

**“The greatest value of a picture is when it forces us to notice what we never expected to see.”**

**John W. Tukey**  
American Statistician

John Wilder Tukey (June 16, 1915 – July 26, 2000) was an American mathematician best known for development of the FFT algorithm and box plot. Born in New Bedford, Massachusetts in 1915, and obtained a B.A. in 1936 and M.Sc. in 1937, in chemistry, from Brown University, before moving to Princeton University where he received a Ph.D. in mathematics.

**MIS771 – Descriptive Analytics and Visualisation**  
**Data Visualisation Module – Week 2**

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

1

**Paul Hodge**  
www.DataVizGuy.com

**Telecommunications** **Consumer Goods** **Life-sciences** **Retail**

1980 1990 2000 2010

**Location (Years)**

10 New York  
1 Lausanne  
1 New Delhi  
3 Hong Kong  
3 Tokyo  
12 Melbourne

Successfully Climbed Mount Fuji at Sunrise with friends.

Ran the New York Marathon two times, fastest in 3hrs 56min

**Focus Area**

- Strategy & Roadmaps
- Business Intelligence
- Project Delivery
- Analytics
- Data Warehouse

Lecturer, Data Visualisation

DEAKIN UNIVERSITY

DATA VIZ GUY  
www.DataVizGuy.com

How as my experience has brought me here? (Identifying themes that transcend domains, that connect with people, and are

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

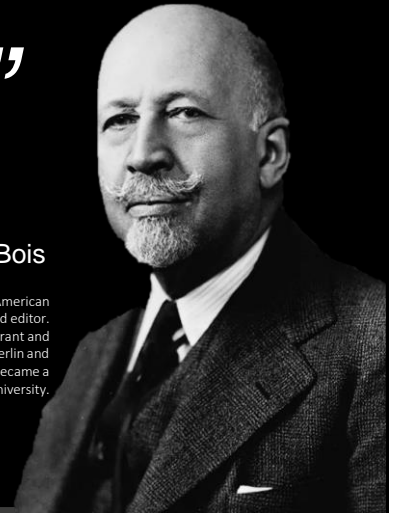
2

“Education must not simply teach work – it must teach Life.”



W. E. B. Du Bois

William Edward Burghardt Du Bois (February 23, 1868 – August 27, 1963) was an American sociologist, socialist, historian, civil rights activist, Pan-Africanist, author, writer and editor. Born in Great Barrington, Massachusetts, Du Bois grew up in a relatively tolerant and integrated community, and after completing graduate work at the University of Berlin and Harvard, where he was the first African American to earn a doctorate, he became a professor of history, sociology and economics at Atlanta University.



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

3

**gates**  
notes

According to the World Bank,

gates  
notes

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Gates Notes (third wave of poverty reduction):  
<https://youtu.be/efnHiz0USQE>

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

4

# Brief Session 1 Recap .....

1. Introduction

2. Framework

3. Engagement (Story | Navigation | Motivation)

Homework....

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL


Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

5

Engaging

Audience Centric

"The two most engaging powers of an author are to make new things familiar, familiar things new"




William Makepeace Thackeray

Informative

Data Centric

"Data is a precious thing and will last longer than the systems themselves"




Tim Berners-Lee

Designed

Experience Centric

"It's not just what it looks like and feels like. Design is how it works."



Steve Jobs

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

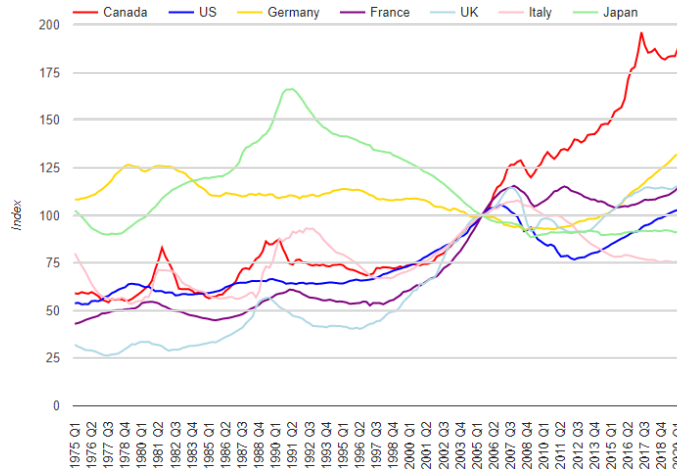
6

8

8

## G7 Real Estate Price Index

An inflation adjusted index of G7 real estate prices.



Source: US Federal Reserve Bank of Dallas, Better Dwelling.

Source: <https://betterdwelling.com/canada-doubled-down-on-real-estate-in-2005-now-its-the-biggest-bubble-the-g7-has-ever-seen-and-its-getting-bigger/>



Scott Berinato

Author, Good Charts and Good Charts Workbook

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

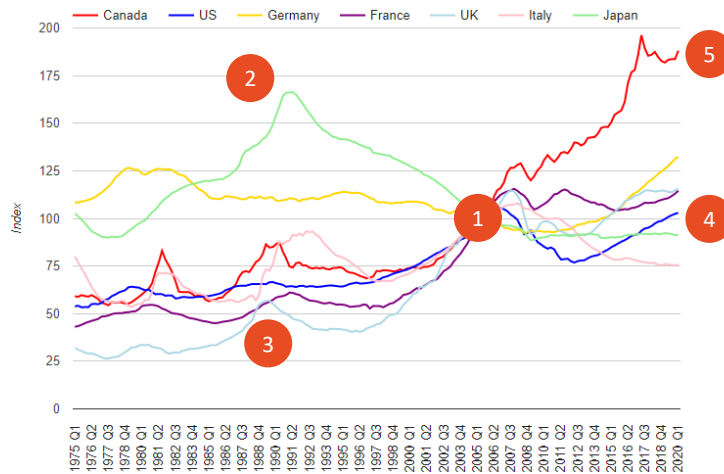
Find the Story in the Data

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

9

## G7 Real Estate Price Index

An inflation adjusted index of G7 real estate prices.



Source: US Federal Reserve Bank of Dallas, Better Dwelling.

Source: <https://betterdwelling.com/canada-doubled-down-on-real-estate-in-2005-now-its-the-biggest-bubble-the-g7-has-ever-seen-and-its-getting-bigger/>

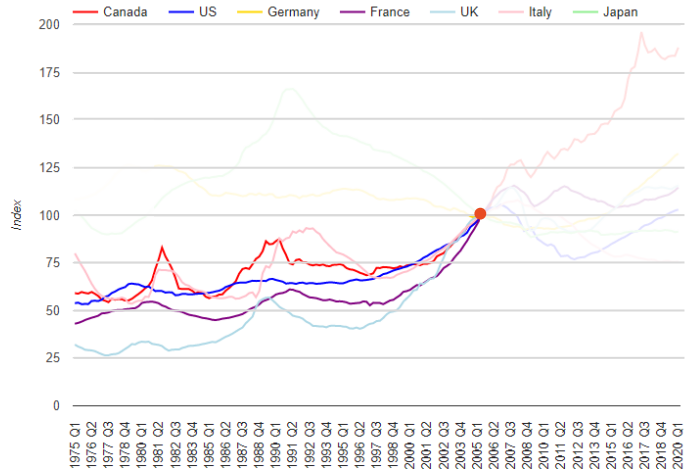
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

10

Home prices saw steady growth for most Countries... compared to 2005

Home prices saw few spikes or dips for 30 years leading up to 2005



Source: US Federal Reserve Bank of Dallas, Better Dwelling.

Source: <https://betterdwelling.com/canada-doubled-down-on-real-estate-in-2005-now-its-the-biggest-bubble-the-g7-has-ever-seen-and-its-getting-bigger/>

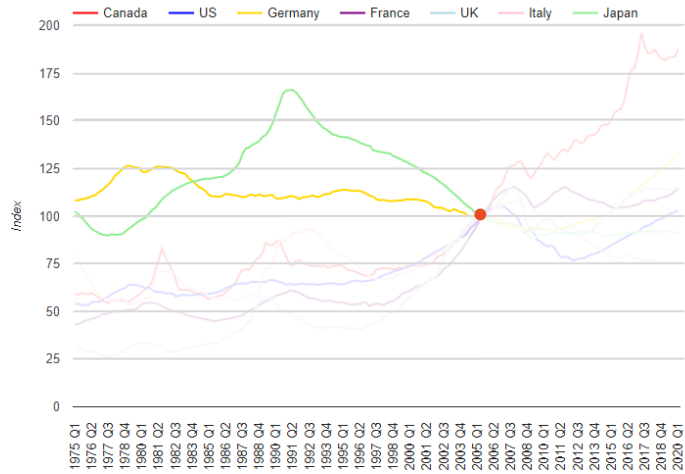
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

11

Mostly Steady growth... except for a spike in Japan

Japan experiences a 30 year bubble that peaked in the early '90s, with Germany homes basically holding their value



Source: US Federal Reserve Bank of Dallas, Better Dwelling.

Source: <https://betterdwelling.com/canada-doubled-down-on-real-estate-in-2005-now-its-the-biggest-bubble-the-g7-has-ever-seen-and-its-getting-bigger/>

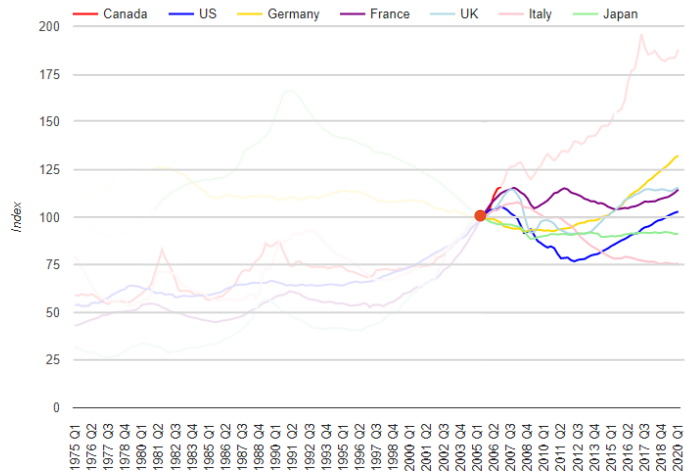
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

12

A slow recovery emerging... after 2005

Half the countries are barely maintaining 2005 value (following the GFC), with Germany doing the best and the US building momentum



Source: US Federal Reserve Bank of Dallas, Better Dwelling.

