Getting Started with Tableau

Hello!

I am Nopphiphat Suraminitkul (Nopi)

Graduate student in Computer Science

Statistical consultant

Overview

- ❑ Learning Objectives
- ❑ Introduction
- ❑ Why Tableau?
- ❑ Workshop Materials
- ❑ Data Visualization Process
- ❑ Practice

1. Learning Objectives

Learning Objectives

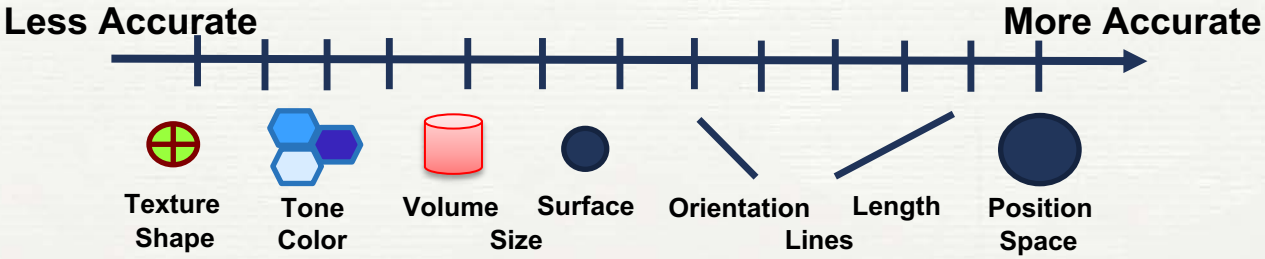
You should leave with:

- Understanding on ways to handle different data types
- Methods for building good visualizations
- Hands-on experience with Tableau Public

2. Introduction to Data Visualization

Introduction

Data visualization is the pictorial representation of dataset(s) and information using visual elements to give us a framework to make sense of masses of information, see and understand trends, detect hidden patterns, identify outliers, and develop new insights in data, which can then be used to make data driven decision.

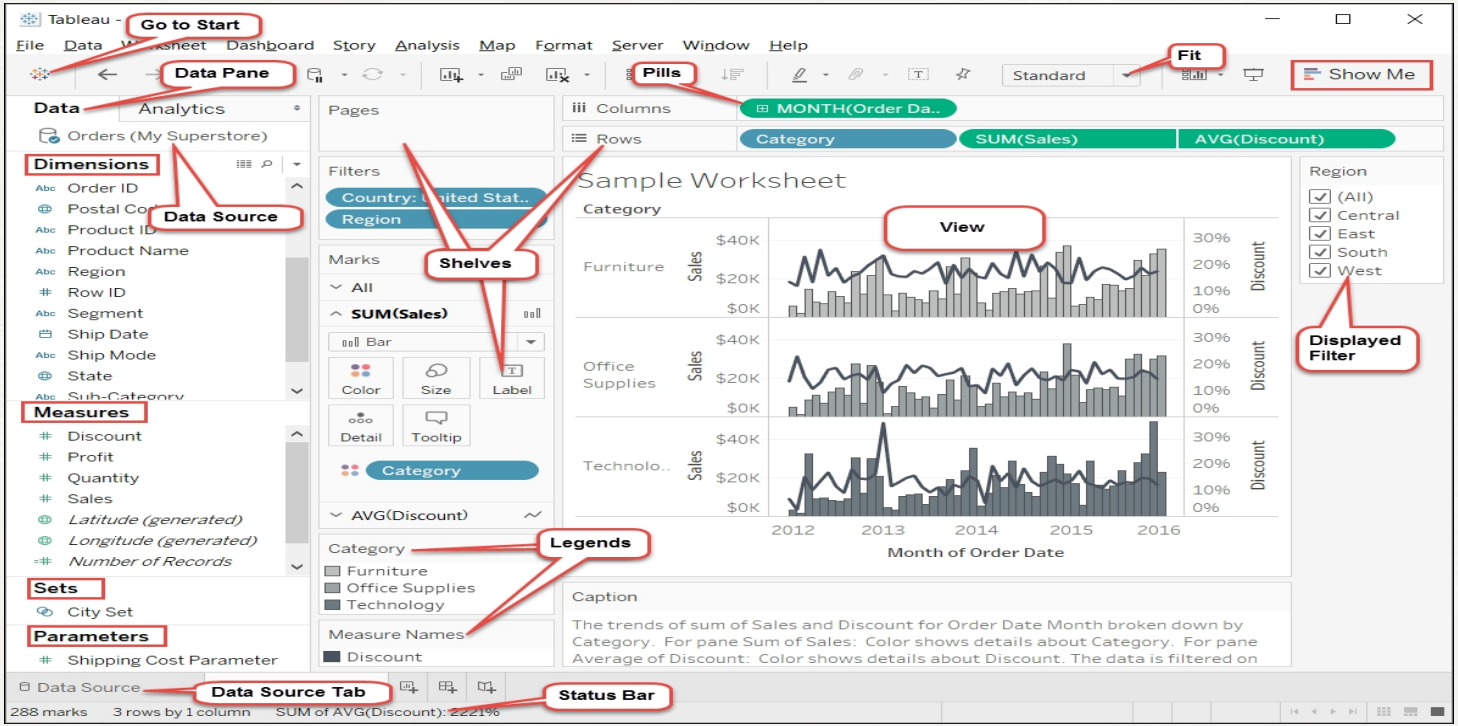


Why Tableau?

