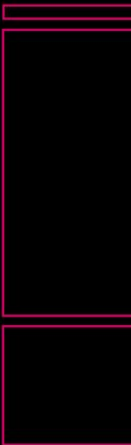




# PROJETO PEDAGÓGICO DO CURSO

MBA ON-LINE EM DATA SCIENCE



# S U M Á R I O

<b>ORGANIZAÇÃO DIDÁTICO-PEDAGÓGICA</b>	<b>3</b>
Projeto Pedagógico do Curso: aspectos gerais	3
Objetivos do Curso	6
Tese de Transformação do Curso	7
Perfil do Egresso	8
Mercado de Trabalho	9
Metodologias Inovadoras	9
Conexão entre as fases e disciplinas	13
Competências e Ferramentas	13
Matriz Curricular	14
Ementas e Bibliografias	16
Design Experience FIAP	44
Processo de Avaliação	44
Projeto Integrador – Startup One MBA FIAP ON	46

# ORGANIZAÇÃO DIDÁTICO-PEDAGÓGICA

## Projeto Pedagógico do Curso: aspectos gerais

### Contexto educacional

O **Centro Universitário FIAP** é uma Instituição de Ensino Superior com atuação principal nos eixos de tecnologia, gestão e inovação. Inserido fisicamente na região com maior densidade tecno-econômica do país, a Região Metropolitana de São Paulo tem mais de 22 milhões de habitantes<sup>1</sup>, e possui relevância e liderança nacional no desenvolvimento da economia nacional.

A Região concentra a maioria das sedes de empresas brasileiras dos mais importantes complexos industriais, comerciais e financeiros, o que cria também grande demanda por profissionais qualificados nas principais áreas de competências e habilidades da Nova Economia.

Nesta Nova Economia, onde os dados são gerados por todos os lados, seja nas redes sociais, nas transações efetuadas, nas localizações geográficas, seja na Internet das Coisas, Web 3.0 e blockchain, é necessário associar inteligência competitiva com novos recursos analíticos para a organização continuar vibrante no seu ecossistema. Um olhar para os dados transforma-os em um novo ativo valioso para as organizações e para a economia como um todo. São usados dados para criar serviços e produtos para fins econômicos.

A Análise de Dados é uma das áreas que mais crescem na área da Tecnologia da Informação. Analytics e Business Intelligence fazem parte deste contexto de busca e uso da informação para tomada de decisões gerenciais. Mais e mais empresas estão aderindo com sucesso ao processo de Transformação Digital e as empresas com maior sucesso no mercado são

---

<sup>1</sup> IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Estimativa da População 2021**. Rio de Janeiro: IBGE, 2021.

aquelas que sabem lidar com os dados. A Ciência de Dados ocupa um lugar de destaque no mundo da tomada de decisões, pois as empresas estão migrando de um processo baseado no sentimento das pessoas para decisões baseadas em dados.

O Cientista de Dados é o profissional responsável por ingerir os dados no ambiente de Big Data, transformar e integrar os dados para torná-lo compreensível para os modelos matemáticos e estatísticos que serão utilizados no processo de tomada de decisões.

Segundo o site Statista.com, o mercado mundial de Data Science em 2022 será de US\$ 274 bilhões de dólares e mantém uma expectativa de crescimento de quase 25% ao ano até 2026, de acordo com a CISION PR Newswire.

Há 10 anos, a **Harvard Business Review** disse que o futuro do Data Science seria superaquecido, sendo a profissão mais promissora do século 21 e corroborando com a estimativa de que os EUA teriam de 140 a 190 mil cientistas até 2019.

Essa previsão, um pouco ousada e bem acima da realidade, se mostra coerente e correta. Nos últimos anos, cada vez mais empresas adotaram a **cultura data-driven**, por conta da explosão na geração e utilização de dados que acontecem todos os dias na internet.

Mais da metade da população do planeta possui acesso ao meio digital e, com isso, a geração de informações relevantes de todas as áreas de atuação cresceu de maneira exponencial, fazendo com que um termo muito utilizado no dia a dia de dados surgisse: **big data**, que representa o alto volume de dados, a grande variedade e a velocidade presente nisso tudo.

Este curso está, portanto, adequado ao mercado de trabalho regional e ao perfil das organizações empregadoras na região, no Brasil e até no exterior. As condições econômicas e sociais de São Paulo são indicadores positivos para a existência de uma instituição de ensino como a FIAP e, especificamente, para a proposição do MBA em Ciência de Dados.

Segundo Ricardo Villate, vice-presidente da IDC Latin America, o maior investimento do setor de TI será em tecnologias denominadas “pilares para a

terceira plataforma”, como Cloud, Big Data/Analytics, Data Science e Mobilidade e Empreendimento social, que concentrarão 58% dos investimentos.

De acordo com a empresa de recrutamento Michael Page (<https://www.michaelpage.com.br/advice/lideran%C3%A7a-e-gest%C3%A3o/integra%C3%A7%C3%A3o-e-engajamento/perspectivas-para-o-mercado-de-ti-em-2021>), a área de dados, é uma das que tem chamado mais atenção na busca por profissionais, dado o alto crescimento para Business Intelligence, Data Science e Analytics. A importância dos dados para o crescimento estratégico e tomada de decisão é tão grande que essa área é a grande aposta para os próximos dois anos. Muitos profissionais do meio acadêmico e de pesquisa, como matemáticos, físicos e estatísticos, estão migrando para o corporativo. Análises da empresa Glassdoor sobre médias salariais de um profissional em ciência de dados variam de R\$ 8.400,00 a R\$ 19.190,00 reais mensais. Isso comprova o aquecimento e justificativa na formação e capacitação de profissionais especialistas para essa área.

O objetivo do curso justifica-se, principalmente, na concepção da importância da análise, descoberta e predição/tratamento de grandes volumes de dados disponíveis em todas as organizações a fim de melhorar a tomada das decisões corporativas, inclusive antecipando tendências e cenários. Este curso possui um viés técnico específico que mescla Matemática Aplicada com Tecnologia e leva o aluno a visualizar oportunidades de criação de Modelos e cenários em Ciência de Dados, utilizando os dados para análises avançadas, incluindo previsão de cenários futuros.

O curso propõe-se a contribuir para a qualificação dos profissionais da área de Ciência de Dados, ampliando sua parcela de participação como agente transformador e reforçando seu comprometimento, principalmente, na cidade de São Paulo e região metropolitana, mas, também, expandindo sua atuação dentro e fora do Brasil.

Nesta modalidade de curso on-line, os alunos realizam atividades extracurriculares como participação em webinars, atividades em grupo nas aulas ao vivo, leitura de artigos para discussão – com uso da metodologia ativa:

Sala de Aula Invertida. Essa modalidade nos permite maior flexibilidade para disponibilizar conteúdos que podem ser consumidos a qualquer momento pelo aluno, respeitando a velocidade, interesse e dedicação individual. Outra grande vantagem é a possibilidade de trabalhar um projeto junto com empresas parceiras que fornecem casos reais para serem resolvidos pelos alunos. Com essa técnica de imersão, o aluno realiza todas as etapas de um processo de Ciência de Dados, desde a concepção e compreensão do problema, definição da arquitetura de solução, criação e operacionalização dos algoritmos de machine learning e inteligência artificial. As mentorias técnicas e de negócio ajudam os alunos a selecionar os conteúdos mais adequados para solução do caso real apresentado e estimula a criatividade e inovação.

### **Cenário Futuro**

Atualmente as empresas e os profissionais de Tecnologia da Informação têm de lidar com grandes massas de dados variados e díspares, e isso só reforça a necessidade de o mercado ter profissionais capacitados para trabalhar com essa demanda massiva. Com a imposição da transformação digital às empresas e o direcionamento cada vez mais voltado para corporações Data-Driven, com necessidades de compreensão do seu universo de dados, geração de visões preditivas e diagnósticas; o profissional Cientista de Dados enfrentará grandes desafios em tratamento de massas gigantescas de dados, tendo como preocupação e responsabilidade, a performance, a governança, ética, aquisição, tratamento e descoberta de insights sobre esse universo variado de informações.

### **Objetivos do Curso**

#### **Objetivo geral:**

O curso tem como objetivo geral formar cientistas de dados com foco analítico capazes de desenvolver soluções técnicas para resolução de problemas complexos e que seja responsável por coletar, gerenciar e transformar em modelos utilizáveis uma grande quantidade de dados não-

estruturados, para que seja possível extrair desse conjunto informações relevantes.

### **Objetivos específicos:**

Com foco em matemática aplicada e tecnologia, este curso abrirá um mundo de possibilidades ao aluno, preparando o profissional para:

- Compreender os modelos estatísticos e analíticos trazidos pela matemática para o suporte e construção de modelos precisos.
- Dar suporte e tomada de decisão, com auxílio de modelos estatísticos selecionados que expliquem um problema ou uma oportunidade de negócio e as incertezas associadas.
- Extrair, organizar e combinar dados, estruturados ou não, numéricos ou textuais, para a modelagem estatística do problema ou da oportunidade de negócio.
- Entender como um ambiente de ecossistema de dados se integra ao mundo da ciência de dados, trazendo benefícios a todo tipo de indústria ou segmento de mercado.
- Explorar, por meio de práticas e cases, os usos, as boas práticas e os melhores modelos em ciência de dados para as principais indústrias.

