



# Exploratory Data Analysis: *from insights to storytelling*

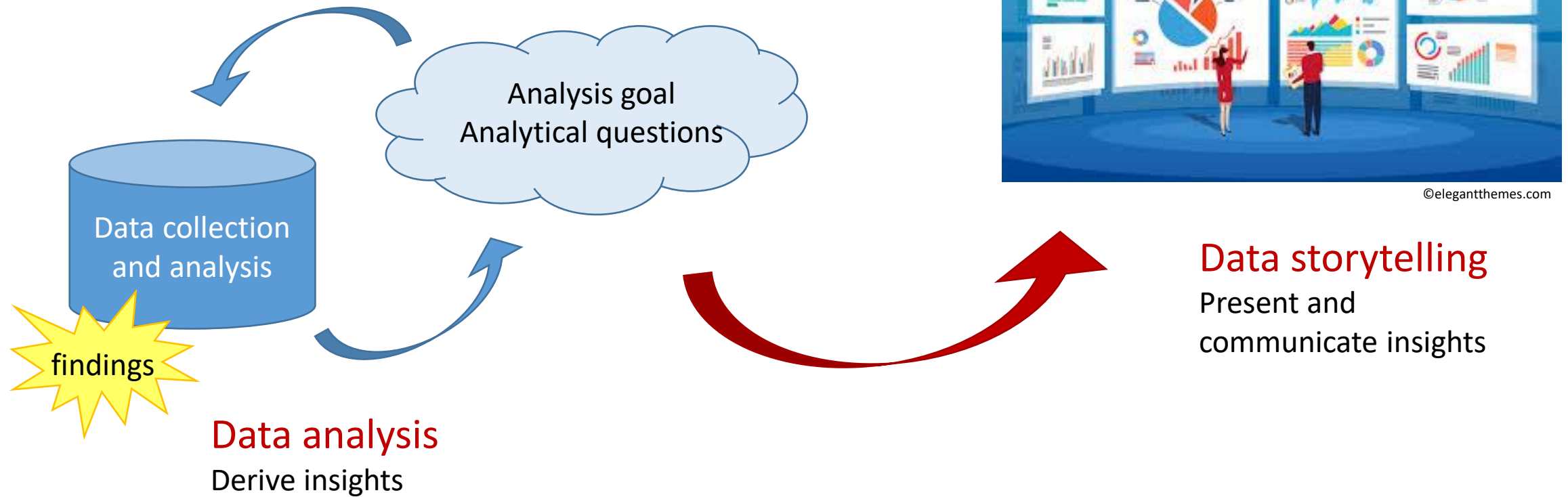
**Patrick Marcel - Verónica Peralta**

**LIFAT - University of Tours – France**

**eBISS 2022 - Cesena**

# Part 2 - Storytelling

# From insights to storytelling

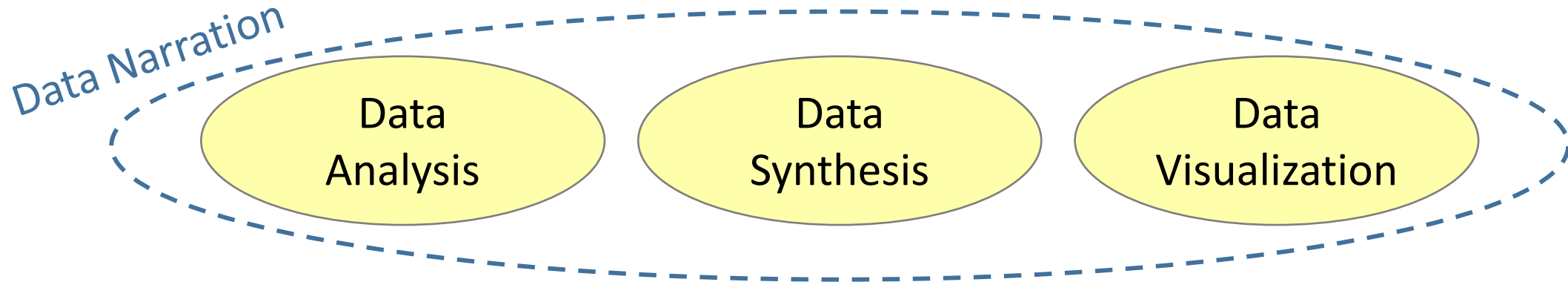


# Outline

- ❑ Part 1: Insights
- ❑ Part 2: Storytelling
  - What is a data narrative?
    - ❑ Definition and examples
    - ❑ Conceptual model
  - Crafting process
    - ❑ Focus on storytelling activities
  - Automation
- ❑ Perspectives

# Data narration = Narrating with data visualization

- ❑ The activity of producing narratives supported by facts extracted from data analysis, using interactive visualizations



📖 J. Hullman, S. Drucker, N. Riche, B. Lee, D. Fisher, E. Adar: "A Deeper Understanding of Sequence in Narrative Visualization", TVCG 19:12, 2013.





📖 S. Carpendale, N. Diakopoulos, N. Riche, C. Hurter: "Data-Driven Storytelling", Dagstuhl Reports 6:2, 2016.

# Data narratives

A **data narrative** is a **structured** composition of **messages** that

- (a) convey **findings** over the **data**, and,
- (b) are typically delivered via **visual means** in order to facilitate their reception by an intended audience.

Based on definitions of narrative and storytelling:

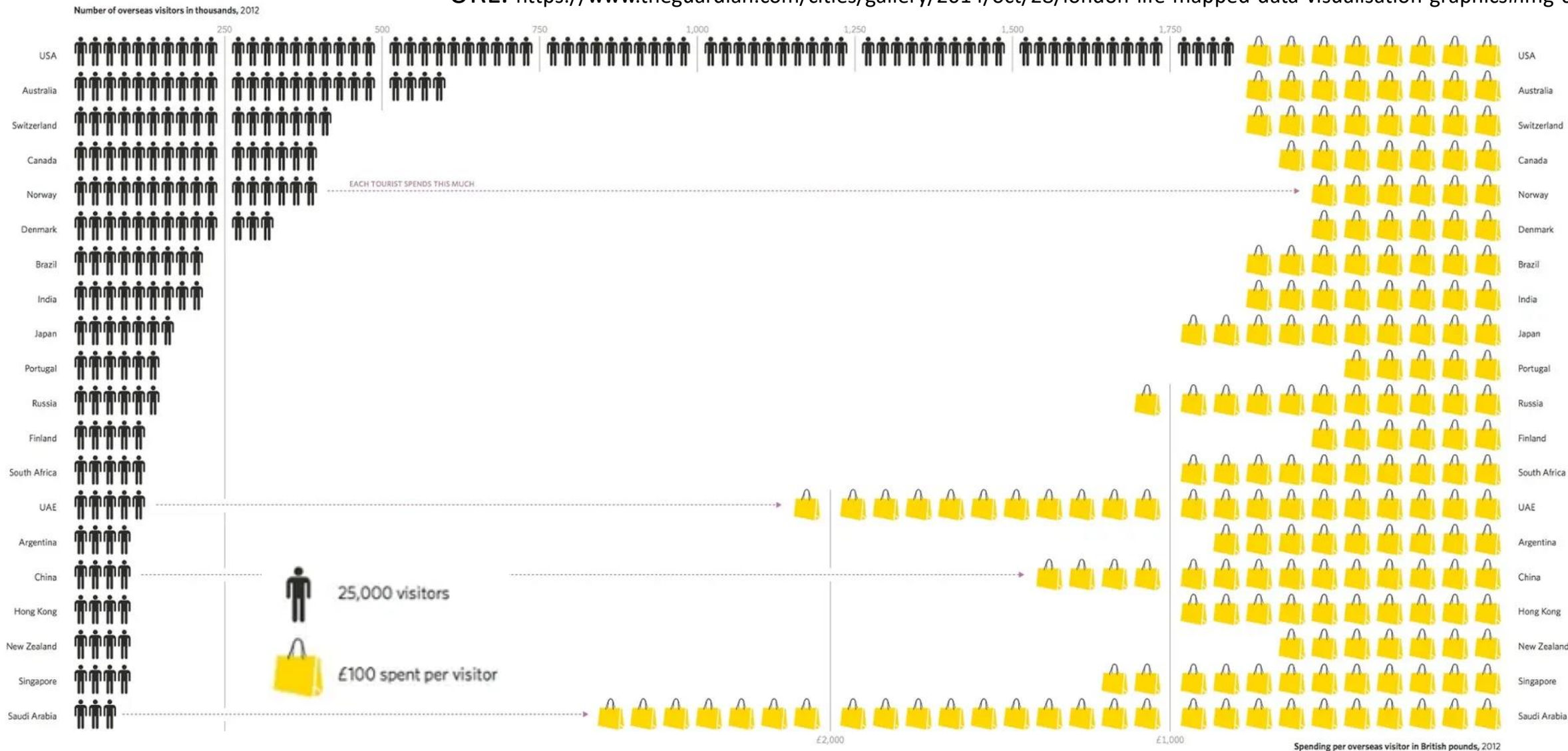
-  D. Elson: “Modeling narrative discourse”, Ph.D. thesis, Columbia University, 2012.
-  S. Chatman: “Story and Discourse: Narrative Structure in Fiction and Film”, Cornell paperbacks, 1980.
-  S. Chen, J. Li, G. Andrienko, N. Andrienko, Y. Wang, P. Nguyen, C. Turkay: “Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling”, TVCG 2018.
-  F. El Outa, M. Francia, P. Marcel, V. Peralta, P. Vassiliadis: “A conceptual model of data narrative for exploratory data analysis”, ER 2020.

# Examples

## Top 20 nations by spending by visitor to London in 2012

Source: The Guardian

URL: <https://www.theguardian.com/cities/gallery/2014/oct/28/london-life-mapped-data-visualisation-graphics#img-6>






# Examples

## Covid-19 situation update worldwide

Source: European Centre for Disease Prevention and Control

URL: <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>

**European Centre for Disease Prevention and Control**  
An agency of the European Union

[Home](#) > [All topics: A to Z](#) > [Coronavirus](#) > [Threats and outbreaks](#) > [COVID-19](#) > [Situation updates on COVID-19](#) > [Situation update worldwide](#)

[Situation updates on COVID-19](#)  
[Situation update for the EU/EEA and the UK](#)  
**[Situation update worldwide](#)**  
[14-day incidence of reported COVID-19](#)  
[COVID-19 country overviews](#)  
[Weekly surveillance report on COVID-19](#)  
[Data collection](#)

## COVID-19 situation update worldwide, as of 28 June 2020

Epidemiological update

[Twitter](#) [Facebook](#) [LinkedIn](#) [Email](#)

*The data presented on this page has been collected between 6:00 and 10:00 CET*

**Disclaimer:** National updates are published at different times and in different time zones. This, and the time ECDC needs to process these data, may lead to discrepancies between the national numbers and the numbers published by ECDC. Users are advised to use all data with caution and awareness of their limitations. Data are subject to retrospective corrections; corrected datasets are released as soon as processing of updated national data has been completed.

[Download today's data](#) [How is the data collected?](#)

