



Tuberculose

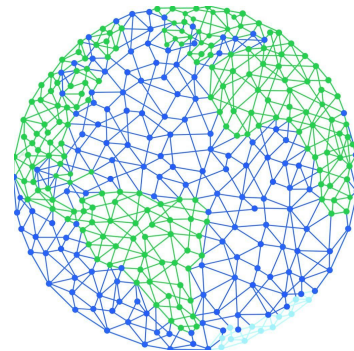
Gabriel Giorisatto De Angelo

Renan Borges Alves

Rebecca Moreira Messias



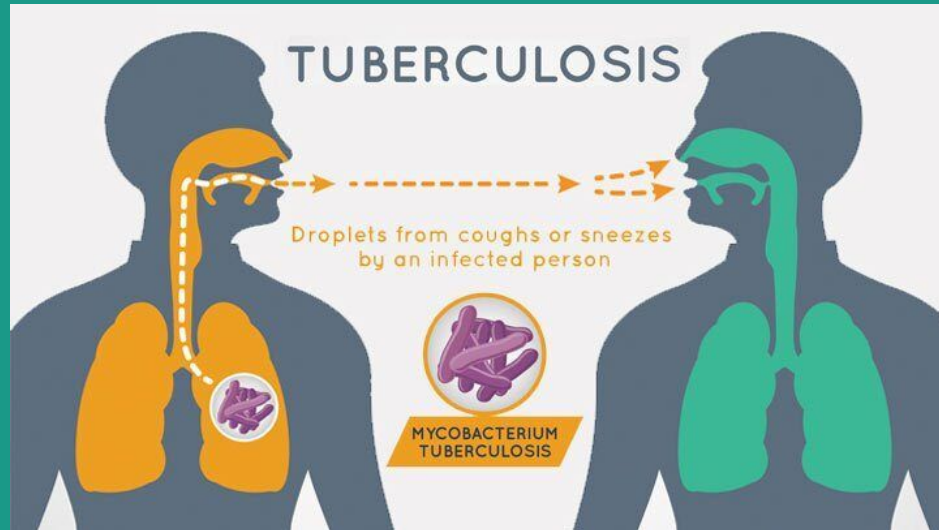
Incidência da Tuberculose



- **Top 10** entre as causas de **morte** no mundo
- Em 2017, a tuberculose causou cerca de **1,3 milhões de mortes**
- Globalmente, a melhor estimativa é que **10 milhões** de pessoas **desenvolveram a doença em 2017**
- Principal maneira de prevenção por **vacinação**

Fonte: https://www.who.int/tb/publications/global_report/en/

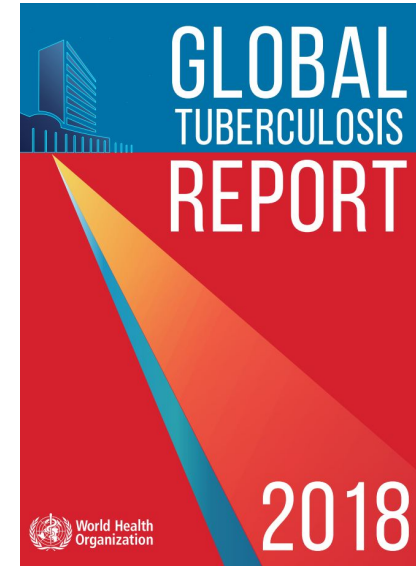
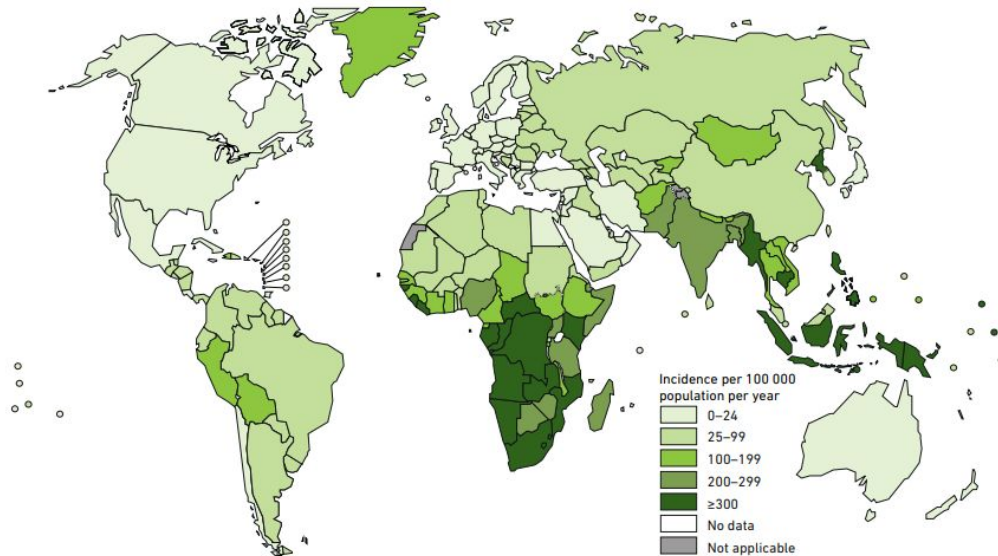
Qual país de prioridade?



