Introduction to Information Visualization

M. Cristina

2023

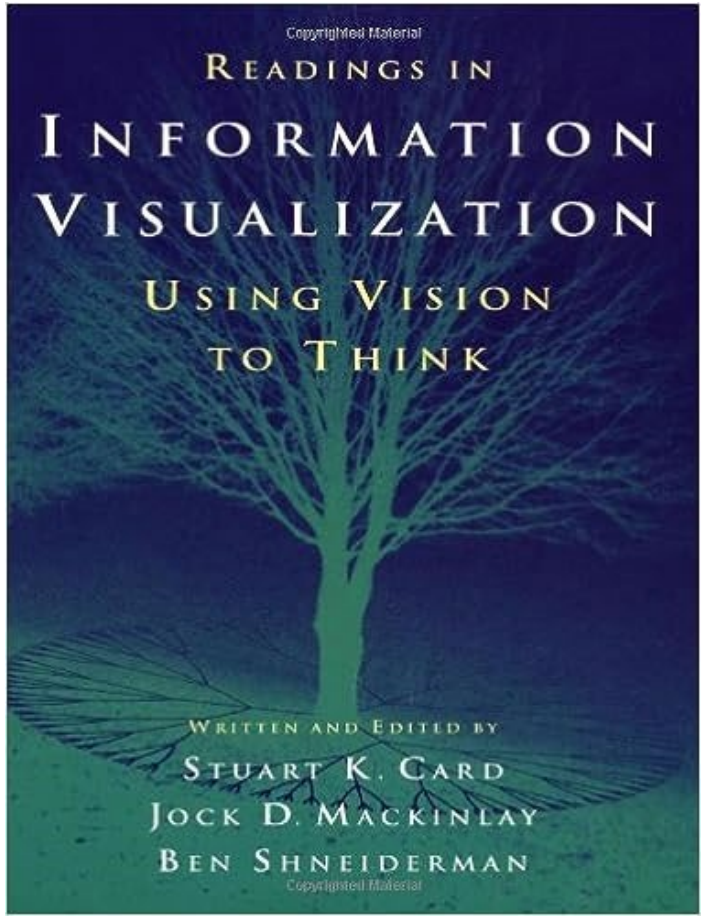
Goals of this class

- Answer two main questions
  - What is Information Visualization?
  - Why use it?

Part I: What is Information Visualization?

A definition

“the use of computer-supported, interactive, visual representations of abstract data to amplify cognition”



Example

- Visualization in action...

- Warranty Analysis | SAS

- http://playground.tensorflow.org/

- **A.I. Experiments: Visualizing High-Dimensional Space**

https://www.youtube.com/watch?v=wvsE8jm1GzE

Key concepts

- In definition and demos

Key concepts

- Computer-based
- Visual representation
- Abstract data
- Interactive
- Amplify cognition

- Can we describe better these key concepts?

“the use of computer-supported, interactive, visual representations of abstract data to amplify cognition”

The InfoVis pipeline

- Which are the elements in a visualization?

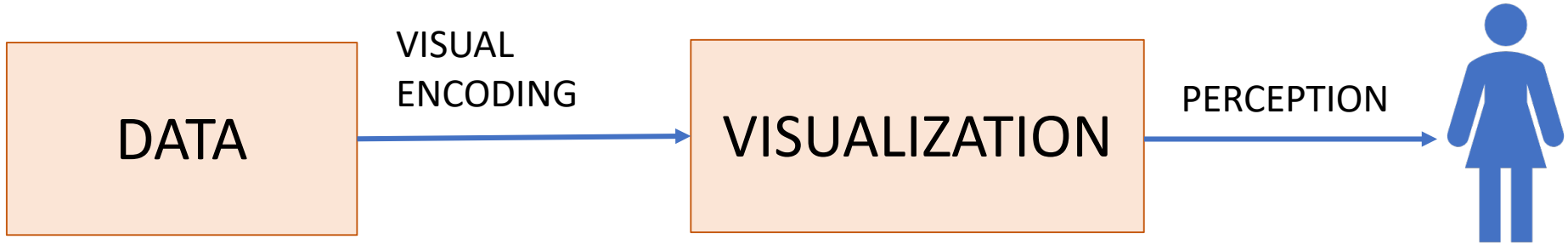
The InfoVis pipeline

- Which are the elements in a visualization?

The data, a visual encoding step, a human who `interprets´ the visual encoding

The InfoVis pipeline

- Which are the elements in creating a visualization?



The InfoVis pipeline

- Which are the elements in a visualization?

The data, a visual encoding step, a human who `interprets´ the visual encoding

What else?

