Data Visualisation
## 1. Course Summary

El jadid Sara <eljadidsara@gmail.com>

Brain refreshment

❑ The science of visualisation: Humans are visual creatures

❑ Visualisation makes data accessible

❑ The four pillars of data visualisation

❑ Choosing the right chart: know your data

❑ Tools and tips: best practices

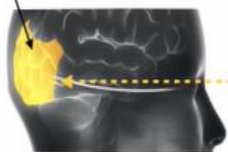❑ The programming part to visualise data

Our brain is a pattern-detecting machine

We are extremely good at detecting **patterns** and **pattern violations**:

- trends
- gaps
- outliers
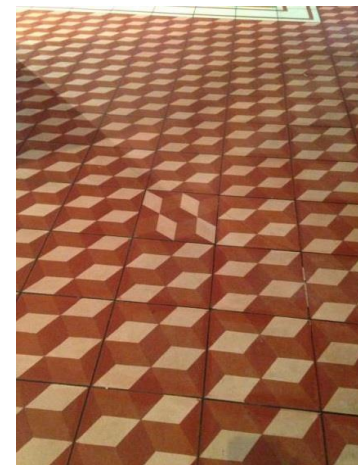
Visualisation makes data accessible

- Show the data
- Induce thinking about the substance
- Avoid distorting what the data has to say
- Present many numbers in a small space
- Make large data sets coherent
- Encourage the eye to make comparisons
- Reveal data at several levels of detail

4

1. Purpose - why this visualization

For the creators:

- **Why** am I doing this visualization?
- Who is it for?
- What do they need to understand?
- What actions do you need to enable?
- How it will be consumed?
- What is the most important takeaway message?
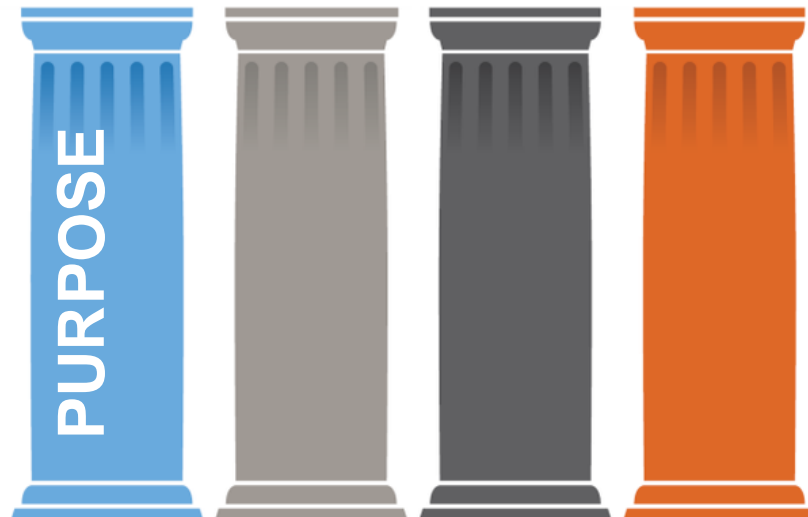
PURPOSE

