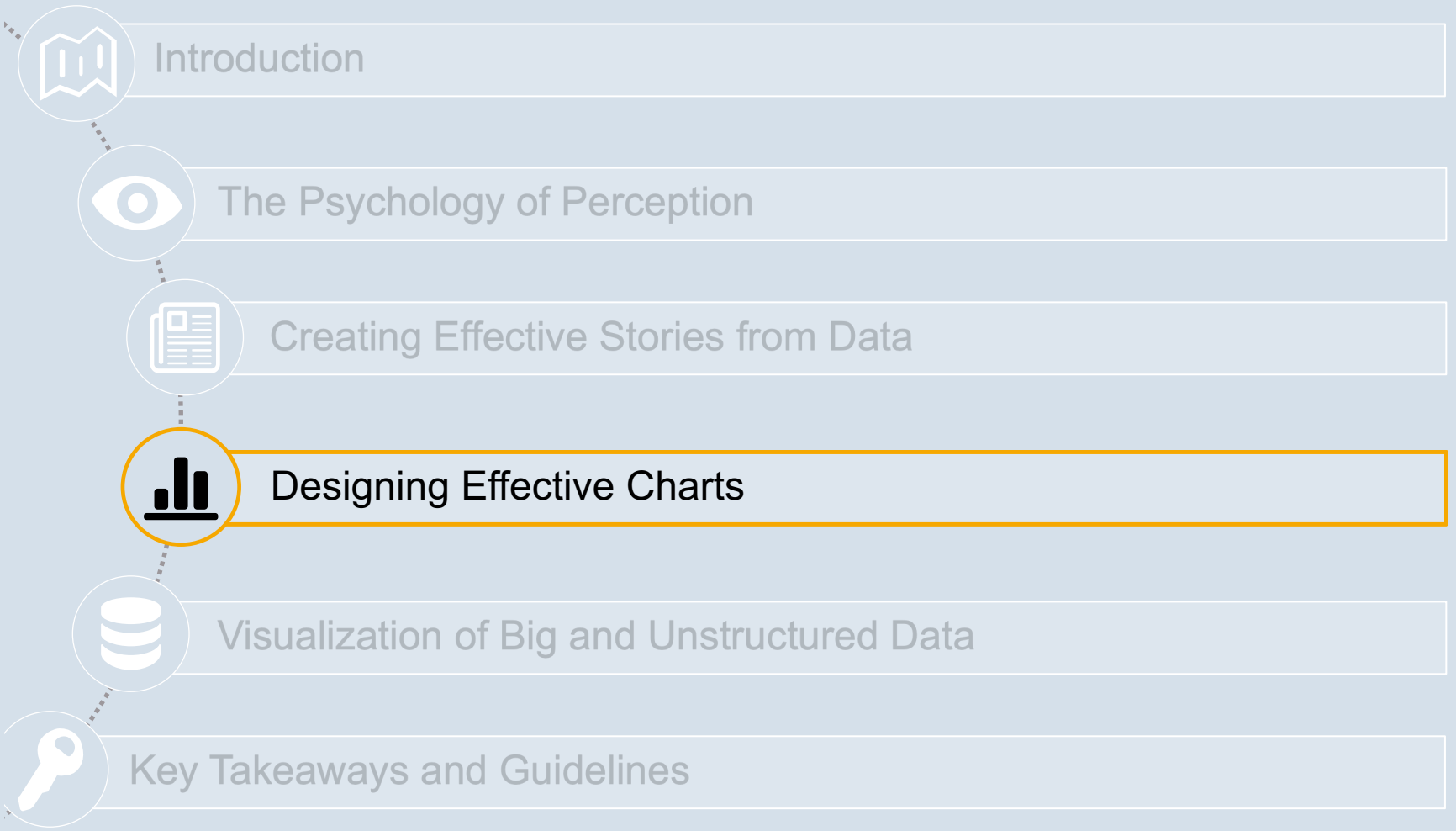
MANNHEIM BUSINESS SCHOOL

DATA VISUALIZATION

PROF. DR. FLORIAN STAHL

Overview – Data Visualization

MANNHEIM BUSINESS SCHOOL



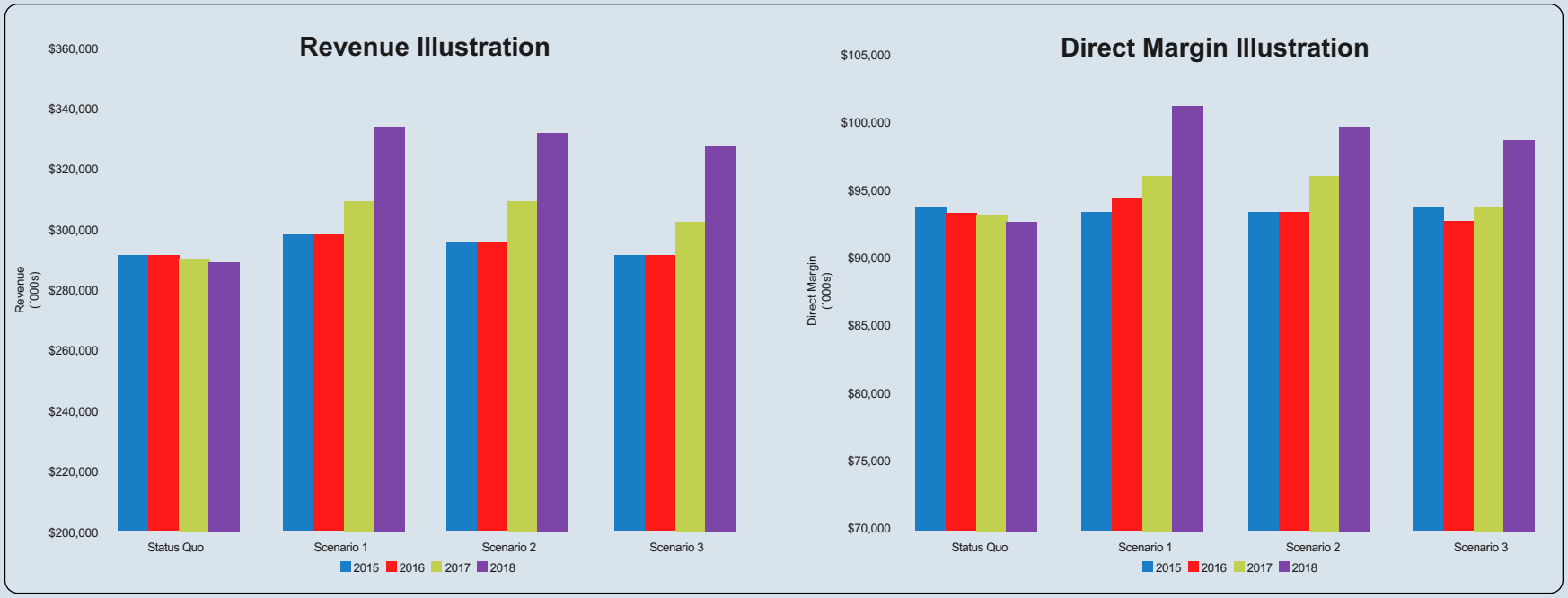
The two sides of visualization

Graphic displays are often very effective at communicating information

Graphic displays are also often not effective at communicating information

The effectiveness of your communication depends on how you visually create your charts!

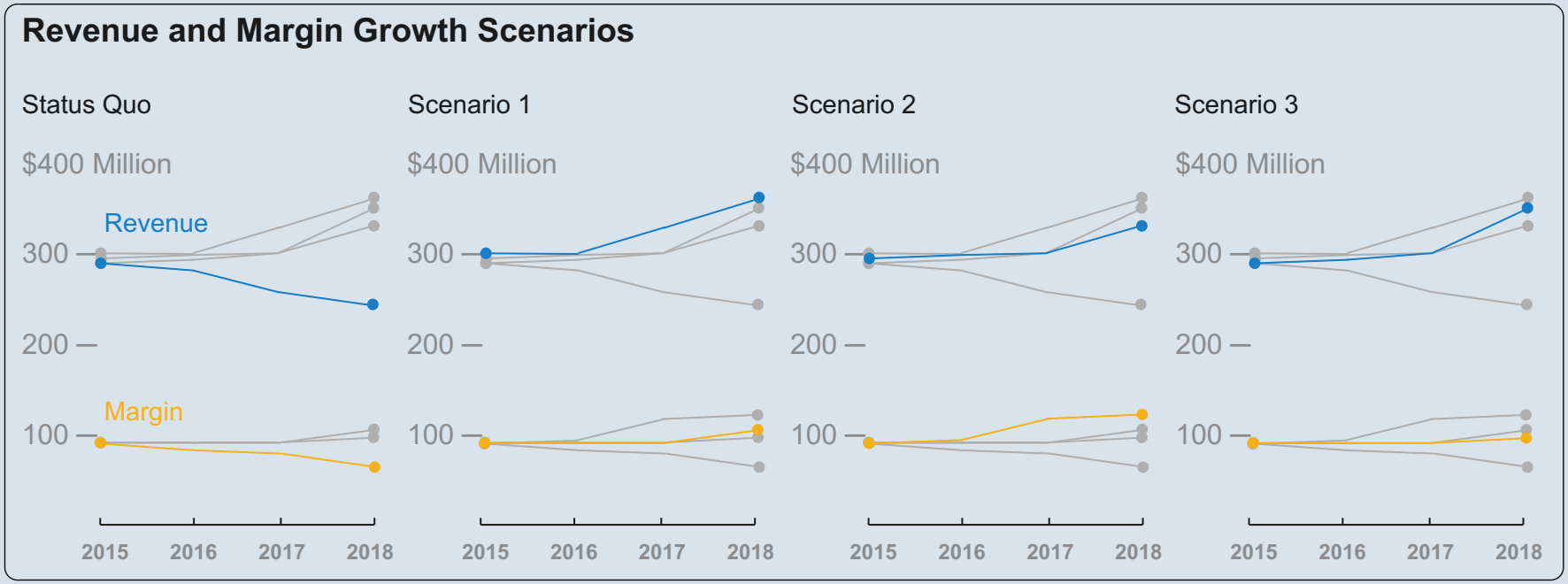
Example 1 – Graphics done wrong



- ✗ Excessive use of colors
- ✗ Wrong depiction for comparing scenarios
- ✗ Inconsistent Scaling

Example 1 – Graphics done right

MANNHEIM BUSINESS SCHOOL



- Intelligent use of colors
- Easy comparability of different scenarios
- Consistent scaling

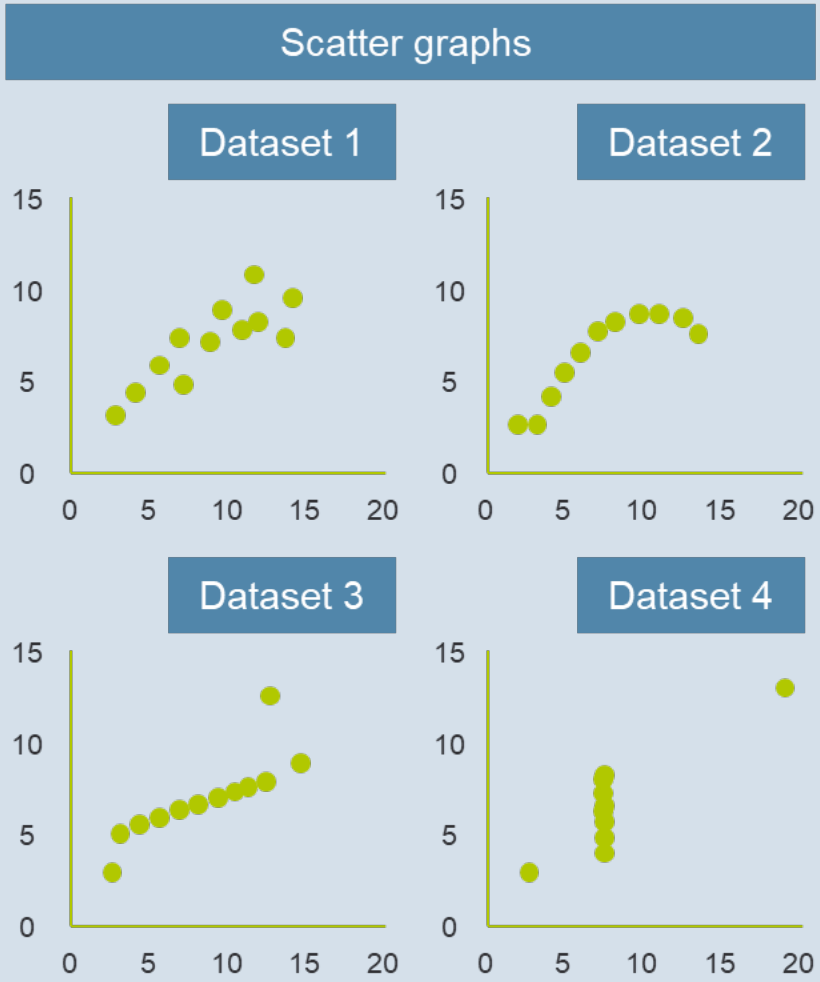
Example 2 – The problem with descriptive statistics

