

A aula interativa do Módulo 4 – Bootcamp Cientista de Dados em breve!

Importante:

- 1) Não se esqueça de acessar a aula com seu e-mail cadastrado no ambiente de aprendizagem do IGTI e seu nome completo. Entrou com os dados errados? Saia da sala e entre novamente com os dados corretos.
- 2) Para sua frequência ser computada, responda a enquete no ambiente de aprendizagem, no horário indicado pelo professor. A enquete ficará disponível por 10 minutos.
- 3) Utilize o Chat para interagir com os colegas durante a aula, a ferramenta Raise Hands para pedir a palavra, e, em caso de dúvidas sobre o conteúdo, utilize o Q&A (perguntas e respostas) para que o professor tutor possa respondê-las.



Informações importantes

- Acesse a aula com o seu nome completo e com o mesmo e-mail utilizado no cadastro do ambiente de aprendizagem do IGTi. Entrou com os dados errados? É só sair da sala e entrar com os dados corretos!
- As presenças das Aulas Interativas são computadas através de uma enquete, que será realizada no ambiente de aprendizagem do IGTi. Para sua frequência ser computada, quando solicitado pelo professor, você deverá ir até a seção “Enquete de presença da Aula Interativa”, localizada na Área Acadêmica da disciplina, e responder a enquete. Essa seção estará logo abaixo da que contiver o link para a Aula Interativa em questão. A enquete ficará no ar por 10 minutos e sua nota estará disponível ao término desse tempo.
- Utilize o chat para interagir com os colegas durante a aula interativa! Em caso de dúvidas sobre o conteúdo, é só postá-las no Q&A que o tutor irá respondê-las! Lembre-se que você ainda pode esclarecer as suas dúvidas nos fóruns disponibilizados no Ambiente de Aprendizagem.



Informações importantes

- Ah! E se você não conseguir assistir a aula interativa, não se preocupe! Sua gravação ficará disponível no Área Acadêmica, juntamente com os slides utilizados pelo professor, em até 24 horas úteis após o término da aula. Você também poderá realizar a atividade de reposição para recuperar os pontos de presença!
- Se você precisar solicitar prorrogação e/ou 2ª oportunidade para entrega de atividades, saiba que isso é realizado somente mediante a apresentação de atestado médico ou de óbito de parentes de 1º grau.
- Para melhor experiência nas aulas interativas, sugerimos que você baixe o aplicativo do Zoom no seu computador.

Técnicas para o Processamento do Big Data

Primeira Aula Interativa

Prof. Túlio Vieira



1ª Aula Interativa

PROF. Túlio Philipe Vieira

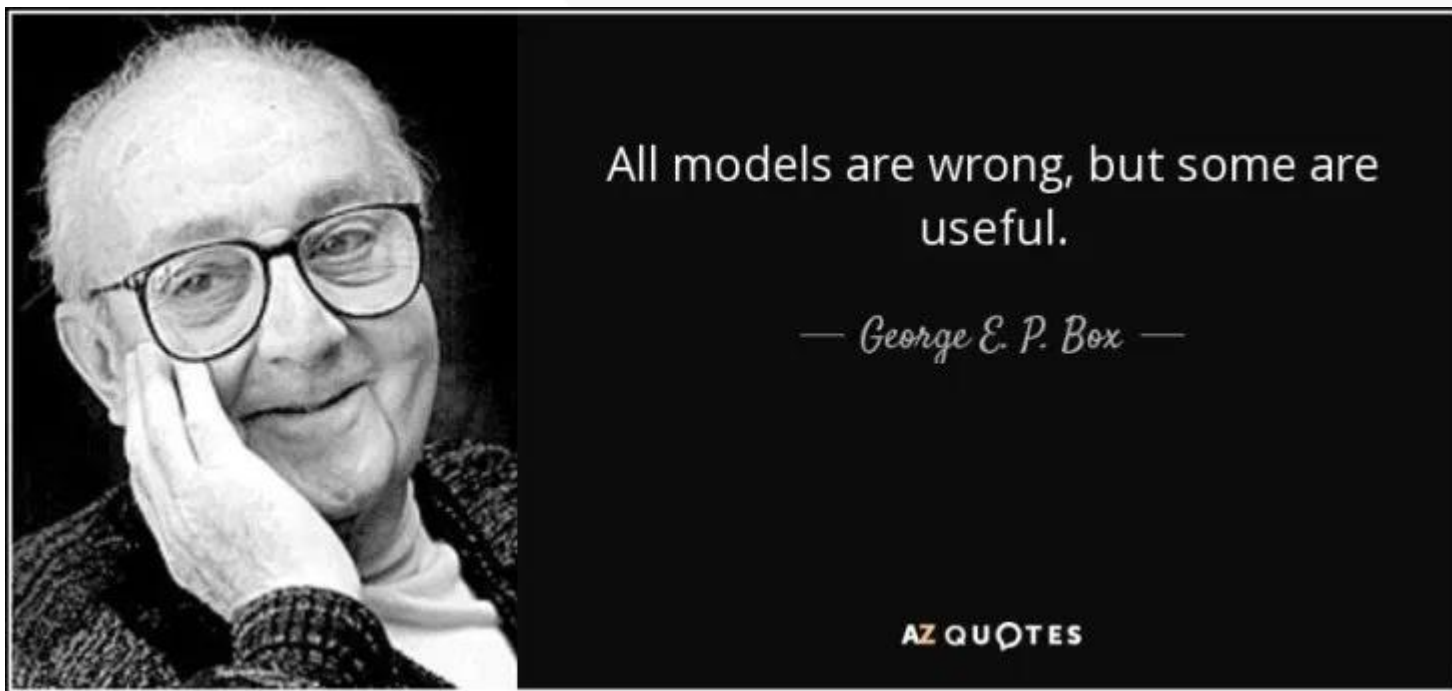


Nesta aula

1 Apresentação.

2 Correlação + PCA.

3 Prática.



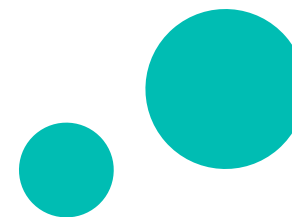


Apresentação

- Engenheiro de controle e automação pela UFMG, 2014;
- Especialista em engenharia da qualidade pela PUC-MG, 2016;
- Mestre em Modelagem Matemática Computacional pelo CEFET-MG, 2017;
- Doutorando em Modelagem Matemática Computacional pelo CEFET-MG, previsão 2022.



