Who lives where?

Jacques Lévy, Jean Coldefy, Sébastien Piantoni, Julien François

Science Advances
AAAS

Manuscript Template

FRONT MATTER

Title

Who lives where?

Counting, Locating, and Observing France's Real Inhabitants

Authors

Jacques Lévy*[1], Jean Coldefy[2], Sébastien Piantoni[3], Julien François[4]

Affiliations

[1] Jacques Lévy* is a professor (H), Ecole Polytechnique fédérale de Lausanne (EPFL), a member of the Chôros research network, and director of the Spatial Intelligence Chair, Polytechnic University Hauts-de-France (UPHF).
https://orcid.org/0000-0003-3182-3731
Jacques.Levy@choros.place

[2] Jean Coldefy is an advisor to the CEO of Transdev, and director of the Mobility and Transitions Program, ATEC ITS France.

[3]Sébastien Piantoni is an engineer in cartography and geomatics at Rheims University (Habiter-URCA) and a researcher at the Spatial Intelligence chair at Polytechnic University Hauts-de-France (UPHF).

[4]Julien François is an engineer in geomatics and data-visualization at Transdev.

Abstract

This contribution proposes answers to the question: who is where? For the first time, massive, reliable phone data collected and verified at a high level of spatial and temporal precision over the entire French territory have been used to address this questioning. The article raises the issue of data quality and relevance and sets out a framework which focuses on the inhabitant, not just the resident or the commuter. The processing of these data allows us to challenge the approximation of a home-stuck dweller and to propose a new approach. Two novel indicators (inhabitants.year and attractivity index) have been designed to account for the reality of individual spatial practices and to encompass the multiple ways in which a country is inhabited. Finally, maps that these data make possible are produced and analyzed. A fresh portrait of France as a complexly inhabited space emerges.

Teaser

Comprehensive data now makes it possible to locate and count the effective population with a high space and time accuracy.

MAIN TEXT

Introduction

It is not easy to know how many people there are in a locality, a district, a city. We have on the one hand censuses, where they are reliable, and on the other hand mobility surveys. The digital world provides new sources, but they remain fragmentary on this point. The main weakness of censuses is that they are based on an implicit assumption: that the location of a respondent's primary residence is predictive of that person's permanent location. This is not the case. Individuals move, sometimes coming from very far away, circulating in many places during the day, and may well still be spend the night elsewhere.

Mobility surveys help enrich the picture, but they too have several weaknesses (see Materials and Methods). This article presents the first set of results that originate from a different starting point: Individuals divide their time between

different places. They allocate their time resources to their various housings, places of work, education, personal relationships, or leisure, as well as to their travel activities. On average, each person spends time in several places each day, generally between 2.5 and 4 per person per day in urbanized societies: about France, see (1). For example, in addition to commuting (2), 40% of working people travel for work in France. An inhabitant is not only a resident.

It is therefore useful to distinguish *residents*, defined by the location of their home, from *inhabitants*, who are defined by the different types of places they practice, their primary home being only one of them. To avoid ambiguities, we will not use the term 'dweller' in the remainder of this text.

In this perspective, the article presents the results of a research that exploited massive cell phone data covering the whole of mainland France for one year. We have established indicators that allow us to answer the two questions: Where are the inhabitants really located? And how does their location differ from the classic population distribution based on permanent homes. This is what the first section of this paper is about. Maps draw a fresh, sometimes spectacular, geography of France.

For personal data protection reasons, the team did not have access to mobile phone users' personal identifiers. We have nevertheless been able to answer one part of the question: who are the inhabitants? In the second part, a self-portrait of the country through their inhabitants' spatial practices shows how lifestyle choices, tourism, or the role of foreign visitors contribute to define specific profiles for the different places, that constitute France, inside and outside its territory.

This first set of results marks the beginning of a vast work program. It is also an invitation to redefine and enrich what can be called *inhabiting*.

Results

Real People in Actual Places

A simple, versatile indicator: the inhabitant.year

Inhabiting is not only sleeping. As censuses suppose people remaining permanently still at their residential address, we more realistically consider that each individual allocates their time in different places.

A inhabitant.year (IY) is a full year stay equivalent. The value of IY_i is [1]:

$$IYi = \sum_{i=1}^{n} \frac{S30i}{17520} \quad [1]$$

where S30*i* is the number of people that stayed for any of the forty-eight 30-minute time slots of a day of a determined year in the *i* spatial unit and 17520 being the number of 30-min slots in a year (see Materials and Methods, below).

This indicator takes into consideration the fact that, most of the time, an individual does not stay in the same place for the whole year, and what is useful to know is how much of his year he/she has spent in a particular place. This measure of *full-year equivalents* also has the advantage of producing efficient results with sources that are not only anonymous but also unidentified. We know that in each location, in a given timeslot, there were *x* people, but this does not tell us anything about these people. We can, however, measure the number of people present in that location and this is a major information.

Using 30 min-stays data all over a year (March 2022-February 2023), we have calculated the number of IY in a determined space in each of the 48,589 spatial statistical units that compose the mainland France territory (see Materials and Methods). The results obtained show a total population of 70,098,618 inhabitants.year divided into 65,121,201 subscribers of national operators (French permanent residents) and 4,977,417 roaming customers (foreign visitors).

Fig. 1-2 shows the spatial distribution of IYs' density. This map appears to be close to a typical density map made from census data. In fact, important differences concern the masses and their distribution.

The Euclidean map (Fig. 1) particularly emphasizes transportation routes, which makes sense because the measurement detects moving people in proportion to the time spent in a small spatial unit. These mobile inhabitants who are missed in the usual census counts now affect the density map.

The cartogram (Fig. 2) has a background map making the areas represented proportional to the number of IYs. This shows the considerable weight of the largest urban areas. Their density reaches particularly high values in the city centers. The maximum density reaches 490,676 IY/sq. km (≈189,400 IY/sq. mi.) in central Paris, where the average residential density is 20,360 people/sq. km (27,346 in Manhattan). These scores are all the more impressive because they do not reflect seasonal overrepresentation: they are full-year equivalents.

The non-metropolitan territory appears demographically residual, which is confirmed by comparison with census data (see below).

