

COURSEWARE

# EDF

## Data Literacy Professional Courseware

Auteur: Michel Dekker

EDF Data Literacy  
Professional Courseware

## Colophon

Title: EDF Data Literacy Professional Courseware

Authors: Michel Dekker

Publisher: Van Haren Publishing, 's-Hertogenbosch

ISBN Hard Copy: 978 94 018 0985 6

Edition: First edition, first print, March 2023

Design: Van Haren Publishing, 's-Hertogenbosch

Copyright: © Van Haren Publishing 2023

For further information about Van Haren Publishing please e-mail us at: [info@vanharen.net](mailto:info@vanharen.net) or visit our website: [www.vanharen.net](http://www.vanharen.net)

No part of this publication may be reproduced in any form by print, photo print, microfilm or any other means without written permission by the publisher.

Although this publication has been composed with much care, neither author, nor editor, nor publisher can accept any liability for damage caused by possible errors and/or incompleteness in this publication.

## **Publisher about the Courseware**

The Courseware was created by experts from the industry who served as the author(s) for this publication. The input for the material is based on existing publications and the experience and expertise of the author(s). The material has been revised by trainers who also have experience working with the material. Close attention was also paid to the key learning points to ensure what needs to be mastered.

The objective of the courseware is to provide maximum support to the trainer and to the student, during his or her training. The material has a modular structure and according to the author(s) has the highest success rate should the student opt for examination. The Courseware is also accredited for this reason, wherever applicable.

In order to satisfy the requirements for accreditation the material must meet certain quality standards. The structure, the use of certain terms, diagrams and references are all part of this accreditation. Additionally, the material must be made available to each student in order to obtain full accreditation. To optimally support the trainer and the participant of the training assignments, practice exams and results are provided with the material.

Direct reference to advised literature is also regularly covered in the sheets so that students can find additional information concerning a particular topic. The decision to leave out notes pages from the Courseware was to encourage students to take notes throughout the material.

Although the courseware is complete, the possibility that the trainer deviates from the structure of the sheets or chooses to not refer to all the sheets or commands does exist. The student always has the possibility to cover these topics and go through them on their own time. It is recommended to follow the structure of the courseware and publications for maximum exam preparation.

The courseware and the recommended literature are the perfect combination to learn and understand the theory.

-- Van Haren Publishing

## Other publications by Van Haren Publishing

Van Haren Publishing (VHP) specializes in titles on Best Practices, methods and standards within four domains:

- IT and IT Management
- Architecture (Enterprise and IT)
- Business Management and
- Project Management

Van Haren Publishing is also publishing on behalf of leading organizations and companies: ASLBiSL Foundation, BRMI, CA, Centre Henri Tudor, Gaming Works, IACCM, IAOP, IFDC, Innovation Value Institute, IPMA-NL, ITSqc, NAF, KNVI, PMI-NL, PON, The Open Group, The SOX Institute.

Topics are (per domain):

### IT and IT Management

ABC of ICT  
ASL®  
CATS CM®  
CMMI®  
COBIT®  
e-CF  
ISO/IEC 20000  
ISO/IEC 27001/27002  
ISPL  
IT4IT®  
IT-CMF™  
IT Service CMM  
ITIL®  
MOF  
MSF  
SABSA  
SAF  
SIAM™  
TRIM  
VeriSM™

### Enterprise Architecture

ArchiMate®  
GEA®  
Novius Architectuur  
Methode  
TOGAF®

### Business Management

*BABOK® Guide*  
BiSL® and BiSL® Next  
BRMBOK™  
BTF  
EFQM  
eSCM  
IACCM  
ISA-95  
ISO 9000/9001  
OPBOK  
SixSigma  
SOX  
SqEME®

### Project Management

A4-Projectmanagement  
DSDM/Atern  
ICB / NCB  
ISO 21500  
MINCE®  
M\_o\_R®  
MSP®  
P3O®  
*PMBOK® Guide*  
Praxis®  
PRINCE2®

For the latest information on VHP publications, visit our website: [www.vanharen.net](http://www.vanharen.net).

## Table of content

	<i>--- Slide number</i>	<i>--- Page number</i>
Reflection		6
Agenda		8
<b>Effective Data Foundation</b>	(2)	11
<b>Introduction</b>	(7)	14
<b>Read Data</b>	(13)	17
What is Data?	(14)	17
Summarize Data	(21)	21
Consume Data	(28)	24
Check your Data	(48)	34
<b>Work with Data</b>	(58)	39
Creating Data	(59)	40
Data Quality	(66)	43
Acquiring & Cleaning	(72)	46
Managing	(94)	57
<b>Analyze Data</b>	(108)	64
Expectations	(111)	66
Thinking shortcuts	(115)	68
Types of Analysis	(135)	78
Analytical skills	(151)	86
<b>Argue with Data</b>	(175)	98
Explore to Explain	(177)	99
CHRTTS	(182)	101
Storytelling	(190)	104
<b>Practice exam info</b>		120
<b>Syllabus</b>		121

## Self-Reflection of understanding Diagram

*‘What you do not measure, you cannot control.’ – Tom Peters*

Fill in this diagram to self-evaluate your understanding of the material. This is an evaluation of how well you know the material and how well you understand it. In order to pass the exam successfully you should be aiming to reach the higher end of Level 3. If you really want to become a pro, then you should be aiming for Level 4. Your overall level of understanding will naturally follow the learning curve. So, it’s important to keep track of where you are at each point of the training and address any areas of difficulty.

Based on where you are within the Self-Reflection of Understanding diagram you can evaluate the progress of your own training.

