Visual Analysis

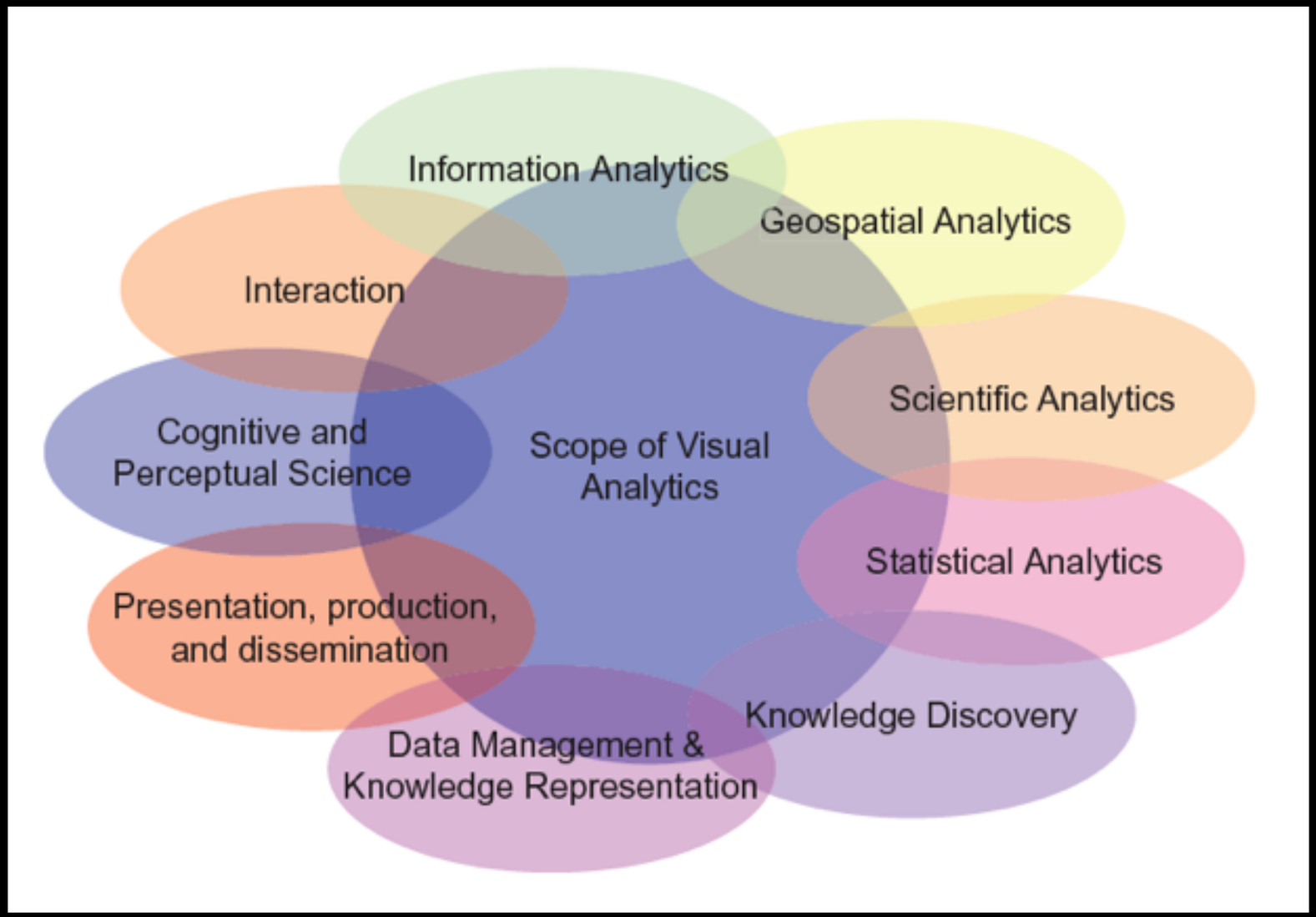
An Introduction to Theory and Tools

Anushka Anand
DSSG 2015 Technical Mentor

6.18.2015

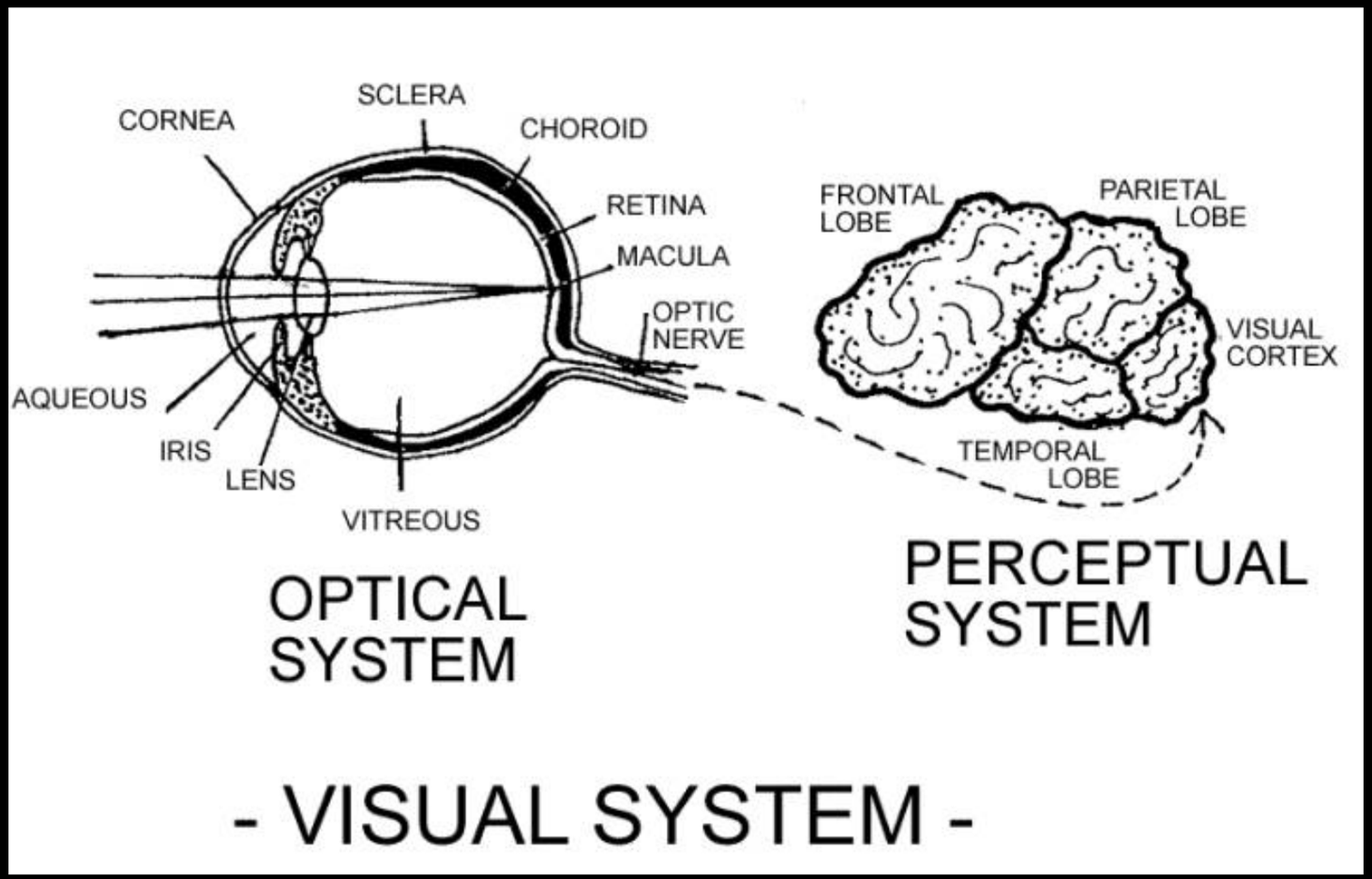
Visual analytics is the science of analytical reasoning facilitated by interactive visual interfaces.

