

Realització d'una base mundial de dades geogràfiques de primer nivell administratiu

Guiu Asbert Rubio

Grau en Geografia, Medi Ambient i Planificació Territorial

Treball de fi de grau

Tutor: Xavier Pons Fernández

Curs 2022-2023

Grau: Geografia, Medi Ambient i Planificació Territorial

Curs acadèmic: 2022-2023

L'estudiant Guiu Asbert Rubio amb NIF 46422755-S

Lliura el seu TFG amb:

Títol Realització d'una base mundial de dades geogràfiques de primer nivell administratiu.

TUTOR: Xavier Pons Fernández

Declaro que el Treball de Fi de Grau que presento és fruit de la meva feina personal, que no copio ni faig servir idees, formulacions, cites integrals o il·lustracions diverses, extrems de cap obra, article, memòria, etc. (en versió impresa o electrònica), sense esmentar-ne de forma clara i estricta l'origen, tant en el cos del treball com a la bibliografia.

Soc plenament conscient que el fet de no respectar aquests termes implica sancions universitàries i/o d'un altre ordre legal.

Signatura de l'Estudiant

Vist-i-plau Tutor/a

Bellaterra, 21 de juny de 2023

Resum

Els mapes estan de moda. Mai havia sigut tan fàcil ni tan ràpid elaborar-ne, ahora l'interès que tenen com a eines de divulgació de dades geogràfiques és d'un gran valor. En un món globalitzat com el nostre, un mapamundi ha de servir per comparar i estudiar diverses parts del món, però sovint una representació a escala estatal mostra dades molt generalitzades sobre territoris molt més diversos i complexos. És per això que una representació mundial del nivell regional de cada país aporta una major precisió en la lectura del context global. De la mateixa manera, no existeix públicament una base de dades actualitzada que aglutini totes les regions del món. El present treball pretén resoldre aquest fet, a partir de la creació de la base de dades, la descripció i discussió dels criteris i processos que s'han seguit, mostrant-ne també la seva evident utilitat. A més, s'ha posat especial atenció a la correcció dels topònims en llengua catalana.

Paraules clau: cartografia, mapamundi, circumscripció administrativa, regió, SIG, base de dades.

Resumen

Los mapas están de moda. Su papel como herramientas de divulgación de datos geográficos es muy valioso, y su proceso creativo nunca había sido tan fácil ni tan rápido. Un mapamundi debe ser útil para comparar y estudiar varias partes del mundo en un mundo cada vez más globalizado, pero con frecuencia un nivel de representación estatal ofrece datos muy generalizados sobre territorios de mayor complejidad y diversidad. Una representación global del nivel regional de cada país fomenta una comprensión más precisa del contexto global. Además, en la actualidad no existe una base de datos accesible para el público general que cubra todas las regiones del planeta. El presente trabajo pretende solventar esta cuestión mediante la creación de la base de datos, la descripción y discusión de los criterios y procesos que se han seguido y la demostración de su evidente utilidad.

Palabras clave: cartografía, mapamundi, demarcación administrativa, región, SIG, base de datos.

Abstract

Maps are currently in style. It has never been so quick or simple to create and share original works. Their potential as a tool for science communication is also quite great. The conventional state-level mapping does not adequately depict the complexity and diversity of each country, revealing instead a less precise data generalization. In a globalized world, a world map must provide sufficient data to compare and study different regions of the world. A regional-level database provides a more comprehensive view of global dynamics because of this. Furthermore, the general public does not have at this moment a complete worldwide and regional database. This study aims to address this issue by creating a new database, outlining the various procedures and standards that were followed, and illustrating its value as a crucial tool.

Keywords: cartography, world map, administrative division, region, GIS, database.

Agraïments

En primer lloc, agrair tot el suport del meu tutor Xavier Pons, que no ha sigut poc.

També a l'Abel Pau, del CREAF, i a la Cristina Cea, del Departament de Geografia de la Universitat Autònoma de Barcelona, per la seva ajuda.

A l'Institut d'Estudis Catalans, i en particular al Sr. Joan Anton Rabella, per facilitar-me l'accés al nomenclàtor mundial.

Finalment a la meua família.

Índex

Objectius i preguntes de recerca.....	7
Metodologia.....	8
Antecedents: Mapes típics del Món de MiraMon.....	12
Criteris d'elaboració de la nova base de dades.....	13
Creació de la capa vectorial.....	14
Obtenció de la cartografia de referència.....	14
Processos d'automatització.....	16
Importació (JSONMM).....	17
Desagrupació dels nivells administratius (VecSelect).....	18
Modificació de l'estructura (GestBD).....	18
Mosaic.....	19
Assignació d'un identificador d'entitat a cada unitat administrativa (MiraDades) i generació d'etiquetes per a tots els polígons elementals (Etiqueta).....	19
Estructuració topològica (LinArc).....	20
Eliminació de micropolígons (MicroPol).....	21
Generació de grups de polígons (Ciclar).....	21
Supressió dels polígons petits (MicroPol).....	22
Generalització morfològica (GenVPol).....	23
Recollida de dades.....	25
Descàrrega dels mapes.....	26
Resultats i conclusions.....	26
Bibliografia i fonts d'informació.....	29
Annex.....	29

Objectius i preguntes de recerca

El meu objectiu principal és la creació d'una base de dades geogràfiques mundials de primer nivell administratiu, el que implica que contindrà informació de tots els països en un àmbit més profund que el nacional: a Espanya se'n diu comunitats autònomes, al Brasil estats i a França regions, però la idea és aproximadament la mateixa. Mitjançant la informació recopilada es podran comparar diferents regions de diferents països en àmbits com població i societat, salut o economia.

La divisió territorial dels estats moderns té orígens i recorreguts diferents. Per posar un parell d'exemples propers, les comunitats autònomes, divisió territorial d'Espanya de primer nivell, tenen orígens en els regnes i principats que formaven les Corones de Castella i d'Aragó, etc, i a França les regions actuals també es poden remuntar a les províncies de l'Antic Règim.

Les mateixes comarques catalanes, vigents en l'actualitat, són ja esmentades -si bé amb canvis notables- en el primer llibre sobre geografia catalana, publicat l'any 1600 per Pere Gil Estalella (1551-1622)¹.

A priori tenim diverses preguntes. La primera decisió que caldrà prendre és la divisió administrativa a emprar. La prioritat serà l'ús del primer nivell administratiu, però hi haurà casos on aquest nivell no serà el més adequat, potser perquè les dades estadístiques regionals fan servir una altra divisió territorial, potser per un elevat nombre de polígons que dificulta la visualització i gestió de la informació de la base de dades.

Una altra incògnita és quina representació fer per aquells estats amb un conflicte territorial, casos com Palestina, Israel, el Sàhara Occidental, Kosovo o Taiwan, on hauré de prendre una decisió a l'hora de representar una narrativa o una altra.

També és rellevant saber quina informació és més valuosa per produir una base de dades de major qualitat. Els polígons contindran dos tipus d'informació: dades més físiques, sobre el perímetre, l'àrea o l'altitud, i dades més socials, com el nombre d'habitants, indicadors econòmics com la renda per càpita, l'índex de desenvolupament humà, la taxa d'atur i d'ocupació, l'esperança de vida o el grau d'alfabetització.

La nova base de dades permetrà una representació instantània d'aquests indicadors, que facilitaran la feina d'investigadors o usuaris curiosos, que s'estalviaran fer la recerca prèvia.

Al meu coneixement, no hi ha ara mateix a internet una base de dades amb aquestes característiques, amb les divisions administratives de primer nivell de tots els països del món així com algun indicador per fer-les comparables. N'hi ha pels països de la Unió Europea, o algun recull parcial que barreja països i regions, però cap com el que em proposo en aquest treball.

Metodologia

Primerament, ha sigut necessari trobar una font de dades de la qual es poguessin obtenir les capes vectorials necessàries per a l'elaboració del mapa. Calia que la llicència no fos privada i ens permetés dur a terme el projecte i en acabat publicar-lo. En el cas dels productes ESRI, encara que tenen una bona base de dades de primer nivell administratiu, la seva llicència és privada i si féssim servir les seves dades, no les podríem presentar. Una bona alternativa és OSM-Boundaries.

OSM-Boundaries és un recurs que utilitza la llicència d'OpenStreetMap, que és un projecte col·laboratiu de creació de mapes amb llicència oberta. Facilita, per tant, l'extracció dels límits administratius dels països del món, extrets de la base de dades d'OpenStreetMap. Aquests límits van des del nacional, al regional i municipal. De cada país n'he extret la totalitat de capes, i posteriorment he fet la selecció per mostrar només la capa que més ens interessa. Ho he fet així perquè no sempre s'empra el mateix nivell i per poder disposar de tots els polígons que té la base de dades existent.

En referència a la toponímia feta servir el referent és el *Nomenclàtor Mundial* de l'Institut d'Estudis Catalans, realitzat entre 2016 i 2021, creat amb la finalitat d'establir una forma d'ús oficial per promoure la normalització toponímica en qualsevol medi en llengua catalana.

El *Nomenclàtor* conté els noms de tots els països, però només algunes regions: quan aquestes són més conegudes o emprades per la població catalana o quan hi ha més relació amb el país referent. Si bé hi tenim les regions d'Espanya (Saragossa, Lleó, Conca) o les de França (Aquitània, Provença, Borgonya), no s'arriba tan profund quan es tracta de països més llunyans. En aquests casos he seguit els criteris definits pel mateix Institut d'Estudis Catalans en l'elaboració del *Nomenclàtor*, publicats a desembre de 2022, on es defineix el camí a seguir pel que fa a topònims exògens amb forma en català, exògens de llengües amb alfabet llatí i exògens de llengües amb alfabet no llatí. Per aquest darrer grup, que és el més complex, proporciona informació sobre quin és el sistema de transcripció que s'ha de seguir en cada cas.

