

INTRODUCTION TO DATA VISUALIZATION

Jie Jane Jian

20 Nov, 2018



OUTLINE

- What is Data Visualization?
- Why Data Visualization?
- Application of Visualization
- How to do Data Visualization?
- Tableau with Example
- D3.JS with Example

WHAT IS DATA VISUALIZATION?

How Many “V”?

MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFPSTWYARN
GFPSVCEILQGKMFPSNDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

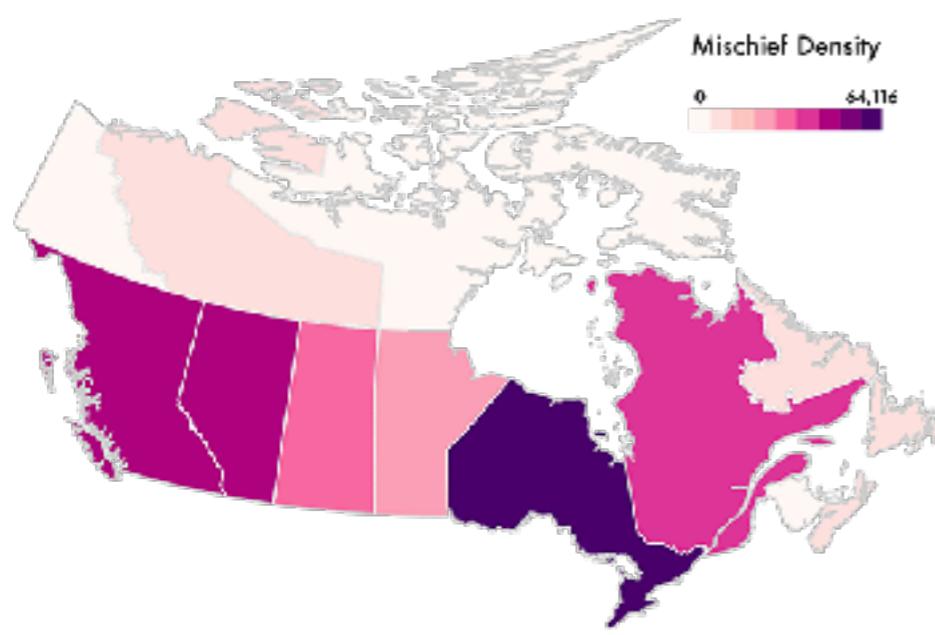
WHAT IS DATA VISUALIZATION?

How Many “V”?

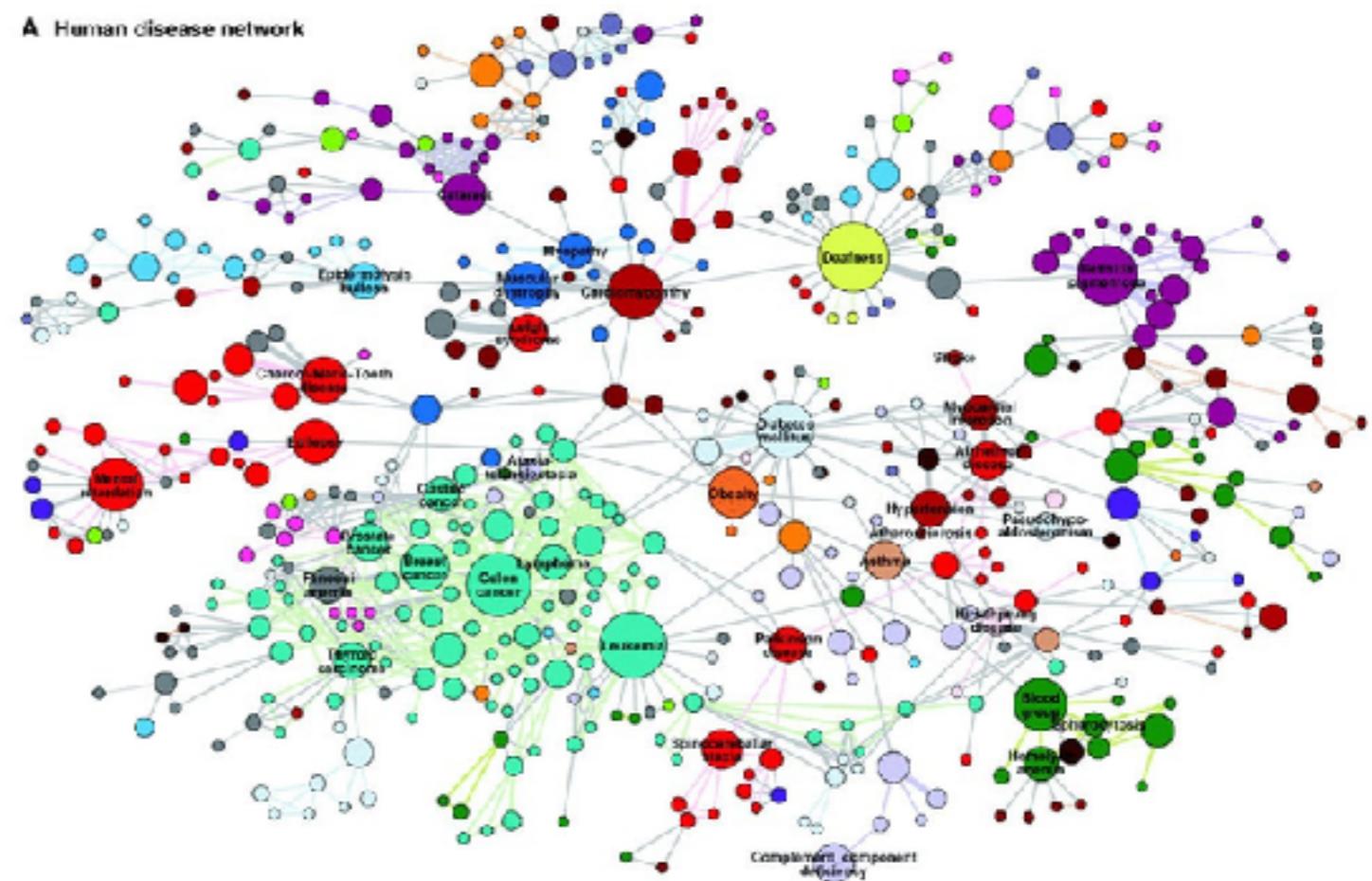
MTHIVLWYADCEQGHKILKMTWYN
ARDCAIREQGHLVKMFNSTWYARN
GFPSSVCEILQGKMFPSNDRCEQDIFP
SGHLMFHKMVPSTWYACEQTWRN

WHAT IS DATA VISUALIZATION?