Thomas & Cook, 2005



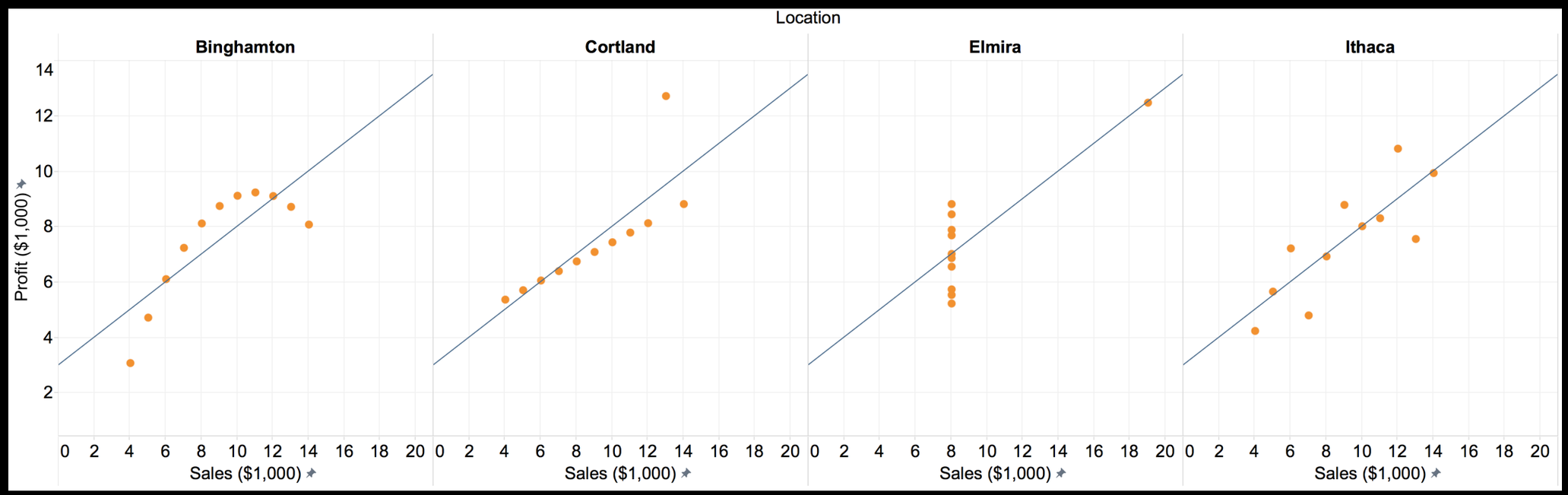
Visual analytics as a highly interdisciplinary field of research. [Keim et al., 2006]

Why Visualization?

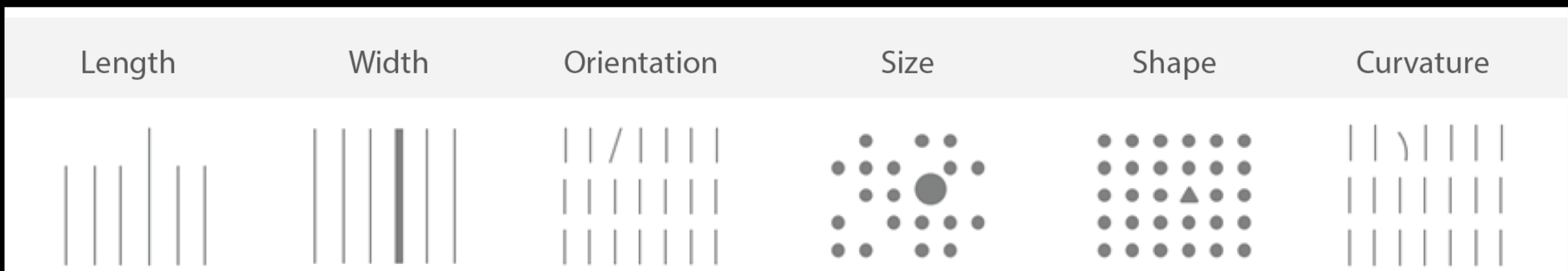


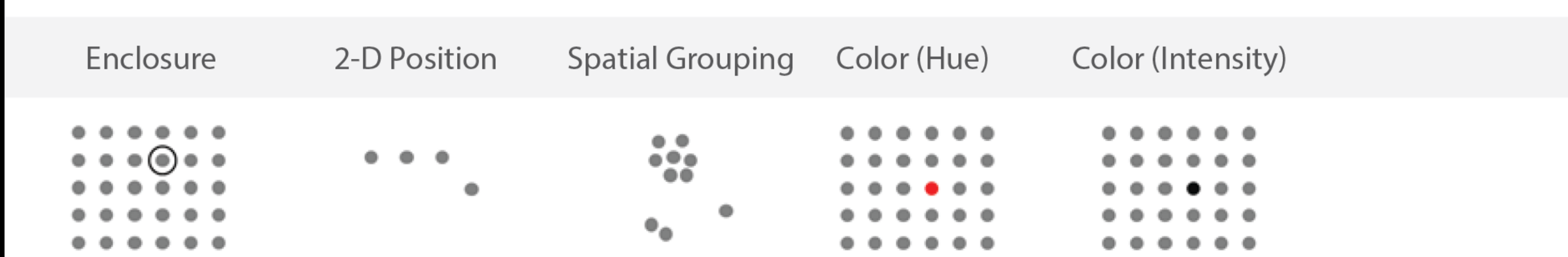
Why Visualization?

	Location							
	Binghamton		**Cortland**		**Elmira**		**Ithaca**	
	Sales ($1,000)	Profit ($1,000)	Sales ($1,000)	Profit ($1,000)	Sales ($1,000)	Profit ($1,000)	Sales ($1,000)	Profit ($1,000)
January	10.00	9.14	10.00	7.46	8.00	6.58	10.00	8.04
February	8.00	8.14	8.00	6.77	8.00	5.76	8.00	6.95
March	13.00	8.74	13.00	12.74	8.00	7.71	13.00	7.58
April	9.00	8.77	9.00	7.11	8.00	8.84	9.00	8.81
May	11.00	9.26	11.00	7.81	8.00	8.47	11.00	8.33
June	14.00	8.10	14.00	8.84	8.00	7.04	14.00	9.96
July	6.00	6.13	6.00	6.08	8.00	5.25	6.00	7.24
August	4.00	3.10	4.00	5.39	19.00	12.50	4.00	4.26
September	12.00	9.13	12.00	8.15	8.00	5.56	12.00	10.84
October	7.00	7.26	7.00	6.42	8.00	7.91	7.00	4.82
November	5.00	4.74	5.00	5.73	8.00	6.89	5.00	5.68



Pre-attentive attributes





How many 9 s?

3 3 0 3 0 1 8 7 6 8 2 1 4 0 3 8 3 7 7 2 0 5 2 3 2 7 0 2 0
7 1 4 6 0 2 1 3 2 7 6 0 2 5 6 3 2 5 7 6 3 3 0 2 0 3 0 7 2
8 7 5 7 2 8 3 8 7 7 8 2 0 7 7 5 2 3 1 1 5 6 3 8 4 7 8 2 0
0 5 0 5 1 6 1 7 5 6 8 0 4 4 6 7 4 7 1 4 0 0 8 4 4 3 0 3 2
2 4 3 1 3 5 4 9 5 0 7 6 0 7 4 3 1 8 2 7 3 4 6 0 2 4 8 2 3
8 6 2 2 6 5 4 6 7 0 7 6 0 0 3 9 0 2 4 7 1 7 2 3 3 5 8 7 0
0 8 4 5 1 3 1 7 6 4 5 4 1 2 4 5 3 3 5 4 9 6 7 7 6 3 4 2 5
4 7 7 0 2 2 0 1 1 7 7 7 0 2 6 6 4 7 5 8 6 1 4 3 7 8 5 4 6
4 3 6 6 4 6 6 2 8 4 8 5 3 7 8 8 1 3 8 5 4 5 7 4 0 3 2 8 4
5 5 0 3 5 3 5 3 8 3 2 3 8 2 3 1 6 2 7 2 4 6 3 6 4 4 3 2 5
4 4 0 2 1 7 2 4 4 7 4 1 9 2 4 5 2 5 0 4 0 0 5 3 6 3 3 6 7
7 4 6 6 8 7 5 7 9 2 0 2 8 8 8 8 3 2 4 2 6 4 0 4 6 3 7 2 1
0 1 7 1 5 9 1 4 2 8 7 3 7 1 4 5 1 8 7 8 0 5 1 7 0 5 8 8 1
2 8 5 2 1 2 8 7 7 6 2 5 6 2 6 4 1 5 1 6 1 2 1 1 0 5 6 4 0
2 1 1 7 7 2 0 0 1 8 7 0 2 9 0 2 8 5 7 8 4 6 0 6 5 0 7 1 2
0 5 2 4 1 5 3 3 1 5 5 1 4 0 1 6 4 3 3 9 8 8 3 4 6 8 4 8 6
7 3 7 5 2 4 0 2 7 6 3 8 5 5 4 5 8 8 7 5 5 6 5 6 7 9 7 7 4
0 3 2 8 1 4 4 6 0 8 2 3 0 1 3 4 6 2 0 5 7 7 3 6 1 8 7 3 5
4 4 8 3 3 3 5 0 1 0 3 8 6 3 2 0 5 0 6 1 3 3 4 3 6 1 5 8 6
1 0 2 2 7 6 3 3 0 8 8 0 3 1 8 8 1 2 1 7 5 2 9 3 5 8 3 2 5