Cal dir que el *Nomenclàtor* no és perfecte, i en més d'una ocasió es porta la contrària. Per exemple al punt 4.3.2 del document dels criteris diu “La resta de noms que presenten diacrítics i lletres que podrien sobtar en català es poden simplificar eliminant els diacrítics i adaptant les lletres a l'alfabet del català, que ja és la pràctica habitual en els mitjans de comunicació”, però, en canvi, manté la majoria de les lletres inexistents a l'alfabet llatí, sense fer-hi cap modificació, el què de retruc genera molts problemes a l'hora de mostrar aquests caràcters. Ja en el mateix web del *Nomenclàtor* hi ha molts topònims que no es visualitzen correctament, i també d'altres que a la seva base de dades utilitzen una lletra no llatina mentre al web es representa amb la seva correspondència de l'alfabet llatí. Els topònims que el *Nomenclàtor* escriu amb aquestes lletres corresponen als països d'Azerbaidjan, Algèria, Afganistan, Kazakhstan, Kirguizistan, Moldàvia, Romania, Turkmenistan, Turquia i Vietnam.

Vietnam

Anam (Regió natural o històrica)	Annamita, serralada (Serra, altiplà, massís, etc.)	Ấp L?c (Nucli de població)
Buôn Ma Thu?t (Nucli de població)	Cà Mau, punta de (Punta, cap, península, etc.)	C?n Tho (Nucli de població)
Cao B?ng (Nucli de població)	Cotxinxina (Regió natural o històrica)	Đà L?t (Nucli de població)
Đà N?ng (Nucli de població)	Đ?ng H?i (Nucli de població)	H? Long, badia de (Golf, badia)
H? Long, badia de (Patrimoni de la UNESCO)	H?i Phòng (Nucli de població)	Hanoi (Capital)
H? Chí Minh (Nucli de població)	H?i An (Nucli de població)	Hu? (Nucli de població)
Indoxina (Regió natural o històrica)	Lào Cai (Nucli de població)	Mekong, boques del (Delta, estuari)
Mekong, el (Riu)	Nam Đ?nh (Nucli de població)	Ng?c Linh, pic (Cim, pic, tossal, turó, etc.)
Nha Trang (Nucli de població)	Parc Nacional Phong Nha-K? Bàng (Patrimoni de la UNESCO)	Phan Xi Pang, muntanya (Cim, pic, tossal, turó, etc.)
Quy Nho'n (Nucli de població)	Sông H?ng, delta del riu (Delta, estuari)	Sông H?ng, riu (Riu)
Tân An (Nucli de població)	Tây Ninh (Nucli de població)	Tonkin (Regió natural o històrica)
Tonquin, golf de (Golf, badia)	Trà Vinh (Nucli de població)	Tuy Hòa (Nucli de població)
Tuyên Quang (Nucli de població)	Vietnam (el) (País, territori, etc.)	Vinh (Nucli de població)
Yên Bái (Nucli de població)		

Objecte 1. Topònims del *Nomenclàtor* del Vietnam, amb caràcters no llatins i d'altres que ni tan sols apareixen.

Font: Elaboració pròpia.

En aquests casos he optat per escriure els topònims assimilant les lletres no llatines a les corresponents en alfabet llatí, tal com defineix el mateix criteri de l'IEC. Per exemple en el cas del quart topònim del Vietnam el primer caràcter ha esdevingut una D llatina.

També hi ha algun cas com les Illes Marshall, on el *Nomenclàtor* defineix alguns noms basant-se en la versió anglesa del topònim, quan la versió original (en llengua marshallesa) és fàcilment adaptable.

Per reunir la informació que omplirà la base de dades mundial he fet servir les fonts que proporciona cada país, ja que les fonts que comparen estadístics entre països sovint no passen del nivell nacional. Així per obtenir registres regionals i provincials he hagut de veure com proporcionava cada país les seves dades. En el cas dels països de la Unió Europea he fet servir l'Eurostat (<https://ec.europa.eu/Eurostat>).

L'Eurostat és un portal web on es representa informació regional a través de les divisions NUTS (Nomenclatura de les Unitats Territorials Estadístiques), que divideixen els estats membres en tres nivells segons el nombre d'habitants de la regió. Això implica que per cada cas ens caldrà emprar un nivell o altre, ja que els setze estats federats d'Alemanya són inclosos dins del nivell NUTS1, en tenir una població d'entre 3 i 7 milions, però els deu comtats de Lituània formen part de NUTS3.

Ha sigut un maldecap cercar informació regional dins els portals estadístics nacionals. Cada país té una organització diferent, amb una interfície única, uns estadístics únics i basant-se en els seus propis criteris. Tots els portals consultats estan inclosos en la bibliografia i les fonts d'informació.

Hi ha països que no actualitzen les dades poblacionals de les primeres divisions administratives des de fa una o dues dècades. Un d'aquests casos, Saint Vincent i les Grenadines, ho justifica pel fet que les seccions censals i les parròquies, que són la primera divisió administrativa, no concorden i, per tant, no saben quanta població viu en cada parròquia.

N'hi ha d'altres que tenen regions amb dades actualitzades, però ajunten valors de diversos districtes, o no mostren els valors de totes les seves regions de primer nivell administratiu; per tant, hi ha un descontrol evident.

El més gratificant per a treballar amb dades regionals ha sigut l'Eurostat i la seva organització en nivells administratius NUTS, que m'ha facilitat no només la cerca de dades sinó també la seva fàcil indexació a la mateixa base de dades. També he trobat altres portals que centralitzaven dades de múltiples països en l'àmbit regional, però sovint el problema ha sigut que la informació representada no era la més recent que es podia trobar als portals estadístics d'aquests països. És el cas del portal interactiu del Banc Mundial (<https://data.worldbank.org/>) on per exemple es podien visualitzar els indicadors d'Afganistan per l'any 2017 quan al mateix web hi havia les dades de 2020. Una eina que m'ha facilitat molt la cerca en l'àmbit de la població ha sigut el web *City Population* de Thomas Brinkhoff, que centralitza les dades poblacionals de tots els països del món en diversos àmbits (estatal, regional, municipal). Així i tot, només permet una visualització i còpia de les dades, la seva descàrrega no es contempla i els noms de cada regió segueixen els criteris propis de la pàgina.

Un altre dels problemes identificats en la descàrrega i indexació dels valors ha sigut la manca d'un criteri únic a l'hora d'anomenar les regions: Hi ha llocs on les regions porten el nom en la llengua original, d'altres que el porten en anglès, d'altres que a vegades van acompanyats del tipus de regió (Regió de Brussel·les, Prov. de Lieja) i a vegades l'ometen. Addicionalment, el nom del tipus de regió pot variar d'un lloc a un altre. Aquest fet és fàcilment resoluble a través de l'ús de codis estàndards com l'ISO 3166, però malgrat la seva llarga trajectòria (van ser publicats per primera vegada l'any 1974), és gairebé impossible trobar una base de dades que els inclogui.

Per tots aquests problemes exposats, he pres la decisió de només afegir un indicador dels que em proposava a la base de dades. Ho he compensat amb la creació de nous camps comparatius com són la densitat de població (creada a partir de la població i de l'àrea, en base als polígons sense generalitzar) o l'altitud topogràfica mediana. Ho he cregut convenient perquè afegir un nou estadístic econòmic o social, per la dificultat de trobar les dades de cada país, implica dedicar-hi uns tres dies de feina, i, al cap i a la fi, la prioritat d'aquest treball no és la de crear una base de dades gegant per no tocar-la mai més. La intenció és fer-la d'una manera que es pugui anar modificant i a mesura que sigui més fàcil trobar estadístics regionals, afegir-hi nous camps.

Antecedents

El primer que cal presentar és el mapa Límits administratius (vec.) de la Col·lecció de mapes típics del Món de MiraMon, elaborat l'any 1998 per ESRI. Representa 2605 unitats de primer nivell administratiu. Té 14 camps descriptius, amb informació molt bàsica: nom i codi de la unitat, tipus d'unitat administrativa, àrea, població en data de 1994 i país, regió i continent associat al polígon. A més, els polígons que pertanyen a la mateixa unitat no es presenten agrupats (per exemple Mallorca i Menorca es presenten en polígons individuals, cadascun amb indicació "Illes Balears" al camp "Name" de la "First-level administrative unit").

Fets notables són que els països dels Balcans no tenen cap classe d'organització territorial, el polígon representat és el del país. Passa el mateix amb Letònia, Estònia, Lituània, Moldàvia, Geòrgia, Azerbaidjan i Armènia. En el primer cas, els anys noranta foren una època molt turbulenta, on es succeïren múltiples guerres, fet que segurament propicià la falta d'informació cartogràfica. Pel segon cas, es tracta dels estats que sortiren de la dissolució de la URSS (Unió de Repúbliques Socialistes Soviètiques) l'any 1991, i, per tant, la falta d'informació cartogràfica es pot associar a la nova reorganització territorial.

El Regne Unit es divideix en els diferents regnes: Gal·les, Escòcia, Irlanda del Nord i Anglaterra, però no hi ha les divisions internes de cada nació. Aquesta és una mancança evident, perquè cada regne té les seves divisions administratives de primer nivell.

Addicionalment, hi ha moltes divisions administratives que avui dia, més de dues dècades després, ja no tenen vigència, bé per la implantació de noves organitzacions territorials o la modificació de les vigents. Aquest fet, juntament amb les mancances que presentava l'antic mapa, són les principals raons que ens han dut al procés d'actualització i creació d'un nou mapa, aquesta vegada d'elaboració pròpia, el que també sumarà a la gran eina que és MiraMon.

Els codis de les unitats que conté la base de dades d'ESRI s'anomenen Global Mapping International (GMI) i Federal Information Processing Standards (FIPS), aquests codis estàndards no són els més habituals en l'àmbit internacional (ho seria, per exemple, el codi ISO), però el mapa el va elaborar ESRI, una empresa nord-americana, i per això va emprar els codis GMI i FIPS, més comuns als Estats Units.