	Dataset 1		Dataset 2		Dataset 3		Dataset 4	
	X1	Y1	X2	Y2	X3	Y3	X4	Y4
observations	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	8,10	14	8,84	8	7,04
	6	7,24	6	6,13	6	6,08	8	5,25
	4	4,26	4	3,10	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
mean	9,00	7,50	9,00	7,50	9,00	7,50	9,00	7,50
variance	11,00	4,13	11,00	4,13	11,00	4,12	11,00	4,12
Cor-relation	0,816		0,816		0,816		0,817	

Anscombe's quartet

- Same means
- Same variances
- Same correlation
- **Same data?**

Example 2 – The help of visualization

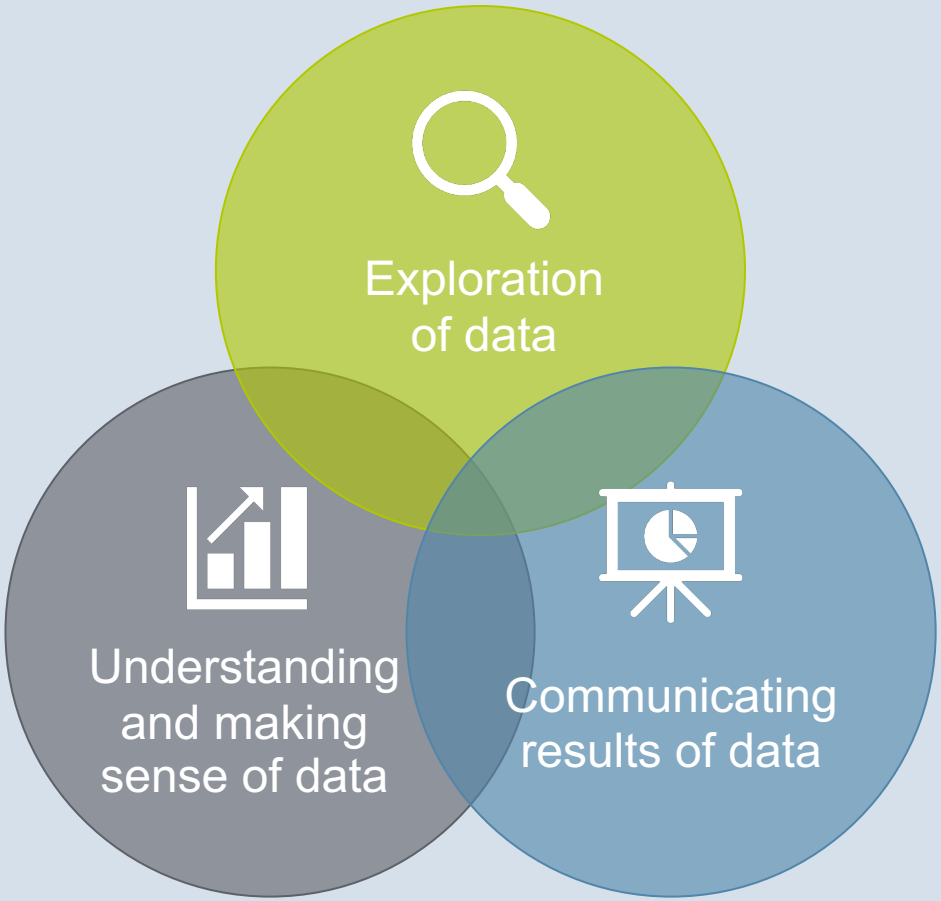


Anscombe's quartet

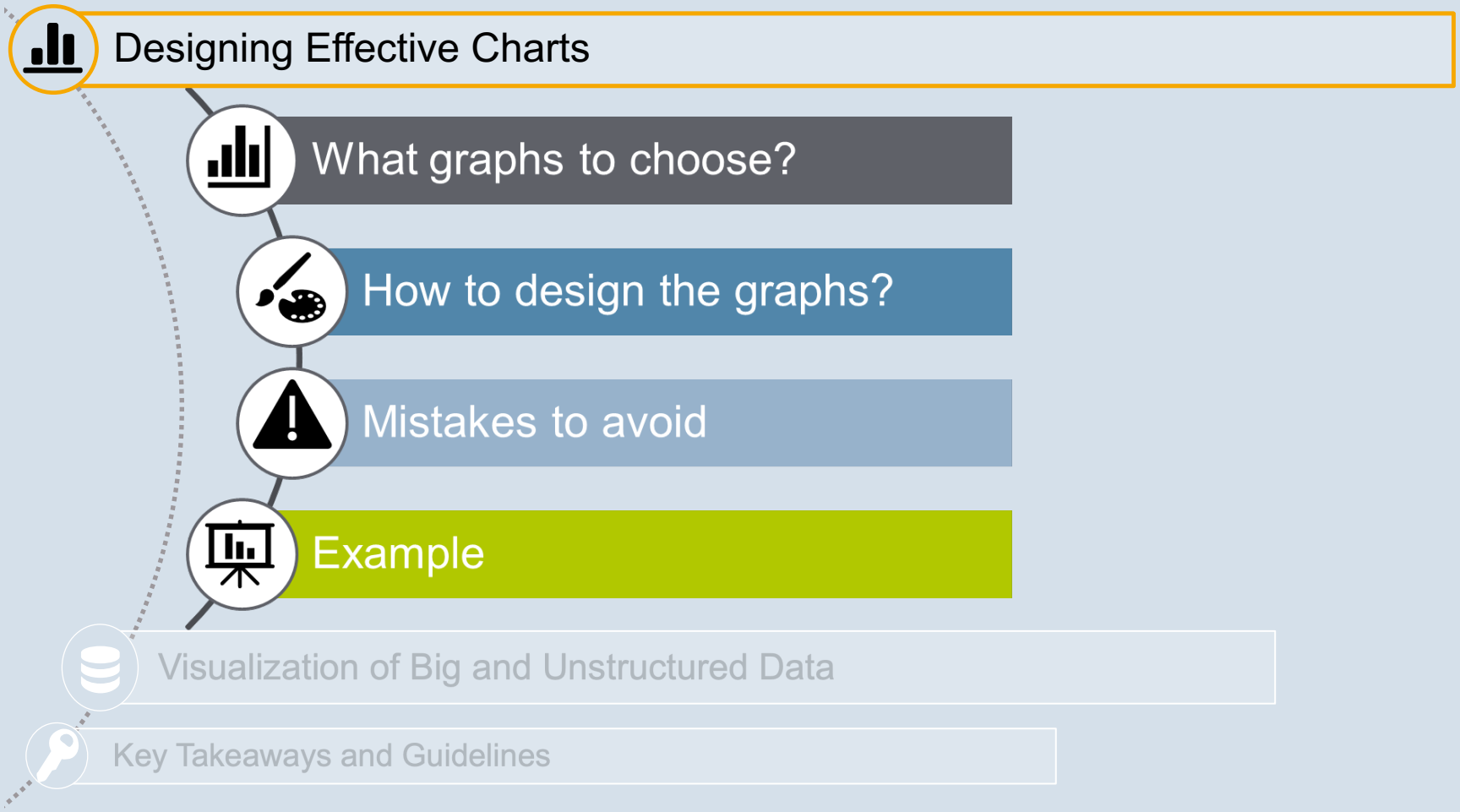
Visualization …

- … tells four different stories
- … helps in understanding the data
- …helps identifying patterns, trends and outliers

