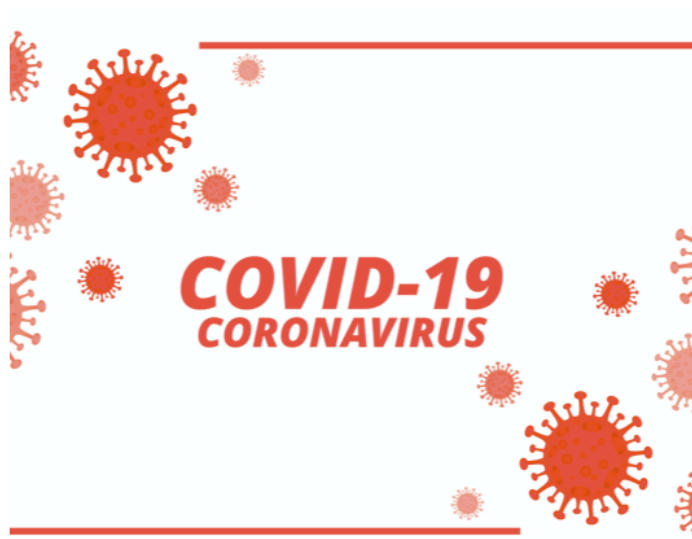




Università degli Studi di Napoli Federico II

FACOLTÀ DI INGEGNERIA
Corso di Laurea Magistrale in Ingegneria Gestionale

Studio del virus COVID-19: Analisi delle modalità e vie di diffusione



Docente:
Elio Masciari

Studente:
Giovanni Officioso
Matricola M62001579

Con il seguente progetto si andrà ad analizzare la diffusione del virus COVID-19 nel territorio nazionale italiano, evidenziando particolari differenze di veicolazione del virus stesso a livello regionale. Pertanto si provvederà in prima istanza a fornire una stima dei dati del contagio elaborando una sintesi circa le regioni più o meno affette; in seconda istanza invece si provvederà ad analizzare le motivazioni relative alla differente diffusione del virus, nonché l'impatto che esso ha avuto nei confronti dello stile di vita della popolazione italiana. Il tutto sarà effettuato tramite raccolta dati e compilazione di databases ad essi funzionali, con conseguente estrazione dei risultati delle analisi tramite opportune queries. L'elaborato sarà strutturato nel seguente modo:

1. elaborazione di una tabella "master" riepilogativa dei dati necessari a sviluppare le analisi previste;
2. esplosione della tabella master in tabelle figlie ad essa afferenti;
3. formulazione di differenti queries per estrapolare i dati funzionali ad analizzare il fenomeno.

Per ottemperare alle richieste è stato sfruttato come relational database management system MySQL.

INDICE

1	CREAZIONE TABELLE	1	
1.1	Cause ed effetti della diffusione del virus		2
1.2	Esplosione tabella master	3	
2	SCHEMA ER	7	
3	POPOLAZIONE E MANIPOLAZIONE TABELLE		8
3.1	Popolazione tabella master	8	
3.1.1	Popolazione tabella Trentino		12
3.1.2	Manipolazione tabella master		14
3.2	Popolazione dati sottotabelle	15	
4	QUERY	20	

1

CREAZIONE TABELLE

Innanzitutto è stata creata una tabella, denominata tabella "**master**", atta a contenere i dati con cui sarà costruito l'intero database in base a cui effettuare le analisi. Siccome la base di dati con cui inizialmente popolare la tabella è improntata su informazioni disponibili a livello provinciale, al fine di creare la tabella master è stato utilizzato il seguente script:

```
1 CREATE TABLE `master` (  
2   `data` TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,  
3   stato CHAR (3) DEFAULT NULL,  
4   cod_reg SMALLINT NOT NULL,  
5   regione VARCHAR(35) NOT NULL,  
6   cod_prov SMALLINT NOT NULL DEFAULT '0',  
7   provincia VARCHAR(50) DEFAULT NULL,  
8   sigla_prov CHAR(2) DEFAULT NULL,  
9   lat DECIMAL(10,8) DEFAULT NULL,  
10  `long` DECIMAL(10,8) DEFAULT NULL,  
11  tot_casi_prov INT DEFAULT NULL,  
12  densità_abitanti_prov SMALLINT DEFAULT NULL,  
13  note_it VARCHAR(50) DEFAULT NULL,  
14  note_en VARCHAR(50) DEFAULT NULL,  
15  PRIMARY KEY (`data`, regione, cod_reg, cod_prov)  
16 );
```

Successivamente è stata modificata la tabella master in maniera conforme al tipo di informazioni da aggiungere.

Per comprendere la struttura finale della tabella, bisogna tenere presente lo spirito con cui è stato svolto l'elaborato: da un lato comprendere la differenza che si è riscontrata in alcune aree del paese in merito sia alle cause che potrebbero aver favorito una diffusione massiccia del virus sia ai fattori che potrebbero aver influito negativamente sul numero di decessi, dall'altro comprendere come siano effettivamente variate, a seguito del lockdown forzato, le esigenze e le consuetudini degli italiani. In virtù di tale scopo, si terrà conto della percentuale di variazione degli acquisti effettuati in merito a:

- le possibili attività ricreative da perpetuare,
- lo sfruttamento di servizi digitali come Netflix ed Amazon Prime,
- la necessità di disporre di una connessione in fibra ottica in grado di garantire performance elevate.

Innanzitutto la tabella dovrà, necessariamente, essere arricchita con informazioni relative all'andamento della diffusione del virus in base ai dati che saranno importati dalla repository ufficiale del governo. Si tenga presente che l'analisi sarà focalizzata a livello regionale e

non provinciale in quanto, per fronteggiare l'emergenza soprattutto nelle regioni più colpite, numerosi pazienti sono stati ospedalizzati e curati in province diverse da quelle di appartenenza. In tale ottica alla tabella master sono state aggiunte le seguenti colonne:

```

1 #Aggiunte tabella master
2 ALTER TABLE `master`
3 ADD COLUMN ricov_sintomi INT,
4 ADD COLUMN terapia_intensiva INT,
5 ADD COLUMN totale_ospedal INT,
6 ADD COLUMN isolamento_domiciliare INT,
7 ADD COLUMN totale_pos INT,
8 ADD COLUMN delta_tot_pos INT,
9 ADD COLUMN nuovi_pos INT,
10 ADD COLUMN dimessi_guariti INT,
11 ADD COLUMN deceduti INT,
12 ADD COLUMN totale_casi_regionali INT,
13 ADD COLUMN tamponi INT,
14 ADD COLUMN casi_testati INT;
```

In seguito si è provveduto ad inserire all'interno della medesima tabella ulteriori colonne per valutare sia le cause ritenute responsabili di una maggiore diffusione del virus, sia gli effetti che la pandemia ha avuto sulle diverse abitudini degli italiani.

1.1 CAUSE ED EFFETTI DELLA DIFFUSIONE DEL VIRUS

Tramite diverse analisi, sono state annoverate come ipotetiche cause di una massiccia diffusione del virus:

- un'età media regionale elevata;
- un numero elevato di case di riposo;
- il numero di accesso alle regioni in termini di passeggeri aerei e traffico su gomma mensile;
- le temperature medie;
- la tipologia di cura.

Un fattore che sicuramente ha influito sulla diffusione del virus, è la percentuale di particolato atmosferico presente nell'aria, correlabile con un tasso di inquinamento maggiore. Infatti diversi studi hanno dimostrato come le zone in cui è presente una concentrazione di particolato più elevata rispetto alla norma, sono più esposte ai fenomeni di diffusione del virus. In virtù dell'esistenza di un'abbondante letteratura volta ad effettuare questa tipologia di analisi, si è preferito dare come assodato questo risultato senza passarlo in rassegna.

Per valutare adeguatamente i fattori fino ad ora elencati, è stata ulteriormente arricchita la tabella master nel modo seguente:

```

1 #Aggiunte tabella master
2 ALTER TABLE 'master'
3 ADD COLUMN età_media DECIMAL (3,1),
4 ADD COLUMN case_riposo INT,
5 ADD COLUMN km_autostrada INT,
6 ADD COLUMN flusso_veicoli INT,
7 ADD COLUMN cura VARCHAR(50),
8 ADD COLUMN data_cura DATE,
9 ADD COLUMN posti_intensiva_reg INT,
10 ADD COLUMN densità_abitanti_reg INT,
11 ADD COLUMN aeroporto VARCHAR(50),
12 ADD COLUMN passeggeri_nazionali INT,
13 ADD COLUMN passeggeri_internazionali INT,
14 ADD COLUMN variazione_acquisti_attrezzi DECIMAL (3, 2),
15 ADD COLUMN variazione_acquisti_libri DECIMAL (3, 2),
16 ADD COLUMN variazione_uso_netflix DECIMAL (3, 2),
17 ADD COLUMN variazione_uso_amazon DECIMAL (3, 2),
18 ADD COLUMN variazione_richiesta_fibra_tim DECIMAL (3, 2),
19 ADD COLUMN variazione_richiesta_fibra_wind3 DECIMAL (3, 2);

```

In tal modo la tabella master è stata interamente conformata a tutti i dati che andranno a popolarla. L'aspetto positivo di tale tabella risiede nella presenza in un'unica vista dell'intero set di dati nonché nella facilità di esecuzione delle queries, evitando il ricorso a costrutti nidificati e di **join** tra tabelle. Tuttavia, in tale vista, vi sarà senza dubbio una parziale ridondanza di dati non identificativi di una singola tupla, in quanto saranno attribuite ad esempio a ciascuna provincia, grandezze riferite alla regione, come, ad esempio, l'estensione della rete autostradale regionale.

1.2 ESPLOSIONE TABELLA MASTER

Al fine di ottimizzare l'esecuzione delle diverse queries, facendo in modo che l'accesso al database fosse il più veloce possibile, sono state create diverse tabelle che raggruppassero diverse serie di dati. Nello specifico è stata dapprima creata una tabella denominata "**regioni**":

```

1 #creazione tabella regioni
2 create table regioni
3 (
4   cod_reg SMALLINT not null,
5   regione VARCHAR (35) not null unique,
6   età_media DECIMAL (3,1),
7   case_riposo SMALLINT,
8   km_autostrada SMALLINT,
9   cura VARCHAR (50),
10  data_cura DATE,
11  posti_intensiva_reg SMALLINT,
12  densità_abitanti SMALLINT,
13  PRIMARY KEY (cod_reg)
14 );

```

In seguito è stata creata una tabella relativa all'andamento del virus nelle diverse regioni denominata "**datareg**":

```

1  #creazione tabella datareg
2  CREATE TABLE datareg
3  (
4  'data' TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,
5  cod_reg SMALLINT NOT NULL,
6  regione VARCHAR (35) NOT NULL,
7  ricov_sintomi INT DEFAULT 0,
8  terapia_intensiva INT DEFAULT 0,
9  totale_ospedal INT DEFAULT 0,
10 isolamento_domiciliare INT DEFAULT 0,
11 totale_pos INT DEFAULT 0,
12 delta_tot_pos INT DEFAULT 0,
13 nuovi_pos INT DEFAULT 0,
14 dimessi_guariti INT DEFAULT 0,
15 deceduti INT DEFAULT 0,
16 totale_casi INT DEFAULT 0,
17 tamponi INT DEFAULT 0,
18 casi_testati INT DEFAULT 0,
19 PRIMARY KEY ('data', cod_reg),
20 FOREIGN KEY (cod_reg) REFERENCES regioni(cod_reg)
21 );

```

Successivamente sono state create due tabelle relative alle province. La prima denominata per l'appunto "**province**", riporta i dati di densità abitativa di ciascuna provincia:

```

1  #creazione tabella province
2  CREATE TABLE province
3  (
4  cod_reg SMALLINT,
5  regione VARCHAR (35),
6  cod_prov SMALLINT NOT NULL,
7  provincia VARCHAR(50) UNIQUE,
8  sigla_prov CHAR(2) UNIQUE,
9  densità_abitanti_prov INT,
10 PRIMARY KEY (cod_prov, provincia),
11 UNIQUE KEY 'prov_codpro' (cod_prov,provincia),
12 FOREIGN KEY (cod_reg, regione) REFERENCES regioni(cod_reg, regione)
13 );

```

l'altra, denominata "**dataprov**", riporta per ogni giorno il totale dei casi constatati:

```

1  #creazione tabella dataprov
2  CREATE TABLE dataprov
3  (
4  'data' TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,
5  cod_reg SMALLINT,
6  regione VARCHAR (35),
7  cod_prov SMALLINT NOT NULL,
8  provincia VARCHAR(50),
9  sigla_prov CHAR(2),
10 tot_casi_prov INT,
11 PRIMARY KEY ('data', cod_prov),
12 FOREIGN KEY (cod_prov, provincia) REFERENCES province(cod_prov, provincia)
13 );

```

In seguito per monitorare gli accessi per via aerea è stata creata la tabella **"aeroporti"** selezionando i principali per traffico aereo commerciale come riportato dal sito [assaeroporti](#):

```

1 #creazione tabella aeroporti
2 CREATE TABLE aeroporti
3 (
4   'data' TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,
5   nome VARCHAR (25),
6   regione VARCHAR (35),
7   cod_reg SMALLINT,
8   provincia VARCHAR(50) REFERENCES province(provincia),
9   cod_prov SMALLINT NOT NULL REFERENCES province(cod_prov),
10  pass_nazionali INT NOT NULL,
11  pass_internazionali INT NOT NULL,
12  PRIMARY KEY ('data', cod_prov),
13  FOREIGN KEY (regione, cod_reg) REFERENCES regioni(regione, cod_reg),
14  FOREIGN KEY (provincia, cod_prov) REFERENCES province(provincia, cod_prov)
15 );

```

Per monitorare gli accessi tramite veicoli che circolano su gomma è stata creata la tabella **"veicoli"**, con l'indicazione del traffico su strada rilevato nei mesi di febbraio e marzo, le cui informazione sono state tratte dal sito dell'[anas](#):

```

1 #creazione tabella Veicoli
2 CREATE TABLE veicoli
3 (
4   'data' TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,
5   regione VARCHAR (35),
6   cod_reg SMALLINT not null,
7   traffico INT NOT NULL,
8   PRIMARY KEY ('data', cod_reg),
9   FOREIGN KEY (regione, cod_reg) REFERENCES regioni(regione, cod_reg)
10 );

