LIFE DATA EPIDEMIOLOGY

HUMAN MOBILITY AND COVID-19 EPIDEMIC

FATIMA RABIA YAPICIOGLU – 2049536

FIKRIYE SARI - 2049592

INTRODUCTION

IS HUMAN MOBILITY A RELEVANT INFORMATION FOR THE COVID-19 SPREADING?

The COVID-19 pandemic has spread rapidly all over the world, affecting many countries to varying degrees. Countries in Europe took different mobility containment measures to curb the spread of COVID-19. The European Commission asked Mobile Network Operators to share on a voluntarily basis anonymized and aggregate mobile data to improve the quality of modelling and forecasting for the pandemic at EU level.

The recent health emergency caused by the COVID-19 pandemic has forced people to change their mobility behaviors, with the reduction of leisure travels and the promotion of teleworking and online educational activities (de Vos, 2020). Among the most applied contagion control measures, those relating to the limitation of travels, the so-called "stay-at- home" order, have become widespread with the aim of avoiding the circulation of the virus in public environments. Public transport has been highly impacted both by government restrictions and travelers' choices (Jenelius and Cebecauer, 2020; Gutiérrez et al. 2020).

Setting Objectives and Goals	Data Acquisation and Exploration	Data Preprocessing	Data Visualization	Statistical Calculations	Gathering Analysis Results

PURPOSE OF THE PROJECT

HUMAN MOBILITY AND COVID-19 EPIDEMIC

One of the most important policies for controlling COVID-19 spread is mobility limitation. However, there are a variety of ways to limit movement through diversified limitations, and it's difficult to forecast how these restrictions would affect virus spread. For policymakers who need to enact effective and timely policies, this is a limitation. Despite the importance of data analysis in understanding this phenomenon, more general models capable of anticipating the impact of various events are also required.

Focus questions of this study are,

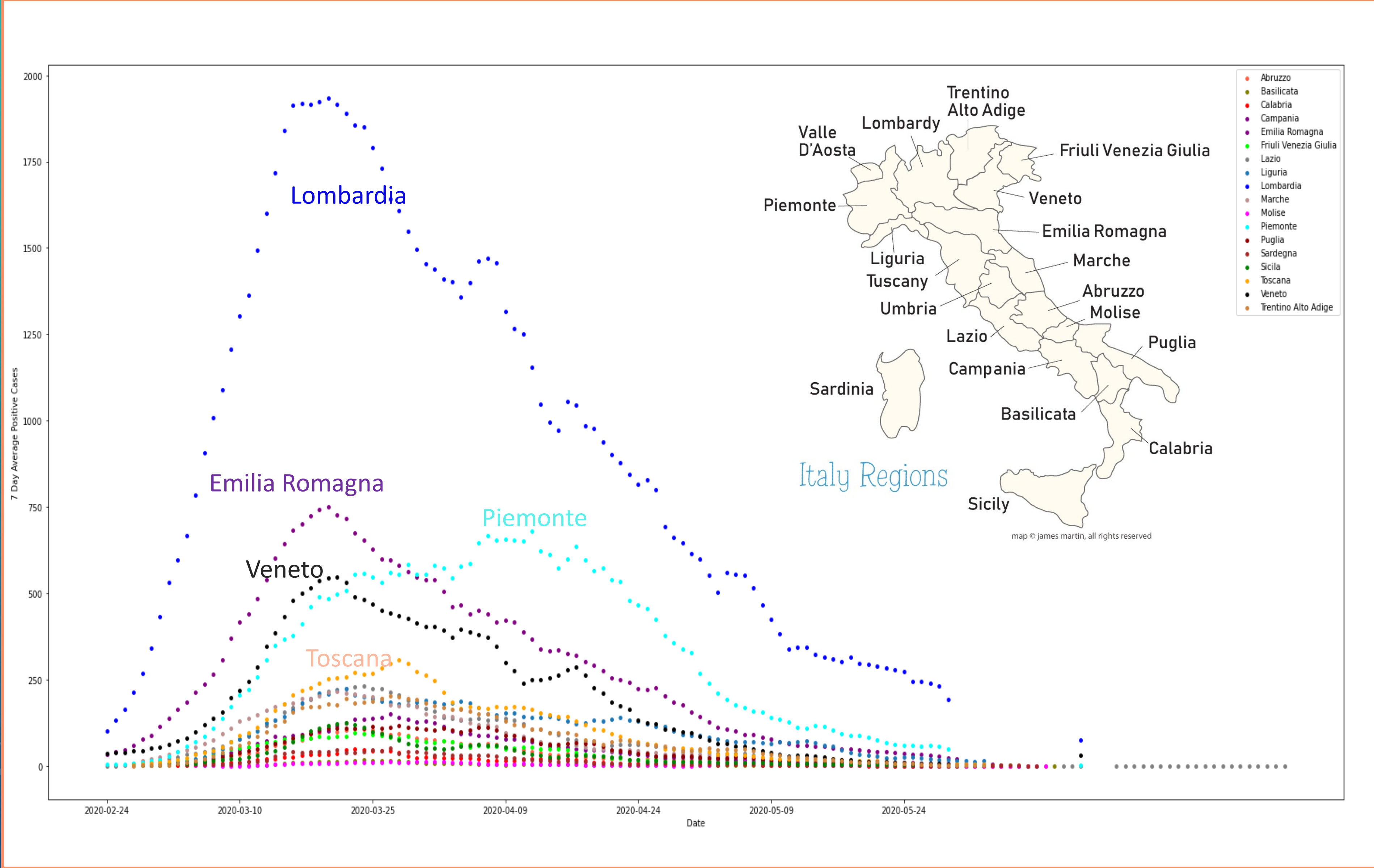
Show how North Italy has both faster dynamics of human mobility & variations and the speed of COVID-19 spreading compared to other regions in Italy.

DATA ACQUISATION & EXPLORATION

WHICH DATA SOURCES WE USED IN ORDER TO ACTUALIZE OUR GOALS

1	Google	**COVID-19 Community Mobility Reports**	**The reports chart movement trends over time by geography, across different categories of places such as retail and recreation, groceries and pharmacies, parks, transit stations, workplaces, and residential**
2	SCIENTIFIC DATA	COVID-19 outbreak response assess mobility changes in Italy	Pepe, Bajardi, Gauvin, Privitera, Lake, Cattuto, Tizzoni, Scientific Data, 230 (2020): Daily origin-destination matrices, daily average degree of users' proximity network.
3	COVIDSTAT INFN	covid19.infn.it Project	It is a project carried out by the CovidStat INFN Working Group.