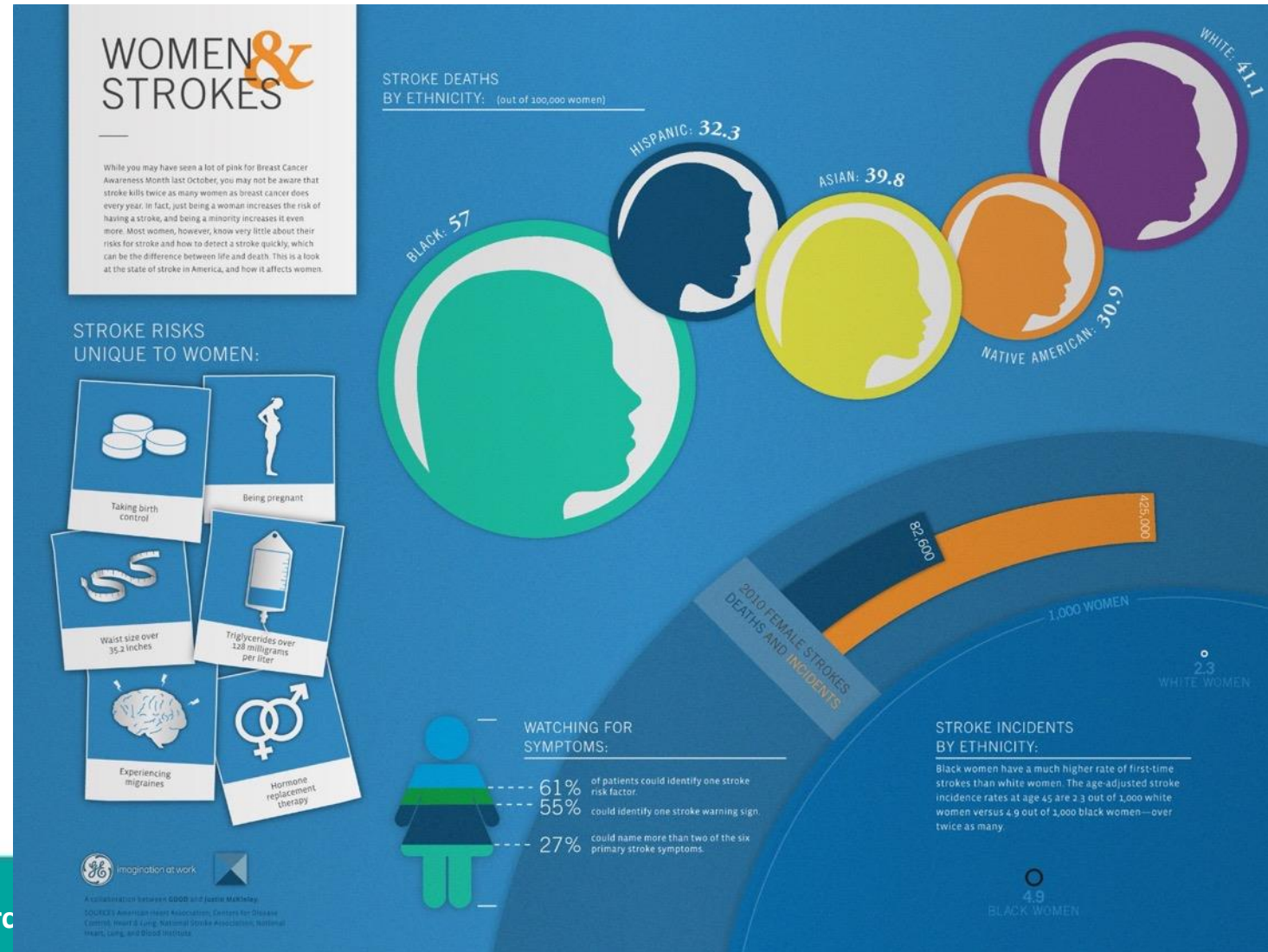


# Examples

## Stroke, a Silent Killer of Women, Facts About Women and Strokes

Source: GOOD

URL: <https://www.good.is/infographics/facts-about-women-and-strokes>

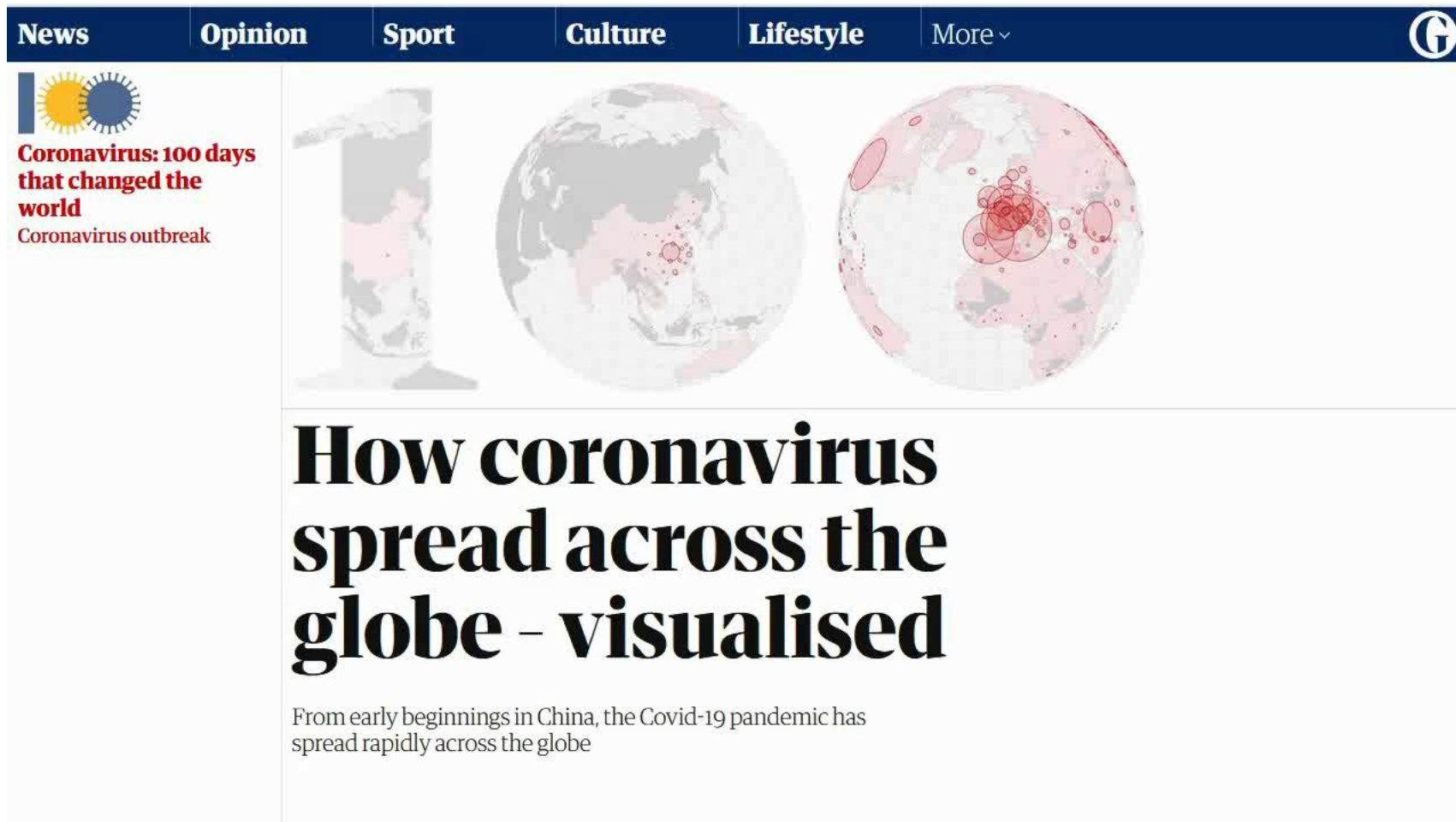


# Examples

## Coronavirus: 100 days that changed the world

Source: The Guardian

URL: <https://www.theguardian.com/world/ng-interactive/2020/apr/09/how-coronavirus-spread-across-the-globe-visualised>



The image is a screenshot of a web article from The Guardian. At the top is a dark blue navigation bar with white text links: 'News', 'Opinion', 'Sport', 'Culture', 'Lifestyle', and 'More v'. A white circular logo with a stylized 'G' is on the right. Below the navigation bar, on the left, is a small graphic of a yellow sun and a blue virus particle, followed by the text 'Coronavirus: 100 days that changed the world' and 'Coronavirus outbreak'. The main content area features a large, stylized '100' where the zeros are globes showing the spread of the virus with red circles of varying sizes. Below this graphic, the title 'How coronavirus spread across the globe - visualised' is written in a large, bold, black font. Under the title, a sub-headline reads: 'From early beginnings in China, the Covid-19 pandemic has spread rapidly across the globe'.

# Examples

## How BuzzFeed News Used Betting Data To Investigate Match-Fixing In Tennis

Source: BuzzFeed News

URL: <https://www.buzzfeednews.com/article/johntemplon/how-we-used-data-to-investigate-match-fixing-in-tennis#.vyKWjpWkn>

### The Tennis Racket

Betting worth billions. Elite players. Vi  
And suspicious matches at Wimbledon  
tennis authorities have kept secret for



**Heidi Blake**  
UK Investigations Editor, UK



Posted on January 17, 2016, at 4:58 p.m. ET



Tweet



Share



Copy

With GIFs.

**Secret files exposing** evidence of w  
of world tennis can today be reveale



**John Templon**  
BuzzFeed News Reporter

Posted on January 17, 2016, at 5:02 p.m. ET

Tweet

Share

Copy

The sport's governing bodies have b  
players – all of whom have ranked i  
and more than half of them will beg

It has been seven years since world t  
evidence about a network of players  
including Wimbledon following a la  
allowed to continue playing.

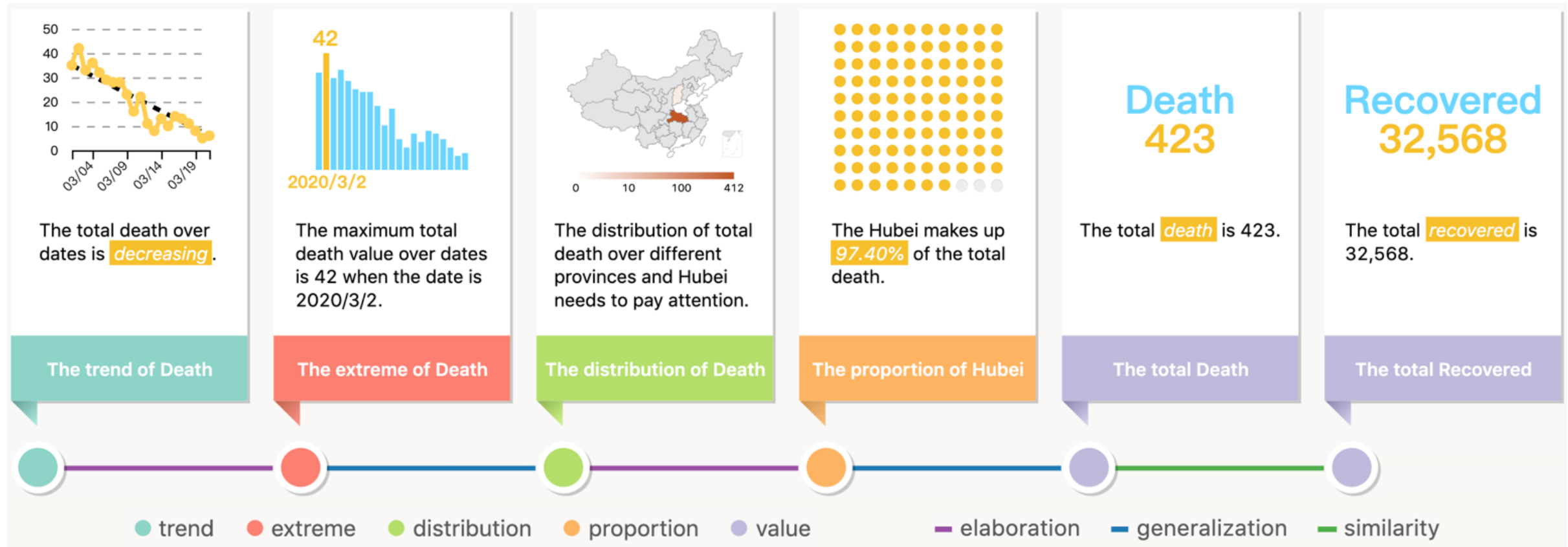
## How BuzzFeed News Used Betting Data To Investigate Match-Fixing In Tennis

**I'm John Templon, an investigative data reporter for BuzzFeed News. I spent the past 15 months analyzing tennis betting data to see if I could figure out whether players were fixing matches.**

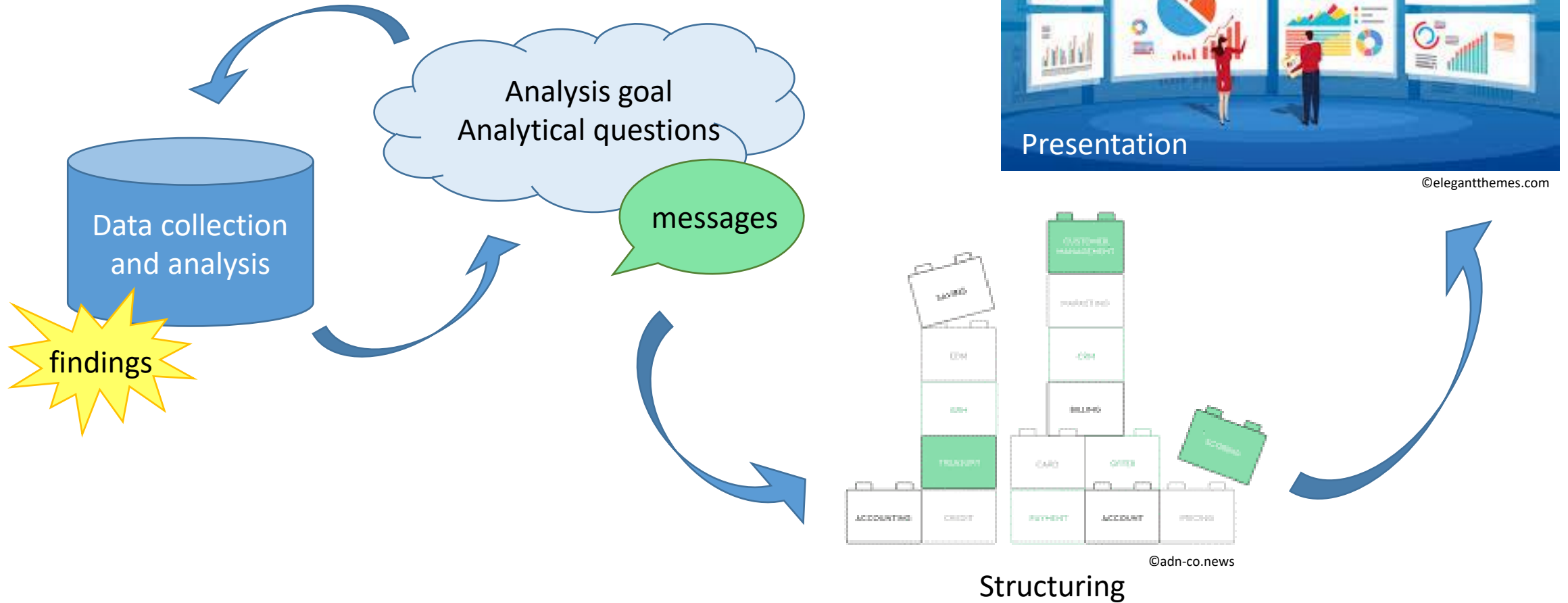
# Examples

## COVID-19 virus propagation in March 2020 in China

📖 D. Shi, X. Xu, F. Sun, Y. Shi, N. Cao: Calliope: “Automatic visual data story generation from a spreadsheet”, TVCG 27:2, 2021.