The InfoVis pipeline

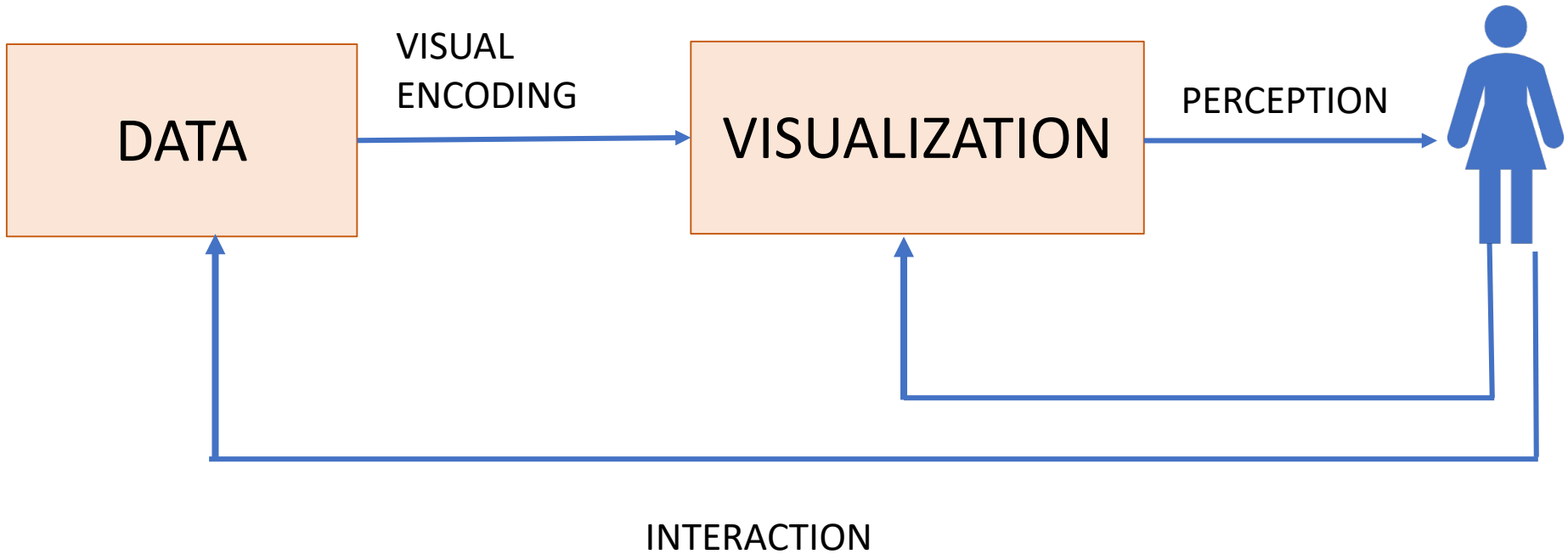
- Which are the elements in a visualization?

The data, a visual encoding step, a human who `interprets´ the visual encoding

The ability to interact with the visualization and with the data!

The InfoVis pipeline

- Which are the elements in creating a visualization?



The key concept

- Computer-based graphical representations and visualizing abstract data
  - Computer-based
  - Visual representations
  - Abstract data – what does it mean?

The key concept

- Computer-based graphical representations and visualizing abstract data
  - Computer-based
  - Visual representations

  - Abstract data – representing visually data that does not necessarily have obvious visual representations

Abstract data vs “physical” data

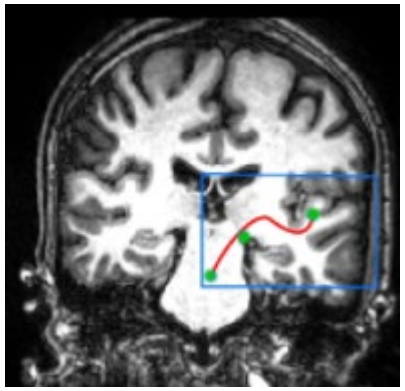
Ex. “physical” data

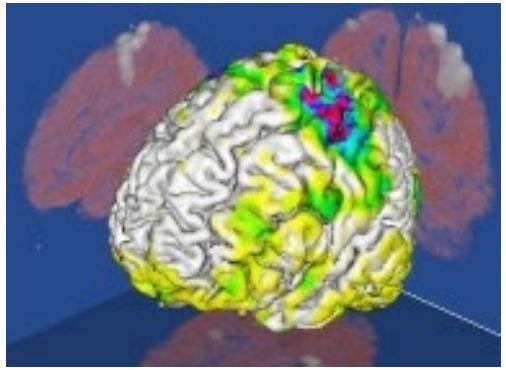
Visualizing MRI & CT Scans in Mixed Reality / VR / AR, Part 2: 3D Volume Rendering – andreasjakl.com

3D Surface and Volume Display 3D confocal volume control and display – HORIBA

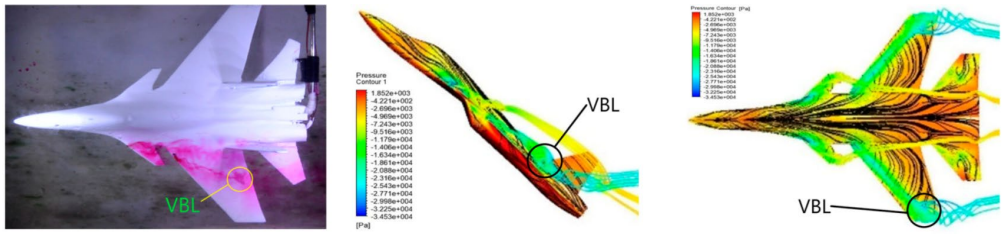
Abstract data vs “physical” data

Scientific Visualization





http://www.cs.rug.nl/svcg/SciVis/DTIVis



Sutrisno; Rohmat, T.A.; Wibowo, S.B.; Iswahyudi, S. Vortex Dynamics Study of the Canard Deflection Angles’ Influence on the Sukhoi Su-30-Like Model to Improve Stall Delays at High AoA. *Aerospace* **2019**, *6*, 12.

