



Especialista em Ciência de Dados para Economia e Finanças

Programa da Formação

Vítor Wilher, Cientista-Chefe da Análise Macro

Resumo

O objetivo desse documento é o de apresentar uma formação completa de conhecimento para pessoas interessadas em aprender a utilizar do absoluto ZERO ferramentas de ciência de dados nas áreas de economia e finanças.

Sumário

1	Objetivo	2
2	Cursos	2
2.1	Fundamentos de Análise de Dados	2
2.2	Programação em R para Análise de Dados	3
2.3	Programação em Python para Análise de Dados	4
2.4	SQL para Economia e Finanças	5
2.5	Produção de Relatórios Automáticos	5
2.6	Produção de Dashboards Automáticos	6
2.7	Estatística para Análise de Dados	6
2.8	Análise de Dados Macroeconômicos e Financeiros	7
3	Certificação	7

1 Objetivo

Torne-se um Cientista de Dados capaz de coletar, tratar, analisar e apresentar dados econômicos e financeiros. Seja treinado nas linguagens de programação mais utilizadas no mercado e aprenda a construir DO ZERO relatórios e dashboards automáticos.

2 Cursos

A Formação em Ciência de Dados para Economia e Finanças será composta por uma sequência de 8 Cursos que se propõem em conjunto a ir do ZERO ao que há de mais moderno em ciência de dados com exemplos das áreas de economia e finanças. Somados, os módulos dos Cursos terão mais de 168h de videoaulas. Veja abaixo, o conteúdo de todos os Cursos:

1. Fundamentos de Análise de Dados
2. Programação em R para Análise de Dados
3. Programação em Python para Análise de Dados
4. SQL para Economia e Finanças
5. Produção de Relatórios
6. Produção de Dashboards
7. Estatística usando R e Python
8. Análise de Dados Macroeconômicos e Financeiros

Essa sequência de cursos foi pensada para que o aluno consiga cumprir as etapas de coleta, tratamento, modelagem e apresentação de dados de forma plena e autônoma. O programa completo dos Cursos é detalhado na sequência.

2.1 Fundamentos de Análise de Dados

O objetivo do curso de Fundamentos de Análise de Dados é o de fornecer uma **porta de entrada** a estudantes e profissionais de mercado que queiram adicionar conhecimentos de ciência de dados aos seus currículos. O programa completo do curso pode ser visto abaixo:

1. Afinal, o que é ciência de dados?
2. Habilidades requeridas pela ciência de dados
3. Como adicionar ciência de dados ao currículo e se tornar um profissional disputado no mercado?
4. Por que o Excel não basta?
5. O que é uma linguagem de programação e por que eu preciso aprender isso?
6. Introdução à linguagem R
7. Introdução à linguagem Python
8. O que são banco de dados?
9. Introdução à linguagem SQL
10. Importação: onde estão os dados?
11. Limpando os dados: por que a vida não é simples?
12. Análise exploratória de dados: usando estatística simples para entender o que está acontecendo com a sua base de dados
13. O que são modelos
14. Aprendendo a fazer uma regressão linear
15. Como apresentar os resultados da sua análise de dados?

16. Construindo sua primeira análise de dados do ZERO
17. Como continuar aprendendo?

2.2 Programação em R para Análise de Dados

O Curso de Programação em R para Análise de Dados percorrerá todo o processo de análise de dados, desde a instalação de programas e pacotes até a automatização das etapas de coleta, limpeza e visualização de dados.

