

dnc
apresenta




Data Science & Machine Learning



Sponsor

Daniel Omar Soria

Conecte-se no 



Consultor de Inteligencia Artificial @IBM

Founder @TryOn



dnc

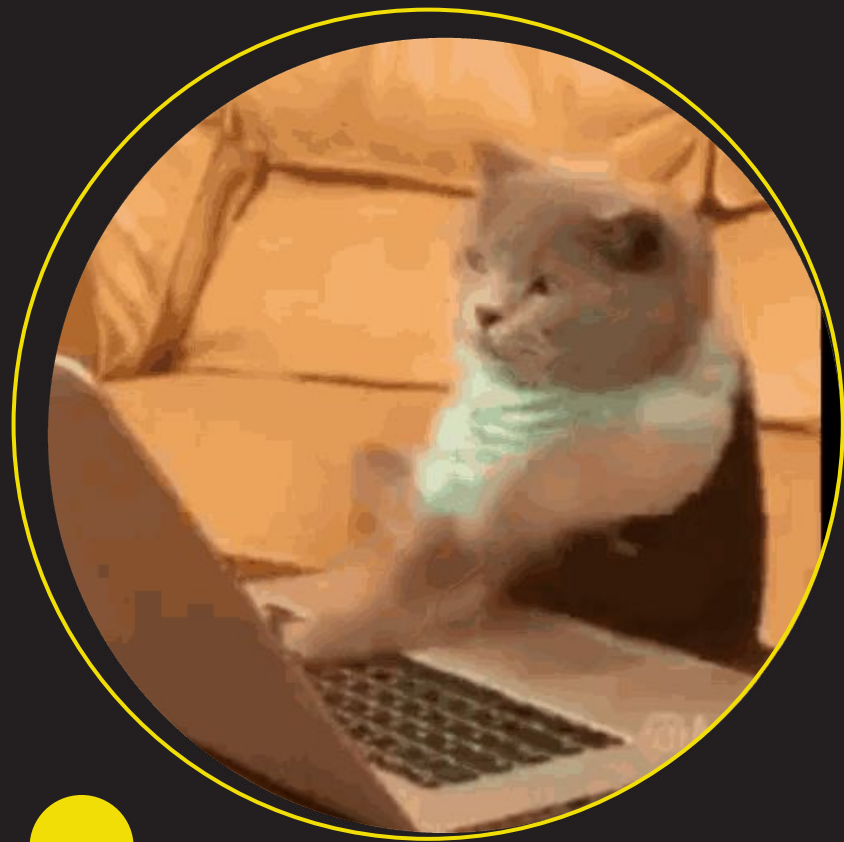
ANTES DE COMEÇARMOS...

IMPORTANTE!

É extremamente importante o
preenchimento do

Formulário de Confidencialidade!

Preencham **agora!**



O DEX

COMEÇOU!!

Bora compartilhar!! Poste um stories e marque:

@dnc.group

@danielomar.soria

@lucas.rana

dnc
apresenta



Data Science & Machine Learning



AGENDA, DAY #1

09h00 – 12h10 // 13h10 – 18h30

Manhã

- Alinhamento de Expectativas
- Sobre o Curso
- Dinâmica Integradora
- Framework **CRISP-DM**
- Boas Práticas de Consultoria
- Roda de Conversa, com @**Carolina Bez** e @**Isabelle Idalgo**

Tarde

- Introdução a **Algoritmo**
- **Dinâmica 1** - Algoritmo
- Introdução a **Diagnóstico**
- **Dinâmica 2** - Diagnóstico
- Gamificação: **Kahoot**
- **Próximos Passos** do Curso



Pilares #Hardwork

- Objetivo Claro
- Colaboração
- Dedicação
- Preparação
- Sem vitimismo
- Overdelivery
- Sorriso no rosto
- Curtindo o processo

Focus in what matters. Show me the money.

Siga @dncgroupbr e @dnc.alunos

Como lido com o curso

- Você **não vai** ter tudo na sua mão!
- Saiba **ouvir** feedback
- Feedback não tem réplica
- O curso é o começo □ Importante é o pós curso



Alinhamento de expectativas



Carga horária **longa!**



Days = **prática** dos conhecimentos vistos nas aulas online



Vocês terão acesso aos slides no **final da aula**

Quem irá te ajudar?



Consultor(a)



Co-host



Staff



Facilitador(a)



Mentor(a)

Além do suporte da equipe dnc.group, claro!!!



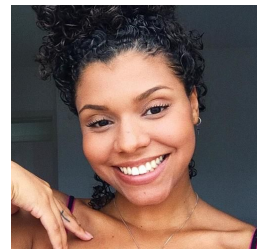
Glaucia Lopes



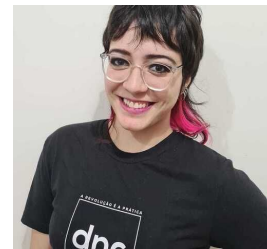
Natália Mota



Caroline Fuji



Juliana Santos



Ana Júlia Souza

Quem irá te ajudar?



Consultor

A pessoa que estará facilitando o Day. É aquele que vai estar apresentando os dias EaD, liderando o Kahoot, apresentando as dinâmicas, etc



Staff / Co-Host

A pessoa que estará atenta aos chats e que estará por trás de toda a mágica do Day.