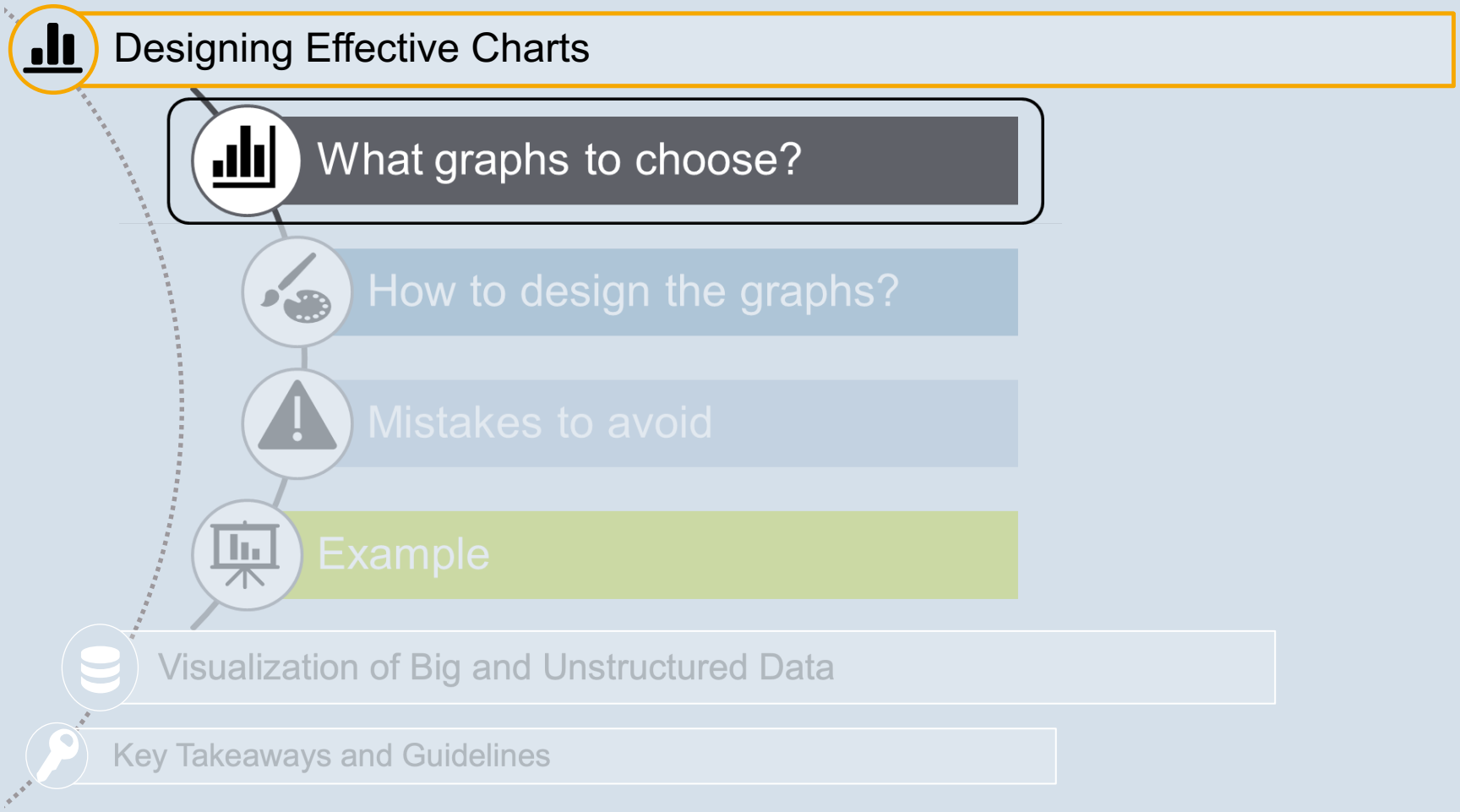
Three objectives of visualizing data



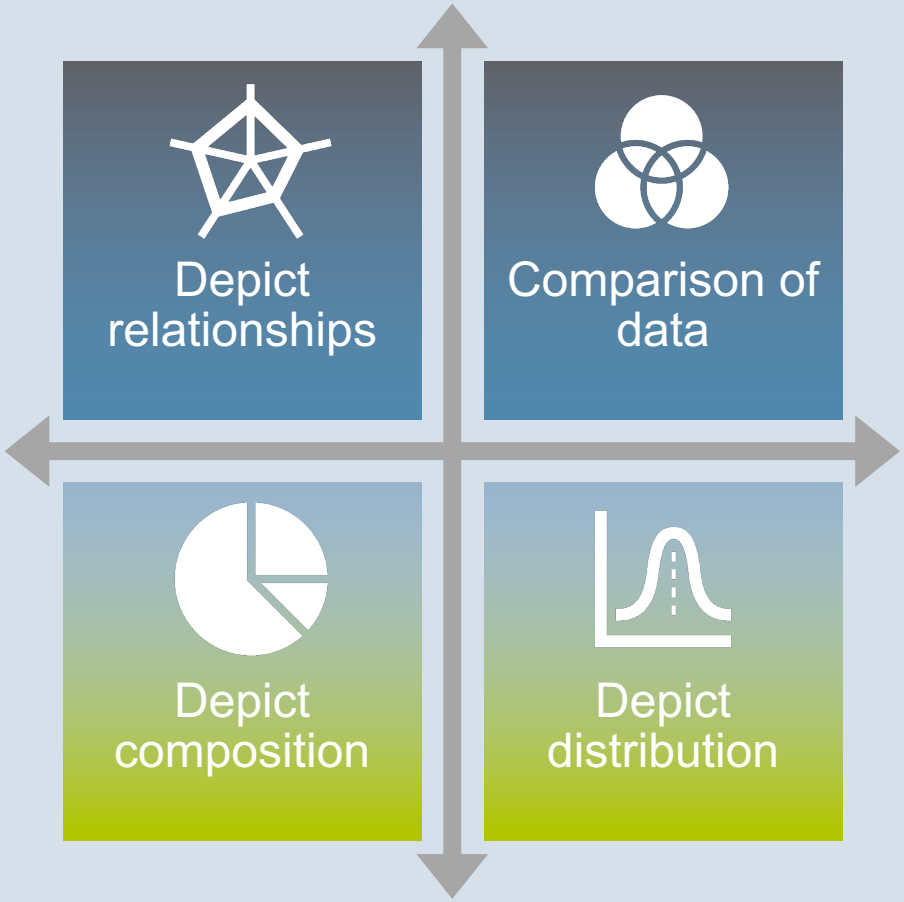
Data Literacy & Data Intuition



Data Literacy & Data Intuition



Choose graphs that underline what you want to tell

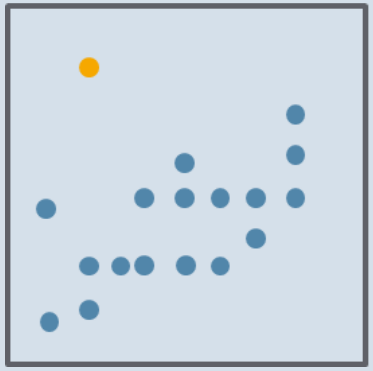


Relationship between data points

MANNHEIM BUSINESS SCHOOL

Depict relationships | Comparison of data | Depict composition | Depict distribution

Scatter chart

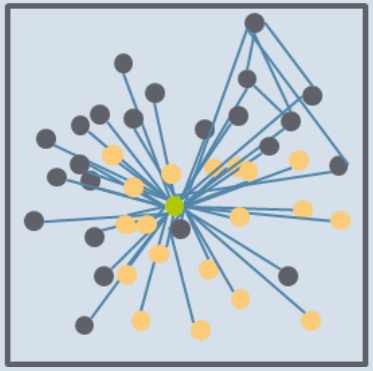


- Relationship between 2+ variables

Bubble chart

- Same as scatter chart
- Adds third variable through size of bubbles

Network chart



- Relationship between objects or individuals

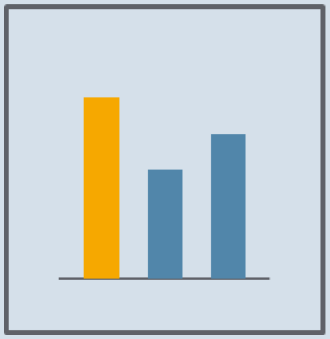
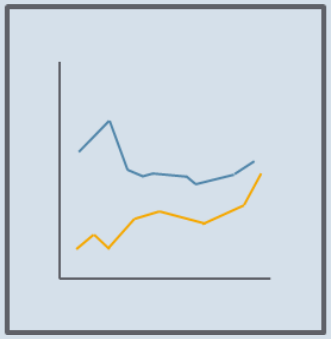
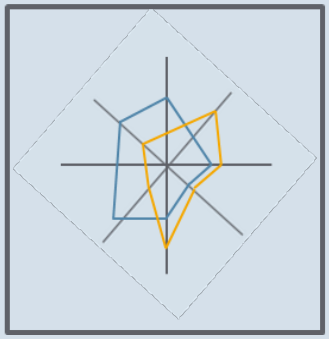
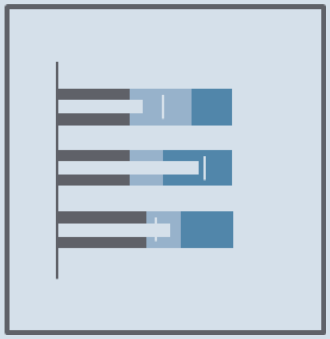
Circular network chart

- Same as network chart
- Includes position and importance of objects

Comparing data points

MANNHEIM BUSINESS SCHOOL

Depict relationships | **Comparison of data** | Depict composition | Depict distribution

Column chart	Bar chart	Line chart	Radar chart	Bullet chart
				
› Simple comparison	› Simple comparison › More subunits possible	› Comparison of data points over time	› Multivariate data in a two-dimensional chart of 3+ variables	› Similar to bar chart › Compares primary measure unit with other units

Composition of data

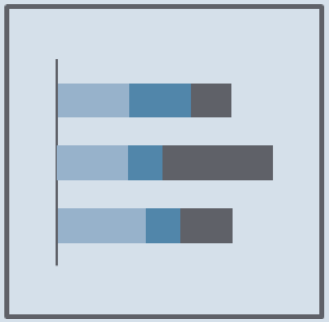
MANNHEIM BUSINESS SCHOOL

Depict relationships | Comparison of data | **Depict composition** | Depict distribution

Pie chart

- Shows simple composition
- Do not use with too many categories!

Stacked chart



- Stapled column or bar chart

Waterfall chart

- Useful for breakdown of a variable in components
- Can also be used for development depiction

Stacked area chart

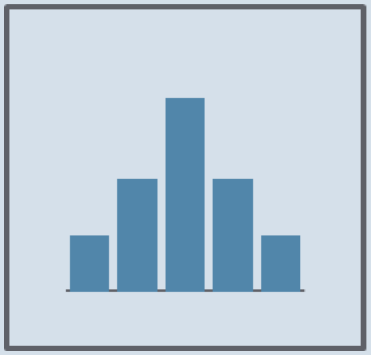
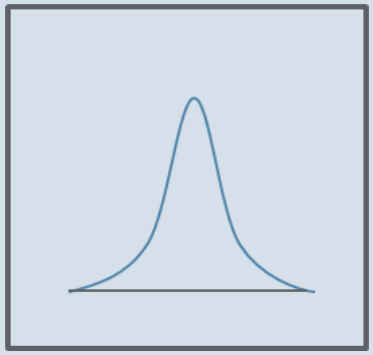
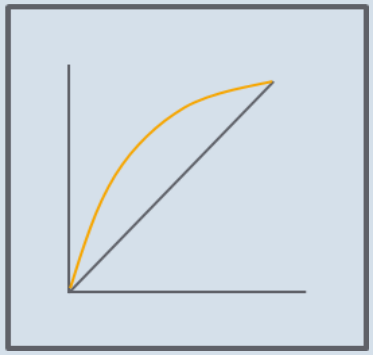
- Can show changes over time

Word/Tag chart

- To highlight frequency or importance of words in text analysis

Distribution of data

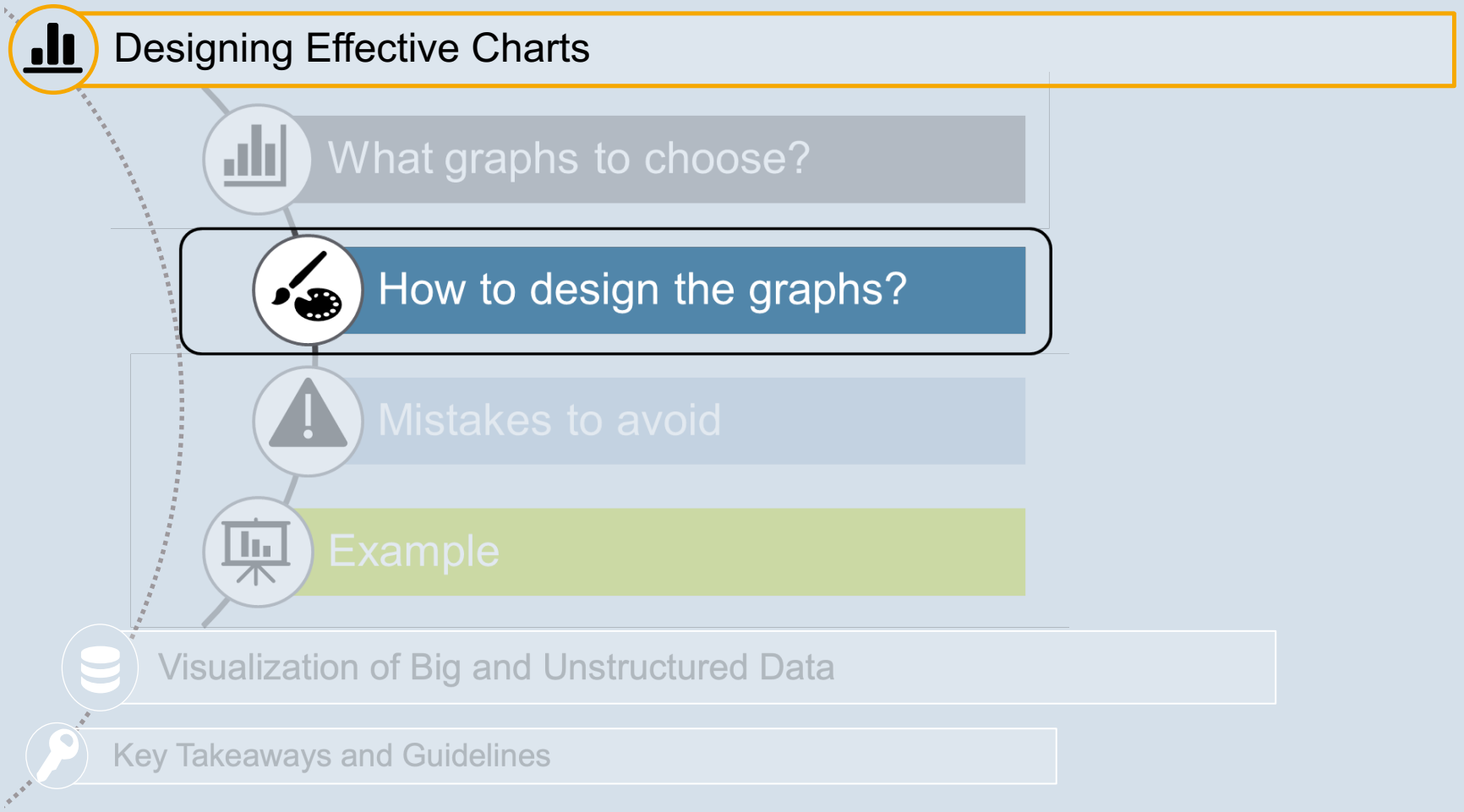
MANNHEIM BUSINESS SCHOOL

Depict relationships	Comparison of data	Depict composition	Depict distribution
Column histogram	Line chart	Double bar chart	Lift chart
			
