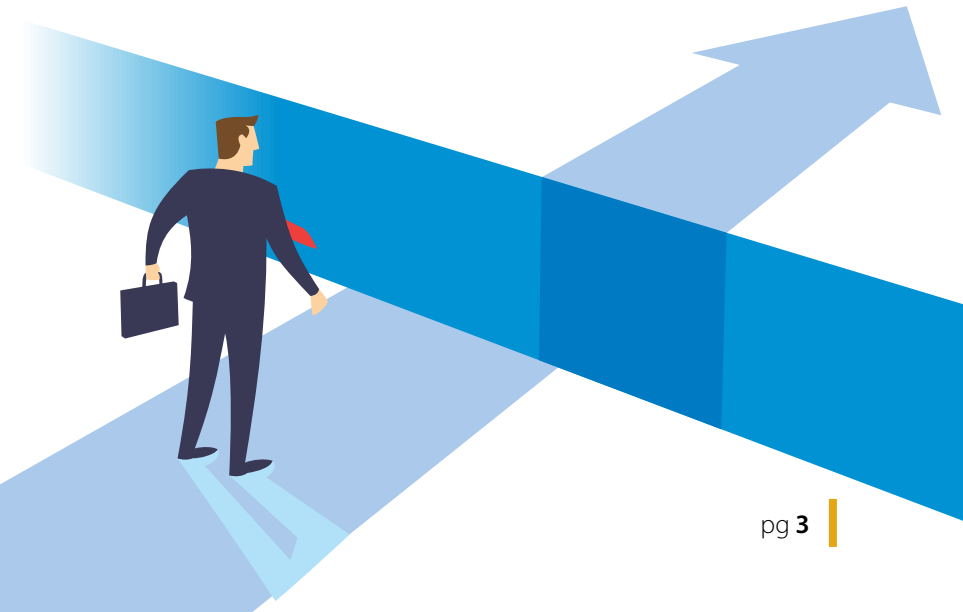
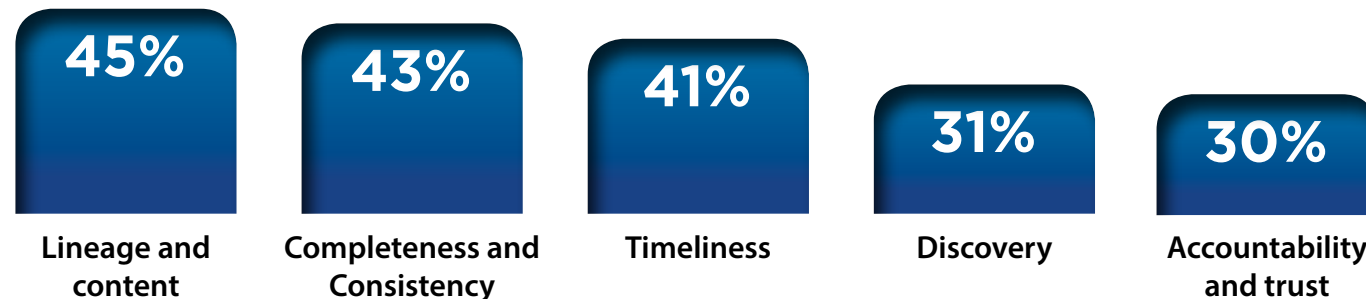
- Information is incomplete
- Information isn't timely
- Information isn't easy to use
- Information isn't easy to discover
- Information isn't trusted

Top 5 important attributes:

- Complete, Consistent Data
- Access to Timely Data
- Data Lineage and Context
- Data Accountability and Trust
- Data Discovery

Gaps are impacting personal efficiencies and effectiveness, increasing organizational risk, and costing time = money.

