

The background of the entire slide is filled with numerous water bubbles of varying sizes. Some bubbles are large and prominent, while others are small and scattered. They have a realistic appearance with highlights and shadows, giving them a three-dimensional look.

ON VA A PARAR LA NOSTRA AIGUA?

Favila Valdés-Bango Martín
Data Analytics
IT Academy

INTRODUCCIÓ

Els darrers anys han estat marcats per una greu sequera que ha afectat tota la Península Ibèrica. A les conques internes de Catalunya la sequera ha estat extrema, fins al punt de provocar l'aplicació de mesures del control del consum de l'aigua per assegurar-ne els subministrament fins la reversió del fenomen. Gràcies a les pluges de març i abril, la sequera va poder donar-se per finalitzada després de 2(3) anys d'haver començat. A arrel d'aquest fenomen meteorològic sorgeix la pregunta que dona títol a aquest treball, on va a parar la nostra aigua?

(REFERENCIAR)

OBJECTIUS

Els objectius d'aquest treball es centren en 3 punts concrets:

- Determinar la relació existent entre el consum domèstic i el de les activitats econòmiques. És a dir quina és superior a cada comarca o si una predomina per sobre de l'altra.
- Explorar els diferents sectors econòmics disponibles. Veure quines variables poden ser d'ús per poder relacionar-les amb el consum d'aigua.
- Establir la relació entre el consum de les activitats econòmiques i els seus sectors.

METODOLOGIA

Les dades emprades en aquest treball han estat obtingudes de web institucional [Dades Obertes de Catalunya](#). Concretament s'han seleccionat 3 *datasets* pertanyents a aquest repositori. Aquests *datasets* han estat [Mapa d'explotacions agrícoles \(DUN\) de Catalunya](#) | [Dades obertes de Catalunya](#), [Consum d'aigua a Catalunya per comarques](#) | [Dades obertes de Catalunya](#) i [Establiments d'allotjament turístic inscrits al Registre de Turisme de Catalunya](#) | [Dades obertes de Catalunya](#).

Un cop obtingudes aquestes dades s'ha procedit a netejar les dades tot emprant la llibreria *Pandas* de *python* (v. 3.13.1). Gràcies a aquesta llibreria s'ha pogut comprovar la capacitat dels *datasets* d'interrelacionar-se, ja que en tots s'hi troba el codi comarcal de les comarques a les quals pertanyen les dades. Un cop observada aquesta capacitat s'ha procedit a treballar amb cadascun dels *datasets* per poder treballar més tard conjuntament.

En el cas del *dataset* relacionat amb el consum de l'aigua, les modificacions han estat nul·les, ja que aquest es proveïa net, només s'ha escurçat allà on era necessari per fer més lleuger l'anàlisi.

Per contra, en el cas del *dataset* de les explotacions agràries s'han hagut de convertir diverses variables numèriques que es trobaven en format *float* a *integer*, ja que això dificultava la combinació dels *datasets*. També s'han agrupat les dades per campanyes (anys) i per codis comarcals, tot sumant les hectàrees de cada explotació. D'aquest manera s'han obtingut les hectàrees explotades per any i comarca. En algun cas s'han fet servir també el tipus d'explotació.

Pel que fa al *dataset* d'allotjaments turístics, és el que presenta de les dades més fragmentades, ja que els seus camps contenen en alguns casos fins el 70% de registres nuls. Tot i així, com resultaven ser variables d'interès s'han emprat igualment (nombre total de places, per exemple). En aquest cas cada registre representa un allotjament

turístic inscrit al Registre de Turisme on resulten interessants, els camps relatius al codi comarcal, el tipus d'establiment i el nombre total de places (tot i presentar una gran fragmentació de les dades). En aquest cas s'ha procedit de forma similar a fer el sumatori a les explotacions agrícoles, tot realitzant el sumatori del nombre total de places per comarca (i en cas de ser necessari per tipus d'establiment). En aquest cas les dades no es relacionen amb un any, ja que el registre s'actualitza freqüentment en funció de les baixes i les noves inscripcions, alhora que és de recent creació (11 d'abril del 2024).

Per poder visualitzar les dades s'ha fet servir un arxiu *geojson* de les [Comarques de Catalunya](#) extret de GitHub. Aquest arxiu s'ha combinat amb els diferents altres *datasets* per poder representar la informació en forma de mapa comarcal de Catalunya.