› Useful when there are only a few categories	› Useful when there are a lot of categories	› To compare the distribution of two different variables	› To visualize how good a predictive model is

Overview of chart types

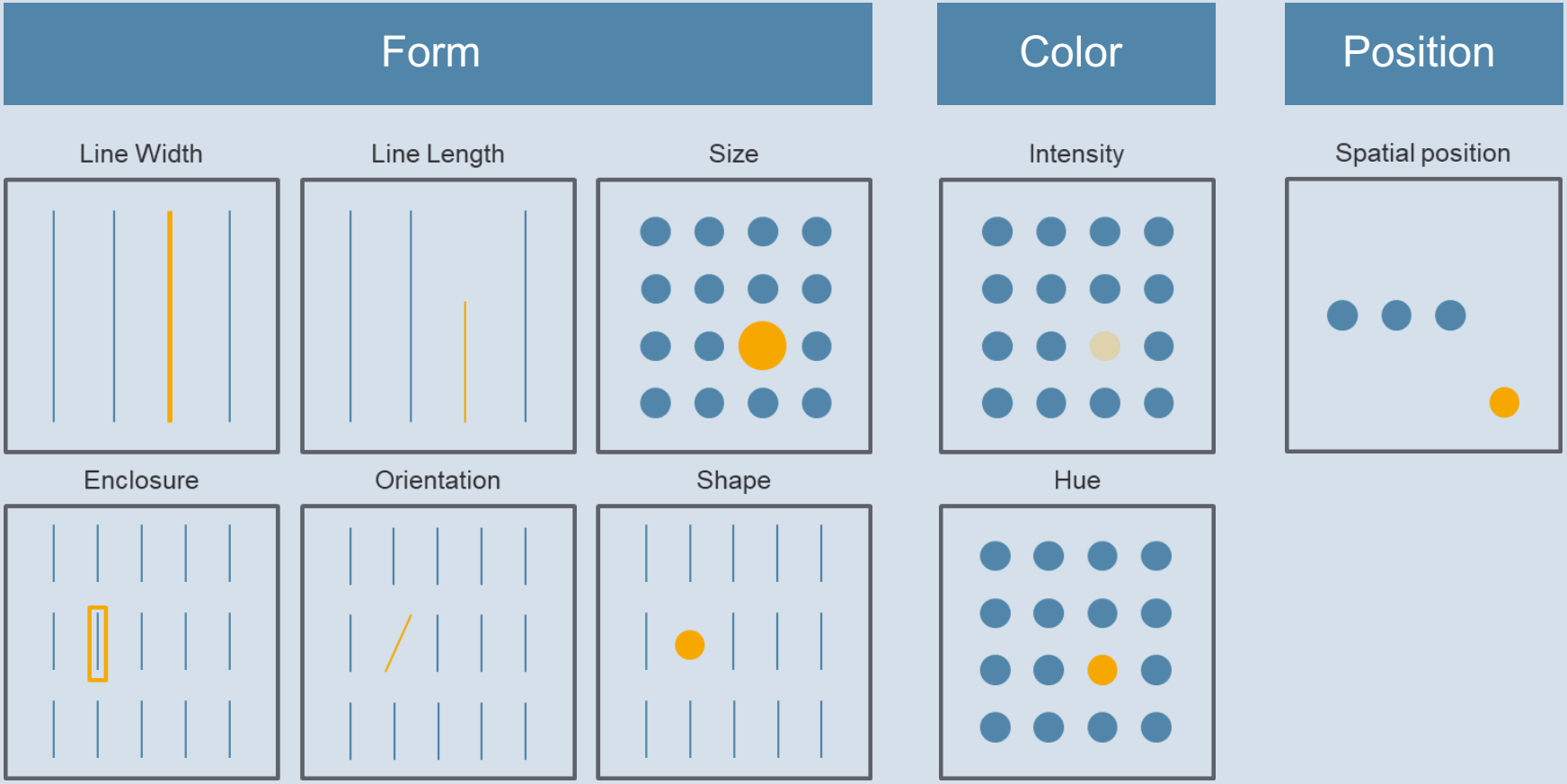
MANNHEIM BUSINESS SCHOOL

Data Literacy & Data Intuition

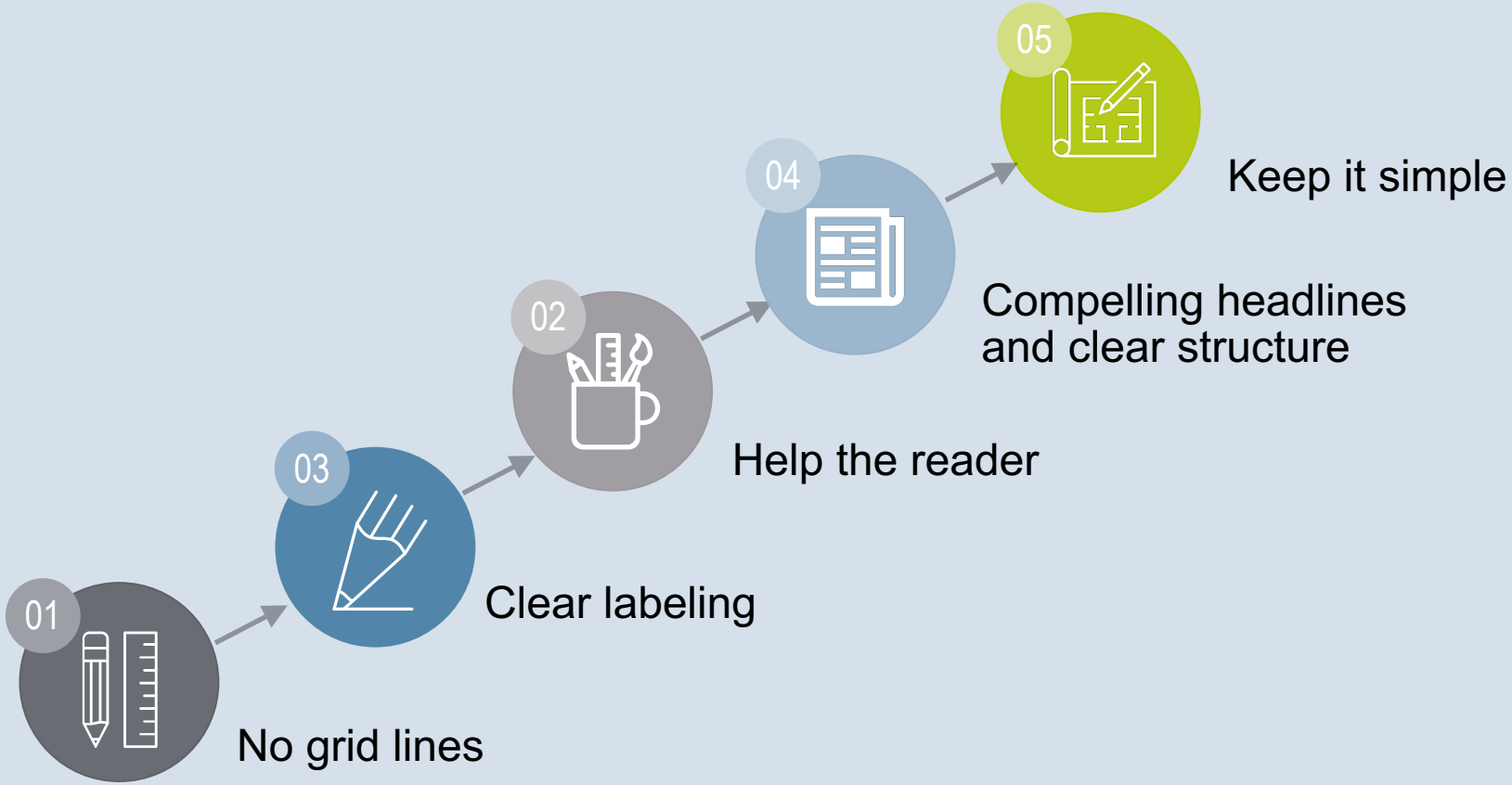


You have a variety of design options

MANNHEIM BUSINESS SCHOOL



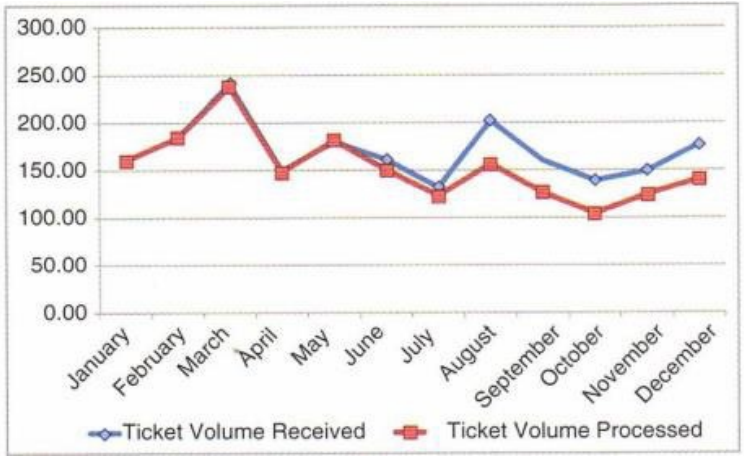
Guidelines and tips for graphical design



Tip 1 & 2 – No grid lines & clear labeling

MANNHEIM BUSINESS SCHOOL

- **Remove chart boarders**
- **Reduce** prominence or remove **grid lines**
- **Remove data markers**
- **Label data** directly and in **consistent** color
- Consider **orientation and spacing** of labels

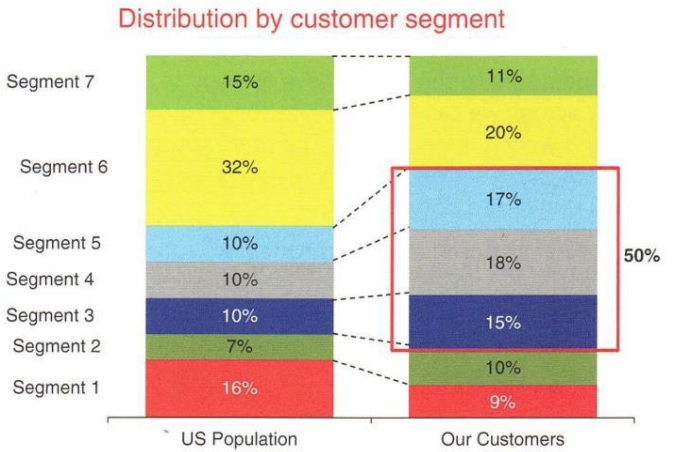


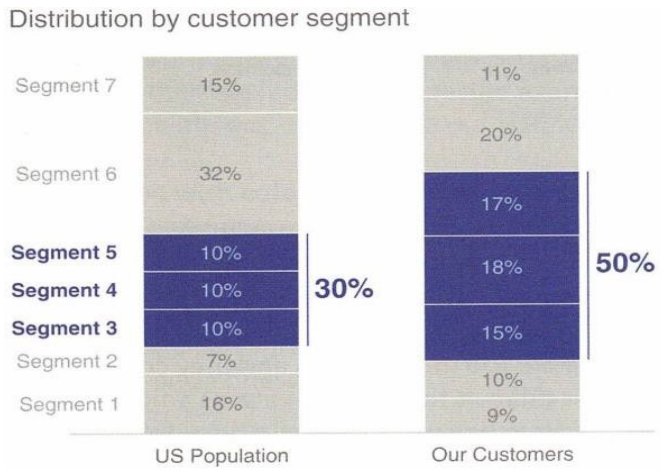
Knaflic (2015) Storytelling

Tip 3 – The Use of Color

MANNHEIM BUSINESS SCHOOL

- Keep it **simple**
- Remember the **meaning of colors**
- Do **not** work with **too bright colors**
- Use color to **draw attention**



