



Anushka Anand

Senior Product Manager, Tableau

aanand@tableau.com

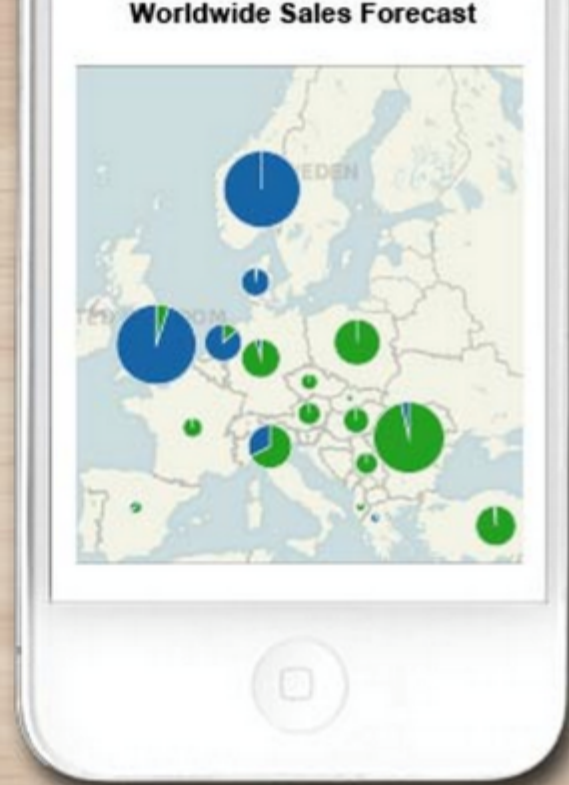
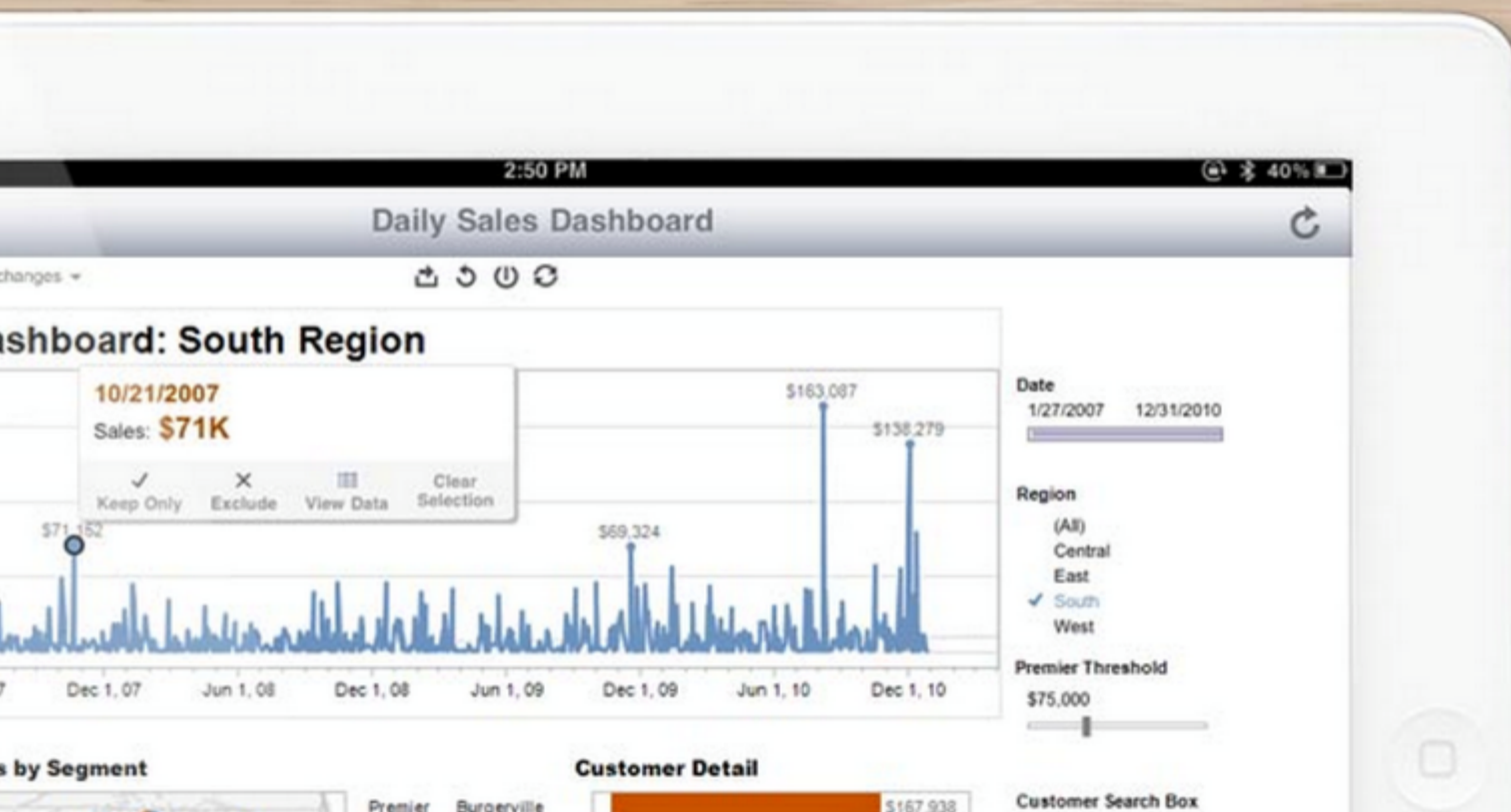
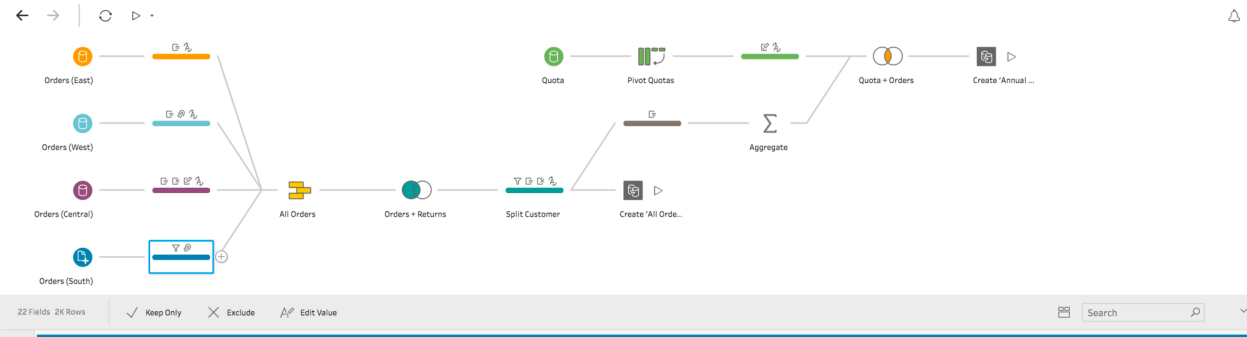


