



Data Visualization

for Analytics and Business Intelligence
A Comprehensive Overview

Dr. Jack Zheng

Version 7.1, Fall 2023

IT 7113 Data Visualization

<http://idi.kennesaw.edu/it7113/>

Overview



This lecture notes provides a high-level overview of data visualization primarily used in business intelligence and analytics. This overview is comprehensive and covers as many aspects as possible, but it keeps them at a high level. More details are provided in additional lecture notes.

Topics:

1. Basic concepts
 - What is data visualization?
 - What are the purposes and types of visualization?
2. What are the related terms and fields? How are they similar or different?
3. Data visualization in business intelligence and analytics
 - Basic data visualization elements and forms (types)
4. Design and development
 - Principles, processes, applications and tools, IT support
5. Learning and career

Sections



Data Visualization Concepts

Definition, basic concepts and terms
Values and benefits of data visualization

4



Related Terms and Fields

Data presentation
Information design
Information visualization, including infographic and illustration
Text visualization
Big data visualization
Computer graphics, reality visualization (VR, AR), scientific visualization
Data art
Business data visualization
Visual analytics

IT 7113 focus

21



Data Visualization in BI/Analytics

Basic visual properties and visual forms/styles used in BI and analytics applications

38



Developing Data Visualizations

The design/development of data visualizations involves
Processes
Principles
Best practices
Tools
IT's role in business data visualization

53



Learning Data Visualization

Skills, jobs, career, and learning resources

63





Data Visualization Concepts

Definition, basic concepts and terms

Values and benefits of data visualization



Data Visualization

Data visualization is the **visual** representation and presentation of **data** for the purpose of enhancing perception and cognition.

- Key terms
 - Data, visual
 - Representation, presentation
 - Purpose: perception and cognition
- Modern systems and software applications also emphasizes the **interactive** feature of this process.

More details about interactivity in IT 7113 module 10



Notable Definitions

- Some notable definitions also share some common features of the four key terms.
- Tableau (<https://www.tableau.com/learn/articles/data-visualization>)
 - Data visualisation is the graphical **representation** of information and data. By using visual elements like charts, graphs and maps, data visualisation tools provide an accessible way to see and understand trends, outliers and patterns in data.
- Andy Kirk, in his book Data Visualisation
 - “The **representation and presentation** of data to facilitate understanding.”
- Stephen Few (<https://www.perceptualedge.com/blog/?p=2636>)
 - Data visualization is a collection of methods that use visual **representations** to explore, make sense of, and communicate quantitative data.
 - Read more comments from users on this page

Essential read: Defining Data Visualisation

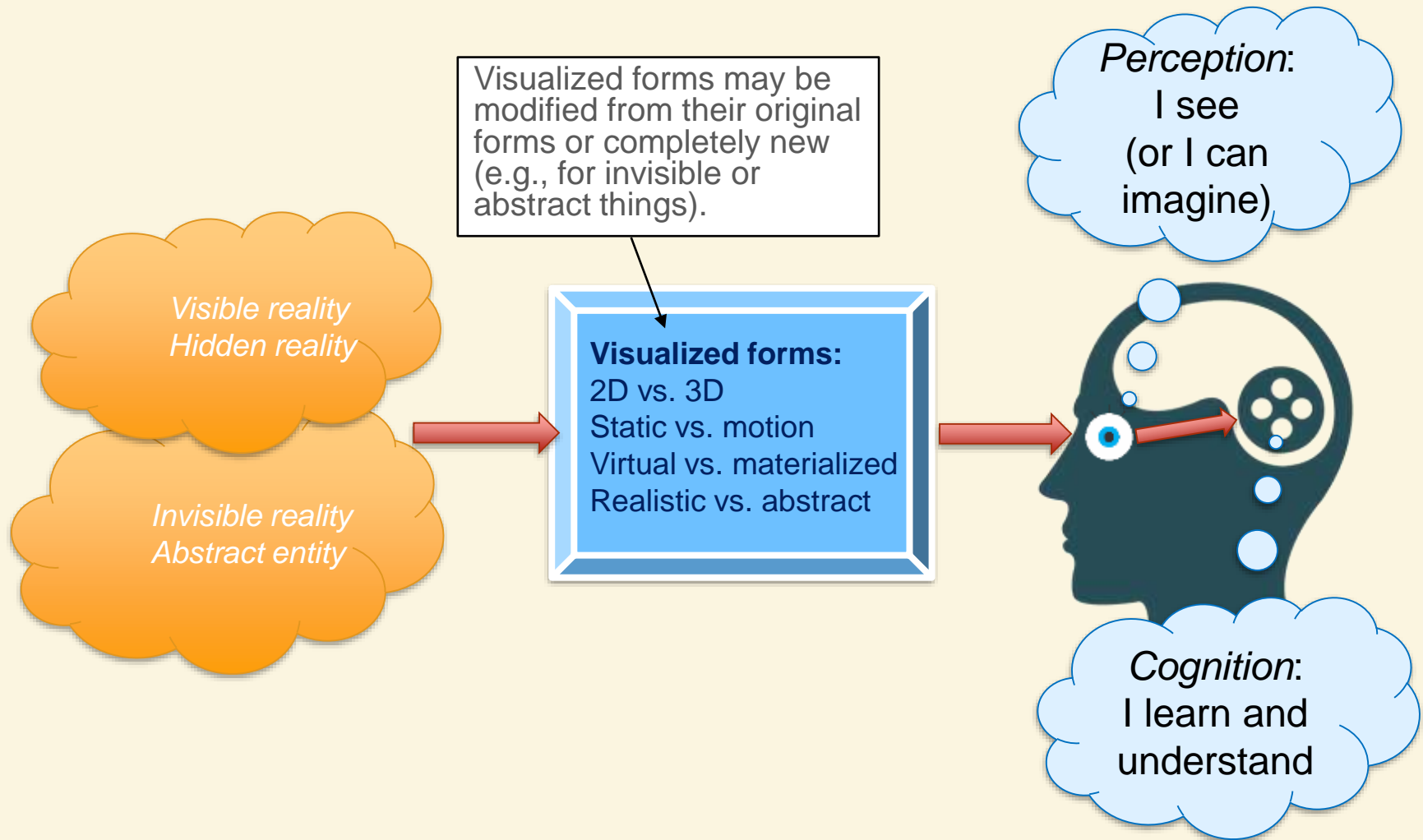
<https://chrisluv.medium.com/defining-data-visualisation-daf71c22ec03>

Data, Visual, Visualization



- Visual (graphic)
 - is related to **vision** (seeing through eyes) - one of the major human senses to interact with the world.
 - meaning it can be seen by human eyes; or can be imagined (even though eyes are closed).
- **Visualization** is the process of forming a visual image of things that can be seen through eyes (and/or imagined in human mind).
- What can be visualized?
 - *Visible reality*: person, animal, building, mountain
 - *Hidden reality*: earth core, blood vessel, universe
 - *Invisible reality*: wind (hint.fm/wind), air, heat, electron, sound, smell, magnetic fields
 - *Abstract entity*: activity, event, idea, hierarchy, process, relationship
- How to visualize them?
 - These realities and entities can be described using data and information.
 - Data are typically numerical values (but can also be qualitative) that describe its associated entity or activity.
- Data itself is abstract. The visualization process creates visible forms to *represent* the meaning of these abstract data.
 - Utilizing a combination of the three elements: spatial substrate (space and area), graphical (visual) elements (shapes and symbols), and graphical properties (visual variables/properties) like size, color, positions, etc. – see next a few slides
- Types/features of visualization forms
 - 2D vs. 3D objects
 - Static vs. motion visuals
 - Virtual (created in computers) vs. materialized (built/displayed in real world)
 - Realistic (using realistic objects) vs. abstract (using abstract shapes)
- The visualization process (and the result) may change the original form of things, or create new forms, for better understanding and communication.
 - For example, using 2D squares to represent 50 states on a map; or using arrows for air flow.

Visualizing “visualization”



Representation



- Representation is an encoding or visual mapping process that connects visuals to data (meanings)

Visual mapping = data + visual

The + here indicates some kind of defined associations (encoding or mapping) between the data and visuals

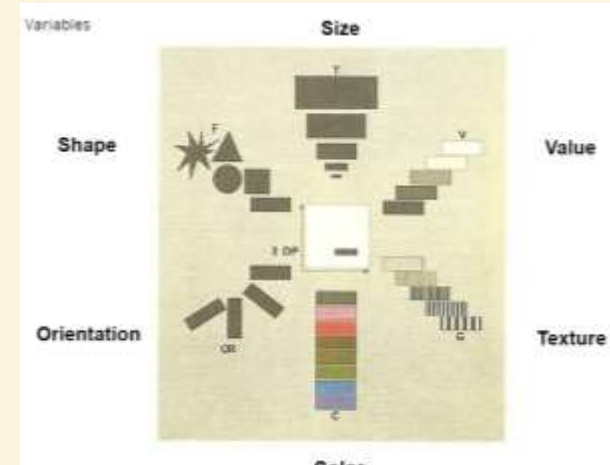


- It directly defines how data is coded or mapped to the three elements
 1. Spatial substrate: The spatial substrate is the space in which we're going to create the visualization.
 2. Graphical (visual) elements:
 - Visual (or graphical) elements will appear in the spatial substrate. They represent data items or entities.
 - Visual elements are the basic building blocks in a visualization (a chart or a diagram).
 3. Graphical properties (visual variables/properties)
 - Visual property, or attribute, or variable, is the “decoration” applied to visual elements to represent data values
 - http://wiki.gis.com/wiki/index.php/Visual_Variables
 - <https://blogs.ifgi.de/digital-cartography/symbols/visual-variables/>
 - A visual property is used to encode different values of a particular dimension of data

Visual properties will be covered in IT 7113 module 2.

The six basic visual properties can be remembered as “**SCOPEs**” - my term - refer to the lecture notes “Visual Encoding with SCOPEs Visual Variables/Properties” <https://www.edocr.com/v/631d1wpb/jgzheng/SCOPEs>

Size
Color
Orientation
Position
Texture
Shape



Bertin's Original Visual Variables in his book “The Semiology of Graphics” (Jacques Bertin, 1967)

[Reference] The Three Elements of Visual Mapping for Information Visualization
<https://www.interaction-design.org/literature/article/visual-mapping-the-elements-of-information-visualization>



Presentation

- Presentation of data is not directly related to data values, but it impacts the overall user experience of the visual and affects perception and cognition as well
- Common presentation perspectives
 - Design choice, like color scheme, icon set, theme, animation, etc.
 - Styling for attention shaping and reasoning, including grouping, sorting, etc.
 - Annotation: extra help info, which may include the design of titles, labels, legends, tool-tips, etc.
- From a software app perspective
 - UI, composition, layout, position
 - Interaction

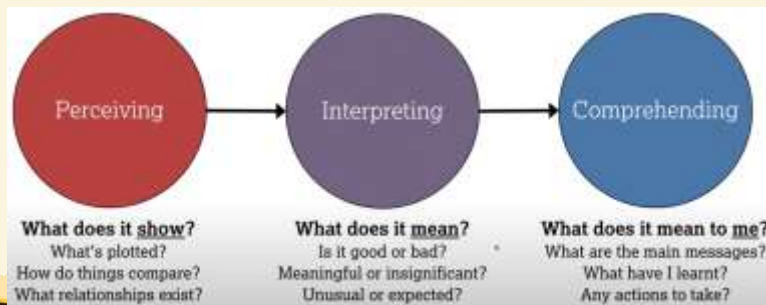
More details about attention shaping and grouping are in IT 7113 module 2



Purposes of Data Visualization

- Visualizing is basically a human physiological and psychological capability, and plays an important role in human information behavior. Enhancing perception and cognition are the two fundamental purposes and values of data visualization
- Perception
 - Perception is the organization, identification, and interpretation of sensory information in order to represent and understand the presented information, or the environment.
<https://en.wikipedia.org/wiki/Perception>
 - The most basic purpose of data visualization is sensory enhancement to impact human perception.
 - Enable fast perception (of patterns) based on instinct.
 - Help to shape the attention and focus on key things.
- Cognition
 - Is the interpretation and comprehension of the perceived information and relate to long term experience and knowledge for sense making.
 - Ease the cognitive load of information processing and exploration (especially those with a space/position factor)
 - Recall or memorize data more effectively (long term memory)
 - Extract/provoke additional (implicit) perspectives and meanings (related to knowledge and experience)
 - Leads to improved explanation, communication, problem solving, decision making, etc.

