COMM1190

Data, Insights and Decisions

Comprehensive Course Notes

UNSW Business School

Covering all Weeks, Lectures, Workshops, Readings and Case Studies

Table of Contents

Week 1: Introduction to Data Analytics

1.1 What Is COMM1190?

COMM1190 Data, Insights and Decisions is a foundational course in business analytics at UNSW. The course builds practical skills in using data to understand business problems, make decisions, and communicate findings. It integrates quantitative methods, statistical thinking, ethical reasoning, and data visualisation through the R programming environment.

The course spans three broad modules: descriptive and exploratory analytics, predictive analytics (regression and classification), and the broader context of data in organisations including ethics and big data. Students are expected to work with real datasets in R workshops throughout the term.

1.2 Core Reading: Banerjee, Bandyopadhyay and Acharya (2013)

Banerjee, A., Bandyopadhyay, T. and Acharya, P. (2013) 'Data Analytics: Hyped Up Aspirations or True Potential?' Vikalpa, 38(4), pp. 1-11.

1.2.1 Defining Analytics

Analytics is defined in practice as any fact-based deliberation leading to insights (diagnostics) and possible implications for planning future action in an organisational setting. It can range from routine tracking of business performance to the diagnosis of root causes of business problems or strategic prediction about future business initiatives.

The common thread across all analytical exercises is that they are fact-driven, based on data gathered as part of business and market information collection.

Business Analytics (Beller and Barnett, 2009)
A set of all the skills, technologies, applications and practices required for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning.

1.2.2 The Four Types of Analytics (Gartner Report, 2012)

Type	Question Answered	Nature	Complexity
Descriptive Analytics	What happened?	Retrospective; monitors and tracks performance over time using dashboards and reporting tools	Low -- foundational; used routinely
Diagnostic Analytics	Why did it happen?	Investigative; uses exploratory data analysis, root cause identification, often engages consultants	Moderate -- requires research design

Predictive Analytics	What is likely to happen?	Forward-looking; uses statistical and data mining techniques to forecast outcomes and identify drivers	High -- requires modelling skills
Prescriptive Analytics	What should be done about it?	Action-oriented; associates decision alternatives with outcome predictions; uses optimisation and simulation	Very high -- rare in practice; 'what if' simulators

Important exam point: The analytics types form a spectrum from low complexity/high frequency (descriptive) to high complexity/low frequency (prescriptive). Each type is dependent on the one before it. Prescriptive analytics is widely regarded as underdeveloped in practice due to database limitations.

1.2.3 Organisational Use Cases of Analytics

The authors identify three key categories of analytics in practice:

- Dashboard applications: Routinely generated metrics to monitor processes across time, including decision support systems and financial health reporting. These fall under descriptive analytics and require the skill of reading relevant facts from figures and connecting them to decision-making.
- Investigative applications: Not routine; exploratory or confirmatory in nature. Typically engaged through consultants. Requires framing research questions, collecting data, and connecting findings to business objectives. Often diagnostic or predictive in nature.
- Prescriptive or scenario-builder applications: Statistical market response models generating 'what if' simulators. These are often partial due to database limitations and must be used cautiously.
- Big Data Analytics: Described as a buzz word in the article (2013). Processing of large volumes of unstructured data. Primarily exploratory in nature. The challenge is filtering signal from noise.

1.2.4 The Analytics Industry

The analytics industry had grown significantly by 2012 -- approximately 14 percent in that year alone -- and was forecast to reach USD 50.7 billion by 2016 (IDC Report, 2011). The business intelligence and analytics market was identified as the fastest-growing segment in software markets.

India emerged as a major hub for analytics outsourcing, holding the largest share of the global outsourcing market (KPO sector). Indian analytics firms delivered USD 375 million out of a total global data analytics outsourcing market of USD 500 million in 2012. Projected growth to 21 percent of the total KPO market by 2015.

Key analytics-consuming sectors: BFSI (Banking, Financial Services and Insurance), Telecom, ITES, FMCG, and Retail.

1.2.5 Required Competencies in Analytics

The authors identify four core competency areas for analytics practitioners:

- Technical skills: Programming, data processing, statistical modelling.

- Business acumen: Understanding industry context, framing relevant business questions, translating findings into strategy.
- Communication and translation: Conveying insights meaningfully to decision-makers, often non-technical stakeholders.
- Domain knowledge: Sector-specific expertise to contextualise findings correctly.

The authors note tension between technical capability and business acumen, particularly in offshore environments where technology skills dominate but business context is limited. The jury remains out on whether specialisation versus integration of these competencies is more effective.

