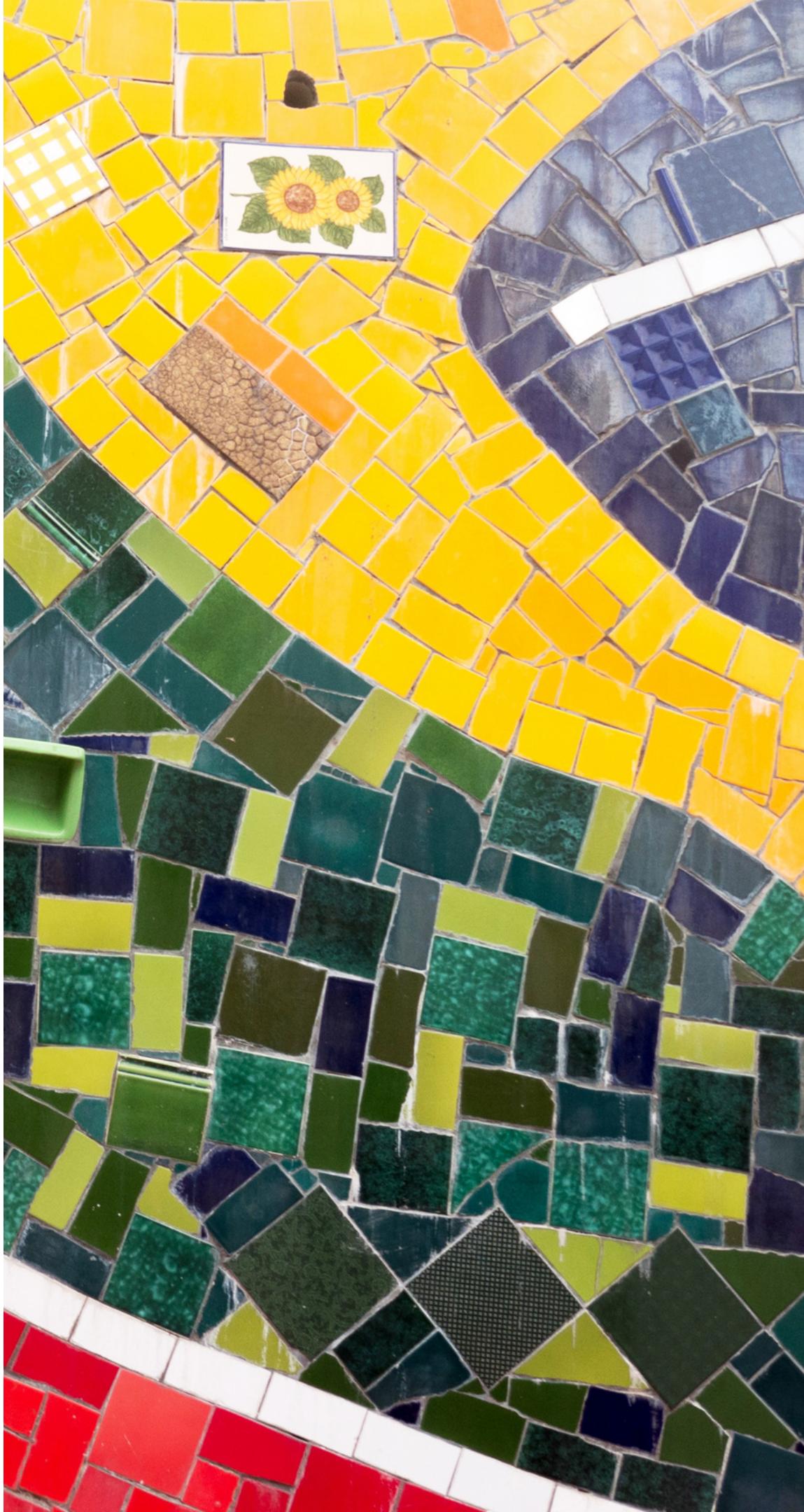


STORYTELLING



PROJETO APLICADO I

ANÁLISE EXPLORATÓRIA DAS RELACÕES ENTRE TAXA DE MORTALIDADE E FATORES SOCIOECONÔMICOS NO BRASIL

MEMBROS DO GRUPO:

Nome: **Adriano Mamoru Takeshita**

TIA: 23022647

E-mail: 10923022647@mackenzista.com.br

Nome: **Camila Vieira**

TIA: 23005432

E-mail: camila.vieira1@mackenzista.com.br / 10923005432@mackenzista.com.br

Nome: **Gabriel Schonenberger de Campos**

TIA: 23011165

E-mail: 10923011165@mackenzista.com.br

Nome: **Gustavo Santiago Zarpelão**

TIA: 23002824

E-mail: 10923002824@mackenzista.com.br

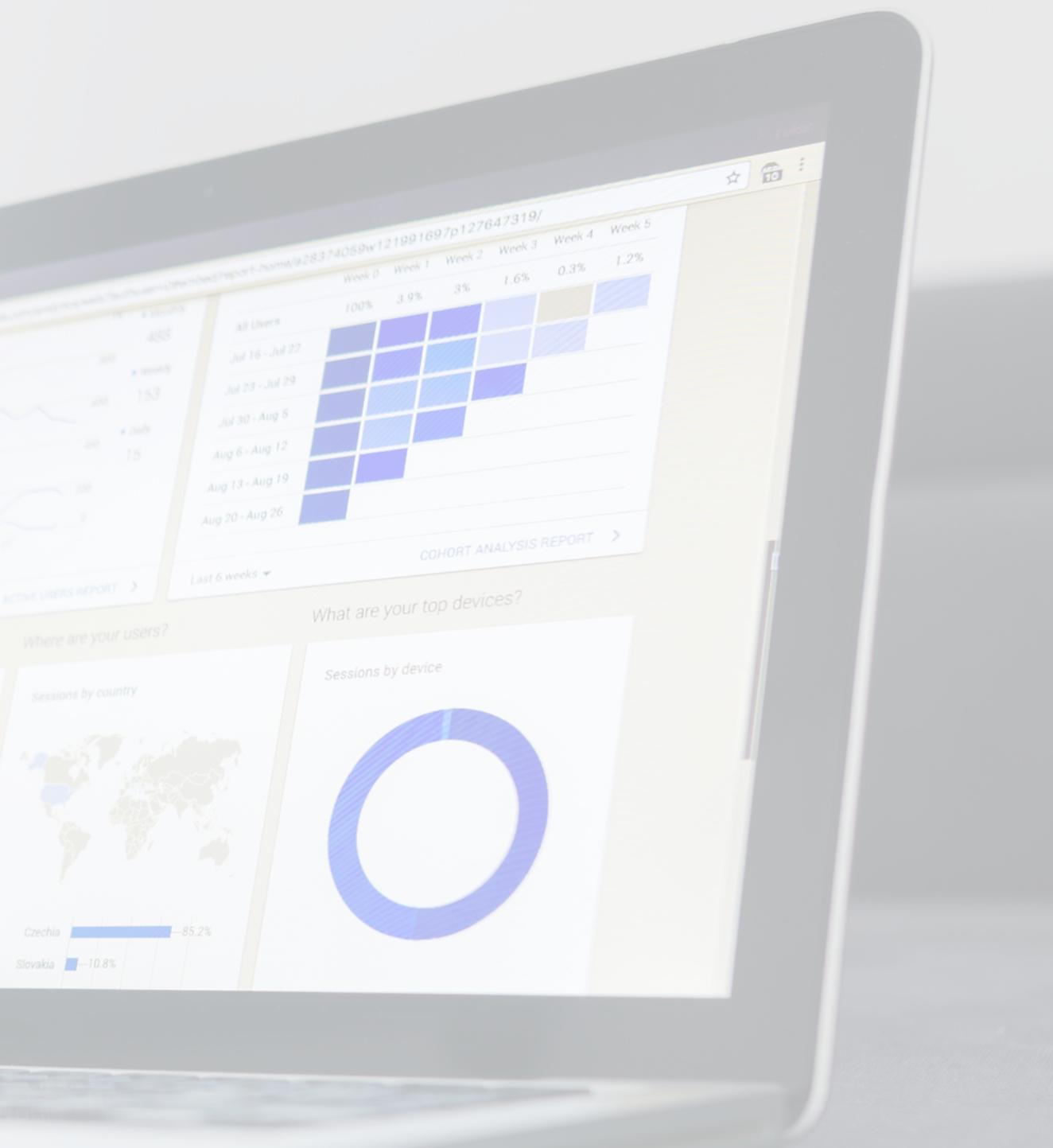
Nome: **Luís Eduardo Alves de Moura da Silva**

TIA: 230094705

E-mail: 10923009470@mackenzista.com.br

O QUE É O PROJETO?

Ao unir a diversidade do Brasil com a qualidade dos dados do IBGE, este projeto busca fornecer insights importantes para o entendimento das relações complexas entre variáveis sociais, econômicas e de saúde, que podem contribuir para o desenvolvimento de estratégias de melhoria e bem-estar da população brasileira.



CONTEXTO



A diversidade populacional, heterogeneidade regional, desigualdades socioeconômicas e desafios em saúde pública criam um **contexto relevante** para a **análise da taxa de mortalidade no Brasil**, revelando a interação entre fatores sociais, econômicos e demográficos.

