Data Analysis Fundamentals with Tableau

Ben Sullins
Bensullins.com
ben@bensullins.com

pluralsight
hardcore developer training

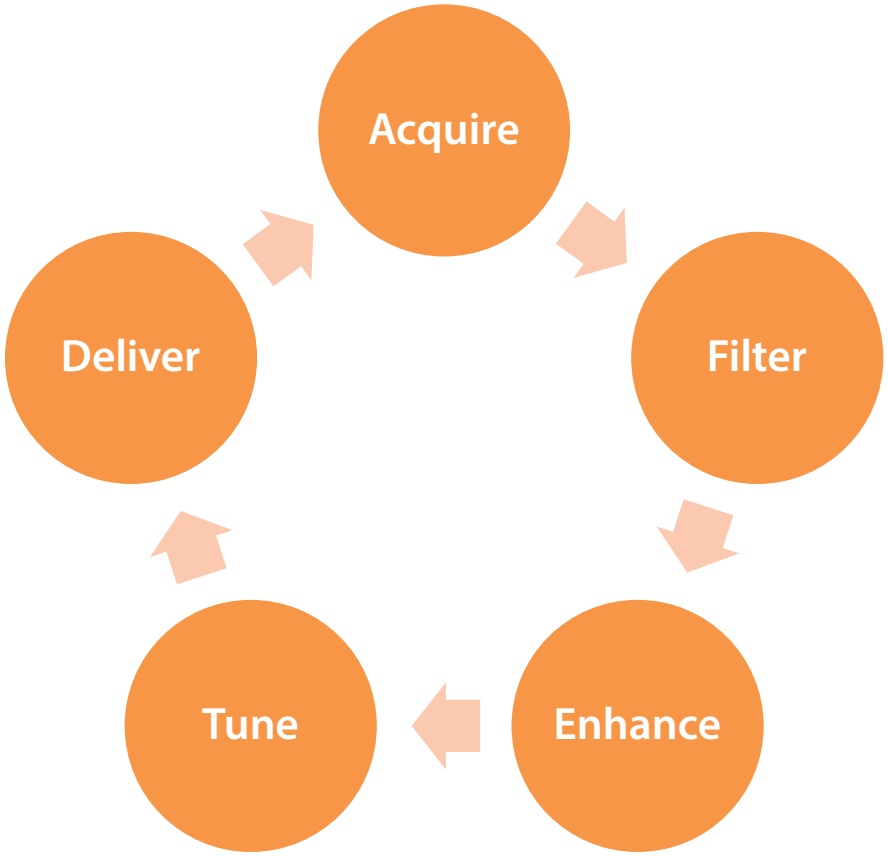
Course Outline

- Introduction to Tableau
- Getting dirty with your data
- Show me the data!
- Adding the intelligence factor
- Painting the whole picture
- Sharing your work

Overview

- Visual analysis
- Visual perception
- Tableau's product family
- Connecting to your data
- Data terminology
- View terminology

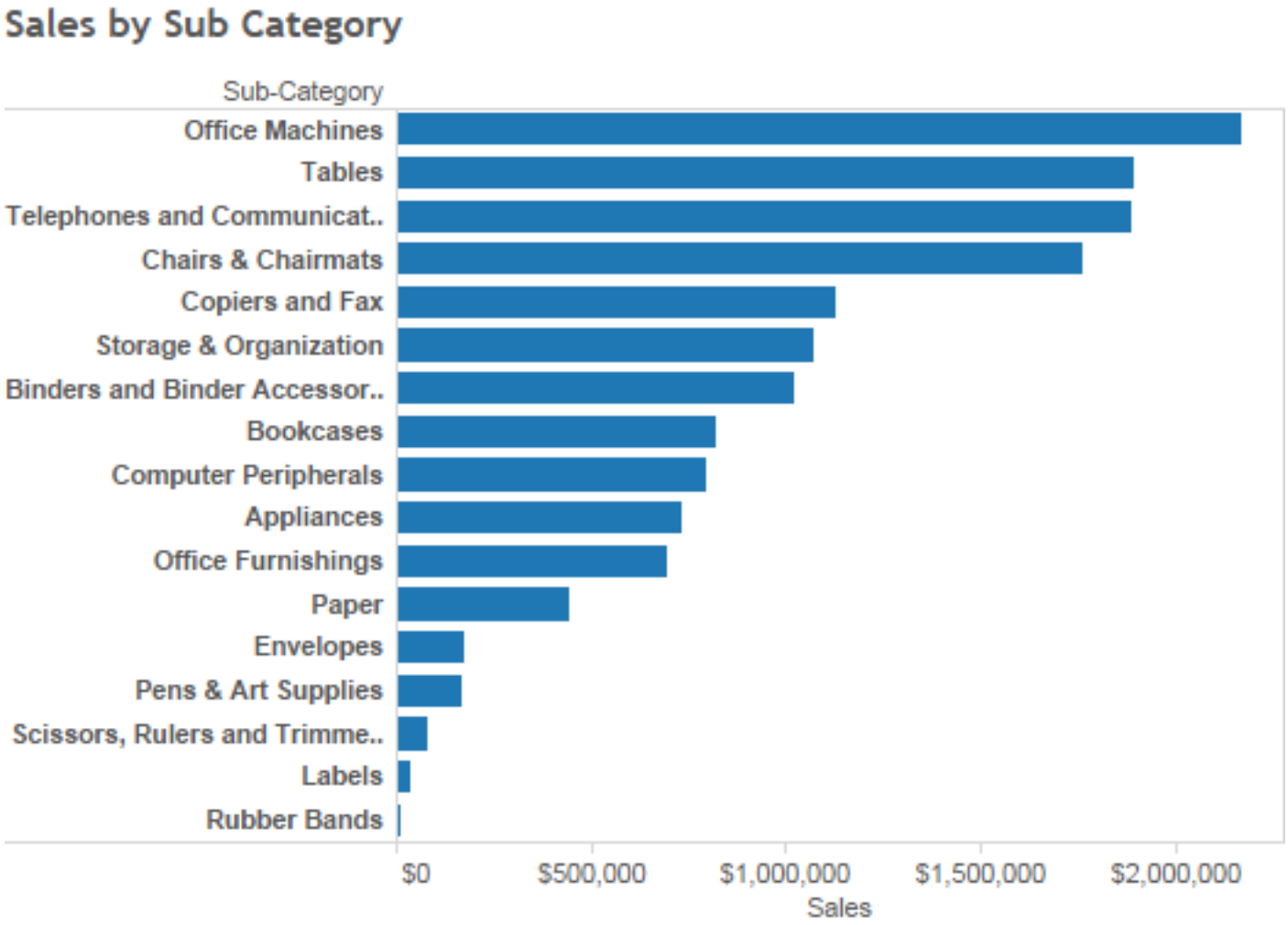
Introduction to Visual Analysis



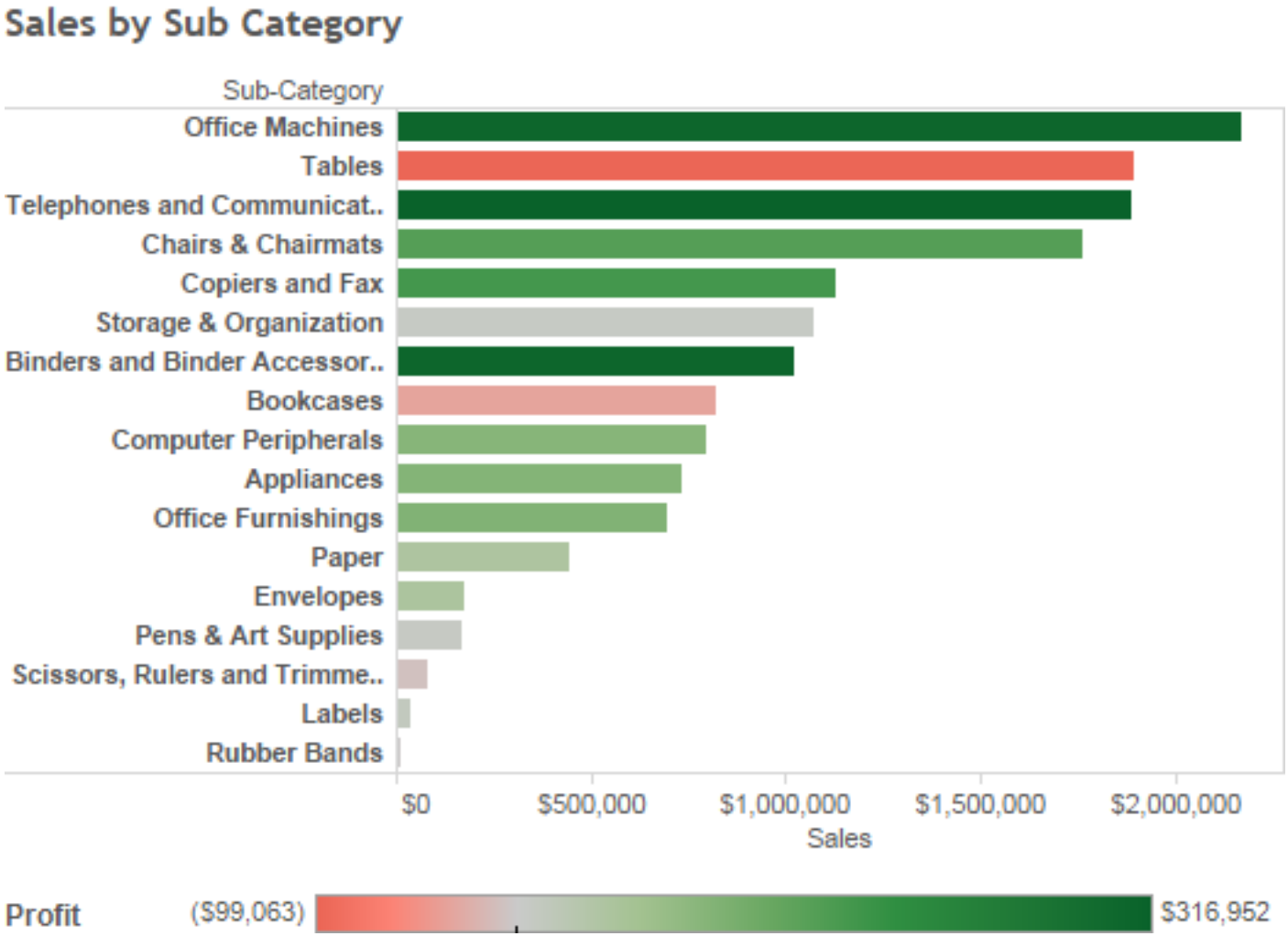
“Visual analysis means exploring data visually.
A story unfolds as
you navigate from one visual summary into
another.”