```

In seguito è stata creata un'ulteriore tabella atta a monitorare la variazione degli usi del popolo italiano, denominata **"usi"**:

```

1 #creazione tabella usi
2 CREATE TABLE usi
3 (
4   'data' TIMESTAMP NOT NULL DEFAULT CURRENT_TIMESTAMP,
5   regione VARCHAR (35),
6   cod_reg SMALLINT,
7   variazione_acquisti_attrezzi DECIMAL (3, 2),
8   variazione_acquisti_libri DECIMAL (3, 2),
9   variazione_uso_netflix DECIMAL (3, 2),
10  variazione_uso_amazon DECIMAL (3, 2),
11  variazione_richiesta_fibra_tim DECIMAL (3, 2),
12  variazione_richiesta_fibra_wind3 DECIMAL (3, 2),
13  PRIMARY KEY (cod_reg),
14  FOREIGN KEY (regione, cod_reg) REFERENCES regioni(regione, cod_reg)
15 );

```


Per tenere conto invece delle province autonome di Trento e Bolzano, è stata creata una tabella d'appoggio denominata "**Trentino**" che andrà popolata con la somma dei casi constatati a livello regionale. Il codice utilizzato per la creazione della tabella è stato:

```

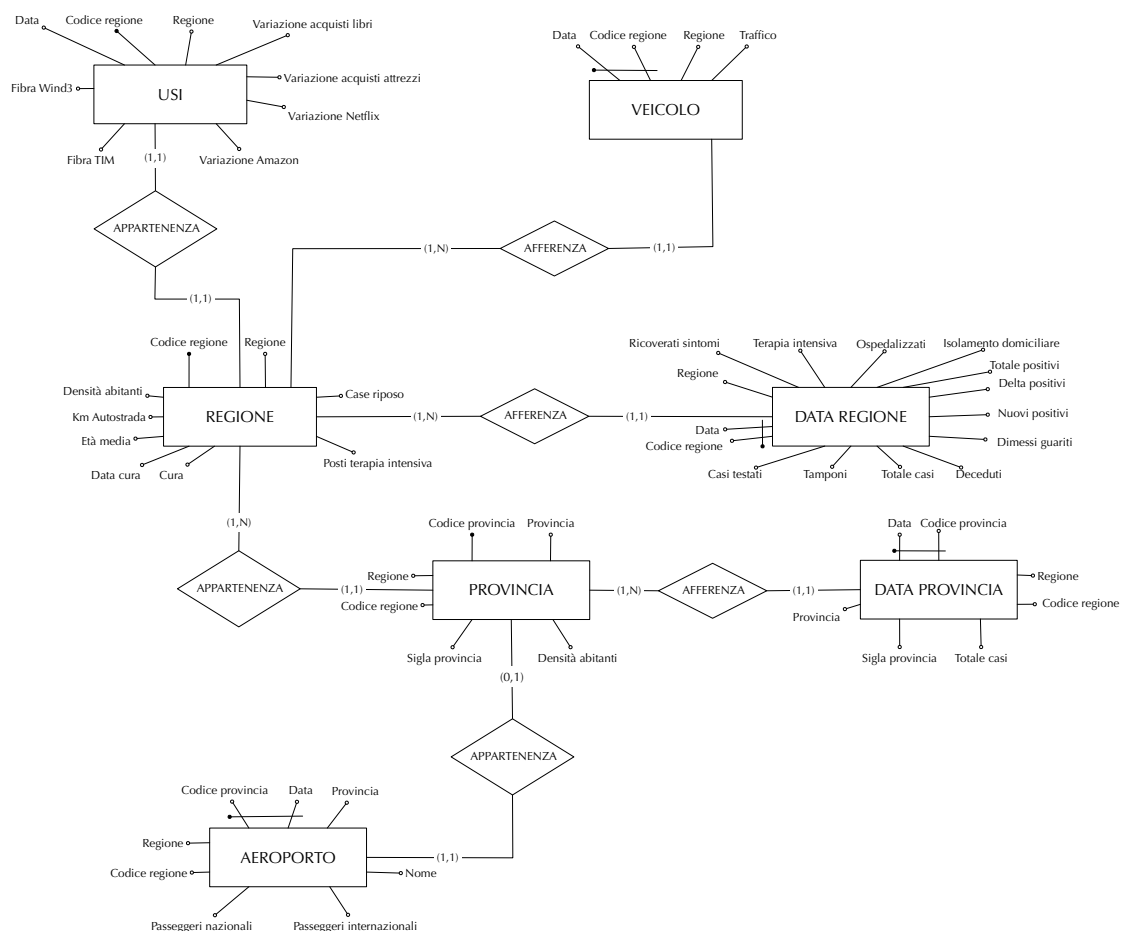
1 #creazione tabella Trentino
2 CREATE TABLE Trentino(
3   'data' timestamp NOT NULL DEFAULT CURRENT_TIMESTAMP,
4   stato char(3) DEFAULT NULL,
5   cod_reg smallint NOT NULL,
6   regione varchar(35) NOT NULL,
7   cod_prov smallint NOT NULL DEFAULT '0',
8   ricov_sintomi int DEFAULT NULL,
9   terapia_intensiva int DEFAULT NULL,
10  totale_ospedal int DEFAULT NULL,
11  isolamento_domiciliare int DEFAULT NULL,
12  totale_pos int DEFAULT NULL,
13  delta_tot_pos int DEFAULT NULL,
14  nuovi_pos int DEFAULT NULL,
15  dimessi_guariti int DEFAULT NULL,
16  deceduti int DEFAULT NULL,
17  totale_casi_regionali int DEFAULT NULL,
18  tamponi int DEFAULT NULL,
19  casi_testati int DEFAULT NULL,
20  età_media decimal(3,1) DEFAULT NULL,
21  case_riposo int DEFAULT NULL,
22  km_autostrada int DEFAULT NULL,
23  flusso_veicoli int DEFAULT NULL,
24  cura varchar(50) DEFAULT NULL,
25  data_cura date DEFAULT NULL,
26  posti_intensiva_reg int DEFAULT NULL,
27  densità_abitanti_reg int DEFAULT NULL,
28  variazione_acquisti_attrezzi decimal(3,2) DEFAULT NULL,
29  variazione_acquisti_libri decimal(3,2) DEFAULT NULL,
30  variazione_uso_netflix decimal(3,2) DEFAULT NULL,
31  variazione_uso_amazon decimal(3,2) DEFAULT NULL,
32  variazione_richiesta_fibra_tim decimal(3,2) DEFAULT NULL,
33  variazione_richiesta_fibra_wind3 decimal(3,2) DEFAULT NULL,
34  PRIMARY KEY('data', cod_reg)
35 );

```

In seguito si è provveduto a popolare le tabelle create a partire da quella master.

2 | SCHEMA ER

Giunti a tal punto, tramite un processo di Reverse Engineering, è stato possibile realizzare uno schema concettuale E/R della base di dati costruita.



3 | POPOLAZIONE E MANIPOLAZIONE TABELLE

3.1 POPOLAZIONE TABELLA MASTER

In seguito, è stato riportato il dataset ufficiale di [github](#) nella tabella master sfruttando il comando di "import" del file in estensione csv:



Figura 1: Import file csv

In tal modo si è provveduto a popolare la tabella master con i dati relativi all'andamento della diffusione del virus a livello provinciale. In seguito è stato importato il dataset relativo alla diffusione del virus a livello regionale seguendo la medesima modalità.

Al fine di completare il dataset relativo alla tabella master, sono state aggiunte le restanti informazioni sfruttando diversi script. Innanzitutto sono state aggiunte le informazioni relative al traffico aereo registrate nelle diverse regioni:

```
1 #febbraio
2 UPDATE
3   'master' m
4 SET
5   aeroporto = 'Alghero',
6   passeggeri_internazionali = 9997,
7   passeggeri_nazionali = 60994
8 WHERE
9   m.'data' LIKE '2020-02-29%'
10  AND m.cod_prov = 90;
11
12 UPDATE
13   'master' m
14 SET
15   aeroporto = 'Ancona',
16   passeggeri_internazionali = 19192,
17   passeggeri_nazionali = 3677
18 WHERE
19   m.cod_prov = 42
20   AND m.'data' LIKE '2020-02-29%';
21 ...
22
23 update
24   'master' m
```

```

25 set
26     aeroporto = 'Verona',
27     passeggeri_internazionali = 127475,
28     passeggeri_nazionali = 60205
29 where
30     m.cod_prov = 23
31     and m.'data' LIKE '2020-02-29%';
32
33 #marzo
34
35 update
36     'master' m
37 set
38     aeroporto = 'Alghero',
39     passeggeri_internazionali = 1854,
40     passeggeri_nazionali = 13397
41 where
42     m.'data' LIKE '2020-03-31%'
43     and m.cod_prov = 90;
44
45 update
46     'master' m
47 set
48     aeroporto = 'Ancona',
49     passeggeri_internazionali = 3741,
50     passeggeri_nazionali = 762
51 where
52     m.cod_prov = 42
53     and m.'data' LIKE '2020-03-31%';
54 ...

```

Successivamente sono stati aggiunti i dati relativi al mese di febbraio in egual modo. Si è pertanto provveduto ad inserire i dati relativi al traffico rilevato nei mesi di febbraio e marzo ¹. Si riporta uno stralcio dello script utilizzato:

```

1
2 #mese febbraio
3 UPDATE
4     'master' m
5 SET
6     flusso_veicoli = 6920
7 WHERE
8     m.'data' LIKE '2020-02-29%'
9     AND m.cod_prov = 0
10    AND m.regione = 'Abruzzo';
11
12 UPDATE
13     'master' m
14 SET
15     flusso_veicoli = 5620
16 WHERE
17     m.'data' LIKE '2020-02-29%'
18     AND m.cod_prov = 0
19     AND m.regione = 'Basilicata';
20 ...
21
22 #mese marzo
23 UPDATE
24     'master' m
25 SET
26     flusso_veicoli = 3226
27 WHERE
28     m.'data' LIKE '2020-03-31%'
29     AND m.cod_prov = 0
30     AND m.regione = 'Abruzzo';
31
32 UPDATE
33     'master' m
34 SET
35     flusso_veicoli = 2513
36 WHERE
37     m.'data' LIKE '2020-03-31%'
38     AND m.cod_prov = 0
39     AND m.regione = 'Basilicata';
40 ...

```

¹ Per i dati aggregati a base mensile come il traffico autostradale o aereo, è stata considerata come data di riferimento l'ultimo giorno del mese.

In seguito sono stati aggiornati diversi campi. Nello specifico con un primo script sono state aggiornate per ciascuna regione:

- l'età media;
- il tipo di cura con cui sono stati trattati i pazienti in gravi condizioni respiratorie,
- la data da cui hanno iniziato ad usare quel tipo di cura,
- i km di autostrada,
- il numero di case di riposo,
- la densità di abitanti,
- i posti in terapia intensiva disponibili.

Lo script è stato il seguente:

```

1  #info regionali#
2  update
3      `master`
4  set
5      età_media = 46.4,
6      case_riposi = 16,
7      km_autostrada = 1447,
8      cura = 'Idrossiclorochina',
9      data_cura = '2020-03-26',
10     densità_abitanti_reg = 121,
11     posti_intensiva_reg = 151
12 where
13     `master`.cod_prov = 0
14     and `master`.regione = 'Abruzzo';
15
16
17 update
18     `master`
19 set
20     età_media = 46.1,
21     case_riposi = 1,
22     km_autostrada = 1064,
23     cura = 'Idrossiclorochina',
24     data_cura = '2020-03-16',
25     densità_abitanti_reg = 56,
26     posti_intensiva_reg = 64
27 where
28     `master`.cod_prov = 0
29     and `master`.regione = 'Basilicata';
30
31 ...

```

L'informazione riguardante la data in cui sono state implementate le diverse tipologie di cura per i pazienti che versavano in condizioni respiratorie critiche, è stata tratta dal *Sole 24 Ore*.

Infine sono state aggiunte alla tabella master le informazioni relative alla variazione degli usi degli italiani, mutate tramite gli open data forniti da [Google](#)

```

1
2 #aggiornamento variazione usi
3 UPDATE
4     `master` m
5 SET
6     variazione_acquisti_attrezzi = 0.56,
7     variazione_acquisti_libri = 0.12,
8     variazione_uso_netflix = 0.38,
9     variazione_uso_amazon = 0.14,
10    variazione_richiesta_fibra_tim = 0.14,
11    variazione_richiesta_fibra_wind3 = 0.08
12 WHERE
13     m.`data` LIKE '2020-05-03%'
14     AND m.cod_prov = 0
15     AND m.regione = 'Abruzzo';
16
17 UPDATE
18     `master` m
19 SET
20     variazione_acquisti_attrezzi = 0.70,
21     variazione_acquisti_libri = 0.12,
22     variazione_uso_netflix = 0.36,
23     variazione_uso_amazon = 0.08,
24     variazione_richiesta_fibra_tim = -0.10,
25     variazione_richiesta_fibra_wind3 = 0.16
26 WHERE
27     m.`data` LIKE '2020-05-03%'
28     AND m.cod_prov = 0
29     AND m.regione = 'Basilicata';
30
31 ...