De cara a analitzar estadísticament les dades s'han servit tests de Shapiro-Wilk per comprovar la normalitat, correlacions d'Spearman i Pearson per veure la correlació lineal entre les diferents variables testades i per últim s'han realitzat regressions lineals per veure com de realista es generar un model a partir de les variables disponibles.

CONSUM D'AIGUA

El consum d'aigua a les troba diferenciat principalment pels seu ús. D'una banda es troba el consum domèstic. Aquest es caracteritza per ser el consum del dia a dia, dutxes, rentar roba, plats, cuinar, regar jardins, omplir piscines, etc. D'altra banda es troba el consum de l'activitat econòmica i de fonts pròpies, que inclou tots aquelles empreses i particular que consumeixen tant aigua de la xarxa com de pous propis, entre d'altres, per poder dur a terme la seva activitat econòmica.

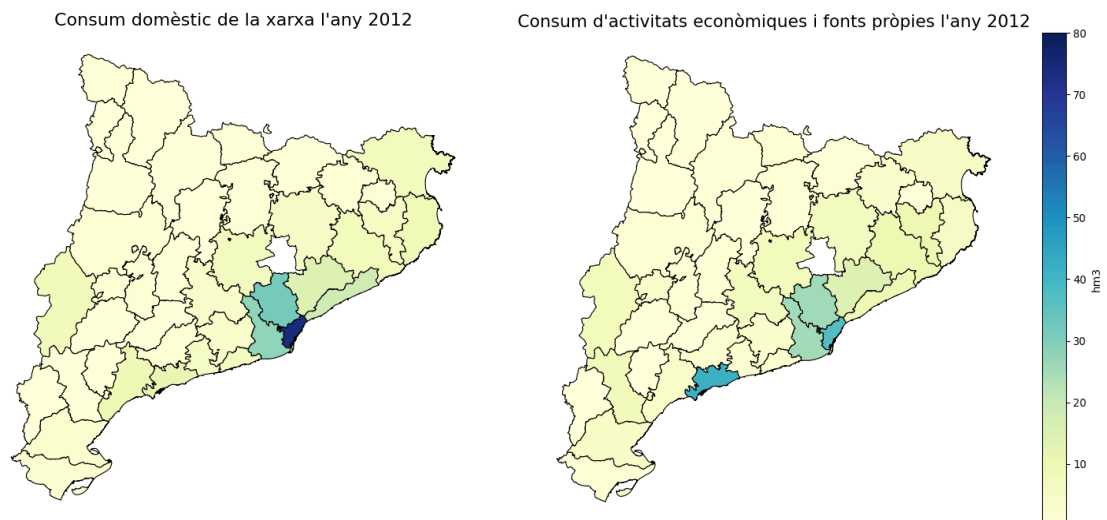


Figura 1. Mapa de les comarques de Catalunya. Representen els consums domèstic (esq.) i d'activitats econòmiques (dta.) canviant amb els anys.

Com es pot observar a la **Figura 1**, el consum domèstic arriba a màxims a l'àrea metropolitana de Barcelona que concentra la major part de la població de Catalunya. Pel que fa al consum de les activitats econòmiques s'inclou també la comarca del Tarragonès, la qual presenta una elevada industrialització que dispararia aquests consums. A grans trets, el consum domèstic és **superior** al de les activitats econòmiques (Taula A1), tot i així que sigui major no implica que se'n faci un ús abusiu, per això

apareixen altres mesures com el consum domèstic per càpita **Figura 2**. En el cas d'aquesta figura sí que ens pot servir per entendre on hi ha un consum més elevat de

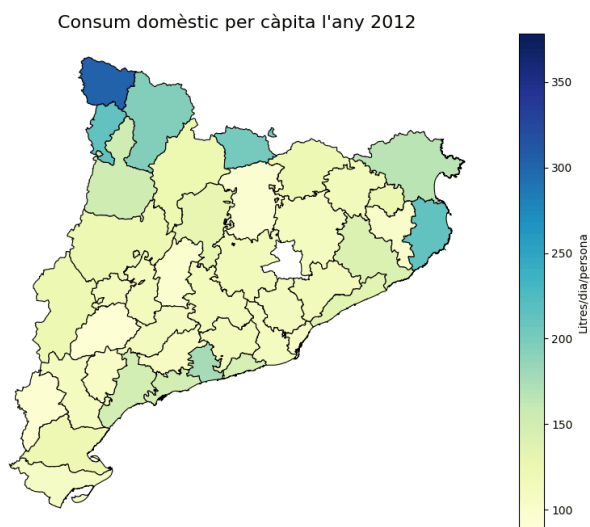


Figura 2. Mapa de les comarques de Catalunya que representa el consum domèstic per càpita canviant amb els anys.

