



Making data meaningful



Data does not create meaning,
people do



The Beginning: Idea



1. A single idea can be the source of many stories.
Which one do you want to tell?

Example: Air pollution

Different paradigms

- a. Political
- b. Social
- c. Human interest
- d. Economic etc.



What is your story?



2. Finding the hidden story

- Talking about things that people overlook
- New interpretations
- Our take – **CLIMATE CHANGE CAN BE THE VALUE ADDITION**



Source it!



- 4. Finding the right sources for your story – which is the relevant data?

Through the use of flowcharts, note-making, mind-mapping



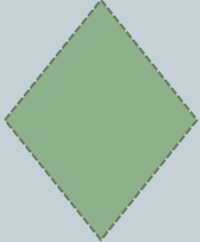
Flowcharts



Central idea/ premise



Logical progression

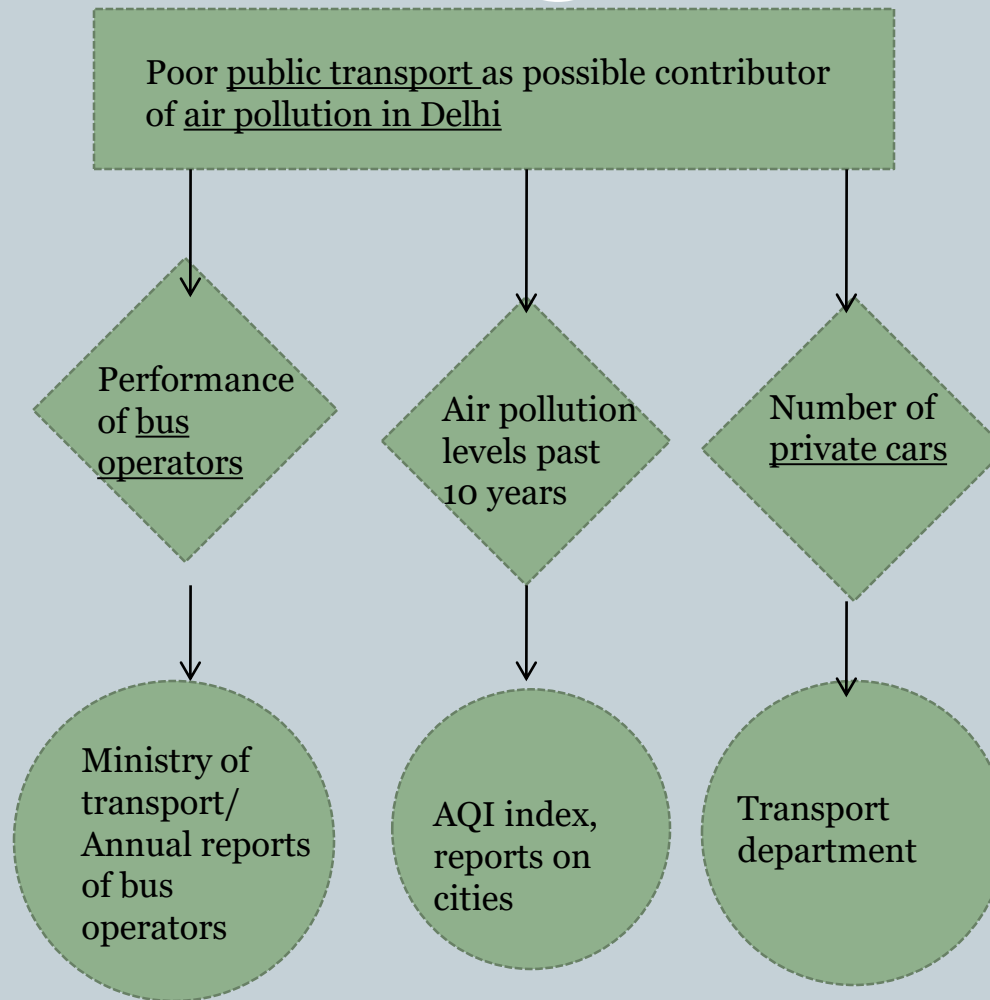


Contextualised data for the story/
What you want the data to say



Data sources

Flowcharts





3 (b) One of the ways of categorising data stories

Outliner stories

Trend stories

Correlation stories



Outliner stories



A value that is different from all the others

- Which city has the highest air pollution levels?
- Why farmers from a particular area are more susceptible to suicides or mortgages

