

INTRODUÇÃO À **ANÁLISE DE DADOS**



Fatec



São José do
Campos
Prof. Jessen Vidal

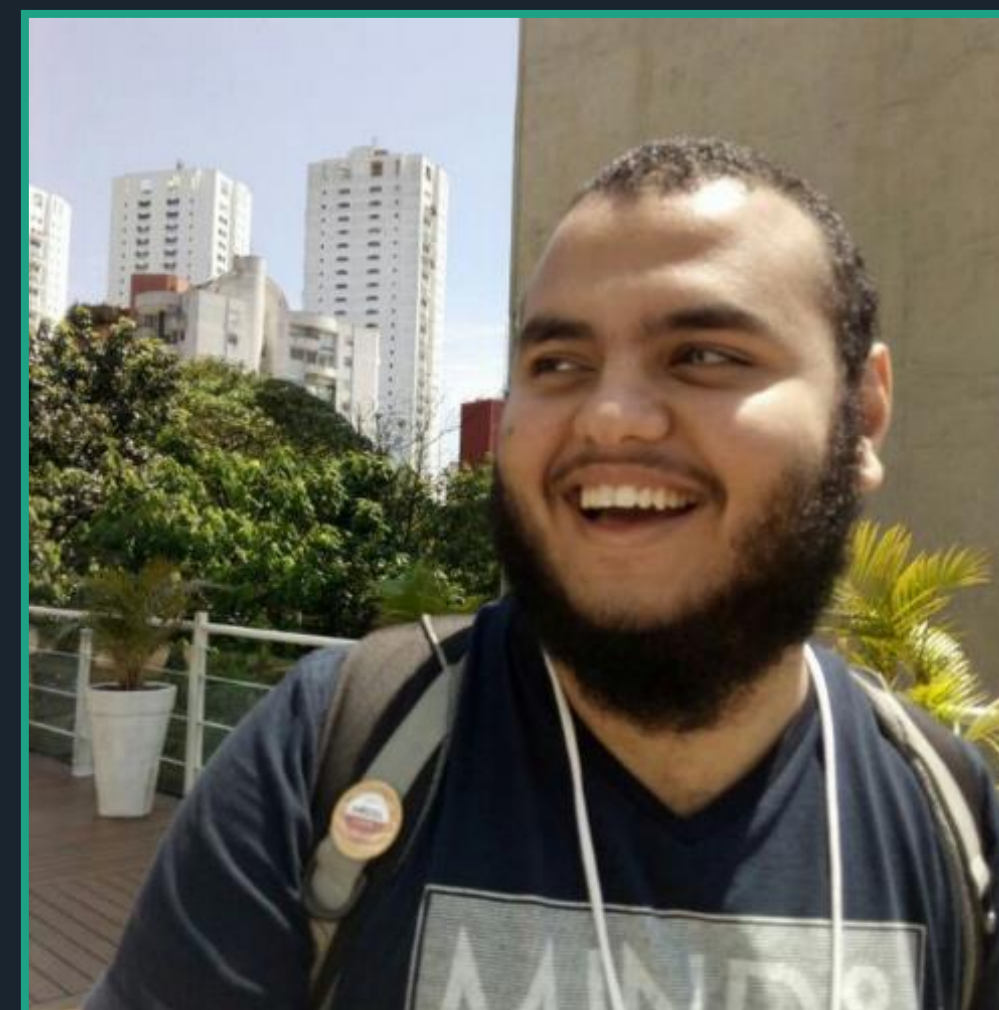
COMPANY

Apresentação

Meet our Awesome Team



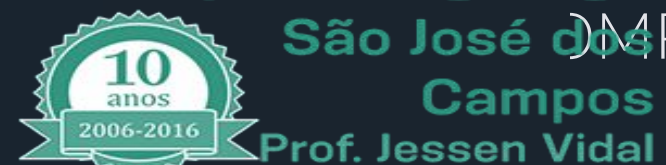
FELIPE CARVALHO



FELIPE MENINO



Fatec



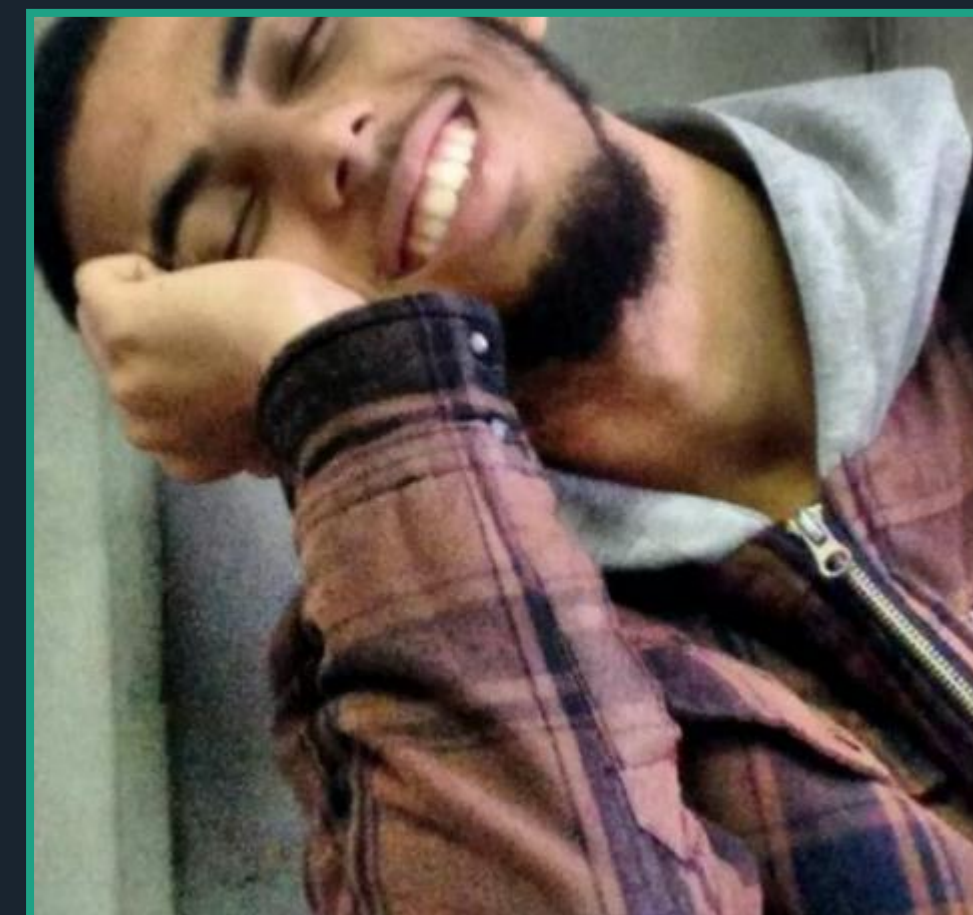
São José dos Campos
Prof. Jessen Vidal

Apresentação

Meet our Awesome Team



DANIELA ZARAMELLO



WESLEY LUIZ



LUCAS AUGUSTO



Disponível no Github

- ✓ Ambiente: <https://github.com/dataAt/ambiente>
- ✓ Apresentação: <https://github.com/dataAt/apresentacao>
- ✓ Códigos e leitura auxiliar:
<https://github.com/dataAt/introducao-analise-de-dados>

Fatec



São José do
Campos
Prof. Jessen Vidal

Divisão do Curso

Parte 01

D	S	T	Q	Q	S	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

Conceitos básicos da análise de dados e aplicações (Python ou R)

Fatec

10 anos
2006-2016

São José do
Campos
Prof. Jessen Vidal

Parte 02

D	S	T	Q	Q	S	S
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

Conceitos básicos da análise de dados e aplicações (Python ou R)

Introdução





Objetivo

Introduzir o processo básico de análise de, utilizando as linguagens de programação R e Python, para aqueles que querem ter o primeiro contato com esta atividade.

Desvendando a Análise de Dados

Sobre os Dados



Dado

É tudo aquilo que é registrado, para posteriormente, ser analisado



Formatos

Cada formato pode apresentar etapas diferentes de organização



DataFrame

São tabelas utilizadas para armazenar dados, podem ser manipulados através de linhas e colunas;



Peso

Entenda que todo dado tem um peso, contexto e impacto ao ser analisado.

