

Data Analytics and Decision Sciences

Course Introduction

Data and Results Visualization

Daniele Loiacono



[https://github.com/Daniele
Loiacono/DRViz2019](https://github.com/DanieleLoiacono/DRViz2019)

Instructor

□ Daniele Loiacono

- ▶ Contact: daniele.loiacono@polimi.it
- ▶ Office: DEIB, room 150
- ▶ Phone: +39 02 2399 3615

□ Research interests

- ▶ Data Mining and Machine Learning
- ▶ Games and Immersive Technologies

□ Materials

- ▶ [Github](#)
- ▶  DataCamp

Logistics

❑ When and where?

- ▶ 15/5/2019, 14:30 – 17:30, Aula Beta
- ▶ 20/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ 22/5/2019, **14:00 – 17:00**, Sala Conferenze DEIB
- ▶ 23/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ 27/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ ~~29/5/2019, 14:30 – 17:30~~
- ▶ 30/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ 5/6/2019, 09:30 – 12:30, TBD

❑ How lectures are organized?

- ▶ Start on time (no usual 15 minutes)
- ▶ Two chunks (1h ~ 1h10m)
- ▶ A 15m break in the middle

Logistics

❑ When and where?

- ▶ 15/5/2019, 14:30 – 17:30, Aula Beta
- ▶ 20/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ 22/5/2019, **14:00 – 17:00**, Sala Conferenze DEIB
- ▶ 23/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ 27/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ ~~29/5/2019, 14:30 – 17:30~~
- ▶ 30/5/2019, 09:30 – 12:30, Sala Seminari DEIB
- ▶ 5/6/2019, **13:30 – 16:30**, TBD

30/5, 13:30 – 16:30
5/6, 09:30 – 16:30

❑ How lectures are organized?

- ▶ Start on time (no usual 15 minutes)
- ▶ Two chunks (1h ~ 1h10m)
- ▶ A 15m break in the middle

How to pass the exam?

- Attending the 70% lectures (~18h)
 - ▶ Remember to sign at the beginning and after the break
- Small visualization project
 - ▶ Team: 1 or 2 students
 - ▶ Subject: either proposed by you or given by instructor
 - ▶ Evaluation: report and/or short presentation of results
- Grade: pass or fail

Course topics

- Introduction and data visualization framework
- Tools overview
 - ▶ Introduction to Pandas (hands-on)
 - ▶ Introduction to Matplotlib (hands-on)
 - ▶ Introduction to Seaborn (hands-on)
- Marks and Channels
 - ▶ Colormaps in Seaborn / Colorbrewer (hands-on)
- General principles
- Tables visualization(hands-on)
- Spatial data (hands-on)
- Networks and Graphs visualization (hands-on)
- Interactive visualization and dashboard design
- Dealing with complexity
 - ▶ Principles
 - ▶ Examples (hands-on)
- Validation

Part 1

Part 2

Part 3

Part 4

Course topics (tentative scheduling)

- Introduction and data visualization framework → 15/5
- Tools overview → 20/5
 - ▶ Introduction to Pandas (hands-on)
 - ▶ Introduction to Matplotlib (hands-on)
 - ▶ Introduction to Seaborn (hands-on)
- Marks and Channels → 20/5 and 22/5
 - ▶ Colormaps in Seaborn / Colorbrewer (hands-on)
- General principles → 22/5
- Tables visualization(hands-on) → 23/5
- Spatial data (hands-on) → 27/5
- Networks and Graphs visualization (hands-on) → 27/5 and 30/5
- Interactive visualization and dashboard design → 30/5 – Lecture 7
- Dealing with complexity → Lecture 7 and Lecture 8
 - ▶ Principles
 - ▶ Examples (hands-on)
- Validation → Lecture 8

Part 1

Part 2

Part 3

Part 4

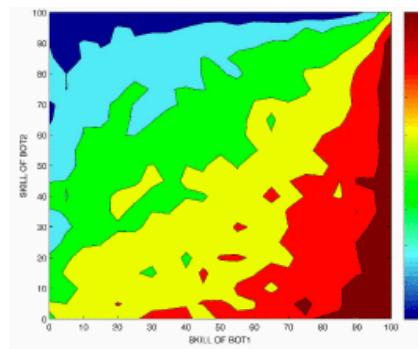
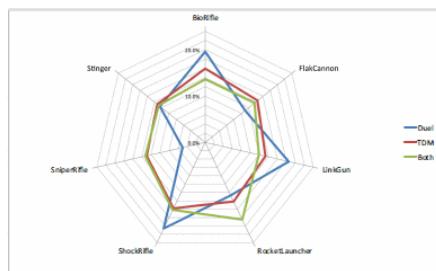
Why this course?