Eg: Rankings



Trend stories



A trend is a pattern through time

- Has the use of public transport declined over the last ten years?
- What is the change in hospitalization charge over the last five years?

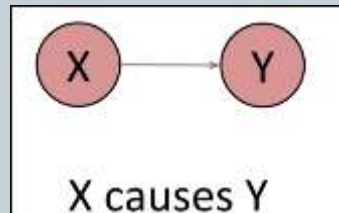


Correlation stories



A correlation is when two variables change together

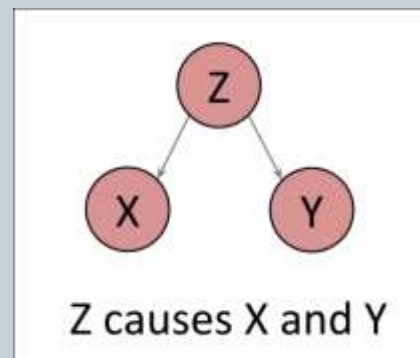
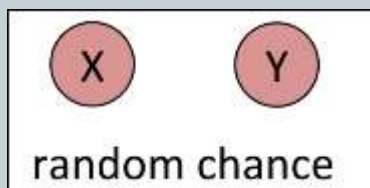
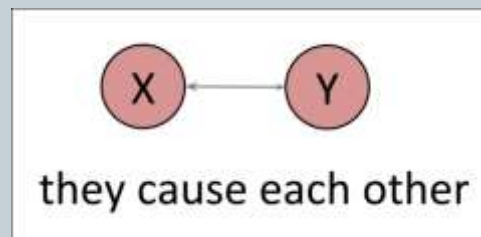
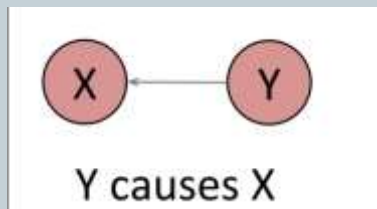
- More smoking causes more cancer.
- People buy more umbrellas when it's raining.