Criteris d'elaboració de la nova base de dades

Hi ha representats els 193 països que formen part de les Nacions Unides, dos que en formen part com a membres observadors: la Santa Seu i Palestina, i tres que en són membres de forma parcial: les Illes Cook, Niue i Kosovo.

Addicionalment, es representen Taiwan i el Sàhara Occidental de manera independent, casos especials perquè no tots els països en reconeixen l'existència, ni formen part de les Nacions Unides. Les illes Fèroe i Groenlàndia també es representen com a elements únics, en ser països independents dins del regne de Dinamarca. Les dependències i territoris d'ultramar del Regne Unit eren considerades entitats diferenciades a la base de dades OSM-Boundaries, però en la nostra base tot forma part del Regne Unit.

En la gran totalitat de països he emprat els polígons de primera divisió administrativa, no ha pogut ser en tots els casos, per dos motius principals: Primer, hi havia països que no disposaven de la capa vectorial que necessitava, és el cas d'Antigua i Barbuda (havien de ser les parròquies i he acabat amb les tres illes principals), Palestina (havien de ser les governacions, però he fet servir els dos territoris principals, Cisjordània i la Franja de Gaza) i les Seychelles (havien de ser els districtes, he fet servir el nivell estatal).

Segon, quan la divisió administrativa de primer nivell era municipal i suposava un nombre de polígons desproporcionadament elevats i hi havia opció d'utilitzar una altra divisió com les regions estadístiques, s'ha fet així. És el cas d'Islàndia, Letònia i Macedònia del Nord. En el cas d'Uganda la primera divisió administrativa són els 136 districtes, però per una major claredat he fet servir les quatre regions, que no tenen un rol administratiu definit.

En la cerca de dades, he fet una classificació en dos àmbits:

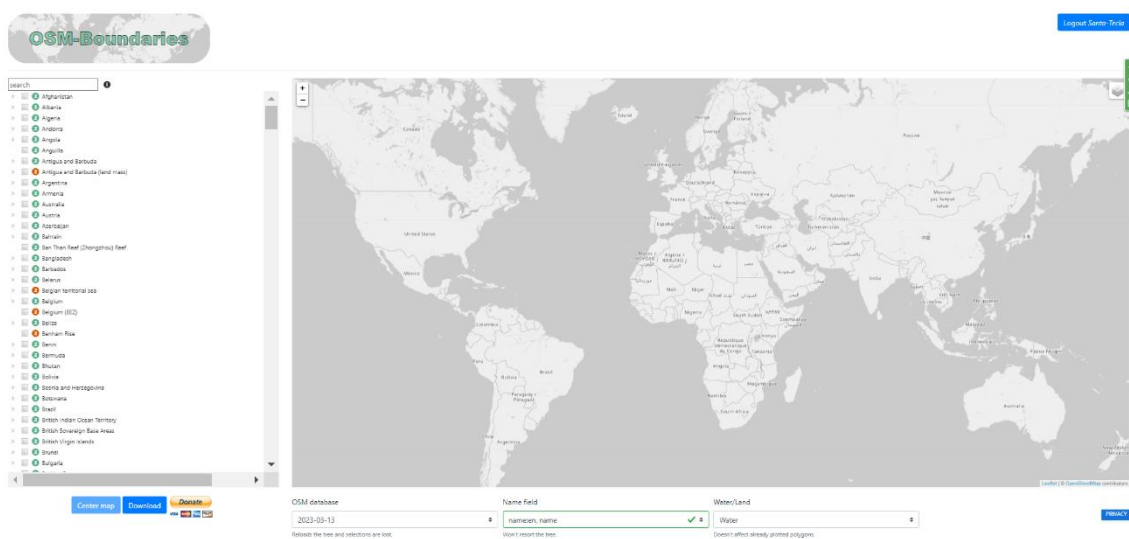
- Dades geogràfiques, que seran totes les que pertanyen a l'àmbit més físic del territori, relatives a l'altitud (màxima, mínima i mediana), al clima o a la temperatura. Són especialment importants perquè no es depèn dels departaments estadístics nacionals per obtenir-les, com que són fitxers ràster, engloben tot l'àmbit d'interès.

- Dades de l'àmbit humà, que seran la població i la densitat de població. En un futur hi ha la possibilitat d'afegir més sèries de dades. El criteri establert és que les dades han de ser disponibles per una gran majoria dels polígons, o bé per tots els polígons d'un sol continent. Tot i això, la prioritat és poder mostrar una visió global d'un indicador.

Creació de la base de dades

Obtenció de la cartografia de referència

L'eina OSM-Boundaries ens ha proporcionat gairebé tota la base cartogràfica necessària per al nostre projecte. En la captura de pantalla de sota s'hi pot veure la interfície gràfica, que ens permet baixar les dades d'una llarga llista de països i entitats territorials, amb un registre temporal que va fins a 2010, i diverses opcions com la inclusió de fronteres marítimes. Per baixar les dades cal autenticació d'usuari.



Objecte 2. OSM-Boundaries (<https://osm-boundaries.com/>) (22-05-2023)

El format dels arxius és GeoJSON, un format per a representar elements geogràfics, que el MiraMon permet importar amb el mòdul JSONMM. Són arxius que contenen polígons sobreposats seguint la lògica de diferents nivells administratius dins del país, des del nivell estatal (nivell 2) fins al municipal. Durant el procés de creació de la base de dades ha calgut fer alguna actualització del mòdul, la primera per la mida dels arxius a importar, la segona perquè els fitxers importats no documentaven correctament si el fitxer de polígons era de grups o no, i una tercera perquè algun dels fitxers contenia informació que no estava vinculada a cap objecte geogràfic, aspecte que ni ens interessava ni té cabuda en el model de dades de la major part de formats SIG.

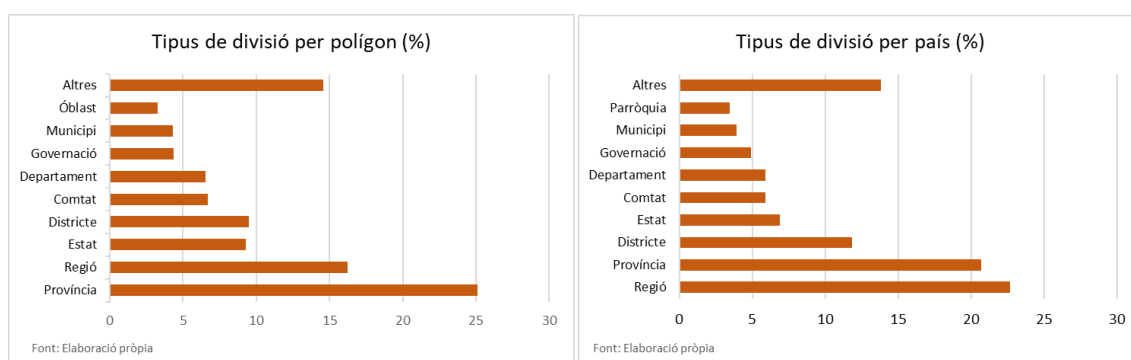
Una vegada descarregats els arxius de les 218 entitats territorials necessàries, i solucionades les incidències presentades en la descàrrega i importació al MiraMon amb l'ajuda de la Cristina Cea i l'Abel Pau, vaig procedir a estudiar cas per cas quina era l'organització territorial, anotant quin nivell de l'arxiu GeoJSON calia agafar, i mirant si hi havia els polígons adequats.

En la majoria dels casos (79,0 %), la base de dades OSM-Boundaries assignava el nivell 4 als polígons de primer nivell administratiu, però hi ha algunes excepcions: vint-i-tres països (11,0 %) utilitzen els polígons del nivell 6, cinc fan servir els de nivell 2 o nacionals (2,5 %), nivell 5 (2,5 %) i nivell 6 (2,5 %), tres utilitzen els de nivell 3 (1,5 %) i finalment n'hi ha un que fa servir el nivell 7 (0,5 %).

Per molt que la base de dades OSM-Boundaries em facilités molt la feina, vaig haver de treballar bé les dades per poder-les fer servir. Dels 201 països, vaig haver de modificar 47 arxius, gairebé un 25 %. Aquestes modificacions van anar des d'un senzill canvi en l'etiqueta del nivell, per treure o posar els polígons pertinents, fins a la creació de la capa administrativa regional a partir dels mateixos municipis o la unió o divisió de polígons per fer les regions adequades. Els casos on la modificació ha sigut més important han sigut nou, sovint coincidint amb països en conflicte (Palestina, Sàhara Occidental, Marroc, Sudan del Sud). A l'annex he inclòs una llista dels canvis efectuats per cada país.

Després de fer un repàs exhaustiu vaig obtenir una llista final de 3373 grups de polígons, que representen un augment del 29,5 % partint dels polígons de la base regional existent a MiraMon (2605 polígons).

L'estudi de les divisions administratives mundials m'ha permès fer una classificació dels 3373 grups de polígons segons el tipus de divisió administrativa. La província representa un 25 % del total. A continuació hi ha la regió (16 %), l'estat (9 %), el districte (9 %) el comtat (6 %), el departament (6 %), la governació (4 %), el municipi (4 %), i l'óblast (3 %). La resta de casos tenen un pes inferior al 3 %, el que representa menys de 100 elements.



Objecte 3. Tipus de divisió administrativa, sobre el total de polígons o dels països.

Font: Elaboració pròpia

Si en comptes de fer l'anàlisi sobre els polígons el fem sobre les 203 entitats territorials (països i estats) el que veiem és diferent. La regió és la divisió administrativa més freqüent i representa un 23 % del total de països. La segueixen la província (21 %), el districte (12 %), l'estat (7 %), el comtat (6 %), el departament (6 %), la governació (5 %), el municipi (4 %) i la parròquia (3 %). Per tant, podem veure que les divisions majoritàries a escala mundial són les regions i les províncies, si bé la tendència és que els països que fan servir províncies en tinguin un major nombre que en el cas de les regions.

