UGANDA CHRISTIAN UNIVERSITY
A Centre of Excellence in the Heart of Africa

Data Science Lifecycle

MSDS-1

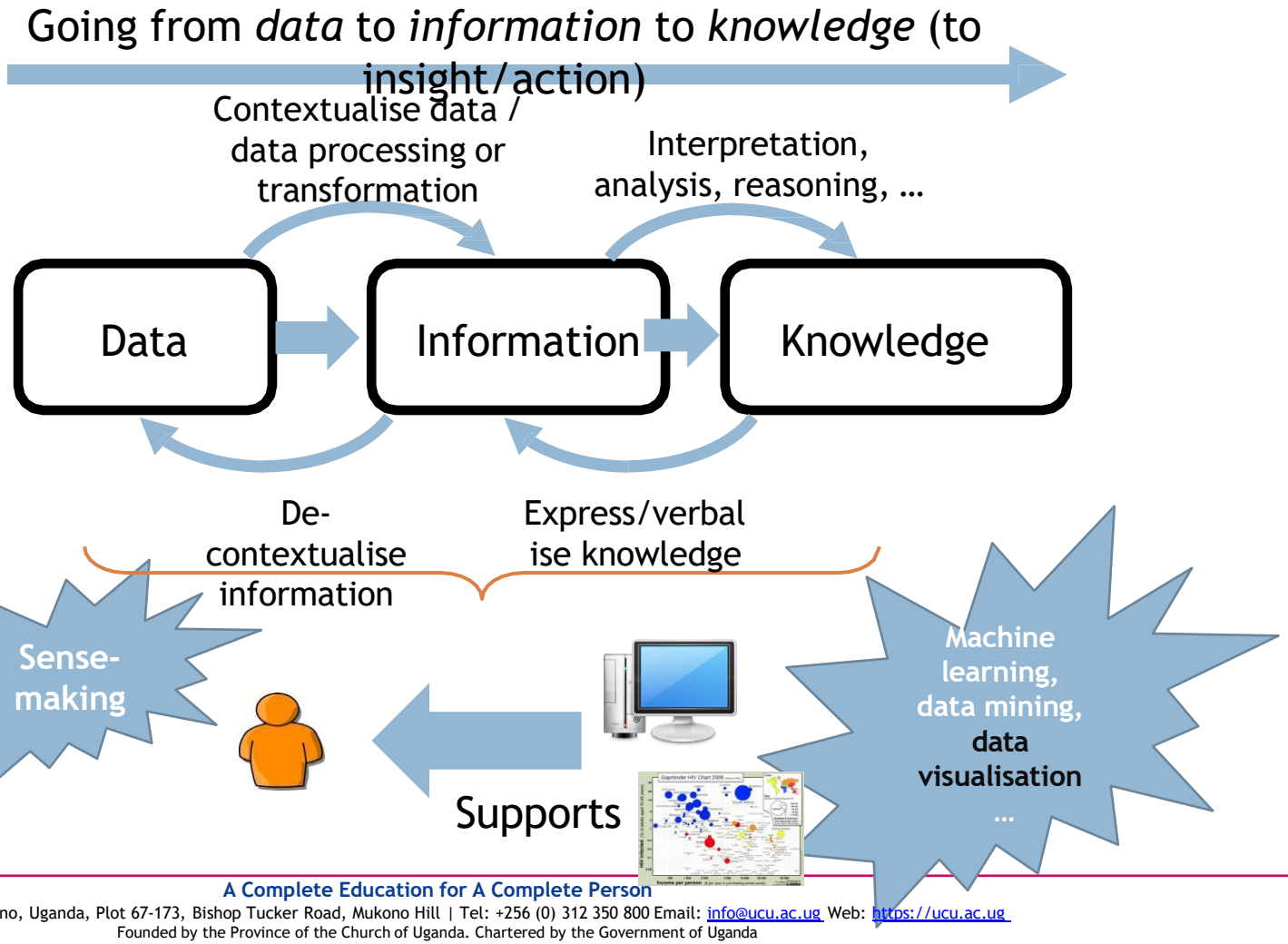
Data Visualization and Storytelling

Ms. Immaculate Kamusiime
ikamusiime@ucu.ac.ug
(MSc, BSc)
Department of Computing & Technology
Faculty of Engineering, Design & Technology

Lecture 3

Using data visualization to tell a story

The Knowledge Cycle



Apply knowledge, e.g. gaining insight, problem-solving, decision-making, etc.

Data storytelling journey

Effective data storytelling involves 2 questions

Effective data storytelling involves 2 questions

- WHAT IS AN INSIGHT?
  - "Intuition is the use of patterns they've already learned, whereas insight is the discovery of new patterns.
    - "— *Gary Klein, in "Seeing what others don't: The remarkable ways we gain insights"*

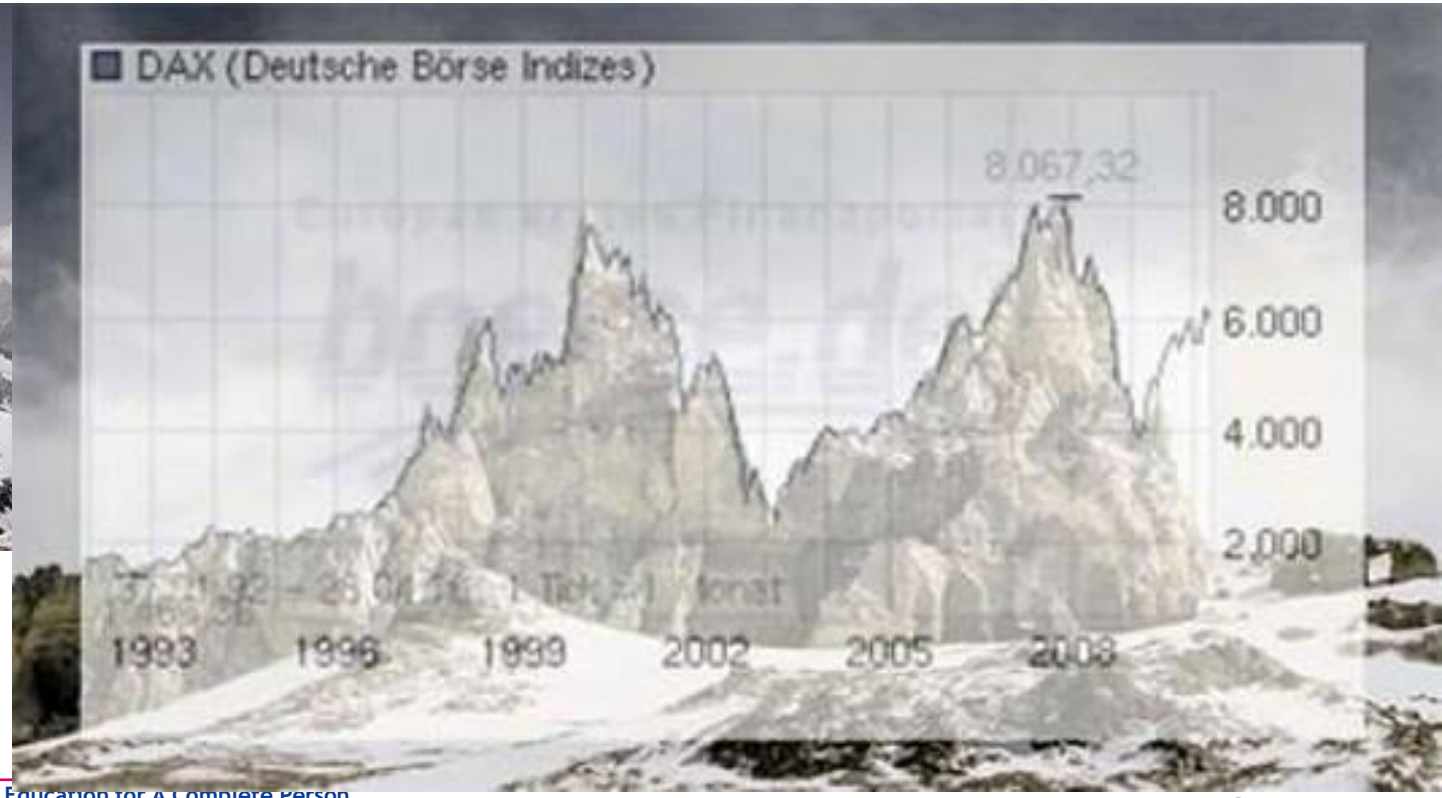
Visual Metaphors

•

http://www.lokeshdhakar.com/2007/08/20/an-illustrated-coffee-guide/

Visual metaphors for real-life phenomena

- Michael Najjar's Variations in Stock market price rendered as mountains (https://www.michaelnajjar.com/)

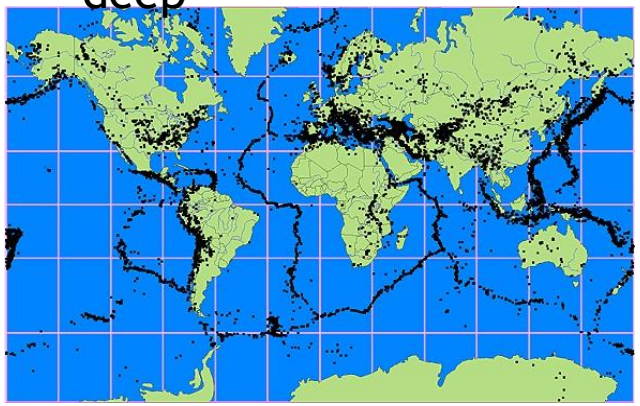


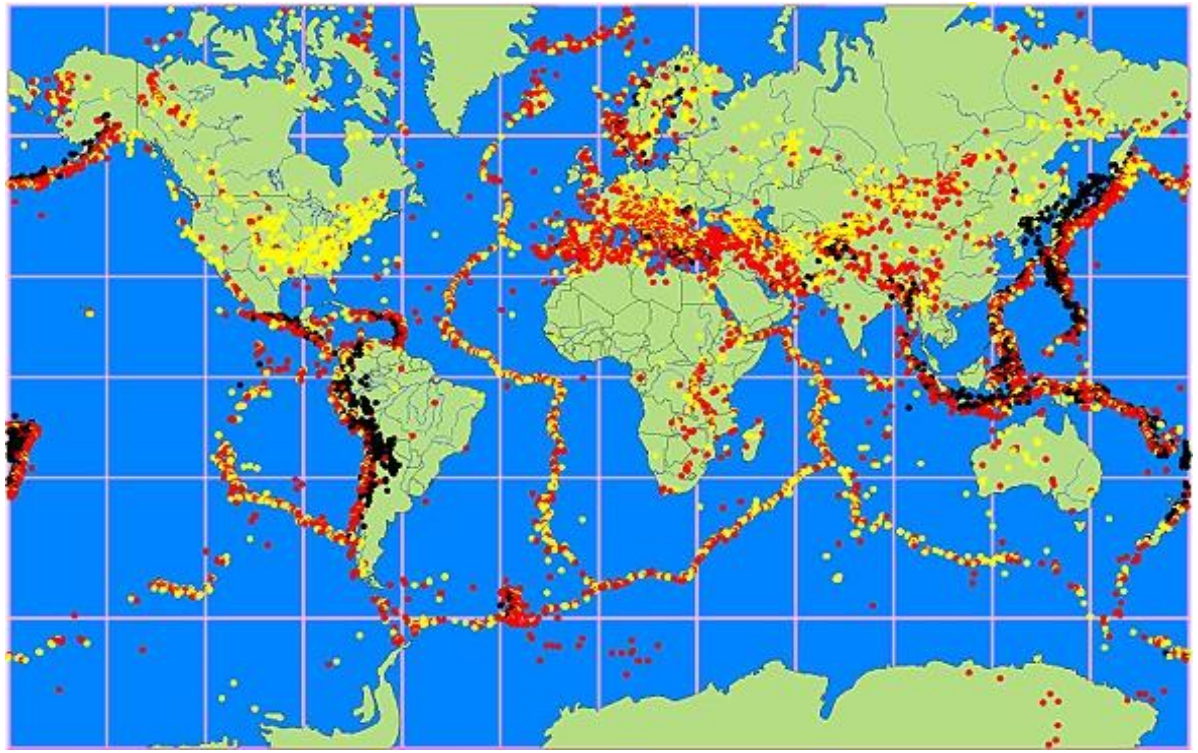
http://infosthetics.com/archives/2010/03/high_altitude_the_stock_market_trends_as_realistic_mountain_ranges.html

Example: earthquakes

Earthquake events organized according to depth (attributes of location of each earthquake - feature)

- shallow
- intermediate
- deep





http://www.physicalgeography.net/fundamentals/2f.html

Distribution of earthquake events that have occurred over the last century

The process of creating visualisations

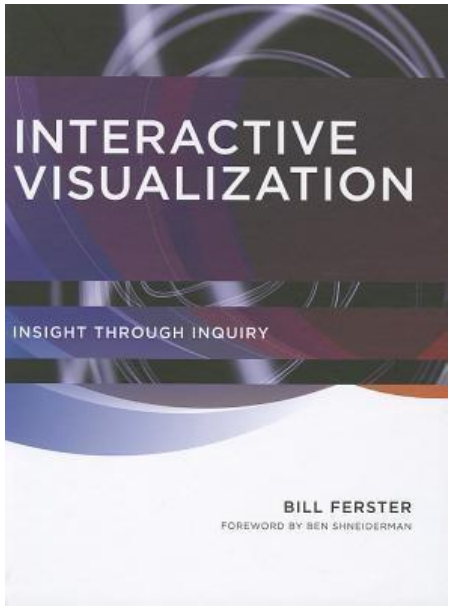
A quick overview of existing frameworks and theories

