

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS**  
**NÚCLEO DE EDUCAÇÃO A DISTÂNCIA**  
**Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data**

**CAMILA MORAIS DE MELO - RA: 1077579**  
**CRISTIANE A. MASSENA - RA: 1143503**  
**DANIEL BICALHO - RA: 1143507**

**ANÁLISE DE DADOS MUNDIAIS SOBRE SUICÍDIO**

**CAMILA MORAIS DE MELO - RA: 1077579**

**CRISTIANE A. MASSENA - RA: 1143503**

**DANIEL M. BICALHO - RA: 1143507**

## **ANÁLISE DE DADOS MUNDIAIS SOBRE SUICÍDIO**

Trabalho de Conclusão de Curso apresentado  
ao Curso de Especialização em Ciência de  
Dados e Big Data como requisito parcial à  
obtenção do título de especialista.

Belo Horizonte

2019

## LISTA DE FIGURAS

|   |    |
|---|----|
| Figura 1: boxplot: média anual do rate de suicídio.....   | 17 |
| Figura 2: histogramas dos dados da média de temperatura e taxa de suicídio.....   | 18 |
| Figura 3: taxa de suicídio por idade.....   | 18 |
| Figura 4: relação entre temperatura média entre países e cidades.....   | 18 |
| Figura 5: Taxa de mortalidade por suicídio por faixa etária, entre 2007 e 2014. Taxas por 100.000 habitantes – Códigos da CID: X60-X84..... | 20 |
| Figura 6: taxa de suicídio por 100 mil habitantes, por sexo, Portugal (2007-2014).....  | 21 |
| Figura 7: Taxa de suicídio por 100 mil habitantes, por sexo, Brasil (1980-2010) (Silva, 2018).....  | 22 |
| Figura 8: Taxa de suicídio por 100 mil habitantes, por sexo, Brasil (1980-2010).....  | 22 |
| Figura 9: gráfico de correlação.....  | 26 |
| Figura 10: gráfico de correlação 2.....   | 27 |
| Figura 11: descrição da página.....   | 28 |
| Figura 12: introdução.....  | 28 |
| Figura 13: problema.....  | 29 |
| Figura 14: suicídio feminino.....   | 29 |
| Figura 15: influência do clima.....   | 30 |
| Figura 16: países com maiores índices por arma de fogo.....   | 30 |
| Figura 17: taxas de crescimento do país com relação ao ano anterior, crescimento da inflação do país e crescimento do GDP per capita.....   | 31 |
| Figura 18: boxplot por faixa etária.....  | 31 |
| Figura 19: crescimento por faixa de idade e tendências.....   | 32 |
| Figura 20: taxa de suicídio por países entre jovens.....  | 32 |
| Figura 22: conclusões.....  | 33 |

## LISTA DE TABELAS

|   |    |
|---|----|
| Tabela 1: levantamento do dataset: WHO - data visualizations dashboard.....   | 8  |
| Tabela 2: levantamento do dataset: IHME - Global Health Data Exchange.....  | 9  |
| Tabela 3: levantamento do dataset: IHME – Codebook.....   | 10 |
| Tabela 4: levantamento do dataset: World Cities.....  | 10 |
| Tabela 5: levantamento do dataset: Kaggle - Counties geographic coordinates.....  | 11 |
| Tabela 6: levantamento do dataset: BigQuery - WEATHER_GSOD.....   | 11 |
| Tabela 7: levantamento do dataset: Global Land Temperature By Country.....  | 12 |
| Tabela 8: levantamento do dataset: Global Land Temperature By Country – World Bank .....  | 12 |
| Tabela 9: Colunas do dataset final – Cidade (ihme_data_city).....   | 15 |
| Tabela 10: Colunas do dataset final – País (ihme_data_city).....  | 16 |
| Tabela 11: Taxa de mortalidade por lesões autoprovocadas por 100.000 hb, países selecionados, 2000-2015 e ordenados pelo IDH de 2015..... | 23 |
| Tabela 12: taxa de suicídio por 100 mil habitantes (2000-2010).....   | 23 |
| Tabela 13 Mortes por lesões autoprovocadas, coeficientes de mortalidade bruta por 100 mil, estimativas de totais globais, 2000-2015.....  | 24 |
| Tabela 14: taxa de suicídio por 100 mil e faixa etária (2000 a 2010).....   | 24 |

## SUMÁRIO

|  |           |
|--|-----------|
| <b>INTRODUÇÃO .....</b>                                | <b>6</b>  |
| <b>1.1 Contextualização .....</b>                      | <b>6</b>  |
| <b>1.2. O problema proposto .....</b>                  | <b>7</b>  |
| <b>2. COLETA DE DADOS .....</b>                        | <b>8</b>  |
| <b>3. PROCESSAMENTO/TRATAMENTO DE DADOS .....</b>      | <b>13</b> |
| <b>4. ANÁLISE E EXPLORAÇÃO DOS DADOS .....</b>         | <b>17</b> |
| <b>4.1. Resultados iniciais .....</b>                  | <b>17</b> |
| <b>4.2. Comparação dos dados .....</b>                 | <b>19</b> |
| <b>5. CRIAÇÃO DE MODELOS DE MACHINE LEARNING .....</b> | <b>26</b> |
| <b>6. APRESENTAÇÃO DOS RESULTADOS .....</b>            | <b>28</b> |
| <b>7. REFERÊNCIAS .....</b>                            | <b>34</b> |

## INTRODUÇÃO

Suicídio consiste no ato deliberado executado pelo próprio indivíduo, visando a morte, de forma consciente e intencional. Este auto-aniquilamento, vivenciado por aquele em situação de vulnerabilidade. O suicídio abrange diferentes classes sociais, idades e raças, e existem grupos que tem números preocupantes como idosos e adolescentes.

Ocorre quando não se encontra formas de lidar com o sofrimento psíquico, visualizando a morte como única alternativa viável, na qual o autoextermínio é apenas o ponto mais alto do sofrimento psíquico, haja vista que a vida de uma pessoa que comete suicídio é comumente marcada por eventos trágicos (KALINA; KOVADLOFF, 1983 apud TEIXEIRA, 2018)

Os números são preocupantes e já é considerado um problema de saúde pública mundial. (SANTOS, 2017). É preciso investigar os motivos que levaram a pessoa a tentar suicídio e as razões psicológicas que conduziram o sujeito a desistir da própria vida. (TEIXEIRA, 2019)

Os índices de suicídio crescem todos os anos e como ressaltam Minayo; Cavalcante, et all. (2010) mundialmente os suicídios matam mais que homicídios e guerras juntos, sendo que no Brasil a média nacional encontra-se por volta de 5,9 a cada 100.000 (cem mil) habitantes. Segundo Sérvio e Cavalcante (2013), e esses números podem ser maior devido à falta de qualidade dos dados referentes ao suicídio no Brasil.

### 1.1 Contextualização

O suicídio é um fenômeno complexo, por si só, e, por essa razão, tem despertado a atenção e o estudo das diversas áreas. A característica que marca a complexidade de se estudar o suicídio refere-se à combinação de fatores, que concorrem para um desfecho fatal, isto é, o caráter multifatorial do fenômeno tratado (MINAYO; CAVALCANTE, 2010).

No Brasil a inconsistência nos dados no sobre suicídio são mencionados em diversos trabalhos como enfatiza Mello (2002) sempre que estatísticas sobre suicídio são apresentadas, alguém se pergunta se elas são mesmo confiáveis. Não haveria casos de suicídio não registrados como tais? Sim, isso é de fato um problema, entre vários outros, quando interpretamos e comparamos esses índices (Mello Jorge et al., 2002).

Diante deste contexto, ainda levantar dados, tratar, garantir sua qualidade e analisar os resultados e auxiliaria na recomendação de programas preventivos.

## 1.2. O problema proposto

Estudos na área da saúde que estimule a discussão sobre qualidade dos dados são extremamente relevantes, o problema proposto neste trabalho é verificar os índices de suicídio em diversos países utilizando as bases de dados da **The Institute for Health Metrics and Evaluation (IHME)**, **World Health Statistics (WHO)**, Datasus e trabalhos qualitativos.

Analisaremos essas bases de dados no período de 1990 a 2017, compilaremos as informações relevantes de grupos de risco: idosos e adolescentes.

## 2. COLETA DE DADOS

Para obter os dados a serem analisados consideramos obter a maioria das informações de organizações oficiais ao invés de utilizar dados já agregados (Ex: Kaggle etc.), inicialmente levantamos os seguintes *datasets*:

- **World Health Statistics (WHO) data visualizations dashboard**

**Fonte:** <http://apps.who.int/gho/data/view.sdg.3-4-data-ctry?lang=en>

**Método de coleta:** Obtido via download direto do site em formato .csv

**Dataset:** WHO\_SDGSUICIDE.csv

**Data de acesso:** 19/06/2019

O dataset está originalmente organizado na seguinte forma:

Tabela 1: levantamento do dataset: World Health Statistics (WHO) data visualizations dashboard

| Nome da coluna/campo | Descrição  | Tipo    |
|----------------------|--|---------|
| Country              | País   | Texto   |
| Year                 | Ano  | Inteiro |
| Rate Both sexes - DC | Rate geral de causa de mortes comuns por 100k de habitantes.     | Decimal |
| Rate Male - DC       | Rate masculino de causa de mortes comuns por 100k de habitantes. | Decimal |
| Rate Female - DC     | Rate feminino de causa de mortes comuns por 100k de habitantes.  | Decimal |
| Rate Both sexes - SU | Rate geral de suicídio por 100k de habitantes.                   | Decimal |
| Rate Male - SU       | Rate masculino de suicídio por 100k de habitantes.               | Decimal |
| Rate Female - SU     | Rate feminino de suicídio por 100k de habitantes.                | Decimal |

\*DC - Probability (%) of dying between age 30 and exact age 70 from any of cardiovascular disease, cancer, diabetes, or chronic respiratory disease

\*SU - Crude suicide rates (per 100 000 population)

- **The Institute for Health Metrics and Evaluation (IHME) - Global Health Data Exchange**

**Fonte:** <http://ghdx.healthdata.org/gbd-results-tool>

**Método de coleta:** Obtido via download direto do site em formato .csv

**Dataset:** <csv>



**Data de acesso:** 20/06/2019

O dataset está originalmente organizado na seguinte forma:

Tabela 2: levantamento do dataset: The Institute for Health Metrics and Evaluation (IHME) - Global Health Data Exchange

| Nome da coluna/campo | Descrição   | Tipo    |
|----------------------|---|---------|
| Measure              | Medida, como utilizamos um filtro na extração este campo tem sempre o valor “deaths”.   | Texto   |
| Location             | País, estado e divisões geográficas.  | Texto   |
| Sex                  | Sexo, valores possíveis: “male”, “female” ou “both”.  | Texto   |
| Age                  | Distribuição entre idade, valores possíveis: <incluir lista>  | Inteiro |
| Cause                | Causa da morte, valores possíveis: <i>self-harm</i> , <i>self-harm by firearm</i> e <i>self-harm by other specified means</i> . | Texto   |
| Metric               | Rate de suicídio  | Decimal |
| Year                 | Ano   | Inteiro |
| Val                  | Valor absoluto que para ser informado por 100k por habitante  | Decimal |
| Upper                | Maior rate no período   | Decimal |
| Lower                | Menor rate no período   | Decimal |

**OBS:** O download foi realizado diretamente pelo site em diversos arquivos no formato .csv e agrupado em um único arquivo via *Python* em nosso código para facilitar a manipulação dos dados.

