Data Visualisation

Associate Professor Goh Wooi Boon

College of Computing and Data Science
Nanyang Technological University

email: *aswbgoh@ntu.edu.sg*

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

1

1

Chapter 1 – Introduction

Contents

- What is Data Visualisation?
- Importance of Data Visualisation
- What will You Learn in this Course

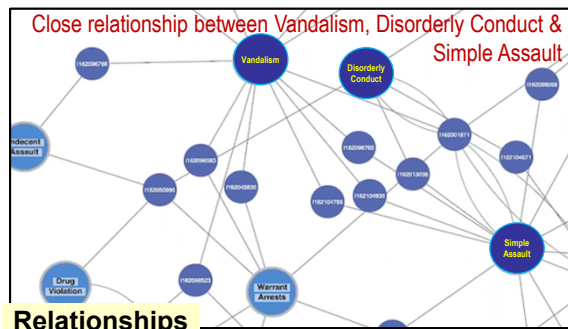**NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE**

© A/P Goh Wooi Boon (CCDS/NTU)

2

2

Introduction

What is Data Visualisation?

- Data visualisation is the **representation** of **data** or **information** in a graph, chart, map or other **visual format**.

- **Visual elements** are used to provide an accessible way to **see** and **understand** trends, outliers, patterns and relationships in the data.

Close relationship between Vandalism, Disorderly Conduct & Simple Assault

Relationships

Crime Visualisation from Cambridge Intelligence.com

Outliers

Image from Analyticsvidhya.com

COVID-19 IN SINGAPORE
TOTAL CASES AND TOTAL DISCHARGED

Trend
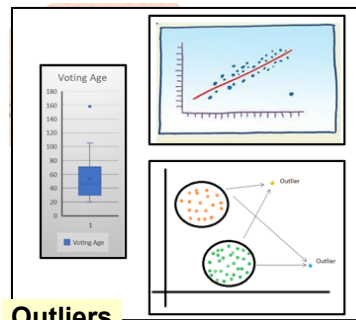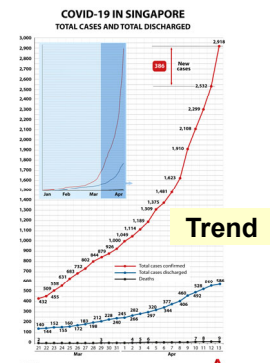
Image from CNA (13 Apr 2020)   3

3

Importance of Data Visualisation

The visual creature

- About half the human brain is directly or indirectly devoted to **visual processing**.

- Humans are much better at remembering **pictures** than words[1]. Students given **visuals** to go along with lectures have also been shown to have **better retention** than students listening to lectures without visual aids[2].

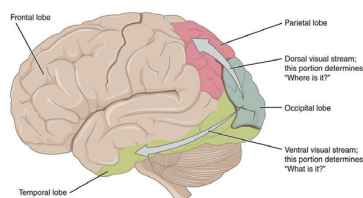- The McGurk effect seems to also suggest that the **visual** sensory modality can **dominate** the auditory modality.
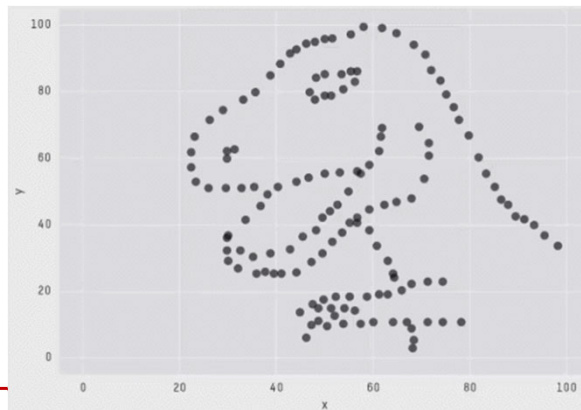

Image from Wikimedia

BBC video on McGurk Effect - https://youtu.be/G-IN8vWm3m0

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

[1] Neural correlates of the episodic encoding of pictures & words - https://www.pnas.org/content/95/5/2703.full

[2] Bui, D.C. & McDaniel, M.A. (2012) – Enhancing learning during lecture ... https://www.sciencedirect.com/science/article/abs/pii/S2211368115000200

4

4

Importance of Data Visualisation

Blinded by statistics

- Basic statistical analysis alone may not provide useful insights into how datasets can have **differing characteristics** despite having **similar statistical** parameters.

12 datasets of 2D points with similar statistical parameters


```
X Mean: 54.2659224
Y Mean: 47.8313999
X SD  : 16.7649829
Y SD  : 26.9342120
Corr. : -0.0642526
```

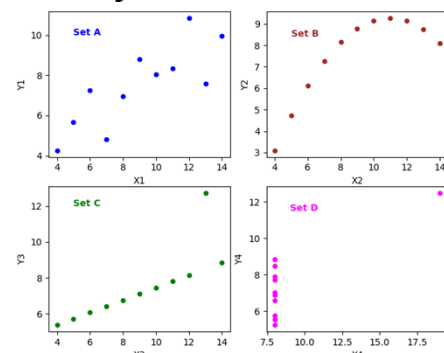DataSaurus Dozen by
Matejka & Fitzmaurice (CHI2017)

5

5

Importance of Data Visualisation

Blinded by statistics

- Basic statistical analysis alone may not provide useful insights into how datasets can have **differing characteristics** despite having **similar statistical** parameters.

