



emergo

# DATA STORYTELLING WORKSHOP

Facilitator: Louis de Roo

# AGENDA

1

**Introduction**

2

**Datadriven**

3

**KPI's**

4

**Data Visualization**

5

**Assignment: create a data story**

6

**Presenting Results**





emergo

**MANAGE YOUR DATA,  
INNOVATE YOUR  
BUSINESS**



**Louis de Roo**

Data Strategy Leader



# ABOUT E-MERGO

- E-mergo was founded in 2001
- +/- **70 co-workers**
- Based in **Delft**
- **Software and consultancy**
- **4 departments:** Data Management, Data Analytics, Low-Code App Development, Data Driven
- **Microsoft Gold partner** (Azure, Power BI)
- **TimeXtender partner of the Year** for Data Warehouse Automation
- **Qlik Elite Partner** for Qlik Data Analytics and Qlik Data Integration
- **Mendix partner** for app development





# E-MERGING TALENT PROGRAM

Almost graduating? Interested in data? Then the E-merging Talent Program is just what you're looking for!

During this program you will be trained into a data expert within 2 years.

**The entire training path has a total value of over €45.000,-**



#### Data Analytics techniques

- Advanced Analytics basics
- Data modelling
- Data visualisation
- KPI's and metrics
- Datawarehouse automation
- Infrastructure networks
- Process modelling



#### Tooling

- Power BI
- Qlik
- TimeXtender
- Azure



#### Soft skills

- Advisory skills
- Presentation skills
- Communicate effectively
- Agile methodology

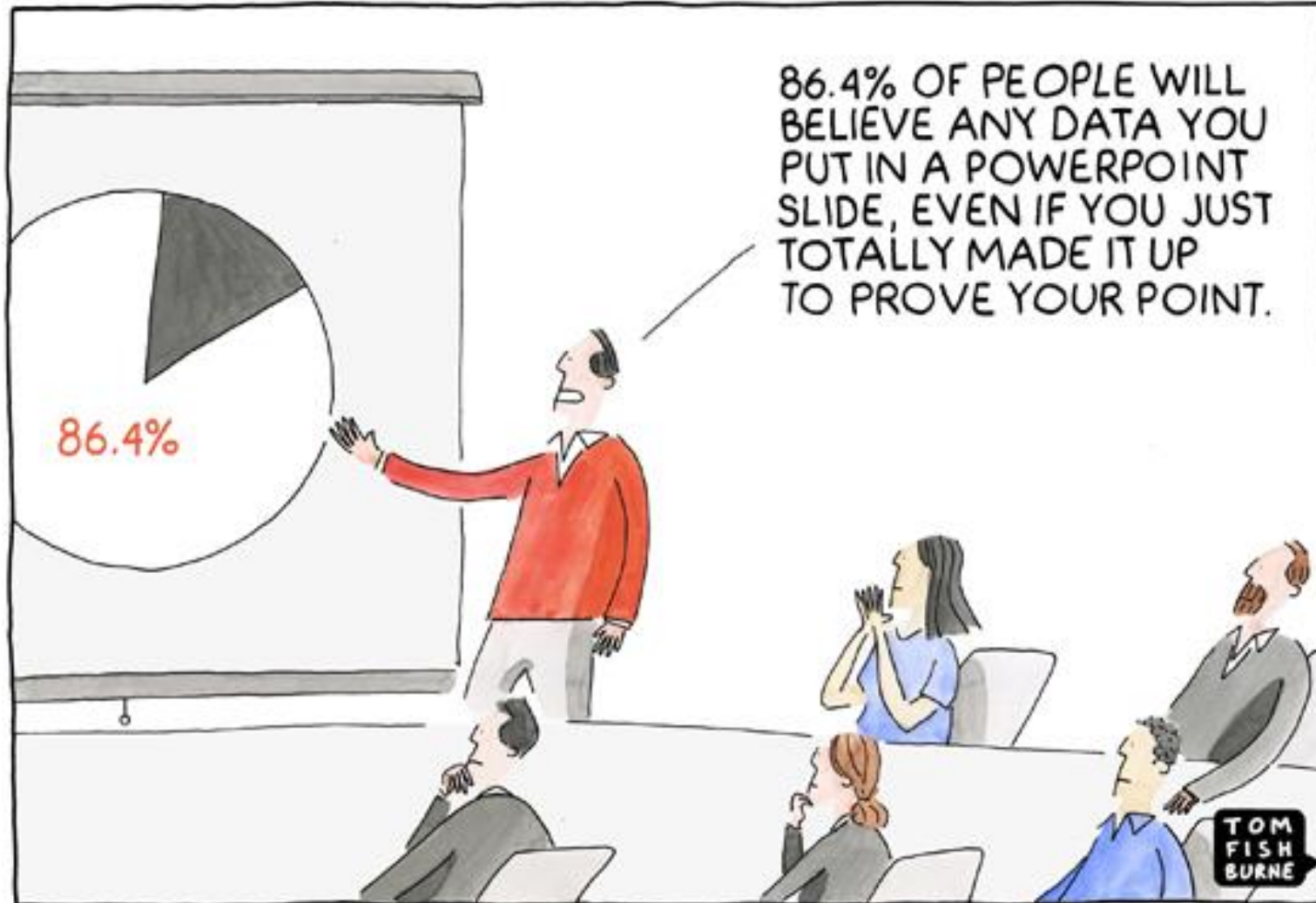


**Scan the QR-code  
for more info!**

Interested? Register via the link or reach out to our recruiter Josita on LinkedIn!

# DATA-DRIVEN

emerge



© marketoonist.com

# DATA-DRIVEN IN 7 STEPS

- Attainable and scalable
- From support to full service
- Tailored approach
- Result-driven





Value

Development

Value



Development

Value



Development

Value



Development



Value



Development

- A Key Performance Indicator is a measurable value that demonstrates how effectively a company is achieving key business objectives

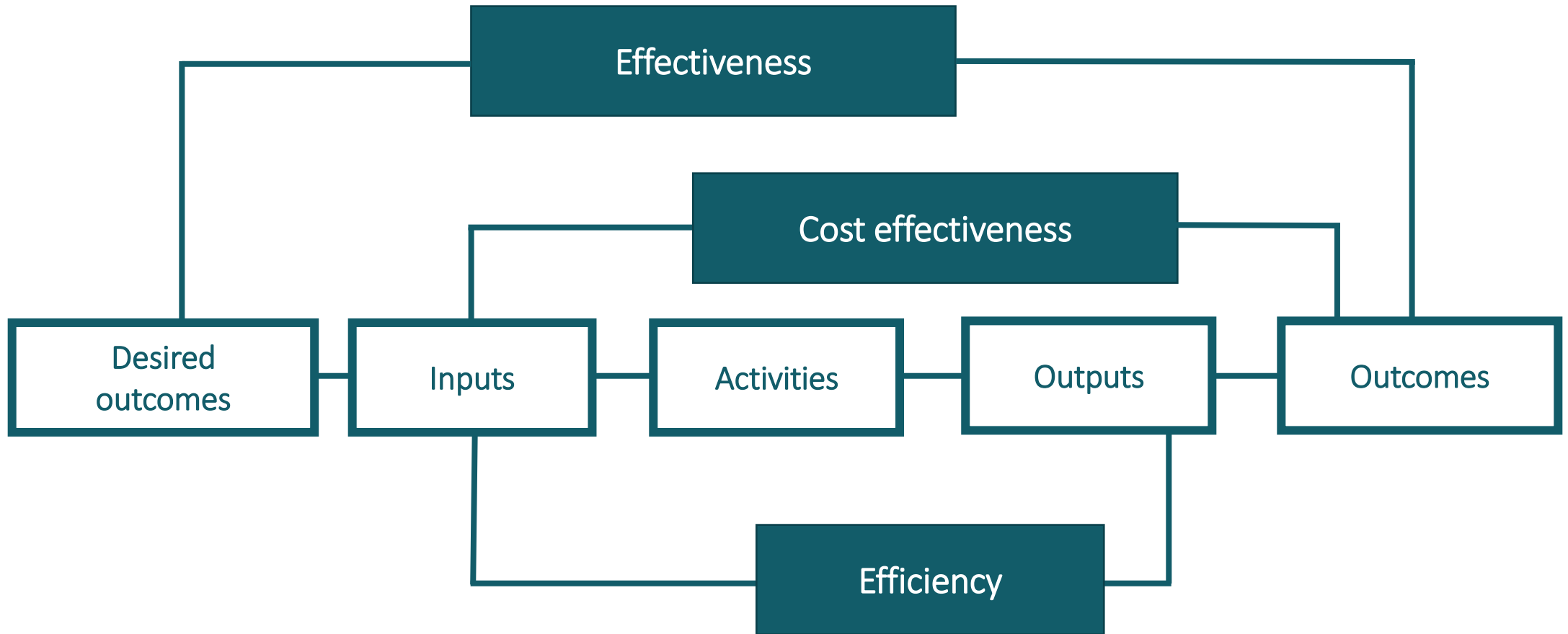


© Scott Adams, Inc./Dist. by UFS, Inc.