L'Institut d'Estudis Catalans defineix la província com: "Unitat administrativa territorial de primer grau pròpia de molts estats", i també "Territori biogeogràfic en què es divideix una regió, caracteritzat per la possessió d'algunes espècies pròpies i que inclou diversos sectors"; aquesta última afirmació defineix els criteris amb què sovint els països estableixen la seva divisió regional, per bé que és el concepte de regió el que s'identifica més a l'existència d'identitats particulars i històriques.

La regionalització dels estats no sempre es basa en criteris identitaris, sovint es basa en les mateixes característiques geogràfiques (el més comú en aquests casos és la creació de regions segons l'orientació, com ara Kazakhstan Nord, Est i Oest, però també segons les barreres físiques naturals com poden ser illes, arxipèlags o atols). Una característica de la regionalització de molts països caribenys és l'ús de les parròquies, característica compartida amb el Principat d'Andorra. Aquesta divisió és herència de les divisions eclesiàstiques, que en la majoria dels països ja han perdut aquest rol, que encara trobem a Andorra, les Barbados, Dominica, Grenada, Jamaica, Saint Kitts i Nevis i Saint Vincent i les Grenadines.

Hi ha divisions que són característiques de regions transfrontereres amb trets culturals compartits. És el cas de les governacions dins del món islàmic, herència de l'imperi otomà, també anomenades valiats. Actualment, són emprades per Bahrain, Egipte, el Iemen, l'Iraq, Jordània, Kuwait, el Líban, Oman, Síria i Tunísia. En el món exsoviètic la regió característica és l'óblast, que encara és vigent a Belarús, Bulgària Rússia i Ucraïna.

Processos d'automatització

Un altre dels objectius del treball era aprendre processos d'automatització que ens permetessin treballar amb molta més agilitat. En cas d'haver de tornar a executar algun pas anterior, l'execució esdevé fàcil i ràpida.

Aquests processos estan estretament relacionats amb una clara organització dels arxius; el fet que els directoris i arxius segueixin un ordre ben definit permet que les línies de comanda funcionin correctament. Escriurem les ordres a la llibreta de Windows i ho convertirem en un fitxer “.bat”, que és l’extensió que permet l’automatisme de les ordres que especifiquem.

- Importació (JSONMM)

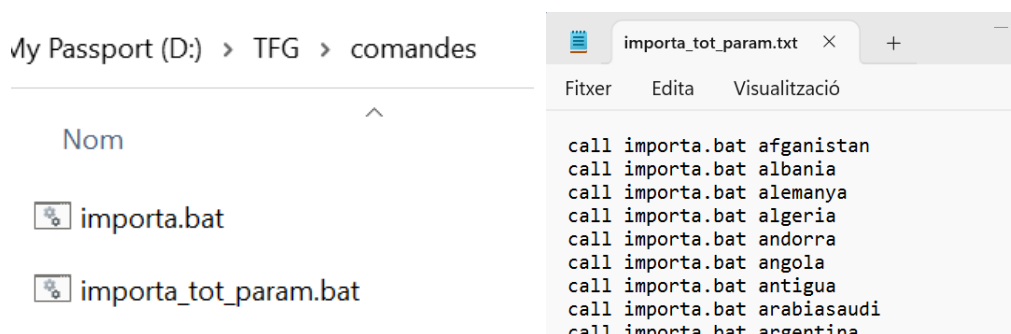
La primera ordre és la que serveix per importar els fitxers GeoJSON al format que en permet l’obertura.

c:\MiraMon\JSONMM 1

D:\TFG\descomprimits\%1.geojson\%1. geojson

D:\TFG\descomprimits\%1. geojson\%1.mmm /DT=c:\MiraMon /SAC

La sintaxi diu el següent: En primer lloc, i seguint l’ordre de les línies, crida l’eina que des de MiraMon executa la importació. En segon lloc, indica on es troba l’arxiu que ha d’importar. El símbol %1 permet, en trobar-se tots els països al mateix directori i seguint l’estructura “nom.geojson” (on “nom” és el nom del país), que amb només aquesta ordre es faci per tots els països, fent servir la llista dels mateixos que es pot veure a la imatge inferior. En tercer lloc, s’indica on s’ha d’escriure la importació del nou arxiu i el format que ha de tenir (.mmm, o Mapa del MiraMon). Atès que un fitxer GeoJSON pot contenir objectes de diferent natura (punts, línies o polígons), la importació es fa a un Mapa del MiraMon, el qual pot tenir barreges de tipus d’objectes vectorials, a més de capes ràster de diferent natura, enllaços a geoserveis, etc).



Objecte 4. Arxiu “importa.bat” on hi ha l’ordre principal i arxiu “importa_tot_param.bat” amb la llista dels països, que crida el primer arxiu i hi substitueix %1 per cada nom indicat. El verb anglès “call” permet cridar el fitxer de tipus .bat a continuació i que, en acabar l’execució, continuï executant altres línies del programa, fent altres crides a “call” en aquest cas. Font: Elaboració pròpia.

- Desagrupació dels nivells administratius (VecSelec)

La segona ordre fa la desagrupació dels paquets de polígons segons el nivell administratiu classificat per la base de dades OSM-Boundaries.

C:\MiraMon\VecSelec

D:\TFG\descomprimits\%1.geojson\%1\Polygon.pol

D:\TFG\mon\niv_4\Lim_adm_%1_nivell_4.pol /COND1_CAMP=ADMIN_LEVEL

/COND1_OP=EQ /COND1_VALOR=4 /AGRUPA_NO_SELEC /EMANCIPA /DT=C:\MiraMon
/SAC

La sintaxi primer crida l'eina de selecció (que permet la desagrupació, ja que cada nivell o grup té un codi diferent), després li diu on hi ha l'arxiu que conté les diferents capes de les quals separar la que ens interessa i, finalment li diu quin nom portarà l'arxiu resultant, juntament amb la condició que ha de complir per a fer-se la selecció (en aquest cas, que el camp "ADMIN_LEVEL" ha de tenir un valor igual a 4) Aquesta ordre es repetirà tants cops com nivells hi hagi, en el meu cas ho he fet des del nivell 2 (nacional) fins al 8 (municipal).

- Modificació de l'estructura de la taula (GestBD)

La tercera i última ordre automatitzada permet modificar l'estructura de les taules (.dbf), emprant el mòdul de gestió de bases de dades de MiraMon (GestBD), concretament l'opció número 12 per moure un o diversos camps consecutius. La seva finalitat és fer que totes les taules de cada país amb què treballem tinguin una mateixa estructura, condició necessària per tal de poder fer un mosaic de tots els països.

C:\Miramon\GESTBD_64 12

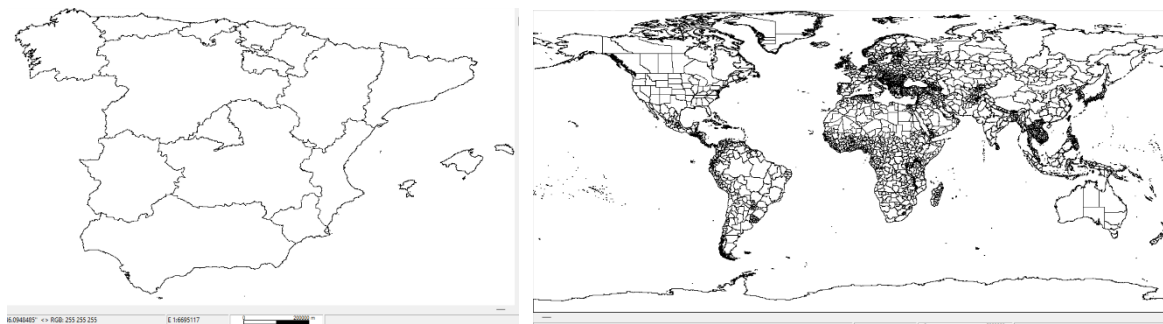
D:\TFG\mon\niv_4\Lim_adm_%1_nivell_4P.dbf ADMIN_LEVEL ADMIN_LEVEL NAME_EN
/SAC

La sintaxi en aquest cas en primer lloc crida el mòdul del MiraMon encarregat d'actuar i, en segon lloc, indica la ubicació de la taula (.dbf) a modificar, el nom del camp afectat, altra vegada el camp afectat (per indicar on comença l'acció) i, finalment, el nom del camp després del qual es desplaçarà el primer camp.

A l'hora de fer el mosaic, una diferent mida dels camps no ocasionava problemes. Sí que ho feia l'ordre dels camps, com ja hem dit, i el tipus del camp. Hi va haver uns pocs casos on diversos països tenien un mateix camp que en un cas era numèric i en l'altre de text (de caràcters). Aquests casos els vaig arreglar manualment i van esdevenir tots ells de text, que era el cas majoritari. Els països en qüestió van ser Eritrea, Palau, Palestina, Saint Lucia, Samoa, Vanuatu, Tuvalu i la Santa Seu.

Mosaic

L'eina per fer el mosaic serveix per unir dues o més capes de polígons. És necessària, per tant, per a aglutinar els 218 arxius en un de sol, que serà la base de dades definitiva.



Objecte 5. Divisió administrativa de primer nivell d'Espanya i mosaic de tot el món. Font: Elaboració pròpia.

A partir d'aquest pas, deixem de treballar amb els països aïllats per fer-ho en conjunt (a l'esquerra, divisió administrativa de primer nivell d'Espanya, les comunitats autònomes, i a la dreta el mosaic de totes les divisions administratives del món).

