
DTU

Mads Paulsen, DTU Management

Data Visualisation and Communication

Contains figures and other relevant content from:
Wilke, C. O. (2019). *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media. https://clauswilke.com/dataviz/.

Mads' Part of 42588 Data & Data Science

- Week 6 – Data Visualisation & Communication
- Week 7 – Spatial Data
- Week 8 – Project 2 Presentations + Data Weighting and Imputation

- What I will do:
  - Prepare and design classes that aim to optimize your learning
  - Be available to answer and discuss any question that you may have
  - Take the feedback I receive seriously and adjust accordingly

- What I will ask of you:
  - Active participation when you come to class
  - Provide feedback so that I can improve my teaching

Learning Objectives

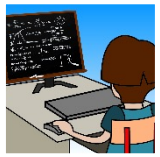
- From kursusbasen:
  - Conduct exploratory data analysis and visualization

- More specifically, after today you should be able to:
  - Explain and apply the concept of data-ink ratio when analysing figures
  - Identify and discuss pros and cons of a data visualisation figure
  - Design a data visualisation figure that supports your storytelling (and tell the story)

Agenda

Figure Fight

Storytelling & Good Practice



Live-coding Session

Storytelling

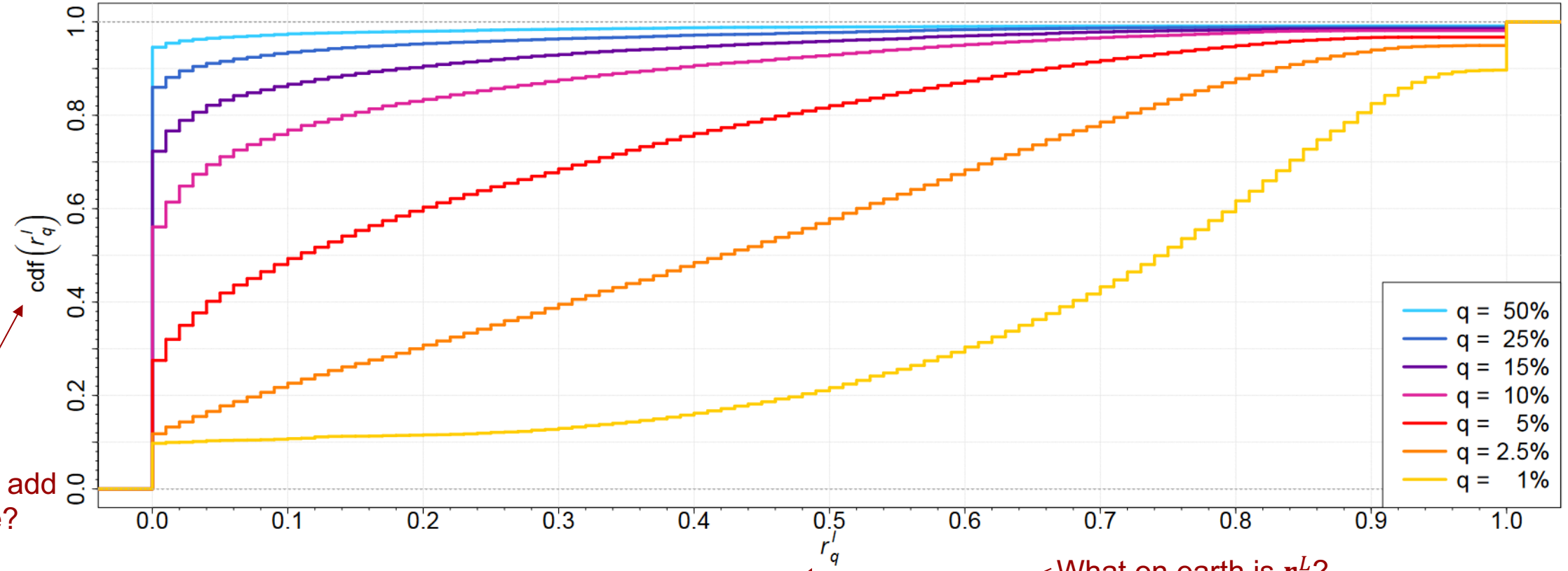
Work on Project 2

Data Visualisation Analysis

Based on Wilke (2019), Chapter 29: *Telling a story and making a point*

Not my proudest figure…

How did I ever think that this was a good way to tell a story to someone else?



Does this add any value?

Fig. 3. Empirical cumulative distribution function of $\mathbf{r}_q^L$. for various values of q.

What on earth is $\mathbf{r}_q^L$?

Why is **r** boldface here but not in label?

What's up with this dot?

Figure from Paulsen, M., Rasmussen, T. K., & Nielsen, O. A. (2018). *Output variability caused by random seeds in a multi-agent transport simulation model*. Procedia computer science, 130, 850-857.

Data visualisation communication

- We use **data visualisations** to:
  - **Obtain insights** about a problem
  - **Communicate insights** to others

- To improve communication we **analyse data visualisations**
  - We analyse the content (e.g. which variables)
  - We analyse the format (e.g. type of visualisation)
    – ”Analysis of visualisations of analysis of data”

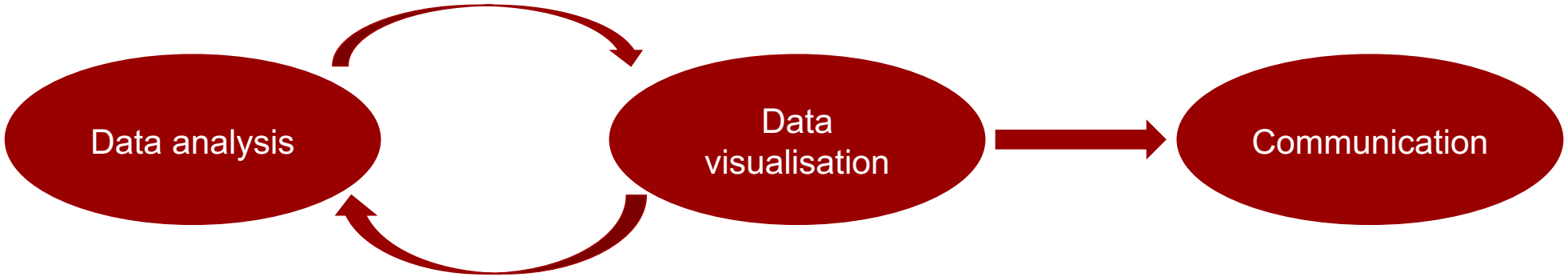
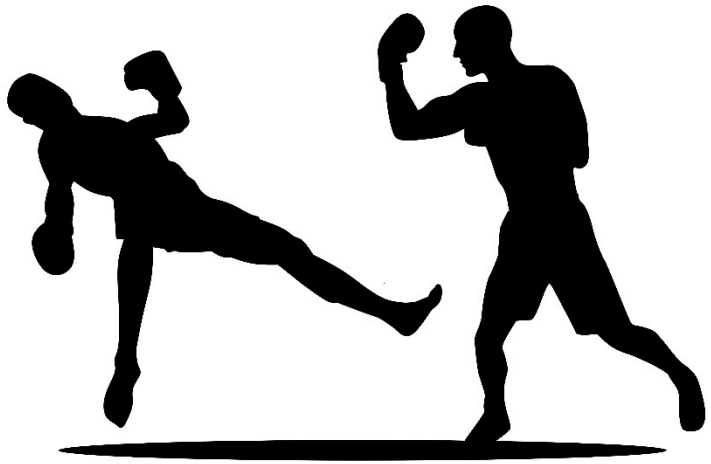


Figure Fight - Format

- I will distribute 4 different figures (A-D) among you
  - They all visualise the same data in different ways

- The game consists of 6 rounds
  - 1st round is individual
  - 2nd round is with your ”figure mates”
  - 3rd-5th rounds are fought 1 vs 1
  - After round 1-5 you will answer survey questions on DTU Learn
    - » Week 6 → Figure Fight → Figure Fight Survey
    - » We will use the collected data later today!

