CIs

ANDREA BOSELLI, LEONARDO PERELLI,
VALENTINA SGARBOSSA & RANDEEP SINGH

are proud to present

Safari Njema 2.0

How Covid changed our movements

Transport modes analysis

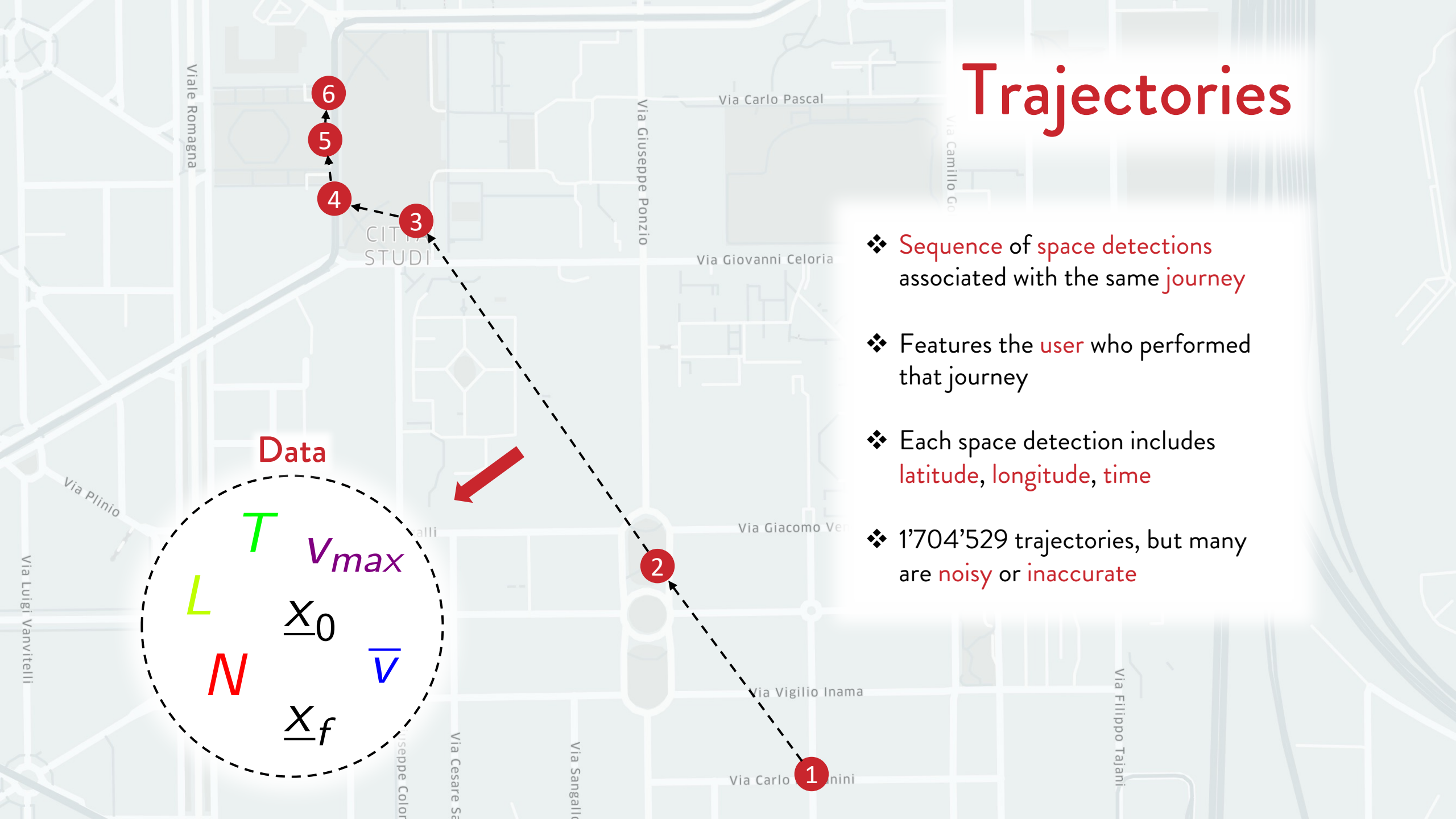
PCA

Clustering

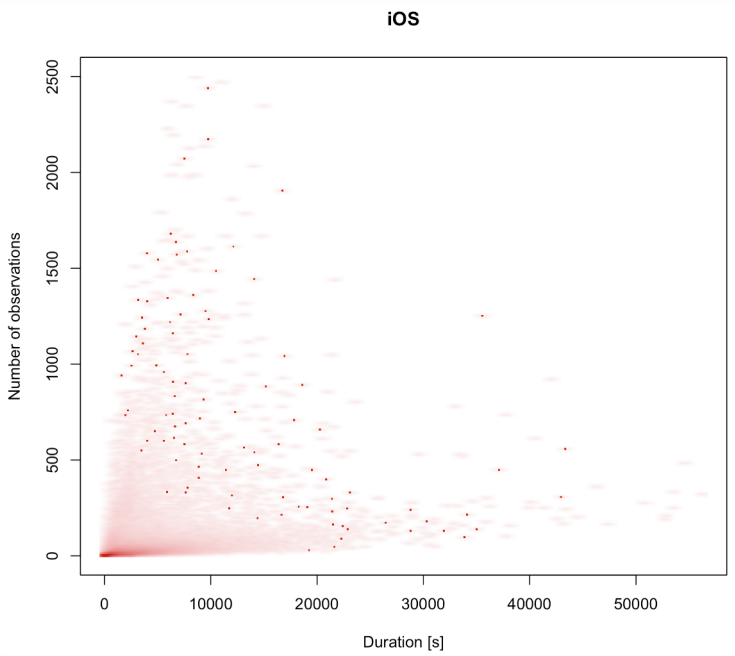
FDA

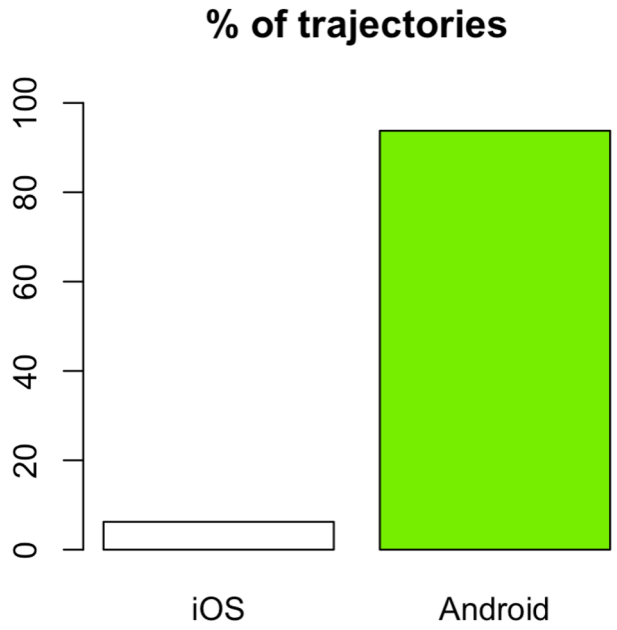
Trajectories

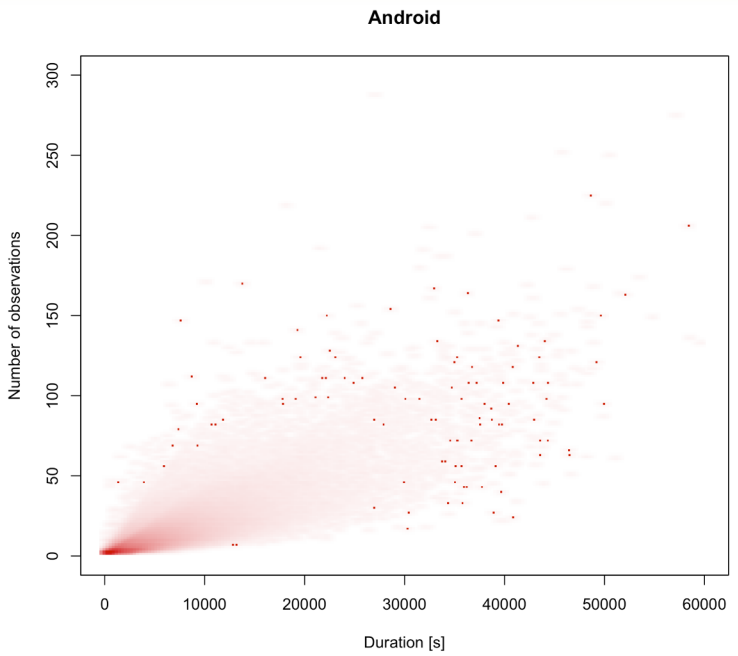
- Sequence of space detections associated with the same journey
- Features the user who performed that journey
- Each space detection includes latitude, longitude, time
- 1'704'529 trajectories, but many are noisy or inaccurate



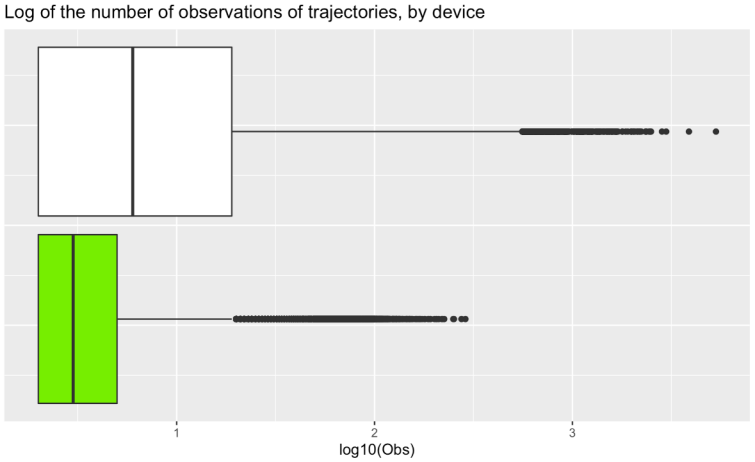