You (will) use data visualization

- Explain data and results...
 - ▶ ... in lectures
 - ▶ ... in presentations
 - ▶ ... in papers
- Analyze data and results...
 - ▶ ... to explore an unknown problem
 - ▶ ... to design a new solution of a problem
 - ▶ ... to assess an existing solution of a problem
- Create a tool to support yours and others decisions.

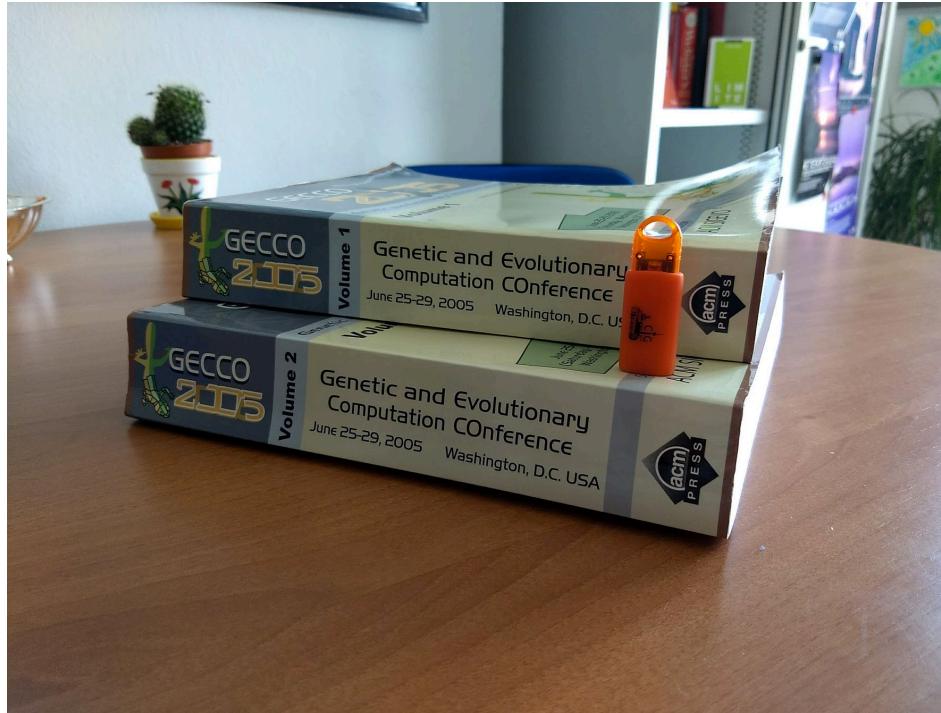
Disclaimer

- ❑ Data visualization is difficult to evaluate and mistakes are common (lack of time, misjudgment, constraints, etc.).



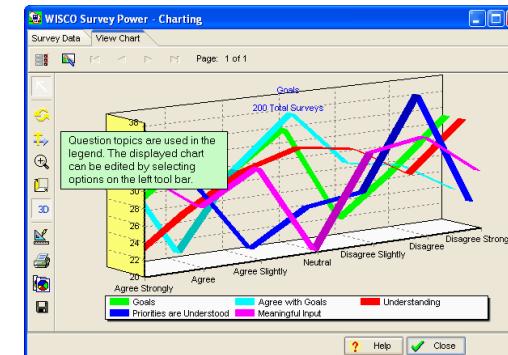
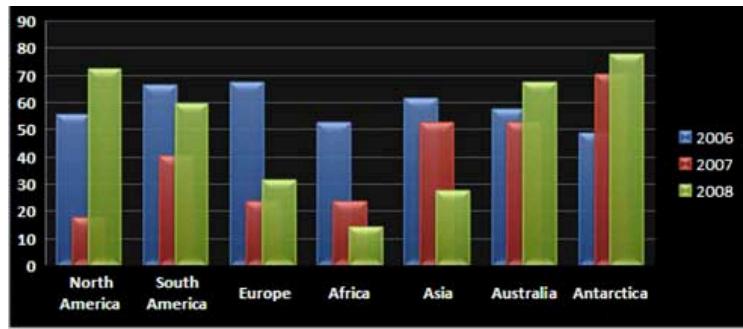
Disclaimer

- ❑ Data visualization is difficult to evaluate and mistakes are common (lack of time, misjudgment, constraints, etc.).
- ❑ Highly problem/domain dependent (also issues keep changing over time).