- In round 6 we pick up on your findings



Survey to be found here on DTU Learn:
Week 6 → Figure Fight → Figure Fight Survey

Figure Fight - Content

- What does the histograms show?

- Are **both distributions** shown in a way that can **be interpreted easily**?
- Are the **two distributions** easily **comparable**? Easily **distinguishable**?
- What seems to be the **data range** for males and females?
- Are the **bin sizes appropriate**, too small or too large?

- For 3rd-5th round only:
  Which one do you **prefer?**

- How is the **data-ink ratio**?
- Is it easy to understand the **legend**? Is it **necessary**?
- Are there any **grid lines**? Are they **helpful**?
- Is the **font size appropriate**?
- Does the figure have a **meaningful caption**?

Figure Fight Takeaways

- All the figures had weaknesses!

- Bin size
  - …
- Grids
  - …
- Caption
  - …
- Stacked / Shaded / Population
  - …
- Hidden data…?
  - …
- Font size
  - …
- Data-ink ratio
  - …
- Trade-offs (e.g. compare vs distinguish)
  - …

Break

- If you want to learn more about *data-ink ratio*, check out this 4 minute YouTube video…

Agenda

Figure Fight

Storytelling & Good Practice

Live-coding Session

Storytelling

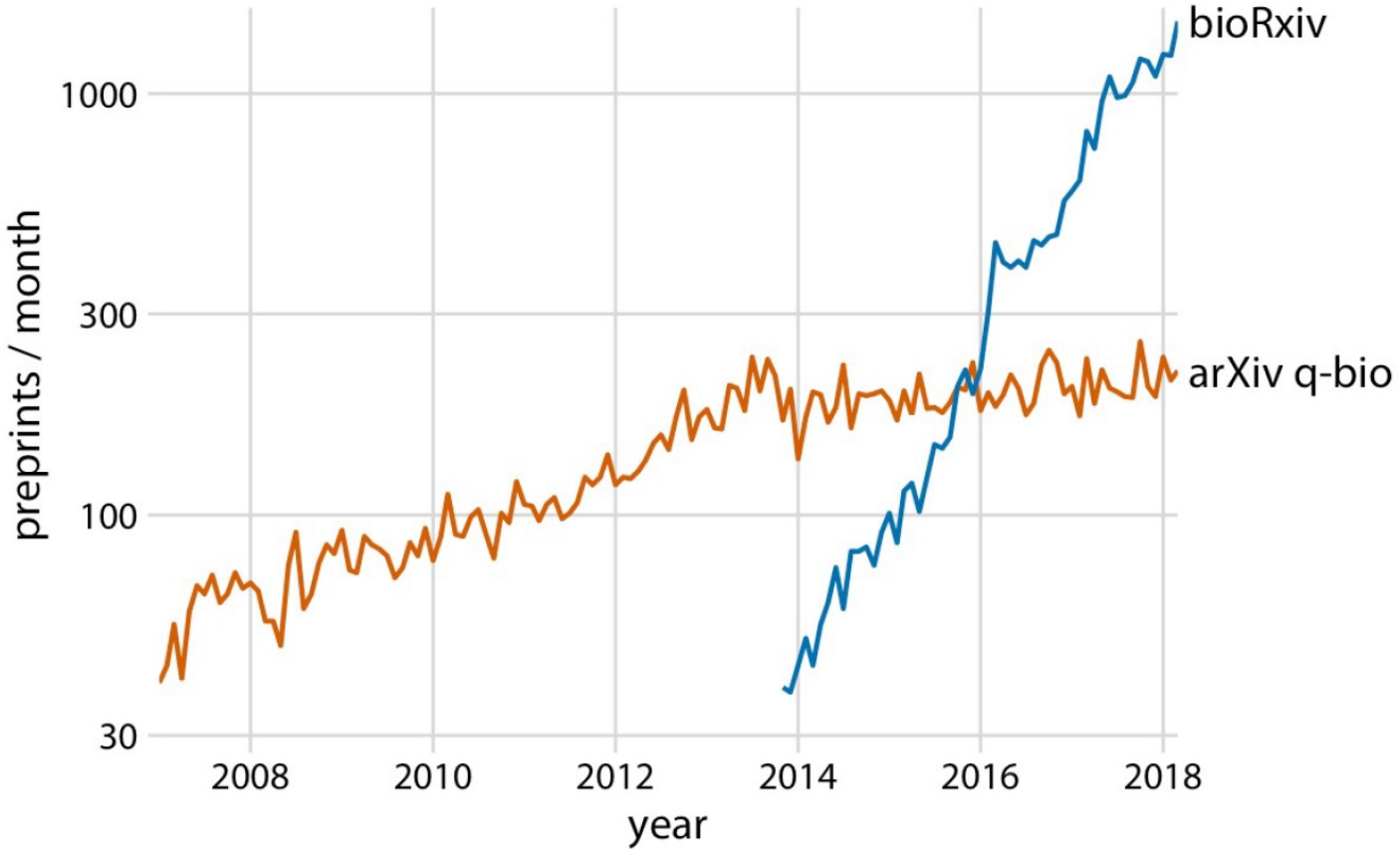
Work on Project 2

Storytelling and Good Practice

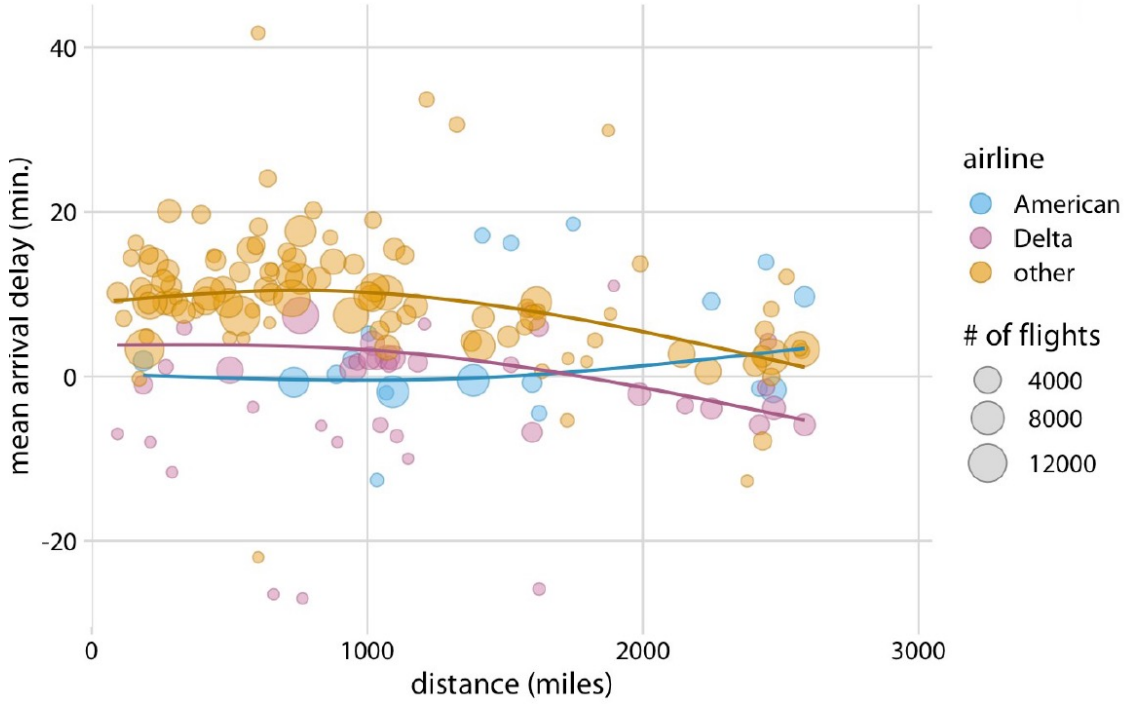
Based on Wilke (2019), Chapter 29: *Telling a story and making a point*

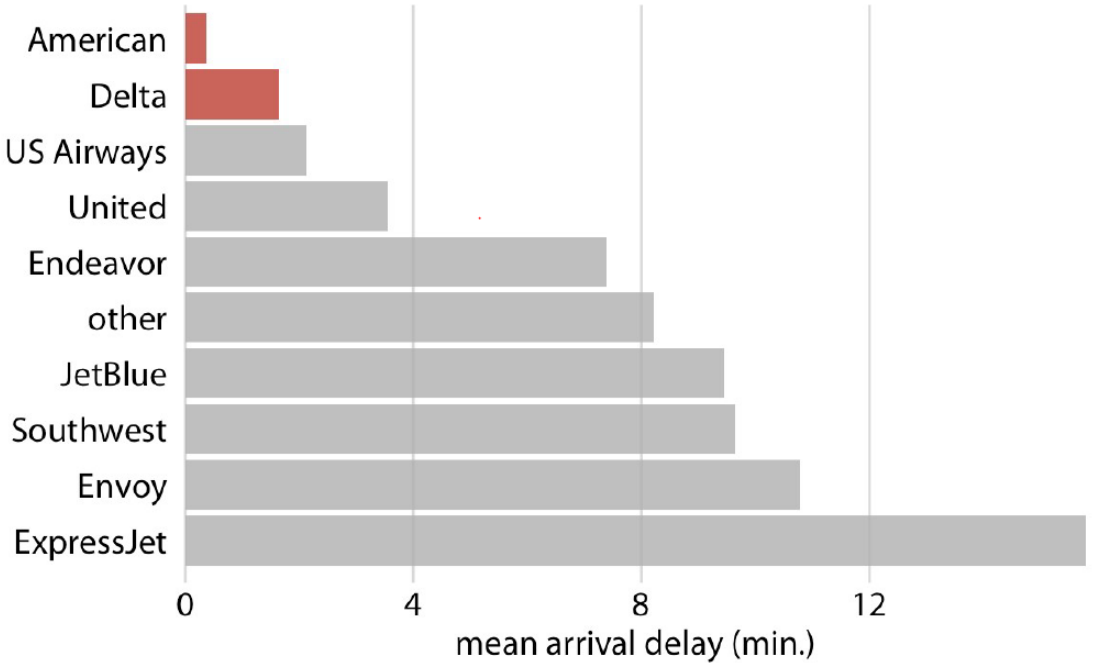
Visualisations guide the story

- Did the growth in bio research stop?



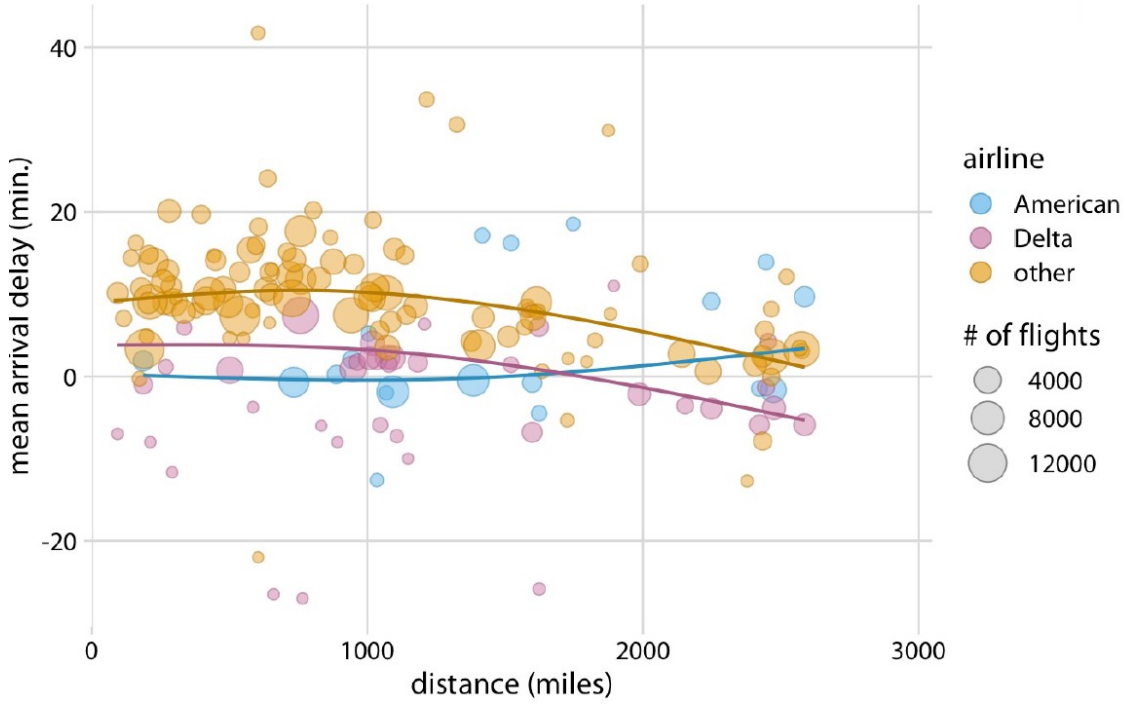
Which one conveys the message best?

