**CONTEMPORARY BUSINESS ANALYTICS**

**Joseph F. Hair, Jr. | Ermanno Affuso | Joshua Lambert| Michail Tsagris | Bradford Tuckfield**

**Executive Abstract**

The field of Business Analytics (BA) has experienced rapid growth in recent years, underscored by the proliferation of academic programs dedicated to the BA discipline. There is a noticeable lag, however, in innovation within the Modern Business Analytics field when compared to the area often identified as its parent discipline, Data Science, which has continually incorporated cutting-edge business AI applications. Existing BA textbooks focused on business administration courses predominantly adhere to the traditional three pillars of Business Analytics: descriptive, predictive, and prescriptive, thereby serving as mere extensions of conventional business statistics textbooks.

Our objective, led by the Hair et al. team, is to redefine the contours of Business Analytics in a transdisciplinary manner consistent with the evolving demands of contemporary e-businesses. Our primary aim, therefore, is to write a comprehensive, accessible textbook encompassing novel business applications of data science—a critical gap currently unaddressed by competitors in the Business Analytics arena. Notably, our team of authors comprises business professors with both academic and industry credentials, underpinned by extensive research, teaching, and textbook publishing experiences in the realm of business.

The textbook we are proposing will feature user-friendly companion software in both R and Python, emphasizing their utility as potent pedagogical resources. This approach will cater to graduate students enrolled in Masters Programs emphasizing Business Analytics as well as practitioners operating within the field. This proposal outlines our rationale, objectives, textbook structure, and competitive advantages as we embark on this mission to redefine and revitalize the discipline of Business Analytics. We note also that our team would also consider a tandem project at a lower level specifically targeted at the undergraduate level.

**Background**

Business Analytics is an interdisciplinary field within business that includes behavioral science, communication, econometrics, statistics, scientific research methods, computational science, information systems, Artificial Intelligence, and various algorithms. Its purpose is to distill valuable insights from structured and unstructured business data, facilitating the analysis, understanding, interpretation, and prediction of business processes and performance. The historical roots of this field originated in the mid-1960s when Decision Support Systems (DSS) began to emerge, providing the initial framework for leveraging data in managerial decision-making. The evolution of DSS along with the explosion of data, primarily digital, laid the essential groundwork for the emergence of Business Analytics.

In 2018, the AACSB's program database indicated over 400 analytics programs were offered by 220 business schools worldwide, a number that has increased substantially since then. For instance, the number of Ph.D. programs in Business Analytics experienced a remarkable 350% growth within the past three years. At the same time, the increased interest in the Master of Science degree in Business Analytics has been exponential as well, noticeably outpacing the demand for traditional Master of Business Administration (MBA) programs, as indicated by our analysis of Google Trends data, which serves as a proxy for consumer interest (student & employer) in these disciplines.

|  |
| --- |
| MSBA’s consumer interest overtime (2004 – 2023) |
| MBA’s consumer interest overtime (2004 – 2023) |

Given the current digital age, we anticipate a continued expansion in consumer interest in Business Analytics, with a likelihood that MBA programs will incorporate a substantial business analytics component, and likely also a minor field.

Some contemporary statisticians, including notable figures like Nate Silver, contend that Business Analytics is essentially a rebranding of statistics, albeit with extensions such as data visualization for effective communication. In the modern context, Business Analytics heavily relies on reporting and dashboarding tools to provide actionable insights to business stakeholders, primarily anchored in three core analytical pillars: descriptive, predictive, and prescriptive analytics. We assert, however, that Business Analytics extends well beyond these boundaries, especially considering the substantial contributions from the parent discipline of Data Science. This extension includes the integration of tools like generative AI, Natural Language Processing, and Recommender Engines, which have become indispensable for modern businesses, particularly those reliant on e-commerce, where substantial returns on investment from AI applications are commonplace. For example, a decade ago MacKenzie et al. (2013) estimated 35% of Amazon purchases stemmed from recommendations. With Amazon's total revenue in 2019 having grown to US $280.5 billion and considering the proportionate revenue share attributed to the Amazon recommender systems from previous data, the AI tool contributed approximately US $98 billion in added value during that year, an amount clearly substantially larger today.

Our aim is to contemporize the field of Business Analytics through the creation of a comprehensive textbook that reconceptualizes the discipline as a transdisciplinary business field, encompassing crucial data science business applications often overlooked in current Business Analytics textbooks.

**Textbook Rationale**

In 2020, Professors Hair and Affuso established the Business Analytics concentration within the existing Ph.D. program in Business at the University of South Alabama. At the time, there were fewer than ten similar programs globally, and it has since grown rapidly (350% growth in the past three years). However, the dearth of teaching materials that cover contemporary topics has led to frustration among instructors and students alike. Instructors often resort to additional readings from journal articles authored by engineers and data scientists, which can prove excessively technical for business students. Furthermore, existing textbooks tend to employ applications in Excel or copious algorithms in R, requiring students to transcribe multiple lines of code from the textbook to the R console, leading to errors that necessitate time-consuming code verification during lab sessions.

