

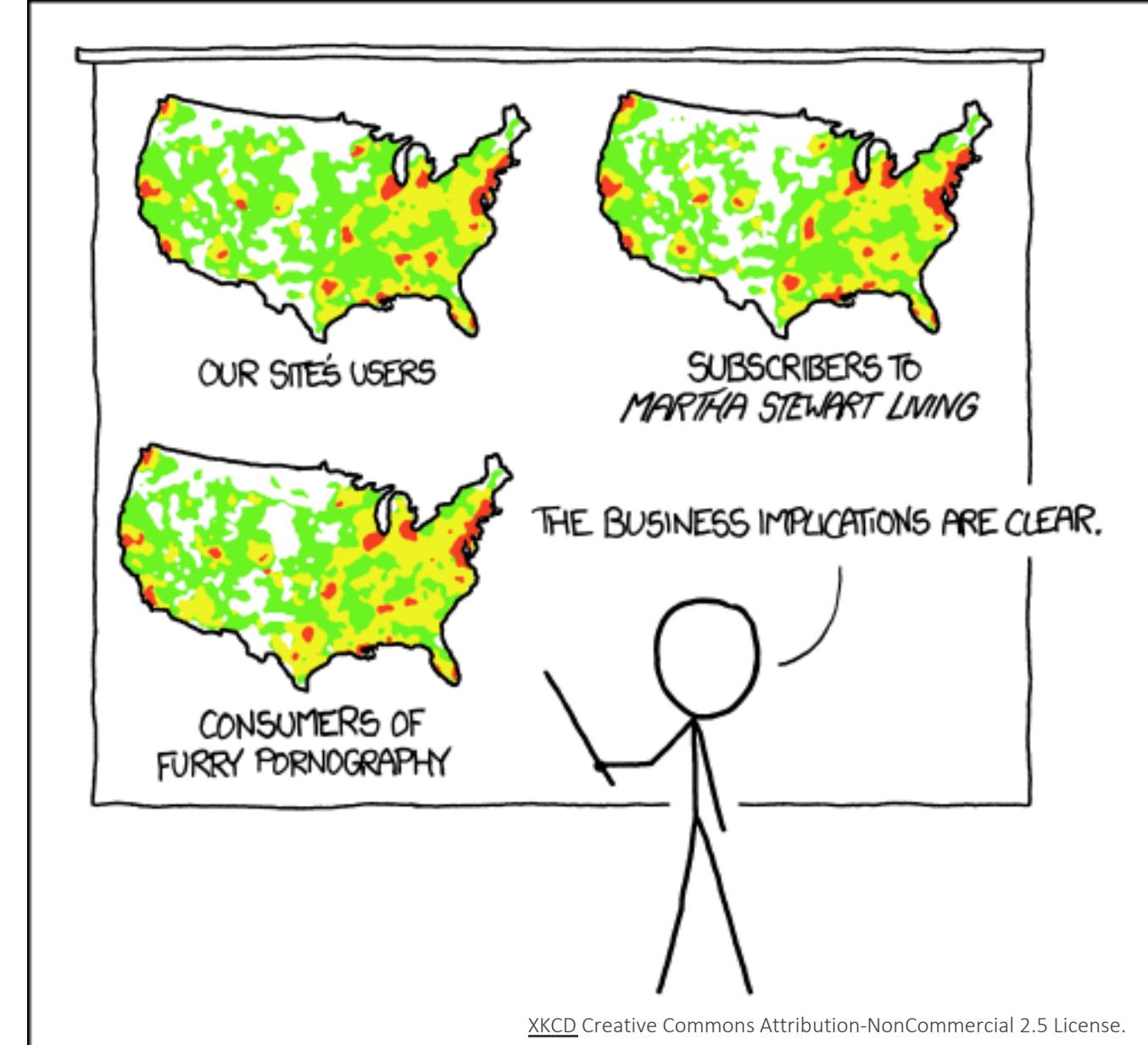
# Introduction to *Urban* Data Science

## Geo-Visualisation

(EPA1316A)

Lecture 6

Trivik Verma



PET PEEVE #208:  
GEOGRAPHIC PROFILE MAPS WHICH ARE  
BASICALLY JUST POPULATION MAPS

# Last Time

- History of Visualisations
- Exploratory Data Analysis
- Types of Visualisations
- Effective Visualisation

# Design Exercise (Time: 10 min)

Q: How Do You Feel about doing science?

Interest	Before	After
Excited (E)	19	38
kind of E	25	30
Ok	40	14
Not great	5	6
Bored	11	12

## Instructions

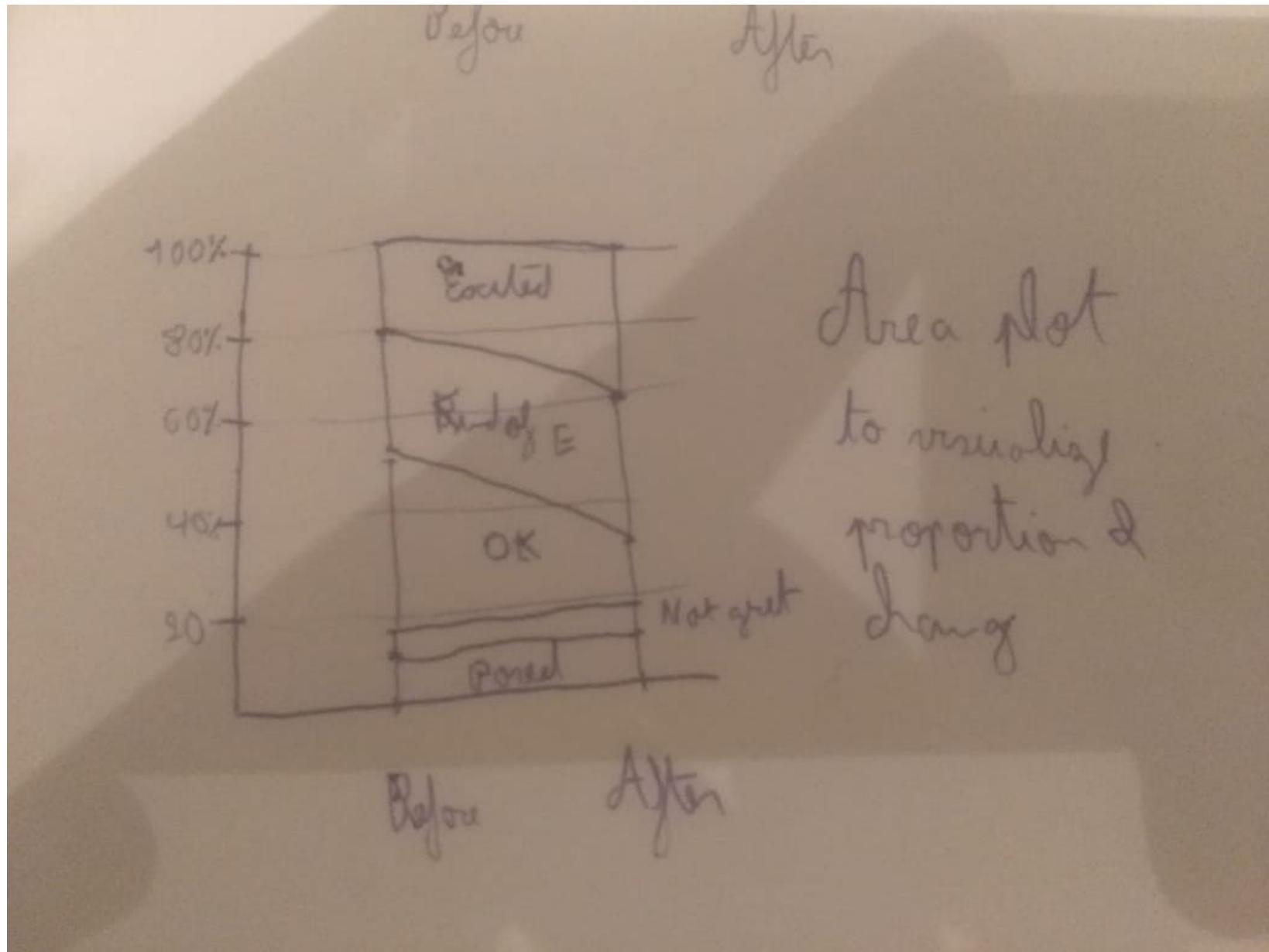
1. **What do you want to do:** Analyse data or Communicate an insight
2. Sketch a visualisation (pen and paper is fine)
3. Take a photo and submit on Assignments in Brightspace under *Visualisation: Design Exercise* (.jpg, .jpeg, .png)
4. Submission deadline tonight by 2330
5. Discussion of some of your submissions follows in Lecture 06.
6. Exercise is **not** graded

After the pilot program,

68%

of kids expressed interest towards science,  
compared to 44% going into the program.

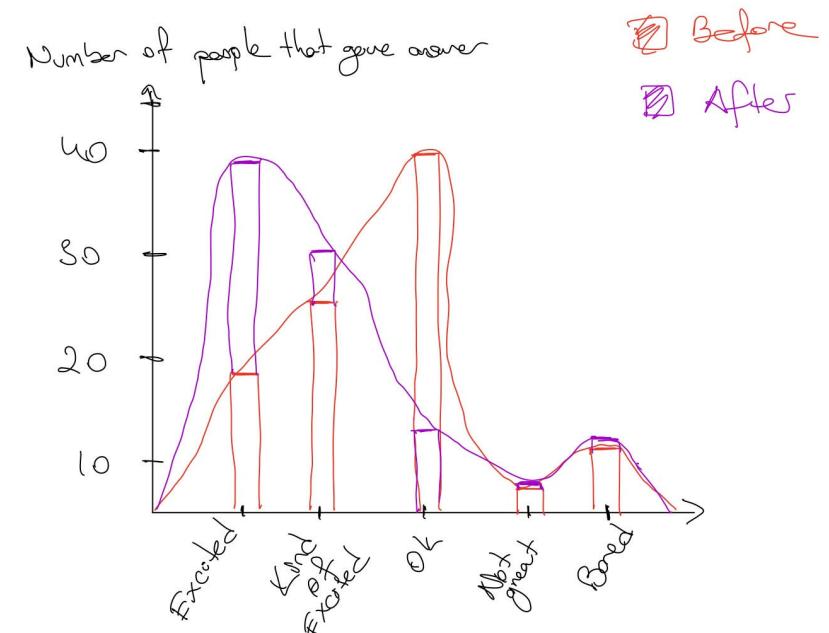
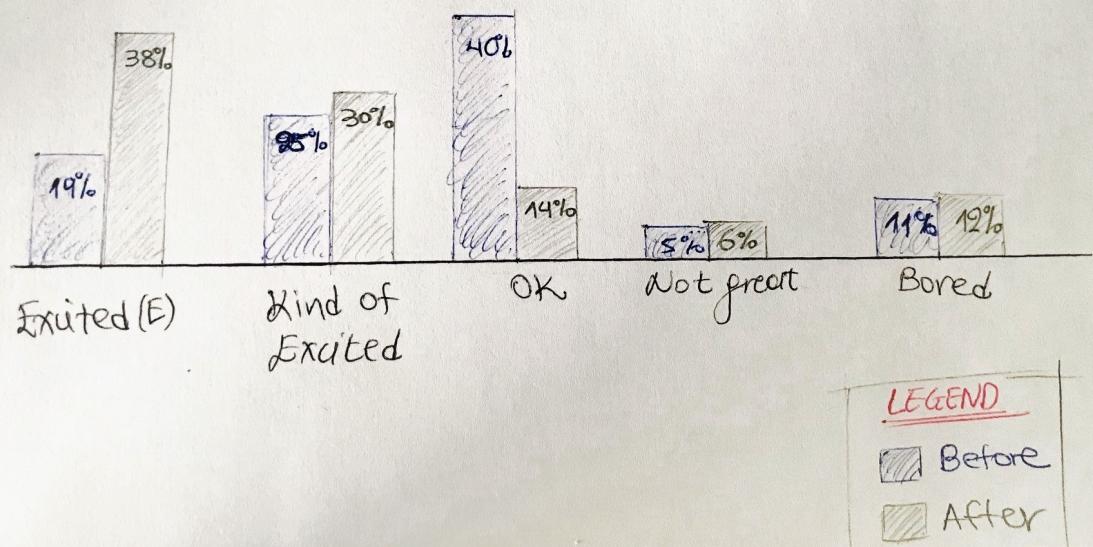
# *Your Submissions*

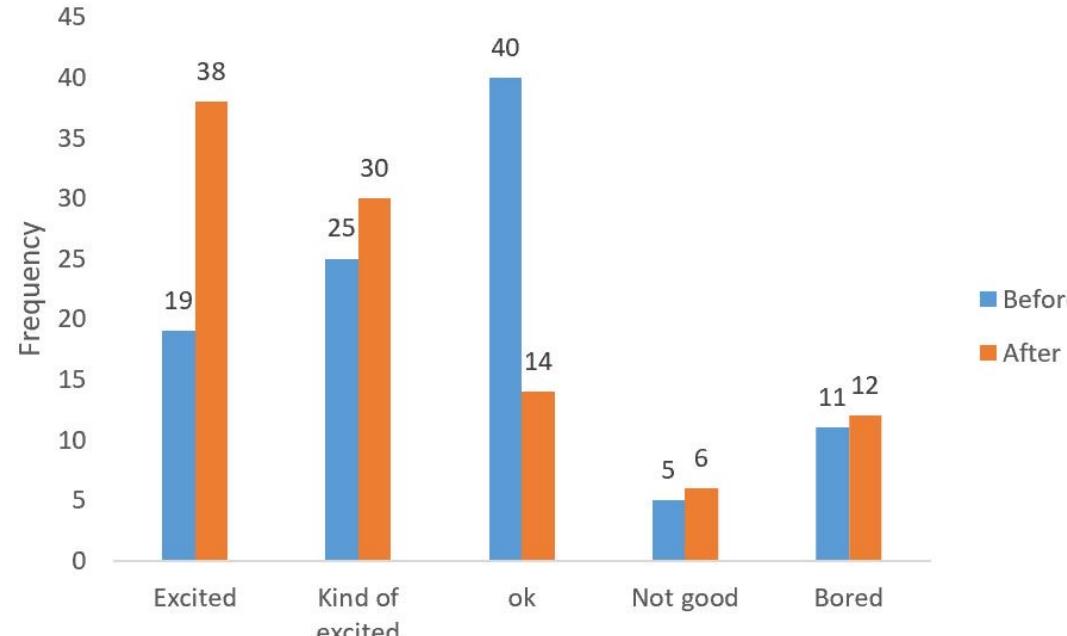


*Last Year*

Q: How do you feel about doing science?

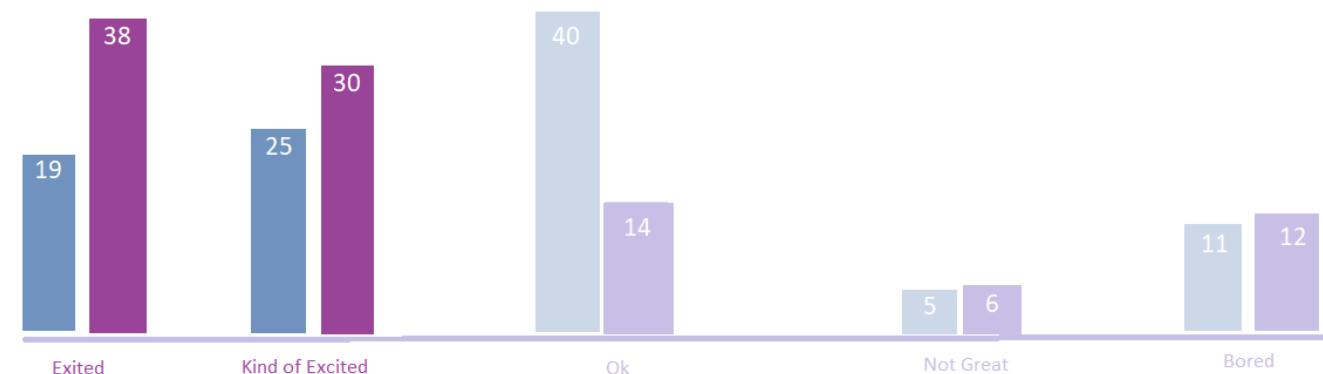
After the pilot program, 68% ( $38\% + 30\%$ ) of kids expressed interest towards science, compared to 44% ( $19\% + 25\%$ ) going into the program.

