

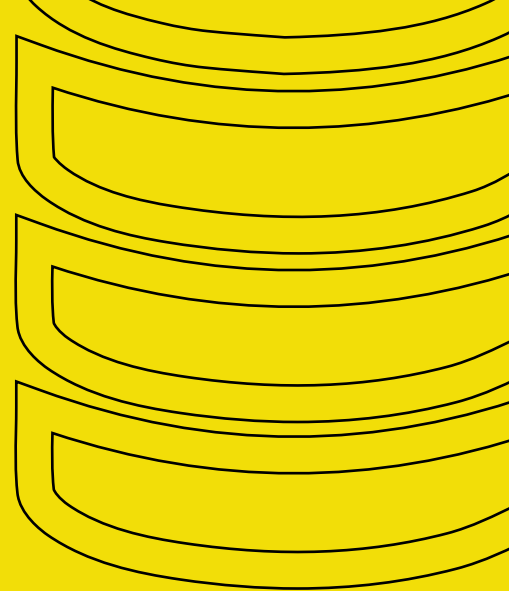
# DATA EXPERT

DATA SCIENCE & MACHINE LEARNING

2022  
**EMENTA  
DO CURSO**



# SUMÁRIO

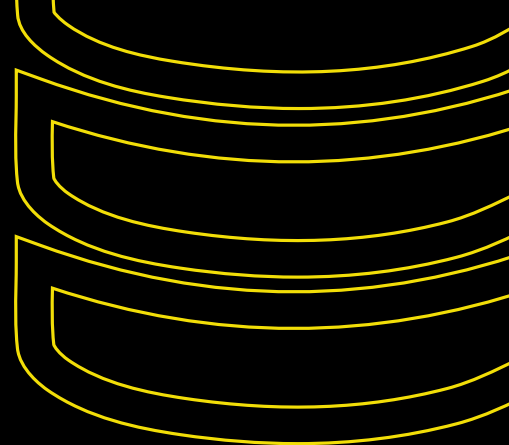


**CLIQUE PARA IR  
ATÉ O CONTEÚDO**

- PÁG. 03** Bem-vindo à DNC Group
- PÁG. 04** Nossa metodologia
- PÁG. 05** Pra quem é o curso?
- PÁG. 06** Após o curso você será capaz de...
- PÁG. 07** Sobre o Data Expert
- PÁG. 08** Cronograma do curso
- PÁG. 09** Modelo de estudos
- PÁG. 10** Projetos do curso
- PÁG. 11** Empresas Parceiras
- PÁG. 12** Regras de avaliação
- PÁG. 13** Pontuação
- PÁG. 14** O que eu ganho ao fazer o curso?
- PÁG. 15** Além disso oferecemos...
- PÁG. 16** Por que Data Science?
- PÁG. 17** Conteúdo Programático
- PÁG. 27** Certificações
- PÁG. 28** Perguntas Frequentes



## BEM-VINDO À DNC GROUP



Somos uma empresa focada em educação complementar que busca entregar o que não é aprendido nas instituições tradicionais.

Nossos exemplos não são de livros! Nós contamos o que vivenciamos em projetos.

Alunos botando a mão na massa, orientados por especialistas realizando projetos para aprender na prática e recebendo suporte pós curso para consolidar o aprendizado - essa é a nossa forma de ensinar!

### **NOSSO MAIOR OBJETIVO É O SUCESSO DE NOSSOS ALUNOS!**

**HARDWORK  
SMILES  
PLURAL  
2ND HOME  
GOODWILL**

Nosso propósito não nos deixa parados e nos leva a sempre buscar inovação.

Vivemos nosso propósito no dia a dia e não há nada que supere os sorrisos dos alunos que impactamos.

Diferentes ideias, gêneros, raças, opiniões e histórias. Isso trará o equilíbrio. Isso trará resultado.

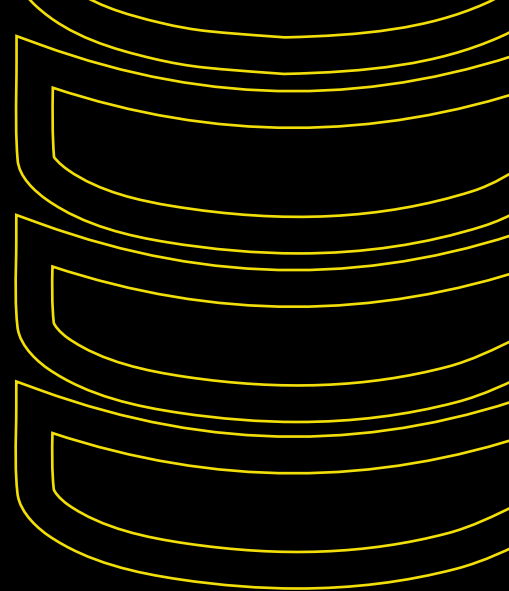
Aqui, se você é a média das 5 pessoas com quem mais convive, nós aumentamos essa média.

