How to do data-driven investigations in 12 memes

(even though you're not an investigative journalist)

CAUTION! Before you get started, here are some questions to consider:

- Does your investigation require anonymity?
- Do you need secure/private communication channels?
- Do you need help to secure your data and communications?

If so, check out **securityinabox.org** and/or look for trainings at **cryptoparty.in!**

- Does your investigation come with legal risks?

Consider seeking legal advice!


CHOOSE YOUR TARGET
NOW
memegenerator.net

Step 1: Define the target

Define the topic/target of your investigation: Who or what are you investigating? Narrow it down!

Investigating “global corruption” is probably too broad and would take you a lifetime. Try investigating, for example, the flow of money in your country instead!

MAP THE TARGET AND FIND ALLIES

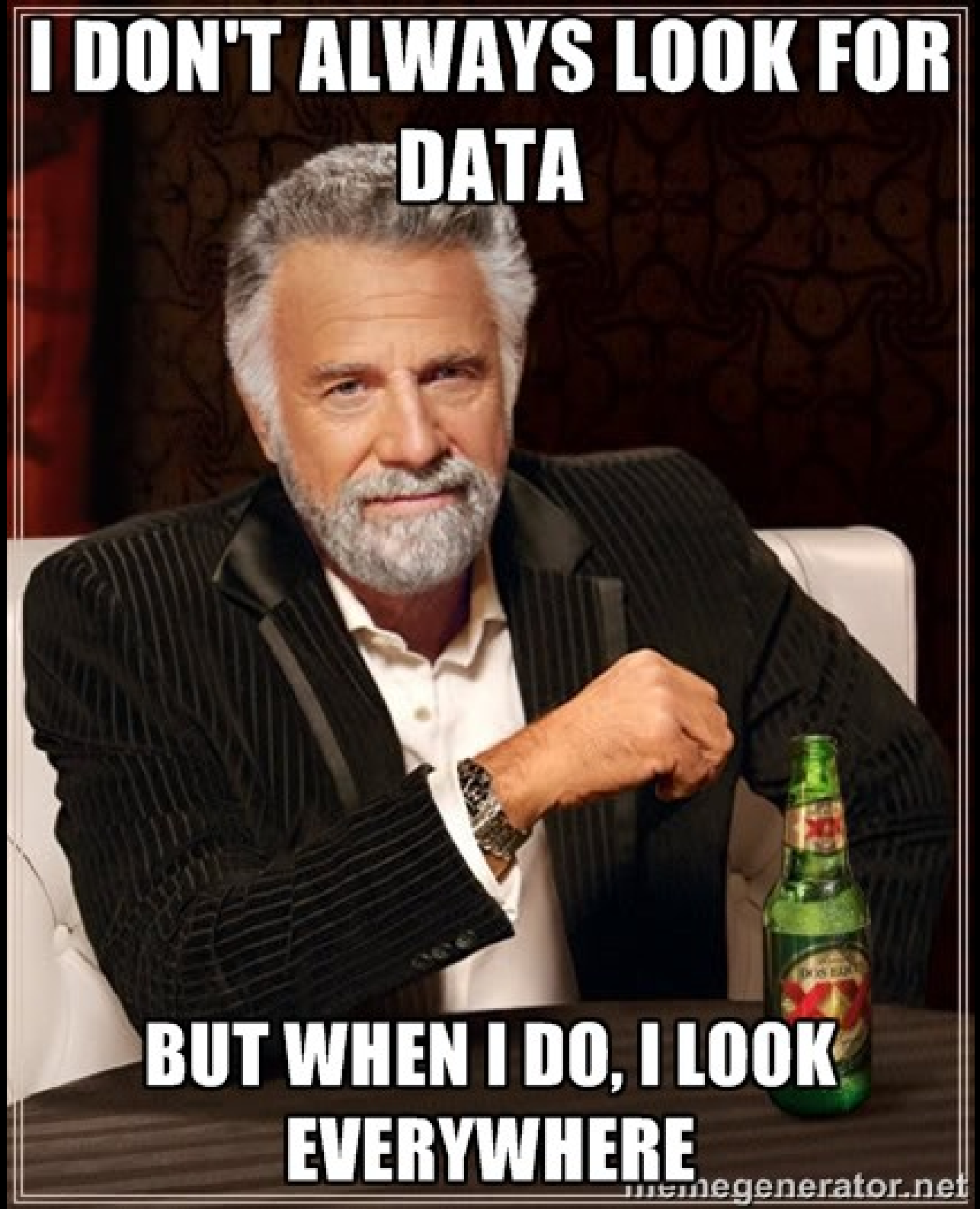
YOU MUST

Step 2: Map the target

Map the target/topic of your investigation:

- Identify actors
- Scope of operations

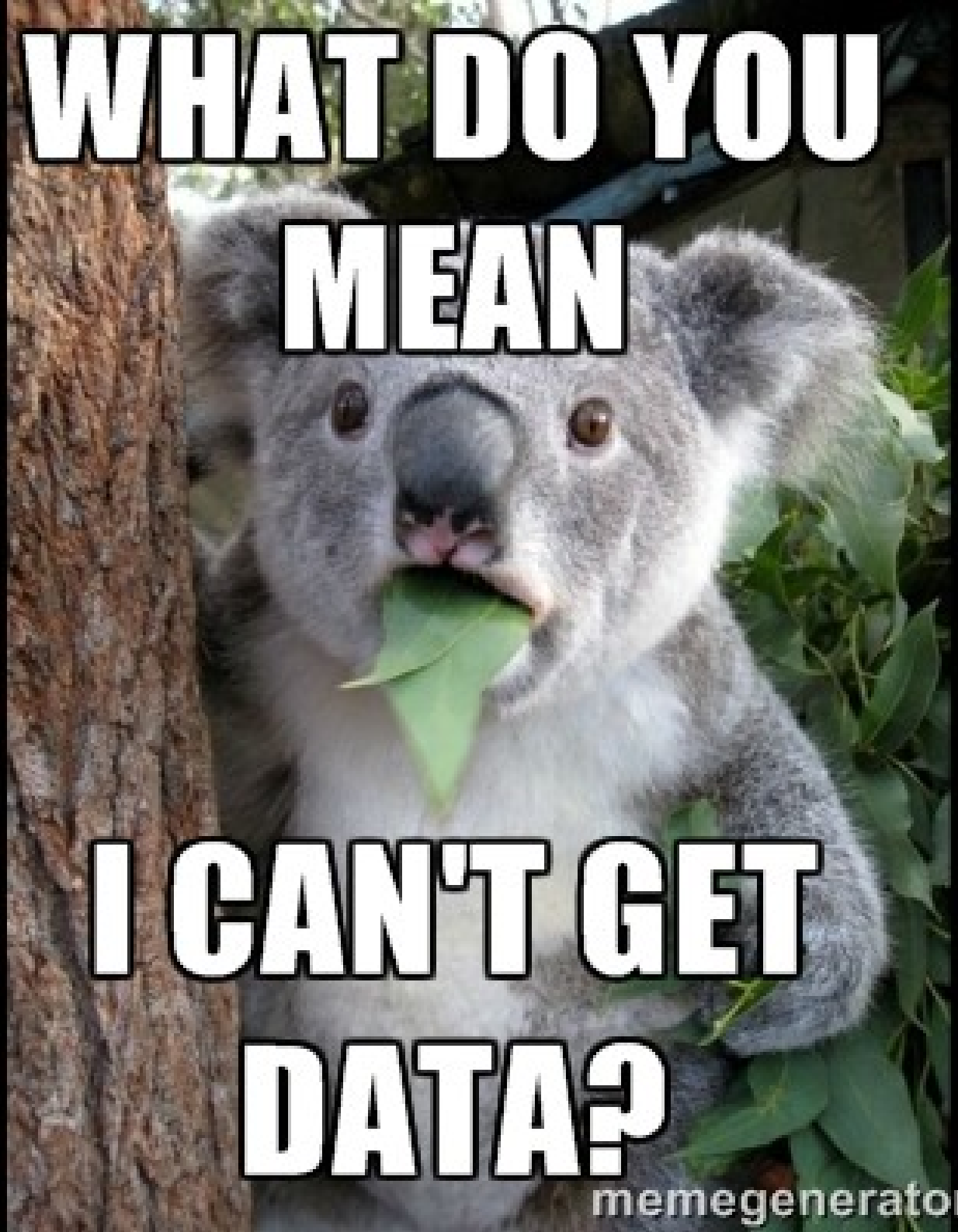
Then map all others working on similar investigations as you - and try to collaborate with them!


