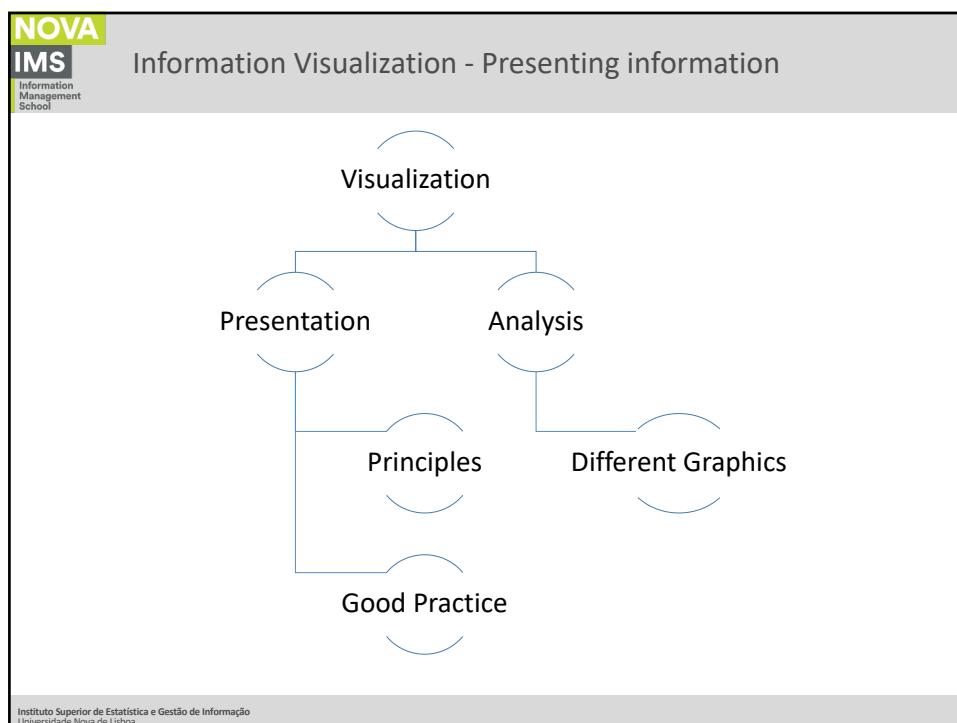


The slide features the NOVA-IMS logo in the top left corner. The title "Data Mining" and subtitle "S4" are centered. Below the title, the text "NOVA-IMS", "Fernando Lucas Bação", and an email address "bacao@isegi.unl.pt" are displayed, along with a URL "http://www.isegi.unl.pt/bacao". At the bottom, the text "Instituto Superior de Estatística e Gestão de Informação" and "Universidade Nova de Lisboa" is present.

1



The slide features the NOVA-IMS logo in the top left corner. The main title is "Information Visualization - Presenting information". A hierarchical diagram is centered, showing "Visualization" at the top, which branches into "Presentation" and "Analysis". "Presentation" further branches into "Principles" and "Good Practice". "Analysis" branches into "Different Graphics".

2

NOVA

IMS
Information
Management
School

Information Visualization – 5 Influential

*Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.**(Dessin par M. Minard, Ingénieur Général des Ponts et Chaussées, à Paris, le 20 Novembre 1869.)*

Les nombreux hommes perdus sont représentés par les longueurs des lignes colorées à savoir : s'ils n'avaient pas été tués pour être morts, il leur aurait fallu plus d'hommes ; ils sont le plus écrit en termes de pertes. Le rouge désigne les hommes qui ont survécu à Russie, le noir ceux qui ne l'ont pas fait.

Les renseignements qui ont servi à dresser la carte ont été tirés des ouvrages de M.M. Chaptal, de Chaptal, de Secretan, de Chambray et du journal intitulé "Archives historiques de l'Armée", depuis le 25 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai superposé sur la carte un déclin à Madelaine auquel regarder vers Orléans et Wittenberg, auquel trajet marche avec l'armée.

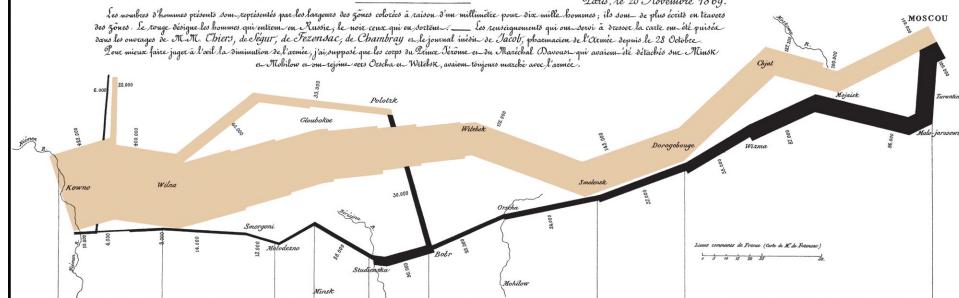
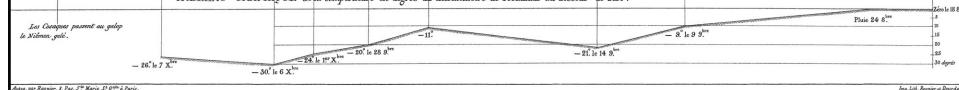


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Source: https://en.wikipedia.org/wiki/Charles_Joseph_MinardInstituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

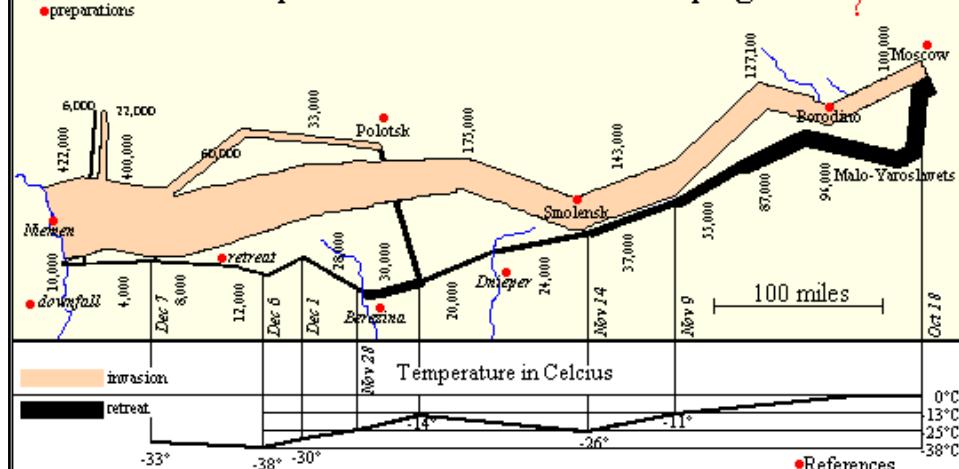
3

NOVA

IMS
Information
Management
School

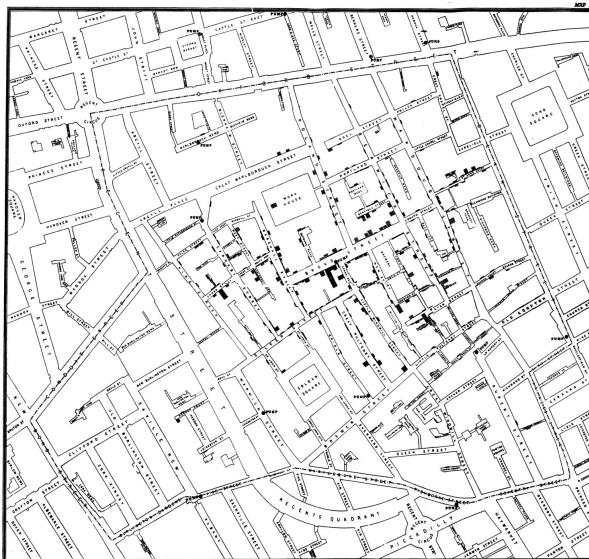
Information Visualization – 5 Influential

Napoleon's 1812 Russian Campaign

Source: https://www.d.umn.edu/~cstroupe/ideas/representational_space_11.htmlInstituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

4

Information Visualization – 5 Influential

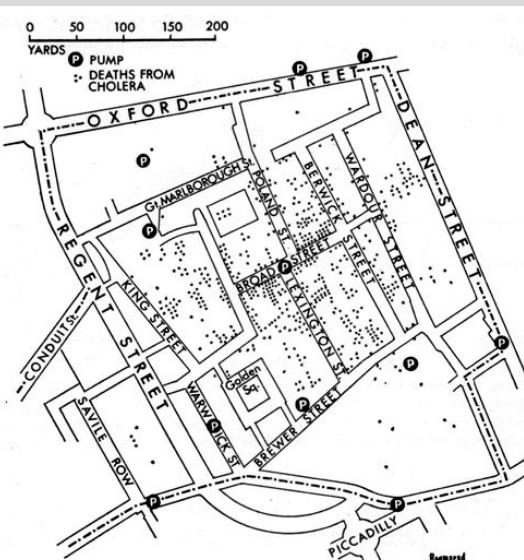


Source: https://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

5

Information Visualization – 5 Influential



Source: <http://www.ph.ucla.edu/epi/snow/mapsbroadsstreet.html>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

6

NOVA
IMS
Information Management School

Information Visualization Strong Points

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNIGIS A3ES iSchools eduniversal

7

NOVA
IMS
Information Management School

Information Visualization – Strong Points

EGG CURRY 4 PEOPLE, MAIN DISH SERVED WITH WHITE RICE & RED BEANS

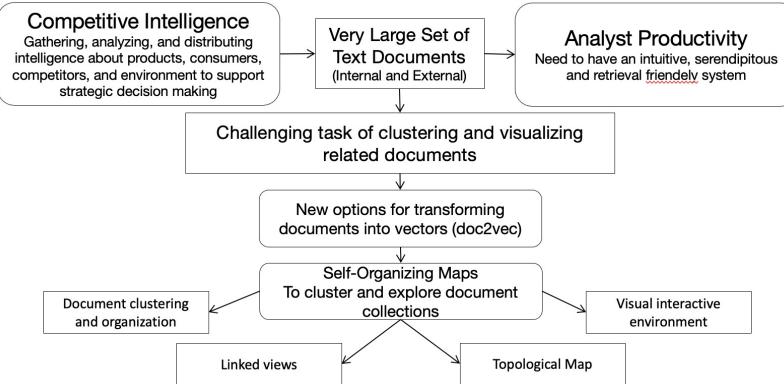
EGGPLANT ROUGAIL 4 PEOPLE, SIDE DISH SERVED HOT OR COLD, GOES WELL WITH MEAT OR EGGS

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

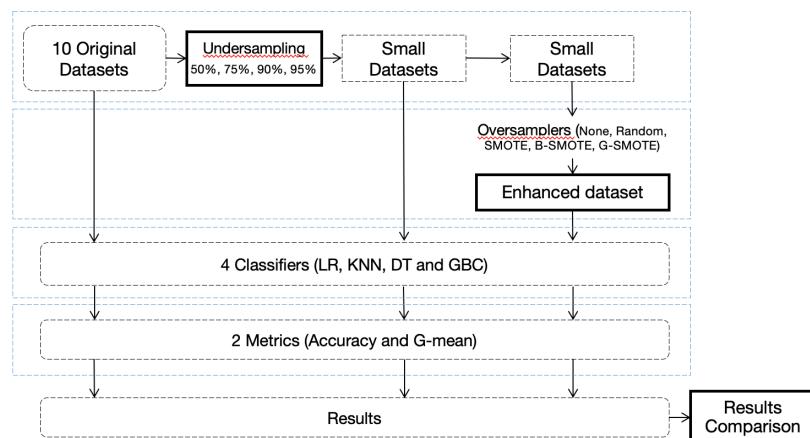
8

Information Visualization – Strong Points

A neural network-based document clustering system, which is a genuine content-addressable memory system, meaning it clusters and stores text documents so they can be retrieved based on their content.

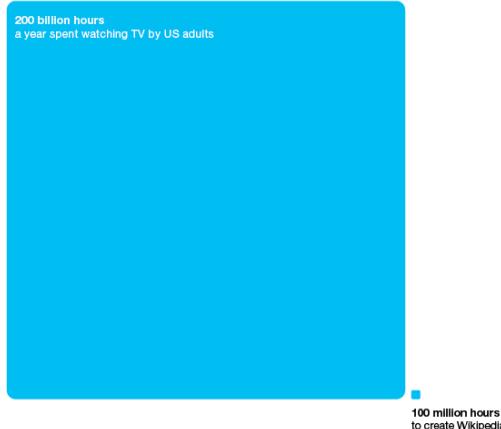


Information Visualization – Strong Points



Information Visualization – Strong Points

Goggle Boxes Hours spent...

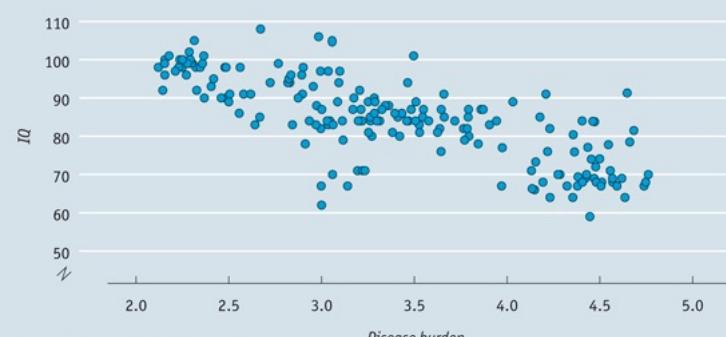


David McCandless // July 10 // source: Cognitive Surplus by Clay Shirky // InformationIsBeautiful.net

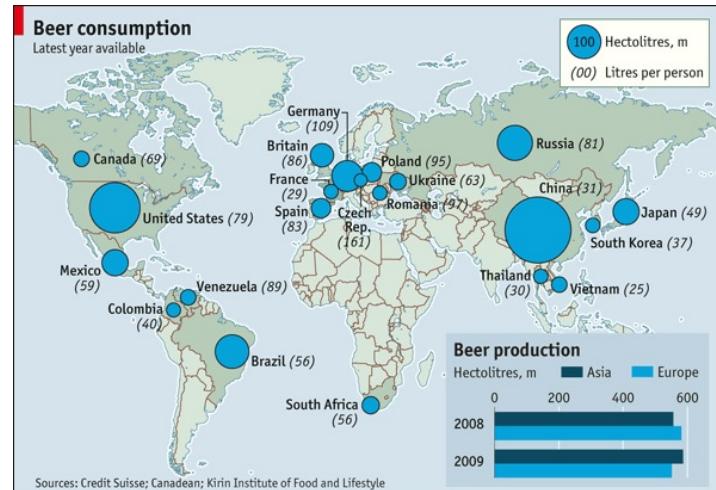
Information Visualization – Strong Points

Intelligence tested

IQ v burden of infectious disease in 184 countries



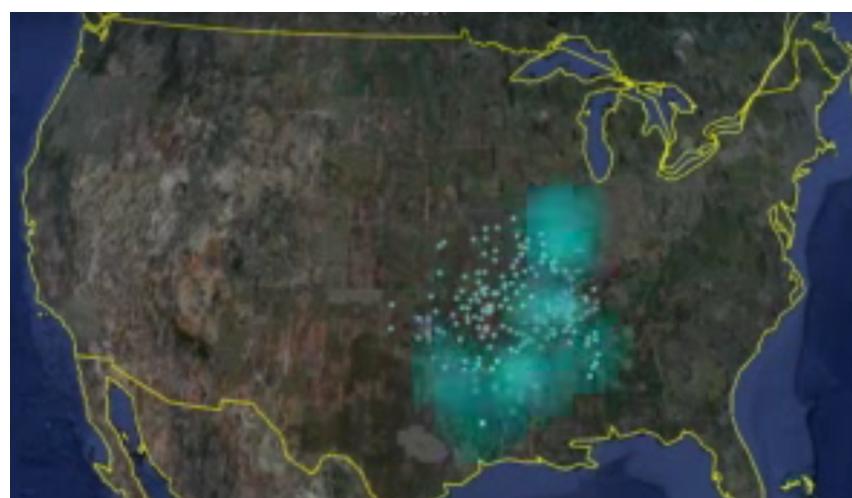
Information Visualization – Strong Points



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

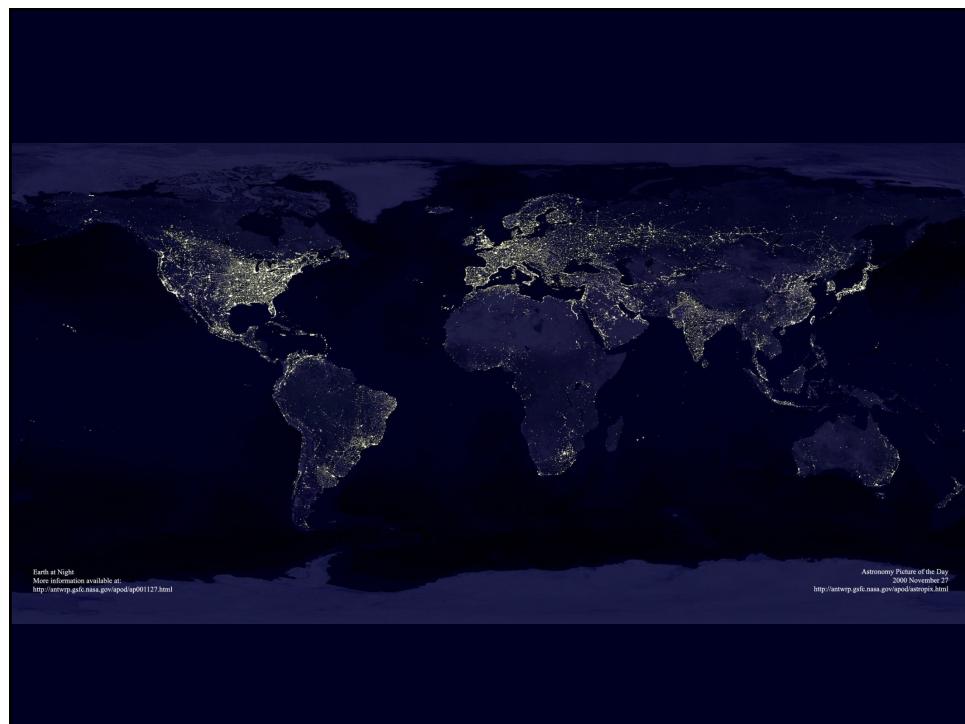
13

Information Visualization – Strong Points



Source: <http://www.youtube.com/watch?v=ZmUSk231WFE&feature=related>
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

14

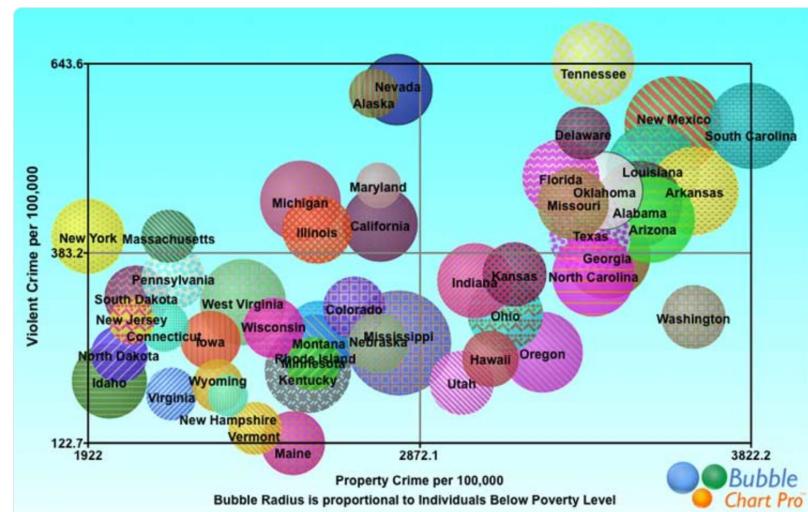


15

A presentation slide for NOVA IMS Information Visualization. The slide has a light gray background. In the top left corner, there is a logo for NOVA IMS, which includes the text "NOVA" in a green box, "IMS" in a black box, and "Information Management School" in a smaller black font below it. The main content area is a white box containing the text "Information Visualization" in large green bold letters, followed by "Please, don't do this!!!" in a slightly smaller green bold font. The bottom of the slide features a green footer bar with the text "Instituto Superior de Estatística e Gestão de Informação" and "Universidade Nova de Lisboa". On the right side of the footer, there are several logos: EQUIS, UNICIS, A3ES, AACSB, iSchools, and eduniversal.

