

1 VISUALISATION DESIGN AND TRANSFORMATION PROCESS

The design process to create the pairs of visualisations for this study followed the *Data Storytelling Framework* developed by Zdanovic et al. [9] and the visual enhancements described in the data storytelling and visualisation literature [2, 5–8]. Zdanovic et al. [9]’s framework is divided into three distinct phases – 1) *Explore the data*, 2) *Craft the visualisation*, and 3) *Tell the story* -- each with individual steps that are necessary to guide the visualisation creation process. While generating conventional visualisations generally involves going through the first two phases – data exploration and visualisation creation -- generating data stories extends this process by performing a more nuanced crafting of the visualisation in the second phase and incorporating storytelling elements in the third phase, enriching the narrative and enhancing the communication of insights.

1.1 Phase 1 – Explore the Data.

As suggested by the names of the first steps of the framework (*1-understand the context* and *2-aggregate and summarise data*), this phase involves gaining a deep understanding of the dataset, and organising and summarising the data in a way that highlights key patterns and insights, preparing it for effective visual representation.

Implementation. These steps had already been completed by Herre et al. [3], who explored the data using conventional visualisations and provided a report in *Our World in Data*. While key insights were extracted, they were presented in written format rather than integrated into the visualisations themselves.

1.2 Phase 2 – Craft the Visualisation.

Zdanovic et al. [9] stated that the next step – *3-eliminate clutter* –, which is part of the second phase, should differ between conventional and DS-enhanced visualisations. For the conventional visualisation, exploratory features discussed by Zdanovic et al. [9] – such as axis labels, gridlines, and the legend – were added where necessary to allow participants to read and explore the graphs. Conversely, for the DS-enhanced visualisations the idea behind decluttering is removing or de-emphasising aspects of the visualisation that do not contribute to the data story [5, 8]. The fourth step of this process was *4-applying aesthetics* [9], which involves refining the visualisation through deliberate choices of colours, fonts, shapes, and layouts to enhance both clarity and engagement. For the case of DS-enhanced visualisations, this step is crucial to ensure that the visualisation not only captures attention but also communicates the data and key insights in a clear, effective, and accessible manner.

Implementation. The original analysis reported in *Our World in Data* by Herre et al. [3] provided a conventional visualisation for each of the four topics outlined. These conventional visualisations were recreated and modified to ensure they aligned with the same intention, the extent of decluttering and aesthetics as used in *Our World in Data*. This included using the same colour palette and text as the original visualisations. For the topics of *Colonisation* and *State Capacity*, two line charts were created, and for the topics of *Territory Control* and *Tax as a percentage of GDP*, two maps were created. Examples of the conventional visualisations for the line chart and the map can be seen in Figure 1-left and Figure 2-left.

For each of these, an *explanatory* visualisation that utilises the DS-enhanced elements discussed by Knaflc [5] and Zdanovic et al. [9] was crafted to form a pair of conventional and DS-enhanced visualisations for the same topic. For the DS-enhanced visualisations, pre-attentive attributes – such as colour – were carefully leveraged in such a way that important data points were highlighted. Specifically, for the map visualisations we used the suggestion from Cairo [1] of using a mix of subdued grey hues with bright colours to make data points of interest stand out.

As a result of this phase, we crafted four pairs of visualisations:

- A-CV and A-DS (Colonisation)
- B-CV and B-DS (Territory Control)
- C-CV and C-DS (State Capacity)
- D-CV and D-DS (Tax Revenue)

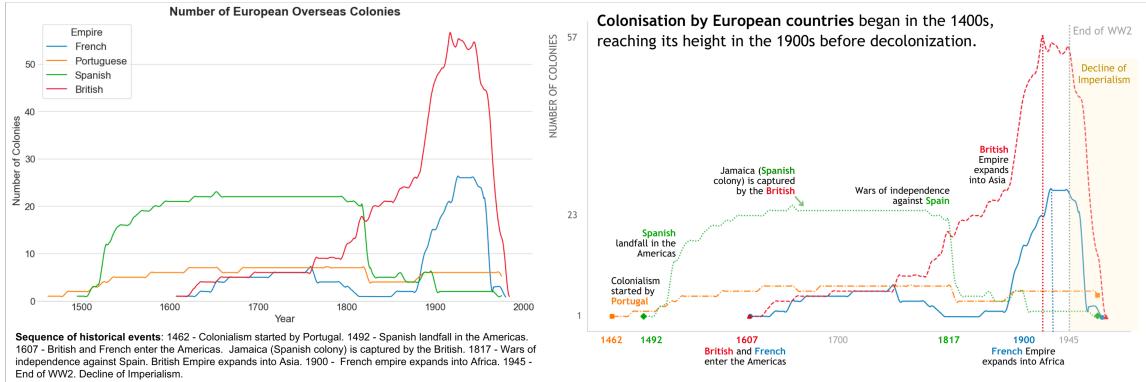


Fig. 1. Visualisation A based on the topic of Colonialism. The pair consists of the conventional visualisation A-CV [left] and the DS-enhanced visualisation A-DS [right].

