



DATA DAY OUT

Prepare & Manage Your Data with Tableau

Graeme Britz

Senior Product Manager, Data Management
Tableau

Agenda

- Visually clean and shape data with Tableau Prep Builder
- Operationalize data prep with Prep Conductor
- Understand data usage, communicate trusted data, and find data with Tableau Catalog

Order Number	Customer Identifier	Order Status	Reque
201766-6300005	C00005	Complete	
201766-6300014	C00014	In Transit	
201766-6300017	C00017	Processing	
201766-6300024	C00024	Complete	
201766-6300026	C00026	Complete	
201766-6300054	C00054	In Transit	
201766-6300055	C00055	Complete	
201766-6300083	C00083	Complete	
201766-6300105	C00105	In Transit	
201766-6300106	C00106	In Transit	
201766-6300115	C00115	Complete	
201766-6300126	C00126	In Transit	
201766-6300135	C00135	Complete	
201766-6300138	C00138	In Transit	
201766-6300141	C00141	Complete	
201766-6300142	C00142	In Transit	
201766-6300144	C00144	In Transit	
201766-6300152	C00152	In Transit	
201766-6300157	C00157	In Transit	
201766-6300161	C00161	In Transit	
201766-6300165	C00165	Processing	
201766-6300171	C00171	In Transit	
201766-6300177	C00177	In Transit	
201766-6300185	C00185	In Transit	
201766-6300187	C00187	Processing	
201766-6300191	C00191	In Transit	
201766-6300197	C00197	In Transit	
201766-6300222	C00222	Complete	
201766-6300224	C00224	In Transit	
201766-6300225	C00225	In Transit	
201766-6300232	C00232	In Transit	
201766-6300235	C00235	Complete	
201766-6300255	C00255	Complete	
2017512-600416	C00416	Complete	
2017515-600416	C00416	Complete	
2017515-600417	C00417	Complete	
2017516-600563	C00563	Complete	
2017516-600566	C00566	Complete	
2017516-600568	C00568	Complete	
2017516-600568	C00568	Complete	
2017516-600571	C00571	Complete	
2017516-600596	C00596	Complete	
2017516-600598	C00598	Complete	
2017516-600611	C00611	Complete	
2017516-600621	C00621	Complete	
2017516-600625	C00625	Complete	
2017516-600631	C00631	Complete	

```

WITH cteSource ( StaffMember, EventDateTime, EventType
AS ( SELECT StaffMember ,
CAST(EventDate AS DATETIME)
+ CAST(EventTime AS DATETIME)
EventType ,
ROW_NUMBER() OVER ( PARTITION BY
ORDER BY EventDate, Even
FROM dbo.StaffHours
),
cteGrouped ( StaffMember, EventDateTime, grp )
AS ( SELECT c.StaffMember ,
c.EventDateTime ,
ROW_NUMBER() OVER ( PARTITION BY c
ORDER BY c.EventDateT
- 1 AS grp
FROM cteSource AS c
LEFT JOIN cteSource AS p
ON p.StaffMember = c.StaffMember
AND p.rn = c.rn - 1
LEFT JOIN cteSource AS n
ON n.StaffMember = c.StaffMember
AND n.rn = c.rn + 1
WHERE ( COALESCE(p.EventType, 'Exit') = '
AND c.EventType = 'Enter'
)
OR ( c.EventType = 'Exit'
AND COALESCE(n.EventType, 'Ent
)
SELECT StaffMember ,
MIN(EventDateTime) AS EventDateTime ,
MAX(EventDateTime) AS ExitDateTime ,
DATEDIFF(SECOND, MIN(EventDateTime), MAX(Eve
CONVERT(CHAR(8), DATEADD(SECOND,
DATEDIFF(SECOND, MIN(Ev
MAX(EventDateTime)),
'00:00:00'), 8) A
FROM cteGrouped
GROUP BY StaffMember ,
grp / 2
HAVING COUNT(*) = 2;

