

DATA SUS

Os dados descritos neste DataCard foram retirados do DataSus (Departamento de Informática do Sistema Único de Saúde), plataforma e sistema de informações de saúde mantido pelo Ministério da Saúde do Brasil. Os dados abrangem os nascimentos ocorridos em todo o território nacional em 2022 e incluem informações sobre as características do parto, dos pais e dos recém-nascidos.

Link do conjunto de dados

<https://s3.sa-east-1.amazonaws.com/ckan.saude.gov.br/SINASC/DNOPEN22.csv>

Autores do cartão de dados

- Luís Carlos M. A. Júnior
- Mariana Aparecida Ferreira
- Marco Antonio J. Ticona

Autoria

Publicadores

Organização publicadora

Departamento de Informática do Sistema Único de Saúde (DataSUS)

Tipos de indústria

- Instituições Médicas
- Intituições acadêmicas

Proprietários de conjunto de dados

Equipe

Nome da equipe: Departamento de Informática do Sistema Único de Saúde

Detalhes do contatos

- **Proprietários do conjunto de dados:** CGIAE/DASNT/SVS
- **Website:** <https://opendatasus.saude.gov.br/dataset/sistema-de-informacao-sobre-nascidos-ivos-sinasc>

Motivações e Intenções

Motivações

Propósitos

- Análise de dados
- Coleta de dados
- Monitoramento

- Pesquisa em Saúde

Domínios de Aplicação

Machine Learning, Classification, Nascidos Vivos, SUS.

Fatores Motivadores

- Armazenamento de grandes volumes de dados de saúde em bancos de dados seguros
- Utilização dos dados por profissionais de saúde, pesquisadores e o público em geral para pesquisas

Aplicações

Uso do conjunto de dados

- Suporte à tomada de decisão
- Padronização dos dados
- Armazenamento de dados
- Disponibilização de Dados

Casos de uso adequados

Caso de uso adequado: Estudo utilizando machine learning da relação dos dados coletados pelo SINASC e o peso do recém-nascido. <https://repositorio.unesp.br/handle/11449/236236>

Caso de uso adequado: Mortalidade infantil entre os anos de 2000 e 2017 em uma cidade do Sul do Brasil: técnicas de mineração de dados. <https://rsdjournal.org/index.php/rsd/article/view/7489>

Espaço de pesquisa e Problemas

O objetivo da utilização da base de dados é criar um modelo de *machine learning* que seja capaz de estimar a probabilidade de um recém-nascido nascer com alguma anomalia congênita, a partir de informações gerais da mãe, pai e outras informações demográficas.

Visão geral do conjunto de dados

Características do conjunto de dados

Categoria	Data
Formato	text/csv
Tipo de dados	Dados Tabulares
Espaço	Bruto
Criado	10/Agosto/2023
Licença	Creative Commons Atribuição
Tamanho do conjunto de dados	848,4 MB
Número de Instâncias	2471519

Nome do Campo	Descrição
DTNASCMAE	Quasi-identificadores
RACACORMAE	Dados sensíveis
SERIESCMAE	Dados sensíveis
IDADEPAI	Quasi-identificadores

Tabela 3: Tabela com as variáveis que são sensíveis/identificadores

Licença e Acesso

Tipo de Licença

CC BY 4.0

Permissões de licença

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

Attribution — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or [technological measures](#) that legally restrict others from doing anything the license permits.

Versionamento e Manutenção

Status do conjunto de dados

Versão	1.0
Última Atualização	10/Agosto/2023
Primeira versão	10/Agosto/2023

Plano de manutenção

Atualização do dataset são publicados no [opendatasus](#)

Dicionário das variáveis do DATASUS:

- **ORIGEM:** Banco de dados origem (1-Oracle, 2-FTP, 3-SEAD)
- **CODETAB:** Código do estabelecimento onde ocorreu o nascimento
- **CODMUNNASC:** Código IBGE do município de nascimento
- **LOCNASC:** Local de nascimento (1-Hospital, 2-Outros estabelecimentos de saúde, 3-Domicílio, 4-Outros, 5-Aldeia indígena)

