

The Human Factor Data Visualization

Benjamin Bach
bbach@inf.ed.ac.uk
<http://benjbach.me>

Knowing you

Who has an engineering background?

Who has a design / graphics background?

Who has a humanities background?

Who has done a visualization course before?

Who has used data visualizations?

Who has created data visualizations?

COLUMBIA THREADNEEDLE RANKINGS AFTER WORLD TRIATHLON LONDON (RACE 6 OF 10)

1	GWEN JORGENSEN	USA	4000
2	KATIE ZAFERES	USA	3645
3	ANDREA HEWITT	NZL	2910
4	SARAH TRUE	USA	2632
5	BARBARA RIVEROS	CHI	2100

1	JAVIER GOMEZ NOYA	ESP	3507
2	MARIO MOLA	ESP	2979
3	FERNANDO ALARZA	ESP	2961
4	RICHARD MURRAY	RSA	2490
5	ALISTAIR BROWNLEE	GBR	2340



MIXED RELAY WORLD CHAMPIONSHIP MEDALLISTS



386,250km

THE TOTAL DISTANCE EXPECTED TO BE COVERED BY ATHLETES TAKING PART AT THE WORLD TRIATHLON HAMBURG. EQUIVALENT DISTANCE FROM EARTH TO THE MOON



€15 MILLION

ESTIMATED VALUE OF BIKES IN THE WORLD TRIATHLON HAMBURG TRANSITION AREA

10,000

AGE-GROUP ATHLETES HAVE SIGNED UP TO RACE THE OLYMPIC AND SPRINT DISTANCE EVENTS OVER THE WORLD TRIATHLON HAMBURG WEEKEND

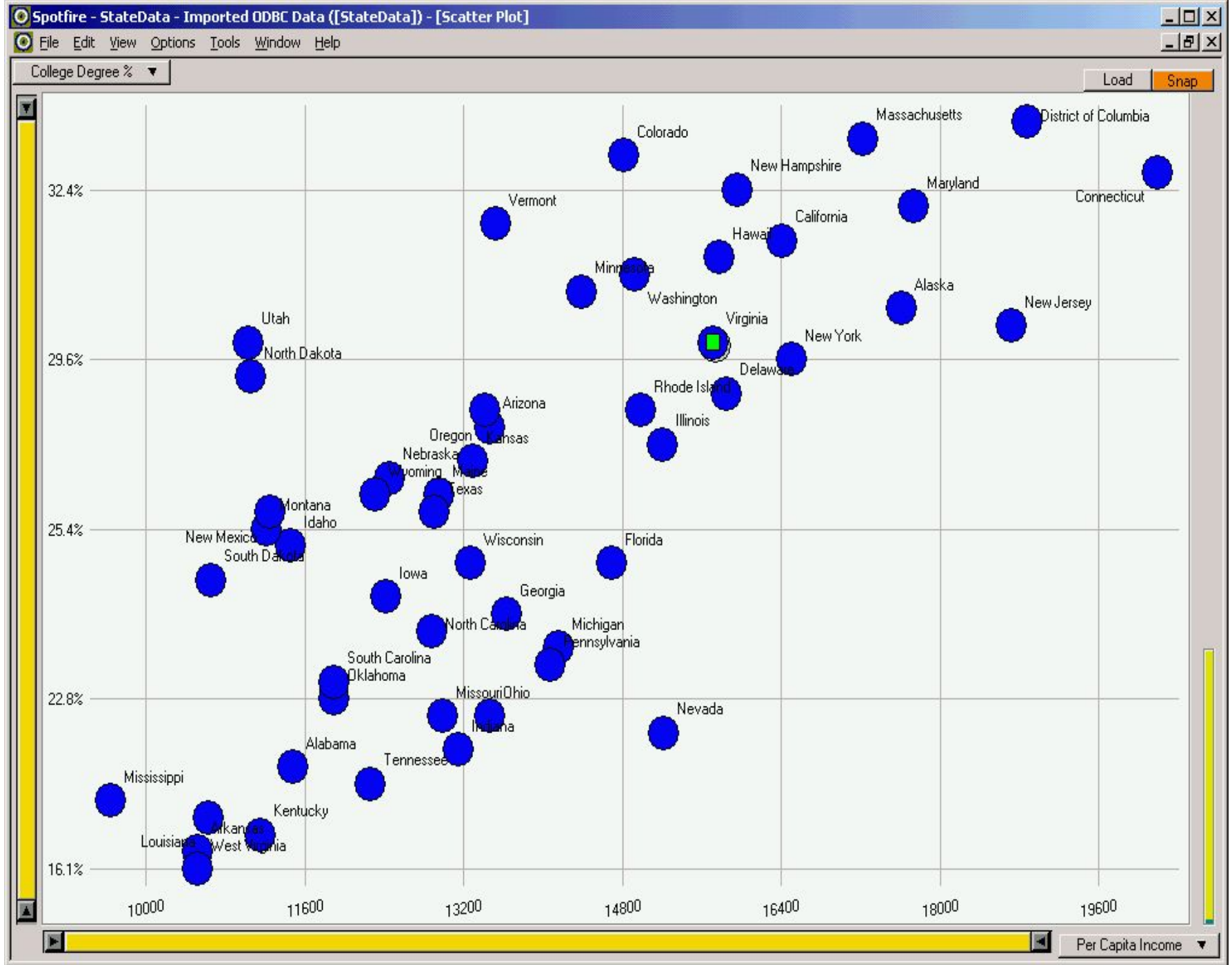


FOOD/DRINK CONSUMED 9,200 BANANAS, 6,700 ORANGES, 8,000 ENERGY BARS, 8,250 LITRES OF ENERGY DRINK, 10,000 LITRES OF WATER, 2,900 PIECES OF CAKE

@CTInvest_tri #CT_Rankings

COLUMBIA THREADNEEDLE RANKINGS 2015 WORLD TRIATHLON HAMBURG PREVIEW

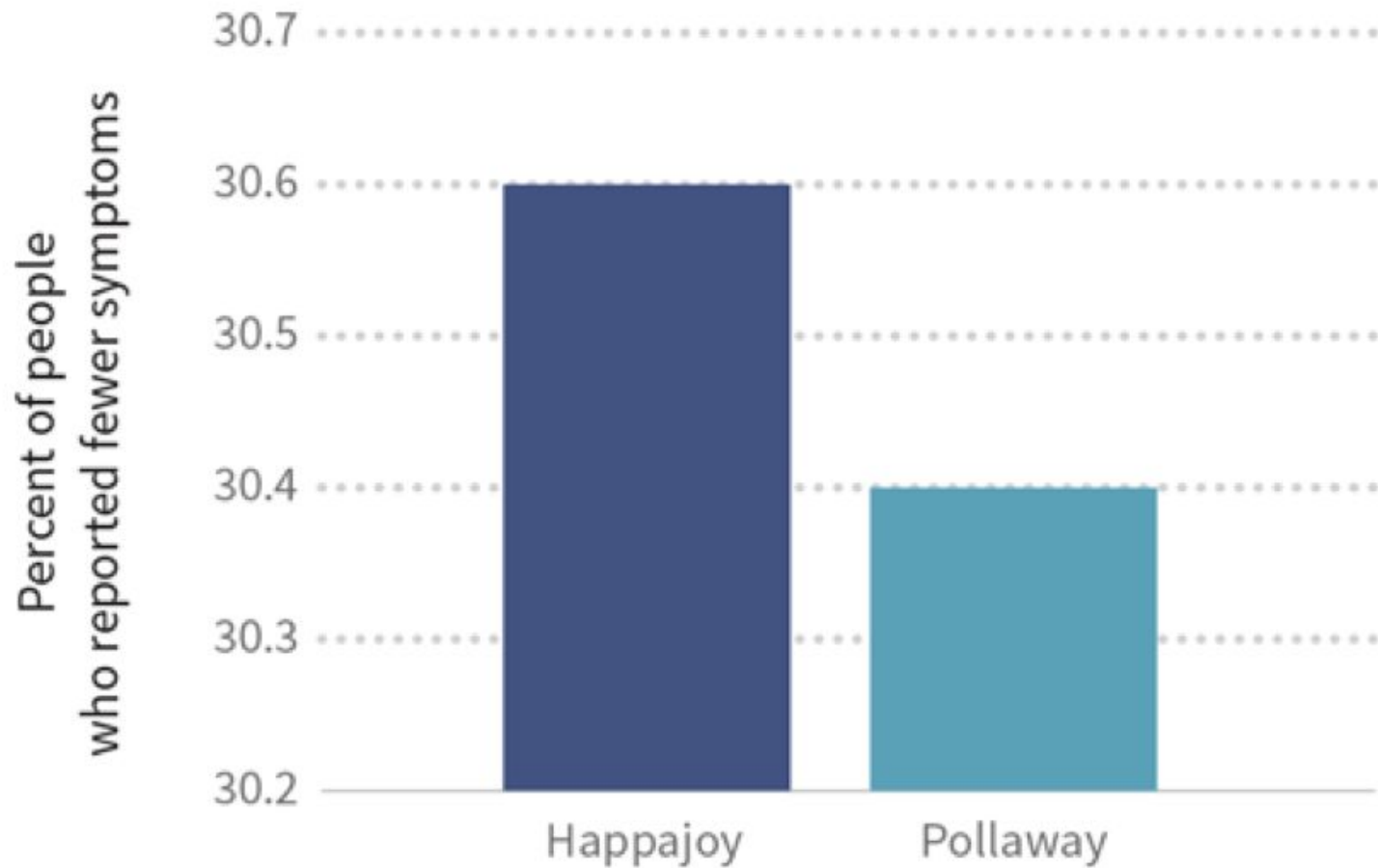




(Possible) Definitions

- *"The use of computer-supported, interactive, **visual representations of abstract data to amplify cognition**"—Stu Card*
- *"... an accessible way to see and understand **trends, outliers, and patterns** in data."—Tableau*
- *".. to help people carry out **tasks more effectively**" — Tamara Munzner*