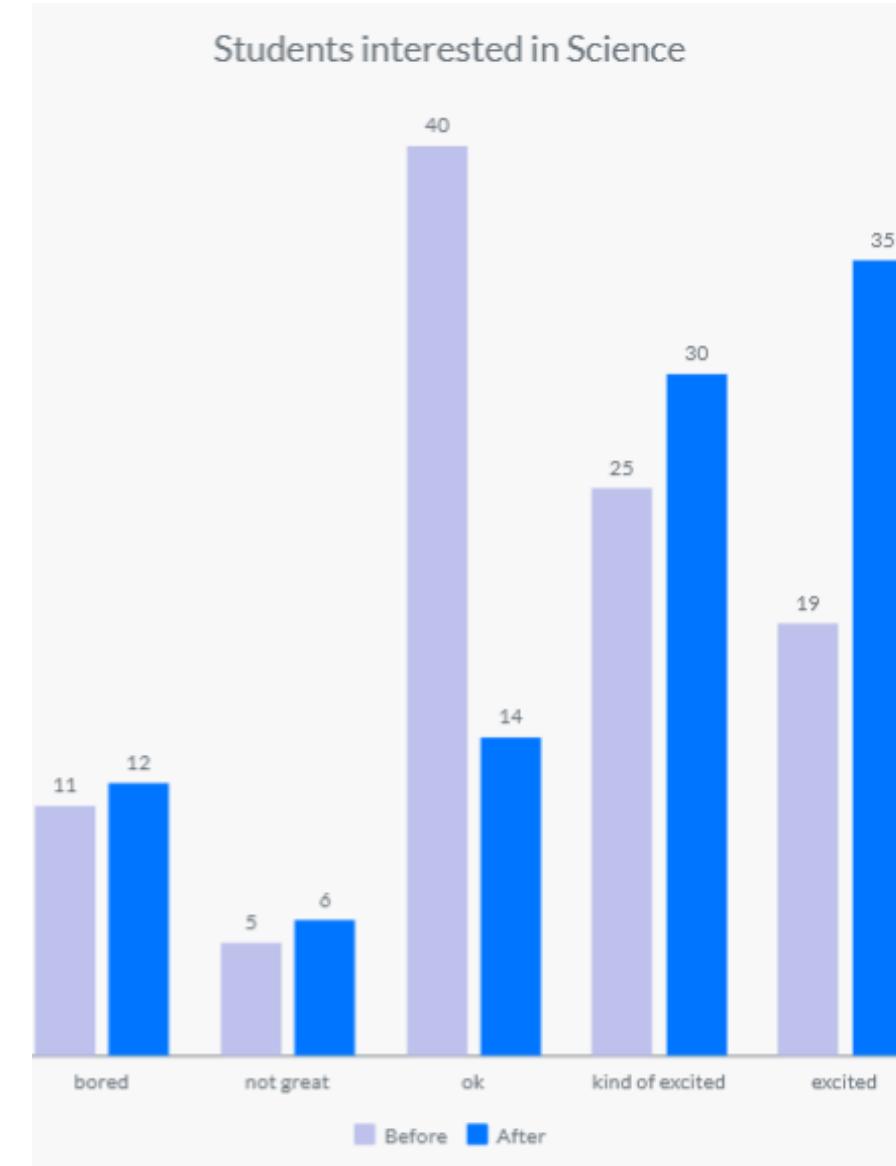


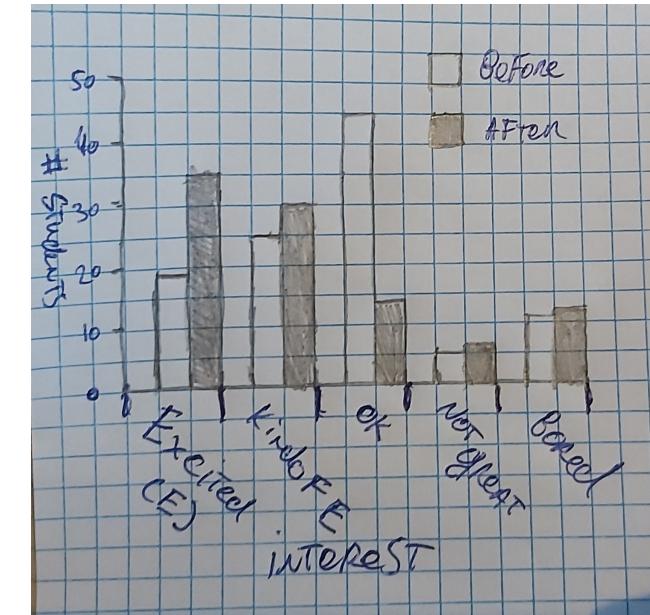
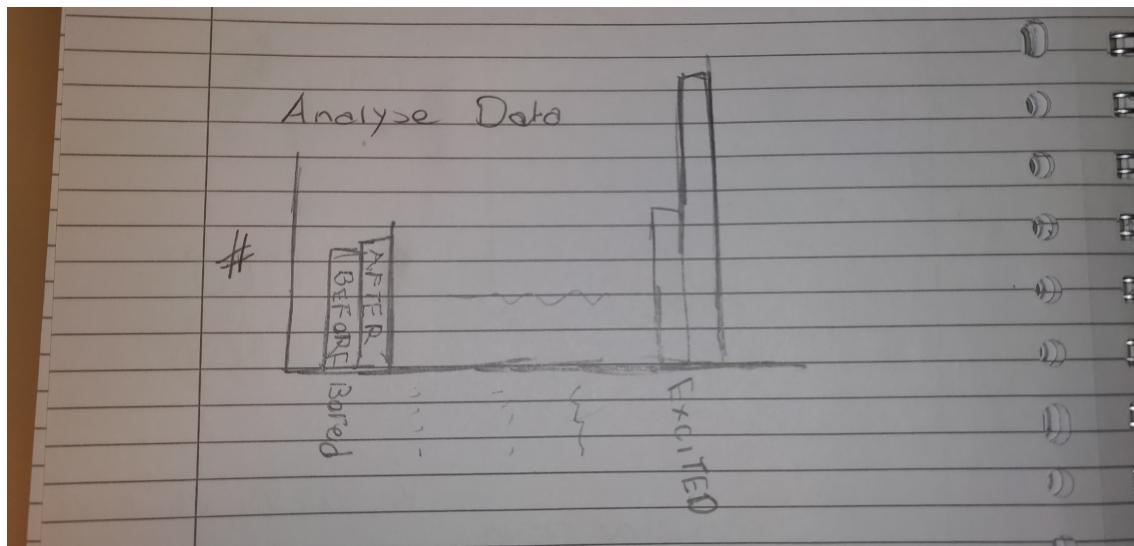


After the pilot program

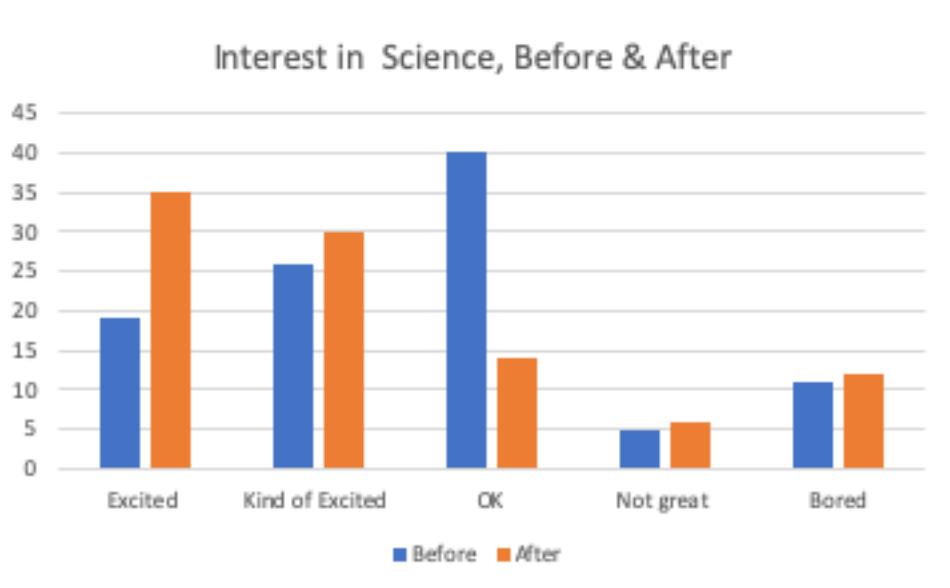
**68%** of kids expressed interest towards science  
compared to 44% going into the program.



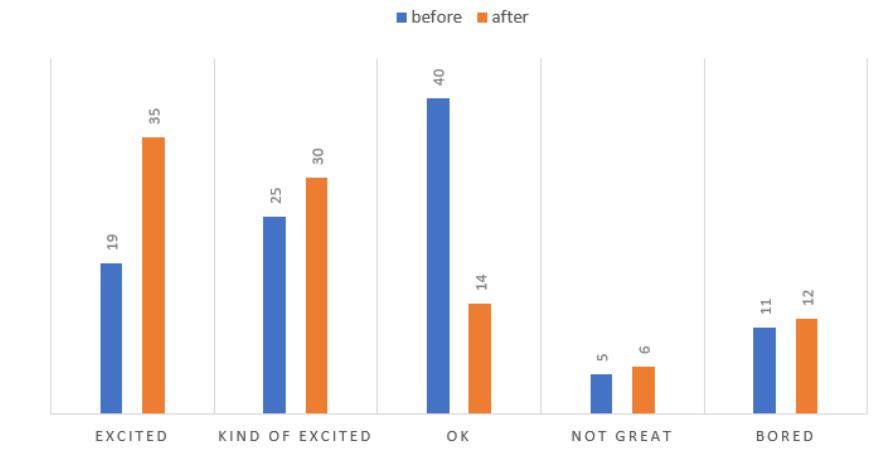


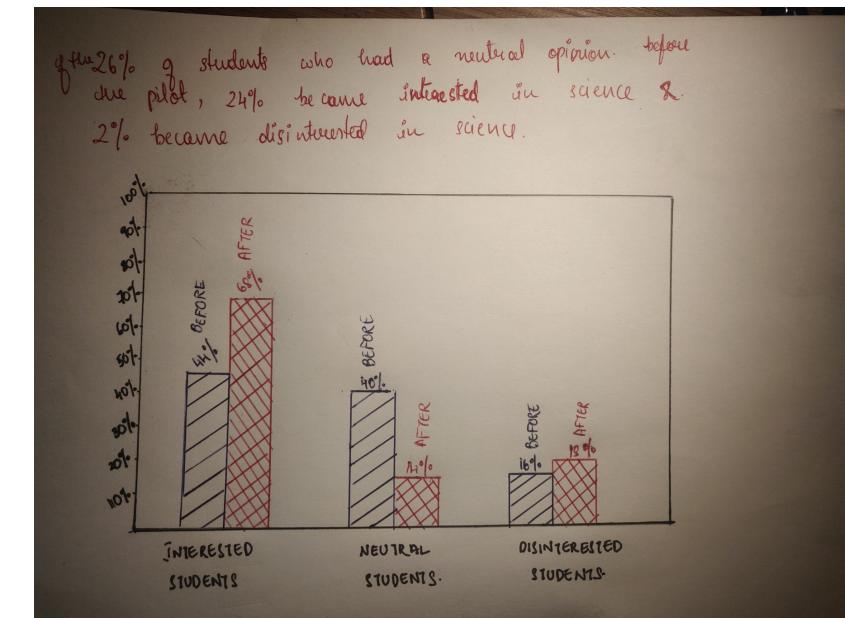
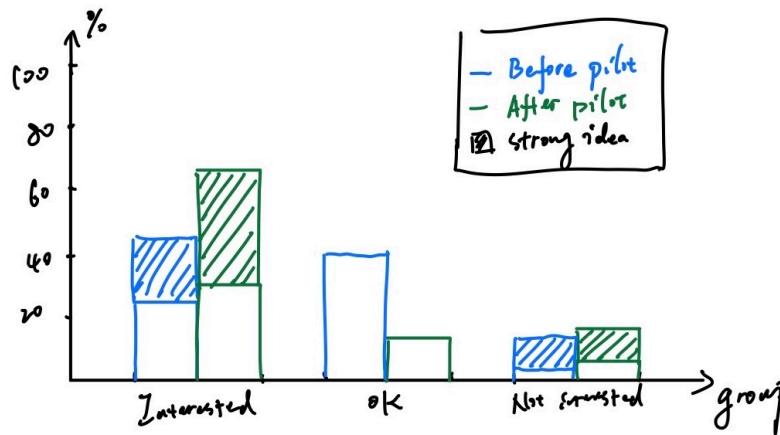


Interest in Science, Before &amp; After

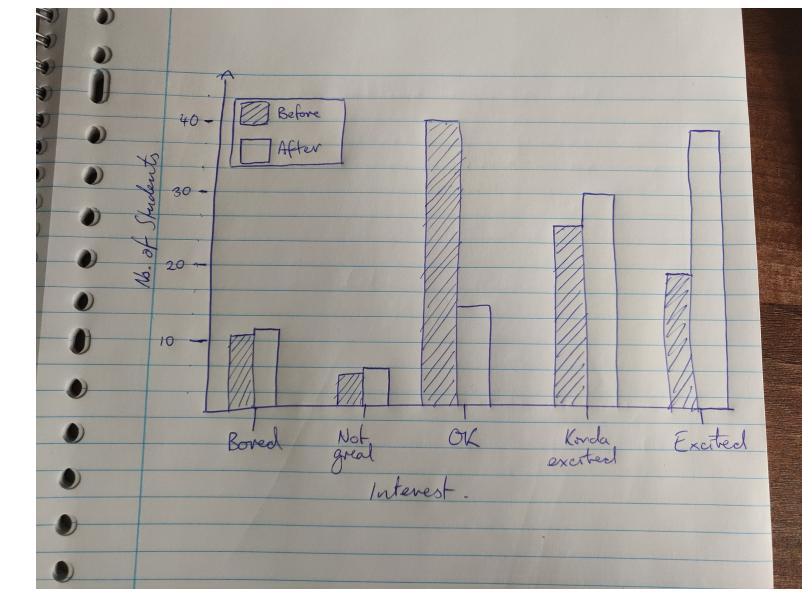
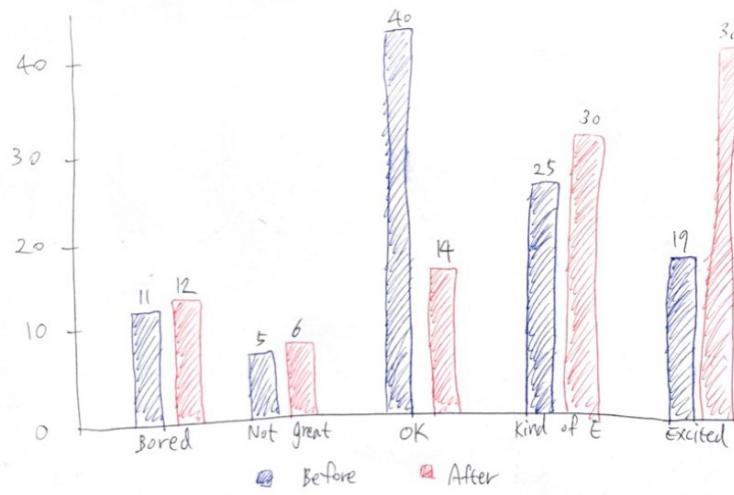


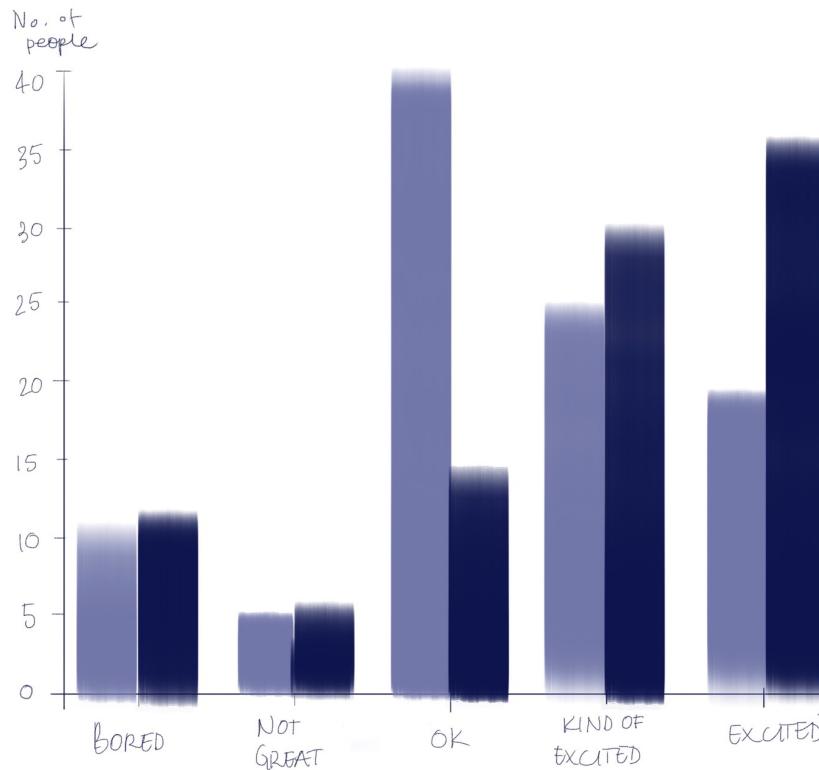
HOW DO YOU FEEL ABOUT DOING SCIENCE?





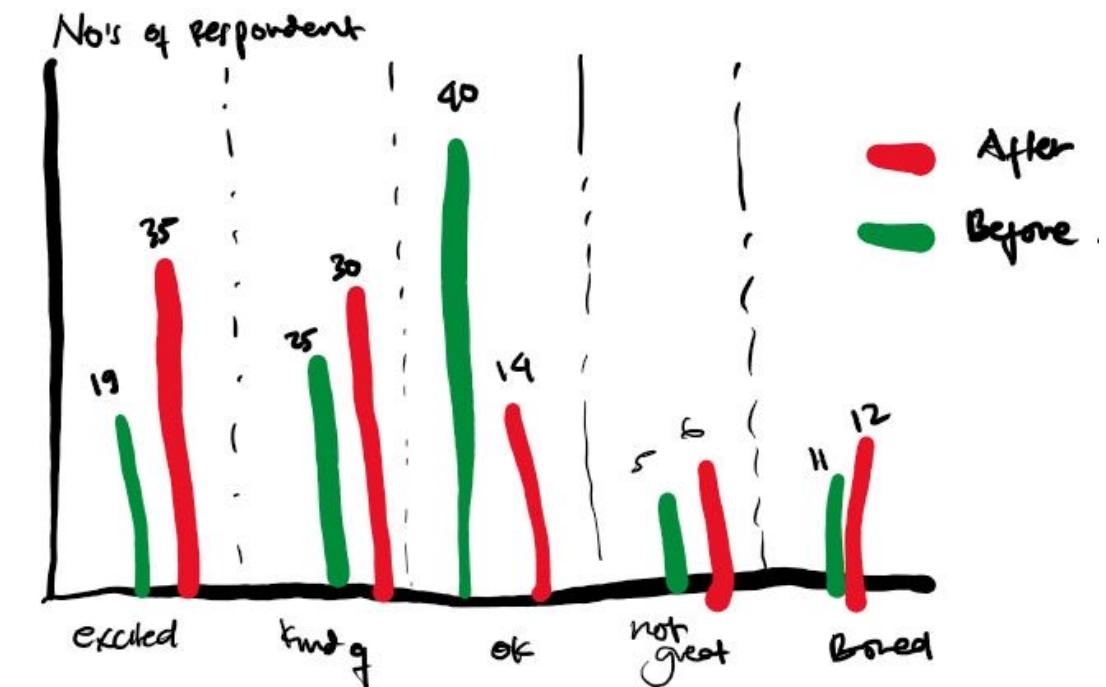
Q: How Do You Feel About Doing Science?