In Table 1, we present a preliminary analysis of contemporary topics in Business Analytics as covered by major competitors in the market.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Textbook** | **Authors** | **Publisher** | **Year** | **Data Manipulation/ Mining** | **Data Visualization** | **Statistical Inference** | **Supervised Learning** | **Unsupervised Learning** | **Deep Learning** | **Natural Language Processing** | **Recommender Systems** |
| Business Analytics, 5th Edition | Camm et al. | Cengage | 2023 | × | × | × | — | — | — | — | — |
| Introduction to Business Analytics, 1st Edition | Richardson and Watson | Cengage | 2023 | × | × | × | — | — | — | — | — |
| Business Analytics: Data Analysis & Decision Making, 7th Edition | Albright and Winston | Cengage | 2019 | × | — | — | — | — | — | — | — |
| Business Analytics, 2nd Edition | Jaggia et al. | McGraw-Hill | 2022 | × | × | × | ×/limited | × | — | — | — |
| Introduction to Management Science and Business Analytics: A Modeling and Case Studies Approach with Spreadsheets, 7th Edition | Hillier and Hillier | McGraw-Hill | 2022 | — | — | — | — | — | — | — | — |
| Modern Business Analytics, 1st Edition | Taddy et al. | McGraw-Hill | 2022 | — | — | × | × | × | × | only Textual Analysis | — |
| Business Analytics, 3rd edition | Evans | Pearson | 2020 | × | × | × | — | — | — | — | — |
| Business Analytics, Solving Business Problems with R | Mishra and Mishra | Sage | 2024 | — | — | — | × | × | × | only Textual Analysis | — |
| Business Analytics | Gordon | Sage | 2023 | × | — | — | — | — | — | — | — |
| Business Analytics, Applied Modelling and Prediction | Abdey | Sage | 2023 | × | × | × | × | — | — | — | — |
| Data Mining and Business Analytics with R | Ledolter | Wiley | 2013 | × | — | × | × | × | — | × | — |
| Business Analytics and Statistics, 1st Edition | Black et al. | Wiley | 2021 | × | × | × | — | — | — | — | — |
| **Contemporary Business Analytics** | **Hair et al.** | **NA** | **2025** | **×** | **×** | **×** | **×** | **×** | **×** | **×** | **×** |

**Table 1. Competing Textbooks: Content Analysis and Comparison.**

**Objectives**

Our primary objective is to prepare a textbook that serves as a powerful pedagogical tool, rooted in real-world case studies and practical applications, rendering the complex discipline accessible to a wide audience, ranging from industry practitioners to graduate students enrolled in Masters and Doctoral Programs. We also envision a condensed, accessible version of the textbook for the undergraduate market.

We aim to bridge the gap between Data Science and Modern Business Analytics by incorporating essential topics currently absent in competing textbooks. Our intention is to provide theoretical foundations for Data Science models applied in the business context, approaching them from a business scholar's perspective rather than that of an engineer or mathematician. To cater to the pedagogical requirements of doctoral students in Business Analytics and Information Systems, we will include concise end-of-chapter or online appendices.

Additionally, one of added values of our proposed new textbook will be the development of companion libraries in both R and Python, making it "bilingual" in terms of computational capability to cater to a broader audience familiar with either syntax. These libraries will articulate numerical results from models in plain English, similar to a scientific article or technical report, enhancing understanding and minimizing reliance on technical statistical software printouts.

**Competitive Advantages**

* The textbook will provide up-to-date information essential for both traditional and digital firms, addressing gaps in existing competing textbooks.
* Open-source companion statistical libraries in R and Python will accommodate users familiar with either syntax – which will significantly reduce the need for extensive coding by students and professors.
* Availability of these companion libraries in the R and Python repositories incentivizes professors to adopt and practitioners to purchase the textbook for a deeper understanding of their applications.
* The lead author, Professor Hair, boasts a remarkable citation count of 388,000+, nearly three times that of the second-ranking business analytics scholar and surpassing the combined citations of scholars ranked #2 to #5. This high visibility and recognition of the Hair et al. brand are expected to drive interest and boost book sales. Moreover, the lead author has published other books focusing on analytical methods, several of which are market leaders.
* The co-authors possess doctoral training in Business and Behavioral Science, coupled with interdisciplinary backgrounds in Computer Science, Engineering, and Statistics, confirming their active scholarship in Business Analytics and Data Science.

Top of Form

**Pedagogical Approach**

*General Overview*

The business disciplines need a new textbook that provides a solid foundation in business analytics techniques and helps students learn how to interpret and communicate analytics results to maximize their strategic value for business. A core feature of our book will be a continuing case study in the relevant chapters. The continuing case will be based on a data set associated with access to and viewing of movies and entertainment, similar to a Netflix type of services provider. The dataset will be adaptable to different types of analytical tools, such as neural networking (NN), natural language processing (NLP), cognitive analytics (CA), machine learning (ML), cluster analysis, multiple regression (both regular & logistic), structural equation modeling (SEM), network analysis, and recommender systems.