Effectiveness of Allergy Medicines



AFFORDABLE CARE ACT ENROLLMENT



IE IN SECOND PLACE WITH \$26.5 MIL, WHILE "MUPP DOW FUT 16,325.00

S&P 500 VS. BITCOIN

■ S&P 500 ■ BITCOIN DAILY PRICE

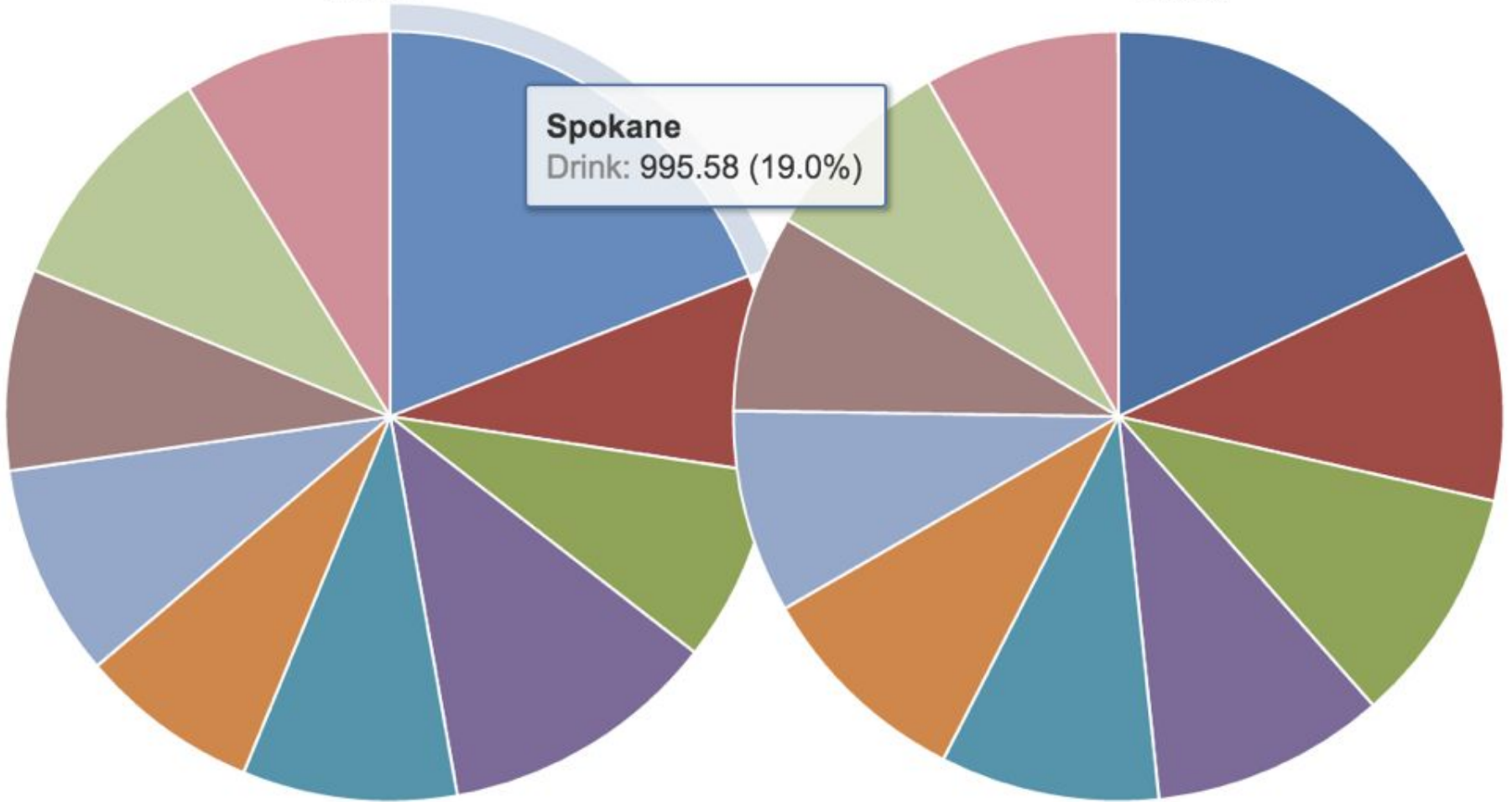


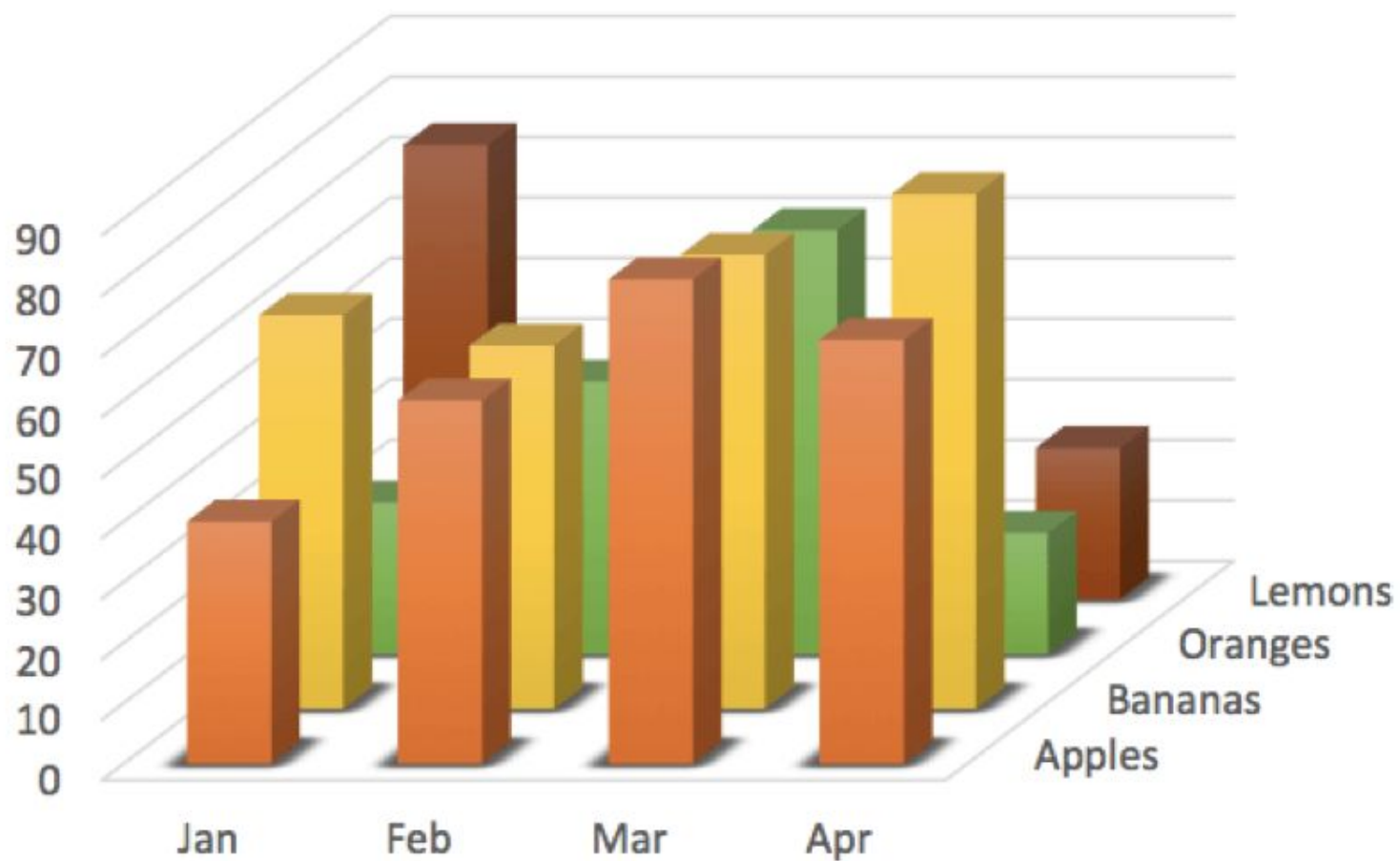
Drink

Food

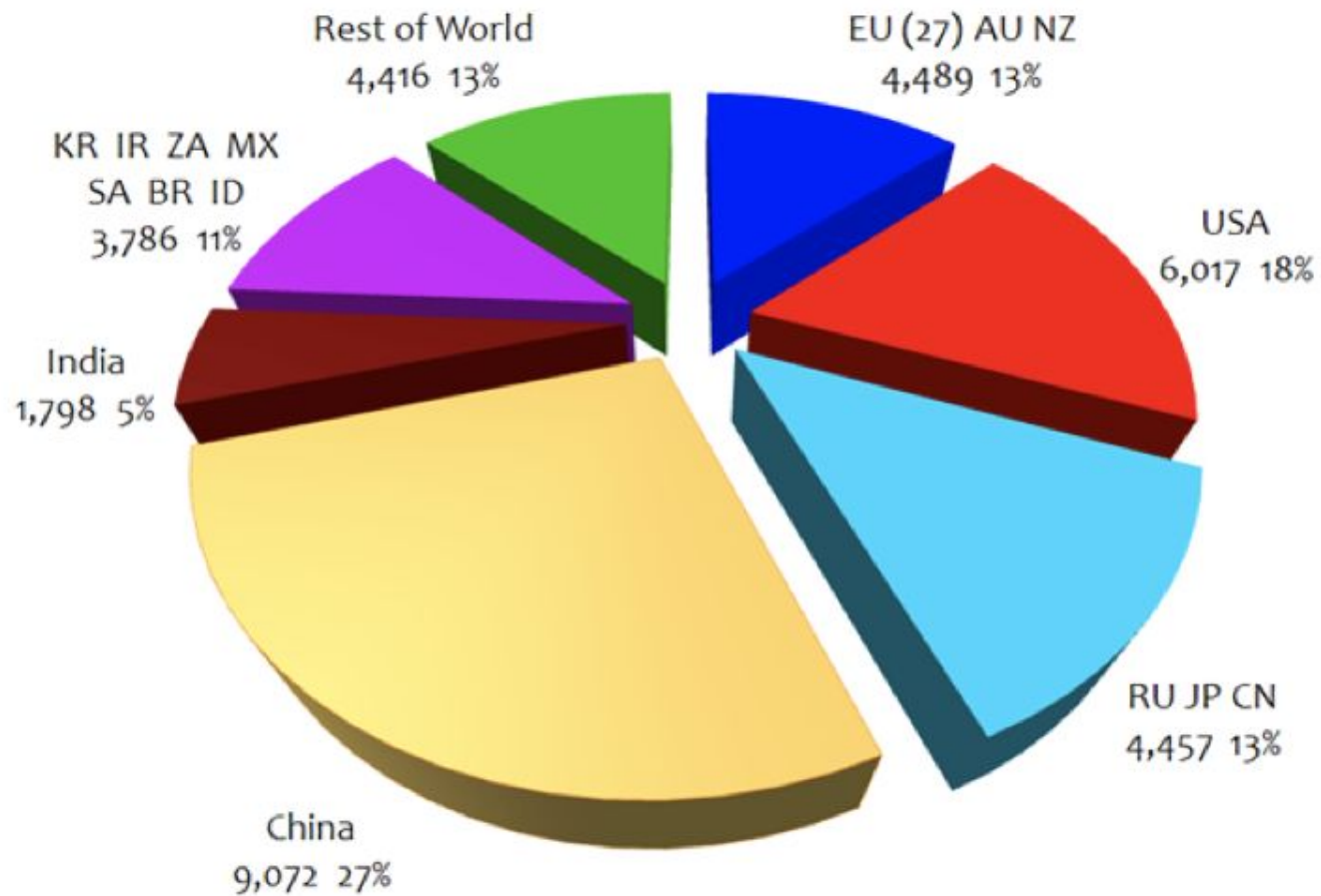
Spokane

Drink: 995.58 (19.0%)



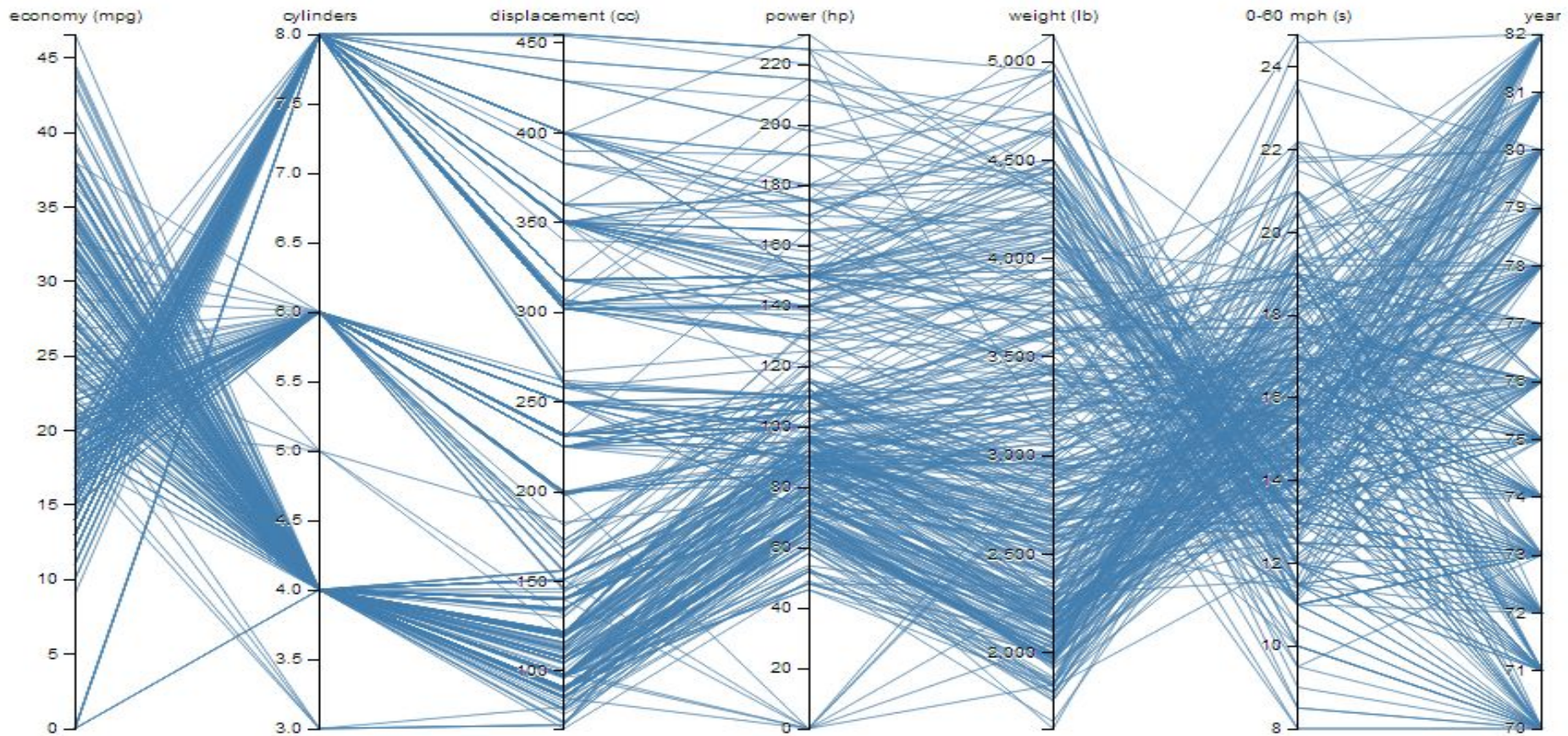


CO2 emissions 2011: '000,000 tonnes



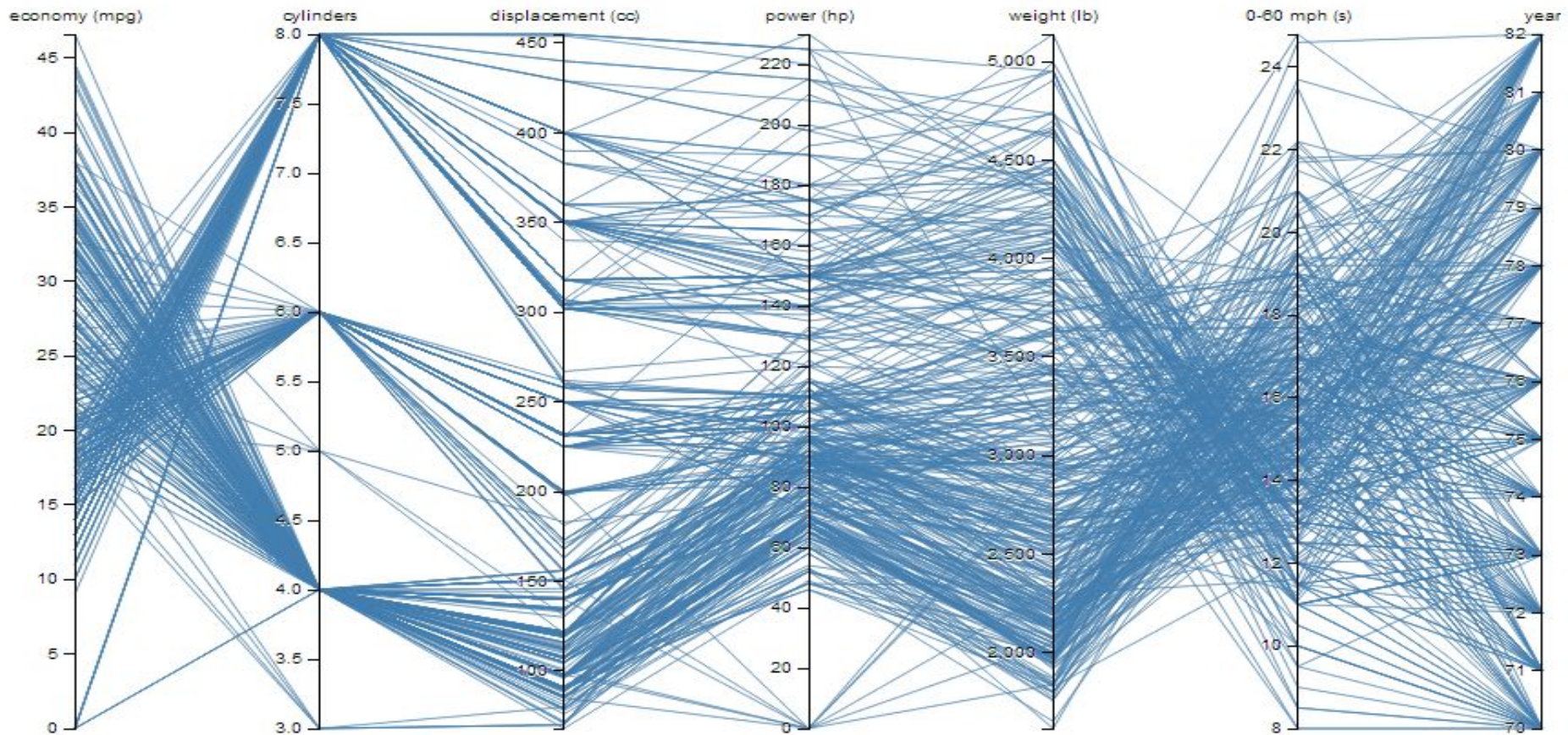
This course: Learning Outcomes 1/3

Interpretation: Correctly interpret non-standard visualizations.