litres d'aigua per persona i dia i entendre que pot estar passant en aquelles comarques on el consum és més elevat. Aquesta mesura sí que ens podria mostrar aquelles regions on s'està produint un consum més abusiu. Resulta que la figura mostra com les comarques amb consums domèstics per càpita més elevats són comarques de segones residència tant de costa com de muntanya. Baix Empordà, Alt Empordà, Cerdanya o Vall D'Aran apareixen com a comarques amb valors elevats.

Per acabar d'analitzar les dades relacionades amb el consum cal destacar que l'existència d'un consum domèstic superior al de l'activitat econòmica no implica que a cada comarca aquest fet es compleixi. Per poder veure aquesta irregularitat en la relació en el consum d'aigua, s'ha construït un índex a partir de la *ratio* entre consum d'activitat econòmica i consum domèstic. Aquest índex representat a la **Figura 3** en un mapa, permet comparar el consum entre les diferents comarques i veure quines comarques tenen un consum domèstic inferior al econòmic. Això podria ajudar a inferir els usos que se li donen o no a l'aigua, on es podria suposar que aquelles que tenen un consum econòmic superior són comarques amb menys població que tenen economies basades en una agricultura amb major demanda hídrica. Tot i així això no es pot demostrar en aquest estudi.

Índex d'activitat econòmica l'any 2012

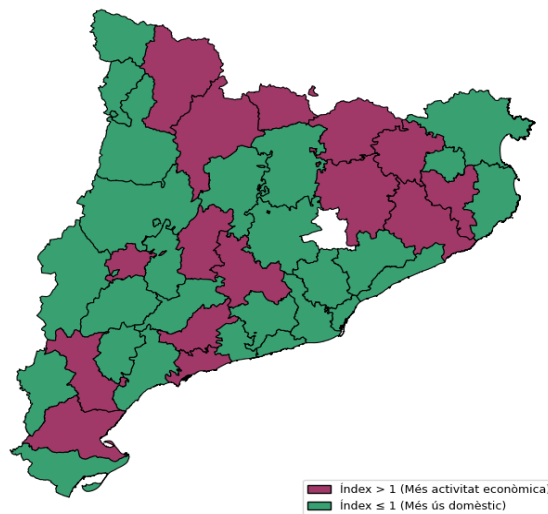


Figura 3. Mapa de les comarques de Catalunya que representa l'índex d'activitat econòmica per càpita canviant amb els anys

EXPLOTACIONS AGRÍCOLES

Una de les possibles activitats econòmiques que podrien presentar un consum elevat a Catalunya és l'agricultura. Aquesta ocupa un gran extensió del territori català, on prop del **46% del territori** és ocupat per explotacions agrícoles. La distribució de la superfície d'aquestes explotacions es pot observar a la **Figura 4**, on podem observar com les comarques de l'interior del territori són les que presenten major superfície destinada a activitats agrícoles.

Hectàrees explotades l'any 2019

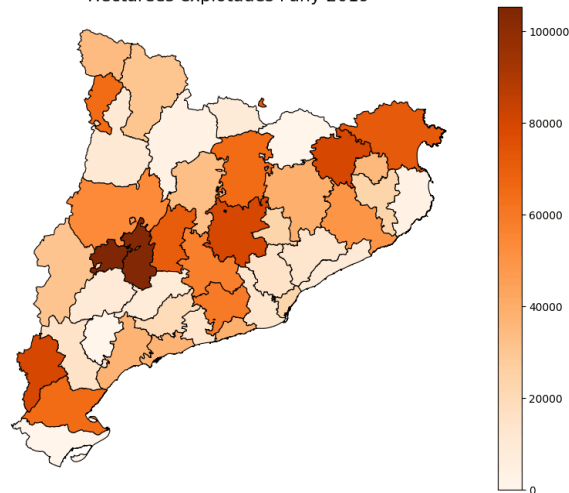


Figura 4. Mapa de les comarques de Catalunya que representa la superfície destinada a l'activitat agrícola canviant amb els anys.