I DON'T ALWAYS LOOK FOR DATA
BUT WHEN I DO, I LOOK EVERYWHERE
memegenerator.net

Step 3: Search for data

Depending on your investigation, you can search for data through the following sources:

- Government sources
- Leaked documents
- Open data sources
- Websites
- Contracts
- Research (interviews, surveys, research papers, etc.)


WHAT DO YOU MEAN
I CAN'T GET DATA?
memegenerator.net

Step 4: Access data

Once you have identified the data sources for your investigation, you can gain access to data through the following ways:

- Freedom of Information (FOIA) requests
- Scraping websites
- Data extraction from documents and/or websites
- Interviews
- Surveys
- Libraries / Archives


DATA
COLLECT IT ALL!
memegenerator.net

Step 5: Collect data

Start off by:

- Defining the questions you want to ask
- Based on your questions, define the fields of data collection

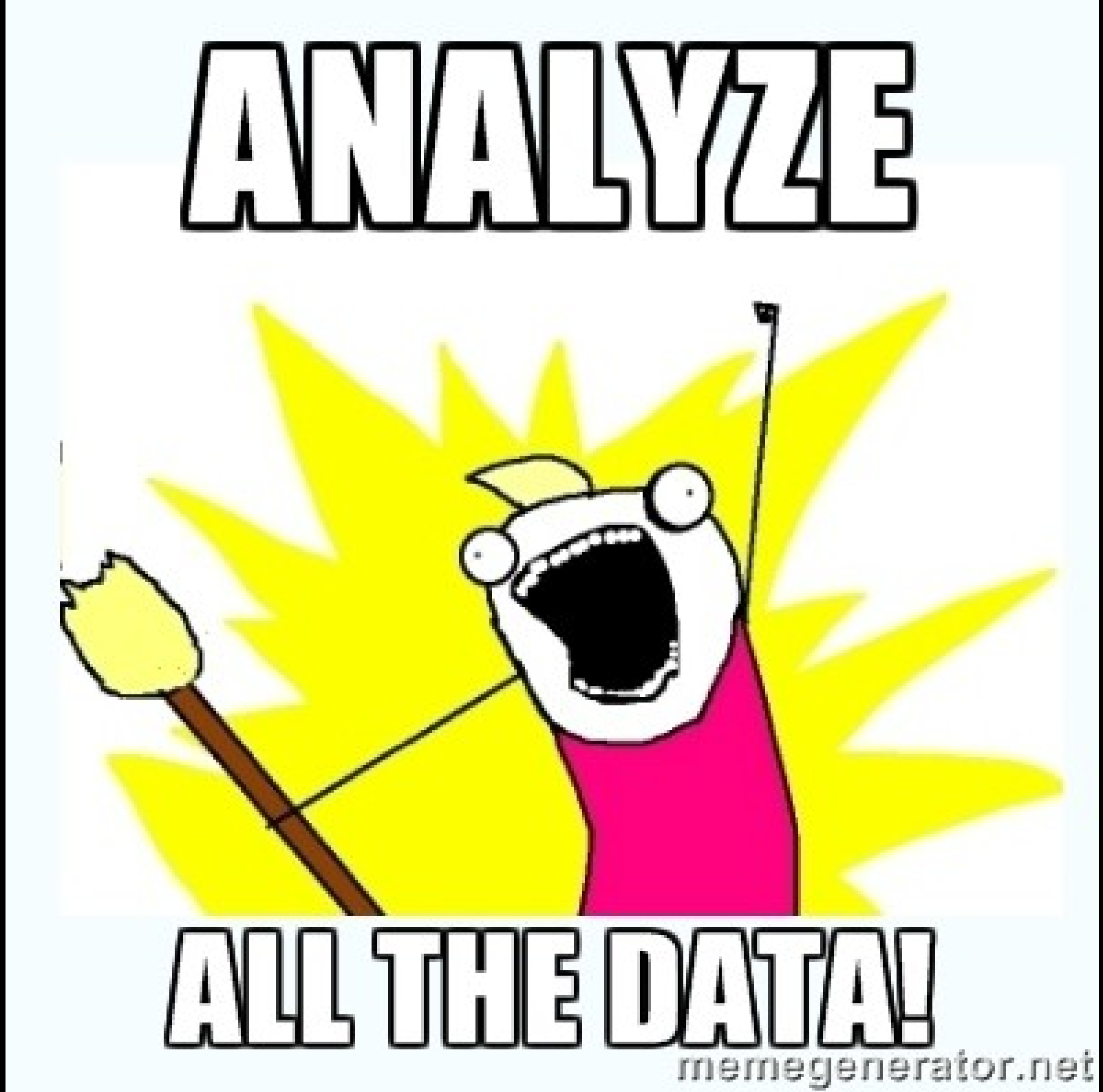
Then collect data through:

- The use of search terms (companies, location, etc.)
- Automation
- Patterns, correlations and clues (later stage of data collection)


ONE DOES NOT
SIMPLY
COLLECT THE DATA
memegenerator.net

Step 6: Process data

- Extract your data from documents
- Make sure you have a standard data format (text cleaning/data normalization)
- Cluster your data
- Make your data machine-readable


ANALYZE
ALL THE DATA!
memegenerator.net

Step 7: Analyze data

- Cross-reference datasets
- Correlate your data
- Match patterns
- Identify clues
- Outline the patterns/clues/questions that could be useful to your investigation
- Re-define your question(s) of investigation!
- Collect more/different data

Data visualizations are useful for the above!


TOO MUCH DATA
NEED TO FIND HIGHLIGHTS
memegenerator

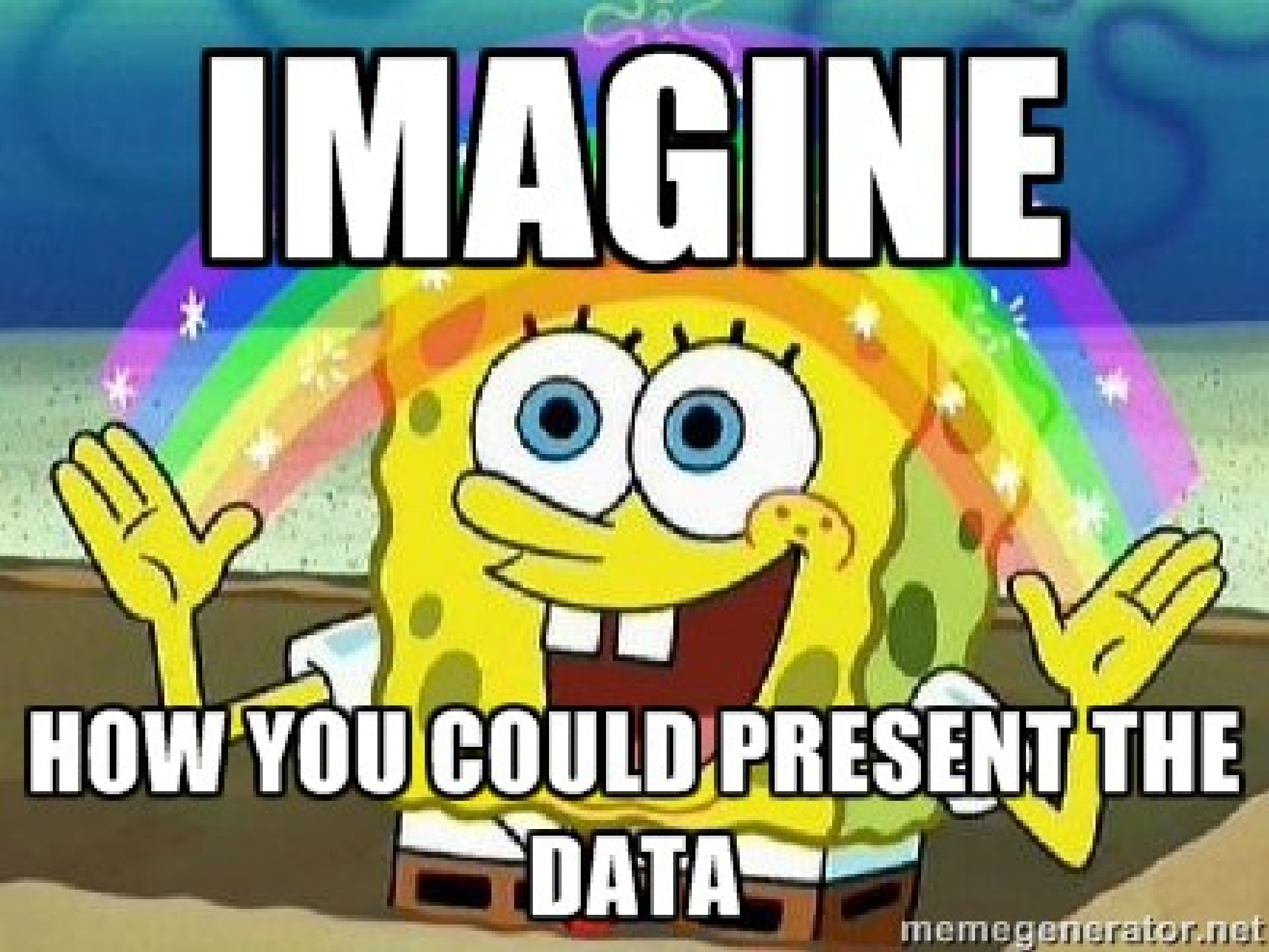
Step 8: Model data

- Identify different types of highlights/trends
- Determine which highlights are essential to the question(s) that you are investigating
- Determine how you want to present your data to the world


DID YOU CHECK
ALL THE FACTS?
memegenerator.net

Step 9: Check facts

- Make sure that your facts are accurate!
- Verify details across multiple sources
- Consider working with a fact-checker


IMAGINE
HOW YOU COULD PRESENT THE DATA
memegenerator.net

Step 10: Tell a story

- Reports, articles, blogs
- Data visualizations
- Film
- Radio
- Animation
- Webdocs
- Comics, memes
- Photos
- Talks/presentations
- Meet-ups

GOT THE DATA

STARTED AN ACTION

Step 11: Take action

Take action based on your data-driven investigation

- Campaigns
- Lobbying
- Legal cases
- Protests
- Tool development
- etc.

Consider collaborating with other groups!


WHAT IF I TOLD YOU
YOU NEED TO START OVER AGAIN
memegenerator.net

Step 12: Evaluate

- Map the outcome of your data-driven investigation and track its impact
- Evaluate whether you would do things differently next time
- You're ready to get started with the next investigation!


Y U NO STARTED INVESTIGATING YET?
memegenerator.net