- **IHME – Codebook**

**Fonte:** [http://ghdx.healthdata.org/sites/default/files/ihme\\_query\\_tool/IHME\\_GB\\_D\\_2017\\_CODEBOOK.zip](http://ghdx.healthdata.org/sites/default/files/ihme_query_tool/IHME_GB_D_2017_CODEBOOK.zip)

**Método de coleta:** Obtido via download direto do site em formato .csv

**Dataset:** data/IHME\_GBD\_2017\_ALL\_LOCATIONS\_HIERARCHIES\_Y2018M11D18.XLSX

**Data de acesso:** 05/08/2019

O dataset está originalmente organizado na seguinte forma:

Tabela 3: levantamento do dataset: IHME – Codebook

| Nome da coluna/campo    | Descrição  | Tipo    |
|-------------------------|--|---------|
| Location_set_version_id | Versão da categorização de hierarquia de países/cidades. | Inteiro |
| Location_id             | Identificador do país ou cidade                          | Inteiro |
| Location_name           | Descrição da localização                                 | Texto   |
| Parent_id               | Identificador da hierarquia pai do país ou cidade        | Inteiro |
| Level                   | Nível: País, Cidade, Geo-localização, etc.               | Inteiro |
| Sort_order              | Ordem de ordenação                                       | Decimal |

- **World Cities**

**Fonte:** <https://creativecommons.org/licenses/by/4.0/>

**Método de coleta:** Obtido via download direto do site em formato .csv

**Dataset:** data/simplemaps\_worldcities\_basicv1.5/worldcities.csv

**Data de acesso:** 12/08/2019

O *dataset* está originalmente organizado na seguinte forma:

Tabela 4: levantamento do dataset: World Cities

| Nome da coluna/campo | Descrição  | Tipo     |
|----------------------|--|----------|
| City                 | Descrição do nome da cidade  | Texto    |
| City_ascii           | Descrição em código ASCII (sem caracteres especiais) do nome da cidade | Texto    |
| Lat                  | Valor da latitude da cidade  | Decimal  |
| Ing                  | Valor da longitude da cidade   | Decimal  |
| Country              | Descrição do nome do país  | Texto    |
| Iso2                 | Descrição do nome do país em código ISO2                               | Texto    |
| Iso3                 | Descrição do nome do país em código ISO3                               | Texto    |
| Admin_name           | Campo sistêmico  | Texto    |
| Capital              | Campo sistêmico  | Texto    |
| Population           | Total da população da cidade   | Numérico |
| Id                   | Identificador do registro da cidade (Campo sistêmico)                  | Numérico |

- **Kaggle - Counties geographic coordinates**

**Fonte:** <https://www.kaggle.com/eidanch/counties-geographic-coordinates>

**Método de coleta:** Obtido via download direto do site em formato .csv

**Dataset:** data/countries.csv

**Data de acesso:** 20/08/2019

O *dataset* está originalmente organizado na seguinte forma:

Tabela 5: levantamento do dataset: Kaggle - Counties geographic coordinates

| Nome da coluna/campo | Descrição                 | Tipo    |
|----------------------|---------------------------|---------|
| country              | Código ISO2 do país       | Texto   |
| latitude             | Informação de latitude    | Decimal |
| longitude            | Informação de longitude   | Decimal |
| NAME                 | Descrição do nome do país | Texto   |

- **BigQuery - WEATHER\_GSOD**

**Fonte:** [https://bigquery.cloud.google.com/table/fh-](https://bigquery.cloud.google.com/table/fh-bigquery:weather_gsod.all?pli=1)

[bigquery:weather\\_gsod.all?pli=1](https://bigquery.cloud.google.com/table/fh-bigquery:weather_gsod.all?pli=1)

**Método de coleta:** via Google Big Query (GBQ).

**Dataset:** fh-bigquery.weather\_gsod.all

**Data de acesso:** 01/09/2019

O *dataset* está originalmente organizado na seguinte forma:

Tabela 6: levantamento do dataset: BigQuery - WEATHER\_GSOD

| Nome da coluna/campo | Descrição                              | Tipo           |
|----------------------|--|----------------|
| year_                | Ano                                    | Inteiro (YYYY) |
| Country_id           | Código ISO do país ou cidade           | Texto          |
| avg_temp             | Média de temperatura do país ou cidade | Decimal        |

- **Global Land Temperature By Country**

**Fonte:** <https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data#GlobalLandTemperaturesByCountry.csv>

**Método de coleta:** Obtido via download direto do site em formato .csv

**Dataset:** data/GlobalLandTemperaturesByCountry.csv

**Data de acesso:** 28/09/2019

O *dataset* está originalmente organizado na seguinte forma:

Tabela 7: levantamento do dataset: Global Land Temperature By Country

| Nome da coluna/campo          | Descrição  | Tipo    |
|-------------------------------|--|---------|
| dt                            | Data do registro, no formato YYYY-MM-DD                      | Data    |
| AverageTemperature            | Média de temperatura   | Decimal |
| AverageTemperatureUncertainty | Intervalo de confiança de 95% em torno da temperatura mínima | Decimal |
| Country                       | Descrição do nome do país                                    | Texto   |

- **Global Land Temperature By Country**

**Fonte:** <https://data.worldbank.org/>

**Método de coleta:** Utilização do framework “**pandas\_datareader**” para ingestão dos dados.

**Dataset:** data/world\_bank\_data.csv

**Data de acesso:** 01/08/2019

O *dataset* está originalmente organizado na seguinte forma:

Tabela 8: levantamento do dataset: Global Land Temperature By Country – World Bank

| Nome da coluna/campo        | Descrição                            | Tipo    |
|-----------------------------|--------------------------------------|---------|
| GDP Growth                  | Crescimento do PIB                   | Decimal |
| GDP Per Capita              | PIB per capita                       | Decimal |
| Inflation - Consumer Prices | Inflação, preço ao consumidor (IPCA) | Decimal |
| Inflation - GDP Deflator    | Inflação como alterador do PIB       | Decimal |

O dataset inicial está disponível em:

<https://drive.google.com/file/d/1GgD1rde7BHrdsP-KwYz4s3ZVluw9iaCd/view?usp=sharing>

### 3. PROCESSAMENTO/TRATAMENTO DE DADOS

Nesta etapa do trabalho iremos tratar os dados coletados, realizando a tratativa dos valores ausentes, duplicados e ajuste do *dataset* para aplicação modelo ML.

Dentro das etapas realizadas neste processo identificou-se algumas tarefas **frequentemente** utilizadas, desta forma criou-se algumas funções para facilitar a aplicação destas mudanças e equalização dos registros:

- A função "**trata\_texto**" é utilizada para retirar caracteres especiais e manter todo o campo em maiúsculo;
- A função "**trata\_nome\_paises**" é utilizada para unificar a nomenclatura dos países, essa função é extremamente importante no momento do merge dos *datasets*.
- A função "**comprimir\_arquivo**" É utilizada para comprimir os *datasets* finais no formato *pickle* para ganhar agilidade no carregamento do mesmo no momento da aplicação do modelo ML.
- A função "**carregar\_arquivo**" É utilizada para carregar os *datasets* armazenados em *pickle* para *dataframe*.

A etapa de tratamento do processo seguiu o seguinte fluxo (macro):

**1. Agrupamento dos dados:** Nesta etapa adequou-se os dados que seriam cruzados para serem equivalentes, por exemplo, no dataset "GlobalLandTemperaturesByCountry.csv" o campo "dt" continha a data do registro de temperatura, mas já o dataset obtido do IHME está dividido por ano, desta forma, agrupou-se os dados de temperatura por ano.

Filtrou-se e agrupou-se os dados do **global\_land\_temp\_by\_country** da seguinte forma:

- Apenas temperaturas após o ano de **1990** serão consideradas, pois o dataset do IHME (ao qual faremos o merge) contém dados após esta data;
- Agrupamos por ano/País os dados de temperatura para possibilitar o merge entre os *datasets*. (Este conjunto de dados será utilizado futuramente para completar os dados de temperatura que faltam do dataset obtido pela base do BigQuery)

**2. Transposição dos dados:** Para os dados de Cidade realizamos a transposição das colunas “**Country**” e “**Geo Region**”, neste *dataset* originalmente continha uma coluna com a informação de “Location” e que em linhas divergentes armazenava a informação do País e Divisões Geográficas etc., realizamos a tratativa para agrupar essa informação em colunas e não mais em linhas e facilitar a aplicação do modelo ML.

**3. Inclusão da informação de Latitude e Longitude:** Realizamos o cruzamento dos dados com intuito de incluir a informação de latitude e longitude no mesmo, essa informação será utilizada no momento da visualização dos dados no *Tableu Public* para montagem dos Mapas.

**4. Tratativa dos dados inexistentes:** Os países **NAMÍBIA, SUL DO SUDÃO E SÃO TOMÉ E PRÍNCÍPIO** necessitaram de ajustes pontuais no *dataset*, pois não continham a informação de latitude e longitude.

**5. Tratativa dos dados duplicados:** Os dados duplicados foram **mantidos** no *datasets* principais, pois, iremos utilizá-los como filtro na apresentação dos resultados e na aplicação do modelo ML, são eles:

- Informação de Sexo (Ambos);
- Informação de Idade (Todas);
- Causa de Morte (Suicídio - sem informação de método);
- Dados de divisões geográficas (*Level* igual à 5).

**6. Inclusão dos dados de Temperatura média:** Realizamos também o merge do *dataset* para obter as informações de temperatura média

(`global_temp_by_country_data`), desta forma agrupou-se o *dataframe* por **ano** e depois realizamos o merge entre os *datasets* (`ihme_data_city` e `ihme_data_country`)

**7. Tratativa dos dados Nulos:** Os países com dados NAN (**faltantes**) são carregados com o valor médio de temperatura do país (`global_land_temp_by_country`) esta tratativa foi necessária, pois **276480** registros não continham dados de temperatura, para avaliar esta situação utilizamos da função `info()` do pandas que nos mostra quantos registros não nulos existem por coluna do *dataframe*.

```
ihme_data_country.info()

<class 'pandas.core.frame.DataFrame'>
Int64Index: 790272 entries, 0 to 790271
Data columns (total 13 columns):
Country          790272 non-null object
Sex              790272 non-null object
Age             790272 non-null object
Cause_of_Death  790272 non-null object
Year            790272 non-null int64
Total           790272 non-null float64
Upper           790272 non-null float64
Lower           790272 non-null float64
Country_id      790272 non-null object
latitude        790272 non-null object
longitude       790272 non-null object
year_           513792 non-null float64
avg_temp        513792 non-null float64
dtypes: float64(5), int64(1), object(7)
memory usage: 84.4+ MB
```

**Documentação original da função “df.info()” do pandas:** “This method prints information about a *DataFrame* including the index dtype and column dtypes, non-null values and memory usage.”