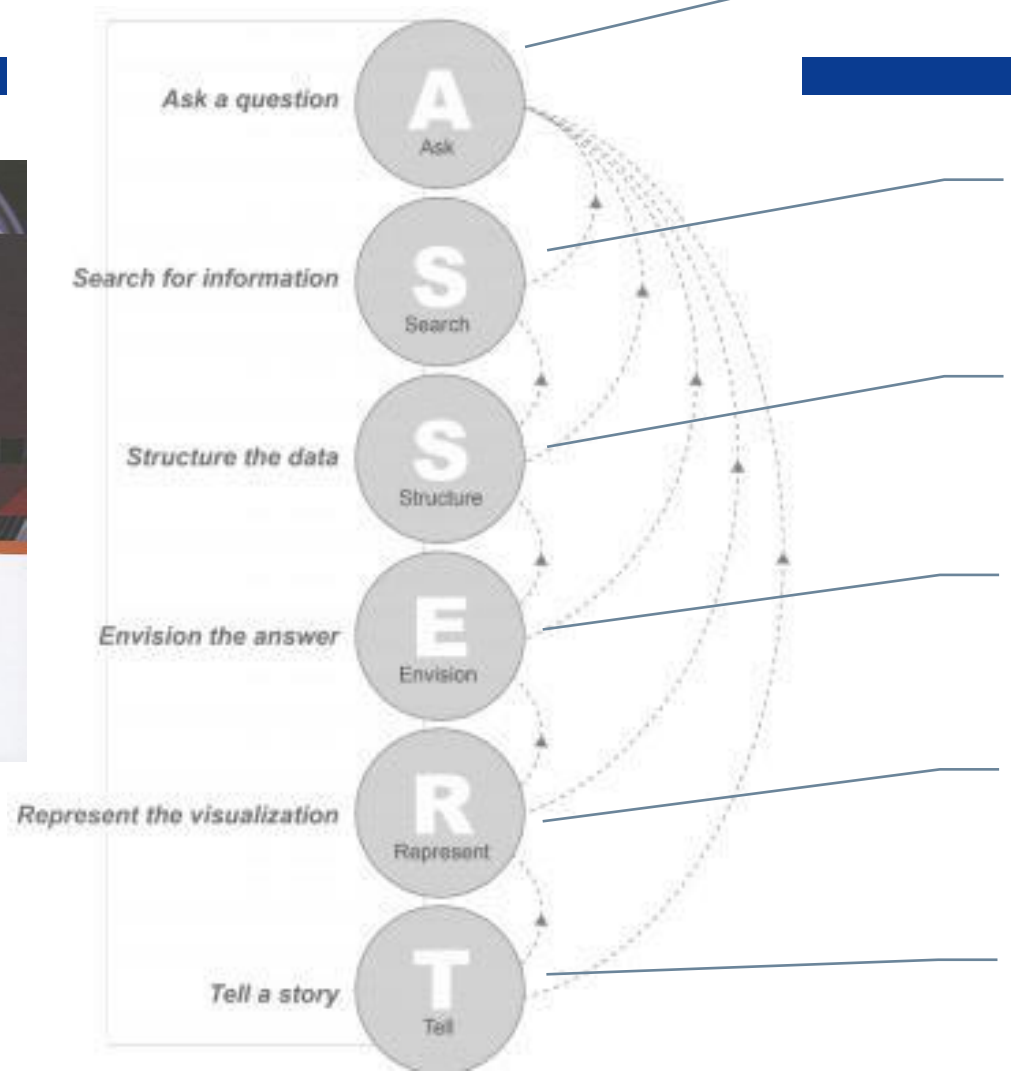
Seven stages of data visualisation

- Ben Fry (2008) in his visualising data book proposes the following process:
  - **Acquire**: obtain the data
  - **Parse**: provide structure and order into categories
  - **Filter**: remove all but the data of interest
  - **Mine**: explore and discern (statistical) patterns
  - **Represent**: chose basic visual model (e.g. bar chart)
  - **Refine**: improve basic representation to make clearer
  - **Interact**: add methods for manipulating the data

"80% perspiration, 10% great idea, 10% output" Simon Rogers, The Guardian

Framework for Developing visualisations





Good visualisations are driven by good questions

Identify suitable sources of data and information (e.g. primary, tertiary)

Information in its raw form needs to be abstracted and structured in order to tell a compelling story

Exploratory process to reveal answers to the questions identified, e.g. following design principles

Visible product of visualisation is effective representative, e.g. takes into account people's cognitive ability

Effective and compelling visualisations tell a story – storytelling is one of the primary ways we make sense of the world

Value of visualization & story telling

Value of visualization & story telling

- Statistics are not enough (Anscombe's quartet dataset)

Value of visualization & story telling

I		II		III		IV	
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

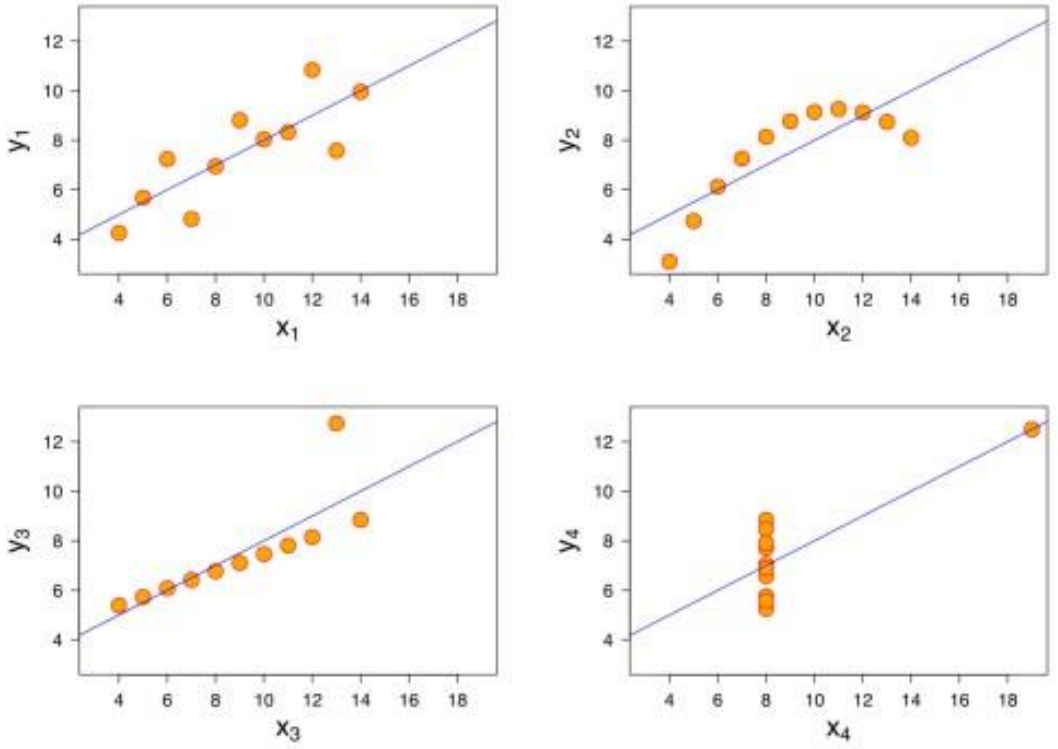
Anscombe's quartet

What can you see in these datasets? Any differences between I, II, III and IV?

Value of visualization & story telling

Anscombe's quartet

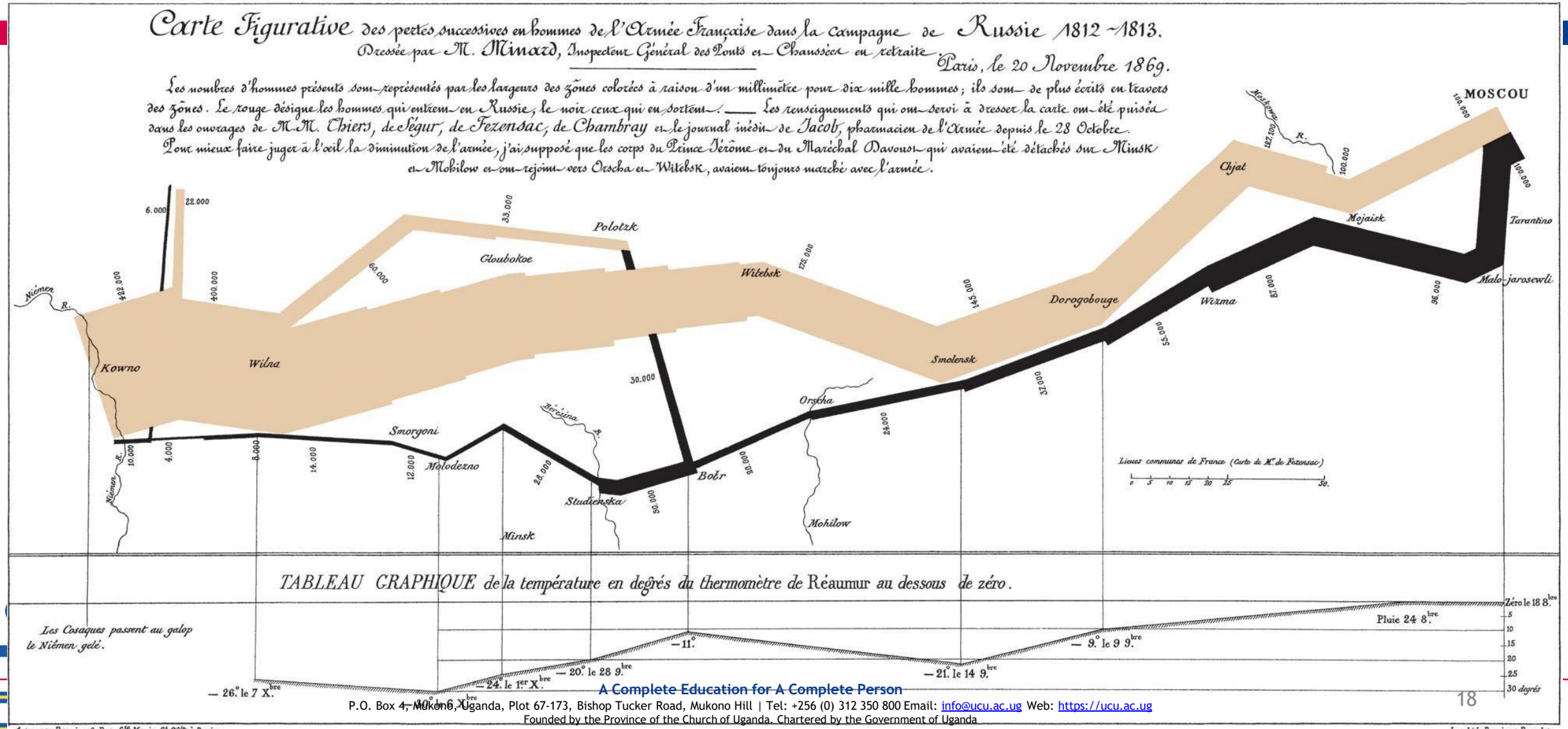
Property	Value
Mean of x in each case	9 (exact)
Sample variance of x i n each case	11 (exact)
Mean of y in each case	7.50 (to 2 decimal places)
Sample variance of y in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between x and y in each case	0.816 (to 3 decimal places)
Linear regression line in each case	$y = 3.00 + 0.500x$ (to 2 and 3 decimal places, respectively)



What can you see in these datasets? Any differences between I, II, III and IV?

*"A computer should make **both** calculations **and** graphs. Both sorts of output should be studied; each will contribute to understanding."* (Anscombe, 1973)

Minard's Napoleon March



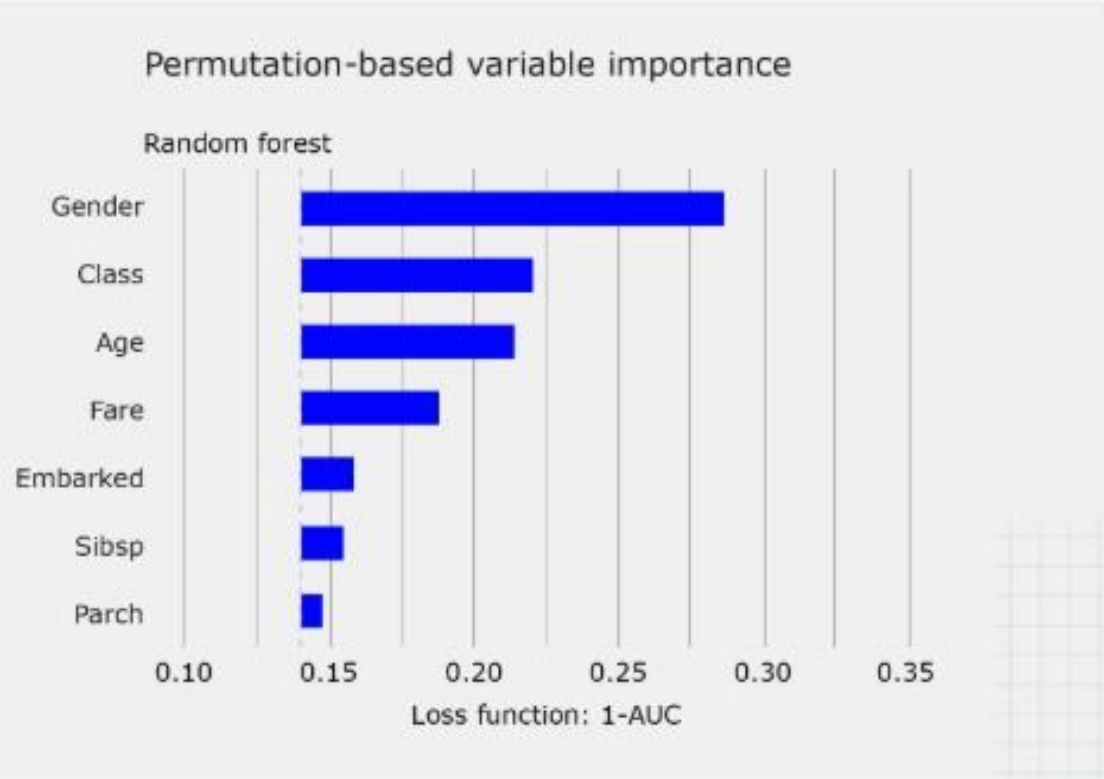
How communicate to the audience.

Types of data visualization

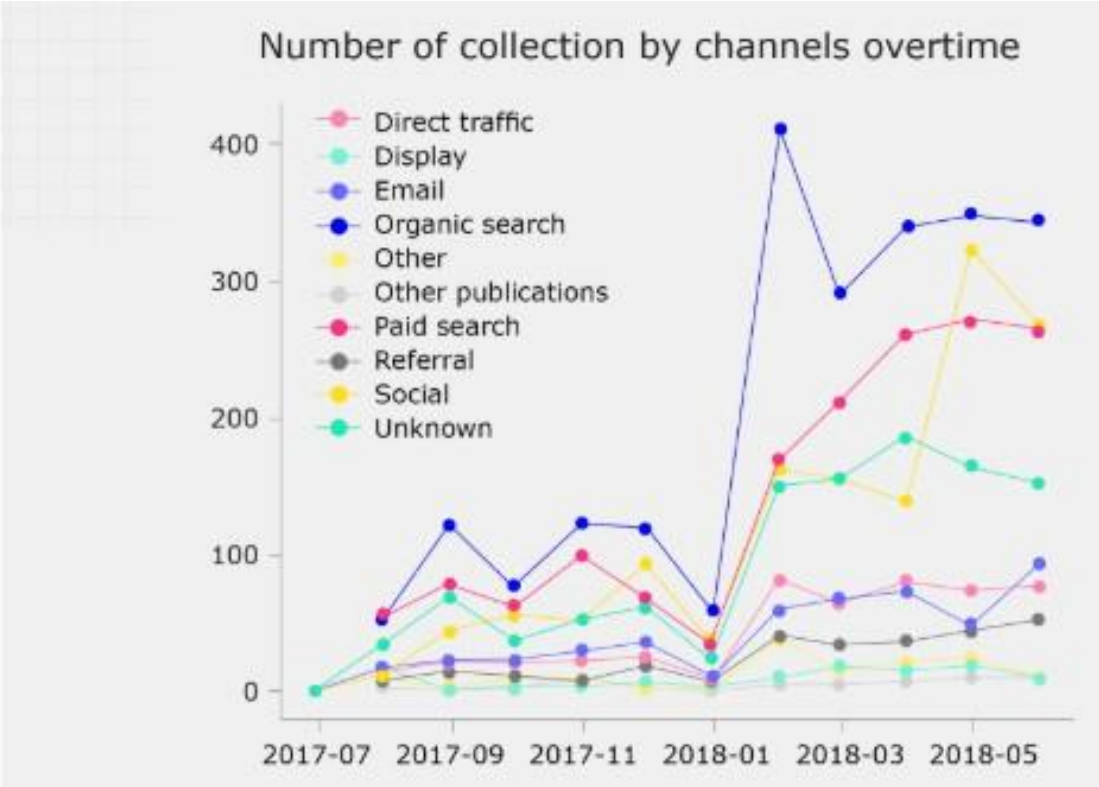
- Exploratory
- Explanatory

How communicate to the audience.