O **IBGE**, uma instituição estatística renomada, foi usado como fonte de dados e desempenha um papel crucial na coleta e divulgação de dados geográficos, contribuindo para o entendimento da realidade brasileira e o desenvolvimento de políticas públicas.



PREMISSAS



**Fonte de dados
confiável (IBGE)**



**Foco nas taxas
de mortalidade**

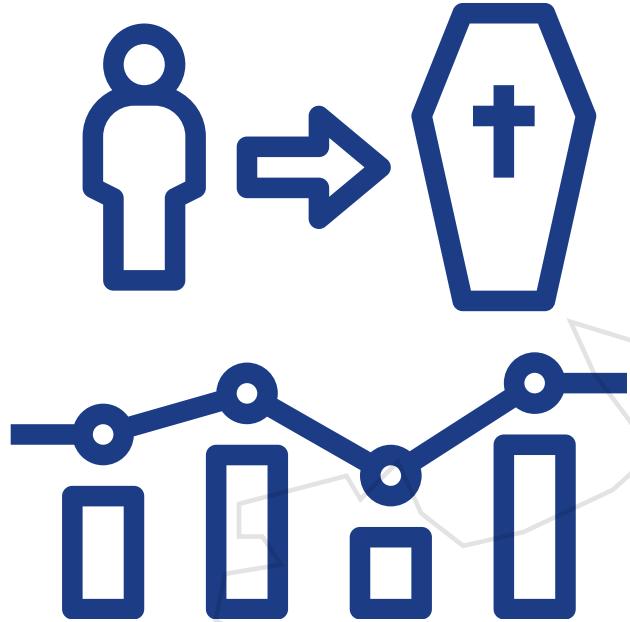


**Pensamento
Computacional**



**Contribuição para
políticas públicas**

OBJETIVOS E METAS



Objetivo principal:

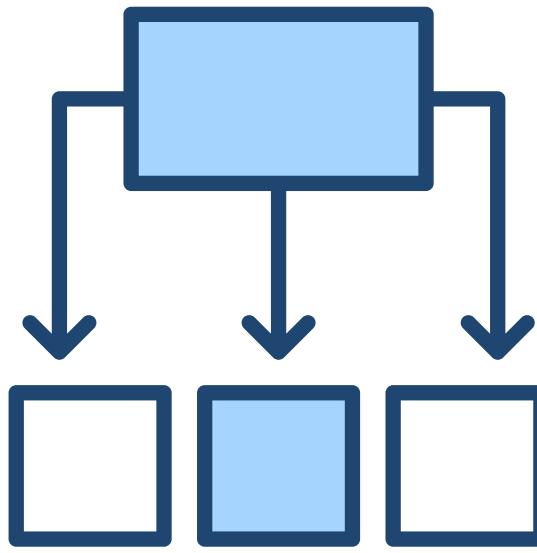
Analizar as relações entre a taxa de mortalidade e diversos fatores sociais, econômicos e demográficos no Brasil.



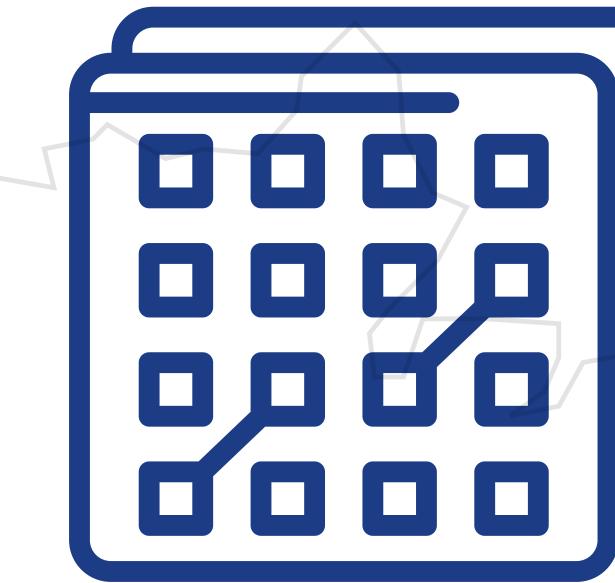
Metas:

Identificar correlações significativas entre a taxa de mortalidade e outras variáveis; compreender como fatores como renda, acesso a serviços básicos, segurança e saúde estão relacionados à taxa de mortalidade.

