Mapping the Third Republic. A Geographic Information System of France (1870–1940).

Gay, Victor

Toulouse School of Economics, Institute for Advanced Studies in Toulouse

1 October 2020

MAPPING THE THIRD REPUBLIC

A GEOGRAPHIC INFORMATION SYSTEM OF FRANCE (1870–1940)

Victor Gay[1]

[1]Toulouse School of Economics (TSE) and Institute for Advanced Study in Toulouse (IAST), University of Toulouse Capitole, Toulouse, France (victor.gay@tse-fr.eu).

ABSTRACT

This article describes a comprehensive geographic information system of Third-Republic France: the TRF-GIS. It provides annual nomenclatures and shapefiles of administrative constituencies of metropolitan France from 1870 to 1940, encompassing general administrative constituencies (*départements, arrondissements, cantons*) as well as the most significant special administrative constituencies: military, judicial and penitentiary, electoral, academic, labor inspection, and ecclesiastical constituencies. It further proposes annual nomenclatures at the contemporaneous *commune* level that map each municipality into its corresponding administrative framework along with its population count. The 901 nomenclatures, 830 shapefiles, and complete reproduction material along with primary sources of the TRF-GIS database are available at https://dataverse.harvard.edu/dataverse/TRF-GIS.

Background & Summary

For seven decades, between the collapse of the Second Empire in 1870 and the establishment of the Vichy Regime in 1940, France remained under a single political regime: the Third Republic. Despite the unprecedented stability of its institutions, France underwent dramatic socio-economic changes during this period. These were the result of several critical junctures, among which the Second Industrial Revolution and the modernization of France (1870–1914), the establishment of free, mandatory, and secular education (1881–1882), the Separation of the Churches and the State (1905), the First World War (1914–1918), the economic and political turmoils of the 1930s, and the defeat by Nazi Germany (1939–1940).[1–4] Research programs aimed at understanding how these experiences shaped the French society increasingly rely on empirical evidence, in part due to the conjunction of two phenomena: the substantial production of administrative statistics under the Third Republic and the recent revival of quantitative history.[5,6] Ongoing advances in optical character recognition (OCR) techniques applied to the digitization of historical archives suggest that this trend will likely intensify in the future.[7]

But a fundamental element empirical researchers typically need in order to process, map, and analyze spatially localized historical data is an underlying geographic frame of reference—some form of geographic information system (GIS). And contrary to the United States (nhgis.org), Great Britain (visionofbritain.org.uk), and Germany (digihist.de), a historical GIS of France has not been produced thus far.[8–10] This lack is especially problematic for quantitative research in the context of Third-Republic France, as statistics then were oftentimes produced by administrations operating at heterogeneous and incompatible levels of aggregation, despite the coordination efforts by the *Statistique Générale de France*—Figure 1 provides a simplified diagram of the complexity of the administrative framework of statistics-producing entities over this time period.[11] As a result, it currently falls onto individual researchers to complete this time-consuming, difficult, and yet, crucial task. Although simple département-level historical shapefiles are routinely produced for the needs of single studies, these are generally insufficiently documented and unavailable to the community. Furthermore, nomenclatures and shapefiles for more fine-grained administrative constituencies remain rare, and, when they exist, are hardly curated: they are generally neither findable due to lacking metadata, nor accessible under appropriate licensing, nor interoperable with other GISs, nor reusable in machine-readable format.[12] A notable exception is LARHRA's cantons shapefiles for 1884 and 1925, which were produced by the manual vectorization of georeferenced historical maps (a description of acronyms used in this article is available in Table 1).[13] Overall, research endeavors in the context of Third-Republic France are bound to be hindered by the lack of a shared frame of reference that would not only ease researchers efforts by generating economies of scale, but also improve reproducibility and interoperability across scientific studies.[14,15]

To alleviate these issues and empower research programs in this context, this article proposes a comprehensive geographic information system of Third-Republic France: the TRF-GIS. It provides annual nomenclatures (codifications and toponymies) along with shapefiles of administrative constituencies of metropolitan France (mainland France and Corsica) from 1870 to 1940—901 nomenclatures and 830 shapefiles in total. It encompasses general administrative constituencies (départements, arrondissements, cantons) as well as the most significant special administrative constituencies: military constituencies (military

regions and subdivisions), judicial constituencies (courts of appeal and of first instance), penitentiary constituencies, electoral constituencies (*circonscriptions*), academic constituencies (academies), labor inspection constituencies, and ecclesiastical constituencies (dioceses). The TRF-GIS further proposes annual nomenclatures at the contemporaneous commune level that map each municipality into its corresponding administrative framework along with its population count.

The construction of TRF-GIS shapefiles uses a methodology different to the manual vectorization of georeferenced historical maps.[7,13,16,17] It first reconstructs the administrative framework contemporaneous communes were embedded in, then aggregates these units into relevant constituencies using modern communes shapefiles. Despite some limitations, this method offers substantial advantages: not only it provides more precise results than existing outputs, but, more importantly, it also enables to circumvent the fundamental problems of non-existence of annual historical maps and of time and resource investments associated with manual vectorization methods.

By providing a comprehensively-curated common frame of reference that encompasses all aspects of the French society, the TRF-GIS will help create the conditions for a better understanding of the dramatic socio-economic changes that France underwent during the seven decades that lasted the Third Republic.

Methods

General methodology

The underlying structure of the TRF-GIS builds on three datasets that are available under open licensing (Figure 2 provides an overview of the construction logic of the TRF-GIS):

1. **INSEE's *Codes Officiel Géographique* (COG) nomenclatures.[18]**

   INSEE's annual COGs provide the official nomenclature of communes and of general administrative constituencies in the geography of a given year. Existing communes are uniquely identified through a five-digit coding scheme, which first two digits correspond to a commune's département and last three to its number within the département, generally assigned in ascending alphabetical order. For instance, the commune of Allonne in the département of Oise (`60`) holds INSEE code `60009`. COGs and their coding scheme were first established in 1943 and have remained stable ever since, barring changes in communes' territorial structures. The construction of the TRF-GIS uses the COGs 2005 and 2011.[19,20]

2. **The *Histoire Administrative des Communes* (HAC) database distributed through cassini.ehess.fr.[21,22]**

   The HAC database provides, for each of the 41,410 communes that ever existed between 1793 and 2005, a unique record that contains a commune's dates of creation and subsequent modifications, the general administrative constituencies it ever belonged to, its INSEE codification, the evolution of its toponymy, and its municipal population counts across censuses—these records can only accessed individually through cassini.ehess.fr. Information herein relies on official administrative acts published in the *Bulletin des Lois* from 1793 to 1931 and in the *Journal Officiel* hereafter.[23,24] Except for a few cases documented in the TRF-GIS source code, communes in the HAC database are effectively characterized in 2005 geography. Communes that no longer existed by 2005 are assigned the INSEE code of their absorbing commune in the HAC database. For instance, absorbed by Beauvais in 1943, the commune of Voisinlieu holds Beauvais' INSEE code `60057`. Former and current communes can nevertheless be uniquely identified through the url query parameter of their individual record on cassini.ehess.fr, which I denominate "Cassini code." It is assigned in ascending alphabetical order from 0 to 41,475 (with a few gaps). For instance, while Beauvais holds Cassini code `3332`, Voisinlieu holds Cassini code `40965`—Voisinlieu's record is available at http://cassini.ehess.fr/cassini/fr/html/fiche.php?select_resultat=40965.

3. **IGN's *GEOFLA®Communes Édition 2011 France Métropolitaine* (GEOFLA) shapefile.[25]**

   IGN's GEOFLA 2011 shapefile provides an official representation of the 36,568 communes of metropolitan France in 2011 geography in polygon and vector forms. It is expressed through an RGF93 Lambert-93 projection system and derived from the geometry of IGN's BD CARTO® with a precision of 1:1,000,000.

TRF-GIS shapefiles

Annual shapefiles use the GEOFLA 2011 as their underlying frame to the exclusion of the territories that did not belong to France during certain years (Figure 3): Alsace-Lorraine from 1871 to 1918, and the communes of La Brigue and Tende throughout the period—these were not annexed by France from Italy until 1947. As a result, shapefiles that constitute the base frame for 1870 and 1919–1940 represent 36,566 of the 36,568 existing communes of 2011 and 34,932 for 1871–1918. Each of these communes is then matched to the general administrative constituencies they belonged to in the relevant year using the HAC database, which initial 2005 geography is converted into 2011 geography using the COGs 2005 and 2011—between these two dates, nine communes disappeared and seven were created.

While nearly all communes of 2011 were created by 1793, some were not until after 1870. For these 737 communes, I assign the constituencies their parent commune belonged to prior to their existence. For instance, the commune of Voisinlieu was created from Allonne in 1930, a parent commune that belonged to the département of Oise between 1870 and 1929. Hence, Voisinlieu is assigned to the département of Oise in the base frames for the shapefiles 1870–1929.

Commune-level polygons are then dissolved to create general administrative constituency-level polygons. The construction of special administrative constituency shapefiles relies on this initial characterization, as these constituencies were generally based on the general administrative framework constituted by départements, arrondissements, and cantons.

TRF-GIS nomenclatures

Annual constituency-level nomenclatures use contemporaneous toponymies and codifications as provided in the HAC database for general administrative constituencies. For special administrative constituency nomenclatures, I rely on historical sources. Annual commune-level nomenclatures are composed only of those communes that existed in the relevant year. For instance, the commune of Voisinlieu only appears in the commune-level nomenclatures 1930–1940. Figure 4 displays the evolution of the number of communes that compose annual commune-level nomenclatures. It ranges from 35,974 communes in 1871 to 38,038 communes in 1939–1940.

Municipal arrondissements of Paris, Lyon, and Marseille

The three building blocs of the TRF-GIS have different treatments of the 45 municipal arrondissements of Paris, Lyon, and Marseille: while the GEOFLA treats them as independent elements, explaining why it contains 36,610 distinct entities representing 36,568 communes, COGs and the HAC database do not. Because they have been stable since 1859, the TRF-GIS keeps the 20 municipal arrondissements of Paris as distinct entities. However, it treats the communes of Lyon and Marseille as single entities because their municipal arrondissements were modified several times since 1870, making it challenging to match these to the GEOFLA.

Timeline

The TRF-GIS encompasses the period of the Third Republic. This political regime was proclaimed following the fall of the Second Empire on September 4, 1870, and later dissolved on July 10, 1940, after the capitulation against Nazi Germany and the advent of the Vichy Regime. Although the Third Republic began in 1870, the general administrative constituencies that prevailed during the regime were only effective by May 1871 and the signature of the Frankfurt treaty, which settled the aftermath of the Franco-Prussian War and the annexation of Alsace-Lorraine by the newly unified German Empire. For interoperability with research programs covering the Second Empire, the TRF-GIS begins in 1870 and therefore encompasses constituencies prevailing by the end of this regime.

Limitations

The TRF-GIS exhibits two limitations. First, its shapefiles rely on delineations of communes in 2011 geography through the GEOFLA 2011. I use this dataset for quality and reproducibility purposes, as it is the earliest geography for which an IGN-produced commune-level shapefile is accessible under open licensing.[25] Moreover, annual shapefiles of commune boundaries for 1870–1940 have not been produced thus far—this is among the endeavors of the ongoing project COMMUNE HIS-DBD, which is scheduled to be completed by the end of 2022.[26] This needs not invalidate the TRF-GIS, for two reasons. First, commune boundaries did not change much since the beginning of the Third Republic: out of the 36,568 communes represented in the GEOFLA 2011, less than seven percent ever underwent territorial modifications (creation, suppression, territorial transfer) between 1870 and 2011 (Figure 5). Second, TRF-GIS shapefiles represent constituencies that are for the most part defined at higher levels of aggregation than communes, thereby limiting potential inaccuracies at their boundaries. Still, this might represent an issue for the larger urban centers as some increased in size over time, absorbing peripheral parcels and hamlets. For instance, the area of the Bois de Vincenne was not absorbed by Paris until 1929. TRF-GIS shapefiles are therefore better suited for usage with regional and national extents—for local extents, specific tools such as ALPAGE for Paris might be more relevant.[27]

The second limitation of TRF-GIS shapefiles is due to some constituencies being at times defined across commune boundaries, generally in the vicinity of urban centers. For instance, the area around the city of Montluçon was divided between two cantons, Montluçon-Ouest and Montluçon-Est, with the commune of Montluçon being subdivided between these two cantons along the Cher river (Figure 6, Panel a). While this poses no issue for the construction of constituency-level nomenclatures, shapefiles and commune-level nomenclatures have communes as their base unit. I circumvent this problem by implementing INSEE's and IGN's common methodology when building their COGs and GEOFLAs, and use the concept of "pseudo-constituency:" whenever part of a constituency is defined across commune boundaries, a pseudo-constituency is created for that specific part, which level of aggregation is the commune itself and which encompasses all constituencies defined within that commune. In the case of Montluçon's cantons, TRF-GIS shapefiles and commune-level nomenclatures assign communes that fall entirely into cantons to either cantons of Montluçon-Ouest (seven communes) and Montluçon-Est

(eight communes), while the commune of Montluçon itself, being subdivided between these two cantons, is assigned to the pseudo-canton of Montluçon (Figure 6, Panel b).

Institutional details

General administrative constituencies

General administrative constituencies (départements, arrondissements, cantons) formed the basic territorial divisions through which the State deployed its administration. During the Third Republic, these constituencies were strictly nested: cantons belonged to a single arrondissement, and arrondissements, to a single département. In each département, a prefect, operating from the chef-lieu, the prefecture, represented the State and supervised its administration within the département. Similarly, a subprefect held office in each arrondissement at the subprefecture. While départements and arrondissements were both administrative units and territorial constituencies, cantons were simple territorial divisions with no administrative prerogatives—except for local justices of the peace—that formed the basis of military and electoral constituencies. This organization had been in place since the National Constituent Assembly in 1789–1791[28] and was later consolidated by the law of August 10, 1871, which stabilized this framework and made départements the central constituency for France's territorial administration.[29] It was only essentially altered by two events: the annexation of Alsace-Lorraine in 1871 and its recovery in 1919, and the arrondissements reform of 1926.[30]

The annexation of Alsace-Lorraine in 1871 and its recovery in 1919 Following the Franco-Prussian War, the Frankfurt treaty of May 10, 1871 (complemented by the convention of October 12, 1871) stipulated the transfer of Alsace-Lorraine to the German Empire. Annexed territories consisted of the département of Bas-Rhin, most of the département of Haut-Rhin, parts of the départements of Moselle and Meurthe, and a few communes of the département of Vosges—1,694 communes in total. Supplementary Table 1 documents the composition of annexed territories along with supporting administrative sources. Boundaries of these territories did not overlap with pre-war limits of general administrative constituencies. As a result, it was necessary to reconfigure these constituencies within the territories of annexed départements that had remained in France: the remaining territories of the département of Haut-Rhin were grouped into the Territoire-de-Belfort and the département of Meurthe was renamed Meurthe-et-Moselle, absorbing the remaining territories of the département of Moselle. The reconfiguration of these territories' arrondissements and cantons between 1871 and 1873 is documented in Supplementary Table 2. Upon the official recovery of Alsace-Lorraine per the Versailles treaty of June 28, 1919, territories of this region kept their administrative structures defined in the aftermath of the Franco-Prussian War: the Territoire-de-Belfort and the département of Meurthe-et-Moselle remained unchanged, and the territories formerly in the départements of Bas-Rhin, Haut-Rhin, and Moselle recovered their pre-war structures, barring changes that occurred under German rule between 1871 and 1918—changes documented in Supplementary Table 3.

The arrondissements reform of 1926 Due to the deterioration of France's budgetary situation after World War I and the relative uselessness of the arrondissement as administrative constituency, the Poincaré government suppressed 106 arrondissements per the decree-law of September 10, 1926, bringing their number to 279.[31] Although four were re-established by 1940, the configuration of arrondissements defined in 1926 remained unchanged until the end of the Third Republic. Territorial changes induced by this reform are documented in Supplementary Table 4.

Overall, except for these two events and the creation of five dozen cantons (documented in Supplementary Table 5), general administrative constituencies remained relatively stable throughout the Third Republic.[32,33] Figure 7 displays the evolution of their number between 1870 and 1940: from 87 to 90 départements, from 279 to 385 arrondissements, and from 2,861 to 3,028 cantons. Figure 8 further displays the geographic configuration of these administrative constituencies at four key dates: 1870, 1871, 1919, and 1926.

Special administrative constituencies

Military constituencies[34,35] The territorial organization of the military prevalent during most of the Third Republic was based on a series of reforms passed in 1873–1874. At its core, the law of "general organization of the army" of July 24, 1873, ensured consistency between military recruitment and military command. It divided the territory into 18 military regions, each further divided into eight subdivisions per the decree of August 6, 1874. Moreover, per the decree of September 28, 1873, one army corps per military region was created. Each army corps was composed of two infantry divisions of two brigades with two regiments—one infantry regiment per subdivision of military region. In each subdivision of military region was located a recruitment bureau, which managed recruitment and mobilization under the authority of the military region command. Although they overlapped geographically, military regions were the relevant administrative structures for recruitment, while army corps were relevant for military command.

