

# Big data, Big insights

*Jeroen van der Leijé*

*[g.j.van.der.leije@liacs.leidenuniv.nl](mailto:g.j.van.der.leije@liacs.leidenuniv.nl)*

Leiden Institute of Advanced Computer Science (LIACS)



**Universiteit  
Leiden**  
The Netherlands

Discover the world at Leiden University

<https://www.youtube.com/watch?v=HLUac2JAtHI>

# Big data vs “Good data”

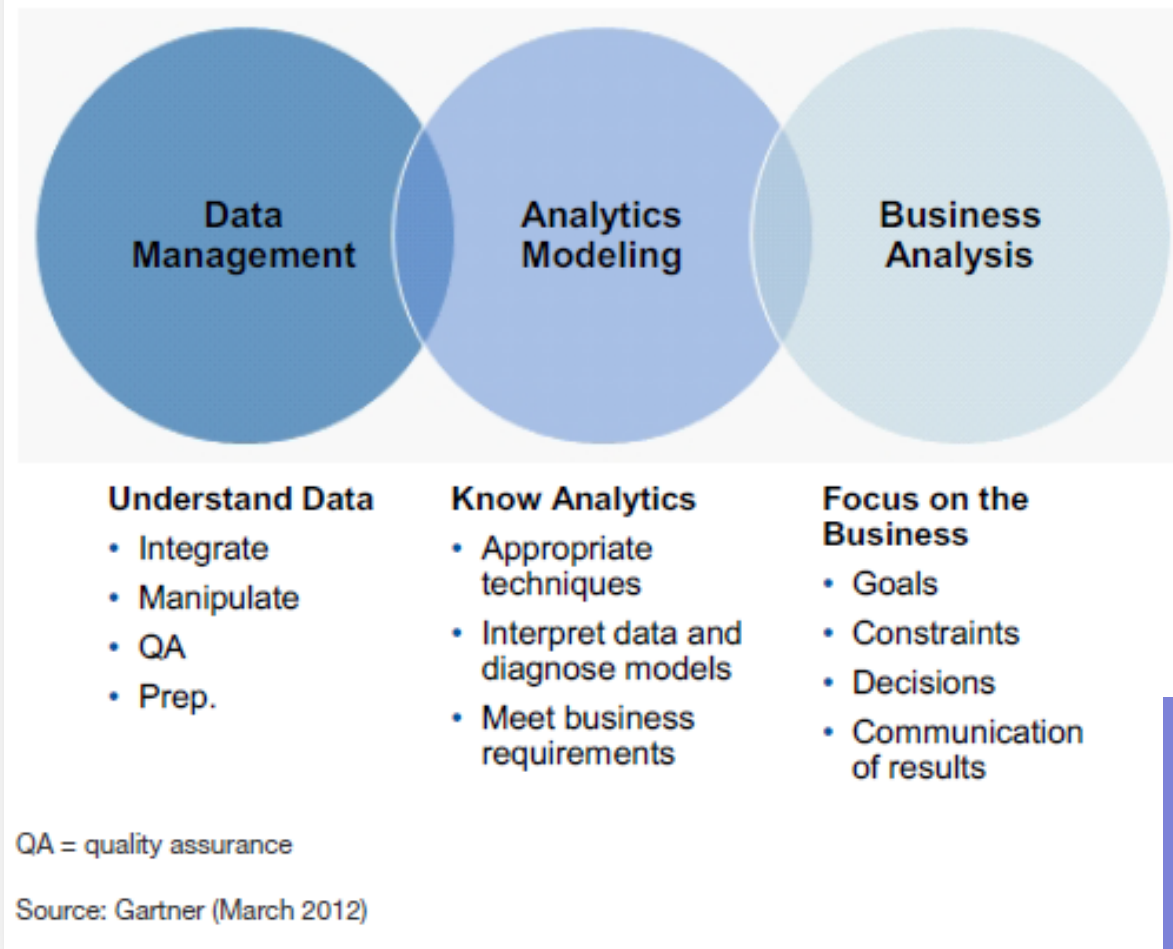
- Big Data gaat niet primair over de grootte van bestanden. Het gaat over de kenmerken van de data.
- “Good Data”: Als de data in een relationele database zit, dan is het “accuraat, opgeschoond en 100% betrouwbaar...”
- Big Data is slecht gedefinieerd, waarschijnlijk inaccuraat en incompleet.
- Maar door het volume van big data worden deze beperkingen statistisch insignificant en verwaarloosbaar.
- Als algoritmes dan functioneel blijken bij de analyses: Big Data

# Big data - Computational turn

- Van causaliteit naar correlatie
- Het inzicht in causale relaties wordt minder belangrijk
- Correlatie (wat werkt wel en wat werkt niet) heeft de prioriteit gekregen: Computational Turn
- Computational Turn biedt mogelijkheden voor een totaal andere aanpak in gezondheid, economie, techniek, recht, gedragswetenschappen, etc.