Most popular type of correlation

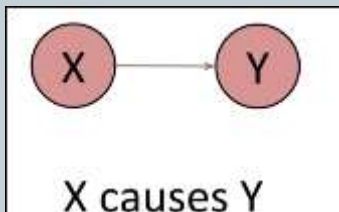
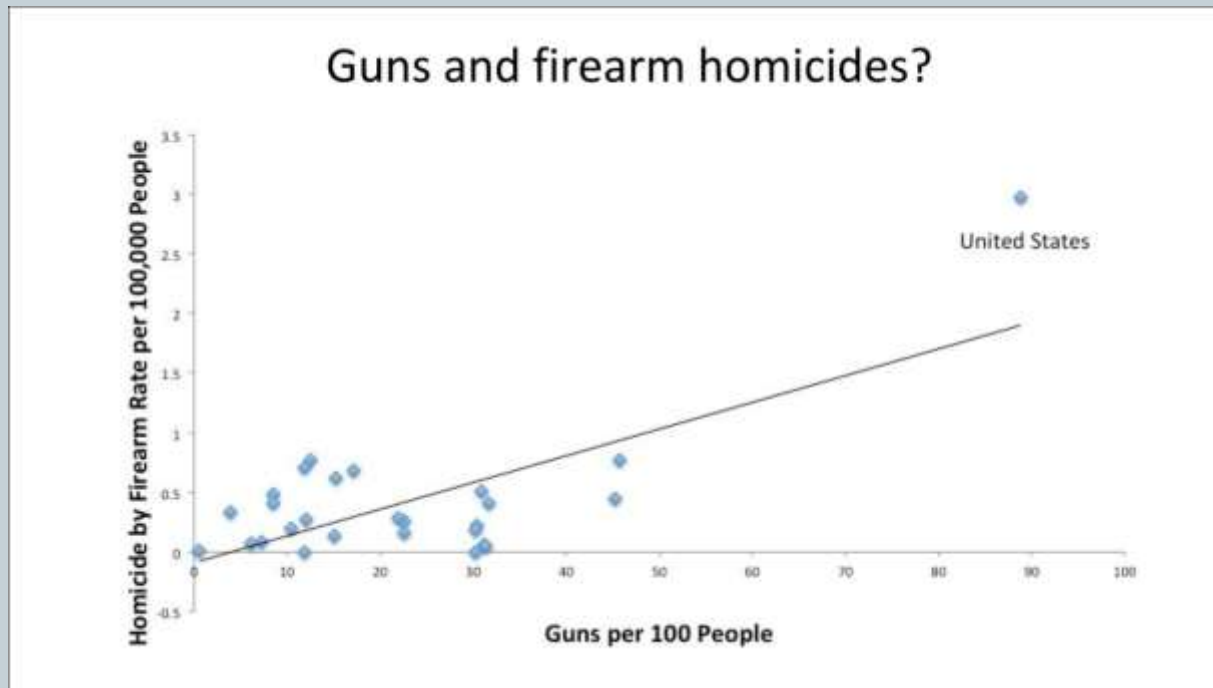


Correlation stories (contd.)



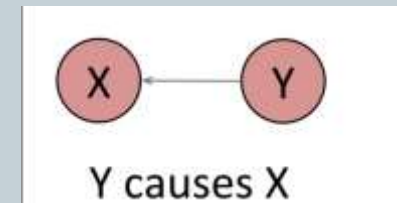


Correlation stories (contd)



If you have a gun, you are likely to use it

If it is a dangerous locality, you will buy a gun

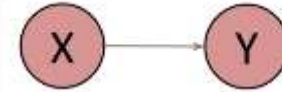
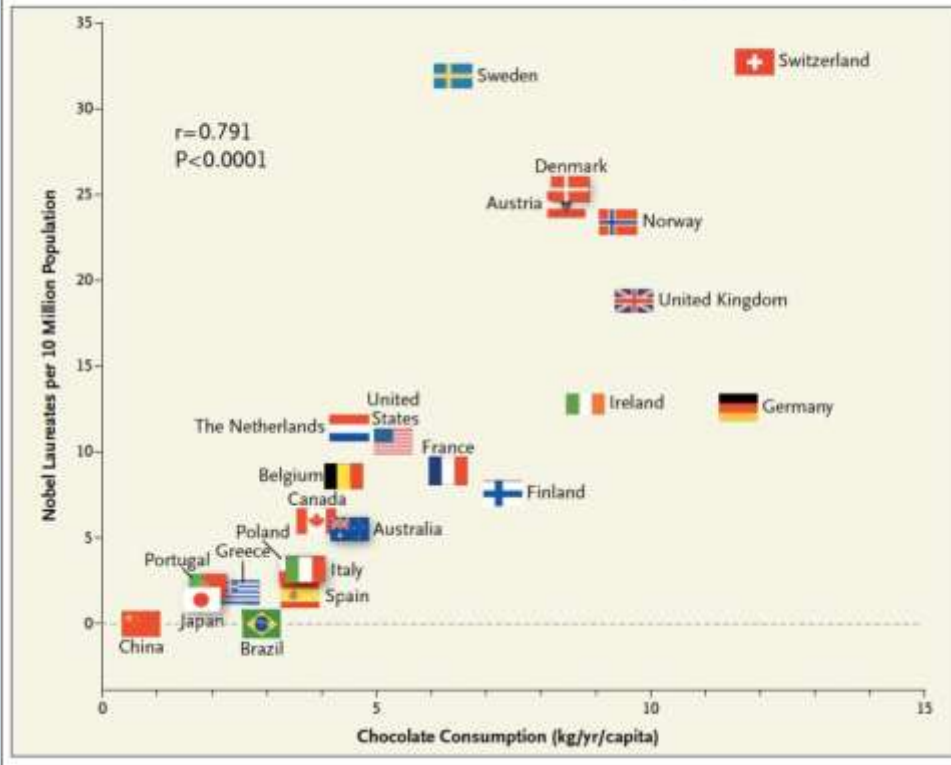




