

DATA VISUALIZATION: INTRODUCTION TO TABLEAU

D-VELOP WORKSHOP SERIES – Summer 2021
Trevor Bonjour

D-VELOP WORKSHOP SERIES – Summer 2021

Jun 9

- Data Visualization: ggplot2

Jun 16

- Data Visualization using Python: Matplotlib and Seaborn

Jun 23

- Exploratory Data Analysis in R

July 7

- Data Visualization using Python: Bokeh (Interactive Plots)

July 14

- Exploring and Visualizing Time Series Data

July 21

- **Data Visualization: Introduction to Tableau**



PURDUE
UNIVERSITY®

Libraries and School
of Information Studies

What will we cover today?

- Motivation
- What is Tableau?
- Tableau Workflow
- Important Components
- Learn by Doing

Visualization Objectives

- Record information
- Analyze data to support reasoning
- Confirm hypotheses
- Communicate ideas to others

Why Visualize

To record information



Why Visualize

To point out interesting things

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPNSDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

Why Visualize

To point out interesting things

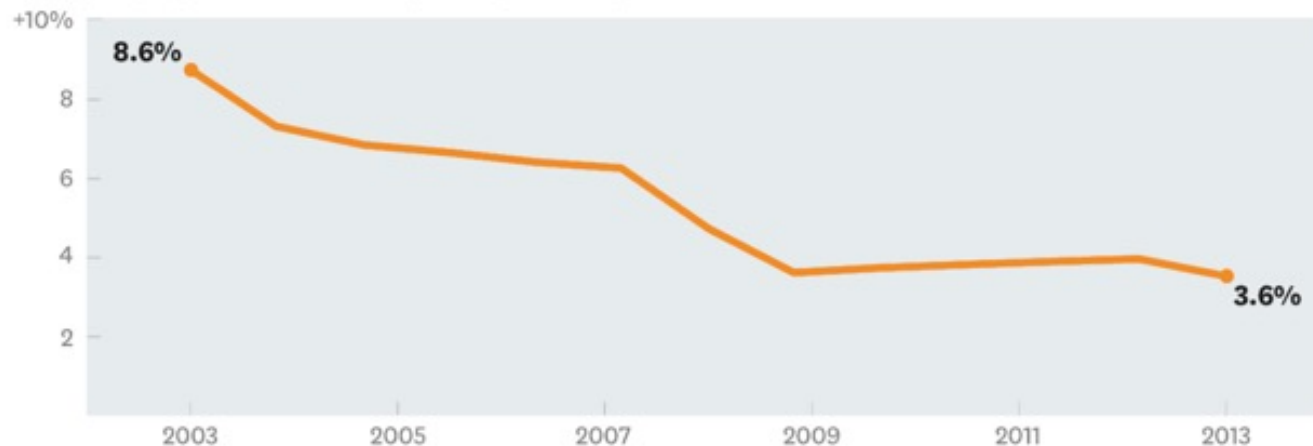
MTHI**V**LWYADCEQGCHKILKMTWYN
ARDCAIREQGH**L**KMFPSSTWYARN
GFPS**V**CEILQGKMFPNSDRCEQDIFP
SGHLMFH**K**M**V**PSTWYACEQTWRN