Buscamos aprender e ensinar; sabemos que ninguém sabe tudo; temos boa vontade; paciência e humildade.



## **NOSSA METODOLOGIA**

Nossa metodologia é nosso diferencial!  
Nós acreditamos que é possível revolucionar a educação de maneira prática e consistente.



### **HANDS ON**

Nosso lema é “A revolução é a Prática”, então todos os nossos treinamentos possuem um approach muito focado na aplicabilidade do conhecimento.

### **PROJETOS REAIS**

Em todos os nossos treinamentos, os alunos desenvolvem um projeto real paralelamente ao curso, para aplicar tudo o que aprenderam nas aulas.



### **DINÂMICAS COM CASES REAIS**

Em todos os exemplos dados no curso, utilizamos a experiência de nossos consultores como um guia para que os alunos aprendam como os conceitos são aplicados na vida real.

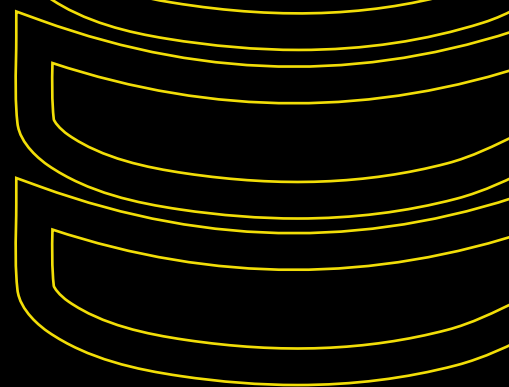
### **ACOMPANHAMENTO PÓS CURSO**

O nosso time de Sucesso do Cliente está focado em atingir os objetivos profissionais dos alunos após o fim do curso, com mentorias e acompanhamento individual.



