



**DIPARTIMENTO DI INGEGNERIA
DELL'INFORMAZIONE**

**CORSO DI LAUREA MAGISTRALE
IN INGEGNERIA INFORMATICA**

**ANALYSIS OF THE ANTI-COVID19
VACCINATION TREND IN
LOMBARDIA, ITALY**

BETTINZANA ALESSANDRO

MATRICOLA 719707

A.A. 2020/2021

INDEX

1. INTRODUCTION	2
2. DATA DESCRIPTION	4
3. DATA ANALYSIS	6
3.1 PRELIMINARY ISSUES	6
3.2 OBTAINING THE COORDINATES	7
3.3 SITUATION AS OF JUNE 17	9
3.1.1 FIRST DOSES ADMINISTERED	9
3.1.2 SECOND DOSES ADMINISTERED	14
3.4 SITUATION AS OF JUNE 22	17
3.5 SITUATION AS OF JUNE 27	22
3.6 SITUATION AS OF JULY 3	27
3.7 SITUATION AS OF JULY 8	32
3.8 SITUATION AS OF JULY 13	37
4. FURTHER DATA ANALYSIS AND FINAL RESULTS	
42	
4.1 TOTAL INCREMENTS OF DOSES ADMINISTERED	42
4.2 MY HOME MUNICIPALITY: PROVAGLIO D'ISEO	48
5. CONCLUSION	49

1. INTRODUCTION

As we all know very well, since the start of 2020 the world has been dealing with the COVID-19 pandemic.

Fortunately, thanks to scientific and technological progress that has exploded in the last thirty years, in less than a year four different vaccines have been developed, tested and officially approved by the major health organizations in the world, especially the European Medicines Agency, also known as EMA.

Since these vaccines were developed in a very short period of time (compared to the years that were required in the past), the available doses were initially very limited: in the first months of the vaccination campaign they have been reserved for health workers and for the older people.

As time went by, the availability increased steadily and more and more people had access to vaccines.

In this report I want to focus on the situation in my home country, and especially in my home region: I'll describe and analyze the vaccination campaign's trend in Lombardia, Italy.

In particular, my primary goal is to observe the response of young people to the vaccine call.

Let's make a general overview: from the third of June, people under the age of 30 and over the age of 12 can access the reservation page of the regional SISS (Sistema Informativo Socio Sanitario) in order to reserve an appointment and get vaccinated.

The majority of the people that book a vaccination appointment are vaccinated about two weeks after the reservation: based on this information, i've constantly monitored dati.lombardia.it, the website where every day is updated the dataset containing the information about the vaccine doses that are administered in every municipality in Lombardia.

In order to make a complete and in-depth analysis, I've downloaded the dataset on five different dates, beginning from the 17th of June, exactly two weeks after the start of the reservation from the young people.

The page from which I downloaded the dataset is the following:

<https://www.dati.lombardia.it/Sanit-/Somministrazioni-dei-vaccini-anti-Covid-19-in-Lomb/xdg8-8qek>

Here are the days that I've chosen:

- 17th of June
- 22nd of June
- 27th of June
- 3rd of July
- 8th of July
- 13th of July

in order to analyze the evolution of the vaccination campaign in the four most significant weeks.

I've also downloaded other two datasets:

- The first contains information about all the municipalities in Lombardia, including the address of the municipal office.
- The second contains information about the population of all the municipalities in Lombardia.

2. DATA DESCRIPTION

The datasets which show the situation of the vaccination campaign are structured in five columns:

- “CODISTAT_COMUNE_DOM”: ISTAT (Istituto Nazionale di Statistica) code of the municipality of residence of the citizen subjected to the vaccine (string feature);
- “COMUNE_DOM”: municipality of residence of the citizen subjected to the vaccine (string feature);
- “PROVINCIA_DOM”: province of residence of the citizen subjected to the vaccine (string feature);
- “TOT_DOSE1”: cumulative total of the first doses administered, in reference to the specific municipality (numeric feature).
- “TOT_DOSE2”: cumulative total of the second doses administered, in reference to the specific municipality (numeric feature).

As can be expected, the main features that will be used in order to extract the information are “TOT_DOSE1” and “TOT_DOSE2”, which will be used to create analysis graphs and calculate numerical statistics.

Another important feature is “COMUNE_DOM”, because its values will be used to merge this datasets with the other two datasets that I’ve downloaded:

- the dataset that contains information about the address of the municipality is essential to try to build maps;
- the dataset that contains information about the population of the municipality is essential to relate the numbers of the administered vaccines with the population of the municipalities and thus obtain data as a percentage.

Let’s look at the last two datasets in order to see their main features.

The first one has ten different features, but the most important are “INDIRIZZO”, “CAP” and “PROVINCIA”, which will be used to obtain the coordinates of the municipalities and then to produce the graphs. Another

important feature is “COMUNE”, which contains the name of the municipality and it allows me to connect this dataset with the others.

The second one has fifteen different features, but the most important one is “Value”, which contains the number of inhabitants of a municipality. Another important feature is “Territorio”, which also contains the name of the municipality and it allows me to connect this dataset with the others.

3. DATA ANALYSIS

3.1 PRELIMINARY ISSUES

Unfortunately, as soon as I started to load the data on my RStudio environment, I noticed a big issue.

Here's how some records appear in the csv dataset:

COMUNE	SINDACO	INDIRIZZO	CAP	PROVINCIA	TELEFONO	FAX	EMAIL	POSTA_CERTIFICATA	INTERNET
RIVOLTA D'ADDA,Fabio Maria Martino Calvi,"P.zza V. Emanuele II, 1",26027,CR,036337701,0363377031,comune@comune.rivoltadadda.cr.it									
CALVIGNANO,Marco Casarini,"Via Roma, 1",27040,PV,0383871120,0383871037,uffici@comune.calvignano.pv.it,comune.calvignano@pec.provincia.pv.it									
MONTICELLI PAVESE,Enrico Berneri,"Via Roma, 11/A",27010,PV,0382722220,0382722236,sindaco@comune.monticellipavese.pavia.it,comune.monticelli@pec.provincia.pv.it									

If we look at the feature “INDIRIZZO”, we can notice that its format is completely inappropriate for a .csv file because of the comma in the middle of the feature.

When I loaded the dataset with the command

```
comuniLombardi<- read.csv("COMUNI_LOMBARDI.csv")
```

this was the result:

..COMUNE	SINDACO	INDIRIZZO	CAP	PROVINCIA
RIVOLTA D'ADDA,Fabio Maria Martino Calvi,"P.zza V. Emanu...			NA	
CALVIGNANO,Marco Casarini,"Via Roma, 1",27040,PV,03838...			NA	
MONTICELLI PAVESE,Enrico Berneri,"Via Roma, 11/A",27010,...			NA	
MESENZANA	Alberto Rossi	Via San Martino	21030 VA	
VIMERCATE,Francesco Sartini,"P.zza UnitÃ d'Italia, 1",20871,...			NA	
INARZO,Fabrizio Montonati,"Via Patrioti, 26",21020,VA,0332...			NA	

Only the few records that have the feature “INDIRIZZO” stored without the comma and the quotation marks were correctly read.

I therefore decided not to use this dataset but to simply rely on the name of the municipality to obtain all the coordinates.

I've then uploaded the other dataset using the same command `read.csv`.

3.2 OBTAINING THE COORDINATES & THE POPULATION

In order to obtain the coordinates of the municipality in Lombardia, we have to use the command *geocode*, which is included in the *ggmap* library.

After I've extracted and concatenated the columns “COMUNE_DOM” and “PROVINCIA_DOM” into another array called *ListaComuni*, i used the function *geocode* in order to extract the coordinates of each municipality in Lombardia:

```
coordinateComuni<-geocode(ListaComuni)
```

	lon	lat
1	9.498437	45.70523
2	9.680215	45.94498
3	9.315577	45.17638
4	11.051154	45.03148
5	9.495219	45.35890
6	10.312234	45.17597
7	8.950689	45.67537
8	9.180719	45.77809
9	8.673041	45.88164
10	9.747621	45.97773

Once the coordinates have been extracted, I've merged with them the main dataset into a new dataset called *datiComuni1706*, which is the one that I'll use for analyzing the situation as of June 17.

Before doing any analysis one more feature is missing in the new dataset: the value of the population in each municipality.

This information is stored in the last dataset that I've described in the introduction part, so I had to add that column into the new dataset, paying attention to the correspondence of the values: before merging I had to sort all the dataset in alphabetical order in order to assure the lack of errors.

I've noticed that in the main dataset there was a row that stored information about vaccines administered in unknown locations: before merging the dataset I also had to manage that row, removing it from my dataset.

Here's how the data look before the start of the analysis:

	i..CODISTAT_COMUNE_DOM	COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione
1184	015146	MILANO	MILANO	701287	297617	9.189982	45.46420	1397715
1133	017029	BRESCIA	BRESCIA	108310	50872	10.211802	45.54155	195102
898	108033	MONZA	MONZA E DELLA BRIANZA	65916	29792	9.274449	45.58450	123983
1498	016024	BERGAMO	BERGAMO	63592	34424	9.677270	45.69826	119684
54	013075	COMO	COMO	47658	20807	-80.132756	25.78013	84808
198	012026	BUSTO ARSIZIO	VARESE	42366	19053	8.853127	45.61189	83121
304	015209	SESTO SAN GIOVANNI	MILANO	41579	18160	9.225687	45.53282	80589
496	012133	VARESE	VARESE	44010	20959	8.825058	45.82060	80039
750	015077	CINISELLO BALSAMO	MILANO	36628	16305	9.214384	45.55836	73537
742	019036	CREMONA	CREMONA	38543	22650	10.022651	45.1325	71523
1392	018110	PAVIA	PAVIA	43051	22809	9.158207	45.18472	70971
228	018177	VIGEVANO	PAVIA	30711	14329	8.835471	45.31549	62108
546	015118	LEGNANO	MILANO	30859	13858	8.914248	45.59834	58938
849	012070	GALLARATE	VARESE	25429	11661	8.792127	45.66236	53960
1108	015182	RHO	MILANO	28685	12062	9.039612	45.53265	49609
165	020030	MANTOVA	MANTOVA	25289	13122	10.791375	45.15642	48523
20	015166	PADERNO DUGNANO	MILANO	24702	9365	9.163981	45.57116	47980
55	097042	LECCO	LECCO	27545	13298	9.397670	45.85657	47415
1304	015081	COLOGNO MONZESE	MILANO	23534	9972	9.278707	45.52860	46687

Now the dataset looks complete. We can find information about:

- the number of vaccine doses administered in the different municipalities;
- the coordinates of every municipality;
- the population of every municipality.

3.3 SITUATION AS OF JUNE 17

To start my analysis I wanted to do a quick overview of the status of the vaccination campaign in Lombardia.

3.1.1 FIRST DOSES ADMINISTERED

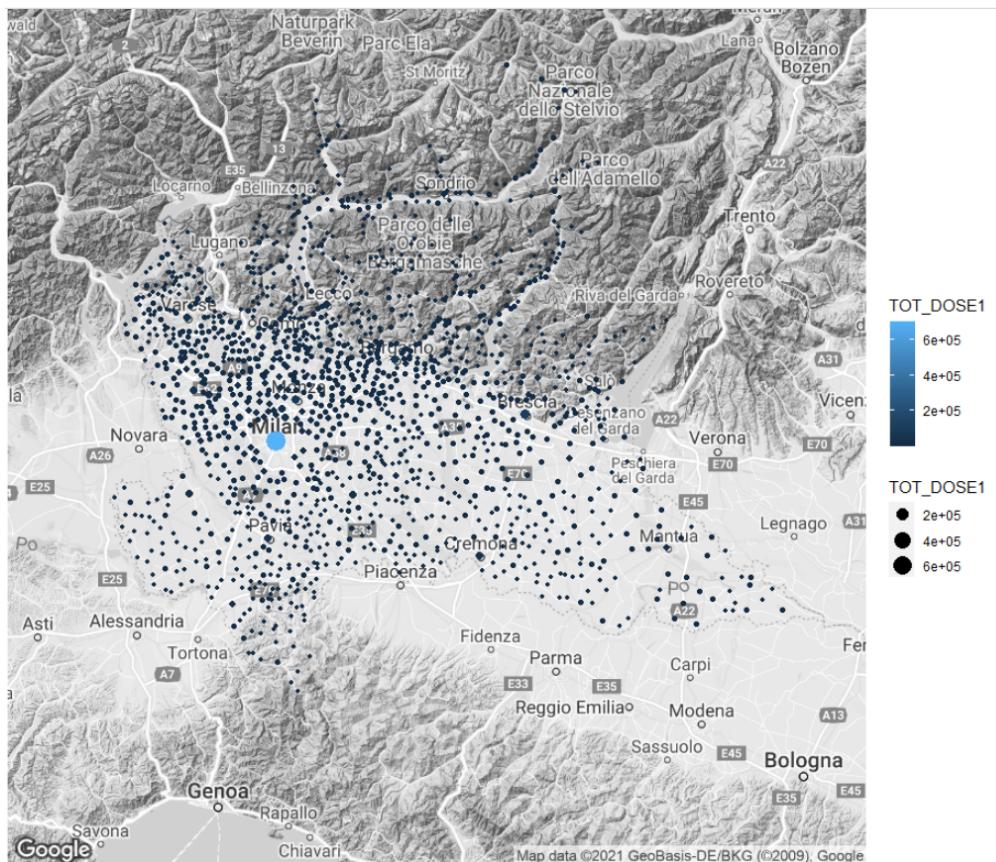
At first I treated only the “TOT_DOSE1” feature, in order to see which municipality has administered the majority of first doses to its inhabitants.

For a better understanding of the data, I have created, using the *qmap* command, some maps that clearly show the information that I’ve extracted from the datasets.

In the first map are shown the amount of first doses administered in the different municipalities.

Commands used:

```
map<-qmap("Lombardia", zoom=8,color="bw")  
map+geom_point(aes(x=lon,y=lat,color=TOT_DOSE1,size=TOT_DOSE1),data=datiComuni1706)
```



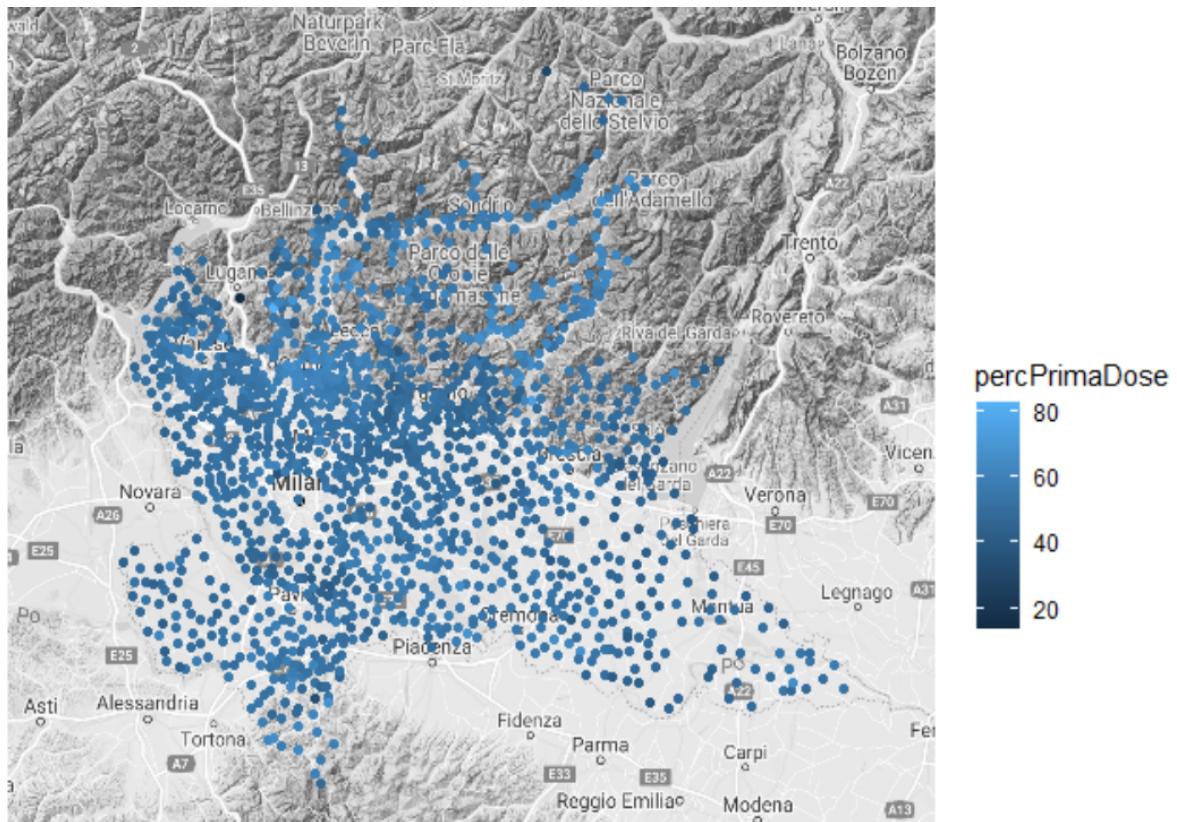
As we can clearly see, the majority of the first doses are concentrated into the larger cities, especially the region's capital Milan: this map isn't very useful in order to understand how the vaccination campaign is going.

