SIT718 Real World Analytics

Lecturer: Dr Ye Zhu

School of Information Technology
Deakin University

Week 1: Introduction to Data Analysis

Learning Outcomes

These are the Learning Outcomes (ULO) for this Unit:

- LO1: Apply knowledge of multivariate functions data transformations and data distributions to summarise data sets. (week 1 – week 4)
- LO2: Analyse datasets by interpreting summary statistics, model and function parameters. (week 5 – week 6)
- LO3: Apply game theory, and linear programming skills and models, to make optimal decisions. (week 7 – week 10)
- LO4: Develop software codes to solve computational problems for real world analytics. (Pracs/Workshops)
- LO5: Demonstrate professional ethics and responsibility for working with real world data. (Assessments)

Learning Outcomes 2

In light of the unit learning objectives, the students will be able to do following practical outcomes:

- meaningfully use (numerical) data ? the operations we apply make sense, the output has a straightforward interpretation and is useful for making decisions;
- critically interpret outputs given by others;
- use data responsibly and professionally;
- understand arithmetic, geometric, harmonic means and the Power means, which generalise the geometric, harmonic and arithmetic mean;
- make better decisions through mathematical methods in optimisation problems;
- do some basic calculation and analysis in R.

UNIT LOGISTICS

1. 1 x 2 hour class and 1 x 2 hour workshop per week.
2. Peer support sessions: https://d2l.deakin.edu.au/d2l/home/965803
3. Maths Mentors: www.deakin.edu.au/maths-mentors
4. IT Help: https://www.deakin.edu.au/students/help/it-help
5. Discussions Forum

No attendance requirement, recordings are available for review

- Contact **Unit Chair** (ye.zhu@deakin.edu.au) for all admin issues, special consideration, request for extension and complaints.
- Contact **Tutors** for all workshops/marking issues:
  - Delaram Pahlevani d.pahlevani@deakin.edu.au
  - Anagi Gamachchi a.gamachchi@deakin.edu.au

Make sure to include the unit name ('SIT718') in the subject line of your email and your student ID in the body of your email.

ASSESSMENTS

1. Assessment 1 Online quizzes x 5 – each 4% total mark
2. Assessment 2 (Problem solving) – 20% total mark
   report in pdf format, software code and/or data
3. Assessment 3 (Problem solving) – 30% total mark
   report in pdf format, software code and/or data
4. Assessment 4 (Examination) – 30% total mark
   2 hours, Online quiz with open book.

No hurdle at all, only get overall 50 scores to pass this unit

Your assignment 2 and 3 will not be assessed if the code is missing or the outputs of the code are inconsistent with the report. You must paraphrase and cite references correctly in your report. High Turnitin score of the report/code may lead to plagiarism investigation by the faculty.

HD questions need extra efforts and independent research.

Textbooks

Go to the **Reading List** in CloudDeakin:

1. Simon James. *An Introduction to Data Analysis using Aggregation Functions in R*. Springer 2016.
2. EMC Education services Editors. *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis, IN: John Wiley and Sons 2015.
3. Operations Research: Applications and Algorithms, by Wayne L. Winston, 2004

These books are free for Deakin students to view online. There are some further optional reading books.

WEEK 1: EXPLORATORY DATA ANALYSIS WITH R

This would introduce data types and processes for cleaning and preparing data for analysis. Working with R and R studio will be demonstrated as well as commands in R to compute input and visualise data. Scatter plot and histograms will be discussed.
This would include the following material:

1. Data types
2. Role of data analyst
3. Working with data
4. Ethics for Data Science
5. Introduction to R
6. Math refresher

Data Types

- *Data* and the more relevant now *Big Data* is creating significant new opportunities for organisations to derive new value and create competitive advantage from their most valuable asset: information.
- For businesses, *Big Data* helps drive efficiency, quality, and personalised products and services, producing improved levels of customer satisfaction and profit.
- For scientific purposes, *Data analytics* enable new avenues of investigation with potentially richer results and deeper insights than previously available.
- In many cases, *Data analytics* integrate structured and unstructured data with real-time feeds and queries, opening new paths to innovation and insight.

Some key concepts

- Data and Big Data overview
- State of the practice in analytics
- Intelligence versus Data Science
- Key roles for the new Big Data ecosystem
- The Data Scientist
- Examples of Data analytics

Big Data Overview

Data is created constantly, and at an ever-increasing rate. Mobile phones, social media, imaging technologies to determine a medical diagnosis, all these and more create new data.

Three attributes stand out as defining Big Data characteristics:

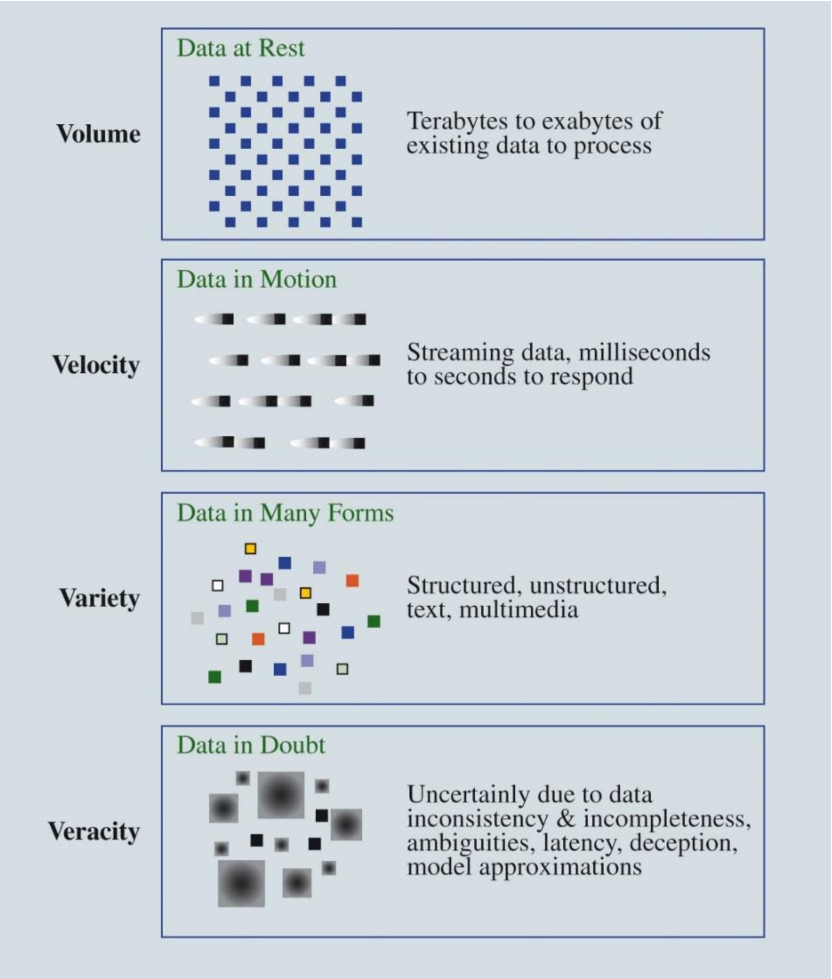
- Huge volume of data: Big Data can be billions of rows and millions of columns.
- Complexity of data types and structures: Big Data reflects the variety of new data sources, formats, and structures, including digital traces being left on the web and other digital repositories for subsequent analysis.
- Speed of new data creation and growth: Big Data can describe high velocity data, with rapid data ingestion and near real time analysis.

Big Data Overview

Another definition of Big Data comes from the McKinsey Global report from 2011:

Big Data is data whose scale, distribution, diversity, and/or timeliness require the use of new technical architectures and analytics to enable insights that unlock new sources of business value.

McKinsey & Co.; Big Data: The Next Frontier for Innovation, Competition, and Productivity.