- Explanatory



- Exploratory



How to tell a story

- Why you should tell a story with your data visualization
- How you can tell a story with your data visualization

How to tell a story

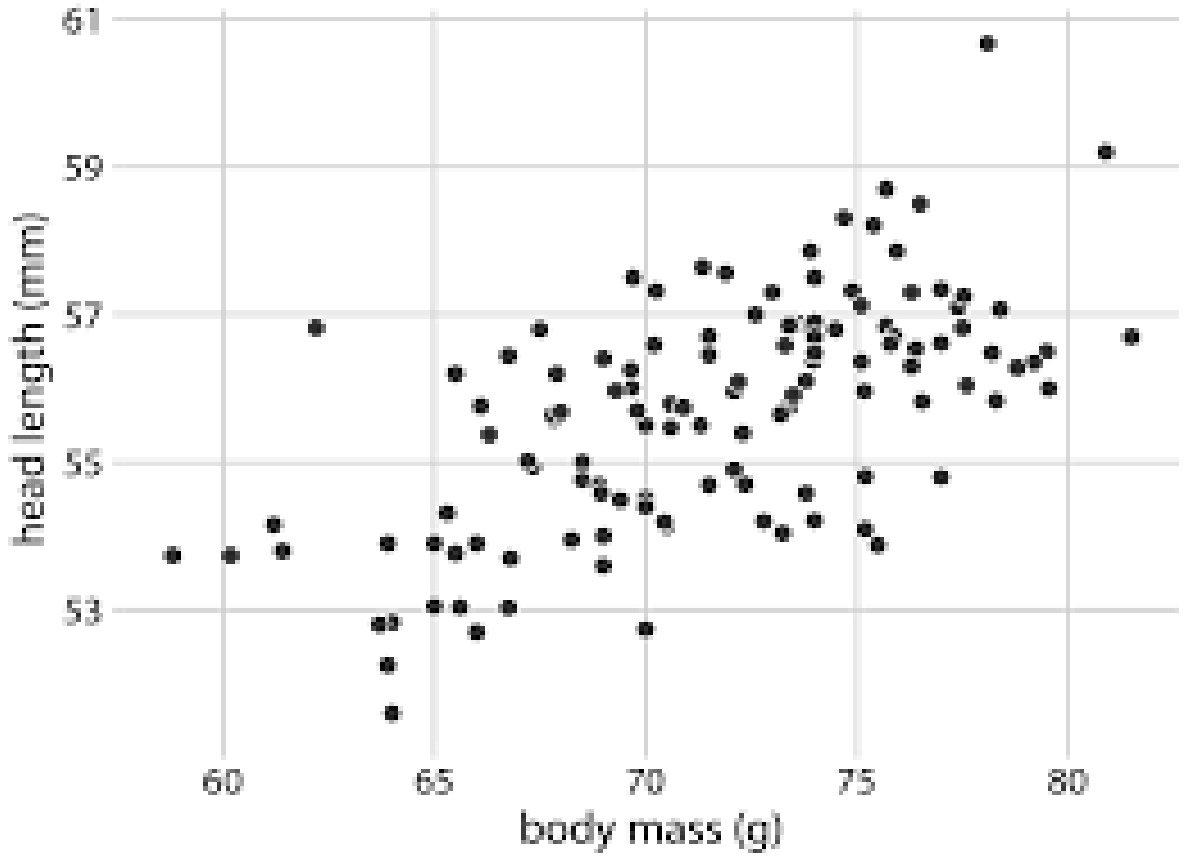
The best exploratory visualization is not necessarily the best explanatory visualization

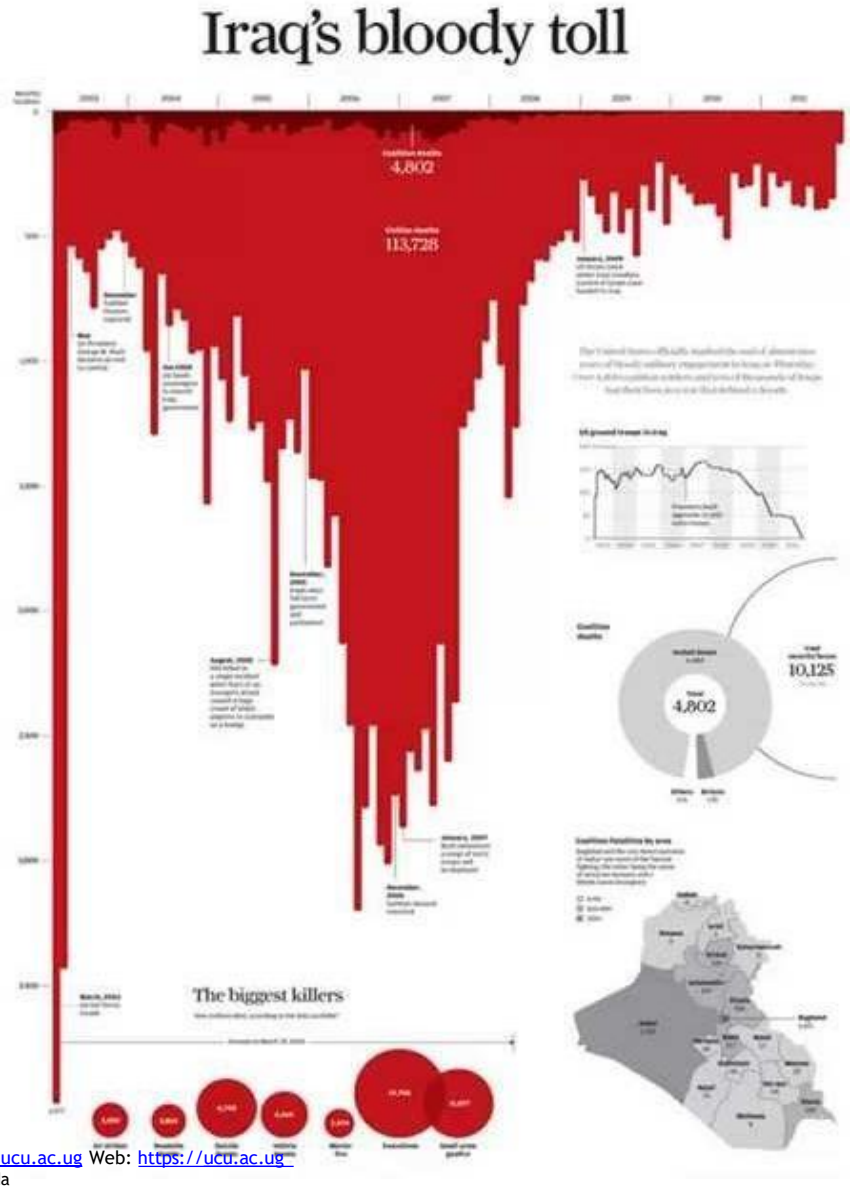
How to tell a story

Explanatory visualizations

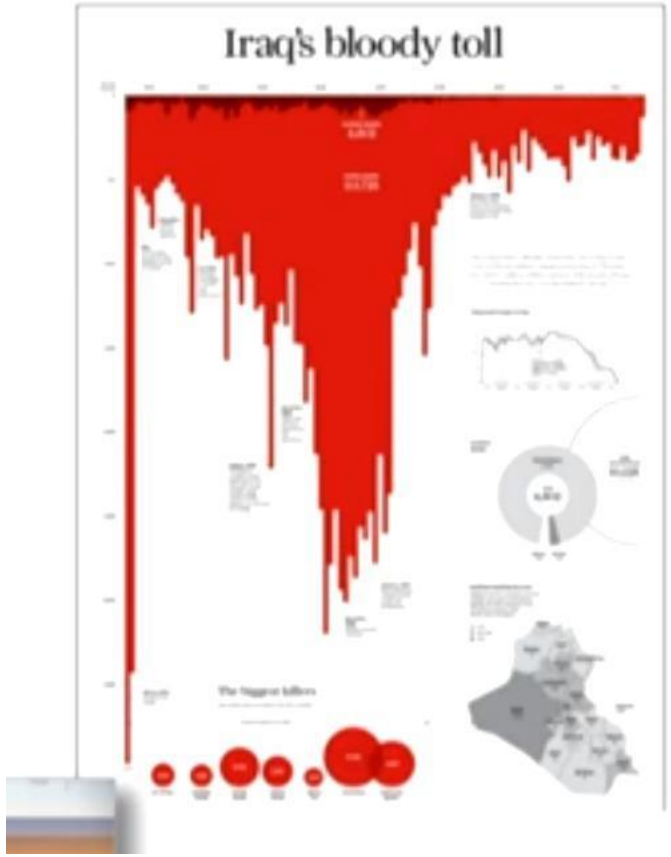
tell a story

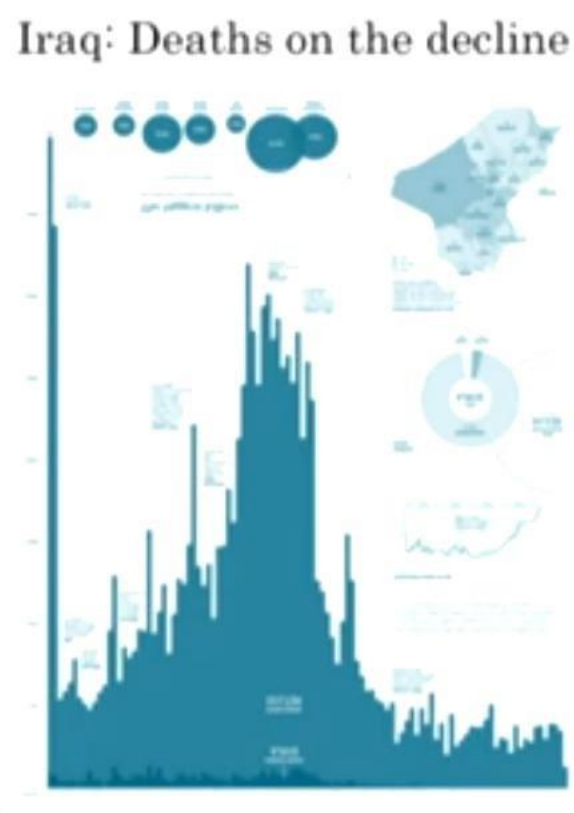
How to tell a story





How to tell a story





How to tell a story

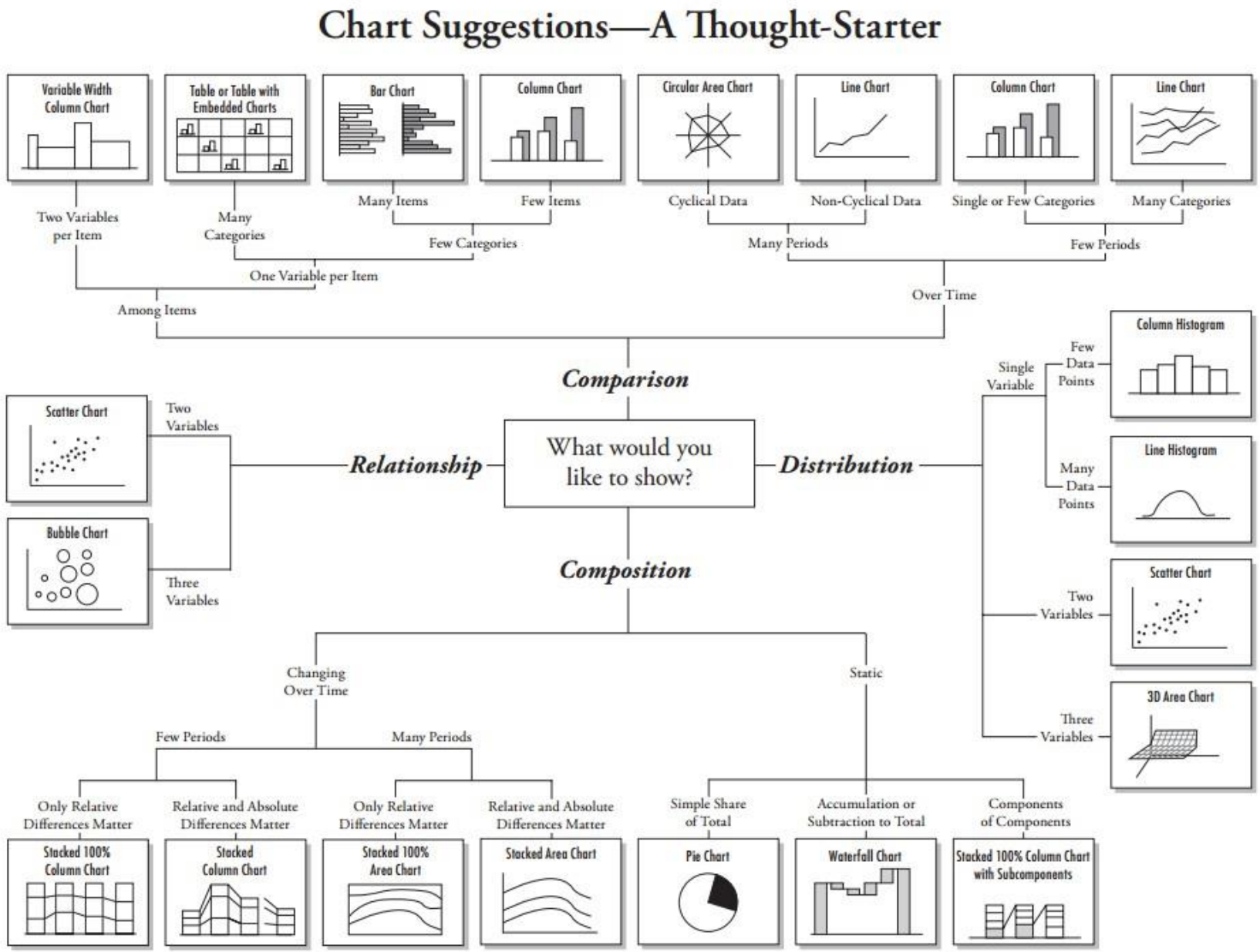
Key point of a story

- Foundation
- Focus
- Forward

How to tell a story# foundation

How to determine the right chart for your use case?