Data Detection





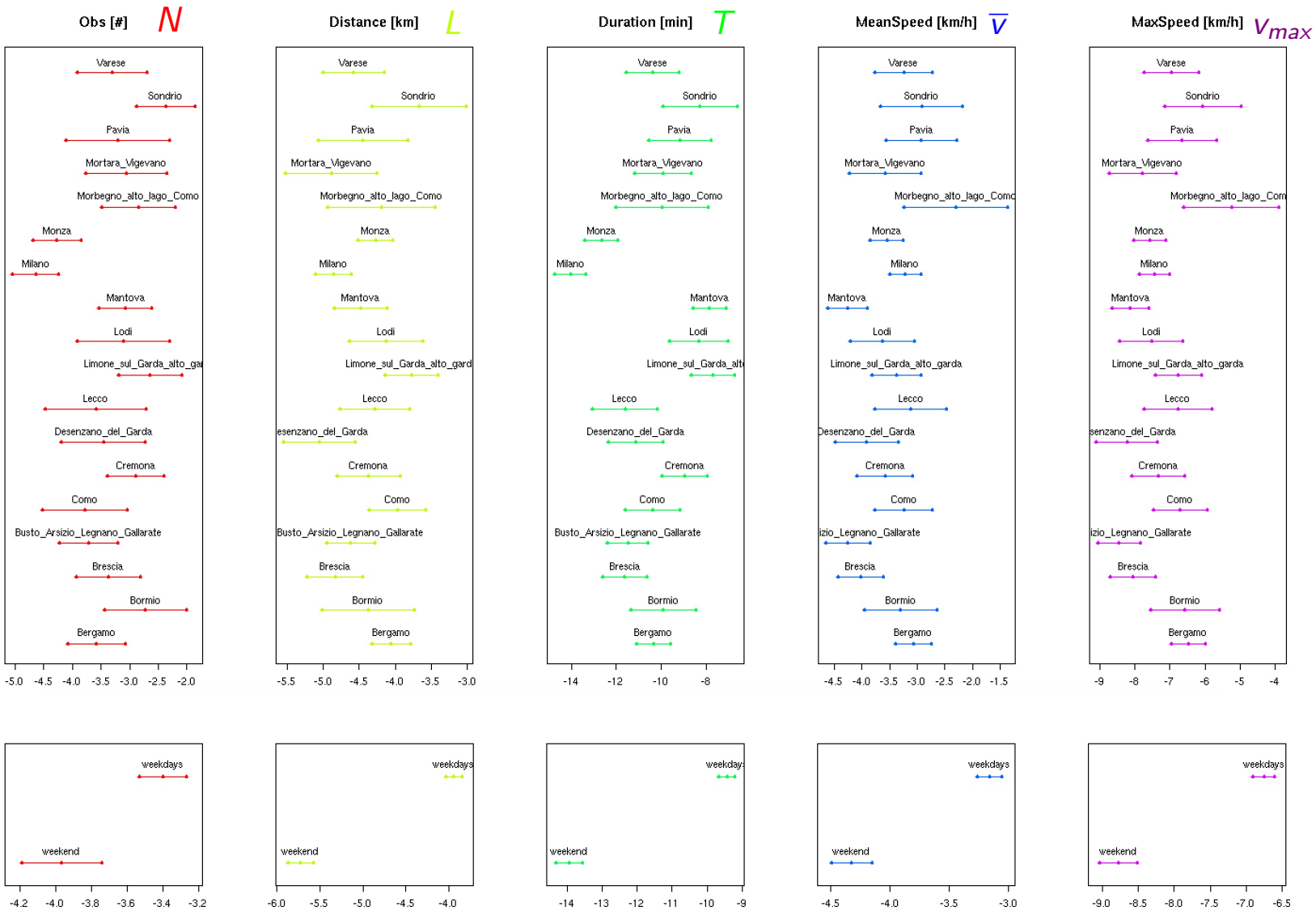


- iOS performs more space detections per trajectory than Android: 23.54 on average
- Represents only 15.5% of the devices, and 6.23% of the trajectories



- Android performs its detections at regular time intervals
- On average 4.88 detections per trajectory: coarse estimate

Trends of change wrt lockdown



Within each group (e.g. zone, time of the week) we computed CIs for the change of mean due to lockdown:

$$\overline{X}_i \sim N_5(\underline{\mu}_i, \frac{\Sigma_i}{n_i}) \quad \overline{X}_f \sim N_5(\underline{\mu}_f, \frac{\Sigma_f}{n_f})$$

