DATA VISUALIZATION: INTRODUCTION TO TABLEAU

D-VELOP WORKSHOP SERIES – Summer 2021

Trevor Bonjour

PURDUE UNIVERSITY | Libraries and School of Information Studies

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
GFPSVCEILQGKMFPSNDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

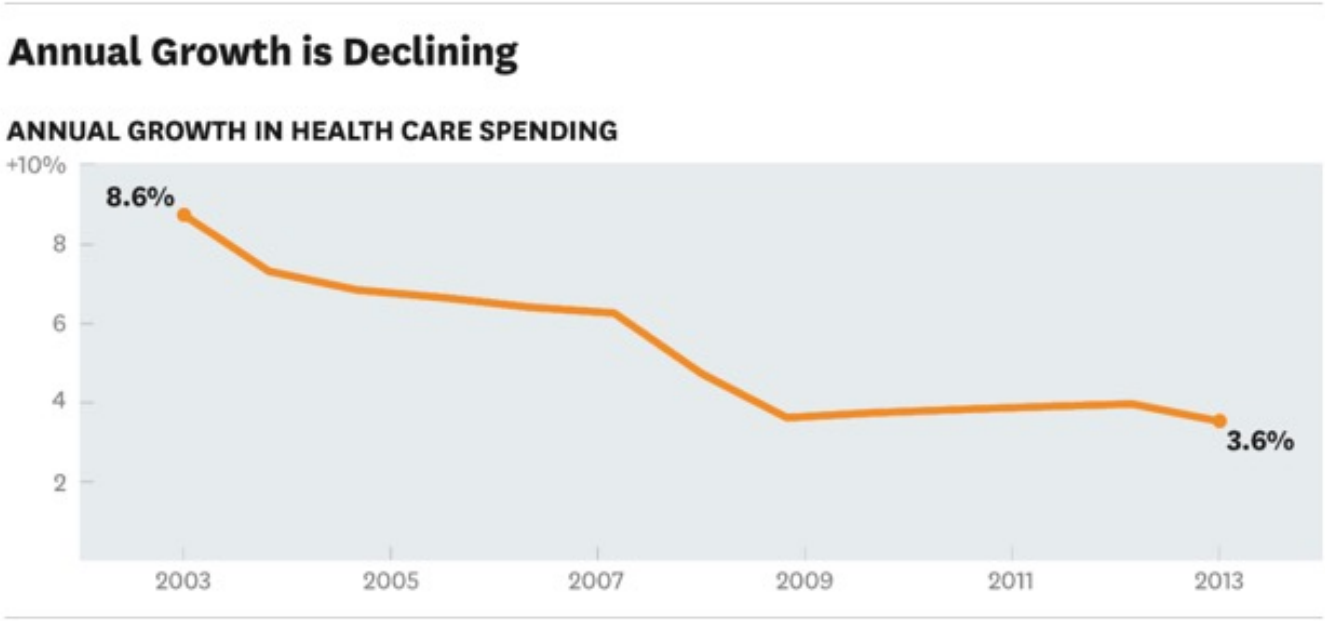
Why Visualize

To point out interesting things

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPSNDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