- Medicine
- Simulation
- Engineering

- Weather forecast
- Biomolecular research
- Cosmology

Abstract data vs “physical” data

Ex. “abstract” data

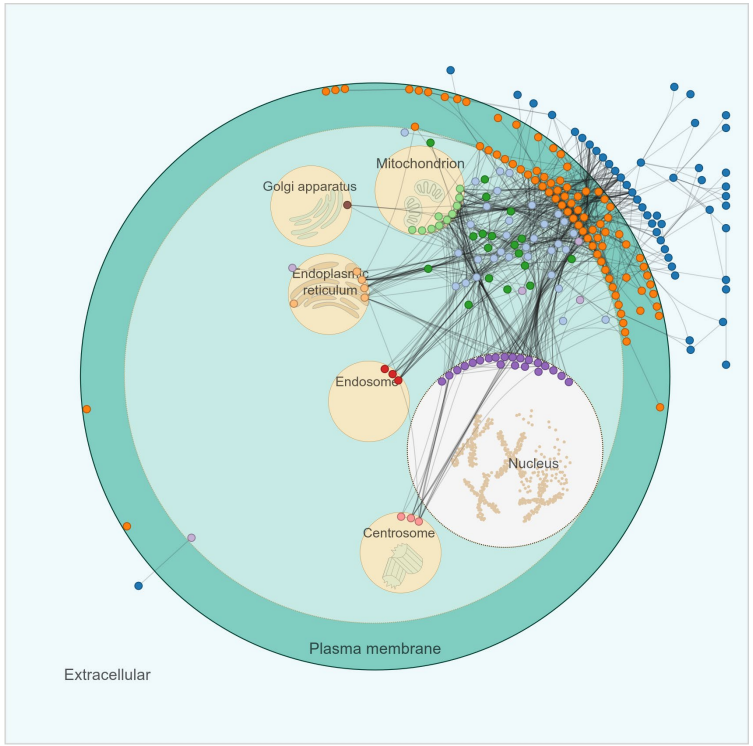
- The 10 Best Data Visualization Examples | Tableau

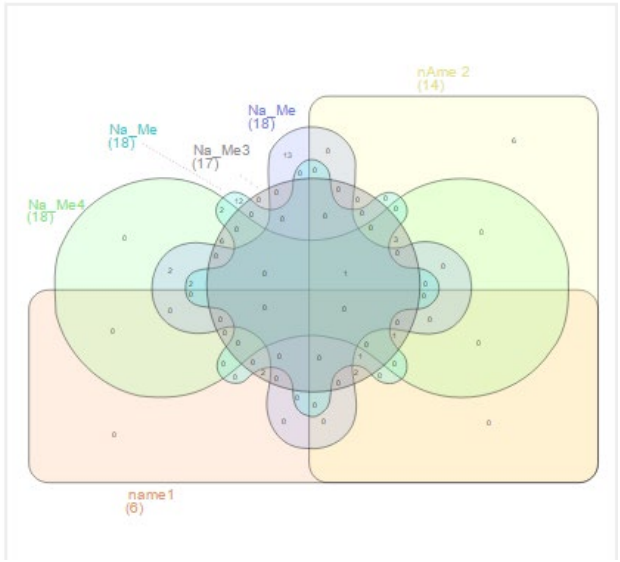
Example: Visual Analysis of Social Networks (Rumour)

Anh Dang, Abidalrahman Moh'd, Evangelos Milios, and Rosane Minghim. 2016. What is in a Rumour: Combined Visual Analysis of Rumour Flow and User Activity. (CGI '16). ACM, New York, NY, USA, 17–20. DOI:https://doi.org/10.1145/2949035.2949040

Example: Proteomics and Cancer

- Masters/PhD project: Visual inspection of protein candidates





Heberle, H. ; Meirelles, G. V. ; Silva, F. R. ; Telles, G. P. ; MINGHIM, R. . InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics**, v. 16, n. (169), 2015.

Highly-cited WoS

HEBERLE, H.; CARAZZOLLE, M. F.; TELLES, G. P.; MEIRELLES, G. V.; MINGHIM, R. CellNetVis : a web tool for visualization of biological networks using force-directed layout constrained by cellular components.
BMC Bioinformatics, v. 18, n. 10, p. 395, 2017

Best Paper BioVis 2017.

Abstract data vs “physical” data

“physical” data

- Result from observation or simulation of physical phenomena
  - there is an `expected´ visual representation
  - `spatial´ attributes are inherent to the data – and drive the creation of a graphical representation
  - `realism´ is a desired property of the visualization

Abstract data vs “physical” data

“abstract” data

- Result from observation of phenomena that are not `physical´
  - a visual representation has to be devised
  - How do you visualize information such as time? duration? Project names? Type of `something´? Topics? No spatial attributes are given with the data!
  - No obvious or straightforward visual encoding
  - Visualization designer has the difficult task of selecting a visual encoding, or creating a novel visual representation

The key concept

- Computer-based graphical representations and visualizing abstract data
  - Computer-based
  - Visual representations
  - Abstract data
  - Interactive – what does it imply?