To increase the quality of the information I've made feature expansions, connecting "TOT_DOSE1" and "popolazione" in order to create the new feature "percPrimaDose", which shows the percentage of first doses administered in relation to the population of the municipality.

With this new feature I was able to create a much more useful and readable map, which shows with different colours the values of the new feature "percPrimaDose".

Commands used:

```
datiComuni1706$percPrimaDose<-(datiComuni1706$TOT_DOSE1/datiComuni1706$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percPrimaDose),data=datiComuni1706)
```



With this new map we can finally see the status of the vaccination campaign in the different municipalities in Lombardia: the brighter the color, the better it's going.

To read this map better it's necessary to also report some portion of the dataset, which will help me to discuss about what we can extract from these numbers.

Here we can see the municipalities with the highest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
DIZZASCO	COMO	491	258	9.100715	45.94522	594	82.65993
PARLASCO	LECCO	101	34	9.344041	46.01849	131	77.09924
MONTE ISOLA	BRESCIA	1231	526	10.075321	45.71157	1646	74.78736
VIGGIU'	VARESE	3508	3183	8.908244	45.86999	5066	69.24595
SALA COMACINA	COMO	335	185	9.168509	45.96575	485	69.07216
VIONE	BRESCIA	438	142	10.444887	46.26500	636	68.86792
MONNO	BRESCIA	353	116	10.339959	46.21166	514	68.67704
MELETI	LODI	304	163	9.835853	45.11916	443	68.62302
CORTENOVA	LECCO	794	339	9.382347	46.00006	1160	68.44828
PIARIO	BERGAMO	712	340	9.920291	45.89046	1044	68.19923

Looking at this table we can clearly see that the municipalities with high percentage of first doses are usually small cities with few inhabitants, with the exception of *Viggiù*, in the province of Varese.

I'm not surprised to see in the table *Monte Isola*, a place that I know very well as I live very close to there: in February Lombardia was in the *yellow zone* and this led to an increase in tourism, especially in the lakeside municipalities.

Monte Isola and a few other municipalities were included in a so-called *vaccine cord*, and had priority in vaccine bookings for older people.

Monte Isola is a lake island with few inhabitants whose mean age is above average: this explains the reason for this very high percentage.

Here instead we can see the municipalities with the lowest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
CAMPIONE D'ITALIA	COMO	251	171	8.971459	45.96995	1869	13.42964
LIVIGNO	SONDRIO	1992	592	10.135732	46.53864	6791	29.33294
ROCCA DE' GIORGI	PAVIA	22	5	9.256065	44.95679	66	33.33333
BARANZATE	MILANO	4250	1651	9.114030	45.52822	11717	36.27208
COMEZZANO-CIZZAGO	BRESCIA	1562	695	9.959941	45.46564	3990	39.14787
CAPPELLA DE' PICENARDI	CREMONA	161	93	10.229879	45.15847	411	39.17275
BLELLO	BERGAMO	29	13	9.571008	45.83765	74	39.18919
PONTI SUL MINCIO	MANTOVA	912	373	10.686071	45.41207	2326	39.20894
PREVALLE	BRESCIA	2637	1119	10.414450	45.54620	6723	39.22356

Here the variety of municipalities is surprising, because we see very small municipalities that enter this table for a handful of doses and bigger municipalities that are far behind the ones in the previous table.

The one that stands out is *Campione d'Italia*.

If you know a bit of Lombardia's geography you can clearly see it in the map: a very dark point isolated from everyone else, because it is the only non-island municipality not connected to others as it's surrounded by Swiss territory.

Let's finally look at the biggest municipalities percentage of first doses:

i..CODISTAT_COMUNE_DOM	COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
1184 015146	MILANO	MILANO	701287	297617	9.189982	45.46420	1397715	50.17382
1133 017029	BRESCIA	BRESCIA	108310	50872	10.211802	45.54155	195102	55.51455
898 108033	MONZA	MONZA E DELLA BRIANZA	65916	29792	9.274449	45.58450	123983	53.16535
1498 016024	BERGAMO	BERGAMO	63592	34424	9.677270	45.69826	119684	53.13325
54 013075	COMO	COMO	47658	20807	-80.132756	25.78013	84808	56.19517
198 012026	BUSTO ARSIZIO	VARESE	42366	19053	8.853127	45.61189	83121	50.96907
304 015209	SESTO SAN GIOVANNI	MILANO	41579	18160	9.225687	45.53282	80589	51.59389
496 012133	VARESE	VARESE	44010	20959	8.825058	45.82060	80039	54.98569
750 015077	CINISELLO BALSAMO	MILANO	36628	16305	9.214384	45.55836	73537	49.80894
742 019036	CREMONA	CREMONA	38543	22650	10.022651	45.1325	71523	53.88896
1392 018110	PAVIA	PAVIA	43051	22809	9.158207	45.18472	70971	60.65999
228 018177	VIGEVANO	PAVIA	30711	14329	8.835471	45.31549	62108	49.44774

In order to understand better this table, I've calculated the mean of the percentage of the first doses administered:

```
> mean(datiComuni1706$percPrimaDose)
[1] 52.79087
```

We can see that, more or less, all the biggest municipalities are in line with the regional's mean: the one that stands out is *Pavia*, with a percentage of above 60%.

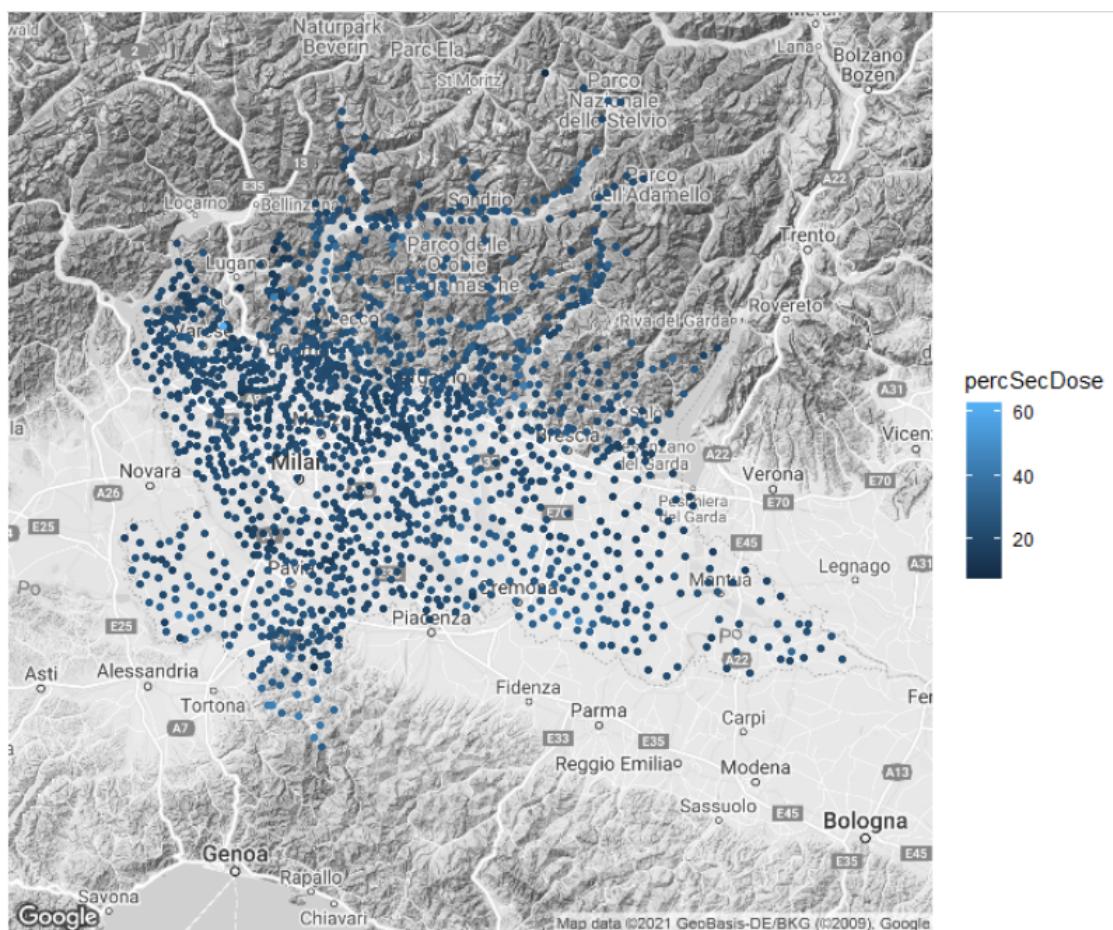
3.1.2 SECOND DOSES ADMINISTERED

Now let's focus on the feature "TOT_DOSE2", in order to see the status of the vaccination campaign, regarding the second doses administered.

Since the map plotted before using just the number of total doses administered didn't show any useful information, I decided to apply another feature expansion and to create the feature "PercSecDose", which represents the percentage of second doses administered in relation to the population of the different municipalities.

Commands used:

```
datiComuni1706$percSecDose<-(datiComuni1706$TOT_DOSE2/datiComuni1706$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percSecDose),data=datiComuni1706)
```



In this map we can see, like we expected, that the average percentage is much lower than the average percentage of the first doses, with the

exception of a few municipalities. Let's see which are the ones with the higher percentages:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
VIGGIU'	VARESE	3508	3183	8.908244	45.86999	5066	69.24595	62.83064
SOSPIRO	CREMONA	1996	1449	10.159339	45.10591	3031	65.85285	47.80600
CINGIA DE' BOTTI	CREMONA	775	546	10.262591	45.07735	1145	67.68559	47.68559
PONTE NIZZA	PAVIA	499	337	9.097394	44.85125	761	65.57162	44.28384
DIZZASCO	COMO	491	258	9.100715	45.94522	594	82.65993	43.43434
TAVERNOLA BERGAMASCA	BERGAMO	1246	871	10.048048	45.71031	2014	61.86693	43.24727
SANTA MARGHERITA DI STAFFORA	PAVIA	277	200	9.254419	44.76483	466	59.44206	42.91845
MEDE	PAVIA	3411	2686	8.735417	45.09723	6263	54.46272	42.88680
PEDESINA	SONDRIO	22	15	9.550299	46.08200	35	62.85714	42.85714
VARZI	PAVIA	1932	1300	9.196187	44.82351	3059	63.15789	42.49755

Here what stands out is again the municipality of *Viggiù*, with a percentage of second doses of 62%, more than 15% higher than any other municipality. We can also find, in fifth place, *Dizzasco*, which was on top of the first doses table.

Now let's see which are the ones with the lower percentages:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
ROCCA DE' GIORGI	PAVIA	22	5	9.256065	44.95679	66	33.33333	7.575758
LIVIGNO	SONDRIO	1992	592	10.135732	46.53864	6791	29.33294	8.717420
VAL REZZO	COMO	67	14	9.103436	46.07172	158	42.40506	8.860759
CAMPIONE D'ITALIA	COMO	251	171	8.971459	45.96995	1869	13.42964	9.149278
CORRIDO	COMO	352	106	9.135677	46.04846	807	43.61834	13.135068
ZELBIO	COMO	81	25	9.179892	45.90499	190	42.63158	13.157895
ROGNANO	PAVIA	265	84	9.089398	45.28845	629	42.13037	13.354531
LIMIDO COMASCO	COMO	1719	516	8.976706	45.69130	3843	44.73068	13.427010
CLAINO CON OSTENO	COMO	277	77	9.093942	46.00521	554	50.00000	13.898917
BARANZATE	MILANO	4250	1651	9.114030	45.52822	11717	36.27208	14.090638

Without any surprises, here we find again some municipalities that we've already found in the bottom of the first doses table, like *Campione d'Italia*, *Livigno* and *Rocca de' Giorgi*.

Let's finally look at the biggest municipalities percentage of second doses:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MILANO	MILANO	701287	297617	9.189982	45.46420	1397715	50.17382	21.29311
BRESCIA	BRESCIA	108310	50872	10.211802	45.54155	195102	55.51455	26.07457
MONZA	MONZA E DELLA BRIANZA	65916	29792	9.274449	45.58450	123983	53.16535	24.02910
BERGAMO	BERGAMO	63592	34424	9.677270	45.69826	119684	53.13325	28.76241
COMO	COMO	47658	20807	-80.132756	25.78013	84808	56.19517	24.53424
BUSTO ARSIZIO	VARESE	42366	19053	8.853127	45.61189	83121	50.96907	22.92201
SESTO SAN GIOVANNI	MILANO	41579	18160	9.225687	45.53282	80589	51.59389	22.53409
VARESE	VARESE	44010	20959	8.825058	45.82060	80039	54.98569	26.18598
CINISELLO BALSAMO	MILANO	36628	16305	9.214384	45.55836	73537	49.80894	22.17251
CREMONA	CREMONA	38543	22650	10.022651	45.13325	71523	53.88896	31.66813
PAVIA	PAVIA	43051	22809	9.158207	45.18472	70971	60.65999	32.13848

In order to understand better this table, I've calculated the mean of the percentage of the second doses administered:

```
> mean(datiComuni1706$percSecDose)
[1] 23.30407
```

We can conclude that, like we said about the first doses, all the biggest cities are in line or slightly above the regional average: *Pavia* is leading again the way with a percentage of 32%, almost 10% higher than the mean.

Now that we have a complete overview of the data, the main goal will be to analyze and deeply understand the evolution of the vaccination campaign in all the municipalities:

- Which ones will stand out from the rest?
- Which ones will have the most increases of first doses?
- Which ones will have the most increases of second doses?
- Which ones will have the worst increases?