Aquesta activitat, però, es troba molt diversificada al llarg del territori. Les comarques del nord destinen les seves extensions a bestiar, que canvia al llarg dels anys, i el sud és més d'agricultura de secà, i es manté més estable en els temps. Aquest fet es pot observar la **Figura 5**. Resulta de rellevància ja que l'ús agrícola que se'n fa és clau en

el consum d'aigua que es realitza. No és igual el consum de conreus de secà que el de regadius, com es pot observar al Pla de gestió del districte de la conca fluvial de Catalunya 2022-2027(ACA, 2021) on s'esmenta el el regadu suposaria prop del 40% del consum totals d'aigua a Catalunya. Ni tan poc es comparable el de la ramadareia extensiva amb la intensiva. Tot i així, el *dataset* del qual provenen aquestes dades, no proveix cap camp que permeti relacionar el consum d'aigua amb l'explotació agrícola. L'única forma que es presenta és l'ús de la dimensió de les explotacions per poder

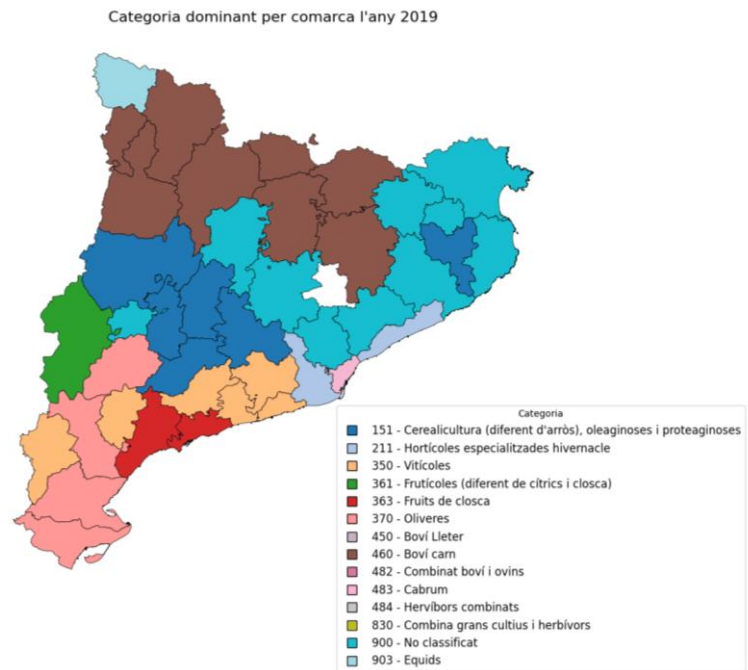


Figura 5. Mapa de les comarques de Catalunya que representa el tipus d'explotació agrícola majoritari, canviant amb els anys.

intentar inferir una possible relació entre la mida de l'explotació i un consum superior d'aigua. Tot i que, no té perquè existir dita relació.

ALLOTJAMENTS TURÍSTICS

Un altre sector de cabdal importància és el turisme. A Catalunya pren una especial rellevància sobretot a la costa, la qual presenta una enorme quantitat d'allotjaments turístics registrats, com es pot observar a la **Figura 6**, amb el nombre de places totals

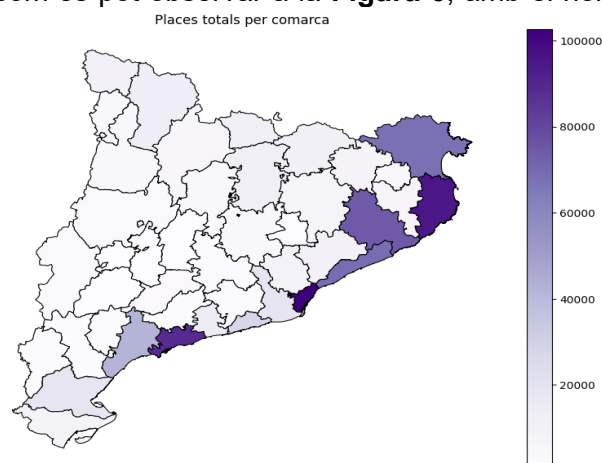


Figura 6. Mapa de les comarques de Catalunya que representa el nombre total de places disponibles en allotjaments turístics.

disponibles a cada comarca. A més a més, el tipus d'allotjament també podria ser un factor clau en el consum tot i que *a priori* no tindria tant de pes com la tipologia de l'explotació agrícola. El que podem trobar en aquest cas és com hi ha una distribució irregular en les comarques de la tipologia dels establiments (**Figura 7**). Resulta que el nord de Catalunya presenta més places destinades a càmpings, que la Catalunya central destina més i occidental destina més places en hotels i que els habitatges d'ús turístic es troben de forma irregular en algunes comarques