Dividido em 42 seções temáticas, o Curso irá capacitar estudantes e profissionais sem experiência em programação a sair do zero e construir análises práticas que resolvam os seus desafios diários envolvendo dados. Abaixo, o programa detalhado:

1. Por que usar uma linguagem de programação para analisar dados?
2. Instalando os programas necessários
3. Instalando pacotes e entendendo as mensagens no Console
4. Introdução ao mundo tidyverse
5. Criando scripts
6. Lidando com objetos no R
7. Estruturas de Dados (data frames, listas, matrizes, vetores)
8. Operadores lógicos e aritméticos
9. O operador pipe
10. Funções
11. Estruturas de repetição e de decisão
12. Introdução à coleta de dados
13. Trabalhando com arquivos .csv
14. Trabalhando com arquivos Excel
15. O que são APIs e como utilizá-las?
16. Consumindo dados brasileiros
17. Trabalhando com SQL
18. Trabalhando com Google Sheets
19. Trabalhando com Google Drive
20. Trabalhando com arquivos JSON
21. Trabalhando com arquivos XML
22. Exportando dados
23. Introdução à Web Scraping
24. Introdução à limpeza de dados
25. Tratamento de dados textuais
26. Renomeando colunas
27. Seleccionando colunas
28. Como criar e alterar colunas
29. Como aplicar filtros em linhas
30. Como fazer operações por grupos de dados
31. Como sumarizar os dados
32. Formato long vs wide
33. Como fazer cruzamento de dados
34. Lidando com datas no R
35. Identificando e tratando dados faltantes
36. Como deflacionar uma série

37. Como fazer ajuste sazonal
38. Introdução à visualização de dados e *storytelling*
39. Produzindo gráficos com ggplot2
40. Produzindo relatórios com RMarkdown
41. Produzindo dashboards
42. Como automatizar o processo de análise de dados?

2.3 Programação em Python para Análise de Dados

O Curso de Programação em Python para Análise de Dados percorrerá todo o processo de análise de dados, desde a instalação de programas e pacotes até a automatização das etapas de coleta, limpeza e visualização de dados.

Dividido em 42 seções temáticas, o Curso irá capacitar estudantes e profissionais sem experiência em programação a sair do zero e construir análises práticas que resolvam os seus desafios diários envolvendo dados. Abaixo, o programa detalhado:

1. Por que usar uma linguagem de programação para analisar dados?
2. Escolhendo uma IDE e trabalhando com o Google Colab
3. Instalando pacotes e entendendo as mensagens
4. Introdução ao pandas e numpy
5. Criando scripts
6. Lidando com objetos no Python
7. Estruturas de Dados (data frames, listas, matrizes, vetores)
8. Operadores lógicos
9. operadores aritméticos
10. Funções
11. Estruturas de repetição e de decisão
12. Introdução à coleta de dados
13. Trabalhando com arquivos .csv
14. Trabalhando com arquivos Excel
15. O que são APIs e como utilizá-las?
16. Consumindo dados brasileiros
17. Trabalhando com SQL
18. Trabalhando com Google Sheets
19. Trabalhando com Google Drive
20. Trabalhando com arquivos JSON
21. Trabalhando com arquivos XML
22. Exportando dados
23. Introdução à Web Scraping
24. Introdução à limpeza de dados
25. Tratamento de dados textuais
26. Renomeando colunas
27. Seleccionando colunas
28. Como criar e alterar colunas
29. Como aplicar filtros em linhas
30. Como fazer operações por grupos de dados
31. Como sumarizar os dados
32. Formato long vs wide

33. Como fazer cruzamento de dados
34. Lidando com datas no Python
35. Identificando e tratando dados faltantes
36. Como deflacionar uma série
37. Como fazer ajuste sazonal
38. Introdução à visualização de dados e *storytelling*
39. Produzindo gráficos
40. Produzindo relatórios
41. Produzindo dashboards
42. Como automatizar o processo de análise de dados?