Hanrahan, Stolte, Mackinlay. “Visual Analysis for Everyone." *Tableau White paper* (2007).

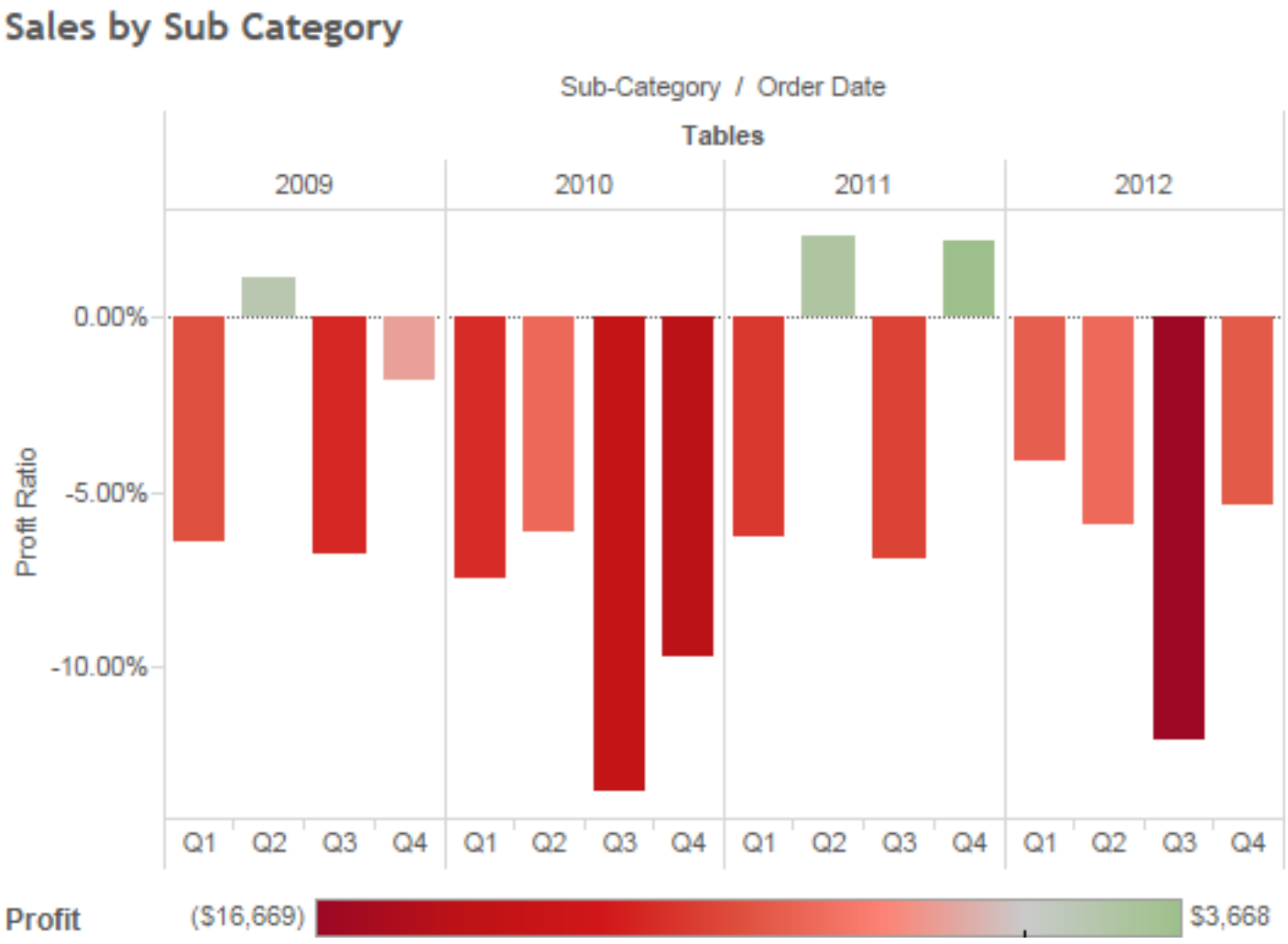
“Visual analysis means exploring data visually…



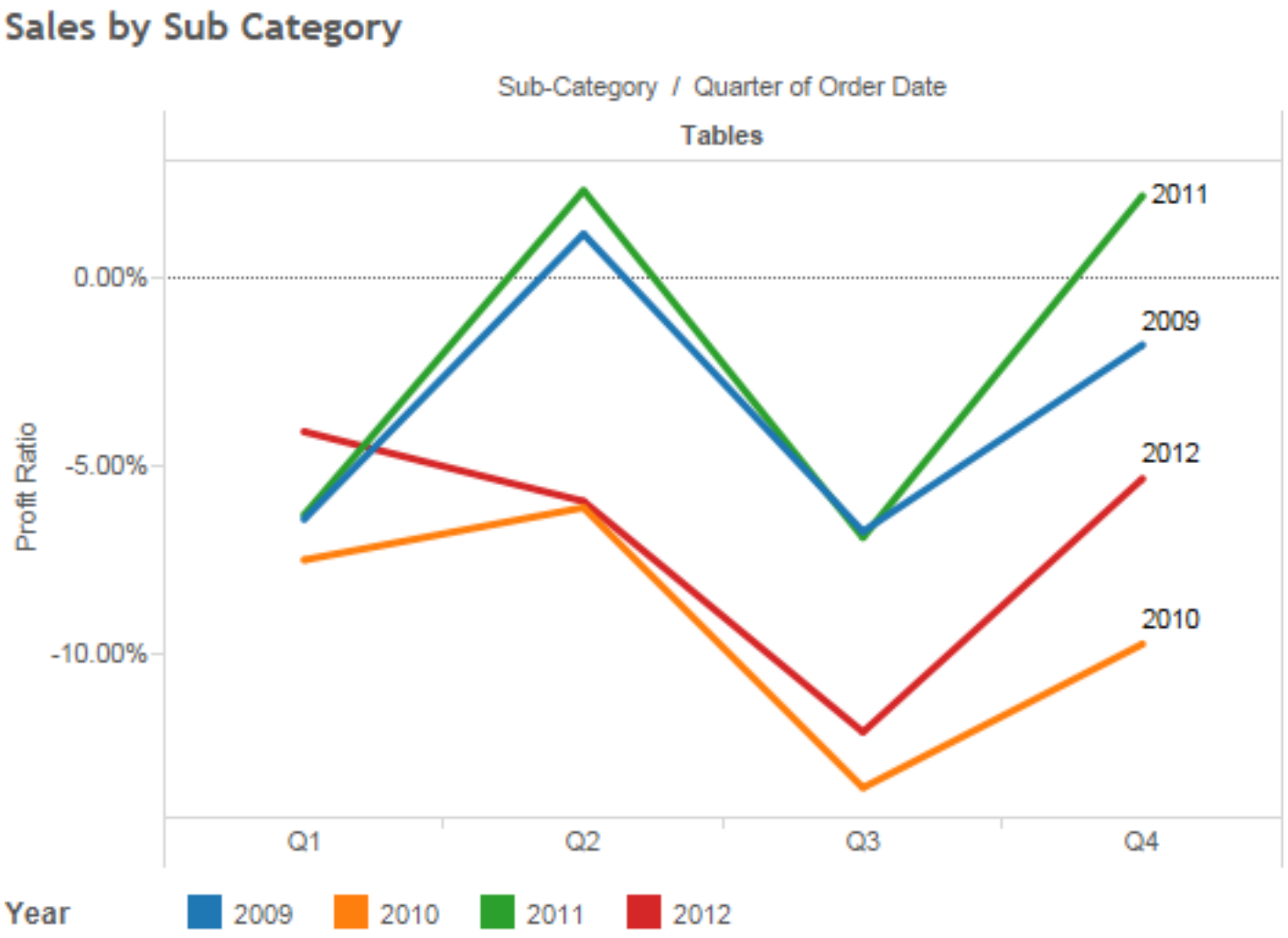
... a story unfolds as
you navigate from one visual summary into another."



Instantly change *what* you are looking at



Instantly change the *way* you are looking at it



“Simplicity is the essence of clear communication”

- Steve Jobs

Count how many times the number “5” appears

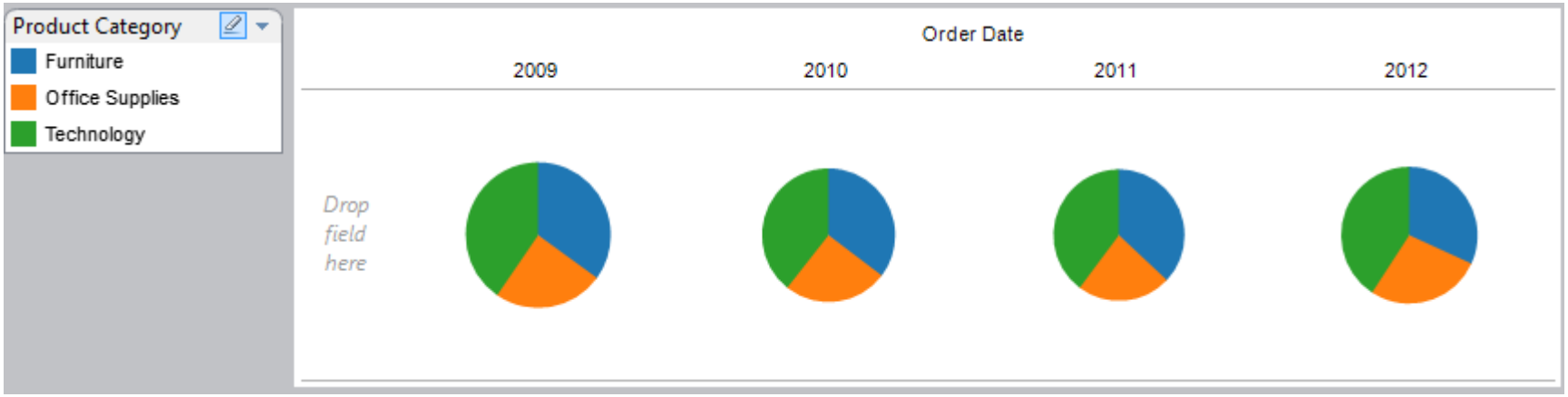
Now count the number of time “5” appears in the same set of numbers

