

# Session 2

## Theory behind visualisations



# Visualization basics & an introduction to Tableau



## Data

Facts and statistics collected together for reference or analysis.



*Data at Scale*

# Visualization basics & an introduction to Tableau



## Data

Facts and statistics collected together for reference or analysis.

## Information

Facts provided or learned about something or someone.



*Data at Scale*

Dr Evgeniya Lukinova

# Visualization basics & an introduction to Tableau



## Data

Facts and statistics collected together for reference or analysis.

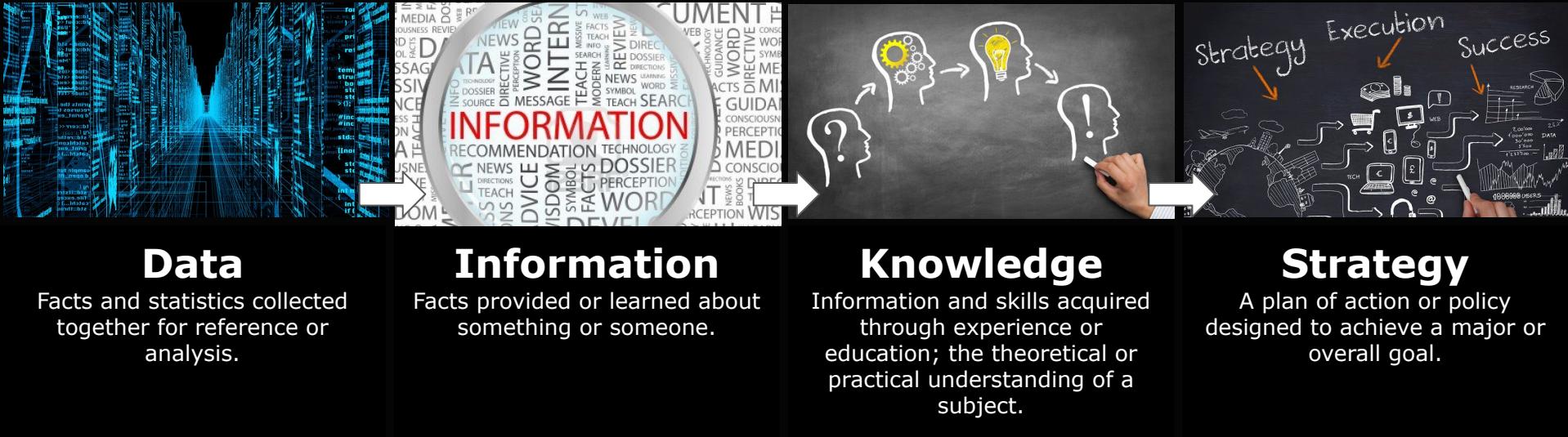
## Information

Facts provided or learned about something or someone.

## Knowledge

Information and skills acquired through experience or education; the theoretical or practical understanding of a subject.

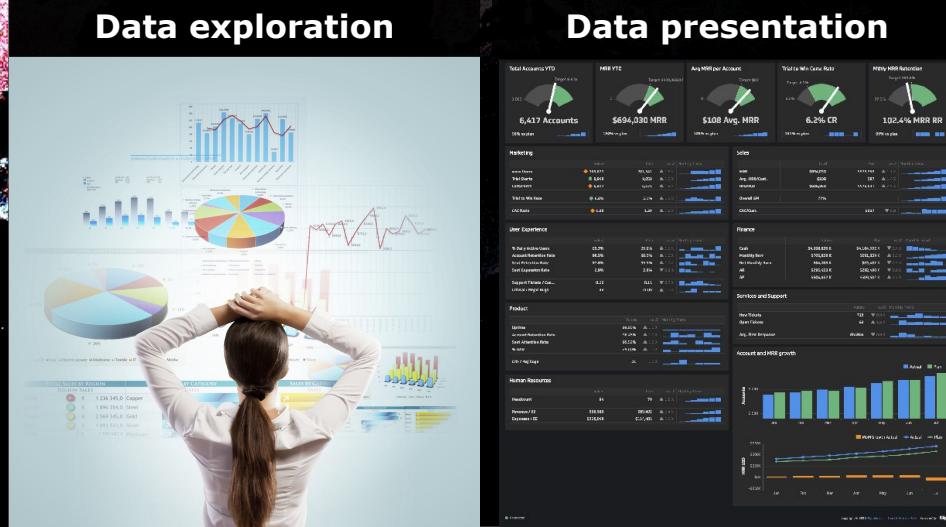
# Visualization basics & an introduction to Tableau

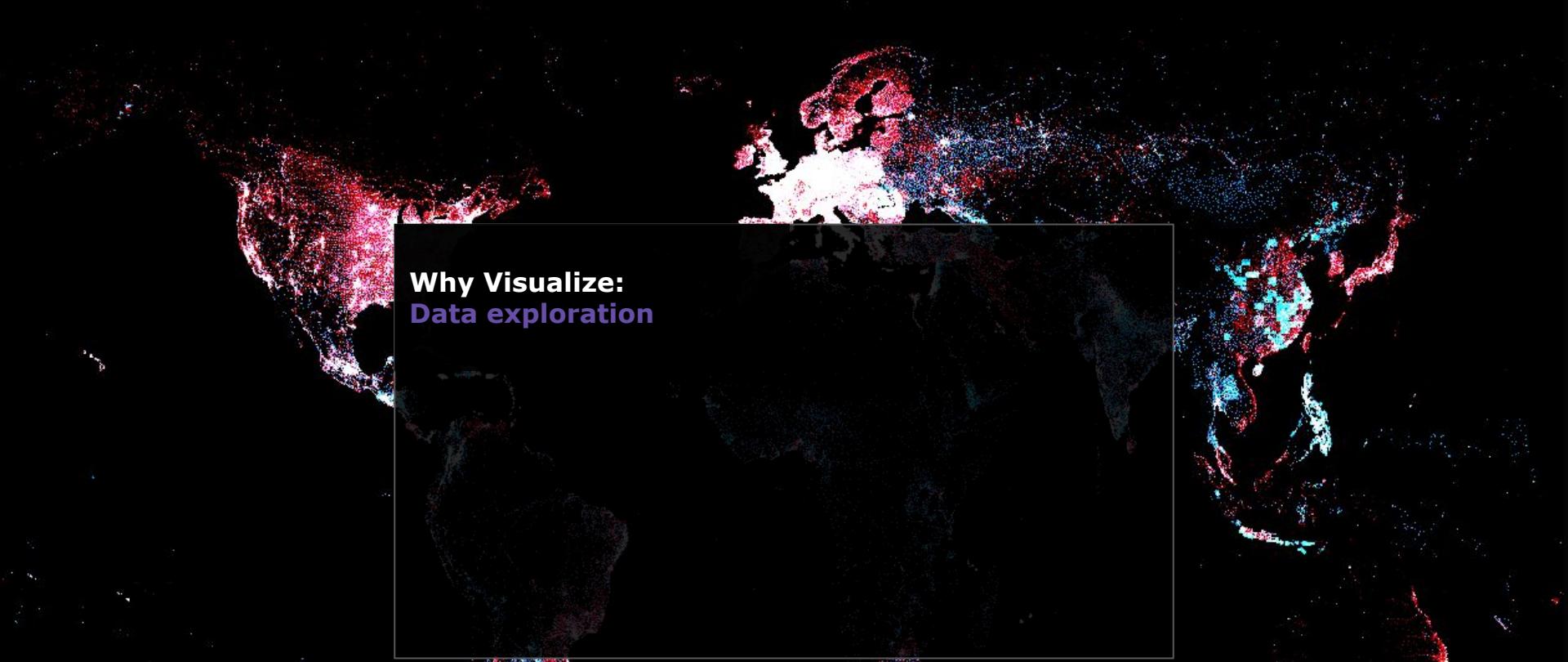


*Data at Scale*

Dr Evgeniya Lukinova

# Visualization basics & an introduction to Tableau





**Why Visualize:  
Data exploration**



*Data at Scale*

Dr Evgeniya Lukinova

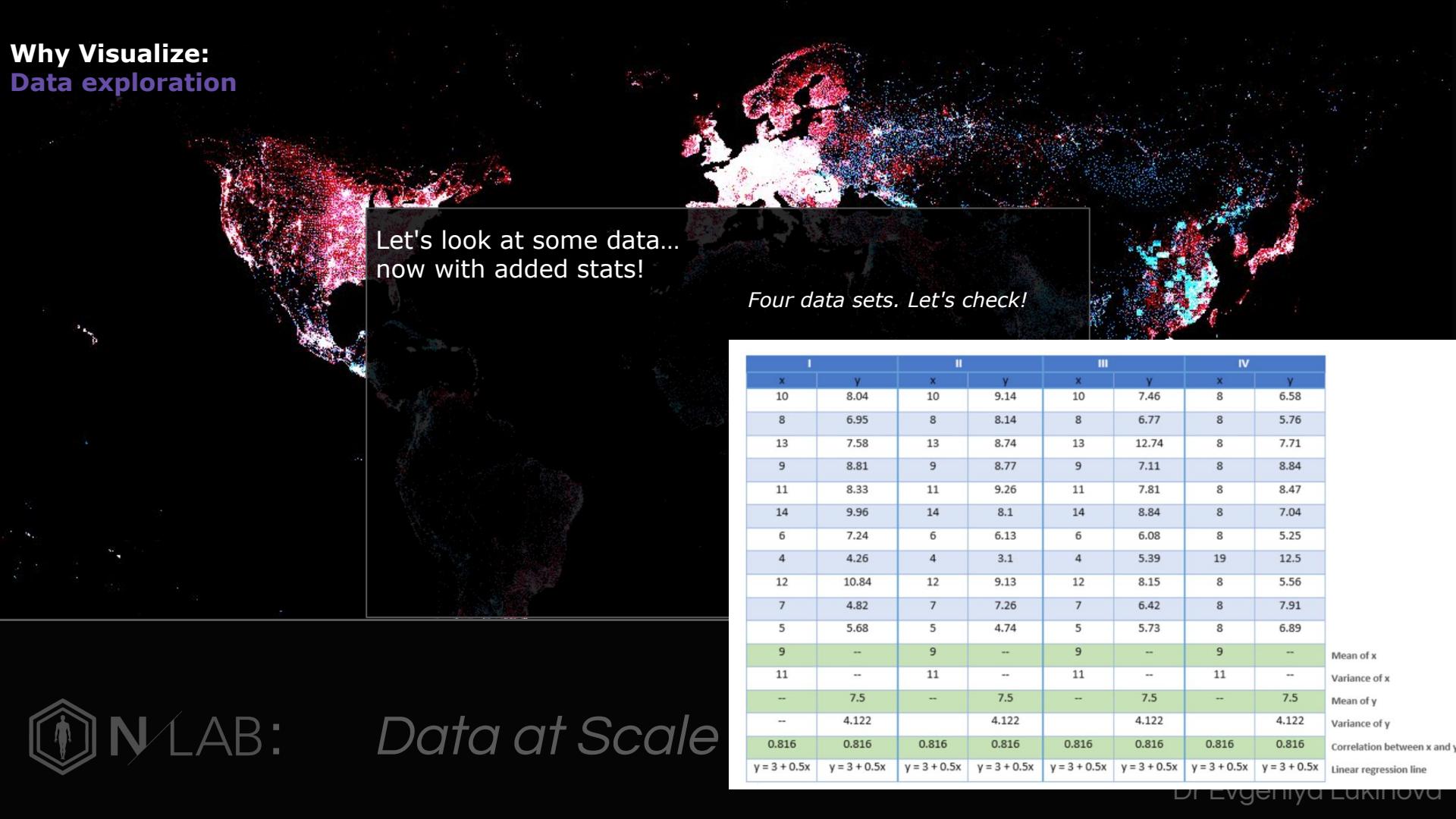
## Why Visualize: Data exploration

Let's look at some data...

*Four data sets. Are they the same?*

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

# Why Visualize: Data exploration



Let's look at some data...  
now with added stats!

Four data sets. Let's check!