<i>Level of Understanding</i>	<i>Before Training (Pre-knowledge)</i>	<i>Training Part 1 (1st Half)</i>	<i>Training Part 2 (2nd Half)</i>	<i>After studying / reading the book</i>	<i>After exercises and the Practice exam</i>
<i>Level 4 I can explain the content and apply it .</i>					
<i>Level 3 I get it! I am right where I am supposed to be.</i>					<i>Ready for the exam!</i>
<i>Level 2 I almost have it but could use more practice.</i>					
<i>Level 1 I am learning but don't quite get it yet.</i>					

(Self-Reflection of Understanding Diagram)

Write down the problem areas that you are still having difficulty with so that you can consolidate them yourself, or with your trainer. After you have had a look at these, then you should evaluate to see if you now have a better understanding of where you actually are on the learning curve.

### Troubleshooting

*Problem areas:*

*Topic:*

---

Part 1

---

---

---

---

---

---

Part 2

---

---

---

---

---

---

You have gone  
through the book  
and studied.

---

---

---

---

You have answered  
the questions and  
done the practice  
exam.

---

---

---

---

---

## Timetable

Start	Duration	Subject	Start slide
09:00	00:15	Intro & Agenda	1
09:15	00:30	What is data?	14
09:45	00:15	Summarize data	21
10:00	00:25	Consume Data	28
10:25	00:20	Chart Types	38
10:45	00:10	BREAK	
10:55	00:25	Check your Data - The Fork	48
11:20	00:40	Check your Data - exercise	52
12:00	00:10	Wrap up	56
12:10	00:50	LUNCH	
13:00	00:10	Intro & Agenda	58
13:10	00:15	Creating	59
13:25	00:20	Data Quality	66
13:45	00:10	BREAK	
13:55	00:55	Aquire & Clean	72
14:50	00:50	Join	84
15:40	00:10	Managing	94
15:50	00:30	KPI Mantras	97
16:20	00:10	Wrap up	106

<b>Start</b>	<b>Duration</b>	<b>Subject</b>	<b>Start slide</b>
09:00	00:15	Intro & Agenda	108
09:15	00:20	Expectations	111
09:35	00:45	Shortcuts	115
10:20	00:40	Data Analysis	135
11:00	00:10	BREAK	
11:10	00:40	Analytical skills	151
11:50	00:10	Wrap up	173
12:00	01:00	LUNCH	
13:00	00:15	Intro & Agenda	175
13:15	00:40	Explore to Explain	177
13:55	00:55	CHRTTS	182
14:50	00:20	Storytelling ARC	190
15:10	00:10	BREAK	
15:20	00:35	Storytelling principles	193
15:55	00:25	Storytelling exercise	212
16:20	00:10	Wrap up	214



# Data Literacy

how to read, work with, analyze & argue with data



COURSEWARE



©2023 - All training materials are sole property of Van Haren Publishing BV and are not to be reproduced in any form or shape without written permission.

## Effective Data Foundation

not-for-profit collective,  
who enables **professionals** to  
**leverage data** to make  
**sustainable business**  
**decisions**



2

©2023 Van Haren Publishing BV.

# DATA

**Analysis / Literacy / Management / Visualization**


3

©2023 Van Haren Publishing BV.



## Certification

Effective Data Foundation



Congratulations to  
**Marly Buitenhuis**  
for achieving the  
**Data Visualization**  
certificate from the Effective Data Foundation.

4

©2023 Van Haren Publishing BV.



## Certification



**60** multiple choice questions

within **60** minutes

at least **65%** correct to pass

5

©2023 Van Haren Publishing BV.



## Training structure



**Read  
data**

**25%**



**Work with  
data**

**25%**



**Analyze  
data**

**25%**



**Argue with  
data**

**25%**

6

©2023 Van Haren Publishing BV.



## Data driven...



## Data literacy

Employees recognizing data in the workplace as an **asset**



Employees **trust** their decisions more when based on data



Employees are **confident** in their data literacy skills



## Introduction

*Data in the hands of a few data experts can be powerful but data at the fingertips of **many is truly transformational***

Brent Dykes - author of Effective Data Storytelling

## Data literacy

The ability to **read**, **work** with, **analyze**, and **argue** with data

Source: Raul Bhargava and Catherine D'Ignazio from MIT and Emerson College



**Read**  
data



**Work**  
with data



**Analyze**  
data



**Argue**  
with data

# The Data Journey

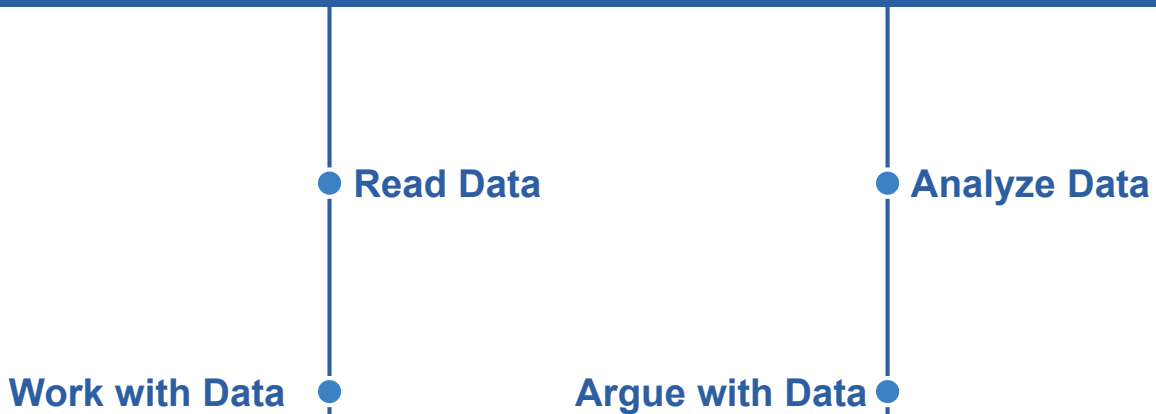


11

©2023 Van Haren Publishing BV.