More details about perception and cognition in IT 7113 module 2



Extended reading/watching
Andy's Webinar

<https://www.youtube.com/watch?v=GVkXbQOzKNs&t=233s>

Visualizing “Data Visualization”



Plain (tabular) data

Car/Bus/Car/Bus/Bus	Product ID	Value	Unit Cost	Unit Price
Blue Sedan	001-001	\$14,400	\$14,400	\$14,400
Black Sedan	001-002	\$14,400	\$14,400	\$14,400
White Sedan	001-003	\$14,400	\$14,400	\$14,400
Blue Sedan	001-004	\$14,400	\$14,400	\$14,400
Black Sedan	001-005	\$14,400	\$14,400	\$14,400
White Sedan	001-006	\$14,400	\$14,400	\$14,400
Blue Sedan	001-007	\$14,400	\$14,400	\$14,400
Black Sedan	001-008	\$14,400	\$14,400	\$14,400
White Sedan	001-009	\$14,400	\$14,400	\$14,400
Blue Sedan	001-010	\$14,400	\$14,400	\$14,400



Visualized data



Image from <http://prezi.com/qvhyfup5z7yz/dashboard-design-making-reports-pop/>

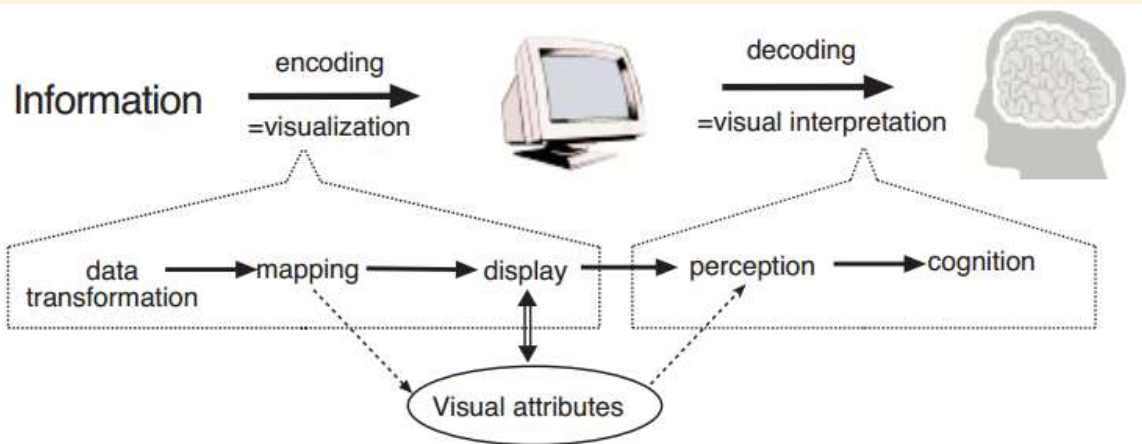


Figure “The Visualization Process” from: A survey, classification and analysis of perceptual concepts and their application for the effective visualisation of complex information
<http://dl.acm.org/citation.cfm?id=1082104>

The cognitive visualization process in human brain and elements of visual mapping is covered with more details in module 2

<https://www.edocr.com/v/e6ql9njin/jgzhen g/data-visual-foundation>

Key readings

- Visual Mapping – The Elements of Information Visualization:
<https://www.interaction-design.org/literature/article/visual-mapping-the-elements-of-information-visualization>
- The Properties of Human Memory and Their Importance for Information Visualization:
<https://www.interaction-design.org/literature/article/the-properties-of-human-memory-and-their-importance-for-information-visualization>

Practical Benefits of Data Visualization

A picture is worth
1000 words (clicks)

- Data visualization helps with higher-level tasks or information behaviors:
 - Exhibitory: showing (presenting) for quick understanding
 - Explanatory: analysis, insight generation, problem solving, decision support
 - Exploratory: interactive information seeking, browsing, exploration, and discovery
 - Communicative: e.g., story telling (impression/persuasion)
- **Note:** data visualization are also used for artistic (beauty) expression and appreciation, entertaining and for fun, which we will not cover in this class.
- More specifically (see examples in the following slides)
 - Direct perception of measurement magnitude (especially for abstract measures) based on size, color, etc.
 - Identify patterns and trends
 - Identify structures and relationships, especially those hard to express in words
 - Quickly detect and focus on area of interest or area of difference
 - Associate with familiar real-world context (e.g., locations and maps)



Perception of Value Magnitude

- Value recognition directly based on visual properties like color, size, position, etc.

Average Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	62	43	46	53	62	70	77	79	73	63	53	45	

Average High Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	72	52	56	64	73	80	86	89	82	73	63	54	

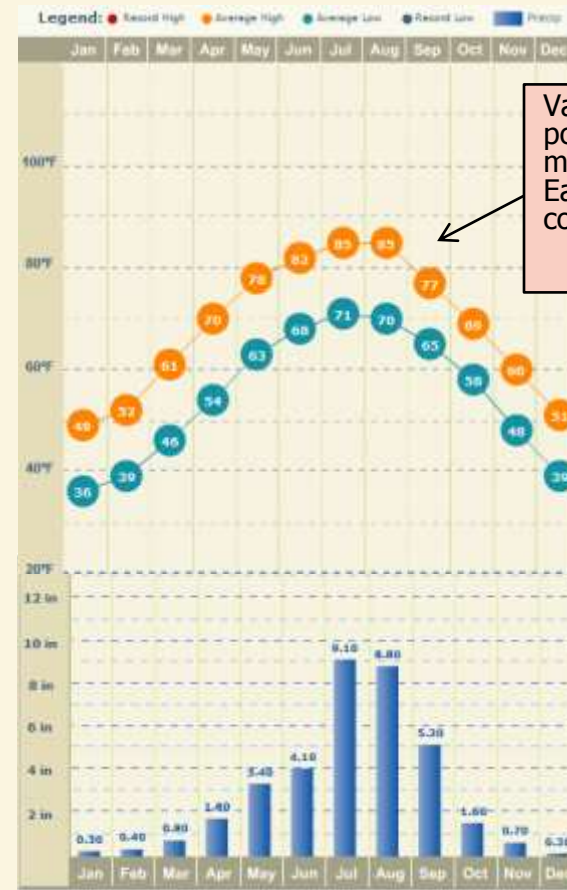
Average Low Temperature													Years on Record: 50
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	52	33	36	43	51	59	67	70	64	53	42	36	

Highest Recorded Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	105	79	80	89	93	95	101	105	102	98	95	84	79

Lowest Recorded Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	-8	-8	5	10	26	37	46	53	55	36	28	3	---

Average Precipitation													Years on Record: 50
in.	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
	50.3	4.7	4.6	5.6	4.1	4	3.7	5.3	3.7	3.6	3	4	4.1

Not everyone has a sharp sense of comparing numbers (especially a lot of numbers)



Identify Trends and Patterns



Average Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	62	43	46	53	62	70	77	79	73	63	53	45	

Average High Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	72	52	56	64	73	80	86	89	82	73	63	54	

Average Low Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	52	33	36	43	51	59	67	70	64	53	42	36	

Highest Recorded Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	105	79	80	89	93	95	101	105	102	98	95	84	79

Lowest Recorded Temperature													Years on Record: 53
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	-8	-8	5	10	26	37	46	53	55	36	28	3	---

Average Precipitation													Years on Record: 10
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
in.	50.3	4.7	4.6	5.6	4.1	4	3.7	5.3	3.7	3.6	3	4	4.1

What's the difference between these two cities? Which one is Atlanta? In 10 seconds?

The visualization adds a shape (an arch) perception that represents a change pattern.

Average Temperature													Years on Record: 13
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	62	43	46	53	63	71	75	78	71	64	54	45	

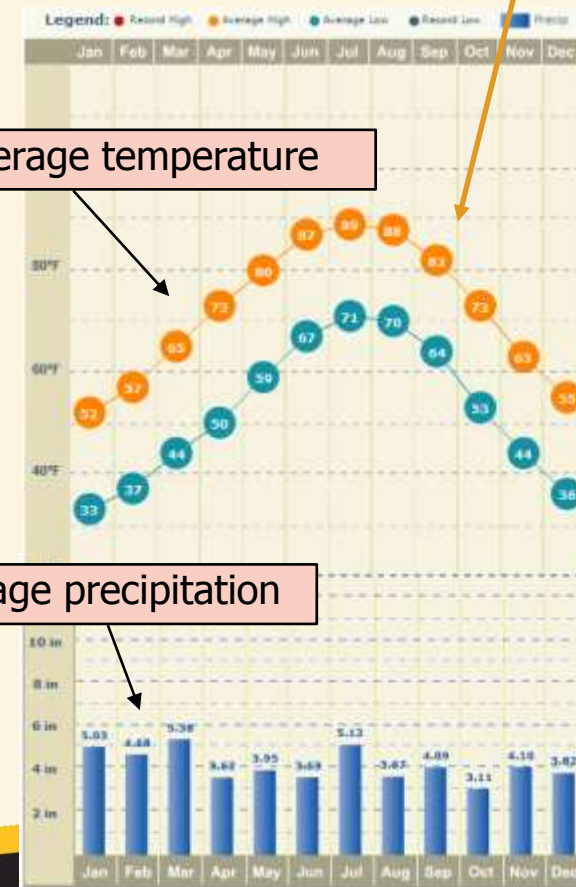
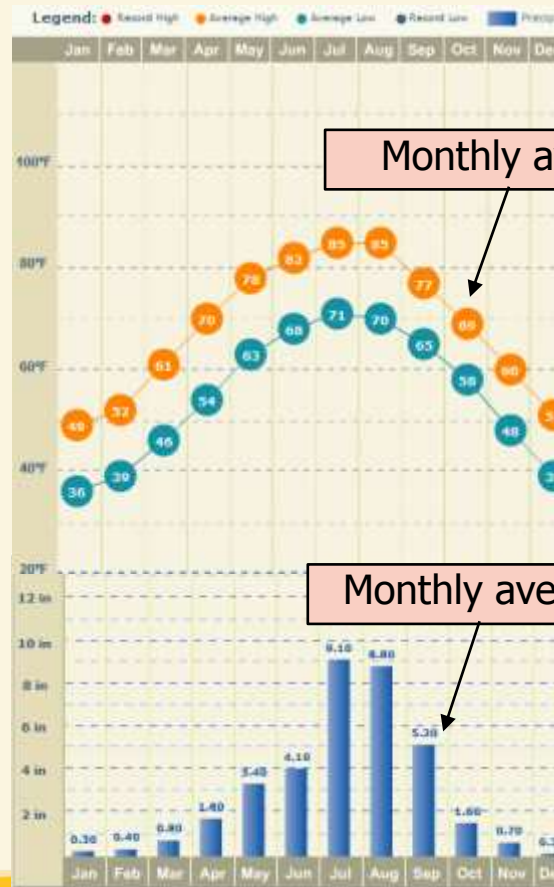
Average High Temperature													Years on Record: 13
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	67	48	51	59	70	77	81	84	76	68	59	50	

Average Low Temperature													Years on Record: 13
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	56	37	41	47	55	63	69	72	71	66	59	49	40

Highest Recorded Temperature													Years on Record: 13
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	104	68	73	84	90	93	104	104	100	91	84	81	77

Lowest Recorded Temperature													Years on Record: 13
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
°F	21	25	27	30	37	49	55	57	55	50	36	32	21

Average Precipitation													Years on Record: 81
YEAR	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
in.	38.8	0.3	0.5	0.8	1.9	3.4	4.3	9.5	10	4.9	1.8	0.6	0.2



Identify Structures/Relationships



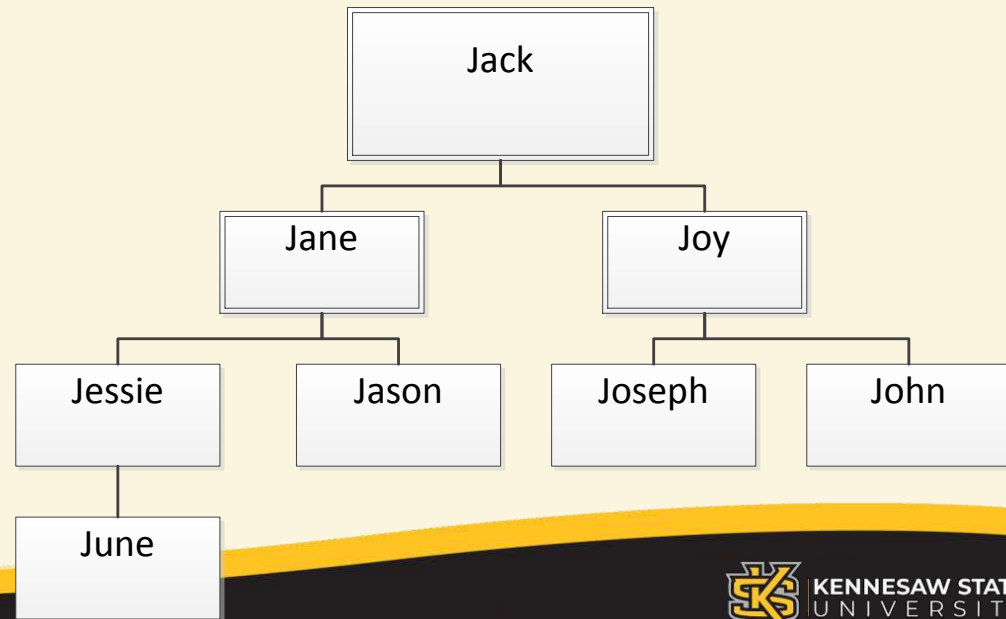
- Who Is Fighting Whom

http://www.slate.com/blogs/the_slatest/2015/10/06/syrian_conflict_relationships_explained.html



- Does June report to Joy?

Employee	Reports to
Jane	Jack
Jessie	Jane
Jason	Jane
John	Joy
Joseph	Joy
Joy	Jack
June	Jessie





Quickly Focus on Area of Interests

- Which stock performed different from others?