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89
9	--	9	--	9	--	9	--
11	--	11	--	11	--	11	--
--	7.5	--	7.5	--	7.5	--	7.5
--	4.122		4.122		4.122		4.122
0.816	0.816	0.816	0.816	0.816	0.816	0.816	0.816
$y = 3 + 0.5x$							

Mean of x  
Variance of x  
Mean of y  
Variance of y  
Correlation between x and y  
Linear regression line



Data at Scale

# Why Visualize: Data exploration

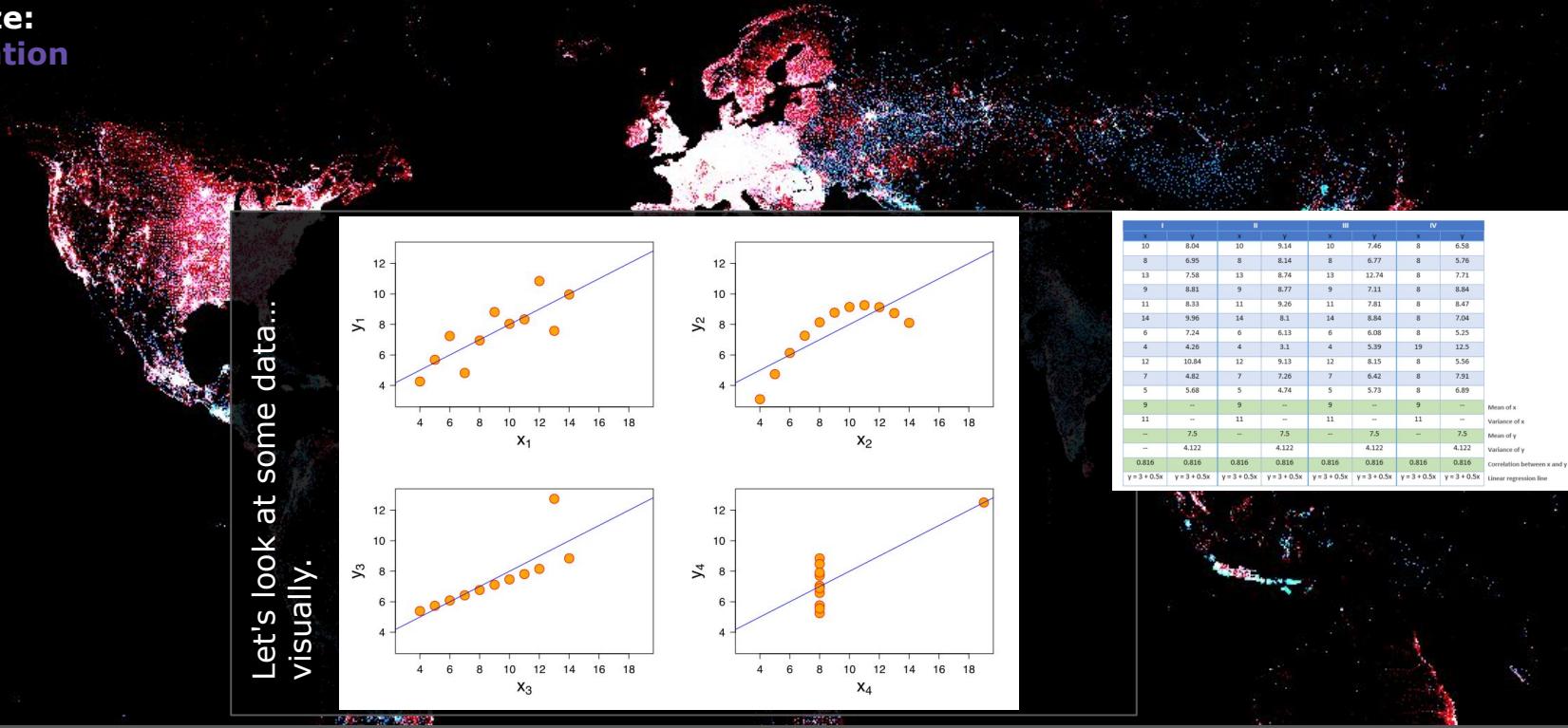


Image source Wikipedia. CC BY-SA 3.0. Anscombe's quartet.

## Why Visualize: Data exploration



*Data at Scale*

## Why Visualize: Data exploration

So

**summary stats < visualization??**



*Data at Scale*

## Why Visualize: Data exploration

So

**summary stats < visualization??**



→ Compress data to focus  
on one data property

→ for the given summary  
statistic the whole data set  
(high dimensional or not) is  
correctly represented (by  
definition)

## Why Visualize: Data exploration

→ need to choose  
the right summary  
statistic

So  
**summary stats < visualization??**



- Compress data to focus  
on one data property
- for the given summary  
statistic the whole data set  
(high dimensional or not) is  
correctly represented (by  
definition)

## Why Visualize: Data exploration

→ need to choose  
the right summary  
statistic

So  
**summary stats < visualization??**

→ Compress data to focus  
on one data property

→ for the given summary  
statistic the whole data set  
(high dimensional or not) is  
correctly represented (by  
definition)

→ Displays a range of data  
dimensions

→ cannot accurately show  
high-dimensional and/or  
large data sets (often must  
visualize summary  
statistics)

## Why Visualize: Data exploration

→ need to choose  
the right summary  
statistic

So  
**summary stats < visualization??**

→ Compress data to focus  
on one data property

→ for the given summary  
statistic the whole data set  
(high dimensional or not) is  
correctly represented (by  
definition)

→ Displays a range of data  
dimensions

→ cannot accurately show  
high-dimensional and/or  
large data sets (often must  
visualize summary  
statistics)

→ need to choose the  
dimensions (features/attributes)  
to visualize

## Why Visualize: Data exploration

→ need to choose  
the right summary  
statistic

So

### summary stats < visualization??

→ Compress data to focus  
on one data property

→ for the given summary  
statistic the whole data set  
(high dimensional or not) is  
correctly represented (by  
definition)

→ Displays a range of data  
dimensions

→ cannot accurately show  
high-dimensional and/or  
large data sets (often must  
visualize summary  
statistics)

→ need to choose the  
dimensions (features/attributes)  
to visualize

→ need to choose the  
right visualization  
(highlights different properties of the  
dataset ≈ summary stats)

## Why Visualize: Data exploration

→ need to choose  
the right summary  
statistic

So

### summary stats < visualization??

→ Compress data to focus  
on one data property

→ for the given summary  
statistic the whole data set  
(high dimensional or not) is  
correctly represented (by  
definition)

→ Displays a range of data  
dimensions

→ cannot accurately show  
high-dimensional and/or  
large data sets (often must  
visualize summary  
statistics)

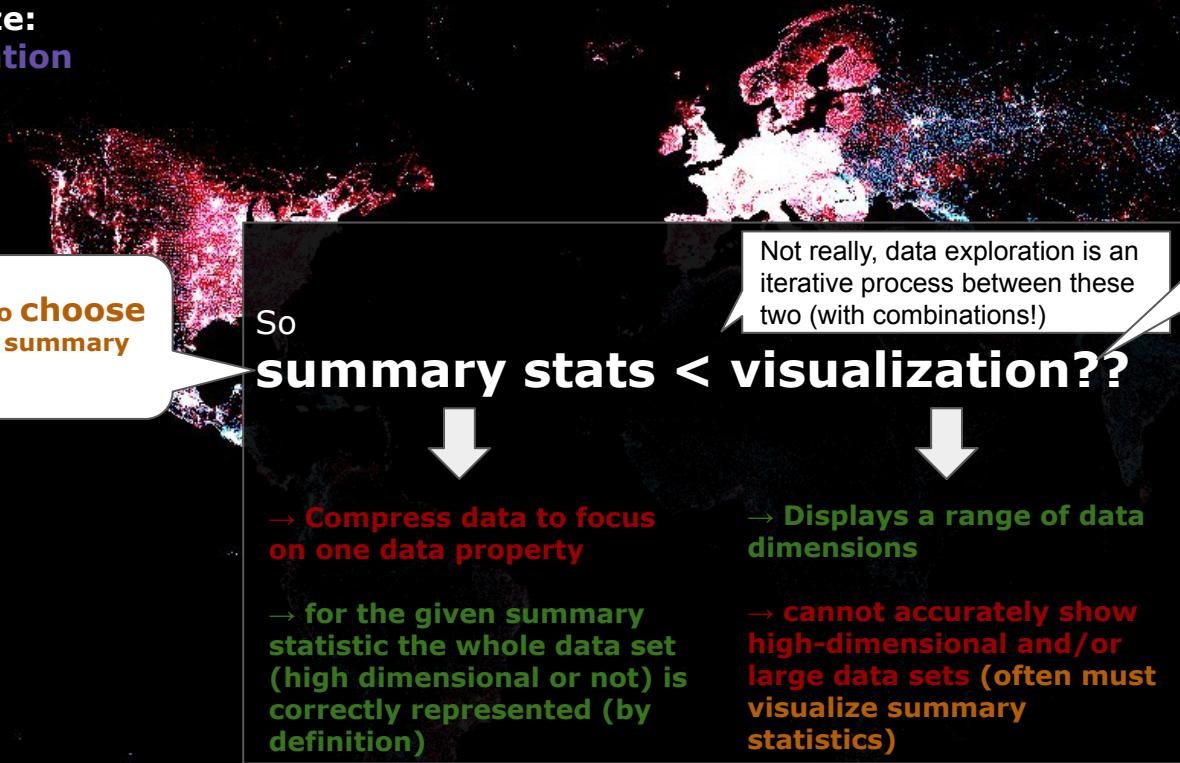
→ need to choose the  
dimensions (features/attributes)  
to visualize

→ need to choose the  
right visualization  
(highlights different properties of the  
dataset ≈ summary stats)

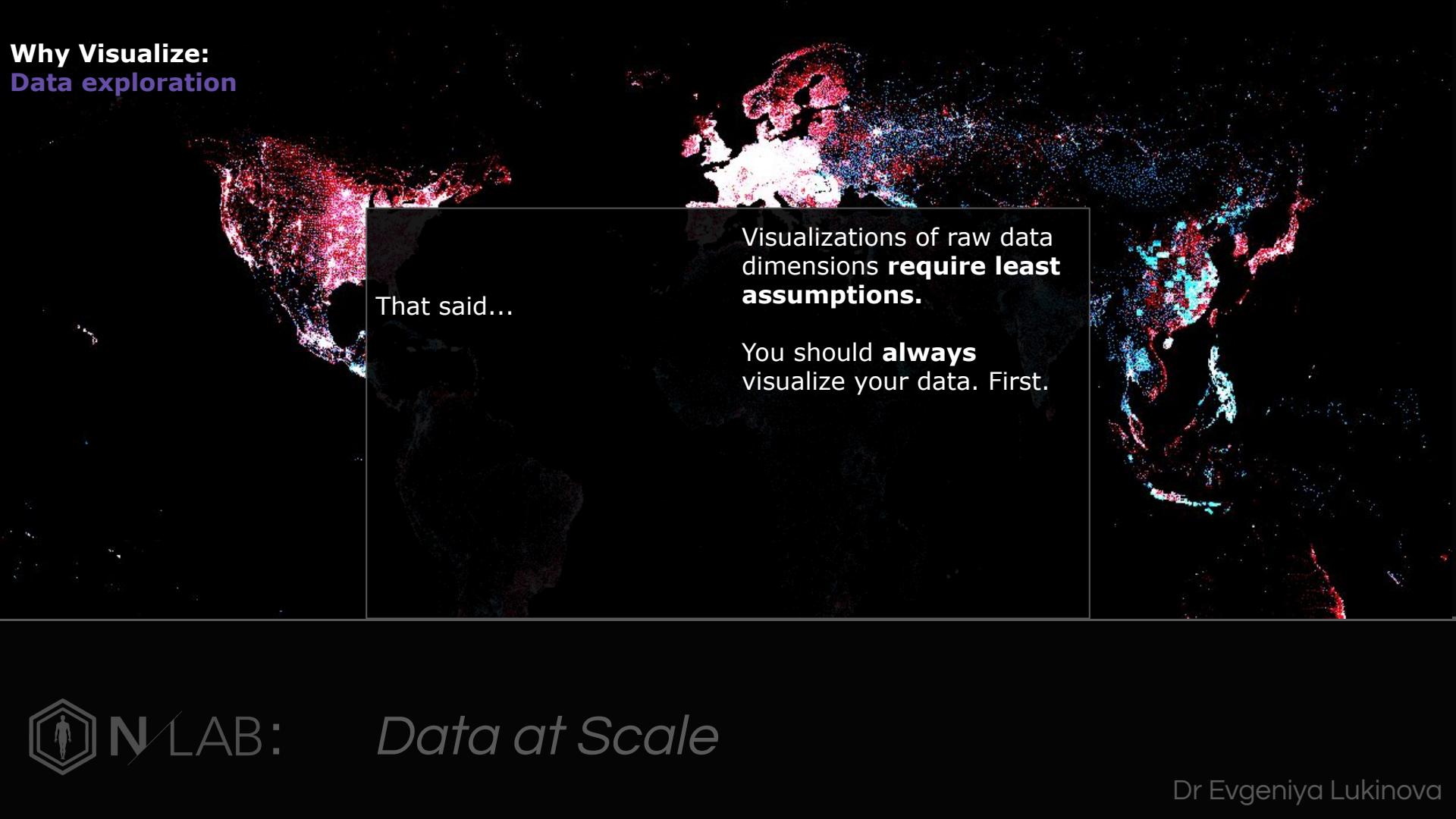
→ Can't visualize every aspect  
at once: both summary stats &  
visualizations compress the  
data for interpretation.