The approach of this book is based on the authors’ many years of conducting and teaching business research and statistical analysis method, as well as the desire to communicate the fundamentals of new and emerging analytical methods to a broad audience. To accomplish this goal, we have limited the emphasis on equations, formulas, Greek symbols, and so forth that are typical of most books and articles. Instead, we explain in detail the basic fundamentals of business analytics following a direct and simple approach. We will place the glossary of terms at the front of each chapter to make it convenient and quick for readers to clarify the meaning of technical concepts introduced in each chapter. The initial section of each chapter, following the glossary, will be an introduction and overview of the chapter topics. Where appropriate, we will provide Rules of Thumb for analytics methods organized in table format, that can be used as general guidelines for understanding, interpreting, and evaluating the results from applying each of the methods, and will lead to citations for the book in journal articles. Finally, in the final chapter we will provide an overview of the elements of relevant software packages including SQL, the R Studio and Python, which can be used not only with exercises in this book, but also in the reader’s own research.

The proposed textbook will introduce students to topics and techniques associated with exploring large datasets consisting of both primary and secondary data, structured and unstructured data, CRM, data visualization, web analytics, and digital data, including social media, sensors, images, and so forth. In every chapter, we will have a practitioner’s spotlight highlighting a case study based on a large dataset to address a strategic business initiative.

An innovative supplement to the textbook will be a constantly updated YouTube channel. The YouTube channel will be updated regularly to expand on the material presented in the textbook. We anticipate developing an authors’ website for interaction with instructors using our textbook and supplements.

We will also work with McGraw-Hill to develop and support McGraw-Hill Connect®, LearnSmart®, and other electronic channels of delivery for the text. Additional Instructor Resources for the book will include Powerpoint slides, Instructor’s manual, and similar traditional materials. We propose to be involved in the preparation of these supplementary materials for the text, but initial versions of the instructor’s manual, test bank and Powerpoint slides will be prepared by individuals contracted directly to others by McGraw-Hill, for our review and approval. Our vision is a *Contemporary Business Analytics* text we anticipate will be about 400 - 500 pages long. We can adjust the pedagogy for either a graduate or undergraduate market, but realistically could easily write two versions of the book – one for the undergraduate market and the other for graduate classes. For chapter-by-chapter details, see the proposed Table of Contents.

*Further pedagogical details*

In the era of Big Data and the emerging field of business analytics, students increasingly must learn be prepared to make decisions in their jobs based on using almost unlimited information. To identify and organize this vast amount of data, they must become familiar with the numerous emerging business analytics methods. In short, they must ensure they are familiar with and master the required data analytics skills to obtain their next job. *Contemporary Business Analytics* will explain a portfolio of business analytics techniques for individuals who are not data scientists but need to be sufficiently familiar with the topic to manage and execute many of the methods discussed in the book. The authors provide a straightforward, hands-on approach to the essential managerial process of analyzing data to make relevant and timely business decisions. The book includes critical topics, such as the increasing role of analytical methods in managerial decision making; technologies impacting the field such as artificial intelligence (AI); machine learning (ML); cognitive analytics (CA); recommender systems; customer relationship management; ethical issues; privacy matters; and more.

Our proposed text will be the only business analytics book with a chapter on recommender systems, and the coverage of unstructured data is much more extensive as well as easier to understand than in other texts. In addition, a primary feature of the book is a realistic continuing case throughout all chapters enabling students to understand and apply business analytics and methodologies applied in the real world. Numerous research examples are included in all chapters, as well as mini cases and exercises.

Business analytics methods convert information into knowledge. Better business knowledge is essential to improved decision making. The book minimizes the emphasis on statistical theory and maximizes coverage of basic as well as selected advanced analytical tools and skills representing a broad range of potential business analytics applications. Applying the concepts and principles presented in this book will enable students to more effectively cope with the fast-paced decision-making environment of business today as well as tomorrow.

An innovative feature referred to as Business Analytics in Action (BAIA) (or Business Analytics Insights (BAI) provides mini cases throughout the chapters on timely and relevant analytics applications. Business Analytics Dashboards emphasizing applied examples of actual research problems and current issues are included in all chapters. Some Business Analytics Dashboard examples summarize actual research studies. Others describe websites that help researchers analyze unstructured data or locate sources of digital and secondary data. Case studies involving applications of research approaches are also included as well as supplements with instructions on how to use statistical software to analyze data.

An example of a very relevant topic not covered in other BA texts but included in our book is sampling. Virtually all unstructured digital and secondary data as well as selected structured datasets explored with business analytics applications rely on datasets consisting of literally millions of data points. None of the competitive texts mention the need for sampling with these data sets, but our book will address it as an example in one or more of the BAIA mini cases. Why is sampling important with large datasets? Because statistical assessment of big data samples is sensitive to sample size – and the result is all relationships examined will be statistically significant. As a result, analysis of most all large datasets should include sampling to determine if relationships are statistically meaningful or simply the result of the millions of data points included in the analysis.