To convey information through graphical representations

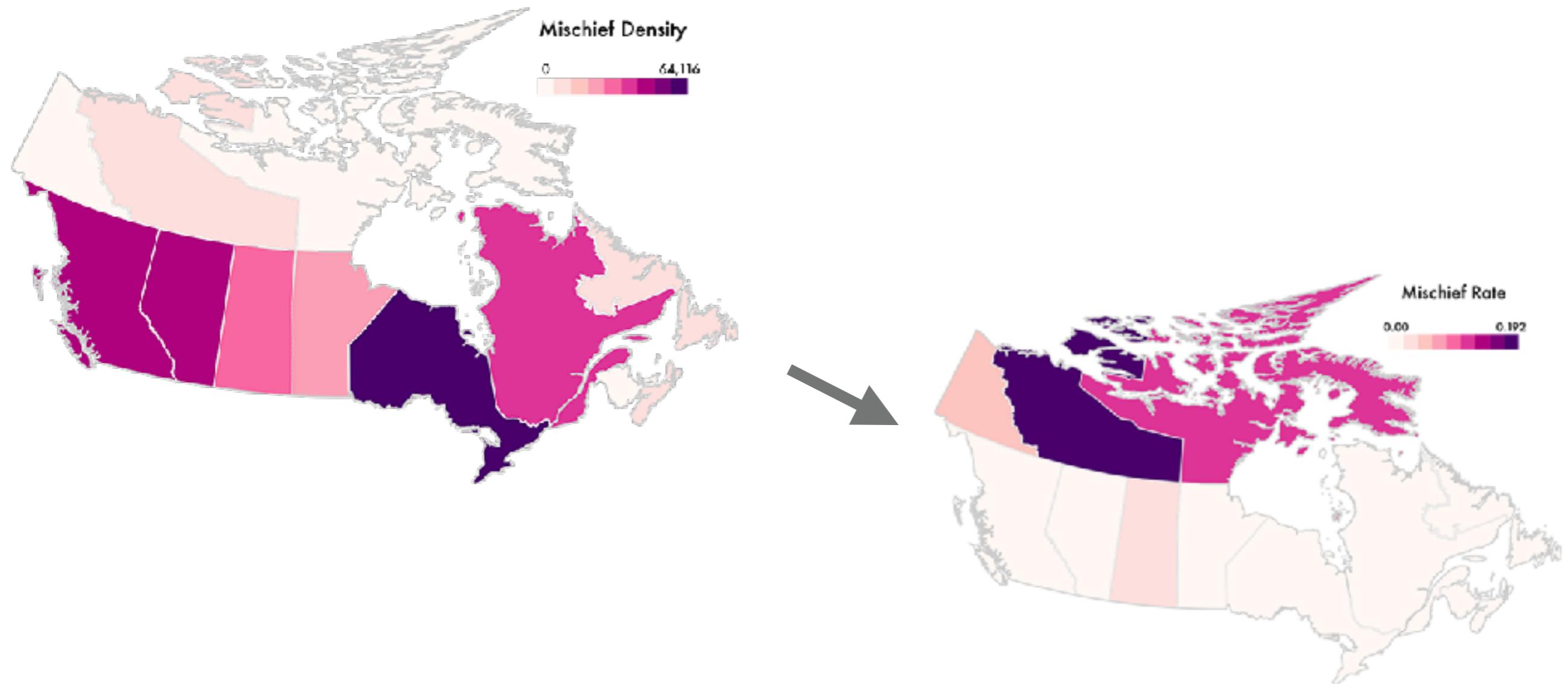


A. Human disease network

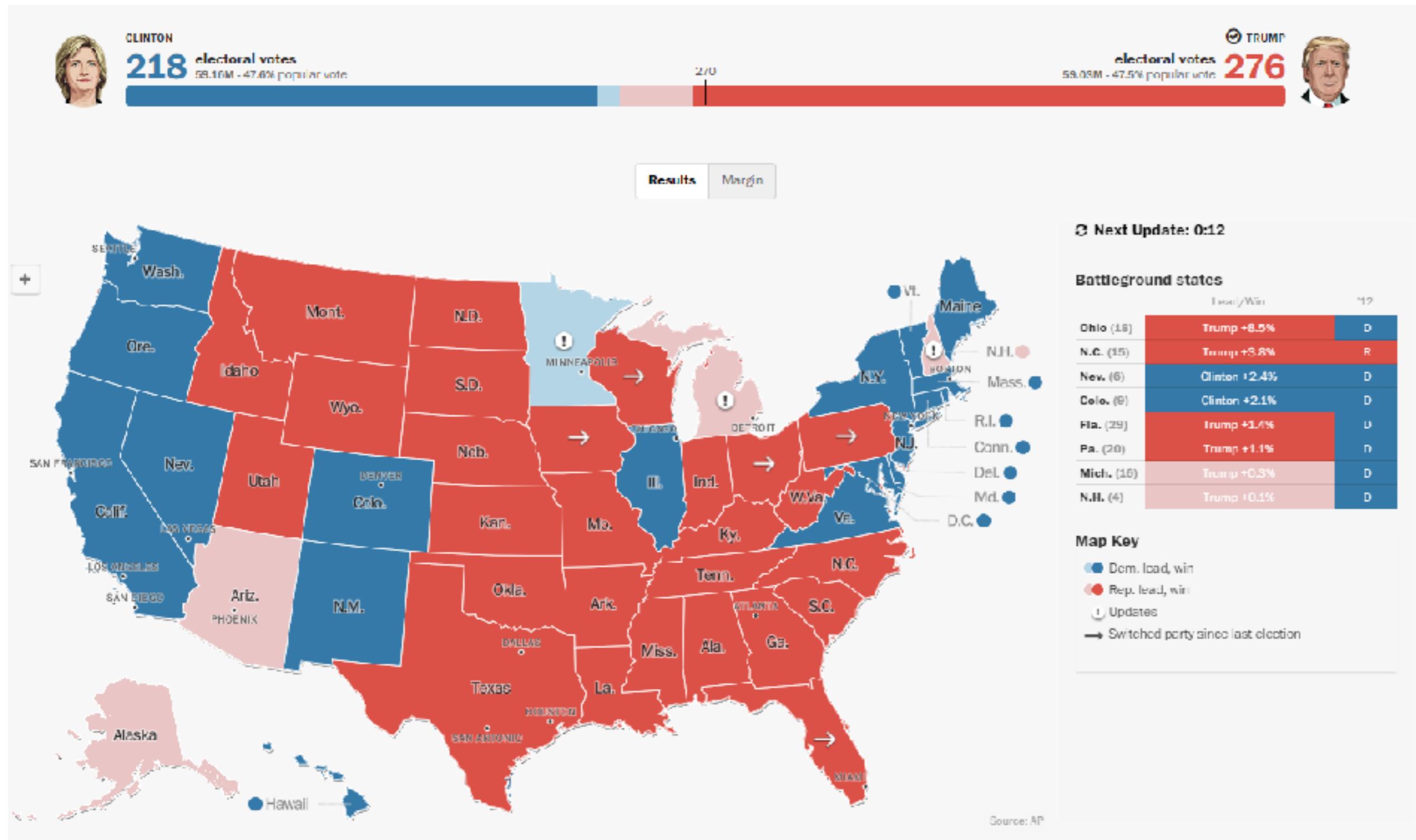


WHAT IS DATA VISUALIZATION?

To convey information through graphical representations



WHAT IS DATA VISUALIZATION?



WHY DATA VISUALIZATION?

Set A		Set B		Set C		Set D	