→ Eventually for presentation  
you'll be deciding on the "best"  
compression (for your story)

# Why Visualize: Data exploration



## Why Visualize: Data exploration



That said...

Visualizations of raw data dimensions **require least assumptions.**

You should **always** visualize your data. First.



*Data at Scale*

## Why Visualize: Data exploration



That said...

Visualizations of raw data dimensions **require least assumptions.**

You should **always** visualize your data. First.

Visualization can play a **critical role** in helping you figure out what the **interesting questions** are.





Figure from: Tor Norretranders' The User Illusion. Visualisation from Stephen Few's Information Dashboard Design



*Data at Scale*

Dr Evgeniya Lukinova

# Why Visualize: Data presentation

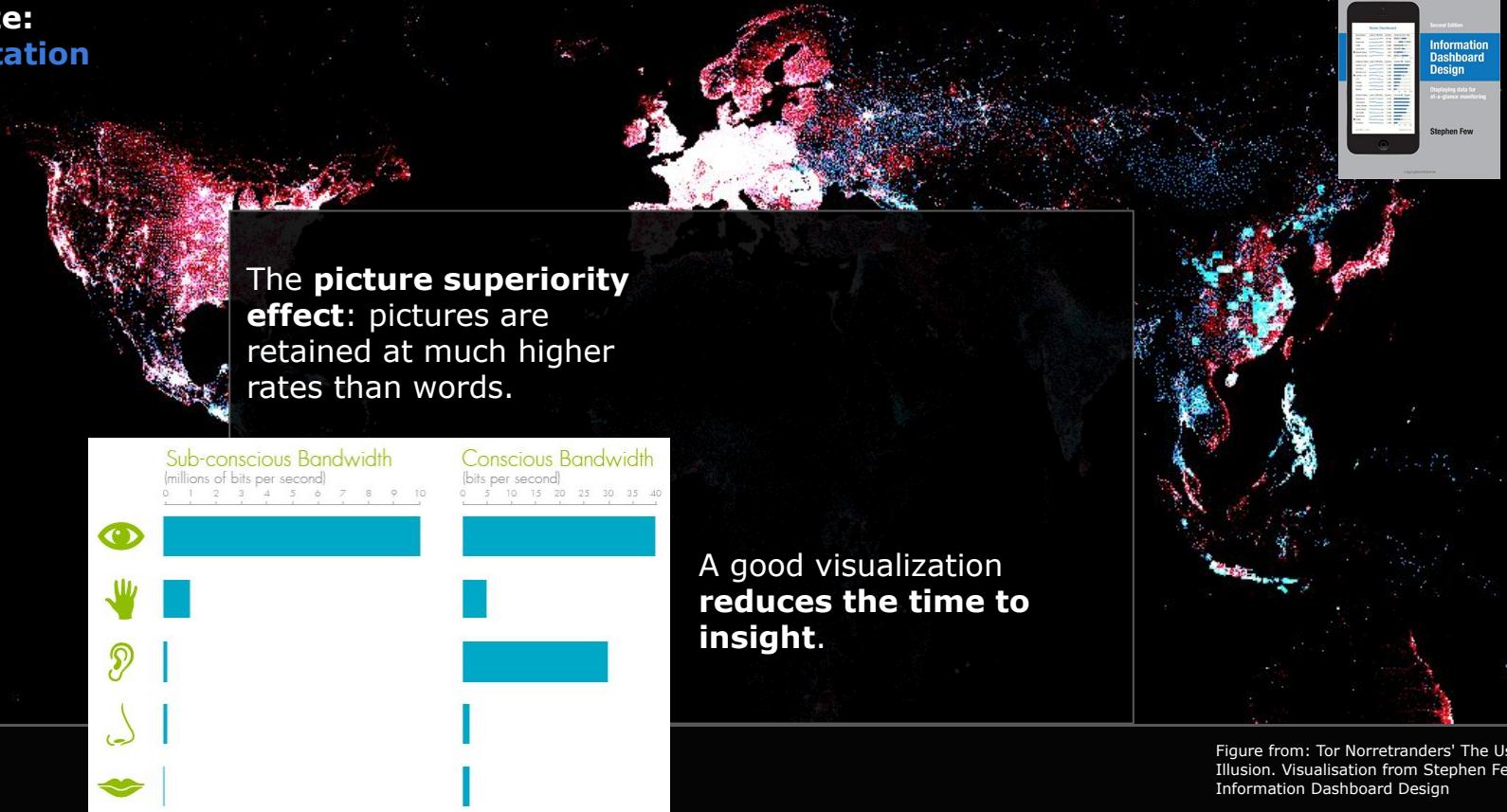


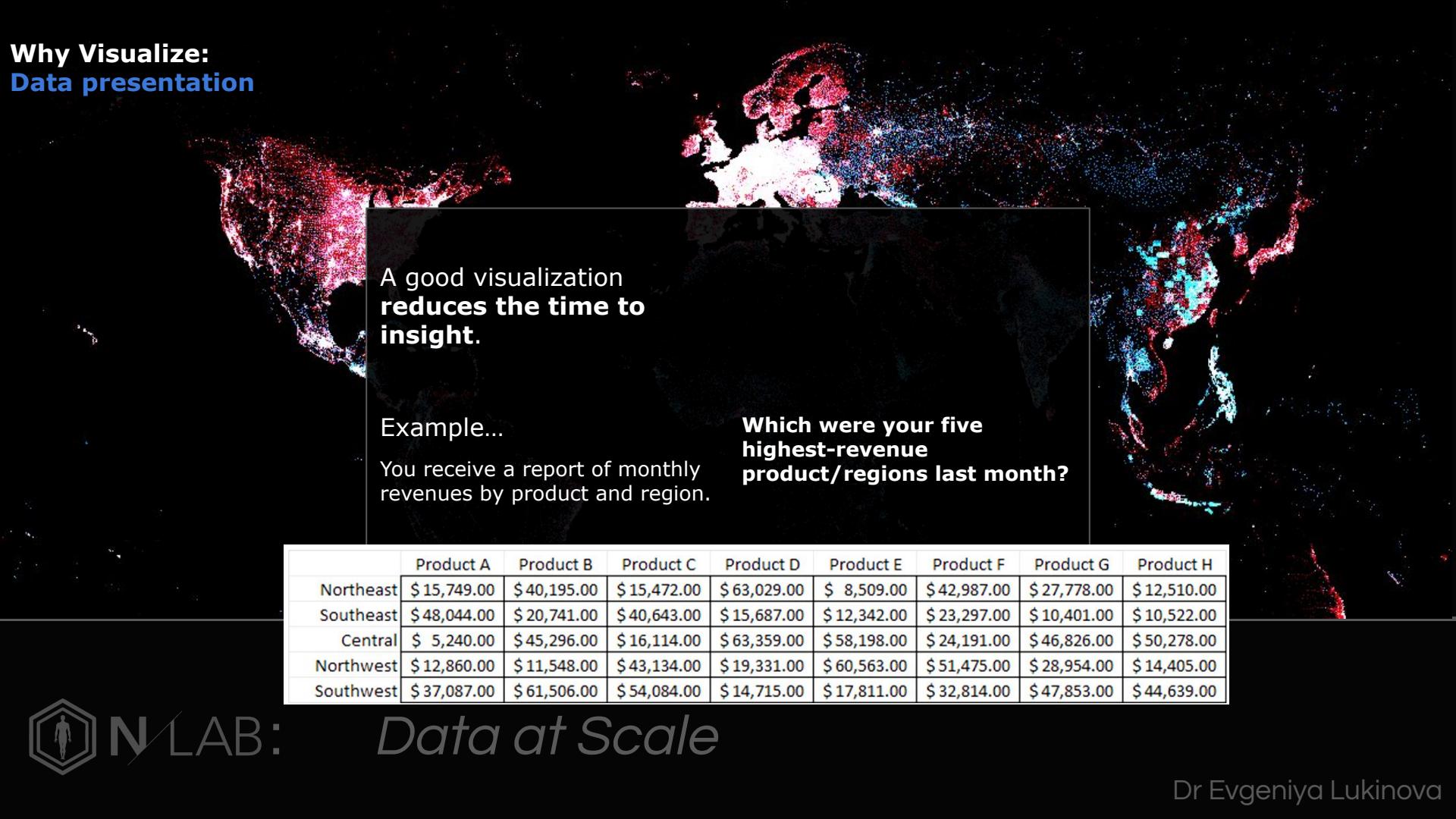
Figure from: Tor Norretranders' The User Illusion. Visualisation from Stephen Few's Information Dashboard Design



*Data at Scale*

Dr Evgeniya Lukinova

# Why Visualize: Data presentation



A good visualization  
**reduces the time to  
insight.**

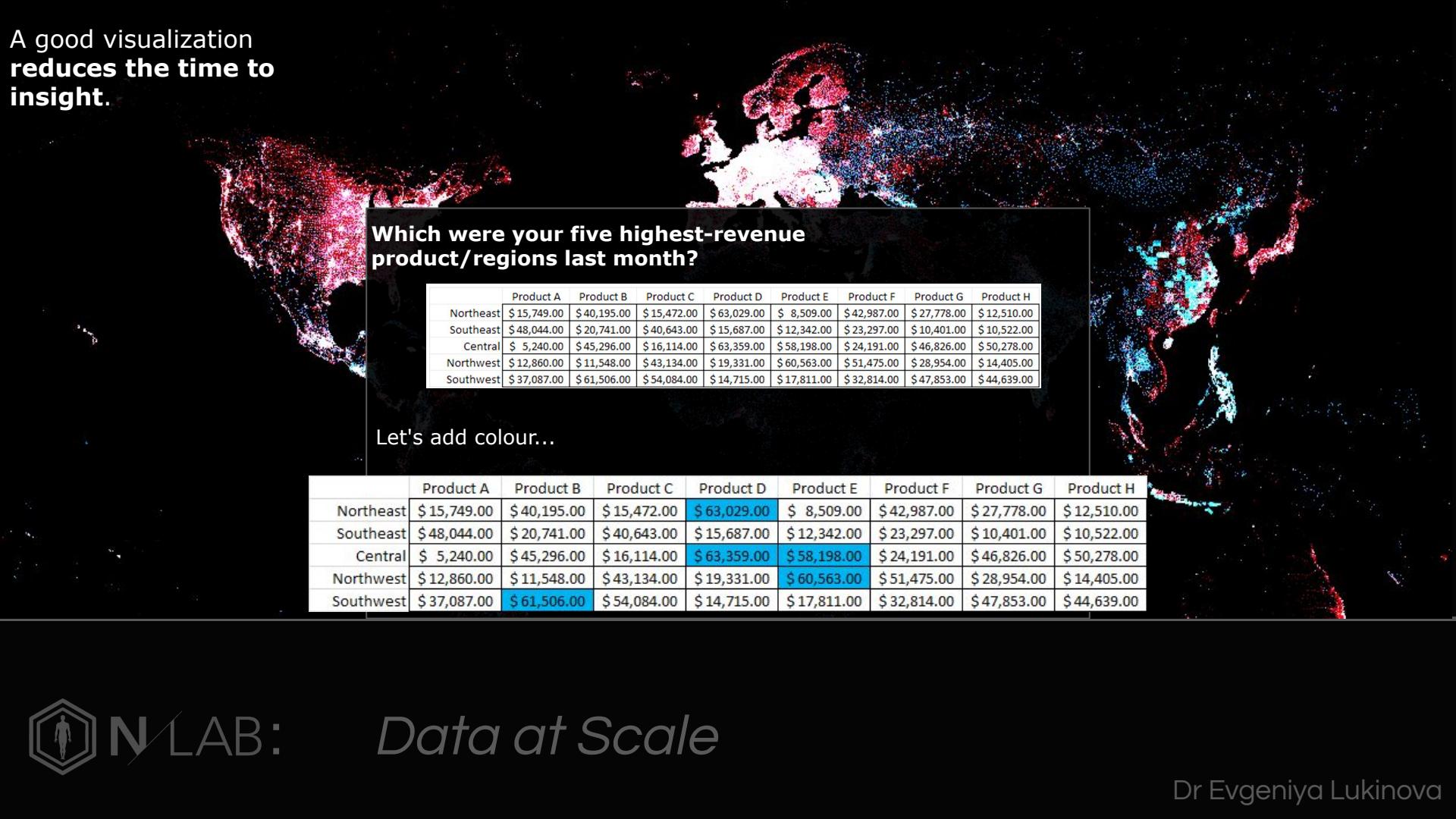
Example...