```

È stata quindi effettuata un'operazione di normalizzazione del codice delle province, in quanto a partire dalla provincia 108, esse non presentavano una numerazione consequenziale e congruente con quella delle province precedenti:

```

1
2 #normalizzazione cod_prov
3 update
4     `master`
5 set
6     cod_prov=104
7 where
8     `master`.cod_prov=108;
9
10 update
11     `master`
12 set
13     cod_prov=105
14 where
15     `master`.cod_prov=109;
16
17 update
18     `master`
19 set
20     cod_prov=106
21 where
22     `master`.cod_prov=110;
23
24 update
25     `master`
26 set
27     cod_prov=107
28 where
29     `master`.cod_prov=111;

```

Quindi è stato possibile inserire per ogni provincia l'informazione circa la densità di abitanti:

```

1
2 #aggiornamento densità province#
3 update
4     'master'
5 set
6     densità_abitanti_prov = 331
7 where
8     'master'.cod_prov = 1;
9
10 update
11     'master'
12 set
13     densità_abitanti_prov = 82
14 where
15     'master'.cod_prov = 2;
16 ...
17
18 update
19     'master'
20 set
21     densità_abitanti_prov = 54
22 where
23     'master'.cod_prov = 107;

```

3.1.1 Popolazione tabella Trentino

Prima di manipolare la tabella master, è stata popolata la tabella "Trentino" con i dati regionali da inserire successivamente nella tabella master. Nello specifico sono stati utilizzati diversi script:

```

1 INSERT INTO 'Trentino'('data', 'ricov_sintomi', 'terapia_intensiva', 'totale_ospedal',
2 'isolamento_domiciliare', 'totale_pos', 'delta_tot_pos', 'nuovi_pos',
3 'dimessi_guariti', 'deceduti', 'totale_casi_regionali', 'tamponi', 'casi_testati')
4 (SELECT m.'data',
5     SUM(m.'ricov_sintomi'),
6     SUM(m.'terapia_intensiva'),
7     SUM(m.'totale_ospedal'),
8     SUM(m.'isolamento_domiciliare'),
9     SUM(m.'totale_pos'),
10    SUM(m.'delta_tot_pos'),
11    SUM(m.'nuovi_pos'),
12    SUM(m.'dimessi_guariti'),
13    SUM(m.'deceduti'),
14    SUM(m.'totale_casi_regionali'),
15    SUM(m.'tamponi'),
16    SUM(m.'casi_testati')
17 FROM 'master' m
18 WHERE m.'cod_prov'= 0
19 AND m.'cod_reg'= 4
20 GROUP BY m.'data');

```

```

1 UPDATE
2     'Trentino' t
3 SET
4     t.'cod_prov'= 0,
5     t.'cod_reg'= 4,
6     t.'stato'='ITA',
7     t.'età_media'= (SELECT m.'età_media'
8 FROM 'master' m
9 WHERE m.'cod_prov'= 0 and m.'cod_reg'=4 and m.'regione'= 'P.A. Trento'
10 and m.'data' like '2020-05-03%'),
11     t.'case_riposo'= (SELECT m.'case_riposo'
12 FROM 'master' m
13 WHERE m.'cod_prov'= 0 and m.'cod_reg'=4 and m.'regione'= 'P.A. Trento'
14 and m.'data' like '2020-05-03%'),
15     t.'km_autostrada'= (SELECT m.'km_autostrada'
16 FROM 'master' m
17 WHERE m.'cod_prov'= 0 and m.'cod_reg'=4 and m.'regione'= 'P.A. Trento'
18 and m.'data' like '2020-05-03%'),
19     t.'flusso_veicoli'= (SELECT m.'flusso_veicoli'
20 FROM 'master' m

```

```

21 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
22 and m.'data' like '2020-05-03%',
23 t.'cura'=(SELECT m.'cura'
24 FROM 'master' m
25 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
26 and m.'data' like '2020-05-03%'),
27 t.'data_cura'=(SELECT m.'data_cura'
28 FROM 'master' m
29 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
30 and m.'data' like '2020-05-03%'),
31 t.'posti_intensiva_reg'=(SELECT m.'posti_intensiva_reg'
32 FROM 'master' m
33 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
34 and m.'data' like '2020-05-03%'),
35 t.'densità_abitanti_reg'=(SELECT m.'densità_abitanti_reg'
36 FROM 'master' m
37 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
38 and m.'data' like '2020-05-03%');

```

```

1 UPDATE 'Trentino' t
2 SET
3     t.'flusso_veicoli'=(SELECT m.'flusso_veicoli'
4 FROM 'master' m
5 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
6 and m.'data' like '2020-02-29%')
7 WHERE
8     t.'data' LIKE '2020-02-29%';
9
10 UPDATE 'Trentino' t
11 SET
12     t.'flusso_veicoli'=(SELECT m.'flusso_veicoli'
13 FROM 'master' m
14 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
15 and m.'data' like '2020-03-31%')
16 WHERE
17     t.'data' LIKE '2020-03-31%';
18
19 UPDATE 'Trentino' t
20 SET
21     t.'variazione_acquisti_attrezzi'=(SELECT m.'variazione_acquisti_attrezzi'
22 FROM 'master' m
23 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
24 and m.'data' like '2020-05-03%'),
25     t.'variazione_acquisti_libri'=(SELECT m.'variazione_acquisti_libri'
26 FROM 'master' m
27 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
28 and m.'data' like '2020-05-03%'),
29     t.'variazione_uso_netflix'=(SELECT m.'variazione_uso_netflix'
30 FROM 'master' m
31 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
32 and m.'data' like '2020-05-03%'),
33     t.'variazione_uso_amazon'=(SELECT m.'variazione_uso_amazon'
34 FROM 'master' m
35 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
36 and m.'data' like '2020-05-03%'),
37     t.'variazione_richiesta_fibra_tim'=(SELECT m.'variazione_richiesta_fibra_tim'
38 FROM 'master' m
39 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
40 and m.'data' like '2020-05-03%'),
41     t.'variazione_richiesta_fibra_wind3'=(SELECT m.'variazione_richiesta_fibra_wind3'
42 FROM 'master' m
43 WHERE m.'cod_prov'='0' and m.'cod_reg'='4' and m.'regione'='P.A. Trento'
44 and m.'data' like '2020-05-03%')
45 WHERE
46     t.'data' LIKE '2020-05-03%';

```

Sono state aggiunte le informazioni relative alla regione del "Trentino-Alto Adige":

```

1 INSERT INTO 'master'('data', 'stato', 'cod_reg', 'regione', 'cod_prov', 'ricov_sintomi',
2     'terapia_intensiva', 'totale_ospedal', 'isolamento_domiciliare', 'totale_pos',
3     'delta_tot_pos', 'nuovi_pos', 'dimessi_guariti', 'deceduti',
4     'totale_casi_regionali', 'tamponi', 'casi_testati', 'età_media', 'case_riposo',
5     'km_autostrada', 'flusso_veicoli', 'cura', 'data_cura', 'posti_intensiva_reg',
6     'densità_abitanti_reg', 'passeggeri_nazionali', 'variazione_acquisti_attrezzi',
7     'variazione_acquisti_libri', 'variazione_uso_netflix', 'variazione_uso_amazon',
8     'variazione_richiesta_fibra_tim', 'variazione_richiesta_fibra_wind3')
9 (SELECT *
10 FROM 'Trentino');

```


Successivamente è stata eliminata la tabella d'appoggio relativa alla regione dei Trentino-Alto Adige.

3.1.2 Manipolazione tabella master

È stata quindi manipolata la tabella master per ripulirla di quelle informazioni poco utili ai fini delle analisi da seguire:

```

1  #cancellazione dati di longitudine, latitudine, stato e note
2  ALTER TABLE `master`
3  DROP COLUMN `lat`,
4  DROP COLUMN `long`,
5  DROP COLUMN `stato`,
6  DROP COLUMN `note_it`,
7  DROP COLUMN `note_en`;
8
9
10 #cancellazione righe in fase di aggiornamento
11 DELETE
12 FROM `master` m
13 WHERE m.provincia LIKE 'In fase di%';
14
15 #omologazione regione Trentino-Alto Adige
16 DELETE
17 FROM `master` m
18 WHERE m.`cod_prov`= 0
19 AND m.`cod_reg`= 4
20 AND m.`regione` LIKE 'P.A.%';

```

Di seguito si riporta l'aspetto finale della tabella master:

[illegible][illegible]

data	cod_reg	regione	cod_prov	provincia	sigla_prov	tot_casi_prov	densita_abitanti_prov	ricov_sintomi	terapia_intensiva	totale_ospedal	isolamento_domiciliare	totale_pos	della_tot_pos
2020-02-24 18:00:00	13	Abruzzo	0	N	NUL	NUL	0	0	0	0	0	0	0
2020-02-24 18:00:00	65	Aquila	65	AQ	AD	0	159	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	13	Abruzzo	67	Teramo	TE	0	NUL	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	13	Abruzzo	68	Pescara	PE	0	259	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	13	Abruzzo	69	Chieti	CH	0	148	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	17	Basilicata	0	NUL	NUL	NUL	0	0	0	0	0	0	0
2020-02-24 18:00:00	17	Basilicata	76	Potenza	PZ	0	55	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	17	Basilicata	77	Matera	MT	0	57	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	18	Calabria	0	NUL	NUL	NUL	0	0	0	0	0	0	0
2020-02-24 18:00:00	18	Calabria	78	Cosenza	CS	0	105	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	18	Calabria	79	Catanzaro	CZ	0	148	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	18	Calabria	80	Reggio di Calabria	RC	0	171	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	18	Calabria	101	Crotone	KR	0	101	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	18	Calabria	102	Vibo Valentia	VV	0	139	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	15	Campania	0	NUL	NUL	NUL	0	0	0	0	0	0	0
2020-02-24 18:00:00	15	Campania	61	Caserta	CE	0	346	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	15	Campania	62	Benevento	BN	0	133	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	15	Campania	63	Napoli	NA	0	2617	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	15	Campania	64	Avellino	AV	0	149	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	15	Campania	65	Salerno	SA	0	222	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	0	NUL	NUL	NUL	10	2	12	0	18	0	18
2020-02-24 18:00:00	8	Emilia-Romagna	33	Piacenza	PO	0	111	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	34	Parma	PR	0	131	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	35	Reggio nell'Emilia	RE	0	42	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	36	Modena	MO	0	262	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	37	Bologna	BO	0	274	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	38	Ferrara	FE	0	131	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	39	Ravenna	RA	0	209	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	40	Faenza	FC	0	106	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	8	Emilia-Romagna	90	Rimini	RN	0	392	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	6	Friuli Venezia Giulia	0	NUL	NUL	NUL	0	0	0	0	0	0	0
2020-02-24 18:00:00	6	Friuli Venezia Giulia	30	Udine	UD	0	106	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	6	Friuli Venezia Giulia	31	Gorizia	GO	0	298	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	6	Friuli Venezia Giulia	32	Trieste	TS	0	113	NUL	NUL	NUL	NUL	NUL	NUL
2020-02-24 18:00:00	6	Friuli Venezia Giulia	93	Pordenone	PN	0	197	NUL	NUL	NUL	NUL	NUL	NUL

Figura 2: Tabella master

3.2 POPOLAZIONE DATI SOTTOTABELLE

Si riportano quindi i diversi script con cui sono stati importati i dati nelle diverse tabelle create, a partire dalla tabella master che li conteneva tutti:

```

1 #insert regioni
2 insert
3   into
4     regioni (cod_reg , regione, età_media , case_riposi ,
5     km_autostrada , cura, data_cura, posti_intensiva_reg,
6     densità_abitanti) (
7     select
8       distinct m.cod_reg , m.regione, m.età_media ,
9       m.case_riposi , m.km_autostrada , m.cura,
10      m.data_cura , m.posti_intensiva_reg ,
11      m.densità_abitanti_reg
12   from
13     master m
14   where
15     m.cod_prov = 0);
16

```

cod_reg	regione	età_media	case_riposi	km_autostrada	cura
1	Piemonte	regione - VARCHAR(35) 3	696	Plaquenil	
2	Valle d'Aosta	46.4	2	143	Remdesivir
3	Lombardia	45.4	678	950	Remdesivir/Idrossiclorochina
4	Trentino-Alto Adige	44.0	54	179	Tocilizumab
5	Veneto	45.9	520	731	Avigan
6	Friuli Venezia Giulia	47.7	72	199	Remdesivir
7	Liguria	49.1	120	644	Tocilizumab
8	Emilia-Romagna	46.3	348	1149	Idrossiclorochina
9	Toscana	47.2	319	1451	Ruxolitinib
10	Umbria	47.2	50	807	Tocilizumab
11	Marche	46.9	51	1261	Tocilizumab
12	Lazio	45.3	207	1228	Lopinavir/Ritonavir
13	Abruzzo	46.4	16	1447	Idrossiclorochina
14	Molise	47.1	6	663	Roacterna
15	Campania	43.0	121	1482	Tocilizumab
16	Puglia	45.1	61	1492	Tocilizumab
17	Basilicata	46.1	1	1064	Idrossiclorochina
18	Calabria	44.8	94	1977	Tocilizumab
19	Sicilia	44.3	39	3914	Tocilizumab
20	Sardegna	47.2	16	2952	Tocilizumab

cura	data_cura	posti_intensiva_reg	densità_abitanti
Plaquenil	2020-03-20	420	172
Remdesivir	2020-04-02	25	39
Remdesivir/Idrossiclorochina	2020-03-15	1260	422
Tocilizumab	2020-03-22	84	79
Avigan	2020-04-08	825	267
Remdesivir	2020-04-01	155	153
Tocilizumab	2020-03-19	251	286
Idrossiclorochina	2020-03-22	962	199
Ruxolitinib	2020-04-06	650	162
Tocilizumab	2020-03-20	96	104
Tocilizumab	2020-03-18	163	162
Lopinavir/Ritonavir	2020-04-03	707	341
Idrossiclorochina	2020-03-29	151	121
Roacterna	2020-03-25	27	69
Tocilizumab	2020-03-15	600	424
Tocilizumab	2020-04-12	306	206

cod_reg	regione	età_media	case_riposi	km_autostrada	cura	data_cura	posti_intensiva_reg	densità_abitanti
1	Piemonte	47.2	608	696	Plaquenil	2020-03-20	420	172
2	Valle d'Aosta	46.4	2	143	Remdesivir	2020-04-02	25	39
3	Lombardia	45.4	678	950	Remdesivir/Idrossiclorochina	2020-03-15	1260	422
4	Trentino-Alto Adige	44.0	54	179	Tocilizumab	2020-03-22	84	79
5	Veneto	45.9	520	731	Avigan	2020-04-08	825	267
6	Friuli Venezia Giulia	47.7	72	199	Remdesivir	2020-04-01	155	153
7	Liguria	49.1	120	644	Tocilizumab	2020-03-19	251	286
8	Emilia-Romagna	46.3	348	1149	Idrossiclorochina	2020-03-22	962	199
9	Toscana	47.2	319	1451	Ruxolitinib	2020-04-06	650	162
10	Umbria	47.2	50	807	Tocilizumab	2020-03-20	96	104
11	Marche	46.9	51	1261	Tocilizumab	2020-03-18	163	162
12	Lazio	45.3	207	1228	Lopinavir/Ritonavir	2020-04-03	707	341
13	Abruzzo	46.4	16	1447	Idrossiclorochina	2020-03-29	151	121
14	Molise	47.1	6	663	Roacterna	2020-03-25	27	69
15	Campania	43.0	121	1482	Tocilizumab	2020-03-15	600	424
16	Puglia	45.1	61	1492	Tocilizumab	2020-04-12	306	206
17	Basilicata	46.1	1	1064	Idrossiclorochina	2020-03-20	64	56
18	Calabria	44.8	94	1977	Tocilizumab	2020-03-30	221	128
19	Sicilia	44.3	39	3914	Tocilizumab	2020-03-20	611	194
20	Sardegna	47.2	16	2952	Tocilizumab	2020-03-29	163	68

Figura 3: Tabella regioni

```

1 #insert datareg
2 insert
3 into
4 datareg ("data", cod_reg , regione, ricov_sintomi ,
5 terapia_intensiva, totale_ospedal ,
6 isolamento_domiciliare , totale_pos, delta_tot_pos,
7 nuovi_pos, dimessi_guariti, deceduti, totale_casi,
8 tamponi, casi_testati )(
9 select
10 distinct m."data" ,
11 m.cod_reg ,
12 m.regione,
13 m.ricov_sintomi ,
14 m.terapia_intensiva ,
15 m.totale_ospedal ,
16 m.isolamento_domiciliare,
17 m.totale_pos ,
18 m.delta_tot_pos ,
19 m.nuovi_pos,
20 m.dimessi_guariti ,
21 m.deceduti ,
22 m.totale_casi_regionali ,
23 m.tamponi ,
24 m.casi_testati
25 from
26 master m
27 where
28 m.cod_prov = 0);
29