Correlation stories (contd)

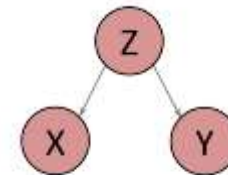


Chocolate and Nobel prizes



X causes Y

Chocolates make
you smarter

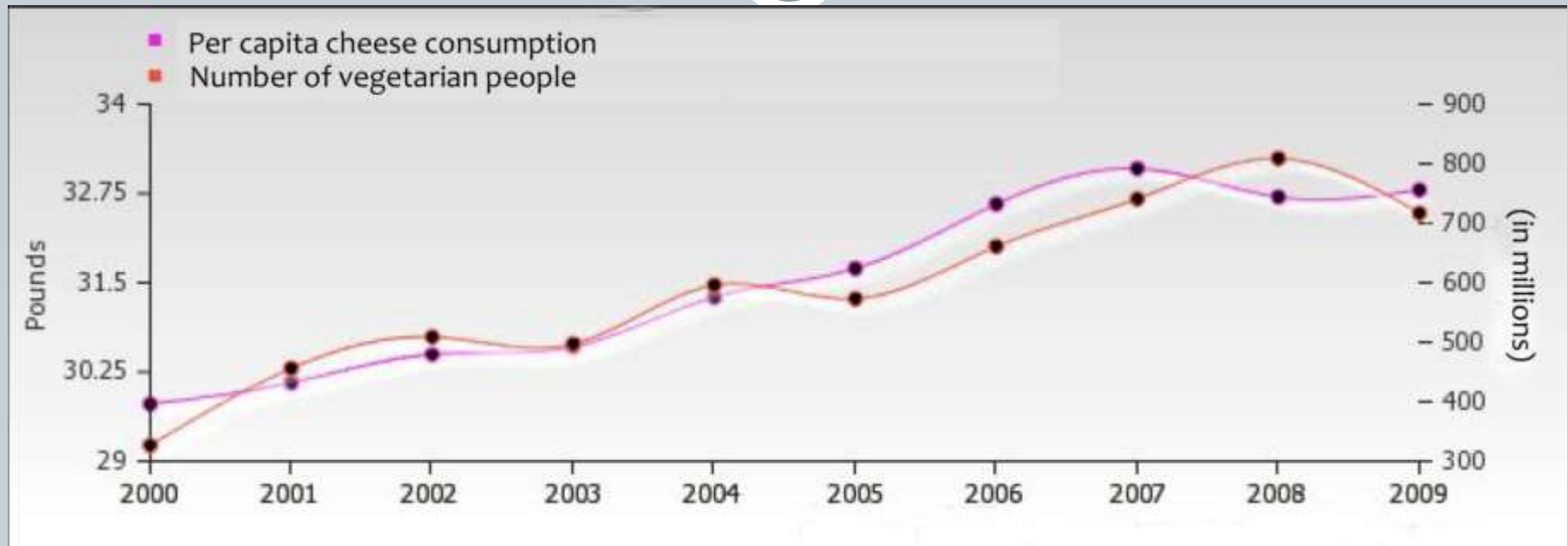


Z causes X and Y

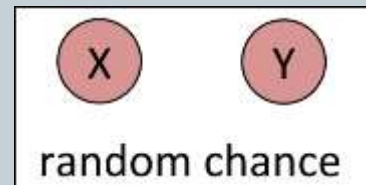
- 1) Higher income makes you eat more chocolate
- 2) Higher incomes can fund better education