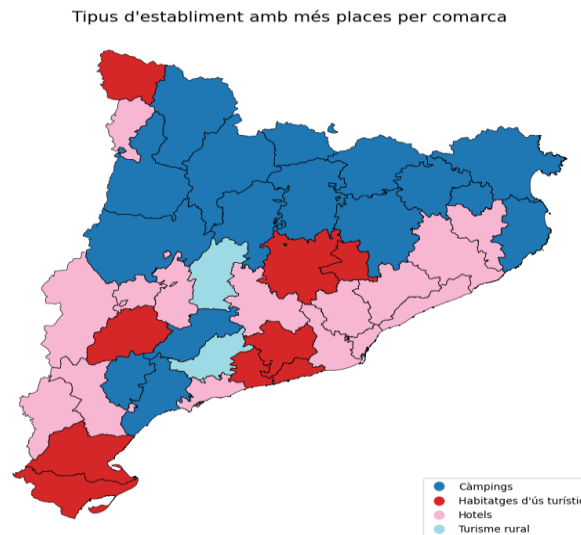


Figura 7. Mapa de les comarques de Catalunya que representa el tipus d'establiment turístic predominant.

Una altra vegada més però, no hi ha una variable que permeti fer comparacions acurades entre el consum del sector i el de l'activitat econòmica en conjunt. Per això s'ha emprat el nombre total de places disponibles als allotjaments turístics per fer la comparació, ja que semblaria la variable més adequada per trobar alguna relació.

RELACIÓ ENTRE CONSUM I SECTORS

Per poder trobar relacions entre el consum de l'activitat econòmica, i els sectors agrari i turístic, s'han emprat diferents variables a correlacionar degut a la manca d'una dada concreta de consum hídric d'aquests sectors. S'ha testat la normalitat de les variables correlacionades, i s'ha observat que en cap dels casos segueixen una normal ja que els p-valors eren inferior a 0,05. Tot i així, s'ha considerat el Teorema del Límit Central per poder fer ús de la correlació de Pearson.

Per observar la possible relació entre el sector turístic i el consum en l'activitat econòmica s'han correlacionat aquest últim amb el nombre de places disponibles per a cada comarca. De manera que s'ha obtingut una Correlació de Pearson de 0,47. Aquesta correlació podria considerar-se com a mitjanament correlacionades, amb un p-valor de 0,002. A partir d'aquesta correspondència s'ha procedit a generar una regressió lineal (**Figura 8**). El model predit és precari ja que té un ajust bastant baix, amb una R^2 de 0,22. Tot i així, tampoc es busca trobar un model fiable a partir del qual predir futurs consums amb el nombre de les places d'allotjament disponibles. Més aviat es busca visualitzar la possible correlació existent entre aquestes dades. En aquest cas s'intueix que en cas de disposar de dades relacionades amb el consum i la ocupació en establiments turístics es podrien generar models capaços de predir la demanda d'aigua

de forma acurada per aquest sector. Cal esmentar que s'han eliminat alguns valors de la taula original, aquells que es troben per sobre de 15 milions de metres cúbics ja que podien considerar-se **outliers** en l'estudi, en no encabir els usos industrials en l'estudi dels sectors.

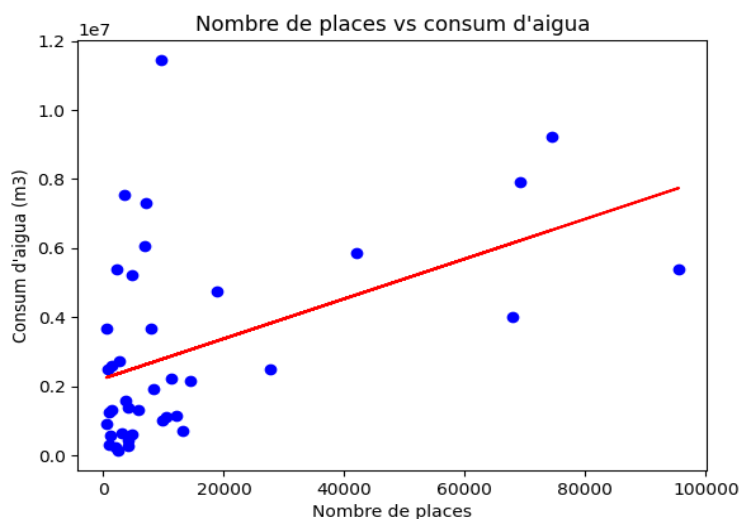


Figura 8. Regressió lineal entre el nombre de places i el consum hídic de l'activitat econòmica i les fonts pròpies.