Volume

Data at Rest

Terabytes to exabytes of existing data to process

Velocity

Data in Motion

Streaming data, milliseconds to seconds to respond

Variety

Data in Many Forms

Structured, unstructured, text, multimedia

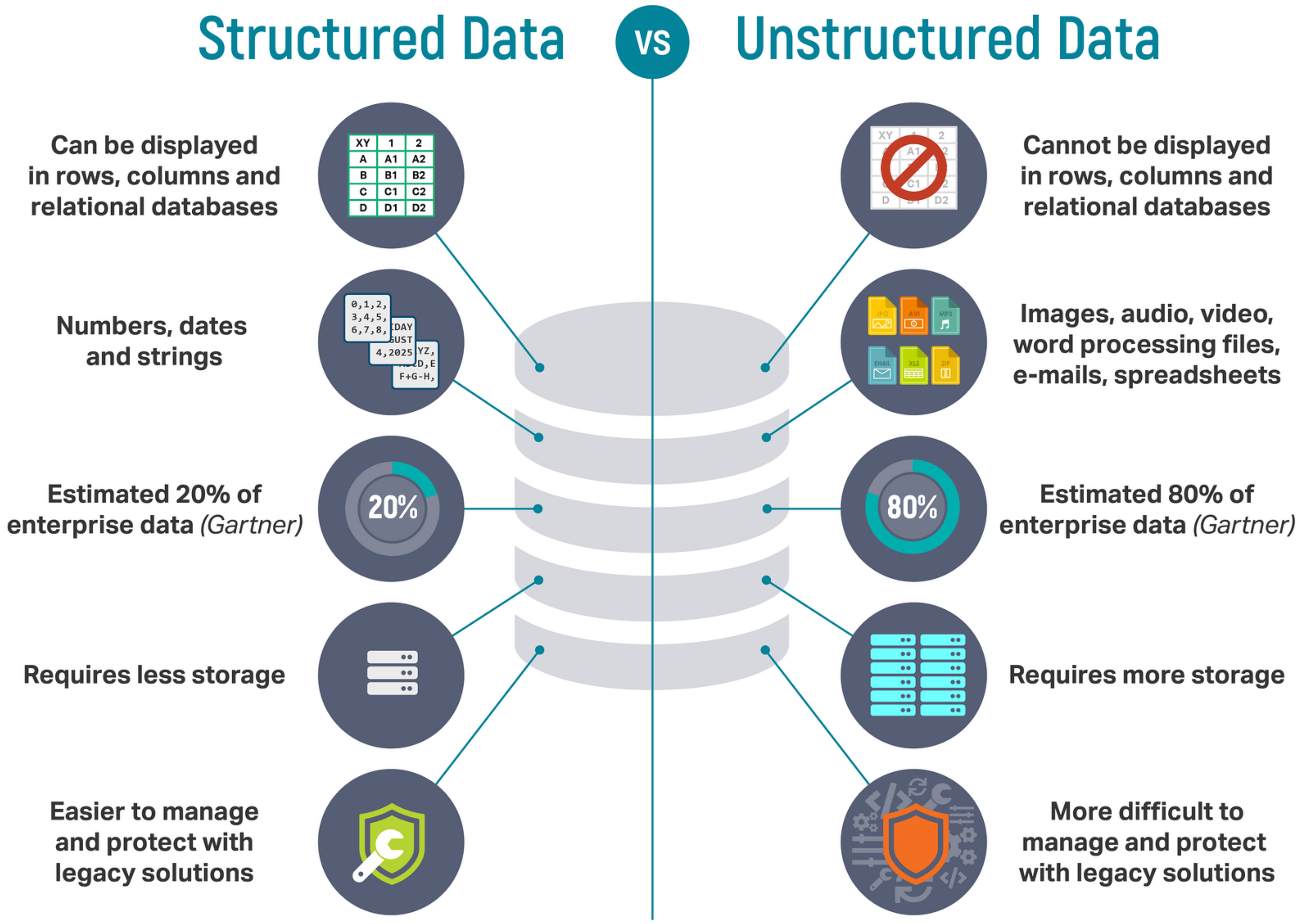
Veracity

Data in Doubt

Uncertainly due to data inconsistency & incompleteness, ambiguities, latency, deception, model approximations

Examples of Data Structures

- Structured data: Data containing a defined data type, format, and structure (that is, transaction data, online analytical processing [OLAP] data cubes, traditional RDBMS, CSV files, and even simple spreadsheets).
- Semi-structured data: Textual data files with a discernible pattern that enables parsing (such as Extensible Markup Language [XML] data files that are self- describing and defined by an XML schema).
- Quasi-structured data: Textual data with erratic data formats that can be formatted with effort, tools, and time (for instance, web clickstream data that may contain inconsistencies in data values and formats).
- Unstructured data: Data that has no inherent structure, which may include text, documents, PDFs, images and video.



Structured Data vs Unstructured Data

Can be displayed in rows, columns and relational databases

Numbers, dates and strings

Estimated 20% of enterprise data *(Gartner)*

Requires less storage

Easier to manage and protect with legacy solutions

Cannot be displayed in rows, columns and relational databases

Images, audio, video, word processing files, e-mails, spreadsheets

Estimated 80% of enterprise data *(Gartner)*

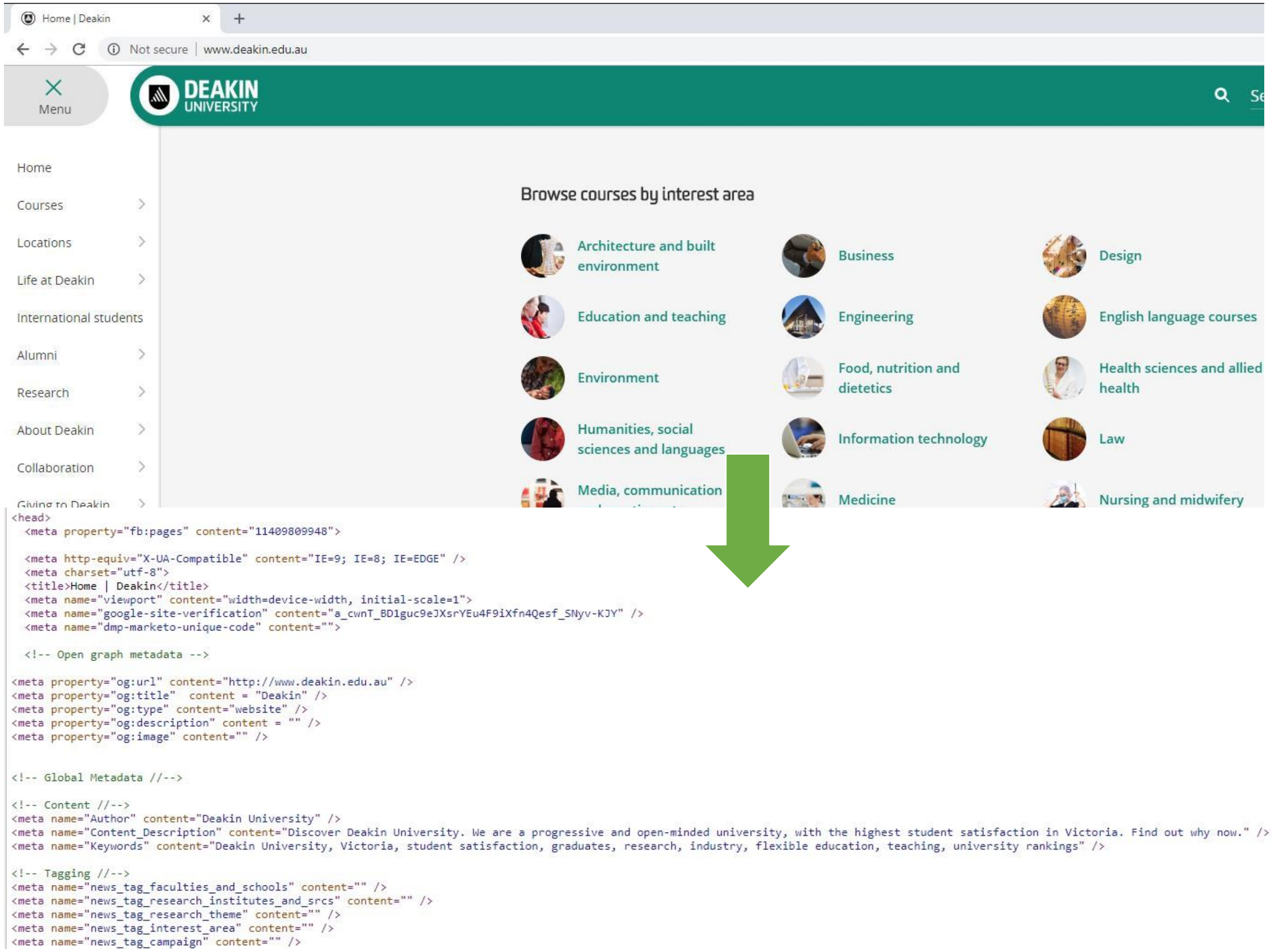
Requires more storage

More difficult to manage and protect with legacy solutions

https://lawtomated.com/structured-data-vs-unstructured-data-what-are-they-and-why-care/

Original website

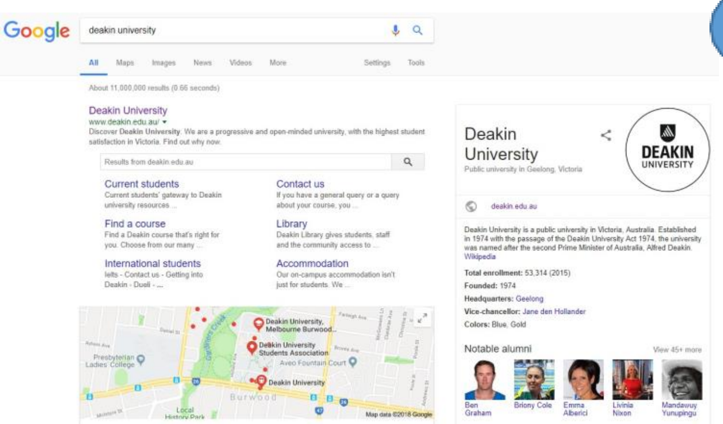
View page source



```
<head>
  <meta property="fb:pages" content="11409809948">

  <meta http-equiv="X-UA-Compatible" content="IE=9; IE=8; IE=EDGE" />
  <meta charset="utf-8">
  <title>Home | Deakin</title>
  <meta name="viewport" content="width=device-width, initial-scale=1">
  <meta name="google-site-verification" content="a_cwnT_BD1guc9eJXsrYEu4F9iXfn4Qesf_SNyv-KJY" />
  <meta name="dmp-marketo-unique-code" content="">

  <!-- Open graph metadata -->

<meta property="og:url" content="http://www.deakin.edu.au" />
<meta property="og:title"  content = "Deakin" />
<meta property="og:type" content="website" />
<meta property="og:description" content = "" />
<meta property="og:image" content="" />


<!-- Global Metadata //-->

<!-- Content //-->
<meta name="Author" content="Deakin University" />
<meta name="Content_Description" content="Discover Deakin University. We are a progressive and open-minded university, with the highest student satisfaction in Victoria. Find out why now." />
<meta name="Keywords" content="Deakin University, Victoria, student satisfaction, graduates, research, industry, flexible education, teaching, university rankings" />

<!-- Tagging //-->
<meta name="news_tag_faculties_and_schools" content="" />
<meta name="news_tag_research_institutes_and_srcs" content="" />
<meta name="news_tag_research_theme" content="" />
<meta name="news_tag_interest_area" content="" />
<meta name="news_tag_campaign" content="" />
```

Figure: Example of semi-structured data

1



Google

deakin university

All Maps Images News Videos More Settings Tools

About 11,000,000 results (0.66 seconds)

Deakin University
www.deakin.edu.au/
Discover **Deakin University**. We are a progressive and open-minded university, with the highest student satisfaction in Victoria. Find out why now.

Results from deakin.edu.au

Current students
Current students' gateway to Deakin university resources ...

Contact us
If you have a general query or a query about your course, you ...

Find a course
Find a Deakin course that's right for you. Choose from our many ...

Library
Deakin Library gives students, staff and the community access to ...

International students
Ielts - Contact us - Getting into Deakin - Duell - ...

Accommodation
Our on-campus accommodation isn't just for students. We ...