Interactive

- Users must be able to interact to change what is being visualized and also how it is being visualized

Key concepts

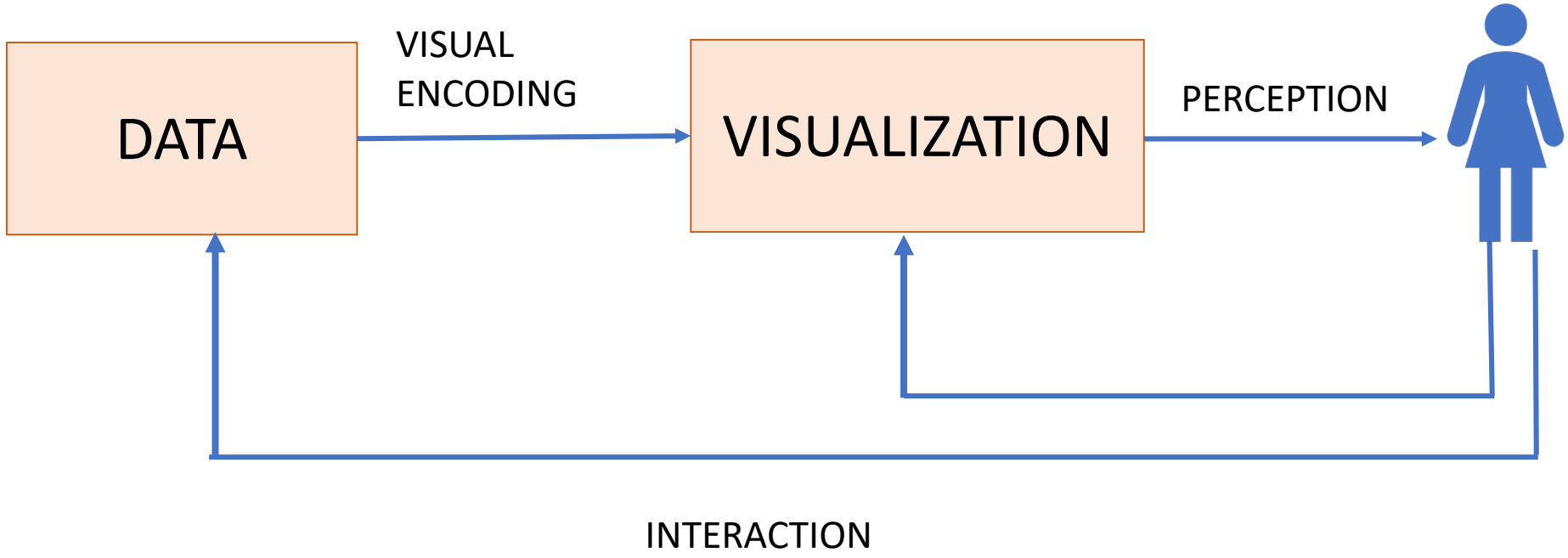
- Computer-based
- Visual representation
- Abstract data
- Interactive
- Amplify cognition

- Can we describe better these key concepts?

“the use of computer-supported, interactive, visual representations of abstract data to amplify cognition”

Interactive

- Users must be able to interact to change what is being visualized and also how it is visualized



The key concept

- Computer-based graphical representations and visualizing abstract data
  - Computer-based
  - Visual representations
  - Abstract data
  - Interactive
  - Amplify cognition – what does it mean?

Cognition

- Cognition refers to the mental processes and activities related to acquiring, processing, understanding, and using information.
- It involves various aspects of thinking, perceiving, remembering, problem-solving, and decision-making.
- Cognition is a complex and multifaceted concept that encompasses a wide range of mental activities, and can be defined in several different ways

Amplify cognition

- Producing tools that help us to think better when solving problems and making decisions
  - Solving problems that require data analysis with less effort, in a shorter time, and more accurately
  - Even, being able to perform analyses that could not be possible without the aid of a computer and a graphical representation

Amplify cognition

- Producing tools that help us to think better when solving problems and making decisions

- Related concept: cognitive artifacts

Amplify cognition

- Producing tools that help us to think better when solving problems and making decisions

- Related concept: cognitive artifacts
  - Tools human build to help them to reason and think
  - Physical or digital objects that help us with ou mental processes

  - Examples?

Exercise

- Try to multiply 34 x 72 mentally in your head
- Is is easy? It is much easier with pen and paper, right?
- Why?

Exercise

- With pen and paper, you follow a number of steps and write down intermediate results
- You off-load your mind by not having to store intermediate information while reasoning

- This is a very importante concept!

Distributed cognition

- To solve a problem, we use external cognitive artifacts: tools that are in the world, to record information that can be accessed when it's needed, without us having to retain it in our mind all the time.
- Our cognitive system is not exclusively made of our brain, our mind, and our sensors: it's also made of the artifacts that we use to store and manipulate information.
- Visualization is a way for us to store information out of our mind and make it accessible through our eyes and also through manipulation with interactive systems.

Another example

- The `game of 15´
- Let´s play!

Another example

- Tic-tac-toe

- Let´s play!

Another example

- Which of the two games is easier to play?

Another example

- They are actually the same problem!
- Problem isomorph: Herbert Simon (Nobel Prize)
- Why playing tic-tac-toe is easier?

Another example

- They are actually the same problem!
- Problem isomorph: Herbert Simon (Nobel Prize)
- Why playing tic-tac-toe is easier?
  - the tic-tac-toe version uses a representation that helps us store some of the information in the world rather than having all this information in our mind. Ultimately, this makes reasoning about the problem much easier.