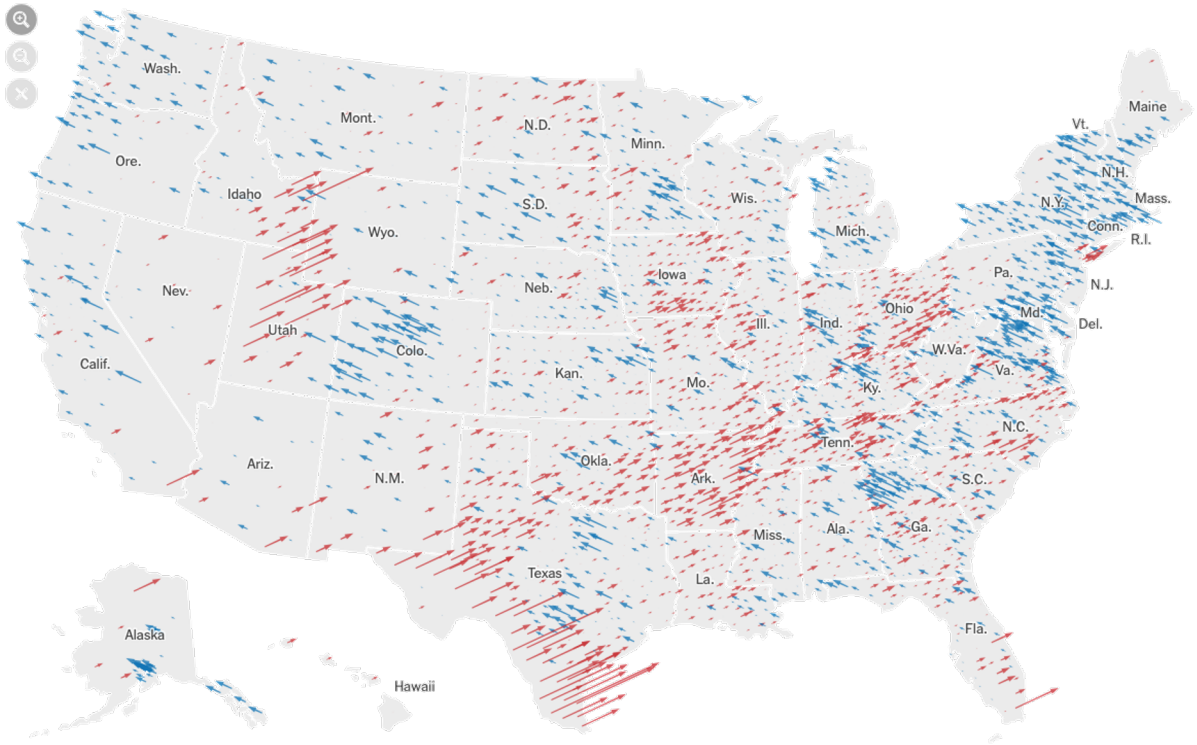
Why Visualize

To communicate information



Why Visualize

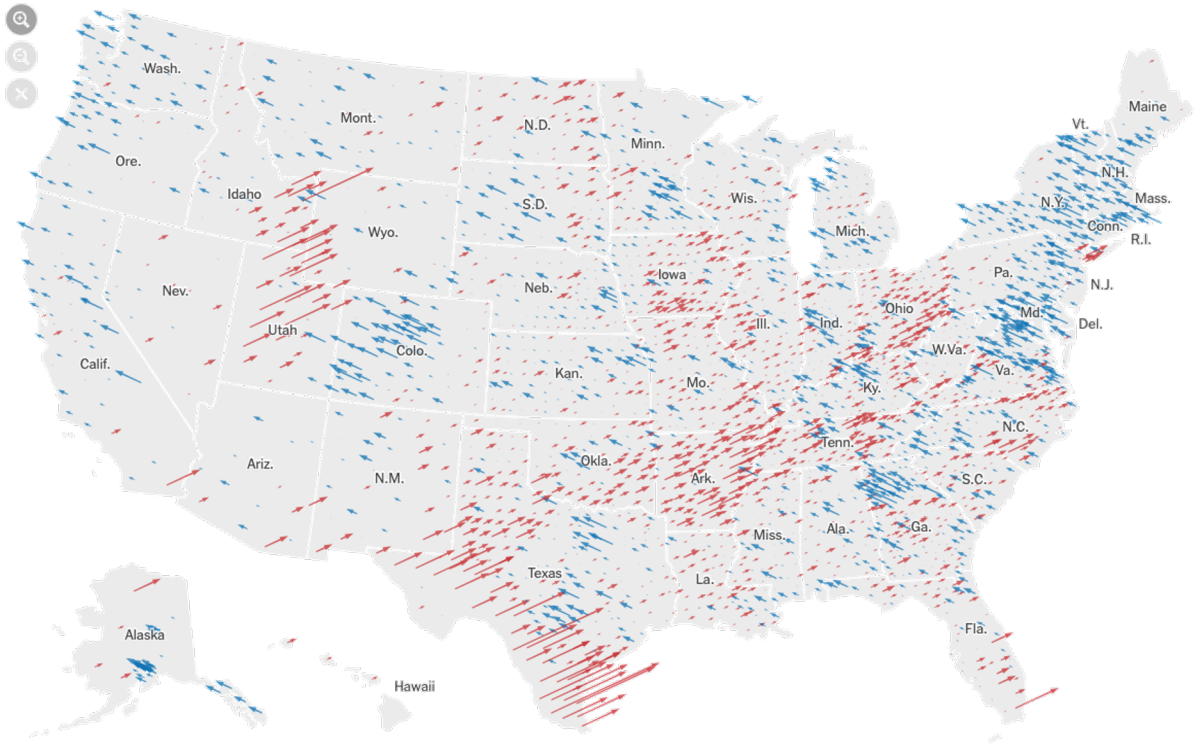
To analyze data



2020 US Elections (NYTimes)

Why Visualize

To analyze data

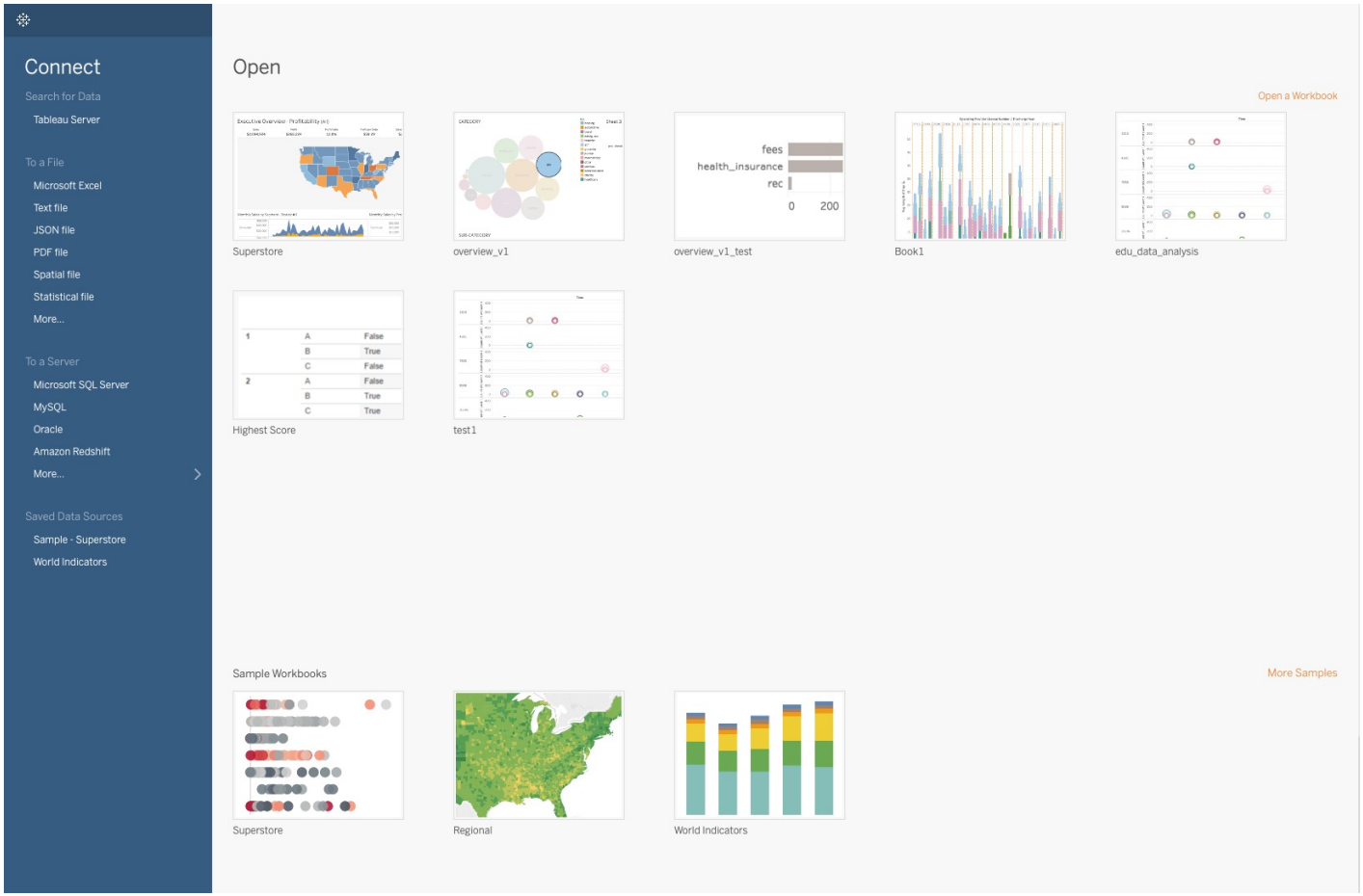


2020 US Elections (NYTimes)

Tableau Workflow

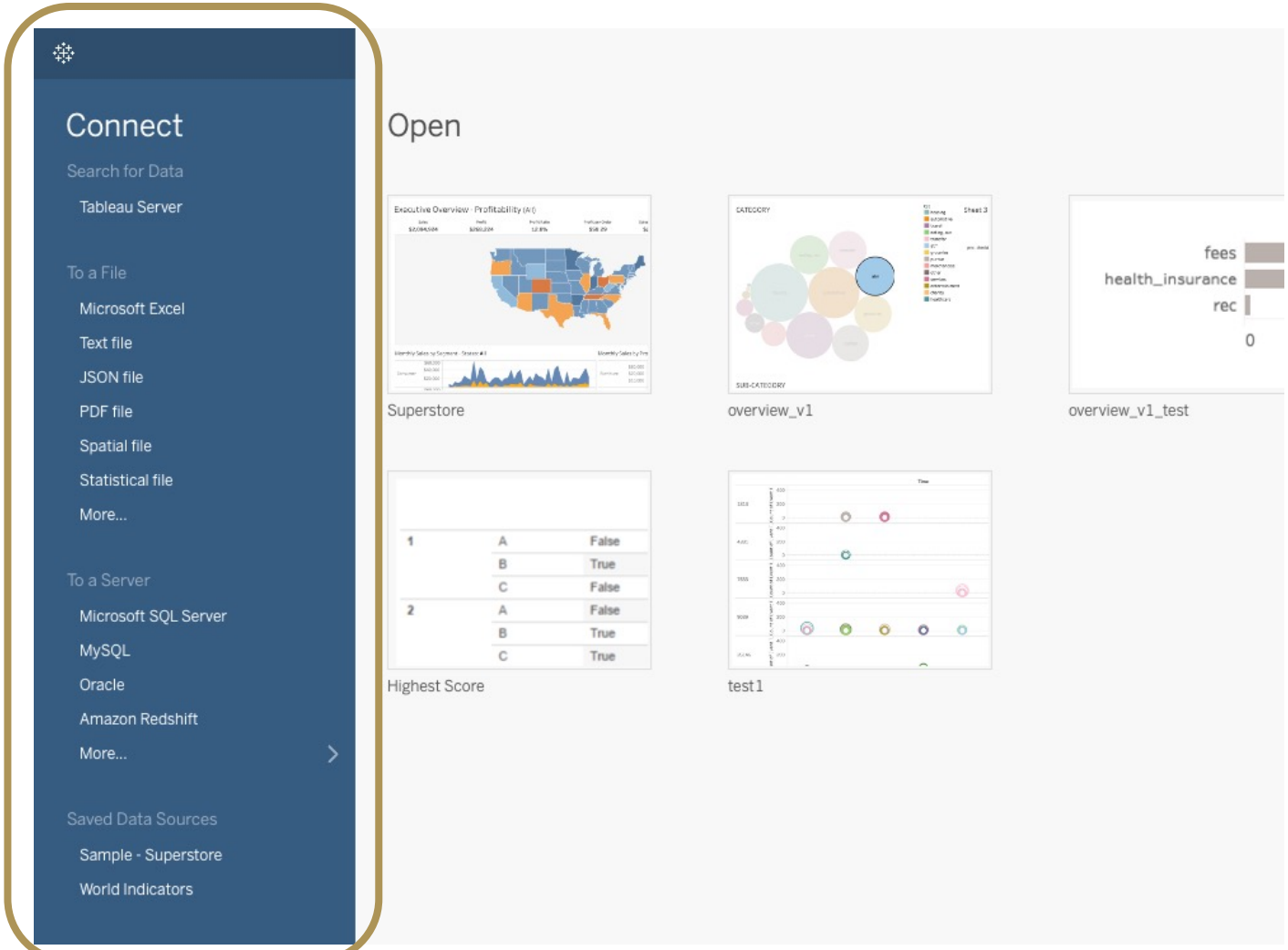
- Data source
- Visualize data in Workspace
- Dashboard or Story

Connect



PURDUE UNIVERSITY | Libraries and School of Information Studies

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
fees
health_insurance
rec
0
overview_v1_test
1 A False
B True
C False
2 A False
B True
C True
Highest Score
test1

Data Sources Types

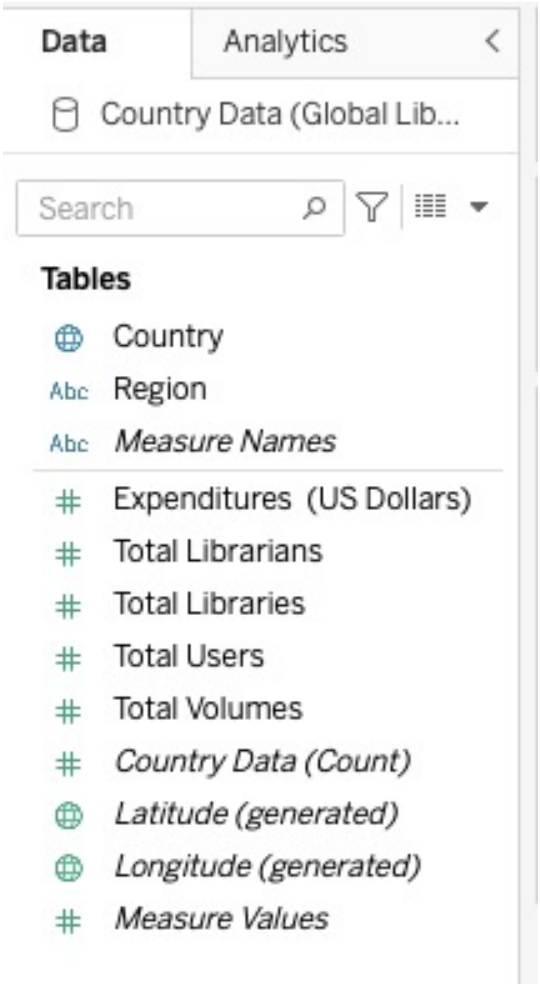
Spreadsheets	Relational Databases	Cloud Data	Other Sources
• Excel or csv file	• MySQL or Oracle	• AWS or Microsoft Azure	• Spatial Files or R

Data Field

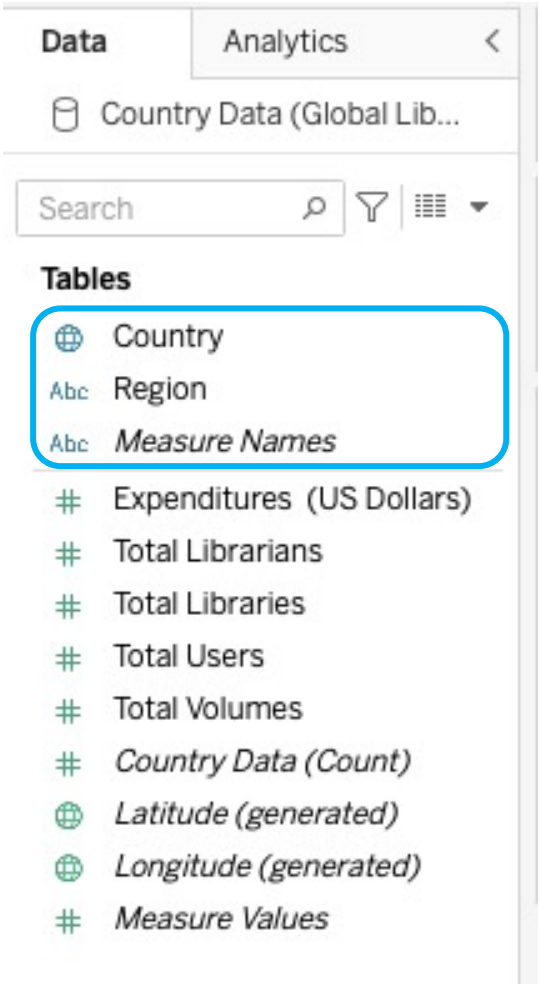
A field, also known as a column, is a single piece of information from a record in a data set.

- 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

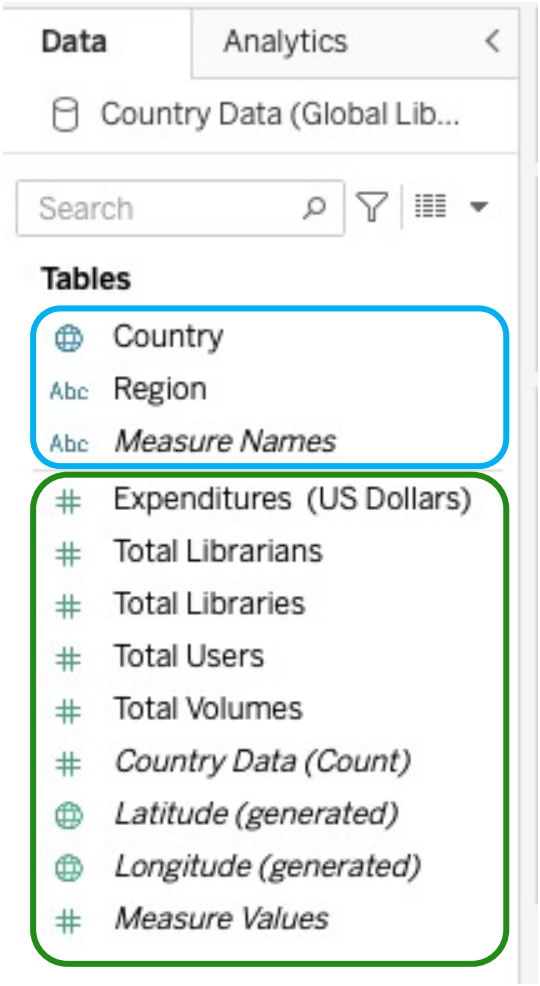

Data
Analytics
Country Data (Global Lib...
Search
Tables
Country
Region
Measure Names
Expenditures (US Dollars)
Total Librarians
Total Libraries
Total Users
Total Volumes
Country Data (Count)
Latitude (generated)
Longitude (generated)
Measure Values