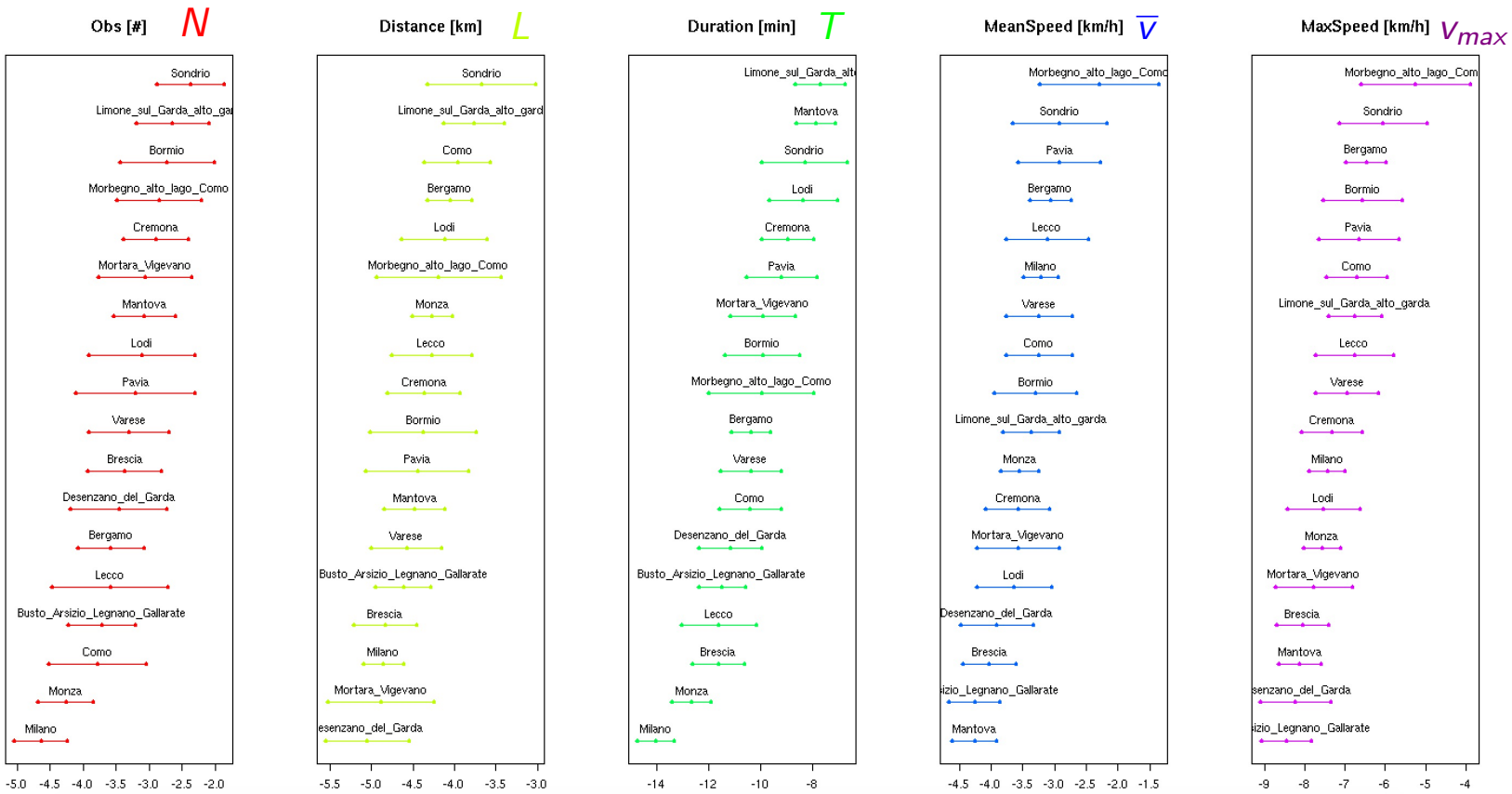
$$\Rightarrow \underline{D} := \overline{X}_f - \overline{X}_i \sim N_5(\underline{\delta}, \Sigma)$$

$$\underline{\delta} = \underline{\mu}_f - \underline{\mu}_i \quad \Sigma = \frac{\Sigma_f}{n_f} + \frac{\Sigma_i}{n_i}$$

$$\Rightarrow CI_{1-\alpha}(\underline{a}'\underline{\delta}) = [\underline{a}'\underline{D} \pm \sqrt{\underline{a}' S_p \underline{a}} \; z_{1-\frac{\alpha}{2k}}]$$

At level 95%, lockdown decreased the means within each of the considered groups. There seems to be a relation between changes in mean and zone: population, area

Trends of change wrt lockdown



Within each group (e.g. zone, time of the week) we computed CIs for the change of mean due to lockdown:

$$\underline{\overline{X}}_i \sim N_5\left(\underline{\mu}_i, \frac{\Sigma_i}{n_i}\right) \qquad \underline{\overline{X}}_f \sim N_5\left(\underline{\mu}_f, \frac{\Sigma_f}{n_f}\right)$$

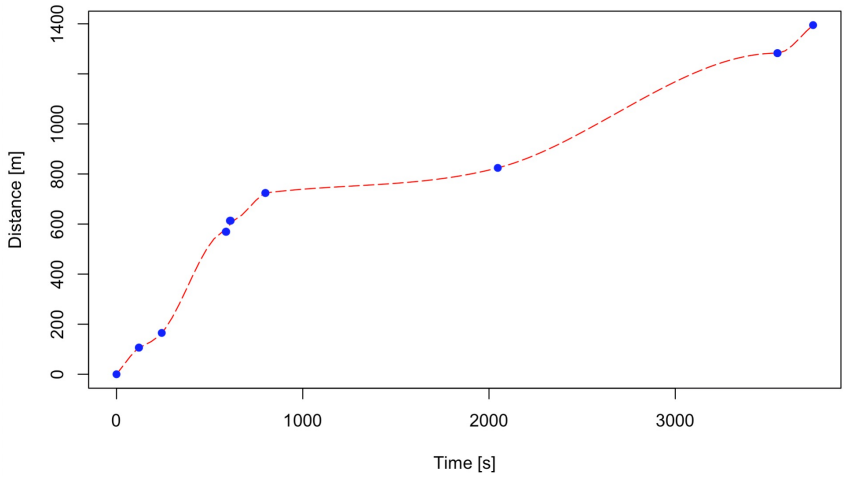
$$\Rightarrow \underline{D} := \underline{\overline{X}}_f - \underline{\overline{X}}_i \sim N_5(\underline{\delta}, \Sigma)$$

$$\underline{\delta} = \underline{\mu}_f - \underline{\mu}_i \qquad \Sigma = \frac{\Sigma_f}{n_f} + \frac{\Sigma_i}{n_i}$$

$$\Rightarrow CI_{1-\alpha}(\underline{a}'\underline{\delta}) = \left[\underline{a}'\underline{D} \pm \sqrt{\underline{a}' S_p \underline{a}}\ \ z_{1-\frac{\alpha}{2k}}\right]$$