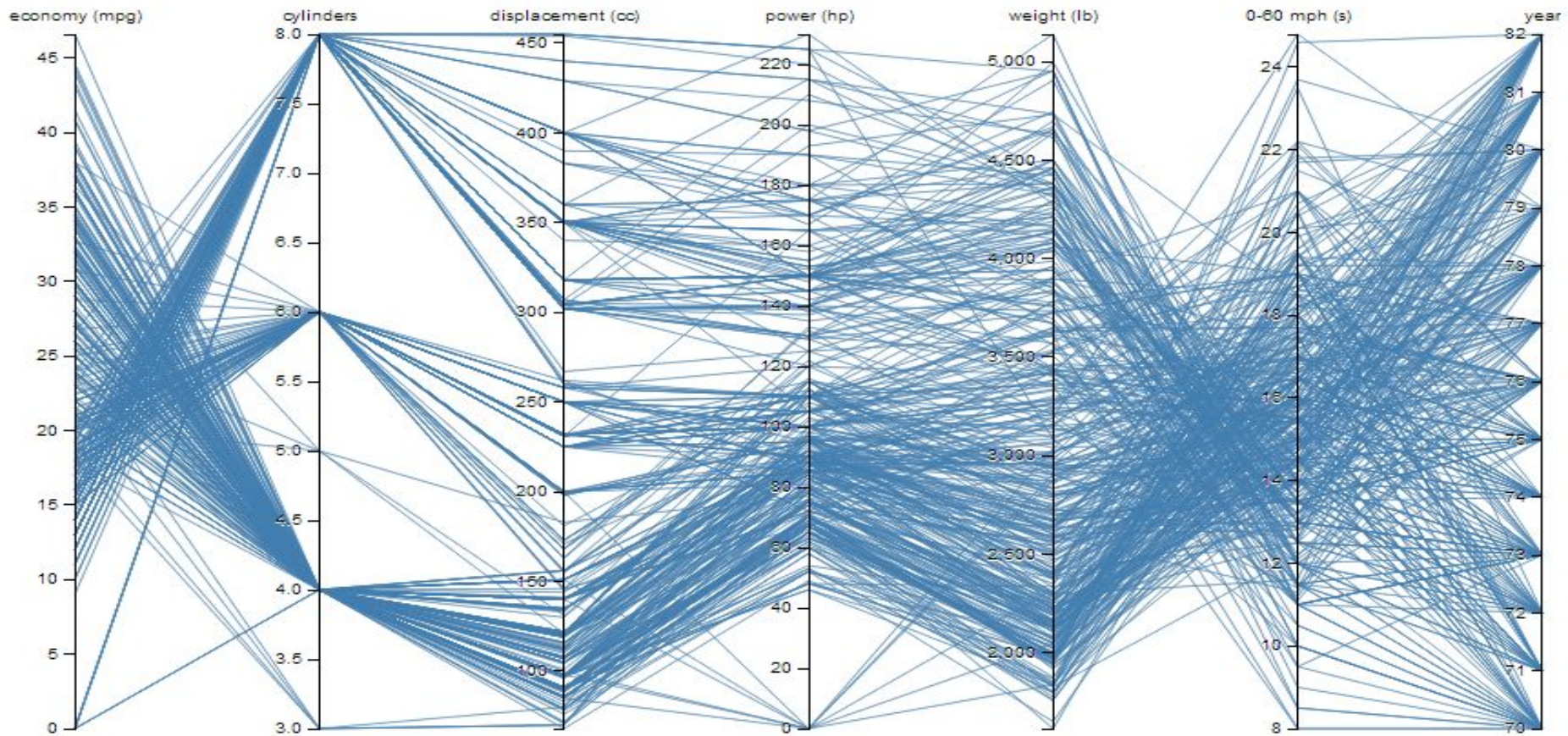
This course: Learning Outcomes 2/3

Critique: Critically judge its visualization design with respect to deception and perception flaws.



This course: Learning Outcomes 3/3

Design: Create an informed (re-)design of a visualization for a specific context and data set.



Outline

1. Foundations
2. Task and Scenarios > **TUTORIAL I**
3. Visualization Techniques I
4. Visualization Techniques II
5. Design > **TUTORIAL II**
6. Sketching session (bring data)
7. Presentation and Storytelling > **TUTORIAL III**
8. Interaction
9. Vis tools
10. Evaluation
11. general Q&A

Assignment III (April 26)

Pick a data set of your choice (at least 20 elements (rows) and 6 attributes (columns)) and create an appropriate visualization by hand or using available tools.

1. **Describe your design rationale** for the visualization (25%):
 - a. why have you chosen this visualization?
 - b. Which alternatives did you consider?

Assignment III

2. **Create an explanatory piece** to communicate context, insights, and take-home messages for this data, using visualizations. This can be

- an infographic (1 A3 page)
- a report (up to 300 words+ min 4 figures),
- or a data comic (2 A4 pages) (75%).

Again, briefly describe your design rationale, your intended audience, and the context in which the story will be read by your audience.

Assignment III

Collect data for 1 week:

- **Personal data:** (arguments, friends, hours spend on something, personal waste..)
- **Public data:** CO2 emissions, public spending, voting polls, ...
- **Observations:** in streets, cafés, ...

Assignment III: 30 items + 5 attributes

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69		5 4-Not Specified	Small Pack	0.44	6/6/05
69		5 4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06

Tutorials & Assignment

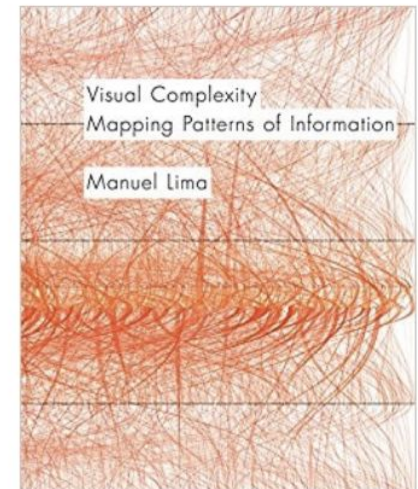
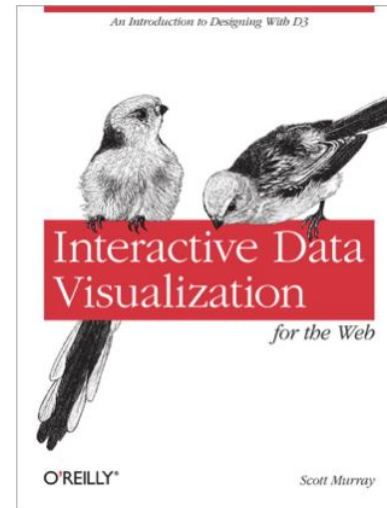
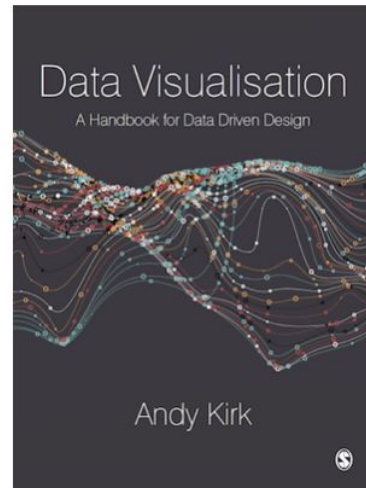
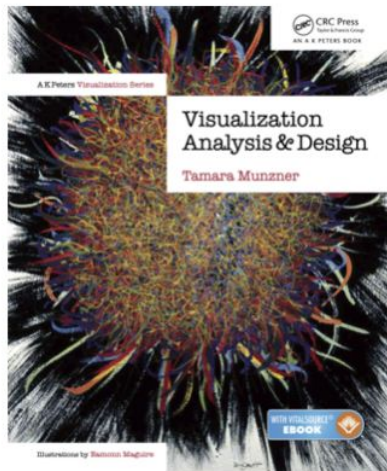
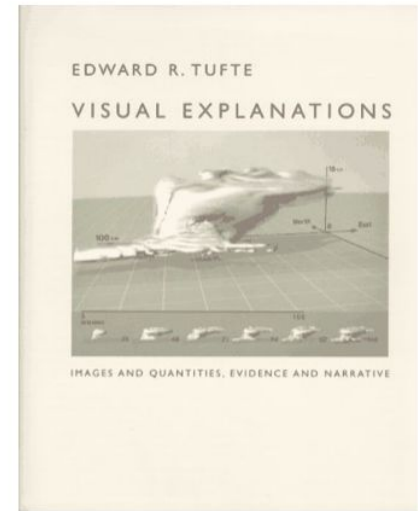
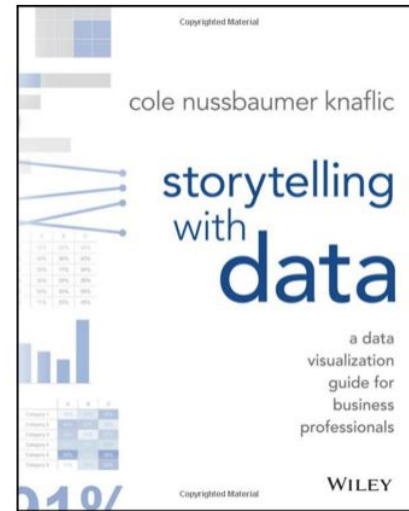
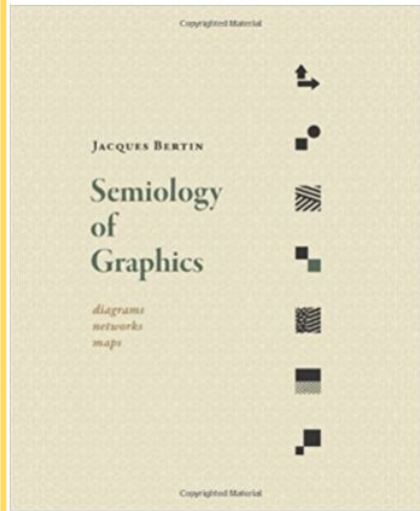
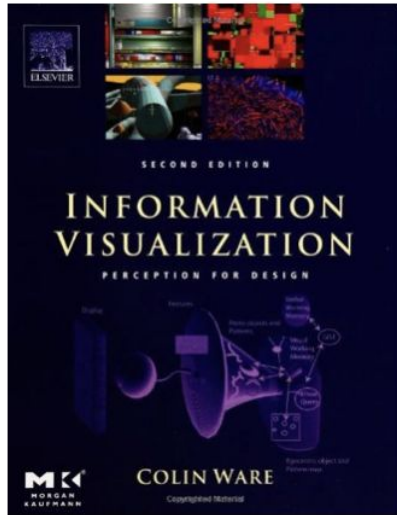
Week 7: **Tasks and Scenarios**

Week 8: **Visualization Techniques**

Week 10: **Storytelling and Presentation**

April 26th: **Assignment**

General Readings



<https://vishubblog.wordpress.com/books/>

1

Foundations of Data Visualization

1. Visual Perception
2. Visual Mapping & Visual Variables

Representations

34

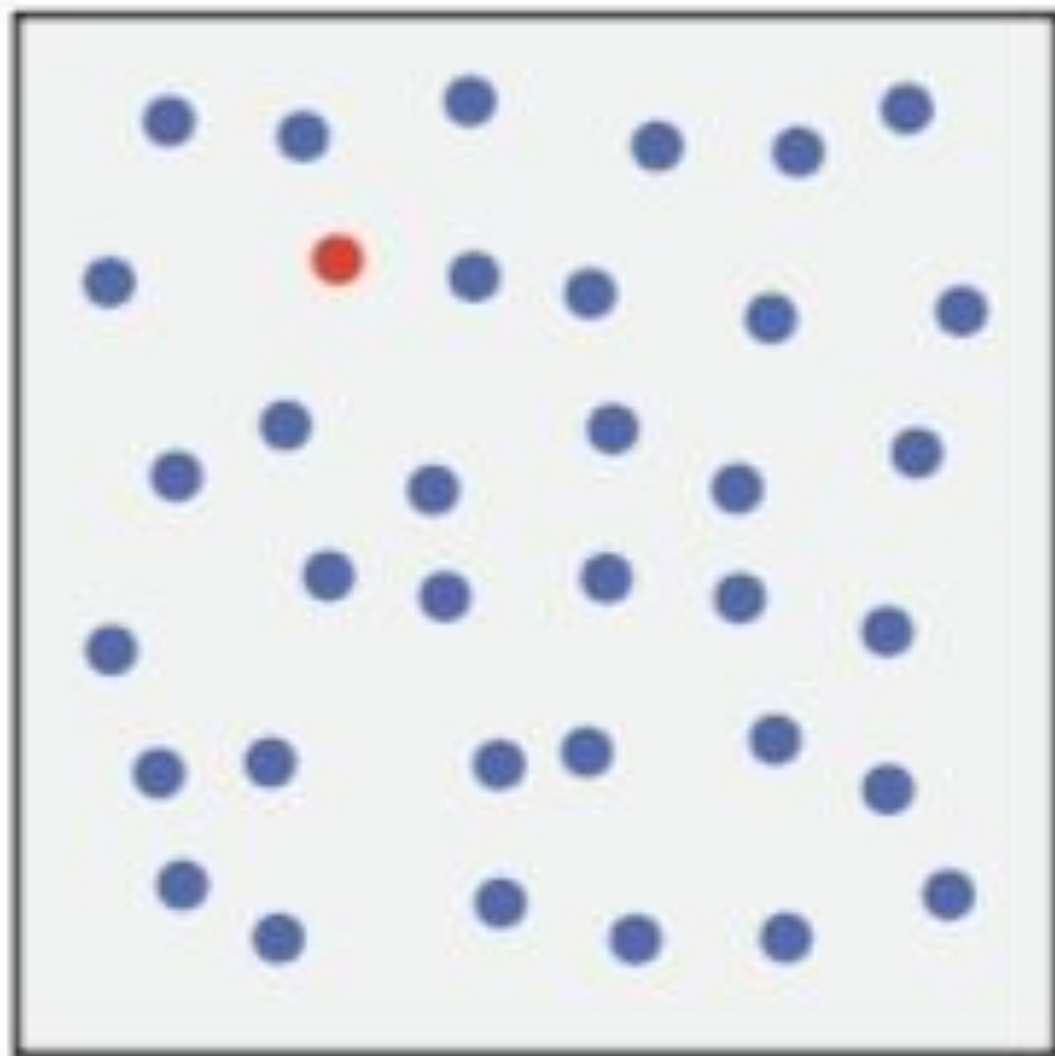
XXXIV

100010

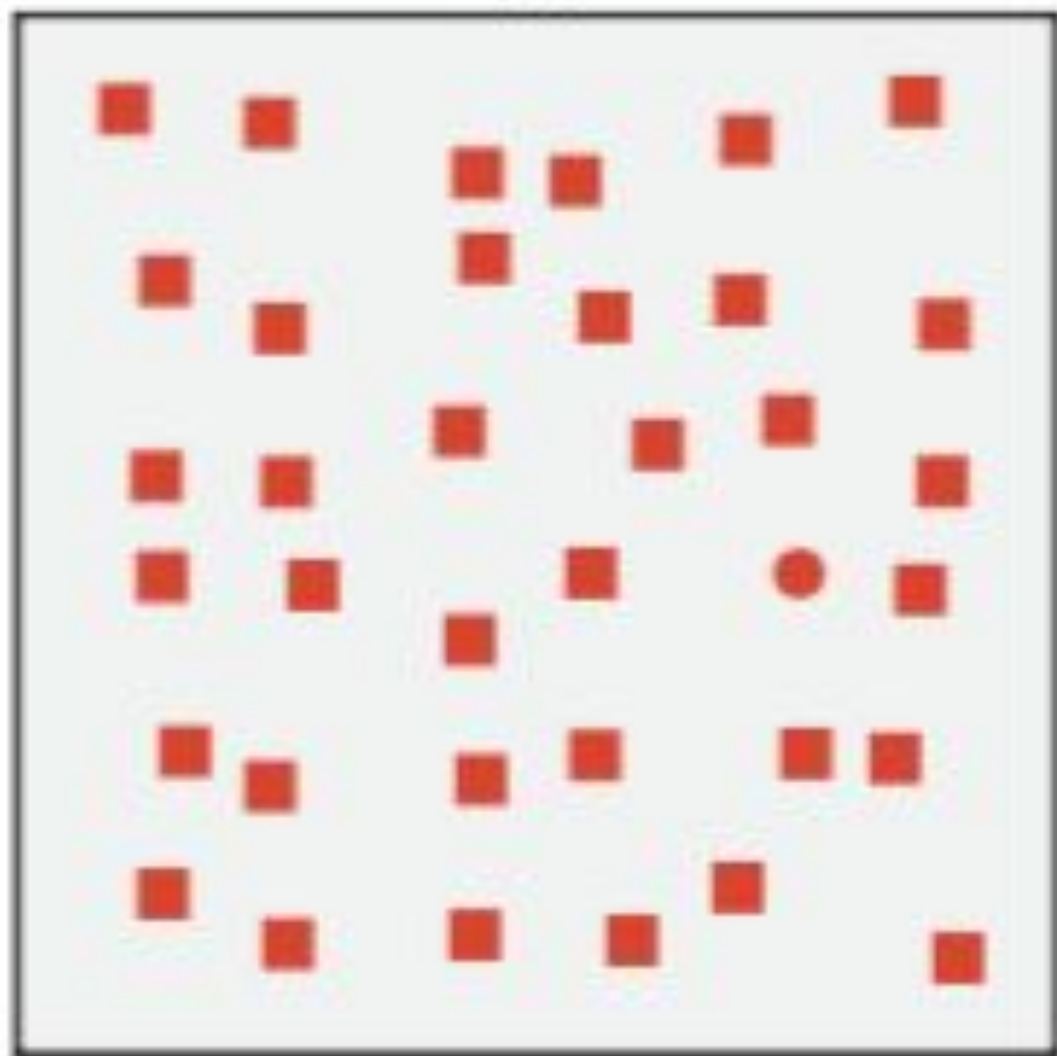
Find every '5'