Data Field



Dimensions

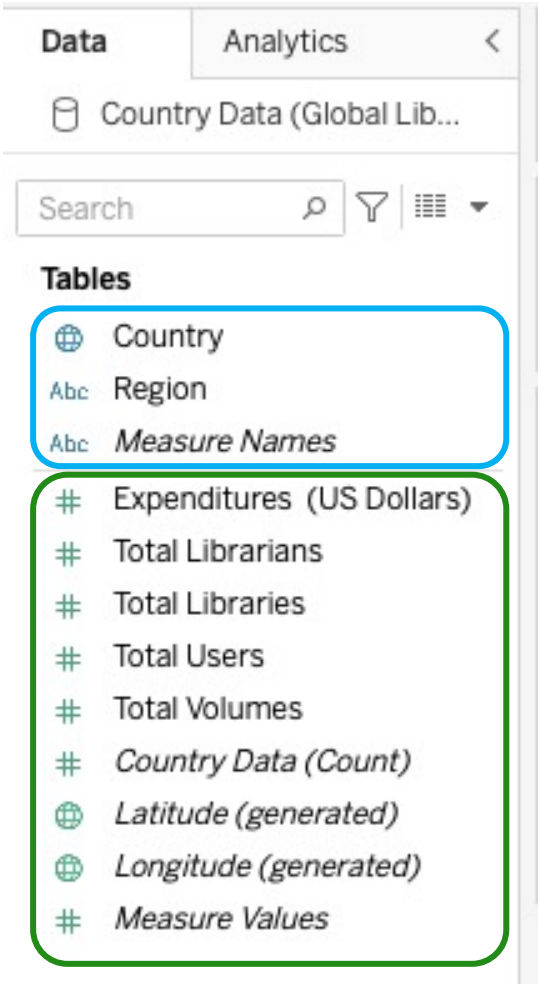
Data Field



Dimensions

Measures

Data Field

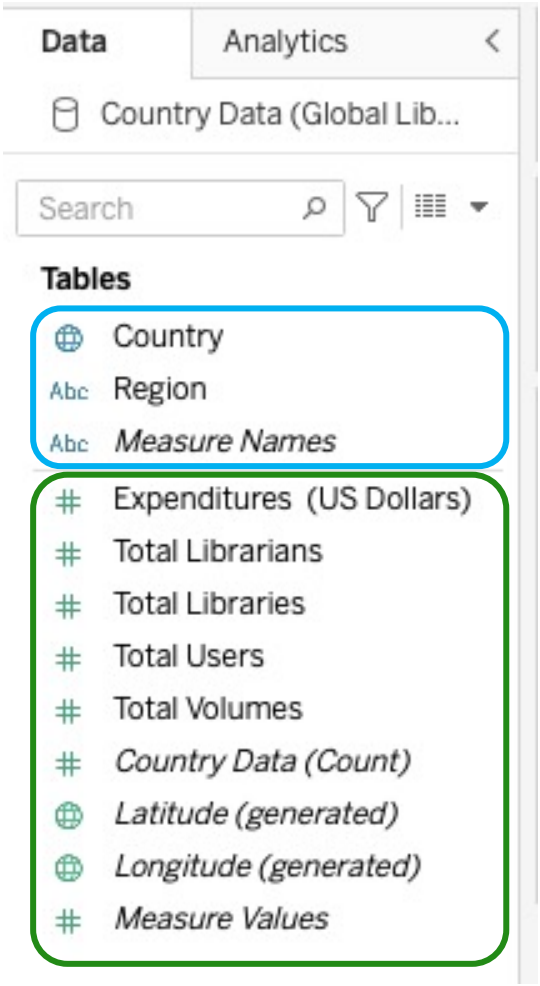


Dimensions

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

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.

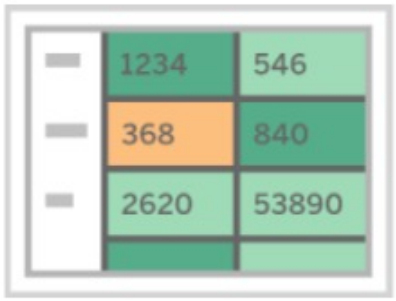
Bar — Compare data across categories.

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

Heat Map — Show the relationship between two factors.

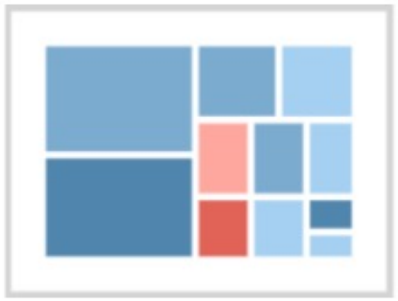
Examples: Segment analysis of target market, or sales leads by individual rep.

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, performan ce 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