7 DAY AVERAGE POSITIVE CASES VERSUS REGIONS



Region by Regions Comparison

We make the following assumptions,

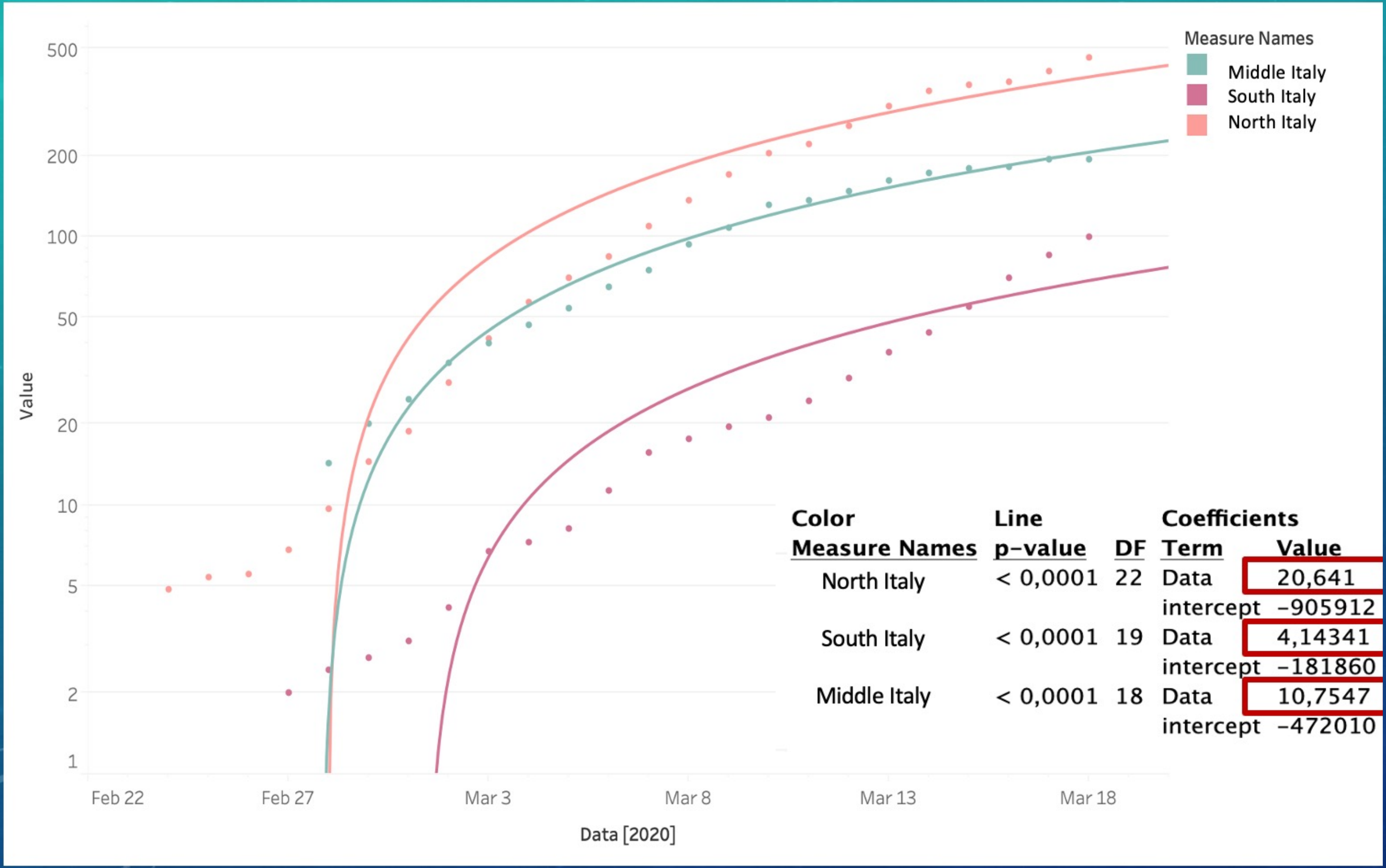
North Italy includes the following regions: Lombardia, Veneto, Trentino Alto Adige, Tuscany, Piemonte, Emiglia Romagnia, and Liguria.

Middle Italy includes regions: Umbria, Marche, Sardegnia, Lazio, Abruzzo

South Italy includes regions: Puglia, Sicilia, Basilicata, Campania, Calabria and Molise.

This graph shows the daily number of positive cases specific to each region. But in order the understand the epidemic is faster in the North Italy we should take the y-axis in the logarithmic scale. Because the intensity of cases depends on the population. It may also be named as normalization.