```

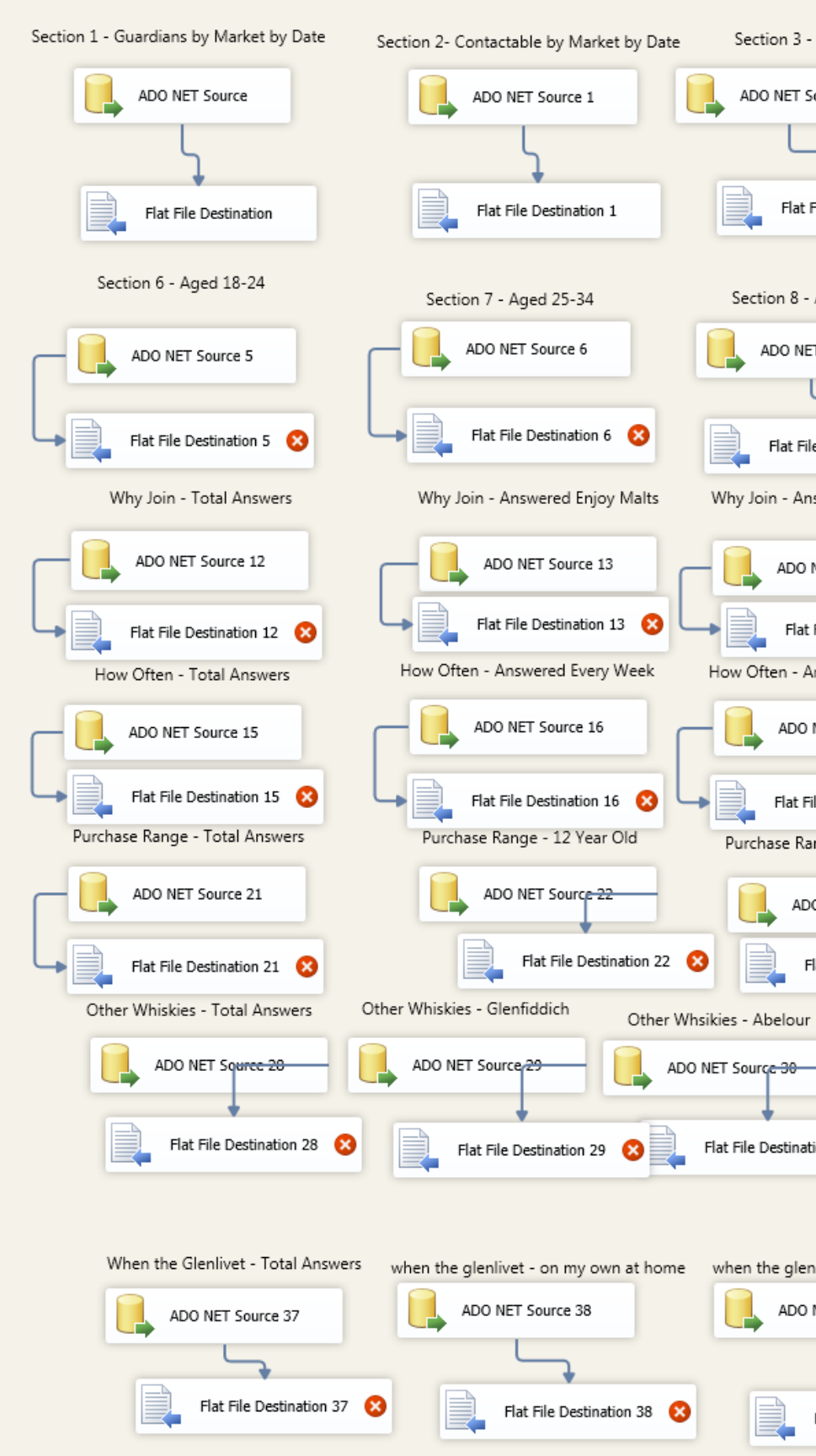


Tableau Prep Builder

Helps people get data ready for analysis

The screenshot displays the Tableau Prep Builder interface for a join operation. At the top, it shows 'Returns (All)' and 'Returns' tabs. The main configuration area is titled 'Orders + Returns' with '29 Fields' and '14K Rows'. Under 'Applied Join Clauses', 'All Orders' and 'Returns' are joined on 'Product ID'. The 'Join Type' is set to 'Inner join'. A Venn diagram shows the relationship between 'All Orders' and 'Returns'. The 'Summary of Join Results' table shows 2,332 rows included from 'All Orders', 275 rows included from 'Returns', and 0 rows excluded from either. The total 'Join Result' is 3,179 rows. A 'Days to Ship' bar chart shows a distribution from 0 to 7. A 'Table Names' list includes 'Orders (South)' and several 'adhoc-mssql2016.test.tsi.l...' entries. A 'Product ID' list shows various alphanumeric strings. A 'Join Clause Recommendations' section is visible at the bottom.