```

data	cod_reg	regione	ricov_sintomi	terapia_intensiva	totale_ospedal
2020-02-24 18:00:00	1	Piemonte	2	0	2
2020-02-24 18:00:00	2	Valle d'Aosta	0	0	0
2020-02-24 18:00:00	3	Lombardia	76	19	95
2020-02-24 18:00:00	4	Trentino-AL	0	0	0
2020-02-24 18:00:00	5	Veneto	12	4	16
2020-02-24 18:00:00	6	Friuli Vene...	0	0	0
2020-02-24 18:00:00	7	Liguria	0	0	0
2020-02-24 18:00:00	8	Emilia-Rom...	10	2	12
2020-02-24 18:00:00	9	Toscana	0	0	0
2020-02-24 18:00:00	10	Umbria	0	0	0
2020-02-24 18:00:00	11	Marche	0	0	0
2020-02-24 18:00:00	12	Lazio	1	1	2
2020-02-24 18:00:00	13	Abruzzo	0	0	0
2020-02-24 18:00:00	14	Molise	0	0	0
2020-02-24 18:00:00	15	Campania	0	0	0
2020-02-24 18:00:00	16	Puglia	0	0	0
2020-02-24 18:00:00	17	Basilicata	0	0	0
2020-02-24 18:00:00	18	Calabria	0	0	0
2020-02-24 18:00:00	19	Sicilia	0	0	0
2020-02-24 18:00:00	20	Sardegna	0	0	0
2020-02-25 18:00:00	1	Piemonte	2	0	2
2020-02-25 18:00:00	2	Valle d'Aosta	0	0	0
2020-02-25 18:00:00	3	Lombardia	79	25	104
2020-02-25 18:00:00	4	Trentino-AL	1	0	1
2020-02-25 18:00:00	5	Veneto	12	7	19

isolamento_domiciliare	totale_pos	delta_tot_pos	nuovi_pos	dimessi_guariti	deceduti	totale_casi	tamponi	casi_testati
1	3	0	3	0	0	3	141	NULL
0	0	0	0	0	0	0	7	NULL
71	166	0	166	0	6	172	1463	NULL
0	0	0	0	0	0	0	4	NULL
16	32	0	32	0	1	33	2200	NULL
0	0	0	0	0	0	0	68	NULL
0	0	0	0	0	0	0	1	NULL
6	18	0	18	0	0	18	148	NULL
0	0	0	0	0	0	0	140	NULL
0	0	0	0	0	0	0	0	NULL
0	0	0	0	0	0	0	16	NULL
0	2	0	2	1	0	3	124	NULL
0	0	0	0	0	0	0	5	NULL
0	0	0	0	0	0	0	0	NULL
0	0	0	0	0	0	0	10	NULL
0	0	0	0	0	0	0	0	NULL
0	0	0	0	0	0	0	0	NULL
0	0	0	0	0	0	0	1	NULL
0	0	0	0	0	0	0	5	NULL
1	3	0	0	0	0	3	141	NULL
0	0	0	0	0	0	0	7	NULL
127	231	65	68	0	9	240	3700	NULL
0	1	1	1	0	0	1	4	NULL
23	42	10	10	0	1	43	3760	NULL
0	0	0	0	0	0	0	89	NULL
0	1	1	1	0	0	1	39	NULL
9	26	8	8	0	0	26	391	NULL

data	cod_reg	regione	ricov_sintomi	terapia_intensiva	totale_ospedal	isolamento_domiciliare	totale_pos	delta_tot_pos	nuovi_pos	dimessi_guariti	deceduti	totale_casi	tamponi	casi_testati
2020-04-19 17:00:00	5	Veneto	1264	184	1448	8762	10210	-234	243	4638	1087	15935	255797	161231
2020-04-19 17:00:00	6	Friuli Venez...	140	25	165	1172	1337	-66	14	1183	225	2745	44622	27544
2020-04-19 17:00:00	7	Liguria	885	101	986	2504	3490	-78	227	2110	928	6528	31651	21103
2020-04-19 17:00:00	8	Emilia-Rom...	3166	289	3455	10097	13552	-32	376	9985	3023	23560	104916	88776
2020-04-19 17:00:00	9	Toscana	852	192	1044	5462	6496	26	135	1239	637	8372	103975	84620
2020-04-19 17:00:00	10	Umbria	108	30	138	298	436	5	4	854	58	1348	25170	16790
2020-04-19 17:00:00	11	Marche	799	86	885	2297	3182	10	48	1780	807	5769	41474	27099
2020-04-19 17:00:00	12	Lazio	1370	185	1555	2766	4321	39	87	1093	341	5755	91807	22702
2020-04-19 17:00:00	13	Abruzzo	317	40	357	1630	1987	16	34	276	258	2521	27791	22424
2020-04-19 17:00:00	14	Molise	27	4	31	184	215	6	10	47	17	279	3584	3482
2020-04-19 17:00:00	15	Campania	590	61	651	2371	3022	-23	41	703	304	4029	48167	34994
2020-04-19 17:00:00	16	Puglia	590	60	650	2136	2786	92	120	427	316	3539	42598	41580
2020-04-19 17:00:00	17	Basilicata	59	8	67	180	247	-15	3	71	24	342	6528	6528
2020-04-19 17:00:00	18	Calabria	142	6	148	696	844	12	24	116	75	1035	23760	22234
2020-04-19 17:00:00	19	Sicilia	522	41	563	1639	2202	31	45	315	200	2717	49772	49772
2020-04-19 17:00:00	20	Sardegna	117	22	139	725	864	-17	17	265	86	1215	14859	13760
2020-04-20 17:00:00	1	Piemonte	3299	301	3600	10957	14557	87	292	4383	2409	21349	99669	73192
2020-04-20 17:00:00	2	Valle d'Aosta	102	8	110	438	548	-14	0	413	127	1088	4799	3910
2020-04-20 17:00:00	3	Lombardia	10139	901	11039	23548	34597	90	735	20008	12376	66991	270486	173333
2020-04-20 17:00:00	4	Trentino-AL	432	59	491	2978	3469	-68	72	1900	616	5984	57592	32436
2020-04-20 17:00:00	5	Veneto	1273	180	1453	8608	10061	-149	192	4954	1112	16727	260810	163757
2020-04-20 17:00:00	6	Friuli Venez...	140	22	162	1028	1190	-147	30	1346	239	2775	45161	28777
2020-04-20 17:00:00	7	Liguria	878	102	980	2516	3496	6	141	2216	957	6669	33007	21979
2020-04-20 17:00:00	8	Emilia-Rom...	3144	283	3427	10095	13522	-30	307	6266	3079	22867	129530	90894
2020-04-20 17:00:00	9	Toscana	850	182	1032	5536	6568	72	135	1272	667	8507	105857	85946
2020-04-20 17:00:00	10	Umbria	113	29	142	282	424	-12	1	967	88	1348	25552	17099
2020-04-20 17:00:00	11	Marche	792	80	872	2340	3212	30	57	1792	622	5826	42782	28006
2020-04-20 17:00:00	12	Lazio	1388	185	1573	2792	4365	44	60	1101	349	5815	97823	23067
2020-04-20 17:00:00	13	Abruzzo	320	38	358	1704	2062	75	91	287	263	2612	27942	22552
2020-04-20 17:00:00	14	Molise	27	2	29	184	213	-2	2	50	18	281	3783	3683
2020-04-20 17:00:00	15	Campania	586	61	647	2372	3019	-3	45	746	309	4074	51090	36153
2020-04-20 17:00:00	16	Puglia	584	66	650	2160	2810	24	38	431	326	3567	44189	43119
2020-04-20 17:00:00	17	Basilicata	59	7	66	176	242	-5	0	76	24	342	6868	6868
2020-04-20 17:00:00	18	Calabria	140	7	147	681	828	-16	3	135	75	1038	24573	22647
2020-04-20 17:00:00	19	Sicilia	526	39	565	1645	2210	8	42	346	203	2789	51373	51373
2020-04-20 17:00:00	20	Sardegna	115	21	136	718	854	-10	13	288	86	1228	15328	14160

Figura 4: Tabella data regioni

```

1 #insert province
2 insert
3 into
4 province(cod_reg , regione, cod_prov , provincia,
5 sigla_prov, densità_abitanti_prov ) (
6 select
7 distinct m.cod_reg , m.regione, m.cod_prov ,
8 m.provincia, m.sigla_prov, m.densità_abitanti_prov
9 from
10 master m
11 where
12 m.cod_prov <> 0 );
13

```

cod_reg	regione	cod_prov	provincia	sigla_prov	densità_abitanti_prov
1	Piemonte	1	Torino	TO	331
1	Piemonte	2	Vercelli	VC	82
1	Piemonte	3	Novara	NO	275
1	Piemonte	4	Cuneo	CN	85
1	Piemonte	5	Asti	AT	142
1	Piemonte	6	Alessandria	AL	118
2	Valle d'Aosta	7	Aosta	AO	39
7	Liguria	8	Imperia	IM	185
7	Liguria	9	Savona	SV	179
7	Liguria	10	Genova	GE	459
7	Liguria	11	La Spezia	SP	249
3	Lombardia	12	Varese	VA	743
3	Lombardia	13	Como	CO	468
3	Lombardia	14	Sondrio	SO	57
3	Lombardia	15	Milano	MI	2063
3	Lombardia	16	Bergamo	BG	485
3	Lombardia	17	Brescia	BS	265
3	Lombardia	18	Pavia	PV	184
3	Lombardia	19	Cremona	CR	203
3	Lombardia	20	Mantova	MN	176
4	Trentino-Alto Adige	21	Bolzano	BZ	72
4	Trentino-Alto Adige	22	Trento	TN	87
5	Veneto	23	Verona	VR	299
5	Veneto	24	Vicenza	VI	317
5	Veneto	25	Belluno	BL	66
5	Veneto	26	Treviso	TV	358
5	Veneto	27	Venezia	VE	345

16	Puglia	71	Foggia	FG	89
16	Puglia	72	Bari	BA	324
16	Puglia	73	Taranto	TA	234
16	Puglia	74	Brindisi	BR	211
16	Puglia	75	Lecce	LE	284
17	Basilicata	76	Potenza	PZ	85
17	Basilicata	77	Matera	MT	57
18	Calabria	78	Cosenza	CS	105
18	Calabria	79	Catanzaro	CZ	148
18	Calabria	80	Reggio di Calabria	RC	171
19	Sicilia	81	Trapani	TP	174
19	Sicilia	82	Palermitano	PA	250
19	Sicilia	83	Messina	ME	192
19	Sicilia	84	Agrigento	AG	142
19	Sicilia	85	Caltanissetta	CL	123
19	Sicilia	86	Enna	EN	84
19	Sicilia	87	Catania	CT	310
19	Sicilia	88	Ragusa	RG	198
19	Sicilia	89	Siracusa	SR	188
20	Sardegna	90	Sassari	SS	64
20	Sardegna	91	Nuoro	NU	37
20	Sardegna	92	Cagliari	CA	345
6	Friuli Venezia Giulia	93	Pordenone	PN	137
14	Molise	94	Isernia	IS	55
20	Sardegna	95	Oristano	OR	53
1	Piemonte	96	Bielle	BI	192
3	Lombardia	97	Lecco	LC	419
3	Lombardia	98	Lodi	LO	294
8	Emilia-Romagna	99	Rimini	RN	392

Figura 5: Tabella province

```

1 #insert dataprov
2 insert into dataprov ('data', cod_reg, regione, cod_prov, provincia, sigla_prov, tot_casi_prov)
3 (select distinct m.'data',
4 m.cod_reg,
5 m.regione,
6 m.cod_prov,
7 m.provincia,
8 m.sigla_prov,
9 m.tot_casi_prov
10 from 'master' m
11 where m.cod_prov <> 0 );
12

```

data	cod_reg	regione	cod_prov	provincia	sigla_prov	tot_casi_prov
2020-03-05 17:00:00	3	Lombardia	13	Como	CO	11
2020-03-05 17:00:00	3	Lombardia	14	Sondrio	SO	4
2020-03-05 17:00:00	3	Lombardia	15	Milano	MI	197
2020-03-05 17:00:00	3	Lombardia	16	Bergamo	BG	537
2020-03-05 17:00:00	3	Lombardia	17	Brescia	BS	155
2020-03-05 17:00:00	3	Lombardia	18	Pavia	PV	151
2020-03-05 17:00:00	3	Lombardia	19	Cremona	CR	406
2020-03-05 17:00:00	3	Lombardia	20	Mantova	MN	26
2020-03-05 17:00:00	4	Trentino-Alto Adige	21	Bolzano	BZ	1
2020-03-05 17:00:00	4	Trentino-Alto Adige	22	Trento	TN	7
2020-03-05 17:00:00	5	Veneto	23	Verona	VR	25
2020-03-05 17:00:00	5	Veneto	24	Vicenza	VI	19
2020-03-05 17:00:00	5	Veneto	25	Belluno	BL	7
2020-03-05 17:00:00	5	Veneto	26	Treviso	TV	89
2020-03-05 17:00:00	5	Veneto	27	Venezia	VE	73
2020-03-05 17:00:00	5	Veneto	28	Padova	PD	175
2020-03-05 17:00:00	5	Veneto	29	Rovigo	RO	5
2020-03-05 17:00:00	6	Friuli Venezia Giulia	30	Udine	UD	11
2020-03-05 17:00:00	6	Friuli Venezia Giulia	31	Gorizia	GO	5
2020-03-05 17:00:00	6	Friuli Venezia Giulia	32	Trieste	TS	5
2020-03-05 17:00:00	8	Emilia-Romagna	33	Piacenza	PC	378
2020-03-05 17:00:00	8	Emilia-Romagna	34	Parma	PR	150

Figura 6: Tabella data province

```

1 --insert aeroporti--
2 INSERT INTO aeroporti
3 (SELECT m.'data', m.aeroporto, m.regione, m.cod_reg, m.provincia, m.cod_prov,
4  m.passeggeri_nazionali, m.passeggeri_internazionali
5  FROM 'master' m
6  WHERE m.passeggeri_nazionali <> 0
7        AND m.passeggeri_internazionali <> 0);
8

