

# MBA EM DATA SCIENCE E ANALYTICS

ementas do curso

segundo semestre de 2021

**MBA**USP  
ESALQ

## **MÓDULO - TENDÊNCIAS EM DATA SCIENCE E ANALYTICS**

### **Data Wrangling**

Manipulação de dados I; Manipulação de dados II; Criação de Projects e de scripts RMarkdown; Projetos de Data Science & Analytics no GitHub; Programação funcional com o pacote purrr e Visualização Avançada de Dados.

### **Análise Espacial**

Shapefiles - apresentação, manipulação de dados, recorte e merge de shapefiles; Objetos Simple Feature; Imagens Raster; Análise Exploratória Espacial; Aplicações de bancos de dados reais de mercado em R.

### **Conceitos básicos para Introdução ao Deep Learning**

Conceito básico de rede neural artificial; Construção de rede neural artificial sem pacotes computacionais.

### **Deep Learning**

Conceitos de otimização de Redes Neurais Artificiais e funcionamentos; Tipos de Feed Forward Neural Networks; Feed Forward Neural Network; Convolutional Neural Networks: aplicação prática em R; Recursive Neural Networks; Introdução à Boltzmann Machines.

### **Coleta de Dados: Crawlers e Web Scraping**

Introdução ao HTML; HTML Parsing; Escrevendo um scraper com R.

### **Text Mining e Sentiment Analysis**

Pré-processamento de textos; Algoritmos de classificação; Detecção de emoções com Naive Bayes; Uso de outros algoritmos para análise de sentimentos.

### **Pesquisa Operacional e Modelos de Otimização e Simulação**

Pesquisa operacional: função-objetivo, parâmetros, variáveis de decisão e restrições em modelos de otimização; Programação linear; Métodos de otimização; Programação inteira; Programação binária; Simulação de Monte Carlo; Aplicações de bancos de dados reais de mercado em Excel.

### **Cloud Computing**

Arquitetura de computação em nuvem; serviços na nuvem e soluções para armazenamento; Casos e aplicações de computação na nuvem; Arquitetura e infraestrutura para big data e analytics; Computação em nuvem; HDFS; Hadoop e MapReduce; Spark, Casos aplicados.

### **Introdução à Segurança de Dados e Cyber Security**

A importância da segurança e da proteção de dados na área de negócios; princípios ligados à segurança da informação e de dados; soluções mais utilizadas para a segurança da informação e de dados; cyber security; governança em TI.

### **Bibliografia do módulo**

ALCANTARA, L. K. Big Data e IoT: Desafios da Privacidade e da Proteção de Dados no Direito Digital. São Paulo: Bok2, 2017.

ALBACH, J. Designing web navigation. O'Reilly Media, 2007.

ANDRADE, E.L. Introdução à pesquisa operacional: métodos e modelos para análise de decisões. Rio de Janeiro: LTC, 2009.

ALMEIDA, E. (2012). Econometria Espacial Aplicada. São Paulo: Alínea.

AMARAL, F.C.N. Data Mining: Técnicas e Aplicações para o Marketing Direto. São Paulo: Editora Berkeley, 2001.

ANSELIN, L. (1995). Local Indicators of Spatial Association-LISA. *Geographical Analysis*, 27(2), 93-115.

ANSELIN, L. & REY, S. J. (2014). *Modern Spatial Econometric in Practice*. Chicago: GeoDa Press.

ASSAF NETO, A. Finanças corporativas e valor. São Paulo: Atlas, 2003.

ASSAF NETO, A. Matemática Financeira e suas Aplicações. 12ª Ed. São Paulo: Atlas, 2012.

BASTOS, D. B.; NAKAMURA, W. T. Determinantes da estrutura de capital das companhias abertas no Brasil, México e Chile no período 2001-2006. *Revista Contabilidade e Finanças*, v. 20, n. 50, p. 75-94, 2009.

BARABÁSI, A. *Linked: how everything is connected to everything else and what it means for business, science, and everyday life*. New York: Plume, 2003

BEKAERT, G.; HARVEY, C. R. Research in emerging markets finance: looking to the future. *Emerging Markets Review*, v. 3, n. 4, p. 429-448, 2002.

BERGSON, L. R. *Gestão e Governança de Dados: promovendo dados como ativo de valor nas empresas*. Rio de Janeiro: Editora Brasport, 2013.

BERRY, M.J.A; LINOFF, G. *Data Mining Techniques For Marketing, Sales and Customer Support*. 3ª ed. New York, USA: John Wiley & Sons, Inc., 2011.

BESANKO, D; BRAEUTIGAM, R. *Microeconomics*. 5ª ed. Wiley, 2014.

BIVAND, R. (2014). SPDEP: Spatial Dependence: Weighting Schemes, Statistics and Models. R package version 1.1-2. <https://cran.r-project.org/web/packages/spdep/index.html>.

BIVAND, R., Keitt, T. & Rowlingson, B. (2018). rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.4-4. <https://cran.r-project.org/web/packages/rgdal/index.html>.

BIVAND, R & RUNDEL, C. (2012). rgeos: Interface to Geometry Engine - Open Source ('GEOS'). R package version 0.5-1. <https://cran.r-project.org/web/packages/rgeos/index.html>.

BRIGGS, A. *Uma História Social da Mídia: De Gutenberg à Internet*. Rio de Janeiro: Editora Jorge Zahar, 2004.

BRITO, G. A. S.; ASSAF NETO, A. Modelo de risco para carteiras de créditos corporativos. *Revista de Administração (RAUSP)*, v. 43, n. 3, p. 263-274, 2008.