Days to Ship	Table Names	Product ID
4	adhoc-mssql2016.test.tsi.lan/Orders/orders_west	OFF-ST-10000129
4	adhoc-mssql2016.test.tsi.lan/Orders/orders_west	OFF-ST-10000129
4	adhoc-mssql2016.test.tsi.lan/Orders/orders_west	OFF-ST-10000129

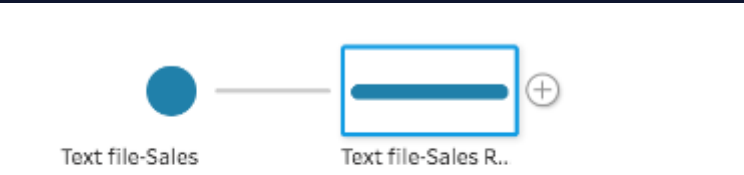
Days to Ship	Table Names	Product ID
0	Orders (South)	FUR-BO-10000112
0	adhoc-mssql2016.test.tsi.l...	FUR-BO-10000330
1	adhoc-mssql2016.test.tsi.l...	FUR-BO-10000468
2	adhoc-mssql2016.test.tsi.l...	FUR-BO-10000711
3		FUR-BO-10000780
4		FUR-BO-10001337
5		FUR-BO-10001519
6		FUR-BO-10001567
7		FUR-BO-10001601
		FUR-BO-10001608
		FUR-BO-10001619
		FUR-BO-10001798

Summary of Join Results
All Orders: 2,332 (Included)
Returns: 275 (Included)
Excluded: 0
Join Result: 3,179



Tableau Prep Design Principles

Visual & Direct



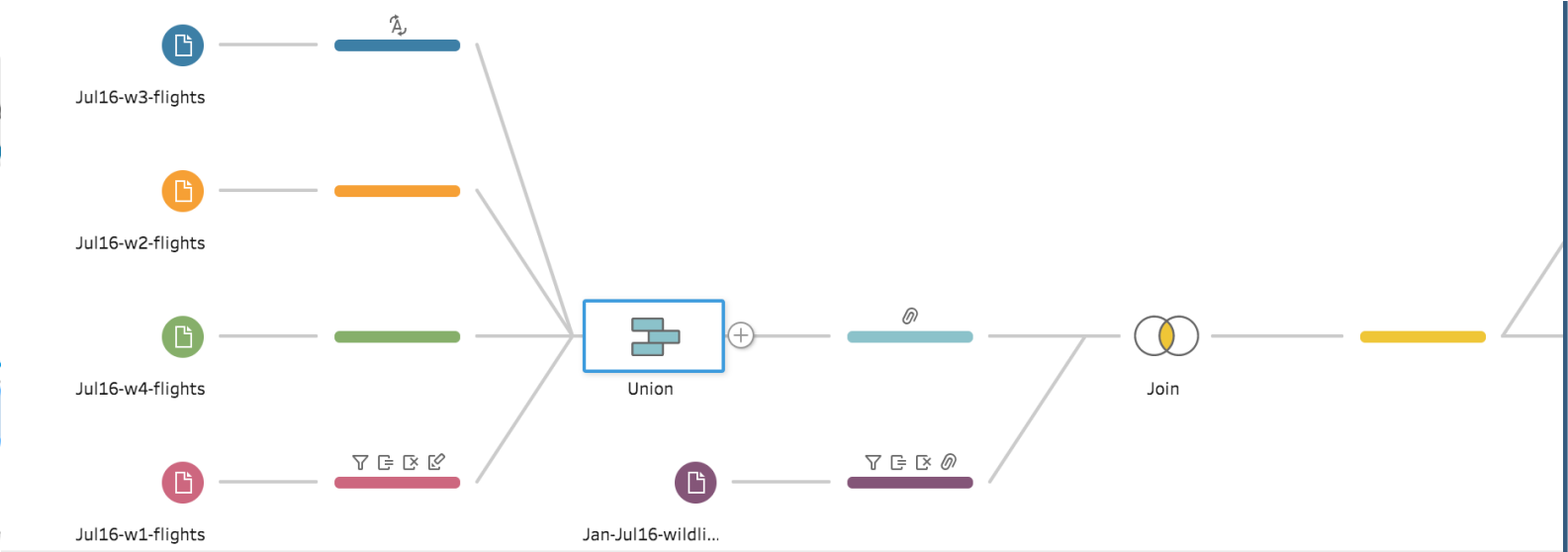
Text file-Sales Results 17 Fields 6K Rows ✓ Keep Only

City 109 State 15

City: Akron, Allentown, Altoona, Andover, Atlantic City, Auburn, Baltimore, Bangor, Bayonne, Belleville, Bethlehem

