

Introduction to Data Science Workshop Materials

Dr Rosie Ridgway

1. Statistics Data and Information

- ? **What are data and information** in 'data science'?

- ? **What stages of information processing** convert raw data into useful information?

- ? What are **official statistics** – what **examples** are offered?

Task: Discuss what might make official statistics more or less reliable? (What barriers exist to collection, obstacles to reporting, measurement problems)

Additional materials to explore: Census Data

- **India Census**
<https://censusindia.gov.in/census.website/data/census-tables>
- **USA Census**
<https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-main.html>
- **UK Census**
<https://www.ons.gov.uk/census>
- **On Open Data:**
Governance lab at NYU thegovlab.org open data project
odimpact.org/periodic-table.html

2. Simple data Visualisation: Population Pyramids

? **What are simple features** of population pyramids?

Task: Go to populationpyramid.net

Explore the tool “Population Pyramids of the World from 1950 to 2100” for different countries.

? **What can we tell** from this data visualisation?

Task: Pick two pyramids (e.g. from different countries or years) to compare. Tell a story about this with your group.

? **Which pyramids** did you pick?

? What are the **features of the two pyramids**?

? What does the **visualisation tell us**?

3. Data Visualisations from simple to complex

? What are some **principles for data visualisation**?

Task: Go to [our world in data](https://ourworldindata.org/age-structure) (<https://ourworldindata.org/age-structure>)

- Look at the population **Age Structure** data.
- **Explore one** of the visualisations we haven't talked about.
- Tell a story about the data using the data visualisation.
- Talk about it with your group.

Task Feedback:

? Which visualisations did you choose?

? Was it useful?

? Was it usable?

? Was it intuitive?

? What were the strengths and limitations of the visualisation?

? Could you tell a story about the data based on the data visualisation?

? What stories did you tell?