```

data	nome	regione
2020-02-29 17:00:00	Torino	Piemonte
2020-02-29 17:00:00	Cuneo	Piemonte
2020-02-29 17:00:00	Genova	Liguria
2020-02-29 17:00:00	Milano Linate/Malpensa	Lombardia
2020-02-29 17:00:00	Bergamo	Lombardia
2020-02-29 17:00:00	Verona	Veneto
2020-02-29 17:00:00	Treviso	Veneto
2020-02-29 17:00:00	Venezia	Veneto
2020-02-29 17:00:00	Trieste	Friuli Venezia Giulia
2020-02-29 17:00:00	Bologna	Emilia-Romagna
2020-02-29 17:00:00	Ancona	Marche
2020-02-29 17:00:00	Firenze	Toscana
2020-02-29 17:00:00	Pisa	Toscana
2020-02-29 17:00:00	Perugia	Umbria
2020-02-29 17:00:00	Roma Ciampino/Fiumicino	Lazio
2020-02-29 17:00:00	Napoli	Campania
2020-02-29 17:00:00	Pescara	Abruzzo
2020-02-29 17:00:00	Bari	Puglia
2020-02-29 17:00:00	Brindisi	Puglia
2020-02-29 17:00:00	Lamezia Terme	Calabria

cod_reg	provincia	cod_prov	pass_nazionali	pass_internazionali
1	Torino	1	127075	186016
1	Cuneo	4	8316	1686
7	Genova	10	44860	22386
3	Milano	15	564293	1630760
3	Bergamo	16	208465	698201
5	Verona	23	60205	127475
5	Treviso	26	81767	126640
5	Venezia	27	81660	529351
6	Trieste	32	27979	17102
8	Bologna	37	122076	468151
11	Ancona	42	3677	19192
9	Firenze	48	23407	144370
9	Pisa	50	98798	166507
10	Perugia	54	5279	8793
12	Roma	58	2169362	666005
15	Napoli	63	242706	360027
13	Pescara	68	18212	17553
16	Bari	72	202675	120511
16	Brindisi	74	121122	27985
18	Catanzaro	79	136771	8212

2020-02-29 17:00:00	Torino	Piemonte	1	Torino	1	127075	186016
2020-02-29 17:00:00	Cuneo	Piemonte	1	Cuneo	4	8316	1686
2020-02-29 17:00:00	Genova	Liguria	7	Genova	10	44860	22386
2020-02-29 17:00:00	Milano Linate/Malpensa	Lombardia	3	Milano	15	564293	1630760
2020-02-29 17:00:00	Bergamo	Lombardia	3	Bergamo	16	208465	698201
2020-02-29 17:00:00	Verona	Veneto	5	Verona	23	60205	127475
2020-02-29 17:00:00	Treviso	Veneto	5	Treviso	26	81767	126640
2020-02-29 17:00:00	Venezia	Veneto	5	Venezia	27	81660	529351
2020-02-29 17:00:00	Trieste	Friuli Venezia Giulia	6	Trieste	32	27979	17102
2020-02-29 17:00:00	Bologna	Emilia-Romagna	8	Bologna	37	122076	468151
2020-02-29 17:00:00	Ancona	Marche	11	Ancona	42	3677	19192
2020-02-29 17:00:00	Firenze	Toscana	9	Firenze	48	23407	144370
2020-02-29 17:00:00	Pisa	Toscana	9	Pisa	50	98798	166507
2020-02-29 17:00:00	Perugia	Umbria	10	Perugia	54	5279	8793
2020-02-29 17:00:00	Roma Ciampino/Fiumicino	Lazio	12	Roma	58	2169362	666005
2020-02-29 17:00:00	Napoli	Campania	15	Napoli	63	242706	360027
2020-02-29 17:00:00	Pescara	Abruzzo	13	Pescara	68	18212	17553
2020-02-29 17:00:00	Bari	Puglia	16	Bari	72	202675	120511
2020-02-29 17:00:00	Brindisi	Puglia	16	Brindisi	74	121122	27985
2020-02-29 17:00:00	Lamezia Terme	Calabria	18	Catanzaro	79	136771	8212
2020-02-29 17:00:00	Trapani	Sicilia	19	Trapani	81	20190	6
2020-02-29 17:00:00	Palermo	Sicilia	19	Palermo	82	297727	75100

Figura 7: Tabella aeroporti

```

1 #insert veicoli
2 INSERT INTO veicoli
3 (SELECT m.'data', m.'regione', m.'cod_reg', m.'flusso_veicoli'
4 FROM 'master' m
5 WHERE m.'flusso_veicoli'<>0);
6

```

data	regione		cod_reg	traffico
2020-02-29 17:00:00	Lazio	➔	12	56902
2020-02-29 17:00:00	Abruzzo	➔	13	6920
2020-02-29 17:00:00	Molise	➔	14	4579
2020-02-29 17:00:00	Campania	➔	15	9356
2020-02-29 17:00:00	Puglia	➔	16	14106
2020-02-29 17:00:00	Basilicata	➔	17	5620
2020-02-29 17:00:00	Calabria	➔	18	7196
2020-02-29 17:00:00	Sicilia	➔	19	5689
2020-02-29 17:00:00	Sardegna	➔	20	7120
2020-03-31 17:00:00	Piemonte	➔	1	2560
2020-03-31 17:00:00	Valle d'Aosta	➔	2	680

Figura 8: Tabella veicoli

```

1 #insert usi
2 INSERT INTO usi
3 (SELECT m.'data', m.'regione', m.'cod_reg', m.'variazione_acquisti_attrezzi',
4 m.'variazione_acquisti_libri', m.'variazione_uso_netflix', m.'variazione_uso_amazon',
5 m.'variazione_richiesta_fibra_tim', m.'variazione_richiesta_fibra_wind3'
6 FROM 'master' m
7 WHERE m.'variazione_acquisti_attrezzi'<> 0 );
8

```

data	regione	cod_reg	variazione_richiesta_fibra_tim	variazione_richiesta_fibra_wind3	variazione_acquisti_libri	variazione_uso_netflix	variazione_uso_amazon
2020-05-03 17:00:00	Piemonte	➔ 1	0.10	0.02	0.06	0.30	0.14
2020-05-03 17:00:00	Valle d'Aosta	➔ 2	-0.26	0.10	0.06	0.30	0.16
2020-05-03 17:00:00	Lombardia	➔ 3	0.14	0.12	0.08	0.30	0.18
2020-05-03 17:00:00	Trentino-Alto Adige	➔ 4	0.48	0.12	0.10	0.24	0.16
2020-05-03 17:00:00	Veneto	➔ 5	0.16	0.08	0.08	0.26	0.16
2020-05-03 17:00:00	Friuli Venezia Giulia	➔ 6	0.06	0.06	0.06	0.24	0.18
2020-05-03 17:00:00	Liguria	➔ 7	0.12	0.12	0.08	0.28	0.14
2020-05-03 17:00:00	Emilia-Romagna	➔ 8	0.20	0.12	0.08	0.30	0.16
2020-05-03 17:00:00	Toscana	➔ 9	0.16	0.08	0.10	0.32	0.16
2020-05-03 17:00:00	Umbria	➔ 10	0.04	0.12	0.08	0.32	0.12
2020-05-03 17:00:00	Marche	➔ 11	0.22	0.06	0.10	0.36	0.14
2020-05-03 17:00:00	Lazio	➔ 12	0.12	0.08	0.10	0.30	0.16
					0.12	0.38	0.14
							0.10

data	regione	cod_reg	variazione_acquisti_attrezzi	variazione_acquisti_libri	variazione_uso_netflix	variazione_uso_amazon	variazione_richiesta_fibra_tim	variazione_richiesta_fibra_wind3
2020-05-03 17:00:00	Piemonte	➔ 1	0.46	0.06	0.30	0.14	0.10	0.02
2020-05-03 17:00:00	Valle d'Aosta	➔ 2	0.32	0.06	0.30	0.16	-0.26	0.10
2020-05-03 17:00:00	Lombardia	➔ 3	0.48	0.08	0.30	0.18	0.14	0.12
2020-05-03 17:00:00	Trentino-Alto Adige	➔ 4	0.56	0.10	0.24	0.16	0.48	0.12
2020-05-03 17:00:00	Veneto	➔ 5	0.52	0.08	0.26	0.16	0.16	0.08
2020-05-03 17:00:00	Friuli Venezia Giulia	➔ 6	0.46	0.06	0.24	0.18	0.06	0.06
2020-05-03 17:00:00	Liguria	➔ 7	0.54	0.08	0.28	0.14	0.12	0.12
2020-05-03 17:00:00	Emilia-Romagna	➔ 8	0.40	0.08	0.30	0.16	0.20	0.12
2020-05-03 17:00:00	Toscana	➔ 9	0.58	0.10	0.32	0.16	0.18	0.08
2020-05-03 17:00:00	Umbria	➔ 10	0.52	0.08	0.32	0.12	0.04	0.12
2020-05-03 17:00:00	Marche	➔ 11	0.60	0.10	0.36	0.14	0.22	0.06
2020-05-03 17:00:00	Lazio	➔ 12	0.58	0.10	0.30	0.16	0.12	0.08
2020-05-03 17:00:00	Abruzzo	➔ 13	0.56	0.12	0.38	0.14	0.14	0.08
2020-05-03 17:00:00	Molise	➔ 14	0.50	0.12	0.38	0.10	0.02	-0.02
2020-05-03 17:00:00	Campania	➔ 15	0.46	0.12	0.44	0.14	0.13	0.02
2020-05-03 17:00:00	Puglia	➔ 16	0.52	0.14	0.42	0.10	0.04	0.06
2020-05-03 17:00:00	Basilicata	➔ 17	0.70	0.12	0.36	0.08	-0.10	0.16

Figura 9: Tabella usi

4 | QUERY

Innanzitutto si è cercato di comprendere l'entità della diffusione del virus. Quindi si è analizzato il totale dei positivi in ciascuna regione:

```

1 SELECT
2     dr.'regione',
3     dr.'cod_reg',
4     dr.'totale_pos',
5     (dr.'totale_pos'/sommapo.'totpo') AS 'Tasso positivi regionale'
6 FROM 'datareg' dr,
7     ( SELECT
8         SUM(dr1.'totale_pos') AS 'totpo'
9     FROM
10         'datareg' dr1
11     WHERE
12         dr1.'data' LIKE '2020-05-03%') AS sommapo
13 WHERE dr.'DATA' LIKE '2020-05-03%'
14 ORDER BY dr.'totale_pos' DESC;

```

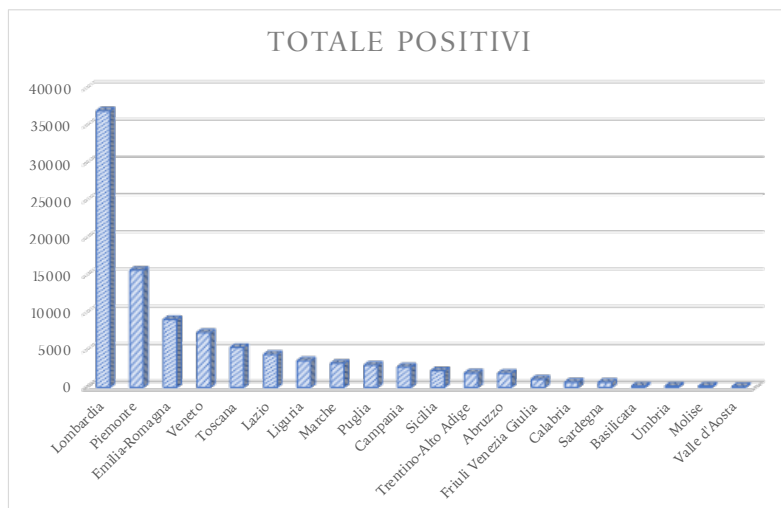


Figura 10: Andamento totale positivi

Come è possibile notare dai grafici, le regioni del centro nord hanno registrato un numero maggiore di positivi, con numeri molto elevati per la Lombardia. Nello specifico all'interno di essa è stato registrato quasi il 40% dei contagi nazionali. Questo dato è comunque da contestualizzare sia con il numero di tamponi effettuato dalla medesima regione, che dalla densità di abitanti

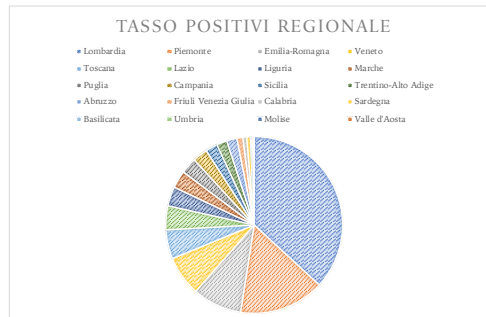


Figura 11: Tasso positivi

della stessa. È infatti plausibile aspettarsi che regioni con un numero elevato di tamponi ed un'elevata densità abitativa possano registrare un maggior numero di contagi.

```

1 SELECT
2   dr.'regione',
3   dr.'totale_pos',
4   dr.'tamponi',
5   (dr.'totale_pos'/dr.'tamponi') AS 'Rapporto positivi/tamponi',
6   r.'densità_abitanti',
7   (dr.'totale_pos'/r.'densità_abitanti') AS 'Rapporto positivi/densità abitanti'
8 FROM 'datareg' dr,
9      'regioni' r
10 WHERE dr.'DATA' LIKE '2020-05-03%'
11 AND r.'cod_reg'=dr.'cod_reg'
12 ORDER BY (dr.'totale_pos'/dr.'tamponi') DESC;
```

Regione	Positivi	Tamponi	Positivi/tamponi	Densità abitanti	Positivi/densità abitanti
Piemonte	15638	172208	9,08%	172	90,92
Lombardia	36926	410857	8,99%	422	87,50
Liguria	3551	54492	6,52%	286	12,42
Marche	3198	64412	4,96%	162	19,74
Emilia-Romagna	9045	197075	4,59%	199	45,45
Abruzzo	1868	40699	4,59%	121	15,44
Puglia	2955	66443	4,45%	206	14,34
Toscana	5328	150914	3,53%	162	32,89
Campania	2726	86498	3,15%	424	6,43
Lazio	4385	150912	2,91%	341	12,86
Molise	181	7075	2,56%	69	2,62
Sicilia	2203	85955	2,56%	194	11,36
Sardegna	689	27737	2,48%	68	10,13
Trentino-Alto Adige	1912	85335	2,24%	79	24,20
Veneto	7299	378202	1,93%	267	27,34
Calabria	702	38835	1,81%	128	5,48
Friuli Venezia Giulia	1087	74990	1,45%	153	7,10
Basilicata	194	14210	1,37%	56	3,46
Valle d'Aosta	109	8100	1,35%	39	2,79
Umbria	183	38823	0,47%	104	1,76