Assignació d'un identificador d'entitat a cada unitat administrativa (MiraDades) i generació d'etiquetes per a tots els polígons elementals (Etiqueta)

Els polígons de la base original són, com hem dit, polígons explícits, sense topologia, i formant grups (per exemple, el de les Balears, compost pels polígons de Mallorca, Menorca, etc). Atès que en els següents procediments prendrem totes les línies de frontera duplicades i les convertirem a línies simples (arcs i nodes) per tornar-les a Ciclar en polígons elementals i, finalment, reagrupar aquests polígons en les unitats administratives originals (com ara les Balears), caldrà assignar un identificador d'entitat a cada unitat administrativa (cosa que fem amb el gestor de taules del MiraMon, el MiraDades, que ens permet crear un nou camp a la taula de polígons (P.dbf) i assignar un valor diferent a cadascun dels 3373 polígons originals). A continuació caldrà generar etiquetes en forma d'entitats de tipus punt per a cada polígon; aquesta darrera acció es fa amb el mòdul Etiqueta del MiraMon, el qual, a més, quan la base presenta polígons que poden estar formant grups genera automàticament una etiqueta a cada polígon elemental (una etiqueta a Mallorca, una a Menorca, etc).

Aquest procés dona lloc a una base de més de 700 000 punts (etiquetes) que contenen no només els atributs inicials (nom del país, de la regió, etc) sinó també l'identificador de grup que permetrà, en el procediment final, tornar a generar els grups de polígons que són les divisions administratives amb què volem treballar, ja sense possibles problemes topològics.

Estructuració topològica (LinArc)

Recordem que els polígons de la base de dades originals són polígons explícits, sense topologia. Això vol dir que en zones de frontera totes les línies i vèrtexs són duplicats, el que suposa un augment notable de la mida de l'arxiu i, encara més important, comporta un risc elevat d'incoherència en haver dues línies que pretenen descriure el mateix.

L'estructuració topològica permet fer anàlisi espacial ben feta, és garantia d'unicitat espacial i no permet encavalcaments o esclatxes entre polígons; això ho aconsegueix mitjançant una estructura arc-node de totes les vores. Al MiraMon l'eina utilitzada és el LinArc.

En ser una base de dades tan gran, reconstruir-la de cop era una feina massa pesada, així que s'ha fet a partir de les regions de les Nacions Unides (Àfrica del Nord, Europa del Sud, etc.) per tal de disminuir la càrrega suportada pel programa i, per tant, poder distribuir-la en diversos processadors de l'ordinador a través de sengles finestres de comandes (consola CMD del Windows), tot paral·lelitzant el procés. Addicionalment, es va dividir la regió d'Europa del Nord en Noruega, Suècia, Finlàndia i la resta, en ser aquests polígons molt més pesats, a causa de l'elevat nombre de vèrtexs de les zones costaneres (fiords) i l'elevadíssim nombre d'illes. Un cop feta l'estructuració topològica de les regions, s'ha fet pels continents i finalment amb aquests s'ha fet l'estructuració final.

A continuació va ser necessari reconstruir, amb l'eina Ciclar del MiraMon, els més de 700 000 polígons elementals (per exemple, les Balears es presentaven en polígons individuals per a Mallorca, Menorca, etc.); a continuació es va assignar, amb l'eina AtriTop del MiraMon, a cada polígon elemental l'etiqueta que havíem preservat en el procediment anteriorment descrit. Amb això cada polígon elemental tenia de nou els atributs temàtics (nom del país, de la regió, etc) d'interès i, a més, l'identificador dels objectes originals per a poder tornar a generar els grups que són les divisions administratives amb què volem treballar

En fer l'estructuració topològica i tornar a fer els polígons a partir dels arcs, els grans sectors de mars i llacs interiors (Mar Caspi, Llac Victòria, etc.), sovint exclosos de qualsevol regió en la base OSM original, van esdevenir polígons, i vaig creure convenient etiquetar-los com a Aigües Continentals. Malgrat tot, la base no representa totes les aigües continentals de mida considerable, ja que hi ha molts països que no les representen cartogràficament com a elements independents, mentre que hi ha casos on una massa d'aigua que comparteix frontera entre diversos països és dibuixada només per un dels països (no sempre un projecte de cartografia col·laborativa com l'OSM aconsegueix seguir criteris del tot

homogenis; tot i això, el nombre d'inconsistències va ser relativament petit). En total són catorze casos que incloem en la nostra base com a elements independents.

Ahora, tres polígons entre Egipte i el Sudan, i entre el Sudan del Sud i Kenya no formen part de cap país en ser considerats territoris en disputa. Així ho he volgut representar amb l'etiqueta Territori en disputa. Aquestes dues inclusions porten el total de polígons de la nostra base de dades a 3390.

Eliminació de micropolígons (MicroPol)

La resta d'esclatxes i imperfeccions que van sorgir de l'estructuració topològica van ser eliminades utilitzant l'eina d'eliminació de polígons, MicroPol. En efecte, tots aquells polígons ciclats *ex novo* que no havien rebut cap etiqueta és que eren o petites esclatxes o petits encavalcaments entre polígons provocats perquè a la base original OSM les fronteres no eren coherents. Es van detectar més de 13 000 petites incoherències d'aquest tipus, com acostuma a passar quan el productor de la informació no utilitza eines de control topològic (i/o utilitza formats sense topologia com el GeoJSON).

El criteri per eliminar totes aquestes incoherències va ser generar un camp de nom ELIMINA, de tipus lògic, i que va prendre un valor FALS en tots els polígons excepte en aquells sense identificar de grup; els polígons que van desaparèixer es van fusionar amb els polígons que compartissin l'arc comú més llarg amb el polígon a eliminar.

Després d'executar l'eina d'eliminació de micropolígons la nostra base queda amb les següents característiques.

- Arcs: 713 060
- Nombre total de vèrtexs: 79 046 138
- Arc amb més i menys vèrtexs: 524 830 i 2 (mediana: 15)
- Polígons: 705 326 (sense comptar el polígon zero)
- Nombre d'illes: 701 283
- Índex de niament màxim: 3
- Nombre màxim d'arcs per polígon: 616 (mediana: 1)
- Nombre màxim d'anells per polígon (prèviament a agrupar-los per regions): 595 (mediana: 1)

Generació de grups de polígons (Ciclar)

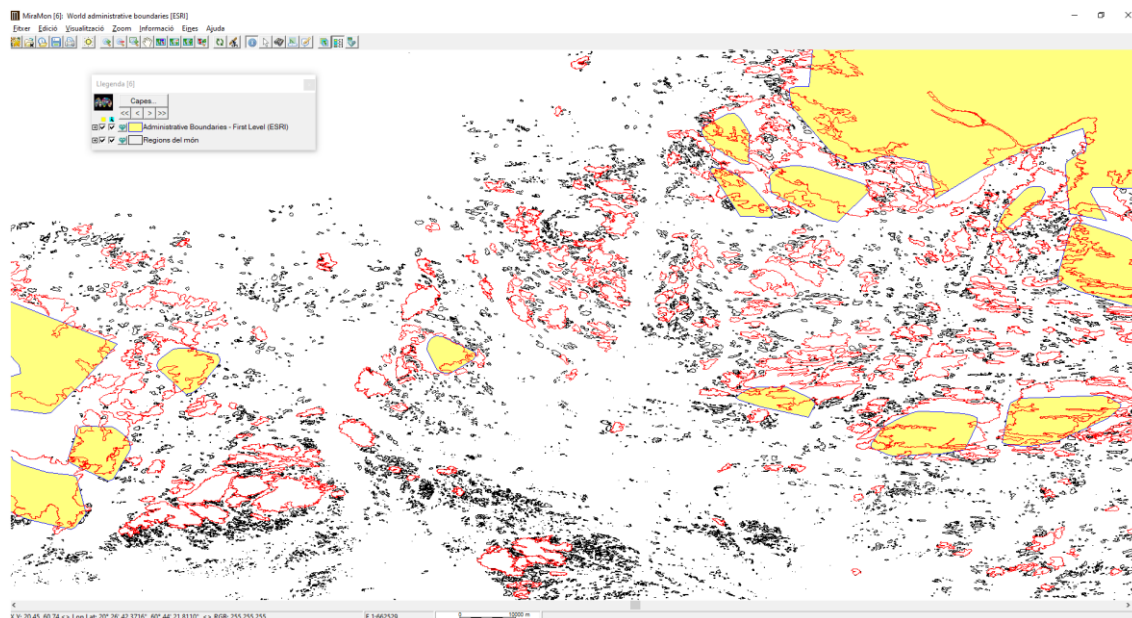
Per tal de generar la base final es van tornar a crear els polígons de grups formats per polígons elementals que tenien un mateix identificador de grup. Aquesta operació es va fer amb el mòdul Ciclar del MiraMon.

La base finalment aconseguida té un gran detall i és, per tant, ideal per a abordar la cartografia de pràcticament qualsevol lloc del món a una escala raonablement gran. Tanmateix, va semblar oportú generar una versió més lleugera a través de generalització cartogràfica per quan no és necessari tan detall a la morfologia de les costes o dels límits administratius, o no cal representar la miríada d'illots petits que hi ha al món.

Supressió dels polígons petits (MicroPol)

L'eina MicroPol es va executar una segona vegada, ara per alleugerar més la base de dades, eliminant aquells polígons massa petits per ser d'interès en una base de dades mundial. El polígon regional més petit correspon a la Santa Seu, que fa al voltant de 49 hectàrees. Aquest ha sigut el criteri per a decidir la mida de polígon mínima a conservar.

La supressió dels polígons va fer que desapareguessin tres regions de la nostra base. La raó essent que la Santa Seu feia unes 49 hectàrees, però en un sol polígon, mentre que les tres regions, tot i superar la mida de la Santa Seu quan agrupades, en desagrupar-les no hi havia ni un sol polígon més gran, per la qual cosa van ser eliminades. Per resoldre el cas, es van recuperar els polígons més grans de cada agrupació, excepte pel cas de les Places de Sobirania, part de la divisió territorial d'Espanya, en el qual s'han recuperat quatre polígons, un per cada grup principal, agrupats posteriorment amb el mòdul Ciclar.