Attentive Processing

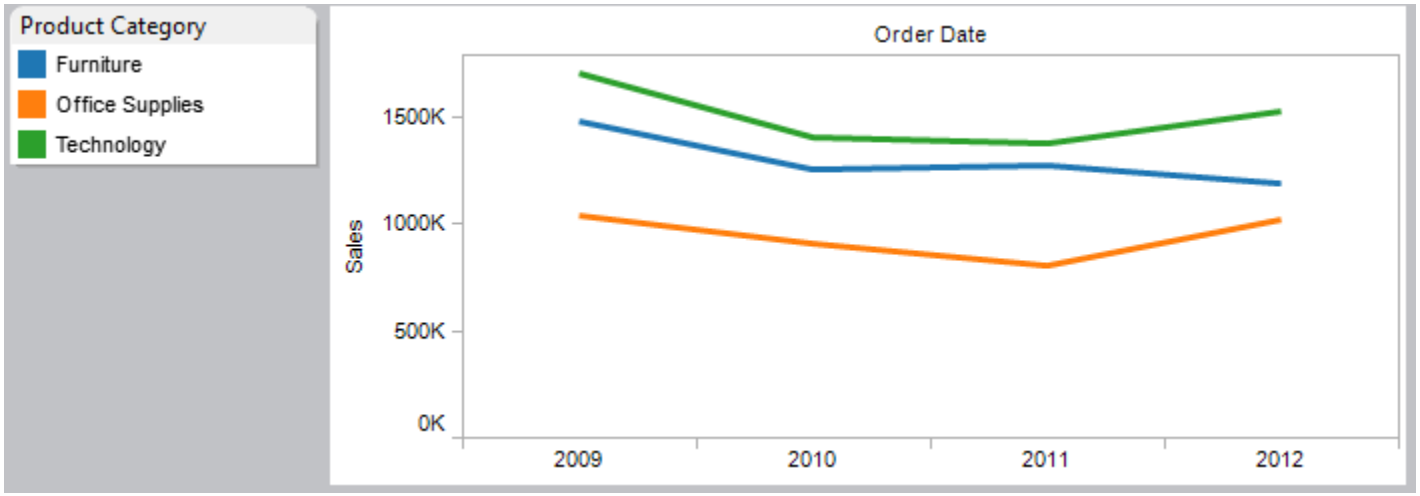
9873497902756479028947286240924060370705702790728032080290073025012702370083740820787202720070832478026027037937757097073779706674620970947027809279797097230972309795927509272797987349726080 27

Pre-attentive Processing

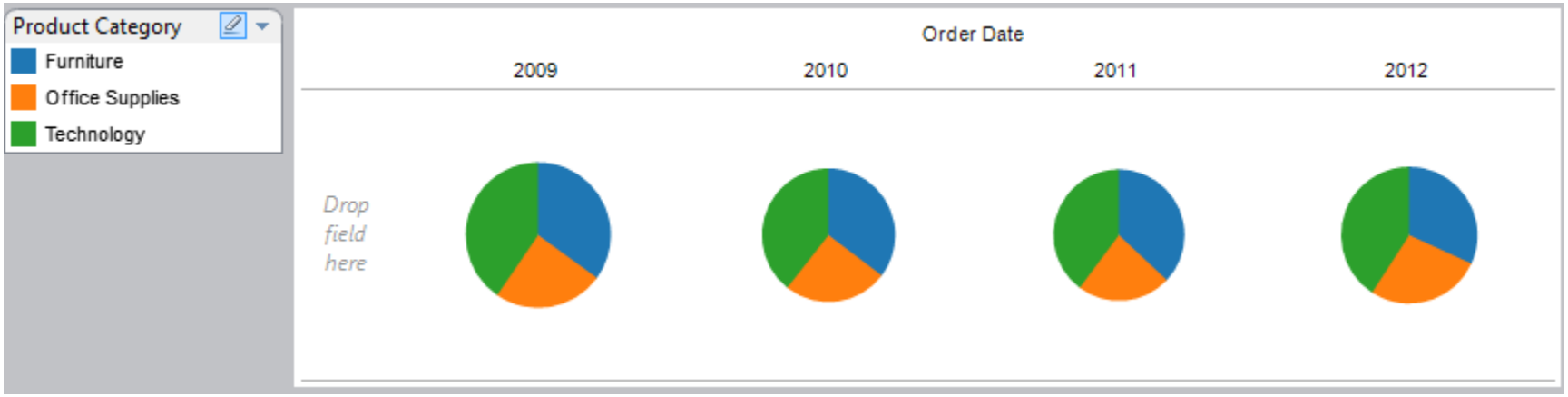
Which product categories are trending up?



Which product categories are trending up?



Attentive



Pre-attentive

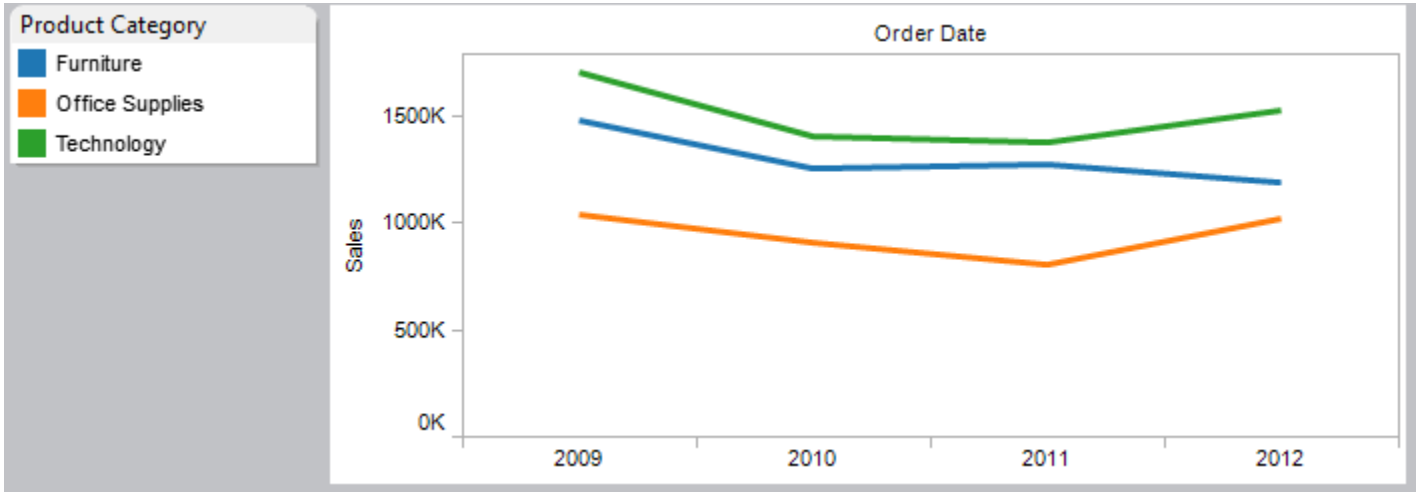


Tableau Product Family

Tableau Desktop
- Author/Create
- Visualize

Tableau Reader
- Free Desktop Client
- Only packaged workbooks
- No data Security

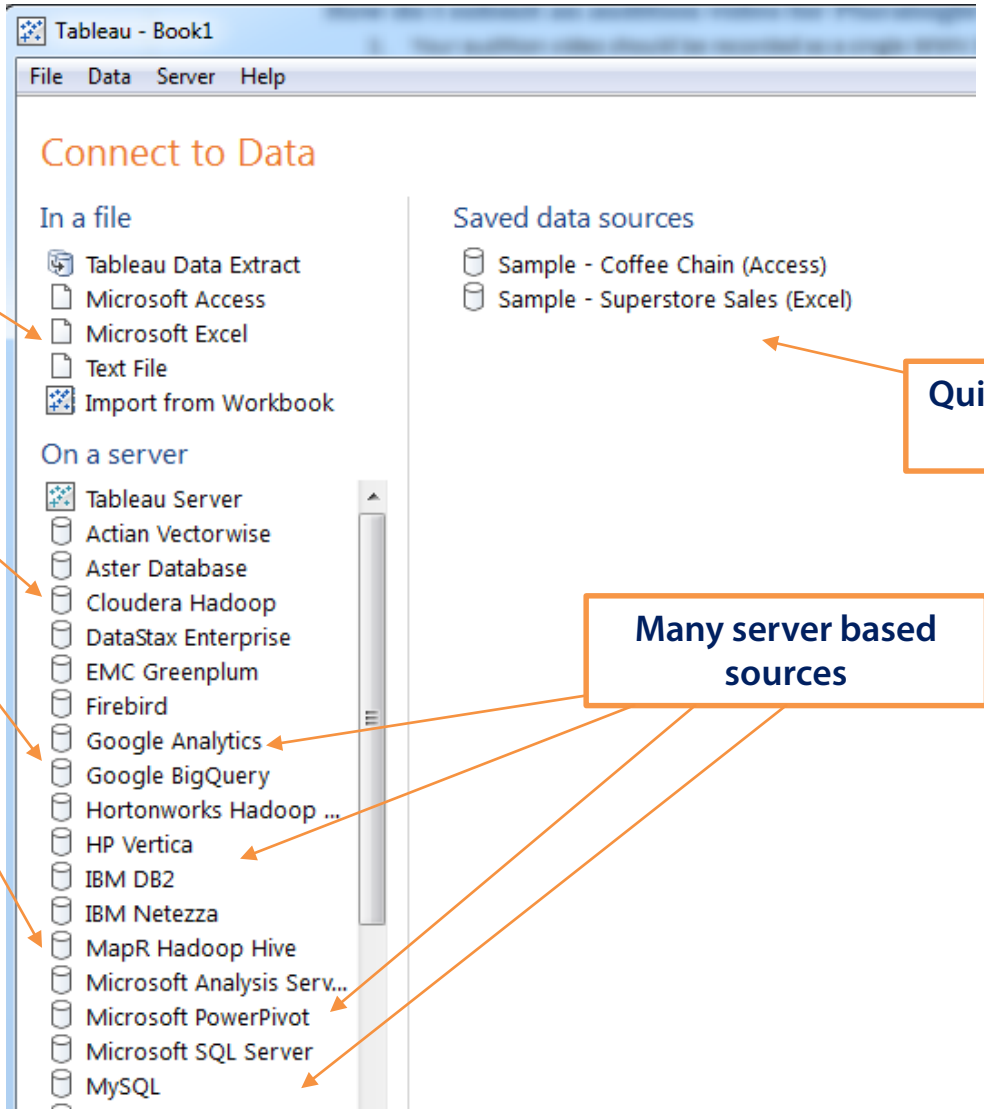
Tableau Public
- Free hosted solution
- No security
- Data source & size limitations