16

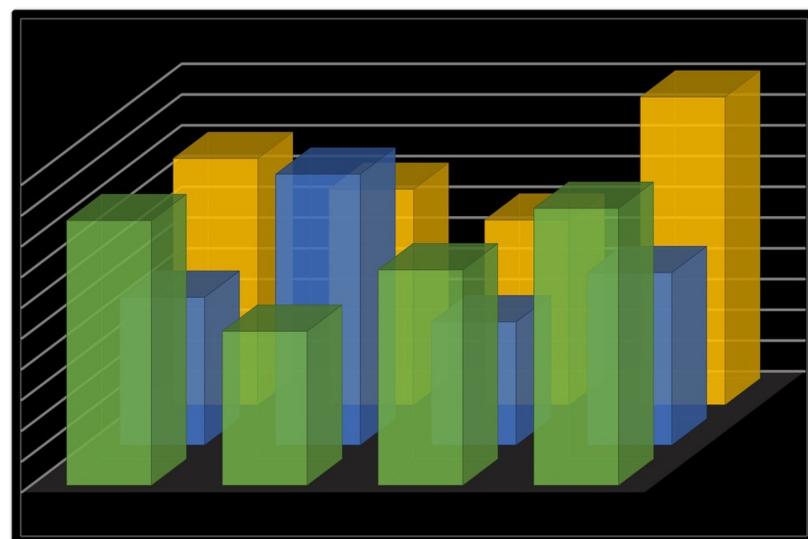
Information Visualization - Presenting information



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

17

Information Visualization - Presenting information



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

18



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

19

Information Visualization Some guidelines

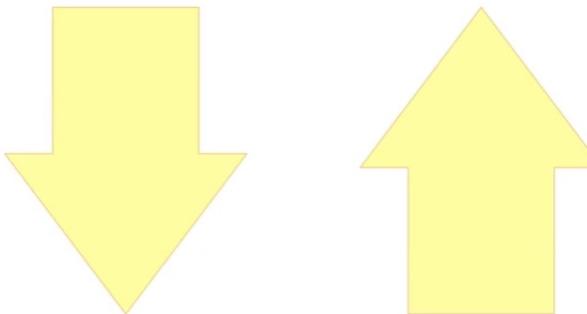
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

20

NOVA
IMS
Information Management School

Information Visualization - Presenting information

Fundamental rule of chart design



Reduce chartjunk
(remove chart elements that are decorative or ornamental)
<http://www.juiceanalytics.com/white-papers-guides-and-more/>

Increase data-ink ratio
(make every pixel tell a story about your data)

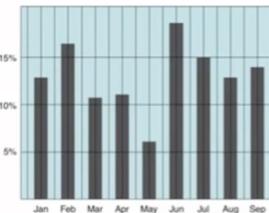
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

21

NOVA
IMS
Information Management School

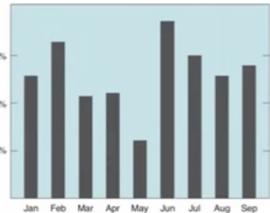
Information Visualization - Improvements (good and bad)

Remove chart-junk



Mês	Dados
Jan	12%
Feb	15%
Mar	10%
Apr	11%
May	6%
Jun	17%
Jul	14%
Aug	12%
Sep	13%

Increase data-ink



Mês	Dados
Jan	12%
Feb	15%
Mar	10%
Apr	11%
May	6%
Jun	17%
Jul	14%
Aug	12%
Sep	13%

Increase data-ink

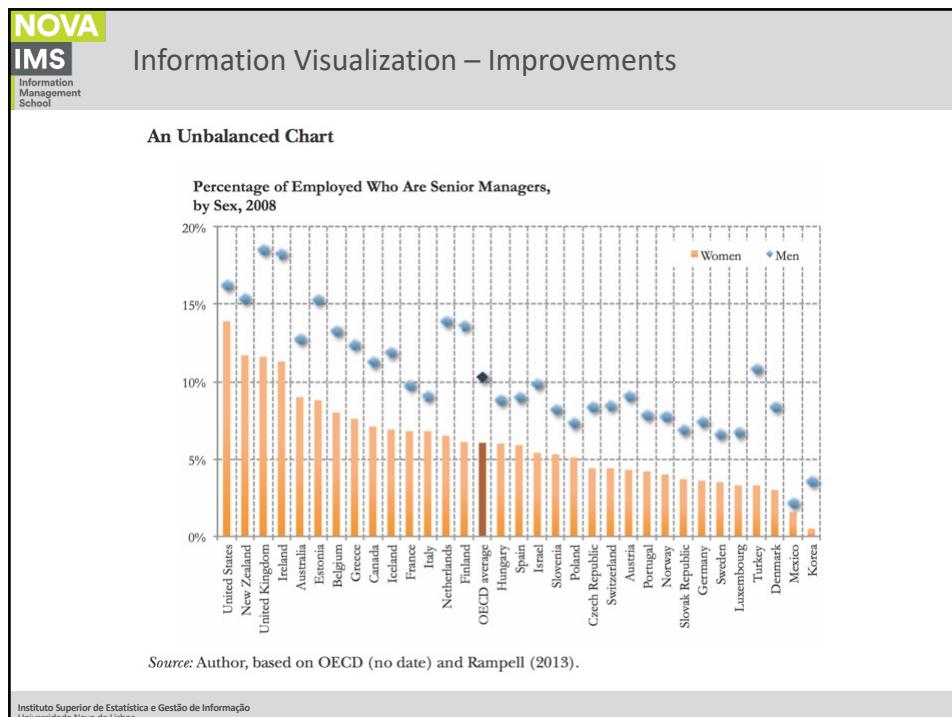


Mês	Dados
Jan	12%
Feb	15%
Mar	10%
Apr	11%
May	6%
Jun	17%
Jul	14%
Aug	12%
Sep	13%

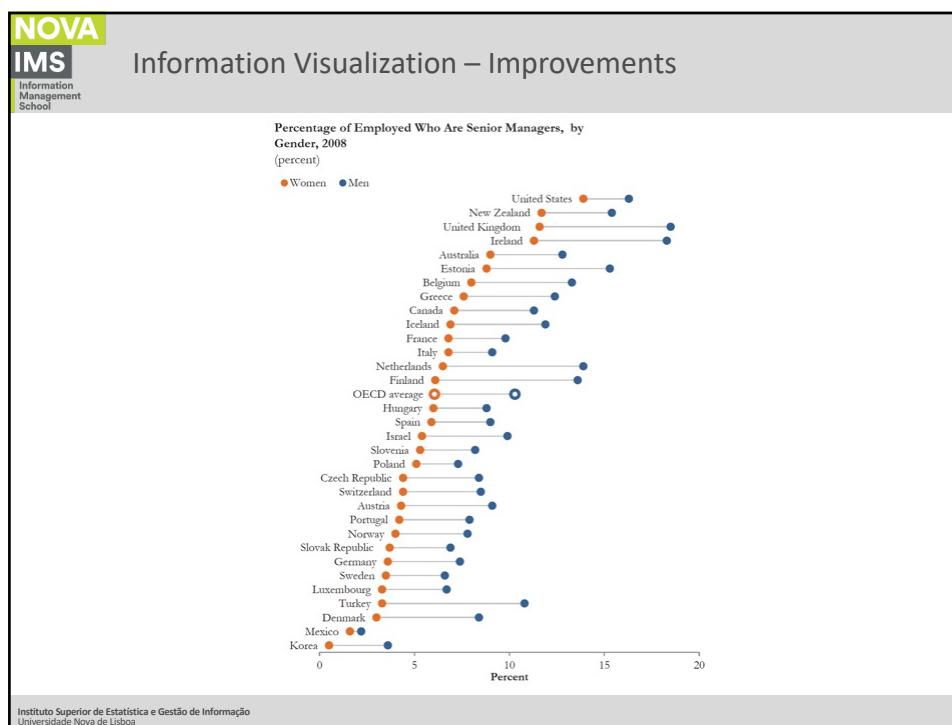
<http://www.juiceanalytics.com/white-papers-guides-and-more/>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

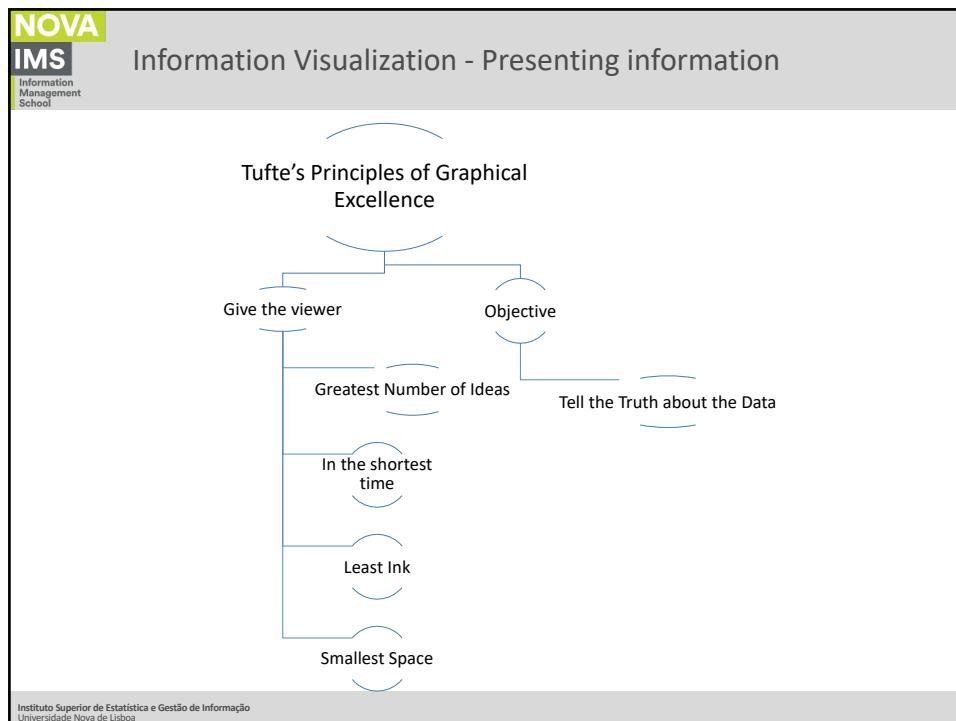
22



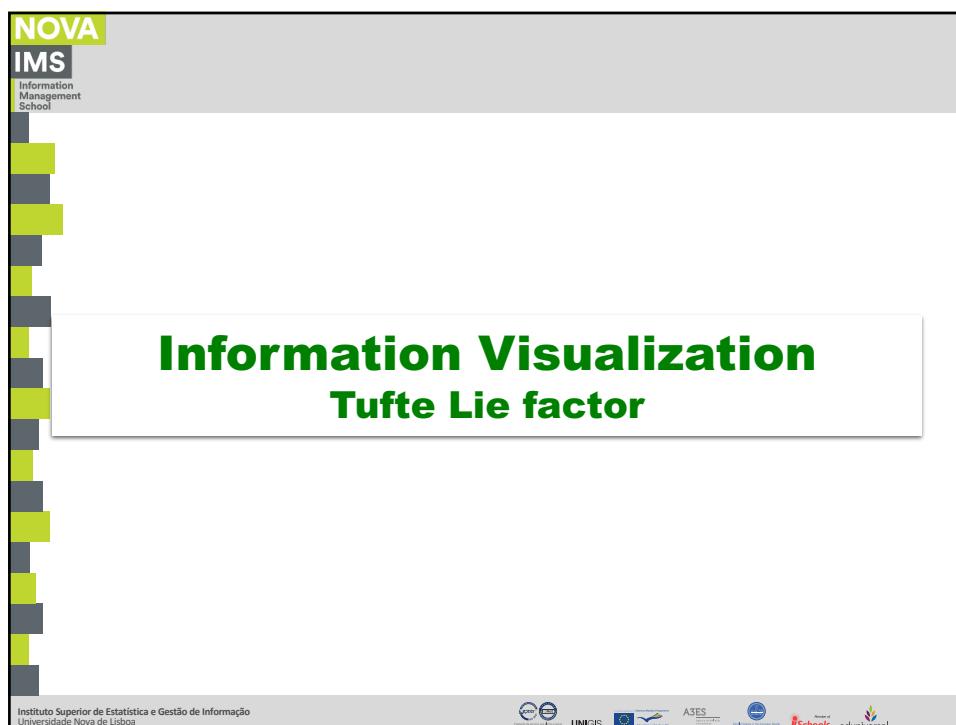
23



24

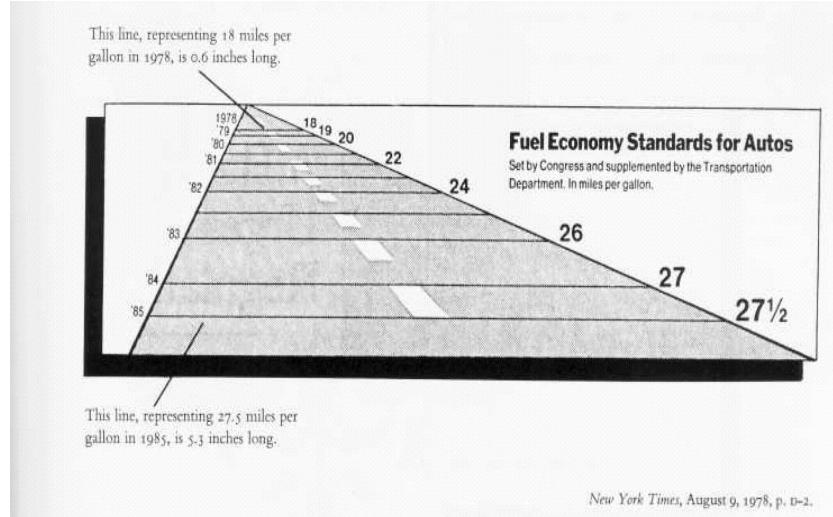


25



26

Information Visualization - Tufte Lie factor



New York Times, August 9, 1978, p. D-2.

(E.R. Tufte, "The Visual Display of Quantitative Information", 2nd edition)

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

27

Information Visualization - Tufte Lie factor

- Ed Tufte of Yale University has defined a “lie factor” as a measure of the amount of distortion in a graph.

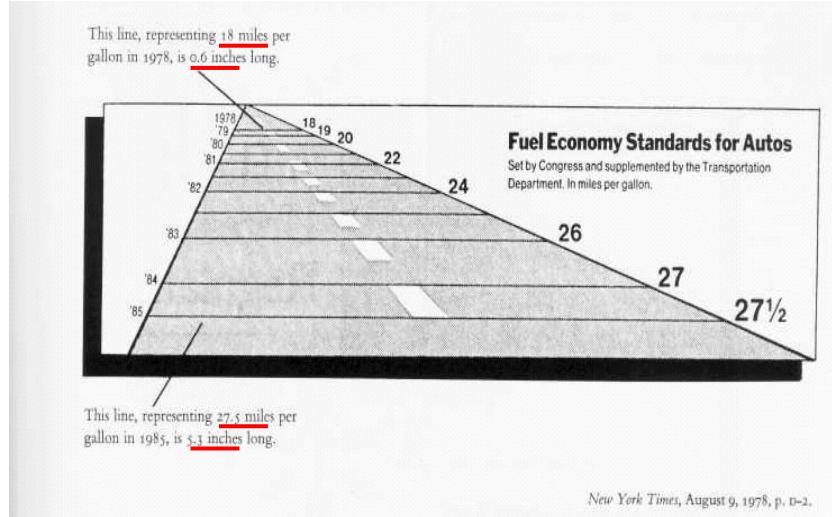
$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}} =$$

- If the lie factor of a graph is greater than 1, the graph is exaggerating the size of the effect.
- Tufte requirement: $0.95 < \text{Lie Factor} < 1.05$

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

28

Information Visualization - Tufte Lie factor



(E.R. Tufte, "The Visual Display of Quantitative Information", 2nd edition)

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

29

Information Visualization - Tufte Lie factor

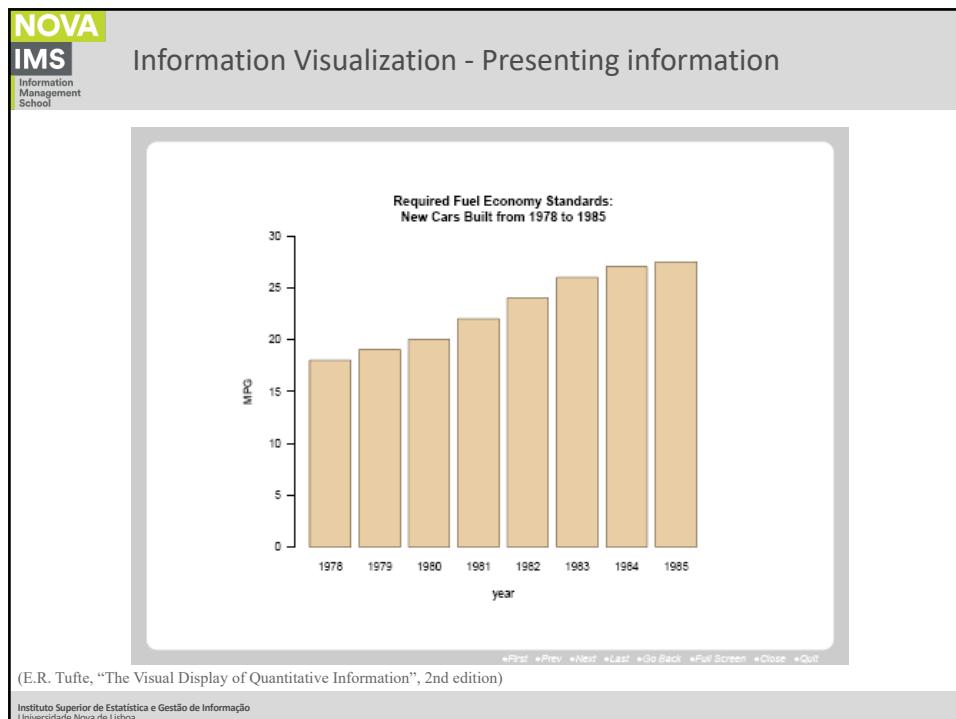
$$\text{Lie Factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}} =$$

$$= \frac{\frac{(5.3 - 0.6)}{0.6}}{\frac{(27.5 - 18.0)}{18}} = \frac{\frac{4.7}{0.6}}{\frac{9.5}{18}} = \frac{7.833}{0.528} = 14.8$$

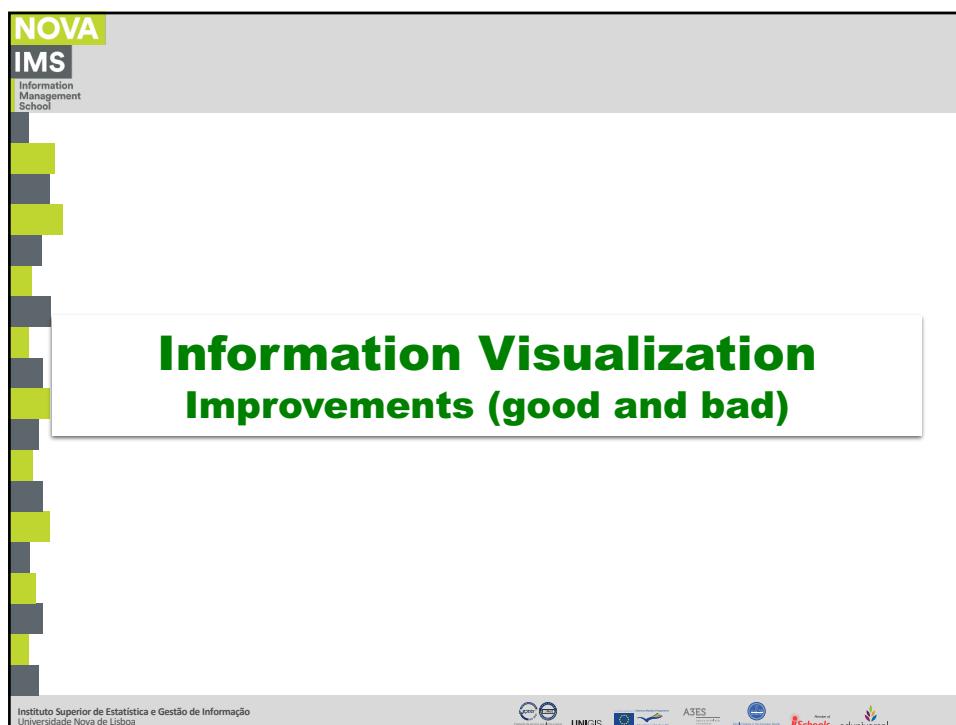
Tufte requirement: $0.95 < \text{Lie Factor} < 1.05$

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

30



31

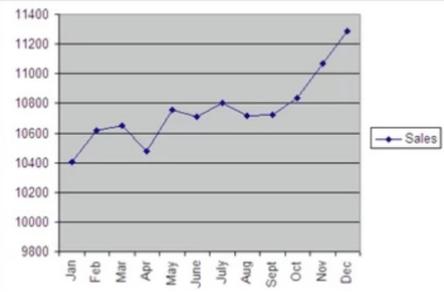


32

NOVA
IMS
Information Management School

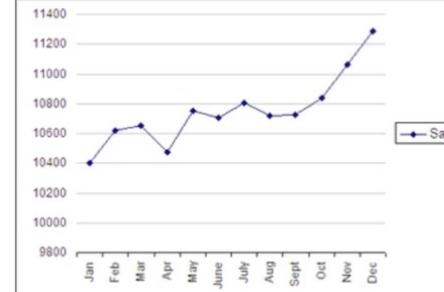
Information Visualization - Improvements (good and bad)

Bad



Month	Sales
Jan	10400
Feb	10600
Mar	10700
Apr	10500
May	10800
June	10700
July	10900
Aug	10700
Sept	10700
Oct	10900
Nov	11100
Dec	11300

Good



Month	Sales
Jan	10400
Feb	10600
Mar	10700
Apr	10500
May	10800
June	10700
July	10900
Aug	10700
Sept	10700
Oct	10900
Nov	11100
Dec	11300

Maximize the contrast between your data and the background.

<http://www.juiceanalytics.com/white-papers-guides-and-more/>

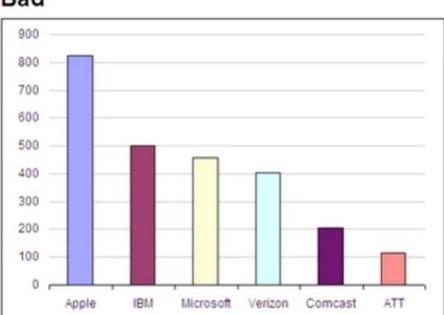
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

33

NOVA
IMS
Information Management School

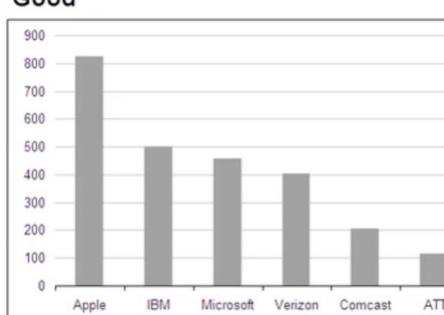
Information Visualization - Improvements (good and bad)

Bad



Company	Value
Apple	820
IBM	500
Microsoft	450
Verizon	400
Comcast	200
ATT	100

Good



Company	Value
Apple	820
IBM	500
Microsoft	450
Verizon	400
Comcast	200
ATT	100

<http://www.juiceanalytics.com/white-papers-guides-and-more/>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

34

NOVA
IMS
Information Management School

Information Visualization – Improvements (good and bad)

Bad

Company	Revenue (approx.)
Abbot Labs	200
Lilly	400
Merck	820
Novartis	450
Pfizer	500
Roche	100

Good

Company	Revenue (approx.)
Merck	820
Pfizer	500
Novartis	450
Lilly	380
Abbot Labs	200
Roche	100

<http://www.juiceanalytics.com/white-papers-guides-and-more/>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

35

NOVA
IMS
Information Management School

Information Visualization – Improvements (good and bad)

Bad

Company	Revenue (approx.)
Apple	400
IBM	250
Microsoft	200
Verizon	180
Comcast	100
ATT	80

Good

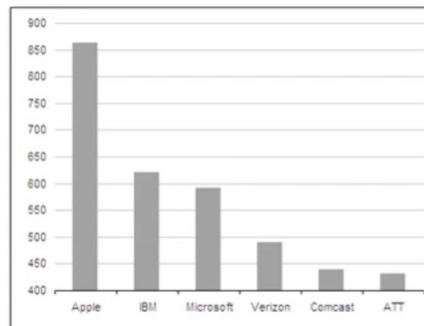
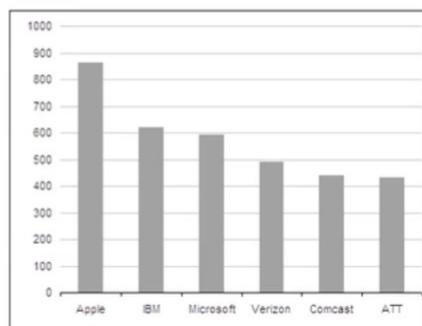
Company	Revenue (approx.)
Apple	800
IBM	500
Microsoft	450
Verizon	380
Comcast	200
ATT	120

<http://www.juiceanalytics.com/white-papers-guides-and-more/>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

36

Information Visualization – Improvements (good and bad)

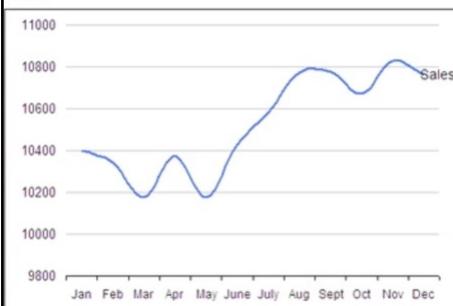
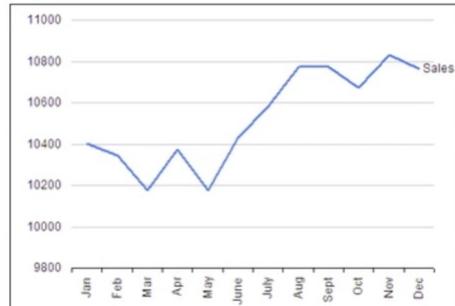
Bad**Good**

<http://www.juiceanalytics.com/white-papers-guides-and-more/>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

37

Information Visualization – Improvements (good and bad)

Bad**Good**

<http://www.juiceanalytics.com/white-papers-guides-and-more/>

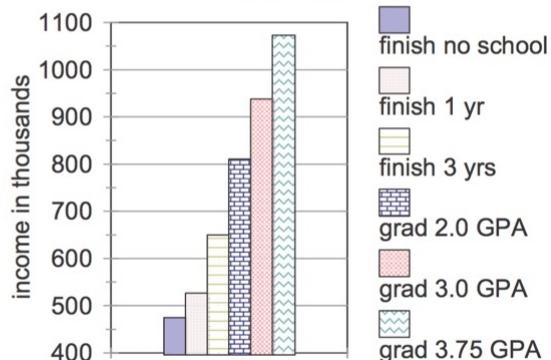
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

38

Information Visualization – Improvements (good and bad)

The Basic Column Chart

Figure 2 Discounted Expected Lifetime Earnings, VN(t')

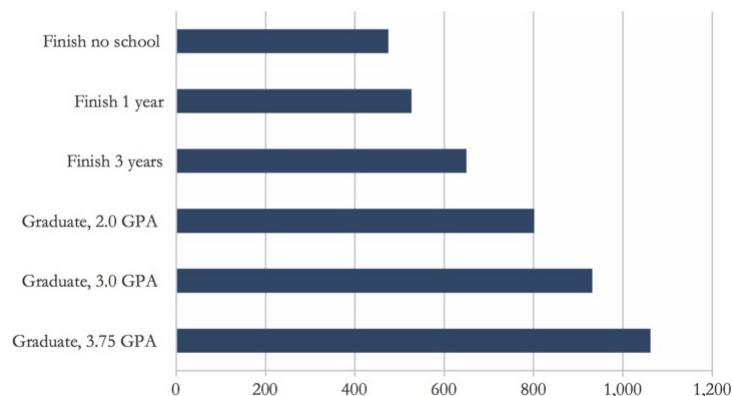


Source: Stinebrickner and Stinebrickner (2013).