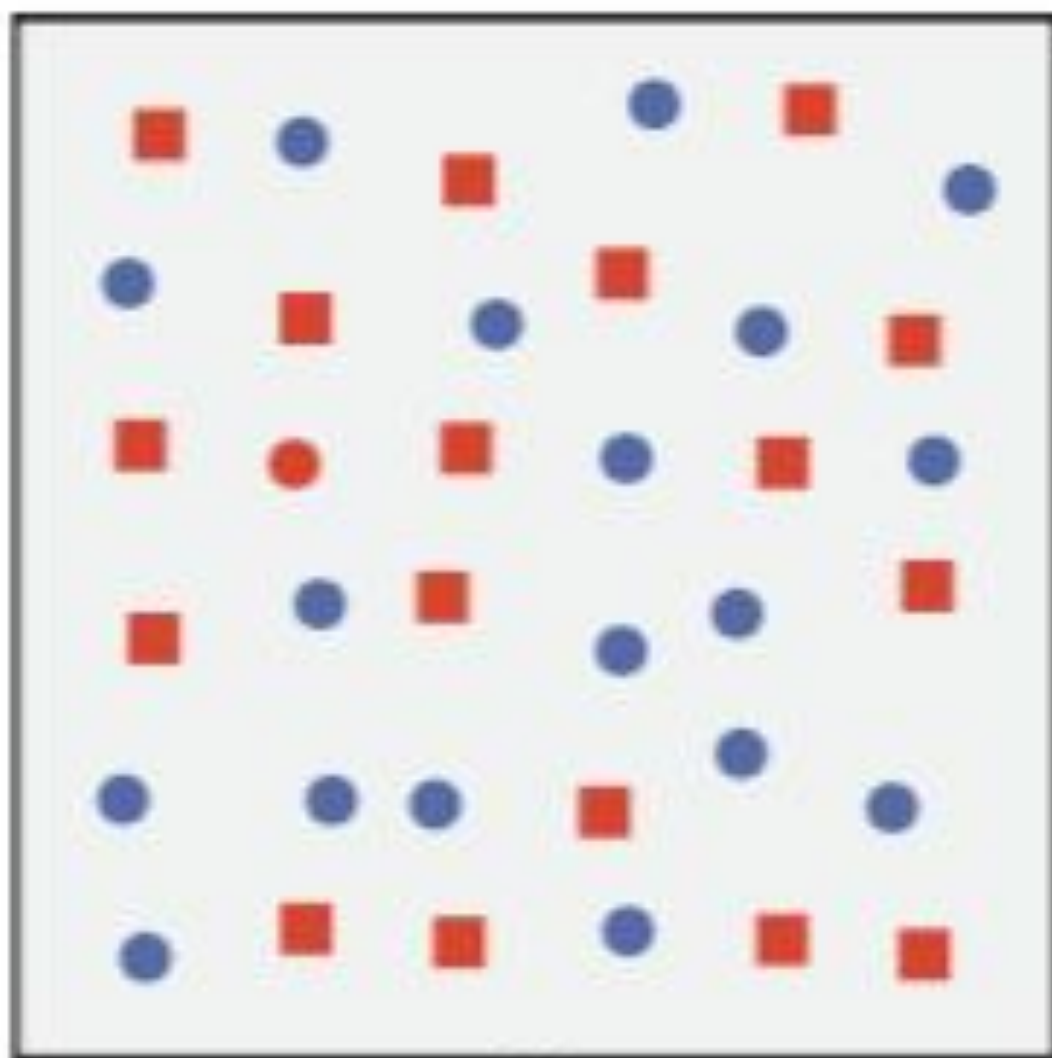
9176867991960386255930486551443
9353652502752141394912668766013
5095444663473545993604636448078
6425620171155906710121069218701
0584448205561385902503573068845
1172102391879794238892669764599
6987227951015948926017759166604
8436204036527029291707742717051
3280506428553158863136868380421
4055135906770329783671748392874

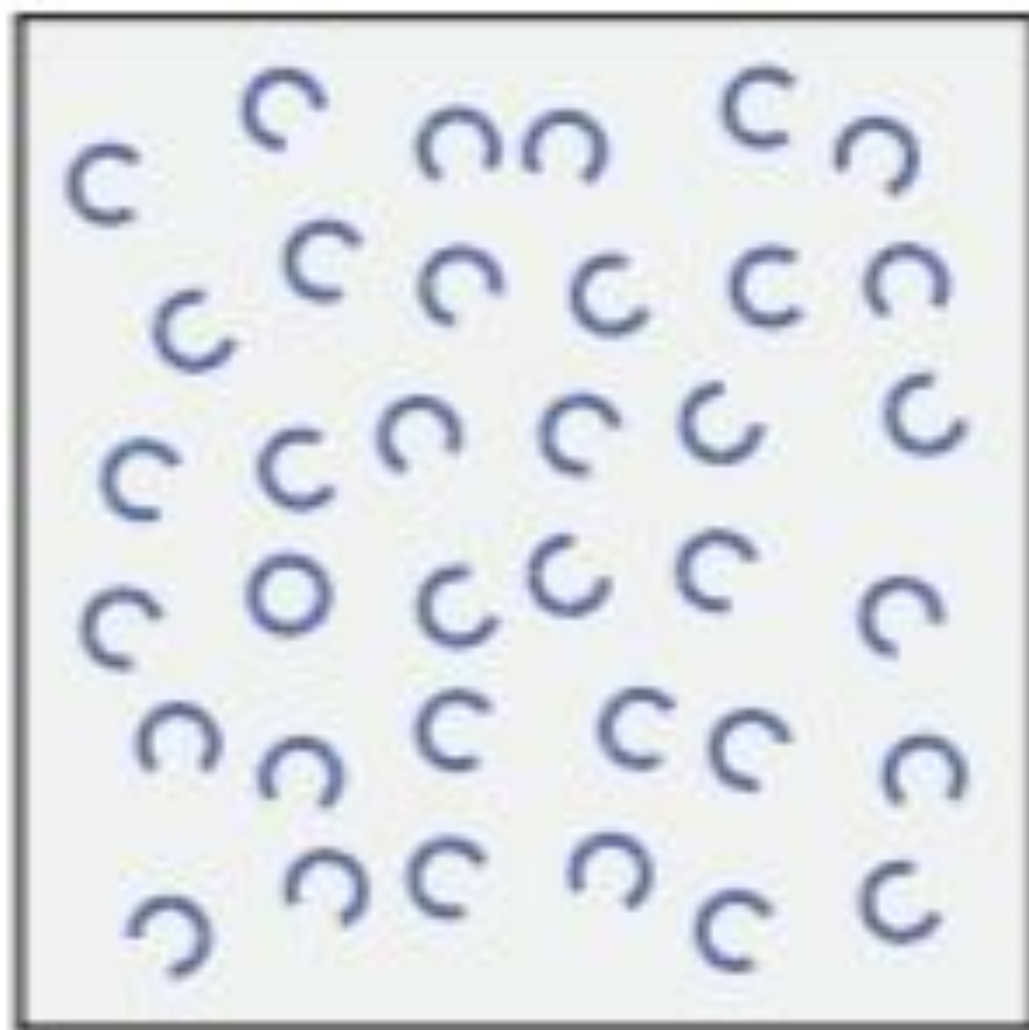
Parallelity & Pre-attentiveness



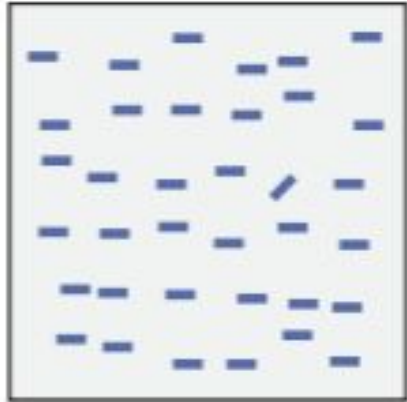
8-7



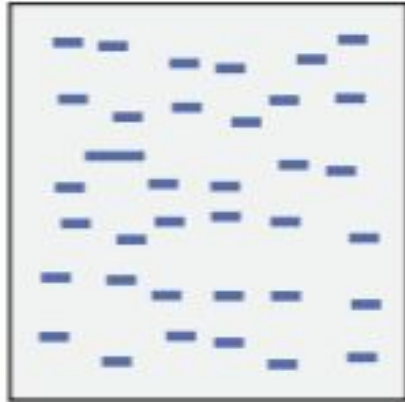




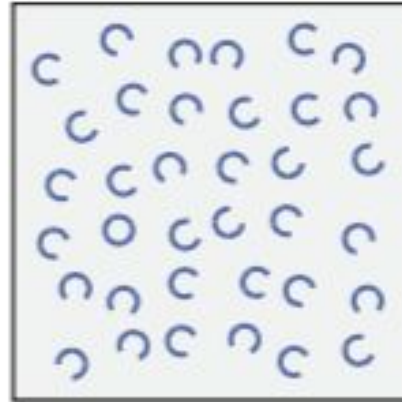
Preattentiveness



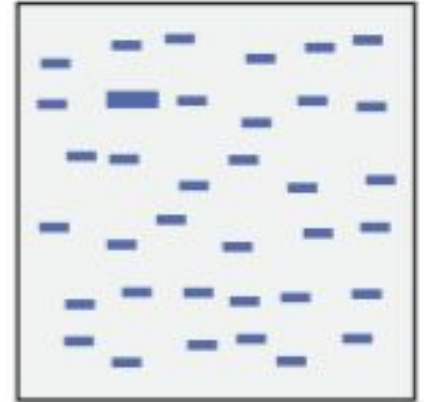
orientation
[16], [17], [18], [19]



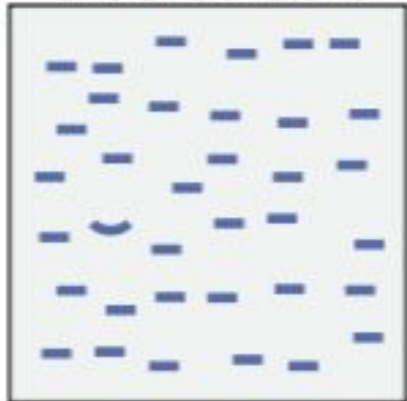
length
[20], [21]



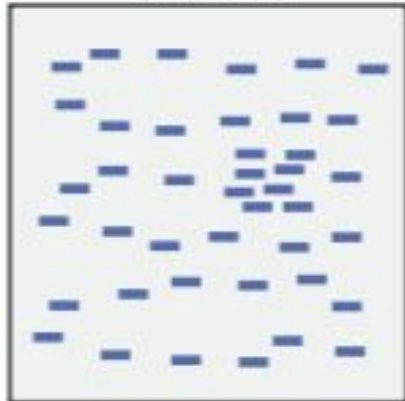
closure
[16]



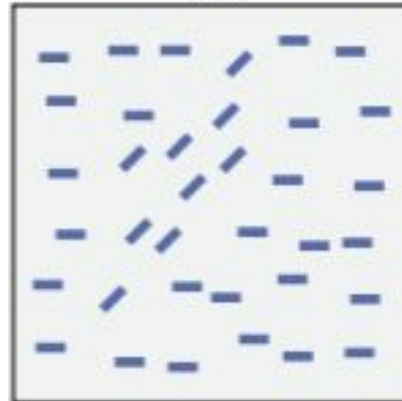
size
[22], [23]



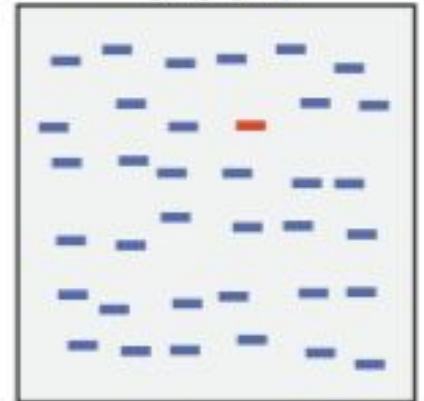
curvature
[21]



density
[23]

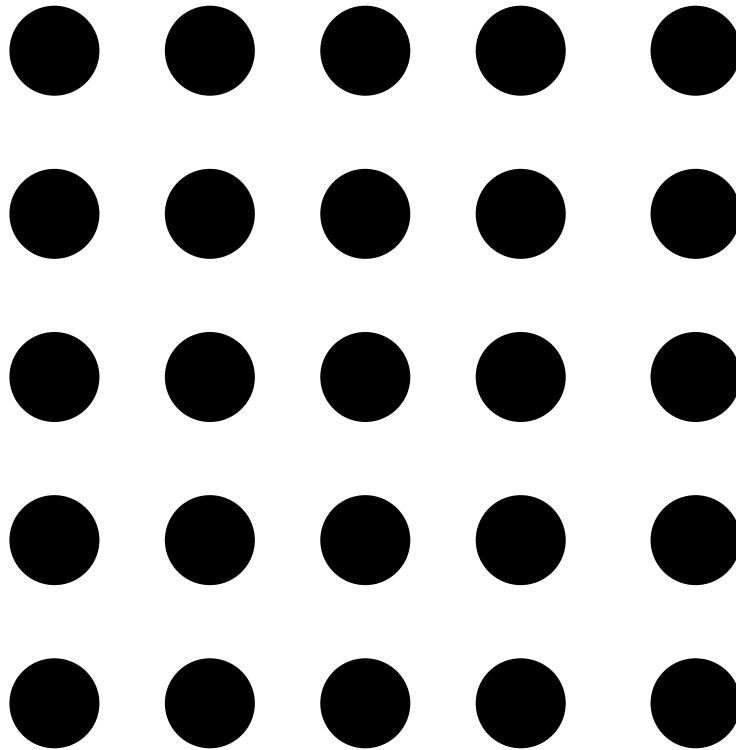


number
[20], [24], [25]

