

# Lesson 1:

Introduction to tableau  
and field types



# Module Outline

- Introduction to Tableau
- Tableau Features
- Tableau Products
- Tableau Desktop Architecture
- Design Flow
- Tableau Navigation
- File Types
- Data Types and Tableau Terminologies
- Data Sources



# Module Outline (Cont...)

- Tableau Calculations
- Sort and Filter
- Tableau Charts
- Formatting
- Dashboard
- Forecasting
- Trend Line

# Tableau introduction :



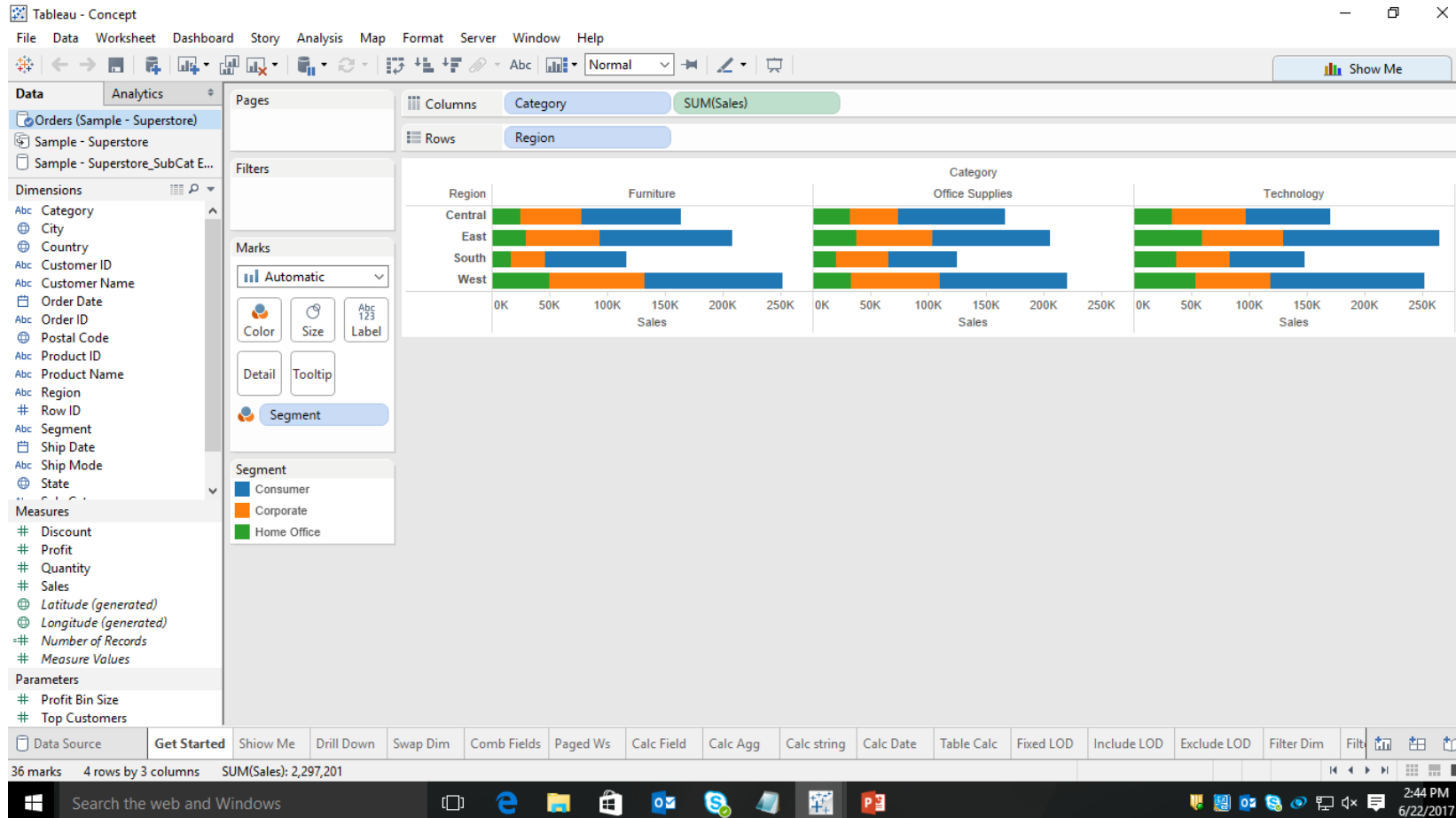
Tableau is a business intelligence (BI) tool that can help you create beautiful and visually-appealing reports, charts, graphs and dashboards using your data. These reports are interactive and can easily be shared with anyone

- **Tableau Products** : Tableau Desktop , Tableau reader , Tableau server , Tableau online , Tableau public
- **Tableau Desktop** : All your development is done here. From creating reports, charts, formatting them, putting them together as a dashboard all the work is done on Tableau Desktop.
- **Tableau Server** : The dashboards you create are shared with other users using Tableau Server. When you upload(publish) a Dashboard to Tableau Server from Tableau Desktop, other users can access those Dashboards by logging on Tableau Server.
- **Tableau Public**: It's a free product from Tableau. Visualization Enthusiast can play with data using Tableau Public. It has its limitation. You can only use limited data(no. of rows), can only connect excel or text file (no database connectivity) & anything you save in Tableau Public will be saved on Tableau Public Sever which anyone can download (no confidentiality).
- **Tableau Online**: Tableau Online is a hosted Tableau Server.
- **Tableau Reader** :It is a free desktop application that you can use to open and interact with data visualizations built in Tableau Desktop. Filter, drill down and discover.



# Introduction to Tableau

- A data visualization tool with many important features
- No scripting required
- Any business user can use the and share the product





# Tableau Features

- Speed of Analysis
- Visual Discovery
- Data Blending
- Real-Time Collaboration
- Centralized Data



# Tableau Products

- Tableau Desktop
- Tableau Server and online
- Publisher

# Workbooks and Sheets

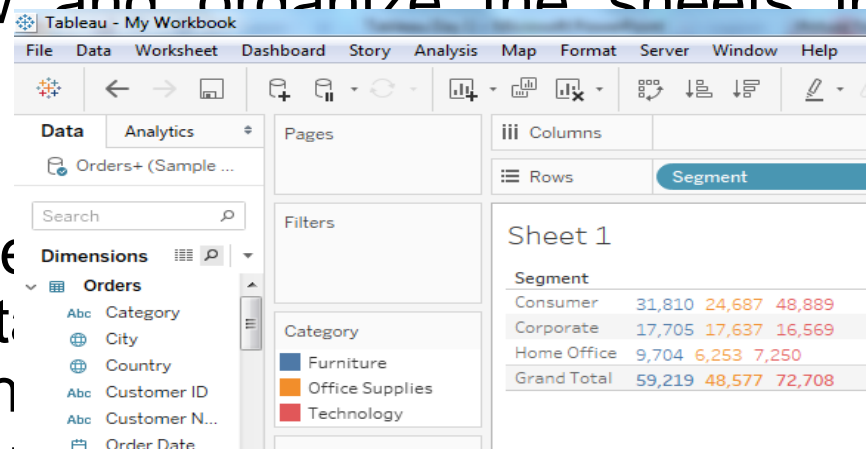


## Workbooks and Sheets

Tableau uses a workbook and sheet file structure, much like Microsoft Excel. A workbook contains sheets, which can either be a worksheet or a dashboard. A worksheet contains a single view along with shelves, legends, and the Data window. A dashboard is a collection of views from multiple worksheets. Within a workbook, you can create new sheets, clear an entire worksheet, duplicate sheets, hide or show a worksheet, and delete a sheet. Tableau has several ways to view and organize the sheets in your workbook.

## Workbooks

Tableau workbook file workbooks. They contain worksheets or dashboards. You can save, and share your results. The workbook is the container for all of your work in Tableau.





# Sheets



Each workbook can contain worksheets and dashboards. A worksheet is where you build views of your data by dragging and dropping fields onto shelves. A dashboard is a combination of several views that you can arrange for presentation or to monitor. The sheets, whether worksheets or dashboards, display along the bottom of the workbook as tabs.

## Creating Worksheets

Create a new worksheet by doing one of the following:

Selecting Worksheet > New Worksheet

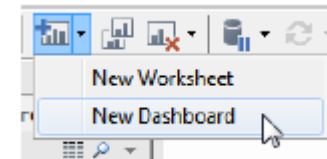
Click the New Worksheet tab at the bottom of the workbook.

## Creating Dashboards

Create a new dashboard by doing one of the following:

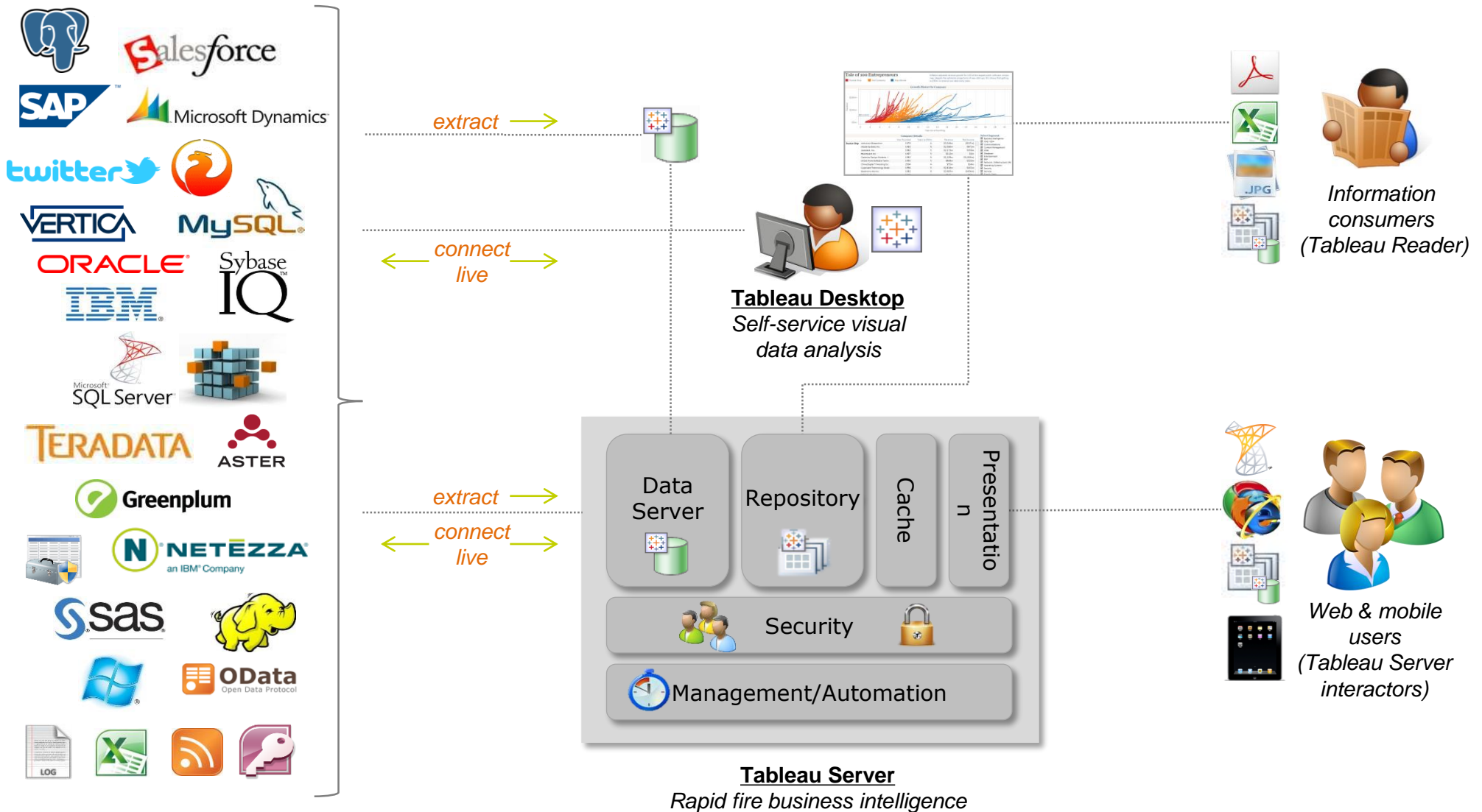
Select Dashboard > New Dashboard

Click the New Dashboard tab at the bottom of the workbook.