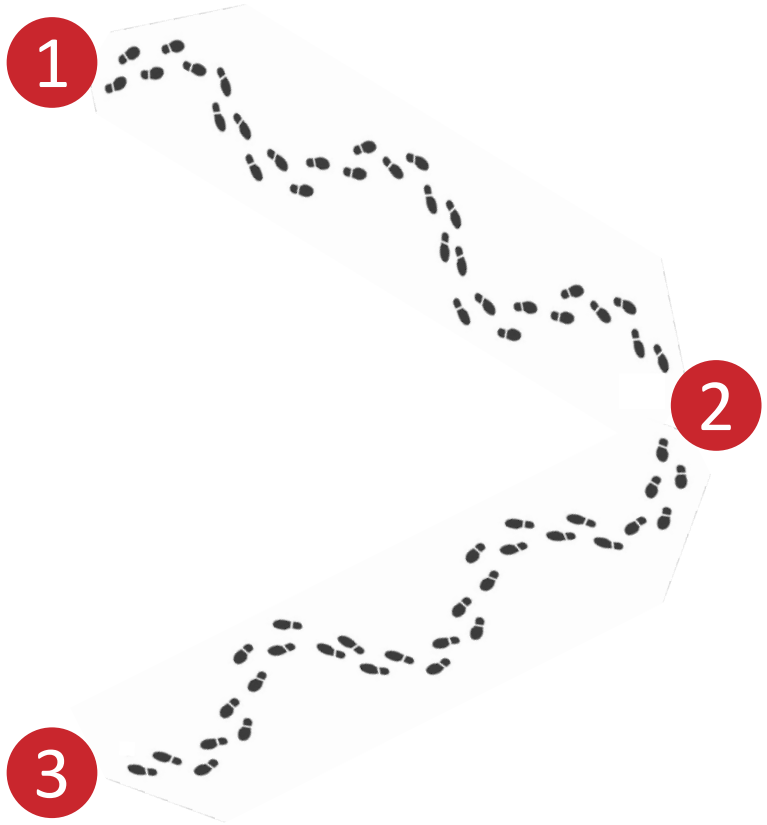
At level 95%, lockdown decreased the means within each of the considered groups. There seems to be a relation between changes in mean and zone: population, area

Transport Mode Detection



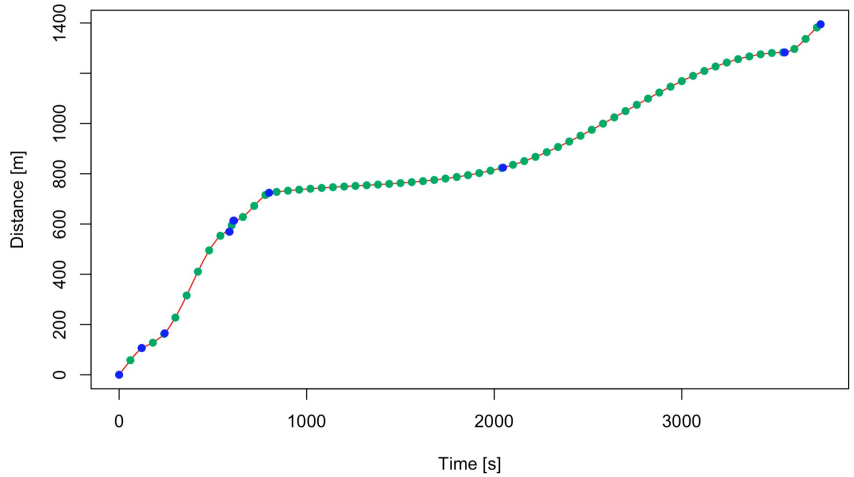
Interpolation

Interpolate the cumulative distance with respect to time of each trajectory with a monotone function



Features extraction

Exploit the refined information about each trajectory to extract further relevant quantities: percentiles for velocity and acceleration



Sampling

- Sample many time-distance couples from each monotone function
- Where necessary, split the trajectory in homogeneous subtravels

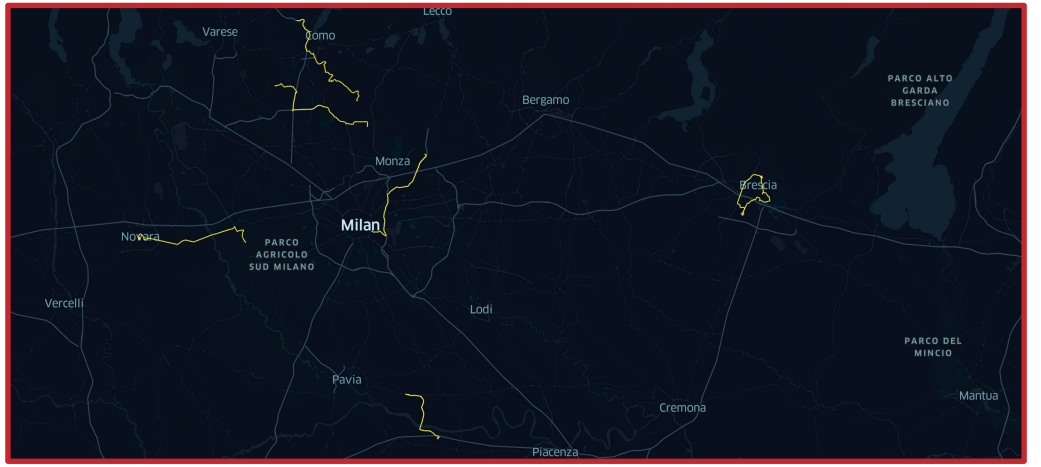
Clustering

Cluster 1

Short or slow travels, mostly done on foot

Cluster 2

Intermediate travels; they are very frequent outside towns



Cluster 3

Long and fast travels; many of them are between different towns or provinces

Some specifics

Data selection

To hold the most accurate trajectories, we performed a strong data filtering:

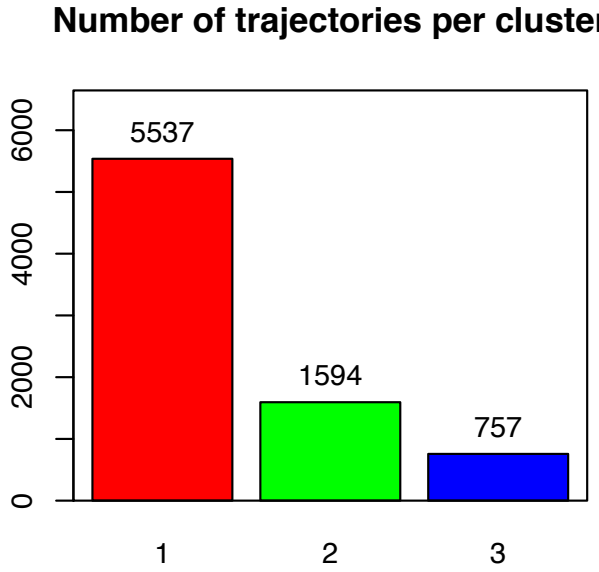
- $N_{det} > 15$
- $10min < T < 10h$
- $\overline{v} < 150 \frac{km}{h}$
- $v_{max} < 220 \frac{km}{h}$
- Only trajectories non containable in a 200m x 200m square

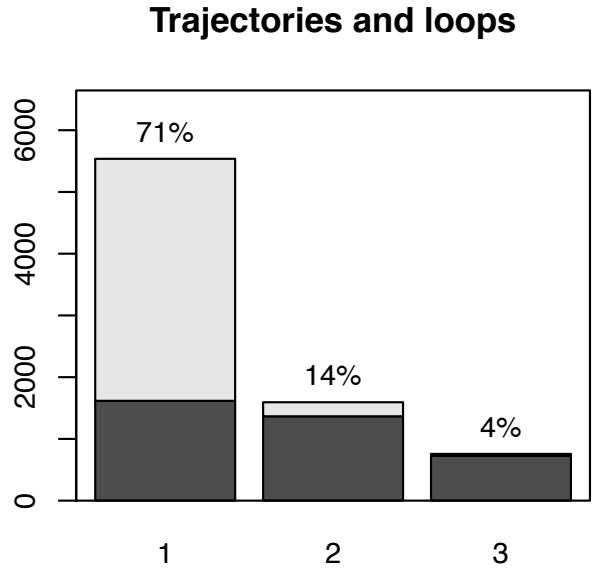
1'704'529 traj. → 7'888 traj.