Tabella 1: Correlazione numero positivi e densità abitativa

Dalla tabella possono essere effettuate diverse considerazioni. Sep-
pur i numeri registrati in Lombardia siano tragici, essa non risulta
l'unica regione ad avere un tasso di positività pari al 9% in quanto de-
tiene tale "record" insieme con il Piemonte, che tra l'altro ha registrato
un tasso di circa 91 positivi per ogni km². Altro aspetto singolare è
il seguente: le regioni del centro-nord hanno registrato una positività
maggiore rispetto a quelle del centro-sud. Tra l'altro correlare la dif-
fusione dei contagi con l'estensione del territorio appare fuorviante:
si evince infatti dalla tabella che un'elevata densità di abitanti non im-
plica strettamente un elevato numero di contagi. Esempio è il caso
della Campania che, pur avendo la maggiore densità di abitanti, ha
registrato solo 6 positivi per km² con un tasso di positività rispetto ai
tamponi effettuati pari solo al 3%. Discorso analogo può essere esteso
alla regione Lazio i cui dati di contagio risultano contenuti rispetto
a quanto ci si sarebbe potuto aspettare dai dati demografici che la
contraddistinguono. Quindi appare lecito pensare che vi siano state
diverse motivazioni a sostegno della differente modalità di diffusio-
ne del virus. Tra queste, certamente, la prontezza con cui le regioni

hanno fatto rispettare le misure di contenimento governative o, come accaduto in Campania, la capacità di farsi precursori di tale direttive. Se, infatti, si analizza l'andamento dei contagi tra le prime tre regioni per tassi di positività e quelle della Campania e della Calabria caratterizzate da tassi molto bassi, è possibile osservare che, nelle regioni del sud, la curva epidemiologica tenda asintoticamente ad essere stabile e ad estinguersi nel tempo. Differente è il discorso delle altre caratterizzate invece da un andamento ancora crescente:

```

1 SELECT
2     dr.`data`,
3     dr.`regione`,
4     dr.`totale_casi`
5 FROM
6     `datareg` dr
7 WHERE
8     dr.`regione`='Lombardia'
9 OR dr.`regione`='Piemonte'
10 OR dr.`regione`='Campania'
11 OR dr.`regione`='Liguria'
12 OR dr.`regione`='Calabria';

```

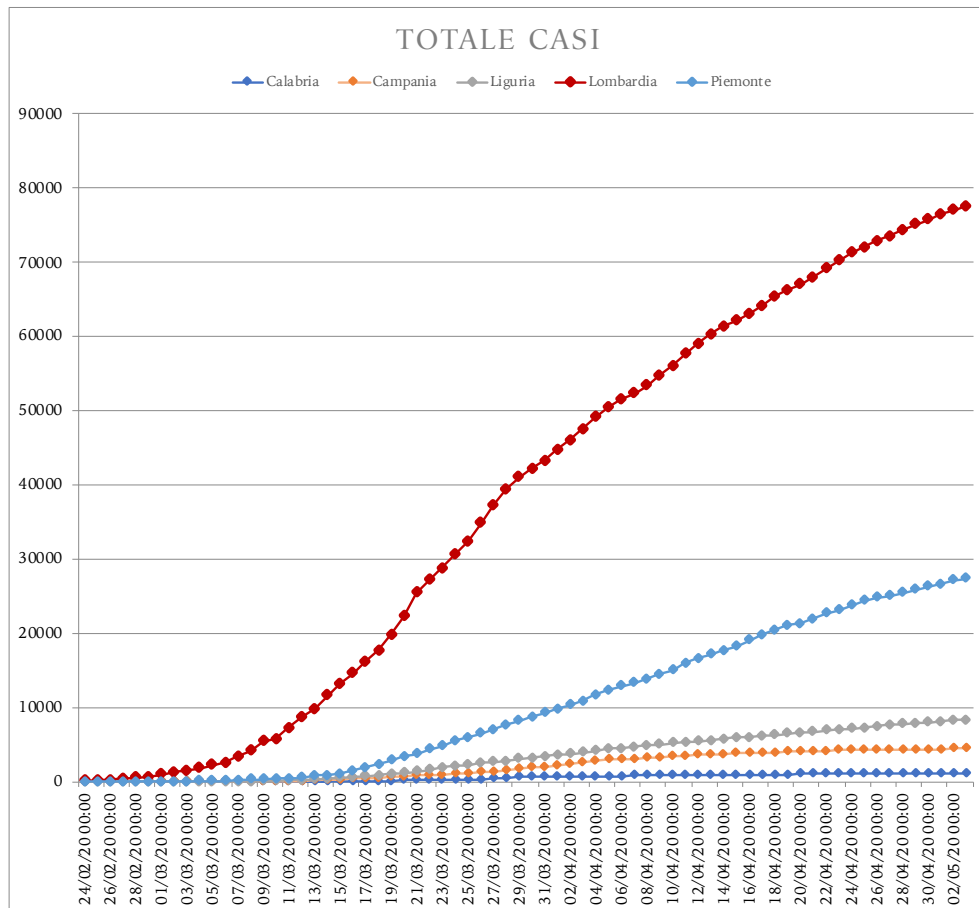


Figura 12: Andamento curva epidemiologica

Si nota come, sebbene il trend nelle regioni del Piemonte e della Lombardia sia in continua crescita, essa non segua più un andamento esponenziale, ma al più lineare o polinomiale avvicinandosi, in en-

trambi i casi, ad un flesso sintomatico di un'inversione di tendenza. Appare opportuno chiedersi, dato il numero di casi elevato, quale sia stata la gravità dei pazienti ricoverati, valutando quanti siano stati i ricoveri in terapia intensiva e la capacità delle strutture sanitarie di fronteggiare l'emergenza. Nello specifico è stata effettuata la seguente richiesta¹:

```

1 SELECT
2   dr.'data',
3   dr.'regione',
4   dr.'terapia_intensiva',
5   r.'posti_intensiva_reg'
6 FROM
7   'datereg' dr,
8   'regioni' r
9 WHERE
10  dr.'cod_reg'=r.'cod_reg'
11  AND dr.'regione'='Campania';

```

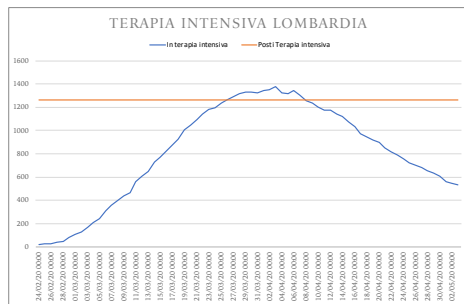


Figura 13: Lombardia



Figura 14: Piemonte

Si nota come la regione Campania sia stata in grado di fronteggiare, lestamente, l'emergenza sanitaria, garantendo in tutta la regione, le opportune cure di cui necessitavano i pazienti che versavano in gravi condizioni di salute. Questo è stato, di sicuro, un elemento che ha garantito sicuramente anche un numero di decessi molto basso all'interno della stessa regione, soprattutto a confronto dei numeri elevati fatti registrare nelle regioni del nord e, particolarmente, in Lombardia particolarmente.



Figura 15: Campania

¹ Si riporta il codice per la sola Campania in quanto per le diverse regioni varia appunto solo il nome delle stesse.

```

1 SELECT
2   dr.'data',
3   dr.'regione',
4   dr.'deceduti',
5   (dr.deceduti/totde.de) AS 'Deceduti reg/naz'
6 FROM
7   'datareg' dr,
8   (SELECT SUM(drl.'deceduti') AS 'de'
9    FROM 'datareg' drl
10   WHERE drl.'data' LIKE '2020-05-03%') AS totde
11 WHERE
12   dr.'DATA' LIKE '2020-05-03%'
13 ORDER BY
14   dr.'deceduti' DESC;

```

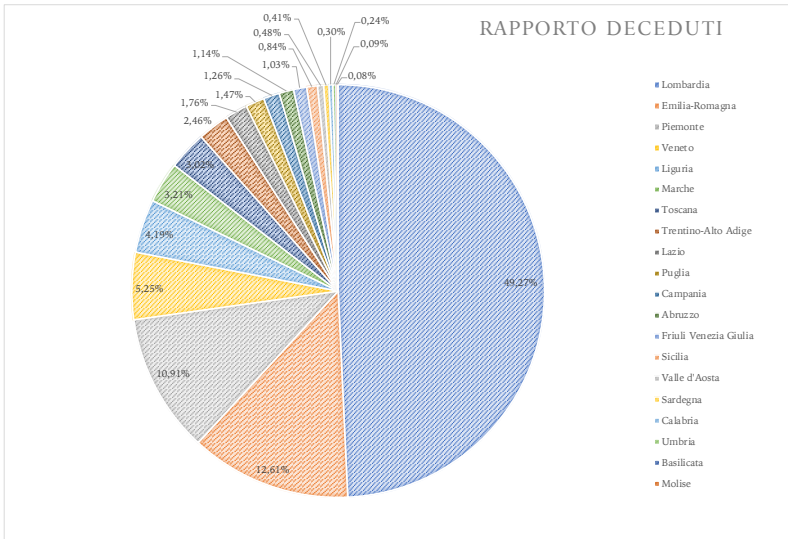


Figura 16: Ripartizione deceduti

È possibile notare come più dell'80% dei decessi sia avvenuto nelle regioni del nord, con la metà dei decessi nazionali registrati nella sola Lombardia. Sorge lecito, quindi, indagare sulle cause che hanno determinato questa disparità. Se da un lato potenzialmente essa può essere attribuita ad un sistema di sanità locale che, come visto, non ha saputo fronteggiare adeguatamente la situazione con posti in terapia intensiva inferiori rispetto a quanto necessario, dall'altro ciò potrebbe essere imputato a varie cause, tra cui è possibile annoverare un'elevata età media della popolazione dato il target di età dei de-

Regione	Deceduti	Deceduti reg/naz
Lombardia	14231	49,27%
Emilia-Romagna	3642	12,61%
Piemonte	3152	10,91%
Veneto	1516	5,25%
Liguria	1209	4,19%
Marche	927	3,21%
Toscana	872	3,02%
Trentino-Alto Adige	710	2,46%
Lazio	508	1,76%
Puglia	424	1,47%
Campania	364	1,26%
Abruzzo	330	1,14%
Friuli Venezia Giulia	297	1,03%
Sicilia	242	0,84%
Valle d'Aosta	138	0,48%
Sardegna	119	0,41%
Calabria	88	0,30%
Umbria	68	0,24%
Basilicata	25	0,09%
Molise	22	0,08%

Tabella 2: Totale deceduti

ceduti, un cospicuo numero di case di riposo, un elevato numero di accessi all'interno della regione o tipologie di cure dallo scarso impatto da un punto di vista clinico. Si è provveduto, in primis, a valutare l'età media regionale. Ci si aspetterebbe, quindi, che le regioni aventi una popolazione di età media superiore, abbiano registrato elevati tassi di deceduti.

```

1 SELECT
2   r.'regione',
3   r.'età_media'
4 FROM
5   'regioni' r
6 ORDER BY
7   r.'età_media' DESC;

```

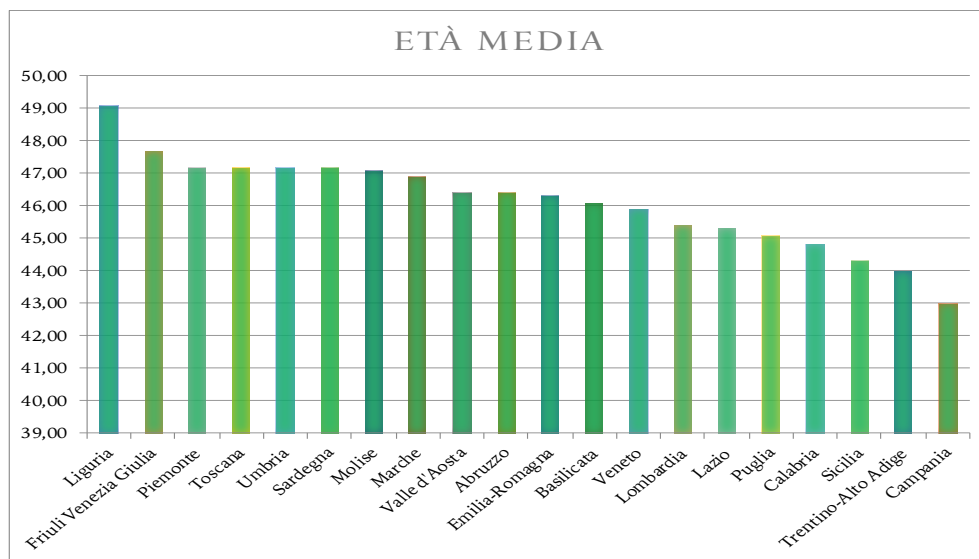


Figura 17: Età media regioni italiane

Dal grafico si evince che sarebbe fuorviante ritenere l'età media elevata una variabile preponderante per una maggiore veicolazione del virus. Basti pensare, infatti, che regioni come l'Emilia-Romagna e la Lombardia hanno un'età media inferiore rispetto al Molise, il quale ha registrato il minor tasso di deceduti su scala nazionale.

Accertata, pertanto, l'impossibilità di correlare il numero dei decessi con l'età media elevata, tenendo comunque presente che il numero maggiore di decessi è stato registrato per persone over 65, ha senso indagare, piuttosto, sul numero di case di riposo presenti in ogni regione. In tal caso sarebbe d'uopo pensare che le regioni con un numero maggiore di case di riposo, siano state più esposte ai decessi di persone anziane che, tramite le strutture adibite ad ospitarle, hanno avuto l'opportunità di interfacciarsi con terzi provenienti dall'esterno delle strutture, configurandosi, essi stessi, come veicolo principale di trasporto e diffusione del virus. Purtroppo eventuali patologie già accertate, la fragilità dell'età ed un sistema di terapia intensiva regionale

al collasso, hanno fornito un canale preferenziale nel determinare il conseguente decesso.

```

1 SELECT
2   r.'regione',
3   r.'età_media',
4   r.'case_riposo',
5   dr.'deceduti'
6 FROM
7   'regioni' r,
8   'datareg' dr
9 WHERE
10  r.'cod_reg'=dr.'cod_reg'
11 AND dr.'data' LIKE '2020-05-03%'
12 ORDER BY
13  r.'case_riposo' DESC;

```

Regione	Età media	Case riposo	Deceduti
Lombardia	45.4	678	14231
Piemonte	47.2	608	3152
Veneto	45.9	520	1516
Emilia-Romagna	46.3	348	3642
Toscana	47.2	319	872
Lazio	45.3	207	508
Campania	43.0	121	364
Liguria	49.1	120	1209
Calabria	44.8	94	88
Friuli Venezia Giulia	47.7	72	297
Puglia	45.1	61	424
Trentino-Alto Adige	44.0	54	710
Marche	46.9	51	927
Umbria	47.2	50	68
Sicilia	44.3	39	242
Abruzzo	46.4	16	330
Sardegna	47.2	16	119
Molise	47.1	6	22
Valle d'Aosta	46.4	2	138
Basilicata	46.1	1	25

Tabella 3: Correlazione case riposo e decessi

Dalla tabella si evince, infatti, che i focolai scoppiati nelle case di riposo potrebbero aver contribuito in maniera rilevante all'incremento del numero di decessi registrati, in quanto solo alle prime 4 regioni, per numero di case di riposo disponibili sul territorio, è possibile imputare l'80% del totale dei decessi registrati. Questo evidenzia, probabilmente, l'assenza di adeguate misure di tutela nei confronti dei centri ospitanti anziani, senza opportune misure di controllo e senza un'adeguata limitazione degli accessi dall'esterno.