GOODHART'S LAW:

"When a measure becomes a target, it ceases to be a good measure."

# VALUE FLOW ANALYSIS





# DATA-DRIVEN WITH KPI'S

- Accessible
- Actionable
- Controllable
- Cost-effective
- Incentivised
- Standardised
- Timely
- Verifiable



# DATA-DRIVEN WITH KPI'S

- Decisions over reporting
- Connect to the user level
- Process over output
- Connect to customer promise



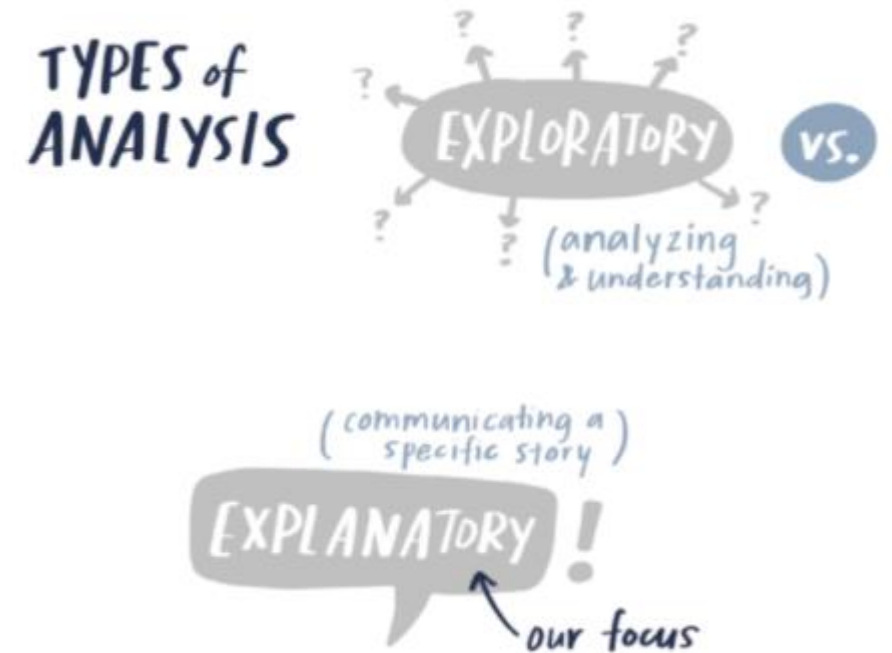
# DATA VISUALIZATION

emergo



# WHAT IS DATA VISUALIZATION?

*Data visualization is the graphical representation of information and data by using visual elements like charts, graphs, and maps. Data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.*





# FRANCIS ANSCOMBE (ANSCOMBE'S QUARTET)

Property	Value
Mean of $x$	9
Sample variance of $x : s_x^2$	11
Mean of $y$	7.50
Sample variance of $y : s_y^2$	4.125
Correlation between $x$ and $y$	0.816
Linear regression line	$y = 3.00 + 0.500x$
Coefficient of determination of the linear regression : $R^2$	0.67

I	
x	y
10.0	8.04
8.0	6.95
13.0	7.58
9.0	8.81
11.0	8.33
14.0	9.96
6.0	7.24
4.0	4.26
12.0	10.84
7.0	4.82
5.0	5.68

# FRANCIS ANSCOMBE (ANSCOMBE'S QUARTET)

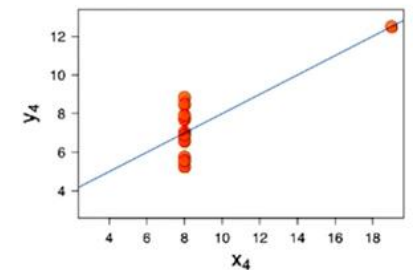
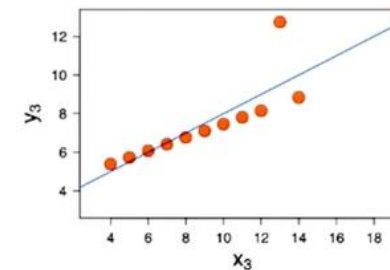
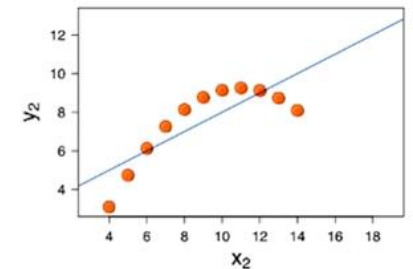
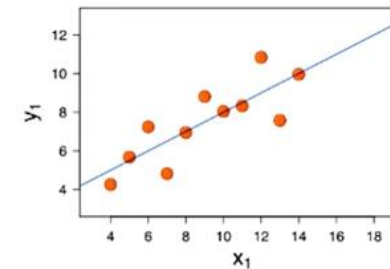
Property	Value
Mean of $x$	9
Sample variance of $x : s_x^2$	11
Mean of $y$	7.50
Sample variance of $y : s_y^2$	4.125
Correlation between $x$ and $y$	0.816
Linear regression line	$y = 3.00 + 0.500x$
Coefficient of determination of the linear regression : $R^2$	0.67

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

# FRANCIS ANSCOMBE (ANSCOMBE'S QUARTET)

Property	Value
Mean of $x$	9
Sample variance of $x : s_x^2$	11
Mean of $y$	7.50
Sample variance of $y : s_y^2$	4.125
Correlation between $x$ and $y$	0.816
Linear regression line	$y = 3.00 + 0.500x$
Coefficient of determination of the linear regression : $R^2$	0.67

I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89



# WHY IS DATA VISUALIZATION IMPORTANT?

emergo

- See the bigger picture
- Identify patterns, outliers etc.
- Present meaningful data
- Sharing understandable insights
- Democratize your data
- Creating simple visuals
- Make better decisions

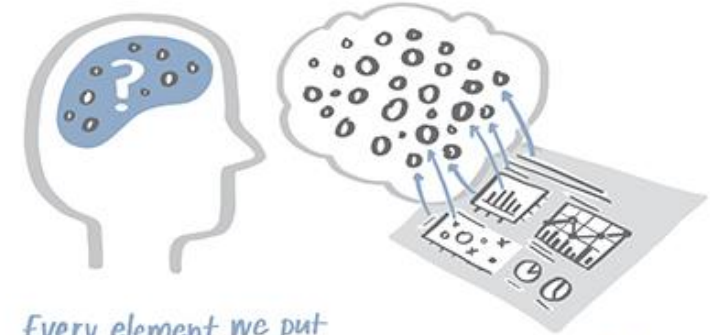


# REDUCE CLUTTER

- Cognitive load
- **Gestalt principles of visual perception**
  - Proximity
  - Similarity
  - Enclosure
  - Closure
  - Continuity
  - Connection

## COGNITIVE LOAD

The MENTAL EFFORT that's REQUIRED  
to LEARN NEW INFORMATION



Every element we put  
on a page or screen  
puts cognitive burden  
on our audience...

so we should take  
care not to include  
things that aren't  
adding information