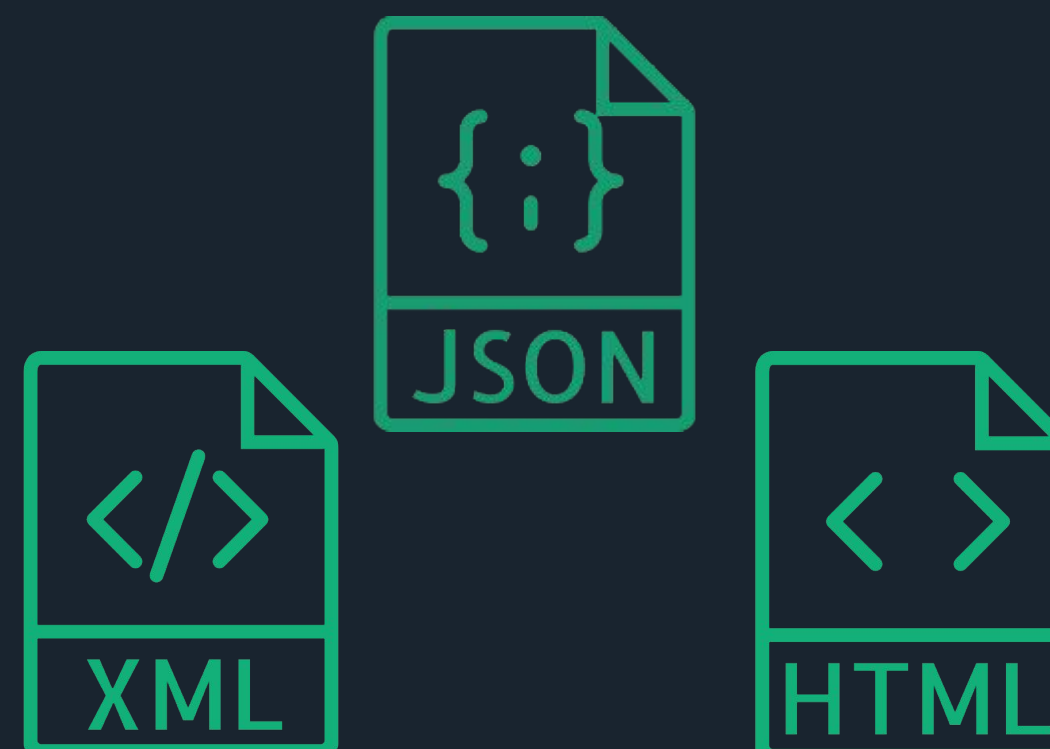
Sobre os dados

Formatos de Dados

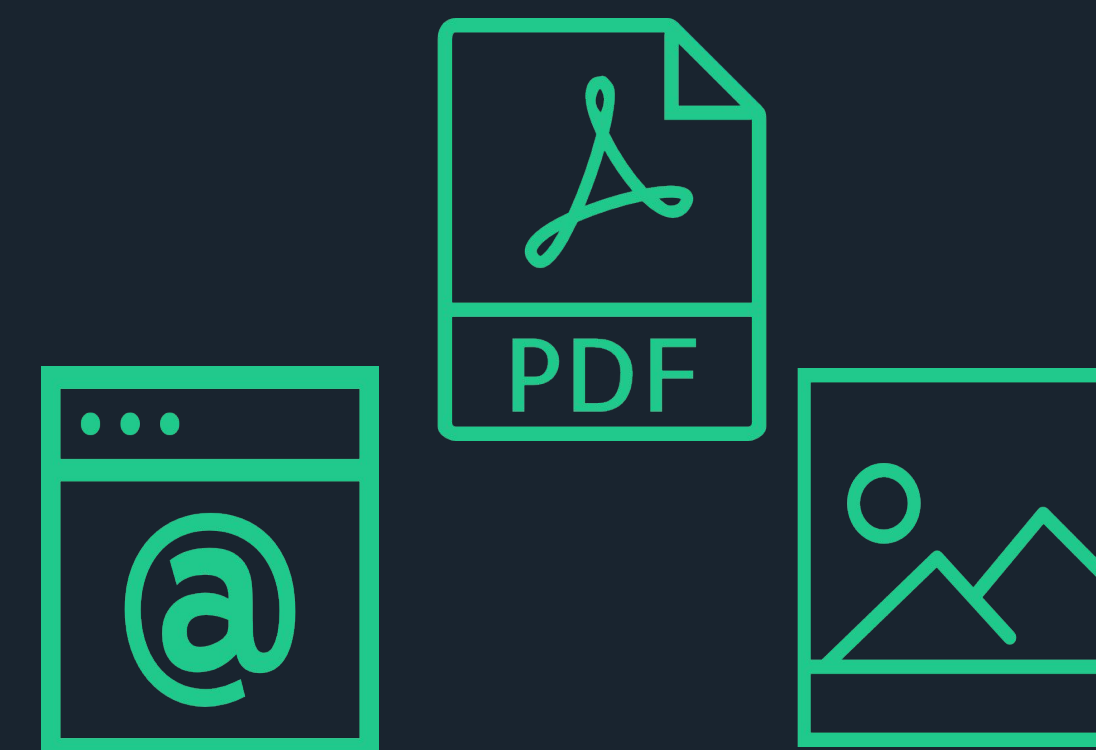
Dados Estruturados



Dados Semiestruturados



Dados não Estruturados



Análise de dados

“

Para John W. Tukey, grande parte da análise de dados é inferencial, ou seja, o ato de extrair informações de uma amostra em relação ao conjunto todo.

→ Não é um conceito novo!

– Tukey fez esta definição em 1961! No artigo *The Future of Data Analysis*

”

Analista de dados



De acordo com artigo publicado pela Dataquest, o analista de dados é um:

→ Agregador de valor para a companhia através respostas obtidas através dos dados, dessa forma, ajudando na tomada de decisão da empresa.

E possui como principais atividades:

Limpeza e organização de dados;

→ Uso de estatística para ter uma visão geral dos dados;

→ Análise de tendências encontradas nos dados;

→ Criação de visualização e *dashboards* para ajudar na interpretação e

→ tomada de decisão da empresa;

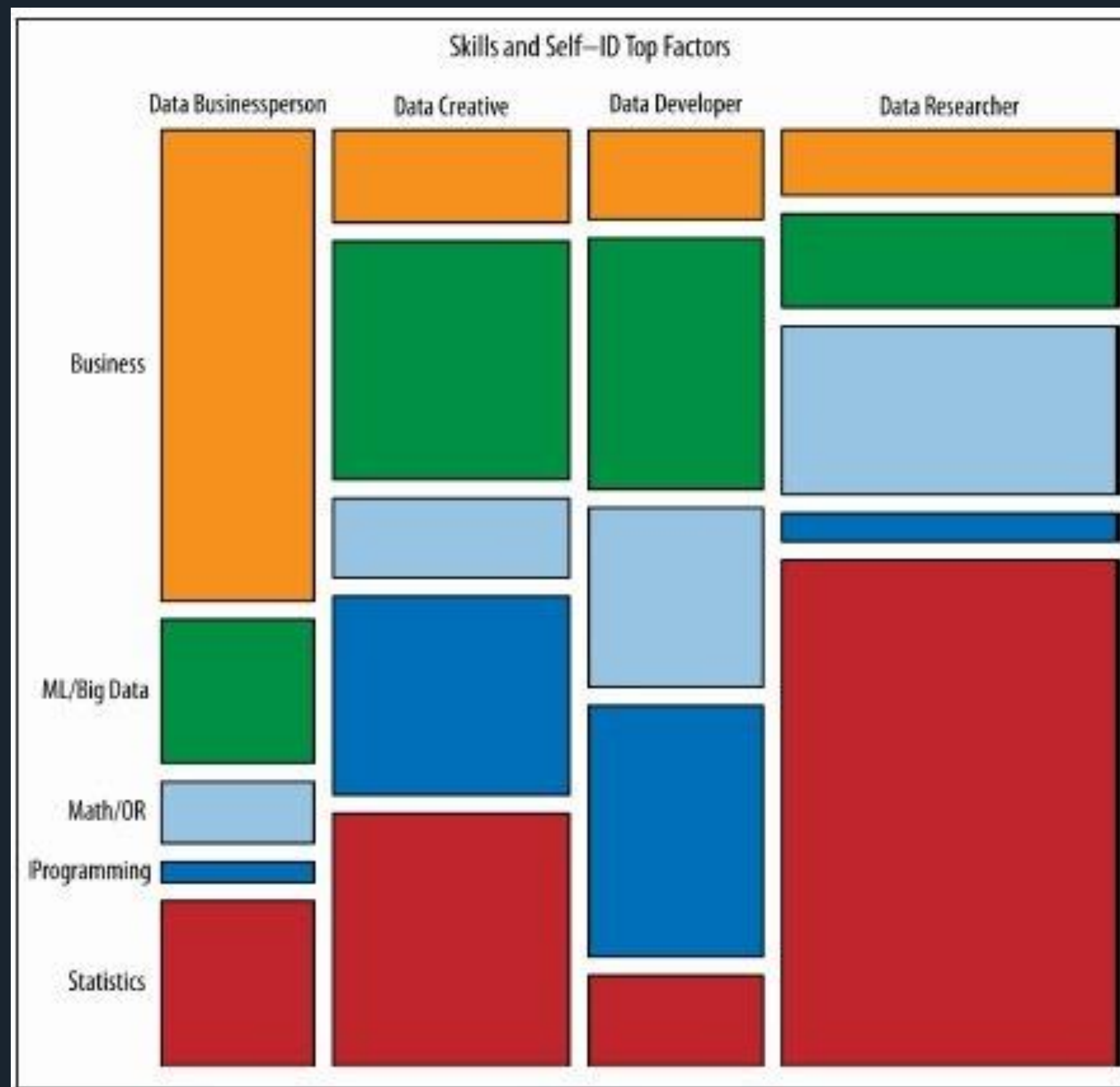
Apresentação dos resultados obtidos pela análise para os clientes.