Altra causa, che potrebbe aver influito pesantemente sull'elevato numero di morti, è il tipo di cura con cui sono stati trattati i pazienti in gravi condizioni respiratorie. Infatti, non essendo presente un vaccino né una cura riconosciuta come adeguata a fronteggiare la nuo-

va epidemia, diversi sono stati i farmaci con cui sono stati trattati i pazienti ed altrettanto differenti sono stati i risultati prodotti.

```
1 SELECT
2   r.'regione',
3   r.'cura'
4 FROM
5   'regioni' r
6 ORDER BY
7   r.'cura';
```

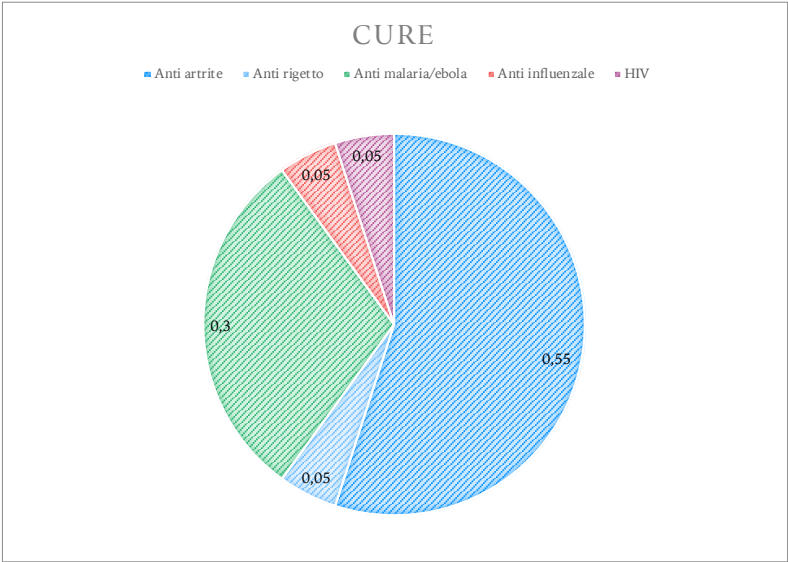


Figura 18: Famiglia di farmaci

I dati raccolti si riferiscono al tipo di cura sperimentale cui hanno aderito, almeno in un primo momento, le diverse regioni. Per poter avere riscontro dell’efficacia di tale farmaco sulla salute dei pazienti, è opportuno correlare la data di adozione delle diverse tipologie cura con il picco dei decessi. Se infatti il picco dei decessi si è verificato a distanza di 2 o 3 giorni dal ricorso ad un determinato farmaco, probabilmente ciò può essere correlato ad un benevolo impatto dello stesso sulle condizioni di salute dei ricoverati. Qualora invece facendo ricorso ad una determinata tipologia di cura, non fosse stato registrato in un breve periodo di tempo il picco di decessi, probabilmente ciò sarebbe sintomatico di una scarsa efficacia della cura testata.

Regione	Cura
Veneto	Avigan
Emilia-Romagna	Idrossiclorochina
Abruzzo	Idrossiclorochina
Basilicata	Idrossiclorochina
Lazio	Lopinavir/Ritonavir
Piemonte	Plaquenil
Valle d’Aosta	Remdesivir
Friuli Venezia Giulia	Remdesivir
Lombardia	Remdesivir/Idrossiclorochina
Molise	Roacterna
Toscana	Ruxolitinib
Trentino-Alto Adige	Tocilizumab
Liguria	Tocilizumab
Umbria	Tocilizumab
Marche	Tocilizumab
Campania	Tocilizumab
Puglia	Tocilizumab
Calabria	Tocilizumab
Sicilia	Tocilizumab
Sardegna	Tocilizumab

Tabella 4: Farmaci testati

```
1 SELECT
2   r.'data_cura' AS 'Data cura',
3   DELTA.'regione',
4   DELTA.'data' AS 'Data Picco',
5   MAX(DELTA.Differenza) AS Picco
6 FROM
7   'regioni' r,
8   (SELECT
9     dr.'data',
10    dr.'regione',
11    dr.'deceduti',
12    dr.'deceduti'-LAG (dr.'deceduti') OVER (PARTITION BY dr.'regione') AS Differenza
13   FROM
14     'datareg' dr
15   ORDER BY
16    dr.'deceduti'-LAG (dr.'deceduti') OVER (PARTITION BY dr.'regione') DESC) AS DELTA
17 WHERE
18   r.'regione'=DELTA.'regione'
19 GROUP BY
20   DELTA.regione
21 ORDER BY
22   Picco DESC;
```

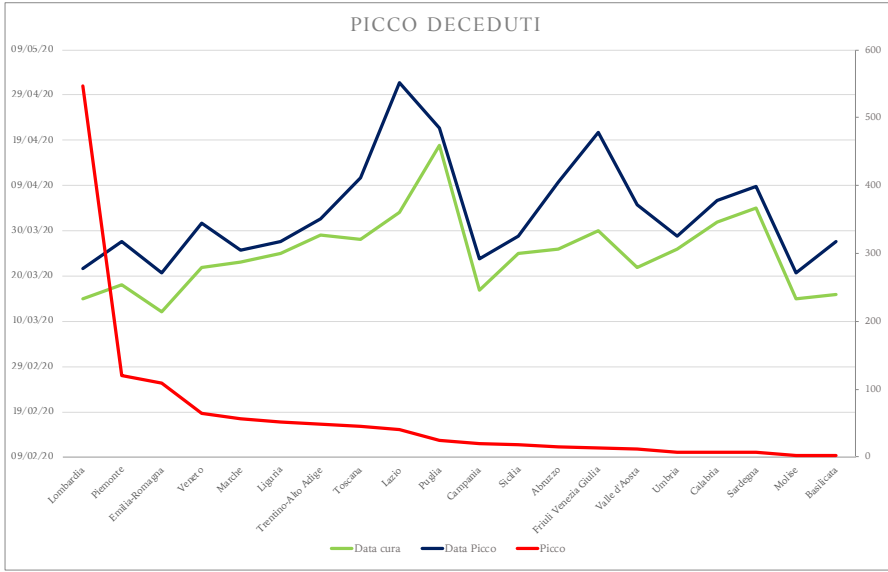


Figura 19: Picco decessi

Come è possibile notare dalla tabella riepilogativa, le regioni che hanno beneficiato maggiormente dell'adozione di un determinato tipo di cura sono quelle che hanno trattato i pazienti con farmaci anti artrite, in particolare quelle che si sono servite del Tocilizumab. Dal grafico si evince come a fare da pioniera nei test del Tocilizumab sia stata la regione Campania. È altresì evidente come, una volta adottata tale tipologia di cura, sia stato possibile raggiungere immedia-

Data cura	Regione	Data Picco	Picco
2020-03-15	Lombardia	2020-03-21	546
2020-03-18	Piemonte	2020-03-27	120
2020-03-12	Emilia-Romagna	2020-03-20	109
2020-03-22	Veneto	2020-03-31	64
2020-03-23	Marche	2020-03-25	56
2020-03-25	Liguria	2020-03-27	51
2020-03-29	Trentino-Alto Adige	2020-04-01	49
2020-03-28	Toscana	2020-04-10	46
2020-04-03	Lazio	2020-05-01	41
2020-04-18	Puglia	2020-04-21	25
2020-03-17	Campania	2020-03-23	20
2020-03-25	Sicilia	2020-03-28	18
2020-03-26	Abruzzo	2020-04-09	15
2020-03-30	Friuli Venezia Giulia	2020-04-20	14
2020-03-22	Valle d'Aosta	2020-04-04	12
2020-03-26	Umbria	2020-03-28	7
2020-04-01	Calabria	2020-04-05	7
2020-04-04	Sardegna	2020-04-08	7
2020-03-15	Molise	2020-03-20	3
2020-03-16	Basilicata	2020-03-27	2

Tabella 5: Data cura e picco decessi

tamente il picco dei decessi, il che ha causato una successiva riduzione del numero dei decessi. Evidenziando quindi l'impatto positivo di tale cura sui diversi pazienti, è stato predisposto dagli enti governativi che, a partire dagli inizi di aprile, tutti i pazienti che versavano in gravi condizioni respiratorie potessero essere trattati con Tocilizumab. Tuttavia, siccome vi era un'evidenza solo empirica e non clinicamente testata della bontà di tale trattamento, le diverse regioni hanno scelto se allinearsi o meno con tali direttive. In tal senso la regione Lazio rappresenta un caso particolare in quanto unica a sottoporre i propri pazienti ad una cura con farmaco per l'HIV. Esso ha, sì, dato dei riscontri positivi sui malati in condizioni non gravi, ma non si è dimostrato in grado di fronteggiare la situazione dei malati in gravi condizioni respiratorie, peggiorando anche, in taluni casi, la loro situazione clinica.

sebbene questi dati possono spiegare il diverso numero di decessi riscontrato nelle varie regioni, per comprendere il diverso tasso di contagio tra i vari fattori va annoverato anche il numero di accessi alle regioni, valutati sia in termini di flusso di veicoli sia di flusso di passeggeri aerei. Siccome il periodo di incubazione del virus varia dai 15 ai 30 giorni, è stato tenuto conto del flusso verificatosi nel mese di febbraio e marzo, periodo in cui si presume possa essere stato maggiormente veicolato il virus.

```

1 SELECT
2   a.'data',
3   a.'nome',
4   a.'regione',
5   a.'provincia',
6   a.'pass_nazionali',
7   a.'pass_internazionali',
8   (a.'pass_nazionali'/TOT.'totnaz') 'Percentuale regionale traffico nazionale aereo',
9   (a.'pass_internazionali'/TOT.'totint') 'Percentuale regionale traffico internazionale aereo'
10  FROM
11   'aeroporti' a,
12   ( SELECT
13     SUM(al.'pass_nazionali') AS totnaz,
14     SUM(al.'pass_internazionali') AS totint
15   FROM
16     'aeroporti' al
17   WHERE
18     al.'DATA' LIKE '2020-02%' ) AS TOT
19  WHERE
20   a.'DATA' LIKE '2020-02%'
21  GROUP BY
22   a.'cod_prov';

```

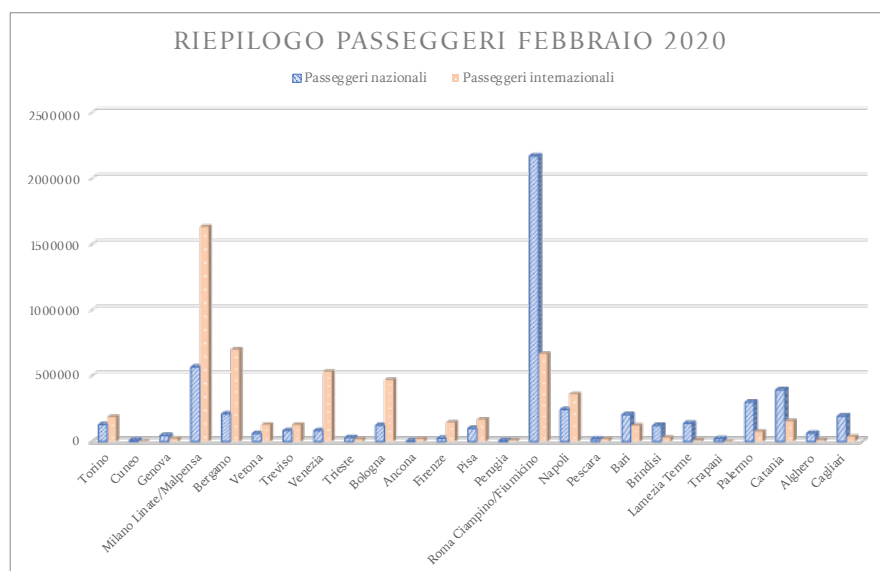



Figura 20: Passeggeri febbraio

Come si può vedere dai grafici, è ipotizzabile che il numero di accessi tramite viaggi aerei, soprattutto internazionali, sia stata una variabile determinante nella veicolazione del virus. Infatti, nel mese di febbraio gli aeroporti della Lombardia sono stati caratterizzati da un elevato numero di passeggeri e non desta meraviglia che il picco di contagi all'interno delle due province più colpite, ossia Milano e Bergamo, sia avvenuto a metà del mese di marzo, coincidente con il lasso di tempo necessario affinché possano essere manifestati i sintomi di infezione da coronavirus. A sostegno di quanto affermato è il dato relativo ai passeggeri del periodo di marzo, in quanto è possibile evidenziare come sia calato complessivamente il traffico aereo, con picchi elevati esclusivamente per gli aeroporti della provincia di Roma, in cui, seppur limitato, è stato comunque mantenuto un elevato numero di accessi tramite voli, rafforzando di gran lunga le modalità di controllo dei passeggeri. Si tenga presente, tra l'altro, che molti dei voli aventi come mete gli aeroporti della Lombardia erano strettamente connessi alle attività produttive progressivamente limitate dalle misure governative nel mese di marzo.

