Anushka Anand

Senior Product Manager, Tableau

aanand@tableau.com

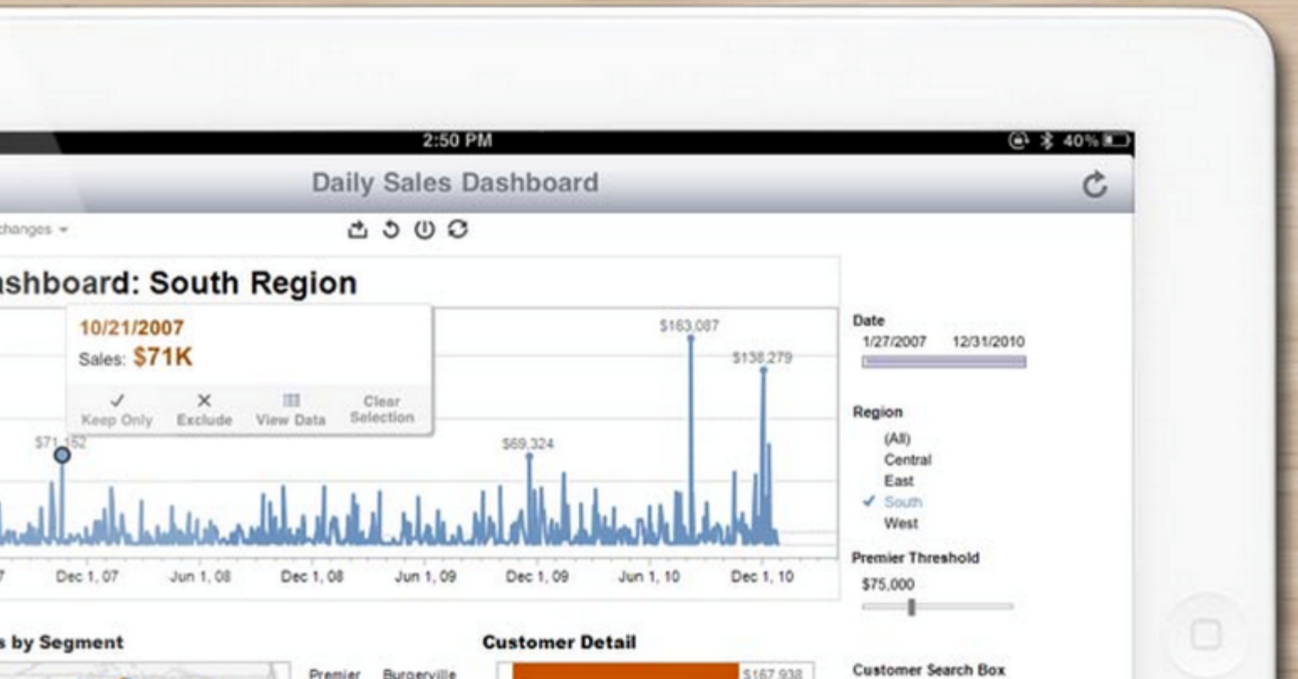
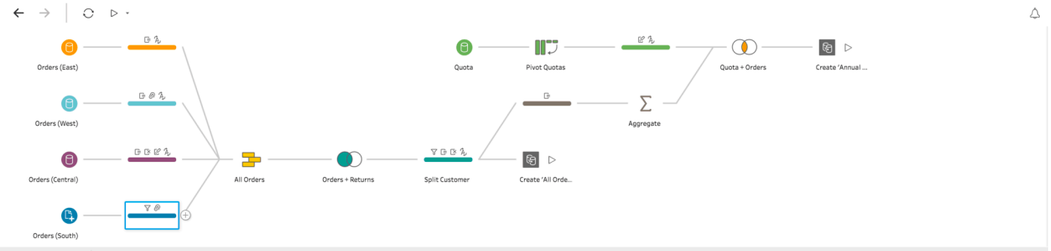
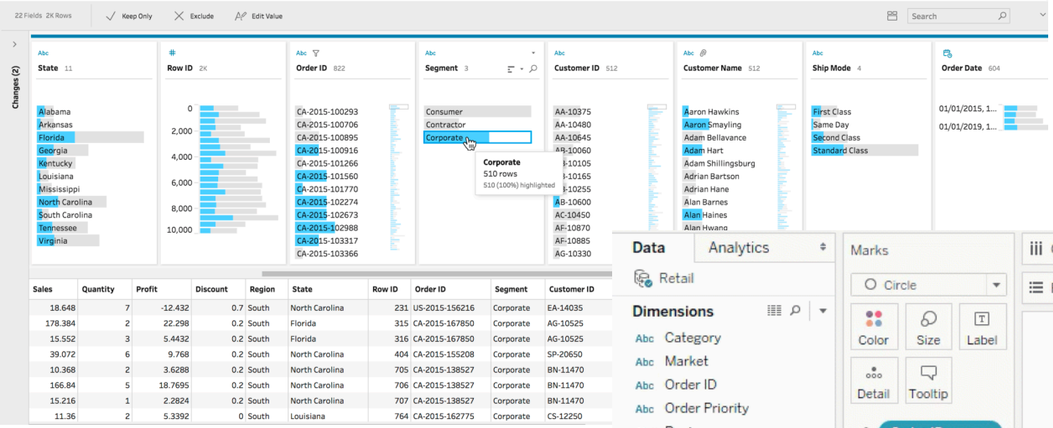
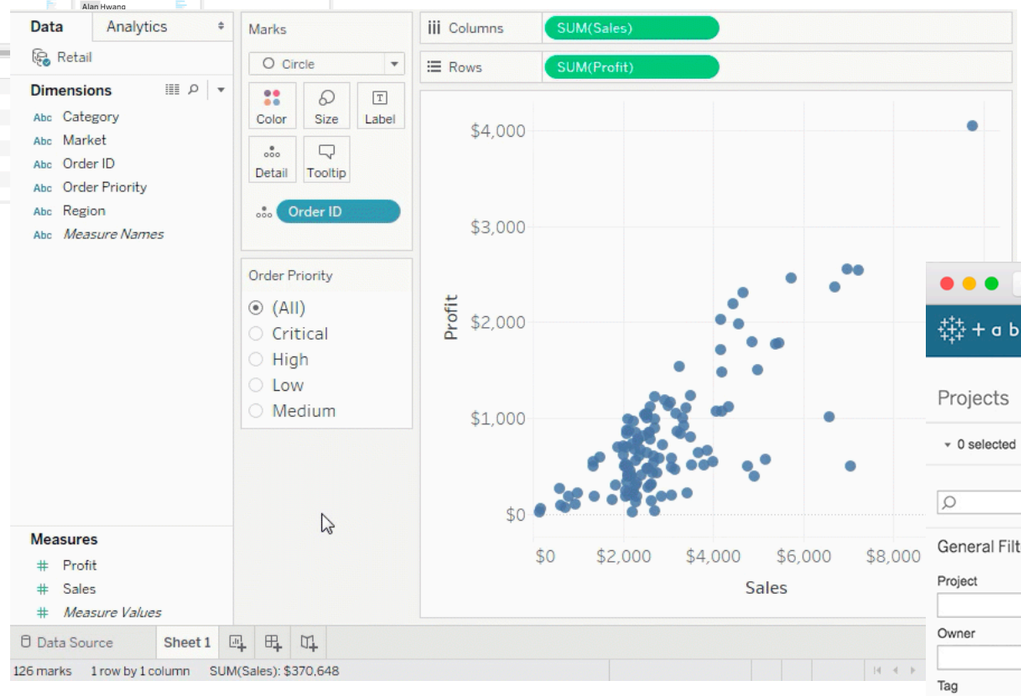