You receive a report of monthly revenues by product and region.

**Which were your five  
highest-revenue  
product/regions last month?**

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

A good visualization  
**reduces the time to**  
insight.



Which were your five highest-revenue product/regions last month?

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

Let's add colour...

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

A good visualization  
**reduces the time to**  
insight.

Which were your five highest-revenue  
product/regions last month?

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

Let's add colour...

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	<b>\$ 63,029.00</b>	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	<b>\$ 63,359.00</b>	<b>\$ 58,198.00</b>	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	<b>\$ 60,563.00</b>	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	<b>\$ 61,506.00</b>	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

Color automatically  
focused your brain –  
no mental calculations /  
comparisons needed!

A good visualization  
**reduces the time to**  
insight.

Which were your five highest-revenue  
product/regions last month?

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast	\$ 15,749.00	\$ 40,195.00	\$ 15,472.00	\$ 63,029.00	\$ 8,509.00	\$ 42,987.00	\$ 27,778.00	\$ 12,510.00
Southeast	\$ 48,044.00	\$ 20,741.00	\$ 40,643.00	\$ 15,687.00	\$ 12,342.00	\$ 23,297.00	\$ 10,401.00	\$ 10,522.00
Central	\$ 5,240.00	\$ 45,296.00	\$ 16,114.00	\$ 63,359.00	\$ 58,198.00	\$ 24,191.00	\$ 46,826.00	\$ 50,278.00
Northwest	\$ 12,860.00	\$ 11,548.00	\$ 43,134.00	\$ 19,331.00	\$ 60,563.00	\$ 51,475.00	\$ 28,954.00	\$ 14,405.00
Southwest	\$ 37,087.00	\$ 61,506.00	\$ 54,084.00	\$ 14,715.00	\$ 17,811.00	\$ 32,814.00	\$ 47,853.00	\$ 44,639.00

For **some applications**, eliminating all distractions  
with a very simple viz can be effective:

	Product A	Product B	Product C	Product D	Product E	Product F	Product G	Product H
Northeast								
Southeast								
Central								
Northwest								
Southwest								

1) However, removing the  
base data removes  
information for reader.

2) Now the info is gone  
perhaps a list would simply  
be better?

A good visualization  
**reduces the time to**  
insight.



How many nines are there...

4	7	7	5	5	2	7	4	7	1
4	9	2	5	7	7	2	6	1	7
1	7	6	9	3	4	7	5	1	2
5	1	6	3	3	8	4	8	6	6
6	5	6	4	9	3	8	9	1	9
3	8	1	5	2	2	3	6	3	9
4	6	4	5	6	3	7	7	9	1
9	1	3	3	6	1	3	3	1	8
8	1	1	8	7	5	8	1	7	4
3	6	9	2	8	9	3	7	5	7
4	4	4	2	8	2	2	9	2	8

A good visualization  
**reduces the time to**  
insight.

How many nines are there...

4	7	7	5	5	2	7	4	7	1
4	9	2	5	7	7	2	6	1	7
1	7	6	9	3	4	7	5	1	2
5	1	6	3	3	8	4	8	6	6
6	5	6	4	9	3	8	9	1	9
3	8	1	5	2	2	3	6	3	9
4	6	4	5	6	3	7	7	9	1
9	1	3	3	6	1	3	3	1	8
8	1	1	8	7	5	8	1	7	4
3	6	9	2	8	9	3	7	5	7
4	4	4	2	8	2	2	9	2	8



*Data at Scale*

Dr Evgeniya Lukinova

A good visualization  
**reduces the time to**  
insight.



So colour can be used to  
**highlight data** and reduce time  
to insight.

While providing the same data.  
**Guides interpretation.**

But....

A good visualization  
**reduces the time to**  
insight.

### Using colour must be done with care...

"Color used well can enhance and beautify, but color used poorly can be worse than no color at all." Maureen Stone

"...avoiding catastrophe becomes the first principle in bringing color to information: Above all, do no harm."  
(Envisioning Information, Edward Tufte, Graphics Press, 1990)

A good visualization  
**reduces the time to**  
insight.

### Using colour must be done with care...

- ~ 8% of men worldwide are colour blind
  - Be nice, **avoid red/green palettes**
  - Blue/orange is a good alternative

A good visualization  
**reduces the time to**  
insight.

Using colour must be done with care...

Not a natural designer? Use "known good" colour ramps.

E.g. In python, see [seaborn](#).  
For printing/white backgrounds:  
[colorbrewer2.org](#)

- For continuous data, **colour ramps** are effective
- For discrete data, always try to limit colours (as a guide think 8 max)

A good visualization  
**reduces the time to**  
insight.

### Using colour must be done with care...

The use of too many colors makes it hard to distinguish, and also requires frequent referencing of the legend.

**Limit yourself to just a few colors** so your audience can actually remember what's what.  
*Be nice :)*

- For continuous data, **colour ramps** are effective
- For discrete data, always try to limit colours (as a guide think 8 max)

Not a natural designer? Use "known good" colour ramps.

E.g. In python, see [seaborn](#).  
For printing/white backgrounds:  
[colorbrewer2.org](#)

A good visualization  
**reduces the time to**  
insight.

### Using colour must be done with care...

The use of too many colors makes it hard to distinguish, and also requires frequent referencing of the legend.

**Limit yourself to just a few colors** so your audience can actually remember what's what.  
*Be nice :)*

- For continuous data, **colour ramps** are effective
- For discrete data, always try to limit colours (as a guide think 8 max)

Not a natural designer? Use "known good" colour ramps.

E.g. In python, see [seaborn](#).  
For printing/white backgrounds:  
[colorbrewer2.org](#)

But the amount of colours and type is quite variable depending on context... i.e. are the size of the shapes varying?  
[c.f. "In Color Perception, Size Matters" Maureen Stone 2012]

A good visualization  
**reduces the time to**  
insight.

### Some more on colour

- Colour has the potential to communicate meaning:  
(more recent work on the communication of emotion [1])
- Bright colours should be reserved for small highlight areas and almost never used as backgrounds



- Colour is relative, not absolute!

[1] Lyn Bartram, Abhishek Patra, and Maureen Stone. 2017. Affective Color in Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1364-1374. DOI: <https://doi.org/10.1145/3025453.3026041>

A good visualization  
**reduces the time to**  
insight.

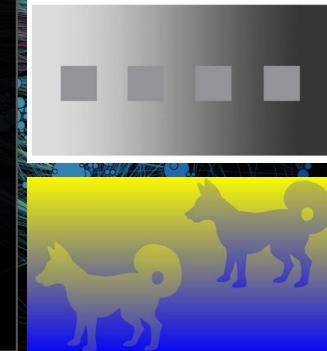
### Some more on colour

- Colour has the potential to communicate meaning:  
(more recent work on the communication of emotion [1])



- Bright colours should be reserved for small highlight areas and almost never used as backgrounds

- Colour is relative, not absolute!  
The squares are the same shade of gray.



[1] Lyn Bartram, Abhishek Patra, and Maureen Stone. 2017. Affective Color in Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1364-1374. DOI: <https://doi.org/10.1145/3025453.3026041>

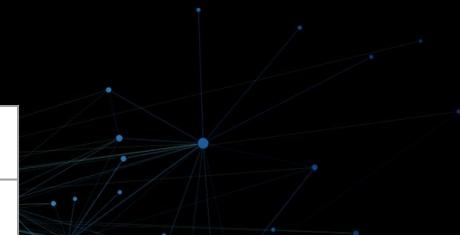
A good visualization  
**reduces the time to**  
insight.

### A final bit on colour...

When using colour in graphs there are three main types of colour schemes to consider (useful w.r.t colourbrewer2.org, seaborn)...

Other standard uses for colour:  
→ Highlight  
→ Alert

Type	Use when
Sequential	ordering values from low to high
Divergent	ordered values + critical mid-point (e.g average or zero)
Categorical / Qualitative	Data is in distinct groups, any colour could appear next to any colour (e.g. colouring counties based on properties).  No implication of magnitude differences between legend classes.



THE USE OF COLOR IN DATA VISUALIZATION

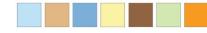
SEQUENTIAL  
color is ordered from low to high



DIVERGING  
two sequential colors with a neutral midpoint



CATEGORICAL  
contrasting colors for individual comparison



HIGHLIGHT  
color used to highlight something



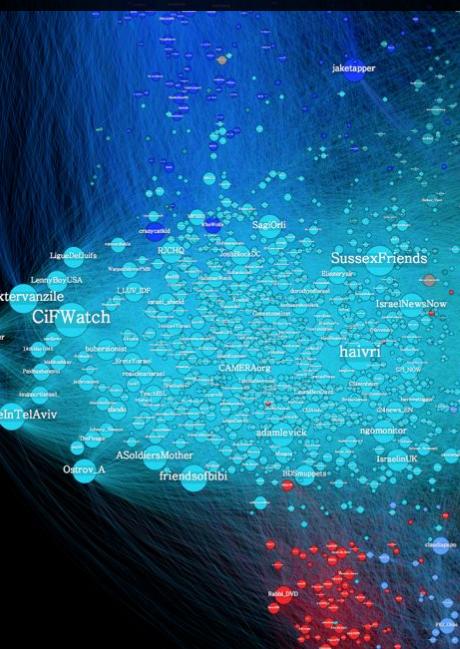
ALERT  
color used to get reader's attention



[2] Cynthia A. Brewer, 1994, "Color Use Guidelines for Mapping and Visualization," Chapter 7 (pp. 123-147) in *Visualization in Modern Cartography*, edited by A.M. MacEachren and D.R.F. Taylor, Elsevier Science, Tarrytown, NY.

Image: The Big Book of Dashboards by Andy Cotgreave, Jeffrey Shaffer, and Victor Grossman (pg 15)

A good visualization  
**reduces the time to insight.** Yes, still this.



Your brain is on autopilot...

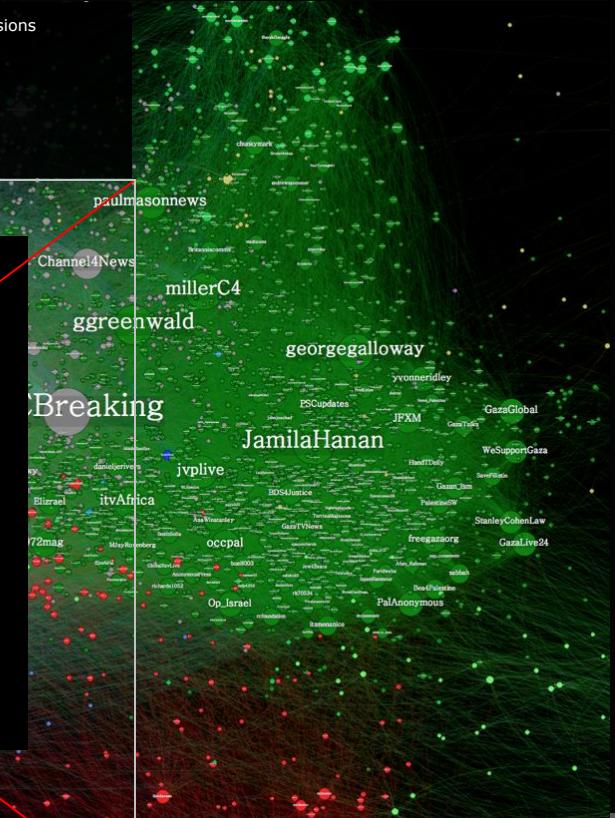
At least a bit...