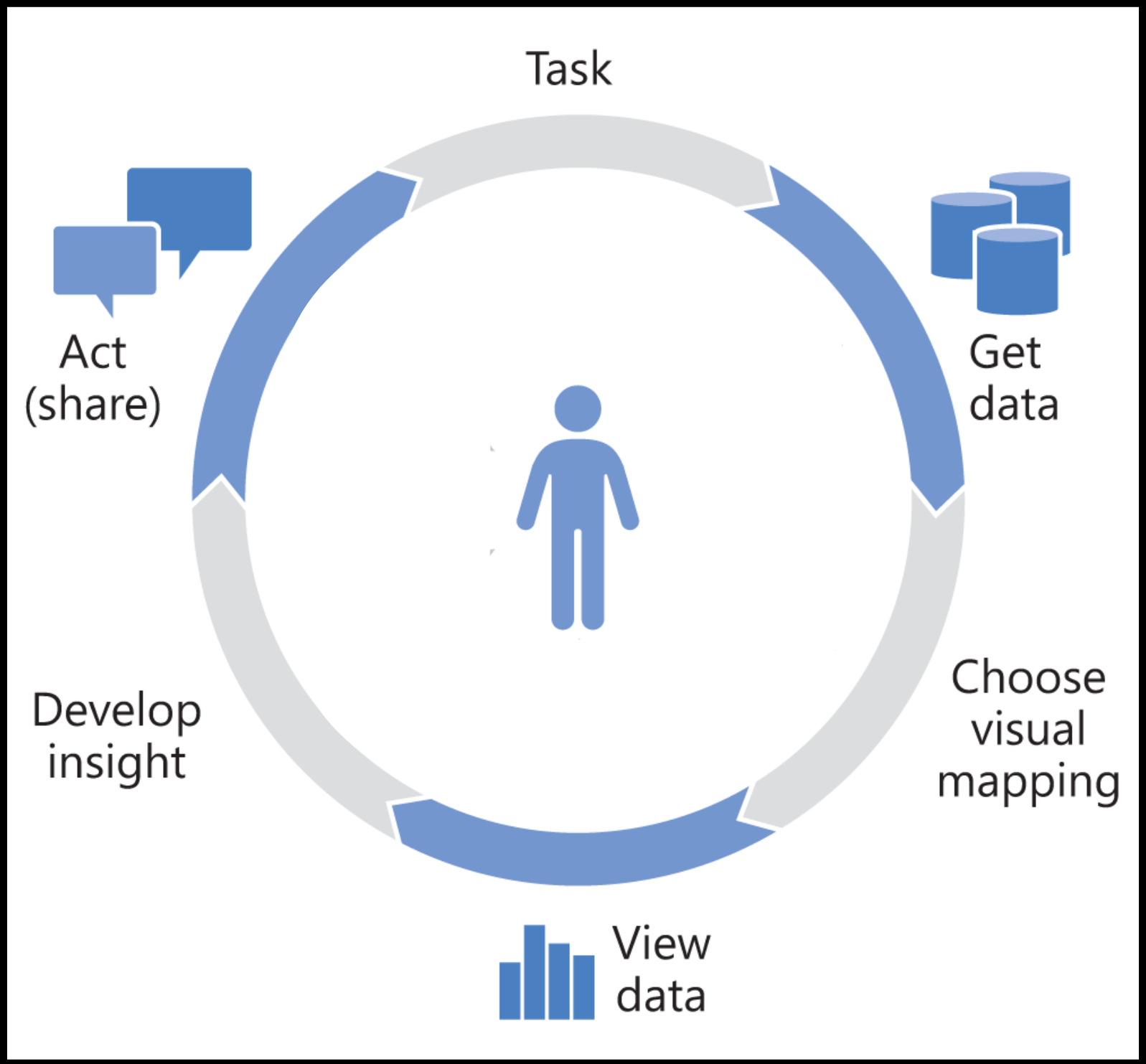
How many 9 s?

3	3	0	3	0	1	8	7	6	8	2	1	4	0	3	8	3	7	7	2	0	5	2	3	2	7	0	2	0
7	1	4	6	0	2	1	3	2	7	6	0	2	5	6	3	2	5	7	6	3	3	0	2	0	3	0	7	2
8	7	5	7	2	8	3	8	7	7	8	2	0	7	7	5	2	3	1	1	5	6	3	8	4	7	8	2	0
0	5	0	5	1	6	1	7	5	6	8	0	4	4	6	7	4	7	1	4	0	0	8	4	4	3	0	3	2
2	4	3	1	3	5	4	**9**	5	0	7	6	0	7	4	3	1	8	2	7	3	4	6	0	2	4	8	2	3
8	6	2	2	6	5	4	6	7	0	7	6	0	0	3	**9**	0	2	4	7	1	7	2	3	3	5	8	7	0
0	8	4	5	1	3	1	7	6	4	5	4	1	2	4	5	3	3	5	4	**9**	6	7	7	6	3	4	2	5
4	7	7	0	2	2	0	1	1	7	7	7	0	2	6	6	4	7	5	8	6	1	4	3	7	8	5	4	6
4	3	6	6	4	6	6	2	8	4	8	5	3	7	8	8	1	3	8	5	4	5	7	4	0	3	2	8	4
5	5	0	3	5	3	5	3	8	3	2	3	8	2	3	1	6	2	7	2	4	6	3	6	4	4	3	2	5
4	4	0	2	1	7	2	4	4	7	4	1	**9**	2	4	5	2	5	0	4	0	0	5	3	6	3	3	6	7
7	4	6	6	8	7	5	7	**9**	2	0	2	8	8	8	8	3	2	4	2	6	4	0	4	6	3	7	2	1
0	1	7	1	5	**9**	1	4	2	8	7	3	7	1	4	5	1	8	7	8	0	5	1	7	0	5	8	8	1
2	8	5	2	1	2	8	7	7	6	2	5	6	2	6	4	1	5	1	6	1	2	1	1	0	5	6	4	0
2	1	1	7	7	2	0	0	1	8	7	0	2	**9**	0	2	8	5	7	8	4	6	0	6	5	0	7	1	2
0	5	2	4	1	5	3	3	1	5	5	1	4	0	1	6	4	3	3	**9**	8	8	3	4	6	8	4	8	6
7	3	7	5	2	4	0	2	7	6	3	8	5	5	4	5	8	8	7	5	5	6	5	6	7	**9**	7	7	4
0	3	2	8	1	4	4	6	0	8	2	3	0	1	3	4	6	2	0	5	7	7	3	6	1	8	7	3	5
4	4	8	3	3	3	5	0	1	0	3	8	6	3	2	0	5	0	6	1	3	3	4	3	6	1	5	8	6
1	0	2	2	7	6	3	3	0	8	8	0	3	1	8	8	1	2	1	7	5	2	**9**	3	5	8	3	2	5