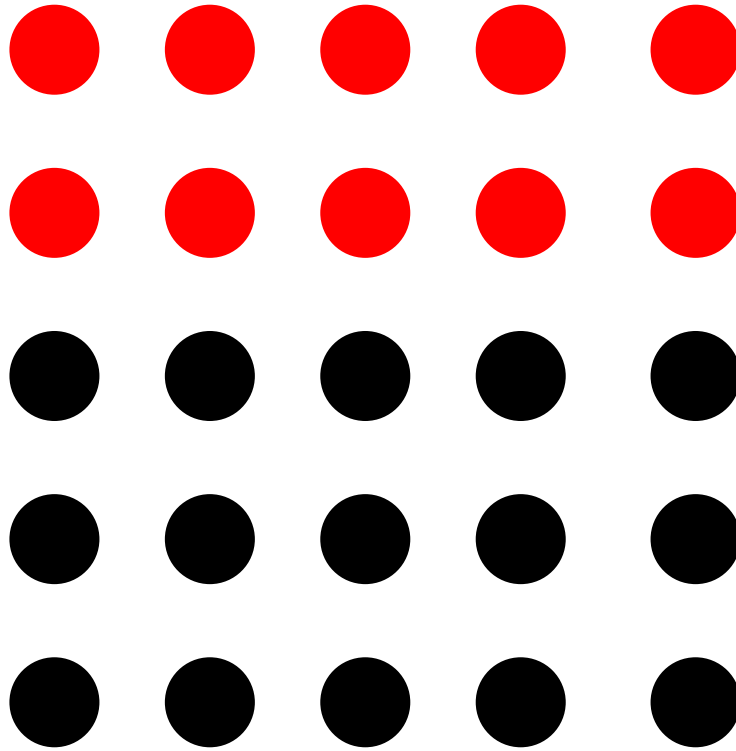


hue
[23], [26], [27], [28], [29]