Source: <https://betterdwelling.com/canada-doubled-down-on-real-estate-in-2005-now-its-the-biggest-bubble-the-g7-has-ever-seen-and-its-getting-bigger/>

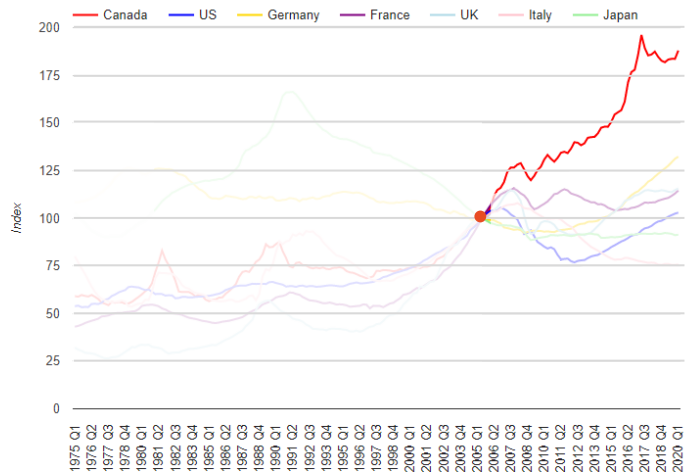
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

13

Is there a Potential New Bubble emerging... especially Canada

While most countries are barely maintaining 2005 value or making modest gains, Canada is racing away



Source: US Federal Reserve Bank of Dallas, Better Dwelling.

Source: <https://betterdwelling.com/canada-doubled-down-on-real-estate-in-2005-now-its-the-biggest-bubble-the-g7-has-ever-seen-and-its-getting-bigger/>

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

14



What goes into a great wine: Ingredients for success: Knowledge, Practice, a splash of natural Aptitude (talent) – We will come back to this.  
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL


Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

15

Engaging

Audience Centric

"The two most engaging powers of an author are to make new things familiar, familiar things new"




William Makepeace Thackeray

Informative

Data Centric

"Data is a precious thing and will last longer than the systems themselves"




Tim Berners-Lee

Designed

Experience Centric

"It's not just what it looks like and feels like. Design is how it works."



Steve Jobs

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL


Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

16



MIS771 – Data Visualisation Module

Descriptive Analytics and Visualisation



Week 1

120mins

60mins

6pm

6:40pm

7pm

8:00pm

8:20pm

9pm

Introduction & History

Engagement Elements

Hands-on Tutorial

Week 2

120mins

60mins

6pm

6:40pm

7pm

8:00pm

8:20pm

9pm

Information Elements

Ethics

Hands-on Tutorial

Week 3

120 mins

60mins

6pm

6:40pm

7pm

8:00pm

8:20pm

9pm

Design Elements

Future

Hands-on Tutorial

Assignment Deadline : TBA

Note: We will discuss more at the end of the Session

Empowerment Promise (together), Principles focused (Tool Agnostic) – We will explore some examples, establish a language, and discuss some practical tips & tricks

Copyright © 2021 All rights reserved ( @DataVizGuy) CONFIDENTIAL

Principles focused – Tool Agnostic


Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

17

Informative

"Data is a precious thing and will last longer than the systems themselves"

Tim Berners-Lee




"Data Centric"

Data

Augment

Charts

(algorithm)



Copyright © 2021 All rights reserved ( @DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

18

Copyright 2021 @DataVizGuy

9

“ Having all the information in the world at our fingertips doesn't make it easier to communicate: it makes it harder ”

Cole Nussbaumer Knaflic  
Author, Speaker, CEO

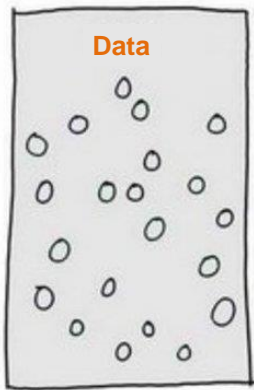
Cole Nussbaumer Knaflic tells stories with data. She is Founder & CEO of storytelling with data, the author of "storytelling with data: a data visualization guide for business professionals" (Wiley, November 2015)



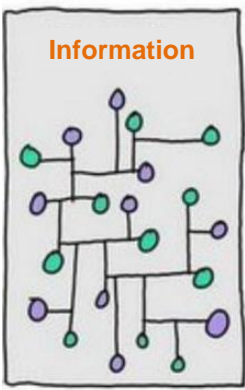
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

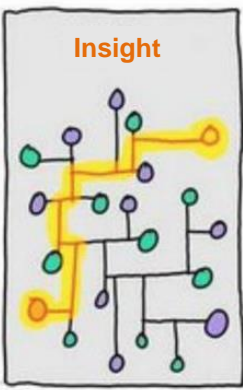
19



**Data** is the raw numbers that we capture according to some agreed to standards.



**Information** is a collection of data points that we can use to understand something about the thing being measured.



**Insight** is gained by analysing data and information to understand what is going on with the particular situation

- Observations
- Objects
- Situations
- Scenarios

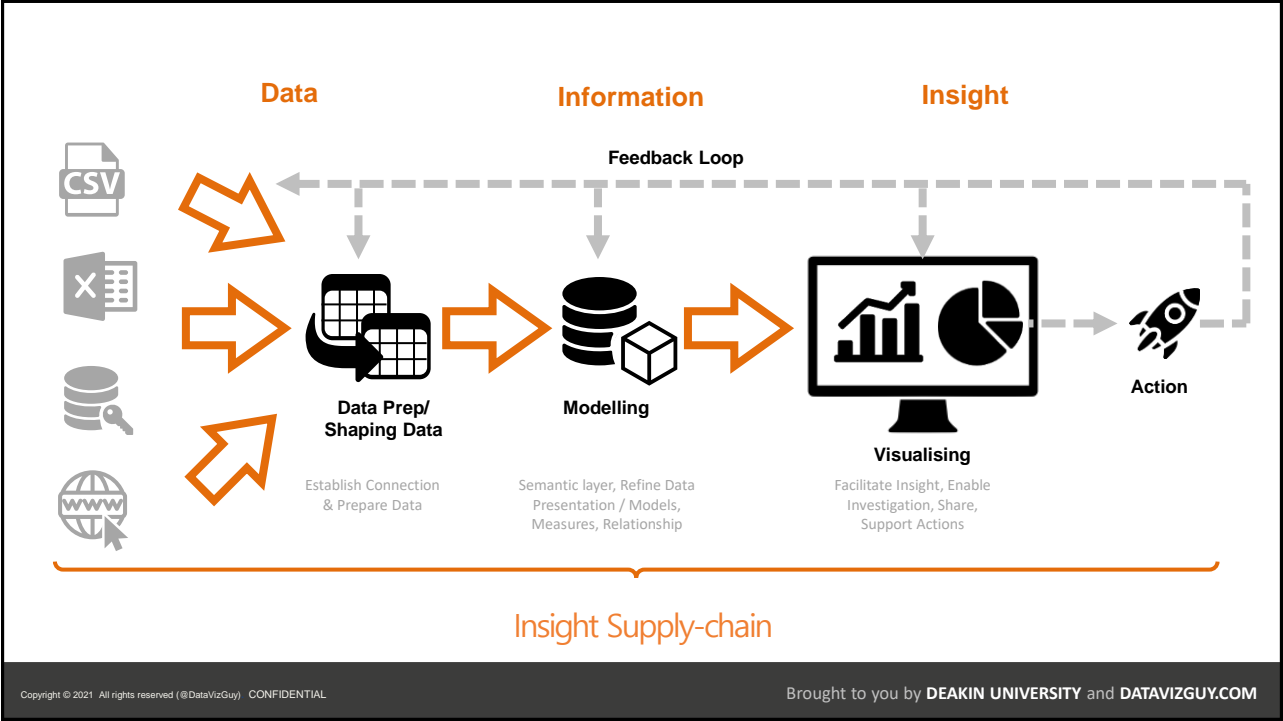


Carly Fiorina, ex CEO of Hewlett-Packard  
“The goal is to turn data into information, and information into insight.”

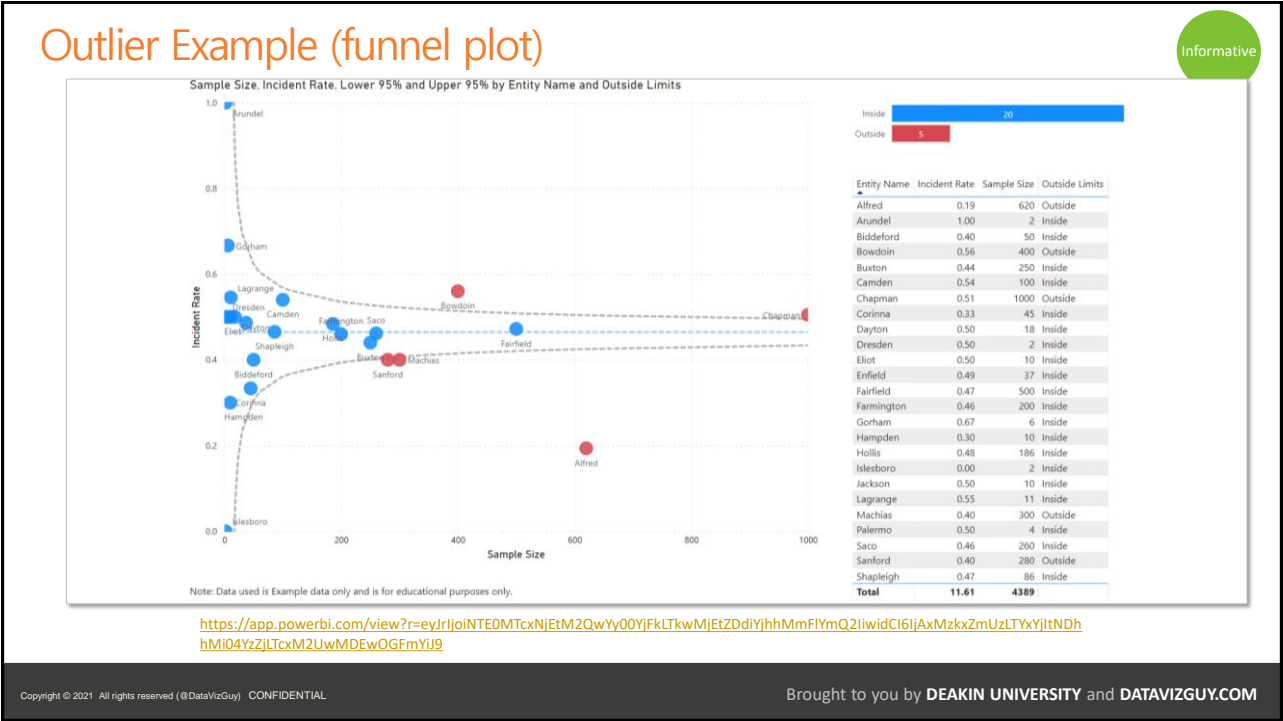
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

