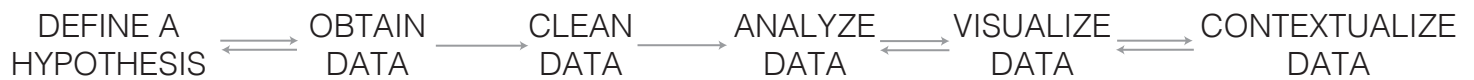


DDJ PATH



SHAPING YOUR IDEA

Come up with a main hypothesis you want to prove/disprove

Illustration vs dataviz (how big is your dataset?)

Ways to generate data story ideas

- look at development over time
- look at geographic distribution patterns
- fact-check statements, follow up on policy goals
- compare/contrast values
- do a time-event-analysis
- go for a network analysis
- borrow from design thinking: what are different user groups for your topic and which are each group's needs?

A datapoint walks into a bar. Or: Data depicts humans

Who and where (device, situation) is your audience?

Do you want to show details or give an overview?

Check what others have done already!



OBTAINING DATA

Who might have such data

Always trace down original/primary source

How was the data collected?

Why was it collected?

Who collected the data?

What are biases of each of the above points?

Where are gaps in your data, what do they mean?

Which are conclusions the dataset it doesn't allow?

When was the data gathered vs when published?

Look up definitions, metadata

Talk to the people who gathered it!

Find a second source for the same data, or another proxy for the same phenomenon you want to show



WORKING WITH DATA

Write a data diary (sources, problems you encountered/how you solved them, definitions, findings)

Come up with 5 to 8 questions you can ask your dataset, have them guide your analysis (don't drown)

Make a copy of your datafile before you start working

Clear all formatting

Document formulas and changes for your future self to understand (reproducibility, transparency)

Give new columns a meaningful titles

Exploratory analysis: figure out the min, max, average, median of your data, define the lower/upper acceptance limit for outliers (IQR rule)

For surprising results: find alternative explanations, sample back against original data

Explain your findings to both someone new to the topic and someone data-savvy



VISUALIZING DATA

What message do you want to convey with the visual?

Does the visual communicate its message effectively by guiding the user well with its visual hierarchy?

Does the visual allow to easy to compare elements, draw connections, recognize trends or patterns?

Does the visual allow you to draw conclusions without looking at every single element?

Does the text on the visual help to make clear to the user

- what the visual shows (e.g. labels)
- why you want to show it to the user (e.g. headline)

What is a question your user would have, looking at the visual?

Double-check:

- Is there new data since you started working on this?
- Units all correct?
- Labels needed?
- Source (in a way that makes finding the original easy)