### **Tese de Transformação do Curso**

O MBA foi dividido didaticamente por fases para proporcionar uma experiência cumulativo-gradativa de aprendizagem, na qual, alinhada às técnicas e metodologias ativas de ensino, proporcionaremos aos alunos uma vivência nos ambientes políglotas que o Cientista de Dados precisará dialogar. Em cada fase o aluno será envolvido ativamente em todos os processos estatísticos, como matemática estatística simples, a definição de arquiteturas híbridas (sejam elas relacionais ou NoSQL), os ambientes atuais de implementação física (OnPrem – Hadoop ou em Cloud), a geração e o trato massivo de dados e, por fim, o enfoque em ciência de dados no que tange ao uso e aplicação de técnicas KDD (Data Mining) e algoritmos em Machine

Learning e Deep Learning. Durante todo o MBA, o aluno será exposto aos métodos de PBL (Problem Based Learning), EBL (Experience Based Learning) e PjEL (Project Base Learning). Acreditamos que essas metodologias ativas, nas quais o aluno se torna parte ativa de seu aprendizado, bem como as experiências em ambientes próximos da realidade e completamente hands-on prepararão o aluno para vencer os desafios trazidos pela cultura data-driven.

## Perfil do Egresso

Com foco em tecnologia e em questões sobre a gestão de negócios, este curso vai abrir diversas possibilidades de alocação profissional. Os alunos estarão aptos a:

- Transformar a vivência corporativa e experiência profissional em soluções de Ciência de Dados.
- Conhecimentos sólidos em estatística descritiva: medidas de tendências, dispersão, histogramas.
- Domínio de linguagens de programação estatística como R e Python.
- Criação e Arquitetura de Banco de Dados relacionais e não relacionais para apoiar soluções de Ciência de Dados.
- Definição de Arquitetura de Dados em Cloud e On-prem para apoiar projetos de Ciência de Dados.
- Aplicar modelos estatísticos diversos para criar modelos preditivos e prescritivos ampliando o uso dos dados nas corporações.
- Definir arquitetura de Operações para esteira de dados (DataOps) e algoritmos estatísticos (MLOps).
- Aplicar técnicas de Governança de Dados para manter a rastreabilidade dos dados e atender os requisitos da LGPD.
- Implementar o processo de ingestão de dados, incluindo curadoria, catálogo e integração utilizando técnicas de EL, ELT e ETL.



## Mercado de Trabalho

O curso foi estruturado com o objetivo de formar os seguintes profissionais para o mundo corporativo:

- Aprofundamento na visão da matemática aplicada que é essencial para o bom funcionamento do cientista de dados.
- Extração, tratamento e trato em grandes volumes de dados para análise;
- Realização de testes de hipóteses, inferência, regressões, conglomerados, análise de associações e correlações sobre os dados;
- Aplicação de técnicas e métodos avançados de machine learning sobre os dados;
- Identificação e aplicação de técnicas de modelos de deep learning sobre os dados;
- Trabalho com redes neurais e convolucionais, na criação de modelos preditivos e prescritivos;
- Compreender como a ciência de dados se dá nas principais indústrias do mercado.

## Metodologias Inovadoras

Todos os cursos do MBA On são entregues por meio de uma jornada composta por 5 fases.

A concepção da jornada é inspirada em PBL (Project Based Learning) que é uma metodologia ativa onde problemas reais são a base do processo de aprendizagem, por isso, a composição das fases nasce da junção de disciplinas da grade curricular do curso que demonstram sinergia entre si, proporcionando ao nosso aluno a aplicação direta dos conteúdos apresentados. Cada fase reúne arcabouço teórico e um conjunto de ferramentas para resolução de problemas que são/serão vivenciados pelos nossos alunos em sua vida profissional real, equilibrando soft e hard skills necessários para formar líderes protagonistas do futuro.

Os conteúdos entregues via plataforma são criados exclusivamente para os alunos do MBA ON e para que atendam todos os requisitos necessários

(atualização, informação relevante de mercado, base teórica consistente e uso de linguagem e forma amigáveis) um time de especialistas participa desde a concepção da ideia até a revisão final.

- Coordenador de curso – identificação da necessidade do conteúdo e definição da ementa e busca de profissionais no mercado.
- Conteudista – responsável pela escrita dos materiais equilibrando base teórica sólida e cases de mercado.
- Professores especialistas – gravação de vídeos em formatos que variam de acordo com o perfil de consumo da persona do curso: videoaula, podcast, talks (TED), painéis de discussão, etc.
- Profissionais de mercado – gravação de vídeos com a visão das empresas em formatos de cases.

Todo projeto de criação de conteúdo é acompanhado e validado pelo coordenador de curso que conta com o apoio do time de professores (mentores) de cada curso. Por isso, é tão importante que além de formação acadêmica, nosso time acadêmico tenha experiência de mercado também.

A partir do desenho da persona do curso, e da geração de indicadores que monitoram os acessos aos mais variados formatos de conteúdo, conseguimos desenhar uma experiência de consumo personalizada para cada um de nossos MBAs ON (no momento da produção do conteúdo). Atualmente contamos com 4 formatos que podem ser combinados entre si: HTML, vídeo, áudio e PDF.

Os quatro formatos se complementam, e trazem ênfases diversificadas. Para os cursos com maior enfoque em *business* temos aulas com aplicações práticas para formatação de modelos/estratégias de negócio e cases para fornecer *benchmarking* aos nossos alunos. A ideia é que eles aprendam com profissionais que já erraram e acertaram na prática, e que hoje são referências no mercado.

Já em nossos cursos técnicos, a ênfase está no hands on. Mostramos como fazer, fazendo.

Os podcasts geralmente são utilizados para trazer informações relevantes do mercado discutindo boas práticas e experiência de carreira.

Para os alunos que ainda preferem estudar de maneira mais tradicional, temos também o formato PDF que organiza o conteúdo por meio dos textos e imagens, formando uma apostila que pode ser baixada, e acessada off-line.

Existem ainda cursos em que percebemos uma maior apreciação dos alunos, por exemplo, por vídeos em detrimento à leitura de textos, nesses casos, o número de páginas diminui, e o total de horas de vídeo aumentam.

Hoje, além de todo o conteúdo digital, entregamos aproximadamente 80 horas de aulas ao vivo nas 5 fases e no Startup One.

As aulas apresentam formatos diferentes que vão desde aulas de experimentação focadas em hands on, dinâmicas e aplicação de conteúdos, bate-papos com profissionais de mercado até mentorias com especialistas das empresas parceiras. Todas as aulas são gravadas e disponibilizadas na plataforma para consumo, caso o aluno não tenha conseguido participar ao vivo.

O Solution Sprint é adotado nas fases 1 a 4 dos cursos como forma de avaliação e experimentação de conteúdo. Nesse formato, convidamos uma empresa parceira para trazer uma dor que será a base do desafio entregue ao aluno, para que ele explorando o conteúdo da fase, crie soluções. Durante o processo de criação de soluções, os alunos recebem mentoria das empresas nas aulas ao vivo agendadas e a validação dessa solução com visão de mercado é feita pelas empresas por meio de *pitches* ou teste de entregáveis no caso dos cursos mais técnicos (validação técnica de aplicações, modelos analíticos, microsserviços, cloud etc).

Por meio dessa metodologia, a retenção e aplicação de conceitos por parte de aluno é avaliada pelo nosso time acadêmico e a validação das soluções pelo mercado é fornecida pelos nossos parceiros em forma de *feedback* ao vivo proporcionando ao aluno a experiência e segurança necessária para que ele realmente replique as soluções no seu cotidiano profissional.

Na fase 5 o aluno já possui conhecimento e maturidade suficiente para atuar em desafios mais complexos e é aí que entram os *challenges*, que reúnem os conteúdos de todo o curso.

Por serem mais complexos e exigirem uma dedicação maior do aluno, além de ser o fechamento da experiência do curso, as empresas parceiras e a FIAP oferecem premiação para os melhores projetos.

A escolha dos melhores projetos é realizada pelos parceiros e a avaliação acadêmica é realizada pelo nosso time. Assim, garantimos que nosso aluno foi munido de base teórica consistente e desenvolveu competências em resolução de problemas. Com isso, pretendemos que ele se torne um profissional capaz de analisar novos cenários e entregar soluções, sendo protagonista da sua carreira.

É importante ressaltar que para o público de MBA, a avaliação precisa ser parte do processo de aprendizado para entregar conhecimento ao aluno valorizando o tempo investido por ele na jornada do curso.

Dentro do portfólio do MBA ON temos cursos com públicos bem diferentes, que demandam não só formatos diferentes de conteúdo, mas também dinâmicas avaliativas e experiências customizadas.

Como variação dos challenges de final de curso atualmente temos os hackathons, CTFs (Capture de Flag – uma competição de conhecimentos muito valorizada pelo mercado de cyber security) e os Datathons.

Durante todo o curso, nossos alunos contam com mentorias que podem ser agendadas sob demanda com nossos professores especialistas, além do contato direto com o time de professores e coordenador do curso, disponíveis sempre que precisarem.