With the Business Analytics Dashboards and Business Analytics in Action examples the text material is truly brought to life! In addition to these pedagogical enhancements, we include online applications and questions at the end of each chapter that provide interactive exercises for students, as well as discussion questions posing analytical issues going beyond just repeating topics covered in the chapters.

Our pedagogy has been developed over many years of conducting and teaching multivariate data analysis, the precursor to business analytics. To bring the concepts to life and make the text more interesting, we focus on a single case throughout the book. The topic of the continuing case is an online movie selection business similar to Netflix.

A notable aspect of *Contemporary Business Analytics* is the inclusion of tailored statistical packages in R and Python. These packages are accompanied by extensive datasets, comprising both real and artificial data, enabling the replication of the business case studies presented in the textbook. The dataset sizes vary depending on the specific case study. To illustrate, for the chapter on recommender systems we utilize the publicly available *MovieLens* dataset. This dataset contains a vast compilation of data, including over 20 million ratings and 465,564 tag applications across 27,278 movies. The data is collected from the period between January 9, 1995, and October 17, 2016. The users included in the dataset were selected randomly and each user had rated at least 20 movies.

Exercises in the BAIA mini cases as well as at the end of selected chapters provide an opportunity for students to further examine the Continuing Case data which ranges from 200 to 20 million observations.

The significant variation in dataset sample sizes stems from the nature of the business analytical problems that need to be addressed. For instance, in the chapters dedicated to statistical inference, employing a large dataset similar to the entire population undermines the purpose of statistical inference itself, which involves drawing conclusions about the population from a data sample. To illustrate this, we will employ resampling techniques to extract small random samples from the extensive dataset. This will demonstrate to the reader that every conclusion we derive regarding the population includes sampling errors and biases.

Based on our extensive experience teaching inferential business statistics, we have observed that many students often fail to recognize they are estimating or forecasting "errors" associated with statistical parameter estimates. Our pedagogical objective in this context is to ensure students understand that the essence of inferential statistics lies in characterizing the sampling error.

Conversely, machine learning techniques, which often utilize the same statistical models employed in inferential statistics, have a distinct goal focused on feature classification, pattern recognition, and predictions. In the chapter dedicated to AI and machine learning topics, we will utilize exceptionally large datasets to ensure the machine learns how to generate precise predictions. Electronic copies of the variables and dataset will be available on our Web site.

The continuing case study makes it easy for readers to become familiar with the business issues associated with the case and to apply decision-making concepts relevant to each chapter. For example, we refer to the case when we discuss analytical method selection procedures as well as when evaluating different sampling approaches. In all the analytics methods chapters we use the case study data to illustrate relevant methods including Geographic Information Systems, Network Analysis, Natural Language Processing, Neural Networks, Machine Learning, Generative AI, and Recommender Systems, among others, and they can be enhanced with relevant emerging Artificial Intelligence systems. Focusing on a single case study throughout the book enables readers to more easily understand the benefits and pitfalls of undertaking and applying analytics approaches to improve business decision making.

The book’s coverage of business analytics methods is more extensive and much easier to understand than in other texts. Step-by-step instructions are included on how to apply statistical software to execute data analytics for all techniques. This enables instructors to spend much less time teaching students how to use the software. It also saves time later by providing a handy reference for students if they forget how to use the software. For instructors who want to expose their students to increasingly available analytical techniques, our book is the first to provide a comprehensive text on the most relevant topics.

The text and all major supplements are organized around the learning objectives listed at the beginning of each chapter. Instead of a single summary of each of the chapters, there will be summaries of major chapter topics associated with each learning objective. This organizational approach makes the book especially readable for students and readily useful for instructors. In short, it delivers value for students, professors, universities, and research organizations.

**Case Studies Design**

As mentioned in the Pedagogical Approach section, the global case study revolves around a hypothetical company, *Asterflix[[1]](#footnote-1)*, a leading company in the entertainment industry. In the introductory section, we will explain the business model of this hypothetical company, outlined in Table 2. We will also explore how the concepts of data science, discussed further in the textbook, can offer effective solutions to the specific business challenges encountered by different divisions within *Asterflix*. We anticipate virtually all readers will have used a movie app, such as Netflix. As a result, they can easily understand and relate to the concepts and variables in the continuing case study.

*Transition from Global to BAIA mini-case:*

Our chapter cases will be derived from real-life case studies, where we will fictionalize the challenges encountered by companies similar to *Asterflix* and propose data-driven analytical solutions. To achieve this, we will monitor job postings and relevant blogs on the Netflix website to identify the professional profiles sought by the company to tackle its business obstacles. For instance, referencing an article by Richard Wellington, Engineering Manager at Netflix, published on October 19, 2019[[2]](#footnote-2), the author highlights some of the challenges faced by Content Finance Engineers at Netflix:

“Here are some of the business problems we’re tackling in our space:

Programming a global catalog years ahead of the release date is hard. What is the right combination of data, intuition, and analytics to help create the best slate possible?