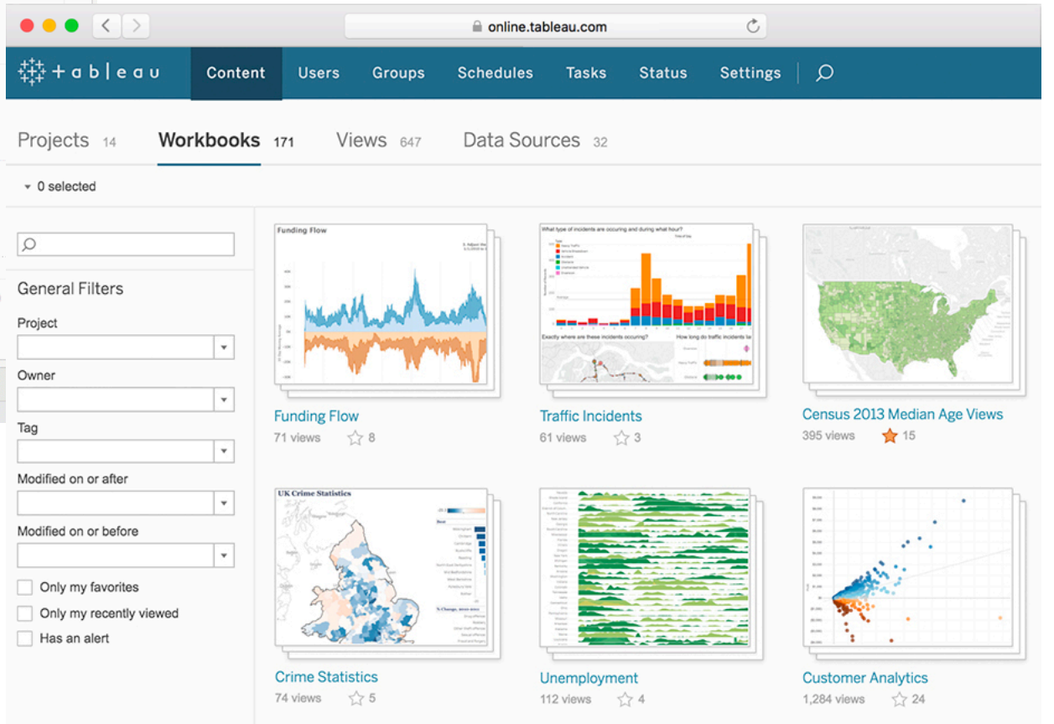
Tableau helps people see and understand their data.





Sales	Quantity	Profit	Discount	Region	State	Row ID	Order ID	Segment	Customer ID
18.648	7	-12.432	0.7	South	North Carolina	231	US-2015-156216	Corporate	EA-14035
178.384	2	22.298	0.2	South	Florida	315	CA-2015-167850	Corporate	AG-10525
15.552	3	5.4432	0.2	South	Florida	316	CA-2015-167850	Corporate	AG-10525
39.072	6	9.768	0.2	South	North Carolina	404	CA-2015-155208	Corporate	SP-20650
10.368	2	3.6288	0.2	South	North Carolina	705	CA-2015-138527	Corporate	BN-11470
166.84	5	18.7695	0.2	South	North Carolina	706	CA-2015-138527	Corporate	BN-11470
15.216	1	2.2824	0.2	South	North Carolina	707	CA-2015-138527	Corporate	BN-11470
11.36	2	5.3392	0	South	Louisiana	764	CA-2015-162775	Corporate	CS-12250