% of Respondents that Have a Gap Between Reality and Importance



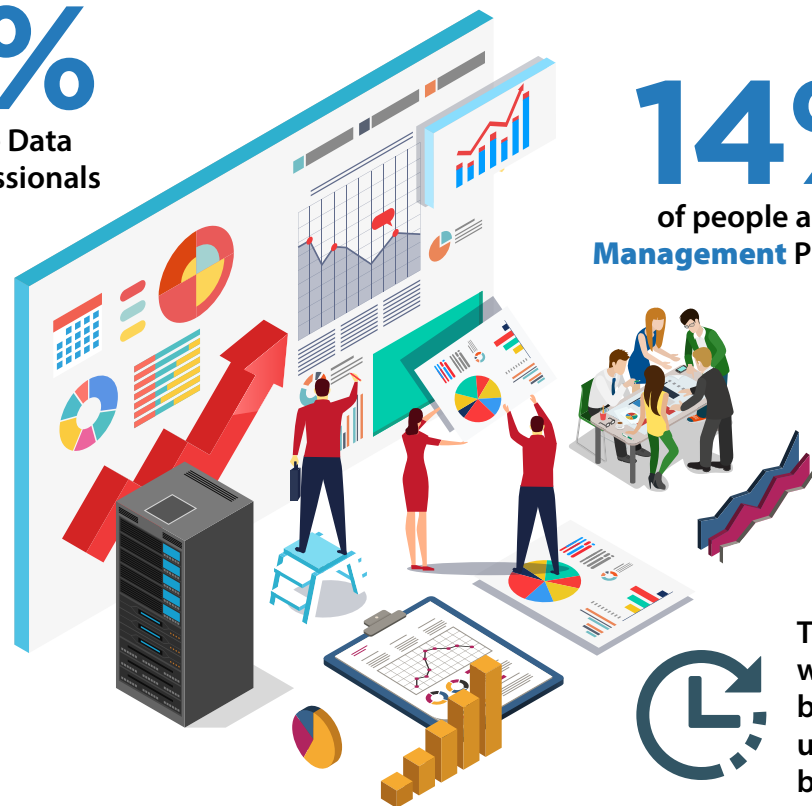
Data professionals on average **make up a third** of every organization's workforce, and most are working more than 50 hours per week

20%

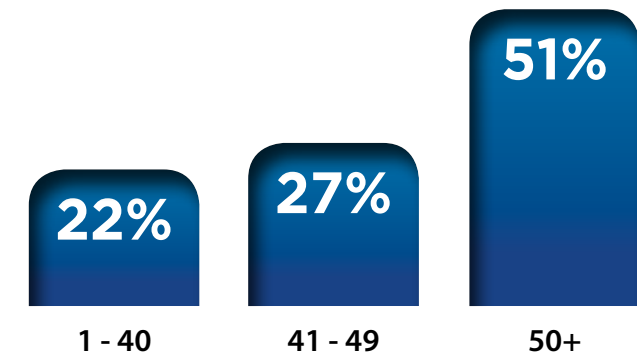
of people are Data
Analysis Professionals

14%

of people are Data
Management Professionals



Average Hours Worked Per Week

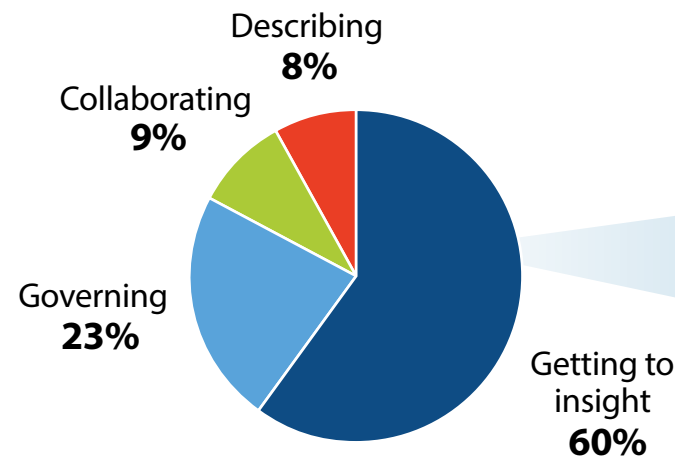


Long hours are symptomatic of data professional's ability to be efficient and effective in data activities.

