

Covid Tracking

Relatório BDAD 2020/2021

G201

André de Jesus Fernandes Flores – up201907001  
Diogo Luís Araujo de Faria – up201907014  
Tiago Andre Batista Rodrigues – up201906807

Contents

[Introdução 3](#_Toc67992698)

[Classes 4](#_Toc67992699)

[Location 4](#_Toc67992700)

[Country 4](#_Toc67992701)

[District 4](#_Toc67992702)

[County 4](#_Toc67992703)

[Parish 4](#_Toc67992704)

[Vaccine 5](#_Toc67992705)

[VaccinationAmount 5](#_Toc67992706)

[Manufacturer 5](#_Toc67992707)

[Strain 5](#_Toc67992708)

[Case 5](#_Toc67992709)

[EmploymentSector 6](#_Toc67992710)

[Ethnicity 6](#_Toc67992711)

[Outcome 6](#_Toc67992712)

[NursingHome 6](#_Toc67992713)

[Hospitalization 6](#_Toc67992714)

[ICUStay 7](#_Toc67992715)

[Ventilation 7](#_Toc67992716)

[Hospital 7](#_Toc67992717)

[Modelo Conceptual 8](#_Toc67992718)

[Esquema Relacional 9](#_Toc67992719)

[Análise a Forma Normal 10](#_Toc67992720)

[Vaccine 10](#_Toc67992721)

[EmploymentSector 10](#_Toc67992722)

[Hospitalization 11](#_Toc67992723)

[EthnicityCOVIDCase e EmployedIn 11](#_Toc67992724)

[Geral 11](#_Toc67992725)

# Introdução

A base de dados visa guardar informação relativa aos casos de COVID-19 num dado país de modo a ser possível analisar a evolução do estado da pandemia ao longo do tempo. Para a esta análise ser o mais detalhada possível a informação geográfica dos casos será dividida por várias unidades administrativas sendo a mais pequena a freguesia. Também se intende providenciar contexto a cada caso de COVID-19 registado através de outra informação relevante.

Uma grande inspiração para este projeto é o [COVID Tracking Project](https://covidtracking.com).

# Classes

## Location

A classe “Location” serve de generalização para os elementos que apresentem como atributos comuns um identificador locID, um name, /caseNumber e /noVaccinated. É uma classe completa e disjunta.

### Country

Esta classe representa um país, que possui um atributo a mais, o /population que representa a população do país, derivado a partir da população de cada distrito, “**District**”.

Cada país tem como suas partes pelo menos um distrito, “**District**”, e pode ser o país de origem de alguma variante de SARS-CoV-2, guardada na classe “**Strain**”.

### District

Esta classe representa um distrito, que possui um atributo a mais, o /population que representa a população do distrito, derivado a partir da população de cada “**County**”.

Cada distrito faz parte de um e um só país, “**Country**”, sendo que tem como suas partes pelo menos um concelho, “**County**”.

### County

Esta classe representa um concelho, que possui um atributo a mais, o /population que representa a população do concelho, derivado a partir da população de cada freguesia, “**Parish**”.

Cada concelho faz parte de um e um só distrito, “**District**”, sendo que tem como suas partes pelo menos uma freguesia, “**Parish**”.

### Parish

Esta classe representa uma freguesia, que possui um atributo a mais, o population que representa a sua população.

Cada freguesia faz parte de um e um só concelho, “**County**”. Associada a cada freguesia pode haver inúmeros casos, “**Case**”, e também inúmeras vacinas administradas, “**Vaccine**”, numa relação que apresenta uma classe de associação, “**VacciantionAmount**”. Em cada freguesia podem não existir ou existir inúmeros hospitais, “**Hospital**”.

## Vaccine

Esta classe representa uma vacina e tem como atributos um identificador vacID, um name, que representa a designação da vacina utilizado, e um doseNumber, que representa o número de doses necessárias.

Cada vacina específica pode existir em várias ou inúmeras freguesias, “**Parish**”, numa relação em que apresenta uma classe de associação, “**VaccinationAmount**”. Cada vacina pode ter sido criada apenas por um ou por uma colaboração de vários fabricantes, “**Manufacturer**”.

## VaccinationAmount

Esta classe é uma classe de associação entre freguesia, “**Parish**”, e vacina, “**Vaccine**”, que representa o número de vacinas de um certo tipo presentas numa certa freguesia e que tem como atributos vaccinationNumber e date, que representa a data em que se verificou a quantidade de vacinas existentes de forma a se poder estudar a variação da mesma em relação aos casos, “**Case**”.

## Manufacturer

Esta classe representa um fabricante de uma determinada vacina e tem como atributos um identificador manId e um name, que representa o nome do fabricante.