Except for the regions of Paris and Lyon, which had specific military governments, delineations of military regions and subdivisions followed canton and arrondissement boundaries so as to maintain balance in terms of population, transportation

networks, topography, and cultural homogeneity. This organization was not fundamentally altered until 1940, though it underwent a series of modifications. Until the early-1920s, the military adapted its territorial organization mostly along the north-eastern border in order to ensure efficient mobilization in case of military conflict with Germany. Among the main modifications, a 20th military region (Nancy) was created in 1898 through the splitting of the 6th military region (Châlons-sur-Marne) as well as a 21st military region (Épinal) by the end of 1913 through the splitting of the 7th military region (Besançon). Following its recovery in 1919, the military organization in Alsace-Lorraine followed the same logic. Modifications from the late 1920s onward were instead dominated by the willingness of the military command of nesting subdivisions of military regions into département boundaries, which was nearly achieved by the mid-1930s. At odds with this system, the military organization that prevailed before the reforms of 1873–1874 did not exhibit such consistency between military recruitment and military command. It simply divided troops into six army corps across 22 divisions along département boundaries—the territorial configuration of military divisions in 1870 is documented in Supplementary Table 6. Figure 9 displays the evolution of the number of military constituencies between 1870 and 1940: from 18 to 20 military regions, and from 100 to 156 subdivisions of military regions.

The construction of annual constituency-level shapefiles for military corps and divisions (1870–1873) and military regions and subdivisions (1874–1940) is based on their characterization in terms of general administrative constituencies. While they overlapped département boundaries between 1870 and 1873, military constituencies generally followed cantons and arrondissement boundaries between 1874 and 1940. For this period, I first match each canton from the TRF-GIS canton-level nomenclature 1874 to its military region and subdivision as per the decree of August 6, 1874. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and military constituencies—the 35 territorial modifications to military constituencies are documented in Supplementary Table 7.

Annual constituency-level shapefiles for military regions and subdivisions (1874–1940) exhibit the same limitations as canton-level shapefiles, as some subdivisions of military regions were defined across canton and commune boundaries. For instance, from 1874 onward, all but one subdivision of the 2nd military region (Amiens) were composed of fractions of the cantons of Saint-Denis and Pantin (Seine), as well as fractions of the 10th, 19th, and 20th municipal arrondissements of Paris. Following the same logic as for canton-level shapefiles, I create pseudo-military constituencies whenever relevant. These pseudo-constituencies were mostly present in the vicinity of the larger urban centers of Paris, Lyon, and Marseille. Figure 10 displays the geographic configuration of military constituencies at four key dates: 1870, 1874, 1919, and 1940.

Judicial and penitentiary constituencies[36] France's judicial organization remained broadly stable from the early nineteenth century to the end of the Third Republic. Ordinary judicial institutions consisted of five nested jurisdictional levels. At the first level, justices of the peace (*justices de paix*) and simple police courts (*tribunaux de simple police*) had cantons as constituency and were headed by the same cantonal judge. Their attributions were limited to minor civil cases that could be resolved by conciliation and to minor criminal offences. The second level was constituted by courts of first instance (*tribunaux de première instance*), the ordinary first-degree jurisdiction for civil cases, which also ruled on appeals to justices of the peace. In criminal matters, its criminal court (*tribunal correctionnel*) ruled on all offenses for which sentences were shorter than five years but exceeded those imposed by simple police courts. Until the decree of September 3, 1926, courts of first instance had arrondissements as constituency—except for the département of Seine, which had only one, and the arrondissement of Puget-Théniers, which had none. These courts were located at the subprefecture, barring a dozen exceptions. In the spirit of the arrondissements reform of 1926, the judicial reform of the same year suppressed 228 courts of first instance. However, this reform soon failed and all but six courts were reinstated in 1930 per the law of August 22, 1929—three courts were later suppressed in 1931–1932. At the third level, assize courts (*cours d'assises*) were autonomous départemental jurisdictions that ruled on crimes. They were only in session for one trimester each year except for the assize court of Paris, which was permanent. The fourth jurisdictional level was constituted by courts of appeal (*cours d'appel*), which ruled on appeals to courts of first instance, both in civil and criminal matters. There were 26–28 courts of appeal, each composed of one (Bastia) to seven (Paris) départements. Their delineations remained stable throughout the Third Republic barring modifications induced by the annexation and recovery of Alsace-Lorraine: the courts of appeal of Metz and Colmar were suppressed in 1871, the later being reinstated in 1919. Finally, the higher jurisdiction was the Court of Cassation (*Cour de Cassation*), which reviewed other courts' rulings in last resort, both in civil and criminal matters. Panels a and b of Figure 11 display the evolution of the number of judicial constituencies between 1870 and 1940: from 26 to 28 courts of appeal, and from 138 to 370 courts of first instance.

The construction of annual constituency-level shapefiles for courts of appeal is straightforward as these followed département boundaries. For courts of first instance, I rely on their delineations along canton boundaries—although courts of first instance were based on arrondissements until 1929, their delineations from 1930 to 1940 were based on the geography of arrondissements prior to the reform of 1926, which left cantons untouched. I first match each canton from the TRF-GIS canton-level nomenclature 1870 to its court of first instance. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and courts of first instance. Modifications that affected courts of appeal and courts of first instance, along with other details, are documented in Supplementary Table 8. Because justices of the peace, police courts, and assize courts had

cantons and départements as constituencies, their geographies and nomenclatures follow that of these general administrative constituencies. Figure 12 displays the geographic configuration of judicial constituencies at four key dates: 1870, 1871, 1926, and 1940.

Besides ordinary judicial institutions, special jurisdictions ruled over litigation specific to some professional groups, such as commercial courts (*tribunaux de commerce*) and labor courts (*conseils de prud'hommes*). Because these only existed in some cities and had heterogeneous jurisdictions, I do not provide nomenclatures and shapefiles for these special jurisdictions. Annual commune-specific nomenclatures do, however, detail their locations.

Finally, linked to the judicial system, the penitentiary system was administered through regional constituencies that comprised several départements from 1871 onward—in 1870, penitentiary constituencies and départements were confounded. Their number decreased over time, from 45 in 1871–1887 to 16 in 1926–1940 (Figure 11, Panel c). These modifications are documented in Supplementary Table 8. Figure 13 displays the geographic configuration of penitentiary constituencies at four key dates: 1871, 1888, 1909, and 1940.

Electoral constituencies[37,38] The Third Republic was a Parliamentarian regime in which the National Assembly, France's Lower House, elected the head of the executive branch (together with the Senate from 1875 onward). Members of the National Assembly, the *députés*, were elected by male citizens aged 21 and older through a single-member system, with one député per electoral constituency, denominated *circonscription*—though the general elections of 1871, 1885, 1919, and 1924 used a block-list system. While block-list elections generally had département-level electoral constituencies, arrondissements formed the basis of electoral constituencies throughout the Third Republic. Its electoral geography was based on the redistricting of 1875, which followed three principles that barely changed during the next seven decades: each arrondissement had at least one député, and therefore constituted at least one circonscription; each additional hundred thousand inhabitants entitled an arrondissement to an additional député; and a circonscription was modified only if its number of députés had to change due to population changes. There was only one major redistricting after 1875: that of 1927 after the return of the single-member system, which was broadly based on the geography of arrondissements prior to the Poincaré reform of 1926. The redistricting of 1889, also after the return to the single-member system, was essentially identical to that of 1875. Still, the population rule for redistricting entailed marginal changes after each census, adjusting constituencies in 10–20 arrondissement every five years—Supplementary Table 9 documents all 13 redistricting laws. Panel a of Figure 14 displays the evolution of the number of circonscriptions during the Third Republic: from 87 to 100 for block-list elections, and from 526 to 598 for single-member elections.

The construction of annual circonscription-level nomenclatures and shapefiles relies on two elements: TRF-GIS annual canton-level nomenclatures and Gaudillère's (1995) *Atlas Historique des Circonscriptions Électorales Françaises*, which provides maps of circonscriptions from 1815 to 1986 for each département, with cantons as underlying frame.[37] Using Gaudillère's 800-page atlas, I manually match each contemporaneous canton to its electoral constituency for each redistricting year. I then expand these configurations to the following years until the next redistricting, taking into account territorial changes to cantons through TRF-GIS annual canton-level nomenclatures. I use the same methodology but at the levels of communes for Seine-et-Oise and municipal arrondissements for Seine, as circonscriptions were generally defined at such levels in these two départements.

Annual circonscription-level shapefiles exhibit the same limitations as canton-level shapefiles, as some circonscriptions were defined across canton and commune boundaries due to the population rule for redistricting. For instance, the city of Montluçon was divided along canton boundaries into the circonscriptions of Montluçon-1 (canton of Montluçon-Est) and Montluçon-2 (canton of Montluçon-Ouest) during single-member elections. Following the same logic as with canton-level shapefiles, I create pseudo-circonscriptions whenever relevant. Figure 15 displays the geographic configuration of electoral constituencies at four key general elections dates: 1876, 1914, 1919, and 1928.

Academic constituencies[39] The administration of public education during the Third Republic was operated through a system of regional academies that was defined in 1854 and that overlapped département boundaries. Since then, each academy was administered by a rector who supervised primary, secondary, and superior education within the academy. He was assisted by several academic inspectors—one per département, two or more in the populous départements of Nord and Seine. The territorial configuration of the 17 academies that existed by 1870 was seldom modified during the Third Republic: in 1871, the Territoire-de-Belfort was transferred to the academy of Besançon; in 1888, the chef-lieu of the academy of Douai was transferred to Lille; in 1919, the recovered département of Moselle was transferred to the academy of Strasbourg; and in 1920, the academy of Chambéry was suppressed and its départements (Savoie and Haute-Savoie) transferred to the academy of Grenoble. Panel b of Figure 14 displays the evolution of the number of academies during the Third Republic, which remained at 16–17.

The construction of annual academy-level shapefiles is straightforward as academies were supra-départemental constituencies that underwent little modifications. I first match each département from the TRF-GIS département-level nomenclature

1870 to its academy per the decrees of August 22nd, 1854, and June 13th, 1860. Then, for each year onward, I update this initial configuration with territorial modifications to both départements and academies—the few territorial modifications to academies are documented in Supplementary Table 10. Panel a of Figure 16 displays the geographic configuration of academic constituencies in 1919.

Labor inspection constituencies[40] The service of labor inspection of the Ministry of Labor was created by the law of May 19th, 1874, which objective was to regulate child labor. However, it was not until the law of December 13th, 1892 that a service of labor inspectors structured the territory in order to control the application of labor regulations. It was initially organized into 11 supra-départemental constituencies, each headed by one divisional inspector. Divisional inspectors were assisted by several départemental inspectors operating in territorial sections—92 in total. Labor inspection constituencies underwent little modifications after 1892: sieges of the 2nd and 4th constituencies were respectively transferred from Châteauroux to Tours and from Bar-le-Duc to Nancy in 1893; the département of Isère, initially split between the 10th and 11th constituency, was reunited into the 11th constituency in 1902; and the département of Somme was transferred from the 5th to the 6th constituency in 1911, then back to the 5th constituency in 1919. After World War I, a 12th constituency was created with Strasbourg as chef-lieu in order to structure labor inspections in the départements of Moselle, Bas-Rhin, and Haut-Rhin. Panel c of Figure 14 displays the evolution of the number of labor inspection constituencies during the Third Republic, which remained at 11–12.

To construct annual labor inspection-level shapefiles, I first match each canton from the TRF-GIS canton-level nomenclature 1892 to its labor inspection constituency per the law of December 13th, 1892—although these constituencies were supra-départemental, cantons of the département of Isère were initially split between the 10th and 11th constituency. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and labor inspection constituencies. Panel b of Figure 16 displays the geographic configuration of labor inspection constituencies in 1919.

Ecclesiastical constituencies[41] Despite the Separation of the Churches and the State in 1905, catholic ecclesiastical constituencies—dioceses—continued to play an important role in the French society, and to produce statistics.[42] Inherited from the geography of roman provinces, most dioceses were created between the second and the fifth centuries. Their delineations were substantially altered twice after the Revolution: by the Concordat of 1801, which broadly nested the diocesan geography into that of the newly created départements, and by the Papal bull *Paternae charitatis* of 1822, which restored some prerevolutionary dioceses. By 1870, the metropolitan territory was divided into 86 dioceses, among which 17 archbishoprics (*archevêchés*) and 69 bishoprics (*évêchés*). Ruled by an archbishop, archbishoprics were metropolitan sees and headed the territories of several subordinate bishoprics. Some episcopal sees were immediately subject to the Holy See—the dioceses of Metz and Strasbourg after 1919. Throughout the Third Republic, the territories of most dioceses followed département boundaries. Still, some followed arrondissements: the diocese of Marseille, which corresponded to the arrondissement of Marseille; the diocese of Reims, which corresponded to the département of Ardennes and the arrondissement of Reims (Marne); the diocese of Fréjus, which corresponded to the département of Var and the arrondissement of Grasse (Alpes-Maritimes) until 1886: and the dioceses of Lille and Cambrai, which split the département of Nord after 1913. The diocese of Carcassonne (Aude) even had an enclave in the département of Ariège through the canton of Quérigut. The diocesan geography of the départements of Savoie and Haute-Savoie was more complex, as it divided their territories along commune boundaries—this configuration is documented in Supplementary Table 11.[43] Diocese boundaries underwent little changes during the Third Republic: beyond those described above, the Territoire-de-Belfort was transferred from the diocese of Strasbourg to that of Besançon after the Franco-Prussian War, and the territories of the département of Meurthe-and-Moselle were attributed to the diocese of Nancy—Supplementary Table 12 documents these changes. Panel d of Figure 14 displays the evolution of the number of dioceses between 1870 and 1940, which remained between 84 and 87.

To construct annual diocese-level shapefiles, I first match each canton from the TRF-GIS canton-level nomenclature 1870 to its diocese as per the Papal bull of 1822, taking into account the creation of the diocese of Laval in 1855. Then, for each year onward, I update this initial configuration with territorial modifications to both cantons and dioceses. I use the same methodology but at the level of communes for the départements of Savoie and Haute-Savoie. Panel c of Figure 16 displays the geographic configuration of dioceses in 1919, together with the extent of metropolitan bishoprics.

Data Records

The TRF-GIS database is composed of two types of data files: annual constituency-level nomenclatures and annual constituency-level shapefiles. A description of each these data files' names, content, formats, sizes, and locations is available in Table 2 for general administrative constituencies, in Table 3 for military constituencies, in Table 4 for judicial and penitentiary constituencies, and in Table 5 for electoral, academic, labor inspection, and ecclesiastical constituencies. Table 5 also describes data files for commune-level nomenclatures. Annual constituency- and commune-level nomenclatures are available in Stata data format (`.dta`) as well as in text delimited format (`.txt`), with the main file-specific metadata replicated as `notes` within `.dta` files and along with separate codebooks for `.txt` files. Each annual constituency-level shapefile is composed of

a (zipped) set of five files: a shape format file (`.sph`), a shape index format file (`.shx`), an attribute format file (`.dbf`), a projection description file (`.prj`), and a character encoding file (`.cpg`). These shapefiles use an RGF93 Lambert-93 projection system, IGN's reference projection system. I further follow IGN's methodology and provide shapefiles both in polygon and vector forms. Shapefiles metadata are also replicated for each constituency in QGIS metadata format (`.qmd`). Furthermore, separate datasets of chefs-lieux coordinates are provided in `.csv` format—these coordinates correspond to the current locations of their city halls.

The TRF-GIS follows FAIR data management practices.[14] Data files are available in the TRF-GIS Dataverse, hosted by the Harvard Dataverse, under a CC-BY 4.0 licence at https://dataverse.harvard.edu/dataverse/TRF-GIS, in which each constituency constitutes a separate dataset—15 datasets in total. Each dataset and datafile is attributed a unique and persistent Digital Object Identifier (DOI) for better findability and citability. Data files are curated through DataCite Metadata Schema 4.3 and use the Library of Congress Subject Heading (LCSH) controlled vocabulary for topic classification.[44]

General administrative constituencies

Départements Annual département-level nomenclatures are composed of 10 variables (Table 6). They include départements' codes, names, and whether they were composed of enclaved territories and islands (e.g., the Enclave des Papes of the département of Vaucluse, enclaved in the territory of the département of Drôme). They also include information on départements' chefs-lieux (prefectures): their INSEE codes, names, whether they are current or former communes (e.g., the prefecture of the département of Ardennes was located in the former commune of Mézières), and their geographic coordinates. For départements codification, I give number `99` to the département of Meurthe in 1870—codifications and toponymies are otherwise standard. Note that although the TRF-GIS treats the Territoire-de-Belfort as a département, it did not achieve this status officially until 1922. Annual département-level shapefiles contain the same variables as département-level nomenclatures.

Arrondissements Annual arrondissement-level nomenclatures are composed of 15 variables (Table 6). They include arrondissements' département code and name as well as arrondissements' codes and names. They also include information on arrondissements' chefs-lieux (subprefectures): their INSEE codes, names, whether they are current or former communes, whether they were common to other arrondissements (e.g., Strasbourg was the subprefecture of the arrondissements of Strasbourg-Campagne and of Strasbourg-Ville), whether they were located outside of their arrondissement (e.g., Strasbourg was located outside of the arrondissement of Strasbourg-Campagne), and their geographic coordinates. Arrondissements one-digit codification uniquely identifies arrondissements within départements and follows the COG 2011. For arrondissements that no longer existed by then, I assign a higher number in ascending alphabetical order within each département and keep this codification constant so that arrondissements can be uniquely identified across time. Moreover, I keep the contemporaneous toponymy of arrondissements, which names generally were that of their subprefectures—Supplementary Table 13 documents the ten toponymic modifications to arrondissements that occurred during the Third Republic. Note that although the TRF-GIS treats municipal arrondissements of Paris and the Territoire-de-Belfort as arrondissements, these territories did not have this status officially. Annual arrondissement-level shapefiles contain the same variables as arrondissement-level nomenclatures.