- Statistician Francis Anscombe[3] used the **Anscombe's Quartet**[3] of four datasets to show why **visualisation** is a **crucial part of data analysis**.

	Set A		Set B		Set C		Set D	
	X1	Y1	X2	Y2	X3	Y3	X4	Y4
	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	8.1	14	8.84	8	7.04
	6	7.24	6	6.13	6	6.08	8	5.25
	4	4.26	4	3.1	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
Correlation Coefficient	0.82		0.82		0.82		0.82	
Std Dev	3.16	1.94	3.16	1.94	3.16	1.94	3.16	1.94

Anscombe's Quartet

[3] Anscombe, F. J. (1973). "Graphs in Statistical Analysis". American Statistician. 27 (1): 17–21
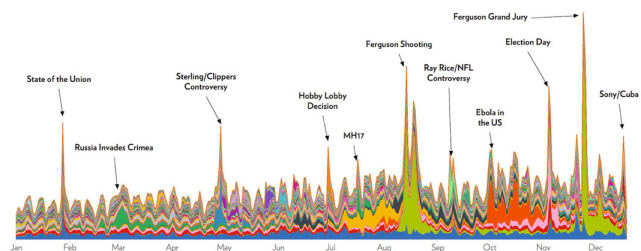
6

6

Importance of Data Visualisation

Seeing insights in big data

- The amount of data being generated is increasing rapidly. Gathering **meaningful insights** from such **big data** is not possible without good data visualisation.
- Good visualization allows users to easily identify relationships and patterns **hidden inside** the large data sets.

Word cloud - 20+ million customers' comments [4]

[4] Big Data Visualization. Value it brings & techniques in requires -
https://www.scnsoft.com/blog/big-data-visualization-techniques

Visualisation by Echelon Insights of what America twitted in 2014

7

7

Importance of Data Visualisation

Telling the story

- Data visualisation can help **communicate the message** embedded in the data.
- Telling a visual story that conveys the **intended** message requires designing clear visualisation that understands how **humans perceive** visual information.
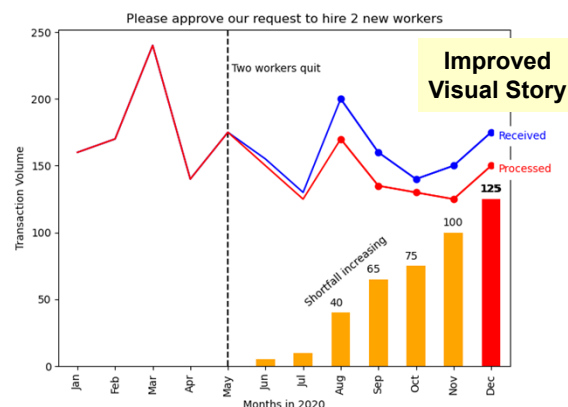
We lost 2 workers in May. We need to persuade management to allow us to hire 2 new workers to cope with the **increasing** processing load [5].

Transactions	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
A	B	C	D	E	F	G	H	I	J	K	L	M
Received	160	170	240	140	175	155	130	200	160	140	150	175
Processed	160	170	240	140	175	150	125	170	135	130	125	150

Excel data of transactions received and processed over year 2020

[5] C.N. Knaflic, Storytelling with data, Wiley (2015)
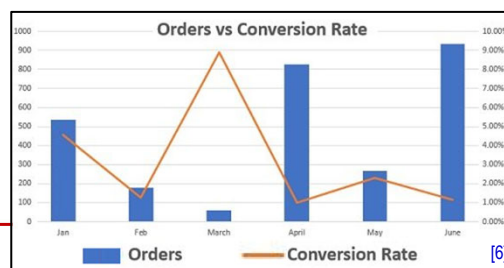
Improved Visual Story

8

8

Visualising Datasets

Is There An Ideal Way?

- There are many ways to visualise a dataset.
- Choosing the best way to visualise the dataset will depend on:
  - The data **types** (e.g. its scale of measure, spatial (geodata), temporal, etc)
  - The **context** and **relationship** within the data (e.g. highest value, part of a whole)
  - The **purpose** of the visualisation (e.g. identify trends, compare values, show composition, communicate a specific message)

[6] Chart from: https://www.optimizesmart.com/what-type-of-chart-to-use-to-compare-data-in-excel/
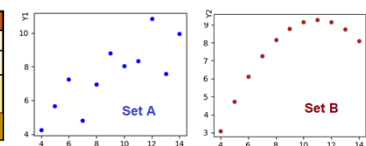
9

9

Visualising Datasets

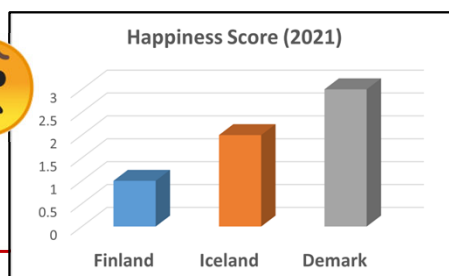### What Is Considered A Good Visualisation?