Questão de Aquecimento

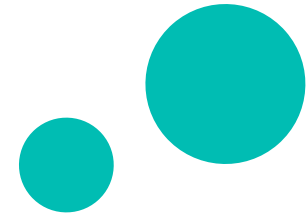


- Classificação Binária:
 - Identificar se o empregado vai ou não continuar no emprego.
 - Dataset: Human Resources Analytics.

	satisfaction_level	last_evaluation	number_project	average_monthly_hours	time_spend_company	Work_accident	promotion_last_5years	sales	salary	left
0	0.38	0.53	2	157	3	0	0	sales	low	1
1	0.80	0.86	5	262	6	0	0	sales	medium	1
2	0.11	0.88	7	272	4	0	0	sales	medium	1
3	0.72	0.87	5	223	5	0	0	sales	low	1
4	0.37	0.52	2	159	3	0	0	sales	low	1



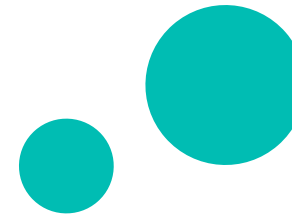
Questão de Aquecimento



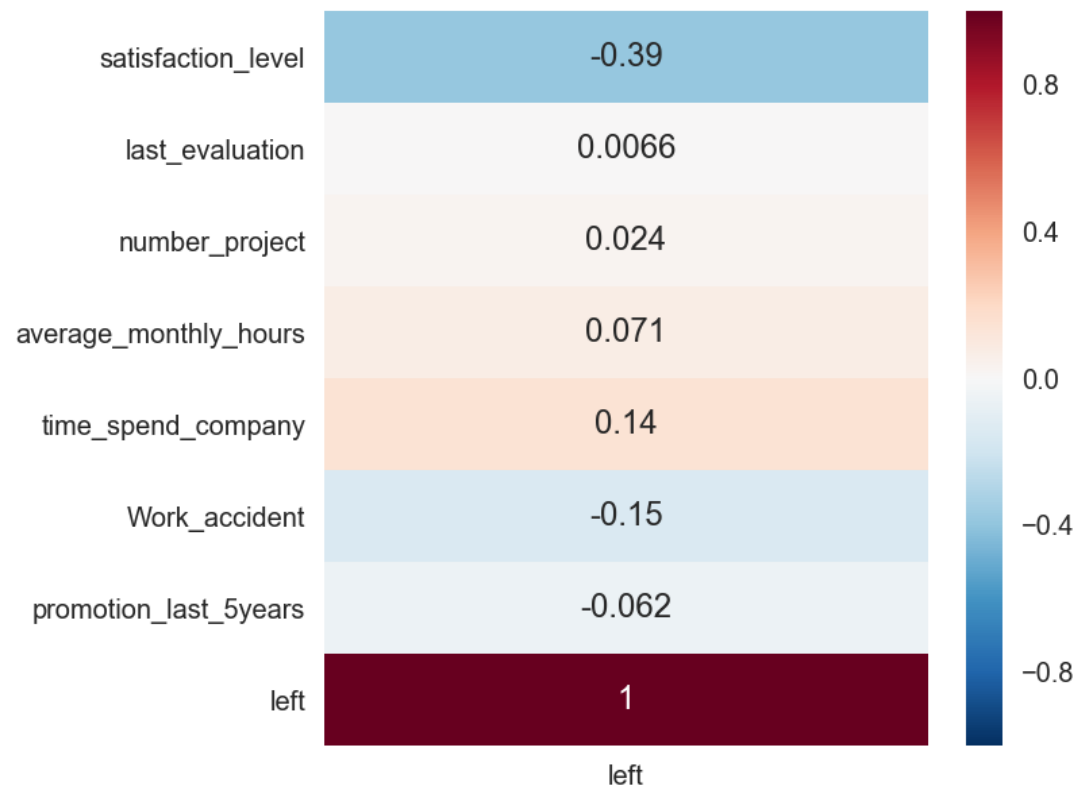
- Visualização e pré-processamento dos dados:
 - Qual tipo de dados estamos trabalhando: valores reais, categóricos ou ambos?
 - Existem correlações entre as variáveis?
 - Existem dados nulos?
 - Existem dados que precisam ser normalizados?
 - Para os “labels”, existem classes não balanceadas?

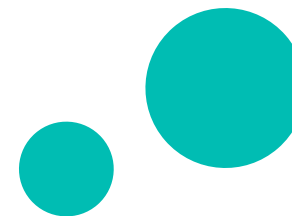


Questão de Aquecimento

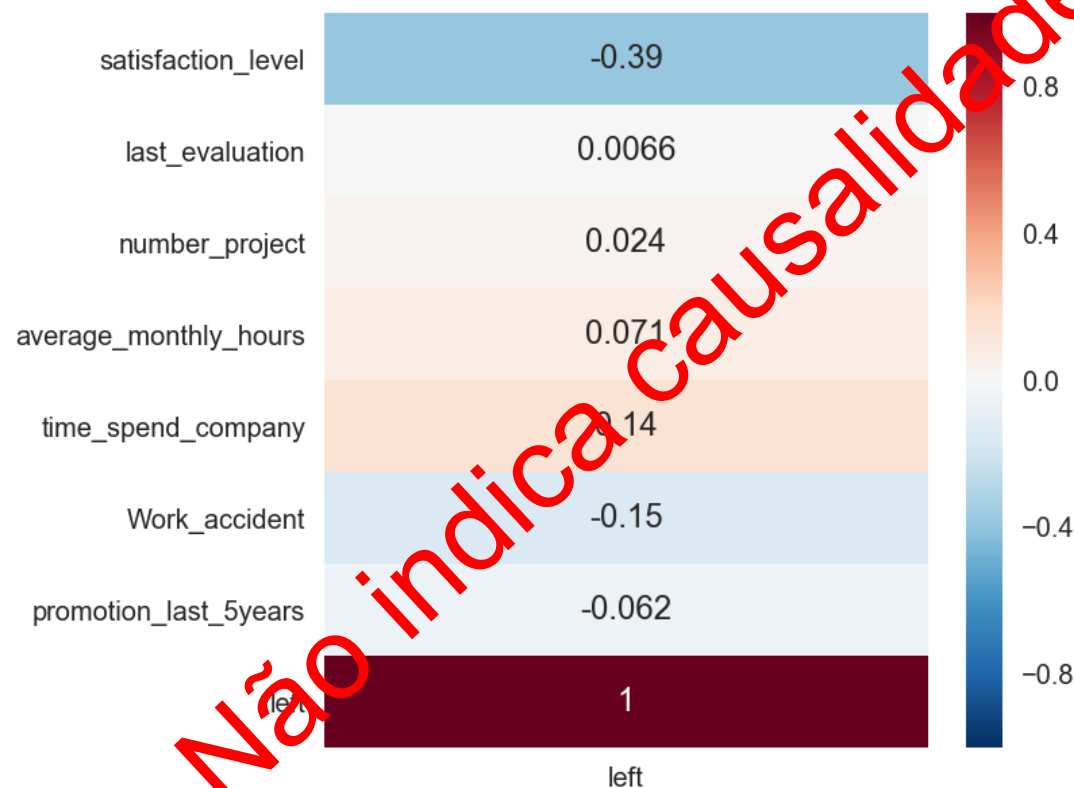


- Correlação:



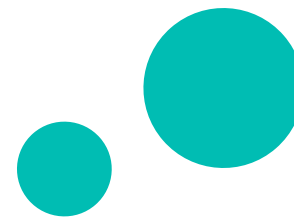


- Correlação:



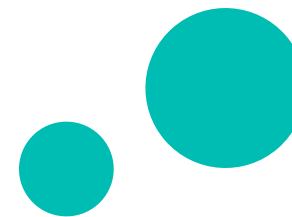


Questão de Aquecimento





Questão de Aquecimento

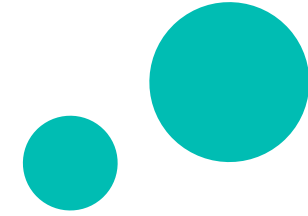


- Correlação:



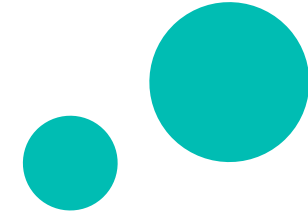


Principal Component Analysis (PCA)

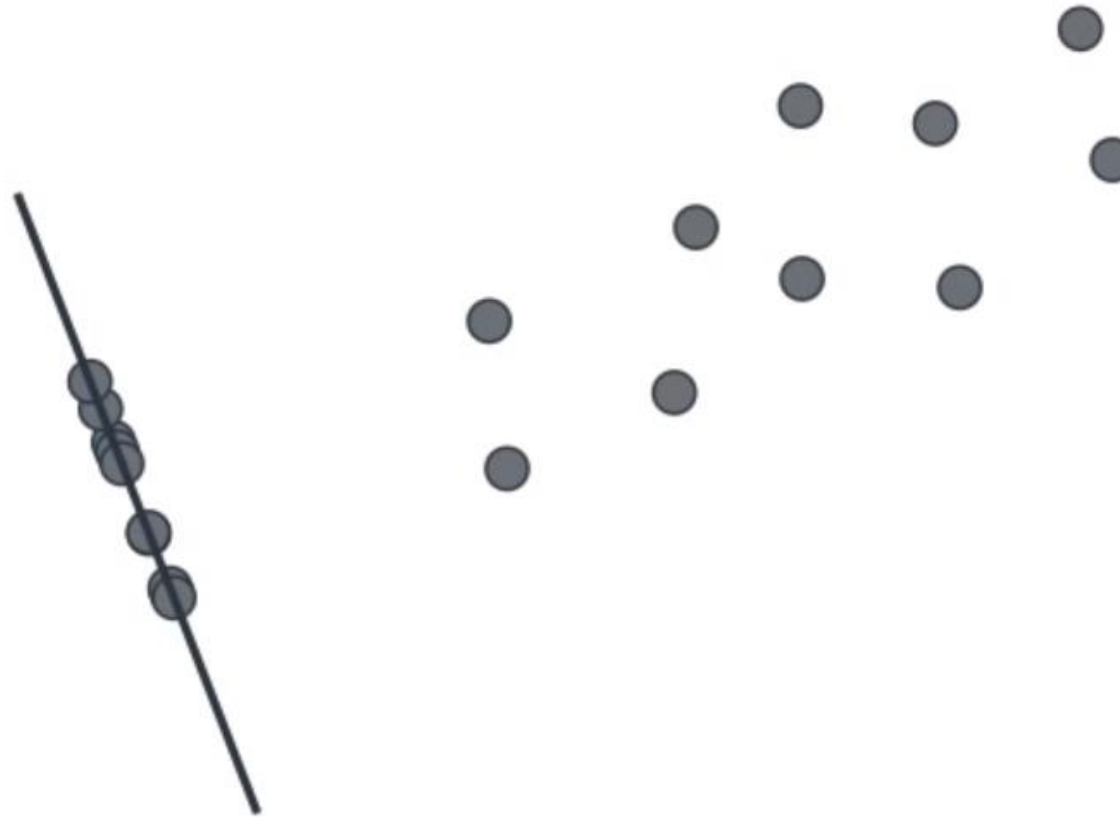
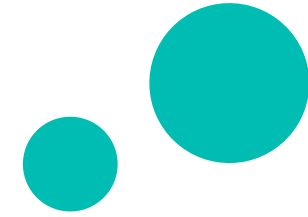


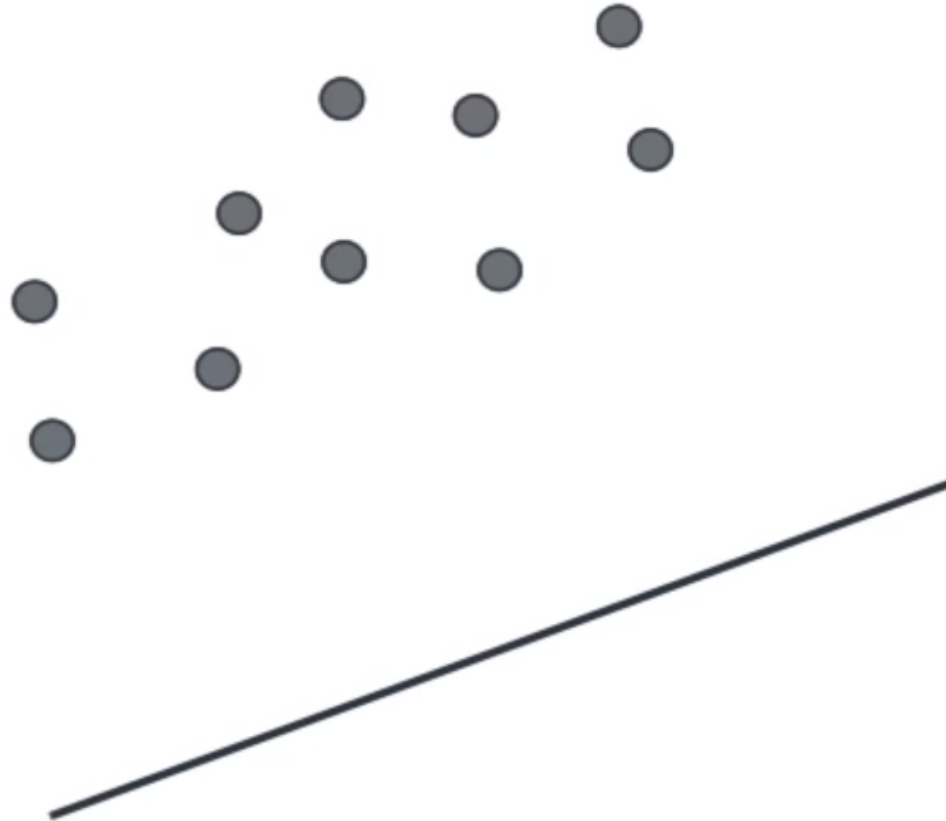
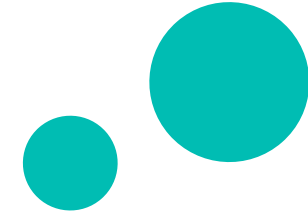