Deakin University, Melbourne Burwood...
Deakin University Students Association
Aveo Fountain Court
Presbyterian Ladies College
Deakin University
Burwood
Local History Park
Map data ©2018 Google

Deakin University
Public university in Geelong, Victoria

DEAKIN UNIVERSITY

deakin.edu.au

Deakin University is a public university in Victoria, Australia. Established in 1974 with the passage of the Deakin University Act 1974, the university was named after the second Prime Minister of Australia, Alfred Deakin. Wikipedia

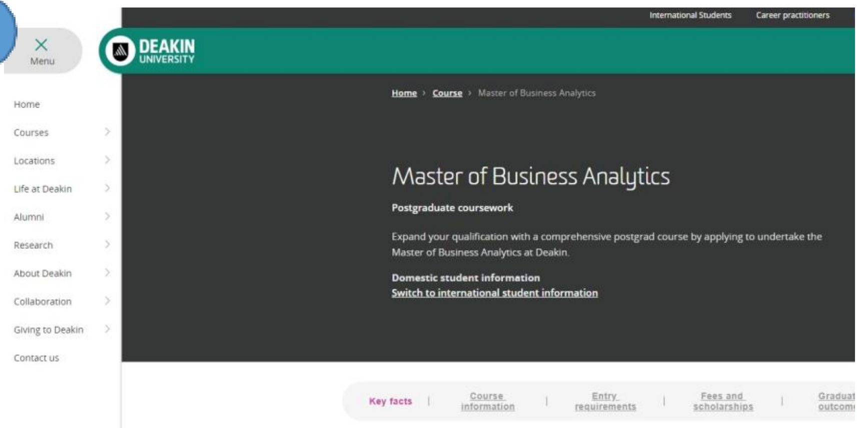
Total enrollment: 53,314 (2015)
Founded: 1974
Headquarters: Geelong
Vice-chancellor: Jane den Hollander
Colors: Blue, Gold

Notable alumni — View 45+ more

Ben Graham, Briony Cole, Emma Alberici, Livinia Nixon, Mandawuy Yunupingu

https://www.google.com.au/

2



International Students — Career practitioners

DEAKIN UNIVERSITY

Menu

- Home
- Courses
- Locations
- Life at Deakin
- Alumni
- Research
- About Deakin
- Collaboration
- Giving to Deakin
- Contact us

Home › Course › Master of Business Analytics

Master of Business Analytics

Postgraduate coursework

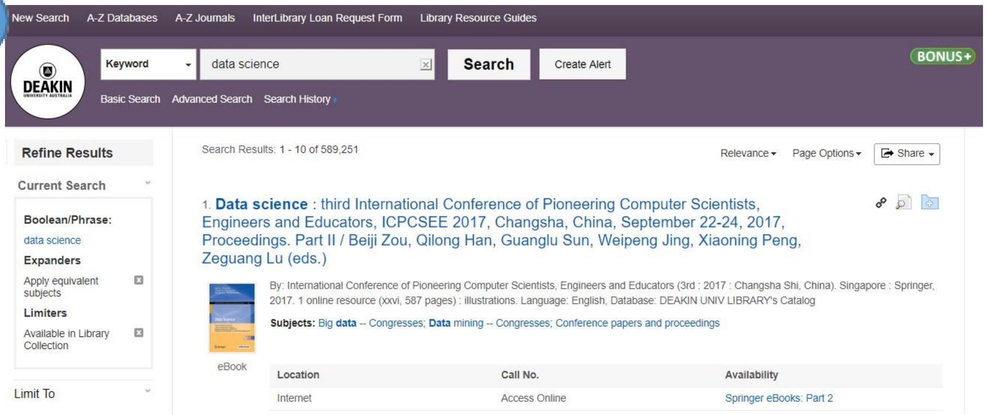
Expand your qualification with a comprehensive postgrad course by applying to undertake the Master of Business Analytics at Deakin.

Domestic student information
Switch to international student information

Key facts | Course information | Entry requirements | Fees and scholarships | Graduat outcom

http://www.deakin.edu.au/course/master-business-analytics

3



New Search   A-Z Databases   A-Z Journals   InterLibrary Loan Request Form   Library Resource Guides

DEAKIN UNIVERSITY AUSTRALIA

Keyword | data science | Search | Create Alert

BONUS+

Basic Search   Advanced Search   Search History

Refine Results

Current Search

Boolean/Phrase:
data science

Expanders
Apply equivalent subjects

Limiters
Available in Library Collection

Limit To

Search Results: 1 - 10 of 589,251   Relevance ▾   Page Options ▾   Share ▾

1. **Data science** : third International Conference of Pioneering Computer Scientists, Engineers and Educators, ICPCSEE 2017, Changsha, China, September 22-24, 2017, Proceedings. Part II / Beiji Zou, Qilong Han, Guanglu Sun, Weipeng Jing, Xiaoning Peng, Zeguang Lu (eds.)

By: International Conference of Pioneering Computer Scientists, Engineers and Educators (3rd : 2017 : Changsha Shi, China). Singapore : Springer, 2017. 1 online resource (xxvi, 587 pages) : illustrations. Language: English, Database: DEAKIN UNIV LIBRARY's Catalog

Subjects: Big **data** -- Congresses; **Data** mining -- Congresses; Conference papers and proceedings

eBook

Location	Call No.	Availability
Internet	Access Online	Springer eBooks: Part 2

https://www.deakin.edu.au/library

Figure: Example of quasi-structured data

Figure: Example of unstructured data.

Data architecture

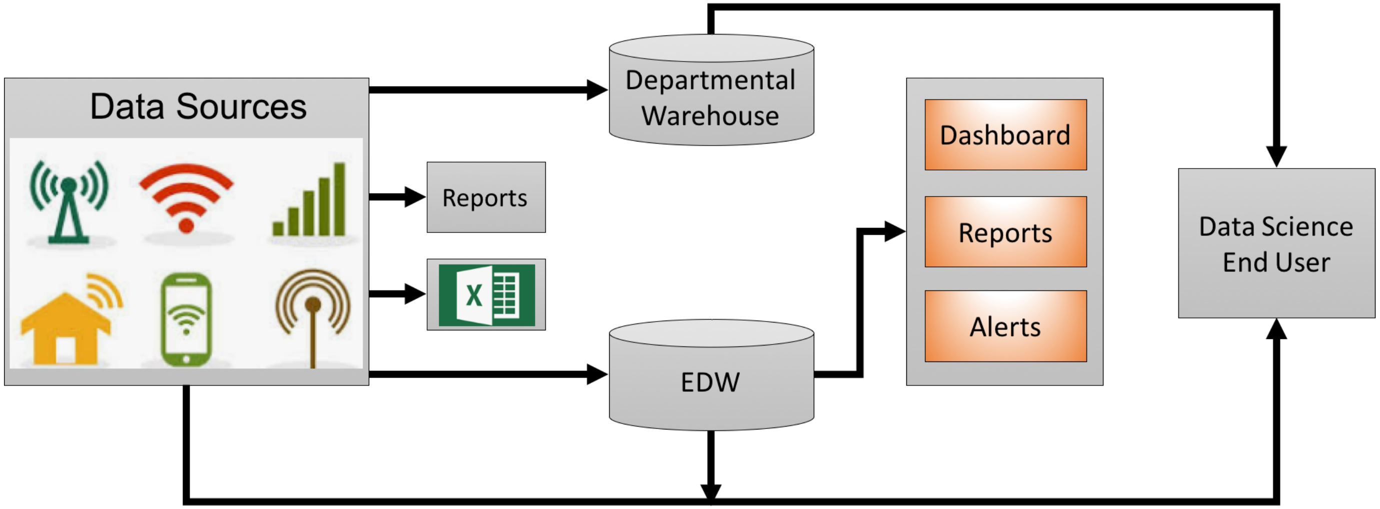


Figure: Typical data architecture.