# Training Agenda



12

©2023 Van Haren Publishing BV.



## Training Agenda



**Read  
data**

**What is Data?**

**Summarize Data**

**Consume Data**

**Check your Data**

13

©2023 Van Haren Publishing BV.



## What is Data?



Data has a better idea

14

©2023 Van Haren Publishing BV.



## Data is not Reality



15

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## What is Data?

Describes a **quality** or **quantity**  
of some **object** or **event**.

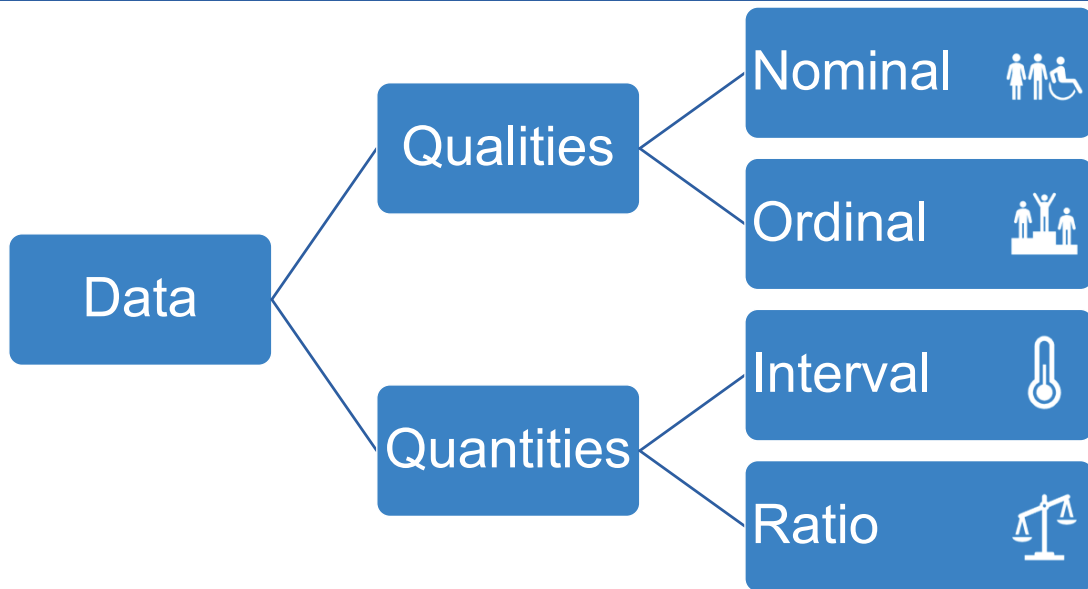
Source: unknown

16

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Quantity or Quality



17

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Types of scales

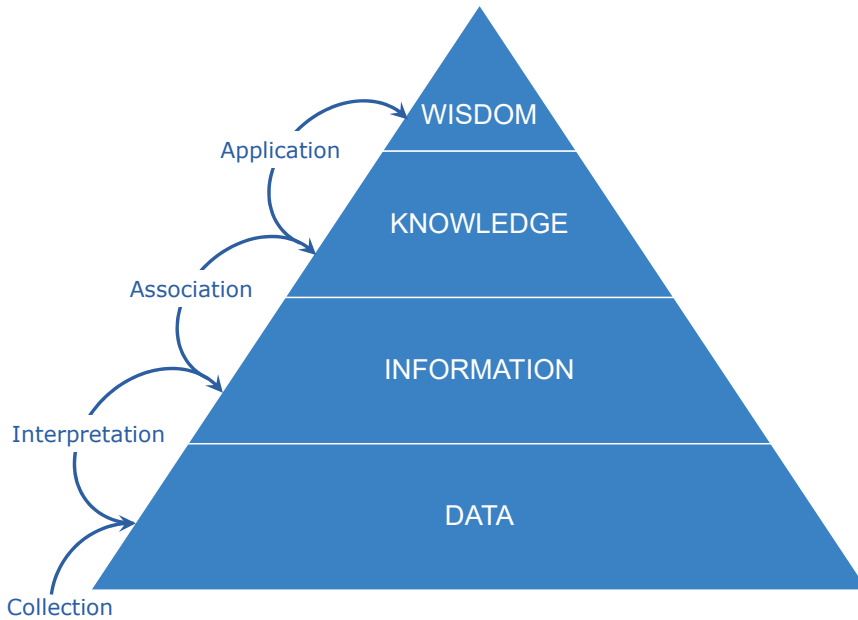


18

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Why do we use Data?

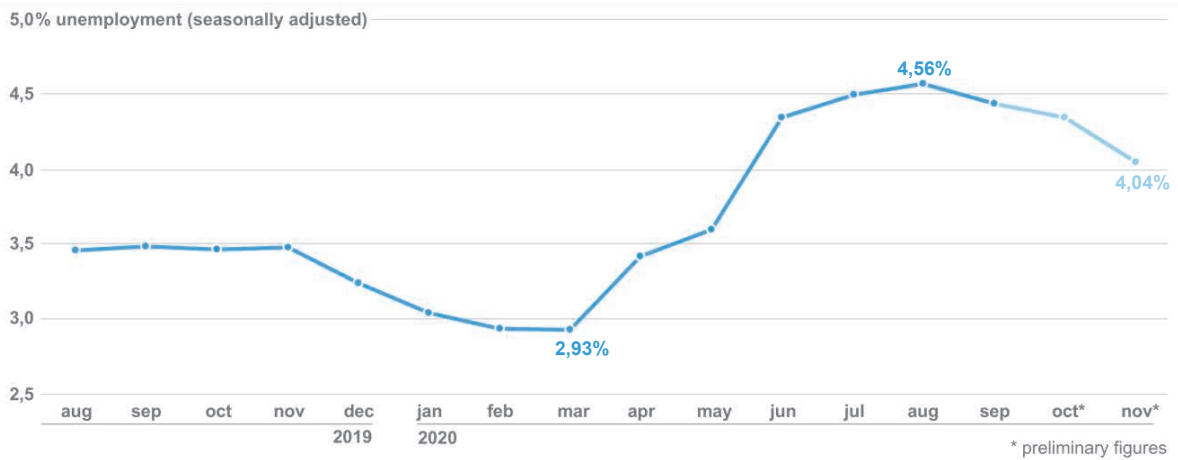


19 Source: Jones, Ben, Data Literacy Fundamentals: Understanding the Power & Value of Data

©2023 Van Haren Publishing BV.