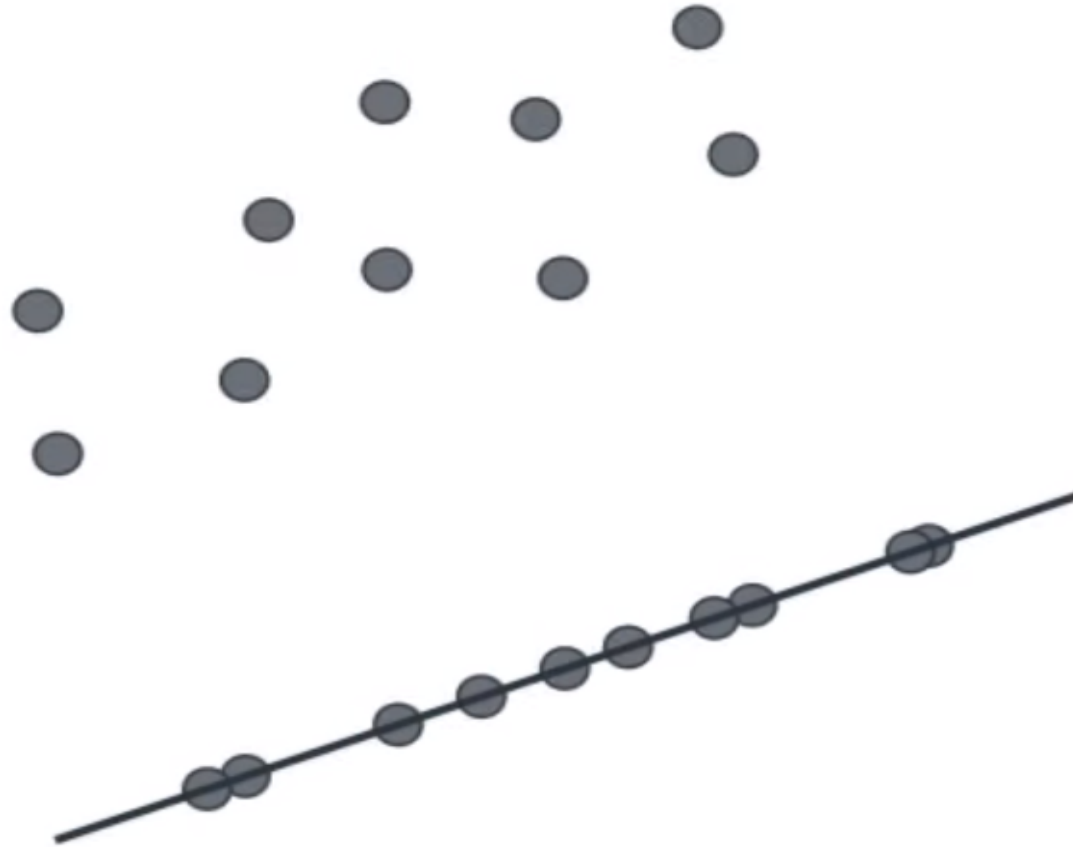
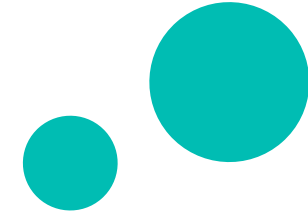
Principal Component Analysis (PCA)

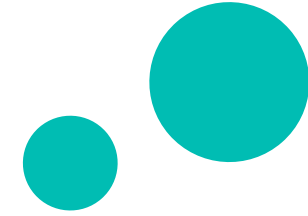


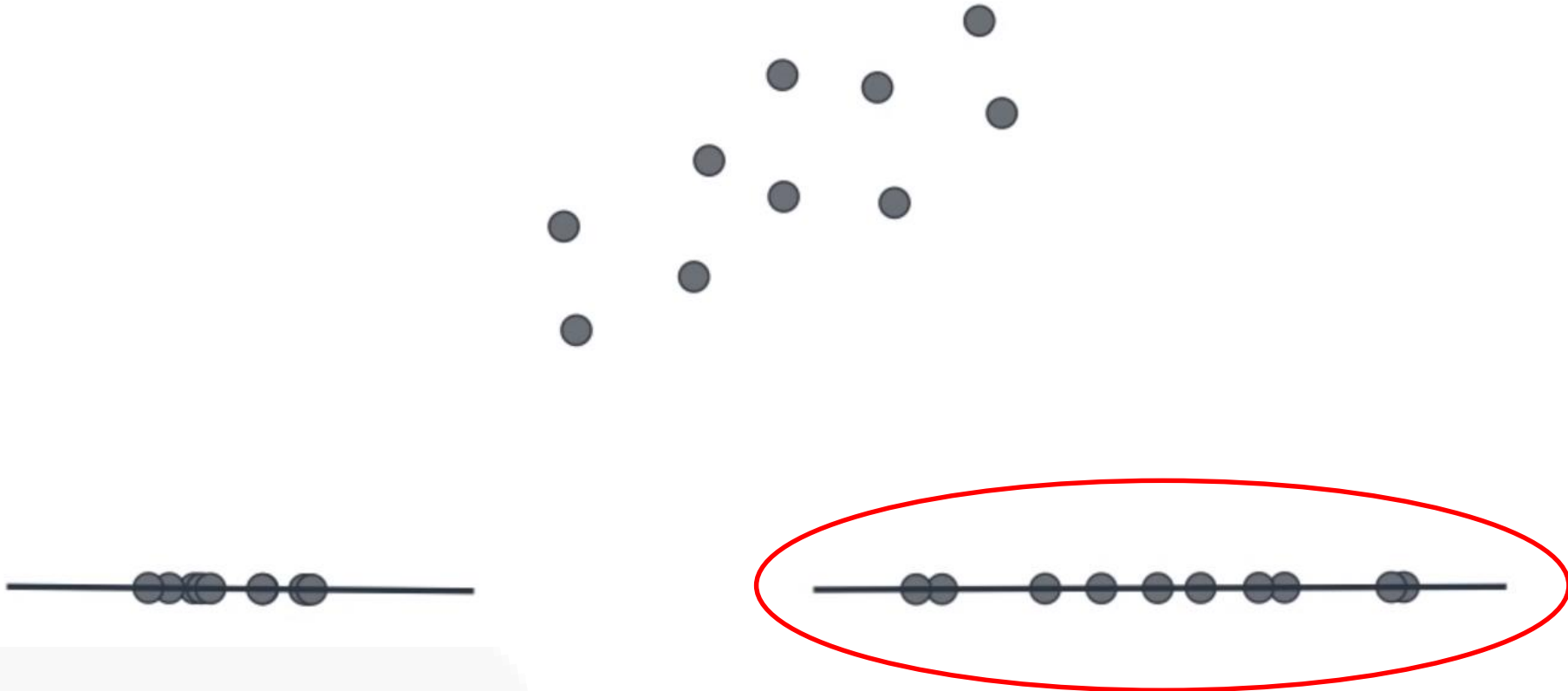
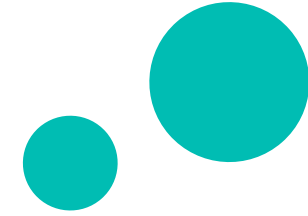
Principal Component Analysis (PCA)



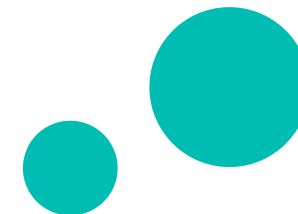






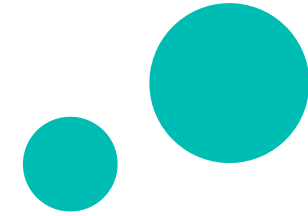


Principal Component Analysis (PCA)