State: Connecticut, Delaware, District of Columbia, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Ohio, Pennsylvania

Product ID	Ship Mode	Shipping Costs
AAA-AA-00000000	Same Day	-99
AAA-AA-00000000	Same Day	-99
AAA-AA-00000000	Same Day	-99
AAA-AA-00000000	Same Day	-99
AAA-AA-00000000	Same Day	-99
AAA-AA-00000000	Same Day	-99
AAA-AA-00000000	Same Day	-99



Union 14 Fields 502K Rows

Inputs

- Transform 1
- Transform 2
- Transform 3
- Transform 4

Union Results Show only mismatched fields

Origin 304, Wheels On 1K, Flight Date 31

ABE, ABI, ABQ, ABR, ABY, ACK, ACT, ACV, ACY, ADK, ADQ, AEX

Wheels On: null, 250, 625, 1,000, 1,375, 1,750, 2,125, 2,500

Flight Date: 06/27/2016, 08/01/2016

Resulting Fields
1 Mismatching fields from 14 resulting fields.

Applied Join Clauses

Clean Notes/Approver = All Orders

Product ID = Product ID

Order ID = Order ID

Join Type: Right join

Click the graphic to change the join type.

Clean Notes/Approver All Orders

Summary of Join Results

Click the bar segments to view the included and excluded values.

Mismatched values

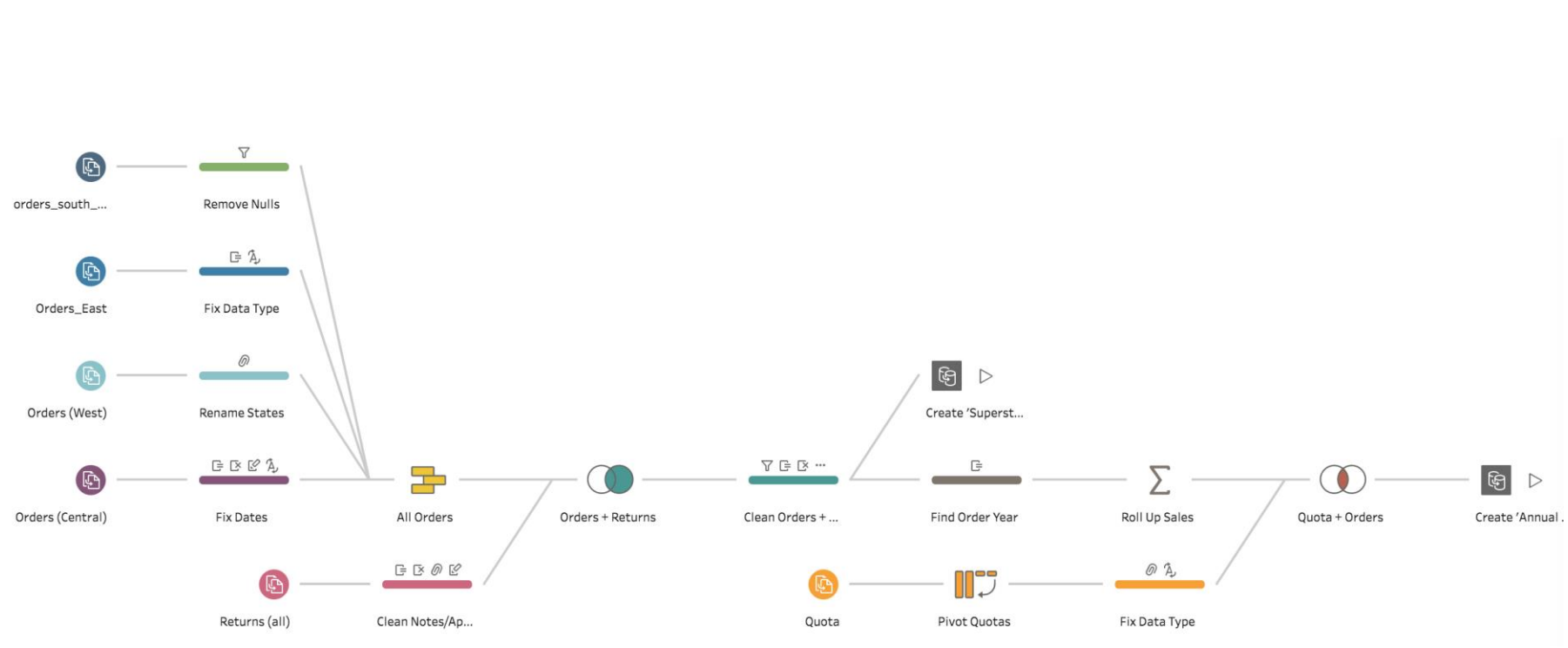
	Included	Excluded
Clean Not...	272	3
All Orders	16,302	0
Join Result	16,302	