# Tableau Architecture





# Design Flow?

Tableau is based on three simple concepts:

- **Connect** - Connect Tableau to any database you want to analyze.
- **Analyze** - Using Tableau you can do all of these things by simply arranging the fields of your data source on a Tableau worksheet. When you drop a field on a worksheet, Tableau queries the data using standard drivers and query languages (like SQL and MDX) and presents a visual analysis of the data.
- **Share** - You can share results with others either by sharing Tableau workbooks with other Tableau users, by pasting results into other applications such as Microsoft Office, by printing to a PDF or by using Tableau Server to publish or embed your views across your organization.

## Visual Analytics

- Based on a visual query language – VizQL
- Unifies computer graphics and databases



## In-Memory Data Engine

- Breakthrough in-memory data engine
- Column storage, architecture aware
- Fast performance with massive data

## Live Query Engine

- Attaches directly to enterprise data stores
- Supports the world's largest databases
- Compatible with data of all formats and sizes





# Visual Analytics Using Tableau

The Key to Making Sense of Data is Visualization



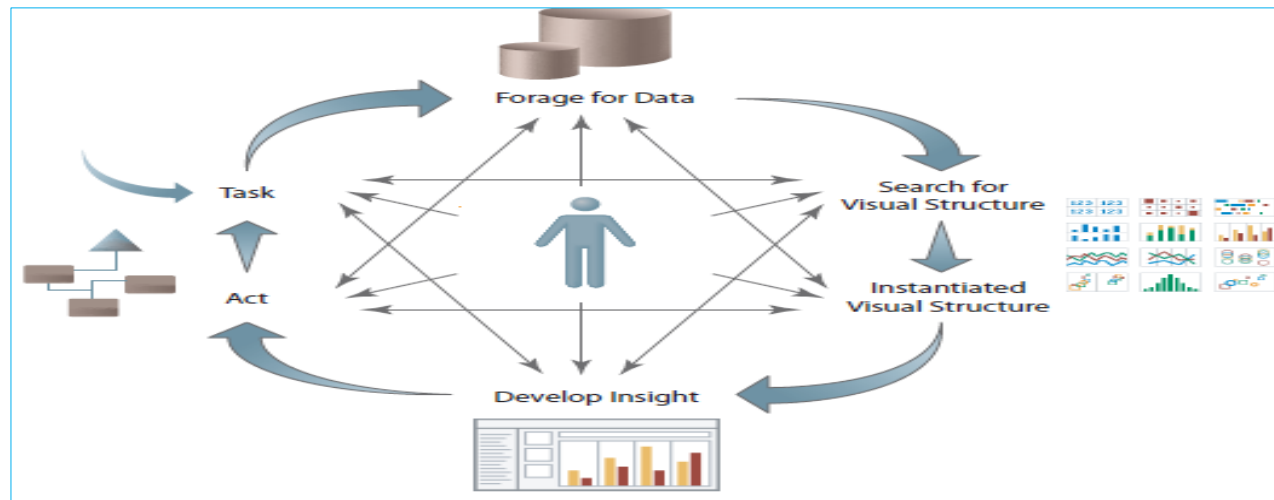


# Visual Analytics Using Tableau (contd...)

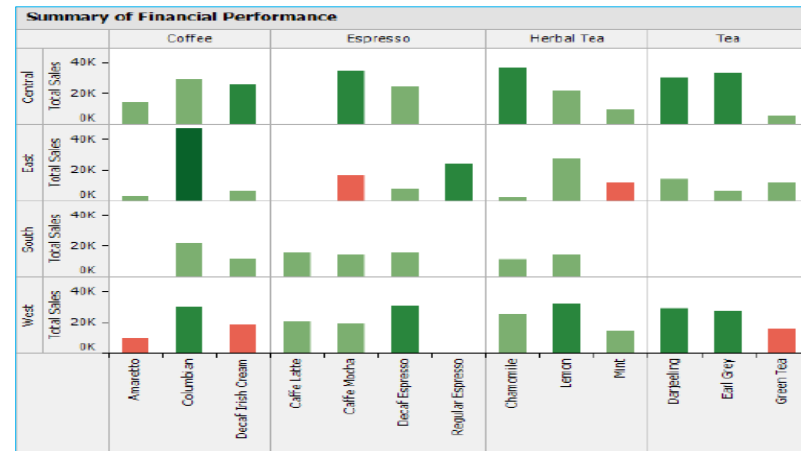
- Visual analysis tools enable power users and business analysts (such as financial, marketing, and sales analysts) to explore data sets visually and identify trends and anomalies. These tools usually work with data stored in memory and expose rich navigational features that let users explore data at the speed of thought. Many also incorporate some form of statistical or predictive analytics.
- Visual analysis tools compress and store data in memory, providing sub-second response times for any action taken against the data (such as filtering, drilling, calculating, sorting, and ranking).
- Visually, analysts point and click to interact with charts, apply filters, and change views.



# Visual Analytics Using Tableau (contd...)







Summary of Financial Performance									
		Central		East		South		West	
		Total Sales	Total Profit	Total Sales	Total Profit	Total Sales	Total Profit	Total Sales	Total Profit
Coffee	Ameretto	\$14,011	5,105	\$2,993	1,009			\$9,265	-1,225
	Columbian	\$28,913	8,528	\$47,386	27,253	\$21,664	8,767	\$30,357	11,253
	Decaf Irish Cream	\$26,155	9,632	\$6,261	2,727	\$11,592	2,933	\$18,235	-1,305
Espresso	Caffe Latte					\$15,442	3,872	\$20,458	7,502
	Caffe Mocha	\$35,218	14,640	\$16,646	-6,230	\$14,163	5,201	\$18,876	4,064
	Decaf Espresso	\$24,485	8,860	\$7,722	2,410	\$15,384	5,930	\$30,578	12,302
Herbal Tea	Regular Espresso			\$24,036	10,062				
	Chamomile	\$36,570	14,434	\$2,194	765	\$11,186	3,180	\$25,632	8,852
	Lemon	\$21,978	6,251	\$27,176	7,901	\$14,497	2,593	\$32,274	13,120
Tea	Mint	\$9,337	4,069	\$11,992	-2,242			\$14,380	4,330
	Darjeeling	\$30,289	10,772	\$14,096	6,497			\$28,769	11,780
	Earl Grey	\$32,881	10,331	\$6,505	3,405			\$27,387	10,425
	Green Tea	\$5,211	1,227	\$11,571	5,654			\$16,063	-7,109







# What can I do with Tableau?

Visual Analytics	Fast	Any Data	Business Integration	Web & Mobile Authoring
				
Forecasting	Parallelized dashboards	Salesforce.com	JavaScript API	Web authoring & editing
Sets and visual groups	Faster quick filters	Google Analytics & Google BigQuery	Data Server Security	iPad, Android authoring & editing
Treemaps, bubble charts, word clouds	Data Engine & Extract performance	Cloudera Impala, Cassandra, HortonWorks, Hadapt, Karmasphere	Server Auditing	Local rendering
New marks card	Fast graphics and calculations		Distributed Data Engine	Subscriptions
Freeform dashboards	Performance recorder	SAP HANA		
Data Blending improvements		Data Extract API		



# Tableau - Latest Features

Usability

Analytics in the Flow

Smart Maps

Data Preparation

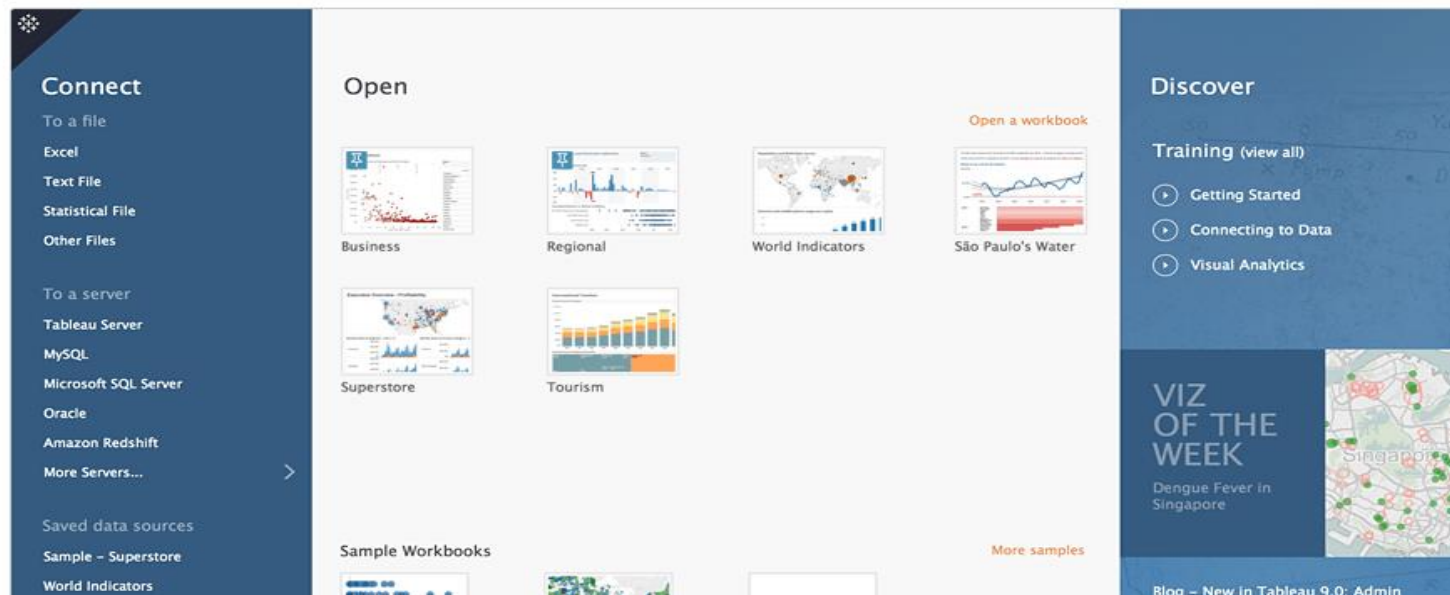
New Server & Online

Faster Performance

## Usability

Exploring and presenting your data is easier, faster, and more beautiful than ever before.