Fig. 1. **France,** Density, I**nhabitants.year, 2022-23, Euclidean Metric.**

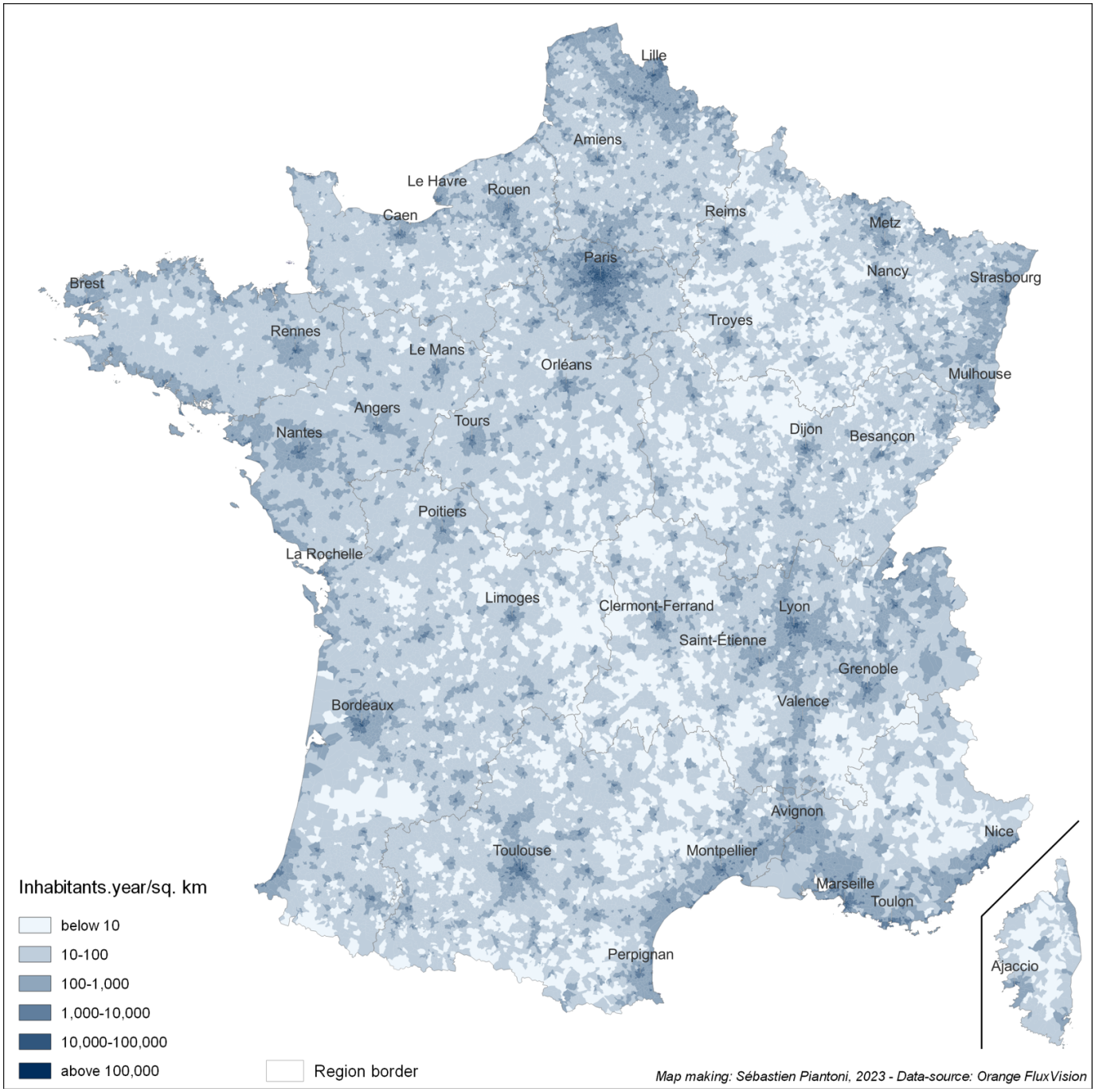
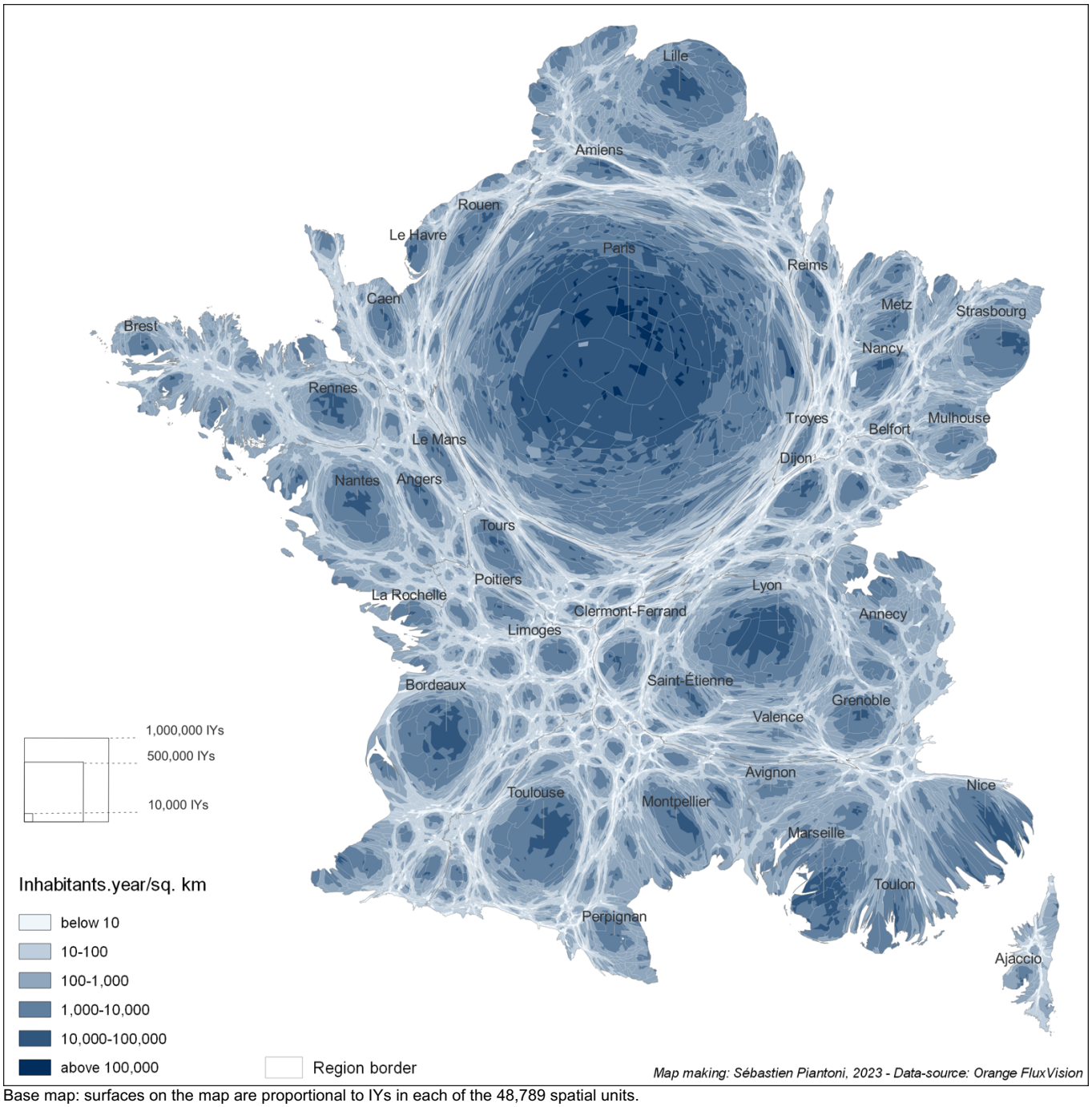


Fig. 2. **France, Density, Inhabitants.year, France, 2022-23, Cartogram**



Base map: surfaces on the map are proportional to IYs in each of the 48,789 spatial units.

Inhabitants vs. Residents: An Index of Attractiveness (A-index)

The comparison between the number of people counted at their primary residence (PR) of a spatial unit as measured by censuses and their actual inhabitants (IY) can be considered a simple and powerful indicator of the attractiveness of a place (see Material and Methods. The A-index [2] is a simple relation between an overall population and its residential component.

[2]

$$\mathbf{A}i = 100\left(\frac{\mathrm{IY}i}{\mathrm{PR}i}\right)$$

The A-index appears as a useful complement to the classic measurement of different types of mobility, on the one hand, and to economic attractiveness (e.g., with jobs or investments), on the other.

It includes those flows of people, goods, or capital but goes beyond it, offering a vision that is both synthetic and comprehensive of all the logics that link one place with all the others. Moreover, the A-index makes sense at all scales, from micro-local to global.

Fig. 3. **France, A-index, Euclidean Metric.**

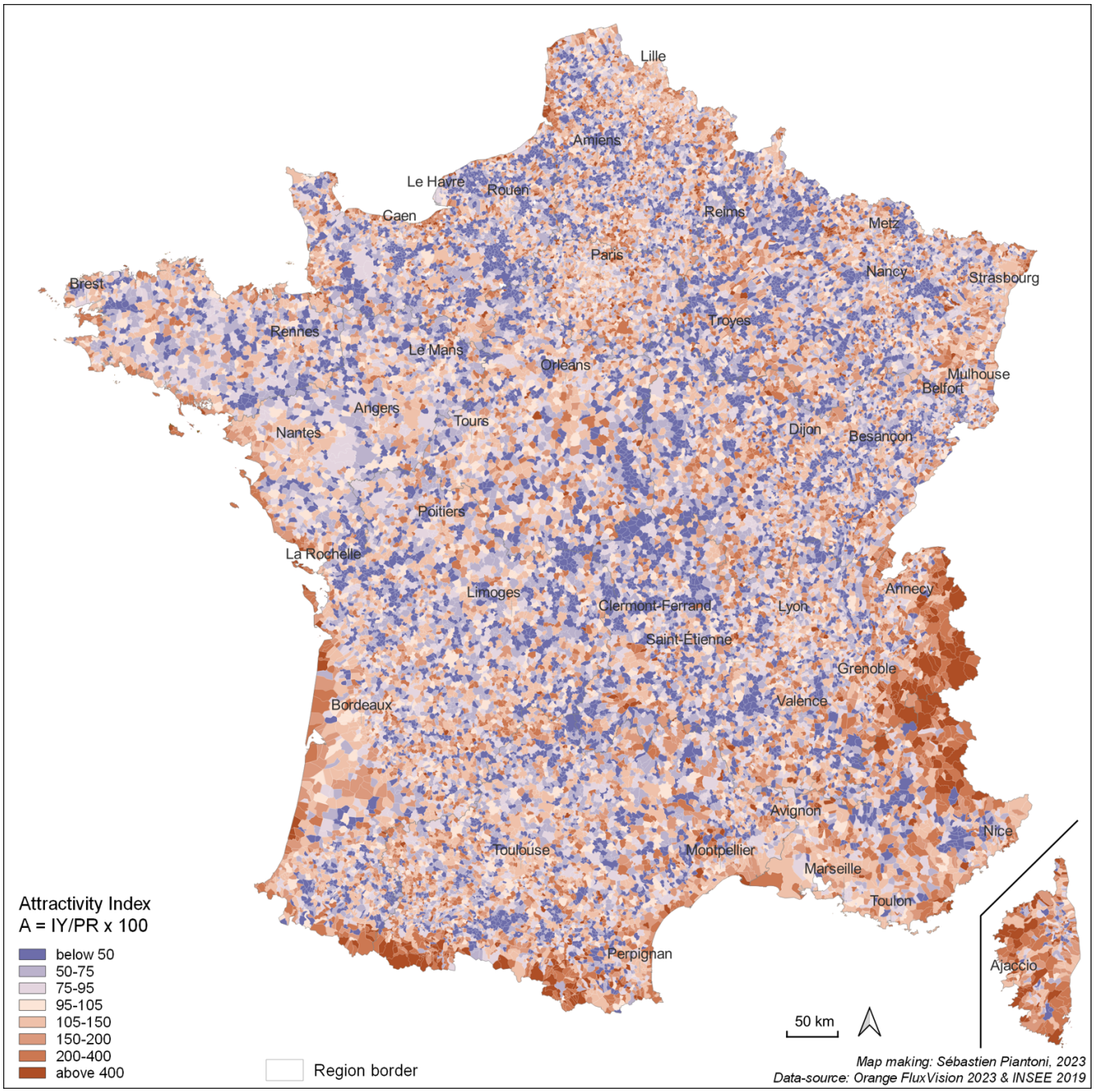
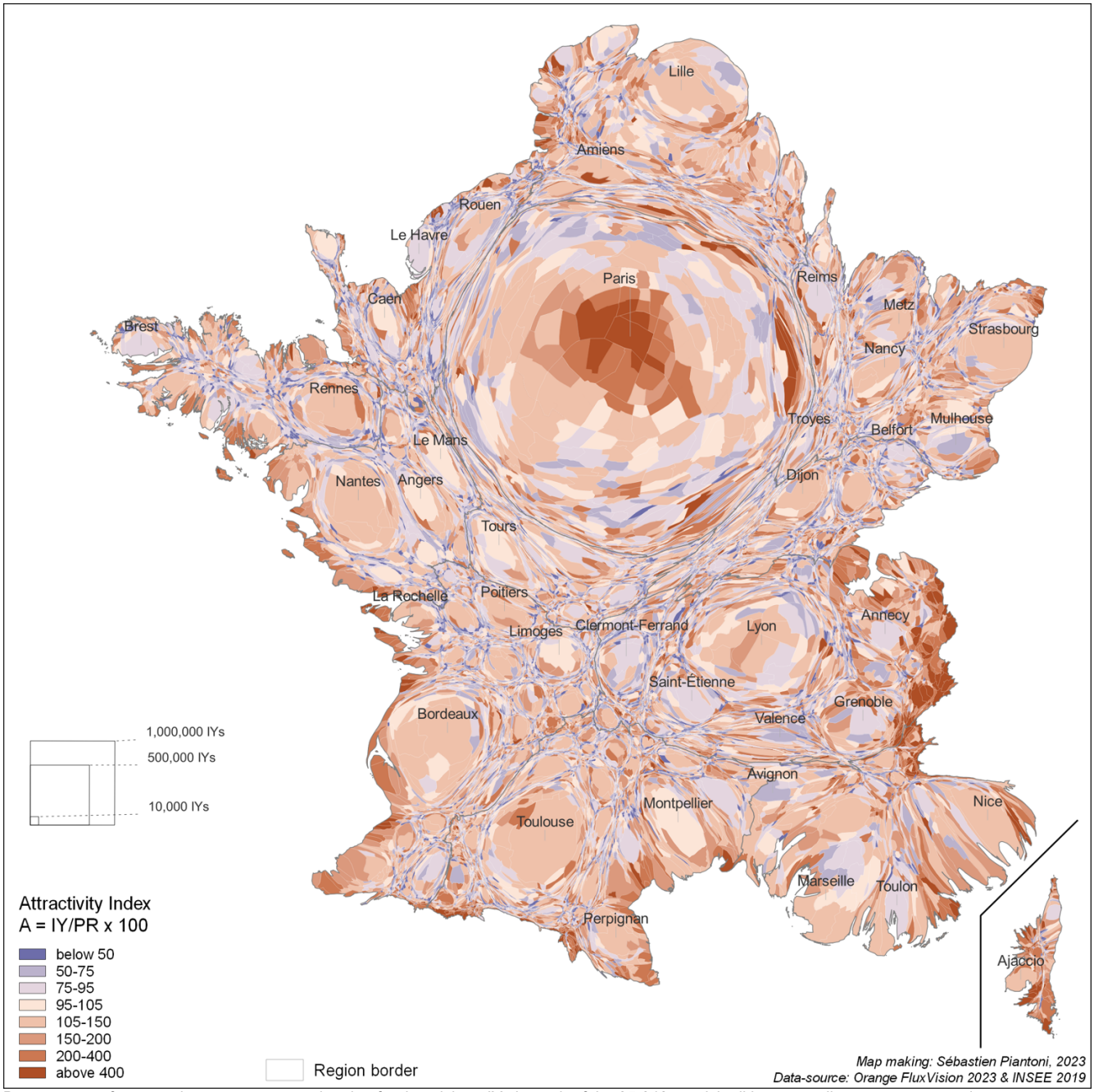


