

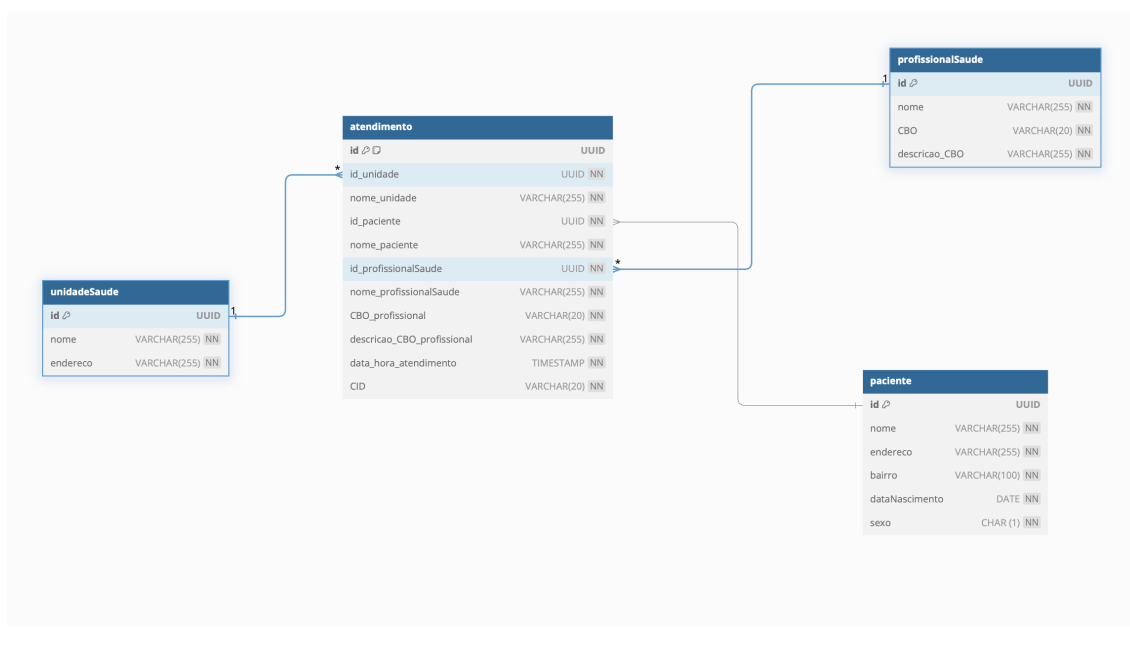
MEMORIAL DESCRITIVO

Diagrama Entidade-Relacionamento (DER)

O Diagrama Entidade-Relacionamento para a base de dados descrita contém quatro entidades principais: **unidadeSaude**, **profissionalSaude**, **paciente**, e **atendimento**, com os seguintes relacionamentos:

1. **unidadeSaude**: Cada unidade de saúde tem um relacionamento com a entidade **atendimento**, onde a chave estrangeira `id_unidade` em **atendimento** referencia a chave primária `id` em **unidadeSaude**.
2. **profissionalSaude**: Relaciona-se com **atendimento** através do `id_profissionalSaude`, onde cada profissional de saúde está vinculado a vários atendimentos.
3. **paciente**: Relaciona-se com **atendimento** através do `id_paciente`, onde cada paciente pode estar associado a múltiplos atendimentos.
4. **atendimento**: A tabela central que registra os atendimentos médicos, com referências às entidades **unidadeSaude**, **profissionalSaude**, e **paciente**.

DER (Visualização Descritiva):



Entidades:

1. **unidadeSaude**:
 - **Descrição**: Tabela que armazena as informações das unidades de saúde onde os atendimentos ocorrem.

- **Atributos:**
 - `id` (UUID, PK): Identificador único da unidade de saúde.
 - `nome` (VARCHAR(255)): Nome da unidade de saúde.
 - `endereco` (VARCHAR(255)): Endereço da unidade de saúde.
- 2. **profissionalSaude:**
 - **Descrição:** Tabela que armazena os dados dos profissionais de saúde, com seus códigos de ocupação (CBO) e uma breve descrição de suas funções.
 - **Atributos:**
 - `id` (UUID, PK): Identificador único do profissional de saúde.
 - `nome` (VARCHAR(255)): Nome do profissional de saúde.
 - `CBO` (VARCHAR(20)): Código Brasileiro de Ocupações (CBO) que define a função do profissional.
 - `descricao_CBO` (VARCHAR(255)): Descrição do cargo baseado no CBO (Ex.: Médico, Enfermeiro, Técnico de Enfermagem).
- 3. **paciente:**
 - **Descrição:** Tabela que armazena os dados dos pacientes atendidos nas unidades de saúde.
 - **Atributos:**
 - `id` (UUID, PK): Identificador único do paciente.
 - `nome` (VARCHAR(255)): Nome do paciente.
 - `endereco` (VARCHAR(255)): Endereço do paciente.
 - `bairro` (VARCHAR(100)): Bairro do paciente.
 - `dataNascimento` (DATE): Data de nascimento do paciente.
 - `sexo` (CHAR(1)): Sexo do paciente (M/F).
- 4. **atendimento:**
 - **Descrição:** Tabela que registra os atendimentos realizados, associando-os aos profissionais de saúde, pacientes e unidades de saúde.
 - **Atributos:**
 - `id` (UUID, PK): Identificador único do atendimento.
 - `id_unidade` (UUID, FK): Referência para a unidade de saúde onde ocorreu o atendimento.
 - `nome_unidade` (VARCHAR(255)): Nome da unidade de saúde onde ocorreu o atendimento.
 - `id_paciente` (UUID, FK): Referência para o paciente atendido.
 - `nome_paciente` (VARCHAR(255)): Nome do paciente atendido.
 - `id_profissionalSaude` (UUID, FK): Referência para o profissional de saúde que realizou o atendimento.
 - `nome_profissionalSaude` (VARCHAR(255)): Nome do profissional que realizou o atendimento.
 - `CBO_profissional` (VARCHAR(20)): Código CBO do profissional de saúde.
 - `data_hora_atendimento` (TIMESTAMP): Data e hora em que o atendimento foi realizado.
 - `CID` (VARCHAR(20)): Classificação Internacional de Doenças (CID) relacionada ao atendimento.