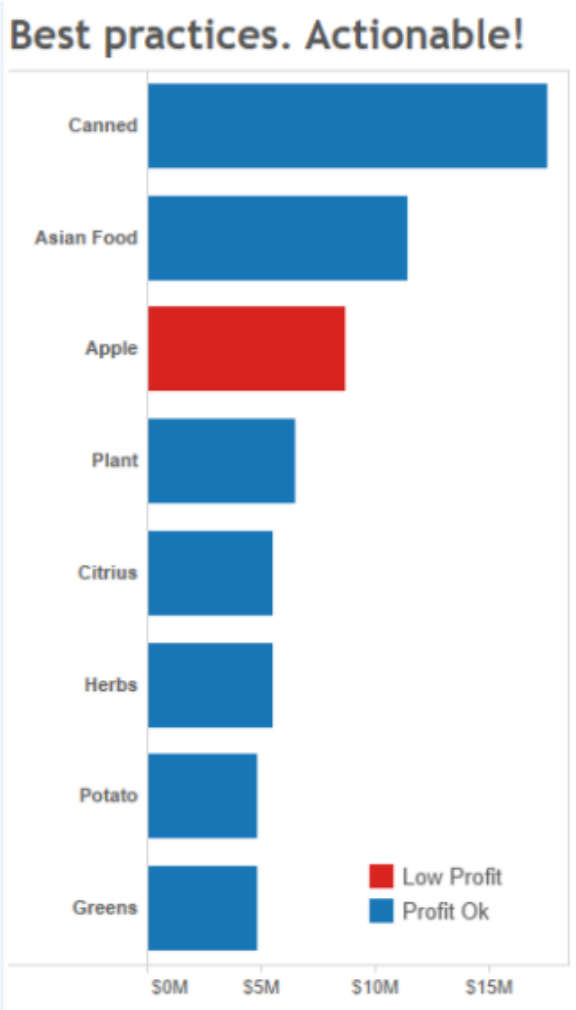


Knaflic (2015) Storytelling

Tip 3 – The Use of Color

MANNHEIM BUSINESS SCHOOL

- Use color to **emphasize points** you want to tell
- Use colors to **underline actions** you want the audience to take



Tip 3 – Number Formatting

MANNHEIM BUSINESS SCHOOL

- Keep it **consistent**
- Try to keep it **simple**
- Use **abbreviations** (K, M, …)
- Make numbers **comparable**
- Include **units** ($, …)
- Use **percentage signs** (%)

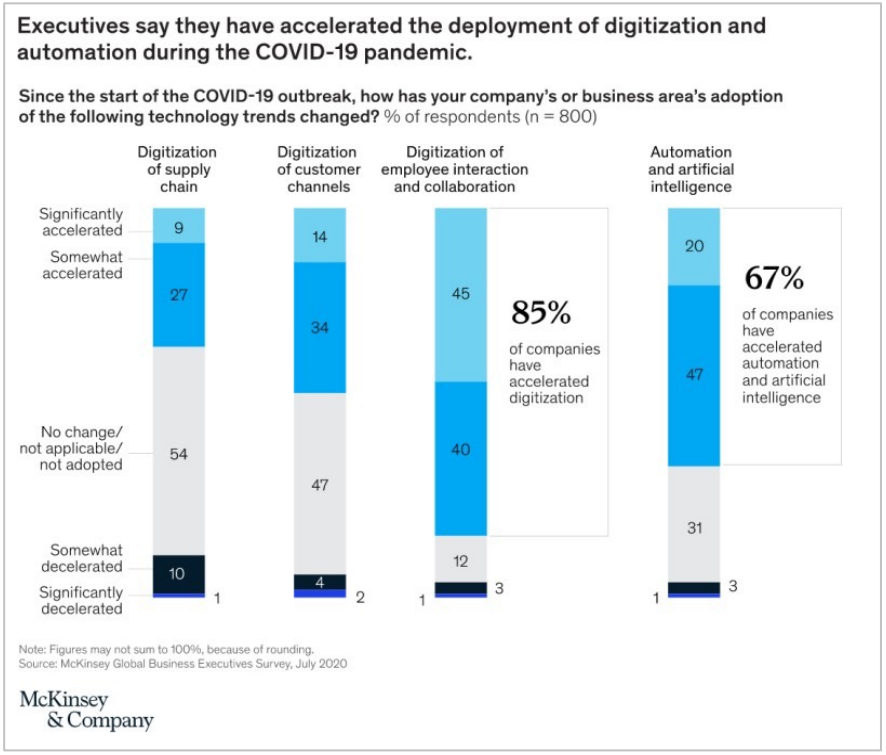
	Qty Sold	Sales	Pct GP
Abc Corp	1497019.00	86509343.13	0.54
Intl Ops	11440509.00	51235507.73	0.09
West Co	4198314.00	36319928.23	0.51
Xyz Inc	5828632.00	31876796.84	0.57

	Qty Sold	Sales	Pct GP
Abc Corp	1,497K	$86.5M	54.3%
Intl Ops	11,441K	$51.2M	9.1%
West Co	4,198K	$36.3M	51.0%
Xyz Inc	5,829K	$31.9M	57.4%

Tip 4 – Clear structure & compelling headline

MANNHEIM BUSINESS SCHOOL

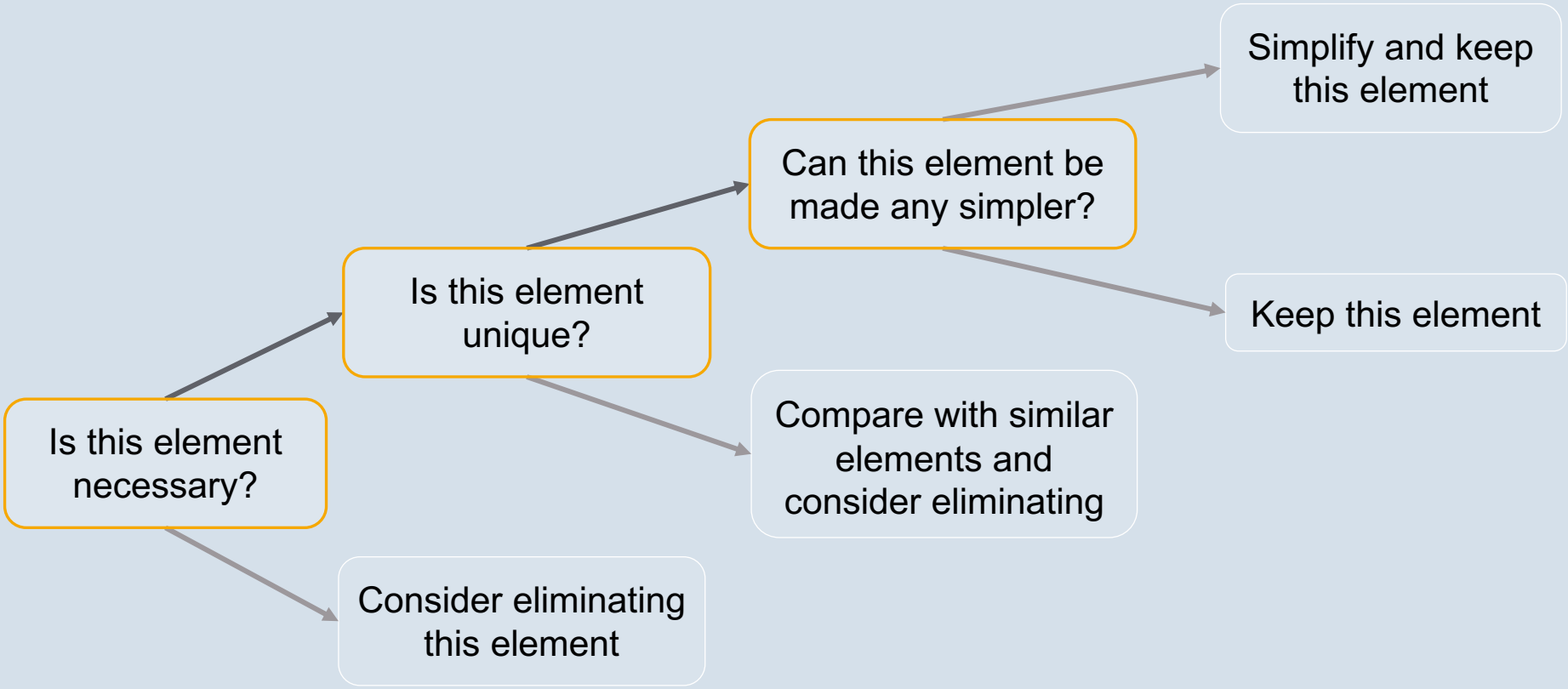
- Focus on **one clear message** per slide/chart
- **Title of the slide** should reflect that key message
- The words on the slide and the visuals should **reinforce one another**
- The order should **reflect natural reading patterns** (left to right, top to bottom)