Why Visualize

To communicate information

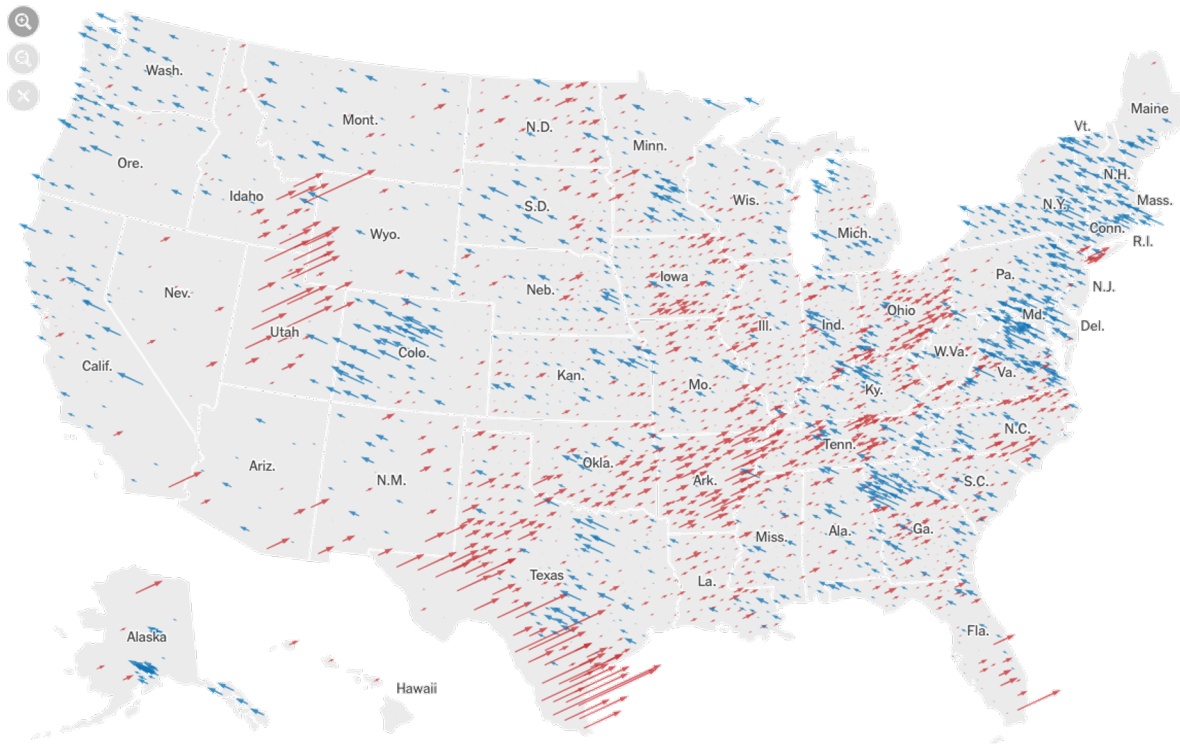
Annual Growth is Declining

ANNUAL GROWTH IN HEALTH CARE SPENDING



Why Visualize

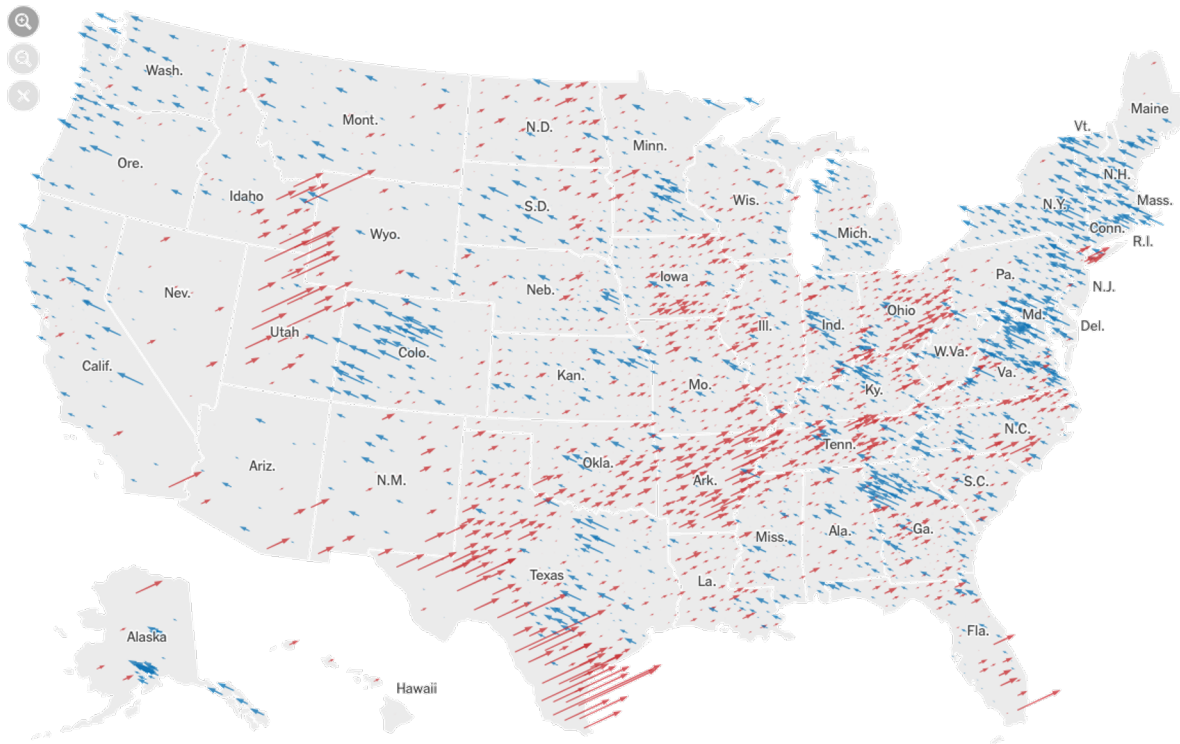
To analyze data



2020 US Elections (NYTimes)

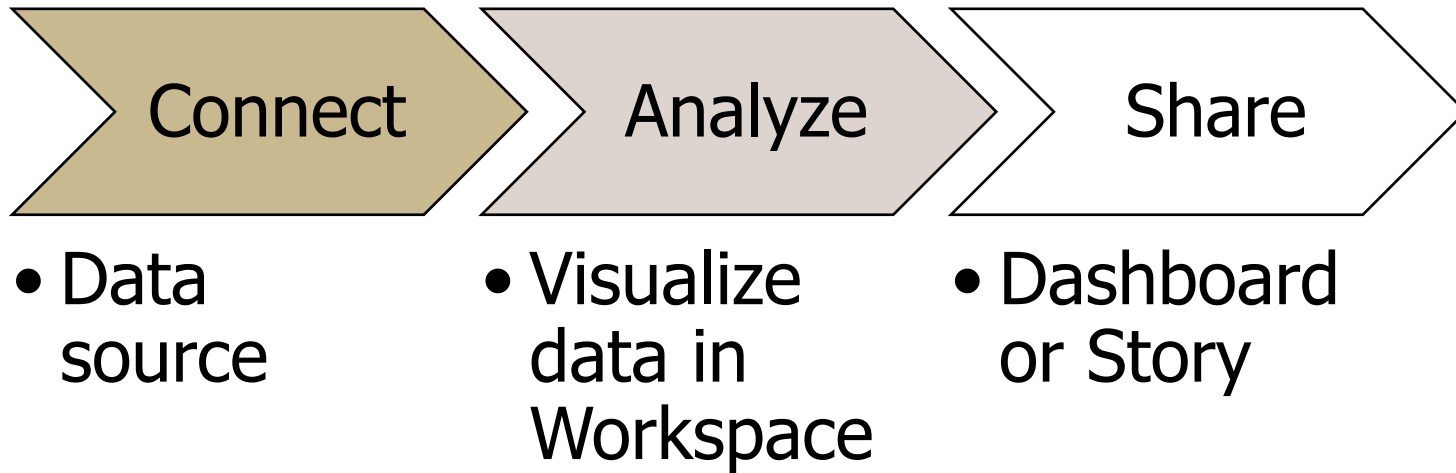
Why Visualize

To analyze data



2020 US Elections (NYTimes)

Tableau Workflow



Connect

Connect

Search for Data

Tableau Server

To a File

- Microsoft Excel
- Text file
- JSON file
- PDF file
- Spatial file
- Statistical file
- More...

To a Server

- Microsoft SQL Server
- MySQL
- Oracle
- Amazon Redshift
- More...

Saved Data Sources

- Sample - Superstore
- World Indicators

Open

Superstore

overview_v1

overview_v1_test

Book1

edu_data_analysis

Highest Score

test1

Sample Workbooks

Superstore


Regional

World Indicators

[Open a Workbook](#)

[More Samples](#)

Connect



Connect

Search for Data

Tableau Server

To a File

- Microsoft Excel
- Text file
- JSON file
- PDF file
- Spatial file
- Statistical file
- More...

To a Server

- Microsoft SQL Server
- MySQL
- Oracle
- Amazon Redshift
- More...

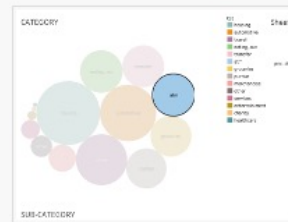
Saved Data Sources

- Sample - Superstore
- World Indicators

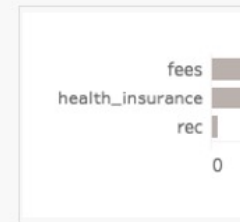
Open



Superstore



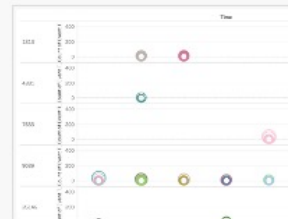
overview_v1



overview_v1_test

1	A	False
	B	True
	C	False
2	A	False
	B	True
	C	True

Highest Score



test1

Data Sources Types

Spreadsheets

- Excel or csv file

Relational Databases

- MySQL or Oracle

Cloud Data

- AWS or Microsoft Azure

Other Sources

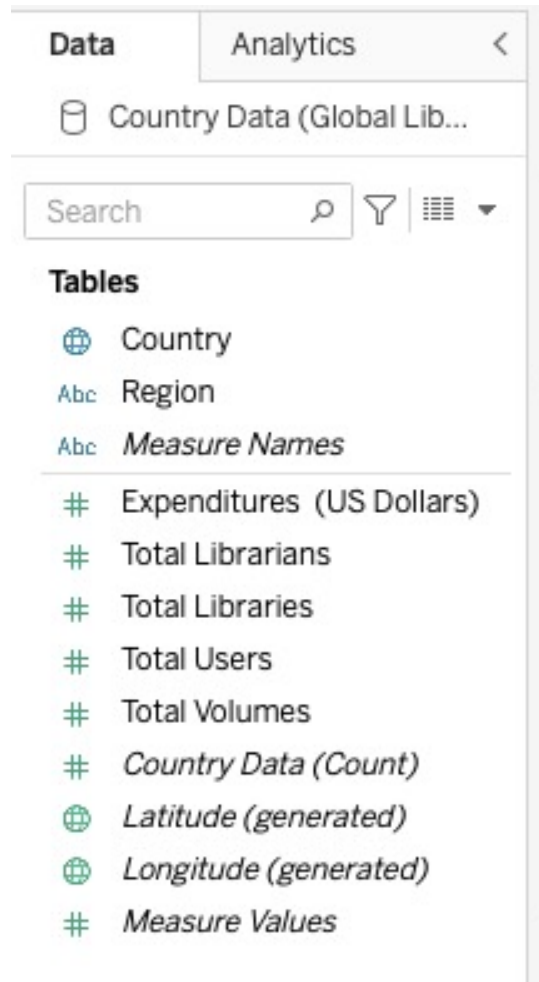
- Spatial Files or R

Data Field

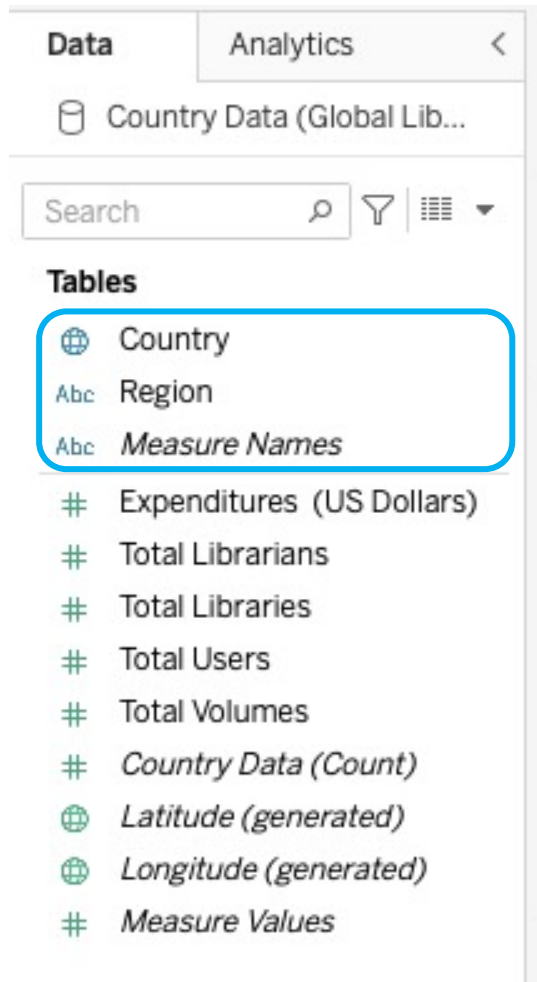
A field, also known as a column, is a single piece of information from a record in a data set.

