TABLEAU



DATA ANALYSIS CAN BE BORING

- Searching for/obtaining data
- Cleaning the data, you must make it "workable"
- Analyzing & finding your answers
- Trying to explain to people who don't understand what you've found

YOU HAVE TO MAKE IT INTERESTING!

You WANT to be able to tell your STORY!

WHAT IS TABLEAU?

A very powerful data visualization software that helps people view and understand data.

- It makes it easier for people to explore and manage their data.
- It can be a central location for data from several sources
- It will change raw data into data that is easy to understand
- It is growing really fast as a key data analytics platform.
- There's a HUGE community of Tableau users
- No coding necessary

POSSIBLE JOB TITLES

- Business Analyst
- BI Engineer
- Data Analyst
- Bl Analyst
- BI Developer

- Verizon
- Lenovo
- Chipotle
- REI
- Charles Schwab

AS A DATA ANALYST

• You are required to manipulate the data in a tabular and systematic form to help solve some challenges that your company might be facing.

 Business analytics, the importance of data visualization and how to utilize Tableau.

BUSINESS INTELLIGENCE

The right information to the right people at the right time to aid in better decision making.

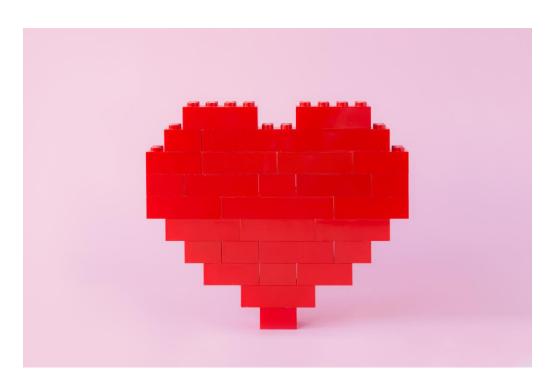
- Business Analytics
- Data mining
- Data visualization
- Data tools
- Infrastructure

EXAMPLES

- Identify ways to increase profit
- Track performance
- Analyze customer behavior
- Compare data with competitors

- Discover issues or problems
- Optimize operations
- Recognize trends
- Predict range of successes

DATA VISUALIZATION



- Uses pictures to represent data
- Most effective way to communicate
- LOTS of tools available
- Allows for quantitative data to be viewed easier & better decisions made quickly

01 8006.52 JPY C +9192.42 +06.85 57 9072.84 AUD F +1437.42 +05.55 .05 8169.19 CHF H +3192.07 +04.55 .61 2591.78 CAD C +6205.12 +07 .60 9217.67 EUR F +5083.11 +05 3.29 7805.51 GBP S +8595.32 +05 3.29 7805.51 GBP S +8595.32 +05 51 86 2244.57 CHF X +9547.24 +05 51 86 2244.57 CHF X +9547.24 +05



WITH TABLEAU

- Can utilize data in various forms (except for Tableau Public)
- No technical or programming experience required
- You can data blend (connect multiple sources for one analysis using joins)
- Analyze in real-time
- Collaborate with other team members

Advantages

- Scripting language
- Data Visualization
- Quick Insights
- Large amounts of data
- Easy Implementation
- Responsive Dashboard

Disadvantages

- Restrictive Imports
- Custom Formatting
- Static Parameter
- Screen Resolution
- Scheduling Reports

