# Dicionário de Dados do Dataset de Sprints

Este documento descreve a estrutura, significado e origem de cada campo presente no dataset utilizado para análise e modelagem de produtividade em sprints de desenvolvimento.

|  |  |  |  |
| --- | --- | --- | --- |
| Campo | Tipo | Descrição | Origem |
| sprint\_id | string | Identificador único da sprint | Simulado ou API |
| data\_inicio | date | Data de início da sprint | Simulado ou API |
| data\_fim | date | Data de fim da sprint | Simulado ou API |
| qtd\_membros | int | Número de desenvolvedores na sprint | Simulado ou API |
| duracao\_dias | int | Duração da sprint em dias corridos | Calculado |
| cartoes\_previstos | int | Total de cartões planejados para entrega | Simulado |
| cartoes\_entregues | int | Total de cartões realmente entregues | Simulado ou API |
| story\_points\_previstos | int | Total de story points planejados | Simulado ou API |
| story\_points\_entregues | int | Total de story points entregues | Simulado ou API |
| tipo\_dominio | string | Tipo de projeto (Web, API, Mobile, Dados, etc.) | Simulado |
| complexidade\_media | float | Complexidade média dos cartões (1 a 5) | Simulado |
| percentual\_bugs | float | % de cartões classificados como bugs | Simulado ou API |
| percentual\_retrabalho | float | % de cartões que voltaram para a lista inicial | Simulado ou API |
| velocidade\_passada | float | Média de entrega (story points) nas 3 sprints anteriores | Calculado |
| produtividade\_estimada | float | Estimativa: story points por desenvolvedor | Calculado |

## Uso do Dataset

Este dataset será utilizado como base de treino para algoritmos de machine learning (Random Forest, Redes Neurais, etc.) com o objetivo de prever a produtividade esperada (story points por desenvolvedor) em novas sprints, considerando variáveis contextuais como tipo de projeto, complexidade e histórico de entregas.