Mentor

A pessoa que estará cuidando do grupo. Essa pessoa irá auxiliar vocês durante todo o curso a entregar um projeto de qualidade e será responsável pela correção dos deliverables (vamos falar mais pra frente).



Facilitador

Temos dois tipos de facilitador: O que estará atenta ao slack para responder às dúvidas técnicas de vocês uma vez ao dia apenas. E os que auxiliam o grupo nas dinâmicas dos Days.



Ferramentas

- **ZOOM** → Encontros (Days)
- **SLACK** → Comunicação: avisos e dúvidas
- **DNC.ALUNOS** → Plataforma Oficial dos Alunos
 - Assistir as aulas (*dnc.cursos*)
 - Avaliação de Soft Skills do grupo (*dnc.softcheck*)
 - Notas, Avaliação e Ranking (*dnc.score*)
 - Central de Ajuda / Dúvidas Frequentes (*dnc.help*)
 - Portal de Contratação / Headhunting (*dnc.hiring*)



**TURMA
DEX01**

SOBRE O CURSO

Prévia

**Day
1**

HOJE

Day 2

Day 3

Day 4

Slot 0

Python 0

Slot 1

Onboarding;
Introdução à Projetos de
Ciência de Dados ;
Estatística Descritiva;
Ambientes;
Bibliotecas (Python 1).

Slot 2

Linguagem de Banco de
Dados;
Data Storytelling;
Metabase;
CRISP-DM: Entendendo o
Problema de Negócio e Dados.

Slot 3

Estatística Inferencial;
Estatística Avançada;
Introdução à Machine
Learning.

Slot 4

Data Cleaning;
CRISP-DM: Preparando os
Dados e Entendendo o
Modelo.

Day 5

Day 6

Day 7

Day 8

Day 9

Slot 5

Regressão: Modelos
Supervisionados;
Classificação: Modelos
Supervisionados.

Slot 6

Clustering: Modelos não
Supervisionados;
Recomendação com
Machine Learning;
CRISP-DM: Avaliando seu
Modelo.

Slot 7

Consumindo Data Lakes;
Introdução a Big Data;
Imbalanced Data Set;
CRISP-DM: Produzindo o
seu Modelo;
Implantação de Modelos.
Modelos Disponíveis em
Cloud.

Slot 8

Replicação (Plus);
Segurança, LGPD E Bias
Carreira e Planos de Estudo
Entrevistas
Cases da Vida Real
Plataformas de Data
Science

Banca



Cronograma

Total: ~370h



240h

ONLINE



60h

PROJETO



72h

AO VIVO



Dia 9 | fechamento

**Banca dos
projetos
(empresa e
gestores)**

12 Certificados

**+ 1 Experiência
no CV**

PREMIAÇÃO

TOP 1
Experiência Internacional





Certificados

ANTES

MARTA LINS
 +XX 99999-9999 marta.lins@email.com.br

EXPERIÊNCIAS

Cargo 1
2018-2019

EMPRESA XYZ
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi

FORMAÇÃO

Curso 1
2016-2020

UNIVERSIDADE XYZ
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi

DEPOIS

MARTA LINS
 +XX 99999-9999 marta.lins@email.com.br

EXPERIÊNCIAS

Cargo 1
2018-2019

EMPRESA XYZ
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi

Cargo 2
2020-2021

DNC.GROUP
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi

FORMAÇÃO

Universidade X
2016-2020

CURSO XYZ
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi

DNC.GROUP
2020-2021

DATA EXPERT
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi

CERTIFICADOS

Python Zero

SQL para Data Science

Estatística com Python

Regressão Linear com Machine Learning

Classificação com Machine Learning

Data Science com CRISP-DM

Clusterização da Machine Learning

Data Science com Sage Maker e Watson Studio

Segurança de Dados e LGPD

Data Expert

Consultores



Samuel Bozzi

Whirlpool



Isabelle Idalgo



Daniel Omar



Peterson Almeida

AmBev



Carolina Bez



Amanda Ferraboli



Pedro Pereira





Sistema de avaliação

Item	Nota Máxima (UN)	Quantidade	Total
Assignment (cases)	90	32	2.880
Deliverable (entregas de projeto)	350	7	2.450
Certificados Intermediários	150	11	1.650
Certificado Final	520	1	520
Banca	1250	1	1.250
Hackathon	250	3	750
Super Challenge Powered By	500	1	500
TOTAL			10.000



Regra de desligamento

Para não ocorrer o desligamento do curso, será necessário minimamente:

Entregar **2 Assignments** por slot.

75% de presença nos encontros.

Entregar os **Deliverables**

SLOT 1

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 2

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 3

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 4

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 5

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 6

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 7

2 assignments
Total: 4 assignments
Gabarito: Day 2

SLOT 8

2 assignments
Total: 4 assignments
Gabarito: Day 2

As verificações serão feitas em toda sexta-feira antes do encontro ao vivo (Days).

A data limite de submissão dos assignments e os deliverables até sexta-feira 12h (meio dia) antes do próximo encontro ao vivo (Days).