ACE	64.32	▼	-0.16	-0.25%
AGILENT	48.64	▲	1.90	+4.07%
ALCOA	16.64	▲	0.20	+1.22%
APPLE	342.41	▲	4.55	+1.35%
AMERISOURCEBER	40.49	▲	0.72	+1.81%
ABBOTT LABS	50.80	▼	-0.29	-0.57%
ADOBE	33.69	▲	0.04	+0.12%
ANALOG DV	38.98	▲	1.64	+4.39%
ARCHER DLS	35.81	▲	0.85	+2.43%
ADP	52.76	▲	0.89	+1.72%
AUTODESK	44.35	▲	1.92	+4.53%
AMEREN	28.68	▲	0.40	+1.41%
AM ELECTRIC	35.53	▲	0.40	+1.14%
AES	12.88	▲	0.41	+3.29%
AETNA	37.02	▲	0.22	+0.6%
AFLAC	53.22	▲	1.20	+2.31%
ALLERGAN	77.66	▲	1.17	+1.53%
AIG	32.35	▲	0.23	+0.72%
APT INVEST	25.87	▲	0.61	+2.41%
ASSURANT	37.65	▲	0.56	+1.51%
AKAMAI	39.94	▲	0.32	+0.81%
AK STEEL	16.18	▲	0.10	+0.62%
ALLSTATE	31.18	▲	0.13	+0.42%
ALTERA	45.63	▲	2.85	+6.66%

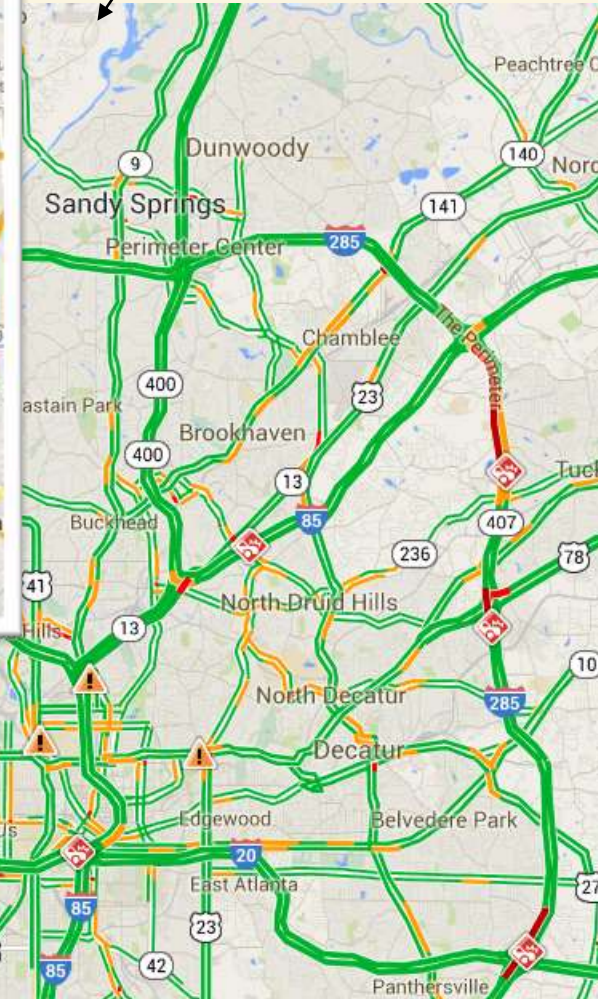
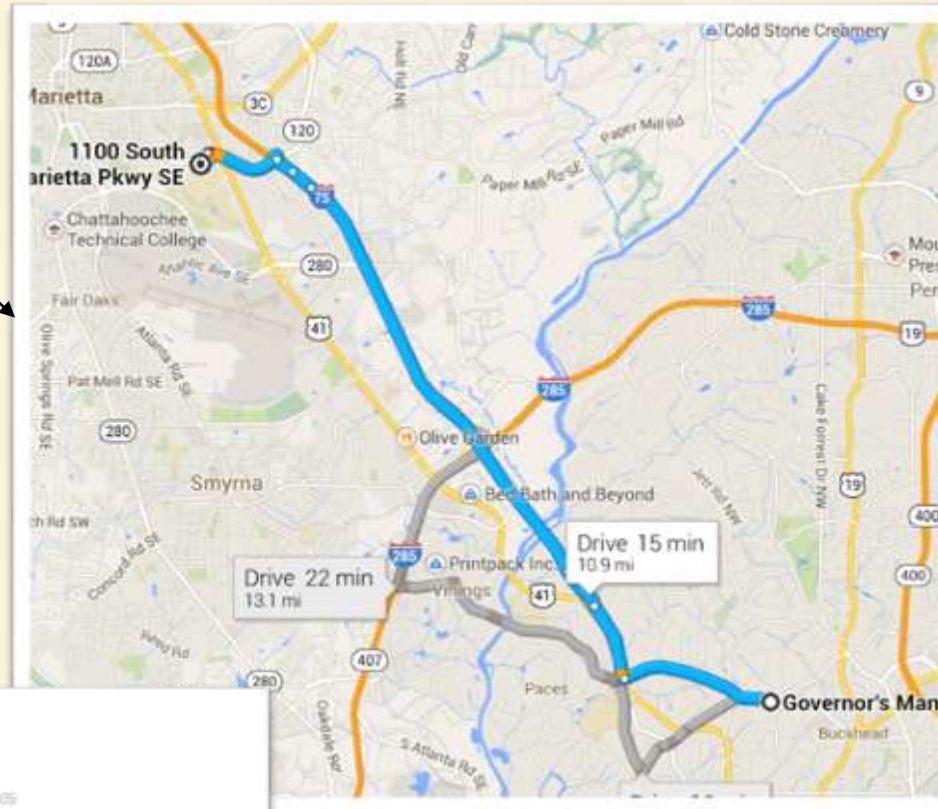
<http://finviz.com/map.ashx>

Associate with Familiar Visual Context

Geo map is the recognizable familiar real world visual context (geo map)

Visual

Text (non visual)



Drive 10.9 mi, 15 min

○ Governor's Mansion
291 W Paces Ferry Rd, Atlanta, GA 30305

- Get on I-75 N from West Paces Ferry Rd NW and Northside Pkwy
- Continue on I-75 N to Marietta. Take exit 263 from I-75 N
- Follow South Marietta Pkwy SE to Technology Pkwy SE
- 1. Merge onto South Marietta Pkwy SE
- 2. Turn left onto Technology Pkwy SE. Destination will be on the right

● 1100 South Marietta Pkwy SE
Marietta, GA 30060



Data Visualization Usages

- Broadly used in almost all activities and industries
 - Business reporting in many segments: retail, investment, manufacturing, logistics, etc.
 - Public communication: media, journalism
 - Public management like city planning, election
 - Economics <http://visualnomics.com>
- Sample use cases (depending on the use case, we may need different tools with different features)
 - Presentation and communication
 - Static presentations in meetings or conferences – PowerPoint
 - Reporting
 - Regular/seasonal reports for casual business users and manager – reports, slides
 - Real-time or near real-time reporting - dashboard
 - Interactive reporting and exploration by power users – interactive reports or dashboard.
 - Executive reporting and decision making - dashboard
 - Analytical
 - Used in the process of analysis, accompanying queries and calculations - Excel
 - Advanced visual driven analysis, often used for research – Power BI/Tableau
 - Monitoring: real-time operational monitoring (driving, manufacturing) - dashboard
 - Public communication/journalism
 - School redistricting <https://www.greatschools.org/school-district-boundaries-map/>
 - Tell a story to the public <https://www.vox.com/2018/1/8/16822374/school-segregation-gerrymander-map> - web
 - Demonstration/simulation: interactive demonstration for complex scenarios - <http://setosa.io/bus/>



A Bigger Impact

- Visual literacy
 - The ability to evaluate the advantages and disadvantages of graphic representations, to improve their shortcomings, to use them to create and communicate insights, or to devise new ways of exploring and representing data, information, or knowledge
 - Enables the group of managers, technical professionals, and users to effectively and efficiently communicate and understand on a common ground.
- Data culture
 - Data visualization helps with creating a data culture, especially in an environment lack of analytical skills.
 - It facilitates collaboration among different groups of people with different professional focus, thus creating better working environment.
 - With modern visual tools, data visualization make data and analytics more accessible to average people. This contributes to data literacy and data democratization.



Related Terms and Fields

Data presentation

Information design

Information visualization, including infographic and illustration

Text visualization

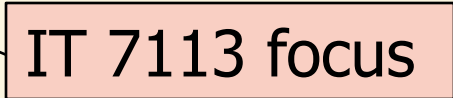
Big data visualization

Computer graphics, reality visualization (VR, AR), scientific visualization

Data art

Business data visualization

Visual analytics

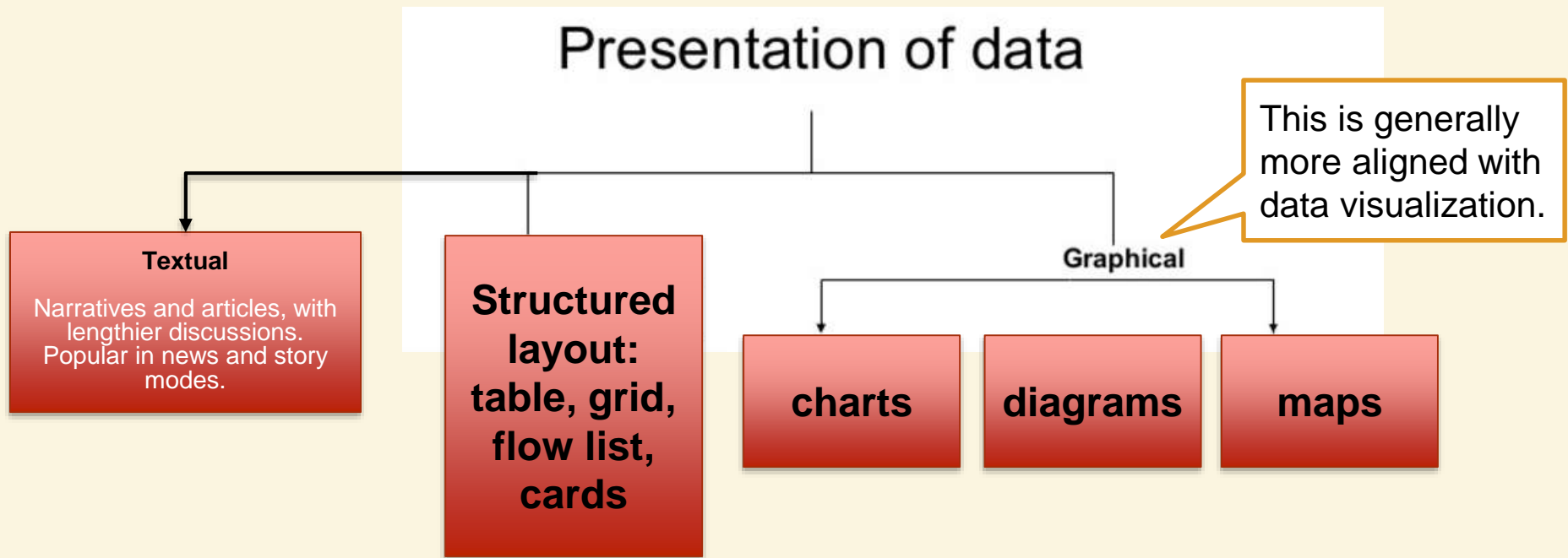


IT 7113 focus



Data Presentation

Data presentation is the method to summarize, organize, and communicate data (raw or analysis results) using a variety of tools. Data can be presented in one of the three forms: text, tables, and/or graphs. The selection of the method of presentation depends on the type of data, method of analysis, and type of information sought from the data.



Reference reading:

- <https://www.toppr.com/guides/economics/presentation-of-data/textual-and-tabular-presentation-of-data/>
- <https://www.toppr.com/guides/business-economics-cs/descriptive-statistics/diagrammatic-presentation-of-data/>



Information Design

- **Information design** is the practice of *presenting* information in a way that fosters an efficient and effective understanding of the information.
 - https://en.wikipedia.org/wiki/Information_design
- These include elements like layout, flow, use of text style, bullets, spacing, etc.
- It shares some similar data visualization design practices, but it typically does not rely on graphics heavily.
- Information also include more types of content than just data.
- Applicable fields
 - Document design, presentation slides, user interface, product design, web design, print media design like books/magazines, and **information visualization/graphics**, media imaging, map, etc.