World Health Organization

FIG. 3.4

Estimated TB incidence rates, 2017

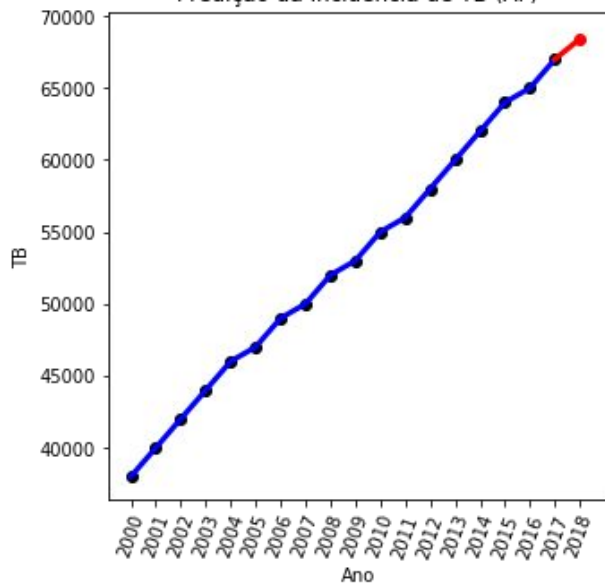


É possível prever as incidências?

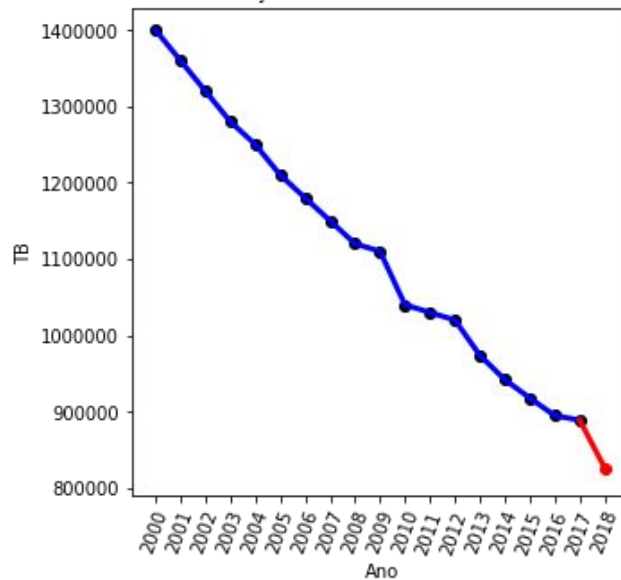


Predição da incidência de TB

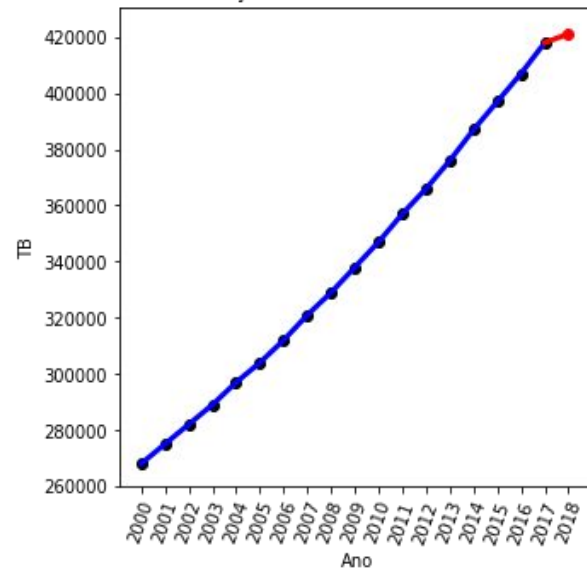
Predição da incidencia de TB (AF)