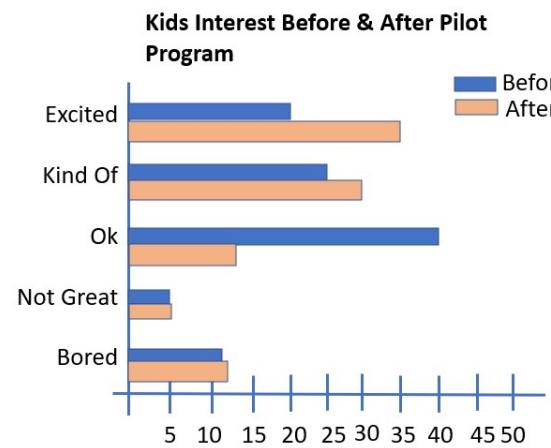
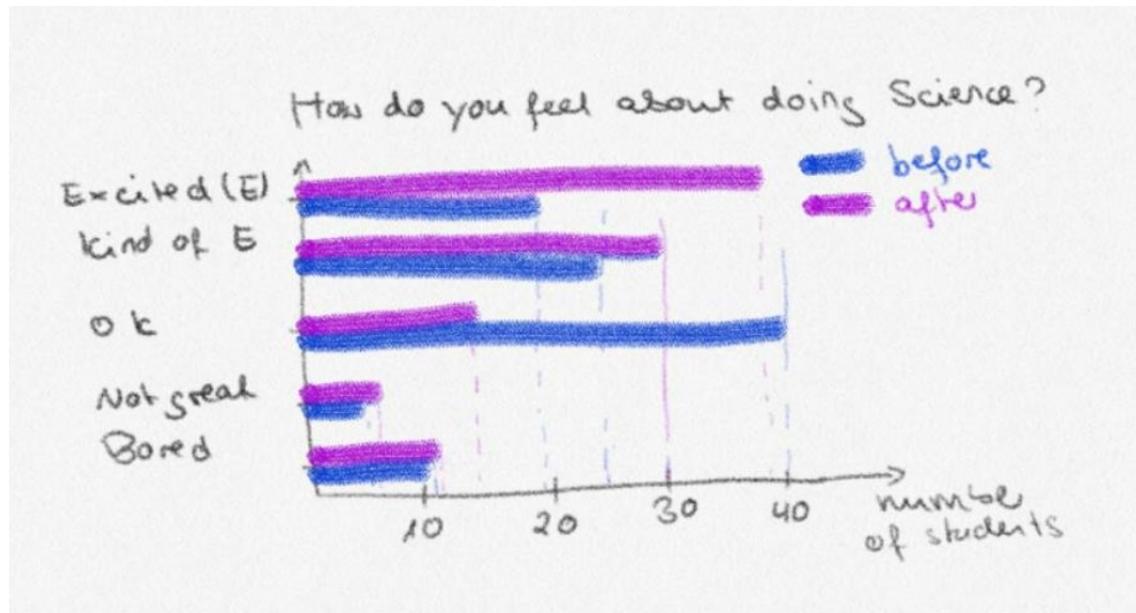
BEFORE

AFTER

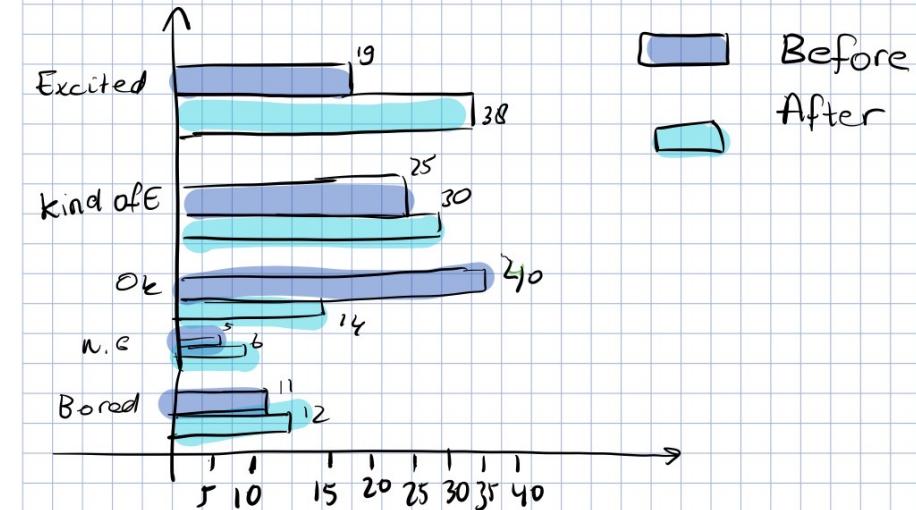


After

Before



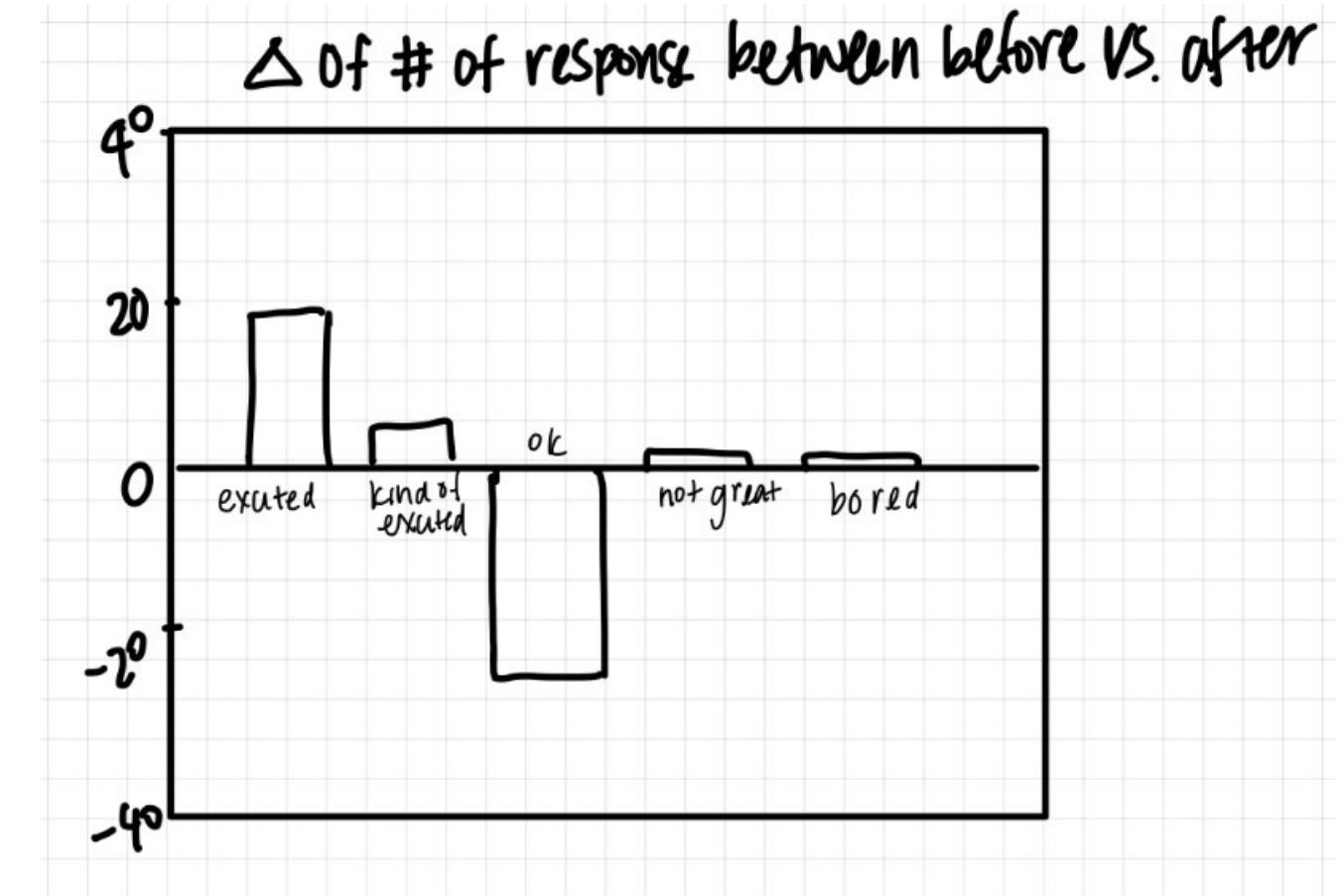
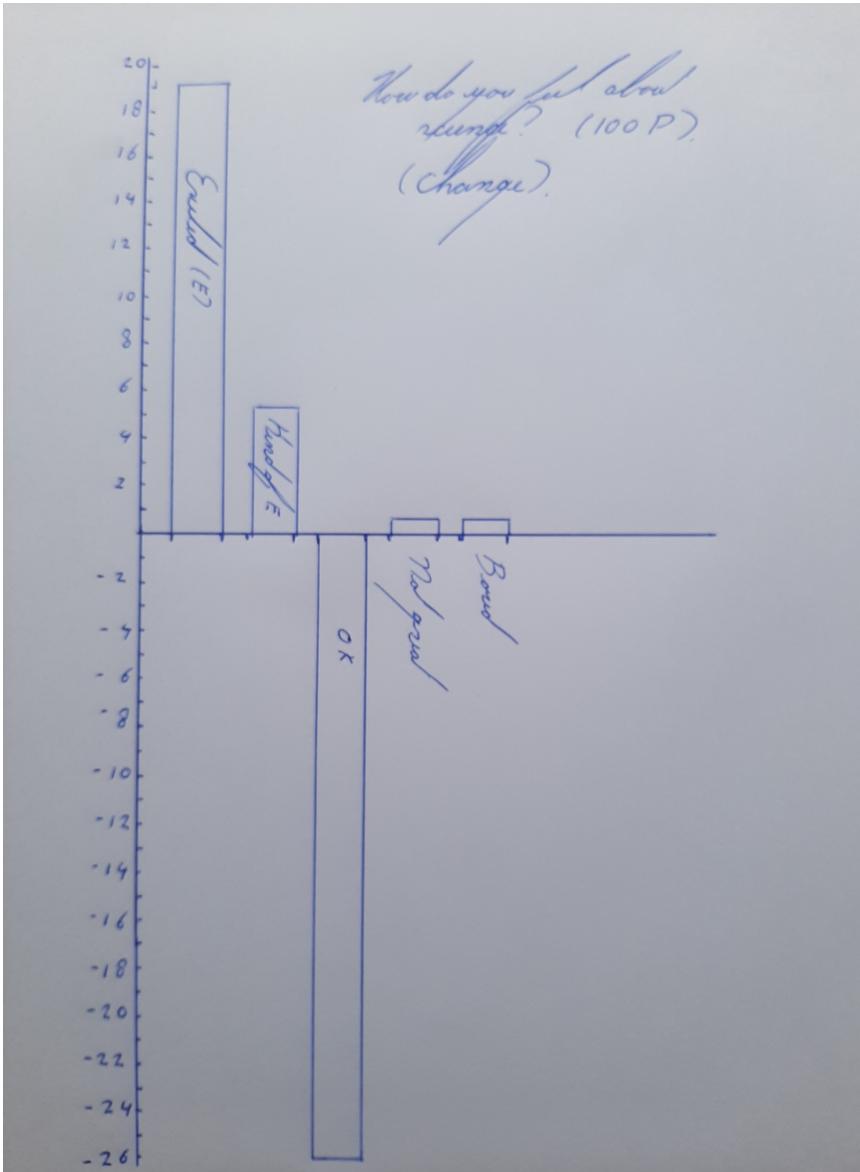
1. Communicate change in perception

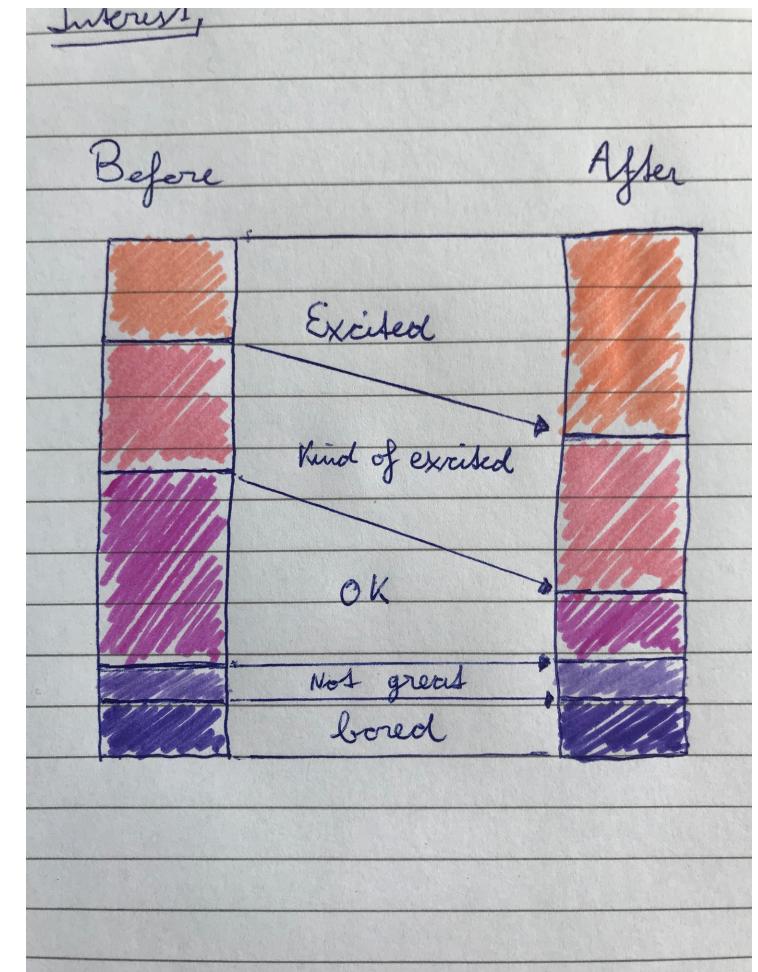
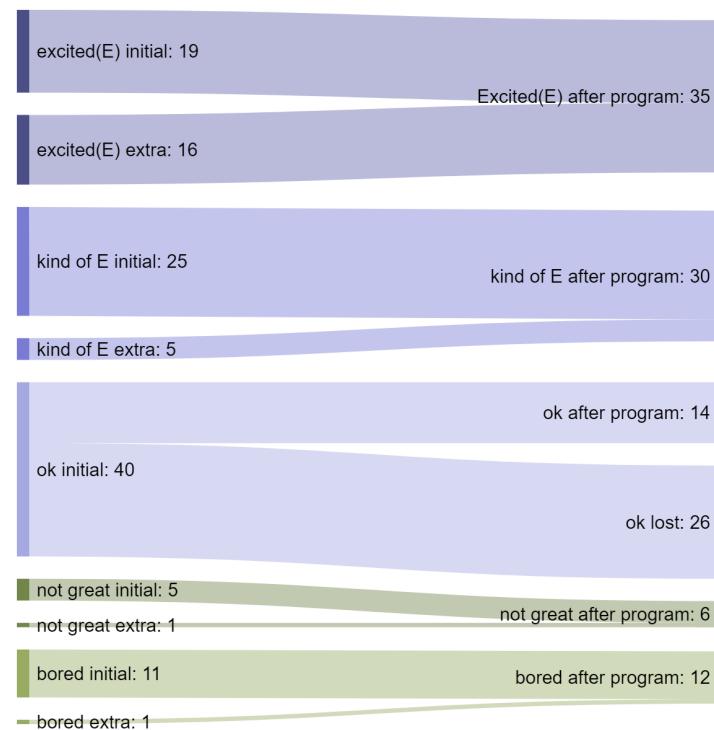
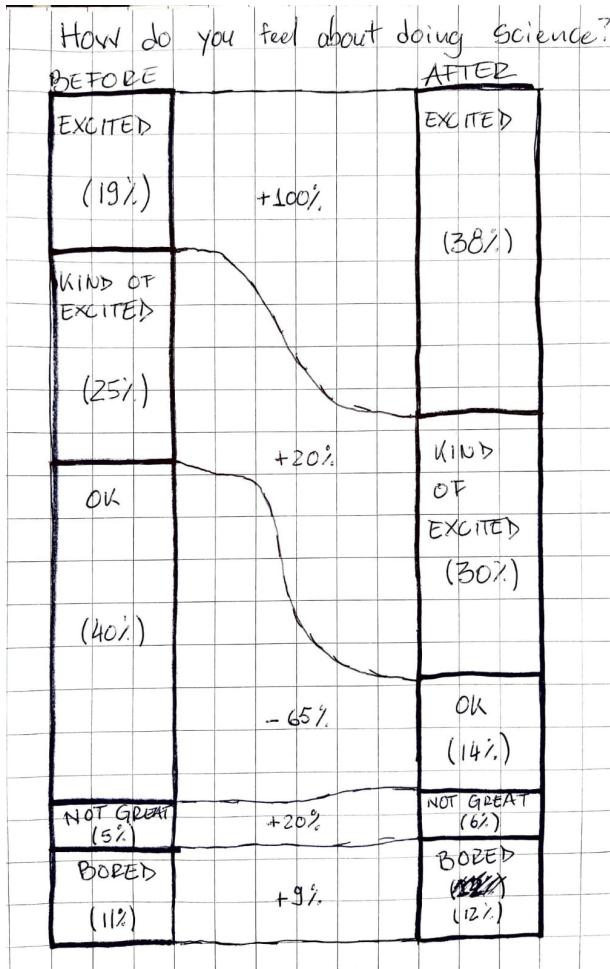