Big Data ecosystems

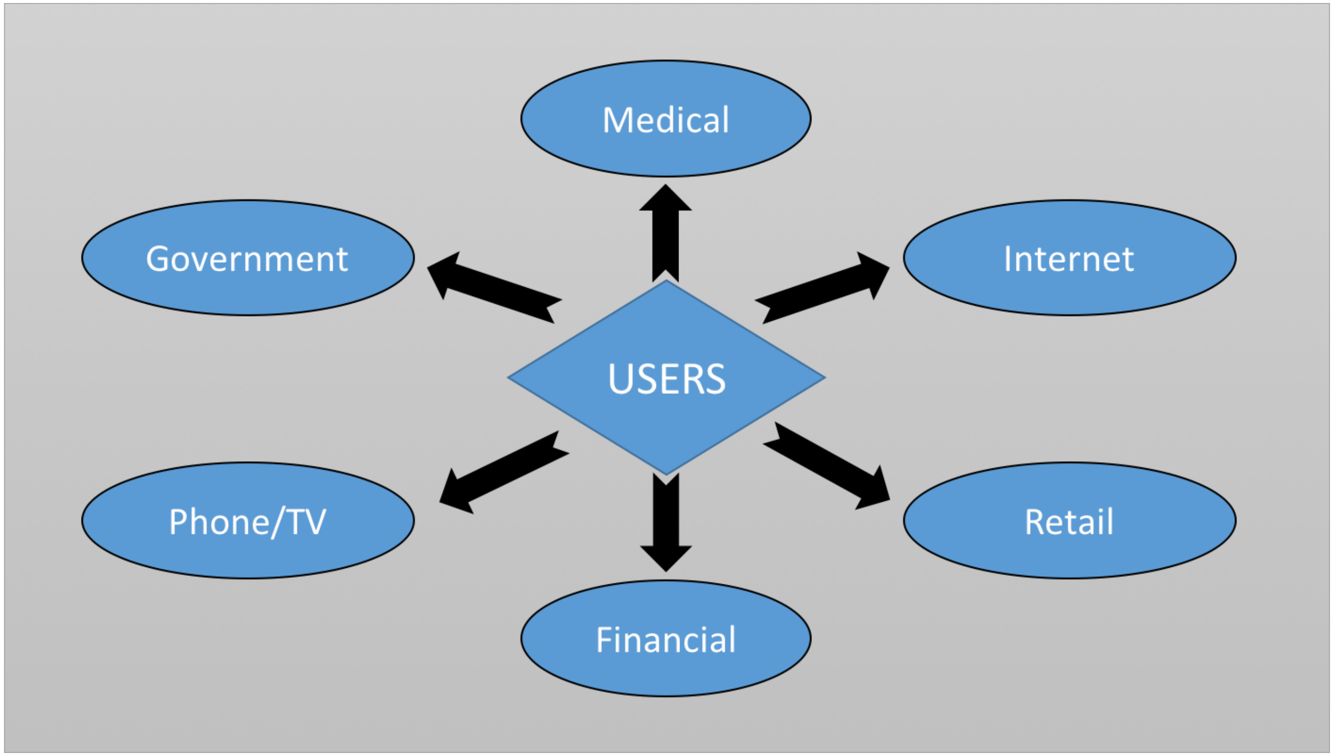
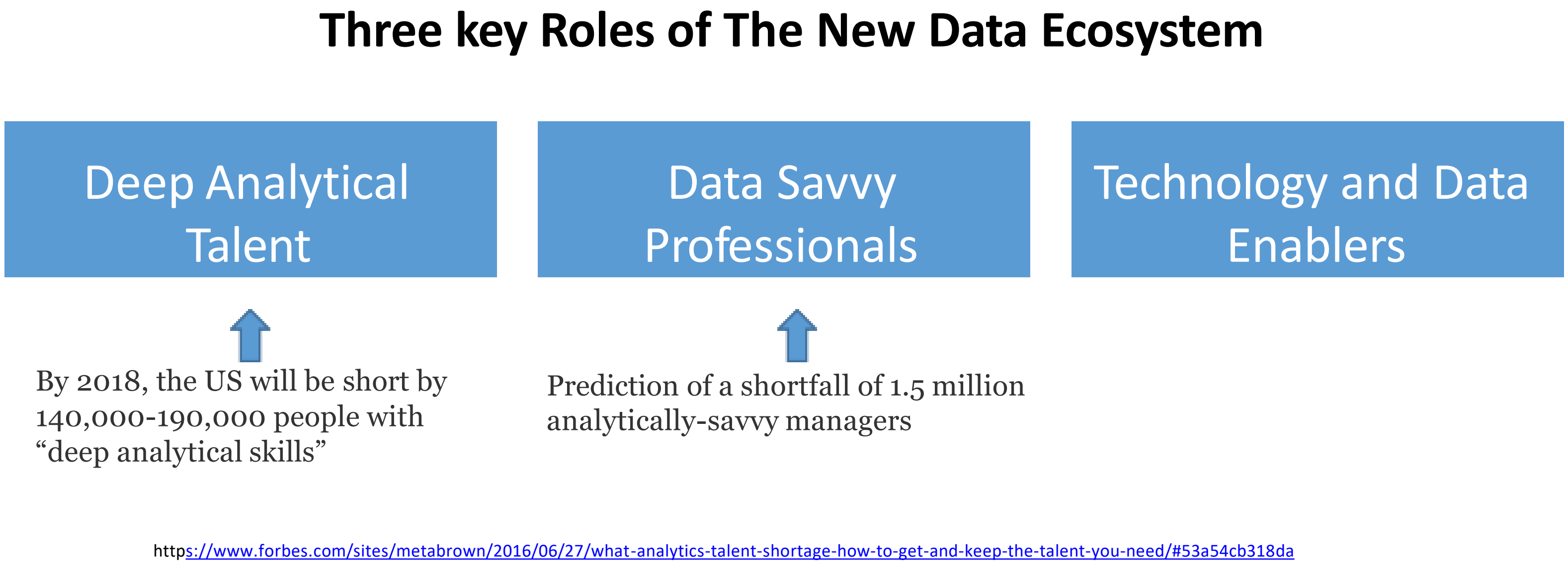


Figure: Emerging Big Data ecosystems.

Big Data ecosystems



Three key Roles of The New Data Ecosystem

Deep Analytical Talent

Data Savvy Professionals

Technology and Data Enablers

By 2018, the US will be short by 140,000-190,000 people with "deep analytical skills"

Prediction of a shortfall of 1.5 million analytically-savvy managers

https://www.forbes.com/sites/metabrown/2016/06/27/what-analytics-talent-shortage-how-to-get-and-keep-the-talent-you-need/#53a54cb318da

Figure: Key roles of the new Big Data ecosystem.

Profile of data analyst

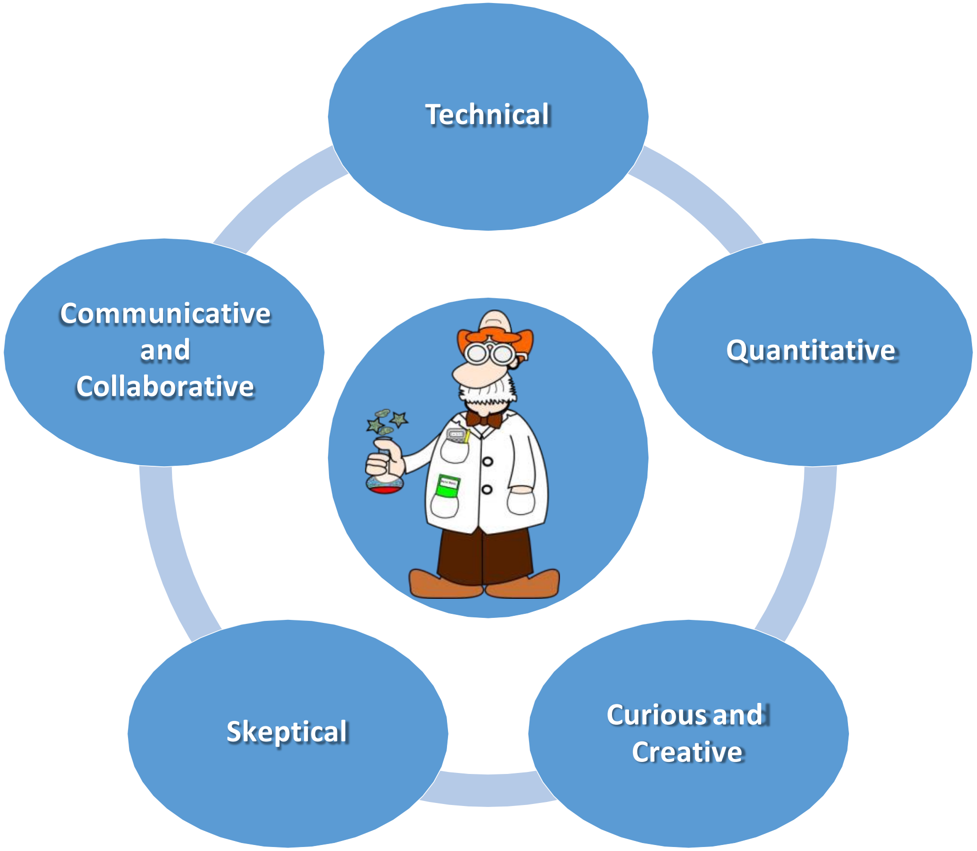


Figure: Profile of data analyst.

Business analytics

- Scientific process of transforming data into insight for making better decisions.
- Used for data-driven or fact-based decision making, which is often seen as more objective than other alternatives for decision making.

Tools of business analytics can aid decision making by:

- Creating insights from data.
- Improving our ability to more accurately forecast for planning.
- Helping us quantify risk.
- Yielding better alternatives through analysis and optimization.

Analytical Methods and Models

Descriptive Analytics:

Descriptive analytics: Encompasses the set of techniques that describes what has happened in the past

Data query: A request for information with certain characteristics from a database.

Data dashboards: Collections of tables, charts, maps, and summary statistics that are updated as new data become available

Data mining: The use of analytical techniques for better understanding patterns and relationships that exist in large data sets.

Predictive Analytics:

Predictive analytics: Consists of techniques that use models constructed from past data to predict the future or ascertain the impact of one variable on another. Survey data and past purchase behavior may be used to help predict the market share of a new product.

Techniques used in Predictive Analytics include:

- Linear regression.
- Time series analysis.
- Data mining is used to find patterns or relationships among elements of the data in a large database; often used in predictive analytics.
- **Simulation** involves the use of probability and statistics to construct a computer model to study the impact of uncertainty on a decision.

Predictive Analytics:

Indicates a best course of action to take:

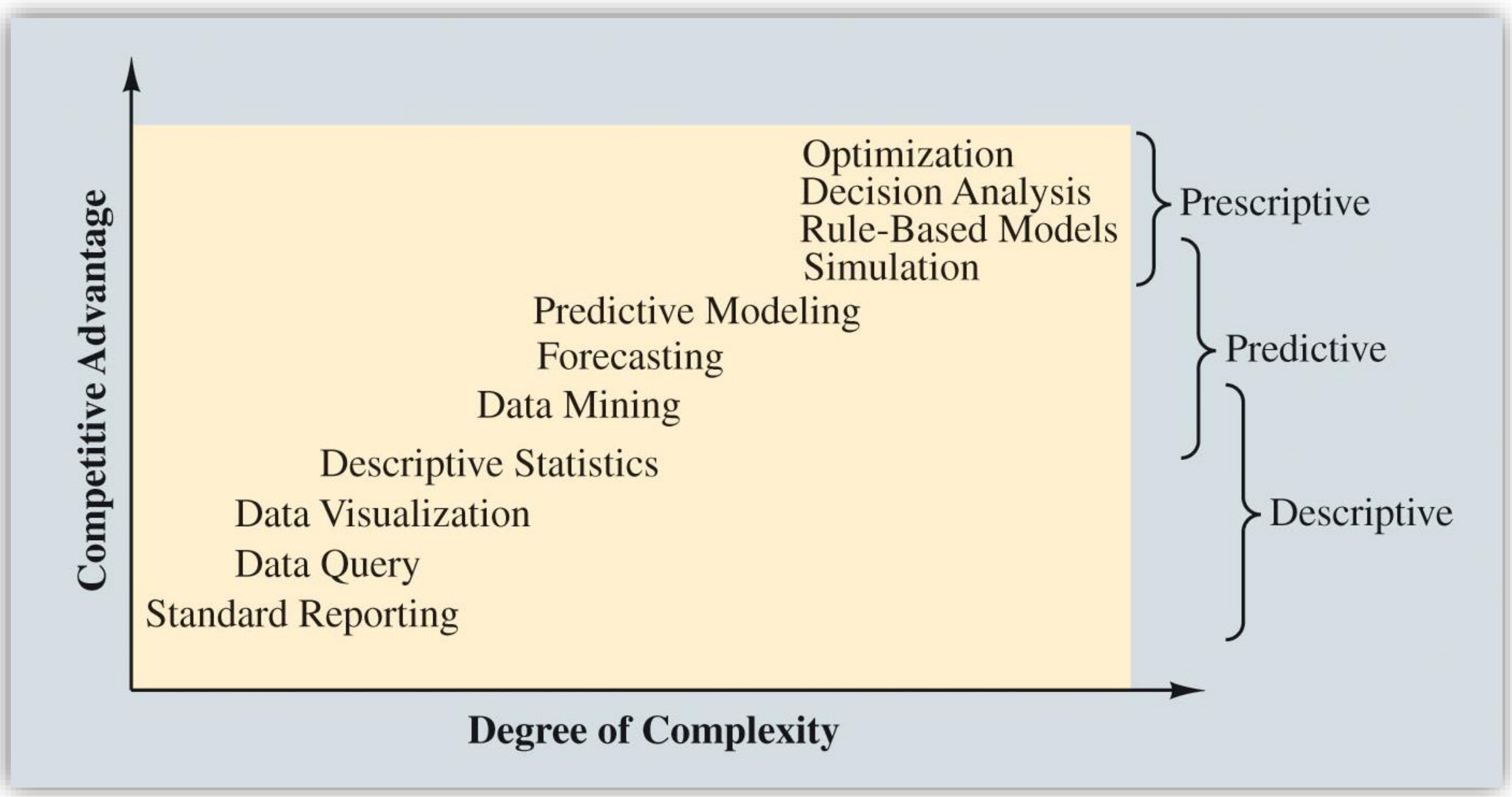
- Provide a forecast or prediction, but do not provide a decision.
- A forecast or prediction, when combined with a rule, becomes a prescriptive model.
- Prescriptive models that rely on a rule or set of rules are often referred to as **rule-based models.**