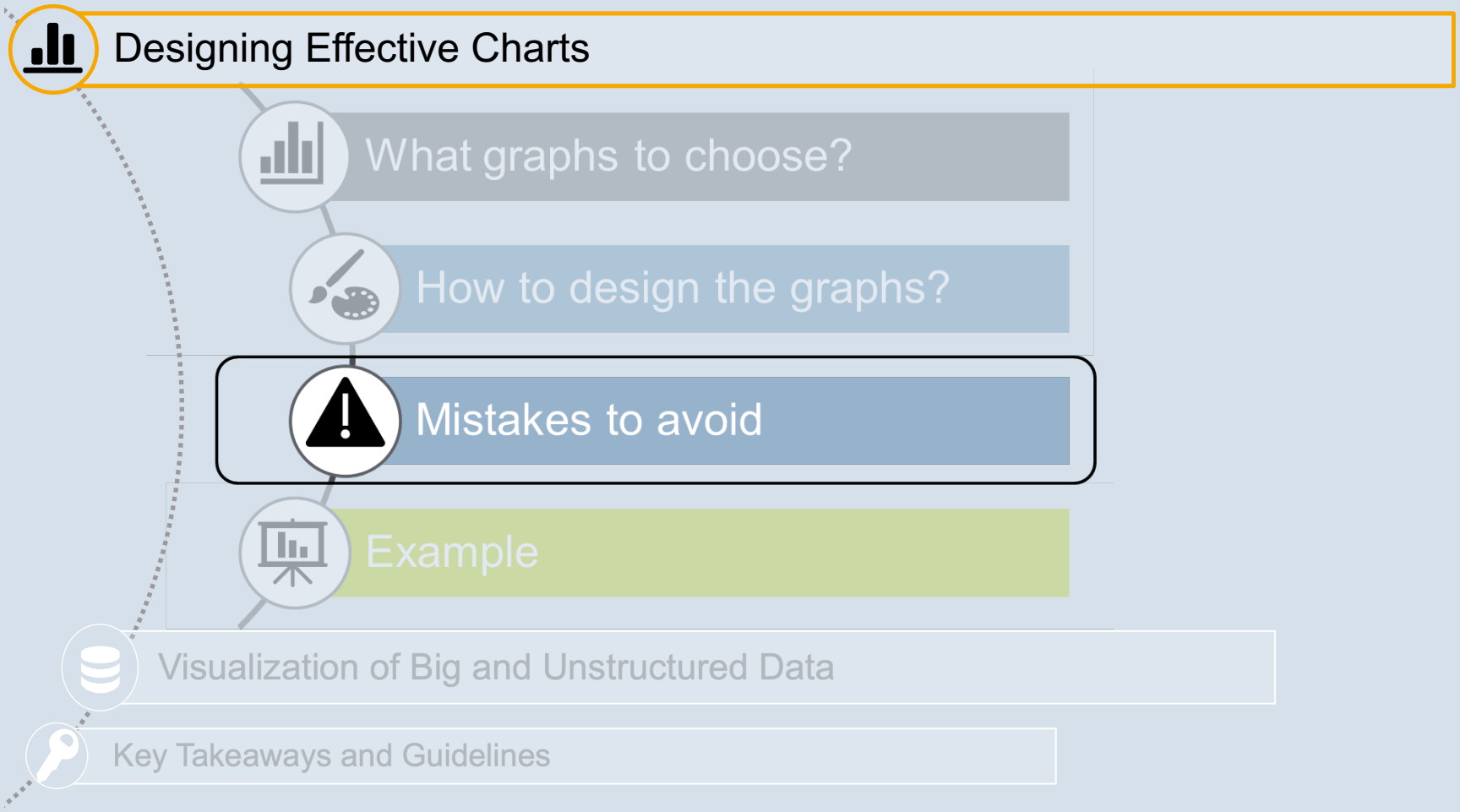
SOURCE: McKinsey.com

Tip 5 – Keep it simple

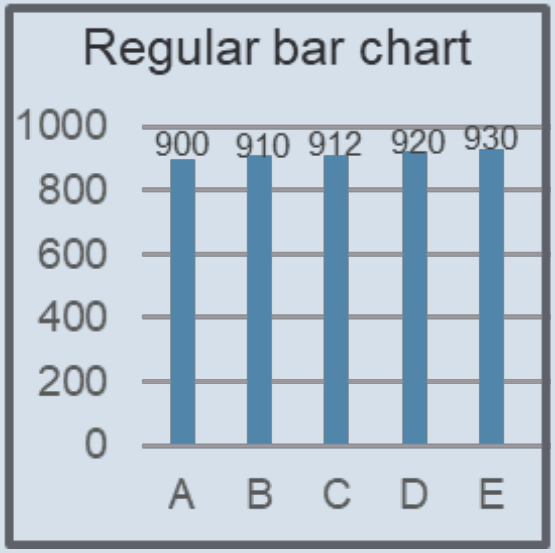
Which elements should you keep?

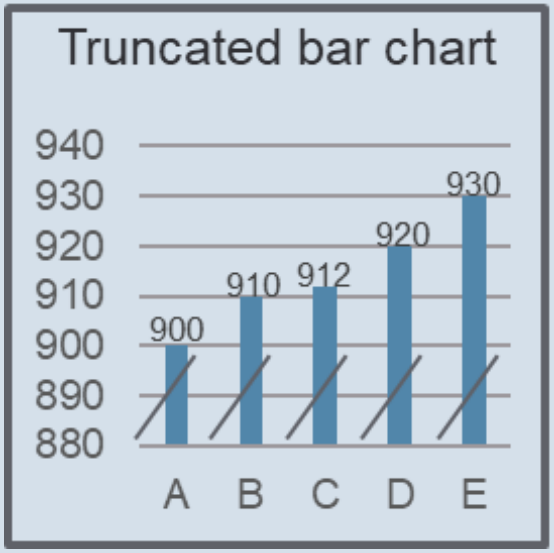


Data Literacy & Data Intuition



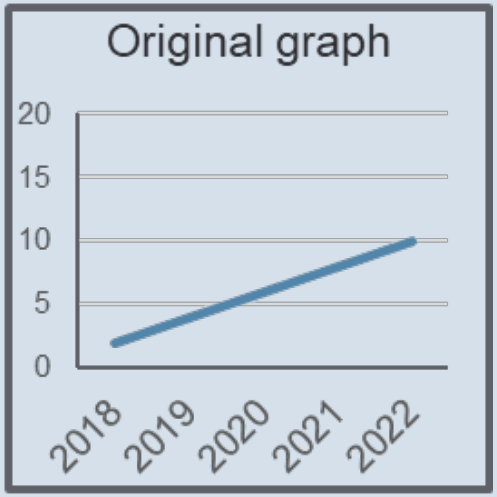
Mistake 1 – Truncated Graphs

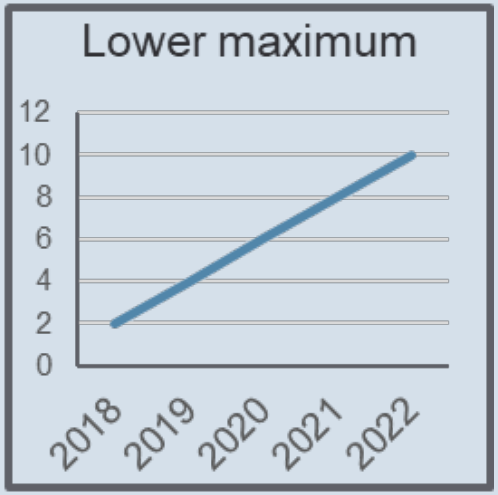


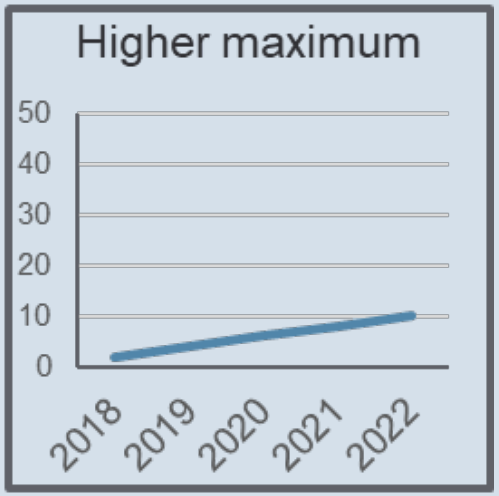


Truncated graphs lead to a distortion of the numbers and an overestimation of factual differences.

Mistake 2 – Adjusted Axis

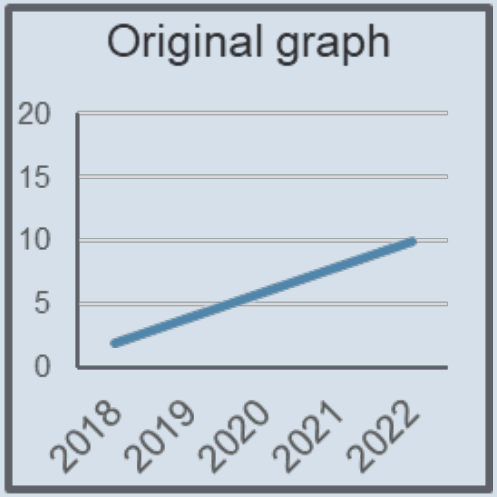


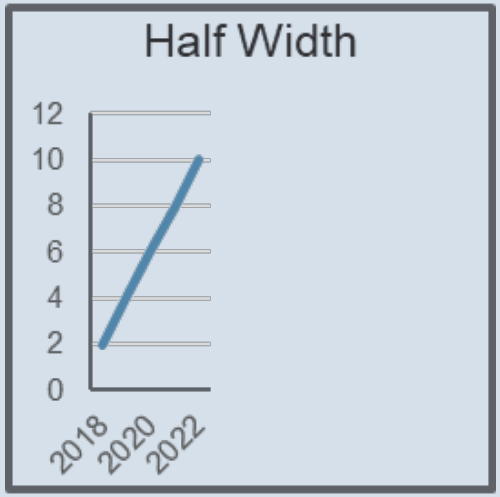


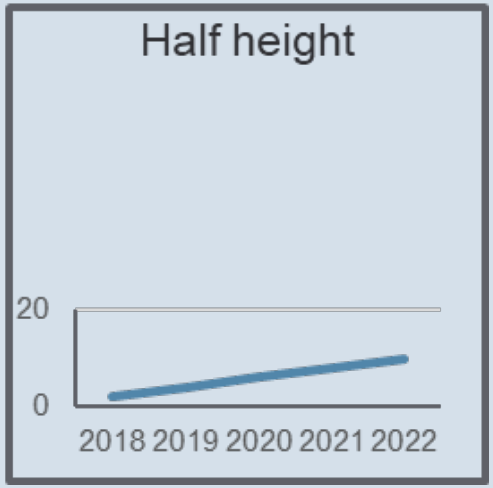


Changing the maximum value of the y-axis can have the same effect on perception as truncated graphs.

Mistake 2 – Adjusted Axis



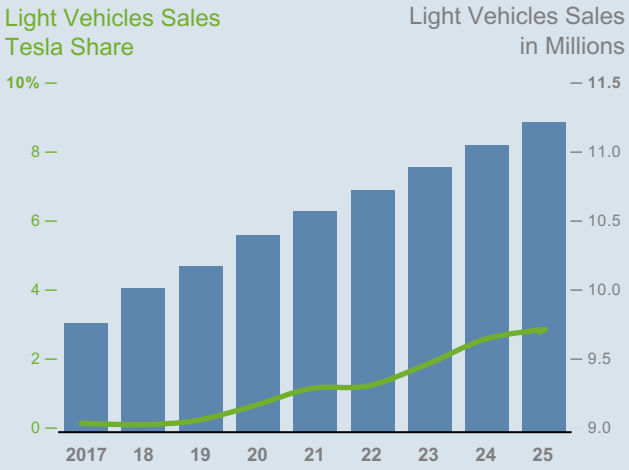


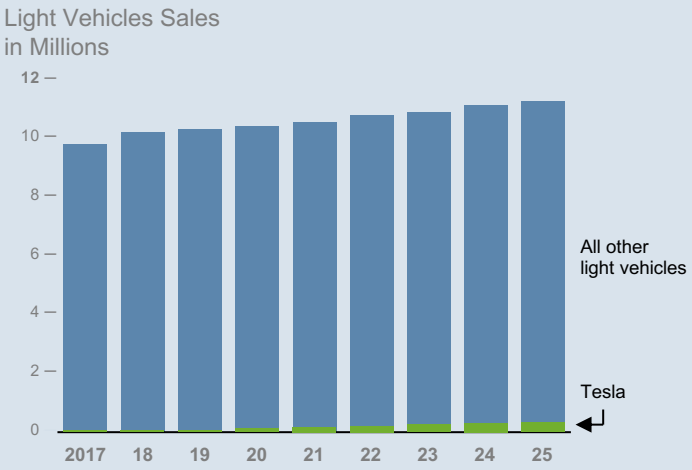


Changing the ratio of the chart can have the same effect on perception as truncated graphs.