BBC Horizon: How You Really Make Decisions

# Documentary

## BBC Horizon: How You Really Make Decisions

(16:22 → 22:23)



A good visualization  
**reduces the time to  
insight.**

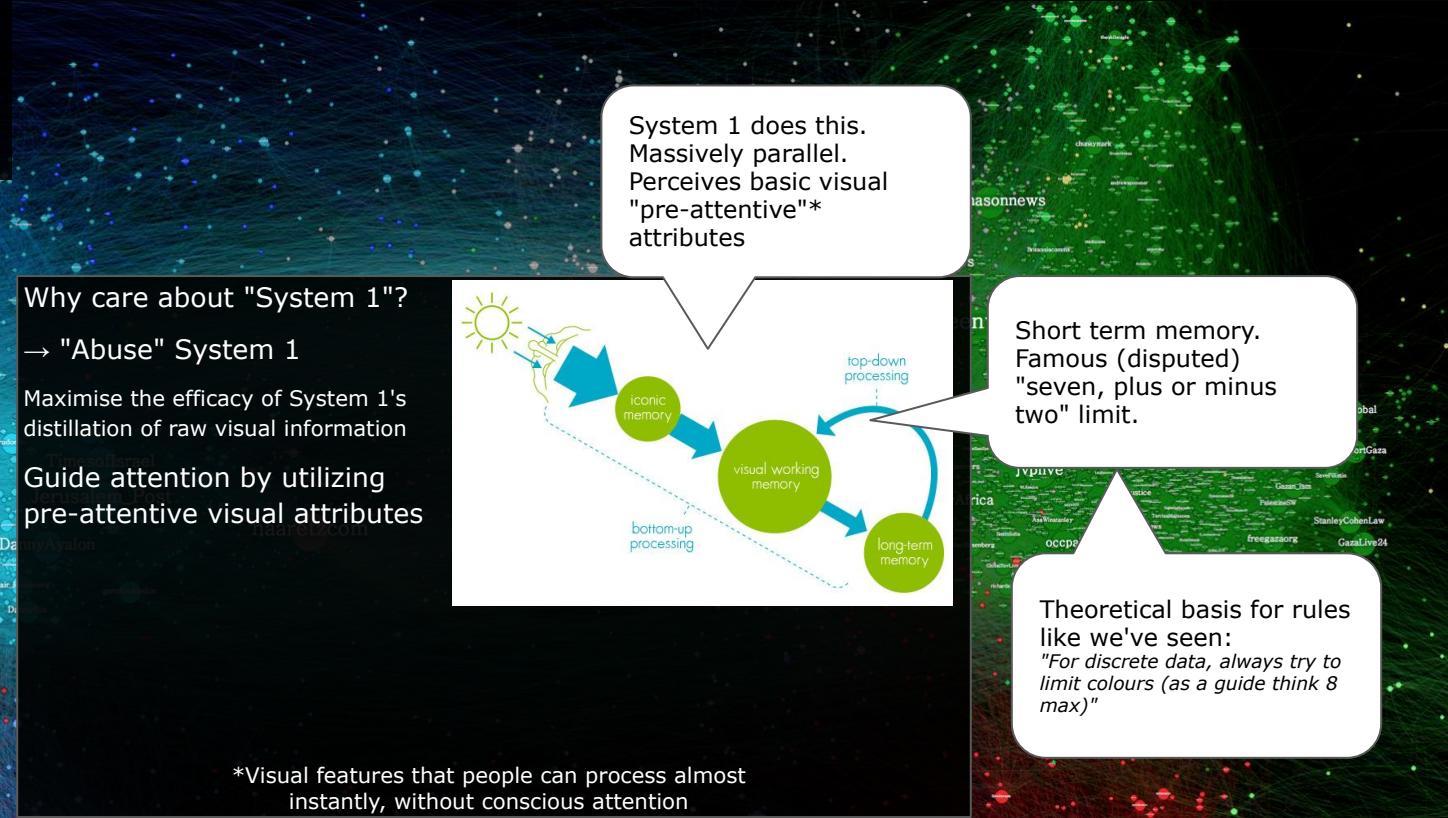
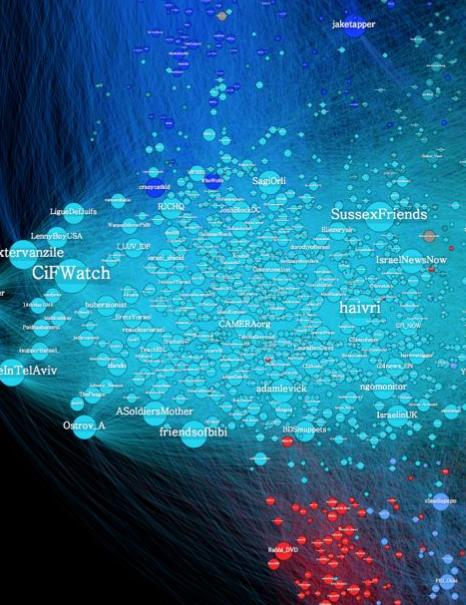


Figure: From Alberto Cairo's The Functional Art

A good visualization  
reduces the time to  
insight.

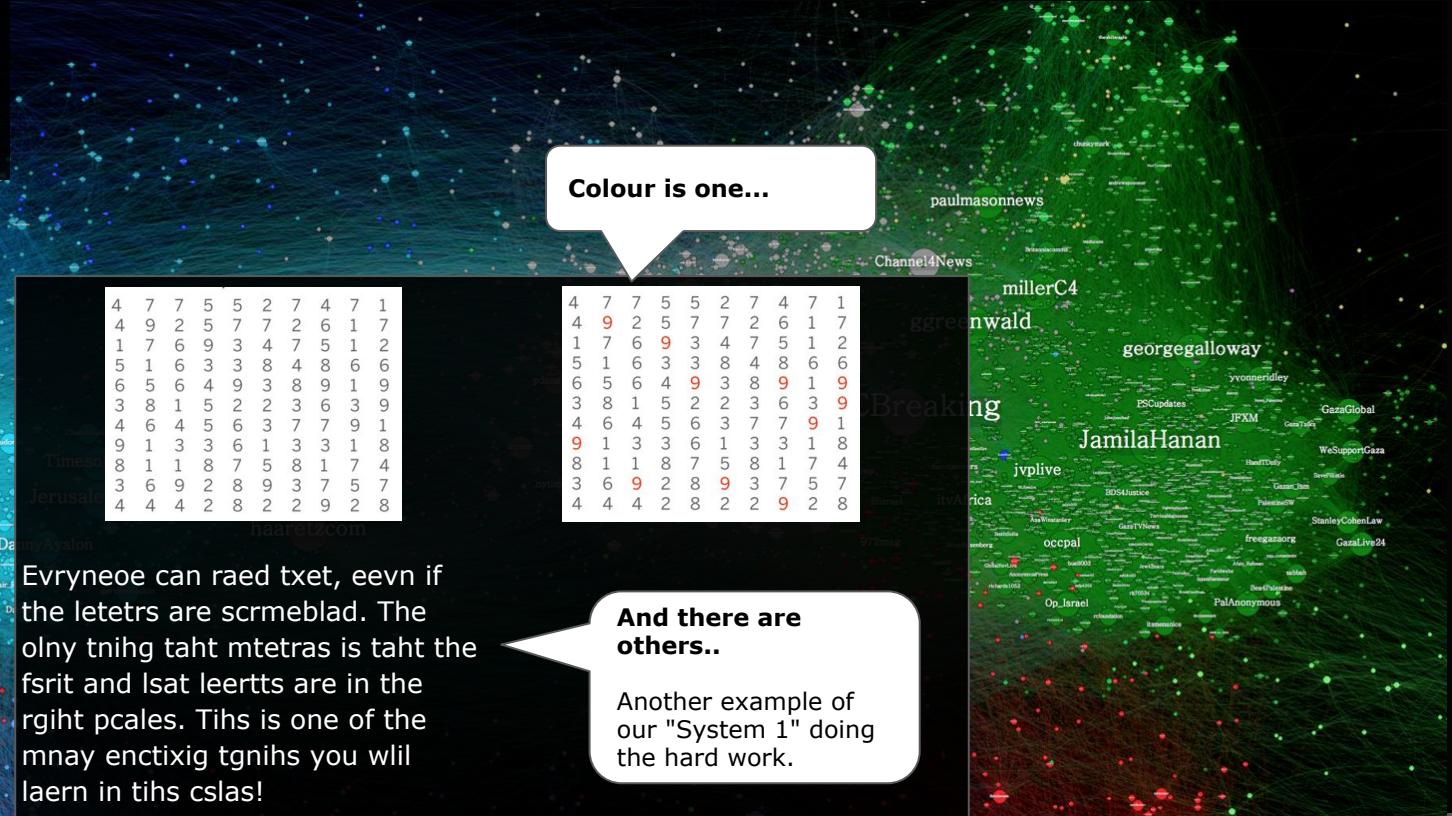
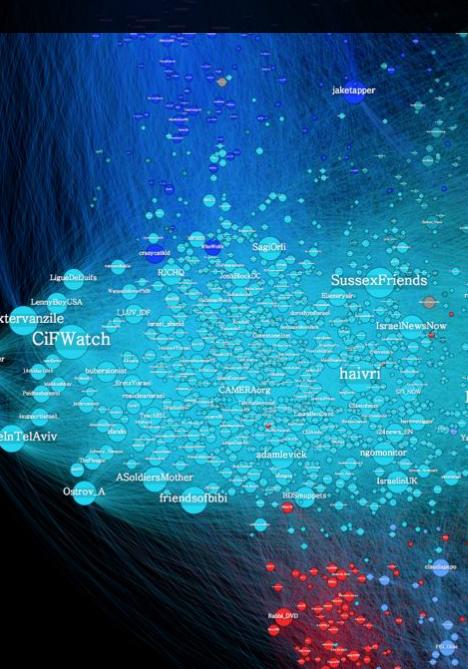
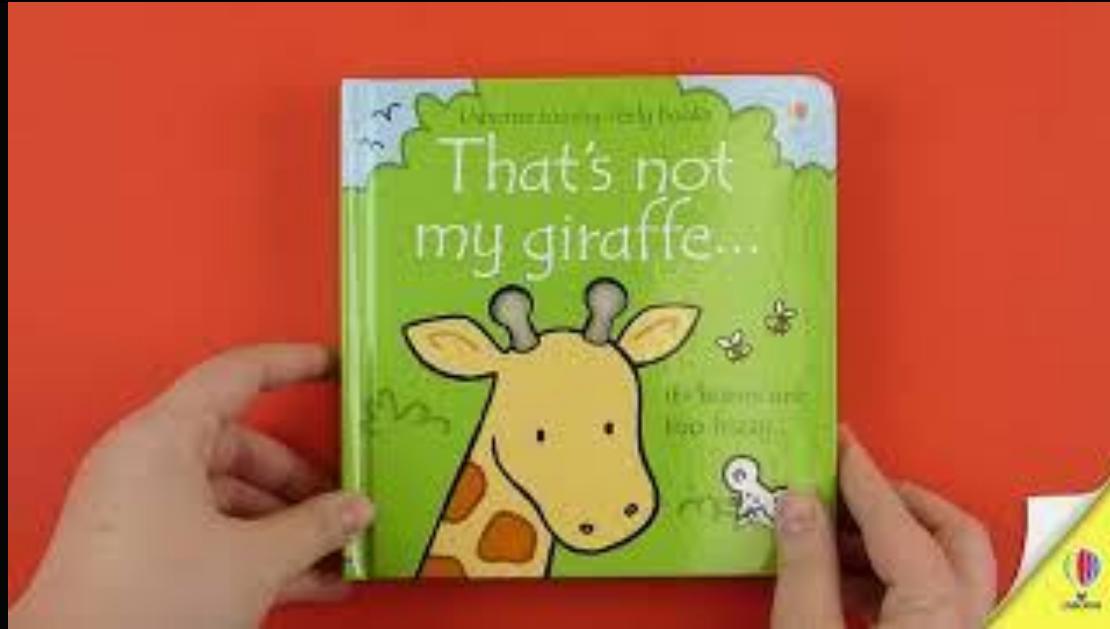
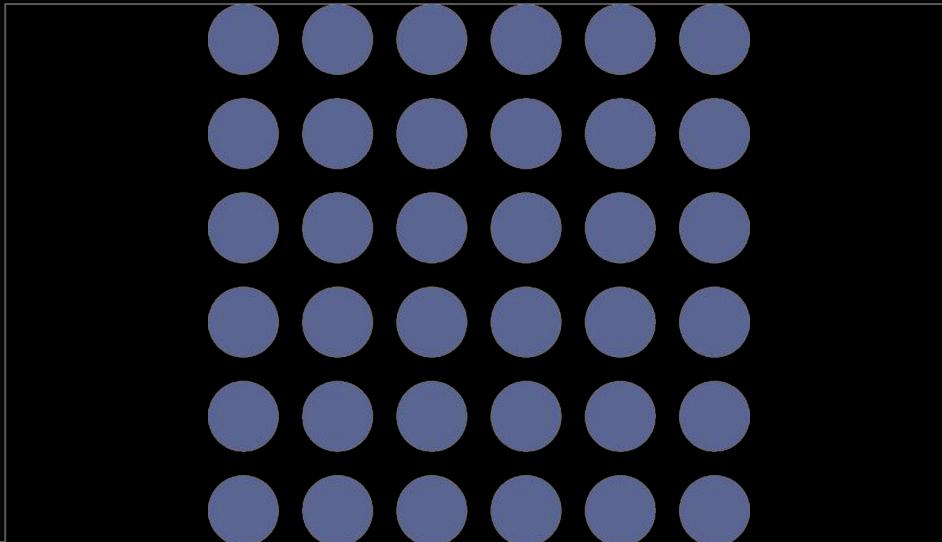