Optimisation models: Models that give the best decision subject to constraints of the situation.

Simulation optimisation: Combines the use of probability and statistics to model uncertainty with optimisation techniques to find good decisions in highly complex and highly uncertain settings.

Decision analysis: Used to develop an optimal strategy when a decision maker is faced with several decision alternatives and an uncertain set of future events. based on the decision maker's attitude toward risk, loss, and other factors.

The Spectrum of Business Analytics



Source: Adapted from SAS

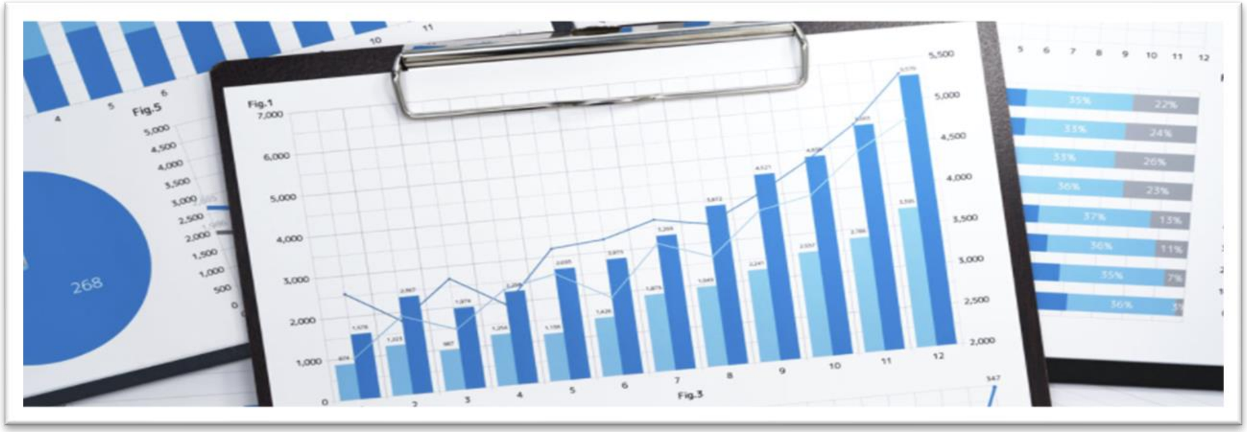
Business Analytics in Practice

Financial Analytics

Use of predictive models to:

- Forecast financial performance.
- Assess the risk of investment portfolios and projects.
- Construct financial instruments such as derivatives.
- Construct optimal portfolios of investments.
- Allocate assets.
- Create optimal capital budgeting plans.

Simulation is also often used to assess risk in the financial sector.



Health Care Analytics

Descriptive, predictive, and prescriptive analytics are used to improve:

- Patient, staff, and facility scheduling.
- Patient flow.
- Purchasing.
- Inventory control.

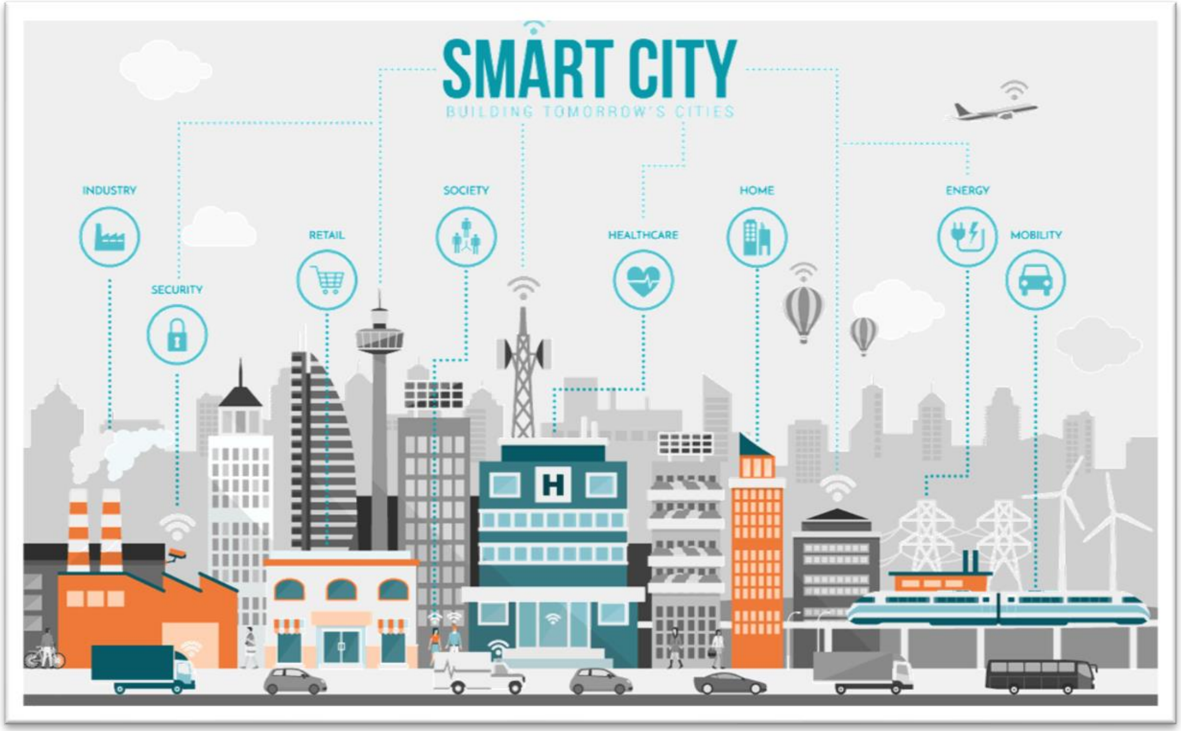
Use of prescriptive analytics for diagnosis and treatment may prove to be the most important application of analytics in health care.

Analytics for Government and Nonprofits

Analytics for government to:

- Drive out inefficiencies.
- Increase the effectiveness and accountability of programs.

Analytics for nonprofit agencies to ensure their effectiveness and accountability to their donors and clients.



Take home message:

- Considered type of data
- Considered the role of data science in modern company framework
- Considered role of data analyst
- What the industry needs is the excellent communication and presentation skills of any data analyst.

Working with Data

- Code of practice for a data analyst
- Personal Data
- Data Protection
- Data Security
- Common sense when working with data

Legal and Ethical Issues

- Increased attention has been paid to ethical concerns around data privacy and the ethical use of models based on data.
- Companies have an obligation to protect the data and to not misuse that data.
- Clients and customers have an obligation to understand trade-offs between allowing their data to be collected, and the benefits they accrue from allowing a company to collect and use that data.
- An agreement must be signed between the customer and the company.
- Stipulations:

  The request for consent to use an individual's data must be easily understood and accessible.

  The intended use of data must be specified.

  Must be easy to withdraw consent.

  The individual has a right to a copy of their data and the right to demand their data be erased.

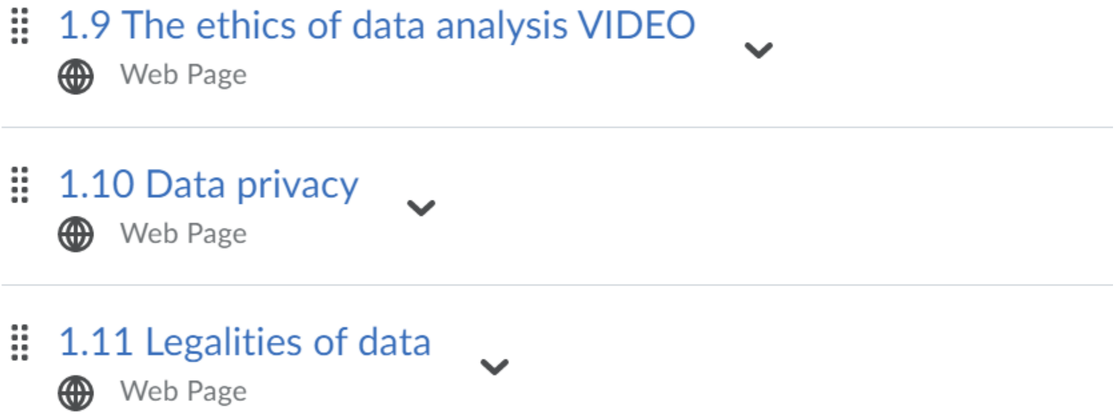
Legal and Ethical Issues (cont.)