METODOLOGIA

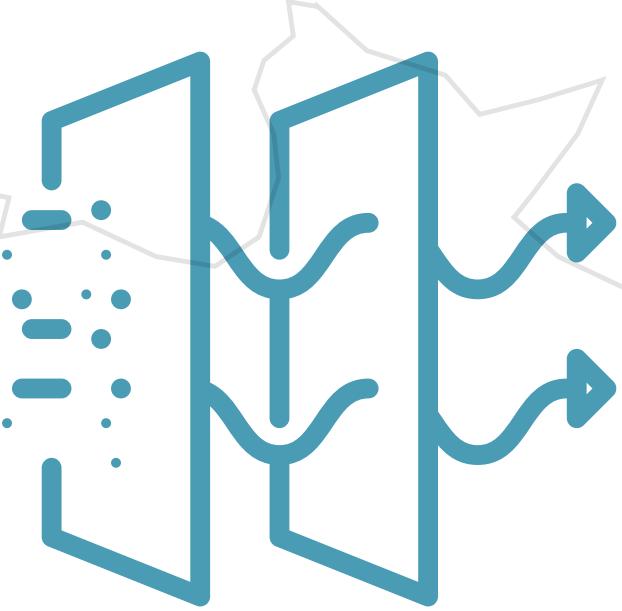


Decomposição:
Iniciamos **dividindo o problema em partes menores** e mais gerenciáveis, considerando variáveis como **renda, educação, acesso a serviços básicos e segurança**.

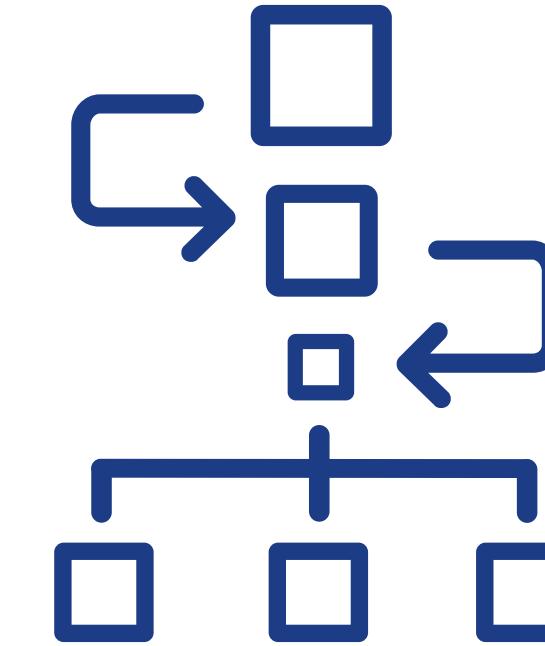


Reconhecimento de Padrões:
Utilizamos técnicas de reconhecimento de padrões para **identificar correlações preliminares entre variáveis, explorando gráficos, tabelas e métricas descritivas**.

METODOLOGIA



Abstração:
Simplificamos a complexidade dos dados, criando indicadores compostos que representam características socioeconômicas mais abrangentes.



Algoritmo:
Focamos em tentar implantar algoritmos de análise exploratória para **quantificar as relações entre as variáveis** e visualizar as tendências.



DESAFIOS

- Despadronização dos Dados
- Dados inconsistentes apresentados de forma
- Lacuna nos dados
- Necessidade de Análise de padrões

DADOS INFOGRÁFICOS



População total brasileira (2020)

211 milhões



Quantidade de pessoas do sexo feminino
107 milhões



Quantidade de pessoas do sexo masculino
104 milhões





Coeficiente de Gini:

O coeficiente de Gini do Brasil está em torno de 53 em 2020, indicando uma alta desigualdade de renda. Quanto mais próximo de 100, maior a desigualdade..

Concentração de Riqueza:

Cerca de 28% da renda total do Brasil está nas mãos dos 1% mais ricos da população.

Pobreza Multidimensional:

Cerca de 25% da população brasileira enfrenta algum tipo de privação em saúde, educação e padrão de vida.



Taxa de Alfabetização (2020):
93%

O que indica que aproximadamente 93%
da população com 15 anos ou mais sabia
ler e escrever.

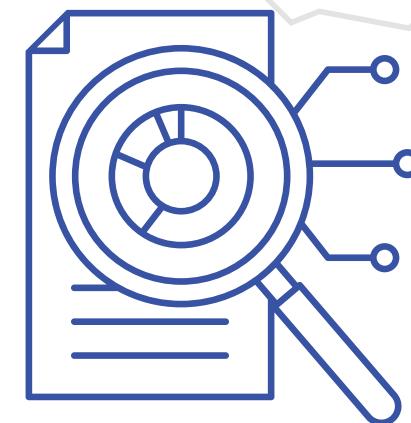
OBJETIVO INICIAL

Nosso principal objetivo foi caracterizar e registrar os datasets que temos à nossa disposição, bem como realizar uma exploração inicial das principais características desses conjuntos de dados.

