

Analytics Governance With Tableau Data Management

Presented By: James Chen – Tableau Solutions Engineering

Today's Agenda

- 1. Discuss Analytics Governance in Practice Today
- 2. Overview of Tableau's Data Management Toolset
- 3. Solving 8 Governance Challenges With Data Management
- 4. Next Steps and QA

Looking Back

These conversations have been going on for decades

By Andrew White | January 20, 2009 | 2 Comments

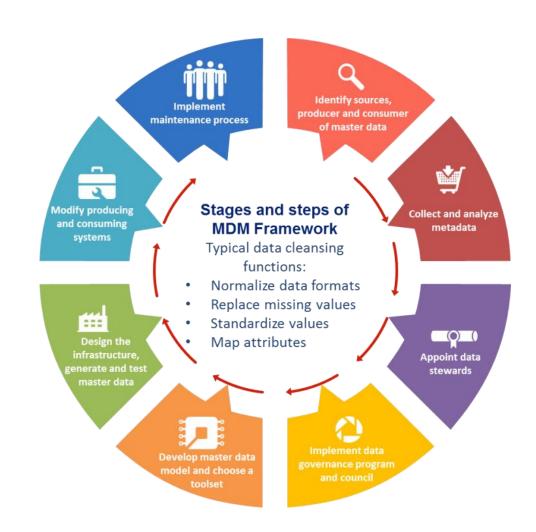
Product Information Management (PIM)

IBM

Customer Data Integration (CDI)

I was sent a link to a story in December, Winning over the MDM Chief Procurement Officer with MDM, (thanks Debbie) that I found interesting. The article is based on some vendor (Bristlecone Inc.) content (not always a bad thing), but the thing that I found most interesting was the source link referred to in this article (see Avoiding the Big Bang Backlash of MDM Implementations). Jessie Chimni, VP of North American Services for Bristlecone, wrote in his source article that MDM consisted originally of Product Information Management, and Customer Data Integration. Now why would I find this interesting?

https://blogs.gartner.com/andrew_white/2009/01/20/where-did-mdm-%E2%80%9Ccome-from%E2%80%9D-in-the-first-place/



Many of the Same Challenges Persist

Defining the Conflict

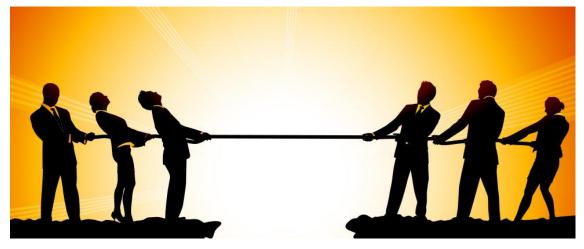
For any business, speed and scale are necessary for success. Business users are driven by the "power of *now*" -- they expect contextual, timely, and comprehensive data instantaneously in order to use business intelligence to make good decisions. For IT, however, matters are not as simple. Issues of governance, security, access control, backlogs, and cost play a large role in IT -- and legitimately so.

From: https://tdwi.org/articles/2017/08/22/biz-all-agility-and-governance-in-harmony.aspx

Common Challenges:

- We need more data
- We need different data
- We need data faster
- We need to experiment with the data*

Business Agility Vs Data Governance





Traditional Approaches Add Hurdles- What Could Go Wrong?



Added Labor:

- Approvals, CAB Review, Change Mgmt
- QA/Certification
- Policy Adherence
- Definitions & Dictionaries
- Data Centralization Requests
- Environmental Promotions

It's Hard to Recover From a Bad Experience with Governance



When the Business Faces:

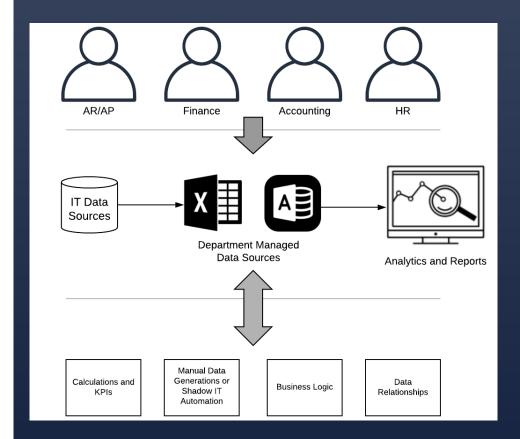
- Overly restrictive governance
- Missed deadlines
- Cumbersome processes
- Arbitrary mandates
- Slow data delivery velocity

They will find another way...