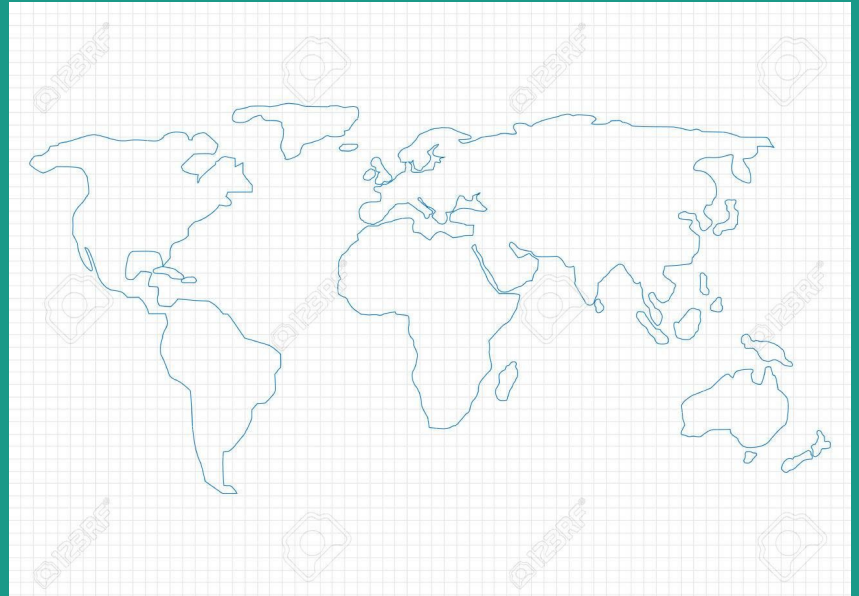
Predição da incidencia de TB (CN)



Predição da incidencia de TB (NG)



Onde coletar os dados geográficos?





GeoNames

The GeoNames geographical database covers all countries and contains over eleven million placenames that are available for download free of charge.