Tableau Public

Can load from 9 different data sources

Tableau Desktop

Can load from 77 different data sources

OPENING TABLEAU PUBLIC

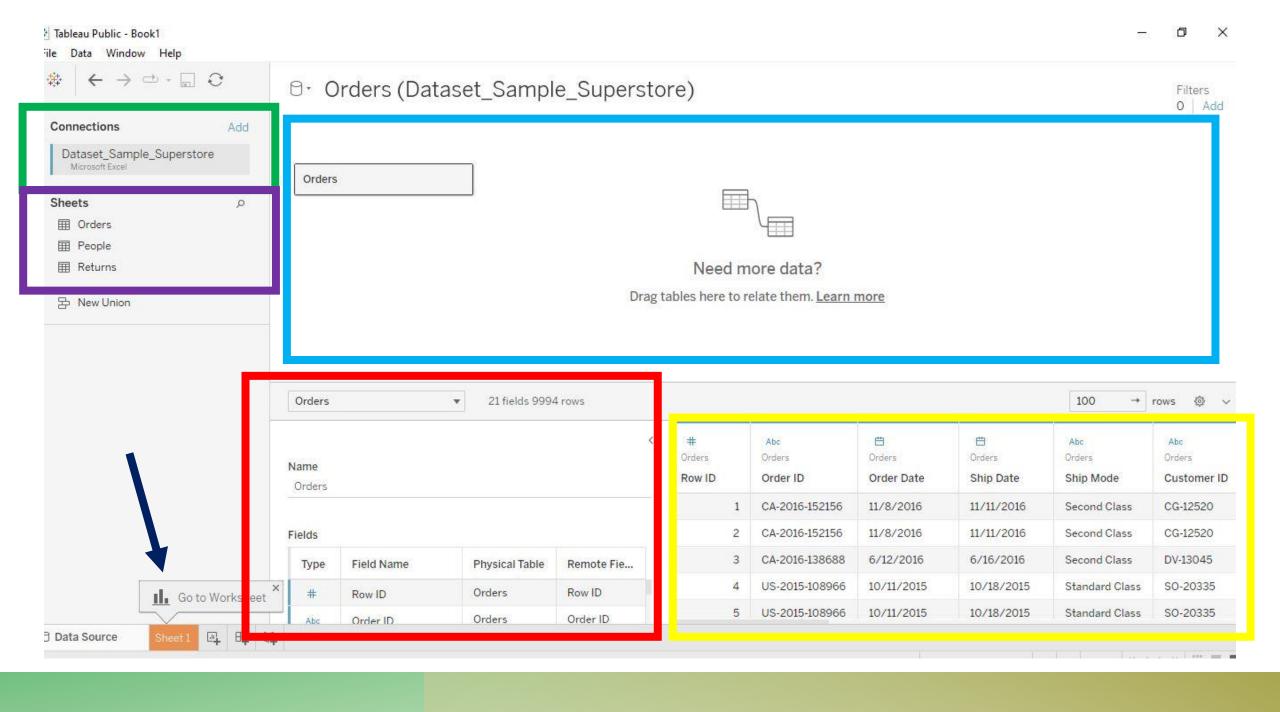
Connect Pane: where various data sources can be connected

 Open Pane: where you can see recently opened files/sample dashboards

• Discover Pane: helps you connect with the Tableau Community and access videos and blogs.