2. Content - what to visualize

- **What** data matters?
- What **relationships** (in the data) matter?
- Informed by the purpose!
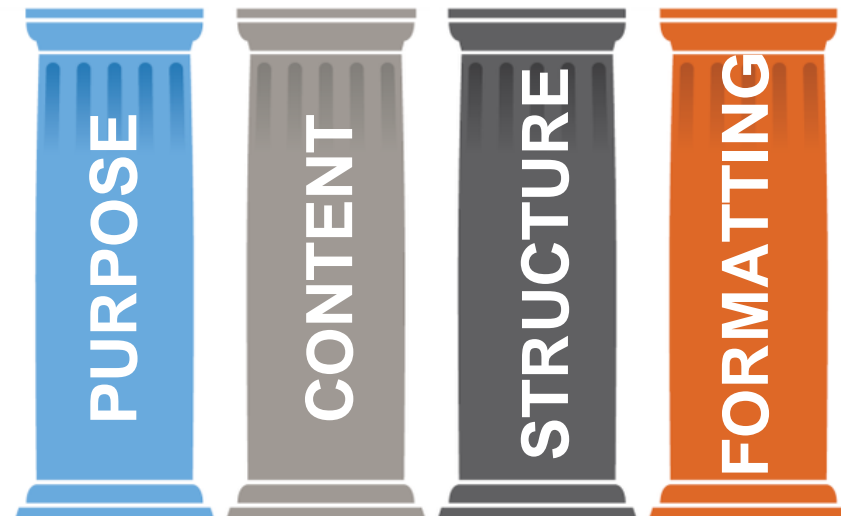- What's *excluded* is as important as what's *included*
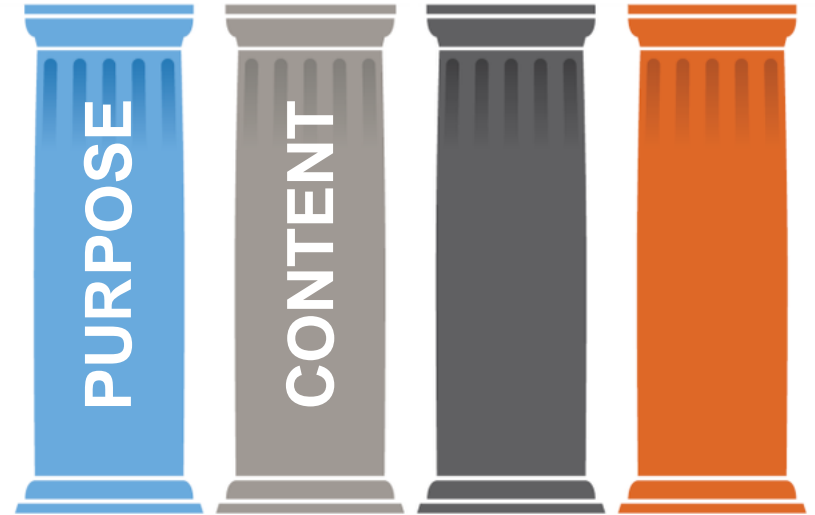
3. Structure - how to visualize it