Join Clause Recommendations

Product ID = Product Name

Smart

Add algorithms to augment human intelligence



Group and Replace

Aircraft: Airline/O... 111

Group values by
Pronunciation

- SKY REGIONAL AIR...
- SKYWEST AIRLINES
- SOUTHWEST AIRLI...**
- SPIRIT AIRLINES
- SUN COUNTRY AIR...
- TACA INTL AIRLINE...
- TAM LINHAUS AER...
- TRANS STATES AIR...
- TURKISH AIRLINES
- UNITED AIRLINES
- UNKNOWN

SOUTHWEST AIRLINE.. replaces:
SOUTHWEST AIRLINES
south west airlines
SOUTHWEST AIR
Southwest Airlines
SOUTHWEST AIRLINES

ABX AIR
AER LINGUS
AEROLITORAL
AEROMEXICO
AIR CANADA
AIR CANADA ROUGE
AIR CHINA CARGO
AIR CHOICE ONE

Integrated

The screenshot shows the Tableau Project Maestro interface. On the left, a 'Connections' panel lists various data sources. A 'Connect' menu is open, showing options for connecting to files (Excel, Text File) or servers (Amazon Aurora, Amazon Redshift, Aster Database, EXASQL, Google Cloud SQL, HP Vertica, IBM BigInsights, IBM DB2, IBM PDA (Netezza), Kognitio, Marklogic, MemSQL, Microsoft SQL Server, Monet DB, MySQL, Oracle, Pivotal Greenplum Database, PostgreSQL, Splunk, SAP Sybase IQ, SAP Sybase ASE, Teradata). In the center, an 'Add Field' dialog is open, showing a search bar and a list of fields including ABS, ACOS, ASCII, ASIN, ATAN, ATAN2, CEILING, CHAR, CONTAINS, and COS. On the right, a bar chart titled 'Units Sold' is displayed, showing a distribution of values. At the bottom, a data table is visible with columns for 'Time', 'Day', 'Sales', 'Product ID', 'Product Name', 'Segment', 'City', and 'State'.

Time	Day	Sales	Product ID	Product Name	Segment	City	State
10 AM	Same Day	-99	0	AA-00000	Consumer	null	nl
10 AM	Same Day	-99	0	AA-00000	Consumer	null	nl
10 AM	Same Day	-99	0	AA-00000	Consumer	null	nl

- Open in Tableau
- Common Tableau Platform
- Publish to Tableau (curation, recommendations)

Operationalize with Prep Conductor

Publish Flows

Publish Flow to Tableau Server

Project
Default

Name
Sales

Description
Contains sales and returns

Tags
sales data Add

Files ⓘ
6 file connections Edit
6 uploaded files
0 direct connections

Cancel Publish

Schedule Refreshes

New Task

Select a schedule to run the flow "Sales".

Daily — Daily at 10:00 PM

First of the month 1:00AM — On the last day of the month at 6:...

Weekday 2:00AM — Weekly at 7:00 PM on Sunday, Monday, T...

Daily — Daily at 10:00 PM

Every half hour — Every 30 minutes

01 — Every hour

Create 'Annual Regio...
http://tpqawin01

Create 'Superstore Sa...
http://tpqawin01

Cancel Create Task

Alerts & See History

Run End	Status	Errors
	Pending	
	Pending	
Oct 13, 2018, 11:57 AM	✓ Succeeded	
Oct 13, 2018, 11:57 AM	✓ Succeeded	
Oct 13, 2018, 11:57 AM	✓ Succeeded	
Oct 13, 2018, 11:57 AM	✓ Succeeded	
Oct 13, 2018, 11:56 AM	✓ Succeeded	
Oct 13, 2018, 11:56 AM	✓ Succeeded	
Oct 13, 2018, 11:56 AM	✓ Succeeded	
Oct 13, 2018, 11:56 AM	✓ Succeeded	
Oct 13, 2018, 11:55 AM	! Failed	1 error