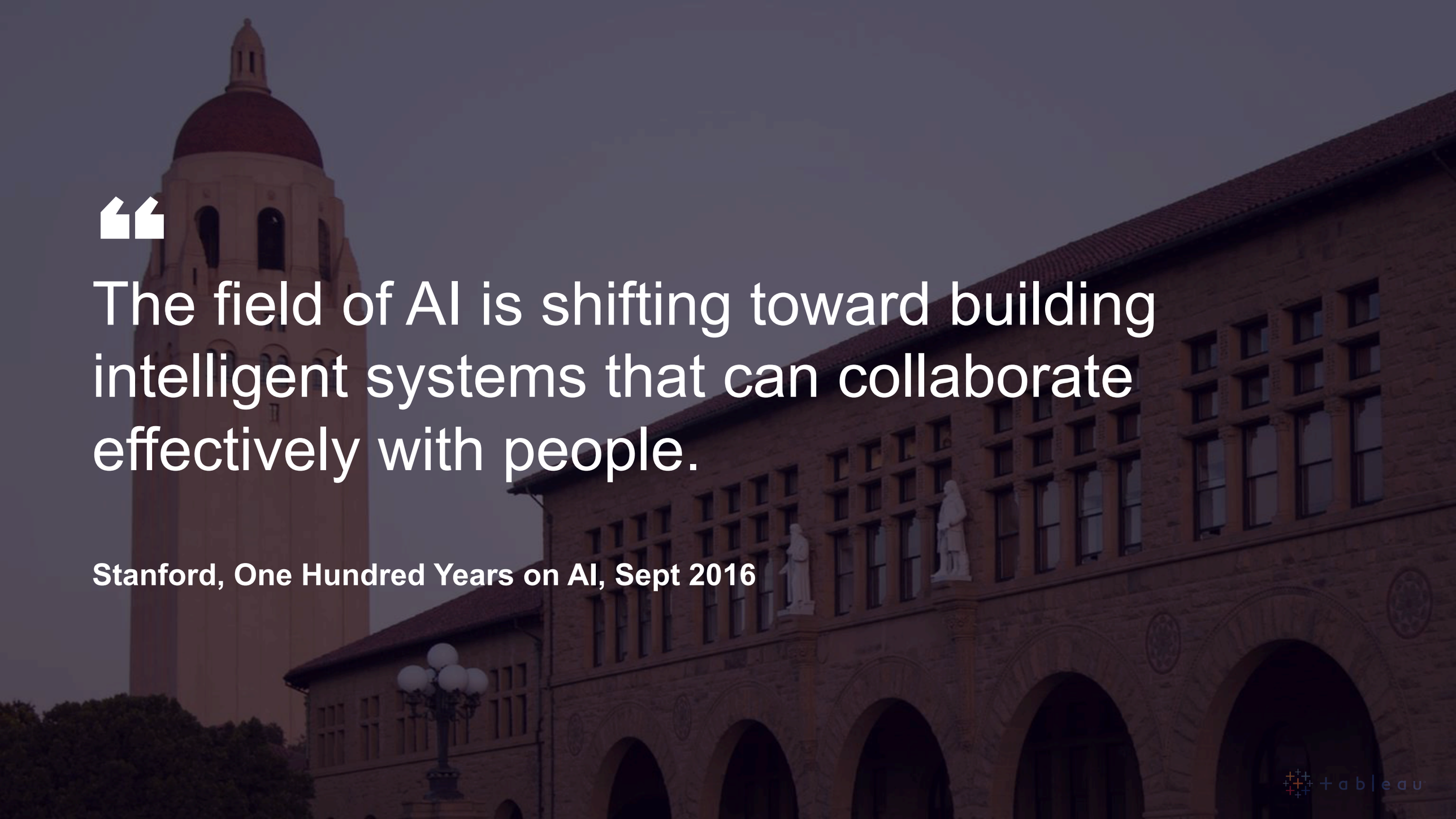




AI will replace the analyst

misconception

AI will ~~replace~~ assist the analyst


“
The field of AI is shifting toward building intelligent systems that can collaborate effectively with people.
Stanford, One Hundred Years on AI, Sept 2016
tableau

Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases

Chris Stolte and Pat Hanrahan
Computer Science Department
Stanford University

Abstract

In the last several years, large multi-dimensional databases have become common in a variety of applications such as data warehousing and scientific computing. Analysis and exploration tasks place significant demands on the interfaces to these databases. Because of the size of the data sets, dense graphical representations are more effective for exploration than spreadsheets and charts. Furthermore, because of the exploratory nature of the analysis, it must be possible for the analysts to change visualizations rapidly as they pursue a cycle involving first hypothesis and then experimentation.

In this paper we present Polaris, an interface for exploring large multi-dimensional databases that extends the well-known Pivot Table interface. The novel features of Polaris include an interface for constructing visual specifications of table-based graphical displays and the ability to generate a precise set of relational queries from the visual specifications. The visual specifications can be rapidly and incrementally developed, giving the analyst visual feedback as they construct complex queries and visualizations.

1. Introduction

In the last several years, large databases have become common in a variety of applications. Corporations are creating large data warehouses of historical data on key aspects of their operations. International research projects, such as the Human Genome Project [15] and Digital Sky Survey [23], are generating massive databases of scientific data.

A major challenge with these databases is to extract meaning from the data they contain: to discover structure, find patterns, and derive causal relationships. The analysis and exploration necessary to uncover this hidden information places significant demands on the human-computer interfaces to these databases. The exploratory analysis process is one of hypothesis, experiment and discovery. The path of exploration is unpredictable, and the analysis requires visually exploring projections of data cubes using linked views of bar charts, scatterplots, and parallel coordinate displays [10].

In this paper we present Polaris, an interface for the exploration of multi-dimensional databases that extends the Pivot Table interface to directly generate a rich, expressive set of graphical displays. Polaris builds tables using an algebraic formalism involving the fields of the database. Each table consists of layers and panes, and each pane may be a different graphic. The use of tables to organize multiple graphs on a display is a technique often used by statisticians in their analysis of data [3][7][28].

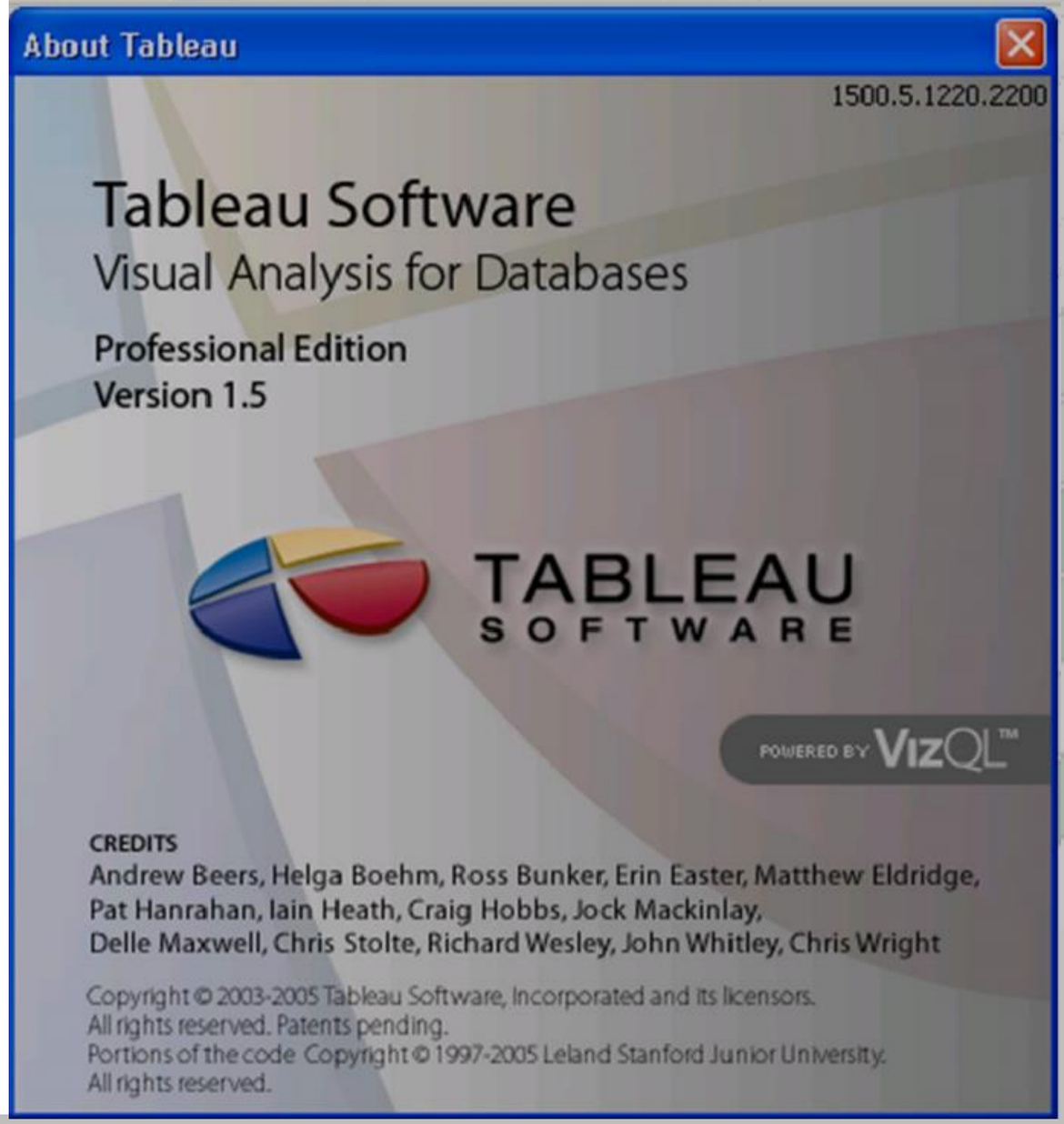
The Polaris interface is simple and expressive because it is built upon a formalism for constructing graphs and building data transformations. We interpret the state of the interface as a visual specification of the analysis task and automatically compile it into data and graphical transformations. This allows us to combine statistical analysis and visualization. Furthermore, all intermediate specifications that can be created in the visual language are valid and can be interpreted to create visualizations. Therefore, analysts can incrementally construct complex queries, receiving visual feedback as they assemble and alter the specifications.