# Behoefte aan data scientists

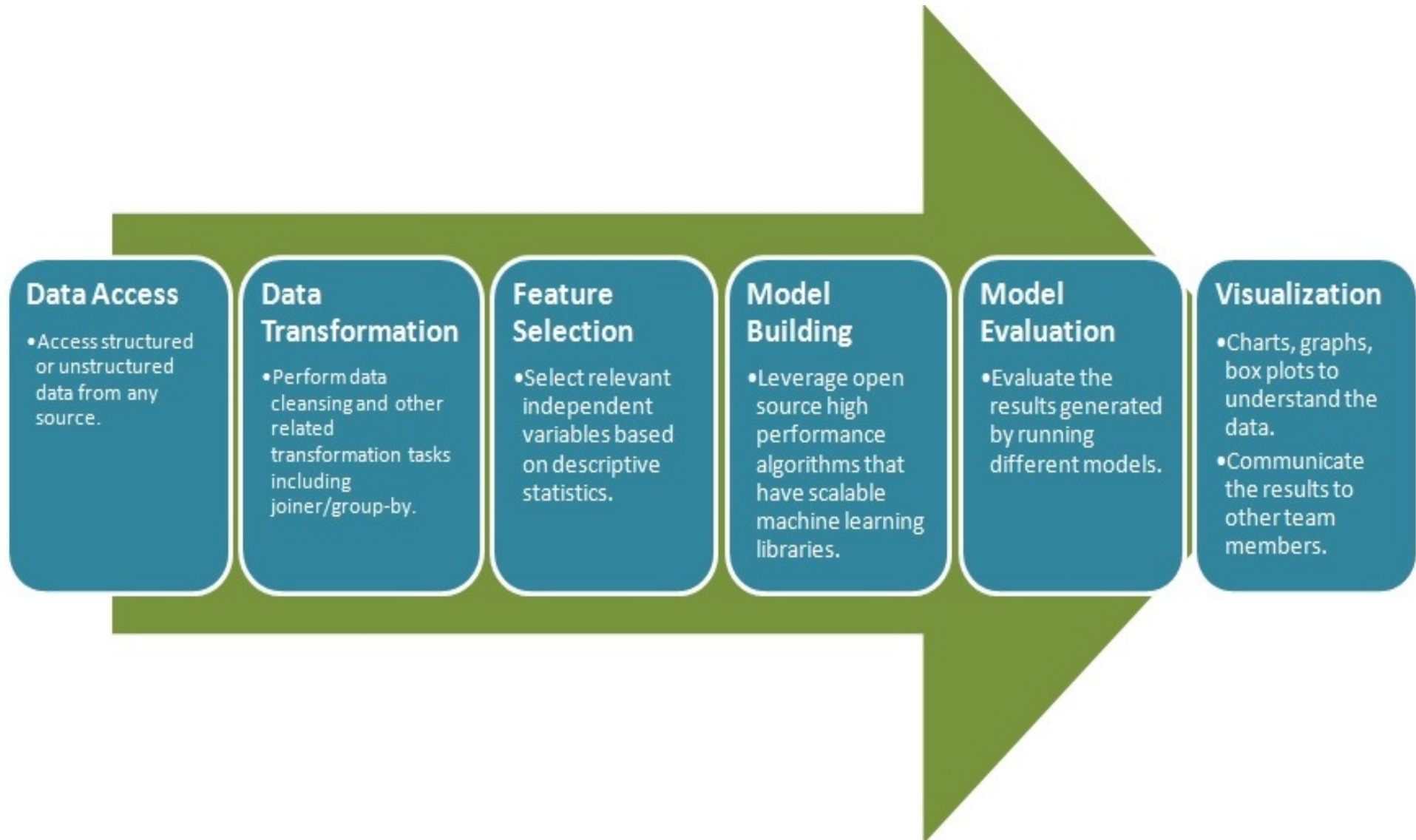
Figure 3. Core Data Scientist Skills



The data scientist has emerged as a new role, distinct from – but with similarities to – those of business intelligence (BI) analysts and statisticians.

