MSA 8020
Data Visualization

Yichen Cheng

Installation

https://www.tableau.com/academic/students to request a free 1-year license

https://www.tableau.com/en-gb/products/trial To start a 14-day trial.

Resources

Free training videos

https://www.tableau.com/learn/training/20203

Gallery

https://public.tableau.com/en-us/gallery/?tab=viz-of-the-day&type=viz-of-the-day

References

https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-connect.htm

https://www.biztory.com/blog/using-r-in-tableau

https://data-flair.training/blogs/tableau-word-cloud/

https://help.tableau.com/current/pro/desktop/en-us/default.htm

Superstore data

Import data

Drag and drop for a first look

Deal with map data

Dashboard

Story

Step 0: Import data

Open Tableau:

- Connect to a file -> Microsoft Excel

- Data Source Tab shows the available sheets -> double click on "Orders"

- Go to Worksheet: Sheet 1

data types

Two major data types:

- Dimension: qualitative, such as a name or date.

- Measure: quantitative numerical data.

Click on sheet1, we can see a list of the variables

- check whether the data type is correct.

- If not, can manually change it.

Step 1: Creating a view

Several element that we can play with

- Columns and rows: specify the major structure

- Filters: filter to focus on part of the data

- Markers: use colors, shapes for extra information

Generate a simple chart

- Sales by order year & category & subcategory

- Columns: Order date, category, subcategory

- Rows: Sales

Add filters and colors

Show profit by colors

- Drag profit to Marks -> Color

Add filters to filter out the interesting subset of the data

- Add order date and sub-category to filters

- Click on sub-category, select "Show filter"

These are two important steps that helps use to finding insights

Add region to Rows

Exercise

Why are these three subcategory not profitable? Because too many discount?

- Repeat the same analysis but use discount instead of profit as the color

- Filter the unprofitable subcategories.

Save a workspace and create a new one

Rename sheet 1 to "Sales by Product/Region"

Duplicate and focus on south region, rename it to "Sales in the South"

- Click on Year(Order Date) -> Show Filter

Step 3: Map view

Create a map that shows sales/profits in south states

- Drag State to Marks -> Detail

- Drag Sales to Marks -> Color

- Filter only South states using Region

- Drag Profit to Marks -> Color

Name it "Profit map"

Step 4: negative profit bar chart

Duplicate the Profit Map Worksheet and name that profit bar chart.

- Click on Show Me and select the horizontal bar chart

- Drag City to Rows

- Filter city with the most negative profit

  – Drag City to Filters

  – Edit Filter -> Bottom 5

- Drill down to the sub-category and identify the trouble-maker

- Verify that on the Profit Map

Exercise

Repeat our analysis for the West region

- California makes a lot profit while Colorado has negative profit

- Figure out the reason

- Play around with the "Show Me" option on the upper right corner.

Dashboard

Help us to combine multiple views in the same page and can view interactively

- Create an empty Dashboard

- Drag Profit map and Sales in the South to it

- Click on Profit Map -> Use as filter

- Remove duplicate card

- Apply one filter to all in the dashboard

    – Apply to Worksheets -> Selected Worksheets -> All on dashboard

https://help.tableau.com/current/guides/get-started-tutorial/en-us/get-started-tutorial-build.htm

https://data-flair.training/blogs/tableau-dashboard/

Story

Put a series of visualization to show the storyline

Emphasis your points/findings

- Sales in the South -> Sales and profit by year

- Duplicate and show only for machine -> Machine sales and profit by year

- Dashboard -> Underperforming items in the South

- Duplicate and show only NC -> Profit in NC, 2014-2017

- Create one page for each year

Save a workbook

Not including the data sources

Including the data sources

Publish to Tableau Public

Rserve – connecting with R

Set up in R

- Rserve library

Set up in Tableau

- Help->Settings and performance->Manage External Service Connection

- Test Connection

Calculation

- SCRIPT_REAL(".arg1/.arg2",

- SUM([Sales]), SUM([Quantity])

-)

https://www.biztory.com/blog/using-r-in-tableau