CLICK ON MICROSOFT EXCEL OPEN SAMPLE - SUPERSTORE

CLICK/DRAG "ORDERS" OVER TO THE WHITE AREA ON THE RIGHT



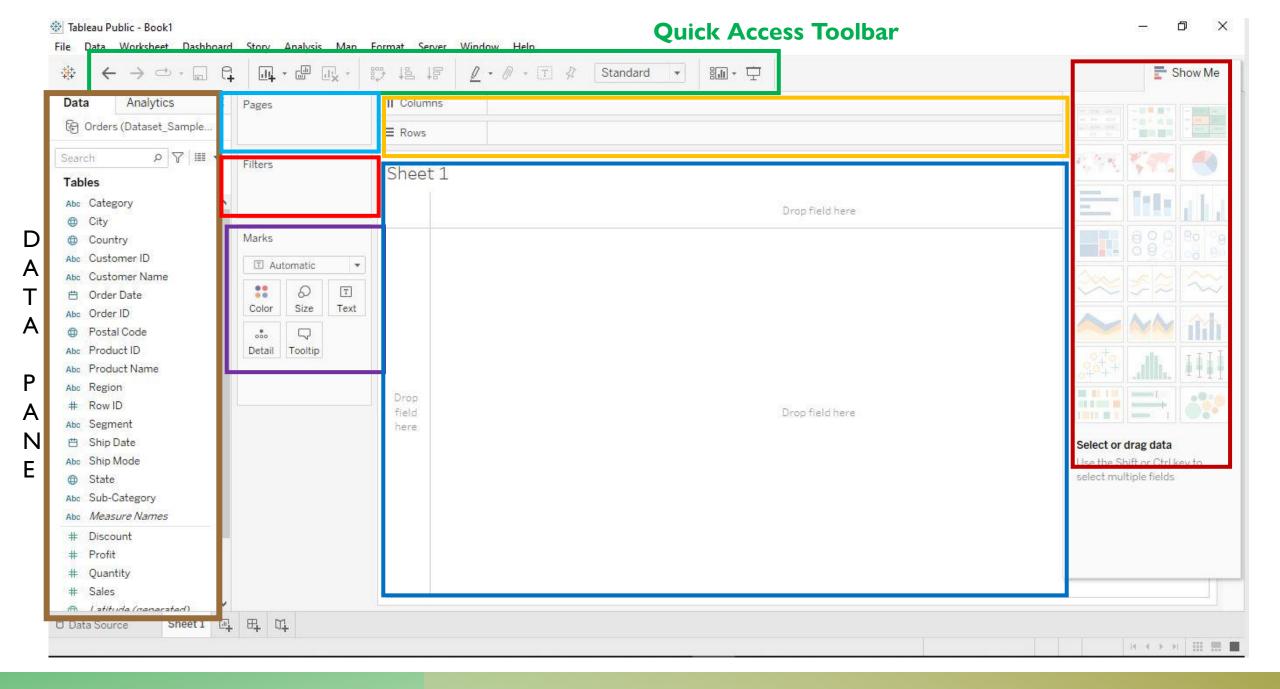


TABLEAU BREAKDOWN

- Each graph should require a **NEW** worksheet & make sure they are named properly
- Dashboards are preps for your STORY (aka presentation)
- Think of each Dashboard as a "slide" where you put your STORY together. Make sure to name them properly
- Your STORY is your presentation of how you want to show your Dashboards (aka slides)

LET'S PRACTICE

- Plot SALES against CATEGORY, SEGMENT & ORDER DATE
- Make it a line graph
- Filter down to Home Office for each segment
- Delete top title
- Make it a bar graph
- Rename it "Home Office Sales"

8 TOOLS TO PREPARE DATA FOR ANALYSIS

- 1. Joins
- 2. Relationships
- 3. Unions
- 4. Data Blending

- 5. Aggregations
- 6. Pivot
- 7. Data Interpreter
- 8. Split

JOINS

Combine data in order to filter and get rid of duplicates

- Left Join
- Right Join
- Inner Join
- Full Outer Join

- Connect to the Sample-Superstore dataset
- Move "Orders" over to the canvas
- Right-click on "Orders" and click Open
- Move "Returns" over to the canvas
 - This creates an inner join BY DEFAULT
- Select the correct join

- Tableau will identify the primary key for joining two or more data sets together
- Left Join: values from LEFT Table & match to RIGHT Table
- Right Join: values from RIGHT Table & match to LEFT Table
- Inner Join: includes common data present in BOTH datasets
- Full Outer Join: includes ALL values from both tables.

RELATIONSHIPS

- Created when you drag more than one sheet over
- Think of it as a contract between two tables
- No joining
- Automatic & context aware
- Flexible

JOINS VERSUS RELATIONSHIPS

UNIONS

Helps bring together data from multiple small files into one large one.

- Connect to dataset
- Drag over needed sheet
- Drag over another to overlap (you'll see "union")
- Overlapping icon

DATA BLENDING

• Tableau combines related data from multiple sources in a single view

- Connect two sources (Superstore & Global Superstore)
- Click on Data table & "Edit Blend Relationships"
- Choose "Segment" and "OK"
- Drag "Profit" from both sets over to Columns
- Drag "Segment" from Global Superstore over to Rows

BLEND

- Used to combine data from different databases that don't support joins
- Data w/in databases are at different levels of detail
- Joins = duplicate rows
- Involves large amount of data

JOIN

- Data format is consistent across ALL sources
- Involves small amounts of data
- Data is at the same level of detail across sources

AGGREGATIONS

- Process of converting a set of values into a single value.
- Done based on Measures & Dimensions

- Aggregating Measures can be added to relational data sources
- Aggregating Dimensions can be done by Minimum, Maximum,
 Count or Count(Distinct)