- Qualitative Field (Dimensions)
 - Describes or Categorizes Data
 - What, when or who
 - Slices the quantitative data
- Quantitative Field (Measures)
 - Numerical Data
 - Provides measurement for qualitative category
 - Can be used in calculations

Data Field

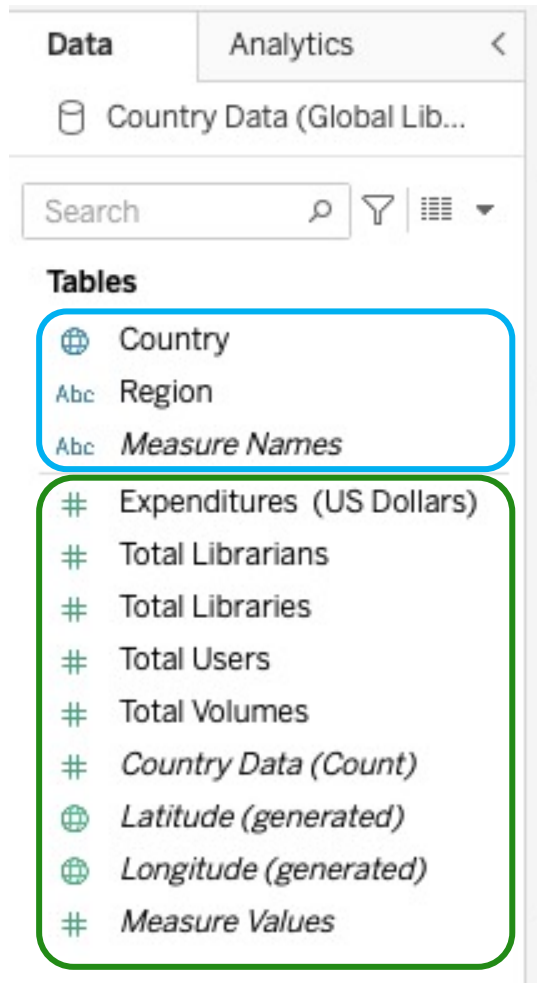


Data Field



Dimensions

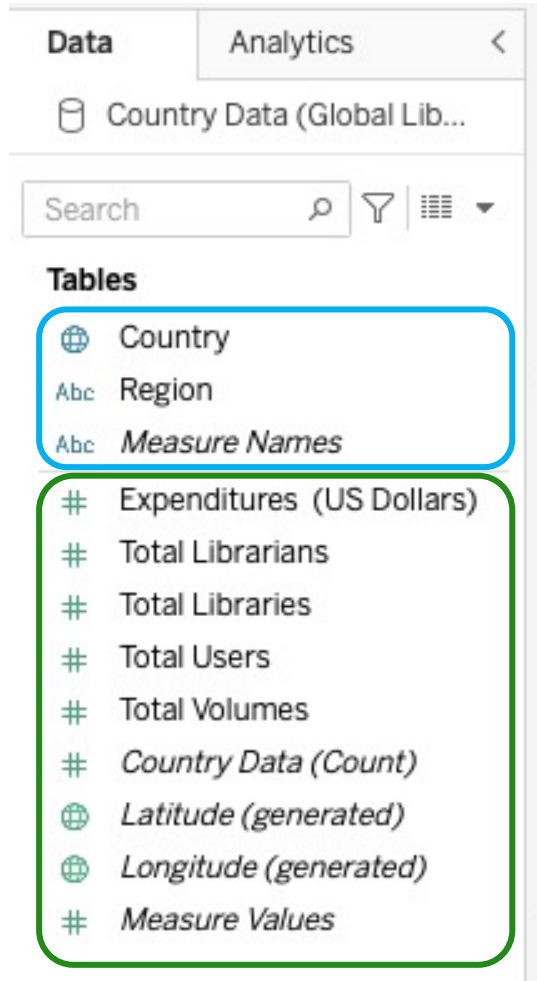
Data Field



Dimensions

Measures

Data Field

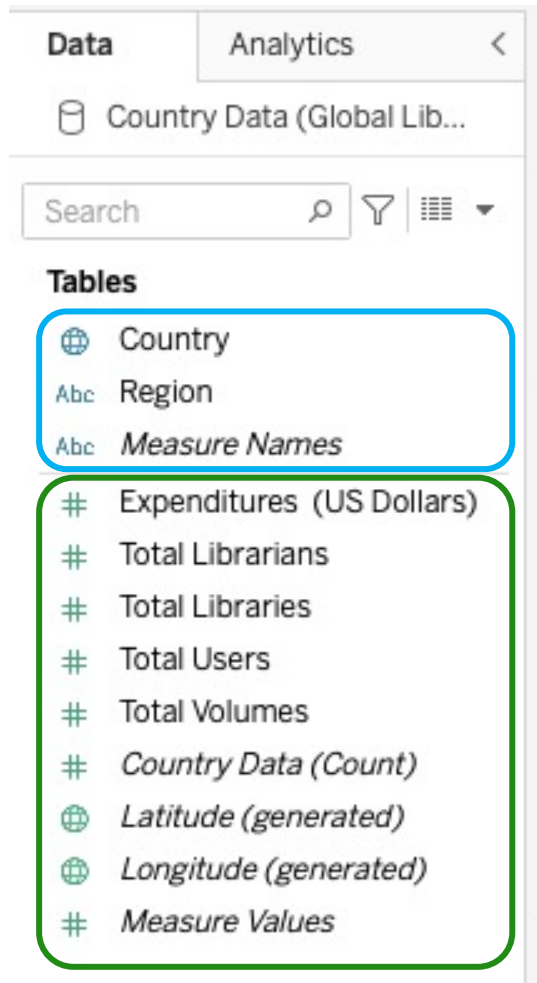


Dimensions

Measures

- By default, aggregated by SUM
- Can be aggregated as average, median, count, or count distinct.

Data Field



Dimensions

- Break down the aggregated total into smaller totals by category.

Measures

- By default, aggregated by SUM
- Can be aggregated as average, median, count, or count distinct.

Data Types

Abc Text or String Values



Discrete Date/Time



Discrete Date



Geographic field -
State or Zip Code



Continuous
Numeric Value



Calculated Field

Chart Types



Line — View trends in data over time.

Examples: Stock price change over a five-year period or website page views during a month.



Bar — Compare data across categories.

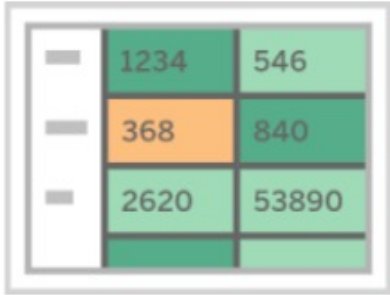
Examples: Volume of shirts in different sizes, or percent of spending by department.



Heat Map — Show the relationship between two factors.

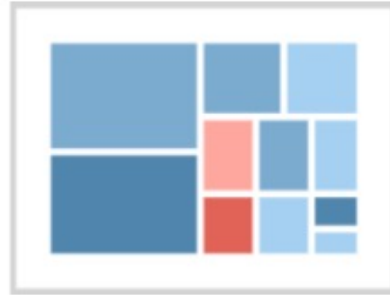
Examples: Segment analysis of target market, or sales leads by individual rep.

Chart Types



Highlight Table — Shows detailed information on heat maps.

Examples: The percent of a market for different segments, or sales numbers in a region.



Treemap — Show hierarchical data as a proportion of a whole.

Examples: Storage usage across computer machines, comparing fiscal budgets between years.



Gantt — Show duration over time.

Examples: Project timeline, duration of a machine's use, availability of players on a team.

Chart Types



Bullet — Evaluate performance of a metric against a goal.

Examples: Sales quota assessment, performance spectrum (great/good/poor).



Scatterplot — Investigate relationships between quantitative values.

Examples: Male versus female likelihood of having lung cancer at different ages



Histogram — Understand the distribution of your data.

Examples: Number of customers by company size, student performance on an exam, frequency of a product defect.

Chart Types



Symbol maps — Use for totals rather than rates. Be careful, as small differences will be hard to see.

Examples: Number of customers in different geographies.



Area maps — Use for rates rather than totals. Use sensible base geography.

Examples: Rates of internet-usage in certain geographies, house prices in different neighborhoods.



Box-and-Whisker — Show the distribution of a set of a data.

Examples: Understanding your data briefly, seeing how data is skewed towards one end, identifying outliers in your data.