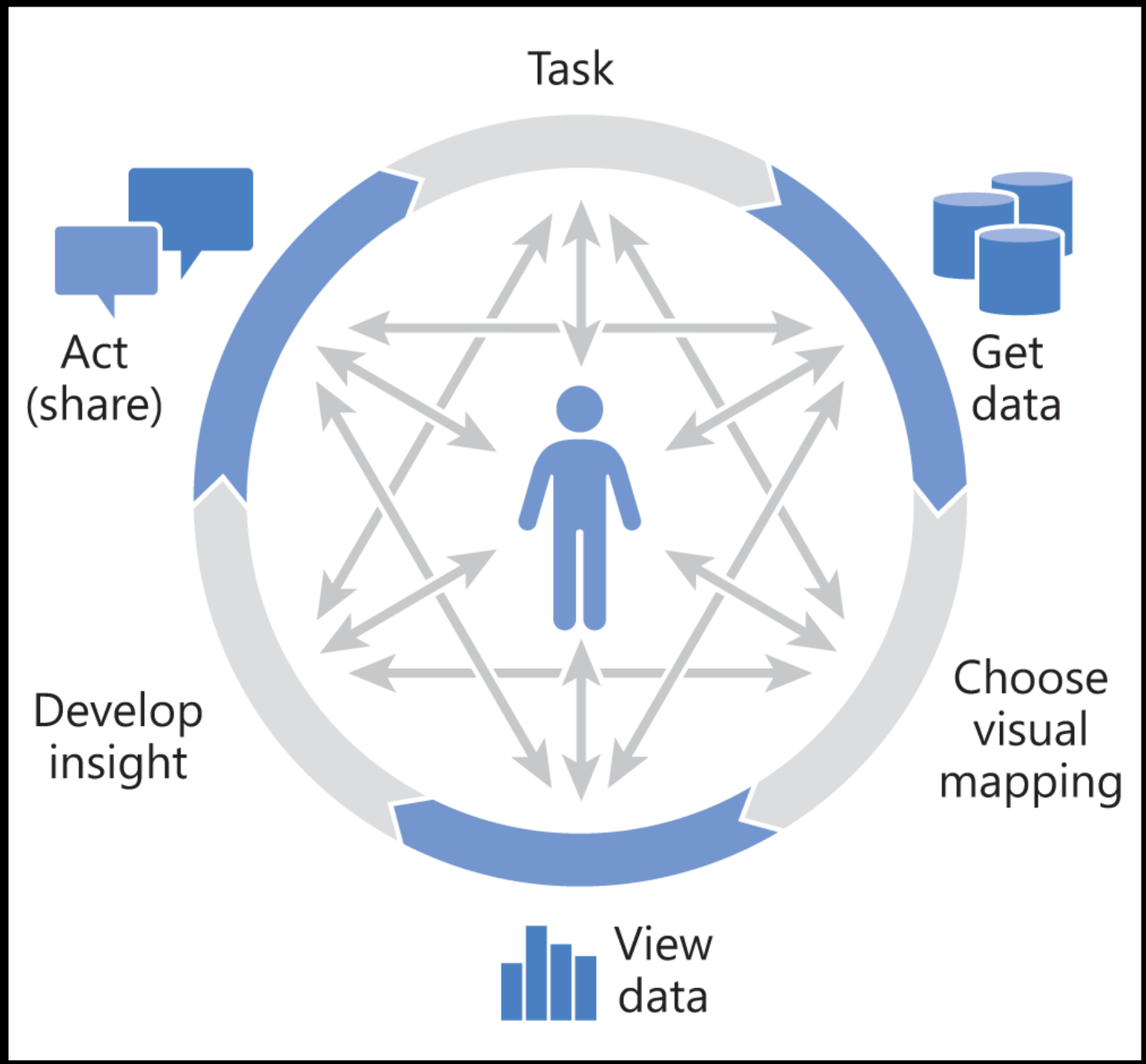
Visualizing Data

		Central	East	South	West
Coffee	Amaretto	$5,104	$1,010		($1,224)
	Colombian	$8,525	$27,256	$8,767	$11,256
	Decaf Irish Cream	$9,635	$2,726	$2,935	($1,307)
	Total	$23,264	$30,992	$11,702	$8,725
Espresso	Caffe Latte			$3,873	$7,502
	Caffe Mocha	$14,642	($6,232)	$5,202	$4,066
	Decaf Espresso	$8,859	$2,411	$5,930	$12,302
	Regular Espresso		$10,065		
	Total	$23,501	$6,244	$15,005	$23,870
Herbal Tea	Chamomile	$14,435	$764	$3,178	$8,854
	Lemon	$6,253	$7,902	$2,593	$13,121
	Mint	$4,069	($2,243)		$4,328
	Total	$24,757	$6,423	$5,771	$26,303
Tea	Darjeeling	$10,769	$6,500		$11,784
	Earl Grey	$10,334	$3,404		$10,426
	Green Tea	$1,227	$5,654		($7,112)
	Total	$22,330	$15,558		$15,098