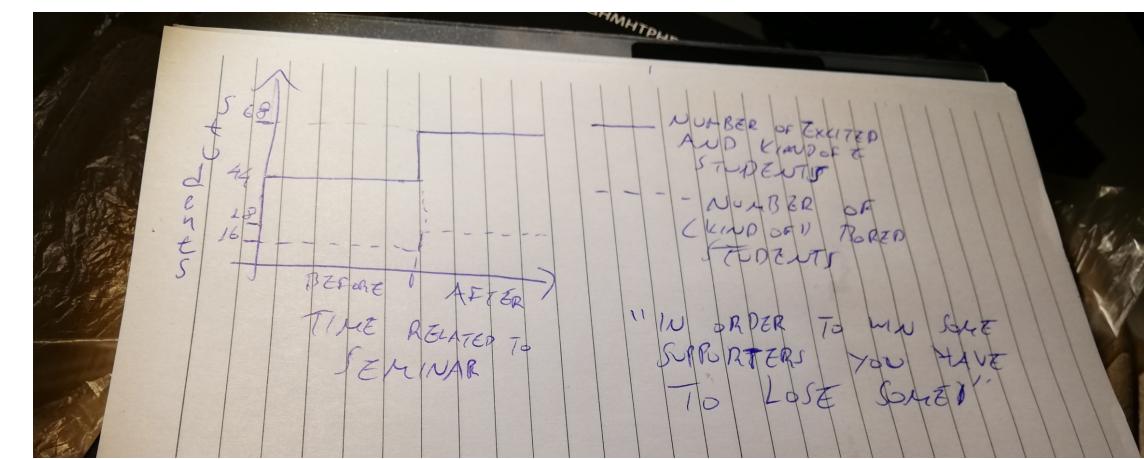
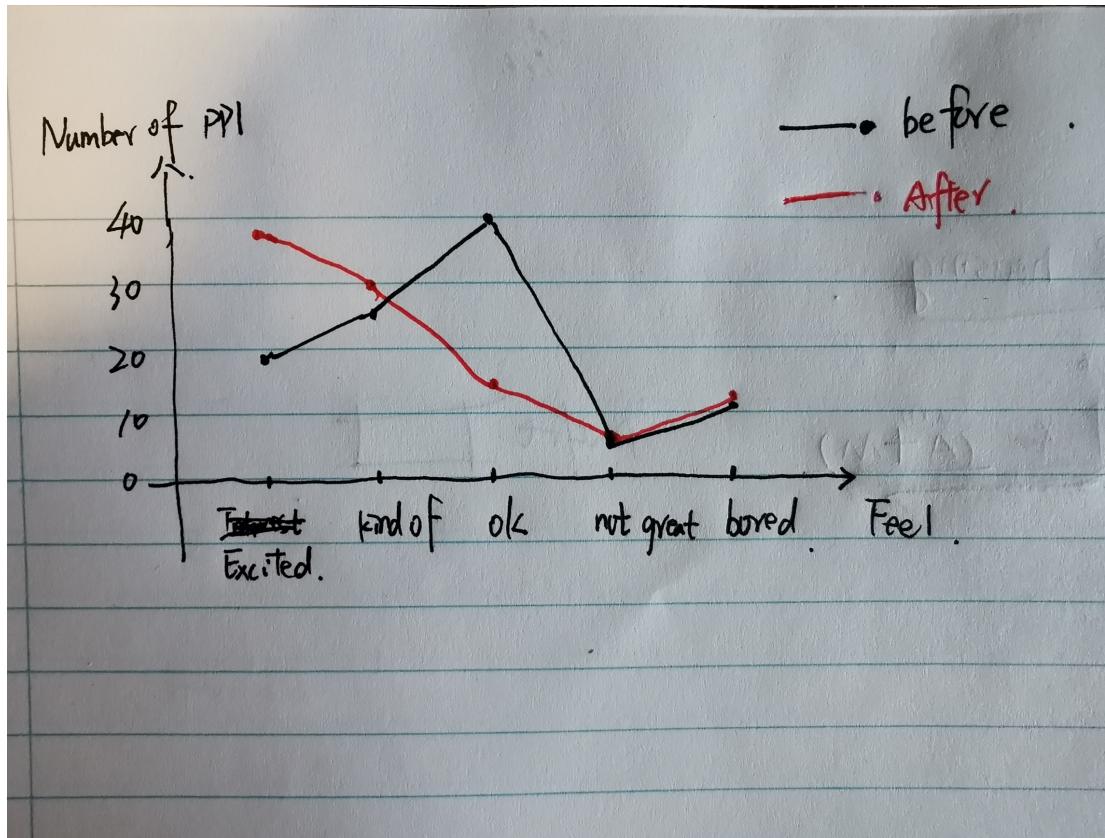
2. Might have problems with scale

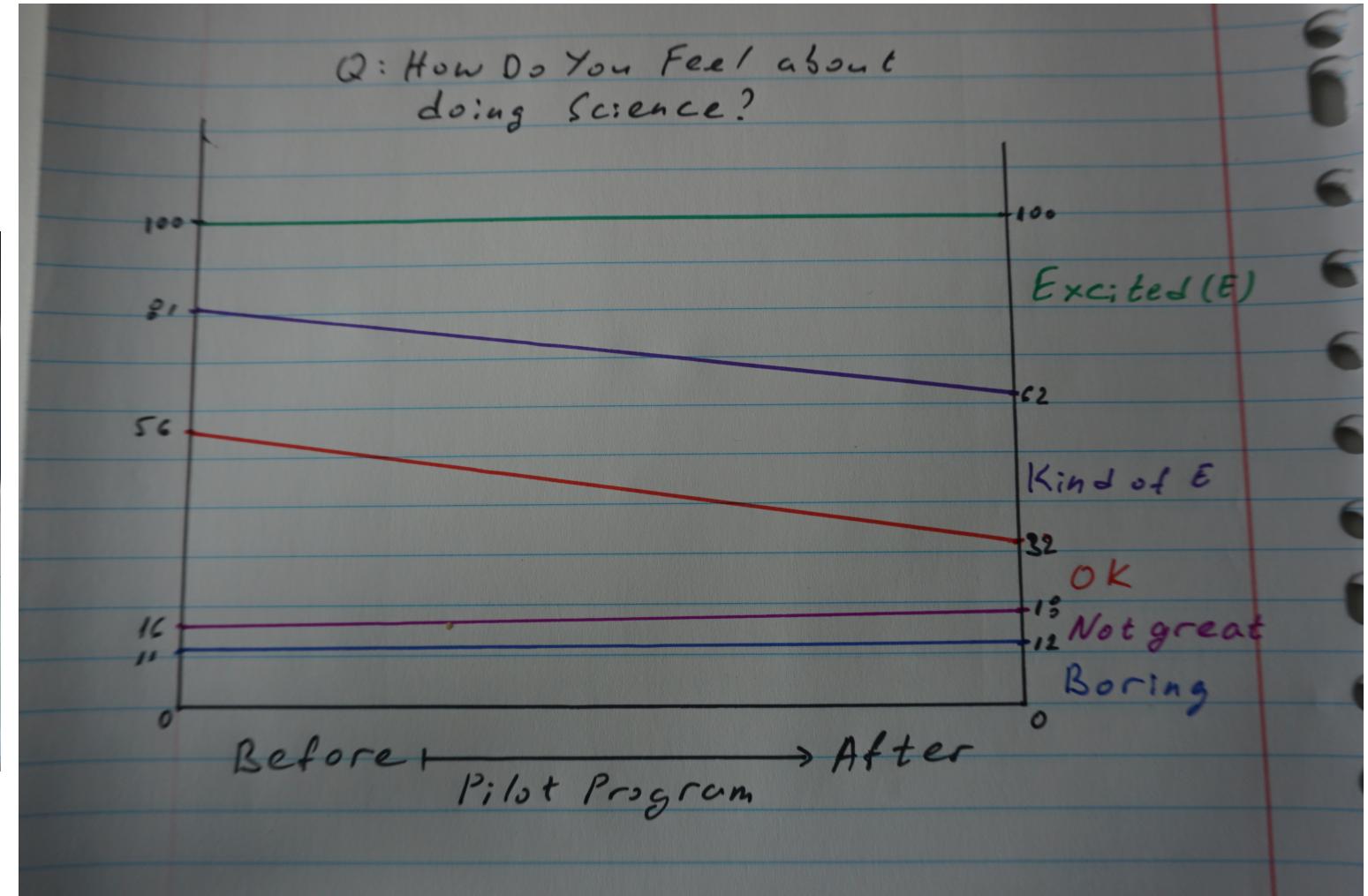
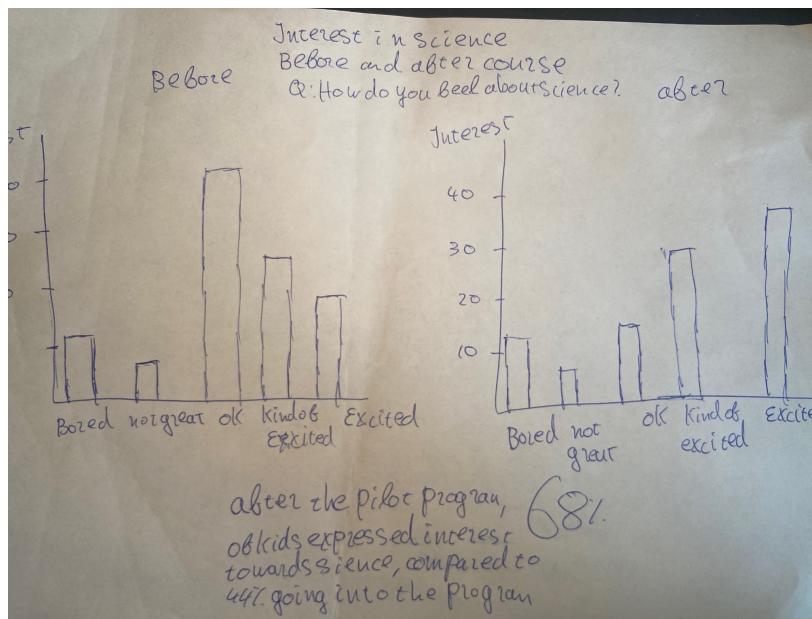
$5 \ll 40$ , not properly implemented.

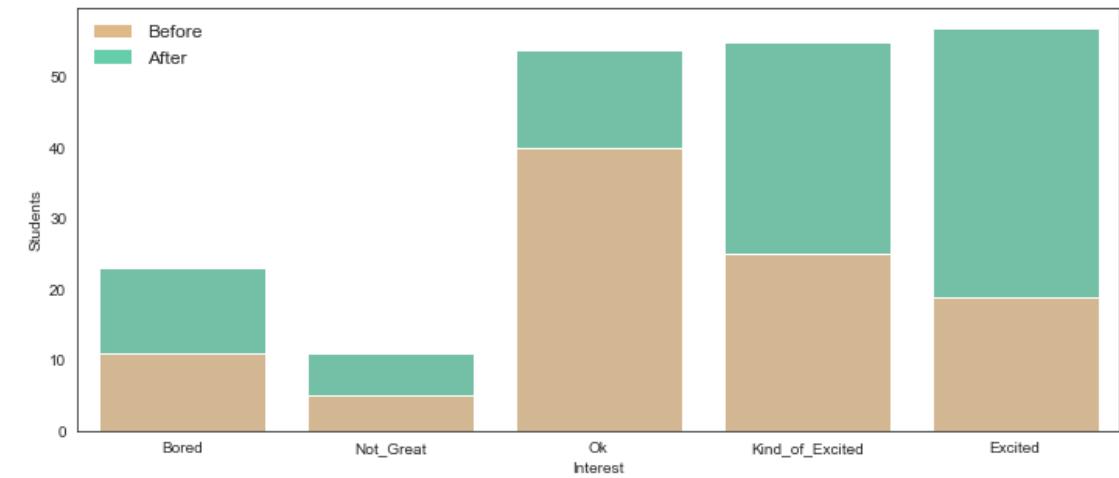
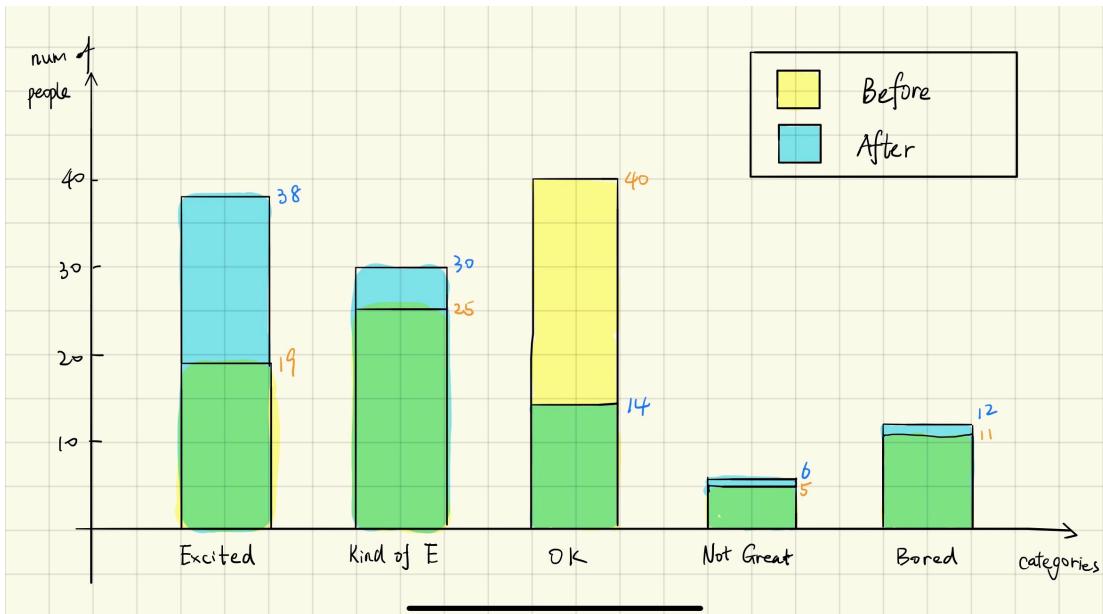
\* Possibly make lollipop representation