Information Visualization – Improvements (good and bad)

The Revised Column Chart

Discounted Expected Lifetime Earnings, VN(t')
(Income in thousands)

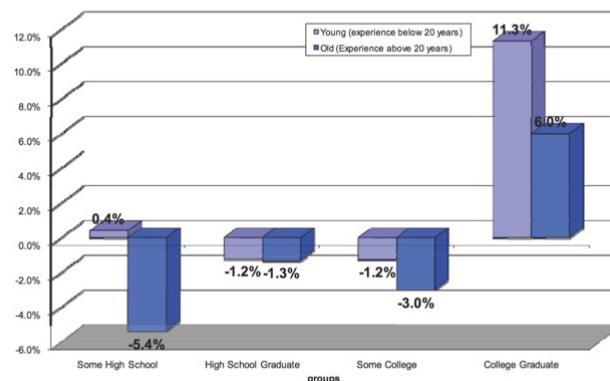


Source: Author's calculations using numbers inferred from text in Stinebrickner and Stinebrickner (2013).

Information Visualization – Improvements (good and bad)

A 3D Chart

Change in real weekly wages of US-born workers by group, 1990-2006



Source: Ottaviano and Peri (2008).

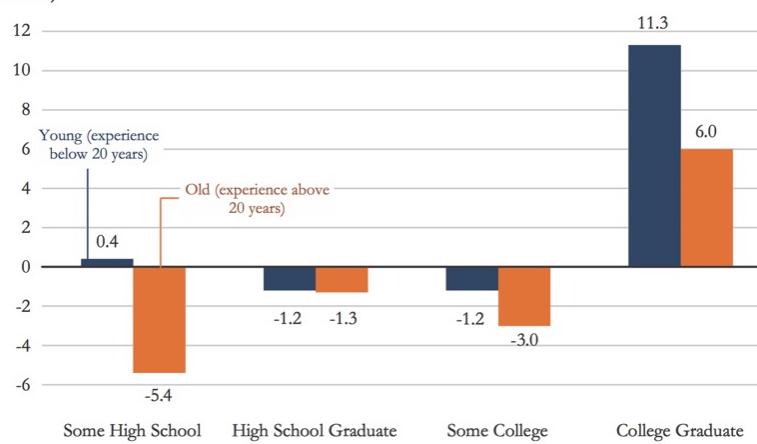
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

41

Information Visualization – Improvements (good and bad)

Flattening a 3D Chart

Change in real weekly wages of US-born workers by group, 1990–2006
(Percent)



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

42

NOVA
IMS
Information Management School

Information Visualization – Improvements (good and bad)

Bad

Good

<http://www.iijiceanalytics.com/white-papers-guides-and-more/>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

43

NOVA
IMS
Information Management School

Information Visualization – Line chart

A Spaghetti Chart

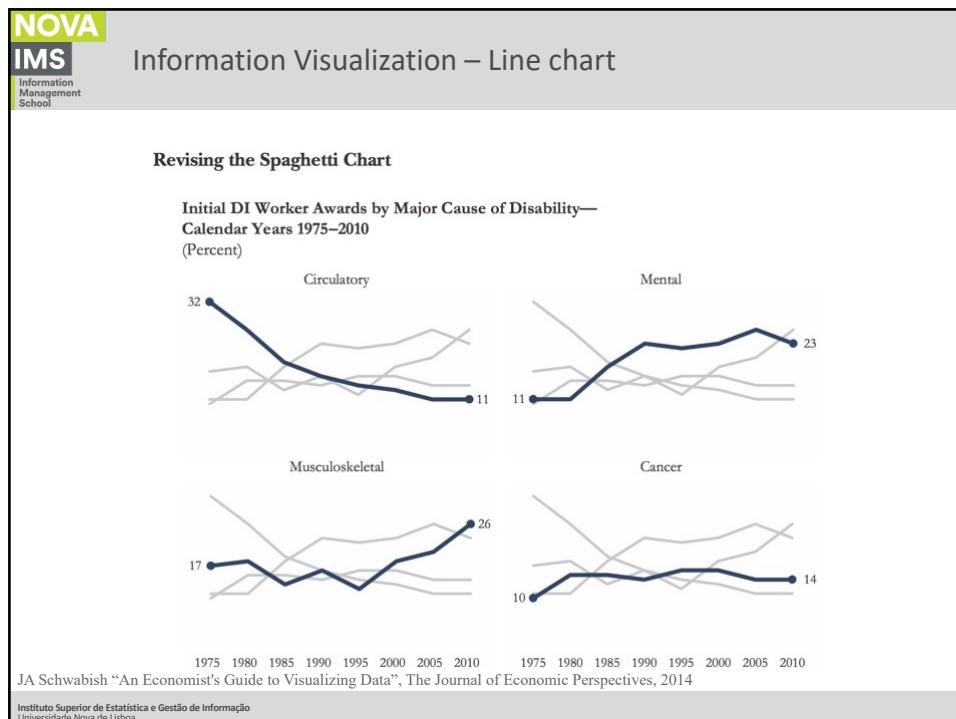
27. Initial DI Worker Awards by Major Cause of Disability—Calendar Years 1975-2010

Year	Mental (%)	Cancer (%)	Circulatory (%)	Musculoskeletal (%)
1975	11	10	33	17
1980	11	15	26	18
1985	18	15	18	14
1990	23	14	16	16
1996	22	15	14	12
2000	23	15	13	18
2005	26	14	11	20
2010	23	14	11	26

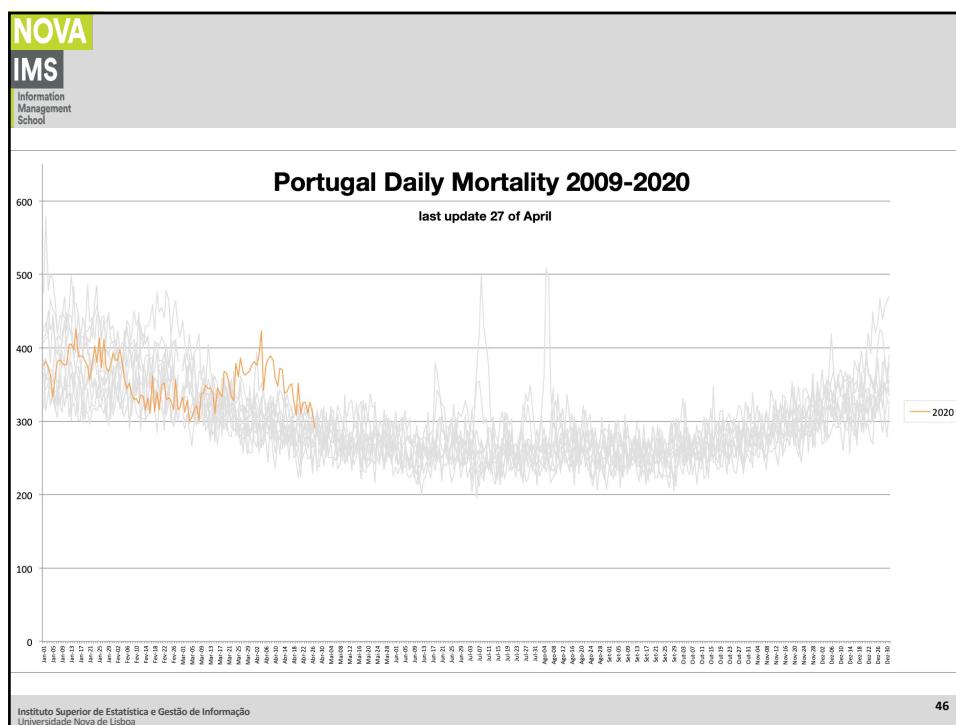
Source: Social Security Advisory Board (2012).
JA Schwabish “An Economist’s Guide to Visualizing Data”, The Journal of Economic Perspectives, 2014

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

44

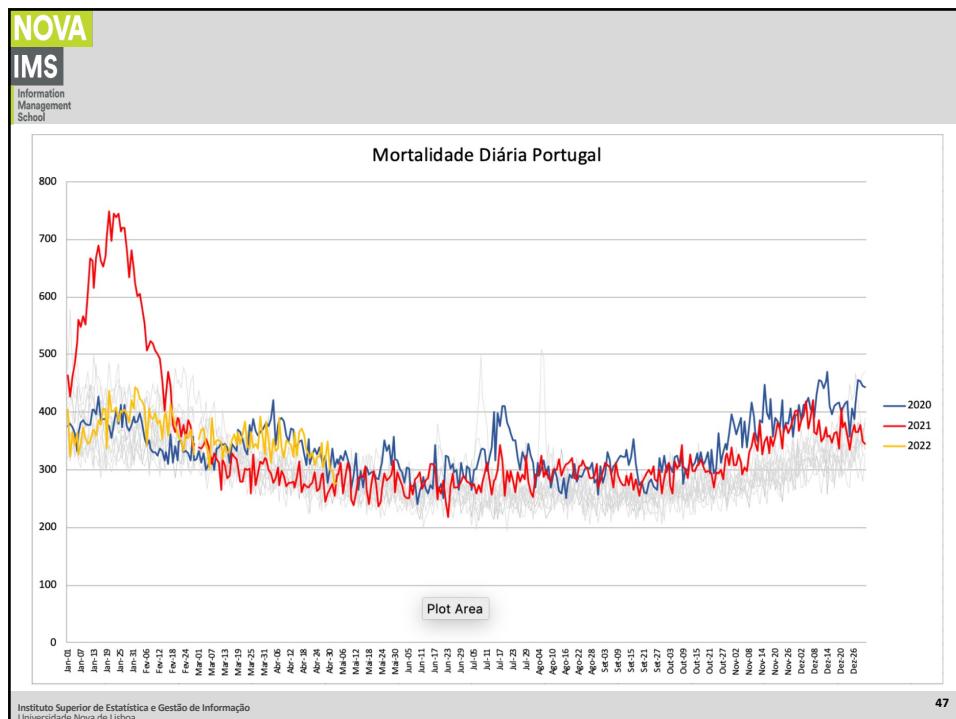


45

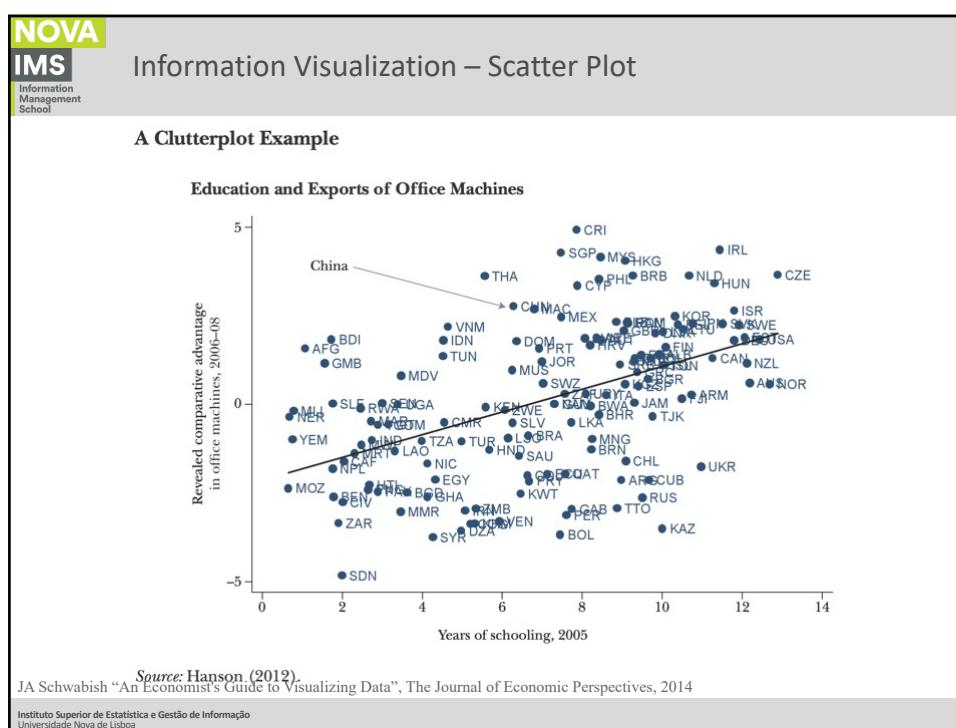


46

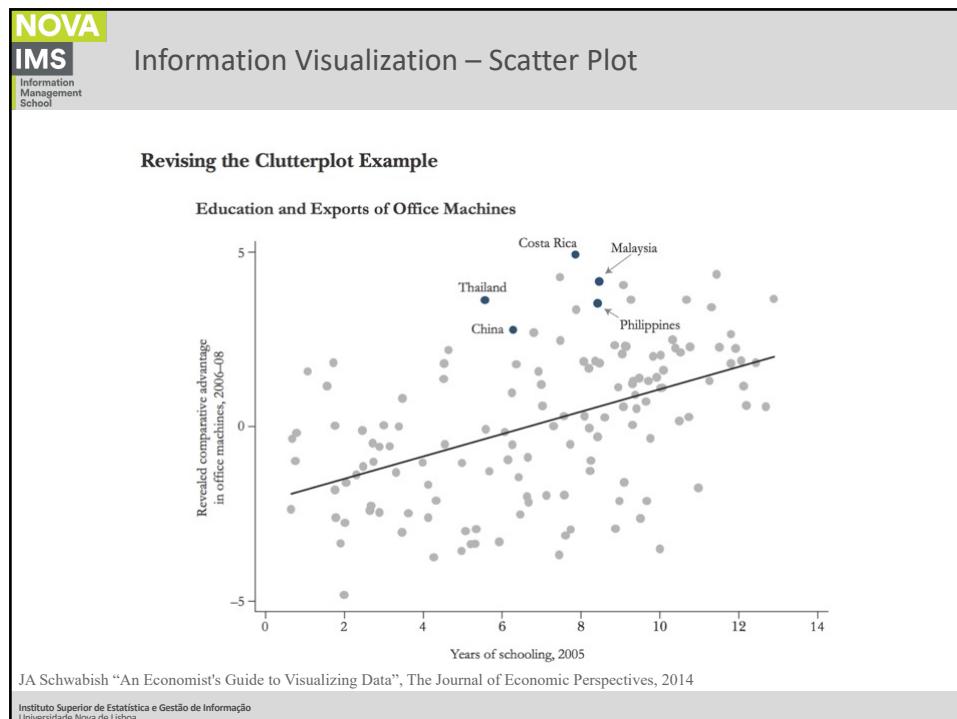
46



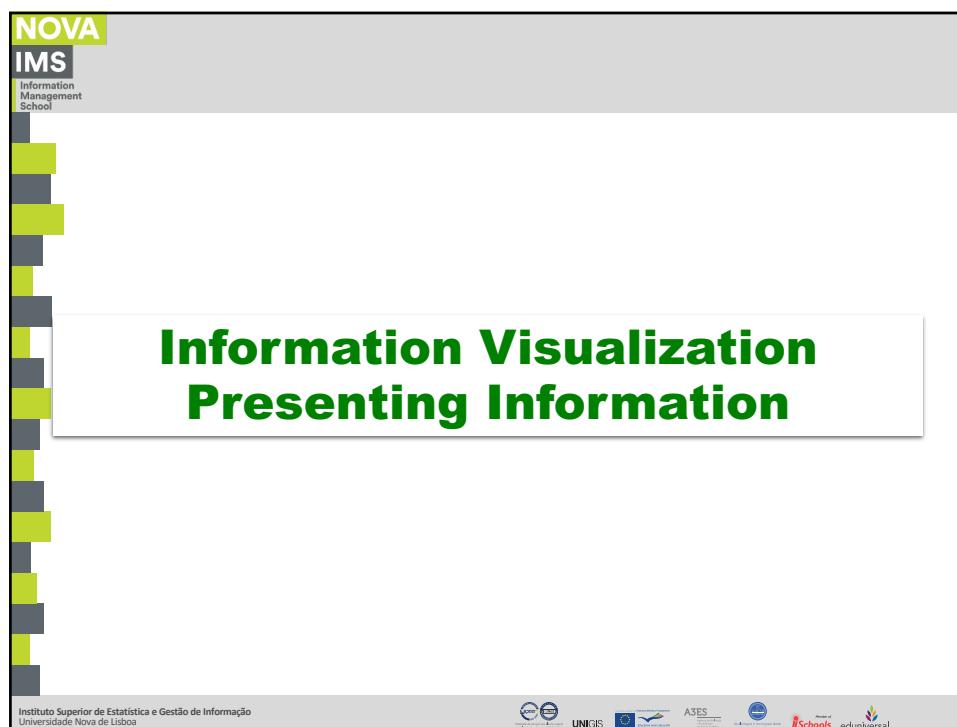
47



48

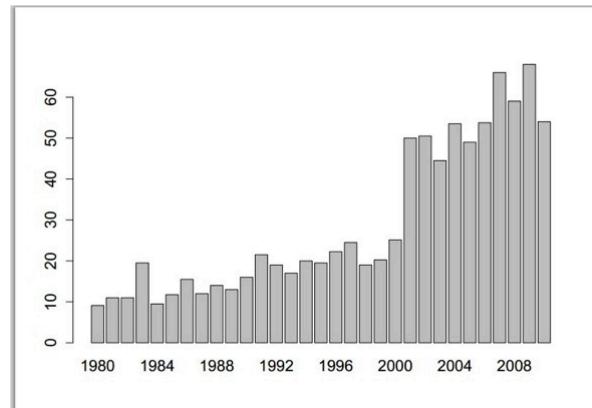


49



50

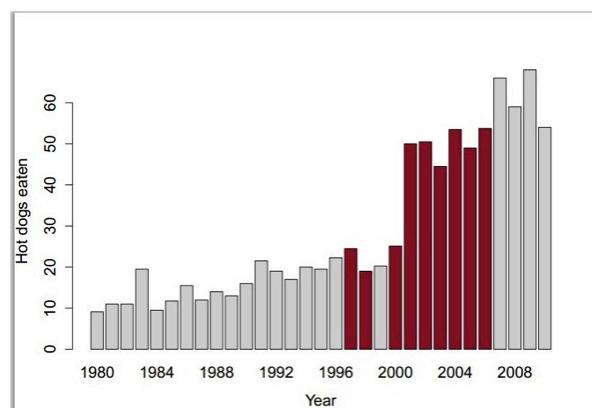
Information Visualization – Bar chart



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

51

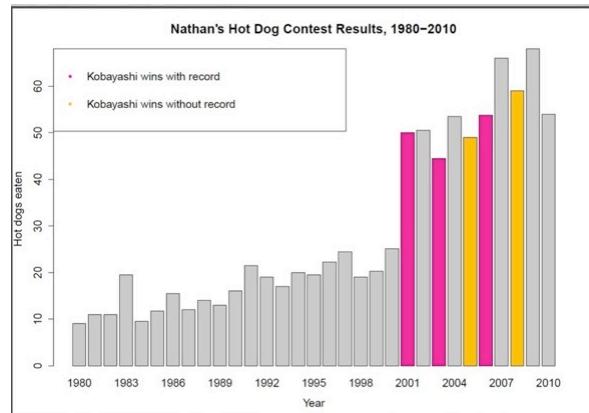
Information Visualization - Bar chart



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

52

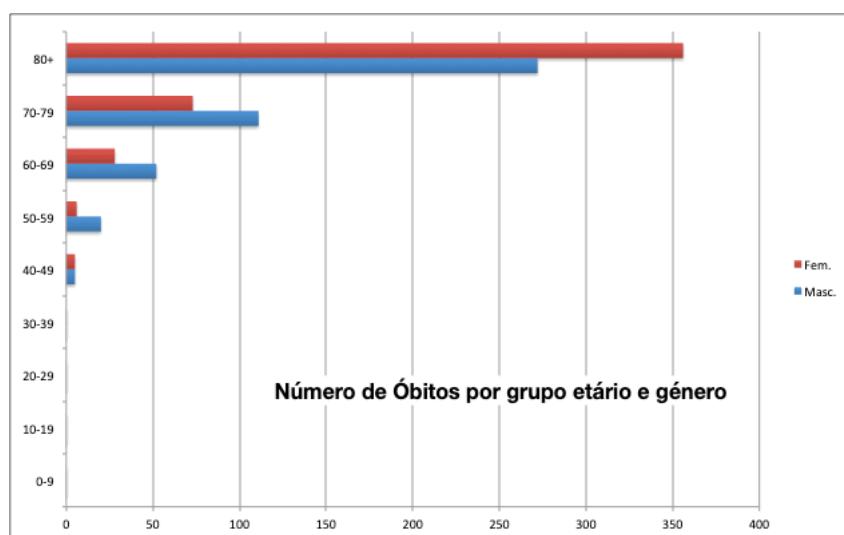
Information Visualization - Bar chart



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

53

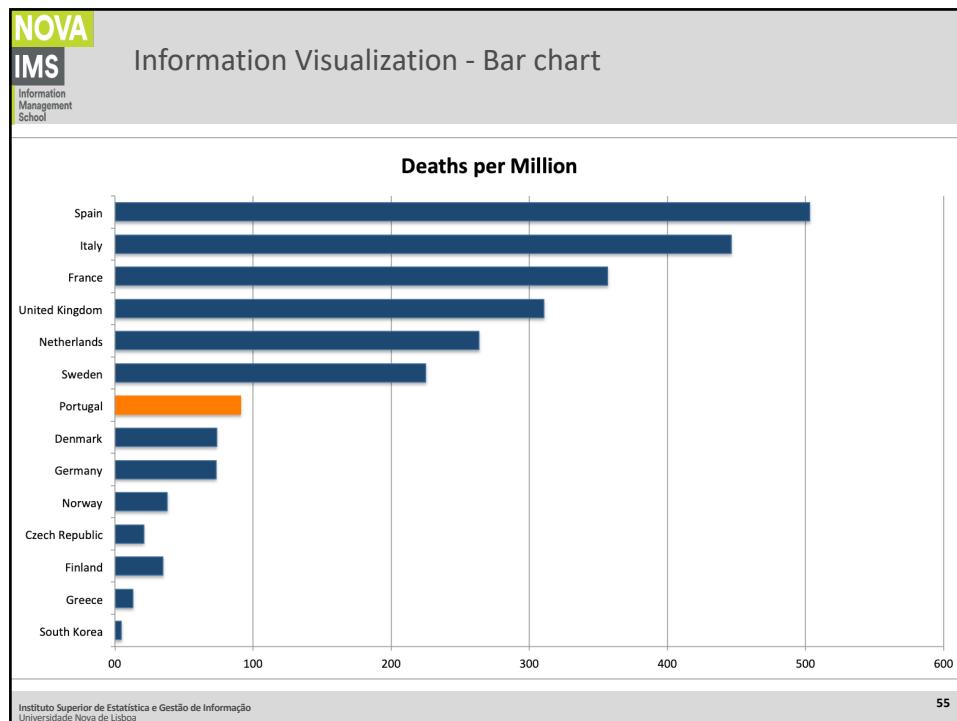
Information Visualization - Bar chart



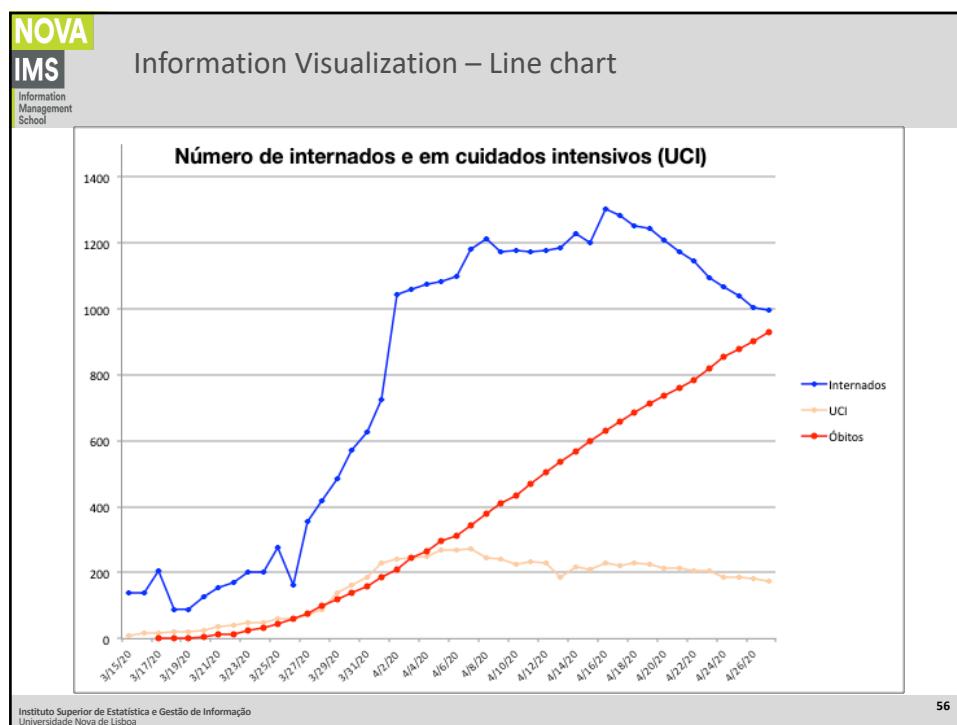
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