2. Related Work

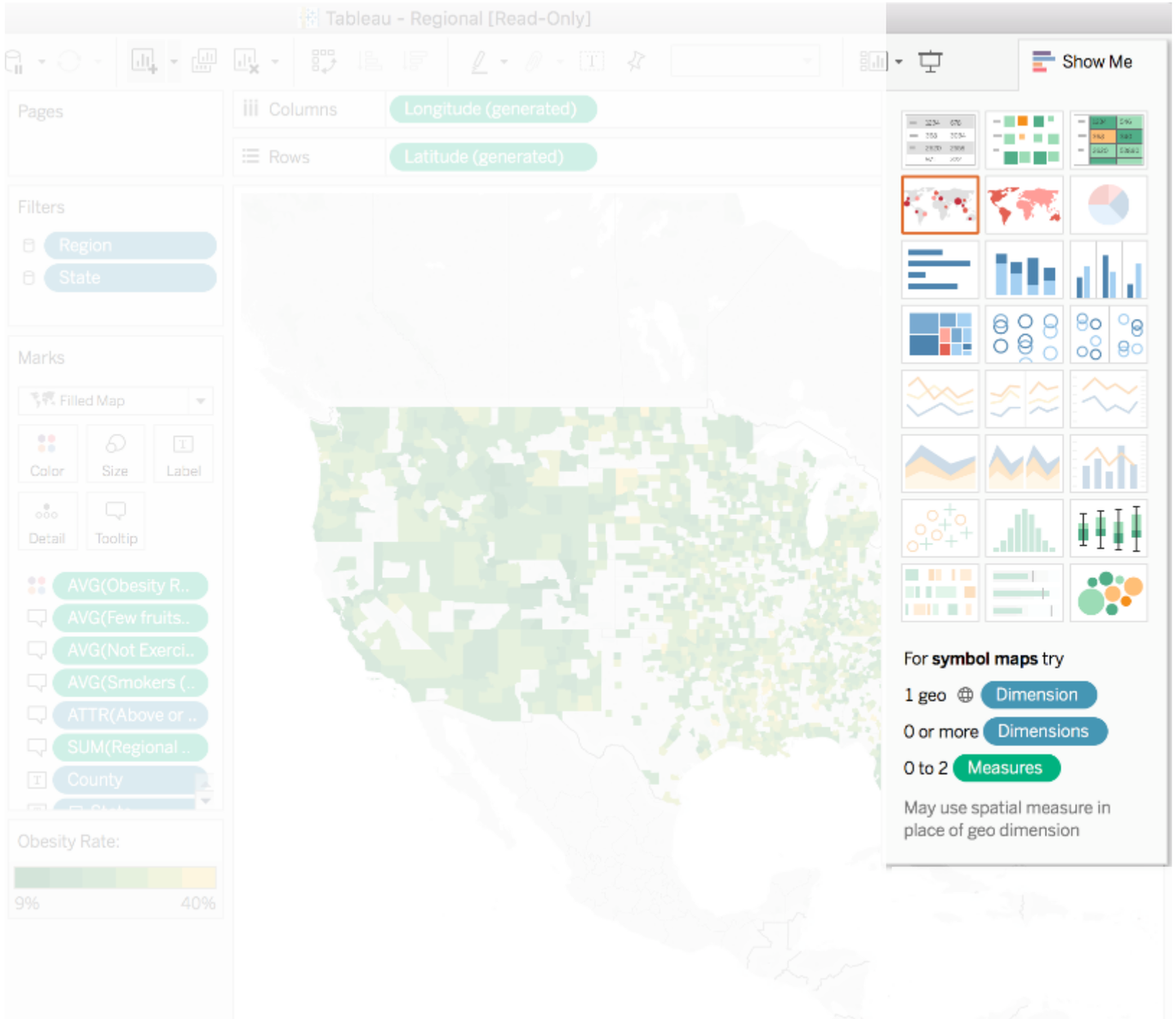
The related work to Polaris can be divided into three categories: formal graphical specifications, table-based data displays and database exploration tools.