BRITO JÚNIOR, I. *Análise do impacto logístico de diferentes regimes aduaneiros no abastecimento de itens aeronáuticos empregando modelo de transbordo multiproduto com custos fixos*. Dissertação (Mestrado em Engenharia de Sistemas Logísticos), Escola Politécnica da Universidade de São Paulo, São Paulo, 2004.

BRITO JÚNIOR, I.; YOSHIZAKI, H. T. Y.; BELFIORE, P. Um modelo de localização e transbordo multiproduto para avaliação do impacto de regimes aduaneiros. *Transportes*, v. 20, n. 3, p.89-98, 2012.

BRUNSDON, C. & CUMBER, L. (2019). *An Introduction to R for Spatial Analysis and Mapping*. Londres: Sage Publications.

CAMM, J. D. et al. *Business Analytics*. 3ª ed. São Paulo: Cengage Learning, 2018.

CAMPBELL, J. Y.; LO, A. W.; MACKINLAY, A. C. *The econometrics of financial markets*. Princeton: Princeton University Press, 1997.

CHEN, H., CHIANG, R. H. L.; STOREY, V. C. *Business Intelligence and Analytics: from Big Data to Big Impact*. *Journal MIS Quarterly*, vol. 36, no. 4. p. 1165-1188, 2012.

CHEN, Y. et al. Big data analytics and big data science: a survey. *Journal of Management Analytics*, v. 3, n. 1, p. 1-42. fev. 2016.

CHISHTI, S. *The FINTECH Book: The Financial Technology Handbook for Investors, Entrepreneurs and Visionaries*. New York, USA: Wiley, 2016.

CHOPRA, S.; MEINDL, P. Gestão da cadeia de suprimentos: estratégia, planejamento e operações. 4. ed. São Paulo: Pearson Prentice Hall, 2011.

CIALDINI, R. As Armas da Persuasão. Sextante, 2012.

CLIFF, A.D., and ORD, J.K., 1973, Spatial autocorrelation (London: Pion).

CLIFF, A.D., and ORD, J.K., 1981, Spatial processes: models and applications (London: Pion).

COLIN, E.C. Pesquisa operacional: 170 aplicações em estratégia, finanças, logística, produção, marketing e vendas. Rio de Janeiro: LTC, 2007.

COSTA, L. G. T. A.; COSTA, L. R. T. A.; ALVIM, M. A. Valuation. 2ª ed. Atlas, 2011.

DAMODARAN, A. Avaliação de Investimentos. Rio: Qualitymark, 1997.

DAVENPORT, T. H.; HARRIS, J. G.; MORISON, R. Analytics at work: smarter decisions, better results. Boston, MA: Harvard Business Press, 2010.

DE FARO, C. Fundamentos de Matemática Financeira. Saraiva, 2006.

ENGE, E. Arte de SEO – Dominando a Otimização dos Mecanismos. Novatec, 2012.

FAWCETT, T.; PROVOST, F. Data Science para negócios. 1ª ed. Alta Books: 2016.

FERNADEZ, G. Data Mining Using SAS Applications. New York: Chapman & Hall/CRC, 2003.

FLATSCHART, Fábio. HTML5 – Embarque Imediato. Rio de Janeiro: Editora Brasport, 2011.

FLEMING, S. Blockchain technology and microservices architecture. USA: eBook Kindle, 2018.

FORTIN, M.J. and DALE, M.R.T. (2005) Spatial Analysis: A Guide for Ecologists. Cambridge University Press, Cambridge, 1-30.

FOTHERINGHAM, A. S.; BRUNSDON, C.; CHARLTON, M. (2002). Geographically Weighted Regression: the analysis of spatially varying relationships. West Sussex: John Wiley and Sons.

GABRIEL, M. Você, Eu e os Robôs. São Paulo: Editora Atlas, 2018.

GABRIEL, M. SEM & SEO: Dominando o marketing de busca. São Paulo: Novatec Editora, 2010.

GETIS, A. & ORD, K. (1992). The Analysis of Spatial Association by Use of Distance Statistics. Geographical Analysis, 24(3), 189-206.

GOERTZEL, B. e PENNACHIN, C. Artificial general intelligence. USA: Springer, 2010.

GRACIOSO, F. Marketing Estratégico. Editora Atlas, 2007.

GRANER, E.A. Como Aprender Estatística. Edições Melhoramentos, 1952.

GRAY, D. A Empresa Conectada. Novatec, 2019.

GRIFFITH, D.A. (2003). Spatial autocorrelation and spatial filtering: gaining understanding through theory and scientific visualization. Berlin: Springer-Verlag.

GUBBI, J.; BUYYA, R.; MARUSIC, S.; PALANISWAMI, M. Internet of Things (IoT): A vision, architectural elements, and future directions, Future Generation Computer Systems, vol. 29, i. 7, p. 1645-1660, September 2013.

HAN, J., KAMBER, M. Data Mining: Concepts and Techniques. New York: Morgan Kaufmann Publisher, 2001.

HIJMANS, R. & VAN ETEN, J. (2014). raster: Geographic data analysis and modeling. R package version 2.9-23. <https://cran.rproject.org/web/packages/raster/raster.pdf>.

HOFFMANN, R. Estatística para Economistas. 4ª. Edição. Thomson Pioneira, 1980.

HOFFMANN, R.; VIEIRA, S. Análise de Regressão. São Paulo: Atual, 2a. ed., 1977.

HOJI, M; DA LUZ, A. E. Gestão Financeira e Econômica. São Paulo: Editora Atlas, 2018.

HUA, H. Mobile marketing management: case studies from successful practices. USA: Productivity Press, 2019.

IEMMA, A.F. Estatística Descritiva. Piracicaba: Fi Sigma Rô Publicações, 1992.

Islam, S. (2018). Hands-On Geospatial Analysis with R and QGIS: A beginner's guide to manipulating, managing, and analyzing spatial data using R and QGIS 3.2.2. Birmingham, UK: Packt Publishing.