En el cas del sector agrari les variables correlacionades han estat el consum d'aigua en l'activitat econòmica i les hectàrees destinades a les explotacions agràries. A diferència del cas anterior el coeficient emprat ha estat Spearman, ja que s'obté un millor resultat en ser més precisa per correlacions determinades entre variables no paramètriques. El resultat del coeficient és de -0,14 tot indicant una correlació negativa molt fluixa amb un p-valor de 0,04. Un cop s'ha generat la regressió lineal (**Figura 9**), s'ha observat que l'ajust R^2 és de 0,01. Per tant aquesta no serveix com a model predictor del consum pel sector agrari i que per a aquesta anàlisi no són dades correlacionables. També pot existir algun problema amb les dades de consum i no estan acabant de incloure l'impacte real de l'agricultura en el consum, fent impossible la seva predicció a partir d'aquesta correlació. En principi, segons la bibliografia consultada, podria ser forta la correlació entre aquestes variables, ja que prop del 40% del consum d'aigua a Catalunya es deu

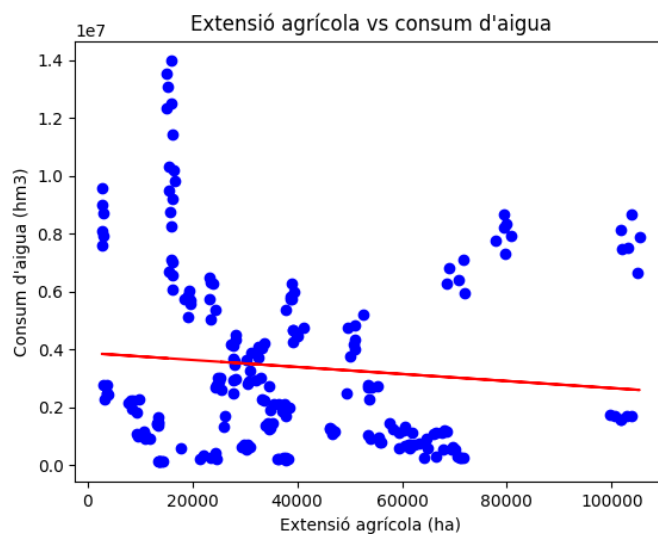


Figura9. Regressió lineal entre l'extensió agrícola i el consum hídic de l'activitat econòmica i les fonts pròpies.

a la indústria agropecuària (regadiu sobretot) i per tant una major extensió de regadiu hauria de generar un major consum d'aigua. Tot i així, no s'ha pogut demostrar.

CONCLUSIÓ

Per concloure el treball el que es pot afirmar és que la relació entre consum domèstic i econòmic varia entre comarques. A més a més, existeix una gran fragmentació de les dades i manca d'informació relativa a alguns sectors que dificulta la interpretació de la anàlisi. Per tant no es pot arribar a determinar amb certesa ni predir la relació que guarden els sectors econòmics exposats amb el seu consum. El sector turístic semblaria presentar una correlació i model més fiable que el sector agrari, tot i així, insuficient per ser fiable. Una millor font d'informació podria ajudar a desenvolupar un model predictiu fiable i a establir correlacions entre les diferents variables.

BIBLIOGRAFIA

- Agència Catalana de l'Aigua (2021).Pla de gestió del districte de la conca fluvial de Catalunya 2022-2027 (DCFC). [3er ciclo de planificación \(2022 - 2027\). Agencia Catalana del Agua](#)

ANNEX

Taula A1. Taula de consum d'aigua en funció de l'ús i l'any.		
Any	Domèstic xarxa (m ³)	Activitats econòmiques i fonts pròpies(m ³)
2012	329587399	264070475
2013	313551779	259345246
2014	309942495	261405426
2015	313989027	261428653
2016	318431495	269429150
2017	321575409	277151797
2018	320815198	269683194
2019	326604221	273181246
2020	345174787	234113745
2021	338432965	247774903
2022	330954411	265598841
2023	322085377	252183425
2024	313935196	233601767



Imatge A1. Mapa comarcal de Catalunya amb les capitals assenyalades.