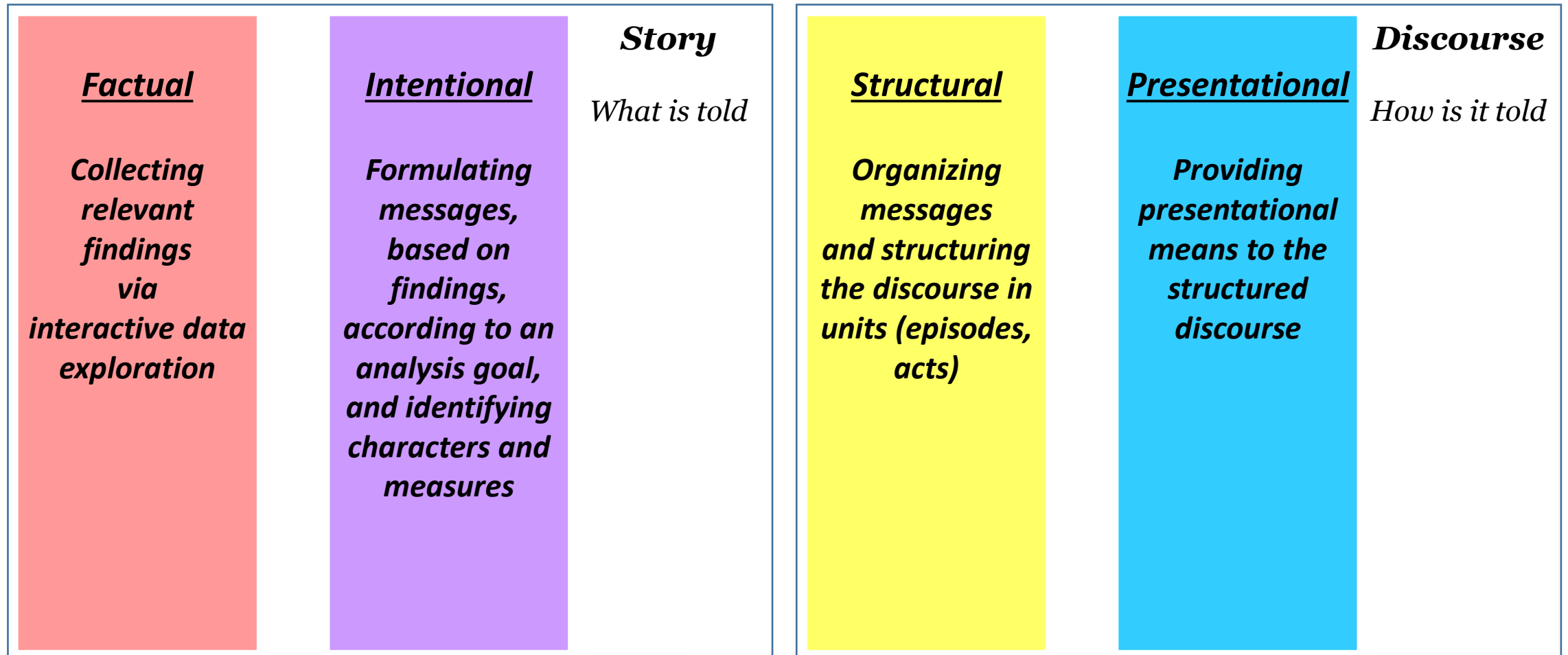
# Towards modeling data narratives





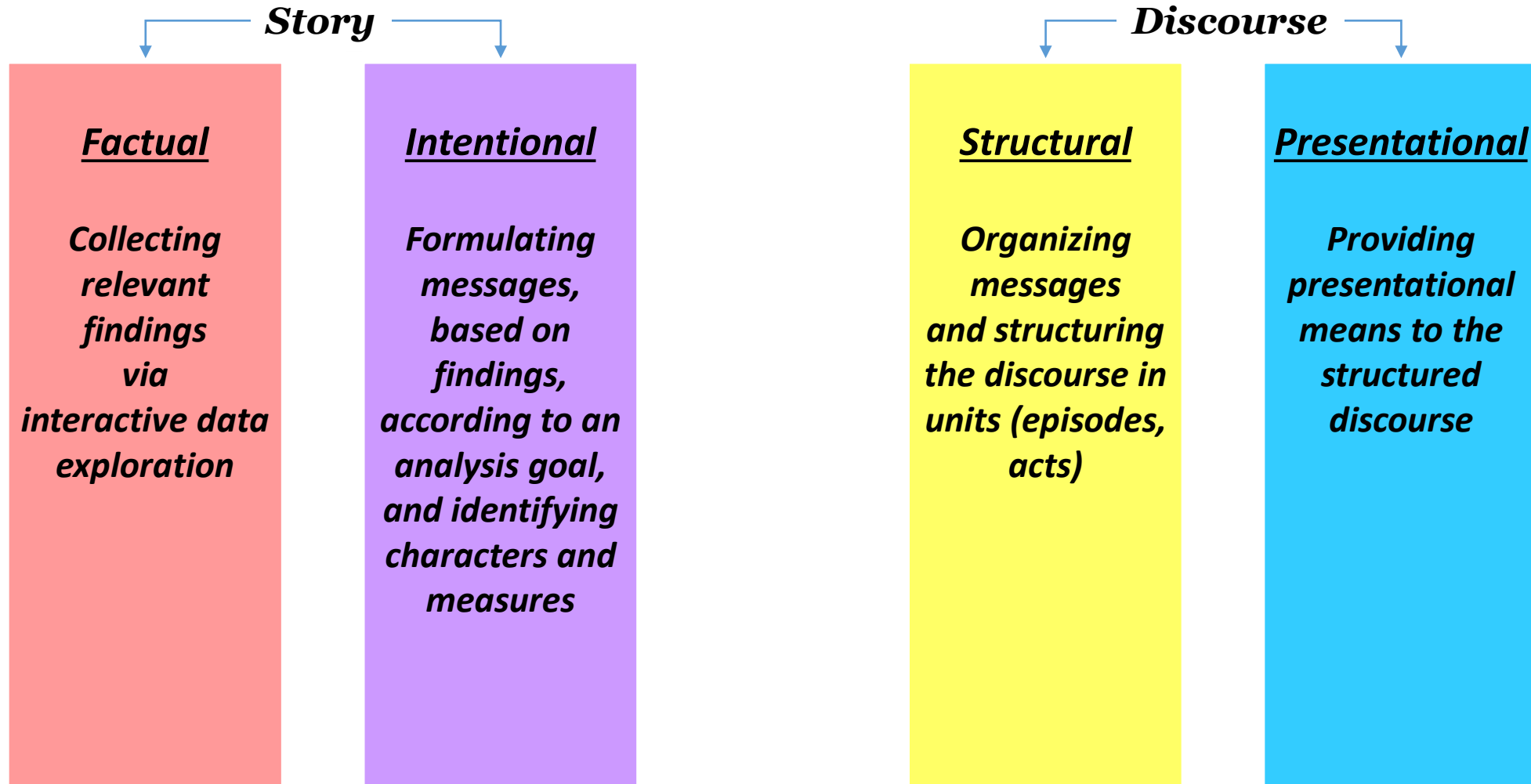
- 📖 S. Chatman: “Story and Discourse: Narrative Structure in Fiction and Film”, Cornell paperbacks, 1980.
- 📖 F. El Outa, M. Francia, P. Marcel, V. Peralta, P. Vassiliadis: “A conceptual model of data narrative for exploratory data analysis”, ER 2020.

# A model in 4 layers





# A model in 4 layers



# The model in practice



**Analytical Question:**  
Relation between ethnicity  
and stroke deaths

Implemented by

## **Collector**

Select ethnicity,  
 $\text{sum}(\text{stroke})/100000$   
from stroke\_deaths  
group by ethnicity

fetches

## **Finding**

Black:57,White:41,  
Asian:39.8,Hispanic:32,  
Native American:30.9

produces

## **Message:**

Uneven distribution of  
stroke death by ethnicity

brings out

## **Dashboard component**



renders

## **Episode**

Stroke death by ethnicity:  
uneven distribution  
Black 57 ; White:41;  
Asian:39.8; Hispanic: 32;  
Native America 30.9

narrates

refers to

Black:57

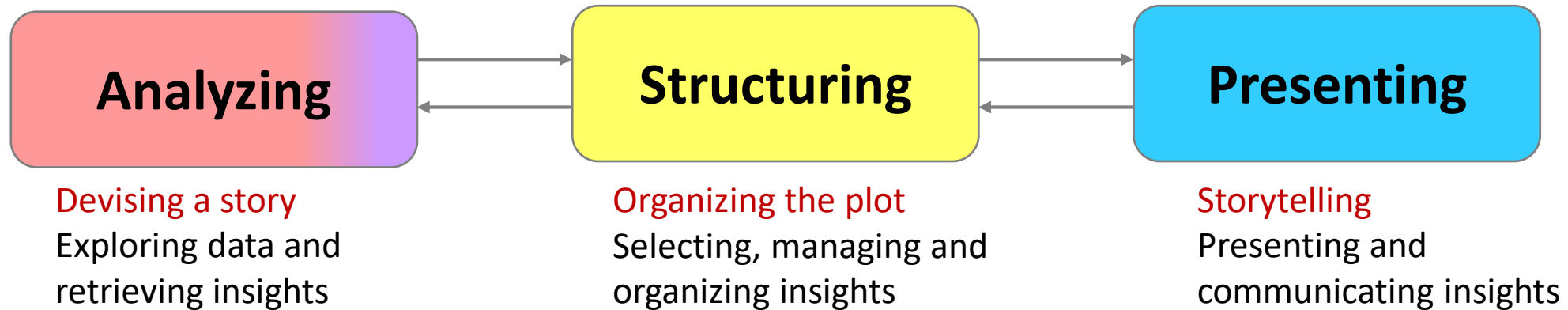
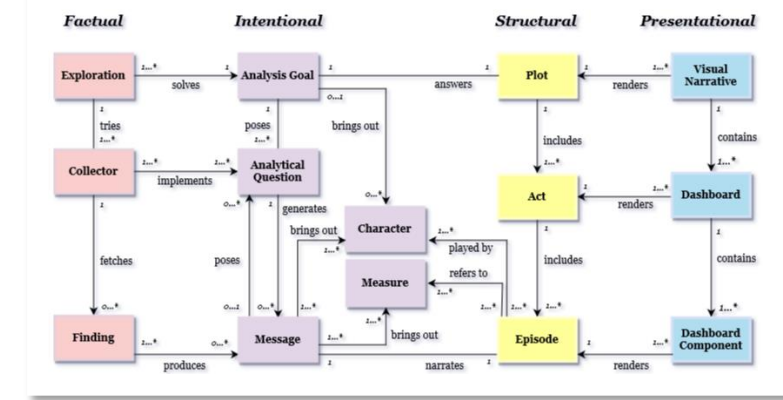
played by

ethnicity

# Crafting process

Focus on storytelling activities

# Crafting process in the literature

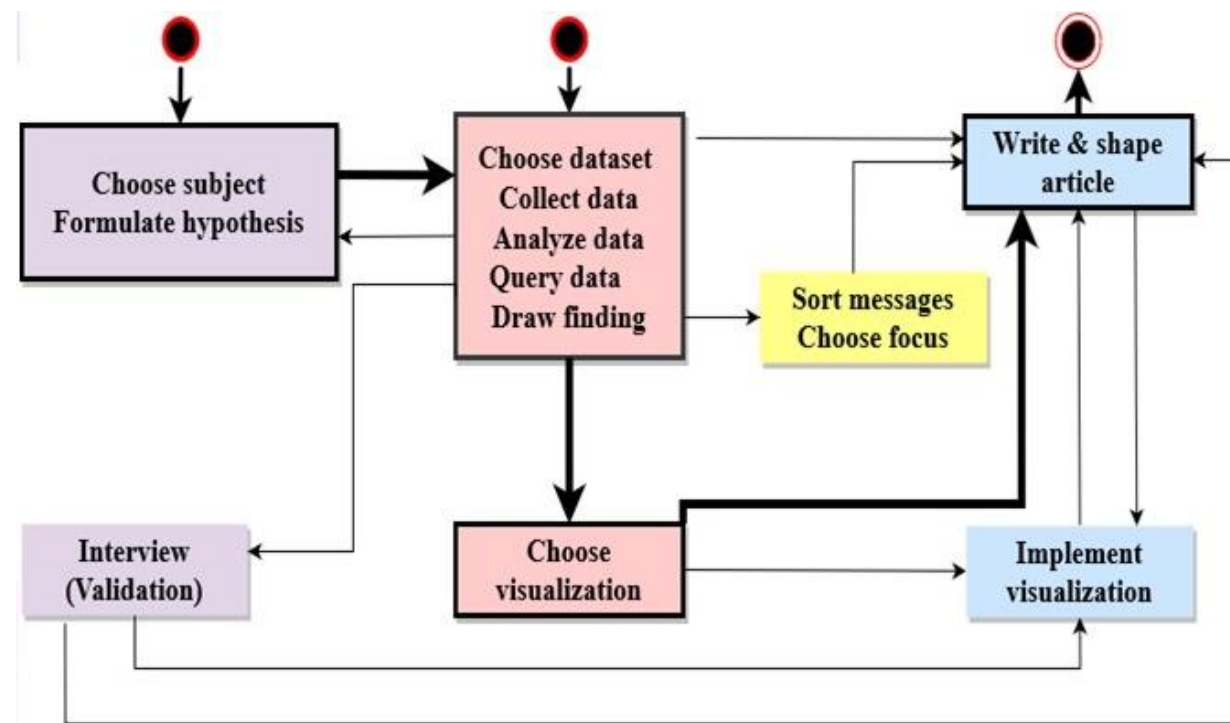


- ❑ No distinction between factual and intentional aspects
- ❑ Few covering of intentional aspects: analytical questions, messages...

📖 S. Chen, J. Li, G. Andrienko, N. Andrienko, Y. Wang, P. Nguyen, C. Turkay: “Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling”, TVCG 2018.

📖 B. Lee, N.H. Riche, P. Isenberg, S. Carpendale: “More than telling a story: Transforming data into visually shared stories”, IEEE Comp. Graph. Appl. 35:5, 2015.

# Crafting process of data journalists – A survey

[illegible]

 M. Chagnoux: “La datavisualisation, double point d’entrée du datajournalisme dans la PQR” (in french), Interfaces numériques 9:3, 2020.