POSITIVE CASES IN LOGARITHMIC SCALE & LINEAR REGRESSION, EXPONENTIAL FACTOR

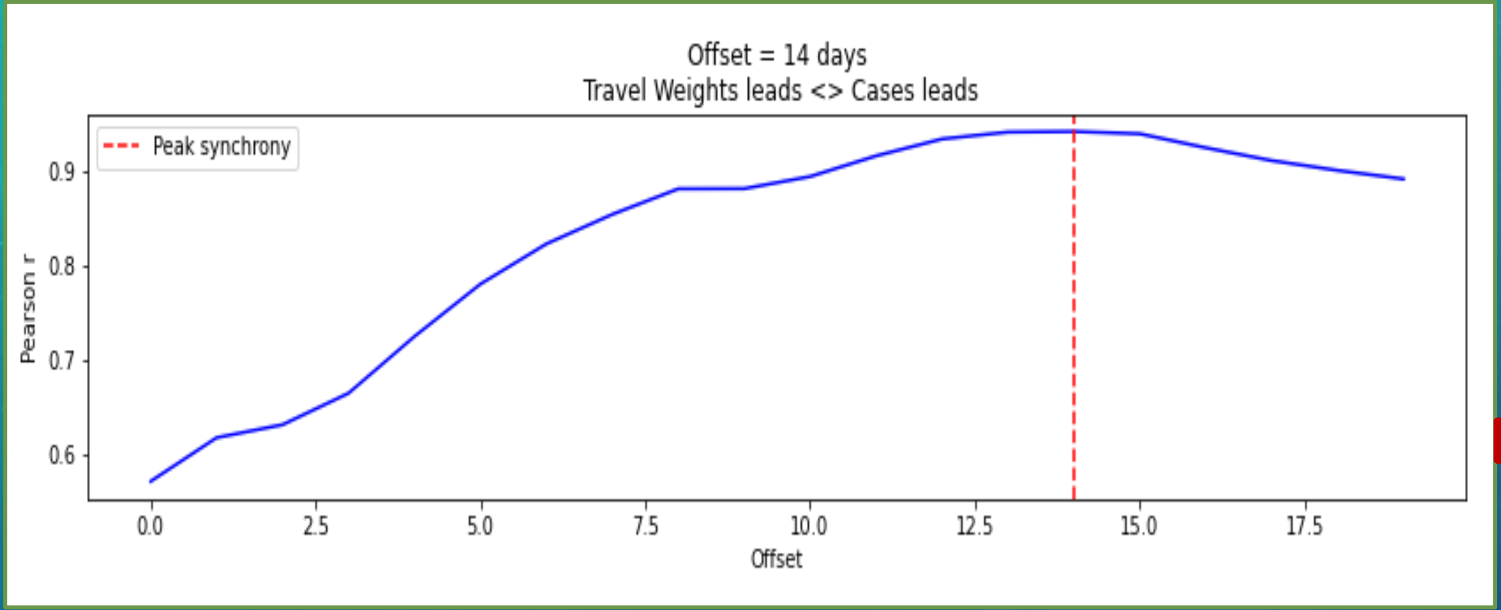


Color Measure Names	Line p-value	DF	Coefficients Term	Value
North Italy	< 0,0001	22	Data	20,641
			intercept	-905912
South Italy	< 0,0001	19	Data	4,14341
			intercept	-181860
Middle Italy	< 0,0001	18	Data	10,7547
			intercept	-472010

We take the Logarithmic scale of the daily number of positive cases in the y-axis. Now we normalized the data depending on each region's population. Then, we fitted a linear line through the plot the following table shows the exponential factors of each line.

We observed that the exponential factor for the North Italy line has the highest value, meaning that it has the greatest slope. The North Italy seems to have a faster dynamics in spreading of COVID-19. Does the human mobility that cause this? Let's see the human mobility dynamics in overall Italy.

LAG ANALYSIS BETWEEN WEIGHT IF FRACTION OF TRAVELLERS AND 7 DAY AVERAGE DAILY POSITIVE CASES



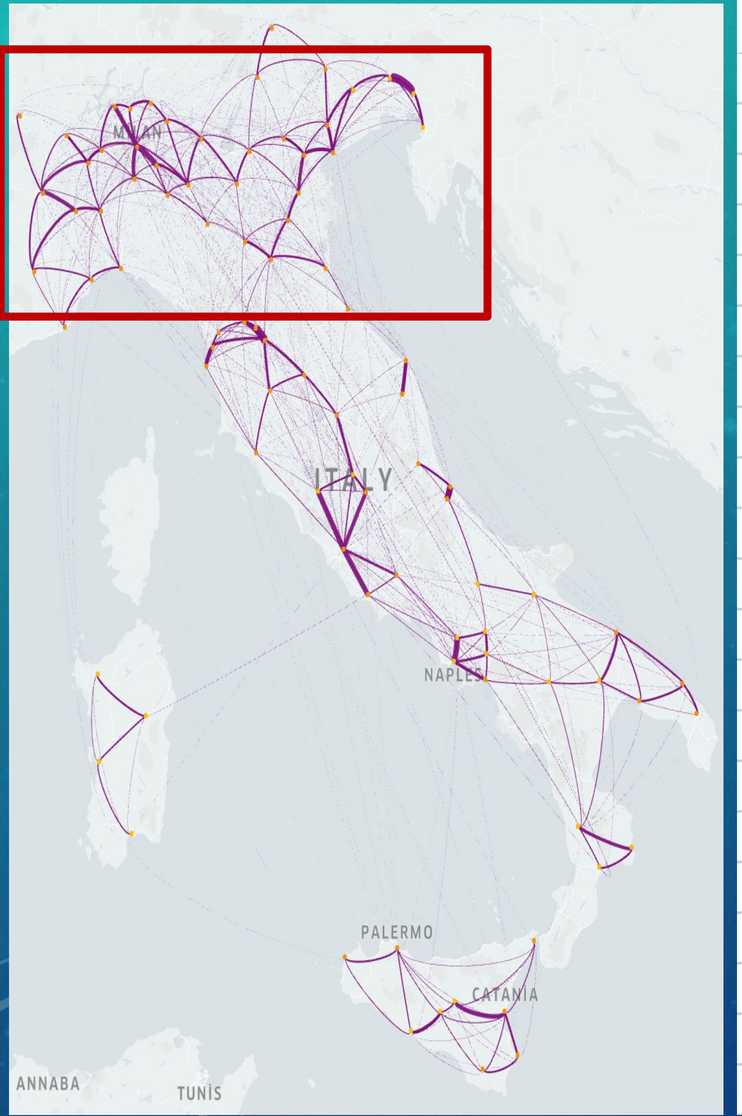
Pearson r:	0.5715219338268225	when the time lag is 0 days
Pearson r:	0.6177354341455426	when the time lag is 1 days
Pearson r:	0.6314799152578151	when the time lag is 2 days
Pearson r:	0.6647652076534701	when the time lag is 3 days
Pearson r:	0.7248484741271228	when the time lag is 4 days
Pearson r:	0.7802648117826267	when the time lag is 5 days
Pearson r:	0.8228057579169747	when the time lag is 6 days
Pearson r:	0.8540476793372336	when the time lag is 7 days
Pearson r:	0.8810706856158124	when the time lag is 8 days
Pearson r:	0.8812483013463698	when the time lag is 9 days
Pearson r:	0.8940719149998057	when the time lag is 10 days
Pearson r:	0.9159404790925623	when the time lag is 11 days
Pearson r:	0.933829743068145	when the time lag is 12 days
Pearson r:	0.9411219421604727	when the time lag is 13 days
Pearson r:	0.9417555856855426	when the time lag is 14 days
Pearson r:	0.9393143640467273	when the time lag is 15 days
Pearson r:	0.9245107795608509	when the time lag is 16 days
Pearson r:	0.9109437471613927	when the time lag is 17 days
Pearson r:	0.9005493446717041	when the time lag is 18 days
Pearson r:	0.8915051723956124	when the time lag is 19 days

Let's see if there is a correlation between the human mobility and log scaled positive number of cases by applying lag analysis.

Notice that the correlation between the two time series becomes less and less positive as the number of lags increases, this tells us that human mobility during a given day is quite predictive of number of positive cases 14 days later, but not predictive of positive cases beyond more than 14 days. After the day 14, it started to decrease.