## 8. Colunas do *dataset* final – Cidade (`ihme_data_city`)

Tabela 9: Colunas do *dataset* final – Cidade (`ihme_data_city`)

| Coluna                | Descrição                                    |
|-----------------------|--|
| <b>City</b>           | Descrição do nome da cidade                  |
| <b>Sex</b>            | Informação do Sexo                           |
| <b>Age</b>            | Idade  |
| <b>Cause_of_Death</b> | Causa da Morte                               |
| <b>Year</b>           | Ano  |
| <b>Total</b>          | Taxa total de suicídio                       |
| <b>Upper</b>          | Maior taxa registrada de suicídio no período |
| <b>Lower</b>          | Menor taxa registrada de suicídio            |

|                     |                     |
|---------------------|---------------------|
|                     | no período          |
| <b>Country</b>      | País                |
| <b>Geo_Region</b>   | Divisão geográficas |
| <b>avg_temp</b>     | Temperatura media   |
| <b>Country_Code</b> | Código ISO2 do país |

### Exemplo:

| City  | Sex  | Age      | Cause_of_Death                     | Year | Total                | Upper                | Lower                 | Country | Geo_Region               | avg_temp | Country_Code |
|-------|------|----------|------------------------------------|------|----------------------|----------------------|-----------------------|---------|--------------------------|----------|--------------|
| OSAKA | Male | 10 to 14 | Self-harm                          | 2007 | 1.0874730617079815   | 1.397241689658009    | 0.8352392415618621    | JAPAN   | High-income Asia Pacific | 18.0     | JA           |
| OSAKA | Male | 10 to 14 | Self-harm by firearm               | 2007 | 0.010181820690978371 | 0.017303236350611567 | 0.0052462847727573975 | JAPAN   | High-income Asia Pacific | 18.0     | JA           |
| OSAKA | Male | 10 to 14 | Self-harm by other specified means | 2007 | 1.0772912410170026   | 1.3831582222388168   | 0.8291851303664991    | JAPAN   | High-income Asia Pacific | 18.0     | JA           |

## 9. Colunas do *dataset* final – País (ihme\_data\_country)

Tabela 10: Colunas do *dataset* final – País (ihme\_data\_city)

| Coluna                | Descrição                                    |
|-----------------------|--|
| <b>Country</b>        | Descrição do nome da cidade                  |
| <b>Sex</b>            | Informação do Sexo                           |
| <b>Age</b>            | Idade  |
| <b>Cause_of_Death</b> | Causa da Morte                               |
| <b>Year</b>           | Ano  |
| <b>Total</b>          | Taxa total de suicídio                       |
| <b>Upper</b>          | Maior taxa registrada de suicídio no período |
| <b>Lower</b>          | Menor taxa registrada de suicídio no período |
| <b>Country_id</b>     | Código ISO2 do país                          |
| <b>Latitude</b>       | Latitude                                     |
| <b>Longitude</b>      | Longitude                                    |
| <b>avg_temp</b>       | Temperatura media                            |

### Exemplo:

| Country | Sex  | Age      | Cause_of_Death                     | Year | Total               | Upper               | Lower               | Country_id | latitude   | longitude | avg_temp |
|---------|------|----------|------------------------------------|------|---------------------|---------------------|---------------------|------------|------------|-----------|----------|
| BRAZIL  | Male | 10 to 14 | Self-harm                          | 2007 | 0.7992355724014127  | 0.8737300393401292  | 0.7205279592114744  | BR         | -14.235004 | -51.92528 | 25.0     |
| BRAZIL  | Male | 10 to 14 | Self-harm by firearm               | 2007 | 0.09451103692346004 | 0.14704596749366405 | 0.06053247559317741 | BR         | -14.235004 | -51.92528 | 25.0     |
| BRAZIL  | Male | 10 to 14 | Self-harm by other specified means | 2007 | 0.7047245354779529  | 0.7799959843901995  | 0.6195507524292094  | BR         | -14.235004 | -51.92528 | 25.0     |

Notebook

passo

a

passo:

<https://colab.research.google.com/drive/1fnJWfhH08KPrtrcCJ5nS4W-UeS02umcA>



## 4. ANÁLISE E EXPLORAÇÃO DOS DADOS

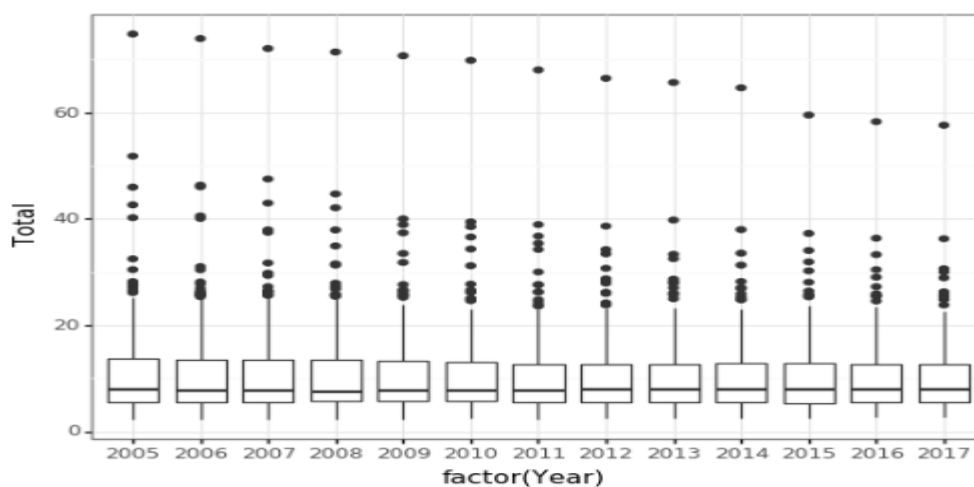
Nesta etapa do trabalho realizou-se a análise dos dados inicialmente coletados e cruzou-se os resultados com alguns registros encontrados em artigos acadêmicos.

### 4.1. Resultados iniciais

Antes de iniciar a montagem da visualização final deste trabalho, foi realizado algumas análises/avaliações gerais do *dataset*, os dados de monitoramento de temperatura média de cidade existem apenas para os países: **INDONÉSIA, ESTADOS UNIDOS, JAPÃO, MÉXICO, BRASIL, REINO UNIDO e SUÉCIA**. No *dataset* com a média de temperatura de países todos contêm os dados completos esperado na análise, desta forma será o principal ponto de partida.

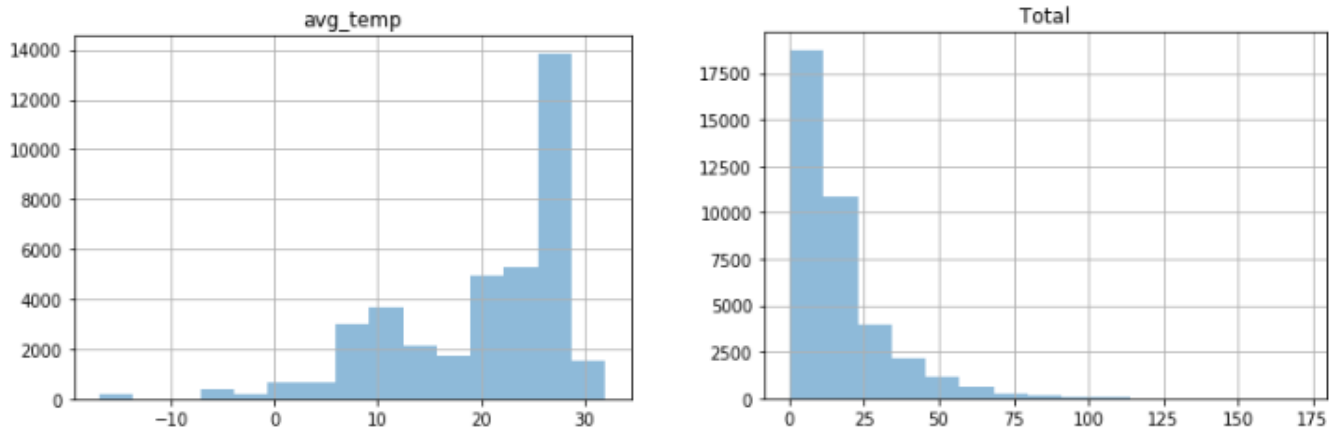
Nossa primeira avaliação, foi elaborado um gráfico **boxplot** para avaliar a média anual do rate de suicídio por **país** e identificar os primeiros países que estavam muito acima da mediana, podemos identificar que para os países fora da curva (*outlier*) ano a ano estão diminuindo suas taxas de suicídio.

Figura 1: box plot: média anual do rate de suicídio



Foi criado também dois histogramas, o primeiro contém os dados da média de temperatura e o segundo foi elaborado a partir da taxa de suicídio (**Total**).

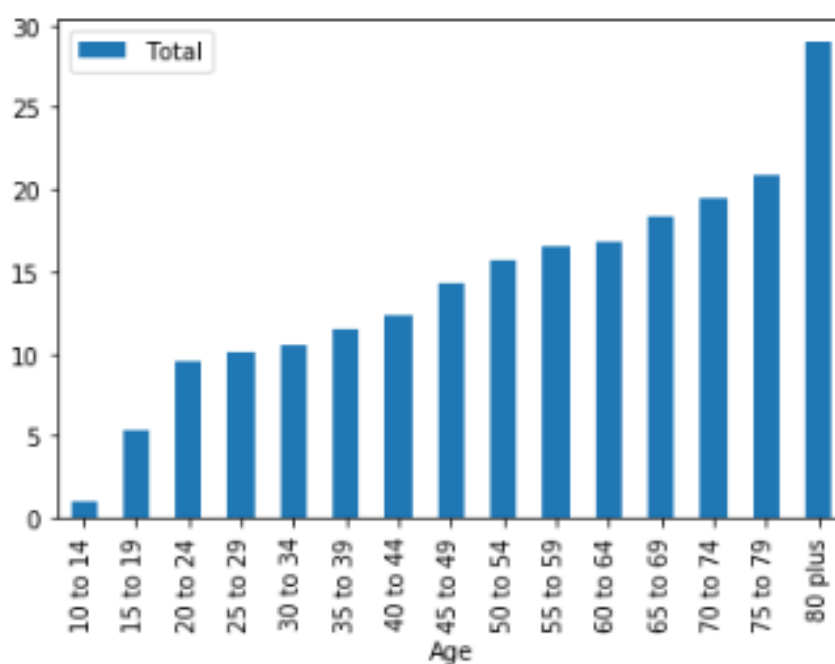
Figura 2: histogramas dos dados da média de temperatura e taxa de suicídio.



Ao avaliar os histogramas identificamos que a média de temperatura dos países possui uma distribuição enviesada à esquerda e já a distribuição do rate de suicídio possui uma distribuição enviesada à direita.

Ao criar um gráfico de barras por taxa de suicídio por idade, podemos ver que a idade com maior risco de suicídio está entre os idosos com idade igual ou maior que **80 anos**.

Figura 3: taxa de suicídio por idade

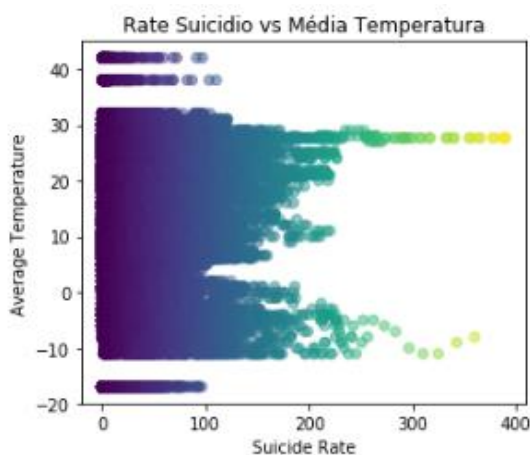


Após o processamento inicial, avalio-se a temperatura média de cada país influenciaria o aumento da taxa de suicídio como informação encontrada em alguns dos artigos acadêmicos estudados, do lado esquerdo há um gráfico de dispersão entre a média de taxa de suicídio e temperatura média entre os dados de **países** e ao lado direito um gráfico de dispersão também entre a média da taxa de suicídio e a média de temperatura entre as **cidades** (Considerando os países: **INDONÉSIA, ESTADOS UNIDOS, JAPÃO, MÉXICO, BRASIL, REINO UNIDO e SUÉCIA**).