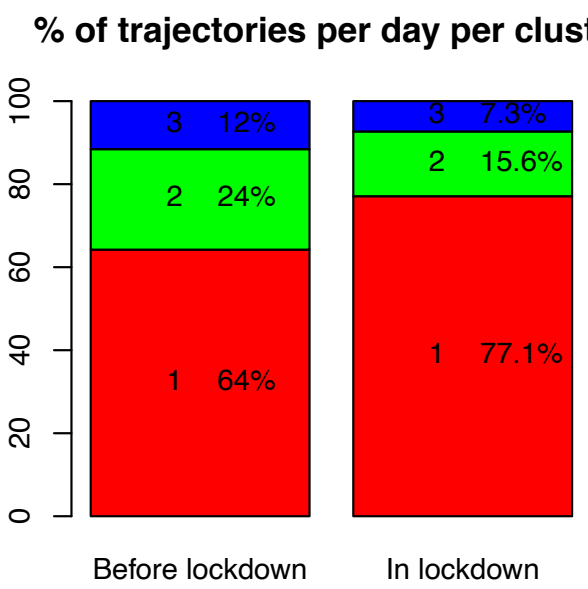
Relevent features for clustering (K-means)

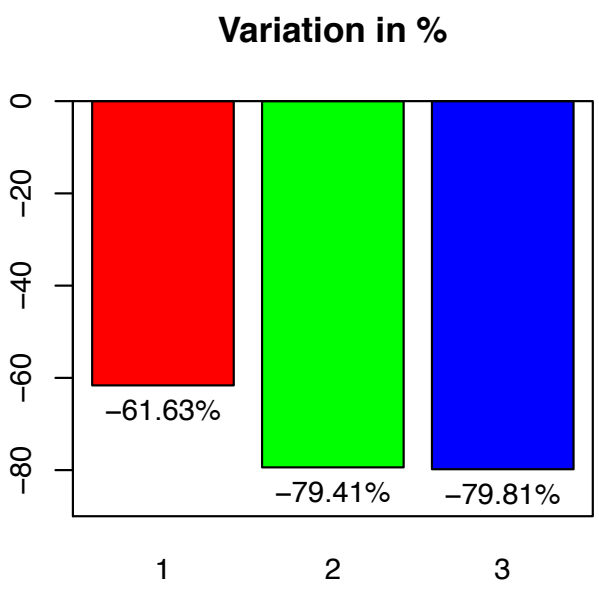
Results

- Cluster1 is dominated by loops: trajectories that start and end roughly in the same place. This is clear signal of walking



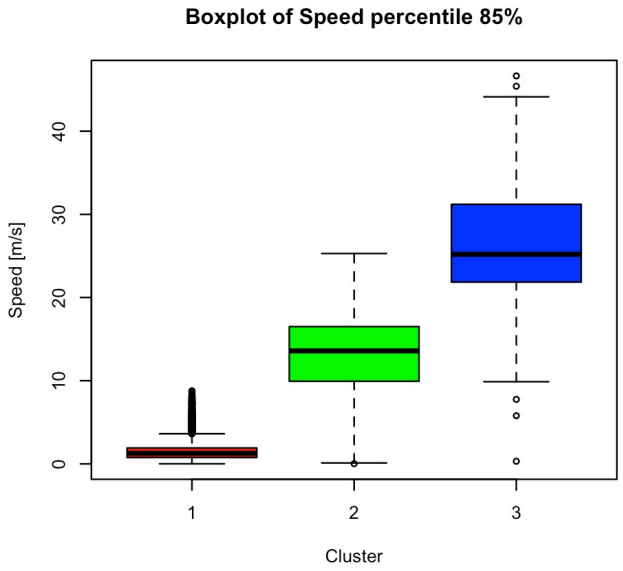


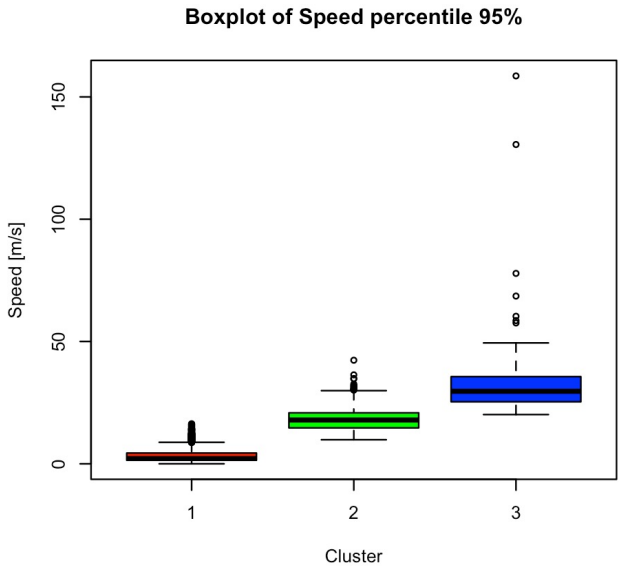


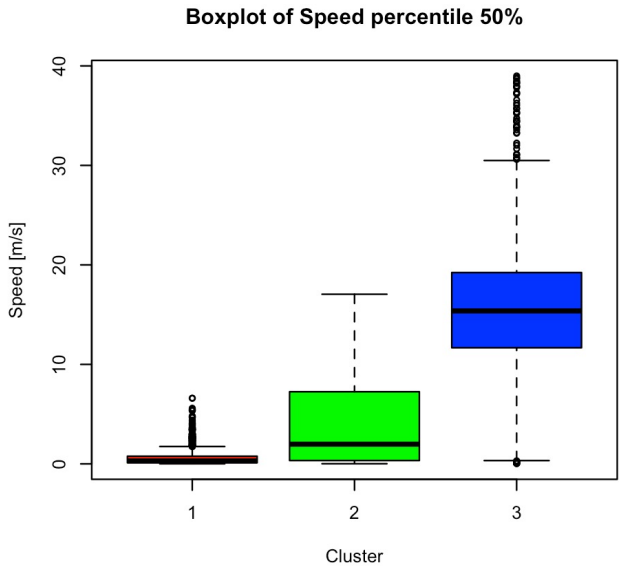


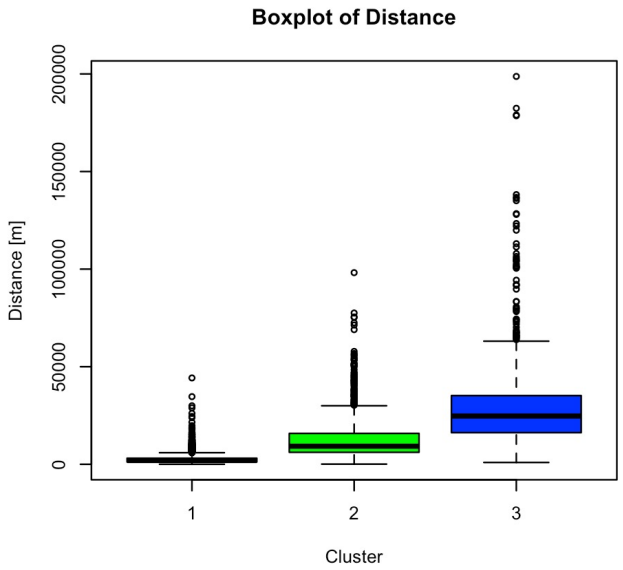
- All types of transport decrease during lockdown. However, the remaining trajectories change composition.
- Walking becomes more frequent during lockdown, while cars and long travels face a stronger decrease.