Demo

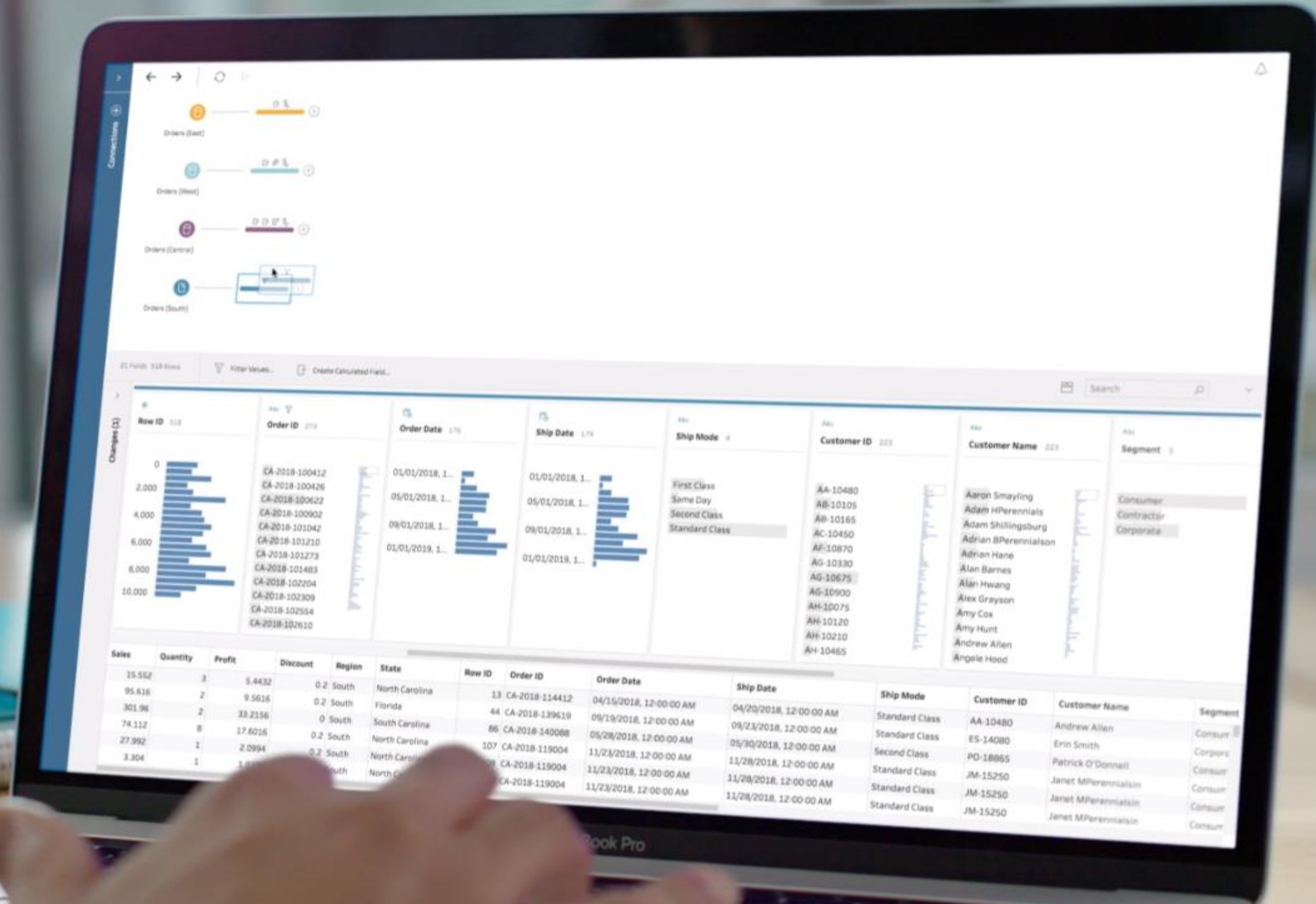


Tableau Catalog

Visibility

Trust

Discovery

Extensibility

Tableau Catalog

Visibility

tableau.com

Explore / Technical Support Official / Support - Case Master

Search

Support - Case Master

Owner: Brett Megaard | Certified by: Brett Megaard | Unencrypted Extract | Jul 24, 2019, 2:44 PM

Intended Audience: Entire company (primarily Core Technical Support and Major Account Care) Business Use: Case Master is the source of truth for information about Technical Support...[read more](#)

New Workbook

Ask Data | Connections 3 | Extract Refreshes 1 | Connected Workbooks 443 | **Lineage**

Fields (623)