Let's use the Sample Superstore Source

- Drag "Sales" to Rows, Columns, or Text (inside Marks box)
- Default is SUM
- Right-click on SUM(Sales)
- Choose Measure(SUM) & select desired aggregation from list (avg)

 You can change the aggregation of any measure by clicking on a measure from the left-hand side, select Default Properties and Aggregation & select one

PIVOT

- Select columns that need to be manipulated & formatting them into a typical dataset
- Need to REMEMBER to pivot from Columns to Rows (tall v. wide)

- Data not formatted exactly in preview pane:
 - Select all columns in the file
 - Select pivot from pop-up menu

DATA INTERPRETER

- Automatically cleans and prepares data
- Available under Connections Pane whenever a unique format is detected.

SPLIT

- String function used to split a column into multiple ones
- Used to make data analysis easier
- Offers both Automatic Split and Custom Split
 - Automatic: done by detecting a common separator
 - Custom: max of 10 new fields based on a separator in original field
 - Can split columns by clicking on column in preview pane and selecting
- Example: Moving "City, State: Denver, CO" to two different cells

DATA TYPES INSIDE TABLEAU

- Text Values
- Date Values
- Date and time Values

- Numerical Values
- Boolean Values
- Geographic Values

DATA ROLES

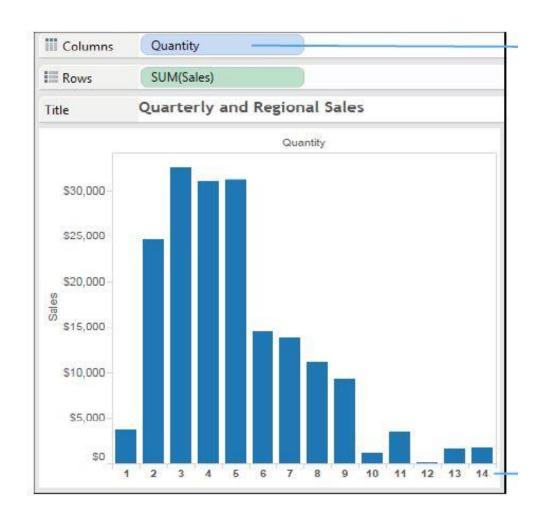
Dimensions v. Measures

- Data is both textual & numerical, automatically segregated into either role
- Dimensions: produce a header when added to row/column, categorical
- Measures: produce an axes when added to row/column, numeric

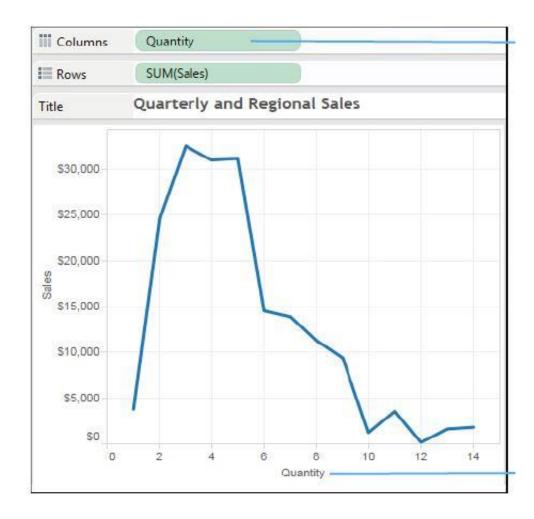
Discrete v. Continuous

• Every field is categorized as either discrete or continuous

Discrete: every quantity as a header at the bottom



Continuous: quantity values are shown in a continuous axis



WORKING W/ METADATA

- Rename: sheet, column names
- Hide: columns you don't want to use
- Unhide
- Sort

WORKING WITH DEFAULT FIELDS

Adding a comment to a field

- Right-click on field
- Select Default Properties
- Select Comment

 You can view the comment when you put your cursor over the field Changing a decimal number to a whole number

- Right-click on field
- Select Default Properties
- Select Number Format
- Select Number (Customer)
- Set Decimal Place to Zero

WORKING WITH DEFAULT FIELDS

Change the Fiscal Year!

