

# Spatial patterns of reported crime in Italy before, during and after the COVID-19 pandemic

Michele Brunelli

February 8, 2026

## Abstract

Some text here later

## 1 Introduction

The analysis of the spatial distribution of crime represents an important topic in applied geography and spatial data science, as criminal phenomena are intrinsically linked to the territorial, social, and institutional contexts in which they occur. Crime, in fact, is not randomly distributed in space, but tends to exhibit patterns of concentration, persistence, and clustering, shaped by socio-economic factors, urban structure, and the opportunities offered by different territorial contexts.

The COVID-19 pandemic constitutes a major shock that has profoundly altered everyday life, social interactions, and economic activities in recent years. Lockdown measures, mobility restrictions, and changes in people's routines have likely affected both the overall volume of reported crimes and their spatial distribution. From a strictly geographical perspective, the pandemic can be interpreted as a natural experiment that allows observation of how crime patterns respond to sudden changes in human behaviour across space.

This project analyses the spatial patterns of reported crimes in Italy over the period 2014–2023, comparing three distinct phases: the pre-pandemic period (2014–2019), used as a baseline; the pandemic phase (2020–2021); and the recovery phase (2022–2023). The aim is to assess whether the observed changes in crime were homogeneous across the national territory or whether they assumed a spatially differentiated configuration, characterised by areas with specific local dynamics.

The analysis is conducted across three different geographical levels. A macro level, defined by five macro-areas dividing Italy into North-East, North-West, Centre, South, and Islands; a meso level, corresponding to the 20 Italian regions; and a micro level, at which the 107 provinces into which the country is divided are analysed. Given that Italy is characterised by territories with markedly different demographic sizes, both absolute numbers of reported crimes and population-normalised crime rates are considered.

From a methodological perspective, the project adopts a reproducible analysis pipeline implemented in Python, integrating official statistical data with geographic boundaries and interactive visualisation tools. The analysis relies on exploratory spatial analysis techniques, including thematic cartography (choropleth maps) and local indicators of spatial association (LISA), aimed at identifying spatial patterns and clusters. To allow users to explore the distribution of crime across space and time, a web-based application has also been developed.

The objective of this work is not to establish causal relationships between the COVID-19 pandemic and crime trends, but rather to provide an exploratory and spatially explicit analysis of the transformations observed during a period of profound social disruption. The project therefore seeks to highlight both persistent—and pre-existing—territorial inequalities and local variations associated with the pandemic period.

## 2 Data

This section describes the data used for the analysis presented above. Official statistics on reported crimes were obtained by querying the SDMX APIs exposed via web services by ISTAT. SDMX (Statistical Data and Metadata eXchange) is a standard widely adopted by statistical institutions to provide access to time series in a standardized and reproducible manner.

Data on crimes reported to the authorities and on crime rates are made available by ISTAT in several formats. For convenience, the data were downloaded in CSV format and subsequently aggregated—one dataset per year—into a single Parquet file. Parquet is a binary, columnar, and typed data format. This structure allows access to only the relevant columns without scanning the entire file each time, thereby minimizing input/output operations. Parquet is widely used in data analysis because it maximizes performance while minimizing both memory usage and storage requirements.

Although the CSV files are described as containing annual data, they include observations for both the reference year and the subsequent year. For this reason, a consistency check was performed on the aggregated Parquet file in order to avoid duplicated values.

For map creation, ISTAT datasets were merged with geographic data provided by Eurostat GISCO. Eurostat supplies geometries in GEOJSON format, organized by territorial levels. The levels used in this analysis are NUTS1 (macro-areas), NUTS2 (regions), and NUTS3 (provinces). ISTAT data rely on the 2006 version of the NUTS classification. As a result, it was necessary to manually map provinces that underwent administrative changes after 2006, most notably in Sardinia.

The data were aggregated into three time periods. The first corresponds to the pre-pandemic period (2014–2019) and reports the six-year average in order to obtain a robust baseline. The second period covers the years of acute pandemic impact (2020–2021) and reports their average. The third corresponds to the post-pandemic recovery phase and reports the average for the subsequent two years (2022–2023).

As mentioned above, the spatial analysis was conducted at three nested geographic levels. The macro level considers five areas of the Italian peninsula: North-East, North-West, Centre, South, and Islands. The meso level corresponds to regions, while the micro level corresponds to provinces.

Both the dataset containing the number of reported crimes and the dataset containing crime rates classify offenses into 55 categories. Although the analytical pipeline and the web-based platform process all crime categories provided by ISTAT, the results discussed in this report focus on a subset of categories that are most relevant to the research question, as shown in Table 1. Crime categories were grouped based on the predominant modality of interaction involved in the offense, distinguishing between crimes requiring direct physical contact (*contact crimes*) and crimes primarily mediated by digital infrastructures (*digital crimes*). Based on changes in daily routines induced by the pandemic, a decrease in contact crimes is expected during the second analyzed period, while an increase in digital crimes is anticipated.

The use of the global Moran's  $I$  index allows for the assessment of spatial autocorrelation in crime rates. In particular, it is expected that some crimes exhibit structured spatial patterns during the pre-pandemic period, whereas these patterns may weaken or change during the pandemic as a consequence of mobility restrictions. Local Indicators of Spatial Association (LISA) are used to identify local clusters and spatial outliers. The analysis aims to assess whether the pandemic altered the spatial distribution of crime clusters, especially for contact and digital crimes, highlighting potential processes of territorial concentration or dispersion.

Table 1: Crime categories discussed in the report (ISTAT codes)

<b>ISTAT code</b>	<b>Category (description)</b>	<b>Group</b>
PICKTHEF	Pickpocketing	Contact
BAGTHEF	Snatch theft	Contact
BURGTHEF	Residential burglary	Contact
STREETROB	Street robbery	Contact
RAPE	Sexual assault	Contact
SWINCYB	Online fraud and cyber scams	Digital
CYBERCRIM	Cybercrime	Digital
PORNO	Child sexual abuse material (CSAM) offences	Digital

### **3 Methodology**

### **4 Results**

### **5 Conclusions**

### **References**