- Analytics professionals have a responsibility to behave ethically.
- This includes protecting data, being transparent about the data and how it was collected, and what it does and does not contain.
- Analysts must be transparent about the methods used to analyze the data and any assumptions that have to be made for the methods used.
- Analysts must provide valid conclusions and understandable recommendations to their clients.
- The American Statistical Association (ASA) and the Institute for Operations Research and the Management Sciences (INFORMS) provide ethical guidelines for analysts.
- “Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations.”

Ethics in Data Analytics - Home Reading

Ethics in Data Analytics Australia introduced in February 2018 the mandatory data breach notification. It means that an organisation that either accidentally loses data, has been hacked and data have leaked, needs to publicly acknowledge to the government what has occurred from an incident perspective, which will have huge brand and reputational sort of impacts.

Read Weekly Resources:



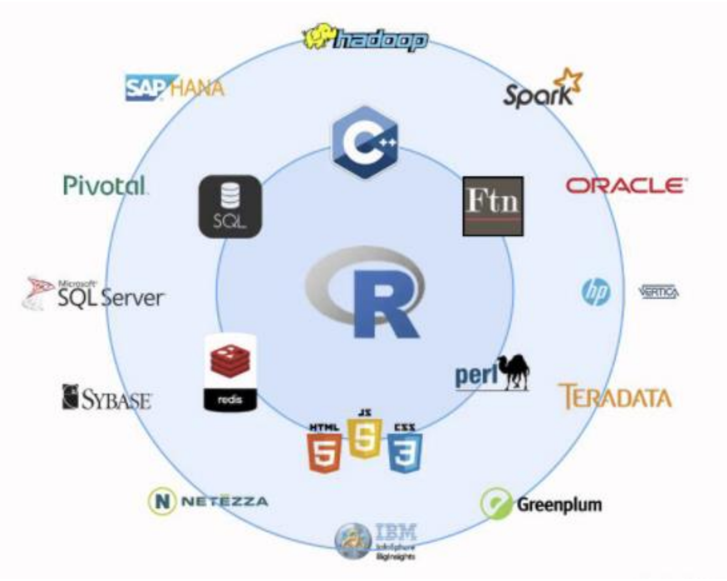
1.9 The ethics of data analysis VIDEO
Web Page

1.10 Data privacy
Web Page

1.11 Legalities of data
Web Page

Introduction to R

Why R? Being an **open source and free** software and having a variety of built in statistical commands, R is the most widely used tool by statisticians and Data analysts. There are a lot of new packages being developed and old packages constantly updated in R to support and handle **Big Data**.



1.Visit https://cran.r-project.org/. At the top you should see three links for downloading R depending on the platform of your choice.

2.Visit https://www.rstudio.com/products/rstudio/download/#download and select the installer for your operating system.

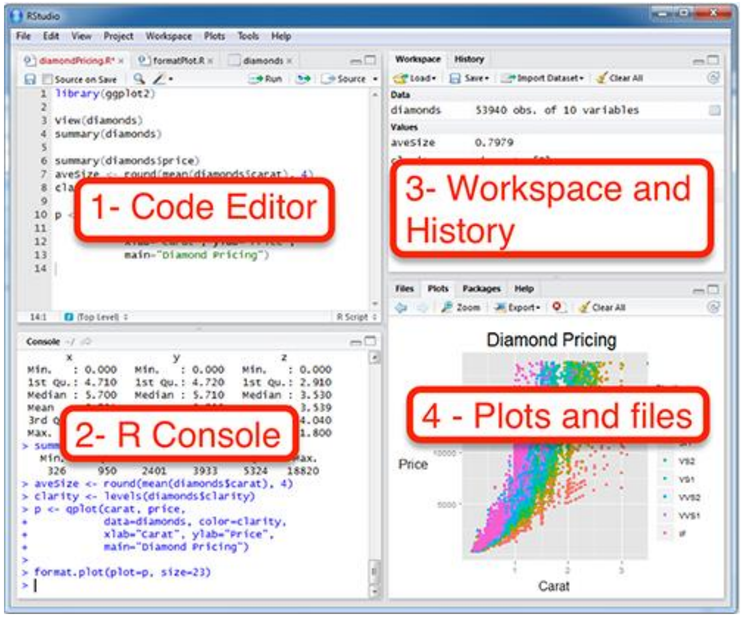
You must have R installed before installing RStudio

Learning R

Basically, please follow the instructions in Workshops and learn with our tutors.

Some other optional learning materials:
https://www.youtube.com/watch?v=_V8eKsto3Ug
https://www.youtube.com/watch?v=eDrhZb2onWY&ab_channel=edureka%21

If you would like to learn advanced skills about R programming, you can join the free online short course at https://www.coursera.org/learn/r-programming/



```
library(ggplot2)

View(diamonds)
summary(diamonds)

summary(diamonds$price)
aveSize <- round(mean(diamonds$carat), 4)
cla

p <


             main="Diamond Pricing")

```

```
     x              y              z
Min.   : 0.000   Min.   : 0.000   Min.   : 0.000
1st Qu.: 4.710   1st Qu.: 4.720   1st Qu.: 2.910
Median : 5.700   Median : 5.710   Median : 3.530
Mean                                       3.539
3rd Q                                      4.040
Max.                                       1.800
> summ
  Min.                                     Max.
   326      950     2401     3933     5324    18820
> aveSize <- round(mean(diamonds$carat), 4)
> clarity <- levels(diamonds$clarity)
> p <- qplot(carat, price,
+            data=diamonds, color=clarity,
+            xlab="Carat", ylab="Price",
+            main="Diamond Pricing")
>
> format.plot(plot=p, size=23)
> |
```

```
diamonds     53940 obs. of 10 variables
aveSize      0.7979
```

1- Code Editor

2- R Console

3- Workspace and History

4 - Plots and files

There is a package called **Swirl** that lets you learn the basics of R in R:
https://swirlstats.com/students.html

There's also a course version of the guide 'A (very) Short Introduction to R' in our week 2 materials:
http://swirlstats.com/scn/A_(very)_short_introduction_to_R.html

Other courses are listed here:
http://swirlstats.com/scn/title.html

Math refresher

The purpose of the blow links (Maths refresher) are to help revise/recall the high school maths, such as done in year 8 to year 11 on some of the topics that will be useful in this unit at various stages.

Refresher for Week 1-2

Algebra: median, mean, matrices, and indices

https://amsi.org.au/ESA_Senior_Years/SeniorTopic1/1_md/SeniorTopic1.html

https://amsi.org.au/ESA_Senior_Years/SeniorTopic2/2_md/SeniorTopic2.html

https://revisionmaths.com/gcse-maths-revision/number

https://www.mathsisfun.com/data/index.html

Operations

Square Root For $a \in R$, where a is positive or zero, the notation $\sqrt{a}$ signifies the square root of a. It is the nonnegative real number whose product with itself gives a; thus $\sqrt{a}\sqrt{a} = a$. The notation $\sqrt{a}$ is sometimes called the principal square root [Stewart et al., 1992]. The negative square root of a is $-\sqrt{a}$.

***Example*:** Let $a = 169$; then $\sqrt{a}$ is 13 since $13 \times 13 = 169$. The number $a = 169$ has the negative square root $-\sqrt{169} = -13$.

Absolute Value For all $a \in R$, if a is positive or zero, then $|a| = a$; if a is negative, then $|a| = -a$.

***Example*:** Let $a = \{5, -1, -6, -14.3\}$; then $|a| = \{5, 1, 6, 14.3\}$.

Greater Than Elements can be ordered using the concepts of greater than and less than. For all $a, b \in R$, b is less than a if for some positive real number c we have $a = b + c$. For such a condition, a is said to be greater than b, which may be written as $a > b$. Alternatively, a is greater than b if $a - b > 0$. Similarly, we may say that b is less than a and write it as $b < a$ if $a - b > 0$.

***Example*:** Let $a = 3$ and $b = -2$. Is a greater than b? To satisfy $b + c = a$, we must have $c = 5$. Since c is a positive real number, we know that $a > b$.

Axioms Let R be a nonempty set with two binary operations:

1. Addition (denoted by $+$), and
2. Multiplication (denoted by juxtaposition).

The nonempty set R is a ring if the following axioms hold:

[R1] Associative law of addition: For any $a, b, c \in R$, $(a+b)+c = a+(b+c)$.

[R2] Zero element: There exists an element $0 \in R$ such that $a + 0 = 0 + a = a$ for every $a \in R$.

[R3] Negative of a: For each $a \in R$ there exists an element $-a \in R$ such that $a + (-a) = (-a) + a = 0$.

[R4] Commutative law of addition: For any $a, b \in R$, $a + b = b + a$.

[R5] Associative law of multiplication: For any a, b, $c \in R$, $(ab)c = a(bc)$.

[R6] Distributive law: For any a, b, $c \in R$, we have

(i) $a(b + c) = ab + ac$, and

(ii) $(b + c)a = ba + ca$.

Subtraction is defined in R by $a - b \equiv a + (-b)$. R is a commutative ring if $ab = ba$ for every $a, b \in R$. R is a ring with a unit element if there exists a nonzero element $1 \in R$ such that $a \cdot 1 = 1 \cdot a = a$ for every $a \in R$.