AVERAGE WEIGHT OF TRAVELLERS FRACTIONS BY REGION – 14.02.20

12.03.20



12.03.20

P1	P2	
Abruzzo	Abruzzo	3,237%
Aosta	Aosta	1,121%
Basilicata	Basilicata	2,071%
Calabria	Calabria	3,317%
Campania	Campania	5,550%
Emilia-Romagna	Emilia-Romagna	7,460%
Friuli-Venezia Giulia	Friuli-Venezia Giulia	4,383%
Lazio	Lazio	5,469%
Liguria	Liguria	3,339%
Lombardy	Lombardy	10,887%
Marche	Marche	2,140%
Molise	Molise	1,032%
Piedmont	Piedmont	7,727%
Puglia	Puglia	5,553%
Sardinia	Sardinia	3,351%
Sardiniaari	Sardiniaari	1,055%
Sicilia	Sicilia	10,207%
Trentino-Alto Adige-Sudti..	Trentino-Alto Adige-Sudti..	2,233%
Tuscany	Tuscany	10,043%
Umbria	Umbria	2,106%
Veneto	Veneto	7,720%

26 March 2020:
New Cases: 6203
7 Day Average: 5651

We prefer to take the 12.03.0 as a reference date, because according to our calculation 14 days are needed for the peak synchrony. Therefore, the effects of mobility will be observable in terms of cases after 14 days. Origin Destination Matrix Fraction of Travelers Percentage of All Italy.

North Italy: 57,235%

South Italy: 27,73%

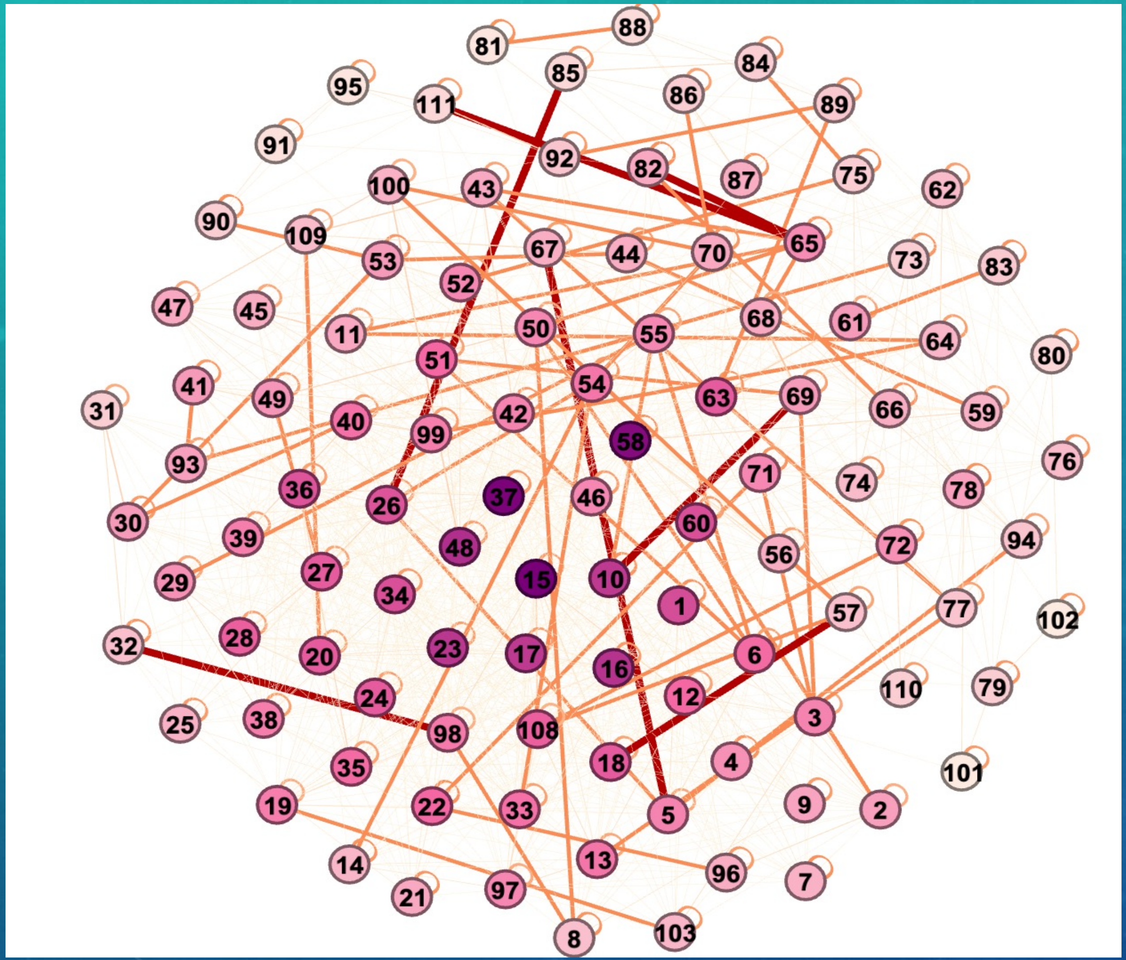
Middle Italy: 15,01%

12 Feb 2020, two weeks before the first cases have been detected, at the very beginning of the pandemics has the visualized weights of the fraction of travelers as shown in the Figure. The cities in the north are strongly connected with each other.

Does this mobility cause the faster spreading of COVID-19 than in the other cities?

CENTRALITY CALCULATIONS ON 14.02.20 MOBILITY NETWORK DATA

14.02.20 – Nodes according degree



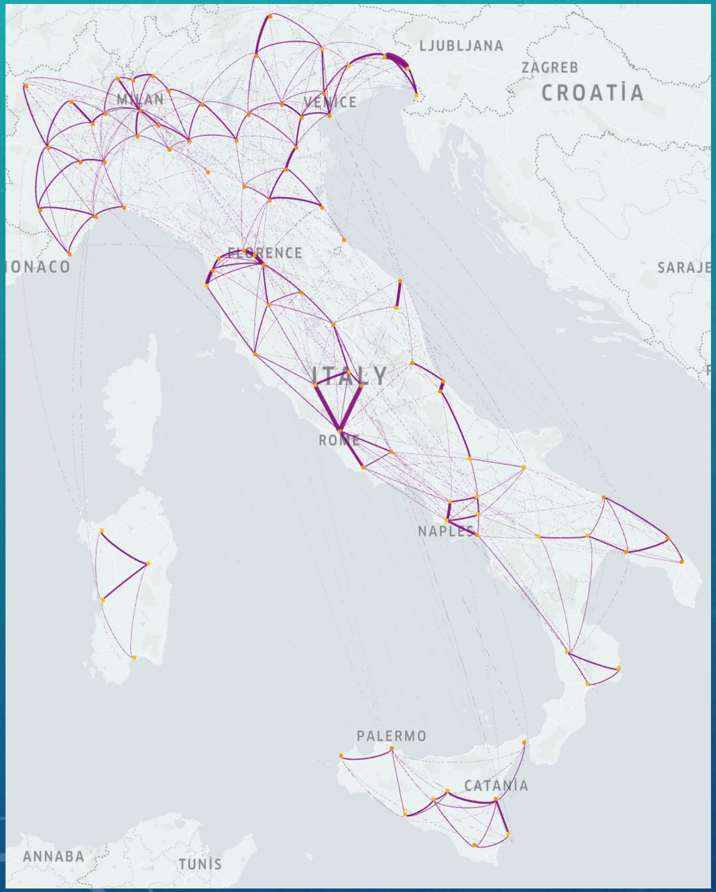
The node 15(Milano), 37(Bologna), 58(Roma), 48(Firenze), 16(Bergamo), 23(Verona), 17(Brescia), 10(Genova), 1(Torino), 34(Parma), 60(Caserta) are the nodes which have the highest number of degrees.