X	Y	X	Y	X	Y	X	Y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.11	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Summary Statistics

$$\mu_X = 9.0 \quad \sigma_X = 3.317$$

$$\mu_Y = 7.5 \quad \sigma_Y = 2.03$$

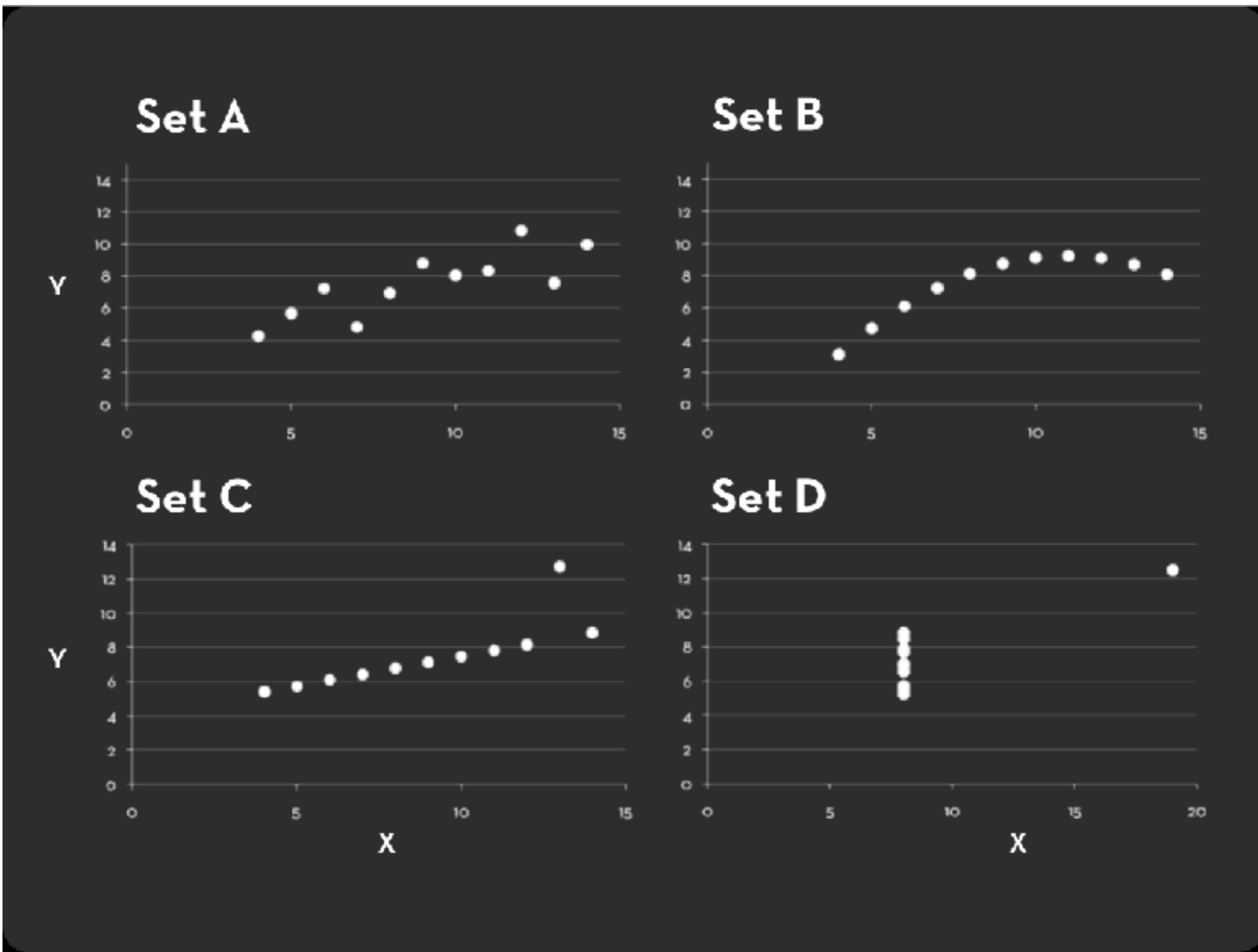
Linear Regression

$$Y = 3 + 0.5 X$$

$$R^2 = 0.67$$

[Anscombe 73]

WHY DATA VISUALIZATION?