Information
design



Information
visualization



Information Visualization

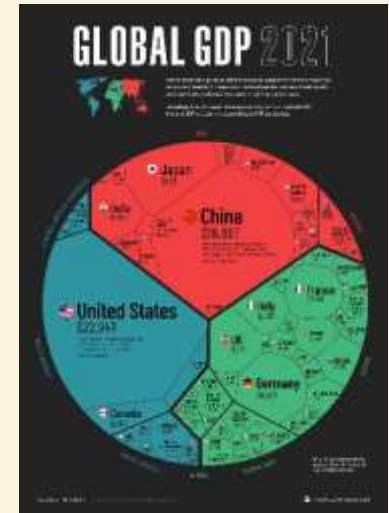
- **Information visualization** is the study of visual representations of information or data to reinforce human cognition. The data include both numerical and non-numerical data, such as text and geographic information.
 - A branch of information design that utilizes graphical elements
 - A very close field, and very often used as the synonym for, or even include, data visualization
 - Often in the form of illustrations and infographics
 - https://en.wikipedia.org/wiki/Information_visualization
- Examples: mall/subway map, more at <https://informationisbeautiful.net>
- See more resource about information visualization
 - <https://www.interaction-design.org/literature/topics/information-visualization>
 - <https://www.informationisbeautifulawards.com/showcase?type=awards>



Infographics



- Infographics is a specific type of information visualization that are usually a mixture of texts, graphics, and data visual forms (charts, diagrams, tables, maps, etc.) to quickly and vividly communicate complex information (multiple variables or dimensions).
 - http://en.wikipedia.org/wiki/Information_graphics
 - <https://visual.ly/blog/11-infographics-about-infographics/>
 - Often used in mass communication (e.g., journalism) and marketing
 - <https://www.business2community.com/digital-marketing/visual-marketing-pictures-worth-60000-words-01126256>
 - <https://www.interaction-design.org/literature/article/information-visualization-who-needs-it>
 - Single block visualization (may be big)
 - <https://www.visualcapitalist.com/visualizing-the-94-trillion-world-economy-in-one-chart/>
 - <https://www.visualcapitalist.com/visualized-world-leaders-in-positions-of-power/>
 - Composite infographics
 - <https://www.easel.ly/blog/endangered-species-infographics/>
- More examples:
 - <https://www.visualcapitalist.com/our-top-21-visualizations-of-2021/>
 - <https://www.visualcapitalist.com>
 - <https://visual.ly/m/design-portfolio/>
 - <https://informationisbeautiful.net> (not all but many are)
 - <http://www.visualisingdata.com> (not all but many are)
 - <http://courses.ischool.berkeley.edu/i247/s18/> (not all but many are)

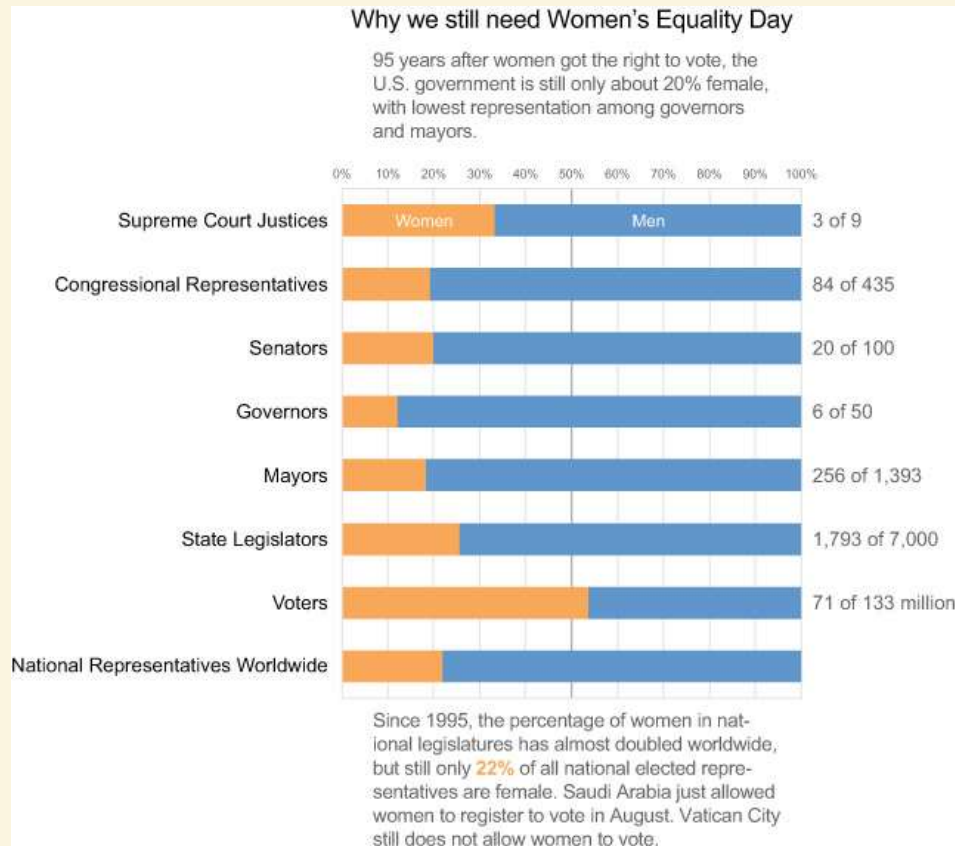


<https://www.visualcapitalist.com>

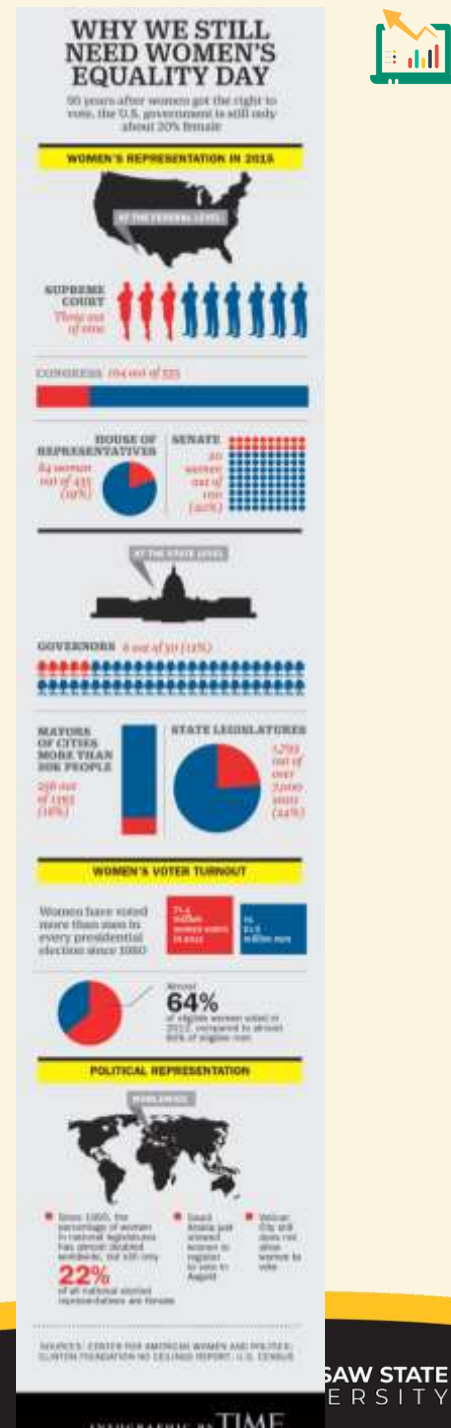


Data Visualization vs. Infographics

<https://www.datarevelations.com/balancing-accurate-and-tone/>



VS.





Infographics vs. Data Visualization

	Infographic	Data visualization
Creation	One time creation and use; mostly created using graphic design tools. Often hand-crafted.	Using data processing or analytical tools; automatically populated from a data source.
Usage	Intended for more casual use (informational) for general people. Presentation only.	Allows interactive exploration and supports analytical needs and decision making.
Data (binding)	Fixed data set or numbers. Information often is more general and can be more qualitative.	Highly quantitative with many measures and metrics. Visualizations are bound to a data source and automatically populated. Data sources can be dynamic.
Visual styles	Utilizes more free forms (non-standard) of visual diagrams or illustrations (illustrational diagrams); emphasizes creativity and artistically expression to communicate or impress casual viewers.	Uses more standard and conventional visualizations that are already familiar to the audience.

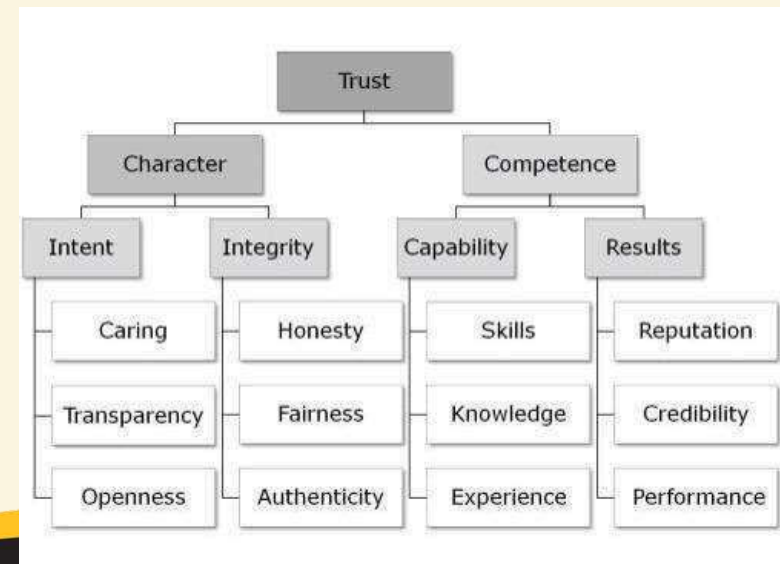
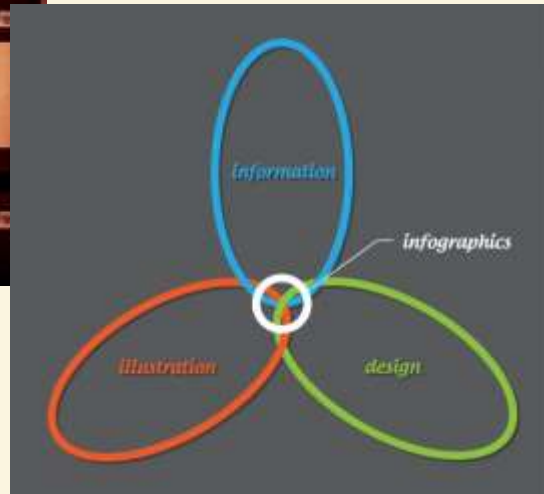
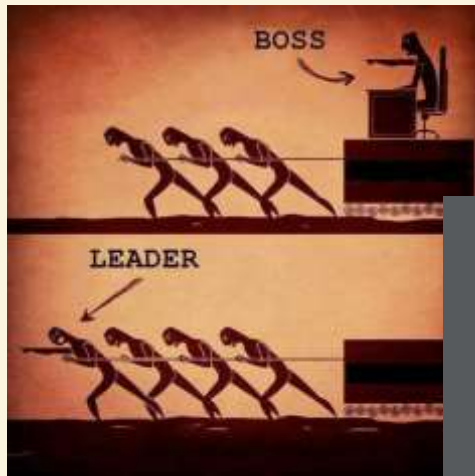
More readings: infographics vs. data visualization

- <https://visage.co/throwdown-data-visualization-vs-infographics/>
- <http://www.jackhagley.com/What-s-the-difference-between-an-Infographic-and-a-Data-Visualisation>

Illustration (of an Idea/Concept)



These can be considered as information visualization but **NOT** data visualization.



More examples of information visualization can be found at

- <https://informationisbeautiful.net>
- <http://dailyinfographic.com/>
- <https://www.visualcapitalist.com/>
- <http://www.cooldailyinfographics.com/>
- <http://blogs.scientificamerican.com/sa-visual/2014/10/14/sa-recognized-for-great-infographics/>
- <https://visual.ly/m/design-portfolio/>
- <http://www.dubberly.com/concept-maps/3x4grid.html>



Text Visualization

- Text visualization is related to text analytics, focusing on its visual presentation part – text as data
 - Articles, documents, web pages, logs, emails, messages, etc.
 - <https://courses.cs.washington.edu/courses/cse512/19sp/lectures/CS512-Text.pdf>
- Common visual forms for text analytics
 - [Word/tag cloud](#), [word tree](#)
 - More: <https://textvis.lnu.se>
- It is also related to information presentation, which is about enhancing of textual information presentation and understanding through decoration of text, using similar visualization techniques found in visualizing data.
 - Examples: <https://chartexpo.com/blog/text-visualization-examples>
- In some sense, text visualization can be related to typography (individual letters and characters)
 - <https://microvis.info/thesis/#foundation>



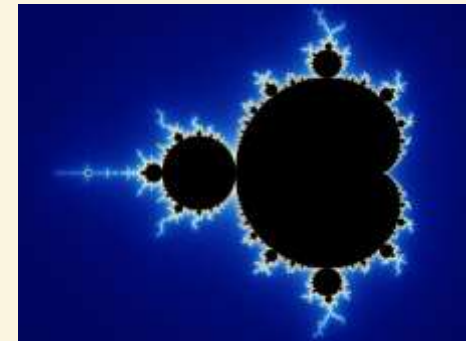
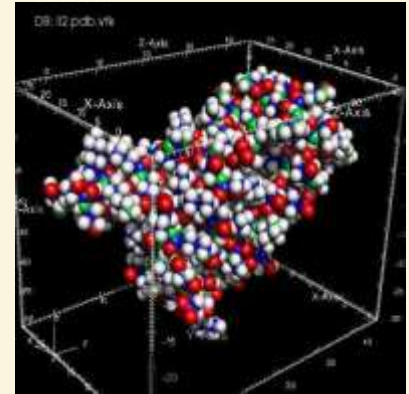
Big Data Visualization