# Concerning activities

## Analyzing

Dataset collection	Goal and question formulation
Data preparation [6], preprocessing [5]	Idea precede data search [1,19]
Collect information/data [6,8,13,15]	Question precede data search [19]
Collect data	Overview data
Choose datasets	Refine subject
Search further datasets	Choose subject
Prepare data	Formulate question
Collector trial	Formulate hypothesis
Explore data [4,7,8,9]	
See aspects of complex data [5]	
Analyze data	Message formulation
Verify hypothesis	Convey interesting information [5]
Visualization trial	Draw conclusion
Choose data visualization [13]	
Make data visible [19]	Message validation
Choose visualization	Cross-checking
Finding formulation	Validation
Fact extraction [17], Choose fact [13]	Interviews
Explore interrelationship [5,8]	
Draw finding	
Draw pattern	
Findings validation	
Discussion	

## Structuring

Determine audience
Conceptualize targeted audience [6]
Select messages
Choose facts [14]
Sort messages
Choose narrative structure
Realization [6]
Tie fact into story [8]
Make story [9]
Construct storyline [13]
Choose focus
Message mapping
Organize information pieces [5]
Organize facts [14]
Fact composition [17]
Organize information pieces [15]

## Presenting

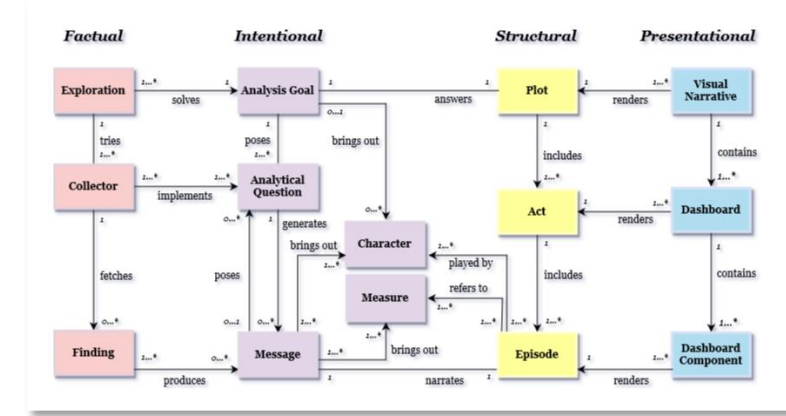
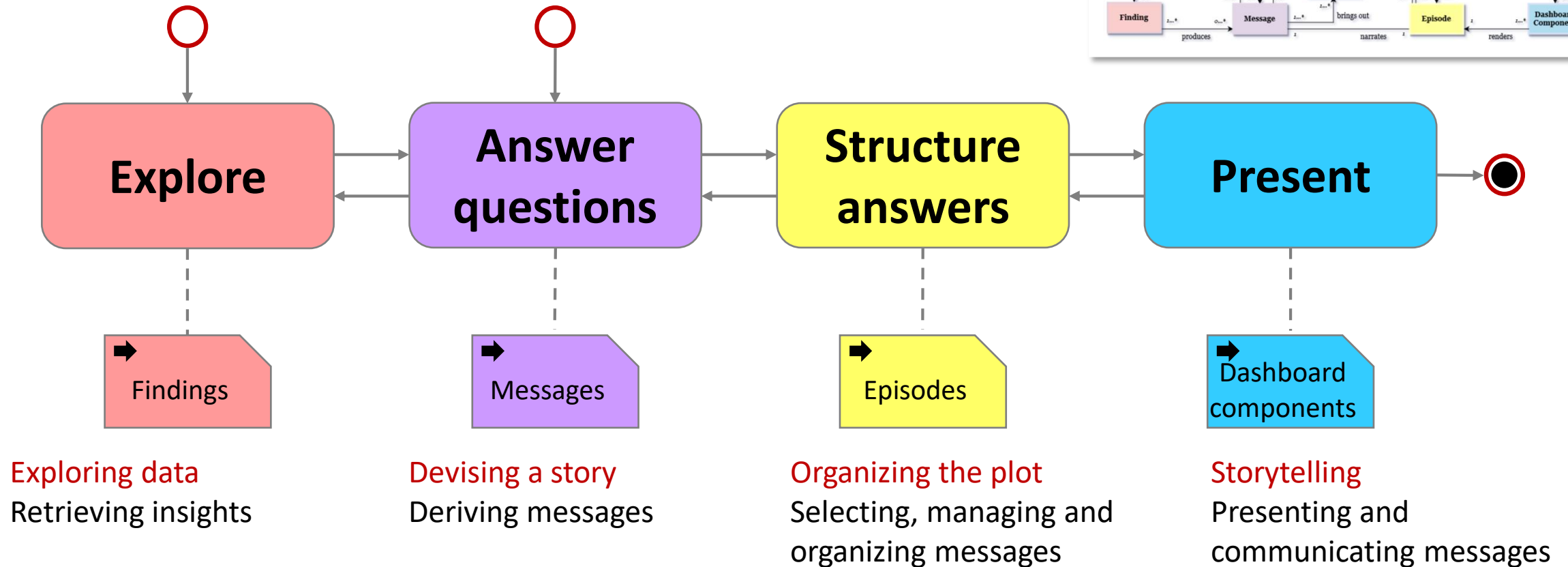
Visual narrative setting
Visualization design [6]
Tell story [9]
Visualize data story [14]
Presentation synthesis [17]
Realize storyboard [13]
Refining presentation [15]
Choose form of article
Shape article
Interactivity choice
Set up interactivity [19]
Dashboard implementation
Visualization development [6]
Convey message through visualization [19], Convey clear message [14]
Visual encoding [15]
Implement visualization
Test visualization
Choose final visualization



# Requirements for a crafting process

- ❑ A comprehensive process should satisfy the following requirements:
  - (R1) cover the activities and the paths identified by the survey with data journalists
  - (R2) cover the activities of the three phases identified from the literature,
  - (R3) allow the free back and forth transition between phases,
  - (R4) clearly delineate the different layers of the conceptual model within its activities.

# A crafting process in 4 phases



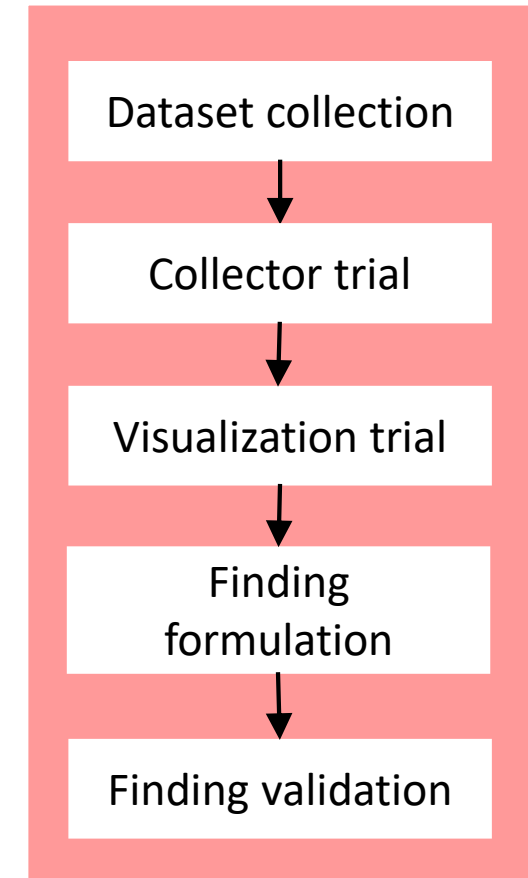
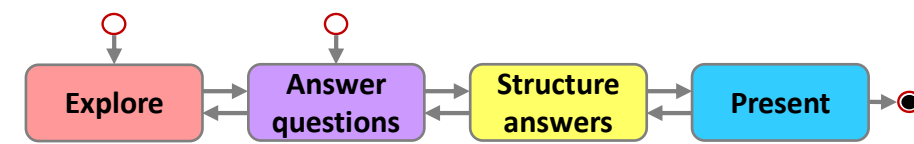
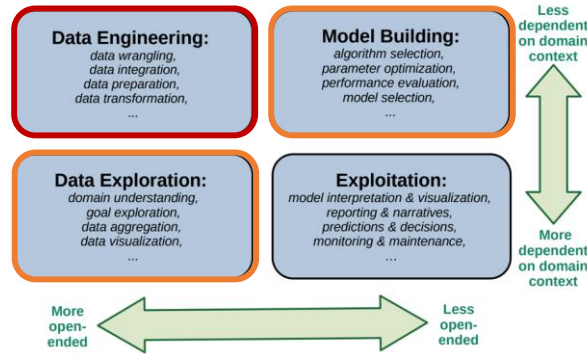
📖 F. El Outa, P. Marcel, V. Peralta, R. da Silva, M. Chagnoux, P. Vassiliadis: “Data narrative crafting via a comprehensive and well-founded process”, ADBIS 2022.

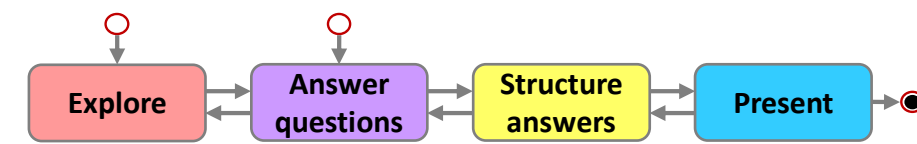
# Explore

## ❑ Concerns many tasks:

- Choice of data sets: source selection, data extraction, integration, preprocessing... (**Data Engineering**)
- Try (and reuse) of collectors: data querying, profiling and mining... (**Data Exploration and Model Building**)
- Try of visualizations: crosstabs, graphics, clusters, statistical tests...
- Formulation of findings (and relationships)
- Validation of findings: crosschecking, statistical tests

## ❑ It's time-consuming (measured in days or even months)





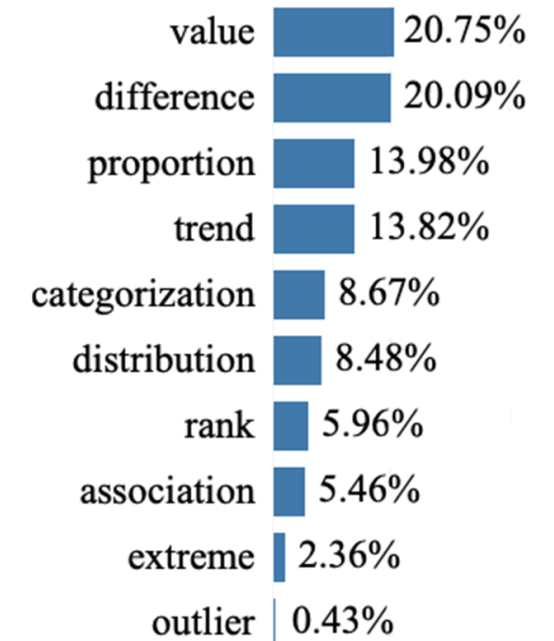
# Explore

## □ Typical types of insights

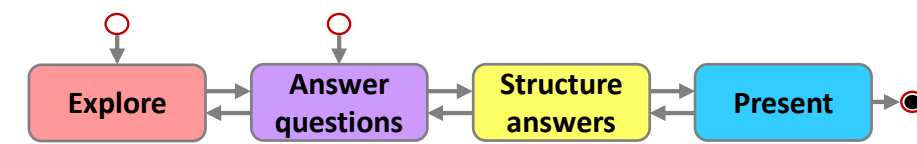
Fact type	Analysis task
Value	Retrieve value
Categorization	Filter
Aggregation	Compute derived value
Extreme	Find extreme
Rank	Sort
Proportion	-
Difference	Compare
Outlier	Find anomalies
Association	Correlate

Survey of 245 examples of fact sheet design

Survey of 230 high-quality animated/motion infographics and data videos



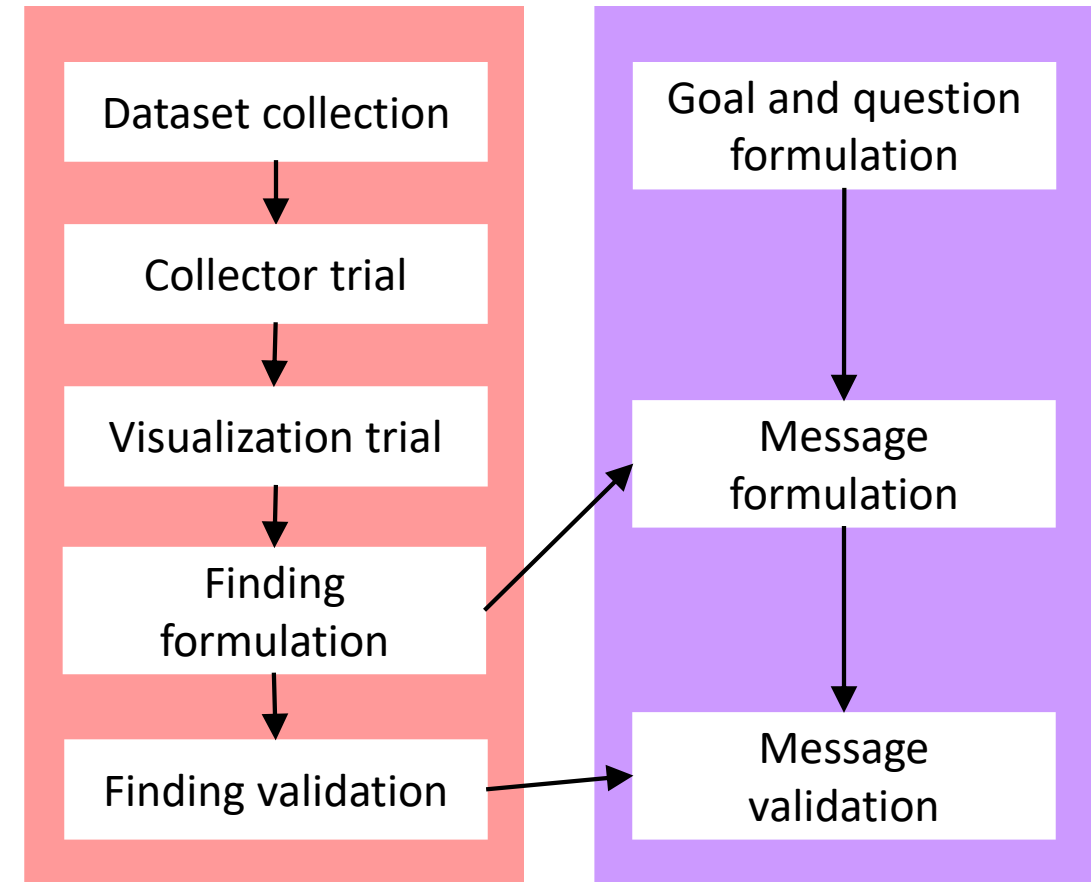
- 📖 Y. Wang, Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, D. Zhang: “DataShot: Automatic Generation of Fact Sheets from Tabular Data”, TVCG 26:1, 2020.
- 📖 D. Shi, X. Xu, F. Sun, Y. Shi, N. Cao: Calliope: “Automatic visual data story generation from a spreadsheet”, TVCG 27:2, 2021.