## Unemployment



$$\text{Unemployment rate} = \frac{\text{Unemployed labor force}}{\text{Labor force}}$$

20 Source: CBS – Statistics Netherlands

©2023 Van Haren Publishing BV.



## Summarize data

*I didn't come here  
to be average*

Source: Michael Jordan



21

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Summarize data



 **6,690 +3**  
inhabitants

 **16.9 km<sup>2</sup>**  
surface

 **18.5K\$**  
average income

 **37.4K\$**  
average income



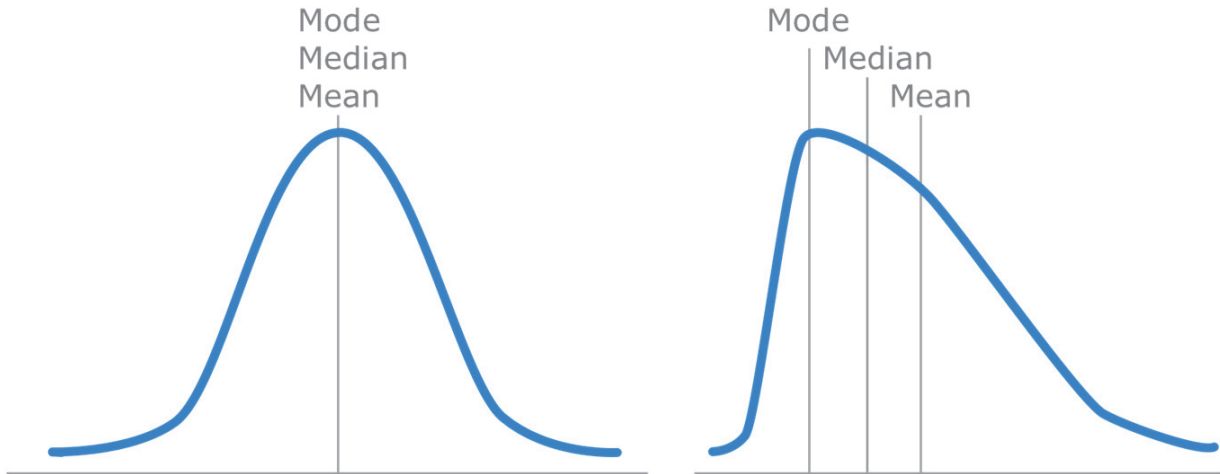
22

Source: [https://en.wikipedia.org/wiki/List\\_of\\_United\\_States\\_counties\\_by\\_per\\_capita\\_income](https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income) & <https://www.sportico.com/personalities/sb/lelew/2022/100-highest-paid-athletes-in-the-world-2022-full-list-1234674901/>

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Summarize data



23

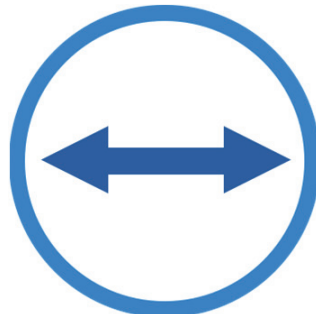
©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Summarize data



**Central  
Tendency**



**Spread**



**Shape**

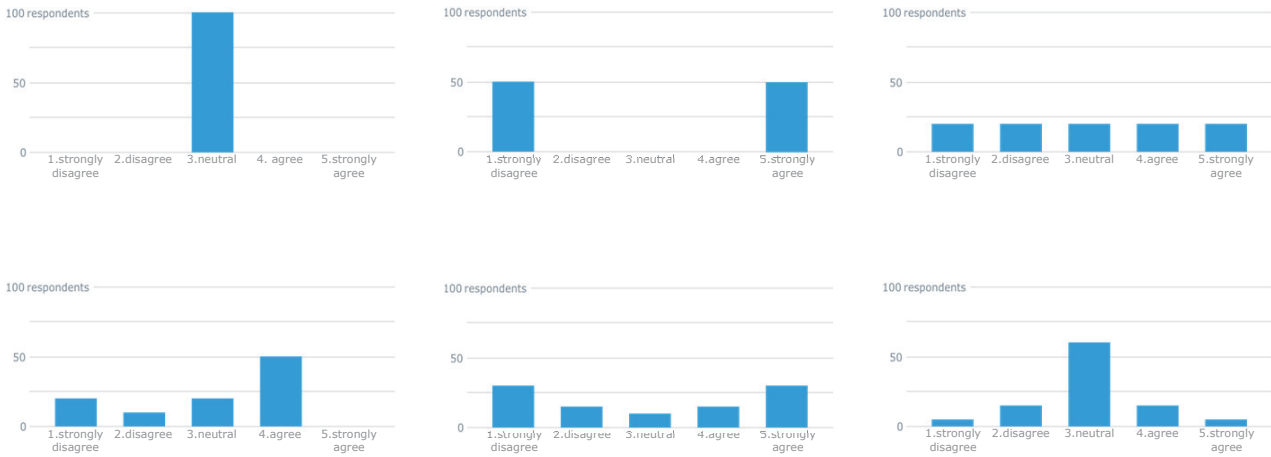
24 Source: Stephen S. Few, Now you see it, 2009