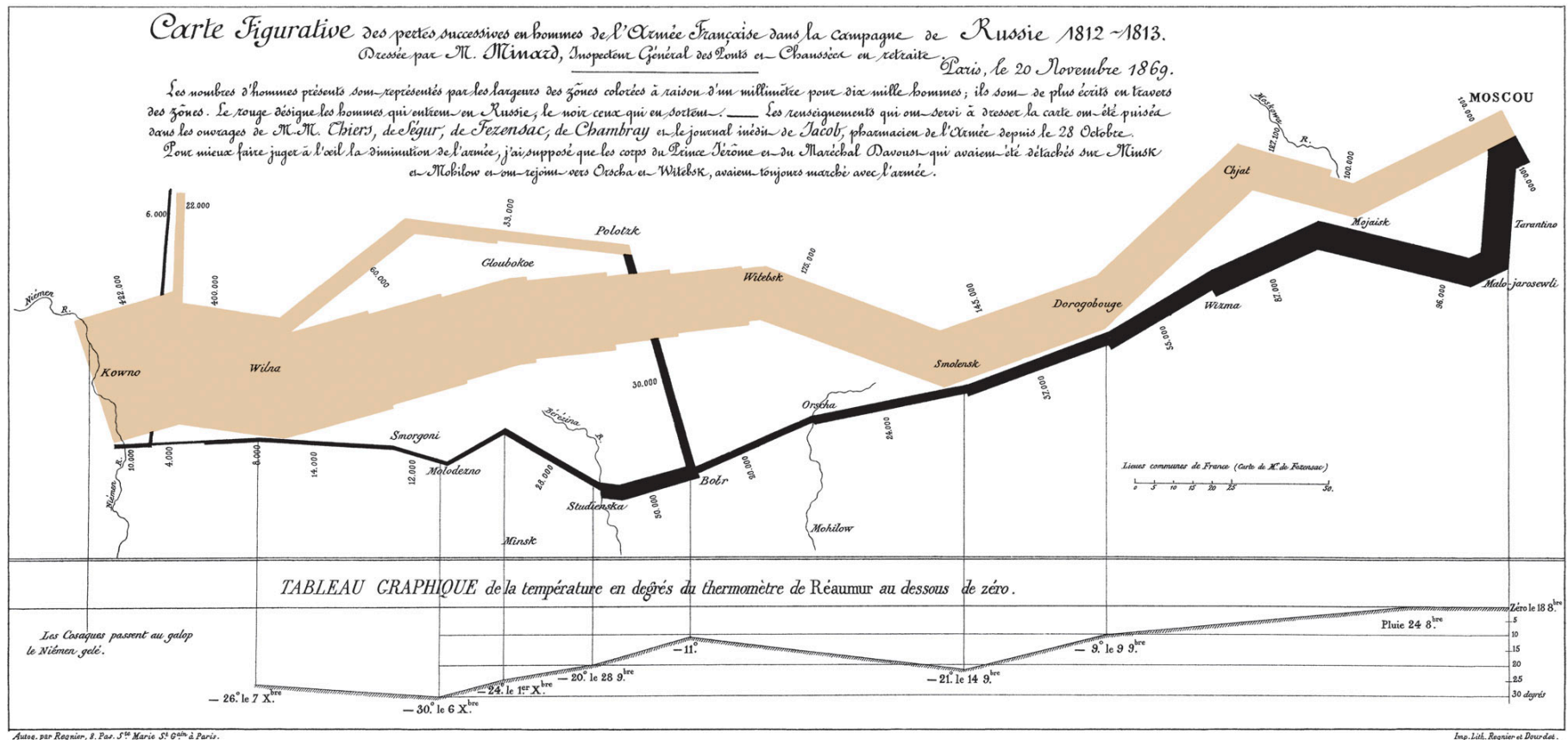
Source: Gartner, Emerging Role of the Data Scientist and the Art of Data Science, Mar. 20, 2012.