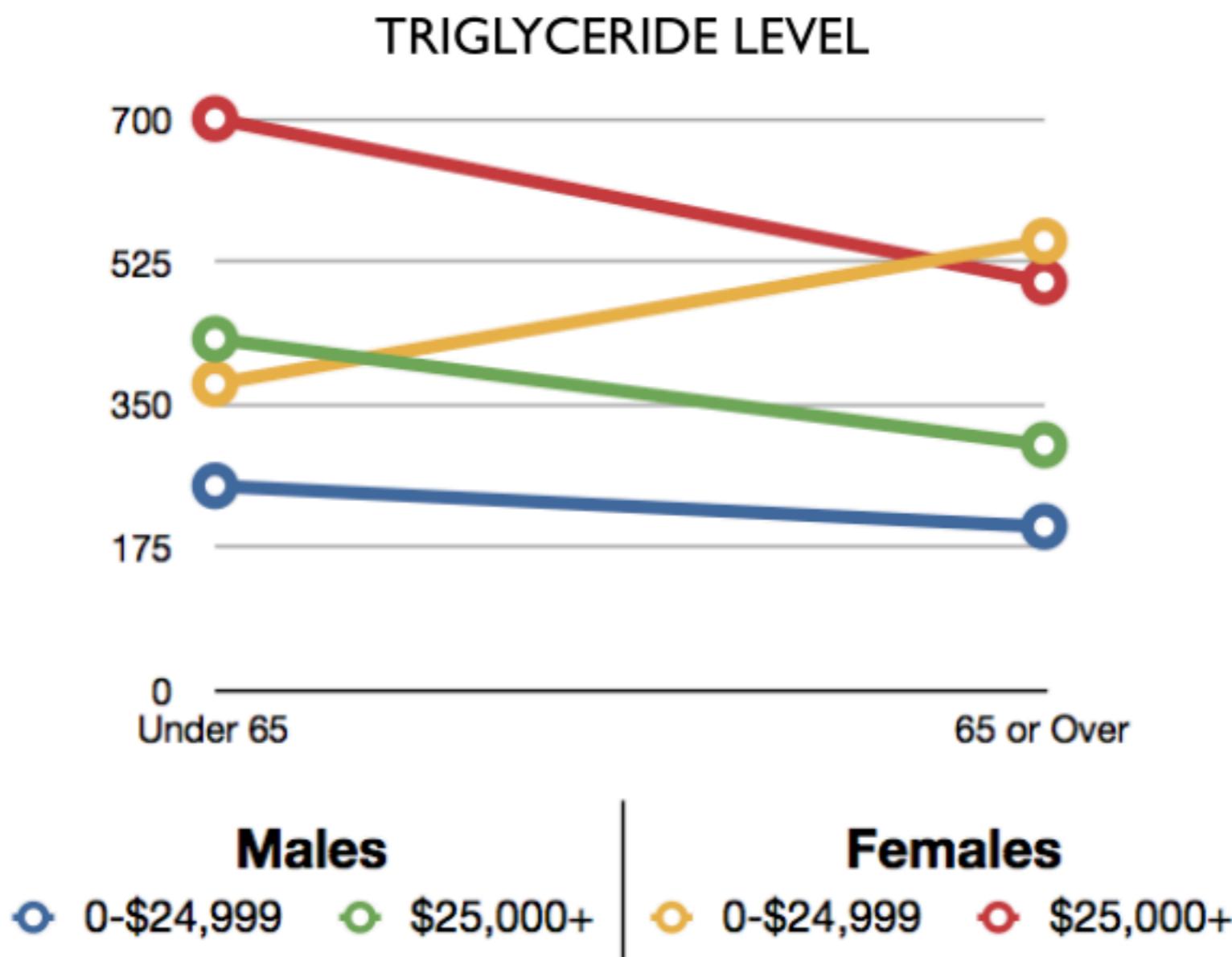
WHY DATA VISUALIZATION?

Which gender or income level group shows different effects of age on triglyceride level?

Income Group	Males		Females	
	Under 65	65+	Under 65	65+
0-\$24,999	250	200	375	550
\$25,000+	430	300	700	500

WHY DATA VISUALIZATION?

Which gender or income level group shows different effects of age on triglyceride level?



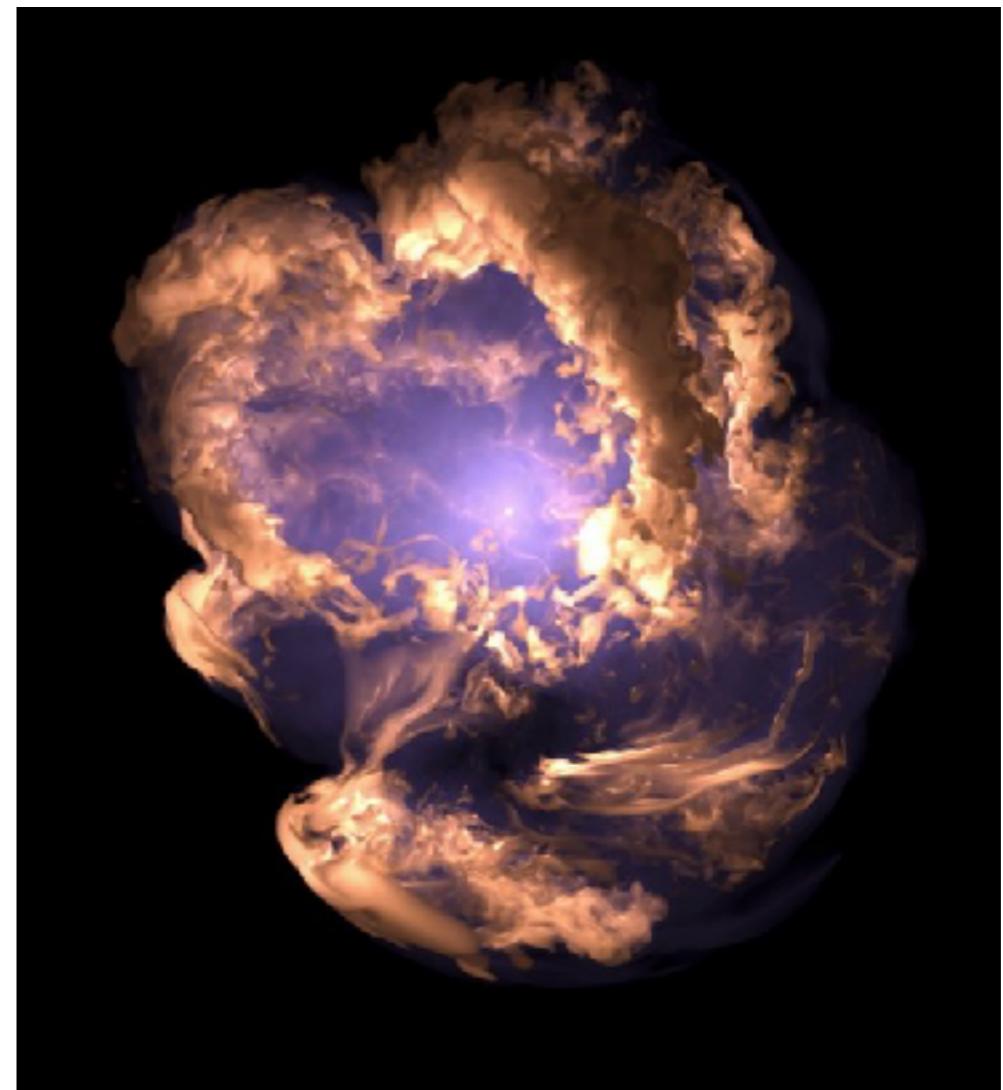
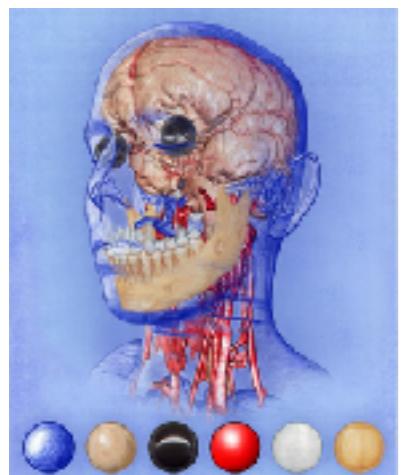
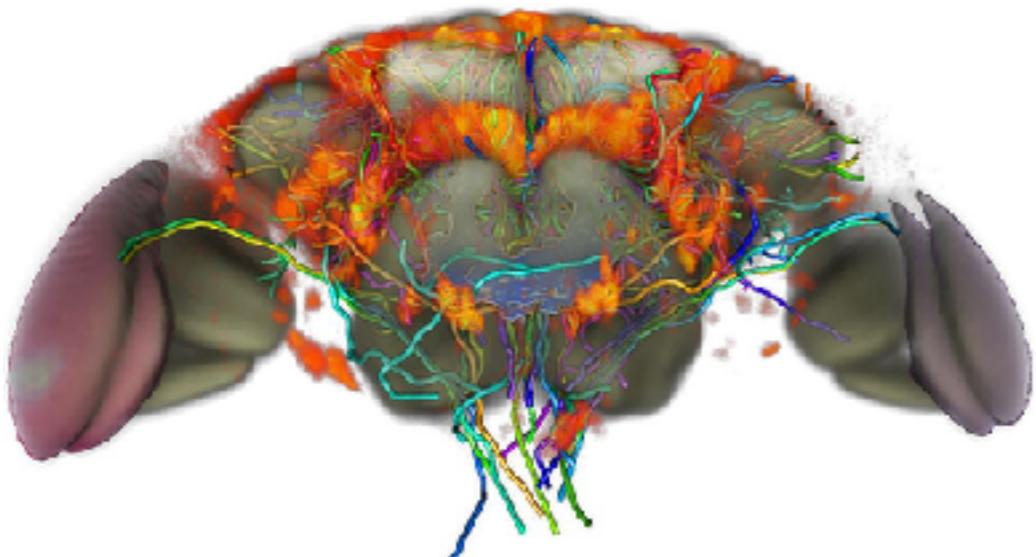
WHY DATA VISUALIZATION?