Keep Gestalt principles of  
visual perception in mind



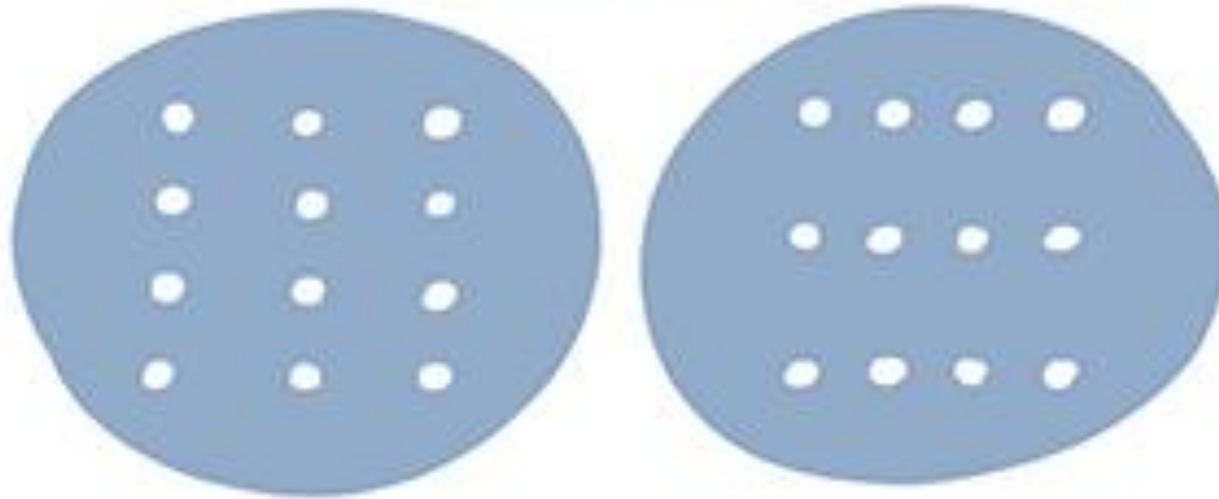
Tip: The visual display of quantitative information by  
Edward tufte



# PROXIMITY

Objects that are close together belong to the same group

PROXIMITY



# SIMILARITY

emergo

Objects with the same colour, form or orientation belong to the same group

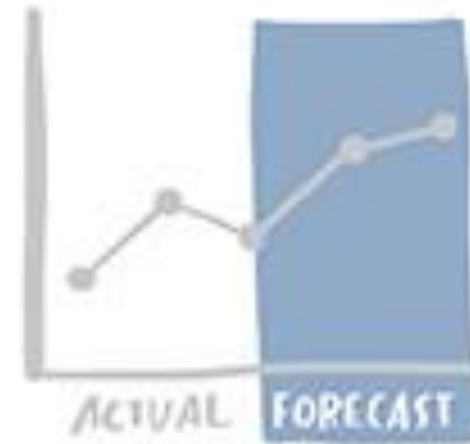
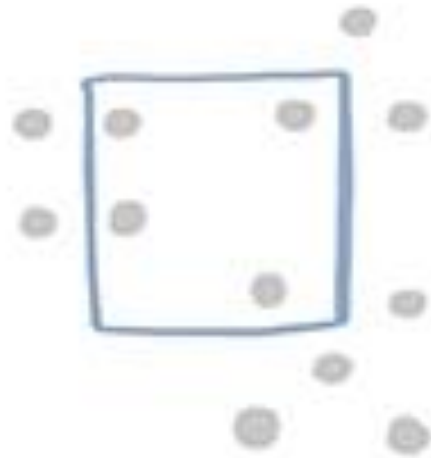
*SIMILARITY*



# ENCLOSURE

Objects with enclosure belong to the same group

ENCLOSURE



# CLOSURE

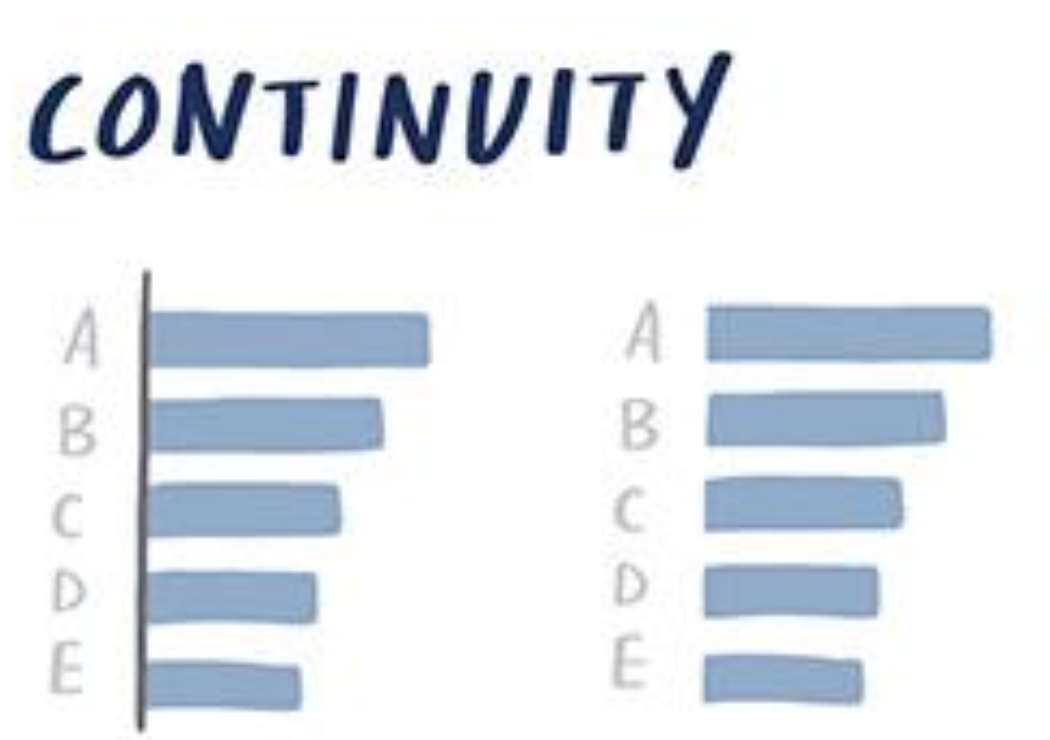
Our brain let us think that certain constructs exist, but they do not

*CLOSURE*



# CONTINUITY

Our brain creates continuity that doesn't exist

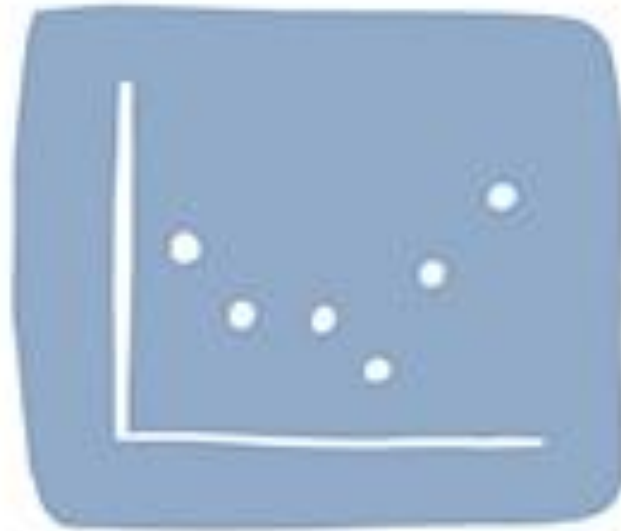




# CONNECTION

Objects that are connected to each other belong to the same group

*CONNECTION*



# WHAT AMOUNT OF CLUTTER IS ACCEPTABLE?

---

- Decimals
- Percentages
- Commas
- Granularity of data



# VISUALIZE YOUR DATA.

Different charts for different messages

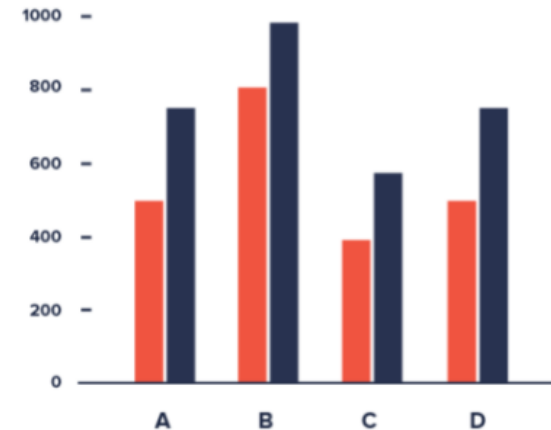
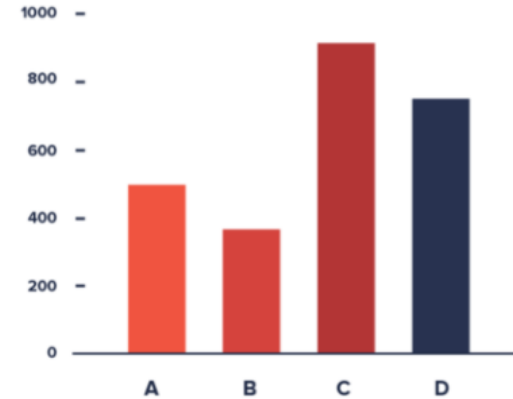
- Viewing comparisons
- Viewing development over time
- Viewing relationship
- Viewing compositions
- Viewing distributions
- Viewing geography
- Viewing data