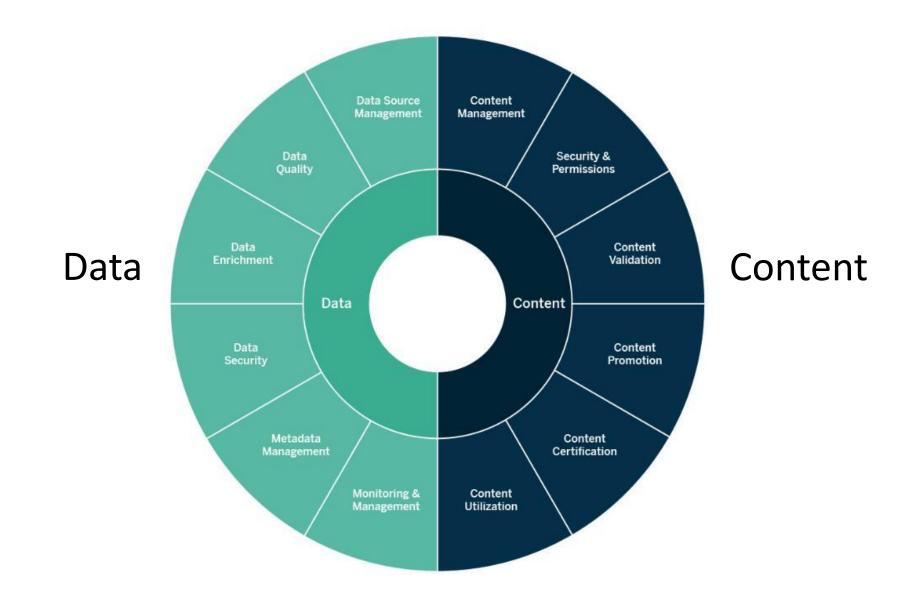
Which Leads to Ungovernable Practices

- Ungoverned spreadsheets + user-operated databases
- Untraceable data curation/preparation
- Ultimately shadow operations

Practical Reality of Departmental Analytics

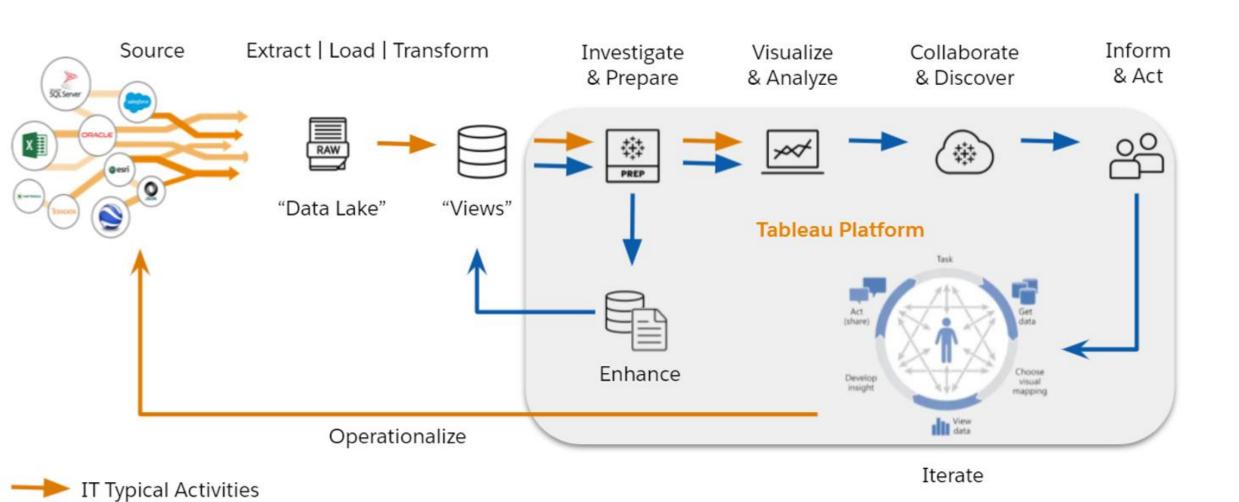


Tableau's Governance Framework has Changed the Paradigm



Data Delivery Lifecycle Across IT and the Business

Business Typical Activities



We are Creating a Convergence

FIND AND UNDERSTAND DATA

Data Cataloging

- Data Lineage
- Discovery
- Certification
- Warnings

CUSTOMIZE YOUR DATA

Data Preparation

- Clean
- Combine
- Reshape
- Quality Insights
- Change Tracking

Self-Service Data Mgmt

- Centralize your data & metadata
- Share it throughout the organization
- Manage compliance

Master Data

- Sync with EDC
- Apply Governance Policies

ACCESS DATA

Discover & Access

- Centralize Access
- Accelerate Queries

USAGE AND POPULARITY

Analytics

- Data Semantics
- Impact from Viz to Table

Simplifying the Lifecycle:

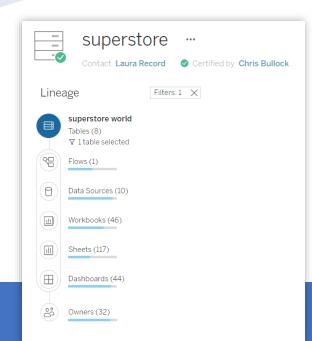
Creation → Oversight → Trust

Govern

Data is vetted + certified, Governance established



Catalog + Metadata API

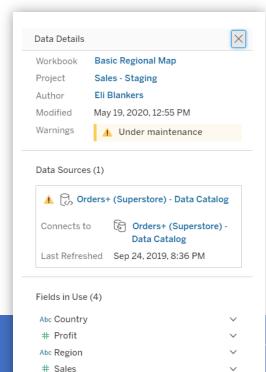


Verify

Cataloged discoverable data, End user trust + empowerment



Data Details

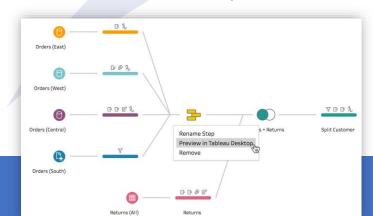


Create

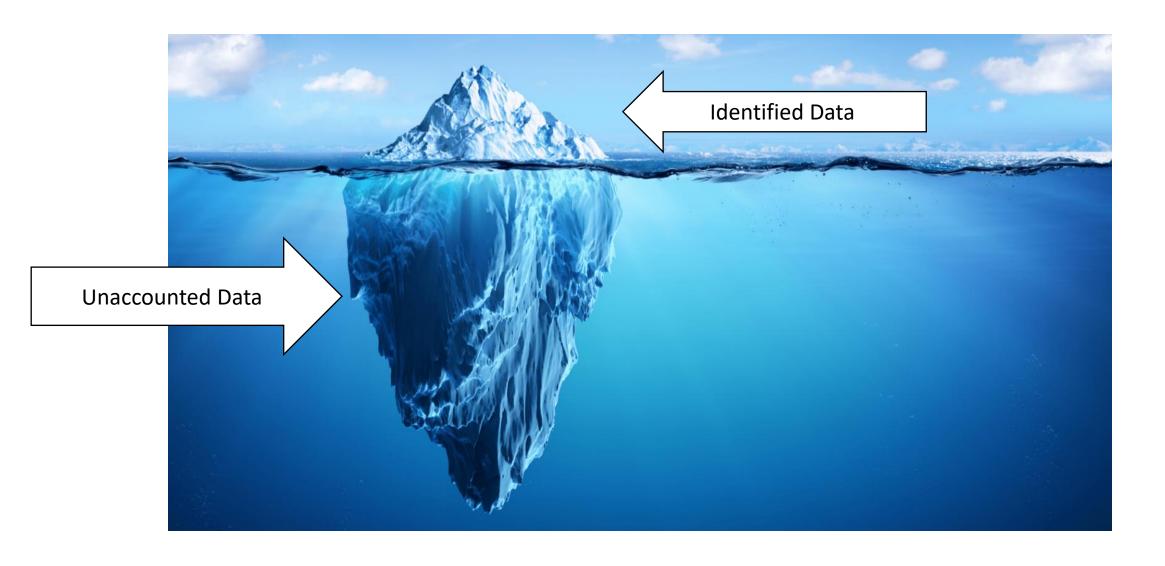
Rapid data development,
Data wrangled and transformed



Tableau Prep



Challenge 1: The Unknown Data Driving Analytics



The Spreadsheet Problem

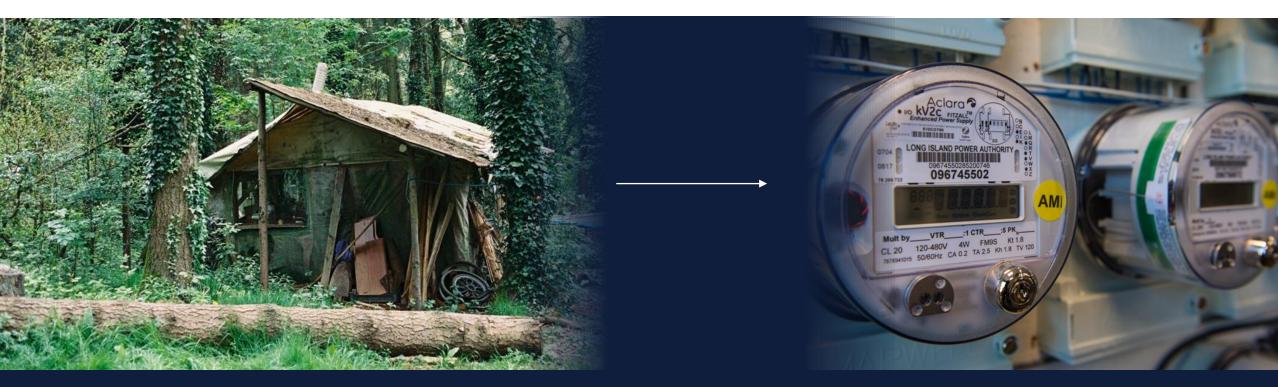


Widespread Spreadsheet Adoption Leads To:

- Data Quality Challenges
- Instant MSOT
- Broken Chain of Custody
- Lack of Auditability/Traceability
- Decentralization of Data
- Off the Grid KPI's/Business Logic/Automation

When The Receipts Are No Longer Optional

- Get ALL of your data on the grid know who did what, when, and where
- Encapsulate your data processes in an auditable/shareable manner
- Repeatable + Reliable + Governable
- Automatic metadata collection



Prep + Catalog Has the Answers

Where Did The Data Come From?