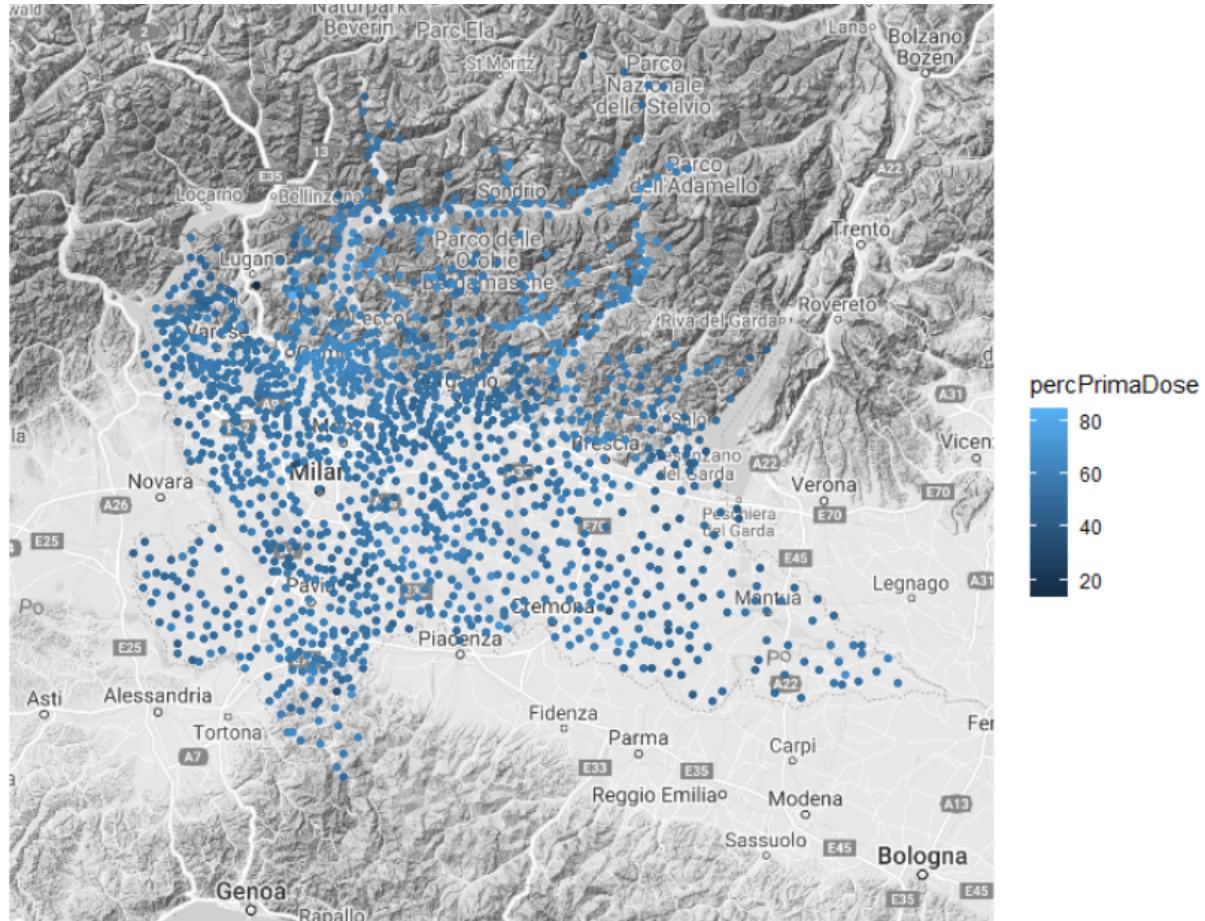
In the next subchapter I'll show the maps and the tables of both first and second doses administered at every day that I've chosen: at the end I'll dedicate a chapter for the last analysis, in which I'll answer all the previous questions.

3.4 SITUATION AS OF JUNE 22

First, let's do feature expansion in order to calculate the “percPrimaDose” feature and to show the map updated with the new data.

Commands used:

```
datiComuni2206$percPrimaDose<-(datiComuni2206$TOT_DOSE1/datiComuni2206$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percPrimaDose),data=datiComuni2206)
```



Looking at this map, we can see that especially in the northern municipalities the colours are slightly brighter, a signal that indicates that the vaccination campaign in these 5 days has progressed quickly.

Now let's see the table of the municipalities with the highest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
DIZZASCO	COMO	502	267	9.100715	45.94522	594	84.51178
PARLASCO	LECCO	104	40	9.344041	46.01849	131	79.38931
MONTE ISOLA	BRESCIA	1246	558	10.075321	45.71157	1646	75.69866
FONTENO	BERGAMO	403	180	10.019890	45.75854	551	73.13975
CINGIA DE' BOTTI	CREMONA	824	563	10.262591	45.07735	1145	71.96507
PREMANA	LECCO	1567	624	9.422898	46.05162	2182	71.81485
PEDESINA	SONDRIO	25	15	9.550299	46.08200	35	71.42857
CORTENOVA	LECCO	828	372	9.382347	46.00006	1160	71.37931
MELETI	LODI	315	171	9.835853	45.11916	443	71.10609
TARTANO	SONDRIO	140	70	9.679080	46.10745	197	71.06599

The assumption made earlier now finds further proof: the municipalities with the highest percentage of first doses are small and with few inhabitants: the top positions are unchanged from 5 days earlier.

Let's see the municipalities with the lowest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
CAMPIONE D'ITALIA	COMO	253	175	8.971459	45.96995	1869	13.53665
LIVIGNO	SONDRIO	2218	649	10.135732	46.53864	6791	32.66087
ROCCA DE' GIORGI	PAVIA	23	6	9.256065	44.95679	66	34.84848
BARANZATE	MILANO	4446	1798	9.114030	45.52822	11717	37.94487
BLELLO	BERGAMO	30	14	9.571008	45.83765	74	40.54054
CAPPELLA DE' PICENARDI	CREMONA	172	95	10.229879	45.15847	411	41.84915
PONTI SUL MINCIO	MANTOVA	979	393	10.686071	45.41207	2326	42.08942
MANERBA DEL GARDA	BRESCIA	2245	969	10.575359	45.55620	5321	42.19132
VAL REZZO	COMO	67	15	9.103436	46.07172	158	42.40506
TORNATA	CREMONA	186	96	10.439034	45.10844	435	42.75862

Here we can see that, like the municipalities with the highest percentage, the bottom positions are unchanged from 5 days earlier, with the municipality of *Campione d'Italia* leading the way with an infamous margin of almost 20%: in 5 days only 2 first doses were administered.

Let's finally look at the biggest municipalities percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
MILANO	MILANO	741879	322970	9.189982	45.46420	1397715	53.07799
BRESCIA	BRESCIA	114260	54425	10.211802	45.54155	195102	58.56424
MONZA	MONZA E DELLA BRIANZA	69789	32931	9.274449	45.58450	123983	56.28917
BERGAMO	BERGAMO	67445	36382	9.677270	45.69826	119684	56.35256
COMO	COMO	49415	22352	-80.132756	25.78013	84808	58.26691
BUSTO ARSIZIO	VARESE	44154	20844	8.853127	45.61189	83121	53.12015
SESTO SAN GIOVANNI	MILANO	43462	20289	9.225687	45.53282	80589	53.93044
VARESE	VARESE	45278	23517	8.825058	45.82060	80039	56.56992
CINISELLO BALSAMO	MILANO	38442	18132	9.214384	45.55836	73537	52.27573
CREMONA	CREMONA	40877	23196	10.022651	45.13325	71523	57.15224
PAVIA	PAVIA	44776	23880	9.158207	45.18472	70971	63.09056
VIGEVANO	PAVIA	32385	15673	8.835471	45.31549	62108	52.14304

In order to understand better this data, I've calculated the mean of the percentage of first doses administered at this day:

```
> mean(datiComuni2206$percPrimaDose)
[1] 55.79913
```

We can see that *Pavia* is still leading the way with a percentage of 8% above average, with *Brescia* and *Como* following 5% below.

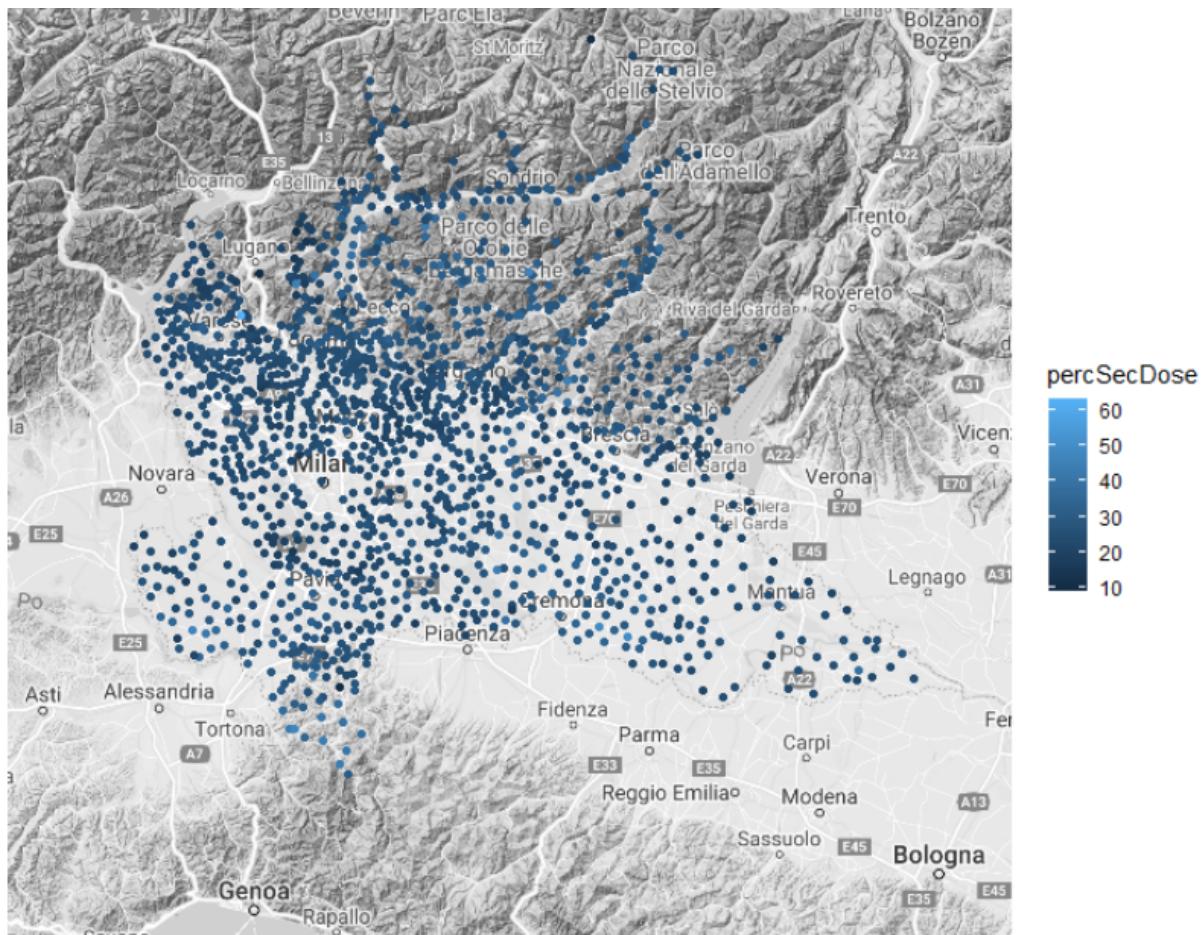
Now let's repeat this process in order to analyze the trend of the second doses administered.

First, let's calculate the "percSecDose" feature and plot the resulting map.

Command used:

```
datiComuni2206$percSecDose<- (datiComuni2206$TOT_DOSE1 / datiComuni2206$popolazione) * 100

map+geom_point(aes(x=lon,y=lat,color=percSecDose),data=datiComuni2206)
```



If we look carefully at this map we can see a slight improvement from 5 days earlier: we have to remember that the increase of second doses administered is due to the fact that the elderly, vaccinated with astrazeneca, had to wait 84 days between the first dose and the second one.

Now let's see the table of the municipalities with the highest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
VIGGIU'	VARESE	3538	3199	8.908244	45.86999	5066	69.83814	63.14647
CINGIA DE' BOTTI	CREMONA	824	563	10.262591	45.07735	1145	71.96507	49.17031
SOSPIRO	CREMONA	2074	1467	10.159339	45.10591	3031	68.42626	48.39987
PONTE NIZZA	PAVIA	504	360	9.097394	44.85125	761	66.22865	47.30618
MEDÈ	PAVIA	3512	2934	8.735417	45.09723	6263	56.07536	46.84656
VALGOGLIO	BERGAMO	317	270	9.913918	45.97610	587	54.00341	45.99659
DIZZASCO	COMO	502	267	9.100715	45.94522	594	84.51178	44.94949
TAVERNOLA BERGAMASCA	BERGAMO	1312	898	10.048048	45.71031	2014	65.14399	44.58788
VARZI	PAVIA	1957	1343	9.196187	44.82351	3059	63.97516	43.90324
SAN BASSANO	CREMONA	1458	925	9.806812	45.24312	2109	69.13229	43.85965

Like with the first doses, the top positions are unchanged from 5 days earlier.

Let's see the municipalities with the lowest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
ROCCA DE' GIORGI	PAVIA	23	6	9.256065	44.95679	66	34.84848	9.090909
CAMPIONE D'ITALIA	COMO	253	175	8.971459	45.96995	1869	13.53665	9.363296
VAL REZZO	COMO	67	15	9.103436	46.07172	158	42.40506	9.493671
LIVIGNO	SONDRIO	2218	649	10.135732	46.53864	6791	32.66087	9.556766
CORRIDO	COMO	387	111	9.135677	46.04846	807	47.95539	13.754647
CLAINO CON OSTENO	COMO	289	82	9.093942	46.00521	554	52.16606	14.801444
ROGNANO	PAVIA	293	96	9.089398	45.28845	629	46.58188	15.262321
BARANZATE	MILANO	4446	1798	9.114030	45.52822	11717	37.94487	15.345225
CARLAZZO	COMO	1475	491	9.157899	46.04933	3102	47.54997	15.828498
LIMIDO COMASCO	COMO	1803	626	8.976706	45.69130	3843	46.91647	16.289357

Here we can see almost the same municipalities as 5 days earlier: the bottom positions are unchanged, with *Campione d'Italia* leading the way alongside the little municipalities of *Rocca de' Giorgi* and *Val Rezzo*.

Finally, let's have a look at the biggest municipalities second doses percentages:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MILANO	MILANO	741879	322970	9.189982	45.46420	1397715	53.07799	23.10700
BRESCIA	BRESCIA	114260	54425	10.211802	45.54155	195102	58.56424	27.89566
MONZA	MONZA E DELLA BRIANZA	69789	32931	9.274449	45.58450	123983	56.28917	26.56090
BERGAMO	BERGAMO	67445	36382	9.677270	45.69826	119684	56.35256	30.39838
COMO	COMO	49415	22352	-80.132756	25.78013	84808	58.26691	26.35600
BUSTO ARSIZIO	VARESE	44154	20844	8.853127	45.61189	83121	53.12015	25.07670
SESTO SAN GIOVANNI	MILANO	43462	20289	9.225687	45.53282	80589	53.93044	25.17589
VARESE	VARESE	45278	23517	8.825058	45.82060	80039	56.56992	29.38193
CINISELLO BALSAMO	MILANO	38442	18132	9.214384	45.55836	73537	52.27573	24.65698
CREMONA	CREMONA	40877	23196	10.022651	45.13325	71523	57.15224	32.43153
PAVIA	PAVIA	44776	23880	9.158207	45.18472	70971	63.09056	33.64755
VIGEVANO	PAVIA	32385	15673	8.835471	45.31549	62108	52.14304	25.23507

In order to understand better this data, I've calculated the mean of the percentage of second doses administered at this day:

```
> mean(datiComuni2206$percSecDose)
[1] 25.43722
```

Even in this table *Pavia* is leading the way, confirming the 8% above average: this time is followed by the municipalities of *Bergamo* and *Varese*, which are almost 5% above the regional average.