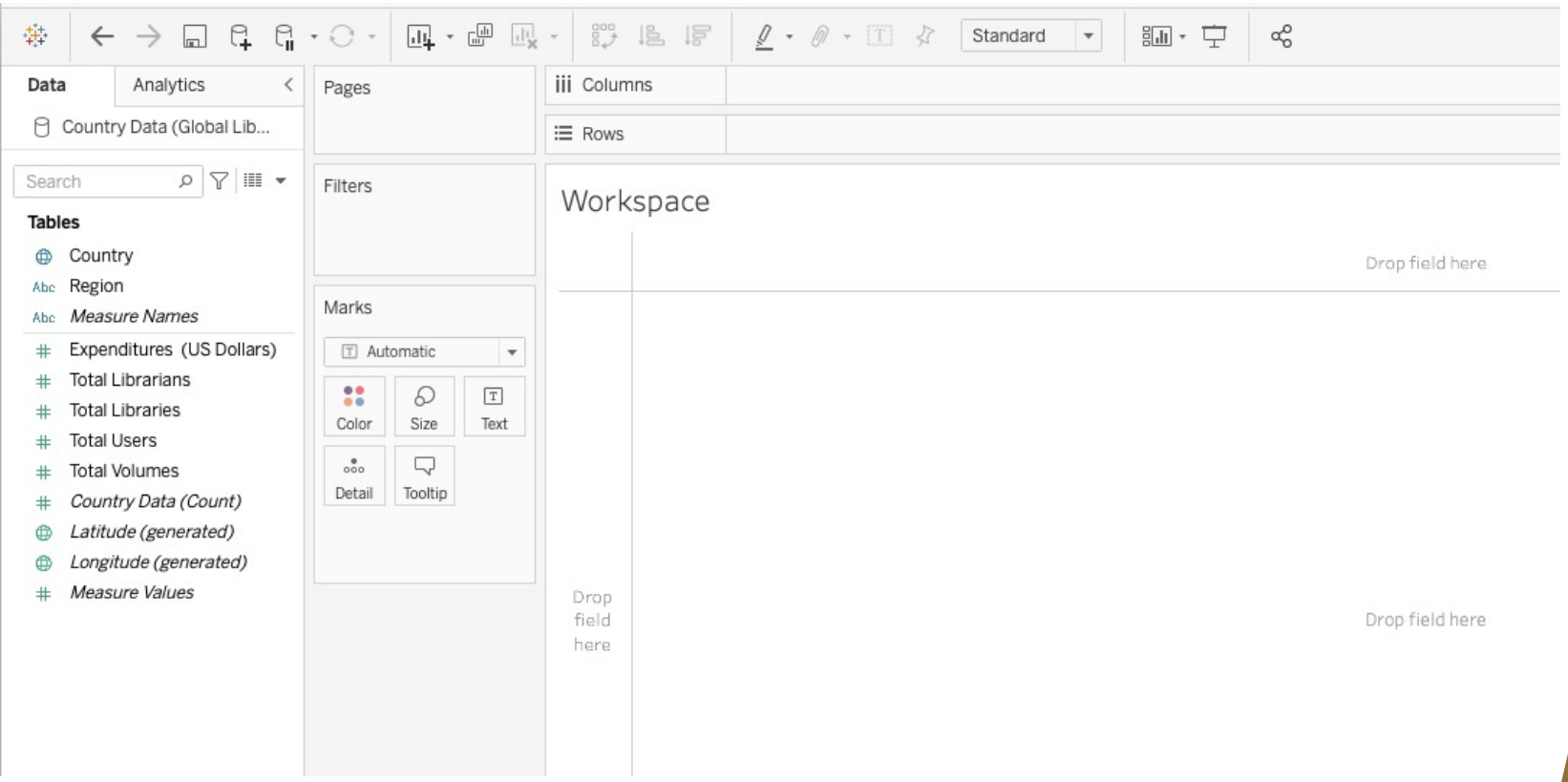


Tableau Workspace

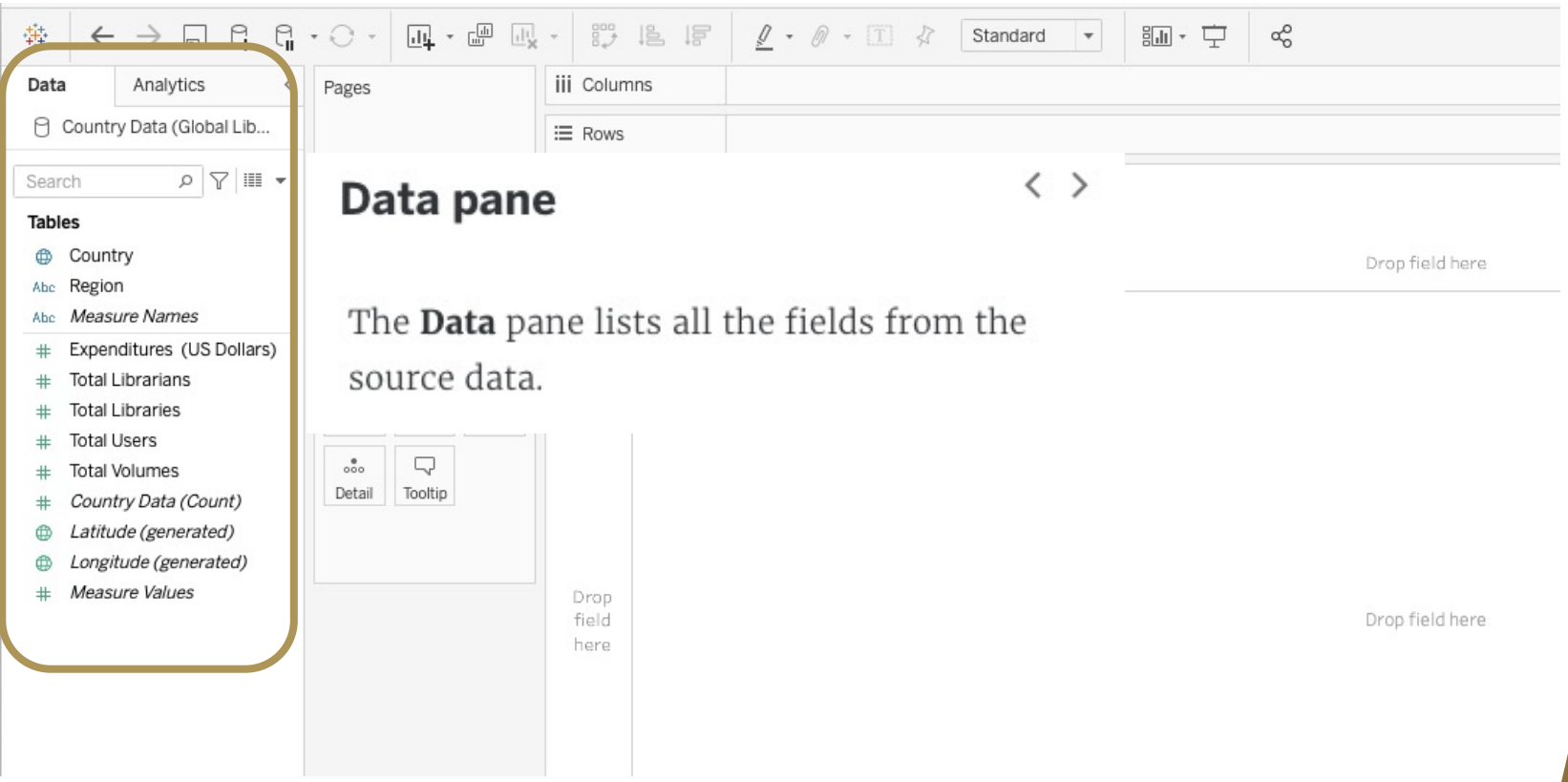


Tableau Workspace

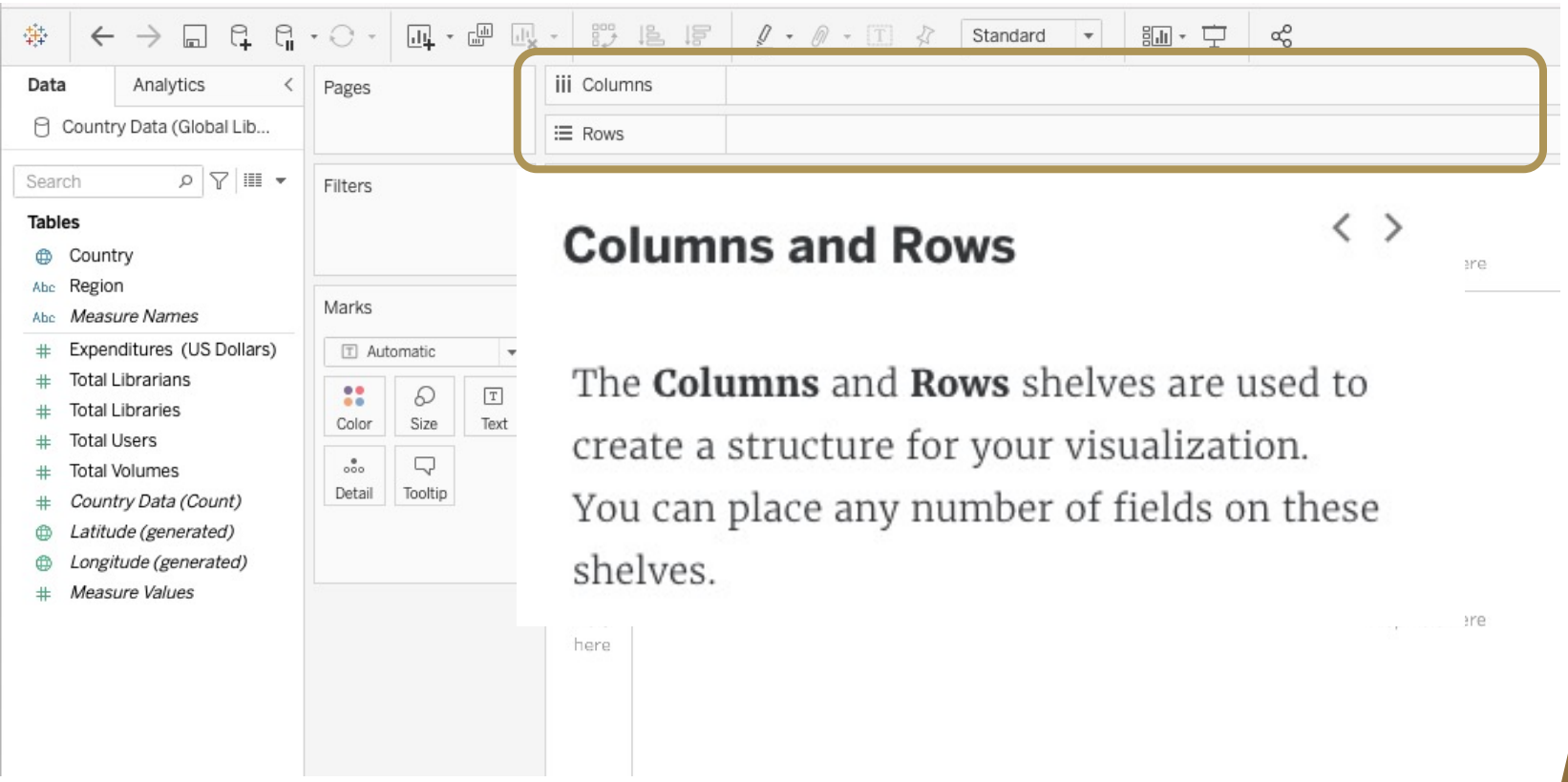


Tableau Workspace

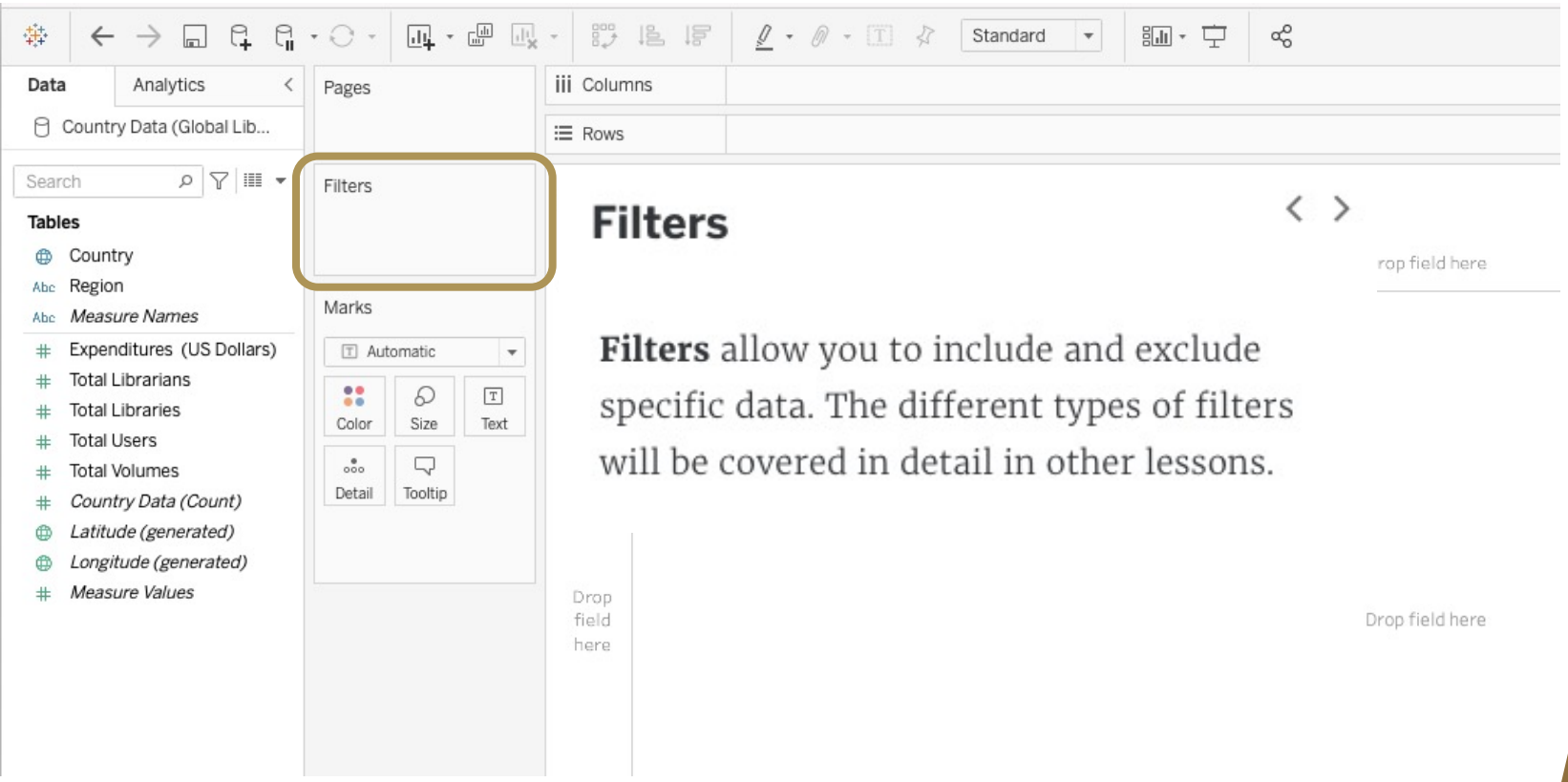