Fig. 4. **France, A-index, Cartogram**.



Base map: surfaces on the map are proportional to foreign visitors IYs in each of the 35,048 municipalities, *arrondissements,* or *quartiers*.

Two comparative maps (Euclidian and cartogram, Fig. 3-4) show the spatial significance of this index. The most populated locations show values significantly above the 100-line. Beyond the density values in a given spatial unit, the absolute value comparison between PR and IY show huge differences as the most populated areas have very few residents, sometimes none.

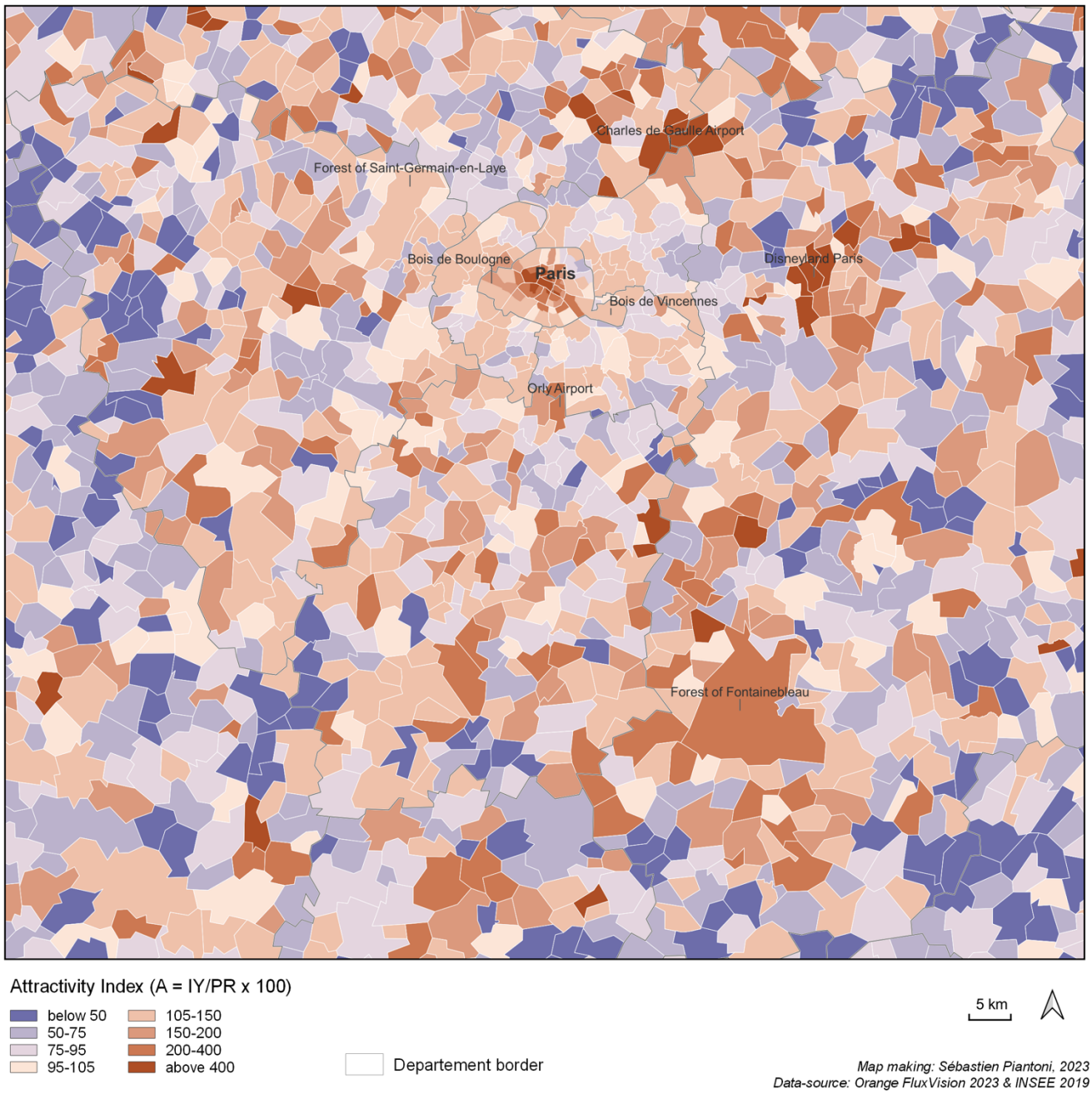
The contrasts between the two series of data (inhabitants vs. residents) at the national level reveals two major features: the polarization on the core areas of urban regions, especially as these are important, and an internal redistribution inside urban areas (see below). We will analyze later a third aspect: the considerable role of tourists-appealing areas.

First, the IY distribution map shows strong regularity in support of a concentric pattern. The core areas appear hyper-populated with values that no density measure had previously identified. The inner suburbs are unevenly populated, but overall, non-residential activities of various types do balance residential mass in many locations.

Conversely, predominantly residential suburban and peri-urban areas show a lower population than in the censuses: their residents spend a significant part of their time outside of their home, while their neighborhoods are unattractive to other inhabitants.