Advantages:

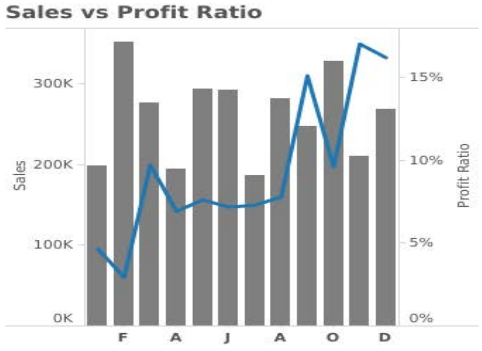
- Superior Visualization Quality with Powerful Computation
- Multiple Data Source Connections
- Easy to Learn and Use
- User-friendly Interface

Tableau Interface

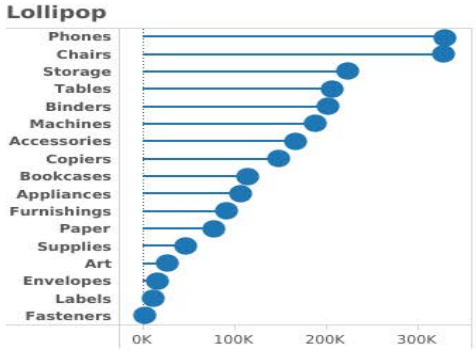


Types of Visualization

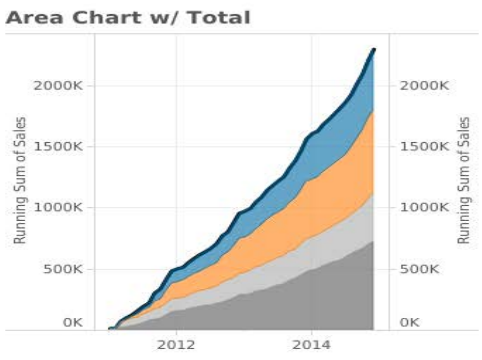
Vertical Bar Chart

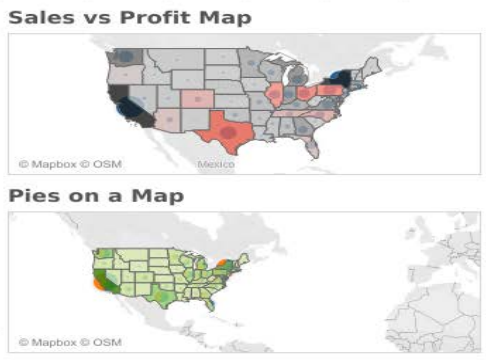


Horizontal Bar Chart



Time Series





Stacked Area Chart

Scatter Plot

Density Map

3. Workshop Materials

Workshop Material: Tableau Public

Installation of Tableau Public

- Access https://public.tableau.com/en-us/s/download
- Enter your email to download the software

(TIP) Making your work less public*

- Register and sign in to the Tableau website
- Profile icon on top-right -> click “Settings”
- Tick:
  - “Allow viz data to be downloaded”
  - “Set my vizzes to hidden”

Workshop Material: Datasets

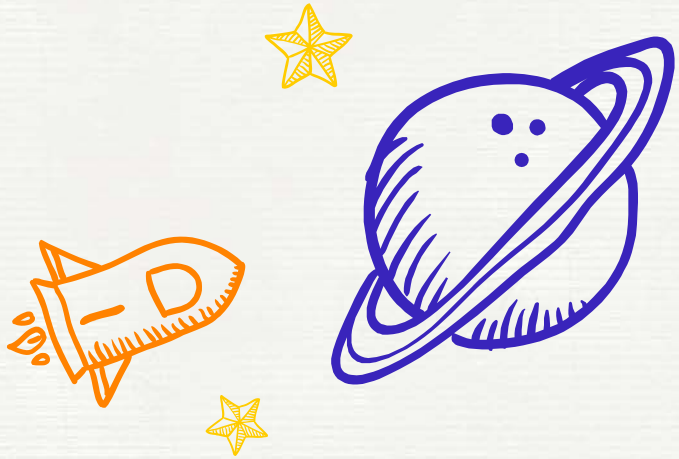
Access: https://www.kaggle.com/shivamb/netflix-shows

- ✗ Dataset: Netflix Movies and TV Shows
- ✗ Click “Download” and unzip the file to your desktop

4. Data Visualization

Data Visualization Process

- Loading your data
- Knowing your data: understand data definition
- Determining what to visualize: what are you looking for
- Pre-processing your data: ensure data is in correct format
- Visualizing your data
- Exporting your data