Spending both too much and too little money on a production is problematic. Either you went over budget, or you didn’t make the content as great as it could have been. What are the real-time systems we can have in place to ensure our productions are financially healthy. How can we best predict, manage, and coordinate budgets and spending?

Treating talent fairly and paying them on time and accurately is exceptionally important to us. Our systems need to be highly accurate and timely in order to continue having a great working relationship with our amazing talent.

Accounting and capitalizing our spend has to be highly accurate, since these are the numbers we report to Wall Street.” (Wellington, 2019).

Whenever possible, we will draw inspiration from these real cases and construct our BAIA mini-case for *Asterflix*. Let's consider a scenario where the Content Finance Engineering division at *Asterflix* aims to determine the ideal number of movie sequels the production team should invest in. Leveraging the authors’ academic expertise in the field of Business, in this case Economics of Entertainment, we will identify this number. To achieve this, we will utilize an extensive dataset encompassing information on movies, costs, revenues, and other determinants of supply and demand (such as movie genre, viewer preferences, talent rankings, etc.). This will enable us to concurrently estimate the Marginal Benefit of producing a sequel (based on the demand for that movie, which diminishes with the viewer's declining marginal utility) and the Marginal Cost of producing the sequel. The multivariate statistical method of PLS-SEM will facilitate the estimation of this model by calculating the viewer's utility (Satisfaction – an unobserved latent process) and determining the optimal number of sequels where the marginal benefit of production equals the marginal cost.

We will apply similar approaches to address various business challenges in other divisions of the company, including accounting, customer care, marketing, HR, business intelligence engineering, and more.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| BUSINESS MODEL | | | | | | | | |
| **Key Partners** | **Key Activities** | **Key Resources** | **Value Propositions** | **Customer Relationships** | **Channels** | **Customer Segments** | **Cost Structures** | **Revenue Streams** |
| Smart TV Manufacturers | Recruiting Software Engineers and Data Scientists | Artificial Intelligence Protocols | Advertisement-Free Contents | Exceptional Customer Service | Online Content Distribution | Film Enthusiasts | Procurement of Content Rights | Monthly subscription fees |
| Interactive Entertainment Firms | Application and Website Upgrades | Streamlined Automation Processes | High-Definition and Ultra HD Streaming | Real-time Chat Support | Android and iOS Apps | Web Series Fans | In-house Content Production | Revenue from original contents |
| Leading Television Companies | Creation of Exclusive Asterflix Content | Machine Learning Models | Cross-Platform Adaptability | Promotional Discounts | Gaming Console Integration and Streaming | Documentary Lovers | AI Investment Costs |  |
| Major Tech Giants (Android, Apple, Microsoft) | Establishing Collaborations with Studios |  | Seamless Termination Process | Active Social Media Engagement (Facebook, Instagram, and more) |  |  | Research and Development Costs |  |
| Data Analytics Powerhouses (Amazon, Google) | Compliance with Legal Regulations |  | Personalized Dynamic Suggestions | User-Friendly Interface |  |  | Data Center Maintenance |  |
|  | Managing Content Filtering |  | Multi-Account Accessibility |  |  |  | Workforce Salaries |  |
| **Table 2. Fictional Business Model.** | | | | | | | | |

**End of the Chapter Pedagogy: Examples**

*Ethical Dilemma*

Business analytics consultants were retained by the state lottery commission to conduct a cluster analysis to gain a better understanding of public attitudes in the state toward the lottery and gambling. The commission plans to use the data to design marketing programs aimed at increasing support and participation in the weekly lottery drawings. The research identified five distinct clusters of citizens. Not surprisingly, the largest cluster (22 percent) was opposed to the lottery and had a low understanding of the economic impact the lottery had on the state. The smallest cluster (13 percent) was made up of people who loved the lottery. This group also had the best understanding of the lottery’s impact on their state. The other three groups fell somewhere in between in both attitude and knowledge—qualified support (25 percent), neutral (27 percent), and nongamblers, who see some benefit for the state (13 percent).

Before the business analytics team begins to work with the marketing group to design the communications strategy for each audience, the group is considering issuing a news release to announce the research findings, focusing on the fact a majority of citizens are in favor of the lottery in hopes of making a positive impression on the state’s legislature. Only one of the lottery commissioners is uncomfortable with the plan because he feels they are overstating the results. What do you think? Is the marketing team misrepresenting the data to the public?

*Review Questions*

1. What is the difference between supervised and unsupervised analytics methods?
2. What is machine learning (ML) and when would a business analytics researcher use it?
3. How does the researcher determine the optimal ML solution?
4. How does machine learning differ from neural networking?
5. Explain the different approaches to executing natural language processing.
6. What types of problems could Cognitive Analytics solve for business analytics professionals?