54

54

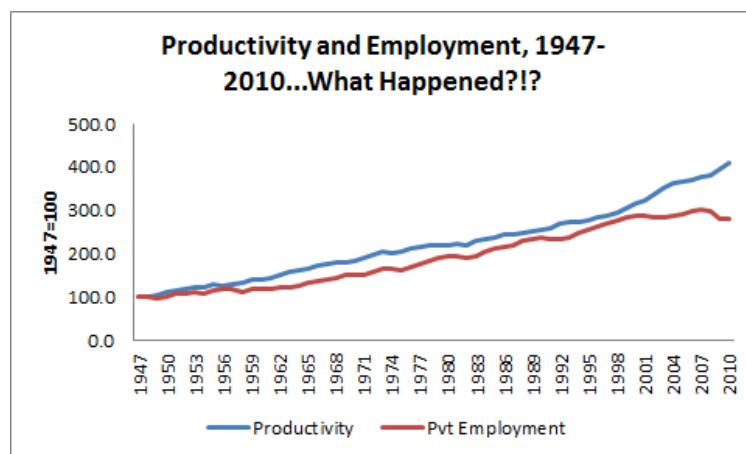


55



56

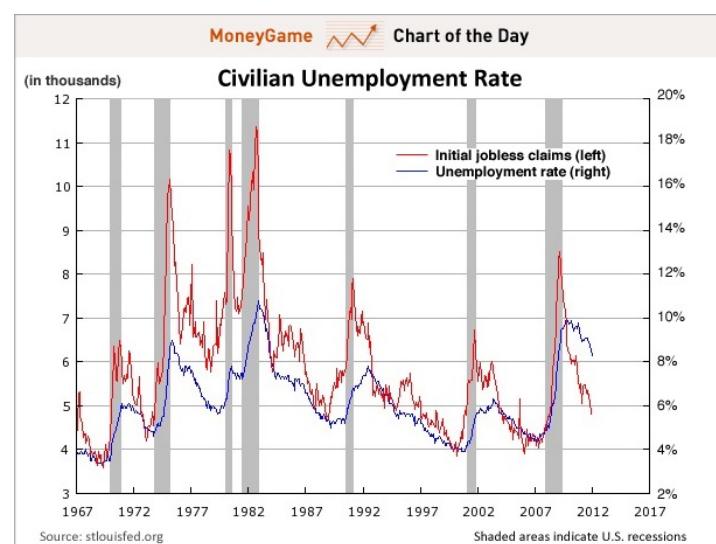
Information Visualization – Line chart



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

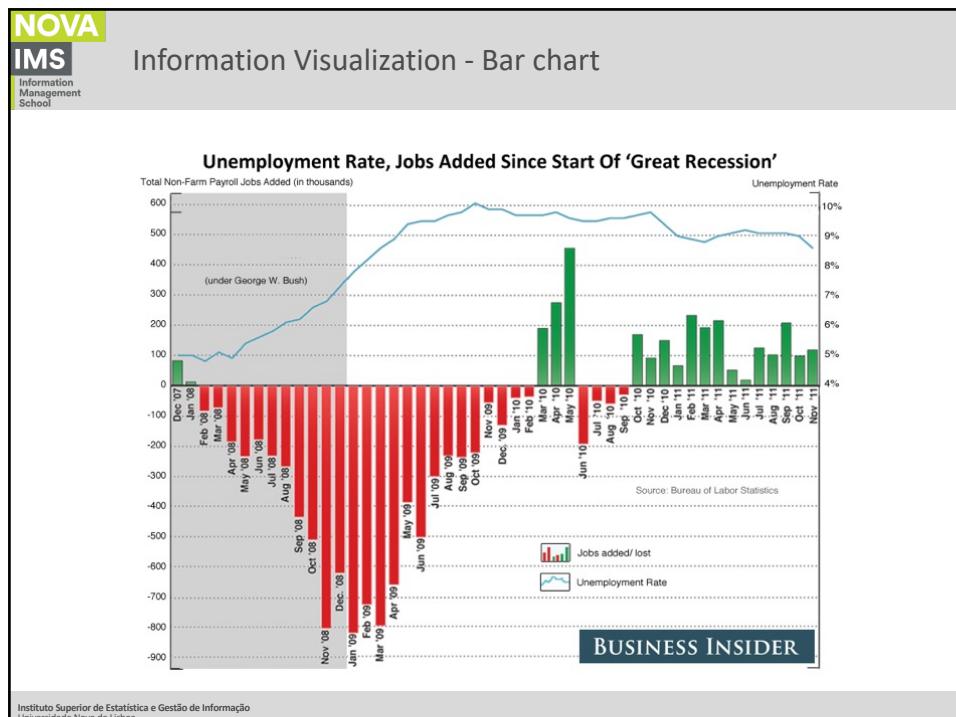
57

Information Visualization - Line chart

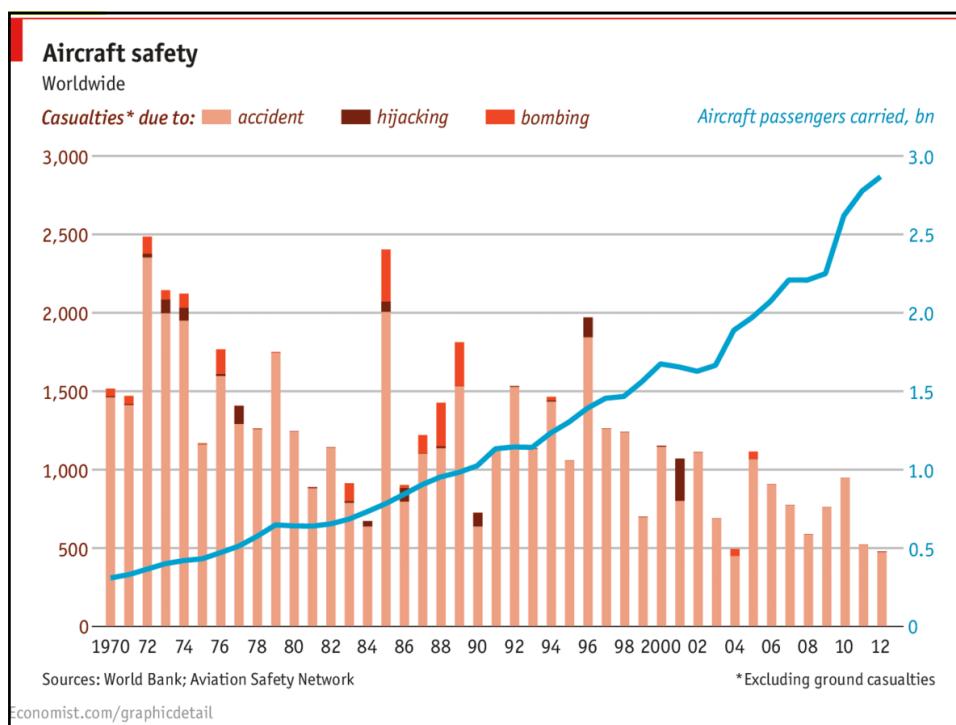


Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

58



59

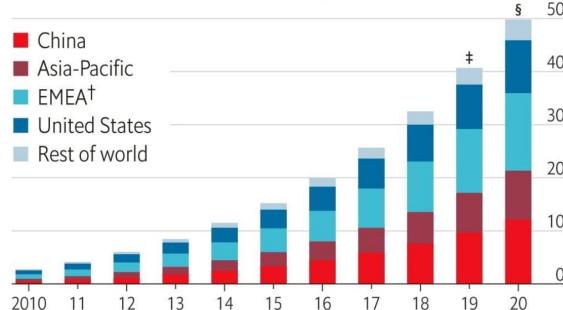


60

Information Visualization – Stacked Bar chart

The world will generate 19trn DVDs worth of data this year and next

Data generated, worldwide, zettabytes*



*1ZB=1 trillion GB †Europe, Middle East and Africa
‡Estimate \$Forecast

Source: IDC, Seagate

The Economist

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

61

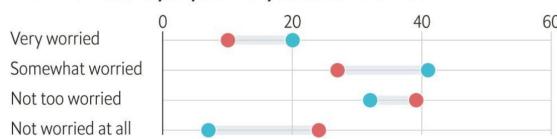
Information Visualization - Stacked Bar chart

Democrats are more worried than Republicans about covid-19

United States, March 2020, %

Democrats Republicans

"How worried are you personally about coronavirus?"



"Have you done any of the following things to prevent the spread of coronavirus?"

- Change travel plans
- Work from home
- Wear a face mask

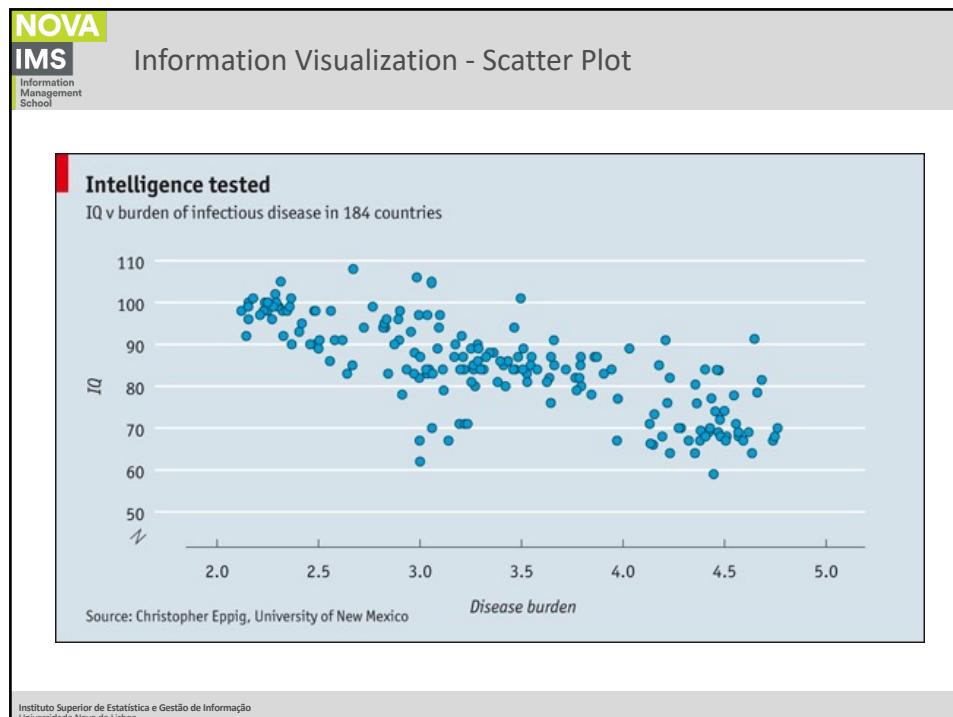
*1,500 adults polled March 8th-10th

Source: YouGov/The Economist

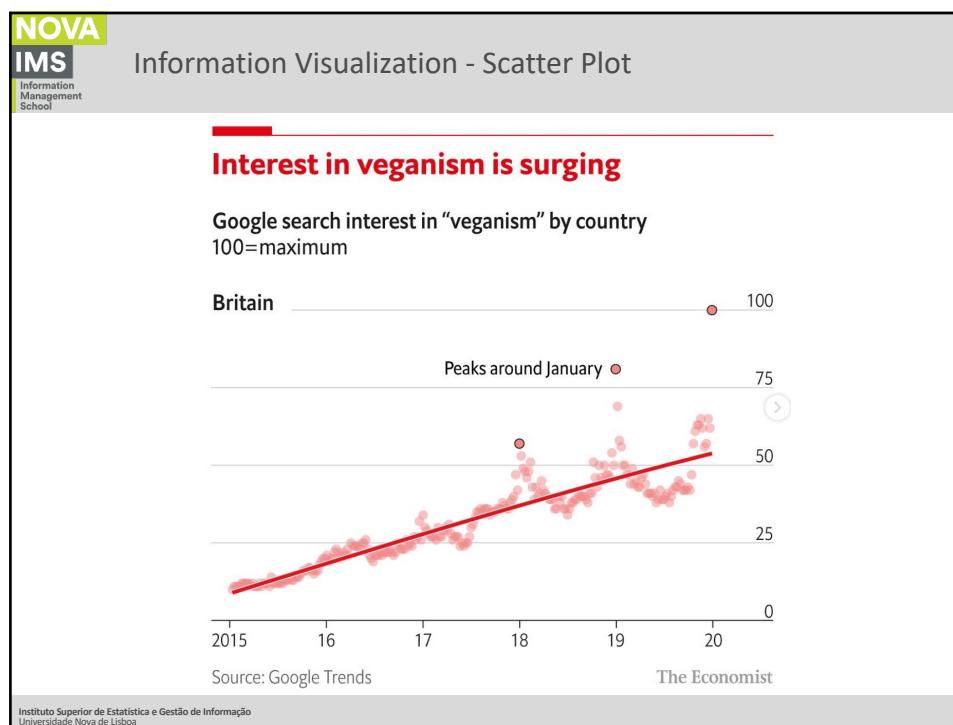
The Economist

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

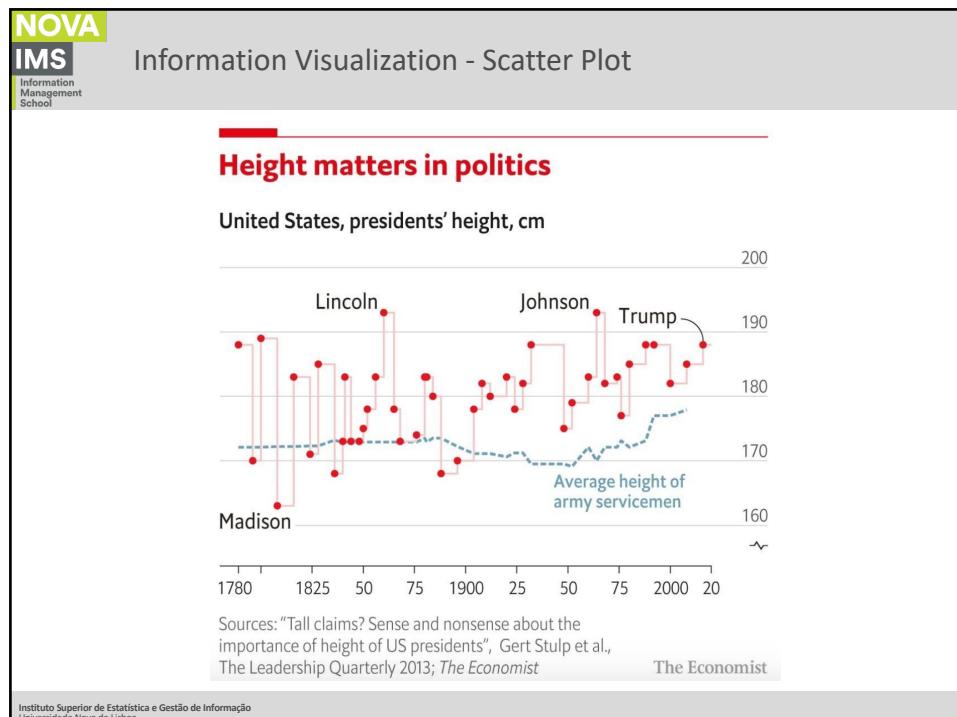
62



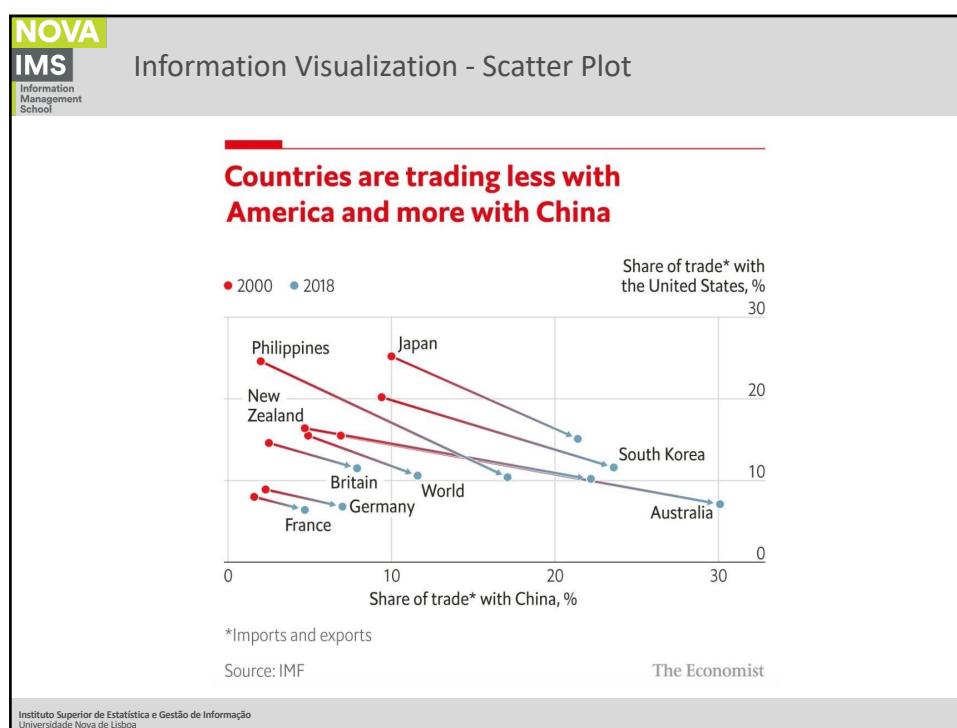
63



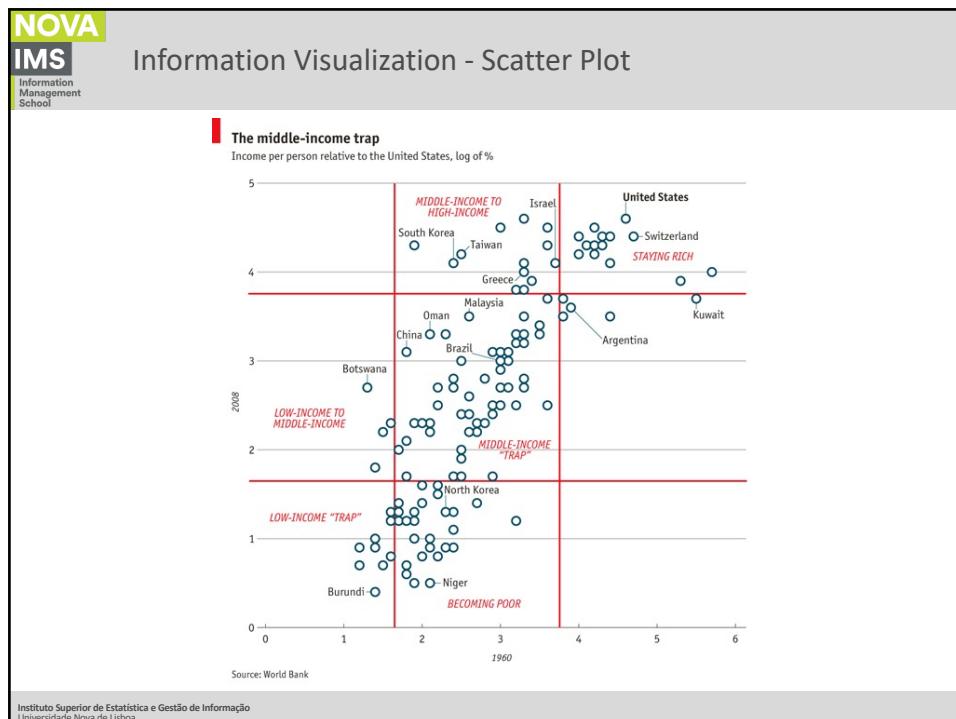
64



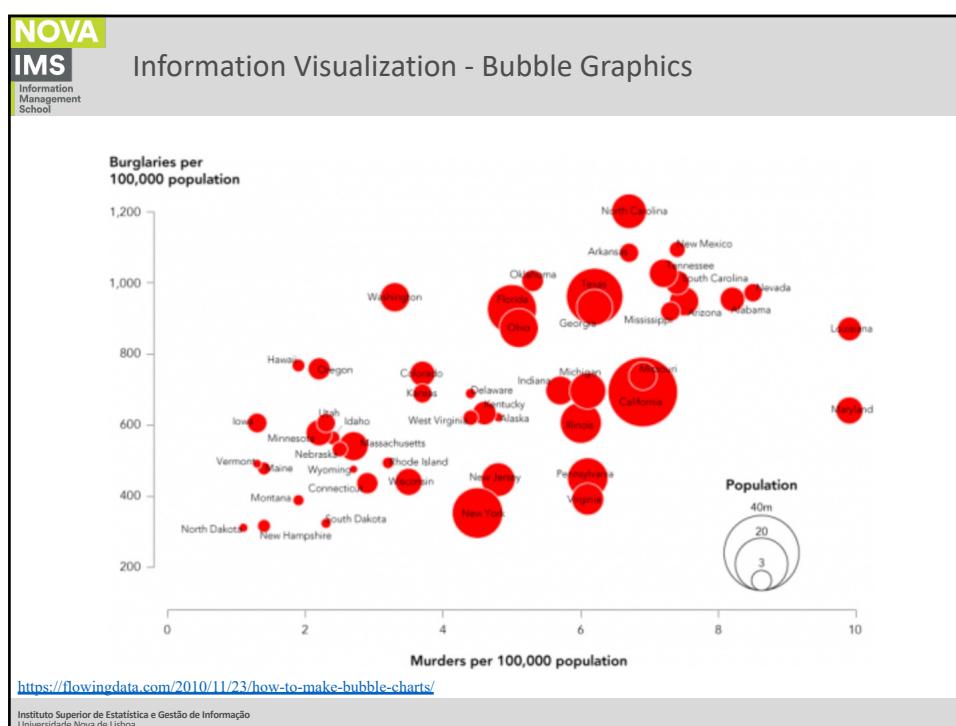
65



66



67



68

Information Visualization - Bubble Graphics

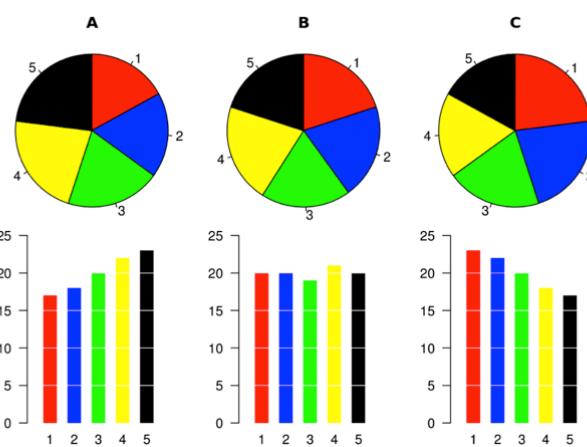


Source: https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

69

Information Visualization – Pie Charts



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

70

NOVA
IMS
Information Management School

Information Visualization – Pie Charts

A Pie Chart, Labeled

Group	Percentage
Group B	40%
Group C	25%
Group D	17%
Group A	10%
Group E	7%
Group F	1%

Pie Chart Alternative: A Bar or Column Chart

Group	Percentage
Group B	40%
Group C	25%
Group D	17%
Group A	10%
Group E	7%
Group F	1%

**Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa**

71

NOVA
IMS
Information Management School

Information Visualization – Pie Charts

Two Pie Charts for Comparison

Shares of Aggregate Income, 1962 and 2007

Aggregate income, by source

Source	1962 (%)	2007 (%)
Social Security	30%	36%
Earnings	28%	29%
Asset income	15%	16%
Private pensions	3%	9%
Government employee pensions	6%	8%
Other	16%	3%

Source: Social Security Administration (2009).