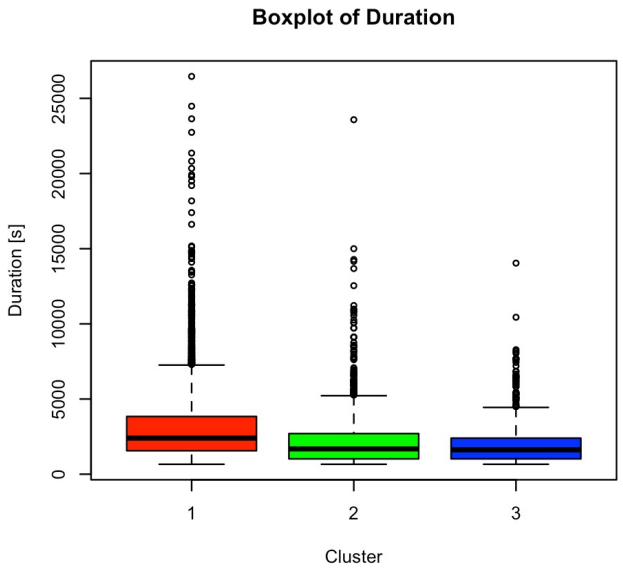
Clusters comparison

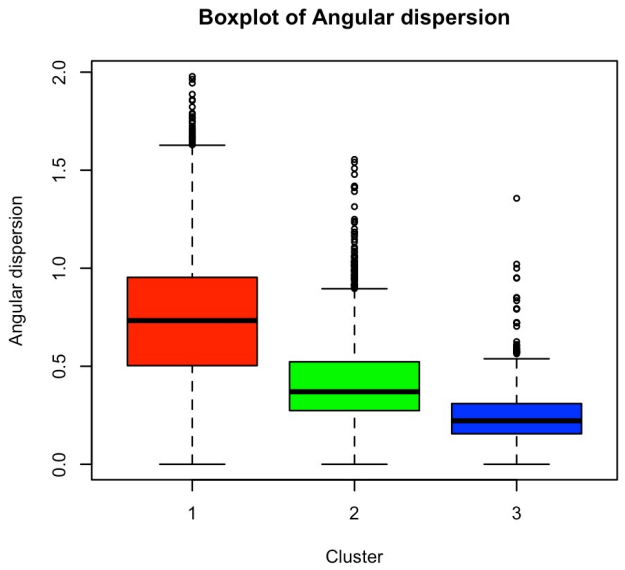






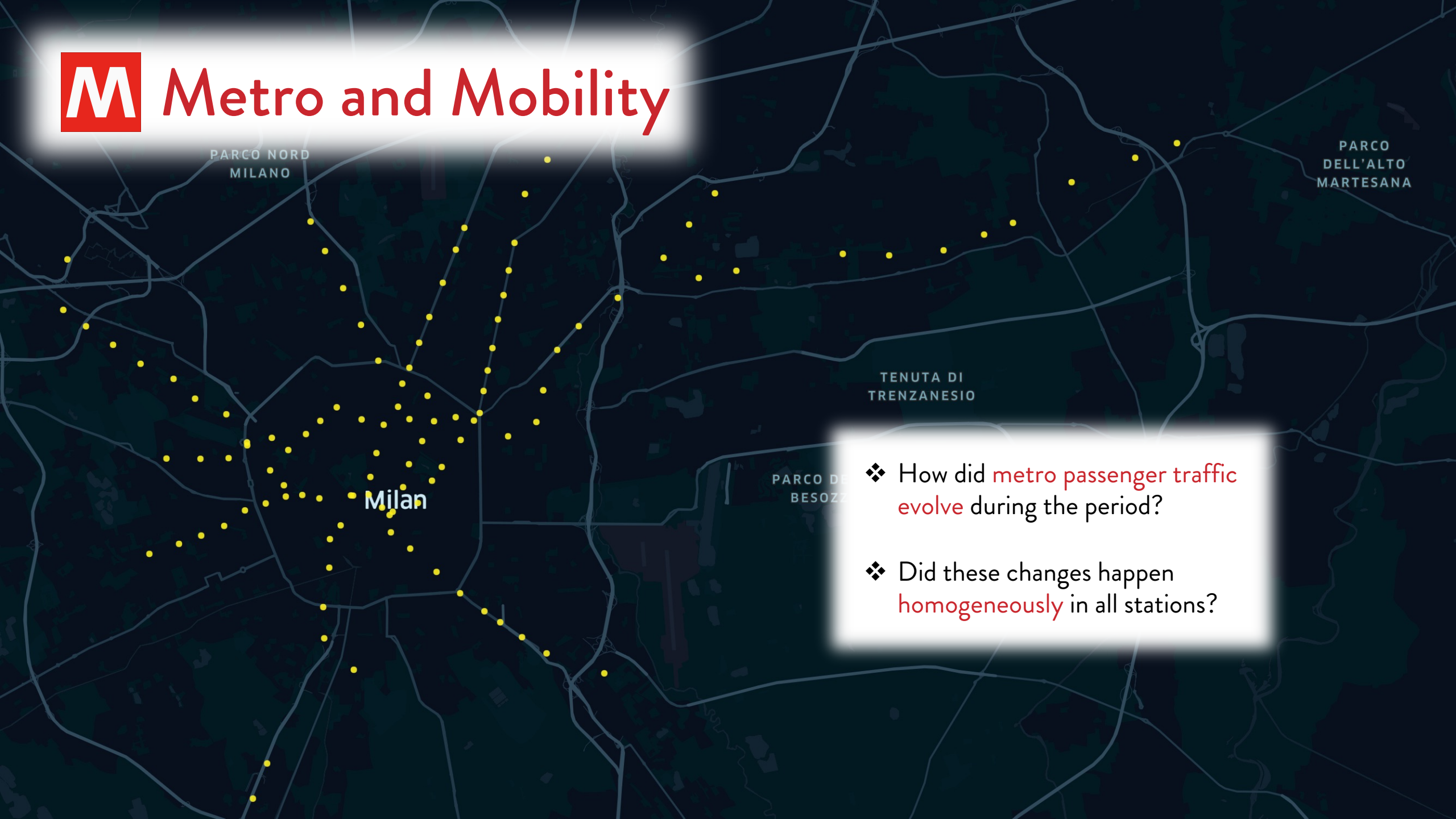






- Angular dispersion: mean of the angular variations between consecutive detections
- As a further confirmation of our hypotheses, angular dispersion tends to be lower as the cluster increases in velocity