Tableau helps
people see
and understand
their data.





State	Row ID	Order ID	Segment	Customer ID	Customer Name	Ship Mode	Order Date
Alabama	0	CA-2015-100293	Consumer	AA-10375	Aaron Hawkins	First Class	01/01/2015, 1...
Arkansas	2,000	CA-2015-100706	Contractor	AA-10480	Aaron Smayling	Same Day	01/01/2019, 1...
Florida	4,000	CA-2015-100895	Corporate	AA-10645	Adam Bellavance	Second Class	
Georgia	6,000	CA-2015-100916		AA-10060	Adam Hart	Standard Class	
Kentucky	8,000	CA-2015-101266		AA-10105	Adam Shillingsburg		
Louisiana	10,000	CA-2015-101560		AA-10165	Adrian Barton		
Mississippi		CA-2015-101770		AA-10255	Adrian Hane		
North Carolina		CA-2015-102274		AA-10600	Alan Barnes		
South Carolina		CA-2015-102673		AA-10450	Alan Haines		
Tennessee		CA-2015-102988		AA-10870	Alan Hwano		
Virginia		CA-2015-103317		AA-10885			
		CA-2015-103366		AA-10330			

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

Dimensions

Category
Market
Order ID
Order Priority
Region
Measure Names

Measures

Profit
Sales
Measure Values



126 marks 1 row by 1 column SUM(Sales): \$370,648

online.tableau.com

Content Users Groups Schedules Tasks Status Settings

Projects 14 Workbooks 171 Views 647 Data Sources 32

0 selected

General Filters

Project
Owner
Tag
Modified on or after
Modified on or before
Only my favorites
Only my recently viewed
Has an alert

Funding Flow

71 views 8

Traffic Incidents

61 views 3

Census 2013 Median Age Views

395 views 15

Crime Statistics

74 views 5

Unemployment

112 views 4

Customer Analytics

1,284 views 24



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

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 ana-

lyst is 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.