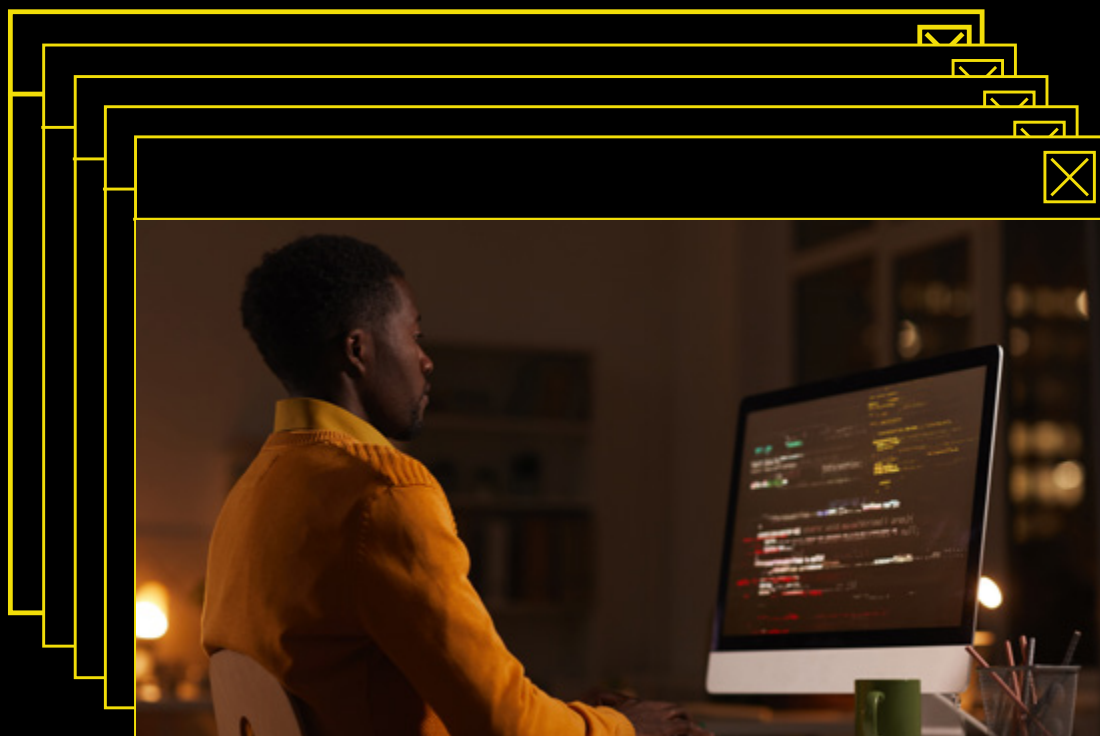


## PRA QUEM É O CURSO?



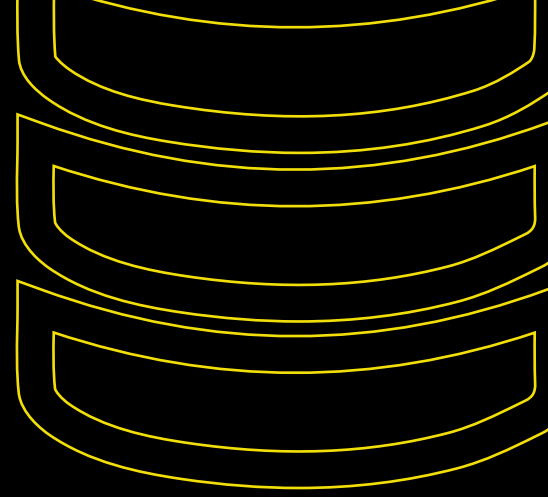
Este curso é recomendado para profissionais que nunca tiveram contato com a área de dados, mas queiram trabalhar com análises e criação de modelos preditivos. Para profissionais do mercado que desejam atuar com análises estatísticas e implementação de algoritmos.

Também é um curso para aqueles que trabalham na área de Tecnologia da Informação e queiram aumentar seu repertório com conhecimentos em Machine Learning. Para Engenheiros, Administradores, Economistas e Analistas de Negócios de segmentos diversos que utilizam dados para tomada de decisão e queiram expandir seu conhecimento nessas análises.





## APÓS O CURSO VOCÊ SERÁ CAPAZ DE:



- Realizar análises preditivas;
- Entender a estrutura de grande banco de dados assim como conectá-los;
- Realizar análises estatísticas;
- Entender comportamento de fenômenos;
- Criar algoritmos através de Machine Learning;
- Gerar insights para seu negócio;
- Compreender linguagem Python;
- Organizar e realizar testes A/B;
- Criar um sistema de recomendação;
- Representar dados na forma de painéis;
- Trabalhar com as bibliotecas Pandas e NumPy;
- Comparar modelos;
- Entender os principais conceitos de Big Data e Spark;
- Saber sobre bancos de dados SQL, como consumir dados, como escrever queries e importar no notebook Jupyter;
- Saber interpretar, limpar os dados e selecionar features importantes para os modelos.
- Definir hiperparâmetros do modelo.





## **SOBRE O DATA EXPERT**

É um programa híbrido de 520 h que tem como principal objetivo a formação de cientistas de dados. Todos os módulos do curso foram baseados no Framework CRISP-DM que é utilizado para execução de projetos de Data Science.

O curso é ministrado pela dnc.group, por meio de aulas expositivas dinâmicas de integração de conhecimento e o desenvolvimento de um projeto durante quase quatro meses com acompanhamento e mentorias de consultores experientes no mercado.

  
**205H**  
**AULAS ONLINE**

  
**80H**  
**AULAS AO VIVO**

  
**175H**  
**PROJETO REAL**

  
**60H**  
**MENTORIA ABERTA**



# CRONOGRAMA DO CURSO

## DAYS

Encontros EAD para realização de dinâmicas de integração de conhecimento e direcionamento geral dos projetos.

## SLOTS

Intervalos entre os DAYS nos quais serão liberadas as aulas online referentes ao conteúdo teórico do curso, com tarefas e entregas a serem submetidos até o próximo encontro.







## MODELO DE ESTUDOS

A parte EAD do treinamento é dada em **14 encontros** espaçados ao longo do programa.

O curso é dividido em **13 slots** sendo que o conteúdo de cada slot deve ser estudado nos intervalos entre os encontros.

Para garantir o máximo aprendizado e aproveitamento de toda a turma, na etapa online será cobrado:

✕

### ASSIGNMENTS

Exercícios individuais por slot.

✕

### DELIVERABLES

Entregas de projeto,  
sendo 1 por slot

✕

### PROVAS

1 por certificação  
intermediária

Para cumprir com essa demanda será necessário ter uma **dedicação** média das **Aulas Online** de 45 a 50 minutos por dia (variável de acordo com cada módulo).

Além disso, deverá ser desenvolvido um **projeto** com uma dedicação de 175 horas até o final do curso. A garantia de um ótimo aproveitamento do curso e do projeto pode ser resumida em 1 palavra:

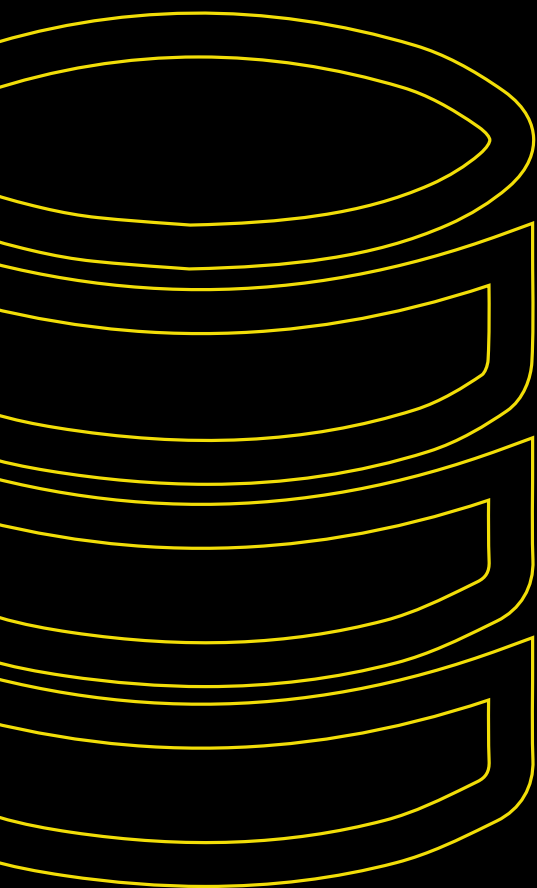
**Comprometimento!** Iremos entregar um conteúdo sólido e bastante denso que necessitará dedicação para absorvê-lo completamente.



