

Caso Estudio

Manejo de datos del covid

ISCH

1 Preambulo

1.1 Objetivo

A partir de los datos de incidencia de COVID-19 en España (2020-2022) proporcionados por el área de vigilancia epidemiológica del CNE (<https://cnecovid.isciii.es/covid19>), se pretende describir la distribución de la epidemia entre grupos de edad y su evolución en el tiempo.

1.2 Importación de los datos

```
ruta.covid = "https://cnecovid.isciii.es/covid19/resources/casos_hosp_uci_def_sexo_edad_provres.csv"
# covid = read.table(ruta.covid, header=TRUE, sep=",") # con paquete base
require(data.table)
covid=fread(ruta.covid)
covid
```

```
##      provincia_iso sexo grupo_edad      fecha num_casos num_hosp num_uci
##      1:           A   H         0-9 2020-01-01         0         0         0
##      2:           A   H        10-19 2020-01-01         0         0         0
##      3:           A   H        20-29 2020-01-01         0         0         0
##      4:           A   H        30-39 2020-01-01         0         0         0
##      5:           A   H        40-49 2020-01-01         0         0         0
##      ---
## 1299026:          ZA   NC        50-59 2022-03-27         0         0         0
## 1299027:          ZA   NC        60-69 2022-03-27         0         0         0
## 1299028:          ZA   NC        70-79 2022-03-27         0         0         0
## 1299029:          ZA   NC         80+ 2022-03-27         0         0         0
## 1299030:          ZA   NC          NC 2022-03-27         0         0         0
##      num_def
##      1:      0
##      2:      0
##      3:      0
##      4:      0
##      5:      0
##      ---
## 1299026:      0
## 1299027:      0
## 1299028:      0
## 1299029:      0
## 1299030:      0
```

2 Formateo de la base

2.1 Creación de una variable temporal mensual

Redondear las fechas a la semana (ver `?round.IDate`).

2.2 Cambiar nombre de las variables

Eliminar el prefijo `num_` de las variables de conteo.

2.3 Formato alargado de la base

Reestructurar la base de manera que la base tenga una fila por cada conteo :

```
##      provincia_iso sexo grupo_edad      fecha      periodo variable value
##      1:           A   H         0-9 2020-01-01 2020-01-01     casos      0
##      2:           A   H        10-19 2020-01-01 2020-01-01     casos      0
##      3:           A   H        20-29 2020-01-01 2020-01-01     casos      0
##      4:           A   H        30-39 2020-01-01 2020-01-01     casos      0
##      5:           A   H        40-49 2020-01-01 2020-01-01     casos      0
##      ---
## 5196116:         ZA   NC        50-59 2022-03-27 2022-03-26       def      0
## 5196117:         ZA   NC        60-69 2022-03-27 2022-03-26       def      0
## 5196118:         ZA   NC        70-79 2022-03-27 2022-03-26       def      0
## 5196119:         ZA   NC          80+ 2022-03-27 2022-03-26       def      0
## 5196120:         ZA   NC          NC 2022-03-27 2022-03-26       def      0
```

3 Resumen de la base de datos

3.1 Resumen por grupos de edad y gravedad

A partir de la base anterior, crear una tabla de frecuencias con el grupo de edad en filas y la gravedad de la enfermedad en columnas:

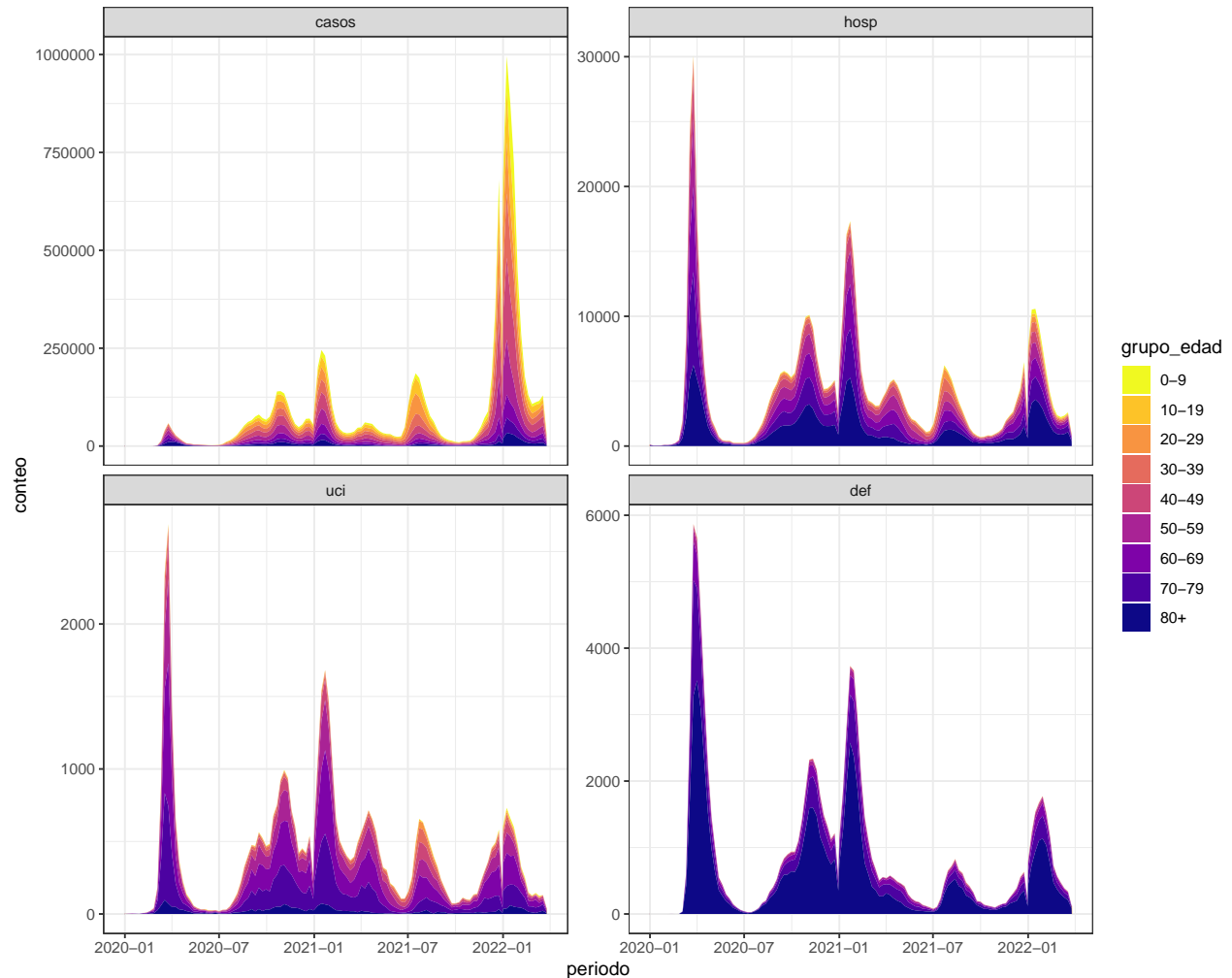
```
##      grupo_edad  casos  hosp  uci  def
## 1:      0-9 1069691  5949  307  24
## 2:     10-19 1534196  5000  314  27
## 3:     20-29 1636255 15466  887 139
## 4:     30-39 1774718 30422 2345 352
## 5:     40-49 2102105 52580 5242 1168
## 6:     50-59 1503049 76392 9983 4013
## 7:     60-69  870345 91550 15273 10528
## 8:     70-79 550944 98057 12544 22293
## 9:      80+ 533630 142980 2479 65366
## 10:      NC   9886  1586  164  207
```

3.2 Resumen por grupos de edad, gravedad y periodo

Agregar la base de datos por grupo de edad (quitar registros sin información sobre la edad), periodo (mes) y gravedad.

```
##      grupo_edad      periodo variable conteo
## 1:      0-9 2020-01-01     casos      0
## 2:     10-19 2020-01-01     casos      0
## 3:     20-29 2020-01-01     casos      0
## 4:     30-39 2020-01-01     casos      0
```

```
##      5:      40-49 2020-01-01      casos      0
##      ---
## 4280:      40-49 2022-03-26      def      1
## 4281:      50-59 2022-03-26      def      7
## 4282:      60-69 2022-03-26      def     12
## 4283:      70-79 2022-03-26      def     19
## 4284:      80+  2022-03-26      def     84
```



4 Juntar información demografica para cálculo de tasas

4.1 Descarga de los datos de población (fuente INE)

Descargar los datos de Población del INE:

```
ruta.ine="https://www.ine.es/jaxi/files/_px/es/csv_bdsc/t20/e245/p08/10/01002.csv_bdsc?nocab=1"
ine=fread(ruta.ine,encoding="UTF-8")
ine
```

```
##      Edad (grupos quinquenales) Españoles/Extranjeros      Sexo      Año
##      1:      TOTAL EDADES      TOTAL Ambos sexos 2022
##      2:      TOTAL EDADES      TOTAL Ambos sexos 2021
##      3:      TOTAL EDADES      TOTAL Ambos sexos 2020
##      4:      TOTAL EDADES      TOTAL Ambos sexos 2019
```

```

##      5:                TOTAL EDADES                TOTAL Ambos sexos 2018
##      ---
## 4946:                100 años y más                Extranjeros    Mujeres 2002
## 4947:                100 años y más                Extranjeros    Mujeres 2001
## 4948:                100 años y más                Extranjeros    Mujeres 2000
## 4949:                100 años y más                Extranjeros    Mujeres 1999
## 4950:                100 años y más                Extranjeros    Mujeres 1998
##
##      Total
##      1: 47.475.420
##      2: 47.385.107
##      3: 47.450.795
##      4: 47.026.208
##      5: 46.722.980
##      ---
## 4946:                199
## 4947:                276
## 4948:                239
## 4949:                189
## 4950:                46