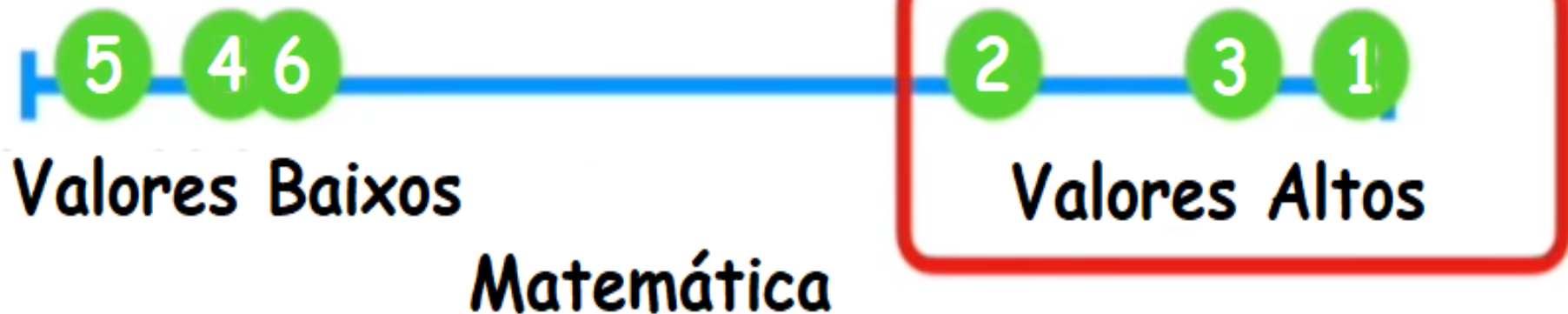


	Estudante 1	Estudante 2	Estudante 3	Estudante 4	...
Matemática	95	88	93	75	...
Português	96	79	98	81	...

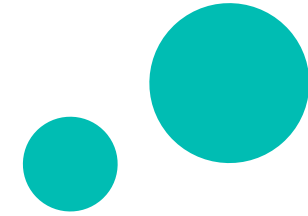
Principal Component Analysis (PCA)



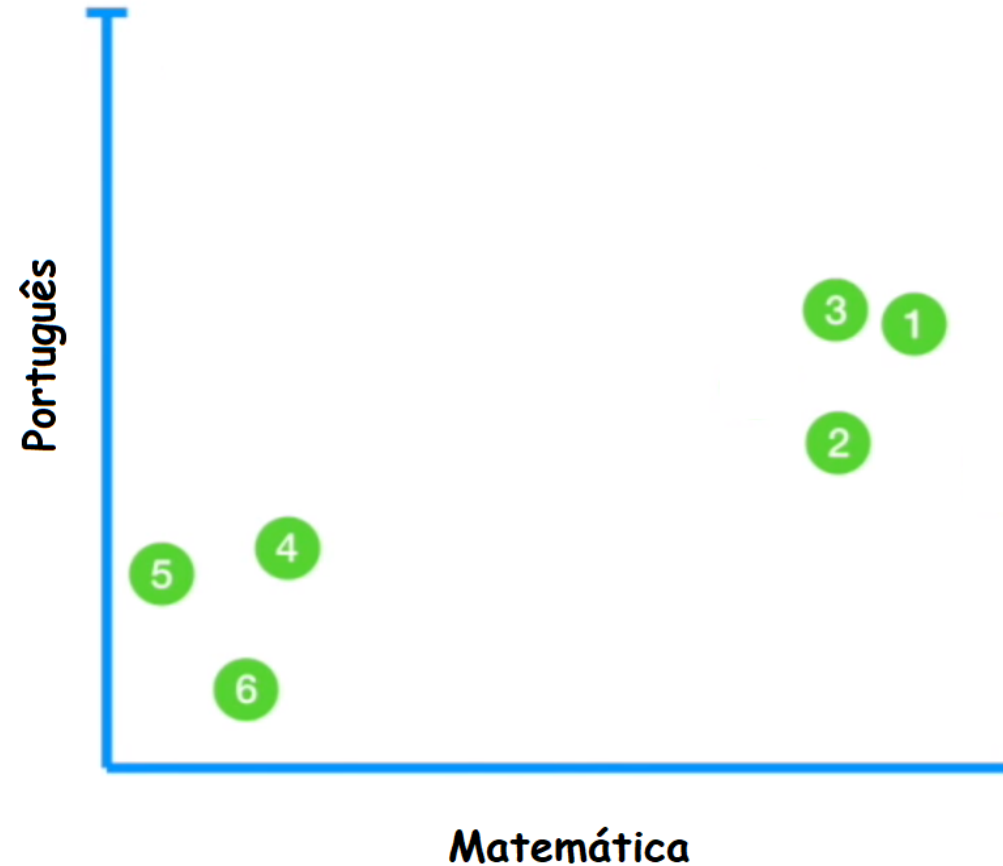
	Estudante 1	Estudante 2	Estudante 3	Estudante 4	...
Matemática	95	88	93	75	...
Português	96	79	98	81	...



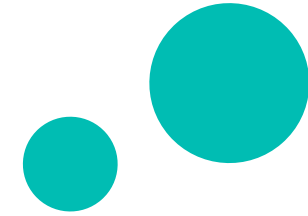
Principal Component Analysis (PCA)



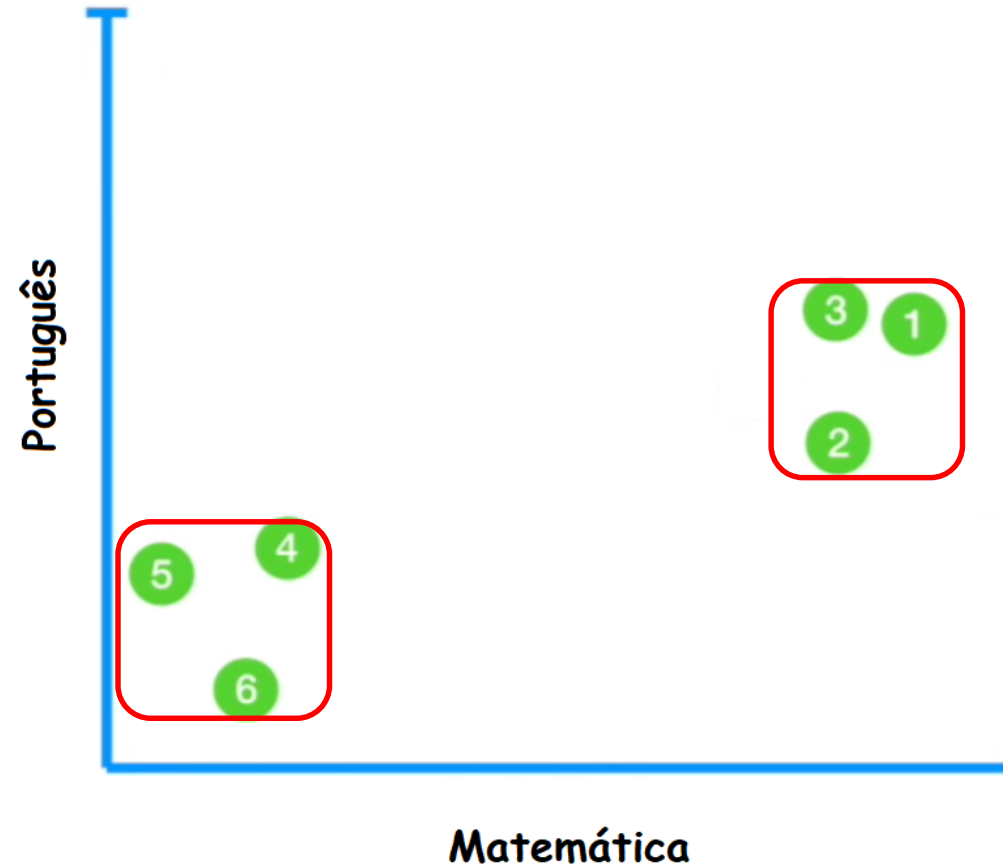
	Estudante	Estudante	Estudante	Estudante	...
	1	2	3	4	...
Matemática	95	88	93	75	...
Português	96	79	98	81	...