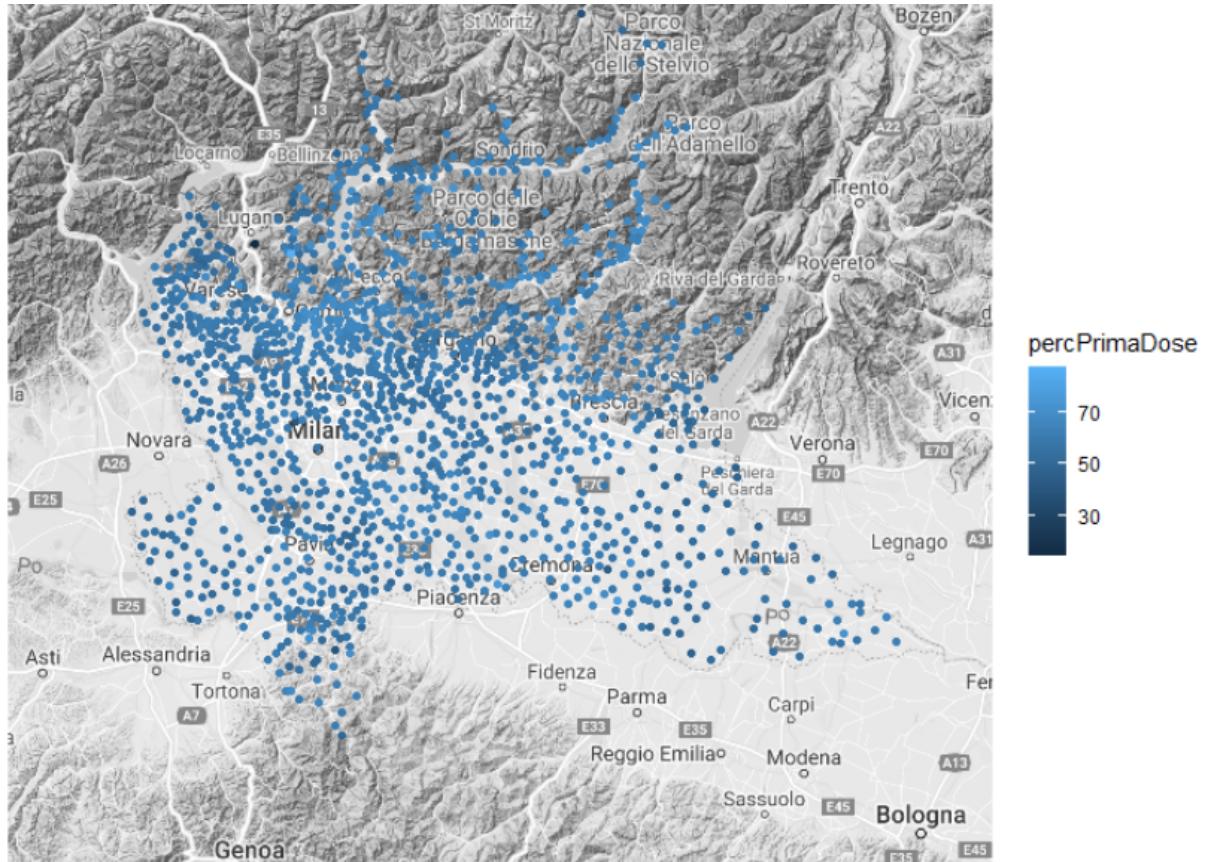
On the other hand, we can see that *Milano* is still a few percentage points below the average.

3.5 SITUATION AS OF JUNE 27

First, let's do feature expansion in order to calculate the “percPrimaDose” feature and to show the map updated with the new data.

Commands used:

```
datiComuni2706$percPrimaDose<-(datiComuni2706$TOT_DOSE1/datiComuni2706$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percPrimaDose),data=datiComuni2706)
```



In this map we can clearly see an improvement, especially from the map plotted using the data from 10 days earlier: I can proudly say that I'm, in my little, a little reason for this improvement, thanks to the first dose received on the 24th of June.

Let's see the municipalities with the highest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
DIZZASCO	COMO	523	269	9.100715	45.94522	594	88.04714
PARLASCO	LECCO	107	44	9.344041	46.01849	131	81.67939
CINGIA DE' BOTTI	CREMONA	896	567	10.262591	45.07735	1145	78.25328
MONTE ISOLA	BRESCIA	1262	564	10.075321	45.71157	1646	76.67072
CUSIO	BERGAMO	163	84	9.603297	45.99079	215	75.81395
MELETI	LODI	335	172	9.835853	45.11916	443	75.62077
PREMANA	LECCO	1639	692	9.422898	46.05162	2182	75.11457
FONTENO	BERGAMO	411	187	10.019890	45.75854	551	74.59165
CORTENOVA	LECCO	861	398	9.382347	46.00006	1160	74.22414
OLTRESSENDA ALTA	BERGAMO	107	50	9.988497	45.93968	146	73.28767

Like 5 and 10 days earlier, here we can see the same municipalities, with *Dizzasco* that is rapidly increasing its percentage and has almost reached 90% of first doses administered.

Now let's see the municipalities with the lowest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
CAMPIONE D'ITALIA	COMO	267	176	8.971459	45.96995	1869	14.28571
LIVIGNO	SONDRIO	2496	750	10.135732	46.53864	6791	36.75453
ROCCA DE' GIORGI	PAVIA	26	7	9.256065	44.95679	66	39.39394
BARANZATE	MILANO	4818	1916	9.114030	45.52822	11717	41.11974
BLELLO	BERGAMO	33	15	9.571008	45.83765	74	44.59459
CAPPELLA DE' PICENARDI	CREMONA	184	96	10.229879	45.15847	411	44.76886
VAL REZZO	COMO	71	16	9.103436	46.07172	158	44.93671
TORRE BERETTI E CASTELLARO	PAVIA	240	94	8.687464	45.07689	525	45.71429
TORNATA	CREMONA	199	99	10.439034	45.10844	435	45.74713
PONTI SUL MINCIO	MANTOVA	1071	409	10.686071	45.41207	2326	46.04471

Even if it is way behind all the other municipalities, we can finally see that the percentage of *Campione d'Italia* is finally starting to increase, since in the period 22nd-27th of June 14 doses were administered in comparison to the 2 administered in the period 17th-21st of June.

Speaking of *Livigno*, its trend is slightly increasing (278 doses in this period vs. only 226 in the last period): I'm very curious about its percentage's evolution.

Finally, let's see the biggest municipalities percentages of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
MILANO	MILANO	794602	344178	9.189982	45.46420	1397715	56.85007
BRESCIA	BRESCIA	122888	56303	10.211802	45.54155	195102	62.98654
MONZA	MONZA E DELLA BRIANZA	73515	34787	9.274449	45.58450	123983	59.29442
BERGAMO	BERGAMO	71828	38195	9.677270	45.69826	119684	60.01471
COMO	COMO	52465	23366	-80.132756	25.78013	84808	61.86327
BUSTO ARSIZIO	VARESE	47263	22194	8.853127	45.61189	83121	56.86048
SESTO SAN GIOVANNI	MILANO	45680	21955	9.225687	45.53282	80589	56.68267
VARESE	VARESE	47954	24460	8.825058	45.82060	80039	59.91329
CINISELLO BALSAMO	MILANO	40184	19403	9.214384	45.55836	73537	54.64460
CREMONA	CREMONA	43766	23480	10.022651	45.13325	71523	61.19150
PAVIA	PAVIA	48074	24520	9.158207	45.18472	70971	67.73753
VIGEVANO	PAVIA	34970	16739	8.835471	45.31549	62108	56.30515

In order to understand better this data, I've calculated the mean of the percentage of first doses administered at this day:

```
> mean(datiComuni2706$percPrimaDose)
[1] 59.67011
```

If we have already discussed previously about *Pavia* leading the way, now we can clearly see that not only *Milano*, but also the biggest municipalities in its hinterland, like *Cinisello Balsamo* and *Sesto san Giovanni*, are struggling to keep up with the regional average.

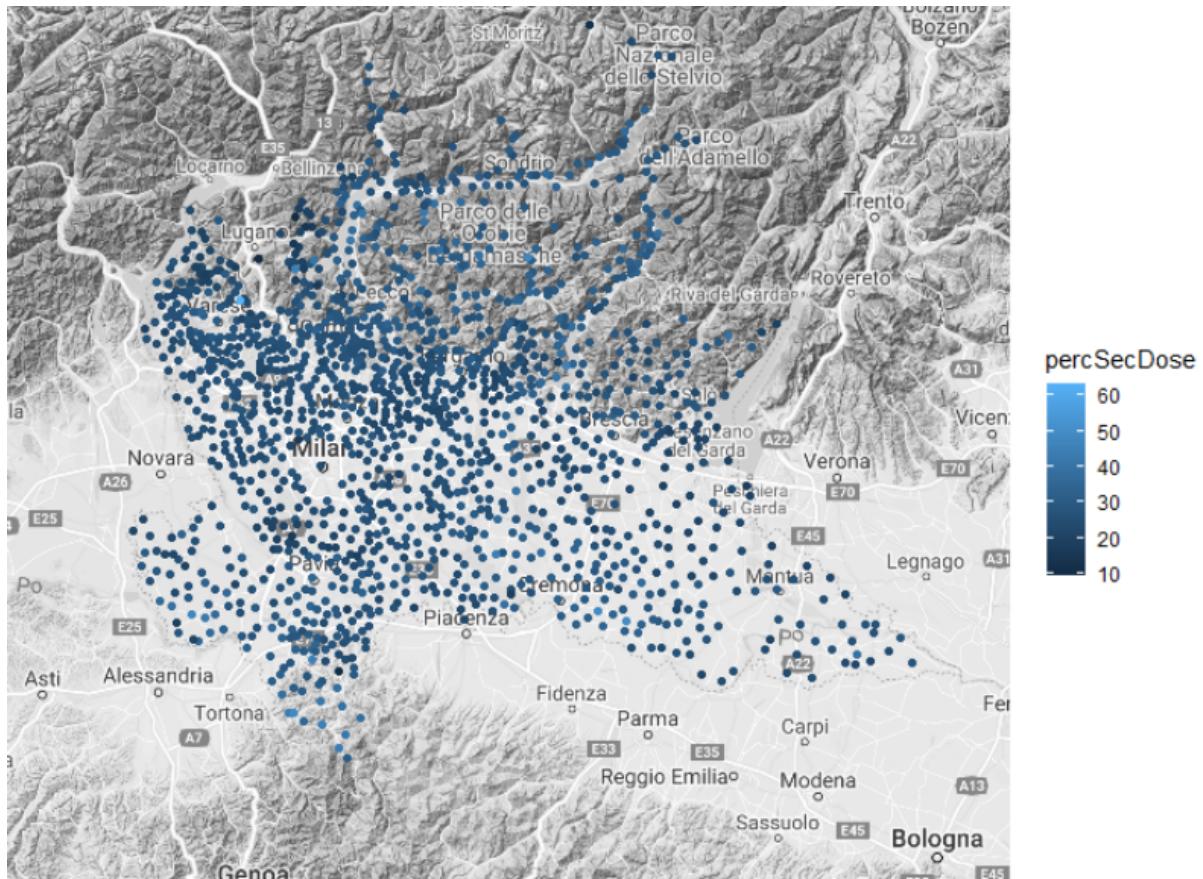
Now let's repeat this process in order to analyze the trend of the second doses administered.

First, let's calculate the "percSecDose" feature and plot the resulting map.

Command used:

```
datiComuni2706$percSecDose<- (datiComuni2706$TOT_DOSE1 / datiComuni2706$popolazione) * 100

map+geom_point(aes(x=lon,y=lat,color=percSecDose),data=datiComuni2706)
```



In this map the differences between 5 days earlier are less visible, so we'll analyze the single municipalities looking at the different tables: from my point of view, i was expecting a result like this, because when I went to receive my first dose the vast majority of the people that were there were also receiving their first doses.

Let's see the municipalities with the highest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
VIGGIU'	VARESE	3596	3211	8.908244	45.86999	5066	70.98302	63.38334
CINGIA DE' BOTTI	CREMONA	896	567	10.262591	45.07735	1145	78.25328	49.51965
PONTE NIZZA	PAVIA	522	373	9.097394	44.85125	761	68.59396	49.01445
SOSPIRO	CREMONA	2188	1477	10.159339	45.10591	3031	72.18740	48.72979
VALGOGLIO	BERGAMO	317	280	9.913918	45.97610	587	54.00341	47.70017
MEDE	PAVIA	3676	2956	8.735417	45.09723	6263	58.69392	47.19783
TAVERNOLA BERGAMASCA	BERGAMO	1378	924	10.048048	45.71031	2014	68.42105	45.87885
CALVIGNANO	PAVIA	79	49	9.161298	44.98327	108	73.14815	45.37037
DIZZASCO	COMO	523	269	9.100715	45.94522	594	88.04714	45.28620
VARZI	PAVIA	2009	1379	9.196187	44.82351	3059	65.67506	45.08009

In order to give credit to my point of view, if we look at this table and at the table representing the data from 5 days earlier, we can clearly see that even in the municipality of *Viggiù* only 12 doses were administered in this period.

Now let's see the municipalities with the lowest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
CAMPIONE D'ITALIA	COMO	267	176	8.971459	45.06995	1869	14.28571	9.41680
VAL REZZO	COMO	71	16	9.103436	46.07172	158	44.93671	10.12658
ROCCA DE' GIORGI	PAVIA	26	7	9.256065	44.95679	66	39.39394	10.60606
LIVIGNO	SONDRIO	2496	750	10.135732	46.53864	6791	36.75453	11.04403
CORRIDO	COMO	421	118	9.135677	46.04846	807	52.16853	14.62206
CLAINO CON OSTENO	COMO	301	85	9.093942	46.00521	554	54.33213	15.34296
ROGNANO	PAVIA	325	102	9.089398	45.28845	629	51.66932	16.21622
CARLAZZO	COMO	1630	505	9.157899	46.04933	3102	52.54674	16.27982
BARANZATE	MILANO	4818	1916	9.114030	45.52822	11717	41.11974	16.35231
ZELBIO	COMO	88	32	9.179892	45.90499	190	46.31579	16.84211

Also here we can see that the bottom positions are unchanged, with the percentage slightly increasing in comparison with the ones from 5 days earlier.

Finally, let's see the biggest municipalities percentages of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MILANO	MILANO	794602	344178	9.189982	45.46420	1397715	56.85007	24.62433
BRESCIA	BRESCIA	122888	56303	10.211802	45.54155	195102	62.98654	28.85824
MONZA	MONZA E DELLA BRIANZA	73515	34787	9.274449	45.58450	123983	59.29442	28.05788
BERGAMO	BERGAMO	71828	38195	9.677270	45.69826	119684	60.01471	31.91320
COMO	COMO	52465	23366	-80.132756	25.78013	84808	61.86327	27.55165
BUSTO ARSIZIO	VARESE	47263	22194	8.853127	45.61189	83121	56.86048	26.70083
SESTO SAN GIOVANNI	MILANO	45680	21955	9.225687	45.53282	80589	56.68267	27.24317
VARESE	VARESE	47954	24460	8.825058	45.82060	80039	59.91329	30.56010
CINISELLO BALSAMO	MILANO	40184	19403	9.214384	45.55836	73537	54.64460	26.38536
CREMONA	CREMONA	43766	23480	10.022651	45.13325	71523	61.19150	32.82860
PAVIA	PAVIA	48074	24520	9.158207	45.18472	70971	67.73753	34.54932
VIGEVANO	PAVIA	34970	16739	8.835471	45.31549	62108	56.30515	26.95144

In order to understand better this data, I've calculated the mean of the percentage of second doses administered up to this date:

```
> mean(datiComuni2706$percSecDose)
[1] 26.81135
```

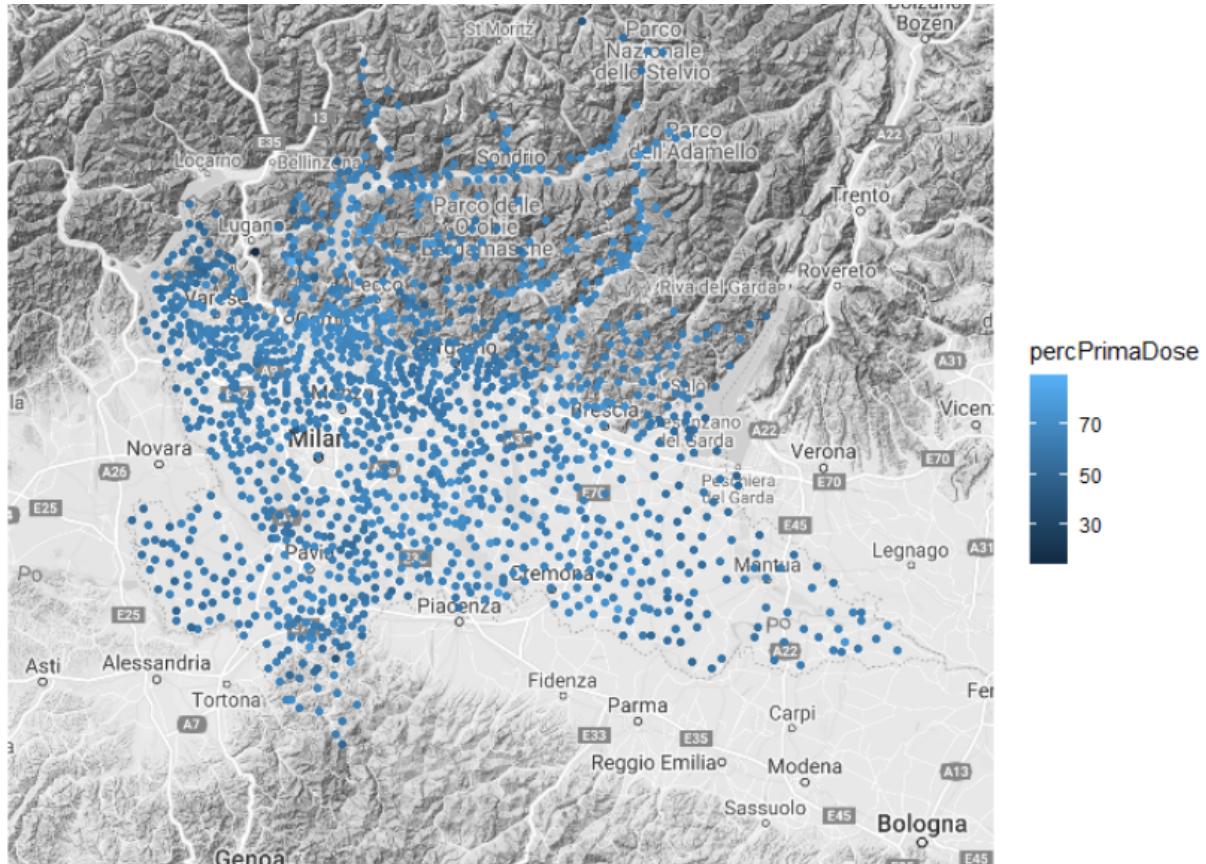
As seen before, *Pavia* is leading the way and *Milano* is the only big municipality with a percentage that stands below the regional average.