Cantons Annual canton-level nomenclatures are composed of 18 variables (Table 6). They include cantons' département and arrondissement codes and names as well as cantons' codes, names, and communal composition. The later variable, which follows INSEE's COGs codification convention, indicates whether a canton was composed of entire communes and/or of fractions of communes (e.g., the canton of Montluçon-Est was composed of eight entire communes and a fraction of the commune of Montluçon). They also include information on cantons' chefs-lieux (*bureaux centralisateurs*): their INSEE codes, names, whether they are current or former communes, whether they were common to other cantons (e.g., Montluçon was the bureau centralisateur of both cantons of Montluçon-Est and Montluçon-Ouest), and their geographic coordinates. Cantons two-digit codification uniquely identifies cantons within départements and follows the COG 2011. For cantons that no longer existed by then, I use the codification in the COG 1954. For those that no longer existed by 1954, I assign a higher number in ascending alphabetical order within each département and keep this codification constant so that cantons can be uniquely identified across time. Moreover, I keep the contemporaneous toponymy of cantons, which names generally included that of their bureau centralisateur—Supplementary Table 14 documents the 206 toponymic modifications to cantons that occurred during the Third Republic. Note that although the TRF-GIS treats municipal arrondissements of Paris as cantons, these territories did not have this status officially. Moreover, the communes of partially annexed cantons that had remained in France after the Franco-Prussian War were not re-assigned to functioning cantons until 1873. As a result, their chefs-lieux was (virtually) outside of France in 1871 and 1872. This was the case for 36 communes of the annexed cantons of Gorze (Moselle), of Vic-sur-Seille and Lorquin (Meurthe), and of Saales and Schirmeck (Vosges).

Because the commune-level GEOFLA 2011 shapefile forms the base frame for annual canton-level shapefiles, and cantons were sometimes defined across commune boundaries, these shapefiles do not entirely match canton-level nomenclatures. Instead, they use the concept of pseudo-cantons. Following INSEE's and IGN's codification conventions, pseudo-cantons are

assigned codes greater than 80 in ascending alphabetical order within départements. Moreover, their toponymy corresponds to the common city infra-communal cantons are aggregated into. For instance, in the case of the Montluçon area displayed in Figure 6, canton-level shapefiles contain the pseudo-canton of Montluçon, which extent corresponds to the commune of Montluçon and which is given pseudo-canton code 80. During the Third Republic, 10–12 percent of cantons were composed of fractions of communes, which translated into 110–120 pseudo-cantons (Figure 7, Panel c). Supplementary Table 15 documents the mapping between cantons and pseudo-cantons—this mapping is also available as a dataset alongside annual canton-level nomenclatures. Annual canton-level shapefiles contain the same variables as canon-level nomenclatures, except that their base units are pseudo-cantons rather than cantons. To reflect this difference, variables prefix in these shapefiles is pct rather than ct.

Special administrative constituencies

Military constituencies (1870–1873) Annual military corps-level nomenclatures are composed of 10 variables (Table 7). They include military corps' codes, names, and composition (entire départements). They also include information on military corps' chefs-lieux: their INSEE codes, names, whether they are current or former communes, and their geographic coordinates. Codifications and toponymies are contemporaneous, with toponymies corresponding to military corps chefs-lieux. Annual military corps-level contain the same variables as military corps-level nomenclatures.

Annual military division-level nomenclatures are composed of 14 variables (Table 7). They include military divisions' corps code and name as well as military divisions' codes, names, and composition (entire départements). They also include information on military divisions' chefs-lieux: their INSEE codes, names, whether they are current or former communes, and their geographic coordinates. Codifications and toponymies are contemporaneous, with toponymies corresponding to military divisions chefs-lieux. Annual military divisions-level shapefiles contain the same variables as military divisions-level nomenclatures.

Military constituencies (1874–1940) Annual military region-level nomenclatures are composed of 10 variables (Table 7). They include military regions' codes, names, and composition. For instance, throughout the Third Republic, the 2nd military region (Amiens) was composed of fractions of arrondissements and of entire cantons. They also include information on military regions' chefs-lieux: their INSEE codes, names, whether they are current or former communes, and their geographic coordinates. Codifications and toponymies are contemporaneous, with toponymies corresponding to military regions chefs-lieux. Because the commune-level GEOFLA 2011 shapefile forms the base frame for military region-level shapefiles, and military regions were sometimes defined across cantons and commune boundaries, these shapefiles do not entirely correspond to military region-level nomenclatures. Instead, these shapefiles use the concept of pseudo-military regions. These are assigned military region codes greater than 90 in ascending alphabetical order, and their names correspond to their territorial composition. For instance, the city of Lyon was divided between the military regions of Grenoble and Besançon from 1874 to 1921. This pseudo-military region is assigned pseudo-region code 90 and name "Grenoble - Besançon." Pseudo-military regions are not assigned a chef-lieu, so all variables with prefix cl are empty for actual pseudo-military regions. Supplementary Table 16 documents the mapping between pseudo-cantons and pseudo-military regions—this mapping is also available as a dataset alongside annual military region-level nomenclatures. Annual military region-level shapefiles contain the same variables as military region-level nomenclatures, except that their base units are pseudo-military regions rather than military regions. To reflect this difference, variables prefix in these shapefiles is pmreg rather than mreg. In addition, the variable pmreg_type indicates whether military regions were effectively pseudo-military regions.

Annual military subdivision-level nomenclatures are composed of 15 variables (Table 7). They include military subdivisions' region code and name as well as military subdivisions' codes, names, and composition. They also include information on military subdivisions' chefs-lieux: their INSEE codes, names, whether they are current or former communes, whether they were common to other subdivisions (e.g., Rouen was the chef-lieu of both military subdivisions of Rouen-Nord and Rouen-Sud), and their geographic coordinates. Codifications and toponymies are contemporaneous, with toponymies corresponding to military subdivisions chefs-lieux. Because the commune-level GEOFLA 2011 shapefile forms the base frame for military subdivision-level shapefiles, and military subdivisions were sometimes defined across cantons and commune boundaries, these shapefiles do not entirely correspond to military subdivision-level nomenclatures. Instead, these shapefiles use the concept of pseudo-military subdivisions. These are assigned military subdivision codes greater than 10 in ascending alphabetical order, and their names correspond to their territorial composition. For instance, the city of Rouen was divided between the military subdivisions of Rouen-Nord and Rouen-Sud from 1874 to 1929. This pseudo-military region is assigned code 11 and name "Rouen (Nord - Sud)." Pseudo-military subdivisions are not assigned a chef-lieu, so all variables with prefix cl are empty for actual pseudo-military subdivisions. Supplementary Table 17 documents the mapping between pseudo-cantons and pseudo-military subdivisions—this mapping is also available as a dataset alongside annual military subdivision-level nomenclatures. Annual military subdivision-level shapefiles contain the same variables as military subdivision-level nomenclatures, except that their base units are pseudo-military subdivisions rather than military subdivisions. To reflect this difference, variables prefix in

these shapefiles is `pmsub` rather than `msub`. In addition, the variable `pmsub_type` indicates whether military subdivisions were effectively pseudo-military subdivisions.

Judicial and penitentiary constituencies Annual court of appeal-level nomenclatures are composed of 9 variables (Table 8). They include courts of appeal's codes and names, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Toponymies are contemporaneous and correspond to courts of appeal chefs-lieux. Courts of appeal two-digit codification uniquely identifies these constituencies over time and is assigned in ascending alphabetical order. Annual court of appeal-level shapefiles contain the same variables as court of appeal-level nomenclatures.

Annual court of first instance-level nomenclatures are composed of 14 variables (Table 8). They include their court of appeal's code and name, as well as courts of first instances' codes and names. They also include information on courts of first instances' chefs-lieux: their INSEE codes, names, whether they were also a subprefecture, and their geographic coordinates. Toponymies are contemporaneous and correspond to courts of first instance chefs-lieux. Courts of first instance two-digit codification uniquely identifies these constituencies over time within their court of appeal and is assigned in ascending alphabetical order. Annual court of first instance-level shapefiles contain the same variables as court of first instance-level nomenclatures.

Annual penitentiary constituency-level nomenclatures are composed of 9 variables (Table 8). They include penitentiary constituencies' codes and names, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Toponymies are contemporaneous and correspond to penitentiary constituencies' chefs-lieux. Penitentiary constituencies two-digit codification uniquely identifies these constituencies over time and is based on the contemporaneous codification defined in the official act of 1871 reorganizing penitentiary constituencies—for 1870, these correspond to département codes; penitentiary constituencies created after 1871 are assigned codes that are ascending in the temporal order of their creation. Annual penitentiary constituency-level shapefiles contain the same variables as penitentiary constituency-level nomenclatures.

Electoral constituencies Annual circonscription-level nomenclatures are composed of 6 variables (Table 9). They include circonscriptions' département code as well as circonscriptions' codes, names, and composition. The later variable indicates whether a circonscription was composed of entire départements, arrondissements, cantons, or communes. Some codes specifically refer to the circonscriptions of Paris. Contrary to other constituencies, electoral constituencies are uniquely identified by their name, which corresponds to their arrondissement name, with the addition of a number for the larger urban centers, or, between 1928 and 1940, to the most important city of the constituency—sometimes two adjoined names, such as "Nantua - Gex." For convenience, I provide a codification in ascending alphabetical order within each département, but it is not necessarily consistent over time due to the changing nature of circonscriptions' boundaries. Likewise, urban circonscriptions which contemporaneous names were composed of the arrondissement name and a number, such as "Saint-Denis-3" in Seine-et-Oise, might not be consistent over time—this is generally the case in the more urban areas in which redistricting was frequent. Moreover, no chefs-lieux are indicated as electoral circonscriptions were not administrative units, but only territorial constituencies—electoral operations were instead coordinated at cantons' bureaux centralisateurs. Each annual circonscription-level nomenclature further indicates three elements among its metadata: the election system (block-list or single-member), whether it was a general election year, and the relevant redistricting year. Annual circonscription-level shapefiles contain the same variables as circonscription-level nomenclatures, except that their base units are pseudo-circonscriptions rather than circonscriptions. To reflect this difference, variables prefix in these shapefiles is `pcirco` rather than `circo`. In addition, the variable `pcirco_type` indicates whether circonscriptions were effectively pseudo-circonscriptions. Note that because shapefiles are stored in associated `.dbf` files, attribute names cannot be more than 10 characters long. Consequently, two field names are abbreviated: `pcrico_name` into `pcrico_nam` and `pcirco_type` into `pcirco_typ`.

Academic constituencies Annual academy-level nomenclatures are composed of 8 variables (Table 9). They include academies' codes and names, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Toponymies are contemporaneous and correspond to academies chefs-lieux. Academies two-digit codification uniquely identifies these constituencies over time and is assigned in ascending alphabetical order—except for the academy of Chambéry, created only in 1860 and suppressed in 1919, which holds academy code `17`. Annual academy-level shapefiles contain the same variables as academy-level nomenclatures.

Labor inspection constituencies (1892–1940) Annual labor inspection-level nomenclatures are composed of 6 variables (Table 9). They include labor inspections codes, as well as information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. Labor inspection constituencies did not have a specific toponymy beyond their 2-digit codification, which was fixed over time. Annual labor inspection-level shapefiles contain the same variables as labor inspection-level nomenclatures.

Ecclesiastical constituencies Annual diocese-level nomenclatures are composed of 14 variables (Table 9). They include the archbishopric's code and name each diocese was attached to, its type (metropolitan see or exempt), as well as dioceses codes, names, and types (archbishoprics or bishoprics). They further include information on their chefs-lieux: their INSEE codes, names, and geographic coordinates. The diocesan toponymy is contemporaneous and corresponds to customary dioceses names, which consisted in the locations of their chefs-lieux (episcopal sees)—except for the diocese of Aire, which episcopal see was transferred to Dax in 1933, and for the diocese of Tarentaise, which name referred the Tarentaise Valley. Different from their customary names, dioceses' formal names were composed of all prerevolutionary dioceses their bishops were entitled. For instance, by 1854, the bishop of Bayeux was entitled the former diocese of Lisieux. As a result, the formal name of the diocese of Bayeux became "Bayeux-Lisieux." These names were sometimes cumbersome: for instance, by 1877, the formal name of the diocese of Avignon was "Avignon-Apt-Cavaillon-Carpentras-Orange-Vaison." Annual diocese-level nomenclatures therefore use dioceses customary names. The formal diocesan toponymy is documented in Supplementary Table 19. Dioceses two-digit codification uniquely identifies these constituencies over time and are assigned in ascending alphabetical order—except for the archbishopric of the dioceses of Metz and Strasbourg after 1919, exempt and subject to the Holy See, which is assigned code `99`. Annual diocese-level shapefiles contain the same variables as diocese-level nomenclatures.

Annual commune-level nomenclatures

Annual commune-level nomenclatures are composed of 69 variables (Table 10). Herein, each commune can be identified through its unique Cassini code from the HAC database as well as through its INSEE code in 2005 geography, which might not be unique if the commune was absorbed by another before 2005—a variable indicates whether the commune existed by then. Communes names are contemporaneous. These nomenclatures include the names and codes of all the above general and special administrative constituencies each commune belonged to in a given year—pseudo-constituencies in the cases of cantons, military regions and subdivisions, and electoral circonscriptions. They also include indicator variables for whether a commune was the chef-lieu of a constituency or the siege of a special jurisdiction (a commercial or labor court).[36] Table 10 describes the 21 additional variables included in annual commune-level nomenclatures that are not present in other constituency-level nomenclatures—the 48 other variables described in Tables 6–9 are not present in Table 10.

To enable users to combine statistics produced by administrations at different and potentially ex-ante incompatible levels of aggregation, commune-level nomenclatures further provide municipal population counts from the HAC database, which reports this information based on population censuses. During the Third Republic, censuses were carried out every five years—except in 1871, when it was postponed to 1872 due to the Franco-Prussian War, and in 1916, when it was cancelled due to the First World War. Three variables relate to population counts: `pop`, the raw municipal population count, `ipop`, the interpolated municipal population count, and `pop_flag`, a municipal population count flag. The `pop` variable is missing for all communes outside of census years (extended missing value `.a`) and for a few communes during census years: when a commune was created during the same year as a census but after it was carried out (`.b`), when the census archive was lost (`.c`) or unreadable (`.d`), and when the commune was not surveyed (`.e`). Except in 1872, when 233 communes were not surveyed (nearly exclusively in the département of Hautes-Pyrénées) and the archives for another 482 communes were lost (exclusively in the département of Vosges), missing population counts during census years remain negligible: they always account for less than 15 communes—less than 3 communes in 9 out of all 13 censuses that were carried out during the Third Republic. Based on these raw population counts, the variable `ipop` provides interpolated municipal population counts for all years and all communes present in annual commune-level nomenclatures. The variable `pop_flag` documents which imputation method is used: none for census years (`1`), linear interpolation for inter-census years (`2`), linear interpolation for census years when the census archive was lost (`3`) or unreadable (`4`), linear interpolation for census years when a commune was not surveyed (`5`), and imputation from the previous (`6`) or next (`7`) census for communes that were created or suppressed during an inter-census year. Imputations remain limited to about 30 communes across all years after 1872, and to about a hundred in 1870–1871.

Finally, annual commune-level nomenclatures have a specific treatment of Paris: it provides both arrondissement- and commune-level information. Parisian arrondissements hold INSEE codes `75101`–`75120`, and the commune of Paris, `75056`. Whenever Paris is the chef-lieu of a constituency, that title falls onto the first arrondissement as well as onto the commune. Whenever it is subdivided into constituencies, the commune of Paris holds the status of pseudo-constituency with corresponding codifications: it holds pseudo-canton code `7580`, pseudo-military region and subdivision code `99`, and pseudo-circonscription code `99`.

Technical Validation

TRF-GIS nomenclatures

Information supporting TRF-GIS nomenclatures documented in secondary sources (the HAC database and other secondary sources[30–41,43]) were thoroughly verified against primary sources. Supplementary Table 20 details the content of each of the 175 primary sources used for this verification procedure: the constituency they are relevant to, their codification (date of official

issue and type), their title, their persistent identifier (PID) assigned by Gallica and the National Library of France (BNF) in archival resource key (ARK) form, and their origin. Primary sources' types take the following forms: law (L), decree (D), order (A, for *arrêté*), ordinance (O), and papal bull (B). Their origins is generally the *Journal Officiel*, though sometimes it is the *Bulletin des Lois*, the *Code Pénitentaire* or *Code des Prisons*, the *Bulletin Administratif de l'Instruction Publique*, and the *Acta Apostolicae Sedis*.[23,24,45–47] All these primary sources are available in PDF format in the repository. This rigorous verification procedure combined with data integrity tests in the TRF-GIS source code enabled to identify 231 inaccuracies or errors in the HAC database and to rectify them. These are documented in Supplementary Table 21.