20

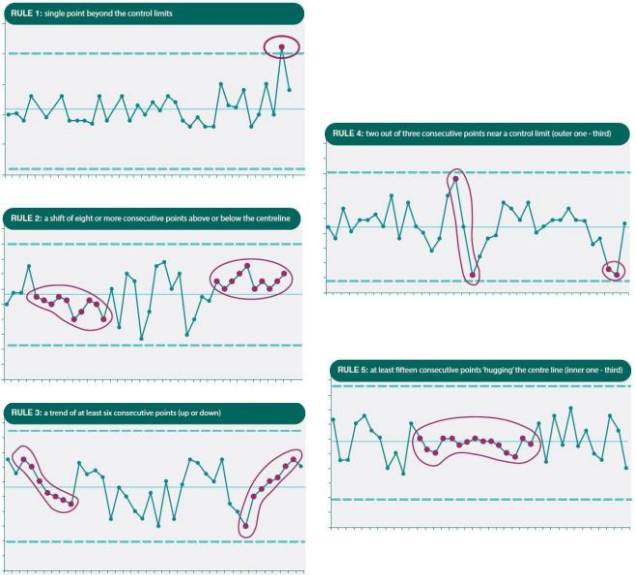


21

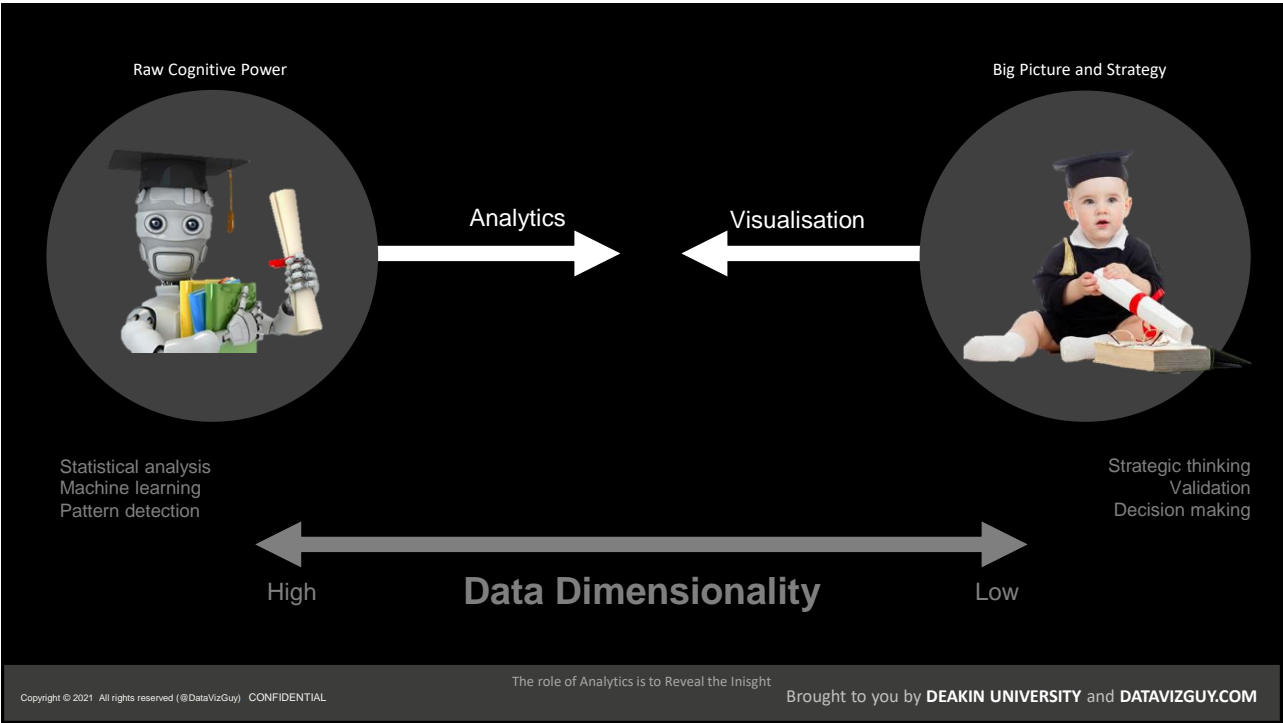


22

# Control Charts / Identifying Patterns



Source: <https://www.cec.health.nsw.gov.au/Quality-Improvement-Academy/quality-improvement-tools/control-charts>



“ The art and practice of visualizing data is becoming ever more important in bridging the human-computer gap to mediate analytical insight in a meaningful way. ”

Edd Dumbill

Edd Dumbill is a technologist, writer and programmer based in California. He is the program chair for the O'Reilly Strata and Open Source Convention Conferences. A veteran of open source, Edd has contributed to various projects, such as Debian and GNOME, and created the DOAP Vocabulary for describing software projects.



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

<https://archive.org/stream/graphicpresentat00brinrich#page/2/mode/2up> Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

25



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Data & Container

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

26

## Data Structure (Basic Components)



**Categorical** (Dimensions, Attributes, Features)

**[Σ] Measure** (Facts, Values, Observations)

**Header Row**  
(Column Headers, Field Names)

**Data Rows**  
(Records, Transactions, Measurements, Events)

<b>Cust ID</b>	<b>Name</b>	<b>Birth Date</b>	<b>Residential Location</b>	<b>Occupation</b>	<b>Years as Customer</b>	<b>Annual Wages</b>	<b>Credit Offered</b>
001	Peter	12-Nov-1962	Australia	Analyst	12.50	\$85,000	Yes
023	Marion	8-Jan-1991	USA	Programmer	2.50	\$45,000	No

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

27

## Data Quality vs. Fit-for-purpose

Informative

- Missing values (empty values or not complete)
- Redundancy of attributes/dimensions
- Consistency (consistent values, type consistency, units of value)
- Duplicates of records (not unique - uniqueness)
- Outliers (data that will distort analysis)
- Noise (simply wrong or invalid data)
- Timing or out of date (old data, not recent enough)

*"Unfortunately you are all likely to be disappointed when you first look at you data, because it won't be as nice and clean as you were hoping"... just remember that "bad" data can also sometime contain valuable insights*



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

28

## Data Elements

### Primary

- **Measurement Data** (eg. 3.14, 42.6)
- **Categorical Data** provides context (eg. Male, Female)
- **Continuous** (eg. 3.1415)
- **Ordinal Data** (eg. 1, 2, 3, 2-Nov-13)
- **Interval Data** (eg. Age 20-25)
- **Time/Date Data** (eg. 15-Nov-2005)
- **Location/Spatial Data** (eg. Melbourne, 3001)
- **Relational** (eg. Parent, Sibling, Level1)

---

### Canvas

- **Background** (eg. Axis, Map, People, Teeth)
- **Legend**
- **Title, Headings & Labels**
- **Citations/References**
- **Notes** (eg. Data inclusions/exclusions)

## Visual Elements

size

color

position

angle

shape

texture

movement

connection

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

29

# Perception Pop Quiz.....

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

30

Pop Quiz 1.

Informative

**A**



**B**



- 1. Is A the *SAME* Length as B?
- 2. Is B *TWICE* the Length as A?
- 3. Is B *THREE* times the Length as A?

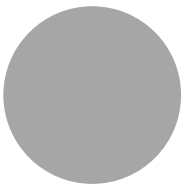
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

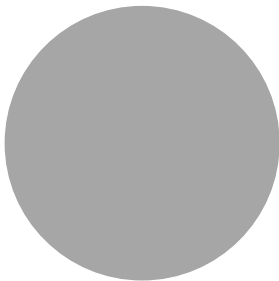
31

Pop Quiz 2.

Informative



**A**



**B**

- 1. Is A the *SAME* Size as B?
- 2. Is B *TWICE* the Size as A?
- 3. Is B *THREE* times the Size as A?

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

32



# Semiology of Graphics

Informative

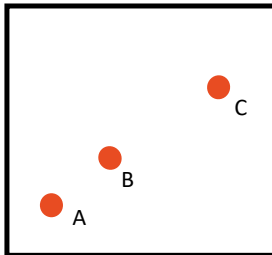


## Jacques Bertin

Jacques Bertin (27 July 1918 – 3 May 2010) was a French cartographer and theorist, known from his book *Semiotique Graphique* (*Semiology of Graphics*), published in 1967. This monumental work, based on his experience as a cartographer and geographer, represents the first and widest intent to provide a theoretical foundation to Information Visualization.

The function of a visual is answering the three following questions:

- What are the components of the data table? What it's all about?
- What are the groups in the data? What is the general level of information?
- What are the exceptions?



1. A, B, C are distinguishable (**N**ominal)
2. B is between A and C (**O**rdinal)
3. BC is twice as long as AB (**Q**uantitative)



Position	N	O	Q
Size	N	O	Q
Value	N	O	q
Texture	N	o	
Color	N	o	
Orientation	N		
Shape	N		

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

33

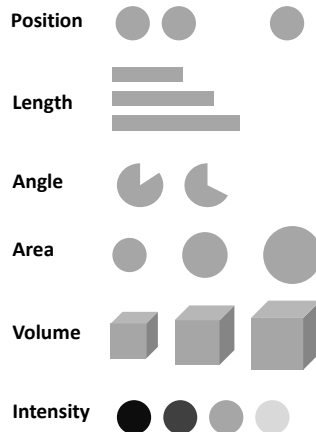
# Graphical Perception

Informative

More Accurate



Less Accurate



Slope



Hue



## William S. Cleveland & Robert McGill

William S. Cleveland and Robert McGill's paper "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods" from the September 1984 edition of the *Journal of the American Statistical Association*.