Tableau Workspace

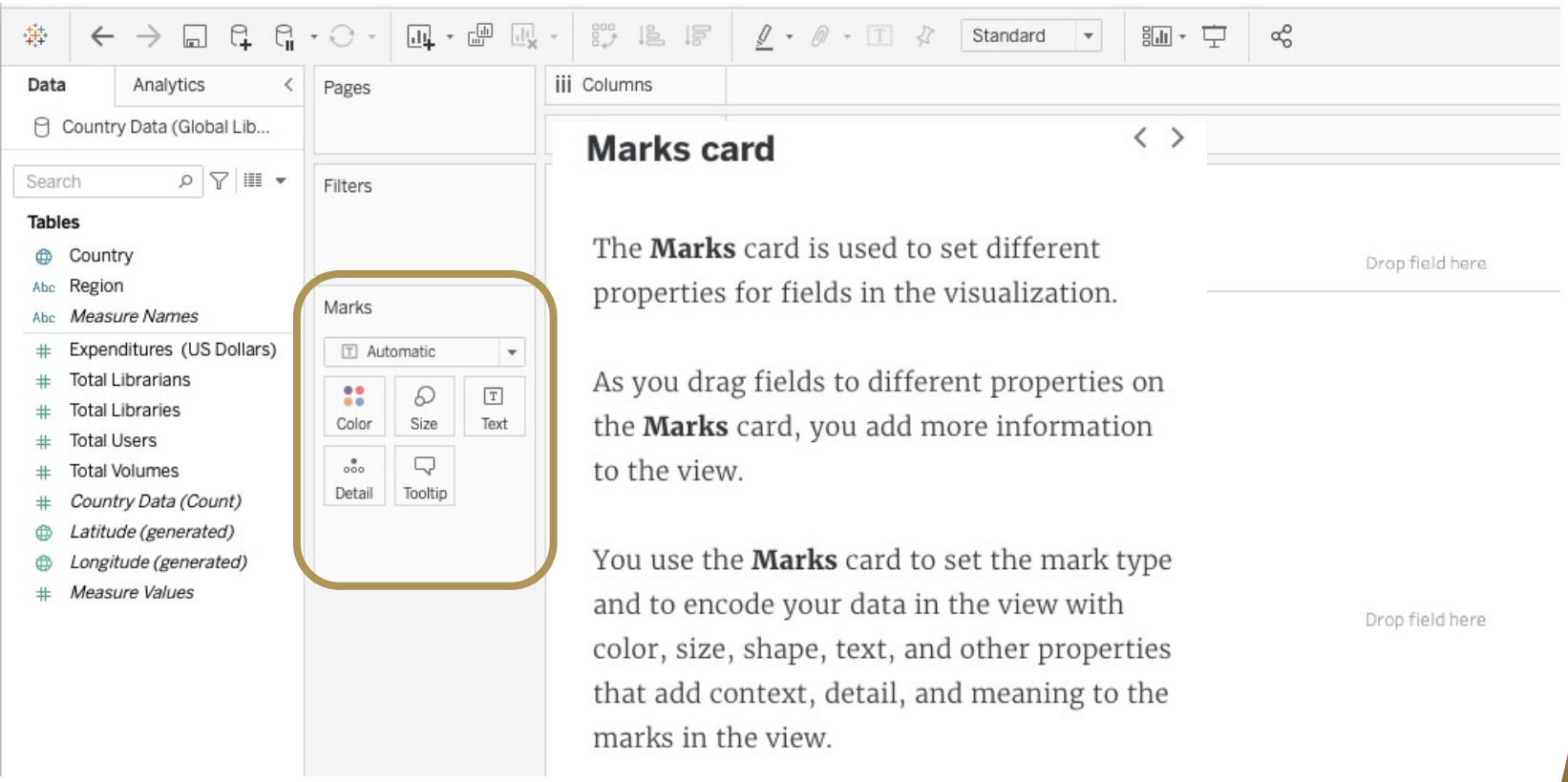
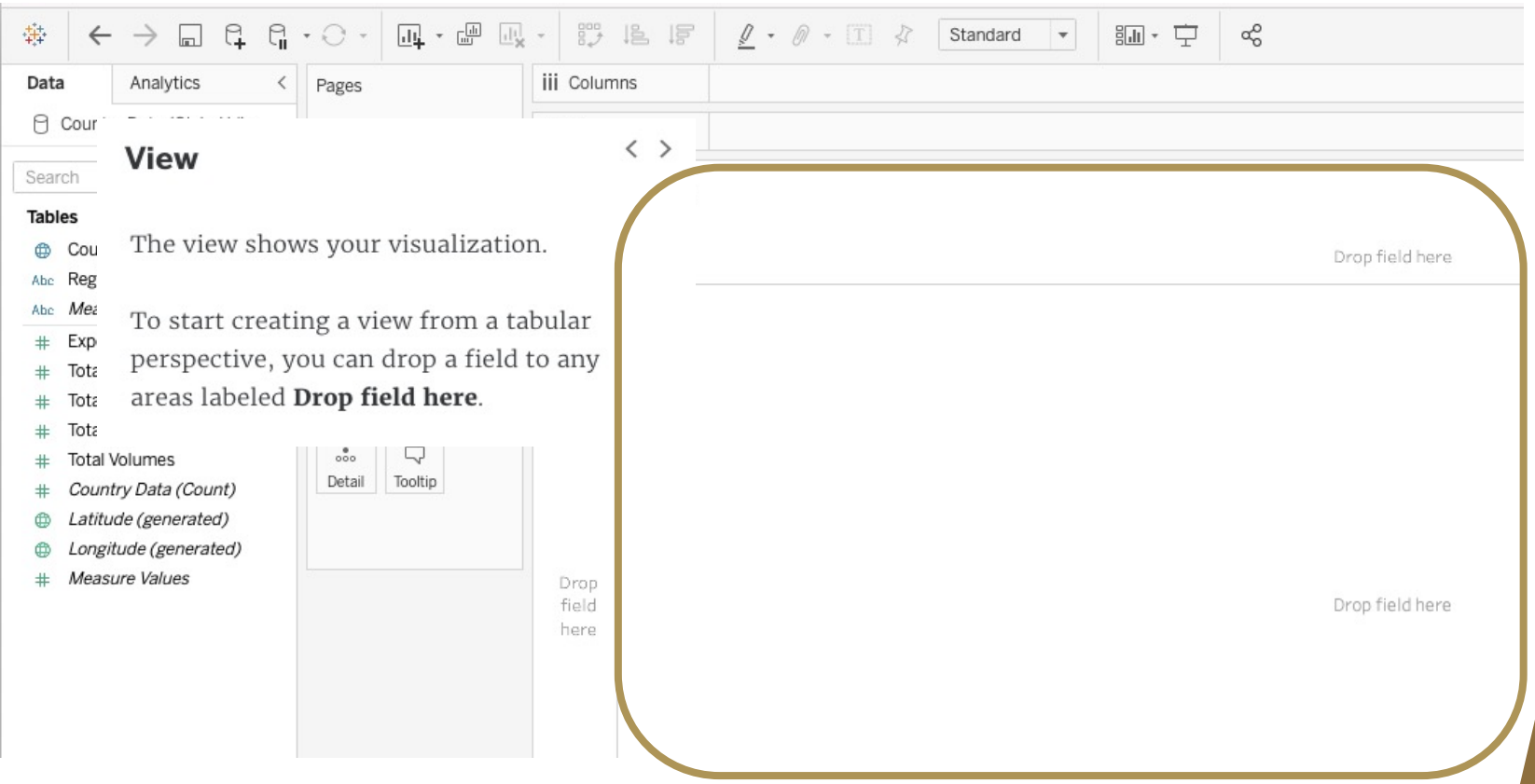


Tableau Workspace

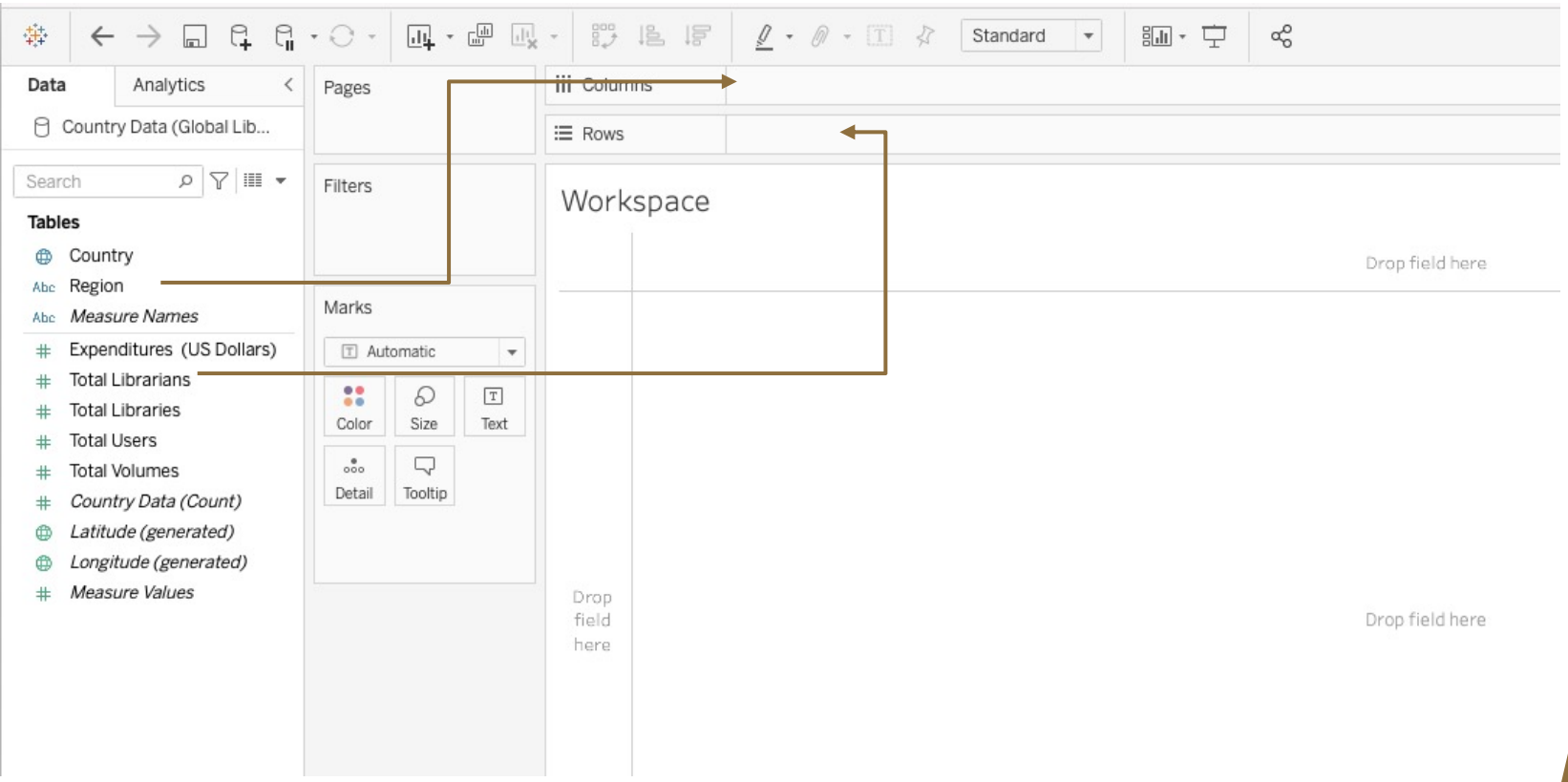


View

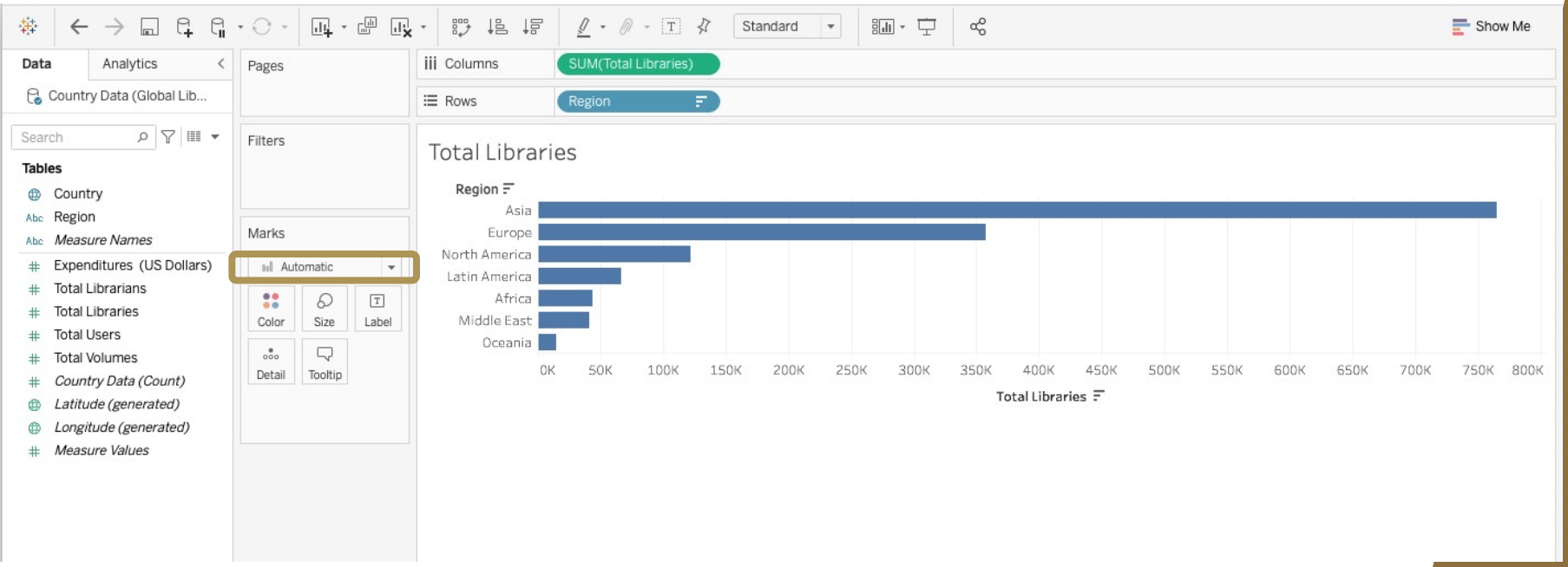
The view shows your visualization.