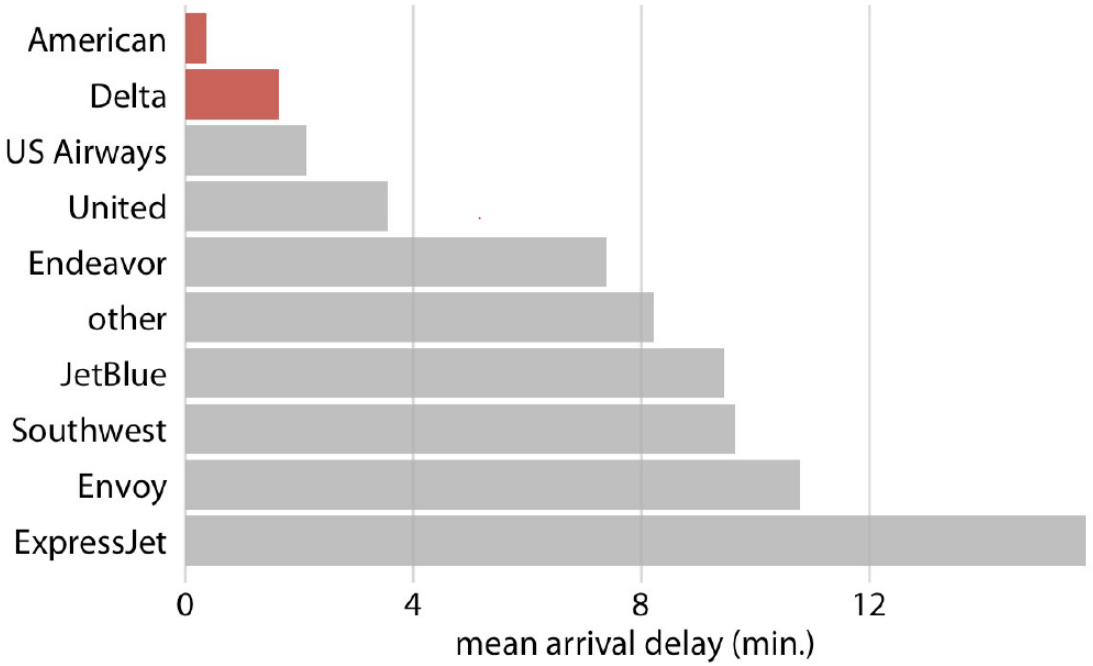




Crystal clear key message

” When you’re trying to show too much data at once you may end up not showing anything.” – Wilke (2019)



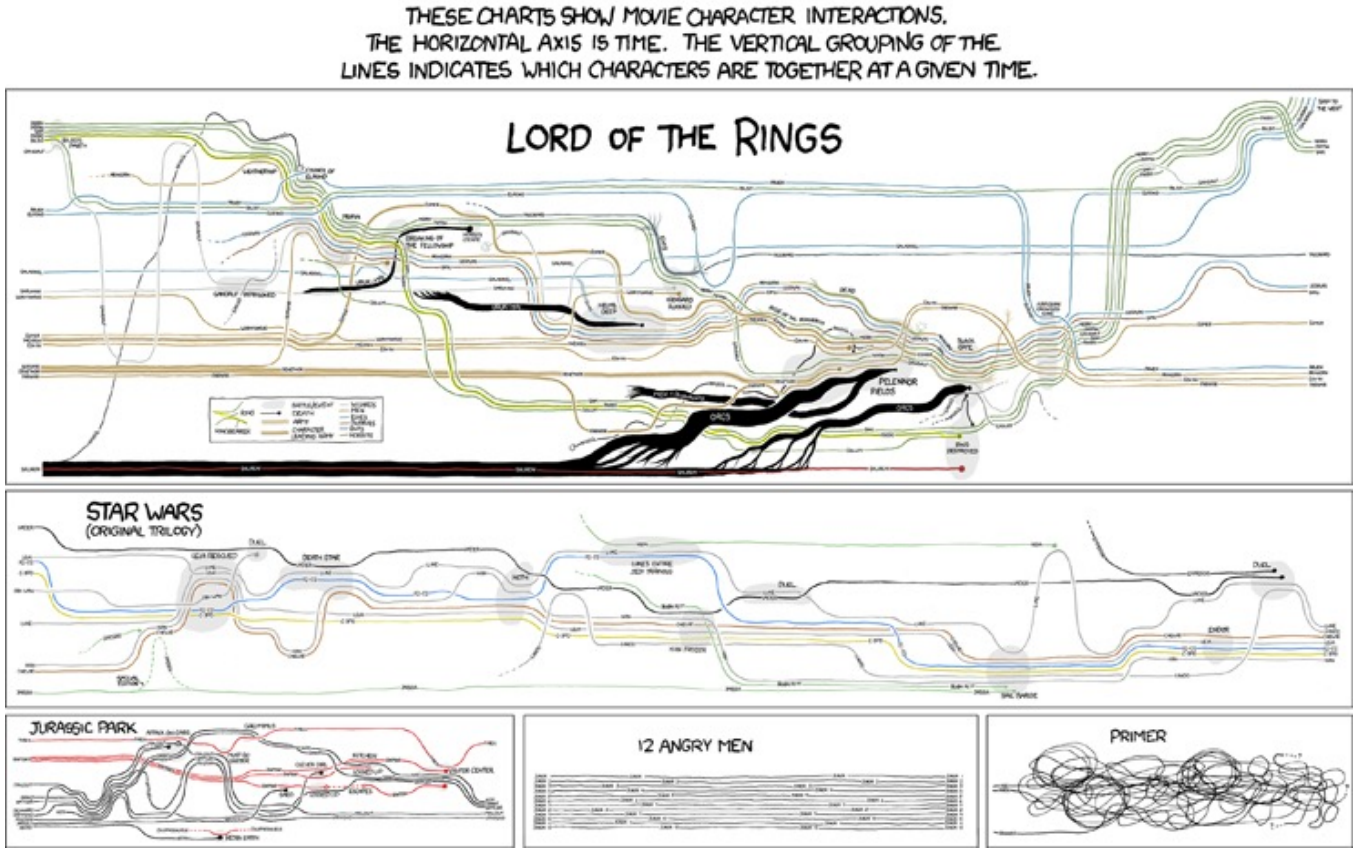


Know your audience

"Never assume your audience can rapidly process complex visual displays"

– Wilke (2019)