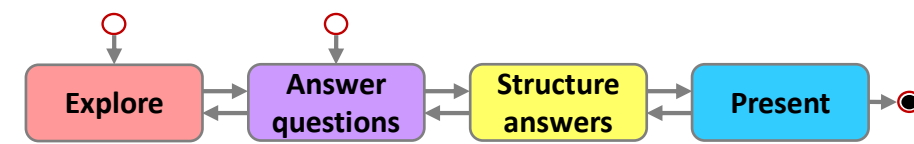


# Answer questions

## ❑ Concerns:

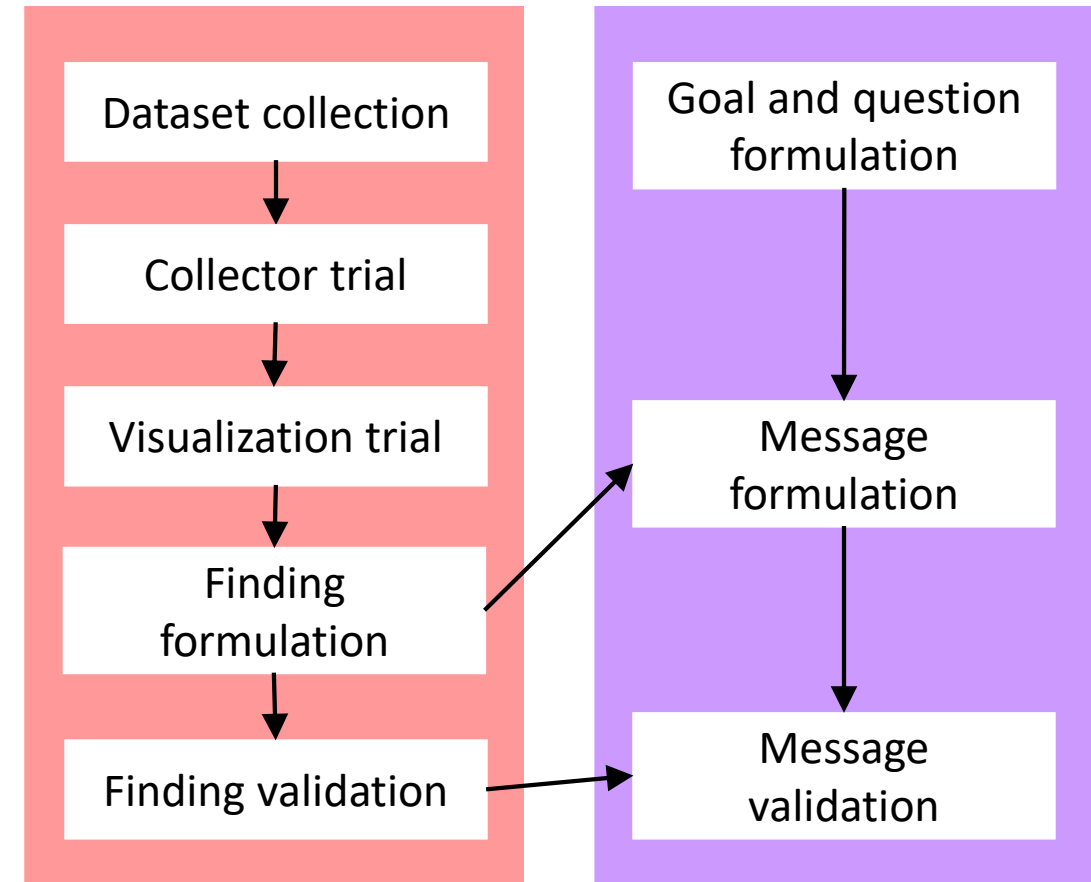
- Formulation of story goal and analytical questions
- Formulation of messages and identification of characters and measures
- Validation of messages





# Answer questions

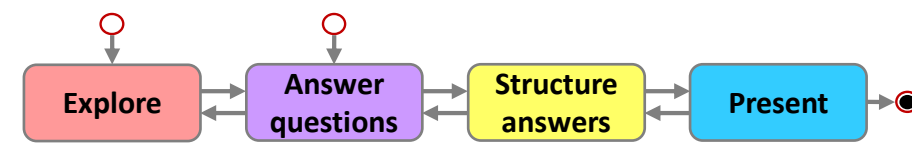
- ❑ **Three ways to start a data narrative:**
  1. having a precise idea on mind
    - Formulate a precise goal and list of questions
  2. having a vague idea refined during exploration
    - Formulate a vague goal and a first question
    - Refine them during data exploration
  3. having no idea before exploring the data
    - Start by exploring data



📖 L. Battle, J. Heer: “Characterizing exploratory visual analysis: A literature review and evaluation of analytic provenance in tableau”, Comput. Graph. Forum 38:3, 2019.

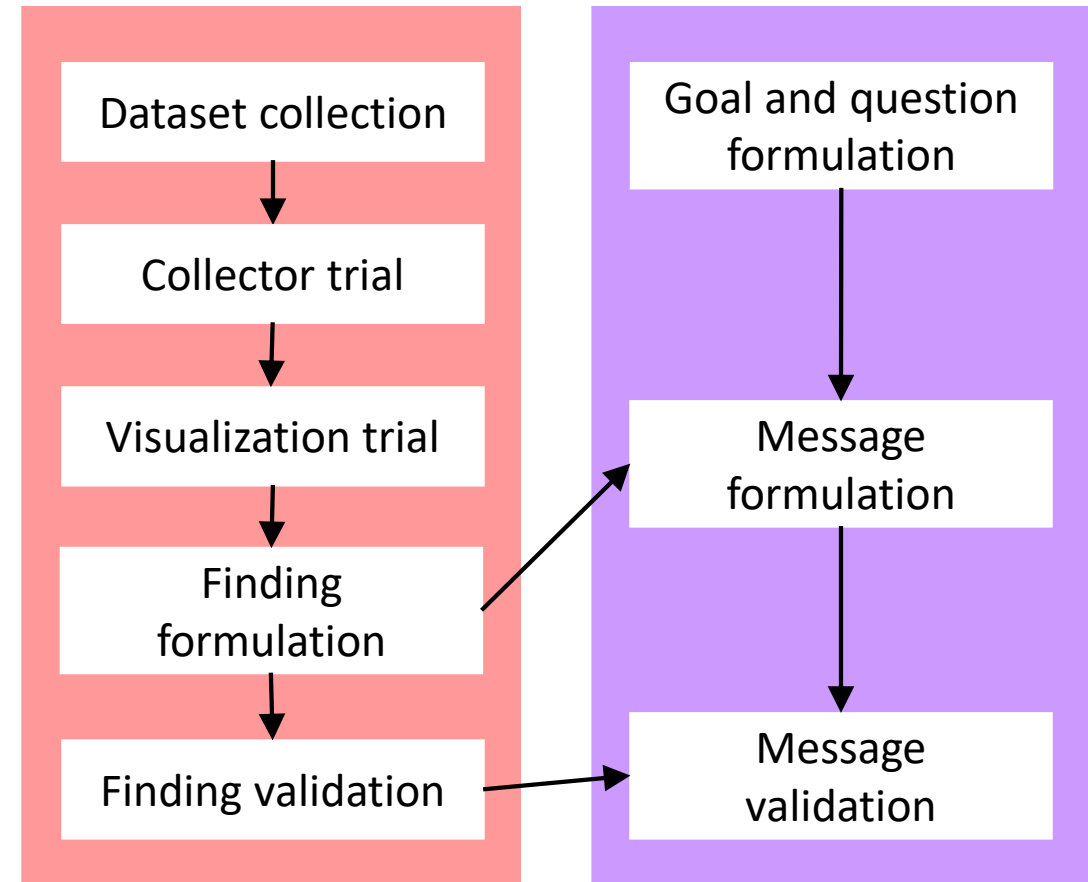
📖 W. Weber, M. Engebretsen, H. Kennedy: “Data stories: Rethinking journalistic storytelling in the context of data journalism”, Studies in Communication Sciences 18, 2018.





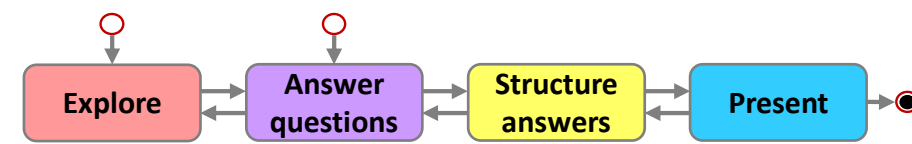
# Answer questions

- ❑ **Findings may lead to:**
  - Formulate messages for the audience: summarizing and pointing findings
  - Identify characters and measures: highlight important information
  - Formulate new analytical questions
- ❑ **Two types of validation:**
  - For findings: crosschecking, statistical tests
  - For messages: comparison with previous works and experts' advice



📖 M. Chagnoux: “La datavisualisation, double point d’entrée du datajournalisme dans la PQR” (in french), Interfaces numériques 9:3, 2020.

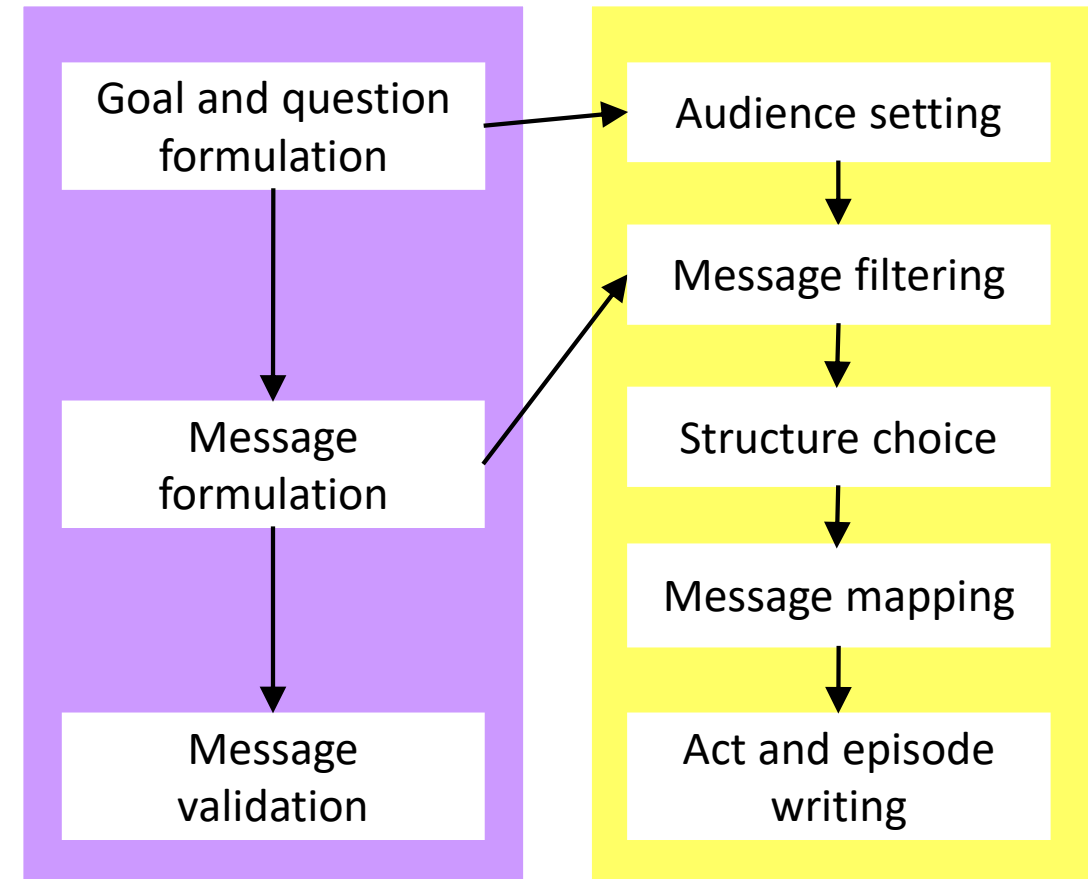
📖 R. Ondzigue, V. Peralta, T. Devogele, F. El Outa, S. Maghendji, E. Ngounou: “A data narrative about tuberculosis pandemics in Gabon”, DARLI-AP 2022.

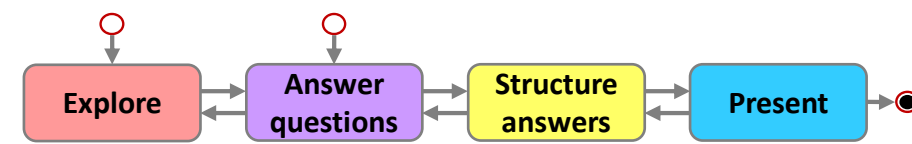


# Structure answers

## ❑ Concerns:

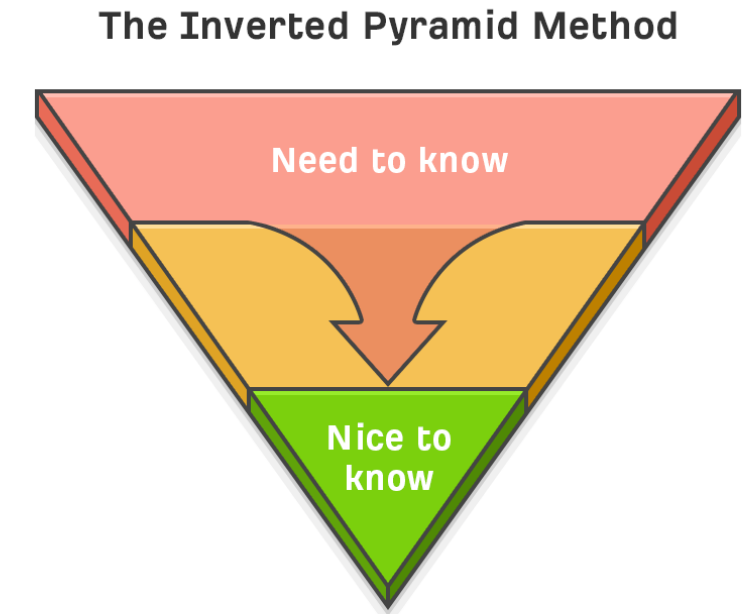
- Determination of the audience
- Selection of messages
- Choice of **narrative structure**
- Mapping of messages to acts
  - ❑ Arrangement of narrator's thoughts into different layers (acts and episodes)
- Writing of acts and episodes
  - ❑ Content, structure, transitions...