What Changes Were Were There Quality Made?

Issues?

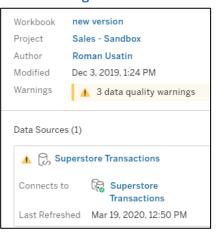
How Was the Business Logic Built?

Where Did The Data Go?

Prep Connection Info



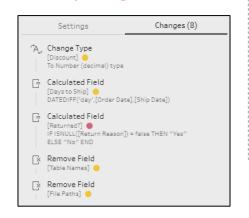
Catalog Data Details



Prep Change Tracking Indicators



Prep Change Details



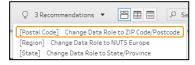
Prep Visual Join Results



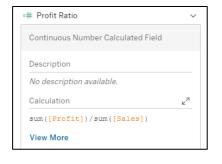
Catalog Quality Warnings



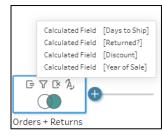
Prep ML Insights



Catalog Calc Details



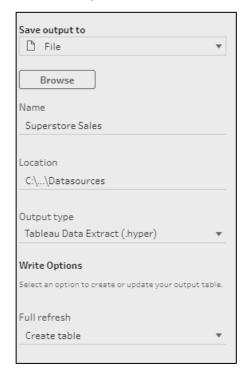
Prep Indicators



Prep Calculation Definitions



Prep Write to File, Database Table, Tableau Server



Challenge 2: Matching Tools with Skillsets



When the Wrong Tool is Selected

- Adoption fails
- Business goes back to the spreadsheet data economy
- Expensive shelf ware
- The same problems persist

Most Data Workers are Analysts, Not Engineers

10.6 Million

LinkedIn Users with SQL Skills*



52.3 Million

LinkedIn Users with Excel Skills*



Prep Builder Makes Self-Service ETL Easy

Abc T Abc @ Row ID 2K Customer ID 512 Customer Name 512 Order ID 822 Ship Mode 4 CA-2015-100293 AA-10375 Standard Class Aaron Hawkins CA-2015-100706 Second Class AA-10480 Aaron Smayling First Class CA-2015-100895 Adam Bellavance AA-10645 Same Day Second Class CA-2015-100916 AB-10060 Adam Hart 4,000 329 rows CA-2015-101266 AB-10105 Adam Shillingsburg CA-2015-101560 Adrian Bartson AB-10165 6.000 CA-2015-101770 AB-10255 Adrian Hane CA-2015-102274 AB-10600 Alan Barnes 8,000 CA-2015-102673 AC-10450 Alan Haines CA-2015-102988 AF-10870 Alan Hwang 10.000 AF-10885 CA-2015-103317 Alan Schoenberger CA-2015-103366 AG-10330 Alan Shonely Order ID Discount Region State Ship Mode Customer ID **Customer Name** Segment 0 South 183 CA-2015-158274 RM-19675 Robert Marley Louisiana Second Class Contractor 0 South Louisiana 184 CA-2015-158274 Second Class RM-19675 Robert Marley Contractor