Como na base de dados se está apenas a considerar fabricantes que realmente já desenvolveram uma vacina e, pelos dados que obtivemos, considera-se que um fabricante vai estar ligado uma e uma só vez a uma vacina, “**Vaccine**”.

## Strain

Esta classe representa cada variante do vírus SARS-CoV-2, tendo como atributos um identificador strainID e designation.

Cada variante é originada em um e um só país, “**Country**”, existindo então uma ligação com a classe mencionada e também pode estar associada a inúmeros casos, “**Case**”.

## Case

Esta classe representa um caso, tendo como atributos um identificador caseID, uma detectionDate e endDate, que representam a data de deteção do case e de fim, respetivamente, e birthYear, que representa o ano de nascimento da pessoa que foi infetada.

Cada caso vai representar alguém infetado com uma e uma só variante do vírus, “**Strain**”, que foi detetado numa só freguesia, “**Parish**”, que representa alguém de uma ou mais etnias, “**Ethnicity**”, e que podem ter inúmeros empregos ou nenhum em certos setores, “**EmploymentSector**” ou estar num lar, “**NursingHome**”. Em cada caso pode ou não acontecer uma e uma só hospitalização, “**Hospitalization**”, visto que de acordo com os dados que obtemos, uma pessoa apenas sai de um hospital apenas se já foi curada ou se faleceu.

## EmploymentSector

Esta classe representa o setor de emprego em que uma pessoa representada por um caso, “**Case**”, trabalha e tem como atributos um identificador sectorId, um name, que representa o nome do setor e contactFrequency, que é dada por um número entre 0 e 5 e que representa a frequência de contacto entre pessoas em diferentes setores.

Cada setor pode estar ligado a nenhum ou inúmeros casos, “**Case**”.

## Ethnicity

Esta classe representa a etnia de alguém representado num caso, “**Case**”, e tem como atributos um identificador, etniID, e um name, que representa o nome da etnia.

Cada etnia diferente pode estar ligada a nenhum ou inúmeros casos diferentes, “**Case**”.

## Outcome

Esta classe representa o resultado de uma infeção do vírus SARS-CoV-2 e tem como atributo um type, que representa o tipo de resultado, sendo que pode ter 2 valores: 0, que significa que o caso, “**Case**”, a que está relacionada acabou em morte com causa de COVID-19, 1, que significa que houve um óbito não relacionado a COVID-19, ou 2, que significa que houve recuperação.

Visto que vários casos, “**Case**”, podem e vão ter apenas um de dois resultados, esta classe está pode estar ligada a nenhum ou inúmeros casos, “**Case**”.

## NursingHome

Esta classe representa um lar e tem como atributos um identificador nursingHomeID e um name, que representa o nome do lar.

Cada lar pode estar associado a nenhum ou inúmeros casos, “**Case**”.

## Hospitalization

Esta classe representa uma hospitalização e tem como atributos um identificador hospStayID, e uma startDate e endDate, que representam a data de início e fim da hospitalização, respetivamente.

Em cada hospitalização podem acontecer ou não inúmeras vezes a necessidade de se estar na unidade de cuidados intensivos, “**ICUStay**”. Uma hospitalização vai estar associada a um ou mais hospitais, “**Hospital**”, visto que existe a possibilidade de um doente ser reencaminhado para diferentes hospitais em certas situações.

## ICUStay

Esta classe representa estar numa unidade de cuidados intensivos e tem como atributos um identificador ICUStayID, e uma startDate e endDate, que representam a data de início e fim em que se está na unidade, respetivamente.

Cada visita à ICU vai ser associada a uma e uma só hospitalização, “**Hospitalization**”, e a zero ou inúmeras vezes à necessidade de ser utilizado um ventilador, “**Ventilation**”.

## Ventilation

Esta classe representa o uso de um ventilador e tem como atributos um identificador ventID, e uma startDate e endDate, que representam a data em que se começou e se acabou de usar um ventilador, respetivamente.

Cada uso de ventilador vai estar associado a uma e uma só estadia numa unidade de cuidados intensivos, “**ICUStay**”, visto que apenas se utilizam ventiladores quando se está na unidade.

## Hospital

Esta classe representa um hospital e tem como atributos um identificador hospitalID e um name, que representa o nome do hospital.

Cada hospital está localizado numa e numa só freguesia, “**Parish**”.