1. *Help us think about data*
2. *Use perception to offload cognition*
3. *Serves as an external aid to augment working memory*
4. *Boosts our cognitive abilities*

APPLICATION OF DATA VISUALIZATION

Volume Visualization

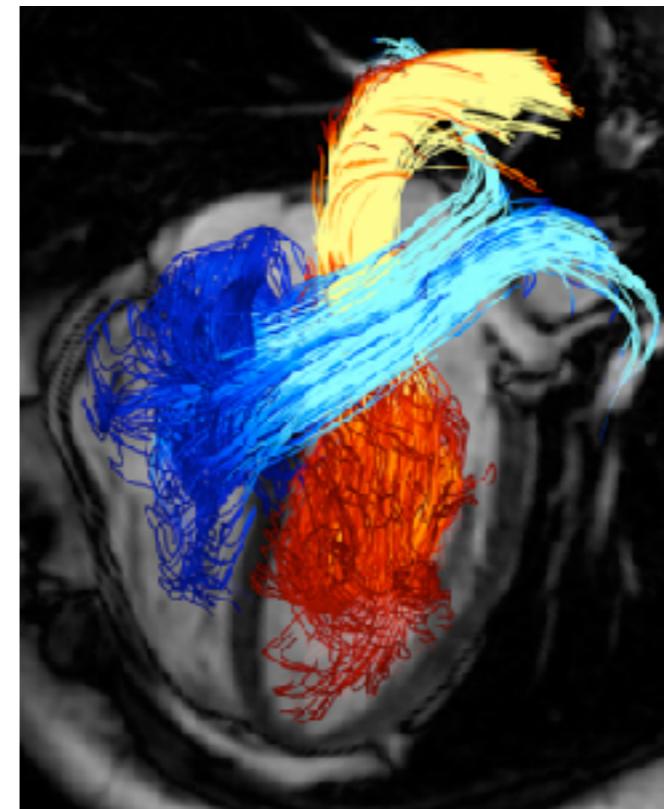
- *Visualization of scalar fields*
- *Important in medicine, biology, geoscience, engineering*



APPLICATION OF DATA VISUALIZATION

Flow Visualization

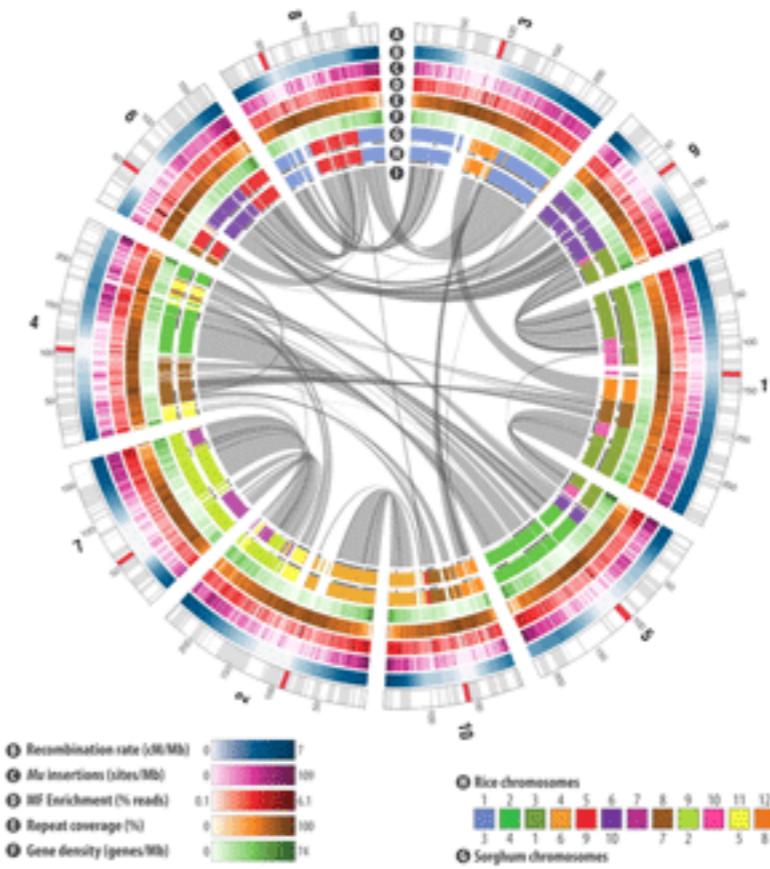
- *Visualization of vector fields*
- *Data typically from computational fluid dynamics (CFD) simulations*