ISMAIL, S.; MALONE, M. S.; YURI, V. G. Organizações Exponenciais. São Paulo: Alta Books, 2015.

JENKIS, H. Cultura da Convergência. São Paulo. Editora Aleph, 2008.

JERKOVIC, J.I. Guerreiro SEO. Novatec, 2010.

JOHNSON, S. Cultura da interface: como o computador transforma nossa maneira de criar e comunicar. Rio de Janeiro. Editora Jorge Zahar, 2001.

KELLY, K. The Inevitable: Understanding the 12 technological forces that will shape our future. New York, USA: Penguin, 2016.

KING, B. Bank 4.0: Banking Everywhere, Never at a Bank. 1ª ed. John Wiley & Sons: 2019.

KIRNER, C.; SISCOOTTO, R. Realidade Virtual e Aumentada: Conceitos, Projeto e Aplicações. Petrópolis – RJ, Livro do Pré-Simpósio, IX Symposium on Virtual Reality. Editora SBC – Sociedade Brasileira de Computação, 2007.

KLEIN, N. Sem logo: a tirania das marcas em um planeta vendido. Rio de Janeiro: Record, 2002.

KINGSNORTH, S. Digital marketing strategy: an integrated approach to online marketing. USA: Kogan Page, 2019.

KRUG, S. Don't make me Think. Pearson Educational. 2010.

LANSLEY & CHESHIRE. (2016). An Introduction to Spatial Data Analysis and Visualisation in R. Geospatial Analysis – A comprehensive guide – 2018. <http://www.spatialanalysisonline.com>

LAURINDO, F. J. B.; ROTONDARO, R. G. (orgs.) Gestão integrada de processos e da tecnologia da informação. São Paulo: Atlas, 2006.

LEWIS-KOH, N., BIVAND, R. & PEBESMA, E. (2011). maptools: Tools for reading and handling spatial objects.. R package version 0.9-5. <https://cran.r-project.org/web/packages/maptools/maptools.pdf>.

LOUCKS, J.; MACAULAY, J.; NORONHA, A.; WADE, M. Digital Vortex: How Today's Market Leaders Can Beat Disruptive Competitors at Their Own Game. 1ª ed. DBT Center Press: 2016.

LUCCI, S. e KOPEC, D. Artificial intelligence in the 21st century. USA: Mercury Learning & Information, 2019.

LUGER, G. Inteligência artificial. São Paulo: Pearson, 2013.

MACHADO, F. N. R. Big Data. O Futuro dos Dados e Aplicações. São Paulo: Editora Érica, 2018.

MAGNANI, E. A internet das coisas. Rio Janeiro: Editora FGV, 2018.

MAHESHWARI, A. Data analytics made accessible. USA: eBook Kindle edition, 2019.

MANSUR, R. Governança da Nova TI: A Revolução. Rio de Janeiro: Editora Ciência Moderna, 2013.

MARQUESONE, R. Big Data/Técnicas e tecnologias para extração de valor dos dados. São Paulo: Casa do Código, 2017.

MARTINS, E. (Org.). Avaliação de Empresas: Da mensuração contábil à econômica. São Paulo: Atlas, 2001.

MAS, J. F., HORTA, M. B., VASCONCELOS, R. N. & CAMBUI, E. C. B. (2019). Análise espacial com R, Feira de Santana (BA): UEFS Editora.

MCCARTHY, J.; WRIGHT, P. Technology as Experience. Cambridge, Massachusetts: The MIT Press, 2004.

MEDEIROS, L. F. de. Inteligência artificial aplicada: uma abordagem introdutória. Curitiba: Editora InterSaberes, 2018.

MEGLIORINI, E.; VALLIN, M. A. Administração Financeira. São Paulo: Pearson 2018.

MENEGHETTI NETO, A. Educação Financeira. Porto Alegre: EDIPUCRS, 2014.

MORAN, P.A.P., 1948, The interpretation of statistical maps, Journal of the Royal Statistics Society, Series B (Methodological), 10, 2, pp. 243 - 251.

MOUGAYAR, W. Blockchain para negócios: Promessa, prática e aplicação da nova tecnologia da Internet. 2017.

NESELLO, P; FACHINELLI, A. C. Big Data: O novo desafio para a gestão. Revista Inteligência Competitiva, v. 4, p. 18-38, 2014.

NIELSEN, J. Usabilidade Móvel. 2ª ed. Elsevier, 2013.

O'BRIEN, J. A. Sistemas da informação e as decisões gerenciais na era da Internet. 3ª ed. São Paulo: Saraiva, 2011.

OLIVEIRA, E. Contabilidade digital. São Paulo: Editora Atlas, 2014.

PALFREY, J.; GASSER, U. Born Digital: Understanding the First Generation of Digital Natives. Basic Books, 2008.

PEBESMA, E., BIVAND, R., COOK, I. & KEITT, T. (2017). sf: Simple Features for R. R package version 0.7-7. <https://cran.rproject.org/web/packages/sf/index.html>.

PHAAL, R., FARRUKH, C. PROBERT, D. Technology roadmapping - A planning framework for evolution and revolution. Technological Forecasting and Social Change, v. 71, p. 5-26, 2004.

PIATKOWSKA, S. J., MESSNER, S. F. & YANG, T. C. (2018). Xenophobic and racially motivated crime in Belgium: exploratory spatial data analysis and spatial regressions of structural covariates. Deviant Behavior, 39(11), 1398-1418.

PROVOST, F; FAWCETT, T. Data Science para Negócios. Rio Janeiro: Editora Alta Books, 2016.