Mistake 2 – Double Y-Axis

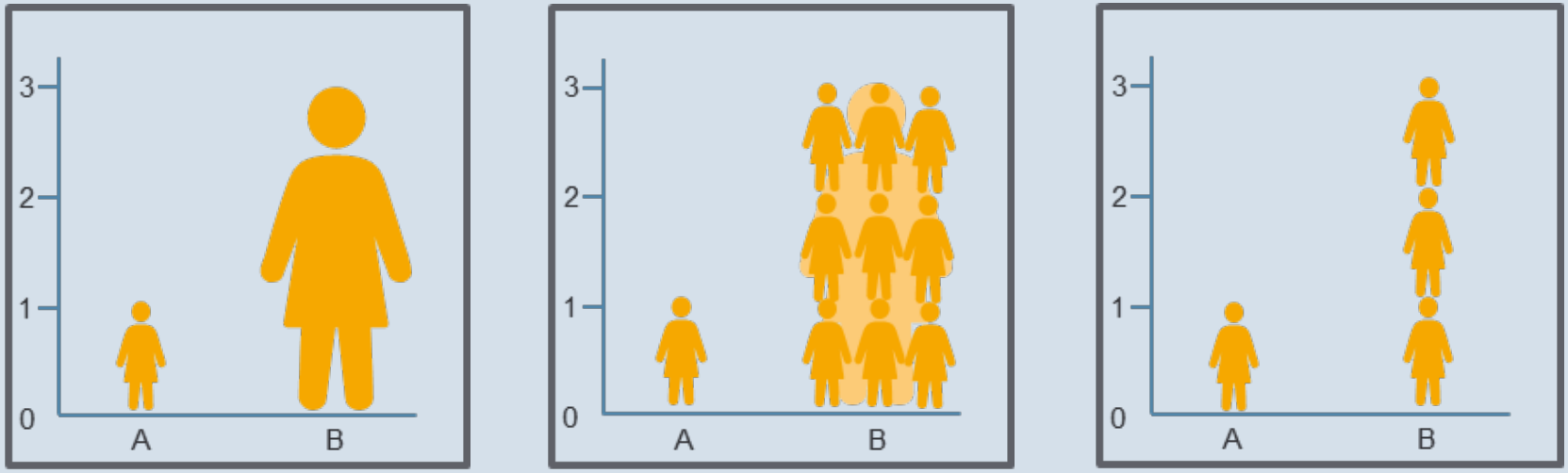
MANNHEIM BUSINESS SCHOOL





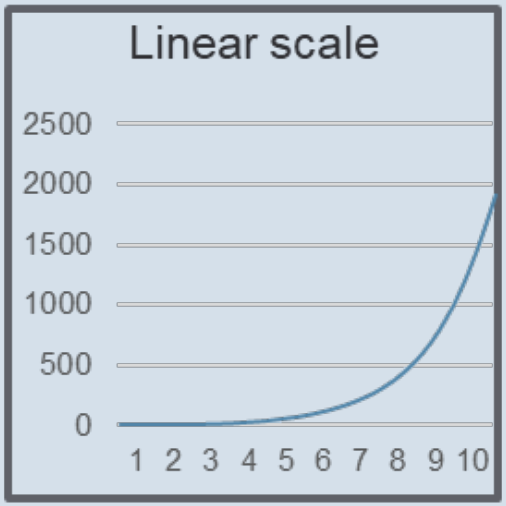
Double y-axis confuse the audience and often convey wrong messages for both measures.

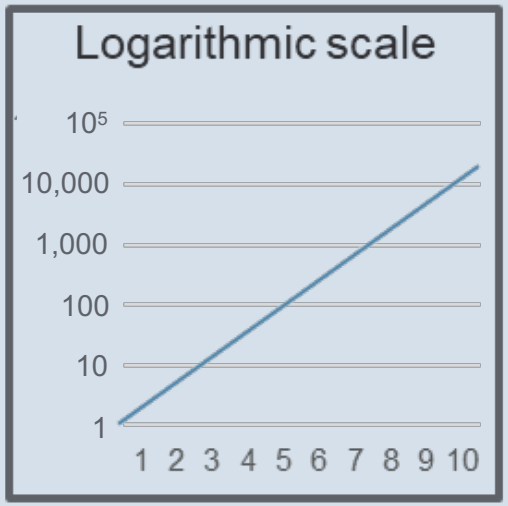
Mistake 3 – Incorrect scaling



Scaling up pictograms results in a perceptually misleading comparison.

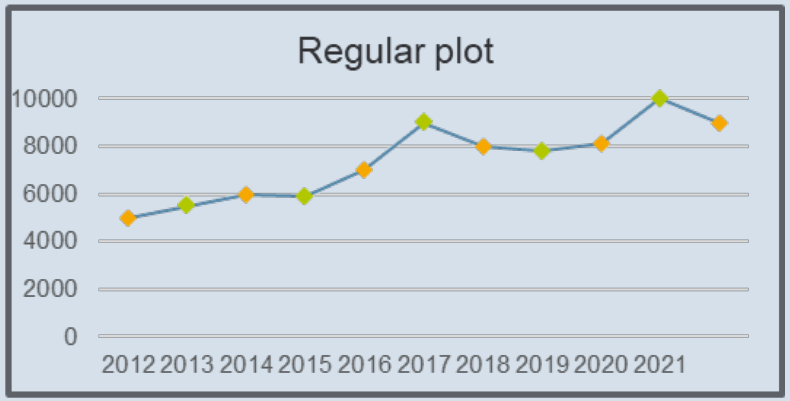
Mistake 4 – Logarithmic Scaling

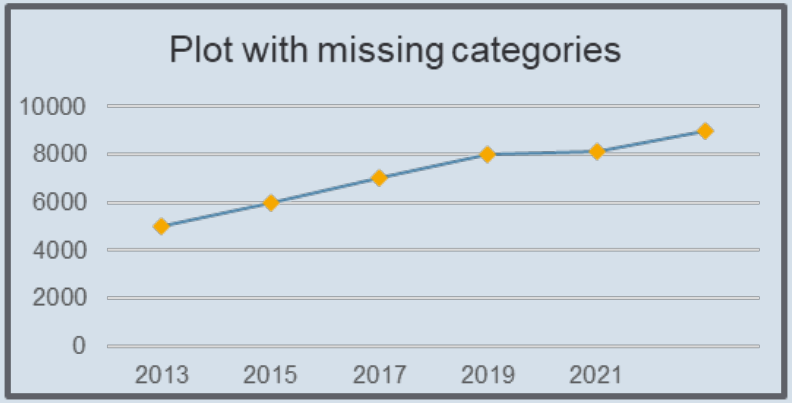




Logarithmic scales are prone to misinterpretation.

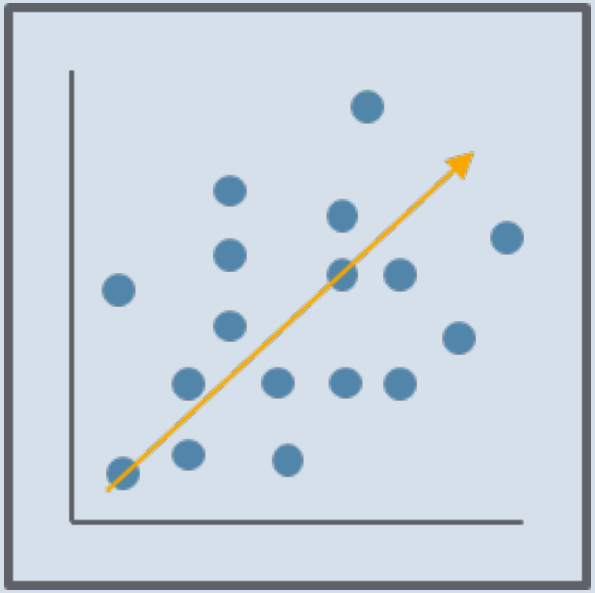
Mistake 5 – Omitting Data





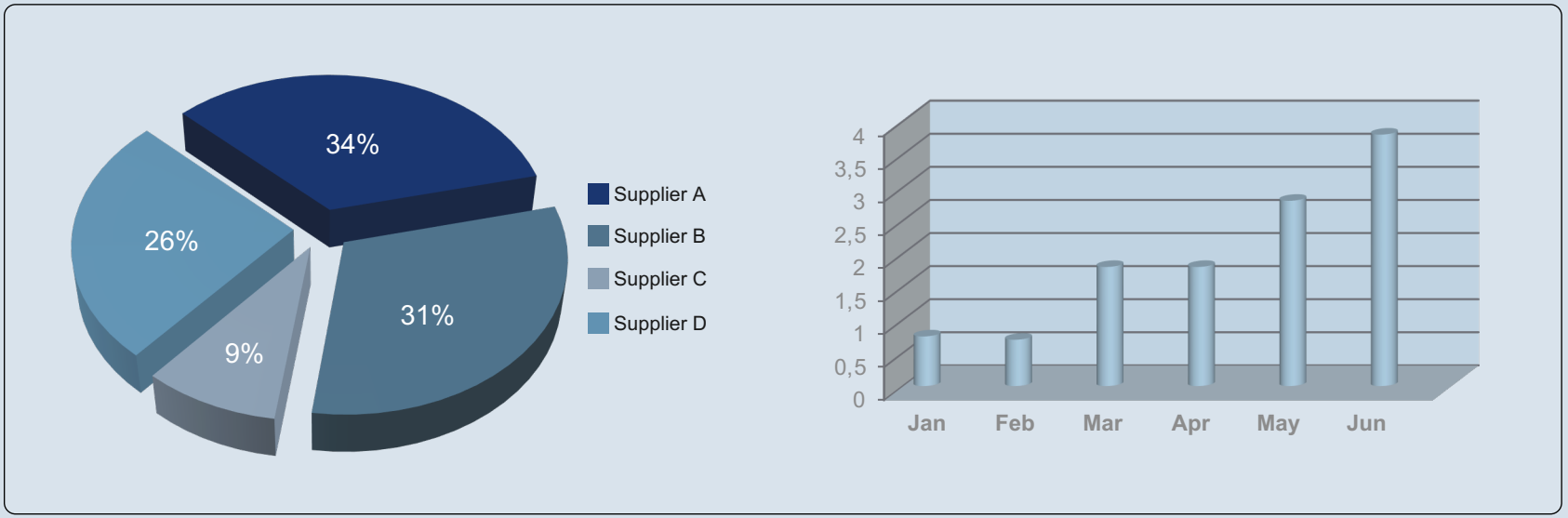
Omitting outliers might lead to wrong conclusions.

Mistake 6 – Simulating Trends



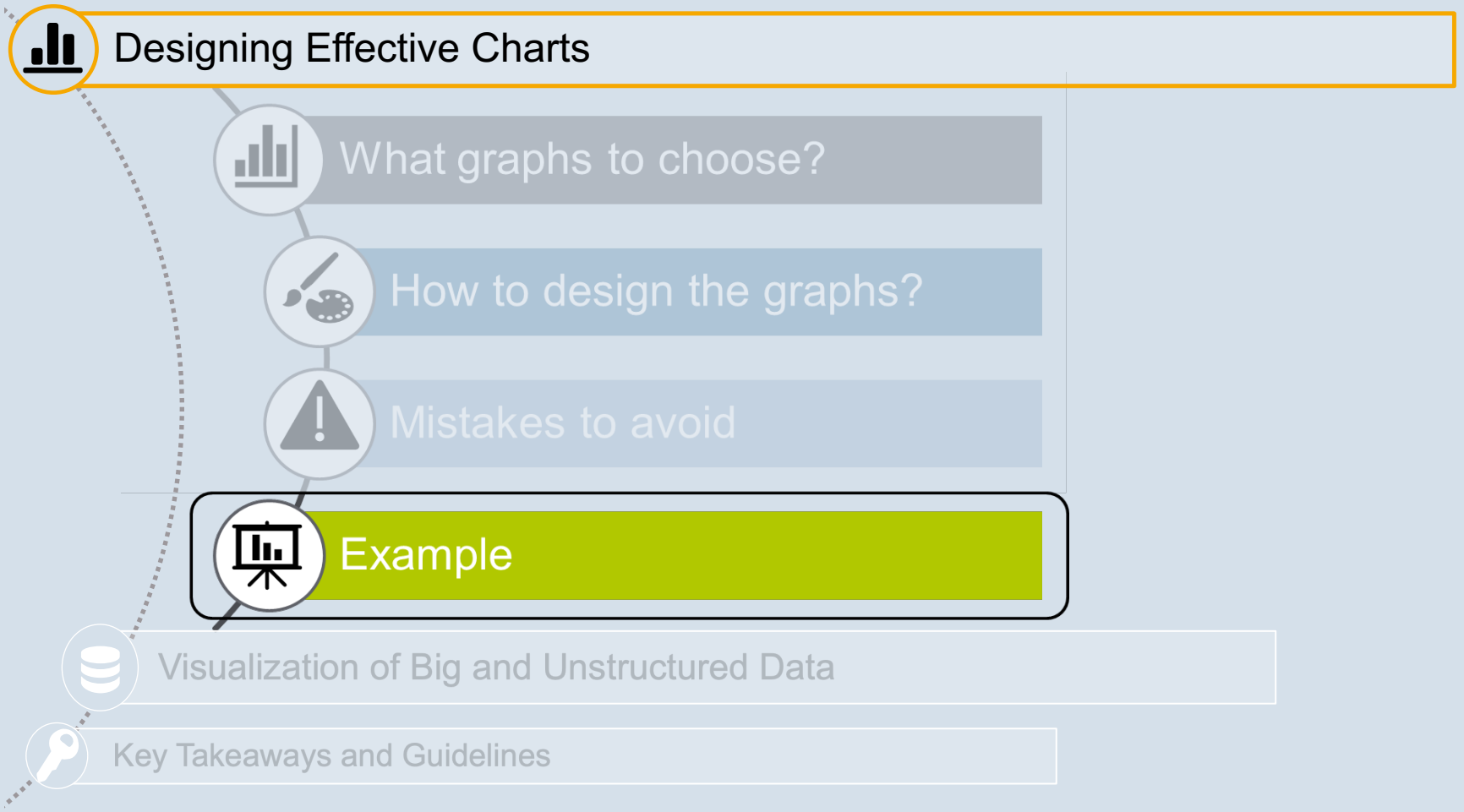
Additional tools such as trend lines should be used for simulate trends that are not there.