LinkedIn - Certificados

Licenças e Certificados

Certificado



Nome *

Data Expert (370h)

Organização emissora *

dnc Dinâmica Group

☒ Esta credencial não expira

Data de emissão

Maio

Sem data de expiração

2020

Código da credencial

x3x3x3x3

URL da credencial

<https://mycourse.app/x3x3x3x3>



credencial



LinkedIn - Experiência

Experiência



Português (perfil principal)

Inglês 2

Cargo *

Data Scientist (Student)

Tipo de trabalho

Temporário

Empresa *

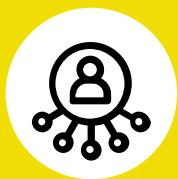


Dinâmica Group



**Não coloquem
a empresa
do projeto.**

INTEGRAÇÃO



Apresentação Pessoal

Nome / Faculdade / Curso / Empresa / Objetivo



<https://linktr.ee/dex01dnc>





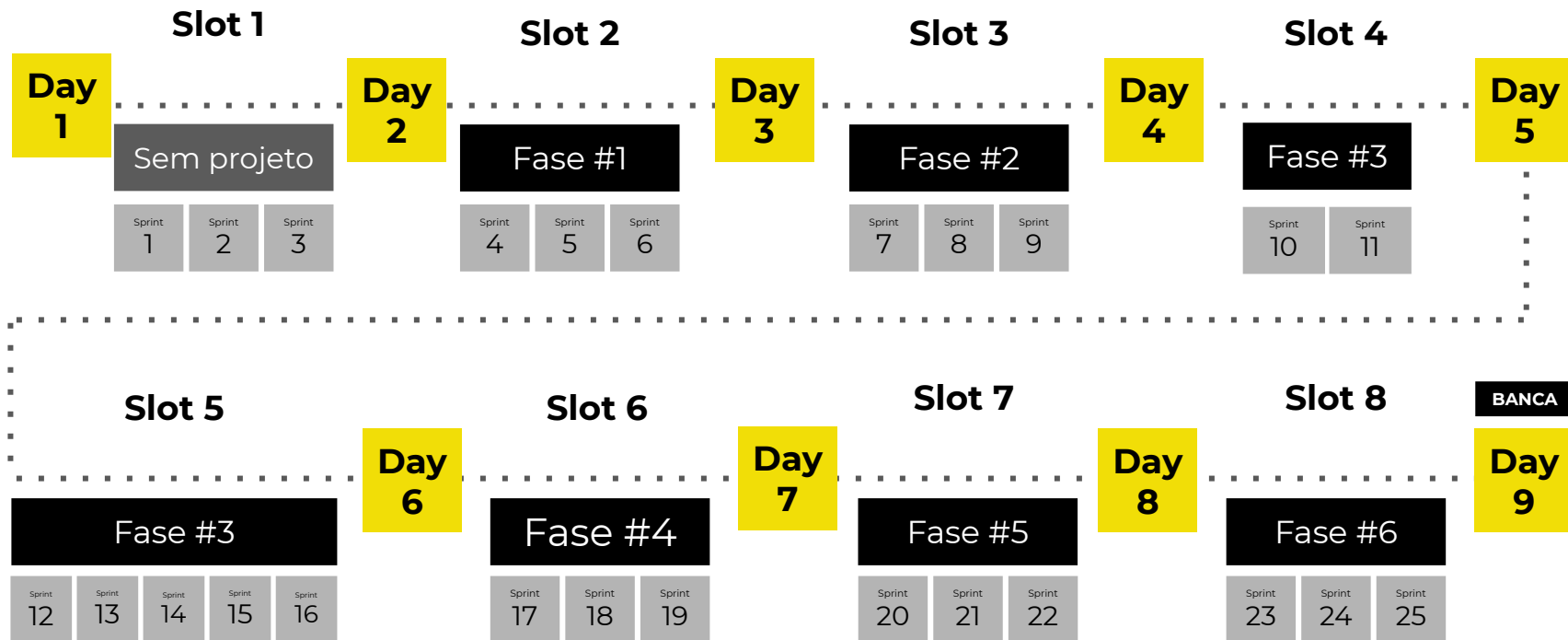
HORA DE DAR UM

BREAK

[15 minutos]