PRIMEIROS TRATAMENTOS



**Caracterização dos
DataSets / download das
tabelas**



**Identificação dos tipos
de dados**



**Adequação de dados
mesclados e exclusão de
dados
perdidos/incorrectos**

**Cálculo de medidas de
dispersão e posição e
exame de distribuição
dos dados**



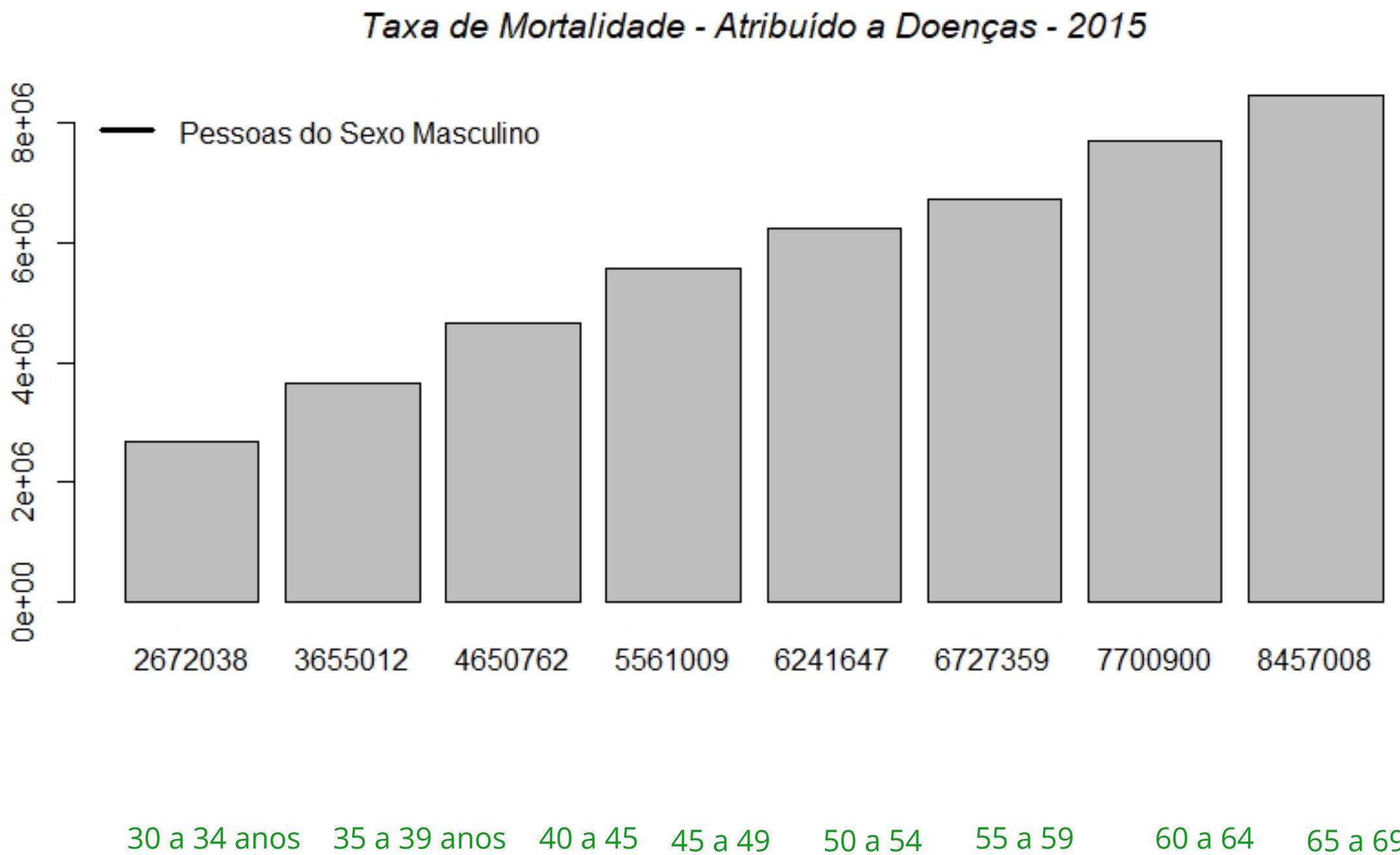
**Compreensão do cenário
dos dados**



Utilizamos o software R para analisar os dados.
Isso incluiu cálculos de **média, mediana, quartil, variância e desvio padrão.**