Use different charts  
for different messages  
Use best practices



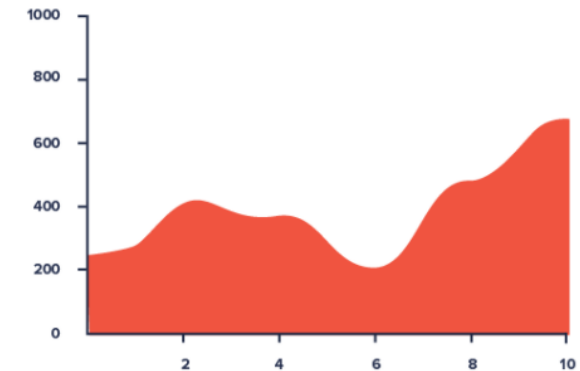
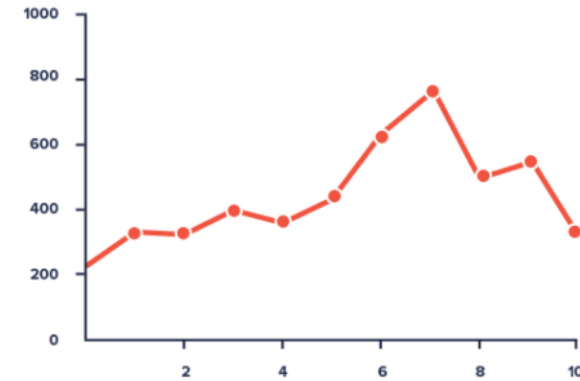
# VIEWING COMPARISONS

- Comparison charts are used to compare values against each other
  - *What product has the highest total sales this year?*



# VIEWING DEVELOPMENT OVER TIME

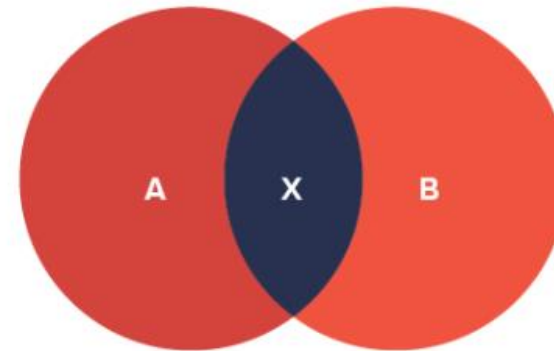
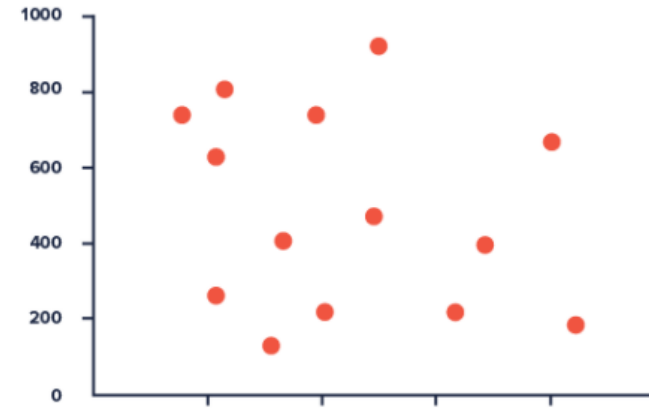
- Viewing development over time: Performance charts provide a quick view of a performance over time or a trend.
  - *What was the sales from Product A in the previous 6 months?*





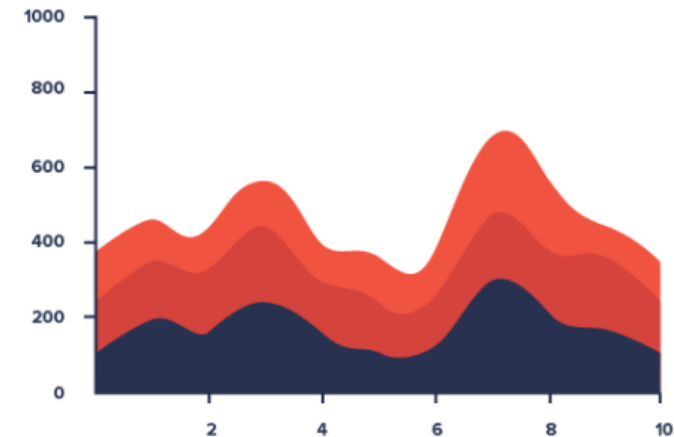
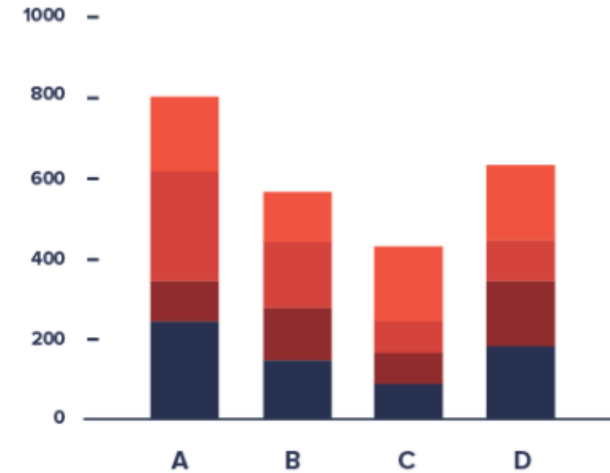
# VIEWING RELATIONSHIPS

- Relationship charts are used to explore how values relate to each other. A relationship chart allows you to find correlations, outliers, and clusters of data.
  - Is there a correlation between advertising spending and sales for our products?*



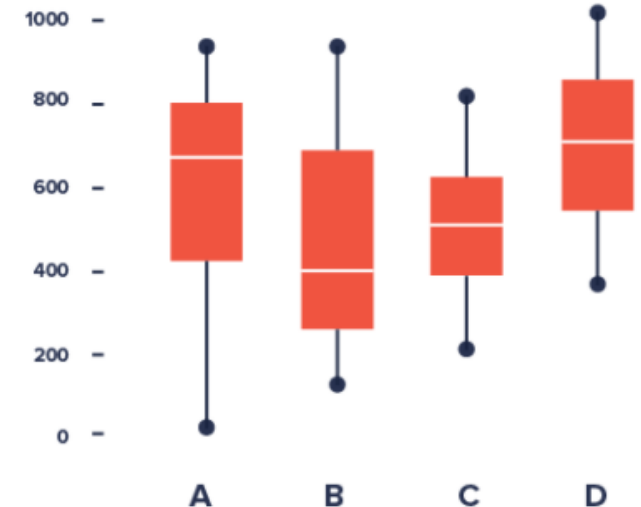
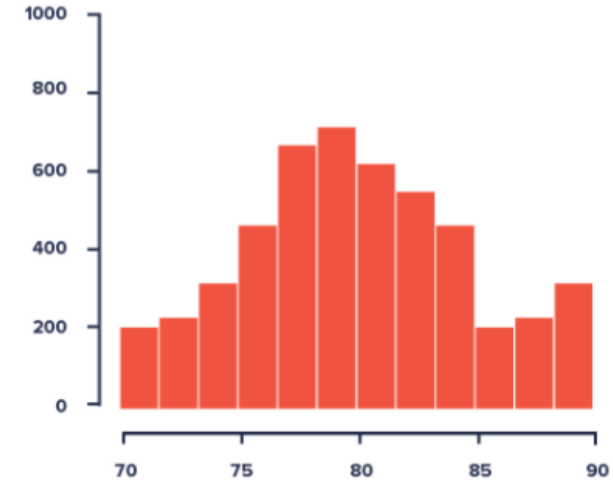
# VIEWING COMPOSITIONS

- Composition charts take a total value and discover what component values make up that total.
  - *What percentages of our total sales come from which regions?*
  - *What is each department's allotment of our total quarterly budget over the past year?*



# VIEWING DISTRIBUTIONS

- Distribution charts are used to explore how the values within data are grouped.
  - *What is the number of customers per age group?*



# VIEWING GEOGRAPHY

- Geographical charts let you visualize your data by geography, displaying your data on a map either as points or areas.
  - *What cities have the highest use of our services?*
  - *Which countries have the most customers?*