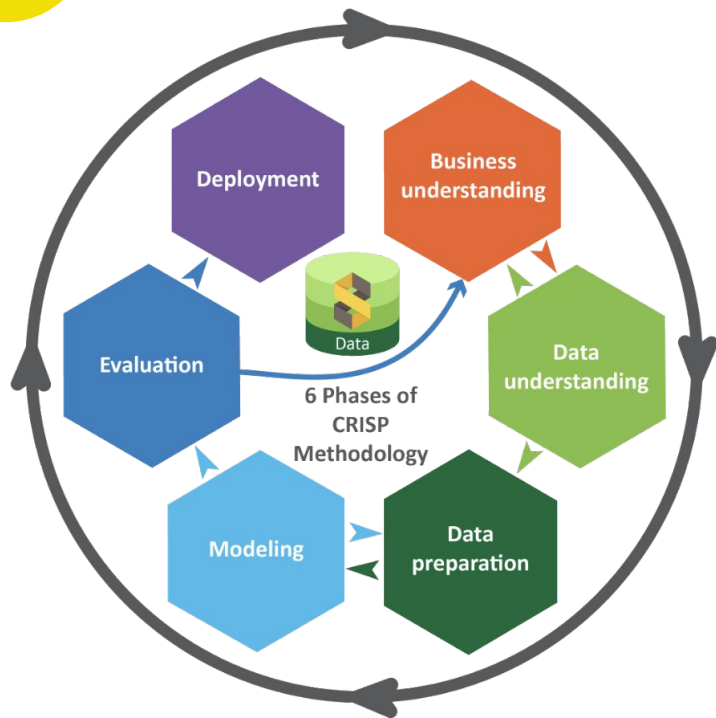
FRAMEWORK CRISP-DM

CRONOGRAMA DO PROJETO





Framework CRISP-DM



#1

Business Understanding

Compreender os objetivos e requisitos do projeto.

#2

Data Understanding

Identificar, coletar e analisar os conjuntos de dados que podem ajudar a cumprir os objetivos do projeto.

#3

Data Preparation

Preparação do conjunto de dados final para a modelagem.

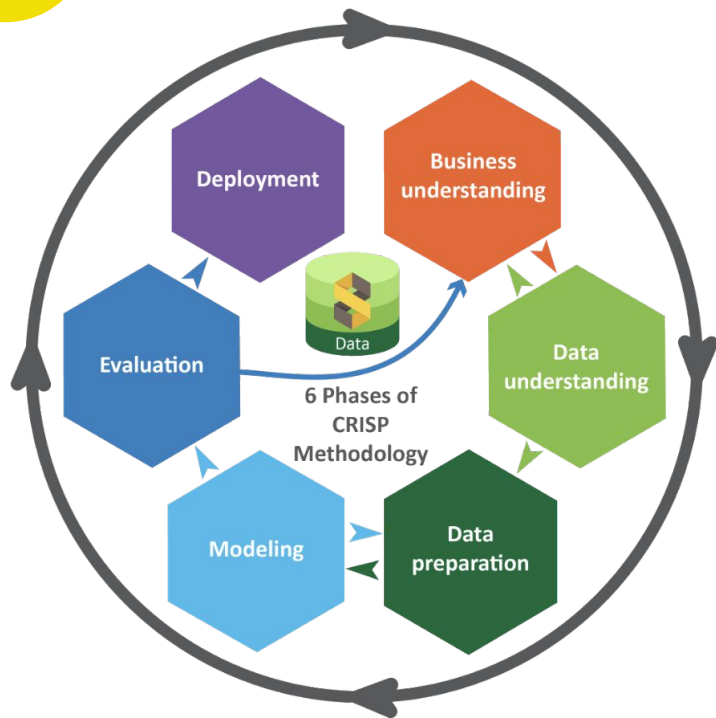
#4

Modeling

Construir e avaliar vários algoritmos com base em diferentes técnicas de modelagem.



Framework CRISP-DM



#5

Evaluation

Analisar mais amplamente qual modelo atende melhor ao negócio e o que fazer com os resultados

#6

Deployment

Garantir a implementação do modelo desenvolvido e o fácil acesso do cliente.



Framework CRISP-DM



**In data we
trust!**

#1

Business Understanding

Compreender os objetivos e requisitos do projeto.

#2

Data Understanding

Identificar, coletar e analisar os conjuntos de dados que podem ajudar a cumprir os objetivos do projeto.

#3

Data Preparation

Preparação do conjunto de dados final para a modelagem.

#4

Modeling

Construir e avaliar vários algoritmos com base em diferentes técnicas de modelagem.

#5

Evaluation

Analisar mais amplamente qual modelo atende melhor ao negócio e o que fazer com os resultados

#6

Deployment

Garantir a implementação do modelo desenvolvido e o fácil acesso do cliente.



Alinhamento de Expectativa de Projeto

- **Não** temos o controle do parceiro!
- A duração, área e complexidade do projeto **variam** de acordo com cada grupo/empresa.
- Independente do tipo de empresa há **sempre** oportunidade de aprendizado.
- 5 à 8% dos projetos são **cancelados** ao longo do treinamento e tá tudo bem!
- Os sponsors podem ter **perfis diferentes**.
- Foco sempre em atingir o **objetivo final** do projeto e nos resultados!
- Vocês receberão auxílio durante todo o projeto de um **mentor**, mas lembre-se, ele não será o líder do projeto.

O projeto é uma PARTE da jornada de aprendizagem



Grupo x sponsor

- **A parte mais difícil do projeto**
- **Engajamento**
 - Confiança
 - Simpatia
 - Agradável
 - Autoridade
 - Seriedade
 - Maturidade
- **Como PEDIR? Convencer...**

Entender

Pedir

Divertir

Conhecer

Compartilhar

Ensinar

Agradar

Agradecer

BOAS PRÁTICAS DE CONSULTORIA



Boas práticas de consultoria

- Postura
- A parte mais difícil do projeto é o contato com o Sponsor
- Cadência de Contato (dnc x sponsor x grupo)
- Abordagens com Sponsor
- Preocupação com resultado e não entrega de deliverable
- Impacto na empresa
- Engajamento
- Papel de cliente e não professor
- Vocabulário
- Apresentação de resultados (caminho)
- Comprometimento com o grupo
- Estudo não significa só assistir aulas
- Somente focamos esforços em cima do que acreditamos ser potencial.