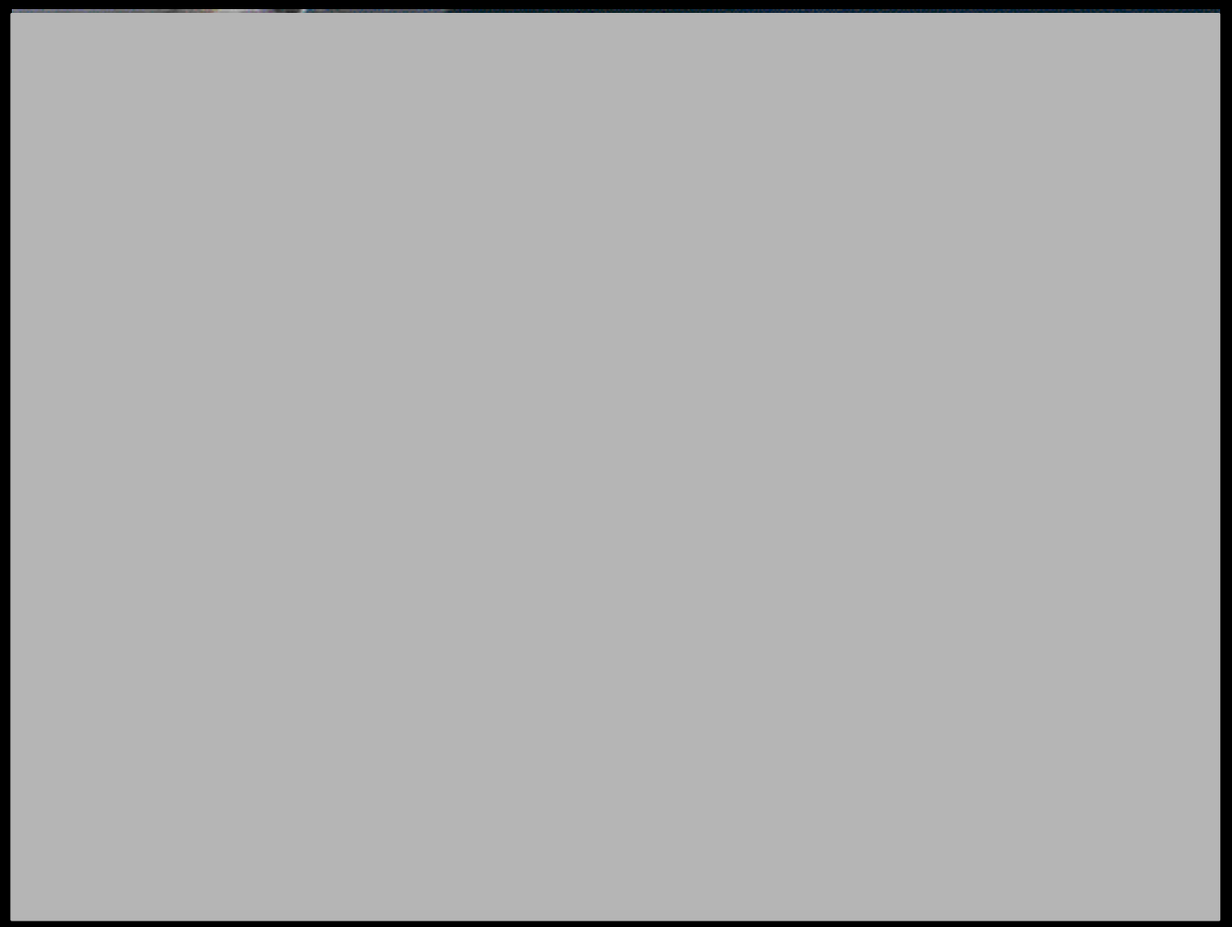
The Cycle of Visual Analysis



The Cycle of Visual Analysis




TL 700828

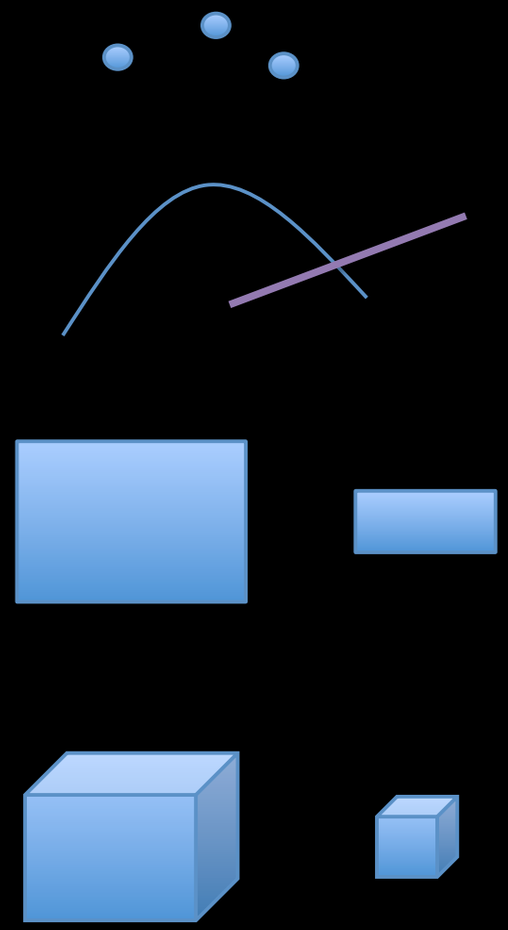


"The basic intent of data analysis is simply stated:
to seek through a body of data ***interesting***
relationships and information to exhibit the results
in such a way as to make them ***recognizable*** to the
data analyzer and recordable for posterity"

- J. W. Tukey, M. B. Wilk
Data Analysis & Statistics: Techniques & Approaches

Semiology of Graphics

Graphical Marks



Retinal Variables

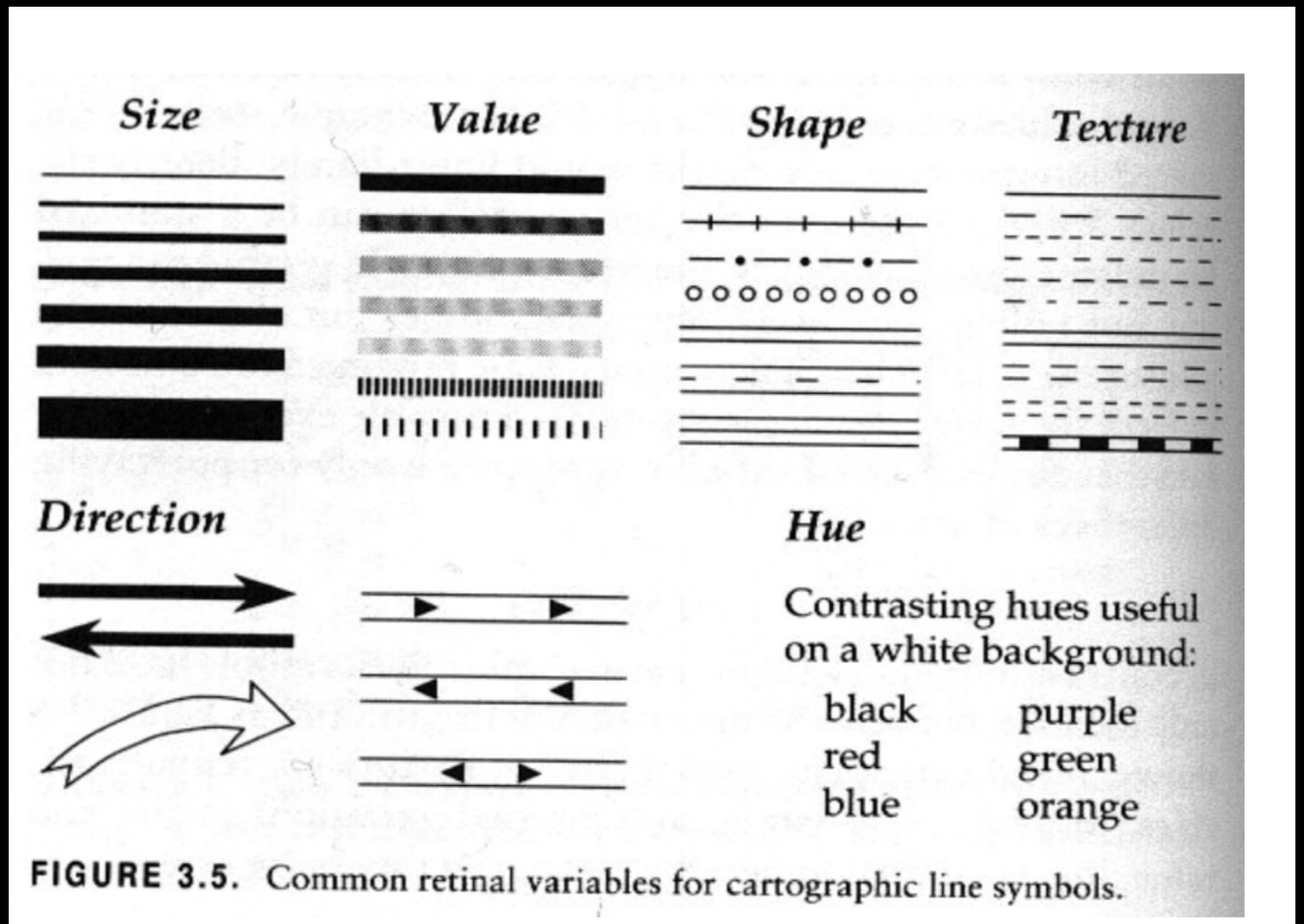


FIGURE 3.5. Common retinal variables for cartographic line symbols.

Jacques Bertin. Semiologie Graphique: Les diagrammes – Les reseaux – Les cartes. reissued by Editions de l'Ecole des Hautes Etudes en Sciences, 1999, 1967.

Visual Channels

- Position
- Size
- Orientation

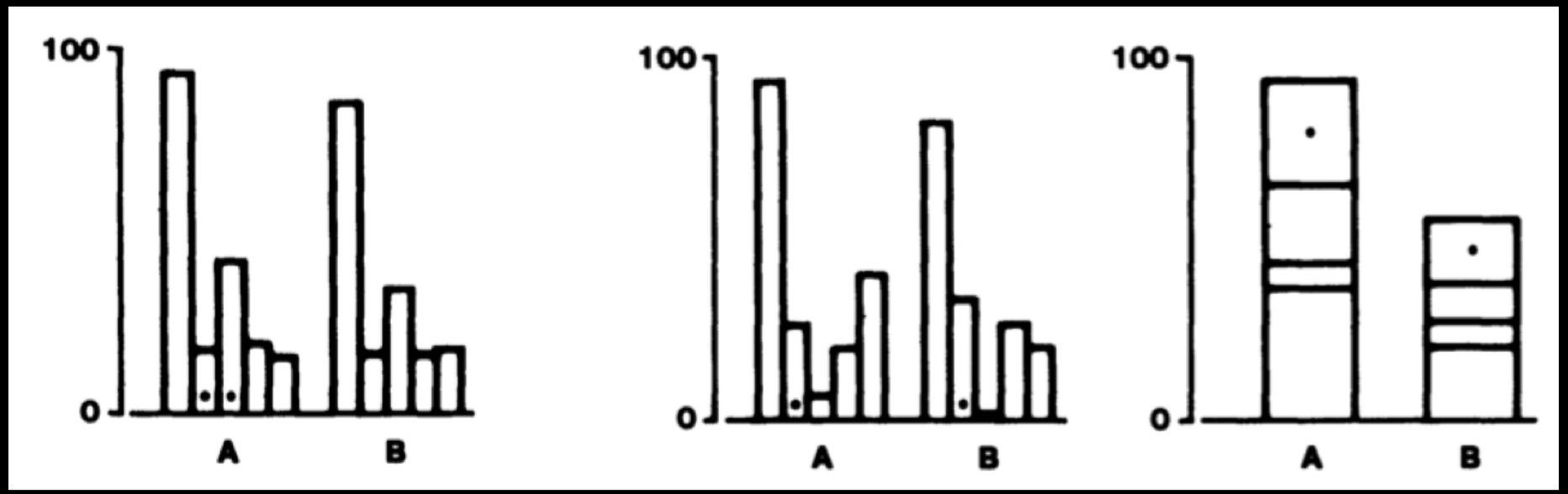
HOW MUCH?