- Right click on field
- Select Default Properties
- Select Fiscal Year Start
- Select Month needed

Changing a decimal number to a whole number

- Right-click on field
- Select Default Properties
- Select Number Format
- Select Number (Customer)
- Set Decimal Place to Zero

YOU CAN CREATE ALIASES!

- Right-click on any column and select Aliases
- Update the alias name and click okay
 - Second Class would be where you would rename

YOU CAN FILTER YOUR DATA

DIMENSION FILTER

- General: selects one or multiple members from the domain
- Wildcard: Matches values based on the presence of a string, the same starting and ending characters, or the same set of characters.
- Condition: Matches values based on field or formula
- Top: Highlights top values by field or formula

PRACTICE

- Connect the SuperStore dataset
- Drag "Orders" over to
 Canvas & go to Worksheet
- Put SUM(Sales) in Columns
- Put Sub-Category in Rows
- Drag Sales to Label

- Click on Entire View
- Click on the sort option
- Drag Sub-Category to Filters area
- Create filter
- You can remove filter by dragging Sub-Category outside of the filters area

MEASURE FILTER

- Range of values: filter to identify sales representatives within a range of sales totals
- At least: filter to identify sales reps with sales ABOVE a certain level
- At most: filter to identify sales reps with sales BELOW a certain level
- Special: filter to identify all values, null values and non-null values

PRACTICE

- Connect the SuperStore dataset
- Drag "Orders" over to Canvas & go to Worksheet
- Put Sub-Category in Rows
- Put Sales in Text
- Resize for visibility
- Sort in descending order

- Drag Sales to filters
- Click next
- Set your filter

 You can remove filter by dragging Sales outside of the filters area

DATE FILTER

- Relative Dates: can specify range of dates that is updated based on the date and time.
- Range of Dates: can specify a range of dates to define a fixed range of dates to filter.
- Discrete Dates: can specify a discrete date value in the dialog box
- Individual Dates: can select individual dates to filter specific dates from the view
- Additional date filter Options: can select null dates, non-null dates or all dates by specifying a start/end date.

PRACTICE

- Connect Sample Superstore
- Drag orders to canvas
- Add Order Date to Rows
- Add Sales to Text
- Expand Order Date to Quarter and Month

- Restructure it by adding
 Year to Column
- Drag Order Date to Filters
- Click on Range of Dates
- Click on Individual Dates

VISUAL FILTER

Filters that are applied straight from Tables, Maps and Charts

PRACTICE

- Connect SampleSuperstore
- Drag orders to canvas
- Sub-Category > Rows
- Sales > Columns & Text
- Sort in descending order

Select Sub-Category & apply inclusion filter.

CONTEXT FILTER

- The ONLY filter that is independent & used for:
 - Improving the performance of Tableau reports
 - Creating a dependent numerical filter
- Any other filter defined alongside this one is a dependent one:
 - It's dependent on the output of the context filter

PRACTICE

- Connect Sample Superstore
- Drag Orders to canvas
- Drag Sub-Category to Rows
- Drag Sales to Columns
- Click on Entire View and sort in descending order

- Add Sales to Label
- Drag Category to Filters
 - Click on Furniture
- Drag Sub-Category to Filters
 - Apply the filter

DATA SOURCE FILTER

- Applied to data sources that are connected to Tableau to make the worksheet work more efficiently.
- The changes are done across ALL worksheets
- Considered an efficient way to limit scope of data for performance/security purposes.

- Extract filters are the same as data source filters, just applied to the extract connection.
- Worksheet filters

Field Types

- Dimensions (Region or product name)
- Measures (Sales or profit)
- Dates (Year or month)

PRACTICE

- Connect Sample Superstore
- Drag orders to canvas
- Add Regions to Rows
- Add Sales to Text
- Drag Region to Color
- Sub-Category to Rows
- Sales to Text & Columns

- Click on Data Source & Add
- Select add in the box
- Click on City & OK & OK again

 Can View entire worksheet filter in the top right corner.