Geonames

Country Codes

| ISO-3166 alpha2 | ISO-3166 alpha3 | ISO-3166 numeric | fips | Country | Capital | Area in km² | Population | Continent |
|-----------------|-----------------|------------------|------|--|---------------------|--------------|-------------|-----------|
| AD | AND | 020 | AN | Andorra | Andorra la Vella | 468.0 | 84,000 | EU |
| AE | ARE | 784 | AE | United Arab Emirates | Abu Dhabi | 82,880.0 | 4,975,593 | AS |
| AF | AFG | 004 | AF | Afghanistan | Kabul | 647,500.0 | 29,121,286 | AS |
| AG | ATG | 028 | AC | Antigua and Barbuda | St. John's | 443.0 | 86,754 | NA |
| AI | AIA | 660 | AV | Anguilla | The Valley | 102.0 | 13,254 | NA |
| AL | ALB | 008 | AL | Albania | Tirana | 28,748.0 | 2,986,952 | EU |
| AM | ARM | 051 | AM | Armenia | Yerevan | 29,800.0 | 2,968,000 | AS |
| AN | ANT | 530 | NT | Netherlands Antilles | Willemstad | 960.0 | 300,000 | NA |
| AO | AGO | 024 | AO | Angola | Luanda | 1,246,700.0 | 13,068,161 | AF |
| AQ | ATA | 010 | AY | Antarctica | | 14,000,000.0 | 0 | AN |
| AR | ARG | 032 | AR | Argentina | Buenos Aires | 2,766,890.0 | 41,343,201 | SA |
| AS | ASM | 016 | AQ | American Samoa | Pago Pago | 199.0 | 57,881 | OC |
| AT | AUT | 040 | AU | Austria | Vienna | 83,858.0 | 8,205,000 | EU |
| AU | AUS | 036 | AS | Australia | Canberra | 7,686,850.0 | 21,515,754 | OC |
| AW | ABW | 533 | AA | Aruba | Oranjestad | 193.0 | 71,566 | NA |
| AX | ALA | 248 | | Åland | Mariehamn | 1,580.0 | 26,711 | EU |
| AZ | AZE | 031 | AJ | Azerbaijan | Baku | 86,600.0 | 8,303,512 | AS |
| BA | BIH | 070 | BK | Bosnia and Herzegovina | Sarajevo | 51,129.0 | 4,590,000 | EU |
| BB | BRB | 052 | BB | Barbados | Bridgetown | 431.0 | 285,653 | NA |
| BD | BGD | 050 | BG | Bangladesh | Dhaka | 144,000.0 | 156,118,464 | AS |
| BE | BEL | 056 | BE | Belgium | Brussels | 30,510.0 | 10,403,000 | EU |
| BF | BFA | 854 | UV | Burkina Faso | Ouagadougou | 274,200.0 | 16,241,811 | AF |
| BG | BGR | 100 | BU | Bulgaria | Sofia | 110,910.0 | 7,000,039 | EU |
| BH | BHR | 048 | BA | Bahrain | Manama | 665.0 | 738,004 | AS |
| BI | BDI | 108 | BY | Burundi | Gitega | 27,830.0 | 9,863,117 | AF |
| BJ | BEN | 204 | BN | Benin | Porto-Novo | 112,620.0 | 9,056,010 | AF |
| BL | BLM | 652 | TB | Saint Barthélemy | Gustavia | 21.0 | 8,450 | NA |
| BM | BMU | 060 | BD | Bermuda | Hamilton | 53.0 | 65,365 | NA |
| BN | BRN | 096 | BX | Brunei | Bandar Seri Begawan | 5,770.0 | 395,027 | AS |
| BO | BOL | 068 | BL | Bolivia | Sucre | 1,098,580.0 | 9,947,418 | SA |
| BQ | BES | 535 | | Bonaire | | 328.0 | 18,012 | NA |
| BR | BRA | 076 | BR | Brazil | Brasilia | 8,511,965.0 | 201,103,330 | SA |
| BS | BHS | 044 | BF | Bahamas | Nassau | 13,940.0 | 301,790 | NA |

Andorra 🌐

country name : [Andorra](#) [\[other languages\]](#)

iso code : AD, 020 and AND

fips code : AN

capital : Andorra la Vella

area : 468.0 km²


population : 84,000

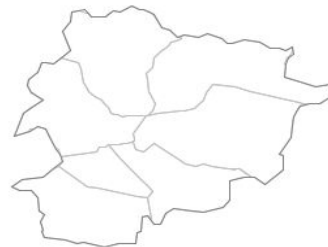
currency : Euro (EUR)

languages : Catalan (ca)

neighbours : [Spain](#), [France](#)

postal code format : AD### [\[postal codes\]](#)

national flag : 



