

*Carlos Eduardo Pires Panzarini*

*Estela de Souza Moraes*

*Iara Lourenço da Silva*

*Kaline Maria de Moura*

*Laís Uchôa Rabelo Mendes*

*Ronaldo Souza Chaves*

## **1. Escolha do dataset**

### **1.1 Justificativa**

<https://catalog.data.gov/dataset/crimes-2001-to-present>

## **2. Gestão do projeto**

### **2.1 Requisitos**

- Limpeza e tratamento dos dados
- Análise exploratória de dados
- Métricas do DW (dimensões e tabelas-fato)
- Criação de um Data Warehouse
- Definição de questões a serem respondidas com machine learning
- Conexão do DW com notebook externo
- Análise de machine learning

### **2.2 Estrutura Analítica do Projeto – abordagem ágil**

#### *1. Definição do projeto*

##### *1.1 Escolha do dataset*

##### *1.2 Levantamento de requisitos*

#### *2. Sprint 1 – Preparação dos dados*

##### *2.1 Modelo lógico*

##### *2.2 Extração e limpeza dos dados*

##### *2.3 Tratamento dos dados*

##### *2.4 Revisão*

##### *2.5 Retrospectiva*

#### *3. Sprint 2 – Data Warehouse*

##### *3.1. Criação do Data Warehouse (PostgreSQL)*

##### *3.2 Validação do DW*

##### *3.3 Teste unitário*

##### *3.4 Apresentação de slides (primeira prévia)*

##### *3.5 Revisão*

*3.6 Retrospectiva*

*3.7 Levantamento das questões a serem respondidas*

*4. Sprint 3 – Machine Learning*

*4.1 Análise exploratória*

*4.2 Conexão do DW com o notebook*

*4.3 Codificação*

*4.4 Testar unitário*

*4.5 Teste de validação*

*4.6 Teste de integração*

*4.7 Revisão*

*4.8 Retrospectiva*

*5. Encerramento*

*5.1 Entrega*

*5.2 Apresentação*

## **2.3 Cronograma**

Definição do projeto

1 semana e meia (18/09 - 29/09)

Sprint 1

1 semana (2/10 - 09/10)

Sprint 2

1 semana (10/10 - 16/10)

Sprint 3

3 semanas (17/10 - 06/11)

Encerramento

2 semanas (07/11 - 20/11)