- Big data visualization usually refers to a visualization with a large number of data points (items and attributes) on a large space.
 - The goal is to see patterns and relationships beyond a few items or aggregated metrics
 - Using more contemporary visualization techniques including real-time changes, animations, rich interactions, etc.
 - Using more illustrative graphics and more artistic visual representation of the data.
- Examples:
 - <http://projector.tensorflow.org>
 - <https://www.mastersindatascience.org/resources/10-cool-big-data-visualizations/>
- Is big data visualization a hype?
 - “Big data visualization is generally not helping humans make novel discoveries.”
https://jeffjonas.typepad.com/jeff_jonas/2016/02/data-visualization-outing-hype.html
- What's the purpose of big data visualization? And what's the effective way to use them?
 - https://pudding.cool/2018/10/city_3d/
 - <https://pudding.cool/2019/07/book-covers/>

Computer Graphics and Scientific Visualization



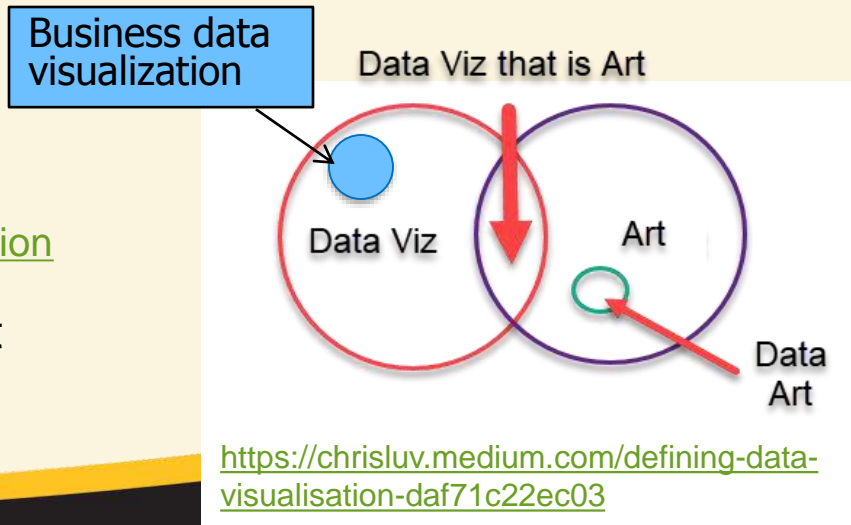
- Computer graphics (CG) is computer generated graphics and image models
 - CG do not feature the use of business or transactional data that are generated from human or organizational activities.
 - Commonly based on computing logic and algorithms.
 - Usually more computing intensive.
- Applications
 - Computer image
 - 3D model, virtual reality
 - Science (scientific visualization)
 - Games and movies
 - Simulation - <https://ncase.me/polygons/>
- Physical science visualization
 - “Primarily concerned with the visualization of complex three-dimensional phenomena (architectural, meteorological, medical, biological, etc.), where the emphasis is on realistic renderings of volumes, surfaces, illumination sources, and so forth.”
 - Visualization (simulation) of complex reality (entities or events, such as universe, sun, explosion, atom, climate, etc.)
 - https://en.wikipedia.org/wiki/Scientific_visualization
- Mathematical model/algorithm visualization/simulation – the visualization created based on math calculations and models
 - <http://acko.net/blog/how-to-fold-a-julia-fractal/>
 - <https://en.wikipedia.org/wiki/Fractal>
 - <https://mathigon.org/world/Fractals>
 - <https://ncase.me/polygons/>



Data Art



- Data Art (or Information Art) is a field where artists express themselves artistically using data as a medium (usually using computers).
- In data art, visualizations of data seek primarily to entertain or produce an aesthetic experience. It is art that is based on data.
 - <https://www.perceptualedge.com/blog/?p=1245>
- Features
 - The input raw material is still data, but the purpose is not to understand or comprehension
 - Often using algorithm (simple or complex) to brush up data (visualize) for self-expression or aesthetics expression
 - The understanding of the product is largely subjective and leave multiple ways of interpretation
- Examples
 - <http://hint.fm/projects/flickr/>
 - <https://nightingaledvs.com/tag/data-art/>
 - <https://www.data-to-art.com>
 - Music visualization
https://en.wikipedia.org/wiki/Music_visualization
- Note: artistic data visualization != data art



Business Data/Information Visualization



IT 7113
focus

- Business data visualization is the data visualization mainly related to business data, and it is used for business activities and purposes.
- **Business** is a general term to describe activities, events, and operations of an entity
 - Business includes many activities directly associated with human, like commerce, public service, education, sports, charity, entertainment, etc.
 - An entity could be an individual, organization, company, government, etc.
 - Or even some natural phenomenon and events that impact human, such as ecosystem, weather, universe, animal, etc. (to some extent)
 - Business data/information records various aspects of these activities and events.
- Main purposes are data exploration, analysis, decision support, monitoring, and communication that is related to business performance.
- Main features of business data
 - *Abstract*: this data mainly describes an activity, pattern, trend, etc.; it does not directly define or create (simulate) a real-world object or phenomenon as close as possible.
 - Mostly *quantitative*
 - Often structured or semi-structured, repeated
 - Can be aggregated in multi-dimensions
 - Directly comprehensible by skilled humans (in a particular “business”)
- Main features of business data visualization
 - It is part of a BI or analytics process and system, especially in self-service systems
 - Using simple, standard, and abstract images (symbol/chart/diagram/map)
 - Highly reused and commonly accepted visualization forms – following standard practices <https://www.ibcs.com/standards>
 - Utilizes data binding techniques to generate visualizations in an automated way (using specialized data visualization software or as part of an analytics software application)



Business Data Visualizations Examples

- Operational reports: communication of results in all kinds of reports (periodical/seasonal or real time) and presentations (e.g. PowerPoint)
 - <https://myit-2019.itdashboard.gov>
- Performance dashboards
 - <https://www.geckoboard.com/learn/dashboard-examples/>
- Visual data exploration and seeking
 - <https://www.productchart.com/smartphones/>
 - <https://finviz.com/map.ashx>
 - <https://www.census.gov/dataviz/>
- Visual analytics
 - <https://www.google.com/publicdata/directory>
- Real-time monitoring (public events, administrative, or operational)
 - <https://www.nytimes.com/interactive/2018/11/06/us/elections/results-dashboard-live.html>
- Presentation of results in statistical analysis, data mining or other advanced analytics.



[Not] Business Data Visuals

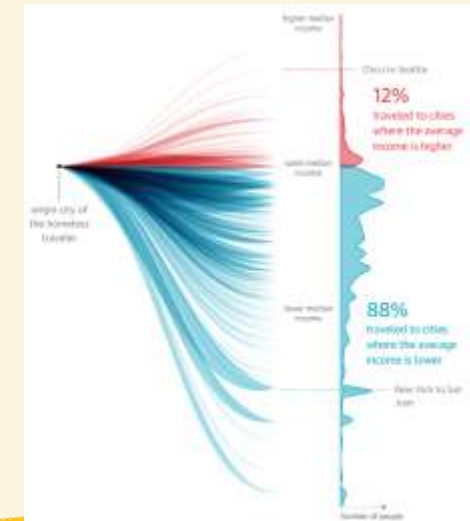


These examples are not really considered to be business data visualization, but still considered as general data visualizations – not the focus of this class

- General data visualization (not typical business data related to business activities)
 - <http://hint.fm/wind/>
 - <http://hint.fm>
 - <http://classes.dma.ucla.edu/Spring13/161/projects/students/david/project-5/html/?/image-gallery/>
 - <https://weather.com/weather/radar/interactive/I/USGA0028:1:US>
- Artistic data visualization: with many artistic decorations – commonly used in journalism style report for public communication
 - <https://www.theguardian.com/us-news/ng-interactive/2017/dec/20/bussed-out-america-moves-homeless-people-country-study>
 - Visual embellishment <http://hci.usask.ca/uploads/173-pap0297-bateman.pdf>
- Not even data visualizations
 - Infographics
 - Mathematical visuals
 - Scientific visualization



<https://ritholtz.com/2012/02/the-beatles-song-keys/>





Visual Analytics

- Visual analytics is "the science of analytical reasoning facilitated by interactive visual interfaces."
 - https://en.wikipedia.org/wiki/Visual_analytics
- **Visual analytics is beyond just visualizing data**
 - Interactive exploratory and analytical processes
 - The major purpose is to discover patterns and relationships
 - Visual analytics does not just visualize raw fact data or a few performance measures; it involves complicated metrics and statistical measures.
- **Example:** <https://www.tradingview.com/chart/>
- **Use visual analytics tools or analytical dashboard – see visual forms in the next section.**

Extended reading: <https://www.sisense.com/blog/data-visualization-and-visual-analytics-seeing-the-world-of-data/>



Comparison of Related Visualization Fields

	Content	Visual Forms/Tools	Purpose/Usage
Business data visualization	Quantitative data related to business activities; metrics, key performance indicators (KPIs)	Standard and common types of charts, diagrams, maps, dashboards	Data exploration, analysis, decision making
General data visualization	General quantitative data, such as natural phenomenon	Also often using creative and stylish charts, diagrams, maps; artistically combining different forms and elements.	Data exploration, cognition, and mass communication
Information visualization	All kinds of information, quantitative and qualitative	Infographics, illustrational diagrams	Information seeking, artistic illustration, casual communication, story telling
Illustration	Processes, structures concepts, ideas	Diagrams, images, graphics	Making the content more vivid and engaging, easier to understand the complexity.
Scientific visualization/simulation	Real world object or phenomenon, mathematical functions and formulas, calculated data based on formulas or rules	Computer generated graphics, 3D virtual reality, animated diagram	Recreate or simulate the real-world object or phenomenon, or visualize an algorithm effect. Demonstrate the effect of scenarios under certain rules.
Visual analytics	Quantitative data; statistical and other metrics	Charts, diagrams, maps, dashboards	Analysis and decision support



IT 7113 focus



Data Visualization in BI/Analytics

Basic visual properties and visual forms/styles used in BI and analytics applications



Data Visualization in BI/Analytics

- Data visualization is an important part of data exploration and decision making. Given the power of visualization, it is only natural to apply the rich communication techniques in the field of BI and analytics.
- Visualization has been considered as a separate field from BI in the early days (prior to 2010), but it quickly brought the traditional business intelligence to life
 - As organizations seek to empower non-technical users to make data-driven decisions, they must consider the powers of data visualization in delivering digestible insights.
 - Visualization tools have become increasingly important to business intelligence, in which people need technology support to make sense of and analyze complex data sets and all types of information.

Dashboards, reporting, end-user self-service, and advanced visualization top the most important technologies and initiatives strategic to BI in 2018.

<https://www.forbes.com/sites/louiscolumbus/2018/06/08/the-state-of-business-intelligence-2018/#b2fca2878289>

Technologies and Initiatives Strategic to Business Intelligence

(Copyright 2018 – Dresner Advisory Services)





Basic Visual Forms/Styles used in BI/Analytics

Form/Style	Description	Typical Types and Examples
Embedded visual	It is embedded in, or directly on top of, texts and other forms of data presentation (table, graphic, etc.).	<ul style="list-style-type: none">• Conditional formatting (visual cues)• Inline chart (Sparkline)
Block visual	It is displayed as an independent visual unit and occupies a larger space. It is often a part of another product such as report or dashboard, appearing together with other content. But sometimes it can become a standalone visual with many data points or enough complexity.	<ul style="list-style-type: none">• Chart• Illustrational diagram• Map (smaller)• Table/card with embedded visuals
Visual app	It is a standalone application. It may consist of a combination of different types of visuals and other content types. It supports more complex interactions, and most interactions are within the application.	<ul style="list-style-type: none">• Dashboard• Visual report/presentation• Infographic (large)• Visual analysis tool (or an analytical dashboard)• Full map (full screen)• (Data-driven) Story

Refer to Zheng (2017) Book Chapter Data Visualization in Business Intelligence (PDF downloadable on the site): https://www.researchgate.net/publication/321804138_Data_Visualization_for_Business_Intelligence

Conditional Formatting

- Conditional formatting
 - Direct formatting on text or numbers using visual properties, embedded in a pre-established presentation
- Example
 - Golf http://www.masters.com/en_US/scores/

HOLE	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	TOTAL
PAR	4	5	4	3	4	3	4	5	4	4	3	5	4	5	3	4	4	4	72
RD 1	4	5	4	3	4	2	4	5	4	5	5	3	4	3	5	3	4	4	71
RD 2	5	4	5	3	4	2	5	4	3	3	4	3	4	3	4	3	4	3	66
RD 3	5	5	3	4	4	3	4	5	4	4	5	3	4	4	5	3	4	5	74
RD 4	4	4	3	4	4	2	3	3	4	4	4	4	5	4	4	3	4	4	67

Legend:
 ○ EAGLE +
○ BIRDIE
□ BOGEY
□ DBL BOGEY -
VIDEO
○ EXAMPLE: BIRDIE WITH VIDEO

CLOSE

- Tag cloud



Sparkline

- A sparkline is a small chart embedded in a context of words, numbers, tables, images, or other type of information.
 - It presents the general shape of the variation in a simple and highly condensed way.
 - <https://chartio.com/blog/new-chart-type-sparklines/>
- Examples
 - <http://omnipotent.net/jquery.sparkline/>
 - <http://www.klipfolio.com/blog/table-component-overview>
 - <https://trumpexcel.com/sparklines/>



Sparkline

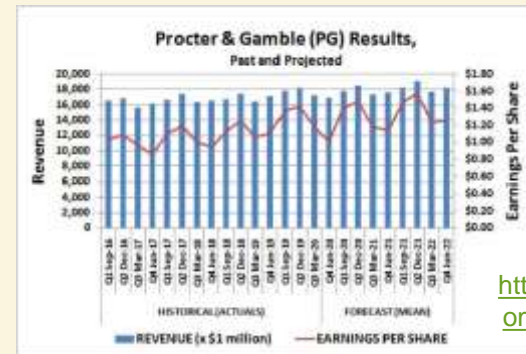
Natural Gas	561.29	-4.10	-0.72%	
Network	273.20	+6.03	+2.26%	
Oil	1,105.34	-0.46	-0.04%	
Oil Services	214.90	-5.60	-2.54%	



Chart



- A chart is a graphical representation of data
 - Chart is a unique combination of symbols (visual elements) with visual properties which directly represents quantitative values
 - <http://en.wikipedia.org/wiki/Chart>



<https://fool.whotrades.com/blog/43151739255>

- Chart vs. Diagram
 - These two terms are very similar; they are often used together or interchangeably in daily life.
 - Chart is more abstractly presented and focuses more on quantitative values.
 - Diagrams also cover a lot of qualitative information like process, concepts, ideas, structures, etc.; they also intergrade more real-world contexts like maps.
 - Diagram is sometimes considered to include chart.