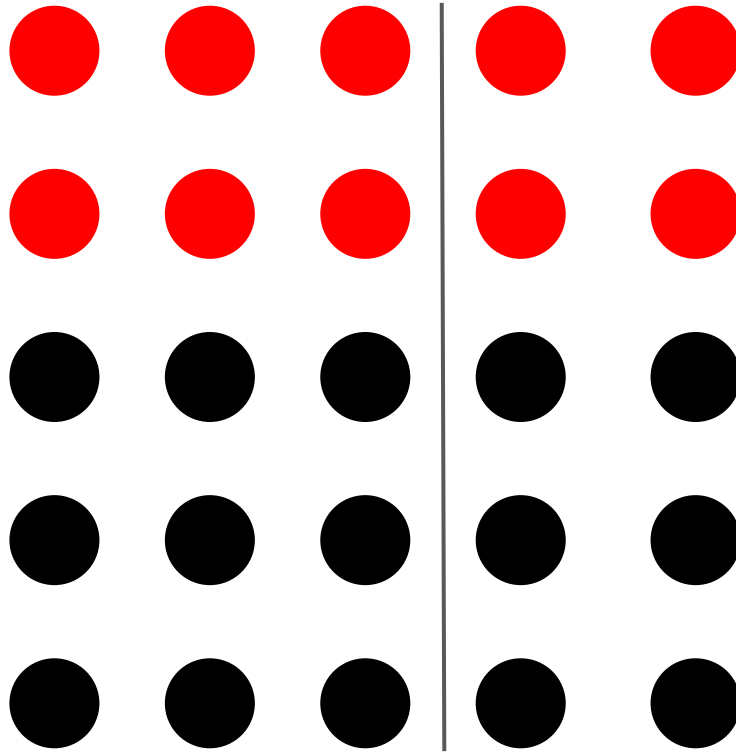
Gestalt Laws



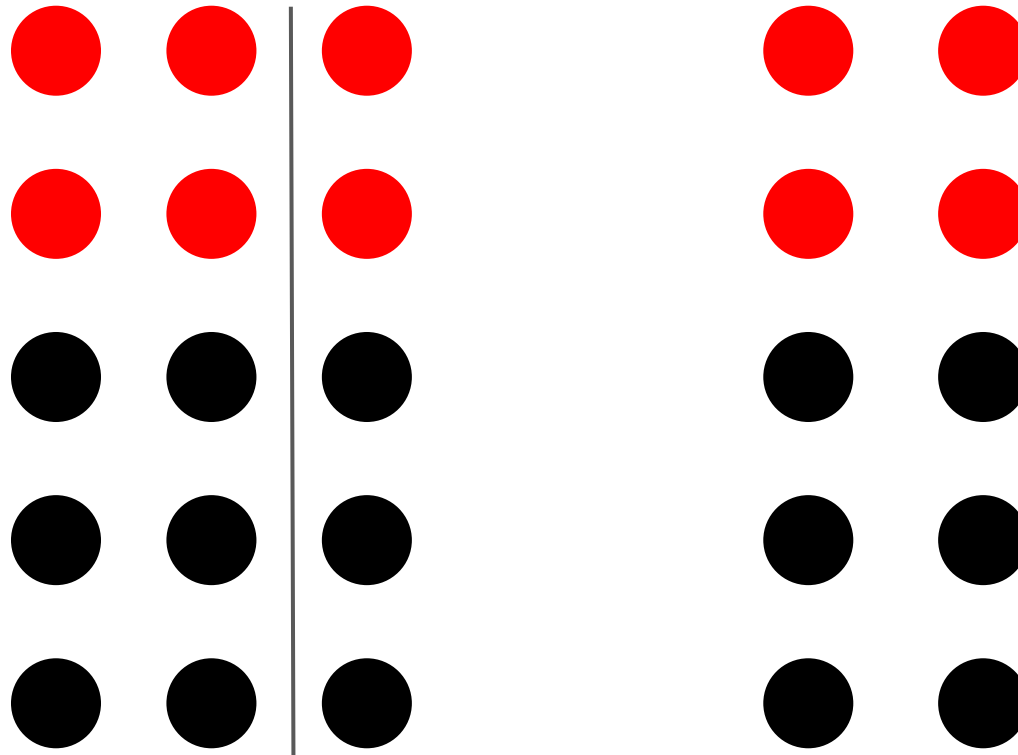
Law of Similarity



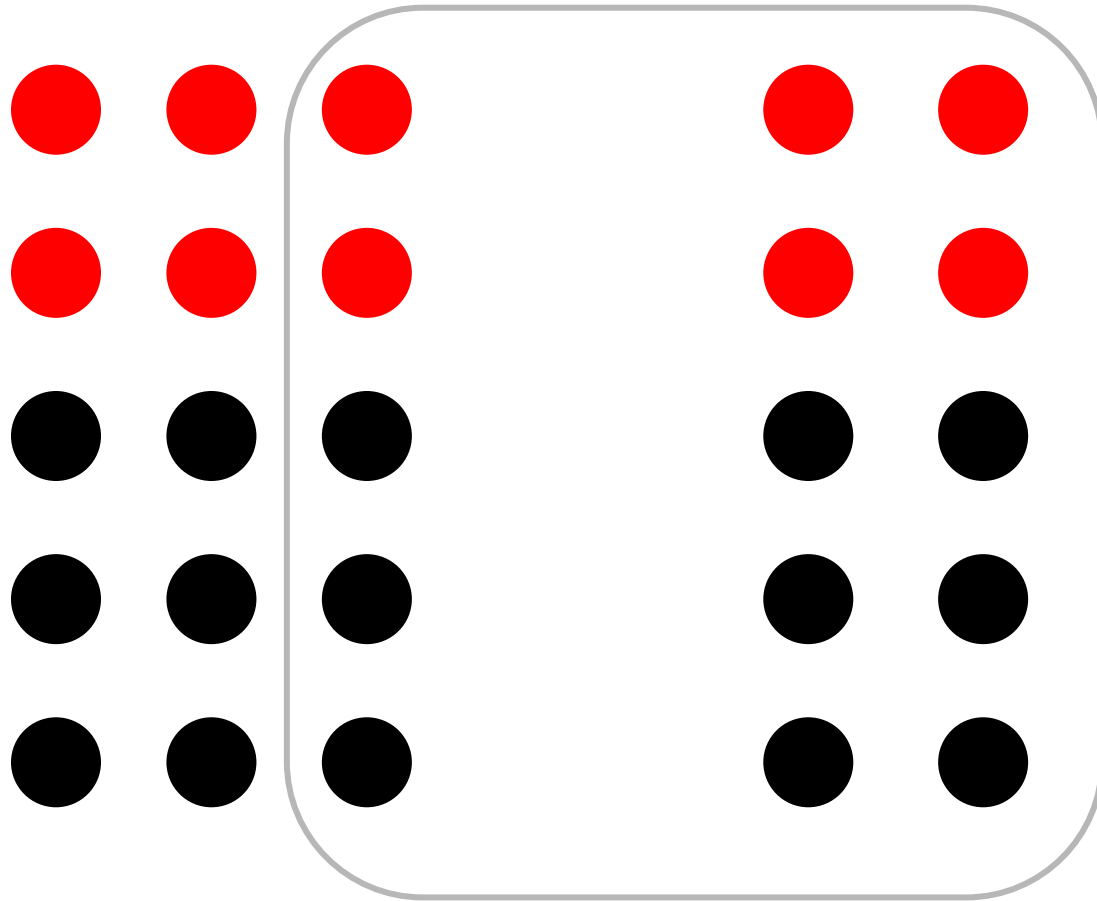
Law of Closure



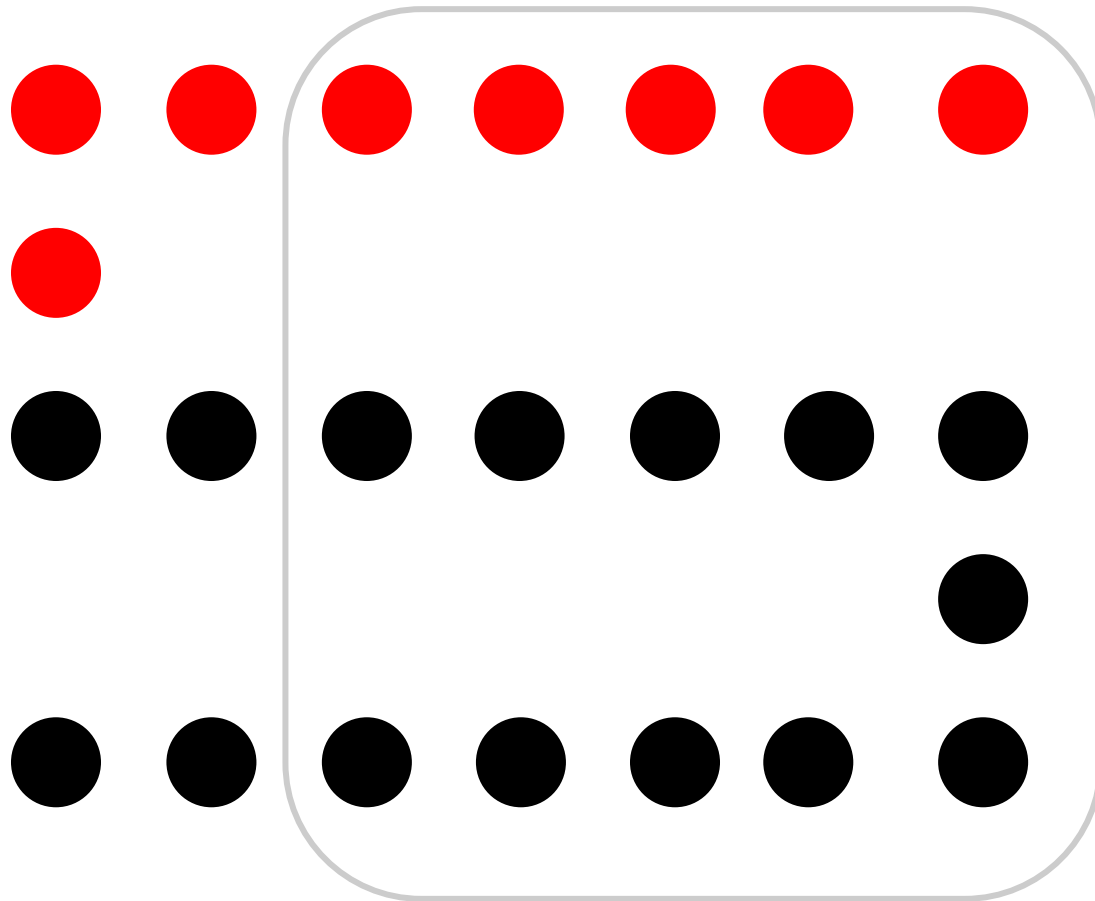
Law of Proximity



Law of Enclosure

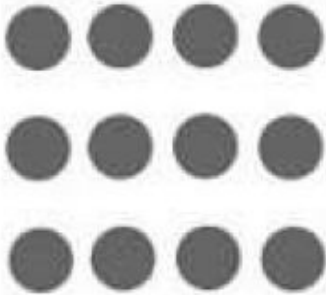


Law of Continuity

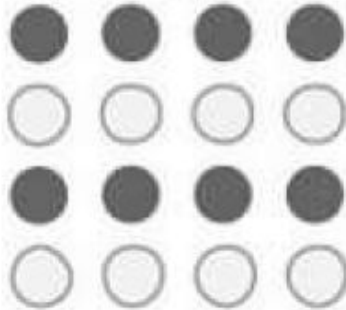


Gestalt Laws

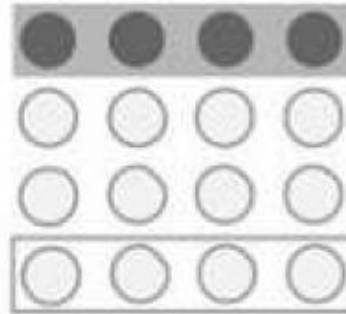
Proximity



Similarity



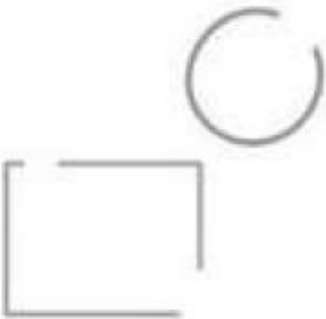
Enclosure



Symmetry



Closure



Continuity



Connection

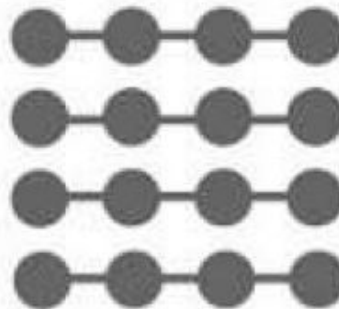


Figure & ground

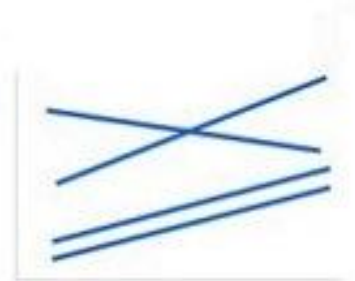


Visual Patterns

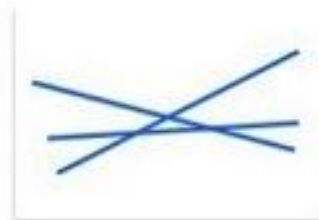
High, low and in between



Non-intersecting and intersecting



Going up, going down and remaining flat



Symmetrical and skewed



Steep and gradual



Wide and narrow



Visual Patterns

Steady and fluctuating



Random and repeating



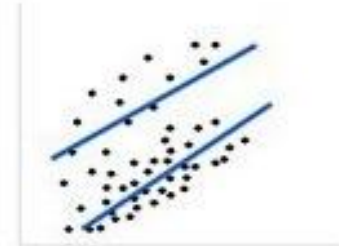
Straight and curved



Clusters and gaps



Tightly and loosely distributed



Normal and abnormal



1.2

Visual Mapping & Visual Variables

Visual Mapping

Data



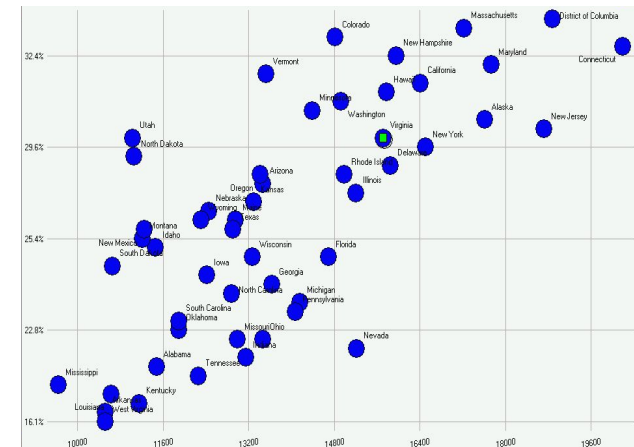
Visualization

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box		7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69		5 4-Not Specified	Small Pack	0.44	6/6/05
69		5 4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06

attribute

cell

item



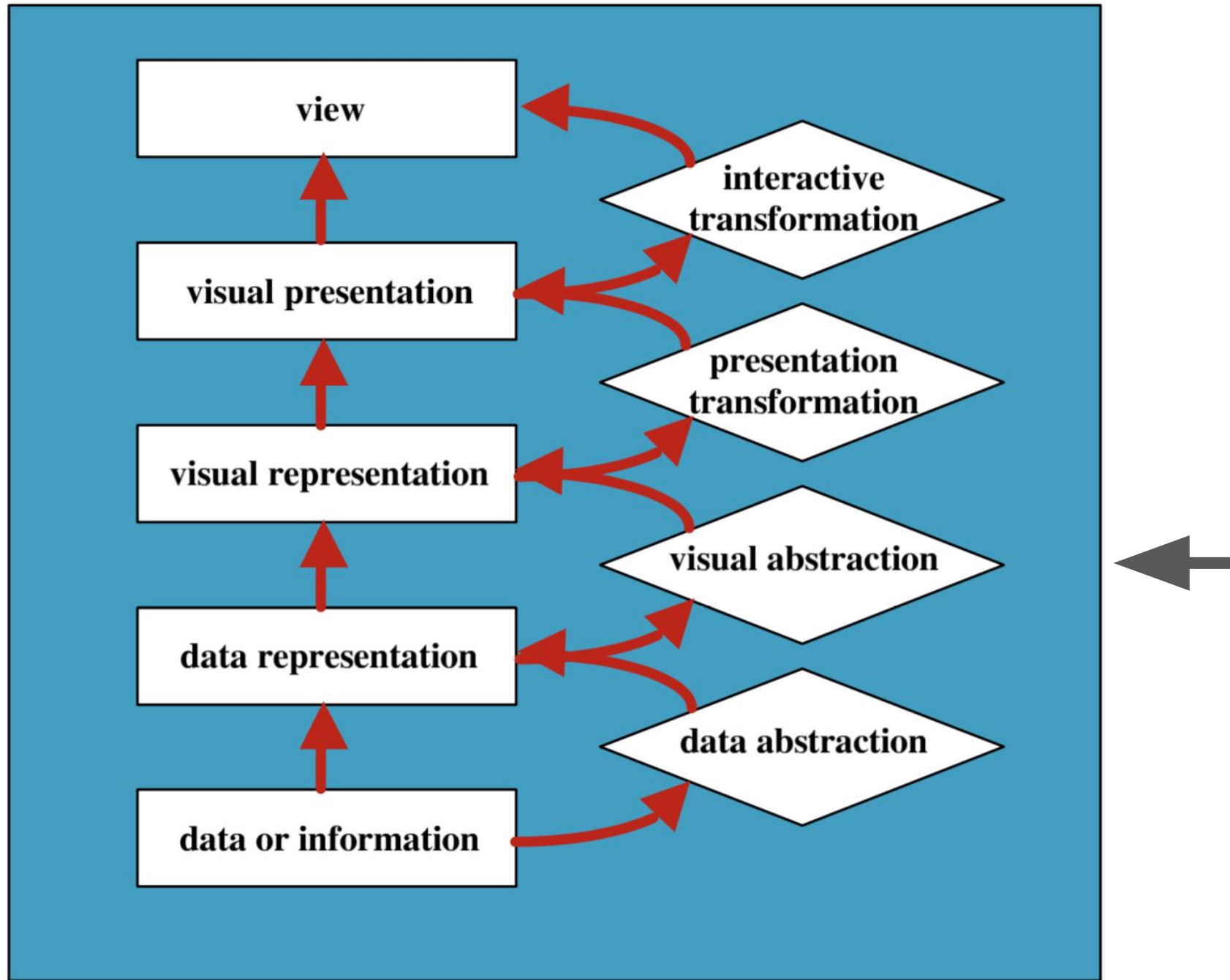


Figure 1: Visualisation pipeline

Data Types

Categorical
(*Qualitative*)



Ordered

Numerical
(*Quantitative*)



A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69		5 4-Not Specified	Small Pack	0.44	6/6/05
69		5 4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06

attribute

item

cell

Visual Mapping

Data



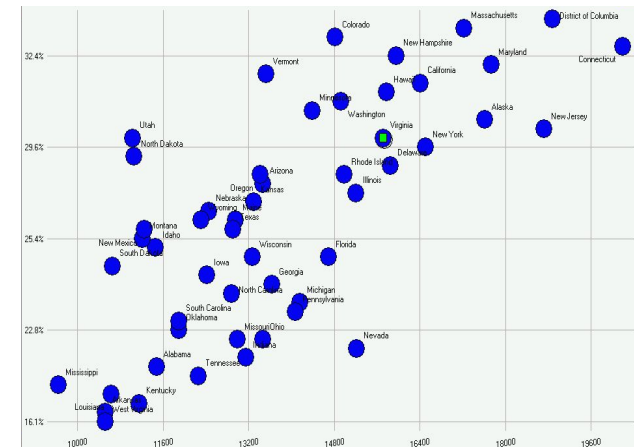
Visualization

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box		7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69		5 4-Not Specified	Small Pack	0.44	6/6/05
69		5 4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06

attribute

cell

item



JACQUES BERTIN

Semiology of Graphics

*diagrams
networks
maps*



Jaques Bertin
(1918–2010)

Visualization Pipeline



Points



Lines

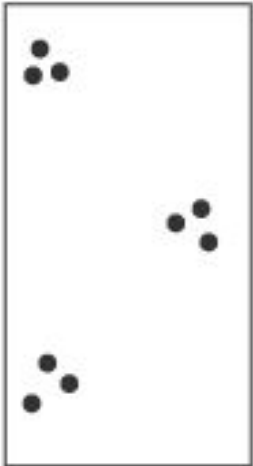


Areas

Visual Variables

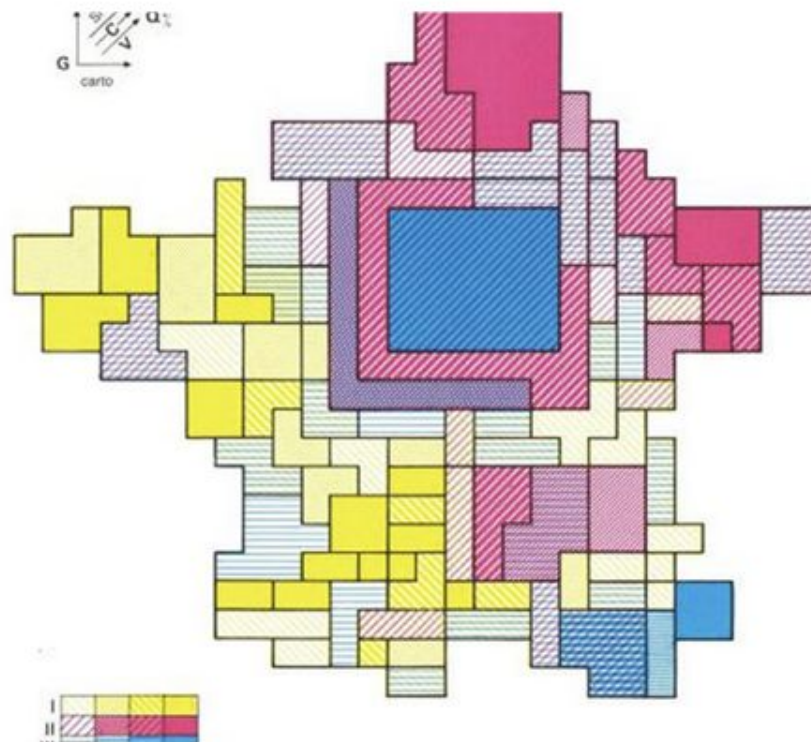
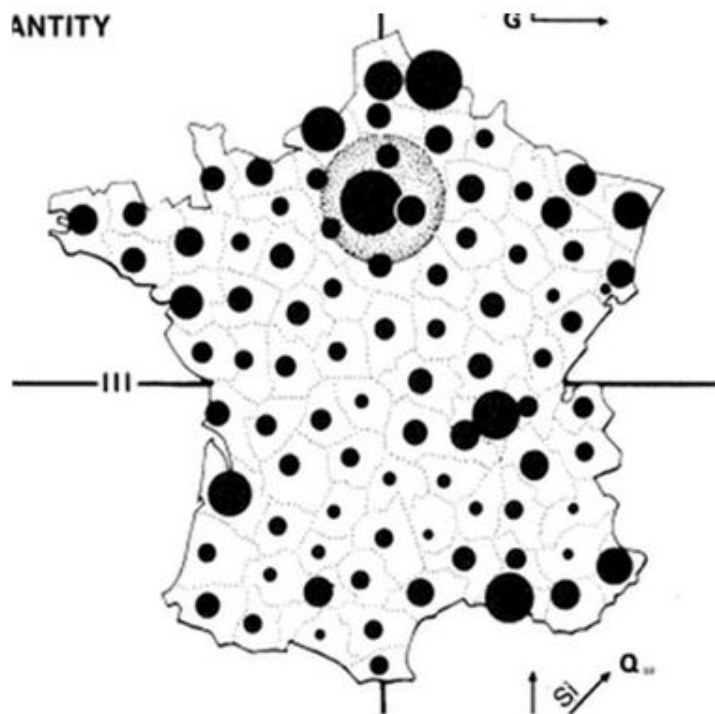
Bertin's Visual Variables

POSITION



Visual Variables

ANTITY



Characteristics of Visual Variables

Selective: *Does it make a mark distinguishable?*

Associative: *Can I group marks?*

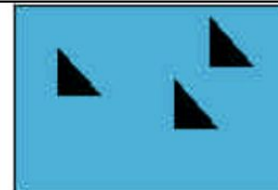
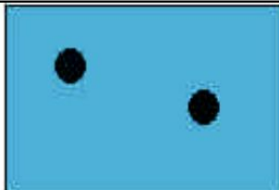
Quantitative: *Can I do maths with it?*

Order: *Does it imply an order on the marks?*

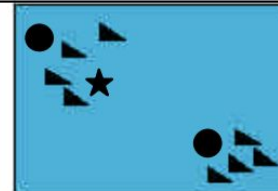
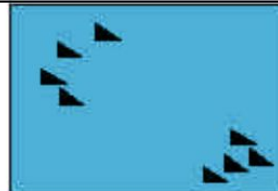
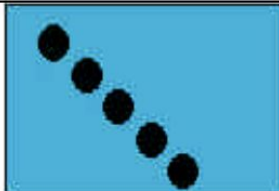
Length: *How many values does the variable has?*