# VIEWING DATA

- Data charts present detailed data rather than a visualization of the data.
  - *What are the records for each transaction for this month?*

	A	B	C
X	\$40	240	48
Y	\$50	200	59
Z	\$60	310	79

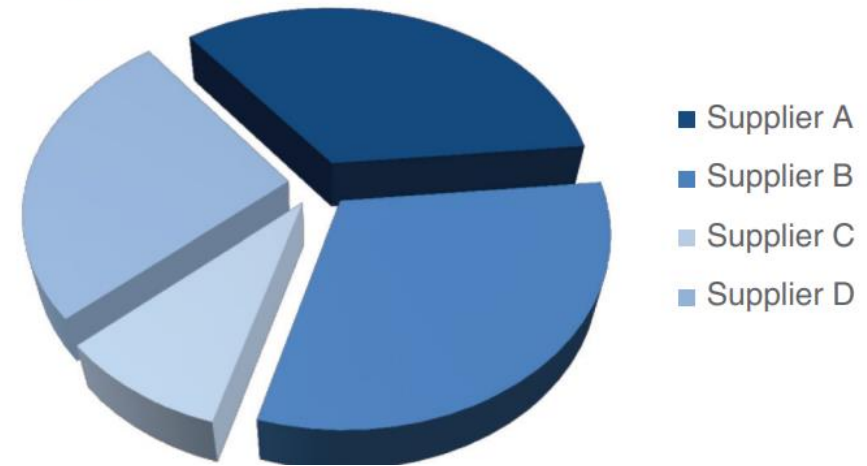
 **180%**



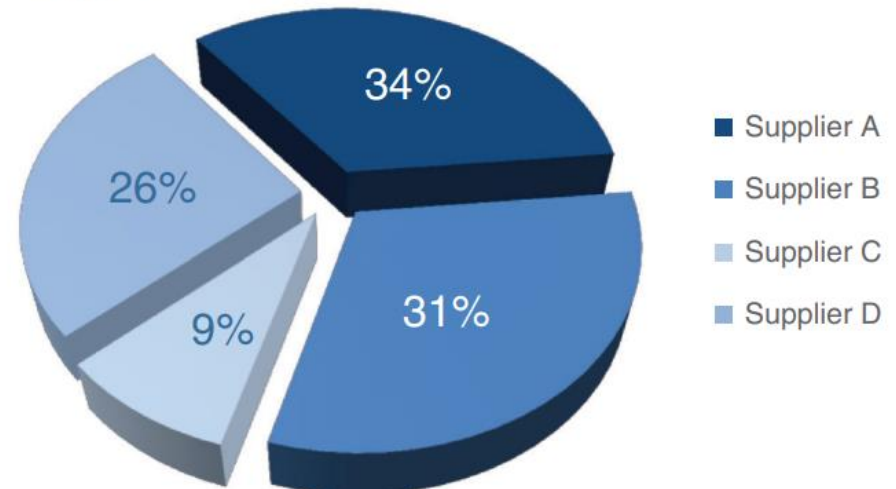
# TO AVOID IN DATA VISUALIZATION

- Why pie charts are evil

Supplier Market Share



Supplier Market Share

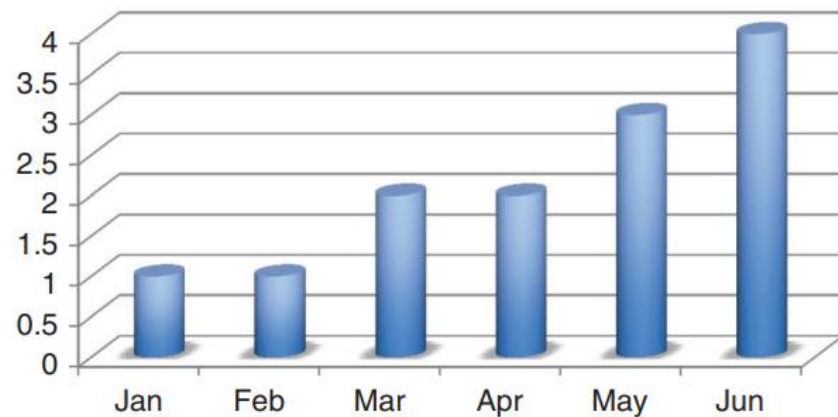


# TO AVOID IN DATA VISUALIZATION

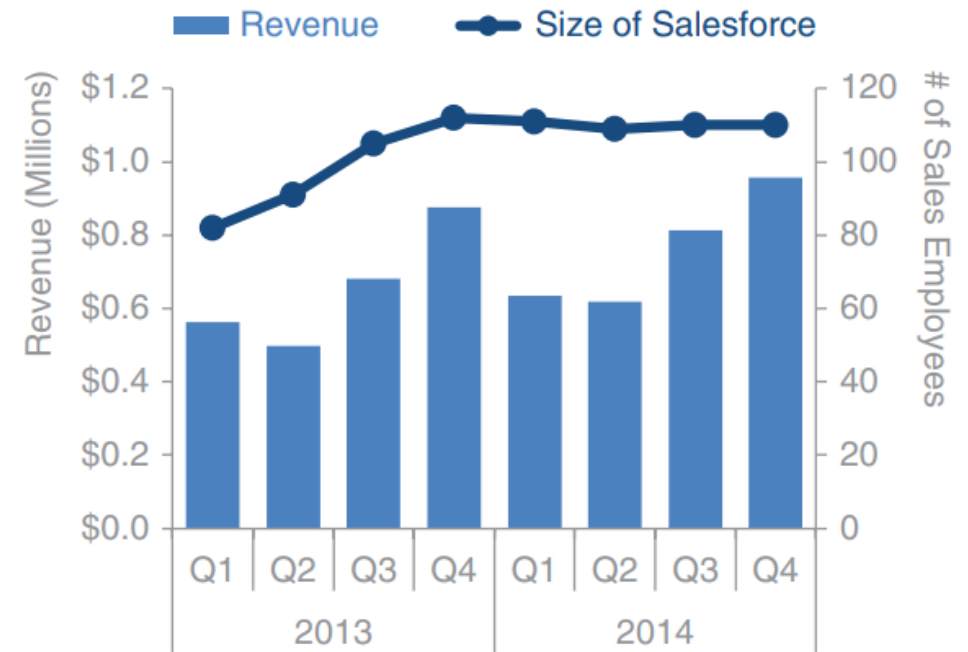
## 3D

- Exception: plotting a third dimension  
(but be careful)