Nomenclatures were also verified against periodical statistical publications of relevant administrations, among which censuses for general administrative constituencies, the *Compte Rendu sur le Recrutement de l'Armée* for military constituencies, the *Compte Général de l'Administration de la Justice Civile et Commerciale* and the *Compte Général de l'Administration de la Justice Criminelle* for judicial constituencies, the *Statistique Pénitentiaire* for penitentiary constituencies, or the *Tableau des Élections à la Chambre des Députés* for electoral constituencies.[48–51] Useful to interested users, an up-to-date listing of the availability of these periodical statistical publications is available at progedo.hypotheses.org/514.[52]

TRF-GIS shapefiles

IGN's GEOFLA 2011 shapefile constitutes the underlying frame of TRF-GIS shapefiles. As discussed above, I use this dataset for quality and reproducibility purposes: it constitutes the official representation of France's communes in 2011 geography and is available under open licensing.[25] Although IGN's BD CARTO® shapefiles are far more precise than the GEOFLA (1:25,000 against 1:1,000,000), these files are not all freely accessible. Their use would therefore substantially alter the possibilities of reproduction and dissemination of TRF-GIS shapefiles. Moreover, their manipulation is more cumbersome: while GEOFLA is available as an ensemble, BD CARTO files are only available separately for each département, and their overall size approaches 2.5 Go against 150 Mo for the GEOFLA. In the end, the lower precision of the GEOFLA need not be very problematic because TRF-GIS shapefiles aim at regional and national extents (see the 'Limitations' section).

Moreover, the general methodology used to construct TRF-GIS shapefiles offers several advantages over the manual vectorization of georeferenced historical maps.[7,13,16,17] First, it yields more precise results than existing outputs, such as LARHRA's cantons shapefiles for 1884 and 1925: historical maps with national extents oftentimes lacked precision and had unspecified projection systems, making the resulting georeferencing potentially approximate, and urban centers, Corsica, and islands challenging to vectorize—historical cantons maps of 1884 and 1925 have a precision of 1:1,250,000 and 1:1,600,000, respectively.[53,54] Inaccuracies that result from these methods are displayed in Figure 17, which plots three randomly selected cantons from TRF-GIS and LARHRA cantons shapefiles 1884. Herein, we can see that five out of twenty communes (which territories did not change historically) are classified in the wrong canton by LARHRA's shapefile—Plougar (Plouvézédé instead of Plouescat), Plougourvest (Plouvézédé instead of Landivisiau), Sibiril (Plouvézédé instead of Saint-Pol-de-Léon), Plouvorn (Taulé instead of Plouvézédé), and potentially Guiclan (Saint-Thégonnec instead of Taulé). Moreover, the commune of Ile-de-Batz (canton of Saint-Pol-de-Léon) is not represented by LARHRA's shapefile. Second, and more importantly, historical maps are generally not available for every year for general administrative constituencies, and are even more rare for special administrative constituencies, which implies that relying uniquely on the vectorization of georeferenced historical maps would leave substantial gaps in a historical GIS of France. Finally, manual vectorization methods are quite time- and resource-consuming, making it very challenging to build annual shapefiles with realistic resources and to verify their accuracy.

Usage Notes

The TRF-GIS database can be accessed through the TRF-GIS Dataverse under a CC-BY 4.0 licence, which is hosted in the Harvard Dataverse and accessible at https://dataverse.harvard.edu/dataverse/TRF-GIS. The reproduction material is available in the "TRF-GIS Reproduction Material" dataset at doi.org/10.7910/DVN/6FCK3W. It contains source data, reproduction codes, and primary archival sources. Although each file can be accessed independently, I recommend downloading the full dataset for reproducibility purposes in order to preserve the integrity of the folders structure.

Code availability

The TRF-GIS database was produced using Stata SE version 16 as well as QGIS version 3.12 for its shapefiles.[55,56] Stata `.do` files and python `.py` files needed to reproduce TRF-GIS data files are available in the "TRF-GIS Reproduction Material" dataset under the MIT open licence. Using a standard processor of 4.00GHz with 16 GB RAM, the total processing time of the master Stata `.do` file is 7 hours, and 8 hours for the python `.py` files constructing shapefiles.

References

1. Weber, E. *Peasants into Frenchmen: The Modernization of Rural France, 1870–1914* (Stanford University Press, Stanford, 1976).
2. Squicciarini, M. Devotion and Development: Religiosity, Education, and Economic Progress in 19th-century France. *American Economic Review* https://cepr.org/active/publications/discussion_papers/dp.php?dpno=13877 (forthcoming).
3. Abramitzky, R., Delavande, A. & Vasconcelos, L. Marrying Up: The Role of Sex Ratio in Assortative Matching. *Am. Econ. Journal: Appl. Econ.* **3**, 124–57, https://www.jstor.org/stable/41288641 (2011).
4. Boehnke, J. & Gay, V. The Missing Men: World War I and Female Labor Force Participation. *Journal of Human Resources* **57**, https://doi.org/10.3368/jhr.57.4.0419-10151R1 (2022).
5. Desrosières, A. *The Politics of Large Numbers. A History of Statistical Reasoning.* (Harvard University Press, Cambridge, 1998 [1993]).
6. Ruggles, S. & Magnuson, D. L. The History of Quantification in History: The JIH as a Case Study. *J. Interdiscip. Hist.* **50**, 363–81, https://doi.org/10.1162/jinh_a_01446 (2019).
7. Ostafin, K., Kaim, D., Siwek, T. & Miklar, A. Historical Dataset of Administrative Units with Social-Economic Attributes for Austrian Silesia 1837–1910. *Scientific Data* **7**, 1–14, https://doi.org/10.1038/s41597-020-0546-z (2020).
8. Manson, S., Schroeder, J., Van Riper, D., Kugler, T. & Ruggles, S. IPUMS National Historical Geographic Information System: Version 15.0 [dataset], http://doi.org/10.18128/D050.V15.0 (2020).
9. Gregory, I. N., Bennett, C., Gilham, V. L. & Southall, H. R. The Great Britain Historical GIS Project: From Maps to Changing Human Geography. *The Cartogr. J.* **39**, 37–49, https://doi.org/10.1179/caj.2002.39.1.37 (2002).
10. Kunz, A. & Boehler, W. HGIS Germany: An Information System on German States and Territories from 1820 to 1914. *Hist. Geogr.* **33**, 145–7, https://ejournals.unm.edu/index.php/historicalgeography/article/view/2937/2416 (2005).
11. Huber, M. Quarante Années de la Statistique générale de la France 1896–1936. *J. de la Société Française de Stat.* **78**, 179–214, http://www.numdam.org/item/JSFS_1937__78__179_0/ (1937).
12. Ruggles, S. The Importance of Data Curation. In *The Palgrave Handbook of Survey Research*, 303–308 (Palgrave Macmillan, London, 2018).
13. Laboratoire de Recherche Historique Rhône-Alpes. Les Cantons Français de 1884 à 1966, http://geo-larhra.ish-lyon.cnrs.fr/?q=atlas-historique/territoires-d-etat/evolution-des-cantons-en-france (2011).
14. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **3**, 1–9, https://doi.org/10.1038/sdata.2016.18 (2016).
15. Christensen, G. & Miguel, E. Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature* **56**, 920–80, https://doi.org/10.1257/jel.20171350 (2018).
16. Perret, J., Gribaudi, P. & Berthelemy, M. Roads and Cities of 18th Century France. *Scientific Data* **2**, 1–7, https://doi.org/10.1038/sdata.2015.48 (2015).
17. Hedefalk, F., Svensson, P. & Harrie, L. Spatiotemporal Historical Datasets at Micro-Level for Geocoded Individuals in Five Swedish Parishes, 1813–1914. *Scientific Data* **4**, 1–13, https://doi.org/10.1038/sdata.2017.46 (2017).
18. Lang, G. Le Code Officiel Géographique (COG) Avant, Pendant et Autour. Tech. Rep., INSEE (2016). http://projetbabel.org/gl/cog_2016.pdf.
19. INSEE. Code Officiel Géographique au 1er Janvier 2005, https://www.insee.fr/fr/information/2560656 (2005).
20. INSEE. Code Officiel Géographique au 1er Janvier 2011, https://www.insee.fr/fr/information/2560625 (2015).
21. Motte, C., Séguy, I. & Théré, C. *Communes d'Hier, Communes d'Aujourd'hui. Les Communes de la France Métropolitaine, 1801–2001. Dictionnaire d'Histoire Administrative* (INED, Paris, 2003).
22. Motte, C. & Vouloir, M.-C. Le Site `cassini.ehess.fr`: Un Instrument d'Observation pour une Analyse du Peuplement. *Bull. du Comité Français de Cartogr.* 68–84, http://lecfc.fr/new/articles/191-article-7.pdf (2007).
23. France. Journal Officiel de la République Française. Lois et Décrets. 1880–1949, https://gallica.bnf.fr/ark:/12148/cb34378481r/date (2015).
24. France. Bulletin des Lois de la République Française. 1794–1931, https://gallica.bnf.fr/ark:/12148/cb32726274t/date (2019).

25. IGN. Geofla®Communes Édition 2011 France Métropolitaine, https://geoservices.ign.fr/documentation/diffusion/telechargement-donnees-libres.html (2013).

26. French Historical GIS, 1700–2020. Administrative units, Populations, Transports, Economy. (forthcoming 2022).

27. Noizet, H., Bove, B. & Costa, L. (eds.) *Paris de Parcelles en Pixels. Analyse Géomatique de l'Espace Parisien Médiéval et Moderne* (Presses Universitaires de Vincennes-Comité d'Histoire de la Ville de Paris, Paris, 2013).

28. Ozouf-Marignier, M.-V. *La Formation des Départements. La Représentation du Territoire Français à la Fin du 18ème Siècle* (Éditions de l'École des Hautes Etudes en Sciences Sociales, Paris, 1989).

29. Gros, D. *Naissance de la Troisième République*, 307–35 (Presses Universitaires de France, Paris, 2014).

30. Masson, J.-L. *Provinces, Départements, Régions. L'Organisation Administrative de la France d'Hier à Demain*, 358–99 (Fernand Lanore, Paris, 1984).

31. Verdier, N. La Réforme des Arrondissements de 1926 : Un Choix d'Intervention entre Espaces et Territoires. In Allorant, P. (ed.) *Les Territoires de l'Administration : Départir, Décentraliser, Déconcentrer*, 107–22, https://halshs.archives-ouvertes.fr/halshs-00467770 (Presses Universitaires d'Orléans, Orléans, 2009).

32. Ozouf-Marignier, M.-V. & Verdier, N. Le Canton d'Hier à Aujourd'hui : Étude Cartographique d'un Maillage. In Lagadec, Y., Le Bihan, J. & Tanguy, J.-F. (eds.) *Le Canton, un Territoire du Quotidien*, 281–95, https://halshs.archives-ouvertes.fr/halshs-00391699 (Presses Universitaires de Rennes, Rennes, 2009).

33. Ozouf-Marignier, M.-V. & Verdier, N. Les Mutations des Circonscriptions Territoriales Françaises. Crise ou Mutation? *Mélanges de l'École Francaise de Rome - Italie et Méditerranée Modernes et Contemporaines* **125**, 271–84, https://doi.org/10.4000/mefrim.1462 (2013).

34. Boulanger, P. *La France Devant la Conscription. Géographie Historique d'une Institution Républicaine, 1914–1922*, 15–37 (Economica, Paris, 2001).

35. Chanet, J.-F. *Vers l'Armée Nouvelle: République Conservatrice et Réforme Militaire, 1871–1879.* (Presses Universitaires de Rennes, Rennes, 2006).

36. Farcy, J. C. *Guide des Archives Judiciaires et Penitentiaires: 1800–1958* (CNRS éditions, Paris, 1992).

37. Gaudillère, B. *Atlas Historique des Circonscriptions Électorales Françaises* (Librairie Droz, Genève, 1995).

38. Marty, T. De l'Espace Électoral à son Découpage: Circonscriptions Uninominales et Circonscriptions de Listes sous la Troisième République Française. *Norois* **226**, 91–106, https://doi.org/10.4000/norois.4606 (2013).

39. Direction des Archives. Série F—F/10 à F/17. In *Etat Sommaire des Versements Faits aux Archives Nationales par les Ministères et les Administrations qui en Dépendent*, 751–9 (Ministère de l'Instruction Publique et des Beaux-Arts, Paris, 1962).

40. Chetcuti, C. & Le Noël, M. Évolution et organisation de l'administration centrale du ministère du Travail de 1887 à 1940. In *Études et Documents pour Servir à l'Histoire de l'Administration du Travail*, 1–138 (Ministère de l'Emploi et de la Solidarité, Paris, 1998).

41. Duquesnoy, J.-P. *Atlas Historique des Diocèses en France* (Archives & Culture, Paris, 2020).

42. Sorrel, C. (ed.) *Les "Matériaux Boulard" Trente Ans Après. Des Chiffres et des Cartes: Approches Sérielles et Spatiales en Histoire Religieuse.* (LARHRA, Lyon, 2013).

43. Deloche, E. *Le Diocèse d'Annecy de la Séparation à Vatican II (1905–1962)* (Académie Salésienne, Annecy, 2017).

44. DataCite Metadata Working Group. DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 4.3, https://doi.org/10.14454/7xq3-zf69 (2019).

45. France. Code Pénitentiaire. 1879–1910, https://gallica.bnf.fr/ark:/12148/cb32743831w/date (2007).

46. France. Code des Prisons. 1845–1876, https://gallica.bnf.fr/ark:/12148/cb32743829b/date (2017).

47. France. Bulletin Administratif de l'Instruction Publique. 1850–1890, https://education.persee.fr/collection/baip (2018).

48. France. Compte Rendu sur le Recrutement de l'Armée (1819–1923), https://gallica.bnf.fr/ark:/12148/cb414027628/date (2015).

49. France. Compte Général de l'Administration de la Justice Civile et Commerciale en France (1831–1935), https://gallica.bnf.fr/ark:/12148/cb345304266/date (2010).

50. France. Compte Général de l'Administration de la Justice Criminelle en France (1827–1934), https://gallica.bnf.fr/ark:/12148/cb412790415/date (2010).

51. France. Statistique Pénitentiaire (1880–1920), https://gallica.bnf.fr/ark:/12148/cb41286437h/date (2011).

52. Brée, S. Les Données Historiques Quantitatives Disponibles en Ligne, https://progedo.hypotheses.org/514 (2020).

53. France. Carte cantonale de la France, https://catalogue.bnf.fr/ark:/12148/cb407750021 (1884).

54. France. Carte cantonale de la France, https://catalogue.bnf.fr/ark:/12148/cb40688588q (1923).

55. StataCorp. Stata Statistical Software: Release 16, https://www.stata.com/ (2019).

56. QGIS.org. QGIS Geographic Information System. Open Source Geospatial Foundation Project, http://qgis.org (2020).

Acknowledgements

The author thanks Anne Alonzo, Yuzuru Kumon, Christophe Lévêque, Imil Nurutdinov, Mohamed Saleh, Jan Stuckatz, Ronan Tallec, and Erik Wang for fruitful discussions. He also gratefully acknowledges funding from the Agence Nationale de la Recherche (ANR) under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

Competing interests

The author declares that he has no relevant or material financial interests that relate to the research described in this article.

Figures & Tables

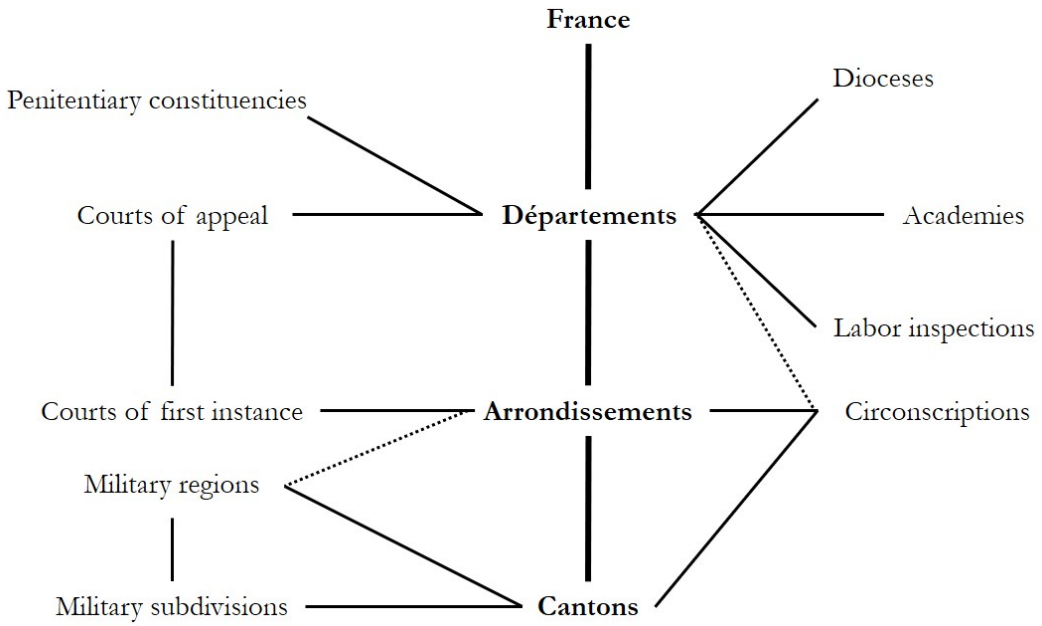


Figure 1. Schematic hierarchy of administrative entities described by the TRF-GIS. Constituencies with a vertical link are strictly nested, with the lower entity being nested into the upper one. Horizontal and diagonal links indicate the entities upon which the geography of a constituency is generally based on—some are based on two different levels, usually depending on the urban-rural status of a location. Dotted lines indicate logic that only applied in certain years.