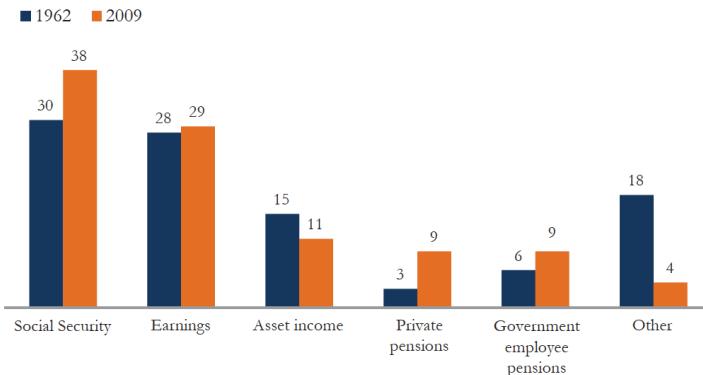
**Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa**

72

Information Visualization – Pie Charts

Alternative to a Pie Chart: A Paired Column Chart

Shares of Aggregate Income, 1962 and 2009
(Percent)



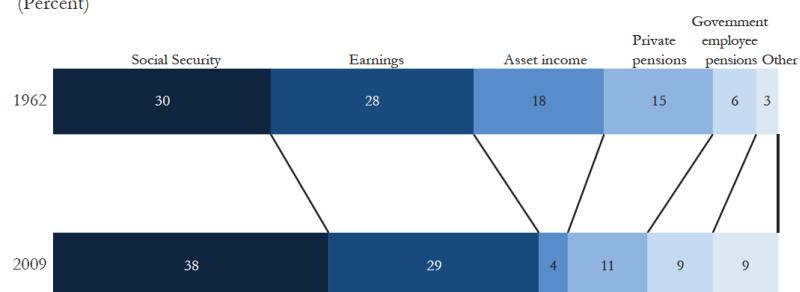
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

73

Information Visualization – Pie Charts

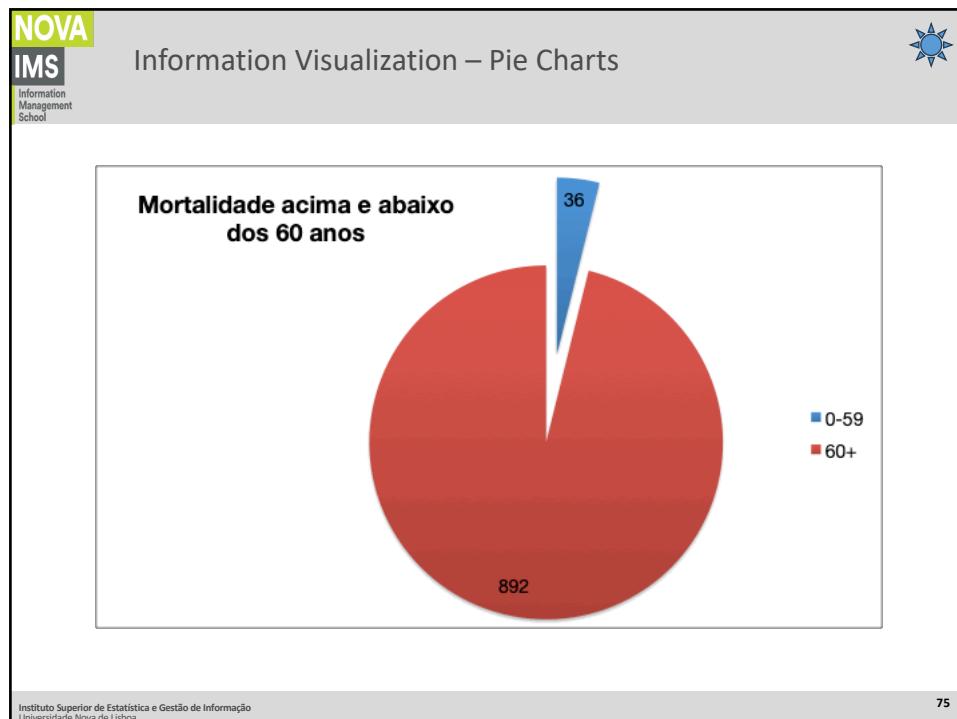
Alternative to a Pie Chart: A Stacked Bar Chart

Shares of Aggregate Income, 1962 and 2009
(Percent)

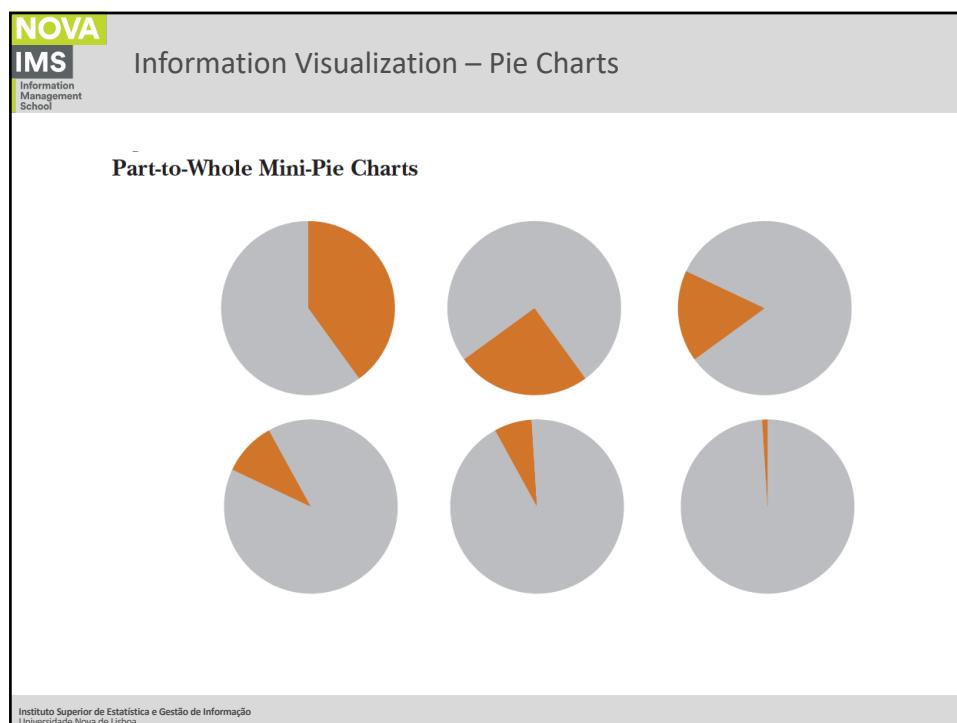


Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

74



75



76

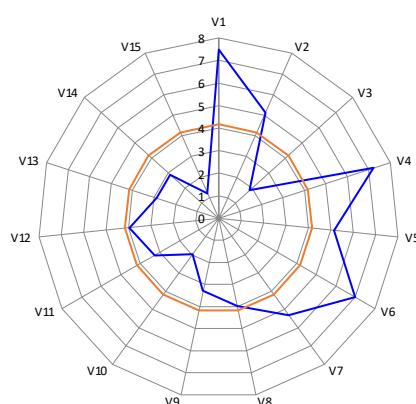
Information Visualization – Pie Charts



Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

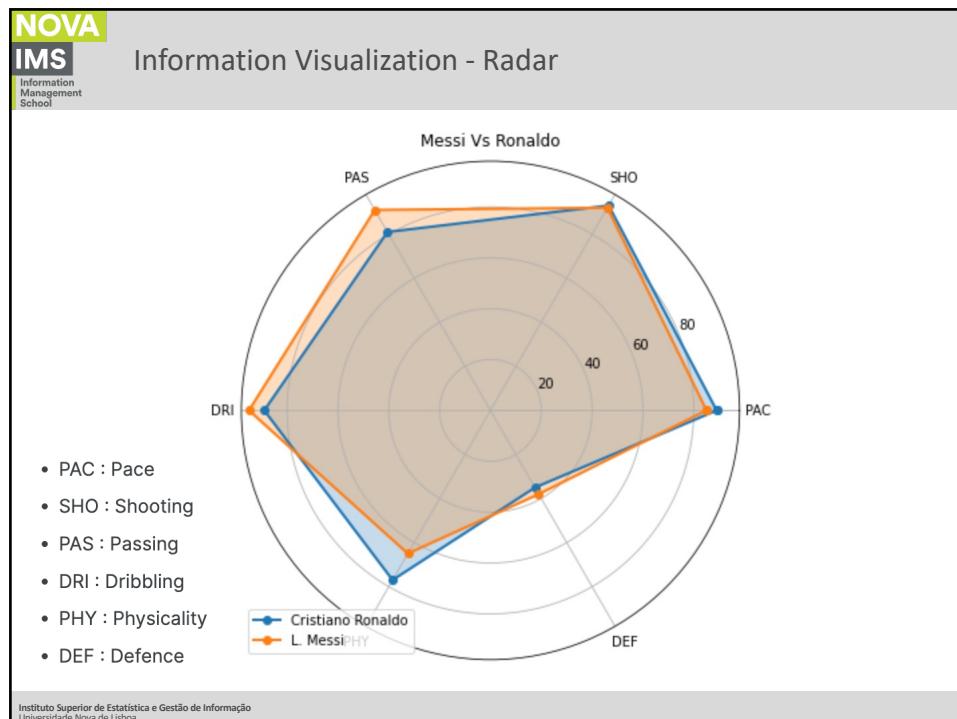
77

Information Visualization - Radar

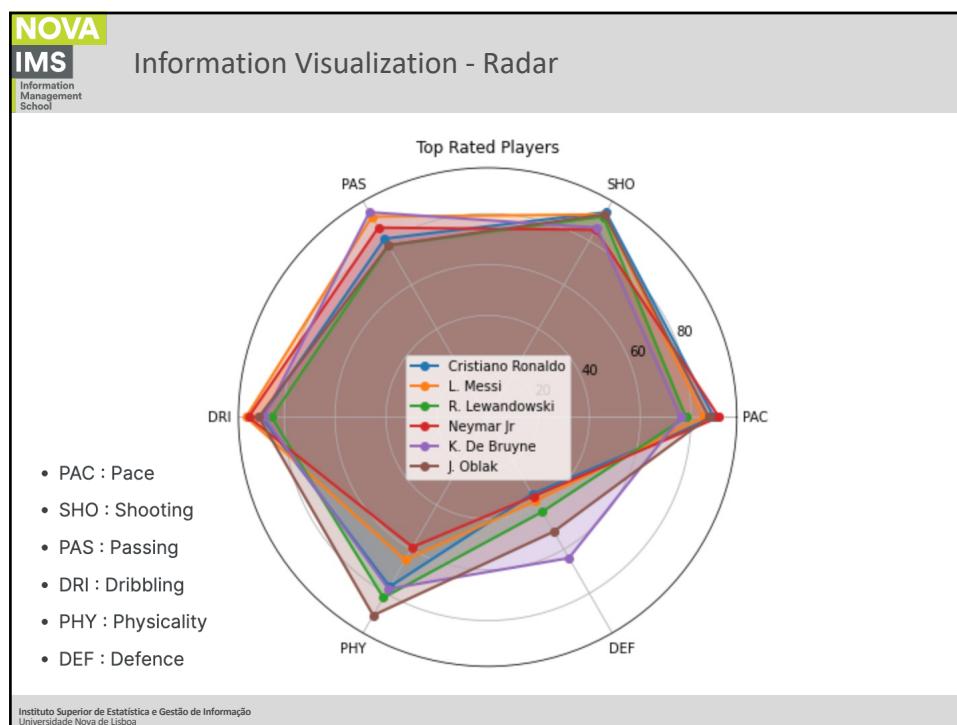


Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

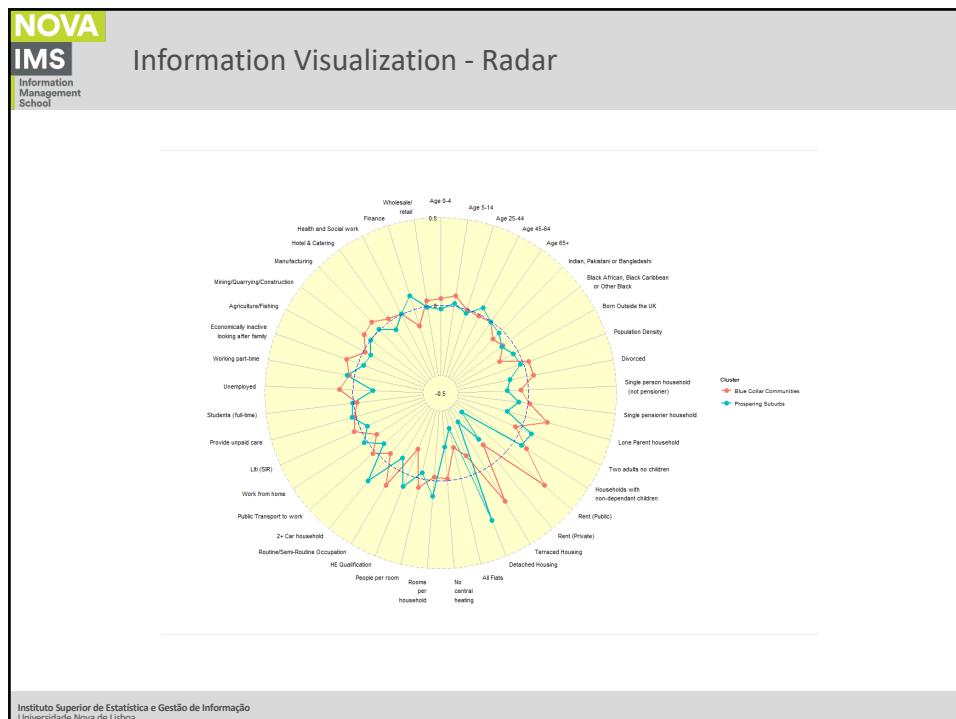
78



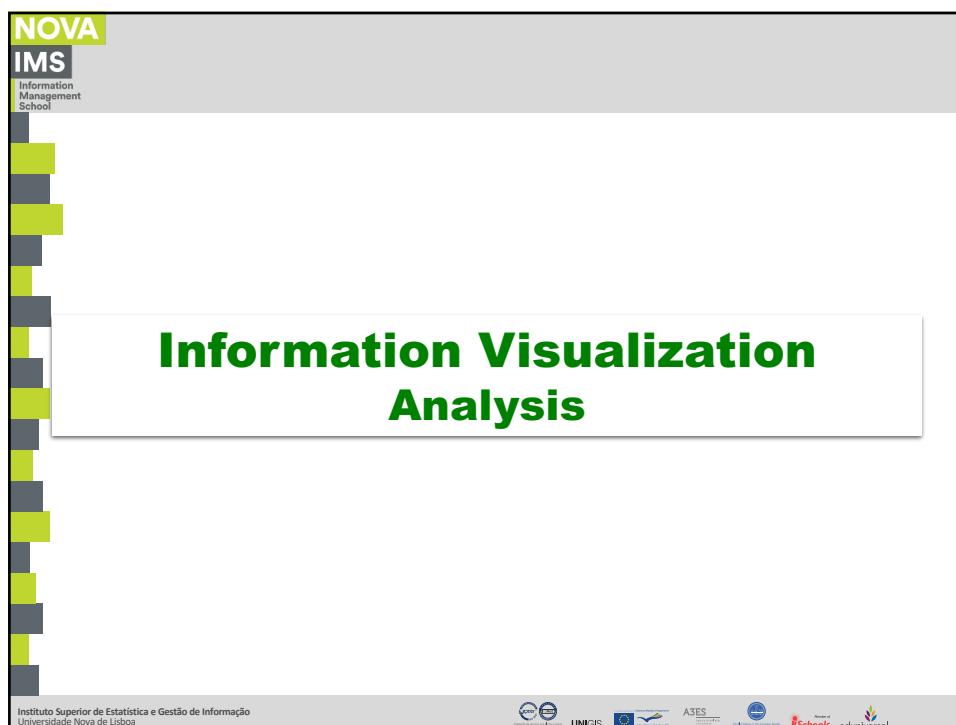
79



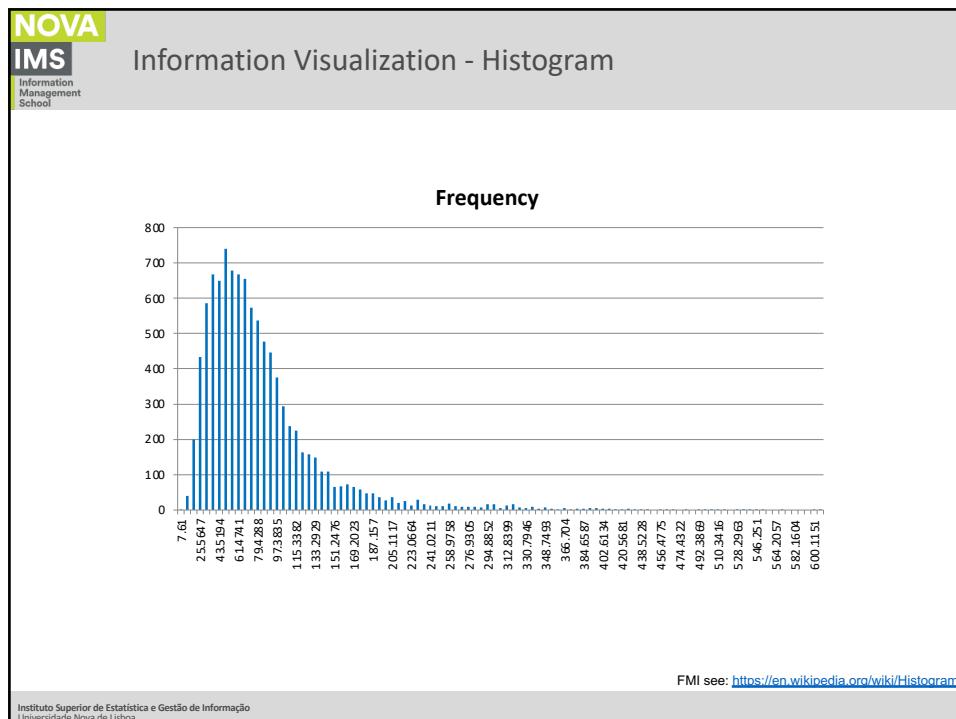
80



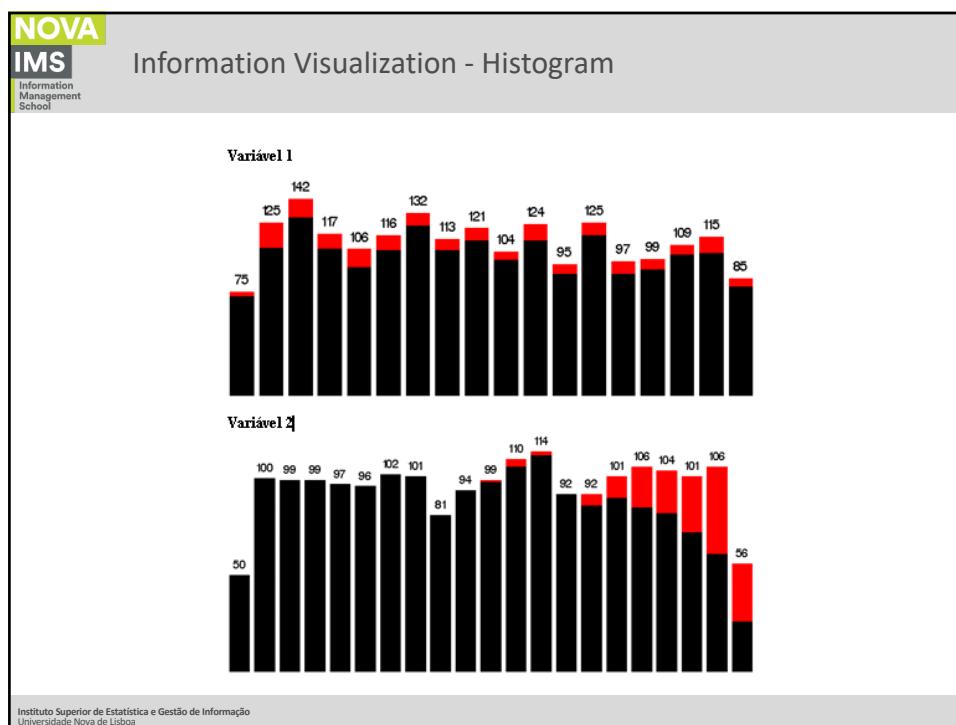
81

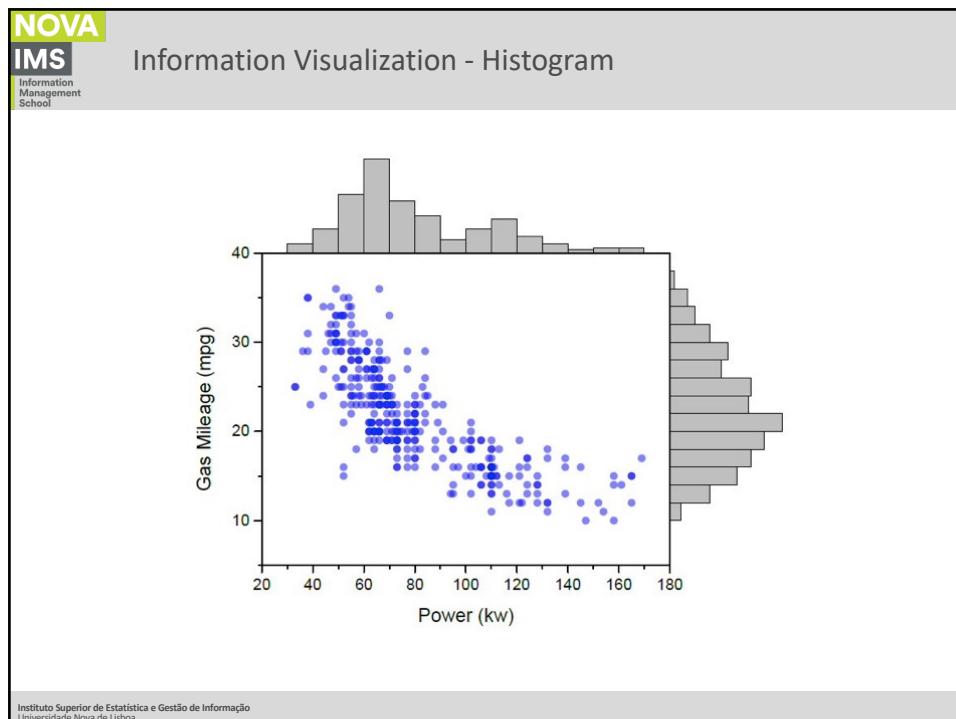


82

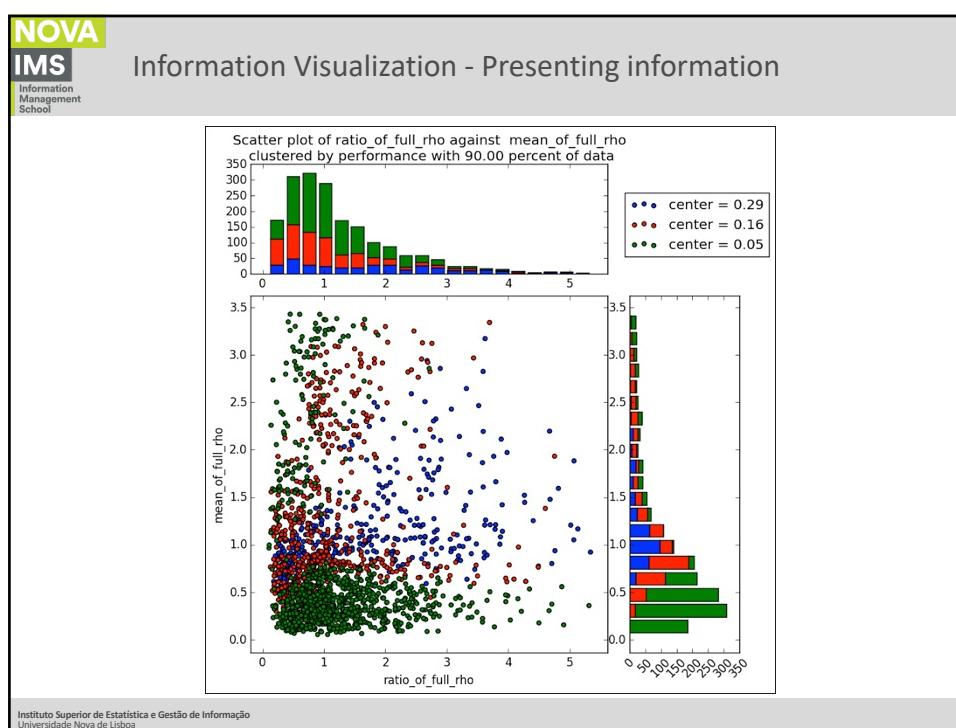


83





85



86

NOVA
IMS
Information Management School

Information Visualization - Boxplot

The diagram illustrates a boxplot with the following labels:

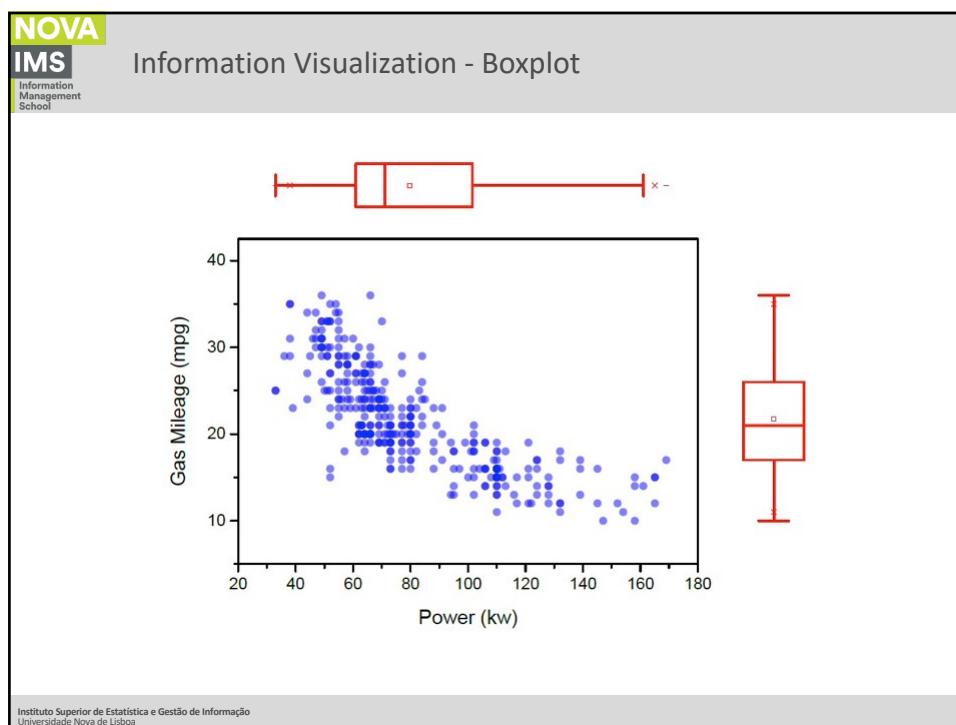
- Q1: Quartile 1, or median of the left data subset after dividing the original data set into 2 subsets via the median (25% of the data points fall below this threshold)
- median: The central vertical line.
- Q3: Quartile 3, median of the right data subset (75% of the data points fall below this threshold)
- IQR: Interquartile-range, $Q3 - Q1$
- Whiskers: Horizontal lines extending from the box to the minimum and maximum values.
- Outlier: A red square point located outside the whiskers.

Outliers: Data points are considered to be outliers if
value $< Q1 - 1.5 \times IQR$ or
value $> Q3 + 1.5 \times IQR$

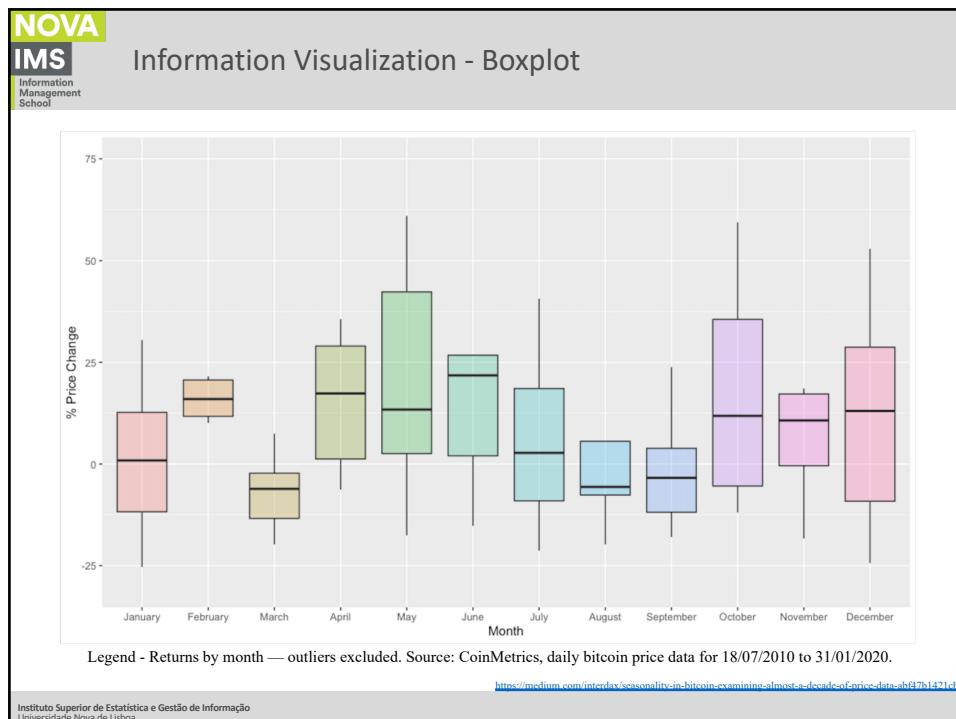
FMI see: https://en.wikipedia.org/wiki/Box_plot

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

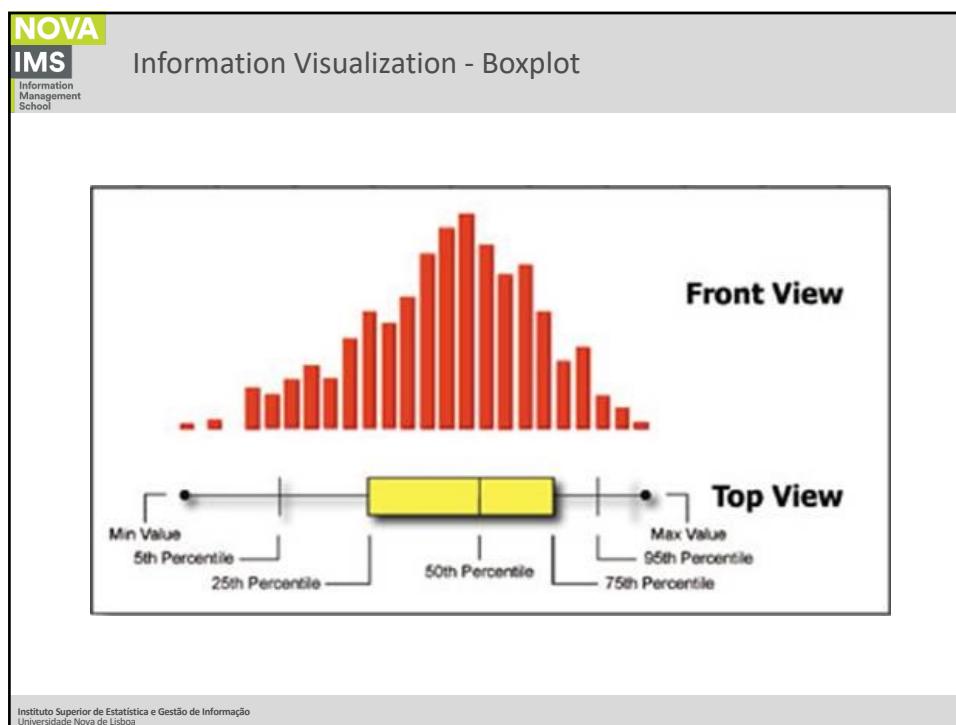
87



88



89



90

NOVA
IMS
Information Management School

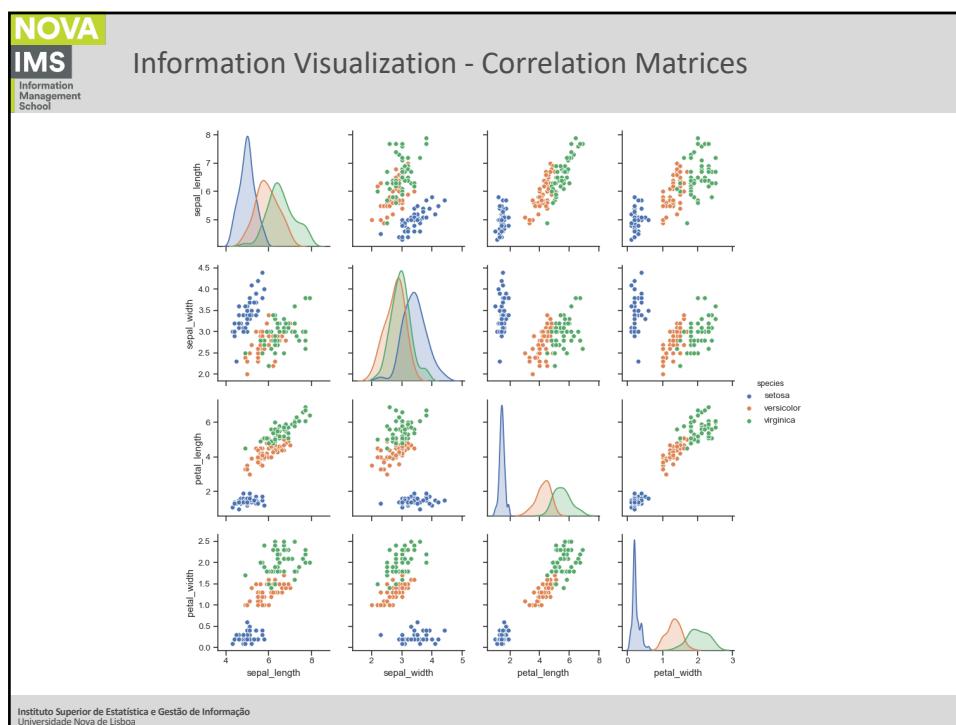
Information Visualization Correlation Matrices

FMI see: <https://builtin.com/data-science/correlation-matrix>

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNIGIS A3ES iSchools eduniversal

91



92

NOVA
IMS
Information Management School

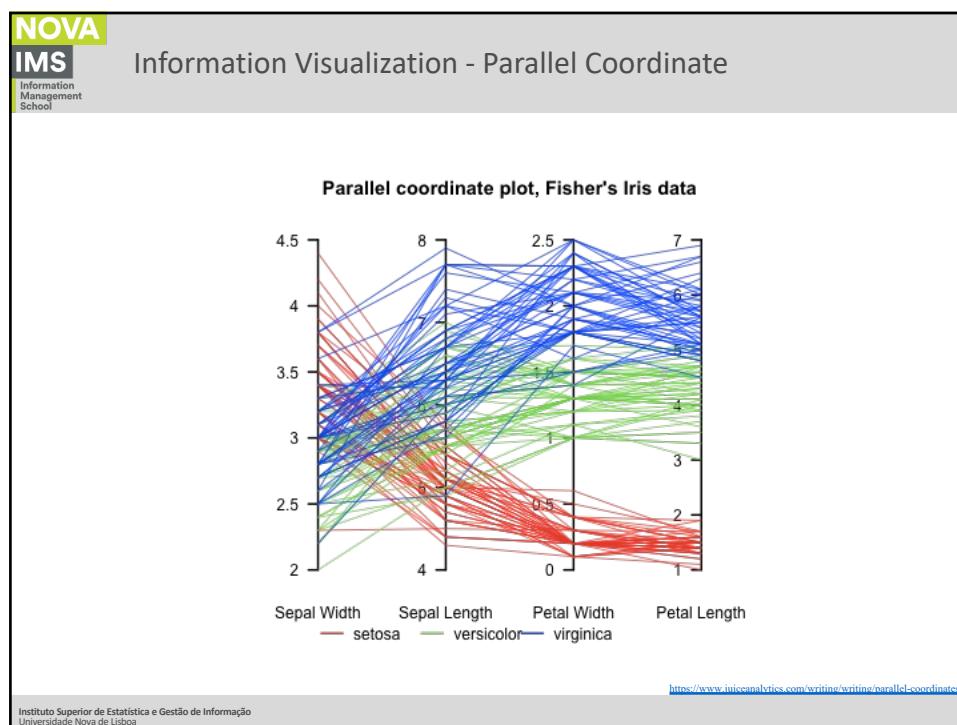
Information Visualization Parallel Coordinates

FMI see: <https://www.iuiceanalytics.com/writing/writing/parallel-coordinates>

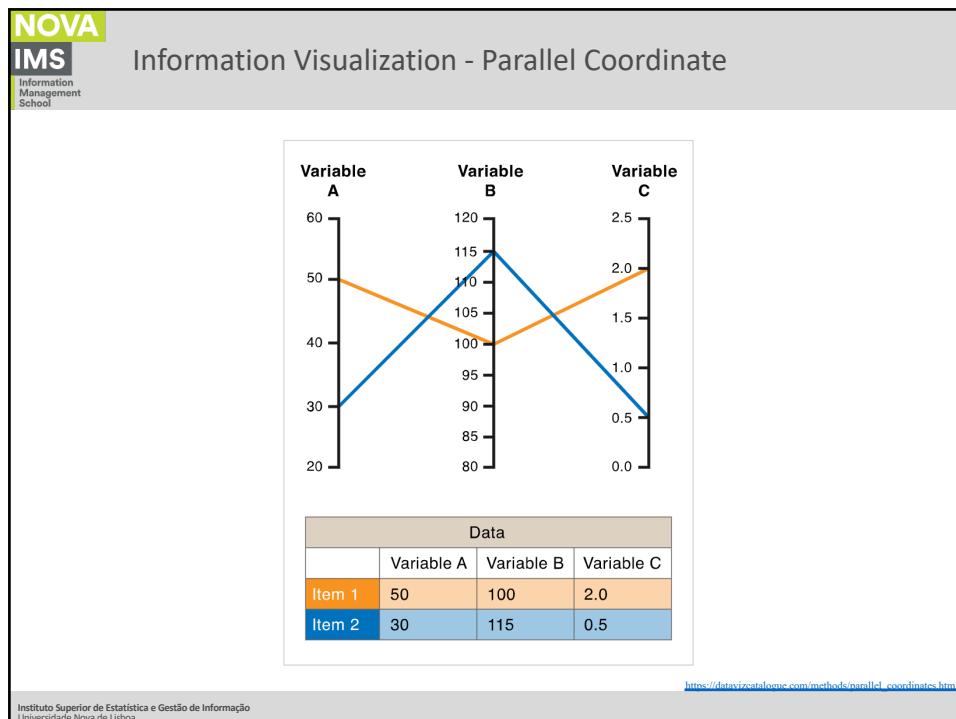
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNIGIS ASES iSchools eduniversal