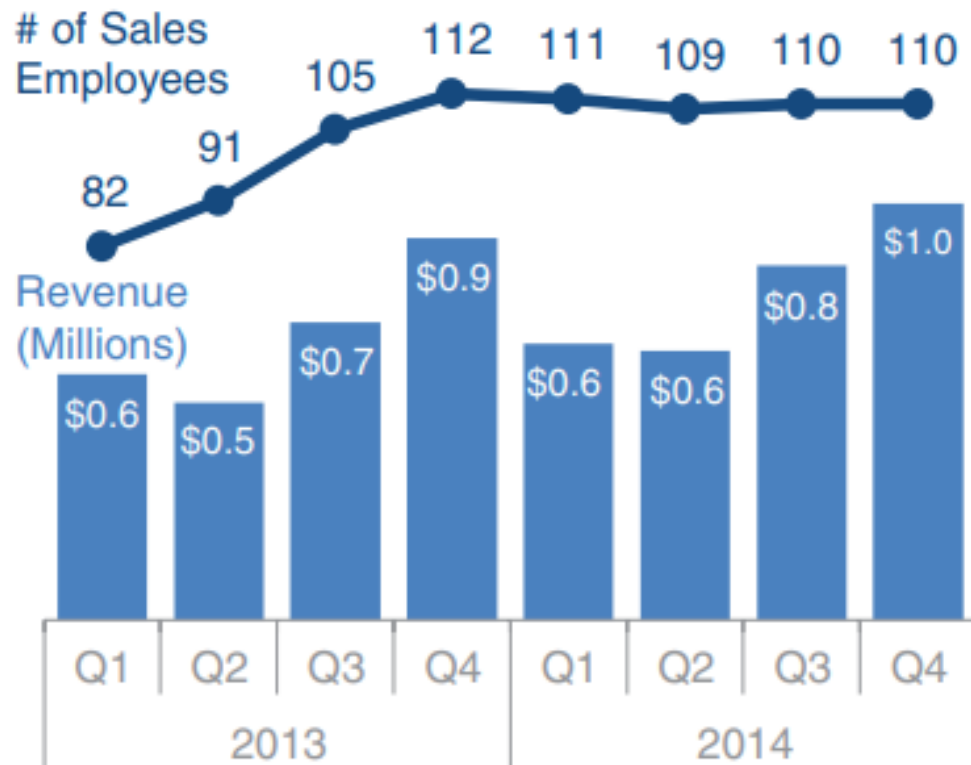
Number of issues



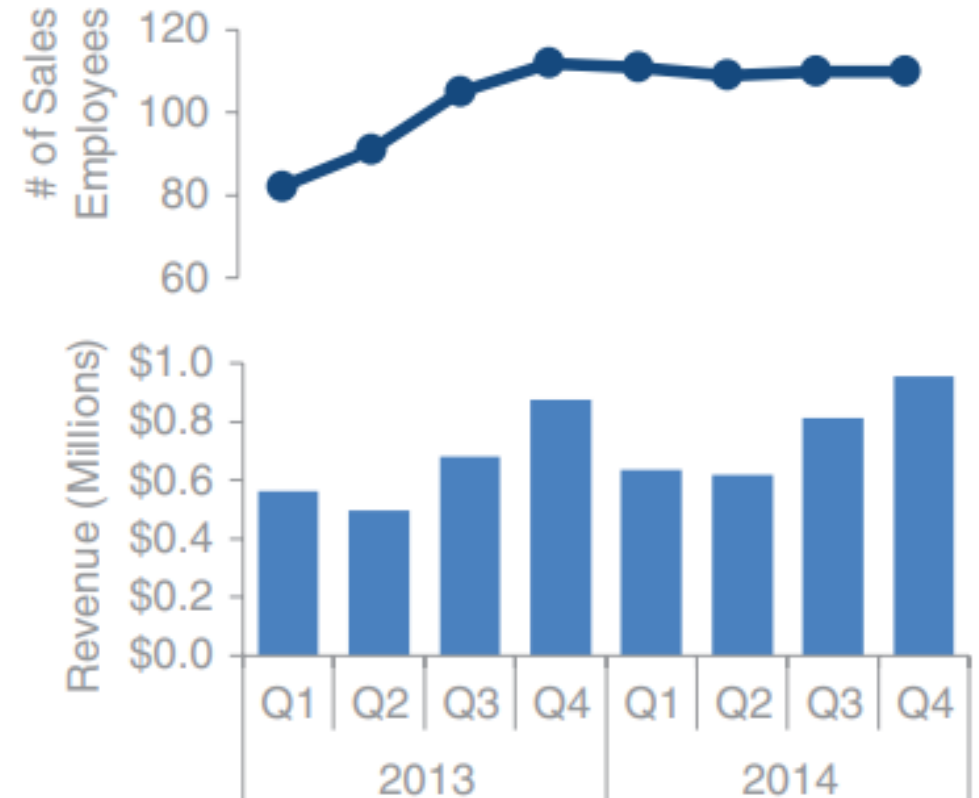
Secondary y-axis



## Alternative 1: label directly



## Alternative 2: pull apart vertically



756395068473

658663037576

860372658602

846589107830

756**3**9506847**3**

65866**3**0**3**7576

860**3**72658602

8465891078**3**0

# PREATTENTIVE ATTRIBUTES

SIGNAL WHERE to LOOK and create VISUAL HIERARCHY  
to help ease the processing of information