- **IDADEMAE:** Idade da mãe
- **ESTCIVMAE:** Situação conjugal da mãe (1-Solteira, 2-Casada, 3-Viúva, 4-Separada judicialmente/divorciada, 5-União estável, 9-Ignorada)
- **ESCMAE:** Escolaridade em anos de estudos concluídos da mãe (1-Nenhuma, 2-1 a 3 anos, 3-4 a 7 anos, 4-8 a 11 anos, 5-12 e mais, 9-Ignorado)
- **CODOCUPMAE:** Código de ocupação da mãe conforme tabela do CBO (Código Brasileiro de Ocupações)
- **QTDFILVIVO:** Número de filhos vivos
- **QTDFILMORT:** Número de perdas fetais e abortos.
- **CODMUNRES:** Código IBGE do município de residência
- **GESTACAO:** Semanas de Gestação (1- Menos de 22 semanas, 2- 22 a 27 semanas, 3- 28 a 31 semanas, 4- 32 a 36 semanas, 5- 37 a 41 semanas, 6- 42 semanas e mais, 9- Ignorados)
- **GRAVIDEZ:** Tipo de gravidez (1- Única, 2- Dupla, 3- Tripla ou mais, 9- Ignorado)
- **PARTO:** Tipo de Parto (1- Vaginal, 2- Cesário, 9- Ignorado)
- **CONSULTAS:** Número de consultas de pré-natal (1- Nenhuma, 2- de 1 a 3, 3- de 4 a 6, 4- 7 e mais, 9- Ignorado)
- **DTNASC:** Data de nascimento dd mm aaaa
- **HORANASC:** Horário de nascimento
- **SEXO:** Sexo do recém nascido (1- M - Masculino, 2- F - Feminino, 0 - I - Ignorado)
- **APGAR1:** Apgar no 1º minuto 00 a a10
- **APGAR5:** Apgar no 5º minuto 00 a 10
- **RACACOR:** Tipo de raça e cor do nascido (1- Branca; 2- Preta; 3- Amarela; 4- Parda; 5- Indígena).
- **PESO:** Peso ao nascer em gramas.
- **IDANOMAL:** Anomalia identificada (1- Sim; 2- Não; 9- Ignorado)
- **DTCADASTRO:** Data do cadastro da DN no sistema
- **CODANOMAL:** Código da anomalia (CID-10)
- **NUMEROLOTE:** Número do lote
- **VERSAOSIST:** Versão do sistema
- **DTRECEBIM:** Data do último recebimento do lote, dada pelo Sisnet.
- **DIFDATA:** Diferença entre a data de Nascimento e data do recebimento original da DN ([DTNASC] – [DTRECORIGA])
- **DTRECORIGA:** Cria-se campo DTRECORIGA e copia os valores de DTRECORIG para esse campo. Se DTRECORIGA = Nulo, copia os valores de DTRECEBIM. Se DTRECEBIM = Nulo, copia os valores de DTCADASTRO
- **NATURALMAE:** Se a mãe for estrangeira, constará o código do país de nascimento.
- **CODMUNNATU:** Código do município de naturalidade da mãe
- **CODUFNATU:** Código UF de naturalidade
- **ESCMAE2010:** Escolaridade 2010. (0 – Sem escolaridade; 1 – Fundamental I (1ª a 4ª série); 2 – Fundamental II (5ª a 8ª série); 3 – Médio (antigo 2º Grau); 4 – Superior incompleto; 5 – Superior completo; 9 – Ignorado).

- **SERIESCMAE:** Série escolar da mãe. Valores de 1 a 8.
- **DTNASCMAE:** Data de nascimento da mãe: dd mm aaaa
- **RACACORMAE:** Tipo de raça e cor da mãe: 1– Branca; 2– Preta; 3– Amarela; 4– Parda; 5– Indígena.
- **QTDGESTANT:** Número de gestações anteriores
- **QTDPARTNOR:** Número de partos vaginais
- **QTDPARTCES:** Número de partos cesáreos
- **IDADEPAI:** Idade do pai
- **DTULTMENST:** Data da última menstruação (DUM): dd mm aaaa
- **SEMAGESTAC:** Número de semanas de gestação.
- **TPMETESTIM:** Método utilizado. Valores: 1– Exame físico; 2– Outro método; 9– Ignorado.
- **CONSPRENAT:** Número de consultas pré-natal
- **MESPRENAT:** Mês de gestação em que iniciou o pré-natal
- **TPAPRESENT:** Tipo de apresentação do RN. Valores: 1– Cefálico; 2– Pélvica ou podálica; 3– Transversa; 9– Ignorado.
- **STTRABPART:** Trabalho de parto induzido? Valores: 1– Sim; 2– Não; 3– Não se aplica; 9– Ignorado.
- **STCESPARTO:** Cesárea ocorreu antes do trabalho de parto iniciar? Valores: 1– Sim; 2– Não; 3– Não se aplica; 9– Ignorado.
- **TPNASCASSI:** Nascimento foi assistido por? Valores: 1– Médico; 2– Enfermagem ou Obstetriz; 3– Parteira; 4– Outros; 9– Ignorado
- **TPFUNCRESP:** Tipo de função do responsável pelo preenchimento. Valores: 1– Médico; 2– Enfermeiro; 3– Parteira; 4– Funcionário do cartório; 5– Outros.
- **TPDOCRESP:** Tipo do documento do responsável. Valores: 1-CNES; 2-CRM; 3- COREN; 4-RG; 5- CPF.
- **DTDECLARAC:** Data da declaração: dd mm aaaa
- **ESCMAEAGR1:** Escolaridade 2010 agregada. Valores: 00 – Sem Escolaridade; 01 – Fundamental I Incompleto; 02 – Fundamental I Completo; 03 – Fundamental II Incompleto; 04 – Fundamental II Completo; 05 – Ensino Médio Incompleto; 06 – Ensino Médio Completo; 07 – Superior Incompleto; 08 – Superior Completo; 09 – Ignorado; 10 – Fundamental I Incompleto ou Inespecífico; 11 – Fundamental II Incompleto ou Inespecífico; 12 – Ensino Médio Incompleto ou Inespecífico.
- **STDNEPIDEM:** Status de DO Epidemiológica. Valores: 1 – SIM; 0 – NÃO.
- **STDNNOVA:** Status de DO Nova. Valores: 1 – SIM; 0 – NÃO.
- **CODPAISRES:** Código do país de residência
- **TPROBSON:** Código do Grupo de Robson, gerado pelo sistema
- **PARIDADE:** Define se é a primeira gravidez ou se teve mais de uma. 1 – Multípara; 0- Nulípara.
- **KOTELCHUCK:** índice de Kotelchuck - Avaliação da assistência pré-natal