Before the pandemics gain speed and show an exponential increase in the daily number of positive cases, the cities in the north have generally the highest degrees and strong weights which shows the fraction of travelers between those.

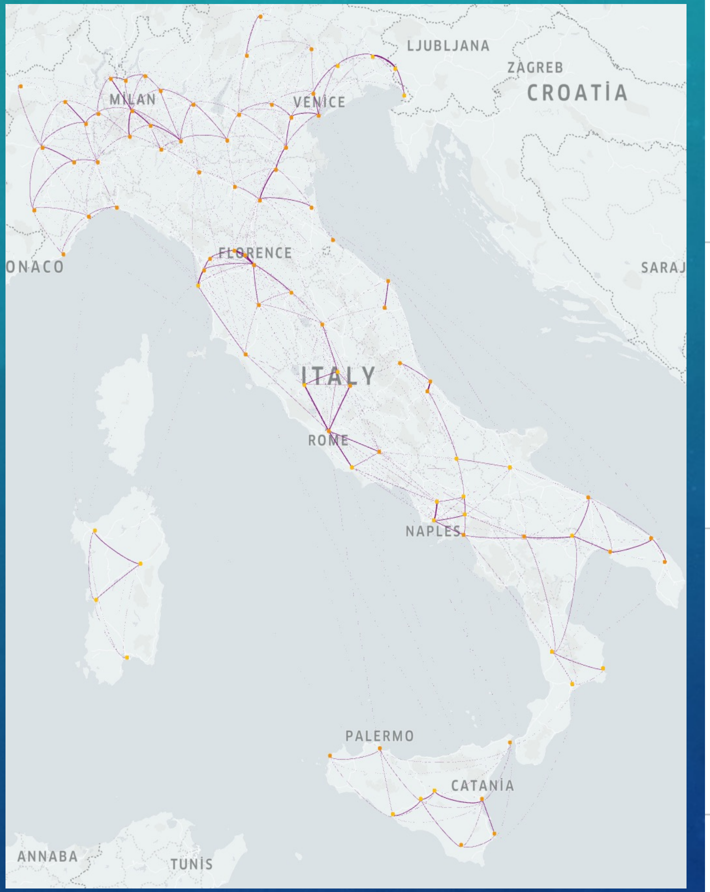
DATA VISUALIZATION & ANALYSIS

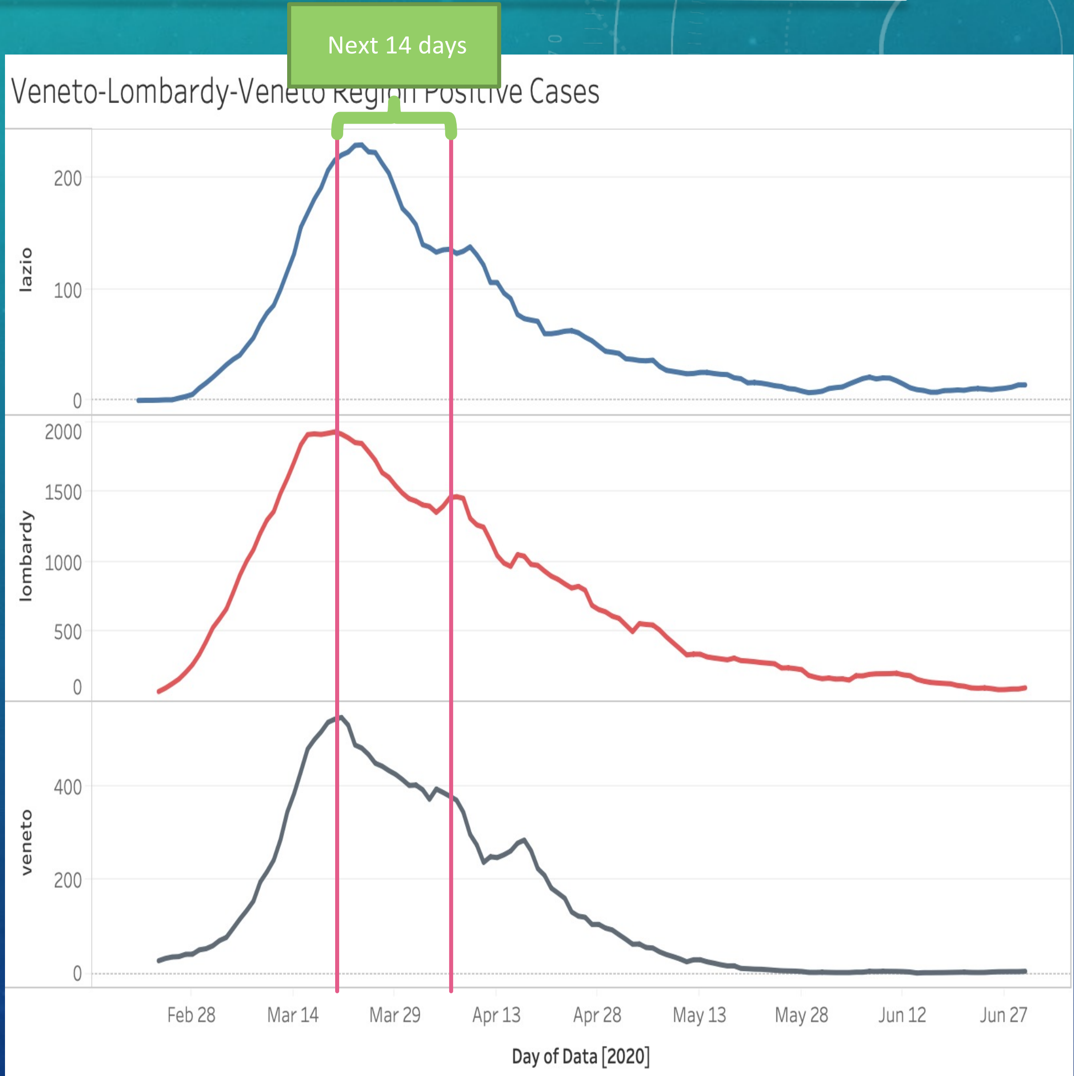
EARLY IN THE COVID-19 RESTRICTIONS IN LODI & PADUA