- A good visualisation shows **connections** within the data that are too complex to explain with words or is **hidden** from non-visual analysis.

	Set A		Set B	
	X1	Y1	X2	Y2
Mean	9.00	7.50	9.00	7.50
Correlation Coefficient	0.82		0.82	
Std Dev	3.16	1.94	3.16	1.94

From Anscombe's Quartet

- A good visualisation makes it **easier** to quickly understand the information presented and grasps meaningful **insights** embedded within the data.
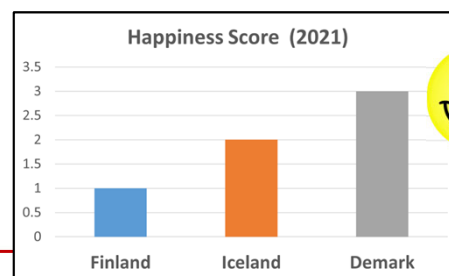
Poor Visual Perception
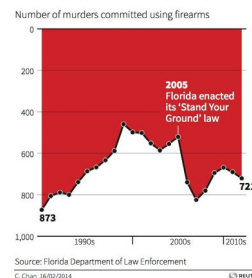

Good Visual Clarity
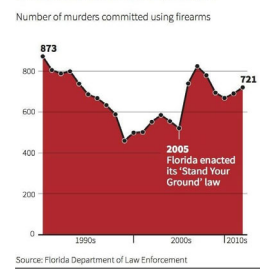
10

10

Ethical Visualisation
Staying True

- Ethics are the set of **moral principles** that guides a person's **decision** and **actions**.
- In data visualisation, there are many opportunities to **manipulate** and **distort** the message a person perceives using misleading or untruthful data representation.
- **Ethical conduct** is an important part of **your responsibility** when you design and present visual information.
- Your **professional integrity** requires you to be mindful of how your audience will interpret the information you present in your visualisation.
- So, communicate **clearly** & **truthfully.**

Reuters Original Version [7]   P.A. Fedewa's Ethical Version [7]

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

[7] BusinessInsider.com, This Chart Shows An Alarming Rise In Florida Gun Deaths After 'Stand Your Ground' Was Enacted - https://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2

11

11

Welcome to Data Visualisation

What will you be learning in this course?

The Person – the theory of human visual perception and its implications in designing effective data visualisation.

The Process – from preparing the data to choosing and implementing the right visualisation solution for displaying and exploring various types of data.

The Practice – solve data visualisation problems through design and code development for different application domains.

The Purpose – learn to design visualisation narratives that effectively communicate your intended purpose and facilitate data analysis.
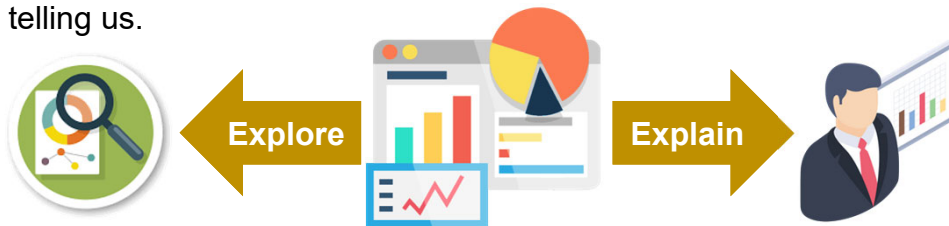
NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

12

12

Summary

Introduction

- Data visualisation is the transformation of data into a **visual form** so that **meaningful information** embedded within it can be made **more accessible**.

- Data visualisation is important to any disciplines or individual who **deals with data** and the need to **extract** or **share insights** from these data.

- Effective data visualisation serves two main purposes. It allows us to **explore** the dataset to gain useful insights and it allows us to **communicate** effectively what the data is telling us.

Explore ← → Explain

Data Visualisation

13

13

References for Introduction

[1]   Neural correlates of the episodic encoding of pictures & words - https://www.pnas.org/content/95/5/2703.full

[2]   Bui, D.C. & McDaniel, M.A. (2012) – Enhancing learning during lecture...   https://www.sciencedirect.com/science/article/abs/pii/S2211368115000200

[3]   Anscombe, F. J. (1973). "Graphs in Statistical Analysis". American Statistician. 27 (1): 17–21

[4]   Big Data Visualization. Value it brings & techniques in requires - https://www.scnsoft.com/blog/big-data-visualization-techniques

[5]   C.N. Knaflic, Storytelling with data, Wiley (2015)

[6]   Chart from: https://www.optimizesmart.com/what-type-of-chart-to-use-to-compare-data-in-excel/

[7]   BusinessInsider.com, This Chart Shows An Alarming Rise In Florida Gun Deaths After 'Stand Your Ground' Was Enacted - https://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2

Note: All online articles were accessed between Jul to Aug 2025

14

14