- Shape
- Color (hue, saturation, lightness)
- Texture

WHAT?

Visual Encoding Principles

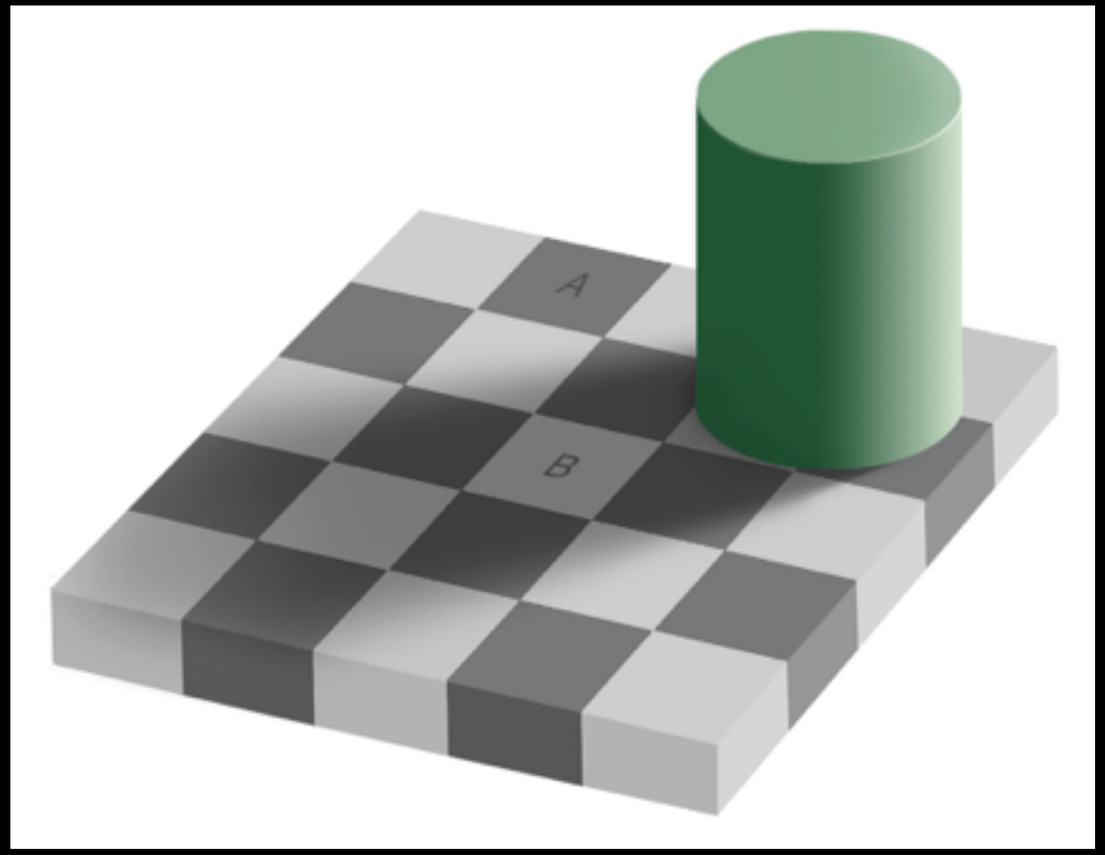
- *Weber's Law*
  The human perceptual system is fundamentally based on relative judgements, not absolute ones

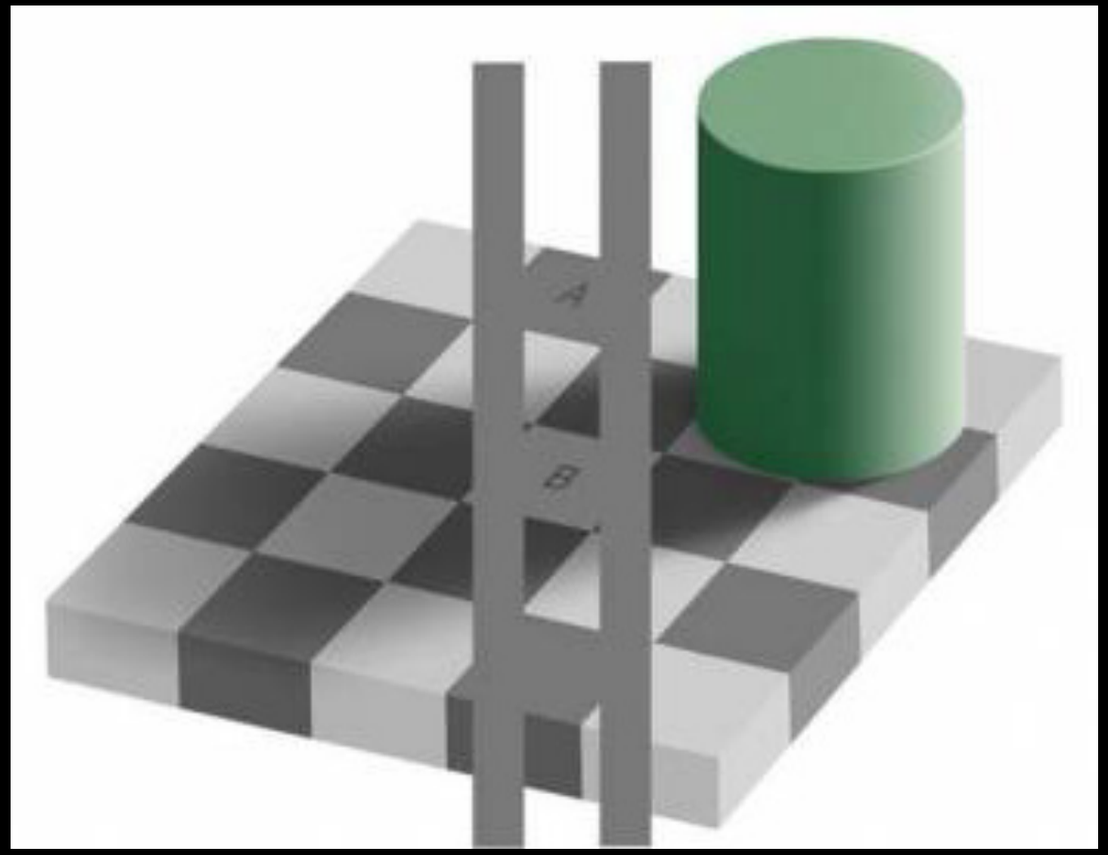


WS Cleveland, R McGill. *Graphical perception: Theory, experimentation, and application to the development of graphical methods.* Journal of the American statistical association 79 (387), 531-554

Visual Encoding Principles

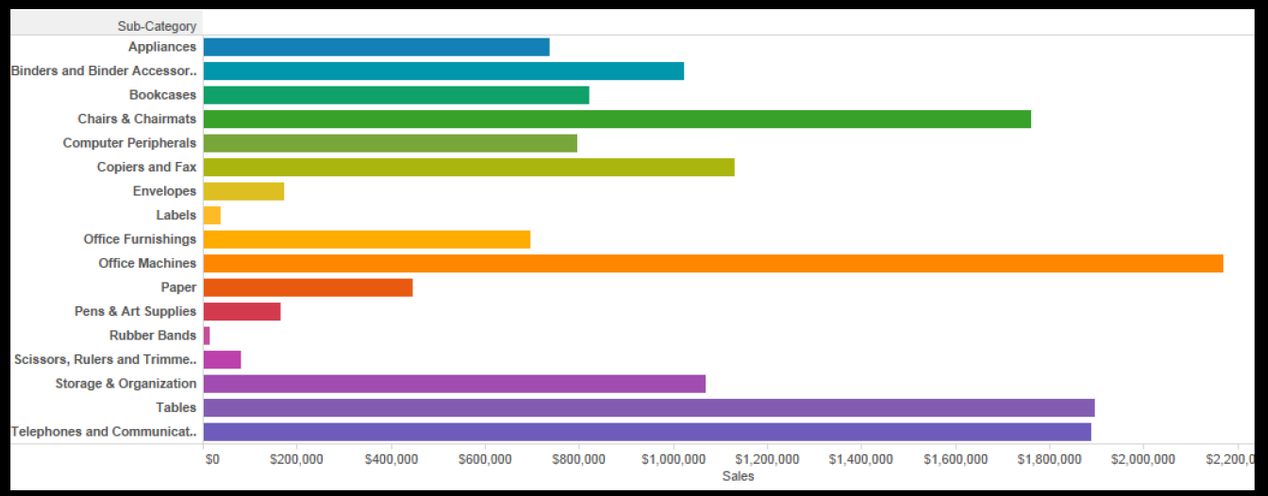
- *Color Constancy*
  An example of subjective constancy - color perception remains relatively constant under varying illumination conditions