New Start Experience • Rich Story Points Formatting • Preview Thumbnails





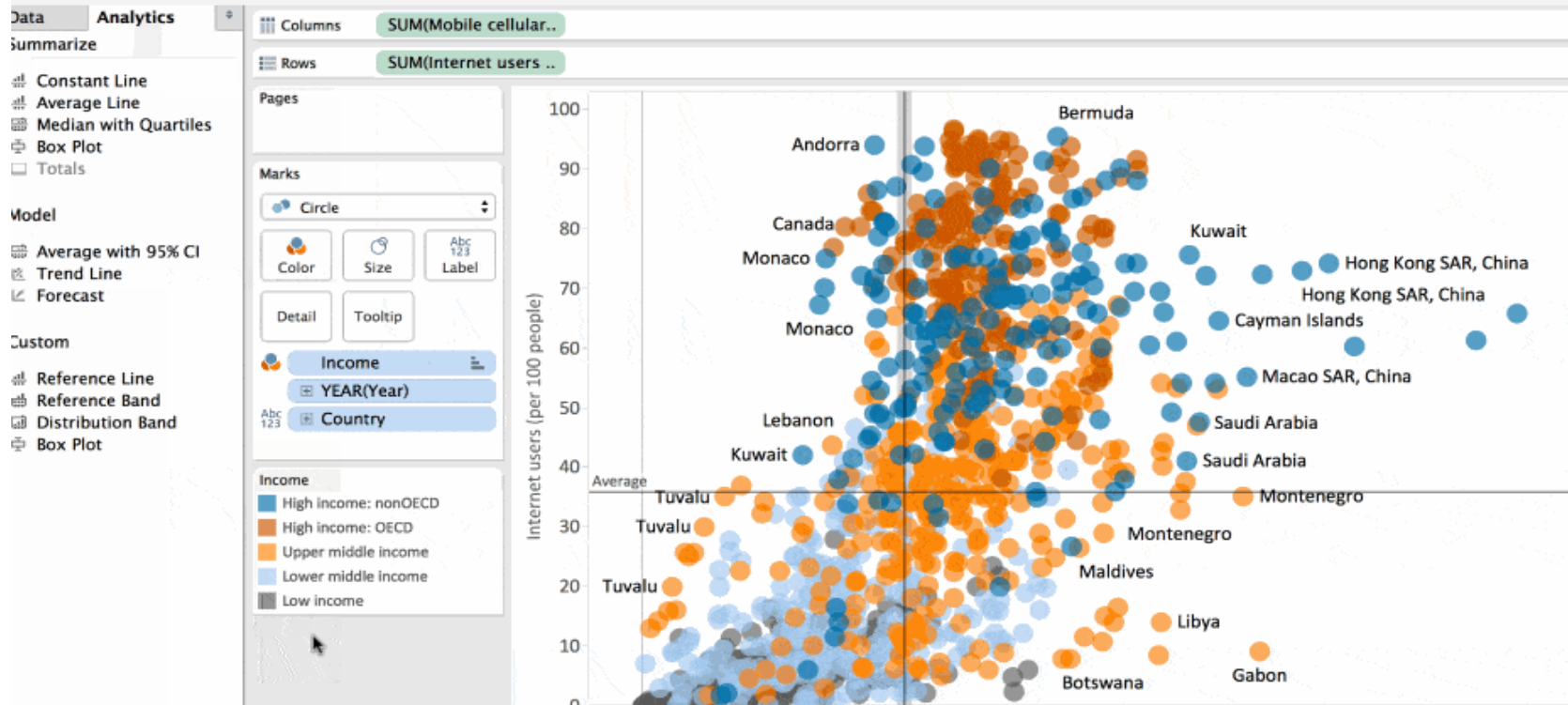


# Tableau - Latest Features (contd...)

## Analytics in the Flow

Tableau gives you instant visual feedback as you experiment with data. Answer deeper questions as the tool fades to the background.

Analytics Pane • Instant Analytics • Fast Marks & Tooltips • Smart Calculations Editor • LOD Expressions





# Tableau - Latest Features (contd...)

## Data Preparation

Connecting to messy spreadsheets with Tableau is fast and easy. Fix data while you connect and skip Excel.

Pivot & Split • Data Interpreter • Metadata Grid • Statistical File Connector • Cloud & Big Data Connections

Country	Program	Year	Pivot field values	
			Pivot	Pivot
≡Abc	≡Abc	Abc	#	
Afghanistan	Child Survival and...	2000		null
Afghanistan	Department of De...	2000		null
Afghanistan	Development Assi...	2000		null
Afghanistan	Economic Support...	2000		null
Afghanistan	Food For Education	2000		null
Afghanistan	Global Health and...	2000		null
Afghanistan	Inactive Programs	2000		null
Afghanistan	Migration and Ref...	2000		11,146,395
Afghanistan	Narcotics Control	2000		null
Afghanistan	Nonproliferation, ...	2000		3,000,000
Afghanistan	Other Active Gran...	2000		2,315,941
Afghanistan	Other Food Aid Pr...	2000		23,632,010
Afghanistan	Other Food Aid Pr...	2000		"

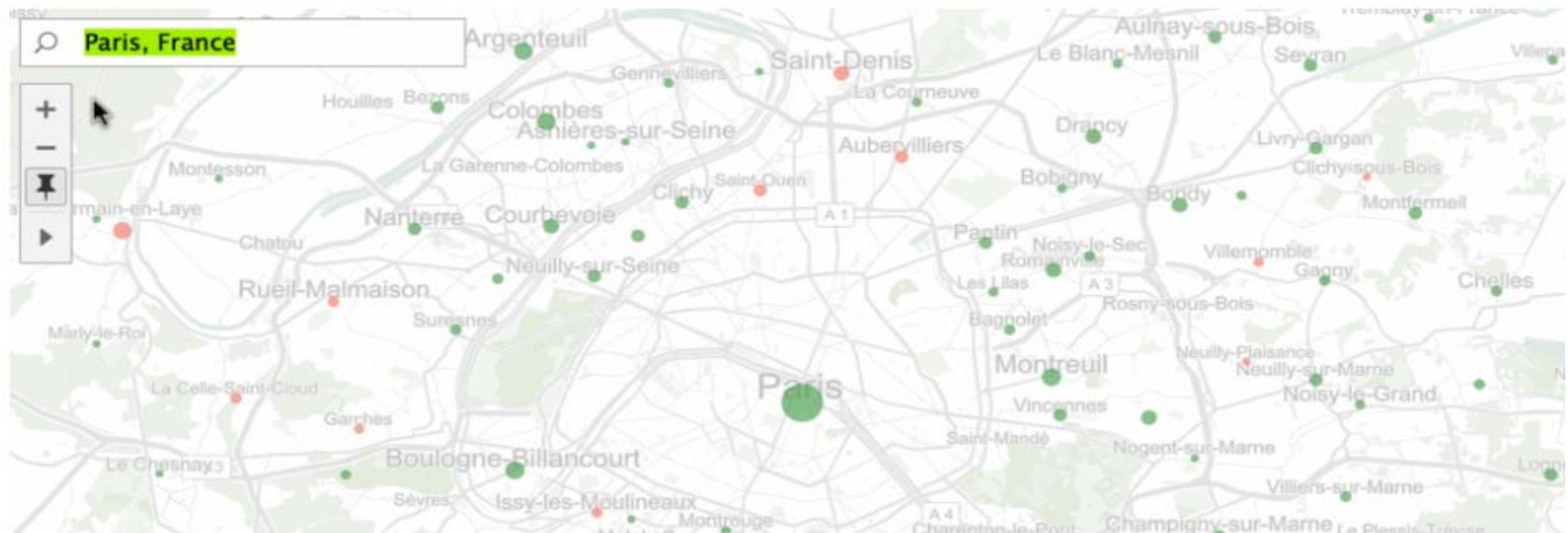


# Tableau - Latest Features (contd...)

## Smart Maps

Answer **geographic questions** more easily. Search in the map or lasso data points. Smoother map interactions mean you stay in the flow while you investigate “where?”

Geographic Search • Fast Navigation & Selection • U.S. Census Data • Expanded Global Data





# Tableau - Latest Features (contd...)

## Faster Performance

Individual **performance improvements in Tableau 9.0** combine to provide unprecedented overall speed increases across workbooks. Caching and consolidation completely eliminate some queries. Other queries run faster on extracts, or process in parallel for quicker load times.

### Consolidated Queries

Big improvements in the Tableau Data Engine lead to significantly faster query results of extracts. Data Engine Vectorization and Parallel Aggregation fully leverage today's multi-core and multi-threaded CPUs.

### Query Fusion

Query Fusion runs a single combined query, replacing multiple individual queries for views that have the same level of detail. This reduces the workload on your database and improves load times.

### Parallel Queries

All independent views have their queries sent to the database simultaneously. These Parallel Queries return data to Tableau faster, meaning your dashboards load faster.

### Reused Queries

Instantly load workbooks and views with cached data from previously returned queries. Tableau Desktop and Tableau Server now avoid re-running the same queries, resulting in near-instant load times.

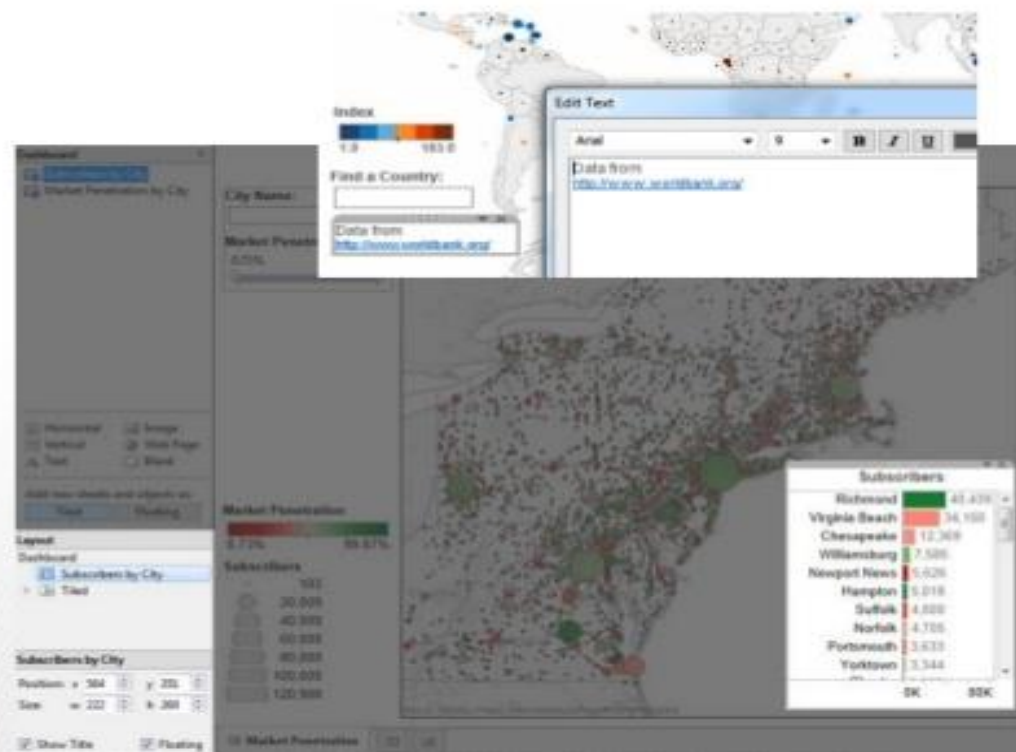


# Tableau - Latest Features (contd...)

## Dashboards

<http://www.tableausoftware.com/new-features/improved-dashboards>

- Pixel-perfect dashboards
- Floating sheets and legends
- Targeted filters
- Hyperlinks in text objects