Second Class

Standard Class

RM-19675

EA-14035

Robert Marley

Erin Ashbrook

Contractor

Corporate

185 CA-2015-158274

231 US-2015-156216

0 South

0.7 South

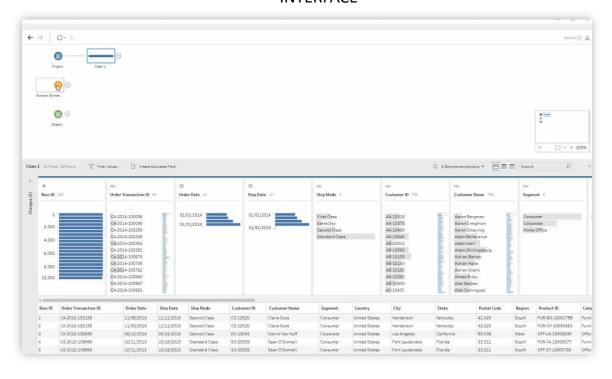
Louisiana

North Carolina

VISUAI DATA

ANALYSIS

DRAG AND DROP INTERFACE



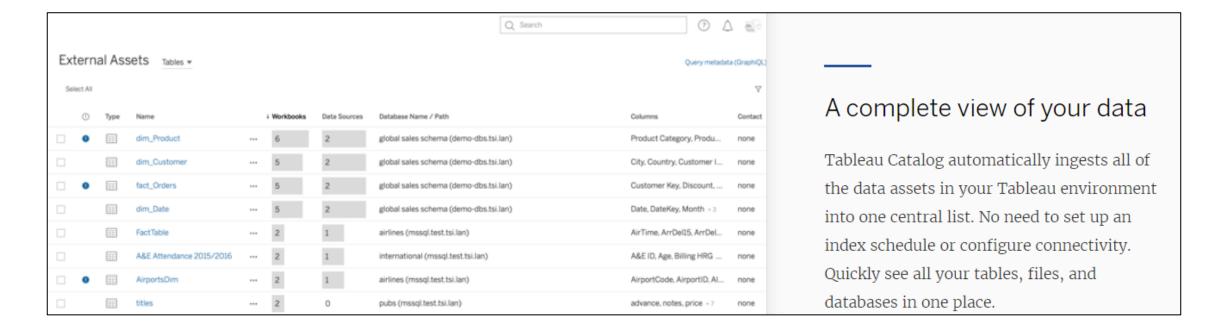
Challenge 3: Finding the Right Data

- How can I find a dataset created from the proper source of truth?
- How would I know if the source database/table is safe to use?
- How will I know if there are issues on the source databases and tables?
- When there are 20 Sales data sources – how will I know which, one is right?

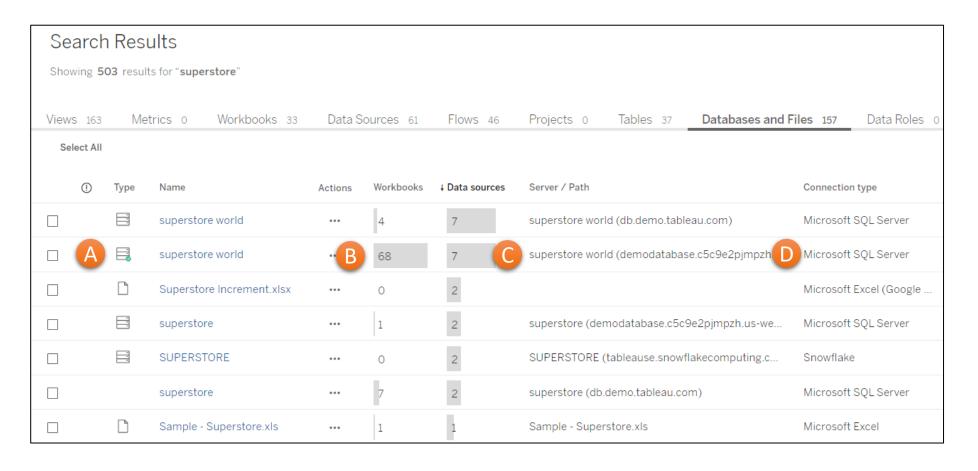


Catalog Makes Finding the Right Data Fast and Easy

- Search, Browse, Sort, Filter external + file-based connections
- Automatic ingestion of metadata from Extracts, Live Connections, and Prep Flows
- Observe data source usage, certifications and warnings



The Right Details to Drill Down



A- Is the data certified or have a quality warning?

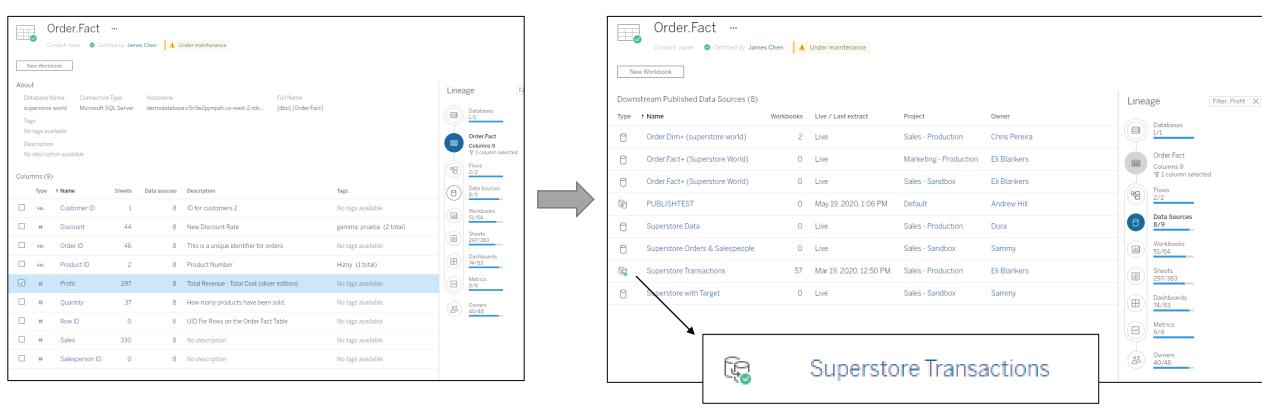
B- Is this data source actively in use?