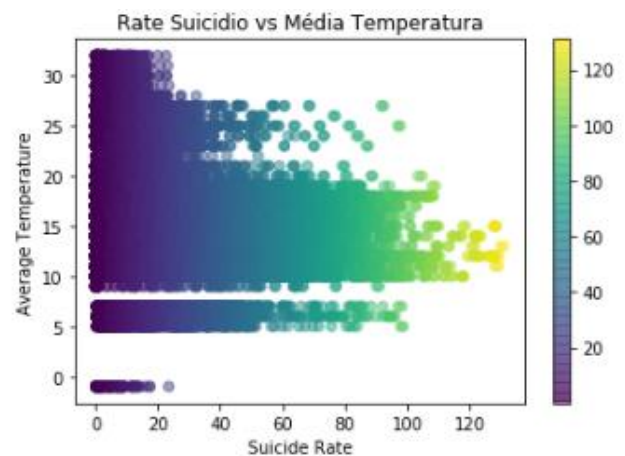
Identificou-se que a relação entre a temperatura média está fortemente ligada nos dados das cidades existentes no *dataset*, porém para os dados coletados de países não possui a mesma relação. Na relação entre as cidades houve

Figura 4: relação entre temperatura media entre paises e cidades

#### Países



#### Cidades

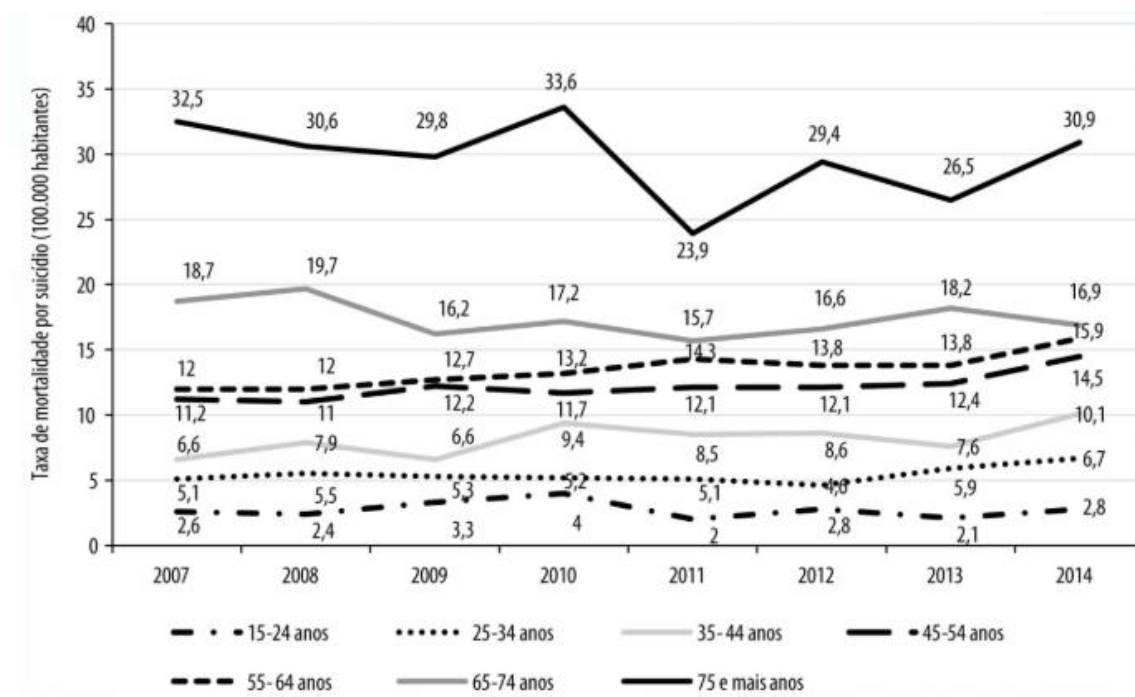


## 4.2. Comparação dos dados

Buscou-se na literatura reforçar as hipóteses levantadas em diversos países e também para comparar os dados obtidos para refutar ou não as teorias levantadas.

No trabalho de Nunes (2018) foi possível levantar as taxas de suicídio por faixa etária entre 2007 a 2017.

Figura 5: Taxa de mortalidade por suicídio por faixa etária, entre 2007 e 2014. Taxas por 100.000 habitantes – Códigos da CID: X60-X84.



Fonte: Instituto Nacional de Estatística (apud NUNES, 2018)

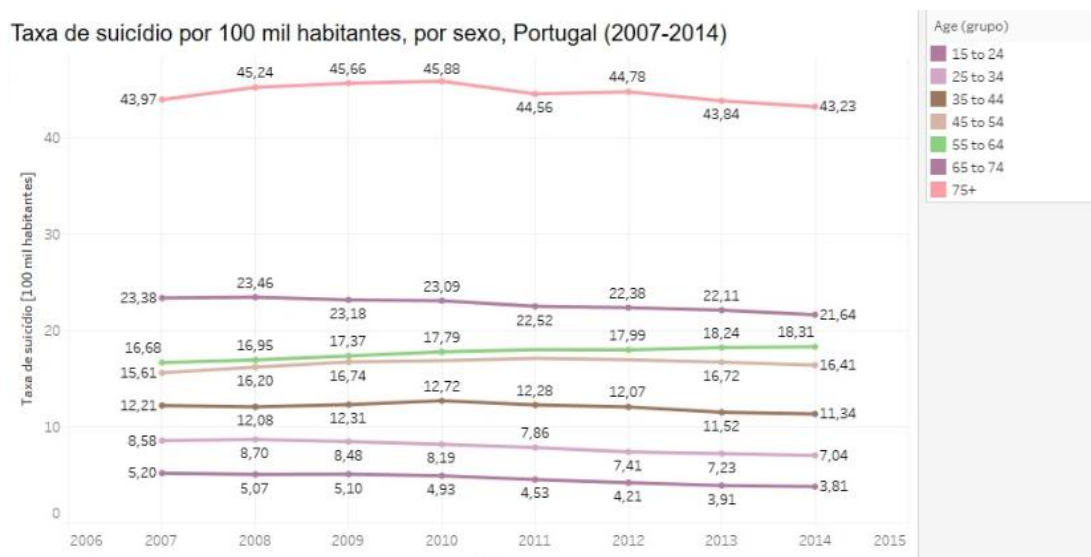
A figura mostra os dados evolutivos sobre a mortalidade por lesão autoprovocada em Portugal, não se consideram os casos de óbitos de crianças com idade inferior a 15 anos como menciona Nunes (2018). O autor ainda avaliou que entre 2007 e 2014, a taxa de mortalidade por suicídio registrou aumento da taxa de suicídio em todas as faixas etárias:

- 15 a 24 anos: pico em 2010, tendo em 2014 se aproximado dos valores de 2007;
- 25 a 34 anos: pico em 2014, tendo ficado 1,6% acima da taxa de 2007;
- 35 a 44 anos: pico em 2010 (9,4%) e em 2014 (10,1% - 3,5% acima da taxa de 2007);
- 45 a 54 anos: pico em 2009 e em 2014 (14,5% - 3,3% acima da taxa de 2007);
- 55 a 64 anos: pico em 2014 (15,9% - 3,9% acima da taxa de 2007);

- 65 a 74 anos: pico em 2008, tendo em 2014 apresentado uma taxa abaixo da registrada em 2007 (-1,8%);
- 75 anos e mais: pico em 2010, tendo em 2014 apresentado uma taxa abaixo da registrada em 2007 (-1,6%). (NUNES, 2018)

A figura 6 a seguir mostra esses dados com a variação com os números da variação descrita, com enriquecimento dos dados, tornando as oscilações mais relevantes.

Figura 6: taxa de suicídio por 100 mil habitantes, por sexo, Portugal (2007-2014)



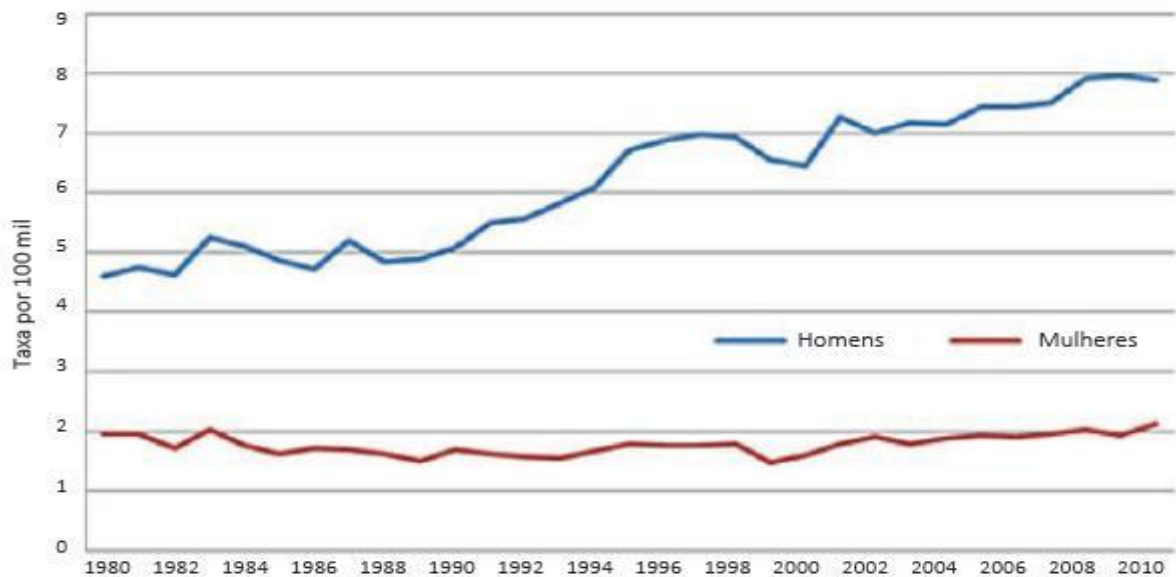
No trabalho de Silva, et all (2018) a análise temporal no período de 1980 a 2010, mais de 196 mil pessoas cometeram suicídio no Brasil. Na década de 1980, a média anual ficou em torno de quatro mil suicídios, passando para seis mil na década seguinte e mais de oito mil na década de 2000.

Ao longo de todo esse período, variaram substancialmente pouco, passando de 3,27 no primeiro ano analisado para 4,95 em 2010. Essa aparente estabilidade, no entanto, omite uma diferença significativa se comparada a incidência na população masculina e feminina. (SILVA; et all, 2018)

Como comprova também pelo trabalho de Silva et all (2018), entre as mulheres, as taxas de suicídio são de fato estáveis ao longo dos anos 1980 a

2010. Entre os homens, ao contrário, o padrão é de um crescimento linear, alcançando um valor quatro vezes superior à taxa das mulheres em 2010.

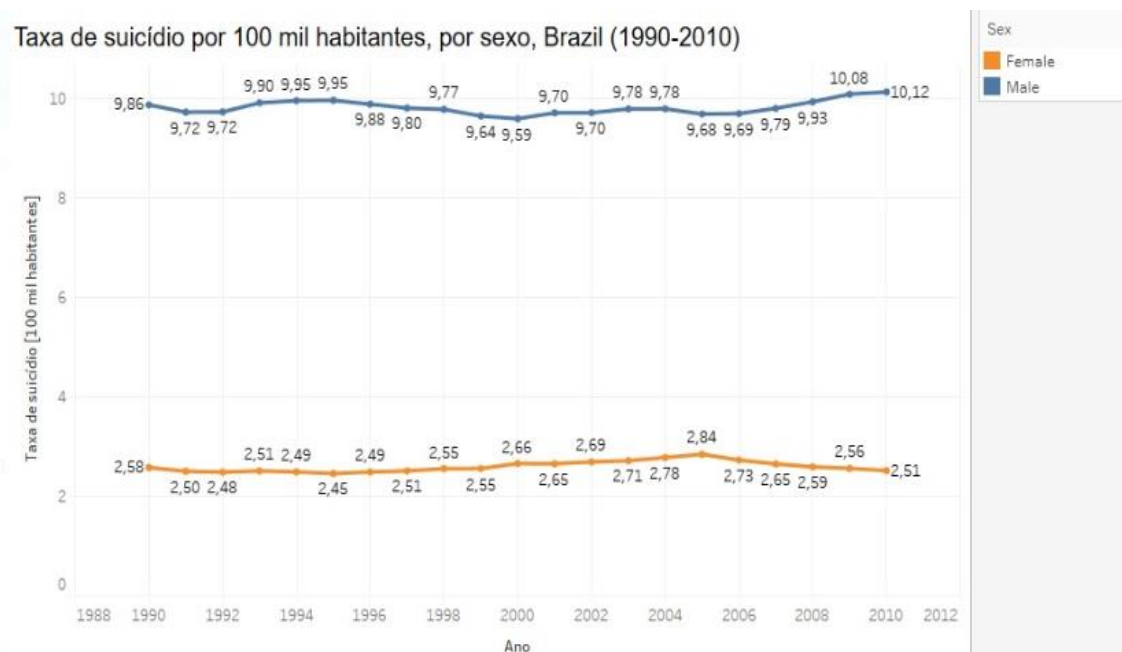
Figura 7: Taxa de suicídio por 100 mil habitantes, por sexo, Brasil (1980-2010)



Fonte: (SILVA; et all, 2018)

O mesmo enriquecimento foi feito na tabela a seguir acrescentou-se mais detalhes ao gráfico.