Type	Name	Sheets	Description
-#	# of Customer Touches	44	Number of customer touches based on the last customer contact field in case history plus the n...
#	# of Customer Touches (old)	79	Number of customer touches based on the las customer contact field in case history. How many ...
Abc	Account ID	31	Account ID 18-character ID of account in SFDC Source: Account.ID
#	Account Key	53	Account Key Æ Surrogate key of account, stored in dimAccount. Æ Source:
Abc	Account Name	436	Account Name Name of Account Source: Account.Name
Abc	Account Owner	151	No description
Abc	Account Type	22	Account Type Name of the REcord Type of the account Source: The name of the Record Type (Re...
Abc	Account_Escalation__c	22	No description
Abc	Account_Message__c	0	No description
#	Account_Message_End_Date__c	0	No description

Lineage

- Databases (1)
- Tables (24)
- Support - Case Master** (Fields: 623)
- Workbooks (471)
- Sheets (3,599)
- Owners (125)

Tableau Catalog

Visibility

Trust

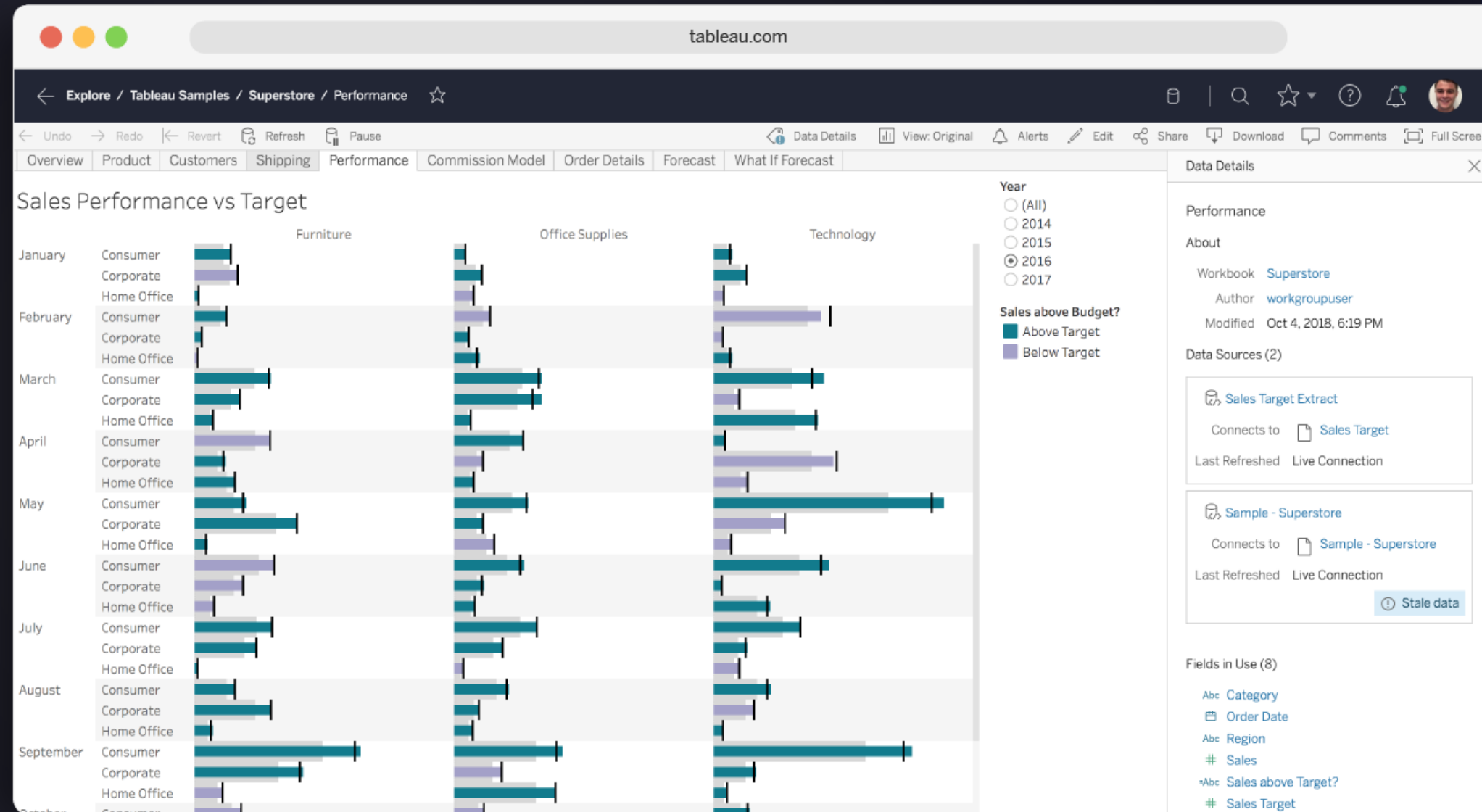


Tableau Catalog

Visibility

Trust

Discovery

The screenshot shows the Tableau Catalog interface. At the top, there's a navigation bar with 'New Workbook' and menu items: 'File', 'Data', 'Worksheet', 'Dashboard', 'Analysis', 'Map', 'Format', 'Help'. The main area is titled 'Connect to Data' and contains a search bar and a table of data sources. A dropdown menu is open over the 'Connection Type' column, showing options: 'Data Sources', 'Databases and Files', and 'Tables'. The 'store' row is highlighted in blue. Below the table, there is a descriptive text block and a 'Connect' button.

Type	Name	Workbooks	Data Sources	Server / Path	Connection Type
Table	EFG_Institution	3	0	sandboxdb_andrew ()	Microsoft SQL Server
Table	RSN_AD_SALES_LEADERBOARD	3	0	RSN_AD_SALES_LEADERBOARD (local co...	Teradata
Table	stores	2	0	pubs (mssql.test.tsi.lan)	Microsoft SQL Server
Table	store	2	3	tpcds (mssql.test.tsi.lan)	Microsoft SQL Server dbeck...
Table	titles	2	0	pubs (mssql.test.tsi.lan)	Microsoft SQL Server
Table	workbooks	2	0	workgroup (datacatstaging)	PostgreSQL
Table	promotion	2	3	tpcds (mssql.test.tsi.lan)	Microsoft SQL Server

This table has information about all of the store locations for our organization. Suggested use includes joining it to the customer and orders tables in order to compare the performance across different stores our company owns.

[Connect](#)

Tableau Catalog

Visibility

Trust

Discovery

Extensibility