## PROJETOS DO CURSO

Os alunos do Data Expert desenvolverão um portfólio de 4 projetos durante o treinamento, referente a problemas reais a serem resolvidos, sendo 3 projetos intermediários e um projeto completo (Capstone).

Os projetos intermediários serão ciclos menores e realizados individualmente enquanto o projeto completo (Capstone) será realizado em grupo durante todo o período do curso.





## EMPRESAS PARCEIRAS

Nossos alunos já desenvolveram projetos de alto impacto em **mais de 200 empresas** em diferentes segmentos do mercado. Confira algumas delas, abaixo:

99jobs.com

ifood

ambev

BASF  
We create chemistry

BAYER

cora

dm  
card

endeavor

Eurofarma

Johnson & Johnson

club  
petro

stone

Nestlé Faz Bem

hand talk

reckitt



Smurfit Kappa

TELE  
CINE

tembici

Unilever

XP inc.



## REGRAS DE AVALIAÇÃO

Durante o treinamento, a avaliação ocorre a todo momento nos seguintes quesitos:

■ PARTICIPAÇÃO DOS ENCONTROS AO VIVO

■ CUMPRIMENTO DOS DELIVERABLES DE PROJETOS

■ CUMPRIMENTO DOS ASSIGNMENTS

■ TRABALHO EM EQUIPE E DEDICAÇÃO

■ PERFORMANCE NAS DINÂMICAS DE GRUPO E BUSINESS CASES

■ RESULTADO DO PROJETO



O não cumprimento de **50% dos assignments e entrega de 100% dos deliverables** no prazo estipulado significa um **desligamento automático** do treinamento.

Além disso, para evitar o desligamento, é necessário ter tirado, no mínimo, **75% dos certificados intermediários** e ter **75% de presença** referente aos dias ao vivo.



## **A PONTUAÇÃO SERÁ FEITA DA SEGUINTE FORMA**

### **ONLINE**

Nota dada aos assignments e deliverables baseado na qualidade e prazo da entrega

### **DIAS AO VIVO**

Nota dada às dinâmicas de grupo

### **APRESENTAÇÃO DO PROJETO**

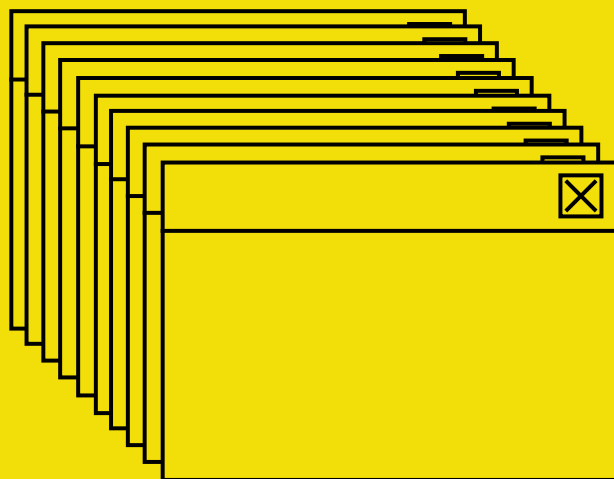
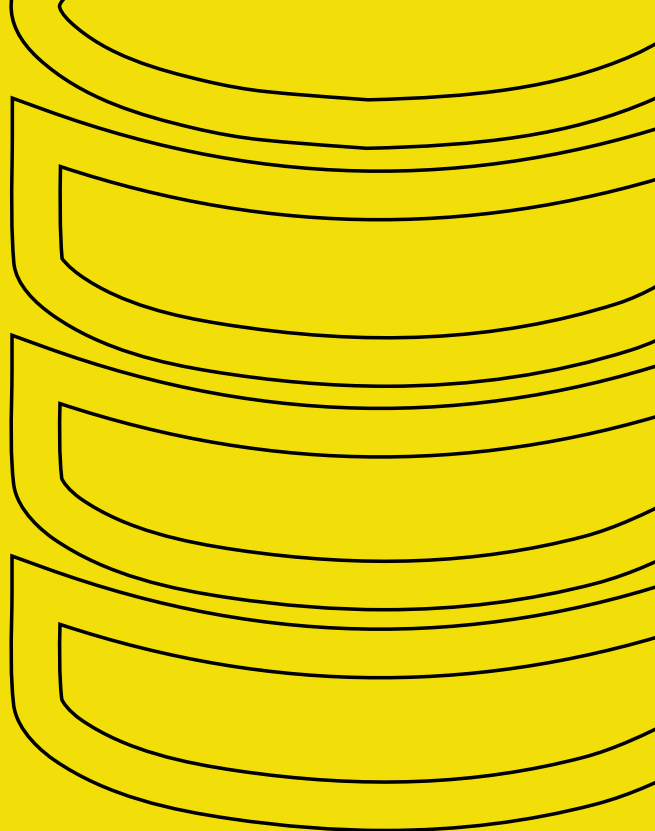
Todos os projetos serão avaliados pela banca