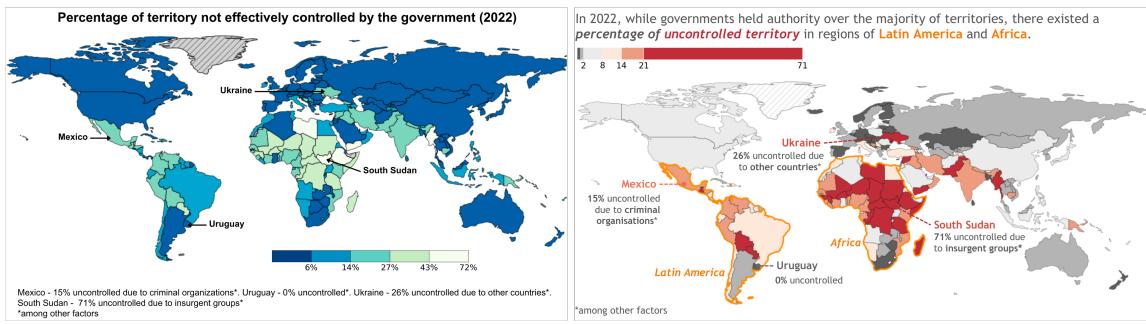


Fig. 2. Visualisation B based on the topic of Territory Control. The pair consists of the conventional visualisation B-CV [left] and the DS-enhanced visualisation B-DS [right].

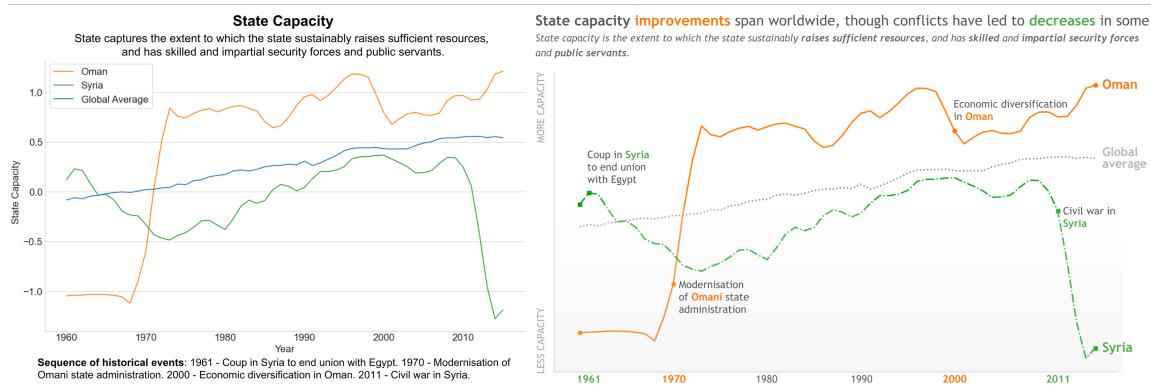


Fig. 3. Visualisation C based on the topic of State Capacity. The pair consists of the conventional visualisation C-CV [left] and the DS-enhanced visualisation C-DS [right].

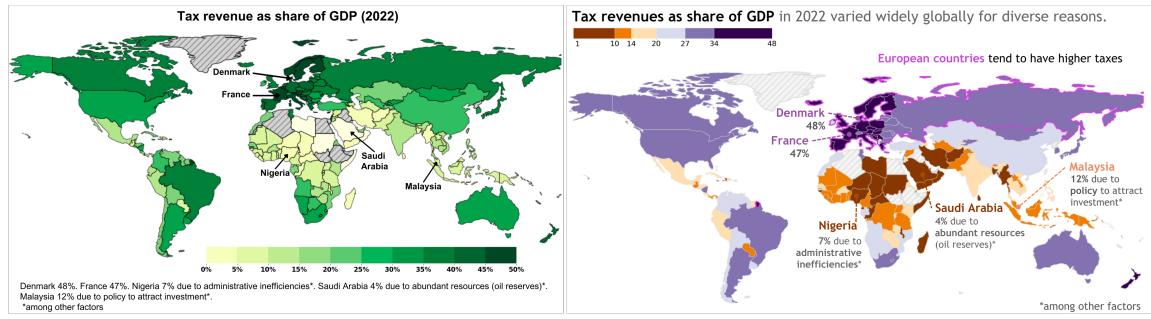


Fig. 4. Visualisation D based on the topic of Territory Control. The pair consists of the conventional visualisation D-CV [left] and the DS-enhanced visualisation D-DS [right].