→

Cientista de dados



Nem mesmo as pessoas que fazem isto tem uma definição concreta



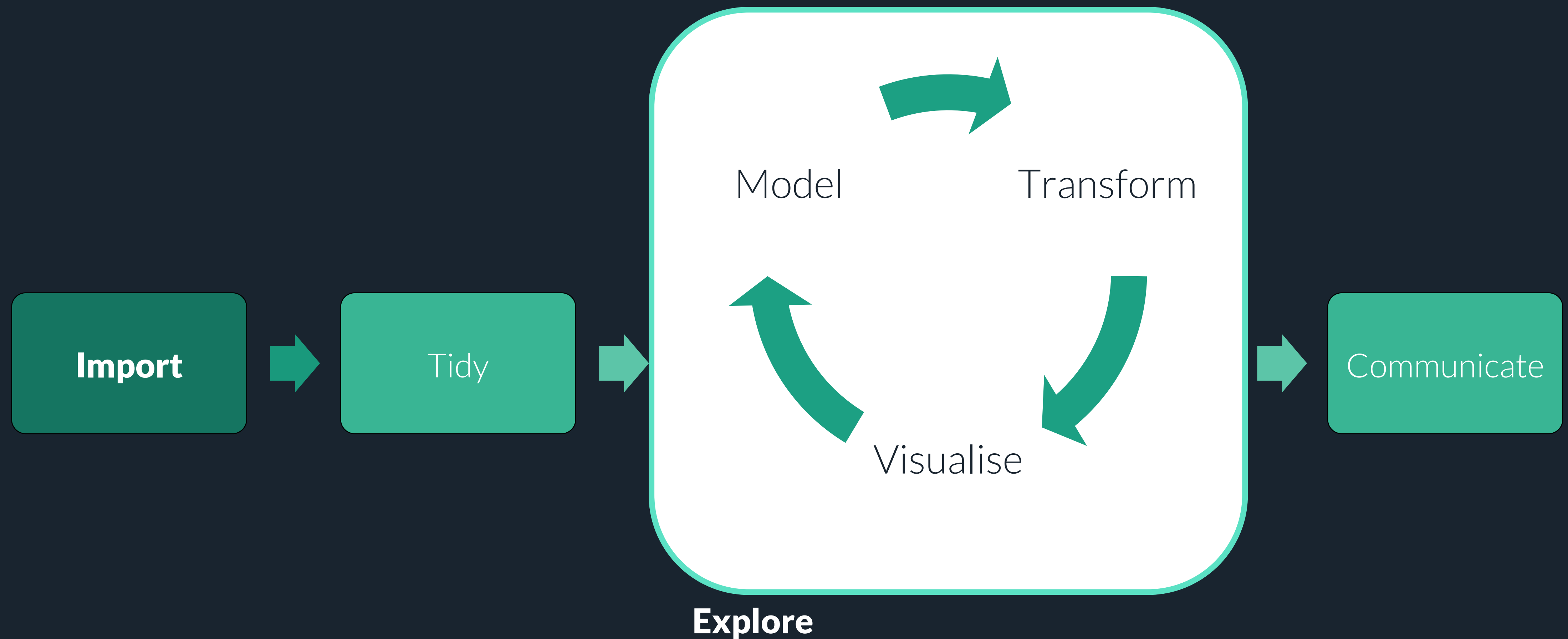
Cientista de dados

Por conta destas múltiplas interpretações, vamos utilizar a definição do kdnuggets

O analista busca resposta para o time e para o negócio. Por outro lado o cientista, cria modelos estatísticos de aprendizado de máquina, visualizações mais elaboradas e gera novas perguntas em relação aos dados.