APPLICATION OF DATA VISUALIZATION

Information Visualization

- *High dimensional data, hierarchical data, textual data, ...*



HOW TO DO DATA VISUALIZATION

A. *Tableau*

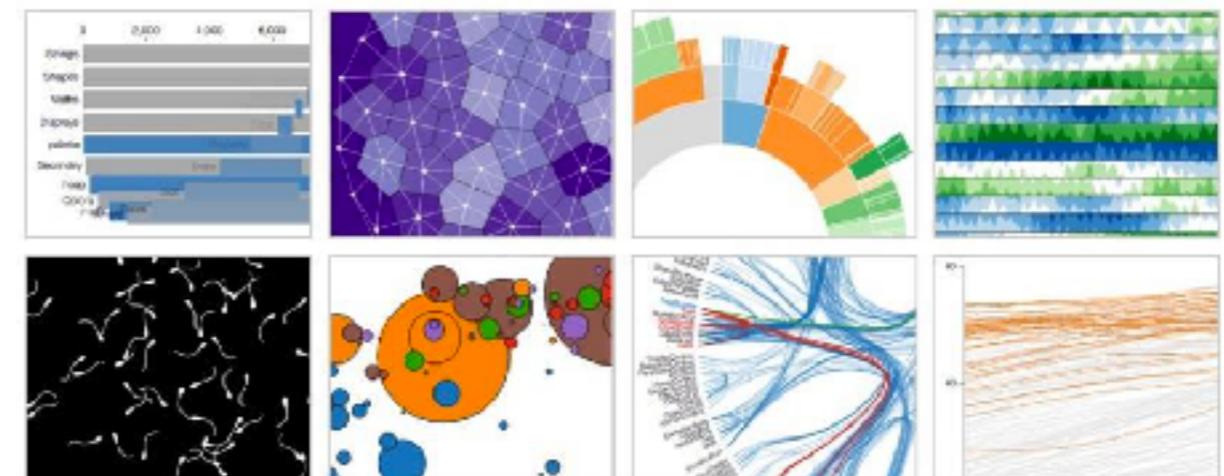


B. *R*

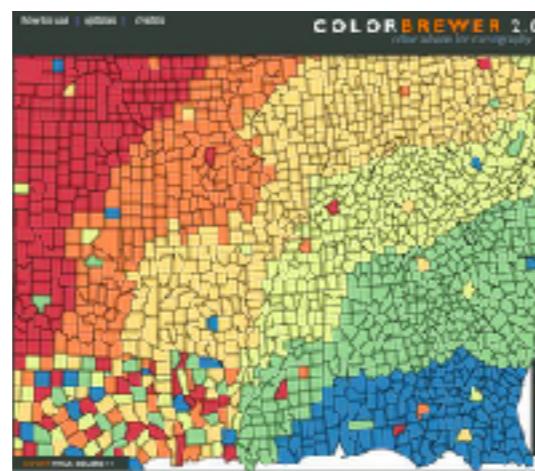


C. *Google Refine*

Data-Driven Documents



D. *Processing*



E. *D3 (JS)*



F. *ColorBrewer*



TABLEAU



- Tableau is a pathbreaking data visualization software for Business Intelligence (BI).
- Tableau have access to many data source.
- Tableau supports interactive visualizations which is called dashboards.