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Summarize data – Likert scale

Mean = 3.0

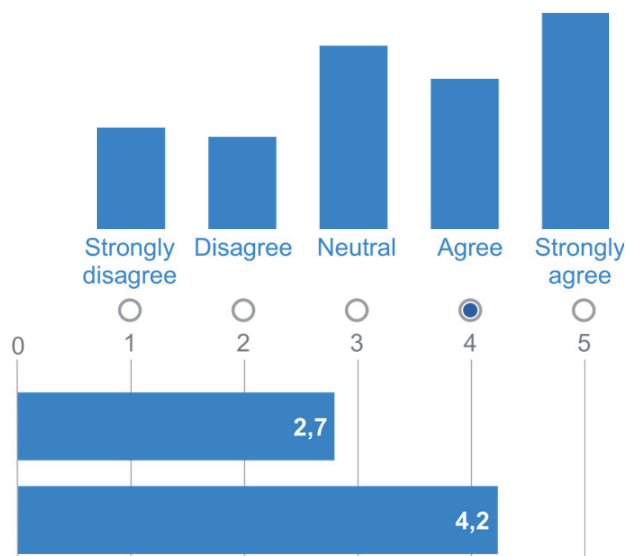


25

©2023 Van Haren Publishing BV.



## Summarize data – Likert scale

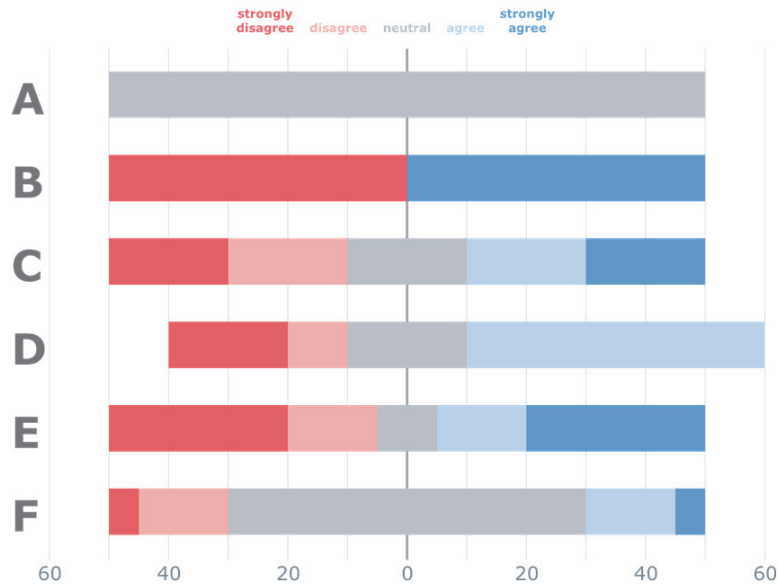


26

©2023 Van Haren Publishing BV.



## Summarize data – Likert scale



27

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Consume data



28

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Consume data



***"We used to use intuition; now we use analytics."***

2014 IBM ad\*

\* and almost everyone else in BI/AI, at some point or another

Effective  
DATA  
Foundation

29

©2023 Van Haren Publishing BV.

## Consume data



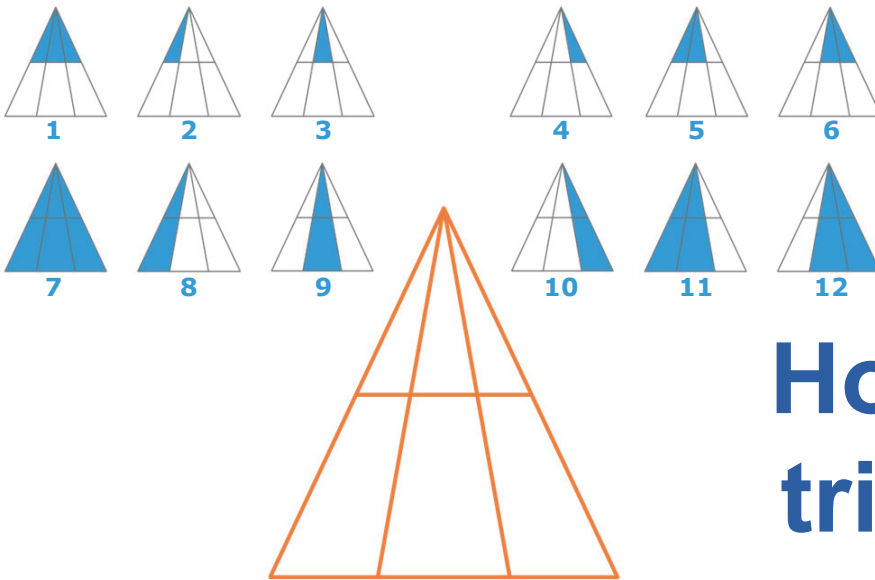
**How does she feel?**

Effective  
DATA  
Foundation

30 Photo by nappy op Pexels.com

©2023 Van Haren Publishing BV.

## Consume data



# How many triangles?

31 Source: Jones, Ben, Data Literacy Fundamentals: Understanding the Power & Value of Data

©2023 Van Haren Publishing BV.



## Thinking, Fast and Slow

### System 1

Fast

Unconscious

Automated

Everyday decisions

Error prone

### System 2

Slow

Conscious

Effortful

Complex decisions

Reliable

32 Source: Daniel Kahneman - 2011 - Thinking, Fast and Slow

©2023 Van Haren Publishing BV.