Diante dos diversos formatos de conteúdo e interações que oferecemos nos MBAs ON, nosso aluno pode escolher a forma que mais se adequa ao seu perfil para se conectar com professores, colegas e profissionais de mercado, criando sua própria rotina de estudos, no seu ritmo e dentro das suas necessidades.

## Conexão entre as fases e disciplinas

Por meio do uso de metodologias ativas e do processo de ensino-aprendizagem PBL (Project Based Learning), procuramos manter uma grande integração entre as disciplinas e as fases, a fim de que o aluno tenha consciência, experiência e vivência de ambientes reais onde um cientista de dados atuaria, percorrendo as principais áreas desse ambiente.

O programa combina conceitos e fundamentos do mundo corporativo com metodologias profissionais que vão mudar a forma de capturar e interpretar dados.

O aluno vai aprender, na prática, a tomar as melhores decisões de algoritmos descritivos, preditivos e prescritivos, com base nos dados disponíveis na organização.

Ele vai testar seus conhecimentos em cases e situações reais de mercado, utilizando técnicas de aprendizagem com base em problemas reais (PBL).

Nosso programa aborda as mais modernas tecnologias de ciência de dados: Databricks, Snowflake, trifacta, jupiter Notebooks, Bases NoSQL (Cassandra, Redis, MongoDB, Neo4J), assim como metodologias, estratégias e modelos de negócios inovadores voltados para o cientista de dados.

## Competências e Ferramentas

Este curso vai preparar, em termos de hard e soft skills, o profissional para:

- implementar áreas de armazenamento de dados com foco em ciência de dados, tanto de forma técnica quanto administrativa;
- dominar temas como multipluralidade tecnológica, sistemas híbridos, ingestão de diversos tipos de dados, além da organização e do uso efetivo desse ambiente;

- definir os diferentes tipos de Arquitetura de Dados, uso de infraestruturas diversas (on premise e cloud) e aplicação de diversos tipos de algoritmos de inteligência de negócios no ambiente.
- desenvolver uma visão pós-sistêmica, ampliando formas de análise e competitividade com base nos dados internos e externos da organização.
- resolver problemas complexos de forma ágil, simples e efetiva.
- utilizar ferramentas e metodologias para apresentação de dados e geração de insights por meio de dashboards e storytellings.
- desenvolver o pensamento crítico e criativo.
- praticar técnicas de comunicação, apresentação e negociação.

Para alcançarmos esses objetivos, utilizaremos ferramentas diversas e atuais dentro do programa de disciplinas, tais como:

- Soluções Big Data (AWS, Azure, HDFS),
- Bases NoSQL (Redis, MongoDB, Cassandra, Neo4J),
- Soluções de Ingestão (Kafka, NiFi, Trifacta),
- Armazenamento de dados (S3, Bases Relacionais),
- Programas para ciência de dados (R, RStudio, Python, JupiterNotebooks), Tensorflow,
- Ferramentas OpenSource para Cientistas de Dados, MLOps solutions OpenSource (DataRobot, RapidMiner, H2O.io).
- Cloud Lakes.
- Ferramental Analítico (PowerBI, Tableau).

## Matriz Curricular

MATRIZ CURRICULAR	
Data Architecture	16h
Data Modeling	16h
Data Integration & Ingestion	16h

Database for Data Science	16h
Statistics Fundamentals	20h
Machine Learning	20h
Deep Learning	20h
Recomendation Systems	16h
Linear Algebra for Data Scientists	12h
R for Data Scientists	16h
Pyhton for Data Scientists	16h
Data Ecosystems – Cloud Infrastructure	16h
Feature Engineering	16h
DataOps & Governance	16h
MLOps	12h
Text & Social Media	20h
Reinforcement Learning	20h
Human-based Data	16h
Data Driven Organizational Strategies	12h
Data Science for Business	12h
Business Analytics	16h
Empreendedorismo e Inovação	20h
<b>CARGA HORÁRIA TOTAL DO CURSO</b>	<b>360h</b>



## Ementas e Bibliografias

Disciplina	Data Architecture
<b>Ementa</b>	
<p>Tem por objetivo prover ao aluno conhecimentos sobre as arquiteturas de dados utilizadas nos diversos tipos de sistemas transacionais e analíticos utilizados hoje em ambientes diversos. E como as diferentes abordagens permitem maior desempenho e eficiência em operações de consulta, leitura, gravação e integrações sistêmicas. O aluno entenderá dentro do contexto de sistemas transacionais o uso de sistemas gerenciadores, arquitetura de banco de dados e sua evolução por meio de tabelas, atributos, chaves estrangeiras e demais constraints, índices, particionamento e linguagem SQL.</p> <p>Sob o aspecto de ciência de dados o aluno será apresentado aos conceitos de Data Warehouse / Data Mart, suas diferenças em relação ao modelo transacional e a justificativa de negócio para sua aplicação.</p> <p>Complementando a disciplina, serão expostos conceitos de Data Lake, Lake House e aplicação e NoSQL.</p> <ul style="list-style-type: none"><li>• Conceitos de Data Warehouse / Data Mart.</li><li>• Principais diferenças em relação ao modelo transacional e a justificativa de negócio para sua aplicação.</li><li>• Conceitos de Data Lake.</li><li>• Lake House.</li><li>• Introdução a NoSQL.</li></ul>	
<b>Bibliografia Básica</b>	
<p>COUGO, P. <b>Modelagem conceitual e projeto de bancos de dados</b>. Rio de Janeiro: Campus, 2009.</p> <p>WHITE, T. <b>Hadoop - The definitive guide</b>. EUA: O'Reilly, 2016.</p>	



### Bibliografia Complementar

ROSS, R.; KIMBALL, M. **The data warehouse toolkit**: guia completo para modelagem dimensional. Rio Janeiro: Campus, 2002.

SADALAGE, P. J.; FOWLER, M. **NOSQL Essencial**. São Paulo: Novatec, 2013.

INMON, B.; LEVINS, M.; SRIVASTAVA, R. **Building the Data Lakehouse**, Estados Unidos, Teachincs Publications, 2020.

### Disciplina

### Data Modeling

### Ementa

Oferecer ao profissional de dados recursos para escolher a estrutura de dados mais adequada à situação sistêmica encontrada, assim como oferecer conhecimentos estruturais para criar modelos entidade-relacionamento (transacional e dimensional) e modelos semiestruturados (JSON).

Serão apresentados conceitos de modelagem por meio de entidades, atributos, relacionamentos e documentos com uma abordagem prática por meio do Oracle Data Modeler e Hackolade.

- Conceitos de modelagem por meio de entidades, atributos, relacionamentos.
- Modelagem de documentos.
- Oracle Data Modeler e Hackolade.

### Bibliografia Básica

COUGO, P. **Modelagem conceitual e projeto de bancos de dados**. Rio de Janeiro: Campus, 2009.

ROSS, R.; KIMBALL, M. **The data warehouse toolkit**: guia completo para modelagem dimensional. Rio Janeiro: Campus, 2002.

### Bibliografia Complementar

SADALAGE, P. J.; FOWLER, M. **NOSQL Essencial**. São Paulo: Novatec, 2013.

## Disciplina Data Integration & Ingestion

### Ementa

Oferecer uma ampla visão sobre as melhores práticas de arquiteturas de ingestão, integração e processamento de dados por meio de conceitos de ETL, ELT, cargas em lote (batch), streaming de dados (integração em tempo real), CDC e ingestão massiva.

Além disso, apresentação de ferramentas de integração como Knime, Sqoop, Flume, NiFi, Kafka, Oracle Golden Gate, Azure Data Factory, GCP Dataflow, etc.

- ETL, ELT, cargas em lote (batch).
- Streaming de dados (integração em tempo real).
- CDC.
- Introdução à ferramentas de integração como Knime, Sqoop, Flume, NiFi, Kafka, Oracle Golden Gate, Azure Data Factory, GCP Dataflow).

### Bibliografia Básica

KIMBALL, R; CASERTA, J. **The Data Warehouse ETL Toolkit**: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data. Estados Unidos: Editora Wiley, 2011.

NERY RODRIGUES MACHADO, F. **Tecnologia e Projeto de Data Warehouse – Uma Visão Multidimensional**. São Paulo: Editora Érica, 2013.

### Bibliografia Complementar

N/C.

### Disciplina

### Statistics Fundamentals

### Ementa

Tem como objetivo fazer com que o aluno se debruce sobre os temas relacionados ao mundo estatístico com enfoque direto em análise descritiva e preditiva, abordando temas essenciais como os métodos associativos, correlacionais, regressivos etc. Além de permitir ao aluno o embasamento de tais métodos por meio de fórmulas matemáticas.

Entrada no mundo estatístico:

- Conceitos Básicos de Estatística.
- Medidas de Tendência Central e Dispersão.
- Gráficos Estatísticos.
- Análise Descritiva & Outliers.
- Percentis e a Frequência Acumulada.
- Análise Exploratória.
- Conceitos Amostragem.
- Probabilidade - Modelos de distribuição.
- Noções de Medidas de Associação e Correlação.
- Processo de Análise Estatística – Testes Estatísticos.
- Teste de hipóteses paramétrico e não paramétrico.
- Processo inferencial (Clássica e Bayesiana).