We will cover charts and their designs in IT 7113 (three modules).

Summary of Chart Categorizations

Choosing charts will be covered in IT 7113 module 3.

- Simple – simply by purpose or data type without much explanation

Abela's version	https://extremepresentation.com/design/7-charts/	This is the most widely referred version with a simple visual itself. But it was criticized by Stephen Few with a lot of details https://www.perceptualedge.com/blog/?p=2080
Camões's version	https://excelcharts.com/classification-chart-types/	Influenced by Abela's version, added evolution (like trend) and profiling. The blog is very brief. He has an updated version with more details in his book "data at work" https://www.dataatworkbook.com .
Juice Analytics	https://www.juiceanalytics.com/chartchooser	Provided as an interactive online chooser with templates for Excel and PowerPoint. Categorized similarly but added an "trend" category.
Schwabish's Graphic Continuum	https://policyviz.com/2014/09/09/graphic-continuum/	A poster style visual presentation covering nearly 90 charts.
Financial Times Visual Vocabulary	https://www.ft.com/vocabulary	Financial Times Visual Vocabulary is based on the Graphic Continuum.
Fraconeri's version	http://experception.net	A quick reference in PDF that considers data types and inputs, rather than purposes.

- Detailed and interactive

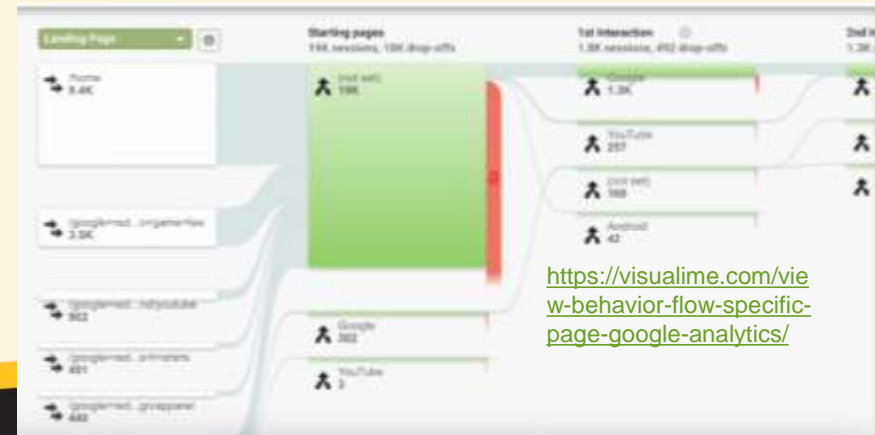
Ferdio	http://datavizproject.com	An interactive resource with a lot of examples. Included diagrams and maps. Categorization by function and a unique category by data inputs. It provides details for each chart. I use it for reference a lot.
Data catalog	http://www.datavizcatalogue.com	An interactive catalog with very detailed description for each chart. Added many smaller and specific categories. I use it for reference a lot.
From Data to Viz	https://www.data-to-viz.com	A classification of chart types based on input data format. It comes in the form of a decision tree. It also provide details for each chart.
Chart make directory	http://chartmaker.visualisingdata.com	This is a community effort to catalog charts by function and show solutions for each major visualization tool (with links to external resources).



Illustrational Diagrams

- Illustrational diagrams
 - Mainly to visualize quantitative as well as qualitative data to illustrate their features, relationships, sequences, etc.
 - Also includes position as a dimension in a logical (virtual) structure, such as network diagram, process diagram, hierarchy diagram, etc.
 - <http://en.wikipedia.org/wiki/Diagram>
 - <https://en.wikipedia.org/wiki/Illustration>
- Common examples used in business information visualization
 - Flow chart: <http://en.wikipedia.org/wiki/Flowchart>
 - Network graph: <https://flowingdata.com/charttype/network-graph/>
 - Tree diagram: http://en.wikipedia.org/wiki/Tree_structure
 - Timeline: <https://datavizcatalogue.com/methods/timeline.html>
 - Structure diagram: http://en.wikipedia.org/wiki/Data_structure_diagram
 - More: <https://datavizproject.com/family/diagram/>
- Some does incorporate quantitative data
 - Sankey diagram

https://en.wikipedia.org/wiki/Sankey_diagram





Maps (Location-based Visualizations)

- Location intelligence (LI) is a business intelligence (BI) tool capability that relates geographic contexts (usually as a dimension) to business data.
 - <http://searchbusinessanalytics.techtarget.com/definition/location-intelligence-LI>
- Location based visualization (map) is the base for location intelligence and plays an important role in business intelligence.
 - Within all of the leading front-end BI tools, interactive maps are replacing or augmenting standard table and chart views of geographic data.
 - involves layering multiple data sets spatially, for easy reference on a map
 - Maps provide context ... Quickly associate data with familiar position/location – added familiarity increase comprehension. (Tableau “The Power of Where”
<https://www.tableau.com/learn/whitepapers/government-mapping>)
 - Maps are widely recognized and understood—even by nontechnical professionals—which helps make the data represented more accessible and understood. (Forbes “The Power of Place”
https://www.forbes.com/forbesinsights/pitney_bowes_power_of_place/)
- Like business intelligence, location intelligence supports analysis and decision making. But for the past 20 years, these two data-centric disciplines have forged independent but parallel paths. Only now are they beginning to converge. The explosion of mobile and IoT devices facilitates the integration of business and location intelligence.
 - The first step toward converging location and business intelligence is plotting business metrics on a map.
 - The next step is the interactive process of location driven visual analytics, utilizing more sophisticated mapping layers and data presentation, even on three-dimensional surface, with the help of VR/AR technologies.
 - <https://www.esri.com/news/arcuser/1012/files/morethanamap.pdf>

Major Map Types



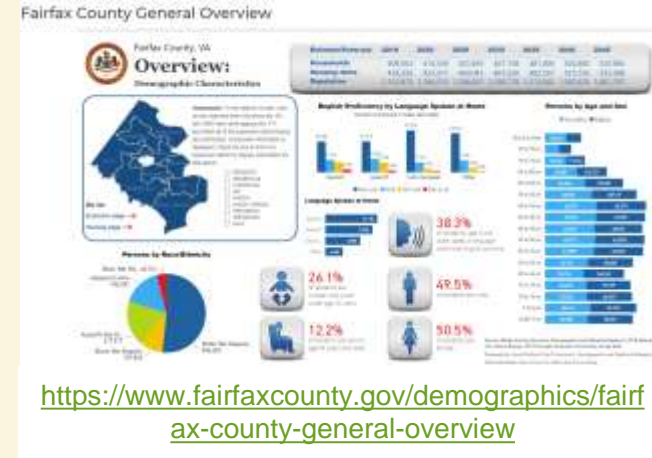
- Geospatial (geo) maps
 - Visualize geo location related data directly on real world maps
 - Data represented as points, areas, paths
 - <http://luminocity3d.org/WorldCity/>
 - <https://maps.google.com/gallery/>
- Abstract illustrational geo map (geo chart)
 - These are conceptually related to geo location but presented in an abstract way – an illustration, rather than accurate geo locations.
 - Geo chart: <https://googletrends.github.io/search-election-election-night/>
 - Sometimes they can even go symbolic like tile grid map: <https://fivethirtyeight.com/features/where-your-state-gets-its-money/>
 - Road/path map: <http://fatalities.safer63and881.com/#highway>
- Contextual map
 - Any data relevant to the positioning in a particular context or space, e.g., building, campus, mall, stadium, a just a space (like a hitting area) etc.
 - <https://www.facilityquest.com/occupancy-utilization-studies/>



Dashboard

A dashboard is a visual-oriented display of the most important data and information needed to achieve defined goals and objectives; consolidated and arranged on a single screen so the information can be viewed at a glance.

Adapted from: Dashboard Confusion, Stephen Few,
http://www.perceptualedge.com/articles/ie/dashboard_confusion.pdf



- Elements of a dashboard

Dashboard = data/information + visual + UI

- Data/information: the most important element
- Visual: data visuals (charts, etc.) provide a high level at-a-glance view
- User interface
 - a clean UI that unifies all elements to work together as a whole
 - supporting interactions as needed

- The Value of Dashboard

- Provides a one-place presentation of critical information, so users can quickly understand data and respond quickly at one place.
 - Saves time over running multiple reports.
- Allows decision makers to see a variety of data that affects their divisions or departments
 - This allows decision makers to focus only on the items over which they have control
 - The dashboard is generally customized for each user
- Allows all users to understand the analytics. For non-technical users, dashboards allow them to participate and understand the analytics process by compiling data and visualizing trends and occurrences.
- More: <http://www.bidashboard.org/benefits.html>

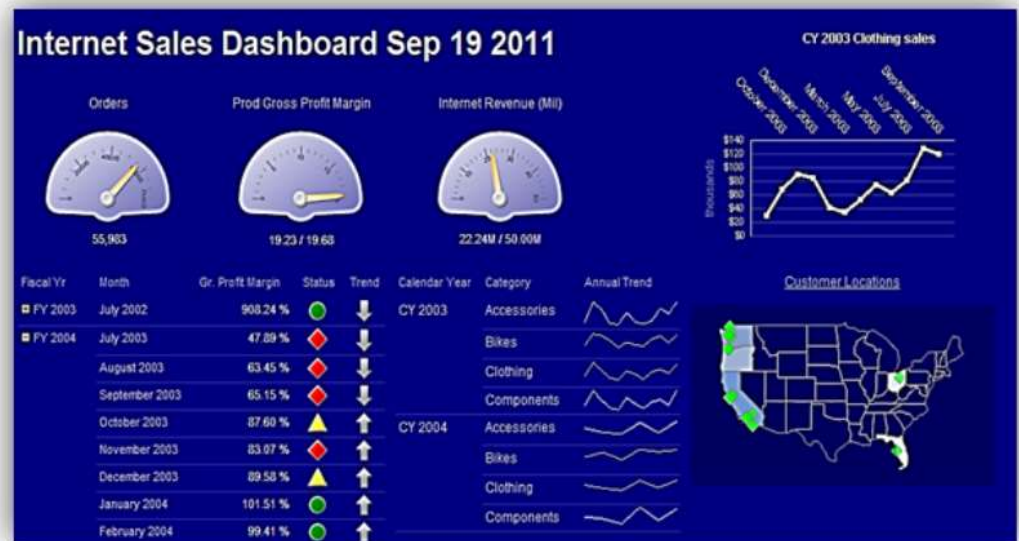
The two dashboard modules in IT 7113 provide more details:
<https://www.edocr.com/v/oekl31vr/jgzheng/Dashboard>

Dashboard vs. Report



- Reports
 - A report is the presentation of detailed data arranged in defined layouts and formats
 - Based on simple and direct queries: they usually involve simple analysis and transformation of data (sorting, calculating, filtering, grouping, formatting, etc.)
- Traditional reports contain detailed data in a tabular format and typically display numbers and text only.
 - It is geared towards people who need data rather than a direct understanding or interpretation of data.
 - Its purpose is mainly for printing (with styling) or exporting (raw data).
- Modern reports can be interactive and visual, but the focus is still on detailed data. Sometimes the distinction is a bit blurred with dashboards in some practical cases.
 - Visual intensive report: <https://datareportal.com/essential-facebook-stats>
 - A report style “dashboard” (or more like a visual intensive interactive report): <https://www.cityhealthdashboard.com/ga/atlanta/city-overview>
 - Magic Quadrant report vs. <https://www.g2.com/categories/data-visualization?segment=all>
 - Dashboard or report? <http://www.crazybikes.com/mrc/CRAZYBIKES.R00090s>

Cat/SubCat/Product	Product #	Color	Std Cost	List Price
Accessories				
Bike Racks				
High Rack - 4-Bike	RA-HY23		\$44.88	\$120.00
Bike Stands				
All-Purpose Bike Stand	ST-1401		\$69.47	\$169.00
Bottles and Cages				
Mountain Bottle Cage	BC-M005		\$3.74	\$9.99
Road Bottle Cage	BC-R005		\$3.36	\$9.99
Water Bottle - 30 oz.	WB-H006		\$1.87	\$4.99
Cleaners				
Bike Wash - Discover	CL-8009		\$2.97	\$7.95
Fenders				
Fender Set - Mountain	FE-6854		\$9.22	\$21.99
Helmets				
Sport-100 Helmet, Black	HL-U509	Black	\$13.09	\$34.99
Sport-100 Helmet, Blue	HL-U509-B	Blue	\$13.09	\$34.99
Sport-100 Helmet, Red	HL-U509-R	Red	\$13.09	\$34.99
Hydration Packs				
Hydration Pack - 70 oz.	HY-1023-70	Silver	\$20.57	\$54.99