1.3 Phase 3 – Telling the Story.

The final step to finalise the DS-enhanced visualisations is *5-facilitating the narrative*. This involves the addition of textual elements to explicitly communicate key insights or provide additional contextual information about the data [5].

Implementation. In the DS-enhanced visualisations, text from Herre et al. [3] was adapted to the form of annotations and explanatory titles. Using the suggestions by Kim et al. [4] and Stokes et al. [7], annotations were carefully placed to guide the audience towards the data points they were describing. Pre-attentive attributes for text, such as bolding particular terms or using colour to highlight to further draw attention towards key terms or data-points [5]. Additionally, a footnote was added to the bottom of the conventional visualisations to provide participants with the same contextual information, ensuring that differences in responses were due to how the information was presented (i.e., DS-enhanced vs. conventional) rather than the amount of information provided. Examples of the DS-enhanced visualisations for the line chart and the map can be seen in Figure 1-right and Figure 2-right.

1.4 Iterative process

Phases 2 and 3 of the process described above were carried out over four iterations. In the first iteration, two researchers applied [9]'s framework to independently generate initial drafts of the four visualisation pairs. During the second iteration, these visualisations were compared, and the researchers discussed and reached a consensus on the key design features to be incorporated. In the third iteration, two additional researchers with experience in data storytelling reviewed the visualisations, providing input on patterns, colour schemes, and text placement to ensure alignment with storytelling principles and to maximise fairness in the comparison between conventional and DS-enhanced visualisations. For instance, this led to the inclusion of a footnote in the conventional visualisations including all the contextual information of the annotations in the DS-enhanced visualisation as part of step 5 (*5-facilitating the narrative*). In the final iteration, feedback from two non-specialists was incorporated to improve aesthetic clarity and ensure accessibility for a broader audience. After these four iterations, the final designs for each of the four visualisation pairs were agreed upon by the entire research team. All of the resulting visualisations are available at [this link](#).

REFERENCES

- [1] Alberto Cairo. 2012. *The Functional Art: An introduction to information graphics and visualization* (1st ed.). New Riders Publishing, USA.
- [2] Vanessa Echeverria, Roberto Martinez-Maldonado, Simon Buckingham Shum, Katherine Chiluiza, Roger Granda, and Cristina Conati. 2018. Exploratory versus Explanatory Visual Learning Analytics: Driving Teachers' Attention through Educational Data Storytelling. *Journal of Learning Analytics* 5, 3 (2018), 73–97. <https://doi.org/10.18608/jla.2018.53.6>
- [3] Bastian Herre, Pablo Arriagada, and Max Roser. 2023. State Capacity. *Our World in Data* (2023). <https://ourworldindata.org/state-capacity>.
- [4] Dae Hyun Kim, Vidya Setlur, and Maneesh Agrawala. 2021. Towards Understanding How Readers Integrate Charts and Captions: A Case Study with Line Charts. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 610, 11 pages. <https://doi.org/10.1145/3411764.3445443>
- [5] Cole Nussbaumer Knaflic. 2015. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- [6] Lindy Ryan. 2016. *The Visual Imperative: Creating a Visual Culture of Data Discovery*. Elsevier Inc., Cambridge, MA.
- [7] Chase Stokes, Vidya Setlur, Bridget Cogley, Arvind Satyanarayan, and Marti A. Hearst. 2023. Striking a Balance: Reader Takeaways and Preferences when Integrating Text and Charts. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (Jan 2023), 1233–1243. <https://doi.org/10.1109/TVCG.2022.3209383>
- [8] Edward R Tufte. 2001. *The visual display of quantitative information*. Vol. 2. Graphics press Cheshire, CT. <https://doi.org/10.4135/9781071812082.n670>
- [9] Dominyk Zdanovic, Tanja Julie Lembcke, and Toine Bogers. 2022. The Influence of Data Storytelling on the Ability to Recall Information. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval* (Regensburg, Germany) (CHIIR '22). ACM, New York, NY, USA, 67–77. <https://doi.org/10.1145/3498366.3505755>