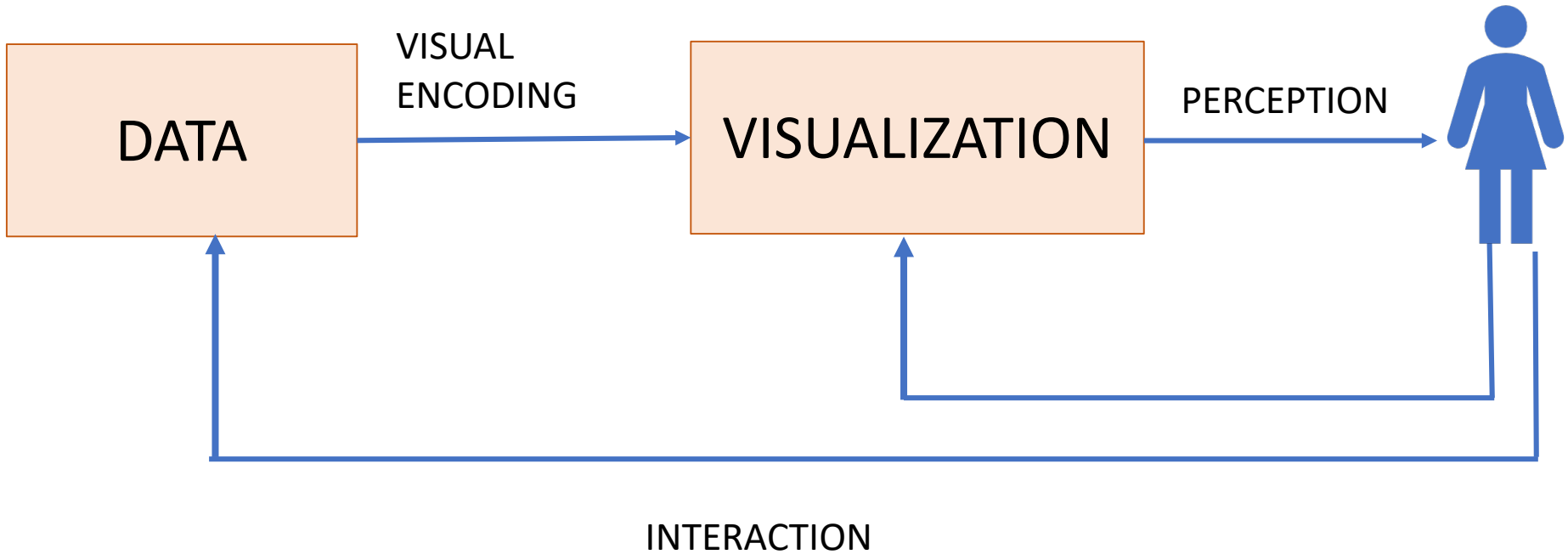
Part II: Why visualize data?

Goals of this class

- Answer two main questions
  - What is Information Visualization?
  - Why use it?