C- Where does this data come from?

D- What kind of connection is it?

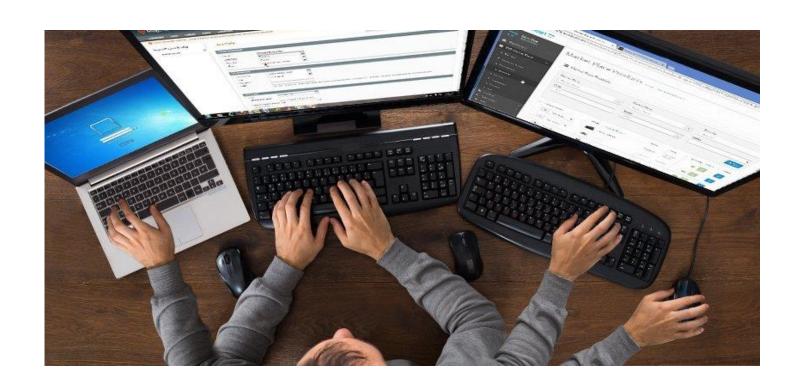
Surf the Lineage to Find Existing Assets

Prevent MSOT by finding the data you are looking for using Catalog's relational lineage



Database → Table → Field → Related Data Sources → Certified Data Source

Challenge 4: Consolidating the User Experience



Eliminate Swivel Chairing



The Dangers of Disjointed Experiences

Why is it important to place key information at the point of usage?

Let's imagine that your vehicle's check engine, TPMS, and low fuel indicators required a separate interface – perhaps a phone app

- Would you check it every time you drive?
- Would you get run out of gas on the side of the road?
- Would you drive a vehicle that is potentially dangerous to operate?
- How much do these probabilities increase?



Certifications and Quality Warnings in Tableau

Databases, Tables, and Data Sources can be certified or have a quality warning applied

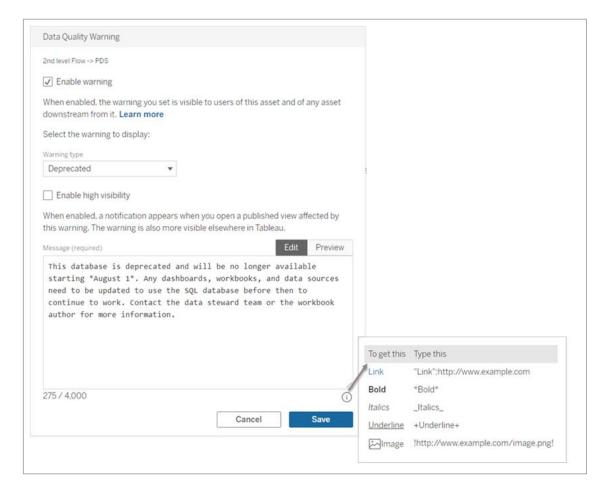
Certifying a Data Source



Certification Icon and Description



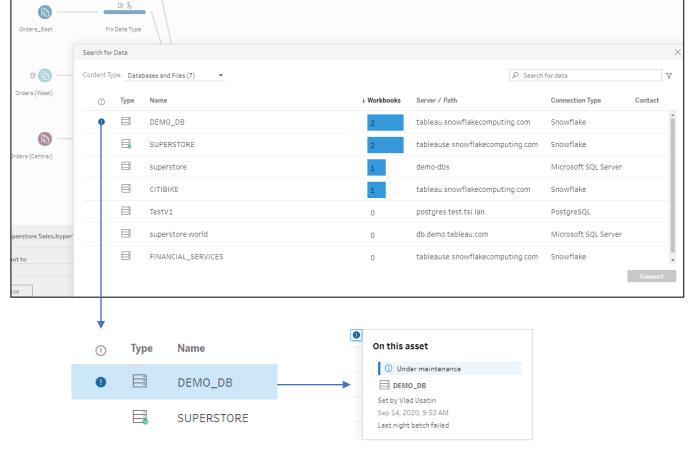
Setting a Quality Warning



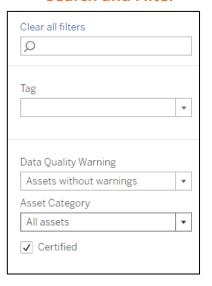
Critical Information At the Point of Consumption