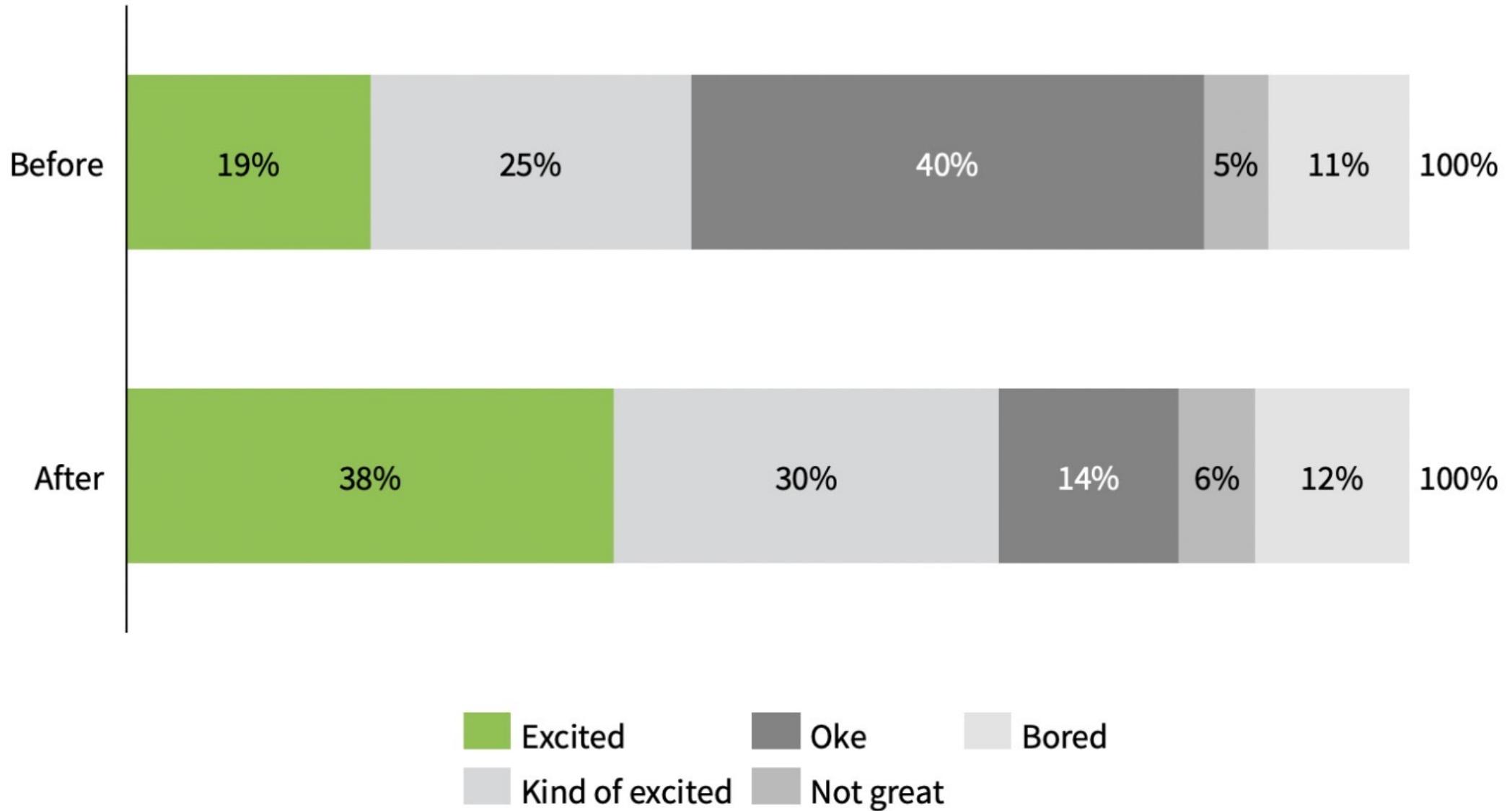








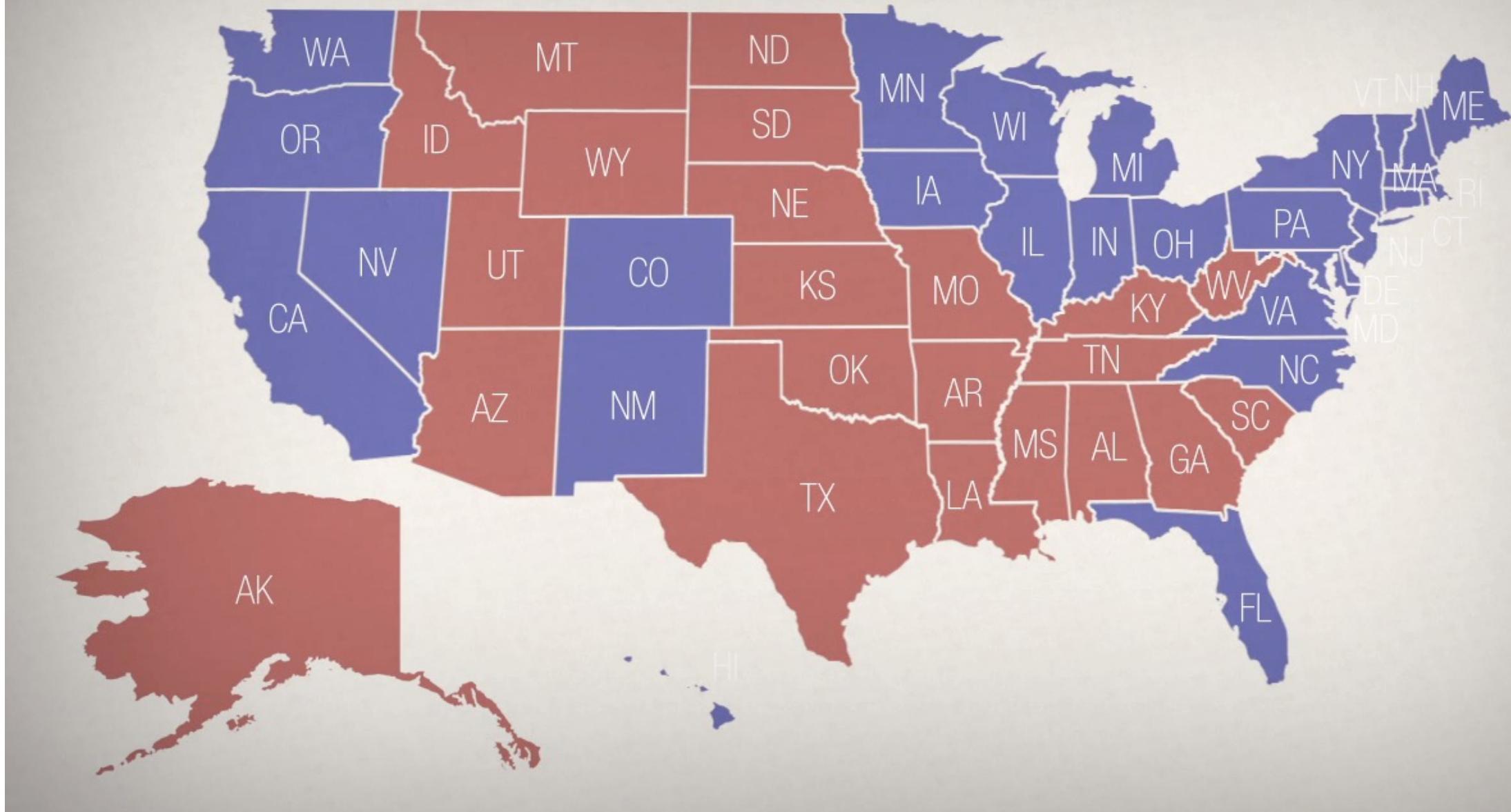




# Today

- Geo-Visualisation
- Dangers of Geo-Vis
- Mapping Data
  - MAUP
  - Choropleths

# 2008 Election



# *Geo-Visualisation*

# Tufte (1983)

*“The most extensive data maps place millions of bits of information on a single page before our eyes. No other method for the display of statistical information is so powerful”*

# MacEachren (1994)

*“Geographic visualization can be defined as the use of concrete visual representations – whether on paper or through computer displays or other media – to make spatial contexts and problems visible, so as to engage the most powerful human information processing abilities, those associated with vision.”*

# Geo Visualisation

- End goal is not to replace the human *in the loop*, but to augment her.
- Augmentation here comes through engaging the **pattern recognition** capabilities that our brain inherently has.
- Combines:
  - Traditional maps
  - Statistical maps
  - Statistical devices of other kind (charts, scatter plots, etc.)
  - **Different roles** in the analysis process...

# A map for everyone

Maps can fulfill several needs

Depending on which one we want to stress, the best map will look very different

MacEachren & Kraak (1997) identify three main dimensions

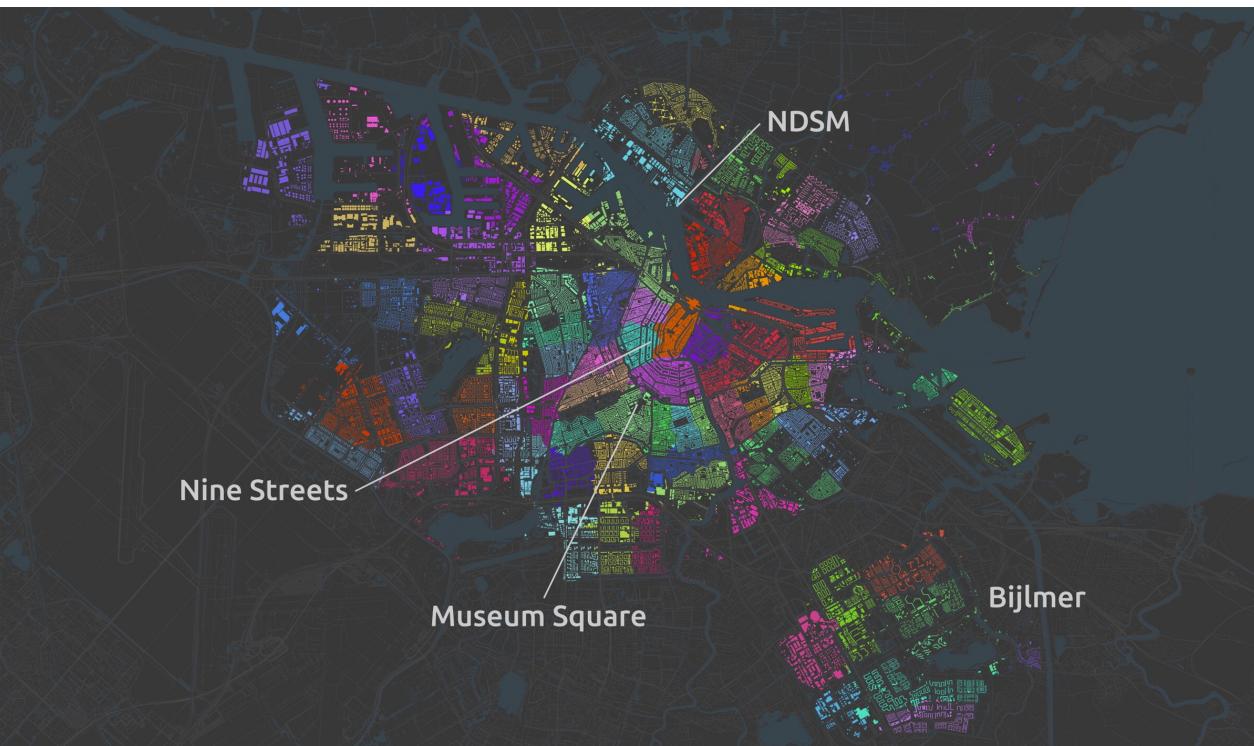
- Knowledge of what is being plotted
- Target audience
- Degree of interactivity

# All plots are composed of

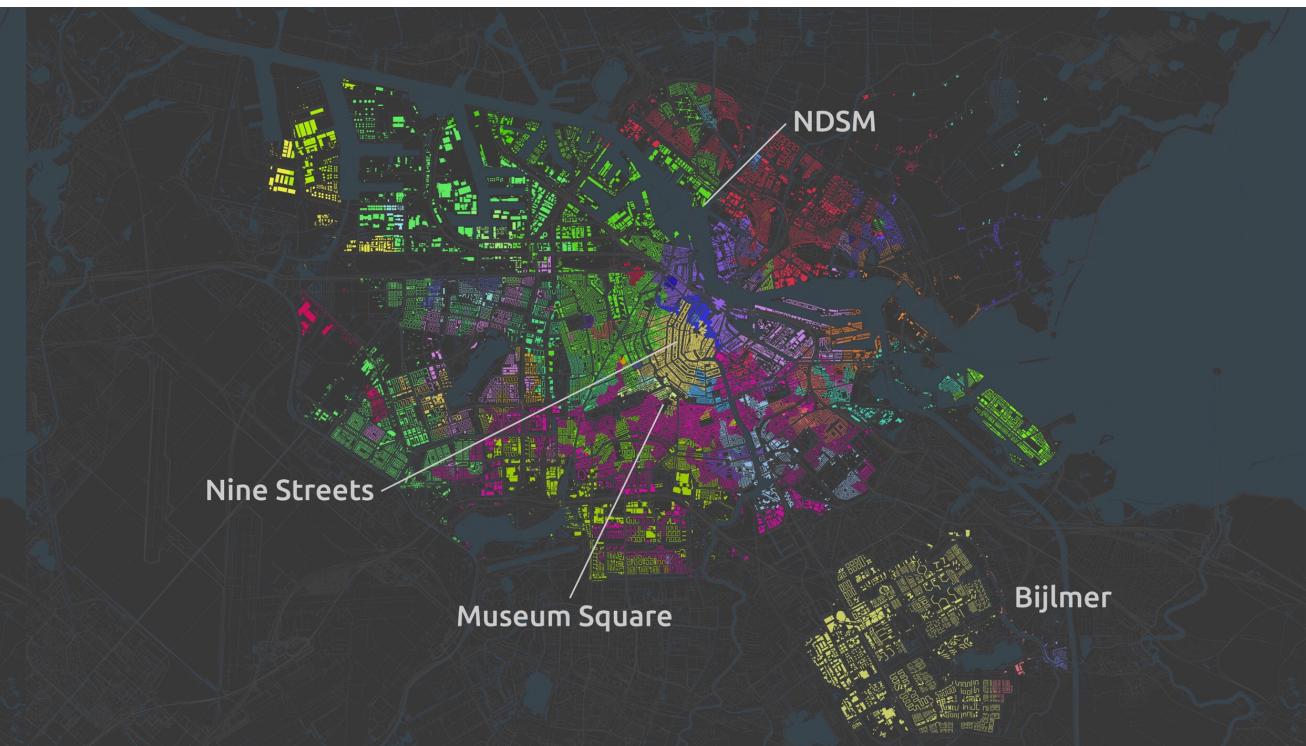
- *Data* that you want to visualize
- *Layers* made up of geometric elements
- *Scales* which map values to aesthetics
- *Systems* of coordinates
- *Facets* and their specification
- *Themes* controlling finer points

Un/known: *fast* and *slow* maps

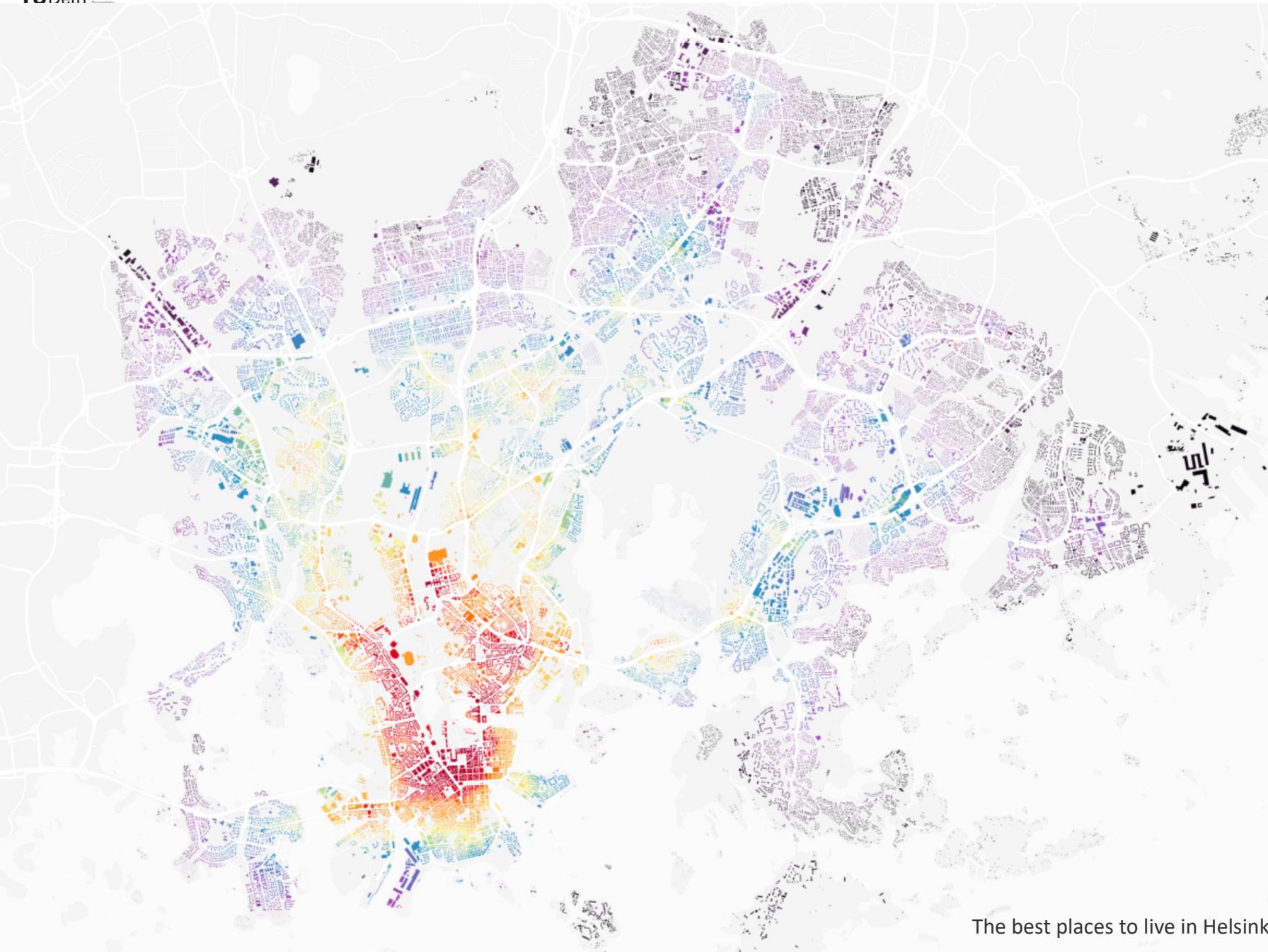
Digit postcodes : fast



Twitter : slow



Audience: *easy* and *hard* maps



*easy map*

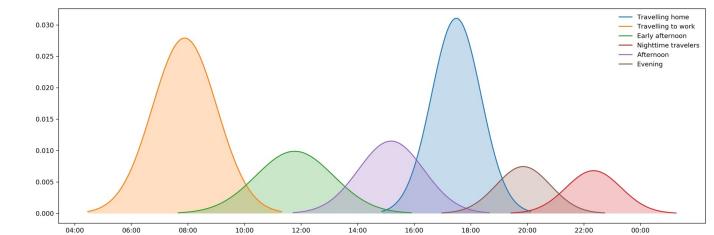
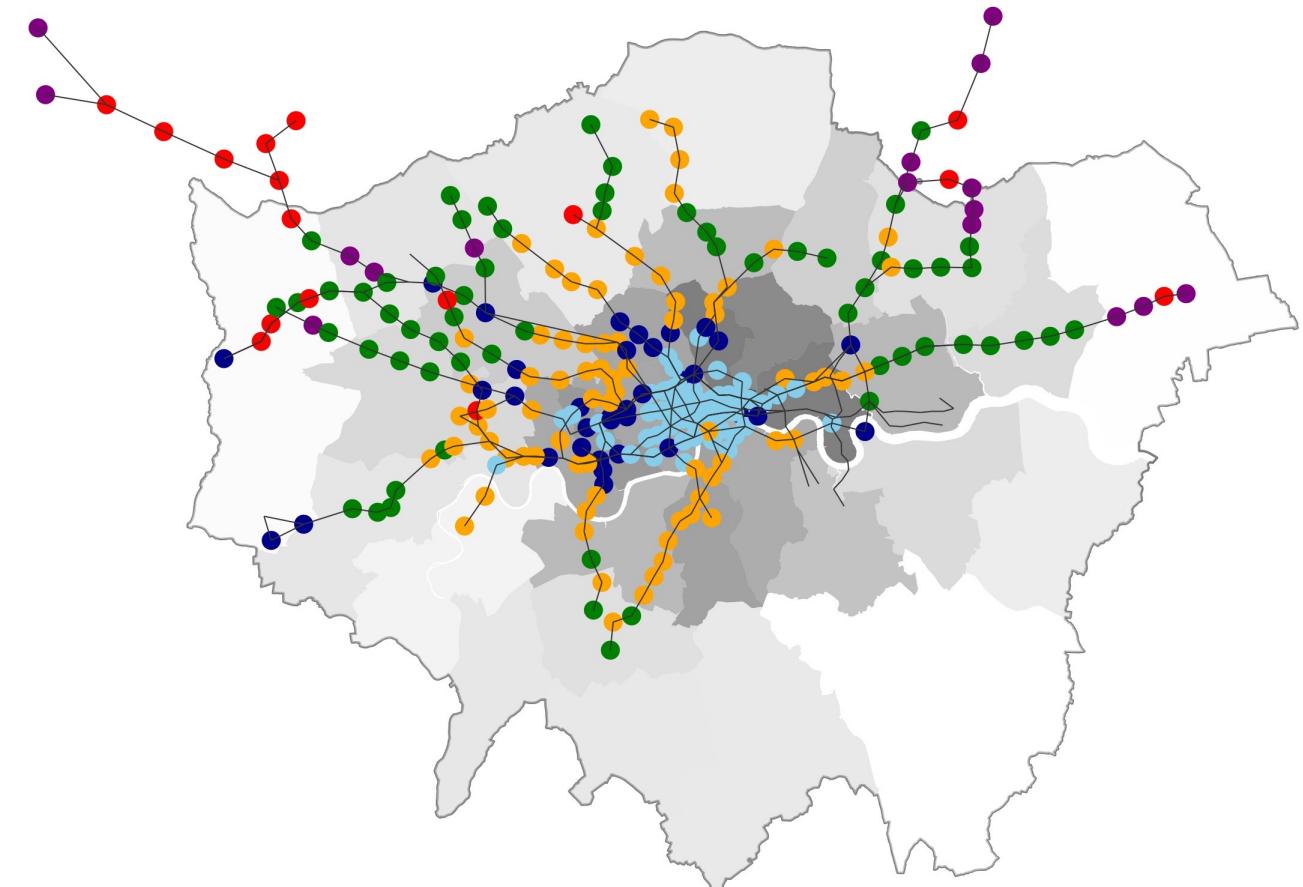
Ambitions Travel Index