Algebraic operations may be presented as a collection of axioms. In all cases assume $a, b, c, d \in R$. The following presents equality axioms:

EQUALITY AXIOMS

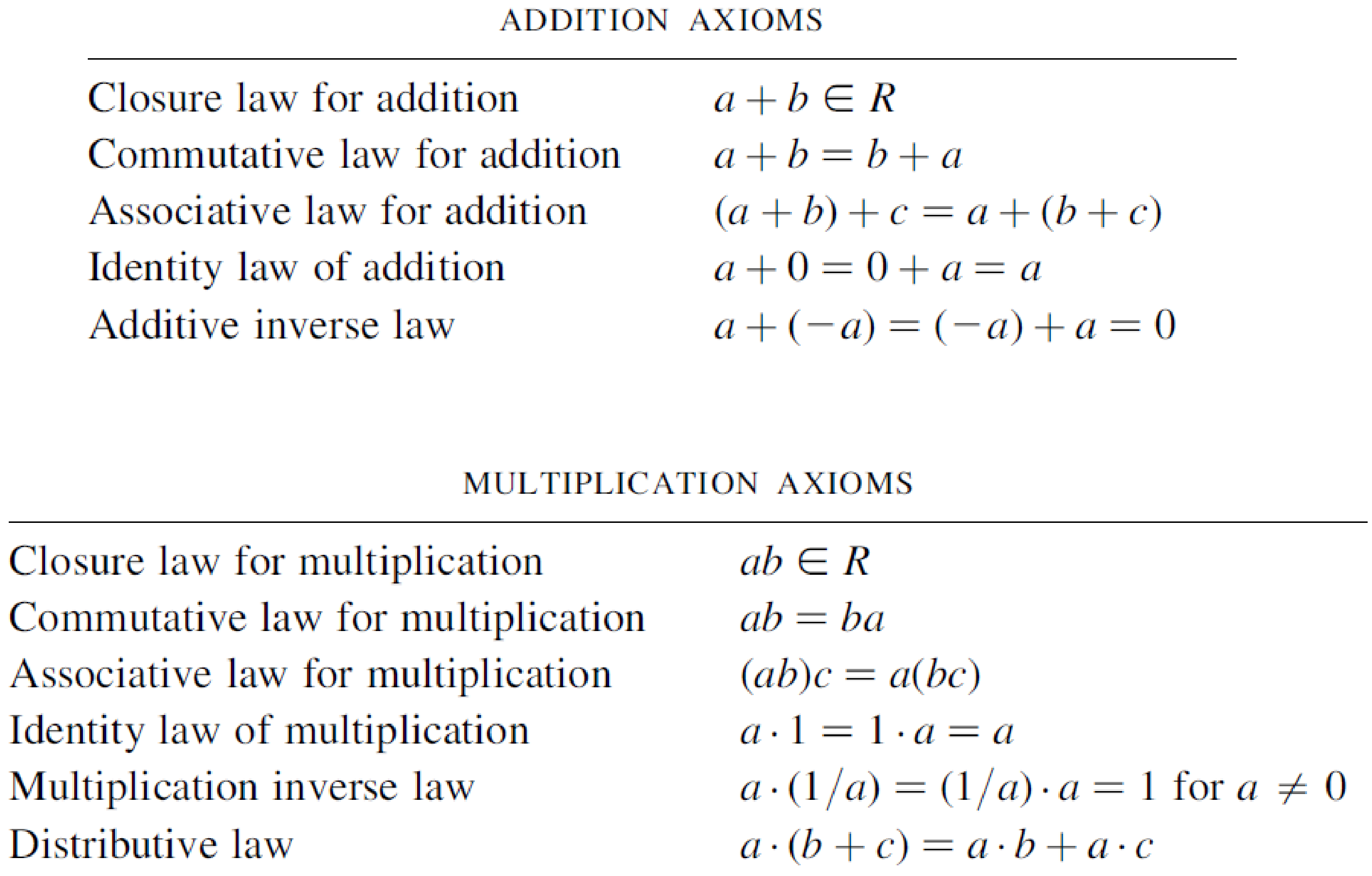
Reflexive law	$a = a$
Symmetric law	If $a = b$, then $b = a$
Transitive law	If $a = b$ and $b = c$, then $a = c$
Substitution law	If $a = b$, then a may be substituted for b or b for a in any expression

Ordering relations obey the following axioms:

ORDER AXIOMS

Trichotomy law	Exactly one of the following is true: $a < b$, $a = b$, or $a > b$
Transitive law	If $a < b$ and $b < c$, then $a < c$
Closure for positive numbers	If $a,b > 0$, then $a + b > 0$ and $ab > 0$

Axioms for addition and multiplication operations are summarized as follows:

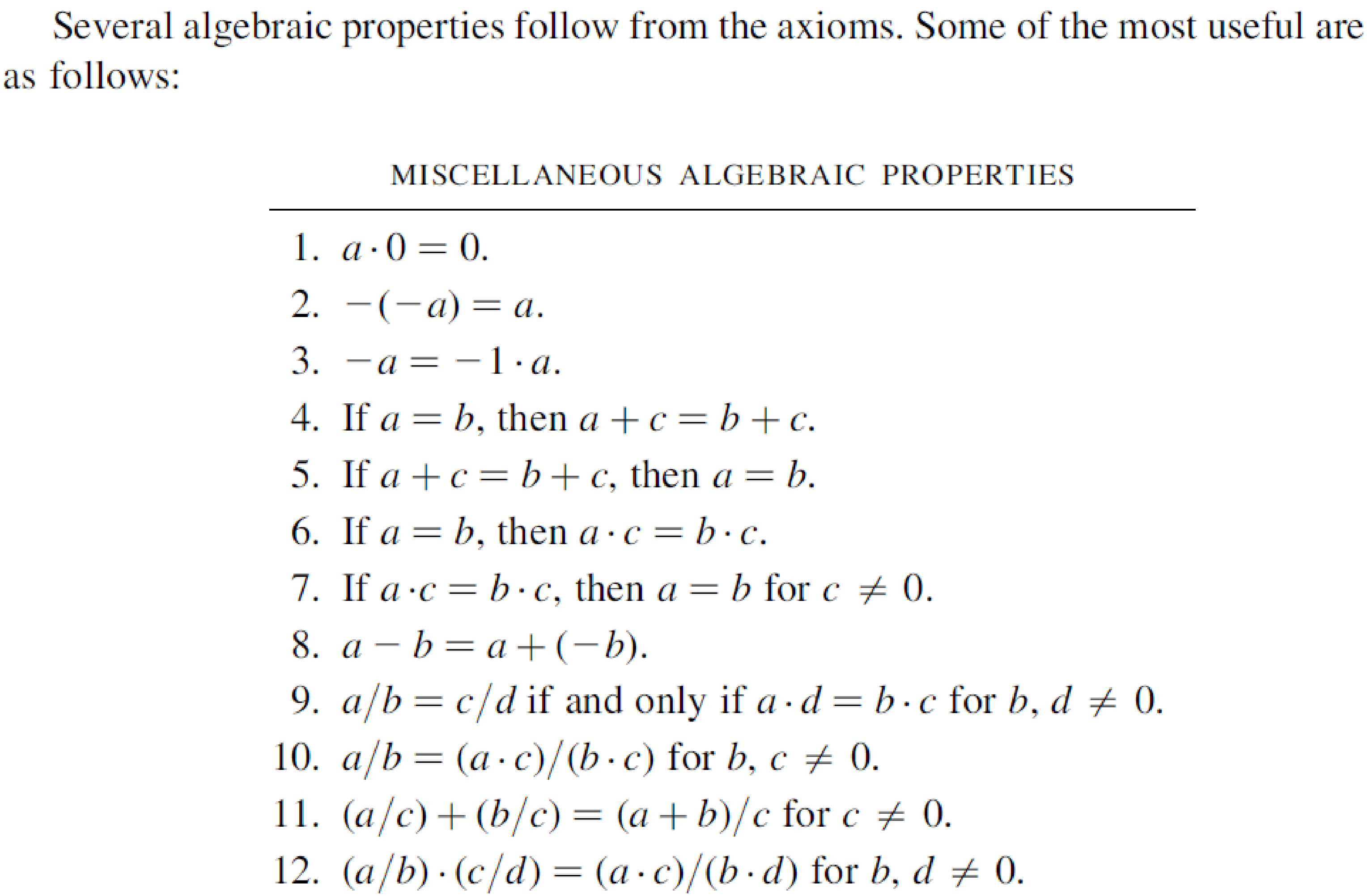


ADDITION AXIOMS

Closure law for addition	$a + b \in R$
Commutative law for addition	$a + b = b + a$
Associative law for addition	$(a + b) + c = a + (b + c)$
Identity law of addition	$a + 0 = 0 + a = a$
Additive inverse law	$a + (-a) = (-a) + a = 0$

MULTIPLICATION AXIOMS

Closure law for multiplication	$ab \in R$
Commutative law for multiplication	$ab = ba$
Associative law for multiplication	$(ab)c = a(bc)$
Identity law of multiplication	$a \cdot 1 = 1 \cdot a = a$
Multiplication inverse law	$a \cdot (1/a) = (1/a) \cdot a = 1$ for $a \neq 0$
Distributive law	$a \cdot (b + c) = a \cdot b + a \cdot c$