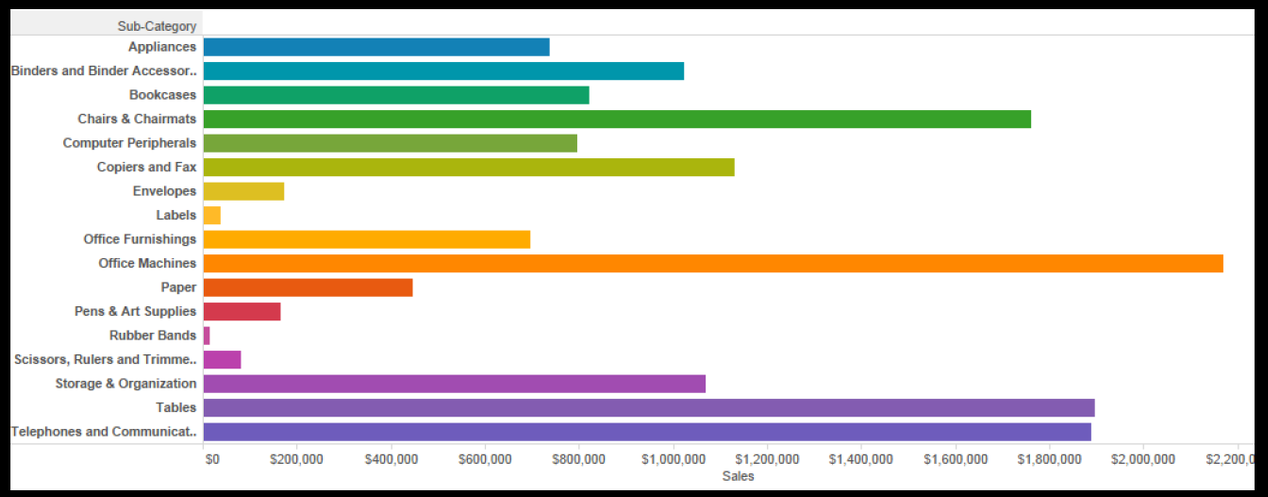
Evaluating Visual Channels

- *Expressiveness*
  The visual encoding should express all, and only, the information in the data

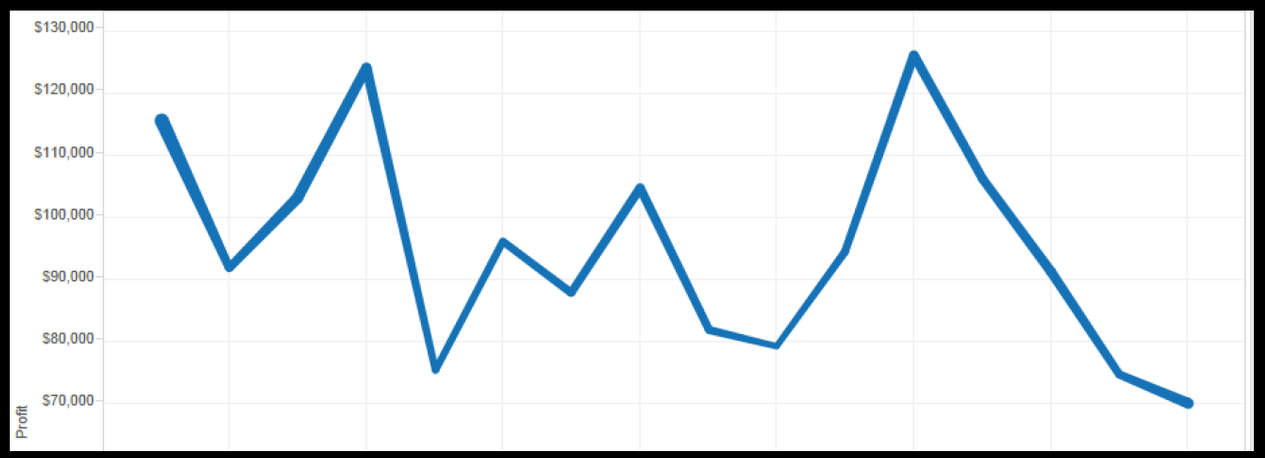


Evaluating Visual Channels

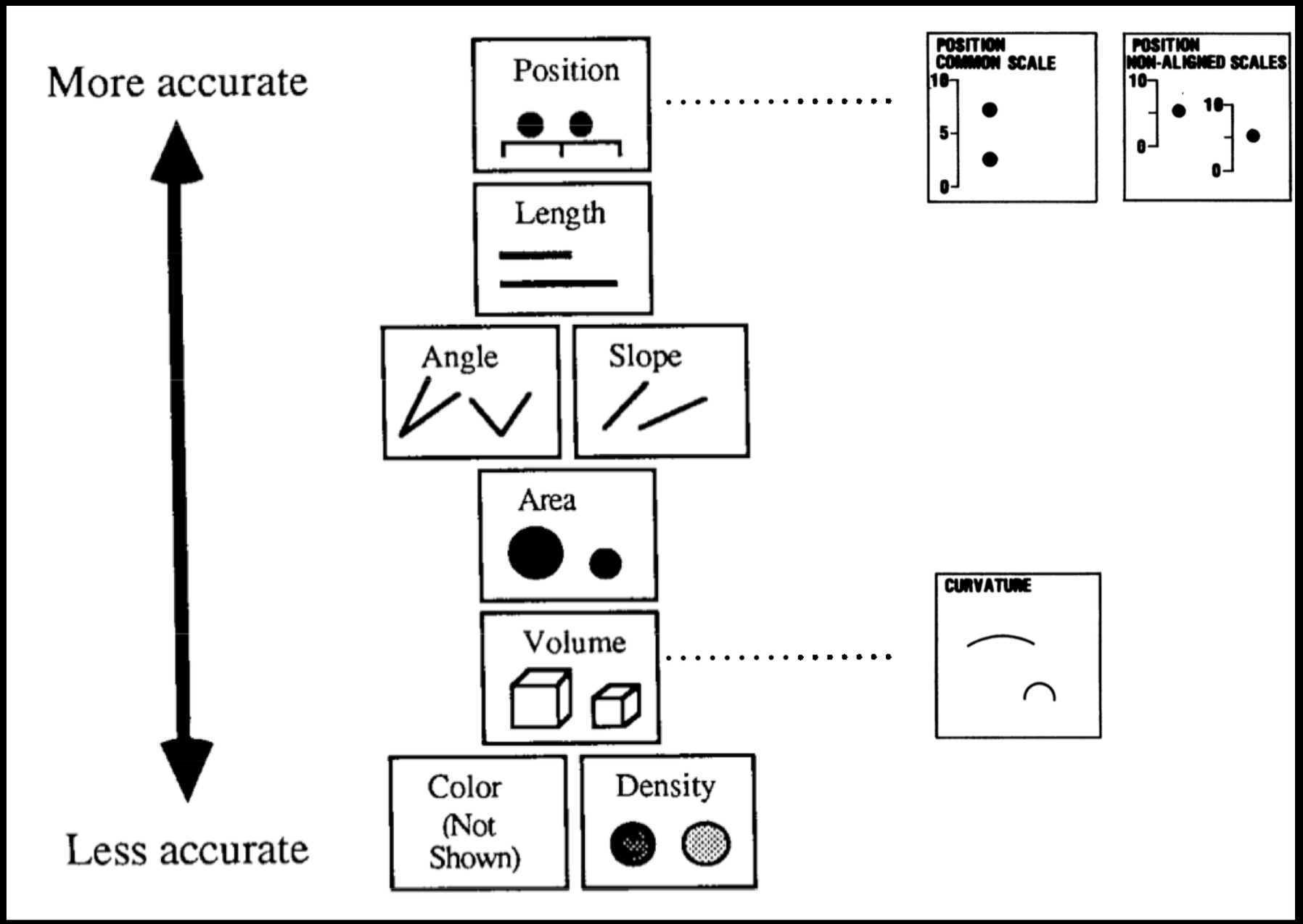
- *Expressiveness*
  The visual encoding should express all, and only, the information in the data