Figure: From Alberto Cairo's The Functional Art

# What other pre-attentive attributes can you think of?



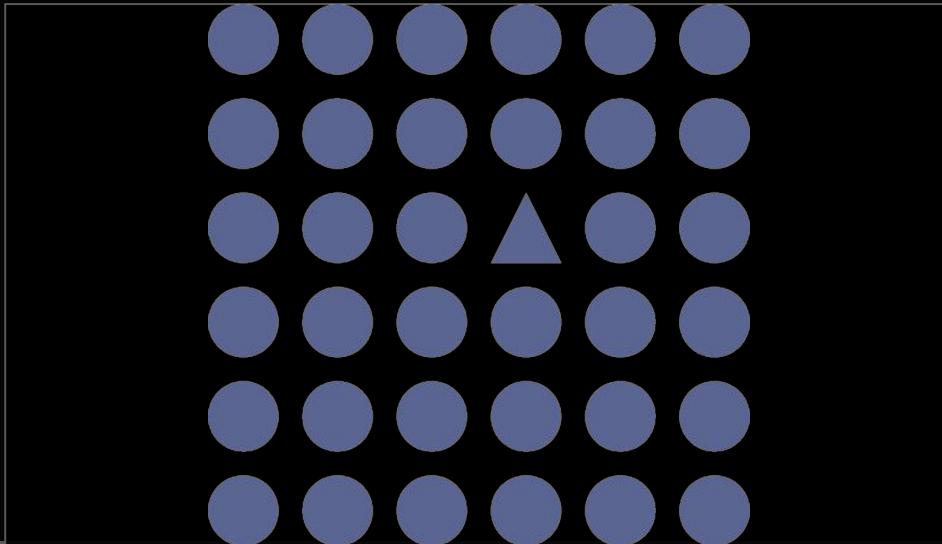
# Shape



*Data at Scale*

Dr Evgeniya Lukinova

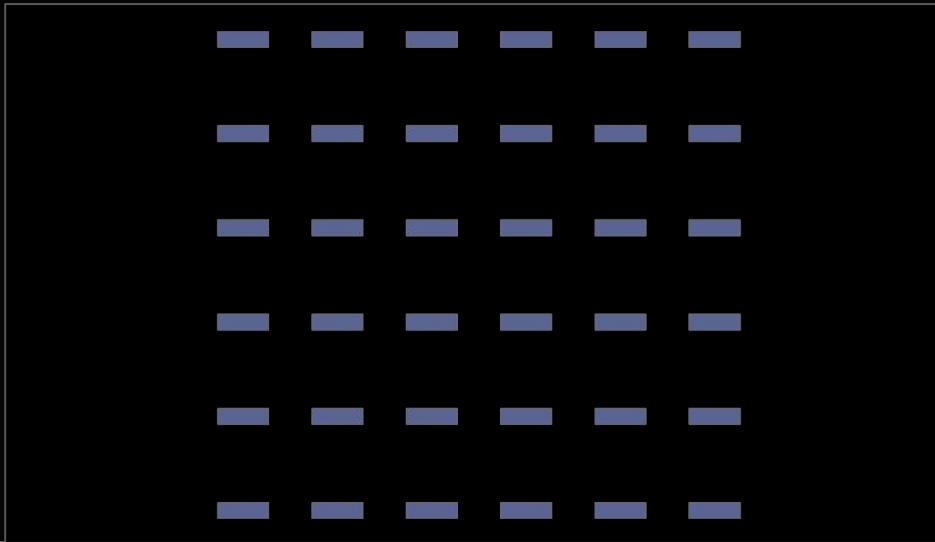
# Shape



*Data at Scale*

Dr Evgeniya Lukinova

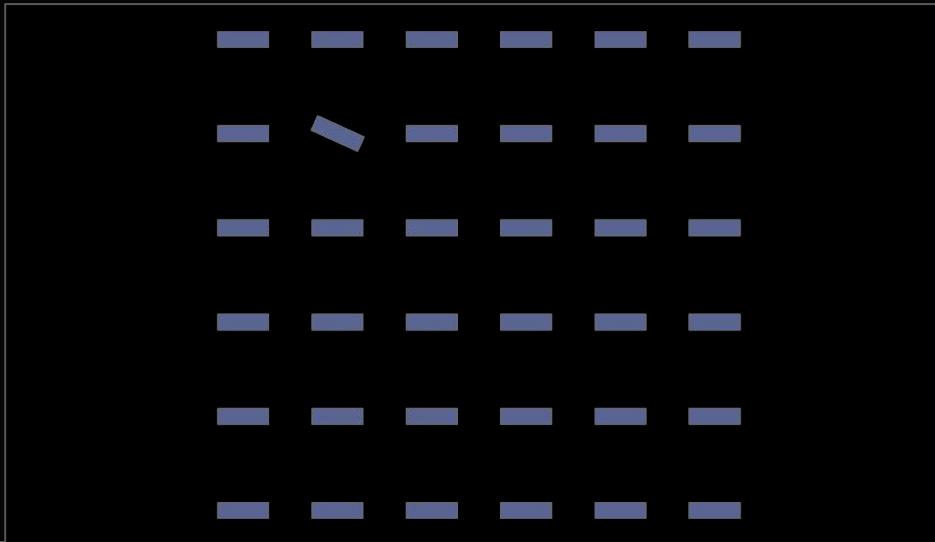
# Orientation



*Data at Scale*

Dr Evgeniya Lukinova

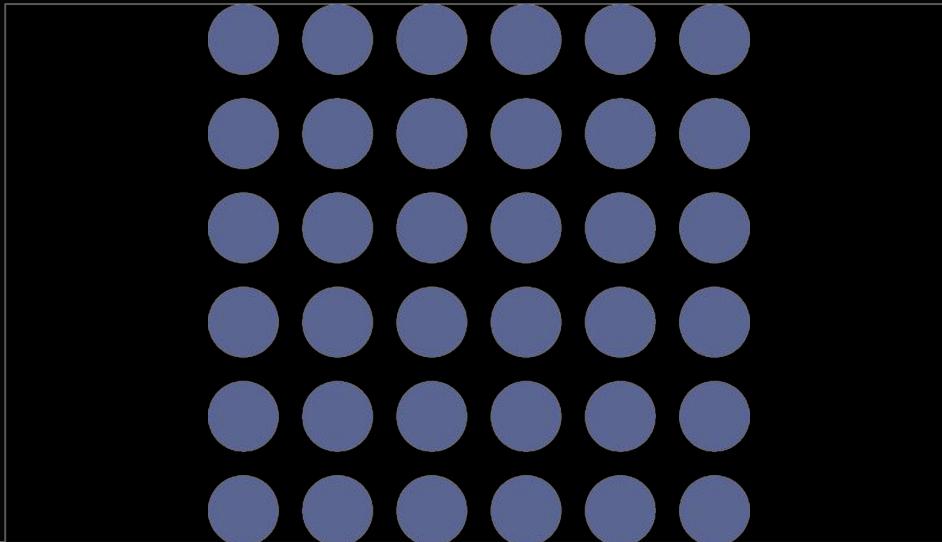
# Orientation



*Data at Scale*

Dr Evgeniya Lukinova

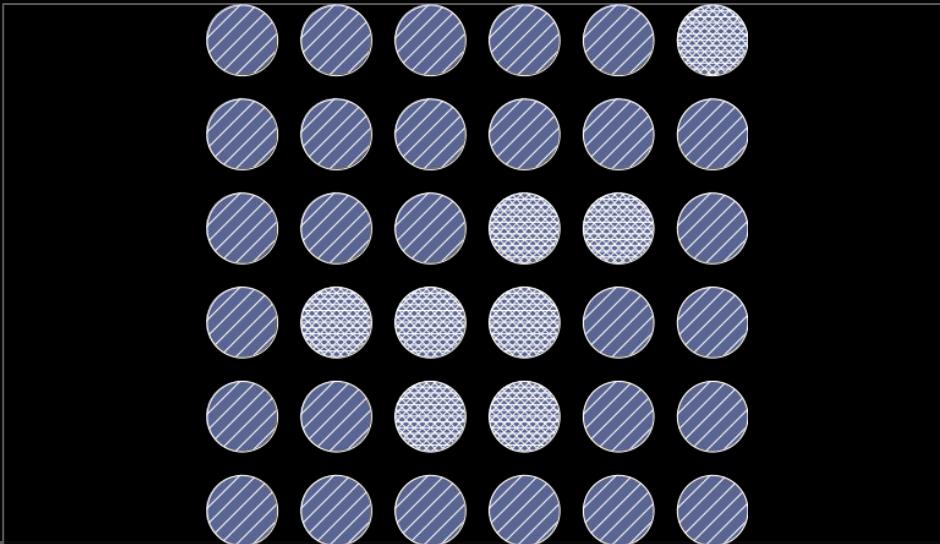
# Texture



*Data at Scale*

Dr Evgeniya Lukinova

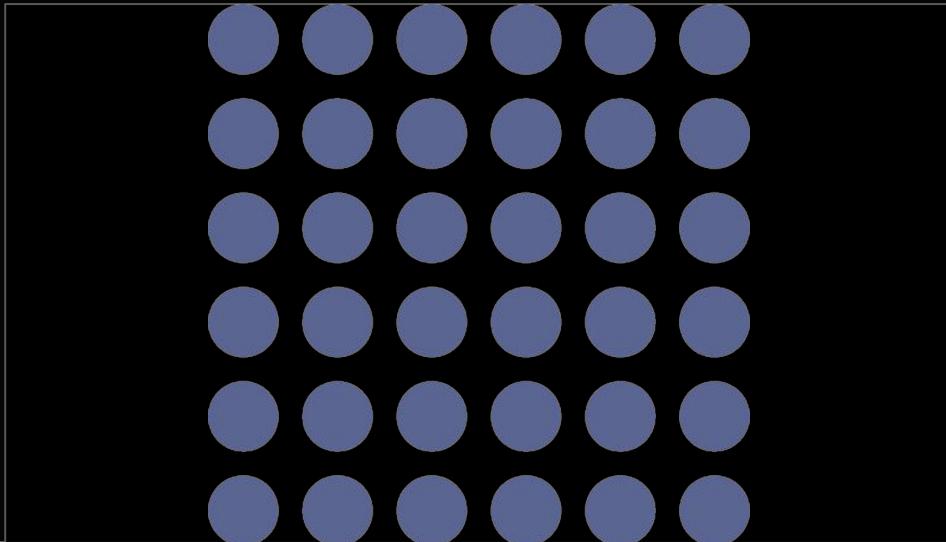
# Texture



*Data at Scale*

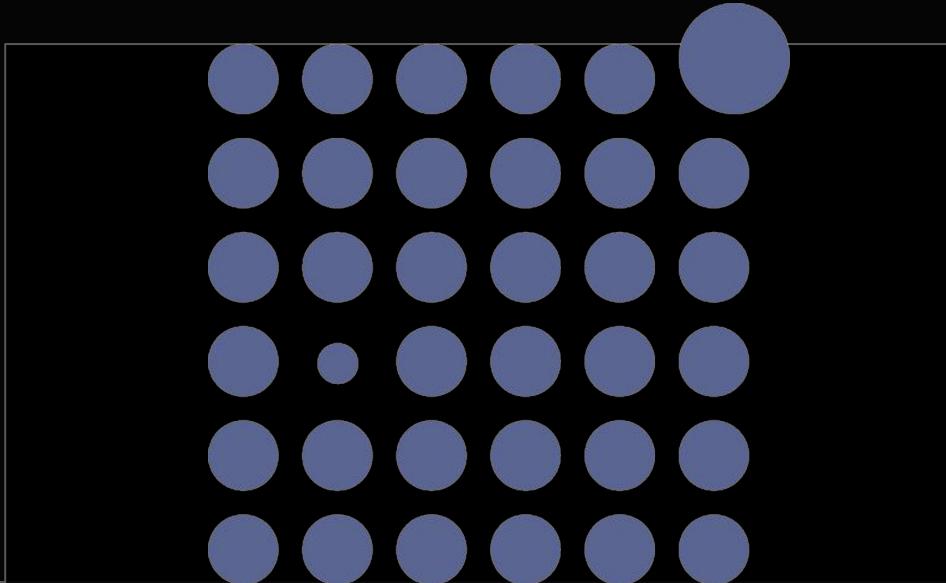
Dr Evgeniya Lukinova

# Size



*Data at Scale*

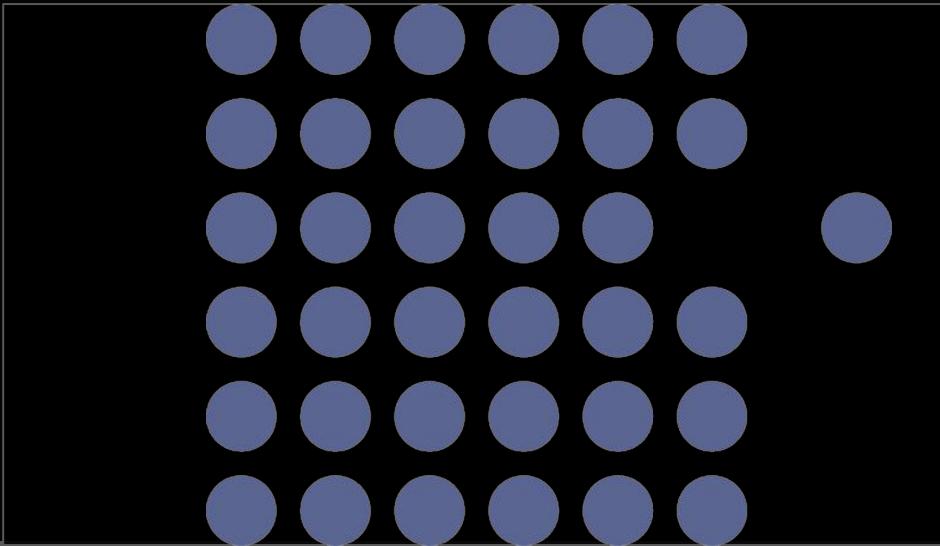
# Size



*Data at Scale*

Dr Evgeniya Lukinova

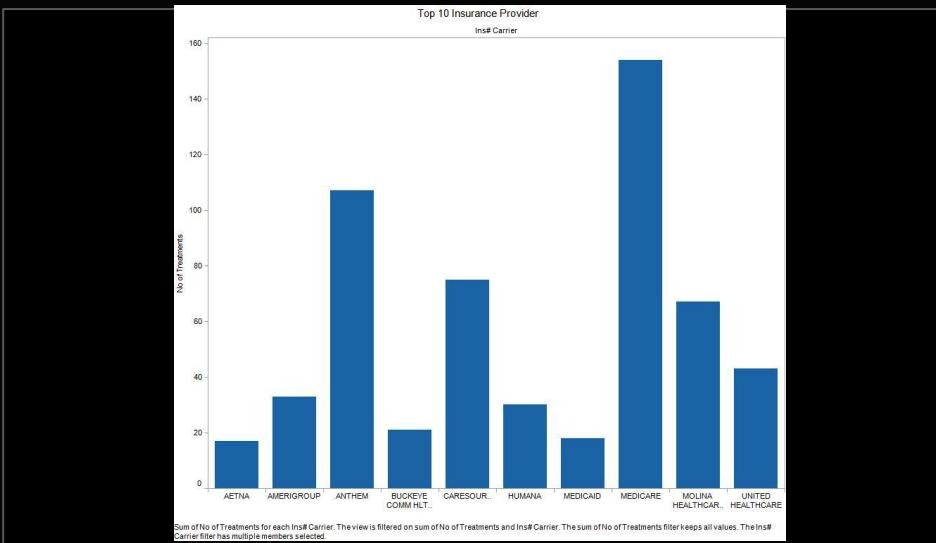
# Position



*Data at Scale*

Dr Evgeniya Lukinova

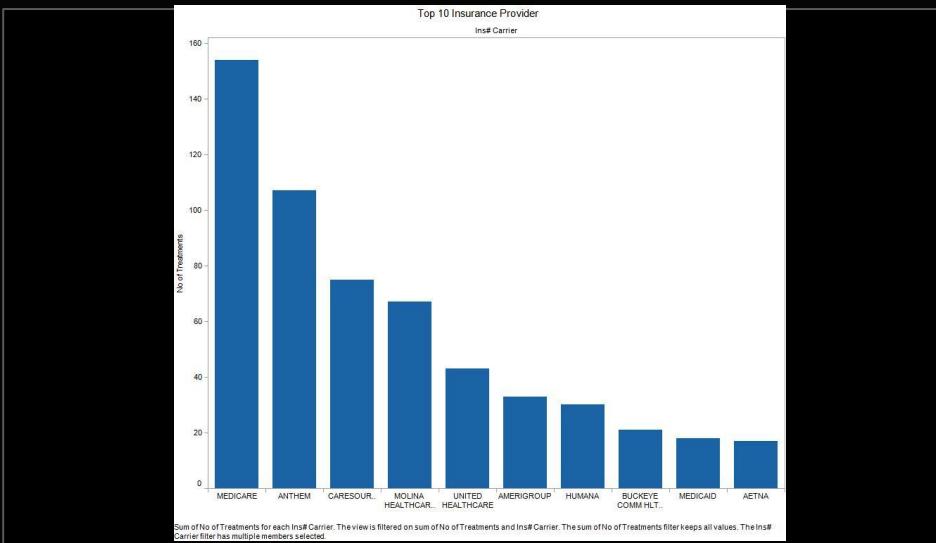
# Order



*Data at Scale*

Dr Evgeniya Lukinova

# Order



*Data at Scale*

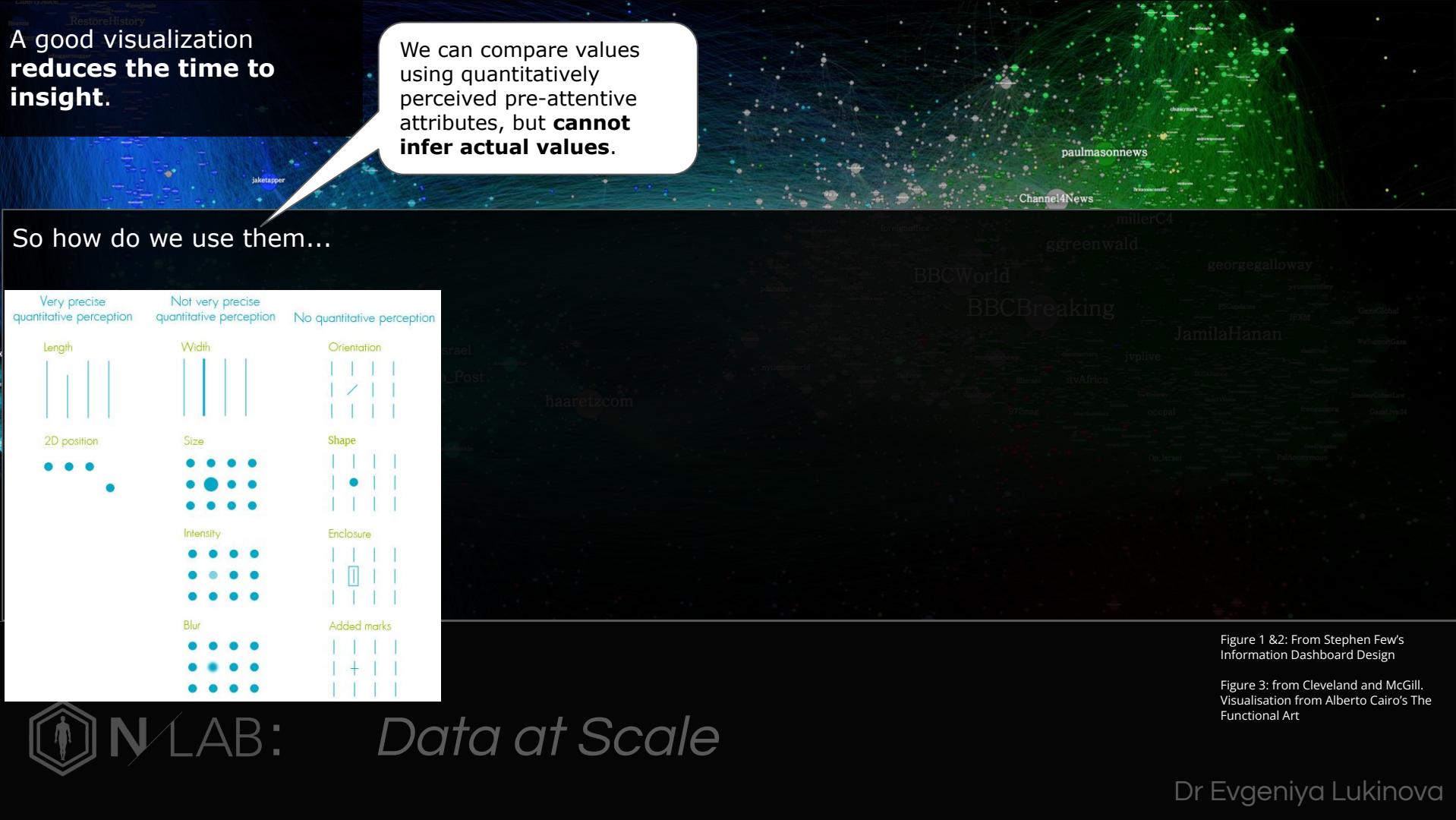
Ok, so hopefully you're convinced pre-attentive attributes are a thing...