### **CERTIFICAÇÕES INTERMEDIÁRIAS**

Nota dada às certificações concluídas

### **HACKATHONS E SUPER CHALLENGE**

Nota dada ao vencedor dos desafios realizados nos encontros ao vivo





## O QUE EU GANHO AO FAZER O CURSO?

Queremos promover uma experiência PRÁTICA e não apenas uma base teórica forte.

Por isso integramos a forte base de negócios com muitos exemplos reais para que cada aluno desenvolva os projetos.



- CURSO COM PROJETOS REAIS
- EXPERIÊNCIA NO CV
- TRABALHO EM EQUIPE
- METODOLOGIA INOVADORA



## **ALÉM DISSO OFERECEMOS:**

### **CERTIFICADO**

Você receberá um certificado do treinamento para enriquecer o seu currículo

### **REFAÇA O CURSO**

Somente aqui você pode refazer o curso quantas vezes quiser por prazo indeterminado

### **PÓS-TREINAMENTO**

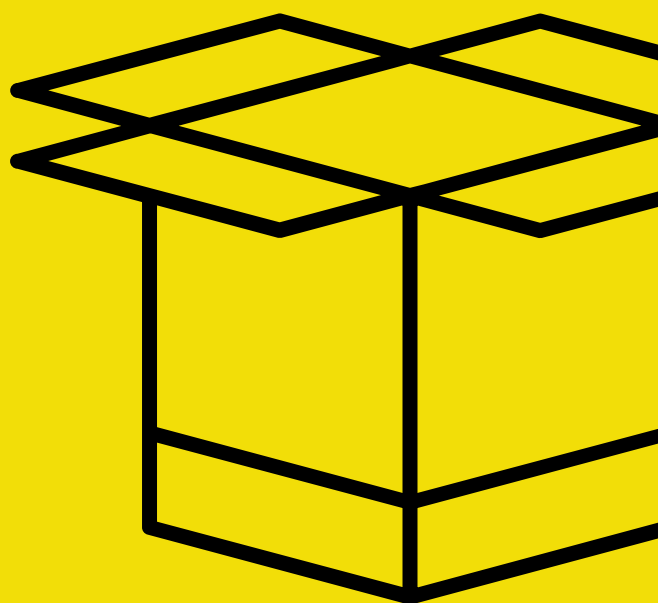
Conte sempre com nossa equipe de suporte e consultoria para tirar dúvidas

### **APLICAÇÕES REAIS**

Todos os desafios dos treinamentos são tirados de casos reais para simular problemas reais

### **MATERIAL DE APOIO ONLINE**

Temos vários materiais de apoio e acompanhamento para o seu aprendizado ser efetivo



### **ACESSO À PLATAFORMA**

Você tem acesso a um ambiente online com diversos materiais exclusivos, descontos e vagas exclusivas

### **SEJA CONTRATADO**

Grandes empresas procuram agora por profissionais como você pela nossa plataforma do Hiring, você topa?

### **HARDWORK TRIBE**

Temos uma forte comunidade de ex-alunos nos mais diversos segmentos de mercado, áreas e cargos. Utilize esta rede de contatos para aumentar seu networking e impulsionar ainda mais sua carreira!



## POR QUE DATA SCIENCE?

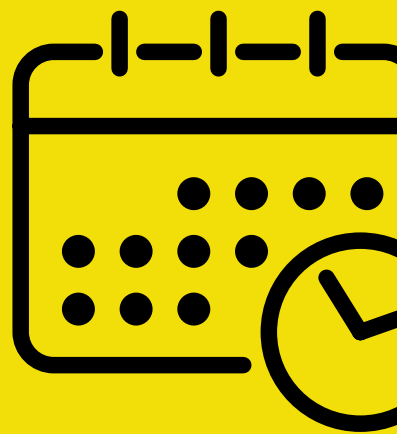
- A profissão de Cientista de Dados possui demanda crescente: De acordo com um relatório da McKinsey & Company, a partir de 2018, os EUA terão cerca de 140.000 a 180.000 cientistas de dados a menos do que o necessário;
- Salários imbatíveis De acordo com o Glassdoor, em 2016 a ciência de dados foi o campo mais bem pago para entrar, com um salário inicial de R\$6.000,00;
- Eles agregam valor ao negócio: ciência de dados está florescendo em todos os campos dos negócios, de TI a saúde, de comércio eletrônico a marketing e varejo. Dado que os dados são o ativo mais valioso para uma empresa, os cientistas de dados desempenham um papel muito importante, pois atuam como consultores de confiança e parceiros estratégicos para sua gestão.