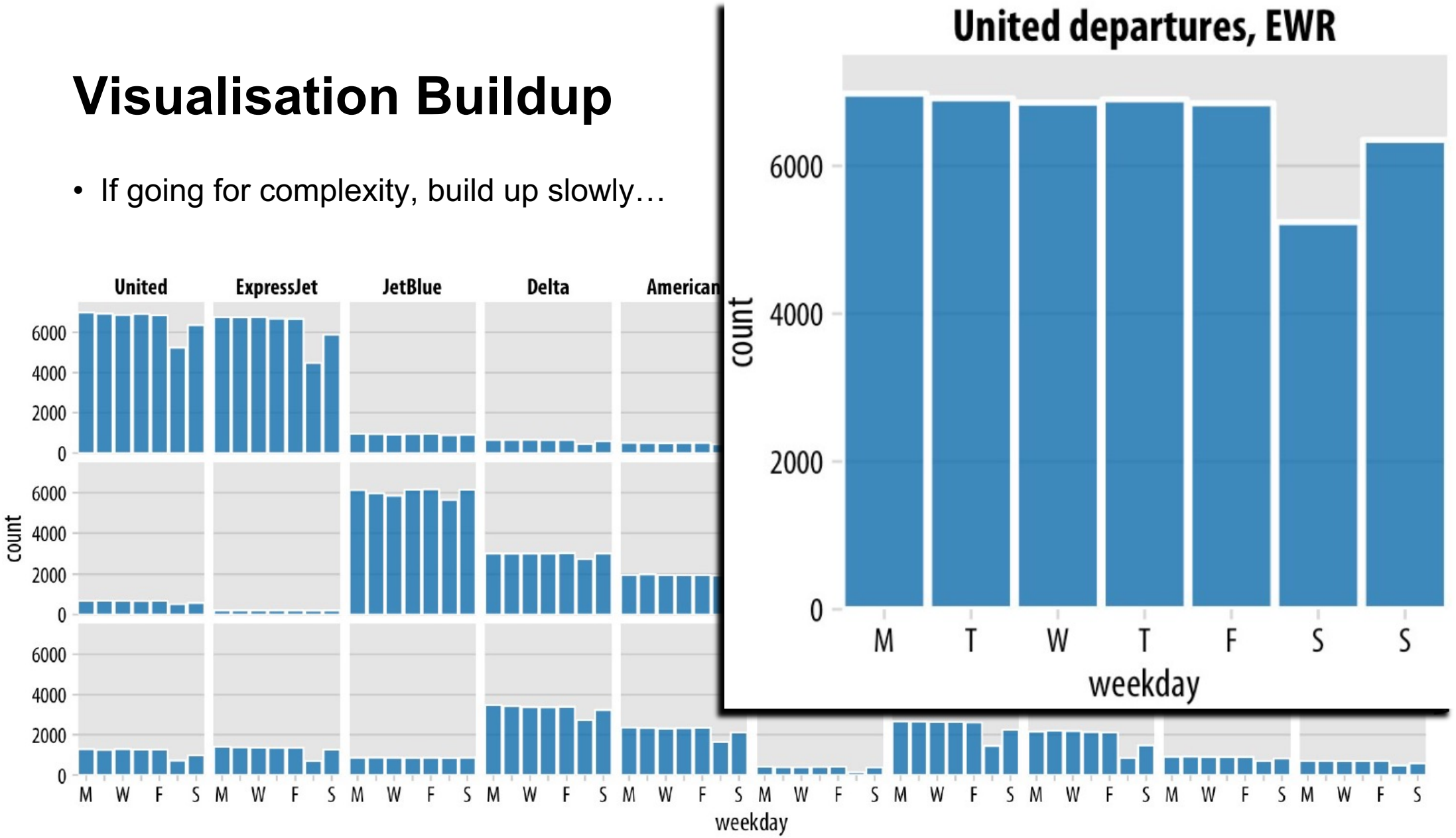
- Your audience…
  - … does not know the data
  - … does not know the underlying work
  - … have never seen the figure before
- Prepare figures accordingly
- Help the reader!