Tableau Server
- Self hosted
- Scalable & Secure
- Enterprise Solution

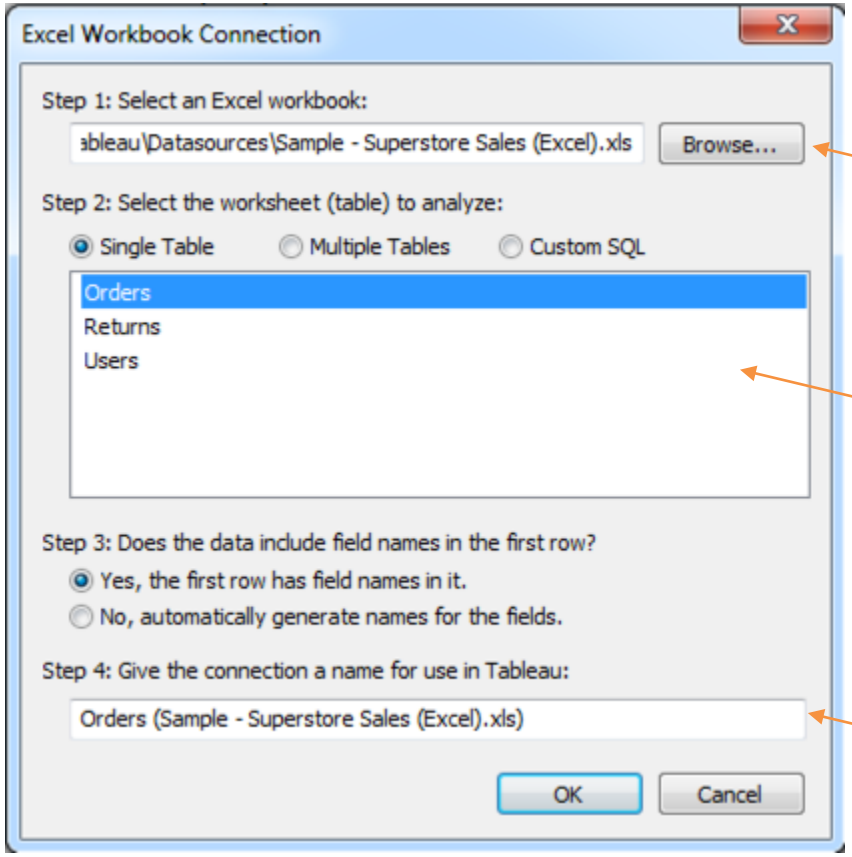
Tableau Online
- Paid SaaS solution
- Data source limitations
- Named User Licensing

http://www.tableausoftware.com/products

Connecting to Data



Setting up the Connection – Single Table

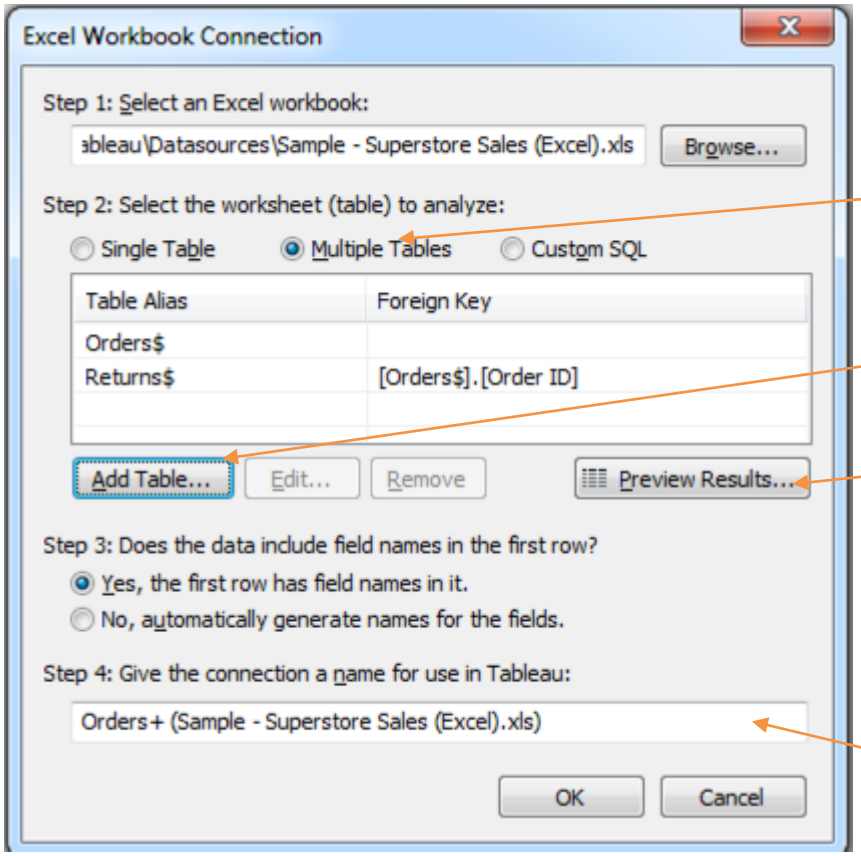


Connection String

Choose a single table

Name the connection

Setting up the Connection – Multiple Tables



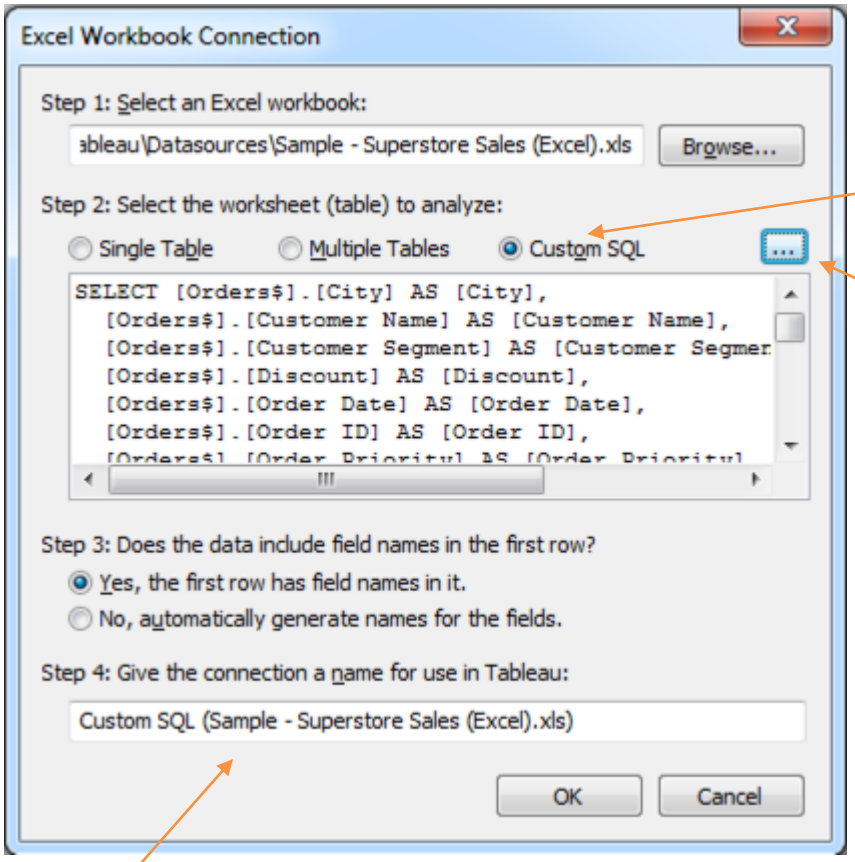
Choose Multiple Tables

Add Table

Preview Results

Name the connection

Setting up the Connection – Custom SQL

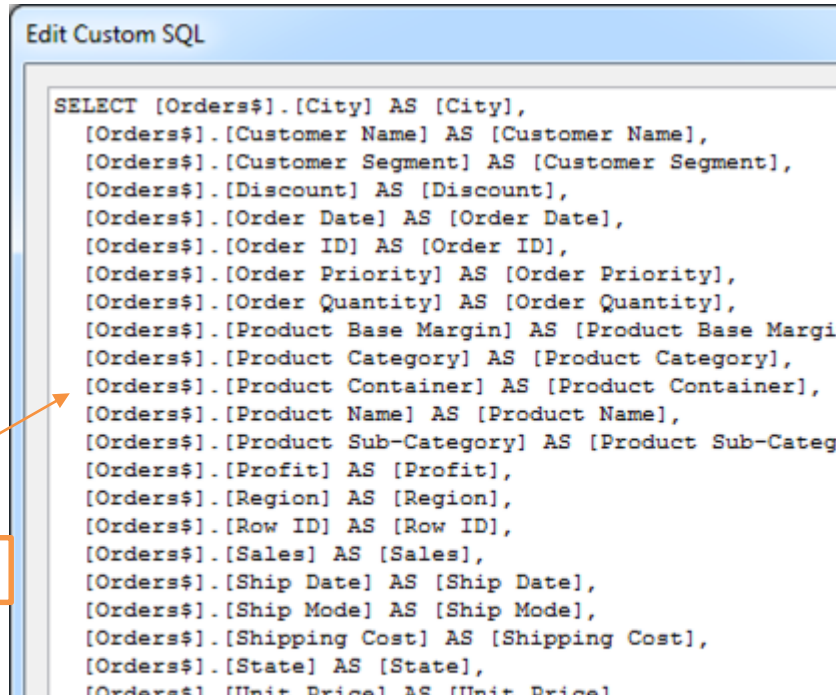


Choose Custom SQL

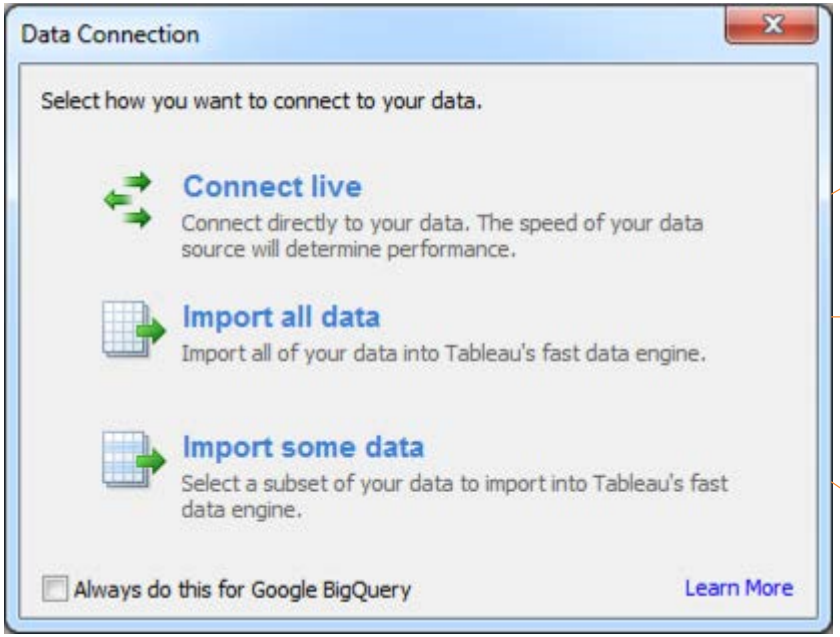
Click the elipse (…)