- *Effectiveness*
  The importance of the attribute should match the salience of the channel



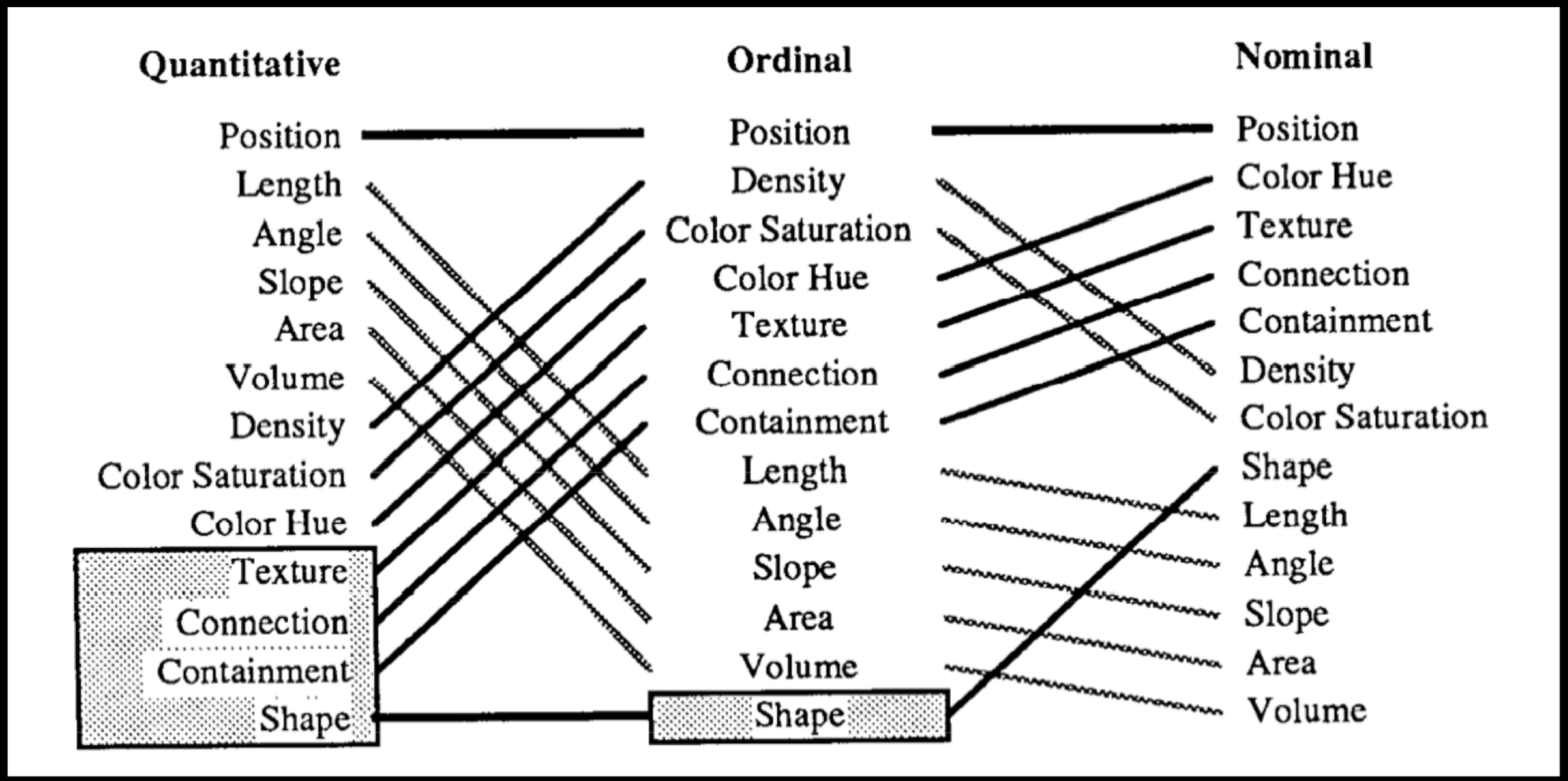
Graphical Perception



William S. Cleveland and Robert McGill. “Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods.” Journal of the American Statistical Association 79:387 (1984), 531–554.

Design Guidelines

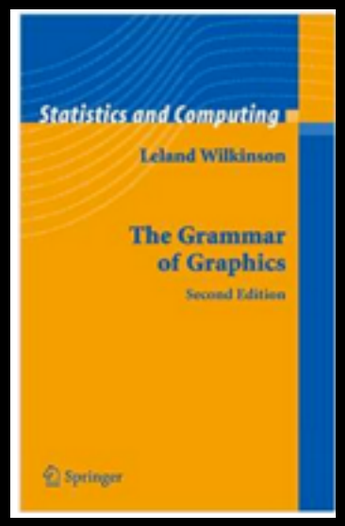
Mapping Data Types to Visual Channels

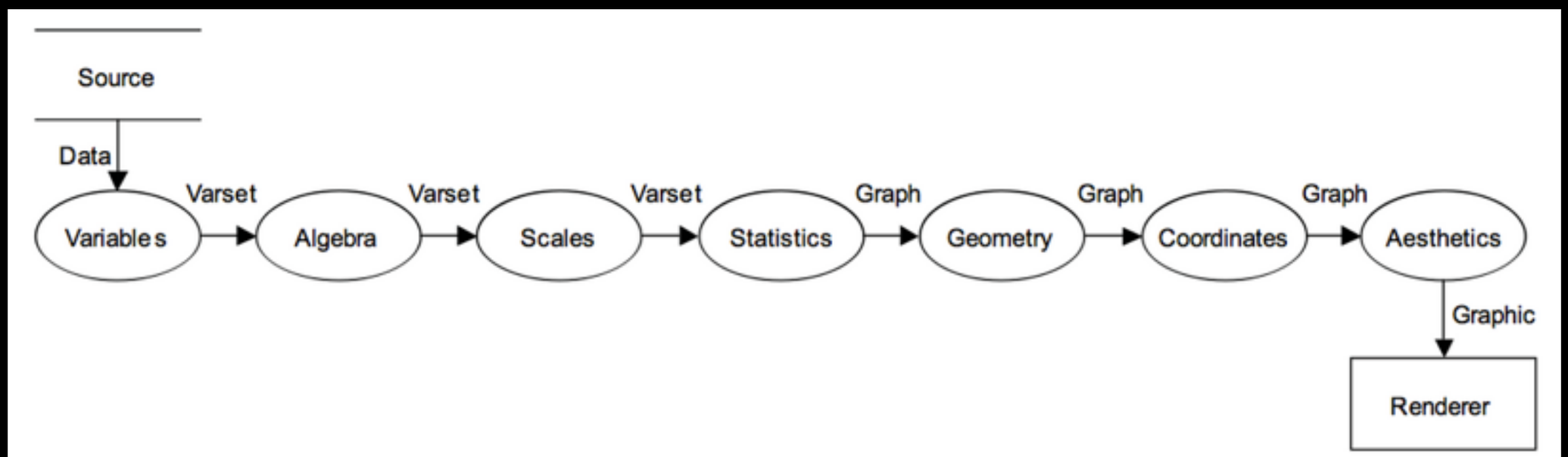


Jock Mackinlay. “Automating the Design of Graphical Presentations of Relational Information.” ACM Trans. on Graphics (TOG) 5:2 (1986), 110–141.

ggplot2

http://ggplot2.org/





Leland Wilkinson's original pipeline diagram (Grammar of Graphics p.24).

Reference Books

- Stephen Few
  Show Me The Numbers. 2004
- Edward Tufte
  The Visual Display of Quantitative Information. 1983
- Colin Ware
  Information Visualization, Perception for Design. 2004
- William Cleveland
  Visualizing Data. 1993
- John Tukey
  Exploratory Data Analysis. 1977
- Card, Mackinlay, Shneiderman: Chapter 1
  Readings in Information Visualization. 1999
- Jacques Bertin
  Semiology of Graphics. 1967