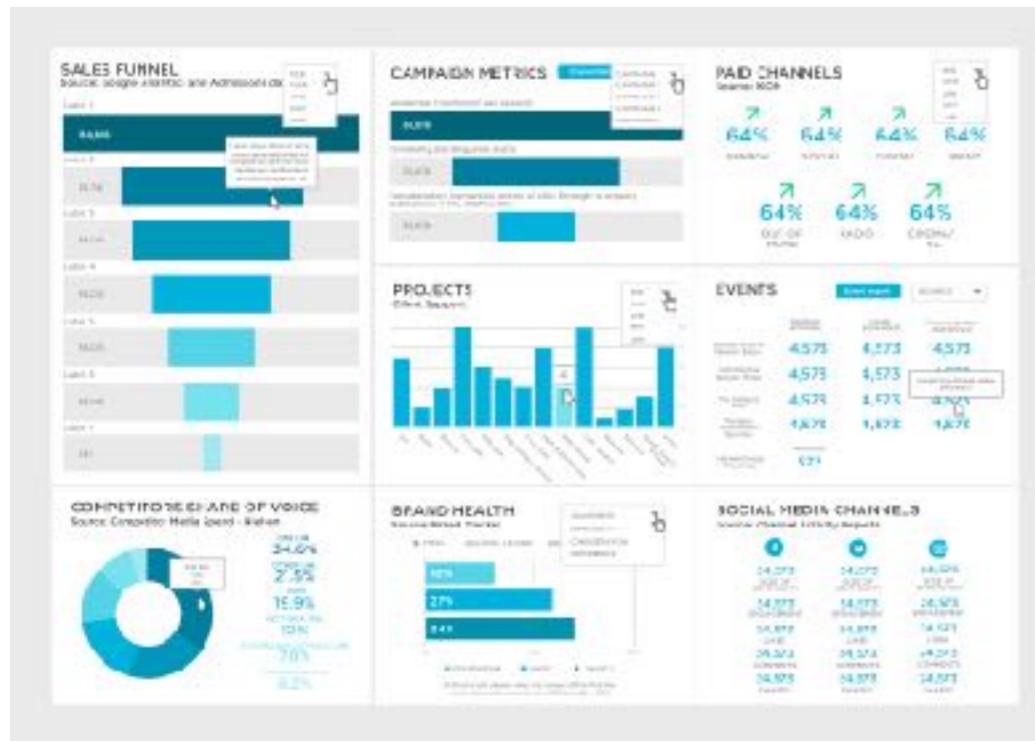
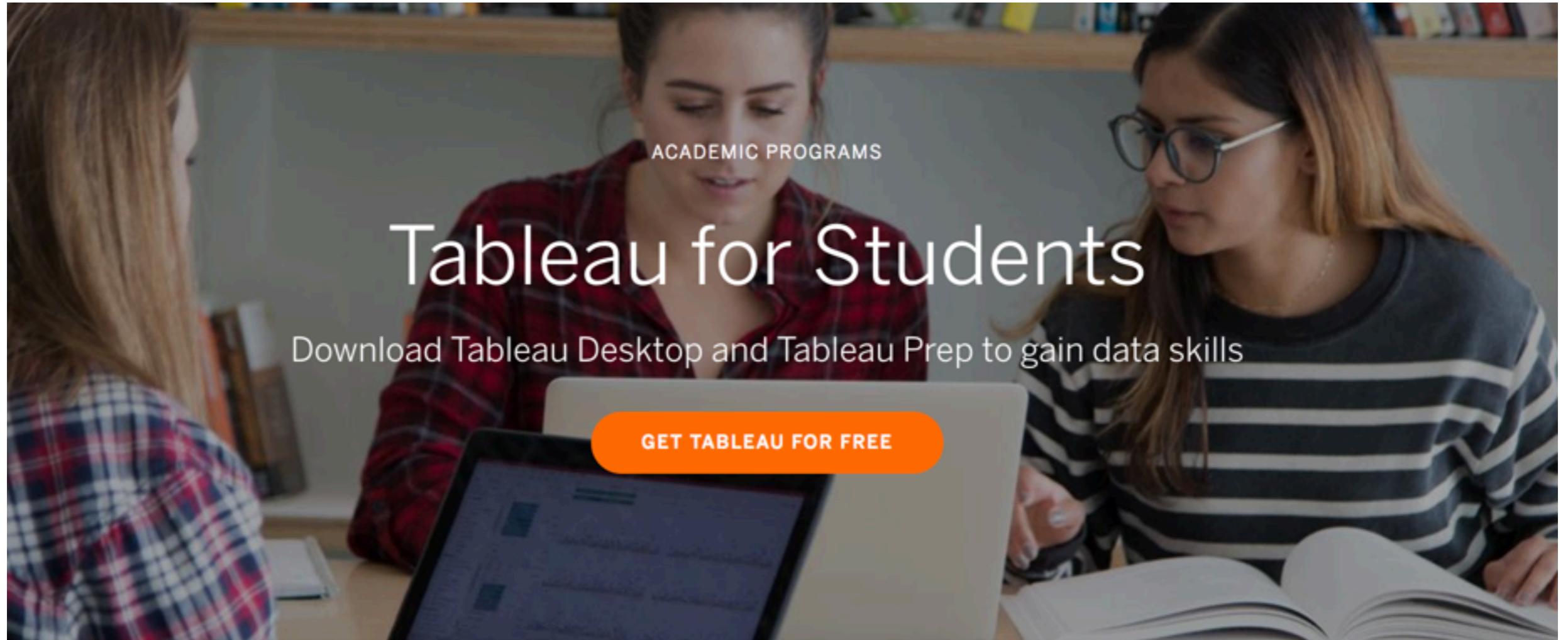


Tableau Download Free for Students for One Year



<https://www.tableau.com/academic/students>

TABLEAU EXAMPLE



ebay data:

- *Customer*
- *Orders*
- *Products*

Tools:

- *Tableau Desktop*
- *Tableau Online*

A	B	C	D	E	F	G	H	I
1 Order ID	Ship Date	Ship Mode	Sales	Discount	Profit	Shipping Cost	Order Date	Order Priority
2 CA-2014-AB10015140-41954	2014-11-13	First Class	\$221.98	0	\$62.15	40.77	2014-11-11	High
3 CA-2014-AB10015140-41954	2014-11-13	First Class	\$341.96	0	\$84.71	25.27	2014-11-11	High
4 CA-2012-AB10015140-40974	2012-03-07	First Class	\$48.71	0.2	\$5.48	11.13	2012-03-06	High
5 CA-2012-AB10015140-40974	2012-03-07	First Class	\$17.94	0	\$4.66	4.29	2012-03-06	High
6 CA-2012-AB10015140-40958	2012-02-25	Standard Class	\$12.62	0.2	-\$2.52	1.97	2012-02-26	Medium
7 CA-2012-AB10015140-40974	2012-03-07	First Class	\$242.94	0	\$64.96	12.98	2012-03-06	High

Order

A	B	C	D	E	F	G	H	I	J
1 Order ID	Customer ID	Customer Name	Segment	Postal Code	City	State	Country	Region	Market
2 CA-2014-AB10015140-41954	AB-100151402	Aaron Bergman	Consumer	73120	Oklahoma City	Oklahoma	United States	Central US	USCA
3 Ca-2014-AB10015140-41954	AB-100151402	Aaron Bergman	Consumer	73120	Oklahoma City	Oklahoma	United States	Central US	USCA
4 Ca-2012-AB10015140-40974	AB-100151404	Aaron Bergman	Consumer	98002	Seattle	Washington	United States	Western US	USCA
5 Ca-2012-AB10015140-40974	AB-100151404	Aaron Bergman	Consumer	98002	Seattle	Washington	United States	Western US	USCA
6 Ca-2012-AB10015140-40974	AB-100151402	Aaron Bergman	Consumer	76117	Arlington	Texas	United States	Central US	USCA
7 Ca-2012-AB10015140-40974	AB-100151404	Aaron Bergman	Consumer	98101	Seattle	Washington	United States	Western US	USCA
8 ID-2012-AB1001527-41439	AB-1001527	Aaron Bergman	Consumer		Ruhan	Ruhel	China	Eastern Asia\\Pacific	
9 ID-2014-AB100157-41815	AB-100157	Aaron Bergman	Consumer		Marius	Queensland	Australia	Oceania	Asia Pacific
10 IS-2014-AB1001545-41950	AB-1001545	Aaron Bergman	Consumer		Marseille	Provence-Alpes-Cote d'Azur	France	Western Europe	Western Europe

Customer

A	B	C	D	E	F
1 Order ID	Product ID	Category	Sub-Category	Product Name	Quantity
2 CA-2014-AB10015140-41954	TEC-PH-5816	Technology	Phones	Samsung Convoy 3	2
3 CA-2014-AB10015140-41954	FUR-BG-5957	Furniture	Bookcases	Sauder Facets Collection	2
4 CA-2012-AB10015140-40974	FUR-CH-4421	Office Supplies	Chairs	Global Push Button	1
5 CA-2012-AB10015140-40974	OFF-AR-5309	Office Supplies	Art	Newell 330	3
6 CA-2012-AB10015140-40958	OFF-ST-307B	Office Supplies	Storage	Akro Stacking Bins	2
7 CA-2012-AB10015140-40974	OFF-ST-3744	Office Supplies	Storage	Carina 42" Hx23 3/4	3
8 ID-2012-AB1001527-41439	TEC-MA-6150	Technology	Machines	StarTech Printer WiFi	0

