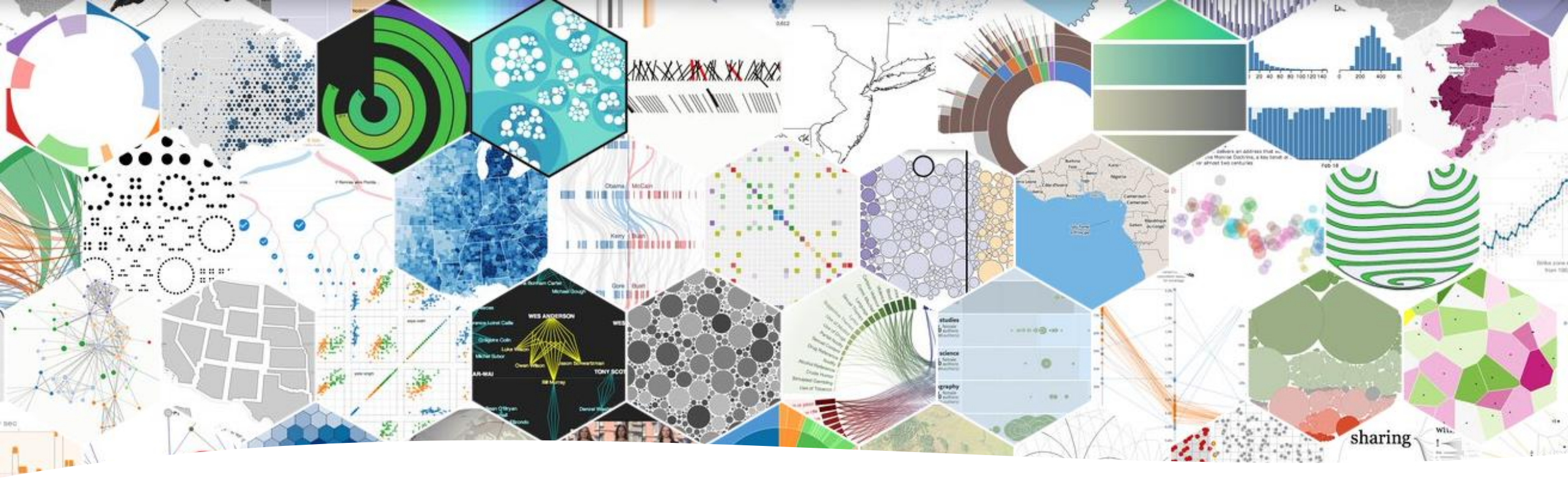
Data Visualization in a nutshell

Presented by Wing Chan

Agenda

- What is Data Visualization?
- Data Visualization Framework
- 3 styles of Data Visualization
- Tools (Tableau vs D3.js)
- Demos