Objecte 6. Obtenció d'una base generalitzada, més lleugera. Font: Elaboració pròpia.

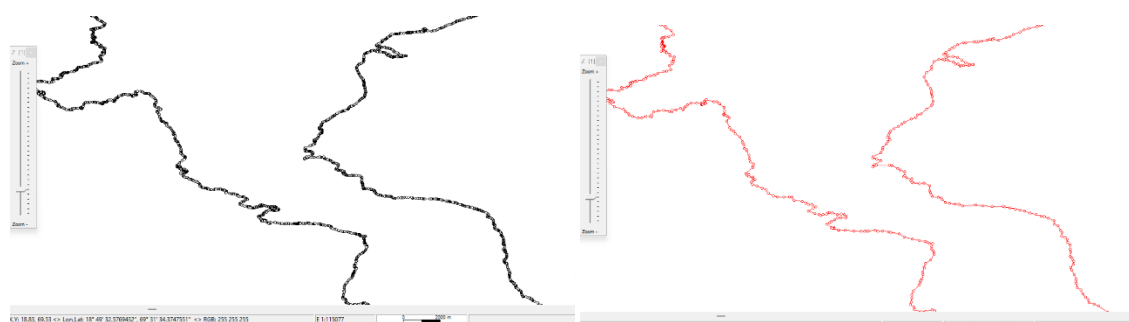
En la captura anterior es pot veure un sector de la costa finlandesa i l'impacte de l'eliminació de micropolígons. En vermell hi ha els polígons que s'han conservat, en negre els que s'han eliminat, i en groc hi ha l'antiga base de dades feta per ESRI l'any 1998. És notable l'augment de precisió cartogràfica respecte la base anterior, fins i tot en la versió generalitzada.

Generalització morfològica (GenVPol)

La següent i última eina que hem aplicat a la base de dades és la generalització de vèrtexs. L'objectiu és fer-la més lleugera simplificant les formes dels polígons a partir de l'eliminació de vèrtexs. En ser una base mundial, no és necessari que les fronteres territorials o la línia de costa tinguin un grau de detall molt elevat. La prèvia estructuració topològica ha permès abordar la generalització sense que apareguin inconsistències en les fronteres entre polígons, en resultar-ne una única frontera, que a la vegada resulta en un fitxer molt menys pesant.

L'eina que fem servir és el GenVPol, que aplica a cada vèrtex un filtre de mitjana mòbil sobre les seves coordenades XY i a continuació aplica l'algoritme de simplificació de vèrtexs de Douglas-Peucker. La primera funcionalitat el que fa és recalcular la posició dels vèrtexs, suavitzant-ne el traçat (sense excedir un primer paràmetre de tolerància) mentre que la segona, basant-se en els canvis de la primera, elimina vèrtexs allà on la distància entre la línia simplificada i l'original és menor a un segon paràmetre de tolerància indicat.

Els paràmetres escollits per a cada opció són: Per a la grandària de la mitjana mòbil 7 vèrtexs amb una tolerància de 0.018 graus (equivalents a uns 2 km) i per a l'algoritme Douglas-Peucker una tolerància de 0.0045 graus (equivalents a uns 500 m).



Objecte 7. Obtenció d'una base generalitzada, més lleugera. Font: Elaboració pròpia.

La generalització en polígons amb una alta complexitat de vèrtexs, tal com es pot veure a les imatges superiors, no acaba de funcionar de la millora manera amb els algoritmes actuals, però el resultat és satisfactori. Només cal veure la gran quantitat de vèrtexs que desapareixen d'una base (en negre, sense generalitzar) a l'altre (en vermell, generalitzat).

Després de fer una primera generalització es va fer una segona generalització, més simple, per tal de suavitzar els arcs. Això es va fer aplicant altra vegada la generalització amb una tolerància de 0.018 graus (~2 km), sense aplicar l'algoritme Douglas-Peucker, per la qual cosa no s'elimina cap vèrtex addicional.

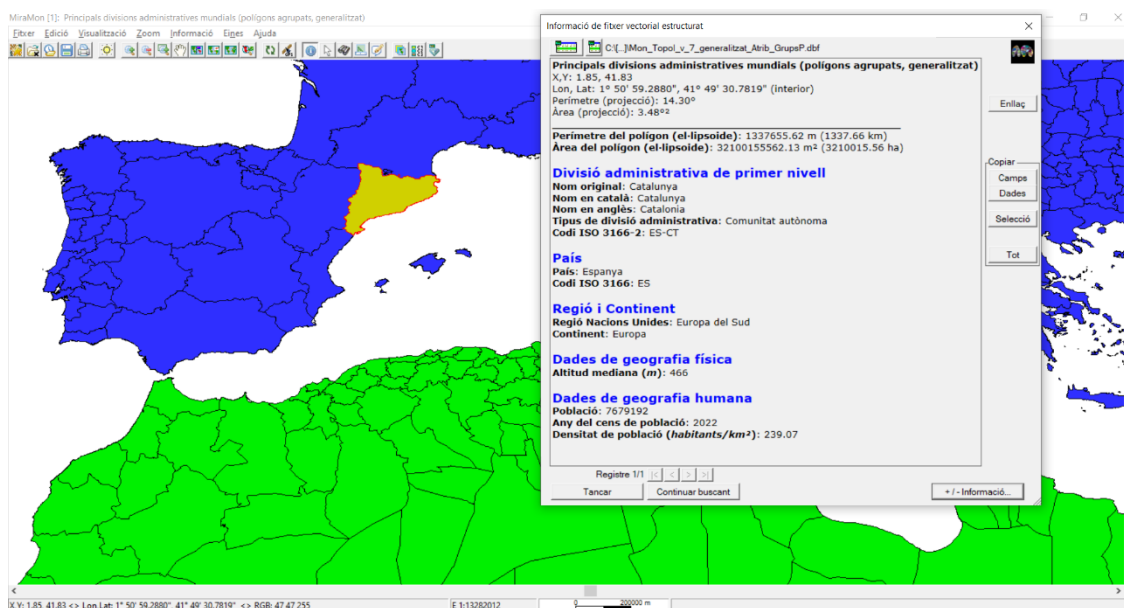
El procés de generalització morfològica ha fet que els vèrtexs totals passin de 69 793 616 a 4 601 389 (fins a quinze vegades menys!), i la base passi d'ocupar 15 MB a 3 MB, el que resulta en una millora en la rapidesa de càrrega, malgrat la seva complexitat.

Recollida de dades

Obtinguda la base de dades definitiva, se li va afegir contingut per donar-li més robustesa. El primer indicador va ser la població, que vaig haver de buscar per diferents fonts estadístiques, tal com he explicat amb detall a la metodologia. A partir d'aquesta dada vaig calcular la densitat de població de cada regió, prenent com a àrea la del polígon sense les vores generalitzades. Si fes servir l'àrea dels polígons generalitzats, la dada no seria correcte en tractar-se d'àrees simplificades.

Per enllaçar les dades ràster amb la base de dades es va fer a partir d'una combinació analítica de capes (CombiCap), que per a cada polígon calcula automàticament els estadístics de les dades que vulguem analitzar, en el nostre cas l'altitud. L'eina CombiCap calcula la mitjana, la desviació estàndard, la variància, el sumatori, el mínim, el màxim, el Rang, el 1r quartil, la mediana i el 3r quartil. En el context d'aquest treball de final de grau, s'ha optat per utilitzar una base d'elevacions mundials simplificada (Col·leccions preferides del MiraMon) i, per tant, s'han obviat altres estadístics que no siguin la mediana, per tal com no serien prou fiables. La intenció és de tornar a realitzar el càlcul sobre la base mundial a 30 m d'Aster.

En paral·lel, es van omplir les metadades d'informació important: resums de contingut i propòsit de la base de dades, i origen i temporalitat de les dades representades, en català i anglès. Aquest procés es va completar amb la modificació de la interfície gràfica de la consulta per localització, amb l'objectiu de fer-la més agradable a la vista de l'usuari.



Objecte 8. Consulta per localització. Font: Elaboració pròpia

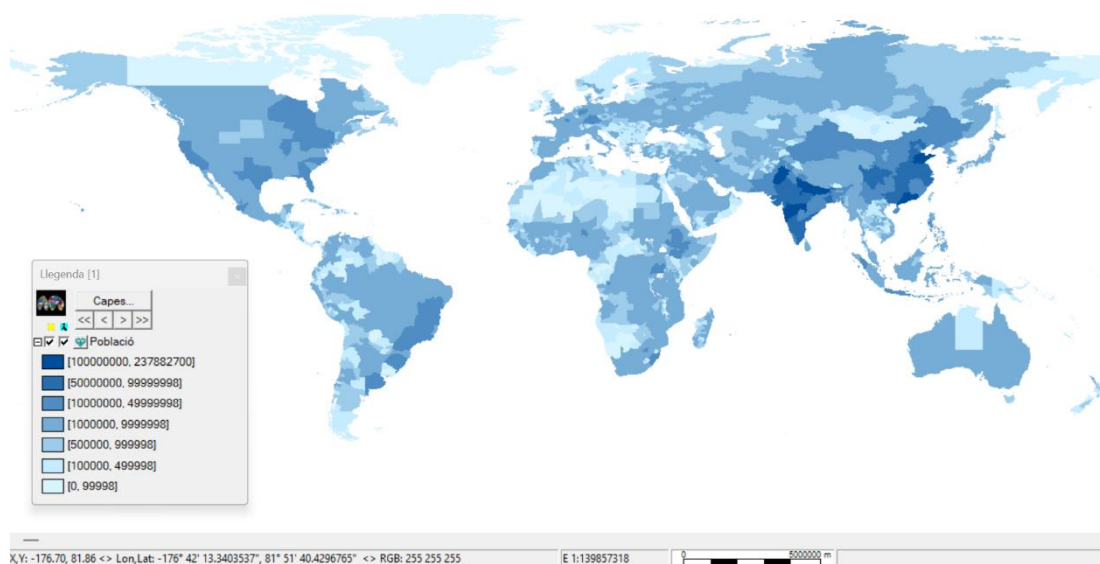
Descàrrega dels mapes

Hem posat a disposició dos mapes, un utilitza els polígons originals, sense generalitzar, i l'altre, més lleuger, utilitza els polígons generalitzats. Ambdós arxius tenen els polígons agrupats i les solucions i processos aplicats fins a la generalització. Es poden consultar a través del següent enllaç:

www.infoambiental.grumets.cat/LimitsAdmMundials/Llegiu-me.txt

Resultats i conclusions

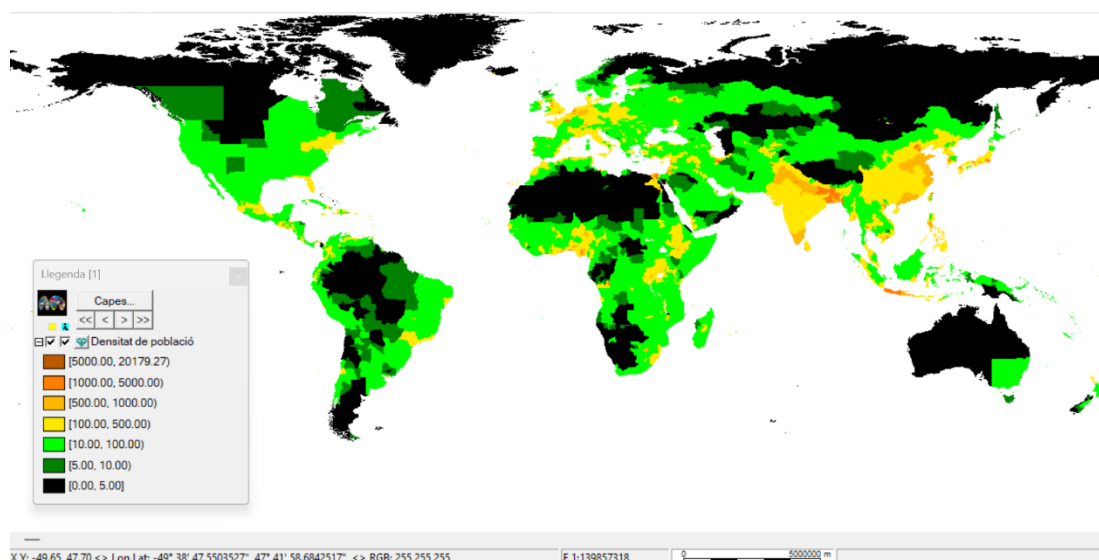
Volia concloure el treball amb tres mapes que es poden realitzar amb la nostra base de dades. El primer és el mapa mundial de població segons la regió.



Objecte 9. Mapa de població, per divisió administrativa de primer nivell. Font: Elaboració pròpia.

A la llegenda hi ha nou classes, la més petita va dels 0 als 99 000 habitants, mentre que l'última va dels 100 milions al valor màxim (237 milions, corresponent a l'estat indi d'Uttar Pradesh). El valor mínim és 0, per les múltiples illes deshabitades que formen regions. Hi ha dos factors que s'hi poden observar: El primer és la correspondència dels grans buits continentals a les zones desèrtiques existents (el Namib, el Sàhara, Atacama, el Gobi i els deserts polars). El segon és la relació entre les zones costaneres i una major població, ben observable arreu, però especialment al Brasil, a l'Amèrica del Nord, al Mediterrani i a l'Àfrica del Sud.

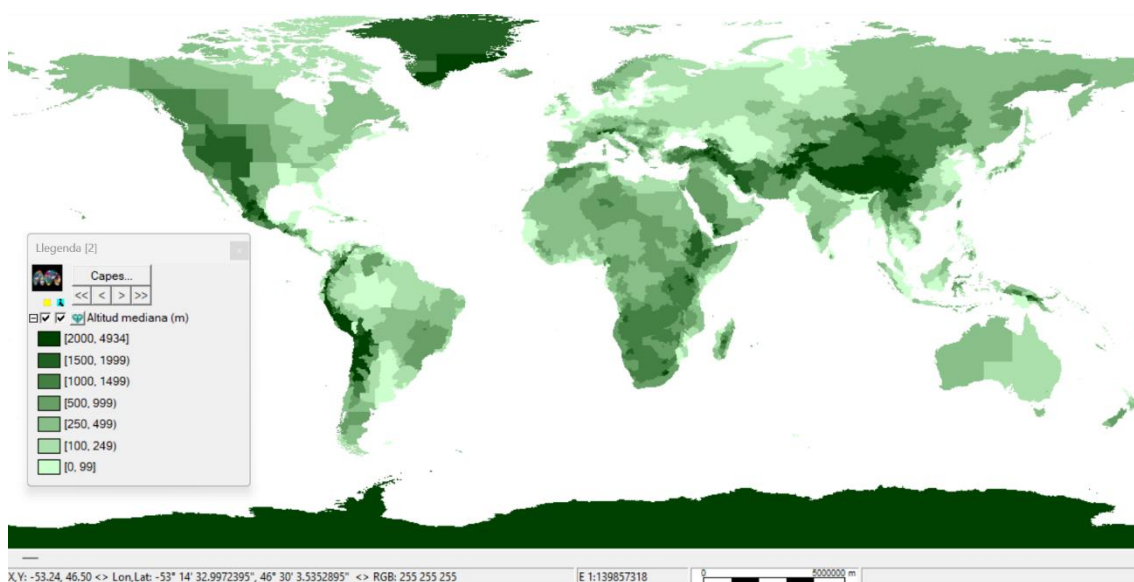
En segon lloc hi ha el mapa del món regional segons la densitat de població (població/km²).



Objecte 10. Mapa de la densitat de població, per divisió administrativa de primer nivell. Font: Elaboració pròpia.

En aquest cas hi ha set classes, que van de la primera categoria (0 a 5 habitants per quilòmetre quadrat) a l'última (5000 al màxim, que en aquest cas són la regió metropolitana de Manila, capital de les Filipines, i la governació de Beirut, capital del Líban, amb vora 20 000 habitants per quilòmetre quadrat cadascun). Es poden veure les zones on més gent viu en un espai més reduït: l'Europa Central (Bèlgica, els Països Baixos i Alemanya), la costa xinesa, el Nord de l'Índia o el Golf de Guinea.

Finalment, el mapa d'altitud mediana en metres, segons la regió.



Objecte 11. Mapa de l'altitud mediana, per divisió administrativa de primer nivell. Font: Elaboració pròpia.

La vista és similar als mapes altimètrics, però facilita una major comprensió del context regional. Hi són representats els elements de relleu més destacats del món: els Andes, les Rocalloses, el Rift Valley africà, l'Himàlaia, el Zagros i els pols.

De moment, a part de visualitzar el món per països, regions o continents, aquestes són les úniques panoràmiques de les quals disposa la nostra base, però espero que en un futur es pugui completar i actualitzar lliurement, partint d'una molt bona base i una feina ben documentada.

En general estic molt content amb el producte que he dut a terme, els objectius definits a l'inici del treball s'han complert satisfactòriament, i si bé hi ha reptes que no s'han pogut completar, com ara la inclusió de més indicadors socioeconòmics i físics, s'han assentat les bases per a poder fer-ho en un futur.

El meu grau d'aprenentatge al llarg del treball ha sigut molt elevat, ja que el meu tutor, en Xavier Pons, no només ha cregut convenient ser el meu tutor de TFG, sinó també fer-me de mentor en un àmbit on té molt de recorregut.

Bibliografia de referència

¹Casassas Simó, Lluís (1988). El debate de la organización territorial de Cataluña. ¿Hacia el Reparto de las minucias del poder?. ERIA, pp.15-26

OSM-Boundaries. (n.d.). <https://osm-boundaries.com/>

Nomenclàtor mundial - Oficina d'Onomàstica - Secció Filològica - Institut d'Estudis Catalans. (n.d.). <https://nomenclator-mundial.iec.cat/>

Nomenclàtor mundial. Criteris per a la fixació de formes. (2022, Desembre). Institut d'Estudis Catalans. https://nomenclator-mundial.iec.cat/docs/criteris_nomenclator_mundial.pdf

World Bank Open Data. (s. f.). World Bank Open Data. <https://data.worldbank.org/>

European Commission. Eurostat. (s. f.). European Commission. Eurostat. <https://ec.europa.eu/eurostat>

City Population - Population Statistics in Maps and Charts for Cities, Agglomeration and Administrative Divisions of all Countries of the World. (s. f.) <http://www.citypopulation.de/>

Portals estadístics:

Albània. <https://www.instat.gov.al/en/>

Alemanya. https://www.destatis.de/EN/Home/_node.html

Andorra. <https://www.estadistica.ad/portal/apps/sites/#/estadistica-ca/>

Aràbia Saudí. <https://datasource.kapsarc.org>

Argentina. <https://censo.gob.ar/>

Aruba. <https://cbs.aw/wp/>

Bahames. <https://www.bahamas.gov.bs/>

Belarús. <https://president.gov.by/en/belarus/numbers/facts>

Colòmbia. <https://www.dane.gov.co/>

Estònia. <https://andmed.stat.ee/en/stat>

Filipines. <https://openstat.psa.gov.ph/>

Grenada. <https://www.finance.gd/>

Islàndia. <https://statice.is/>

Jamaica. <https://statinja.gov.jm/>

Kenya. <https://www.knbs.or.ke/>

Montenegro. <https://www.monstat.org/>

Noruega. <https://www.ssb.no/en>

Països Baixos. <https://www.cbs.nl/en-gb>

Panamà. <https://minerpa.com.pa/>

Annex: Llista d'entitats territorials, nombre de polígons, nivell de l'arxiu GEOJSON i informació dels canvis i criteris.