# CONTEÚDO PROGRAMÁTICO



## SLOT 1



Onboarding do Curso;  
Python Zero;  
Introdução a Projetos de  
Ciência de Dados;  
Scrum I: Introdução;  
Scrum I: Papéis;  
Carreira I: Construção da  
Jornada;

## SLOT 3



Estatística Descritiva;  
SQL - Linguagem de Banco  
de Dados;  
Metabase: Visualização de  
Dados com SQL;  
Scrum III: Épicos e Stories;  
Scrum III: Gráficos e Reports;

## SLOT 2



CRISP-DM: Business  
Understanding e Data  
Understanding;  
Bibliotecas com Python;  
Ambientes (Anaconda e GIT);  
Scrum II: Eventos;  
Scrum II: Artefatos;  
Carreira II: O Mercado na Área de  
Ciência de Dados;

## SLOT 4



Data Storytelling;  
Mongo DB;  
Estatística Inferencial;

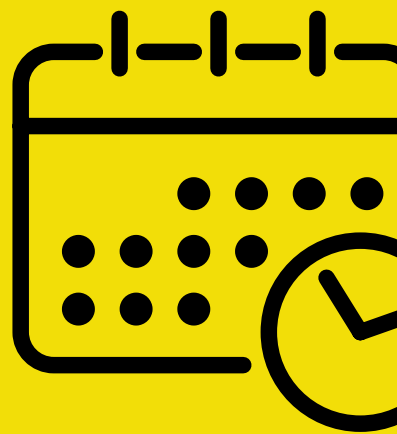
## SLOT 5



CRISP-DM: Data Preparation;  
Introdução a Machine Learning;  
Data Cleaning;  
Carreira III: Frameworks para  
Desenvolver Soft Skills;



# CONTEÚDO PROGRAMÁTICO



## SLOT 6



Data Wrangling;  
Estatística Avançada;

## SLOT 8



Classificação: Modelos  
Supervisionados de Machine  
Learning

## SLOT 10



Recomendação: Modelos Não-  
Supervisionados de Machine  
Learning;  
CRISP-DM: Evaluation;

## SLOT 12



Imbalanced Data Set;  
Implantação de Modelos;  
Modelos Disponíveis em Cloud;  
Carreira IV: Preparação para  
Processos Seletivos;

## SLOT 7



CRISP-DM: Modelling;  
Regressão: Modelos  
Supervisionados de Machine  
Learning;

## SLOT 9



Clusterização: Modelos Não-  
Supervisionados de Machine  
Learning;  
Apresentação de Projetos - Parte 1;

## SLOT 11



CRISP-DM: Deployment;  
Introdução a Big Data;

## SLOT 13



Replicação;  
Segurança e LGPD;  
Plataformas de Data Science (AWS  
e Sage Maker);  
Apresentação - Parte 2;



## **SLOT 1 PYTHON ZERO**

- Variáveis;
- Operações;
- Condicionais;
- Loops;
- Funções;
- Classes;
- Listas, Tuplas e Dicionários;
- Tratamento de erros;
- Scripts.



## **SLOT 1 INTRODUÇÃO A PROJETOS DE CIÊNCIA DE DADOS**

- Dados, Cultura e Técnica;
- Introdução a CRISP-DM;
- Metodologias Científicas;
- Machine Learning Canvas.



## **SLOT 1 CARREIRA I: CONSTRUÇÃO DA JORNADA**

- Jornada de carreira;
- Trilhas e áreas;
- Multinacional x startup;
- Roadmap;
- CV e LinkedIn;
- O que não fazer.



## **SLOT 1 SCRUM I: INTRODUÇÃO**

- Detalhes do framework DNC;
- Visão geral do processo;
- Introdução a Scrum Board no Trello;
- Introdução ao Scrum Board no Jira;
- Dicas de Scrum Board;



## **SLOT 1 SCRUM I: PAPÉIS**

- Características do time Scrum;
- Papéis e responsabilidades Product Owner, Development Team e Scrum Master;



## **SLOT 2 CRISP-DM**

- Business Understanding;
- Data Understanding.



## **SLOT 2 BIBLIOTECAS COM PYTHON**

- Numpy e Pandas;
- Matplotlib e Seaborn.



