

Apresentação do Autor

Meu nome é **João Vitor Ferreira Freitas**, sou estudante de **Análise de Dados** e apaixonado por transformar informações em **insights estratégicos**. Tenho experiência com **Google Sheets, SQL, Python e Looker Studio**, aplicando técnicas de **análise exploratória, visualização de dados e machine learning** para resolver problemas reais.

Este projeto foi desenvolvido como parte do curso de **Analista de Dados**, com foco em um desafio prático: **a superlotação de leitos hospitalares** no Brasil.

O objetivo foi demonstrar como **dados públicos podem ser analisados e visualizados** de forma clara, trazendo valor para a gestão em saúde e apoiando decisões estratégicas.

Minha missão como futuro Analista de Dados é unir **tecnologia e análise** para criar soluções inteligentes que ajudem pessoas, empresas e instituições a tomarem melhores decisões.

Contatos:

- E-mail: freitasfj1999@gmail.com
- LinkedIn: <https://www.linkedin.com/in/jo%C3%A3o-vitor-ferreira-freitas-37837a20b/>
- Telefone: (35) 99997-1999

Superlotação de Leitos Hospitalares

I - Descrição Detalhada do Problema

A superlotação hospitalar é um fenômeno recorrente e preocupante no sistema de saúde brasileiro. Esse cenário ocorre quando a demanda por leitos, especialmente em unidades de terapia intensiva (UTIs), ultrapassa a capacidade instalada disponível nos hospitais públicos e privados. Em situações de crise sanitária, como durante a pandemia da COVID-19, epidemias de H1N1 ou surtos de dengue, a pressão sobre o sistema de saúde se intensifica, resultando em falta de leitos para pacientes em estado grave.

As consequências da superlotação são diversas: aumento das taxas de mortalidade, demora no atendimento a casos críticos, transferência inadequada de pacientes e sobrecarga física e emocional de profissionais da saúde. Além disso, a má distribuição geográfica dos leitos contribui para acentuar desigualdades regionais, deixando populações vulneráveis em maior risco.

II - Importância e Relevância do Problema

O impacto da superlotação hospitalar transcende a área da saúde, atingindo dimensões sociais, econômicas e de gestão pública.

- **Social:** o acesso a um leito pode ser determinante para a sobrevivência de pacientes em condições críticas. A superlotação representa risco direto à vida e compromete o direito fundamental à saúde.
- **Econômica:** a pressão sobre hospitais gera custos adicionais com internações prolongadas, uso de insumos e realocação emergencial de recursos.
- **Gestão Pública:** a superlotação dificulta o planejamento estratégico dos sistemas de saúde, prejudicando a tomada de decisão e a distribuição adequada de recursos financeiros e humanos.

Assim, trata-se de um problema de alta relevância nacional, que exige soluções baseadas em evidências e sustentadas por análises de dados confiáveis.

III - Contribuição da Análise de Dados para a Solução

A análise de dados e o uso de modelos de machine learning representam alternativas estratégicas para mitigar os efeitos da superlotação hospitalar. A partir da coleta e tratamento de informações de bases públicas, é possível:

- **Prever a demanda por leitos** em diferentes períodos, considerando fatores sazonais (doenças respiratórias no inverno, por exemplo).
- **Identificar padrões de internação**, como tempo médio de permanência, principais causas de hospitalização e perfis demográficos mais afetados.
- **Apoiar a alocação eficiente de recursos**, orientando gestores na distribuição de leitos, profissionais e insumos em períodos de crise.
- **Criar dashboards interativos** que possibilitem o monitoramento em tempo real da ocupação hospitalar, facilitando a tomada de decisão.

Dessa forma, a ciência de dados se apresenta como uma ferramenta essencial para reduzir os impactos da superlotação hospitalar, promovendo maior eficiência no sistema de saúde e, principalmente, salvando vidas.

Levantamento das Bases de Dados e Plano de Coleta

I - Objetivo da Etapa

Definir quais bases de dados públicas serão utilizadas para analisar o problema da superlotação hospitalar, quais variáveis são relevantes e como essas informações poderão ser coletadas e integradas para a análise.

II - Fontes de Dados Públicas

DATASUS – SIH/SUS (Sistema de Informações Hospitalares)

- **O que contém:** Registros de internações hospitalares financiadas pelo SUS.
- **Variáveis relevantes:**
 - Quantidade de internações por município/hospital.
 - Causas da internação (CID-10).
 - Tempo médio de permanência.
 - Taxa de ocupação hospitalar.
- **Uso no projeto:** Identificar padrões de internação e medir a pressão sobre o sistema de saúde em diferentes regiões.

DATASUS – CNES (Cadastro Nacional de Estabelecimentos de Saúde)

- **O que contém:** Informações sobre a infraestrutura hospitalar no Brasil.
- **Variáveis relevantes:**
 - Número de leitos clínicos e de UTI por hospital.
 - Localização dos hospitais.
 - Serviços oferecidos.
- **Uso no projeto:** Calcular a disponibilidade de leitos por região e comparar com a demanda.

SIVEP-Gripe (Sistema de Vigilância Epidemiológica da Gripe)

- **O que contém:** Dados de casos graves de síndrome respiratória aguda (SRAG).
- **Variáveis relevantes:**
 - Número de casos notificados por semana epidemiológica.
 - Evolução dos casos (alta, óbito, internação prolongada).
 - Faixa etária e comorbidades.

- **Uso no projeto:** Correlacionar aumento de casos gripais com ocupação hospitalar e prever sazonalidade.

IBGE – Censos e PNAD Contínua

- **O que contém:** Informações demográficas e socioeconômicas
- **Variáveis relevantes:**
 - População por faixa etária e município.
 - Indicadores sociais (renda, escolaridade).
- **Uso no projeto:** Normalizar dados hospitalares por população e identificar perfis de maior risco.

Open DataSUS (Plataforma Centralizada)

- **Link:** <https://opendatasus.saude.gov.br/>
- **Uso no projeto:** Repositório oficial de dados abertos do SUS, que consolida as bases mencionadas acima.

III - Estratégia de Coleta e Integração

1. **Coleta:** Download dos conjuntos de dados em formato CSV/Parquet disponíveis no Open DataSUS.
2. **Tratamento:**

- Padronização de colunas (datas, códigos de doenças, unidades hospitalares).
- Limpeza de inconsistências (valores nulos, duplicados).

3. Integração:

- Unir dados do **CNES** (capacidade) com **SIH/SUS** (demandas).
- Relacionar **SIVEP-Gripe** (doenças sazonais) com períodos de maior ocupação.
- Ajustar as análises conforme o tamanho populacional do município (IBGE).

4. Análise Exploratória (próxima etapa):

- Taxa de ocupação = Internações / Leitos disponíveis.
- Tendência temporal de ocupação hospitalar.
- Regiões mais críticas e sazonalidades.

Análise Exploratória dos Dados (EDA)

I - Objetivo da EDA

A análise exploratória de dados tem como finalidade:

- **Compreender a estrutura dos dados coletados** (dimensões, variáveis e consistência).
- **Identificar padrões e sazonalidades** na ocupação hospitalar.
- **Revelar fatores críticos** que contribuem para a superlotação.
- **Preparar insights visuais** que orientarão as próximas etapas (modelagem preditiva e dashboards).

II - Métricas Relevantes

Capacidade x Demanda

- **Taxa de Ocupação de Leitos:**

$$\text{Taxa de Ocupação} = \frac{\text{Número de Internações}}{\text{Número de Leitos Disponíveis}} \times 100$$

- Identificar hospitais/regiões que operam constantemente acima de 80% (considerado crítico pela OMS).

Padrões de Internação

- **Tempo médio de permanência (dias).**
- **Principais causas de internação (CID-10).**
- **Distribuição por faixa etária e sexo.**

Sazonalidade

- Evolução temporal de internações hospitalares.
- Correlação com surtos de doenças respiratórias (dados do **SIVEP-Gripe**).
- Comparação entre meses do ano (ex.: aumento no inverno).

Indicadores Regionais

- **Leitos disponíveis por 100 mil habitantes (ajustado com dados do IBGE).**
- **Diferença entre capitais e interior.**
- **Ranking das regiões mais críticas.**

III - Visualizações Propostas

- 1. Linha do tempo (série temporal):**
 - Ocupação hospitalar mês a mês.
 - Comparação entre regiões/estados.
- 2. Mapa de calor (heatmap):**
 - Ocupação por município.
 - Identificação de “hotspots” de superlotação.
- 3. Gráfico de barras:**
 - Principais causas de internação (CID-10).

- Distribuição de leitos por tipo (clínicos x UTI).

4. **Boxplot:**

- Distribuição do tempo de permanência dos pacientes.

5. **Dashboard no Looker Studio:**

- Indicadores principais (taxa de ocupação, tempo médio de internação, leitos disponíveis por 100 mil habitantes).
- Filtros por estado, hospital, faixa etária.

IV - Insights Esperados

- Identificação de regiões que enfrentam **crise crônica de leitos**.
- Sazonalidade clara em internações relacionadas a **doenças respiratórias**.
- Evidência de desigualdade regional na distribuição de leitos.
- Base para criação de **modelos preditivos** de demanda hospitalar.

Visualização dos Resultados no Looker Studio

I - Objetivo

Construir um **dashboard interativo** que permita:

- Monitorar a taxa de ocupação hospitalar em tempo real ou histórico.
- Comparar regiões, hospitais e períodos.
- Facilitar a tomada de decisão dos gestores de saúde.

II - Indicadores-Chave (KPIs)

1. **Taxa média de ocupação de leitos (%).**
 - Visão geral do sistema.
 - Alerta quando passar de 80% (nível crítico).

2. **Número de leitos disponíveis vs. leitos ocupados.**
 - Segmentação por tipo: Clínicos x UTI.
3. **Tempo médio de internação (dias).**
 - Comparação entre regiões e doenças.
4. **Top 5 causas de internação (CID-10).**
 - Foco nas doenças que mais pressionam o sistema.
5. **Leitos disponíveis por 100 mil habitantes.**
 - Ajustado com dados do IBGE.

III - Estrutura do Dashboard

Página 1 – Visão Geral

- **Scorecards (cartões):**
 - Taxa média de ocupação.
 - Leitos disponíveis totais.
 - Tempo médio de permanência.
 - **Gráfico de linhas:** Ocupação hospitalar ao longo do tempo.
-

Página 2 – Comparação Regional

- **Mapa geográfico:** Ocupação por estado/município (heatmap).
 - **Tabela dinâmica:** Ranking dos estados mais críticos (ocupação > 90%).
 - **Filtro interativo:** Seletor de estado/município.
-

Página 3 – Perfil das Internações

- **Gráfico de barras:** Top 5 causas de internação (CID-10).
 - **Boxplot:** Tempo de permanência por faixa etária.
 - **Segmentação por gênero e idade.**
-

Página 4 – Projeções (Machine Learning)

- **Gráfico de linhas com previsão:** Demanda futura por leitos (ex.: próximos 3 meses).
- **Indicador de alerta:** Projeções acima da capacidade instalada.

IV Recursos do Looker Studio a Utilizar

- **Filtros interativos** (por período, estado, hospital, faixa etária).
- **Campos calculados** para criar indicadores como “Taxa de Ocupação” e “Leitos por 100 mil habitantes”.
- **Cores condicionais** (verde = normal, amarelo = atenção, vermelho = crítico).
- **Tooltip/ajuda** explicando fórmulas e interpretações.

V - Benefícios da Visualização

- **Gestores de saúde** podem identificar regiões críticas rapidamente.
- **Tomada de decisão baseada em evidências.**
- **Capacidade de antecipar crises** de superlotação com base em tendências.
- **Material de impacto** para apresentar no portfólio profissional.

Conclusão e Insights Finais

A análise realizada sobre a **ocupação de leitos hospitalares** no Brasil permitiu identificar padrões importantes que reforçam a necessidade de políticas públicas mais eficientes na gestão da saúde.

Estados mais críticos

- Estados como **Bahia, Rio de Janeiro e São Paulo** apresentam **taxas de ocupação mais elevadas** e maior risco de superlotação.
- Em alguns casos, o **tempo médio de permanência acima de 30 dias** agrava a baixa rotatividade dos leitos.

Principais doenças que pressionam leitos

- **Doenças respiratórias** (pneumonia e asma).
 - **Doenças crônicas** (diabetes tipo 2 e insuficiência cardíaca).
 - **Doenças infecciosas agudas** (como diarreias).
- Essas condições concentram a maior parte das internações, com forte impacto em períodos sazonais.

Períodos de maior risco (sazonalidade)

- Meses de **inverno** (junho a agosto) apresentam aumento significativo nas internações por doenças respiratórias.

- Períodos de **epidemias sazonais** (como surtos de influenza) elevam ainda mais a taxa de ocupação.

Recomendações de gestão

- **Redistribuição de recursos:** alocar leitos de forma dinâmica, priorizando estados mais críticos.
- **Abertura de leitos temporários** em períodos sazonais de maior risco.
- **Monitoramento contínuo via dashboards**, permitindo decisões mais rápidas e baseadas em evidências.
- **Investimento em prevenção:** campanhas de vacinação e controle de doenças crônicas podem reduzir a pressão sobre os hospitais.

✓ Assim, concluímos que o **uso de dados públicos aliado a ferramentas de análise e visualização** é fundamental para compreender a realidade da saúde no Brasil e propor soluções que aumentem a eficiência e reduzam o risco de superlotação hospitalar.

Autorização LGPD

Eu, João Vitor Ferreira Freitas, portador da Cédula de Identidade RG nº MG-24.391.362, inscrito no CPF sob o nº 017.584.666-93, autorizo a cessão do meu projeto em favor da Semantix, bem como a divulgação do meu nome como autor responsável pelo projeto, uma vez que será possível incluir esse trabalho em meu portfólio de trabalho. Nesse sentido, autorizo também a divulgação dos meus contatos (telefone: (35) 99997-1999 e e-mail: freitasfj1999@gmail.com) para a Semantix, tão somente para uso interno com finalidade única de contato em decorrência da elaboração do projeto mencionado.