About Tableau

1500.5.1220.2200

Tableau Software Visual Analysis for Databases

Professional Edition
Version 1.5



TABLEAU
SOFTWARE

POWERED BY VizQL™

CREDITS

Andrew Beers, Helga Boehm, Ross Bunker, Erin Easter, Matthew Eldridge, Pat Hanrahan, Iain Heath, Craig Hobbs, Jock Mackinlay, Delle Maxwell, Chris Stolte, Richard Wesley, John Whitley, Chris Wright

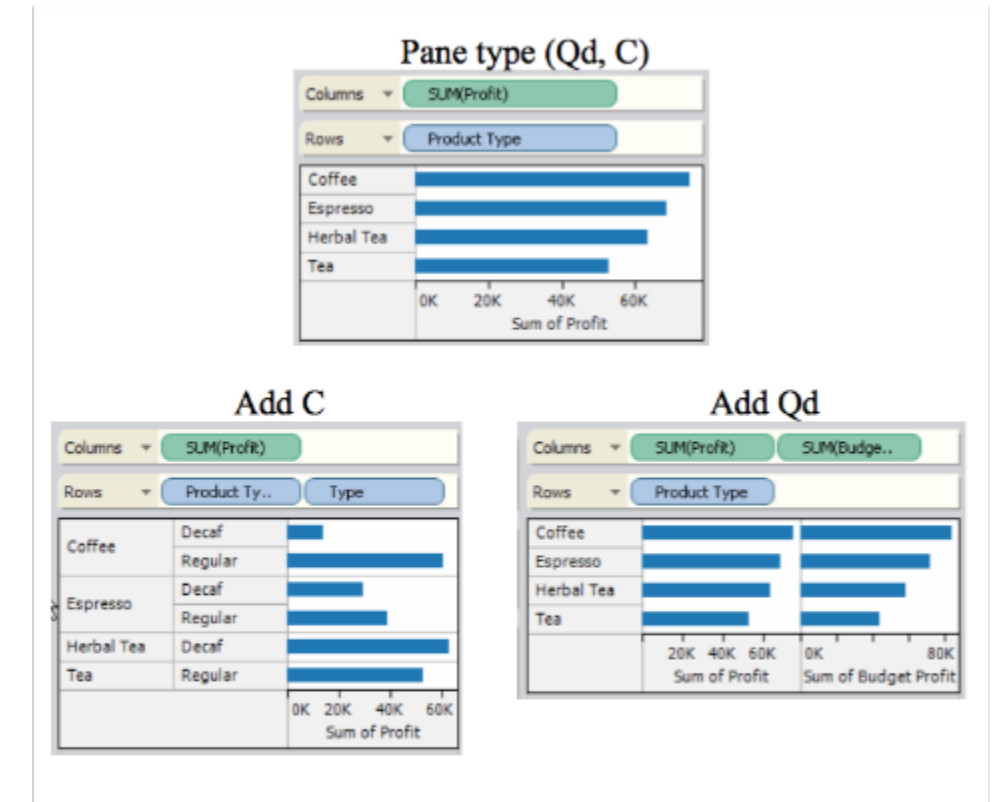
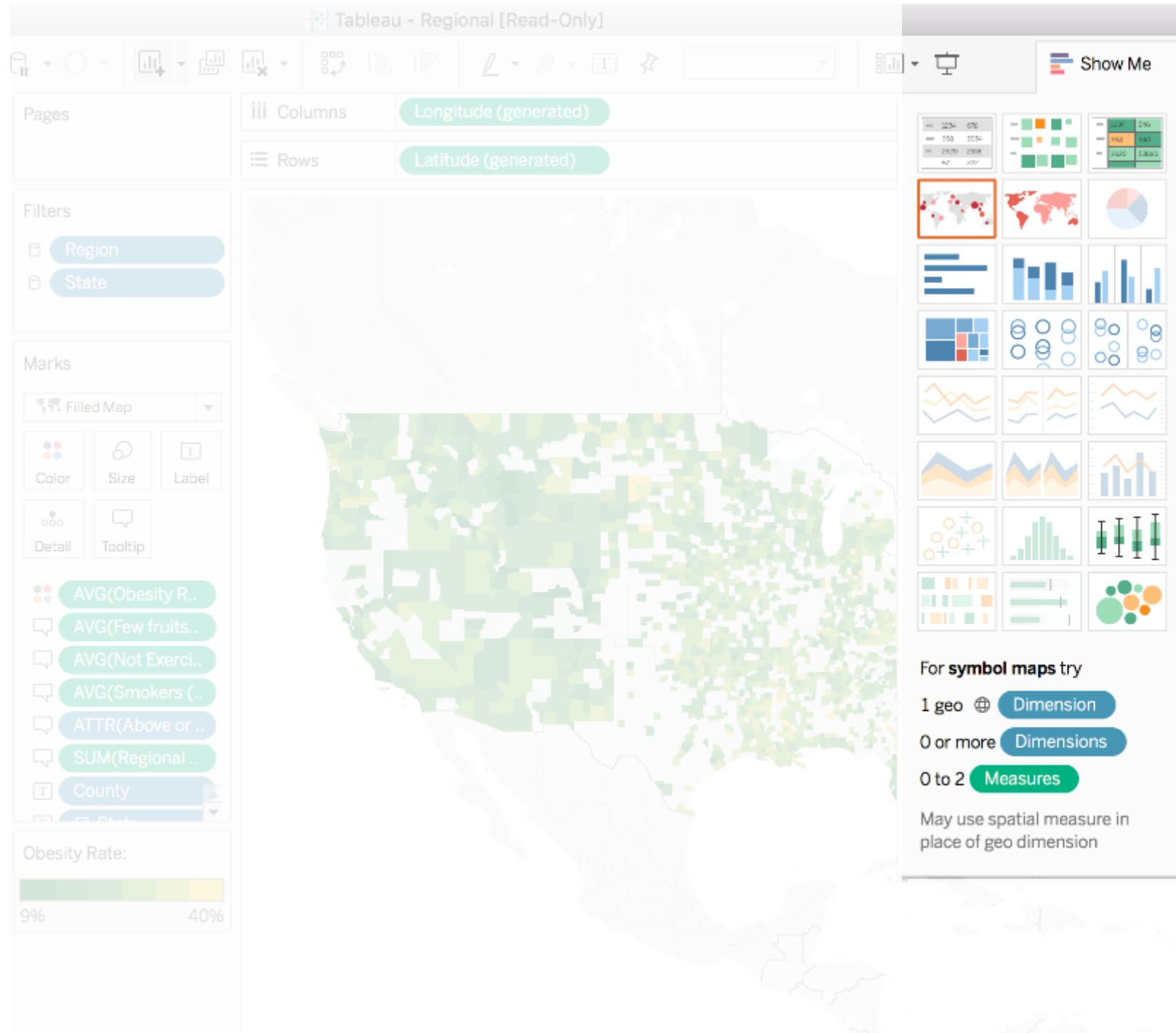
Copyright © 2003-2005 Tableau Software, Incorporated and its licensors.