3.6 SITUATION AS OF JULY 3

First, let's do feature expansion in order to calculate the “percPrimaDose” feature and to show the map updated with the new data.

Commands used:

```
datiComuni0307$percPrimaDose<-(datiComuni0307$TOT_DOSE1/datiComuni0307$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percPrimaDose),data=datiComuni0307)
```



This map shows that the vaccination campaign is going very well, in fact if we look back at the first map we can clearly see that the colors here are much brighter than before, with a lot of municipalities with percentages of first doses above 60/70%.

Let's see the municipalities with the highest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
DIZZASCO	COMO	530	326	9.100715	45.94522	594	89.22559
PARLASCO	LECCO	110	48	9.344041	46.01849	131	83.96947
CINGIA DE' BOTTI	CREMONA	928	585	10.262591	45.07735	1145	81.04803
MELETI	LODI	354	182	9.835853	45.11916	443	79.90971
PREMANA	LECCO	1693	767	9.422898	46.05162	2182	77.58937
MONTE ISOLA	BRESCIA	1277	602	10.075321	45.71157	1646	77.58202
PEDESINA	SONDRIO	27	15	9.550299	46.08200	35	77.14286
CALVIGNANO	PAVIA	83	51	9.161298	44.98327	108	76.85185
COLLEBEATO	BRESCIA	3465	1637	10.213636	45.58396	4542	76.28798
CUSIO	BERGAMO	164	88	9.603297	45.99079	215	76.27907

Here I can finally welcome another municipality in the province of Brescia in this chart: *Collebeato* joins *Monte Isola* in the top positions of the percentages of first doses administered.

Now let's see the municipalities with the lowest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
CAMPIONE D'ITALIA	COMO	271	181	8.971459	45.96995	1869	14.49973
ROCCA DE' GIORGI	PAVIA	26	9	9.256065	44.95679	66	39.39394
LIVIGNO	SONDRIO	2719	807	10.135732	46.53864	6791	40.03829
BARANZATE	MILANO	5090	2133	9.114030	45.52822	11717	43.44115
BLELLO	BERGAMO	33	18	9.571008	45.83765	74	44.59459
VAL REZZO	COMO	71	19	9.103436	46.07172	158	44.93671
PEGLIO	COMO	87	44	9.261677	46.19376	182	47.80220
ZELBIO	COMO	91	37	9.179892	45.90499	190	47.89474
BRUSIMPIANO	VARESE	560	240	8.889302	45.94934	1163	48.15133
MASCIAGO PRIMO	VARESE	145	66	8.782525	45.92177	300	48.33333

Unfortunately, if we look at the previous table we can see that in *Campione d'Italia* only 4 doses were administered in this period: I'm expecting to see it also at the bottom of the final analysis.

Finally, let's see the biggest municipalities percentages of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
MILANO	MILANO	846879	380750	9.189982	45.46420	1397715	60.59025
BRESCIA	BRESCIA	129705	60650	10.211802	45.54155	195102	66.48061
MONZA	MONZA E DELLA BRIANZA	78310	37382	9.274449	45.58450	123983	63.16189
BERGAMO	BERGAMO	76261	39883	9.677270	45.69826	119684	63.71863
COMO	COMO	54570	26727	-80.132756	25.78013	84808	64.34534
BUSTO ARSIZIO	VARESE	50852	24462	8.853127	45.61189	83121	61.17828
SESTO SAN GIOVANNI	MILANO	48171	23741	9.225687	45.53282	80589	59.77367
VARESE	VARESE	50328	27144	8.825058	45.82060	80039	62.87935
CINISELLO BALSAMO	MILANO	42292	20858	9.214384	45.55836	73537	57.51118
CREMONA	CREMONA	45260	25325	10.022651	45.13325	71523	63.28034
PAVIA	PAVIA	49819	26567	9.158207	45.18472	70971	70.19628
VIGEVANO	PAVIA	36495	18711	8.835471	45.31549	62108	58.76055

In order to understand better this data, I've calculated the mean of the percentage of first doses administered at this day:

```
> mean(datiComuni0307$percPrimaDose)
[1] 62.96875
```

In comparison to the previous tables, we can see that the trends are still going: *Pavia* is leading the way and *Milano* and its biggest municipalities are struggling below the regional average.

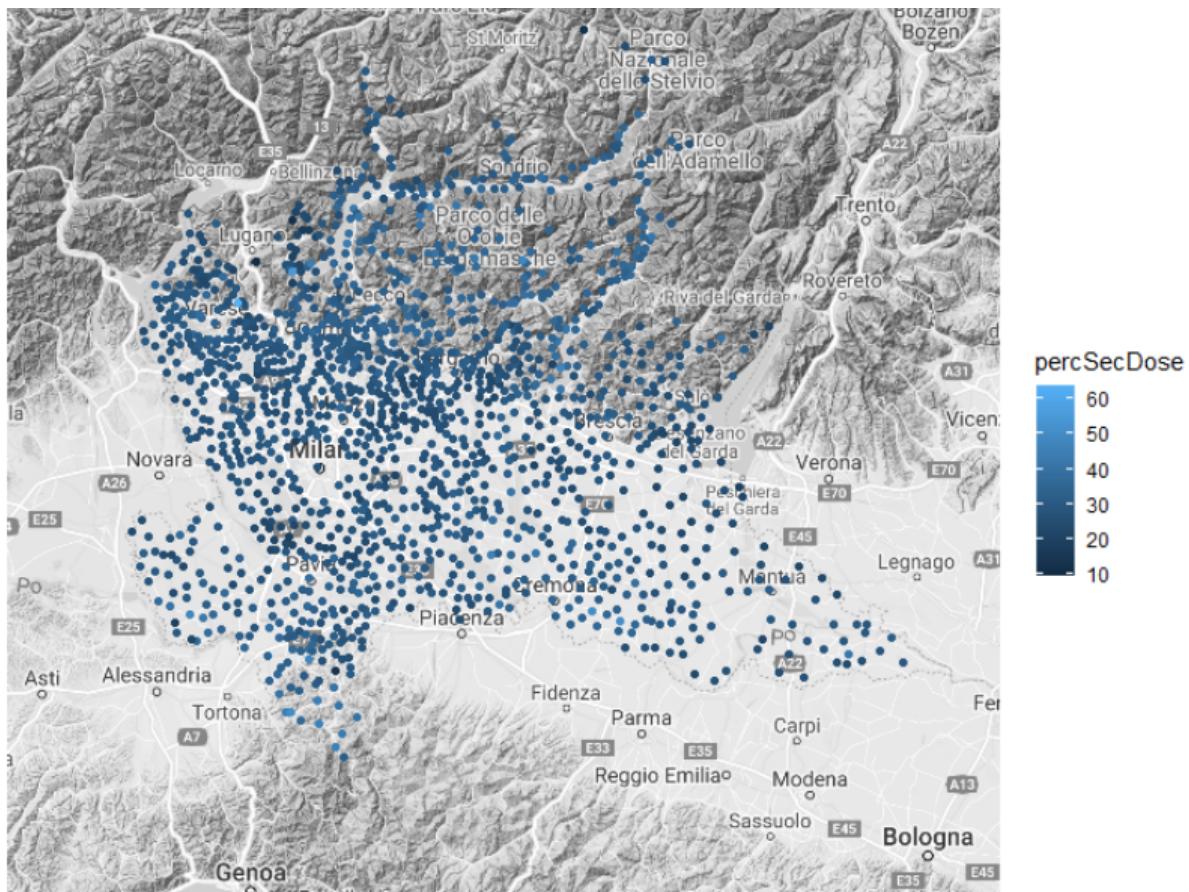
Now let's repeat this process in order to analyze the trend of the second doses administered.

First, let's calculate the “percSecDose” feature and plot the resulting map.

Command used:

```
datiComuni2706$percSecDose<- (datiComuni2706$TOT_DOSE1 / datiComuni2706$popolazione)*100

map+geom_point(aes(x=lon,y=lat,color=percSecDose),data=datiComuni2706)
```



As we have seen in the previous second doses map, we can't clearly see the increase in the percentages: we have to look at the tables in order to make some statement about our data.

Let's see the municipalities with the highest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
VIGGIÙ	VARESE	3642	3224	8.908244	45.86999	5066	71.89104	63.63995
DIZZASCO	COMO	530	326	9.100715	45.94522	594	89.22559	54.88215
CINGIA DE' BOTTI	CREMONA	928	585	10.262591	45.07735	1145	81.04803	51.09170
PONTE NIZZA	PAVIA	543	384	9.097394	44.85125	761	71.35348	50.45992
SOSPIRO	CREMONA	2241	1525	10.159339	45.10591	3031	73.93599	50.31343
VALGOGLIO	BERGAMO	319	282	9.913918	45.97610	587	54.34412	48.04089
MEDE	PAVIA	3932	2982	8.735417	45.09723	6263	62.78141	47.61297
TAVERNOLA BERGAMASCA	BERGAMO	1468	954	10.048048	45.71031	2014	72.88977	47.36842
CALVIGNANO	PAVIA	83	51	9.161298	44.98327	108	76.85185	47.22222
VARZI	PAVIA	2067	1406	9.196187	44.82351	3059	67.57110	45.96273

The trend observed with the dataset of 27th June is confirmed: in *Viggiù*, which is leading the way from the start, only 13 second doses were administered in this period.

Now let's see the municipalities with the lowest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
CAMPIONE D'ITALIA	COMO	271	181	8.971459	45.96995	1869	14.49973	9.684323
LIVIGNO	SONDRIO	2719	807	10.135732	46.53864	6791	40.03829	11.883375
VAL REZZO	COMO	71	19	9.103436	46.07172	158	44.93671	12.025316
ROCCA DE' GIORGI	PAVIA	26	9	9.256065	44.95679	66	39.39394	13.636364
CORRIDO	COMO	448	134	9.135677	46.04846	807	55.51425	16.604709
ROGNANO	PAVIA	348	111	9.089398	45.28845	629	55.32591	17.647059
CARLAZZO	COMO	1716	562	9.157899	46.04933	3102	55.31915	18.117344
BARANZATE	MILANO	5090	2133	9.114030	45.52822	11717	43.44115	18.204319
TORRE BERETTI E CASTELLARO	PAVIA	255	101	8.687464	45.07689	525	48.57143	19.238095
ZELBIO	COMO	91	37	9.179892	45.90499	190	47.89474	19.473684

Finally, let's see the biggest municipalities percentages of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MILANO	MILANO	846879	380750	9.189982	45.46420	1397715	60.59025	27.24089
BRESCIA	BRESCIA	129705	60650	10.211802	45.54155	195102	66.48061	31.08630
MONZA	MONZA E DELLA BRIANZA	78310	37382	9.274449	45.58450	123983	63.16189	30.15091
BERGAMO	BERGAMO	76261	39883	9.677270	45.69826	119684	63.71863	33.32359
COMO	COMO	54570	26727	-80.132756	25.78013	84808	64.34534	31.51472
BUSTO ARSIZIO	VARESE	50852	24462	8.853127	45.61189	83121	61.17828	29.42939
SESTO SAN GIOVANNI	MILANO	48171	23741	9.225687	45.53282	80589	59.77367	29.45936
VARESE	VARESE	50328	27144	8.825058	45.82060	80039	62.87935	33.91347
CINISELLO BALSAMO	MILANO	42292	20858	9.214384	45.55836	73537	57.51118	28.36395
CREMONA	CREMONA	45260	25325	10.022651	45.13325	71523	63.28034	35.40819
PAVIA	PAVIA	49819	26567	9.158207	45.18472	70971	70.19628	37.43360
VIGEVANO	PAVIA	36495	18711	8.835471	45.31549	62108	58.76055	30.12655

In order to understand better this data, I've calculated the mean of the percentage of second doses administered up to this date:

```
> mean(datiComuni0307$percSecDose)
[1] 29.39833
```

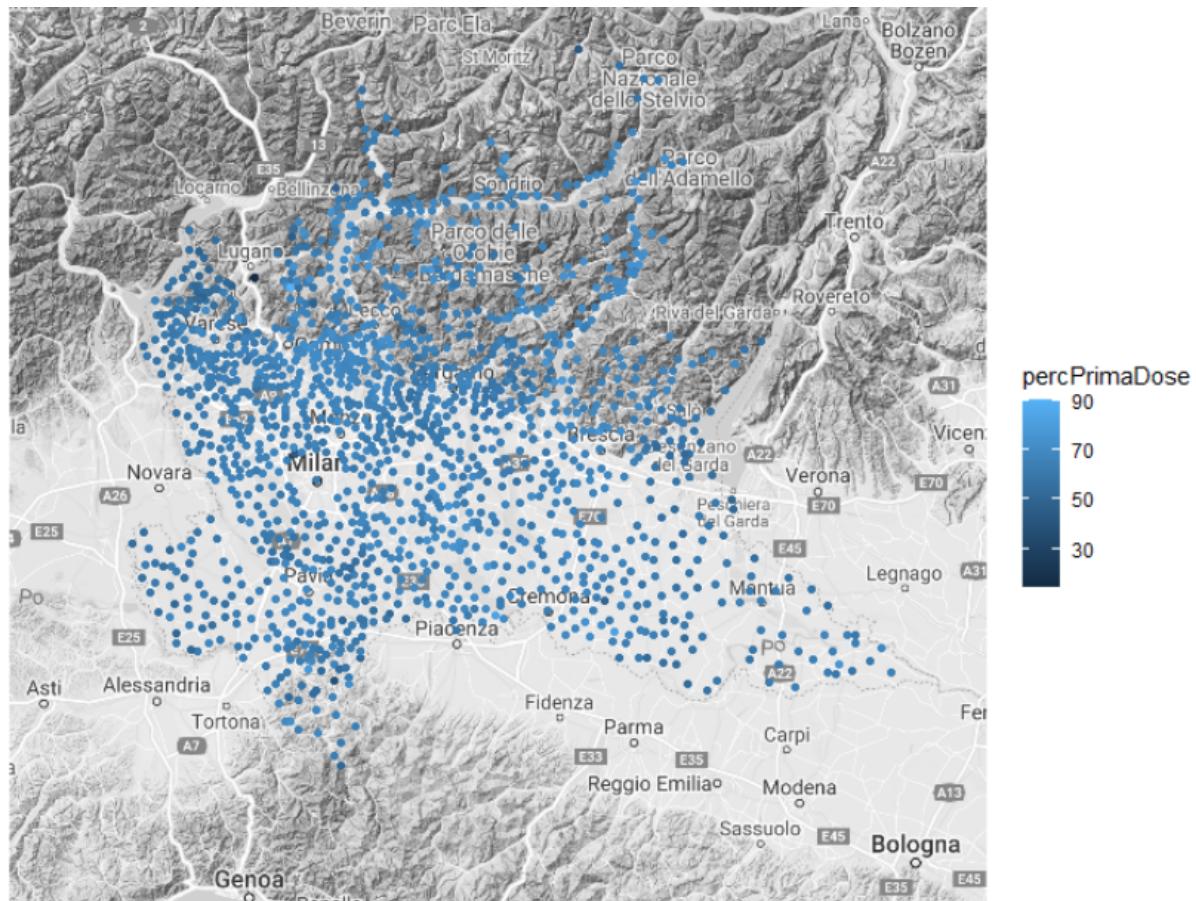
Here we can see that, unlike the first doses, only the municipality of *Milano* is struggling to keep up with the regional average, while almost all of the other biggest cities are in line or above average.