- **How** can the most important data and relationships be revealed the best?
- **Choose** meaningful layout and axes!
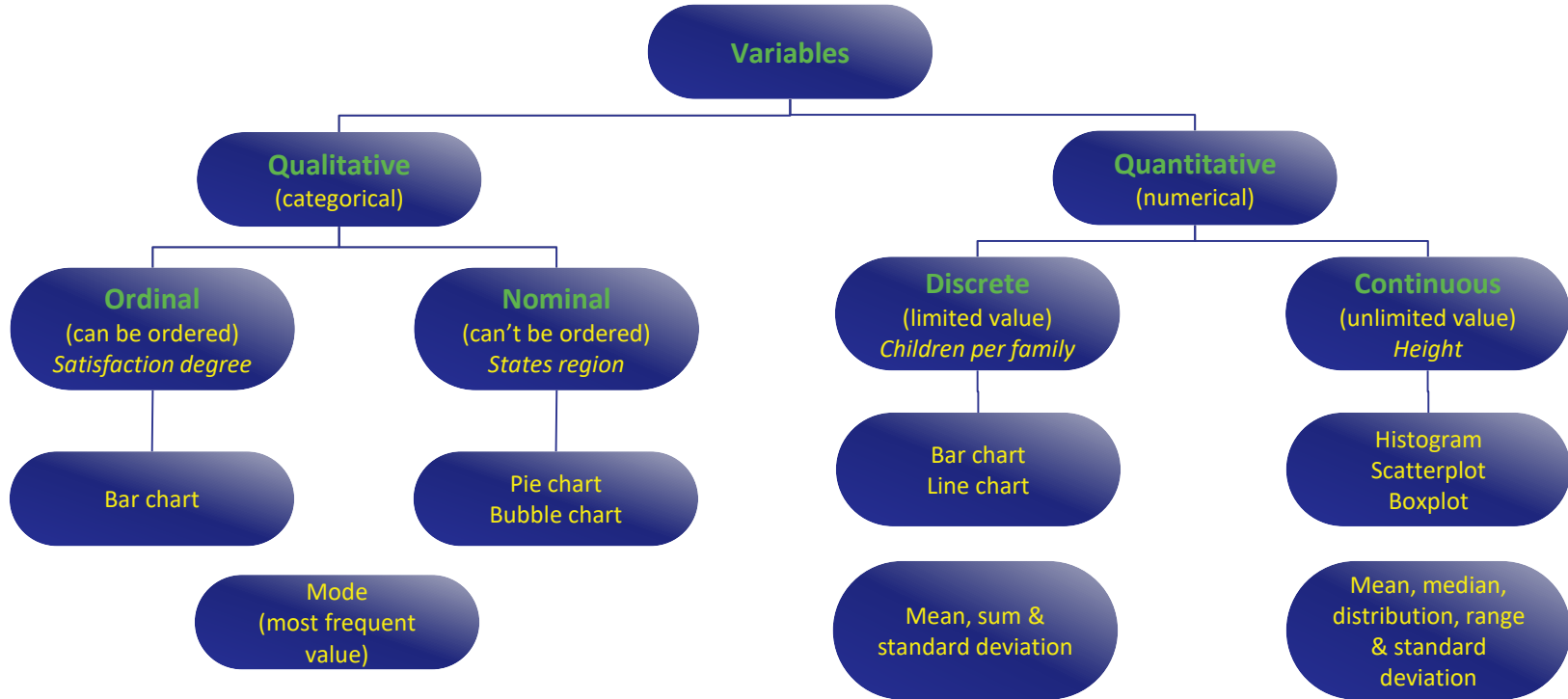- Use both axes (both, not three..)
- Informed by purpose and content!
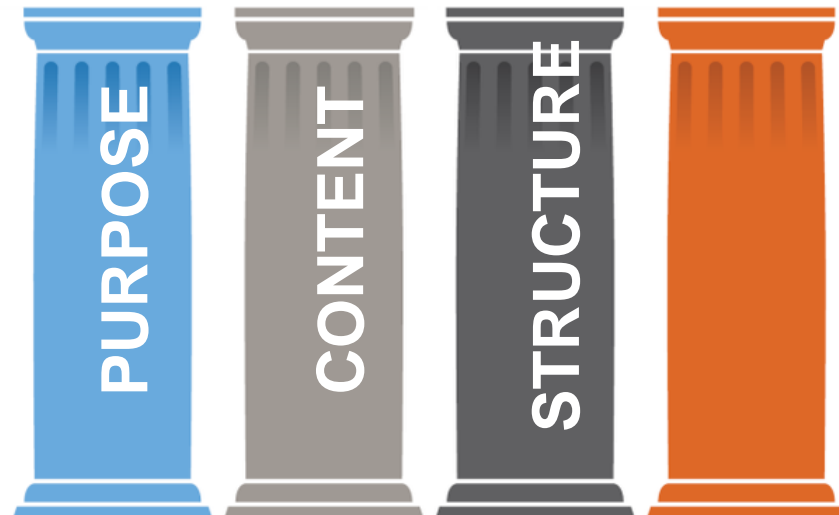
4. Formatting - how to make it appealing

- **How** it should look and feel?
- How will it be consumed?
- Makes data and relationships accessible
- Makes importance visible
- Informed by purpose, content and structure!