RODA DE CONVERSA



Consultora
Carolina Bez

"O futuro do
Profissional de
Dados"



Consultora
Isabelle Idalgo



VAMOS

ALMOÇAR?

ALGORITMO

ESTRUTURA DE RACIOCÍNIO

Como criar algoritmos e
rotinas antes de codar



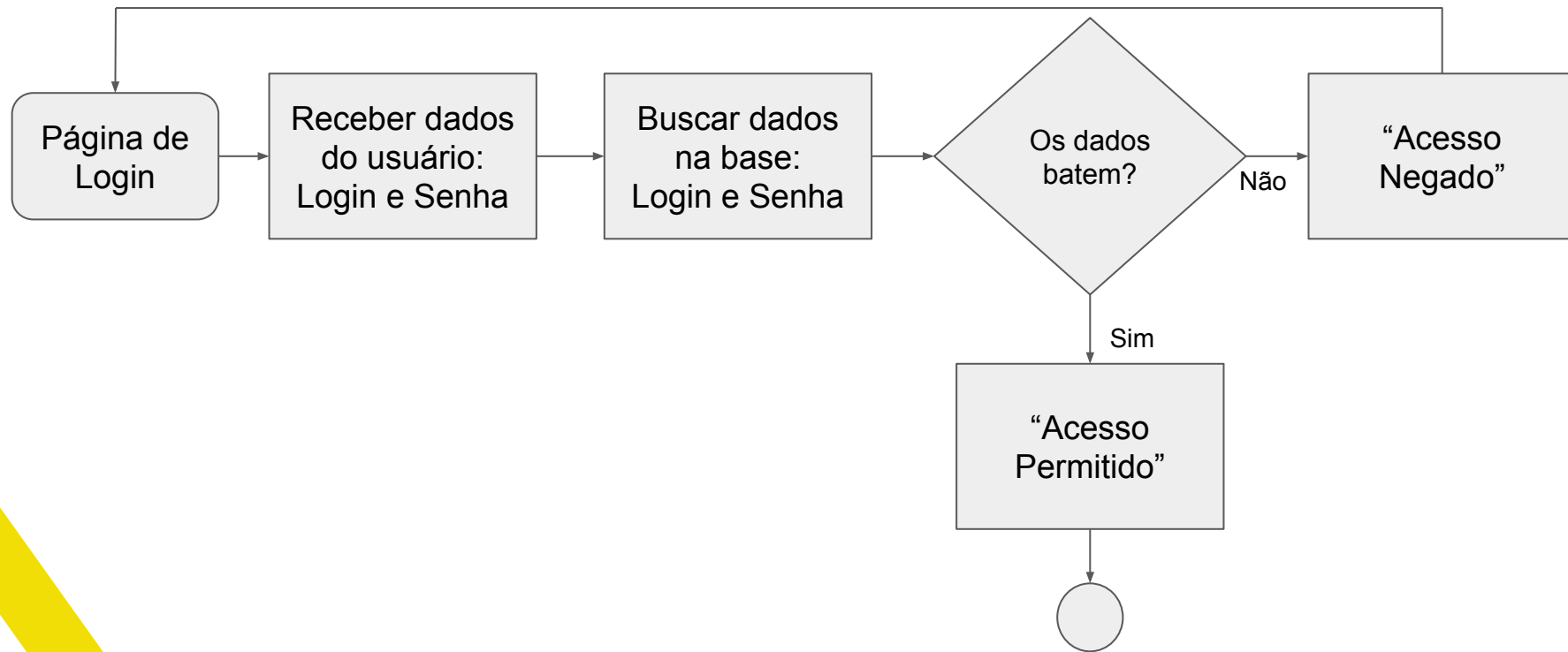
O que é um algoritmo?

**“Conjunto das regras e
procedimentos lógicos
perfeitamente definidos que
levam à solução de um problema
em um número finito de etapas.”**