Catalog Puts the Metadata in the Creator and Consumer's Flow

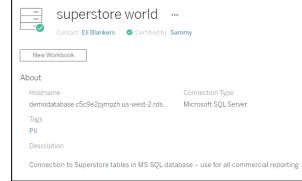
Data Catalog Within Prep and Desktop



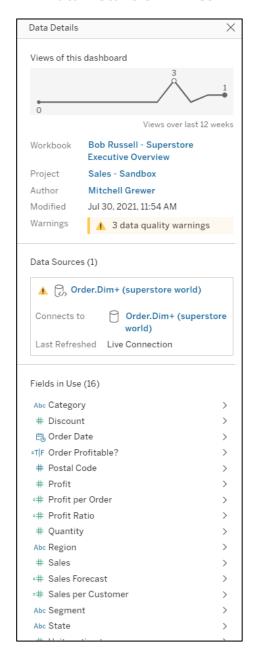
Search and Filter



Web Authoring From Catalog



Data Details on Vizzes



Challenge 5: Keeping Metadata Synced

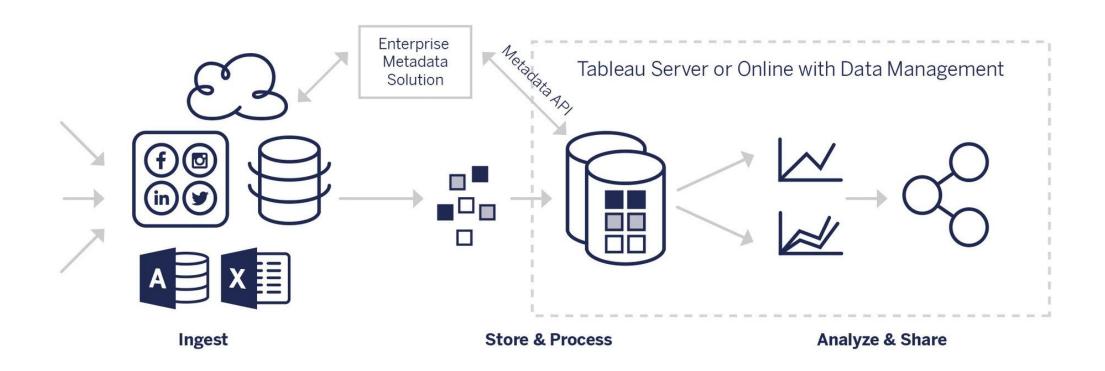
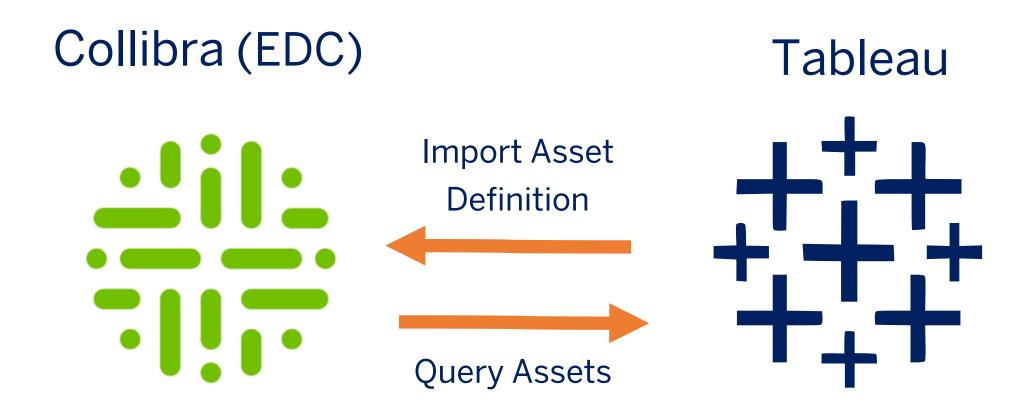


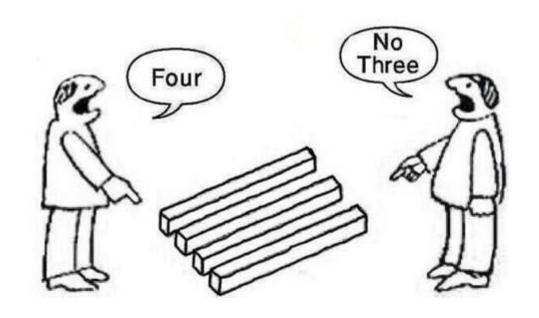
Tableau to EDC Via Metadata API



EDC Metadata to Tableau Via REST API



Challenge 6: Establishing Data Trust



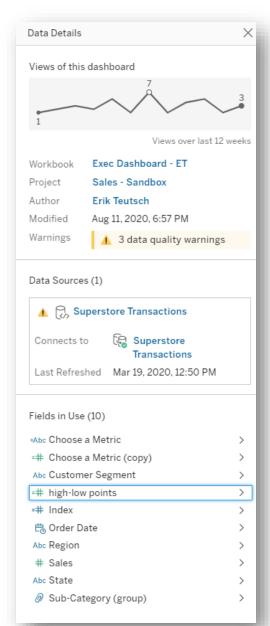
Do your executives sound like this?

Do your users trust what they see?

Say it's February 28.
WOULD YOU DRINK THIS MILK?