To start creating a view from a tabular perspective, you can drop a field to any areas labeled **Drop field here**.

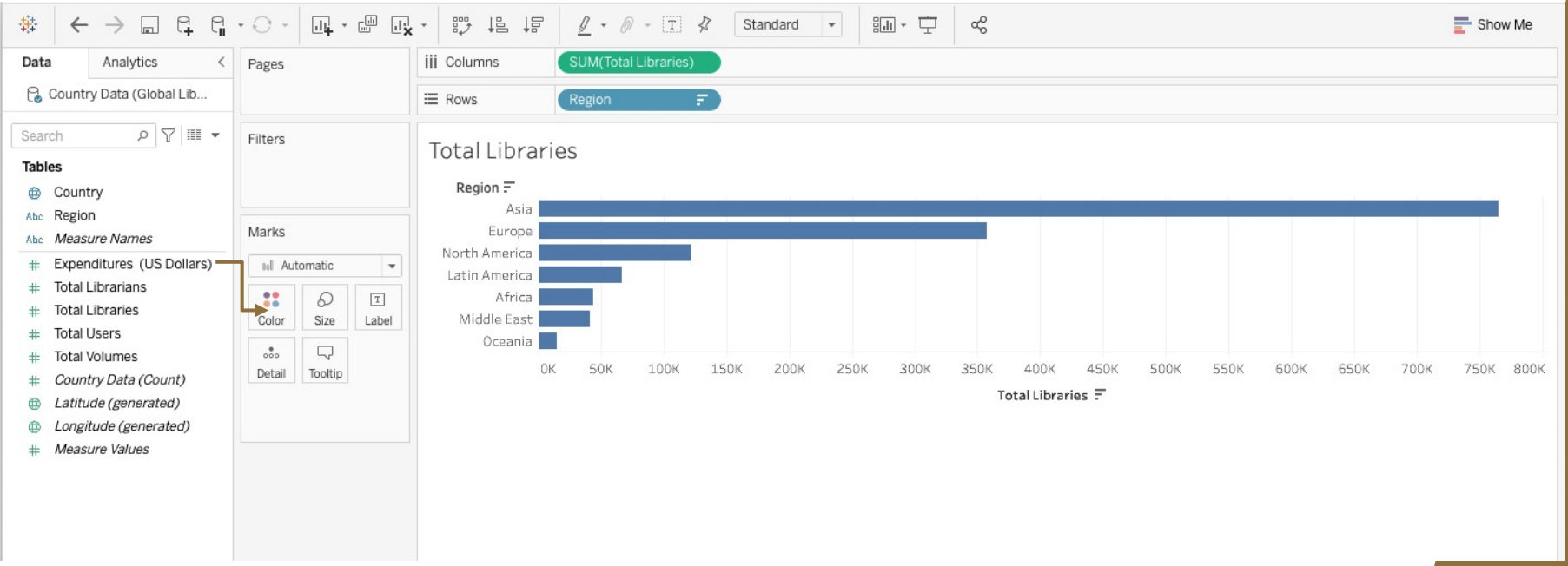
Visualization



Visualization

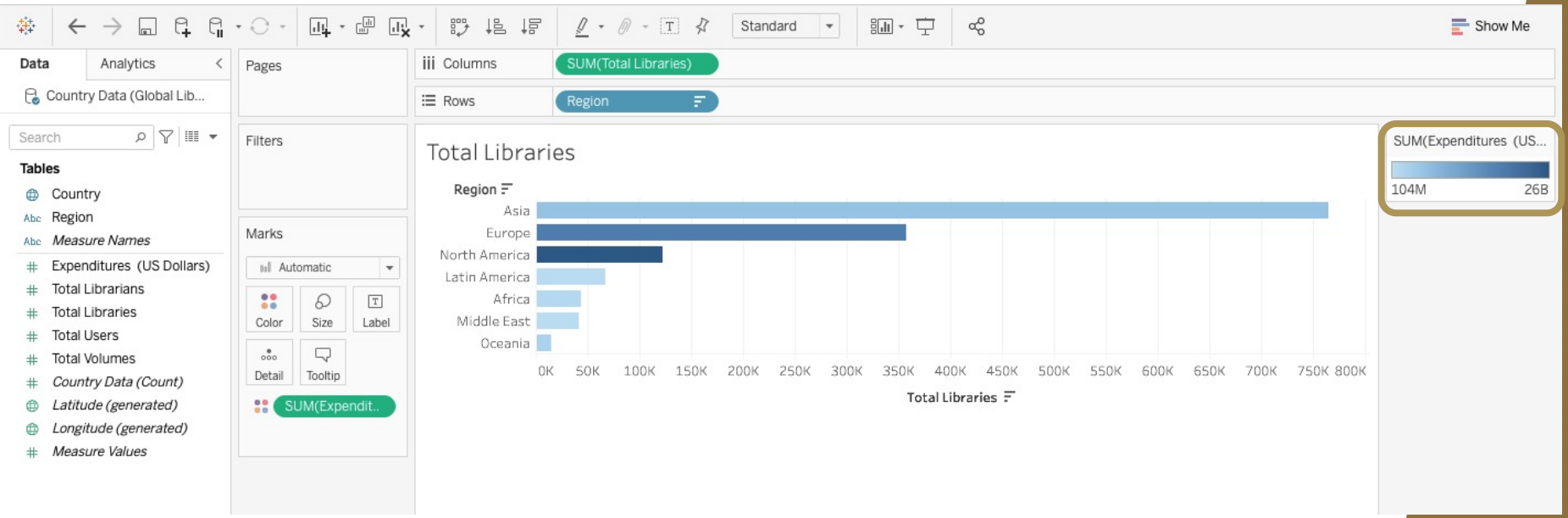


Visualization

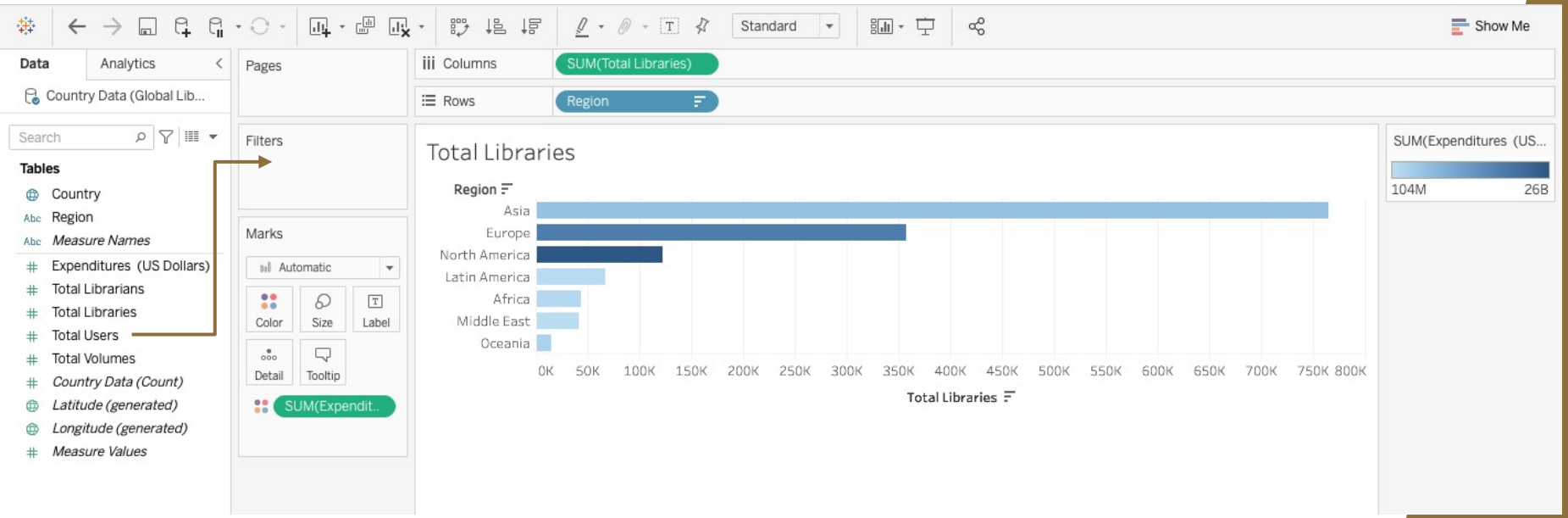


PURDUE UNIVERSITY | Libraries and School of Information Studies

Visualization