All rights reserved. Patents pending.

Portions of the code Copyright © 1997-2005 Leland Stanford Junior University.

All rights reserved.

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.

FileDataWorksheetDashboardStoryAnalysisMapFormatServerWindowHelp

ConnectionsAdd

Tableau CustomersMicrosoft SQL Server

DatabaseTC16

Table

Schedule

Tableau Doctor

Customer Details

Training

Zen Masters

Attendees

Certification

Customer Details

Employees

Expo Hall

Hotels

Keynote

Kiosks

Partners

Reception

Sales Meetings

Schedule

Sessions

Site

Tableau Doctor

Tableau Foundation

Training

Vendors

Workshops

Zen Masters

New Union

Conference Registration Keynote

Connection

Live

Extract

3Recommended

Filters0Add

Sort: RelevanceFilter: None

TC16 Customer Sessions

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:AttendeesCustomer DetailsSessionsZen Masters

Frequently used fields:Attendee TypeContent TypeSession TitleCompany

TC 16 Tableau Doctor Schedule

TC16 Tableau Doctor signups and schedule

Author: Daniel Horn

Tables:AttendeesTableau DoctorEmployeesCustomer Details

Frequently used fields:Attendee NameEmployee NameDate TimeCompany

TC16 Workshops

TC16 Workshop Attendance

Author: Cliff Renberg

Tables:AttendeesWorkshopsEmployeesCustomer Details

Frequently used fields:Attendee NameEmployee NameDate TimeCompany

TC16 Customer Sessions

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:AttendeesCustomer DetailsSessionsZen Masters

Frequently used fields:Attendee TypeContent TypeSession TitleCompany

TC 16 Tableau Doctor Schedule

TC16 Tableau Doctor signups and schedule

Author: Daniel Horn

Tables:AttendeesTableau DoctorEmployeesCustomer Details

Frequently used fields:Attendee NameEmployee NameDate TimeCompany

TC16 Workshops

TC16 Workshop Attendance

Author: Cliff Renberg

Tables:AttendeesWorkshopsEmployeesCustomer Details

Frequently used fields:Attendee NameEmployee NameDate TimeCompany

Sheet1Attendees by Sales RegionTC16 Registration

The interface displays a data flow diagram with the following components:

- Donors** (orange icon) connects to **Clean 4** (orange bar).
- Donations** (blue icon) connects to **Clean 3** (blue bar).
- Projects** (orange icon) connects to **Clean 2** (orange bar).
- Clean 4** and **Clean 3** join at **Join 1** (Venn diagram).
- Clean 2** and **Join 1** join at **Join 3** (Venn diagram).
- Teachers** (green icon) joins **Join 3** at **Join 4** (Venn diagram).
- Join 4** connects to **Clean 6** (yellow bar).

A **Recommendations** dialog box is open, showing the following content:

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)

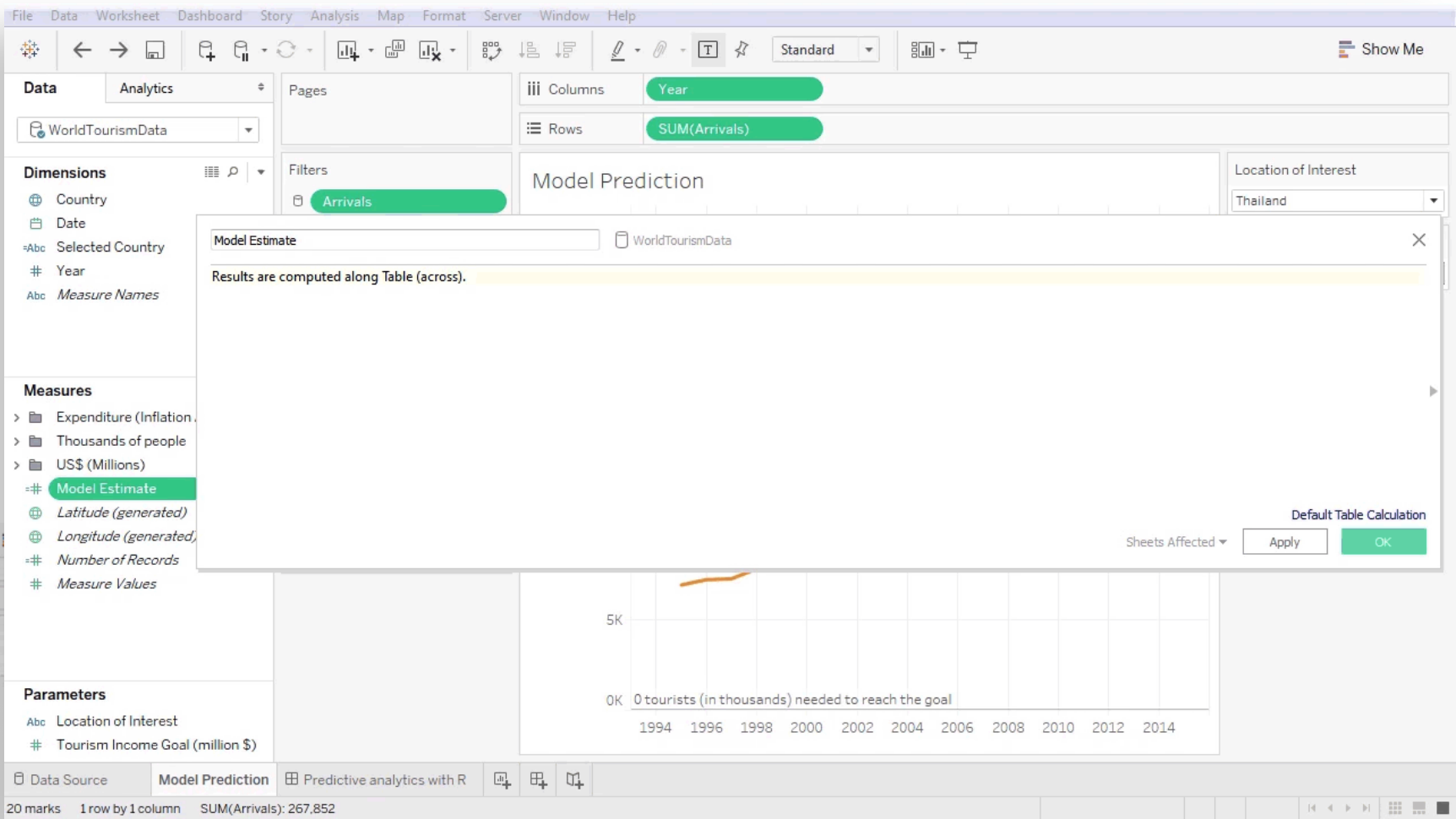
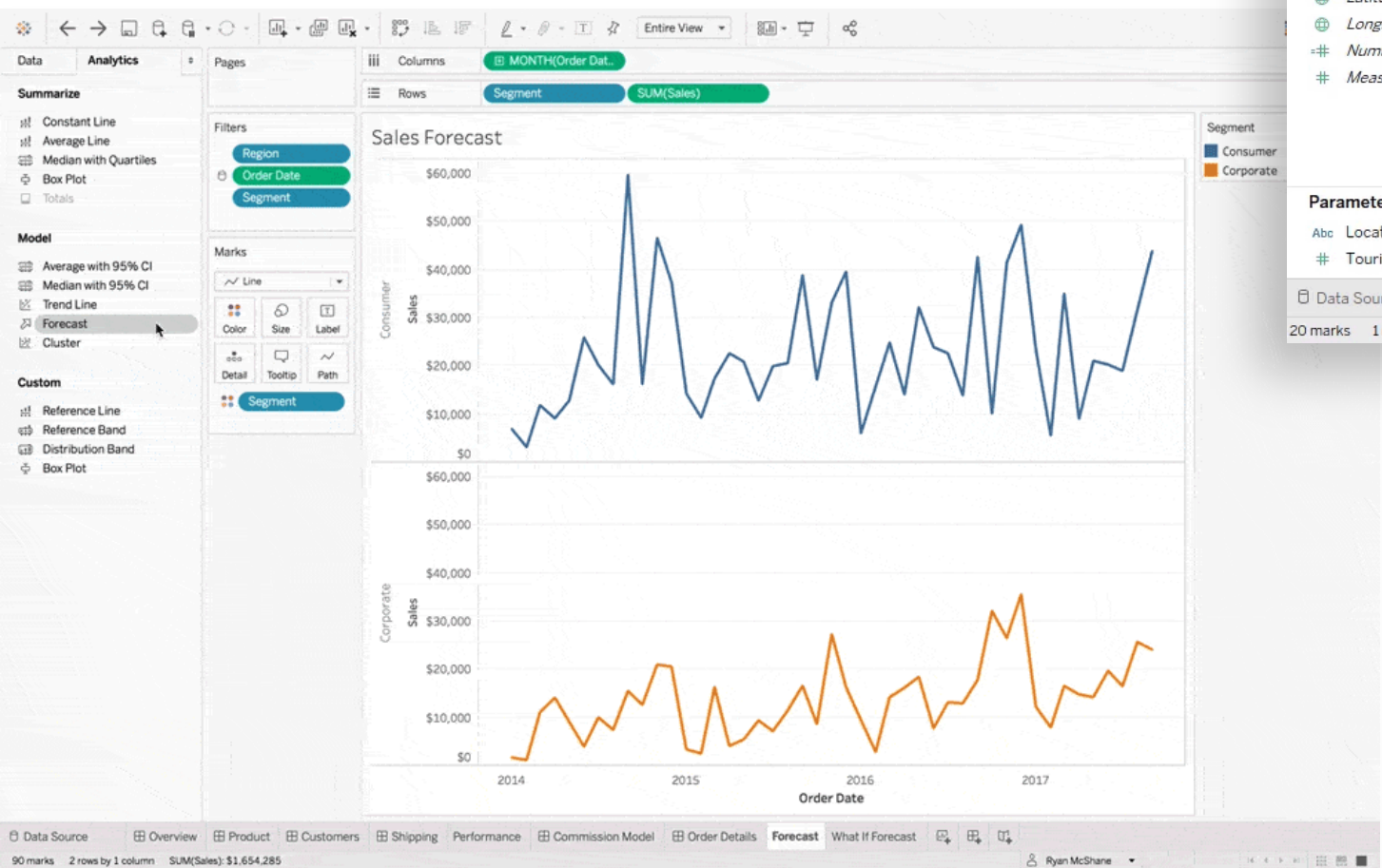
- !other

Apply

The bottom section shows the **Clean 4** data table with 5 fields and 984K rows, marked as **Sampled**. The fields and their data are:

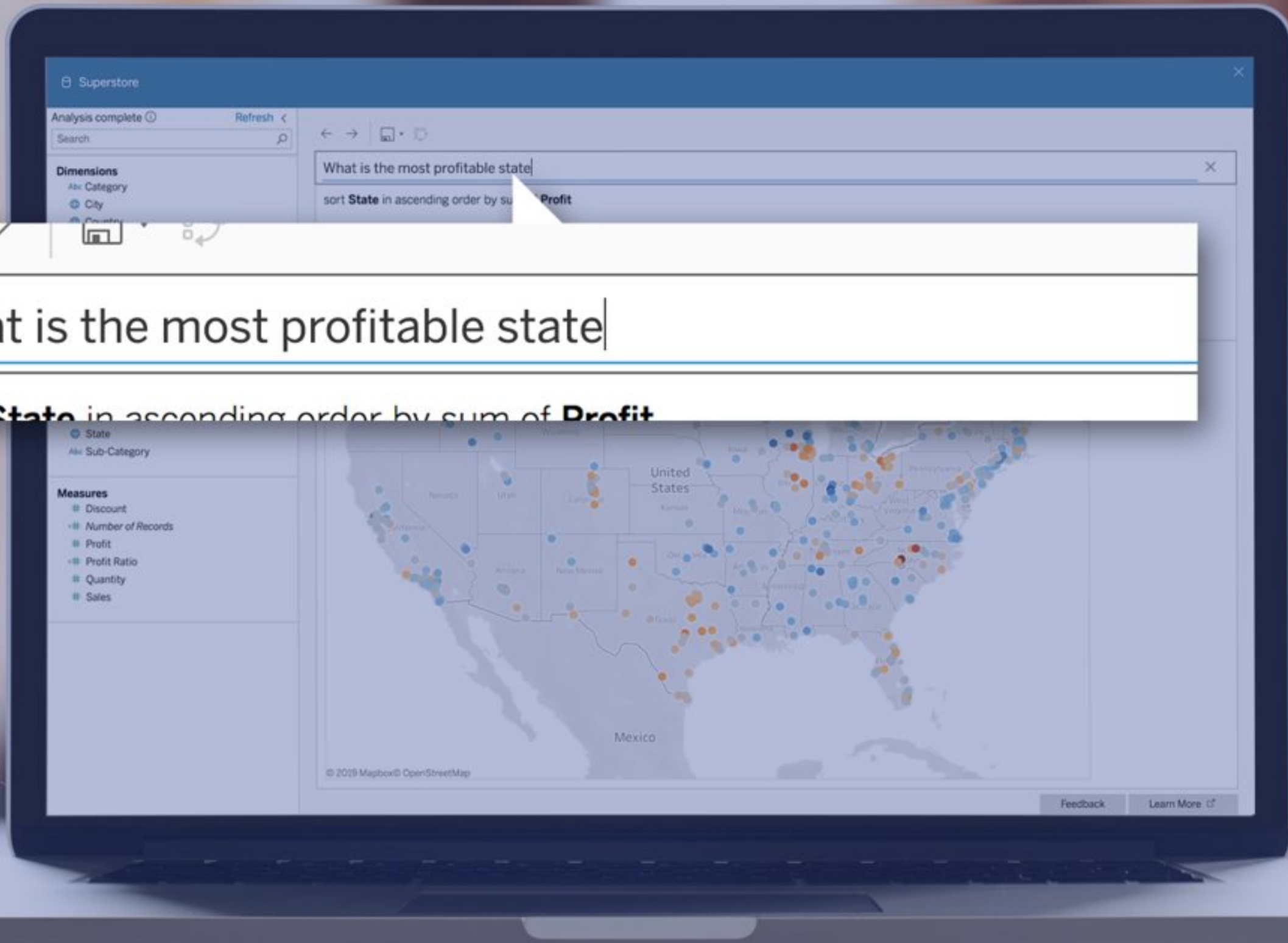
- Donor ID** (984K rows): 00000ce845c00cbf068..., 00002783bc5d108510..., 00002d44003ed46b06..., 00002eb25d60a09c31..., 0000300773fe015f87..., 00004c31ce07c22148..., 00004e32a448b4832e..., 00004fa20a986e60a4..., 00005454366b6b914f..., 0000584b8cdaaaa6b3..., 00005f52c98eeaf92b2..., 00006084c3d92d904a...
- Donor City** (14K rows): null, Abbeville, Abbot, Abbotsford, Abbott, Abbottstown, Abercrombie, Aberdeen, Aberdeen Proving G..., Abernathy, Abilene, Abingdon
- Donor State** (52 rows): Alabama, Alaska, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, District of Columbia, Florida, Georgia, Hawaii
- Donor Is Teacher** (2 rows): No, Yes
- Donor Zip** (958 rows): null, 100, 250, 400, 550, 700, 850, 1,000