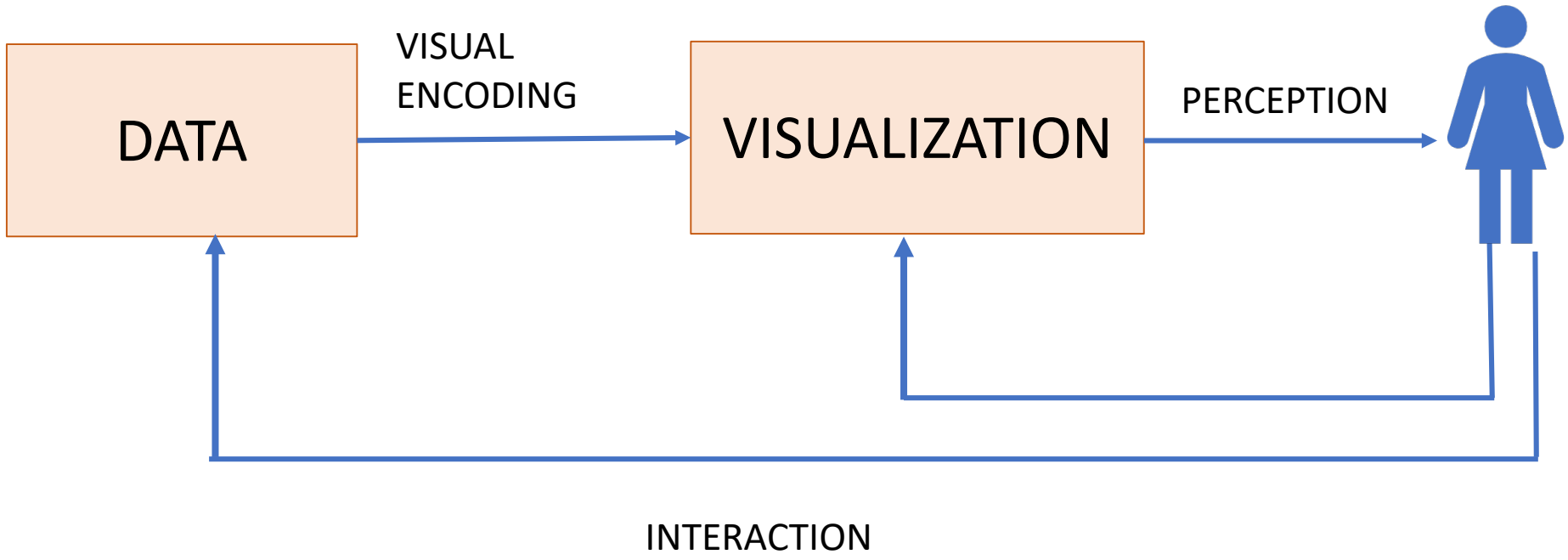
Information visualization pipeline

Users interact to change what is being visualized and also how it is visualized

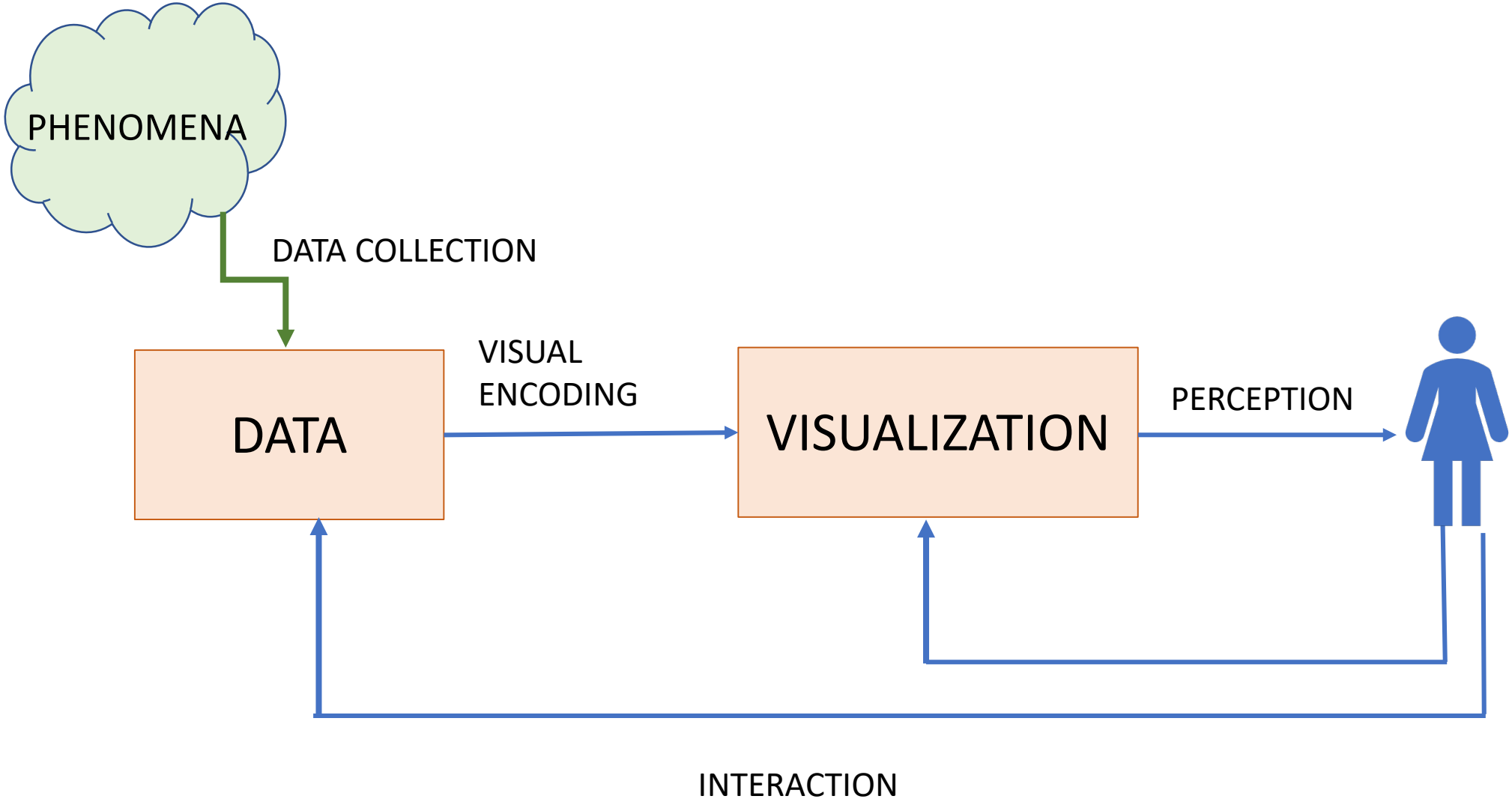


Information visualization pipeline

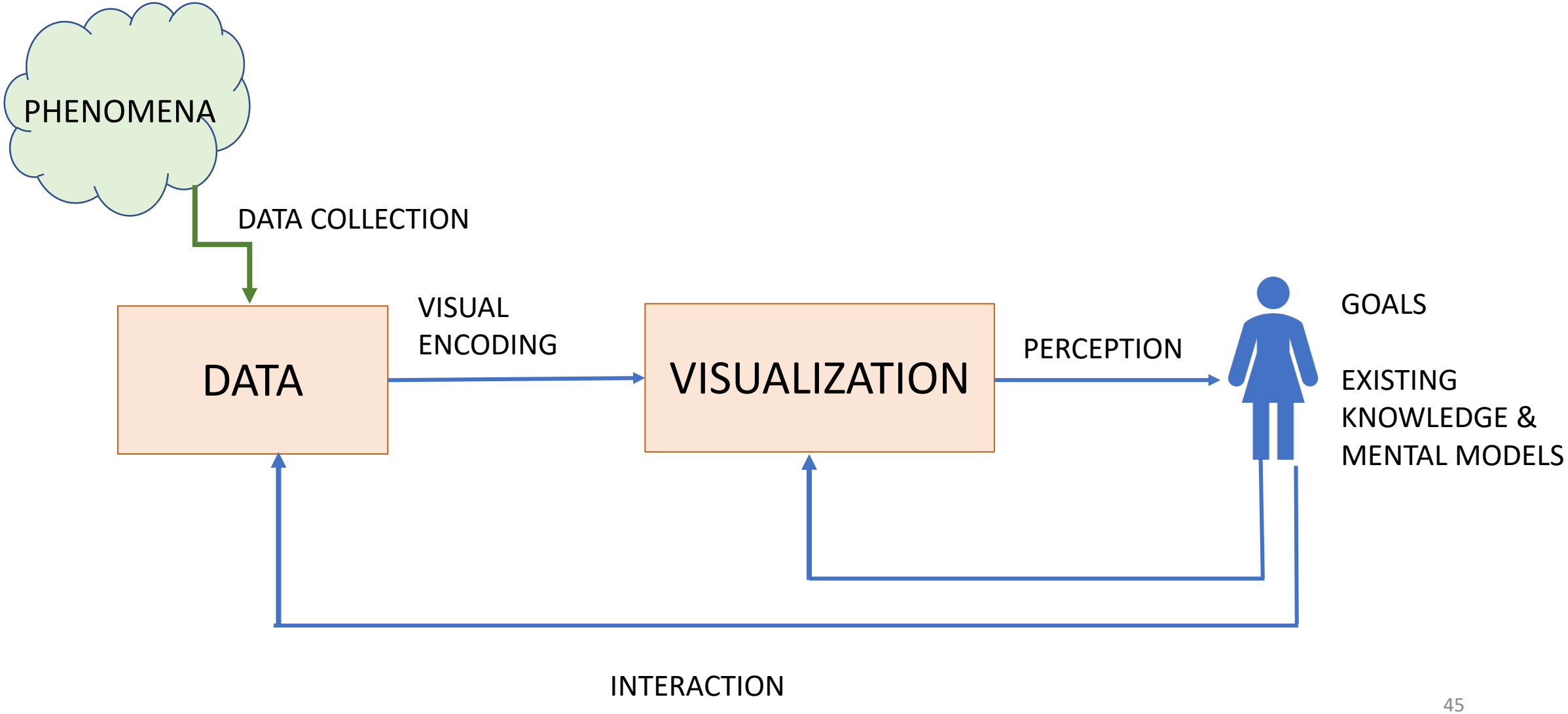
Why are we interested in data?