Oxford Languages

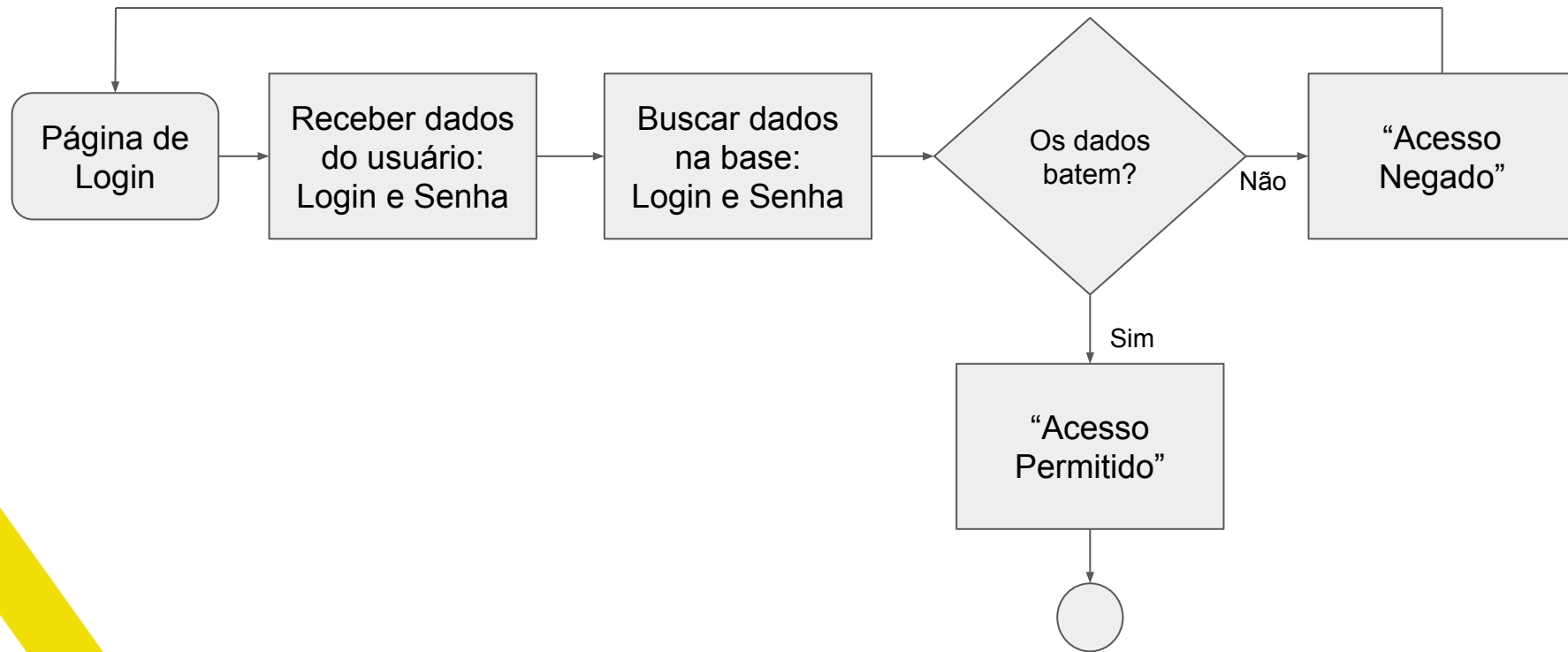


O que é um algoritmo?





O que é um algoritmo?

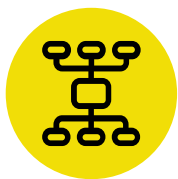




DINÂMICA #1

ALGORITMO

**[Lógica de programação +
Fluxograma de processo]**



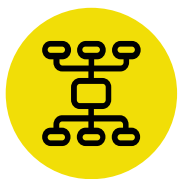
Dinâmica #1 - Algoritmo

As certificações intermediárias ao longo do DEx ocorrem mediante a uma prova. Para realiza-las, o aluno deve logar com seus dados de CPF e código da turma na respectiva prova. Se ele acertar 70% ou mais da avaliação, recebe a certificação e pontua no ranking. Caso não obtenha o resultado mínimo, só poderá realizar uma segunda tentativa em 7 dias, além de não receber pontos no ranking.

Desenvolva um algoritmo para **registrar os dados** relevantes no **processo** de certificação.



Tempo:
1h10 para
desenvolvimento do
fluxograma



Dinâmica #1 - Instruções

Ao término do processo é importante que seja possível retirar da base de dados:

- Pontuação feita a cada tentativa;
- Tempo de realização da prova;
- Alunos certificados;
- Pontuação no ranking;
- Qual a turma do aluno;



Dinâmica #1 - Observações

Observações:

- Todas as respostas da prova devem ter alguma alternativa selecionada.
- O certificado é exibido ao final da prova e apresenta o nome do aluno e a data da certificação.
- O Ranking só registra a primeira tentativa.



Dinâmica #1 - Entrega

Vocês precisam entregar um fluxograma do processo, para isso podem utilizar o software que desejarem:

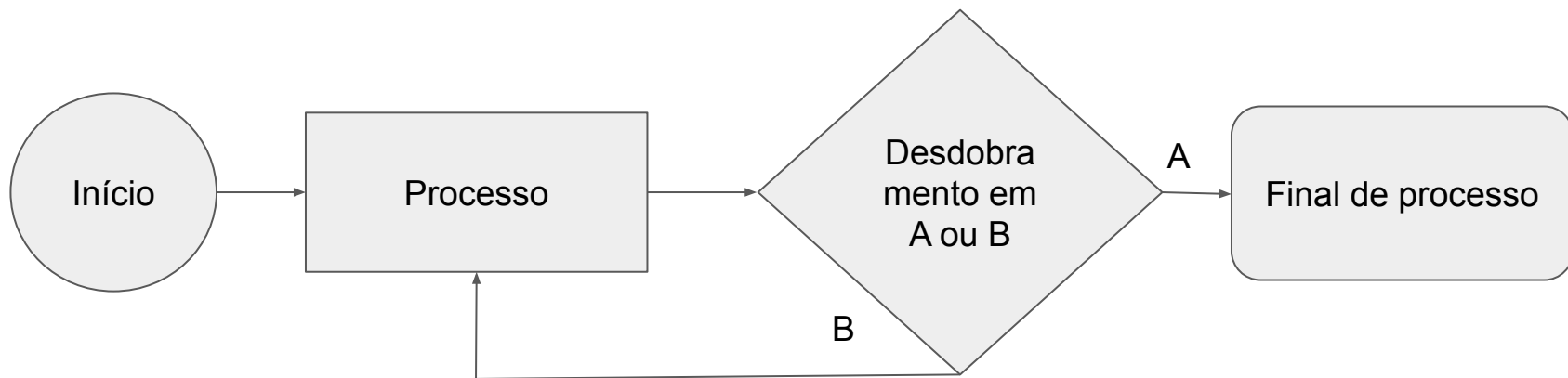
- Miro;
- Google Slides;
- Google Sheets;
- Paint;