## Consume Data

5'2"

Height Napoleon

~~157cm~~  
Height Napoleon

168cm

Height Napoleon

165cm

Avg height



33 Source: James Gillray - 1805

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Consume data

**"We used to use  
intuition;  
then we used  
analytics.**

**Now we use both"**

Ben Jones

34

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## BIG numbers



Every hour: **55 million** PET bottles sold in the world

35 Source: <https://graphics.reuters.com/ENVIRONMENT-PLASTIC/0100B275155/index.html>

©2023 Van Haren Publishing BV.



## BIG numbers



**1.3 billion**  
bottles every  
day

36 Source: <https://graphics.reuters.com/ENVIRONMENT-PLASTIC/0100B275155/index.html>

©2023 Van Haren Publishing BV.



## How solid is your opinion?

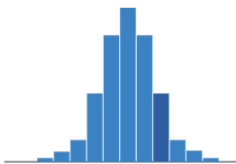
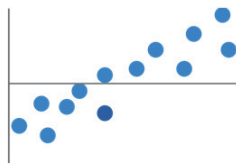
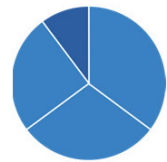


37 Source: Ben Jones

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Read carefully

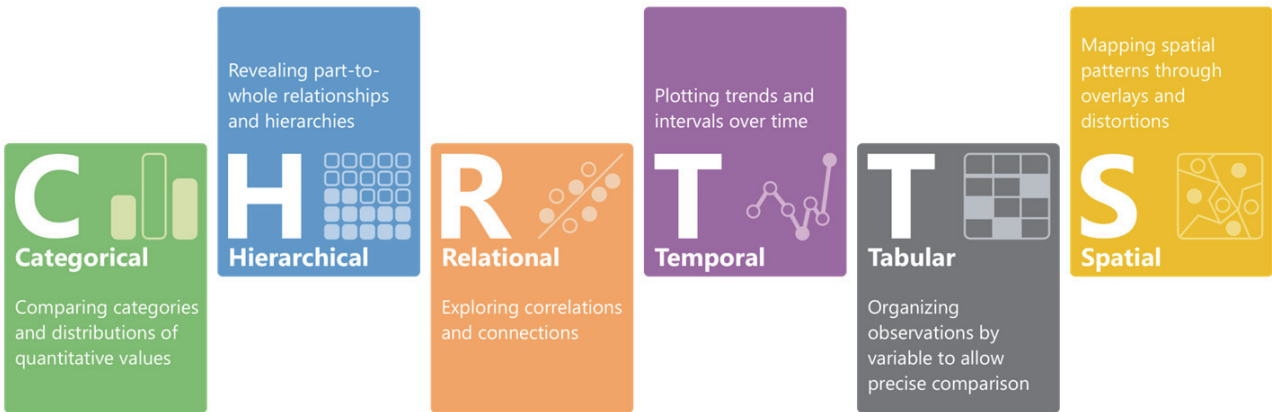


38 Source: Stephen S. Few, Show me the Numbers, 2004

©2023 Van Haren Publishing BV.

Effective  
DATA  
Foundation

## Read carefully

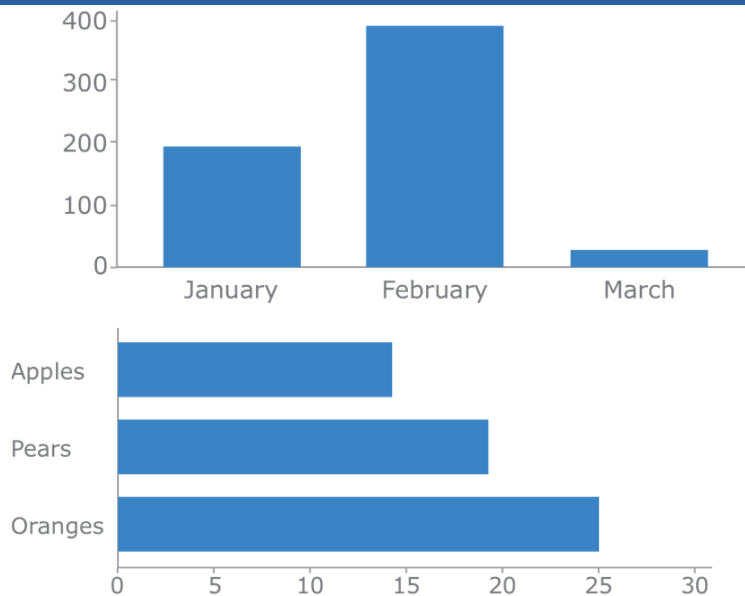


39 Based on: Data Visualization, Andy Kirk

©2023 Van Haren Publishing BV.



## Read carefully



40

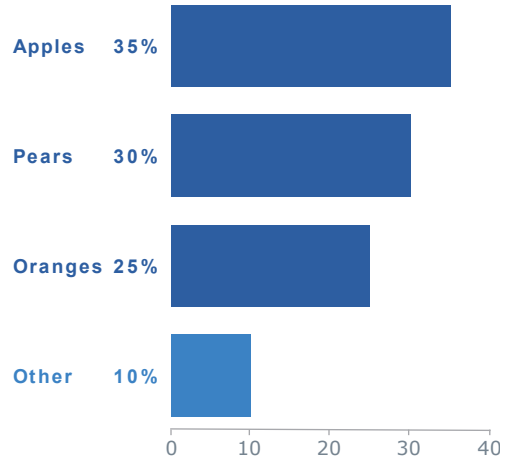
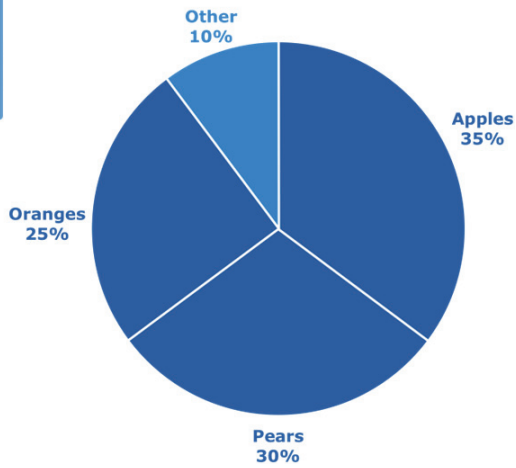
©2023 Van Haren Publishing BV.