3.7 SITUATION AS OF JULY 8

First, let's do feature expansion in order to calculate the “percPrimaDose” feature and to show the map updated with the new data.

Commands used:

```
datiComuni0807$percPrimaDose<-(datiComuni0807$TOT_DOSE1/datiComuni0807$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percPrimaDose),data=datiComuni0807)
```



Here we can clearly see that almost all of the municipalities have a percentage of first doses administered above 60%: the only reason why the legend is showing values between less than 30% and 90% is for that single dark blue point that tells us that *Campione d'Italia* is still struggling.

Let's see the municipalities with the highest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
DIZZASCO	COMO	537	353	9.100715	45.94522	594	90.40404
PARLASCO	LECCO	110	59	9.344041	46.01849	131	83.96947
MELETI	LODI	369	206	9.835853	45.11916	443	83.29571
CINGIA DE BOTTI	CREMONA	942	612	10.262591	45.07735	1145	82.27074
MONTE ISOLA	BRESCIA	1294	1086	10.075321	45.71157	1646	78.61482
PREMANA	LECCO	1710	854	9.422898	46.05162	2182	78.36847
SAN BASSANO	CREMONA	1646	1017	9.806812	45.24312	2109	78.04647
CALVIGNANO	PAVIA	84	51	9.161298	44.98327	108	77.77778
CUSIO	BERGAMO	167	102	9.603297	45.99079	215	77.67442
PEDESINA	SONDRIO	27	15	9.550299	46.08200	35	77.14286
ALBAREDO PER SAN MARCO	SONDRIO	219	133	9.590034	46.10298	286	76.57343
CORTENOVA	LECCO	887	502	9.382347	46.00006	1160	76.46552
COLLEBEATO	BRESCIA	3470	1843	10.213636	45.58396	4542	76.39806

While *Collebeato* has lost a few positions, we can finally see the first municipality with a percentage of first doses above 90%: *Dizzasco*, in the province of Como, with only 57 inhabitants that haven't received the vaccine. If we think that children under the age of 12 can't receive it, we can see that *Dizzasco* is quickly approaching the end of the first dosage.

Now let's see the municipalities with the lowest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
CAMPIONE D'ITALIA	COMO	273	197	8.971459	45.96995	1869	14.60674
ROCCA DE' GIORGI	PAVIA	26	9	9.256065	44.95679	66	39.39394
LIVIGNO	SONDRIO	2855	907	10.135732	46.53864	6791	42.04094
BARANZATE	MILANO	5223	2434	9.114030	45.52822	11717	44.57626
VAL REZZO	COMO	72	24	9.103436	46.07172	158	45.56962
BLELLO	BERGAMO	34	19	9.571008	45.83765	74	45.94595
ZELBIO	COMO	92	48	9.179892	45.90499	190	48.42105
BRUSIMPIANO	VARESE	564	288	8.889302	45.94934	1163	48.49527
AGRA	VARESE	195	107	8.772689	46.03470	401	48.62843
MASCIAGO PRIMO	VARESE	146	83	8.782525	45.92177	300	48.66667

As we've seen in the map, here we can observe that the *Campione d'Italia* percentage of first doses is 25% below any other municipality in Lombardia.

Finally, let's see the biggest municipalities percentages of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
MILANO	MILANO	868953	432206	9.189982	45.46420	1397715	62.16954
BRESCIA	BRESCIA	130360	67759	10.211802	45.54155	195102	66.81633
MONZA	MONZA E DELLA BRIANZA	80446	41043	9.274449	45.58450	123993	64.88470
BERGAMO	BERGAMO	78551	43225	9.677270	45.69826	119684	65.63200
COMO	COMO	55090	31534	-80.132756	25.78013	84808	64.95849
BUSTO ARSIZIO	VARESE	51964	27789	8.853127	45.61189	83121	62.51609
SESTO SAN GIOVANNI	MILANO	49902	26116	9.225687	45.53282	80589	61.92160
VARESE	VARESE	50997	30073	8.825058	45.82060	80039	63.71519
CINISELLO BALSAMO	MILANO	43910	22464	9.214384	45.55836	73537	59.71144
CREMONA	CREMONA	45570	27951	10.022651	45.13325	71523	63.71377
PAVIA	PAVIA	50185	29239	9.158207	45.18472	70971	70.71198
VIGEVANO	PAVIA	37446	20727	8.835471	45.31549	62108	60.29175

In order to understand better this data, I've calculated the mean of the percentage of first doses administered at this day:

```
> mean(datiComuni0807$percPrimaDose)
[1] 64.31121
```

Here we can see that all the biggest municipalities have slowed down a bit with the pace of the administration, because even *Pavia* has lost part of its margin on the regional average.

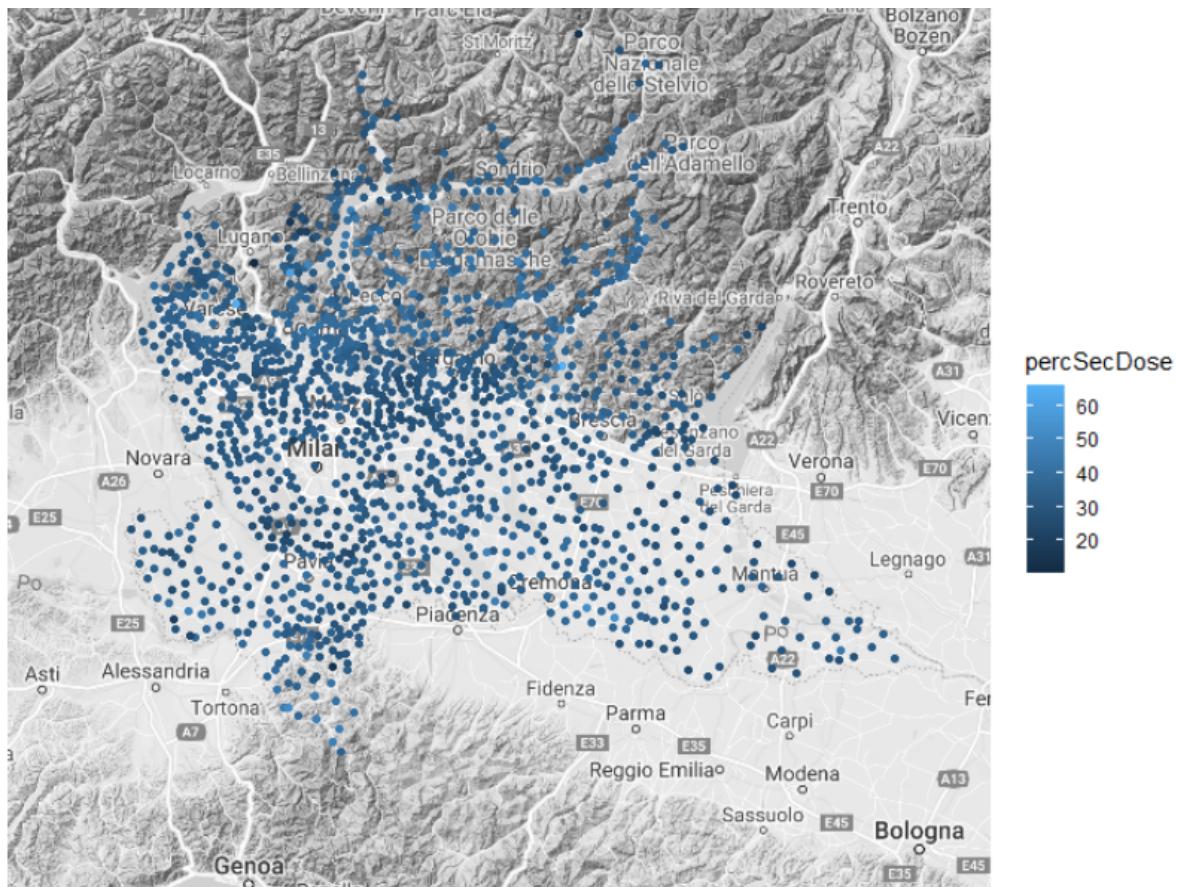
Now let's repeat this process in order to analyze the trend of the second doses administered.

First, let's calculate the "percSecDose" feature and plot the resulting map.

Command used:

```
datiComuni0807$percSecDose<-(datiComuni0807$TOT_DOSE1/datiComuni0807$popolazione)*100

map+geom_point(aes(x=lon,y=lat,color=percSecDose),data=datiComuni0807)
```



The main thing that stands out is that we don't see the 10% value on the legend anymore, a symptom that the percentages are slowly growing.

Let's see the municipalities with the highest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MONTE ISOLA	BRESCIA	1294	1086	10.075321	45.71157	1646	78.61482	65.97813
VIGGIU'	VARESE	3662	3250	8.908244	45.86999	5066	72.28583	64.15318
DIZZASCO	COMO	537	353	9.100715	45.94522	594	90.40404	59.42761
CINGIA DE' BOTTI	CREMONA	942	612	10.262591	45.07735	1145	82.27074	53.44978
SOSPIRO	CREMONA	2253	1598	10.159339	45.10591	3031	74.33190	52.72187
PONTE NIZZA	PAVIA	556	393	9.097394	44.85125	761	73.06176	51.64258
SALA COMACINA	COMO	355	238	9.168509	45.96575	485	73.19588	49.07216
TAVERNOLA BERGAMASCA	BERGAMO	1524	984	10.040848	45.71031	2014	75.67031	48.85799
VALGOGLIO	BERGAMO	320	284	9.913918	45.97610	587	54.51448	48.38160
SAN BASSANO	CREMONA	1646	1017	9.806812	45.24312	2109	78.04647	48.22191

For the first time since the start of this analysis, we don't see *Viggiù* on the top of this table. Now *Monte Isola* is leading the way, supporting my previous hypothesis of the *vaccine cord*: I think that this huge increase on the percentage of the second doses (5 days earlier was only 36%) leads back to the first doses of astrazeneca administered at the beginning of this Spring.

Now let's see the municipalities with the lowest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
CAMPIONE D'ITALIA	COMO	273	197	8.971459	45.96995	1869	14.60674	10.54040
LIVIGNO	SONDRIO	2855	907	10.135732	46.53864	6791	42.04094	13.35591
ROCCA DE' GIORGI	PAVIA	26	9	9.256065	44.95679	66	39.39394	13.63636
VAL REZZO	COMO	72	24	9.103436	46.07172	158	45.56962	15.18987
TORRE BERETTI E CASTELLARO	PAVIA	263	102	8.687464	45.07689	525	50.09524	19.42857
CORRIDO	COMO	472	157	9.135677	46.04846	807	58.48823	19.45477
CARLAZZO	COMO	1778	611	9.157899	46.04933	3102	57.31786	19.69697
BARANZATE	MILANO	5223	2434	9.114030	45.52822	11717	44.57626	20.77324
ROGNANO	PAVIA	350	133	9.089398	45.28845	629	55.64388	21.14467
SAN CIPRIANO PO	PAVIA	241	102	9.283639	45.10637	472	51.05932	21.61017

Like we've seen in the map, finally the percentage of *Campione d'Italia* is above 10%, but still way below any other municipalities.

Finally, let's see the biggest municipalities percentages of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MILANO	MILANO	868953	432206	9.189982	45.46420	1397715	62.16954	30.92233
BRESCIA	BRESCIA	130360	67759	10.211802	45.54155	195102	66.81633	34.73004
MONZA	MONZA E DELLA BRIANZA	80446	41043	9.274449	45.58450	123983	64.88470	33.10373
BERGAMO	BERGAMO	78551	43225	9.677270	45.69826	119684	65.63200	36.11594
COMO	COMO	55090	31534	-80.132756	25.78013	84808	64.95849	37.18281
BUSTO ARSIZIO	VARESE	51964	27789	8.853127	45.61189	83121	62.51609	33.43198
SESTO SAN GIOVANNI	MILANO	49902	26116	9.225687	45.53282	80589	61.92160	32.40641
VARESE	VARESE	50997	30073	8.825058	45.82060	80039	63.71519	37.57293
CINISELLO BALSAMO	MILANO	43910	22464	9.214384	45.55836	73537	59.71144	30.54789
CREMONA	CREMONA	45570	27951	10.022651	45.13325	71523	63.71377	39.07974
PAVIA	PAVIA	50185	29239	9.158207	45.18472	70971	70.71198	41.19852
VIGEVANO	PAVIA	37446	20727	8.835471	45.31549	62108	60.29175	33.37251

In order to understand better this data, I've calculated the mean of the percentage of second doses administered up to this date:

```
> mean(datiComuni0807$percSecDose)
[1] 32.8105
```

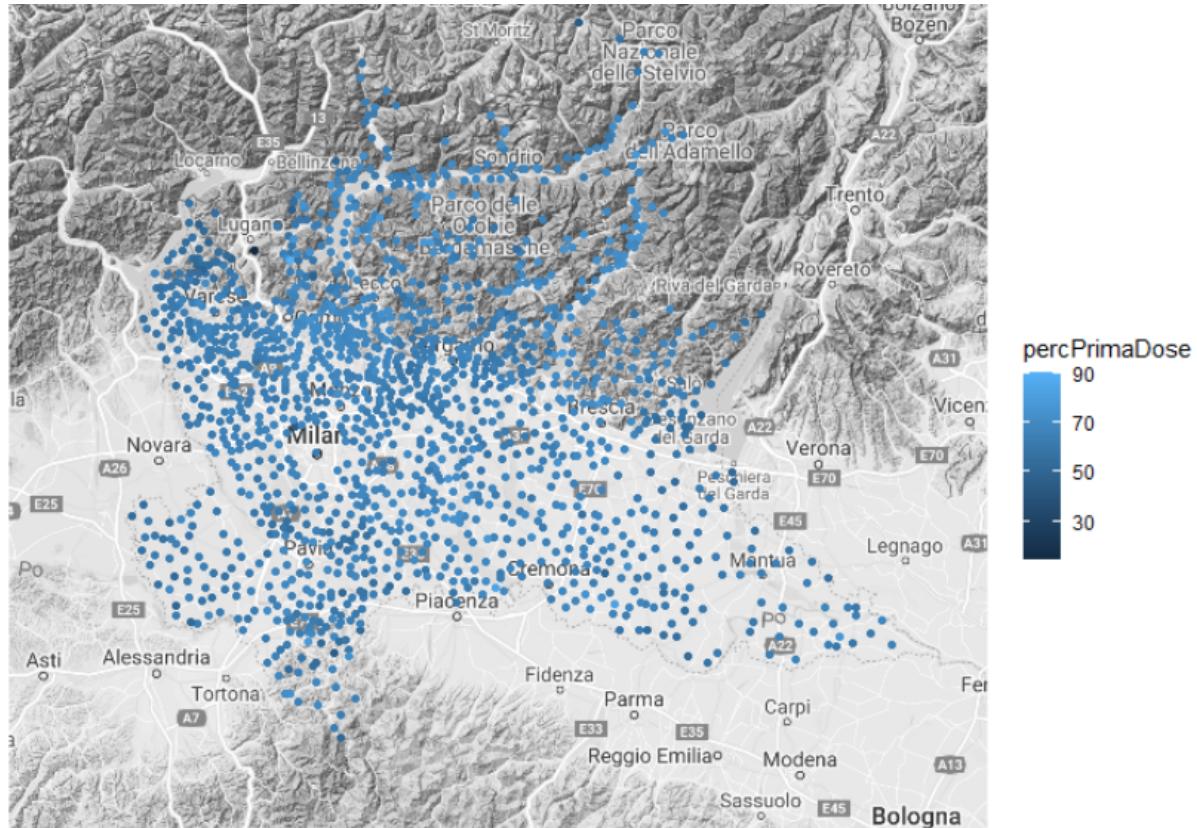
Unlike the first doses, in this table we can see that the biggest cities are constantly increasing their percentages, with only *Milano* still under the regional average.