1.2.6 Counter-Views and Limitations of Analytics

The authors acknowledge significant criticisms and limitations of analytics as a discipline:

- Greenfield limitation: Analytics largely uses past data and cannot provide reliable insights for entirely new initiatives or markets where no historical data exists.
- Black swan events: Mainstream analytical tools are poorly suited to predicting high-profile, rare, hard-to-predict events (Taleb, 2010). These outlier risks carry the largest threats.
- Creativity risk: Over-reliance on analytics may suppress heuristic judgment and out-of-the-box thinking that sometimes yields better outcomes.
- 'Snob value' adoption: Many organisations adopt analytics due to social or competitive pressure rather than genuine strategic need; Bloomberg Survey (2011) found analytics was still not used in major parts of decision-making.
- Steep learning curve: Implementation requires significant time, investment, and cultural change. If not smooth, it frustrates policy-makers and undermines confidence.
- Employee threat perceptions: Frontline staff may feel like data entry operators; mid-level managers may see analytics as a threat to their power as instinctive decision-makers.
- Ethical misuse: Precision marketing, tracking of web behaviour, and data monetisation can infringe on consumer privacy. The Tom Tom GPRS speed trap case (Netherlands) is cited as an example where analytics enabled ethically questionable government action.
- Data security vulnerabilities: Server crashes, data breaches, outsourcing risks, and disruptive technology changes can render analytics infrastructures deficient.

1.2.7 Big Data: The New Frontier

The article distinguishes big data as a specific and emerging form of analytics. Key characteristics: volume (quintillion bytes of data generated daily by 2012), variety (texts, photographs, videos, tweets, blogs), velocity (real-time data generation), and the shift from structured corporate data to unstructured social and digital data.

The challenge is to extract meaningful patterns or correlations from unstructured data. The risk is that the noise is increasing faster than the signal -- finding a pattern may be like searching for a needle in a haystack.

The authors are cautiously optimistic: social media data may provide early signals of market behaviour if properly linked to actual consumer behaviour. However, the analytical prowess required to link disparate data sources, glean information from new formats, and integrate that into traditional models remains a significant challenge.

> Critical assessment note: The article argues that both extreme scepticism and excessive euphoria about analytics are inappropriate. Analytics is best understood as 'pervasive ether' in the organisation -- supporting and guiding decision-making without overriding common sense acumen.

Week 2: Data Visualisation in R

2.1 Overview of R and RStudio

R is an open-source programming language and environment for statistical computing and graphics. COMM1190 uses R as its primary analytical tool throughout workshops. RStudio is the standard IDE (Integrated Development Environment) used alongside R.

Key features of R include: a broad library of packages for statistical analysis, data manipulation and visualisation; active community development; and tight integration between data processing and graphical output.

2.2 Loading and Exploring Data in R

The fundamental workflow in R begins with reading data into the environment, then exploring its structure before proceeding to visualisation or modelling.

Core R Functions for Data Loading and Exploration
read.csv() -- reads a comma-separated file into a data framenrow() -- returns number of rowsncol() -- returns number of columnssubset() -- extracts rows satisfying a conditioncolSums() -- calculates column-wise sums for numeric datahead() / tail() -- previews first or last rows of a data frame

2.2.1 Workshop Dataset: Weekly Ice Cream Sales

The workshop dataset tracks weekly ice cream sales by customer type (Staff vs Student) and by flavour group across 26 weeks (April to September). The dataset contains 52 rows (26 weeks x 2 customer types) and 19 columns (week, month, customer type, and 15 flavour columns).

Flavour groups were constructed for analysis:

- Fruit: Apricot, Banana, Cherry Almond, Ginger, Lime Coconut, Mango, Pure Coconut, Red Bean
- Caramel: Salted Caramel
- Chocolate: Chocolate, Mint Choco
- Tea/Coffee: Chai Tea, Green Tea, Coffee
- Nut: Hazelnut, Pistachio

2.2.2 Key R Code for Subsetting and Aggregating

Extracting data for a specific customer group:

staff <- subset(data, Customers == 'Staff')This returns only rows where the Customers column equals 'Staff'. The resulting data frame contains 26 rows (one per week).

Aggregating sales by flavour group:

staff_fruit <- subset(staff, select = c(Apricot, Banana, Cherry.Almond, Ginger, Lime.Coconut, Mango, Pure.Coconut, Red.Bean))staff_fruit_sum <- sum(staff_fruit)This selects only the fruit flavour columns for staff, then sums all values across all weeks and flavours.