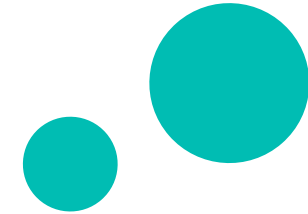
Principal Component Analysis (PCA)



	Estudante	Estudante	Estudante	Estudante	...
	1	2	3	4	...
Matemática	95	88	93	75	...
Português	96	79	98	81	...

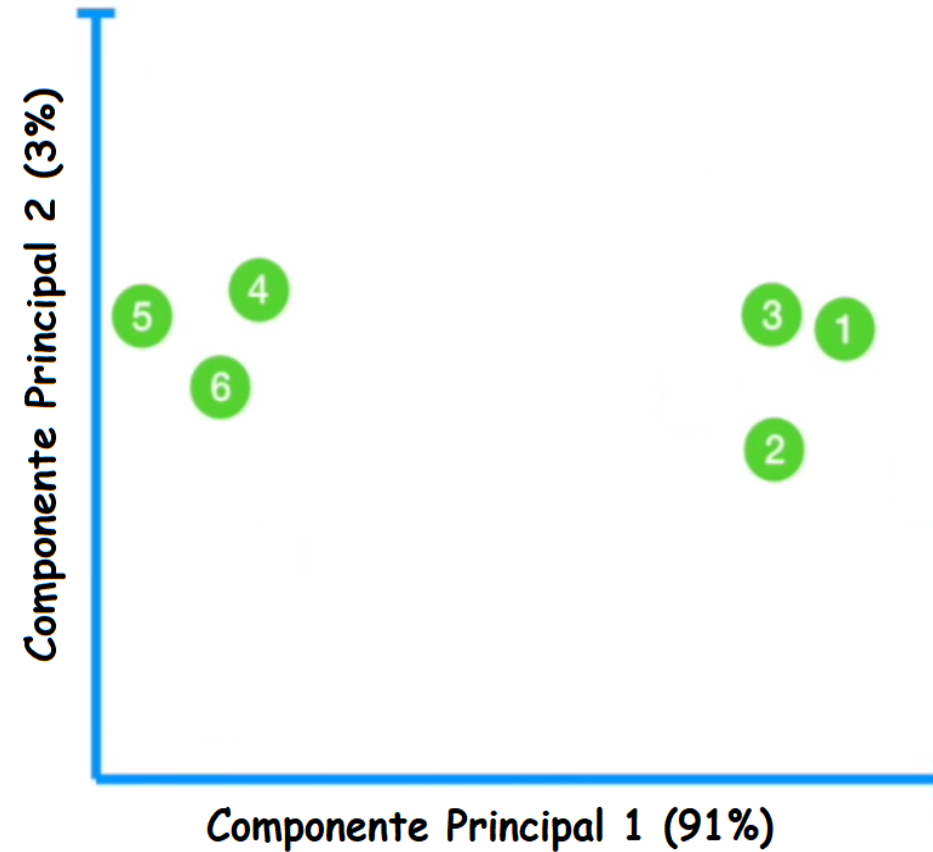


Principal Component Analysis (PCA)

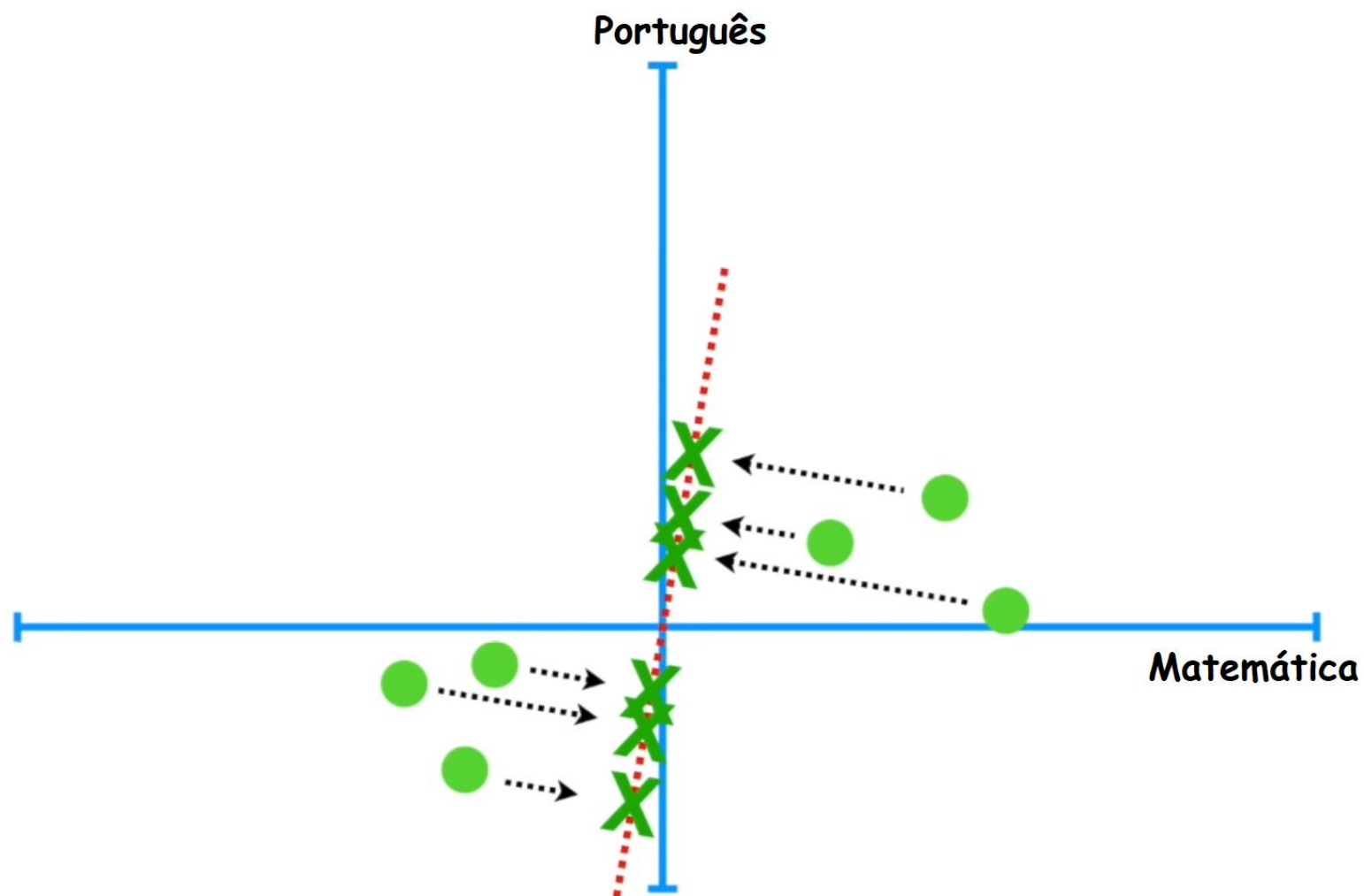


	Estudante	Estudante	Estudante	Estudante	...
	1	2	3	4	...
Matemática	95	88	93	75	...
Português	96	79	98	81	...