# Structure answers

- ❑ **Many types of structures:**
  - Sequencing (transitions): hierarchical, parallel
  - Ordering (timelines): visual metaphor, scale, layout
- ❑ **Ex. Inverted Pyramid**
  - Striking messages first
  - Details for interests readers



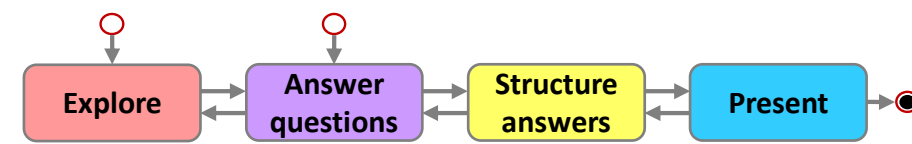
© <https://ahrefs.com/blog/seo-copywriting/>

ahrefs

📖 R. Kosara: “An argument structure for data stories”, EuroVis 2017.

📖 J. Hullman, R. Kosara, H. Lam: “Finding a clear path: Structuring strategies for visualization sequences”, Comput. Graph. Forum 36:3, 2017.

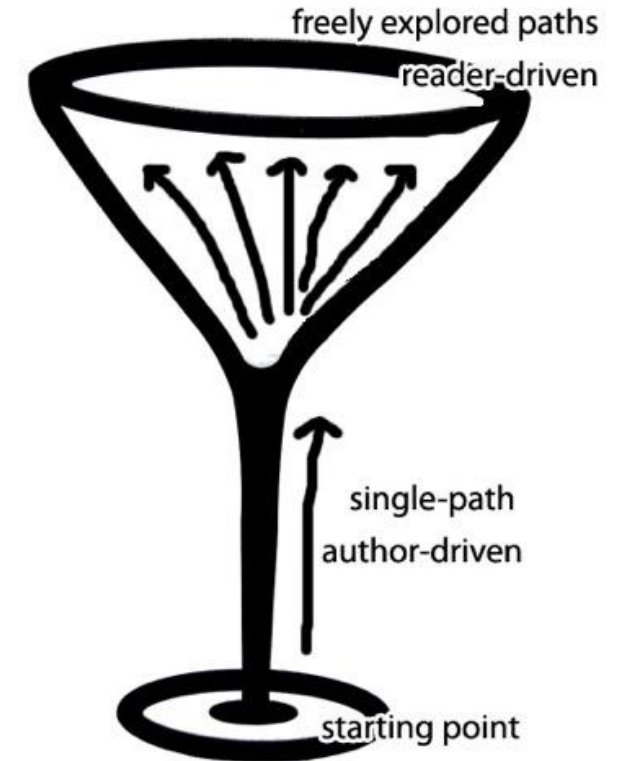
📖 M. Brehmer, B. Lee, B. Bach, N.H. Riche, T. Munzner: “Timelines revisited: A design space and considerations for expressive storytelling”, TVCG 23:9, 2017.



# Structure answers

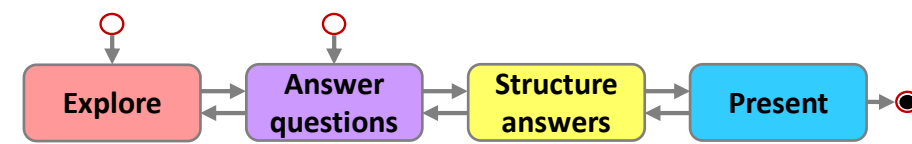
## □ Level of interactivity:

- Author-driven: linear pass, explanations, no interactivity
  - Ex. structure : Inverted Pyramid (largely used by data journalists)
  - Ex. visual narrative : Infographics, slideshows, data videos...
- Reader-driven: no ordering, no explanations
  - Ex. structure : Hierarchical structure with links (largely used in BI)
  - Ex. visual narrative : Portfolios, querying interfaces
- Hybrid
  - Ex. structure : Martini glass
  - Ex. visual narrative : eBooks, web pages



📖 E. Segel, J. Heer: “Narrative visualization: Telling stories with data”, TVCG 16(6), 2010.

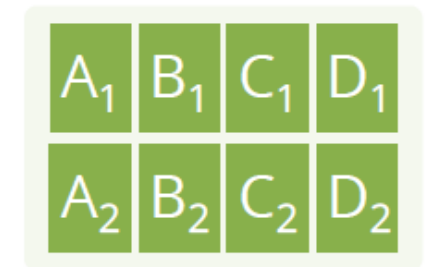
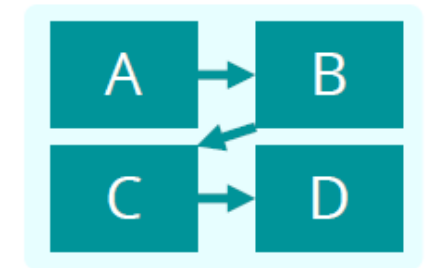
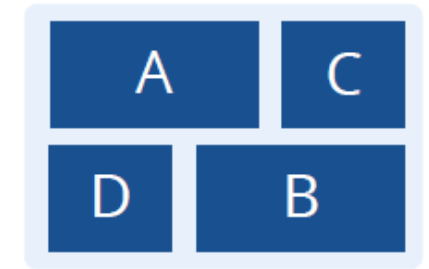
📖 W. Weber, M. Engebretsen, H. Kennedy: “Data stories: Rethinking journalistic storytelling in the context of data journalism”, Studies in Communication Sciences 18, 2018.



# Structure answers

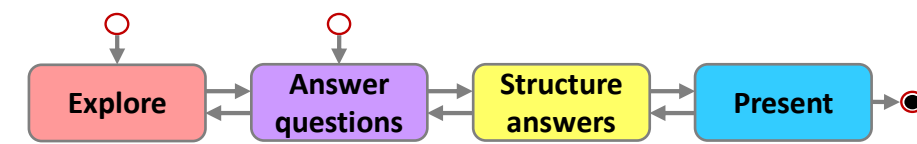
## ❑ Content structures

- How acts (and latterly dashboards) are typically structured
  - ❑ Random order (52%)
  - ❑ Sequential order (26%)
  - ❑ Multiple series (22%)



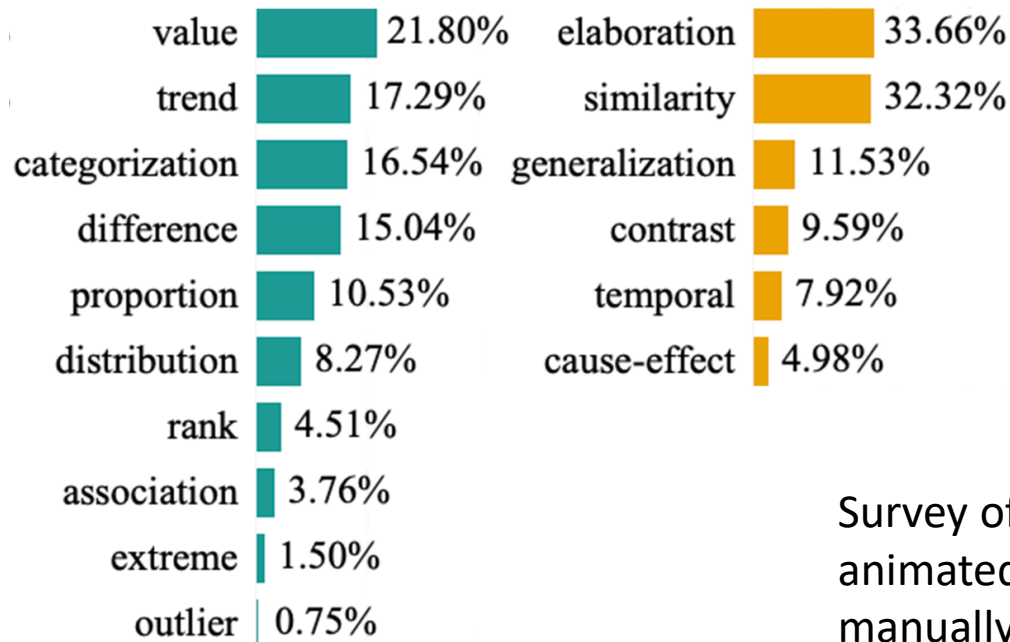
Survey of 245 examples of fact sheet design

📖 Y. Wang, Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, D. Zhang: “DataShot: Automatic Generation of Fact Sheets from Tabular Data”, TVCG 26:1, 2020.



# Structure answers

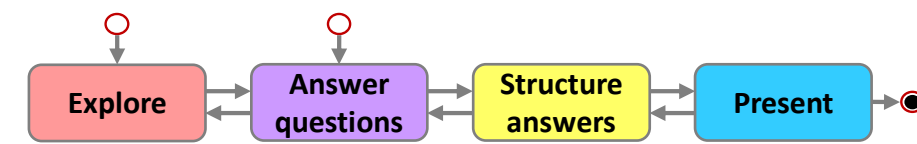
## □ Starting points and transitions



Survey of 230 high-quality animated/motion infographics and data videos, manually segmented into 4186 episodes.

Fact Types	$r_s$	$r_t$	$r_c$	$r_a$	$r_e$	$r_g$
Value	45.6	8.9	0.0	4.2	26.8	14.5
Difference	41.6	6.7	0.0	5.8	31.1	14.8
Proportion	52.1	7.3	0.0	5.2	22.4	13.0
Trend	34.7	9.4	8.2	7.1	28.2	12.4
Categorization	37.7	3.4	0.0	3.4	47.5	7.8
Distribution	49.0	12.1	0.0	4.4	22.3	12.1
Rank	43.8	11.7	0.0	6.6	34.3	3.6
Association	31.0	5.6	15.1	7.1	26.2	15.1
Extreme	51.8	5.6	0.0	3.7	25.9	13.0
Outlier	20.0	10.0	0.0	10.0	40.0	20.0





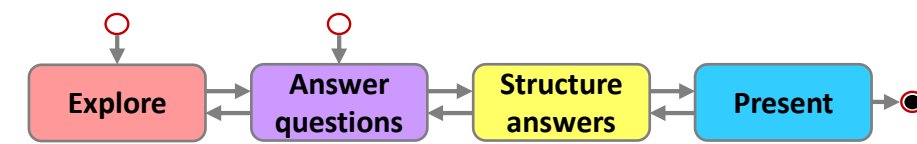
# Structure answers

## □ Cognitive cost of transition types

Temporal > (Dimension | Measure) > Granularity

Representation	Relevant Transition Types	Relevant Interactions
Dependent variable	Comparative – Measure walk	Sort, Derive, Navigate (Distortion), Coordinate (small multiples)
Independent variable	Comparative – Dimension walk	Filter (independent variable, such as with query widget), Navigate (scroll, pan), Coordinate (small multiples)
Time	Temporal	Filter (direct selection, slider), Coordinate (small multiples)
Hierarchical relation	Granularity	Filter (direct selection, query widget, slider), Navigate (overview & detail, zoom, semantic zoom), Derive (aggregate)

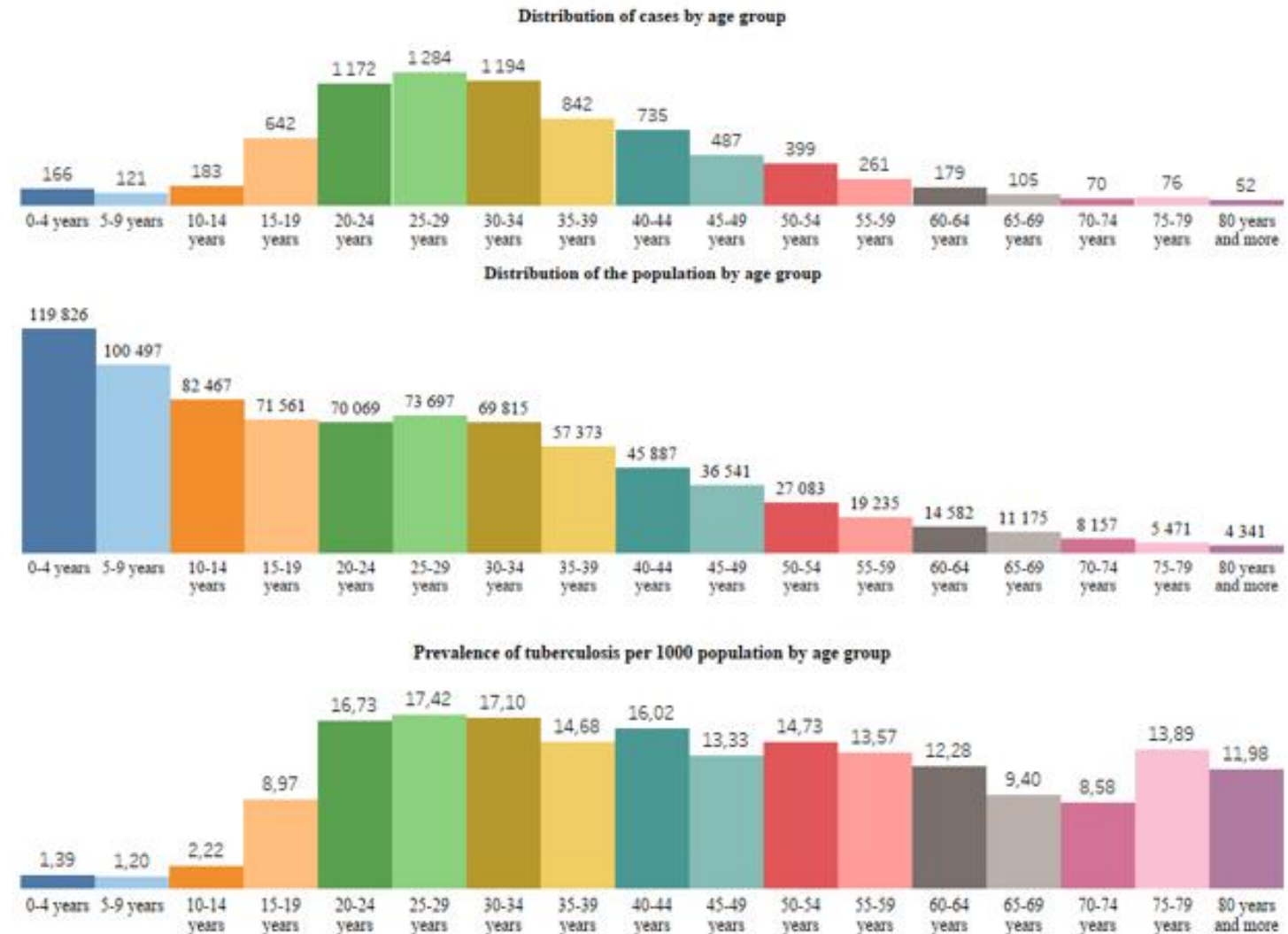
 J. Hullman, S. Drucker, N. Riche, B. Lee, D. Fisher, E. Adar: “A Deeper Understanding of Sequence in Narrative Visualization”, TVCG 19:12, 2013.



