

# JORNADA

## DATA SCIENCE

## DESENVOLVA SUA AUTONOMIA ANALÍTICA

Seja reconhecido e esteja entre os melhores profissionais de dados do Brasil.



Aprenda a analisar dados com autonomia e tenha um dos perfis mais desejados e bem pagos do mercado.

## Ω Sobre o curso

Se você deseja se tornar um Cientista de Dados altamente capacitado, independente e confiante, a Jornada Ômega Data Science é a formação completa que você precisa. Com nossa metodologia exclusiva, cuidadosamente desenvolvida e validada por nossos experientes professores, você atingirá o patamar da **Autonomia Analítica**.

**Autonomia analítica** significa que você terá as habilidades necessárias para identificar a melhor estratégia de análise de dados para solucionar qualquer problema. Você será capaz de pensar de forma inovadora, adaptar soluções e conectar ideias de forma apropriada e eficaz.

Com a Jornada Ômega Data Science, você se tornará um profissional data-driven, reconhecido pelo seu perfil analítico e independente. Além disso, você terá a oportunidade de se destacar no mercado de trabalho, alcançando novas oportunidades e conquistando o sucesso que você merece! Tudo isso por um valor acessível e justo, para que você possa investir em seu futuro sem comprometer suas finanças.

Não perca mais tempo e junte-se à Jornada Ômega Data Science hoje mesmo!

## Ω Da tecnologia para a autonomia

Atualmente, muitos Cientistas de Dados estão empenhados em aprender as mais recentes e diversas tecnologias para darem o próximo passo em suas carreiras, entretanto, grande parte deles não dispõem da autonomia necessária para ir além do que as tecnologias *commodities* são capazes de oferecer. A maioria dos problemas reais de empresas reais não se encaixa facilmente nos famosos tipos de aprendizagem supervisionada e não supervisionada. Portanto, para se destacar no mundo da Ciência de Dados você precisa ir além e aprender a pensar usando dados. Isso é o que chamamos de **Autonomia Analítica!**

O objetivo da Jornada Ômega Data Science é apresentar um novo mundo para pessoas que querem analisar dados de forma **autônoma e segura**. Nós não vamos ensinar você apenas como usar as diversas ferramentas da estatística e do aprendizado de máquina existentes! Nós vamos te ensinar como elas foram criadas, como seus criadores pensaram no momento da sua criação! Isso vai elevar drasticamente o seu nível de consciência na hora de usá-las e capacitar você a pensar novas formas de uso das técnicas existentes, bem como, da necessidade de criar novas técnicas para resolver problemas reais de forma completamente inovadora.



### Nossa metodologia está baseada em três pilares



Matemática



Probabilidade  
e Estatística



Programação



## Ω Esta Jornada é para você que quer

- ✓ Evoluir exponencialmente em sua carreira.
- ✓ Resolver mais problemas em menos tempo.
- ✓ Ser reconhecido pelo seu perfil analítico/quantitativo.
- ✓ Ser mais eficiente na execução de suas tarefas.
- ✓ Elevar o valor analítico do seu trabalho.
- ✓ Fazer uma transição de carreira, mas não tem tempo para uma nova faculdade ou MBA.
- ✓ Ter a possibilidade de ganhar em dólar trabalhando de casa.

### Para quem é essa Jornada?

Pessoas que querem analisar dados de forma **autônoma e segura** para criar soluções inovadoras em qualquer área de negócio.

## Ω Projetos práticos

Para tornar o aprendizado prático durante a Jornada você desenvolverá 23 projetos que vão desde implementações computacionais até a criação e aplicação de modelos estatísticos e técnicas de aprendizagem de máquina. Os projetos foram cuidadosamente escolhidos para elevar o seu nível de **autonomia analítica** ao mesmo tempo em que compõem um diversificado portfólio para enriquecer seu curriculum e concorrer as melhores vagas de Cientistas de Dados.



### Receba Feedback personalizado dos seus projetos

Além de praticar com os projetos você poderá por meio da Comunidade Ômega Data Science no discord receber feedback personalizado sobre a sua solução além de discutir com colegas e professores possibilidades para melhorias.

### Marque sua presença nas redes sociais

Estar presente nas redes sociais demonstrando o seu conhecimento é fundamental para receber as melhores propostas de emprego! Pensando nisso, nossos projetos foram criados de forma a mostrar o seu desenvolvimento profissional. São projetos não triviais que vão realmente te destacar no mundo da Ciência de Dados evidenciando a sua **autonomia analítica**.