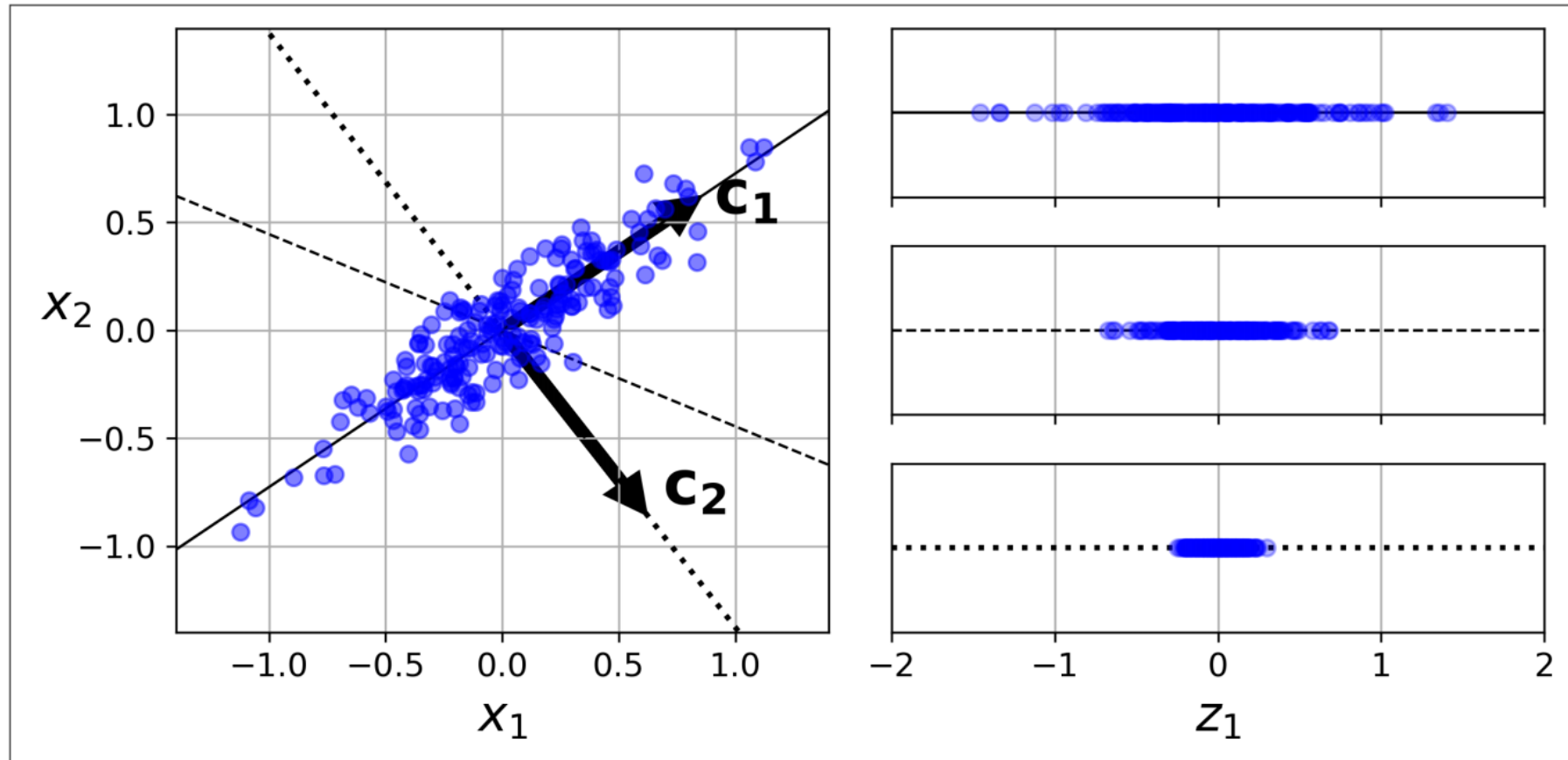
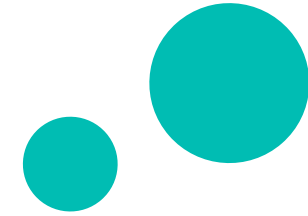
.

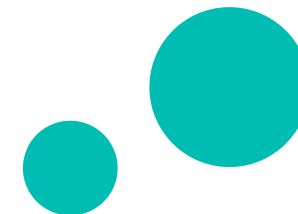


Principal Component Analysis (PCA)