Name the connection

Enter SQL



Data Connection Options

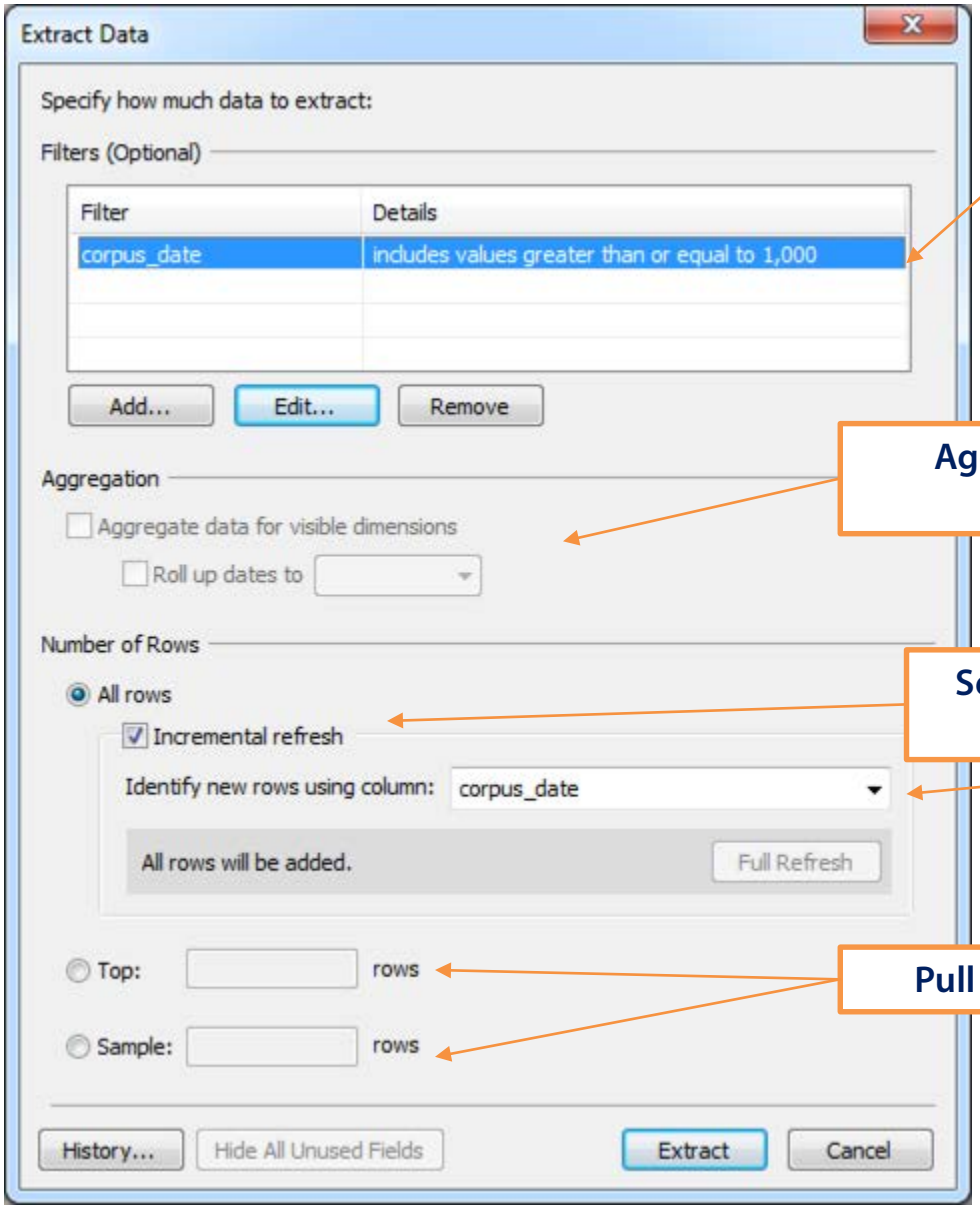


Real-time results directly from data source but could be slow depending on source.

All data is imported to Tableau's own Columnar format called a Tableau Data Extract (.tde)

Some data is imported and you have more options for future refreshes of the data source. If I'm importing data I often choose this option.

Importing Some Data



Filter down to relevant data set

Aggregate for increased performance

Setup Incremental data refreshes using dates or integers

Pull a sample data set only

The Data Pane

Our Connections (no limit to how many)

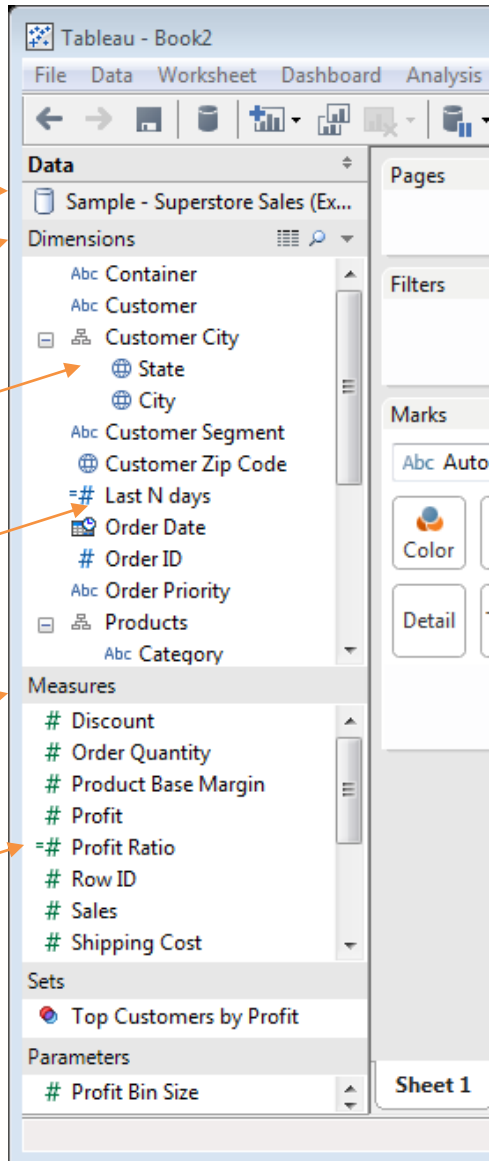
Dimensions (context / level of detail)

Hierarchy (drill-down attributes)

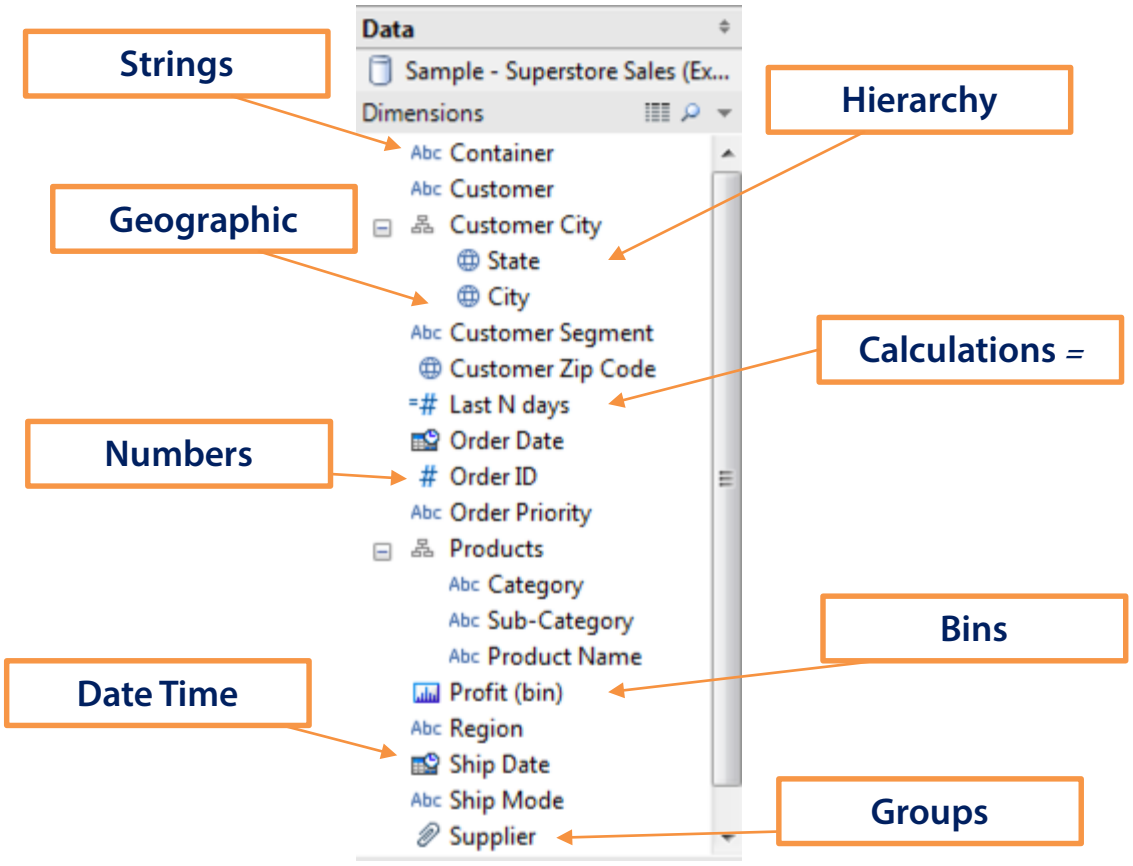
A Calculation (equal sign indicator)

Measures (aggregates)