Tableau Workspace

The screenshot displays the Tableau Workspace interface. At the top is a toolbar with navigation and visualization icons. Below the toolbar, the interface is divided into several panes:

- Data Pane (Left):** Shows the data source 'Country Data (Global Lib...)' and a list of fields under the 'Tables' section:
 - Country
 - Region
 - Measure Names
 - Expenditures (US Dollars)
 - Total Librarians
 - Total Libraries
 - Total Users
 - Total Volumes
 - Country Data (Count)
 - Latitude (generated)
 - Longitude (generated)
 - Measure Values
- Columns and Rows Shelves (Top Right):** Both are currently empty.
- Marks Card (Bottom Left):** The mark type is set to 'Automatic'. Below it are buttons for 'Color', 'Size', 'Text', 'Detail', and 'Tooltip'.
- Filters Pane (Middle Left):** Currently empty.
- Workspace (Center):** A large area for creating visualizations, with 'Drop field here' prompts.

Tableau Workspace

The screenshot displays the Tableau Workspace interface. On the left, the **Data** pane is highlighted with a yellow border. It shows a search bar and a list of fields under the **Tables** section. The fields include: Country, Region, Measure Names, Expenditures (US Dollars), Total Librarians, Total Libraries, Total Users, Total Volumes, Country Data (Count), Latitude (generated), Longitude (generated), and Measure Values. The main workspace area is titled **Data pane** and contains the text: "The **Data** pane lists all the fields from the source data." Below this text, there are two buttons labeled **Detail** and **Tooltip**. To the right of the buttons, there is a placeholder text "Drop field here". The top of the interface shows the **Columns** and **Rows** shelves, and a toolbar with various icons.

Tableau Workspace

The screenshot shows the Tableau Workspace interface. On the left is the 'Data' pane with a search bar and a list of tables including 'Country', 'Region', 'Measure Names', and various measures like 'Expenditures (US Dollars)', 'Total Librarians', 'Total Libraries', 'Total Users', 'Total Volumes', 'Country Data (Count)', 'Latitude (generated)', 'Longitude (generated)', and 'Measure Values'. The main area is divided into 'Columns' and 'Rows' shelves, which are highlighted by a yellow callout box. Below these shelves are the 'Filters' and 'Marks' shelves. The 'Marks' shelf currently shows 'Automatic' and has buttons for 'Color', 'Size', 'Text', 'Detail', and 'Tooltip'. The top toolbar contains various icons for navigation, data manipulation, and visualization.

Columns and Rows

The **Columns** and **Rows** shelves are used to create a structure for your visualization. You can place any number of fields on these shelves.

Tableau Workspace

Data | Analytics | Pages

Country Data (Global Lib...)

Search

Tables

- Country
- Region
- Measure Names
- Expenditures (US Dollars)
- Total Librarians
- Total Libraries
- Total Users
- Total Volumes
- Country Data (Count)
- Latitude (generated)
- Longitude (generated)
- Measure Values

Filters

Marks

Automatic

Color Size Text

Detail Tooltip

Columns

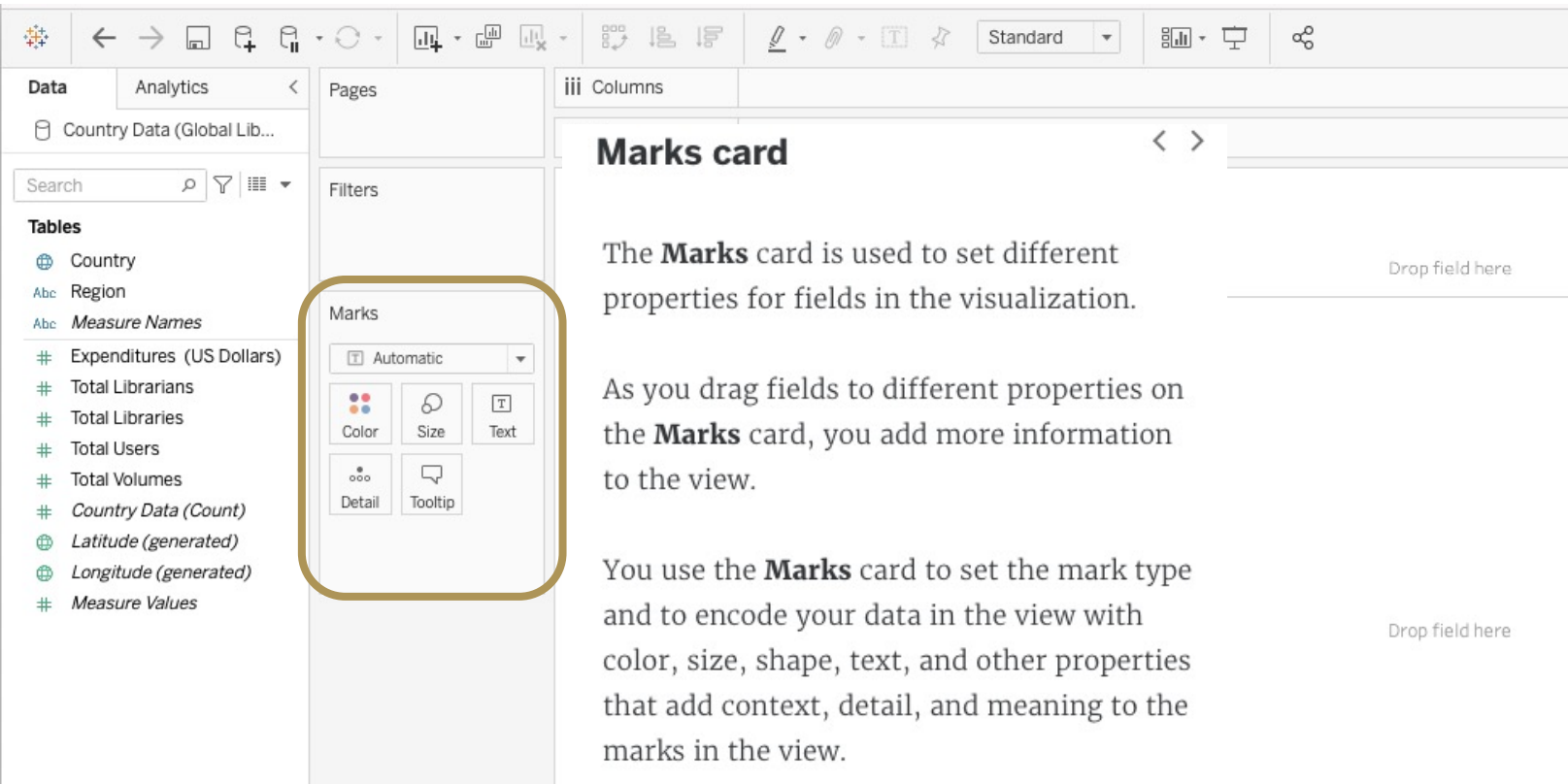
Rows

Filters

Filters allow you to include and exclude specific data. The different types of filters will be covered in detail in other lessons.

Drop field here

Tableau Workspace



The screenshot shows the Tableau Workspace interface. On the left, the 'Data' pane displays a list of tables: Country, Region, Measure Names, Expenditures (US Dollars), Total Librarians, Total Libraries, Total Users, Total Volumes, Country Data (Count), Latitude (generated), Longitude (generated), and Measure Values. The 'Marks' card is highlighted with a yellow border. The 'Columns' shelf is empty. The 'Marks card' section on the right contains the following text:

Marks card

The **Marks** card is used to set different properties for fields in the visualization.

As you drag fields to different properties on the **Marks** card, you add more information to the view.

You use the **Marks** card to set the mark type and to encode your data in the view with color, size, shape, text, and other properties that add context, detail, and meaning to the marks in the view.

Drop field here

Drop field here

Tableau Workspace

The screenshot displays the Tableau Workspace interface. At the top is a toolbar with various icons for navigation and visualization. Below the toolbar, the 'Data' pane on the left shows a list of tables under the heading 'Tables'. The 'View' pane on the right is titled 'View' and contains a large area labeled 'Drop field here' at the top and bottom. The 'Columns' shelf is visible above the view area. The 'Data' pane lists the following tables:

- Cou
- Reg
- Mez
- Exp
- Totz
- Totz
- Totz
- Total Volumes
- Country Data (Count)
- Latitude (generated)
- Longitude (generated)
- Measure Values

The 'View' pane also includes a 'Detail' button and a 'Tooltip' button. The 'Columns' shelf is currently empty.