# Tableau - Latest Features (contd...)

## Any Data

Total access to on-premise and cloud data, wherever it lives

### Connect to Data

#### In a file

- Tableau Data Extract
- Microsoft Access
- Microsoft Excel
- Text File
- Import from Workbook

#### On a server

- Tableau Server
- Amazon Redshift
- Aster Database
- Cloudera Hadoop
- Cloudera Impala
- EMC Greenplum
- Firebird
- Google Analytics
- Google BigQuery
- Hortonworks Hadoop Hive
- HP Vertica
- IBM DB2
- IBM Netezza
- MapR Hadoop Hive
- Microsoft Analysis Services
- Microsoft PowerBI
- Microsoft SQL Server
- MySQL
- Oracle
- Oracle Exadata
- Panorama
- PostgreSQL
- Progress OpenEdge
- Salesforce
- SAP HANA
- SAP HANA Business Warehouse
- SAP Sybase IQ
- Teradata
- Windows Azure Marketplace DataMarket
- Other Databases (ODBC)



- Salesforce.com
- Google Analytics
- Google BigQuery
- Cloudera Impala
- DataStax Enterprise (Cassandra)
- Hortonworks
- Hadapt
- Extract API



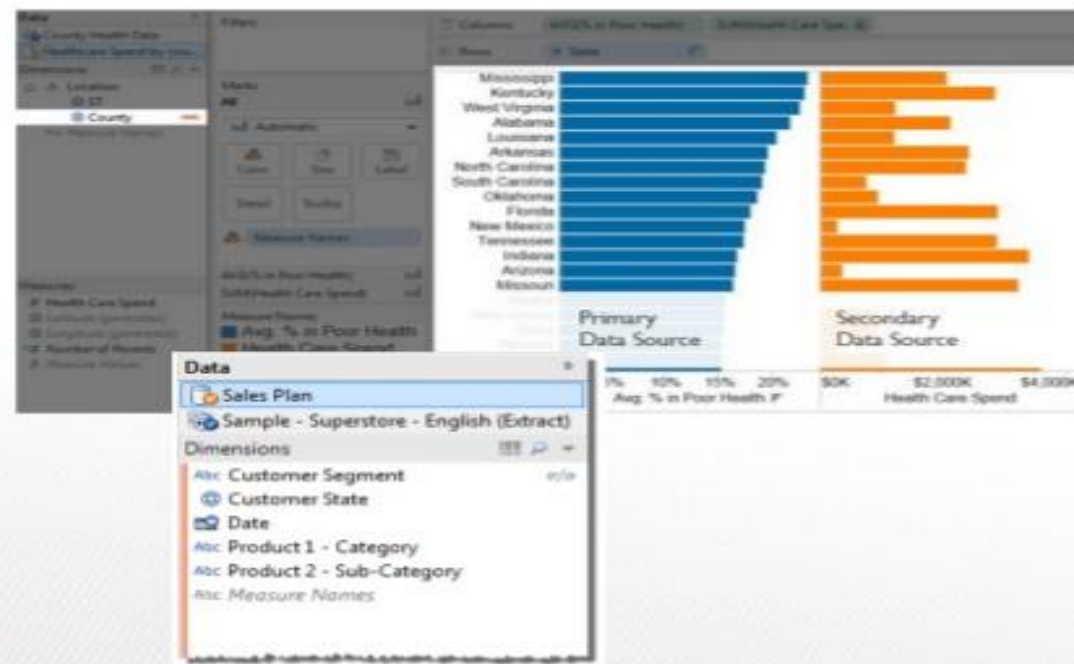


# Tableau - Latest Features (contd...)

## Blend Data

<http://www.tableausoftware.com/new-features/data-blending>

- Easier to use blended data in your analysis
- No longer need the linked field in the view
- Dimensions from secondary data sources can also be filtered





# Tableau - Latest Features (contd...)

## Integrate with Business Applications

<http://www.tableausoftware.com/new-features/javascript-api>

- Embedding Views
- Custom User Interfaces
- Web Site Integration
- Write-back and What-if





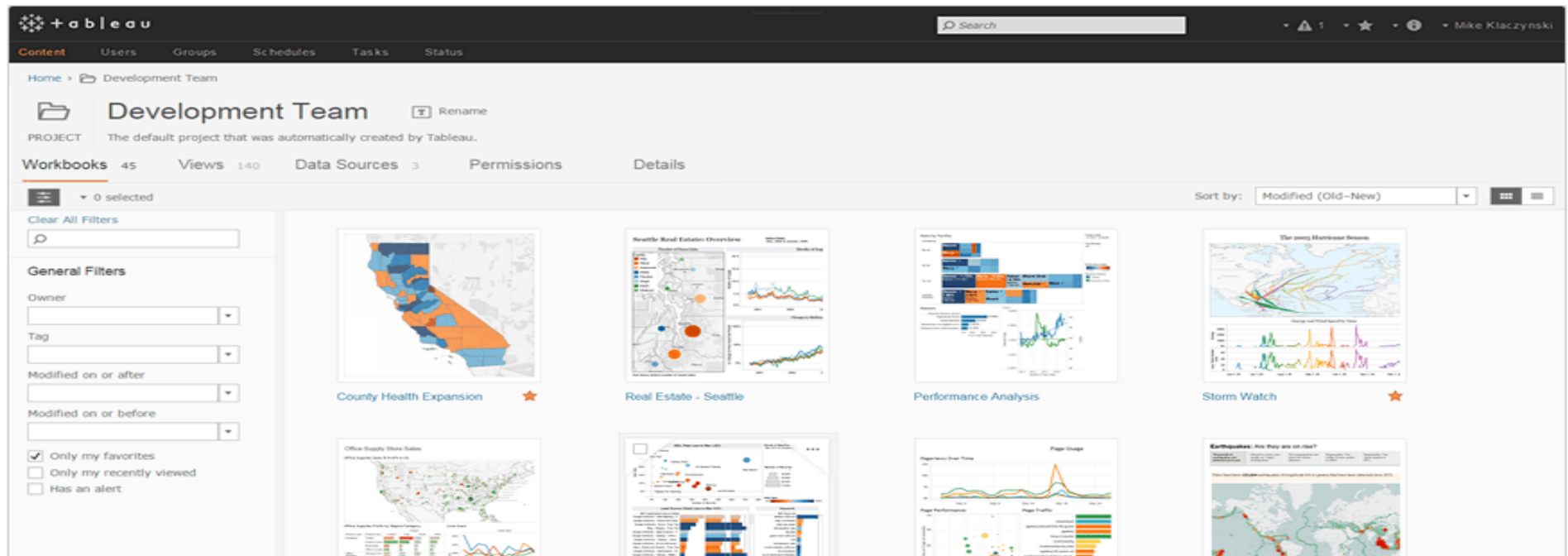


# Tableau - Latest Features (contd...)

## New Server & Online

Tableau Server has been rebuilt from the ground up to be faster, more scalable, and more extensible—all while sporting a new look and more intuitive navigation.

Intuitive Interface • Visual User Permissions • Enterprise Grade • Upgrade Resources



# Dimensions & Measures



## Understanding Data Fields

The data in all data sources are categorized into fields such as Customer, Sales, Profit, Temperature, etc. These fields are made from the columns in your data source. When you connect to a data source with Tableau, the fields are displayed along the left side of the workbook in the Data window. The fields are what you will use to build views of your data. Each field is automatically assigned a data type (such as integer, string, and date) and a pair of data roles.

The Data window organizes fields into three areas:

- **Dimensions** – Fields that typically hold discrete qualitative data. Examples of dimensions include dates, customer names, and customer segments.
- **Measures** – Fields that typically hold numerical data that can be aggregated. Examples of measures include sales, profit, number of employees, temperature, frequency, and pressure.
- **Parameters** – Parameters are dynamic values that can replace constant values in calculations, filters, and reference lines. For example, you may create a calculated field that returns true if Sales is greater than \$500,000 and otherwise return false. You can replace the constant value of “500,000” in the formula with a parameter.

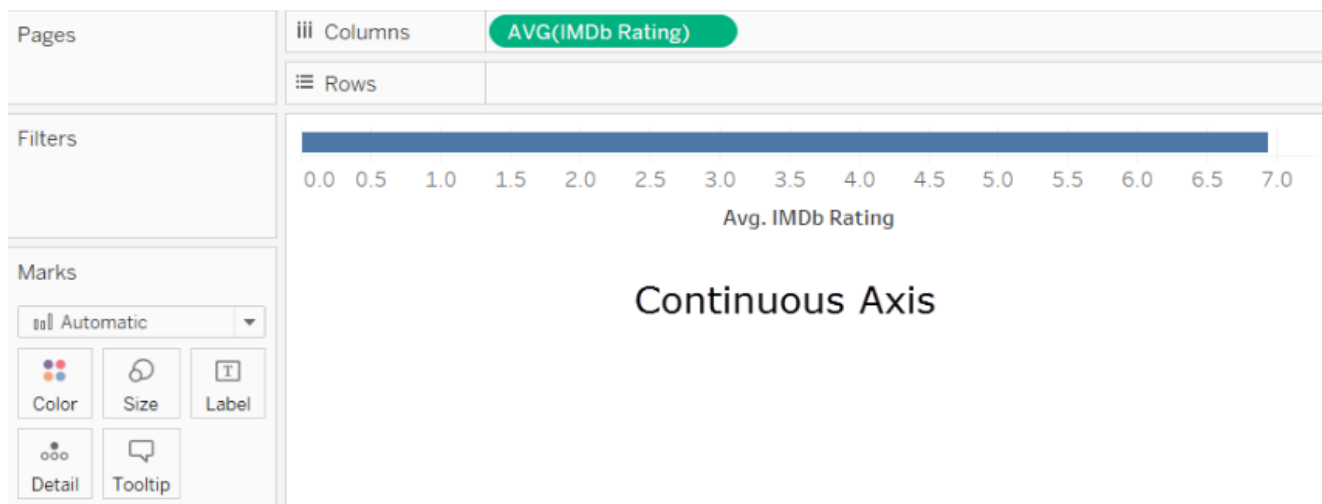


# Discrete And Continuous

## Continuous

In Tableau, fields can be either continuous or discrete. When you drag a field from the **Measures** area to **Columns** or **Rows**, the values are continuous by default and Tableau creates an axis. When you drag a field from the **Dimensions** area of the **Data** pane to **Columns** or **Rows**, the values are discrete by default and Tableau creates column or row headers.

## Continuous fields produce axes





# Contd...

## Discrete fields create headers

If a field contains values that are names, dates, or geographical locations—anything other than numbers—Tableau assigns that field to the **Dimensions** area of the **Data** pane when you first connect to a data source. Tableau treats the values as discrete.