Information visualization pipeline



Information visualization pipeline



Why visualize data?

Visualization systems use data, visual representations, and interaction to help people get a better understanding of some phenomena.

Why visualize data?

Visualization systems use data, visual representations, and interaction to help people get a better understanding of some phenomena.

What are the purposes of visualization?

Why visualize data?

Explanatory

Confirmatory

Exploratory

Why visualize data?

Explanatory: the visualization conveys a narrative about the data. Someone has analyzed it already and wants a reader/observer to get the major findings.

- Purpose is to communicate something to somebody else

Confirmatory

Exploratory

Why visualize data?

Explanatory

Confirmatory: someone has a hypothesis about a phenomenon, wants to verify whether it holds in the data.

- Purpose is to analyze

Exploratory

Why visualize data?

Explanatory

Confirmatory

Exploratory: someone is trying to understand a phenomena, and the visualization provides means to explore the data. A user explores the data to raise hypotheses about the behavior of the phenomenon, that can be verified by other means.

- Purpose is to analyze

Explanatory

- How Americans Think About Climate Change, in Six Maps - The New York Times (nytimes.com)

Confirmatory/Exploratory

- Crossfilter (square.github.io)
- Aula Visualização de Dados - Dra. Rosane Minghim/Irlanda (2021-08-19 at 06:10 GMT-7) - Google Drive

Why use a graphical representation?

Why use a graphical representation?

- Humans are visual animals!

- Which ways can we communicate information?
  - Verbally (spoken, written), Numerically, Graphically

Why use a graphical representation?

- Which ways can we communicate information?
  - Verbally (spoken, written), Numerically, Graphically
  - Crossfilter (square.github.io)

Why use a graphical representation?

- Which ways can we communicate information?
  - Verbally (spoken, written), Numerically, Graphically

- Why is processing graphical information `easier´ in many scenarios?

Why use a graphical representation?

- Which ways can we communicate information?
  - Verbally (spoken, written), Numerically, Graphically
- Why is processing graphical information `easier´ in many scenarios?
- Sequential vs parallel processing
  - When communicating information verbally, we are forced to process information sequentially
  - When observing a chart, were are consuming and processing the information in a parallel fashion all at once

Why use a graphical representation?

- Why not use numbers? e.g., summary statistics?