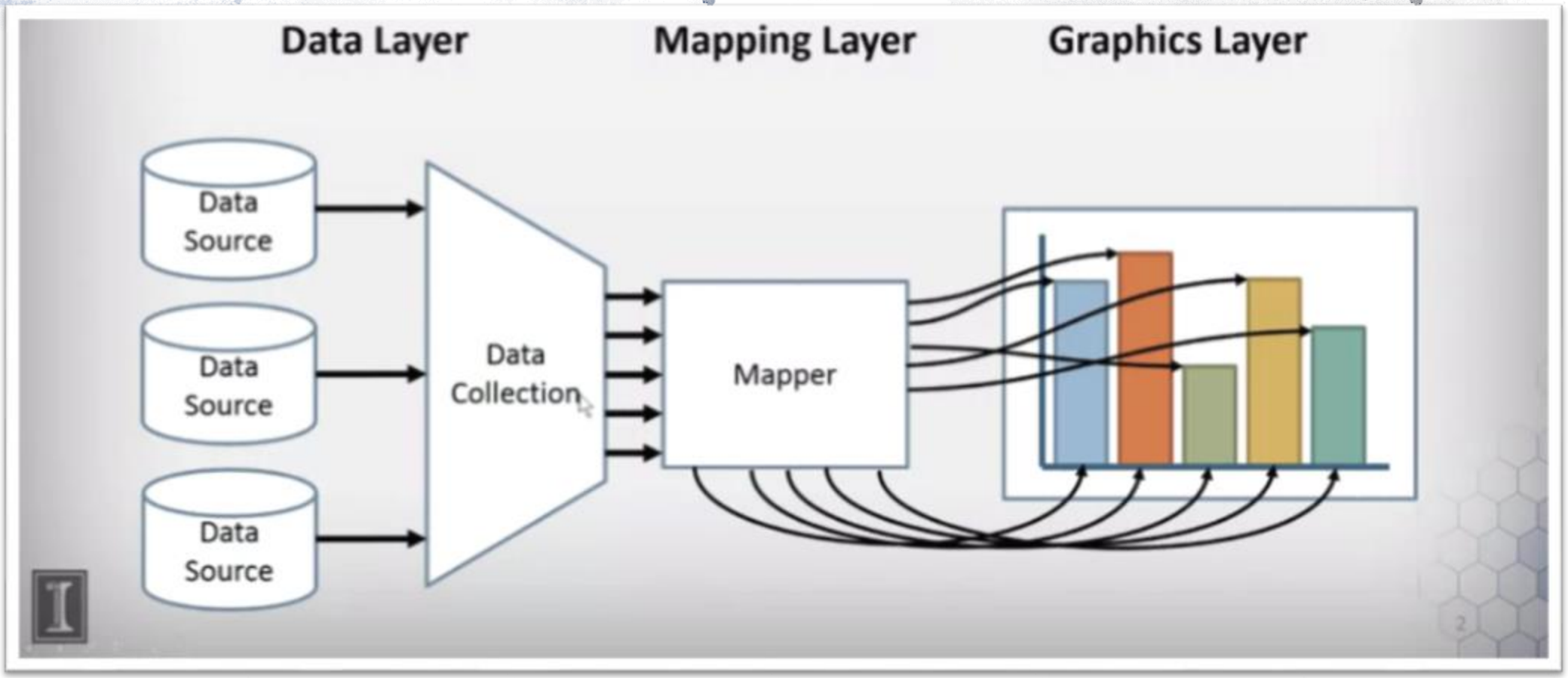


What is Data Visualization?

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

In the world of Big Data, data visualization tools and technologies are essential to analyze massive amounts of information and make data-driven decisions.

Data Visualization Framework



Data Layer

- Purposes:
  - Locating and obtaining data
  - Importing data in proper format
  - Relating data for proper correspondence
  - Data analysis and aggregation
- Data Types
  - Discrete vs Continuous
  - Ordered vs Unordered

	Discrete (no between values)	**Continuous** (values between)
Ordered (values are comparable)	**Ordinal,** e.g. size: S,M,L,XL,… **Quantitative,** e.g. counts: 1,2,3,…	**Fields,** e.g. altitude, temperature
Unordered (values not comparable)	**Nominal,** e.g. shape: ☐○△ **Categories,** e.g. nationality	**Cyclic values,** e.g. directions, hues

Mapping Layer

- Purposes
  - Associating appropriate geometry with corresponding data channels
  - Data analysis and algorithms
- Perceptual Accuracy

Quantitative	Ordinal	Nominal
Position	Position	Position
Length	Density	Hue
Angle	Saturation	Texture
Slope	Hue	Connection
Area	Texture	Containment
Volume	Connection	Density
Density	Containment	Saturation
Saturation	Length	Shape
Hue	Angle	Length
	Slope	Angle
	Area	Slope
	Volume	Area
		Volume

Graphics Layer

- Purposes:
  - Conversion of geometry into displayable image
  - Decorations
  - Managing interaction
- Common Charts
  - Bar Chart, Line Chart, Scatter Plots and Gantt Chart
- When to use?

Dep.	Quantitative Continuous	Bar	Line
	Quantitative Discrete	Bar	Bar
Ind.	Quantitative Continuous	Gantt	Scatter
	Nominal or Q. Discrete	Table	Gantt
		Nominal or Q. Discrete	Quantitative Continuous
		Independent	

3 styles of Data Visualization

1. Interactive Visualization
2. Presentation Visualization
3. Interactive Storytelling

Interactive Visualization

- Used for discovery
- Intended for a single investigator or collaborators
- Pre-renders based on input
- Prototype quality
- Example: Tableau public
- https://public.tableau.com/profile/wing.chan7063#!/vizhome/MoviesDashboard_15619067238430/Dashboard

Presentation Visualization

- Used for communication
- Intended for large group or mass audience
- Does not support user input
- Highly polished
- Example: Electric cars sales
- https://cleantechnica.com/2019/05/24/visualization-2012-2019-us-electric-car-sales-this-is-a-must-see/

Interactive Storytelling

- Presentations via interactive webpages
- Narrative, using data to tell a story
- Invites viewers to explore further
- A medium for data journalism
- Example: The Economist: Big Mac Index
- https://www.economist.com/comment/2602607

Tools

- **Tableau** - a Business Intelligence Tool used for data visualization that enables you to create interactive visualizations in form of dashboards, worksheets to gain business insights. It allows non-technical users to easily create customized dashboards that provide insight to a broad spectrum of information. And, Tableau software founded in 2003. In 2019, it is acquired by Salesforce.

- **D3.js** - D3 stands for Data-Driven-Documents and is a JavaScript library for dynamic and interactive data visualizations in web browsers. It was created in 2011 and is used on many websites. It provides web developers with a toolset for utilizing APIs and SVG graphics to create visualizations across the web.

D3.js

```
<html>
<script src="https://d3js.org/d3.v5.min.js"/>
<body>
    <div id="chart"></div>

    <script>
    var data = [10, 20, 30];

    d3.select('#chart')
        .selectAll('div')
        .data(data)
        .enter()
            .append('div')
            .text(function(d, i){ return d; });
    </script>
</body>
</html>
```

Tools

Tool	Tableau	D3.js
Target User	Business Analysts	Developers
Cost	Expensive	Free
Learning curve	Low	High
Interface	Desktop app or Web portal	JavaScript library – load data and generate SVG or Canvas graphics
Advantage	• Great for data exploration • Short development time	• Integrate better on websites • Highly customizable
Disadvantage	• Expensive • Not easy to integrate with websites	• Long development time • Require web development skill

Demos

- Data: Titanic Survival dataset from Kaggle

  https://www.kaggle.com/c/titanic/data

- Demo #1: using Tableau

  https://public.tableau.com/profile/wing.chan7063#!/vizhome/TitanticDashboard/Dashboard

- Demo #2: using D3.js

  https://wingchanatibsa.github.io/cs498dv/index.html


Questions?