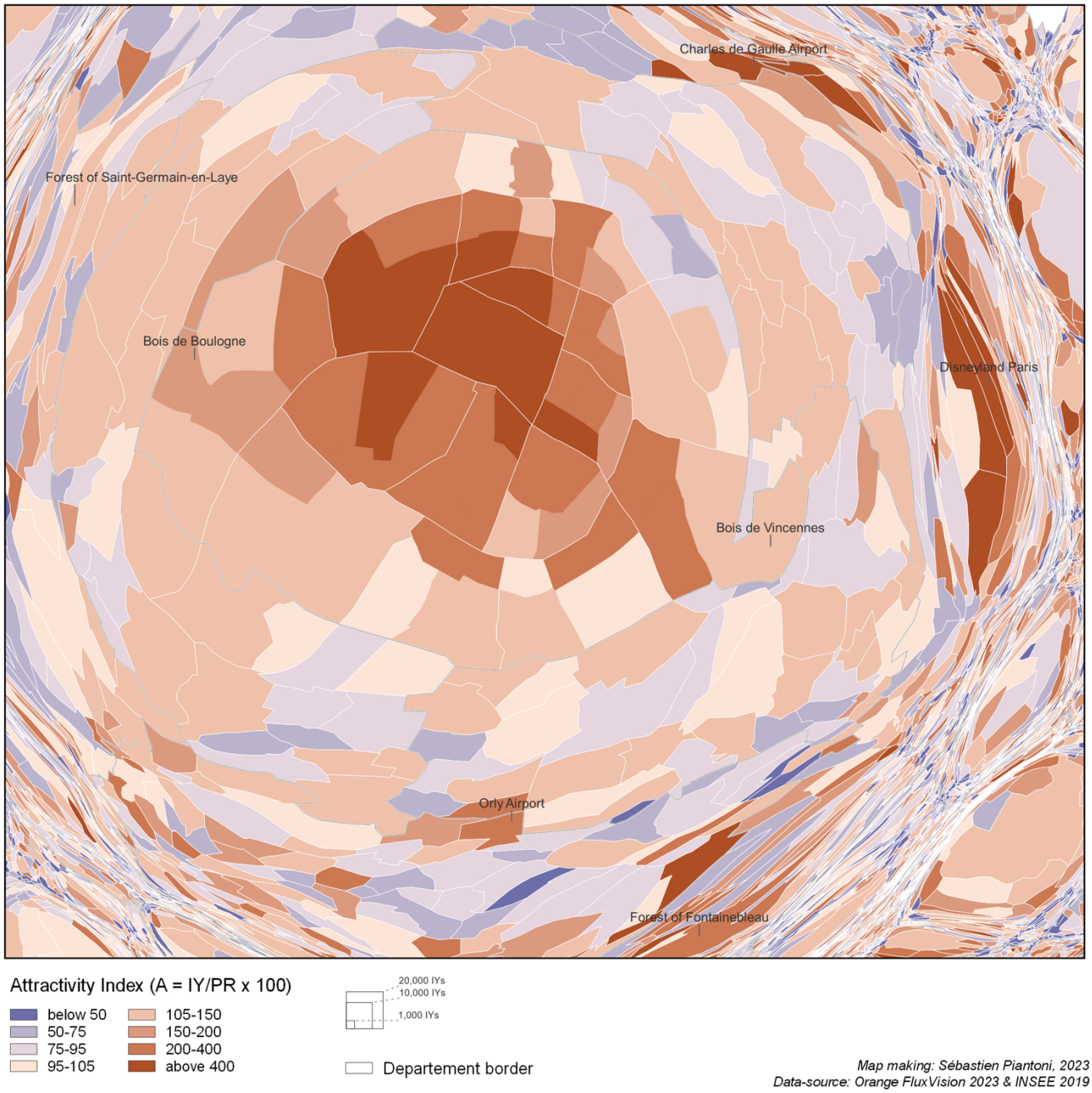
In large urban areas, the differentiation of spaces according to the two measurement systems (residents and IYs) generates a mosaic-style pattern. In each urban area, places with a IY surplus are next, often very close to locations with a deficit compared to census population. The places that have the most visible positive contrast are the employment areas in the historic centers, commercial or cultural attractions, mobility hubs and axes, and recreational areas (urban parks, forests, and theme parks).
The zoom-in maps (Figures 5-6) show the situation in Paris urban area and in Lyon region.

Fig. 5. **Paris Metropolitan Area, A-index, Euclidean Metric.**



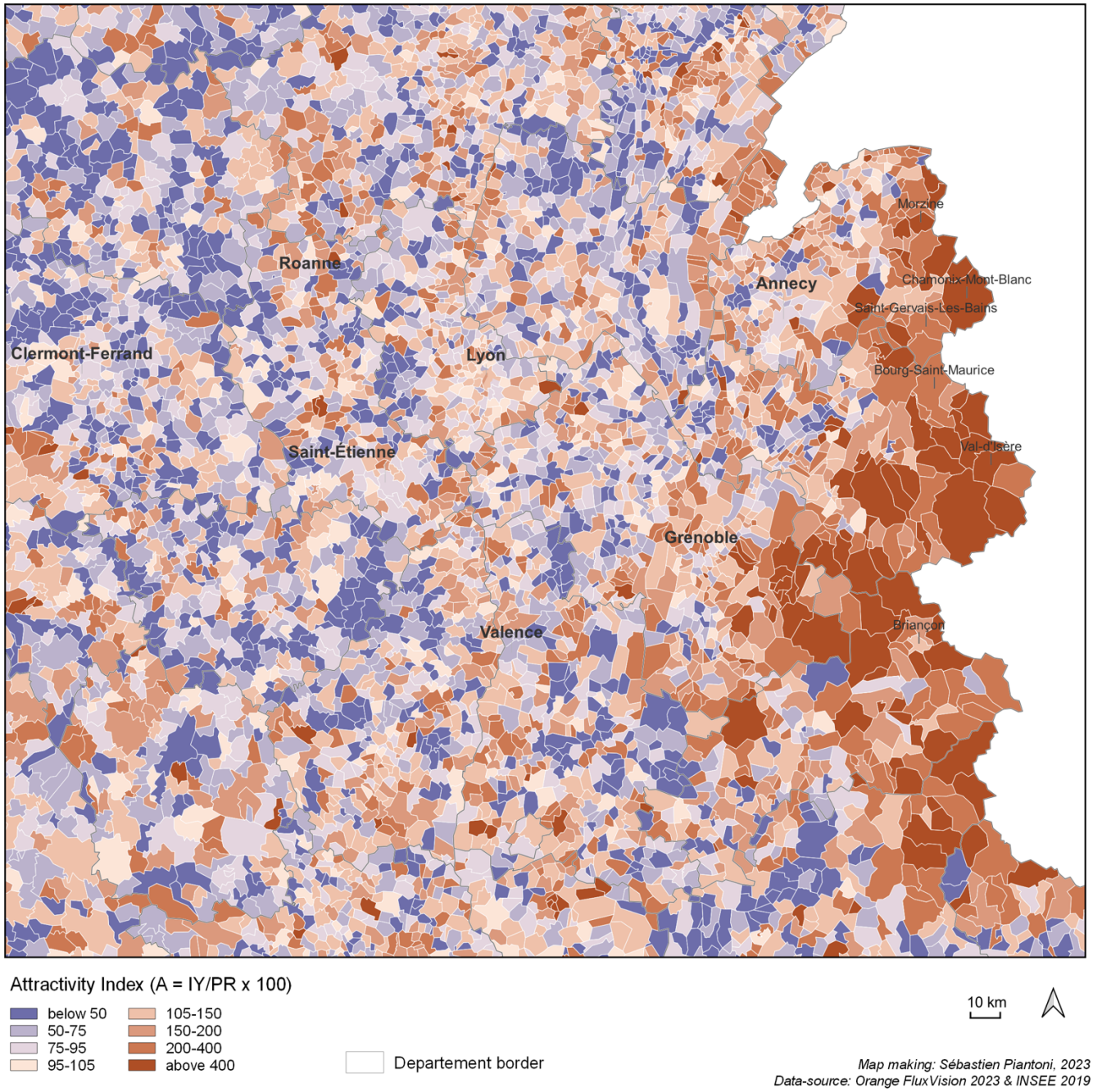
In the Paris metropolitan area, roughly the Île-de-France region (Fig. 5-6), the Charles-de-Gaulle and Orly airports or the Disneyland Paris resort as well as public forest (Fontainebleau, Saint-Germain-en-Laye) stand out in particular.

Fig. 6. **Paris Metropolitan Area, A-index, Cartogram.**



Base map: surfaces on the map are proportional to foreign visitors IYs in each of the 35,048 municipalities, *arrondissements,* or *quartiers*.