Guide to data type and how to graph them

Variables

- **Qualitative** (categorical)
  - **Ordinal** (can be ordered) *Satisfaction degree*
    - Bar chart
    - Mode (most frequent value)
  - **Nominal** (can't be ordered) *States region*
    - Pie chart / Bubble chart
- **Quantitative** (numerical)
  - **Discrete** (limited value) *Children per family*
    - Bar chart / Line chart
    - Mean, sum & standard deviation
  - **Continuous** (unlimited value) *Height*
    - Histogram / Scatterplot / Boxplot
    - Mean, median, distribution, range & standard deviation

Visualisation 1 – El jadid Sara

Do's and don'ts in data design & visualization

- Use one color to present each category. 👍 *DOs*

- Order data sets using logical hierarchy. 👍 *DOs*

- Use high contrast color combinations such as Red/Green or Blue/Yellow. 👎 *DON'Ts*

- Use callouts to highlight important or interesting information. 👍 *DOs*

- Use 3D charts. 👎 *DON'Ts*

- Visualize your data in a way that it's easy for readers to compare values. 👍 *DOs*

- Add chart junk. 👎 *DON'Ts*

- Use more than 6 colors in a single layout. 👎 *DON'Ts*

- Use icons to enhance comprehension. 👍 *DOs*

- Use italic, bold or underline text. 👎 *DON'Ts*

The programming approach to visualize data

- Generate plots from data according to their type (discrete, continuous …)
- Manage plot settings
- Produce plots from data in a data frame
- Modify and customize a plot
- Create complex and fancy plot

You managed to create

- 1D plots: Bar plots for discrete variables
- 1D plots: density plots and boxplots for continuous variables
- 1.5D: Layers & Time series plots
- Plotting a summary
- 2D: Scatterplots
- Advanced customized and fancy plot : Bubble plot and an interactive plot
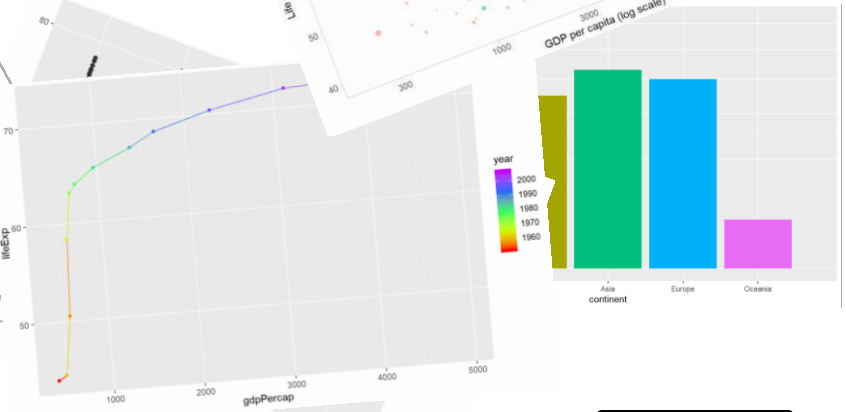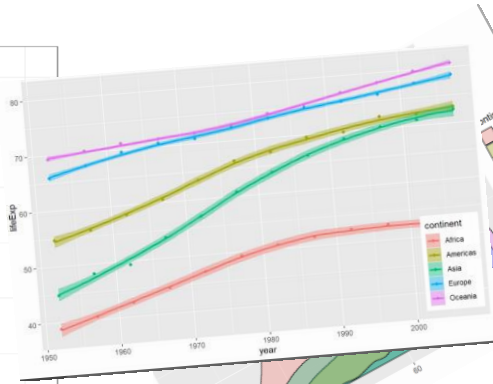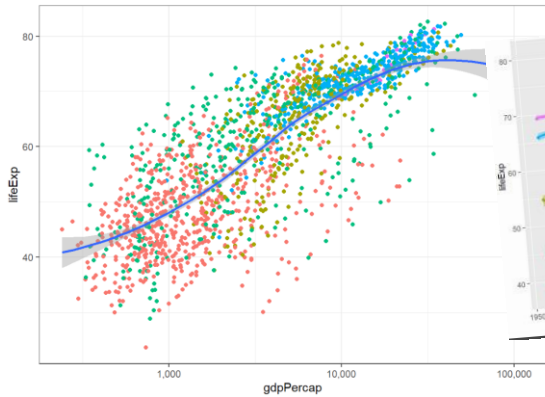
Data Visualisation

A picture is worth a thousand of words

Thank you