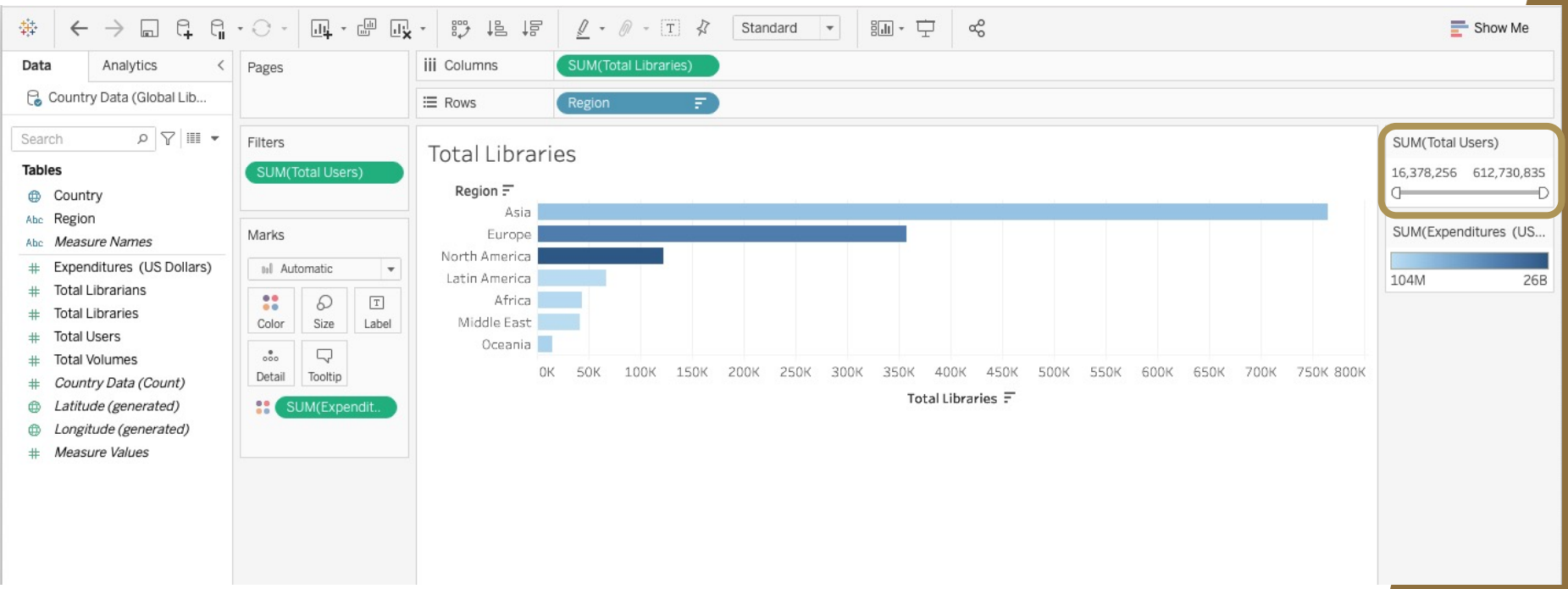


Visualization



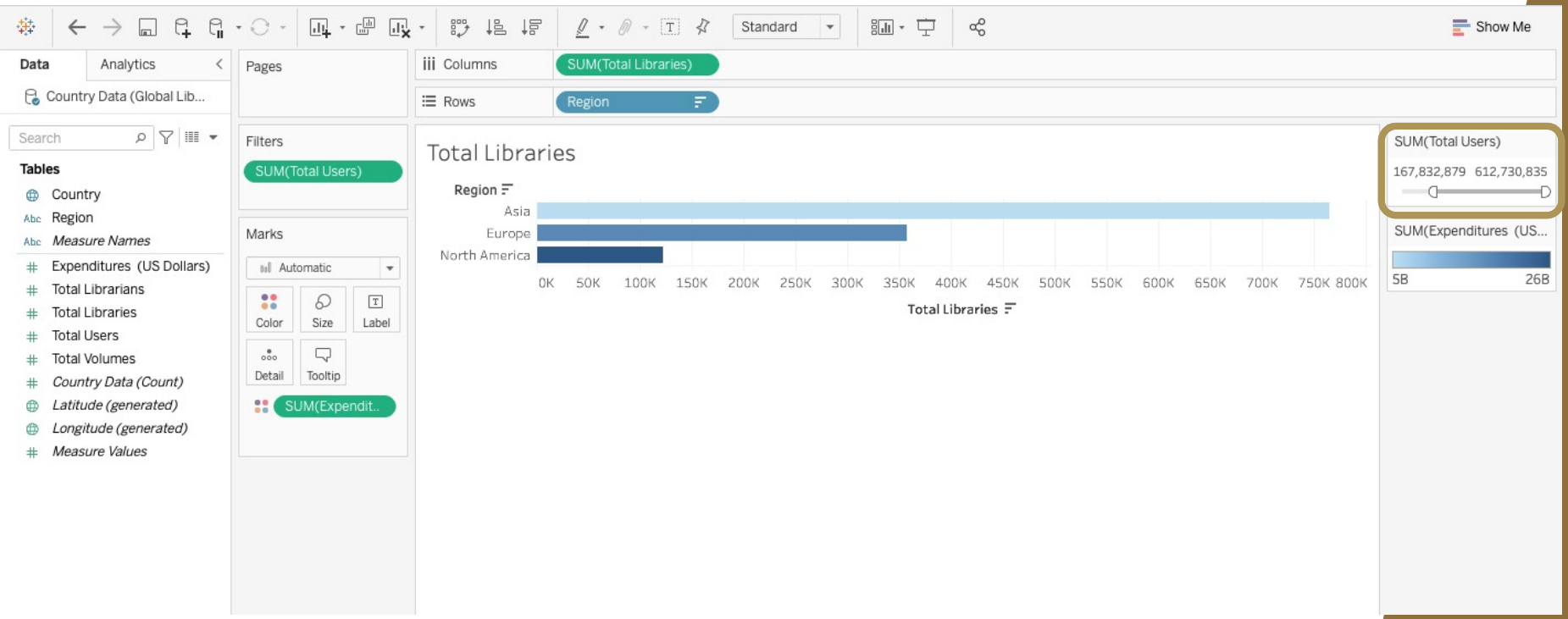
PURDUE UNIVERSITY | Libraries and School of Information Studies

Visualization



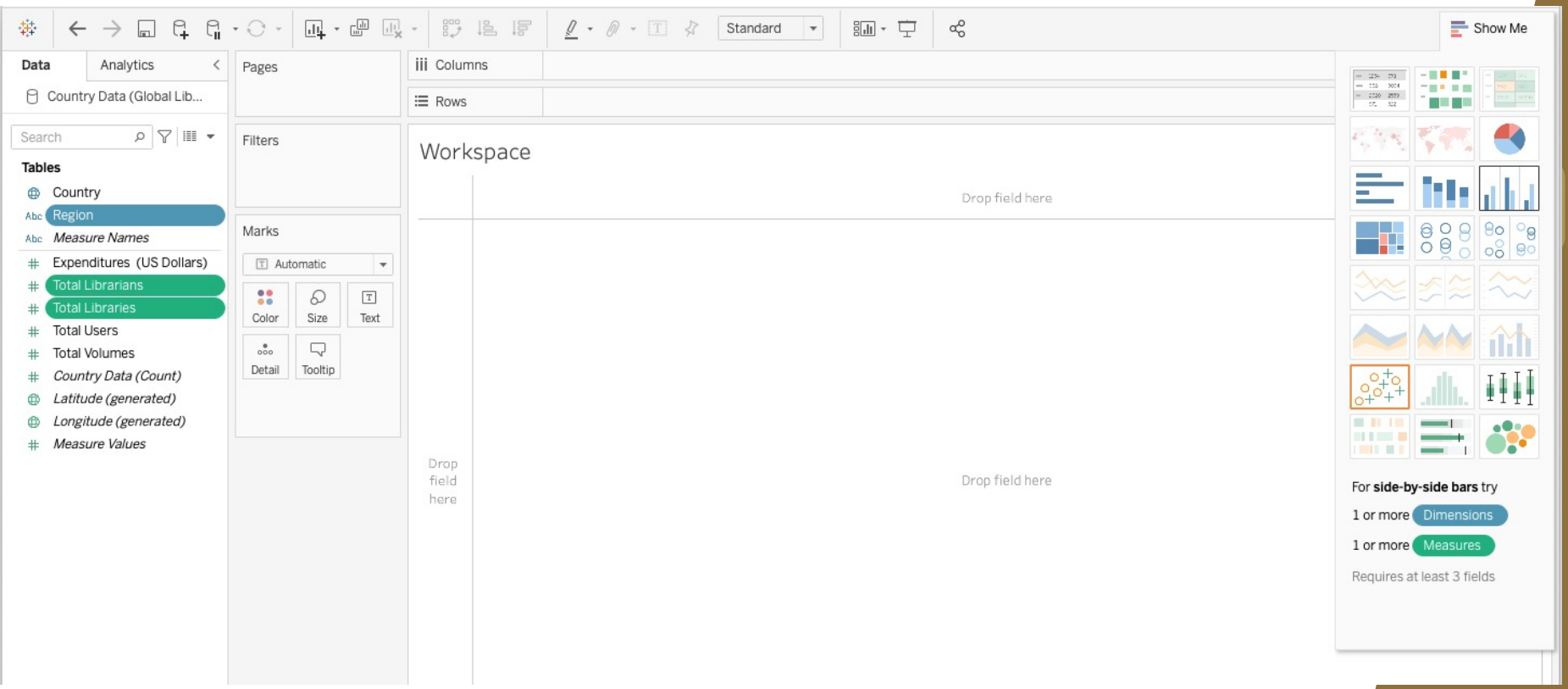
PURDUE UNIVERSITY | Libraries and School of Information Studies

Visualization



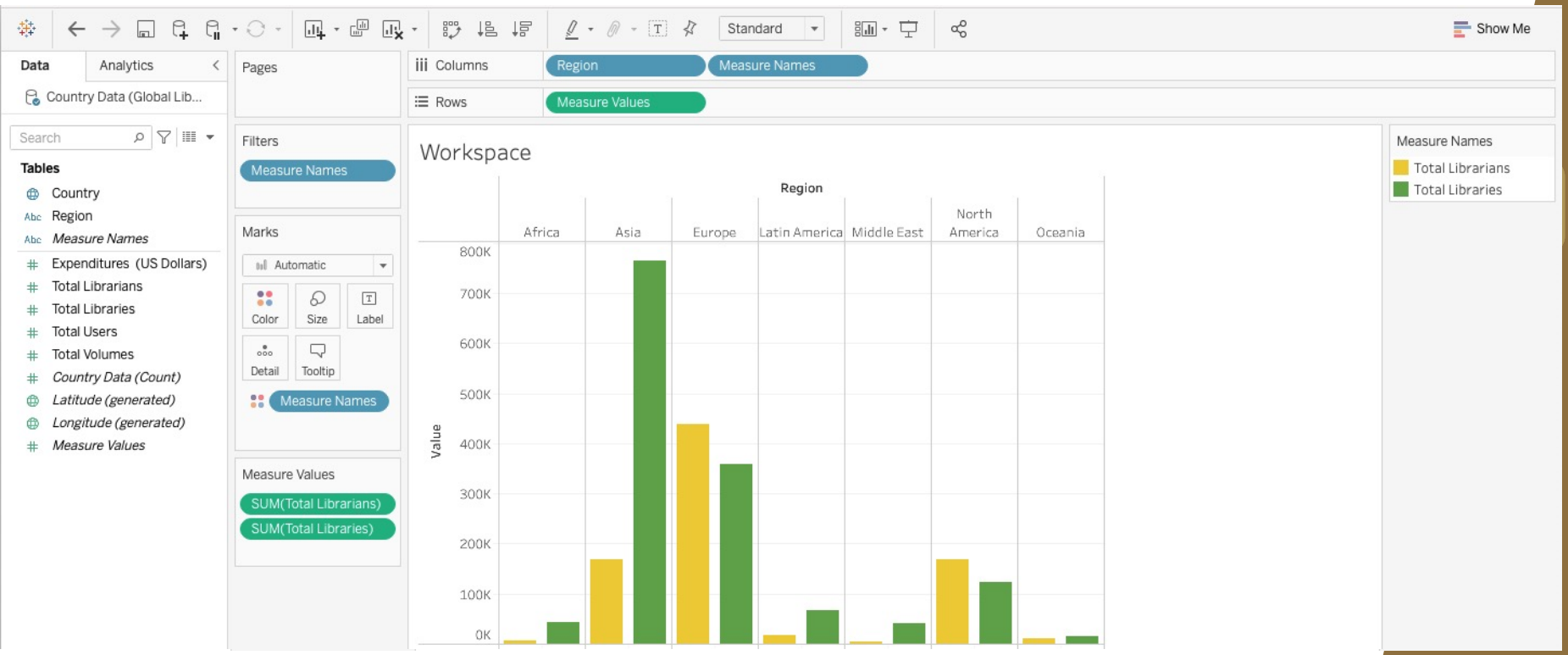
PURDUE UNIVERSITY | Libraries and School of Information Studies