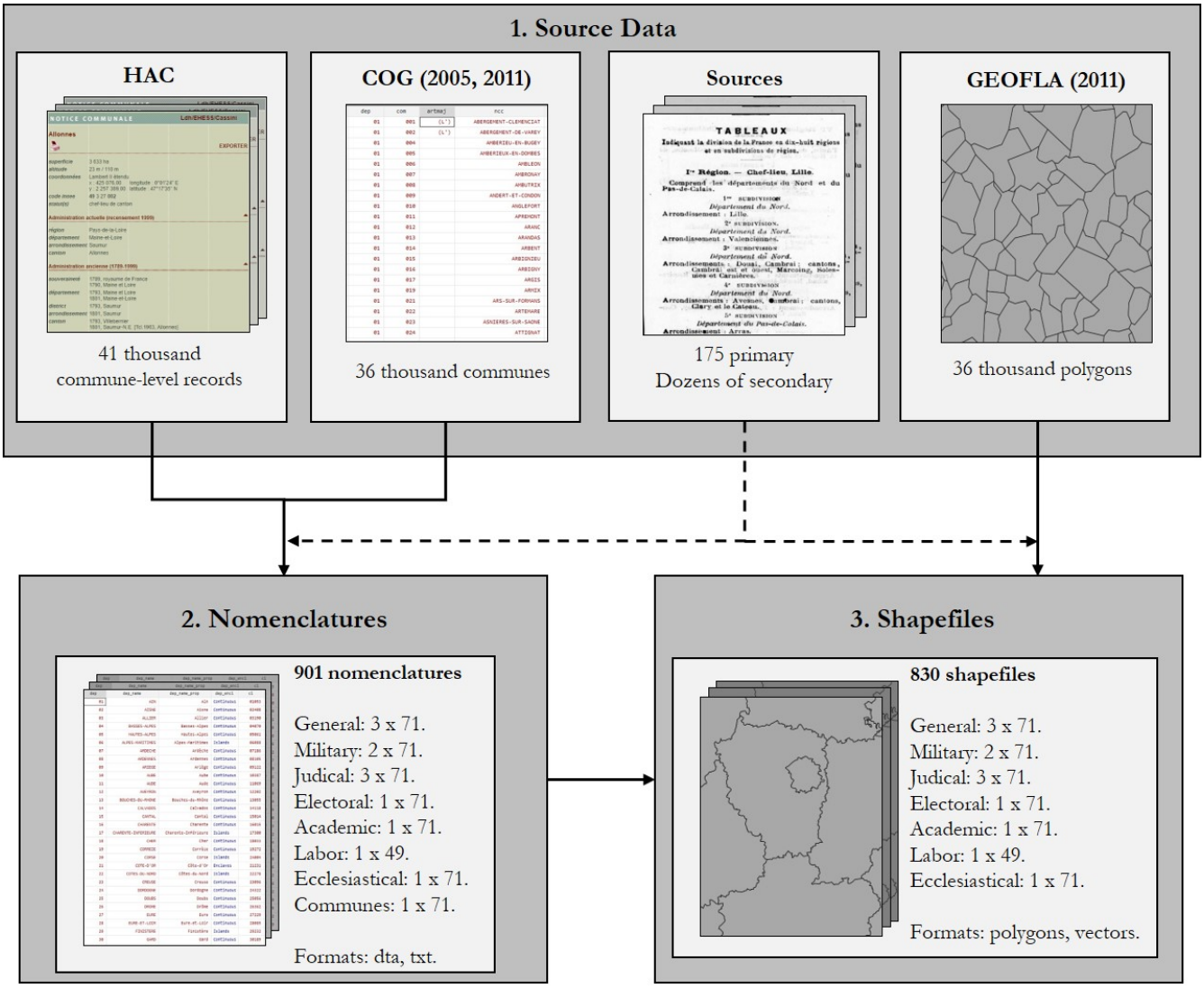


Figure 2. Construction logic of the TRF-GIS. *HAC* corresponds to the *Histoire Administrative des Communes* database distributed through cassini.ehess.fr.[21,22] *COG (2005, 2011)* corresponds to INSEE's *Codes Officiels Géographiques* nomenclatures 2005 and 2011.[18–20]. *GEOFLA (2011)* corresponds to IGN's *GEOFLA® Communes Édition 2011 France Métropolitaine* (GEOFLA) shapefile.[25]

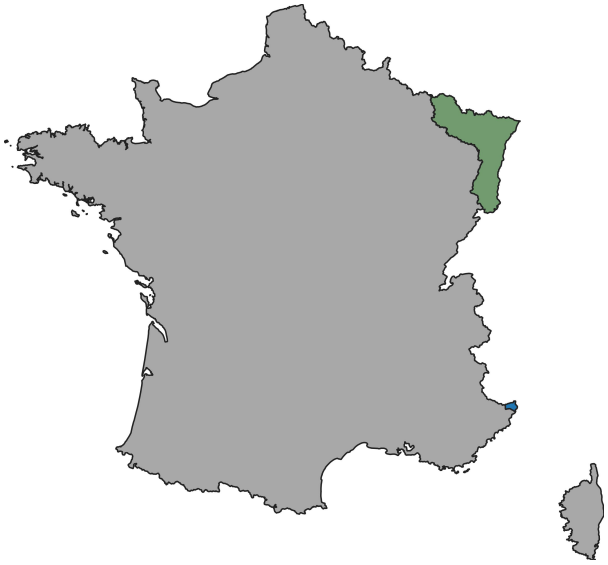


Figure 3. Missing territories during the Third Republic compared to IGN's GEOFLA 2011: Alsace-Lorraine (1871–1918; in green), La Brigue and Tende (1870–1940; in blue).

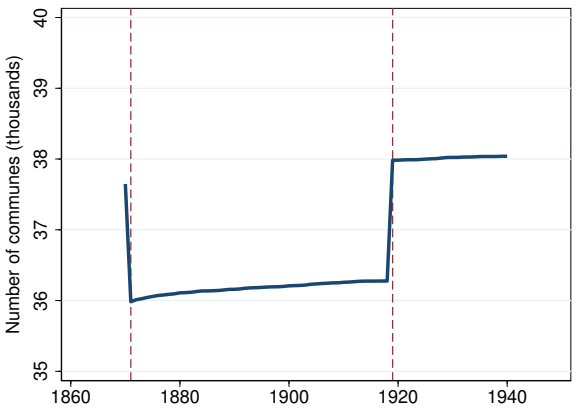


Figure 4. Number of communes in TRF-GIS commune-level nomenclatures (in thousands). Red vertical dashed lines indicate the annexation (1871) and recovery (1919) of Alsace-Lorraine.

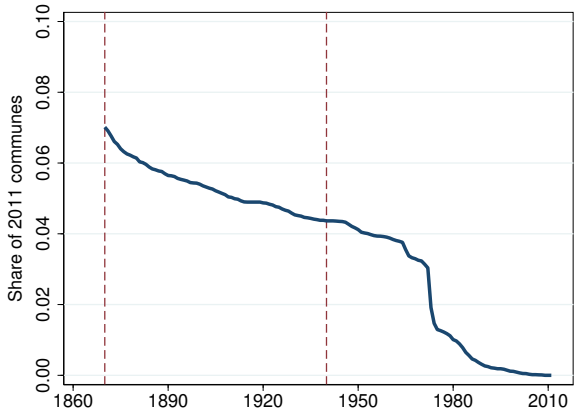
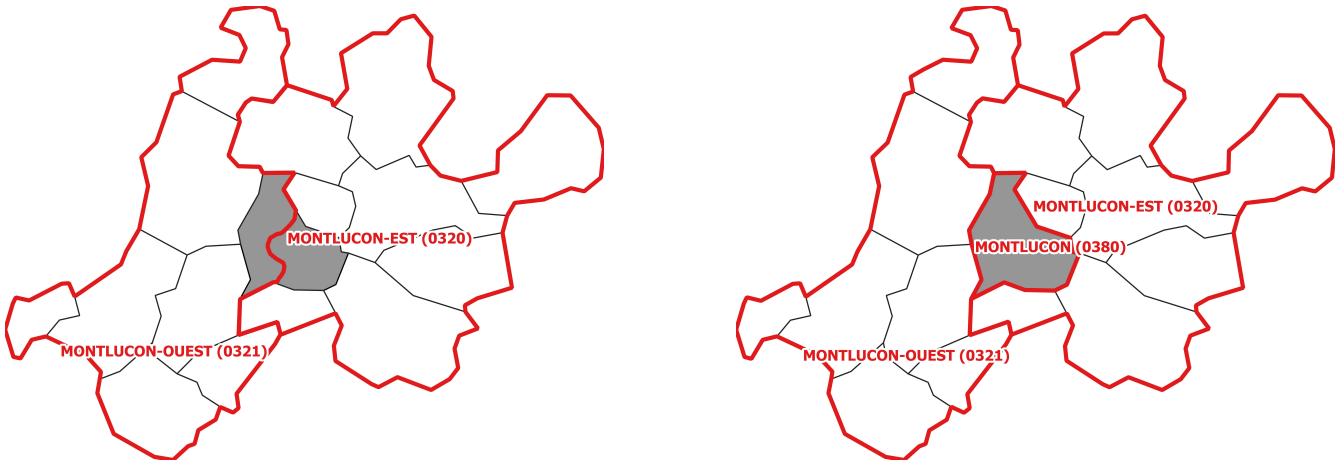
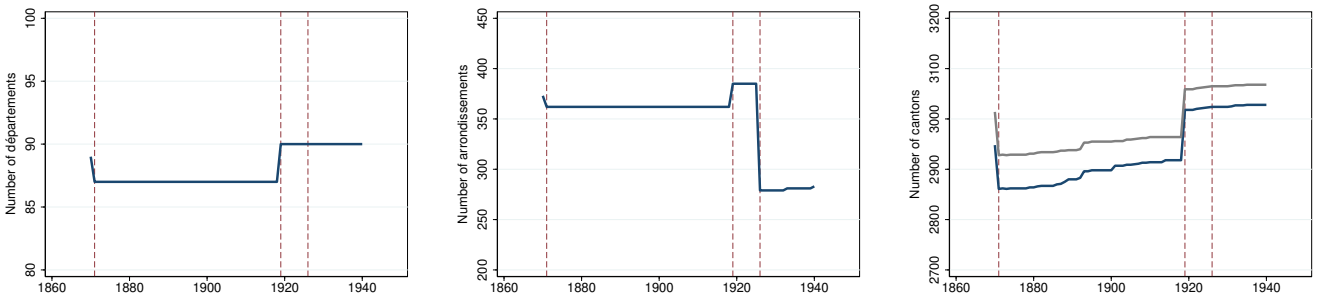


Figure 5. Share of communes of 2011 that underwent territorial modifications since a given year (out of 36,568). Red vertical dashed lines indicate the proclamation (1870) and dissolution (1940) of the Third Republic.



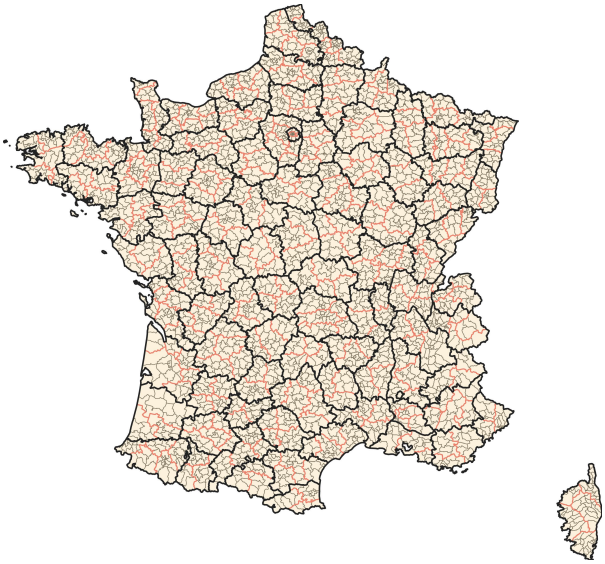
(a) Cantons **(b)** Pseudo-cantons

Figure 6. Cantons and pseudo-cantons of the Montluçon area (1870–1940). Red edges indicate cantons (a) and pseudo-cantons (b) boundaries. Black edges indicate commune boundaries. The shaded area corresponds to the commune of Montluçon.

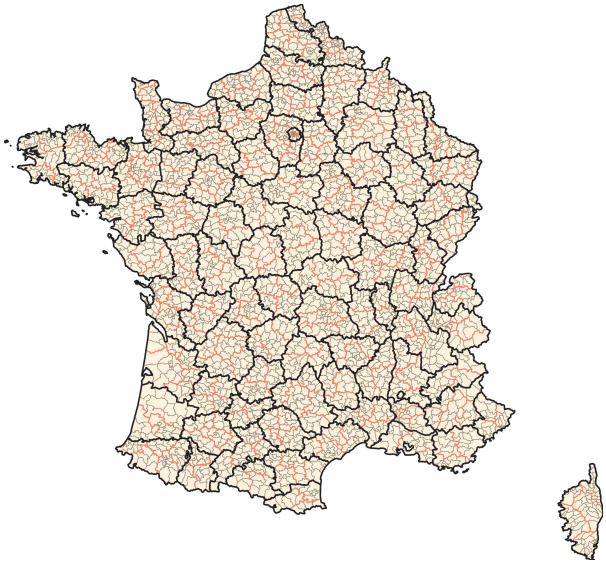


(a) Départements **(b)** Arrondissements **(c)** Cantons and pseudo-cantons

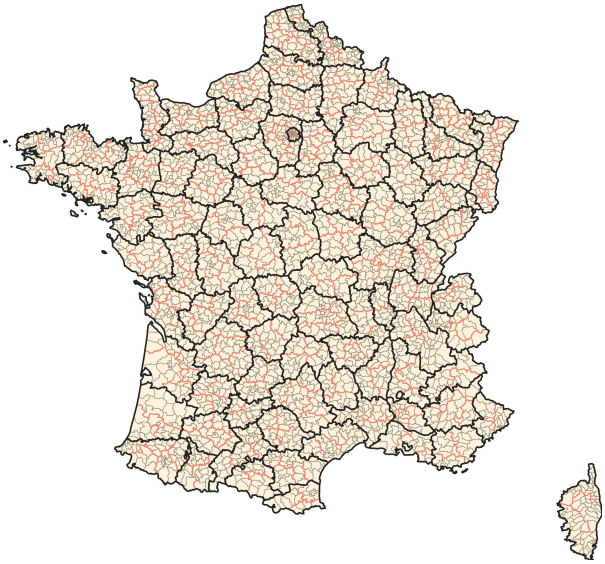
Figure 7. Number of départements, arrondissements, and cantons (1870–1940). Red vertical dashed lines indicate the annexation (1871) and recovery (1919) of Alsace-Lorraine, as well as the arrondissements reform (1926). In Panel c, the blue line denotes cantons, and the gray line, pseudo-cantons.



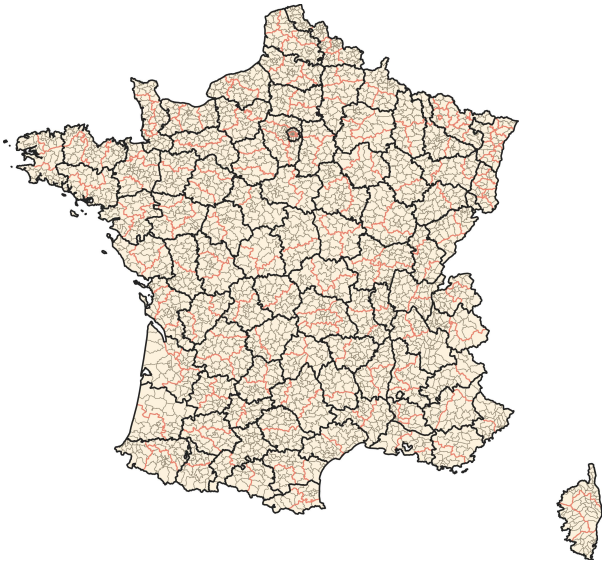
(a) General administrative constituencies in 1870



(b) General administrative constituencies in 1871

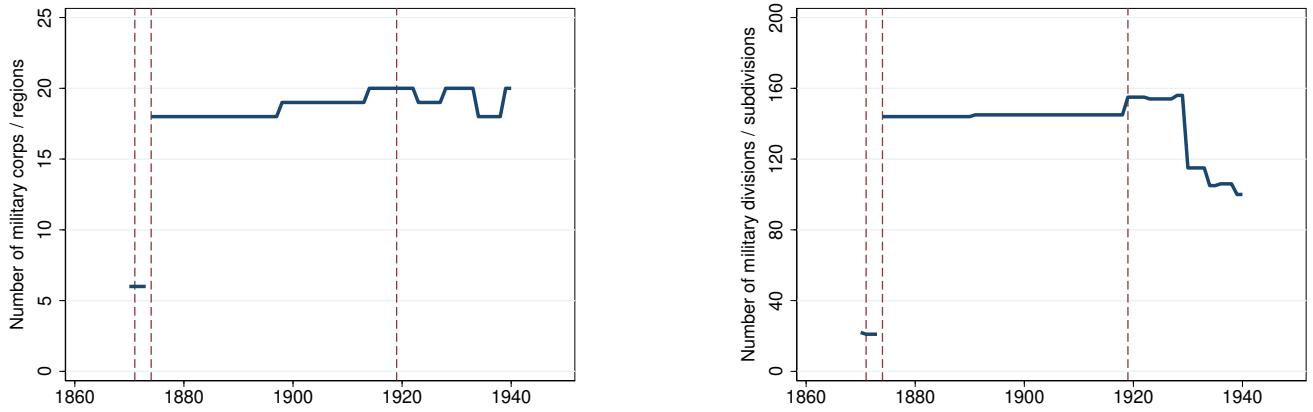


(c) General administrative constituencies in 1919



(d) General administrative constituencies in 1926

Figure 8. General administrative constituencies in 1870, 1871, 1919, and 1926. Black lines delineate départements, red lines, arrondissements, and gray lines, cantons.