END

Bring the attention of your audience over a key concept using icons or illustrations

References

- Garrett, J., Blaine, A., & Schechter, G. (2020, January 31). Getting Started with Data Visualization. Retrieved from osf.io/s4w5b
- https://www.youtube.com/playlist?list=PLWPirh4EWFpGXTBu8IdLZGJCUeTMBpJFK

Step 1: Loading your data

- ✗ Open Tableau Public
- ✗ Click on “File” -> “Open…”
- ✗ Navigate to the location of the file (e.g. Desktop)
- ✗ “Select” the file

(Optional) Under “Files” in the left column,

- ✗ [Tick] Cleaned with Data Interpreter, which is designed to help take care of cleaning and transforming the data into the proper format for analysis in Tableau.

Step 2: Knowing your data

✗ What you do know / What is being measured ?

- **show_id**: Unique ID for every movie or tv show
- **type**: Identifier – a movie or tv show
- **title**: Title of the movie or tv show
- **country**: Country where the movie or show was produced
- **date_added**: Date it was added on Netflix
- **release_year**: Actual release year of the movie or show
- **rating**: TV rating of the movie / show
- **duration**: Total duration – in minutes or number of seasons
- **listed_in**: Genre

Step 2: Knowing your data

✗ What data types are we dealing with ?

Data Type	Icon
String values (Text)	Abc
Integer values (Numbers)	
Date values (DD/MM/YYYY or MM/DD/YYYY)	
Date & Time values	#
Boolean values (True or False; relational)	T|F
Geographic values (Region, Postal code etc.)	
Cluster group or mixed values	

Step 3: Determining what to visualize

- ✗ **What do you not know ?**
  - Quick overview of data

- ✗ **Can the information be represented without visualization?**
  - *Not all data needs to be visualized*

- ✗ **Does audience needs to see pattern to understand data?**
  - Preferences over time

Step 4: Modification

- Other library workshop

Step 5: Visualization

- Content Type on Netflix
- Content added over year by Type

Pie: Content Type on Netflix

- ✗ Drag “Type” under Dimensions to “Rows”
- ✗ Drag “Measure Values” to “Text” under Marks
- ✗ Select display “Entire View”
- ✗ From the Marks card drop-down menu, select “Pie”
- ✗ Drag “Type” to “Label” under Marks
- ✗ Drag “Measure Values” to “Label” under Marks
- ✗ Click on the inverted triangle icon on the green pill under “Measure Values” and click “Quick Table Calculation” and select “Percent of the total”

Line: Content added over year

- ✗ Drag “Date Added” under Dimensions to “Columns”
- ✗ Drag “Type” under Dimensions to “Rows”
- ✗ Drag “Number of Records” to “Text”
- ✗ From the Marks card drop-down menu, select “Line”
- ✗ Drag “Type” to “Color”
- ✗ Drag “Number of Records” to “Label”
- ✗ Click on “Label” and click “Select” under “Marks to Label”
- ✗ Right-click on the points you want to show label, and “Mark Label” and “Always Show”

Map: Content by Different Countries

- Drag “Country” to the view
- Drag “Number of Records” to “Color”
- Drag “Country” to “Label”
- Click on “Label” and select “Selected” under “Marks to Label”
- Right-click the country with darker color, and select “Always Show” under “Mark Label”

Bar Chart

- Simple Bar Chart (Vertical vs Horizontal): rectangular column with size as data in measures
  - Ctrl + W or Swap Icon
  - Raw Data Drill Down (Sub-Category drag on another dimension)
  - Group by Header (Select headers and Group)
- Colored Bar Chart
  - 1 Blue pill in column + 1 green pills in rows + 1 measure to color
  - Create set (Top 5) and color the bars by set
- Stacked Bar Chart
  - 1 Blue pill in column + 1 Green pill in rows + 1 dimension to color
- Leveled Bar Chart
  - Two blue pills in columns + 1 green pills in rows
- Interactive Bar Chart
  - Insert Filters

Line Chart

- Line Chart
  - A measure and a dimension are taken along the two axes of the chart area
  - The pair of values for each observation becomes a point and the joining of all these points create a line showing the variation or relationship between the dimensions and measures chosen. -> Show Caption
- Date Line Chart
  - (Date Icon) “DATE” to columns and measures to rows
- Date Line Details
  - Dimension “Category” to Color
  - Measure to Label (Show maximum)
- Continuous (GREEN) vs Discrete (BLUE)
  - Modify the type of dates
  - Which year, which month, which date has highest value (continuous)

Scatter Plots

- Scatter Chart
  - 2 Measures in columns and rows
  - For category wise, drop the category into Shape/Color (Modify size)
- Interactive Scatter Chart
  - Filters
- Drop Lines
  - Click the point + Drop line (Edit Drop line, labels “automatic”, drop when “always”)

Word Maps

- Drag text dimension to “Text”, Drag measure to “Size”, Set the Drop down to text
  Coloring, Bold
- Interactive Word Map