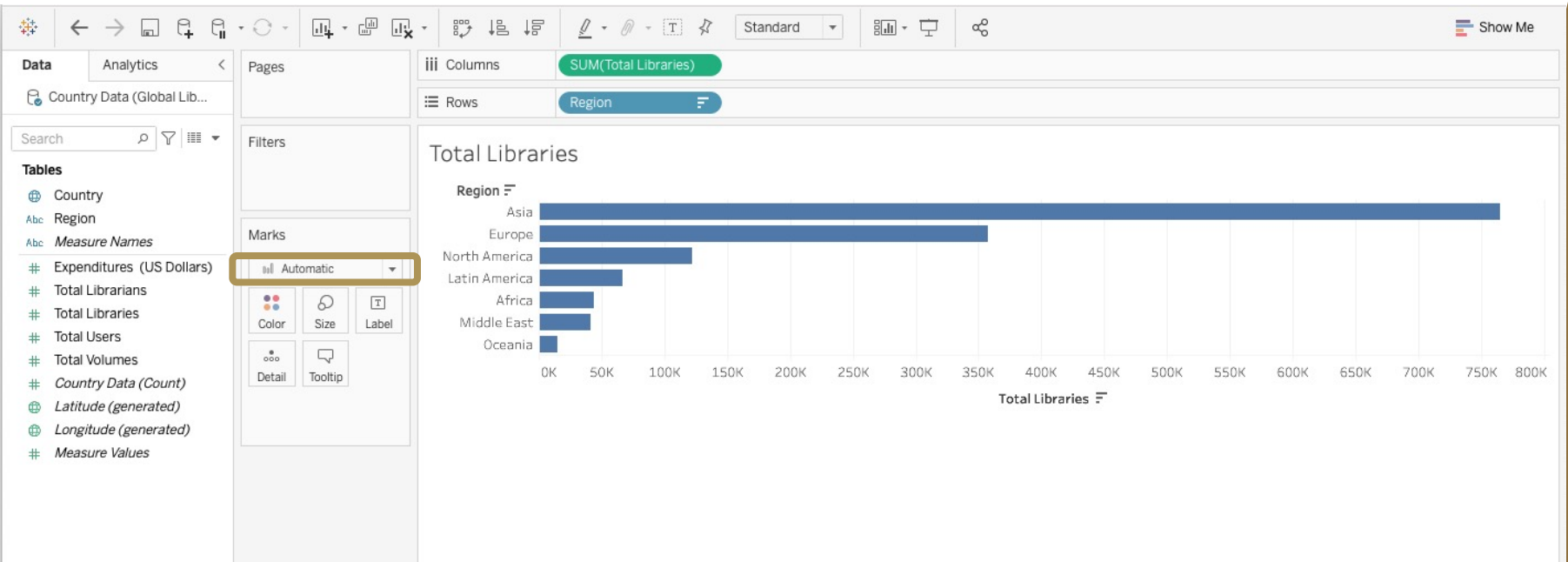
Visualization

The screenshot displays the Tableau interface with the following components:

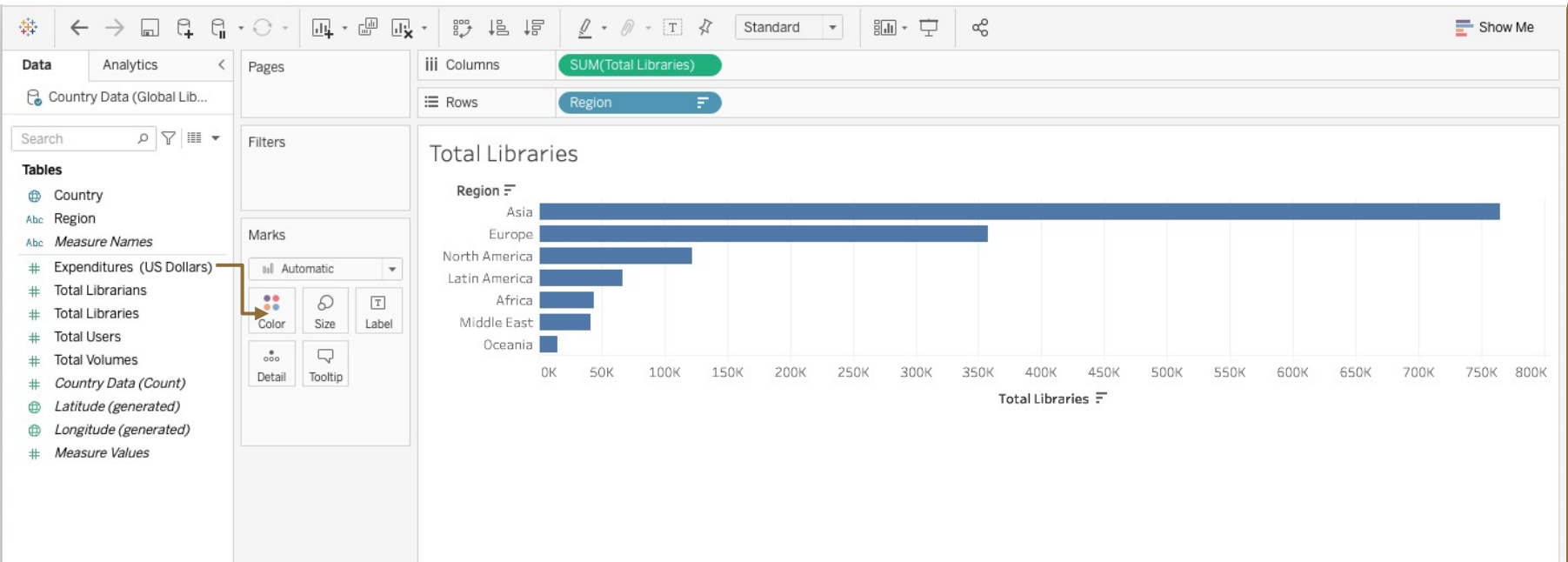
- Top Bar:** Includes navigation icons, a 'Standard' view dropdown, and sharing options.
- Data Pane (Left):**
 - Country Data (Global Lib...)**: The selected data source.
 - Search**: A search bar with a magnifying glass icon.
 - Tables**: A list of fields including 'Country', 'Region', 'Measure Names', 'Expenditures (US Dollars)', 'Total Librarians', 'Total Libraries', 'Total Users', 'Total Volumes', 'Country Data (Count)', 'Latitude (generated)', 'Longitude (generated)', and 'Measure Values'.
- Columns Shelf:** Labeled 'Columns' at the top, it contains the 'Country' field.
- Rows Shelf:** Labeled 'Rows' at the top, it contains the 'Expenditures (US Dollars)' field.
- Marks Card:** Labeled 'Marks' at the top, it shows 'Automatic' as the mark type. Below it are icons for 'Color', 'Size', 'Text', 'Detail', and 'Tooltip'.
- Workspace:** The main area for building the visualization, currently showing a blank grid with 'Drop field here' prompts.

Arrows indicate the drag-and-drop process: one arrow points from 'Country' in the Tables list to the Columns shelf, and another points from 'Expenditures (US Dollars)' in the Tables list to the Rows shelf.