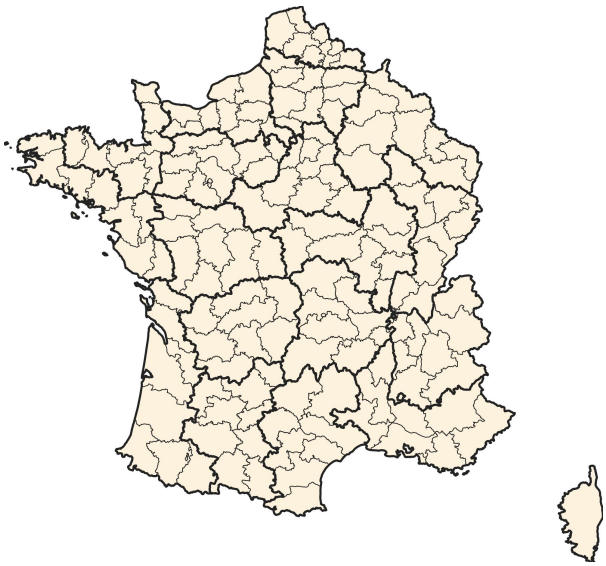


(a) Military corps and regions

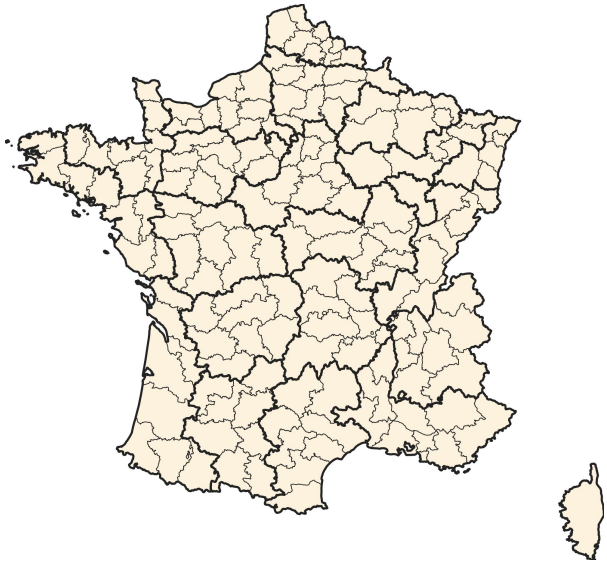
(b) Military divisions and subdivisions

Figure 9. Number of military constituencies (1870–1940). Between 1870 and 1873, military corps and military divisions; between 1874 and 1940, military regions and subdivisions of military regions. Red vertical dashed lines indicate the military reform of 1874, and the annexation (1871) and recovery (1919) of Alsace-Lorraine.

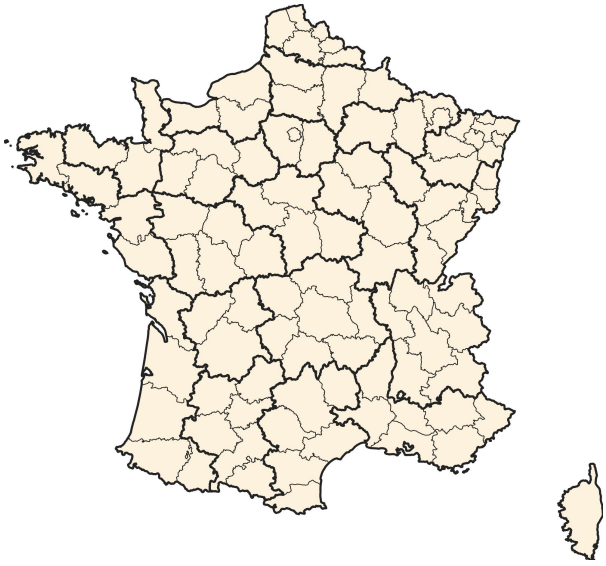
(a) Military constituencies in 1870



(b) Military constituencies in 1874

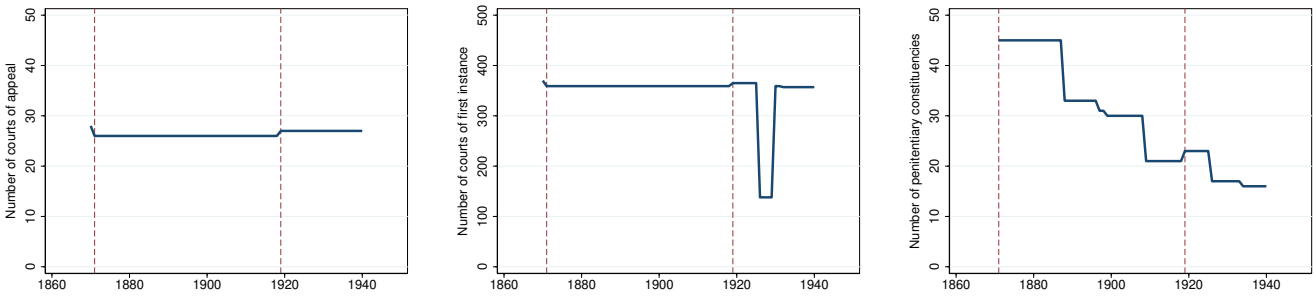


(c) Military constituencies in 1919



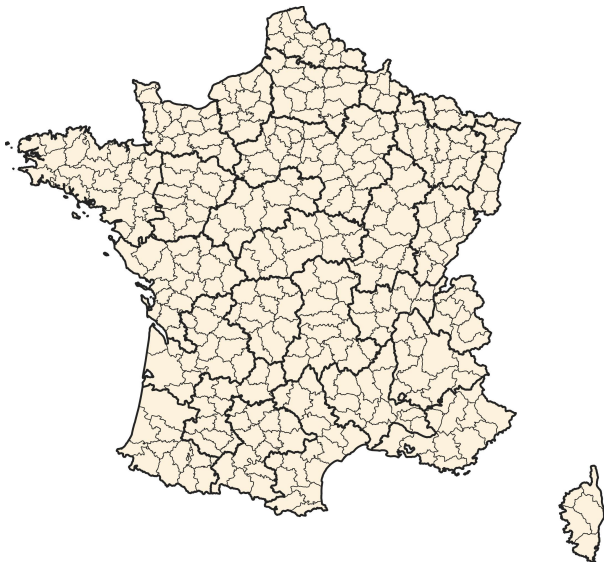
(d) Military constituencies in 1940

Figure 10. Military constituencies in 1870, 1874, 1919, and 1940. Thick lines delineate military corps (1870) and military regions (1874, 1919, 1940) and thin lines, military divisions (1870) and subdivisions of military regions (1874, 1919, 1940).

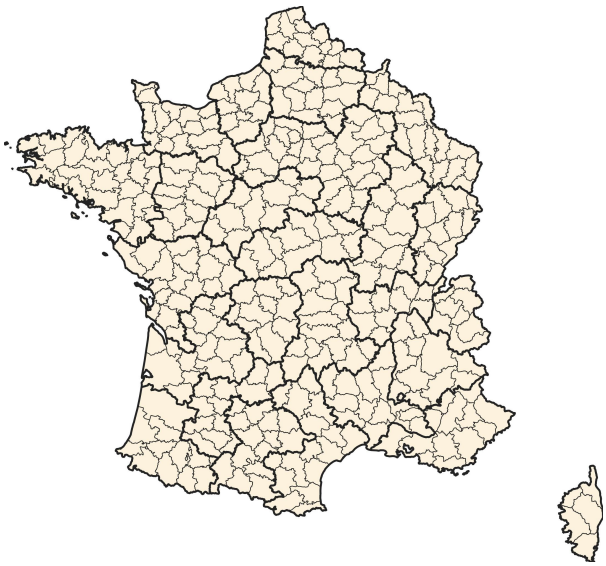


(a) Courts of appeal **(b)** Courts of first instance **(c)** Penitentiary constituencies

Figure 11. Number of judicial (1870–1940) and penitentiary constituencies (1871–1940). Red vertical dashed lines indicate the annexation (1871) and recovery (1919) of Alsace-Lorraine.



(a) Judicial constituencies in 1870

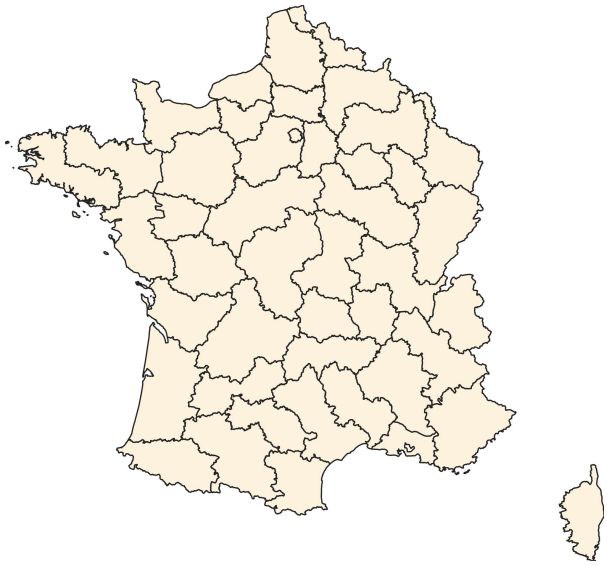


(b) Judicial constituencies in 1871

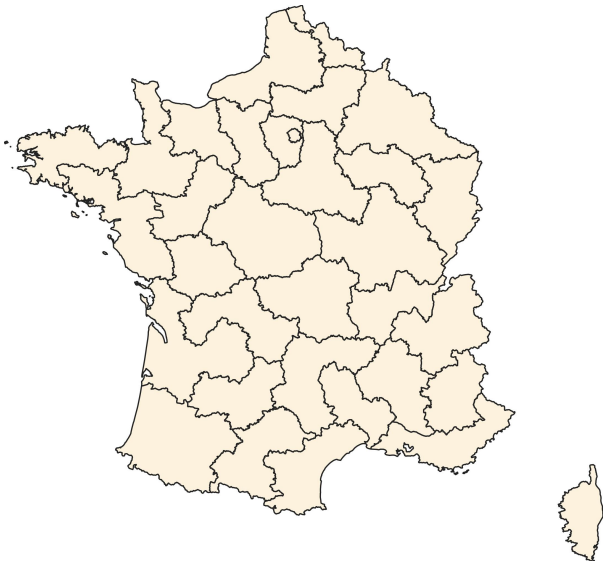
(c) Judicial constituencies in 1926

(d) Judicial constituencies in 1940

Figure 12. Judicial constituencies in 1870, 1871, 1926, and 1940. Thick lines delineate courts of appeal and thin lines, courts of first instance.



(a) Penitentiary constituencies in 1871

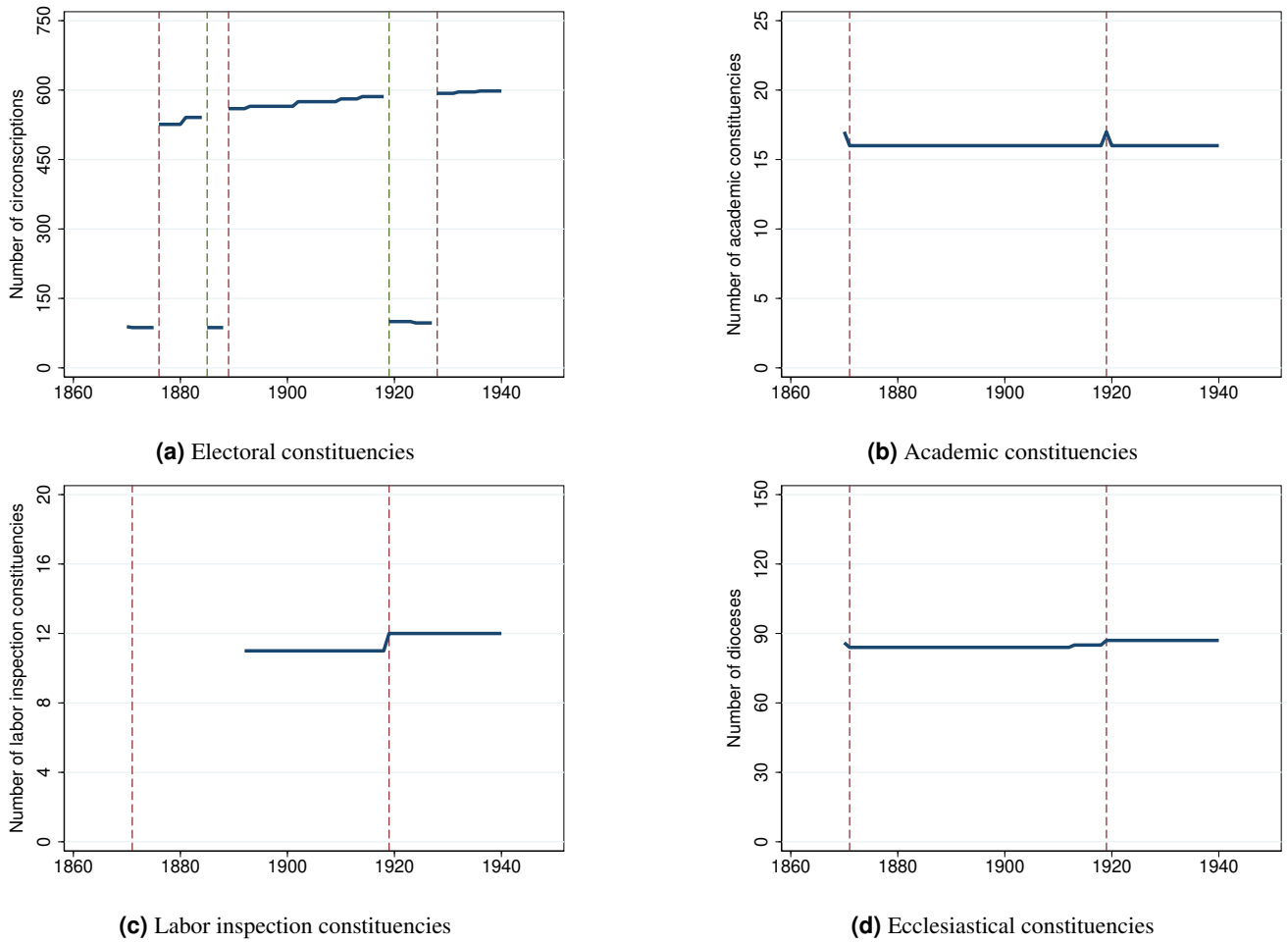


(b) Penitentiary constituencies in 1888

(c) Penitentiary constituencies in 1909

(d) Penitentiary constituencies in 1940

Figure 13. Penitentiary constituencies in 1871, 1888, 1909, and 1940.