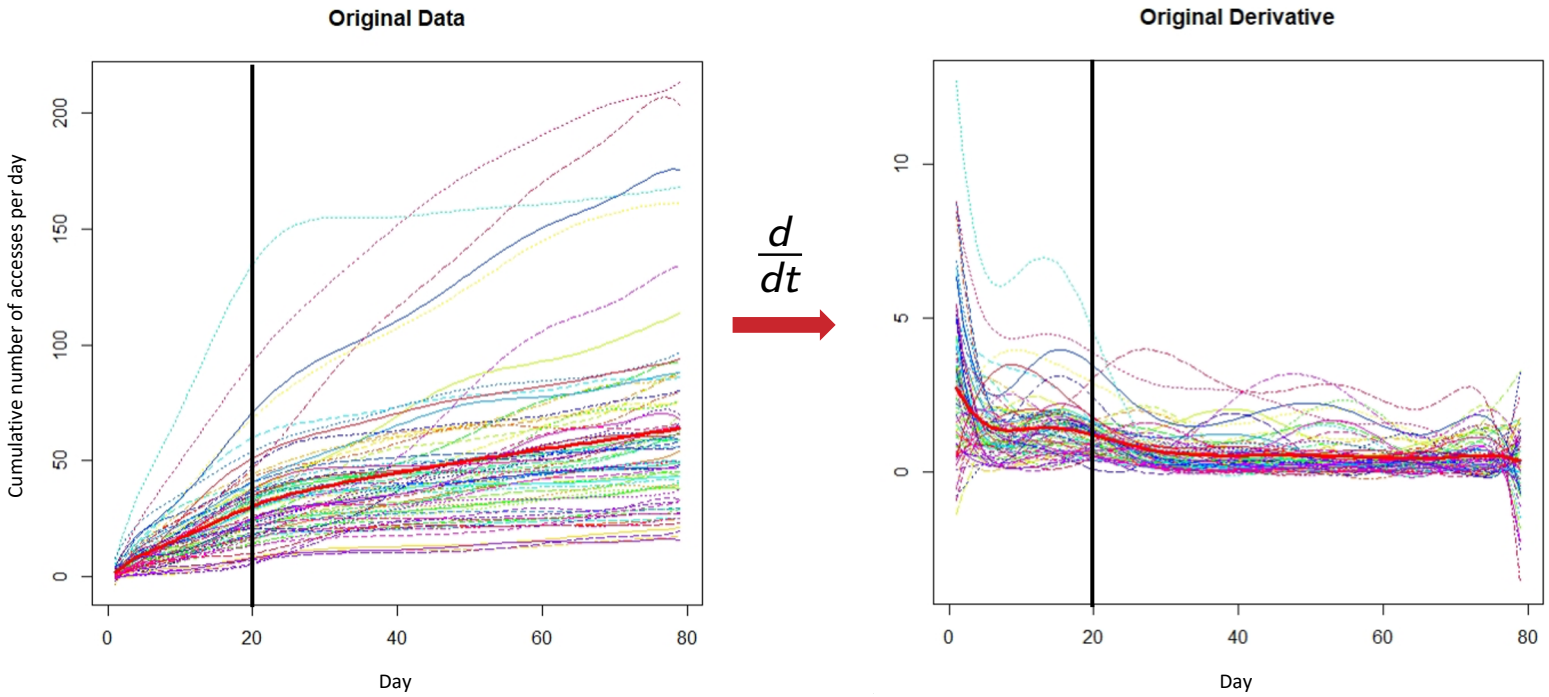
Metro and Mobility

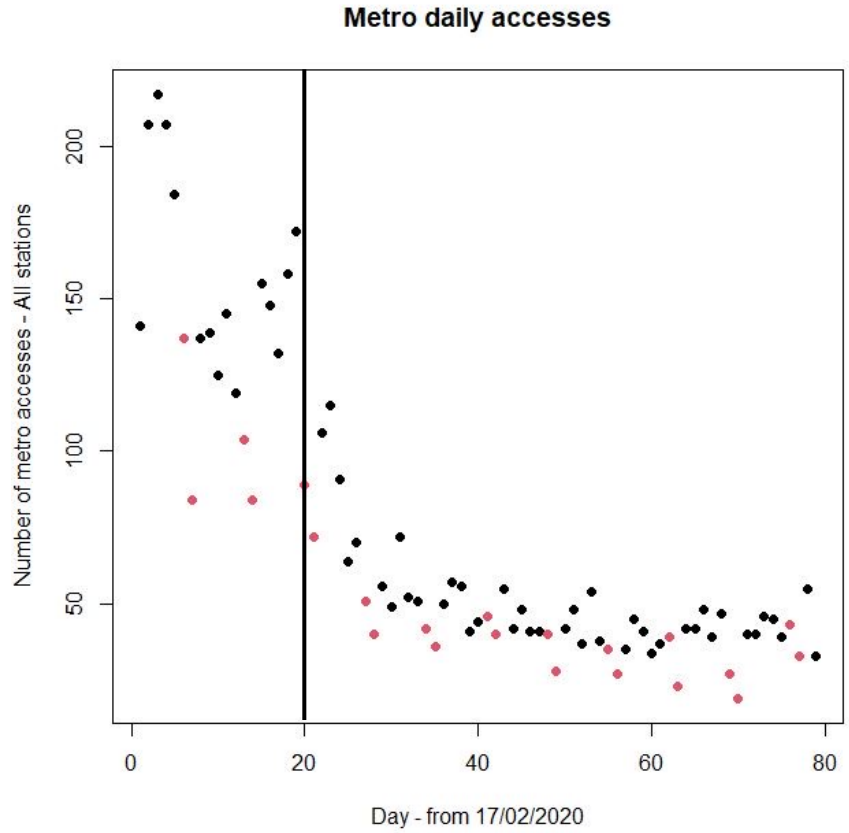


- How did metro passenger traffic evolve during the period?
- Did these changes happen homogeneously in all stations?

Smoothing of cumulative arrivals

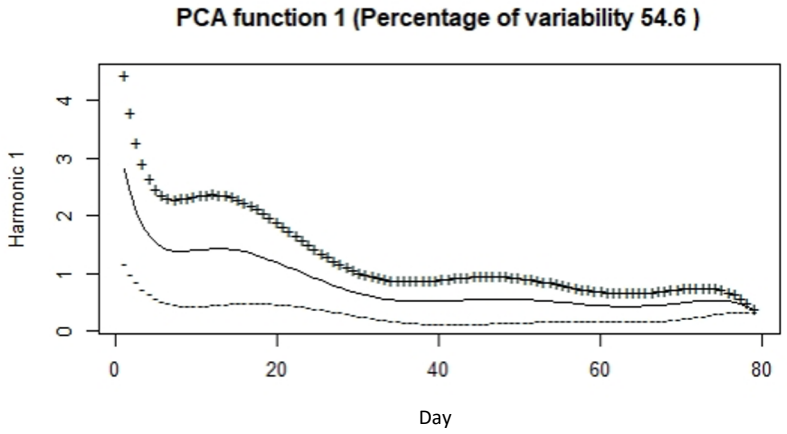
- Prominent decrease at day 20, corresponding to the beginning of lockdown
- Weekends (red points) always record a decrease: work commuting has a major impact on the number of passengers



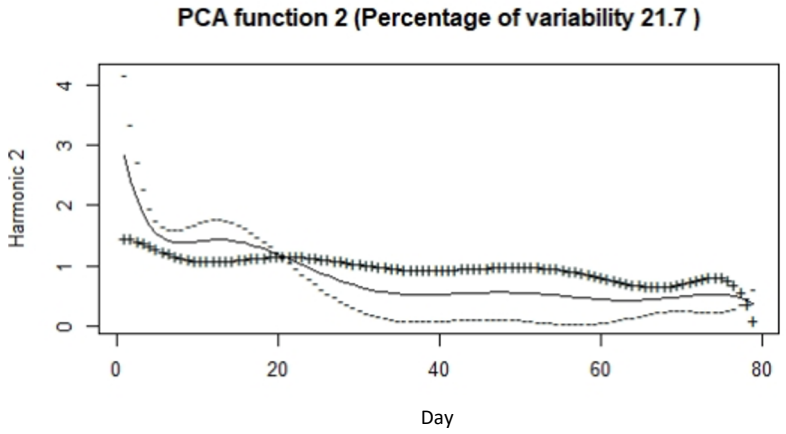


- We investigate differences in density of metro use between stations
- Focus on first derivative