*Discussion and Thinking Activities*

1. Describe a research problem that would benefit from a machine learning (ML). What are the variables, and what would you hope to achieve?
2. How could a business analytics researcher use cluster analysis to improve the validity of a research problem? Give an example to clarify your answer.
3. Describe a research problem that would require the use of sentiment analysis in combination with natural language processing.
4. Thinking of your research, describe a research problem that should use structural equation modeling. If you decide to apply the PLS-SEM approach, how will this influence the design of your study?
5. *Business Analytics Application:* Run a random forest model using the twelve perceptions variables from the continuing case database to select the most important features that explain customer care satisfaction. Prepare a brief report on your findings.
6. *Business Analytics Application:* Run a cluster analysis using movie genre, movie rater gender, and movie rater age (X2, X13, X14) from the continuing case database and identify age group features by gender and movie genre. Prepare a brief report on your findings.
7. *Business Analytics Application:* Run a neural networking analysis using the twelve perception variables from the customer survey database. Prepare a brief report on your findings.
8. *Business Analytics Application:* Run a cluster analysis using the three variables related to customer perceptions of employees from the customer survey database. Develop a three-cluster solution instead of the two-cluster solutions reported in the chapter. Prepare a brief report on your findings.
9. *Business Analytics Application*: For the continuing case database develop a cognitive analytics solution using Convoluted Neural Network in *Google Colab* that portrays the relationship between actors’ facial features and viewer rating. Prepare a brief report on your findings.

*Online Exercises*

1. Use the key words “multivariate” and “multivariate data analysis” with the Google search engine (http://www.google.com). Prepare a brief report on what you found.
2. Use your favorite Internet search engine. Type in the key words “neural networks”, “machine learning,” and “natural language processing.” Prepare a brief report on the similarities and differences you found about these methods and how they might be used to solve business analytics problems.
3. Execute two Google searches. Use the key words “Python” and “Tableau” separately. Report similarities and differences in the two visualization software platforms.
4. Execute a Google trend search (https://trends.google.com/trends/) for the past 12 months using the key words “Python”, “R”, and “Tableau” separately. Discuss the time and geographic trends.
5. Consider a business scenario where a company is evaluating the effectiveness of a new marketing campaign aimed at increasing customer engagement. Using appropriate statistical techniques, companion library of your choice (R or Python) and the companion marketing dataset, compute and report the numerical estimate of the campaign's impact on customer acquisition or retention. Discuss the statistical and business significance of this estimate and prescribe strategic a potential strategic marketing policy.
6. Imagine you are tasked with designing a recommendation system for a well-known e-commerce platform. How would you assess the effectiveness of the recommendation system in terms of user satisfaction and increased sales? Discuss the key metrics and evaluation methods you would employ to evaluate the system's performance and effectiveness. Provide details on how these metrics can optimize the recommendation system, thereby enhancing user experience and boosting business profits.

**Students Prerequisites.**

Our textbook exposes students to contemporary applications of Business Analytics. It is designed for students enrolled in Masters’ Programs and practitioners. Since we are covering basic statistical concepts, students will not need to have acquired prerequisites beyond high school calculus.

**R Library Companion.**

The users will be able to access the libraries through the R packages[[3]](#footnote-3) repository[[4]](#footnote-4) making it publicly available to over 2 million R users all over the world − hence incentivizing the purchase of the textbook. The library can be loaded in any R environment (Windows, Macintosh, UNIX, Linux, and Anaconda[[5]](#footnote-5)).

**Python Library Companion.**

The users will be able to access the libraries through the open and available Github[[6]](#footnote-6) Repository system. The textbook will have a repository that will act as a companion and be a singular location for all code related to specific sections. This will include things such as data files and Jupyter notebooks[[7]](#footnote-7) in addition to any other dependencies that are required to satisfy the various coding methodologies. Because of the inclusion of Jupyter notebooks, we also can implement the code in the browser natively using Google Colab[[8]](#footnote-8). Additionally, the repo will be installable as a self-contained library within a user’s python environment. To do so will be available through the Github repo as well.

