**Possíveis Variáveis-Alvo**

1. **Causa do Óbito (CAUSABAS ou CAUSABAS\_O)**:
   * Você pode tentar prever a causa do óbito com base em características demográficas e sociais. Isso pode ajudar a entender se existem padrões associados a causas específicas.
2. **Faixa Etária (idade)**:
   * A partir da idade, você pode criar faixas etárias (por exemplo, jovens, adultos, idosos) e tentar prever a faixa etária com maior risco de suicídio. Isso pode ser importante para direcionar políticas de saúde mental.
3. **Escolaridade (ESC)**:
   * Pode-se prever se uma pessoa tem um nível de escolaridade que a coloca em risco maior de suicídio. Esse modelo pode ajudar a direcionar programas de educação e conscientização.
4. **Sexo (SEXO)**:
   * Você pode analisar e prever a distribuição de suicídios entre os sexos. Isso pode ajudar a entender se existe um padrão de risco mais elevado em um dos gêneros.
5. **Estado Civil (ESTADO CIVIL)**:
   * A partir do estado civil, você pode tentar prever se certas condições (casado, solteiro, divorciado) estão associadas a um maior risco de suicídio.

**Variáveis Independentes**

Para treinar seu modelo, você pode usar as seguintes variáveis independentes:

* **Estado (estado)**: Pode influenciar a disponibilidade de serviços de saúde mental e o suporte social.
* **Ano (ano)**: Analisar tendências ao longo dos anos pode revelar se certos períodos têm um aumento no risco.
* **Data de Nascimento (DTNASC)**: Pode ser utilizada para calcular a idade diretamente ou inferir outras características.
* **Sexo (SEXO)**: Como mencionado, pode ser uma variável independente para várias previsões.
* **Raca/Cor (RACA/COR)**: Pode fornecer insights sobre disparidades raciais em saúde mental.
* **Escolaridade (ESC)**: Como mencionado, pode ser um preditor significativo do risco.
* **Ocupação (OCUP)**: A ocupação pode influenciar o estresse e, consequentemente, o risco de suicídio.
* **Local Ocorrido (LOCAL OCORRIDO)**: O local pode dar informações sobre ambientes de maior risco.
* **Assistência Médica (ASSISTMED)**: Indica se a pessoa recebeu assistência, o que pode ser um preditor de prevenção de suicídios.
* **Causa Básica (CAUSABAS, CAUSABAS\_O)**: Pode ajudar a entender se fatores de saúde pré-existentes estão correlacionados com suicídios.

**Exemplo de Abordagem**

1. **Escolha da Variável Alvo**: Por exemplo, você pode escolher **faixa etária** como a variável alvo para prever.
2. **Pré-processamento**: Transformar variáveis categóricas em variáveis numéricas (por exemplo, usando one-hot encoding) e normalizar as variáveis numéricas.
3. **Treinamento do Modelo**: Utilizar um modelo como KNN, árvore de decisão ou regressão logística para prever a variável alvo escolhida.
4. **Avaliação do Modelo**: Usar métricas apropriadas como precisão, recall e F1-score para avaliar a eficácia do modelo.