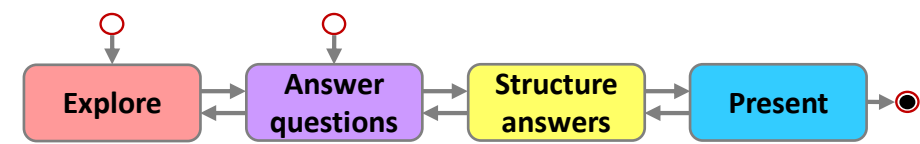
# Structure answers

## □ Example:

- Analysis of tuberculosis cases per age group



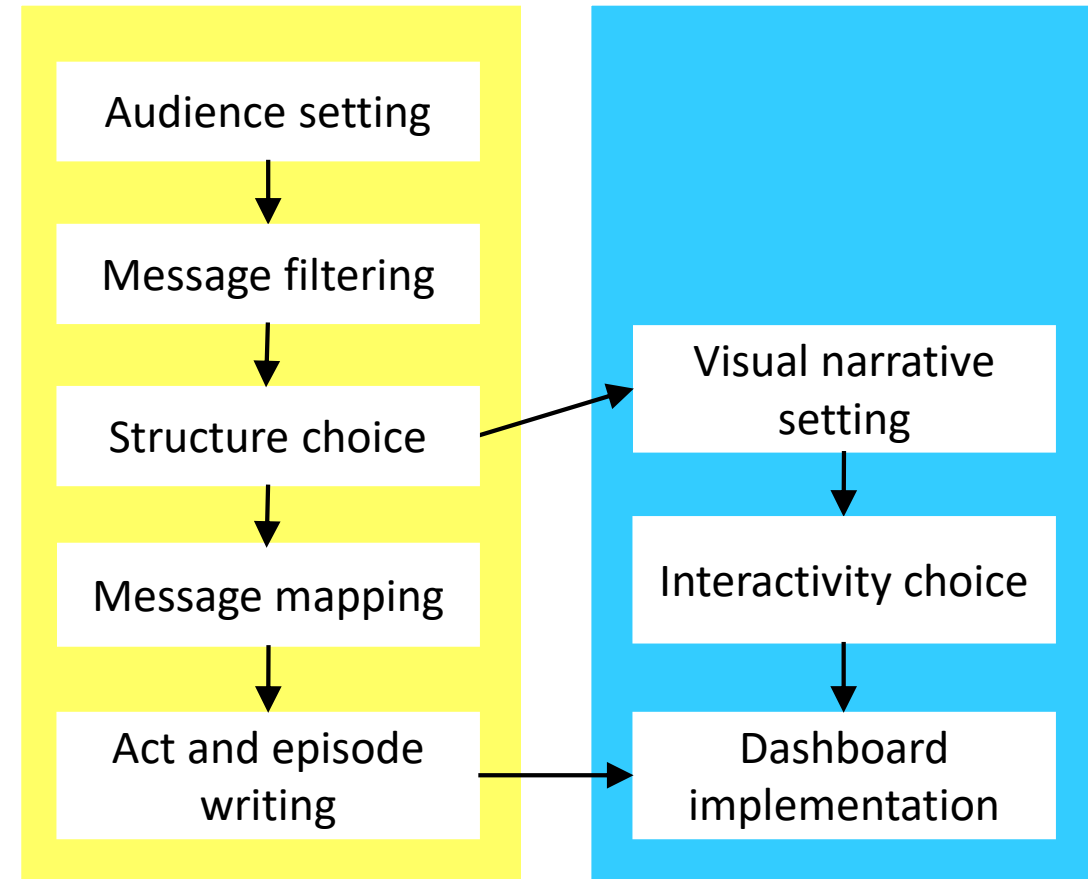
R. Ondzigue, V. Peralta, T. Devogele, F. El Outa, S. Maghendji, E. Ngounou: “A data narrative about tuberculosis pandemics in Gabon”, DARLI-AP 2022.

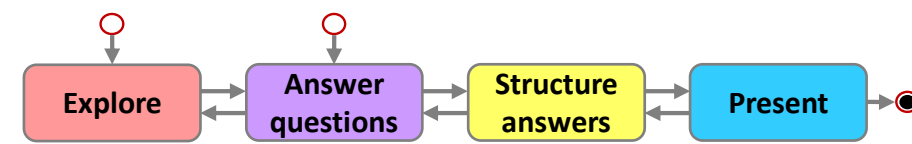


# Present

## ❑ Concerns:

- Setting of visual narrative
- Choice of interactivity implementation
- Dashboard implementation



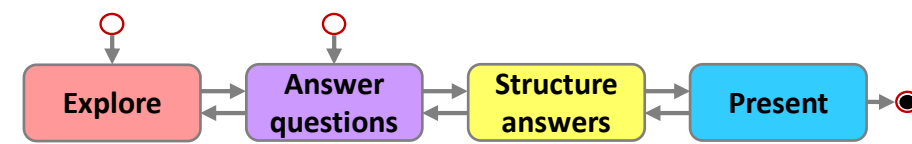


# Present

## ❑ Many visual narrative styles:

- magazine-style (images embedded in a page of text),
- comic strip (multiple images organized as a linear path),
- partitioned poster (multiple images organized in a loose order),
- flow chart (a diagram that depicts a process, system or algorithm),
- slide show (a series of still images in a prearranged sequence),
- video/animation

📖 E. Segel, J. Heer: “Narrative visualization: Telling stories with data”, TVCG 16(6), 2010.

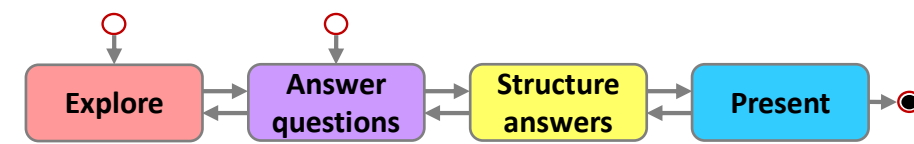


# Present

## ❑ Implementation of dashboards and components

- Choice of the right components (e.g. position, size, shape, color)
- Manual organization (specific cognitive tasks) vs. automatic derivation (according to the nature of data)
- Learning most adapted visualizations per insight type

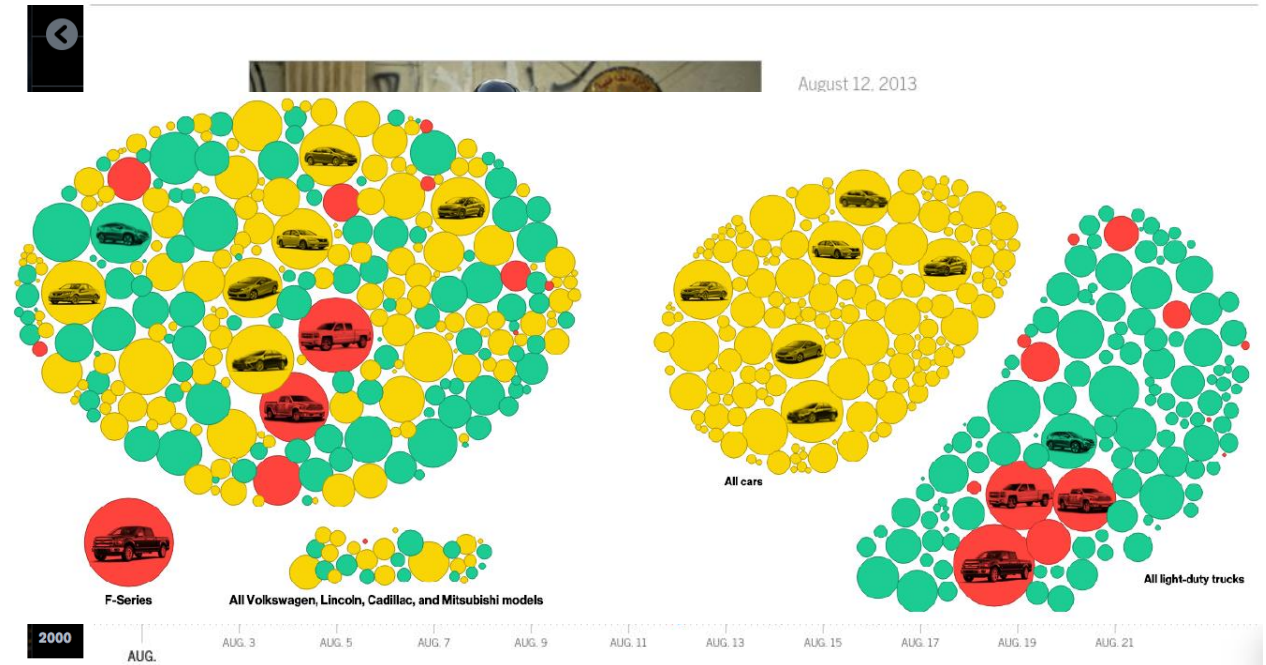
- 📖 A. Lavalley, A. Mate, J. Trujillo, S. Rizzi: “Visualization requirements for business intelligence analytics: A goal-based, iterative framework”, RE 2019.
- 📖 R. Brath, M. Peters: “Dashboard design: Why design is important”, DM Direct 2004.
- 📖 T. Blount, L. Koesten, Y. Zhao, E. Simperl: “Understanding the use of narrative patterns by novice data storytellers”, CHIRA 2020.
- 📖 M. Golfarelli, S. Rizzi: “A model-driven approach to automate data visualization in big data analytics”, Inf. Vis. 19:1, 2020.
- 📖 Y. Wang, Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, D. Zhang: “DataShot: Automatic Generation of Fact Sheets from Tabular Data”, TVCG 26:1, 2020.



# Present

- ❑ **Techniques for communicating in a creative way**
  - Communicating narrative and explaining data
  - Linking separated story elements
  - Enhancing structure and navigation
  - Providing controlled exploration

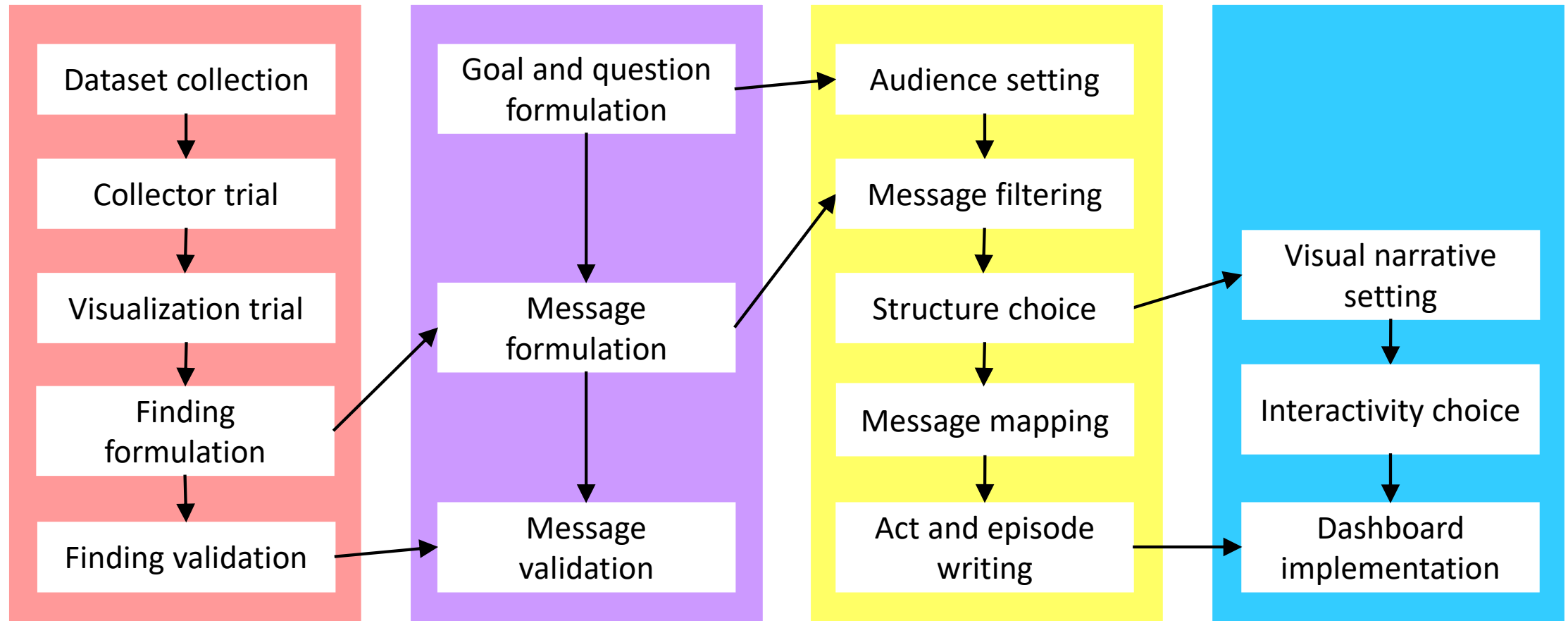
## Timeline: Egypt in Turmoil



- 📖 C.D. Stolper, B. Lee, N. Henry Riche, J. Stasko: “Emerging and recurring datadriven storytelling techniques: Analysis of a curated collection of recent stories”, Tech. rep., 2016.
- 📖 A. L. G. “The Evolution of Israeli Politics”. <http://www.economist.com/blogs/economist-explains/2015/03/economist-explains-11>, 2015.
- 📖 M. C. Klein. “Bubble to Bust to Recovery”. <http://www.bloomberg.com/dataview/2014-02-25/bubble-to-bust-to-recovery.html>, 2014.
- 📖 A. America. “Egypt in Turmoil”. <http://america.aljazeera.com/articles/timeline-egypt-inturmoil0.html>, 2013.
- 📖 A. Pearce, B. Migliozi, D. Ingold. “Scientific Proof that Americans are Completely Addicted to Trucks”. <http://www.bloomberg.com/graphics/2015-auto-sales/>, 2015.