Figure: From Alberto Cairo's The Functional Art



## *Data at Scale*

Dr Evgeniya Lukinova



A good visualization reduces the time to insight.

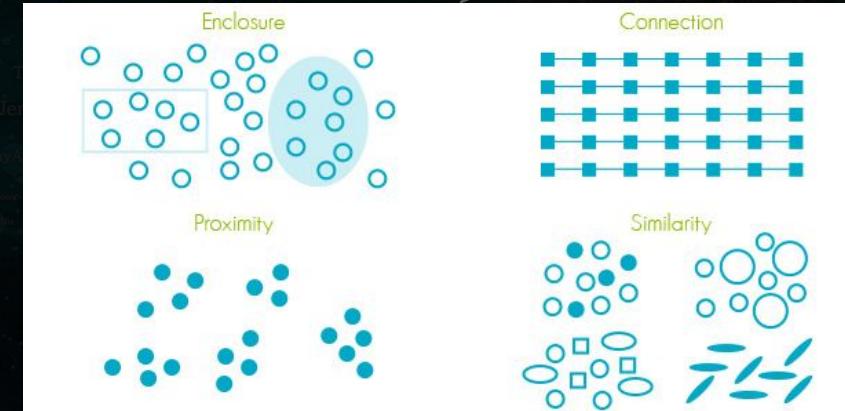


**Order** is another important pre-attentive attribute for highlighting relationships.

A good choice of visualisations can effortlessly and immediately highlight desired insights.

So how do we use them...

### Highlighting relationships



	Very precise quantitative perception	Not very precise quantitative perception	No quantitative perception
Length		Width	
2D position	•••	Size	•••
Intensity	•••	Orientation	/
Blur	•••	Shape	•
Added marks	+	Enclosure	

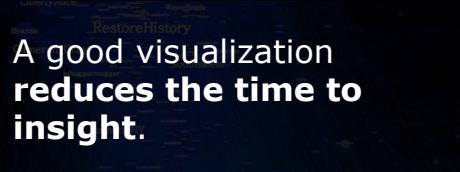
Figure 1 & 2: From Stephen Few's Information Dashboard Design

Figure 3: from Cleveland and McGill. Visualisation from Alberto Cairo's The Functional Art



*Data at Scale*

Dr Evgeniya Lukinova



A good visualization  
reduces the time to  
insight.



More analytical, higher in the chart (more standard).



**Compromise between accuracy and visual interest required for the particular story.**

So how do we use them...