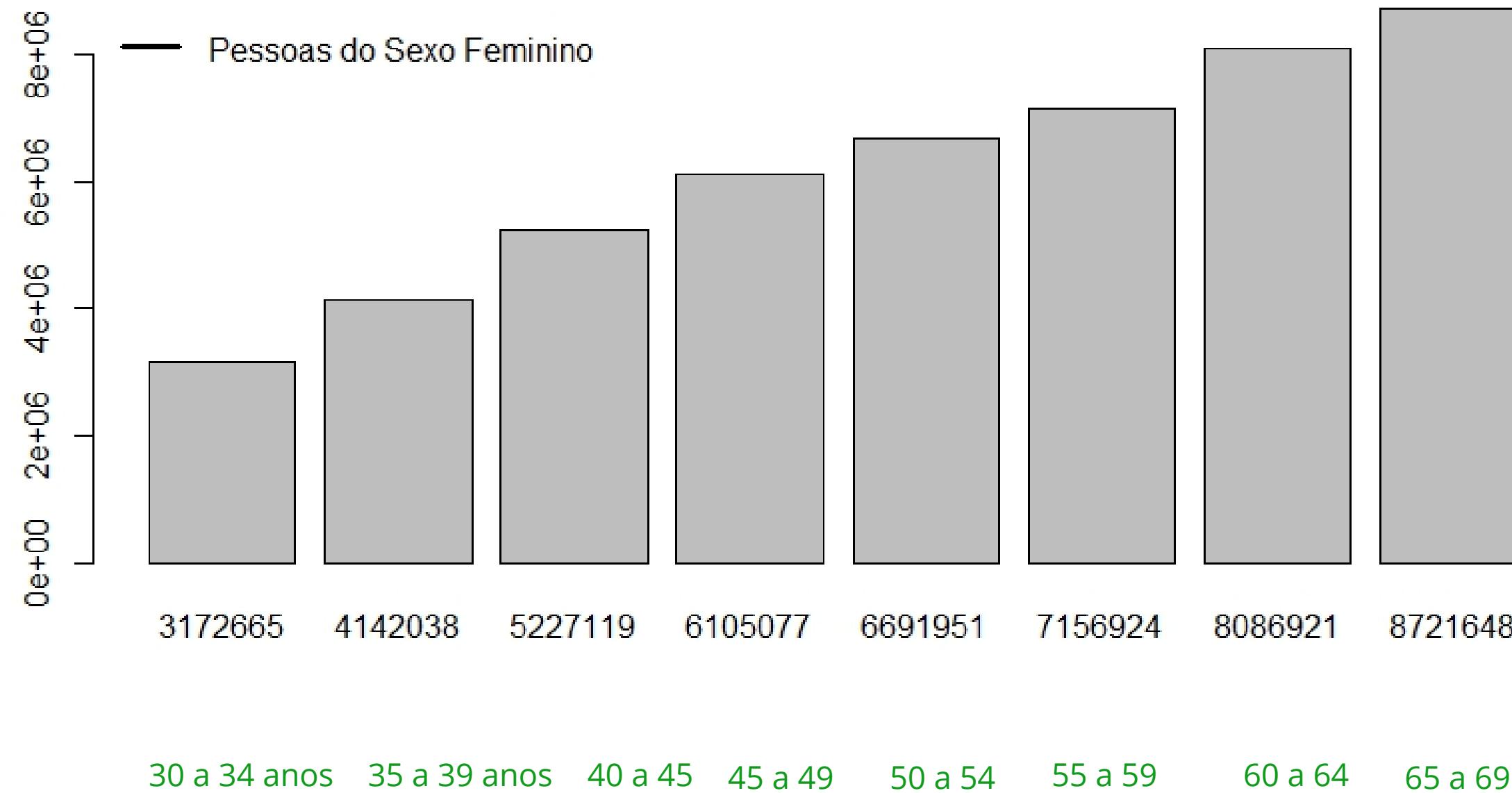
```
> mean(npessoas_doenca$Homens_2015)
[1] 5708217
> mean(npessoas_doenca$Mulheres_2015)
[1] 6163043
> mean(obito_doenca$Homens_2015)
[1] 20484.5
> mean(obito_doenca$Mulheres_2015)
[1] 15869.88
> median(npessoas_doenca$Homens_2015)
[1] 5901328
> median(npessoas_doenca$Mulheres_2015)
[1] 6398514
> median(obito_doenca$Homens_2015)
[1] 17595.5
> median(obito_doenca$Mulheres_2015)
[1] 14187
> quantile(npessoas_doenca$Homens_2015)
  0%   25%   50%   75%   100%
2672038 4401825 5901328 6970744 8457008
> quantile(npessoas_doenca$Mulheres_2015)
  0%   25%   50%   75%   100%
3172665 4955849 6398514 7389423 8721648
> quantile(obito_doenca$Homens_2015)
  0%   25%   50%   75%   100%
3247.00 7216.00 17595.50 32437.75 43673.00
> quantile(obito_doenca$Mulheres_2015)
  0%   25%   50%   75%   100%
2982.00 6778.75 14187.00 23875.00 32882.00
> var(npessoas_doenca$Homens_2015)
[1] 3.917679e+12
> var(npessoas_doenca$Mulheres_2015)
[1] 3.631635e+12
> var(obito_doenca$Homens_2015)
[1] 243342268
> var(obito_doenca$Mulheres_2015)
[1] 122251824
> # usando a Função "sd" para encontrar o Desvio Padrão de maneira individual
> sd(npessoas_doenca$Homens_2015)
[1] 1979313
> sd(npessoas_doenca$Mulheres_2015)
[1] 1905685
> sd(obito_doenca$Homens_2015)
[1] 15599.43
> sd(obito_doenca$Mulheres_2015)
[1] 11056.75
> |
```

RESULTADOS DA ANÁLISE EXPLORATÓRIA



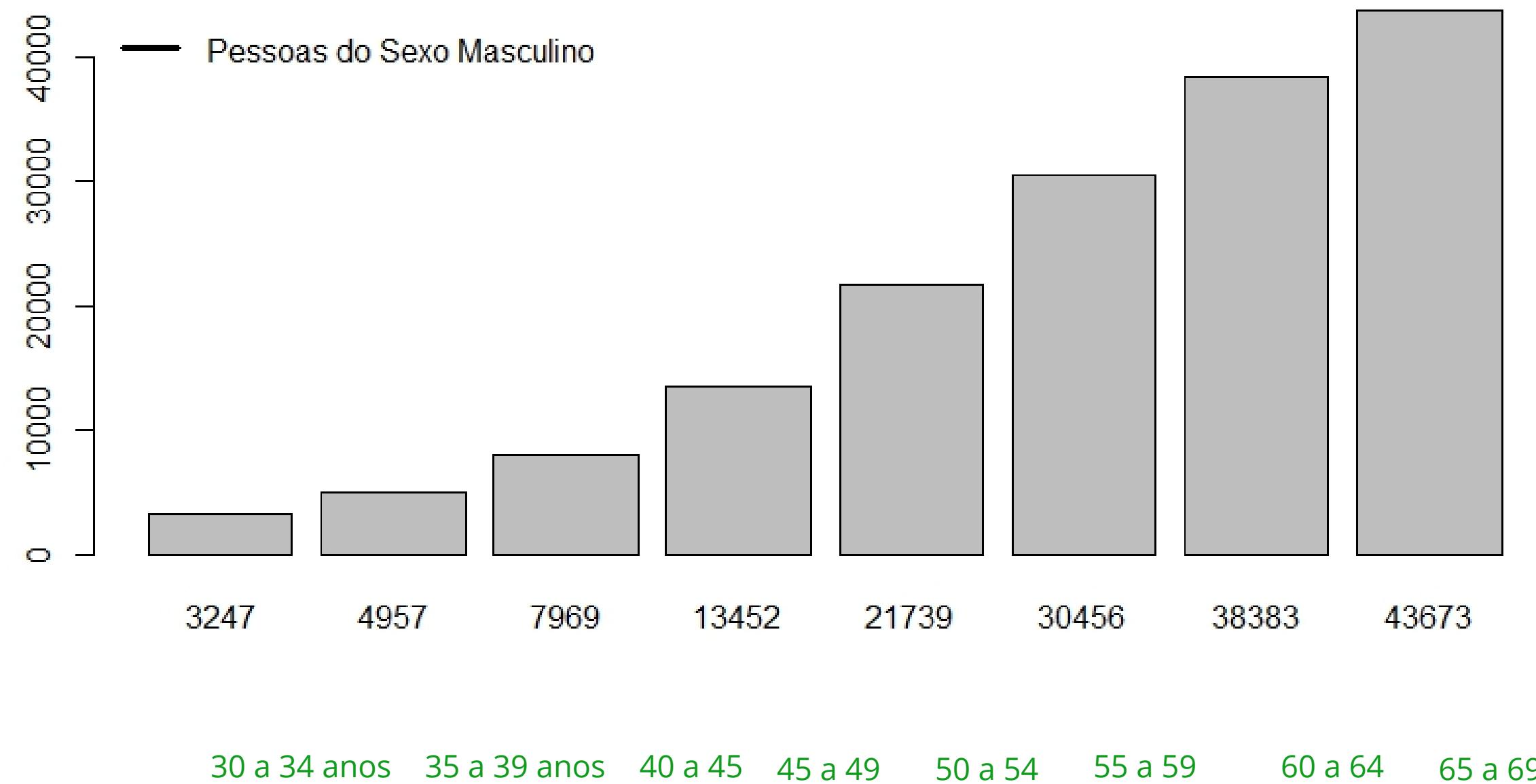
RESULTADOS DA ANÁLISE EXPLORATÓRIA

Taxa de Mortalidade - Atribuído a Doenças - 2015

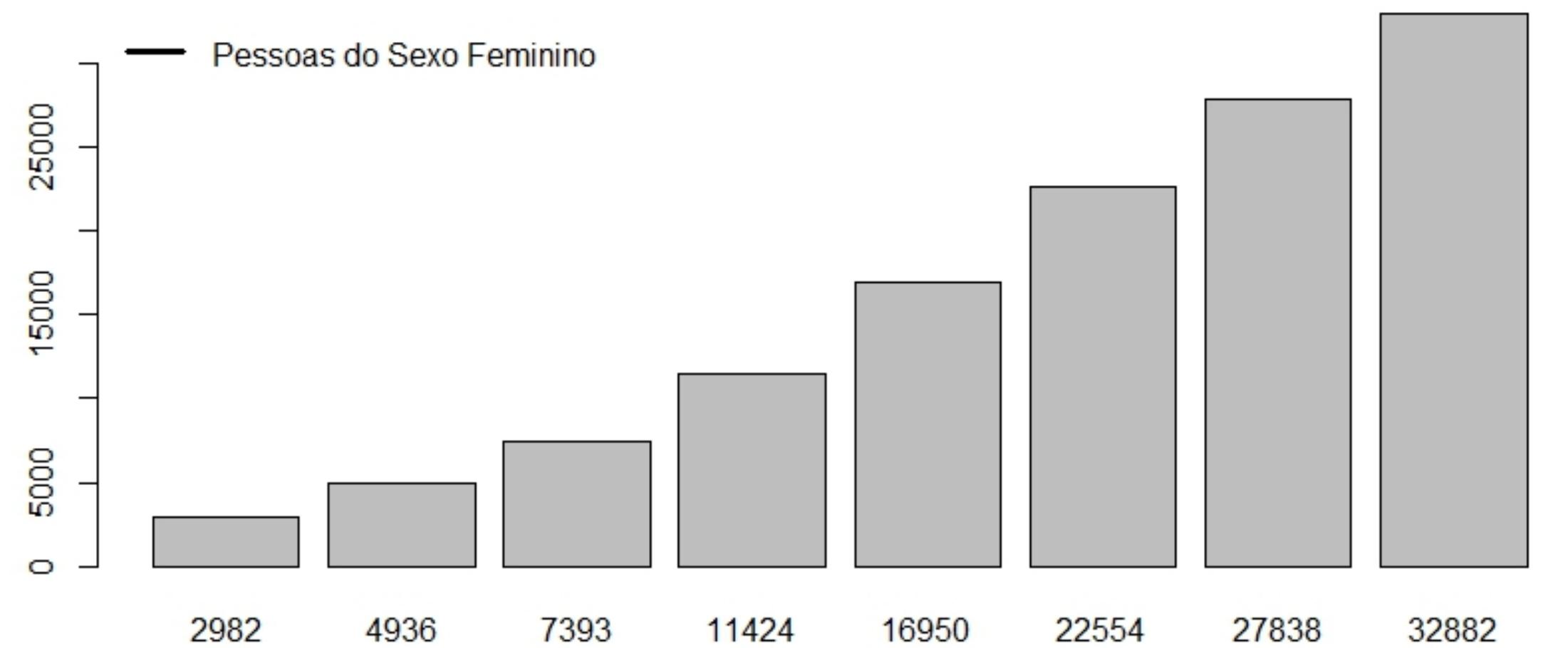


RESULTADOS DA ANÁLISE EXPLORATÓRIA

Taxa de Mortalidade - Atribuído a Óbito - 2015



Taxa de Mortalidade - Atribuído a Óbito - 2015



30 a 34 anos 35 a 39 anos 40 a 45 45 a 49 50 a 54 55 a 59 60 a 64 65 a 69

CONCLUSÕES





O desenvolvimento deste projeto sobre a taxa de mortalidade no Brasil foi marcado por desafios significativos na estruturação do dataset para análises de correlação. Enfrentando obstáculos na coleta, integração e formatação de dados heterogêneos de diversas fontes, a complexidade desse processo foi evidente. O uso do R para análises estatísticas também trouxe dificuldades técnicas notáveis, especialmente na conexão de conjuntos de dados distintos. No entanto, continuaremos a superar esses desafios, atrás de insights. **Este trabalho destaca a importância da qualidade na estruturação de dados para análises precisas, sublinhando a necessidade contínua de melhorar a acessibilidade e padronização dos dados na área da saúde.**