Correlation stories (contd)



There is no way to
correlate the to
parameters



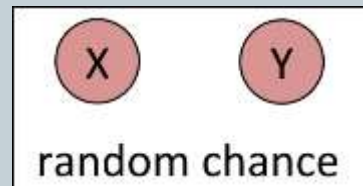


Correlation stories (contd)



The number of drowning cases increases in town X
when the consumption of ice cream increases

There is no way to
correlate the to
parameters





How to smart search on Google



Smart Search



“ ”

Returns results with the exact phrase

Sleep habits of Asians: 1,54,000 results

“Sleep habits of Asians”: 6 results



Smart search



site:

Allows you to search a phrase within a website

"PM 2.5 levels" of major cities globally `site:www.who.int/`
(A search for **`site:downtoearth.org.in`** will work, but
`site: downtoearth.org.in` won't.



Smart search



+ or **-**

Includes or excludes a phrase from search

“PM 2.5 levels” of major cities globally + Lagos
pandas -site:wikipedia.org



Smart search



Add an asterisk as a placeholder for any
unknown or wildcard terms

"a * saved is a * earned"



Smart search



Separate numbers by two periods without spaces to see results that contain numbers in a range

camera \$50..\$100



Smart search



related

Find sites that are similar to a web address you already know

`related:time.com`



Smart search



Or

Find pages that might use one of several words

forest or forestland or tree cover



Smart search



Filetype:



Includes or excludes a phrase from search

PM 2.5 levels of major cities globally filetype:pdf

PM 2.5 levels of major cities globally filetype:xls



Smart search



Google Advanced Search

Find pages with...

all these words:	heha in ho ha	To do this in the search box:	Type the important words: <code>tri-colour rat terrier</code>
this exact word or phrase:			Put exact words in quotes: <code>"rat terrier"</code>
any of these words:			Type OR between all the words you want: <code>miniature OR standard</code>
none of these words:			Put a minus sign just before words that you don't want: <code>-rodent, -"Jack Russell"</code>
numbers ranging from:		to	Put two full stops between the numbers and add a unit of measurement: <code>10...20 kg, £300...£500, 2010...2011</code>

Then narrow your results by...

language:	any language	Find pages in the language that you want.
region:	India	Find pages published in a particular region.
last update:	past 24 hours	Find pages updated within the time that you specify.
site or domain:		Search one site (like: <code>wikipedia.org</code>) or limit your results to a domain like: <code>.edu .org</code> or <code>.gov</code>
terms appearing:	anywhere in the page	Search for terms in the whole page, page title or web address, or links to the page you're looking for.
SafeSearch:	Show most relevant results	Test SafeSearch whether to filter sexually explicit content.
file type:	any format	Find pages in the format that you prefer.
usage rights:	not filtered by licence	Find pages that you are free to use yourself.

Advanced Search



Some more tricks



Pdf to Excel

www.pdf tables.com

Other softwares are
Tabula (free) and Nitro

PM 2.5 levels of major cities globally filetype:pdf



Interview the data



- Why was the data created and is it reliable?
- Who has created this data? (sources are sprouting like mushrooms)
- How was it created? sample size, duration etc
- Just like you would interview a source carefully to make sure that the information is good



Hard truths



- Only a small amount of raw data will be used finally for your story
- You will have to prioritise (choose between) raw data
- Always give a context... even if it 'seems' unnecessary
- Whenever you are using big numbers, try to give a reference for correlation (1,484 km²)
- Places to look for side stories:
Money, operation size, outreach, impact
Court judgements, laws
Case studies, geographical locations



Hard truths (contd)



- Look at percentages
- Look at comparable data to see the aberration
- Mean (average), median (middle value) and mode (most frequent value)
- Always double check your final analysis... especially to ensure that no biases have crept in
- Also try to answer the WHY of the final analysis