## **SLOT 2 AMBIENTES**

- Instalação de pacotes;
- Introdução ao Git;
- Introdução ao Anaconda;
- Introdução ao Jupyter;



## **SLOT 2 CARREIRA II: O MERCADO NA ÁREA DE CIÊNCIA DE DADOS**

- Carreira em Ciência de Dados;
- Outputs e requisitos comuns para cientistas de dados;
- Perguntas comuns em processos seletivos;
- Mapeamento de Lacun;
- Organização de Tempo;



## **SLOT 2 SCRUM II: EVENTOS**

- Processos e artefatos;
- Cone da incerteza;
- Sprint, Sprint Planning, Daily Scrum, Sprint Review e Sprint Retrospective;
- Produtividade;
- Roleplay: Daily Meeting;
- O que é DoD e DoR?



## **SLOT 2** **SCRUM II:** **ARTEFATOS**

- Product view;
- Product backlog
- Sprint backlog
- Incremento.



## **SLOT 3** **ESTATÍSTICA** **DESCRITIVA**

- Frequências e Medidas;
- Probabilidades e Amostragem;
- Correlação.



## **SLOT 3** **LINGUAGEM** **DE BANCO DE** **DADOS**

- Linguagem SQL;



## **SLOT 3** **METABASE:** **VISUALIZAÇÃO DE** **DADOS COM SQL**

- Dashboards com Metabase.



## **SLOT 3** **SCRUM III:** **ÉPICOS** **E STORIES**

- Épico, user, stories e task;
- Grooming/Priorização;
- Exemplo de decomposição;
- Montando Story Map do zero;
- Story para backlog priorizado - Jira;



## **SLOT 3**

# **SCRUM III: GRÁFICOS E REPORTS**

- Técnicas de estimativa;
- Burndown chart;
- Velocity chart;
- Rodando Scrum do Zero;



## **SLOT 4**

# **DATA STORY- TELLING PARA VISUALIZAR DADOS**

- Storytelling com dados;
- Gráficos.



## **SLOT 4**

# **MONGO DB**

- Bancos NoSQL (MongoDB).



## **SLOT 4**

# **ESTATÍSTICA INFERENCIAL**

- Estimativa;
- Testes de hipóteses;
- População e amostra;
- Probabilidade e Distribuições.



## **SLOT 5**

# **CRISP-DM**

- Data preparation.



## **SLOT 5**

# **INTRODUÇÃO À MACHINE LEARNING**

- Algoritmos Supervisionados;
- Algoritmos Não Supervisionados.



## **SLOT 5 DATA CLEANING**

- Estruturação dos Dados;
- Tratamento de Nulos;
- Tratamento de Outliers;
- Tratamento de Datas.



## **SLOT 5 CARREIRA III: FRAMEWORKS PARA DESENVOLVER SOFT SKILLS**

- Frameworks para Desenvolver Soft Skills.



## **SLOT 6 DATA WRANGLING**

- Operações Matemáticas;
- Operações Categóricas;
- Operações de Texto;
- Pipeline.



## **SLOT 6 ESTATÍSTICA AVANÇADA**

- Correlação e Regressão;
- Séries temporais.



## **SLOT 7 CRISP-DM**

- Modelling.



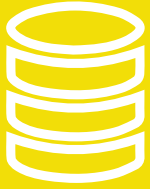
## **SLOT 7 REGRESSÃO: MODELOS SUPERVISIONADOS DE MACHINE LEARNING**

- Regressão Linear Simples/Múltipla;
- Métricas de avaliação.



## **SLOT 8 CLASSIFICAÇÃO: MODELOS SUPERVISIONADOS DE MACHINE LEARNING**

- Regressão Logística;
- K-Nearest Neighbours;
- Decision Trees;
- Support Vector Machines.

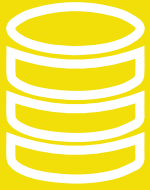


## **SLOT 9**

# **CLUSTERING:**

### **MODELOS NÃO-SUPERVISIONADOS DE MACHINE LEARNING**

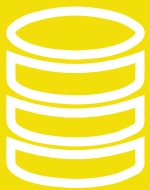
- K-means;
- DBSCAN;
- Mean-shift;
- Content-based filtering;
- Collaborative filtering;



## **SLOT 9**

# **APRESENTAÇÃO DE PROJETOS I**

- Identidade visual;
- Diagramação;
- Design x Persona;
- Fluxo de raciocínio;
- Vendendo seu peixe.

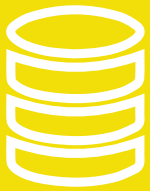


## **SLOT 10**

# **RECOMENDAÇÃO:**

### **MODELOS NÃO-SUPERVISIONADOS DE MACHINE LEARNING**

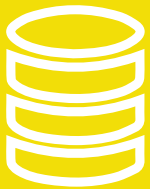
- Filtros Colaborativos,
- Matriz de Relacionamento
- Similaridade Produto/ cliente;
- Redes Neurais para Encoding.



## **SLOT 10**

# **CRISP-DM**

- Evaluation.



## **SLOT 11**

# **CRISP-DM**

- Deployment.



## **SLOT 11**

# **INTRODUÇÃO À BIG DATA**

- Apache Spark;
- Hadoop.