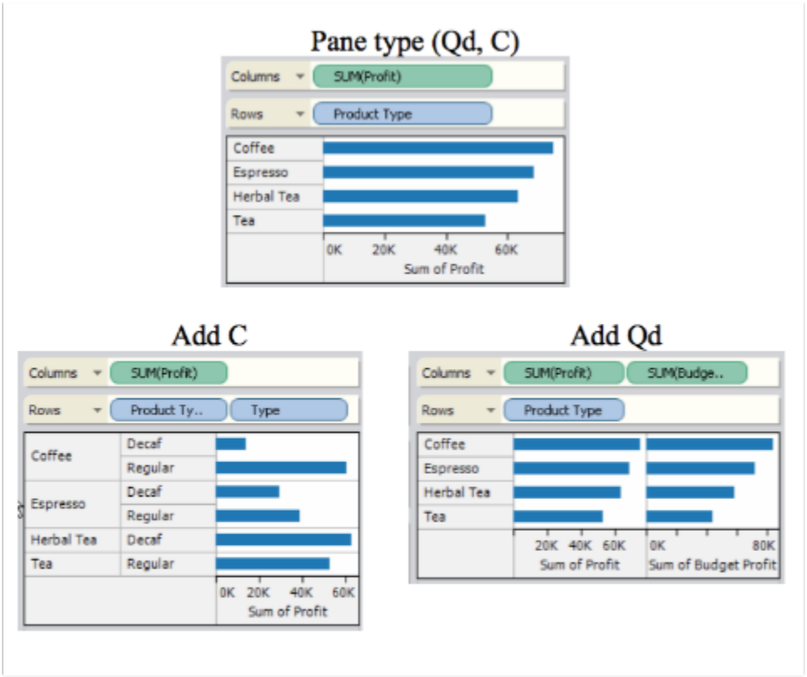
2.1. Formal Graphical Specifications

Bertin's *Semiology of Graphics* [4] is one of the earliest attempts at formalizing graphing techniques. Bertin developed a vocabulary for describing data and the techniques for encoding data in a graphic. One of his important contributions is the identification of the retinal variables (position, color, size, etc.) in which data can be encoded. Cleveland [7][8] used theoretical and experimental results to determine how well people can use these different retinal properties to compare quantitative variations.



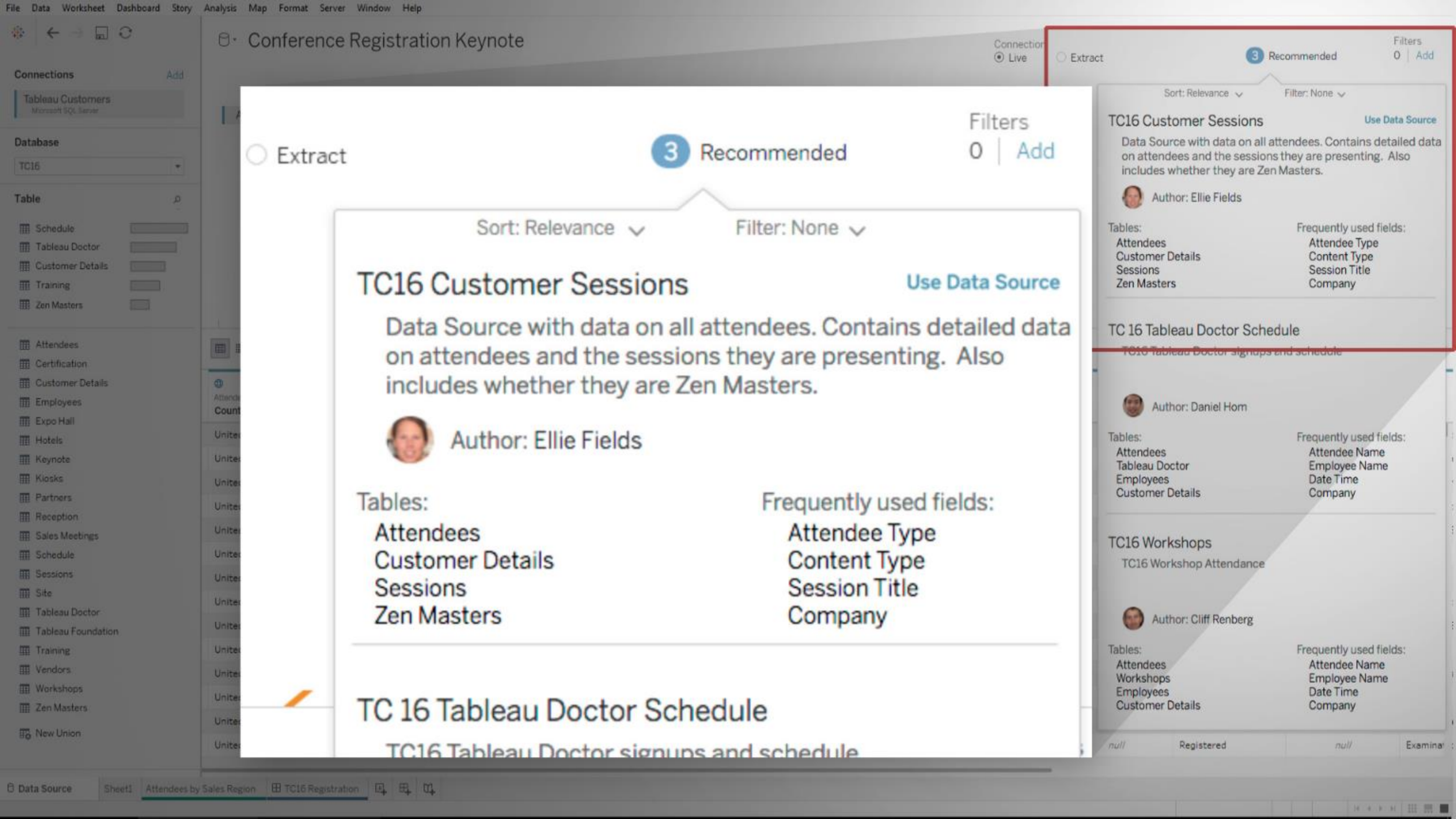
Automatic Visualization





Jock Mackinlay, Pat Hanrahan, and Chris Stolte. 2007. Show Me: Automatic Presentation for Visual Analysis. IEEE Transactions on Visualization and Computer Graphics 13, 6 (November 2007), 1137-1144.

tableau


Extract
3 Recommended
Filters
0 | Add
Sort: Relevance
Filter: None
TC16 Customer Sessions
Use Data Source
Data Source with data on all attendees. Contains detailed data on attendees and the sessions they are presenting. Also includes whether they are Zen Masters.
Author: Ellie Fields
Tables:
Attendees
Customer Details
Sessions
Zen Masters
Frequently used fields:
Attendee Type
Content Type
Session Title
Company
TC 16 Tableau Doctor Schedule
TC16 Tableau Doctor signups and schedule
Author: Daniel Hom
Tables:
Attendees
Tableau Doctor
Employees
Customer Details
Frequently used fields:
Attendee Name
Employee Name
Date Time
Company
TC16 Workshops
TC16 Workshop Attendance
Author: Cliff Renberg
Tables:
Attendees
Workshops
Employees
Customer Details
Frequently used fields:
Attendee Name
Employee Name
Date Time
Company