The screenshot displays a GraphQL IDE interface with three main sections: a query editor on the left, a JSON response viewer in the center, and a metadata sidebar on the right.

GraphQL Query (Left):

```
1 {
2   databaseTables {
3     downstreamSheetsConnection {
4       totalCount
5       nodes {
6         name
7       }
8     }
9
10    downstreamWorkbooksConnection {
11      totalCount
12      nodes {
13        id
14      }
15    }
16
17    downstreamFlowsConnection {
18      totalCount
19      nodes {
20        name
21      }
22    }
23
24    downstreamDatasourcesConnection {
25      totalCount
26      nodes {
27        name
28      }
29    }
30  }
31 }
32 }
```

JSON Response (Center):

```
{
  "data": {
    "databaseTables": [
      {
        "downstreamSheetsConnection": {
          "totalCount": 17,
          "nodes": [ ]
        },
        "downstreamWorkbooksConnection": {
          "totalCount": 5,
          "nodes": [
            {
              "id": "09379ea1-b9b4-296d-8c4f-b595a5d14822"
            },
            {
              "id": "ab826b5b-30e6-6bc3-a52d-47ddf02bde4a"
            },
            {
              "id": "ad305dfe-7457-62d6-60cc-55e31dc99b03"
            },
            {
              "id": "d2c90de1-481c-8e3a-95b9-555d0d0cba0e"
            },
            {
              "id": "d9766356-52e4-20bb-c23c-17ab5494db44"
            }
          ]
        },
        "downstreamFlowsConnection": {
          "totalCount": 0,
          "nodes": [ ]
        },
        "downstreamDatasourcesConnection": {
          "totalCount": 2,
          "nodes": [
            {
              "name": "SuperStore Orders (Master)"
            },
            {
              "name": "Global SuperStore"
            }
          ]
        }
      }
    ]
  }
}
```

Metadata Sidebar (Right):

- Search DatabaseTable...
- table that is contained in a database
- IMPLEMENTS
- Table
- Warnable
- FIELDS
- id: ID! - Unique identifier used by the metadata API. Not...
- vizportalId: String! - Vizportal ID of this table, for use in client-to-ser...
- luid: String! - Locally unique Identifier used for the REST API ...
- name: String - Name shown in server and desktop clients
- IsEmbedded: Boolean - True if this table is embedded in Tableau content
- schema: String - Name of table schema
- fullName: String - Fully qualified table name
- description: String - User modifiable description of this table
- contact: TableauUser - Contact for this table
- IsCertified: Boolean - True if this database table has been marked as ...

Catalog Demo

Questions?