Modelagem dos dados

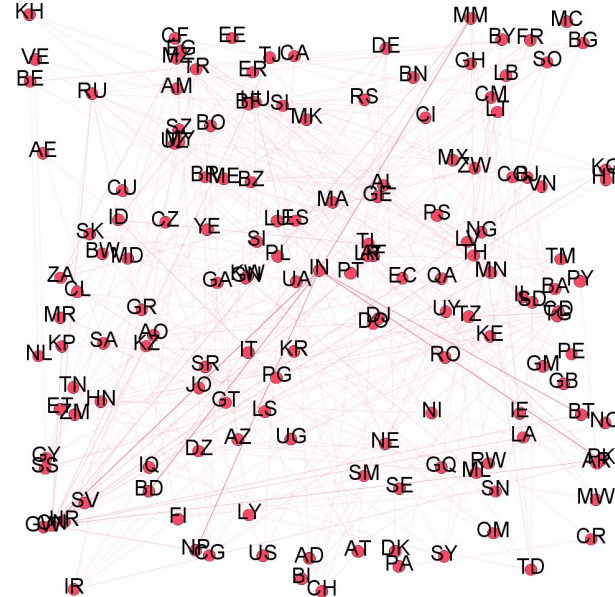
Gp Gephi

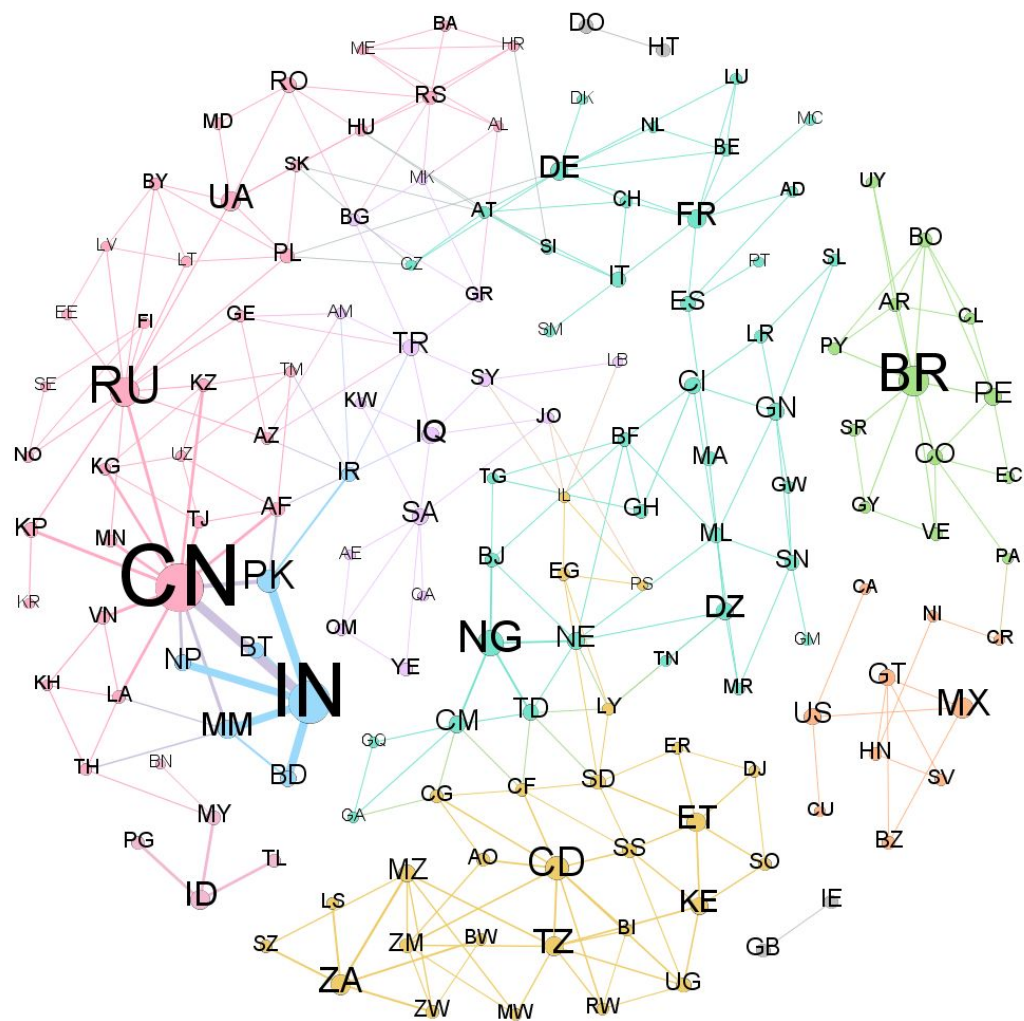
Modelagem dos dados (Gephi)

Comma ▾ Edges table ▾ UTF-8

Preview:

| Source | Target | Weight | Type |
|--------|--------|--------|------------|
| AF | CN | 895097 | Undirected |
| AF | IR | 80541 | Undirected |
| AF | PK | 606319 | Undirected |
| AF | TJ | 74628 | Undirected |
| AF | TM | 71013 | Undirected |
| AF | UZ | 93090 | Undirected |
| AL | GR | 921 | Undirected |
| AL | ME | 528 | Undirected |