INTERACTIVE FILTER

- Allows users to interact with the data via Drop-down, List,
 Radio Button or Checkbox
- Can be applied to: numbers, dates and strong

PRACTICE

- Connect SampleSuperstore
- Drag orders to canvas
- Drag Sub-Category to Rows
- Sales to Columns & Label
- Sort in Descending order

 Select Sub-Category and click on show filter

LET'S DO SOME SORTING

- Based on data course order, field, ABC order or manual
- On a measure, can be applied through a dimensional sort

GROUPS

- Aggregate the data of dimension members
- Example: top 5 sales reps by total salary are in a single row
- Once grouped, can be used inside of any filters

SETS

- Custom fields that define a subset of data based on some conditions
- Example: customers with sales over a certain level
- CONSTANT sets: do not change after they are created
- COMPUTED sets: types of sets that change with the change in data
- COMBINED sets: two sets combined to compare different members
- BUILDING sets: created through manual selection, condition, ranking

BINS

- Equal-sized containers that store data values that correspond to or fit within the bin size.
- Can use any discrete field to create them
- Right click on field, select Create and then Bins
 - Create new field name & designate size
 - Press OK

HIERARCHIES

- An arrangement where entities are present at various levels
- Created by adding one dimension as a level under the principal one
- Example: Category & Sub-Category in Rows and SUM(Sales) in Text
- Can be used as dimensions
 - Right-click, click Hierarchy > Create Hierarchy > Add Hierarchy
 - > Add dimensions

CREATING CHARTS

- Easier to understand then tables with numbers
- Click on "Show Me" button on upper right corner

Working with Sample-Superstore dataset

Orders sheet on Canvas

TEXT TABLE & CROSS TABLE

- Drag Sub-Category to Rows
- Drag Region to Columns
- Drag Sales to Text under Marks

HIGHLIGHT TABLES

- Region > Columns
- Sub-Category > Rows
- SUM(Sales) > Text
- Change drop-down to Square
- Sales > Colors in Marks Card
- Edit Colors > Select color palette

HEAT MAPS

- 2-D representation of data that uses colors to convey info
- As the density of records per mark increases, the color intensity increases.

- Sub-Category > Rows & Region > Columns
- Marks > Square
- SUM(Sales) on Color in Marks Card and edit it

STACKED BAR CHART

- Used to highlight one measure against several other values
- Each category is divided into subcategories for detailed analysis
- Can split bar into categories

- Sub-category > Columns & SUM(Sales) > Rows
- Region > Color under Marks Card

LINE CHART

- Shows trend of data over a period of time
- Normally used when there's a date data type

- Order Date > Columns
- Sales > Rows (Marks type should change automatically to line)
- Date can be discrete or continuous (Right-click on Order Date)

SPARKLINE

- Densely populated line chart that help in understanding anomalies
- "Anomalies" in the data are noticed when the number of data points increases in a line chart.
- Order Date > Columns, click on it and select Day
- Sales > Rows

SCATTER PLOT

- Used to show the correlation between two measures
- Create magic quadrants and identify the relationships between measures.
- Both axeses will have numerical fields
- Other fields can be added via color, size or shape
- Sales > Columns & Profit > Rows
- Sub-Category > Detail
- Adjust colors and shapes

TREE MAP

• Displays hierarchical items in rectangular boxes consisting of a large rectangle divided up into smaller ones to represent sub-categories.

- Region > color & Select Mark Type as square
- Sales > Size & State > Detail & Text

BUBBLE CHART

- The size of the circles denotes the quantity of the metric.
- Region > Color & Select circle for mark type
- Sales > Size
- State > Detail & Text

WORD CLOUD

- Displays frequency of words in a text by making the size of each word proportional to its frequency.
- They can have meta-data associated with them.
- Region > Color & Mark as Text
- Sales > Size
- State > Detail & Text

COMBINED AXIS CHART

- Use multiple mark types in the same sheet!
- Order Date > Columns & select month in date value
- Sales > rows
- Profit on top of Sales Axis ON CHART until two scale symbol appears

DUAL AXIS CHART

- Consists of two independent axes that are on top of each other
- Illustrates relationship between the two variables
- Used to compare multiple measures of the same category.
- Types of charts:
 - Bar vs. line
 - Bar in bar
 - Bar vs. circle
 - Lollipop