Disclaimer

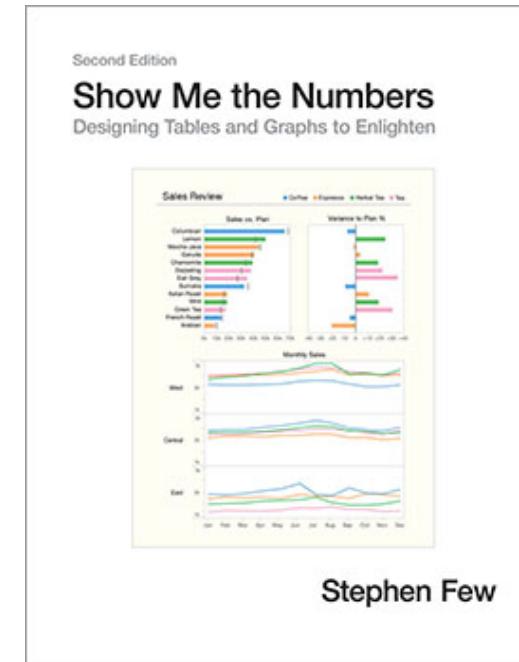
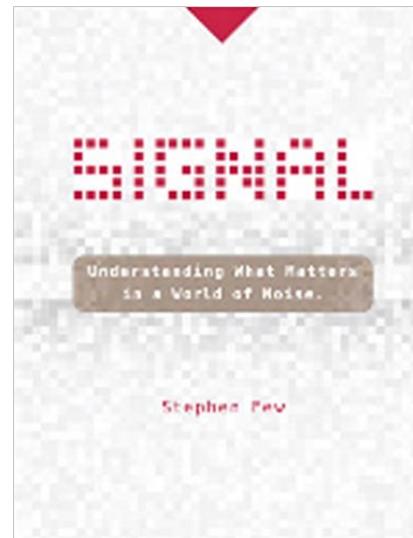
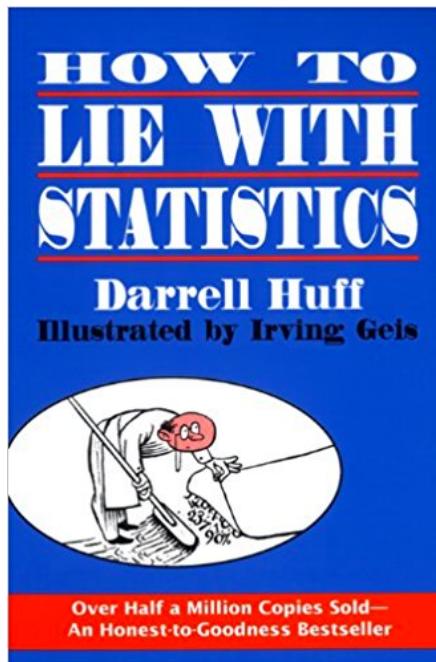
- ❑ Data visualization is difficult to evaluate and mistakes are common (lack of time, misjudgment, constraints, etc.).
- ❑ Highly problem/domain dependent (also issues keep changing over time).
- ❑ Probably you won't become a data viz master, but (hopefully) you'll learn to be more effective and (at least) avoid this:



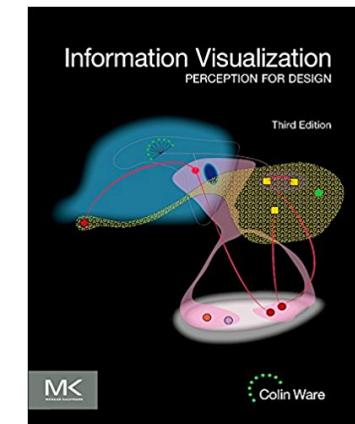
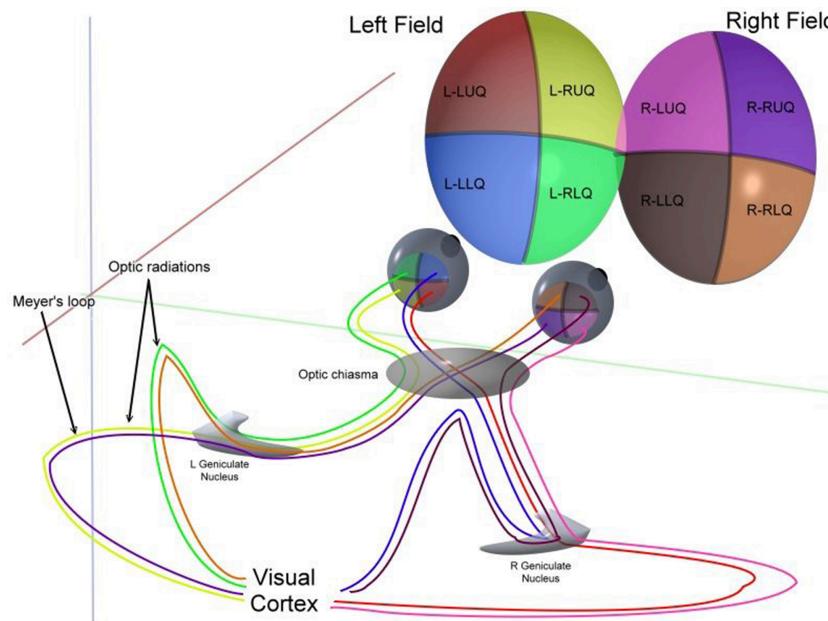
What you will not learn?