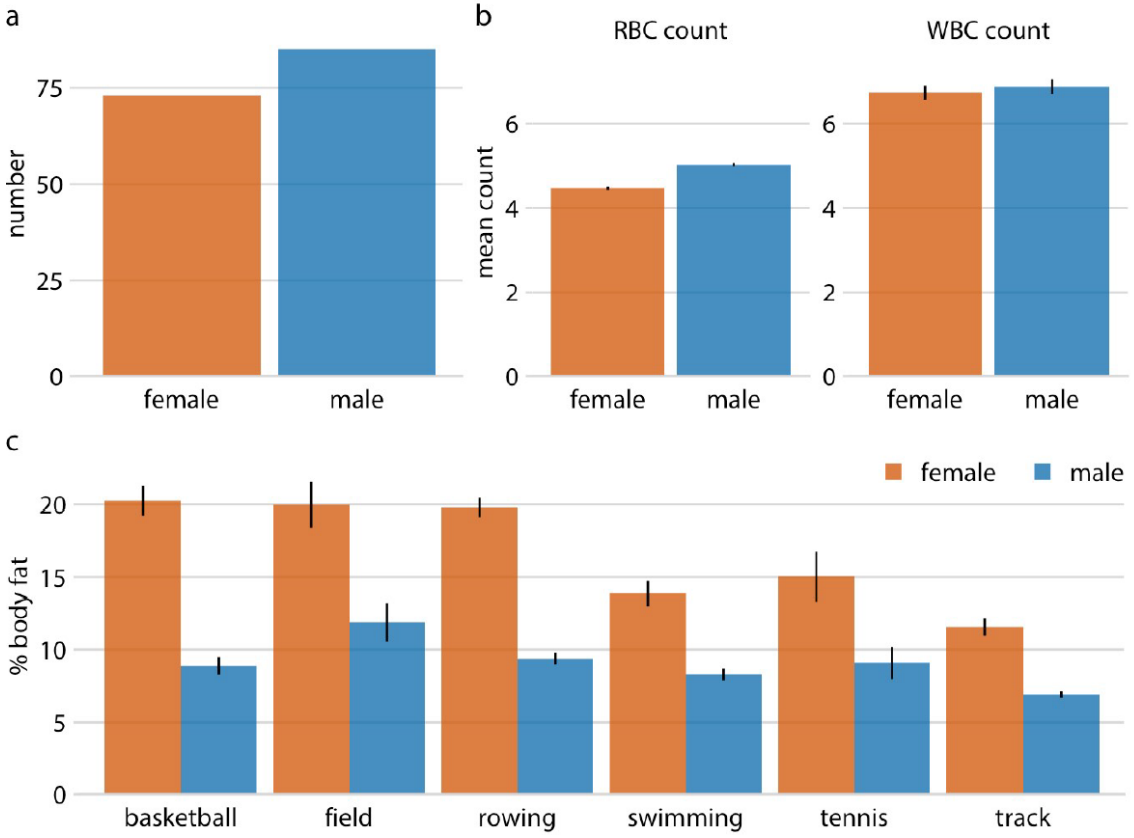
https://xkcd.com/657/

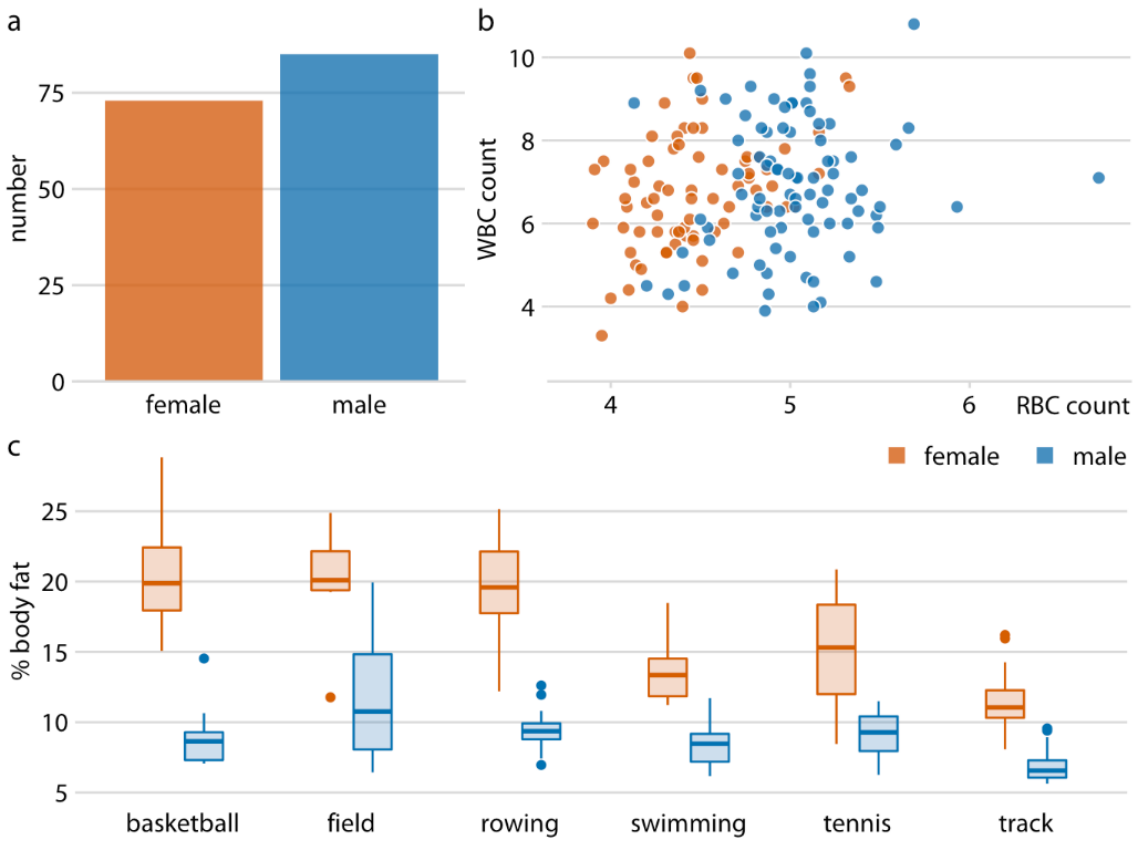
Visualisation Buildup

- If going for complexity, build up slowly…



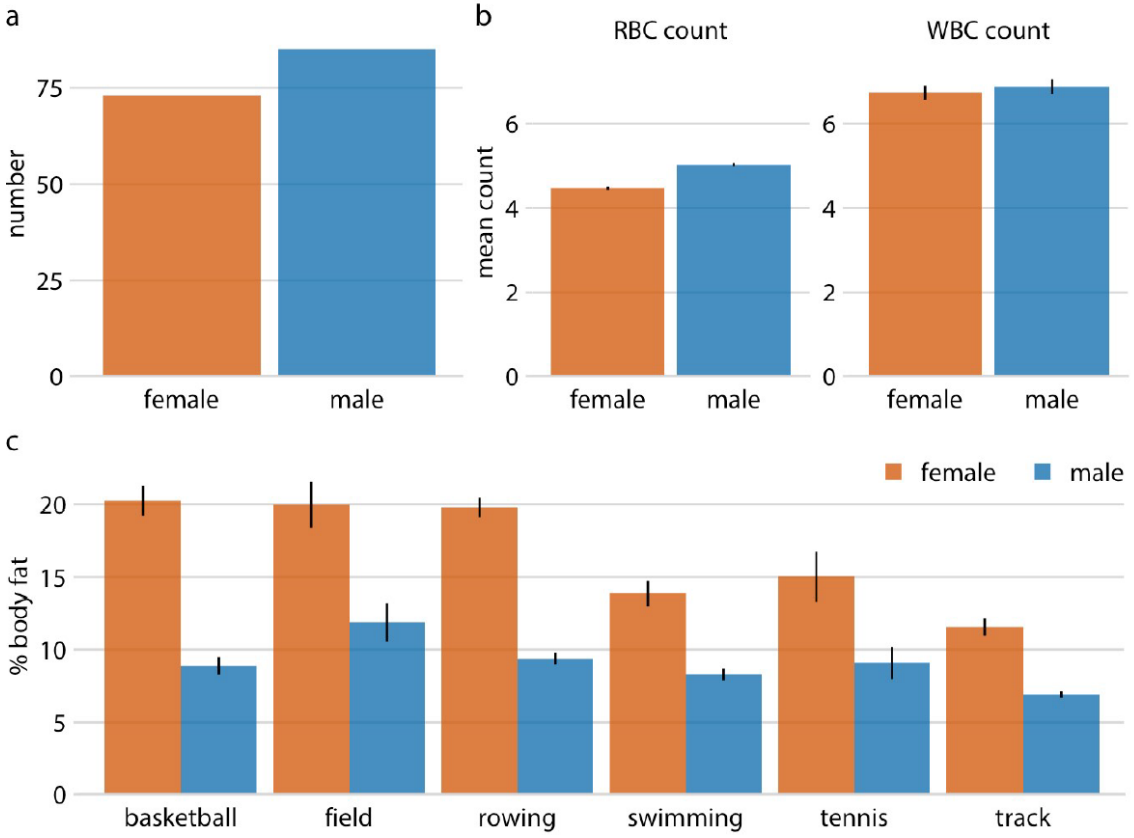
Which one is most appealing?

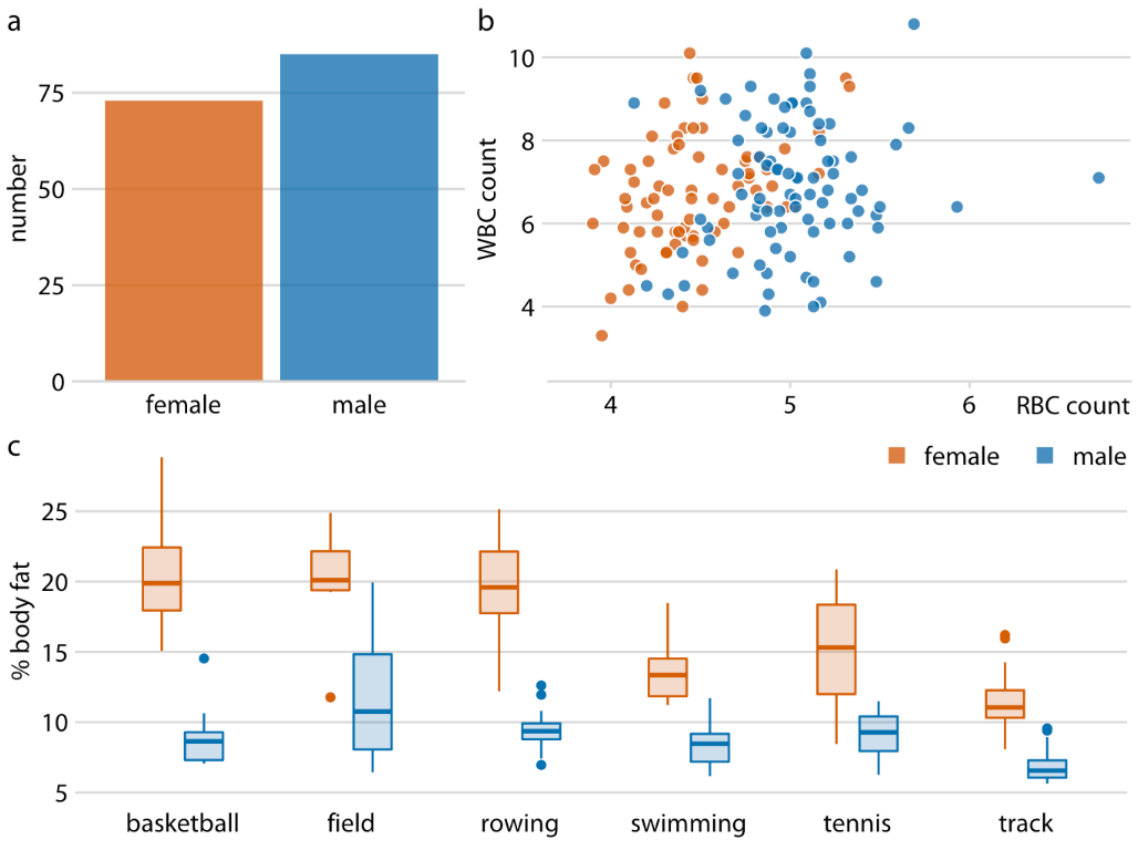




Avoid repetition

"When preparing a presentation or report, aim to use a different type of visualization for each distinct analysis." – Wilke (2019)





Memorable figures?