Visual Analysis/Exploration Tool

- A visual analysis/exploration tool is similar to a dashboard in that it uses visualizations intensively to drive data exploration or analysis (visual analytics).
 - Some consider it a kind of dashboard; some consider it a bit different
- Key characteristics of a visual analysis tool:
 - The visualization is usually a single (or very few) component (a chart or a map) that occupies a big portion of the screen as the main UI component, with a large number of data points visualized.
 - It is highly interactive and usually provides abundant settings and configurations (for adjusting factors and parameters) including filtering or sorting options. In fact, the number of setting combinations can be quite big.
 - The visualization may be more complex with multiple visual layers
 - It is not to visual key metrics, but to visualize patterns, trends, and other complex relationships among data.
 - It fits on one screen, but there may be scroll bars and zooming options.
 - It is primarily used for intensive data exploration or analysis, used by data analysts and researchers.
- Examples
 - Exploration
 - <https://finviz.com/bubbles.ashx>
 - <https://finviz.com/map.ashx>
 - <https://www.productchart.com>
 - Google public data explorer
<https://www.google.com/publicdata/directory>
 - [https://www.gapminder.org/tools/#\\$chart-type=bubbles](https://www.gapminder.org/tools/#$chart-type=bubbles)
 - <https://ourworldindata.org/explorers/global-food>
 - <https://ourworldindata.org/explorers/coronavirus-data-explorer>
 - More analysis intensive
 - <https://www.tradingview.com/chart/>
 - <https://stockcharts.com/h-sc/ui?s=XOM>
 - Map based:
 - <https://d5t6zpljmdkvz.cloudfront.net/>
 - <http://luminocity3d.org/WorldCity>

<https://www.tradingview.com/chart/>



Story



- Stories are predefined and *scripted* interaction and interpretation of data visualizations
 - Or, the data and visualizations drive the story
 - By allowing users to interact with data presented in a clearly-visual manner, a data-intensive ‘story’ becomes visible.
- Some examples:
 - <https://storymaps.arcgis.com/stories/f74a8fbad837435b8e901cc9c04aa345>
 - <https://public.tableau.com/profile/natcen.social.research#!/vizhome/WhatwillBrexitmeantotheUK/Home>
 - <https://projects.fivethirtyeight.com/2020-swing-states/>
 - <https://ourworldindata.org/energy-offshoring>

We do not cover story in depth in IT 7113, but it can be a very good research topic for class project.

Interactivity

IT 7113 module 10 covers more in depth on this topic, explaining common types and features of interactivity in data visualization:
<https://www.edocr.com/v/l0pp3ral/jgzheng/visual-interactivity>

- Interactivity is the functionality provided by the (visualization) system/application to let users interact with the visualization or the system/application through a user interface.
 - So, the visualization itself becomes dynamic based on user actions, providing different views of data.
 - Interactivity is an important aspect of data exploration and analysis, as both are interactive processes.
 - Visual interactivity focuses on the interactions in using various forms of data visualizations (charts, maps, dashboards, etc.)
 - Interactivity is also essential in visual analytics where discoveries are driven by intensive interactions.
- Why interactive?
 - Enable multiple perspectives
 - Static visuals can offer only pre-composed “views” of data, so multiple static views are needed to present a variety of perspectives on the same information. A fixed image is ideal when alternate views are neither needed nor desired, and required when publishing to a static medium, such as print. - Quotes from chapter 1 of the book “Interactive Data Visualization for the Web” by Scott Murray.
 - Reduce complexity
 - The number of views can grow significantly in many cases because of the multi-dimensionality of the data. Presenting all of them is impossible. Even presenting multiple of them maybe cluttered and crowded.
 - Interactivity enables a more prioritized and focused view in a limited space.
 - Ease cognitive load
 - The number of items and data presented at one time may be overwhelming for a user; interaction features can help user focus.
 - Enables customization and exploration
 - Dynamic, interactive visualizations can empower people to explore the data for themselves.
 - Encourage engagement with the data
 - With animated transitions and well-crafted interfaces, some visualizations can make exploring data feel more like playing a game or telling a story. Interactive visualization can be a great medium for engaging an audience who might not otherwise care about the topic or data at hand.
 - Make visualizations smart or tell a story: <http://www.dataversity.net/fact-fiction-smart-data-visualization-tells-tale/>

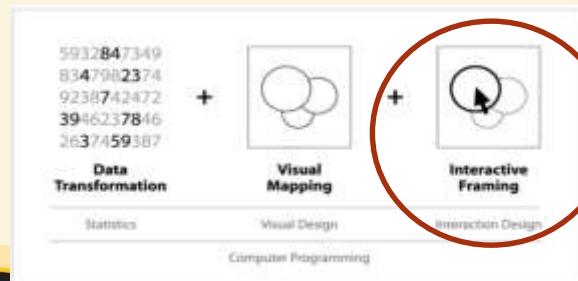


Image from

<https://www.slideshare.net/tillnagel/nagel-unfolding-thecityworkshops>



Developing Data Visualizations

The design/development of data visualizations involves

- Processes
- Principles
- Best practices
- Tools

IT's role in business data visualization



Data Visualization Design Process

- A design process involves a defined set of design considerations and tasks
- Why do we need a process?
 - A repeatable and defined design process embodies maturity in design capability and experience
 - greatly facilitate the design efficiency and effectiveness
 - reduce the randomness and puts tasks in order
 - serve as a guide and a checklist to plan and manage the whole project
- The following is a basic simple process for the most often scenario: we need to present all data in just one chart of a common type. This class will focus on this kind of scenario.
 1. Requirement analysis: set project goals and contexts
 2. Choose one basic chart type (general or industry specific) based on a number of factors (mainly purposes and data features)
 3. Representation design: determining visual data coding, involving visual mapping and visual properties (SCOPEs)
 4. Presentation design: apply perceptual and attention shaping best practices (for example, pre-attentive processing and Gestalt principles) to make charts more effective and efficient

We will cover more about design process in IT 7113 module 4.

There are other ways to define the designing process or factors. For example:

- A 3-Step Approach To Data Visualization <https://digitalimpact.io/getting-started-a-3-step-approach-to-data-visualization/>
- A 5-step guide to data visualization <https://www.elsevier.com/connect/a-5-step-guide-to-data-visualization>
- The Data Visualization Design Process: A Step-by-Step Guide for Beginners <https://depictdatastudio.com/data-visualization-design-process-step-by-step-guide-for-beginners/>
- Andy Kirk's 4 stages: <https://www.youtube.com/watch?v=GVkXbQOzKNs&t=754s> or Andy Kirk's book "Data at Work" Chapter 2
- Design process for information visualization <https://www.interaction-design.org/literature/article/how-to-design-an-information-visualization>



Basic Design Principles and Guidelines

- The **CASE** principles

<u>C</u>larity	The chart delivers the message and makes the point clearly.
<u>A</u>ccuracy	Avoid data visual distortion and disinformation.
<u>S</u>implicity	Perceptually easy to locate and identify key data and other information.
<u>E</u>legance	Visual quality to attract audience and sustain that sentiment and interest – Andy Kirk.

CASE

- Also follow conventions, design patterns, and standards.
 - Standards and conventions are great to promote consistency and perception alignment.
 - For example, IBCS <https://www.ibcs.com/standards/>

We will cover principles in IT 7113 module 5.



Dashboard Design Principles

In addition to general data visualization design principles, dashboards share many principles and practices of general UI design (on usability):

1. Meeting the goals and objectives (focus on data and insights)
 - Focus on data, instead of visual – “Dashboards are not an appropriate venue for artistic impression.” - Stephen Few
 - All visuals and data needs to be relevant and directly support the objectives of the dashboard
2. Clarity and effectiveness
 - Effective visuals that clearly reveal the insight and deliver the message
 - Artistic expression in Data Visualization: use more memorable, less abstract, real world iconic representations
3. Simplicity and efficiency
 - Simple and clear: use the simple design to meet the objectives and deliver messages
 - Easy to understand, explore, and interact
4. Consistency
 - Design as a whole: dashboard level design beyond single chart
 - Be consistent for the complete application, and consistent with the other organization applications and cultures. multiple charts, pages
 - Follow conventions, norms, and standards

We will cover dashboard design principles in IT 7113 module 9
<https://www.edocr.com/v/9oqqrzoo/jgzheng/designing-dashboards>



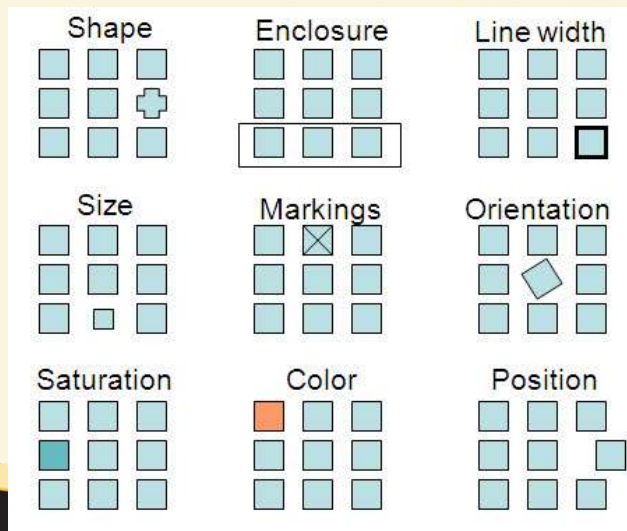
Standards, Conventions, and Guidelines

- Standards and conventions are great to promote consistency and perception alignment.
- Two theories and design practices are widely referenced (next two slides)
 - Pre-attentive processing
 - Gestalt principles
- Are there any standards in an industry or a particular sector?
 - It is difficult and to make and require standards in design. They are more like guidelines.
 - For example, IBCS <https://www.ibcs.com/standards/>
- If not, it's always good to establish organizational guidelines, and follows conventions and best practices
 - For example, European Environment Agency (eea.europa.eu) has a set of usability guidelines for improving visualisations
<https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts>
 - <https://xdgov.github.io/data-design-standards/>

Pre-Attentive Processing



- Any visual processing of that item prior to the act of selection can be called “preattentive”.
 - Wolfe et. al. 2010 <https://www.researchgate.net/publication/245623481>
- Preattentive processing can help to rapidly draw the focus of attention to a target with a unique visual feature (i.e., little or no searching is required in the preattentive case).
 - Healey, 2005, <https://www.csc2.ncsu.edu/faculty/healey/PP/index.html>
- The technique is commonly used in many fields involving visual designs, including:
 - UX/UI and interaction design
 - Data/information visualizations (charts, maps, dashboards)
 - Web design, product design, shelf display, painting, etc.
- Basic technique: using contrast to differentiate the part that needs to draw attention
 - <http://kenhirakawa.com/significance-of-contrast/>
 - <https://www.coursera.org/lecture/dataviz-design/strategic-use-of-contrast-sDV6C>



Some examples by using various distinctive visual properties

Pre-Attentive Processing are covered with more details in module 2 and emphasized throughout the course.

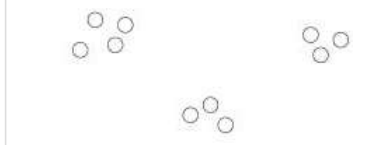
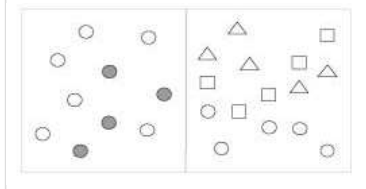
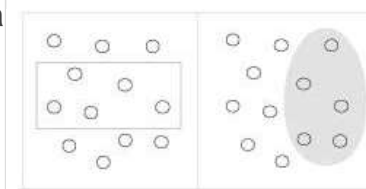

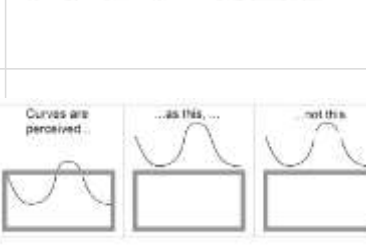
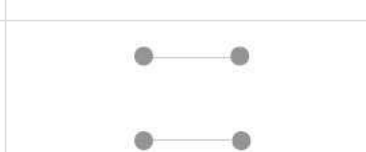
Extended reading:
“What shall we do with the preattentive processing stage”
<https://www.researchgate.net/publication/245623481>

More examples in “Perception in Visualization” by
Christopher Healey
<https://www.csc2.ncsu.edu/faculty/healey/PP/index.html>

Gestalt Principles of Perception

- Gestalt principles describes a set of ways how human perceives images and how visual information are identified and related from images.
 - These principles have profound implication on many fields involving visual designs, including:
 - UX/UI and interaction design
 - Data/information visualizations (charts, maps, dashboards)
 - Web design
 - Graphical design
 - Painting/Photography
 - Shelf display, catalog, form, etc.
 - See how it is applied in general visual designs
- <https://www.webfx.com/blog/web-design/gestalt-principles-applied-in-design/>

Gestalt Principles are covered with more details in module 2 and emphasized throughout the course.