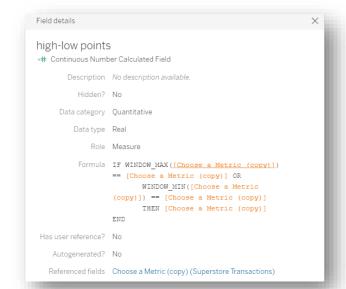
What verification information do you provide today?

Data Details Provides Consumers With the Receipts





- Where did this data come from?
- Are the underlying data sources certified?
- Are there Data Quality Warnings on these data sources?
- When was the data last refreshed?
- What fields were used from which tables?
- How were the calculations and KPI's created?



Challenge 7: Managing Sensitive Data

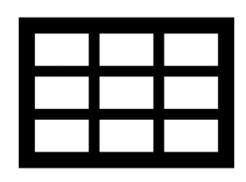


Mistakes Happen Often With Ungoverned Data

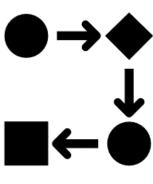


Tableau ProServ PII/PHI Data Manager to the Rescue

PII / PHI Data Manager



Find PII / PHI Content



Act on PII Content



Email Notifications



Audit Data sources & Dashboards

Search → Tag → Flag → Remove → Analyze → Automate

Say Goodbye to Sensitive Data Messes



PII/PHI — Data Manager

Challenge 8: Communicating Impact

Uncommunicated and under-communicated changes to source data lead to confusion, incorrect dashboards, and ultimately a deterioration of data trust

This can take many forms:

- Fields added/removed from tables
- Field/KPI definitions changed
- Data models altered (joins/filters/scope/tables)
- Database/schema/table deprecation
- Data under maintenance

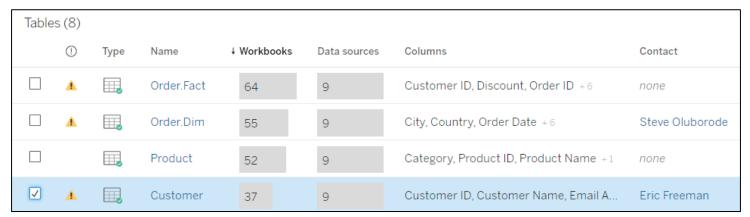


What Happens When Changes Are Uncommunicated?



Impact Analysis and Downstream Notification

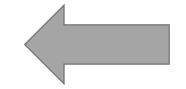
Filter by Database, Table, or Field



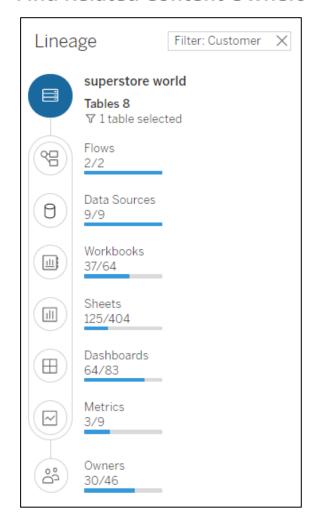


Proactively Alert All Downstream Owners



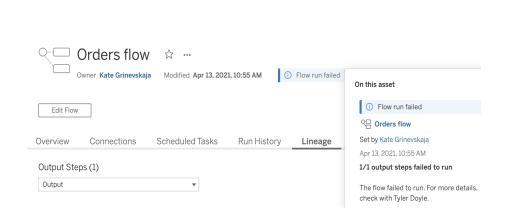


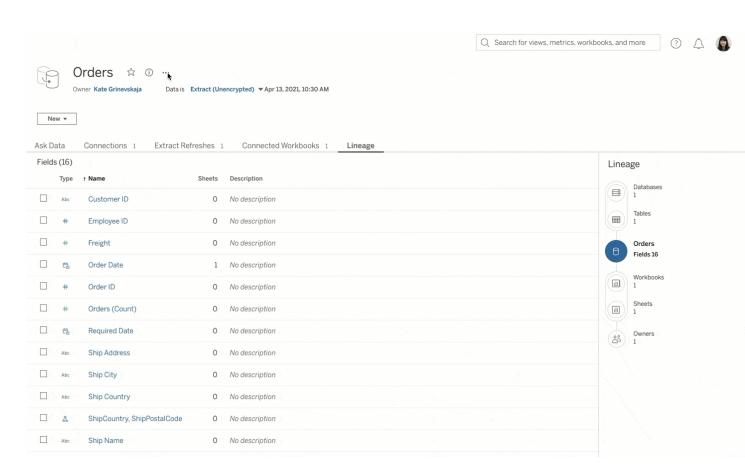
Find Related Content Owners



Automatic Data Quality Warnings

Tableau communicates when data sources encounter problems via Automatic Data Quality Warnings for extract and Prep flow failures





Recap

Whew – that was a lot! What did we just discuss:

- The misalignments of traditional approaches to analytics data governance
- How Tableau Data Management solves numerous hard-to-solve governance challenges
- The value for data workers + consumers of having governance baked directly into their analytics experience

Let's Hear Your Questions

