📊 Data Science Workflow Canvas*

Start here. The sections below are ordered intentionally to make you state your goals first, followed by steps to achieve those goals. You're allowed to switch orders of these steps!

Title: **ANÁLISE DE DADOS DE TEMPERATURA AMBIENTE VERSUS LOGS DE ALERTA EM UM DATA CENTER**

1 Problem Statement
What problem are you trying to solve?
What larger issues do the problem address?

- Todas as vezes que são registradas altas temperaturas (maior que 33 graus Celcius), é temos problemas de hardware nos servidores, dentro do nosso data center?

2 Outcomes/Predictions
What prediction(s) are you trying to make?
Identify applicable predictor (X) and/or target (y) variables.

- Resultados: verificar existência de relação entre alta temperatura e logs de alerta com erro (Critical)

- Variáveis preditoras: atributos de temperatura dos registros.

- Variáveis alvos: atributos de alerta registrados

3 Data Acquisition
Where are you sourcing your data from?
Is there enough data? Can you work with it?

- Logs de temperatura e de alertas de 30 servidores de um data center em Manaus, coletados entre 2014 e 2020

4 Modeling
What models are appropriate to use given your outcomes?

- PCA (principal component analisys)

- Árvore de decisão

5 Model Evaluation
How can you evaluate your model's performance?

- Árvore de decisão:
  - Índice de acuracidade;
  - classification_report;

6 Data Preparation
What do you need to do to your data in order to run your model and achieve your outcomes?

- Indexar os datasets de temperaturas e de alertas, por nome de servidor e data de registro de log.

- Combinar os datasets

- Usar One-Hot Encoding para tratar atributos categóricos relevantes

✅ Activation

When you finish filling out the canvas above, now you can begin implementing your data science workflow in roughly this order.

1 Problem Statement → 2 Data Acquisition → 3 Data Prep → 4 Modeling → 5 Outcomes/Preds → 6 Model Eval

* **Note:** This canvas is intended to be used as a starting point for your data science projects. Data science workflows are typically nonlinear.