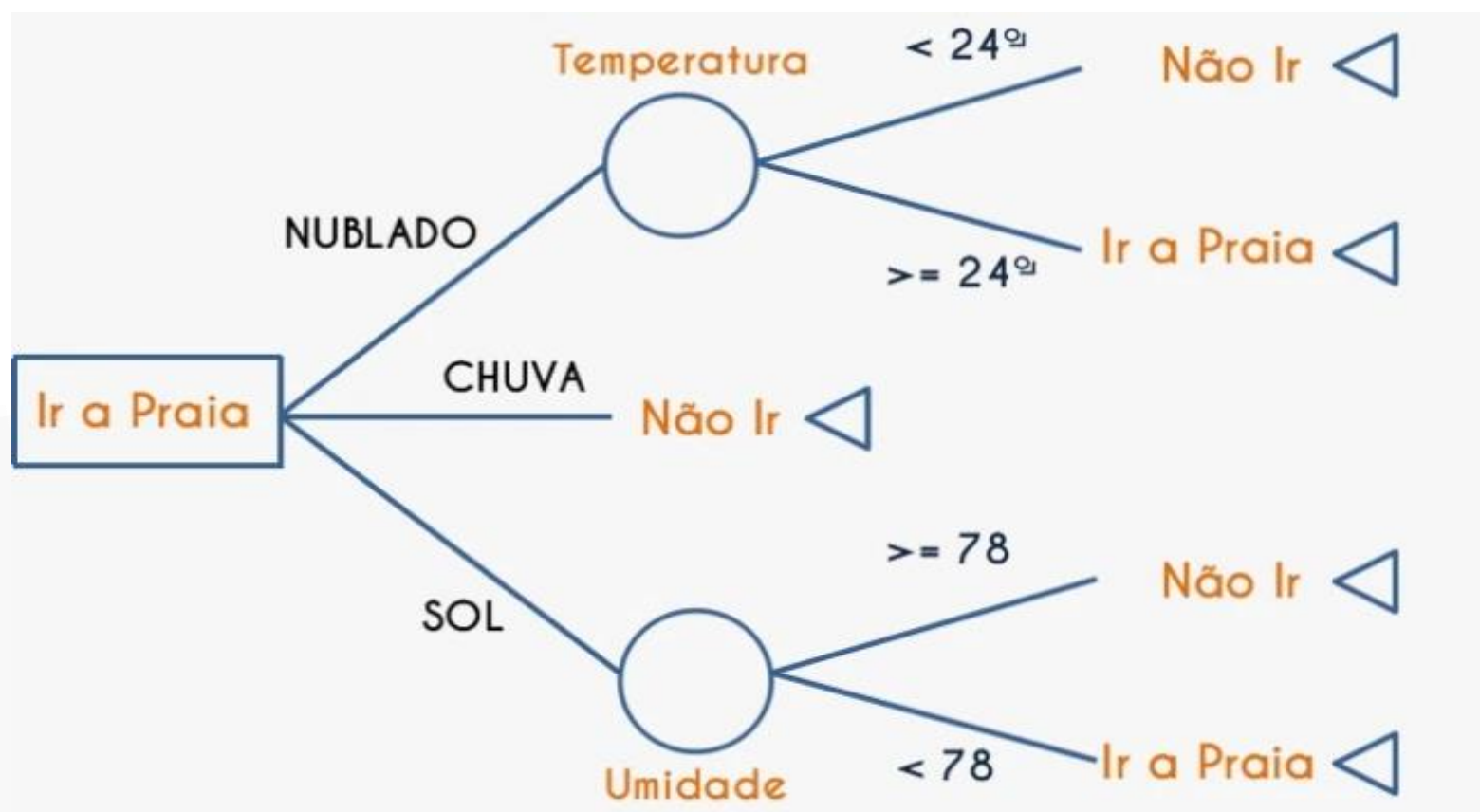


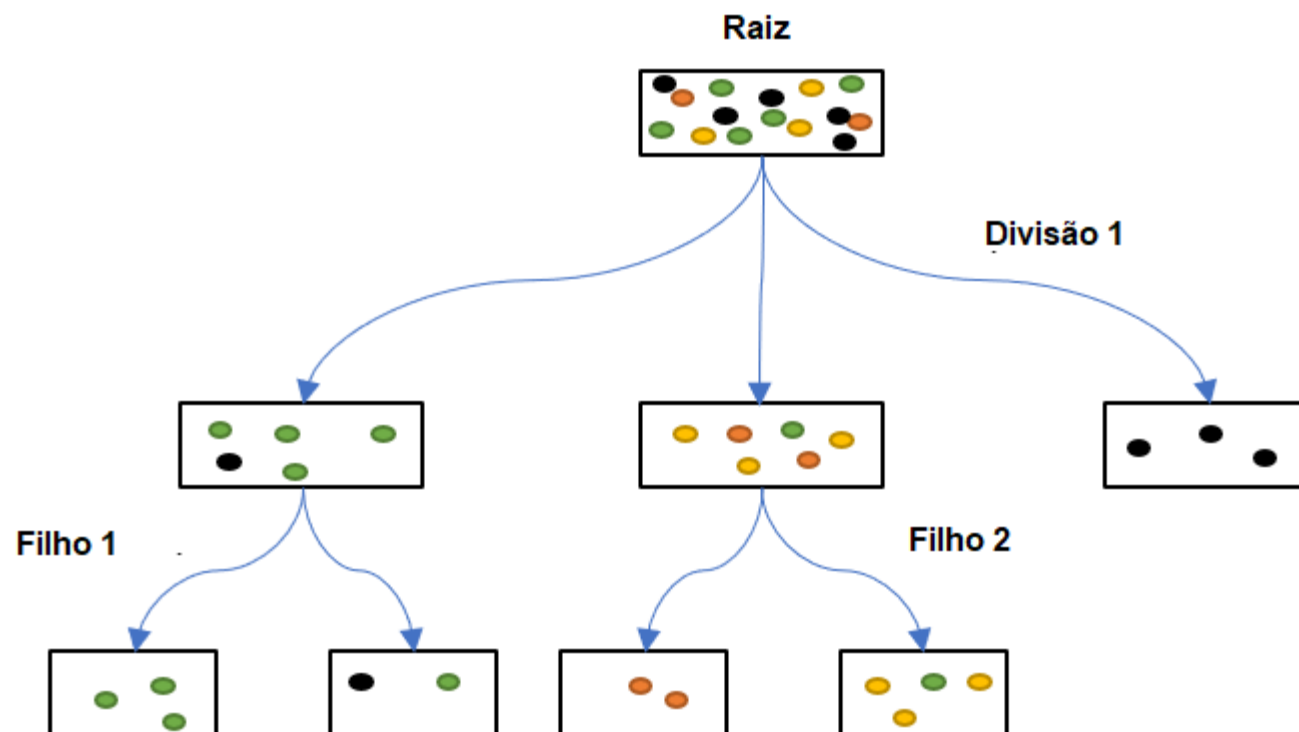
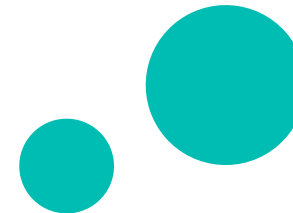
Principal Component Analysis (PCA)





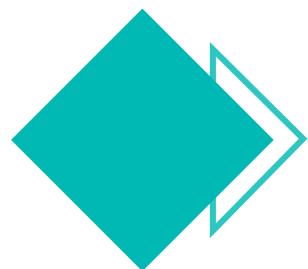
Vamos à Praia?



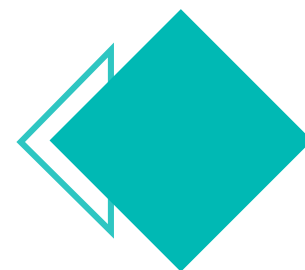
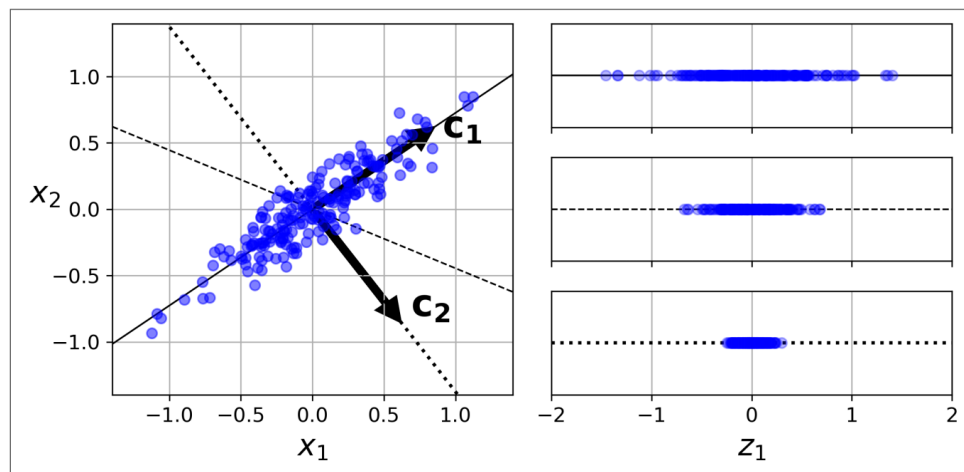


iGTi

Exemplo



colab





Na Aula de Hoje

1 Correlação+PCA.

2 Prática.



Próxima Aula

1 Dúvidas.

2 Desafio.