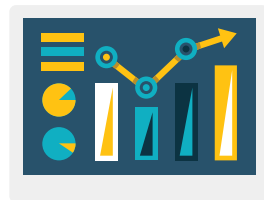
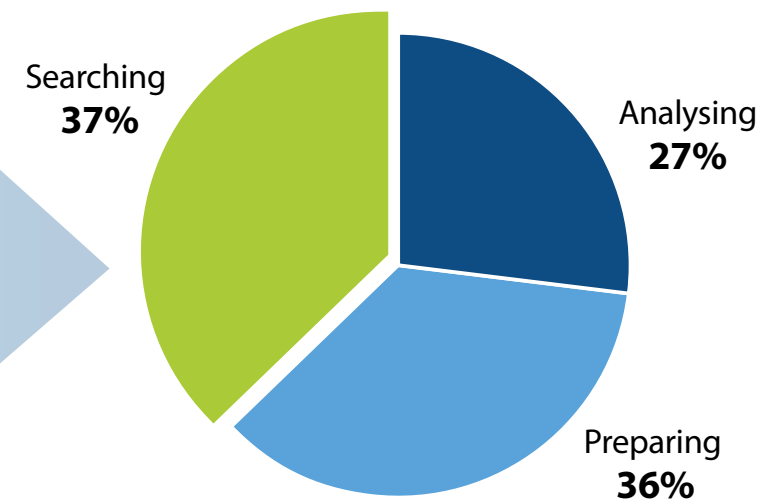
The population of people working with data on a daily basis is increasing as the use of data in organizations becomes more pervasive in business processes.

Data professionals are spending more of their time governing, searching and preparing data than they are on extracting value