ROMANO, R. R. Tecnologia: O futuro do seu negócio passa por aqui. 3ª. Edição. São Paulo: Posigraf, 2012.

SAFKO, L. A Bíblia da Mídia Social. Blucher, 2010.

SALEH, K.; SHUKAIRY, A. Otimização de conversão – A Arte e a Ciência de converter prospects em clientes. Novatec, 2019.

SAMANEZ, C. P. Matemática Financeira: uma aplicação à análise de investimentos. Editora Pearson/Prentice Hall, 2008.

SAYLOR, M.J. The Mobile Wave: How Mobile Intelligence Will Change Everything. Perseus Books/Vanguard Press, 2012.

SCHAHRAM, D. Smart Cities: The Internet of Things, People and Systems New York, USA: Springer, 2017.

SCHWAB, K. A. Quarta Revolução Industrial. 1ª ed. Edipro: 2015

SCOTT, D. Marketing e comunicação na era digital. São Paulo: Évora, 2015.

SIEGEL, E. Análise preditiva. 1ª ed. Alta Books: 2017

SINCLAIR, B. IoT Inc: How your company can use the IoT to win in the outcome economy. New York, USA: McGraw-Hill Education, 2017.

SMASHING, Magazine. The Mobile Book. Smashing Media GmbH, 2012.

STONE, B. A Loja de Tudo - Jeff Bezos e A Era da Amazon. Intrínseca, 2014.

SWAN, M. Blockchain: blueprint for a new economy. USA: O'Reilly Media, 2015.

TAPSCOTT, D.; TAPSCOTT, A. Blockchain Revolution. 1ª ed. SENAI-SP: 2017.

TAPSCOTT, D. Economia digital: promessa e perigo na era da inteligência em rede São Paulo: Makron, 2007.

\_\_\_\_\_ WILLIAMS, A.D. Wikinomics. Como a colaboração de massa pode mudar o seu negócio. São Paulo: Nova Fronteira, 2007.

\_\_\_\_\_ Macrowikinomics: reiniciando os negócios e o mundo. Rio de Janeiro: Elsevier, 2011.

TAURION, C. Big Data. Rio Janeiro: Editora Brasport, 2013.

TELLES, A. A revolução das mídias sociais. M.Books, 2011.

TENNEKES, M. (2018). tmap: Thematic Maps in R. Journal of Statistical Software, 84(6), 1-39, doi: 10.18637/jss.v084.i06.

TIEFELSDORF, M., GRIFFITH, D. A. & BOOTS, B. (1999). A Variance-Stabilizing Coding Scheme for Spatial Link Matrices. Environment and Planning A, 31(1), 165-180.

TORRES, C. A bíblia do marketing digital. Novatec, 2018.

TURBAN, E.; MCLEAN, E.; WETHERBE, J. Information Technology for Management: Transforming Organizations in the Digital Economy. 7th edition, Wiley, 2009.

TURBAN, E.; VOLONINO, L. Tecnologia da informação para gestão: em busca do melhor desempenho estratégico e operacional. 8ª ed. Porto Alegre: Editora Bookman, 2013.

TURCHI, S.R. Estratégias de Marketing Digital e E-commerce. 2ª ed. São Paulo: Editora Gen, 2018.

ULRICH, F. Bitcoin: A moeda na era digital. São Paulo: Instituto Ludwig Von Mises Brasil, 2014.

VAZ, C. A. Os 8 Ps do Marketing Digital: O Seu Guia Estratégico de Marketing Digital. 1ª ed. Novatec, 2011.

WEIL, R. L, SCHIPPER, K, FRANCIS J. Financial Accounting: An Introduction to Concepts, Methods and Uses. 14ª ed, 2013.

WILLIAMS, A.D., TAPSCOTT, D. Radical Openness: Four Unexpected Principles for Success. TED Conferences, 2013.

WROBLEWSKI, L. Mobile First. A Book Apart LLC, 2011.

## **MÓDULO - DATA SCIENCE E BIG DATA NO AMBIENTE DE NEGÓCIOS**

### **Big Data e o Processo de Tomada de Decisão**

Fundamentos do Big Data; Ciclo de Vida dos Dados; O que são Estratégias de Dados; Estruturando um case Data Lake; Perfis e responsabilidades de profissionais Big Data; Casos de sucesso do uso de Big Data; Analytics como ferramenta de decisão aplicando Big Data; Introdução a Clusterizações de computadores; Os conceitos fundamentais para compreender o Framework Hadoop; Como Hadoop garante Segurança e Privacidade dos dados para demandas de negócio; Demonstrações de Serviços e Gerenciadores de um ambiente Hadoop.

### **Data Mining**

Visão geral sobre sistemas de suporte à decisão e inteligência de negócio; entendimento sobre os três principais frameworks para ciências de dados - KDD, CRISP-DM e SEMMA e, em especial, sobre preparação de dados; estudo sobre as principais tarefas de mineração de dados e suas três etapas: exploração, definição de modelo e verificação; a mineração de dados e suas utilizações em estudos de casos reais.

### **Business Intelligence e Data Visualization**

Gráficos adequados a cada tipo de variáveis; Conceito de Business Intelligence; Simplificação e classificação de dados; Mapeamento de dados geograficamente; Personalização de dados e unidades; Análise de dados a partir de cálculos rápidos provenientes de tabelas; Criar segmentações (Filtros) no Power BI; Aplicações de bancos de dados reais em Microsoft Power BI; Criação de templates profissionais em Power BI e elaboração e customização de Dashboards em consonância com o manual das marcas de uma empresa.