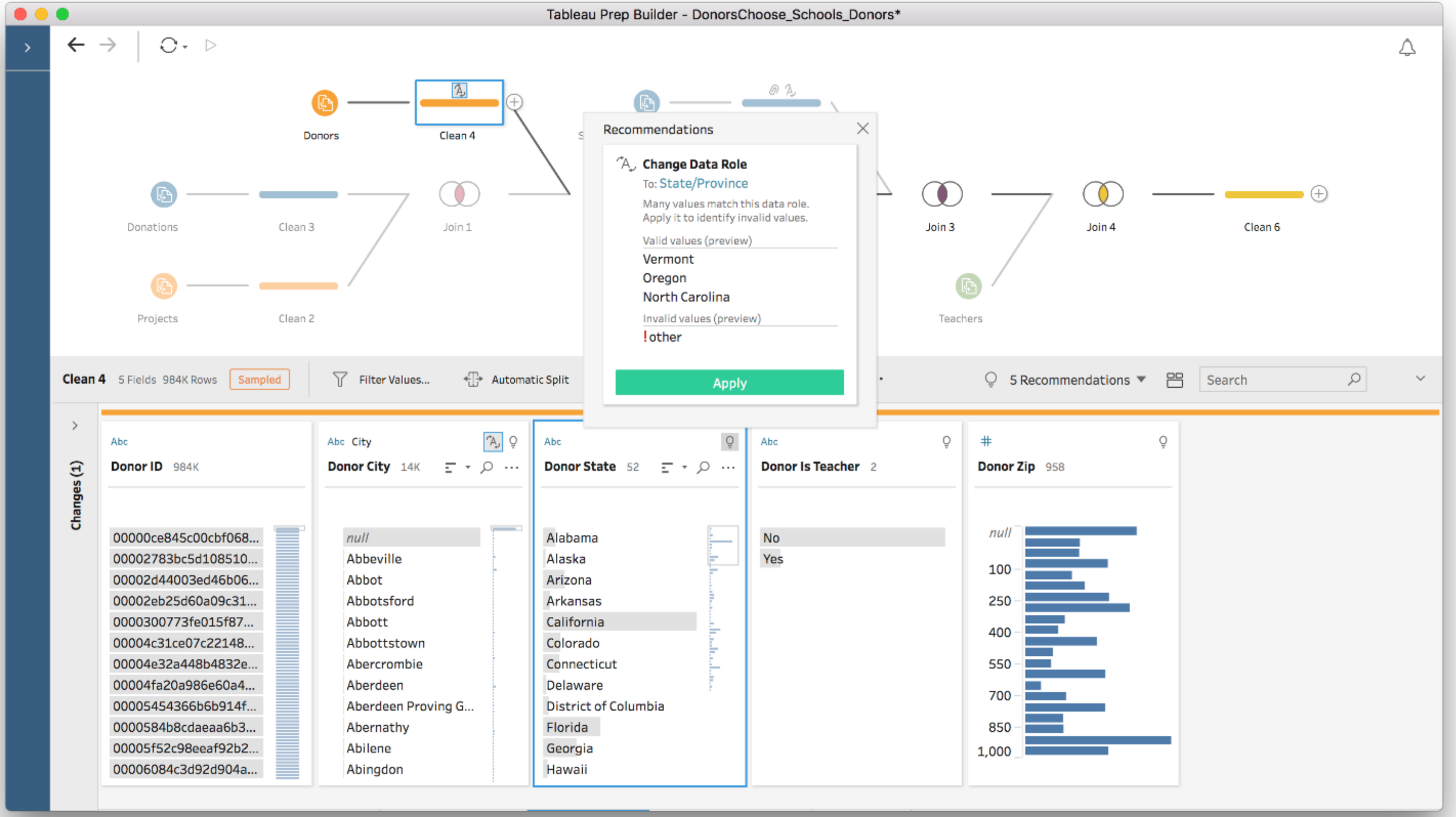
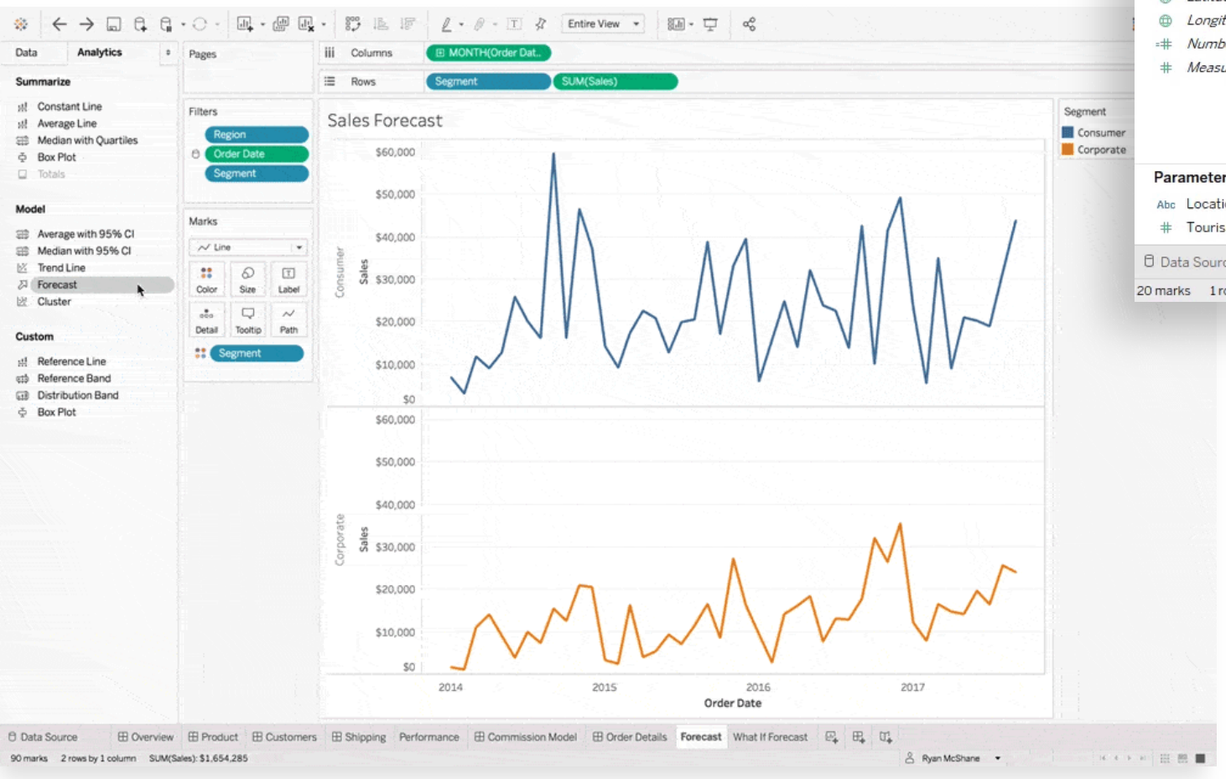
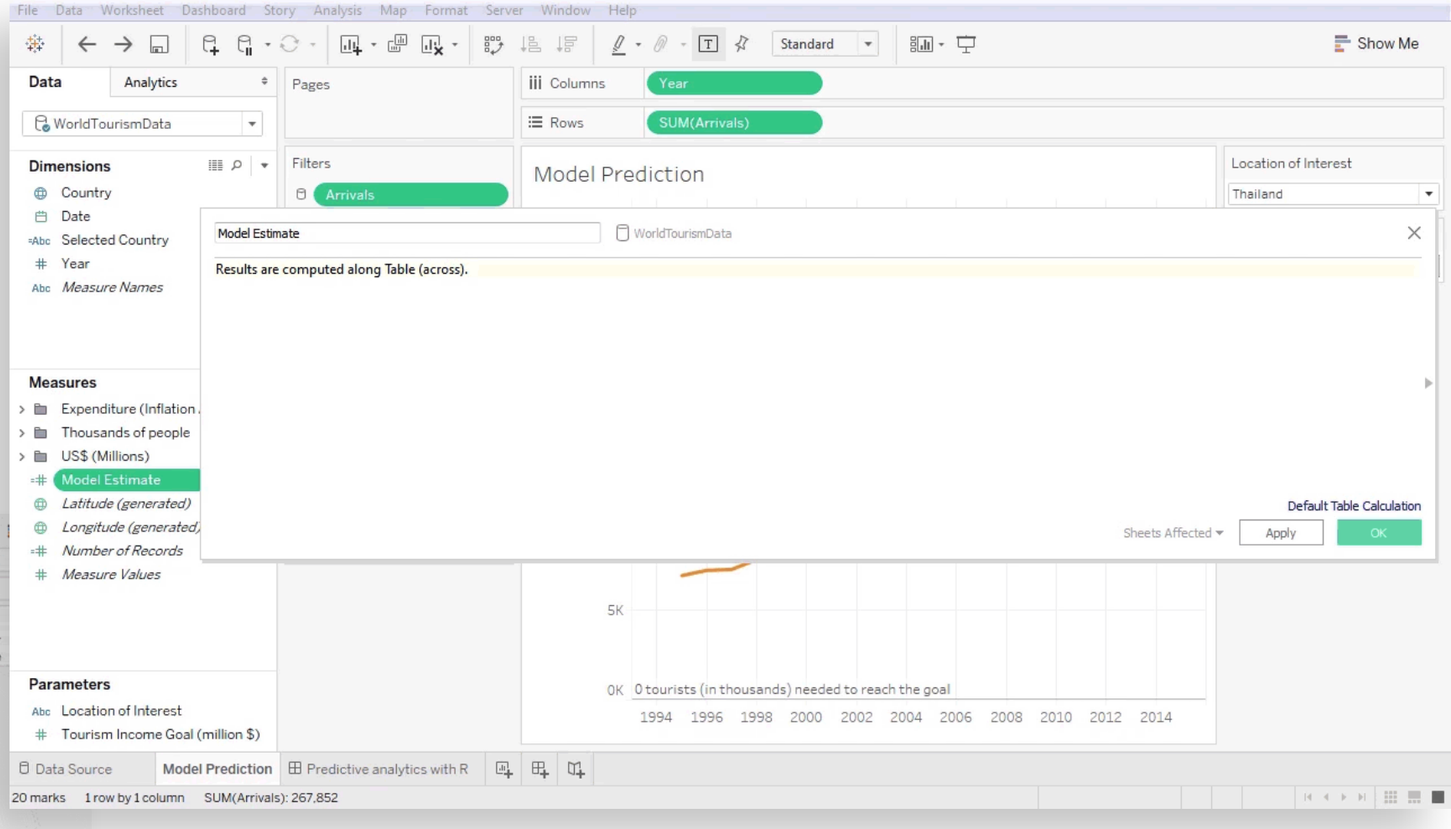