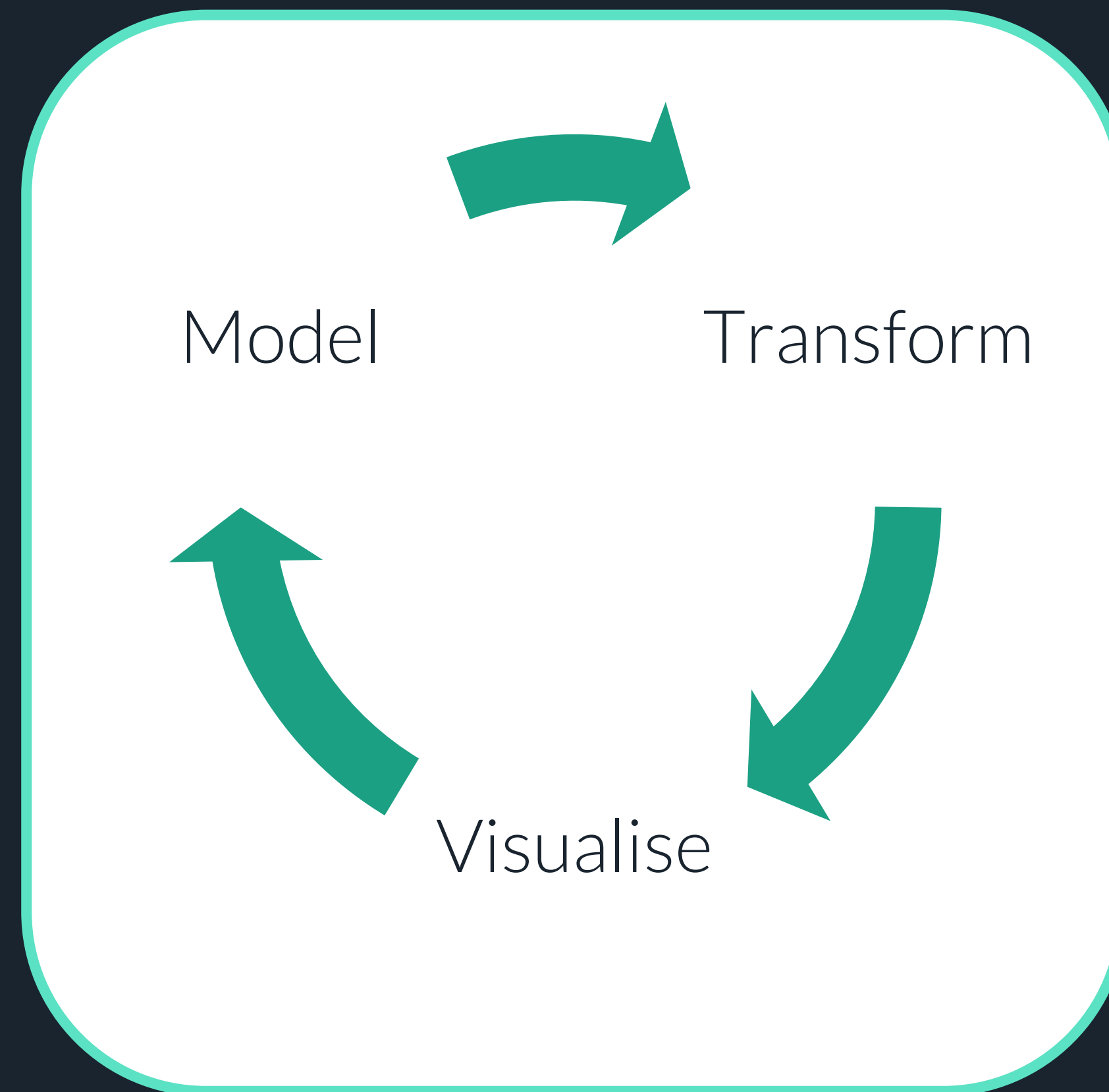


O Processo da **Análise de dados**



O Processo da Análise de dados

Análise exploratória de dados



Fazendo a Análise de dados



Fatec



São José do
Campos
Prof. Jessen Vidal

COMPANY



Ambiente



Markdown



CellToolbar

Simple spectral analysis

An illustration of the [Discrete Fourier Transform](#)