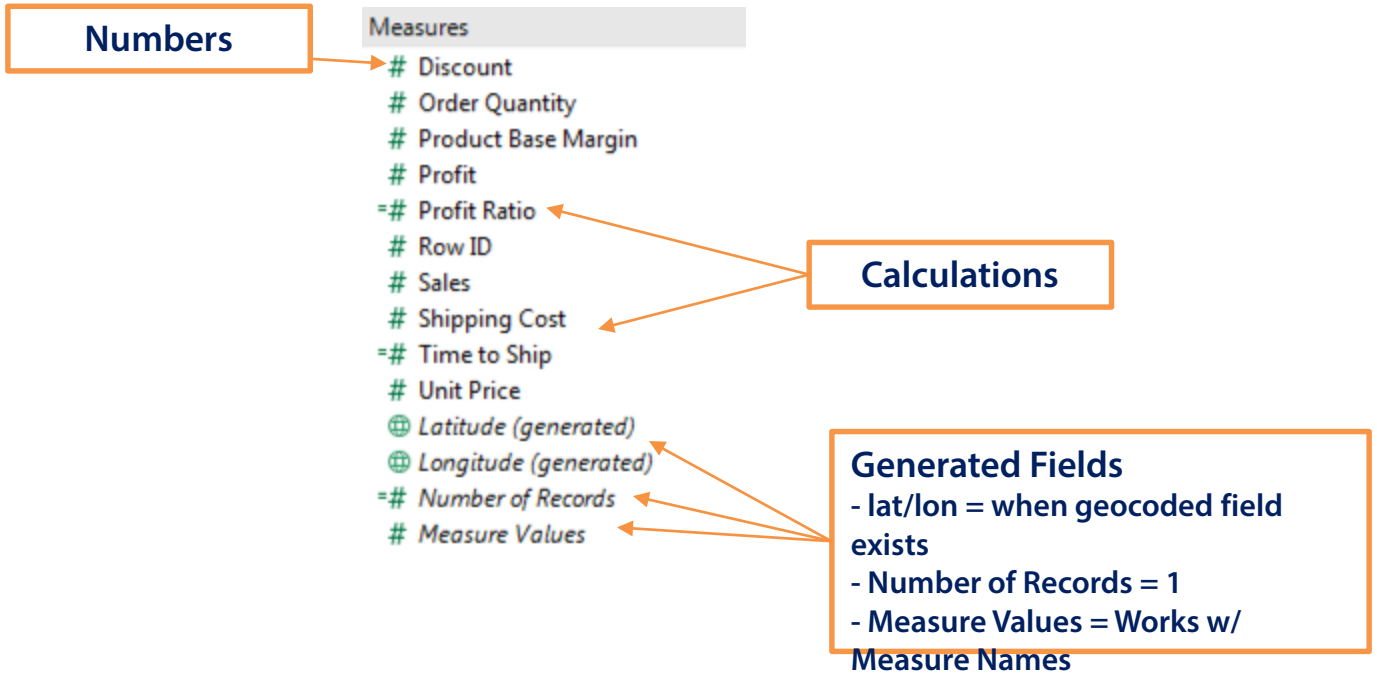
A Calculation (equal sign indicator)



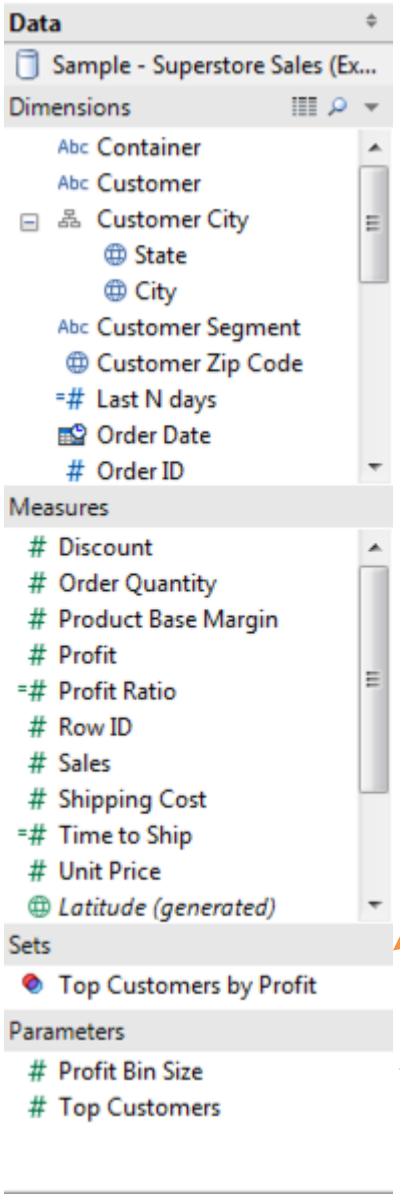
Data Terminology - Dimensions



Data Terminology - Measures



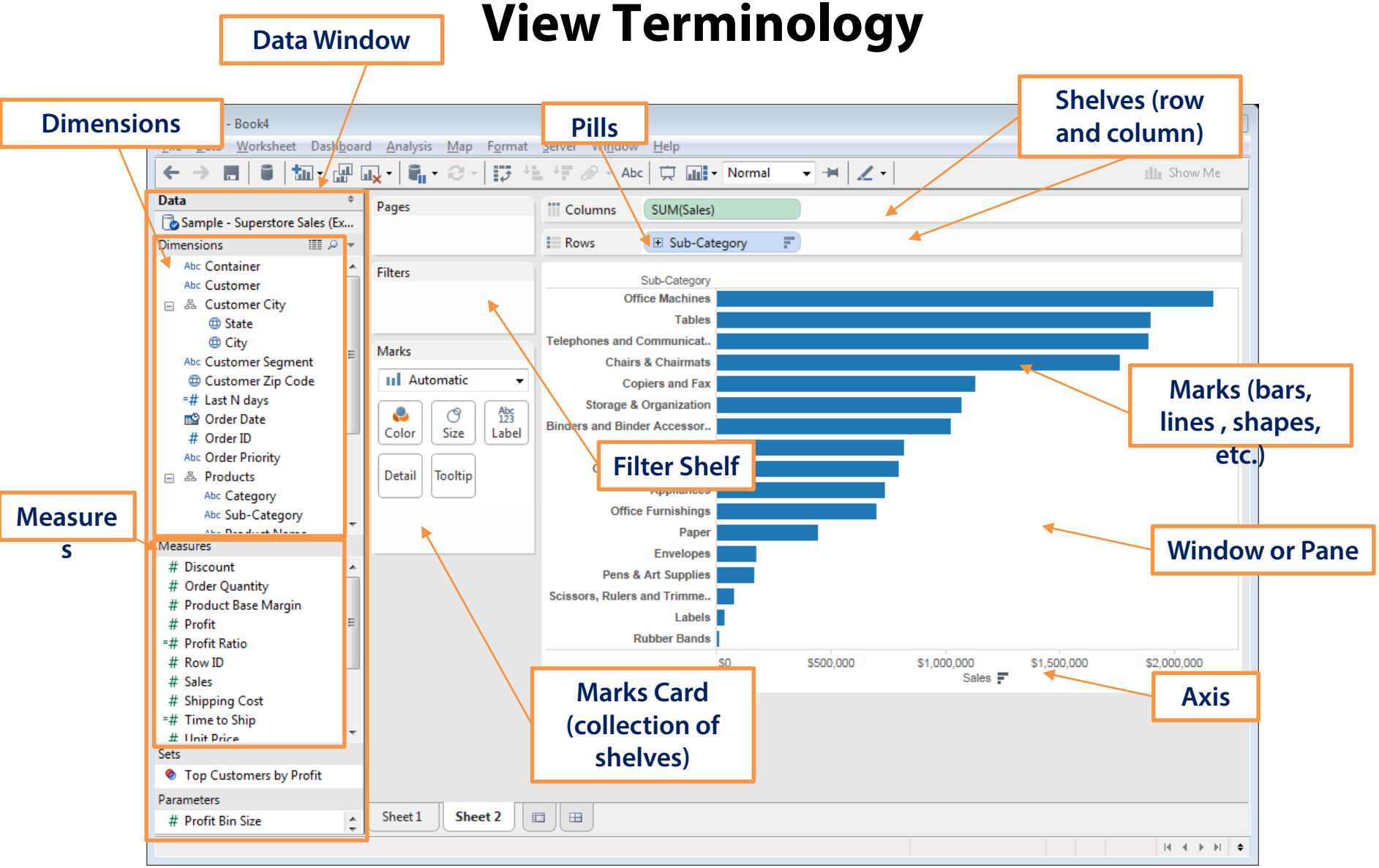
Data Terminology – Sets & Parameters



Sets
- collections of dimension members

Parameters
- Independent variables that allow user input

View Terminology



Summary

- Visual analysis
- Visual perception
- Tableau's product family
- Connecting to your data
- Data terminology
- View terminology

Getting dirty with your data

Ben Sullins
Bensullins.com
ben@bensullins.com

pluralsight
hardcore developer training

Overview

- Connecting to your data
- Joins and Blending
- Filtering
- Sorting
- Hierarchies
- Grouping
- Sets
- Data Types
- Dates
- Measure Names / Values

Connecting

Joins and Blending

Filtering

Sorting

Hierarchies

Grouping

Sets

Data Types

Date Values

Measure Names / Values

Show me the data!

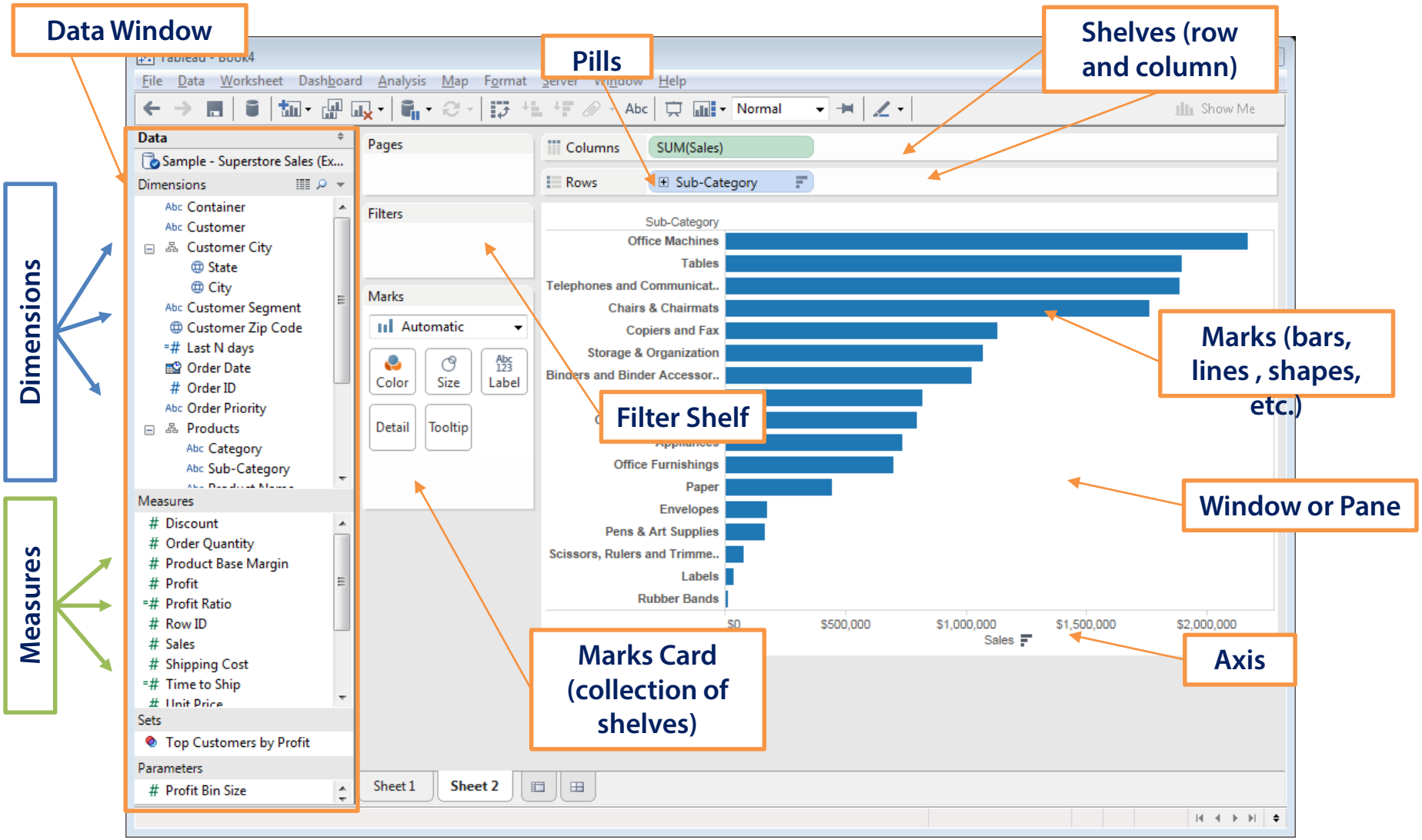
Ben Sullins
Bensullins.com
ben@bensullins.com

