

Introduction to Big Data and its relevance to CAs

Amar Nathwani

Agenda

- What is Big Data
- Steps to follow in using Big Data
- Role and relevance to Members in Business and Practice
- Going forward

What is Big Data

Or when do you have a Big Data problem?

Big Data is characterised by:

1. Volume
2. Velocity
3. Variety

Using Big Data

Big Data become valuable when we can extract insights from it to support ***evidence based decision making.***

Moneyball

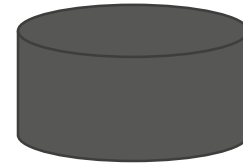
Using Big Data

- 1. What question do you want answered?***
2. Identify your data sources
3. Wrangle the data
4. Run your analysis
5. Communicate the results

Using Big Data

Data Sources

1. Internal vs external
2. Structured vs unstructured
3. Different repositories and formats



Using Big Data

Wrangling

1. Need to extract the data from the different sources
2. Correct the data
3. Map relationships
4. Combine the data

Wrangling is the most time consuming part of the process

Using Big Data

The nutritional value of Weetbix from 2 different websites

Messy Data

Nutrition Facts

Serving Size: 2 biscuits (30 g)

per serve	
Kilojoules	448 kJ
Calories	107 kcal
Protein	3.7 g
Fat	0.4 g
Saturated Fat	0.1 g
Carbohydrate	20 g
Sugar	1.1 g
Fibre	3.3 g
Sodium	87 mg
Potassium	102 mg

Calories	527	Sodium	1,215 mg
Total Fat	13 g	Potassium	254 mg
Saturated	2 g	Total Carbs	90 g
Polyunsaturated	4 g	Dietary Fiber	8 g
Monounsaturated	2 g	Sugars	21 g
Trans	0 g	Protein	14 g
Cholesterol	0 mg		
Vitamin A	16%	Calcium	16%
Vitamin C	33%	Iron	24

Using Big Data

Analysis

1. Visualisation
2. Statistical methods
3. Machine learning

Using Big Data

Analysis

Visualisation

Empired Ltd Example

Using Big Data

Analysis

Statistical analysis

Common methods include

- Regression analysis - Linear, multivariate, logical etc
- Decision trees
- K-means clustering
- Markov chains

Using Big Data

Machine Learning

An AI concept - Computer trains itself without needing pre-defined rules

1. Provide training data to system
2. System finds patterns in the data
3. Enter data you want analysed to get the answers you want

Google Translate

Using Big Data

Analysis

The completeness and quality of the data being analysed is important



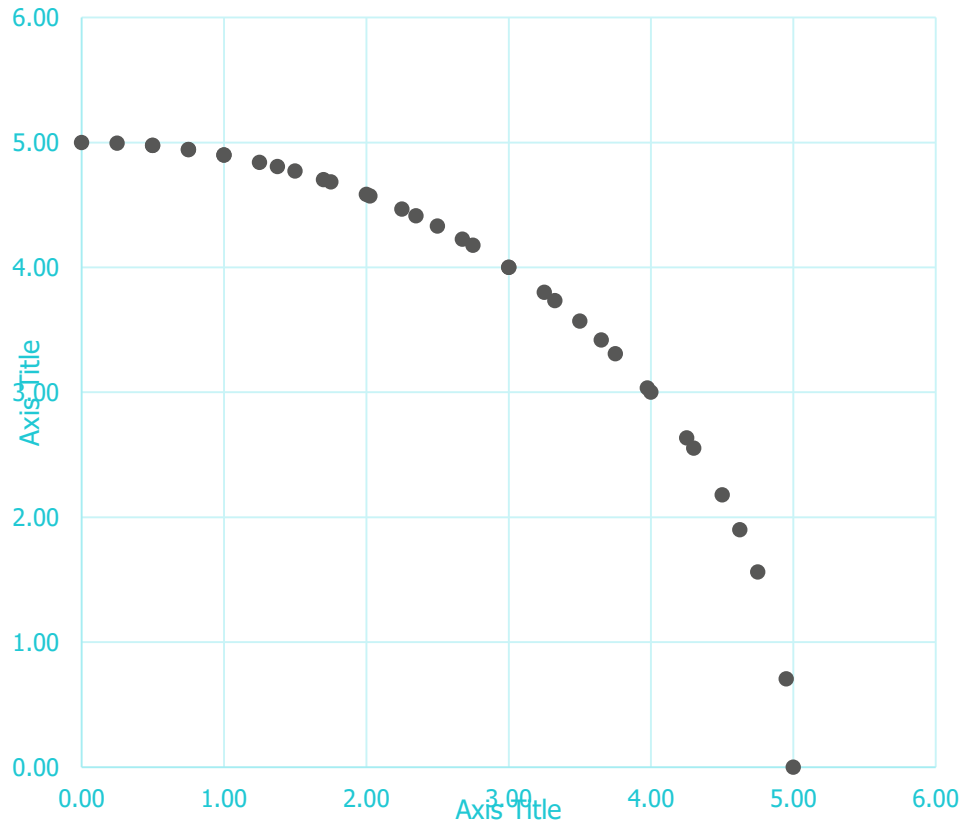
Using Big Data

Communicating the results

x	0.50	0.75	1.00	1.38	1.70	2.03	2.35	2.68
	3.00	3.33	3.65	3.98	4.30	4.63	4.95	5.00
	4.75	4.50	4.25	4.00	3.75	3.50	3.25	3.00
	2.75	2.50	2.25	2.00	1.75	1.50	1.25	1.00
	0.75	0.50	0.25					
y	4.97	4.94	4.90	4.81	4.70	4.57	4.41	4.22
	4.00	3.73	3.42	3.03	2.55	1.90	0.71	0.00
	1.56	2.18	2.63	3.00	3.31	3.57	3.80	4.00
	4.18	4.33	4.47	4.58	4.68	4.77	4.84	4.90
	4.94	4.97	4.99					

Using Big Data

Communicating the results



Using Big Data

Visualisation tools

- Lots of software tools to present data in a visual form
- Tableau, Qlik and Power BI are some of the better known ones
- Best visualisation to date

Using Big Data

Big Data tools



Image sourced from the Gartner Magic Quadrant for BI and Analytics Platforms 2017 report

Role and Relevance of Big Data to CAs

Detail text, 20pt

Role and relevance to us

- Technology is automating routine accounting tasks
- We are increasingly valued on our ability to deliver insights
- Mindset will need to change from pulling reports together to thinking about the business and asking questions
- Our advantages include:
 - Analytical training
 - Channels of communication with all areas of the business
 - Responsibility for internal controls
- Accountants will need good soft skills to challenge established practice and communicate findings
- Helps future proof our careers

CA ANZ WA Big Data Working Group

Going forward

- Have a play and experiment with some of the tools
- Take small steps – don't rush into a large project
- Find out how your sector is using big data/ big data tools
- This field is moving too quickly for one person to keep up with
 - Working group to share ideas, war stories, tips, organise training, get speakers etc?
- Questions/ comments

Amar Nathwani

Nexia

amar.nathwani@nexiaperth.com.au

08 9463 2463