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Big Secret: Data Visualisation are not always about Accuracy

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

34

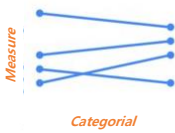
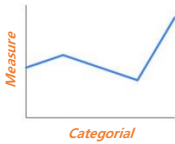
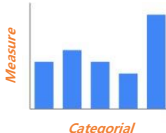
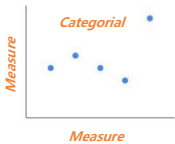
Informative

# Mapping of Categories and Measures

91%

Measure / Categorical

Categorical	Categorical		
	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	25%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%



Cole Nussbaumer

Cole Nussbaumer Knaflic is the author of Storytelling with Data: A Data Visualization Guide for Business Professionals and writes the popular blog storytellingwithdata.com. She's worked at and with some of the most data-driven companies on the planet, including Google, Adobe, Genentech, JPMC, Target, and the World Bank.

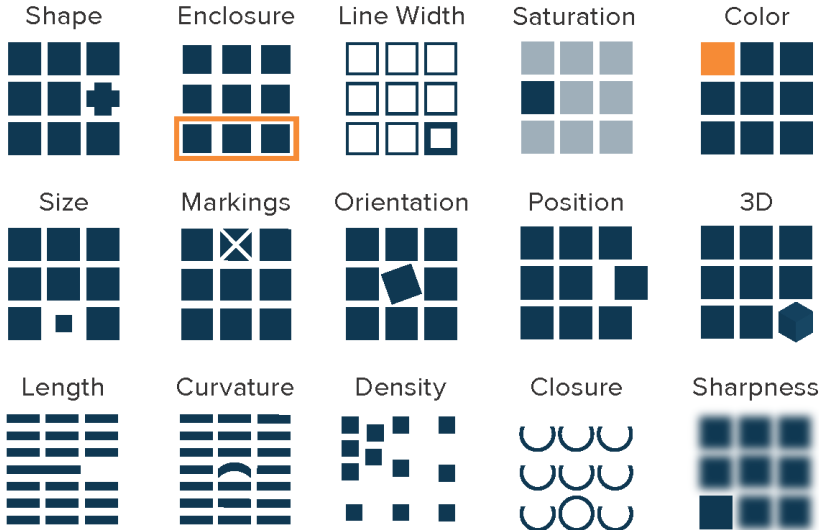
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

35

Informative

# Plus Pre-Attentive Attributes



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

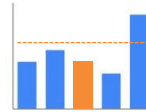
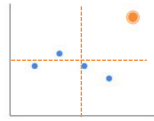
Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

36

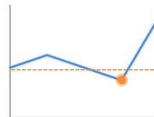
## Mapping Pre-attentive Attributes

Informative

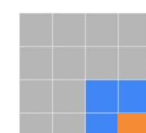
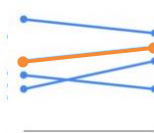
91%



	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%



	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%



Book: *Storytelling with Data*



**Cole Nussbaumer**

Cole Nussbaumer Knaflic is the author of *Storytelling with Data: A Data Visualization Guide for Business Professionals* and writes the popular blog [storytellingwithdata.com](http://storytellingwithdata.com). She's worked at and with some of the most data-driven companies on the planet, including Google, Adobe, Genentech, JPMC, Target, and the World Bank.

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

37

## Tables as Visualisations

#SIDEBAR

*Categorical*

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

*Measure*

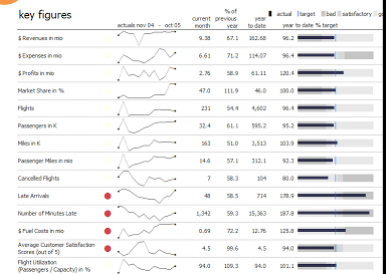
1 Focus on the Data

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

2 Highlight what matters

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	35%	17%	34%
Category 4	30%	29%	26%
Category 5	55%	30%	58%
Category 6	11%	25%	49%

3 Use Sparklines (small-multiples)



Improve data-ink ratio

Highlight the insights

Reveal the patterns

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Accessibility note: use contrast and reduce complexity

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

38

# Insight Game.....

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

39

## Insight Game











Informative

W List of cities in ASEAN by popul... x +

en.wikipedia.org/wiki/List\_of\_cities\_in\_ASEAN\_by\_population

### Largest cities proper [\[edit\]](#)

This list features the most populous cities in ASEAN. Population figures were taken from within the city proper only. See the article on each city for sources. Myanmar data is the least reliable and subject to revision. It is important to note, that the general area of Metro Manila has been divided into its political cities and is not counted as one urban city in this ranking. Therefore the metropolis' true population is not reflected in this ranking (it would otherwise be first or second in population size, depending on the definition of Jakarta's true city area).

Ranking	Name	Image	Country	Population	Year (Census or Estimate)
1	Jakarta		 Indonesia	10,135,030	2014E
2	Bangkok		 Thailand	8,750,600	2017C
3	Ho Chi Minh City		 Vietnam	8,993,082	2019C
4	Hanoi		 Vietnam	8,053,663	2019C
5	Singapore		 Singapore	5,399,000	2014E

Data Source: [https://en.wikipedia.org/wiki/List\\_of\\_cities\\_in\\_ASEAN\\_by\\_population](https://en.wikipedia.org/wiki/List_of_cities_in_ASEAN_by_population)
<https://app.powerbi.com/view?r=eyJrIjoiaMTRlbnJmM3OGYtMGU0ZS00NmRlLWEyMTMtNWQ1N2U0MGU0NzhkiiwidCI6IjAxMzZxZmUzLTlYyYjItNDhhMi04YzZlTcxM2UwMDEwOGFmYjI9>

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Hint: First seek to Understand the Motivation, then the Supporting Questions?

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

40

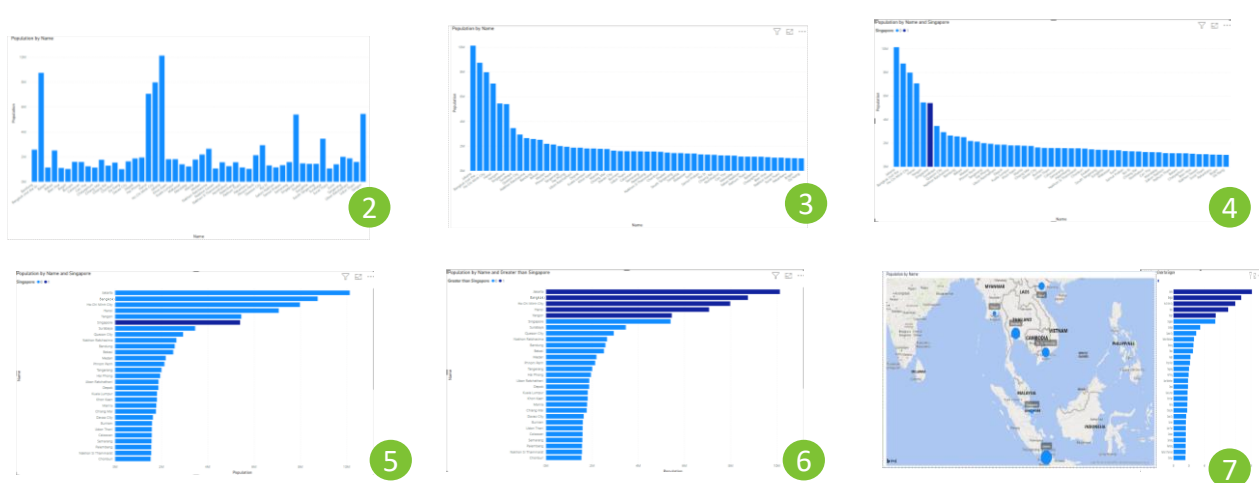
# Insight Game

1

Data Source: [https://en.wikipedia.org/wiki/List\\_of\\_cities\\_in\\_ASEAN\\_by\\_population](https://en.wikipedia.org/wiki/List_of_cities_in_ASEAN_by_population)

Hint: First seek to Understand the Motivation, then the Supporting Questions?

Informative

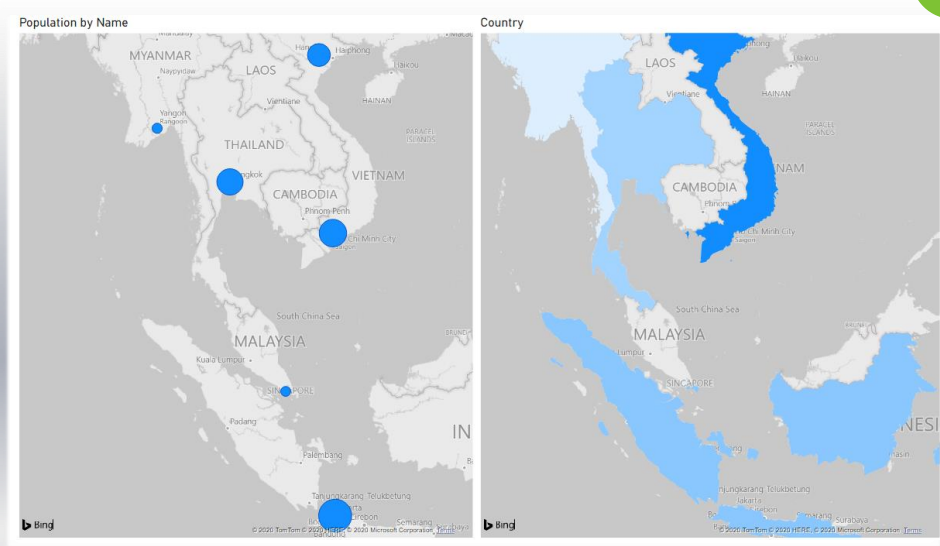


Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

41

# #SIDEBAR



Data Source: [https://en.wikipedia.org/wiki/List\\_of\\_cities\\_in\\_ASEAN\\_by\\_population](https://en.wikipedia.org/wiki/List_of_cities_in_ASEAN_by_population)

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Sidebar: Visual Perception, as applied to maps

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

42

## Algorithms - Simple Statistics

Informative

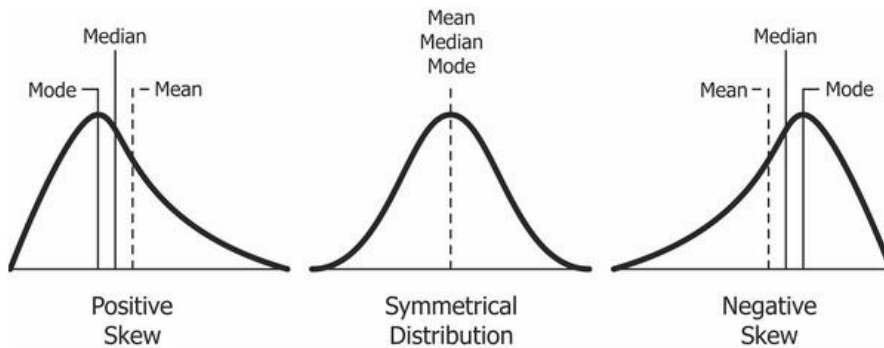
13, 18, 13, 14, 13, 16, 14, 21, 13  
sorted list: 13, 13, 13, 13, 14, 14, 16, 18, 21