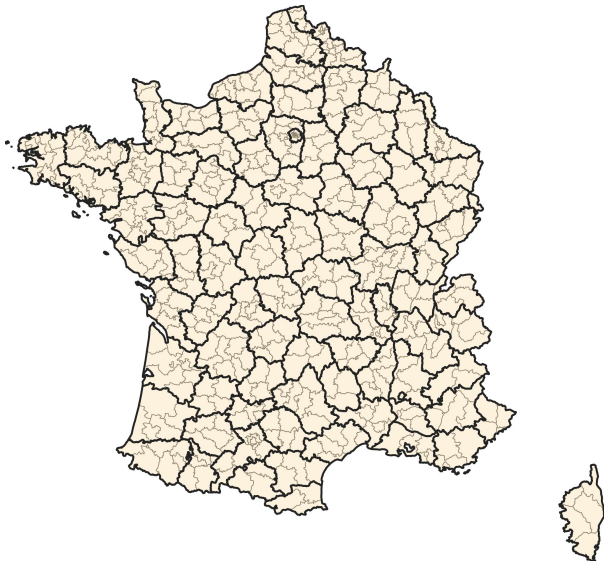
(a) Electoral constituencies

(b) Academic constituencies

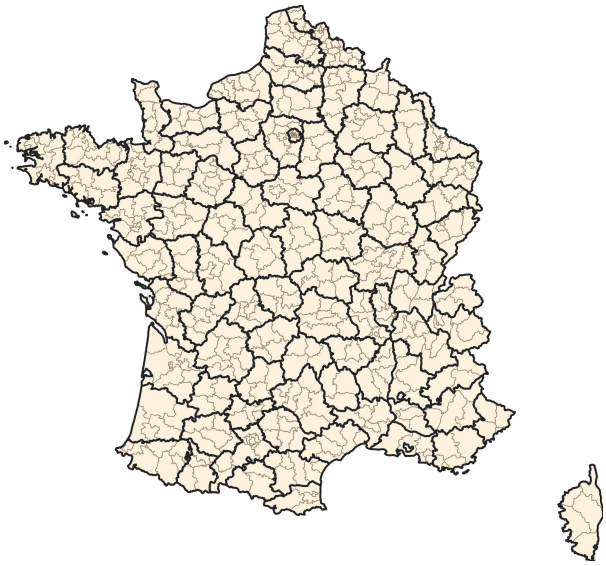
(c) Labor inspection constituencies

(d) Ecclesiastical constituencies

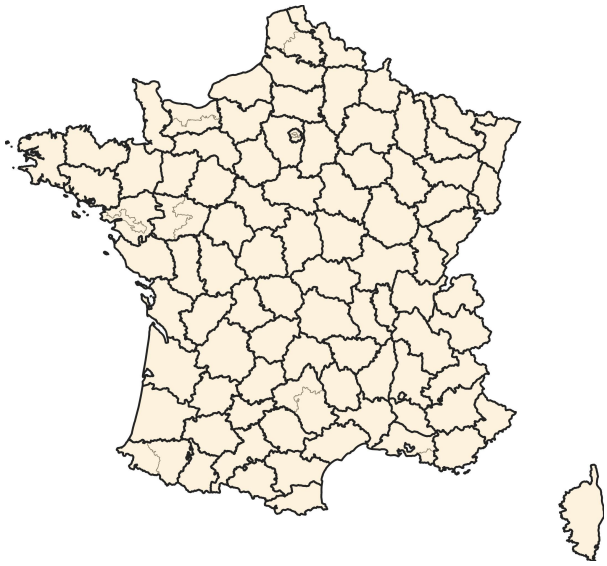
Figure 14. Number of electoral, academic, and ecclesiastical constituencies. In Panel a, red vertical dashed lines indicate shifts from a block-list electoral system to a single-member electoral system, and green vertical dashed lines indicate the reverse. In Panels b, c, and d, red vertical dashed lines indicate the annexation (1871) and recovery (1919) of Alsace-Lorraine.



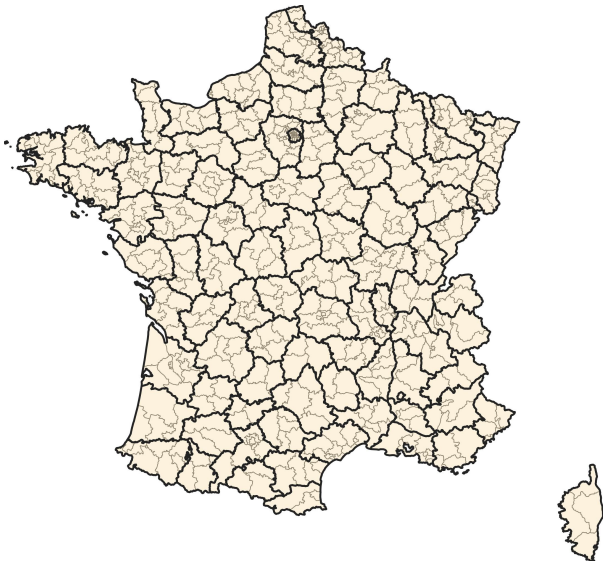
(a) Electoral constituencies in 1876



(b) Electoral constituencies in 1889

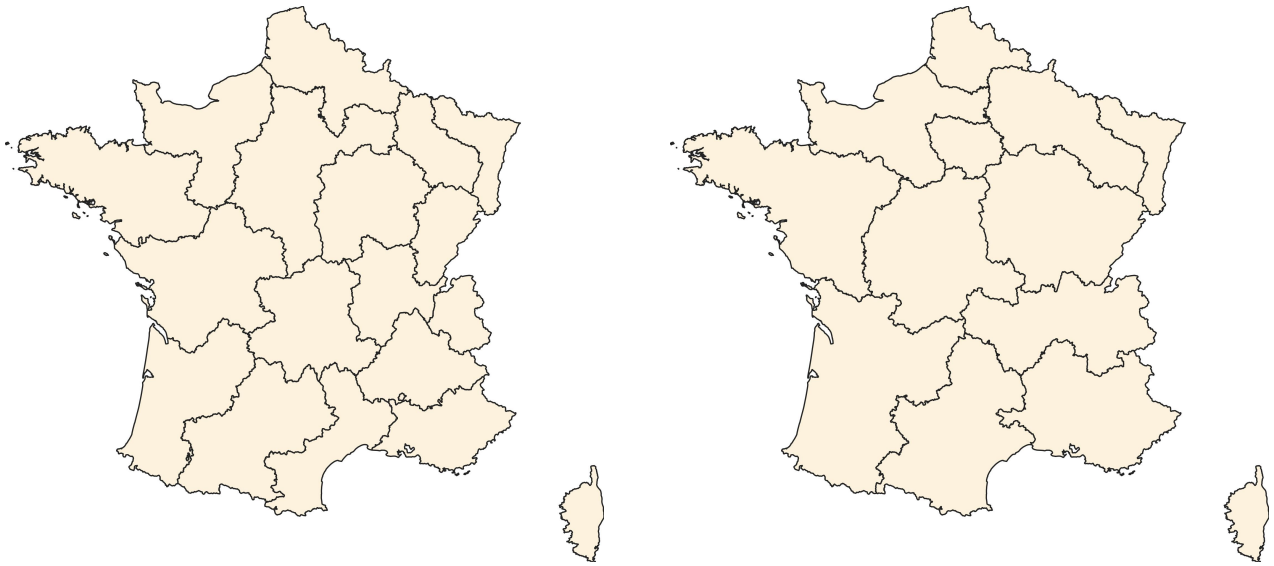


(c) Electoral constituencies in 1919



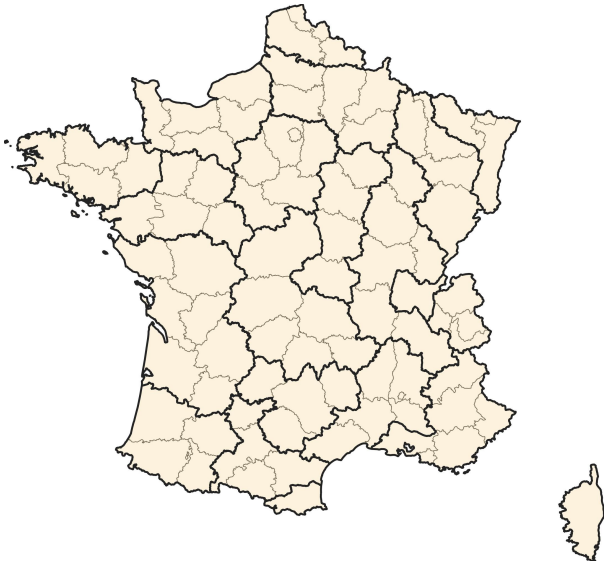
(d) Electoral constituencies in 1928

Figure 15. Electoral constituencies (circonscriptions) in 1876, 1889, 1919, and 1928. Thicker lines delineate départements.



(a) Academic constituencies in 1919

(b) Labor inspection constituencies in 1919



(c) Ecclesiastical constituencies in 1919

Figure 16. Academic, labor inspection, and ecclesiastical constituencies in 1919. In panel c, Thick lines delineate the extent of metropolitan bishoprics and thin lines, dioceses.

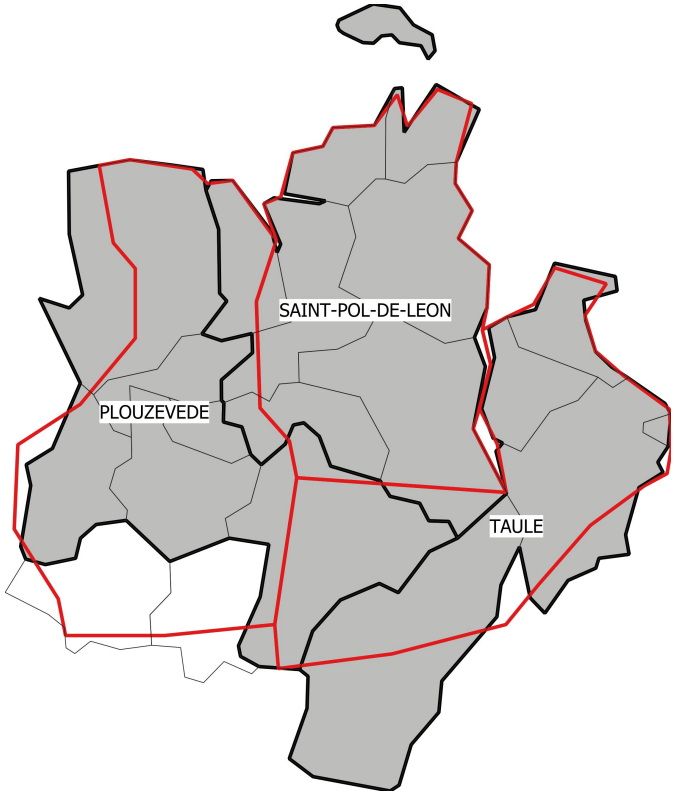


Figure 17. Comparison of TRF-GIS and LARHRA cantons shapefiles 1884.[13] TRF-GIS cantons have thick black boundaries while LARHRA's cantons have thick red boundaries. Thin boundaries correspond to communes in the GEOFLA 2011. The three cantons are Plouzévédé, Saint-Pol-De-Léon, and Taulé in the département of Finistère (29).

Acronym	Definition	Description
TRF-GIS	Third-Republic France Geographic Information System	Database
INSEE	Institut National de la Statistique et des Etudes Economiques	National statistics bureau of France
COG	Code Officiel Géographique	Officiel geographic code
HAC	Histoire Administrative des Communes	Database of historical nomenclatures for general constituencies
IGN	Institut Géographique National	National bureau of geographical information of France
GEOFLA	(no meaning)	IGN shapefile of France
BD CARTO	Base de Données Cartographique	IGN shapefile of France
RGF93	Réseau Géodésique Français 1993	Projection system used by IGN
LARHRA	Laboratoire de recherche historique Rhône-Alpes	Laboratory in Lyon 2 university
COMMUNE HIS-DBD	COllaborative Micro Mapping of UNExploited HIStorical District-Boundary Data)	Research project
ALPAGE	AnaLyse diachronique de l'espace urbain PArisien : Approche GEomatique	Research project
PID	Persistent identifier	Long-lasting reference to a document
ARK	Archival Resource Key	PID for archival resources
CC-BY	Creative Commons (CC) license with attribution (BY)	Licensing category

Table 1. Description of acronyms used in the article.

Constituency	Content	Name	Format	Size	Dataset DOI
Départements	Nomenclatures	COG_DEPARTEMENTS_YYYY	.dta	∼20 kB	10.7910/DVN/ULQYM5
	Nomenclatures	COG_DEPARTEMENTS_YYYY	.txt	6 kB	10.7910/DVN/ULQYM5
	Codebook and metadata	COG_DEPARTEMENTS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/ULQYM5
	Shapefiles (polygon)	DEPARTEMENTS_YYYY	GIS	∼700 kB	10.7910/DVN/ULQYM5
	Shapefiles (vector)	LIMITES_DEPARTEMENTS_YYYY	GIS	∼700 kB	10.7910/DVN/ULQYM5
	Prefectures coordinates	CHEFLIEU_DEPARTEMENTS_YYYY	.csv	4 kB	10.7910/DVN/ULQYM5
	Shapefiles metadata	METADATA_DEPARTEMENTS	.qmd	4 kB	10.7910/DVN/ULQYM5
Arrondissements	Nomenclatures	COG_ARRONDISSEMENTS_YYYY	.dta	∼70 kB	10.7910/DVN/98ERPB
	Nomenclatures	COG_ARRONDISSEMENTS_YYYY	.txt	∼30 kB	10.7910/DVN/98ERPB
	Codebook and metadata	COG_ARRONDISSEMENTS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/98ERPB
	Shapefiles (polygon)	ARRONDISSEMENTS_YYYY	GIS	∼1.70 MB	10.7910/DVN/98ERPB
	Shapefiles (vector)	LIMITES_ARRONDISSEMENTS_YYYY	GIS	∼1.70 MB	10.7910/DVN/98ERPB
	Subprefectures coordinates	CHEFLIEU_ARRONDISSEMENTS_YYYY	.csv	21 kB	10.7910/DVN/98ERPB
	Shapefiles metadata	METADATA_ARRONDISSEMENTS	.qmd	4 kB	10.7910/DVN/98ERPB
Cantons	Nomenclatures	COG_CANTONS_YYYY	.dta	∼870 kB	10.7910/DVN/U85RSZ
	Nomenclatures	COG_CANTONS_YYYY	.txt	∼360 kB	10.7910/DVN/U85RSZ
	Codebook and metadata	COG_CANTONS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/U85RSZ
	Mapping cantons to pseudo-cantons	COG_CANTONS_YYYY_MAPPING	.dta	∼80 kB	10.7910/DVN/U85RSZ
	Shapefiles (polygon)	CANTONS_YYYY	GIS	∼9.00 MB	10.7910/DVN/U85RSZ
	Shapefiles (vector)	LIMITES_CANTONS_YYYY	GIS	∼9.00 MB	10.7910/DVN/U85RSZ
	Bureaux centralisateurs coordinates	CHEFLIEU_CANTONS_YYYY	.csv	169 kB	10.7910/DVN/U85RSZ
	Shapefiles metadata	METADATA_CANTONS	.qmd	4 kB	10.7910/DVN/U85RSZ

Table 2. Summary of TRF-GIS data files for general administrative constituencies. YYYY denotes a year between 1870 and 1940. *GIS* files are a set of five files in .shp, .shx, .dbf, .prj, and .cpg format. *Size* refers to the size of individual data files once unzipped.

Constituency	Content	Name	Format	Size	Dataset DOI
Military corps (1870–1873)	Nomenclatures	COG_MIL_CORPS_YYYY	.dta	13 kB	10.7910/DVN/2KMLSE
	Nomenclatures	COG_MIL_CORPS_YYYY	.txt	1 kB	10.7910/DVN/2KMLSE
	Codebook and metadata	COG_MIL_CORPS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/2KMLSE
	Shapefiles (polygon)	MIL_CORPS_YYYY	GIS	205 kB	10.7910/DVN/2KMLSE
	Shapefiles (vector)	LIMITES_MIL_CORPS_YYYY	GIS	205 kB	10.7910/DVN/2KMLSE
	Chefs-lieux coordinates	CHEFLIEU_MIL_CORPS_YYYY	.csv	1 kB	10.7910/DVN/2KMLSE
Military divisions (1870–1873)	Nomenclatures	COG_MIL_DIVISIONS_YYYY	.dta	17 kB	10.7910/DVN/2H2JVT
	Nomenclatures	COG_MIL_DIVISIONS_YYYY	.txt	2 kB	10.7910/DVN/2H2JVT
	Codebook and metadata	COG_MIL_DIVISIONS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/2H2JVT
	Shapefiles (polygon)	MIL_DIVISIONS_YYYY	GIS	∼360 kB	10.7910/DVN/2H2JVT
	Shapefiles (vector)	LIMITES_MIL_DIVISIONS_YYYY	GIS	∼360 kB	10.7910/DVN/2H2JVT
	Chefs-lieux coordinates	CHEFLIEU_MIL_DIVISIONS_YYYY	.csv	1 kB	10.7910/DVN/2H2JVT
Military regions (1874–1940)	Nomenclatures	COG_MIL_REGIONS_YYYY	.dta	∼20 kB	10.7910/DVN/SQPEUW
	Nomenclatures	COG_MIL_REGIONS_YYYY	.txt	2 kB	10.7910/DVN/SQPEUW
	Codebook and metadata	COG_MIL_REGIONS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/SQPEUW
	Mapping	COG_MIL_REGIONS_YYYY_MAPPING	.dta		10.7910/DVN/SQPEUW
	Shapefiles (polygon)	MIL_REGIONS_YYYY	GIS	∼370 kB	10.7910/DVN/SQPEUW
	Shapefiles (vector)	LIMITES_MIL_REGIONS_YYYY	GIS	∼370 kB	10.7910/DVN/SQPEUW
	Chefs-lieux coordinates	CHEFLIEU_MIL_REGIONS_YYYY	.csv	1 kB	10.7910/DVN/SQPEUW
Military subdivisions (1874–1940)	Nomenclatures	COG_MIL_SUBDIVISIONS_YYYY	.dta	∼ 30 kB	10.7910/DVN/TAMQAG
	Nomenclatures	COG_MIL_SUBDIVISIONS_YYYY	.txt	∼10 kB	10.7910/DVN/TAMQAG
	Codebook and metadata	COG_MIL_SUBDIVISIONS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/TAMQAG
	Mapping	COG_MIL_SUBDIVISIONS_YYYY_MAPPING	.dta		10.7910/DVN/TAMQAG
	Shapefiles (polygon)	MIL_SUBDIVISIONS_YYYY	GIS	∼1.10 MB	10.7910/DVN/TAMQAG
	Shapefiles (vector)	LIMITES_MIL_SUBDIVISIONS_YYYY	GIS	∼1.10 MB	10.7910/DVN/TAMQAG
	Chefs-lieux coordinates	CHEFLIEU_MIL_SUBDIVISIONS_YYYY	.csv	∼10 kB	10.7910/DVN/TAMQAG
Military corps and regions	Shapefiles metadata	METADATA_MIL_REGIONS	.qmd	4 kB	
Military divisions and subdivisions	Shapefiles metadata	METADATA_MIL_SUBDIVISIONS	.qmd	4 kB	

Table 3. Summary of TRF-GIS data files for military constituencies. YYYY denotes a year between 1870 and 1940. *Mapping* corresponds to mappings from military regions/subdivisions to pseudo-military regions/subdivisions. *GIS* files are a set of five files in .shp, .shx, .dbf, .prj, and .cpg format. *Size* refers to the size of individual data files once unzipped.