# Behoefte aan ketenbenadering



# Behoefte aan visualisaties 1

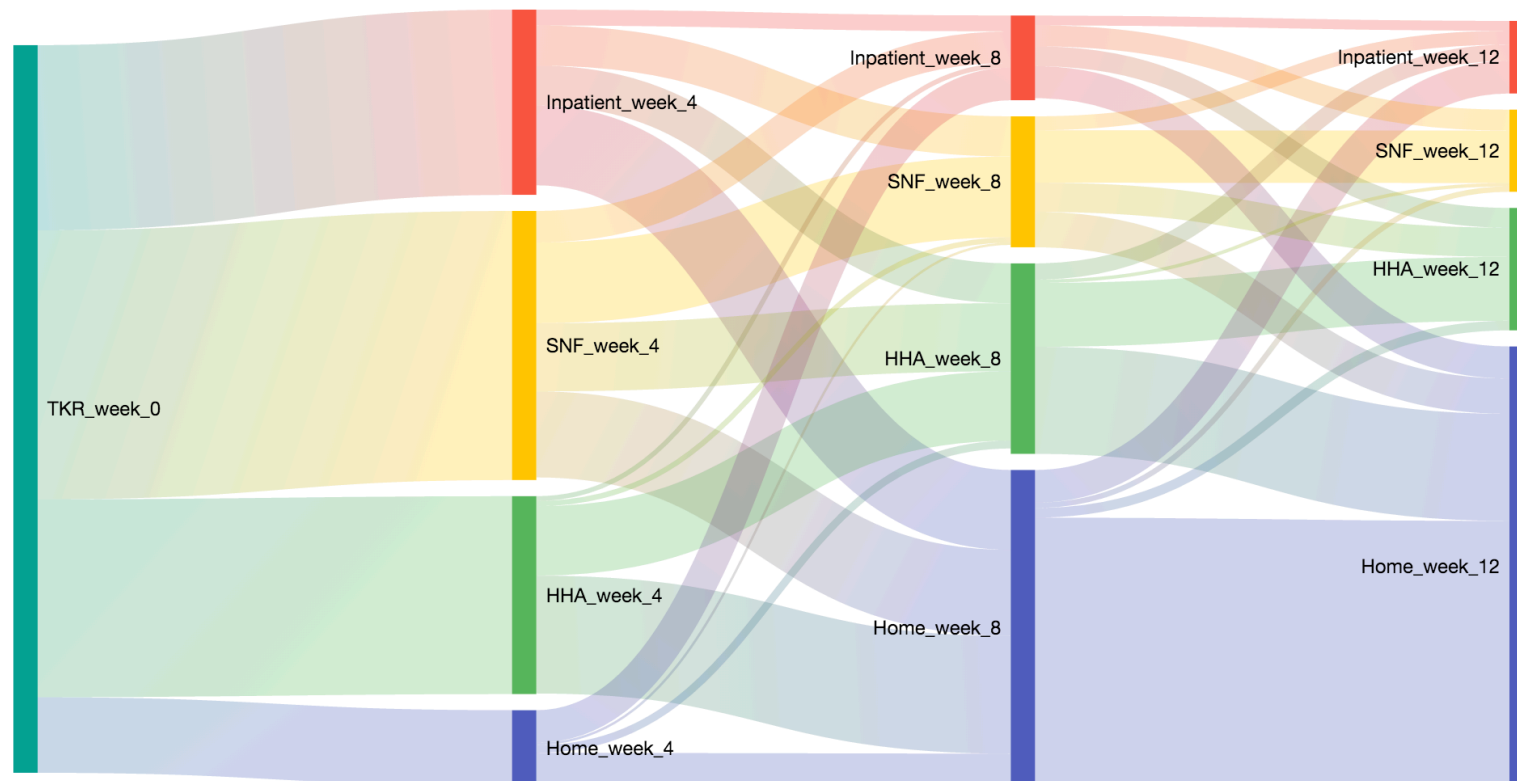
Charles Minard's overzicht van Napoleon's Russische veldslagen in 1812:



Weergave in 2 dimensies van 6 typen data: aantal manschappen, afstand, temperatuur, locatie, richting, datum.