### **Tecnologia da Informação (TI) e Inovação Tecnológica**

Como obter ganhos a partir da adoção das novas tecnologias digitais?; O papel da Tecnologia da Informação (TI) na transformação digital; O uso do TI como arma competitiva e de alavancagem de negócios; O relacionamento do líder digital x CIO; Shadow IT; Alinhamento estratégico de TI; Governança de TI e Planejamento Estratégico. Estruturação da área de T.I

### **Engenharia de Dados**

Introdução ao banco de dados e SQL; Engenharia de dados x ciência de dados; Conceitos de bancos de dados; Sistemas de gerenciamento de bancos de dados; Introdução à modelagem de bancos de dados relacionais; Construindo um diagrama de entidade-relacionamento; Níveis de abstração; Introdução a álgebra relacional; Restrições de integridade referencial e tipos de normalização; Exercícios práticos.

### **Metodologias Ágeis**

Fundamentos da Gestão de Projetos; O que são metodologias ágeis?; Conhecendo as principais metodologias ágeis e onde aplicá-las; Conceitos e principais diferenças: Scrum, Kanban, Lean, Agile Development; Agile Team Organization (Squads, Chapters, Tribes, Guilds); Cases e exemplos práticos em empresas nacionais e internacionais.

### **Disciplina: Analytics e Gestão de Riscos**

Ementa: Modelos de gestão de risco operacional e compliance; Basileia III; modelos de risco de crédito; modelos de risco legal e regulamentar; modelos de risco social e ambiental; riscos reputacionais; riscos cibernéticos; casos reais e práticos.

### **Análise da Conjuntura Econômica em Cenários de Tecnologias Disruptivas**

Fundamentos de política econômica e seus instrumentos; Principais variáveis macroeconômicas e suas relações; PIB - Produto Interno Bruto, taxa de juros, taxa de câmbio e balanço de pagamentos; Sistema financeiro nacional; Impactos do cenário mundial na economia brasileira; Variáveis que impactam o consumo: renda, crédito, emprego e confiança do consumidor; Avaliação de cenário econômico e de tendências

### **Legislação no Ambiente Digital**

Direito Digital; Princípios e responsabilidades - LGPD (Lei nº 113709/2018); Implicações da LGPD nos negócios das empresas varejistas; Transparência x privacidade; Segurança da informação; Cenários de usos e aplicações da LGPD no mercado de consumo.

### **Tópicos Especiais em Tecnologia de Negócios**

Temas contemporâneos em ciência de dados e analytics.

### **Bibliografia do módulo**

- ANDERSON, C. A cauda longa. Rio de Janeiro: Campus, 2006.
- ALBRIGHT, A. C.; WINSTON, W. L. Business analytics: data analysis and decision making. 5. ed. Stamford: Cengage Learning, 2015.
- BACHA, C.J.C.; Entendendo a Economia Brasileira. 2ª ed. Campinas: Alínea, 2011.
- BATEMAN, T. S.; SNELL S. A. Administração: construindo vantagem competitiva. São Paulo: Atlas: 1998.
- BAWENS, M; KOSTAKIS, V. Network society and future scenarios for a collaborative economy. London: Polgrave Pivot, 2014
- BENYON, D. Interação Humano-Computador. 1ª ed. Pearson, 2011.
- BESANKO, D.; DRANOVE, D.; SHANLEY, M. & SCHAEFER, S. Economics of Strategy. 4th Ed. John Wiley & Sons, 2007
- BLOK, M. Compliance e governança corporativa: atualizado de acordo com a Lei Anticorrupção Brasileira (Lei 12.846) e o Decreto-Lei 8.421/2015. Rio Janeiro: Freitas Bastos, 2017.
- BOTELHO, D.; ZOUAIN, D. M. Pesquisa quantitativa em administração. São Paulo: Atlas, 2006.
- BRITO, F; CÂMARA, J. B. Democratização e Gestão Ambiental. Em Busca do Desenvolvimento Sustentável. 3ª Edição, Petrópolis, RJ, Vozes, 1998.
- BROWN, T. Design thinking: Uma metodologia poderosa para decretar o fim das velhas ideias; Rio de Janeiro: Elsevier, 2010.
- BURDEA, G.; COIFFET, P. Virtual Reality Technology. 2ª ed. Wiley, 2003.
- CAVALCANTI, C. (Org.) Meio Ambiente, Desenvolvimento Sustentável e Políticas Públicas. 3ª Edição, São Paulo, Cortez, Recife. Fundação Joaquim Nabuco, 2002.
- CASE, S. A Terceira Onda da Internet: Como reinventar os negócios na era digital. 1ª ed. Alta Books, 2018.
- CHRISTENSEN, C. O dilema da inovação: quando novas tecnologias levam as empresas ao fracasso. 1ª ed. Makron Books, 2011.
- DYER, J; CHRISTENSEN, C. M; GREGERSEN, H. DNA do inovador - dominando as 5 habilidades dos inovadores de ruptura. São Paulo: HSM, 2012.
- DAMODARAN, A. Gestão Estratégica do Risco: uma referência para a tomada de riscos empresariais. Trad. Felix Nonnenmacher. São Paulo: Artmed Editora, 2009.
- DIXIT, A.K.; NALEBUFF, B.J. The Art of Strategy: a Game Theorist's guide to success in business and life. New York: W. W. Norton & Company, 2008.



DYER, J; CHRISTENSEN, C. M; GREGERSEN, H. DNA do inovador - dominando as 5 habilidades dos inovadores de ruptura. São Paulo: HSM, 2012.

GEEST, Y. V. et al. Organizações exponenciais: por que elas são 10 vezes melhores, mais rápidas e mais baratas que a sua (e o que fazer a respeito). São Paulo: HSM Editora, 2015. GRAY, D. A Empresa Conectada. Novatec, 2019.

GREENBERG, B. A.; GOLDSTUCKER, J. L.; BELLENGER, D. N. What techniques are used by marketing researchers in business? *Journal of Marketing*, v. 41, n. 2, p. 62-68, 1977.

