# Contrução da Tabela de Feature Engineering

CREATE OR REPLACE TABLE `ivory-enigma-402100.CrimeStudy.01\_crime\_features` AS

SELECT

    IFNULL(clearance\_status, 'UNKNOWN\_STATUS') AS clearance\_status,

    SAFE\_CAST(IFNULL(FORMAT\_DATE('%Y-%m-%d', SAFE\_CAST(clearance\_date AS DATE)), 'UNKNOWN\_DATE') AS STRING) AS clearance\_date,

    IFNULL(district, 'UNKNOWN\_DISTRICT') AS district,

    -- Extrai o dia da semana (nome completo) como uma string

    SAFE\_CAST(FORMAT\_DATETIME('%A', timestamp) AS STRING) AS day\_of\_week\_name,

    -- Extrai o ano da data de conclusão como um número inteiro

    SAFE\_CAST(IFNULL(EXTRACT(YEAR FROM SAFE\_CAST(clearance\_date AS DATE)), -1) AS INT64) AS clearance\_year,

    -- Caso o tipo primário do crime seja nulo, atribui 'UNKNOWN'

    IFNULL(primary\_type, 'UNKNOWN\_CRIME\_TYPE') AS crime\_type

FROM

    `bigquery-public-data.austin\_crime.crime`

WHERE

    -- Garante que apenas registros com tipo primário e timestamp válidos sejam considerados

    primary\_type IS NOT NULL

    AND timestamp IS NOT NULL;

d

# Criação dos Modelos

## Modelo 1 ( crime\_type\_predictor )

CREATE OR REPLACE MODEL `ivory-enigma-402100.CrimeStudy.02\_crime\_type\_predictor`

OPTIONS(

    model\_type='LOGISTIC\_REG', -- Correção aqui

    labels=['crime\_type']

) AS

SELECT

    clearance\_status,

    clearance\_date,

    district,

    day\_of\_week\_name,

    clearance\_year,

    crime\_type

FROM

    `ivory-enigma-402100.CrimeStudy.01\_crime\_features`

WHERE

    crime\_type != 'UNKNOWN\_CRIME\_TYPE'; -- Excluindo a categoria 'UNKNOWN' para um treinamento mais focado

## Modelo 2 ( district\_predictor )

CREATE OR REPLACE MODEL `ivory-enigma-402100.CrimeStudy.03\_district\_predictor`

OPTIONS(

    model\_type='LOGISTIC\_REG', -- Correção aqui

    labels=['district']

) AS

SELECT

    clearance\_status,

    clearance\_date,

    day\_of\_week\_name,

    clearance\_year,

    crime\_type,

    district

FROM

    `ivory-enigma-402100.CrimeStudy.01\_crime\_features`

WHERE

    district != 'UNKNOWN\_DISTRICT'; -- Excluindo a categoria 'UNKNOWN'

# Tabelas Finais para Análise

SELECT

    predicted\_crime\_type,

    -- Inclua as colunas de entrada para entender a previsão

    clearance\_status,

    clearance\_date,

    district,

    day\_of\_week\_name,

    clearance\_year

FROM

    ML.PREDICT(MODEL `ivory-enigma-402100.CrimeStudy.02\_crime\_type\_predictor`,

               (

                   SELECT

                       clearance\_status,

                       clearance\_date,

                       district,

                       day\_of\_week\_name,

                       clearance\_year

                   FROM

                       `ivory-enigma-402100.CrimeStudy.01\_crime\_features`

                   -- Adicione aqui quaisquer filtros para os dados que você quer prever

                   -- Por exemplo, para prever em dados que não foram usados no treinamento:

                   -- WHERE RAND() < 0.2 -- Assumindo que você separou 20% para teste

               ))

LIMIT 100; -- Limite para não exibir muitas linhas