Proximity	Objects that are close together are perceived as a group.	
Similarity	Objects that share similar attributes (e.g., color or shape) are perceived as a group.	
Enclosure	Objects that appear to have a boundary around them (e.g., formed by a line or area of common color) are perceived as a group.	
Closure	Open structures are perceived as closed, complete, and regular whenever there is a way that they can be reasonably interpreted as such.	
Continuity	Objects that are aligned together or appear to be a continuation of one another are perceived as a group.	
Connection	Objects that are connected (e.g., by a line) are perceived as a group.	



Data Visualization and IT

- Modern advanced and interactive visualizations are driven by the need for technology and tool support (design, development/programming, automation, delivery, and administration).
- Michael Friendly offers some key points how IT drives the modern data visualization (<https://www.interaction-design.org/literature/article/information-visualization-a-brief-20th-and-21st-century-history>)
 - The field of information (data) visualization has broadened to encompass many new forms of data, data structure and problem solving.
 - Highly *interactive* computing systems have been developed and are in common use. This is compared to early command-driven systems which used compiled, batch processing.
 - The information visualization field has begun to implement cognitive and perceptual aspects of displaying data in addition to delivering simple static visualizations which were aesthetically pleasing.
 - Data with large volumes of dimensions can be better explored and analyzed.

Data Visualization Tools



- Modern visualizations are largely dependent or enabled by visualization tools. Visualization products have been evolving fast, and there is increasing overlap. But they generally fall into three major categories.
- Standalone tools
 - They are specifically designed to produce stunning visualizations, and they can work with multiple platforms and data sources.
 - Some of them are growing to more full stack analytics tools.
 - They can be desktop based and/or cloud based
 - Examples: Tableau, Power BI, QlikView, Dundas, Spotfire, SAP Lumira, etc.
 - Cloud: Google Data Studio, <http://idashboards.com>, <http://www.klipfolio.com>
- Embedded tools
 - Broader analytics, business intelligence, and reporting platforms (and even advanced spreadsheet programs) that often incorporate visualization capabilities. These products can address more complex data platform needs and often provide wide-ranging capabilities but may require more training in order to exploit their full potential. In some cases, IT may need to be looped in to assist in integrating these tools with underlying data and related applications.
 - Examples like SSRS, IBM, Oracle, MicroStrategy, SAP Crystal, and others.
 - Microsoft Excel, Google Docs Spreadsheet <http://www.benlcollins.com/spreadsheets/dynamic-dashboard-in-google-spreadsheets/>
- Developer-oriented visualization libraries and APIs
 - These tools are offered as programming libraries or services for general applications (web, mobile, etc.).
 - These tools can be useful when the visualization requires complete customization, substantial interactivity, or for developing a framework that allows you to reuse code.
 - Examples
 - Programming library: D3, dotNetCharting, Telerik, Nevron, amCharts, etc.
 - Web API: Google Charts (<https://developers.google.com/chart/>)
 - Programming language capabilities: R, Python
- More tools
 - <http://selection.datavisualization.ch>
 - <https://www.g2crowd.com/categories/data-visualization>
 - <http://www.creativeblog.com/design-tools/data-visualization-712402>
 - <http://www.computerworld.com/article/2506820/business-intelligence/chart-and-image-gallery-30-free-tools-for-data-visualization-and-analysis.html>
 - <https://bigdata-madesimple.com/review-of-20-best-big-data-visualization-tools/>

We will cover these topics with more details in IT 7113 module 11.

The Industry

- Why are there so many visualization tools?
 - <https://www.linkedin.com/pulse/why-so-many-visualization-bi-tools-adam-roderick/>
- Industry consolidation (as with the BI industry consolidation)
 - Atlassian acquired Chartio (2022)
 - Airtable acquired Bayes (2021)
 - Google bought Looker (2020)
 - Salesforce bought Tableau (2019)

The global data visualization market is expected to register a CAGR of over 9% during the forecast period (2020 - 2025).
<https://www.mordorintelligence.com/industry-reports/data-visualization-applications-market-future-of-decision-making-industry>



<https://www.g2.com/categories/data-visualization>



Learning Data Visualization

Skills, jobs, career, and learning resources

Skills in Data Visualization Dev.



- Data visualization draws knowledge and experience from multiple fields including computing, business, and design.
- Most important
 - Visualization design: charts, diagrams, maps, etc.
 - UI and interaction design
 - Knowledge of the dev tool
 - Business domain knowledge
- Highly useful
 - Communication, story telling
 - Programming/scripting
 - Data literacy, statistics
- Very helpful
 - Data models
 - Data preparation
 - Analytics methods
 - Artistic design
 - Information behavior



Data Visualization: Sample Real Jobs

Data Visualization Analyst (originally posted on LinkedIn):

- Responsible for the management of database analysis projects in support of business initiatives.
- Data visualization (DV) expertise to design, develop and implement clear, interactive and succinct visualizations by processing and analyzing large quantities of (un)structured data.
- Candidate should have ability to turn raw data into compelling, lively stories, enriched with powerful, clear visualizations.
- These visualizations would also provide end-users an ability to discover relationships within related data in fresh and innovative ways.
- Updates visualization items as defined by department, in accordance with system protocol and requests from relevant departments.
- Serves as a liaison between business stakeholders and technology resources to optimize processes and designed visualization functionality.
- Assists with user acceptance testing for new information dashboards and/or analytical systems.

CONA (Coke One North America) is searching for a BI Reporting Analysis Intern. In this position, you will use your technical skills to participate in the development of strategic reporting dashboards and visualization of data solutions to support various business operations. We are looking for candidates with knowledge of BI, reporting and dashboard tools knowledge (preferably Tableau), and understanding of SQL.

In this role, you will work to create comprehensive metrics on some of Coca-Cola Bottling highest-profile initiatives such as the visualization of operational trends and ongoing operations efficiency reporting through the merging of different data sources (SAP, ServiceNow, HPE ALM etc.). You will participate in data analytics projects and the delivery of advanced visual dashboards, contribute to team discussions around reporting strategies, and provide expertise in the development, usability and presentation of data.

Responsibilities:

- Work closely with stakeholders to identify KPIs, metrics, and business insights needs
- Serve as subject matter expert for internal clients performing analysis and creating visual reporting
- Design data models including customized business dimensions, facts and measures to support scorecard, and self-service reporting needs
- Build a data pipeline to enable data integration from different data sources. Drive collection of new data, and refinement of existing data sources
- Utilize BI Software (like Tableau) and SQL to create visual dashboards and reporting solutions that provide meaningful insight to operational results, trends and are actionable
- Work directly with system owners and IT development teams to implement recommendations and work through data and data sources to support reporting needs

The Data Visualization Analyst will be responsible for understanding the strategic needs of the business and translating high-level objectives into the development of visual data analysis and dashboards to support the category management and product strategy teams. The candidate will need to understand how to create and manipulate large data sets and use various visualization tools to meet the needs of needs of their customers. To ensure adoption by the business, this position **will be required** to ensure the quality of each dashboard release, data refresh and adhere to a regular refresh and dashboard publishing schedule.

<https://www.linkedin.com/jobs2/view/12915000>



Data Visualization Trends

- Public communication with intensive visualizations - used creatively in many public media like
 - Journalism (US News Election coverage)
 - Government report (<https://www.usaspending.gov/#/explorer>, <https://itdashboard.gov>)
- More common and easy interactive maps
- Visualization intensive stories (narrative with creative and interactive data visualizations)
 - <https://www.vox.com/2018/1/8/16822374/school-segregation-gerrymander-map>
 - <https://www.jato.com/station-wagons-are-disappearing-but-wait-theres-hope/>
- Dashboards and visualizations in more types of display media and interfaces
 - Mobile friendly visualizations
 - Interactive super big displays
 - VR/AR environments

Interesting read from Elijah Meeks

<https://medium.com/nightingale/2019-was-the-year-data-visualization-hit-the-mainstream-d97685856ec>



Learning Data Visualization

- IT 7113 Data Visualization (MSIT)
 - <http://zheng.kennesaw.edu/teaching/it7113>
 - An elective course in the KSU MSIT and “certificate on data analytics and technology”
<http://ccse.kennesaw.edu/it/programs/cert-dm.php>
 - Open educational resources at <http://idi.kennesaw.edu/it7113/>
- Lecture notes serials
 - <https://www.edocr.com/user/jgzheng/collection/datavisualizationlecturenotes>
- DATA 3230 Data Visualization
- Other good courses
 - UC Davis on Coursera
<https://www.coursera.org/specializations/data-visualization>
 - University of Washington CSE 442/512:
<https://courses.cs.washington.edu/courses/cse442/>
<https://courses.cs.washington.edu/courses/cse512/>



Key Readings

- The Role of Data Visualization in Business Intelligence:
<https://www.highcharts.com/blog/post/role-data-visualization-business-intelligence/> - this is a very practical article for some quick reading
- Defining Data Visualisation <https://chrisluv.medium.com/defining-data-visualisation-daf71c22ec03>
- Data Visualization for Human Perception (by Stephen Few):
<https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception> - this is a very conceptual and intensive reading
- Zheng (2017) Book Chapter Data Visualization in Business Intelligence (PDF downloadable on the site):
https://www.researchgate.net/publication/321804138_Data_Visualization_for_Business_Intelligence (note: this chapter generally corresponds to the lecture notes but the lecture notes is more recent and updated. Use it together with the lecture notes).



Additional Good Resources

- Data Visualisation - A Game of Decisions with Andy Kirk
<https://www.youtube.com/watch?v=GVkXbQOzKNs>
- What Is Data Visualization? Definition, Examples, And Learning Resources <https://www.tableau.com/learn/articles/data-visualization>
- What is data visualization:
<https://www.perceptualedge.com/blog/?p=2636>
- Tegarden (1999) CAIS Business Information Visualization (a bit aged but still classic):
<http://aisel.aisnet.org/cgi/viewcontent.cgi?article=2483&context=cais>
- Some general Wikipedia resources
 - https://en.wikipedia.org/wiki/Data_visualization
 - https://en.wikipedia.org/wiki/Information_visualization
 - <https://en.wikipedia.org/wiki/Infographic>
 - https://en.wikipedia.org/wiki/Scientific_visualization
 - [https://en.wikipedia.org/wiki/Visualization_\(graphics\)](https://en.wikipedia.org/wiki/Visualization_(graphics))
 - https://en.wikipedia.org/wiki/Visual_analytics

General Resources



- Influencers
 - Stephen Few <http://www.perceptualedge.com/>
 - Edward Tufte <https://www.edwardtufte.com/tufte/>
 - Ben Shneiderman https://en.wikipedia.org/wiki/Ben_Shneiderman
 - Andy Kirk <https://www.visualisingdata.com/about/>
 - Jorge Camoes <https://excelcharts.com/>
 - Jeffrey Heer <https://homes.cs.washington.edu/~jheer/>
 - Mike Bostock <https://bost.ocks.org/mike/>
 - Scott Murray <http://alignedleft.com>
- Tableau Zen Masters
 - Mark Jackson <http://ugamarkj.blogspot.com>
 - Jeffrey Shaffer <https://www.dataplusscience.com/insights.html>
 - Andy Kriebel <http://www.vizwiz.com>
 - Ryan Sleeper <https://www.ryansleeper.com>
 - Adam McCann <http://duelingdata.blogspot.com>
 - The Flerlage twins <https://www.kenflerlage.com>
 - Luke Stanke <https://public.tableau.com/app/profile/stanke>
 - Jacob Olsufka <https://public.tableau.com/app/profile/datavizard>
 - Matt Chambers <https://www.sirvizalot.com>
- Books
 - Andy Kirk, Data Visualisation: A Handbook for Data Driven Design
 - Jorge Camoes, Data at Work <https://www.dataatworkbook.com>
 - Stephen Few, Show Me the Numbers, <https://www.amazon.com/dp/0970601972/>
 - “Information Dashboard Design” 2nd, by Stephen Few, 2013, <http://www.amazon.com/gp/product/1938377001/>
 - <https://clauswilke.com/dataviz/>
 - Designing Data Visualizations, by Julie Steele, Noah Iliinsky, O'Reilly, 2011
- Communities and organizations
 - <https://www.datavisualizationsociety.org>
 - <http://www.visualizing.org/>
 - <http://www.interaction-design.org/>
 - <http://www.storytellingwithdata.com>
 - <http://flowingdata.com>
 - <https://ourworldindata.org>
- News, media, and magazines
 - <https://nightingaledvs.com>
 - <https://www.interaction-design.org/literature/topics/information-visualization>
 - <https://www.techtarget.com/searchbusinessanalytics/resources/Data-visualization>
 - <http://understandinggraphics.com/>
 - <https://visage.co/blog/>
 - <https://informationisbeautiful.net/>
 - TED videos: https://www.ted.com/playlists/56/making_sense_of_too_much_data
- Company resources
 - <https://www.tableau.com/learn/articles/data-visualization>
 - https://www.sas.com/en_us/insights/big-data/data-visualization.html
 - <https://material.io/design/communication/data-visualization.html>
 - <http://blog.visual.ly/>
 - <https://www.darkhorseanalytics.com>
 - <https://playfairdata.com>