HARGROVES, K. & SMITH, M. (Editors). *The Natural Advantage of Nations: Business Opportunities, Innovation and Governance in the 21st Century*. Hardback: Earthscan/James&James, 2005.

HUMBLE, J.; MOLESKY, J.; O'REILLY, B. *Lean Enterprise*. 1ª ed. O'Reilly Media, 2015.

ISMAIL, S; MALONE, M; VAN GEEST, Y. Organizações Exponenciais, Por que elas são 10 vezes melhores, mais rápidas e mais baratas que a sua (e o que fazer a respeito). São Paulo: Editora HSM, 2015.

JOHNSON, M. W. et al. HBR's 10 must read on business model innovation. São Paulo: Harvard Business Review Brasil, 2019.

KELLEY, T. *As 10 faces da inovação estratégias para turbinar a criatividade*; Rio de Janeiro: Campus/Elsevier, 2007.

MAGALDI, S.; Neto, J.S. *O Novo Código da Cultura: Vida ou Morte na Era Exponencial*. 1ª ed. Gente, 2019.

MAIA, N. B.; MATOS, L.; BARELLA, N. *Indicadores ambientais: Conceitos e aplicações*. São Paulo, Educ/Coped/INEP, 2001.

MEIRA, S. *Novos negócios inovadores de crescimento empreendedor no Brasil*. Casa da Palavra, 2013.

MELO, A; ABELHEIRA, R. *Design Thinking & Thinking Design: metodologia, ferramentas e uma reflexão sobre o tema*. São Paulo: Novatec, 2015.

MILES, I. The development of technology foresight: A review. *Technological Forecasting & Social Change*, 77. pp. 1448–1456, 2010.

MUELLER, A. C; GUIDO, S. *Introduction to Machine Learning with Python a Guide for Data Scientists*. New York, USA: O'Reilly Media, 2016.

OSTERWALDER, A.; PIG, Y. *Business Model Generation - inovação em modelos de negócios*. Rio Janeiro: Editora Alta Books, 2011.

REIS, D. *Gestão da Inovação Tecnológica*. 2ª ed. São Paulo: Manole 2018.

RIBEIRO, L. *Estudos avançados no marketing sustentável*. São Paulo: Pearson Education, 2017.

RIES, E. *A Startup Enxuta*. 1ª ed. Leya Brasil, 2012.

RIES, E. *The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses*. Crown Business, 2010.

ROAM, D. *Desenhando negócios*. Rio Janeiro: Editora Campus – RJ, 2011.

ROGER D. L.; SERRA, A. C. C. *Transformação digital: repensando o seu negócio para a era digital*. São Paulo: Editora Autêntica Business, 2017.

SCHWAB, K. *A quarta revolução industrial*. São Paulo: Editora Edipro, 2016.

SCHWABER, K; BEEDLE, M. *Agile Software Development with Scrum*. São Paulo: Prentice Hall, 2002.

STICKDORN, M. et al. *Isto é Design Thinking de serviços: fundamentos, ferramentas, casos*. Porto Alegre: Bookman, 2014.

SUTHERLAND, J; SCHWABER, K. *The Scrum Guide*. 11/2017. Disponível em: <https://www.scrumguides.org/docs/scrumguide/v2017/2017-Scrum-Guide-Portuguese-Brazilian.pdf>. Acesso em: 10/06/2019.

TAPPSCOT, D. *A hora da geração digital*. São Paulo: Editora Agir, 2010.

TEIXEIRA, T. *Startups e inovação: direito no empreendedorismo (entrepreneurship law)*. São Paulo: Manole, 2017.

TIDD, J.; BESSANE, J.; PAVITT, K. *Gestão da Inovação*. 3. ed. São Paulo: Bookman, 2008.

TIGRE, P. B. *Gestão da Inovação: a economia da tecnologia no Brasil*. Rio de Janeiro: Elsevier, 2006.

VIDAL, A. *Agile Think Canvas*. Rio de Janeiro: Editora Brasport, 2017.

WERKEMA, C. *Criando a Cultura Lean Seis Sigma*. 3. ed. Rio de Janeiro: Elsevier, 2012.

## **MÓDULO - EXPLORATORY ANALYTICS, PREDICTIVE ANALYTICS & MACHINE LEARNING**

### **Estruturas de Bancos de Dados, Tipos de Variáveis e Escalas de Mensuração**

Bancos de Dados; observações e variáveis; estruturas longitudinais e transversais; variáveis qualitativas e quantitativas; escalas de mensuração de variáveis; variáveis nominais e ordinais; variáveis métricas.

### **Introdução ao Machine Learning**

Machine learning e reconhecimento de padrões; conceito de supervisão e não supervisão; técnicas não supervisionadas e relação com escalas de mensuração de variáveis; técnicas supervisionadas e relação com escalas de mensuração de variáveis; casos reais e práticos.

### **Unsupervised Machine Learning - Clustering**

Análise de conglomerados e de agrupamentos; padronização de variáveis e procedimento z-scores; medidas de distância e de similaridade; clusters hierárquicos; clusters não hierárquicos k-means; esquemas de aglomeração; dendrograma; escalonamento multidimensional; aplicações de bancos de dados reais de mercado em R.

### **Unsupervised Machine Learning - Análise Fatorial e PCA**

Análise fatorial; principal componente analysis (PCA); construção de fatores; matriz de correlações; escores e cargas fatoriais; mapa de cargas; criação de rankings em datasets; aplicações de bancos de dados reais de mercado em R.