pluralsight
hardcore developer training

Overview

- Terminology Review
- Basic Charting
- Mapping
- Formatting
- Advanced Charting

Terminology / User Interface



Basic Charting

Mapping

Formatting

Advanced Charting

Summary

- Terminology Review
- Basic Charting
- Mapping
- Formatting
- Advanced Charting

The Intelligence Factor

Ben Sullins
Bensullins.com
ben@bensullins.com

pluralsight
hardcore developer training

Overview

- Basic Calculations
- Table Calculations
- Parameters
- Stats Functions

Basic Calculations

Table Calculations

Parameters

Stats Functions

Summary

- Basic Calculations
- Table Calculations
- Parameters
- Stats Functions

Painting the whole picture

Ben Sullins
Bensullins.com
ben@bensullins.com

pluralsight
hardcore developer training

Overview

- Dashboards Defined
- Basic Dashboarding
- Formatting
- Actions

Dashboards Defined

Visual display
of the
most important information needed to
achieve one or more objectives
which
fits entirely on a single screen
so it can be
monitored at a glance

Stephen Few, “Dashboard Confusion” 2004

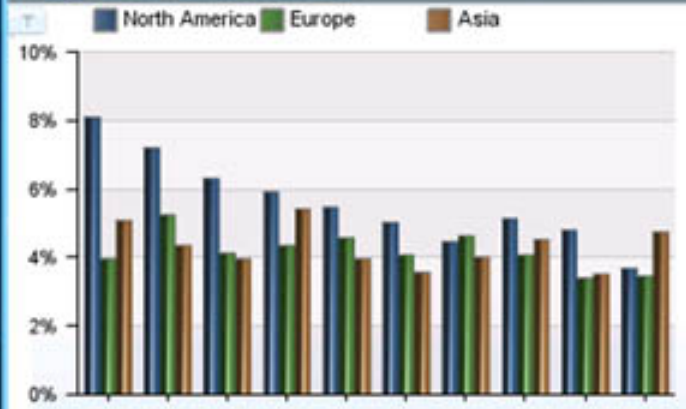
Examples

Layout and Style | Insert | My Workbench

Regional Performance

Win/Loss Trend	Region	KPP	Rep Days	Win / Loss %	Win / Loss Target
	North America		1869	37.06%	
	Europe		609	33.03%	
	Asia		2168	10.81%	

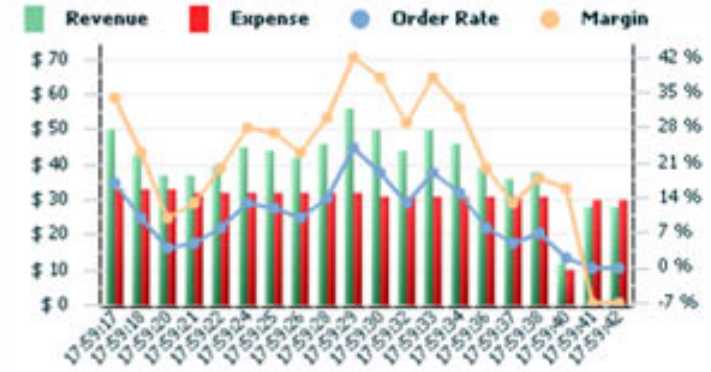
Regional Market Share



Product Performance

Customer Region	Product Line	Growth	Win / Loss %	Shipping Cost
Asia	Alpha	6%	6.6%	6,964,426.54
	Charger	7%	20.1%	2,424,143.07
	Nova	7%	22.8%	2,168,947.17
Asia - Summary		**7%**		
Europe	Alpha	21%	29.0%	1,302,856.25
	Charger	26%	24.5%	1,487,554.56
	Nova	29%	41.0%	700,060.55
Europe - Summary		**25%**		

Advertising Performance



Time Period

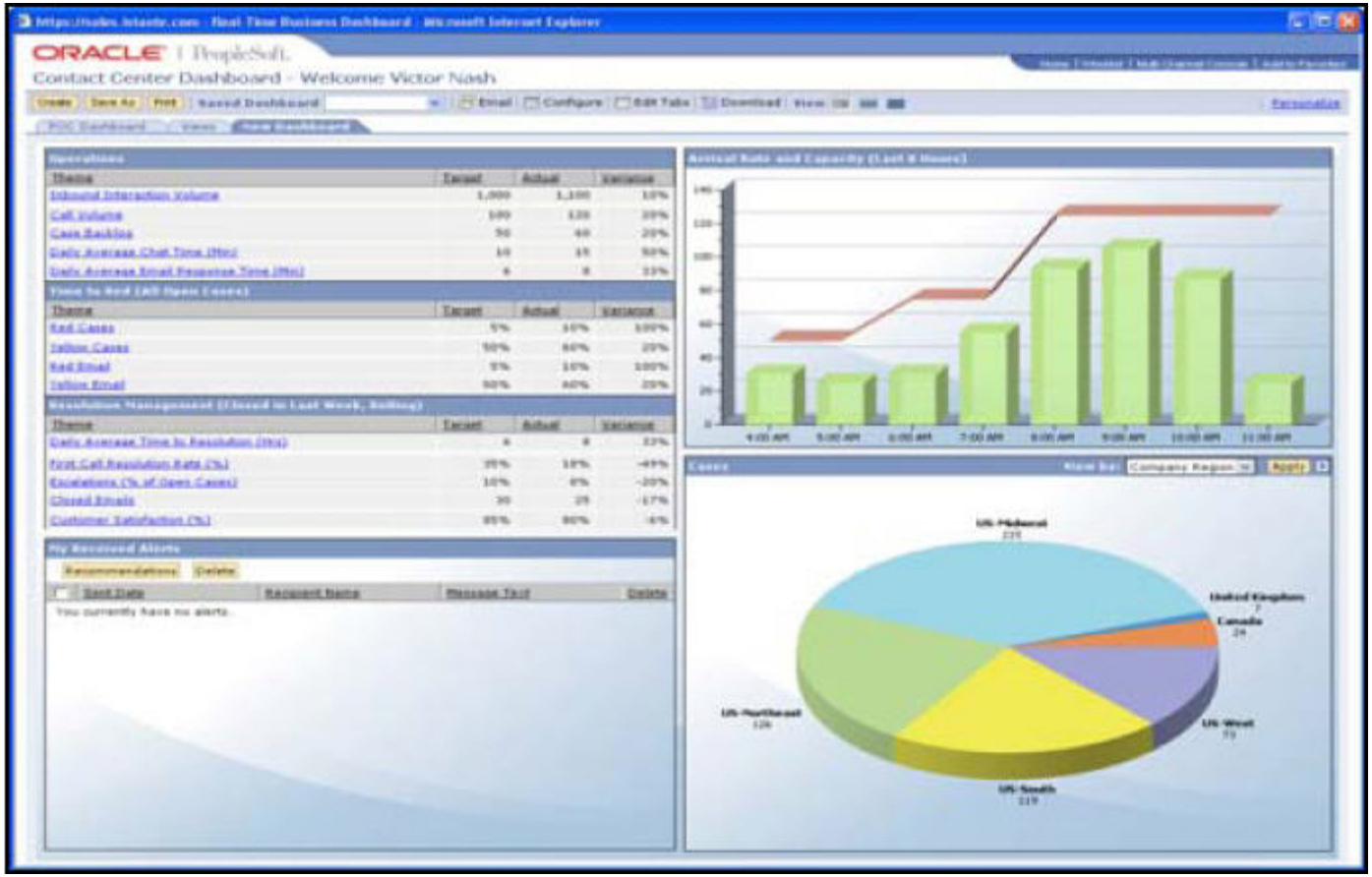
2008-Q3 2008-Q4 2009-Q1 2009-Q2 2009-Q3 2009-Q4 2010-Q1 2010-Q2 2010-Q3 2010-Q4