Fig. 7. **Lyon Region, A-index, Euclidean Metric.**

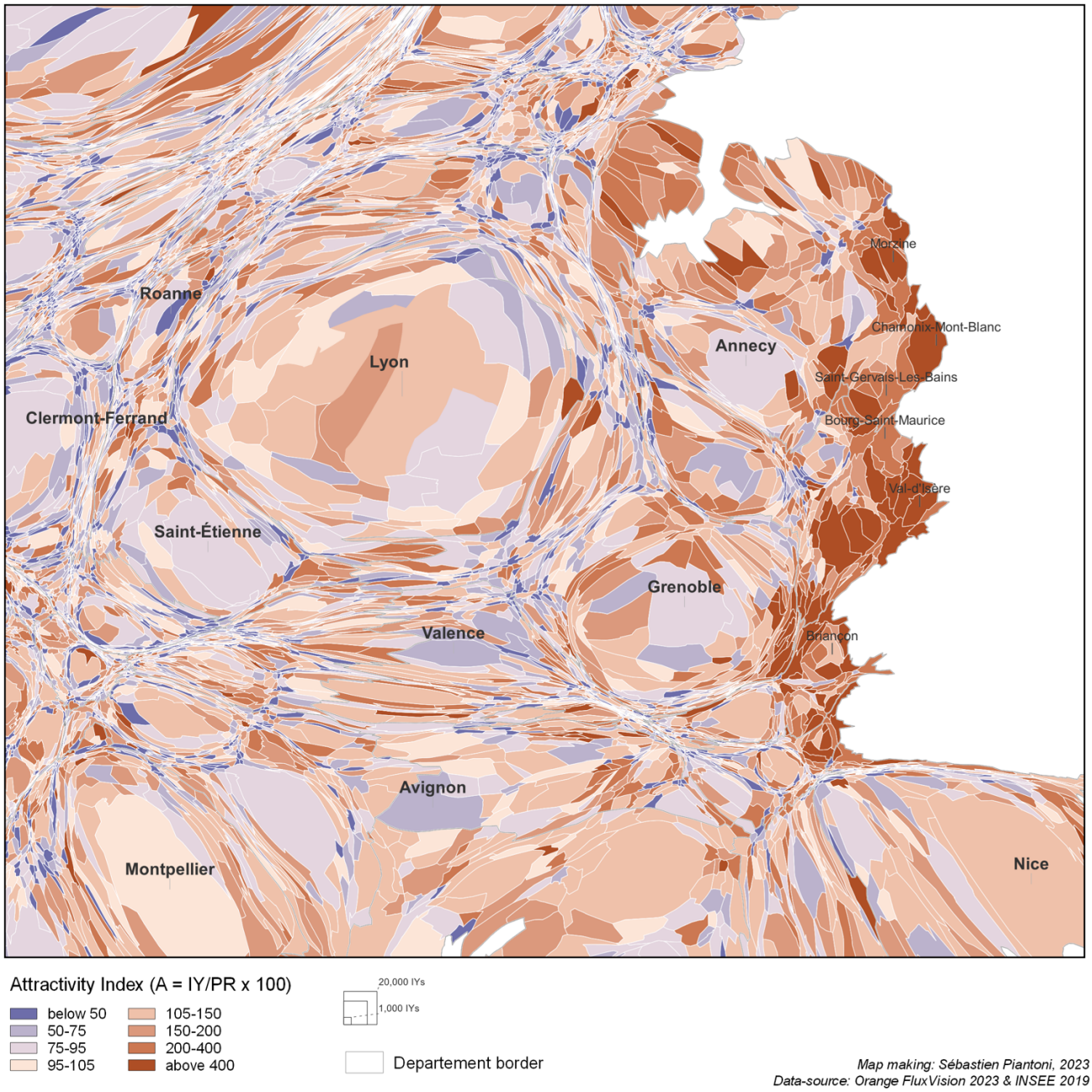


In the Lyon region, Auvergne-Rhône-Alpes (Fig 7-8), the weight of the Lyon core urban area is less marked, competing with suburban industrial districts, business parks, shopping malls, or leisure areas. This phenomenon is also visible in Grenoble, Clermont-Ferrand, and Annecy, and even more in Saint-Étienne a 'rust belt' city that is struggling to bounce back in a post-industrial context. The blue dominant of the residential suburbs, exurbs, and remote countryside appear as a 'background noise', while, on the other hand, the red zone of Alpine resorts forms a spectacular cluster.

The case of Roanne helps us understand how these new data make more visible the strengths and weaknesses of a given place. This city has an urban area of 141,000 residents (2020 census), but, in the same territory, to only 134,000 IYs. This means that it is slightly more repulsive than attractive. Simultaneously, the core municipality contains only 15,000 IYs, or only 44% of its 34,000 census-counted residents. On the other hand, the peri-urban area shows a higher

population (98,000) of inhabitants.years than of residents (81,000). This reflects the fact that employment and commercial activities tend to flee the historical center and settle outside. This is the typical situation in which the stagnation of the overall urban area is correlated with a loss of substance in its central area, where non-residential functions are tendentially evaporating. By crossing our two indicators (IY and A-index), we understand that it is not a simple coincidence: the single cause is a declining attractiveness that affects, in two ways, the whole area and its central district. As we will see later ('Large Cities' Cores Upgraded'), this is a problem common to most small and medium-sized French cities.

Fig. 8. **Lyon Region, A-index, Cartogram.**



Base map: surfaces on the map are proportional to foreign visitors IYs in each of the 35,048 municipalities, *arrondissements,* or *quartiers*.

A Country's Self-Portrait

The data allows us to explore in detail the different types of localities that make up France. Attractive city centers are experiencing centripetal movements as centrifugal forces affect enlarged low-density areas. Mobility turns out to be a

central feature of the inhabiting process. The major role of foreigners and tourists is made visible. Exploring places leads to focus on the diversity of spatial lifestyles.

Paris area and Large Cities' Cores Upgraded

If we adopt districting of *aires d'attraction des villes* ('city attraction areas'*)* set up by Insee, the French statistical office, Paris urban area has 14.790 million IYs (vs. 13.114 million residents) and the core area (the municipality of Paris) 3.663 million IYs (vs. 2.166 million residents). The IYs count thus reinforces both the weight of France's largest urban area and that of its core.

Paris' core area has an A-index of 169.1. The core municipalities of the six other areas with more than one million IYs have an average A-index of 120, those of 12 areas with 500,000-1 million IYs a 101.3 average. and those of the 92 areas of 100,000-500,000 IYs a 101.2 average. This distribution means that centers of larger cities are, by and large, more attractive than those of smaller ones. Said differently, we can observe positive feedback of the overall inhabiting population of an urban area to the attractiveness of its core area.

The attractiveness of areas characterized by a strong urban character (density + diversity, reinforced by urban mass) is thus confirmed by the new data. City centers turn out to be much more inhabited than the census would suggest because they combine and cumulate residence, employment, and all other types of activity. By and large, the classic model of the European city is confirmed: a compact built-up area used by a large array of practices, from residential (and it connected services) to any kind of industries, as diverse as movie theaters and redesigned, low-pollution manufacturing plants.

In French big cities, the *hyper-cores* that get the most out of this game, are always the historic centers, which are generally located at the geometric center of the built-up area. This is a confirmation of a revival of the Burgess concentric model (3) based on a core-periphery pattern. This model has been disputed for a while by two others, the Hoyt (4) sector (i. e. quadrant) model and the Harris and Ullman (5) multiple-nuclei model. The concentric approach is regaining strength due to economical, sociological, and political rearrangement of urban societies (6, 7). Our survey provides new evidence of this emerging process.

Trans-Urban: Movement as a Part of the Inhabiting Process

Four types of places have much more IY than residents: infrastructures of mobility, such as airports, railway stations, and inter-city transportations axes; business districts and business parks; shopping centers and, in general, commercial areas; leisure spots, such as forests, mountains, coastlines, and theme parks.

The *Métro-Boulot-Dodo* ('Subway-Work-Bed') motto was characteristic of the century-long Fordist era. It is challenged by the diversification of daily lives. Professional life is no longer reduced to a workstation but encompasses either an intrinsic mobile component (passenger transportation, storage, shipping, and delivery of goods), or a large number of locations for meetings with customers, partners, colleagues from other companies or organizations, seminars, conferences, field work, especially since professional missions include a significant element of creativity (6,7).

Moreover, thanks to the absolute and relative transportation cost decrease, daily life has become more mobile: shopping, supervising children's multiple activities, sport, cultural consumption, discovery of new places take up time and space in the schedule (8, 9).

These changes, which have been measurable for decades by a multitude of other tools and studies, and which are becoming undeniable thanks to this work, make the classic census approach still very useful but insufficient to describe the current relations between population and geography. Those who wish to understand urban and regional rationales can no longer be satisfied with the implicit postulate of an imaginary immobility in the place of residence and a mobility reduced to commuting between an exclusive home and a single workplace.

The importance of non-commuting, non-work-related mobility invites us to take these new mobility issues seriously. Inhabiting means *staying*, but also *crossing*. Trips and all aspects of 'motility' (10) that goes with them (preparation, impact, technical mastery, travel imaginations) should no longer be seen as a suspension of inhabiting but as another way of making and living an overall *inhabiting agency*.

Beyond commuters' daily networks, which remain powerful, a massive multi-pace intra-city and inter-city space is emerging. It includes, as destination points, tourism-based places. The facilities and the flows that embody this mobility is also part of the contemporary human 'habitats'.

Maps (Fig. 1-4 and 7-8) show the pivotal importance of tourism in the contemporary geographic layout of France. It is as if a part of the lived spaces were propelled elsewhere to the point of constituting significant concentrations of population, comparable to the usual cities. Moreover, the geography of foreigners (Fig. 9) shows that it is also through international attractiveness that these locations are fed.

Of 663 mainland France, non-crossborder urban areas, 28 only have an A-index above 200, meaning they have at least twice as many inhabitants (IY) as permanent residents (PR). These 28 urban areas are all, without exception, tourist spots, seaside or mountain resorts, and tourism-driven medium-sized cities that are located either on the Atlantic and Mediterranean coast and islands, or in the Alps.

Here is the list:

Val-d'Isère, Saint-Tropez, Chamonix-Mont-Blanc, Morzine, La Flotte, Bourg-Saint-Maurice, Carnac, Barcelonnette, Saint-Gervais-les-Bains, Quiberon, Cavalaire-sur-Mer, Le Grau-du-Roi, Bormes-les-Mimosas, Crozon, L'Île-Rousse, Calvi, Modane, L'Île-d'Yeu, Porto-Vecchio, Saint-Jean-de-Monts, Trouville-sur-Mer, Arcachon, Le Pouliguen, , Briançon, Port-Saint-Louis-du-Rhône, Capbreton, Ciboure, Agde.

These locations are neither central, nor suburban, nor peri-urban, nor exurban. Underlining the weight and the geographical distinctiveness of tourism-driven areas, we can speak of a *trans-urban* space as a separate and alternative style of urbanization. This observation is rendered by the Latin prefix *trans*, which expresses the displacement from one place to another. The trans-urban rationale refers to a set of processes that project urban phenomena into new locations. Instead of considering it as a seasonal event, the notion of the year-inhabitant allows it to be treated as a permanent reality, which certainly varies during the year, but whose impact can be compared to standard urban settlements.

Social voids

In areas of low population density, the differences are relatively small. However, the comparison between inhabiting and residential maps shows a general deficit of inhabitants in low-density areas. Empty places are even emptier. The 'Empty Diagonal' ('*Diagonale du vide'*), an area of deep demographic decline nationwide that runs from the Ardennes to the Pyrenees and also includes the Southern Alps can easily been detected on maps (Fig. 1-4). Moreover, there is a vast area located in the northwestern quarter, which combines low densities and low attractiveness. All these regions turn out to be even more depopulated because low residential densities are aggravated by the absence of attractivity, therefore by the absence of non-residence-driven activities. This is all the more true as commerce has moved out of small towns to concentrate in shopping malls. Conversely, polarization towards a reduced number commercial and employment attractive areas plays its role in these less dense areas. These zones can also benefit from the crossing of major mobility routes, which is even more significant when they are located, in low-density areas, like in the southeast of the Paris Basin, from Brie to Morvan.

In all the medium-sized and large cities, a distinction appears between exclusively residential suburban or peri-urban areas that are losing inhabitants and other parts where suburban employment and leisure centers are gaining. This generates a small-scale redistribution of polarities within urban areas which gives its mosaic style to A-index maps (Fig. 3-8).