# Read carefully

Revealing part-to-whole relationships and hierarchies

**H** Hierarchical



41

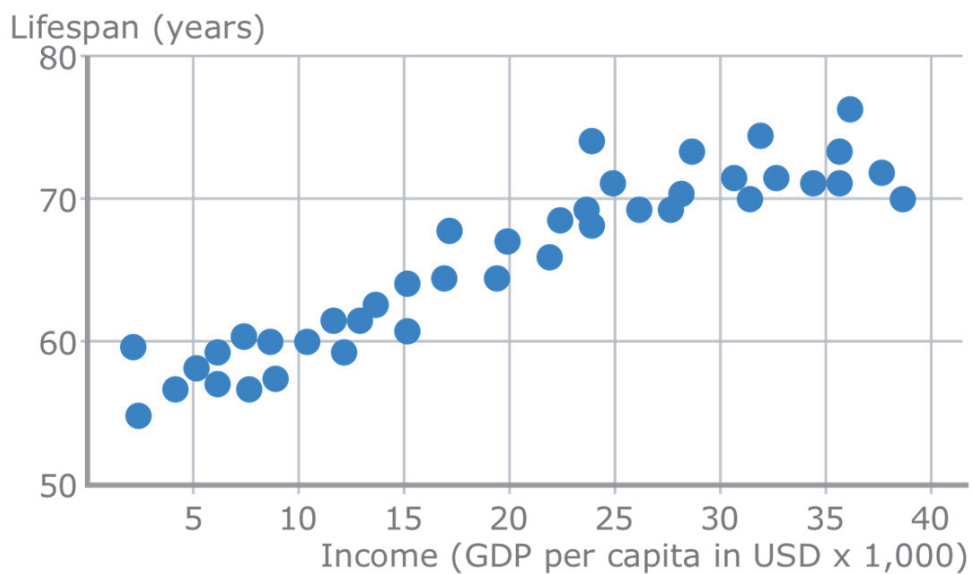
©2023 Van Haren Publishing BV.

Effective DATA Foundation

# Read carefully

**R** Relational

Exploring correlations and connections



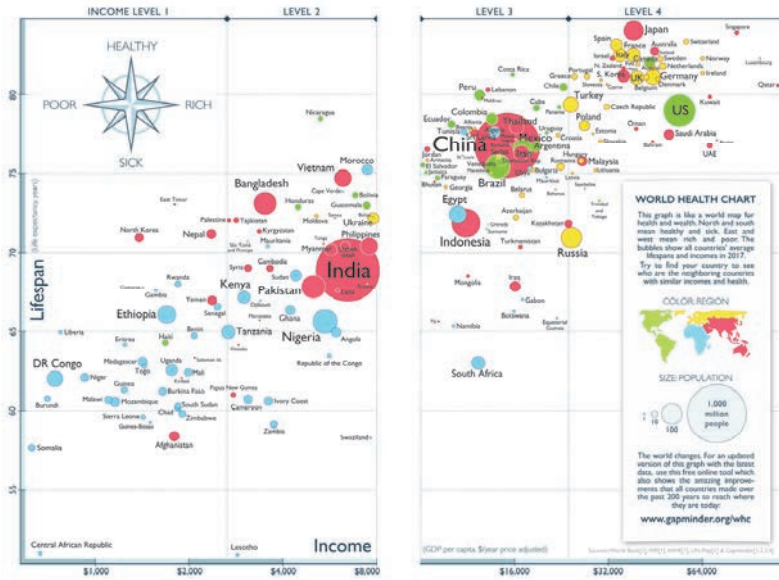
42

©2023 Van Haren Publishing BV.

Effective DATA Foundation

# Read carefully

**R**  
Relational  
Exploring correlations and connections



43 Source: Factfulness, Hans. Ola & Anna Rosling, 2018

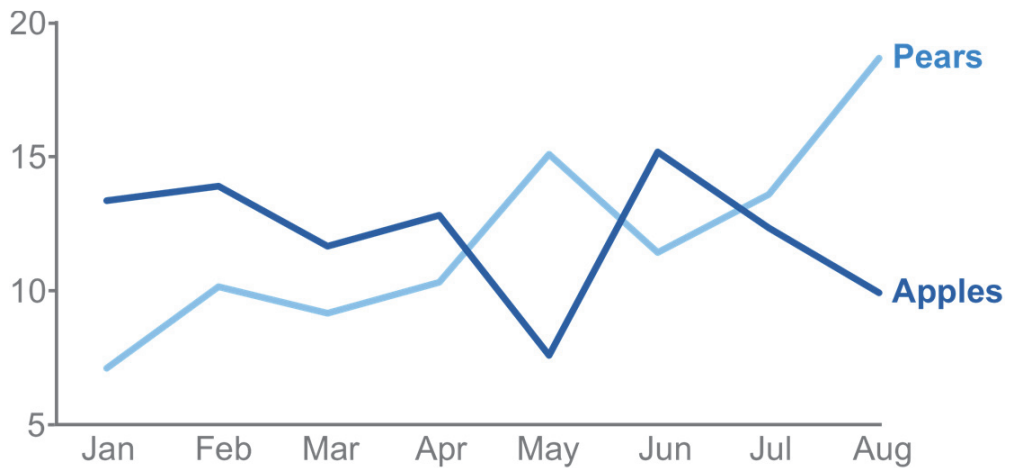
©2023 Van Haren Publishing BV.