Mean/Average: 15 (sum of values / count)

Median: 14 (middle number)

Mode: 13 (most common)

Range: 8 (difference between highest and lowest)



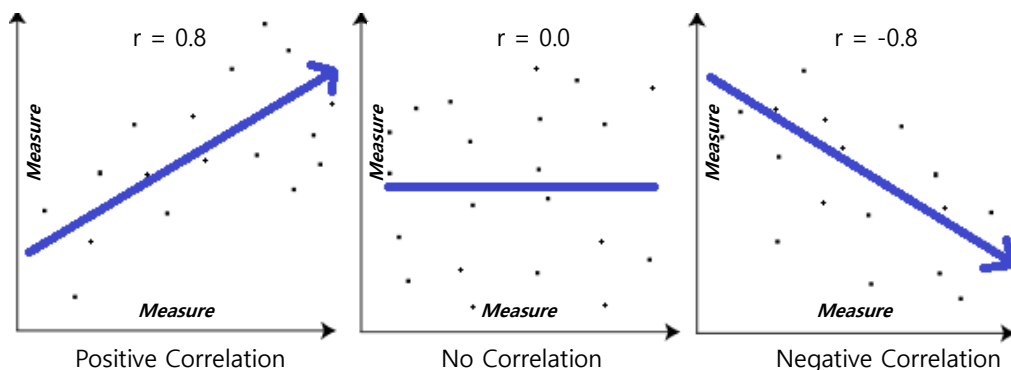
Side Bar: Detecting Outliers using 1. z-score calculations or 2. Copyright © 2021. All rights reserved (@DataVizGuy). CONFIDENTIAL exceeding a distance of 1.5 times the interquartile range (IQR), etc.

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

43

## Algorithms - Correlations

Informative



$$r = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} = \frac{Cov(X, Y)}{S_x S_y}$$

Values between 0.7 and 1.0 (-0.7 and -1.0) indicate a strong positive (negative) linear relationship via a firm linear rule.

The Pearson correlation coefficient is a measure of the strength and direction of the linear relationship between two variables/measures.

Copyright © 2021. All rights reserved (@DataVizGuy). CONFIDENTIAL

Sidebar: Intro Causal vs Correlation, and Confounding Variables (ask for examples)

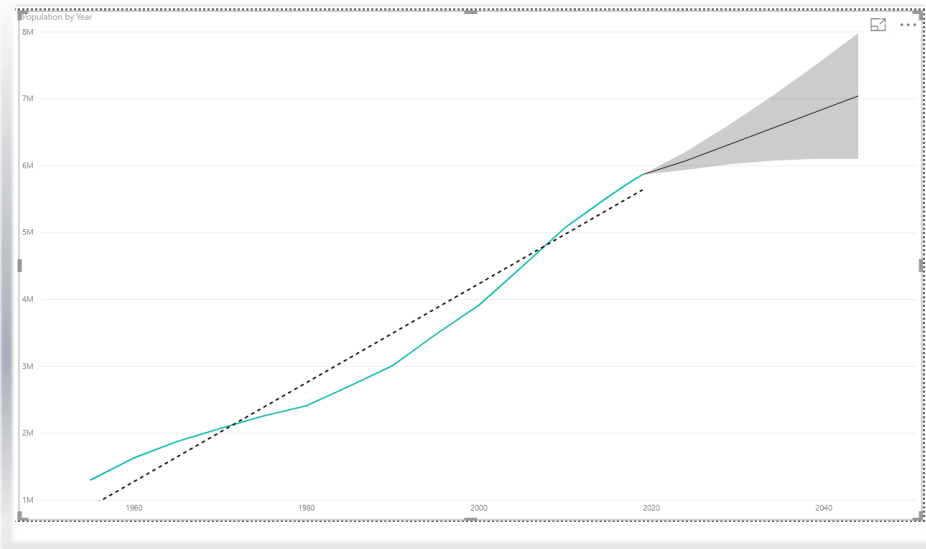
Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

44

# Algorithms – Basic Tool Features

Informative

Understand what the Algorithm is doing



Data Source: <http://www.worldometers.info/world-population/singapore-population/>

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

45

# Chart Selection – Where to Start?

Design



Arc Diagram



Area Graph



Bar Chart



Box & Whisker Plot



Brainstorm



Bubble Chart



Bubble Map



Bullet Graph



Calendar



Candlestick Chart



Chord Diagram



Choropleth Map



Circle Packing



Connection Map



Density Plot



Donut Chart



Dot Map



Dot Matrix Chart



**Severino Ribecca**  
Severino Ribecca is graphic designer who specialises in developing visual systems through the use of information design, a branch of graphic design that deals with data visualisation, diagram design and communicating information visually in a clear & effective manner.

[www.DatavizCatalogue.com](http://www.DatavizCatalogue.com)

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

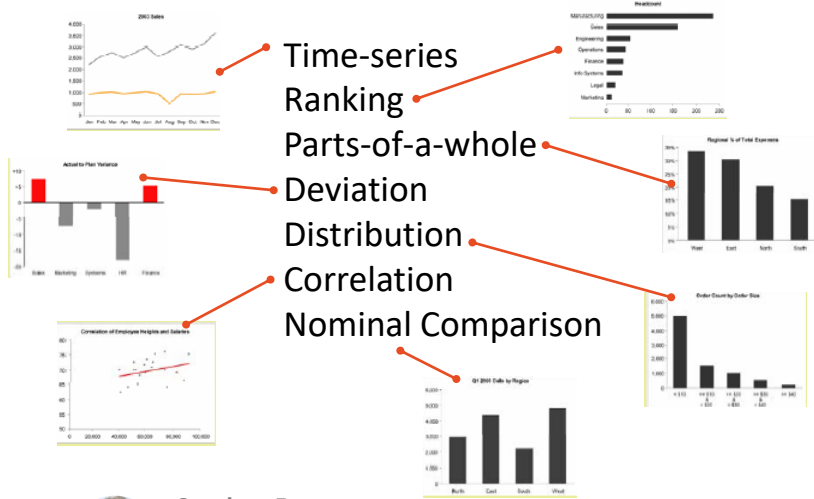
Making sense of the Variety

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

46

## Chart Selection - Seven Basic Chart relationships

Informative



**Stephen Few**

Stephen Few is an innovator, consultant, and educator in the fields of business intelligence (a.k.a. data warehousing and decision support) and information design. Through his company, Perceptual Edge, he focuses on the effective analysis and presentation quantitative business information. Book: Show me the numbers, 2004

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Basic Chart Selection

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

47

## Chart Selectors

Informative



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

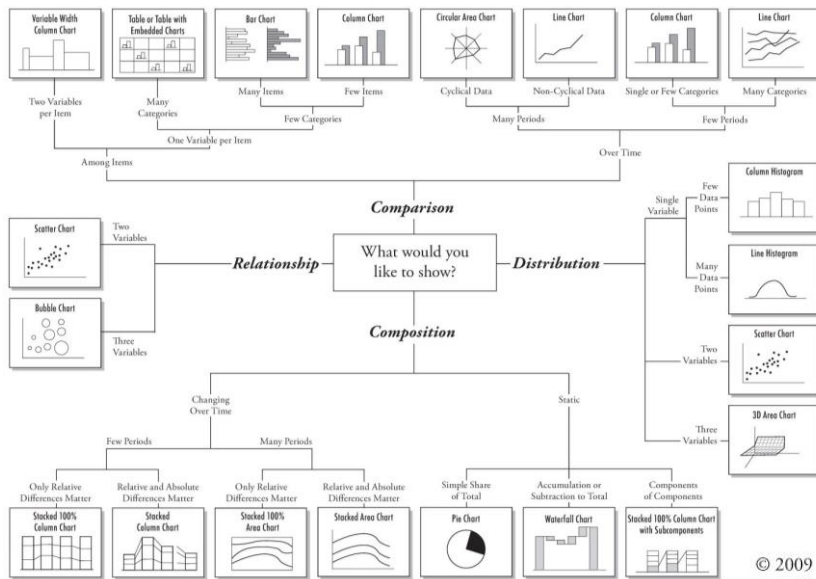
Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

48



## Chart Selection – “Chart Suggestions” (a thought starter)

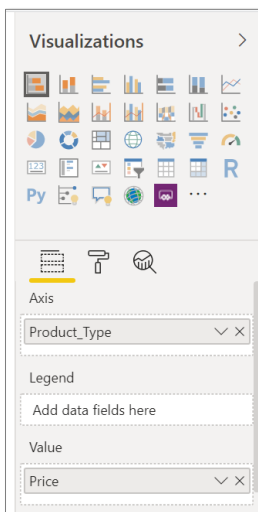
Informative



Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

49

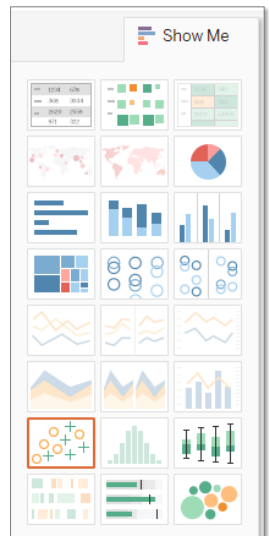
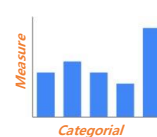
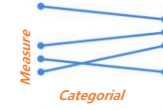
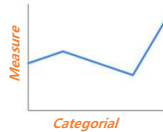
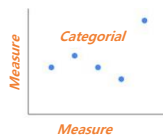
## Chart Selection - Visual Suggestions



**91%**  
Measure / Categorical

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	10%	10%	7%
Category 4	55%	30%	58%
Category 5	11%	25%	49%

	A	B	C
Category 1	15%	22%	42%
Category 2	40%	36%	20%
Category 3	10%	10%	7%
Category 4	55%	30%	58%
Category 5	11%	25%	49%