Figura 8: Taxa de suicídio por 100 mil habitantes, por sexo, Brasil (1980-2010)



Os dados apresentados no ano de 2015 em países de diferentes continentes, culturas, sistemas políticos e econômicos mostra que o suicídio guarda fortes características locais.

Logo, para além de fatores biológicos e neurofisiológicos que têm sido identificados no plano individual, diferenças observadas inclusive entre países de tipos de desenvolvimento próximos, além de compartilharem sistemas políticos e econômicos e proximidade geográfica fortalecem as dimensões sociais e culturais envolvidas nestes agravos. (RIBEIRO, MOREIRA, 2018).

Tabela 11: Taxa de mortalidade por lesões autoprovocadas por 100.000 hb, países selecionados, 2000-2015 e ordenados pelo IDH de 2015.

| Países         | IDH<br>2015 | CM<br>2000 | CM<br>2005 | CM<br>2010 | CM<br>2015 |
|----------------|-------------|------------|------------|------------|------------|
| Austrália      | 0,939       | 13,40      | 11,90      | 12,10      | 11,80      |
| Alemanha       | 0,926       | 14,60      | 13,60      | 13,50      | 13,40      |
| Dinamarca      | 0,925       | 15,50      | 14,70      | 11,70      | 12,30      |
| Holanda        | 0,924       | 9,70       | 9,80       | 9,80       | 11,90      |
| Irlanda        | 0,923       | 12,50      | 11,50      | 11,60      | 11,70      |
| Canadá         | 0,920       | 12,40      | 12,60      | 12,70      | 12,30      |
| Estados Unidos | 0,920       | 10,80      | 11,70      | 13,00      | 14,30      |
| Suécia         | 0,913       | 15,10      | 15,70      | 14,60      | 15,40      |
| Japão          | 0,903       | 24,40      | 24,70      | 24,30      | 19,70      |
| França         | 0,897       | 20,20      | 19,20      | 18,40      | 16,90      |
| Bélgica        | 0,896       | 22,60      | 20,50      | 20,30      | 20,50      |

Segundo estes dados de Ribeiro e Moreira (2018), o Brasil possui taxas bem menores que outros importantes países sul-americanos, como a Argentina e o Chile. Igualmente, com relação ao México e aos demais países da América do Norte. Alguns países tem taxas assustadoras mesmo em períodos menores como 2000 a 2010.

Tabela 12: taxa suicidio por 100 mil habitantes (2000-2010)

#### Taxa de suicídio por 100 mil habitantes (2000-2010)

| Country       | Year  |       |       |       |       |       |       |       |       |       |       |
|---------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|               | 2000  | 2001  | 2002  | 2003  | 2004  | 2005  | 2006  | 2007  | 2008  | 2009  | 2010  |
| AUSTRALIA     | 13,77 | 13,13 | 12,70 | 12,31 | 12,01 | 11,79 | 11,67 | 12,02 | 12,24 | 12,32 | 12,36 |
| BELGIUM       | 23,22 | 22,99 | 22,58 | 22,08 | 21,27 | 21,22 | 20,59 | 20,88 | 21,29 | 21,22 | 21,00 |
| CANADA        | 14,14 | 14,02 | 13,85 | 13,93 | 13,61 | 13,68 | 13,30 | 13,39 | 13,40 | 13,35 | 13,12 |
| DENMARK       | 16,76 | 16,14 | 15,96 | 15,32 | 14,76 | 13,94 | 14,15 | 14,14 | 13,59 | 13,46 | 13,10 |
| FRANCE        | 22,29 | 22,09 | 21,88 | 21,58 | 20,47 | 20,42 | 20,15 | 19,87 | 19,95 | 20,29 | 20,21 |
| GERMANY       | 16,32 | 16,01 | 15,95 | 15,89 | 15,24 | 14,91 | 14,52 | 14,34 | 14,35 | 14,49 | 14,52 |
| IRELAND       | 13,89 | 13,62 | 13,28 | 12,48 | 12,46 | 12,34 | 12,18 | 12,16 | 12,39 | 12,90 | 11,80 |
| JAPAN         | 24,75 | 25,05 | 25,26 | 26,09 | 25,96 | 26,40 | 25,96 | 25,91 | 25,69 | 25,52 | 25,28 |
| SWEDEN        | 16,28 | 16,22 | 15,79 | 15,53 | 16,00 | 15,54 | 15,43 | 15,43 | 15,49 | 15,50 | 14,95 |
| UNITED STATES | 12,51 | 12,73 | 13,00 | 13,11 | 13,00 | 13,32 | 13,49 | 13,57 | 13,72 | 13,86 | 13,80 |



Ainda para a população global, na tabela de Ribeiro e Moreira (2018) são apresentadas as taxas de mortalidade bruta para a série resumida de 2000 a 2015 por sexo e faixa etária. São observadas as grandes diferenças por sexo, com o maior peso do suicídio entre homens e que se avolumam acima dos 30 anos. Em todas as condições, o padrão observado é o da redução de taxas ao longo do período.

Tabela 13: Mortes por lesões autoprovocadas, coeficientes de mortalidade bruta por 100 mil, estimativas de totais globais, 2000-2015.

**Tabela 2** Mortes por lesões autoprovocadas, coeficientes de mortalidade bruta por 100 mil, estimativas de totais globais, 2000-2015.

| Sexo      | Faixa etária         | 2000  | 2005  | 2010  | 2015  |
|-----------|----------------------|-------|-------|-------|-------|
| Total     |                      | 12,22 | 11,61 | 11,23 | 10,73 |
| Masculino | Total (todas idades) | 15,55 | 14,67 | 14,16 | 13,61 |
| Feminino  |                      | 8,84  | 8,51  | 8,26  | 7,80  |
| Masculino | 0-28 dias            | 0,00  | 0,00  | 0,00  | 0,00  |
|           | 1-59 meses           | 0,00  | 0,00  | 0,00  | 0,00  |
|           | 5-14 anos            | 1,23  | 1,14  | 1,1   | 1,10  |
|           | 15-29 anos           | 17,12 | 15,6  | 14,84 | 14,06 |
|           | 30-49 anos           | 20,91 | 18,87 | 17,62 | 16,54 |
|           | 50-59 anos           | 25,98 | 23,47 | 21,72 | 19,74 |
|           | 60-69 anos           | 30,53 | 27,44 | 25,65 | 23,80 |
|           | 70+ anos             | 47,11 | 44,47 | 41,89 | 40,98 |
| Feminino  | 0-28 dias            | 0,00  | 0,00  | 0,00  | 0,00  |
|           | 1-59 meses           | 0,00  | 0,00  | 0,00  | 0,00  |
|           | 5-14 anos            | 1,13  | 1,07  | 1     | 0,92  |
|           | 15-29 anos           | 13,00 | 12,15 | 11,46 | 10,34 |
|           | 30-49 anos           | 9,47  | 8,81  | 8,21  | 7,65  |
|           | 50-59 anos           | 10,62 | 9,89  | 9,81  | 9,36  |
|           | 60-69 anos           | 14,50 | 13,57 | 13,32 | 12,92 |
|           | 70+ anos             | 22,25 | 21,33 | 20,48 | 19,74 |

Fonte: WHO, 2018.

Realizou-se um desdobramento dos anos e um enriquecimento das taxas de suicídio entre os anos de 2000 a 2010

Tabela 14: taxa de suicídio por 100 mil e faixa etária (2000 a 2010).

Taxa de suicídio por 100 mil habitantes sexo e faixa etária (2000-2010)

| Sex    | Age (grupo) | Year  |       |       |       |       |       |       |       |       |       |       |
|--------|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
|        |             | 2000  | 2001  | 2002  | 2003  | 2004  | 2005  | 2006  | 2007  | 2008  | 2009  | 2010  |
| Female | 10 to 14    | 1,03  | 1,01  | 0,99  | 0,98  | 0,97  | 0,95  | 0,94  | 0,94  | 0,92  | 0,90  | 0,89  |
|        | 15 to 24    | 6,15  | 5,98  | 5,88  | 5,81  | 5,71  | 5,63  | 5,50  | 5,39  | 5,31  | 5,19  | 5,09  |
|        | 25 to 34    | 6,47  | 6,30  | 6,18  | 6,09  | 5,99  | 5,88  | 5,69  | 5,55  | 5,46  | 5,35  | 5,22  |
|        | 35 to 44    | 7,19  | 7,03  | 6,92  | 6,83  | 6,72  | 6,61  | 6,42  | 6,28  | 6,16  | 6,01  | 5,86  |
|        | 45 to 54    | 8,98  | 8,82  | 8,70  | 8,60  | 8,46  | 8,32  | 8,10  | 7,94  | 7,83  | 7,70  | 7,55  |
|        | 55 to 64    | 10,76 | 10,54 | 10,40 | 10,27 | 10,12 | 9,97  | 9,72  | 9,50  | 9,34  | 9,14  | 8,98  |
|        | 65 to 74    | 14,21 | 13,89 | 13,68 | 13,53 | 13,27 | 13,03 | 12,70 | 12,43 | 12,21 | 11,97 | 11,78 |
|        | 75+         | 19,14 | 18,79 | 18,61 | 18,47 | 18,21 | 17,96 | 17,60 | 17,34 | 17,14 | 16,89 | 16,74 |
|        | All Ages    | 5,71  | 5,61  | 5,55  | 5,51  | 5,45  | 5,40  | 5,29  | 5,21  | 5,15  | 5,07  | 4,99  |
| Male   | 10 to 14    | 1,82  | 1,78  | 1,74  | 1,72  | 1,69  | 1,69  | 1,67  | 1,65  | 1,62  | 1,59  | 1,57  |
|        | 15 to 24    | 17,04 | 16,86 | 16,64 | 16,35 | 16,12 | 16,00 | 15,82 | 15,70 | 15,54 | 15,24 | 15,00 |
|        | 25 to 34    | 23,65 | 23,49 | 23,33 | 23,08 | 22,84 | 22,71 | 22,37 | 22,14 | 21,83 | 21,32 | 21,00 |
|        | 35 to 44    | 26,64 | 26,53 | 26,29 | 26,03 | 25,73 | 25,50 | 25,07 | 24,70 | 24,35 | 23,83 | 23,42 |
|        | 45 to 54    | 30,15 | 30,06 | 29,84 | 29,65 | 29,27 | 29,14 | 28,67 | 28,25 | 27,85 | 27,37 | 26,94 |
|        | 55 to 64    | 31,71 | 31,56 | 31,27 | 31,10 | 30,74 | 30,55 | 30,16 | 29,75 | 29,43 | 29,12 | 28,87 |
|        | 65 to 74    | 39,01 | 38,66 | 38,27 | 37,98 | 37,48 | 37,07 | 36,55 | 36,07 | 35,74 | 35,41 | 35,16 |
|        | 75+         | 58,92 | 58,41 | 58,00 | 57,77 | 57,12 | 56,58 | 56,01 | 55,53 | 55,20 | 54,77 | 54,59 |
|        | All Ages    | 17,73 | 17,72 | 17,66 | 17,58 | 17,45 | 17,45 | 17,28 | 17,17 | 17,04 | 16,82 | 16,67 |



A partir das comparações de todos os datasets e artigos podem-se criar um estudo próprio onde elaborou-se alguns comparativos que está disponível em:

<https://public.tableau.com/profile/mae.jemison.bigdata.project#!/vizhome/TCC-PucMinasAnlisededadosobresuicidio/TCC?publish=yes>

Todos os dados levantados, datasets estão disponíveis no notebook da equipe: <https://colab.research.google.com/drive/1fnJWfhH08KPrtrcCJ5nS4W-UeS02umcA>

## 5. CRIAÇÃO DE MODELOS DE MACHINE LEARNING

Foi coletado os dados e após realizar a correlação de Pearson optou-se por eliminar um dos índices inflacionários, pois possuem alta correlação e não vai contribuir com o modelo.