Análise estatística



Análise Estatística

- Analisar em qual parcela da população a incidência de tuberculose sempre esteve maior.
- Dividir a população por gênero e por faixa etária
- Dados de 1980 até 2017
- Fonte de dados: [World Health Organization](#) (csv disponível no [github](#))

Correlação dos dados

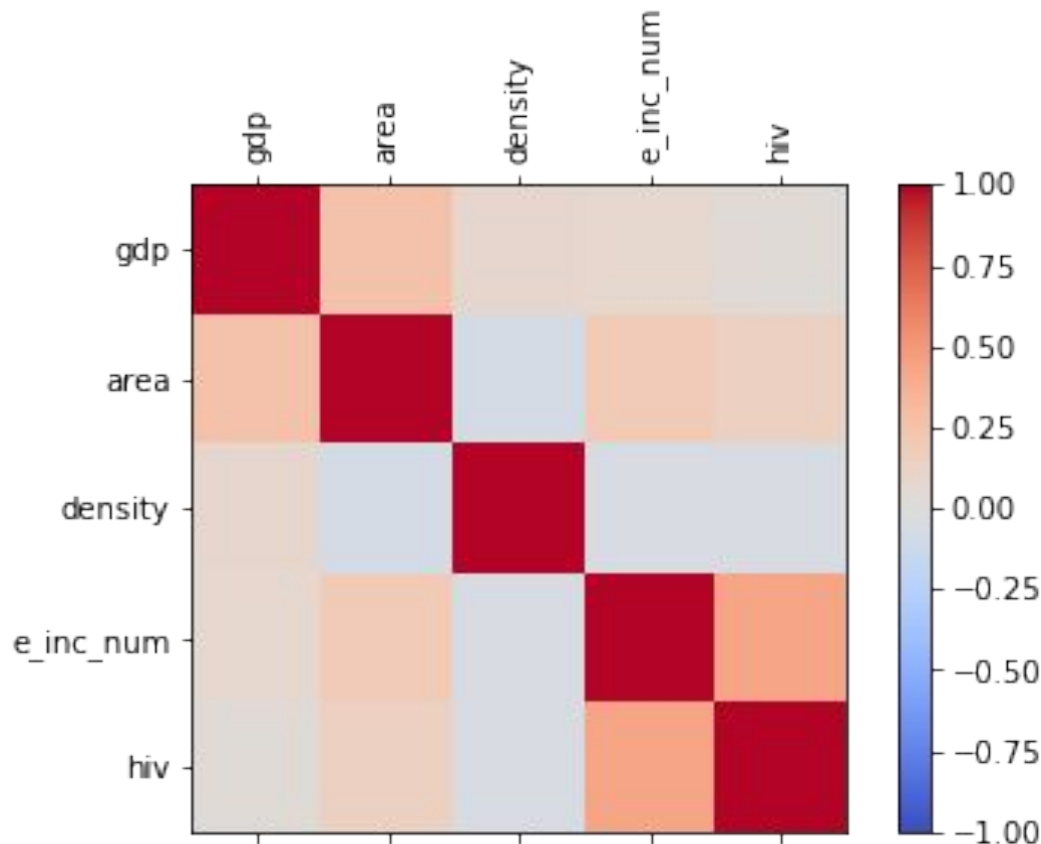
gdp: PIB