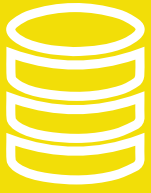


## **SLOT 12**

# **IMBALANCED DATA SET**

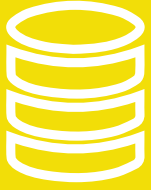
- Oversample e undersample;
- SMOTE;
- ADASYN;
- Near Miss Undersampling.





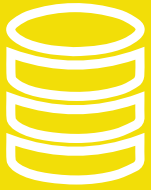
## **SLOT 12** **IMPLANTAÇÃO** **DE MODELOS**

- Deploy como API;
- Aplicação Web.



## **SLOT 12** **MODELOS** **DISPONÍVEIS** **EM CLOUD**

- Conceito em nuvem e open source;
- Análise de dados;
- Orquestradores;
- ML flow.



## **SLOT 12** **CARREIRA IV:** **PREPARAÇÃO PARA** **PROCESSOS SELETIVOS**

- Os processos seletivos e funil;
- Onde achar vagas e como se aplicar;
- Pitch;
- Dinâmica e cases;
- Frameworks de resolução de case;
- Entrevista;
- Negociar job offer;
- Próximos passos de carreira.



## **SLOT 13** **APRESENTAÇÃO** **DE PROJETOS II**

- Narrativa e Storytelling;
- DataViz em apresentações;
- Técnicas de apresentação;
- Postura e Oratória;
- One page report;
- Big numbers e fechamento;
- Montando do zero um ppt de impacto.



## **SLOT 13 REPLICAÇÃO**

- Ciclo de experimentação e automação;
- Noções de CI/CD.



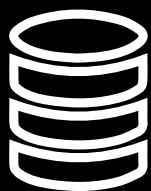
## **SLOT 13 SEGURANÇA E LGPD**

- Anonimização e Pseudoanonimização de Dados;
- Introdução a Criptografia;
- segurança em projetos de dados.

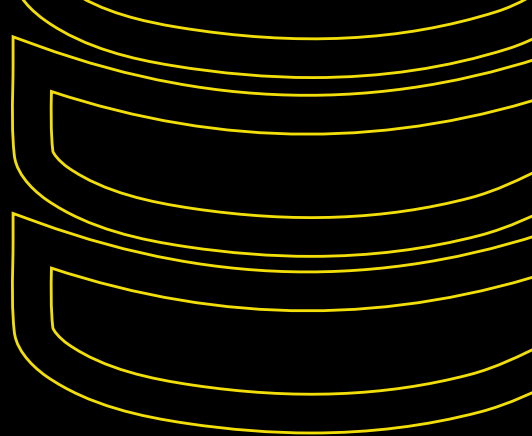


## **SLOT 13 PLATAFORMAS E DATA SCIENCE**

- Apresentando AWS e Sage Maker;
- Serviços para data engineering.



# 15 CERTIFICAÇÕES



**Python  
Fundamentals**



**SQL for  
Data Science**



**Statistics  
with Python**



**Machine Learning  
Cloud Services**



**Data Cleaning and  
Data Wrangling**



**Classification  
Algorithm in Machine  
Learning**



**Linear Regression in  
Machine Learning**



**Data Protection and  
Information Security**



**Clustering in  
Machine Learning**



**Machine Learning  
in Recommender  
Systems**



**CRISP-DM**



**Scrum  
Fundamentals**



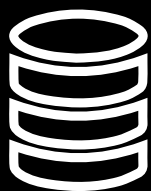
**Career Preparation**



**Presentation  
Techniques**



**Data Expert  
(certificado final)**



# PERGUNTAS FREQUENTES



## O que é abordado no treinamento?

O treinamento aborda programação em Python, do nível básico a funções avançadas ao longo do curso, modelos de machine learning, análises estatísticas com python, técnicas de balanceamento de dados, LGPD e aplicação do framework Crisp-DM no projeto prático.

## Quais são os Pré-requisitos para curso?

Esse curso não exige nenhum pré-requisito em específico.

## Qual a duração do curso?

O curso acontece durante cerca de 9 meses, com 14 encontros ao vivo e assignments online semanais.

## É necessário saber inglês para realizar o curso?

É desejável um inglês no mínimo intermediário, pois grande parte da nomenclatura do curso é dada em inglês. Logo ter conhecimento em inglês ajudará na absorção do aprendizado.

## O curso é indicado para profissionais de todas as áreas?

Este curso é recomendado para profissionais que nunca tiveram contato com a área de dados, mas queiram trabalhar com análises e criação de modelos preditivos. Para profissionais do mercado que desejam atuar com análises estatísticas e implementação de algoritmos

# DATA EXPERT

DATA SCIENCE & MACHINE LEARNING

2022  
**EMENTA  
DO CURSO**

**ACOMPANHE NOSSAS  
REDES SOCIAIS**

CLIQUE NOS ÍCONES PARA ACESSAR