How to tell a story# foundation



http://www.extremepresentation.com/uploads/documents/choosing_a_good_chart.pdf

How to tell a story #focus

8 5 3 4 3 9 3 5 3 7 2 9 1 2

7 2 8 3 5 4 6 7 3 4 9 6 5 4

3 4 9 8 2 9 6 8 5 1 1 9 2 1

2 3 5 8 2 4 7 8 9 3 4 4 6 9

3 9 2 5 4 6 7 2 6 8 9 8 7 3

Count the number of 5s

How to tell a story #focus

Count the number of 5s #color

How to tell a story #focus

Find the 8s

9 9 9 9 9 9 9 9 8 9 9 9 9
9 9 9 9 8 9 9 9 9 9 9 8 9
9 9 8 9 9 9 9 8 9 9 9 9 9

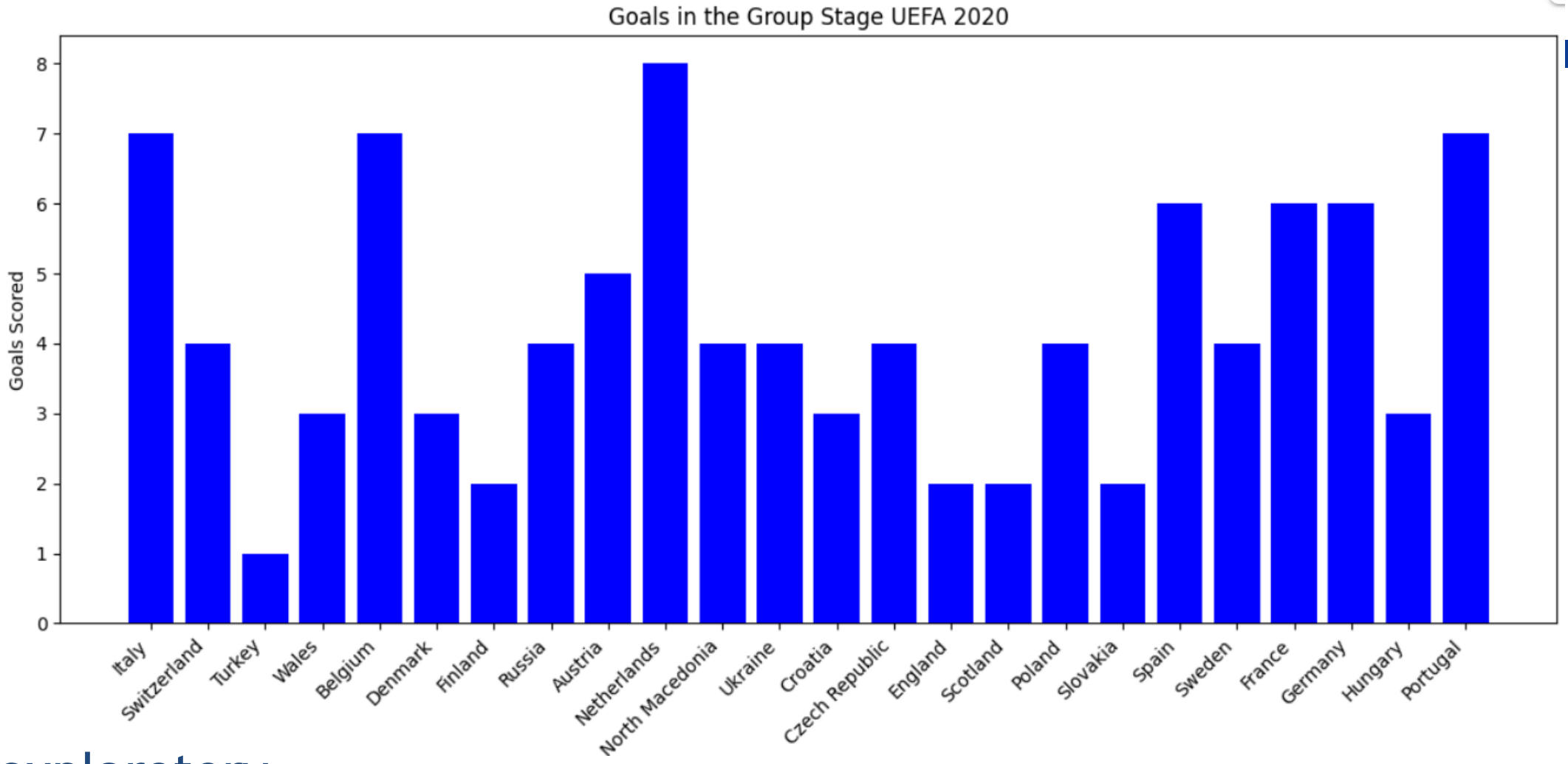
Find the 1s

9 9 9 9 9 9 9 9 1 9 9 9 9
9 9 9 9 1 9 9 9 9 9 9 1 9
9 9 1 9 9 9 9 1 9 9 9 9 9

Count the number of 8s & 1s #shape

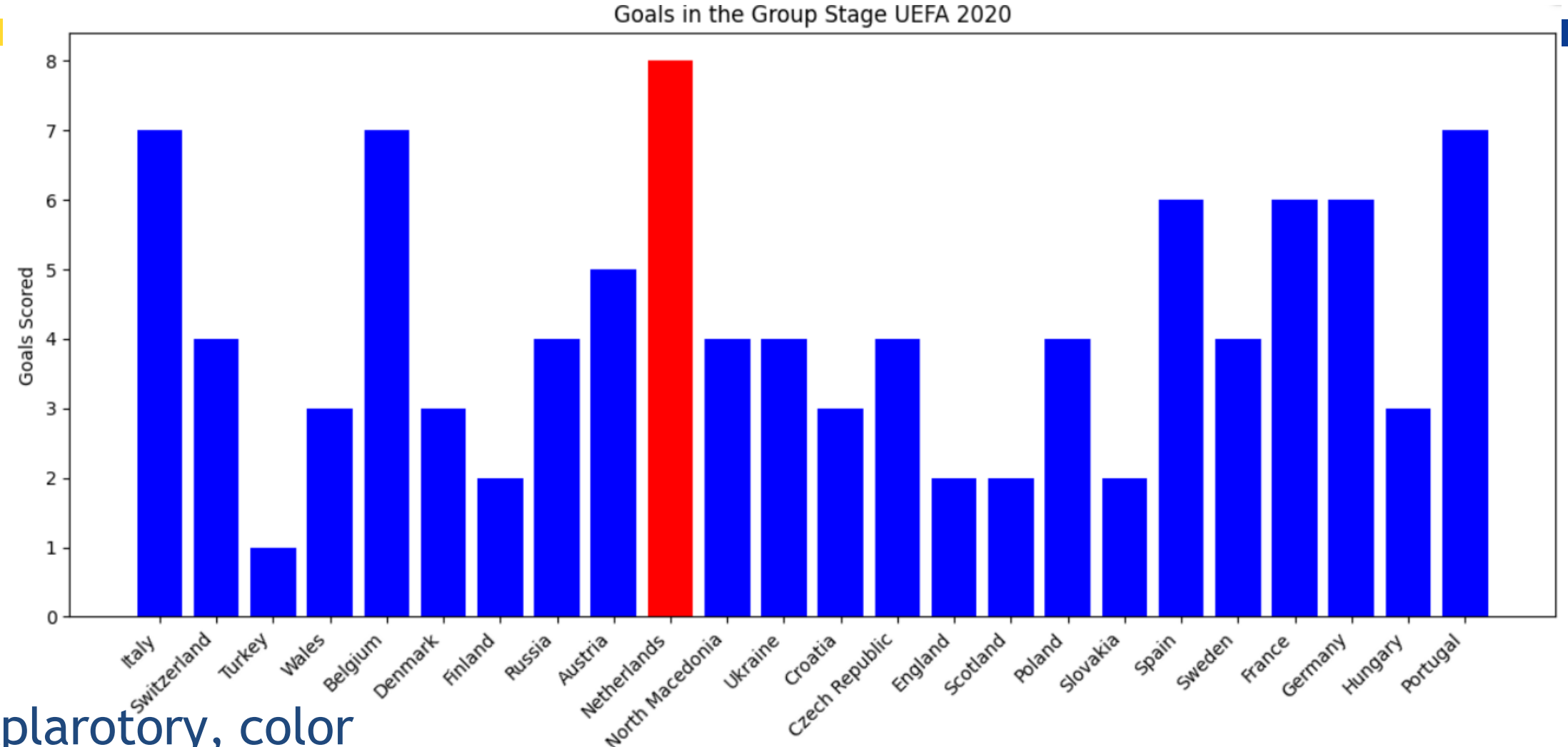
P.O. Box 4, Mukono, Uganda, Plot 67-173, Bishop Tucker Road, Mukono Hill | Tel: +256 (0) 312 350 800 Email: info@ucu.ac.ug Web: https://ucu.ac.ug
Founded by the Province of the Church of Uganda. Chartered by the Government of Uganda