The Day Before Restrictions, 20.03.20



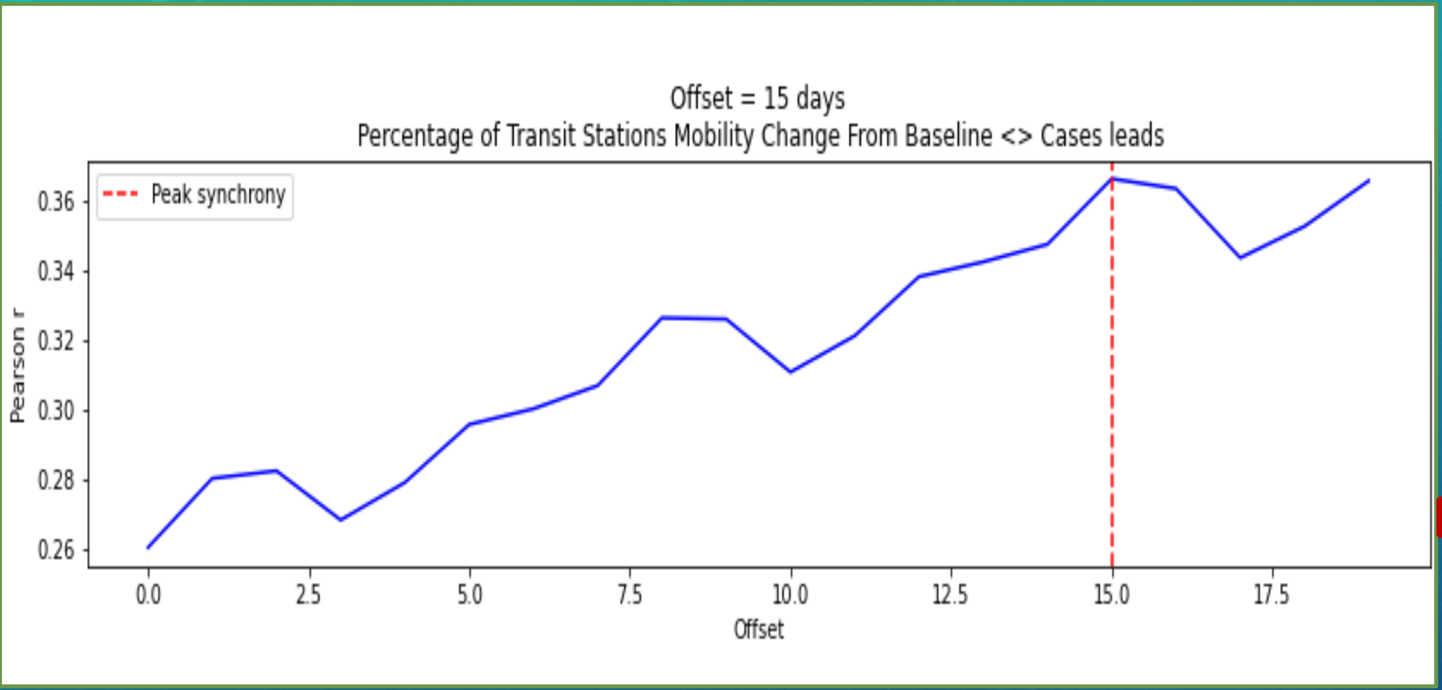
Next Day, 21.03.20





The trends of sum of lazio (iss_bydate_lazio_positivi), sum of lombardy (iss_bydate_lombardia_positivi) and sum of veneto for Data Day. The view is filtered on Data Day, which ranges from February 20, 2020 to June 30, 2020.

VARIATIONS IN MOBILITY: LAG ANALYSIS BETWEEN TRANSIT PERCENTAGE CHANGE FROM BASELINE AND 7 DAY AVERAGE DAILY POSITIVE CASES

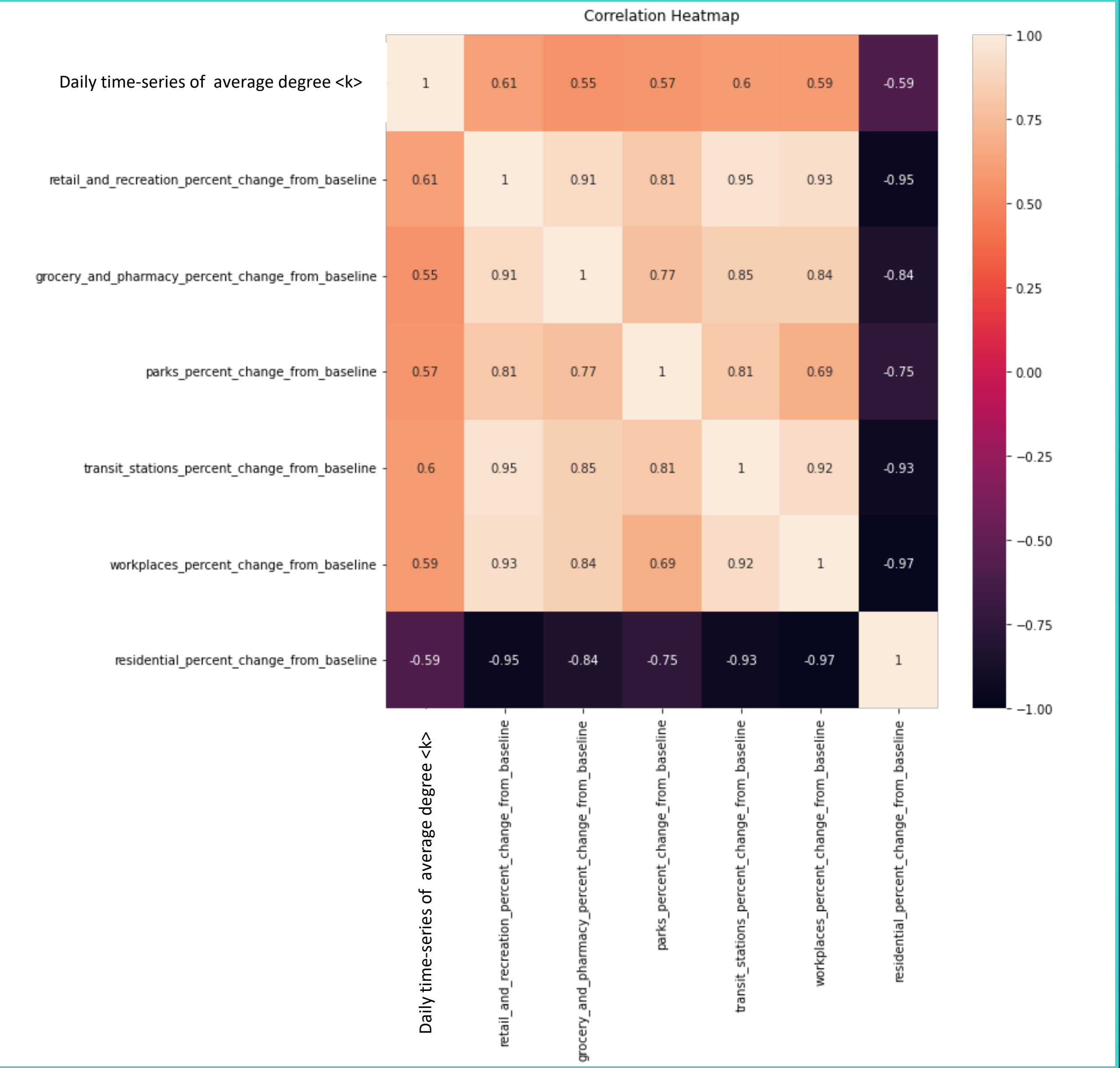



```
Pearson r:  0.26040902281426426 when the time lag is  0  days
Pearson r:  0.28021167876175523 when the time lag is  1  days
Pearson r:  0.2823861902898285 3 when the time lag is  2  days
Pearson r:  0.2682986487665757 when the time lag is  3  days
Pearson r:  0.2790642856034291 when the time lag is  4  days
Pearson r:  0.29566469649417065 when the time lag is  5  days
Pearson r:  0.3001647325861599 when the time lag is  6  days
Pearson r:  0.306813230064884 when the time lag is  7  days
Pearson r:  0.32620769054419946 when the time lag is  8  days
Pearson r:  0.3258688708944837 when the time lag is  9  days
Pearson r:  0.31068020447431655 when the time lag is  10  days
Pearson r:  0.3210696487206448 when the time lag is  11  days
Pearson r:  0.33800954326117455 when the time lag is  12  days
Pearson r:  0.34225883448411965 when the time lag is  13  days
Pearson r:  0.3472909674284202 when the time lag is  14  days
Pearson r:  0.3660727046415668 when the time lag is  15  days
Pearson r:  0.3632567139899392 when the time lag is  16  days
Pearson r:  0.34342431721935157 when the time lag is  17  days
Pearson r:  0.35244828990806604 when the time lag is  18  days
Pearson r:  0.36548390114710466 when the time lag is  19  days
```