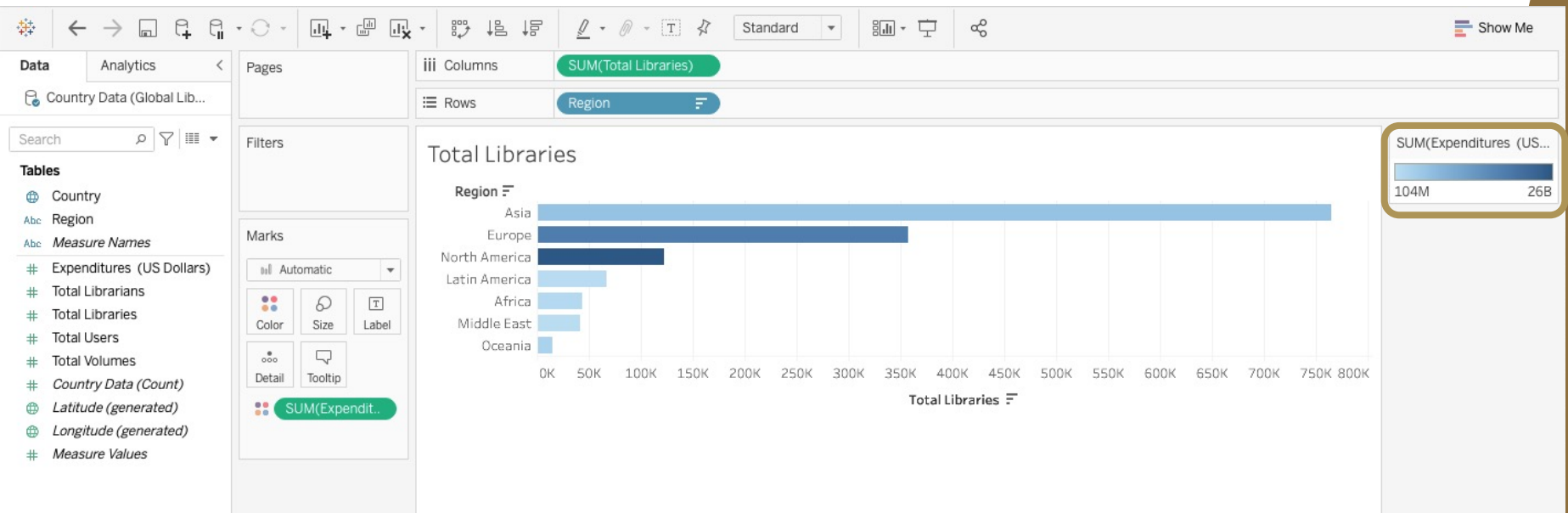
Visualization



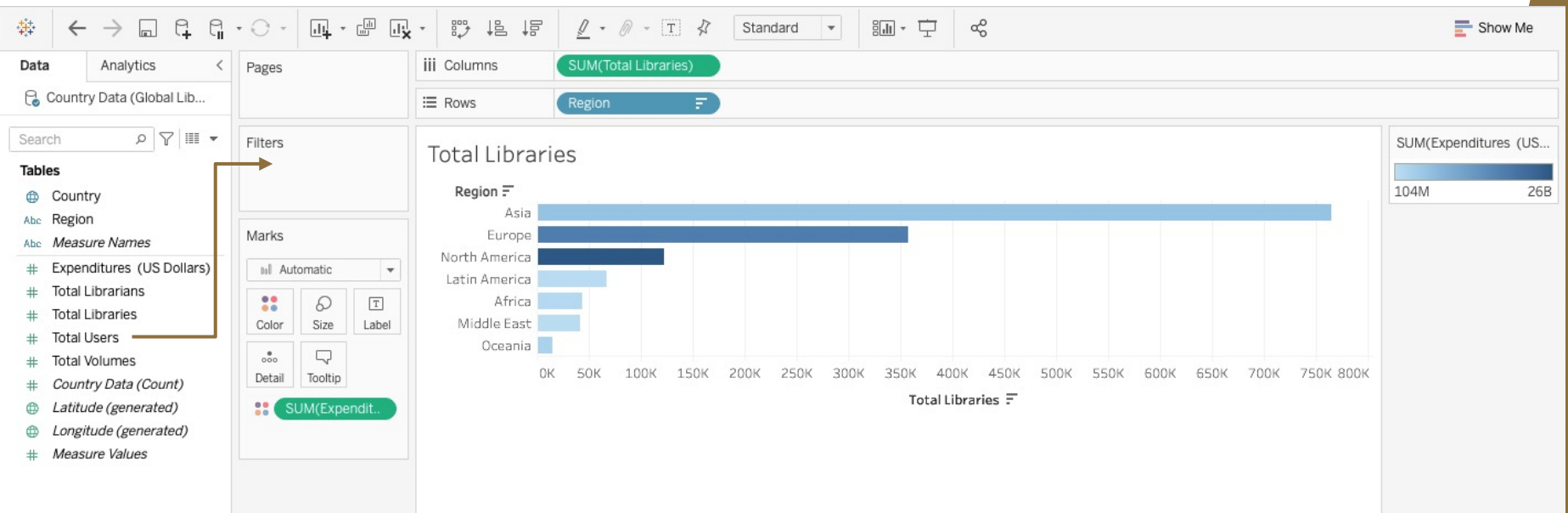
Visualization



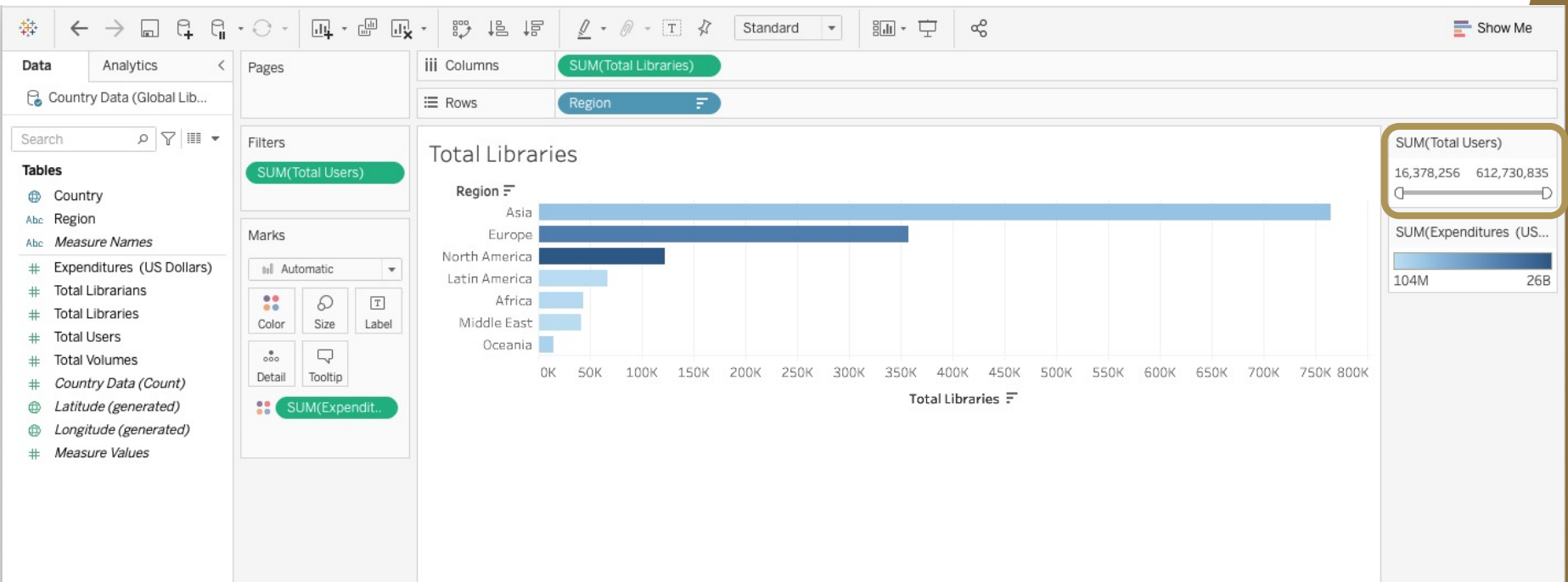
Visualization



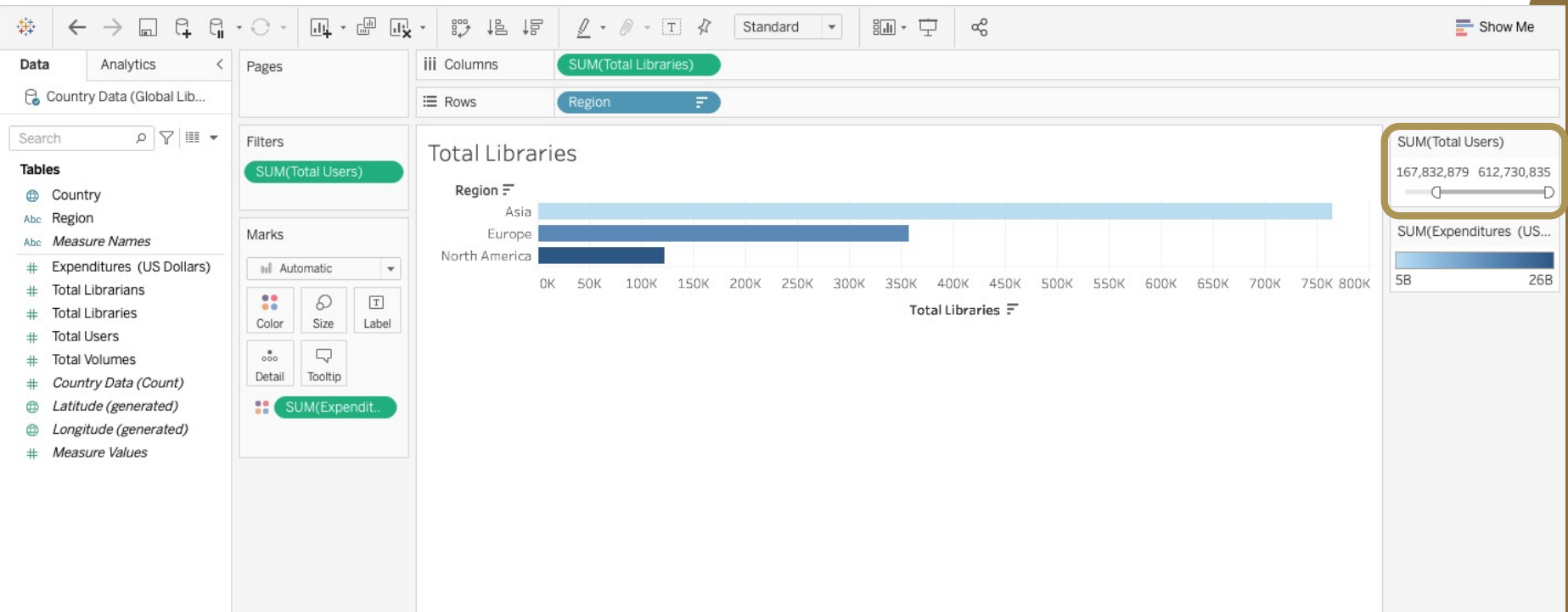
Visualization



Visualization



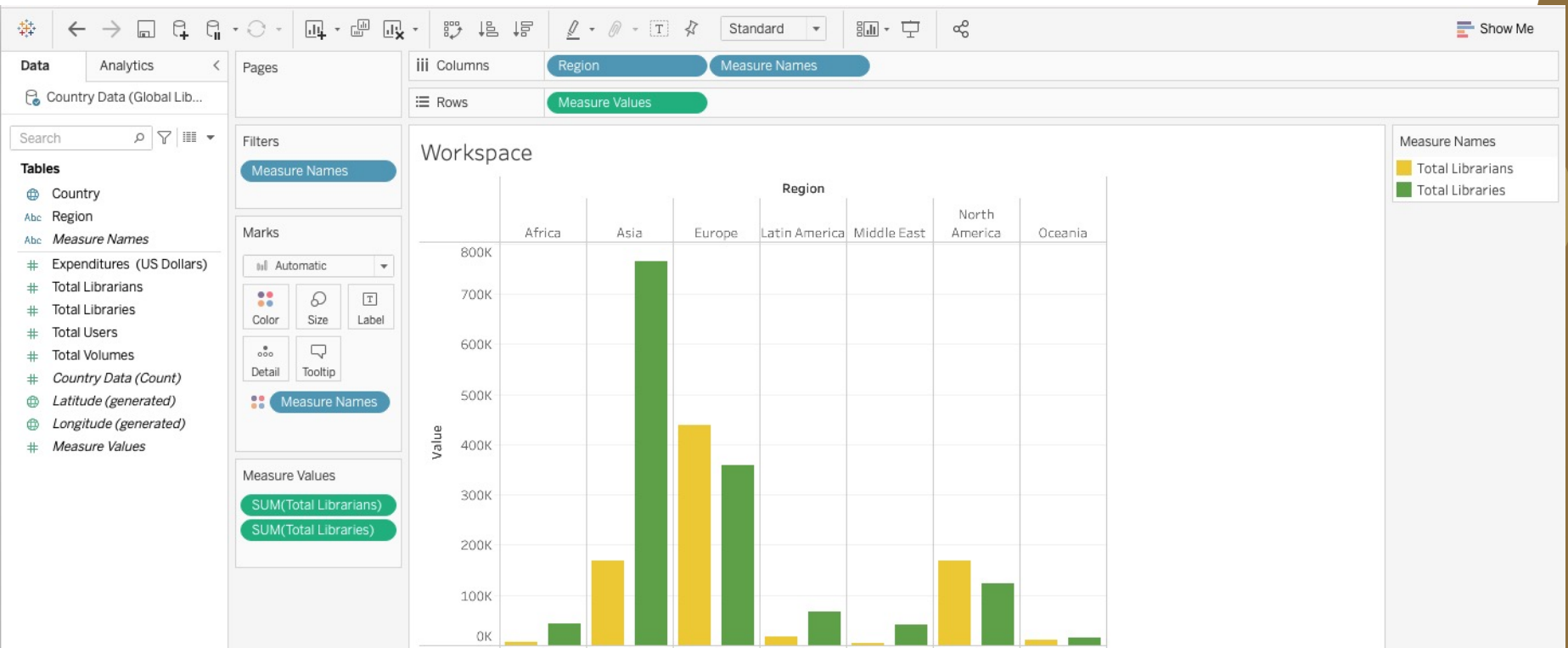
Visualization



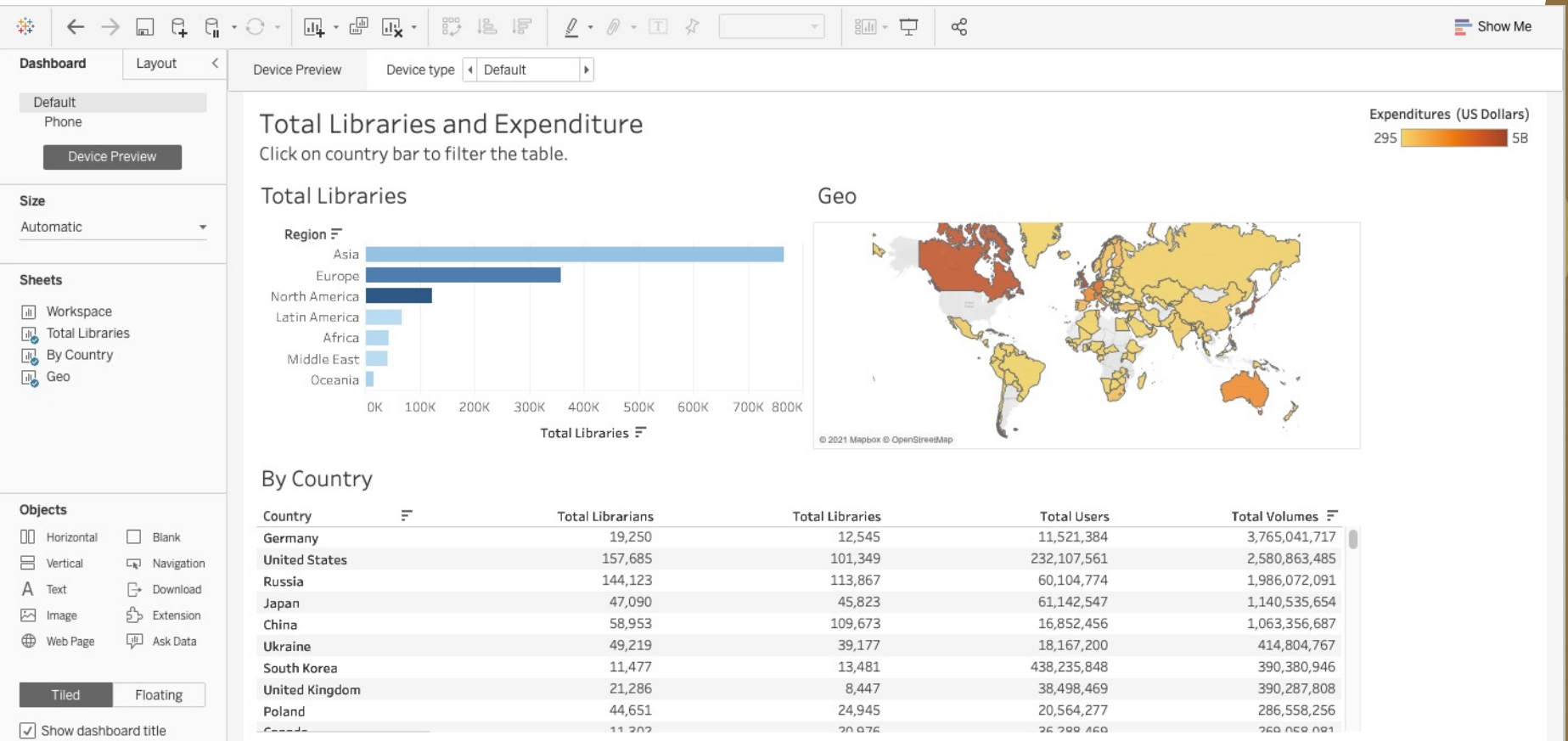
Visualization

The screenshot shows the Tableau Desktop interface with the 'Country Data (Global Lib...)' dataset loaded. The 'Columns' shelf is empty, and the 'Rows' shelf is empty. The 'Marks' card is set to 'Automatic'. The 'Filters' shelf is empty. The 'Tables' list on the left includes 'Country', 'Region', 'Measure Names', 'Expenditures (US Dollars)', 'Total Librarians', 'Total Libraries', 'Total Users', 'Total Volumes', 'Country Data (Count)', 'Latitude (generated)', 'Longitude (generated)', and 'Measure Values'. The 'Workspace' area contains two 'Drop field here' prompts. A 'Show Me' panel on the right displays various chart templates and a 'For side-by-side bars try' section with buttons for 'Dimensions' and 'Measures'.