Para entender o sistema, vocês terão acesso à um vídeo de prova:

<https://www.loom.com/share/920de636d73f4e86bac65e6231948758>



Dinâmica #1 - Elementos





RESOLUÇÃO

[DINÂMICA #1]

COMO FAZER UM DIAGNÓSTICO

[Aprendendo a fazer as perguntas certas]



Como fazer um diagnóstico

Qual o problema do negócio e como os dados podem resolvê-lo?



Benchmark



**Causa Raiz e
Objetivo**



Dados



Como fazer um diagnóstico



Benchmark

- Como são os números do mercado?
- Estamos com resultados condizentes?
- Quais são as soluções trazidas pelos concorrentes?
- Quais são as inovações na área?



Como fazer um diagnóstico



**Causa Raiz e
Objetivo**

- Qual o nosso problema?
- Por que esse problema acontece?
- Qual a causa geradora dele?
- Como podemos resolver isso?
- O que indica que esse problema foi resolvido?



Como fazer um diagnóstico



Dados

- Quais são os dados necessários para resolver esse problema?
- Possuímos esses dados ou será necessário registrá-los?
- Como podemos obtê-los?



HORA DE DAR UM

BREAK

[15 minutos]





DINÂMICA #2

DIAGNÓSTICO

**[Descobrindo o problema da
empresa!]**



Contexto

Utilizaremos o
dnc.softcheck

A Razzle Dazzle, um e-commerce de vestuário, está enfrentando problemas com o faturamento. A empresa possui uma taxa de conversão média de 2%, ou seja, a cada 100 clientes que visitam o site, apenas 2 realizam uma compra, além disso, cerca de 10% dos clientes realizam uma segunda compra.

Com o objetivo de contornar esta situação, a Razzle Dazzle contratou seus serviços de consultoria para realizar um diagnóstico da empresa e determinar qual o principal problema ela tem enfrentado.



Instruções

Vocês devem:

1. Definir o problema enfrentado pela empresa e a causa raiz dele.
2. Determinar o objetivo da consultoria - qual o nome do projeto que será realizado?
3. Definir cinco dados mais relevantes que ajudem a contornar o problema.
4. Estabelecer métrica de sucesso - como a empresa saberá que atingiu o objetivo da consultoria?



Timelog

1h30

30 min

15 min

10 min

15 min

10 min

05 min

05 min

Definição da causa
raiz

Dados necessários

Feedback

Definição do
problema

Objetivo
do projeto

Métrica de
Sucesso

Softcheck



RESOLUÇÃO

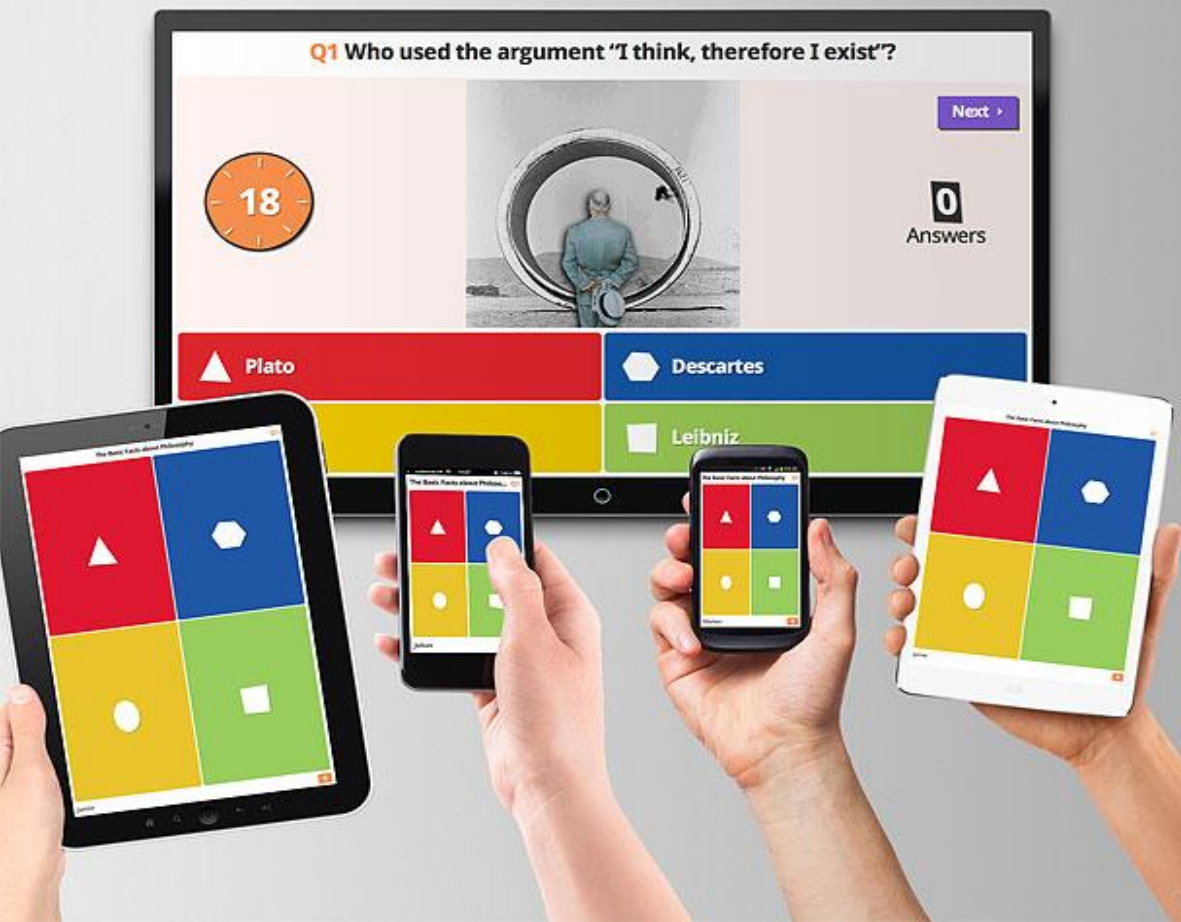
[DINÂMICA #2]