A Place for Strangers

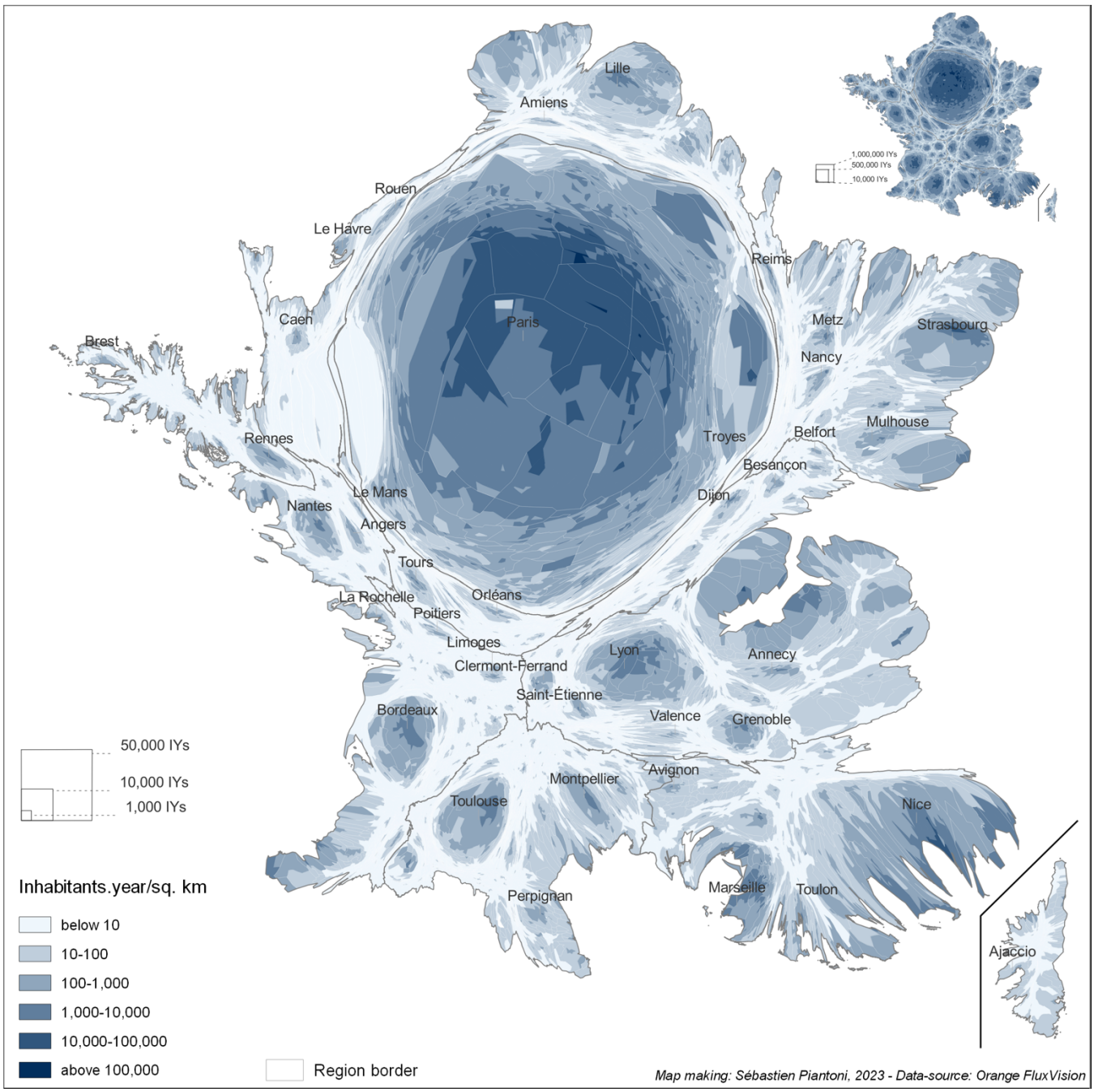
Our datasets use a distinction between two types of users, customers of national operators, and customers of foreign operators using roaming services in mainland France. 'Nationals' users are 'French residents', which means they have an address in France. They are either French citizens, or foreigners who live in France. Conversely, some French expatriates who come back for a while in their country of origin appear among the roamers.

The 65.1 million 'nationals' that appear in our data are in fact a balance. Surveys (11-13) show that an estimated 800,000 French IYs have been traveling abroad or have been border workers during this period. In the opposite direction, there are people not easy to evaluate, probably 100,000 to 300,000 IY, who can be I undocumented migrants who are flying under radars but who are cell-phone users, and II temporary foreign residents (in business stays or in second houses). Those two groups swell the numbers of 'nationals'. Taking into account the difference +800,000 – 200,000, we get 65.7 million IY, which is close to the official population of metropolitan France (65.6 to 65.8 million) as estimated in mainland France by the French statistical institute Insee between March 2022 and February 2023.

The first big surprise comes from the number of foreign residents present on the French territory, close to 4,977 million IYs, which is much more (at least 2 times more) than the number of IYs estimated from different tourism surveys. According to official data and considering both overnight stays, with paid or free accommodation, and day trips (12, 14), a volume of just under 2 million IY is obtained for 2022-2023. How can this substantial difference be explained? We will come back to this question below, but this dramatic difference already shows that another, partly unknown France, is being unveiled thanks to these data.

In mainland France, those around 5 million IYs composed of customers of non-French operators provide a good proxy to approach non-resident foreigners. They are composed of visitors performing business travels, cross-border trips, or vacation tourism. In addition, there are also about 5.2 million foreign residents in the same territory, who generate approximately 5.148 million IYs. In total there are 4.977 + 5.148 = 10.125 million foreigners living in metropolitan France. If the 2.5 million IYs corresponding to current French citizens born foreign nationals (naturalized immigrants) are added, we obtain a total of 12.75 million foreigners or having a foreign origin. They represent 18% of the total population of effective inhabitants. Like its European neighbors, France turns out to be a global place, a concentrate of human globality, where almost one in every five inhabitants brings their share of otherness to the host society.

Fig. 9. **Density, France, Foreign Visitors (IY) 2022-23, Cartogram.**
Top right: Density, Inhabitants.year, France, 2022-23, Cartogram.



Base map: surfaces on the map are proportional to IYs in each of the 48,789 spatial units.

This cartogram (Fig. 9) is directly comparable to the general cartogram of IYs (Fig. 2, duplicated, top right). Here, the base-map shows the specific spatial distribution of the foreign visitors. It clearly demonstrates their extreme concentration in large cities and tourist areas, especially Paris metropolitan region and in particular its core area. Counted in IYs, foreign visitor sometimes reach astonishing densities, as on the banks of the Seine, in Marais or in the Opera district, where they can exceed 100,000 inhabitants/sq. km and in some cases outnumber the French residents. In about 2,000 spatial units out of the 50,000 that cover the French territory, they exceed 1,000 inhabitants/sq. km, which can be considered as an urban density threshold. The other major urban areas come next, with an advantage for those located close to the border, such as Lille and Strasbourg, or Toulouse, which hosts the main Airbus group site in Europe. The tourist areas are in the lead, with a preeminence of the French Riviera, a region that the French tourists are tendentially neglecting. Finally, special spots such as mobility hubs, theme parks and even places of pilgrimage (such as Lourdes) are doing well. This confirms that tourist locations are largely fueled by international attractiveness.

In short, foreign visitors alone possess a considerable populating power. Compared to the general map designed in the same way (Fig. 2), the map of foreign visitors overcolors France's spatial contrasts. Lots of people converge to the big cities and tourist places, while, conversely, a large part of the country remains empty.

A Global France

Where do these foreigners come from? Once again, the data surprise us. The Table 1 shows a relationship between France and the rest of the world that is not exactly what we might have expected. The visitors' origins are sorted in descending order of the top 175 countries that send visitors, in proportion to their population and classified by continental block.

The 2022-3 context in which these data were collected includes two powerful features. On the one hand, the last phase of the CoviD-19 pandemic that kept Asian tourists, especially Chinese, away from Europe. in 2019, before the Pandemic, Chinese tourists were more than two million in France, which means about 30 000 IYs. They disappeared in 2020 and had not yet returned to Europe in early 2023. On the other hand, the war in Ukraine has certainly strengthened in a multidimensional way the links between Eastern and Western Europe, including France. For the same reason, Russia was almost cut off from Europe, hence its degraded rank beyond 100th place.

Moreover, the table shows the intense links between metropolitan France and its former colonial empire. Overseas France (particularly Guadeloupe, Martinique, Réunion, and Mayotte) and the former French colonies occupy a prominent place in this ranking.

With this in mind, an original configuration of French globality can be observed.

The most notable point in this table is the compact score of all European countries. The emergence of the Europe as a consistent space, viewed from France, does not come under the same rationale as for long-term immigration. Some countries very clearly dominate among the origins of foreigners who have settled more permanently in France (15). These locations refer to past (southern Europe, mid and late 20th century) and present (Maghreb and Sub-Saharan Africa, late 20th and early 21st centuries) development differentials between these countries and France. This pattern is not found here. Nor is there any evidence of a tourist 'sunbelt'. Northern or continental Europe is not more attracted to France than Mediterranean Europe.

Lastly, there is no clear 'visa effect'. In addition to Europeans, nationals of most of the Americas and developed countries in the Asia-Pacific region are exempt from visa requirements to enter the Schengen area, and therefore France. This does not seem to bring them any additional proximity.

The spatial configuration of foreign visitors' locations leads to the assumption that these inhabitants are not so different from French residents. There motivations appear to be diverse (work, leisure, personal relations), as if Europeans were circulating in their own single country. As a result, a hypothesis can be formulated: Europe operates as an interactive, intertwined space in which each region (in this case the national areas) is of interest to the others for different reasons. In the absence of significant differentiation based on access time between European countries, we can speak of Europe as a place, a space for which internal distances are of little relevance (16).