3.7 SITUATION AS OF JULY 13

First, let's do feature expansion in order to calculate the “percPrimaDose” feature and to show the map updated with the new data.

Commands used:

```
datiComuni1307$percPrimaDose<-(datiComuni1307$TOT_DOSE1/datiComuni1307$popolazione)*100  
map+geom_point(aes(x=lon,y=lat,color=percPrimaDose),data=datiComuni1307)
```



Here we can see the last map showing the percentages of first doses: if we look back at the first map we can see that the vaccination campaign has made a lot of progress.

Two municipalities that stand out are *Campione d'Italia* and *Dizzasco*, which is represented by the light blue point to the right of the one representing *Campione d'Italia*.

Let's see the municipalities with the highest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
DIZZASCO	COMO	537	386	9.100715	45.94522	594	90.40404
PARLASCO	LECCO	110	74	9.344041	46.01849	131	83.96947
MELETI	LODI	370	224	9.835853	45.11916	443	83.52144
CINGIA DE BOTTI	CREMONA	942	646	10.262591	45.07735	1145	82.27074
MONTE ISOLA	BRESCIA	1298	1119	10.075321	45.71157	1646	78.85784
CALVIGNANO	PAVIA	85	56	9.161298	44.98327	108	78.70370
CUSIO	BERGAMO	169	109	9.603297	45.99079	215	78.60465
PREMANA	LECCO	1712	1072	9.422898	46.05162	2182	78.46013
SAN BASSANO	CREMONA	1654	1082	9.806812	45.24312	2109	78.42579
PEDESINA	SONDRIO	27	16	9.550299	46.08200	35	77.14286

Now let's see the municipalities with the lowest percentage of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
CAMPIONE D'ITALIA	COMO	273	214	8.971459	45.96995	1869	14.60674
LIVIGNO	SONDRIO	2942	1142	10.135732	46.53864	6791	43.32204
ROCCA DE' GIORGI	PAVIA	29	10	9.256065	44.95679	66	43.93939
BARANZATE	MILANO	5301	2983	9.114030	45.52822	11717	45.24196
VAL REZZO	COMO	72	35	9.103436	46.07172	158	45.56962
ZELBIO	COMO	92	60	9.179892	45.90499	190	48.42105
BRUSIMPIANO	VARESE	565	326	8.889302	45.94934	1163	48.58126
AGRA	VARESE	195	122	8.772689	46.03470	401	48.62843
MASCIAGO PRIMO	VARESE	146	90	8.782525	45.92177	300	48.66667
PEGLIO	COMO	90	44	9.261677	46.19376	182	49.45055

Finally, let's see the biggest municipalities percentages of first doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose
MILANO	MILANO	879583	498377	9.189982	45.46420	1397715	62.93007
BRESCIA	BRESCIA	130877	77847	10.211802	45.54155	195102	67.08132
MONZA	MONZA E DELLA BRIANZA	81933	45956	9.274449	45.58450	123983	66.08406
BERGAMO	BERGAMO	79357	49058	9.677270	45.69826	119684	66.30544
COMO	COMO	55248	36582	-80.132756	25.78013	84808	65.14480
BUSTO ARSIZIO	VARESE	52230	32090	8.853127	45.61189	83121	62.83611
SESTO SAN GIOVANNI	MILANO	51129	29553	9.225687	45.53282	80589	63.44414
VARESE	VARESE	51185	33515	8.825058	45.82060	80039	63.95007
CINISELLO BALSAMO	MILANO	44978	24540	9.214384	45.55836	73537	61.16377
CREMONA	CREMONA	45472	31221	10.022651	45.13325	71523	63.57675
PAVIA	PAVIA	50301	32823	9.158207	45.18472	70971	70.87543
VIGEVANO	PAVIA	38009	23202	8.835471	45.31549	62108	61.19824

In order to understand better this data, I've calculated the mean of the percentage of first doses administered at this day:

```
> mean(datiComuni1307$percPrimaDose)
[1] 64.87331
```

As we saw in the table created with the data of 5 days earlier, the biggest cities are slowing down with the increasing of the percentages.

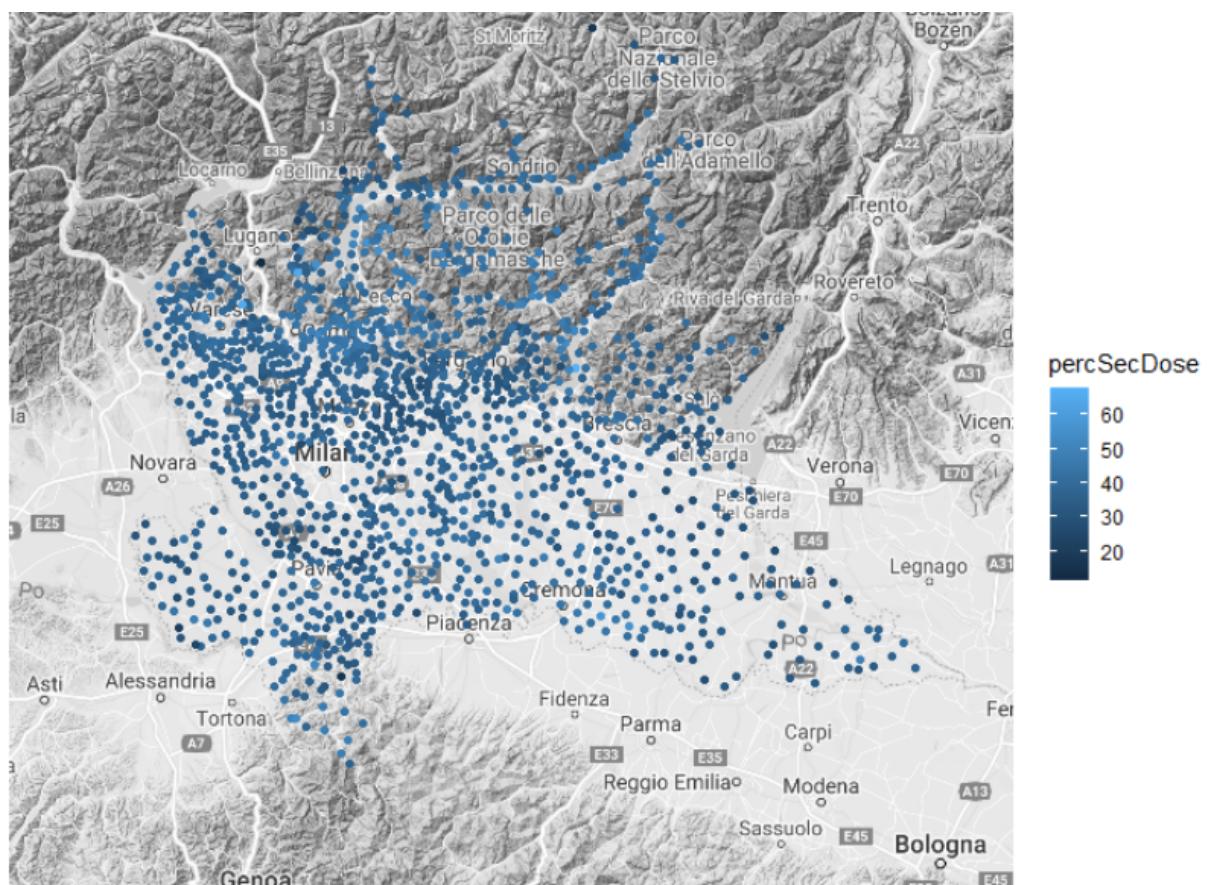
Now let's repeat this process in order to analyze the trend of the second doses administered.

First, let's calculate the "percSecDose" feature and plot the resulting map.

Command used:

```
datiComuni2706$percSecDose<-(datiComuni2706$TOT_DOSE1/datiComuni2706$popolazione)*100
```

```
map+geom_point(aes(x=lon,y=lat,color=percSecDose),data=datiComuni2706)
```



From this map stands out the municipality of *Monte Isola*, represented by the light blue point in the middle of Lake Iseo.

Let's see the municipalities with the highest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MONTE ISOLA	BRESCIA	1298	1119	10.075321	45.71157	1646	78.85784	67.98299
DIZZASCO	COMO	537	386	9.100715	45.94522	594	90.40404	64.98316
VIGGIU'	VARESE	3667	3279	8.908244	45.86999	5066	72.38452	64.72562
SOSPIRO	CREMONA	2245	1731	10.159339	45.10591	3031	74.06796	57.10986
PALARSCO	LECCO	110	74	9.344041	46.01849	131	83.96947	56.48855
CINGIA DE' BOTTI	CREMONA	942	646	10.262591	45.07735	1145	82.27074	56.41921
CORTENOVA	LECCO	888	648	9.382347	46.00006	1160	76.55172	55.86207
SALA COMACINA	COMO	356	270	9.168509	45.96575	485	73.40206	55.67010
DORIO	LECCO	240	173	9.318343	46.10175	321	74.76636	53.89408
PONTE NIZZA	PAVIA	563	407	9.097394	44.85125	761	73.98160	53.48226

As we expected, *Monte Isola* is still leading the way with a percentage of above 67%.

In order to appreciate this number I'm going to remind you the mean of the first doses:

```
> mean(datiComuni1307$percPrimaDose)
[1] 64.87331
```

These numbers are proof that the *vaccine cord* set up at the end of February has paid off very well.

Now let's see the municipalities with the lowest percentage of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
CAMPIONE D'ITALIA	COMO	273	214	8.971459	45.96995	1869	14.60674	11.44997
ROCCA DE' GIORGI	PAVIA	29	10	9.256065	44.95679	66	43.93939	15.15152
LIVIGNO	SONDRIO	2942	1142	10.135732	46.53864	6791	43.32204	16.81637
TORRE BERETTI E CASTELLARO	PAVIA	264	104	8.687464	45.07689	525	50.28571	19.80952
VAL REZZO	COMO	72	35	9.103436	46.07172	158	45.56962	22.15190
SUARDI	PAVIA	313	142	8.742770	45.03490	599	52.25376	23.70618
PEGLIO	COMO	90	44	9.261677	46.19376	182	49.45055	24.17582
SAN CIPRIANO PO	PAVIA	241	115	9.283639	45.10637	472	51.05932	24.36441
ROGNANO	PAVIA	353	160	9.089398	45.28845	629	56.12083	25.43720
BARANZATE	MILANO	5301	2983	9.114030	45.52822	11717	45.24196	25.45874

Finally, let's see the biggest municipalities percentages of second doses administered:

COMUNE_DOM	PROVINCIA_DOM	TOT_DOSE1	TOT_DOSE2	lon	lat	popolazione	percPrimaDose	percSecDose
MILANO	MILANO	879583	498377	9.189982	45.46420	1397715	62.93007	35.65655
BRESCIA	BRESCIA	130877	77847	10.211802	45.54155	195102	67.08132	39.90067
MONZA	MONZA E DELLA BRIANZA	81933	45956	9.274449	45.58450	123983	66.08406	37.06637
BERGAMO	BERGAMO	79357	49058	9.677270	45.69826	119684	66.30544	40.98961
COMO	COMO	55248	36582	-80.132756	25.78013	84808	65.14480	43.13508
BUSTO ARSIZIO	VARESE	52230	32090	8.853127	45.61189	83121	62.83611	38.60637
SESTO SAN GIOVANNI	MILANO	51129	29553	9.225687	45.53282	80589	63.44414	36.67126
VARESE	VARESE	51185	33515	8.825058	45.82060	80039	63.95007	41.87334
CINISELLO BALSAMO	MILANO	44978	24540	9.214384	45.55836	73537	61.16377	33.37096
CREMONA	CREMONA	45472	31221	10.022651	45.13325	71523	63.57675	43.65169
PAVIA	PAVIA	50301	32823	9.158207	45.18472	70971	70.87543	46.24847
VIGEVANO	PAVIA	38009	23202	8.835471	45.31549	62108	61.19824	37.35751

In order to understand better this data, I've calculated the mean of the percentage of second doses administered up to this date:

```
> mean(datiComuni1307$percSecDose)
[1] 37.46741
```

The only municipalities that are below the regional average are, like in the previous dates, *Milano* and the biggest municipalities in its hinterland: *Cinisello Balsamo* and *Sesto San Giovanni*.

On the other hand, *Pavia*, *Como* and *Cremona* are leading the way with percentages above the average by a good margin (more than 5%).

4. FURTHER DATA ANALYSIS AND FINAL RESULTS

4.1 TOTAL INCREMENTS OF DOSES ADMINISTERED

In this chapter I'm going to show the further data analysis that I've done to better understand all of the previous tables: we've seen that, except for a few exceptions, all the municipalities that were on top (or at the bottom) of the tables at the beginning of my analysis stayed there on every data update.

I've decided to take a step further with the feature expansion: I've created a big dataset, called *percComuni*, in which I've stored all the features "TOT_DOSE1", "TOT_DOSE2", "percPrimaDose", "percSecDose" from the different datasets.

Once the dataset has been expanded, I calculated the increases in percentages for both the first and second doses:

- First I calculated the difference between the percentages from one date to the previous one, in order to see the individual increments;
- After that I calculated the total increments, subtracting the percentages of June 17 from the ones of July 13.

Commands used for dealing with first doses percentages:

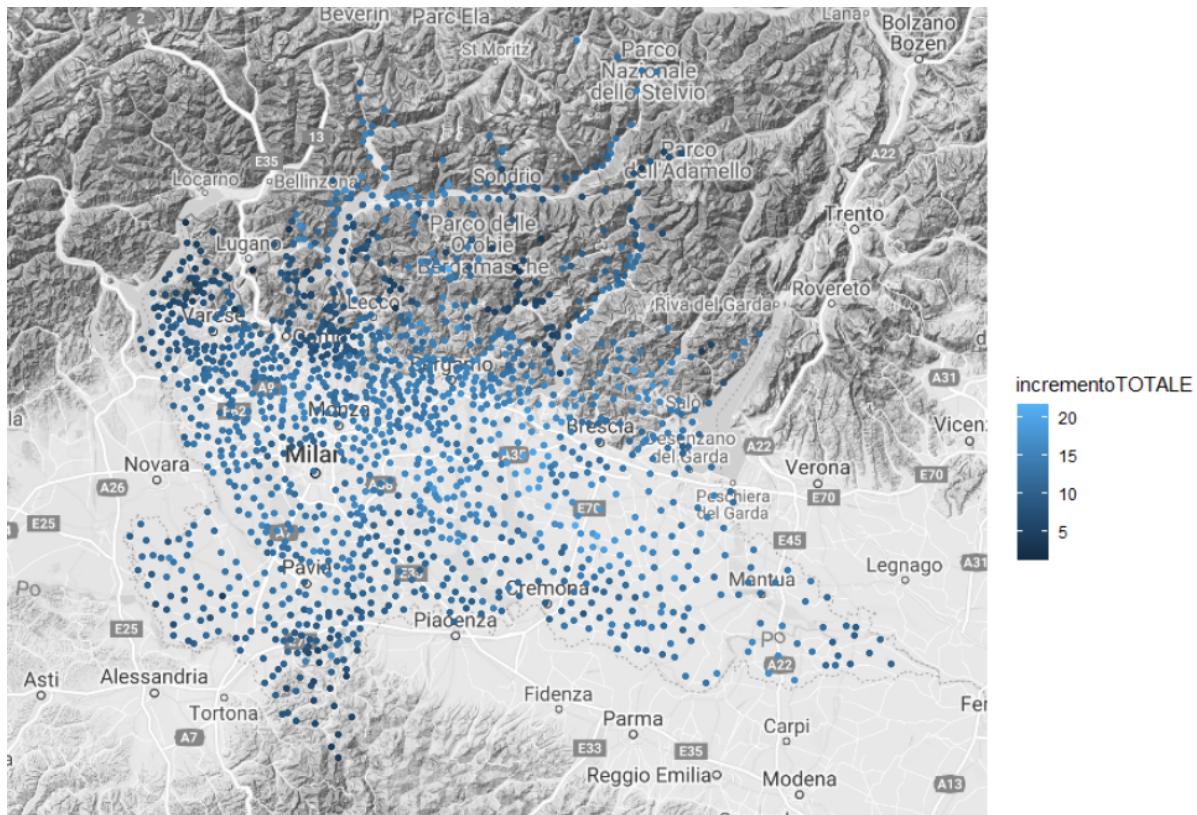
```
percComuni$incremento1<-percComuni$percPrimaDose_2206-percComuni$percPrimaDose  
percComuni$incremento2<-percComuni$percPrimaDose_2706-percComuni$percPrimaDose_2206  
percComuni$incremento3<-percComuni$percPrimaDose_0307-percComuni$percPrimaDose_2706  
percComuni$incremento4<-percComuni$percPrimaDose_0807-percComuni$percPrimaDose_0307  
percComuni$incremento5<-percComuni$percPrimaDose_1307-percComuni$percPrimaDose_0807  
percComuni$incrementoTOTALE<-percComuni$percPrimaDose_1307-percComuni$percPrimaDose
```