### **Unsupervised Machine Learning - Análise de Correspondência Simples e Múltipla**

Tabelas de frequência; massas de frequências; resíduos frequentistas; teste chi-quadrado; associação entre variáveis e entre categorias; coordenadas de categorias e mapas perceptuais; clusterização e rankings de observações a partir de variáveis qualitativas; aplicações de bancos de dados reais de mercado em R.

### **Supervised Machine Learning - Análise de Regressão Simples e Múltipla**

Mapa de correlações; modelos de regressão simples; estimação por mínimos quadrados; coeficiente de ajuste  $R^2$ ; teste F, testes t, modelos de regressão múltipla; procedimento Stepwise; multicolinearidade; teste de Shapiro-Francia; procedimento de Box-Cox; modelos não lineares; intervalos de confiança e previsão; aplicações de bancos de dados reais de mercado em Excel e em R.

### **Supervised Machine Learning - Modelos Logísticos Binários e Multinomiais**

Conceitos de probabilidade e chance de ocorrência de eventos; curva sigmoide S e modelo probabilístico; estimação por máxima verossimilhança; modelos de risco de crédito e de probabilidade de default; intervalos de confiança e previsão; conceito de cutoff; confusion matrix; análise de sensibilidade; indicadores de acurácia, sensibilidade e especificidade; curva ROC; índice GINI; modelos multinomiais; curvas spline; aplicações de bancos de dados reais de mercado em Excel e em R.

### **Supervised Machine Learning - Modelos para Dados de Contagem**

Dados de contagem; conceito de exposição; estimação por máxima verossimilhança; modelo Poisson; modelo binomial negativo; overdispersion; teste de Cameron e Trivedi; intervalos de confiança e previsão; modelagem para eventos raros; zero-inflated models; aplicações de bancos de dados reais de mercado em Excel e em R.

### **Supervised Machine Learning - Modelagem Multinível**

Generalized linear multilevel models (GLMM); conceitos e definições; caracterização de aninhamento e agrupamento nos dados; hierarquias e contextos em modelagem de dados; efeitos contextuais; efeitos aleatórios de intercepto e de inclinação entre contextos; vantagens dos modelos multinível e incremento real de desempenho preditivo; modelos hierárquicos lineares (HLM); inserção correta de variáveis individuais e contextuais em modelos multinível; aplicações de bancos de dados reais de mercado em R.

### **Outros Modelos de Machine Learning**

Árvores de decisão; bagging, boosting; random forest; redes neurais; técnicas de validação; treino, validação e teste, out-of-time, overfitting, validação cruzada; aplicações de bancos de dados reais de mercado em R.

### **Bibliografia do módulo**

- AGRESTI, A. Categorical data analysis. 3. ed. Hoboken: John Wiley & Sons, 2013.
- ALBERGARIA, M.; FÁVERO, L. P. Narrow replication of Fisman and Miguel's (2007a) 'Corruption, norms, and legal enforcement: evidence from diplomatic parking tickets'. *Journal of Applied Econometrics*, v. 32, n. 4, p. 919-922, 2017.
- ALBUQUERQUE, J. P. A.; FORTES, J. M. P.; FINAMORE, W. A. Probabilidade, variáveis aleatórias e processos estocásticos. Rio de Janeiro: Interciência, 2008.
- ALCALDE, A.; FÁVERO, L. P.; TAKAMATSU, R. T. EBITDA margin in Brazilian companies: variance decomposition and hierarchical effects. *Contaduría y Administración*, v. 58, n. 2, p. 197-220, 2013.
- BAKKE, H. A.; LEITE, A. S. M.; SILVA, L. B. Estatística multivariada: aplicação da análise fatorial na engenharia de produção. *Revista Gestão Industrial*, v. 4, n. 4, p. 1-14, 2008.
- BALAKRISHNAN, P. V.; COOPER, M. C.; JACOB, V. S.; LEWIS, P. A. A study of the classification capabilities of neural networks using unsupervised learning: a comparison with k-means clustering. *Psychometrika*, v. 59, n. 4, p. 509-525, 1994.
- BOLFARINE, H.; BUSSAB, W. O. Elementos de amostragem. São Paulo: Edgard Blücher, 2005.
- BOLFARINE, H.; SANDOVAL, M. C. Introdução à inferência estatística. Rio de Janeiro: Sociedade Brasileira de Matemática, 2001.
- BONETT, D. G. Varying coefficient meta-analytic methods for alpha reliability. *Psychological Methods*, v. 15, n. 4, p. 368-385, 2010.
- CHARNET, R.; BONVINO, H.; FREIRE, C. A. L.; CHARNET, E. M. R. Análise de modelos de regressão linear: com aplicações. 2. ed. Campinas: Editora da UNICAMP, 2008.
- CHATTERJEE, S.; JAMIESON, L.; WISEMAN, F. Identifying most influential observations in factor analysis. *Marketing Science*, v. 10, n. 2, p. 145-160, 1991.
- CHEN, M. H.; IBRAHIM, J. G.; SHAO, Q. M. Maximum likelihood inference for the Cox regression model with applications to missing covariates. *Journal of Multivariate Analysis*, v. 100, n. 9, p. 2018-2030, 2009.
- DEMIDENKO, E. Mixed models: theory and applications. New York: John Wiley & Sons, 2005.
- DEUS, J. E. R. Escalamiento multidimensional. Madrid: Editorial La Muralla, 2001.
- DEVORE, J. L. Probabilidade e estatística para engenharia. São Paulo: Thomson Pioneira, 2006.
- ESPINOZA, F. S.; HIRANO, A. S. As dimensões de avaliação dos atributos importantes na compra de condicionadores de ar: um estudo aplicado. *Revista de Administração Contemporânea (RAC)*, v. 7, n. 4, p. 97-117, 2003.
- EVERITT, B. S.; LANDAU, S.; LEESE, M.; STAHL, D. Cluster analysis. 5. ed. Chichester: John Wiley & Sons, 2011.