Mistake 7 – Redundant 3D Perspective



3D perspectives are often just used for aesthetic reasons and lead to the distortion of actual values.

Data Literacy & Data Intuition

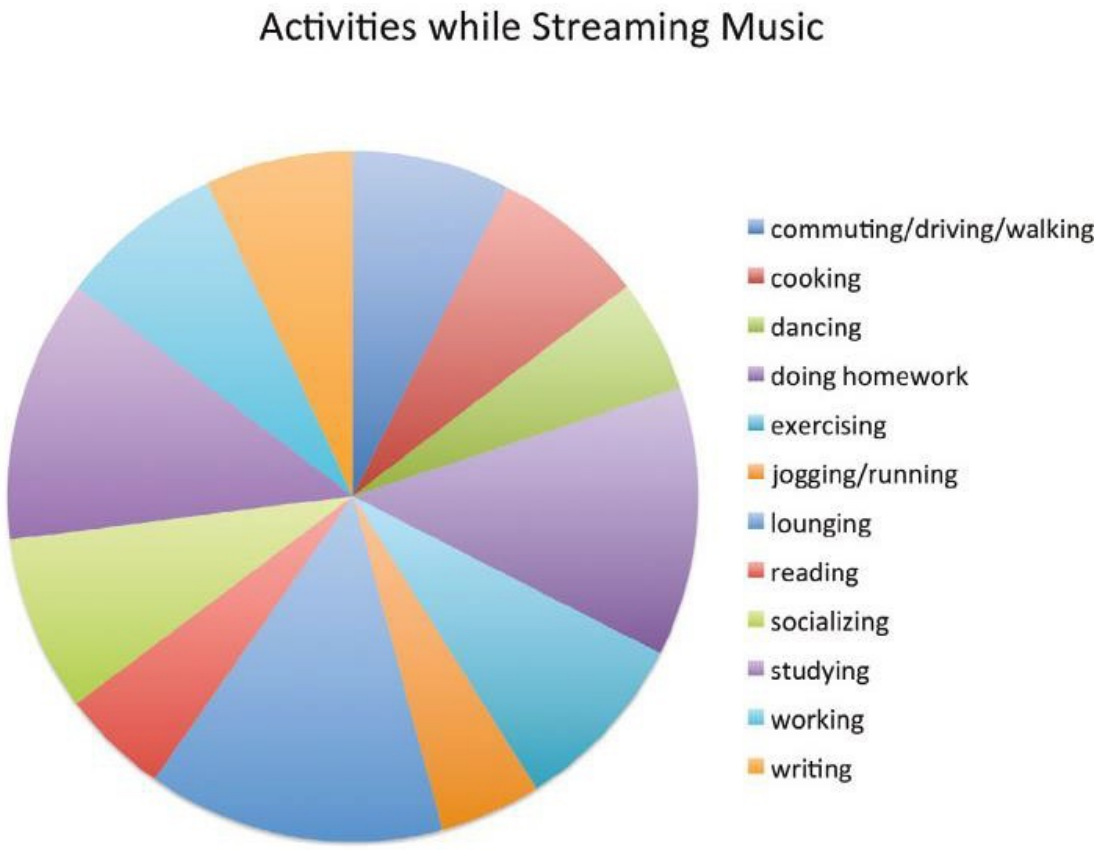


Example from Berinato (2016) “Good Charts”

MANNHEIM BUSINESS SCHOOL

- Pie chart the wrong choice here
- Too many colors used
- Uncertain labeling
- No clear comparison possible
- Overwhelming

Another candidate to improve …



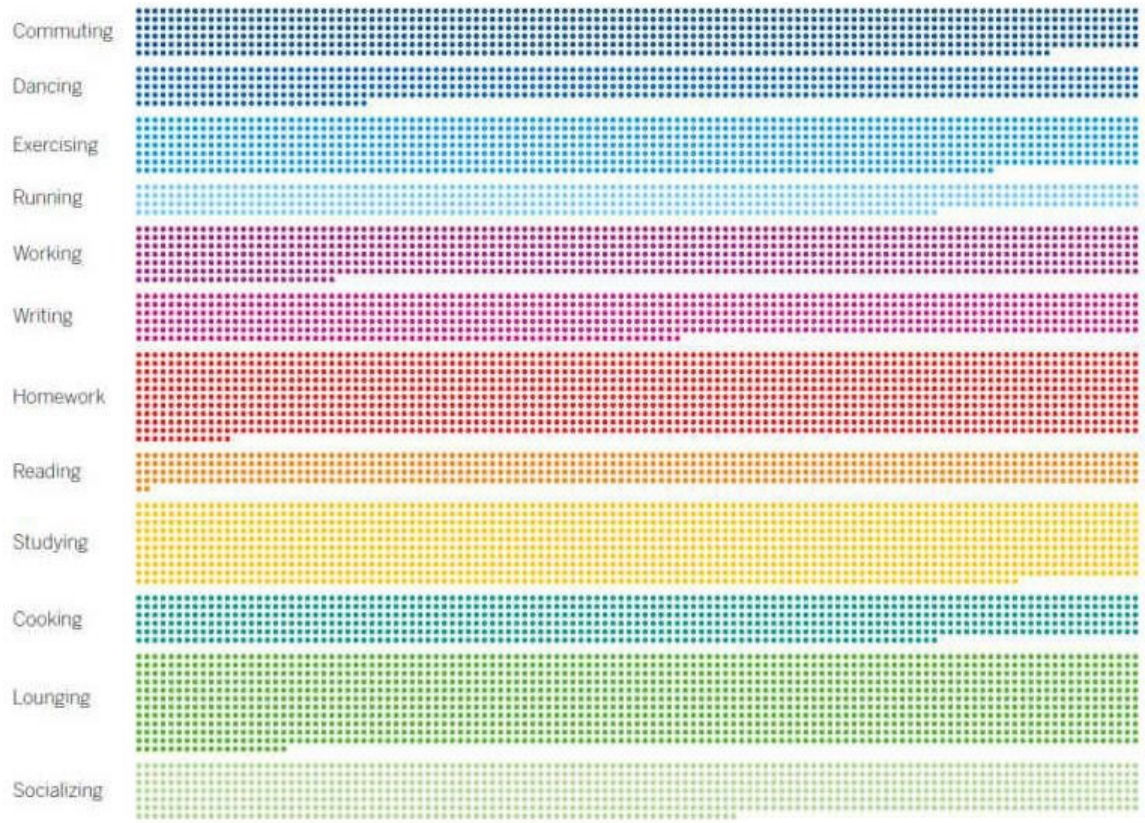
Example from Berinato (2016) “Good Charts”

MANNHEIM BUSINESS SCHOOL

- Clear labeling
- Better graph used

- Too many colors used
- No clear comparison possible
- Overwhelming

From a more effective prototype …



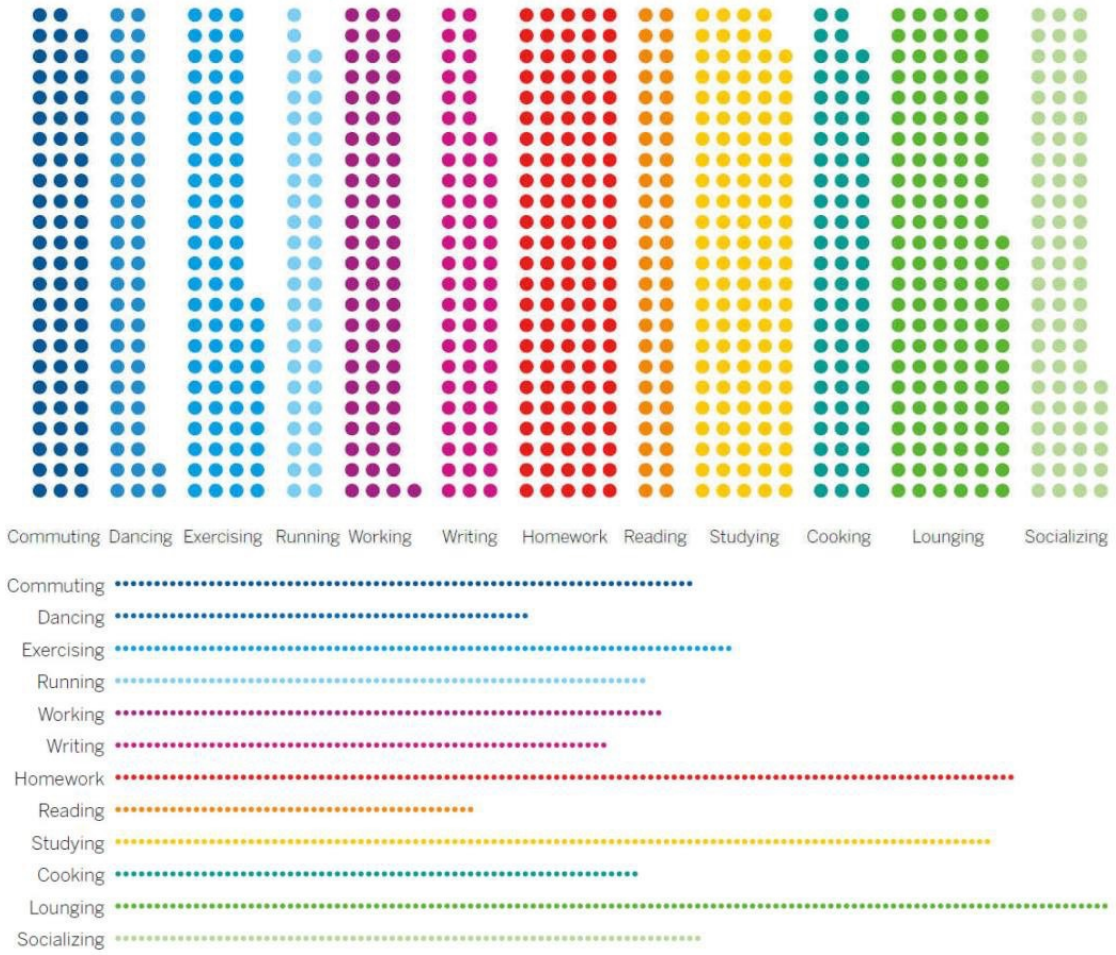
Example from Berinato (2016) “Good Charts”

MANNHEIM BUSINESS SCHOOL

- Clear labeling
- Better graph used
- Comparison possible

- Too many colors used
- Upper graph still without clear message

… over refinements …



Example from Berinato (2016) “Good Charts”

MANNHEIM BUSINESS SCHOOL

- Easy to draw clear messages from
- Good use of colors
- Neat design
- Easy comparison possible
- Can be acted upon

… to a visualization with impact