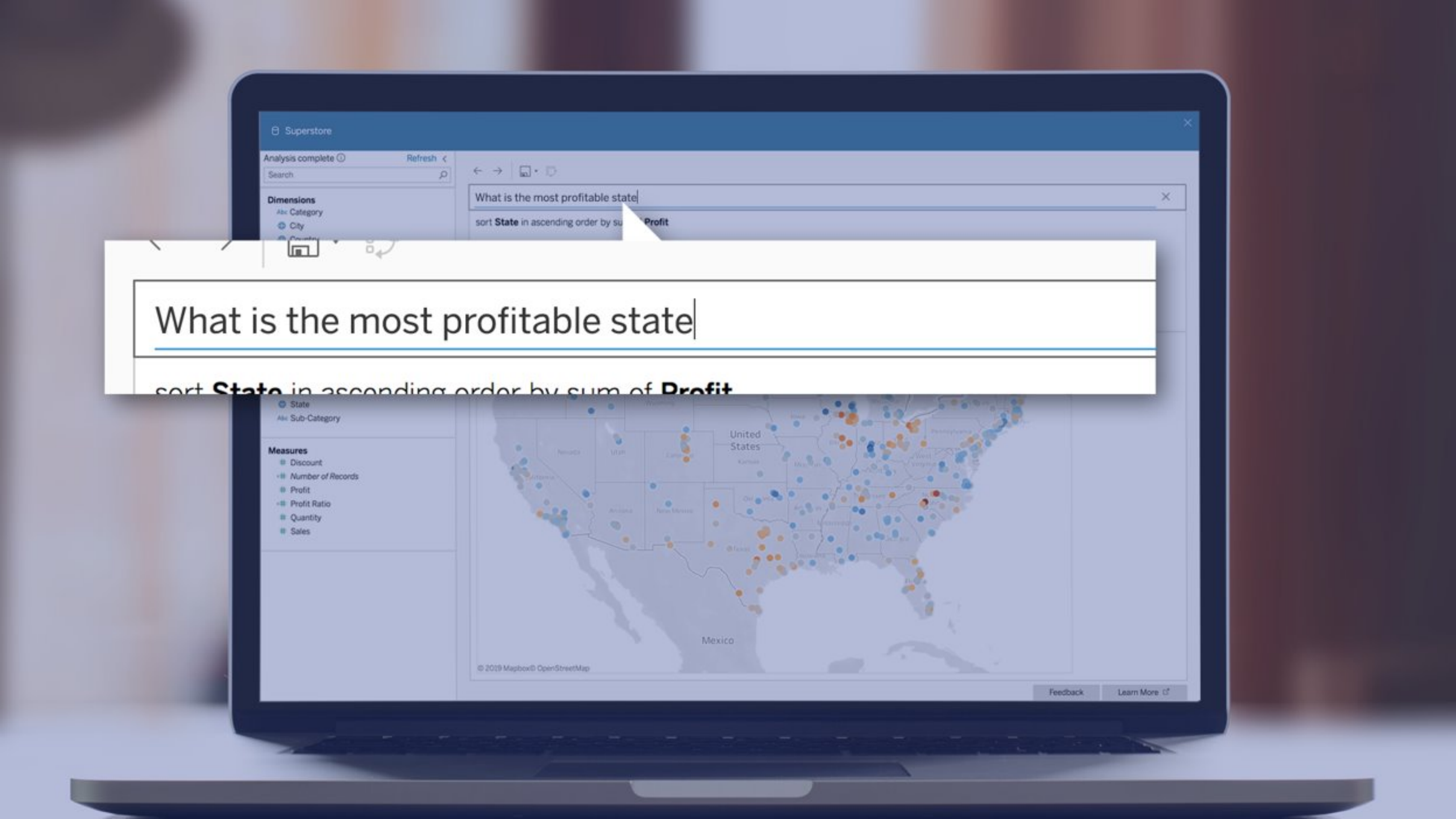
Tableau Prep Builder - DonorsChoose_Schools_Donors*
Donors
Clean 4
Donations
Clean 3
Join 1
Projects
Clean 2
Join 3
Join 4
Clean 6
Teachers
Recommendations
Change Data Role
To: State/Province
Many values match this data role.
Apply it to identify invalid values.
Valid values (preview)
Vermont
Oregon
North Carolina
Invalid values (preview)
!other
Apply
Clean 4 5 Fields 984K Rows Sampled
Filter Values...
Automatic Split
5 Recommendations
Search
Changes (1)
Donor ID 984K
Donor City 14K
Donor State 52
Donor Is Teacher 2
Donor Zip 958
00000ce845c00cbf068...
00002783bc5d108510...
00002d44003ed46b06...
00002eb25d60a09c31...
0000300773fe015f87...
00004c31ce07c22148...
00004e32a448b4832e...
00004fa20a986e60a4...
00005454366b6b914f...
0000584b8cdaeaa6b3...
00005f52c98eeaf92b2...
00006084c3d92d904a...
null
Abbeville
Abbot
Abbotsford
Abbott
Abbottstown
Abercrombie
Aberdeen
Aberdeen Proving G...
Abernathy
Abilene
Abingdon
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
District of Columbia
Florida
Georgia
Hawaii
No
Yes
null
100
250
400
550
700
850
1,000