Aplicando um modelo de regressão, iremos avaliar e tentar prever os valores de suicídio de um intervalo de tempo

Usando a métrica do erro quadrático médio para verificar a performance do modelo. A seguir apresenta-se os gráficos de correlação.

Figura 9: gráfico de correlação

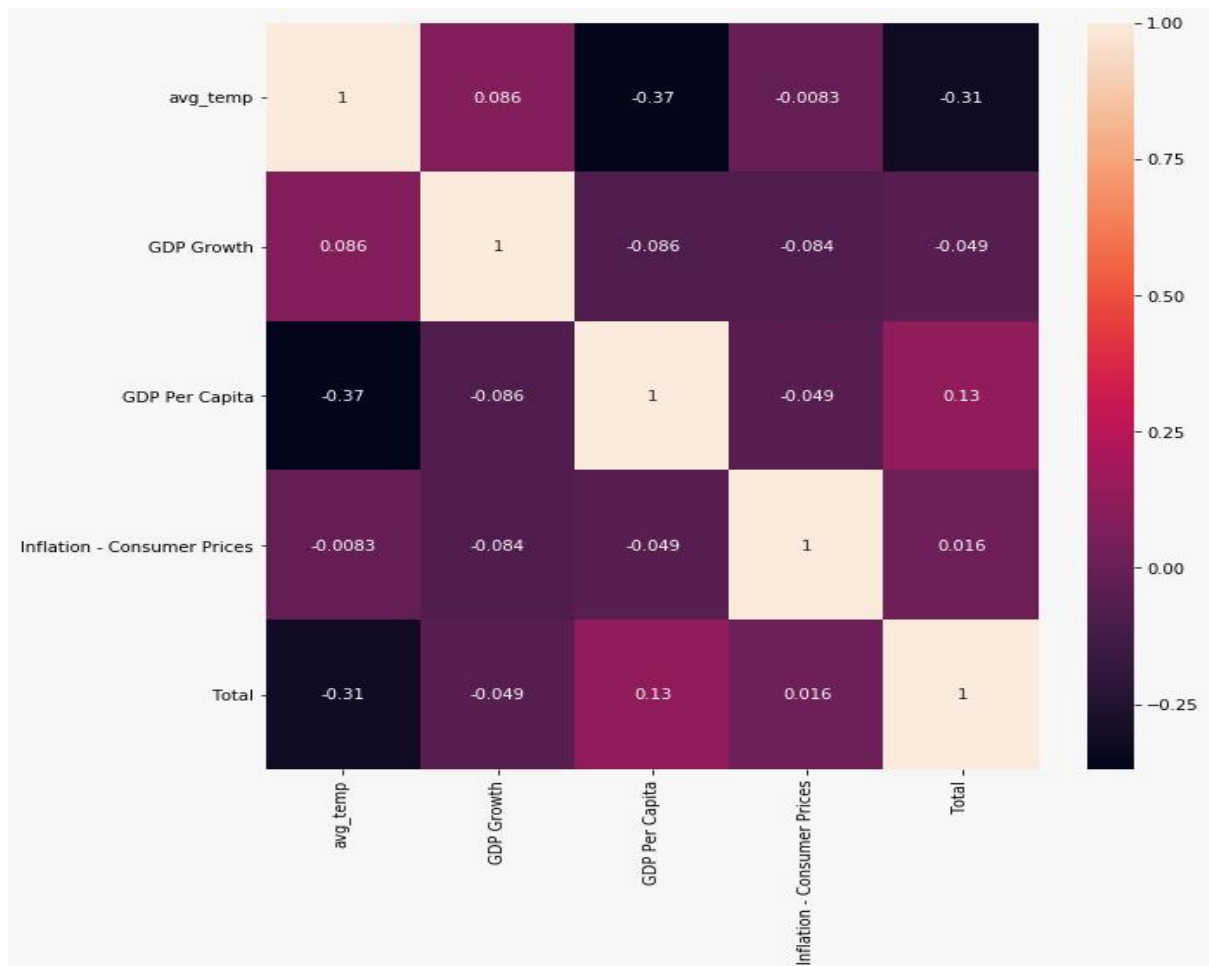
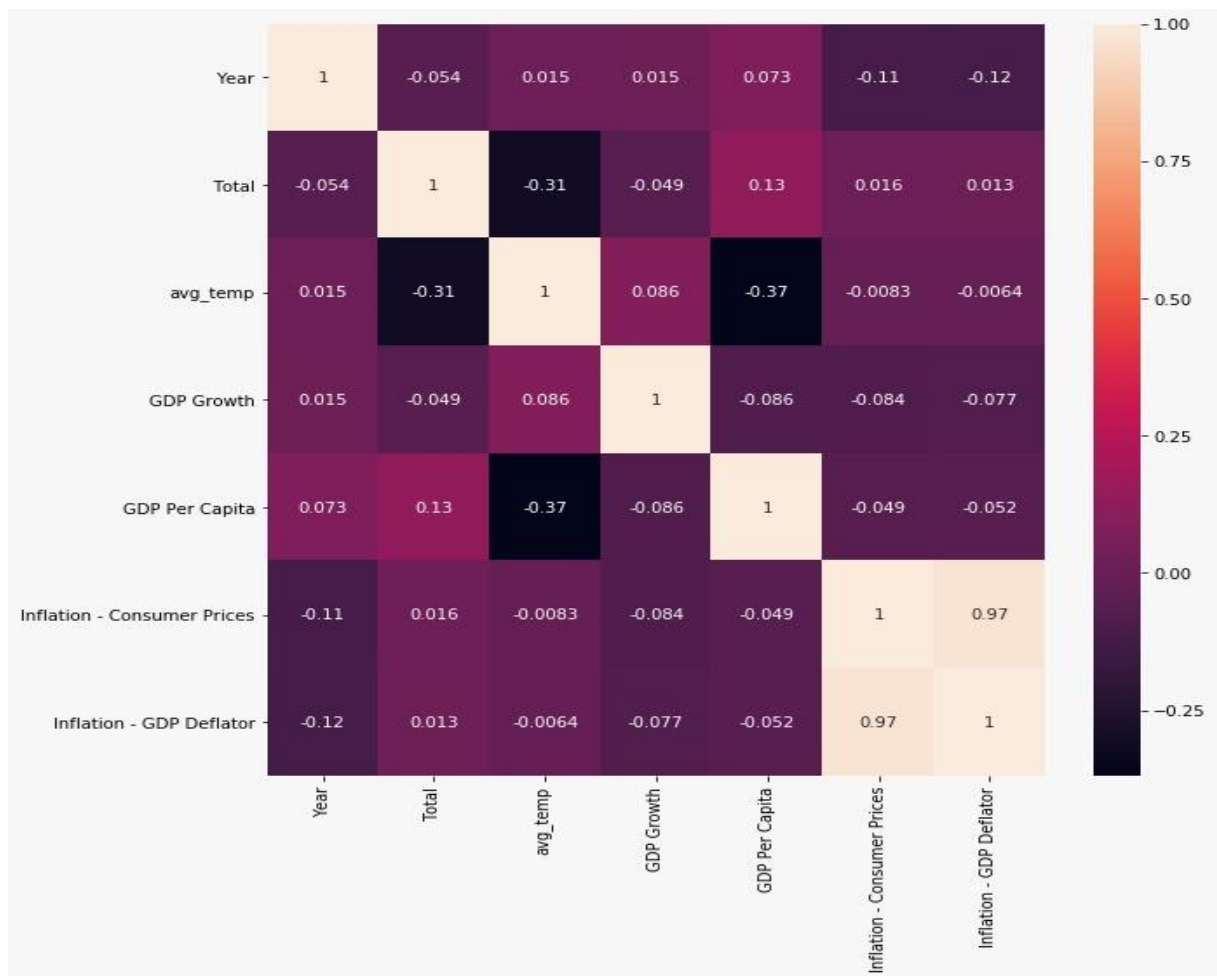


Figura 10: gráfico de correlação 2




## 6. APRESENTAÇÃO DOS RESULTADOS

Os resultados obtidos foram disponibilizados no Tableau onde apresentou-se os resultados e uma descrição dos dados obtidos comparando com a literatura quando possível.

Há uma breve descrição na introdução onde apresenta-se o trabalho.

Figura 11: descrição da página



### Trabalho de Conclusão de Curso

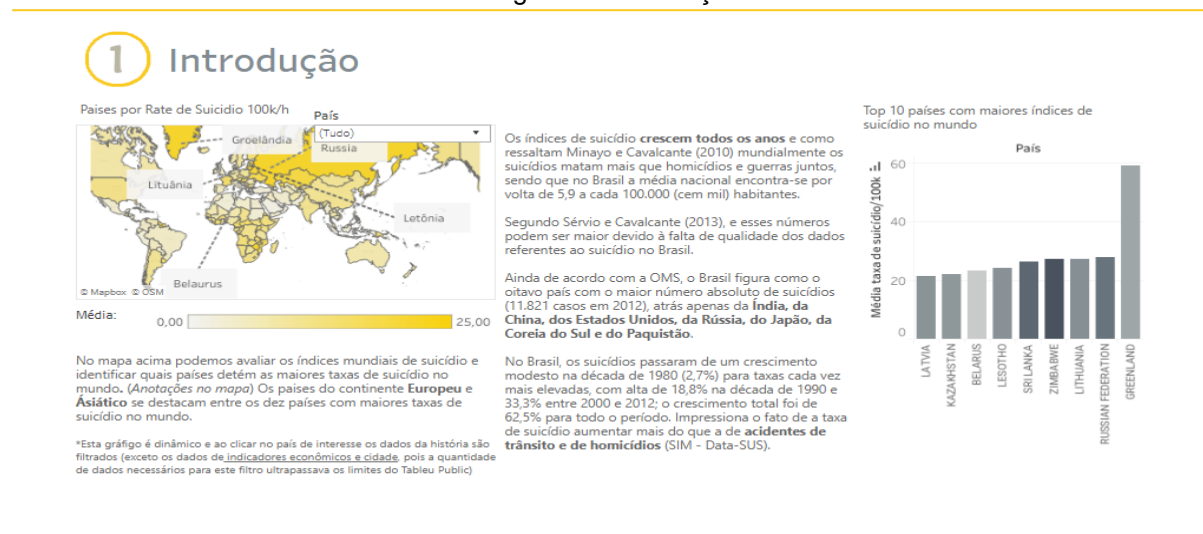
## ANÁLISE DE DADOS MUNDIAIS SOBRE SUICÍDIO

Chama-se suicídio todo caso de morte que resulta direta ou indiretamente de um ato positivo ou negativo praticado pela própria vítima, ato que a vítima sabia dever produzir este resultado. (Durkheim em o suicídio)

Segundo a Organização Mundial de Saúde (OMS), mais de **800 mil pessoas** se matam todos os anos no mundo, sendo que aproximadamente 75% dos casos acontecem em países de baixa e média renda. O suicídio é a segunda principal causa global de morte de pessoas entre **15 e 29 anos**. Apesar da qualidade razoável de dados disponíveis sobre o tema - refém da sensibilidade do assunto, da ilegalidade do ato em certos contextos sociais, da subnotificação e dos erros de classificação (não se raro atribui outras causas de morte a eventos de suicídio) -, fica patente que se trata de um grande problema para a saúde pública atual.