Notice that the correlation between the two time series becomes less and less positive as the number of lags increases, this tells us that human mobility during a given day is quite predictive of number of positive cases 14 days later, but not predictive of positive cases beyond more than 14 days. After the day 14, it started to decrease.

STATISTICAL CALCULATIONS

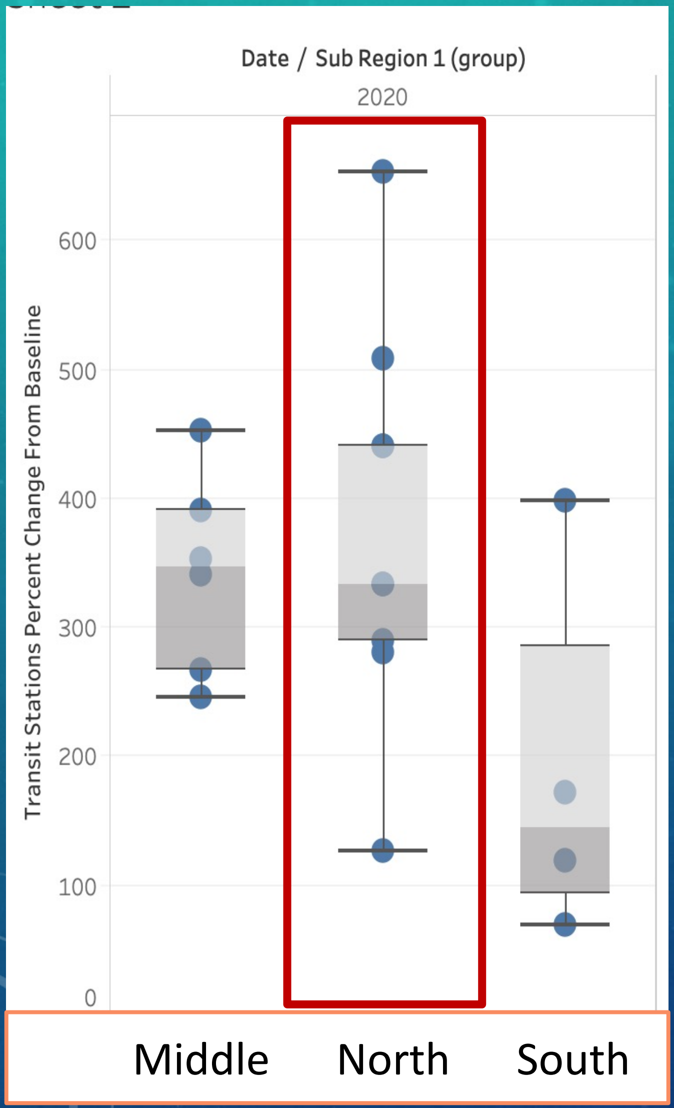
VARIATIONS IN TERMS OF MOBILITY AND THEIR CORRELATION ANALYSIS WITH PROXIMITY NETWORK DATA.