Visualization



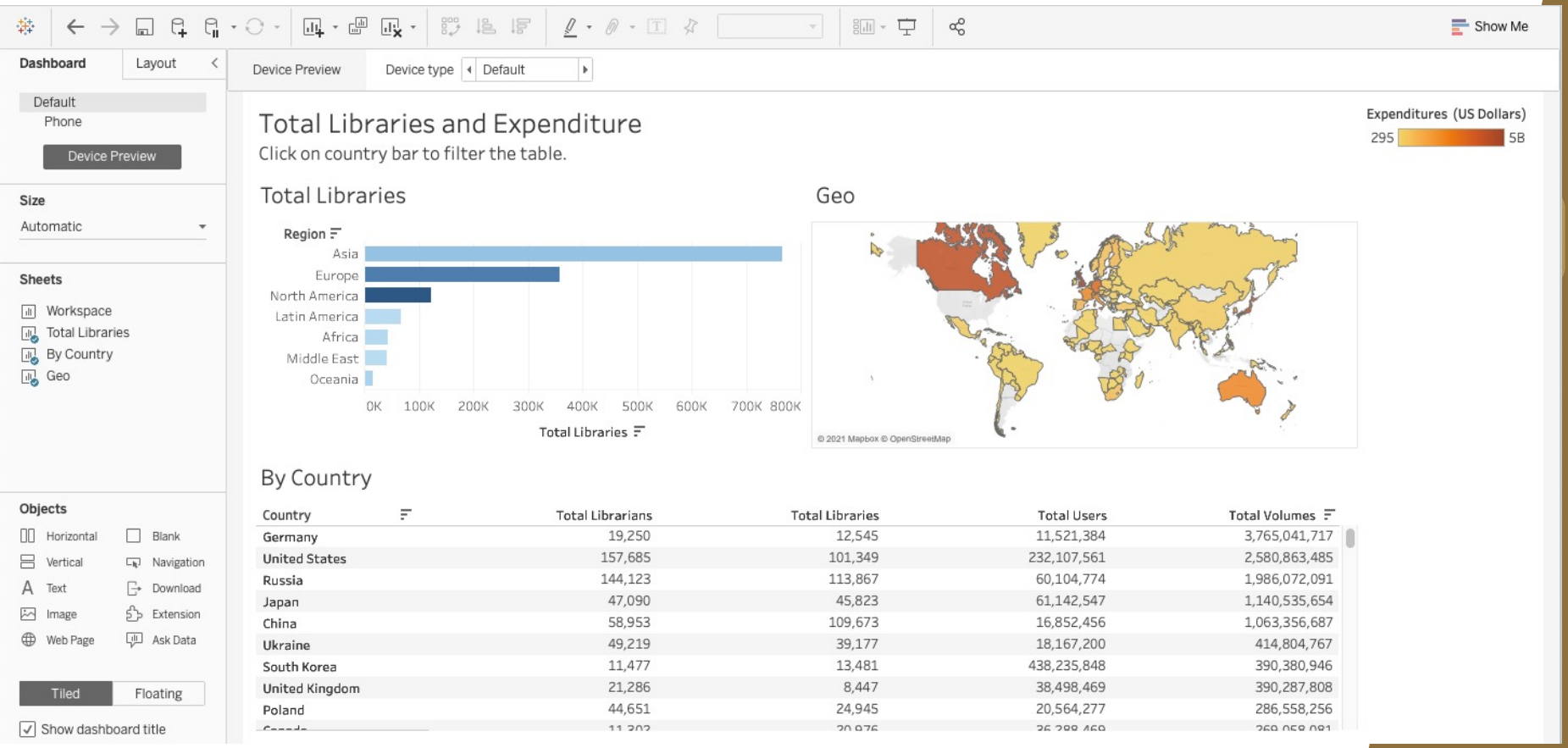
PURDUE UNIVERSITY | Libraries and School of Information Studies

Visualization



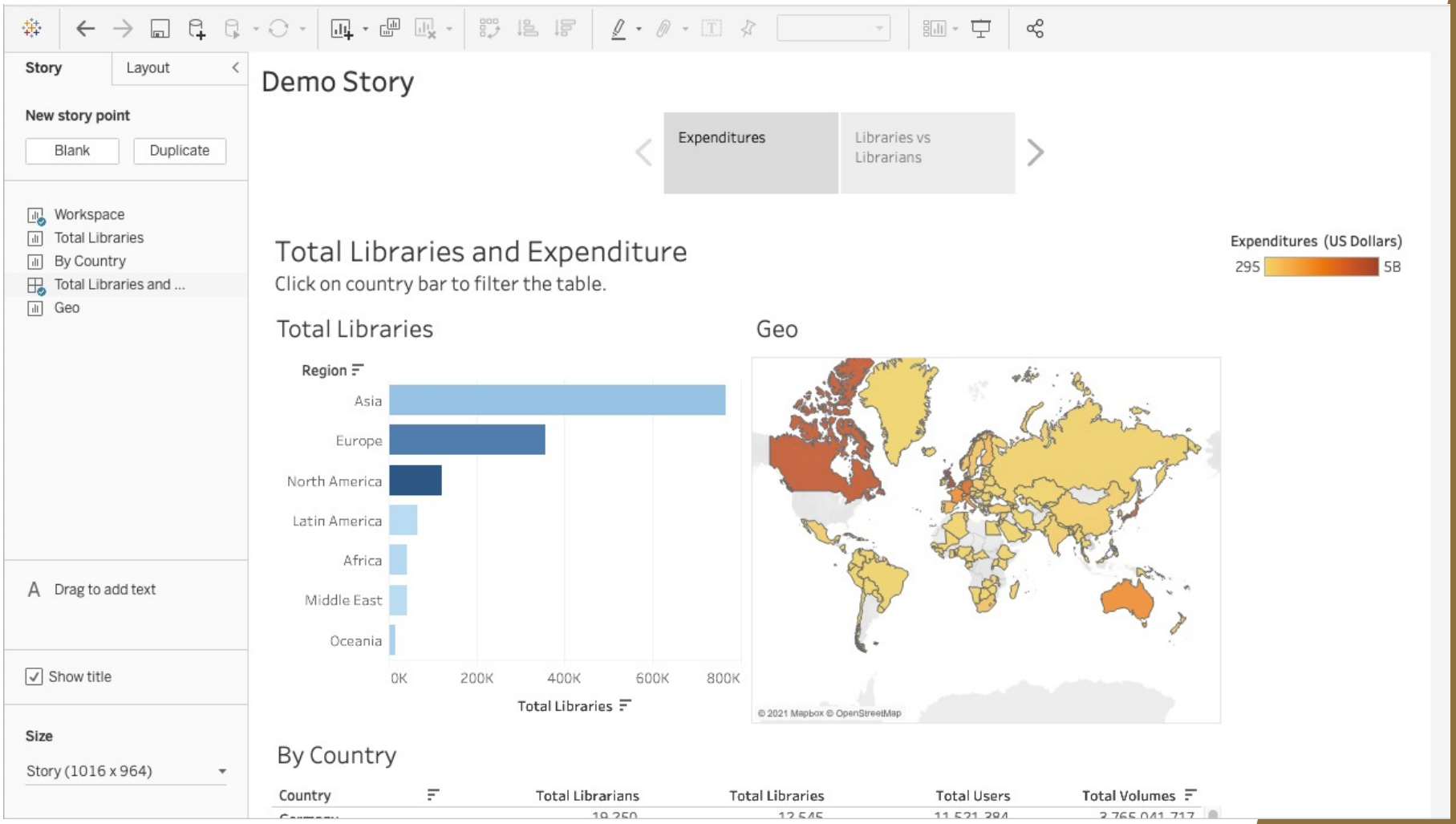
PURDUE UNIVERSITY | Libraries and School of Information Studies

Dashboard

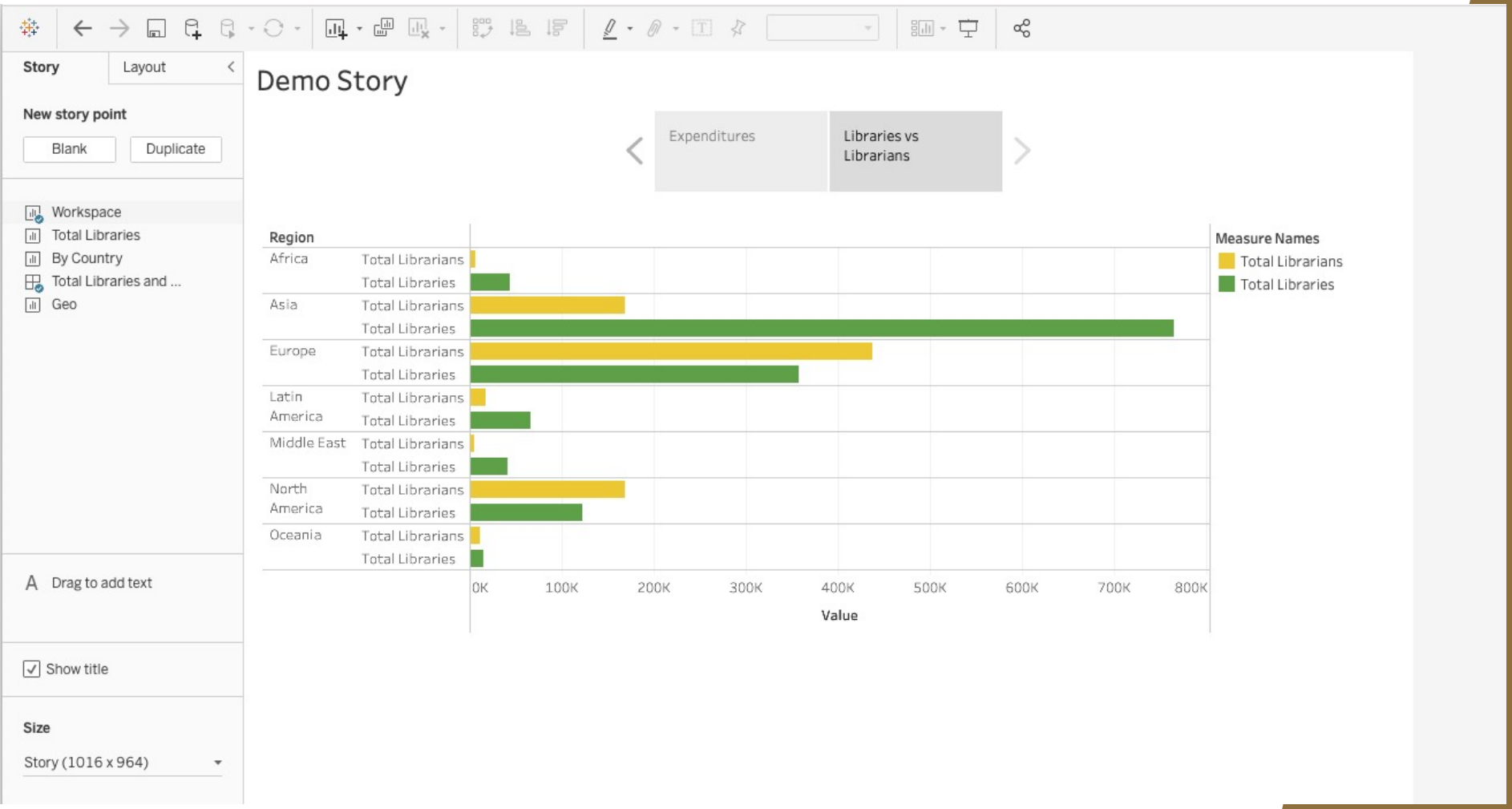


PURDUE UNIVERSITY | Libraries and School of Information Studies

Story


Story
Layout
New story point
Blank
Duplicate
Workspace
Total Libraries
By Country
Total Libraries and ...
Geo
Drag to add text
Show title
Size
Story (1016 x 964)
Demo Story
Expenditures
Libraries vs Librarians
Total Libraries and Expenditure
Click on country bar to filter the table.
Expenditures (US Dollars)
295
5B
Total Libraries
Region
Asia
Europe
North America
Latin America
Africa
Middle East
Oceania
0K
200K
400K
600K
800K
Total Libraries
Geo
© 2021 Mapbox © OpenStreetMap
By Country
Country
Total Librarians
Total Libraries
Total Users
Total Volumes

Story


Story
Layout
New story point
Blank
Duplicate
Workspace
Total Libraries
By Country
Total Libraries and ...
Geo
Drag to add text
Show title
Size
Story (1016 x 964)
Demo Story
Expenditures
Libraries vs Librarians
Region
Africa
Asia
Europe
Latin America
Middle East
North America
Oceania
Total Librarians
Total Libraries
0K
100K
200K
300K
400K
500K
600K
700K
800K
Value
Measure Names
Total Librarians
Total Libraries

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

PURDUE UNIVERSITY | Libraries and School of Information Studies