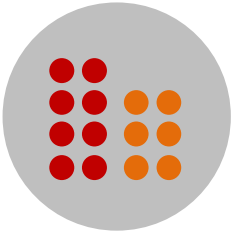
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Basic Chart Selection, keep it simple

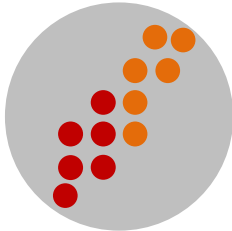
Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

50

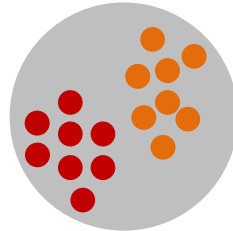
# Insight Patterns to support Key Decisions



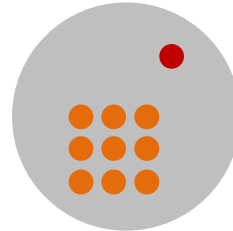
Comparison



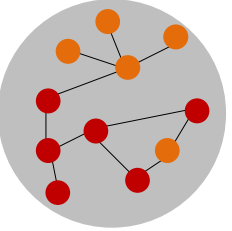
Trend



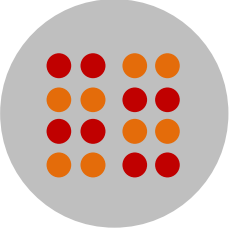
Cluster / Proximity



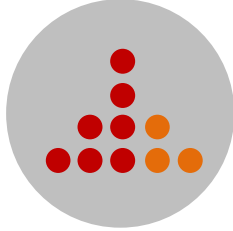
Outlier



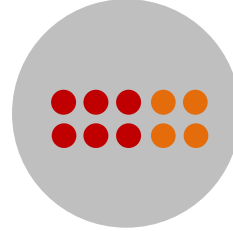
Relationship



Correlation



Distribution



Progress

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL













Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

51

# Explore vs. Explain







































































































































































































































































































































































































## Use these charts to Explore

Complex charts can be fascinating and look impressive, but they are often obscure the main point.



## Use these charts to Explain

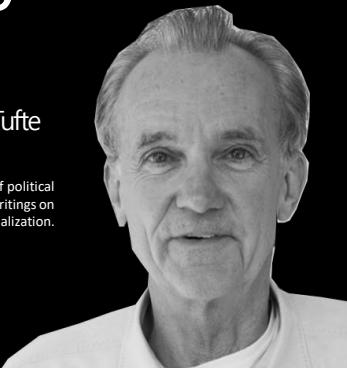
Everyone easily processes and understand the bar chart, pie, line chart, and perhaps the scatter plot

“  
There are two goals when presenting  
data: convey your story and  
establish credibility ”

Edward Tufte

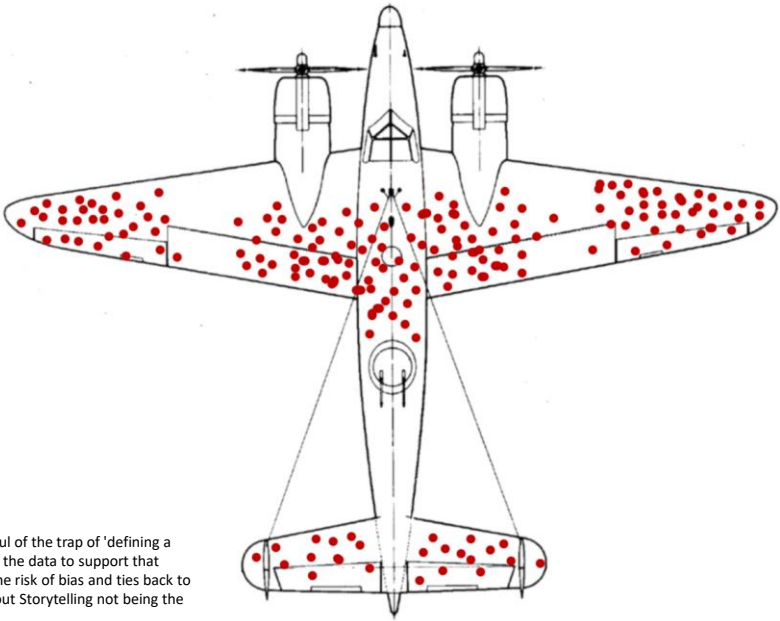
Edward Rolf Tufte (born 1942) is an American statistician and professor emeritus of political science, statistics, and computer science at Yale University. He is noted for his writings on information design and as a pioneer in the field of data visualization.



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

53



Informative



Warning: Also be careful of the trap of 'defining a story first, then finding the data to support that story'. This increases the risk of bias and ties back to what we discussed about Storytelling not being the methodology.

Source: en.wikipedia.org  
(Image shows hypothetical data)

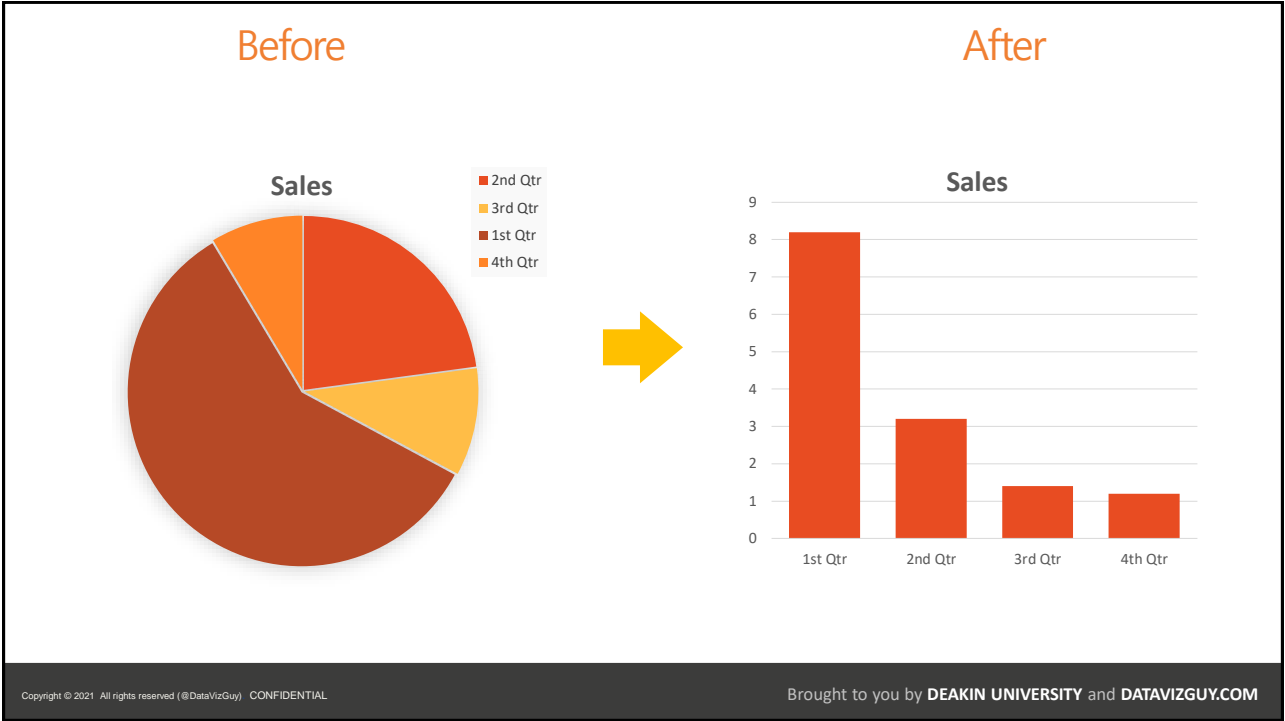
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