Highlighting relationships

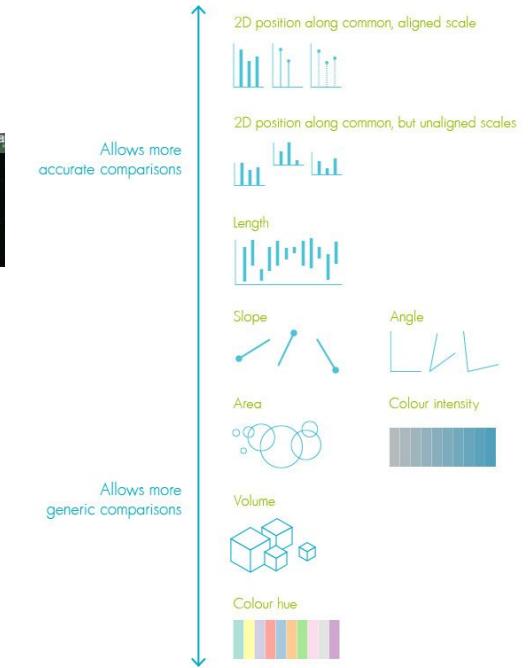
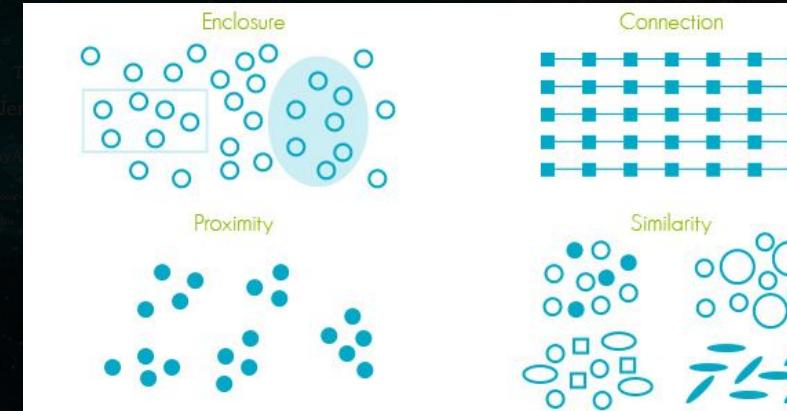
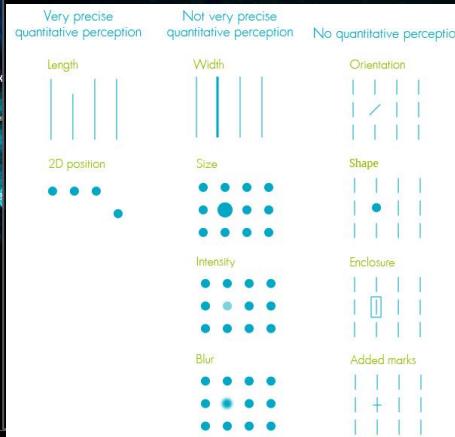


Figure 1 & 2: From Stephen Few's Information Dashboard Design

Figure 3: from Alberto Cairo's The Functional Art



*Data at Scale*

Dr Evgeniya Lukinova

**A good visualization  
reduces the time to  
insight.**

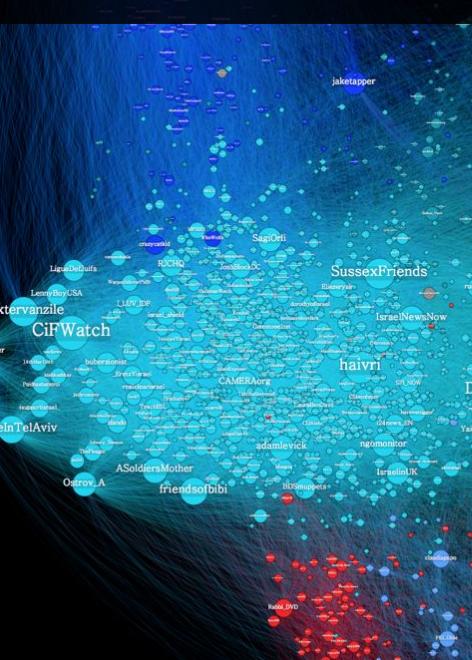


Figure: From Alberto Cairo's The Functional Art



# *Data at Scale*

Dr Evgeniya Lukinova

A good visualization  
**reduces the time to  
insight.**



## So what else can System 1 do?



Figure: From Alberto Cairo's The Functional Art

A good visualization  
**reduces the time to**  
**insight.**

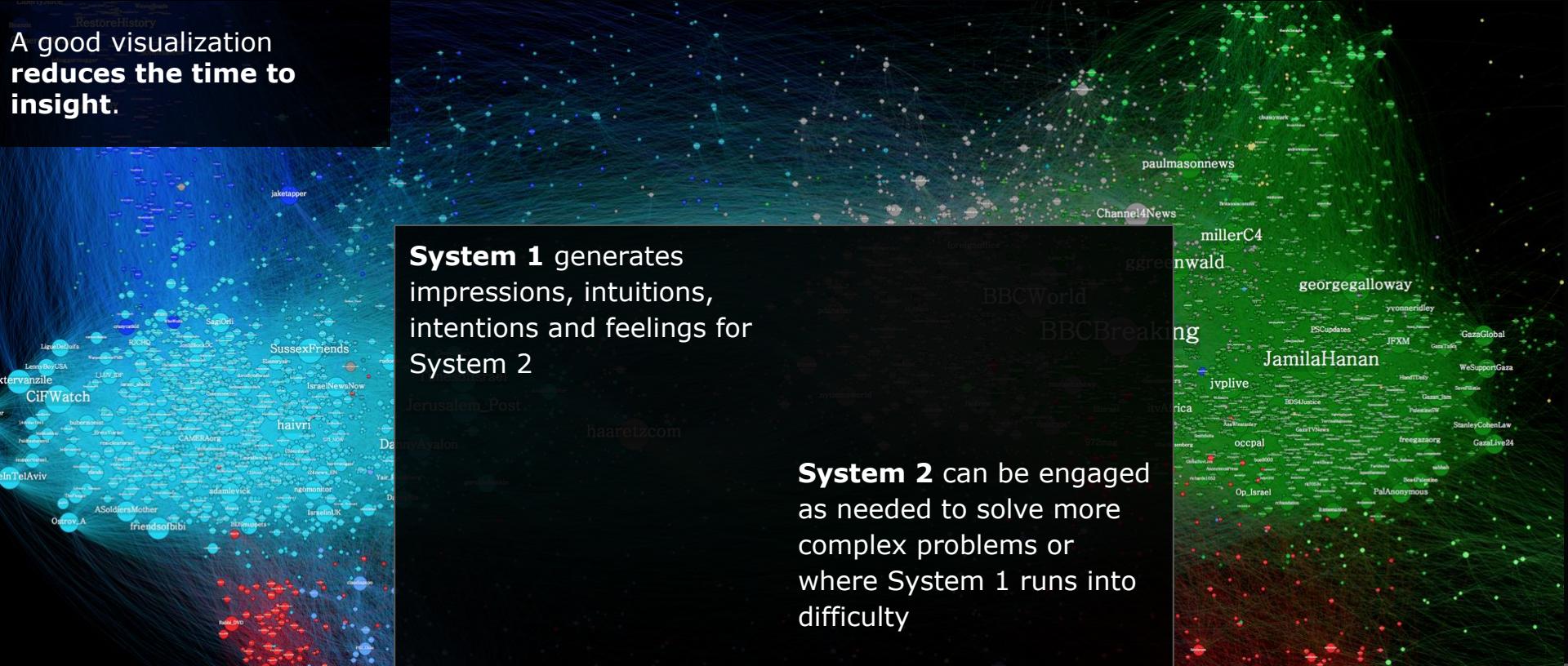


Figure: From Alberto Cairo's The Functional Art

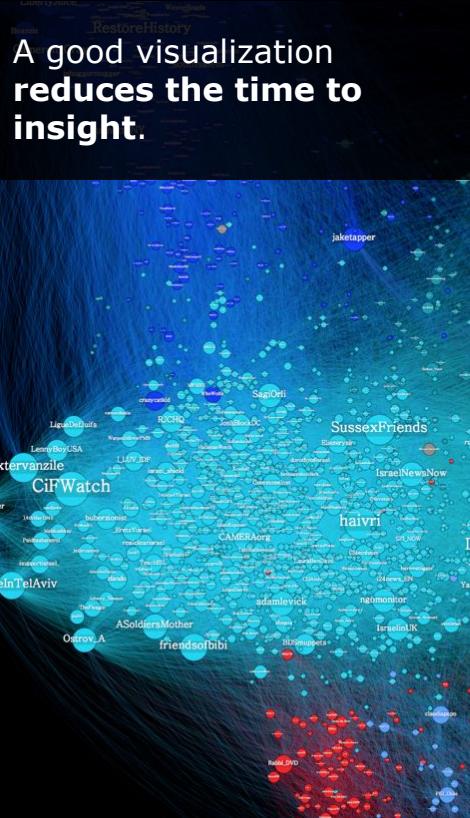


# NLAB:

# *Data at Scale*

Dr Evgeniya Lukinova

A good visualization  
**reduces the time to**  
**insight.**



**Bad impression (e.g. too complex, I'll never understand it) from System 1...**

## **Impressions in System 1 affect the conclusions of System 2.**

→ Presentation impacts the way data is perceived.

→ Mood and emotions  
impact critical thinking

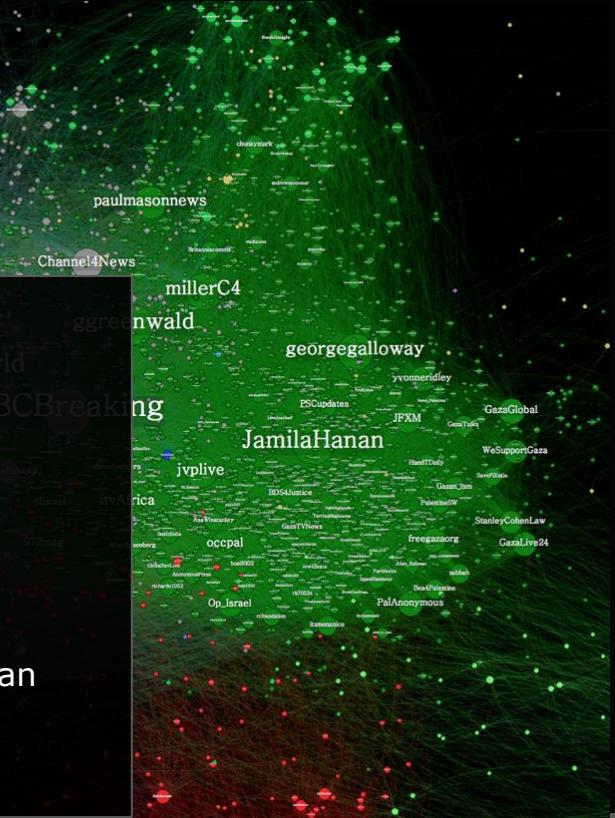
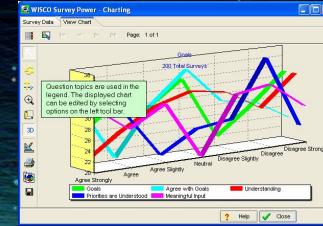


Figure: From Alberto Cairo's The Functional Art



## *Data at Scale*

Dr Evgeniya Lukinova

# Summary

**Why now?** No excuse not to make nice visualizations. Ever. Not even at the start.

## Data exploration vs. Data presentation (focus).

**However, never forget:**  
"Above all else, show the data! Graphics is  
*intelligence made visible*"

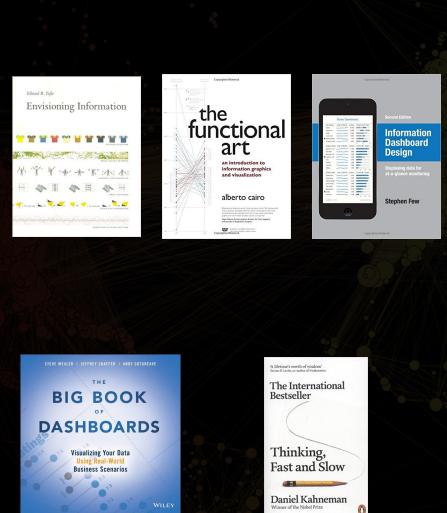
Edward Tufte

Understanding and using  
pre-attentive attributes

"an explanation should be  
as simple as possible, but  
no simpler" Einstein, obviously

We'll take a look at  
**chart types** and  
more general **design  
considerations** later

## Interested in reading more?



[1] Lyn Bartram, Abhishek Patra, and Maureen Stone. 2017. Affective Color in Visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 1364-1374. DOI: <https://doi.org/10.1145/3025453.3026041>)

[2] Cynthia A. Brewer, 1994, "Color Use Guidelines for Mapping and Visualization," Chapter 7 (pp. 123-147) in *Visualization in Modern Cartography*, edited by A.M. MacEachren and D.R.F. Taylor, Elsevier Science, Tarrytown, NY.

[Research on colour by Maureen Stone](#)

[Good blog: The science behind data visualisation](#)