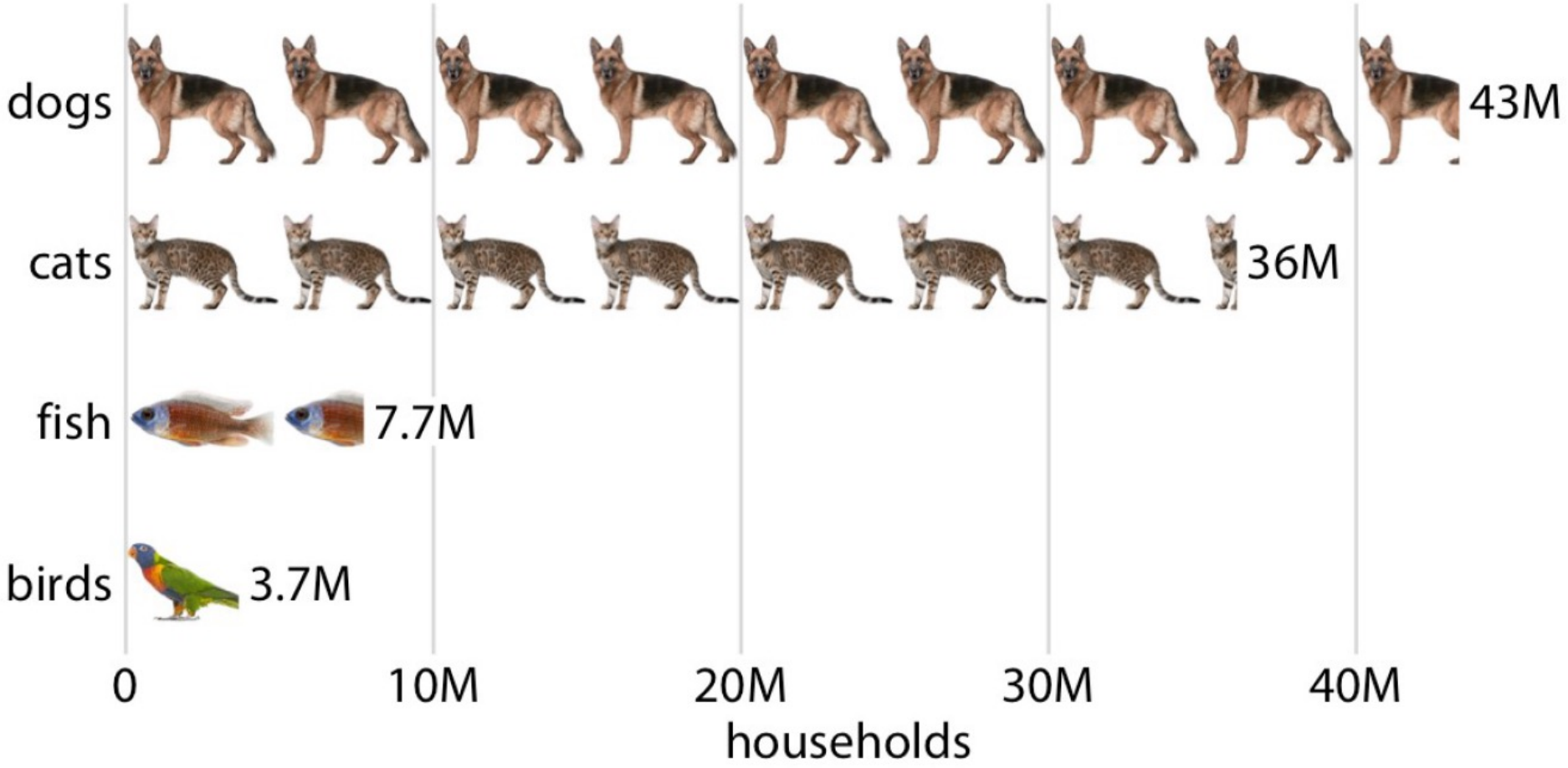
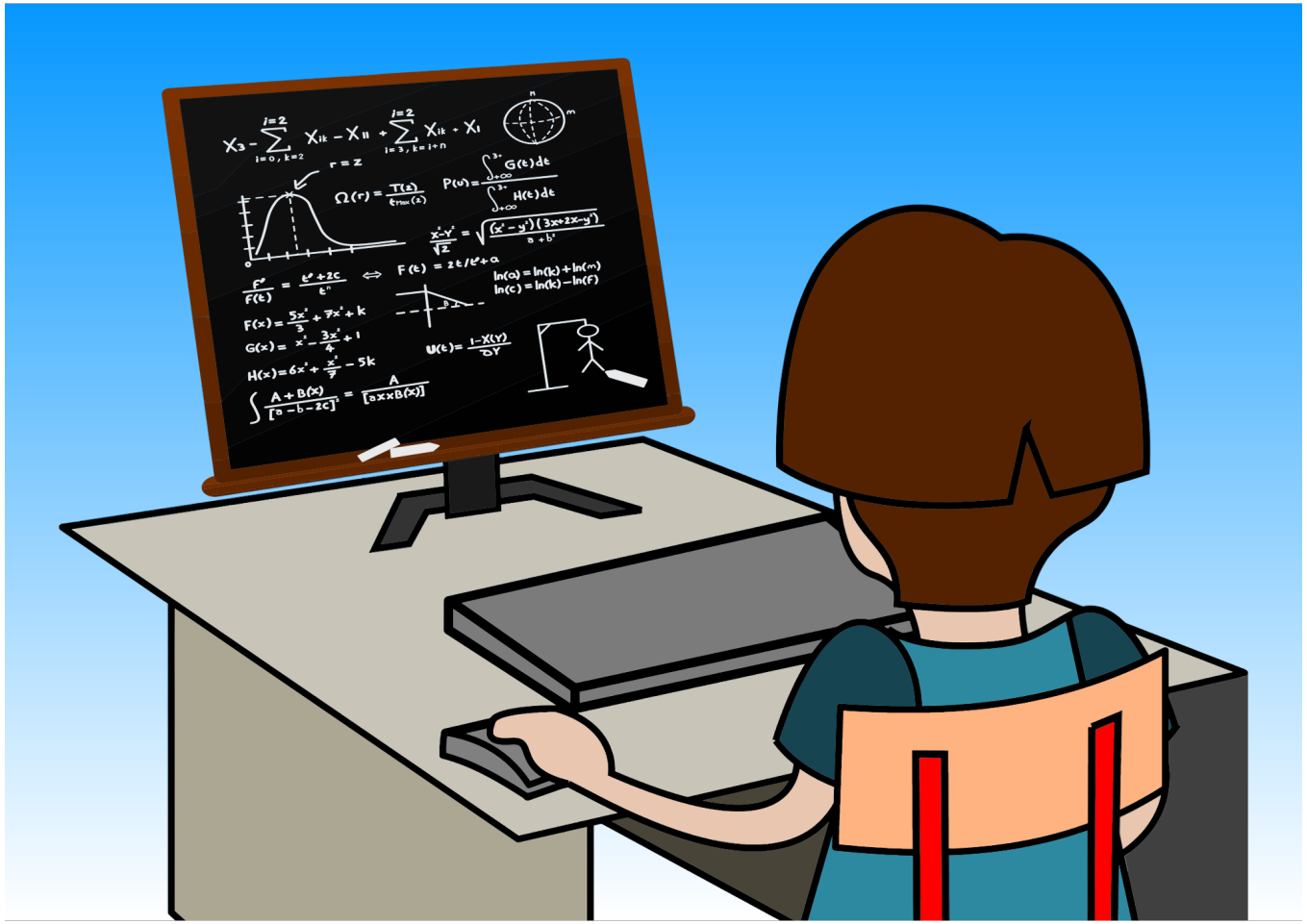