# Read Carefully

Plotting trends and intervals over time

**T**  
Temporal



44

©2023 Van Haren Publishing BV.



## Read carefully

**T**  
**Tabular**  
 Organizing observations by variable to allow precise comparison

	Jan	Feb	Mar	Apr	May	Jun
product 1	267	357	587	320	268	398
product 2	365	387	401	406	421	404
product 3	554	582	561	551	583	600
<b>Total</b>	<b>1.186</b>	<b>1.326</b>	<b>1.549</b>	<b>1.277</b>	<b>1.272</b>	<b>1.402</b>

	Jan	Feb	Mar	Apr	May	Jun
<b>product 1</b>	267	357	587	320	268	398
<b>product 2</b>	365	387	401	406	421	404
<b>product 3</b>	554	582	561	551	583	600
<b>Total</b>	<b>1.186</b>	<b>1.326</b>	<b>1.549</b>	<b>1.277</b>	<b>1.272</b>	<b>1.402</b>

45

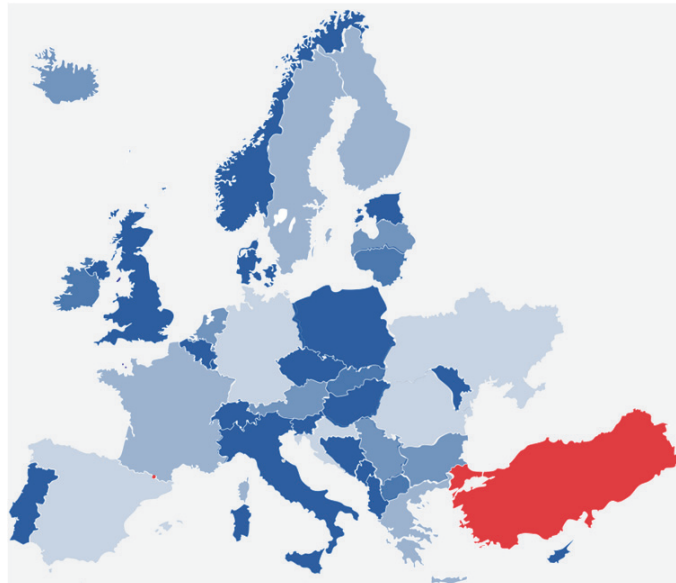
©2023 Van Haren Publishing BV.



## Read carefully

Mapping spatial patterns through overlays and distortions

**S**  
**Spatial**



46

©2023 Van Haren Publishing BV.



## Read carefully

Mapping spatial patterns through overlays and distortions



2020 U.S. presidential election



Last update: 7th November 2020


@karim\_douieb



47

©2023 Van Haren Publishing BV.

## Check your Data: Data Claim Checklist





-  Why does the data matter to you?
-  Does the claim match the data?
-  Does the claim seem plausible?
-  What comparison needs to be made?
-  Who's saying it?
-  How was the data gathered?
-  What's missing?
-  Is the data being distorted?

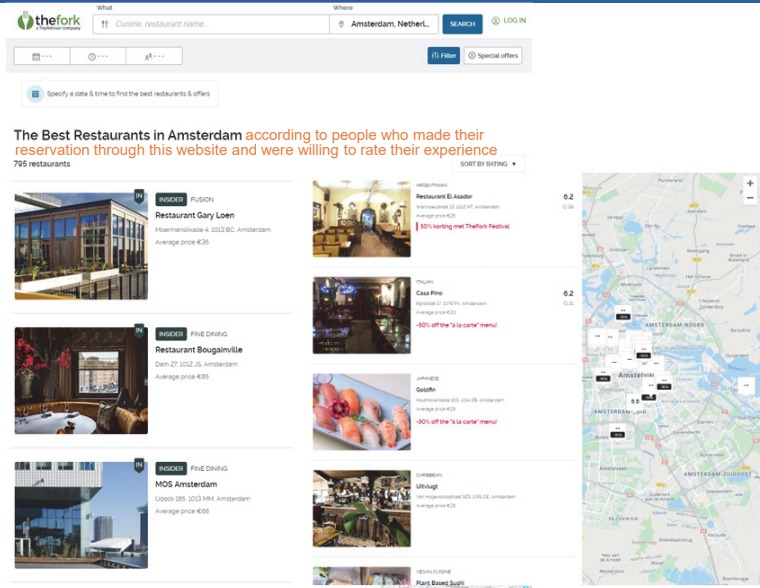


48 Photo by Glenn Carstens-Peters on Unsplash

©2023 Van Haren Publishing BV.

# Check your Data: Data Claim Checklist





-  Why does the data matter to you?
-  Does the claim match the data?
-  Does the claim seem plausible?
-  What comparison needs to be made?




49

©2023 Van Haren Publishing BV.

# Check your Data: Data Claim Checklist

-  Why does the data matter to you?
-  Does the claim match the data?
-  Does the claim seem plausible?
-  What comparison needs to be made?

## ☰ How is my rate calculated?

When you leave a review, you evaluate 3 criterias:

- The quality of the dishes which corresponds to 50% of the rate.
- The treatment and the service, which corresponde to 25% of the rate.
- The ambience of the restaurant, which corresponds to 25% of the rate.

Trending Art

How can I cancel

**...TheFork uses the Bayesian method for calculating the average score...**

Your opinion is important to us! Concentrate on your gastronomic experience and think what would you like to read about the restaurant when you book with **TheFork**.

How can I modify

Can I use my Yu

Information About The Reviews

About The Reviews



50

©2023 Van Haren Publishing BV.