

**Tera**

# Aula #13

## Módulo 2

André Silveira, 17/mar/2018

# Intro

[Revelation, a Visual Poem](#)

[R2D3: A visual introduction to machine learning](#)

# Como será?

## MÓDULO 3

## ANÁLISE E VISUALIZAÇÃO DE DADOS

|                 |               |   |
|-----------------|---------------|---|
| <b>AULA 7</b>   | 8 /Mar        | Orientação a Objetos & Análise de dados com NumPy                     |
| <b>AULA 8</b>   | 9 a 12 /Mar*  | Statistical Thinking in Python (Part 1 e 2)                           |
| <b>AULA 9</b>   | 13 /Mar       | Análise de dados com Pandas   |
| <b>AULA 10</b>  | 15 /Mar       | Visualização de dados com matplotlib                                  |
| <b>AULA 11</b>  | 13 a 16 /Mar* | Exploratory Data Analysis   |
| <b>AULA 12</b>  | 17 /Mar       | How to get data: SQL Queries & APIs                                   |
| <b>AULA 13</b>  | 17 /Mar       | Análise exploratória & Apresentação de insights                       |
| <b>AULA 14</b>  | 20 /Mar       | Execução de projetos: Análise exploratória & Apresentação de insights |
| <b>EVENTO 1</b> | 22 /Mar       | Visita técnica a empresa  |

• Aula EAD



# Como será?

- Blend **teoria** e **prática**
- NumPy, Pandas, SQL-Query&API, Matplotlib, Statistics
- **Análise exploratória**
- Data Science & Machine Learning **for Business**
- Insumos para aula **prática** de **terça-feira**



# Como será?

- Blend **teoria e prática**
- NumPy, Pandas, SQL-Query&API
- **Análise exploratória**
- Data Science & Machine Learning
- Insumos para aula **prática** de **terceira**

Racional

Método

Contexto

Valor



# Agenda

1. Warmup
2. Intro
3. Como será?
4. Flow
5. Análise guiada, Aula 13
6. Considerações sobre o projeto, Aula 14
7. Decompression e feedback



# Flow

- Diferenciar **tipos de análises** pelo método, contexto de negócio e equipe de trabalho
- Conhecer **obstáculos e limitações comuns de bases reais** internas e externas a empresa
- Aplicar técnicas de avaliação, correção e **organização do dataset** para trabalho
- Aplicar técnicas de como **interpretar variáveis** originais e criar variáveis secundárias
- Conhecer abordagens que guiam a **descoberta um data-insight**
- Reconhecer aspectos da apresentação que zelam pela **percepção de valor**, impacto e melhoria contínua





## Contexto





# Contexto

- Em quais situações surge a demanda por uma análise?
- Qual a melhor forma de usar a sua energia e entregar valor de forma eficiente?

# Obtenção dos dados

# Verificação e correção de ruídos

- Sejam realistas, quanto tempo eu vou levar para começar a criar algo novo?
- Meu dataset está consistente?

# Intervalo

# Exploração

→ O que importa neste dataset? Por que importa?

# Exploração

→ Qual hipótese vou avaliar?

# Resultados

- O que esse data-insight vai mudar na rotina das pessoas?



# Apresentação

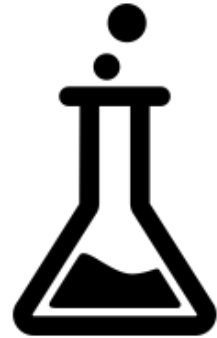
- O que é esperado por quem vai atender a apresentação?
- Você vê o impacto gerado pelo data-insight? Se não, por que não?
- Que ação esse data-insight gerou?



# Projeto: inputs



Google Analytics



...mind the time!

Obrigado!



# Material suplementar

- INMET: <http://www.inmet.gov.br/portal/index.php?r=home2/index>
- GA Demo account: <https://analytics.google.com/analytics/web/demoAccount>
- Google Merchandise Store: <https://shop.googlemerchandisestore.com/>
- Paper Campos do Jordao:  
<https://link.springer.com/article/10.1007/s10064-003-0191-8>



# Material suplementar

**BDMEP - Série Histórica - Dados Diários**

Período - Data início (dd/mm/aaaa) :  fim :

Região :  (OU) Estado :

Selecionar Variáveis :

- ☐ Precipitação(mm)
- ☐ Temp Máxima(°C)
- ☐ Insolação(horas)
- ☐ Evaporação do Piche(mm)
- ☐ Umidade Relativa Média(%)
- ☐ Temp Mínima(°C)
- ☐ Temperatura Compensada Média(°C)
- ☐ Velocidade Vento Média(mps)