# Behoefte aan visualisaties 2

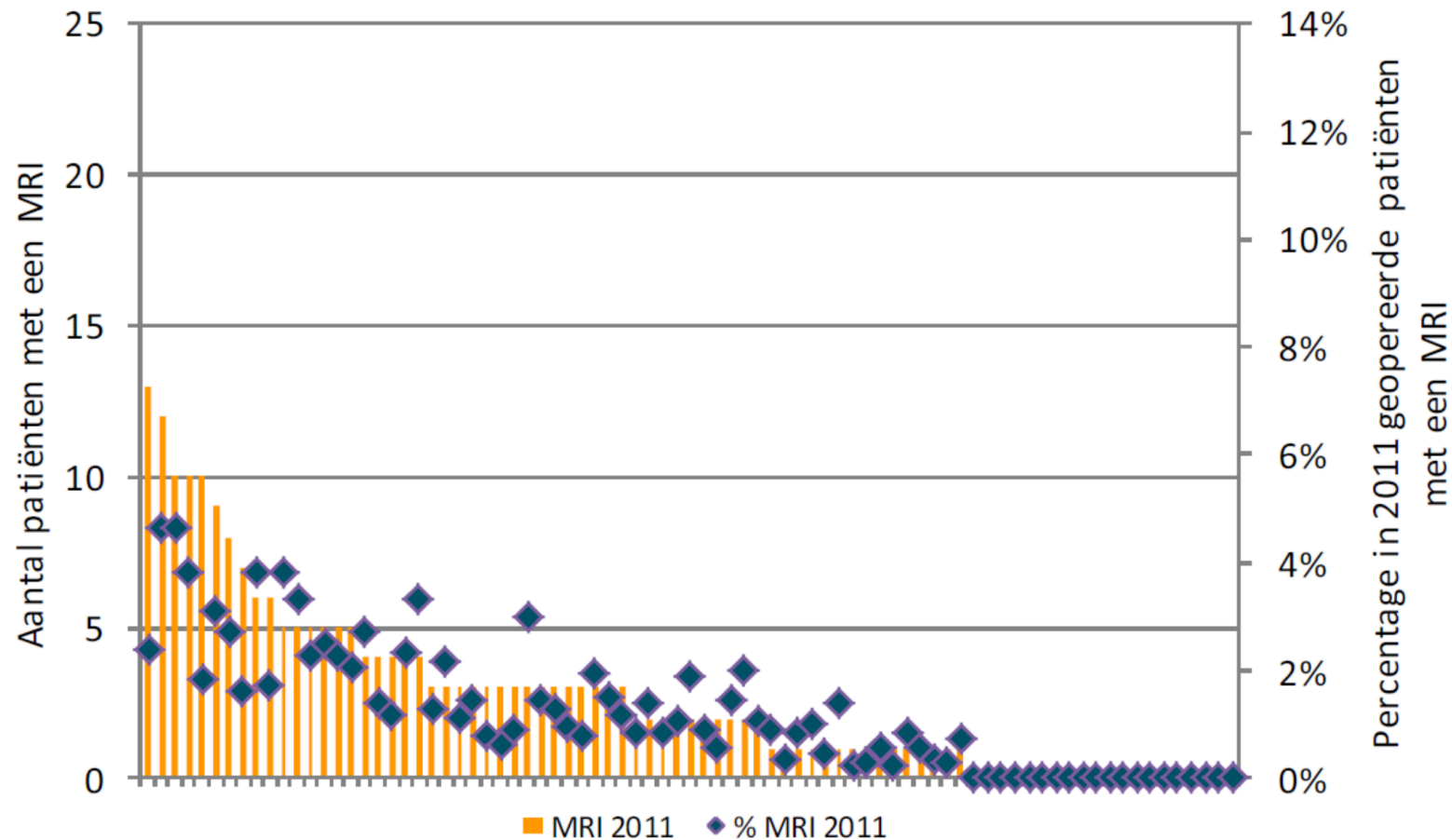
Charles Minard's map legde de basis voor de (Matthew) Sankey diagrammen:



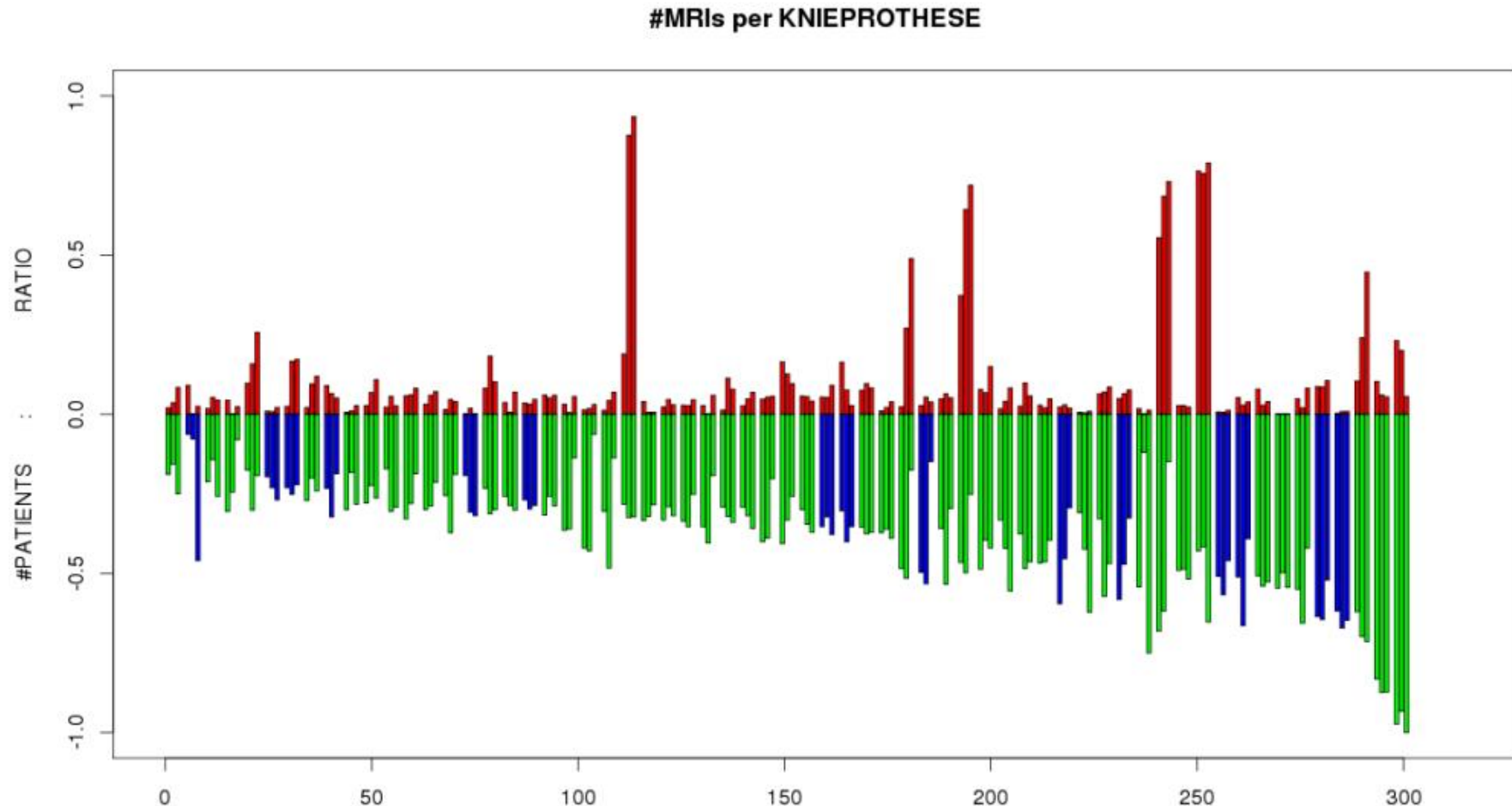
Total Knee Replacement, Skilled Nursing Facilities, Home Health Agencies  
(bron: <http://blog.accordionhealth.com/> )



# 1e, 2e, n<sup>de</sup> orde visualisaties




# 1e, 2e, n<sup>de</sup> orde visualisaties



# Voorstellingsvermogen & visie

- Science fiction: startpunt voor toekomst
- Latente behoefte: laat me iets zien, dan weet ik wat ik niet wil...
- Rol van beleidsmakers: think beyond
- Big Data analytics keten: verzamelen, opschonen, opslaan en toegankelijk maken



SUB6\_BRKRM\_04  
15:08:01

EVALUATING STRATEGIES...

**Contact:** *Jeroen van der Leijé, Business Liaison*

| +31 (0) 624 416 805 | [g.j.van.der.leije@liacs.leidenuniv.nl](mailto:g.j.van.der.leije@liacs.leidenuniv.nl) | [liacs.leidenuniv.nl/business](http://liacs.leidenuniv.nl/business) |

Discover the world at Leiden University