Commands used for dealing with second doses percentages:

```
percComuni$incremento2_1<-percComuni$percSecDose_2206-percComuni$percSecDose  
percComuni$incremento2_2<-percComuni$percSecDose_2706-percComuni$percSecDose_2206  
percComuni$incremento2_3<-percComuni$percSecDose_0307-percComuni$percSecDose_2706  
percComuni$incremento2_4<-percComuni$percSecDose_0807-percComuni$percSecDose_0307  
percComuni$incremento2_5<-percComuni$percSecDose_1307-percComuni$percSecDose_0807  
percComuni$incremento2TOTALE<-percComuni$percSecDose_1307-percComuni$percSecDose
```

Now it's time to show these increments in new dedicated maps and also to show the final table, in order to see which municipalities did the best in these weeks and which did the worse, starting with the first doses.

```
map+geom_point(aes(x=lon,y=lat,color=incrementoTOTALE), data=percComuni)
```



In this map we see the total increments of the percentages of first doses administered in the different municipalities: I'm expecting to see at the top of the tables some municipalities located in the southern part of Brescia's province and one municipality located in the Alpi Orobie Bergamasche.

Let's see the table of the municipalities with the highest increments:

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale
PIAZZOLO	BERGAMO	55.17241	77.01149	21.83908
COMEZZANO-CIZZAGO	BRESCIA	39.14787	60.82707	21.67920
BASSANO BRESCIANO	BRESCIA	47.03592	68.41194	21.37603
PAVONE DEL MELLA	BRESCIA	48.32589	68.82440	20.49851
MURA	BRESCIA	47.08926	67.39974	20.31048
ORZIVECCHI	BRESCIA	51.44270	71.72300	20.28030
CASTELCOVATI	BRESCIA	40.88648	61.08850	20.20202
CASTREZZATO	BRESCIA	41.55582	61.70899	20.15316
PRALBOINO	BRESCIA	48.21301	67.47677	19.26376
POMPIANO	BRESCIA	47.43590	66.66667	19.23077
QUINTANO	CREMONA	50.96481	70.03405	19.06924
MELLO	SONDRIO	47.39084	66.45367	19.06283

As we've seen in the previous map, this chart is led by the municipality of *Piazzolo*, located in the Alpi Orobiche, followed by a lot of municipalities located in the southern part of Brescia's province.

Now let's see the table of the municipalities with the lowest increments:

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale
CAMPIONE D'ITALIA	COMO	13.42964	14.60674	1.177100
VALGOGLIO	BERGAMO	51.95911	54.85520	2.896082
TRONZANO LAGO MAGGIORE	VARESE	54.85437	57.76699	2.912621
VIGGIU'	VARESE	69.24595	72.38452	3.138571
VAL REZZO	COMO	42.40506	45.56962	3.164557
MORTERONE	LECCO	55.17241	58.62069	3.448276
VALBONDIONE	BERGAMO	62.27181	65.82150	3.549696
SANTA MARGHERITA DI STAFFORA	PAVIA	59.44206	63.30472	3.862661
GROMO	BERGAMO	60.42918	64.37768	3.948498
AGRA	VARESE	44.63840	48.62843	3.990025
MONTE ISOLA	BRESCIA	74.78736	78.85784	4.070474

Here we can find a very heterogeneous scenario: on one side we see *Campione d'Italia*, that despite the name it has proved to be the worst municipality in the response of young people to the vaccine; on the other hand we see municipalities like *Viggiù* and *Monte Isola* which started this analysis with the highest percentage of first doses administered and thus the increments were expected to be low.

Finally, let's see the biggest municipalities increments:

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale
MILANO	MILANO	50.17382	62.93007	12.756249
BRESCIA	BRESCIA	55.51455	67.08132	11.566770
MONZA	MONZA E DELLA BRIANZA	53.16535	66.08406	12.918707
BERGAMO	BERGAMO	53.13325	66.30544	13.172187
COMO	COMO	56.19517	65.14480	8.949627
BUSTO ARSIZIO	VARESE	50.96907	62.83611	11.867037
SESTO SAN GIOVANNI	MILANO	51.59389	63.44414	11.850253
VARESE	VARESE	54.98569	63.95007	8.964380
CINISELLO BALSAMO	MILANO	49.80894	61.16377	11.354828
CREMONA	CREMONA	53.88896	63.57675	9.687793
PAVIA	PAVIA	60.65999	70.87543	10.215440
VIGEVANO	PAVIA	49.44774	61.19824	11.750499

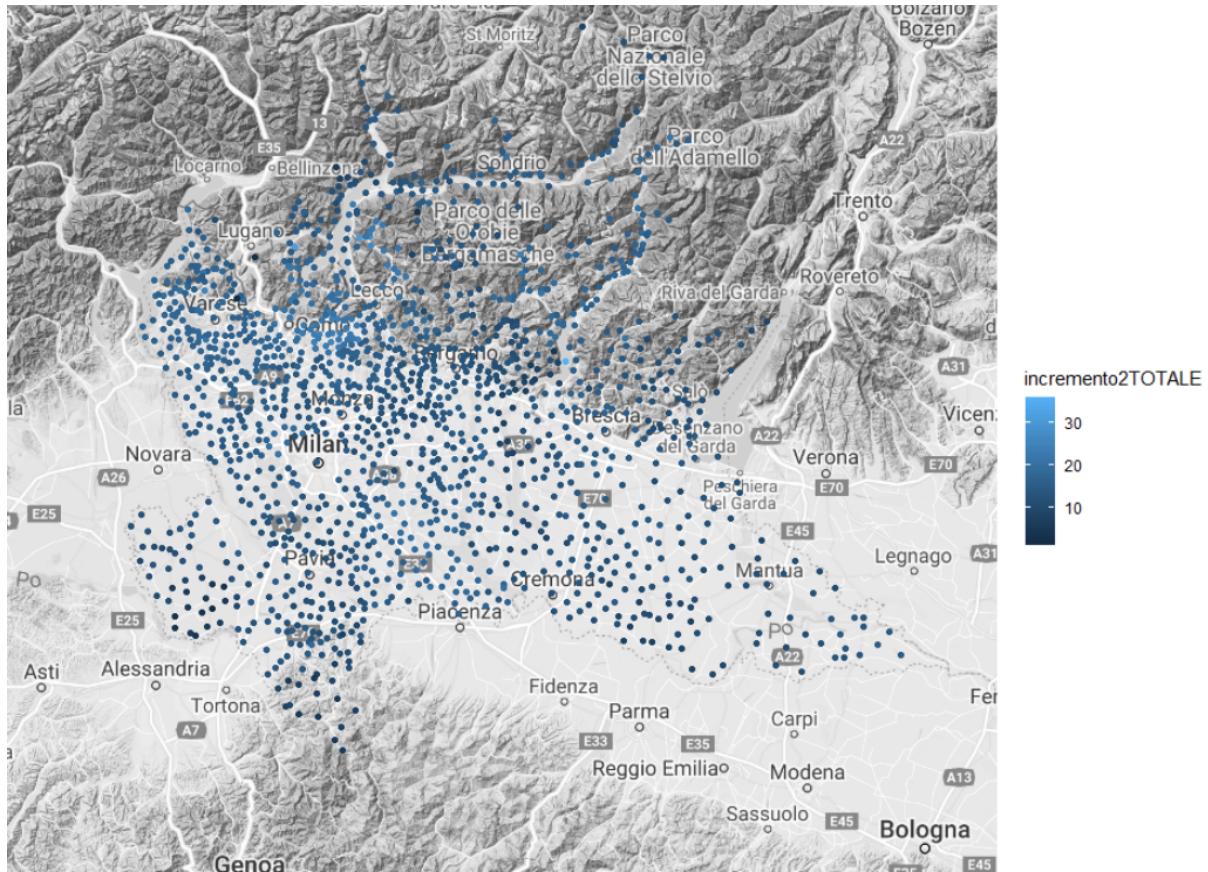
In order to better understand this values we have to compare them to the regional average:

```
> mean(percComuni$incrementoTOTALE)
[1] 12.08244
```

This data is surprising, because in all of the previous tables I've remarked that *Milano* was struggling to keep up with the regional average, while instead it is, alongside *Monza* and *Bergamo*, the only municipality with an increment above average: slowly but surely it will close the gap with the other biggest municipalities like *Brescia*, *Como* and *Pavia*.

To end this analysis I'm going to also show the increments concerning the second doses administered.

```
map+geom_point(aes(x=lon,y=lat,color=incremento2TOTALE), data=percComuni)
```



As we discussed in the previous chapter, the brightest point is the one representing the municipality of *Monte Isola*, which had a very high increase of second doses administered between July 3 and July 8.

Let's see the table of the municipalities with the highest increases:

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale	percSecondaDoseInizio	percSecondaDoseFine	incrementoSecondaDoseTotale
MONTE ISOLA	BRESCIA	74.78736	78.85784	4.070474	31.95626	67.98299	36.02673
PONNA	COMO	67.53247	74.89177	7.359307	17.31602	48.91775	31.60173
PARLASCO	LECCO	77.09924	83.96947	6.870229	25.95420	56.48855	30.53435
CRANDOLA VALSASSINA	LECCO	66.54545	72.36364	5.818182	24.00000	51.27273	27.27273
CORTENOVA	LECCO	68.44828	76.55172	8.103448	29.22414	55.86207	26.63793
LAMBRUGO	COMO	62.17079	69.11413	6.943336	20.59058	46.52833	25.93775
ALSERIO	COMO	62.58879	71.03394	8.445146	20.67877	46.48777	25.80900
BRINZIO	WARESE	63.24461	71.35615	8.111534	24.08112	49.30292	25.22180
ALZATE BRIANZA	COMO	65.56112	73.96746	8.406341	23.46683	47.85148	24.38465
MERONE	COMO	60.77608	70.51409	9.738013	20.01977	44.34009	24.32032

It's clear that the municipalities with the highest increases are the ones that started this analysis with a high percentage of first doses administered: we can see that these two increments are inversely proportional.

Now let's see the table of the municipalities with the lowest increases:

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale	percSecondaDoseInizio	percSecondaDoseFine	incrementoSecondaDoseTotale
GALLIAVOLA	PAVIA	49.13295	56.64740	7.514451	35.260116	36.41618	1.156069
VIGGIU'	VARESE	69.24595	72.38452	3.138571	62.830636	64.72562	1.894986
CAMPIONE D'ITALIA	COMO	13.42964	14.60674	1.177100	9.149278	11.44997	2.300696
TORRE BERETTI E CASTELLARO	PAVIA	42.66667	50.28571	7.619048	17.142857	19.80952	2.666667
PEDESINA	SONDRIO	62.85714	77.14286	14.285714	42.857143	45.71429	2.857143
VALEGGIO	PAVIA	52.28426	56.85279	4.568528	35.532995	38.57868	3.045685
FRASCAROLO	PAVIA	52.44880	62.95637	10.507569	27.248442	30.54319	3.294746
VILLA BISCOSSI	PAVIA	49.15254	59.32203	10.169492	42.372881	45.76271	3.389831
PARZANICA	BERGAMO	46.08939	55.86592	9.776536	36.871508	41.06145	4.189944
GAMBARANA	PAVIA	52.47525	62.37624	9.900990	33.663366	38.11881	4.455446

Except for *Viggiù*, which started this analysis with high percentages of first and second doses administered, and few others, like *Pedesina*, in this table we can find the municipalities that had the worst response to the vaccination campaign: low percentages at the start and low increments during the analysis period.

Once again *Campione d'Italia* confirm his negative trend.

Finally, let's see the biggest municipalities increments:

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale	percSecondaDoseInizio	percSecondaDoseFine	incrementoSecondaDoseTotale
MILANO	MILANO	50.17382	62.93007	12.756249	21.29311	35.65655	14.363443
BRESCIA	BRESCIA	55.51455	67.08132	11.566770	26.07457	39.90067	13.826101
MONZA	MONZA E DELLA BRIANZA	53.16535	66.08406	12.918707	24.02910	37.06637	13.037271
BERGAMO	BERGAMO	53.13235	66.30544	13.172187	28.76241	40.89861	12.227198
COMO	COMO	56.19517	65.14480	8.949627	24.53424	43.13508	18.600840
BUSTO ARSIZIO	VARESE	50.96907	62.83611	11.867037	22.92201	38.60637	15.684364
SESTO SAN GIOVANNI	MILANO	51.59389	63.44414	11.850253	22.53409	36.67126	14.137165
VARESE	VARESE	54.98569	63.95007	8.964380	26.18598	41.87334	15.687352
CINISELLO BALSAMO	MILANO	49.80894	61.16377	11.354828	22.17251	33.37096	11.198444
CREMONA	CREMONA	53.08896	63.57675	9.6887793	31.66813	43.65169	11.983558
PAVIA	PAVIA	60.65999	70.87543	10.215440	32.13848	46.24847	14.109989
VIGEVANO	PAVIA	49.44774	61.19824	11.750499	23.07110	37.35751	14.286404

In order to better understand this values we have to compare them to the regional average:

```
> mean(percComuni$incremento2TOTALE)
[1] 14.16335
```

Like we saw with the first doses, once again *Milano* shows some encouraging signals, with an increment slightly above the average. Also *Pavia* had a very good increase, leading the way with almost half of its population fully immunized.

4.2 MY HOME MUNICIPALITY: PROVAGLIO D'ISEO

The last thing I wanted to analyze was the results of the municipality in which I live: *Provaglio d'Iseo*.

comune	provincia	percPrimaDoseInizio	percPrimaDoseFine	incrementoPrimaDoseTotale	percSecondaDoseInizio	percSecondaDoseFine	incrementoSecondaDoseTotale
PROVAGLIO D'ISEO	BRESCIA	49.8684	66.83751	16.96911	20.63998	35.53124	14.89126

To better understand this data, I have to report all the final means that I've calculated: first doses administered, second doses administered, first doses increase, second doses increase.

```
> mean(percComuni$percPrimaDoseFine)
[1] 64.87331
> mean(percComuni$percSecondaDoseFine)
[1] 37.46741
> mean(percComuni$incrementoTOTALE)
[1] 12.08244
> mean(percComuni$incremento2TOTALE)
[1] 14.16335
```

I can proudly say that I live in a good municipality: seeing that high increment in the percentage of first doses administered confirm the high expectation that I had, because I knew that a lot of young people had received their first doses, but until now I couldn't give a value to this data.

5. CONCLUSION

Doing this analysis was very stimulating: I have always been fascinated by the world of data analysis, and being able to apply the techniques learned in class to analyze in detail the progress of the most important vaccination campaign that modern history has ever seen was a great opportunity.

Seeing these data, I feel confident in saying that herd immunity is not as far away as many imagine: it will be essential to raise awareness among the part of the population that has not yet been vaccinated.