How to tell a story #focus



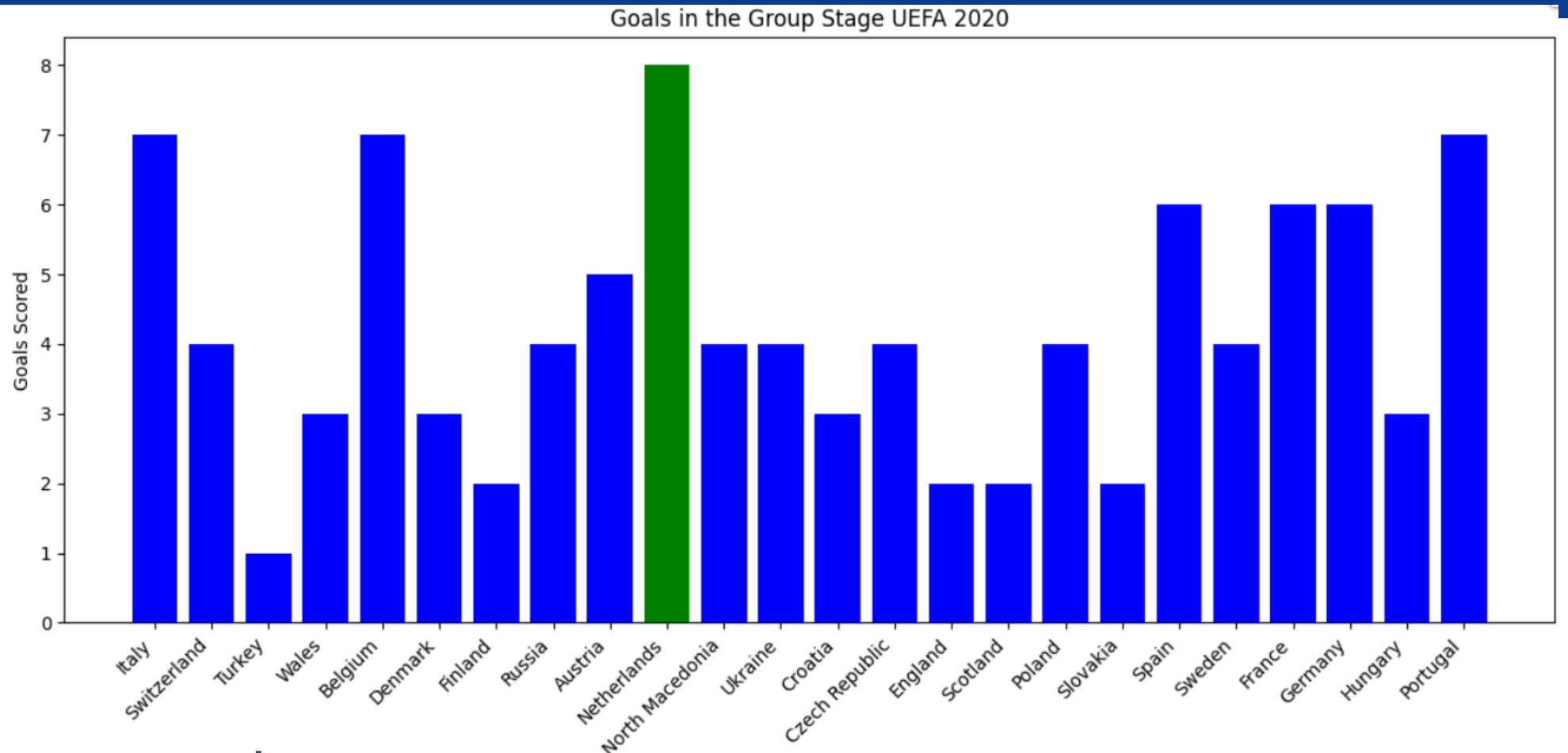
#explarotory

How to tell a story #focus



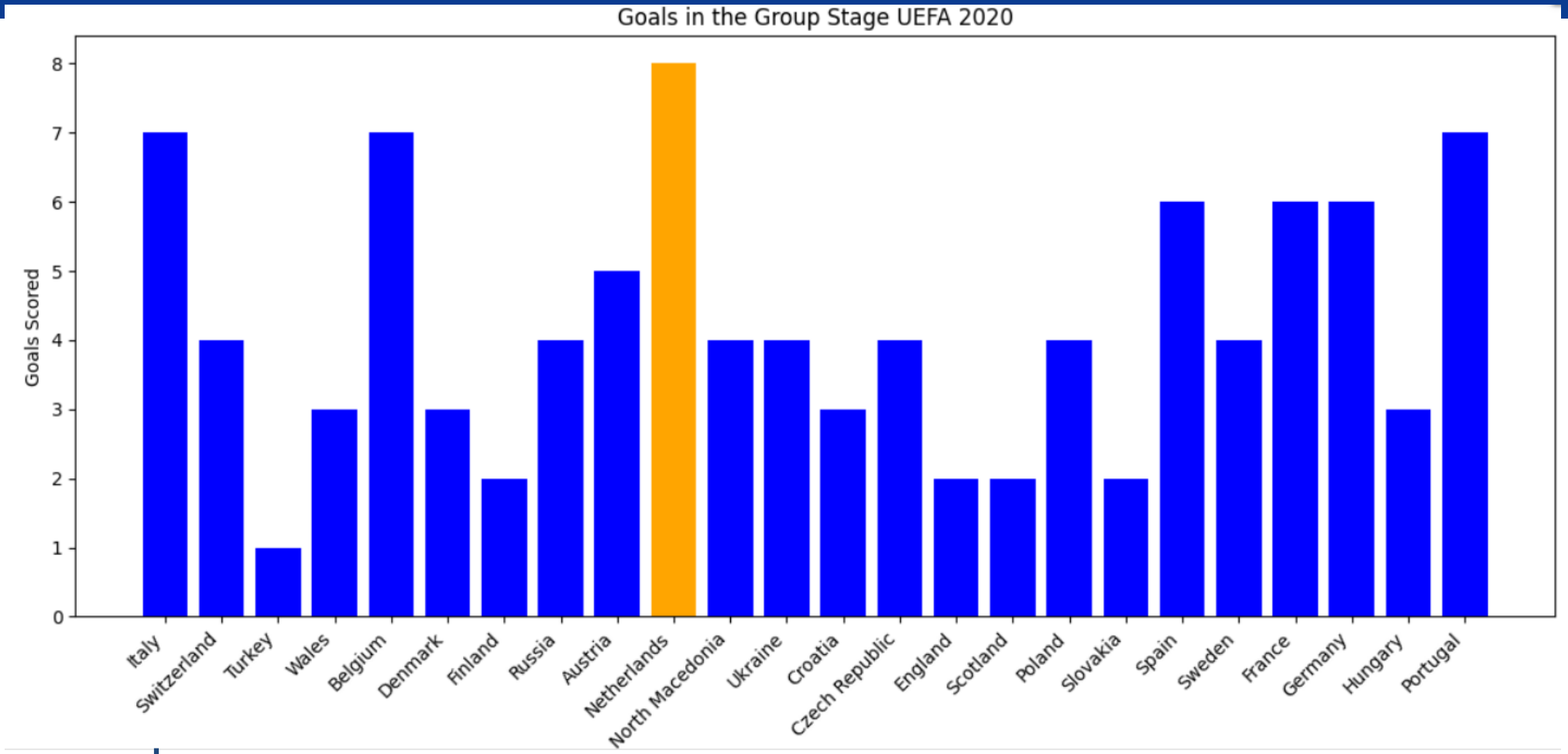
#explarotory, color

How to tell a story #focus



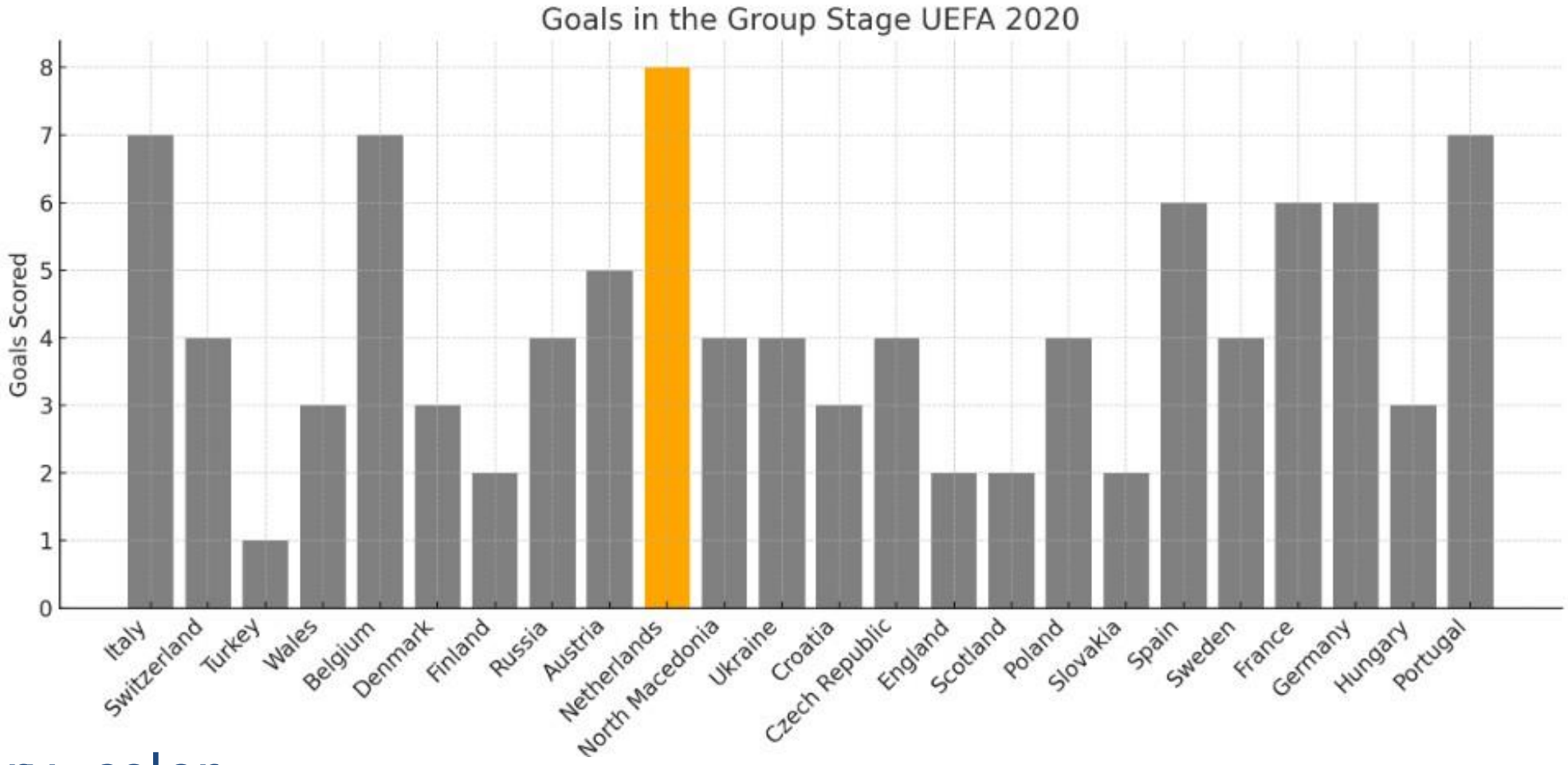
#explarotory, color

How to tell a story #focus



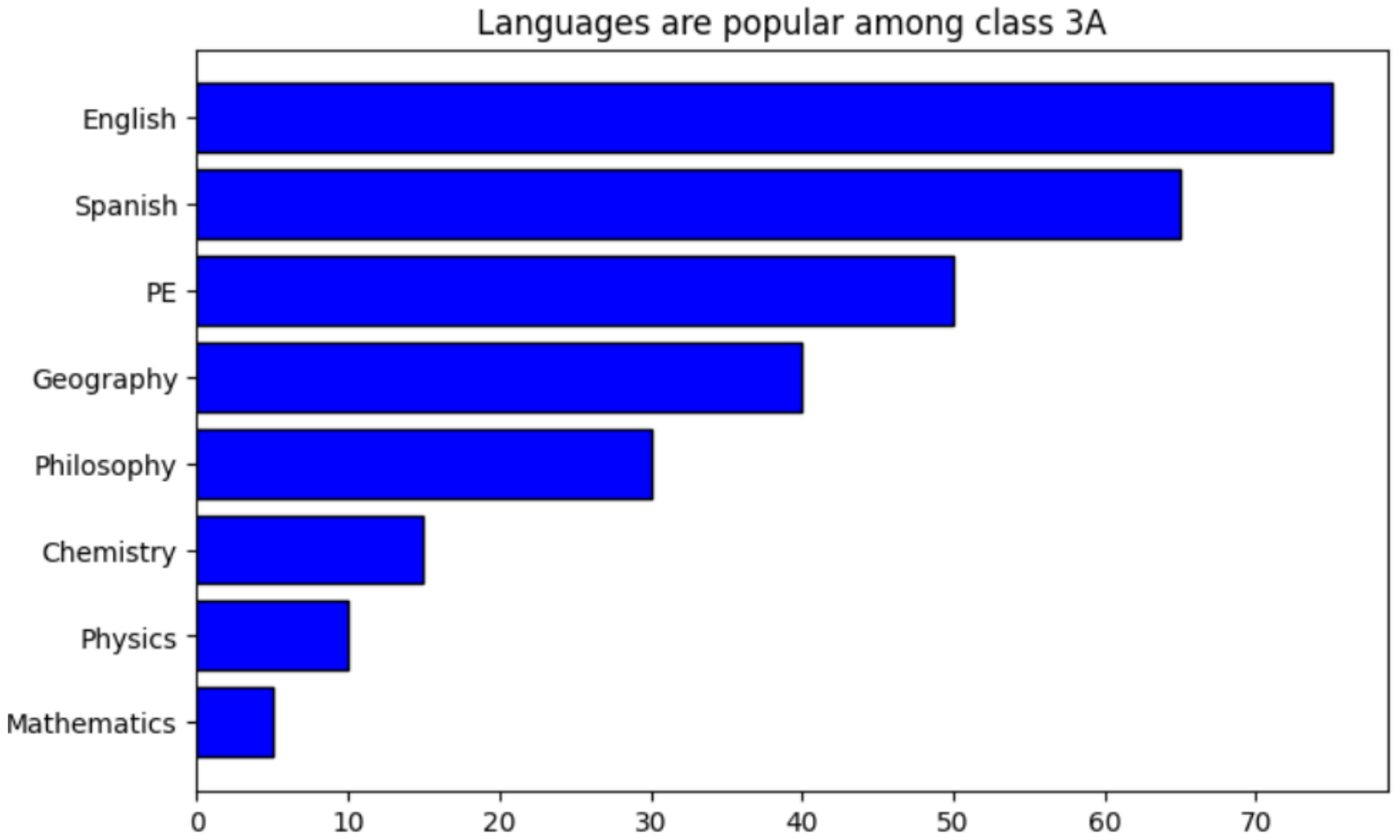
#explarotory, color

How to tell a story #focus



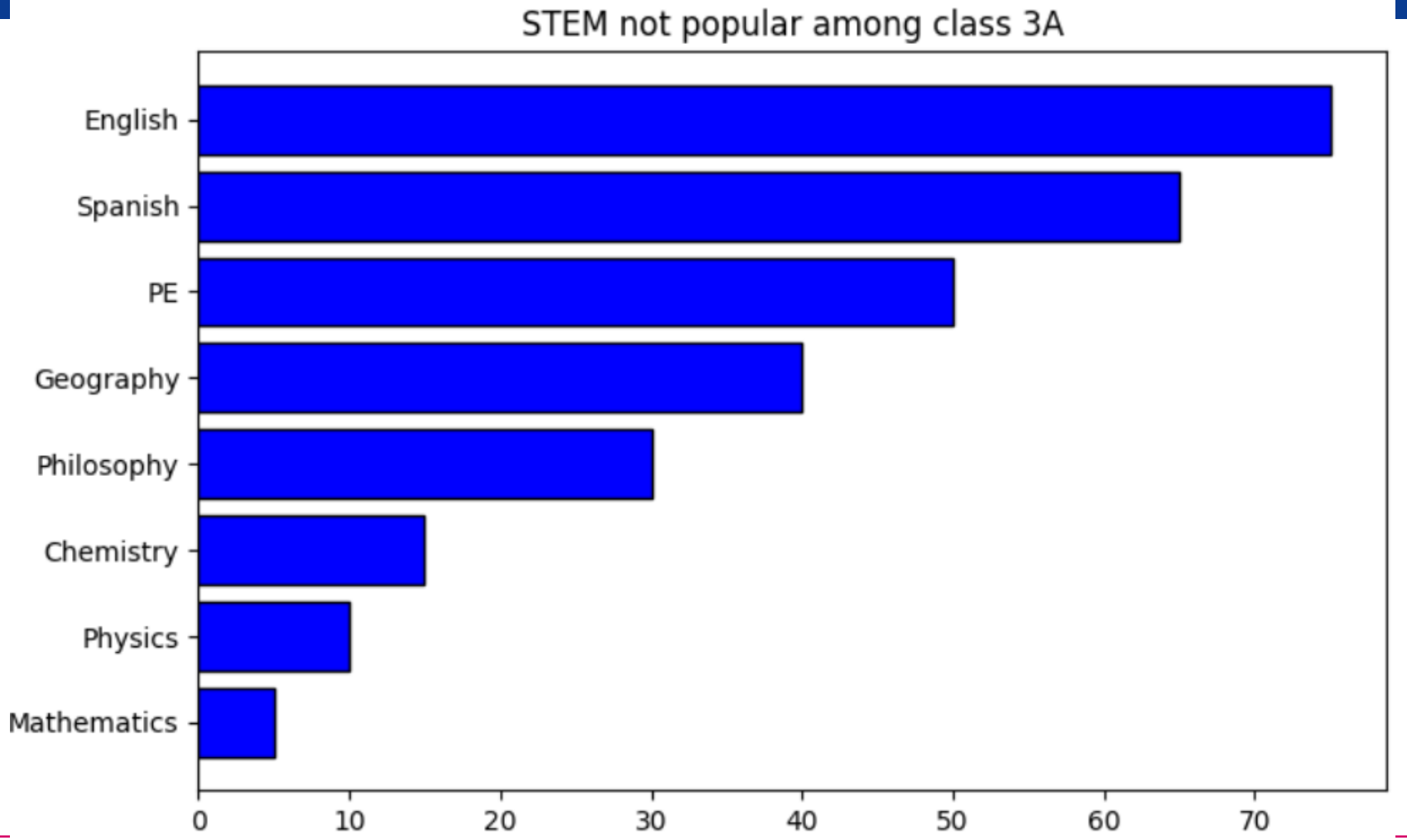