Visual Variable: Position

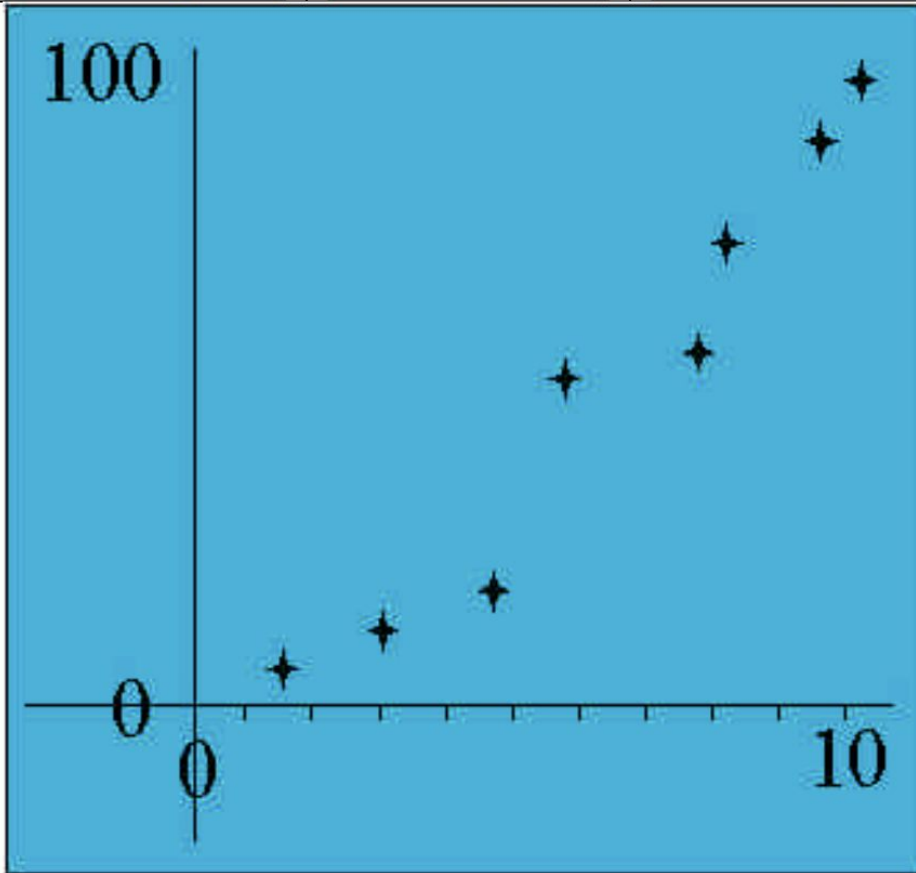
selective



associative




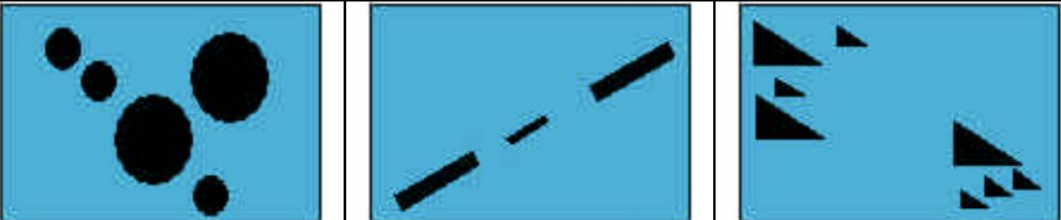
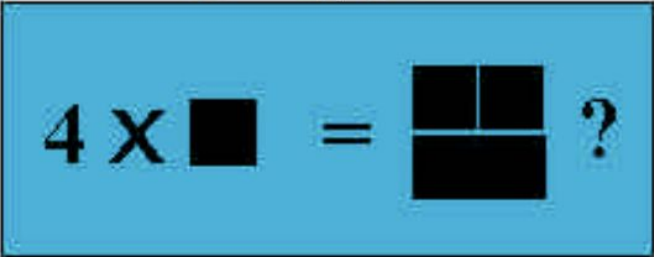
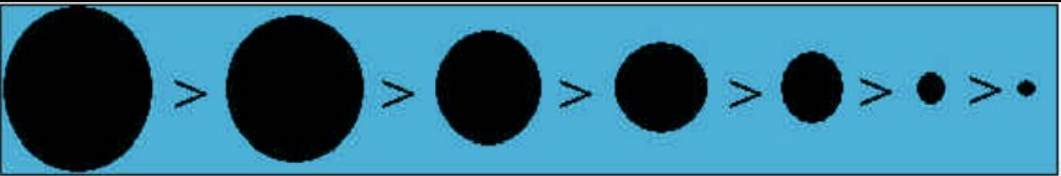
quantitative



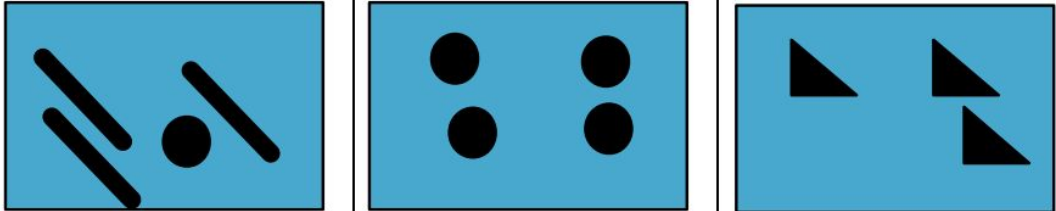
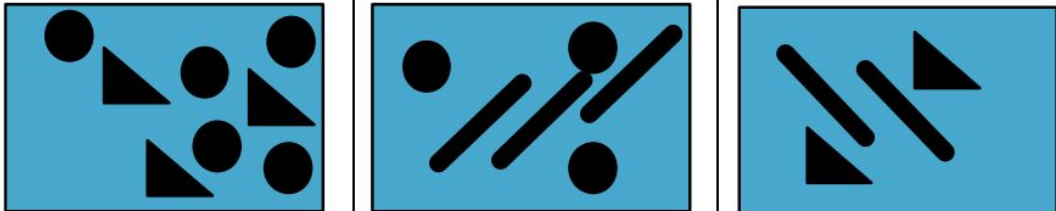


order

length

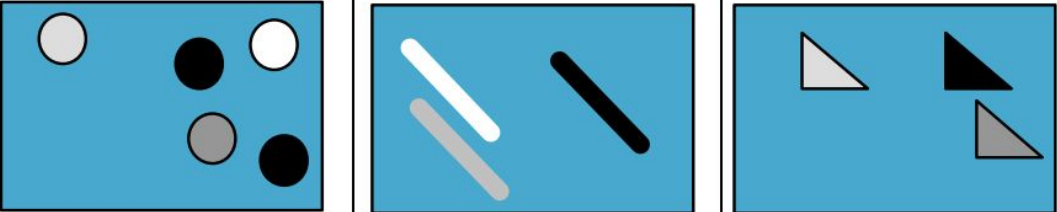
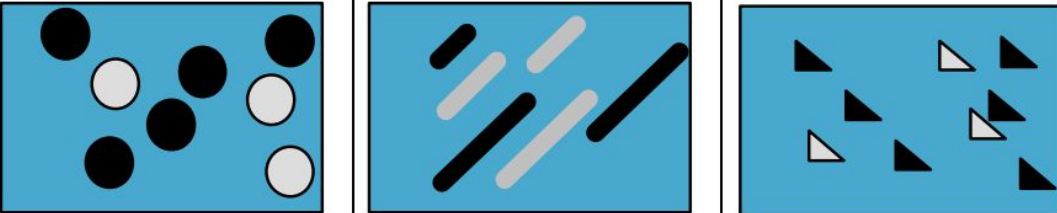
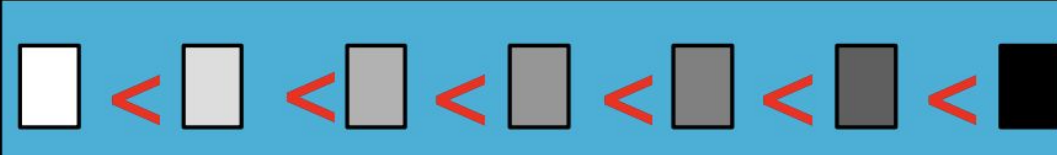
Visual Variable: Size

	selective	
	associative	
	quantitative	
	order	
	Length	<ul style="list-style-type: none"> • theoretically infinite but practically limited • association and selection ~ 5 and distinction ~ 20

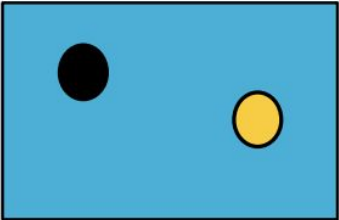
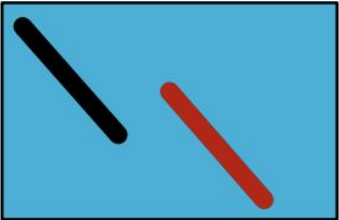
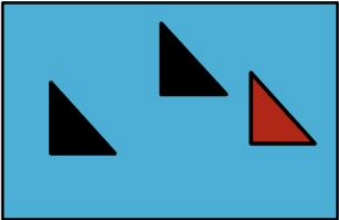
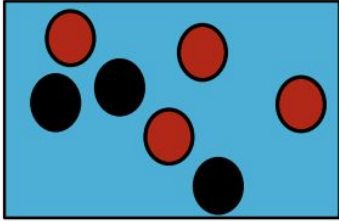
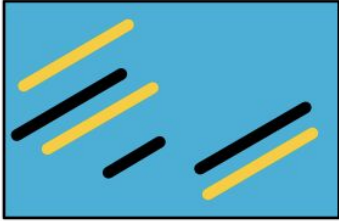
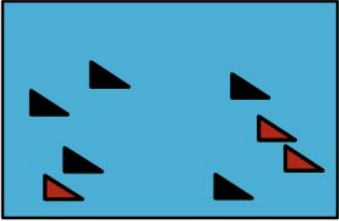
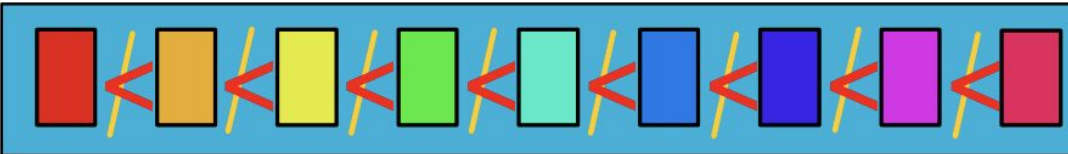

Visual Variable: Shape

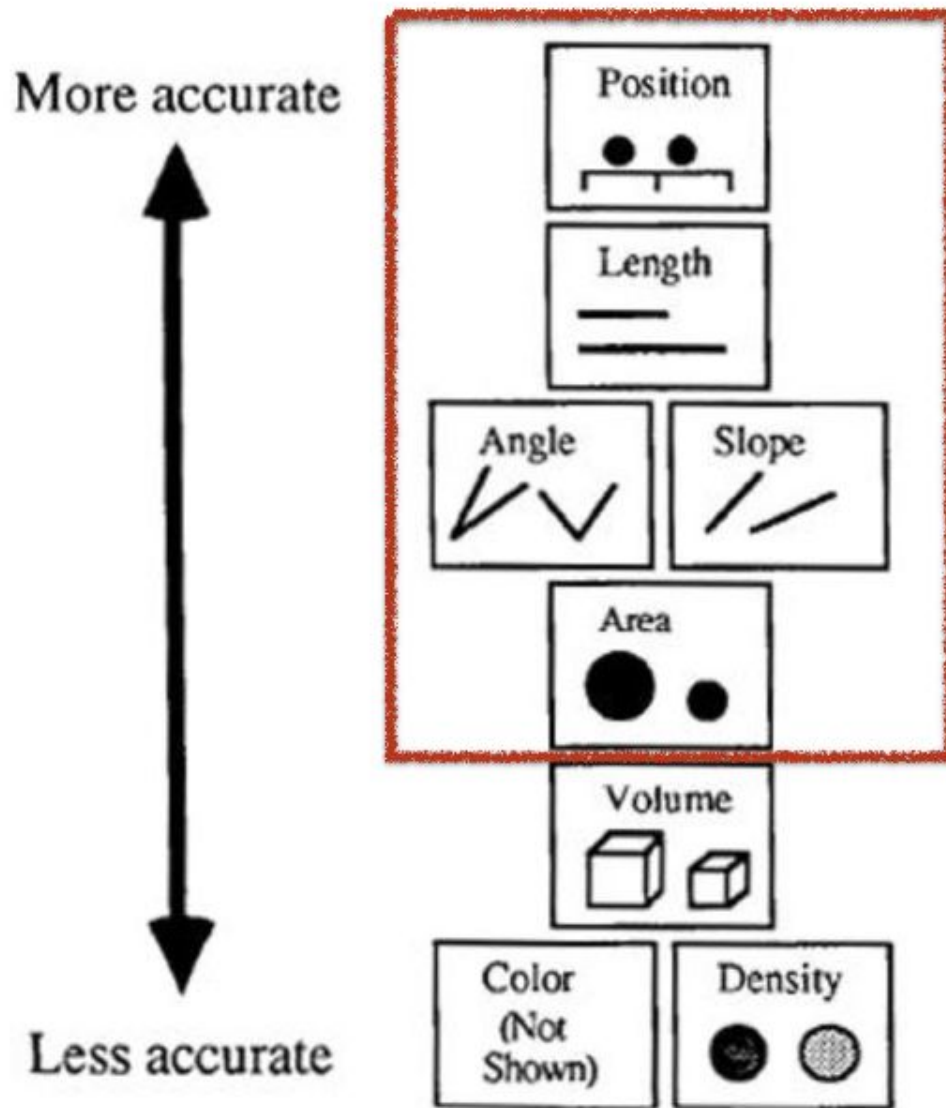
	selective	
	associative	
	quantitative	
	order	
	length	 <p>theoretically infinite</p>

Visual Variable: Value

✓	selective	
✓	associative	
≠	quantitative	
✓	order	
✓	length	<ul style="list-style-type: none"> theoretically infinite but practically limited association and selection ~ < 7 and distinction ~ 10

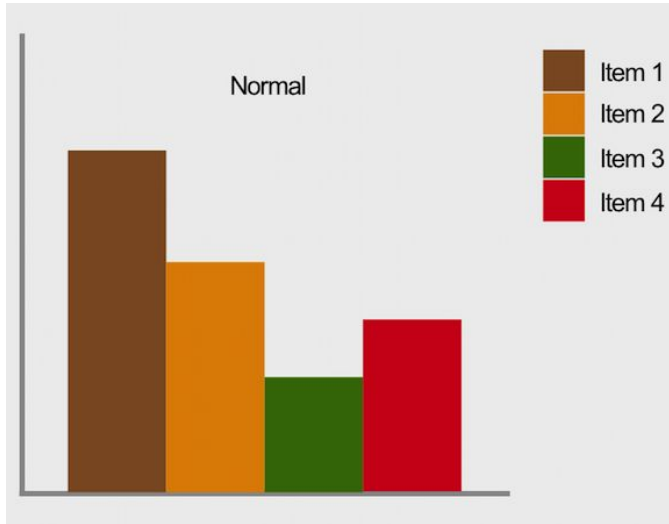
Visual Variable: Colour

✓	selective	  
✓	associative	  
≠	quantitative	
≠	order	
✓	length	 <ul style="list-style-type: none"> theoretically infinite but practically limited association and selection ~ < 7 and distinction ~ 10