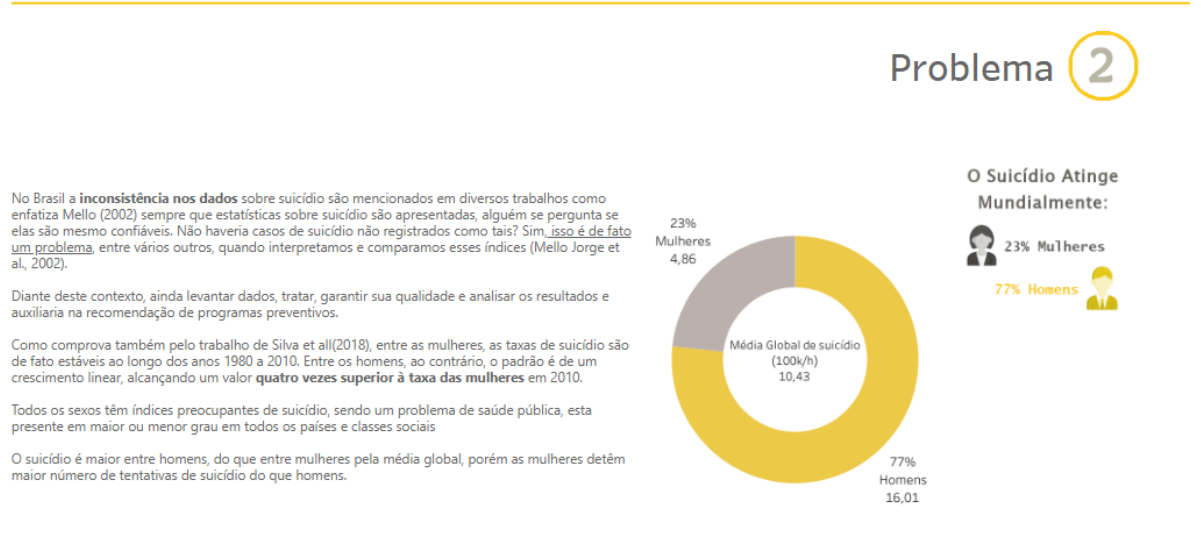
Na imagem a seguir, há uma introdução onde apresenta-se o mapa mundi e os países com maiores taxas de suicídio no mundo.

Figura 12: introdução



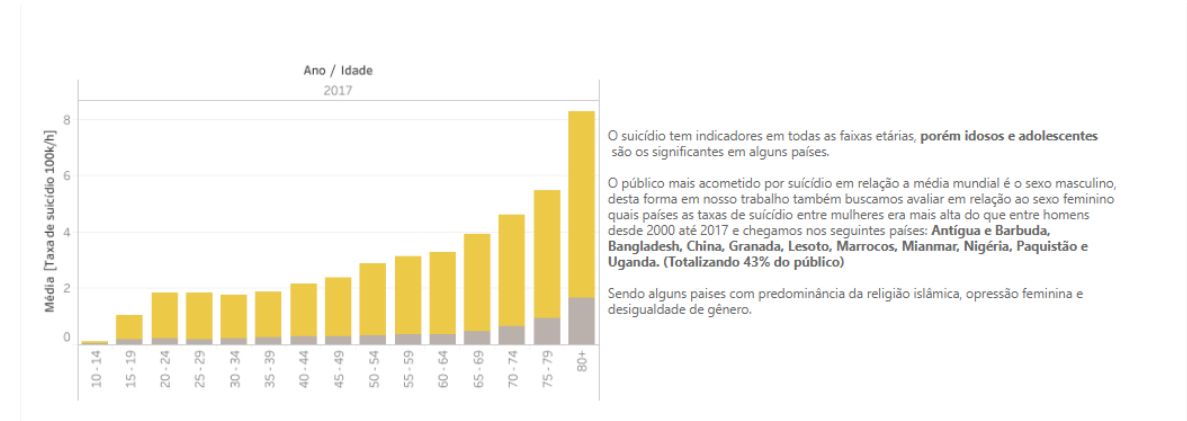
No problema descreveu-se os suicídios em segmentado por sexo, homens destacam-se pelo maior índice de suicídios enquanto as mulheres têm maiores tentativas.

Figura 13: problema



Destacou-se o número de países onde o suicídio masculino é maior que o feminino

Figura 14: suicídio feminino



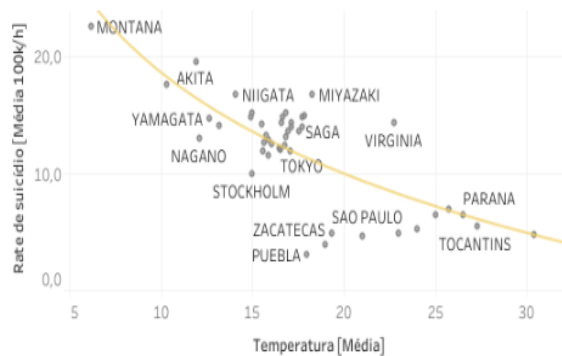
No impacto mundial analisou-se se o clima influenciaria o suicídio, porem notou-se que não houve relação direta.

Figura 15: influência do clima

### 3 Impacto Mundial

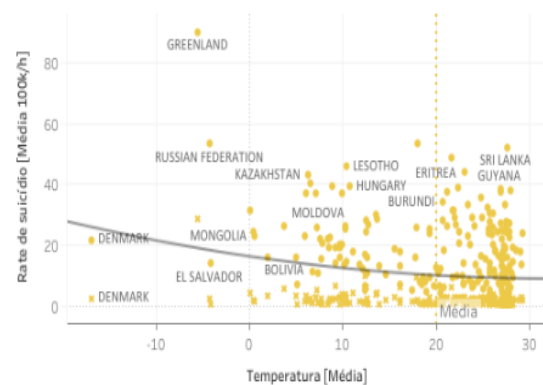
Já foi comprovado que a violência aumenta em situações de extremo calor, porém se mostrou equilibrado em relação ao suicídio. Identificou-se que a relação entre a baixa temperatura está ligada ao crescimento das taxas de suicídio por 100k/h em algumas cidades existentes no dataset estudado, porém para os dados coletados de países não encontramos a mesma relação, mas existe uma tendência mínima.

Rate de Suicídio por Média de Temperatura - Cidade



\*\*Dados apenas dos países: Indonésia, Estados Unidos, Japão, México, Brasil e Reino Unido

Rate de Suicídio por Média de Temperatura - País



O uso de armas de fogo foi analisado a figura 15 onde aborda os países com maiores índices.

Figura 16: países com maiores índices por arma de fogo

Países com maiores índices de suicídio por arma de fogo



Ingestão de pesticidas, enforcamento e armas de fogo estão entre os métodos mais comuns de suicídio em nível global. Os países em destaque têm os maiores índices.

O conhecimento dos métodos de suicídio mais utilizados é importante para a elaboração de estratégias de prevenção que têm se mostrado eficazes, como a restrição de acesso aos meios de suicídio.

Dentre os países analisados, o top 6 países que detém as maiores taxas de suicídio por arma de fogo no mundo são: Groelândia, Geórgia, Estados Unidos, Uruguai, Finlândia e Suíça.

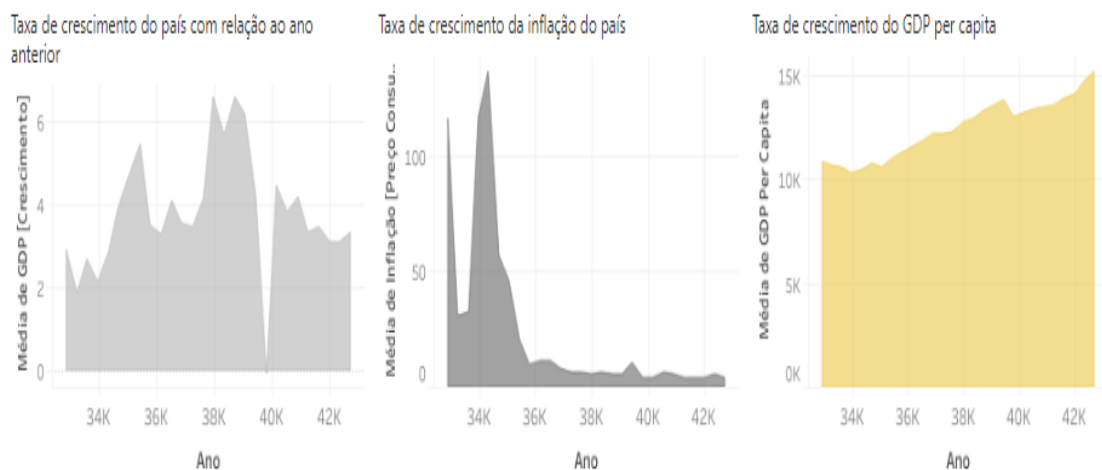
Foi analisado também se o fator econômico influencia ou não o suicídio, os índices foram consistentes em todas as economias.

Figura 17: taxas de crescimento do país com relação ao ano anterior, crescimento da inflação do país e crescimento do GDP per capita

Os resultados obtidos são semelhantes a literatura, porém, pode-se criar um comparativo entre anos anteriores, e os números são assustadores e preocupantes, reforçando que é reforçar a importância deste debate em níveis mundiais.

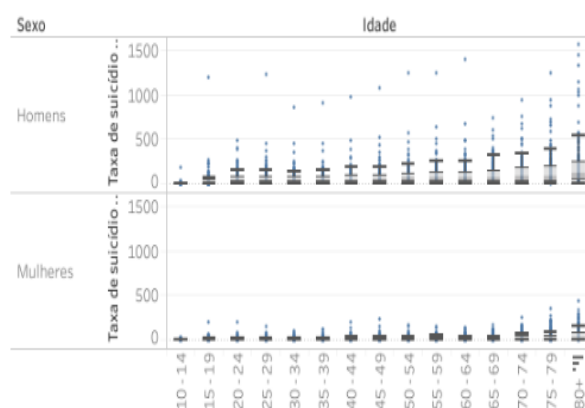
Países com economias estáveis ou não tiveram destaque, porém os índices de suicídio são preocupantes em **todas as economias**.

Crises econômicas tiveram destaque, onde pode-se ter um motivador para quadros depressivos que consequentemente acarreta em suicídio. Como mencionou Fraga (2019) A combinação entre estresse emocional devido à recessão e a problemas financeiros pode levar a quadros suicidas. Alguns grupos sociais e etários são mais atingidos que outros, e a predisposição a determinados transtornos também é um dado relevante. (FRAGA, 2019)



Na análise por idades os jovens e idosos tem níveis preocupantes como segmentado por faixa etária e sexo.

Figura 18: boxplot por faixa etária



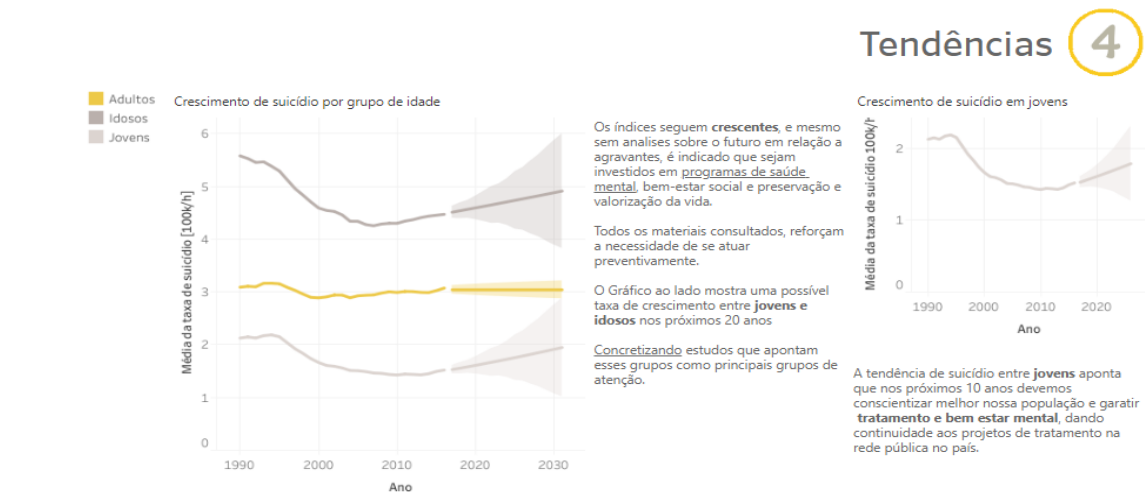
No **boxplot** ao lado podemos avaliar as variações entre as taxas de suicídio por sexo, os homens apresentam diversos **outliers** e que merecem atenção para compreendermos a motivação desta alteração, para o sexo feminino temos uma média abaixo dos homens, porém que também apresentam alguns outliers.