#explarotory, color

How to tell a story #focus



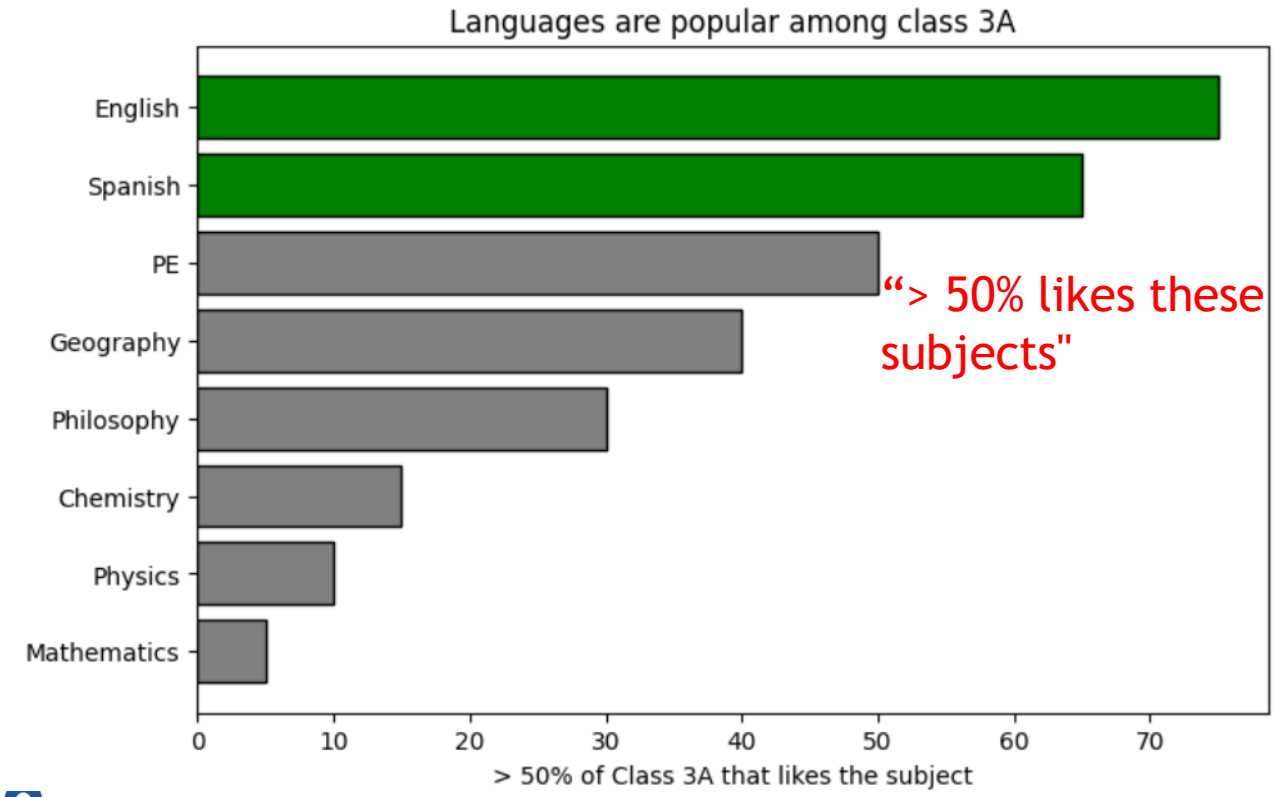
#explarotory, title

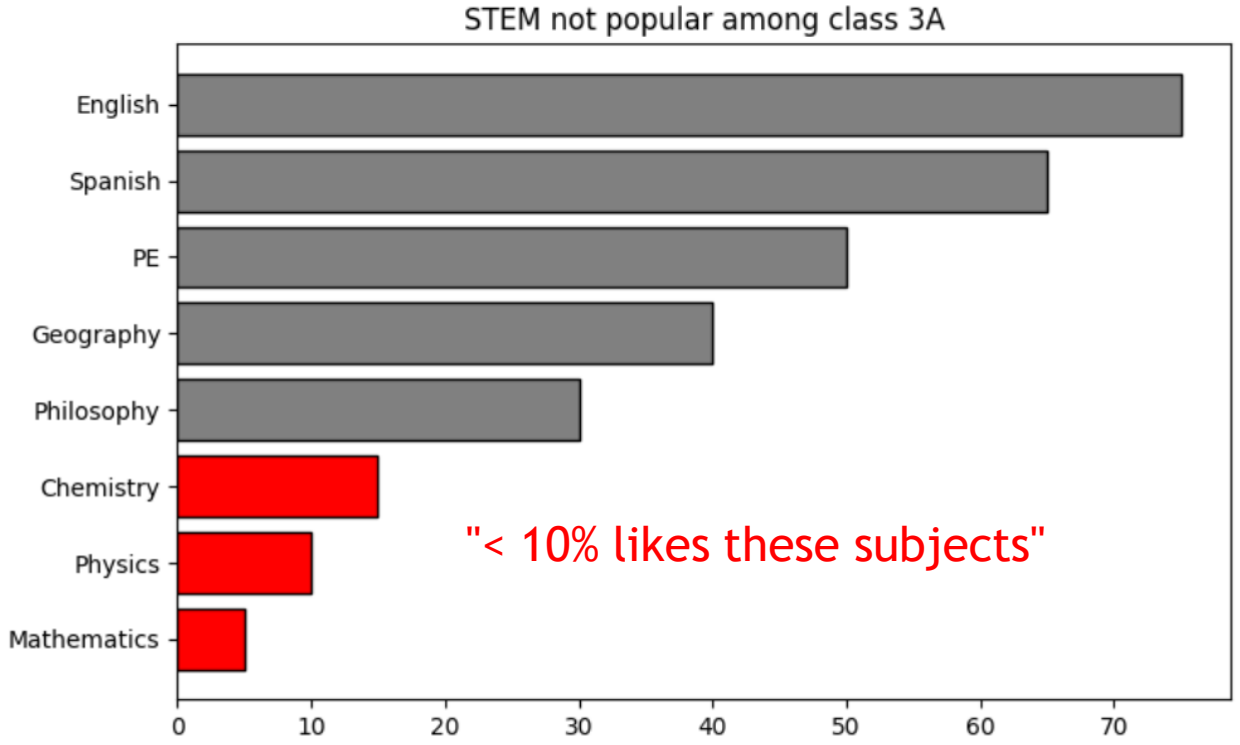
How to tell a story #focus



#explarotory, title

How to tell a story# focus



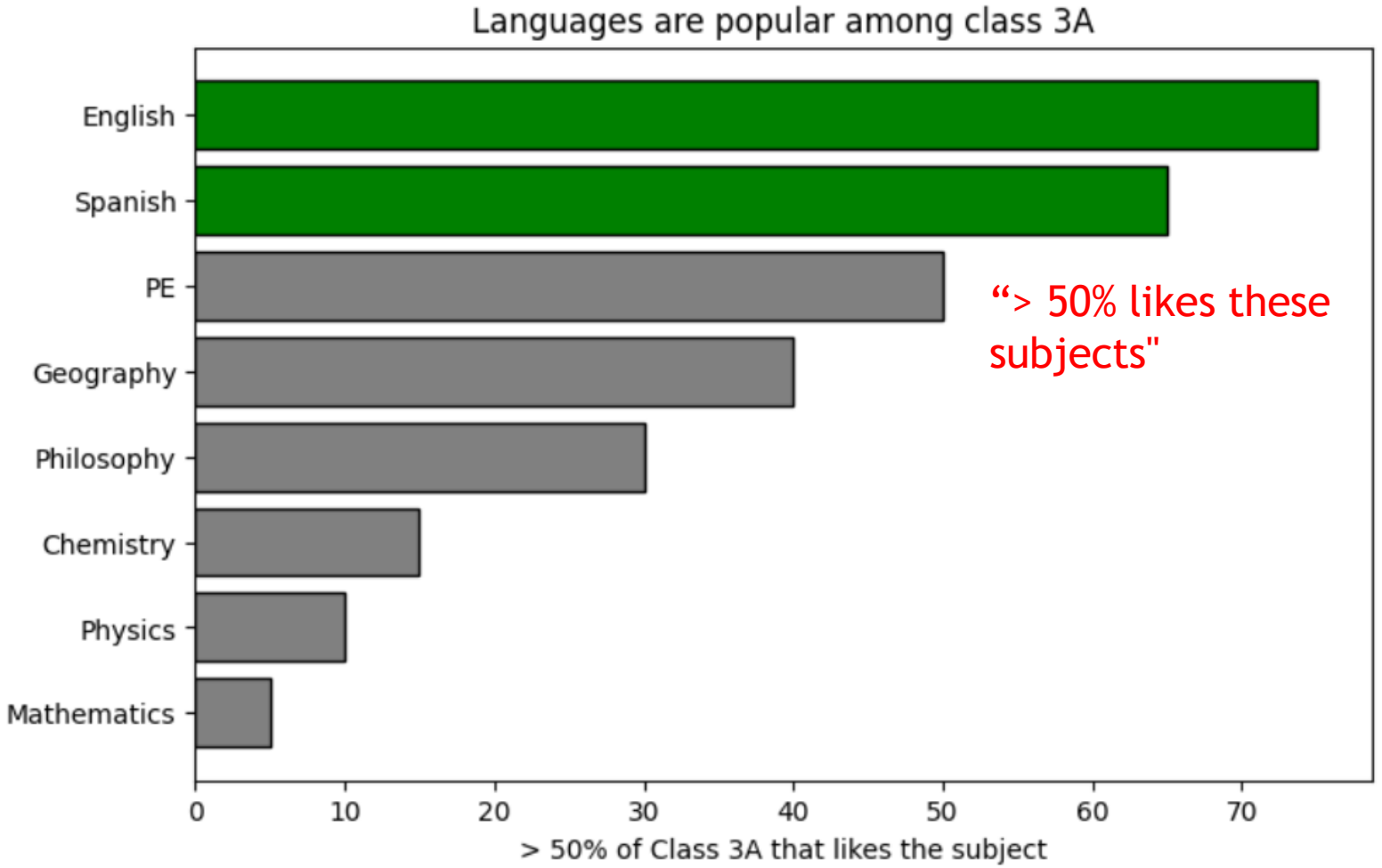


#explarotory, title, color

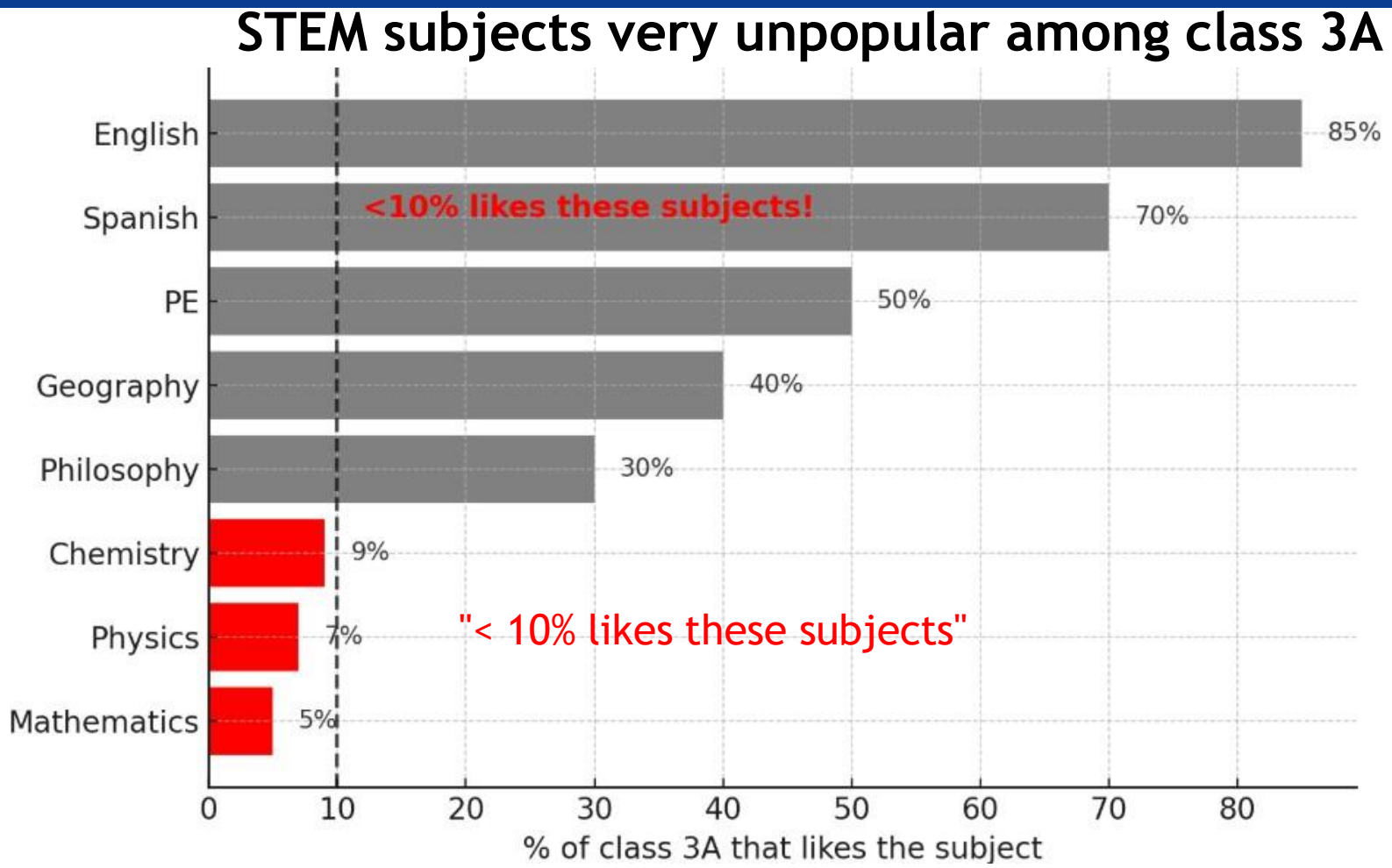
How to tell a story #Forward

What is the take away from your visualization?