FABRIGAR, L. R.; WEGENER, D. T.; MacCALLUM, R. C.; STRAHAN, E. J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, v. 4, n. 3, p. 272-299, 1999.

ESPINOZA, F. S.; HIRANO, A. S. As dimensões de avaliação dos atributos importantes na compra de condicionadores de ar: um estudo aplicado. *Revista de Administração Contemporânea (RAC)*, v. 7, n. 4, p. 97-117, 2003.

EVERITT, B. S.; LANDAU, S.; LEESE, M.; STAHL, D. *Cluster analysis*. 5. ed. Chichester: John Wiley & Sons, 2011.

FABRIGAR, L. R.; WEGENER, D. T.; MacCALLUM, R. C.; STRAHAN, E. J. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, v. 4, n. 3, p. 272-299, 1999.

HAIR JR., J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. *Análise multivariada de dados*. 6. ed. Porto Alegre: Bookman, 2009.

HALL, D. B. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*, v. 56, p. 1030-1039, 2000.

LANDAU, S.; EVERITT, B. S. *A handbook of statistical analyses using SPSS*. Boca Raton: Chapman & Hall / CRC Press, 2004.

LANE, W. R.; LOONEY, S. W.; WANSLEY, J. W. An application of the Cox proportional hazards model to bank failure. *Journal of Banking & Finance*, v. 10, n. 4, p. 511-531, 1986.

LINOFF, G. S.; BERRY, M. J. A. *Data mining techniques: for marketing, sales, and customer relationship management*. 3. ed. Indianapolis: John Wiley & Sons, 2011.

LISBOA, E.F.A. *Pesquisa operacional*. Disponível em . Acesso em 28 de Setembro de 2010.

LOMBARDO, R.; BEH, E. J.; D'AMBRA, L. Non-symmetric correspondence analysis with ordinal variables using orthogonal polynomials. *Computational Statistics & Data Analysis*, v. 52, p. 566-577, 2007.

LONG, J. S.; FREESE, J. *Regression models for categorical dependent variables using Stata*. 2. ed. College Station: Stata Press, 2006.

PAMPEL, F. C. *Logistic regression: a primer*. Thousand Oaks: Sage Publications, 2000.

PARDOE, I. *Applied regression modeling*. 2. ed. Hoboken: John Wiley & Sons, 2012.

RAYKOV, T.; MARCOULIDES, G. A. *An introduction to applied multivariate analysis*. New York: Routledge, 2008.

REIS, E. *Estatística multivariada aplicada*. 2. ed. Lisboa: Edições Sílabo, 2001.

RIGAU, J. G. Traducción del término 'odds ratio'. *Gaceta Sanitaria*, v. 16, p. 35, 1990.

RODRIGUES, M. C. P. Potencial de desenvolvimento dos municípios fluminenses: uma metodologia alternativa ao IQM, com base na análise fatorial exploratória e na análise de clusters. *Caderno de Pesquisas em Administração*, v. 9, n. 1, p. 75-89, 2002.

RODRIGUES, P. C.; LIMA, A. T. Analysis of an European union election using principal component analysis. *Statistical Papers*, v. 50, n. 4, p. 895-904, 2009.

SANTOS, M. A.; FÁVERO, L. P.; DISTADIO, L. F. Adoption of the International Financial Reporting Standards (IFRS) on companies' financing structure in emerging economies. *Finance Research Letters*, v. 16, n. 1, p. 179-189, 2016.

SAPORTA, G. *Probabilités, analyse des données et statistique*. Paris: Technip, 1990.

SARAIVA Jr, A. F.; TABOSA, C. M.; COSTA, R. P. Simulação de Monte Carlo aplicada à análise econômica de pedido. *Produção*, v. 21, n. 1., p.149-164, 2011.

TURKMAN, M. A. A.; SILVA, G. L. *Modelos lineares generalizados: da teoria à prática*. Lisboa: Edições SPE, 2000.

VALENTIN, J. L. *Ecologia numérica: uma introdução à análise multivariada de dados ecológicos*. 2. ed. Rio de Janeiro: Interciência, 2012.

VIEIRA, S. *Estatística básica*. São Paulo, Cengage Learning, 2012.

VITTINGHOFF, E.; GLIDDEN, D. V.; SHIBOSKI, S. C.; McCULLOCH, C. E. *Regression methods in biostatistics: linear, logistic, survival, and repeated measures models*. 2. ed. New York: Springer-Verlag, 2012.

- VUONG, Q. H. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, v. 57, n. 2, p. 307-333, 1989.
- WU, Z. et al. Optimization designs of the combined Shewhart CUSUM control charts. *Computational Statistics and Data Analysis*, v. 53, n. 2, p. 496-506, 2008.
- WULFF, J. N. Interpreting results from the multinomial logit: demonstrated by foreign market entry. *Organizational Research Methods*, v. 18, n. 2, p. 300-325, 2015.
- XIE, F. C.; WEI, B. C.; LIN, J. G. Assessing influence for pharmaceutical data in zero-inflated generalized Poisson mixed models. *Statistics in Medicine*, v. 27, n. 18, p. 3656-3673, 2008.
- XIE, M.; HE, B.; GOH, T. N. Zero-inflated Poisson model in statistical process control. *Computational Statistics & Data Analysis*, v. 38, n. 2, p. 191-201, 2001.

**MBAUSP**  
**ESALQ**

Rua Alexandre Herculano, 120, Vila Monteiro  
Piracicaba/SP CEP: 13.418-445

[queroinfos@pecege.com](mailto:queroinfos@pecege.com)

+55 19 3377 0937 +55 19 3377 0940

    /mbauspesalq