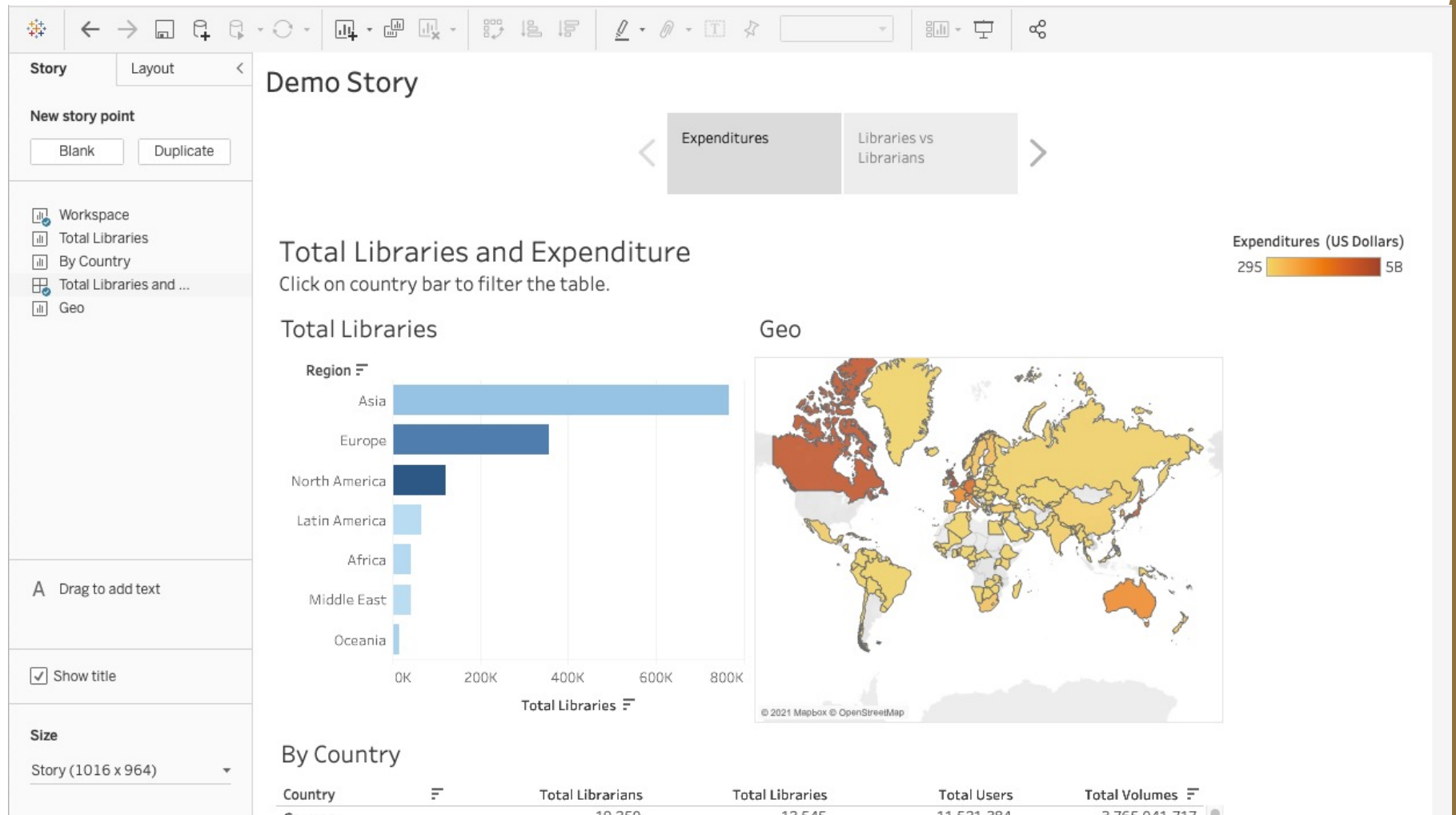
Visualization



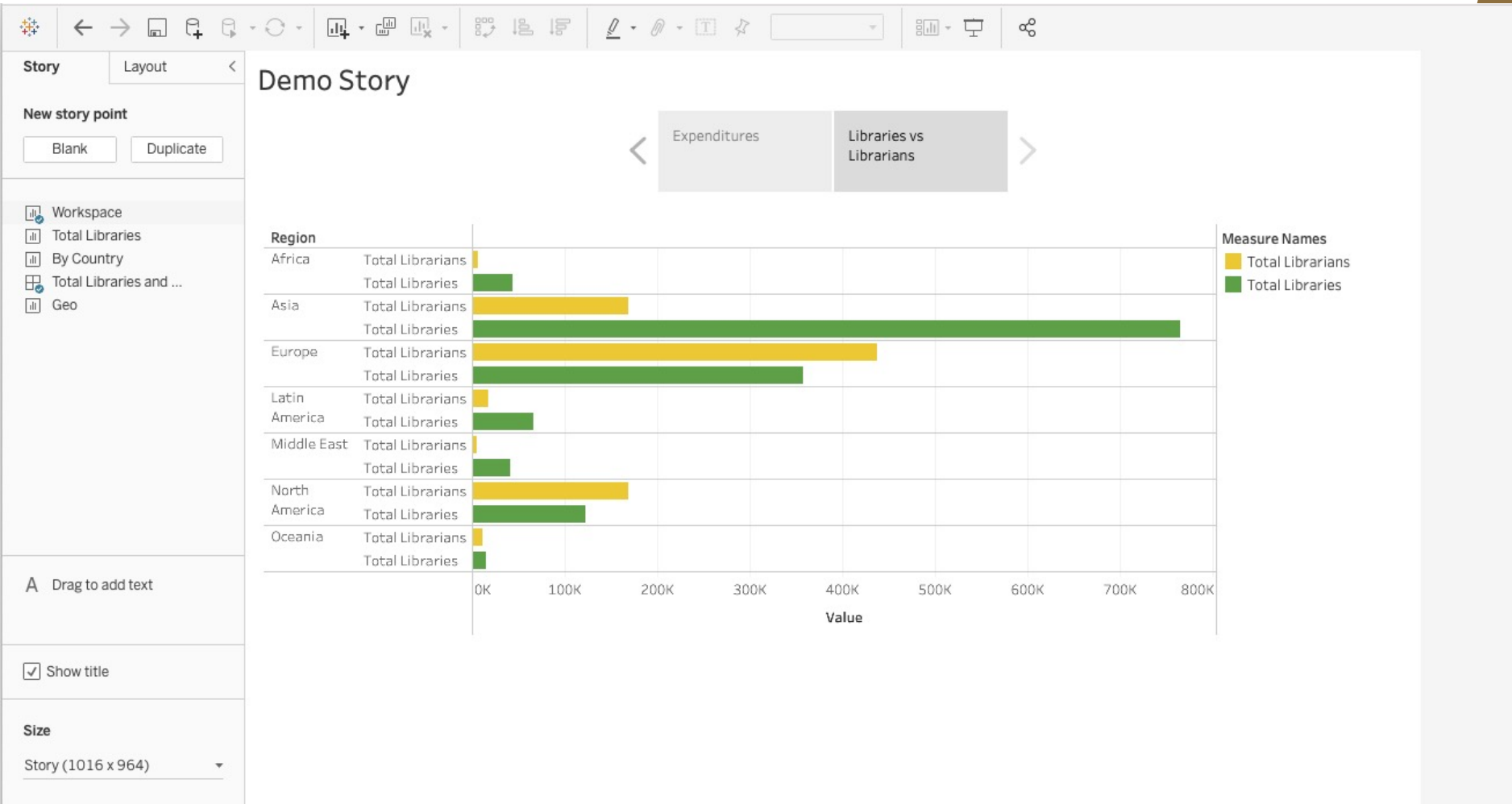
Dashboard



Story



Story



Share

- Saved File
 - .twb or .twbx
- Publish to Tableau Server (Secure)
- Publish to Tableau Public (Unsecure)

LEARN BY DOING

To access the videos and material from the workshop series please visit:
<https://guides.lib.purdue.edu/d-velop>