How to tell a story #Forward #take away

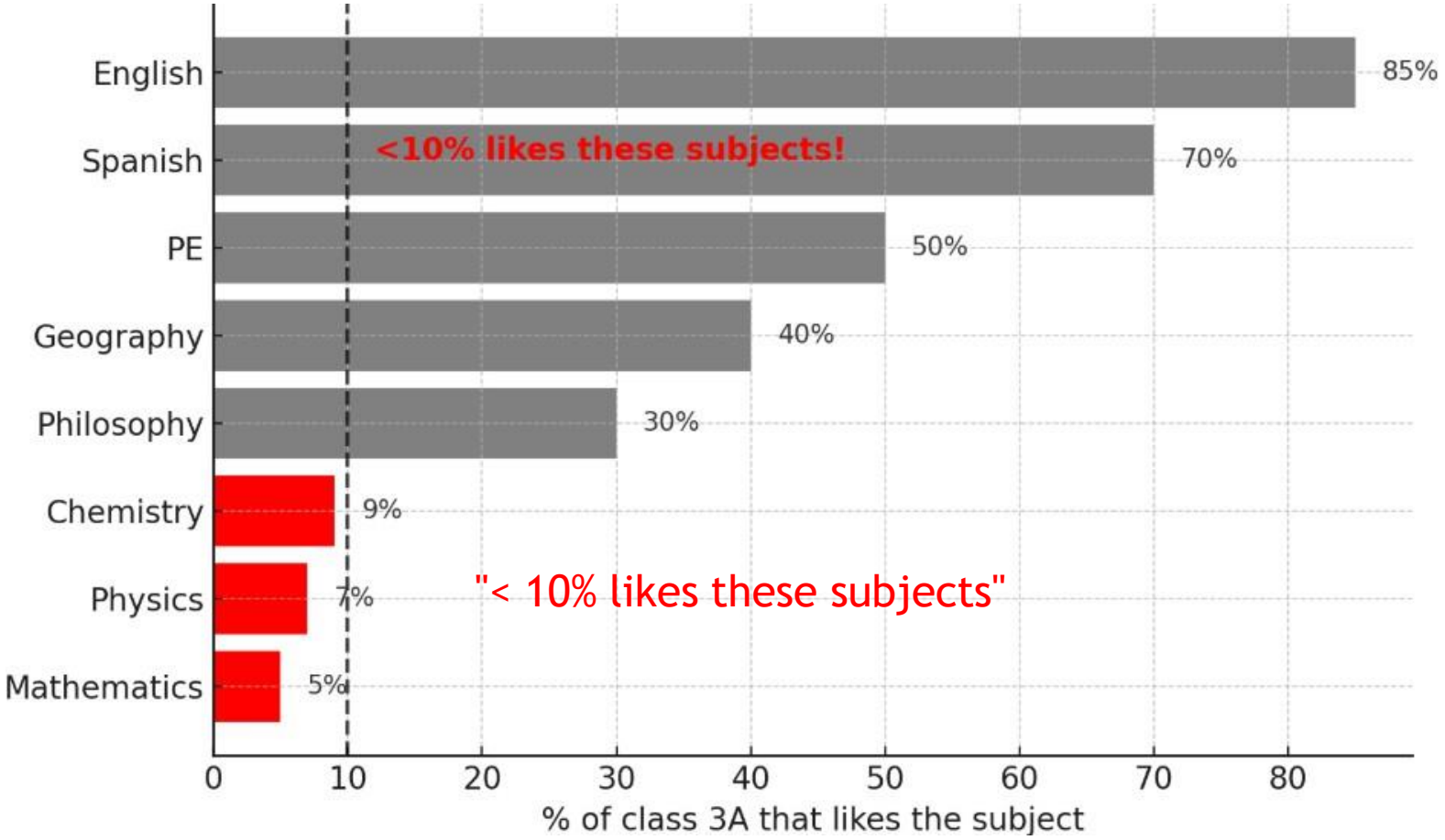


How to tell a story #Forward #take away



How to tell a story #Forward #take away

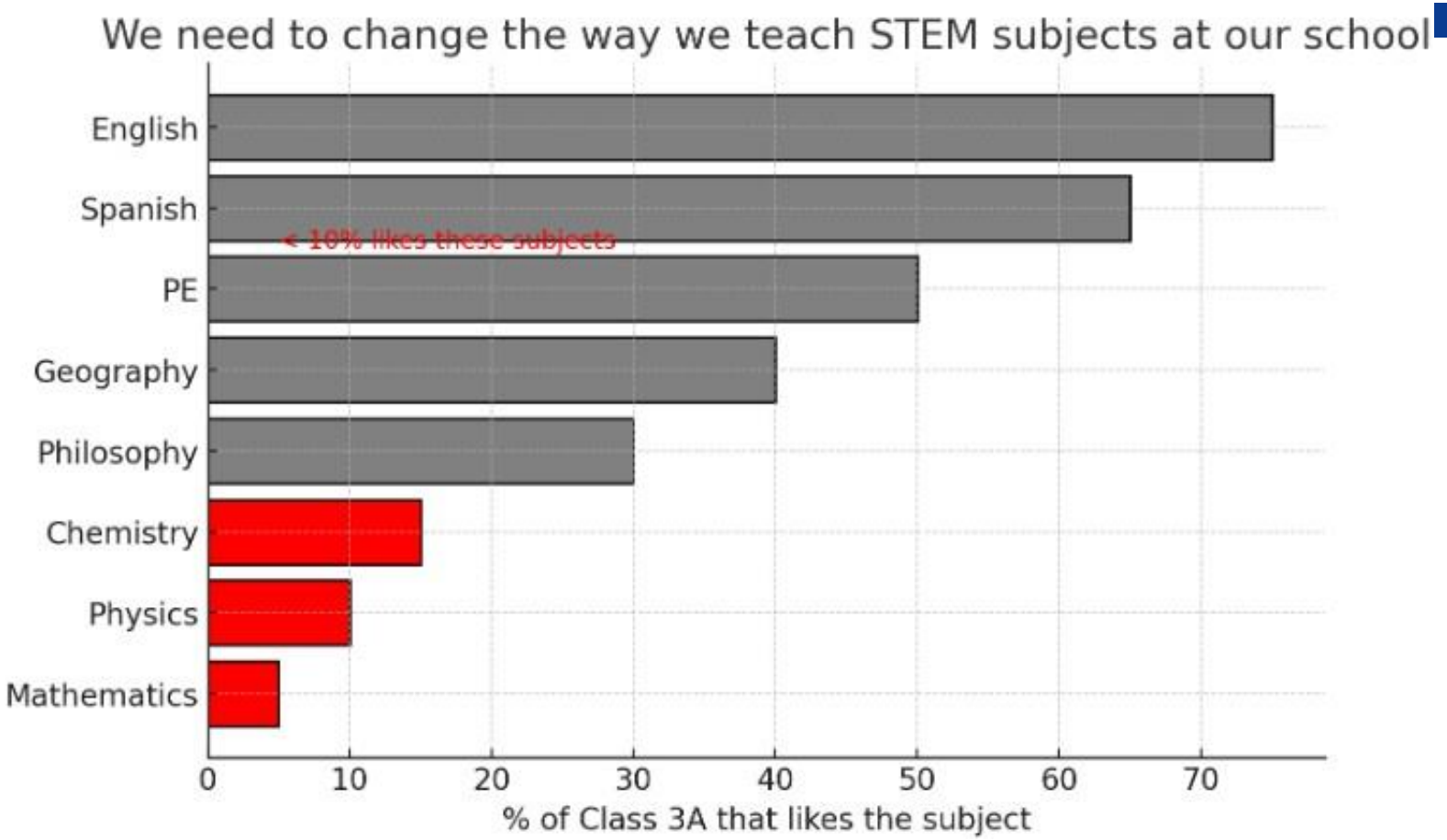
We need to change the way we teach STEM subjects at our school



How to tell a story #Forward #take away

Forward

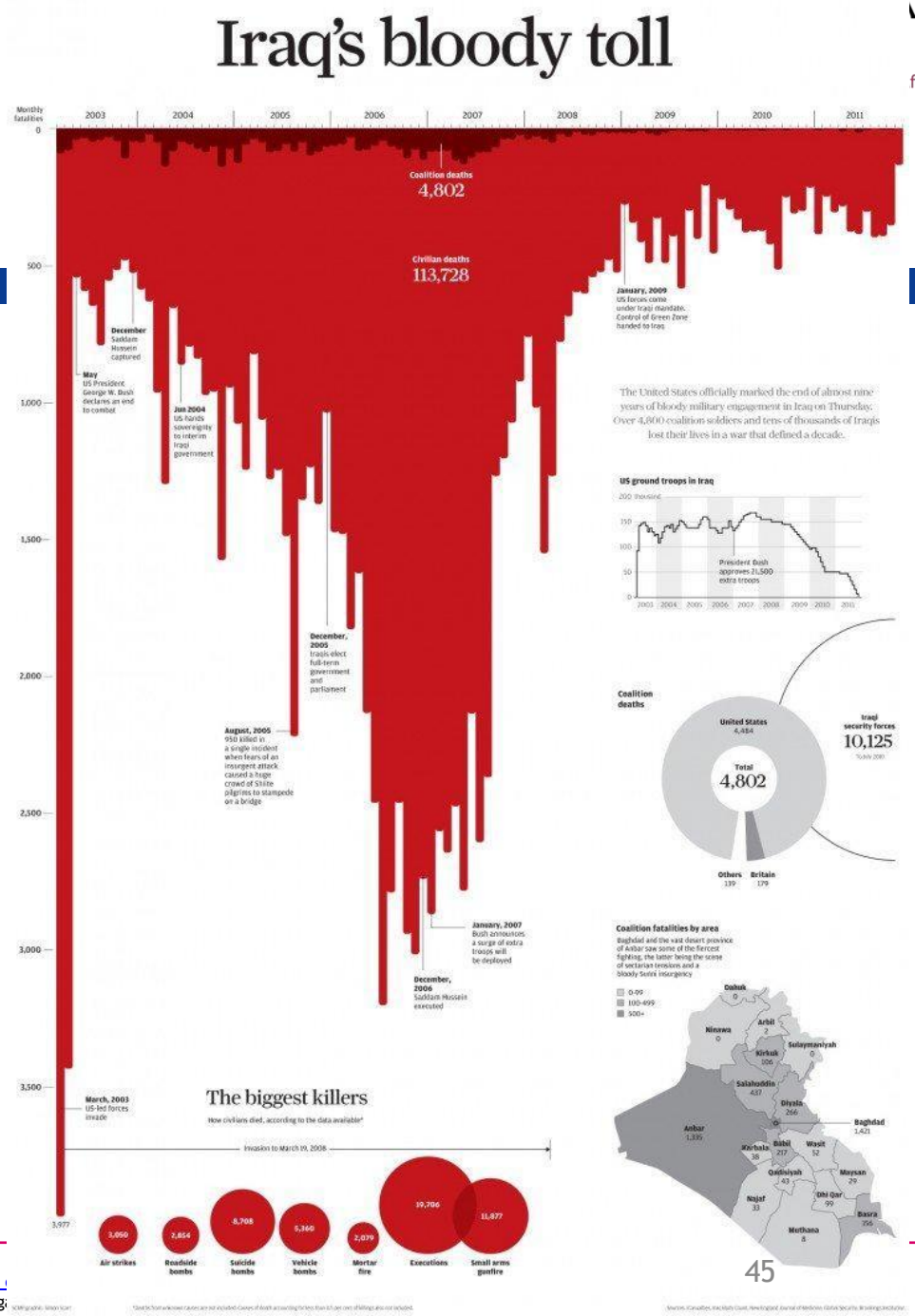
- Put the message first
- Support the conclusion
- Emphasize the point



"Our STEM subjects are heavily disliked. As a school, we must take action and investigate " "how we can make these subjects more fun and enjoyable for our students."

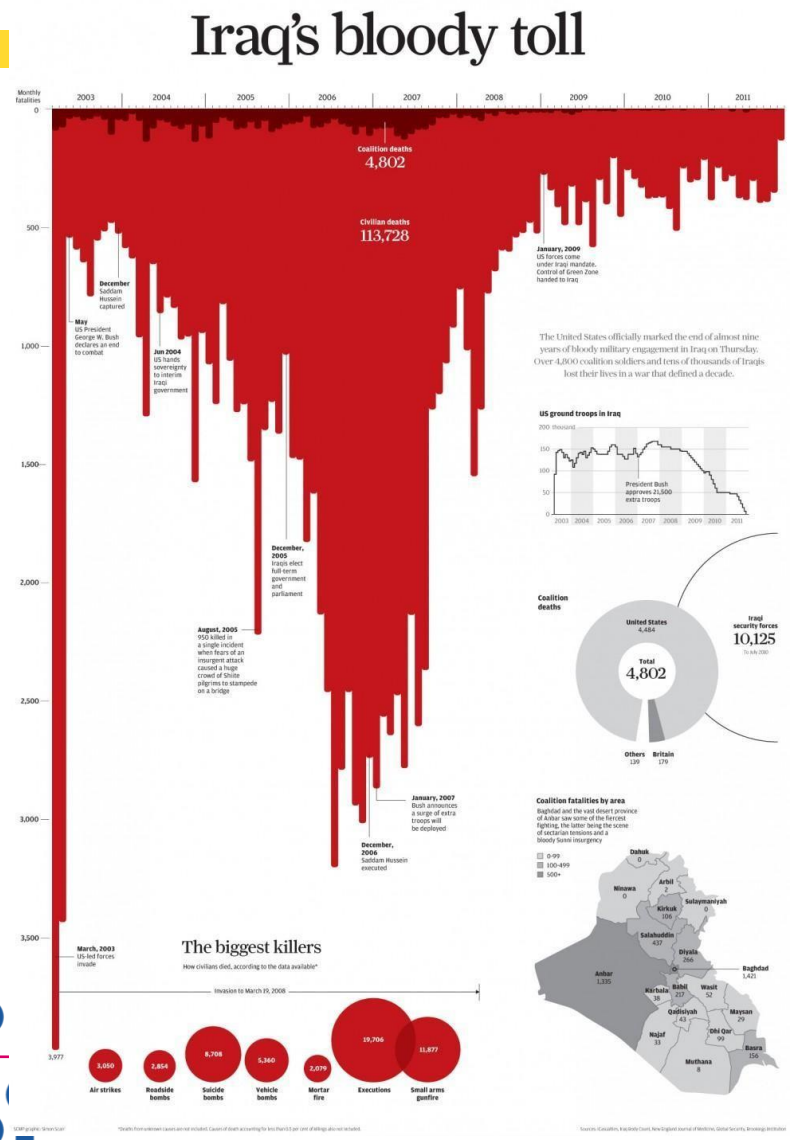
Visualisations for communication

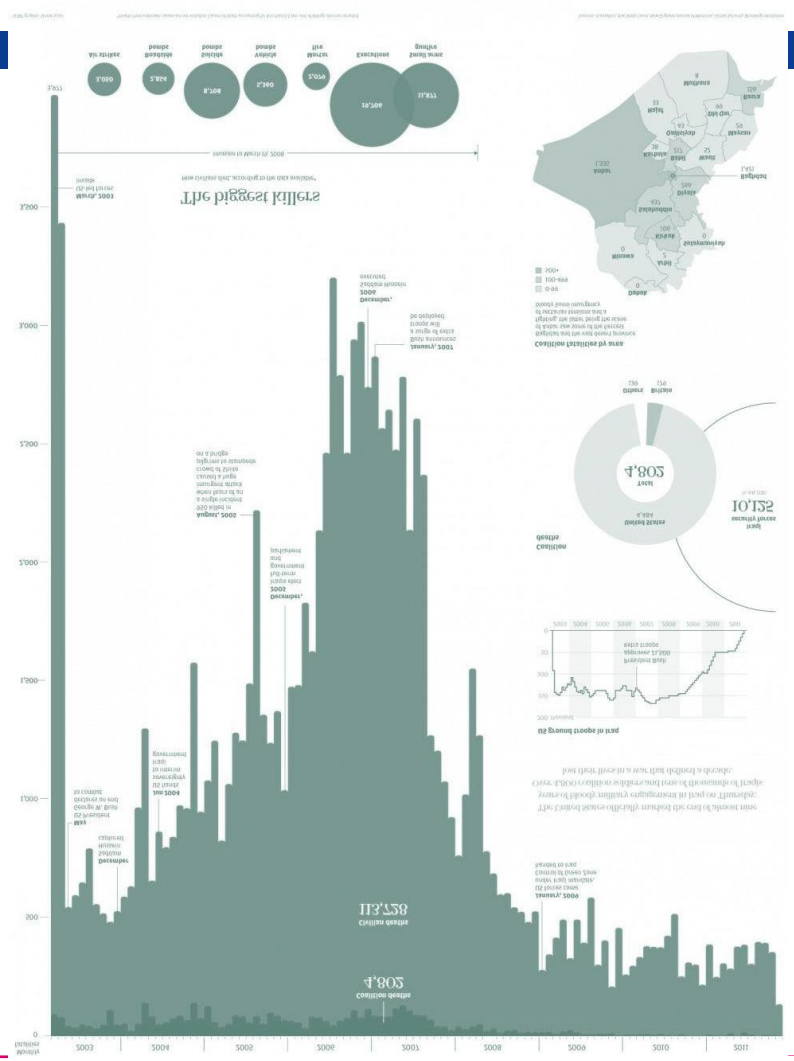
- The title + subject of discussion
- Use of the colour red vs grayscale
- High number of casualties
- Use of narratives along the timeline
- Associated graphs in the dashboard
- Map grounding the topic in reality



Iraq's Bloody Toll, by Simon Scarr, published in the South China Morning Post on December 17, 2011

Visualisations for communication





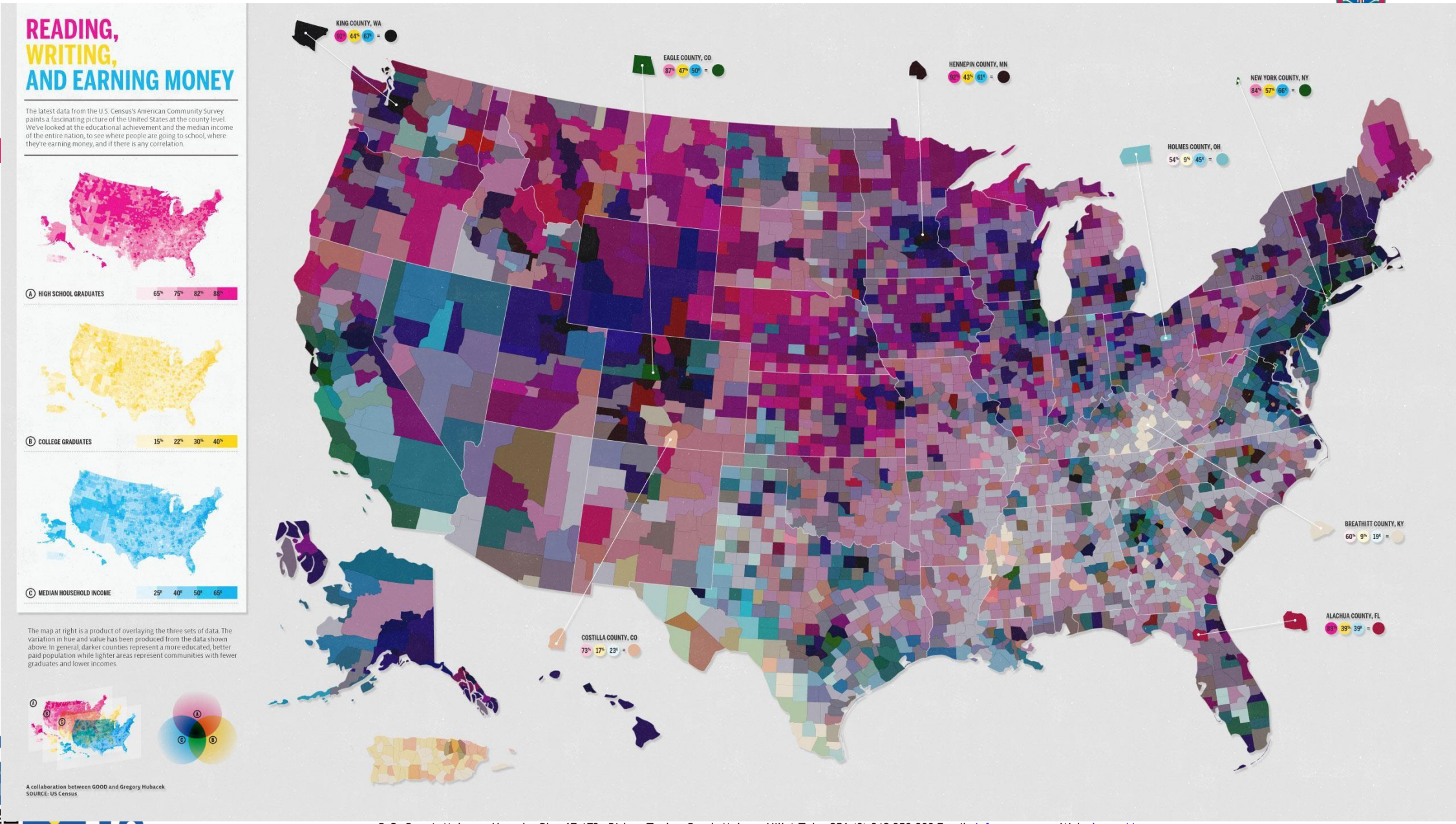
Iraq: Deaths on the decline

Activity: Analysing visualisations

Task: Study Visualisations

- The following slides contain examples from the real-world
- Look at the following two figures and analyse the visualisations
- Can you try to read the visualisations to understand what they are trying to convey?
- Can you think of any issues you can find with these visualisations?