Como por exemplo, entre jovens de 20 à 24 anos, as taxas de suicídio feminino na Groelândia são extremamente altas em relação a média mundial.

As taxas de suicídio dividido em faixas etárias e sexo, foram semelhantes ao mencionados na literatura.

No que se refere ao crescimento por faixa etária, há um número crescente nas tendências para este grupo.

Figura 19: crescimento por faixa de idade e tendências



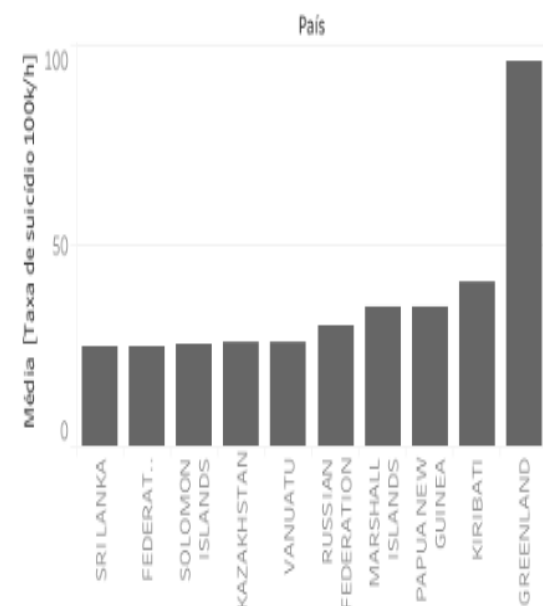
Ainda sobre os jovens, países como Groenlândia, Kiribati, Papua Nova Guiné e Ilhas Marshall tem números preocupantes. Provavelmente o isolamento geográfico possa ser um fator a ser investigado.

Figura 20: taxa de suicídio por países entre jovens

O suicídio entre os jovens é um assunto muito importante e que merece grande atenção, com este objetivo listamos ao lado os países com maior número de suicídio entre os jovens (10 a 25 anos).

Os principais países são: Groenlândia com **taxas altíssimas**, Kiribati, Papua-Nova Guiné e Ilhas Marshall

Taxa de suicídio por países entre jovens





Nas conclusões foram levantadas as recomendações para trabalhos, e a importância de pesquisas na área.

Figura 20: conclusões

## 5 Conclusões

Em nosso trabalho levamos diversos indicadores mundiais em relação do assunto, embora os dados sejam mais ricos quando analisados país individualmente ou mesmo estados ou cidades nosso trabalho tenta comprovar algumas teorias indicadas em trabalhos acadêmicos e em sua principal essência tentamos contribuir mundialmente com os indicadores aqui apresentados.

Com este trabalho pode-se levantar:

- Países com maiores taxas de suicídio mundiais desde 1990, sendo o top 5: **Groelândia, Rússia, Lituânia, Zimbábue e Sri Lanka.**

**Sendo a maioria destes países pertencentes aos continentes Asiático e Africano.**

- Países onde a taxa de suicídio entre mulheres está acima da taxa de suicídio masculino: **Antígua e Barbuda, Bangladesh, China, Granada, Lesoto, Marrocos, Mianmar, Nigéria, Paquistão e Uganda.**

- Países onde a taxa de suicídio feminino é maior que a taxa de suicídio mundial (12,5), sendo o top 10: **Japão, Bielorrússia, Estônia, Cuba, Sri Lanka, Suíça, Letônia, Lituânia, Cazaquistão e África do Sul.**

- Comprovamos o que encontramos nos trabalhos acadêmicos em relação as grupos de faixa idade que merecem atenção, sendo **idosos e jovens.**

- O suicídio entre os **jovens** é um assunto muito importante e que merece grande atenção, com este objetivo listamos também neste trabalho os países com maior número de suicídio entre os jovens (15 a 24 anos).

- As tendências de crescimento da taxas de suicídio por grupo de idade também foram apresentados neste trabalho.

- A classificação de causa de morte mundial no dataset do IHME contém apenas duas variáveis (**Morte por Arma de Fogo** ou **Outras causas**) desta forma apresentamos também os índices mundiais obtidos, mas para enriquecimento deste dado recomendamos a análise de datasets específicos por país, como o **DATASUS** ele contém categorias detalhadas pela causa de morte.

- Com este trabalho podemos considerar que a **temperatura média** do país esta fracamente ligada ao crescimento taxas de suicídio do país, porém ao avaliar este dado à nível de algumas cidade foi possível compreender que há uma relação mais forte e que a aplicação de um modelo machine learning neste nível traria grandes contribuições mundiais e este será o nosso próximo passo.

O nosso trabalho procura também auxiliar a meta definida na **Organização Panamericana de Saúde (OPAS):**

### Meta 3.4:

Até 2030, reduzir em um terço a mortalidade prematura por doenças crônicas não transmissíveis via prevenção e tratamento, e promover a **saúde mental e o bem-estar**

• A probabilidade de morrer por diabetes, câncer, doenças cardiovasculares e doenças crônicas pulmonares entre 30 e 70 anos caiu para 18% em 2016, abaixo dos 22% em 2000. Adultos em países de baixa e baixa-média renda enfrentaram os maiores riscos – quase o dobro da taxa para adultos em países de alta renda. O número total de mortes por doenças crônicas não transmissíveis está aumentando devido ao crescimento populacional e ao envelhecimento.

• Quase **800 mil mortes por suicídio** ocorreram em 2016, com a taxa mais elevada na Região Europeia (15,4 por cada 100.000 habitantes).

Fonte: [https://www.paho.org/bra/index.php?option=com\\_content&view=article&id=5676:organizacao-mundial-da-saude-divulga-novas-estatisticas-mundiais-de-saude&Itemid=843](https://www.paho.org/bra/index.php?option=com_content&view=article&id=5676:organizacao-mundial-da-saude-divulga-novas-estatisticas-mundiais-de-saude&Itemid=843)



### Informação

Última atualização: 17/10/2019

Autores: Camila M. Melo, Cristiane A. Massena e Daniel M. Bicalho  
Pós-graduação Lato Sensu em Ciência de Dados e Big Data (PUC Minas - Virtual)  
A literatura consultada está disponível no trabalho da equipe.

### Referências datasets principais:

WHO: <http://apps.who.int/gho/data/view.sdg.3-4-data-ctry?lang=en>

IHME: <http://ghdx.healthdata.org/gbd-results-tool>

World Bank Open Data: <http://data.worldbank.org>



CAMILA M. MELO - RA: 1077579  
CRISTIANE A. MASSENA - RA: 1143503  
DANIEL M. BICALHO - RA: 1143507



## 7. REFERÊNCIAS

Associação Brasileira de Psiquiatria. **Suicídio: informando para prevenir.** – Brasília: CFM/ABP, 2014.

CICOGNA, J.I.R.; HILLESHEIM, D; HALLAL, A.L.M.C. Mortalidade por suicídio de adolescentes no Brasil: tendência temporal de crescimento entre 2000 e 2015. **J. bras. psiquiatr.**, Rio de Janeiro , v. 68, n. 1, p. 1-7, Mar. 2019. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0047-20852019000100001&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0047-20852019000100001&lng=en&nrm=iso). Acesso em: 08 Jul 2019.

FRAGA, O. Existe alguma ligação entre crise econômica e aumento nas taxas de suicídio? Disponível em: <https://g1.globo.com/mundo/noticia/2019/07/15/existe-alguma-ligacao-entre-crise-economica-e-aumento-nas-taxas-de-suicidio.ghtml>. Acesso em: 08 Jul 2019.

GIRIANELLI, V. R. et al. Qualidade das notificações de violências interpessoal e autoprovocada no Estado do Rio de Janeiro, Brasil, 2009-2016. **Cad. saúde colet.**, Rio de Janeiro , v. 26, n. 3, p. 318-326, July 2018 . Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1414-462X2018000300318&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-462X2018000300318&lng=en&nrm=iso) . Acesso em: 08 Jul 2019.

SANTOS, Erick D. G. M. et al . Suicídio entre idosos no Brasil: uma revisão de literatura dos últimos 10 anos. **Psicol. Conoc. Soc.**, Montevideo , v. 9, n. 1, p. 205-220, jun. 2019 . Disponível em: <[http://www.scielo.edu.uy/scielo.php?script=sci\\_arttext&pid=S1688-70262019000100205&lng=es&nrm=iso](http://www.scielo.edu.uy/scielo.php?script=sci_arttext&pid=S1688-70262019000100205&lng=es&nrm=iso)>. Acesso em: 08 Jul 2019.

IHGE - Global Health Data Exchange. disponível em <http://ghdx.healthdata.org/gbd-results-tool>. Acesso em 20 de junho de 2019.

MINAYO, M. C. S., Cavalcante, F. G., Mangas, R. M. N.,; Souza, J. R. As Autópsias psicológicas sobre suicídio de idosos no Rio de Janeiro. *Ciência & Saúde Coletiva*, 17(10), 2012.

NUNES, Alexandre Morais. Suicídio em Portugal: um retrato do país. **J. bras. psiquiatr.**, Rio de Janeiro, v. 67, n. 1, p. 25-33, Mar. 2018

RIBEIRO, José M; MOREIRA, Marcelo R. Uma abordagem sobre o suicídio de adolescentes e jovens no Brasil. **Ciênc. saúde coletiva**, Rio de Janeiro , v. 23, n. 9, p. 2821-2834, Set. 2018. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-81232018000902821&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232018000902821&lng=en&nrm=iso) . . Acesso em: 08 Jul 2019.

RIBEIRO, N M et al. Análise da tendência temporal do suicídio e de sistemas de informações em saúde em relação às tentativas de suicídio. **Texto contexto - enferm.**, Florianópolis, v. 27, n. 2, e2110016, 2018. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0104-07072018000200310&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-07072018000200310&lng=en&nrm=iso) . Acesso em: 08 Jul 2019.

SERVIO, Selena Mesquita Teixeira; CAVALCANTE, Ana Célia Sousa. Retratos de autópsias Psicossociais sobre suicídio de idosos em Teresina. **Psicol. cienc. prof.**, Brasília , v. 33, n. spe, p. 164-175, 2013 . Disponível em: <[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1414-98932013000500016&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-98932013000500016&lng=en&nrm=iso)>. Acesso em: 08 Jul 2019.

SILVA, B. F.A.; et al . O suicídio no Brasil contemporâneo. **Soc. estado.**, Brasília , v. 33, n. 2, p. 565-579, Aug. 2018 . Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-69922018000200565&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-69922018000200565&lng=en&nrm=iso) . Acesso em: 08 Jul 2019.

TEIXEIRA, S.M.O. O Método de Autópsia Psicossocial como Recurso de Investigação acerca do Suicídio. **Psic.: Teor. e Pesq.**, Brasília , v. 34, e34434, 2018 . Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-37722018000100533&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-37722018000100533&lng=en&nrm=iso) . Acesso em: 08 Jul 2019. .