2.4 SQL para Economia e Finanças

O objetivo da disciplina é o de fazer uma introdução a bancos de dados e à linguagem SQL, tendo como pano de fundo exemplos envolvendo dados econômicos e financeiros. O programa completo do Curso segue abaixo:

1. Por que aprender SQL?
2. Bancos de dados
3. SQLite
4. Select
5. Where
6. Group by e Order by
7. Instruções Case
8. Join
9. Design de banco de dados
10. Gerenciando dados
11. Construindo um banco de dados econômicos e financeiros

2.5 Produção de Relatórios Automáticos

Aprenda a criar relatórios e apresentações do zero através de uma das linguagens de programação mais utilizadas no mercado, R ou Python. O curso é totalmente prático e aborda a produção de relatórios e apresentações do zero: da extração dos dados, passando pelo desenvolvimento do documento até a compilação. Serão desenvolvidos projetos de curso aplicados de relatórios e apresentações com as linguagens de programação R e Python. A seguir, o programa detalhado:

Aula 01 - Boas vindas e apresentação do curso Aula 02 - Relatórios no ciclo de análise de dados Aula 03 - Como produzir relatórios do zero? Aula 04 - Quais ferramentas usar para produção de relatórios? Aula 05 - Preparando o ambiente de programação Posit Cloud Aula 06 - Projetos de curso de relatórios e apresentações Aula 07 - Guia de primeiros passos usando o Quarto Aula 08 - Como formatar o layout do relatório? Aula 09 - Como formatar o conteúdo do relatório? Aula 10 - Como criar apresentações com o Quarto? Aula 11 - Como formatar o layout da apresentação? Aula 12 - Como formatar o conteúdo da apresentação? Aula 13 - Projeto de curso 1: Relatório de Análise da Inflação Aula 14 - Projeto de curso 2: Apresentação sobre DataViz Aula 15 - Como funciona a automatização de um código? Aula 16 - Qual é o maior desafio para automatização de um código? Aula 17 - Quais ferramentas podem ser usadas para automatização? Aula 18 - Automatizando um projeto de curso de relatório do curso Aula 19 - Monitoramento, erros e manutenção

2.6 Produção de Dashboards Automáticos

Aprenda a criar dashboards estáticas e interativas do zero com as linguagens de programação R ou Python. Você irá escolher qual das duas linguagens utilizar no seu projeto. O curso é totalmente prático e aborda a produção de dashboards estáticas e interativas do zero: da extração dos dados, passando pelo desenvolvimento da dashboard até a publicação. Serão desenvolvidos projetos de curso aplicados de dashboards com as linguagens de programação R e Python. A seguir, o programa detalhado:

Aula 01 - Boas vindas e apresentação do curso Aula 02 - Dashboards no ciclo de análise de dados
Aula 03 - Como produzir dashboards do zero? Aula 04 - Quais ferramentas usar para produção de dashboards? Aula 05 - Preparando o ambiente de programação Aula 06 - Projetos de Curso de Dashboards Aula 07 - Produzindo Dashboards Estáticas Aula 08 - Produzindo Dashboards Dinâmicas Aula 09 - Projetos de Curso de Dashboards Projeto de Curso 1: Dashboard de Análise da Inflação Projeto de Curso 2: Dashboard de Economia Internacional Aula 10 - Publicação de Dashboards O que é o deploy de uma dashboard? Opções e serviços disponíveis: Shinyapps.io e github pages Deploy de uma dashboard no Shinyapps.io - Conta de usuário - Instalação de pacotes - Configuração de projeto e token - Processo de publicação - Testes, erros e soluções comuns
Aula 11 - Como funciona a automatização de um código? Aula 12 - Qual é o maior desafio para automatização de um código? Aula 13 - Quais ferramentas podem ser usadas para automatização? Aula 14 - Automatizando um projeto de curso de dashboard Aula 15 - Monitoramento, erros e manutenção

2.7 Estatística para Análise de Dados

O curso se dividirá em 13 seções agrupadas em três grandes áreas. Na primeira parte é vista uma introdução à exploração de dados, onde serão destrinchados conceitos estatísticos básicos. Na segunda parte será abordada a inferência estatística, enquanto na última parte serão introduzidos modelos estatísticos básicos. O diferencial do curso está na intensa aplicação do R e do Python em todas as seções do Curso.

As aulas serão divididas entre aulas teóricas, com a discussão de conceitos estatísticos e aplicações em R e Python dos conceitos vistos. Além disso, também serão propostos laboratórios para que os alunos pratiquem os conceitos vistos em ambas as linguagens.