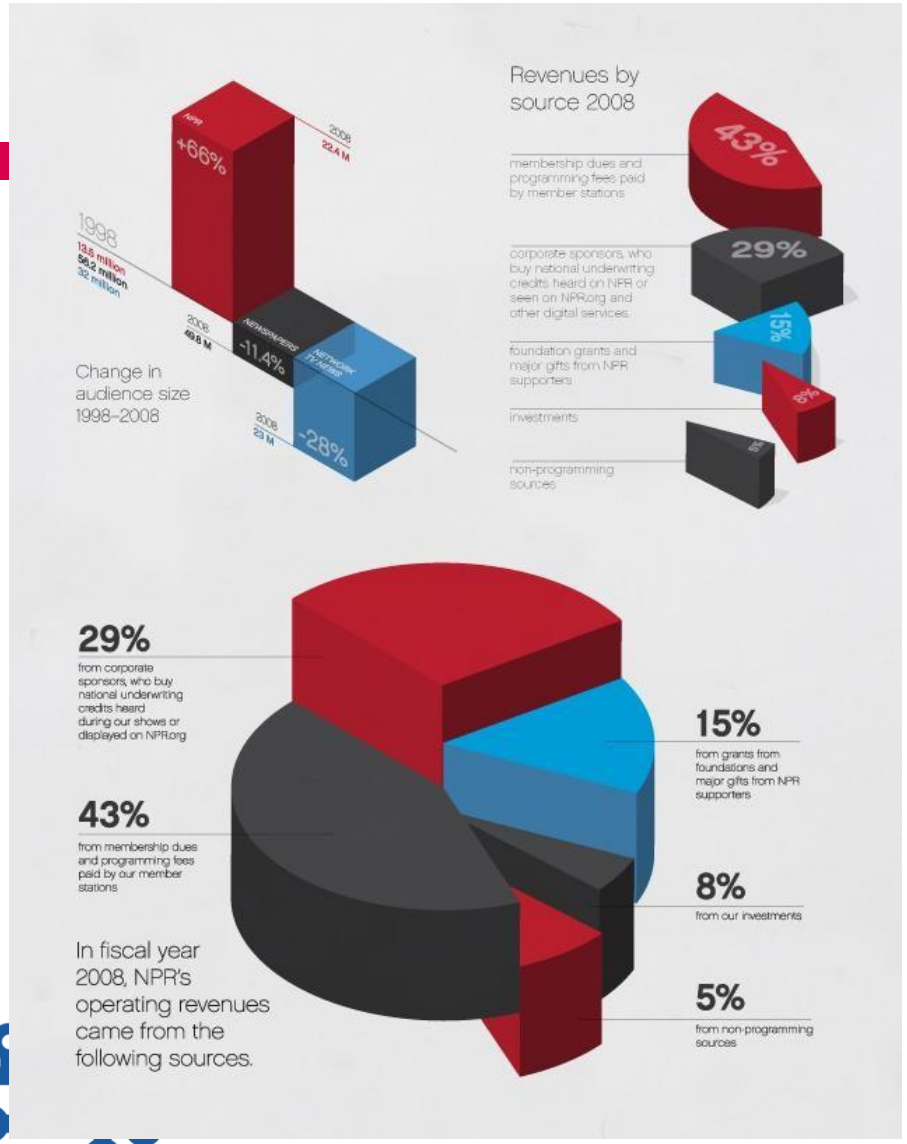
Visualisation 1 – can you identify any issues?

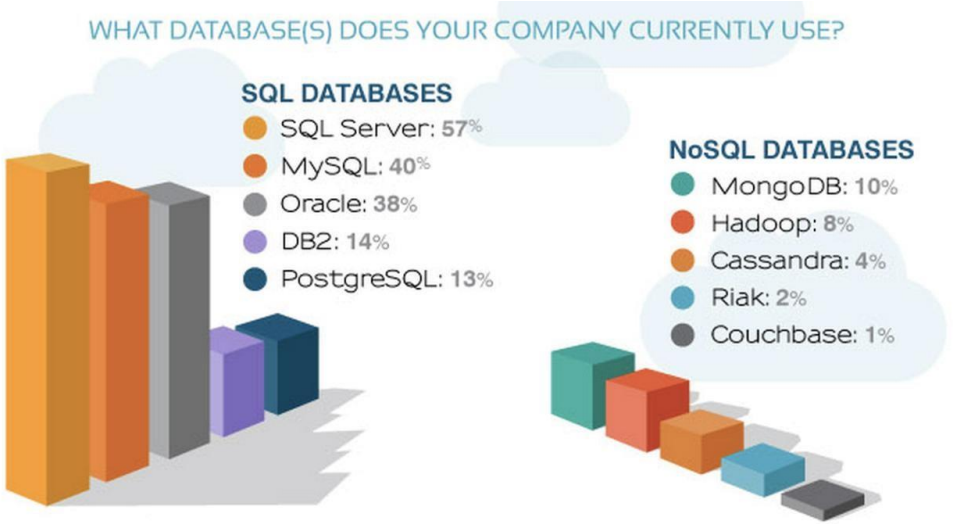


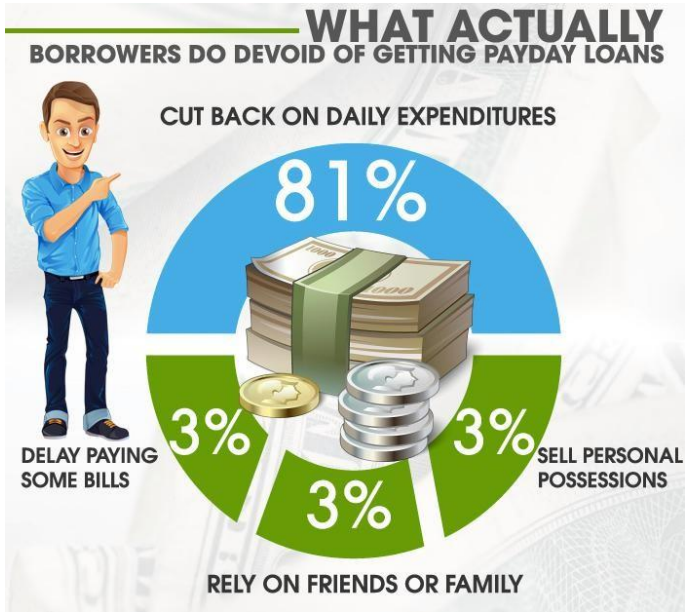
https://www.good.is/infographics/america-s-richest-counties-and-best-educated-countries

Wooclap: DSA24DV

Visualisation 2 – can you identify any issues?







Wooclap: DSA24DV

http://porostocky.prosite.com/65117/1
http://viz.wtf/post/107628689945/all-your-database-are-belong-to-us

Summary

- Visualisations help us understand data better
- Visualisation help us communicate events, and tell stories
- We can learn a lot from existing visualisations
  - Best practices, choices in visualisations
  - Design guidelines
  - But also, what to avoid, and how visualisations can be inaccessible
  - It’s important to be careful, and there’s a value in iterations

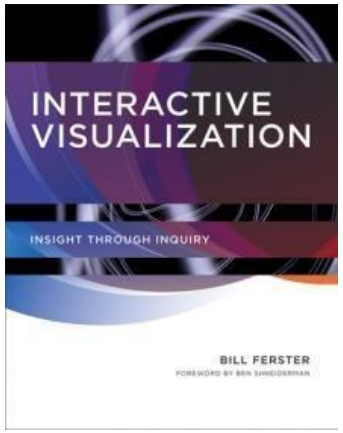
- Next Part – Creating Visualisations

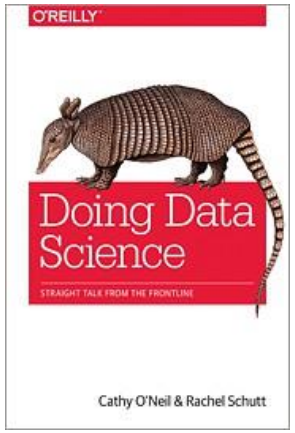
Further resources

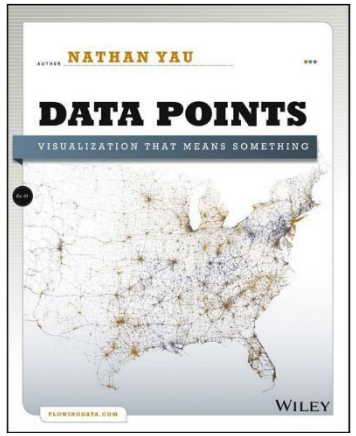
- Online visualisation galleries
  - http://guides.library.duke.edu/vis_types
  - https://github.com/mbostock/d3/wiki/Gallery
  - http://www.informationisbeautiful.net/
  - http://www.datavis.ca/gallery/

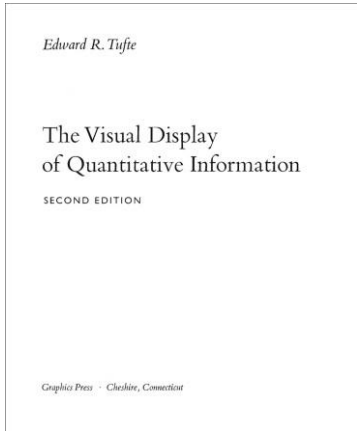
- Online resources
  - https://www.interaction-design.org/encyclopedia/data_visualization_for_human_perception.html
  - http://selection.datavisualization.ch/
  - http://guides.library.duke.edu/content.php?pid=355157&sid=2904817
  - http://www.billiondollargraphics.com/GraphicsCheatSheet_GMG.pdf
  - https://www.evl.uic.edu/aej/424/week04.html

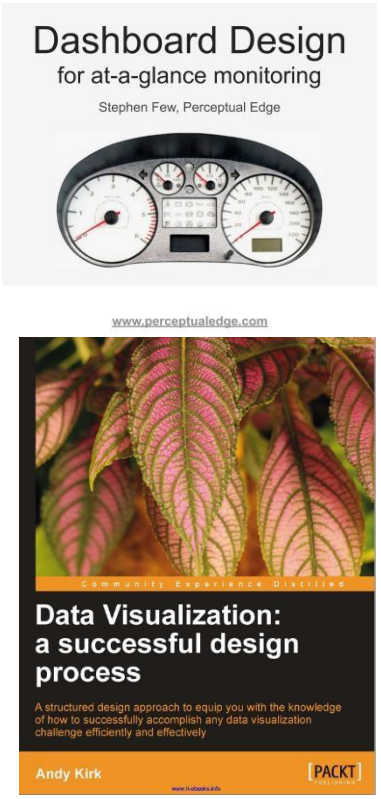
Reference books











Data visualization tools

Overview of popular data visualization tools

- Excel
- Tableau
- Power Bi

Introduction to Tableau

- Tableau is a web-based software tool used to transform data and create visualizations. Many data analysts use Tableau to create informative visualizations and web-accessible data dashboards.
- In this module, you will learn how to get started creating visualizations in Tableau Public, a free version of the Tableau software.
- Objective: Create visualizations using Tableau.

Introduction to Tableau

Lab

- This lab introduces Tableau Public, a free version of the Tableau product. The lab requires the registration and download of the free Tableau application.
  - Part 1: Download and install the Tableau Public Desktop Application
  - Part 2: Navigating the Tableau interface
  - Part 3: First Look at the data

Introduction to Tableau

Lab Instructions

Part 1: Download and install the Tableau Public Desktop

Step 1: Download Tableau.

1. Navigate to the Tableau website using the link https://www.tableau.com/products/public/download. If desired, review the system requirements.
2. Select **DOWNLOAD TABLEAU PUBLIC** from the button in the center of the screen.
    1. Fill in the registration information, including First Name, Last Name, business or school email, and select the Country/Region from the drop down.
    2. Review the Privacy Statement and then select Download the App again. The download should begin immediately; **if not, follow the instructions to try again for Mac or Windows.**

Introduction to Tableau

- Background/Scenario
- Visualizations are an important component of data analysis. Through visualization, comparisons and correlations can be shown in a manner that is clear and understandable. One of the most popular tools used by data analysts is Tableau. This lab uses the Tableau Public product to perform some preliminary analysis using a prepared data set containing information about movies distributed during the last decade.
- Required Resources throughout the labs include:.
  - PC with internet access
  - Tableau Desktop App
  - Dataset

Introduction to Tableau

- Step 2: Install Tableau.

1. Navigate to the download location and install the Tableau Public Desktop file. Refer to the
- generic instructions below for your operating system, if necessary:
    1. **Windows**: Run the installer and follow the prompts.
    2. **Mac**: Open the Disk image file (.DMG) and double-click the installer package (.PKG) to
    - start the installation.

2. On the Welcome screen, ensure that the product is the Tableau Desktop Public Edition, and read and accept the licensing agreement to continue the installation. Accept the default settings and click Install.

Introduction to Tableau

Part 2: Navigating the Tableau Interface

- Step 1: Explore the Tableau interface.
- In this step, you will explore the Tableau interface, where the start page gives you several options.
- Start Tableau if needed after completing the installation. You will see the start page, which has three panes: Connect, Open, and Discover.
  - Note: you can click the Tableau icon in the upper left corner to toggle between the start page and the authoring workspace page.
  - Under Connect on the start page, you can connect to your data in various formats.
  - On the right side of the start page, there is a Discover area that contains links to videos and sample data sets as well as training information. You can investigate these links to access additional information about Tableau.

Uganda Christian University
P.O. Box 4 Mukono, Uganda
Tel: 256-312-350800
https://ucu.ac.ug/ Email: info@ucu.ac.ug.
@ugandachristianuniversity @UCUniversity
@UgandaChristianUniversity

Department of Computing & Technology
FACULTY OF ENGINEERING, DESIGN AND TECHNOLOGY
Tel: +256 (0) 312 350 863 | WhatsApp: +256 (0) 708 114 300
@ucucomputeng @ucu_ComputEng
https://cse.ucu.ac.ug/ Email: dct-info@ucu.ac.ug