ORIENTATION



SHAPE



LINE LENGTH



LINE WIDTH



SIZE



CURVATURE



ADDED MARKS



ENCLOSURE



HUE



INTENSITY



SPATIAL POSITION



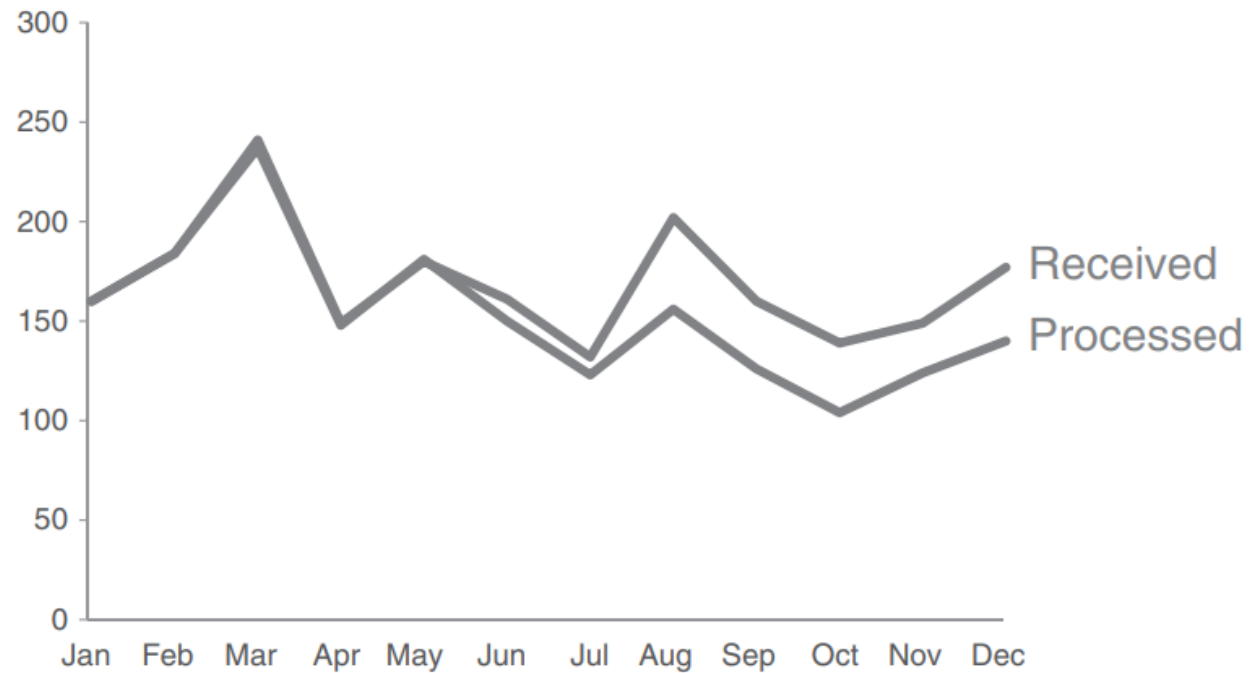
MOTION





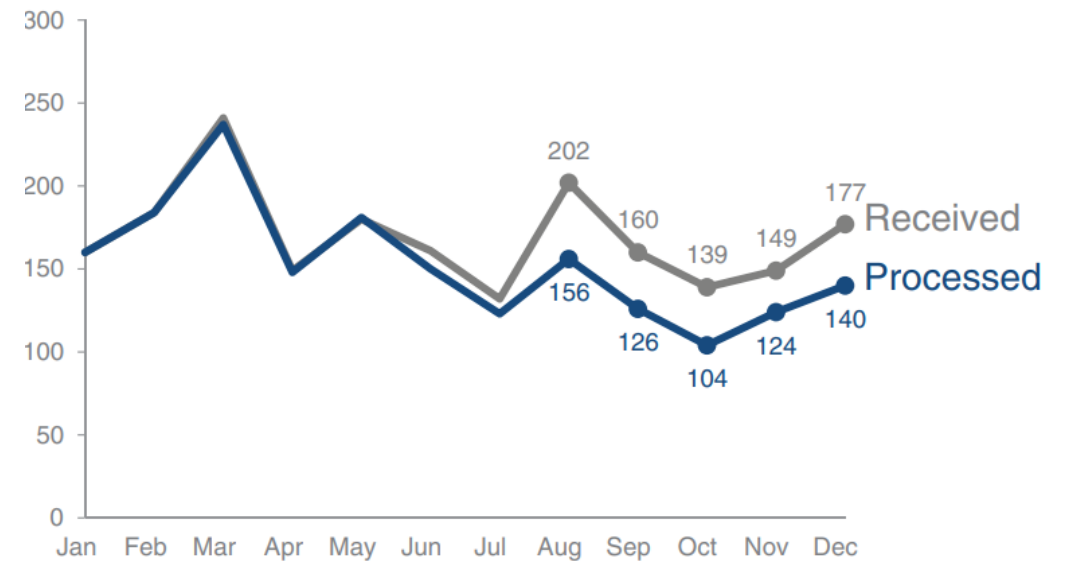
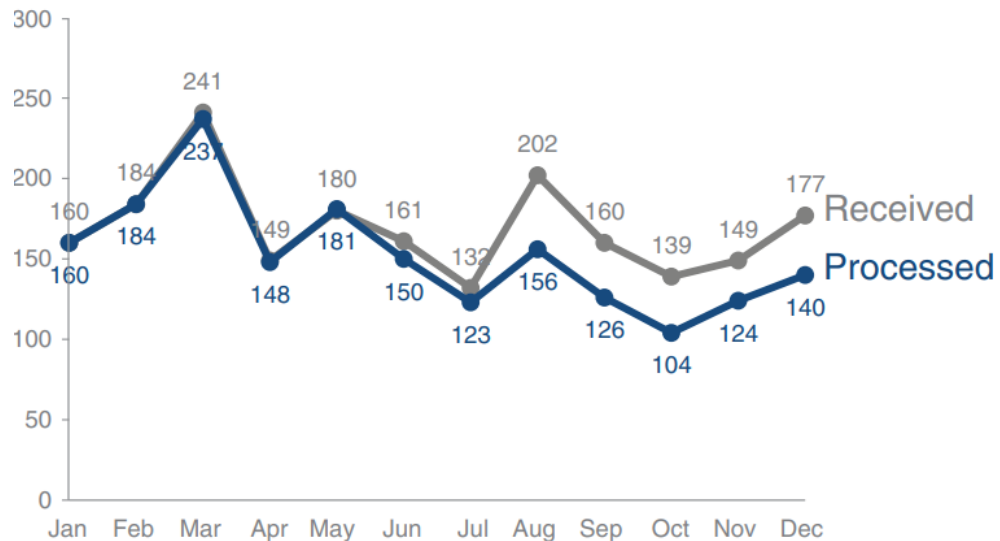
# PRE-ATTENTIVE ATTRIBUTES IN VISUALS

- *What is the story?*
- *Where do you want your audience to focus?*



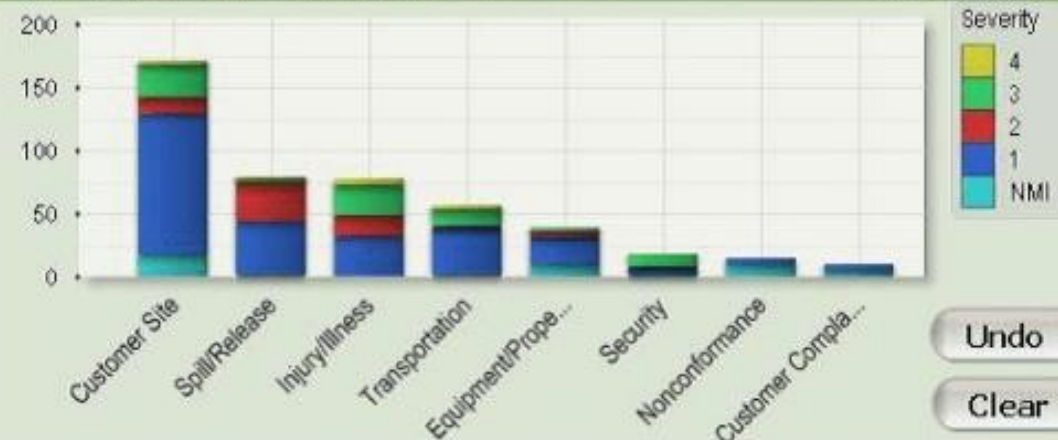
# PRE-ATTENTIVE ATTRIBUTES IN VISUALS

- *What is the story?*
- *Where do you want your audience to focus?*



## Synchronized Data Views: Drill-Down by Category, Severity, Cause, Site, Time, Status

Incidents by Category &amp; Severity



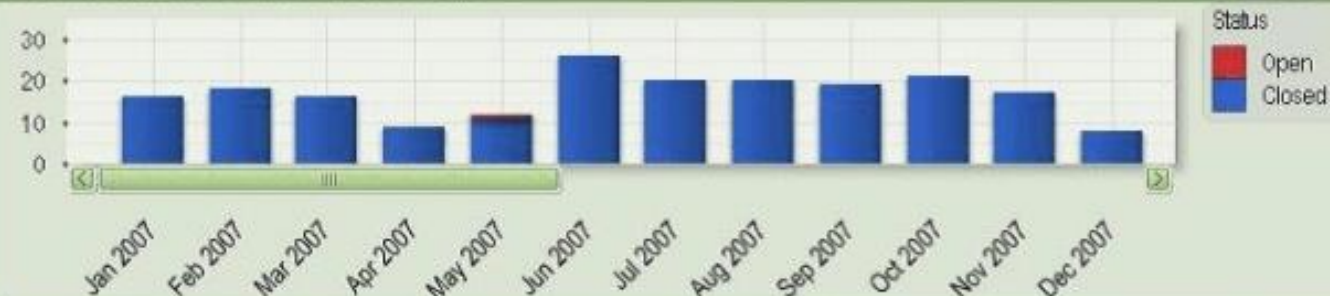
Contributing Cause



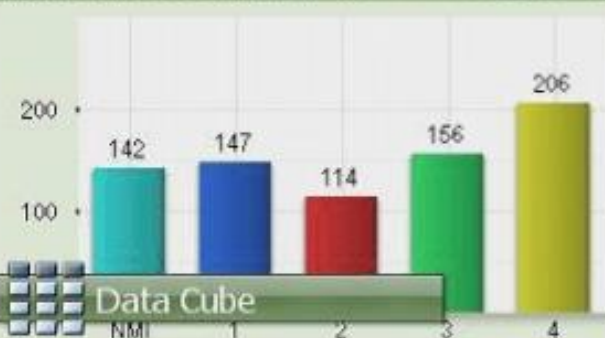
Root Cause



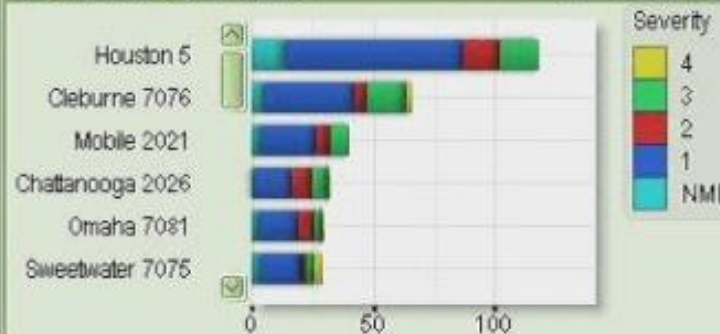
Incident Trends and their Status



Incidents Severity, by ...

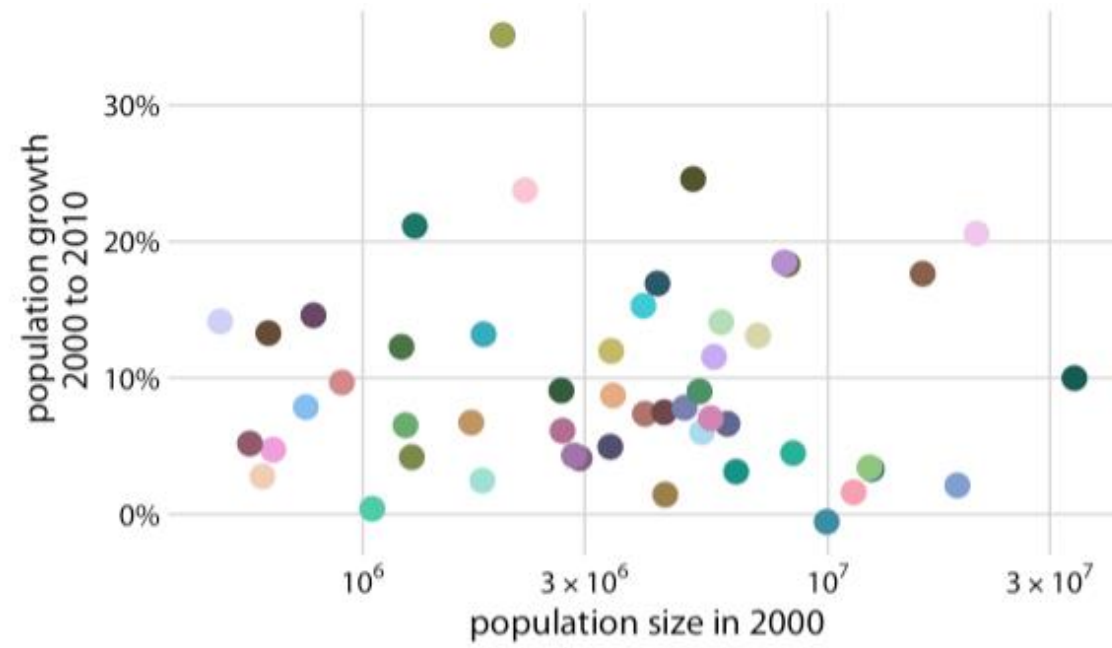


Incidents by Facility



Incidents By Categor...