[source]

## Greater London, TfL stations

*hard map*

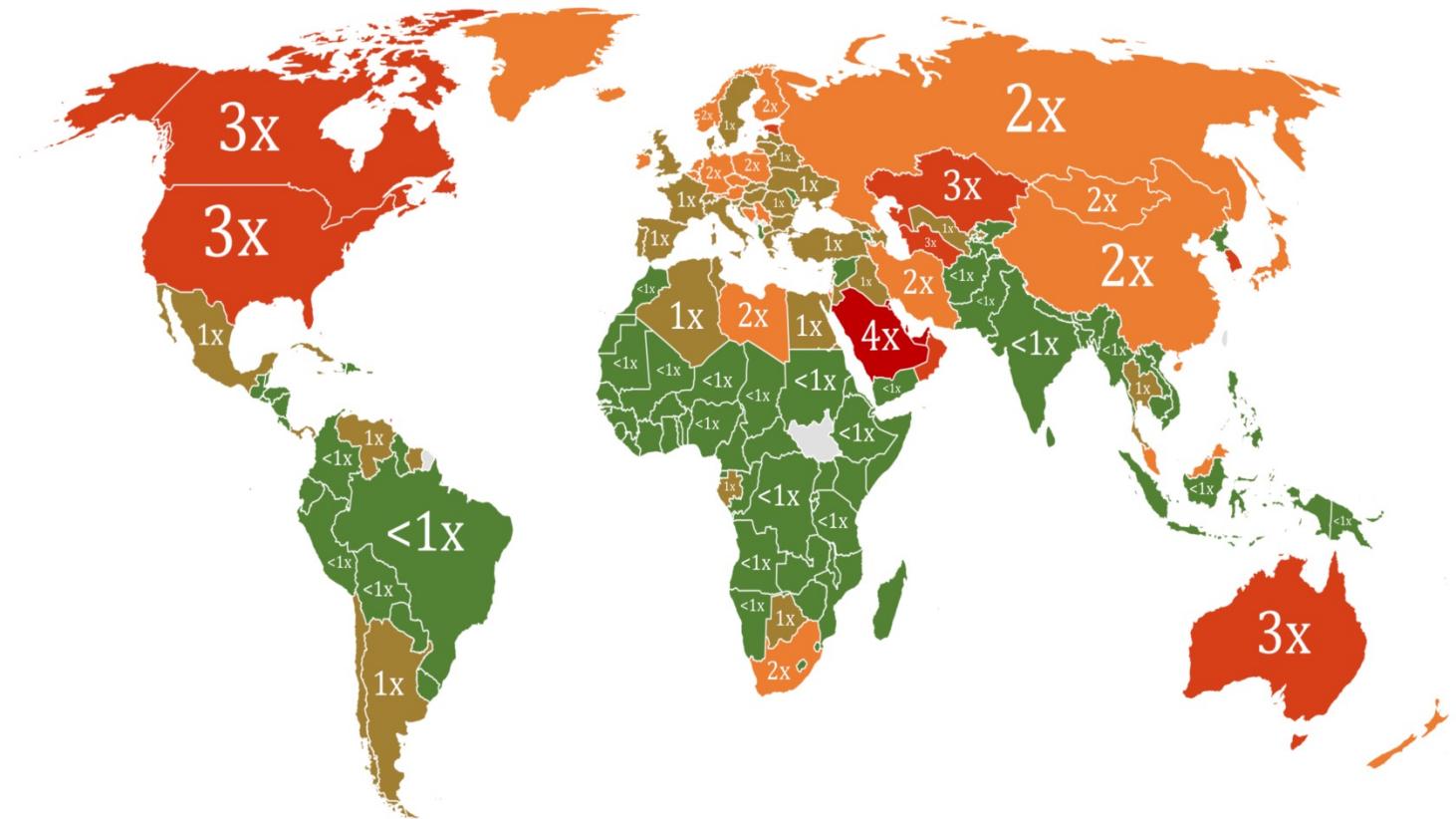


Interaction: *one or many* maps in one

# More Than Their Fair Share:

Which countries produce a greater proportion of global CO<sub>2</sub> emissions than their proportion of the global population?

*static map*



Ratio of Global Emissions Contribution to Share of Global Population

TRANSPORT FOR LONDON Plan a journey Status updates Maps Fares Help & contacts More Search ★

Urban planning and construction Our role in planning Planning with WebCAT WebCAT planning tool

# WebCAT

Address or co-ordinates

 X Go

Access level (PTAL) Time mapping (TIM)

TIM: a new measure, looking at how far you can travel in a given journey time.

Map key - Travel Time

< 15 mins	15 - 30 mins
30 - 45 mins	45 - 60 mins
60 - 75 mins	75 - 90 mins
90 - 105 mins	105 - 120 mins
120 - 135 mins	135 - 150 mins

Change travel time bands +

Map layers

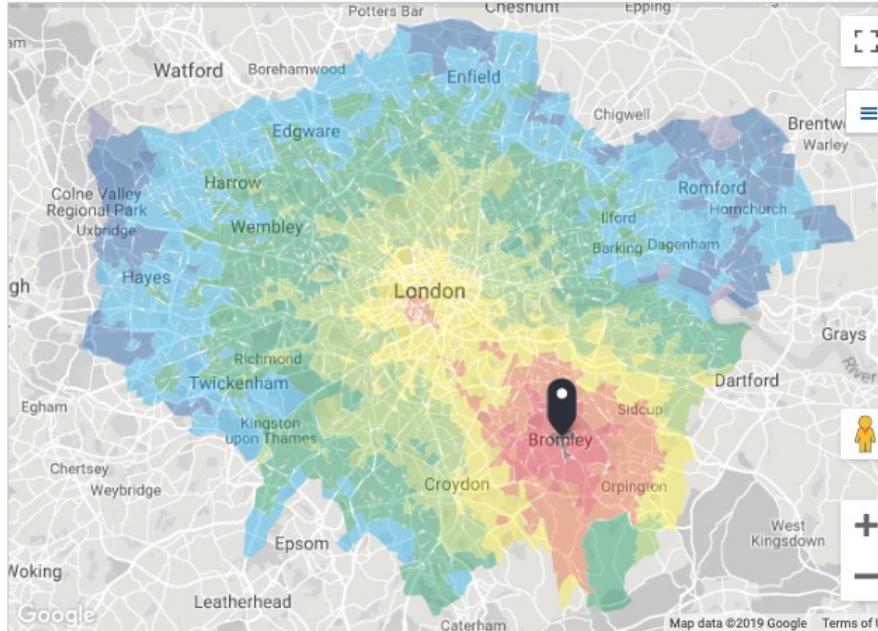
Travel Times

Scenario

Base Year

Mode

All public transport modes



You can click anywhere on the map to change the selected location.

## TIM output for Base Year

Scenario: **Base Year** Mode: **All public transport modes**, Time of day: **AM peak**, Direction: **From location**

### Bromley

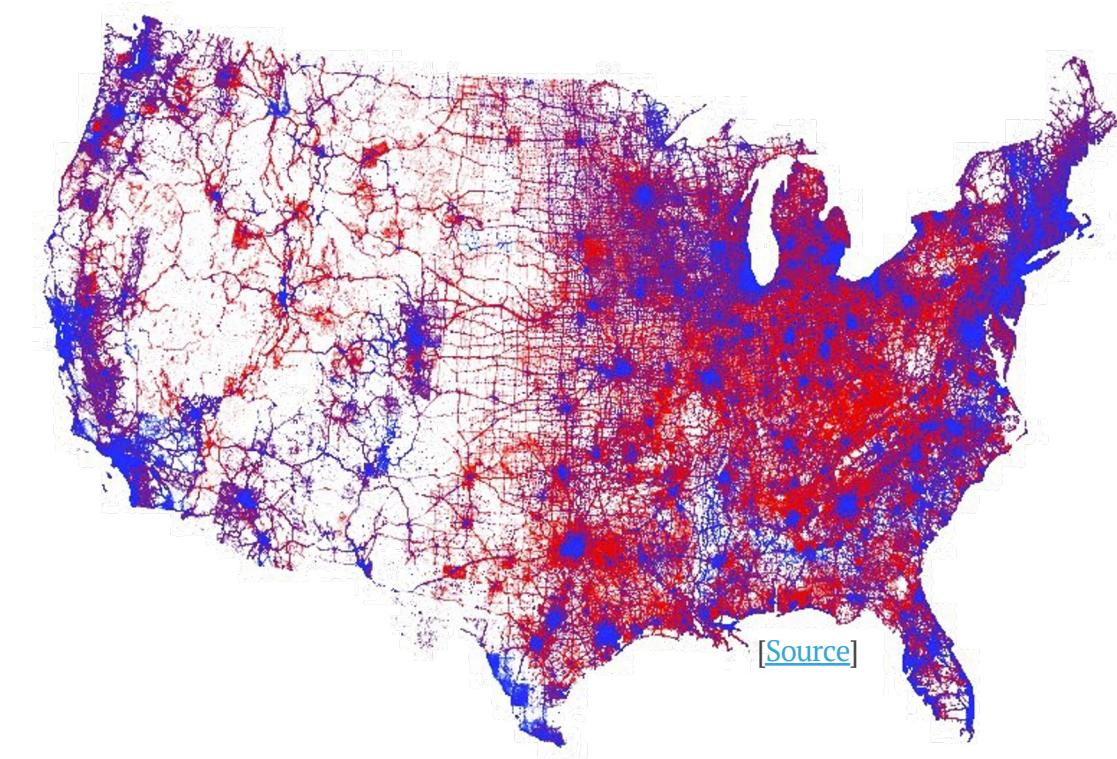
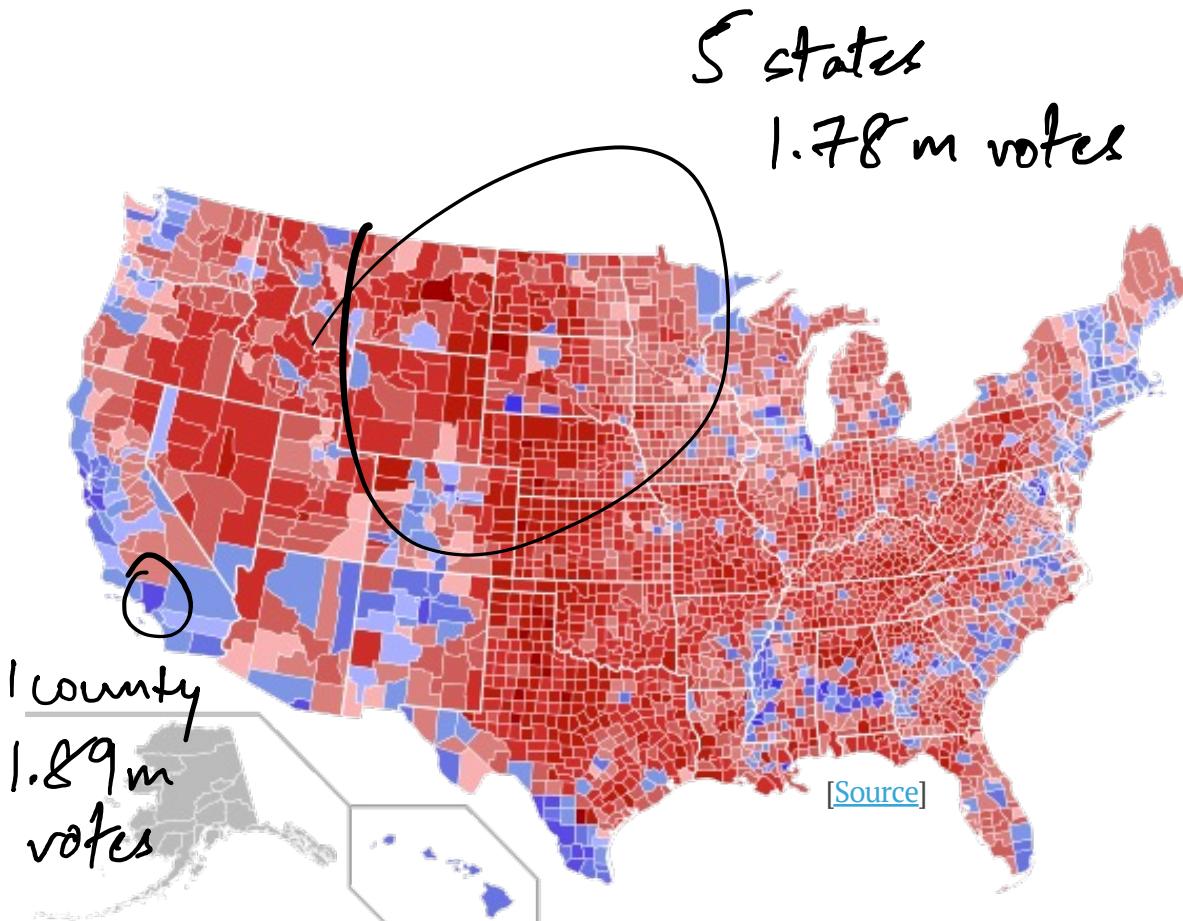
Bromley, UK

Easting: **540120**, Northing: **169366**

# interactive map

# *Dangers* of GeoVisualisation

## 2016 Presidential Election in the USA



# *How to lie with maps\**

The human brain is so good at picking up patterns...  
... that it finds them even where they don't exist!

***Patternicity*** (Shermer, 2008) The tendency to find meaningful patterns in meaningless noise

***Apophenia*** (Konrad, 1958) The experience of seeing patterns or connections in random or meaningless data

# How to *be truthful* with maps

“With great power comes great responsibility”

Statistics to the rescue!!!

- Complement and enhance visuals
- Help disentangling **true** from **spurious** patterns
- **Reciprocity:** GeoVis can also enhance statistics and make them more useful

# How to *be truthful* with maps

- What is the story?
- Who is the audience?
- Did the visual communicate the story?
- What story was not told?
- How are the design elements used?

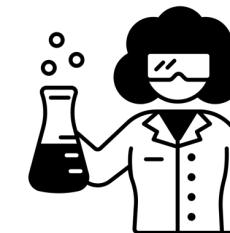
# Break



CHILL



WALK



COFFEE OR TEA

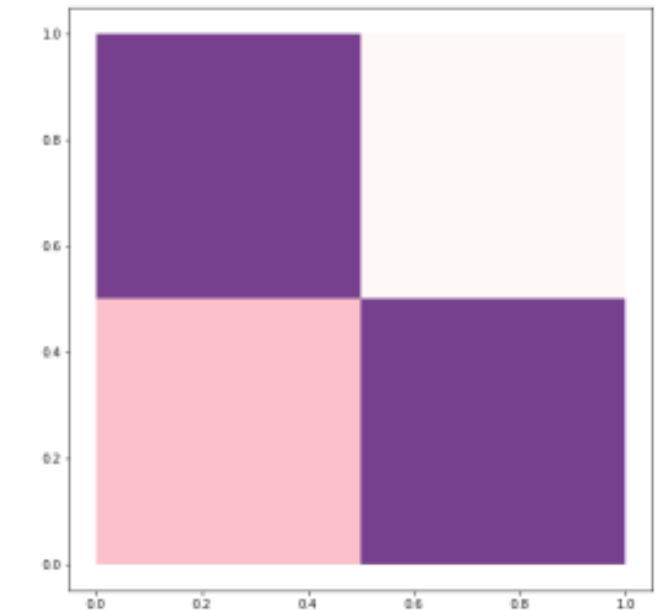
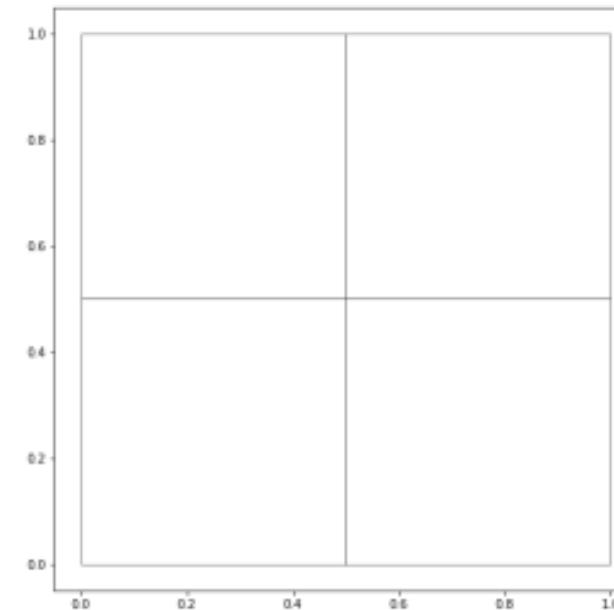
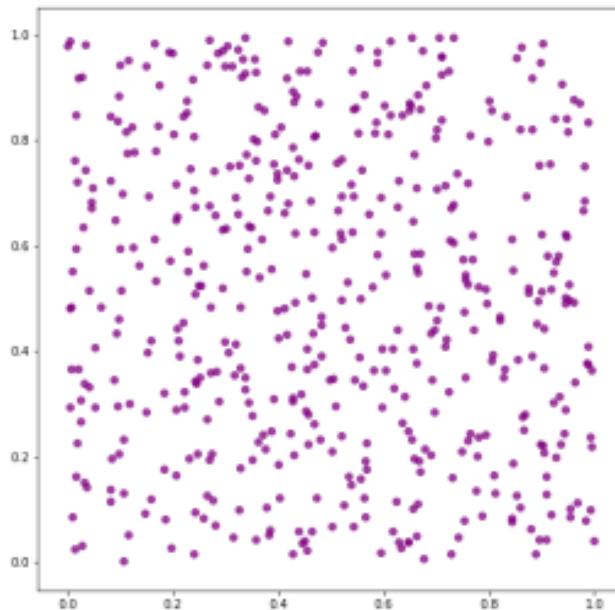