ENTITAT	N_POLIGONS	NIVELL	INFO
Afganistan, l'	34	4	
Albània	12	6	
Alemanya	16	4	
Algèria	58	4	
Andorra	7	7	
Angola	18	4	
Antàrtida, l'	1	2	
Antigua i Barbuda	3	4	
Aràbia Saudí, l'	13	4	
Argentina, l'	24	4	
Armènia	11	4	
Austràlia	15	4	
Àustria	9	4	
Azerbaidjan, l'	14	4	He borrat un polígon repetit (Zangezur Oriental) i afegit al nv.4 una regió de nv.3
Bahames, les	33	8	
Bahrain	4	4	
Bangladesh	8	4	
Barbados, les	11	6	
Belarús	7	4	
Bèlgica	11	6	He afegit un polígon que era al nv.4 (Brussel·les) i eliminat un altre (Liège French Community)
Belize	6	4	
Benín	12	4	
Bhutan	20	4	
Bolívia	9	4	
Bòsnia i Hercegovina	3	4	
Botswana	10	4	Districtes urbans no inclosos en la capa nv.4
Brasil, el	27	4	
Brunei	4	4	
Bulgària	28	6	
Burkina Faso	13	4	
Burundi	18	4	
Cambodja	25	4	
Camerun, el	10	4	
Canadà, el	13	4	
Cap Verd	22	6	
Colòmbia	33	4	
Comores, les	3	4	
Congo, el	12	4	
Cook, illes	11	4	

Corea del Nord	13	4	
Corea del Sud	17	4	
Costa d'Ivori, la	14	4	
Costa Rica	7	4	
Croàcia	21	6	
Cuba	16	4	
Dinamarca	5	4	
Djibouti	6	4	Illes Maskali incloses en la província de Djibouti (repetició de polígon)
Dominica	10	4	
Egipte	27	4	
El Salvador	14	4	
Emirats Àrabs Units, els	7	4	
Equador, l'	24	4	
Eritrea	6	4	
Eslovàquia	8	4	
Eslovènia	12	5	
Espanya	20	4	La capa nv.4 tenia un polígon addicional (Província eclesiàstica)
Estats Units, els	56	4	Hi havia uns quants polígons mal classificats al nv.4
Estònia	15	6	
Eswatini	4	4	
Etiòpia	13	4	
Fèroe, illes	6	4	País dins del Regne de Dinamarca
Fiji, les	5	4	
Filipines, les	17	3	L'arxiu de descàrrega no estava ben classificat al web, s'ha de descarregar Philipines Maritime Boundary
Finlàndia	19	6	
França	26	4	La capa nv.4 tenia tres polígons addicionals (regions eclesiàstiques i província apostòlica), he passat els polígons ultramar de nv.3 a nv. 4, la capa nv. 3 tenia vuit polígons repetits (Saint-Martin i Saint Pierre i Miquelon, les regions d'ultramar ja existents al nv.4 i la França Pròpia)
Gabon, el	9	4	
Gàmbia	6	4	Els polígons de Districte de Banjul i dels municipis Banjul i Kanifig estaven mal classificats
Geòrgia	12	4	Ossètia del sud crea incongruències entre les delimitacions regionals, però no hi ha altra opció a utilitzar
Ghana	16	4	
Grècia	14	5	He mogut la comunitat monàstica de nv.3 a nv.5

Grenada	7	6	
Groenlàndia	6	4	País dins del Regne de Dinamarca, 1 polígon de més Pituffik que he inclòs a Avanaata
Guatemala	22	4	
Guinea	8	4	Hi havia dues illes repetides
Guinea Equatorial	8	4	Hi havia un polígon repetit (Elobey Chico)
Guinea-Bissau	9	4	
Guyana	10	4	
Haití	10	4	
Hondures	18	4	
Hongria	20	6	He eliminat del nv.4 molts polígons de divisions eclesiàstiques
Iemen, el	22	4	
Illes Marshall, les	24	4	He eliminat dos polígons del nv.4
Illes Salomó, les	10	4	
Índia, l'	36	4	
Indonèsia	38	4	
Iran, l'	31	4	
Iraq, l'	19	4	
Irlanda	26	6	
Islàndia	8	5	
Israel	6	4	
Itàlia	20	4	Hi havia 10 polígons (illes) repetits
Jamaica	14	6	Hi havia un polígon de més al nv.4
Japó, el	47	4	Eliminades les províncies històriques del nv.4
Jordània	12	4	
Kazakhstan, el	20	4	
Kenya	47	4	
Kirguizstan, el	9	4	
Kiribati	3	4	
Kosovo	7	4	Els polígons de districte no tenien cap nivell administratiu assignat, els hi he posat el nv.4
Kuwait	6	4	
Lao	18	4	
Lesotho	10	5	
Letònia	6	4	
Líban, el	9	4	
Libèria	15	4	
Líbia	22	4	
Liechtenstein	11	8	
Lituània	10	4	
Luxemburg	12	6	He tret un polígon del nv.6 (Germany-Luxemburg Condominium)
Macedònia del Nord	8	4	

Madagascar	22	4	
Malàisia	16	4	Hi havia polígons (illes) repetits al nv.4
Malawi	28	4	Hi havia quatre polígons al nv.4 que no eren districtes
Maldives, les	21	4	
Mali	11	4	
Malta	5	4	Hi havia polígons (illes) repetits al nv.4
Marroc, el	12	4	He arreglat els límits amb el Sàhara Occidental, que he hagut de crear
Maurici	10	4	Polígon Outer islands inclou tot allò que no és Maurici (tres illes), assignades a la capa nv.4
Mauritània	13	4	
Mèxic	32	4	Hi havia 2 illes repetides dins el nv.4
Micronèsia	4	4	
Moçambic	11	4	
Moldàvia	37	4	
Mònaco	1	2	
Mongòlia	22	4	
Montenegro	25	6	
Myanmar	15	4	
Namíbia	14	4	
Nauru	14	6	
Nepal	7	3	
Nicaragua	17	4	
Níger	8	4	
Nigèria	37	4	
Niue	13	8	Oficialment 14 polígons per la partició de la capital en dos
Noruega	14	4	L'any 2024 seran 15 comtats
Nova Zelanda	18	4	
Oman	11	4	
Països Baixos, els	16	4	He passat els territoris d'ultramar de nv.3 a nv.4
Pakistan, el	7	4	Hi havia 2 polígons de més al nv.4
Palau	16	4	
Palestina	2	4	He hagut de crear el polígon de Cisjordània
Panamà, el	14	4	
Papua Nova Guinea	22	4	
Paraguai, el	18	4	
Perú, el	26	4	He modificat la regió de Lima per excloure-hi el municipi de Lima
Polònia	16	4	
Portugal	20	6	Hi havia polígons (illes) repetits, Les Açores i Madeira han passat de nv.4 a nv.6
Qatar	8	4	

Regne Unit, el	83	3	9 Anglaterra, 6 Irlanda del Nord, 22 Gal·les, 32 Escòcia; + 3 Dependències de la Corona + 11 Territoris d'ultramar, He arreglat polígons (illes) repetits
República Centreafricana, la	17	4	
República Democràtica del Congo, la	26	4	
República Dominicana, la	32	4	
Romania	42	4	
Ruanda	5	4	
Rússia	83	4	He reassignat Crimea i Sebastòpol a Ucraïna
Sàhara Occidental	1	2	He creat el polígon a partir d'altres
Saint Kitts i Nevis	14	6	
Saint Lucia	10	4	
Saint Vincent i les Grenadines	6	6	
Samoa	11	4	Hi havia 11 polígons (illes) repetits al nv.4
San Marino	9	8	
Santa Seu, la	1	2	
São Tomé i Príncipe	7	6	
Senegal, el	14	4	
Sèrbia	25	6	
Seychelles, les	1	2	
Sierra Leone	5	4	
Singapur	5	6	He inclòs un polígon dins del districte pertinent (Pedra Branca al Districte Sud-oriental)
Síria	14	4	
Somàlia	18	5	He fet una unió de dos polígons
Sri Lanka	9	4	
Sud-àfrica	9	4	
Sudan del Sud, el	13	4	L'Àrea d'Abyei l'he exclòs del polígon on era
Sudan, el	18	4	
Suècia	21	4	He eliminat un polígon del nv. 4
Suïssa	26	4	
Surinam, el	10	4	
Tadjikistan, el	5	4	
Tailàndia	77	4	
Taiwan	22	4	Hi havia 54 illes repetides al nv.4
Tanzània	31	4	Hi havia dues illes repetides (Mbudya i Bongoyo)
Timor-Leste	14	4	
Togo	5	4	
Tonga	5	4	
Trinidad i Tobago	15	4	
Tunísia	24	4	
Turkmenistan, el	6	4	

Turquia	81	4	
Tuvalu	8	8	
Txad, el	23	4	
Txèquia	14	6	
Ucraïna	27	4	Hi havia 2 polígons repetits al nv.4 (Crimea i Sebastòpol)
Uganda	4	3	Hi havia camps de refugiat que he tret del nv. 3
Uruguai, l'	19	4	
Uzbekistan, l'	14	4	He canviat un polígon del nv.3 al nv.4
Vanuatu	6	4	
Veneçuela	25	4	
Vietnam, el	63	4	
Xile	16	4	Hi havia 3 illes repetides al nv.4
Xina, la	33	4	He canviat Macau i Hong kong del nv.3 al nv.4
Xipre	1	2	
Zàmbia	10	4	
Zimbàbue	10	4	
Total	3373		