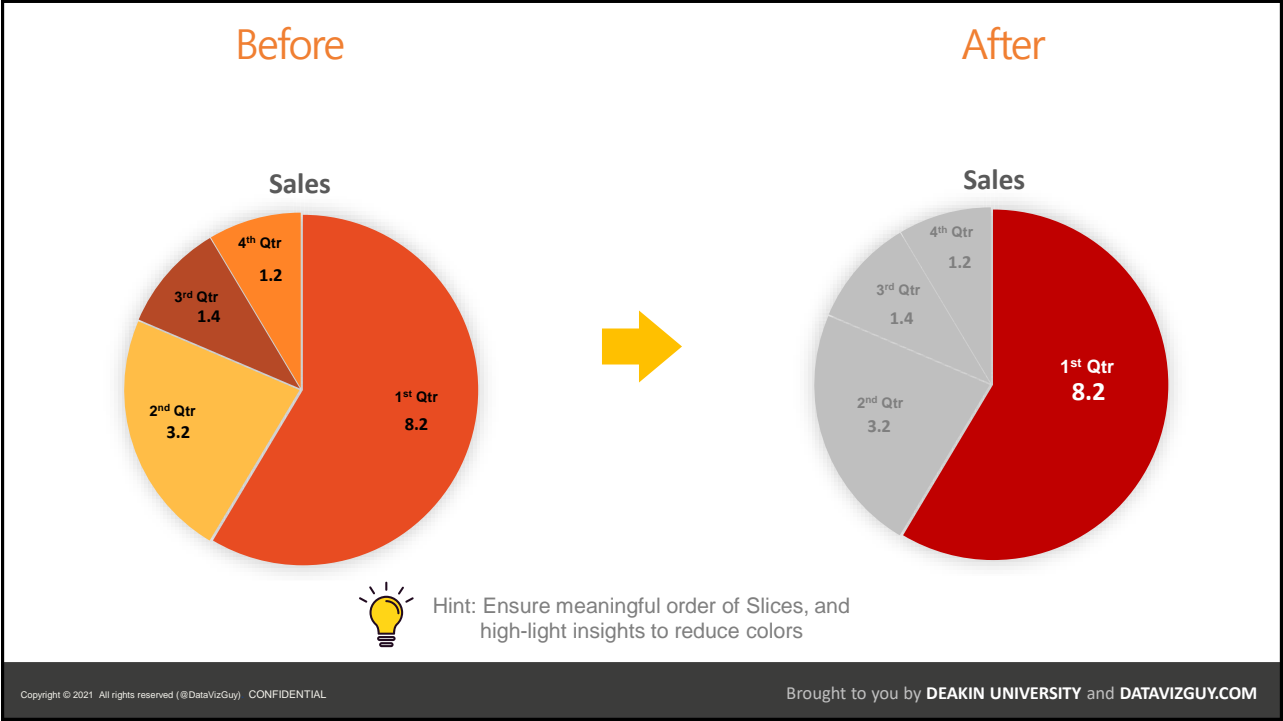
54



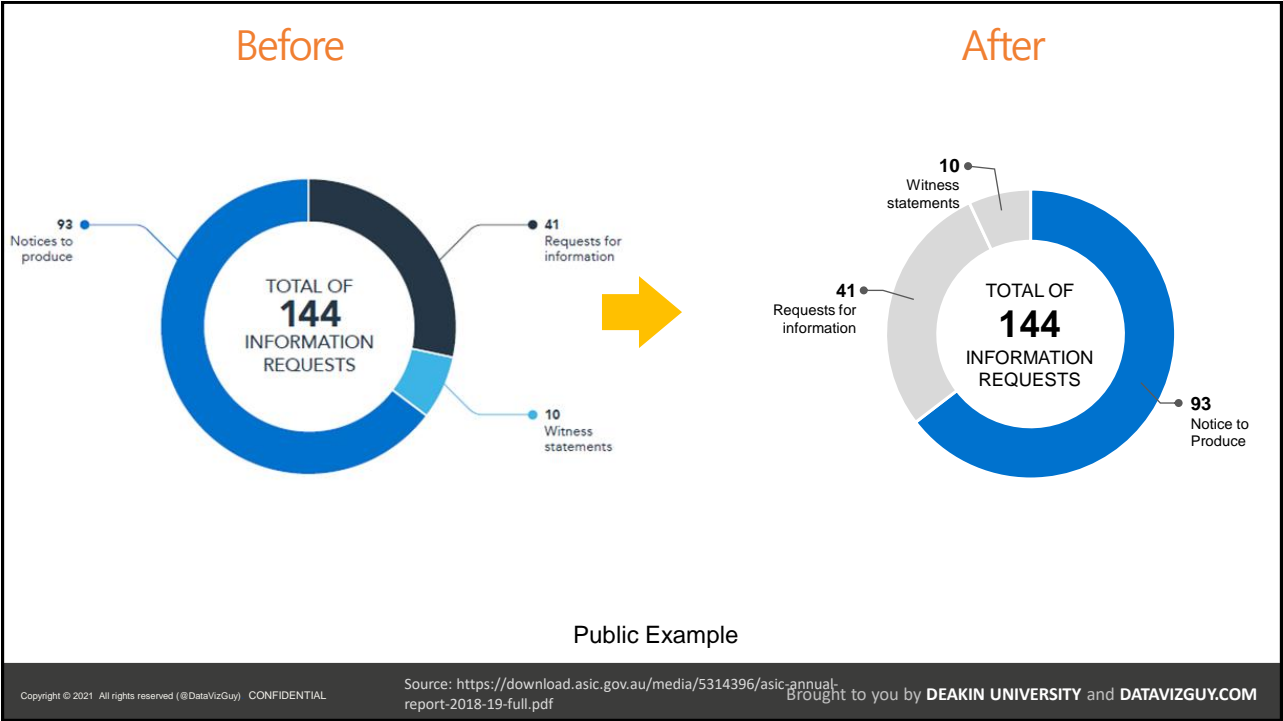
55



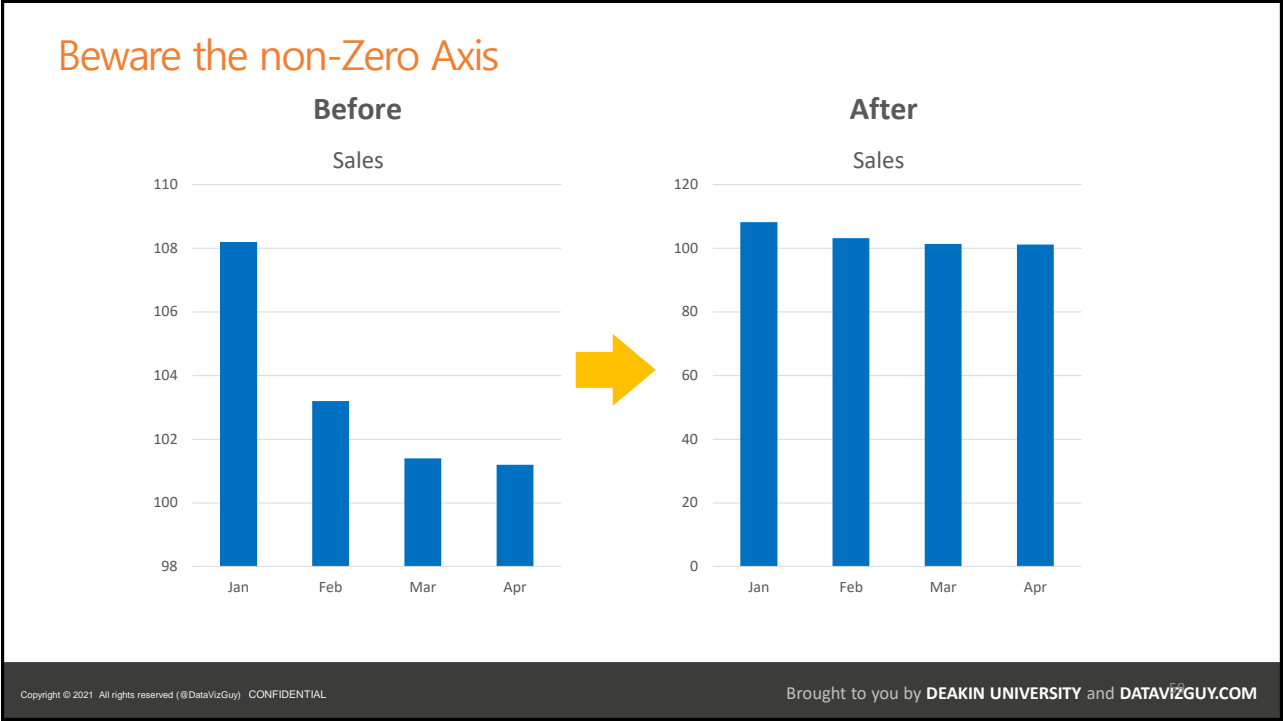
56



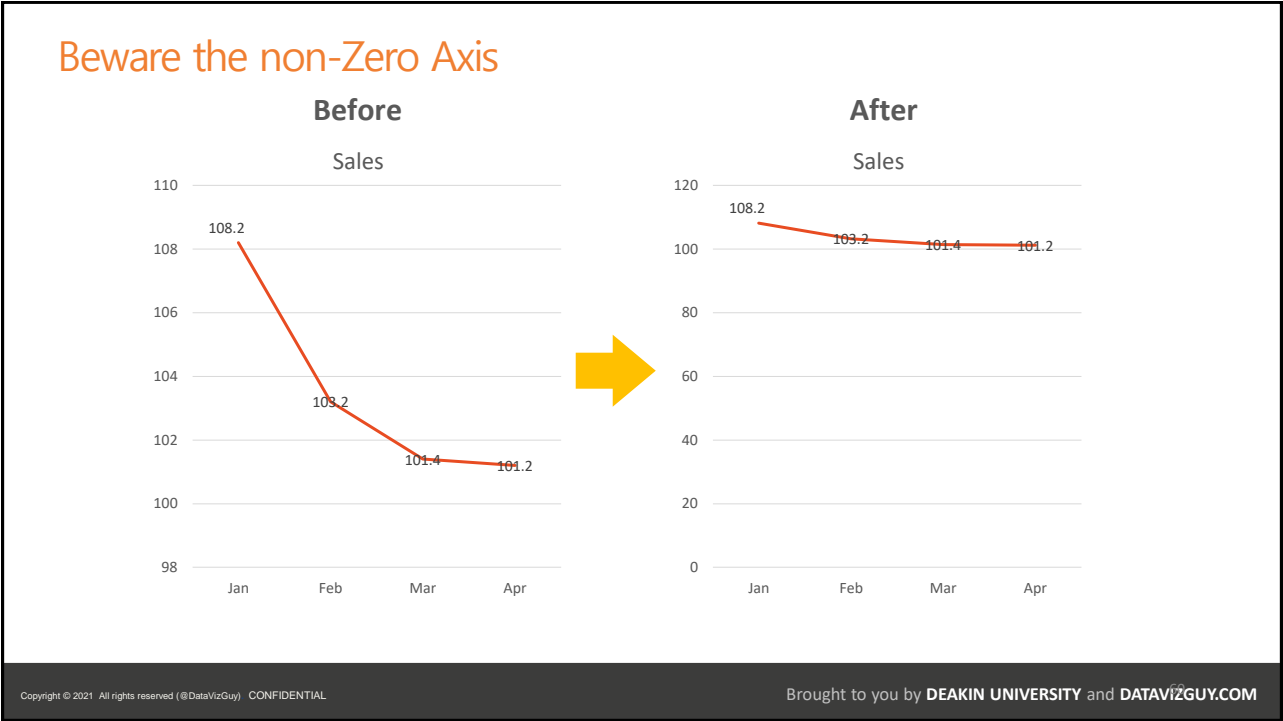
57



58



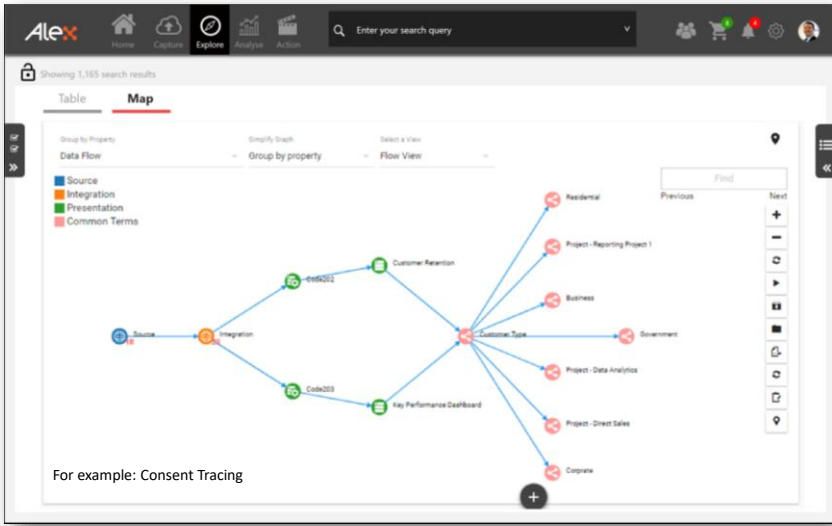
59



60

# Data Governance and Stewardship

Informative



**Alex Solutions**  
Proudly Australian  
Owned and Developed  
[www.AlexSolutions.com.au](http://www.AlexSolutions.com.au)