Tableau creates headers when you drop a discrete field on **Columns** or **Rows**. The individual values for a discrete field become the row or column headings. (Because such values are never aggregated, no new field values are created as you work with your view, so there is no need for an axis)

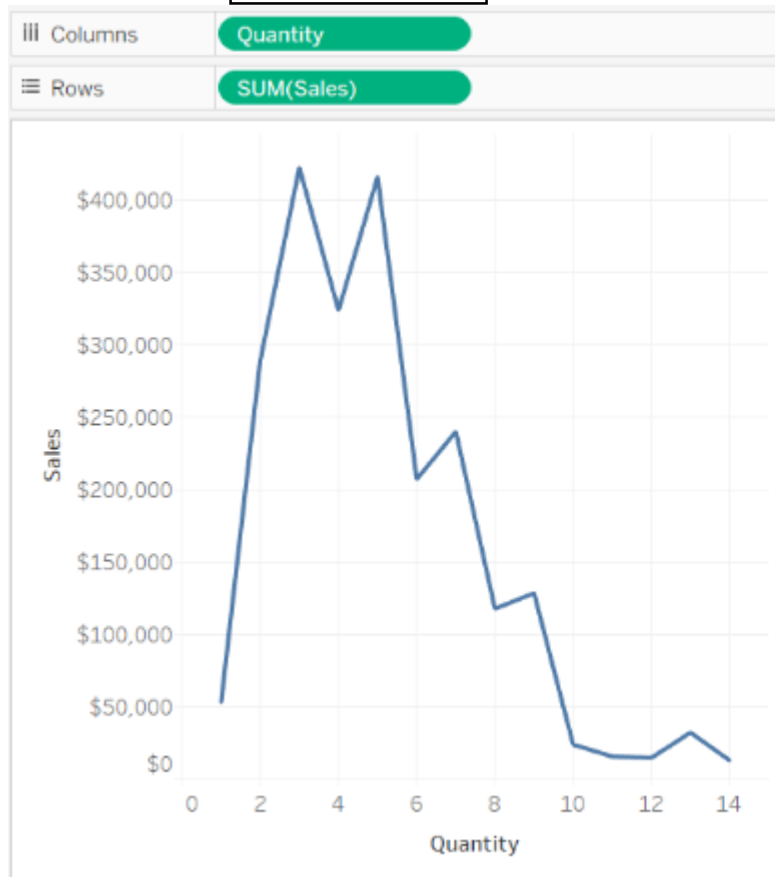
## Recognizing the difference

If a field is continuous, the background color is green. If it is discrete, the background color is blue. Background color does not indicate dimension vs. measure—it indicates continuous vs. discrete

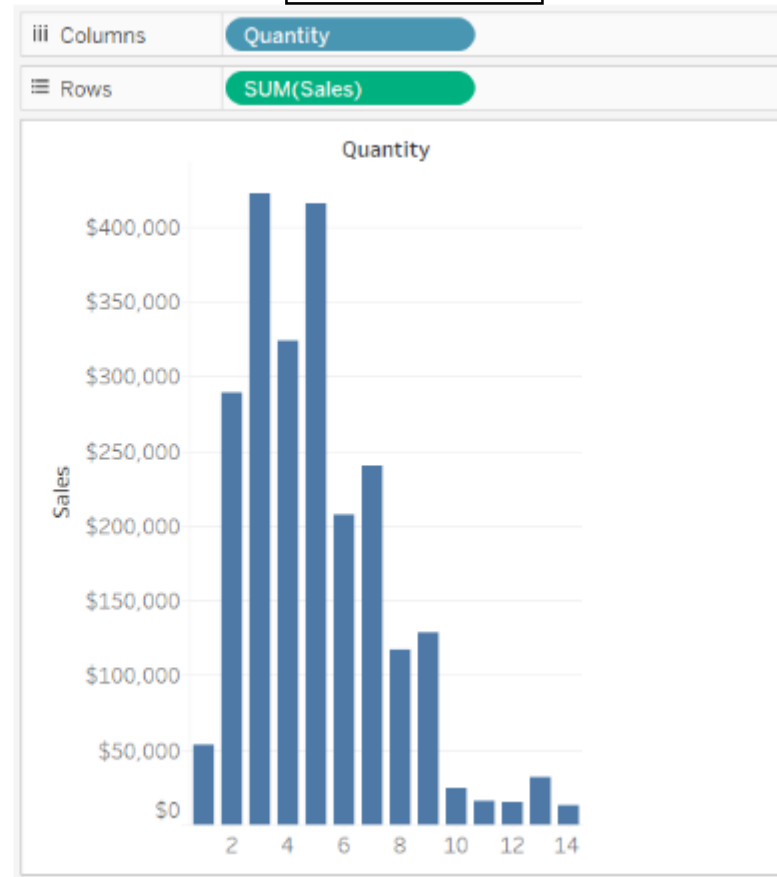
# Contd...



**Continuous**



**Discrete**

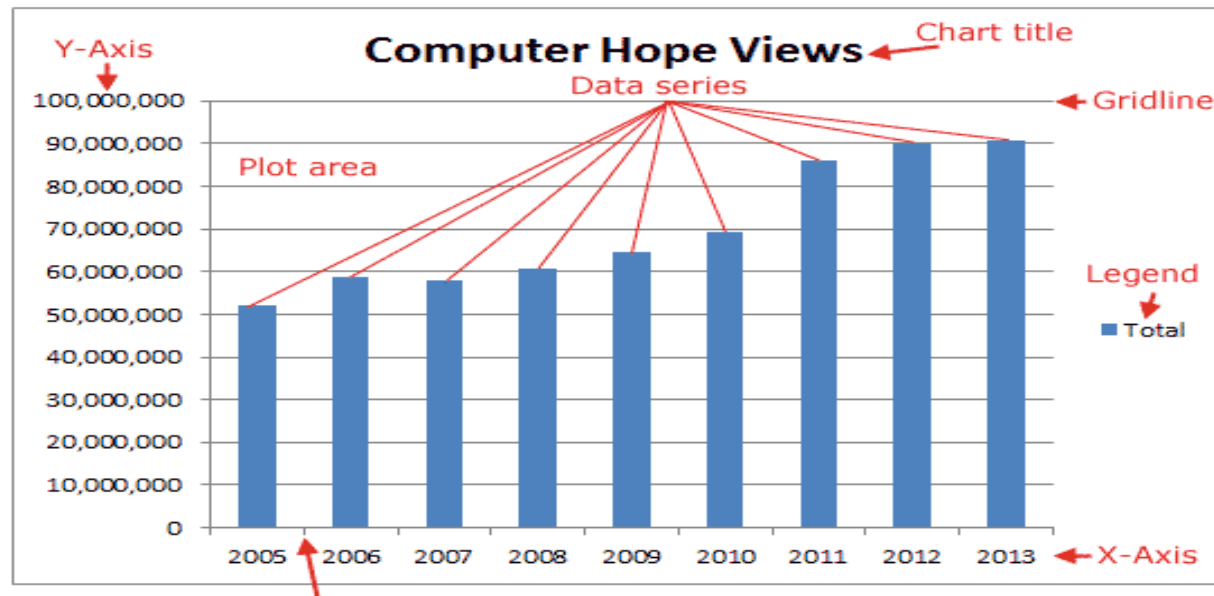




# Chart Overview :

- In general, a **chart** is a graphic representation of data. Charts allow users to see what the results of data to better understand and predict current and future data
- In the picture below, is an example picture of a column chart of unique visitors Computer Hope has received between the years of 2000 and 2006. As can be seen in this example, you can immediately see a gentle increase of users without reading

## Bar Chart Overview



# Extracting Data to Data Engine ( In-Memory )



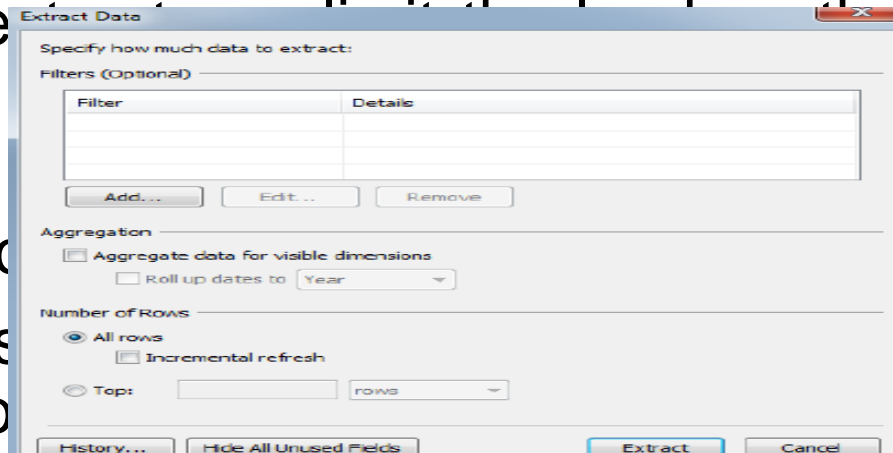
## ➤ Extracting Data to Data Engine

Extracts are saved subsets of a data source that you can use to improve performance, upgrade your data to allow for more advanced capabilities, and analyze offline. You can create an extract by defining filters and limits that include the data you want in the extract.

Extracts can Improve performance. For file based data sources such as Excel or Access, a full extract takes advantage of the Tableau data engine. For large data sources, a filtered extract can be used to reduce the amount of data that is loaded into the server when you only need a subset of the data.

Add functionality to your extract as the need arises. For example, you can add the ability to compute counts, sums, and averages.

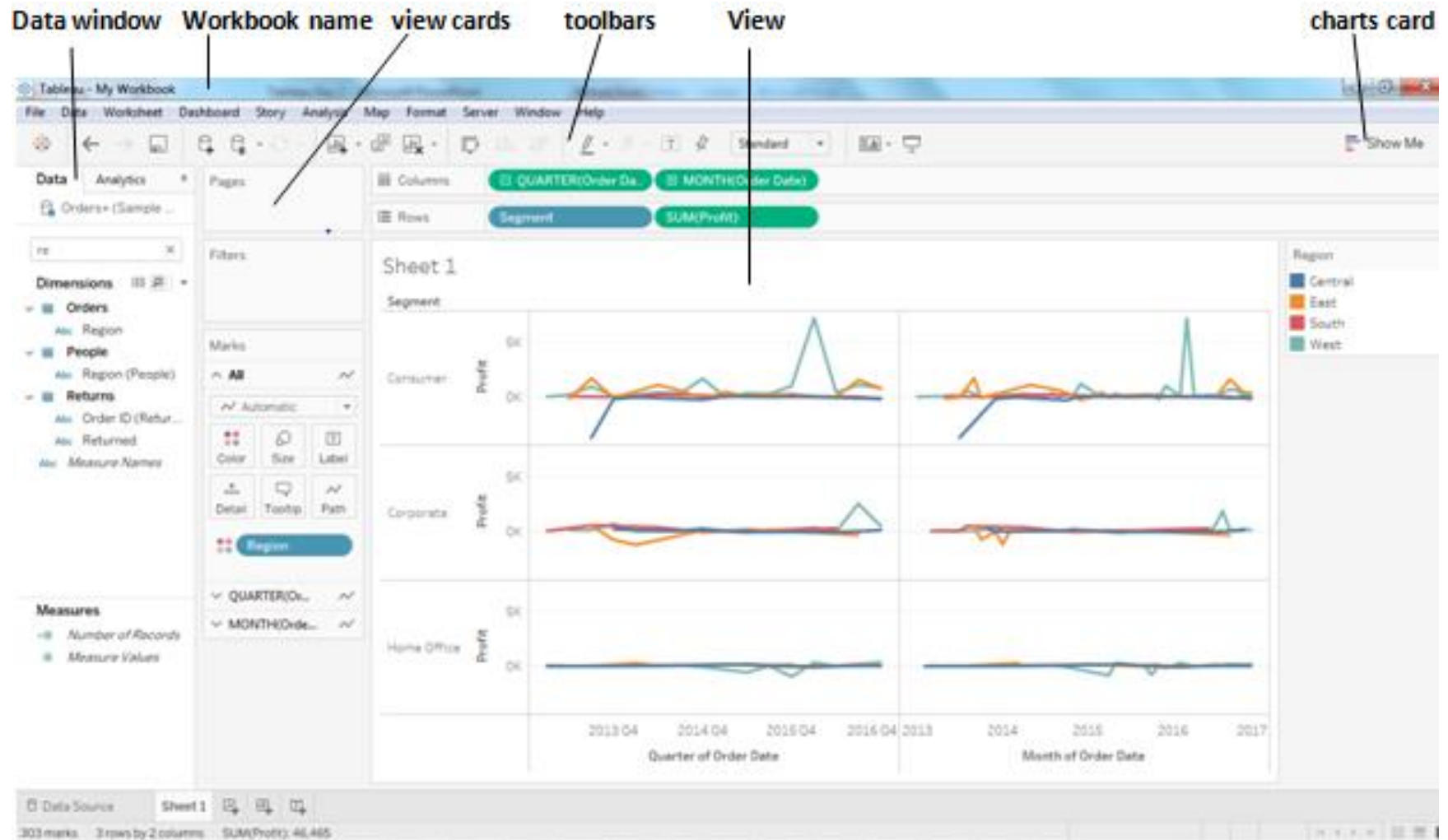
Provide offline access to your data by creating an extract. This allows you to need to access your data without having to connect to the data source. You can also extract the relevant data to a local data source.





# Tableau Workspace

## ➤ Tableau Workspace







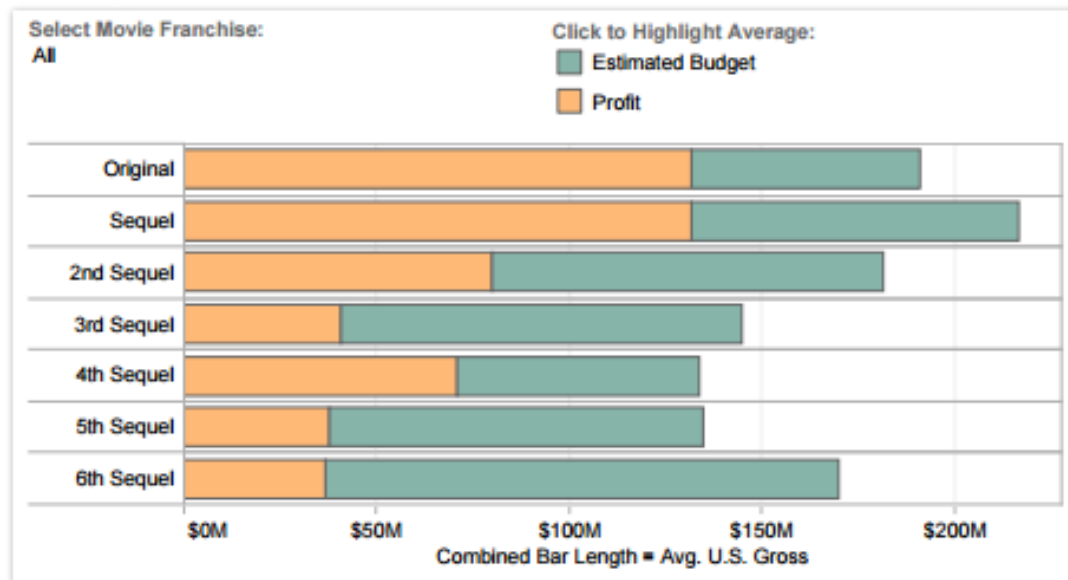
# Charts Types :

- 1. Bar chart
- 2. Line chart
- 3. Pie chart
- 4. Map
- 5. Scatter plot
- 6. Gantt chart
- 7. Bubble chart
- 8. Histogram chart
- 9. Bullet chart
- 10. Heat map
- 11. Highlight table
- 12. Treemap
- 13. Box-and-whisker plot

# Bar Chart



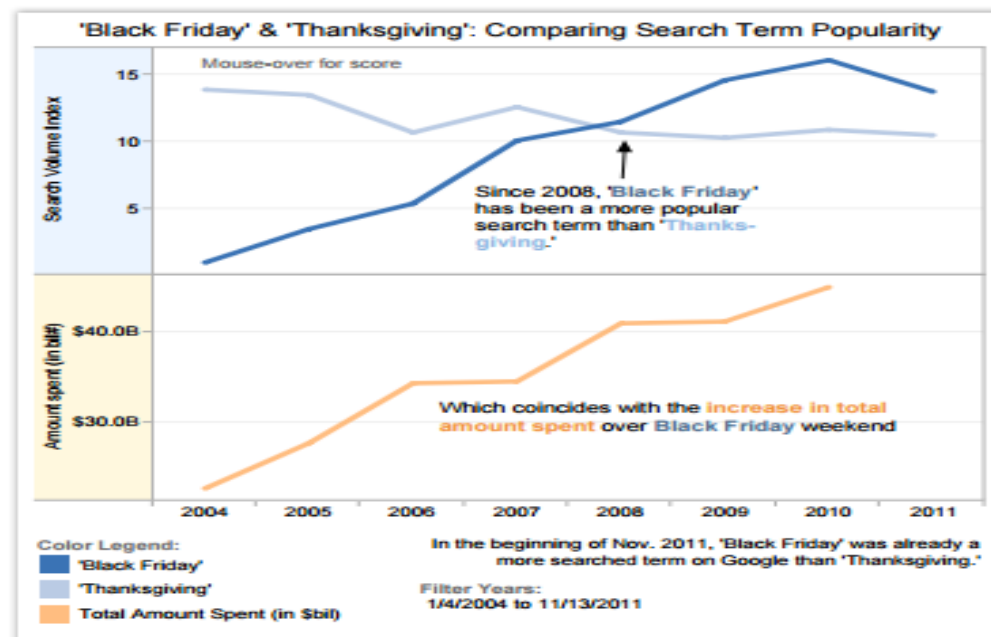
Bar chart Bar charts are one of the most common ways to visualize data. Why? It's quick to compare information, revealing highs and lows at a glance. Bar charts are especially effective when you have numerical data that splits nicely into different categories so you can quickly see trends within your data. When to use bar charts: • Comparing data across categories. Examples: Volume of shirts in different sizes, website traffic by origination site, percent of spending by department.





# Line chart

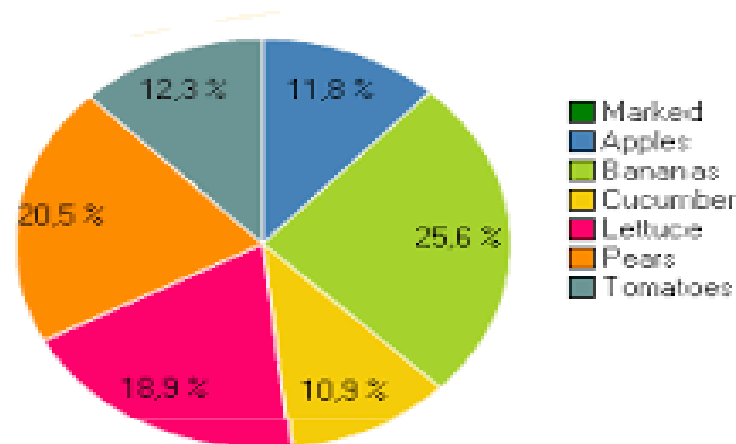
Line charts are right up there with bars and pies as one of the most frequently used chart types. Line charts connect individual numeric data points. The result is a simple, straightforward way to visualize a sequence of values. Their primary use is to display trends over a period of time. When to use line charts: • Viewing trends in data over time. Examples: stock price change over a five year period, website page views during a month, revenue growth by quarter.





# Pie Chart :

Pie chart Pie charts should be used to show relative proportions – or percentages – of information. That's it. Despite this narrow recommendation for when to use pies, they are made with abandon. As a result, they are the most commonly mis-used chart type. If you are trying to compare data, leave it to bars or stacked bars. Don't ask your viewer to translate pie wedges into relevant data or compare one pie to another. Key points from your data will be missed and the viewer has to work too hard. When to use pie charts: • Showing proportions. **Examples:** Percentage of budget spent on different departments, response categories from a survey, breakdown of how Americans spend their leisure time.



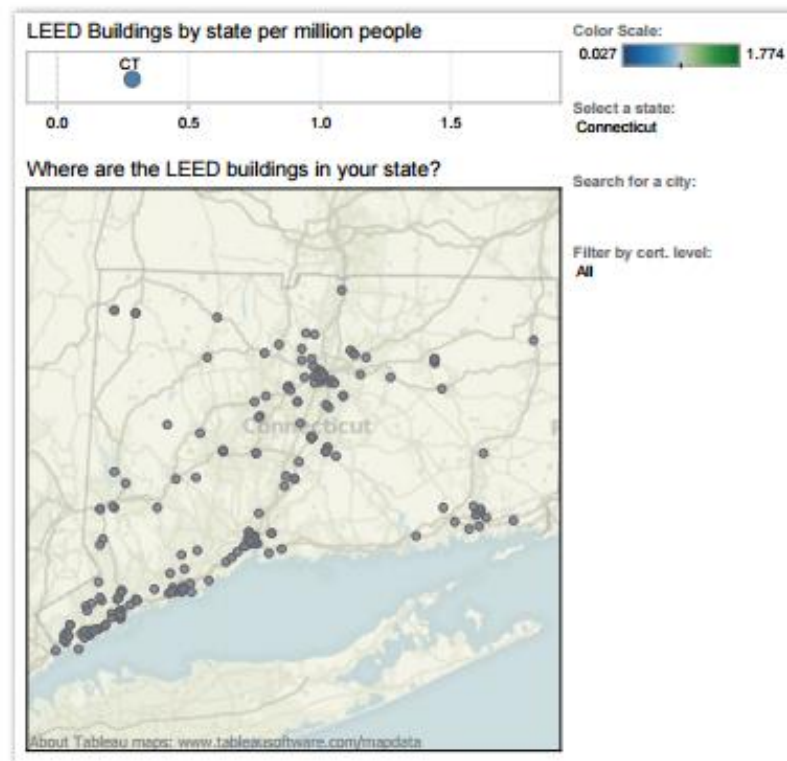
# Map



When you have any kind of location data – whether it's postal codes, state abbreviations, country names, or your own custom geo coding – you've got to see your data on a map. You wouldn't leave home to find a new restaurant without a map (or a GPS anyway), would you? So demand the same informative view from your data. When to use maps:

- Showing geo coded data.

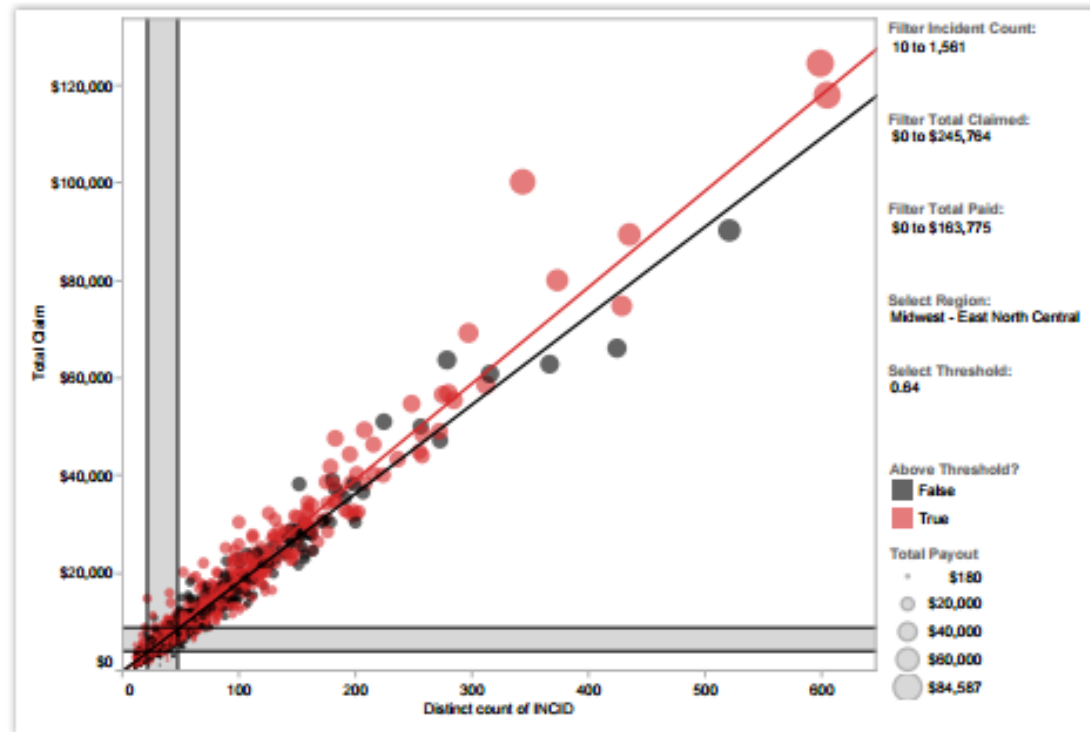
**Examples:** Insurance claims by state, product export destinations by country, car accidents by zip code, custom sales territories.





# Scatter Plot

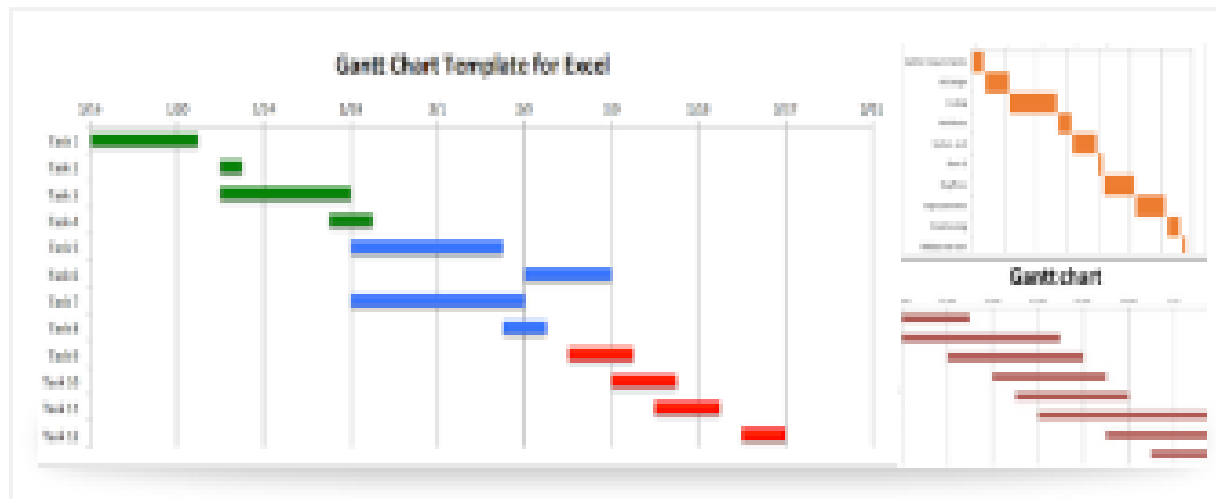
Looking to dig a little deeper into some data, but not quite sure how – or if – different pieces of information relate? Scatter plots are an effective way to give you a sense of trends, concentrations and outliers that will direct you to where you want to focus your investigation efforts.





# Gantt Chart

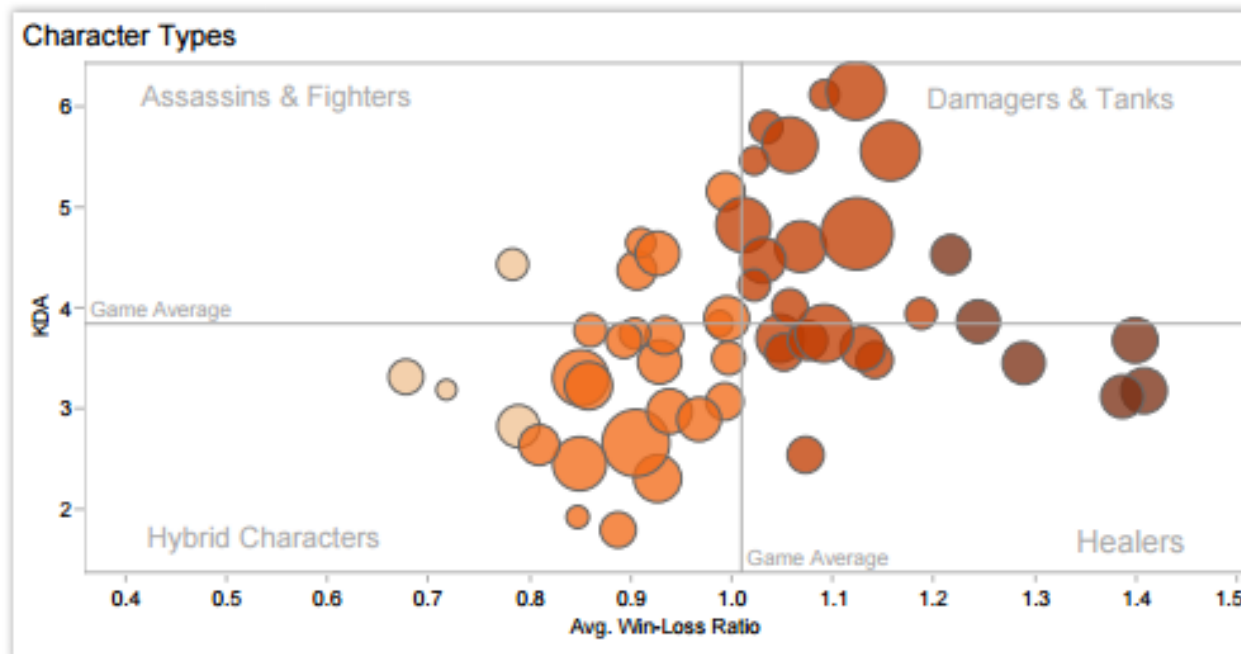
Gantt charts excel at illustrating the start and finish dates elements of a project. Hitting deadlines is paramount to a project's success. Seeing what needs to be accomplished – and by when – is essential to make this happen. This is where a Gantt chart comes in. While most associate Gantt charts with project management, they can be used to understand how other things such as people or machines vary over time. You could use a Gantt, for example, to do resource planning to see how long it took people to hit specific milestones, such as a certification level, and how that was distributed over time.





# Bubble Chart

Bubbles are not their own type of visualization but instead should be viewed as a technique to accentuate data on scatter plots or maps. Bubbles are not their own type of visualization but instead should be viewed as a technique to accentuate data on scatter plots or maps. People are drawn to using bubbles because the varied size of circles provides meaning about the data.



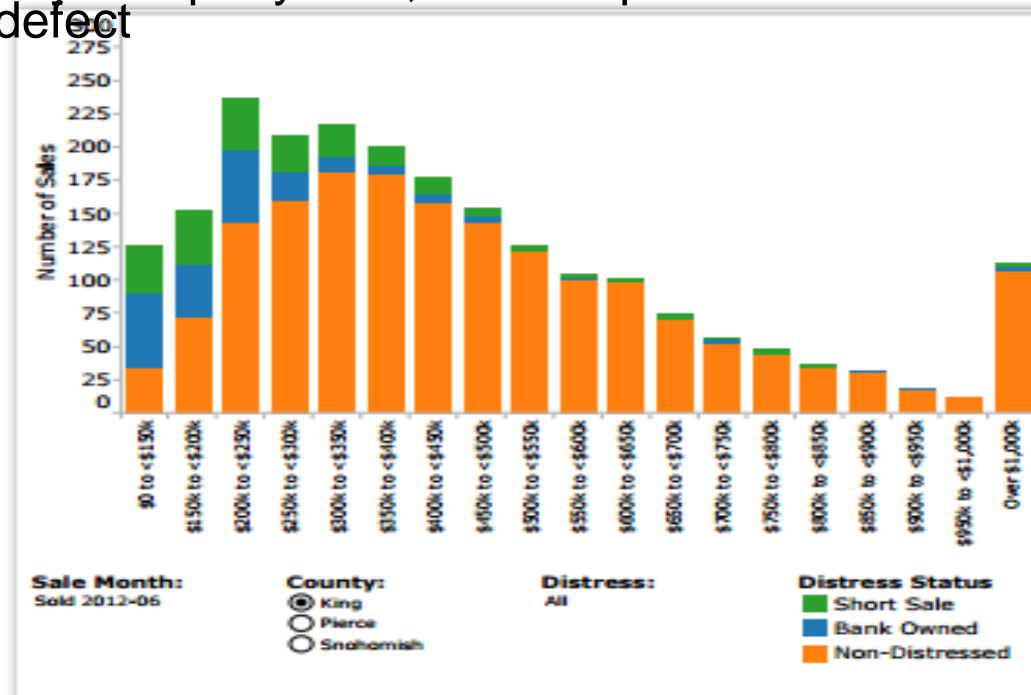




# Histogram Chart

Use histograms when you want to see how your data are distributed across groups. Say, for example, that you've got 100 pumpkins and you want to know how many weigh 2 pounds or less, 3-5 pounds, 6-10 pounds, etc. By grouping your data into these categories then plotting them with vertical bars along an axis, you will see the distribution of your pumpkins according to weight. And, in the process, you've created a histogram. At times you won't necessarily know which categorization approach makes sense for your data. You can use histograms to try different approaches to make sure you create groups that are balanced in size and relevant for your analysis. When to use histograms:

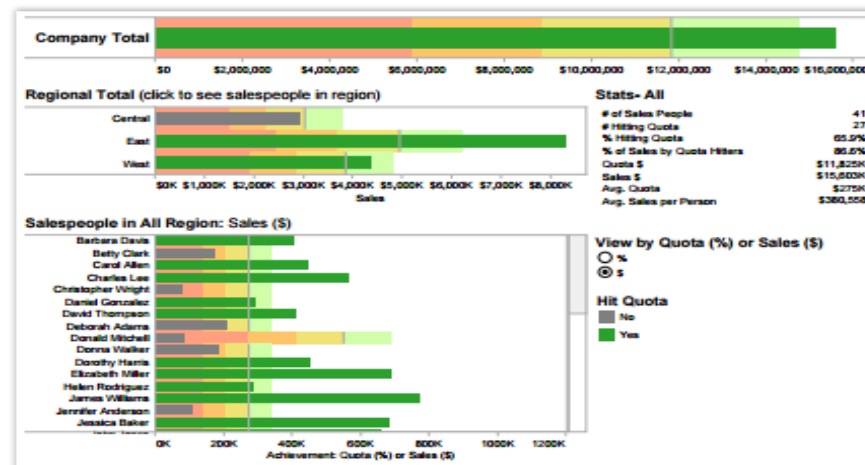
- Understanding the distribution of your data. Examples: Number of customers by company size, student performance on an exam, frequency of a product defect





# Bullet Chart

When you've got a goal and want to track progress against it, bullet charts are for you. At its heart, a bullet graph is a variation of a bar chart. It was designed to replace dashboard gauges, meters and thermometers. Why? Because those images typically don't display sufficient information and require valuable dashboard real estate. Bullet graphs compare a primary measure (let's say, year-to-date revenue) to one or more other measures (such as annual revenue target) and presents this in the context of defined performance metrics (sales quota, for example). Looking at a bullet graph tells you instantly how the primary measure is performing against overall goals (such as how close a sales rep is to achieving her annual quota). When to use bullet graphs: • Evaluating performance of a metric against a goal. Examples: sales quota assessment, actual spending vs. budget, performance spectrum (great/good/poor).





# Heat Map

Heat maps are a great way to compare data across two categories using color. The effect is to quickly see where the intersection of the categories is strongest and weakest. When to use heat maps:

- Showing the relationship between two factors. Examples: segmentation analysis of target market, product adoption across regions, sales leads by individual rep.
- Also consider:
- Vary the size of squares. By adding a size variation for your squares, heat maps let you know the concentration of two intersecting factors, but add a third element. For example, a heat map could reveal a survey respondent's sports activity preference and the frequency with which they attend the event based on color, and the size of the square could reflect the number of respondents in that category.

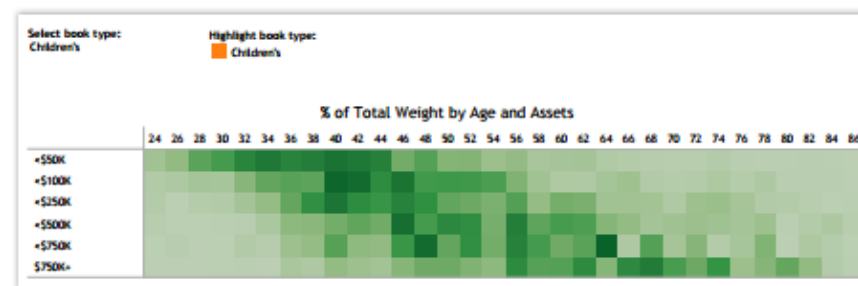


Figure 16: Who buys the most books?

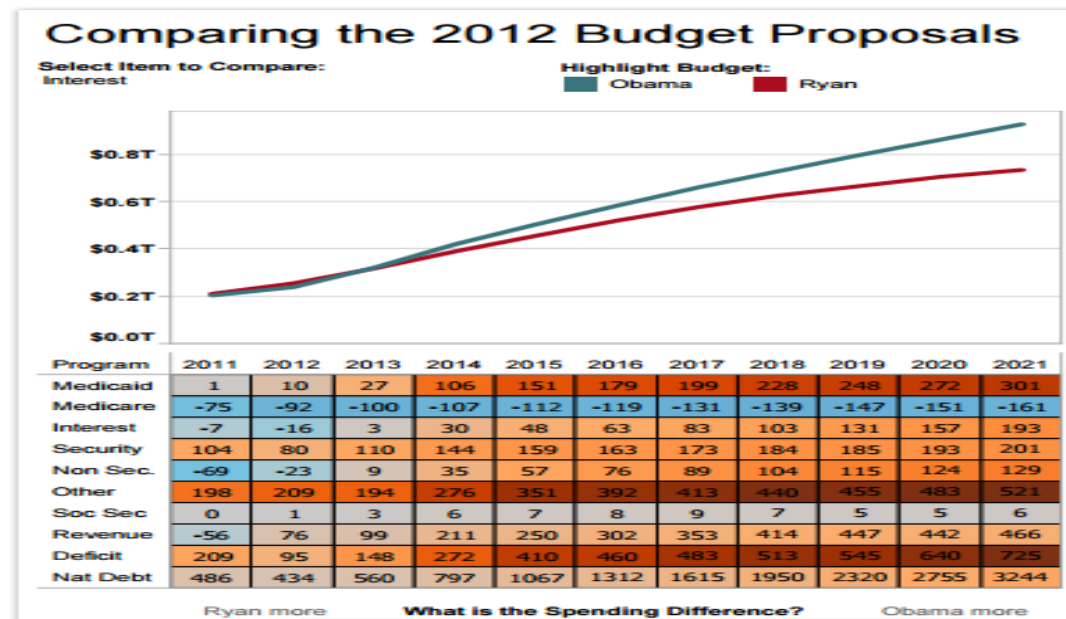
In this **market segmentation** analysis, the heat map reveals a new campaign idea. High-income households of people in their sixties buy children's books. Perhaps it's time for a new grandparent-oriented campaign?



# Highlight table

Highlight tables take heat maps one step further. In addition to showing how data intersects by using color, highlight tables add a number on top to provide additional detail. When to use highlight tables:

- Providing detailed information on heat maps. Examples: the percent of a market for different segments, sales numbers by a reps in a particular region, population of cities in different years.
- Also consider:
- Combine highlight tables with other chart types: Combining a line chart with a highlight table, for example, lets a viewer understand overall trends as well as quickly drill down into a specific cross section of data.





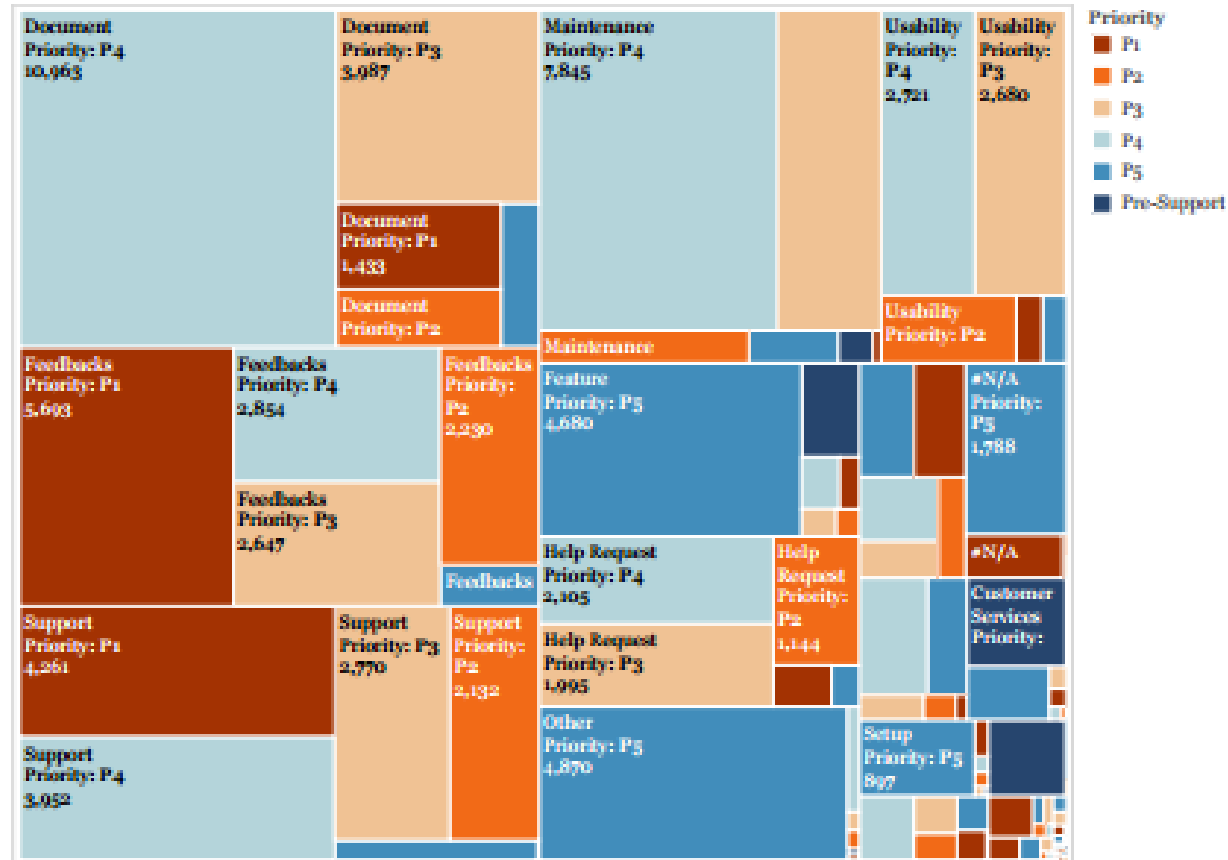
# Tree map

- Looking to see your data at a glance and discover how the different pieces relate to the whole? Then tree maps are for you. These charts use a series of rectangles, nested within other rectangles, to show hierarchical data as a proportion to the whole.
- As the name of the chart suggests, think of your data as related like a tree: each branch is given a rectangle which represents how much data it comprises. Each rectangle is then sub-divided into smaller rectangles, or sub-branches, again based on its proportion to the whole. Through each rectangle's size and color, you can often see patterns across parts of your data, such as whether a particular item is relevant, even across categories. They also make efficient use of space, allowing you to see your entire data set at once.
- When to use tree maps:
  - Showing hierarchical data as a proportion of a whole: Examples: storage usage across computer machines, managing the number and priority of technical support cases, comparing fiscal budgets between years

# Contd...



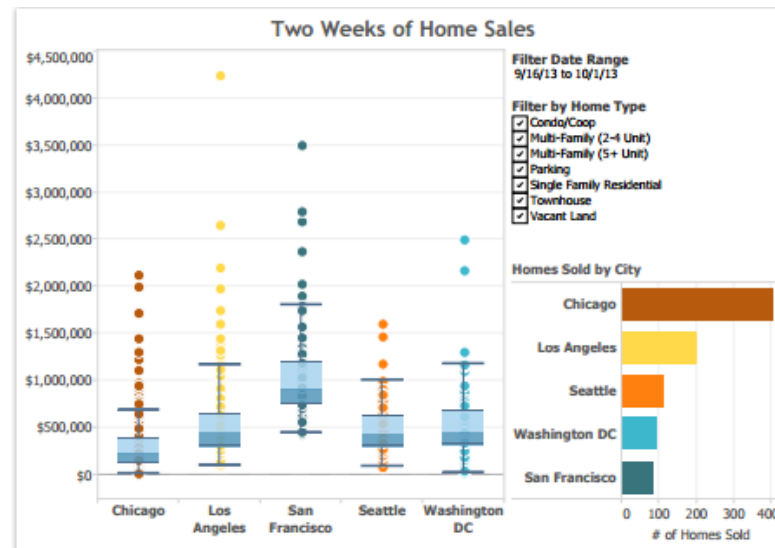
## Support Case Overview





# Box-and-whisker Plot

- **Box-and-whisker Plot** Box-and-whisker plots, or box plots, are an important way to show distributions of data. The name refers to the two parts of the plot: the box, which contains the median of the data along with the 1st and 3rd quartiles (25% greater and less than the median), and the whiskers, which typically represents data within 1.5 times the Inter-quartile Range (the difference between the 1st and 3rd quartiles). The whiskers can also be used to also show the maximum and minimum points within the data. When to use box-and-whisker plots:
- Showing the distribution of a set of a data: Examples: understanding your data at a glance, seeing how data is skewed towards one end, identifying outliers in your data.





# Q&A