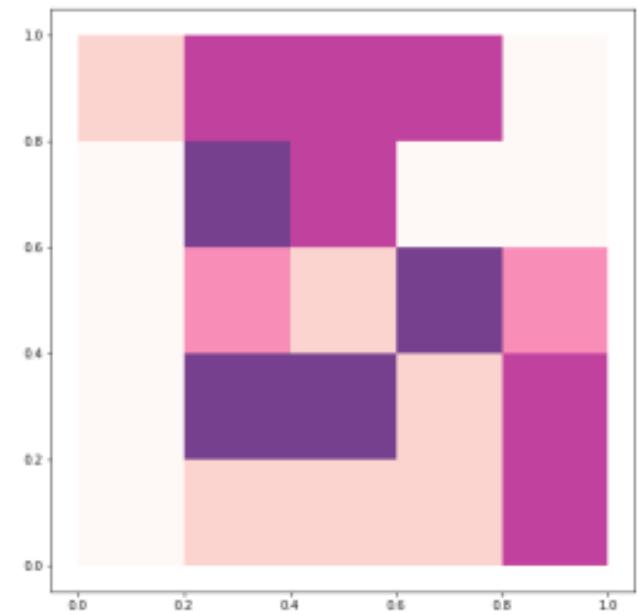
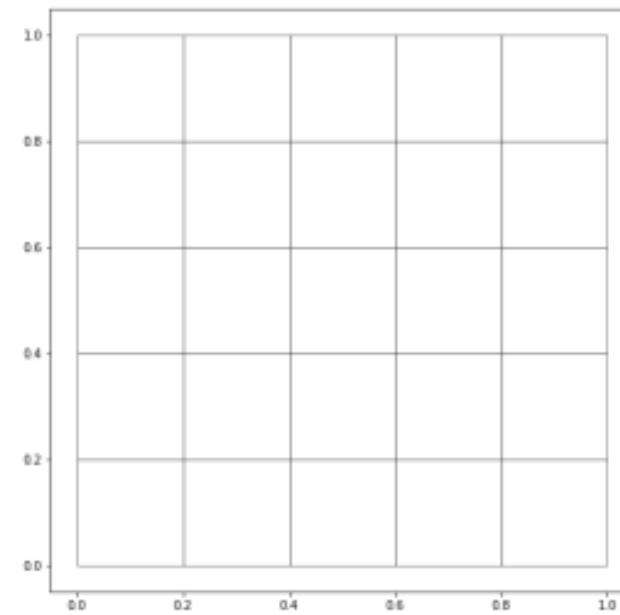
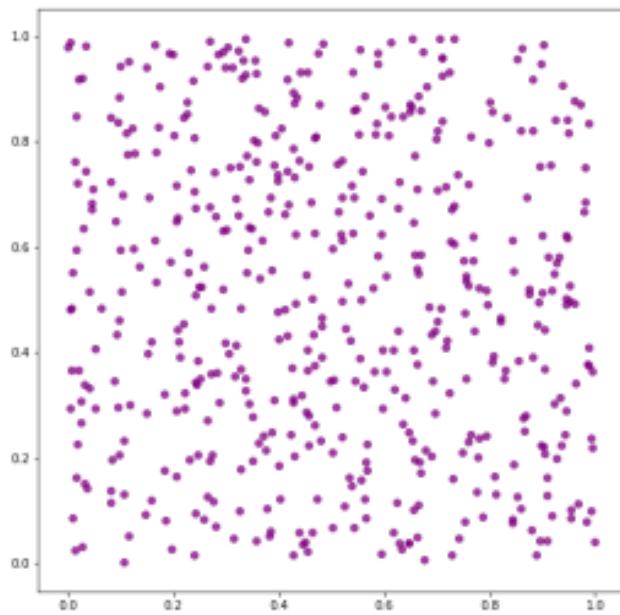
MAKE FRIENDS

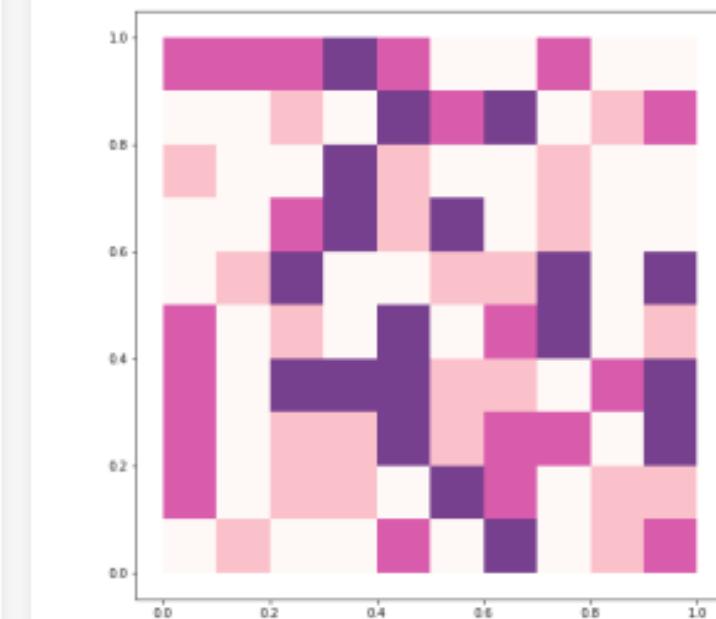
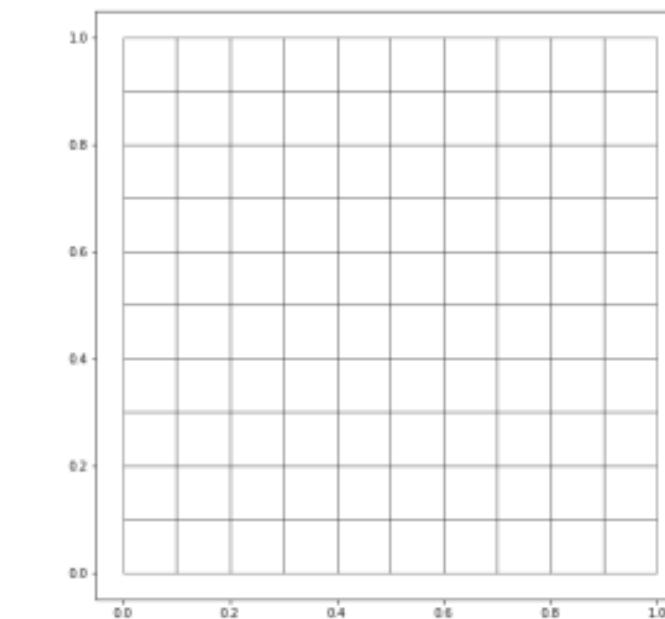
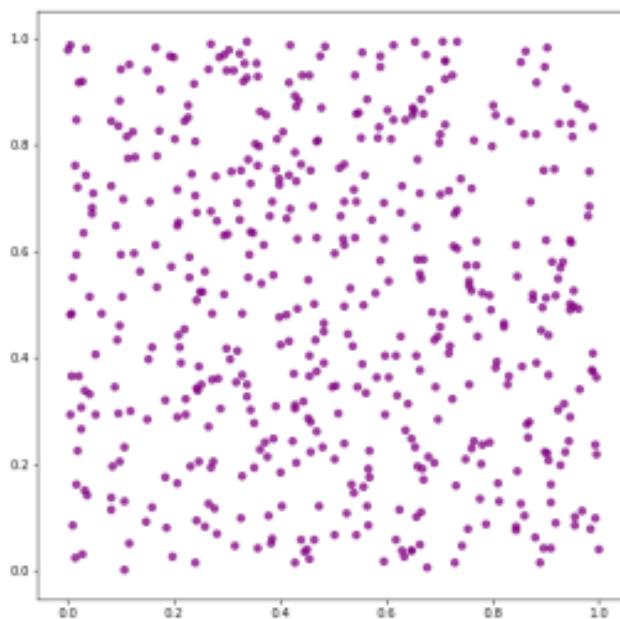
# Making good data maps

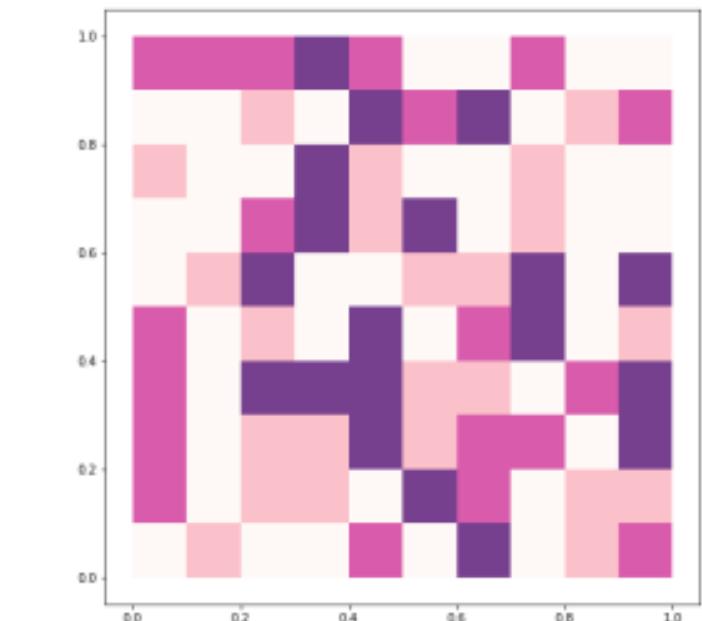
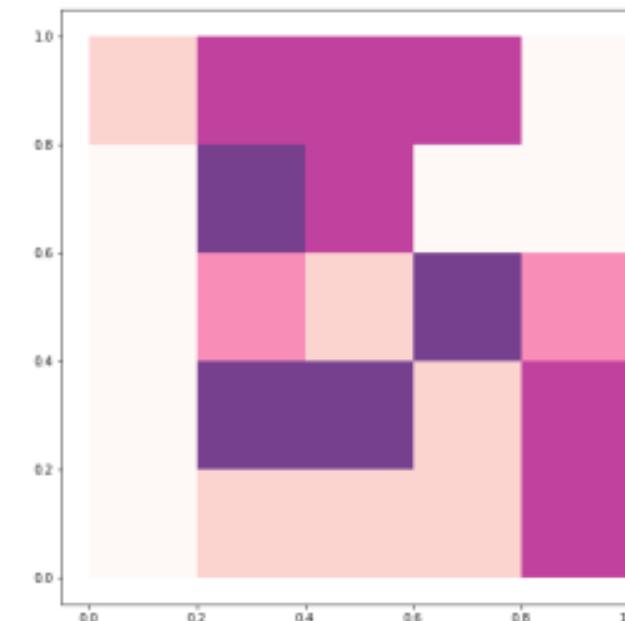
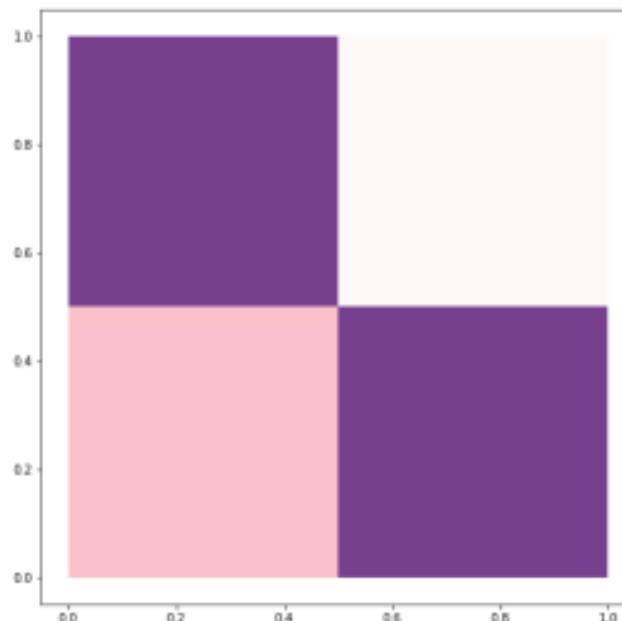
- MAUP
- Choropleths

# Modifiable Areal Unit Problem (Openshaw, 1984)







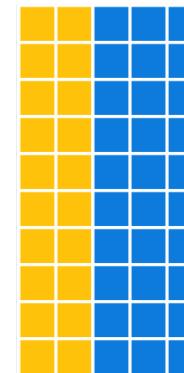


# Why is it a problem?

Gerrymandering: drawing different maps for electoral districts produces different outcomes

**Disproportionate Outcomes  
“gerrymandering”**

50 Precincts  
60% Blue  
40% Yellow

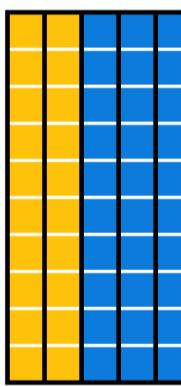


**5 DISTRICTS**  
5 Blue  
0 Yellow  
**BLUE WINS ALL**

**5 DISTRICTS**  
3 Yellow  
2 Blue  
**YELLOW WINS MAJORITY**

**Proportionate Outcomes**

50 Precincts  
are to be  
apportioned  
into  
5 districts,  
10 precincts  
per district.



**5 DISTRICTS**  
3 Blue  
2 Yellow

**5 DISTRICTS**  
3 Blue  
2 Yellow

**Blue and yellow win in proportion to their voting**

# MAUP

**Scale and delineation mismatch** between:

- Point-based measures
  - Underlying process (e.g. individuals, firms, shops)
- Aggregated in space
  - Unit of measurement (e.g. neighborhoods, regions, etc.)

In some cases, it can **seriously mislead** analysis on aggregated data  
(e.g. [FLINT, MI](#))

Always keep MAUP in mind when exploring aggregated data!!!

# Choropleths

# Choropleths

*Thematic map in which values of a variable are encoded using a colour gradient of some sort*

- Counterpart of the histogram
- **Values are classified** into specific colours: value → bin
- **Information loss** as a trade off for **simplicity**

# Classification Choices

- No. of bins
- How to bin?
- Colours

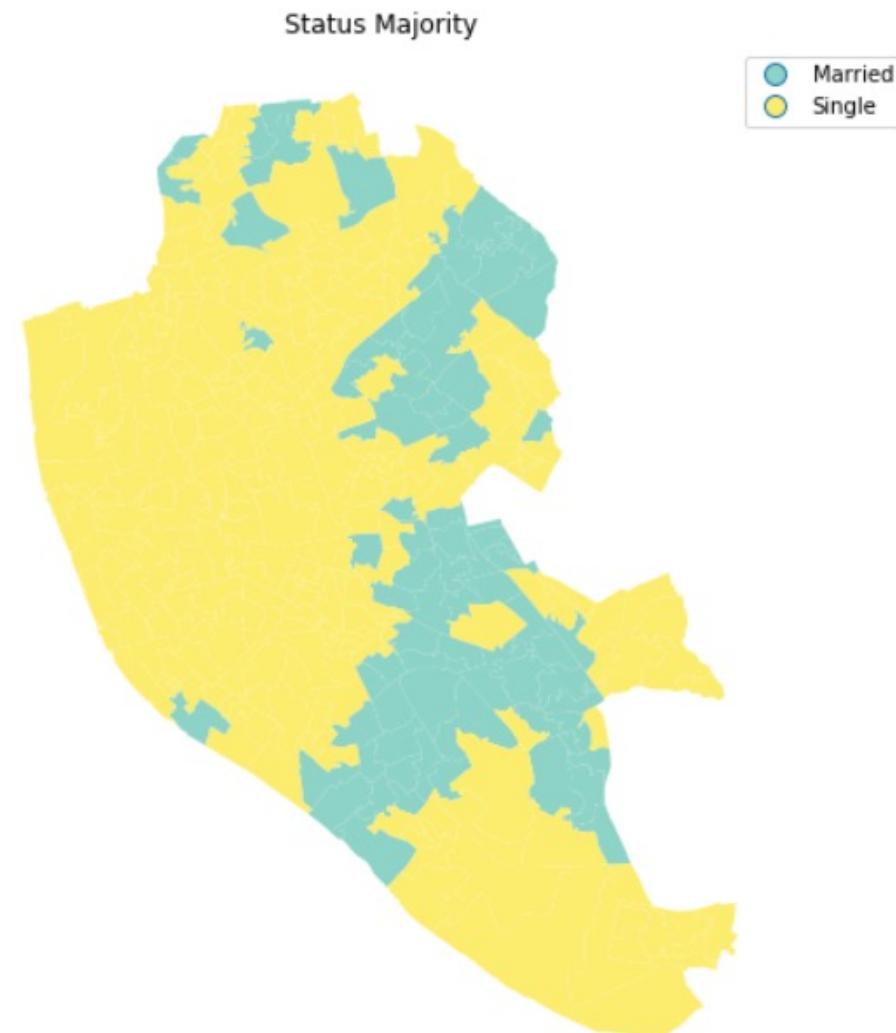
# How many bins

- Trade-off: detail vs cognitive load
- Exact number depends on purpose of the map
- Usually not more than 12

# How to bin?

# Unique values

- Categorical data
- No gradient (reflect it with the colour scheme!!!)
- Examples: Religion, country of origin...