- Order date > columns (set by Date Value months)
- Sales > Rows
- Profit next to Sales in Rows
- Right-click on Profit axis & select Dual Axis
- Change axis range of Profit, right-click on profit and select Synchronize axis
- Change type of Mark for Sales to bar

FUNNEL CHART

- Normally used in marketing and sales
- Helps present sales, profit, and revenue at different stages
- Created for single measure value and multiple measure value
- Sales > Rows & select affregate type as SUM
- Region > Color
- SUM(Sales) > Size & put in descending order
- Convert Standard View to Entire View
- Add Region & SUM(Sales) to Label box

LET'S TALK ABOUT MAPS

MAPS

- Provide more context for the dashboard making it easier for users to spot trends in the data.
- There are Symbol Maps and Filled Maps
- State > Details (notice long. And Lat come up automatically)
- Change Marks to Map
- Profit > Color

PANNING, ZOOMING & SELECTING

- Top left corner of a map, there's a toolbar and search box
- The toolbar selects marks on the map, pans & zooms in/out
- To zoom in, select zoom area option and drag to wanted area
- Zooming creates fixed ranges, click the Pushpin to reset.
- Holding SHIFT, you can drag to move around the map
- Dashed outline tools can be used to select specific areas on map

USING MAPS TO FILTER

CAN CREATE ACTIONS ON A MAP THAT FILTER DATA ON A MAP

MAP LAYERING

- Change the appearance of your map by clicking on MAP at the top and choosing "Map Layers"
- You can change the background style
- Hide/Show map layers like land borders, etc.
- ADD data layers

CREATING CUSTOM TERRITORIES

A GROUP IS CREATED BASED ON THE SELECTION OF CUSTOM TERRITORIES EITHER MANUALLY OR BY A CALCULATED FIELD.

MODIFYING LOCATIONS

TABLEAU MIGHT FAIL TO RECOGNIZE THE LOCATION NAMES SO YOU CAN GO IN AND SET THEM MANUALLY.

GEOCODING

- Process of converting text-based description of a location into coordinates.
- Using an external CSV is possible incase Tableau can't identify the geographical field automatically.

ADDING A BACKGROUND IMAGE MAKES THE DASHBOARD MORE APPEALING

MAP SEARCH

- You can search for a specific location on your map
- Postal Code, Continent, City, Country and State/Province
- Located in upper left corner

CALCULATIONS IN TABLEAU

BASIC EXPRESSIONS

- Simple calculated fields created using the "Create Calculated Field" option
- Can be either a dimension or a measure
- Can be created either at a row or aggregated level
- Aggregated level calculations use all fields to create the aggregated field.

LEVEL OF DETAILS

- Powerful functions that alter the level of detail
- Choices can be made between aggregation and granularity
- Three different expressions: include(), exclude() and fixed ()
- Include increases aggregation, exclude increases granularity and fixed increases independence.

FIXED LEVEL OF DETAIL EXAMPLE

FIXED level of detail expressions compute a value using the specified dimensions, without reference to the dimensions in the view

We are going to create a chart that shows the interval between a customer's first purchase date and any subsequent purchase

- Open Tableau, connect to Sample Superstore and pull over Orders to canvas, create a new worksheet
- Click on Analysis and go to Create Calculated Field
- Create 2 calculated fields: a FIXED level of detail expression and a date subtraction
 - FIRST PURCHASE DATE: {FIXED [Customer Name] : MIN([Order Date])}
 - DAYS SINCE FIRST PURCHASE: DATETRUNC('day', [Order Date])-DATETRUNC('day', [FIRST PURCHASE DATE])

- Drag Days Since First Purchase from Measures to Dimensions (it's automatically a Measure because it contains a number)
- Drag Days Since First Purchase to Columns, click on it and choose Continuous
- Drag Sales to Rows and change aggregate to AVG
- Click on Sales and Choose a Running Total Quick Calculation
- Drag First Purchase Date to Color
- Click on it in Color field and change it to Quarter
- Click on the dots next to the Quarter field and click colors

IT'S INTERACTIVE! CLICK ON THE DIFFERENT YEARS IN THE LEGEND

INCLUDE LEVEL OF DETAIL EXAMPLE

- Compute values using the specified dimensions in addition to whatever dimensions are in the view.
- Useful when you want to calculate at a fine level of detail in the database and then re-aggregate and show at a coarser level of detail in your view. Fields based on INCLUDE level of detail expressions will change as you add or remove dimensions from the view.