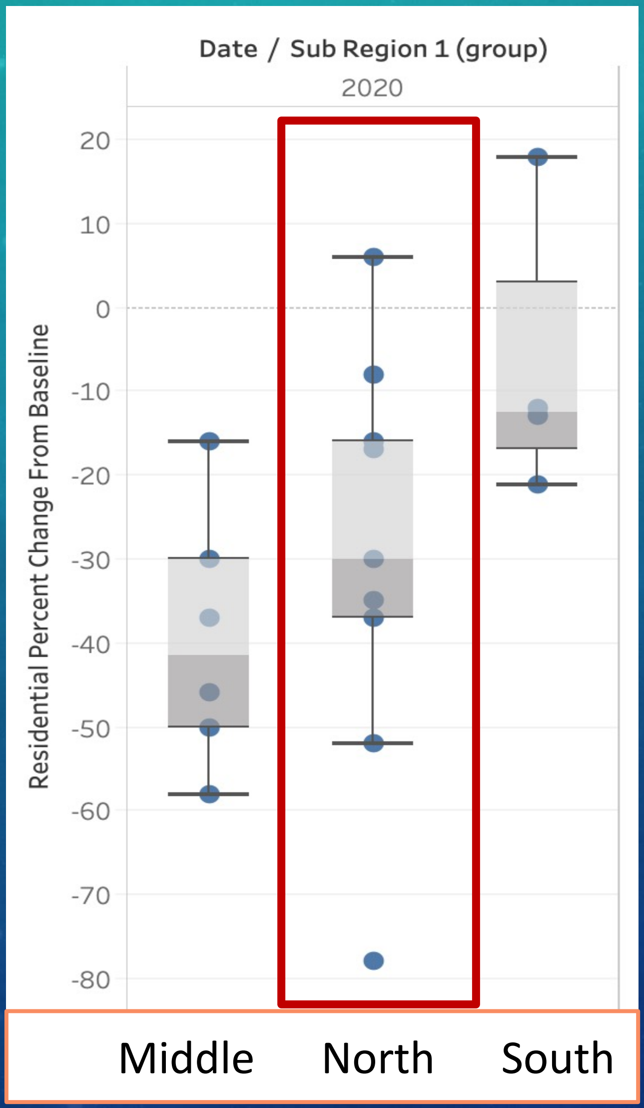


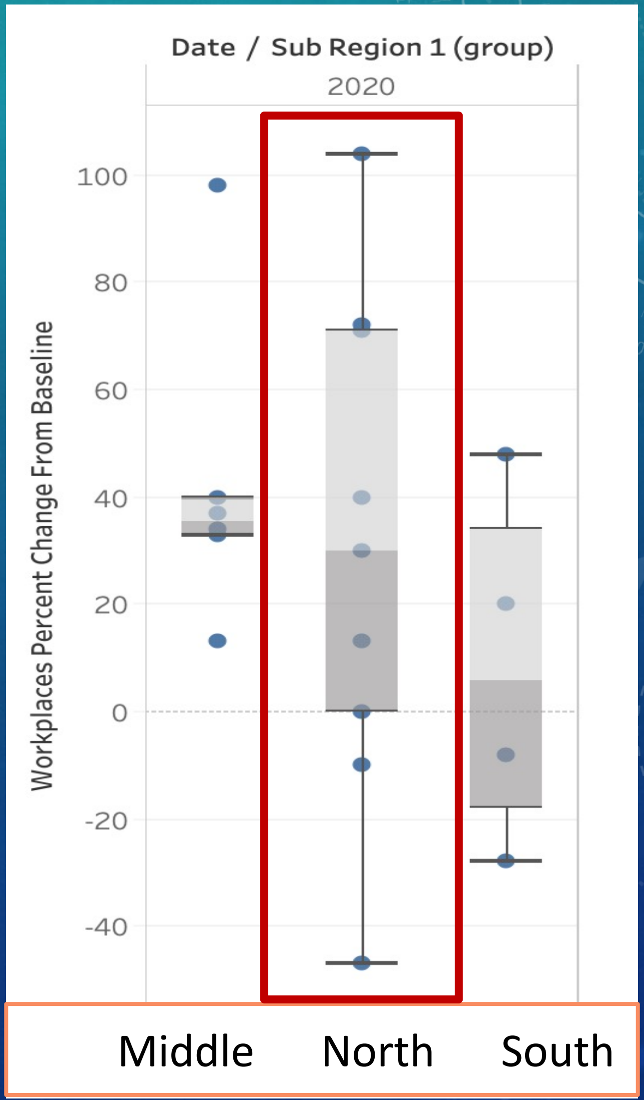
Correlation analysis; Pearson's correlation coefficient between time-series of mobility reductions reported by Google and daily time-series of average degree <k> of the proximity network. The time scale including the time 18.01.2020 and 27.03.20

DATA VISUALIZATION & ANALYSIS

VARIATIONS FROM THREE DIFFERENT PERSPECTIVES BETWEEN 15-02-2020 & 23-02-20, BEFORE THE QUARANTINE







CONCLUSIONS

North Italy has both the highest dynamics of mobility and the covid-19 spreading. In this sense North Italy has the highest exponential factor when compared to other regions in Italy (according to the regions we compiled).

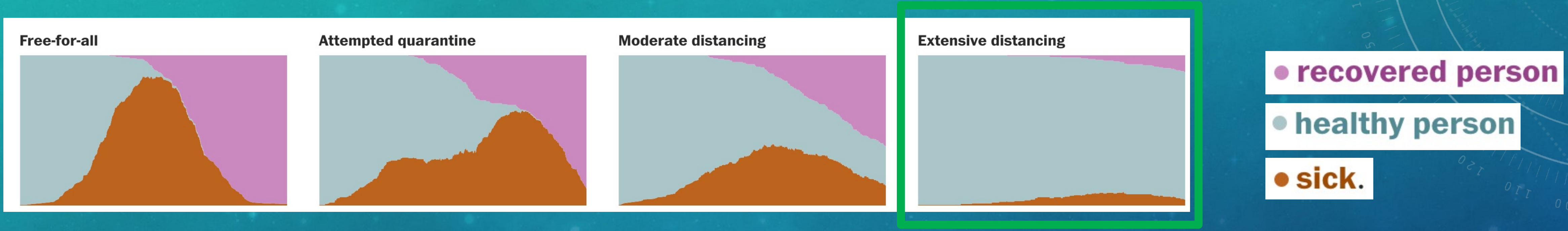
According to lag analysis there is a correlation between the human mobility and daily number of covid-19 cases with 14 days shift.

The average degree of the proximity network and variations in human mobility is highly correlated.

North Italy shows the highest rates also in the variations in human mobility.

REFERENCES

1-)Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve", By Harry Stevens March 14, 2020
https://www.washingtonpost.com/graphics/2020/world/corona-simulator/



we let just one of every eight people move.
Italy is closing all of its restaurants.

"Even with different results, moderate social distancing will usually outperform the attempted quarantine, and extensive social distancing usually works best of all."

2-) https://kepler.gl/demo

3-) Coşkun H, Yıldırım N, Gündüz S. The spread of COVID-19 virus through population density and wind in Turkey cities. Sci Total Environ. 2021 Jan 10;751:141663. doi: 10.1016/j.scitotenv.2020.141663. Epub 2020 Aug 11. PMID: 32866831; PMCID: PMC7418640.
4-) Cartenì, A., Di Francesco, L., & Martino, M. (2020). How mobility habits influenced the spread of the COVID-19 pandemic: Results from the Italian case study. *The Science of the total environment*, *741*, 140489. https://doi.org/10.1016/j.scitotenv.2020.140489
5-) https://covid19.infn.it/,
6) https://www.google.com/covid19/mobility/
7) Pepe, Bajardi, Gauvin, Privitera, Lake, Cattuto, Tizzoni, Scientific Data, 230 (2020): Daily origin-destination matrices, daily average degree of users' proximity network.
8) https://covid19.apple.com/mobility

THANK YOU
QUESTIONS ARE WELCOME!