MODELO -- RELATÓRIO DE PREVISÃO DE ATENDIMENTOS POR BAIRRO

1. Introdução

Este relatório apresenta o desenvolvimento de um modelo de previsão de atendimentos médicos por bairro, utilizando dados de atendimentos, pacientes, profissionais de saúde e unidades de saúde. A análise é realizada com foco em previsões de séries temporais, utilizando a plataforma Driverless AI para otimizar os resultados. O objetivo é avaliar a demanda de atendimentos por bairro em diferentes intervalos de tempo (diário, semanal e mensal).

2. Descrição da Base de Dados

A base de dados utilizada contém informações provenientes de várias tabelas:

- **Tabela Atendimento:** Contém os registros de atendimentos médicos, incluindo dados como data, hora, e unidade de atendimento.
- **Tabela Paciente:** Armazena os dados demográficos dos pacientes, como nome, endereço, bairro, e data de nascimento.
- **Tabela Profissional de Saúde:** Contém dados sobre os profissionais de saúde, como nome e cargo.
- **Tabela Unidade de Saúde:** Armazena informações sobre as unidades de saúde onde os atendimentos ocorreram.

Relacionamentos

Os relacionamentos entre as tabelas foram estabelecidos com base no ID do paciente e na unidade de atendimento, permitindo a integração dos dados de atendimento e os dados demográficos para análises mais robustas.

3. Metodologia

3.1 Tratamento de Dados

1. **União das Tabelas:** Os dados das tabelas de atendimentos foram unidos com os dados da tabela de pacientes, utilizando o `id_paciente` como chave, para incluir as informações de bairro no dataset principal.
2. **Utilizei a linguagem python no experimento:**

python

```
df_unificado = pd.merge(df_atendimentos, df_paciente, left_on='id_paciente', right_on='id')
```

3. **Agrupamento dos Dados:** Foram criados três datasets para análise em diferentes granularidades:
 - Diário: Agrupado por dia.
 - Semanal: Agrupado por semanas.
 - Mensal: Agrupado por meses.
4. **Normalização de Outliers:** Nos datasets diários e semanais, os outliers que apareciam apenas uma vez foram substituídos pela mediana, enquanto os dados mensais não passaram por este tratamento devido à amostra reduzida.
5. **Divisão dos Dados:** Cada dataset foi dividido em treino e teste:

- **Diário e Semanal:** 75% treino - 25% teste.
- **Mensal:** 75% treino - 25% teste.

3.2 Modelo Utilizado

O modelo escolhido foi de **Regressão Linear**, com o **RMSE (Root Mean Squared Error)** como métrica principal de avaliação. A plataforma Driverless AI foi configurada com as seguintes prioridades:

- **Acurácia:** 7
- **Tempo:** 5
- **Interpretabilidade:** 7

4. Resultados

4.1 Previsões Diárias

- **Erro médio (RMSE):** O erro encontrado para as previsões diárias foi significativamente maior em alguns bairros. Por exemplo, o bairro **Montenegro** apresentou um erro de x,xx, enquanto o bairro **Cardoso** teve um erro elevado de y,yy, sugerindo grande variabilidade nos dados diários.

4.2 Previsões Semanais

- **Erro médio (RMSE):** As previsões semanais apresentaram melhor consistência. O bairro **Montenegro** teve um erro de z,zz unidades, enquanto o bairro **Catolé** apresentou maior variabilidade, com um RMSE mais elevado, devido à maior dispersão dos dados reais.



Figura 1-exemplo--01

4.3 Previsões Mensais

- **Erro médio (RMSE):** As previsões mensais apresentaram o menor RMSE (0,tt), mas devido ao número reduzido de casos de teste, há uma limitação na generalização dos resultados. O bairro **Silveira** teve a menor diferença de X unidades, enquanto o bairro **Ouro Velho** apresentou o maior erro com x,yy unidades.



Figura 2- Exemplo 2

5. Discussão

Os resultados mostram que as previsões diárias tendem a ser mais imprevisíveis, com maiores variações no erro, o que pode ser explicado pela volatilidade dos dados em períodos curtos. Já as previsões semanais e mensais demonstram maior estabilidade, com destaque para a análise mensal, que apesar do menor RMSE, requer mais dados para uma avaliação confiável.

A escolha do modelo de **Regressão Linear** foi adequada para este cenário inicial, mas poderia ser interessante testar outros modelos, como **ARIMA**, especialmente para capturar padrões de sazonalidade e melhorar a acurácia em séries temporais de períodos mais longos.

6. Conclusão

Este experimento mostrou que as previsões para o número de atendimentos variam de acordo com a granularidade temporal dos dados. As previsões mensais se mostraram mais precisas, mas carecem de maior volume de dados para validação. No futuro, seria interessante testar modelos mais complexos e realizar experimentos com uma base de dados expandida, além de refinar o tratamento de outliers para obter resultados mais robustos.

7. Próximos Passos

1. Testar modelos adicionais, criar agrupamentos para aplicar ARIMA.
2. Aumentar o período de coleta de dados para capturar sazonalidades.
3. Refinar o tratamento de outliers e melhorar a qualidade das previsões diárias.