Create a visual of total sales per customer per region

- Click on Analysis and Select Create Calculated Field
 - Name is SalesPerCustomer
 - {INCLUDE [Customer Name] : SUM([Sales]) }
- Place SalesPerCustomer on ROWS and aggregate it as an AVG
- Put Region in Columns shelf

NOW

Drag SALES over to Rows

Shows the difference between the sum of sales (somewhere between \$390k and 700k per region) AND the avg sales per customer (between 750 and 1100 per region)

EXCLUDE LEVEL OF DETAIL EXAMPLE

- prevent the calculation from using one or more of the dimensions present in the view.
- useful for 'percent of total' or 'difference from overall average' scenarios. They are comparable to such features as Totals and Reference Lines.
- cannot be used in row-level expressions (where there are no dimensions to omit), but can be used to modify either a view level calculation or anything in between (that is, you can use an EXCLUDE calculation to remove dimension from some other level of detail expression).

- Go to Analysis and Select "Create Calculated Field"
 - Name is ExcludeRegion
 - {EXCLUDE [Region] : SUM([Sales])}
- Move Region and Sales to Columns
- Order Date to Rows and make sure it's by MONTH

This breaks out the sum of sales by region and month

Drag ExcludeRegion over to Color

Shades the view to show total sales by month w/o regional component

Table calculations

- Allow transforming values at the detail level of visualization only.
- Executed based on a tabular format with fields in rows and columns.

Ad-hoc calculations

• Temporary calculations that are carried out only for current visualizations.

KEY PERFORMANCE INDICATOR TABLE

- > Sub-Category to Rows, Region to Columns and Sales to Text
- ➤ Create a Calculated Field "KPI"
 - ➤ IF SUM ([SALES]) > 25000 THEN "ABOVE BENCHMARK" ELSE "BELOW BENCHMARK" END
- ➤ Change Mark card to Shape from drop down
- ➤KPI > Shape
- ➤ Click Shape and choose KPI: Above (green) and Below(red)
- ➤ Change Sums from Text to Detail

EXTRAS

- Tooltips display when you put the mouse over one or more marks in view
- Utilize the Analytics tab for Modeling & Summarizing
- Use trend lines in predicting of given data (linear, logarithmic, exponential and polynomial)
- Forecasting depends on the number of historical data points available
- Clustering groups data points together and separating them from other dissimilar data objects

YOUR DASHBOARD

THE INTERFACE

- Default = the device that the dashboard will be created
- Size = size of your presentation screen
- Sheets = shows the names of your Worksheets
- Objects = how to place sheets and add other helpful items

OBJECTS ON YOUR DASHBOARD

- Horizontal/Vertical: how to display your sheets and other objects
- Text: creates a text box
- Image: insert outside images (like a pic of a matplotlib plot!)
- Webpage: creates a web page interface
- Blank: creates blank box
- Navigation: provides users with ability to build navigation buttons
- Download: dashboard can be downloaded in PDF format
- Extensions: third-party allowing customized visualizations
- Ask Data: Al functionality from Tableau Server

- Click and drag sheets over to the canvas to use
- Tiled: Fixed Location & Floating: unfixed location
- You can highlight data & add URL links

BEST PRACTICES

- Choose appropriate visualizations
- Line charts: best for seeing change over time
- Maps: best for seeing geographical data
- Bar Charts: best for seeing highest or lowest value
- It is difficult to read long names vertically
- Choose appropriate size: avoid scroll bars & scrunched views

YOUR STORYBOARD

- Groups of dashboards and visuals arranged to tell your story
- They are dynamic and interactive!
- Left side has created dashboards, text option and size
- Layout tab has Navigator styles
- The grey boxes designate slides & provide area for you to make notes
- Drag visualizations and dashboards as appropriate

• Your work gets saved automatically to Tableau Public