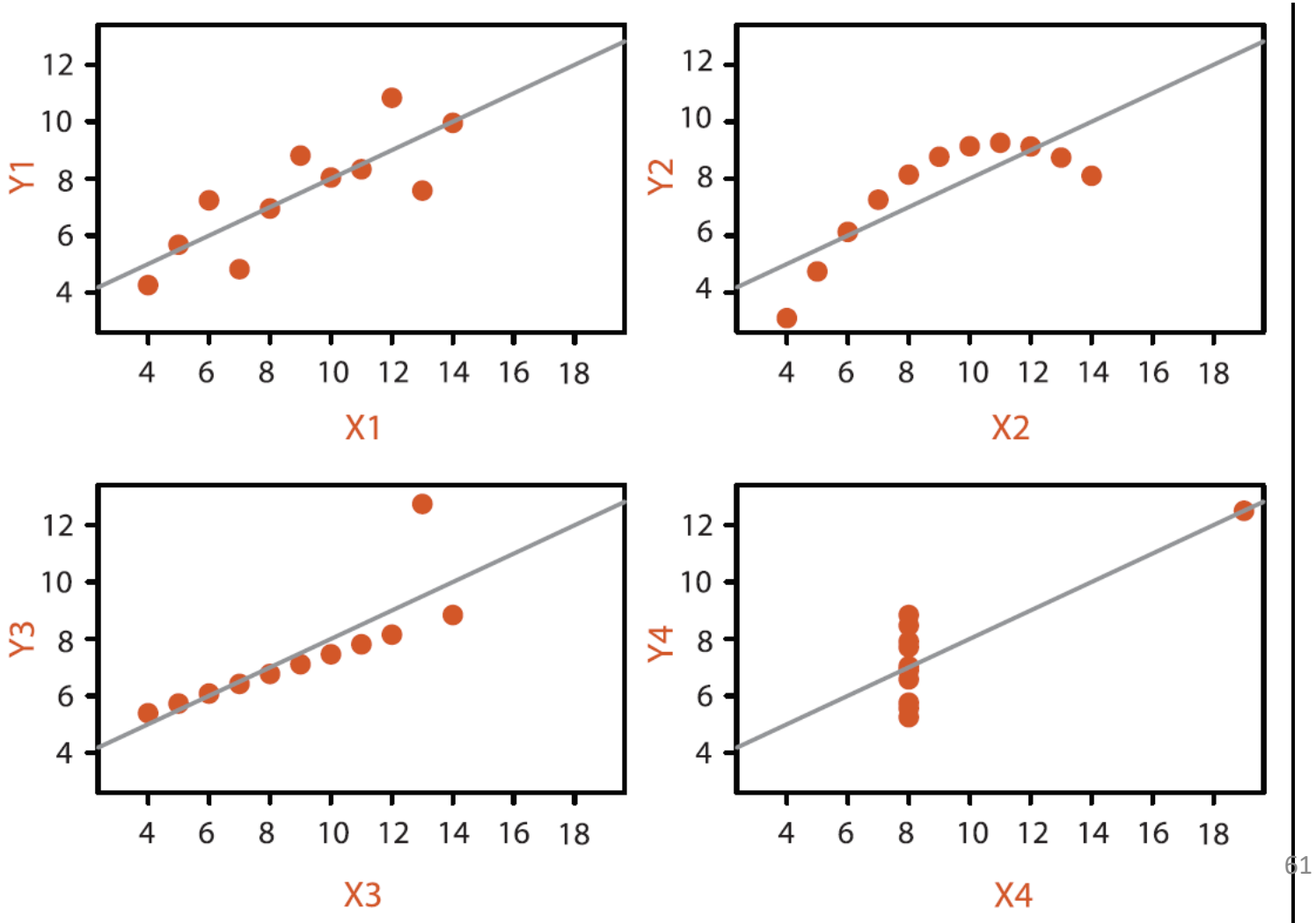
Summary statistics

Anscombe's Quartet: Raw Data

	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

Ascombe´s quartet

we visualize data,
we can perceive
information that is
hidden in the statistics



Why use computers to visualize data?

- You may not! (see this THE PROJECT — Dear Data (dear-data.com)

- But,
  - With a computer you can visualize a lot of data using automated algorithms, and
  - Computers also allow interacting with digital graphical representations, which is very useful

Why interaction is important?

What is the advantage of interacting with graphical representations?

Why interaction is important?

- When people analyze data visually, they often have several questions!
  - The ability to update a graphical representation as we interact helps us answer multiple questions that can hardly be answered by a single chart
- Typically we interact using the standard devices: screen, monitor, keyboard, mouse, maybe a touch device
- Many other devices and interaction modes are possible:
  - small screens, tablets, large touch screens, tabletops, large screen spaces, virtual reality devices

How do you assess the quality of a visualization?

- How do you know if you have a `good´ visualization?
- How do you know that a visualization is better than another one?
- Is it subjective? (people have different tastes...)

How do you assess the quality of a visualization?

- Some visual representations are actually much better than others at communicating some information
- If your goal is to communicate information to somebody else effectively, some visual representations are clearly better because they are easier to interpret, and easier and faster, and more accurate to read. (read this text!)

How do you assess the quality of a visualization?

- Also important: visual representations must be faithful to the data

- 5 examples of bad data visualization | Blog Jotform

- Misleading Data Visualization – how to avoid them? - Management Weekly

Misrepresentation of data!

How do you assess the quality of a visualization?

- When assessing the quality of a visual representation, you have to consider what is the intent, what problem I am trying to solve, what information am I trying to convey.
- And then figure out how to convey it and whether one visual representation is better than other at conveying this information.

How do you assess the quality of a visualization?

Designing effective visualizations requires two main expertises:

(1) A knowledge of the design space and how to explore it effectively.

(2) The ability to compare different solutions in an effective way, which requires knowledge of how human perception of graphical representation works.

Course

A vision of the design space, what visual representations are available for a certain type of data and problems, and

Some knowledge about human perception so you can start reasoning about whether and why a given visual representation may be more effective than another.

Atividade 1

- Analisar uma visualização explanatória

- Entrega: relatório de 2 a 3 páginas (PDF)
  - Identificação do(s) aluno(s)
  - Descrição da fonte da visualização, bem como a origem dos dados
  - Justificar a escolha dessa visualização (PG)
  - Descrever brevemente quais representações visuais foram usadas
  - Resumir os principais pontos comunicados (conclusões)
  - Opinar sobre a efetividade da visualização explanatória para o público alvo

Atividade 1

- Analisar uma visualização explanatória
- Grad: grupos de 2 ou 3, esta visualização: Plastic Waste Pollution data visualisation on Behance
- Posgrad: individual, visualização à escolha
- Entrega no eDisciplinas (até 05-09)

Atividade 1

- Sugestões de possíveis fontes (PG)
  - New York times, Washington Post, Financial Times
  - O Estado de São Paulo
  - Explorable Explanations
  - The 30 Best Data Visualizations of 2023 [Examples] (visme.co)