Results presentation

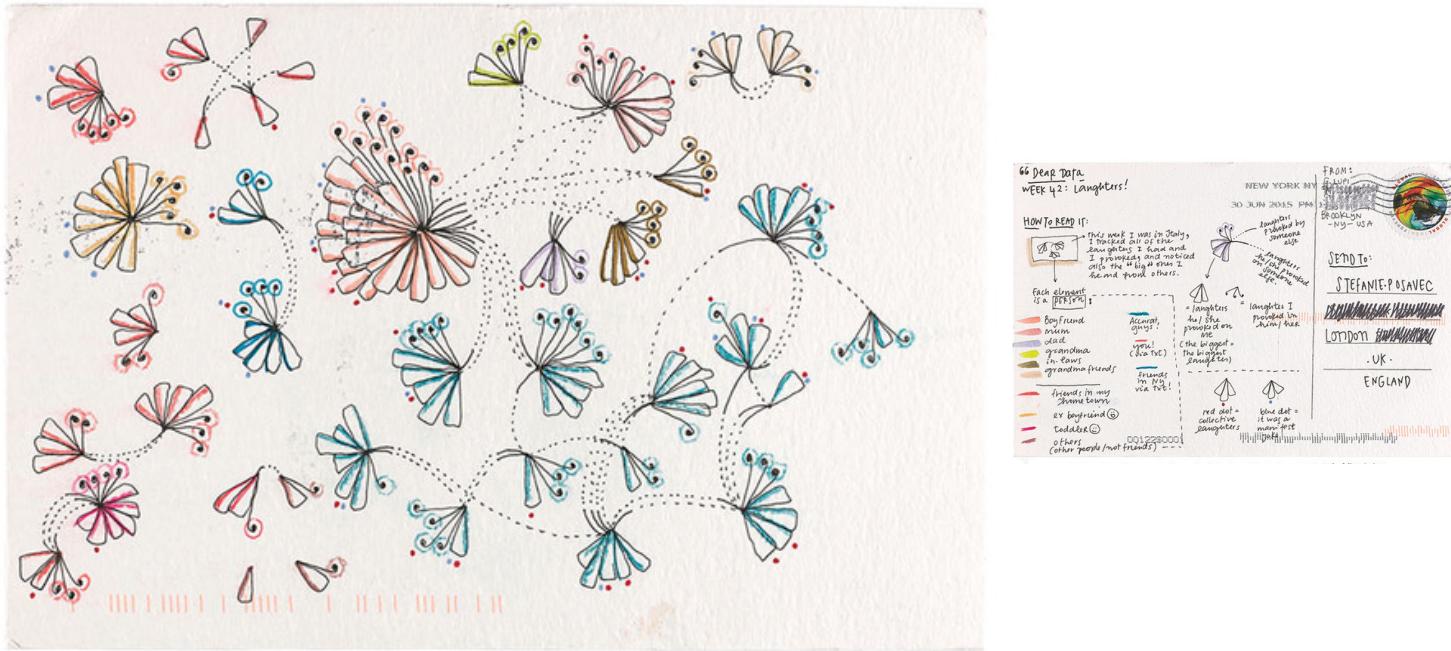
- ❑ How to compare and rank results?
- ❑ How to design a table?
- ❑ What better means? (yes, you do need statistics ☺)