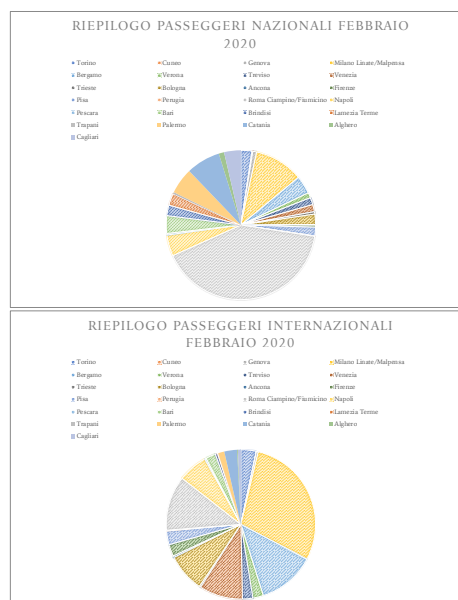


Figura 21: Tasso viabilità aerea regionale

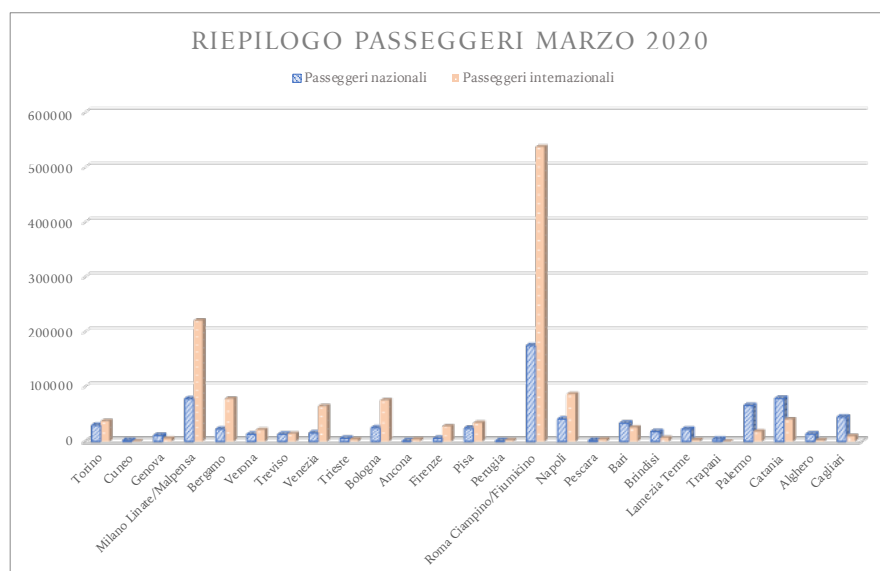


Figura 22: Passeggeri marzo



Figura 23: Tasso viabilità aerea regionale

Analogo discorso può essere affrontato per il traffico autostradale, inerente pertanto ai mezzi su gomma. Anche tale dato sembra confermare la tesi per cui le regioni che hanno presentato un numero di veicoli maggiore, abbiano anche avuto un numero di contagi più elevato. È possibile anche conoscere la saturazione della rete autostradale, fornendo una stima del numero di veicoli circolanti per l'estensione della rete autostradale espressa in km. Tuttavia è giusto ritenere questo indicatore solo come parzialmente valido in quanto ci sono regioni che, avendo una rete autostradale poco estesa, presentano comunque un'elevata saturazione. Così come ci sono regioni collocate in aree geografiche strategiche, come l'Umbria, che fanno da snodo di parte del traffico veicolare per cui presenteranno un elevato indice di saturazione.

```

1 SELECT
2     v.'data',
3     v.'regione',
4     r.'km_autostrada',
5     v.'traffico',
6     ( v.'traffico' / r.'km_autostrada' ) AS "Saturazione_autostrada"
7 FROM
8     'veicoli' v,
9     'regioni' r
10 WHERE
11     r.'cod_reg' = v.'cod_reg'
12     AND v.'data' LIKE "2020-02-%"
13 GROUP BY
14     v.'cod_reg'
15 ORDER BY
16     v.'traffico' DESC;

```

Data	Regione	km autostrada	Traffico	Saturazione autostrada
2020-02-29	Lazio	1228	56902	46.3371
2020-02-29	Lombardia	950	26053	27.4242
2020-02-29	Marche	1261	16957	13.4473
2020-02-29	Emilia-Romagna	1149	16368	14.2454
2020-02-29	Umbria	807	15766	19.5366
2020-02-29	Puglia	1492	14106	9.4544
2020-02-29	Friuli Venezia Giulia	199	11345	57.0101
2020-02-29	Campania	1482	9356	6.3131
2020-02-29	Veneto	731	8968	12.2681
2020-02-29	Toscana	1451	8512	5.8663
2020-02-29	Liguria	644	7420	11.5217
2020-02-29	Calabria	1977	7196	3.6399
2020-02-29	Sardegna	2952	24119	8.1719
2020-02-29	Abruzzo	1447	6920	4.7823
2020-02-29	Piemonte	696	5742	8.2500
2020-02-29	Sicilia	3914	5689	1.4535
2020-02-29	Basilicata	1064	5620	5.2820
2020-02-29	Molise	663	4579	6.9065
2020-02-29	Valle d'Aosta	143	1620	11.3287

Tabella 6: Traffico febbraio

Data	Regione	km autostrada	Traffico	Saturazione autostrada
2020-03-31	Lazio	1228	26132	21.2801
2020-03-31	Lombardia	950	10102	10.6337
2020-03-31	Marche	1261	7956	6.3093
2020-03-31	Puglia	1492	7153	4.7942
2020-03-31	Emilia-Romagna	1149	7125	6.2010
2020-03-31	Umbria	807	6953	8.6159
2020-03-31	Friuli Venezia Giulia	199	5830	29.2965
2020-03-31	Campania	1482	4320	2.9150
2020-03-31	Veneto	731	4220	5.7729
2020-03-31	Toscana	1451	4106	2.8298
2020-03-31	Calabria	1977	3510	1.7754
2020-03-31	Sardegna	2952	3230	1.0942
2020-03-31	Abruzzo	1447	3226	2.2294
2020-03-31	Liguria	644	2854	4.4317
2020-03-31	Sicilia	3914	2745	0.7013
2020-03-31	Piemonte	696	2560	3.6782
2020-03-31	Basilicata	1064	2513	2.3618
2020-03-31	Molise	663	2175	3.2805
2020-03-31	Valle d'Aosta	143	680	4.7552

Tabella 7: Traffico marzo

Si nota come il lockdown abbia avuto un effetto molto sentito sulla saturazione delle autostrade. Infatti si prenda ad esempio la regione più trafficata, ossia il Lazio: è possibile osservare un decremento di più della metà dei veicoli nel mese di marzo rispetto al numero registrato nel mese di febbraio, periodo in cui non erano ancora attive le limitazioni degli spostamenti imposte dal lockdown. Tenendo sempre conto del discorso precedentemente effettuato, è lampante come alcune regioni, quali ad esempio l'Umbria o le Marche, abbiano un'importanza centrale come snodo del traffico autostradale, data la loro collocazione geografica.

Infine è stato possibile valutare gli effetti che il lockdown ha avuto in termini di variazioni di usi, consumi ed esigenze sul popolo italiano sotto forma di scostamento di una serie di parametri nella finestra temporale che va da marzo a maggio, rispetto alla valutazione dei medesimi parametri valutati nel periodo che va da novembre a febbraio.

Primi parametri valutati sono stati:

1. la variazione in percentuale degli ordini di attrezzature sportive effettuati tramite piattaforme e-commerce;
2. la variazione in percentuale degli ordini di libri effettuati tramite piattaforme di e-commerce.

```

1 SELECT
2   u.'regione',
3   u.'variazione_acquisti_attrezzi',
4   u.'variazione_acquisti_libri'
5 FROM
6   'usi' u;

```

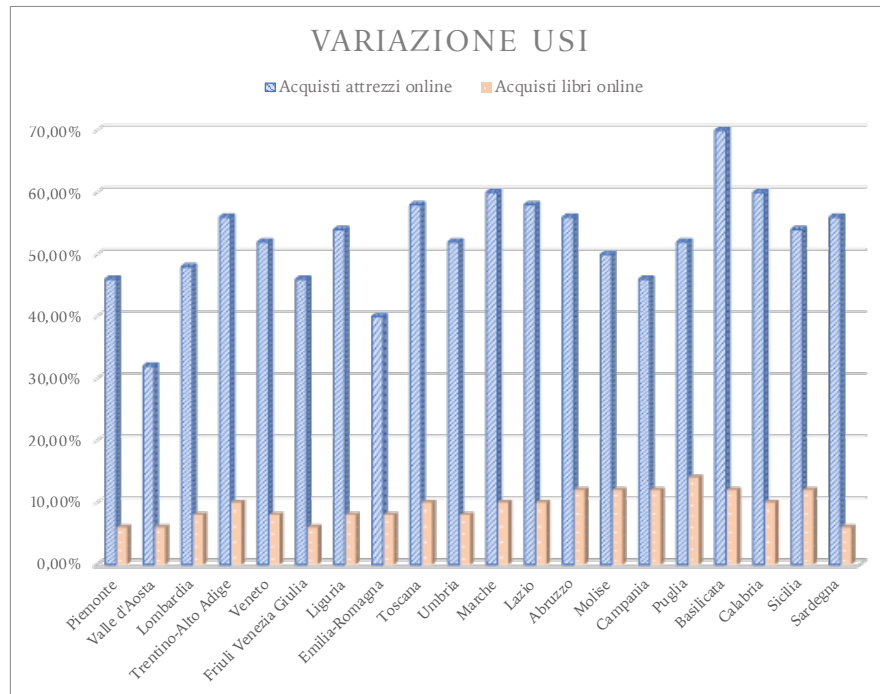


Figura 24: Variazione acquisti attrezzi e libri

In ogni regione si è assistito un incremento degli acquisti sia delle attrezzature sportive che dei libri. Con il lockdown forzato le persone hanno avuto del tempo da dedicare a se stessi curando il proprio aspetto fisico o coltivando attività cadute un po' in disuso col tempo e con la tecnologia, come la lettura. Ad esempio la Basilicata è la regione che si è contraddistinta per il maggiore incremento negli acquisti sia di attrezzi che di libri.

Successivamente è stata valutata la variazione della sottoscrizione di abbonamenti alle piattaforme di Netflix ed Amazon Prime.

```

1 SELECT
2   u.'regione',
3   u.'variazione_uso_netflix',
4   u.'variazione_uso_amazon'
5 FROM
6   'usi' u;

```

Anche in tal caso si è assistito ad un forte incremento della sottoscrizione di nuovi abbonamenti, in particolare per quanto riguarda la piattaforma di Netflix. È vero che la piattaforma di Amazon Prime fornisce sia un servizio di consegna veloce sia una piattaforma di

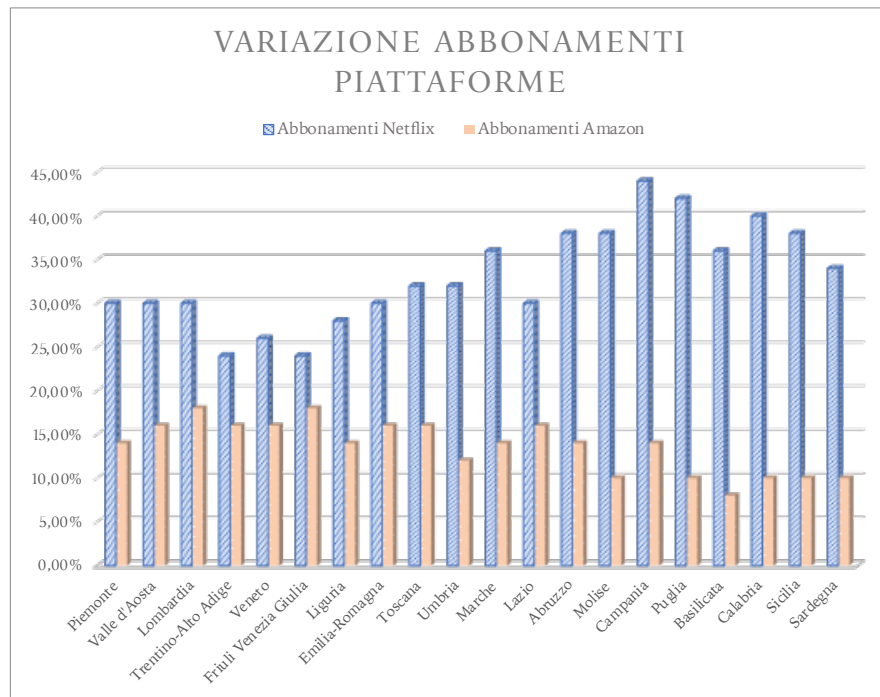


Figura 25: Variazione sottoscrizione abbonamenti

contenuti in streaming. Ma, i costi contenuti, la possibilità di poter condividere un abbonamento, nonché il catalogo di contenuti offerti in continuo aggiornamento sono stati elementi che hanno portato ad un incremento maggiore degli abbonamenti sottoscritti per la piattaforma di Netflix, piuttosto che di Amazon Prime. Ad influire su un incremento contenuto della sottoscrizione di nuovi abbonamenti alla piattaforma di Amazon Prime, ha contribuito anche l'impossibilità di garantire la consegna veloce per i beni non di prima necessità.

Infine, data la necessità di eseguire il lavoro in forma smart e quindi di avere a disposizione una connessione affidabile e veloce, è stata valutata anche la variazione della richiesta di abilitazione di servizi di fibra ottica pervenuta ai due principali gestori dei servizi di telefonia fissa nazionale ovvero: TIM e Wind3, frutto della fusione delle compagnie di Wind Infostrada e 3.

```

1 SELECT
2   u.'regione',
3   u.'variazione_richiesta_fibra_tim',
4   u.'variazione_richiesta_fibra_wind3'
5 FROM
6   'usi' u;

```

È possibile notare come in concomitanza con un aumento della richiesta di informazioni per l'abilitazione della fibra internet relativa ad uno dei due gestori, si è registrato un decremento di richiesta, almeno in alcune regioni, per l'altra compagnia telefonica. Un forte

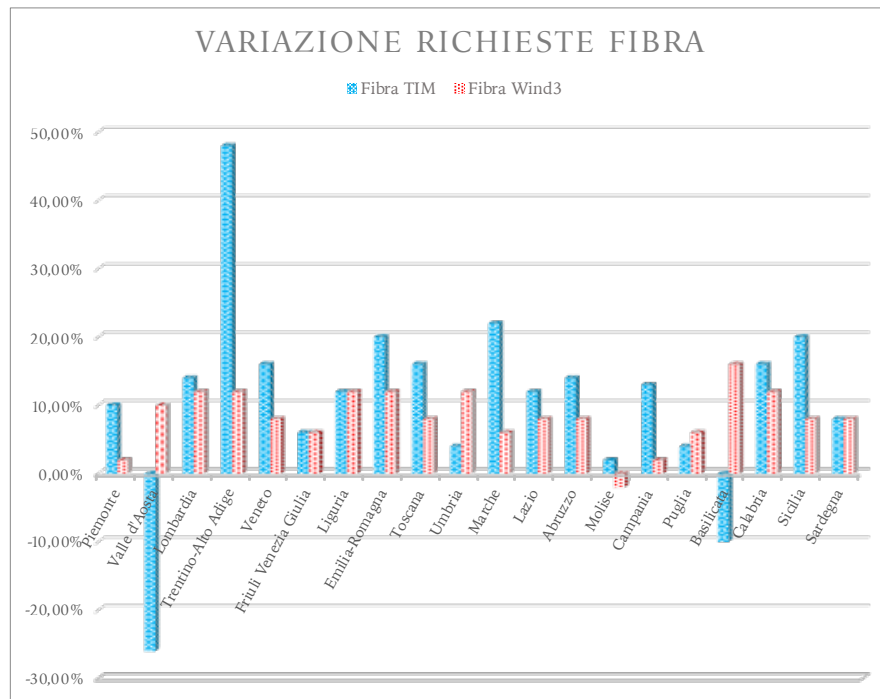


Figura 26: Variazione richiesta fibra

incremento si è registrato nella regione del Trentino-Alto Adige sia per la richiesta di abilitazione di nuovi servizi sia per chiedere informazioni relativamente ad un guasto verificatosi agli impianti nel periodo di marzo.