```

4.2 Formateo de la base INE

Filtrar y formatear la base del INE para poder juntar los datos de población de 2021 (por grupos de edad) a la base agregada de covid anterior.

```

##      grupo_edad_ine origen      sexo año      total      pob grupo_edad
## 1:      0-4 años  TOTAL Ambos sexos 2021 1.887.689 1887689      0-9
## 2:      5-9 años  TOTAL Ambos sexos 2021 2.275.778 2275778      0-9
## 3:     10-14 años  TOTAL Ambos sexos 2021 2.526.140 2526140     10-19
## 4:     15-19 años  TOTAL Ambos sexos 2021 2.417.918 2417918     10-19
## 5:     20-24 años  TOTAL Ambos sexos 2021 2.382.448 2382448     20-29
## 6:     25-29 años  TOTAL Ambos sexos 2021 2.548.736 2548736     20-29
## 7:     30-34 años  TOTAL Ambos sexos 2021 2.786.314 2786314     30-39
## 8:     35-39 años  TOTAL Ambos sexos 2021 3.250.538 3250538     30-39
## 9:     40-44 años  TOTAL Ambos sexos 2021 3.923.707 3923707     40-49
## 10:    45-49 años  TOTAL Ambos sexos 2021 3.945.845 3945845     40-49
## 11:    50-54 años  TOTAL Ambos sexos 2021 3.699.275 3699275     50-59
## 12:    55-59 años  TOTAL Ambos sexos 2021 3.422.655 3422655     50-59
## 13:    60-64 años  TOTAL Ambos sexos 2021 3.007.236 3007236     60-69
## 14:    65-69 años  TOTAL Ambos sexos 2021 2.480.872 2480872     60-69
## 15:    70-74 años  TOTAL Ambos sexos 2021 2.199.993 2199993     70-79
## 16:    75-79 años  TOTAL Ambos sexos 2021 1.767.146 1767146     70-79
## 17:    80-84 años  TOTAL Ambos sexos 2021 1.294.253 1294253      80+
## 18:    85-89 años  TOTAL Ambos sexos 2021   991.099  991099      80+
## 19:    90-94 años  TOTAL Ambos sexos 2021   446.746  446746      80+
## 20:    95-99 años  TOTAL Ambos sexos 2021   112.699  112699      80+
## 21: 100 años y más  TOTAL Ambos sexos 2021    18.020   18020      80+
##      grupo_edad_ine origen      sexo año      total      pob grupo_edad

```

4.3 Juntar la base INE con la base agregada de covid

```

##      grupo_edad      periodo variable conteo      pob
## 1:      0-9 2020-01-01      casos      0 4163467
## 2:      0-9 2020-01-08      casos      1 4163467
## 3:      0-9 2020-01-15      casos      0 4163467

```

```
## 4:      0-9 2020-01-22  casos      0 4163467
## 5:      0-9 2020-01-29  casos      0 4163467
## ---
## 4280:    80+ 2022-02-26   def      370 2862817
## 4281:    80+ 2022-03-05   def      297 2862817
## 4282:    80+ 2022-03-12   def      248 2862817
## 4283:    80+ 2022-03-19   def      207 2862817
## 4284:    80+ 2022-03-26   def       84 2862817
```

4.4 Cálculo de las tasas

Calcular tasas de incidencia mensuales (por 100.000 habitantes) por grupos de edad y por categoría de gravedad.

$$\text{tasa} = \frac{\text{conteo}}{\text{poblacion}} \times 100.000$$

4.5 Resumen de las tasas

Representar mediante una tabla el máximo anual de incidencia y mortalidad por COVID según grupos de edad:

##	grupo_edad	casos_2020	casos_2021	casos_2022	def_2020	def_2021	def_2022
## 1:	0-9	203.77	1272.62	3319.81	0.05	0.05	0.07
## 2:	10-19	361.97	1585.92	2640.08	0.04	0.02	0.04
## 3:	20-29	391.89	2365.68	2642.81	0.20	0.10	0.06
## 4:	30-39	312.73	1934.67	2720.47	0.28	0.17	0.17
## 5:	40-49	301.69	1780.24	2657.22	0.71	0.41	0.30
## 6:	50-59	298.95	1439.26	1980.12	2.68	1.59	1.12
## 7:	60-69	251.18	774.00	1076.40	10.57	5.94	3.44
## 8:	70-79	243.17	460.63	970.22	43.81	20.87	8.67
## 9:	80+	369.04	540.06	1182.61	122.78	90.05	39.68