Average % of Weekly Time by Activity

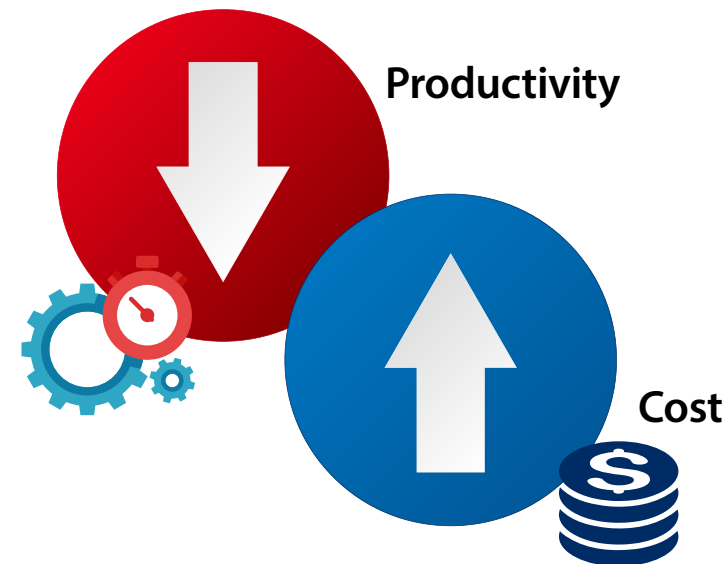
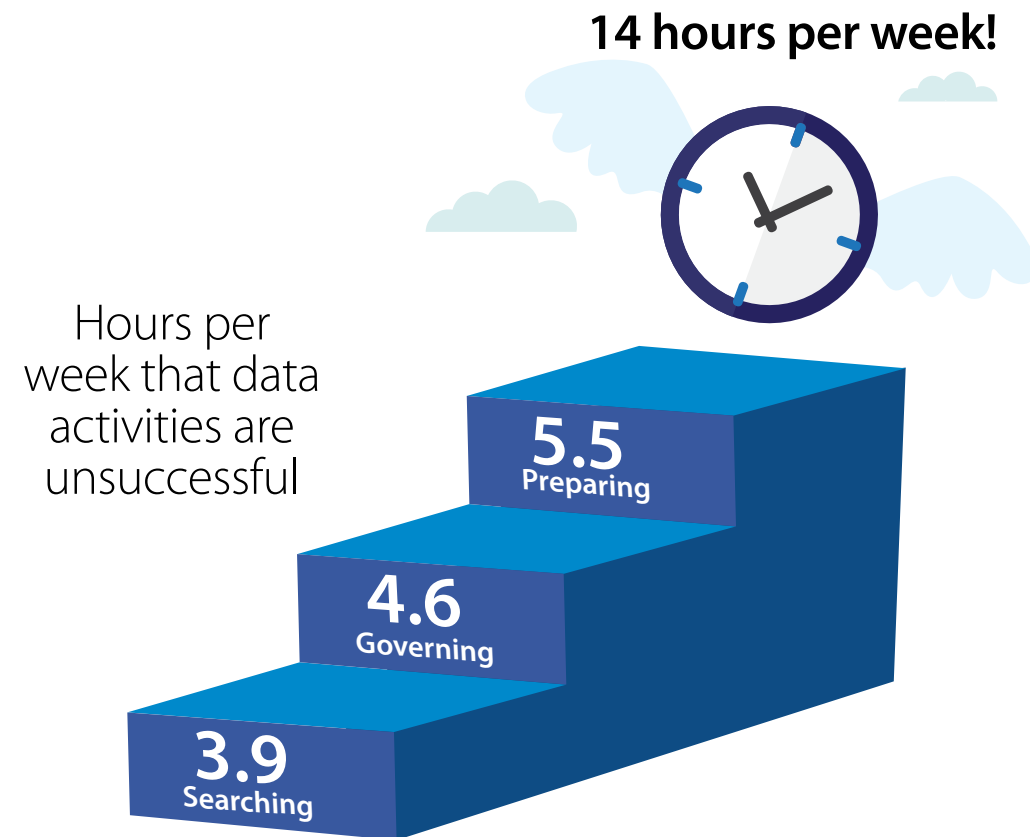


Average % of Time Spent Getting to Insight



Information value is created through analytics and insight, but data professionals are still spending more time searching and preparing data than on analysis for insight.

Data professionals are wasting on average, 30% of their time because they cannot find, protect or prepare data



As data volume and variety continues to grow, resources will become more ineffective because data and information assets will be harder to find.

Data professionals also need to stop re-creating the wheel with every new data analysis project

Knowledge of how and where data is used, by whom, and what information assets already exist will help professionals stop re-creating the wheel, and give them more time back for value added activities.



DATA PROFESSIONALS ON AVERAGE WASTE ANOTHER 20% OF THEIR TIME EVERY WEEK
(or 10 hours) building information assets that already existed.

The current state of data intelligence and knowledge is costing people time, and organizations money

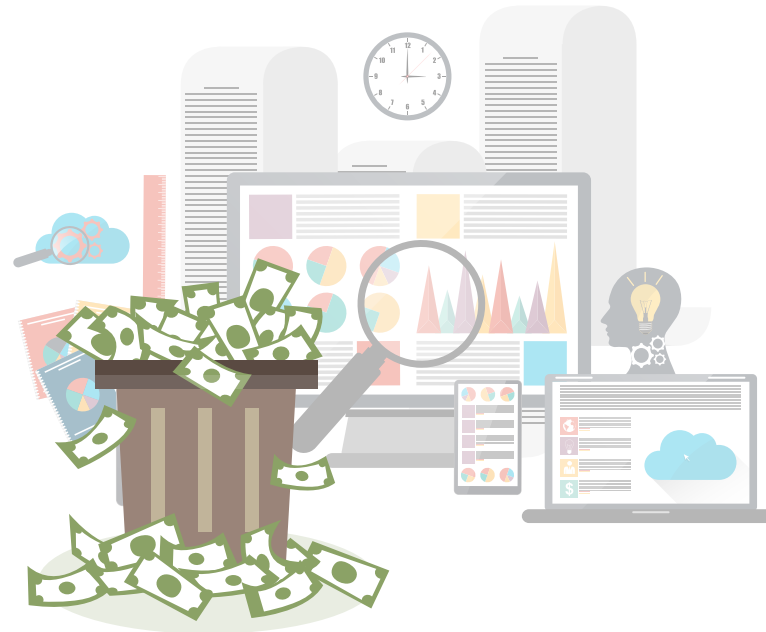
Data professionals are losing **50% of their time** every week: **30%** unsuccessful on data activities, plus **20%** of their time building what already exists.