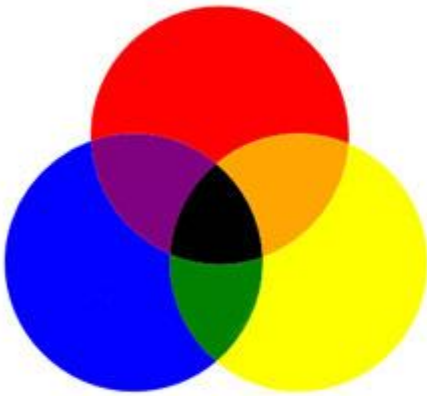


Mackinlay, APT (A Presentation Tool), 1986

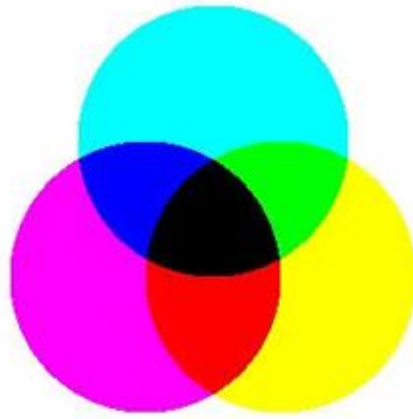
Redundant encoding



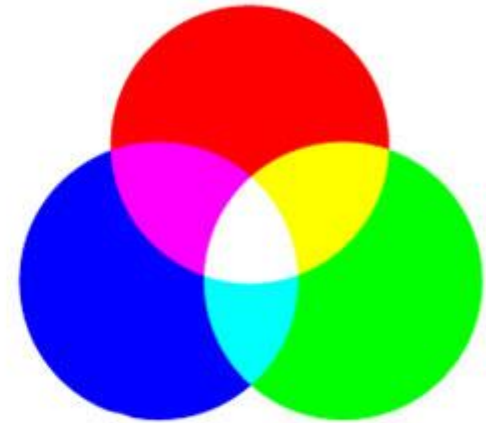
Color Models



RYB



CMYK

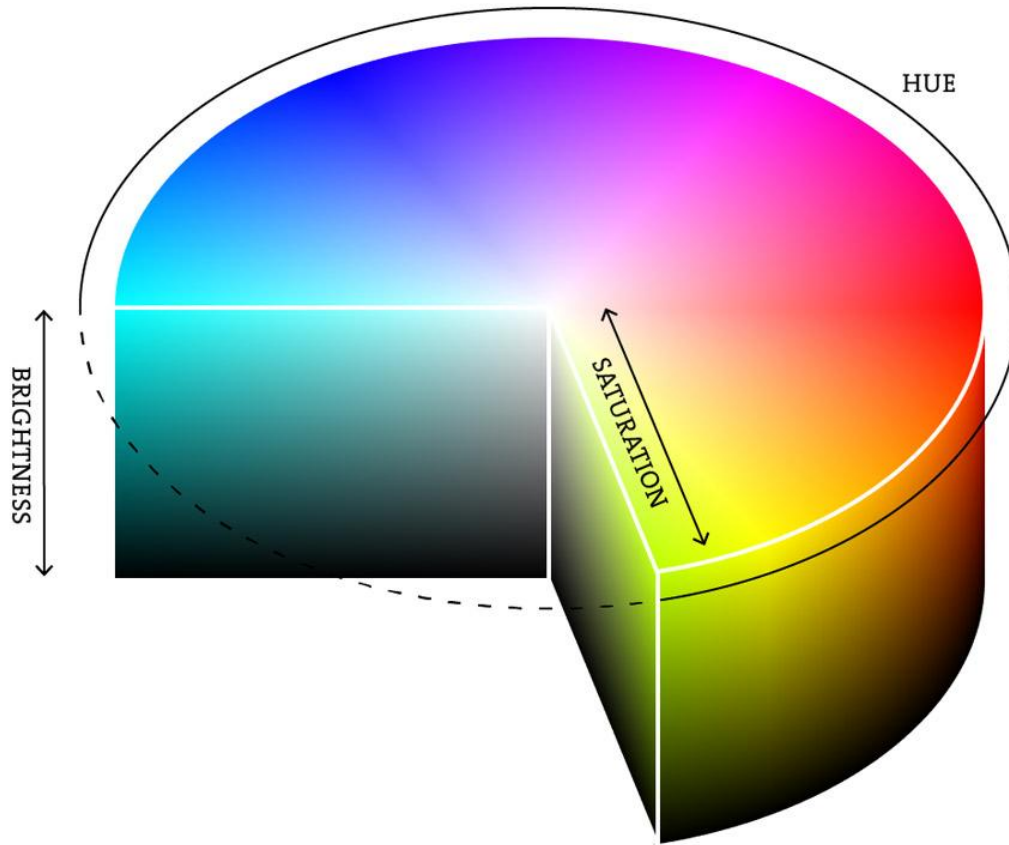


RGB

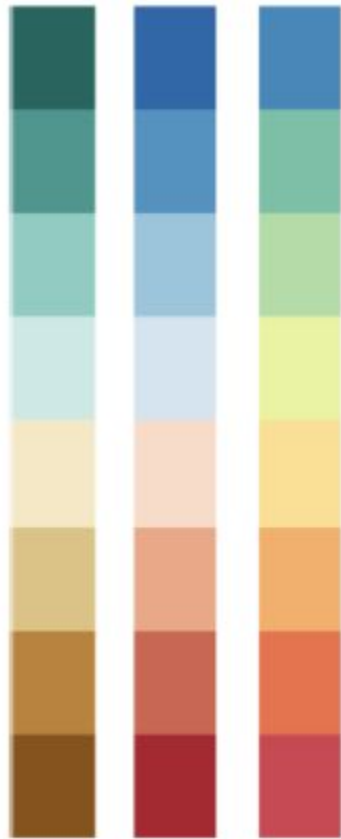
Subtractive
Models

Additive
Models

HSB Model



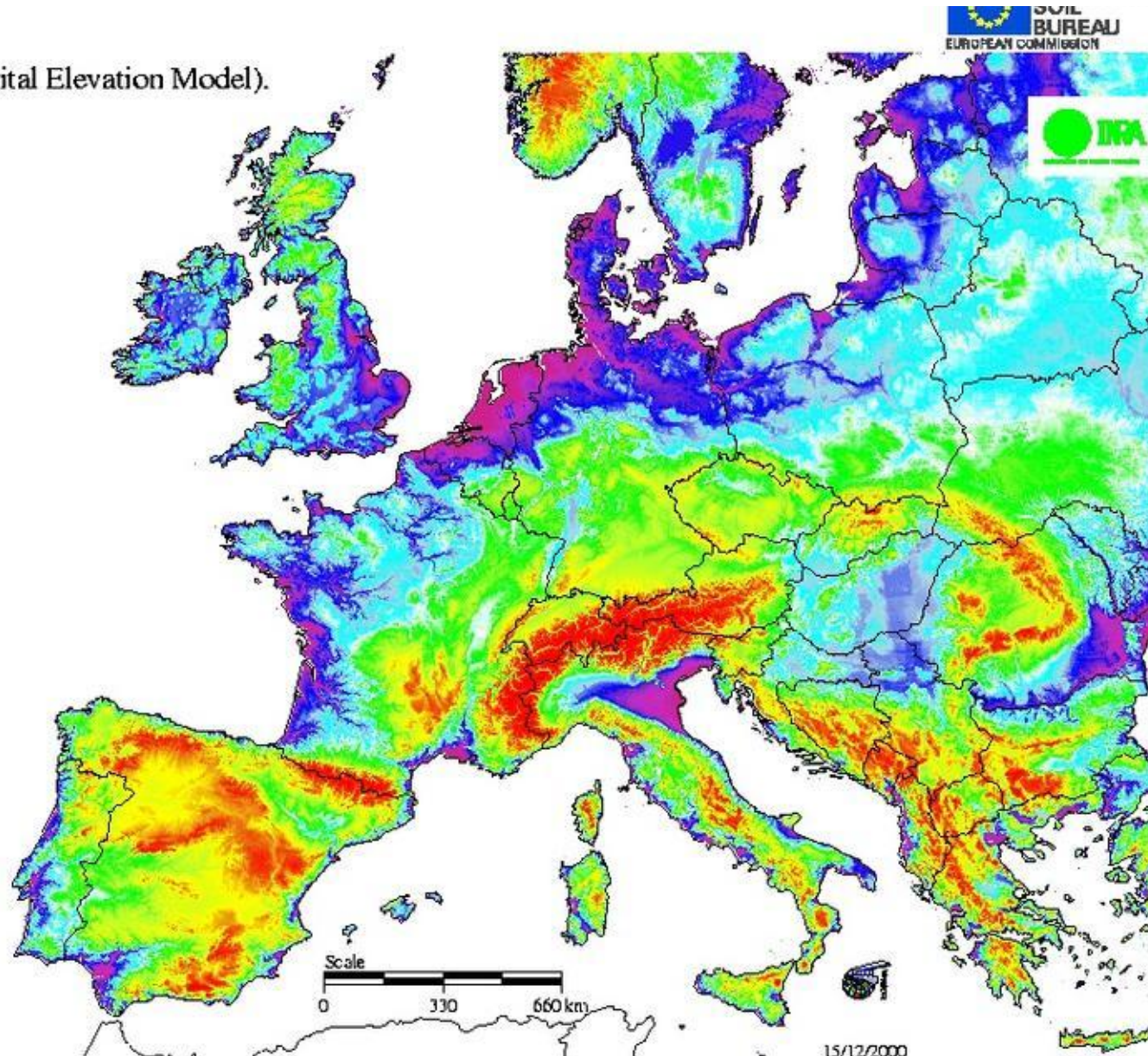
Color Scales



diverging

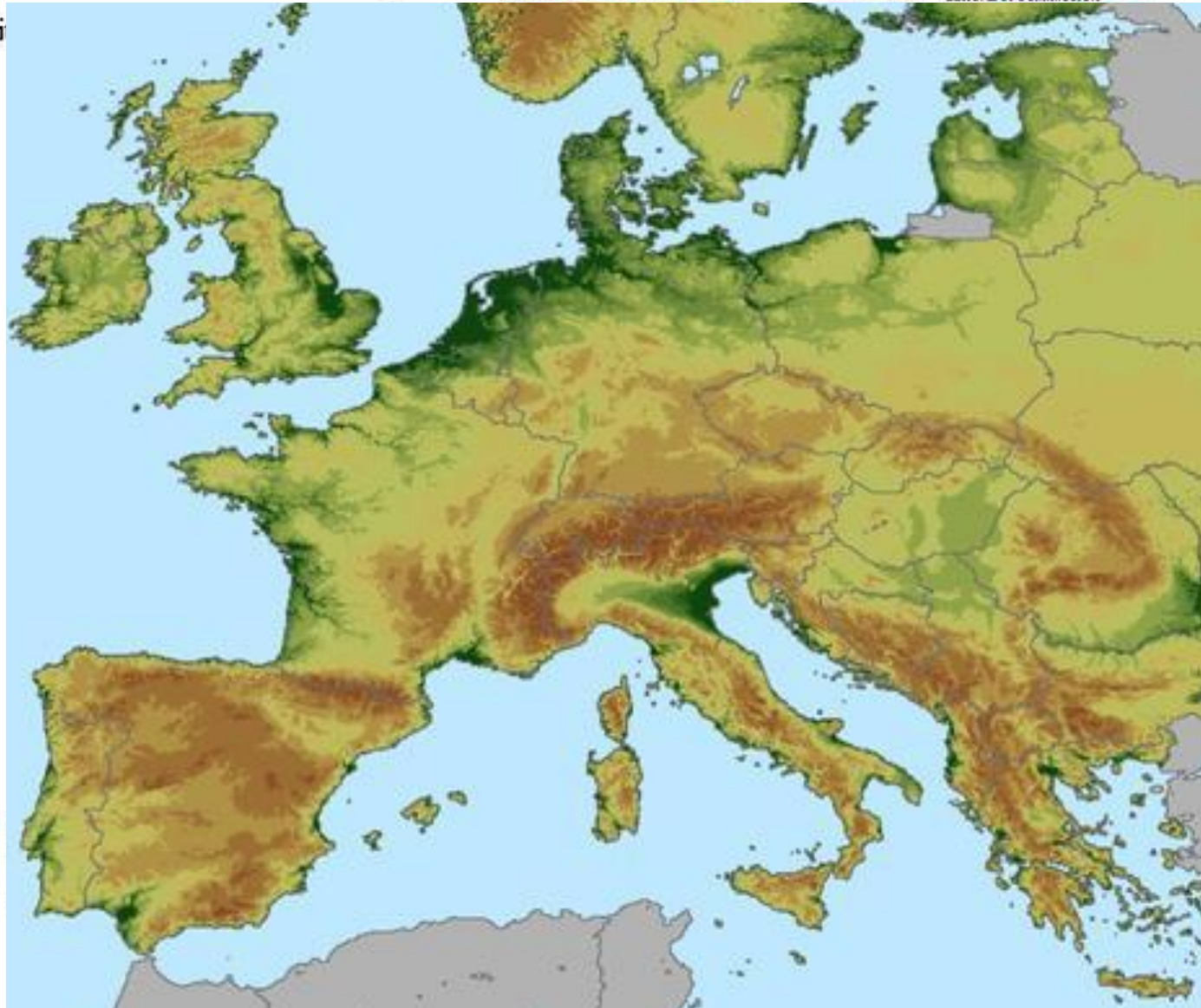
Rainbow Colormap

Digital Elevation Model).



Rainbow Colormap

Digi

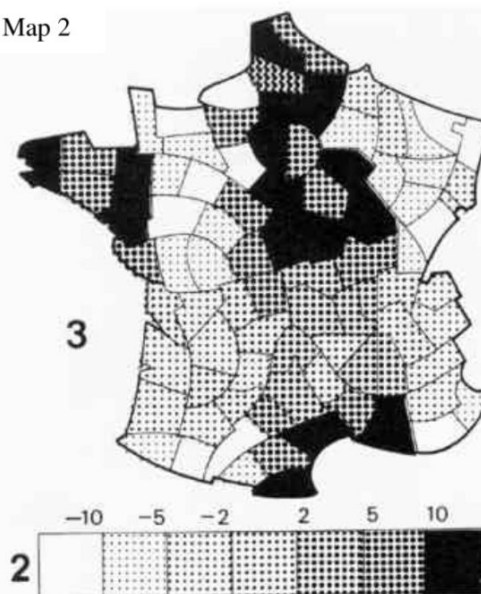


Rainbow Scale Considerations

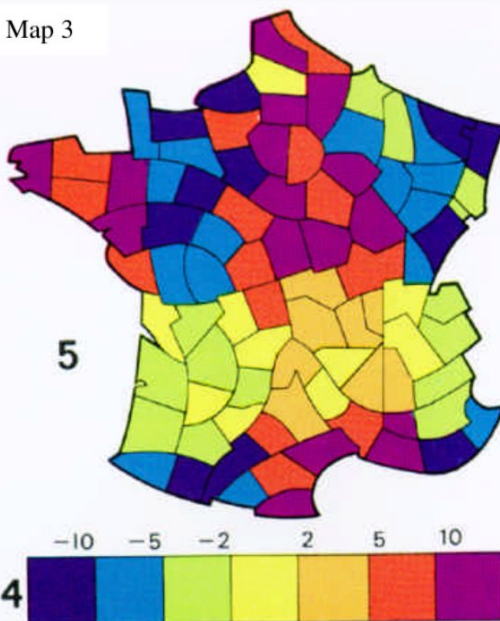
Map 1



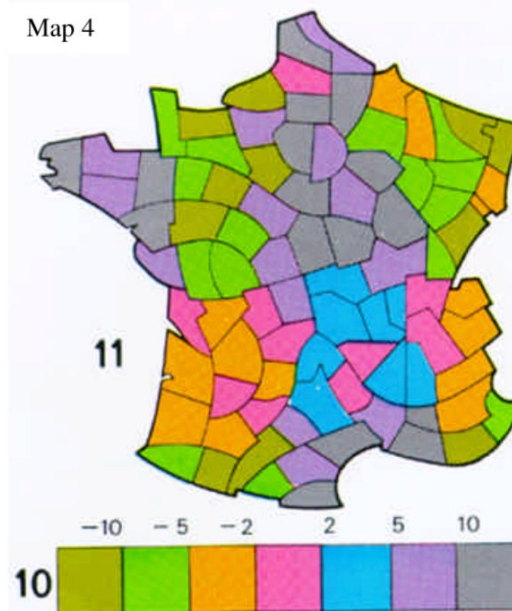
Map 2



Map 3



Map 4



Further Readings

1. Information Visualization—Perception for Design: *Stu Card*
2. Semiology of Graphics: *Jaques Bertin*