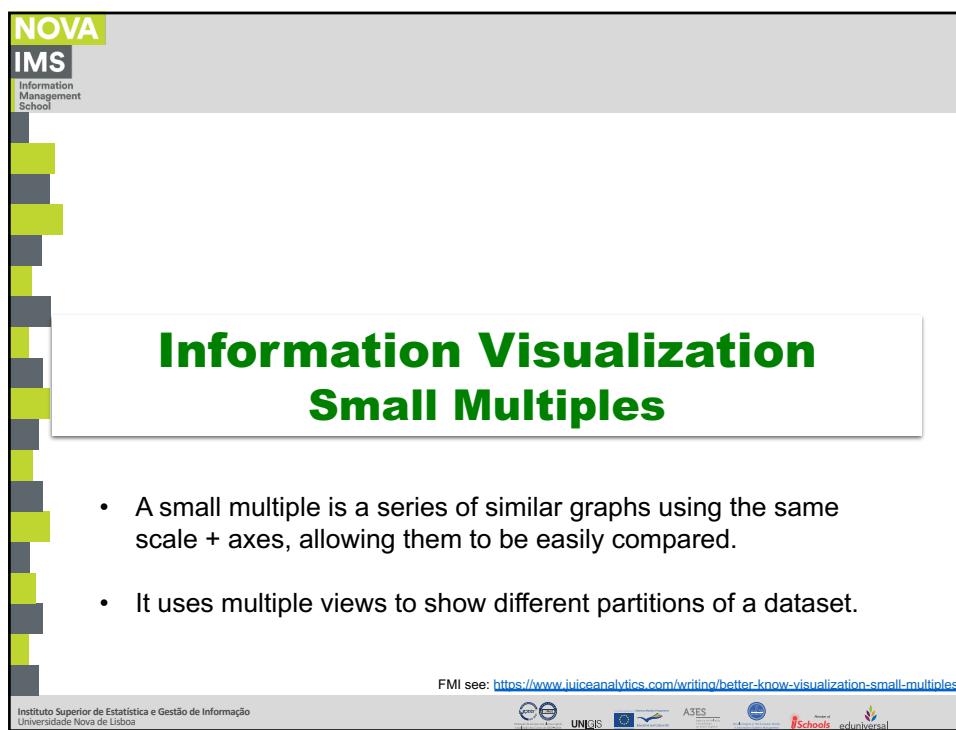
93



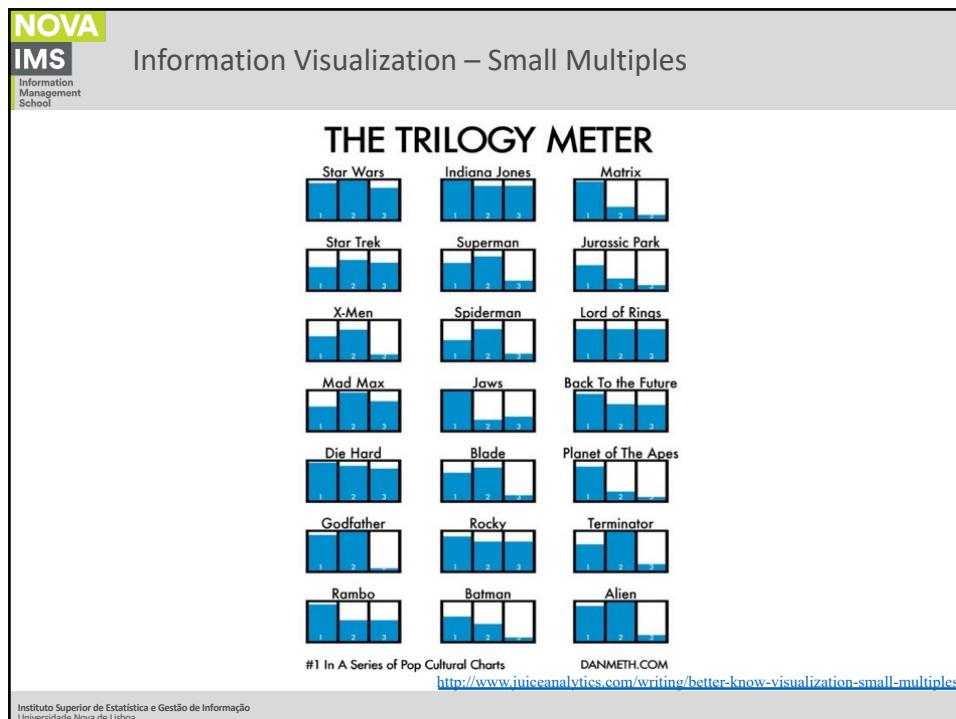
94



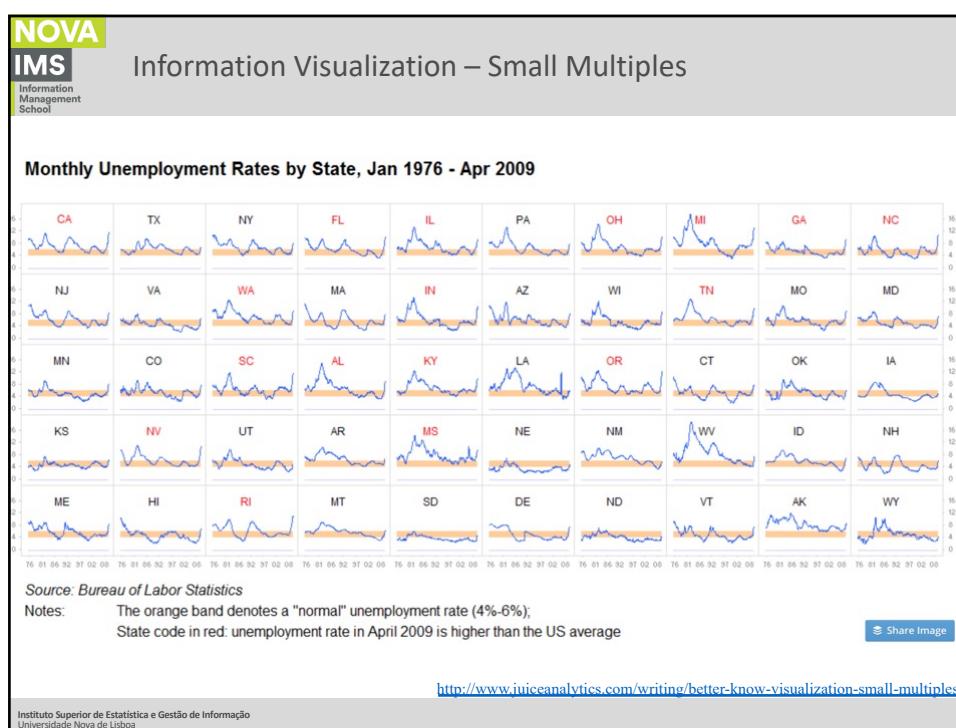
95



96



97



98

NOVA
IMS
Information Management School

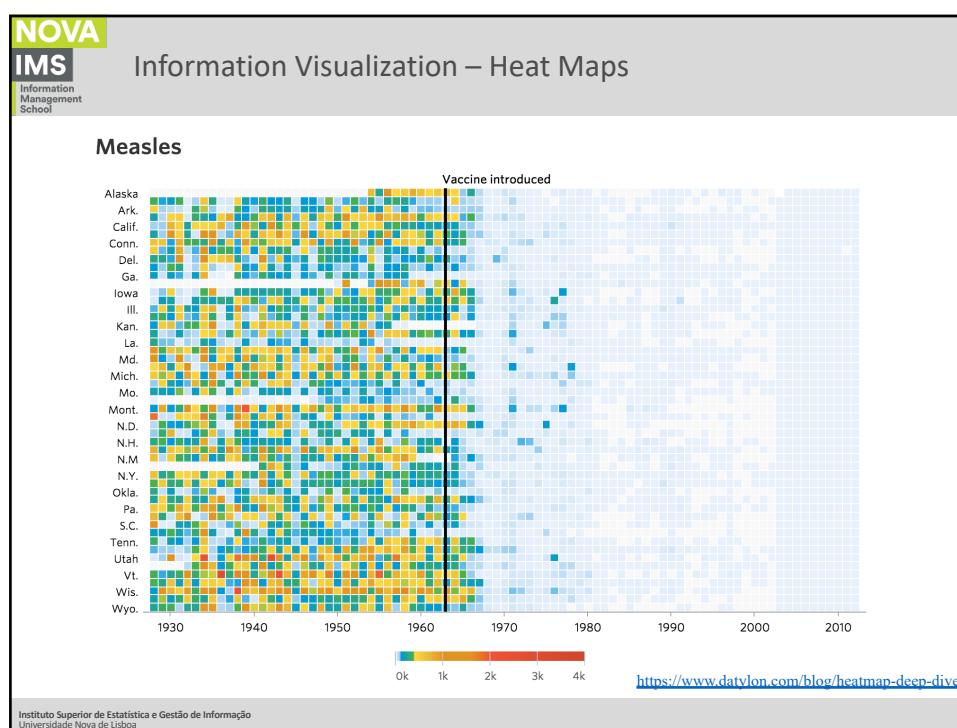
Information Visualization Heat Maps

FMI see: <https://www.datylon.com/blog/heatmap-deep-dive>

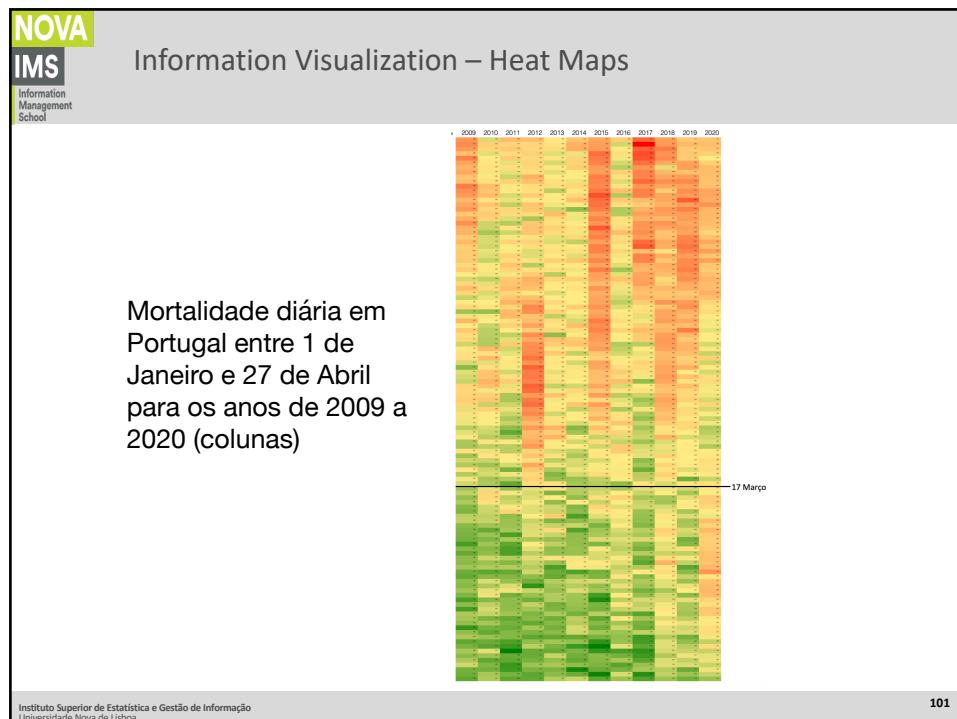
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNIGIS A3ES iSchools eduniversal