Predictive Analytics



What is the most profitable state

sort **State** in ascending order by sum of **Profit**



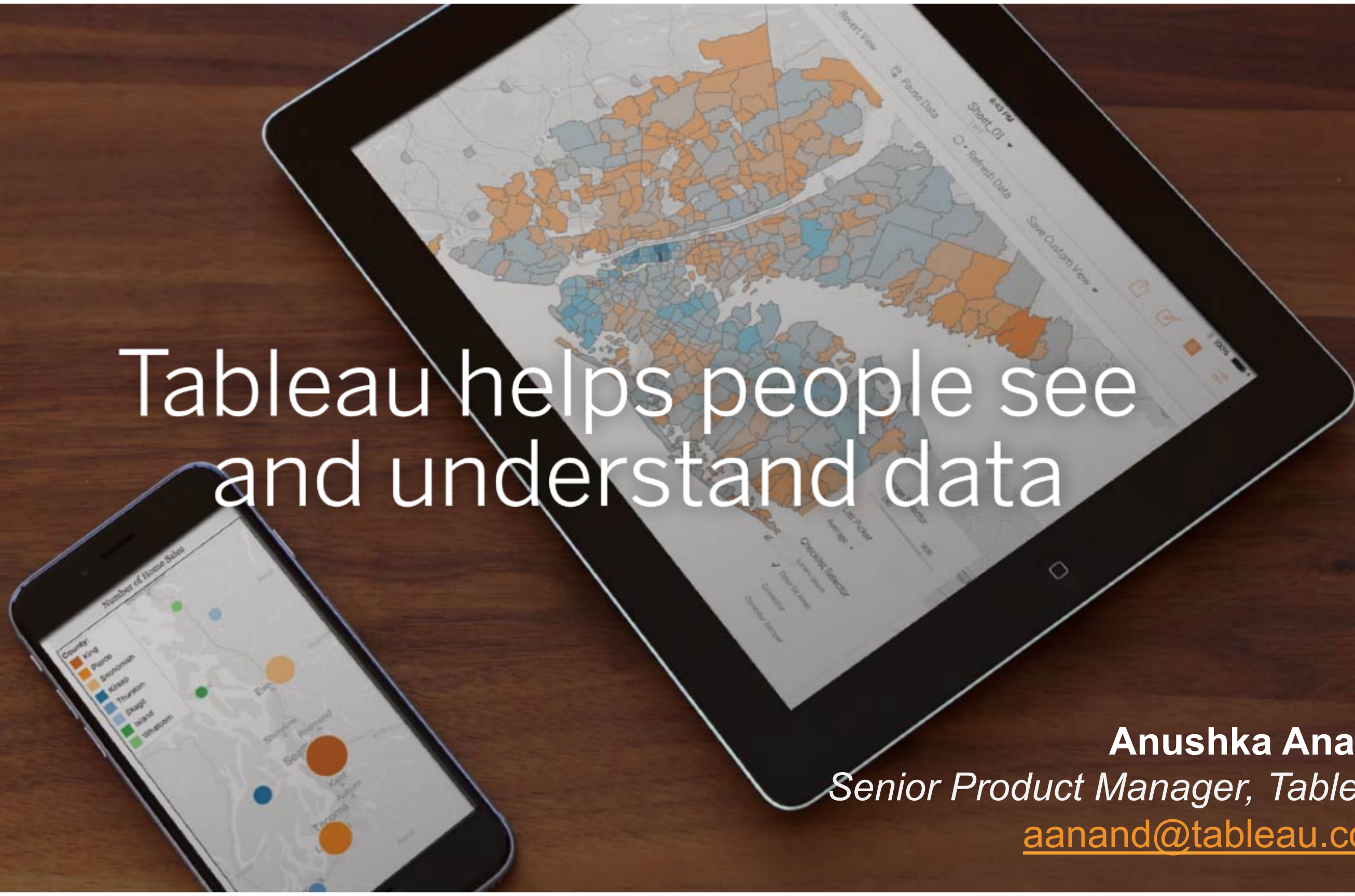
A tablet and a smartphone are shown on a wooden surface. The tablet displays a Tableau map of Washington state with counties colored in shades of orange and blue. The smartphone displays a Tableau bubble chart titled 'Number of Home Sales' with a legend for counties: King (orange), Pierce (light orange), Snohomish (yellow), Kitsap (blue), Thurston (light blue), Skagit (light green), Island (green), and Whatcom (dark green).

Tableau helps people see
and understand data

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