Figure 29.9: Number of households having one or more of the most popular pets, shown as an isotype graph.

Live coding of our results!



Storytelling Exercise

- Please find instructions for the activity on DTU Learn:
  - Week 6 → Storytelling

Break

“Never assume your audience can rapidly process complex visual displays” – Wilke (2019)

”When you’re trying to show too much data at once you may end up not showing anything.” – Wilke (2019)

“When preparing a presentation or report, aim to use a different type of visualization for each distinct analysis.” – Wilke (2019)

Agenda

Figure Fight

Storytelling & Good Practice

Live-coding Session

Storytelling

Work on Project 2

Work on Project 2

It All Begins with a Good Visualisation….

- You have been working on something for half a year, and are to present it to
  - The board
  - Your manager
  - Decision makers
  - Politicians
  - A reviewer

- You have built the entire analysis into one key figure…

  - *“Sorry… What did that figure show again?”*

Time for Project Work…

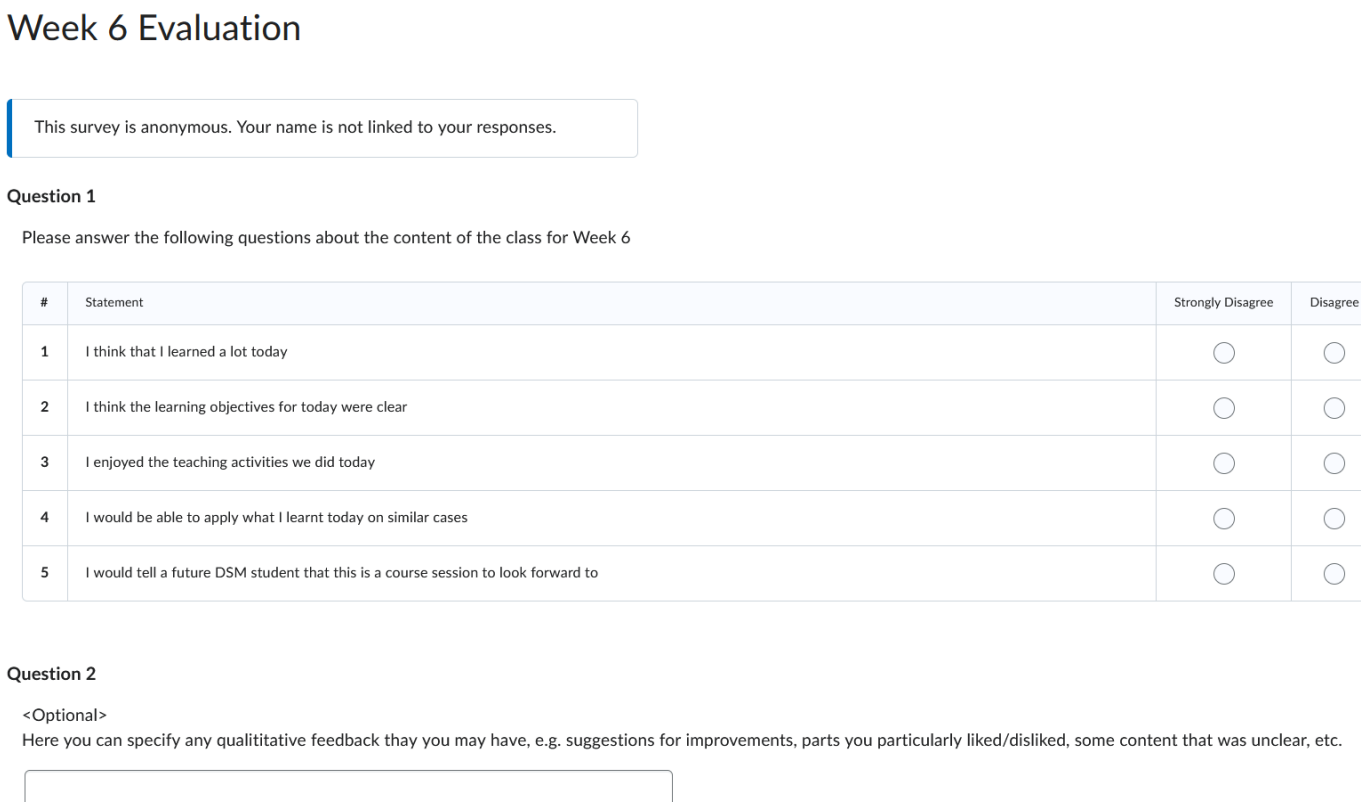
- You can find a lot of content in Wilke (2019)
- Read the content when relevant for your project work, such as:
  - Database of data visualisation types (Chap. 5)
  - Specific chapters on visualising amounts (Chap. 6), histograms and density plots (Chap. 7), and proportions (Chap. 10)
  - Data-ink ratio (Chap. 23)
  - Support a story with data visualisations (Chap. 29)
  - Label sizes(!) (Chap. 24)

- Check the project description on DTU Learn and align your work accordingly

- Presentations on the 20th – For those with DSE-related conflicts, let met know ASAP (when we start the group work). I will then prepare a schedule that fits

Feedback on Today's Class

- Evaluation form ready on DTU learn for today's class
- Week 6 → Feedback → Week 6 Evaluation
- Will improve teaching for next week and future DSM students



Week 6 Evaluation

This survey is anonymous. Your name is not linked to your responses.

Question 1

Please answer the following questions about the content of the class for Week 6

#	Statement	Strongly Disagree	Disagree
1	I think that I learned a lot today	○	○
2	I think the learning objectives for today were clear	○	○
3	I enjoyed the teaching activities we did today	○	○
4	I would be able to apply what I learnt today on similar cases	○	○
5	I would tell a future DSM student that this is a course session to look forward to	○	○

Question 2

<Optional>
Here you can specify any qualititative feedback thay you may have, e.g. suggestions for improvements, parts you particularly liked/disliked, some content that was unclear, etc.