2008-Q3 - 2010-Q4

Future Chips INTERNATIONAL

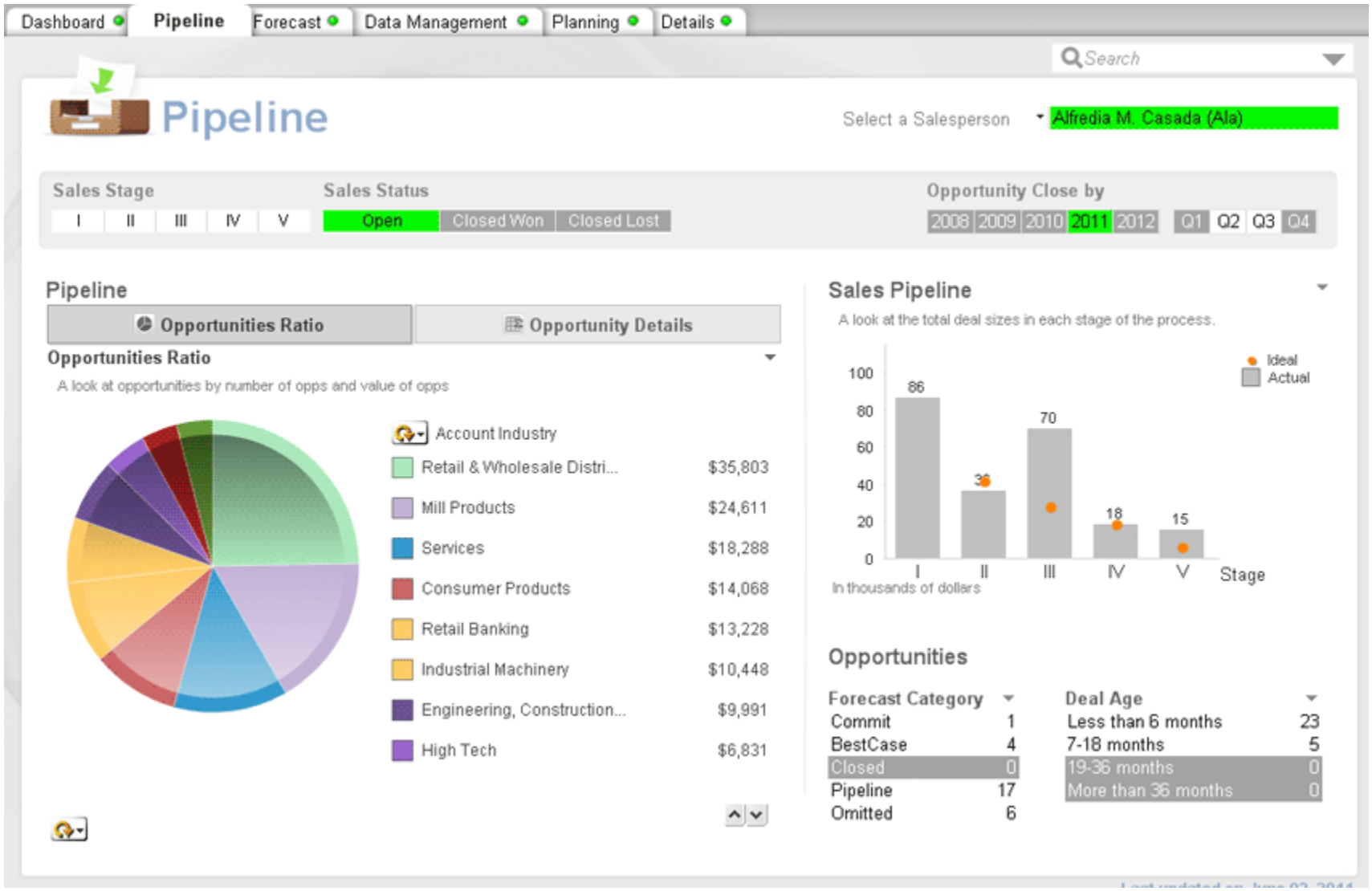
cognos

Microstrategy



Oracle PeopleSoft

Business Objects


Dashboard
Pipeline
Forecast
Data Management
Planning
Details
Search
Pipeline
Select a Salesperson
Alfredia M. Casada (Ala)
Sales Stage
I II III IV V
Sales Status
Open
Closed Won
Closed Lost
Opportunity Close by
2008 2009 2010 2011 2012
Q1 Q2 Q3 Q4
Pipeline
Opportunities Ratio
Opportunity Details
Opportunities Ratio
A look at opportunities by number of opps and value of opps
Account Industry
Retail & Wholesale Distri... $35,803
Mill Products $24,611
Services $18,288
Consumer Products $14,068
Retail Banking $13,228
Industrial Machinery $10,448
Engineering, Construction... $9,991
High Tech $6,831
Sales Pipeline
A look at the total deal sizes in each stage of the process.
Ideal
Actual
100
80
60
40
20
0
86
36
70
18
15
I II III IV V
Stage
In thousands of dollars
Opportunities
Forecast Category
Commit 1
BestCase 4
Closed 0
Pipeline 17
Omitted 6
Deal Age
Less than 6 months 23
7-18 months 5
19-36 months 0
More than 36 months 0

QlickView

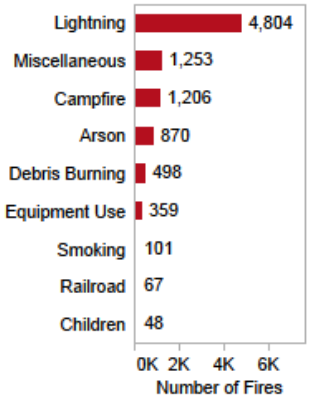
U.S. Forest Fire Hot Spots, 2002-2012

Designed by: DataRemixed

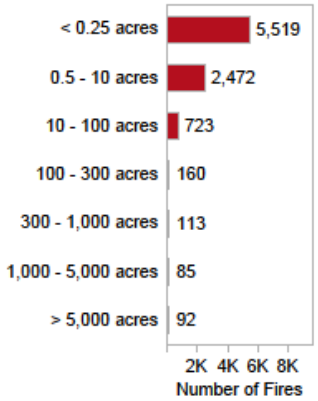
10 years of U.S. forest fire data reveals that most fires burn less than a quarter of an acre, occur in the western part of the country during the summer months and are caused by lightning.
Clicking on a bar filters all other views.

Select a Year:
2007

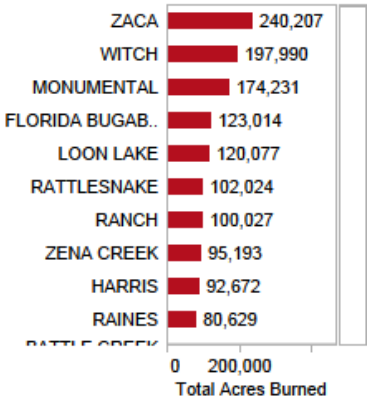
What was the cause?



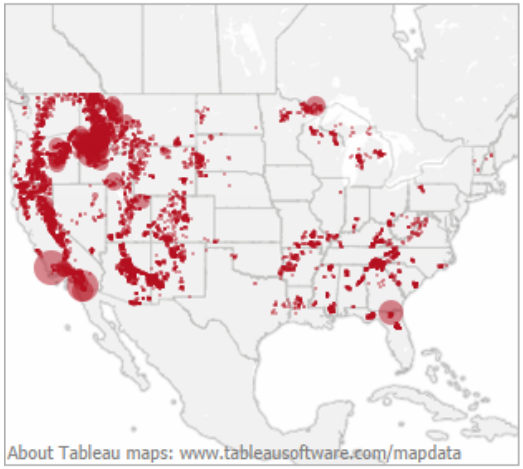
How big is the fire?



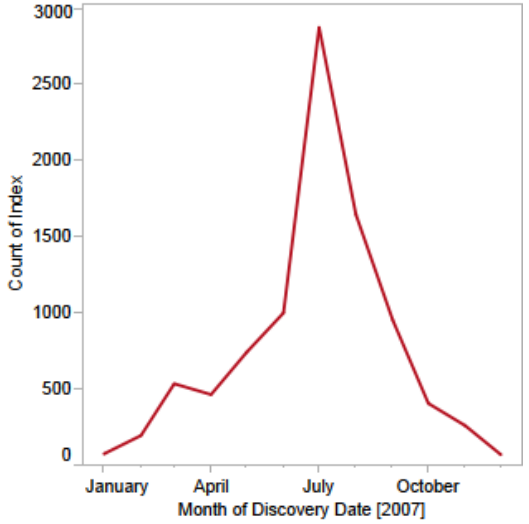
Largest fires by location



Where do they occur?



What time of the year do they occur?



Source: https://fam.nwcg.gov/fam-web/

tableau

Tableau

Now | Network | ERP | Data Warehouse | Web Site | Email

CIO Dashboard

Help

(As of December 19, 2004, 3:35 PM)

System Availability (Last 30 days)

(Actual; Acceptable)

Last 12 Months	System	Availability %
	Network	98.5%
	ERP	97.9%
	Data Warehouse	93.2%
	Web Site	98.5%
	Email	100.0%
	HR	96.7%
	Problem Tracking	94.7%

85% 90% 95% 100%

Hardware % of Capacity

(Actual; Good; Excessive; Critical)

CPU	Today	Overall	79%
Storage	Last 12 Mo.	Today	63%

0% 50% 100%

Daily Network Traffic (Kilobytes)

Daily mean for last 6 months — Daily mean for last 7 days — Yesterday

200K 150K 100K 50K 0K

12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12

AM PM

Response Time (Distribution in seconds)

(Acceptable)

0 1 2 3 4 5 6 7 8 9 10 11 12

DW ERP Web HR

Key Non-System Metrics

(Actual; Good; Excessive; Critical)

Year-to-Date	Metric	% of Target	Actual
	Expenses YTD		$2,456.5K
	Customer Satisfaction		3.25 of 4
	Level 1 Problems		87

0% 50% 100% 150%

Major Project Milestones

Project	Milestone	Due Date	Days Until/ Past Due
ERP Upgrade	Full system test	01/10/06	22
Add services data to DW	ETL coding	12/15/05	-3
Upgrade mainframe OS	Prepare plan	12/23/05	4
Disaster recovery site	Install hardware	01/08/06	20
Budgeting system	Hire team	12/06/05	-13
Web site face-lift	Move into production	01/05/05	17

5 Top Projects in the Queue (Sorted by priority)

	Project	Status	Funding Approved	Sched. Start
1	Professional services module	Pending available staff	X	05/10/06
2	Upgrade MS Office	Cost-benefit analysis		02/15/06
3	Failover for ERP	Preparing proposal		06/02/06
4	Upgrade data warehouse HW	Evaluating options	X	04/15/06
5	Executive dashboard	Vendor assessment		07/01/06

Critical Events (Next 14 Days)

Event	Group Responsible	Date
Full system maintenance outage from 9-11 PM	G. Jones	12/21/05
Present hardware upgrade proposal to CEO	Self/M. Smith	12/22/05
Tom visiting from Asia office	Self	12/23/05
Prepare quarterly financial for public announcement	Self	01/04/06
Present revised information strategy to steering comm.	J. Kane	01/06/05

Basic Dashboarding

Formatting

Actions

Summary

- Dashboards Defined
- Basic Dashboarding
- Formatting
- Actions

Sharing your work

Ben Sullins
Bensullins.com
ben@bensullins.com

pluralsight
hardcore developer training

Overview

- Overview of Sharing Options
- Tableau Reader
- Tableau Public
- Tableau Server
- Tableau Online

Tableau Sharing Options

Tableau Desktop
- Author/Create
- Visualize

Tableau Reader
- Free Desktop Client
- Only packaged workbooks
- No data Security

Tableau Public
- Free hosted solution
- No security
- Data source & size limitations

Tableau Server
- Self hosted
- Scalable & Secure
- Enterprise Solution

Tableau Online
- Paid SaaS solution
- Data source limitations
- Named User Licensing

http://www.tableausoftware.com/products

Tableau Reader

Tableau Public

Tableau Server

Tableau Online

Summary

- Overview of Sharing Options
- Tableau Reader
- Tableau Public
- Tableau Server
- Tableau Online