Espera-se que ao final do Curso o aluno esteja apto a aplicar estatísticas descritivas, análise exploratória de dados, inferência estatística e modelagem básica tanto em R quanto em Python. O programa detalhado pode ser visto abaixo:

1. Preparando o ambiente;
2. Dados Univariados: vetores, estatísticas descritivas e dados categóricos;
3. Dados Bivariados: amostras independentes, estruturas de dados no R/Python e dados emparelhados;
4. Dados Multivariados: data frames, aplicando funções sobre dados e dados externos;
5. Gráficos Multivariados: gráficos básicos, ggplot2 e matplotlib;
6. População: variáveis aleatórias discretas e diferentes tipos de distribuição;
7. Inferência Estatística;
8. Intervalos de Confiança;
9. Testes de Significância;
10. Qualidade do Ajuste;

11. Regressão Linear;
12. Análise de Variância;
13. Extensões: modelos lineares generalizados e modelos não lineares.

2.8 Análise de Dados Macroeconômicos e Financeiros

O objetivo do Curso é proporcionar aos alunos uma imersão no mundo dos dados econômicos e financeiros através do uso de uma linguagem de programação. Ao longo do Curso, o aluno aprenderá a fazer coleta automatizada de dados a partir de diversas fontes como SIDRA/IBGE, Banco Central, Ipeadata, OCDE, FMI, Banco Mundial, dentre outros. Além disso, o aluno aprenderá a construir scripts que limpam esses dados e constroem insights interessantes sobre a evolução das séries ao longo do tempo. Por fim, será apresentada uma metodologia geral do processo de modelagem e previsão de variáveis econômicas e financeiras.

Ao longo do Curso, o aluno terá contato com séries macroeconômicas como PIB, IPCA, Concessões de Crédito, IGPs, CPI norte-americano, IBC-Br, Vendas no Varejo, Serviços, Produção industrial, dentre outras. Entre as séries financeiras que serão vistas, estão o CDI, Selic, juros reais, juros de Títulos Públicos, Câmbio, CDS, índices de commodities, dentre outras.

O programa completo do Curso:

1. Por que usar uma linguagem de programação para analisar dados econômicos e financeiros?
2. Coletando dados do SIDRA, Banco Central, Ipeadata, OCDE, FMI, Banco Mundial e outras fontes;
3. Fazendo limpeza nos dados e desenvolvendo scripts automatizados;
4. Análise exploratória de dados: sazonalidade, estacionariedade, outliers, boxplots, histogramas, densidade e outras coisas;
5. Vamos modelar: como construir um modelo de séries temporais? Um resumo das principais abordagens disponíveis (Modelos ARIMA, Regressões, VAR, VEC, ARCH/GARCH e outros);
6. Testar, treinar e avaliar: como construir um modelo de previsão para séries temporais? Desenvolvendo uma metodologia completa de previsão de variáveis econômicas e financeiras
7. Projeto de Curso 1: construção de modelo de previsão para a inflação brasileira medida pelo IPCA
8. Projeto de Curso 2: Análise de impulso resposta em séries financeiras
9. Projeto de Curso 3: construindo uma Dashboard em shiny do início ao fim

O curso fecha a primeira parte da Formação em Ciência de Dados para Economia e Finanças, possibilitando ao aluno aplicar todos os conhecimentos aprendidos ao longo das disciplinas anteriores em projetos reais, que podem ser adicionados ao seu portfólio pessoal, servindo de cartão de visitas para disputar empregos que exigem conhecimentos de análise de dados e de programação.

3 Certificação

O aluno terá acesso a 8 Certificados ao longo da Formação Especialista em Ciências de Dados para Economia e Finanças, totalizando uma carga horária de 540h, que é composta pelas videoaulas e atividades necessárias para concluir os Cursos.

- Carga horária total: 540h
- Total de horas em vídeo: 168h

- Número de aulas: 182