**Proposed Textbook Structure**

1. **Introduction to Business Analytics**
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Data Science and Business Analytics
   4. Differences and Similarities. Why these disciplines are important?
   5. Generative AI and Business Analytics
   6. Descriptive Analytics
   7. Predictive Analytics
   8. Prescriptive Analytics
   9. Data structure
   10. Data frames
   11. Matrices
   12. Arrays
   13. Pivot Tables
   14. Chapter Summaries organized by chapter objectives.
2. **Data Management and Manipulation**
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Dealing with various data type inputs
   4. Merging Data
   5. Storing data
   6. Manipulating data types
   7. Geographic Data
   8. Generative AI and Data Manipulation
   9. Chapter Summaries organized by chapter objectives.
3. **Foundations of Statistics**
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Elements of probability theory
   4. Statistical distributions
   5. Inferential Statistics
   6. Supervised vs. Unsupervised Approaches
   7. Hypothesis Testing
   8. Data Exploration
   9. A/B Testing (Tuckfield)
   10. Chapter Summaries organized by chapter objectives.
4. **Econometric Methods for Business Analytics** 
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Regression Analysis
   4. Statistical Optimization
   5. Logistic Regression
   6. Probit Models
   7. Panel Data Models
   8. Time Series Models and Forecasting
   9. Chapter Summaries organized by chapter objectives.
5. **Advanced Econometric Methods for Business Analytics** 
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Censoring and Survival Analysis
   4. Causal Models
   5. Event Studies and Difference-in-Differences Models
   6. Heckman and Endogenous switching regime models
   7. Matching Algorithms
   8. Penalized Regression Models: Ridge, LASSO, and elastic net
   9. Partial Least Squares
   10. Structural Equation Models
   11. Generative AI and Econometrics
   12. Chapter Summaries organized by chapter objectives.
6. **Supervised Machine Learning** 
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Data splitting, validation
   4. Classification vs. Regression
   5. Confusion Matrix
   6. Model Metrics
   7. Decision Trees
   8. Random Forests
   9. XGboost
   10. Gaussian
   11. NN
   12. Hyper Parameterization
   13. Overfitting, Bias variance tradeoff and Regularization
   14. Cross validation and Bootstrapping
   15. K-Nearest Neighbors (KNN) algorithm
   16. Chapter Summaries organized by chapter objectives.
7. **Visualization of Multivariate Data** 
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Dimensionality Reduction
   4. Principal Component Analysis (PCA)
   5. Robust PCA
   6. Sparse PCA (LASSO based PCA)
   7. Principal Coordinate Analysis (PCoA)
   8. Kernel PCA
   9. Penalized PCA
   10. t-Distributed Stochastic Neighbor Embedding (t SNE)
   11. Chapter Summaries organized by chapter objectives.
8. **Unsupervised Machine Learning** 
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Statistics vs. Machine Learning
   4. K Means Clustering
   5. PAM (K Medoids) clustering
   6. Hierarchical Clustering
   7. Gaussian Mixture Model clustering (EM algorithm)
   8. Density-Based Spatial Clustering of Applications with Noise (DBSCAN)
   9. Chapter Summaries organized by chapter objectives.
9. **Deep Learning**
   1. Chapter Objectives
   2. Chapter Introduction and Overview
   3. Introduction to Neural Networks
   4. Single unit and activation functions
   5. Architecture Shallow vs Deep
   6. Basic Training Algorithms, Stochastic Gradient Descent, Minibatch
   7. Convolutions and Pooling layer
   8. Transfer Learning
   9. Graph neural networks: Applications, graph convolutions
   10. PyTorch
   11. Transformer Architectures
   12. Chapter Summaries organized by chapter objectives.
10. **Natural Language Processing**
    1. Chapter Objectives
    2. Chapter Introduction and Overview
    3. Audio and Video Data: Speech Recognition
    4. Tokenization in NLP
    5. Stemming and Lemmatization
    6. Stop Words Removal
    7. Term Frequency-Inverse Document Frequency (TF-IDF)
    8. Keyword Extraction in NLP
    9. LDA Models
    10. Word Embeddings
    11. Sentiment Analysis
    12. LLMs
    13. Architecture of ChatGPT
    14. Chapter Summaries organized by chapter objectives.
11. **Recommendation Systems**
    1. Chapter Objectives
    2. Chapter Introduction and Overview
    3. What are Recommendation Systems
    4. AI and Recommendation Systems
    5. Performance measures
    6. Sparsity of Data and Time Varying Data
    7. Modelling process
    8. Solutions Improvement and Clustering
    9. Collaborative Filtering
    10. Singular Value Thresholding
    11. Clustering based recommendation systems
    12. Content based recommendation systems
    13. Predictive recommender engine
    14. Hybrid Recommendation Systems
    15. Chapter Summaries organized by chapter objectives.
12. **Elements of SQL, R, and Python** 
    1. Chapter Objectives
    2. Chapter Introduction and Overview
    3. SQL and relational db structures
    4. Overview of the R Studio platform
    5. Overview of the Python platform
    6. Web Scraping
    7. Chapter Summaries organized by chapter objectives

**How Contemporary Business Analytics by Hair et al., extends the Teddy et al. text**

* **Regression.** Our book will include probit regression, a topic not covered by Taddy et al.
* **Unsupervised Learning.** Our book will have extensive coverage of unsupervised learning, including k-means clustering, which is not covered by Taddy et al.
* **Programming Languages.** Our book contains coverage of Python and SQL in addition to the R platform – the only language covered by Taddy et al.
* **Econometrics and causal inference.** Our deep coverage of econometrics will provide a more comprehensive treatment of causal inference, including more valid approaches than the causal inference methods in Taddy et al. For example, we extend the LASSO method covered by Taddy et al. by including ridge and the elastic net, a better compromise between ridge and LASSO.
* **NLP.** Our book includes deeper coverage of NLP, including coverage of the architecture of GPT, which is not covered by Taddy et al.
* **Visualization.** Our book includes t-SNE, which is not included in Taddy et al.
* **Data management.** Our book includes practical topics related to data management, including how to merge tables and run basic SQL commands, which is more than is covered by Taddy et al.
* **Deep Learning.** Our coverage of deep learning is much deeper than Taddy et al., including graph neural networks and transfer learning.
* **Recommendation systems.** Our book contains extensive coverage of recommendation systems, a topic that is not covered by Taddy et al.