HORA DO

KAHOOT

[30 minutos]

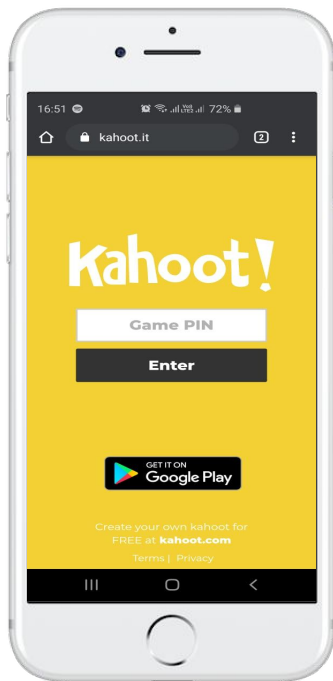


Kahoot

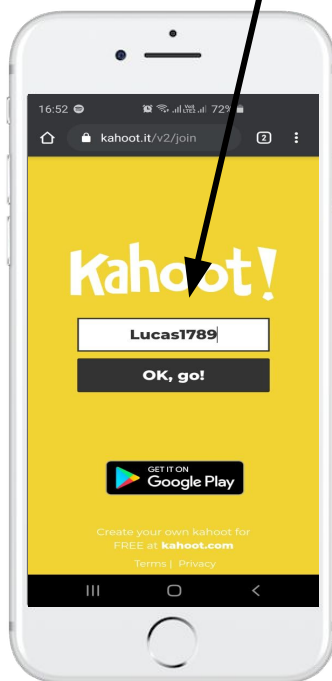
JOGO PARA
FIXAR CONCEITOS

<https://kahoot.it/>

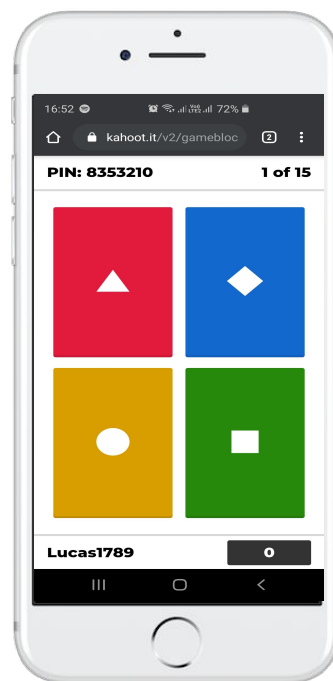
Como funciona?



Passo 1



Passo 2



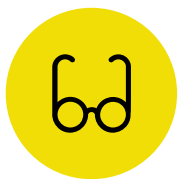
Passo 3

FECHAMENTO

PESQUISA DE SATISFAÇÃO

Para ter acesso aos slides do curso, preencha a pesquisa que se encontra no link abaixo:

<https://linktr.ee/dex01dnc>



Próximos passos

- Os grupos, temas dos projetos e mentores serão divulgados para dia 16/07 para o primeiro deliverable.
- O primeiro slot de aulas online estará disponível dia 28/06 e o fim desse primeiro slot é dia 16/07.
- **AVISO 1:** devido a pandemia os Kits podem atrasar um pouco para chegar. Estamos fazendo o máximo para entregar até o Day 02.
- **AVISO 2:** não esqueçam de preencher o formulário de compliance para que vocês possam ser alocados num projeto. Deadline: **hoje!!**



OBRIGADO!



©dnc group. Todos os direitos reservados.



Tirando dúvidas



PARTE 1

Perguntas no chat

PARTE 2

Levante a mão