Products

TABLEAU EXAMPLE - DESKTOP



Connect data

The screenshot shows the Tableau Desktop application window. On the left, a dark sidebar titled 'Connect' lists various data source options: 'To a File' (Microsoft Excel, Text file, JSON file, PDF file, Spatial file, Statistical file, More...), 'To a Server' (Tableau Server, Microsoft SQL Server, MySQL, Oracle, Amazon Redshift, More...), and 'Saved Data Sources' (Sample - Superstore). The main workspace is titled 'Tableau - Book1'. It features three sections: 'Open' (containing thumbnails for 'Aug17', 'Graph', 'Graph', 'Graph', and 'Book2'), 'Discover' (with links to 'Training', 'Viz of the Week', 'Resources', 'Blog', and 'Forums'), and 'Sample Workbooks' (with thumbnails for 'Sample - Superstore' and 'More Samples'). A prominent orange button at the bottom right says 'Update to 2018.3 Now'.

TABLEAU EXAMPLE - DESKTOP



Select data

The screenshot shows the Tableau Desktop interface with the title bar "Tableau - Book1".

Connections: sales (Microsoft Excel)

Sheets:

- Use Data Interpreter
Data Interpreter might be able to clean your Microsoft Excel workbook.
- Customer
- Orders
- Products
- New Union

Drag sheets here (An orange arrow points from the "Products" sheet towards this area.)

Data Source: Sheet 1 (highlighted with an orange oval)

TABLEAU EXAMPLE



one more step to active Tableau Online

Tableau Online: Login

<https://online.tableau.com/> ▾

Tableau Online Secure Login Page. Sign In to Tableau Online.

You've visited this page 3 times. Last visit: 19/11/18

Your free trial

Reasons to start with Tableau Online:
Connect to data sources ...

Tableau Online Help

Tableau Online, by Tableau
Software®, is a secure, cloud ...

[More results from tableau.com >](#)

Login

Tableau Online Secure Login Page.
[Sign In to Tableau Online.](#)

Online Support

Free Training Videos. Administrative
Overview ...



Your site is waiting for you.

[Activate My Site](#)

D3.JS

Data-Driven Documents



D3.js

Make a webpage

index.html

We can start by defining a simple web page, which has a header (`h1`) and an `svg` element that will hold our visualization. In the style tags, we can add CSS styling for both elements defined in the HTML.

```
<head><meta charset="utf-8">  
<title>D3 Example</title>  
<style>  
h1 {  
    font-family: sans-serif;  
    text-align: center;  
}  
  
svg {  
    display: block;  
    margin-left: auto;  
    margin-right: auto;  
    border: 1px solid black;  
}  
  
</style>  
<script src="https://d3js.org/d3.v4.min.js" charset="utf-8"></script>  
</head>  
  
<body>  
<h1>D3 Example</h1>  
  
<svg class="chart" width="500px" height="500px">  
</svg>  
  
</body>
```

CSS Style

HTML

D3.js

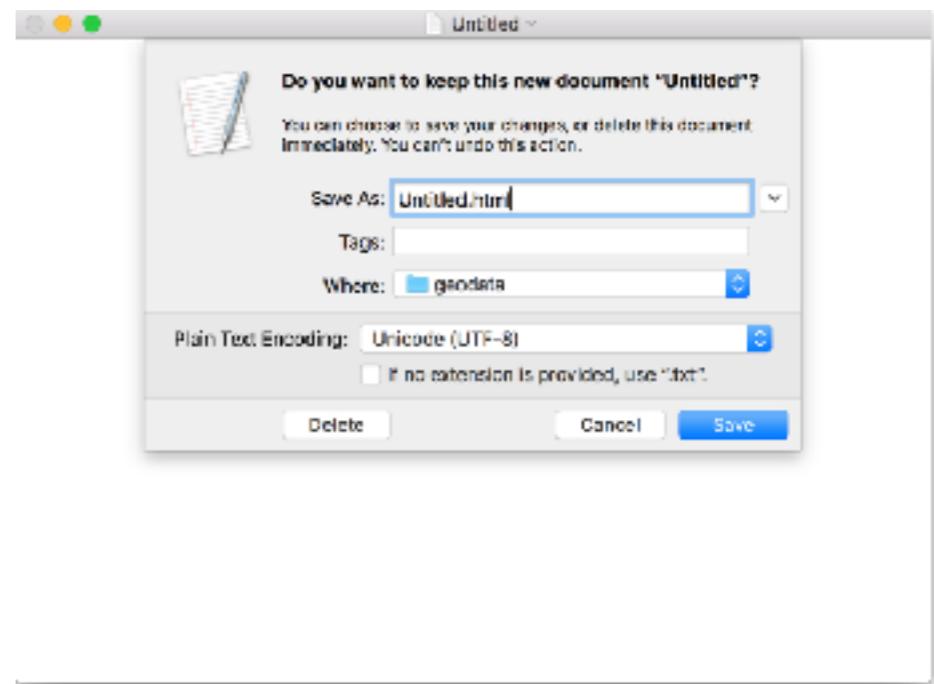
Steps to create a webpage

Step 1: Text Editor

Open your favourite text editor.

TextEdit (Mac); Notepad (PC)

Step 3: Save HTML page



Step 2: Write some HTML

```
<html>
<header><title>This is title</title></header>
<body>
<h3>Let's learn D3!</h3>
</body>
</html>
```

Step 4: View HTML Page in Browser

D3.js

DOM Selection

Script tag:

```
<script src='https://d3js.org/d3.v4.min.js'></script>
```

```
<html>
  <head>
    <title>Learn D3 in 5 minutes</title>
  </head>
  <body>

    <h3>Let's learn D3!</h3>

    <script src='https://d3js.org/d3.v4.min.js'></script>

    <script>
      d3.select('h3').style('color', 'darkblue');
      d3.select('h3').style('font-size', '24px');
    </script>

  </body>
</html>
```

REFERENCE

- Munzner, Tamara. *Visualization analysis and design*. AK Peters/CRC Press, 2014.
- Reading Packages from 2018 ZJU summer school on visual analytics