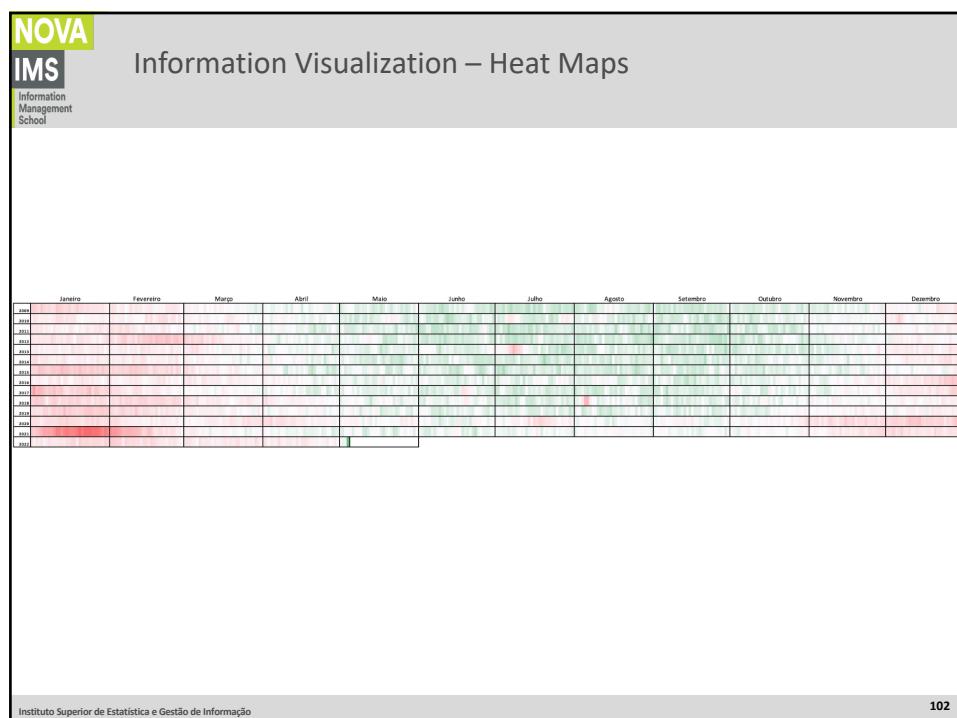
99



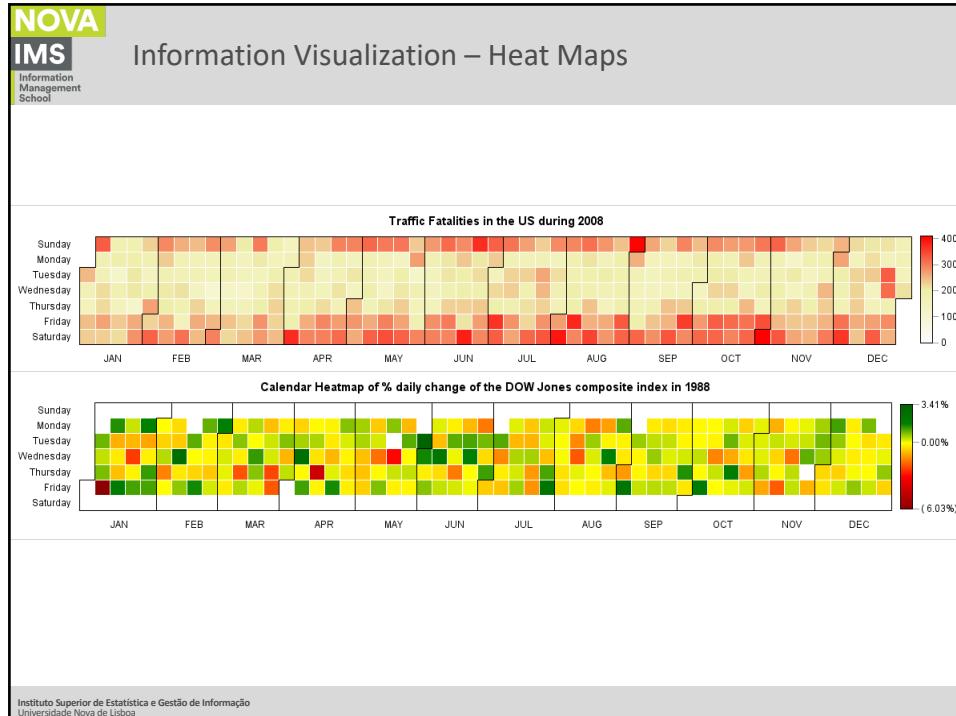
100



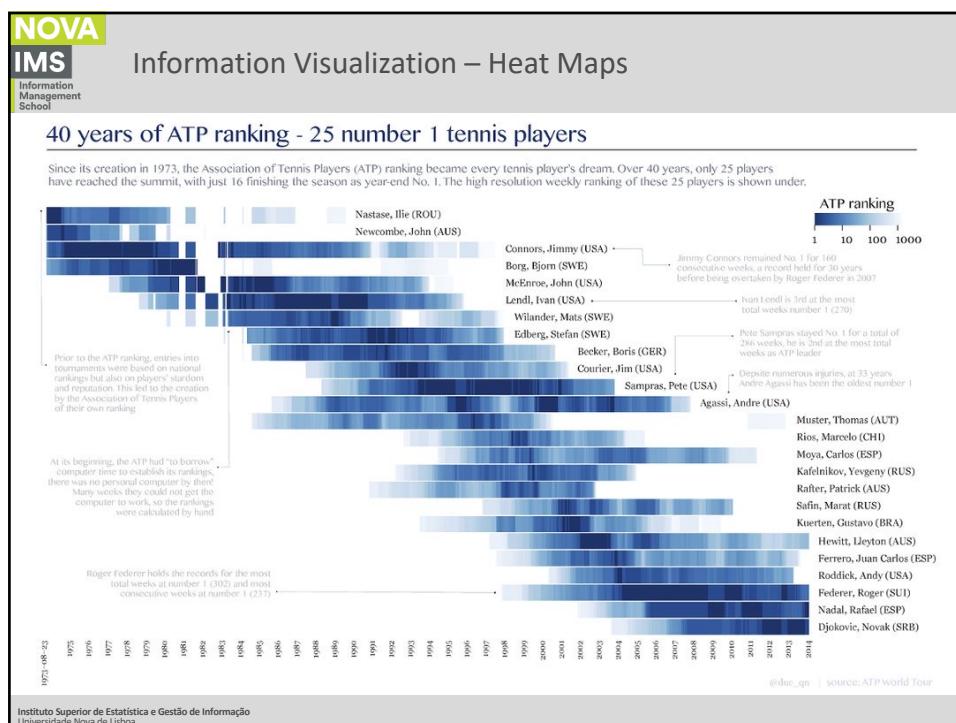
101



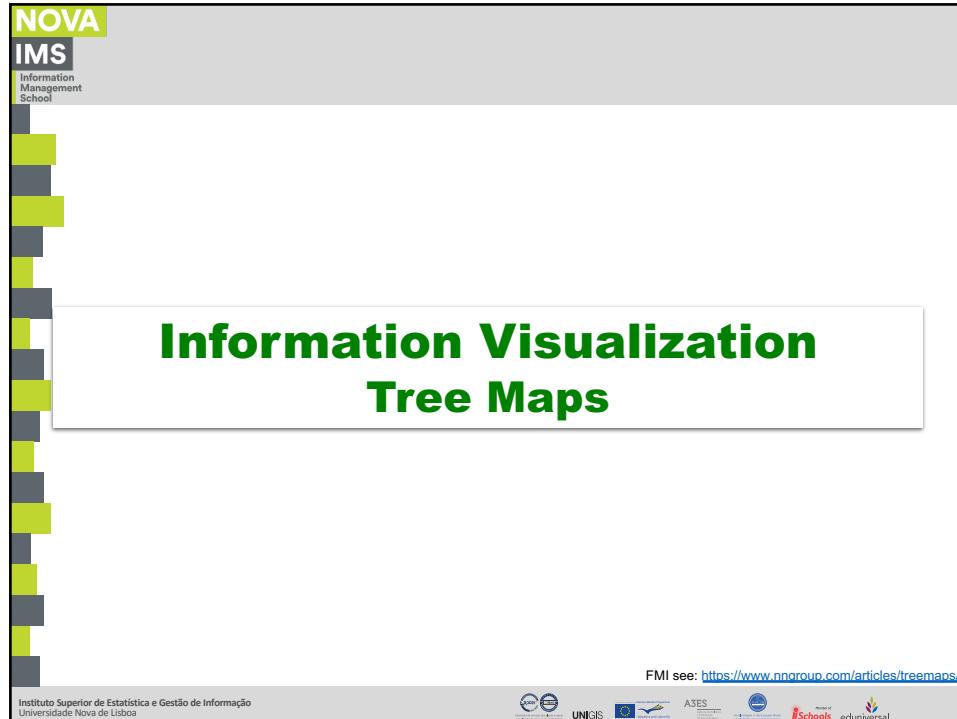
102



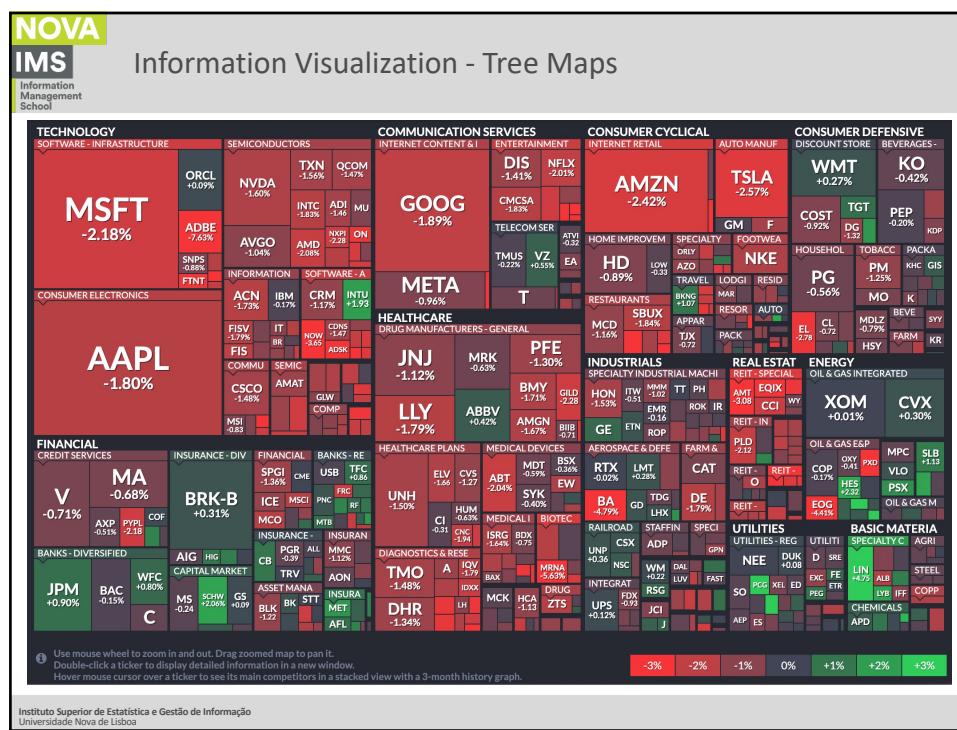
103



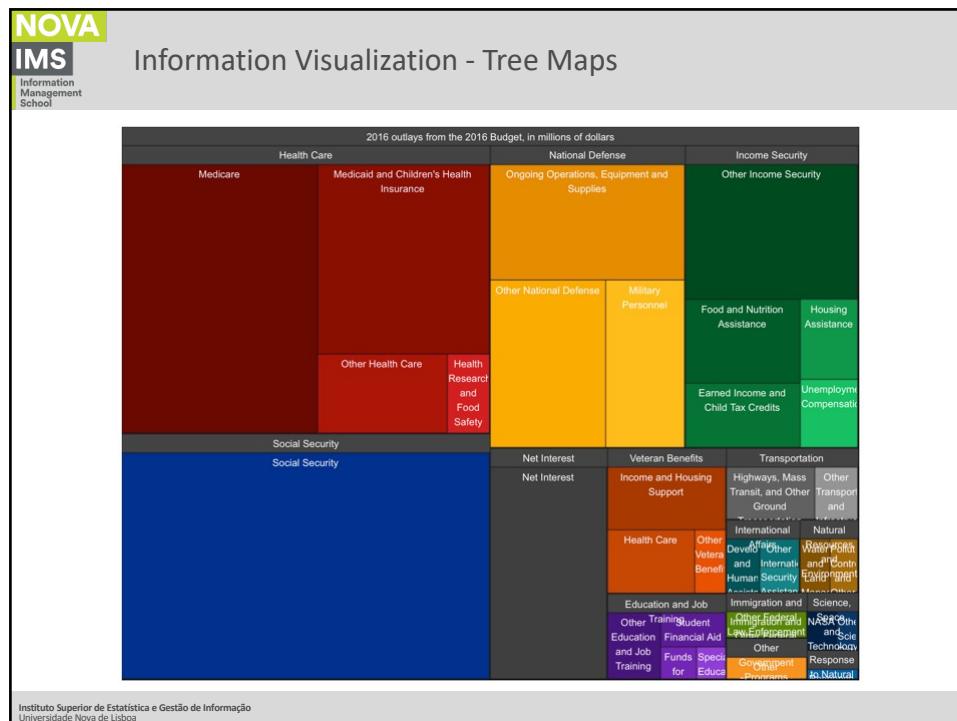
104



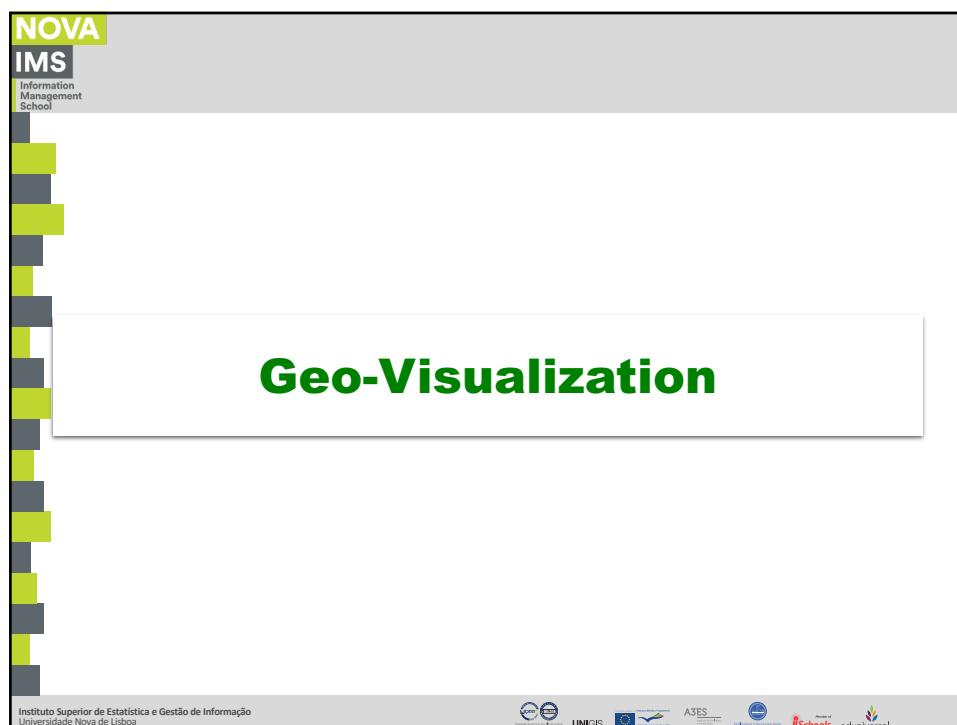
105



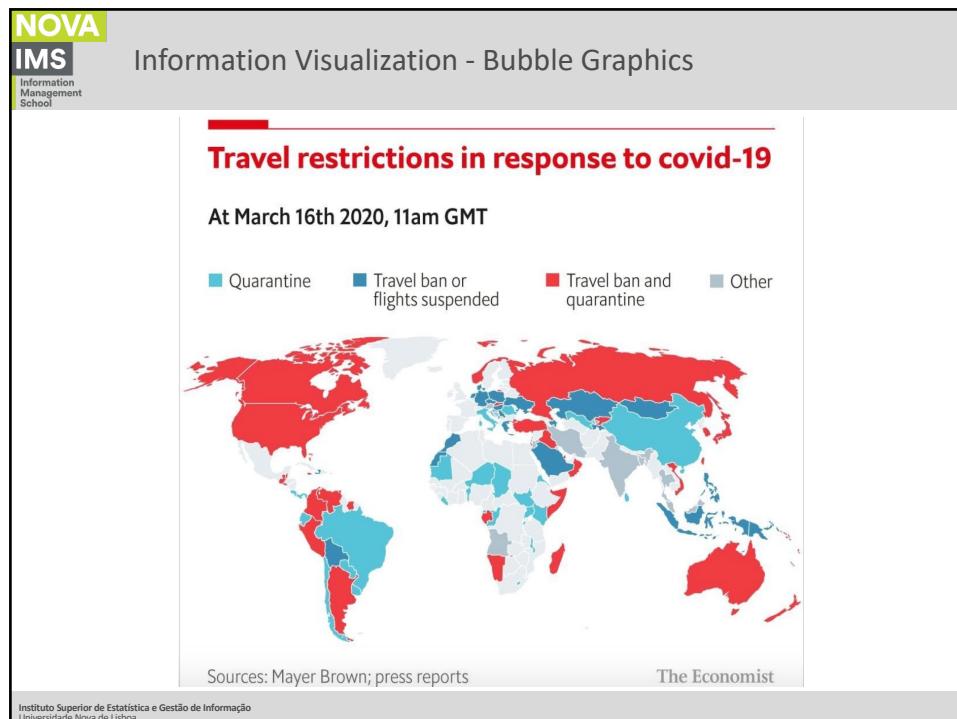
106



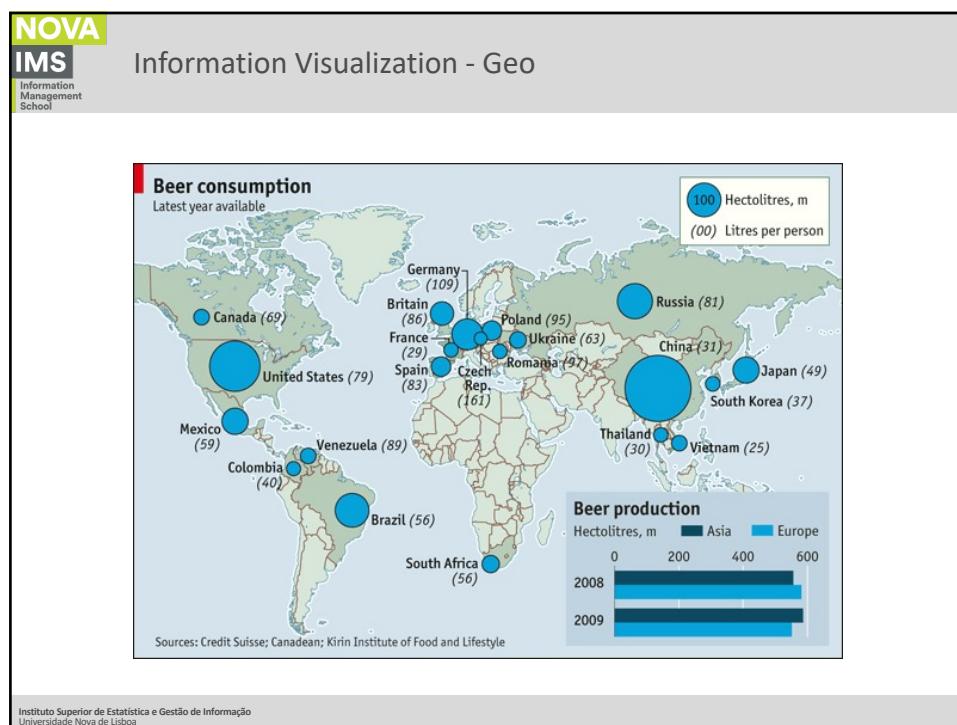
107



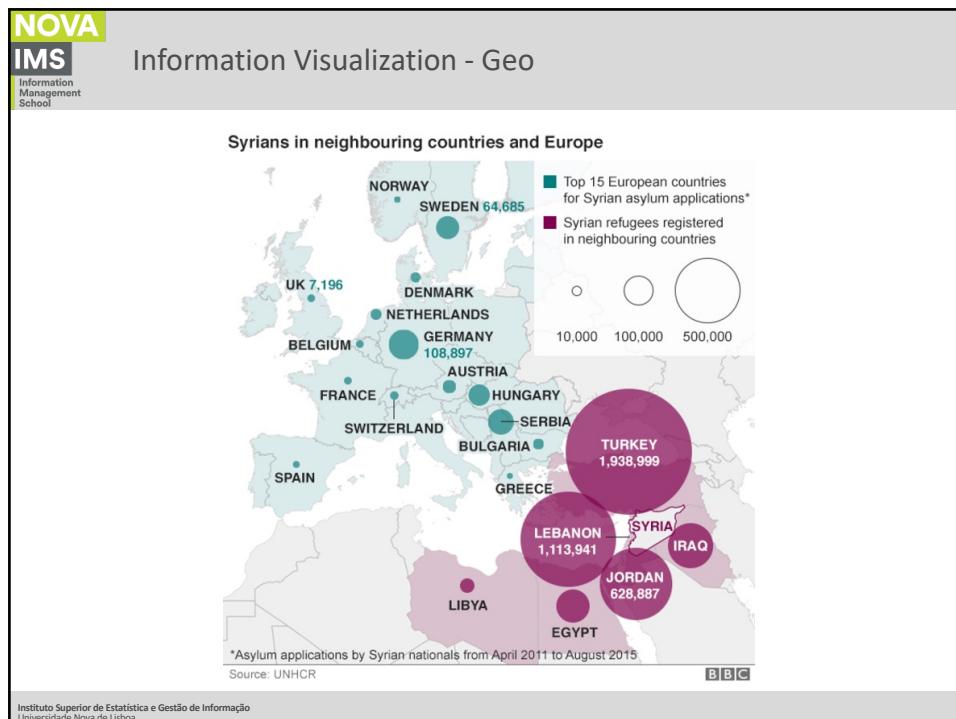
108



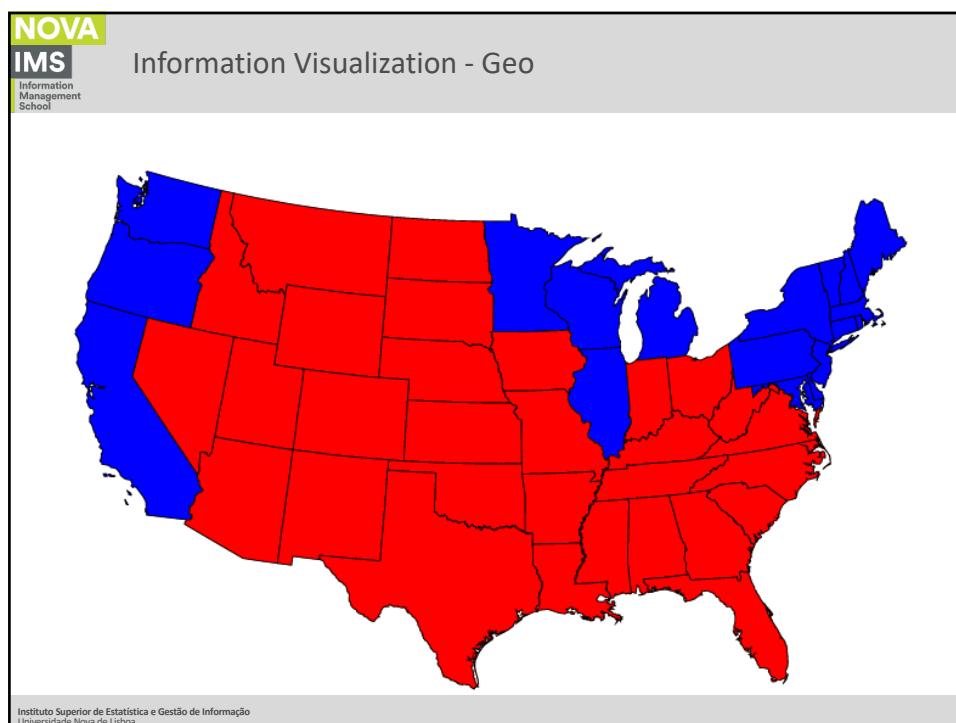
109



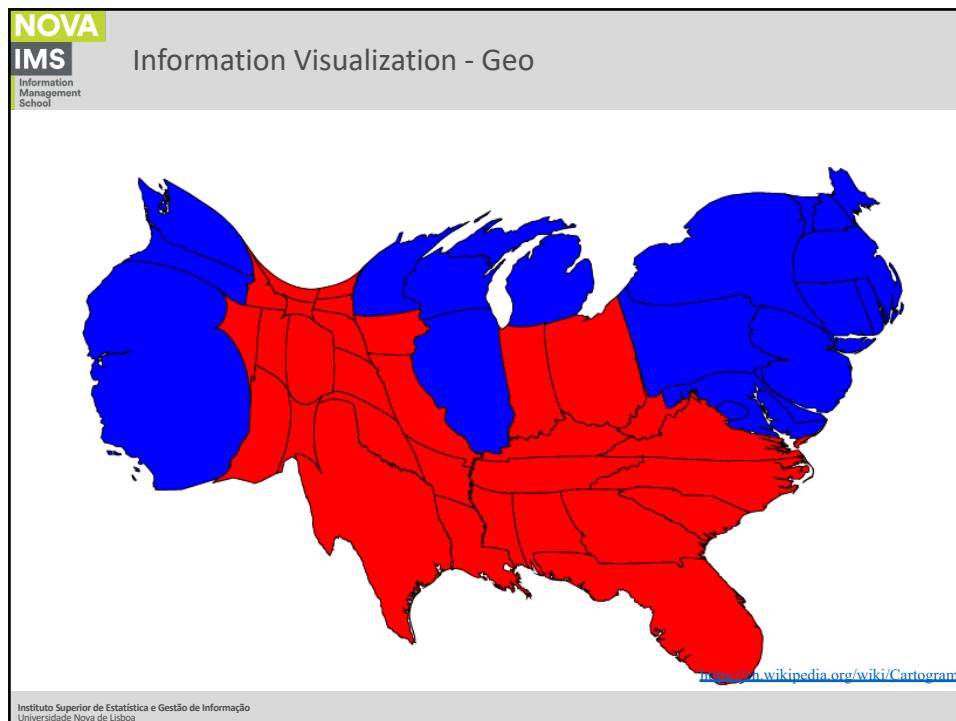
110



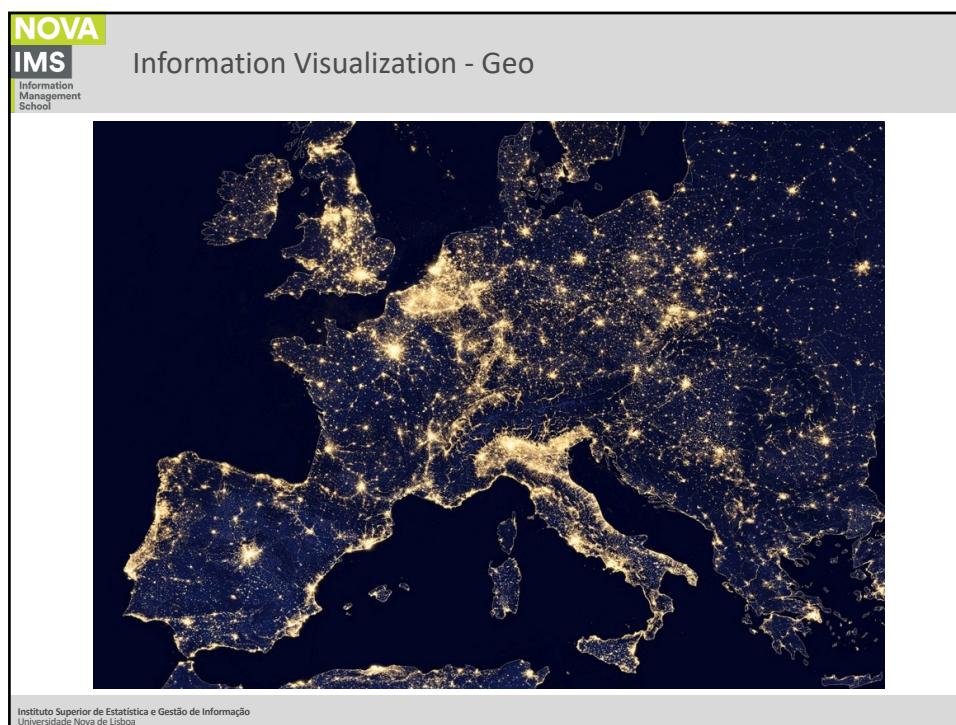
111



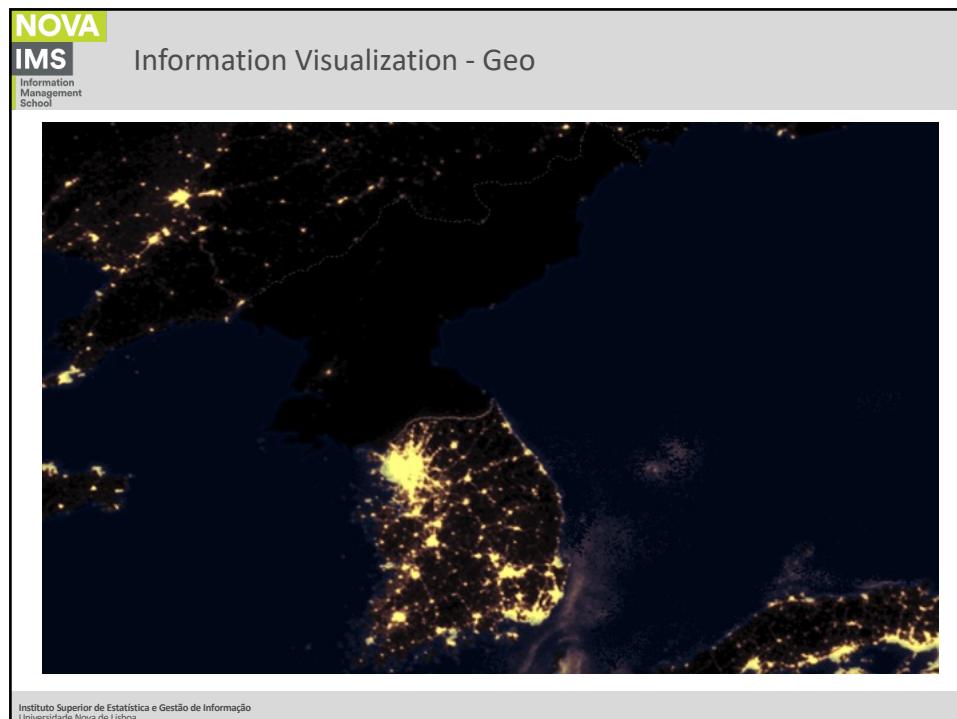
112



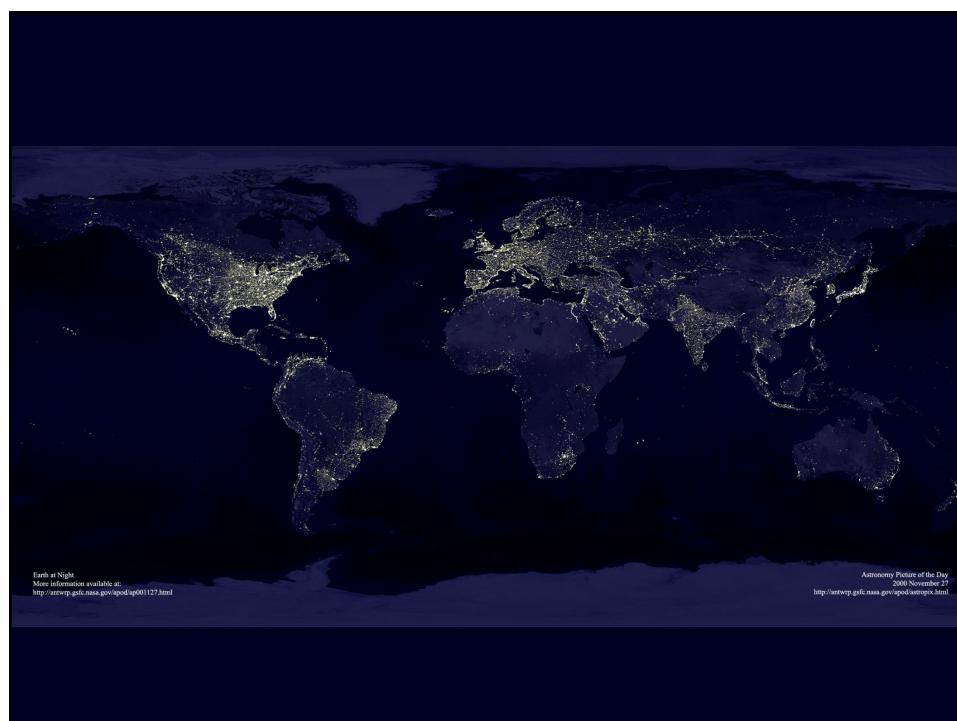
113



114



115



116

NOVA
IMS
Information Management School

Information Visualization Linked Views

FMI see: https://courses.ischool.berkeley.edu/i247/f05/readings/Baldonado_MultipleViews_AV100.pdf

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNIGIS ASES iSchools eduniversal

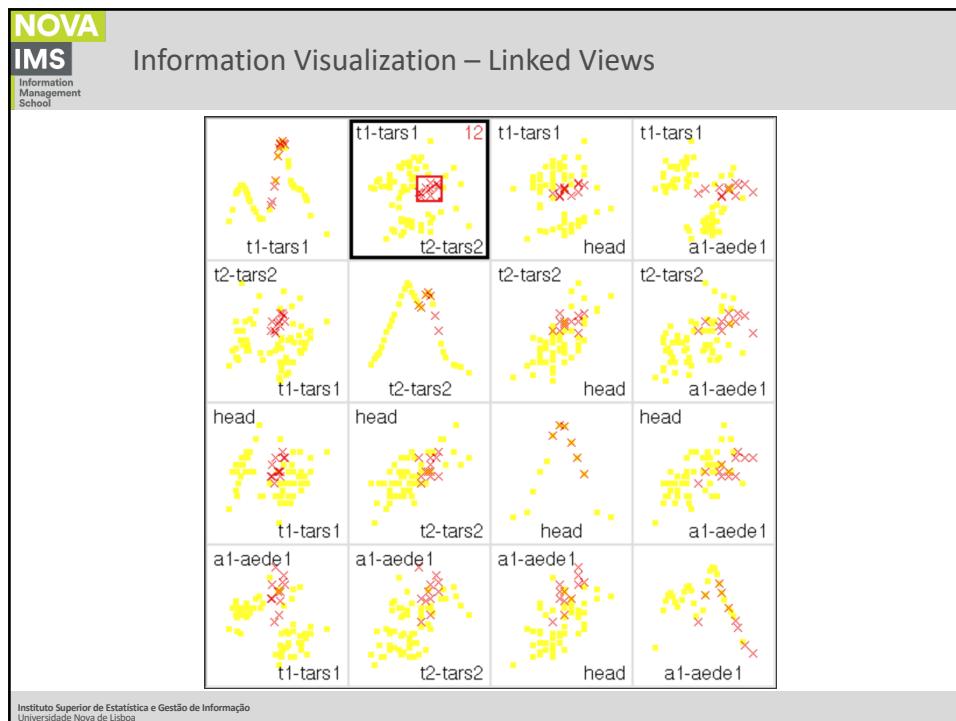
117

NOVA
IMS
Information Management School

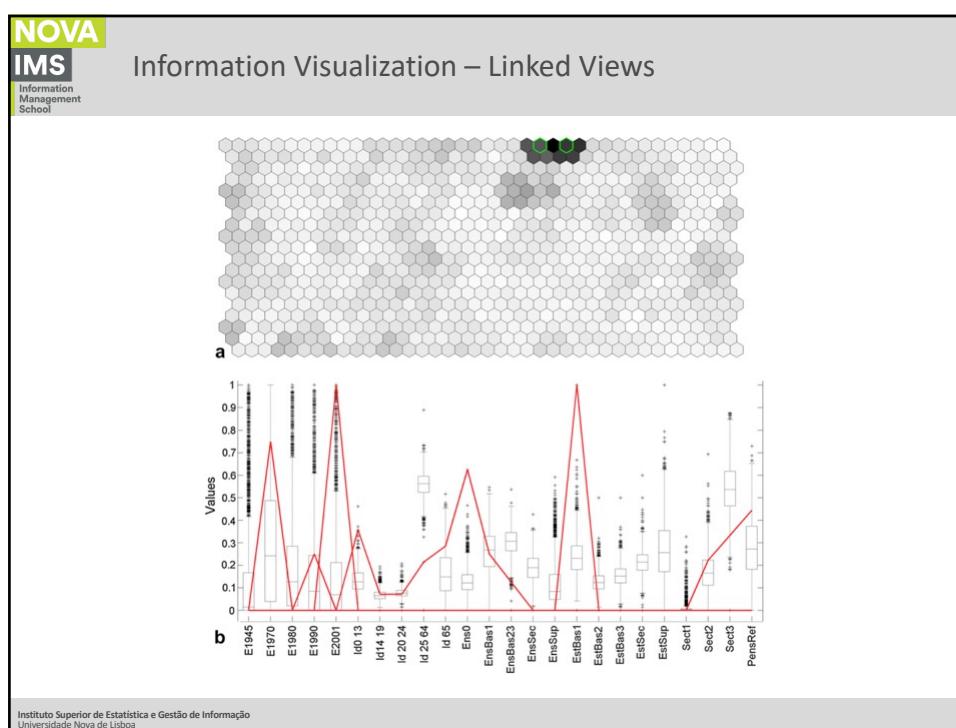
Information Visualization – Linked Views

Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

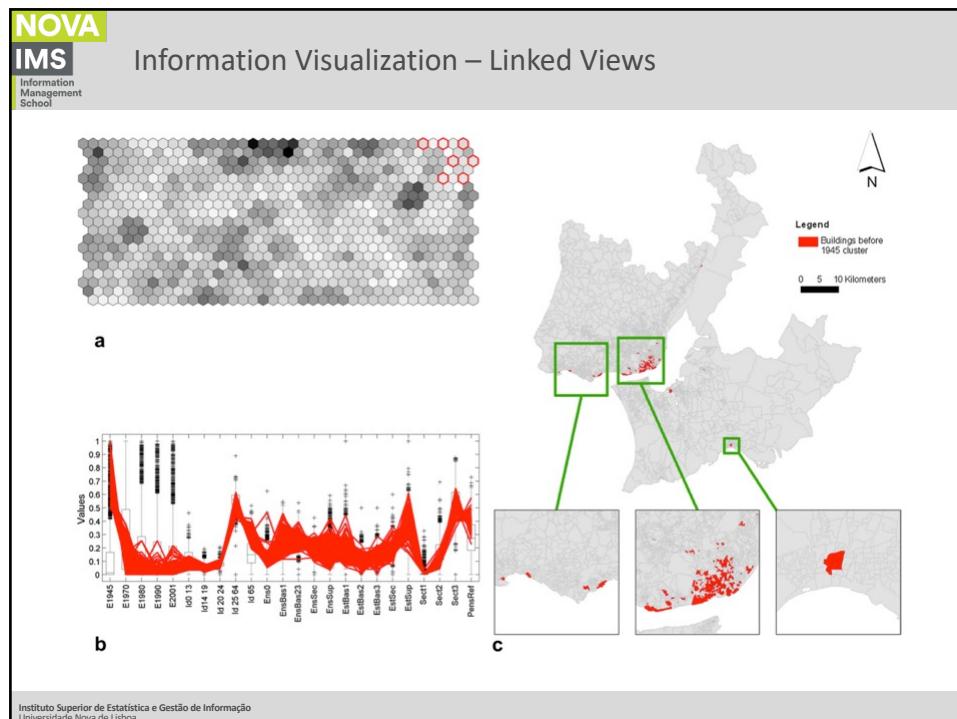
118



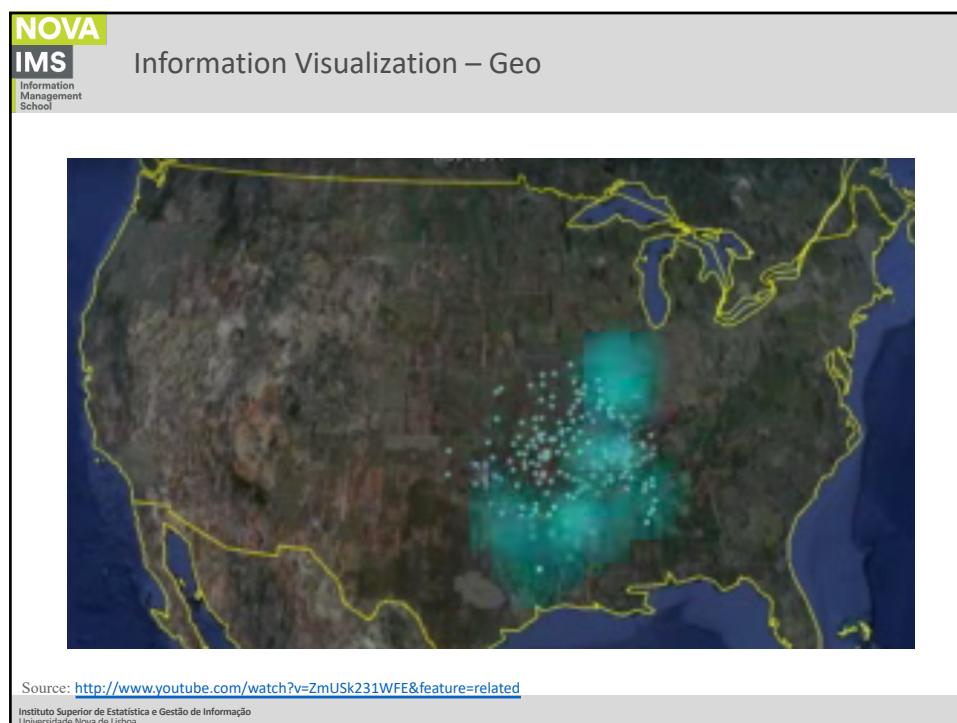
119



120



121



122

122

122



Questions?

123

123