Constituency	Content	Name	Format	Size	Dataset DOI
Courts of appeal	Nomenclatures	COG_COURTS_APPEAL_YYYY	.dta	13 kB	10.7910/DVN/QV6ZBS
	Nomenclatures	COG_COURTS_APPEAL_YYYY	.txt	2 kB	10.7910/DVN/QV6ZBS
	Codebook and metadata	COG_COURTS_APPEAL_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/QV6ZBS
	Shapefiles (polygon)	COURTS_APPEAL_YYYY	GIS	∼400 kB	10.7910/DVN/QV6ZBS
	Shapefiles (vector)	LIMITES_COURTS_APPEAL_YYYY	GIS	∼400 kB	10.7910/DVN/QV6ZBS
	Chefs-lieux coordinates	CHEFLIEU_COURTS_APPEAL_YYYY	.csv	2 kB	10.7910/DVN/QV6ZBS
	Shapefiles metadata	METADATA_COURTS_APPEAL	.qmd	4 kB	10.7910/DVN/QV6ZBS
Courts of first instance	Nomenclatures	COG_COURTS_INSTANCE_YYYY	.dta	∼50 kB	10.7910/DVN/VE2OJS
	Nomenclatures	COG_COURTS_INSTANCE_YYYY	.txt	∼20 kB	10.7910/DVN/VE2OJS
	Codebook and metadata	COG_COURTS_INSTANCE_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/VE2OJS
	Shapefiles (polygon)	COURTS_INSTANCE_YYYY	GIS	∼1.70 MB	10.7910/DVN/VE2OJS
	Shapefiles (vector)	LIMITES_COURTS_INSTANCE_YYYY	GIS	∼1.70 MB	10.7910/DVN/VE2OJS
	Chefs-lieux coordinates	CHEFLIEU_COURTS_INSTANCE_YYYY	.csv	20 kB	10.7910/DVN/VE2OJS
	Shapefiles metadata	METADATA_COURTS_INSTANCE	.qmd	4 kB	10.7910/DVN/VE2OJS
Penitentiary constituencies	Nomenclatures	COG_PENITENTIARY_YYYY	.dta	∼15 kB	10.7910/DVN/XQUXBO
	Nomenclatures	COG_PENITENTIARY_YYYY	.txt	∼3 kB	10.7910/DVN/XQUXBO
	Codebook and metadata	COG_PENITENTIARY_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/XQUXBO
	Shapefiles (polygon)	PENITENTIARY_YYYY	GIS	∼450 kB	10.7910/DVN/XQUXBO
	Shapefiles (vector)	LIMITES_PENITENTIARY_YYYY	GIS	∼450 kB	10.7910/DVN/XQUXBO
	Chefs-lieux coordinates	CHEFLIEU_PENITENTIARY_YYYY	.csv	2 kB	10.7910/DVN/XQUXBO
	Shapefiles metadata	METADATA_PENITENTIARY	.qmd	4 kB	10.7910/DVN/XQUXBO

Table 4. Summary of TRF-GIS data files for judicial and penitentiary constituencies. YYYY denotes a year between 1870 and 1940. *GIS* files are a set of five files in .shp, .shx, .dbf, .prj, and .cpg format. *Size* refers to the size of individual data files once unzipped.

Constituency	Content	Name	Format	Size	Dataset DOI
Circonscriptions	Nomenclatures	COG_CIRCONSCRIPTIONS_YYYY	.dta	∼40 kB	10.7910/DVN/L2LGDW
	Nomenclatures	COG_CIRCONSCRIPTIONS_YYYY	.txt	∼10 kB	10.7910/DVN/L2LGDW
	Codebook and metadata	COG_CIRCONSCRIPTIONS_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/L2LGDW
	Mapping	COG_CIRCONSCRIPTIONS_YYYY_MAPPING	.dta		10.7910/DVN/L2LGDW
	Shapefiles (polygon)	CIRCONSCRIPTIONS_YYYY	GIS	∼1 MB	10.7910/DVN/L2LGDW
	Shapefiles (vector)	LIMITES_CIRCONSCRIPTIONS_YYYY	GIS	∼1 MB	10.7910/DVN/L2LGDW
	Shapefiles metadata	METADATA_CIRCONSCRIPTIONS	.qmd	4 kB	10.7910/DVN/L2LGDW
Academies	Nomenclatures	COG_ACADEMIES_YYYY	.dta	12 kB	10.7910/DVN/PMVTFU
	Nomenclatures	COG_ACADEMIES_YYYY	.txt	2 kB	10.7910/DVN/PMVTFU
	Codebook and metadata	COG_ACADEMIES_YYYY_CODEBOOK	.txt	1 kB	10.7910/DVN/PMVTFU
	Shapefiles (polygon)	ACADEMIES_YYYY	GIS	∼300 kB	10.7910/DVN/PMVTFU
	Shapefiles (vector)	LIMITES_ACADEMIES_YYYY	GIS	∼300 kB	10.7910/DVN/PMVTFU
	Chefs-lieux coordinates	CHEFLIEU_ACADEMIES_YYYY	.csv	1 kB	10.7910/DVN/PMVTFU
	Shapefiles metadata	METADATA_ACADEMIES	.qmd	4 kB	10.7910/DVN/PMVTFU
Labor inspections (1892–1940)	Nomenclatures	COG_INSPECTIONS_YYYY	.dta	10 kB	10.7910/DVN/OSOFAO
	Nomenclatures	COG_INSPECTIONS_YYYY	.txt	2 kB	10.7910/DVN/OSOFAO
	Codebook and metadata	COG_INSPECTIONS_YYYY_CODEBOOK	.txt	1 kB	10.7910/DVN/OSOFAO
	Shapefiles (polygon)	INSPECTIONS_YYYY	GIS	∼250 kB	10.7910/DVN/OSOFAO
	Shapefiles (vector)	LIMITES_INSPECTIONS_YYYY	GIS	∼250 kB	10.7910/DVN/OSOFAO
	Chefs-lieux coordinates	CHEFLIEU_INSPECTIONS_YYYY	.csv	1 kB	10.7910/DVN/OSOFAO
	Shapefiles metadata	METADATA_INSPECTIONS	.qmd	4 kB	10.7910/DVN/OSOFAO
Dioceses	Nomenclatures	COG_DIOCESES_YYYY	.dta	23 kB	10.7910/DVN/A66YJK
	Nomenclatures	COG_DIOCESES_YYYY	.txt	8 kB	10.7910/DVN/A66YJK
	Codebook and metadata	COG_DIOCESES_YYYY_CODEBOOK	.txt	2 kB	10.7910/DVN/A66YJK
	Shapefiles (polygon)	DIOCESES_YYYY	GIS	∼700 kB	10.7910/DVN/A66YJK
	Shapefiles (vector)	LIMITES_DIOCESES_YYYY	GIS	∼700 kB	10.7910/DVN/A66YJK
	Chefs-lieux coordinates	CHEFLIEU_DIOCESES_YYYY	.csv	3 kB	10.7910/DVN/A66YJK
	Shapefiles metadata	METADATA_DIOCESES	.qmd	4 kB	10.7910/DVN/A66YJK
Communes	Nomenclatures	COG_COMMUNES_YYYY	.dta	∼ 24.0 MB	10.7910/DVN/LZTZWE
	Nomenclatures	COG_COMMUNES_YYYY	.txt	∼ 12.0 MB	10.7910/DVN/LZTZWE
	Codebook and metadata	COG_COMMUNES_YYYY_CODEBOOK	.txt	7 kB	10.7910/DVN/LZTZWE

Table 5. Summary of TRF-GIS data files for electoral, academic, labor inspection, and ecclesiastical constituencies, and commune-level nomenclatures. `YYYY` denotes a year between 1870 and 1940. *Mapping* corresponds to mappings from circonscriptions to pseudo-circonscriptions. *GIS* files are a set of five files in `.shp`, `.shx`, `.dbf`, `.prj`, and `.cpg` format. *Size* refers to the size of individual data files once unzipped.

Nomenclature	Variable	Label	Values
Common to all	dep	Département id	2 digits
	dep_name	Département name (capitalized)	String
	dep_name_prop	Département name (proper)	String
Départements	dep_encl	Département with enclaves	Continuous (0), Enclaves (1), Islands (2)
Arrondissements	depar	Département-arrondissement id	3 digits
	ar	Arrondissement id	1 digit
	ar_name	Arrondissement name (capitalized)	String
	ar_name_prop	Arrondissement name (proper)	String
	cl_multi	Arrondissement chef-lieu common	Unique (0), Common (1)
	cl_out	Arrondissement chef-lieu outside	Inside (0), Outside (1)
Cantons	depar	Département-arrondissement id	3 digits
	ar_name	Arrondissement name (capitalized)	String
	ar_name_prop	Arrondissement name (proper)	String
	depct	Département-canton id	4 digits
	ct	Canton id	2 digits
	ct_name	Canton name (capitalized)	String
	ct_name_prop	Canton name (proper)	String
	ct_type	Canton communal composition	See table notes
	cl_multi	Canton chef-lieu common	Unique (0), Common (1)
Common to all	cl	Constituency chef-lieu id	5-digit INSEE code
	cl_name	Constituency chef-lieu name (capitalized)	String
	cl_name_prop	Constituency chef-lieu name (proper)	String
	cl_act	Constituency chef-lieu actual commune	Current (1), Former (2)
	cl_x	Constituency chef-lieu latitude	Lambert-93 projection
	cl_y	Constituency chef-lieu longitude	Lambert-93 projection

Table 6. Variables in annual general administrative constituency-level nomenclatures. Values for ct_type: Only entire commune(s) (1), Fraction of one commune and entire commune(s) (2), Fractions of several communes and entire commune(s) (3), Fraction of one commune (4), Fractions of several communes (5).

Nomenclature	Variable	Label	Values
Common to 1870–1873	mcor	Military corps id	1 digit
	mcor_name	Military corps name (capitalized)	String
	mcor_name_prop	Military corps name (proper)	String
Military corps (1870–1873)	mcor_type	Military corps type	See table notes
Military divisions (1870–1873)	mcordiv	Military corps-division id	3 digits
	mdiv	Military division id	2 digits
	mdiv_name	Military division name (capitalized)	String
	mdiv_name_prop	Military division name (proper)	String
	mdiv_type	Military division type	See table notes
Common to 1874–1940	mreg	Military region id	2 digits
	mreg_name	Military region name (capitalized)	String
	mreg_name_prop	Military region name (proper)	String
Military regions (1874–1940)	mreg_type	Military region type	See table notes
Military subdivisions (1874–1940)	mregsub	Military region-subdivision id	4 digits
	msub	Military subdivision id	2 digits
	msub_name	Military subdivision name (capitalized)	String
	msub_name_prop	Military subdivision name (proper)	String
	msub_type	Military subdivision type	See table notes
	cl_multi	Military subdivision chef-lieu common to other	Unique (0), Common (1)
Common to all	cl	Constituency chef-lieu id	5-digit INSEE code
	cl_name	Constituency chef-lieu name (capitalized)	String
	cl_name_prop	Constituency chef-lieu name (proper)	String
	cl_act	Constituency chef-lieu actual commune	Current (1), Former (2)
	cl_x	Constituency chef-lieu latitude	Lambert-93 projection
	cl_y	Constituency chef-lieu longitude	Lambert-93 projection

Table 7. Variables in annual military constituency-level nomenclatures. Values for mreg_type and msub_type: Only entire départements (1), Fractions of départements, entire arrondissements (2), Fractions of arrondissements, entire cantons (3), Fractions of arrondissements, entire communes (4), Fractions of communes (5).

Nomenclature	Variable	Label	Values
Common to all courts	cour	Court of appeal id	2 digits
	cour_name	Court of appeal name (capitalized)	String
	cour_name_prop	Court of appeal name (proper)	String
Courts of first instance	courinst	Court of appeal-first instance id	4 digits
	inst	Court of first instance id	2 digits
	inst_name	Court of first instance name (capitalized)	String
	inst_name_prop	Court of first instance name (proper)	String
	cl_ar	Court of first instance chef-lieu subprefecture	Not subprefecture (0), Subprefecture (1)
Penitentiary constituencies	pris	Penitentiary constituency id	2 digits
	pris_name	Penitentiary constituency name (capitalized)	String
	pris_name_prop	Penitentiary constituency name (proper)	String
Common to all	cl	Constituency chef-lieu id	5-digit INSEE code
	cl_name	Constituency chef-lieu name (capitalized)	String
	cl_name_prop	Constituency chef-lieu name (proper)	String
	cl_act	Constituency chef-lieu actual commune	Current (1), Former (2)
	cl_x	Constituency chef-lieu latitude	Lambert-93 projection
	cl_y	Constituency chef-lieu longitude	Lambert-93 projection

Table 8. Variables in annual judicial and penitentiary constituency-level nomenclatures.

Nomenclature	Variable	Label	Values
Circonscriptions	dep	Département id	2 digits
	dep_circo	Département-circonscription id	4 digits
	circo	Circonscription id	2 digits
	circo_name	Circonscription name (capitalized)	String
	circo_name_prop	Circonscription name (proper)	String
	circo_compo	Circonscription composition	See table notes
Academies	acad	Academy id	1 digit
	acad_name	Academy name (capitalized)	String
	acad_name_prop	Academy name (proper)	String
Labor inspections (1892–1940)	insp	Labor inspection id	2 digits
Dioceses	arch	Archbishopric id	2 digits
	arch_name	Archbishopric name (capitalized)	String
	arch_name_prop	Archbishopric name (proper)	String
	arch_type	Archbishopric type	Metropolitan See (1), Exempt (2)
	archdioc	Archbishopric-diocese id	4 digits
	dioc	Diocese id	2 digits
	dioc_name	Diocese name (capitalized)	String
	dioc_name_prop	Diocese name (proper)	String
	dioc_type	Diocese type	Archbishopric (1), Bishopric (2)
Common to academies, labor inspections, and dioceses	cl	Constituency chef-lieu id	5-digit INSEE code
	cl_name	Constituency chef-lieu name (capitalized)	String
	cl_name_prop	Constituency chef-lieu name (proper)	String
	cl_x	Constituency chef-lieu latitude	Lambert-93 projection
	cl_y	Constituency chef-lieu longitude	Lambert-93 projection

Table 9. Variables in annual electoral, academic, labor inspection, and ecclesiastical constituency-level nomenclatures. Values for circo_compo: Entire département (1), Entire arrondissement(s) (2), Entire canton(s) (3), Entire commune(s) (4), Fraction(s) of commune(s) (5), Paris arrondissement (6), Fraction of Paris arrondissement (7).

Variable	Label	Values
cassini	Cassini id	Numeric (1–5 digits)
insee	INSEE id	5 digits
com_name	Commune name (capitalized)	String
com_name_prop	Commune name (proper)	String
com_act	Actual commune code	Current (1), Former (2)
pop	Municipal population (census)	Numeric (see table notes)
ipop	Municipal population (interpolated)	Numeric
pop_flag	Municipal population flag	Numeric (see table notes)
cl_dep	Département chef-lieu	Not prefecture (0), Prefecture (1)
cl_ar	Arrondissement chef-lieu	Not subprefecture (0), Subprefecture (1)
cl_ct	Canton chef-lieu	Not bureau centralisateur (0), Bureau centralisateur (1)
cl_mcor or cl_mreg	Military corps or region chef-lieu	Not chef-lieu (0), Chef-lieu (1)
cl_div or cl_msub	Military division or subdivision chef-lieu	Not chef-lieu (0), Chef-lieu (1)
cl_cour	Court of appeal chef-lieu	Not chef-lieu (0), Chef-lieu (1)
cl_inst	Court of first-instance chef-lieu	Not chef-lieu (0), Chef-lieu (1)
trib_com	Commercial court	Not court siege (0), Court siege (1)
trib_prud	Labor court	Not court siege (0), Court siege (1)
cl_pris	Penitentiary constituency chef-lieu	Not chef-lieu (0), Chef-lieu (1)
cl_acad	Academy chef-lieu	Not chef-lieu (0), Chef-lieu (1)
cl_insp	Labor inspection chef-lieu	Not chef-lieu (0), Chef-lieu (1)
cl_dioc	Diocese chef-lieu	Not chef-lieu (0), Chef-lieu (1)

Table 10. Variables in annual commune-level nomenclatures (beyond those described in Tables 6–9). Extended missing value codes for pop: Not census year (.a), Commune nonexistent (.b), Archive lost (.d), Archive unreadable (.d), Commune not surveyed (.e). Values for pop_flag: Census data (1), Interpolated: not census year (2), Interpolated: archive lost (3), Interpolated: archive unreadable (4), Interpolated: commune not surveyed (5), Imputed: previous census (6), Imputed: next census (7). cl_insp only in nomenclatures 1892–1940.