# PROGRAMAÇÃO PARA CIÊNCIA DE DADOS



## Aprendendo a programar

- Instalando e configurando o ambiente de trabalho.
- Estruturas de dados.
- Estruturas de programação.
- Funções.



## Importação e exportação de dados

- Leitura e escrita de arquivos de texto pleno.
- Objetos em R.
- DCF, JASON e HTML.
- Planilhas eletrônicas.
- Bancos de dados relacionais.



## Organização de dados

- Manipulação e visualização de dados em R.
- Visão geral do tidyverse.
- data.table e SQL.
- Dados tabulares com tibble.
- Importação de dados com readr.
- Arrumação de dados com tidyr.



## Manipulação de dados

- Visão geral do dplyr.
- Ordenação, filtro e seleção.
- Transformação de dados.
- Medidas descritivas, agregações e junções.
- Prática com dplyr.



## Visualização de dados

- Visualização de dados com ggplot2.
- A lógica da gramática dos gráficos.
- Análise univariada para variáveis qualitativas e discretas.
- Análise bivariada para variáveis qualitativas e discretas.
- Análise univariada para quantitativas.
- Análise bivariada para quantitativas.
- Anotações e opções de tema.
- Análise bivariada com quantitativa e qualitativa.
- Construção de mapas.



## Manipulação de datas, strings e fatores

- Expressões regulares com stringr.
- Prática de expressões regulares.
- Variáveis categóricas com forcats.
- Prática com variáveis categóricas.
- Datas com lubridate.
- Prática com datas (1).
- Prática com datas (2).
- Programação funcional com purrr.
- Prática de programação funcional.

$$\frac{x}{a} + \frac{y}{b} = 1$$

# MATEMÁTICA, PROBABILIDADE E ESTATÍSTICA PARA CIÊNCIA DE DADOS



## Cálculo Diferencial e Integral

- Funções, limites e continuidade.
- Derivadas, gradiente e hessiano.
- Dissecando a Regressão linear simples.
- Integrais.
- Aplicações.

## Álgebra Matricial

- Vetores e escalares.
- Matrizes.
- Operações com matrizes.
- Cálculo vetorial e matricial.
- Dissecando a Regressão linear múltipla.
- Matrizes esparsas.
- Sistemas lineares.
- Autovalores e autovetores.
- Dissecando a Regressão ridge.



## Métodos numéricos

- Sistemas de equações não-lineares.
- Método do gradiente descendente.
- Método de Newton.
- Diferenciação numérica.
- Extrapolação de Richardson.
- Implementações computacionais.
- Integração Numérica.
- Método de Gauss.
- Aproximação de Laplace.
- Métodos de Monte Carlo.
- Otimização numérica.
- Programação linear.
- Programação quadrática.
- Programação não-linear.



## Projetos integradores

- Criando um classificador binário.
- Implementando o algoritmo kmeans.
- Implementando uma rede neural.



## Pensando como um estatístico

- Tipos de estudos.
- A importância da amostragem.
- Estatística descritiva e exploratória.



## Probabilidades

- Definições de probabilidade.
- Probabilidade condicional.
- Teorema de Bayes e aplicações.



## Variáveis aleatórias

- Modelando a realidade.
- Tipos de variáveis aleatórias.
- Descrição probabilística de variáveis aleatórias.
- Modelos probabilísticos.
- Esperança e variância.
- Covariância e correlação.



## Inferência estatística

- O problema da inferência estatística.
- Pensando probabilisticamente.
- O pensamento frequentista.
- Aprendendo a interpretar intervalo de confiança e p-valor.
- Testando hipóteses e tomando decisões.
- Princípios da modelagem estatística.



# MODELANDO A REALIDADE E TOMANDO DECISÕES



## Modelagem estatística

- O que é um modelo de regressão?
- Regressão linear simples.
- Método de Mínimos quadrados e suas propriedades.
- Distribuição amostral.
- ANOVA.
- Construindo bandas de confiança e predição.



## Regressão linear múltipla

- Notação matricial.
- Estimacão e Inferência.
- ANOVA para regressão linear múltipla.
- Lidando com covariáveis categóricas.
- Alguns alertas no uso do modelo de regressão.



## Como sei se meu modelo funciona?

- Análise de diagnóstico.
- Análise de resíduos.
- Métodos gráficos.
- Testes de hipóteses.
- Detectando observações atípicas.
- Pontos de alavancagem e influência.
- Multicolinearidade.