area: Território geográfico

density: Densidade demográfica

e_inc_num: Incidência de TB

hiv: incidência HIV





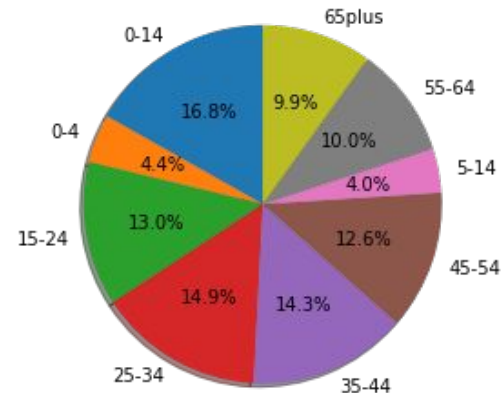
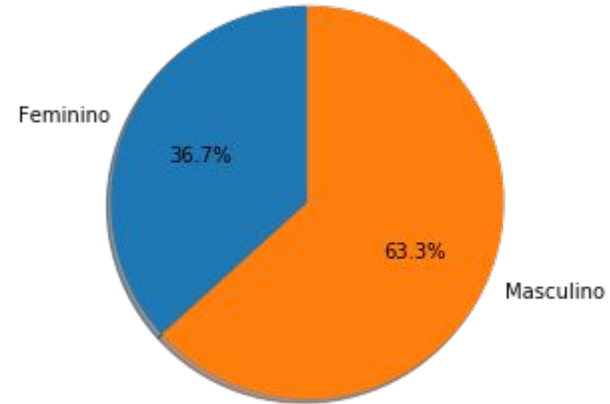
Análise Estatística

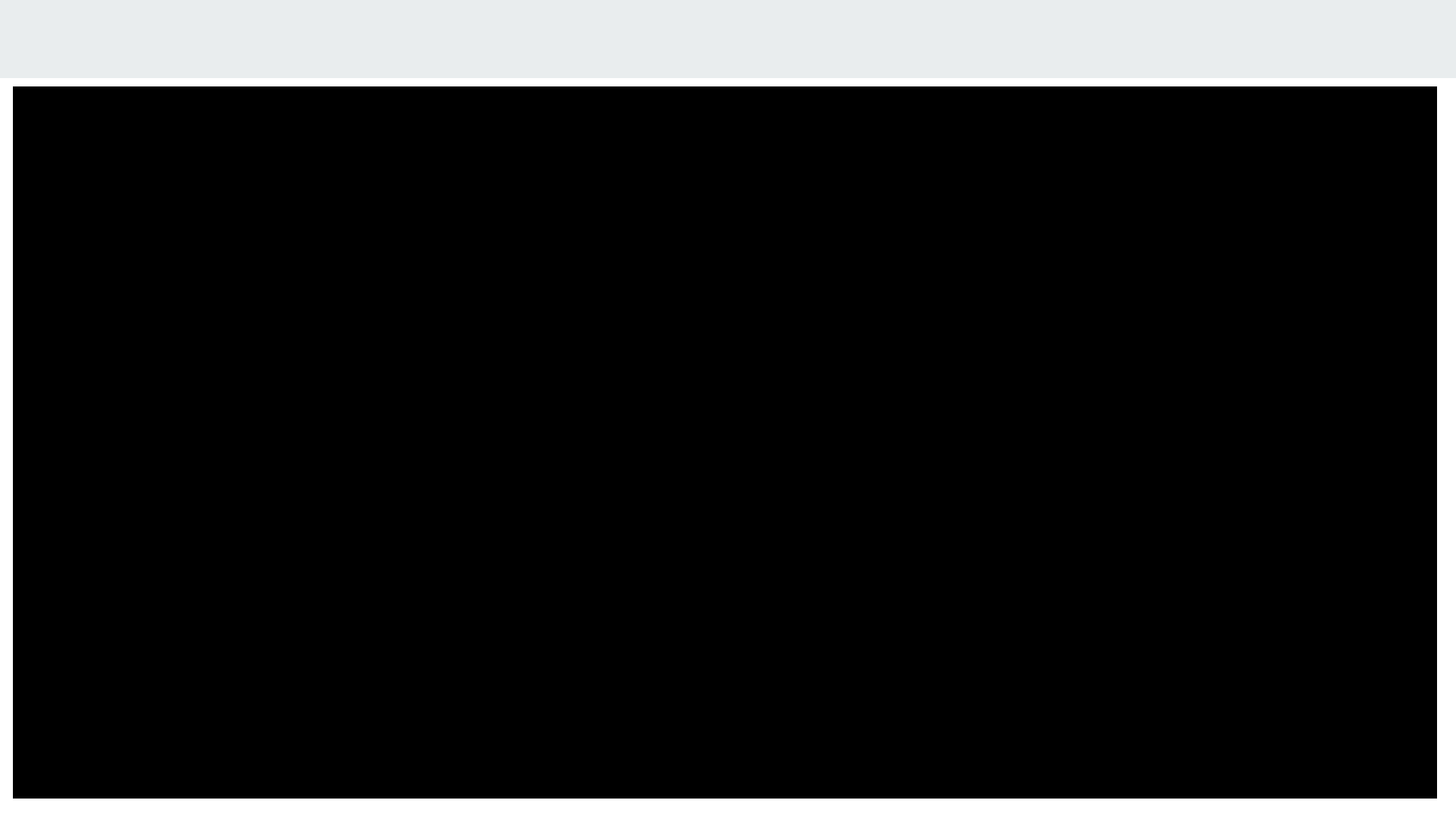
Foi utilizado o arquivo CSV com a incidência mundial da tuberculose com relação ao grupo de idade e sexo