- Intervalo de confiança.

### Bibliografia Básica

MORETTIN, P. A.; BUSSAB, W. de O.; **Estatística Básica**; São Paulo, Editora Saraiva, 2017.

MOORE, D. S.; MCCABE, G. P.; DUCKWORTH, W. M.; SCLOVE, S. L. **Prática da Estatística Empresarial Como Usar Dados para Tomar Decisões**. São Paulo, 2003.

### Bibliografia Complementar

AMARAL, F. **Fundamentos da programação R**. São Paulo: eBook Kindle, 2018.

GROLEMUND, G. **R for Data Science**. Sebastopol: O'Reilly Media, 2017.

LAROSE, C. D.; LAROSE, D. T. **Data Science using Python and R**. Hoboken: Wiley-Blackwell, 2019.

TEETOR, P. **R Cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics**. Sebastopol: O'Reilly Media, 2011.

## Disciplina

## Machine Learning

### Ementa

Fornecer uma visão dos principais algoritmos de classificação, agrupamento e otimização, possibilitando o entendimento dos conceitos fundamentais, do funcionamento dos métodos, melhores aplicações e desenvolvimento. Desenvolver habilidades que permitam a análise de casos de negócios.

- Casos de uso de ML.
- Métodos de Aprendizagem (Supervisionado, não supervisionado, reforço).

- Regressões.
- Conglomerados.
- Validação de Modelos.
- Métricas.
- Métodos Ensemble.
  - random forest.
  - Adaboost.
  - Xgboost.
  - lightGBM.
  - Catboost.
  - hitogramGradientBoost.
  - OPF.
- Outros Modelos de Machine Learning:
  - Knn.
  - SVM.
  - Naive Bayes.
  - Decision Trees.
- Análise de Conglomerados:
  - Kmeans.
  - Dbscan.
  - Hclust.

### Bibliografia Básica

CARVALHO, L. A. V. de. **Datamining A Mineração de Dados no Marketing, Medicina, Economia, Engenharia e Administração**. Rio de Janeiro: Editora Ciência Moderna, 2005.

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

### Bibliografia Complementar

DEAN, J. **Big Data, Data Mining, and Machine Learning: Value Creation for**

Business Leaders and Practitioners. Hoboken: John Wiley & Sons, 2016.  
 LUGER, G. **Inteligência Artificial**. 6. ed. São Paulo: Pearson Education, 2013.

**Disciplina**

**Deep Learning**

**Ementa**

Fornecer uma visão sobre os principais conceitos sobre redes neurais, deep learning, do funcionamento das principais variações, melhores aplicações e desenvolvimento. Desenvolver habilidades que permitam a análise de cases de negócios.

- Redes Neurais Perceptron.
- Combinação de regiões não lineares.
- Multi layer perceptron (MLP).
- Feedforward.
- Redes Neurais Backpropagation.
- Regra da Cadeia Overfitting e Underfitting.
- Regularização L1 e L2, Dropout.
- Funções de ativação, Learning rate.
- Redes Convolucionais.
- Redes Recorrentes.
- Redes Não Supervisionadas.
- Generate Adversarial Networks.
- Transfer Learning, BERT, SOM.
- Deep Q-Learning.
- Sistemas neurais complexos.
- Tuning de hiperparâmetros.



### Bibliografia Básica

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

PEI, J; HAN, J; KAMBER, M. **Data Mining Concepts and Techniques**. São Paulo, 2012.

### Bibliografia Complementar

CARVALHO, L. A. V. **DataMining: a mineração de dados no marketing, medicina, economia, engenharia e administração**. Rio de Janeiro: Editora Ciência Moderna, 2005.

### Disciplina

### Recommendation Systems

### Ementa

Fornecer uma visão dos principais conceitos e técnicas de Sistemas de Recomendação bem como o entendimento de suas propriedades e principais aplicações destes sistemas no mercado. Possibilitar a visão de uma arquitetura para Sistemas de Recomendação utilizando os conceitos, algoritmos, propriedades e métricas de avaliação de eficácia.

Conceitos sobre a disciplina Sistemas de Recomendação e aplicações de mercado.

- Conceito.
- Tipos de Sistemas de Recomendação.
- Propriedades: cold start, aleatoriedade, previsibilidade, novidade, cobertura, conhecimento do usuário e serendipidade.

- Aplicações de mercado.
- Estudar o uso dos Sistemas de Recomendação em diferentes negócios (Varejo, Streaming, Redes Sociais, Mercado Financeiro e Notícias).
- Arquitetura, técnicas e algoritmos para Sistemas de Recomendação.
- Técnicas e Algoritmos.
- Popularidade.
- Tempo de permanência.
- Navegação.
- Busca por palavras chaves.
- Características do item.
- Preferências e histórico do usuário.
- Recomendação baseada em conhecimento.
- Recomendação baseada em conteúdo.
- Pré-processamento, distância e similaridade.
- Filtro colaborativo.
- Agrupamento.
- Raciocínio baseado em casos.
- Abordagem híbrida.
- Formas de mesclar recomendação.
- Aplicação de ranqueamento.
- Estratégias de recomendação de acordo com o negócio e calibragem do modelo.
- Validação.
- Indicadores para aferir a eficácia de um sistema de recomendação.
- Próximos passos e tendências para Sistemas de Recomendações.

### Bibliografia Básica

AGGARWAL, C. C.; **Recommender Systems: The Textbook**. Estados Unidos: Editora Springles – eBook, 2016.

PEI, J; HAN, J; KAMBER, M. **Data Mining Concepts and Techniques**. São



Paulo, 2012.

### Bibliografia Complementar

GORAKALA, S. K.; **Building Recommendation Engines**. Estados Unidos: Editora Packet, 2016.

### Disciplina

### Database for Data Science

### Ementa

Nesta disciplina traremos luz aos bancos de dados NOSQL, seu nascimento, propostas de técnicas e de negócios que visam responder, sua diversidade arquitetural e componentes internos. Falaremos das principais bases de gestão de conteúdo semi-estruturado como: bases chave-valor, document-based, bases colunares e graph-based. Ênfase em bases NOSQL e NEWSQL:

- Arquitetura, Instalação e Funcionamento das principais bases NOSQL
- Bases Key-Value – Redis:
  - Criação de pares chave-valor simples e dos tipos compostos como SETs, SETs ordenados, Listas, Hashs e Georeferenciados
- Bases Column-Store – Cassandra:
  - Criação de Keyspaces, ColumnFamilies e execução de CRUD e suas características.
- Bases Document-Based – MongoDB:
  - Criação de bancos de dados, documentos, importação de JSONs, criação de índices e execução de CRUD.
- Bases Graph-Based – Neo4J:
  - Criação de bancos de dados, nós, arestas e propriedades, bem como a aplicação de CRUD e mecanismos de consulta usando Cypher.
- Introdução a Bases NewSQL



### Bibliografia Básica

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

FOWLER, M. **NoSQL essencial**: um guia conciso para o mundo emergente da persistência poliglota. São Paulo: Novatec, 2013.

### Bibliografia Complementar

HARRISON, G. **Next Generation Databases**: NoSQL, NewSQL, and Big Data. New York: Apress, 2015.

McCREARY, D.; KELLY, A. **Making sense of NoSQL**: a guide for managers and the rest of us. Shelter Island: Manning Publications, 2013.

**Disciplina****Linear Algebra for Data Scientists****Ementa**

Nesta disciplina, dedicaremos um tempo precioso para tratar de um dos pilares da ciência de dados: A matemática aplicada, onde veremos, álgebra linear, teoria dos Conjuntos, vetores, matrizes, determinantes, logaritmo, introdução ao cálculo numérico & funções. Ênfase em Álgebra Linear e Cálculo Numérico:

- Teoria dos Conjuntos.
- Vetores.
- Matrizes.
- Determinantes.
- Logaritmo.
- Introdução ao Cálculo numérico / Funções.

**Bibliografia Básica**

BOLDRINI, J. L.; COSTA, S.I.R.; RIBEIRO, V. L., WETZLER, H.G., **Álgebra Linear**. São Paulo: Harper-Row, 1980.

CALLIOLI, C. A; DOMINGUES, H. H.; COSTA, R.C.F. **Álgebra Linear e Aplicações**. 4ª ed. São Paulo: Atual, 1983.

**Bibliografia Complementar**

BOULOS, P. & CAMARGO, I. **Introdução à geometria analítica no espaço**. São Paulo: Makron Books, 1997.

CAROLI, A. de; CALLIOLI, C. A.; FEITOSA, M. O. **Matrizes, vetores, geometria analítica**. São Paulo: Novel, 1994.

DANIEL, J.W. **Álgebra linear aplicada**. Prentice-Hall do Brasil, 1986.

LEON, S.J. **Álgebra linear com aplicações**. Rio de Janeiro: LTC, 1998.

LIPSCHUTZ, S. **Álgebra linear**. São Paulo: McGraw-Hill, 1971. NOBLE, B. & STEINBRUCH, A. & WINTERLE, P. **Álgebra linear**. São Paulo: McGraw-Hill, 1987.

SANTOS, N. M. dos. **Vetores e matrizes**. Rio de Janeiro: Ao Livro Técnico, 1974.

STRANG, G. Linear algebra and its applications. San Diego: Harcourt.

WINTERLE, P. **Vetores e geometria analítica**. São Paulo: McGraw Hill, 2000.

## Disciplina

## R for Data Scientists

### Ementa

Tem como objetivo apresentar de forma prática os conceitos básicos da linguagem R para análise de dados, com foco na manipulação de dados, uso de bibliotecas, sua arquitetura e geração de gráficos.

- Introdução ao R.
- RStudio.
- Sintaxe e Tipos de Dados.
- Funções e Pacotes.
- Arquivos e I/O.
- Trabalhando com dados temporais.
- Limpeza e Normalização de Dados.
- Gráficos em R.
- Data transformation.
- Workflow: scripts.
- Exploratory Data Analysis.
- Tibbles.
- TidyData.
- String and Factors.
- Pipe.
- Functions and Vectors.



### Bibliografia Básica

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

MENEZES, N. N. C. **Introdução à programação com Python Algoritmos e lógica de programação para iniciantes**. São Paulo: Novatec, 2021.

### Bibliografia Complementar

TEETOR, P. **R Cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics**. Sebastopol: O'Reilly Media, 2011.

### Disciplina

### Python for Data Scientists

### Ementa

Tem como objetivo permitir ao aluno dominar os fundamentos da escrita de scripts Python, usando os principais elementos de script do Python, como variáveis e estruturas de controle de fluxo, listas de dados, ler e gravar arquivos. A disciplina trará o conceito da criação de funções Python para reutilização de código, usar as principais bibliotecas Python para várias áreas e sua integração com outras aplicações:

- Introdução ao Python.
- Anaconda (Jupyter Notebooks).
- Sintaxe & Tipos de Dados.
- Funções & Pacotes.
- Arquivos e I/O.
- Trabalhando com dados temporais.
- Programação funcional.

- Numpy.
- Pandas.
- Conceito de Objetos.
- Estruturas e modelos.
- Scikit-Learn & Scipy.
- Matplotlib.
- Statsmodels & Visualização.

### Bibliografia Básica

MENEZES, N. N. C. **Introdução à Programação com Python**: algoritmos e lógica de programação para iniciantes. São Paulo: Novatec, 2021.

### Bibliografia Complementar

FORBELLONE, A. L. V.; EBERSPACHER, H. F. **Lógica de Programação**: a construção de algoritmos e estrutura de dados. 3. ed. São Paulo: Pearson Prentice Hall, 2012.

KINSLEY, H.; MCGUGAN, W. **Introdução ao Desenvolvimento de Jogos em Python com PyGame**. São Paulo: Novatec, 2015.

LOPES, A.; GARCIA, G. **Introdução a Programação**: 500 Algoritmos. Rio de Janeiro: Campus, 2002.

PAYNE, B. **Ensine seus filhos a programar**. São Paulo: Novatec, 2015.

PIVA Jr., D.; NAKAMITI, G. S.; ENGELBRECHT, A. M. **Algoritmos e Programação de Computadores**. Rio de Janeiro: Elsevier, 2012.

**Disciplina****Data Ecosystems - Infrastructure****Ementa**

Trabalhar com dados e seu ecossistema são de extrema importância para analistas e cientistas de dados.

- Conceitos ao redor de Big Data e Lakehouses.
- Data Sensing – Internal/External Sources, Softwares and Algorithms
- Working with Collections.
- Data Wrangling.
- Data Analysis.
- Storage – Cloud and OnSite & Other Models.
- Examples of Architectures – DataBricks, Snowflake, Trifacta and Data Robot.

**Bibliografia Básica**

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

TAURION, Cezar. **Big Data**. Rio de Janeiro: Brasport, 2013.

**Bibliografia Complementar**

BLUE GRANITE. **Data Lakes in a Modern Data Architecture**. 2018. Disponível em: <<https://www.blue-granite.com/data-lakes-in-a-modern-data-architecture-ebook>>. Acesso em: 26 jul. 2022.

GOOGLE CLOUD. **Leitura de imagens com Cloud Vision API**. 2019. Disponível em: <<https://cloud.google.com/vision/>>. Acesso em: 26 jul 2022.

GORELIK, A. **The Enterprise Big Data Lake**. Sebastopol: O'Reilly Media Inc., 2019.

ORAM, A.; GIDLEY, S. **Data Lake Maturity Model**. Sebastopol: O'Reilly Media Inc., 2019.

ROSS, R.; KIMBALL, M. **The data warehouse toolkit: guia completo para modelagem dimensional**. Rio Janeiro: Campus, 2002.

SADALAGE, P. J.; FOWLER, M. **NOSQL Essencial**. São Paulo: Novatec, 2013.

INMON, B.; LEVINS, M.; SRIVASTAVA, R. **Building the Data Lakehouse**.

Estados Unidos: Technics Publications, 2020.

## Disciplina

## Feature Engineering

### Ementa

Esta disciplina disserta sobre formas e métodos avançados em Data Mining e Feature Engineering.

- Missing Data Inputation.
- Feature Scaling.
- Unbalance Data.
- Downsampling.
- Upsampling (SMOTE).
- Redução de Dimensionalidade.
- PCA.
- LDA.
- T-SNE.
- Feature Extraction (non-structured data).
- Feature Store.
- Feature Selection Methods.
- Métodos de Filtro.
- Métodos Embarcados.
- Métodos Wrapper.
- Backward Feat. Elimination.
- Forward Feat. Construction.





### Bibliografia Básica

KIMBALL, R.; ROSS, M. **The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling**. 2ª ed., John Wiley & Sons, 2002.

PEI, J; HAN, J; KAMBER, M. **Data Mining Concepts and Techniques**. São Paulo, 2012.

### Bibliografia Complementar

GRISHAM, R. **Computational Linguistics – an Introduction**. Cambridge: Cambridge University Press, 1992.

KAO, A. and POTEET, S.R. (eds.) **Natural Language Processing and Text Mining**. New York: Springer, 2006.

LUGER, G.F. **Cognitive Science**. [s.i]. Academic Press, 1994.

MANNING, C.D. and SCHFITZE, H. **Foundations of Statistical Natural Language Processing**. Cambridge, MA: MIT Press, 1999.

OTHERO, G. de A. **Introdução à linguística computacional**. Comunicação apresentada na III Semana Acadêmica da Faculdade de Letras da Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS, 2003.

### Disciplina

**MLOps**

### Ementa

Esta disciplina trata os conceitos e práticas de desenvolvimento estruturado, controlado e rápido de modelos em Machine Learning, consolidando-se como um conjunto de práticas que combina Machine Learning, DevOps e Engenharia

de Dados, a fim de apoiar na implantação e manutenção dos sistemas de ML em produção de maneira confiável e eficiente. Ênfase na aplicabilidade MLOps para entregas em Machine Learning:

- Proposta MLOps.
- Equipes e Projetos.
- Data and Models Pipeline.
- MLOps Management (Validation, Data Rules, Data Quality, Model Quality, Versioning).

### Bibliografia Básica

AOIFE, D; NAMEE, B. M.; KELLEER, J. D. **Fundamentals of Machine Learning for Predictive Data Analytics Algorithms, Worked Examples, and Case Studies.** EUA: The MIT Press, 2015.

GIFT, N.; DEZA, A. **Practical Mlops: Operationalizing Machine Learning Models,** Estados Unidos: O'Reilly, 2021.

MUNIZ, A.; IRIGOYEN, A.; SANTOS, R.; MOUTINHO, R. **Jornada DevOps.** 2ª ed. São Paulo: Editora Brasport, 2020.

TREVELL, M.; **Introducing Mlops: How to Scale Machine Learning in the Enterprise.** Estados Unidos: O'Reilly, 2020.

### Bibliografia Complementar

RAJ, E.; **Engineering MLOps: Rapidly build, test, and manage production-ready machine learning life cycles at scale.** Estados Unidos: Packet Publishings, 2021.



<b>Disciplina</b>	<b>Text &amp; Social Media</b>
<b>Ementa</b>	
<p>Esta disciplina aborda a geração de dados humanos e textos.</p> <ul style="list-style-type: none"> <li>• Difusão da Informação.</li> <li>• Mineração de Textos.</li> <li>• Análise de Emoções, BoW.</li> <li>• Teoria de Grafo.</li> <li>• Técnicas de Embeddings.</li> <li>• Vetorização.</li> <li>• Topologias.</li> </ul>	
<b>Bibliografia Básica</b>	
<p>BARGER, C.; <b>O Estrategista em Mídias Sociais</b>, São Paulo: Editora DVS, 2013.</p> <p>MARTINO, L. M. S. <b>Teoria das Mídias Digitais</b>. São Paulo: Vozes, 2015.</p>	
<b>Bibliografia Complementar</b>	
<p>N/C.</p>	

**Disciplina**

**Reinforcement Learning**

**Ementa**

Esta disciplina aborda sistemas de Aprendizagem por Reforço:

- Fundamentals in Reinforcement Learning.
- Sequential Decision-Making.
- Métodos de Aprendizagem Sample-Based.
- Processo de decisão Markov.
- Value Functions e Equações de Bellman.
- Predições e controle com Funções de Aproximação.
- Programação Dinâmica.
- Sistemas em Reinforcement Learning.

**Bibliografia Básica**

AOIFE, D; NAMEE, B. M.; KELLEER, J. D. Fundamentals of Machine Learning for Predictive Data Analytics Algorithms, Worked Examples, and Case Studies. EUA: The MIT Press, 2015.

VALDATI, A. de B. Inteligência artificial - IA. Curitiba: Editora Contentus, 2020.

**Bibliografia Complementar**

LAPAN, M; **Deep Reinforcement Learning Hands-On: Apply modern RL methods, with deep Q-networks, value iteration, policy gradients, TRPO, AlphaGo Zero and more.** Estados Unidos: Packet Publishings, 2018.



<b>Disciplina</b>	<b>Human-Based Data</b>
<b>Ementa</b>	
<p>Entraremos no mundo de trato com dados humanos como NLP em ambientes de dados gerados por Humanos:</p> <ul style="list-style-type: none"> <li>• Algoritmos em linguagem natural.</li> <li>• Geoprocessing de Dados.</li> <li>• Image and Sound processing Fundamentals.</li> </ul>	
<b>Bibliografia Básica</b>	
<p>FITZ, P. R.; <b>Geoprocessamento sem Complicação</b>, São Paulo: Editora Oficina de Textos, 2008.</p> <p>GONZALEZ, R. C.; WOODS, R. E.; <b>Processamento Digital de Imagens</b>. São Paulo: Editora Pearson Universidades, 2009.</p>	
<b>Bibliografia Complementar</b>	
<p>BHAVSAR, K.; DANGETI, P.; <b>Natural Language Processing with Python Cookbook</b>. Estados Unidos: Editora Packt Publishings, 2017.</p> <p>CUBAS, M. G.; TAVEIRA, B. D. de A.; <b>Geoprocessamento: fundamentos e técnicas</b>, São Paulo: Editora Intersaberes, 2021.</p>	

<b>Disciplina</b>	<b>DataOps &amp; Governance</b>
<b>Ementa</b>	
<p>Esta disciplina aborda a Governança de Dados tendo DataOps como aliado para</p>	

permitir identificar as diversas fontes de dados, montar um catálogo de dados, permitir a criação do Dado Mestre e garantir a rastreabilidade dos dados na organização.

- DataOps.
- Governança de Dados.
- Dado Mestre (MDM).
- Qualidade de Dados.
- Catálogo de Dados.
- Integração de Dados.
- Ferramentas de apoio.

### **Bibliografia Básica**

DONDA, D.; **Guia Prático de Implementação da LGPD**. São Paulo: Editora Labrador, 2020.

REGO, B. L. **Simplificando a governança de dados: governe os dados de forma objetiva e inovadora**. Rio de Janeiro: Brasport, 2020.

### **Bibliografia Complementar**

BARBIERI, C.; **Governança de Dados: Práticas, conceitos e novos caminhos**. São Paulo: Editora Alta Books, 2019.

**Disciplina****Data-Driven Organizational Strategies****Ementa**

Esta disciplina tem como objetivo identificar as estratégias de negócio que facilitam uma empresa adotar os dados como ferramenta básica para o processo de tomada de decisão e estabelecer o processo de transformação digital. Demonstra casos reais e técnicas como gamificação para atingir os objetivos organizacionais.

- Transformação Digital e Pensamento Estratégico.
- Transformação dos Negócios e Geração de Valor.
- Modelo de Maturidade e Framework para Transformação Digital.

**Bibliografia Básica**

SÁTYRO, W. C.; SACOMANO, J. B.; GONÇALVES, R. F.; SILVA, M. T. da; BONILLA, S. H. **Indústria 4.0: conceitos e fundamentos**. São Paulo: Blucher, 2018.

ROGERS, DAVID L. **Transformação digital: Repensando o seu negócio para a era digital**. Belo Horizonte: Autêntica Business, 2017.

**Bibliografia Complementar**

SCHWAB, K.; **A Quarta Revolução Industrial**. São Paulo: Editora Edipro, 2018.

**Disciplina****Data Science for Business****Ementa**

Nesta disciplina abordaremos questões sobre a utilização da ciência de dados nos negócios, como principais processos analíticos em vendas, Análise de mercado e principais KPIs, Controles e análises em retail, Algoritmos e modelos específicos. Processos importantes em agro e seus KPIs, principais indicadores para Health especificamente - Saúde, Algoritmos e modelos específicos, Projetos em ciência de dados para área de negócios. Discussões e provocações ao redor do uso da Ciência de Dados em diversas industrias e áreas do mercado:

- Análise de mercado e principais KPIs para Telco.
- Controles e análises em Retail and Finance.
- Algoritmos e modelos específicos para Manufacturing.
- Processos importantes em Agro e seus KPIs.
- Principais indicadores para Health especificamente – Saúde.
- Algoritmos e modelos específicos, projetos em ciência de dados para área de negócios (Vendas e Marketing).

**Bibliografia Básica**

FOREMAN, J. W.; **Data smart: usando Data Science para transformar informação em insight**, São Paulo: Editora Alta Books, 2016.

PROVOST, F.; FAWCETT, T.; **Data science para negócios: O que você precisa saber sobre mineração de dados e pensamento analítico de dados**. São Paulo: Editora Alta Books, 2016.





## Bibliografia Complementar

HURLEY, R.; **Predictive Analytics**. Estados Unidos: ISBN-10 1952191009, 2020.

## Disciplina

## Business Analytics

## Ementa

Fundamentos Competitivos para Business Intelligence (Evolução Analytics, Data Driven Economy), alinhamento da estratégia de negócios à gestão de dados, Frameworks para indicadores de desempenho (BSCs x Dashboards); Re-Visão sobre Arquitetura em Analytics (Camadas e seus conceitos, Arquitetura de Dados: Data Warehouse, banco de dados e Data Lake, Integração de dados – estruturados e não estruturados (in review));

Gerando Insights (Métodos e Conceitos de Visualização de Dados, Cognição humana, porque não fomos feitos para interpretar textos e números, senso comum e técnicas de apresentação);

Técnicas de Visualização (Usando atributos de pré-atenção, Uso de cores em apresentações, Cuidados estéticos, Diagrama de Gutemberg e o Padrão F);

Tipos, estilos e usos de gráficos (Barras, linhas e áreas, Pie, donut e tremps, Gráficos combinados, bubble charts, pareto e waterfall, Gantt, gráficos de densidade e mapas, Geoanálise); Técnicas de Desenho de Painéis (Dashboards, Storytelling).

- Business Intelligence & Data Driven Orgs.
- Re-Visão sobre Arquitetura em Analytics.
- Gerando Insights.
- Técnicas de Visualização.
- Tipos, estilos e usos de gráficos.
- Técnicas de Desenho de Painéis (Dashboards, Storytelling).



### Bibliografia Básica

BARBIERI, C. **BI2- Business Intelligence Modelagem & Qualidade**. Rio de Janeiro: Elsevier, 2011.

HAIR, J. F. Jr.; ANDERSON, R. E.; TATHAM R. L.; BLACK, W. C. **Análise Multivariada de Dados**. 5. ed. Porto Alegre: Bookman, 2005.

LARSON, R.; FARBER, B. **Estatística Aplicada - 4ª edição**. São Paulo: Pearson Universidades, 2009.

MAHESHWARI, A. **Data analytics made accessible**. 2020. *E-book*.

Disponível em: <https://www.amazon.com.br/Data-Analytics-Made-Accessible-English-ebook/dp/B00K2I2JL8>. Acesso em: 26 jul. 2022.

### Bibliografia Complementar

MOORE, S. D.; MCCABE, G. P.; DUCKWORTH, W. M.; SCLOVE, S. S. **Estatística Empresarial: como usar dados para tomar decisões**. Rio de Janeiro: LTC, 2006.

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

### Disciplina

### Empreendedorismo e Inovação

### Ementa

Introdução ao empreendedorismo inovador e aos modelos de criação de novas empresas emergentes. Apresentação de métodos e ferramentas para ideação. Técnicas e ferramentas de validação de negócios e análise de mercado. Noções sobre intraempreendedorismo e modelos internos de inovação. Modelos empreendedores para criação, testes e evolução de propostas de valor. Modelos

e ferramentas de prototipação de negócios. Noções sobre ecossistemas empreendedores e de inovação. Técnicas de storytelling e formatação de apresentações (pitch).

### Bibliografia Básica

CARVAJAL J. J.; SANCHEZ, W. M.; et al. **Empreendedorismo, Tecnologia e Inovação**. São Paulo: Editora Livrus, 2015.

RIES, E. **A startup enxuta: como os empreendedores atuais utilizam a inovação contínua para criar empresas extremamente bem-sucedidas**. São Paulo: Editora Lua de Papel, 2012.

### Bibliografia Complementar

BESSANT, J. R.; TIDD, J. **Inovação e empreendedorismo**. Porto Alegre: Bookman, 2009.

COZZI, A; JUDICE, V; DOLABELA, F. **Empreendedorismo de base tecnológica spin-off: criação de novos negócios a partir de empresas constituídas, universidades e centros de pesquisa**. São Paulo: Elsevier Academic, 2012.

DRUCKER, P. F. **Inovação e espírito empreendedor (entrepreneurship): prática e princípios**. São Paulo: Cengage Learning, 2014.

GOVINDARAJAN, V; TRIMBLE, C. **Beyond the idea how to execute innovation in any organization**. ST: Martin's Press, 2013.

OSTERWALDER, A., PIGNEUR, Y. **Businnes Model generation: The handbook for visionaries, game changers, and challengers**. New Jersey: Wiley 2010.



## Design Experience FIAP

O MBA on-line em Data Science foi estruturado em fases para que, didaticamente e do ponto de vista do ensino-aprendizagem, fosse efetivo e pudesse abranger uma série de *skills (hard e soft)*, possibilitando ao aluno uma grande interação e integração entre o conteúdo de caráter técnico e sua base a partir do prisma de negócios, tendo como pano de fundo situações e cases reais.

Além disso, estimulamos os alunos a desenvolverem trabalhos em conjunto, incentivando o networking inter e extraclasse. Buscamos, com isso, a prática e o desenvolvimento das relações interpessoais e profissionais no âmbito estatístico e científico, almejando a resolução de desafios complexos.

Propomos, também, como desafio, a criação de uma startup (projeto Startup One), em que o aluno será convidado a exercitar questões relacionadas à comunicação, negociação, apresentação e gestão da emoção.

Um de nossos focos principais é apoiar o aluno no desenvolvimento de uma sociedade fraternal, de apoio e ajuda mútuos, possibilitando o desenvolvimento acelerado e colaborativo do grupo.

Pensando em manter o aluno atualizado com as últimas tendências e ouvir os profissionais do mercado, foi criado um podcast especificamente voltado para analisar o mundo dos dados e a profissão do cientista de dados. O Data On Air, disponível nas principais plataformas de streaming, trata a ciência de dados e todo o universo dos dados de uma maneira simples e descomplicada, trazendo casos, técnicas e abordagem de negócio para envolver o aluno no dia a dia do profissional desta área. Acreditamos que esses conteúdos favorecem o acesso do aluno ao mercado de trabalho.

## Processo de Avaliação

Como o aluno é estimulado a realizar uma imersão em Ciência de Dados, utilizando todas as ferramentas e técnicas do mundo corporativo, uma empresa parceira é convidada a oferecer casos reais para que os alunos busquem soluções apropriadas para resolver os problemas apresentados. O

conteúdo on-line disponibilizado serve de base para que os alunos possam propor soluções, mas não limita o aprendizado do aluno que é constantemente orientado a buscar alternativas viáveis, dentro ou fora do material disponibilizado. Aulas ao vivo complementam o conteúdo on-line, fornecem atualização de técnicas e ferramentas de mercado e permite que o aluno amplie sua visão sobre Big Data e Ciência de Dados. Por meio de mentorias técnicas, os professores/mentores orientam os alunos quanto a possibilidade de colocar em prática as ideias trazidas por eles e guiam os estudos para facilitar a aquisição do conhecimento. Por meio das mentorias de negócio, os alunos têm a possibilidade de interagir com a empresa parceira, entender o problema de negócio, discutir alternativas e apresentar sugestões inovadoras que são validadas pelos profissionais do mercado.

Ao longo das fases o aluno desenvolve um projeto completo, desde a concepção, identificação da necessidade, desenho e implementação da arquitetura, criação e manutenção de algoritmos e modelos estatísticos para solucionar o desafio apresentado.

Ao final do curso o aluno tem, além de um portfólio completo com um projeto real de ponta a ponta, a experiência profissional do Cientista de Dados.

A avaliação se dá naturalmente e de forma progressiva em cada uma das fases. O aluno recebe uma nota técnica atribuída pelos professores/mentores em cada uma das fases baseada nos entregáveis do projeto. No final do curso o aluno apresenta a sua solução para o problema de negócio com a visão holística da solução.

O sistema de avaliação é dividido entre **atividades a distância (AD)** e **atividades presenciais obrigatórias (APO)**, totalizando 100 pontos. Para aprovação, o aluno deve alcançar nota superior ou igual a 70 pontos.

As **atividades a distância (AD)** representam 60 pontos (60%) da nota final, e são compostas por avaliações dissertativas e práticas, realizadas individualmente e/ou em grupo, por meio da plataforma FIAP ON. Para o desenvolvimento de atividades práticas a distância, serão disponibilizadas duas formas de acesso:

- O aluno utilizar o próprio computador, pois o conteúdo disponibilizado contém o passo a passo de instalação dos softwares necessários, e que são gratuitos;
- O aluno utilizar, presencialmente, no horário de sua preferência, as estruturas de laboratórios da FIAP, que contam com os softwares necessários para as práticas.

As **atividades presenciais obrigatórias (APO)** representam 40 pontos (40%) da nota final e são compostas por avaliação objetivas, projetos e dinâmicas, pautados na relação prática e teórica, envolvendo todas as unidades curriculares abordadas até a data do encontro que ocorre ao final do curso, sendo assim, tais atividades integram e avaliam todos os conhecimentos e habilidades das disciplinas apresentadas e, por esta razão, os 40 pontos aqui mencionados são usados para compor essas disciplinas (quarenta para cada uma delas).

A respeito das notas finais: caso o aluno obtenha a nota inferior a 70 pontos, estará automaticamente reprovado na disciplina e deverá realizar uma avaliação substitutiva, de forma a substituir a nota deficitária.

## **Projeto Integrador – Startup One MBA FIAP ON**

O Startup One é integrado aos cursos por meio da disciplina de Empreendedorismo e Inovação, ministrada em todos os cursos de MBA da FIAP. Para a modalidade do MBA on-line (FIAP ON), os conteúdos serão disponibilizados em três (3) fases/períodos distintos do curso, além de encontros on-line ao vivo para mentorias individuais e em grupo com intuito de apoiar os alunos na jornada. O *framework* da disciplina, composto por seu conteúdo, materiais e dinâmicas, foram desenvolvidos com a utilização dos conceitos de *Design Thinking* e *Lean Startup*, aplicando conhecimentos específicos de acordo com a necessidade e respeitando os limites da aplicação de cada método, dado a carga horária.

A disciplina caracteriza-se pela orientação aos alunos de MBA para elaborarem, ao longo do curso, um projeto (plano de negócio prático) para a criação de uma Startup, configurando o trabalho final do curso. Este trabalho final (ou projeto) substitui o TCC (Trabalho de Conclusão de Curso) e é entregue ao final do curso, podendo ser executado em grupos de até quatro (4) alunos.

O projeto pode ser inscrito no Startup One – ST1, competição que ocorre ao final de cada ciclo do MBA FIAP.

### **Objetivos da disciplina:**

- Conceituar os elementos básicos do empreendedorismo.
- Discutir as características principais dos empreendedores, bem como sua aplicação na criação de startups.
- Capacitar o aluno a entender a jornada de um empreendedor, desde a identificação e validação do problema, desenvolvimento da solução, criação e validação do protótipo, análise financeira do empreendimento e apresentação resumida da solução (pitch).

### **Estrutura**

Os conteúdos dos cursos MBA on-line são separados em fases, onde cada uma tem um propósito e direcionamento. Os conteúdos relacionados ao tema Startup também são apresentados em fases, que por sua vez se tornam disponíveis para os alunos em três momentos dentro do curso.

São no total três (3) fases, 6 (seis) capítulos, 6 (seis) mentorias em grupo/aula ao vivo, 6 (seis) atividades intermediárias que não possuem nota (apenas feedback), e uma atividade final (entrega final do Startup One – TCC) organizados da seguinte forma:

## **STO 1 – Fase 1 do Startup One MBA ON**

### **Conteúdos da fase 1**

#### **Capítulo 0 – Welcome to Startup One**

Capítulo de instruções iniciais e boas-vindas ao Startup One, modelo integrado de desenvolvimento do trabalho final dos MBAs da FIAP. É baseado na metodologia de Project Base Learning e busca a convergência entre as disciplinas, capacidades e atitudes dos alunos para estimular suas jornadas de aprendizado de maneira inovadora.

#### **Capítulo 1 – Introdução ao empreendedorismo inovador**

Este capítulo trata a evolução das tecnologias exponenciais e o fato de que o empreendedorismo vem passando por uma transformação social nos últimos anos e as startups de base tecnológica se tornaram uma opção atraente como alternativa de investimento e carreira, para empreendedores e até mesmo para grandes empresas. Por meio de processos bem definidos e estruturados, e somado à flexibilidade e a um crescimento rápido e contínuo, as startups contribuem significativamente com o desenvolvimento econômico.

#### **Capítulo 2 – Como nascem as boas ideias**

Neste capítulo, serão relatadas histórias de startups bem-sucedidas, o padrão que encontramos em boas ideias, tipos de análise, cuidados que se deve tomar ao empreender e entrar em um novo mercado, o que é disfunção e como é a jornada do empreendedor.

#### **Atividades da fase 1**

##### Atividade 1 – Identificando Oportunidade



Desafio e objetivo: Dar início ao projeto de startup. Nesta atividade desafiamos os alunos a buscarem identificação com áreas e segmentos de mercado para, a partir disso, identificar problemas e oportunidades deste segmento. Instruímos os alunos a buscarem pelo menos três (3) grandes problemas para ser base de uma análise futura mais profunda, focada em pesquisa mais detalhada para compreender o problema na sua essência.

### **Mentorias em Grupo/Aula ao vivo (on-line)**

#### Mentoria 1 - Aula inaugural do Startup One

Neste encontro on-line nos reunimos pela primeira vez com os alunos para contextualizar sobre o programa Startup One e ajudar nas direções iniciais dos projetos. Explicamos a jornada do Startup One, as fases e entradas de conteúdos sobre o tema durante o curso, as dinâmicas das mentorias em grupo/aulas ao vivo, o apoio dos professores como pontos focais do curso e agenda de mentorias individuais. Além disso, estruturamos palestras de 15 minutos junto dos professores que falarão sobre temas relacionados a problemáticas de diversos segmentos de mercado para inspirar os alunos e dar dicas de como buscar problemas de grande escala e alto impacto. Como segunda fase deste encontro, estimulamos a formação de grupos multidisciplinares entre diferentes cursos para iniciar a jornada de inovação do Startup One.

#### Mentoria 2 – Validar problema

Nesta mentoria o professor busca despertar nos alunos a importância da validação do problema, os desafios existentes nesta fase do negócio e dicas para não ter análises com vieses que possam gerar riscos para o projeto.

## STO 2 – Fase 2 Startup One MBA ON

### Conteúdos da fase 2

#### Capítulo 3 – Business Canvas

O início de um bom negócio começa com um rascunho do modelo de negócio. A partir da ideia levantada, agora é preciso compreender sua estruturação como negócio. Neste capítulo, o objetivo é compreender a lógica do *Business Model Canvas* e aplicá-lo no projeto do Startup One, permitindo registrar e comunicar decisões do projeto, bem como extrair *insights* e hipóteses para validação.

#### Capítulo 4 – Como testar e evoluir sua ideia de negócio

Neste capítulo, falamos sobre como validar e desenvolver as ideias nos negócios em empresas e startups sustentáveis, o que é um MVP e quais são as abordagens e desafios na etapa inicial do empreendedor, sua escala e tração. Entre esses desafios está o da prototipação.

#### Atividades

##### Atividade 2 – Canvas

Desafio e objetivo: A partir da identificação do problema, iniciar o mapeamento de um modelo de negócio inicial, incluindo formação de equipe e suas habilidades, definição refinada do problema e do cliente, análise de tamanho do mercado, análise de concorrente, mapeamento da proposta de valor por meio do Canvas Proposta de Valor e modelo de negócio por meio do *Business Model Canvas*.

### Atividade 3 – Prototipação

Desafio e objetivo: Efetivamente tirar a ideia do papel, criar um protótipo conceitual da solução por meio de ferramentas de prototipação apresentadas nos conteúdos na plataforma FIAP ON e nas aulas ao vivo, a fim de validar hipóteses da solução e saber se realmente ela gera valor para os clientes.

### **Mentorias em Grupo/Aula ao vivo**

#### Mentoria 3 – Canvas

Nesta mentoria, os professores trazem análises diferentes de modelos de negócio para ajudar os alunos a compreender os prós e contras de cada modelo, os riscos envolvidos de cada um deles com o objetivo de ajudar os alunos a modelarem e validações a estrutura de escala, análise de dependências de fornecedores, estruturas de custo e receita.

#### Mentoria 4 - Validar MVP

Nesta mentoria os professores discutem sobre estratégias de validação de hipóteses sobre a solução da startup por meio do desenvolvimento do mínimo produto viável (MVP), além de abordar detalhes sobre como garantir uma validação não tendenciosa e que por consequência garanta a evolução do projeto.

## **STO 3 – Fase 3 Startup One MBA ON**

### **Conteúdos da fase 3**

### **Capítulo 5 - Análise financeira**

Este capítulo traz uma análise de projeção de resultados financeiros que precisam ser monitorados para avaliar os investimentos realizados em grandes empresas ou mesmo em startups. Além disso, apresenta a ferramenta para projetar as entradas e saídas e identificar se o modelo de negócio da nossa startup.

## Capítulo 6 - Storytelling & pitches

Esse capítulo mostra como o poder de contar histórias pode auxiliar a jornada do empreendedor e o que é e como construir um bom *pitch* para sua startup.

### Atividades

#### Atividade 4 – Análise Financeira

Desafio e objetivo: usando uma planilha como modelo, realizar uma projeção financeira para compreender e validar hipóteses de negócio por meio do racional de entradas, saídas e análise de DRE projetados para cinco (5) anos.

#### Atividade 5 – Pitch

Desafio e objetivo: Desenvolver um pitch do negócio, explicando como o problema foi identificado, o tamanho deste mercado, modelo de negócio e estrutura financeira, as validações de hipóteses do problema, modelo de negócio e solução, assim como composição de equipe e definição de *roadmap* da startup. O tempo deste *pitch* é de cinco (5) minutos e deve ser gravado e entregue em vídeo.

### Atividade 6 – Entrega final (TCC)

Desafio e objetivo: Consolidar todas as entregas de atividades anteriores em um único arquivo (*template*), demonstrando todos os aspectos da jornada de desenvolvimento da startup, partindo desde a identificação do problema, evolução do modelo de negócio, solução, validação de hipóteses, análise financeira e *pitch*.

### **Mentorias em Grupo/Aula ao vivo**

#### Mentoria 5 - Financeiro

Os professores apoiam os alunos a identificar custos fixos, variáveis, linhas de produtos e serviços, precificação e projeção financeira. A partir da análise financeira, é possível identificar se o modelo de negócio e até mesmo o negócio em si pode ser promissor e escalável.

#### Mentoria 6 – Pitch e dicas para entrega final

Os professores ajudam os alunos a compreender como aplicar o conceito de *storytelling* na definição e apresentação de um *pitch*, a estruturar argumentos de venda e se prepararem para objeções que possam surgir durante apresentação para uma banca. Informações sobre as expectativas da entrega final do Startup One são passadas, para que os alunos consigam evoluir os projetos ao ponto que as avaliações acadêmicas estejam de acordo com a proposta do programa.



## Desafios para entrega final do projeto (TCC)

A partir da entrega final do projeto, o desempenho do grupo de alunos na disciplina Empreendedorismo e Inovação é avaliado segundo critérios comuns estabelecidos:

### A ideia

<b>É relevante?</b>	O problema a ser resolvido é relevante para o público alvo?
<b>É uma solução?</b>	A proposta apresentada é uma solução adequada ao problema? (Considerar o protótipo apresentado)
<b>É viável?</b>	É viável de ser aplicada? (Considerar o ambiente do mercado)
<b>Foi validado?</b>	A ideia e a solução foram validadas? (Considerar as evidências apresentadas)
<b>Aplicou o conteúdo do MBA?</b>	O conteúdo do MBA foi aplicado durante o desenvolvimento da ideia?

### Inovação e uso da tecnologia

<b>É inovador?</b>	O produto ou serviço desenvolvido é inovador?
<b>Qualidade do projeto?</b>	O projeto foi bem detalhado e com profundidade?

### Apresentação e entrega

<b>Entregáveis adequados?</b>	O grupo detalhou todos os itens obrigatórios do <i>pitch deck</i> ?
<b>Qualidade do pitch?</b>	Com Clareza/Design/Objetividade/Convencimento?

### Avaliação Geral

Qual sua avaliação geral para a solução apresentada?

A média desses critérios representam a avaliação acadêmica para a obtenção da nota final da disciplina, constituindo-se de obrigação legal ao final do ano letivo de MBA.

## Competição Startup One

Neste mesmo formulário de avaliação do projeto final há também a possibilidade de o professor indicar ou não o projeto da Startup para a competição do Startup One. Importante ressaltar que a participação na competição não tem impacto na avaliação acadêmica realizada pelos professores. O grupo de alunos também tem a opção de não participar da competição sem nenhum prejuízo na nota da disciplina Empreendedorismo e Inovação ou do Trabalho de Conclusão de Curso (TCC).

A avaliação dos projetos indicados ao “TOP30” (30 melhores projetos do ciclo) é realizada por um grupo de professores designados pela Diretoria do MBA da FIAP. Este grupo escolhe, com a utilização de critérios específicos, a seleção de trinta projetos que passarão para uma segunda fase.

Na segunda fase de avaliação, as trinta startups escolhidas internamente pela equipe de Professores FIAP são submetidas a uma banca externa de avaliação, composta por empreendedores, investidores, gestores de empresas, parceiros e demais convidados, com o intuito de isentar a avaliação e de também submeter os alunos a uma situação mais próxima da realidade do mercado (não há influência da FIAP neste processo). Os projetos selecionados compõem o TOP10 (10 melhores projetos do ciclo) que são submetidos a uma segunda fase de avaliação, recebendo mentorias e treinamentos específicos para aprimorarem seus projetos e ficarem aptos para a apresentação do projeto (*Pitch*) para uma banca externa final que escolhe a startup ganhadora da competição.