Gartner 2016 Cool Vendor Magic Quadrant FORRESTER

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

61

#SIDEBAR

### Simplified Checklist

Do's

✓

- ❑ Start by considering your audience
- ❑ Understand your data, including potential bias
- ❑ Try different charts, but start with the basics (see what works in your situation)
- ❑ Show all the data (show patterns)
- ❑ High-light the exceptions
- ❑ High-light the groups
- ❑ Use Axis starting at Zero
- ❑ Use sorting on Bar Charts (ranking)
- ❑ Use Labels on Data (place the legends as close to data as possible)

Don'ts

✗

- ❑ Don't forget the story (what are you trying to explain)
- ❑ Don't use charts with no purpose
- ❑ Don't use overly complex charts
- ❑ Overuse Pie Charts
- ❑ Use 3D (avoid)
- ❑ Use Stacked Bar Charts (avoid)
- ❑ Overuse of color (be careful of color blindness)
- ❑ Logarithmic Axis (avoid)
- ❑ Don't forget to check your data sources
- ❑ Don't always follow the rules

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Accessibility note: use contrast and reduce complexity

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

62

Copyright 2021 @DataVizGuy

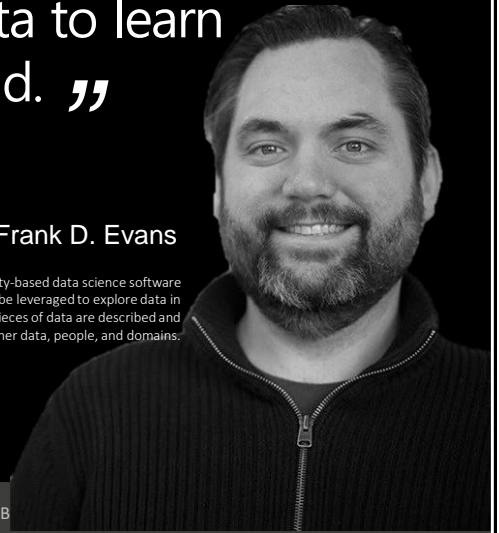
31



“ We’re not trying to learn about data,  
We’re trying to use data to learn  
about the world. ”

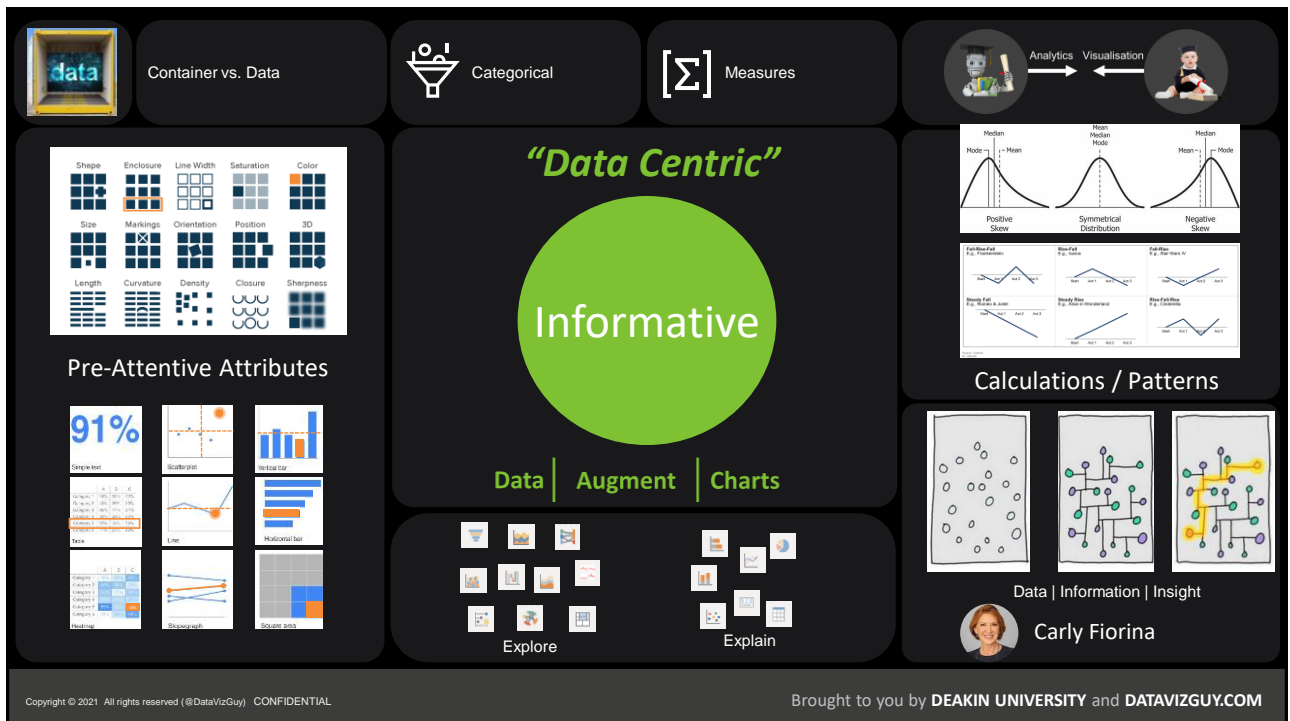
Frank D. Evans

Frank D. Evans is a Data Scientist with Exaptive, an Oklahoma City-based data science software company. Frank focuses on how data-driven applications can be leveraged to explore data in innovative ways. He specializes in data networks—where pieces of data are described and interacted with according to their relationships to other data, people, and domains.



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

63



64





What goes into a great wine: Ingredients for success: Knowledge, Practice, natural Aptitude (talent)  
Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL


Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

65

Engaging

Audience Centric

"The two most engaging powers of an author are to make new things familiar, familiar things new"




William Makepeace Thackeray

Informative

Data Centric

"Data is a precious thing and will last longer than the systems themselves"




Tim Berners-Lee

Designed

Experience Centric

"It's not just what it looks like and feels like. Design is how it works."




Steve Jobs


Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM




66



# Looking forward to continuing the conversation...



**Paul Hodge**  
@DataVizGuy


Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

67


## MIS771 – Data Visualisation Final Assignment

\* Note: Data will change from previous Trimester



### Submitted via Wordpress

- 1 Describe your project**  
**Review Assessment Data:**
  - **Select the Attributes you'd might like to build a story around**
  - **Add your own analysis and/or supporting data**
- 2 Answer a range of questions related to:**
  - **Information**
  - **Design**
  - **Engagement**
- 3 Capture Screenshots of an interactive Dashboard, plus embedded Link to Live Dashboard, with Explanations of work and Interactions (provide link to on-line Tableau dashboard)**



Hints: Review the Scoring Rubric for details on how the assignment will be scored. Students should also consider the professional appearance and presentation in order to be scored well (eg. Tables of contents, page numbers. Citations, etc). Finally keep in mind that the Dashboard design in the assignment should be designed for your Audience, not to simply support your analysis.

**Assignment Deadline: TBA**

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

68

Tutorial



*“Data visualization is the representation of information or concepts using visual elements in order to inform or entertain”*

Art with Intent

  
May-2021

Here is a new key for your class (for your use only):

Landing Page: <http://www.tableau.com/tft/activation>

Desktop Key: **TCNG-C734-B8C0-A4D2-951B**

**Instructions:** Click on the link above and select **Get Started**. On the form, enter your university email address for **“Business email”**; and under **“Organization”**, please input the name of your school.

Valid Through: 4-Jul-2021

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL 7:50p

Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

69

Tutorial



*“Data visualization is the representation of information or concepts using visual elements in order to inform or entertain”*

Art with Intent





IMPORTANT

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL


Brought to you by DEAKIN UNIVERSITY and DATAVIZGUY.COM

70



Copyright 2021 @DataVizGuy

35

# Looking forward to continuing the conversation...



**Paul Hodge**  
@DataVizGuy



Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

71

“ The greatest value of a picture is when it forces us to notice what we never expected to see. ”



**John W. Tukey**  
American Statistician

John Wilder Tukey (June 16, 1915 – July 26, 2000) was an American mathematician best known for development of the FFT algorithm and box plot. Born in New Bedford, Massachusetts in 1915, and obtained a B.A. in 1936 and M.Sc. in 1937, in chemistry, from Brown University, before moving to Princeton University where he received a Ph.D. in mathematics.

MIS771 – Descriptive Analytics and Visualisation  
Data Visualisation Module – Week 1

Copyright © 2021 All rights reserved (@DataVizGuy) CONFIDENTIAL

Brought to you by **DEAKIN UNIVERSITY** and **DATAVIZGUY.COM**

72

Copyright

Published by:  
H&S Advisory Services Pty Ltd  
ACN 617 466 538  
ABN 14 617 466 538

This publication is subject to copyright. Other than for the purposes of and subject to the conditions prescribed under the Copyright Act 1968 (as amended), no part of it may in any form or by any means (electronic, mechanical, microcopying, photocopying, recording or otherwise) be reproduced, stored in a retrieval system or transmitted without prior written permission of the copyright owner.

Enquiries should be addressed to H&S Advisory Services Pty Ltd.

Warning

Unauthorised reproduction in whole or in part is an infringement of copyright. H&S Advisory Services Pty Ltd will actively pursue any breach of its copyright. Distribution of this material is strictly limited to attendees of the related training course and only with prior writing agreement from H&S Advisory Services Pty Ltd.

Disclaimer

All training materials are published and distributed on the basis that the publisher is not responsible for the results of any actions taken by users of information contained in this training manual on the basis of information contained in this manual nor for any error in or omission from this manual.

H&S Advisory Services Pty Ltd does not accept any responsibility whatsoever for misrepresentation by any person whatsoever of the information contained in this training material and expressly disclaims all and any liability and responsibility to any person, whether a reader of this training materials or not, in respect of claims, losses or damage or any other matter, either direct or consequential arising out of or in relation to the use and reliance, whether wholly or partially, upon any information contained or products referred to in this manual.



E: [hodge.paul@datavizguy.com](mailto:hodge.paul@datavizguy.com)  
W: [www.DataVizGuy.com](http://www.DataVizGuy.com)  
T: @DataVizGuy

Copyright 2020 @DataVizGuy

Limited License granted to:

