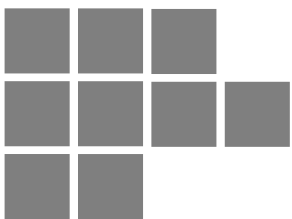




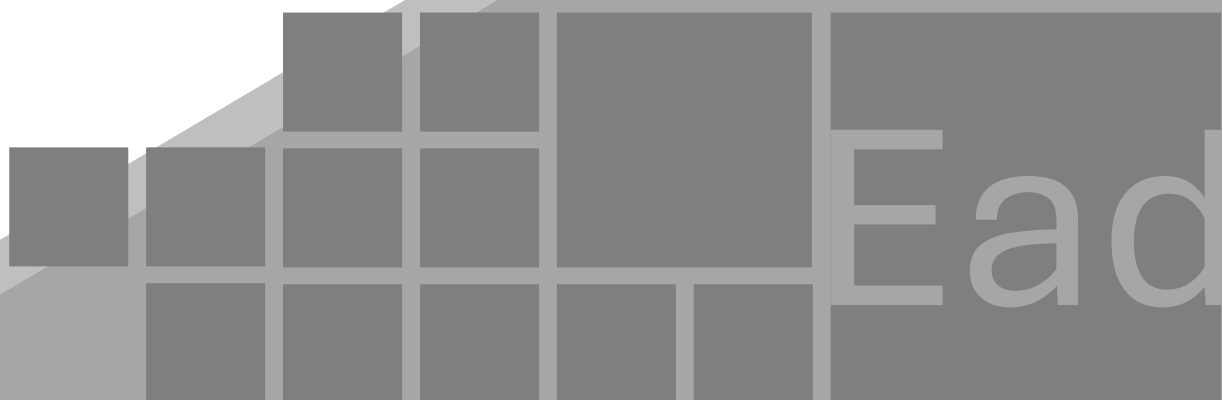
Universidade Presbiteriana

Mackenzie

CIÊNCIA DE DADOS



Projeto **Aplicado** 01

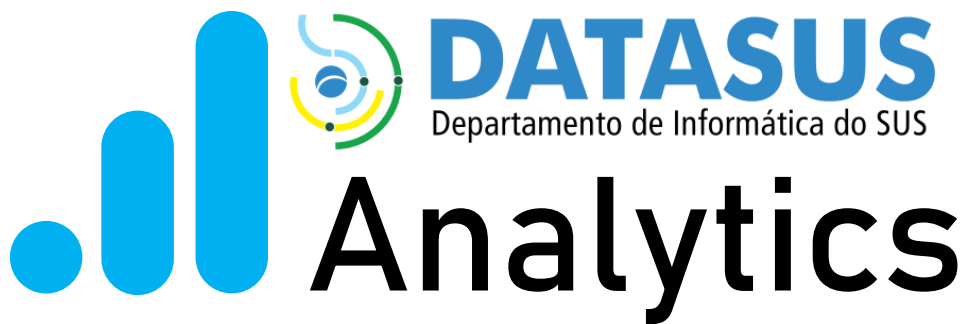
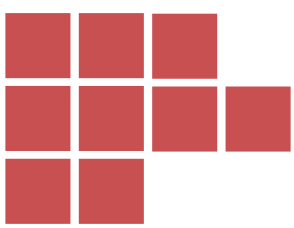


Curso Ciência de Dados

Universidade Presbiteriana Mackenzie

2º semestre – ano letivo: 2025

CIÊNCIA DE DADOS



Este Projeto tem como objetivo principal propor uma análise crítica através de um estudo prático do banco de dados do DataSUS. O estudo é baseado no uso de técnicas e ferramentas desenvolvidas no curso de Ciência de Dados da Universidade Presbiteriana Mackenzie com a orientação do professor Lucas Cerqueira Figueiredo.

Sumário

Introdução:.....	4
Capítulo 1 - Contexto do Estudo:.....	5
Capítulo 2 - Referências de Aquisição do Dataset:.....	6
Capítulo 2.1 - Origem dos Dados:	7
Capítulo 2.2 - Limitações de Uso:	7
Capítulo 2.3 - Período da Coleta:	7
Capítulo 2.4 - Descrição da Origem:.....	8
Capítulo 2.5 - Datasets em vigor:	8
Capítulo 3 - Considerações Finais:	9
Apêndice:	10
Link Repositório ONLINE:	12
Lista de Imagens:	13
Referências Bibliográficas:.....	13



Introdução:

O Sistema Único de Saúde (SUS) configura-se como a maior rede pública de saúde do Brasil, garantindo acesso universal e gratuito a serviços de saúde em todos os níveis de atenção, desde a atenção primária até procedimentos de alta complexidade. A escolha do SUS como objeto de estudo se justifica pela vasta disponibilidade de dados públicos e confiáveis, acessíveis através do DataSUS. Esta plataforma governamental centraliza e disponibiliza informações cruciais sobre a saúde da população brasileira, incluindo dados epidemiológicos, assistenciais, demográficos e socioeconômicos. A riqueza e a granularidade desses dados oferecem um potencial significativo para a realização de análises complexas e aprofundadas, permitindo a identificação de padrões, tendências e relações que podem subsidiar a formulação de políticas públicas mais eficazes, a otimização da gestão dos serviços de saúde e a melhor compreensão dos determinantes sociais da saúde no país. Além disso, a transparência e a acessibilidade dos dados do DataSUS fomentam a pesquisa científica independente e a participação da sociedade civil no acompanhamento e na avaliação do sistema de saúde.

Capítulo 1 - Contexto do Estudo:

O pré-natal configura-se como um pilar fundamental na assistência à saúde materno-infantil, desempenhando um papel crucial na detecção precoce e no manejo adequado de potenciais complicações gestacionais, como a prematuridade, que podem impactar significativamente tanto a saúde da mãe quanto a do recém-nascido.

Apesar de sua reconhecida importância, a cobertura e a qualidade do acompanhamento pré-natal no Brasil apresentam heterogeneidades significativas, frequentemente associadas às disparidades socioeconômicas e à infraestrutura de saúde disponível nas diferentes regiões do país. Áreas com recursos limitados e menor acesso a serviços de saúde tendem a apresentar índices inferiores de adesão ao pré-natal, o que, por sua vez, pode contribuir para elevadas taxas de mortalidade materna e fetal, evidenciando a urgência de investigar e compreender os fatores que influenciam o acesso e a efetividade desse cuidado essencial.



Figura 1 - Pre-Natal

Capítulo 2 - Referências de Aquisição do Dataset:

Os dados utilizados neste estudo serão provenientes de bases de dados públicas e confiáveis, geridas pelo Ministério da Saúde através do Departamento de Informática do SUS (DataSUS) e pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

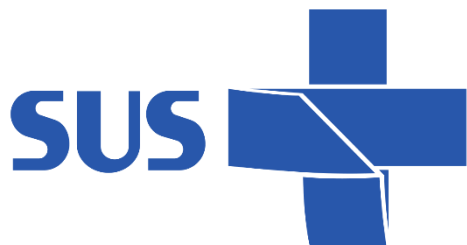


Figura 2 - SUS



Figura 3 - IBGE

Capítulo 2.1 - Origem dos Dados:

Sistema de Informações sobre Nascidos Vivos (SINASC) e Sistema de Informações sobre Mortalidade (SIM): Geridos pelo DataSUS, órgão responsável pela coleta, processamento e disseminação de informações sobre saúde no Brasil. Os dados são originados de Declarações de Nascido Vivo (DNV) e Declarações de Óbito (DO) preenchidas em todo o território nacional por profissionais de saúde em hospitais, maternidades e outros estabelecimentos de saúde, além de registros de óbitos em cartórios.

Infraestrutura Hospitalar: Dados provenientes de levantamentos e pesquisas do IBGE, como o Perfil dos Municípios Brasileiros (Munic) e outras pesquisas relacionadas à saúde e infraestrutura. O IBGE é o principal produtor e disseminador de informações estatísticas e geográficas do Brasil.

Capítulo 2.2 - Limitações de Uso:

Os dados do DataSUS e do IBGE são públicos e de acesso gratuito para fins de pesquisa e estudo. No entanto, é fundamental observar as notas e a documentação de cada base de dados para compreender as especificidades da coleta, as possíveis limitações de cobertura e qualidade, e as recomendações para a interpretação dos resultados. É importante citar adequadamente a fonte dos dados em qualquer publicação ou divulgação dos resultados.

Capítulo 2.3 - Período da Coleta:

A coleta dos dados abrangerá um período específico a ser definido, buscando a maior abrangência temporal possível para identificar tendências e variações relevantes. A seleção do período considerará a disponibilidade e a comparabilidade dos dados nas diferentes bases, visando uma análise robusta e significativa.

Capítulo 2.4 - Descrição da Origem:

DataSUS (SINASC e SIM): O DataSUS é uma organização governamental vinculada ao Ministério da Saúde, responsável por produzir e disseminar informações essenciais para a formulação, acompanhamento e avaliação de políticas de saúde no Brasil. Os dados do SINASC e SIM são gerados no contexto do registro administrativo de eventos vitais (nascimentos e óbitos) em todo o país, sendo cruciais para a vigilância epidemiológica e a gestão da saúde.

IBGE (Infraestrutura Hospitalar): O IBGE é uma fundação pública federal responsável por produzir informações estatísticas, geográficas e ambientais do Brasil. Os dados sobre infraestrutura hospitalar são gerados através de censos, pesquisas amostrais e levantamentos específicos, visando fornecer um panorama detalhado das características e da distribuição dos recursos de saúde no território nacional.

Capítulo 2.5 - Datasets em vigor:

Sistema de Informações sobre Nascidos Vivos (SINASC): Contém informações detalhadas sobre todos os nascimentos ocorridos no Brasil, incluindo características do recém-nascido (peso ao nascer, idade gestacional), da mãe (idade, escolaridade, local de residência) e do pré-natal (número de consultas realizadas). A proposta deste dataset é fornecer dados para o monitoramento da natalidade, das características dos nascidos vivos e da assistência pré-natal.

Sistema de Informações sobre Mortalidade (SIM): Registra informações sobre todos os óbitos ocorridos no Brasil, incluindo dados sobre o falecido (idade, sexo, causa da morte codificada pela CID-10), o local da ocorrência e informações sobre a mãe em casos de óbito fetal ou materno. A proposta é fornecer dados para a vigilância da mortalidade, a identificação de causas de morte e a avaliação de políticas de saúde.

Infraestrutura Hospitalar (IBGE): Apresenta dados sobre o número de hospitais e maternidades por município, a quantidade de leitos disponíveis, a

distribuição de profissionais de saúde e outras características da infraestrutura de saúde. A proposta é fornecer informações sobre a disponibilidade e a distribuição dos recursos de saúde no país.

Esses datasets registram fenômenos relacionados à natalidade, mortalidade e à organização do sistema de saúde, permitindo analisar a relação entre a realização do pré-natal, a incidência de complicações na gestação (inferida a partir dos dados de nascimento e óbito) e a disponibilidade e características da infraestrutura hospitalar.

Capítulo 3 - Considerações Finais:

Ao longo deste semestre, o aprofundamento nos conhecimentos do curso de ciência de dados proporcionará as ferramentas necessárias para abordar o tema proposto. Com a justificativa da escolha do DataSUS como fonte de dados robusta e a contextualização da importância do pré-natal, o estudo possui uma base sólida de dados para o seu desenvolvimento.

Todo o processo de coleta, tratamento, análise e visualização dos dados, utilizando as bases e os indicadores públicos, será meticulosamente documentado e disponibilizado em um repositório no GitHub, garantindo a transparência e a replicabilidade da pesquisa. Espera-se que, ao final deste semestre, este trabalho contribua para uma melhor compreensão das dinâmicas envolvidas, possa oferecer insights valiosos para aprimorar as políticas e práticas de saúde materno-infantil no país e contribuir com a comunidade científica, tanto de dados como da área da saúde.

Apêndice:

Sistema Único de Saúde (SUS): Principal sistema público de saúde do Brasil, que oferece acesso universal e gratuito a serviços de saúde em todos os níveis de complexidade.

DataSUS: Departamento de Informática do SUS, órgão do Ministério da Saúde responsável pela coleta, processamento e disseminação de informações e dados sobre saúde no Brasil.

Dados Epidemiológicos: Informações sobre a ocorrência, distribuição e controle de doenças em populações.

Dados Demográficos: Informações sobre a população, como tamanho, distribuição geográfica, idade, sexo e taxa de natalidade.

Dados Socioeconômicos: Informações sobre as condições sociais e econômicas da população, como renda, escolaridade e acesso a serviços básicos.

Granularidade: Nível de detalhe dos dados. Dados com alta granularidade são mais detalhados e específicos.

Determinantes Sociais da Saúde: Fatores sociais, econômicos, culturais e ambientais que influenciam a saúde das pessoas e das populações.

Pré-natal: Acompanhamento médico e de outros profissionais de saúde durante a gravidez, essencial para a saúde da mãe e do bebê.

Prematuridade: Nascimento do bebê antes de completar 37 semanas de gestação.

Heterogeneidades: Diferenças ou variações significativas.

Mortalidade Materna: Morte de mulheres durante a gravidez, parto ou puerpério (período pós-parto) por causas relacionadas à gestação.

Mortalidade Fetal: Morte do feto antes ou durante o parto.

Dataset: Conjunto de dados estruturado para análise.

Ministério da Saúde: Órgão do governo federal responsável pela formulação e gestão das políticas de saúde no Brasil.

Instituto Brasileiro de Geografia e Estatística (IBGE): Órgão responsável pela produção e disseminação de informações estatísticas, geográficas e ambientais do Brasil.

Sistema de Informações sobre Nascidos Vivos (SINASC): Sistema de informação do DataSUS que coleta dados sobre todos os nascimentos ocorridos no Brasil.

Sistema de Informações sobre Mortalidade (SIM): Sistema de informação do DataSUS que coleta dados sobre todos os óbitos ocorridos no Brasil.

Declarações de Nascido Vivo (DNV): Documento preenchido no momento do nascimento, contendo informações sobre o recém-nascido e a mãe.

Declarações de Óbito (DO): Documento preenchido no momento do óbito, contendo informações sobre o falecido e a causa da morte.

Perfil dos Municípios Brasileiros (Munic): Pesquisa do IBGE que coleta informações sobre as características dos municípios brasileiros, incluindo dados sobre saúde.

Vigilância Epidemiológica: Conjunto de ações que visam o conhecimento e a detecção de mudanças nos fatores determinantes e condicionantes da saúde individual e coletiva, com a finalidade de recomendar e adotar medidas de prevenção e controle de doenças e agravos.

CID-10: Classificação Internacional de Doenças e Problemas Relacionados à Saúde, 10ª revisão, utilizada para codificar as causas de morte.

Software: Software é uma coleção de programas que dizem a um computador como executar tarefas específicas.

Repositório: Um repositório de software é um local de armazenamento de onde pacotes de software podem ser recuperados e instalados em um computador.

GitHub: Plataforma online para armazenar e compartilhar códigos, documentos e outros arquivos de projetos, facilitando a colaboração e a transparência na pesquisa.

Insights: Compreensões ou revelações importantes obtidas a partir da análise dos dados.

Link Repositório ONLINE:

<https://github.com/GrupoMackenzie/ProjetoAplicado01-DataScience-Mackenzie-2025>

Lista de Imagens:

FIGURA 1 – PRE-NATAL: UNIQUE BABY. [Ilustração sobre a importância da consulta pré-natal]. Unique Baby, s.d. Disponível em: <https://unique-baby.com.br/blog/gravidez/consulta-pre-natal/>. Acesso em: 20 mar. 2025.

FIGURA 2 – SUS: Wikipedia. [logotipo SUS]. Wikipedia. Disponível em: https://pt.wikipedia.org/wiki/Sistema_%C3%9Anico_de_Sa%C3%BAde. Acesso em: 20 mar. 2025.

FIGURA 3 – IBGE: Instituto Brasileiro de Geografia e Estatística (IBGE). [logotipo SUS]. Instituto Brasileiro de Geografia e Estatística (IBGE). Disponível em: <https://www.infoescola.com/geografia/instituto-brasileiro-de-geografia-e-estatistica-ibge/>. Acesso em: 20 mar. 2025.

Referências Bibliográficas:

DATA SUS. Site da govbr, 2025. Disponível em: <https://datasus.saude.gov.br/>. Acesso em: 10 de março de 2025.

IBGE. Site da govbr, 2025. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 10 de março de 2025.

Curso de Git e GitHub [20 Horas]. Site do CURSOEMVIDEO, 2025. Disponível em: <https://www.cursoemvideo.com/curso/curso-de-git-e-github/>. Acesso em: 10 de março de 2025.