| country | iso2 | iso3 | iso_num | year | measure | unit | age_group | sex | best | lo | hi |
|-------------|------|------|---------|------|---------|------|-----------|-----|-------|-------|-------|
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 0-14 | a | 7500 | 4300 | 11000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 0-14 | f | 3600 | 1400 | 5800 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 0-14 | m | 4000 | 1500 | 6400 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 0-4 | f | 1800 | 0 | 4400 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 0-4 | m | 2200 | 0 | 5400 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 15-24 | f | 7100 | 0 | 17000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 15-24 | m | 5900 | 0 | 15000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 15plus | a | 60000 | 34000 | 86000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 15plus | f | 31000 | 12000 | 50000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 15plus | m | 29000 | 11000 | 47000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 25-34 | f | 6300 | 0 | 16000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 25-34 | m | 5700 | 0 | 14000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 35-44 | f | 5200 | 0 | 13000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 35-44 | m | 5300 | 0 | 13000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 45-54 | f | 4600 | 0 | 11000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 45-54 | m | 4200 | 0 | 10000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 14-May | f | 1800 | 0 | 4400 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 14-May | m | 1800 | 0 | 4300 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 55-64 | f | 4000 | 0 | 9800 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 55-64 | m | 3600 | 0 | 8900 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 65plus | f | 3400 | 0 | 8400 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | 65plus | m | 4100 | 0 | 10000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | all | a | 67000 | 41000 | 93000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | all | f | 34000 | 15000 | 53000 |
| Afghanistan | AF | AFG | 4 | 2017 | inc | num | all | m | 33000 | 15000 | 51000 |
| Albania | AL | ALB | 8 | 2017 | inc | num | 0-14 | a | 45 | 37 | 52 |
| Albania | AL | ALB | 8 | 2017 | inc | num | 0-14 | f | 22 | 17 | 26 |
| Albania | AL | ALB | 8 | 2017 | inc | num | 0-14 | m | 23 | 18 | 28 |
| Albania | AL | ALB | 8 | 2017 | inc | num | 15plus | a | 530 | 450 | 620 |
| Albania | AL | ALB | 8 | 2017 | inc | num | 15plus | f | 180 | 140 | 220 |
| Albania | AL | ALB | 8 | 2017 | inc | num | 15plus | m | 350 | 280 | 430 |
| Albania | AL | ALB | 8 | 2017 | inc | num | all | a | 580 | 490 | 670 |
| Albania | AL | ALB | 8 | 2017 | inc | num | all | f | 200 | 160 | 240 |
| Albania | AL | ALB | 8 | 2017 | inc | num | all | m | 380 | 300 | 450 |
| Algeria | DZ | DZA | 12 | 2017 | inc | num | 0-14 | a | 3900 | 2800 | 5100 |
| Algeria | DZ | DZA | 12 | 2017 | inc | num | 0-14 | f | 1900 | 1100 | 2600 |
| Algeria | DZ | DZA | 12 | 2017 | inc | num | 0-14 | m | 2100 | 1300 | 2900 |
| Algeria | DZ | DZA | 12 | 2017 | inc | num | 0-4 | f | 1000 | 120 | 1900 |
| Algeria | DZ | DZA | 12 | 2017 | inc | num | 0-4 | m | 1200 | 150 | 2300 |

Análise Estatística

Alguns resultados preliminares obtidos dizem respeito a proporção das incidências com relação ao sexo e idade





Obrigado!

<https://github.com/RebeccaMM/data-science-eHealthy>