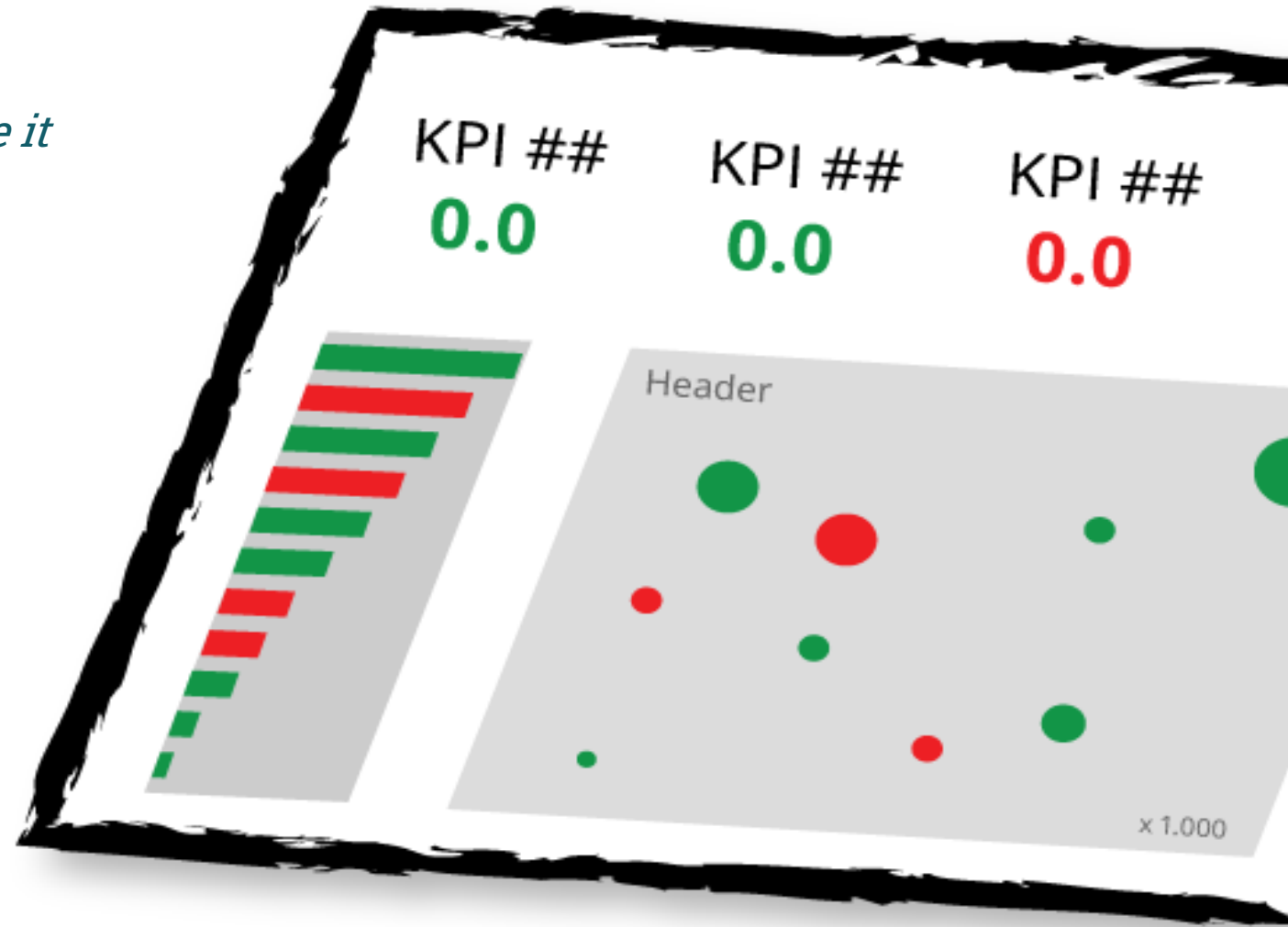


state

- |                      |                |                |
|----------------------|----------------|----------------|
| Alabama              | Kentucky       | North Dakota   |
| Alaska               | Louisiana      | Ohio           |
| Arizona              | Maine          | Oklahoma       |
| Arkansas             | Maryland       | Oregon         |
| California           | Massachusetts  | Pennsylvania   |
| Colorado             | Michigan       | Rhode Island   |
| Connecticut          | Minnesota      | South Carolina |
| Delaware             | Mississippi    | South Dakota   |
| District of Columbia | Missouri       | Tennessee      |
| Florida              | Montana        | Texas          |
| Georgia              | Nebraska       | Utah           |
| Hawaii               | Nevada         | Vermont        |
| Idaho                | New Hampshire  | Virginia       |
| Illinois             | New Jersey     | Washington     |
| Indiana              | New Mexico     | West Virginia  |
| Iowa                 | New York       | Wisconsin      |
| Kansas               | North Carolina | Wyoming        |

# USE OF COLOR IN DATA VISUALIZATION

- Resist the use of color!  
*Only use color when meaningful and use it consistently*
- Use one basic shade  
*(Grey, blue)*
- Use attention grabbing colors  
*(Green, orange and red)*
- Do not use brandcolors



# COLOR GRADIENT

## Country Level Sales Rank Top 5 Drugs

Rainbow distribution in color indicates sales rank in given country from #1 (red) to #10 or higher (dark purple)

Country	A	B	C	D	E
AUS	1	2	3	6	7
BRA	1	3	4	5	6
CAN	2	3	6	12	8
CHI	1	2	8	4	7
FRA	3	2	4	8	10
GER	3	1	6	5	4
IND	4	1	8	10	5
ITA	2	4	10	9	8
MEX	1	5	4	6	3
RUS	4	3	7	9	12
SPA	2	3	4	5	11
TUR	7	2	3	4	8
UK	1	2	3	6	7
US	1	2	4	3	5

## Top 5 drugs: country-level sales rank

RANK		1	2	3	4	5+
COUNTRY   DRUG		A	B	C	D	E
Australia		1	2	3	6	7
Brazil		1	3	4	5	6
Canada		2	3	6	12	8
China		1	2	8	4	7
France		3	2	4	8	10
Germany		3	1	6	5	4
India		4	1	8	10	5
Italy		2	4	10	9	8
Mexico		1	5	4	6	3
Russia		4	3	7	9	12
Spain		2	3	4	5	11
Turkey		7	2	3	4	8
United Kingdom		1	2	3	6	7
United States		1	2	4	3	5



# ASSIGNMENT: CREATE A DATA STORY

- Go to <https://www.e-mergo.nl/erd/>
- Use the provided data set
- Identify your intended user
- Come up with one KPI for your user
- Come up with a way of visualizing your KPI
- Present your KPI in a 3-minute pitch



# E-MERGING TALENT PROGRAM

Almost graduating? Interested in data? Then the E-merging Talent Program is just what you're looking for!

During this program you will be trained into a data expert within 2 years.

**The entire training path has a total value of over €45.000,-**



#### Data Analytics techniques

- Advanced Analytics basics
- Data modelling
- Data visualisation
- KPI's and metrics
- Datawarehouse automation
- Infrastructure networks
- Process modelling



#### Tooling

- Power BI
- Qlik
- TimeXtender
- Azure



#### Soft skills

- Advisory skills
- Presentation skills
- Communicate effectively
- Agile methodology



**Scan the QR-code  
for more info!**

Interested? Register via the link or reach out to our recruiter Josita on LinkedIn!



emergo

# CONTACT

Elektronicaweg 16a  
2628 XG Delft

085 016 04 11  
[info@e-mergo.nl](mailto:info@e-mergo.nl)  
[www.e-mergo.nl](http://www.e-mergo.nl)