Human visual perception mechanisms

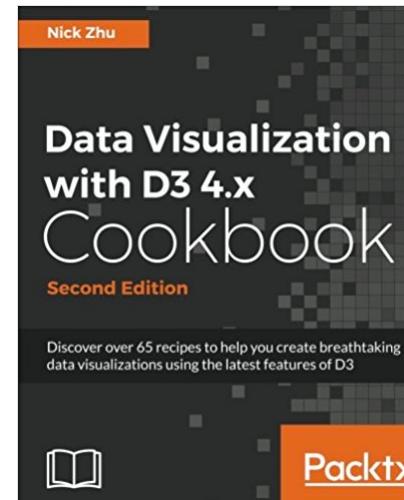
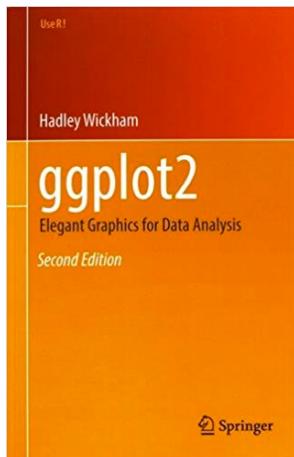
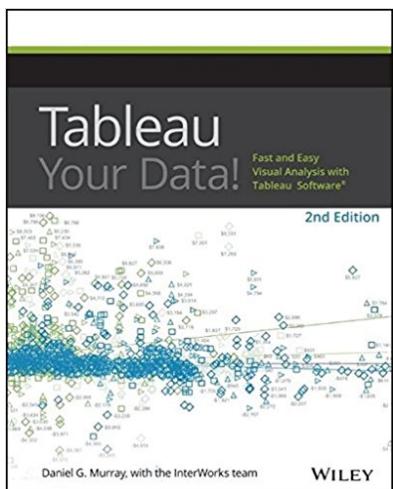


Design beautiful data visualization



From <http://www.dear-data.com/>

Master one or more specific tools



Resources and software

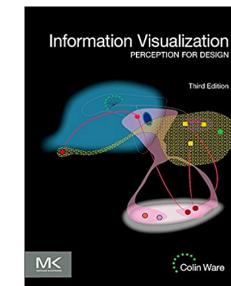
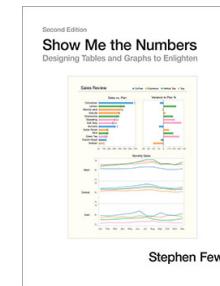
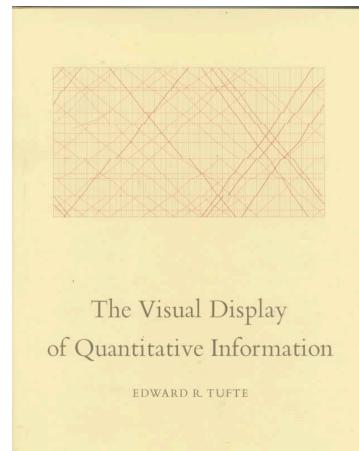
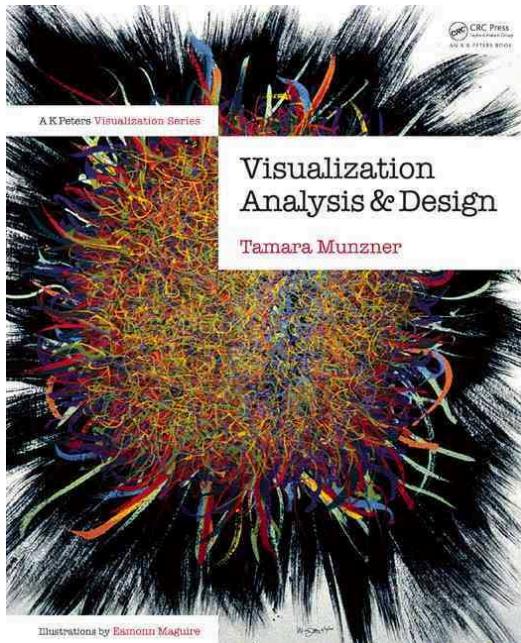
□ Web resources

- ▶ <https://datavizcatalogue.com/>
- ▶ <http://junkcharts.typepad.com/>
- ▶ <https://python-graph-gallery.com/>
- ▶ <http://www.thefunctionalart.com/>
- ▶ <http://www.perceptualedge.com/>
- ▶ <http://chartporn.org/>

□ Software to install

- ▶ [Anaconda](#) (Matplotlib, Seaborn, Pandas, Pynum)
- ▶ [Tableau Public](#)

Bibliography



Most of the materials used in this course are based on or taken from Tamara Munzner's book and slides.