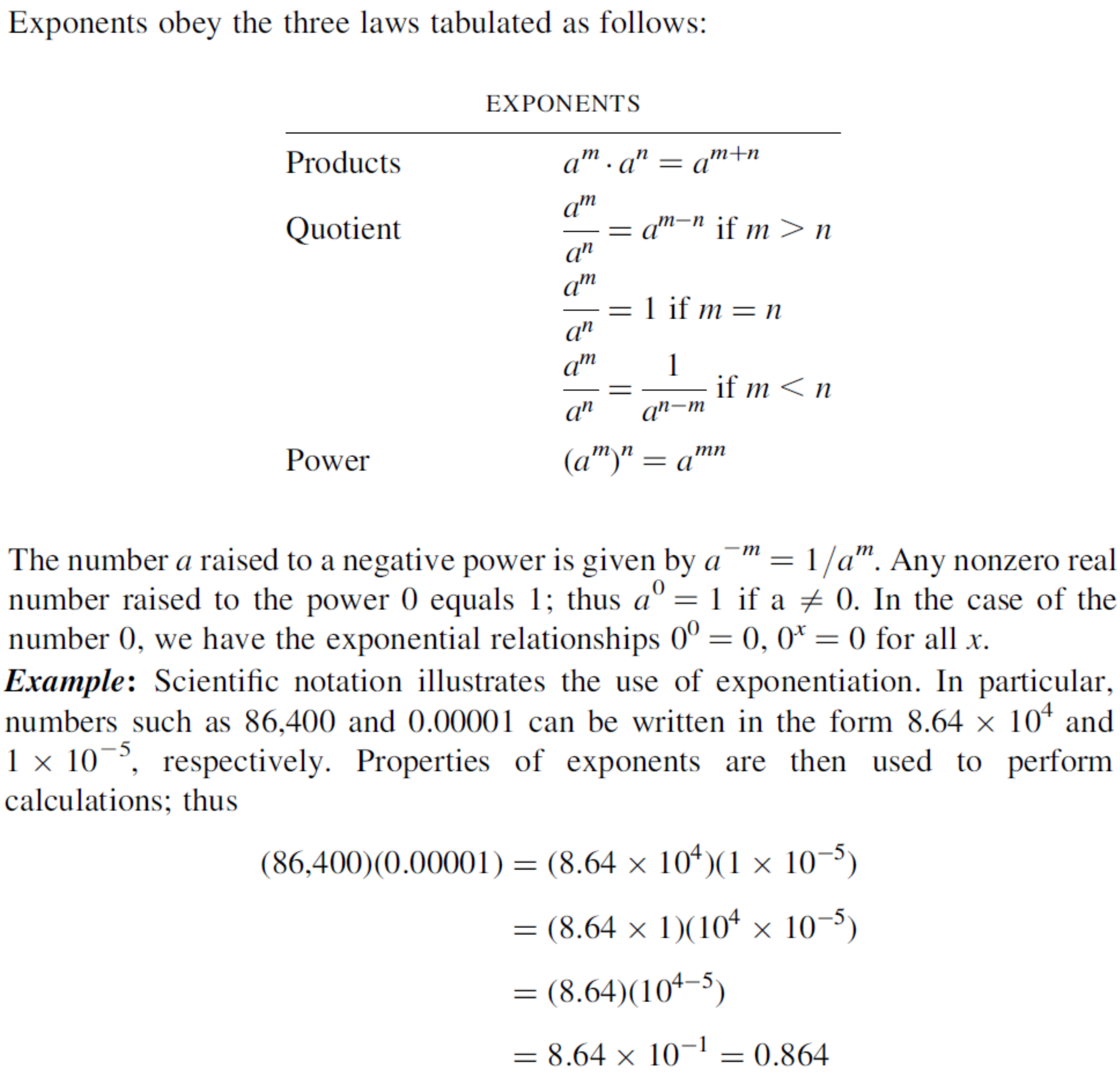


Several algebraic properties follow from the axioms. Some of the most useful are as follows:

MISCELLANEOUS ALGEBRAIC PROPERTIES

1. $a \cdot 0 = 0$.
2. $-(-a) = a$.
3. $-a = -1 \cdot a$.
4. If $a = b$, then $a + c = b + c$.
5. If $a + c = b + c$, then $a = b$.
6. If $a = b$, then $a \cdot c = b \cdot c$.
7. If $a \cdot c = b \cdot c$, then $a = b$ for $c \neq 0$.
8. $a - b = a + (-b)$.
9. $a/b = c/d$ if and only if $a \cdot d = b \cdot c$ for $b, d \neq 0$.
10. $a/b = (a \cdot c)/(b \cdot c)$ for $b, c \neq 0$.
11. $(a/c) + (b/c) = (a + b)/c$ for $c \neq 0$.
12. $(a/b) \cdot (c/d) = (a \cdot c)/(b \cdot d)$ for $b, d \neq 0$.

Exponents



Exponents obey the three laws tabulated as follows:

EXPONENTS

Products	$a^m \cdot a^n = a^{m+n}$
Quotient	$\dfrac{a^m}{a^n} = a^{m-n}$ if $m > n$
	$\dfrac{a^m}{a^n} = 1$ if $m = n$
	$\dfrac{a^m}{a^n} = \dfrac{1}{a^{n-m}}$ if $m < n$
Power	$(a^m)^n = a^{mn}$

The number a raised to a negative power is given by $a^{-m} = 1/a^m$. Any nonzero real number raised to the power 0 equals 1; thus $a^0 = 1$ if $\mathrm{a} \neq 0$. In the case of the number 0, we have the exponential relationships $0^0 = 0$, $0^x = 0$ for all x.

***Example*:** Scientific notation illustrates the use of exponentiation. In particular, numbers such as 86,400 and 0.00001 can be written in the form 8.64×10^4 and 1×10^{-5}, respectively. Properties of exponents are then used to perform calculations; thus

$$\begin{aligned}(86{,}400)(0.00001) &= (8.64 \times 10^4)(1 \times 10^{-5}) \\ &= (8.64 \times 1)(10^4 \times 10^{-5}) \\ &= (8.64)(10^{4-5}) \\ &= 8.64 \times 10^{-1} = 0.864\end{aligned}$$

Roots

The solution of the equation

$$b^n = a$$

may be written formally as the n^{th} root of a equals b, or

$$\sqrt[n]{a} = b$$

A real number raised to a fraction p/q may be written as the q^{th} root of a^p, or

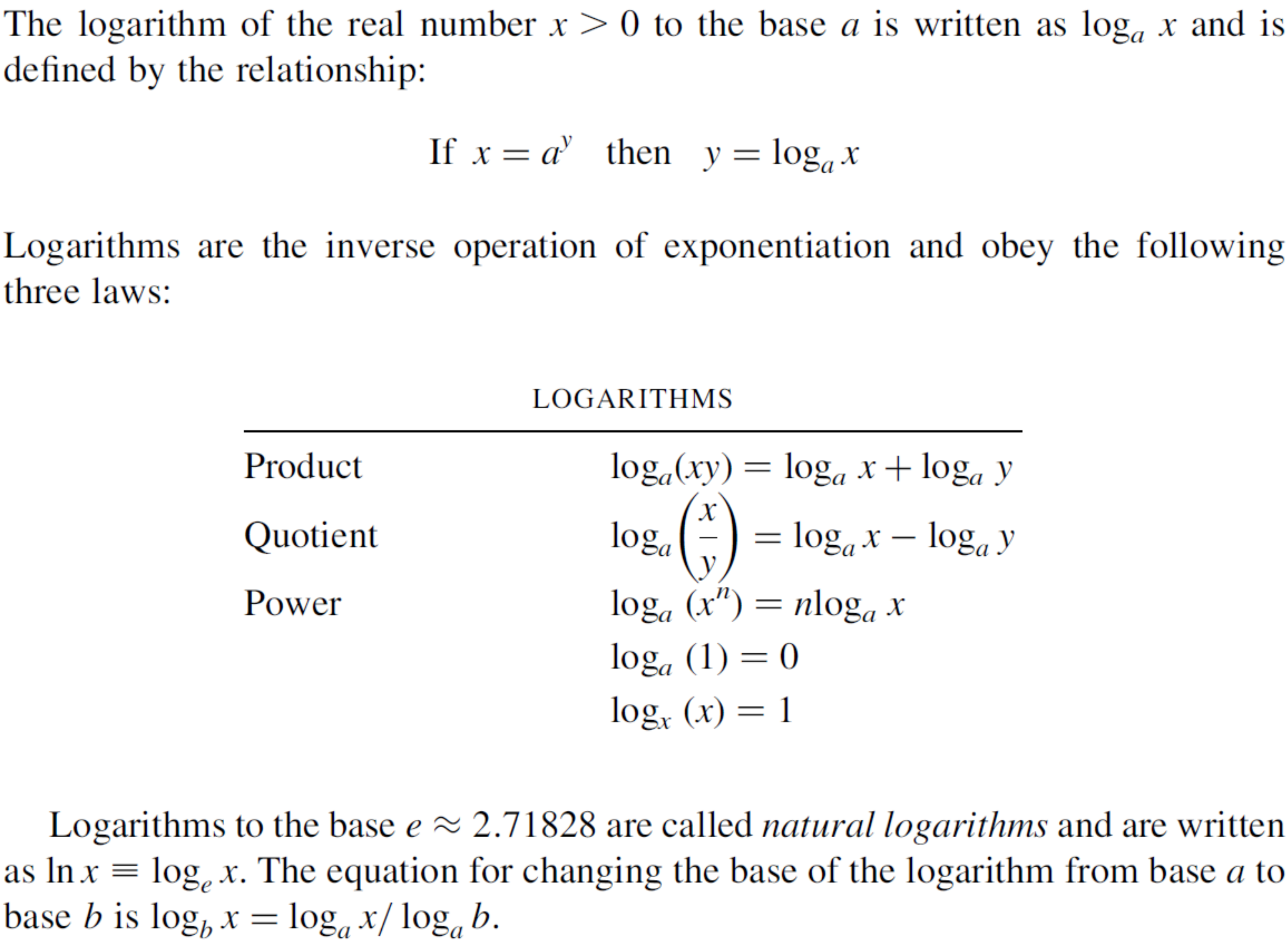
$$a^{p/q} = \sqrt[q]{a^p}$$

A summary of relations for roots is as follows:

ROOTS

	$a^{1/x} = \sqrt[x]{a}$
Product	$\sqrt[x]{ab} = \sqrt[x]{a}\sqrt[x]{b} = a^{1/x}b^{1/x}$
Quotient	$\sqrt[x]{\frac{a}{b}} = \frac{\sqrt[x]{a}}{\sqrt[x]{b}} = \frac{a^{1/x}}{b^{1/x}}$
Power	$(a^{1/y})^x = a^{x/y} = \sqrt[y]{a^x}$
	$\sqrt[x]{\sqrt[y]{a}} = (a^{1/y})^{1/x} = a^{1/xy}$

LOGARITHMS



The logarithm of the real number $x > 0$ to the base a is written as $\log_a x$ and is defined by the relationship:

$$\text{If } x = a^y \quad \text{then} \quad y = \log_a x$$

Logarithms are the inverse operation of exponentiation and obey the following three laws:

LOGARITHMS

Product	$\log_a(xy) = \log_a x + \log_a y$
Quotient	$\log_a\left(\frac{x}{y}\right) = \log_a x - \log_a y$
Power	$\log_a (x^n) = n\log_a x$
	$\log_a (1) = 0$
	$\log_x (x) = 1$

Logarithms to the base $e \approx 2.71828$ are called *natural logarithms* and are written as $\ln x \equiv \log_e x$. The equation for changing the base of the logarithm from base a to base b is $\log_b x = \log_a x / \log_a b$.

Your Task

Your task Read the Guide to Big Data and the Australian Privacy Principles and discuss 'risk points' that may have potential legal implications to organisations and ways to overcome legal issues that may arise.

References

1.Data to Decisions CRC, 2017, Law and policy. Retrieved 18 July 2019
2.IBM Big Data Analytics Hub 2014, Infographics Animations The Four V's of Big Data. Retrieved 13 March 2018.
3.Business Analytics | 4th Edition. Jeffrey D. Camm/James J. Cochran/Michael. Chapter 1.

Further Reading:

DATA SCIENCE CODE OF PROFESSIONAL CONDUCT
Read Data Science Australia's Data Science Code of Professional Conduct,
http://www.datascienceassn.org/code-of-conduct.html