# Modelo Conceptual

Diagram, engineering drawing

Description automatically generated

# Esquema Relacional

**Country**(locID, name, caseNumber, noVaccinated, population)

* locID -> name, caseNumber, noVaccinated, population

**District**(locID, name, caseNumber, noVaccinated, population, countryID -> Country)

* locID -> name, caseNumber, noVaccinated, population, countryID

**County**(locID, name, caseNumber, noVaccinated, population, districtID -> District)

* locID -> name, caseNumber, noVaccinated, population, districtID

**Parish**(locID, name, caseNumber, noVaccinated, population, countyID -> County)

* locID -> name, caseNumber, noVaccinated, population, countyID

**Vaccine**(vacID, name, doseNumber)

* vacID -> name, doseNumber
* name->doseNumber

**VaccinationAmount**(locID -> Parish, vacID -> Vaccine, vaccinationNumber)

* locID, vacID -> vaccinationNumbe

**Manufacturer**(manID, name)

* manID -> name

**VaccineManufacturer**(vacID -> Vaccine, manID -> Manufacturer)

**Strain**(strainID, designation, countryID -> Country)

* strainID -> designation, countryID

**Ethnicity**(ethniID, name)

* ethniID -> name

**EthnicityCOVIDCase**(caseID -> COVIDCase, ethniID -> ethnicity)

**NursingHome**(nursingHomeID, name)

* nursingHomeID -> name

**COVIDCase**(caseID, detectionDate, endDate, birthyear, outcome, parishID -> Parish, strainID -> Strain, nursingHomeID -> NursingHome)

* caseID -> detectionDate, endDate, birthyear, outcome, parishID, strainID, nursingHomeID

**EmploymentSector**(sectorID, name, contactFrequency)

* sectorID -> name, contactFrequency
* name -> contactFrequency

**EmployedIn**(caseID -> COVIDCase, sectorID -> EmplymentSector)

**Hospitalization**(hospStayID, startDate, endDate, hospitalID -> Hospital, caseID->COVIDCase)

* hospStayID -> startDate, endDate, hospitalID, caseID
* caseID-> hospStayID, startDate, endDate, hospitalID

**ICUStay**(ICUStayID, startDate, endDate, hospStayID -> Hospitalization)

* ICUStayID -> startDate, endDate, hospStayID

**Ventilation**(ventID, startDate, endDate, ICUStayID -> ICUStay)

* ventID -> startDate, endDate, ICUStayID

**Hospital**(hospitalID, name, parishID -> Parish)

* hospitalID -> name, parishID

## Análise a Forma Normal

### Vaccine

Na relação Vaccine, em name -> doseNumber, devido ao facto de que o atributo do lado esquerdo (name) nem é uma chave nem superchave, pode-se considerar uma violação BCNF e, também, devido a que o atributo do lado direito (doseNumber) não é um atributo primo, pode-se considerar uma violação à 3NF.

Uma possível solução, seria decompor a relação em:

* VaccineA(name, doseNumber), com name -> doseNumber
* VaccineB(vacID, name), com vacID -> name

Estas, já se encontram na BCNF, e por conseguência na 3NF, visto que apenas apresentam uma relação não trivial em que do lado esquerdo está a sua chave primária.

No entanto, neste contexto, não se utilizam as relações decompostas, visto que vacID é um atributo utilizado apenas de forma a simplificar a identificação da relação, sendo mais simples do que uma string, que neste contexto pode ser longa, e permitindo uma menor existência de erros no povoamento.

### EmploymentSector

Na relação EmploymentSector, em name -> contactFrequency, pelas mesmas razões apresentadas para a relação anterior, pode-se considerar uma violação BCNF e 3NF.

Uma possível solução, seria a decomposição em:

* EmploymentSectorA(name, contactFrequency), com name -> contactFrequency
* EmploymentSectorB(sectorID, name), com sectorID -> name

Tal como apresentado anteriormente, estas relações encontram-se na BCNF, e por sua vez na 3NF, sendo que apenas apresentam uma relação não trivial com uma chave primária do lado esquerdo.

Também, como antes referido, não se utilizam estas relações decompostas devido a razões de simplicidade e de diminuição de probabilidade de erros, visto que o atributo name é uma string, relativamente grande e com caracteres maiúsculos, minúsculos e com acentos, aumentando a possibilidade de erros em povoamento.

### Hospitalization

Na relação Hospitalization, apresentam-se duas dependências funcionais, em que uma no lado esquerda tem uma chave primária e outra um superchave, podendo então se considerar a relação como estando em BCNF e, por consequência, em 3NF.

### EthnicityCOVIDCase e EmployedIn

Estas relações não apresentam nenhuma dependência funcional não trivial, pelo que podem ser consideradas com estando em BCNF e, por sua vez, em 3NF.

### Geral

As relações restantes apresentam apenas uma dependência funcional, em que do lado esquerdo está sua chave primário, podendo, assim, ser consideras como estando em BCNF e, por consequência, em 3NF.