Predictive Analytics





tableau


Superstore
Analysis complete
Refresh
Search
Dimensions
Category
City
What is the most profitable state
sort State in ascending order by su Profit
What is the most profitable state
sort State in ascending order by sum of Profit
State
Sub-Category
Measures
Discount
Number of Records
Profit
Profit Ratio
Quantity
Sales
United States
Mexico
© 2019 Mapbox © OpenStreetMap
Feedback
Learn More

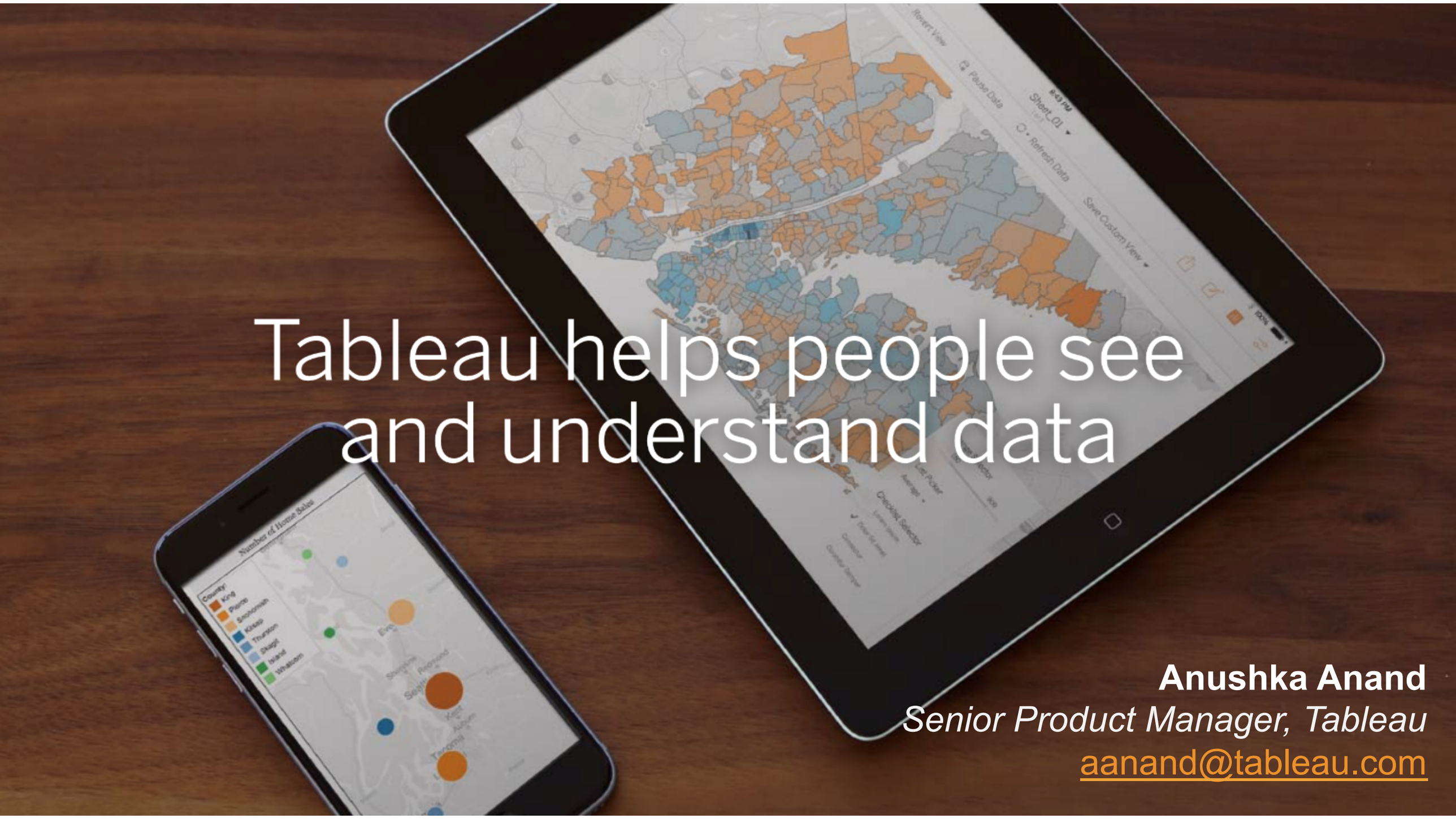


Tableau helps people see and understand data

Anushka Anand
Senior Product Manager, Tableau
aanand@tableau.com