In brief, these data draw a global geography of French visitors' attractiveness. They provide an image of different levels of attractivity of the French territory. They suggest new ways to define a world-scale nearness.

Table 1. **Foreign visitors in France: Ranking of the 150 top origin countries, IYs per capita.**

RANK / REGION	1-25	26-50	51-75	76-100	101-125	126-150
Western Europe	Monaco, Luxembourg, Switzerland, Belgium, Andorra, Spain, Germany, UK, Portugal, Sweden, Malta, Ireland, Italy, Austria	Denmark, Norway, Iceland, Finland				
Eastern, Balkan Europe	Lithuania, Poland, Romania, Bulgaria, Slovenia, Ukraine, Czech R	Latvia, Estonia, Slovakia, Hungary, Croatia, Serbia, Greece, Moldova, Montenegro	Cyprus, Albania, Bosnia-Herzegovina, N Macedonia	Belarus		
North Africa; Middle East	UAE	Tunisia, Algeria, Israel, Kuwait, Qatar, Morocco, Lebanon,	Georgia, Armenia, S Arabia	Bahrein, Turkey, Oman,	Azerbaijan, Jordan,	Egypt, Iran, Libya, Iraq
China; South, South-East, East Asia; Oceania	Réunion-Mayotte,	French Polynesia, Singapore	New Caledonia, Australia, South Korea, New Zealand	Malaysia, Thailand, Japan, Philippines, Sri Lanka	Wallis and Futuna, Cambodia, Brunei, Indonesia, Vietnam	China, India, Maldives, Pakistan, Laos, Mongolia
Americas	Guadeloupe, Martinique	Saint Pierre and Miquelon,	Canada, US, Chile, Uruguay, Mexico, Peru	Panama Colombia, Costa Rica Argentina, Brazil, Trinidad and Tobago, Barbados, Antigua and Barbuda	Salvador, Paraguay, Suriname, Guatemala, Puerto Rico, Cuba, Dominican R, Venezuela, Ecuador	St. Kitts and N, Grenada, St. Vincent and G, St. Lucia, Jamaica, Bolivia
Sub-Saharan Africa		Mauritius, Comoros	Senegal, Congo R, Ivory Coast, Mali, Gabon, Guinea, Cameroon, Cape Verde	Benin, Togo, Burkina Faso, Mauritania, Djibouti, Madagascar, Seychelles, Guinea B	Congo RD, E Guinea, Central African R, São Tome and Príncipe, Chad, Niger, Sierra Leone, Namibia, Botswana	Gambia, Ghana, Rwanda, Kenya, Nigeria, Mozambique, Liberia
Russia and Central Asia					Russia, Kazakhstan, Kyrgyzstan	Uzbekistan, Afghanistan

Discussion

A revolution in progress

The results presented in this article are only the first phase of a research project that includes at least five phases. A mobility approach, the construction of a typology of spatial units, the redefinition of relevant urban areas, and the exploration of rhythms (day, week, year) are still in progress. A comparison of pre- and post-Pandemic situations thanks to the processing of two distinct datasets (2019-20 and 2022-3) will allow for the analysis of such a massive, disruptive event. Future data processes of the same dataset will take advantage of contingency tables that provide limited but

useful information on personal profiles. All of these 'worksites' will be based on the same simple principle: we can know how to connect people and places and explore the richness of what being an inhabitant means.
The first results suggest that it is time for a paradigm shift. The quality of data provided by telephone operators can certainly be further improved to reach the same level of reliability as the censuses. The available data already allow for a thorough processing, and this would be even more the case if a dialogue between operators, statistical offices and researchers were to take place.
The very principle of censuses – counting people at home as a sufficient proxy to give a reasonably accurate picture of an inhabited space – deserves substantial enrichments. It is time to find out how and where the individuals spend their days, their months, and their years, while obviously respecting their privacy. The use of high quality, comprehensive data to explore effective realities of inhabiting is not only a complement for the classic censuses, but it could also (and probably should) become a solid baseline for all those who are concerned with the spatial dimension of the social world. This quest is no longer a utopia, it becomes a reachable horizon for statistical offices, for scholars, and for the public.

Materials and Methods

A Fresh Approach to Emerging Data

A long Tradition: Censuses and an Increasingly Elusive Reality

For a long time, knowledge of the relationship between humans and places was based on censuses that relied on the postulate of immobile inhabitants, supposed to stay all the yearlong in their homes. What might have seemed an acceptable approximation in the 19th century gradually became less and less relevant to describe people's actual relationship to places. Statistical data gradually included studies on mobility, yet often restricted to a single, daily trip to a fixed workplace (17). Statistical data are enriched by sample surveys on mobility, but they generally study daily travel and tourism separately. Finally, when they are more comprehensive and include all travel motives, they lack completeness by being reduced to one urban area and they lack precision by dividing the region studied into vast heterogeneous zones.

State of the art: Promising Answers and Missing Questions

After probative initial results in the early 2010s, the use of phone data to approach human presence in a determined area has shown its relevance. Today, the state of the art, to large extent, is about methodological issues, exploring two major issues: I quality and sampling assessment: the reliability and representativeness of phone data in comparison of baseline information of human presence or movement provided by censuses or by mobility surveys (18-25); and II thematic issues at local scale: scholars and experts have used phone data in combination with other baseline datasets to build up a more sophisticated approach to mobility and urban configurations: (26) on large Spanish cities; (27) on Washington, DC; (28) on Seoul; (29) on Helsinki and (30) on Czech mobility) or to the spatial impact of exceptional events, such as the 2020-22 Pandemic (31-32).
Data deriving from mobile phone use has become massively available, with rapid changes on their substance and their quality. Phone operators provide 'events' (when a phone matches between an antenna), but originally these events were defined by calls, while now any turned-on phone generates usable information. Turned-off phones and non-equipped people are not forgotten anymore thanks to more efficient socio-demographic adjustments. Besides, the location techniques coupling antennas and places are getting very accurate. An alternative method, the collection of GPS-traced activities is another way, very promising too, to locate users.
These works are all stimulating. Several studies have focused on the relevance of the data by benchmarking them against the usual statistical sources. The residential population provided by censuses or other pre-existing information are taken as the ground truth. Phone data are then mobilized to measure daily or seasonal variations.
This step was undoubtedly necessary. It appears to be insufficient today, because if the aim is to verify that phone data provide the same type of information as the censuses, it is not clear what added value they would have. Furthermore, phone data that are now much more reliable as their providers have learned the various biases that could have undermined research.
The current challenge is therefore to design a framework that will make the most of this novel information.

The Datasets

The datasets used in this project were provided by Orange FluxVision, an Orange Business product.

The set provides the number of presences in 48,789 spatial units (French statistical Iris which combine small municipalities and subdivisions of the most populated ones). They host an average of ≈1400 inhabitants. In these units, our data measure human presence in each of the 48 30-minute time slots, for 3 types of days in 4 different types of months (February, August, September, and November) in March 2022-Ferbruary 2023 period for 2 types (French residents and roamers) of users, or 57.6 million elementary information items.
This dataset is totally anonymous. The absence of personal identifiers makes impossible any individual tracking.

Data Adjustment and Relevance

Two types of adjustments have been carried out by Orange Business before the data have been delivered to *La France habitée* taskforce.

1 Coupling antenna-data to spatial units. Matching the antenna network with the spatial units requires specific adjustment work. Orange Business acknowledges in a "*Note méthodologique*" (unpublished, issued on January 14, 2022) the significant probability of false positives and false negatives, especially in low density areas, where the antenna network is also the least dense. In response to this problem, the operator makes adjustments that take into account all the cell phone equipment and telephone traffic in the area concerned. In cooperation with Orange Business, the taskforce made a detailed adjustment to further refine the reliability at the hourly level.
Since these initial discrepancies are random, the distortions remain very limited after adjustment. By making various cross-checks with independent information on the same units (residential population, mobility, activities), we have estimated that significant residual errors could affect ≈2% of the units. They confirm Orange Business own estimations. Above all, they do not affect the geographical configuration of IYs, even on a micro scale.

2 Weighting data to curb limits of the datasets. Market-share of the operator (Orange: 35.4% in mainland France), phones off, no-phone people, dual-phone people have been taken into account by the operator in the delivered data.
The potential error estimation is <±1 %. There is no distortion on the distribution of values in spatial units.

Two complementary treatments have been carried out by *La France habitée* taskforce.

3 Weighting people presence. Using the distinction between the number of people entering a spatial unit and the overall number of those staying in the same unit, we were able to accurately measure the time spent by the persons than have stayed less than 30 minutes. The potential error estimation is <±0.01 %. There is no distortion on the distribution of values in spatial units.

4 Estimating the actual all-year presence. The different types of days (working days, Saturdays, Sundays) and months have been taken into account to adjust the data. The potential error estimation is <±0.5%. There is no distortion on the distribution of values in spatial units.

5 Finally, for privacy reasons, Orange Business deleted information about presences of less than 20 people in a spatial unit in a given time slot. This operation may result in a slight underestimation of the effective population in the micro-units (<100 inhabitants) that host only 0.3% of the country's population. This bias can therefore be considered negligible.

Asking Questions to the Data

The relative dissatisfaction with pioneer studies stems from the type of questions that have been asked of the data. For our part, we focused on fundamental questions for which these new data could constitute a relevant resource. These questions are the following: How many people *actually* live in the French territory? How are they spatially distributed? How does this new information modify the geography of population: settlement, density, urbanity? How can this study on France help us for comparable approaches applied carried out elswhere?

A Novel Methodology: IY and A-index

A IY is a full year stay equivalent. The value of IY_i in a determined space is calculated as follows [1]:

[1]

$$\mathrm{IY}i = \sum_{i=1}^{n} \frac{\mathrm{S30}i}{17520}$$

where S30*i* is the number of people that stayed for any of the forty-eight 30-minute time lapses of a day of a determined year in the *i* spatial unit.