\$1.7M

per year for every 100 employees in US organizations

€1.1M

per year for every 100 employees in EU organizations



\$103M

per org per year in the US

Based on the average number of employees per organization in the survey population.

€49M

per org per year in Europe

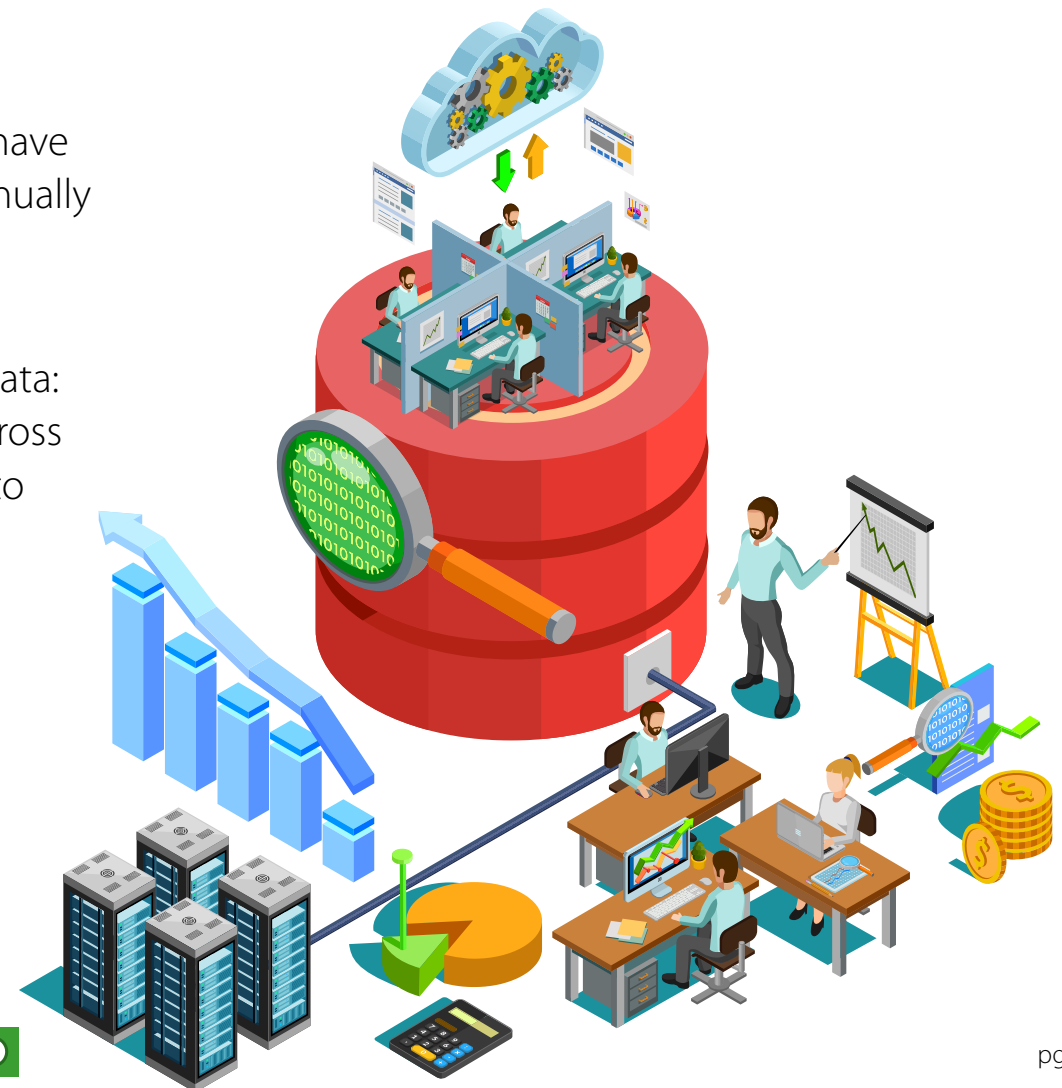
The business case for improving data professional productivity shouldn't only be about the money, but also about giving employees up to 50% of their time back.