**Authors**

|  |  |  |
| --- | --- | --- |
|  | **Joseph F. Hair, Jr.** | Joe is the Director, Ph.D. Program in Business Administration, Mitchell College of Business, University of South Alabama, U.S.A. Joe is ranked #1 globally in Multivariate Data Analysis, SEM, and PLS-SEM, and his career citations exceed 388,000, H-Index is 127 and I-10 483, and ScholarGPS rank #1 in Marketing among all academic scholars in the United States. Joe has co-authored over 100 editions of his books, including *Multivariate Data Analysis*, Cengage Learning, U.K., 8th edition 2019; *A Primer on Partial Least Squares Structural Equation Modeling*, Sage, 3rd edition, 2022; *Advanced Issues in Partial Least Squares Structural Equation Modeling*, Sage, 2nd edition 2023; and *Essentials of Marketing Analytics*, McGraw-Hill, 2nd edition 2024. He has published 160+ articles in scholarly journals such as the *Journal of Marketing Research*, *Journal of Academy of Marketing Science*, *Organizational Research Methods*, *Harvard Business Review, European Journal of Marketing*, *European Management Journal*, and others. |
|  | **Ermanno Affuso** | Ermanno is a Professor of Economics and Chief Scientific Officer of the South Alabama Center for Business Analytics, Real Estate, and Economic Development. He received his PhD in Applied Economics from Auburn University. Dr. Affuso has shared his expertise through teaching engagements at various educational institutions within the United States and internationally. His dedication to research and education has earned him multiple prestigious awards, including the Fulbright Specialist Award in 2019 and the Fulbright Scholar Award in 2023. Dr. Affuso's professional focus centers on Econometrics, Data Science, and Business Analytics, particularly emphasizing their applications in Sustainability. |
|  | **Joshua Lambert** | Josh is a Sr. Machine Learning Engineer for Factset Data Systems where he works on Graph based data solutions and Generative AI. He received a PhD in Security Studies from the University of Central Florida where his dissertation utilized spatial contagion modeling and machine learning to forecast military coups globally. He has published articles in Marine Policy, Foreign Policy Analysis, Journal of Peace Research, Water, Small Wars and Insurgencies, and International Journal of the Commons. |
|  | **Bradford Tuckfield** | Bradford is a data scientist and consultant. He received his PhD from the Wharton School of the University of Pennsylvania. He is the author of *Dive Into Algorithms* and other books on data science and machine learning. He holds two patents related to NLP and risk analysis. |
|  | **Michail Tsagris** | Michail is an Assistant Professor at the Department of Economics of the University of Crete and an Adjunct Professor in the Department of Mathematics and Statistics at the University of New Brunswick Saint John. Prior to these he held contract positions at the University of Crete, the American University of the Middle East, the University of Nottingham and Nottingham Trent University. He received his PhD in statistics from the University of Nottingham and his BSc & MSc in statistics from the Athens University of Economics and Business. Michail has published more than 50 papers in journals, conference proceedings and book chapters, and has (co-)developed 17 R packages. His publications include notable journals such as *Journal of the American Statistical Association*, *Statistics and Computing*, and *Bioinformatics*. |

**References**

Association to Advance Collegiate Schools of Business (AACSB, 2018) – url:<https://www.aacsb.edu/insights/articles/2018/02/an-insatiable-demand-for-analytics-talent-is-reshaping-business-education> - last access: 9/24/2023

MacKenzie, I., C. Meyer, and S. Nobb. 2013. “ How retailers can keep up with consumers”, url: <https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers> - last access: 9/24/2023

Wellington, R. 2019. “Content Finance Engineering at Netflix”, url: <https://jobs.netflix.com/blog/content-finance-engineering-at-netflix> - last access: 10/23/2023/

1. The company name can be changed. We chose this name because it was the closest to Netflix free of trademark. [↑](#footnote-ref-1)
2. <https://jobs.netflix.com/blog/content-finance-engineering-at-netflix> (retrieved on 10/23/2023) [↑](#footnote-ref-2)
3. <https://cran.r-project.org/web/packages/> [↑](#footnote-ref-3)
4. <https://cran.r-project.org/web/packages/available_packages_by_date.html> [↑](#footnote-ref-4)
5. <https://www.anaconda.com/> [↑](#footnote-ref-5)
6. <https://github.com/> [↑](#footnote-ref-6)
7. <https://jupyter.org/> [↑](#footnote-ref-7)
8. <https://colab.google/> [↑](#footnote-ref-8)