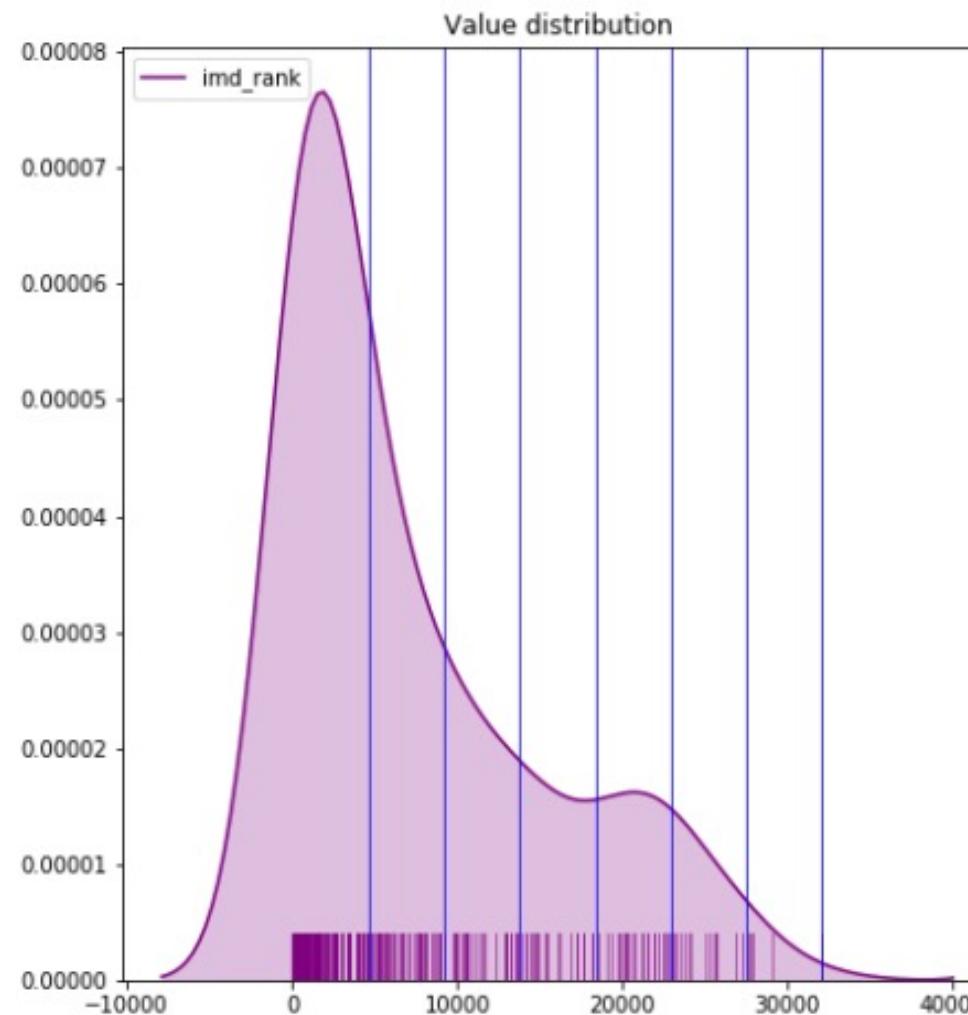


Images taken from: Arribas-Bel, D. (2019). A course on geographic data science. *Journal of Open Source Education*, 2(16), 42.

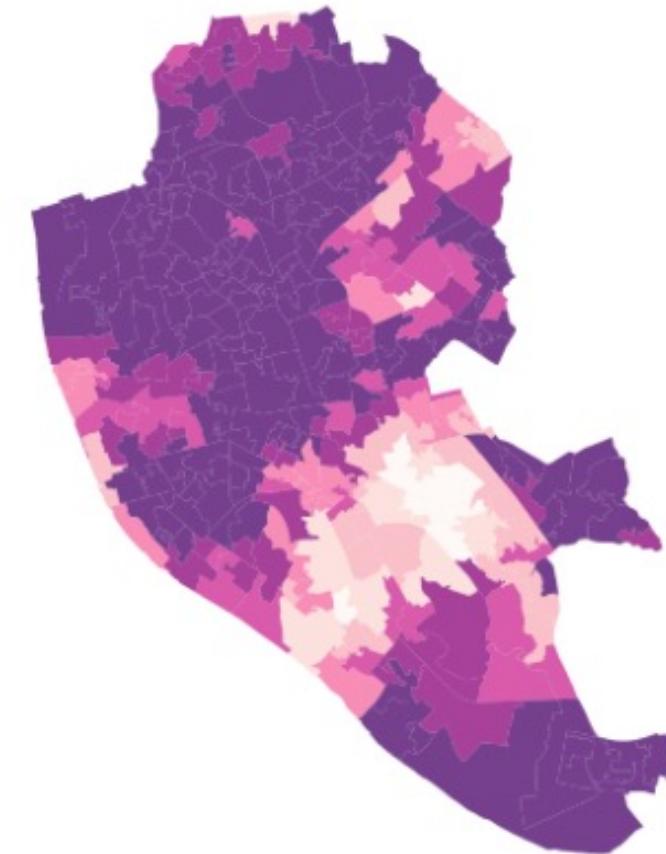
# Equal interval

- Take the **value** span of the data to represent and split it equally
- **Splitting** happens based on the **numerical value**
- Gives more weight to outliers if the distribution is skewed

## equal\_interval



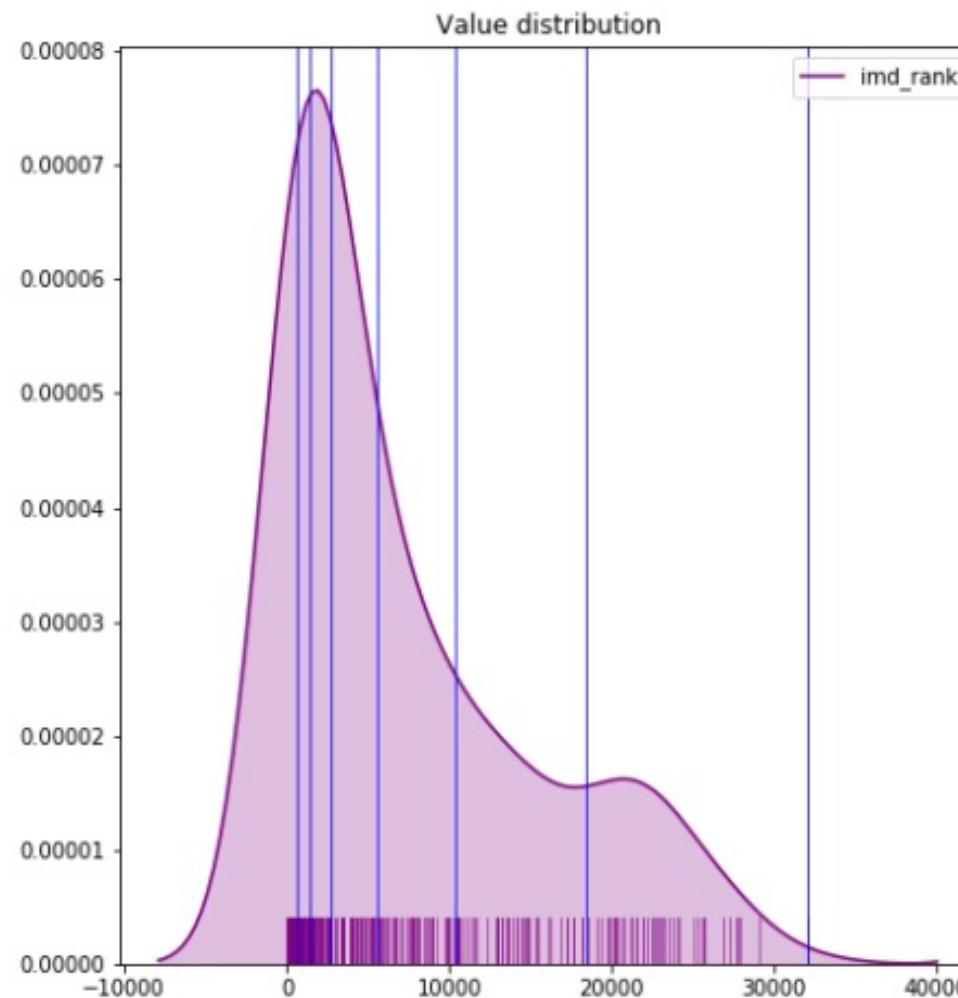
Geographical distribution



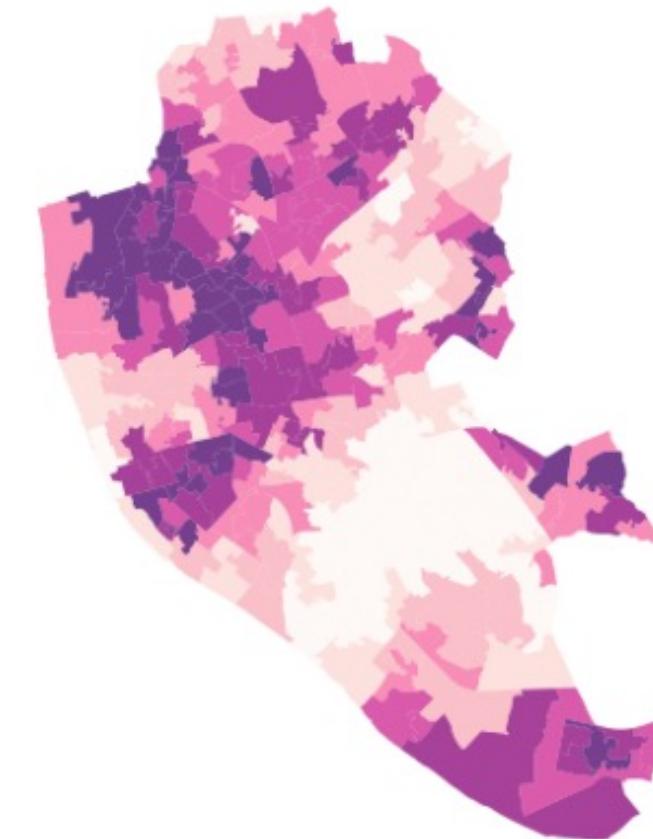
# Quantiles

- Regardless of numerical values, split the distribution keeping the same number of values in each bin
- **Splitting** based on the **rank** of the value
- If distribution is skewed, it can put very different values in the same bin

## quantiles

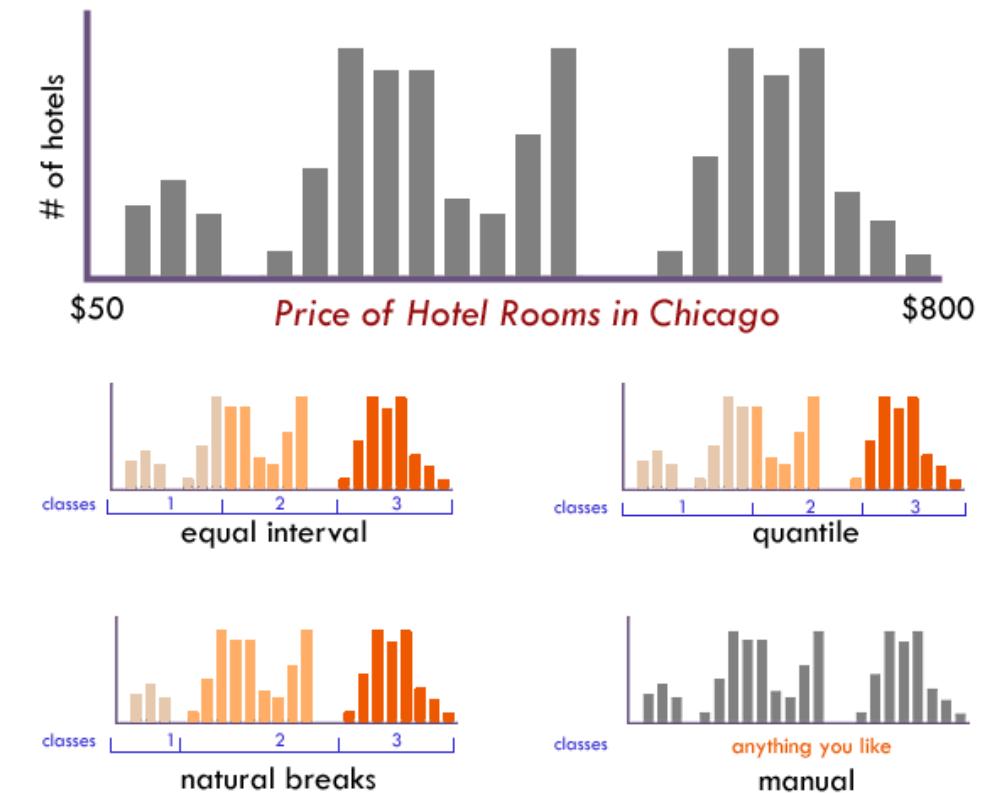


Geographical distribution



# + Other..

- Unique
- Equal Interval
- Quantiles
- Fisher-Jenks or Natural breaks
- Manual



[\[Important – please read me: Basics of Data Classification on Maps\]](#)

# Colour Schemes

Align with your purpose

- **Categories**, non-ordered



- Graduated, **sequential**



- Graduated, **divergent**



TIP: check [ColorBrewer](#) for guidance

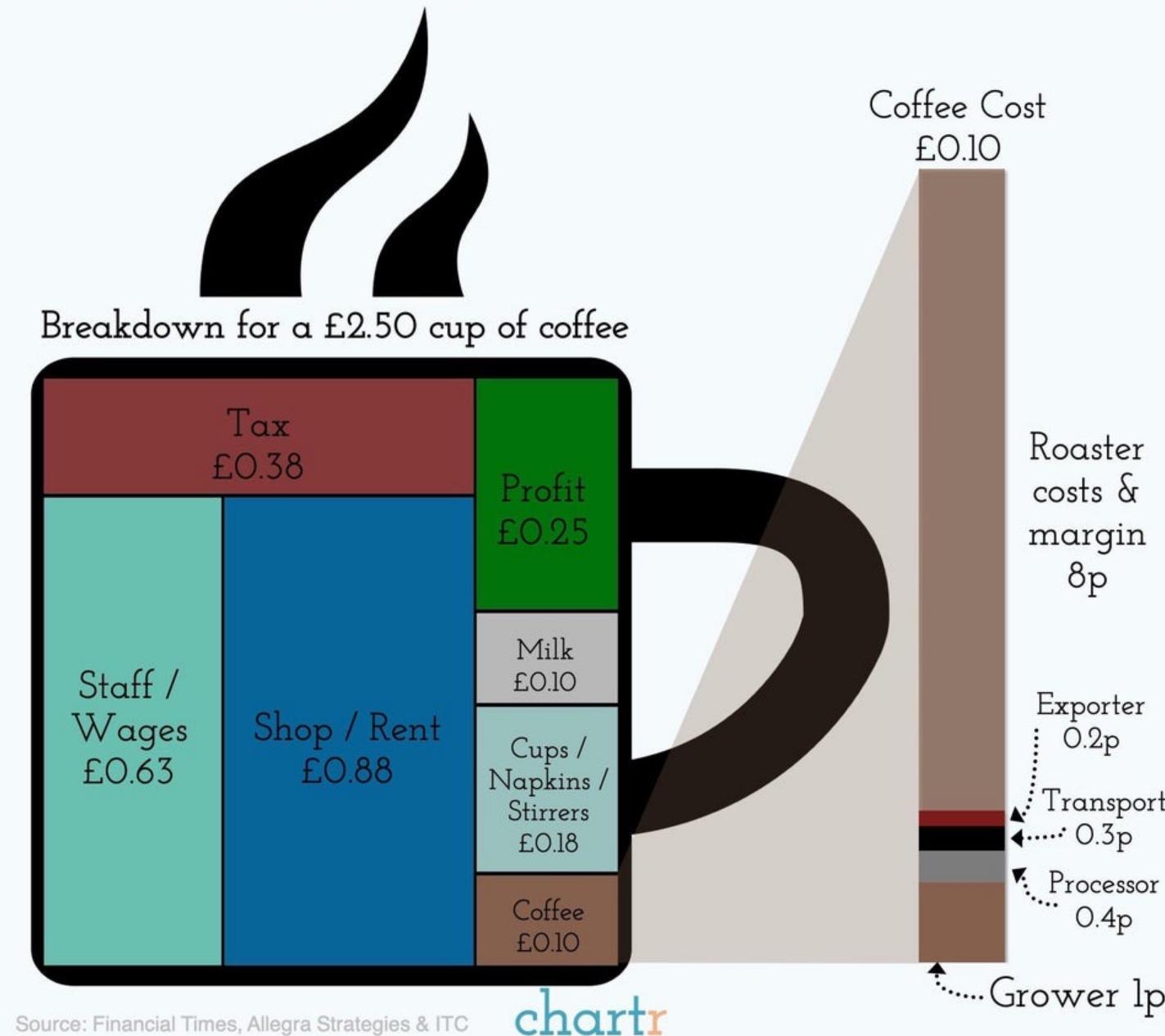
# Tips

- Think of the purpose of the map
- Explore by trying different classification alternatives
- Combine (Geo)-visualisation with other statistical devices

# The Economics behind Coffee

## Recap

- Visualization of statistical data is a recent phenomenon.
- Needs of the audience are key.
- Data can tell more than one story at a time.
- Its power comes from engaging and augmenting the human in the loop, rather than replacing her.
- Its power can be misused, so think twice.



# For next class..



**Finish** Labs to practice programming



**Complete** Homework and review your peers' work



**Check** Assignment contents and due date



**See** "To do before class" for next lecture (~ 1 hour of self study)



**No Discussion** next week (instead lab hours)



**Post** questions on the **Discussion** forum on Brightspace