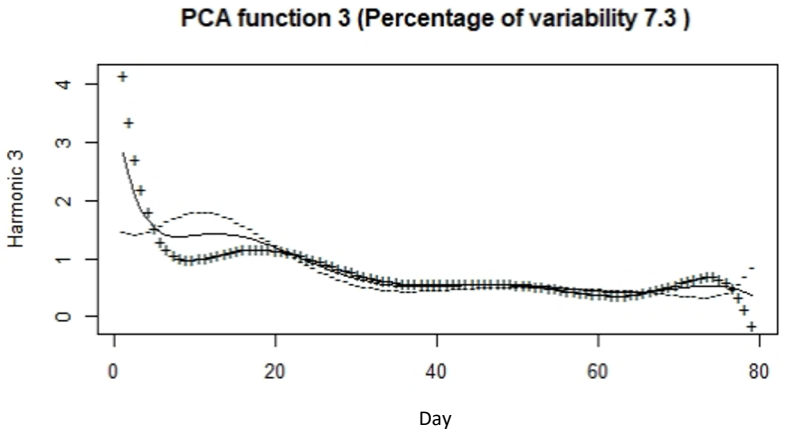
Functional PCA ⟶ Which are the main reasons of variability?



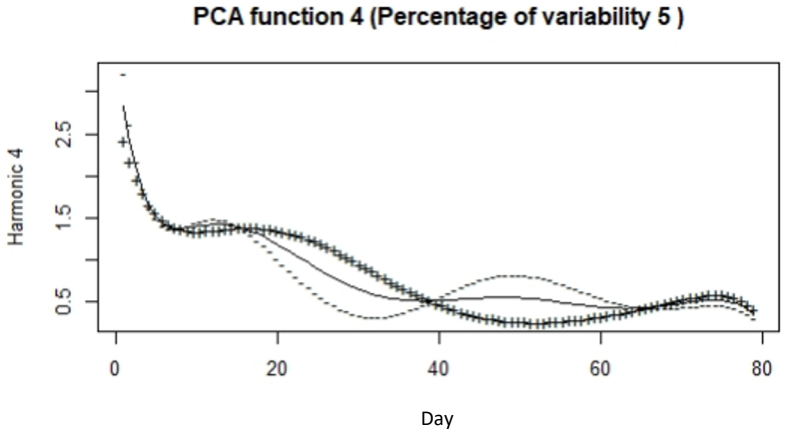
Overall number of passengers



Steeper decrease and lower level after entering lockdown VS milder variation



Different short term reactions to lockdown



Different clocks for variation

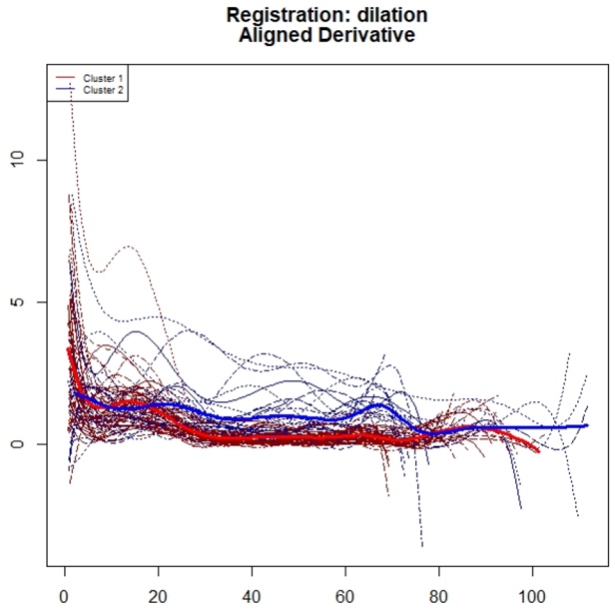
Functional K-means

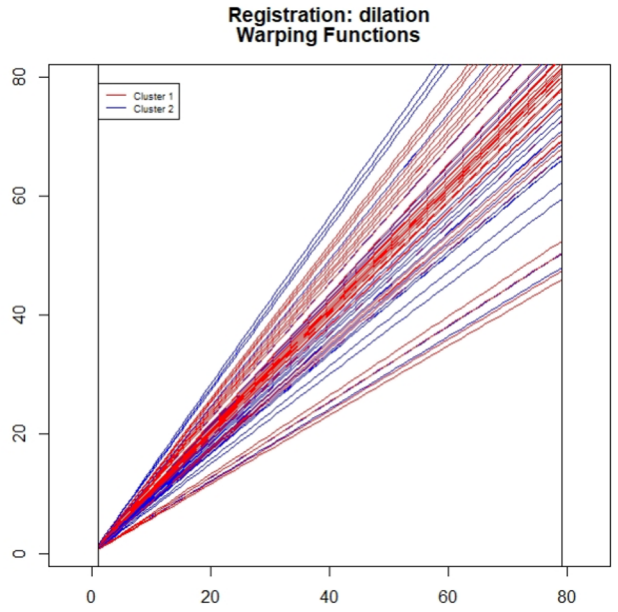
Can we identify groups of similar metro use profiles across stations?

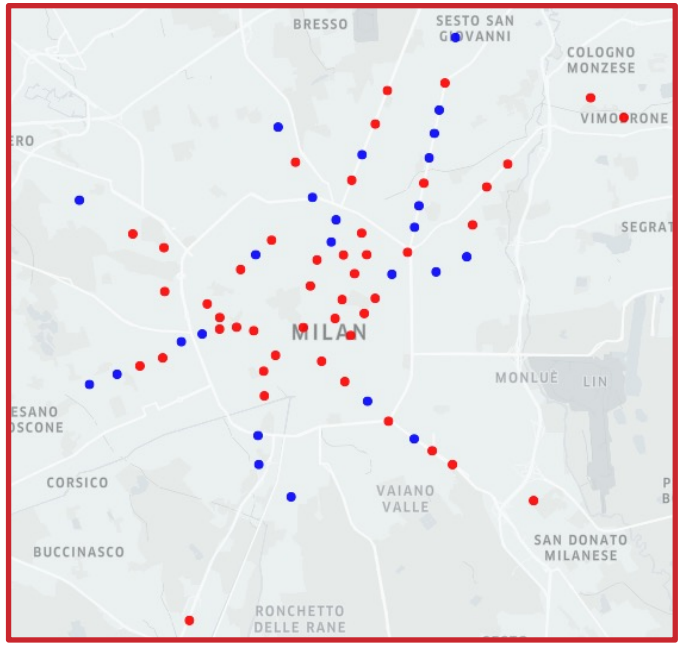
We perform K-means with 2 clusters, and dilation

- Heavy decrease, late recovery
- Less relevant decrease, earlier recovery

- Stations in the city center are definitely in the red group: offices can easily switch to remote working
- Outer stations are mixed: this may be due to the difference in residential or industrial area and income
- Blue stations probably feature smaller businesses, that returned to work earlier







References

- Safari Njema. https://www.safari-njema.polimi.it/
- Kepler. https://kepler.gl/
- Ugarov, Artem. 2019. *Automatic transport mode detection based on Global Positioning System data - Analysis of informal mobility in the city of Maputo*. Master thesis, Department of Mathematics, Politecnico di Milano.