The principle is simple: we add up all the people present in a place of every single half-hour period of a year and we divide the result by the number of half-hours that a year comprises (17520). This allows us to define a population made up of people who, all together, have spent a certain number of years or fractions of a year in this place.

The relation between the number of primary-residents (PR) of a spatial unit *i* as measured by censuses (33) and their actual inhabitants (IY) can be considered a simple and powerful indicator of the attractiveness of a place. The attractivity index (A-index) is calculated as follows [2]:

[2]

$$\mathbf{A}i = 100\left(\frac{\mathrm{IY}i}{\mathrm{PR}i}\right)$$

The A-index shows for any spatial unit *i* the propensity of a whole reference population (in our case, global) to be more present in a location than the residents would be if they stayed permanently in that location.

Cartograms

The base map of a Euclidean map is obtained through a homology of angles, lines, and surfaces between the referred reality and its cartographic transformation. In a cartogram, the base map is built up using a series of located data, such as population or any other continuous variable. In this project, the cartograms' base maps are drawn from information provided by the datasets (general population's IYs and foreign visitors' IYs). To create these base maps, the ScapeToad opensource application (http://scapetoad.choros.place) has been used. For Fig. 2 and 9, the base map data are the Iris spatial units (48,789). For Fig. 4, 6, and 8, the base map data are French municipalities, *arrondissements*, and *quartiers* (35,048).

References

1. MTECT, "Enquête sur la mobilité des personnes 2018-2019" (Ministère de la Transition Ecologique et de la Cohésion des Territoires, 2019); https://www.statistiques.developpement-durable.gouv.fr/resultats-detailles-de-lenquete-mobilite-des-personnes-de-2019.

2. M. Pearce et al., "Enquête nationale mobilité et modes de vie", (Forum Vies Mobiles, Paris 2020); https://www.actu-environnement.com/media/pdf/news-35081-Enquete-mobilite-francais-forum-vie-mobiles-2020.pdf.

3. E. W. Burgess, "The Growth of the City: An Introduction to a Research Project" in R. E. Park et al., Eds., *The City*, (University of Chicago Press, Chicago, 1925), pp. 47-62.

4. C. D. Harris, & E. L. Ullman, The Nature of Cities. *Annals of the American Academy of Political and Social Science*. **242**, 7-17 (1945).

5. H. Hoyt, *The Structure and Growth of Residential Neighborhoods in American Cities*, Washington, DC: Federal Housing Administration (1939).

6. R. Florida, *Cities and the Creative Class* (Routledge, London, 2005).

7. J. Lévy, "The Triumph of Urbanity and Spatial Justice" in *OECD Regional Outlook 2019. Leveraging Megatrends for Cities and Rural Areas* (OECD, Paris, 2019), pp. 229-239.

8. N. Yau, "A Day in the Life of Americans", *FlowingData* (2015); https://flowingdata.com/2015/12/15/a-day-in-the-life-of-americans/.

9. IPUMS, *Time Use Atus Project* (University of Minnesota Minneapolis, 2021); https://www.atusdata.org/atus/about_atus.shtml

10. V. Kaufmann, M. Bergman, Max & D. Joye, Motility: Mobility as Capital. *International Journal of Urban and Regional Research*, **28.4**, 745-756 (2004); https://doi.org/10.1111/j.0309-1317.2004.00549.x.

11. Données et études statistiques, "Comment les Français se déplacent-ils en 2019" (Ministère de la transition écologique et de la cohésion des territoires, Paris, 2019) ; https://www.statistiques.developpement-durable.gouv.fr/comment-les-francais-se-deplacent-ils-en-2019-resultats-de-lenquete-mobilite-des-personnes.

12. S. Marchand, Ed., "Memento du tourisme" (Ministère de l'économie et des finances, Paris, 2019); https://www.entreprises.gouv.fr/files/files/directions_services/etudes-et-statistiques/stats-tourisme/memento/2018/MEMENTO_TOURISME_Edtion2018-WEB.pdf.

13. I. Debouzy & A. Reffet-Rochas, "Travailleurs frontaliers : six profils de 'navetteurs' vers la Suisse" (Insee, Paris, 2022) ; https://www.insee.fr/fr/statistiques/6444379.

14. A. Mainguené et al.,."Été 2022: une fréquentation touristique au-dessus de son niveau d'avant-crise" (Insee, Paris, 2022) ; https://www.insee.fr/fr/statistiques/6537925.

15. M. Bodier et al., 2023. "Immigrés et descendants d'immigrés" (Insee, Paris, 2023); https://www.insee.fr/IMMFRA.pdf.

16. J. Lévy, "Inhabiting" in R. Lee et al., Eds., *The Sage Handbook of Human Geography* (Sage, London, 2014), pp. 45-68.

17. Insee, "Mobilités professionnelles en 2019 : déplacements domicile-lieu de travail" (Insee, Paris, 2022); https://www.insee.fr/fr/statistiques/6454112.

18. P. Deville et al., 2014. Dynamic population mapping using mobile phone data. *PNAS*, **111**, 45, 15888-15893 (2014); https://doi.org/10.1073/pnas.1408439111.

19. F. Ricciato et al., "Estimating population density distribution from network-based mobile phone data", Joint Research Centre, European Commission, Brussels (2015); https://publications.jrc.ec.europa.eu/repository/handle/JRC96568.

20. P. Bonnel, Passive mobile phone dataset to construct origin-destination matrix: potentials and limitations. *Transportation Research Procedia* **11**, 381-398 (2016); https://pdf.sciencedirectassets.com/308315/1-s2.0-S2352146515X00093/1-s2.0-S2352146515003233/main.pdf.

21. D. Wu et al., "Urban Population Distribution Characteristics Analysis Method based on Mobile Phone Data", *Proceedings of the 5th ICAMCS* (2016); https://www.atlantis-press.com/proceedings/icamcs-16/25854911.

22. T. Horanont et al., Resembling Population Density Distribution with Massive Mobile Phone Data. *Data Science Journal*, **17** (2018); http://doi.org/10.5334/dsj-2018-024.

23. B. Sakarovitch et al., Estimating the Residential Population from Mobile Phone Data, an Initial Exploration, *Economie et Statistique/Economics and Statistics*, **505-506**, 109-132 (2019);

24. M. Fekih, et al., A Data-Driven Approach of Origin-Destination Matrix Construction from Cellular Network Signaling Data: A Case Study of Lyon Region (France)", *Transportation* **48**, 1671-1702 (2020); https://link.springer.com/article/10.1007/s11116-020-10108-w.

25. F. Ricciato et al., Towards a methodological framework for estimating present population density from mobile network operator data. *Pervasive and Mobile Computing*, **68**, 101263 (2020); https://doi.org/10.1016/j.pmcj.2020.101263.

26. T. Louail et al., "From Mobile Phone to the Spatial Structure of Cities". *Sci Rep* 4, 5276 (2014); https://doi.org/10.1038/srep05276.

27. K. Sparks et al., "*Modeling Building Use and Population Distribution Opportunity Using Open Geosocial Data in Urban Areas" (*US Department of Energy, Washington, DC, 2019); https://www.osti.gov/servlets/purl/166601.

28. J.-H. Lee et al., Spatiotemporal distributions of population in Seoul: joint influence of ridership and accessibility of the subway system. *EPJ Data Science*, **10**, 41 (2021), https://doi.org/10.1140/epjds/s13688-021-00298-3.

29. Bergroth, et al., A 24-Hour Population Distribution Dataset Based on Mobile Phone Data from Helsinki Metropolitan Area. *Sci Data*, **9**, 39 (2022); https://doi.org/10.1038/s41597-021-01113-4.

30. M. Halás, et al., Population movements based on mobile phone location data: the Czech Republic. *Journal of Maps*, **17**, 116-122 (2021), https://doi.org/10.1080/17445647.2021.1937730.

31. L. Galiana et al.,. "Retour partiel des mouvements de population avec le confinement", *Insee-Statistiques et Études (*Insee, Paris*,* 2020) ; https://www.insee.fr/fr/statistiques/4635407.

32. R. Levin et al., Insights into population behavior during the COVID- 19 pandemic from cell phone mobility data and manifold learning". *Nat Comput Sci*, 1, 588-597 (2021); https://doi.org/10.1038/s43588-021-00125-9.

33. Insee, *Recensement 2019* (Insee, Paris, 2022); https://www.insee.fr/fr/information/6444222.

Acknowledgments

The Team

This project is led by *La France Habitée* research taskforce, directed by Jean Coldefy and Jacques Lévy. The group includes Sébastien Piantoni, Julien François, Jorge Cabrera, Marie-Emmanuelle Huilo, and Stéphane Gallardo.
The project benefited from the precious support of Boris Beaude and Ogier Maitre (STS Lab, Lausanne University) for the data preparation and processing.
Multiple and fruitful interactions with Kévin Grossetti, from Orange Business, were greatly beneficial to the achievement of the project.

Funding

The project was funded by a consortium, created on March 16, 2022, including the following entities:
Transcité (NPO);
Chôros (non-profit research network);
Spatial Intelligence Chair, UPHF (academic research lab);
DGITM-MTECT (French ministerial department);
La Fabrique de la Cité (NPO).

Author Contributions

The following persons participated in the production of this article:
Jacques Lévy: approach, framework, conceptualization, discussion, writing, editing.
Jean Coldefy: approach, framework, conceptualization, discussion, editing.
Sébastien Piantoni: conceptualization, data processing, mapping, discussion, editing.
Julien François: data processing, discussion, editing.

The scientific proofreading carried out by Enka Blanchard was extremely helpful.

Competing Interests

Authors declare that they have no competing interests.

Data and Materials Availability

All data needed to evaluate the conclusions in the paper are present in the paper. Comprehensive data provided by Orange Business that have been used in this research can be provided by the authors pending scientific review and a completed material transfer agreement.