## Medidas corretivas e seleção de modelos

- Transformação e ponderação.
- Regressão robusta.
- Critérios de qualidade de ajuste.
- Escolhendo o melhor modelo.
- Algoritmos para seleção de modelos.



## Extensões do modelo de regressão

- Modelos aditivos generalizados.
- Regressão com efeitos suaves.
- Prophet (inteligência artificial do facebook).
- Construindo um sistema de monitoramento de vendas.



## Modelos de regressão para dados binários e extensões

- Regressão logística.
- Especificação, estimacão e inferência.
- Interpretando os coeficientes de regressão.
- Diagnóstico e seleção de modelos.
- Estudo de caso - Breast cancer detection.
- Regressão com penalização.



## Modelos de regressão para dados de contagem e extensões

- Modelos para dados de contagem.
- Especificação, estimacão e inferência.
- Diagnóstico e seleção de modelos.
- Modelos para superdispersão.
- Modelos com excesso de zeros.
- Estudo de caso - Validando campanhas de marketing.



## Aprendizado de máquina

- Tipos de aprendizado.
- Tarefas comuns em ML.
- Técnicas de validação.
- O balanço vício e variância.



## Pré-processamento de dados

- Organização e qualidade de dados.
- Dados faltantes.
- Dados desbalanceados.
- Transformação de dados.
- Padronização de escalas.
- Partição em treino, teste e validação.
- Métodos de seleção de covariáveis.



## Algoritmos de aprendizagem supervisionada

- Regressão linear.
- Regressão logística.
- K vizinhos mais próximos.
- Naive Bayes.
- Análise discriminante.
- Árvores de classificação e regressão.
- Métodos ensembles.
- Random forest.
- XGBoost.
- Máquinas de vetores de suporte.
- Regressão com regularização.
- Modelos supervisionados com o caret.



## Algoritmos de aprendizagem não supervisionada

- Motivação e intuição dos componentes principais (PCA).
- A matemática dos componentes principais (PCA).
- Implementação e aplicações de redução de dimensão usando PCA.
- Análise de agrupamentos.
- Medidas de dissimilaridades.
- Métodos de agrupamento.
- Avaliação da qualidade do agrupamento.
- Implementação computacional e aplicações.



## Suporte

Oferecemos suporte diferenciado para a Jornada Ômega Data Science contendo:



### Fórum no Discord

onde os alunos e professores podem discutir sobre assuntos dos cursos e dos projetos.



### Tira dúvida em grupos

via nossos grupos no Telegram e Discord.



## Correção dos projetos para construir seu portfólio



Aluno conclui o projeto e encaminha via Discord para os professores.



Professor avalia o projeto e sugere melhorias.



Após o feedback do professor, aluno pública em suas redes sociais.

## Com quem você irá aprender?

A Ômega Data Academy conta com professores altamente qualificados com pesquisa na área de Ciência de Dados e diversas contribuições para a literatura especializada. Em particular a Jornada Ômega Data Science é ministrada pelos professores Walmes Marques Zeviani e Wagner Hugo Bonat.

### Walmes Marques Zeviani



Doutor em Estatística & Experimentação Agropecuária, professor do Departamento de Estatística desde 2010 & programador R. Tem experiência no ensino de Estatística e Data Science, já ministrou inúmeros cursos de R em instituições de ensino e pesquisa desde 2008.

[Lattes](#)

[Linkedin](#)

### Wagner Hugo Bonat



Doutor em Matemática, professor do Departamento de Estatística desde 2010 & programador R. Tem experiência no ensino de Estatística e Data Science e atua como consultor desenvolvendo soluções em R. É coordenador do programa de Especialização em Data Science & Big Data da UFPR. É membro efetivo de programas de pós-graduação com orientações de mestrado e doutorado.

[Lattes](#)

[Linkedin](#)

## Sobre a Ômega Data Academy

A Ômega Data Academy é uma empresa inovadora da área de educação online que busca formar Cientistas de Dados com forte fundamentação estatística e matemática alinhados com as demandas atuais e futuras do mercado de dados.

## Nossa missão

Nossa missão é impactar carreiras, profissionais e negócios com uma formação íntegra e fundamentada, com ensino humano, didático e aplicado orientado aos resultados e autonomia profissional do aluno.

[escola.omegadatascience.com.br](https://escola.omegadatascience.com.br)