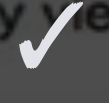
$$X_k = \sum_{n=0}^{N-1} x_n e^{-j2\pi kn/N}$$

In [2]: `from scipy`
`rate, x =`



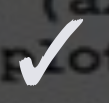
Documentos interativos;

And we can easily view

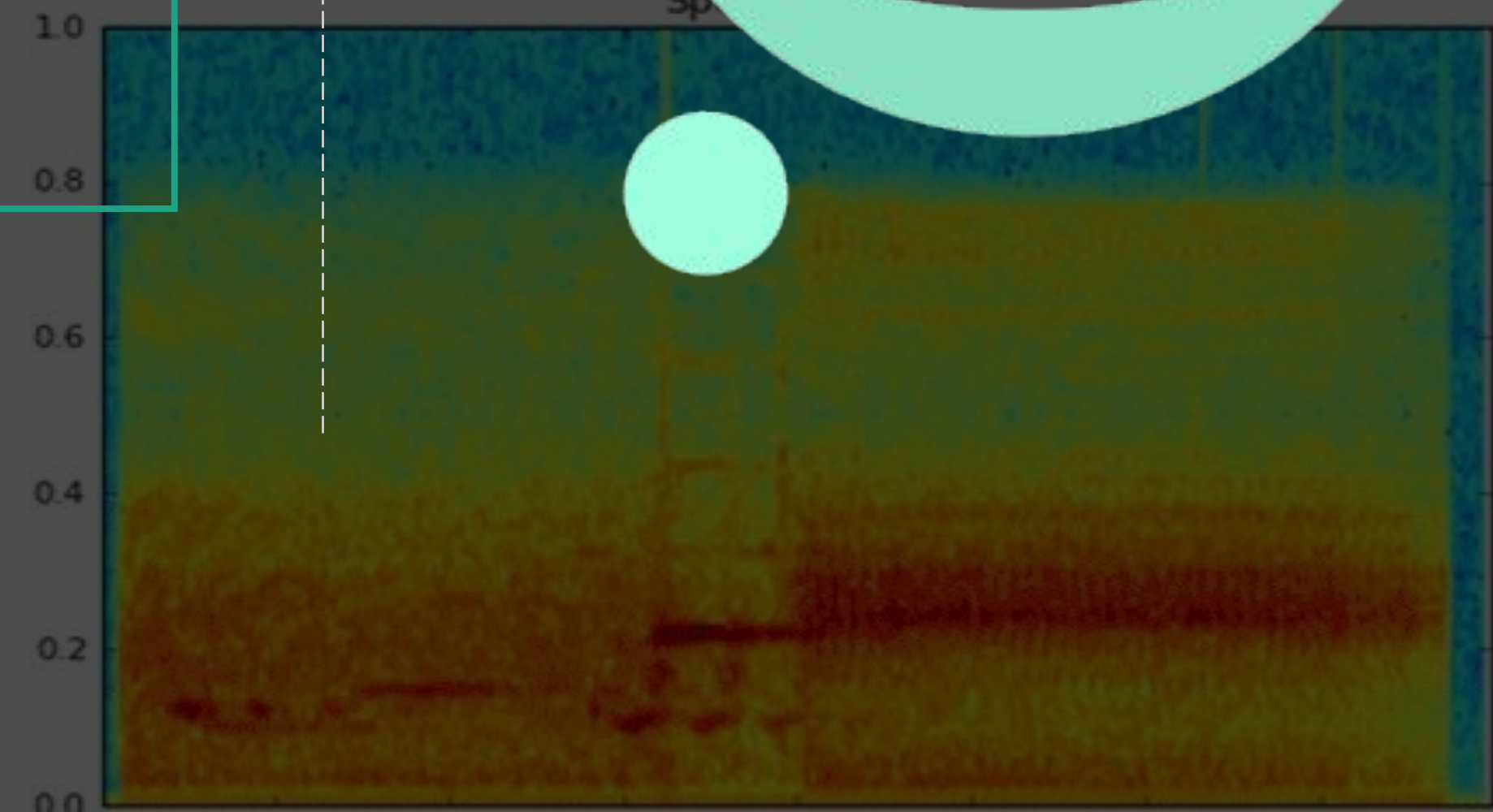
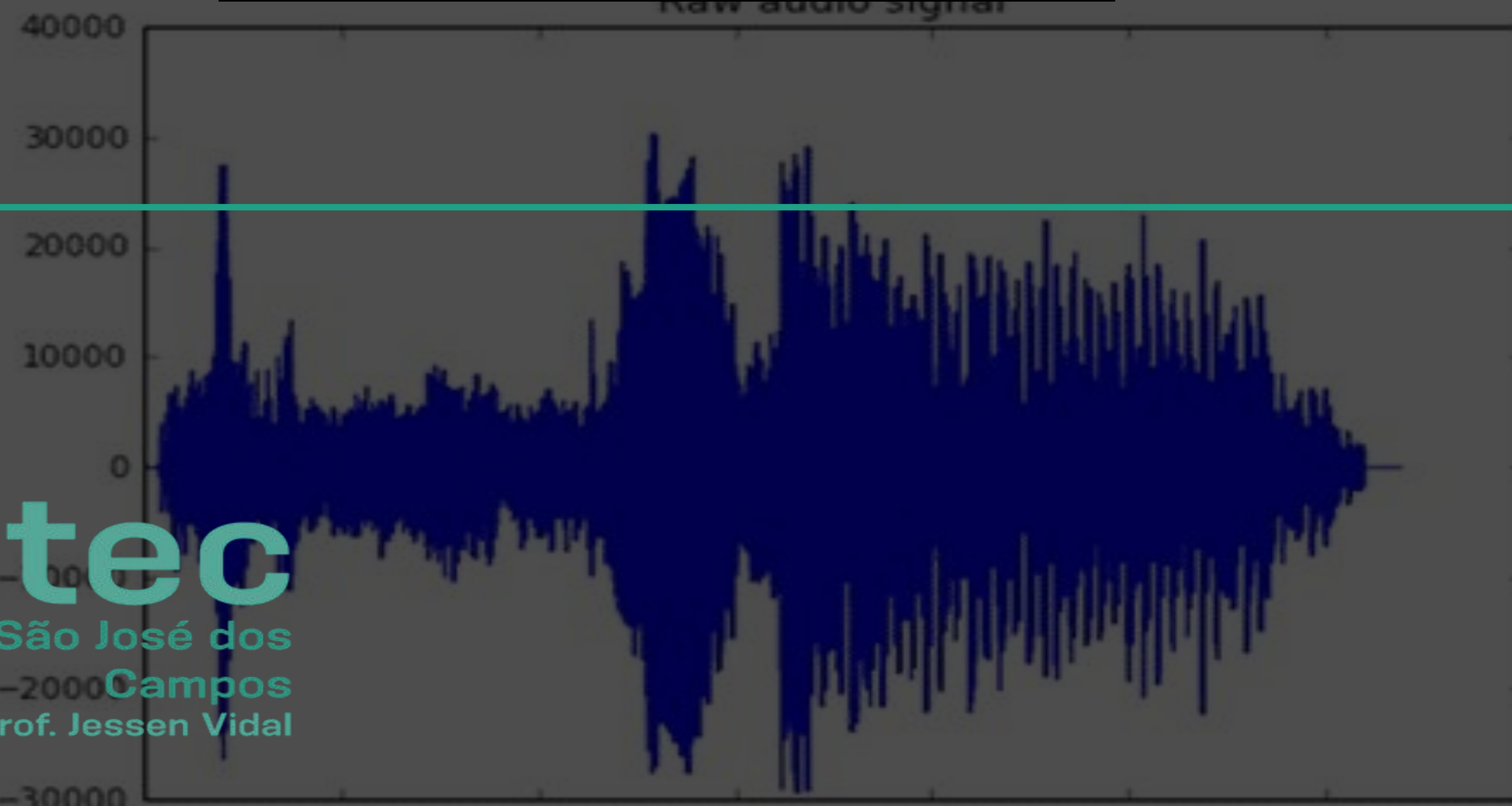


Apresentação bonita e amigável!

In [5]: `fig, (ax1`
`ax1.plot(`
`ax2.specg`



Permite a fácil distribuição dos resultados de análises.



Fatec

São José dos Campos
2006-2016
Prof. Jessen Vidal

Porque? R VS Python



Contexto ! Não existe a melhor linguagem para todos os casos, mas sim para certo contexto