# Overall activities for data narrative crafting



 F. El Outa, P. Marcel, V. Peralta, R. da Silva, M. Chagnoux, P. Vassiliadis: “Data narrative crafting via a comprehensive and well-founded process”, ADBIS 2022.

# Automation


and tools

# Automatic generation of data narratives


## □ Some tools:

- DataShot: Generation of a **fact sheet** from tabular data
- Calliope: Generation of a **sequence of fact sheets** from a spreadsheet
- AutoClips: Generation of a **video** from data facts
- Cinecubes: Generation of a **slideshow** from an OLAP query

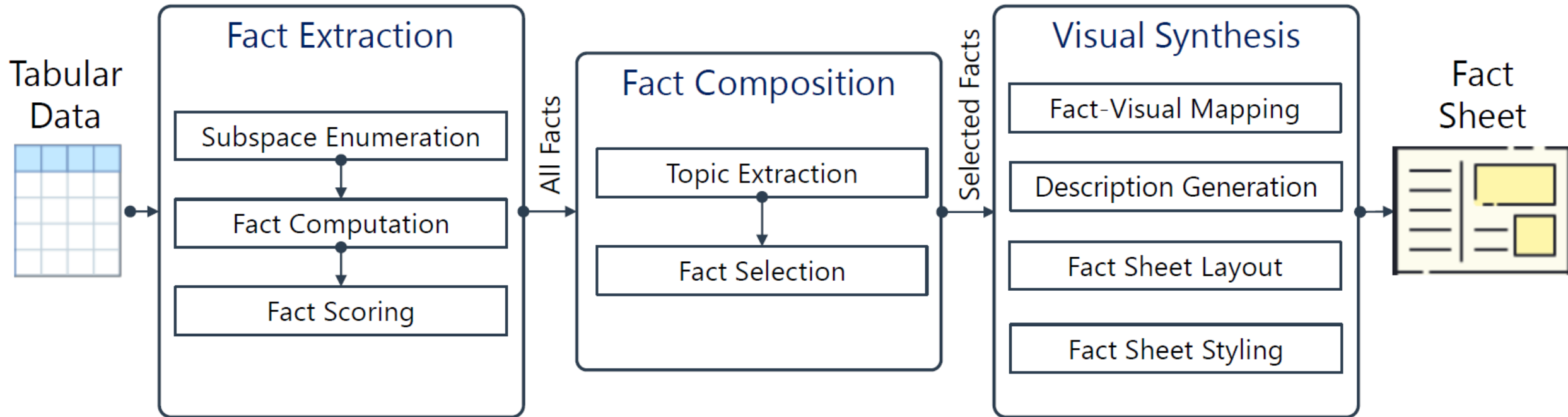
 Y. Wang, Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, D. Zhang: “DataShot: Automatic Generation of Fact Sheets from Tabular Data”, TVCG 26:1, 2020.

 D. Shi, X. Xu, F. Sun, Y. Shi, N. Cao: Calliope: “Automatic visual data story generation from a spreadsheet”, TVCG 27:2, 2021.

 D. Shi, F. Sun, X. Xu, X. Lan, D. Gotz, and N. Cao: “AutoClips: An Automatic Approach to Video Generation from Data Facts”, Comput. Graph. Forum 40:3, 2021.

 D. Gkesoulis, P. Vassiliadis, P. Manousis: “Cinecubes: Aiding data workers gain insights from OLAP queries”, Inf. Syst. 53, 2015.

# DataShot pipeline



Y. Wang, Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, D. Zhang: "DataShot: Automatic Generation of Fact Sheets from Tabular Data", TVCG 26:1, 2020.

# DataShot: *Fact extraction*

*fact := {type, parameters, measures, subject, score}*

*subject := {context, breakdowns, focus}*

## ❑ **Subspace Enumeration:**

- BUC algorithm to explore the data and enumerate the subspaces in a top-down order.

## ❑ **Fact composition:**

- Search for different types of facts (value and aggregation facts for each measure)

## ❑ **Fact scoring:**

- Importance score based on its impact and relevance
  - ❑ Significance score: importance w.r.t. statistical significance (ex: a sharp increase or decrease)
  - ❑ Impact score: how general the data fact sheet is (ratio of records in the subspace/focus)

# DataShot: *Fact composition*

- ❑ **Topic extraction**

- All the facts with the same filters for context or focus

- ❑ **Fact selection**

- Density-based top-k algorithm, balancing fact diversity and fact importance

# DataShot: *Visual synthesis*

## ❑ **Fact-Visual Mapping:**

- Use of a decision tree model (learned from 793 exemplary infographic elements)
- Input: fact types and data types; output: visualization type and options

## ❑ **Description Generation:**

- Templates including subjects, measurements, dimensions, and fact details with additional statistical indicators for each fact type

## ❑ **Fact sheet layout:**

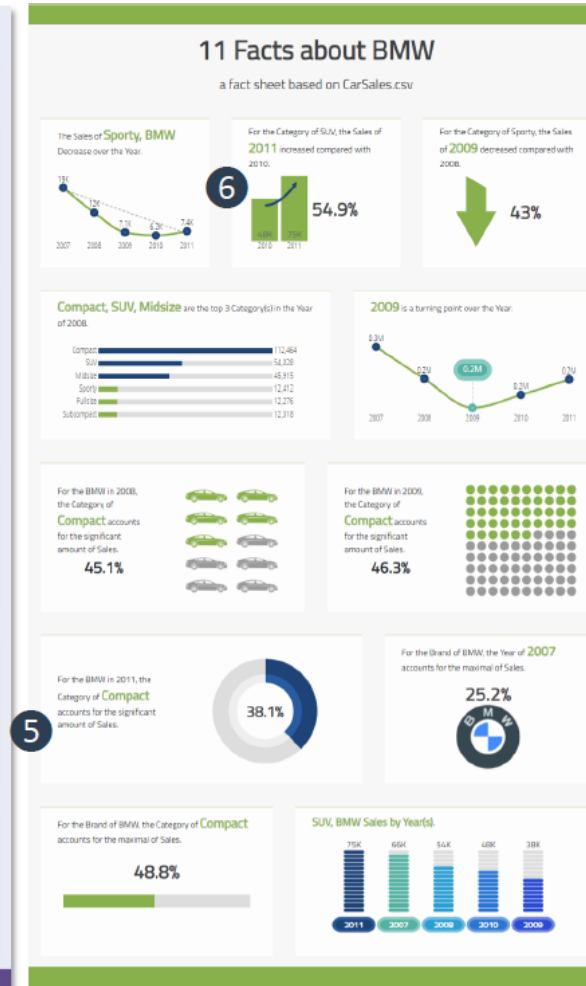
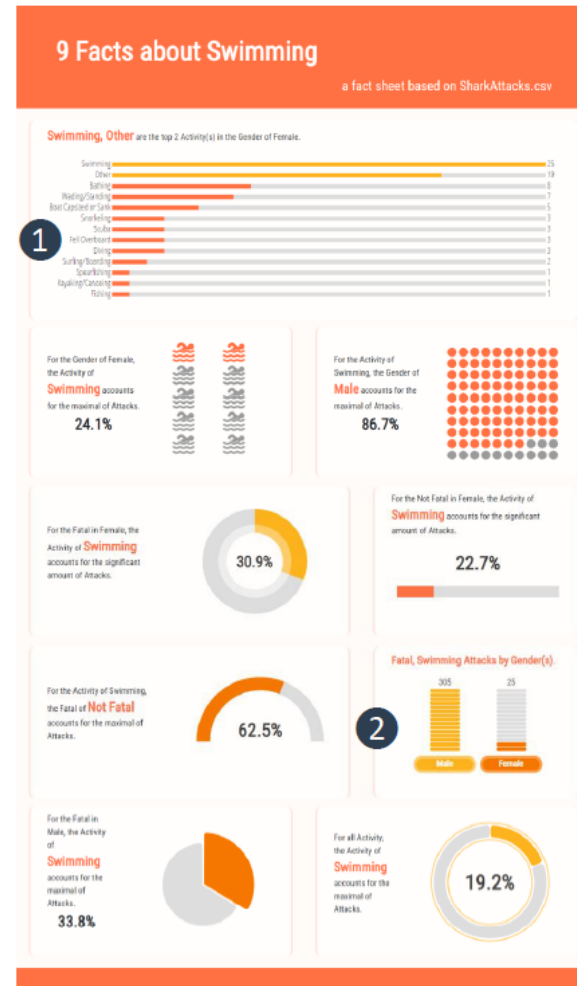
- Tiled layout to arrange data facts
- Fixed height, adjusted width

## ❑ **Fact sheet styling:**

- Supports three styles, including font, title, and embellishments, to diversify the fact sheet presentation



# DataShot: Examples of generated fact sheets



Y. Wang, Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, D. Zhang: "DataShot: Automatic Generation of Fact Sheets from Tabular Data", TVCG 26:1, 2020.

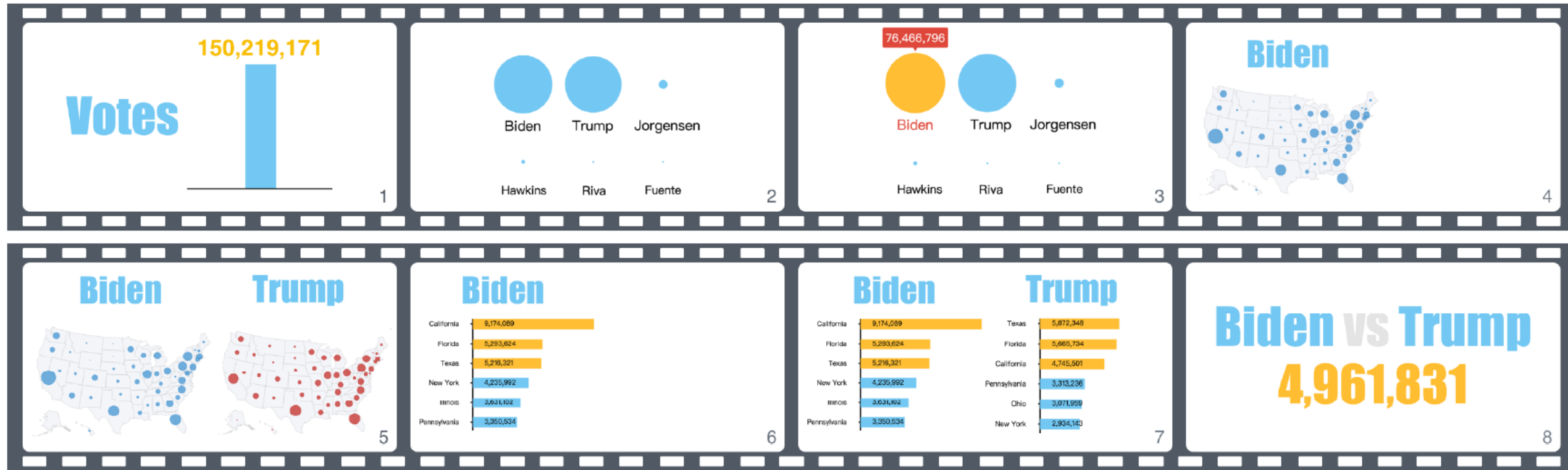
# Calliope: *Example of data story*



D. Shi, X. Xu, F. Sun, Y. Shi, N. Cao: Calliope: “Automatic visual data story generation from a spreadsheet”, TVCG 27:2, 2021.



# AutoClips: *Example of video story*



 D. Shi, F. Sun, X. Xu, X. Lan, D. Gotz, and N. Cao: "AutoClips: An Automatic Approach to Video Generation from Data Facts", Comput. Graph. Forum 40:3, 2021.

# Structure challenge

- ❑ **All automatic generation tools use linear structure**
- ❑ **There is place for**
  - Interactivity (e.g. Martini glass)
  - Including the audience in the loop

# Perspectives

# In summary

- ❑ **Some answers were brought to the 2015 tutorial's perspectives**
- ❑ **In terms of:**
  - Automation
  - Interestingness
  - Languages
  - In-DB engine support

# Next hot topics in EDA and data narration

- ❑ **Integrated approaches to**
  - Explore and analyze datasets
  - And then craft, share, query, reuse data narratives
- ❑ **Explanability**
  - Explanations of insights
  - Data narratives as explanations
- ❑ **Personalization**
  - Leveraging user's preferences, background knowledge, intentions
  - Monitoring their learning curves
  - Personalization of data narratives



# With a little help from...



Marie  
Chagnoux



Thomas  
Devogele



Matteo  
Golfarelli



Nicolas  
Labroche



Stefano  
Rizzi



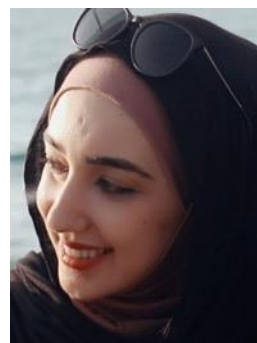
Raphaël  
da Silva



Panos  
Vassiliadis



Alexandre  
Chanson



Faten  
El Outa



Matteo  
Francia



Lucile  
Jacquemart



Raymond  
Ondzigue Mbenga

*Join-us we are hiring!*



Patrick.Marcel@univ-tours.fr  
Veronika.Peralta@univ-tours.fr