Data catalog software is emerging to help organizations close the gaps on data discovery and integrity

Data information management solutions aren't new. Organizations have been using spreadsheets, commercial and bespoke solutions to manually track data origins, define business glossaries, manage ownership, stewardship, metadata and usage information.

Manual solutions are falling short in the age of big and distributed data: the volume, velocity, variety, and distribution of data and analysis across on-premises, in cloud and hybrid environments make it impossible to manually capture and maintain data catalogs.

Data cataloging software solutions are emerging, automating the process of data discovery, applying machine learning, and turning to the crowd for added context.



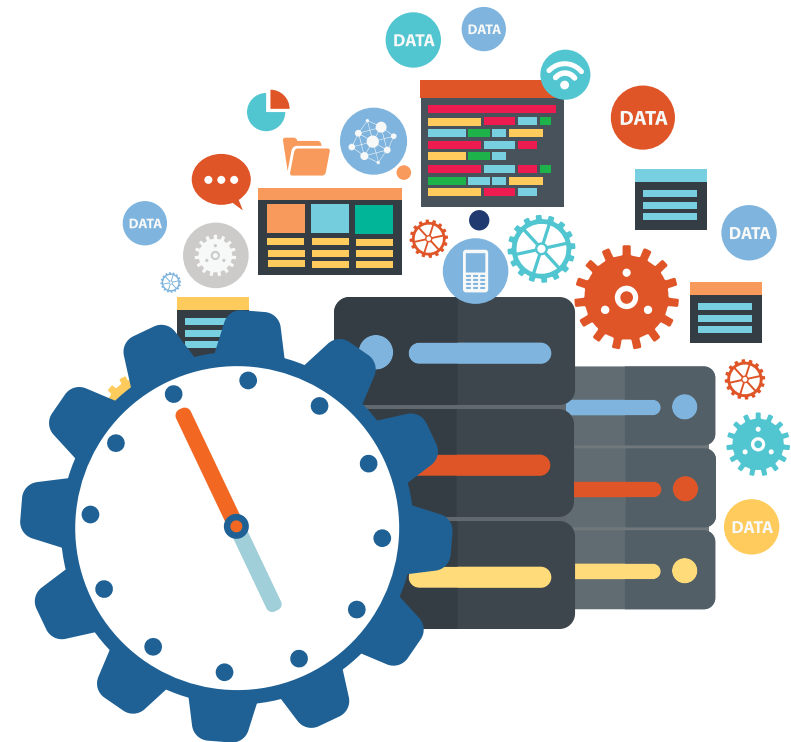
Discovery feeds the data intelligence that holds the knowledge required for professional productivity

Data discovery is important for business; connect the dots between business drivers and data attributes to uncover your own gaps.

Data professionals are working longer, but not smarter: invest in data intelligence processes and solutions to increase data knowledge.

- 50% of a data professional's time is unproductive!
 - » 30% of their time they do not have the data knowledge required find, prepare and govern data
 - » 20% of their time is wasted from a lack of knowledge

Time equals money: Build a business case to invest in data discovery.



Survey Demographics

